

Université des Sciences et Technologies de Lille 1
Institut de Recherches Interdisciplinaires, USR3078
Ecole Doctorale Biologie-Santé

THÈSE DE DOCTORAT

pour l'obtention du grade de docteur de l'Université de Lille 1

Analyse protéomique des complexes associés aux membres de la famille CBX (HP1 et Polycomb) dans des cellules humaines

Discipline : aspects moléculaires et cellulaires de la biologie

Spécialité : biochimie et biologie moléculaire

présentée par **Julien VANDAMME**

dirigée par le Pr Pierre-Olivier ANGRAND

Soutenue le 23 septembre 2009 à Lille

Le jury était composé de :

Rapporteurs :

Dr Francesca PALLADINO
Dr Laurent DELVA

Directeur de Recherche - CNRS, UCBL ENS - Lyon
Chargé de Recherche - Inserm, Université de Bourgogne - Dijon

Examineurs :

Dr Emmanuelle BOUVERET
Dr Christian ROLANDO
Dr Bernard VANDENBUNDER
Pr Pierre-Olivier ANGRAND

Chargée de Recherche – CNRS, LISM IBSM - Marseille
Directeur de Recherche – CNRS, Univ. Sciences Technologies - Lille
Directeur de Recherche – CNRS, Univ. Sciences Technologies - Lille
Professeur, Université des Sciences et Technologies - Lille



Avant propos

Pour commencer, je remercie vivement Francesca Palladino et Laurent Delva d'avoir accepté d'être les deux rapporteurs de ce travail et Emmanuelle Bouveret d'en être un des examinateurs.

Je profite de cette page pour exprimer ma sincère gratitude à celui et à celle qui m'ont choisi... puis m'ont enseigné, encadré, écouté, conseillé et soutenu pendant ces 3 années. Leur gentillesse et leur disponibilité sont à la hauteur de leur savoir. Pierre-Olivier Angrand et Pamela Völkel, veuillez trouver ici tous les deux mes remerciements.

La recherche en biologie n'est malheureusement pas gratuite, je remercie les deux directeurs successifs de l'IRI, Bernard Vandebunder et Ralf Blossey d'avoir assuré, entre autres, les financements de notre équipe et le CNRS ainsi que le Ministère de l'Enseignement et de la Recherche (via l'Université des Sciences et Technologie de Lille 1) pour mon financement personnel. D'un point de vue matériel, je me dois de remercier Yvan de Launoit, qui a accepté d'héberger notre équipe pendant toute la durée de ma thèse à l'Institut de Biologie de Lille, nous permettant d'avoir accès aux différents équipements ainsi qu'à un environnement scientifique de qualité.

Pendant ma thèse, notre équipe s'est agrandie avec les arrivées de 3 post-doctorants : Patricia Pellegrino de Souza, Claire Rosnoblet et Nicolas Fourré. Leurs expériences et leurs collaborations aux projets de l'équipe ont été non négligeables pour la réussite de ma thèse.

J'ai une pensée particulière pour Adeline Page et Christian Rolando qui m'ont chaleureusement accueilli et donné accès au Centre Commun de Mesures de Spectrométrie de Masse de Lille.

Je n'oublierai pas les discussions scientifiques (ou pas) avec David Bernard et Arnaud Augert, mes voisins de paillasse. Je souhaite sincèrement au second d'avoir la même réussite que le premier.

Pour leurs conseils techniques et leur disponibilité, je citerai Capucine Van Rechem, Cindy Degerny, Caroline Smet, Elisabeth Ferreira, Jean-Luc Baert, Didier Monté et Bruno Genet en particulier, qui ont su combler mes nombreuses lacunes expérimentales avec bienveillance.

De plus, pour leur soutien permanent, je remercie bien évidemment ma famille, mes ex-boss : Louise Deltour, Andras Paldi et Fabienne Parker (qui m'ont également transmis leur passion pour la recherche) et mes ami(e)s : Thomas, les « Lyonnais » (dont Caroline qui souhaite voir son prénom dans mes remerciements), les « BioGeCeM 13 » (dont Béa pour le service dernière minute) et les « footeux du samedi ».

Pour terminer, je tiens à remercier chaleureusement toutes celles et tous ceux qui n'auront pas cru en moi et qui m'auront donné la force et l'envie d'y arriver.

Enfin, si je t'ai oublié(e) où que tu estimes le mériter, tu peux toujours écrire ton nom ici:

« J'ai fait ce que je pouvais, à défaut d'avoir accompli ce que je voulais » [Alix Mathurin]

Titre

Analyse protéomique des complexes associés aux membres de la famille CBX (HP1 et Polycomb) dans des cellules humaines

Résumé

Dans le noyau des cellules eucaryotes, l'ADN est associé aux histones pour former la chromatine. Les extrémités terminales des histones peuvent subir un grand nombre de modifications post-traductionnelles réversibles (méthylation, acétylation, phosphorylation, ubiquitinylation...). Certaines de ces marques épigénétiques définissent des domaines chromatiniens particuliers du noyau. Les triméthylations des lysines 9 et 27 de l'histone H3 (H3K9me3 et H3K27me3) correspondent respectivement à des domaines d'hétérochromatine constitutive et d'hétérochromatine facultative. Les protéines qui possèdent un chromo-domaine sont capables de se lier à des lysines méthylées, et permettent de recruter des complexes qui ont une activité enzymatique ou mécanique sur la chromatine. Les protéines de la famille CBX (ChromoBoX) possèdent toutes un chromodomaine, elles sont subdivisées en 2 groupes: les protéines du groupe HP1 (CBX1, CBX3 et CBX5) et celles du groupe Polycomb (CBX2, CBX4, CBX6, CBX7 et CBX8) qui se fixent respectivement sur H3K9me3 et H3K27me3, respectivement. Afin de mieux comprendre comment s'organise la chromatine au niveau des régions hétérochromatiques, nous avons purifié, en condition native, les complexes associés à ces 8 protéines dans des cellules humaines en culture. Pour ce faire, nous avons opté pour la purification d'affinité en tandem (TAP). Les protéines co-éluées ont été identifiées par spectrométrie de masse. Nos résultats confirment que, d'une part les protéines du groupe Polycomb sont associées au complexe PRC1 (Polycomb Repressive Complex 1). Toutefois, ces protéines sont mutuellement exclusives au sein du complexe PRC1, indiquant qu'il existe plusieurs complexes PRC1 de composition distincte dans la cellule. D'autre part, de nouveaux partenaires associés aux protéines du groupe HP1 ont été identifiés. La mise au point et le développement de la technologie TAP sur des cellules humaines en culture nous ont permis de purifier des complexes associés à la chromatine. Cette technique demeure un outil biochimique efficace pour la purification de complexes protéiques.

Mots clefs : purification TAP, HP1, Polycomb, PRC1, complexes protéiques

Thèse préparée à : Institut de Recherches Interdisciplinaires – CNRS USR 3078

Chromatinomics group
Parc Scientifique de la Haute Borne
50 Avenue Halley BP 70478
F-59658 Villeneuve d'Ascq Cedex

Title

Proteomic analysis of CBX-containing complexes (HP1 and Polycomb) from human cells

Abstract

In the nucleus of eukaryotic cells, DNA is wrapped around histones to form chromatin. The terminal ends of histones may undergo many reversible post-translational modifications (methylation, acetylation, phosphorylation, ubiquitinylation...). Some of these epigenetic marks define specific chromatic regions in the nucleus. The tri-methylation of lysines 9 and 27 of histone H3 (H3K9me3 and H3K27me3) correspond respectively to constitutive and facultative heterochromatin. Chromo-domain containing proteins can bind to methylated lysines, and can recruit complexes having enzymatic or mechanical activity on chromatin. All the members of the CBX (ChromoBoX) family contain a chromodomain, they are divided into 2 groups: the HP1 group (CBX1, CBX3 and CBX5) and the Polycomb group (CBX2, CBX4, CBX6, CBX7 and CBX8) able to bind to H3K9me3 and H3K27me3, respectively. To better understand how the chromatin is organized in heterochromatic regions we purified, in native condition, the complexes associated with these 8 proteins from human cells in culture. To this end, we opted for the tandem affinity purification (TAP). The co-eluted proteins were identified by mass spectrometry. Our results confirm that firstly, Polycomb group proteins are involved in PRC1 complex (Polycomb Repressive Complex 1). However, these proteins are mutually exclusive in the PRC1 complex, indicating that several PRC1 of distinct compositions co-exist in the cell. On the other hand, new partners associated with HP1 group were identified. The development of the TAP technology to cultured human cells allowed us to purify complexes associated with chromatin. This technique remains an effective tool for the biochemical purification of protein complexes.

Keywords: TAP purification, HP1, Polycomb, PRC1, multiprotein complexes

SOMMAIRE

Avant propos.....	3
Résumé	4
Abstract.....	5
SOMMAIRE	7
TABLE DES ILLUSTRATIONS.....	11
INTRODUCTION	12
Généralités sur la chromatine et l'épigénétique	15
1 La chromatine.....	15
1.1 Empaquetage de l'ADN dans le noyau et domaines chromatinien.....	15
1.2 Organisation du nucléosome	17
2 L'épigénétique	18
2.1 « Décoration » du matériel génétique	18
2.1.1 Méthylation de l'ADN	18
2.1.1.1 Déméthylation de l'ADN.....	19
2.2 Les modifications post-traductionnelles des histones	19
2.2.1 La méthylation des histones	21
2.2.1.1 Méthylation des arginines	22
2.2.1.1.1 Les PRMT.....	22
2.2.1.1.2 La déméthylation des arginines	22
2.3 La méthylation des lysines	22
2.3.1.1 Les KMT (lysine méthyl-transférases)	22
2.3.1.2 Les KDM (lysine déméthylases)	23
2.3.1 L'acétylation	23
2.3.2.1 La désacétylation	23
2.3.3 La phosphorylation, les kinases	24
2.3.4 Les ubiquitin-ligases	25
2.3.5 SUMO ligases.....	25
2.3.6 ADP-ribosylation.....	26
2.3.7 L'isomérisation des prolines (pas de liaison covalente).....	26
2.3.8 Hypothèse du « code histone »	27
2.4 Les variants d'histones.....	28
2.4.1 Histones H2A	28
2.4.2 Les variants de H2B ont une préférence sexuelle	29
2.4.3 Les variants de H3 : l'hypothèse du code-barre d'histones	29
2.5 Les ARN non codants	30
3 Les effecteurs.....	30
3.1 Lysines acétylées.....	31
3.2 Les lysines méthylées.....	31
3.2.1 La superfamille Royal	31
3.2.1.1 Le chromodomaine	31
3.2.1.2 Le double-tudor domaine	32
3.2.1.3 Le tandem-tudor domaine.....	32
3.2.1.4 MBT	32
3.2.2 PHD (Plant HomeoDomain)	33
3.3 Les phosphorylations	33
3.3.1 14-3-3.....	33
3.3.2 BRCT.....	33
3.4 Domaines de liaison à l'ubiquitine.....	33
3.5 SIM (SUMO-Interacting Motifs)	33
3.6 Domaines de liaison à l'ADP-ribose.....	34
3.7 Les résidus non modifiés = nouveau point de contact avec la chromatine	34
3.7.1 Le PHD de AIRE (AutoImmune REgulator)	34
3.7.2 Le domaine ADD (ATRX-DNMT-DNMT3L) riche en cystéines.....	34
3.7.3 WD40	34
3.8 Domaines de liaison à l'ADN méthylé	34
4 Cancer et thérapie épigénétique/épidrogues.....	35

Introduction sur les CBX	37
LES PROTEINES DE LA SOUS-FAMILLE HP1	38
1 Généralités.....	38
1.1 Variégation par effet de position.....	38
1.2 Structure.....	38
1.3 Phosphorylations.....	40
2 HP1 et l'hétérochromatine péricentrique.....	40
2.1 Paradigme d'assemblage/de propagation de l'hétérochromatine.....	41
2.2 La voie ARNi impliquée dans l'hétérochromatine.....	42
3 Partenaires protéiques et fonctions biologiques.....	43
3.1 Réplication et/ou la réparation de l'ADN.....	43
3.2 Répresseurs transcriptionnels.....	44
3.3 Enveloppe nucléaire.....	44
3.4 Contrôle du cycle cellulaire.....	44
4 Répression des gènes euchromatiques.....	45
5 Activation transcriptionnelle.....	46
5.1 HP1 régule l'expression des gènes euchromatiques.....	46
5.2 HP1 est essentiel pour l'expression des gènes hétérochromatiques.....	47
6 Fonctions télomériques des HP1.....	47
7 Dynamique des HP1 – Réversibilité.....	48
8 Conclusions.....	48
LES PROTEINES CBX POLYCOMB	50
1 Introduction.....	50
1.1 Découverte des Polycomb, une histoire de Nobelisés.....	50
1.2 Mémoire de l'identité cellulaire.....	50
2 Les PcG s'organisent en complexes pour réprimer les gènes.....	51
2.1 Le complexe PRC1.....	51
2.1.1 Structure.....	51
2.1.2 Fonction.....	52
2.2 Le complexe PRC2.....	52
2.2.1 Structure.....	52
2.2.2 Fonction.....	53
2.3 Le complexe PhoRC.....	54
2.4 Modèle de recrutement des complexes PcG.....	55
2.4.1 Chez la drosophile.....	55
2.4.2 Chez les mammifères.....	55
2.4.3 Le modèle n'est pas universel.....	56
3 Le noyau du PRC1 à la loupe.....	56
3.1 Les homologues Pc.....	57
3.2 Les homologues Psc.....	58
3.3 Les homologues de Ring.....	59
3.4 Les homologues de Ph.....	60
4 Les gènes cibles des PcG.....	60
4.1 Gènes Hox et régulateurs du développement.....	61
4.2 Cellules souches et domaines bivalents.....	62
5 Marques épigénétiques associées aux PRC1.....	62
5.1 Les « corps Polycomb » et la sumoylation.....	62
5.2 Ubiquitinylation de H2A.....	63
5.3 Machinerie ARNi.....	63
5.4 Déméthylases de H3K27me3.....	63
6 Conclusions.....	64
Les méthodes d'étude des interactions protéine-protéine et purification de complexes.....	66
1 Le double-hybride (DH).....	66
1.1 Principe de la technologie DH.....	66
1.2 Limitations.....	67
2 La purification d'affinité en tandem (TAP) : nouvel outil pour l'exploration des protéomes.....	68
2.1 Principe de la technologie TAP.....	68
2.2 Limitations.....	69
2.3 Les étiquettes TAP pour la purification de complexes à partir de cellules mammifères.....	70
3 Bilan et choix de la méthode.....	72
OBJECTIFS DE LA THÈSE	74

MATÉRIELS ET MÉTHODES	76
1 Clonage GATEWAY (recombinaison <i>in vitro</i>)	76
2 Amplification PCR.....	77
3 Vecteurs rétroviraux	77
4 Vecteurs d'expression	77
5 Culture cellulaire.....	78
6 Infection rétrovirale pour l'obtention de lignée stable.....	78
7 Transfection des cellules	78
8 Mesure de l'activité luciférase.....	79
9 Western blot	79
10 Immunodétection sur cellules fixées	80
11 Immunoprécipitation	80
12 GST-pulldown	80
13 Chromatographie d'exclusion	81
14 Purification TAP	81
15 Analyse de spectrométrie de masse.....	82
RESULTATS ET DISCUSSION	84
<u>Partie 1 : Les protéines CBX Polycomb (CBX2, 4, 6, 7 et 8)</u>	84
1 Etude de l'activité répressive des protéines CBX Polycomb	84
1.1 Système rapporteur exogène et sur-expression d'une protéine fusionnée à GAL4.	84
1.2 Système rapporteur endogène et expression inductible d'une protéine fusionnée à GAL4.	88
1.2.1 Système rapporteur chromatinien, inductible par un analogue de l'oestrogène	88
1.2.2 Système rapporteur chromatinien, inductible par un analogue de la progestérone.....	91
2 Génération de lignées cellulaires exprimant de manière stable les effecteurs-TAP	92
2.1 Principe d'infection avec des vecteurs viraux.....	92
2.2 Choix du modèle cellulaire	94
2.3 Validation des lignées stables et localisation des protéines CBX-TAP	95
2.4 Discussion.....	96
3 Evaluation des masses moléculaires des complexes associés aux protéines CBX Polycomb	98
3.1 Principe de la chromatographie d'exclusion stérique	98
3.2 Résultats	99
3.3 Discussion.....	100
4 Purification TAP des protéines CBX Polycomb	101
4.1 Principe.....	101
4.2 Résultats	101
5 Identification des partenaires protéiques purifiés par TAP	103
5.1 Principe.....	103
5.2 Protéines co-éluées avec les CBX-TAP	105
5.2.1 Les protéines du PRC1	109
5.2.2 La caséine kinase 2 est-elle un nouveau membre du PRC1 ?	110
5.2.3 Partenaires spécifiques à chaque protéine CBX	111
5.3 Discussion.....	111
6 Validation des interactions avec les membres du PRC1	113
6.1 Co-immunoprécipitations (CoIP) : interactions indirectes.....	113
6.1.1 Principe	113
6.1.2 Résultats	113
6.2 GST-pulldown : interactions directes	115
6.2.1 Principe	115
6.2.2 Validation d'interaction directe entre les protéines PCGF et PHC avec les CBX	116
6.2.3 Discussion	117
7 CBX4 et les 14-3-3	118
7.1 Localisation d'un site de fixation mode 1 dans la séquence de CBX4	119
7.2 Etude de l'interaction de CBX4 avec les 14-3-3 et de la phosphorylation de CBX4.....	119
7.2.1 Identification d'un site de phosphorylation <i>in vivo</i>	119
7.2.2 Validation de l'interaction directe de 14-3-3 avec CBX4.....	121
7.2.3 Importance de la phosphorylation de CBX4 sur la composition du PRC1 associé.....	124
7.2.4 Activité répressive de CBX4-S164A	124
7.3 Discussion.....	125
8 Purification TAP de protéines non-CBX présentes dans le PRC1	126
8.1 Résultats	127
8.1.1 peu de CBX	128
8.1.2 Le coeur du PRC1 est conservé.....	128

8.1.3	RYBP et YAF2.....	128
8.1.4	La caséine kinase 2.....	129
8.1.5	pas de RNF2 avec RING1 ?.....	129
8.2	Discussion.....	129
Partie 2 : Les protéines CBX HP1 (CBX1, 3 et 5).....		132
1	Etude de l'activité répressive des protéines CBX HP1	132
1.1	Système rapporteur exogène et sur-expression d'une protéine fusionnée à GAL4	132
1.2	Système rapporteur endogène et expression inductible d'une protéine fusionnée à GAL4	133
1.3	Discussion.....	134
2	Validation des lignées cellulaires et localisation des protéines HP1-TAP.....	134
2.1	Purification TAP	136
2.2	Discussion.....	139
Partie 3 : Les contaminants		146
CONCLUSIONS		154
PERSPECTIVES		158
ANNEXES		162
	Annexe 1: Listes des peptides identifiés dans les purifications TAP.....	155
	Annexe 2: Peptides appartenant aux protéines Polycomb.....	200
	Annexe 3: Article médecine/sciences	210
	Annexe 4: Article BMC Cell Biology.....	214
REFERENCES BIBLIOGRAPHIQUES		240

TABLE DES ILLUSTRATIONS

Figure 1 – Représentation schématique des différents états de compaction de l'ADN	16
Figure 2 – Structure tridimensionnelle d'un nucléosome	17
Figure 3 – Représentation des principales modifications post-traductionnelles.....	27
Figure 4 - Contrôle épigénétique du destin cellulaire.....	35
Figure 5 - Arbre phylogénétique des protéines CBX humaines et représentation de leur structure.	37
Figure 6 - Représentation schématique des protéines HP1 humaines.....	38
Figure 7 – Modèle actuel pour décrire l'établissement de l'hétérochromatine péricentrique	42
Figure 8 - Modèle actuel pour décrire l'établissement de l'hétérochromatine facultative.....	55
Figure 9 - Représentation schématique des protéines composant le cœur du PRC1	57
Figure 10 – Principe de la technique dite du « double hybride »	67
Figure 11 - Principe de la purification d'affinité en tandem	68
Figure 12 - Carte du vecteur de destination pIE-NGAL4-GW	84
Figure 13 - Système rapporteur luciférase exogène.....	85
Figure 14 - Activité répressive des protéines CBX Polycomb sur 3 promoteurs différents.....	86
Figure 15 - Activité répressive de protéines décrites comme appartenant au complexe PRC1	87
Figure 16 - Représentation schématique du système rapporteur luciférase « chromatinien »	89
Figure 17 - Carte du vecteur de destination pRH-N-DLO-GAL4-GW	89
Figure 18- Activité répressive des protéines CBX Polycomb sur le rapporteur chromatinien.....	90
Figure 19 - Carte du vecteur de destination pRN-N-DLP-GAL4-GW.....	92
Figure 20 - Principe de l'infection virale pour l'obtention de lignées stables TAP	93
Figure 21 - Carte du vecteur de destination pRP-NTAP-GW.....	94
Figure 22 - Immunomarquage des protéines CBX dans les lignées stables HeLa-TAP.....	96
Figure 23 - Principe de la chromatographie d'exclusion	98
Figure 24 - Courbe de calibration de la colonne Superose 6 10/300 GL.....	99
Figure 25 - Analyse par western blot des fractions obtenues après chromatographie d'exclusion.....	100
Figure 26.1 - Gels SDS-PAGE TAP des protéines CBX Polycomb colorés au bleu de coomassie colloïdal.....	101
Figure 27 - Suivi de la purification de la protéine TAP-CBX6 par western blot anti-CBP.....	102
Figure 28 - Etapes successives pour l'identification par spectrométrie de masse des protéines présentes dans un gel de polyacrylamide.....	103
Figure 29 - Peptides de CBX4 identifiés par MS-MS au cours de la purification TAP	111
Figure 30 - Principe de la co-immunoprécipitation (co-IP).....	113
Figure 31 - Co-IP des protéines RNF2, PCGF6 et BMI1 avec CBX8	114
Figure 32 - Co-IP de la protéine RNF2 avec toutes les protéines CBX Polycomb.....	114
Figure 33 - Co-IP des protéines RING1 et RNF2 avec toutes les CBX Polycomb.....	115
Figure 34 – Principe du GST-pulldown pour valider une interaction <i>in vitro</i>	116
Figure 35 - Production des protéines recombinantes en système procaryote.....	116
Figure 36 - GST-pulldown entre les CBX et les membres des familles PCGF et PHC.....	117
Figure 37 - Identification d'un site putatif mode1 de fixation des protéines 14-3-3 dans la séquence de CBX4 ..	119
Figure 38 - Digestions théoriques de la protéine CBX4 (iso2) avec les protéases trypsine ou Glu-C	120
Figure 39 – Co-immunoprécipitations de CBX4 et CBX4-S164A avec 14-3-3 ε.....	122
Figure 40 - Coloration au bleu de coomassie des protéines recombinantes produites en système procaryote...	122
Figure 41 - GST pulldown pour tester l'interaction entre 14-3-3 et CBX4 non phosphorylé	123
Figure 42 - GST pulldown pour tester l'interaction entre 14-3-3 et CBX4 et CBX8.....	123
Figure 43 – Co-immunoprécipitation de RNF2 et PCGF6 avec CBX4 sauvage et mutée	124
Figure 44- Activité répressive de CBX4-S164A sur 3 promoteurs différents dans un système exogène.....	125
Figure 45 - Purification TAP avec les protéines PCGF2 et RING1	126
Figure 46 - Activité répressive des protéines HP1 sur 2 promoteurs différents	133
Figure 47 - Activité répressive des protéines HP1 sur un rapporteur chromatinien.....	134
Figure 48 – Immunomarquage des protéines HP1-TAP dans les lignées stables HeLa-TAP	135
Figure 49 - Coloration au bleu de coomassie colloïdal des gels HP1-TAP.....	136
Figure 50 - Couverture de séquence des protéines HP1 qui ont servi d'appât pour les purifications TAP	138
Figure 51 - Résumé des protéines majoritaires identifiées lors des purifications TAP des 3 isoformes HP1	139
Figure 52 – Bilan des protéines associées au PRC1 identifiées avec les CBX Polycomb	155
Figure 53 – Bilan des protéines majoritaires co-éluées avec les 3 isoformes HP1 par purification TAP	157
Tableau 1 - Modifications post-traductionnelles des histones.....	20
Tableau 2 - Protéines Polycomb et complexes associés chez la drosophile et l'homme	60
Tableau 3 - Tableau des différents oligonucléotides utilisés	76
Tableau 4 – Protéines co-éluées avec les CBX Polycomb-TAP.....	109
Tableau 5 – Résumé des membres du PRC1 identifiés dans les purifications TAP.....	110
Tableau 6 – Protéines co-éluées avec RING1-TAP et PCGF2-TAP	127
Tableau 7 – Protéines co-éluées avec les HP1.....	137
Tableau 8 – Liste des protéines considérées comme contaminants	148

Liste des abréviations

4-OHT: 4-hydroxytamoxifen
ADN: acide désoxyribo-nucléique
ADP: adénosine di-phosphate
ARN: acide ribo-nucléique
ARNi: interférence par ARN
ATM: ataxia-telangiectasia mutated
ATR: ATM and Rad3-related
BAC: chromosome bactérien artificiel
CaCl₂: chlorure de calcium
CAF: chromatin assembly factor
CBX: chromobox protein
CENP: centromeric protein
CDYL: chromodomain Y-like
CHD: chromo helicase DNA-binding domain
Chromo: chromatin organization modifier domain
CK: caséine kinase
CMV: Cytomegalovirus
Co-IP: co-immunoprécipitation
CpG: dinucléotide cytosine-guanine
CtBP: c-terminal binding protein
DA: domaine d'activation
DBD: DNA binding domain
DH: double-hybride
DLO: domaine de liaison aux œstrogènes
DLP: domaine de liaison à la progestérone
DMEM: dulbecco's modified eagle medium
Dnmt: methyl-transférase d'ADN
DTT: dithiothréitol
EMBL: european molecular biology laboratory
EDTA: ethylenediamine tetra-acetic acid
EGTA: ethylene glycol-bis(2-aminomethyl)-N,N',N',N'-tetraacetic acid
Esc: extra sex comb
E(z): enhancer of zeste
FDA: food and drug administration
FPLP: fast protein liquid chromatography
GFP: green fluorescent protein
GNAT: Gcn5-related N-acetyl-transferase
GST: glutathione-S-transférase
HA: hémagglutinine
HCl: chlorure d'hydrogène
HDAC: histone deacetylase
HEK: human embryonic kidney
HOAP: origin recognition complex-associated protein
Hox: homéotique
HP1: heterochromatin protein 1
HTH: helix turn helix
Ig: immunoglobuline
IRES: internal ribosome entry site
JmjC: domaine Jumonji C
KAP: KRAB associated protein
KAT: lysine acetyl-transférase
KDM: lysine déméthylase
KMT: lysine méthyl-transférase
KRAB: krüppel-associated box
L3MBTL: Lethal 3 malignant brain tumor
LAP: localization and affinity purification
LTR: Long Terminal Repeat
MART: mono-ADP-ribosyl-transférase
MBD: methyl-CpG-binding domain
MDC: mediator of DNA damage checkpoint protein
MgCl₂: chlorure de magnésium
MPP: M-phase phosphoprotein
Msk: mitogen- and stress-activated kinase
MYST: Monocytic leukemia zinc finger protein
m/z: masse divisée par la charge

NaCl: chlorure de sodium
 NAD⁺: nicotinamide adenine dinucléotide
 NaF: fluorure de sodium
 Na₃VO₄: orthovanadate de sodium
 NPC: nuclear pore complex
 ORC: origin recognition complex
 ORF: open reading frame
 PADI: peptidyl arginine dé-aminase
 PARG: poly-ADP-ribose glycohydrolase
 PARP: poly-ADP-ribose polymerases
 PAZ: polyubiquitin-associated zinc binding
 PBS: phosphate buffer saline
 PcG: polycomb
 PCGF: protéines Posterior sex comb
 Pcl: polycomb like
 PCR: polymerase chain reaction
 PEI: polymère d'éthylènimine
 Ph: polyhomeotic
 PHC: protéines Polyhoméotic
 PIKK: phosphatidylinositol-3-OH-kinase like family
 PKA: protéine kinase A
 PP: phosphatase
 PRC: polycomb repressive complex
 PRE: polycomb response element
 PRMT: protéine arginine méthyl-transférase
 Psc: posterior sex comb
 PTGS: post-transcriptional gene silencing
 RBBP: retinoblastoma-binding protein
 RING: protéines RING
 RMN: résonance magnétique nucléaire
 ROW: relative-of-WOC
 RU486: mifépristone
 RZPD: deutsches ressourcenzentrum für genomforschung
 SAHF: senescence associated heterochromatic foci
 SAM: sterile alpha motif
 Scm: sex comb on midleg
 SCML: sex comb on midleg-like
 SDS: sodium dodécyl sulfate
 SENP: sentrin specific protease
 SIR: silent information regulatory
 SUUR: suppressor of under-replication
 SUV: suppressor of variegation
 Su(z): suppressor of zeste
 SVF: sérum de veau fetal
 TAP: tandem affinity purification
 TBP: TATA-box binding protein
 TDG: thymidine DNA glycosylase
 TEV: tobacco etch virus
 TGS: transcriptional gene silencing
 TIF: transcriptional intermediary factor
 TiO₂: hydroxide de titanium
 TK: thymidine kinase
 Trx: trithorax
 UAS: Upstream Activation Sequence
 ULP: ubiquitin-like protein specific protease
 USP: ubiquitin specific protease
 UTX: ubiquitously transcribed tetratricopeptide repeat, X chromosome
 VEGF: vascular endothelial growth factor
 WOC: without children
 YY: yin-yang

INTRODUCTION

GENERALITES SUR LA CHROMATINE ET L'EPIGENETIQUE

1 La chromatine

1.1 *Empaquetage de l'ADN dans le noyau et domaines chromatiniens*

Il est admis que l'information génétique de tous les organismes vivants est portée par leur séquence d'ADN. L'ADN siège principalement dans le noyau de tous les eucaryotes, il est constitué de 2 brins de bases nucléiques complémentaires formant la célèbre double hélice (Watson JD, 1953, Nature) (Figure 1A). Chez l'Homme, si on mettait bout à bout les hélices d'ADN d'une seule cellule, on obtiendrait un « fil » d'une longueur proche de 2 mètres. Le noyau de la cellule, mesurant environ 6µm de diamètre, a donc mis au point une stratégie de compaction afin de réduire d'un facteur estimé à 10 000 la taille de l'ADN (Ball P, Nature, 2003). En effet, ce dernier est d'abord enroulé autour d'une structure globulaire composée de 8 protéines basiques : les histones. Environ 146 paires de bases nucléiques entourent l'octamère constitué de 2 de chacune des histones H2A, H2B, H3 et H4 (Arents G, 1991, PNAS). L'association « 146pb d'ADN + 8 histones » est appelée nucléosome, et constitue l'unité fondamentale de la chromatine. La chromatine correspond à un assemblage nucléoprotéique dynamique, réparti de manière hétérogène dans le noyau des cellules. Les nucléosomes permettent un premier état de compaction de l'ADN qui, associé aux histones, ressemble à un « collier de perles » (Figure 1B). Un second niveau de compaction, avec l'incorporation de l'histone de liaison H1 (Allan J, 1986, J Mol Biol), révèle une fibre ayant un diamètre de 30nm appelée nucléofilament (Figure 1C), et permettant de réduire la taille de l'ADN d'un facteur 40. À partir de cette fibre, plusieurs architectures différentes peuvent être observées :

- l'**euchromatine** : elle est répartie de manière homogène au sein du nucléoplasme, elle correspond aux gènes qui sont fortement actifs, c'est-à-dire que l'ADN est facilement accessible par la machinerie de transcription. Il s'agit de chromatine décondensée (Figure 1E) même si le facteur de condensation de l'ADN est de l'ordre de 1000. Les gènes localisés dans ces régions sont répliqués en début de phase S.

- l'hétérochromatine est une région fortement condensée (Figure 1F) et transcriptionnellement silencieuse qui se réplique tardivement lors de la phase S. Elle est localisée principalement en périphérie du noyau et du nucléole. Elle est beaucoup plus dense que l'euchromatine (facteur de compaction = 10000) et est subdivisée en 2 grandes catégories :

- l'**hétérochromatine constitutive** est majoritairement présente au niveau des centromères et des télomères des chromosomes. Ces régions sont riches en ADN non codants comme : les séquences satellites, répétées et parasites. Elles sont identiques dans toutes les cellules de l'organisme.

- l'**hétérochromatine facultative** correspond aux régions chromosomiques qui peuvent-être à la fois transcrites (euchromatine) ou réprimées (hétérochromatine), en fonction de l'état de différenciation au

sein d'une cellule unique ou alors en fonction d'un type cellulaire à l'autre au sein d'un même organisme.

C'est lors de la mitose, plus particulièrement au cours de la métaphase, que l'ADN est le plus compacté. Dans cette configuration, l'ADN est inaccessible aux machineries de transcription et de réplication (Figure 1F).

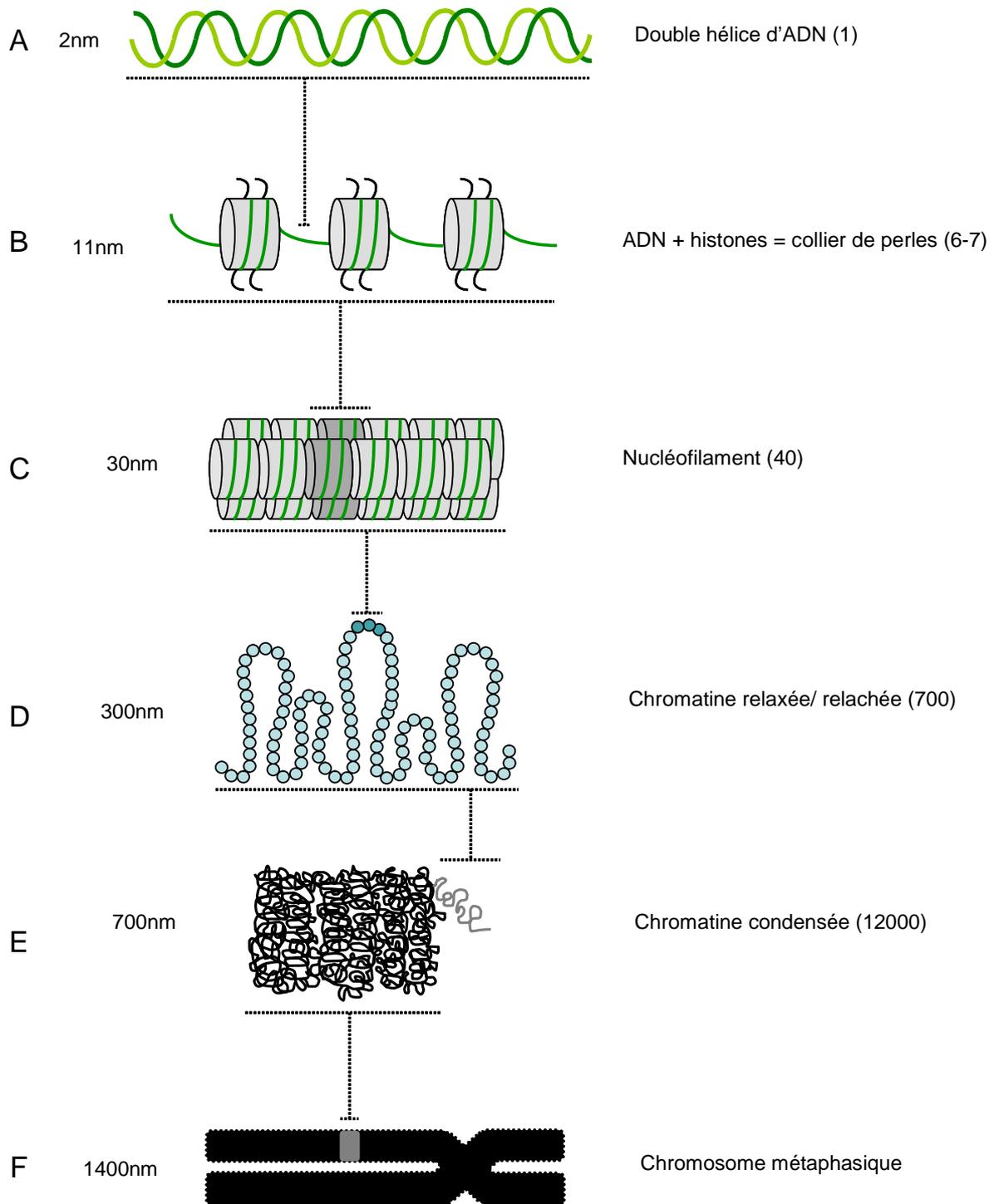


Figure 1 – Représentation schématique des différents états de compaction de l'ADN dans les cellules eucaryotes (**facteur de compaction**)

1.2 Organisation du nucléosome

Les différentes protéines histones ont été très conservées au cours de l'évolution des eucaryotes, reflétant leur importance dans l'organisation des génomes et soulignant leur rôle primordial dans la formation de la chromatine. Ce sont les 2 histones H3 et H4 constituant le cœur de l'octamère qui sont les plus conservées. Les histones sont de petites molécules ayant de faibles poids moléculaires, comprenant entre 102 et 135 acides aminés. Il existe 5 sous-types d'histones : H1, H2A, H2B, H3 et H4. Les familles de gènes qui codent pour les histones sont également très conservées et retrouvées dans de nombreuses espèces. La plupart des organismes pluricellulaires contiennent plusieurs copies de ces 5 gènes car la quantité (en poids) d'histones nécessaires pour assurer la synthèse d'une nouvelle chromatine lors de la réplication est approximativement la même que celle de l'ADN. Ainsi, les histones sont synthétisées au début de la phase S du cycle cellulaire, correspondant à des taux de transcription élevés.

Les histones sont constituées de trois hélices α reliées entre elles par deux courtes boucles L1 et L2. Cette structure, appelée « *histone fold* », permet la dimérisation des histones. D'une manière générale, les histones portent toutes un domaine globulaire central hydrophobe et un domaine amino-terminal chargé positivement, hydrophile, peu structuré et flexible (Luger K, 1997, Nature) :

- Les domaines globulaires hydrophobes interagissent entre eux pour former le cœur du nucléosome. Compte tenu du caractère extrêmement basique des histones, riches en acides aminés chargés positivement (lysine et arginines), le nucléosome n'est stable qu'en présence d'ADN (acide) ou dans des solutions à forte concentration en sels ($>1M$ NaCl).
- Les extrémités N-terminales des histones, constituées de 15 à 30 résidus, ont été baptisées « queues » d'histones et jouent un rôle primordial dans la structure du nucléosome, le remodelage de la chromatine et les événements cellulaires.

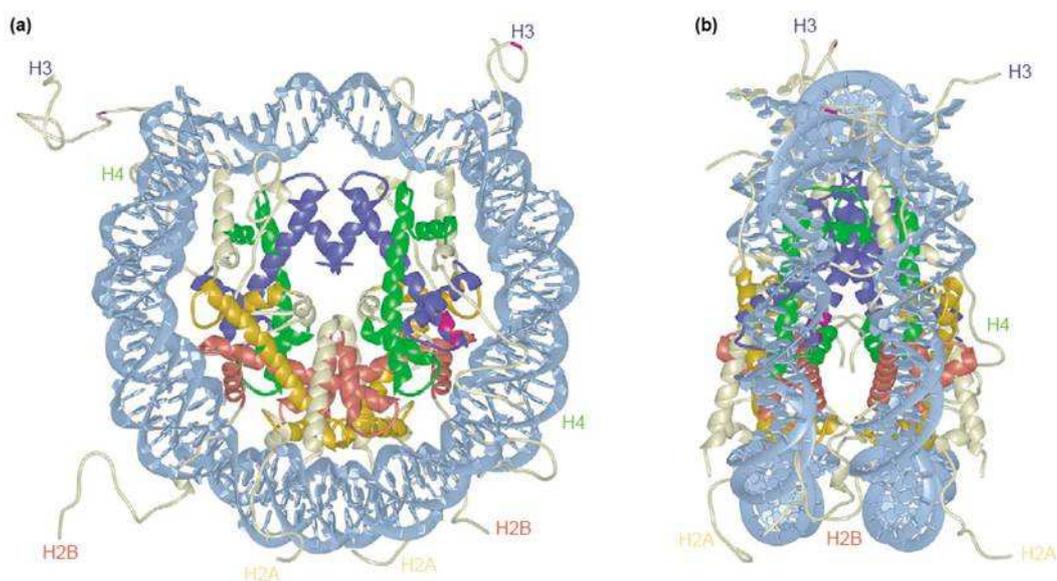


Figure 2 – Structure tridimensionnelle d'un nucléosome (d'après Luger K, 2003, Current Opin Genet Dev). Les histones H2A, H2B, H3 et H4 sont respectivement colorés en jaune, rouge, bleu et vert. Les queues d'histones sont représentées en blanc et l'ADN en bleu clair (a) vue de face (b) vue de côté

La structure en haute résolution du cœur d'un nucléosome (Figure 2) montre les détails de l'interaction histones-ADN, qui est confinée au squelette phospho-diester du filament d'ADN au niveau de la surface interne de la super-hélice. Un ensemble de contacts se fait toutes les 10 paires de bases où le petit sillon de la double hélice est dirigé vers l'intérieur. Des interactions électrostatiques et des liaisons hydrogènes avec les phosphates de l'ADN, aussi bien que les contacts non polaires avec les groupes désoxyriboses, sont observés. Le manque de contacts avec les bases de l'ADN est en accord avec le manque de spécificité de séquence de l'interaction ADN-histone, évidence pour la capacité de l'octamère d'histones à emballer n'importe quel ADN virtuel. Notons que les queues des histones sont exposées à l'extérieur du nucléosome, une particularité qui va s'avérer cruciale pour l'organisation de la chromatine.

2 L'épigénétique

Historiquement, le mot épigénétique a été utilisé pour décrire les événements qui ne pouvaient pas être expliqués par les principes de la génétique (Waddington CH, 1942, Endeavour). D'une manière générale, il correspond à la relation existante entre le génotype et le phénotype, capable de changer le résultat final de l'expression d'un gène sans modifier la séquence de son ADN. C'est ainsi que l'on peut expliquer comment toutes les cellules d'un organisme pluricellulaires, qui contiennent strictement le même génome, possèdent des morphologies et des fonctions cellulaires distinctes. L'épigénétique couvre de nombreux processus impliquant des modifications covalentes et non-covalentes de l'ADN et des histones ainsi que des petits ARN non codants. Les caractéristiques épigénétiques sont héréditaires au cours de la division cellulaire.

Nous utiliserons la nomenclature « Brno » pour signifier les différentes modifications post-traductionnelles des histones dans la suite du manuscrit (Turner BM, 2005, Nat Struct Mol Biol). Le Tableau 1 regroupe l'ensemble des protéines impliquées dans les modifications d'histones qui ont été décrites dans des cellules de mammifères.

2.1 « Décoration » du matériel génétique

2.1.1 Méthylation de l'ADN

Chez les mammifères, la méthylation de l'ADN est confinée majoritairement à l'addition d'un groupement méthyle sur le 5^{ème} carbone d'une cytosine précédant une guanosine (dinucléotide CpG). Cette modification covalente est connue pour avoir une fonction essentielle dans le génome, par son implication dans des processus tels que la régulation transcriptionnelle, la stabilité chromosomique (X inactif) ou l'empreinte parentale (Bird A, 2002, Genes Dev). Les régions riches en « GC », plus connues sous le nom d'îlots CpG, sont en général résistantes à la méthylation et sont associées à la plupart des gènes humains actifs. La méthylation de l'ADN intervient dans la répression transcriptionnelle de régions spécifiques au sein du génome, mais son mécanisme d'action reste

méconnu. Chez les mammifères, il existe 3 classes de méthyltransférases de l'ADN : Dnmt1, Dnmt2 et Dnmt3.

Dnmt1 est considérée comme une méthyltransférase de maintien car elle méthyle préférentiellement l'ADN hémi-méthylé, permettant d'assurer la modification de l'ADN néo-synthétisé au cours de la réplication. Dnmt2 est une méthyltransférase des petits ARN de transfert (Goll MG, 2006, Science). La classe Dnmt3 contient 3 membres, Dnmt3a, 3b et 3L. Les 2 premières possèdent une activité enzymatique et sont capables de méthyle *de novo* l'ADN nu ou hémi-méthylé (Okano M, Cell, 1999). Elles sont très fortement exprimées dans les cellules souches embryonnaires tandis qu'elles sont très faiblement exprimées dans des cellules somatiques différenciées (Pradhan S, 2003, Clin Immunol). Ces enzymes possèdent des cibles communes et des cibles qui leur sont spécifiques, en particulier certains gènes soumis à l'empreinte parentale (Kato Y, 2007, Hum Mol Genet). Dnmt3L qui ne possède pas d'activité méthyltransférase est considérée davantage comme un régulateur de l'établissement de la méthylation (activateur des 2 autres – Suetake I, 2004, JBC) plutôt que son catalyseur (Bourc'his D, 2001, Science). Cependant elle possède un rôle primordial dans l'établissement de la méthylation *de novo* de toutes les régions différentiellement méthylées.

2.1.1.1.1 Déméthylation de l'ADN

Au cours de la réplication semi-conservative de la matrice, seules des cytosines non-méthylées sont incorporées au brin néo-synthétisé, entraînant une déméthylation passive de 50% de l'ADN. Cependant, une déméthylation active faisant intervenir des m⁵C désaminases, a également été décrite. Des cycles de méthylation/déméthylation sont observables au niveau des promoteurs de multiples loci actifs dans des cellules de mammifères (Kangaspeka S, 2008, Nature / Métivier R, 2008, Nature). Les déméthylations sont assurées par les mêmes enzymes qui effectuent les méthylation, à savoir Dnmt3a et Dnmt3b. Ces 2 enzymes sont capables d'assurer la conversion d'une cytosine méthylée en thymine. Ensuite, les glycosylases TDG (thymidine DNA glycosylase) et MBD4 (methyl-binding domain) retirent spécifiquement la thymine pour la remplacer par une cytosine afin de réparer le mésappariement (T:G) au sein de la séquence ADN.

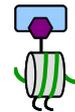
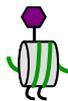
2.2 Les modifications post-traductionnelles des histones

Les histones ne sont pas que de simples protéines d'empaquetage de l'ADN, elles sont également des structures moléculaires qui participent à la régulation de l'expression des gènes. En effet, les modifications post-traductionnelles des histones ont été découvertes au milieu des années 1960, leur influence sur la transcription a immédiatement été suggérée (Allfrey VG, 1964, PNAS). Depuis, une soixantaine d'acides aminés différents ont pu être décrits comme pouvant subir une modification post-traductionnelle. L'observation de ces modifications, principalement au niveau des extrémités amino-terminales des histones, a mis en exergue le rôle de ces modifications :

- elles affectent les interactions ADN-histones ou les histones entre elles et peuvent modifier la structure chromatinienne, par exemple : l'acétylation de la lysine 16 de l'histone H4 a un effet négatif sur la formation de la fibre de 30nm (Shogren-Knaak M, 2006, Science)
- elles créent des sites d'ancrage pour les protéines qui se fixent sur la chromatine

Certaines modifications ont été associées à l'extinction des gènes, d'autres à leur activation. La gamme des modifications qui ont été rapportées dans divers organismes est très variée : méthylation, acétylation et phosphorylation sont les plus étudiées, d'autres modifications comme l'ubiquitinylation, la sumoylation, la déamidation, l'ADP-ribosylation semblent également influencer les nucléosomes.

Toutes ces modifications sont dynamiques et peuvent changer rapidement, elles sont réalisées par des enzymes possédant des domaines catalytiques dédiés à l'ajout ou à la suppression d'une modification bien spécifique sur un résidu donné. Par conséquent, les histones ne sont probablement pas les substrats exclusifs de ces enzymes.



Histone (résidu)	modif	dépose la modification	efface la modification	lecture de la modif	
H3	ac	KAT4/TAF1, KAT7/MYST2, KAT9/ELP3, KAT13A-D/SRC1-ACTR-P160-CLOCK		Bromo ?	
		CUL4A + CUL4B + DDB1 + DDB2 + ROC1			
	R2	0			WD40
		me	PRMT4	JMJD6 (me2)	Tudor ?
	T3	ph	Haspin		14-3-3 ?
		0			PHD, MBT
	K4	me1	KMT7/SET7-9	KDM1/LSD1/BHC110, KDM5B/JARID1B/PLU1	dbles chromo
		me2	KMT2A/MLL1, KMT2B/MLL2, KMT2C/MLL3, KMT2D/MLL4, KMT2E/MLL5, KMT2F/hSET1A, KMT2G/hSET1B, KMT2H/ASH1, KMT8/RIZ1 ?	KDM1/LSD1/BHC110, KDM5A/JARID1A/RBP2, KDM5B/JARID1B/PLU1, KDM5C/JARID1C/SMCX, KDM5D/JARID1D/SMCY	dbles chromo, WD40
		me3	KMT2A/MLL1, KMT2B/MLL2, KMT2C/MLL3, KMT2D/MLL4, KMT2E/MLL5, KMT2F/hSET1A, KMT2G/hSET1B, KMT2H/ASH1, KMT8/RIZ1 ?	KDM5A/JARID1A/RBP2, KDM5B/JARID1B/PLU1, KDM5C/JARID1C/SMCX, KDM5D/JARID1D/SMCY	dbles/tandem tudor, PHD
	R8	me	PRMT7		Tudor ?
	K9	ac	KAT2A/hGCN5, KAT2B/PCAF, KAT12/TFIIIC90		Bromo ?
				KDM3B/JHDM2b, KDM4A/JMJD2A/JHDM3A, KDM4B/JMJD2B, KDM4C/JMJD2C/GASC1	
		me1	KMT1D/EuHMTase/GLP, KMT1E/ESET/SETDB1	KDM1/LSD1/BHC110, KDM3A/JHDM2a	MBT
		me2	KMT1A/SUV39H1, KMT1B/SUV39H2	KDM1/LSD1/BHC110, KDM3A/JHDM2a, KDM4D/JMJD2D	Chromo, MBT
		me3	KMT1A/SUV39H1, KMT1B/SUV39H2, KMT1C/G9a	KDM4D/JMJD2D	Chromo, PHD
	S10	ph	IKK- α , PKA, Msk1/2, Rsk2, Aurora B	MKP-1	14-3-3
	T11	ph	Chk1, DLK/ZIP		14-3-3 ?
	K14	ac	KAT2A/hGCN5, KAT2B/PCAF, KAT6A/MOZ/MYST3, KAT6B/MORF/MYST4, KAT12/TFIIIC90		Bromo ?
	K18	ac	KAT2A/hGCN5, KAT2B/PCAF, KAT12/TFIIIC90		Bromo ?
	R17	me	PRMT4		Tudor ?
R26	me	PRMT4		Tudor ?	

	K27	me1				
		me2	KMT6/EZH2	KDM6A/UTX, KDM6B/JMJD3	Chromo	
		me3		KDM6A/UTX, KDM6B/JMJD3	Chromo	
	S28	ph	Aurora B, MSK1, MSK2		14-3-3	
	K36	me1		KDM2A/JHDM1a/FBXL11, KDM2B/JHDM1b/FBXL10		
		me2		KDM2A/JHDM1a/FBXL11, KDM2B/JHDM1b/FBXL10, KDM4A/JMJD2A/JHDM3A, KDM4B/JMJD2B, KDM4C/JMJD2C/GASC1		chromo barrel
		me3		KDM4A-4B-4C		chromo barrel, PHD
	K79	me2/3	KMT4/DOT1L		Tudor	
	R128	me	PRMT4		Tudor ?	
	R129	me	PRMT4		Tudor ?	
R131	me	PRMT4		Tudor ?		
R134	me	PRMT4		Tudor ?		
H4		me	PRMT8, PRMT9iso4			
		ac	KAT4/TAF1, KAT13A-D/SRC1-ACTR-P160- CLOCK			
		ub	CUL4A + CUL4B + DDB1 + DDB2 + ROC1			
	S1	ph	CKII		14-3-3 ?	
	R3	me	PRMT4, PRMT5	JMJD6 (me2)	Tudor ?	
	K5	ac	KAT1/HAT1, KAT5/TIP60/PLIP, KAT7/MYST2		Bromo ?	
	K8	ac	KAT5/TIP60/PLIP, KAT7/MYST2		Bromo ?	
	K12	ac	KAT1/HAT1, KAT5/TIP60/PLIP, KAT7/MYST2		Bromo ?	
	K16	ac	KAT5/TIP60/PLIP, KAT8/MYST1/HMOF	SirT2	Bromo ?	
	K20	0				
me1		KMT5A/Pr-SET7-8			dbble/tandem tudor, MBT	
me2					dbble/tandem tudor, MBT	
H2A		ac	KAT5/TIP60/PLIP		Bromo ?	
		me	PRMT9iso4			
	K5	ac	KAT3A/CBP, KAT3B/P300		Bromo ?	
	K119	ub	Ring1A, Ring1B, 2A-UB	USP3-22-16-21, 2A-DUB		
H2B		ac	KAT2A/hGCN5, KAT2B/PCAF		Bromo ?	
		ar	PARP1, PARP2	PARG ?		
	K12	ac	KAT3A/CBP, KAT3B/P300		Bromo ?	
	S14	ph	Mst1		14-3-3 ?	
	K15	ac	KAT3A/CBP, KAT3B/P300		Bromo ?	
H2AX	S139	ph	PIKK		BRCT	

Tableau 1 - Modifications post-traductionnelles des histones et enzymes impliquées chez l'Homme

2.2.1 La méthylation des histones

Les histones peuvent être méthylées sur 2 résidus distincts, d'un côté les arginines, d'un autre les lysines. Ces différentes méthylations sont déposées ou retirées par tout un panel d'enzymes possédant des activités spécifiques pour un résidu et/ou un état de méthylation particuliers.

2.2.1.1 Méthylation des arginines

2.2.1.1.1 Les PRMT

Les PRMT (Protéine Arginine Methyl-Transférases) catalysent le transfert d'un ou 2 groupements méthyle sur les arginines, la di-méthylation peut-être symétrique ou asymétrique (Kouzarides T, 2002, *Curr Opin Genet Dev*), elles sont conservées de la levure à l'homme. Chez les mammifères, on dénombre 11 PRMT, 9 possèdent une activité méthyl-transférase identifiée. D'une manière générale, la méthylation des arginines des histones H3 et H4 corrèle principalement avec une activité transcriptionnelle (Lee DY, 2005, *Endocr Rev*). Cependant, la méthylation d'histones par PRMT5 au niveau des promoteurs de certains gènes suppresseurs de tumeurs est responsable de leur extinction (Pal S, *Mol Cell Biol*, 2004). Enfin, la plupart de ces enzymes catalysent la méthylation de protéines non histones comme par exemple PRMT4 sur CBP/p300 (Xu W, 2001, *Science*).

2.2.1.1.2 La déméthylation des arginines

La méthylation des arginines peut-être remplacée par une autre modification : la citrullination. En effet, l'enzyme PADI4 (Peptidyl Arginine Dé-Iminase 4) est capable de convertir les arginines mono-méthylées (ou non modifiées) en citrulline par un processus appelé déimination ou déméthylimination (Cuthbert GL, 2004, *Cell* / Wang Y, 2004, *Science*). PADI4 n'est pas capable de modifier une arginine di-méthylée, son activité préviendrait donc la di-méthylation. La citrullination est considérée comme une modification post-traductionnelle des histones jouant un rôle dans la transcription (Thompson PR, 2006, *ACS Chem Biol*). JMJD6 (PTDSR1) est capable de déméthyliser H3R2me2 et H4R3me2 (Chang et al, 2007, *Science*), donc les méthylation d'arginine sont réversibles.

2.3 La méthylation des lysines

2.3.1.1.1 Les KMT (lysine méthyl-transférases)

Les lysines d'histones sont la cible de mono-, di- ou tri-méthylation par des méthyl-transférases lysine-spécifiques (KMT). Les KMT sont majoritairement des protéines contenant un domaine SET capable de catalyser la méthylation des lysines. KMT4 (DOT1L), qui méthyle H3K79, est l'exception qui confirme la règle car elle ne possède pas de domaine SET (Feng Q, 2002, *Curr Biol*). Des études par immuno-précipitation de la chromatine ont permis de cartographier les différentes méthylation des lysines (Barski A, 2007, *Cell*) : les lysines 4, 36 de l'histone H3 sont associées à l'activation transcriptionnelle tandis que les lysines 9 et 27 de l'histone H3 ainsi que les lysines 20 de l'histone H4 corrélaient avec la répression des gènes. Ces observations correspondent à des domaines chromatiniens qui sont enrichis en certaines méthylation de lysines : H3K9me3 et H4K20me3 définissent l'hétérochromatine constitutive, H3K27me3 l'hétérochromatine facultative, enfin H3K4me3 et H3K36me3 sont préférentiellement situées au niveau de l'euchromatine. Cependant, il faut nuancer cette vision car tout n'est pas noir ou blanc, en effet la marque H3K79me3 est retrouvée aussi bien au niveau des promoteurs actifs de l'euchromatine (Steger DJ, 2008, *Mol Cell Biol*) que des promoteurs silencieux de l'hétérochromatine (Jones B, 2008, *PLoS Genet*). Aussi au niveau des « domaines

bivalents », pendant le développement des cellules souches, H3K4me3 et H3K27me3 co-existent au niveau des promoteurs de certains gènes (Bernstein BE, 2006, Cell). Ni l'établissement, ni l'interprétation de ces marques n'ont pour le moment été résolus. Cependant il y a fort à parier qu'elles sont connectées pour contrôler l'expression des gènes d'une manière épigénétique.

2.3.1.1.2 Les KDM (lysine déméthylases)

Longtemps considérée comme irréversible, car la demi-vie des histones méthylées est très proche de la demi-vie des histones (Byvoet P, 1972, Arch Biochem Biophys), la méthylation des histones peut en réalité être très dynamique, en particulier H3K9me3 (Saccani S, 2002, Genes Dev) et H3K27me3 (Plath K, 2003, Science). KDM1/LSD1 (Lysine specific demethylase 1) fut la première KDM spécifique d'histone à avoir été caractérisée. Elle est capable de déméthyliser les lysines 4 et 9 mono- et di-méthylées de l'histone H3 via l'oxydation des amines (Shi Y, 2004, Cell / Metzger E, 2005, Nature). Cette réaction nécessite la présence d'une lysine protonée, c'est pourquoi les lysines tri-méthylées ne peuvent pas être le substrat de cette enzyme.

Une superfamille de protéines contenant le domaine Jumonji (JmjC) est en mesure de retirer les 3 états de méthylation possibles sur les lysines, par hydroxylation du groupement méthyle. C'est ainsi qu'une dizaine de protéines spécifiques de la déméthylation des lysines 4, 9, 27 et 36 de l'histone H3 ont été décrites chez l'humain, représentant seulement le tiers de toutes les protéines contenant un domaine JmjC au sein du protéome. Les déméthylases de certaines modifications comme par exemple H3K27me1, ou les 3 degrés de modification H3K79 n'ont pas encore été identifiées, suggérant l'activité de protéines à JmjC encore méconnues, l'intervention d'une nouvelle classe d'enzymes ou tout simplement une irréversibilité de ces modifications, en particulier pour la transmission des marques épigénétiques à travers les divisions cellulaires.

2.3.1 L'acétylation

Elle affecte uniquement les lysines et est invariablement associée à une activation de la transcription. Les KAT ou Lysine-Acetyl Transférases d'histones catalysent ces réactions. Cette superfamille regroupe 3 familles de protéines : GNAT (Gcn5-related N-acetyl-transferase) (H3 et H4), MYST (Monocytic leukemia zinc finger protein) et CBP/p300 (4 histones) (Sternier DE, 2000, Microbiol Mol Biol Rev). La plupart de ces enzymes sont capables de fixer un résidu acétyle sur de multiples lysines, par exemple KAT1 (HAT1), la première à avoir été identifiée chez la levure (conservée chez les humains), est capable de modifier les lysines 5 et 12 de l'histone H4 *in vitro* (Kleff S, 1995, JBC / Parthun MR, 1996, Cell). La validation *in vivo* de l'acétylation de H4K12 par KAT1 a été mise en évidence aux environs des cassures double-brin de l'ADN (Qin S, 2006, Mol Cell Biol). KAT1 joue également un rôle dans l'assemblage des nucléosomes en acétylant les histones H4 nouvellement synthétisées au sein du cytoplasme.

2.3.1.1 La désacétylation

Elle est catalysée par les Histones-DésAcétylases (HDAC), cette réaction enzymatique a bien entendu un effet opposé à l'acétylation, elle corrèle généralement avec une répression transcriptionnelle. Les 18 HDACs de mammifères sont groupées en 4 classes en fonction de leur

homologie avec les HDACs de levure, leur activité enzymatique et leur localisation cellulaire (Gregoretta, 2004, J Mol Biol). Elles agissent sous forme de complexes multiprotéiques distincts (NuRD, SIN3, Co-Rest) capables de réprimer la transcription de gènes spécifiques, en particulier certains suppresseurs de tumeur dans les cellules cancéreuses. Les HDACs jouent donc un rôle dans le développement tumoral puisque leurs inhibiteurs, les HDACi (butyrate de sodium, trichostatine A, trapoxine, MS-275), peuvent causer la réversion de cellules transformées (Ginsburg E, 1973, PNAS / Altenburg BC, 1976, Exp Cell Res / Boffa LC, 1978, JBC). Les HDACi peuvent bloquer la prolifération cellulaire et induire une entrée en apoptose en réactivant les gènes suppresseurs de tumeurs (Cameron EE, 1999, Nat Genet).

2.3.2 La phosphorylation, les kinases

Les sérines et thréonines des histones peuvent être la cible de kinases. L'ajout d'un groupement phosphate, possédant une charge négative, neutralise la charge basique locale de l'histone, modifiant son affinité avec l'ADN. Le résidu phosphorylé le plus étudié est la sérine 10 de l'histone H3, sa modification conduit généralement à une activation transcriptionnelle. Une mutation de la sérine 10 diminue l'activation des gènes ciblés par KAT2 (PCAF) (Lo WS, 2000, Mol Cell), suggérant la participation de cette modification à l'activation transcriptionnelle par certaines KAT.

La phosphorylation globale des histones H3 correspond à une marque d'entrée en mitose des chromosomes (Gurley LR, 1978, Eur J Biochem). L'histone H3 peut être phosphorylée sur les sérines 10 et 28 (Goto H, 1999, JBC) mais également sur les thréonines 3 et 11 (Shoemaker CB, 1980, JBC / Preuss U, 2003, NAR). Ces 4 sites ont été décrits comme étant phosphorylés au niveau de la chromatine mitotique, et cette caractéristique est conservée de la levure à l'Homme. Les enzymes impliquées dans la phosphorylation de l'histone H3 sont :

- pour S10: Rsk2 (Sassonne-Corsi P, 1999, Science), la Protéine kinase A (PKA) (Salvador LM, 2001, JBC / Frödin M, 2000, EMBO J) et IKK- α (Yamamoto Y, 2003, Nature)
- pour S10 et S28: Aurora B (Goto H, 2002, Genes Cells) et Msk1/Msk2 (Mitogen- and Stress-activated Kinase) (Soloaga A, 2003, EMBO J)
- pour T3: Haspin kinase (Dai J, 2005, Genes Dev)
- pour T11: Dlk/ZIP (death-associated protein-like kinase) (Preuss U, 2003, NAR) et Chk1 (Shimada M, 2008, Cell).

En plus de la mitose, H3S10ph est présent pendant l'interphase en réponse à l'activation des voies de signalisation impliquant les MAP kinases, PKA ou NF-KB.

La caséine kinase II (CKII) phosphoryle H4S1 *in vitro* et *in vivo* chez la levure suite à des dommages à l'ADN, elle serait impliquée dans la réparation suite à une cassure double brin (Cheung WL, 2005, Curr Biol). En réponse à des stimuli apoptotiques, H2B est phosphorylée sur S14 par Mst1 (mammalian sterile twenty) (Cheung WL, 2003, Cell).

Phosphatases

Une famille de phosphatases semble impliquée dans les processus de déphosphorylation des histones : l'enzyme Gcl7 / PP1 (phosphatase 1) agit sur les histones H1 et H3 (Davie JR, 1998, J Cell

Bioch / Spencer VA, 1999, Gene / Davie JR, 2000, Prog Nucl Acid Res Mol Biol), tandis que PP2A cible H3S10ph chez la drosophile (Nowak SJ, 2003, Mol Cell Biol) et PP4 cible la phosphorylation de H2AX sur la sérine 139 (Nakada S, EMBO Rep, 2008 / Chowdury D, 2008, Mol Cell). Enfin, la MAP kinase phosphatase-1 (MKP-1) est responsable de la déphosphorylation spécifique de H3S10ph lors de l'activation des cellules endothéliales humaines par le VEGF et la thrombine (Kinney CM, 2008, AM J Physiol Cell Physiol).

2.3.3 Les ubiquitin-ligases

Bien que leur rôle n'ait pas encore été complètement élucidé, certains résidus lysine peuvent subir une mono-ubiquitylation, à distinguer de la poly-ubiquitylation qui sert au marquage des protéines à dégrader par le protéasome 26S. Cette mono-ubiquitylation implique la fixation d'un peptide de 76 acides aminés (8,6kDa) sur une lysine, faisant intervenir successivement 3 enzymes : E1 pour l'activation (ubiquitin activating enzyme), E2 pour la conjugaison (ubiquitin-conjugating enzyme) et E3 pour la ligation (ubiquitin-protein isopeptide ligase). Ces dernières sont caractérisées par la présence d'un domaine RING. Seules les histones H2A et H2B ont été décrites comme étant mono-ubiquitylées sur les résidus K119 et K120 respectivement chez les mammifères (Osley MA, 2006, Brief Funct Genomic Proteomic). Ces 2 modifications ont des effets opposés sur la transcription. Les protéines Ring1A, Ring1B et 2A-UB (Zhou W, 2008, Mol Cell) sont responsables de la modification de H2A, associée à une répression transcriptionnelle (de Napoles M, 2004, Dev Cell) tandis que le co-activateur transcriptionnel RNF20 (BRE1) catalyse la réaction sur H2B (Kim J, 2005, Mol Cell). D'autre part, le complexe composé des protéines CUL4A, CUL4B, DDB1, DDB2 et ROC1 (RBX1) est capable d'ubiquityler les histones H3 et H4, affectant la stabilité du nucléosome et leur éviction de la chromatine suite à un dommage à l'ADN (Wang H, 2006, Mol Cell).

Ubiquitines protéases

Les ubiquitines peuvent être retirées par une classe de thiol-protéases connues sous le nom d'ubiquitine-spécifique protéases (USP). Jusqu'à présent, seules USP3 (Nicassio F, 2007, Curr Biol) et USP22 (Zhang XY, 2008, Cell Cycle / Zhang XY, 2008, Mol Cell) ont été décrites comme pouvant désubiquityler les deux histones H2A et H2B. En revanche, d'autres enzymes telles que USP16 (Ubp-M), USP21 et 2A-DUB (MYSM1) sont spécifiques de l'histone H2A (Joo HY, 2007, Nature / Nakagawa T, 2008, Genes Dev / Zhu P, 2007, Mol Cell).

2.3.4 SUMO ligases

Le processus est très similaire à celui de l'ubiquitylation, faisant intervenir plusieurs enzymes en cascade (E1, E2 et E3) pour assurer la fixation covalente d'un ou plusieurs motifs SUMO (Small Ubiquitin like Modifier) sur une lysine accepteuse. SUMO est composé d'une centaine d'acides aminés (12kDa). L'histone H4 a été la première histone de mammifère à avoir été décrite comme pouvant être sumoylée *in vitro* et *in vivo*, cette modification catalysée par Ubc9 est associée à une répression transcriptionnelle (Shiio Y, 2003, PNAS) à travers le recrutement d'HDAC et de HP1 (Heterochromatin Protein 1). Cependant les 4 histones constituant le nucléosome sont sumoylées chez la levure, bien que la modification ne semble pas être égale entre les différents homologues, et

que les toutes les lysines n'aient pas été identifiées: K6-7, K16-17 pour H2B, K126 pour H2A et les 5 lysines N-terminales de H4 (Nathan D, 2006, Genes Dev).

SUMO protéases

Il existe une famille de SUMO-spécifique protéases appelées ULP (Ubiquitin-like protein specific protease) chez la levure ou SENP (SENtrin-spécifique Protéases) chez les mammifères (Mukhopadhyay D, 2007, TRENDS in Bioch Sci). Aucune de ces protéines n'a, à ce jour, été impliquée dans le retrait d'un motif SUMO sur les histones, mais compte tenu de leur activité et de leur localisation nucléaire principalement, il est probable que ces protéines jouent un rôle au niveau de l'organisation de la chromatine en ciblant les histones modifiées.

2.3.5 ADP-ribosylation

Identifiées depuis 1968 sur les histones (Nishizuka Y, 1968, JBC), les ADP-ribosylations ciblent les arginines, les acides glutamiques et les acides aspartiques. Elles sont catalysées par les MARTs (mono-ADP-ribosyl-transférases) et les PARPs (Poly-ADP-ribose polymerases) qui convertissent un NAD⁺ (nicotinamide adenine dinucléotide) en poly-ADP-ribose (Amé JC, 2004, Bioessays / Faraone-Mennella MR, 2005, Biochem Cell Biol). On en dénombre 18 chez les humains, elles possèdent des structures et des fonctions très différentes au sein de la cellule. H2B et H1 sont les substrats histones préférentiels *in vivo* de PARP1 et PARP2 (Huletsky A, 1989, JBC), mais H4 peut, dans une moindre mesure, être également modifiée (Golderer G, 1991, Biochem J). Ces modifications entraînent une relaxation du nucléofilament.

Les PARG

La PARG (poly-ADP-ribose glycohydrolase) est la principale enzyme qui catabolise l'ADP-ribosylation en coupant les liaisons ribose-ribose (Kraus WL, 2003, Cell). Une autre enzyme appelée ARH3 ou ADPRHL2 (ADP-ribosylhydrolase like 2) est également capable de jouer ce rôle (Mueller-Dieckmann C, 2006, PNAS). De plus, les phosphodiésterases, qui hydrolysent les liaisons phosphodiesters, peuvent également potentiellement dégrader cette modification. Cette modification est associée à la réparation des dommages à l'ADN aussi bien qu'à l'induction de la mort cellulaire. Elle affecte également le remodelage de la chromatine, l'acétylation des histones et même l'empreinte parentale à travers la modulation de la méthylation de l'ADN (Klenova E, 2005, Cell Cycle).

2.3.6 L'isomérisation des prolines (pas de liaison covalente)

Au sein d'une chaîne polypeptidique, les prolines existent sous la conformation *cis* ou *trans*. Catalysée par une peptidyl-prolyl-isomerase, ce changement engendre un changement de conformation de la proline, ayant des conséquences sur les repliements des acides aminés voisins. Chez la levure, l'enzyme Fpr4 est responsable de la *cis-trans* isomérisation de la proline P38 de l'histone H3 (Nelson CJ, 2006, Cell). La « bonne » isomérisation de cette proline est indispensable pour avoir une méthylation de H3K36 par KMT3 (SET2). Par conséquent, cette modification a été considérée comme une nouvelle modification potentielle des histones bien qu'elle n'implique pas l'ajout d'un motif par liaison covalente.

2.3.7 Hypothèse du « code histone »

Comme nous l'avons défini dans les précédents paragraphes, plusieurs modifications peuvent être impliquées dans un événement unique (réparation après une cassure double brin de l'ADN) ou nécessaires à la régulation de la transcription. Une modification unique est rarement la seule responsable d'un événement cellulaire. Dès lors, il a été suggéré depuis 2000 que la combinaison de toutes les modifications post-traductionnelles des histones constituerait un code, le « code histone » (Strahl BD, Nature, 2000).

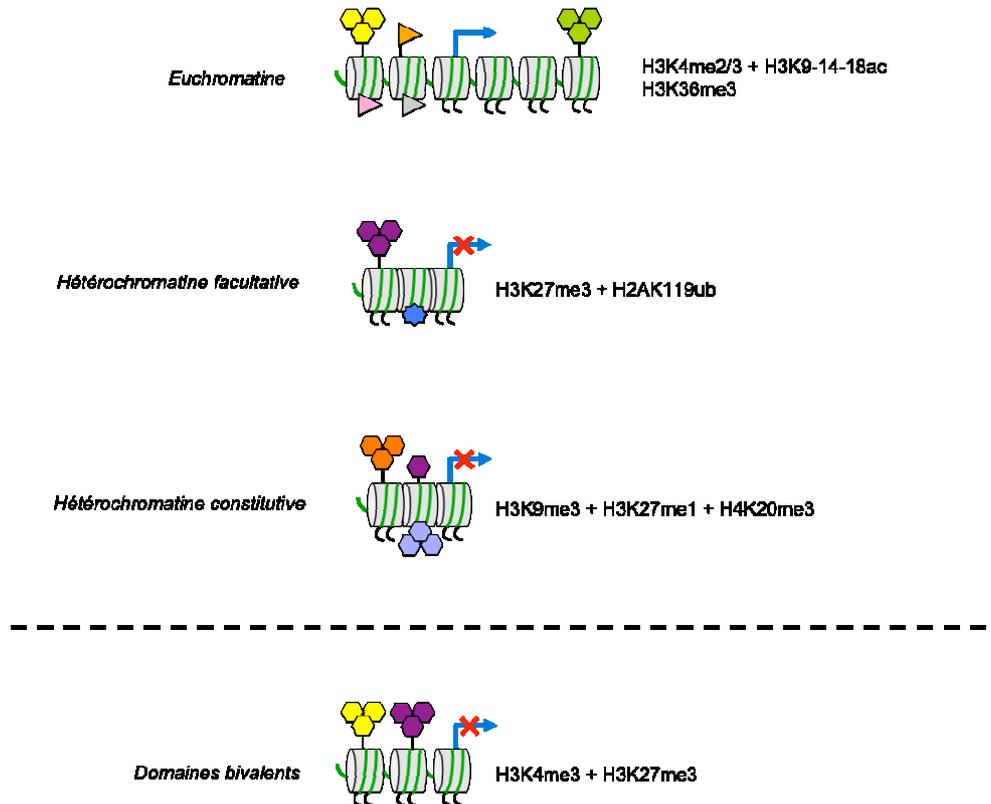


Figure 3 – Représentation des principales modifications post-traductionnelles admises comme participant au code histone au niveau de l'euchromatine, des hétérochromatines et des domaines bivalents

Les différentes modifications décrites précédemment peuvent parfois intervenir sur le même acide aminé, il existe donc une compétition entre certaines modifications. A contrario, certaines modifications sont complémentaires, la première étant indispensable pour l'établissement de la seconde. Enfin, certaines modifications empêchent l'accès des résidus environnants, empêchant leur modification donc certaines modifications sont mutuellement exclusives.

Le nombre de combinaisons possibles sur une unique queue d'histone est déjà très élevé, donc la « complexité » du code au sein d'un nucléosome apparaît extrêmement sophistiquée. C'est pourquoi certains chercheurs, ayant réalisé des études des modifications d'histones à l'échelle de génomes (levure, drosophile), ont remis en question l'existence de ce code (Liu CL, 2005, PLoS Biology / Schübeler D, 2004, Genes Dev / Dion MF, 2005, PNAS). Cependant, d'autres études similaires: chez la levure (Pokholok DK, 2005, Cell), chez la souris (Mikkelsen TS, 2007, Nature) et chez l'Homme (Barski A, 2007, Cell) ont permis de renforcer l'idée que ce code doit bien exister.

L'invalidation de certaines des enzymes responsables de ces modifications ne permet pas toujours d'obtenir des animaux *viables*, soulignant leur importance au sein de l'organisme. Aussi, de nombreuses modifications, présentes de manière stable (héritable comme le sont supposées la méthylation de l'ADN ou certaines modifications d'histones (Dodd IB, 2007, Cell)) ou transitoire, sont spécifiques d'un événement biologique particulier ou d'une région chromatiniennne définie. Le rôle des différentes enzymes impliquées pour marquer le code n'est donc pas redondant. Bien entendu, les histones ne sont probablement pas les substrats exclusifs de ces enzymes, et comme les protéines non-histones peuvent également subir des modifications post-traductionnelles, le concept du « code protéine » a même été suggéré (Sims RJ 3rd, 2008, Nat Rev Mol Cell Biol). En effet, les modifications post-traductionnelles sont responsables de la régulation de l'activité de nombreuses protéines comme par exemple le facteur de transcription p53 qui, comme les histones, peut subir un grand nombre de modifications post-traductionnelles, par les mêmes enzymes. Son acétylation par des KAT permet son activation, qui stabilise sa fixation à l'ADN et favorise son interaction avec ses co-facteurs (Luo J, 2004, PNAS). Sa méthylation peut activer (K372me1 par SET9) ou réprimer (K370me1 par SMYD2, K382me1 par PR-SET7) son activité en fonction de l'acide aminé modifié. D'autres modifications post-traductionnelles d'histones non encore décrites seront très probablement à prendre en considération dans le futur. Récemment la formylation des histones, localisée au niveau des queues ou du noyau a représenté une des modifications « émergentes » qui pouvait prévenir d'une méthylation ou une acétylation des lysines (Wisniewski JR, 2008, NAR).

2.4 Les variants d'histones

La plupart des histones, briques essentielles à la formation du nucléosome, possèdent des variants qui permettent non seulement une variabilité dans la composition des nucléosomes, mais également d'étendre les possibilités du code histone. De nombreux variants des histones H2A, H2B et H3 ont été découverts et les rôles biologiques de certains ont été définis avec précision. Aucun variant de l'histone H4 n'a pour le moment été identifié. Les fonctions biologiques dans lesquelles ces variants ont été impliqués concernent aussi bien l'activation que la répression transcriptionnelle, en passant par la détection des dommages à l'ADN et la mise en place des centromères. Les variants d'histones ont un grand potentiel pour altérer la structure des nucléosomes *via* leur propre séquence ou les modifications post-traductionnelles spécifiques qu'ils peuvent également subir.

2.4.1 Histones H2A

Elles possèdent au moins 5 variants chez les mammifères qui ont des fonctions très diversifiées. H2A.X possède une séquence très proche de H2A, cependant il contient un motif unique SQ au niveau de son extrémité C-terminale. Suite à une cassure double-brin de l'ADN ou un dommage à l'ADN, la sérine de ce motif (S139) est phosphorylée par la famille PIKKs (phosphatidylinositol-3-OH-kinase like family) dont les 3 membres majoritaires sont: ATM (ataxia-telangiectasia mutated), ATR et DNA-PKc (Burma S, 2001, JBC / Wang H, 2005, J Cell Physiol). Les H2A.X phosphorylées forment des foci distincts au niveau des chromosomes interfasciculaires, représentant les sites de cassures double-brin de l'ADN.

D'un côté les H2A.Z sont localisés au niveau de régions où les gènes sont réprimés, en particulier au niveau de l'hétérochromatine péricentrique (Rangasamy D, 2003, EMBO J). De plus, ils participent à la ségrégation des chromosomes (Rangasamy D, 2004, Nat Struct Mol Biol). D'un autre, ils sont retrouvés dans les régions promotrices de gènes transcriptionnellement actifs (Brickner DG, 2007, Plos Biol). Ils contiennent 5 lysines acétylables : 4, 7, 11, 13 et 15 (Dryhurst D, 2004, Biochem Cell Biol), les lysines 120 et 121 peuvent être ubiquitinylées par Ring1B (Sarcinella E, 2007, Mol Cell Biol), marque de leur localisation au niveau de l'hétérochromatine facultative.

MacroH2A est également retrouvé au niveau de la chromatine condensée du chromosome X inactif (Barr-body) chez les femelles mammifères (Costanzi C, 1998, Nature), dans les foci d'hétérochromatine associés à la sénescence (Zhang R, 2005, Dev Cell), les régions riches en îlots CpG méthylés dont de nombreux gènes soumis à l'empreinte parentale (Choo JH, 2006, Hum Mol Genet). Chez les mammifères, 2 gènes séparés codent pour macroH2A1 et macroH2A2 et le premier peut générer 2 variants par épissage alternatif : macroH2A1.1 et macroH2A1.2. Bien que leur rôle précis n'ait pas encore été décrit, des modifications post-traductionnelles de macroH2A ont été identifiées : l'ubiquitinylation de K115, la méthylation de K17, K122 et K238 et la phosphorylation de T128 (Chu F, 2006, MCP).

Enfin H2ABbd (Barr-body deficient), comme son nom le suggère, n'est pas retrouvé au niveau du chromosome X inactif, il a même été rapporté qu'il était présent dans les régions riches en H4 acétylées, suggérant un rôle dans l'activation transcriptionnelle (Chadwick BP, 2001, J Cell Biol).

2.4.2 Les variants de H2B ont une préférence sexuelle

Le variant spH2B est spécifique du sperme, il a été découvert dans un complexe associé aux télomères (Gineitis AA, 2000, J Cell Biol).

Le variant hTSH2B est, quant à lui, spécifique des testicules et du sperme (Zalensky AO, 2002, JBC). Il possède une caractéristique remarquable, il possède une phénylalanine en lieu et place de la sérine 14 de H2B, qui est phosphorylée par la kinase Mst1 pendant l'apoptose (Kimmins S, 2005, Nature). Peut-être que hTSH2B prévient les cellules du sperme de rentrer en apoptose.

H2BFWT est exclusivement exprimé dans les testicules, semble être enrichi au niveau des télomères (Churikov D, 2004, Genomics), ce variant pourrait être un marqueur de l'identité télomérique nécessaire pour la transmission des structures télomériques à travers des générations successives.

2.4.3 Les variants de H3 : l'hypothèse du code-barre d'histones

Un des variants les plus spécialisé est la protéine CENP-A (centromeric protein 1, Cen-H3), sa position au niveau des chromosomes détermine l'identité centromérique (Blower MD, 2001, Nat Cell Biol / Sullivan KF, 2001, Curr Opin Genet Dev). Il est phosphorylé par Aurora B au niveau de S7 lors de la mitose, et sa localisation est distincte des H3 phosphorylées en S10 (Zeitlin SG, 2001, J Cell Biol), il peut également être ADP-ribosylé (Saxena A, 2002, JBC).

H3.3 est retrouvé au niveau de l'euchromatine, son recrutement au niveau de loci transcriptionnellement actifs a notamment été suivi en temps réel dans des cellules humaines (Janicki SM, 2004, Cell). Une forme phosphorylée (S31) a été décrite, présente à proximité des centromères

pendant la métaphase (Hake SB, 2005, PNAS). H3.1 (K9me2, K14ac) et H3.2 (K27me2, K27me3) sont respectivement associés à l'hétérochromatine constitutive et à l'hétérochromatine facultative (Hake SB, 2006, JBC). Ces 3 variants de l'histone H3, retrouvés dans des régions chromatiniennes spécifiques, font l'objet de l'hypothèse du code-barre d'histones (Hake SB, 2006, PNAS).

Les variants d'histones permettent à la cellule de changer la composition en protéines des nucléosomes, permettant d'accroître les possibilités au niveau des modifications post-traductionnelles, mais également de modifier la structure du nucléosome, modifiant son interaction avec l'ADN, à tel point qu'après le « code histone » en 2000, le « code nucléosome » a également été suggéré en 2006 (Bernstein E, 2006, Biochem Cell Biol). Il permettrait de couvrir à la fois les variants d'histones et leurs modifications post-traductionnelles respectives pour expliquer les différents états chromatiniens.

2.5 Les ARN non codants

Depuis quelques années, il est devenu évident que les ARN, en particulier les non codants, jouaient un rôle dans le contrôle des phénomènes épigénétiques multiples (Bernstein E, 2005, Genes Dev). L'exemple le plus connu réside dans le mécanisme de compensation de dosage réalisé par le Xist pour l'inactivation d'un des 2 chromosomes X chez les mammifères femelles. Chez la plupart des eucaryotes, 2 voies d'interférence par ARN (ARNi) co-existent, les PTGS (post-transcriptional gene silencing) et les TGS (transcriptional gene silencing) qui agissent pendant ou après la transcription. Ces ARN agissent très souvent de concert avec différents composants associés à la chromatine et à la machinerie de méthylation de l'ADN pour assurer une extinction stable des gènes. Les TGS (Xist, repeat-associated siRNA) peuvent éteindre certains gènes sur du long-terme, et même à travers les divisions cellulaires, donc ils peuvent être considérés comme une marque épigénétique, contrairement aux PTGS (microRNA, siRNA...).

3 Les effecteurs

Comment obtenir un événement biologique à partir d'une simple modification ? Deux modèles co-existent, le premier met en lumière la stabilisation ou la déstabilisation (phosphorylation, acétylation) des propriétés physico-chimiques d'interaction entre l'ADN et les histones (Cosgrove MS, 2004, Nat Struct Mol Biol). Le second modèle fait intervenir des « effecteurs » qui seraient capables de lire les modifications post-traductionnelles, facilitant les événements biologiques en aval *via* le recrutement ou la stabilisation des complexes agissant sur la chromatine (Seet BT, 2006, Nat Rev Mol Cell Biol / Ruthenburg AJ, 2007, Mol Cell).

En effet, à l'image du « code génétique » qui est admis et reconnu, au cours duquel l'information portée par l'ADN est lue par l'ARN polymérase au cours de la transcription pour former l'ARN messenger, qui lui-même est lu par les ribosomes pour synthétiser les protéines. Il est aujourd'hui évident que les différentes modifications post-traductionnelles présentes au niveau des nucléosomes, qui constituent le « code histone », sont lues par des protéines possédant des domaines protéiques particuliers, capables de déchiffrer et d'interpréter le code en interagissant avec d'autres protéines ou en possédant une activité enzymatique particulière (Cheung P, 2000, Cell). Plusieurs domaines

protéiques sont capables de se fixer spécifiquement sur des histones modifiées, d'une manière dépendante à la modification et/ou à la séquence.

3.1 Lysines acétylées

Bien que certaines lysines acétylées affectent les propriétés physiques des nucléosomes (Wolffe AP, 1999, NAR) ou perturbent leur association (Shogren-Knaak M, 2006, Science), elles peuvent être interprétées par les effecteurs possédant un **bromodomaine** (Dhalluin C, 1999, Nature). Les bromodomains sont retrouvés dans des protéines qui sont associées à la chromatine, en particulier certaines KAT et les composants de certains complexes de remodelage de la chromatine. Leur présence joue un rôle clé dans l'activité de ces protéines (Zeng L, 2002, FEBS Lett). La KAT de levure Gcn5p se fixe sur la lysine 16 acétylée de l'histone H4, et compte tenu de son activité, cela suggère qu'elle participe à la diffusion de l'acétylation sur la chromatine (Owen DJ, 2000, EMBO J). Plusieurs protéines contiennent de multiples bromodomains, c'est le cas de TAF1, une de protéines nécessaire au recrutement de la machinerie de transcription. TAF1, qui possède un tandem de bromodomains, est capable de se fixer sur une deux lysines acétylées simultanément, en particulier les lysines 5 et 12 de l'histone H4 (Jacobson RH, 2000, Science).

3.2 Les lysines méthylées

Comme détaillé précédemment, les lysines peuvent-être mono-, di- et tri-méthylées *in vivo* en fonction de l'activité antagoniste des KMT et des KDM. D'un point de vue fonctionnel, la méthylation des histones a été associée tantôt à une activation (H3K4me3), tantôt à une répression transcriptionnelle (H3K9me3, H3K27me3). Avec l'ajout d'un ou plusieurs groupements méthyles, l'hydrophobicité de la lysine augmente de manière inversement proportionnelle à sa capacité à créer des ponts hydrogènes. Deux classes de repliements protéiques sont capables de lier les lysines méthylées : d'une part les membres de la superfamille Royal (Maurer-Stroh S, 2003, Trends Biochem Sci), d'autre part les protéines à doigts PHD (Pena PV, 2006, Nature / Li H, 2006, Nature).

3.2.1 La superfamille Royal

3.2.1.1 Le chromodomaine

Il reconnaît les di- et tri-méthyl-lysines au niveau de H3K9 et H3K27. Le domaine CHROMO (CHRomatin Organization MOdifier domain) des protéines HP1 est capable de se fixer à H3K9me3 et H3K9me2, ce qui n'est pas le cas de H3K9 non modifié (Jacobs SA, 2002, Science / Nielsen PR, 2002, Nature). En revanche le chromodomaine de la protéine Polycomb (drosophile) cible de manière spécifique H3K27me3 (Fischle W, 2003, Genes Dev / Min J, Genes Dev, 2003). Toutes les protéines possédant un chromodomaine ne sont pas capables d'interagir avec les lysines méthylées, suggérant que les séquences extérieures à ce domaine sont cruciales pour la reconnaissance ou la fixation.

Il est pensé que la fonction première de fixation du chromodomaine sur les méthyl-lysines qui sont associées avec l'hétérochromatine répressive est de stabiliser passivement la conformation dense de la fibre chromatinienne par attachement des nucléosomes entre eux (Nielsen AL, 2001, Mol Cell /

Zhao T, 2000, JBC). Cette fonction de « pontage » n'a pas été confirmée, et l'intervention d'autres protéines dans la répression dépendante de HP1 et Polycomb est plus que probable.

Les hélicases d'ADN appelées CHD (Chromo Helicase DNA-binding) régulent l'assemblage des nucléosomes et sont mobilisées au niveau des régions transcriptionnellement actives (Lusser A, 2005, Nat Struct Mol Biol). Elles possèdent un double chromodomaine à leur extrémité N-terminale et celui de CHD1 possède une affinité pour H3K4me2/me3, aucune pour cette même lysine non modifiée (Flanagan JF, 2005, Nature).

3.2.1.2 Le double-tudor domaine

La lysine déméthylase KDM4A (JMJD2A/JHDM3A) possède, en plus de ses domaines JmjC et JmjN indispensables à son activité catalytique sur H3K9 et H3K36, un double-tudor domaine (Klose RJ, 2006, Nature). Ce dernier est en mesure de reconnaître H3K4me3, suggérant un rôle dans la régulation des régions euchromatiques. Des études complémentaires *in vitro* suggèrent qu'il est également capable de se fixer sur H4K20me3 (Kim J, 2006, EMBO Rep), marque retrouvée au niveau de l'hétérochromatine constitutive, des études *in vivo* sont nécessaires pour valider ce résultat.

3.2.1.3 Le tandem-tudor domaine

53BP1 (p53 binding protein 1) est une protéine qui est recrutée au niveau des cassures double brin suite à des dommages à l'ADN (Huyen Y, 2004, Nature). Cette protéine contient un tandem-tudor domaine capable de se fixer sur H4K20me1 et H4K20me2, mais pas sur H4K20 non modifié, ni H4K20me3 (Botuyan MV, 2006, Cell).

Les double- et tandem-tudor domaines possèdent des repliements qui leur sont spécifiques, expliquant la différence d'affinité entre ces 2 domaines de nature voisine.

3.2.1.4 MBT

SCML2 (Sex Comb on Midleg-like 2) appartient à la famille Polyhoméotique, impliquée dans la régulation négative des gènes qui contrôlent le développement. Cette protéine possède 2 domaines MBT (Malignant Brain Tumor) qui ont une structure proche des chromodomains (Sathyamurthy A, 2003, JBC). Des études récentes ont mis en évidence une affinité préférentielle de ces domaines pour des lysines mono-méthylées, sans spécificité de séquence (Santiveri CM, 2008, J Mol Biol).

L3MBTL1 (Lethal 3 malignant brain tumor 1) contient 3 répétitions du motif MBT, c'est un répresseur transcriptionnel (Boccuni P, 2003, JBC) capable de compacter la chromatine *in vitro* (Trojer P, 2007, Cell). Il est recruté préférentiellement au niveau des lysines 20 mono-méthylées des histones H4, de plus il interagit directement avec KMT5A (PR-SET7), la mono-méthyl-transférase de H4K20, qui semble être une protéine clé pour la répression induite par L3MBTL2 (Kalakonda N, 2008, Oncogene).

3.2.2 PHD (Plant HomeoDomain)

BPTF est une sous-unité du complexe NURF, un des principaux complexes capable de remodeler la chromatine. Un des deux doigts PHD de BPTF interagit exclusivement avec H3K4me3, et pas d'autres modifications (Wysocka J, 2006, Nature), permettant le recrutement de NURF et l'activation transcriptionnelle. L'autre domaine PHD reconnaît H3K4me2 ou H3K4me3 mais pas la forme nue ou mono-méthylée (Li H, 2006, Nature). Les PHD de ING2 (Pena PV, 2006, Nature) et TAF3 (Vermeulen M, 2007, Cell) reconnaissent également H3K4 triméthylé.

3.3 Les phosphorylations

3.3.1 14-3-3

Les protéines 14-3-3 constituent une famille de protéines ubiquitaires régulatrices possédant des modules de fixation sur les sérines ou les thréonines phosphorylées qui sont très conservées (Yaffe MB, 1997, Cell). Leur association avec l'histone H3 a été démontrée au niveau des phospho-sérines 10 et 28 (Macdonald N, 2005, Mol Cell).

3.3.2 BRCT

La phosphorylation de la sérine 139 du variant H2A.X est reconnue par les domaines BRCT de la protéine MDC1 (Mediator of DNA damage checkpoint protein 1) suite à des dommages à l'ADN (Stucki M, 2005, Cell).

3.4 Domaines de liaison à l'ubiquitine

Jusqu'à présent, aucune protéine n'a été impliquée dans la reconnaissance des mono-ubiquitylation de H2A et H2B. Cependant, le protéome des cellules eucaryotes regorge de protéines possédant des domaines capables de se lier sur ce type modifications (mono- ou poly-). En effet, plus de 15 domaines différents ayant cette particularité existent (Hurley JH, 2006, Biochem J), donc une ou plusieurs protéines sont potentiellement capables de se fixer sur les histones ubiquitylées.

HDAC6 possède un doigt de zinc PAZ (polyubiquitin-associated zinc binding) capable de reconnaître *in vitro* une poly-ubiquitylation et non une mono- (Hook SS, 2002, PNAS).

3.5 SIM (SUMO-Interacting Motifs)

Bien que le motif SUMO soit de structure proche de l'ubiquitine, les domaines de liaison à cette dernière sont incapables de reconnaître SUMO. En revanche, une vingtaine de protéines contenant le domaine SIM, divisées en 2 classes (I = RanBP2 et II=PIAS) en fonction de la structure tridimensionnelle du motif de reconnaissance, sont capables de reconnaître spécifiquement le motif SUMO (Hecker CM, 2006, JBC). Aucun lien avec les histones modifiées n'a été décrit pour le moment.

3.6 Domaines de liaison à l'ADP-ribose

Une étude récente a permis de révéler la présence de plusieurs dizaines de protéines capables de se lier à l'ADP-ribose dans le protéome de cellules humaines (dont celles possédant un domaine macro comme le variant d'histone macroH2A (Kustatscher G, 2005, Nat Struct Mol Biol)). Leur spécificité de substrat n'ayant pas été déterminée, nous ne pouvons pas conclure à une interaction directe avec les histones, mais nous ne pouvons pas exclure l'existence de telles protéines compte tenu des processus biologiques dans lesquels certaines sont impliquées : topoisomérases d'ADN, hélicases d'ADN, protéines à doigts de zinc... (Gagné JP, 2008, NAR).

3.7 Les résidus non modifiés = nouveau point de contact avec la chromatine

3.7.1 Le PHD de AIRE (AutoImmune REgulator)

Ce domaine interagit directement avec H3K4 non méthylée (me0) et active les promoteurs contenant des niveaux faibles de H3K4me3 (Org T, 2008, EMBO Rep). Le PHD de BHC80, un co-répresseur transcriptionnel qui appartient au même complexe que KDM1 (LSD1), se fixe spécifiquement sur les queues des histones H3 non modifiées, la lysine 4 en particulier (Lan F, 2007, Nature).

3.7.2 Le domaine ADD (ATRX-DNMT-DNMT3L) riche en cystéines

Il est très proche du domaine PHD et possède une très forte affinité pour H3K4me0, comme l'ont démontrées les études réalisées avec Dnmt3L. Ce qui a permis de proposer un lien entre l'absence de méthylation de H3K4 et la méthylation *de novo* de l'ADN des mammifères (Ooi SK, 2007, Nature).

3.7.3 WD40

La définition des sites d'ancrage du domaine WD40 est très controversée car il semble capable de se fixer aussi bien sur les histones H3 nues (H3R2) (Couture JF, 2006, Nat struct Mol Biol / Han Z, 2006, Mol cell / Ruthenburg AJ, 2006, Nat Struct Mol Biol / Schuetz A, 2006, EMBO J) que modifiées (H3K4me2) (Wysocka J, 2005, Cell) ainsi que sur les histones H4 (Song JJ, 2008, Genes Dev). Les protéines à domaine WD40 révèlent ainsi une diversité au niveau de leur recrutement sur le nucléosome.

3.8 Domaines de liaison à l'ADN méthylé

Les protéines à domaine MBD (Methyl-CpG-Binding Domain) sont capables de reconnaître et de se fixer spécifiquement sur l'ADN méthylé. Elles constituent un site d'ancrage pour des enzymes ayant une activité sur les histones comme les KAT par exemple, qui permet la désacétylation du promoteur du gène ciblé et une répression transcriptionnelle. Les effecteurs subissent également des modifications post-traductionnelles : MBD2 peut-être méthylée par PRMT5, ce qui inhibe sa capacité à se fixer sur l'ADN méthylé et à réprimer la transcription (Tan CP, 2006, Mol Cell Biol).

La Figure 3 présente d'une manière schématique la mise en place du code histone et ses conséquences sur le destin cellulaire. Les modifications des histones modulent directement ou indirectement, *via* les effecteurs, la structure de la chromatine, ce qui permet de contrôler les fonctions cellulaires associées à l'ADN, telles que la transcription, la réplication ou l'extinction des gènes. La combinaison des différents états chromatinien de chacun des gènes de la cellule va conduire le destin cellulaire : prolifération, apoptose ou différenciation.

Les mécanismes épigénétiques semblent donc essentiels au développement harmonieux d'un organisme et une dérégulation ou une expression anarchique des enzymes impliquées dans ces phénomènes peuvent entraîner des pathologies, en particulier le cancer, ouvrant la voie à de nouveaux traitements, les thérapies épigénétiques (Wang GG, 2007, TRENDS in Molecular Medicine).

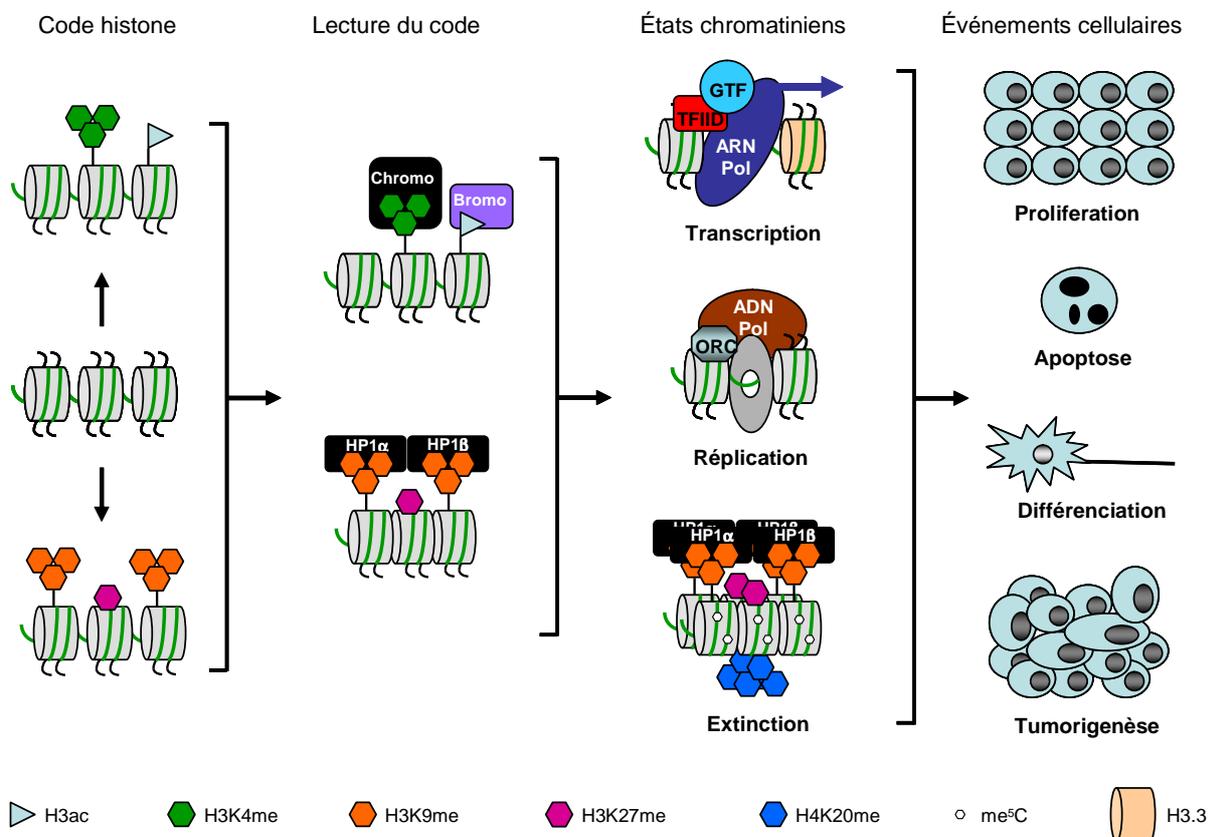


Figure 4 - Contrôle épigénétique du destin cellulaire. Les modifications post-traductionnelles des histones sont reconnues par des domaines protéiques particuliers, permettant le recrutement de complexes ayant des activités spécifiques sur la chromatine, activant ou réprimant les gènes ciblés.

4 Cancer et thérapie épigénétique/épidrogues

Beaucoup de gènes codant pour les enzymes qui modifient les histones sont réarrangés, amplifiés ou mutés dans de nombreux cancers (Schneider R, 2002, Trends Biochem Sci).

Le faible niveau d'ADN méthylé dans les tumeurs, en comparaison au niveau de méthylation de l'ADN dans des tissus sains fut l'une des premières altérations épigénétiques observée dans les cancers humains (Feinberg AP, 1983, Nature). Au cours du développement d'un néoplasme, le degré global

d'hypo-méthylation de l'ADN génomique augmente autant que la lésion progresse d'un état bénin de prolifération à un cancer invasif (Fraga MF, 2004, Cancer Res). La contribution de l'hypo-méthylation déstabilise les chromosomes, réactive les éléments transposables et efface les empreintes génétiques. Inversement, certaines régions non méthylées en cellules saines, deviennent hyper-méthylées dans les cancers, en particulier les îlots CpG présents au niveau des promoteurs des gènes suppresseurs de tumeurs, afin de les inactiver : Rb (Greger V, 1989, Hum Genet), p16INK4a (Merlo A, 1995, Nat Med) et BRCA1 (Esteller M, 2000, J Natl Cancer Inst). Ces hyper-méthylations sont spécifiques du type de cancer (Costello JF, 2000, Nat Genet).

L'hyper-méthylation des îlots CpG dans les régions promotrices des gènes suppresseurs de tumeurs dans les cellules cancéreuses est associée à une combinaison de marques sur les histones : désacétylation des histones H3 et H4, perte de la tri-méthylation de H3K4, et gain de la méthylation de H3K9 et tri-méthylation de H3K27 (Jones PA, 2007, Cell). La présence d'histones H3 et H4 hyper-méthylées et hypo-acétylées au niveau de certains CpG entraîne l'extinction du gène, en dépit de l'absence d'une hyper-méthylation de l'ADN (Richon VM, 2000, PNAS). Ces modifications post-traductionnelles des histones s'accumulent au fur et à mesure du développement de la tumeur. Les modifications sont une conséquence, mais les enzymes qui modifient les histones en sont la cause, et un grand nombre de ces enzymes possède des expressions anarchiques dans les tissus tumoraux (Ozdogan H, 2006, BMC Genomics), comme KMT6 (Varambally S, 2002, Nature), KMT4 (Okada Y, 2005, Cell), KMT2B, NSD3 ou la déméthylase KDM4C (Cloos PA, 2006, Nature).

Le point crucial des maladies épigénétiques réside dans la réversibilité des modifications, en effet les maladies induites par des mutations dans la séquence d'ADN ne sont pas réversibles alors que, nous l'avons vu précédemment, la grande majorité des modifications épigénétiques sont dynamiques. Il apparaît aujourd'hui plus évident et plus simple d'arriver à modifier le comportement d'une enzyme particulière plutôt que de modifier la séquence ADN par thérapie génique.

Il est désormais admis que de nombreuses maladies et pathologies du comportement résultent du dysfonctionnement des gènes. Bien que le cancer soit le plus étudié, d'autres pathologies comme les maladies auto-immunes, l'asthme, le diabète de type 2, les désordres métaboliques, et l'autisme sont la conséquence d'une expression aberrante de certains gènes. La fonction des gènes peut être altérée par un changement dans la séquence ADN ou au niveau de son programme épigénétique en absence d'une mutation. Avec l'aide des épidroques, il est maintenant possible de contrecarrer les profils d'expression géniques aberrants associés à différentes maladies. Plusieurs de ces drogues ciblant les enzymes responsables de la méthylation de l'ADN et de la désacétylation des histones ont été approuvées par la FDA (Food and Drug Administration) américaine et sont actuellement en essais cliniques avec un bon pronostic pour la régression tumorale. Comprendre les machineries épigénétiques et les différents rôles de leurs composants dans des maladies spécifiques est essentiel pour le développement de thérapies épigénétiques. A tel point qu'un projet international pour l'étude de l'épigénome humain a vu le jour en août 2008 à l'initiative de tous les « grands noms » du domaine (AHEAD, Nature, 2008). Aussi, à l'échelle européenne, le consortium EPITRON, réunit plusieurs laboratoires qui, en collaboration avec les industries pharmaceutiques, essaient de développer des épidroques.

Introduction sur les CBX

Les protéines des familles HP1 et Polycomb sont codées par une classe de gènes connus sous le nom de gènes chromobox (CBX). Chez les mammifères, il y a 3 protéines distinctes dans la famille HP1: CBX5 (HP1 α), CBX1 (HP1 β) et CBX3 (HP1 γ) et 5 protéines dans la famille Polycomb : CBX2, CBX4, CBX6, CBX7 et CBX8 (Figure 4). Les protéines CBX sont généralement associées aux mécanismes d'extinction de nombreux gènes et dans la compaction de la chromatine. Les HP1 et les Polycomb sont respectivement associées aux hétérochromatines constitutive et facultative.

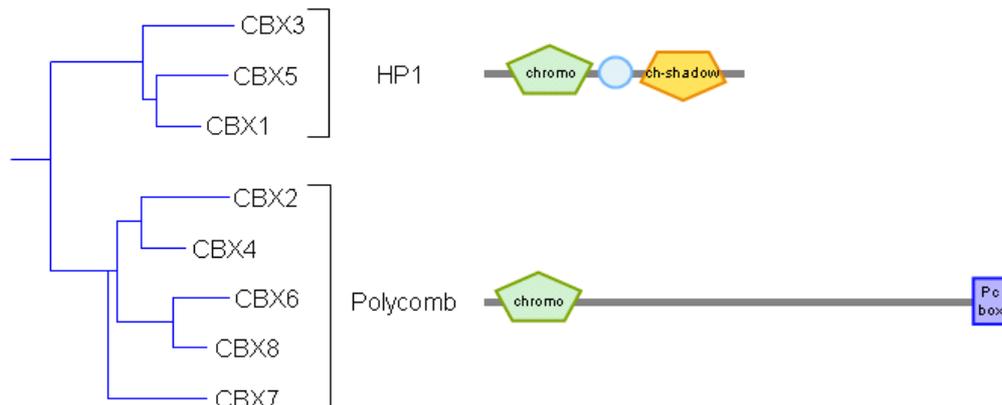


Figure 5 - Arbre phylogénétique des protéines CBX humaines et représentation de leur structure. Alignement réalisé avec CLUSTALW sur le chromodomaine.

Les protéines à chromodomaine

Le chromodomaine a originellement été défini comme une région de 37 résidus homologues partagés par 2 polypeptides chez la drosophile : HP1 et Polycomb (Paro R, 1991, PNAS). Ce motif est identique à 77% entre les 2 protéines. Cependant, il a été graduellement agrandi à 50 acides aminés et un grand nombre de protéines à chromodomaine possédant une région plus large de similarité ont pu être regroupées en différentes familles. Chez l'Homme, elles regroupent : les HP1 et les Polycomb, mais également les méthyltransférases de la famille SUV, les hélicases d'ADN de la famille CHD, et d'autres protéines comme : SMARCC1 appartenant au complexe de remodelage de la chromatine SWI-SNF, RBBP1 (retinoblastoma-binding protein 1), la sumo protéase SENP3 (sentrin specific protease 3), CDYL1 et 2 (chromodomain Y-like) et MPP8 (M-phase phosphoprotein 8) entre autres.

La substitution des chromodomains entre HP1 et Polycomb (protéines chimériques) modifie leurs profils de localisation nucléaire, suggérant que les chromodomains jouent un rôle dans la reconnaissance et/ou la fixation des cibles spécifiques de ces 2 familles de protéines (Platero JS, 1995, EMBO J). Des séquences codant pour des protéines à chromodomaine ont été retrouvées dans les génomes d'animaux et de plantes, suggérant que le chromodomaine a un rôle structural très conservé. Dans des expériences d'espèces croisées, le chromodomaine de HP1 β murin peut fonctionnellement remplacer celui du HP1 de *S pombe* (Wang G, 2000, Mol Cell Biol), et l'expression de HP1 α humain peut récupérer la létalité d'une souris possédant un mutant homozygote du gène *Su(var)2-5*, qui code pour HP1 α (Norwood LE, 2004, Gene).

Chromodomaines et extinction des gènes

La fonction d'extinction des gènes de HP1 dépend de l'interaction entre son chromodomaine et la lysine 9 di- ou tri-méthylée de l'histone H3 (Platero JS, 1995, EMBO J / Jacobs SA, 2002, Science). La poche hydrophobe du chromodomaine fournit un environnement approprié pour la reconnaissance de ce résidu modifié. La méthylation de H3K27, quant à elle, sert de site d'ancrage pour le chromodomaine de Polycomb (Cao R, 2002, Science). La discrimination entre les 2 marques répressives a été analysée. Bien que les chromodomaines soient structurellement proches, les repliements au niveau du site de reconnaissance sont distincts, permettant de discriminer entre les lysines 9 et 27 méthylées de l'histone H3. Polycomb interagit avec un nombre plus important de résidus en aval de la lysine méthylée (LATKAARK₂₇S), tandis que les HP1 reconnaissent plus volontiers les acides aminés voisins directs (QTARK₉S) (Fischle W, 2003, Genes Dev).

LES PROTEINES DE LA SOUS-FAMILLE HP1

1 Généralités

1.1 Variégation par effet de position

Les propriétés répressives ont été attribuées à l'hétérochromatine sur la base que chez la drosophile, les gènes de régions euchromatiques (actifs) qui sont transloqués à proximité de l'hétérochromatine péricentrique acquièrent une variégation de leur profil d'expression (Muller HJ, 1930, J Genet). Ce phénomène a été appelé « effet de position » et des mutations qui l'affectent positivement ou négativement ont été isolées. En particulier les mutations du gène Su(var)2-5 chez la drosophile, qui code pour HP1, sont de forts suppresseurs de la variégation. HP1 elle-même a été retrouvée enrichie au niveau des domaines hétérochromatiques (Eissenberg JC, 1990, PNAS / James TC, 1986, Mol Cell Biol).

1.2 Structure

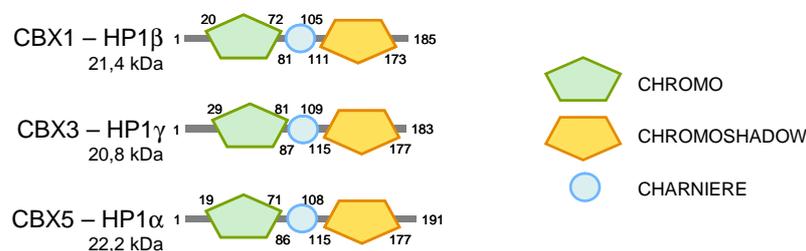


Figure 6 - Représentation schématique des protéines HP1 humaines

Les HP1 sont des protéines phylogénétiquement conservées, retrouvées dans divers organismes eucaryotes à l'exception de la levure bourgeonnante (*S. cerevisiae*), dans laquelle la variégation dépend des protéines du complexe d'extinction SIR (Silent Information Regulatory) (Moazed D, 2001, Mol Cell). HP1 possède des homologues dans divers organismes, de *S. pombe* à l'Homme (Singh PB,

1991, NAR). *Neurospora crassa* et *S. pombe* (Swi6) contiennent un homologue unique de HP1, *Dictyostelium discoideum* et *C. elegans* (HPL-1 et HPL-2) en contiennent deux. *Drosophila melanogaster* possède trois principales isoformes des protéines HP1 (HP1a, HP1b et HP1c), qui sont exprimées de manière ubiquitaire chez la mouche adulte, et 2 isoformes minoritaires qui sont tissus-spécifiques (HP1d/rhino dans les ovaires et HP1e dans les testicules). Cette classification est fondée sur la similarité de séquence en acides aminés et la structure de leurs domaines (Vermaak D, 2005, PLoS Genet). Comme chez la drosophile, les souris et les humains possèdent également 3 principales isoformes codées par des gènes distincts: HP1 α (CBX5), HP1 β (CBX1) et HP1 γ (CBX3) (Li Y, 2002, PNAS / Jones DO, 2000, Bioessays). Les protéines HP1 sont relativement petites (environ 180 acides aminés et 15 à 35kDa, Figure 5).

A défaut des protéines Polycomb, les HP1 possèdent en plus d'un chromodomaine amino-terminal, un domaine chromoshadow carboxy-terminal (Aasland R, 1995, NAR / Koonin EV, 1995, NAR). Les deux domaines sont séparés par une région charnière, moins conservée, riche en acides aminés chargés. Il est remarquable que le chromodomaine reste monomérique en solution tandis que le chromoshadow dimérise facilement dans les mêmes conditions (Brasher SV, 2000, EMBO J). La dimérisation des domaines chromoshadow entraîne la formation d'oligomères homo- ou hétéro-typiques (Ye Q, 1997, JBC / Brasher SV, 2000, EMBO J / Nielsen AL, 2001, Mol Cell / Wang G, 2000, Mol Cell Biol). L'homo- et l'hétéro-dimérisation des HP1 nécessite la présence d'un pentapeptide conservé (PxVxL) présent dans le domaine chromoshadow, qui est également trouvé dans de nombreuses protéines interagissant avec les HP1, certaines sont susceptibles de diriger les HP1 vers une région chromosomique spécifique (Cowieson NP, 2000, Curr Biol / Smothers JF, 2001, Mol Cell Biol). Des études de fonctionnalité inter-espèces dans lesquelles HP1 murin a été exprimé dans *S. pombe* ont montré que les fonctions espèce-spécifiques de HP1 sont régulées par le domaine chromoshadow et non le chromodomaine (Platero JS, 1995, EMBO J). La région charnière contient la séquence en acides aminés la plus variable entre toutes les HP1, tant au niveau d'espèces différentes que de même espèce. Cette région est très flexible et la cible de modifications post-traductionnelles, en particulier la phosphorylation (Badugu R, 2005, Chromosoma / Koike N, 2000, FEBS Lett / Zhao T, 2001, JBC).

Bien que leurs séquences soient identiques à 65%, les HP1 codent pour des protéines possédant différents profils de localisation (Vermaak D, 2005, PLoS Genet). Chez la drosophile, HP1b est localisé dans l'euchromatine et l'hétérochromatine, tandis que HP1c apparaît exclu de l'hétérochromatine centromérique. Dans des cellules mammifères, HP1 α et HP1 β sont principalement localisées au niveau de l'hétérochromatine, alors que HP1 γ est observée au niveau des eu- et hétérochromatines (Minc E, 2000, Cytogenet Cell Genet). Chaque isoforme HP1 est régulée par des modifications post-traductionnelles, comme l'acétylation, la phosphorylation, la méthylation, l'ubiquitinylation et la sumoylation, d'une manière similaire aux histones (Lomberk G, 2006, Nat Cell Biol). Ainsi, les localisations spécifiques des HP1 sont probablement contrôlées par leurs partenaires protéiques, mais également *via* leurs modifications post-traductionnelles. Enfin, les isoformes HP1 de mammifères montrent une dynamique différente au cours de la différenciation et du cycle cellulaire (Furuta K, 1997, Chromosoma / Hayakawa T, 2003, J Cell Sci).

1.3 Phosphorylations

- Chez la drosophile

La phosphorylation de HP1 est nécessaire pour son ciblage vers l'hétérochromatine, et peut-être même aussi pour la formation de cette dernière (Eissenberg JC, 1994, JBC). Trois sites de phosphorylation par la Caséine Kinase 2 ont été identifiés : S15, S199 et S202, qui sont nécessaires pour la formation de l'hétérochromatine (Zhao T, 1999, JBC).

- Chez les mammifères

HP1 α et HP1 γ existent sous différentes formes phosphorylées, devenant hyper-phosphorylées à la mitose. Au contraire, HP1 β reste sous une forme unique au cours du cycle cellulaire. HP1 α est retrouvé principalement au niveau des centromères, HP1 β est distribué largement sur tout le chromosome et HP1 γ est localisé principalement au niveau de l'euchromatine (Minc E, 1999, Chromosoma). Comme pour les drosophiles, leur localisation varie au cours du cycle cellulaire et HP1 β , qui se dissocie de l'hétérochromatine quand l'histone H3 est phosphorylée en début de mitose, se réassocie une fois que la phosphorylation est retirée en fin de mitose (Murzina N, 1999, Mol Cell). Dans HP1 β et HP1 γ , mais pas dans HP1 α , T169 et S172 se situent dans un contexte similaire aux sites de phosphorylation C-terminaux de HP1 drosophile par CK2.

D'autre part, la phosphorylation de la sérine 83 de HP1 γ définit une sous-population de cette isoforme qui est exclusivement localisée au niveau de l'euchromatine (Lomberg G, 2006, Nat Cell Biol).

Lors de dommages à l'ADN, les cassures double-brin de l'ADN mobilisent HP1 β , une phosphorylation de la thréonine 51 accompagne cette mobilisation. La thréonine 51 est localisée au niveau du chromodomaine. La déphosphorylation de ce résidu (médiée par l'inhibition de CK2) assure le relargage de HP1 β à partir des sites de mobilisation (Ayoub N, 2008, Nature). Ainsi les réparations ne seraient pas guidées exclusivement par des modifications post-traductionnelles des histones mais également par les modifications des protéines non-histones.

2 HP1 et l'hétérochromatine péricentrique

La chromatine péricentrique est majoritairement constituée de séquences d'ADN répétées qui sont adjacentes aux centromères. C'est une forme d'hétérochromatine au même titre que les régions télomériques, demeurant condensée et relativement inaccessible durant le cycle cellulaire. Dans ces régions les gènes sont majoritairement réprimés à travers la mitose et la méiose. Il est remarquable de noter que, parmi différentes espèces, les séquences nucléotidiques des centromères n'ont pas été conservées. Ainsi, la fonction des centromères n'est pas définie par la séquence en ADN mais plutôt par une organisation chromatinienne particulière. Cette dernière, une fois établie, doit être maintenue de manière stable à travers toutes les divisions. L'importance de l'organisation de la chromatine plutôt que la séquence en ADN fournit un paradigme sur un domaine chromatinien défini de manière épigénétique.

La formation d'hétérochromatine implique l'assemblage de structures spécialisées. L'extinction des gènes hétérochromatiques est supposée résulter de la compaction des nucléosomes vers une

chromatine compacte et dense qui ne permet pas l'accès à l'ADN par la machinerie de transcription (Grewal SI, 2003, Science). Cependant, cette vision de longue date a été remise en cause par des études du système d'ARN interférence (ARNi). La transcription d'ARN non codants à partir de régions répétées d'ADN hétérochromatique a été observée dans différents organismes : levure (Motamedi MR, 2004, Cell / Verdel A, 2004, Science), nématode (Grishok A, 2005, Genes Dev / Sijen T, 2003, Nature), drosophile (Pal-Bhadra M, 2004, Science), plante (Baulcombe D, 2004, Nature) et mammifère (Maison C, 2002, Nat Genet). Ces ARN semblent participer à l'assemblage de la chromatine et à l'extinction des gènes car ils sont indispensables pour la mise en place de l'hétérochromatine péricentrique. Cela suggère qu'une transcription active à partir des éléments répétés pourrait être une étape préliminaire à l'assemblage de l'hétérochromatine.

HP1 est une des protéines non-histone les plus abondantes au niveau de l'hétérochromatine péricentrique, là où elle est majoritairement localisée (Guenatri M, 2004, J Cell Biol). Cet enrichissement nécessite la présence de la marque H3K9me3 catalysée par SUV39H (Bannister AJ, 2001, Nature / Lachner M, 2001, Nature / Peters AH, 2001, Cell).

2.1 Paradigme d'assemblage/de propagation de l'hétérochromatine

La plus connue des fonctions de HP1 est donc la formation d'hétérochromatine : un modèle implique une boucle de recrutement basée sur la méthylation de H3K9. HP1 est recruté au niveau des H3K9 tri-méthylées par SUV39H1 (Lachner M, 2001, Nature / Stewart MD, 2005, Mol Cell Biol). A son tour, HP1 recrute d'autres protéines SUV39H1 qui propagent la méthylation de H3K9 aux nucléosomes adjacents, permettant le recrutement de nouvelles molécules HP1, et ainsi de suite. Ce modèle a été baptisé « hétérochromatinisation », il a été étendu à la méthylation de l'ADN car les 2 protéines : HP1 et SUV39H1 interagissent avec les méthyltransférases d'ADN (Fuks F, 2003, NAR). Il est à noter que, dans certains cas, la méthylation de H3K9 précède la méthylation de l'ADN (Tamaru H, 2001, Nature / Lehnertz B, 2003, Curr Biol), supportant l'idée que ces molécules participent à une boucle de recrutement pour l'extinction des gènes. Il faut ajouter que les méthyltransférases SUV4-20H1 et SUV4-20H2 interagissent également avec les 3 isoformes HP1 dans des cellules embryonnaires murines, elles catalysent la tri-méthylation de la lysine 20 de l'histone H4 (Schotta G, 2004, Genes Dev), une marque caractéristique de l'hétérochromatine constitutive selon le code histone.

Bien que les relations existantes entre H3K9me3, HP1 et SUV39H soient évidentes, il a été décrit que HP1 et SUV39H pouvaient fonctionner indépendamment l'une de l'autre chez la drosophile et dans des cellules humaines (James TC, 1989, Eur J Cell Biol / Greil F, 2003, Genes Dev / Fanti L, 2003, Genetica), soulignant la non universalité du modèle admis.

Le maintien des modifications de la chromatine, en particulier l'hypoacétylation et la méthylation, est nécessaire pour la fonction des centromères puisque la perte de H3K9me3 dans des souris déficientes en Suv39h ou l'hyperacétylation suite à un traitement avec un HDACi perturbe la localisation des HP1 au niveau de l'hétérochromatine péricentrique, d'où une mauvaise ségrégation des chromosomes (Maison C, 2002, Nat Genet / Taddei A, 2005, EMBO Rep)

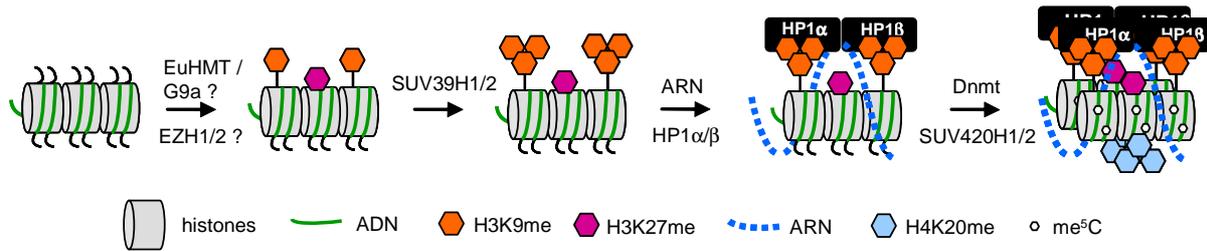


Figure 7 – Modèle actuel pour décrire l'établissement de l'hétérochromatine péricentrique

2.2 La voie ARNi impliquée dans l'hétérochromatine

- Levure

S. cerevisiae ne possède pas de HP1, qui est essentiel pour l'extinction des gènes chez les métazoaires. La levure de fission, *S. pombe*, ne possède pas d'hétérochromatine cytologiquement définie mais ses régions péricentriques, télomériques et silencieuses partagent les mêmes propriétés génétiques et biochimiques, notamment un enrichissement en H3K9me, avec l'hétérochromatine des eucaryotes. Chez *S. pombe*, une connexion entre la voie ARNi et la formation d'hétérochromatine a été bien étudiée. La délétion de Ago, Dicer1, Tas3 ou Rdp1, qui sont des composants de la voie ARNi, entraîne la perte de : la méthylation de H3K9, la fixation de Swi6 (orthologue de HP1) et l'extinction des gènes au niveau de l'hétérochromatine péricentrique (Volpe TA, 2002, Science). Les éléments répétés à l'intérieur de l'hétérochromatine centromérique sont transcrits par la Polymérase II en ARN double-brins qui sont clivés par la ribonucléase Dicer pour former des siARN (short interfering ARN) (Elgin SC, 2003, Curr Biol). Ces répétitions sont suffisantes pour générer de l'hétérochromatine au niveau de sites ectopiques, de manière ARNi-dépendante (Hall IM, 2002, Science / Partridge JF, 2002, Curr Biol). Les siARN, en association avec Ago1, forment le complexe RITS (RNA-induced initiation of transcriptional gene silencing) qui recrute d'autres protéines pour initier l'hétérochromatinisation (Cam HP, 2005, Nat Genet).

- Drosophile

De même, chez *D melanogaster*, des mutations dans les gènes qui codent pour les composants de la voie ARNi provoquent une mauvaise localisation des HP1 et la suppression de l'extinction des gènes à proximité des centromères (Pal-Bhadra M, 2004, Science). Alors qu'il est évident qu'une relation existe entre la formation d'hétérochromatine et l'extinction des gènes par les voies ARNi, un rôle direct des ARNi dans le ciblage de la formation d'hétérochromatine n'a pas encore été démontré. Il semblerait que 2 voies distinctes agissent : une première dans les lignées germinales et une autre au niveau de l'embryon précoce. La seconde nécessite la présence préliminaire d'hétérochromatine dans les noyaux des cellules somatiques pour assurer le maintien de la répression des gènes (Huisinga KL, 2008, BBA).

- Mammifères

Il est important de déterminer s'il existe un lien similaire entre le système ARNi et la formation d'hétérochromatine et la maintenance chez les eucaryotes supérieurs. Dans les cellules de

mammifères, la présence d'ARN est requise pour l'assemblage de la chromatine condensée par HP1, son domaine de liaison à l'ARN se situe au niveau de la charnière entre le chromodomaine et le chromoshadow (Maison C, 2002, Nat Genet / Muchardt C, 2002, EMBO Rep). La délétion du Dicer dans des cellules souches embryonnaires murines cause une mauvaise localisation des HP1 ainsi que l'accumulation de transcrits à partir des éléments répétés centromériques sans un effet prononcé sur les modifications d'histones ou la méthylation de l'ADN (Fukagawa T, 2004, Nat Cell Biol / Kanellopoulou C, 2005, Genes Dev / Murchison EP, 2005, PNAS). Ces données suggèrent que Dicer participe au maintien de la structure de l'hétérochromatine et à l'extinction des gènes au niveau des centromères. Cependant, il n'a pas été déterminé si un processus comparable de formation d'hétérochromatine ARNi-dépendante existait chez les mammifères. Il a été proposé que des structures dédiées assurent la réplication de ces domaines, appelées pHDB : pericentric Heterochromatin Duplication Bodies (Quivy JP, 2004, EMBO J).

3 Partenaires protéiques et fonctions biologiques

Bien que les protéines HP1 soient localisées au niveau de la chromatine, elles ne se fixent pas directement sur l'ADN mais au niveau des histones. Elles sont également capables d'interagir avec de nombreux autres partenaires protéiques.

3.1 Réplication et/ou la réparation de l'ADN

Chez la drosophile, le chromodomaine de HP1 interagit directement avec le complexe ORC (Origin Recognition Complex) indispensable à l'initiation de la réplication chez les eucaryotes (Pak DT, 1997, Cell / Shareef MM, 2001, Mol Biol Cell). Une mutation qui mimerait un état phosphorylé de la région charnière entraîne la perte d'interaction avec ORC1 et HOAP (Origin Recognition complex-Associated Protein), alors qu'elle augmente la localisation au niveau de l'hétérochromatine (Badugu R, 2005, Chromosoma). Les 2 facteurs sont capables de se fixer directement sur l'ADN, pour le moment leur fonction n'est pas clairement définie, à savoir si elles recrutent HP1 ou si elles interagissent avec un HP1 déjà fixé sur la chromatine. Aussi, l'interaction entre HP1 et SUUR (Suppressor of Under-Replication) est cruciale pour le recrutement de cette dernière, qui est responsable de la réplication tardive de l'hétérochromatine péricentrique (Pindyurin AV, 2008, J Cell Sci).

HP1 α murin et le complexe PreRC (ORC1, ORC2, CDC6) co-localisent au niveau de l'hétérochromatine et interagissent en double-hybride (Auth T, 2006, Exp Cell Res).

Enfin, HP1 s'associe également avec des protéines impliquées dans la réplication et la réparation de l'ADN, comme CAF1 (Chromatin Assembly Factor 1). Ce dernier dépose les histones H3 et H4 au niveau des brins d'ADN néo-synthétisés des régions eu- et hétéro-chromatiques pendant la réplication et la réparation de l'ADN (Gaillard PH, 1996, Cell / Kaufman PD, 1995, Cell). HP1 est indispensable pour l'association de CAF1 avec la chromatine (Murzina N, 1999, Mol Cell).

3.2 Répresseurs transcriptionnels

Les HP1 mammifères interagissent avec TIF1 α et TIF1 β (transcriptional intermediary factors) (Le Douarin B, 1996, EMBO J). L'interaction de TIF1 β avec des protéines à domaine KRAB (Krüppel-Associated Box), un domaine particulier commun à de nombreux répresseurs transcriptionnels (Friedman JR, 1996, Genes Dev / Moosmann P, 1996, NAR), suggère qu'un complexe constitué de HP1-TIF1-KRAB pourrait-être recruté au niveau des loci définis par les domaines de liaison à l'ADN des facteurs de transcription de la famille Krüppel (Ryan RF, 1999, Mol Cell Biol). TIF1 α et TIF1 β répriment la transcription à travers une activité histone désacétylase (Nielsen AL, 1999, EMBO J). TIF1 β interagit avec des chromodomaines dimérisés d'HP1 β *via* son motif PxVxL. L'activité de répresseur transcriptionnel de TIF1 β dépend de cette interaction (Nielsen AL, 1999, EMBO J), qui est essentielle pour la délocalisation de TIF1 β à partir de l'euchromatine vers l'hétérochromatine qui accompagne la différenciation de cellules endodermiques (Cammass F, 2002, J Cell Sci). Le recrutement des protéines HP1 par le complexe KRAB-TIF1 β à des loci spécifiques à l'intérieur du génome, à travers la formation de structures proches de l'hétérochromatine, pourrait éteindre la transcription des gènes ciblés. Enfin, le co-répresseur transcriptionnel KAP1 (KRAB Associated Protein 1) est associé directement aux HP1 *via* son domaine chromoshadow (Lechner MS, 2000, Mol Cell Biol). KAP1 coordonne la désacétylation des histones *via* le recrutement du complexe NuRD-HDAC (Schultz DC, 2001, Genes Dev) et la méthylation de H3K9 *via* la méthyl-transférase SETDB1 (Schultz DC, 2002, Genes Dev), permettant le recrutement d'autres HP1. L'interaction de HP1 avec l'isoforme TIF1 δ a été rapportée dans un contexte biologique particulier. En effet, ces deux protéines sont retrouvées au niveau de loci dispersés à l'intérieur du chromocentre centromérique et au sein du nucléoplasme environnant dans les cellules testiculaires qui ont complété la méiose. Cette interaction (*via* le domaine PcVxL de HP1) et sa localisation suggèrent que ces protéines joueraient un rôle pendant les phases post-méiotiques de la spermatogenèse (Khetchoumian K, 2004, J Biol Chem).

3.3 Enveloppe nucléaire

Les protéines HP1 ont été localisées au niveau de la périphérie nucléaire, cela pourrait être expliqué par leur interaction avec le récepteur nucléaire de la lamine B et/ou avec la localisation de l'hétérochromatine centromérique (Ye Q, 1996, JBC / Ye Q, 1997, JBC / Kourmouli N, 2000, EMBO J / Polioudaki H, 2001, EMBO Rep). La méthylation de H3K9 est détectée à la périphérie nucléaire des cellules murines et humaines démontrant que cette modification est enrichie au niveau de l'hétérochromatine périphérique (Cowell IG, 2002, Chromosoma).

3.4 Contrôle du cycle cellulaire

Dans des cellules mammifères, Rb (Retinoblastoma) est responsable de la répression des gènes nécessaires pour le passage G1-S au cours du cycle cellulaire, un événement qui a lieu suite à la fixation et la séquestration du facteur de transcription E2F. Rb interagit avec SUV39H et HP1, permettant leur recrutement au niveau des promoteurs ciblés par E2F pour établir la répression (Vandel L, 2001, Mol Cell Biol / Young AP, 2004, Oncogene). En effet, HP1 et la méthylation de H3K9

ont été localisés au niveau du site d'initiation de la transcription du gène de la *cycline E* (Nielsen SJ, 2001, Nature), et la phosphorylation de Rb, dépendante du cycle cellulaire, abolit son interaction avec Suv39H1 (Vandel L, 2001, Mol Cell Biol).

La liste des partenaires protéiques des HP1 grandit chaque jour et un grand nombre d'entre eux sont associés à la chromatine : enzymes modifiant les histones (SUV39H, SUV420H), régulateurs transcriptionnels (TIF1), facteurs de réplication (ORC, CAF1) ou impliqués dans le cycle cellulaire (Rb), protéines nucléaires structurales (histones, INCENP) (Hiragami K, 2005, Cell Mol Life Sci / Li Y, 2002, PNAS).

4 Répression des gènes euchromatiques

Bien que les HP1 soient concentrées dans l'hétérochromatine péricentrique dans la plupart des organismes, elles sont également retrouvées dans l'euchromatine. En effet chez la drosophile, la distribution des HP1 sur les chromosomes polytènes n'est pas limitée aux régions d'hétérochromatine péricentrique, elles sont également présentes sur environ 200 sites à l'intérieur des bras euchromatiques et au niveau des télomères (Fanti L, 2003, Genetica). Leur fixation peut être corrélée aussi bien à une répression (Grewal SI, 2003, Science / James TC, 1989, Eur J Cell Biol) qu'à une activation (Piacentini L, 2003, J Cell Biol) des gènes.

Au sein d'une souris transgénique dans laquelle un transgène rapporteur a été intégré à l'euchromatine, les protéines HP1 α et HP1 γ sont recrutées au niveau du promoteur viral fusionné au transgène lors de sa répression induite par une hormone (Ayyanathan K, 2003, Genes Dev). La répression corrèle avec la relocalisation du locus contenant le transgène au niveau de régions où la chromatine est condensée. De plus, la répression est maintenue pendant plusieurs générations.

Dans des cellules quiescentes (G_0), HP1 γ et plusieurs histones méthyltransférases co-précipitent avec le complexe répresseur E2F6, soulignant le rôle joué par HP1 dans la répression des gènes ciblés par E2F et Myc (Ogawa H, 2002, Science). De plus, SUV39H humain a également été décrit comme étant nécessaire à la répression des gènes ciblés par E2F dans des myoblastes différenciés (Ait-si-ali S, 2004, EMBO J).

HP1 γ n'est pas significativement corrélé avec la méthylation de H3K9me dans les cellules de drosophile et humaines (Greil F, 2003, Genes Dev / Minc E, 1999, Chromosoma). D'une manière similaire les deux isoformes de HP1 chez le ver *Caenorhabditis elegans*, Hpl-1 et Hpl-2, ont des phénotypes distincts lorsqu'ils sont mutés et possèdent des profils de localisation qui sont partiellement spécifiques, avec une grande dispersion au niveau des sites de fixation sur l'euchromatine (Schott S, 2006, Dev Biol). Ces résultats suggèrent que des mécanismes d'extinction des gènes euchromatiques peuvent être indépendants de la méthylation de H3K9, et spécifiques d'une isoforme particulière.

Pour résumer, HP1 peut initier la formation d'une structure chromatinienne répressive une fois qu'il est recruté au niveau de promoteurs spécifiques à l'intérieur de l'euchromatine. La structure

chromatinienne résultante ressemble à l'hétérochromatine, au moins en termes de transcription, de localisation nucléaire et de transmission aux générations suivantes.

5 Activation transcriptionnelle

En plus de la répression des gènes, HP1 est impliqué dans la régulation positive de l'expression des gènes dans des zones aussi bien euchromatiques qu'hétérochromatiques.

5.1 HP1 régule l'expression des gènes euchromatiques

Le nombre grandissant de publications établissant un lien entre HP1 et l'activation des gènes au niveau de l'euchromatine démontre le potentiel paradoxal de ces protéines. Les chromosomes polytènes de *Drosophile* sont enrichis en HP1 au niveau de certains gènes différentiellement régulés au cours du développement et des « puffs » induits par choc thermique (Piacentini L, 2003, *J Cell Biol*). A l'aide d'immuno-précipitations de la chromatine, il a été démontré que HP1 se fixe sur les puffs seulement après l'induction, et qu'elles se fixent spécifiquement sur la région codante et non au niveau des promoteurs. Les localisations de HP1 au niveau de l'euchromatine sont ARN-dépendantes car une perte de la fixation des HP1 est observée après un traitement à la ribonucléase. Le rôle de ces ARN n'a pas été clairement identifié, à savoir s'ils interagissaient directement avec HP1 ou s'ils favorisaient la fixation de HP1 sur les queues d'histones ou d'autres protéines.

Dans des cellules érythroïdes murines, il a été rapporté la présence de HP1 γ et H3K9me3 au niveau de l'euchromatine (Vakoc CR, 2005, *Mol Cell*). La méthylation de H3K9 est proportionnelle à l'activité transcriptionnelle et sa présence est dépendante de l'élongation par l'ARN polymérase. Dans des cellules drosophiles et humaines, il a été rapporté que HP1 s'associait avec l'hétérochromatine d'une manière ARN-dépendante (Maison C, 2002, *Nat Genet* / Muchardt C, 2002, *EMBO Rep*). L'association avec un ARN semble jouer un rôle important dans la multi-fonctionnalité des HP1 dans différents organismes.

Chez la drosophile, la mutation de Su(var)2-5 entraîne une diminution de la transcription de plusieurs gènes qui sont habituellement ciblés par HP1, ce qui suggère que ces protéines jouent un rôle dans l'activation transcriptionnelle. Il est à souligner que cette activation est indépendante de l'activité de la méthyltransférase SUV39H puisque son invalidation ne modifie pas l'expression de ces mêmes gènes (Cryderman DE, 2005, *Dev Dyn*). Cependant d'autres enzymes responsables de la méthylation de H3K9 ont été identifiées comme pouvant agir au niveau de l'euchromatine : les homologues de G9a chez la drosophile (Stabell M, 2006, *NAR*) et de SETDB1 chez le nématode (Andersen EC, 2007, *Development*) et chez la drosophile (Seum C, 2007, *PLoS Genet*), bien qu'elles aient principalement été associées à une répression transcriptionnelle.

HP1c de drosophile interagit avec les protéines à doigt de zinc WOC (without children) et ROW (Relative-of-WOC) qui sont des facteurs de transcription putatifs qui co-localisent avec la forme active de la polymérase II et H3K4me3. Ils régulent positivement un ensemble de gènes localisés dans des cellules du système nerveux en particulier (Font-Burgada J, 2008, *Genes Dev*).

Ces résultats ont démontré que HP1 est impliquée dans la régulation positive de certains gènes euchromatiques au niveau d'un certain nombre de loci. Ce type de résultats a été répété dans des cellules en culture par la déplétion de HP1 (De Lucia F, 2005, NAR). Plus récemment, des expériences de localisation à haute résolution ont confirmé que HP1 est associée avec de la chromatine transcriptionnellement active chez la drosophile (de Wit E, 2007, PLoS Genet / Johansson AM, 2007, PLoS Genet).

5.2 HP1 est essentiel pour l'expression des gènes hétérochromatiques

Bien qu'il soit accepté que les gènes transloqués à proximité de l'hétérochromatine sont réprimés par les diffusions de HP1 et de la méthylation des histones, chez la drosophile il existe également des gènes intégrés dans les séquences répétées des régions centromériques qui sont toujours actifs. Des analyses à l'échelle du génome ont confirmé que certains gènes présents dans la chromatine péricentrique associée à HP1 et Su(var)3-9 sont transcriptionnellement actifs (Eissenberg JC, 1990, PNAS). Il a été proposé que la structure chromatinienne compacte de l'hétérochromatine était nécessaire pour le maintien des niveaux d'expression de ces gènes (Eissenberg JC, 1990, PNAS / Clegg NJ, 1998, Genome / Hearn MG, 1991, Genetics / Lu BY, 2000, Genetics). En effet, lorsque ces gènes sont transloqués dans des régions euchromatiques, ils acquièrent un phénotype d'effet de position en devenant réprimés (Wakimoto BT, 1990, Genetics).

6 Fonctions télomériques des HP1

Une analyse cytogénétique a démontré que HP1 est un des composants de tous les télomères chez la drosophile et prévient leur fusion, en encapsulant les extrémités chromosomiques. La perte de HP1 dans les mutants Su(var)2-5 entraîne de multiples fusions entre télomères, provoquant une configuration anormale des chromosomes des neuroblastes, dans les disques imaginaux, et les cellules méiotiques mâles (Fanti L, 1998, Mol Cell). Aussi, ces fusions de télomères médiées par l'attachement des extrémités des ADN créent des ponts entre les chromosomes pendant l'anaphase, engendrant des larges cassures chromosomiques. Ces observations impliquent que HP1 est une protéine de coiffage des télomères qui facilite leur stabilité (Perrini B, 2004, Mol Cell). Le chromodomaine de HP1 n'est pas nécessaire à la localisation de la protéine et la stabilité des télomères, suggérant que la localisation télomérique ne repose pas sur la méthylation de H3K9. C'est la région charnière qui semblerait interagir directement avec l'ADN télomérique.

Aussi, HP1 est aussi impliquée dans le contrôle de la transcription des transposons télomériques et de l'élongation des télomères. Dans les mutants hétérozygotes HP1, les télomères montrent une remarquable élongation au cours du temps, et la transcription de TART et HeT-A, qui sont des éléments assimilés à des rétrotransposons non-LTR, est augmentée (Savitsky M, 2002, Mol Cell Biol). Dans les larves qui n'ont plus de HP1 ou possédant un chromodomaine muté, la transcription des transposons télomériques est augmentée (Perrini B, 2004, Mol Cell). D'une manière intéressante, un chromodomaine intact est indispensable pour la méthylation de H3K9 au niveau des télomères. Ainsi, la présence du chromodomaine de HP1 et de la méthylation de H3K9, alors qu'ils sont dispensables pour la stabilité des télomères, sont requis pour la suppression de la transcription des transposons

télomériques et l'élongation des télomères. Le rôle de HP1 au sein des télomères est assuré par 2 types de fixation : le coiffage télomérique repose sur l'interaction directe avec l'ADN tandis que le contrôle de la transcription dépend de l'interaction entre le chromodomaine et H3K9me.

Il semblerait que les HP1 soient impliquées dans le métabolisme des télomères chez les mammifères. Différents homologues HP1 sont localisés au niveau des télomères des chromosomes métaphasiques de mammifères (Aagaard L, 2000, J Cell Sci / Koering CE, 2002, EMBO Rep / Minc E, 1999, Chromosoma). Par exemple, HP1 α s'associe avec Ku70, qui régule les télomères de mammifères (Song K, 2001, JBC). De plus, la sur-expression de HP1 α et HP1 β perturbe l'association de hTERT (l'unité catalytique de la télomérase) sur les télomères, induisant la fusion des télomères dans des cellules humaines. Chez la souris, la réduction des 3 protéines HP1 cause une élongation anormale des télomères (Garcia-Cao M, 2004, Nat Genet) comme chez la drosophile avec HP1a.

7 Dynamique des HP1 – Réversibilité

Initialement, l'hétérochromatine « a été considérée initialement » comme une structure relativement compacte, rigide et inerte. Cependant, plusieurs protéines comme les HP1 démontrent une dynamique dans leur fixation sur la chromatine (Cheutin T, 2003, Science / Festenstein R, 2003, Science / Dialynas GK, 2007, J Cell Sci), suggérant que l'hétérochromatine pourrait être une structure dynamique et malléable, dans laquelle l'accès à l'ADN ne serait pas forcément aboli.

La réversibilité de la fixation de HP1 au niveau de H3K9me3 a été mise en évidence par la phosphorylation de H3S10 par Aurora B lors de l'entrée en mitose (Fischle W, 2005, Nature / Hirota T, 2005, Nature / Mateescu B, 2004, EMBO Rep) qui libère HP1 de la chromatine. Une autre étude a montré qu'une diminution de l'expression de la kinase JIL1, également responsable de la phosphorylation de H3S10, entraîne une propagation de la méthylation de H3K9 et de HP1 au niveau de sites ectopiques (Zhang W, 2006, Development).

Enfin, les déméthylases des familles KDM3 (Yamane K, 2006, Cell) et KDM4 (Klose RJ, 2006, Nature), responsables de la suppression de la marque H3K9me3 sont des enzymes jouant un rôle dans la régulation de la fixation des HP1 sur la chromatine.

8 Conclusions

Les protéines HP1 sont probablement les protéines non histones les plus étudiées depuis leur découverte en 1986 (James TC, 1986, Mol Cell Biol). Aujourd'hui, une quantité significative d'informations est disponible à propos de leur structure ou de leurs propriétés biochimiques. Les HP1 sont préférentiellement associées à l'hétérochromatine péricentrique, mais sont également retrouvées au niveau des télomères et dans de nombreux sites euchromatiques. En fonction du contexte chromosomique (ADN répétés \neq euchromatine, promoteur \neq région codante) et de leurs partenaires, les HP1 peuvent initier : une répression, associée ou non à H3K9me3, ou une activation transcriptionnelle. Les interactions avec un grand nombre de partenaires différents et de l'ARN ainsi que leurs modifications post-traductionnelles contribuent à la diversité de leurs fonctions.

Cependant plusieurs questions restent sans réponse, en particulier concernant les complexes recrutés par les HP1 et leur rôle dans la cellule. La diversité des partenaires protéiques et la spécificité des

isoformes des protéines HP1 suggère que ces protéines sont impliquées dans de nombreux processus cellulaires. Avec le haut degré de similarité entre les 3 isoformes, les facteurs qui influencent ces différences restent pour le moment méconnus. Il est devenu essentiel d'identifier les protéines qui interagissent avec ces différentes isoformes afin de mieux comprendre comment une régulation positive ou négative de la transcription peut avoir lieu au niveau des régions dans lesquelles les HP1 sont recrutées.

LES PROTEINES CBX POLYCOMB

1 Introduction

1.1 *Découverte des Polycomb, une histoire de Nobelisés*

Lorsque William Bateson s'intéresse de près à des animaux et des plantes « atypiques », il constate que certaines parties sont capables de prendre la place d'autres. Il définira en 1894 ce phénomène par le terme « *homeosis* ». Quelques années plus tard, le généticien Thomas H. Morgan adopte la drosophile comme modèle pour ses études phénotypiques. Il reçoit le prix Nobel en 1933 pour avoir démontré que les chromosomes étaient le support physique de l'information génétique. La drosophile devient alors une espèce modèle et à partir de 1910, il sélectionne et analyse un grand nombre de mutants homéotiques. En 1927, Calvin B. Bridges isole une mutation « bithoraxoïde » (btx) responsable de la transformation du premier segment en troisième segment, permettant aux drosophiles d'avoir 3 paires de pattes. Il utilise pour la première fois l'expression « mutation homéotique ». De nombreuses autres mutations de ce genre ont été découvertes par la suite, dont la célèbre « antennapedia » pour laquelle les drosophiles ont des pattes en lieu et place des antennes. Même si un grand nombre de ces phénotypes résultent de la mutation des gènes homéotiques (Hox) qui codent pour des facteurs de transcription responsables du bon développement des métazoaires, un certain nombre sont le résultat de la mutation de gènes spécifiques regroupés sous le terme « Polycomb » (PcG). Il est proposé que ces gènes interagissent de manière synergique pendant le développement afin d'assurer la fonction spécifique des protéines Polycomb. Des mutations dans les gènes PcG entraînent l'expression ectopique des gènes homéotiques, provoquant des transformations des segments et des structures du corps. Edward B. Lewis, un des élèves de Thomas H. Morgan, propose le modèle suivant : chaque segment du corps de la drosophile est spécifié par une combinaison unique de gènes homéotiques qui sont exprimés spécifiquement dans ce segment particulier. Aussi, il soumet l'hypothèse que les gènes du groupe PcG participent à un système répresseur de ces gènes (Lewis EB, 1978, Nature). Il recevra le prix Nobel en 1995 pour ses travaux sur la génétique du développement et la description des mécanismes impliqués dans le contrôle du développement des animaux.

1.2 *Mémoire de l'identité cellulaire*

Le maintien de la répression est un mécanisme clé pour conserver les profils d'expression des gènes, établis au cours du développement par des facteurs spécifiques, qui sont responsables de l'identité de chacune des cellules d'un organisme pluricellulaire. Une fois que la cellule a atteint son état différencié, elle doit maintenir les profils d'expression pendant toute sa vie, même après une division cellulaire. Une sorte de « mémoire » est requise pour permettre aux cellules filles de conserver les caractéristiques de la cellule mère, représentée par l'ensemble des gènes activés et réprimés qui va être spécifique à chaque type cellulaire. Ce phénomène ne peut pas être guidé essentiellement par l'ADN, il est principalement épigénétique et établi par des structures qui peuvent être transmises à travers la division cellulaire pour permettre le ré-établissement des profils d'expression. La répression

par les PcG est un de ces phénomènes épigénétiques, car dans différentes lignées cellulaires, possédant le même génome, les mêmes complexes ciblent des gènes distincts.

Des analyses génétiques ont mis en évidence la présence d'un système antagoniste qui implique les protéines du groupe Trithorax (trxG) qui marquent les gènes ciblés par les PcG pour leur activation transcriptionnelle (Poux S, 2002, Development). En l'absence de trxG, un gène homéotique peut devenir réprimé par l'action des PcG, même dans des cellules dans lesquelles il a été activé précédemment. C'est pour cela que les protéines PcG ont été associées à l'hétérochromatine facultative. Le rôle principal des complexes trxG réside dans l'inhibition de la répression par les PcG, plutôt que dans l'initiation ou l'activation de la transcription (Klymenko T, 2004, EMBO Rep). Des orthologues des PcG et des trxG ont été identifiés chez le nématode *C. elegans* (Ross JM, 2003, Dev Cell), chez les plantes et chez les vertébrés (Alvarez-Venegas R, 2003, Curr Biol / Schumacher A, 1997, Trends Genet). Ainsi le système de mémoire de l'identité cellulaire semble avoir été conservé au cours de l'évolution.

Les modifications post-traductionnelles des histones, la méthylation de l'ADN et la présence d'ARN non codants jouent un rôle non négligeable dans la distinction des gènes à réprimer (par les PcG) ou à activer (par les trxG). Ces marques épigénétiques influencent la structure de la chromatine et sont nécessaires pour le maintien de la répression ou de l'expression des gènes.

2 Les PcG s'organisent en complexes pour réprimer les gènes

Les protéines PcG sont structurellement et fonctionnellement diverses et forment de larges complexes multimériques. Chez *D. melanogaster*, l'extinction par les PcG implique 3 types de complexes qui semblent travailler de concert. Ils ont été baptisés PRC1, PRC2 et PhoRC. Ces complexes possèdent des activités catalytiques pour modifier post-traductionnellement les queues des histones et sont impliqués dans la répression transcriptionnelle de gènes cibles en altérant la structure chromatiniennne locale. Des orthologues de toutes les protéines PcG sont présentes chez les mammifères.

2.1 Le complexe PRC1

2.1.1 Structure

- *Drosophile*

La purification biochimique du PRC1 à partir de la drosophile a permis d'identifier ses 4 principaux composants (Shao Z, 1999, Cell) : Polycomb (Pc), Posterior sex combs (Psc), polyhomeotic (Ph) et Ring, mais également une quantité plus faible de Sex comb on midleg (Scm). Des protéines additionnelles ont également été co-purifiées : Zeste, des facteurs associés à TBP (TATA-box binding protein) comme TAFII250, TAFII110, TAFII85 et TAFII62, ainsi que des éléments d'autres complexes protéiques, tels que MI2, SIN3A et SMRTER (Saurin AJ, 2001, Nature). L'association avec les protéines sus-citées pourrait-être artéfactuelle, mais la présence de facteurs du promoteur TAF est intrigante, supposant une interaction directe entre les complexes PcG et des complexes associés à des promoteurs de gènes. Aucune expérience n'a permis de valider cette interaction *in vivo*.

Certains gènes PcG ont des homologues qui pourraient fonctionner en tant qu'alternative dans différents tissus, différents stades développementaux ou même au niveau de différents gènes cibles au sein de la même cellule. Par exemple, il existe 2 gènes *ph*, qui codent pour les protéines Polyhomeotic proximal (Php) et Polyhomeotic distal (Phd) dont les fonctions précises n'ont pas été définies (Dura J, 1987, Cell).

- Mammifère

Le complexe PRC1 de mammifère a été isolé à partir de cellules HeLa en purifiant une protéine de fusion exprimée de manière exogène (Levine SS, 2002, Mol Cell Biol). Les 4 principaux composants décrits pour la drosophile ont été retrouvés, en revanche aucune protéine TAF n'a pu être détectée. Chez les souris et les humains, toutes les protéines qui composent le PRC1 possèdent plusieurs homologues, qui pourraient fonctionner comme alternative à l'isoforme « principale » à différents stades du développement ou dans différents tissus. Les complexes purifiés contiennent les CBX2, 4, 6, 7 et 8, PHC1, PHC2 et PHC3, RING1 et RNF2, BMI1 et son homologue Mel-18. Ils sont respectivement les homologues de Pc, Ph, Ring et Psc de la drosophile. La présence de plusieurs homologues et la dépendance de la stœchiométrie par rapport à la protéine cible ou l'étiquette utilisée pour la purification renforce l'idée que les PRC1 purifiés sont composés de plusieurs complexes légèrement différents.

2.1.2 Fonction

D'un point de vue fonctionnel, le chromodomaine de Pc et ses homologues CBX se fixe spécifiquement sur la lysine 27 tri-méthylée de l'histone H3. Des études *in vitro* prouvent que cette interaction est 5 fois plus forte avec un peptide tri-méthylé par rapport à des peptides mono- ou di-méthylés (Fischle W, 2003, Genes Dev). Ring chez la drosophile et RING1 et RNF2 chez les mammifères possèdent un domaine RING, qui agit comme E3 ubiquitine-ligase, capable de mono-ubiquitiner la lysine 119 de l'histone H2A (Wang H, 2004, Nature / de Napoles M, 2004, Dev Cell / Cao R, 2005, Mol Cell). La présence de BMI1, qui contient également un domaine RING, favorise l'activité catalytique de RING1 (Buchwald G, 2006, EMBO J / Li Z, 2006, JBC). PRC1 catalyse H2AK119ub1, qui réprime l'élongation transcriptionnelle en inhibant l'ARN polymérase II (Zhou W, 2008, Mol Cell / Stock JK, 2007, Nat Cell Biol). *In vitro*, le PRC1 prévient l'activité de remodelage du complexe SWI/SNF (Shao Z, 1999, Cell) et est capable de condenser la chromatine en l'absence des queues d'histones (Francis NJ, 2004, Science). La fonction précise de Ph au sein du complexe PRC1 n'a pas encore été caractérisée, mais il a été proposé que Ph pourrait promouvoir la propagation des complexes PcG (Kim CA, 2002, Nat Struct Biol).

2.2 Le complexe PRC2

2.2.1 Structure

- Drosophile

Des purifications biochimiques du PRC2 à partir de drosophiles ont révélé la présence de E(z) (Enhancer of zeste), Su(z)12 (Suppressor of zeste 12), Esc (Extra sex comb) ou son homologue Escl

et P55 (RBAP46/RBAP48) (Czermin B, 2002, Cell / Muller J, 2002, Cell / Cao R, 2002, Science / Kuzmichev A, 2002, Genes Dev / Ng J, 2000, Mol Cell Biol). Cette dernière est une protéine de fixation aux histones qui est également associée avec CAF1 et d'autres complexes de remodelage de la chromatine (Polo SE, 2006, Curr Opin Genet Dev / Taylor-Harding B, 2004, Mol Cell Biol). Plusieurs espèces moléculaires de complexes contenant E(z) ont été détectées. L'espèce la plus majoritaire est un complexe d'environ 600kDa, universellement accepté comme le complexe PRC2. Un second complexe d'environ 1MDa contient la protéine Pcl (Polycomb-like) (Tie F, 2003, Mol Cell Biol). D'autres complexes plus grands et moins caractérisés sont présents à des stades tardifs du développement, contenant la désacétylase SIR2 (Furuyama T, 2004, Curr Biol).

- Mammifères

Chez les mammifères, la situation est plus compliquée. Les complexes PRC2, PRC3 et PRC4 ont été caractérisés biochimiquement, ils diffèrent par la présence de différentes isoformes de EED, l'homologue de Esc (Kuzmichev A, 2004, Mol Cell / Kuzmichev A, 2005, PNAS). PRC4 est accumulé lorsqu'EZH2, l'homologue de E(z), est sur-exprimée dans des cellules en culture. Il contient une isoforme EED qui est exprimée uniquement dans des cellules souches non différenciées et SIRT1, l'homologue de mammifère de SIR2. Le knock-out de SUZ12 réduit le niveau de EZH2, indiquant que la formation du complexe stabilise la protéine EZH2 (Pasini D, 2004, EMBO J). Le PRC2 peut contenir la protéine PHF1, l'homologue de Pcl (Sarma K, 2008, Mol Cell Biol / Cao R, 2008, Mol Cell Biol), deux autres homologues existent cependant chez les mammifères : PHF19 et MTF1. Enfin, un complexe homologue au PRC2 précédemment décrit contenant EED, SUZ12 et EZH1, le paralogue de EZH2 a été purifié (Margueron R, 2008, Mol Cell / Shen X, 2008, Mol Cell).

2.2.2 Fonction

Chacun de ses composants contribue à la capacité du PRC2 à se fixer sur les nucléosomes et à méthyler la lysine 27 de l'histone H3 *via* le domaine SET de E(z) (Ketel CS, 2005, Mol Cell Biol), qui ne possède pas d'activité méthyl-transférase lorsqu'elle est isolée (Czermin B, 2002, Cell / Muller J, 2002, Cell). E(z) est catalytiquement inactive *in vitro* si elle n'est pas associée aux autres membres du complexes PRC2, qui sont responsables de la fixation aux histones/nucléosomes ou de l'augmentation de l'activité enzymatique (Cao R, 2004, Curr Opin Genet Dev / Nekrasov M, 2005, EMBO Rep / Tie F, 2007, Mol Cell Biol). *In vivo*, la tri-méthylation de H3K27 est caractéristique des gènes ciblés par les PcG (Schwartz YB, 2006, Nat Genet), mais E(z) est également responsable des mono- et di-méthylations de plus de 50% des H3K27 dans le génome de la drosophile (Ebert A, 2004, Genes Dev). EED est nécessaire pour tous les états de méthylation de H3K27, suggérant que plusieurs complexes contenant EED mono-méthylent H3K27 (Montgomery ND, 2005, Curr Biol).

Chez la drosophile, Pcl est nécessaire pour l'extinction des gènes homéotiques, et est retrouvé au niveau des sites Polycomb sur les chromosomes polytènes (Papp B, 2006, Genes Dev / Tie F, 2003, Mol Cell Biol). L'homologue de mammifère, PHF1, favorise l'activité de EZH2 (Sarma K, 2008, Mol Cell Biol). Cela indique que le « gros » complexe est présent au niveau des gènes ciblés par les PcG et qu'il est responsable de la tri-méthylation de H3K27. EZH1 possède également une activité H3K27 méthyl-transférase et réprime la transcription *in vivo* (Margueron R, 2008, Mol Cell / Shen X, 2008,

Mol Cell). Les complexes PRC2-EZH2 et PRC2-EZH1 auraient des fonctions complémentaires et non-redondantes, le premier agirait dans des cellules en prolifération tandis que le second serait préférentiellement présent dans des cellules différenciées quiescentes.

En présence de l'histone H1, PRC2, mais pas PRC3, méthyle H1K26, qui possède un contexte en acides aminés proche de celui de H3K27 (Kuzmichev A, 2004, Mol Cell).

Enfin, les souris déficientes en *Ezh2* ne sont pas viables (O'Carroll D, 2001, Mol Cell Biol) et les cellules privées de EED ont tendance à se différencier spontanément (Boyer LA, 2006, Nature) soulignant l'importance de ces protéines dans le développement.

2.3 Le complexe PhoRC

Seul ce troisième complexe contient une protéine capable de se lier directement à l'ADN, et qui semble essentielle pour la répression par les PcG : Pho (pleiohomeotic), elle possède un paralogue Phol (Brown JL, 1998, Mol Cell / Fritsch C, 1999, Development / Brown JL, 2003, Development). Ces 2 protéines sont les homologues du facteur de transcription mammifère Yin-Yang 1 (YY1), ainsi dénommé car il possède à la fois des fonctions d'activateur et de répresseur transcriptionnel. Bien qu'il ait été rapporté qu'il interagit avec les complexes PRC1 (Levine SS, 2002, Mol Cell Biol) et PRC2 (Satijn DP, 2001, Mol Cell Biol / Poux S, 2001, Genes Dev / Wang L, 2004 Mol Cell / Liu H, 2007, Genes Dev) chez les drosophiles et chez les mammifères, Pho recrute le PRC2 au niveau de certains gènes (Caretti G, 2004, Genes Dev). Pho a été identifié dans 2 autres types de complexes.

Dans le premier il est associé au complexe INO80 impliqué dans le remodelage de la chromatine, qui est connu pour être impliqué dans les mécanismes PcG (Klymenko T, 2006, Genes Dev). Dans le second complexe, PhoRC, jouant un rôle dans la répression des gènes homéotiques, et contenant une protéine à domaine MBT, Sfmbt. Le domaine MBT est retrouvé dans de nombreux homologues de mammifères. Sfmbt de drosophile est considérée comme une véritable protéine PcG, nécessaire pour l'extinction par les PcG. Son domaine MBT se fixe spécifiquement sur les lysines mono- et di-méthylées de H3K9 et H4K20 (Klymenko T, 2006, Genes Dev).

Au moins au niveau des gènes HOX de drosophile, les 3 complexes se fixent simultanément au niveau d'une région possédant une séquence spécifique en ADN appelée PRE (Polycomb Response Element) (Papp P, 2006, Genes Dev). Le ciblage de PRC1 et PRC2 au niveau des PRE des gènes HOX est dépendant de Pho, comme cela a été démontré dans des animaux mutants en Pho (Wang L, 2004, Mol Cell) ou lors de la perturbation du site de fixation à l'ADN (Klymenko T, 2006, Genes Dev). PhoRC cible un grand nombre de sites génomiques (196 PRE), où il co-localise avec PRC1 et PRC2, établissant son implication dans la répression par les PcG au cours du développement de la drosophile. Le site de fixation de Pho sur l'ADN au niveau des PRE a été défini ainsi : $G^C/A^C/GCCAT^T/C TT$ (Oktaba K, 2008, Dev Cell). Bien que les PRE soient nécessaires et suffisants pour le recrutement des complexes PRC chez la drosophile, de tels éléments n'ont pas encore été identifiés chez les vertébrés ni chez les plantes.

2.4 Modèle de recrutement des complexes PcG

2.4.1 Chez la drosophile

Dans le modèle « simple », le complexe PhoRC se fixe au niveau des PRE, il recrute le PRC2 qui catalyse la méthylation des lysines 27 des histones H3 environnantes. Ces modifications constituent des sites d'ancrage pour le complexe PRC1 qui reconnaît H3K27me3 via le chromodomaine de Pc. Le PRC1 réalise l'ubiquitinylation de H2AK119 et compacte la chromatine, ce qui réprime la transcription des gènes ciblés.

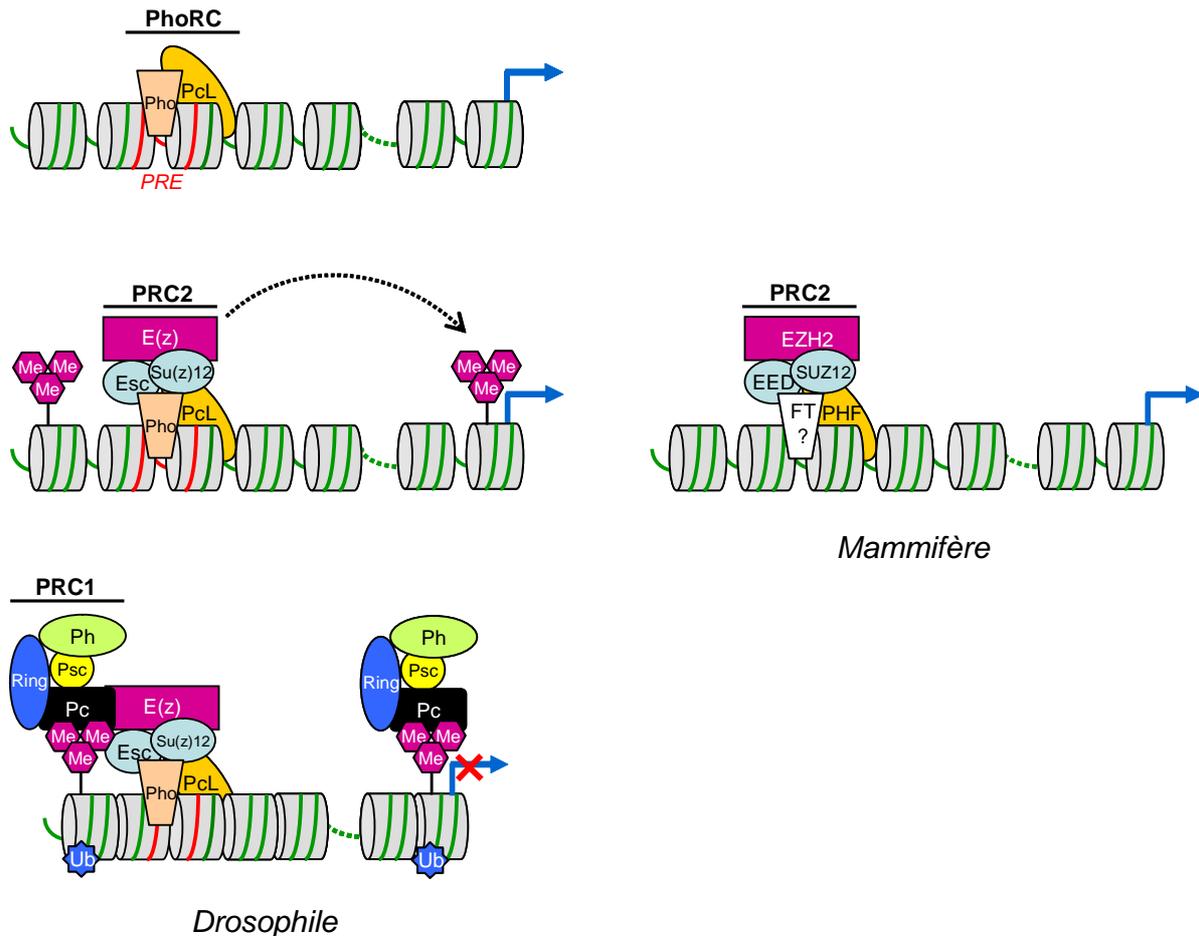


Figure 8 - Modèle actuel pour décrire l'établissement de l'hétérochromatine facultative au niveau des gènes homéotiques chez la drosophile. Recrutement séquentiel des différents complexes Polycomb: PhoRC au niveau des PRE (Polycomb Response Elements), puis PRC2 qui catalyse la tri-méthylation de H3K27, qui constitue un site de fixation pour le PRC1, via la protéine Pc. Les PRE n'existent pas chez les mammifères mais les complexes PRC2 et PRC1 sont conservés.

2.4.2 Chez les mammifères

Des PRE n'ont pas encore été identifiés, l'interaction entre les protéines YY1 et YY2, homologues de Pho et Phol, avec les PRC n'a pas été associée à la répression par les Polycomb. Cependant le recrutement du PRC2 au niveau des promoteurs des gènes cibles entraîne une méthylation de H3K27. Les homologues de Pcl, dont certains interagissent directement avec des facteurs de transcription (communication personnelle Boulay G & Leprince D) pourraient cibler les PRC2. PRC1 est supposé coopérer avec le PRC2 pour maintenir/établir la répression en bloquant la transcription

via la mono-ubiquitylation de H2A par RNF2 et en influençant la compaction de la chromatine dans des cellules souches embryonnaires murines (Levine SS, 2004, Trends Biochem Sci / Stock JK, 2007, Nat Cell Biol / van der Stoop P, 2008, PLoS One).

2.4.3 Le modèle n'est pas universel

Certaines études ont permis de nuancer ce modèle de recrutement des PRC. Par exemple, au niveau du X inactif l'ARN *Xist* peut recruter le PRC1 en absence du PRC2, et des régions déplétées en H3K27me3 peuvent être ciblées par le PRC1 (Schoeftner S, 2006, EMBO J).

3 Le noyau du PRC1 à la loupe

Un mécanisme commun d'évolution est la duplication d'un gène et des divergences ultérieures des séquences codantes ou des éléments régulateurs. Les gènes PcG ont probablement subi de nombreux événements de duplication au cours de leur histoire évolutive. Les invertébrés existants comme la drosophile possèdent des copies uniques des protéines PcG, à une exception près pour les Psc. A l'opposé, les espèces vertébrées possèdent de multiples paralogues de la plupart des PcG. L'exemple le plus frappant de l'expansion des PcG concerne la famille Pc.

Les domaines clefs des homologues PcG sont hautement conservés entre les organismes distants d'un point de vue évolutif mais également parmi les paralogues dans un organisme donné. Cependant, en dehors des domaines clefs, les protéines PcG ont parfois divergé significativement par rapport à leur équivalent chez la drosophile, et même entre paralogues. Ceci laisse penser que les paralogues possèdent des profils d'expression et des fonctions spécialisés. Un des challenges important est de comprendre la signification fonctionnelle de l'émergence des paralogues et comment cela influence la composition du PRC1, le ciblage génomique et/ou les mécanismes de répression transcriptionnelle.

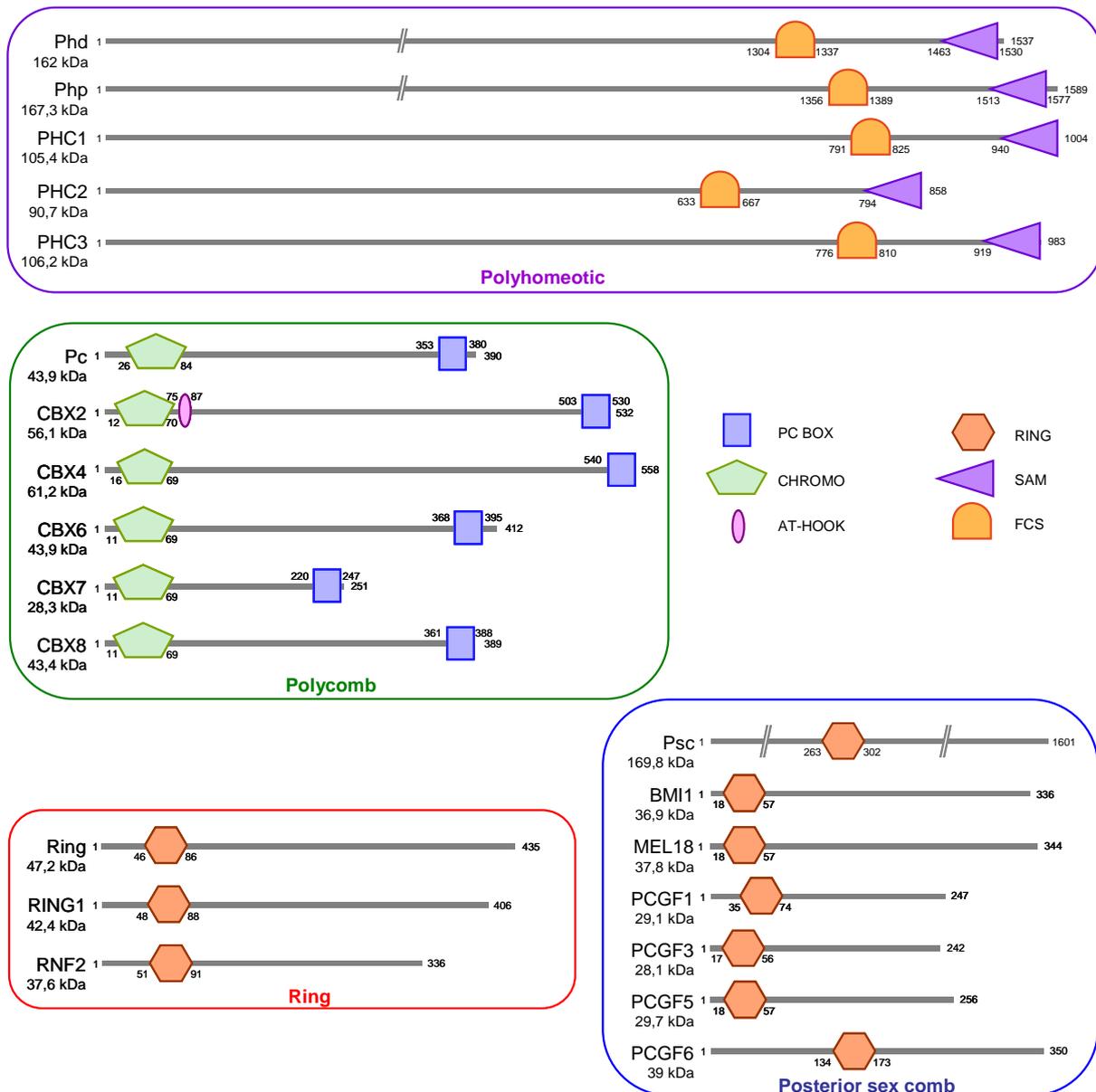


Figure 9 - Représentation schématique des protéines composant le cœur du PRC1 chez la drosophile (minuscules) et leurs homologues chez l'homme (majuscules)

3.1 Les homologues Pc

Représentée par un gène unique chez les invertébrés, ont été identifiés à ce jour cinq homologues Pc chez les mammifères, plus connus sous le nom de CBX. Les protéines homologues à Pc possèdent toutes un chromodomaine dans leur extrémité N-terminale et une « boîte Polycomb » (Pc box) dans leur extrémité C-terminale (Muller J, 1995, EMBO J / Schoorlemmer J, 1997, EMBO J). Le chromodomaine reconnaît spécifiquement H3K27me3. La Pc box est un motif d'une quinzaine d'acides aminés nécessaire pour la répression transcriptionnelle des gènes cibles et pour l'interaction avec Ring, la sous-unité catalytique active du PRC1. Des différences entre les paralogues mammifères existent au niveau de leur taille, de la structure de leurs domaines et de leurs propriétés biochimiques: ces protéines pourraient avoir certaines fonctions communes et d'autres divergentes.

Elles affectent de manière différente la régulation du cycle cellulaire. La sur-expression de CBX7 ou CBX8 mais pas CBX4 permet à des fibroblastes humains et murins d'échapper à la sénescence répliquative en partie en raison de la répression du locus INK4a-ARF (Gil J, 2004, Nat Cell Biol / Dietrich N, 2007, EMBO J). Cependant, elles réalisent cela dans un contexte de complexes PRC1 distincts : CBX8 dépend de Bmi1 (orthologue Psc) pour se fixer sur le locus INK4a-ARF pour prolonger la durée de vie de la cellule, alors que l'échappement de la sénescence induit par CBX7 est indépendant de Bmi1. Il serait nécessaire de démontrer si les complexes PRC1-CBX7 contiennent préférentiellement un autre orthologue Psc (Tableau 2) pour réaliser la répression CBX7-dépendante de ce locus suppresseur de tumeur. CBX2 contrôle également négativement le locus INK4A/ARF dans les cellules lymphoïdes et les cellules souches embryonnaires murines, la mécanistique n'a pas été détaillée (Coré N, 2004, Oncogene).

Les protéines CBX interagissent spécifiquement avec d'autres protéines non-Polycomb. CBX4 est le seul membre de la famille qui se lie avec le co-répresseur transcriptionnel CtBP (C-terminal binding protein) (Sewalt RG, 1999, Mol Cell Biol). Ce dernier interagit avec des facteurs de transcription qui pourraient cibler les complexes contenant CBX4 vers des séquences spécifiques d'ADN. CBX4 est également le seul des orthologues Pc à posséder une activité E3 SUMO ligase à ce jour (Kagey MH, 2003, Cell). Toutes les cibles de CBX4 SUMO n'ont pas été identifiées, cependant la sumoylation de plusieurs régulateurs transcriptionnels, dont CtBP, est assurée par CBX4 (Rosic A, 2006, Mol Cell).

Des études de BiFC (Bimolecular fluorescence complementation) ont permis de déterminer les distributions sub-nucléaires des protéines CBX associées à la chromatine dans des cellules vivantes (Vincenz C, 2008, PNAS). Différentes CBX se fixent sur des régions distinctes de la chromatine. Ni le chromodomaine, ni la Pc box ne sont essentiels pour l'association avec la chromatine ou RNF2 dans des cellules COS1 (African green monkey kidney fibroblast-like). De plus, cette étude révèle que les sites de fixation des CBX sur les autosomes n'est pas superposable à la marque H3K27me3 dans des cellules souches embryonnaires humaines et murines. Notons pour finir que la protéine CBX6 a été très peu étudiée, bien qu'elle possède une structure proche des autres membres de la famille CBX.

3.2 Les homologues Psc

Mel-18 et Bmi1 (2 des 6 orthologues Psc présentes chez les mammifères) sont également des paralogues non redondants, en dépit de 63% d'identité de séquence en acides aminés. Les souris déficientes en Bmi1 et Mel-18 démontrent des phénotypes similaires mais uniques (Akasaka T, 1996, Development / van der Lugt NM, 1994, Genes Dev), et seulement 30% des gènes régulés par Bmi1 sont co-régulés par Mel-18 et vice-versa (Wiederschain D, 2007, Mol Cell Biol). Dans certains cas, Mel-18 et Bmi1 ont des activités opposées sur la régulation du cycle cellulaire. Les souris transgéniques E μ -myc possèdent dans leur génome l'oncogène myc placé sous le contrôle de l'amplificateur de la chaîne lourde μ d'immunoglobuline, elles développent naturellement des malignités lymphoïdes précoces (Adams JM, 1985, Nature). L'insertion rétrovirale de Bmi1 dans ces souris accélère l'apparition des lymphomes (Haupt Y, 1991, Cell) = activité oncogénique, alors que des cellules d'une autre souris transgénique, sur-exprimant Mel-18, arrêtent leur cycle cellulaire au stade G1/S = activité suppresseur de tumeur (Sparmann A, 2006, Nat Rev Cancer). De plus, la

suppression de l'expression de Mel-18 dans des fibroblastes favorise l'émergence de tumeurs lorsqu'elles sont injectées dans des souris (Kanno M, 1995, EMBO J / Tetsu O, 1998, Immunity). Cependant, la relation entre ces 2 paralogues est complexe. La répression de Bmi1 et Mel-18 par des interférences à l'ARN dans des lignées cellulaires cancéreuses révèle paradoxalement des conséquences similaires, à savoir l'inhibition de la croissance cellulaire (Wiederschain D, 2007, Mol Cell Biol). De plus, Mel-18 peut agir en tant que suppresseur de tumeur en réprimant Bmi1 dans des cellules cancéreuses (Guo WJ, 2007, Cancer Res). Ces divergences constituent un des points importants à résoudre pour les futures thérapies anti-cancéreuses. Bien que Bmi1 et Mel-18 forment des complexes PRC1 stables de compositions similaires, seul Bmi1 a été décrit comme étant capable de réguler de manière positive l'ubiquitinylation de H2AK119 (Cao R, 2005, Mol Cell). Une étude structurale suggère que Bmi1 augmente l'activité de RNF2 en stabilisant l'interaction entre la E2 ligase, RNF2 (E3 ligase) et le substrat H2A (Li Z, 2006, JBC). Bmi1 possède une région de reconnaissance de la surface de H2A qui n'a pas été conservée dans Mel-18. De plus, Bmi1 contrôle l'auto-ubiquitinylation de RNF2 au niveau de lysines spécifiques pour générer des chaînes atypiques qui sont nécessaires pour la mono-ubiquitinylation de H2A. Par conséquent, la présence de différents homologues Psc dans le PRC1 pourrait modifier son activité catalytique *in vivo*.

3.3 Les homologues de Ring

Les homologues de Ring chez les vertébrés, RING1 et RNF2, possèdent également des divergences fonctionnelles. Bien qu'elles partagent de longues séquences fidèlement conservées, les souris déficientes en RING1 et RNF2 ont des phénotypes radicalement différents (del Mar Lorente M, 2000, Development / Voncken JW, 2003, PNAS). Les souris hétérozygotes pour RING1 exposent des transformations homéotiques classiques et un défaut au niveau du squelette, alors que les souris hétérozygotes RNF2 ne montrent aucun phénotype au niveau du squelette. Cependant, RNF2 est essentiel pour une gastrulation normale et les embryons qui en sont déficients ne survivent pas plus de 10,5 jours. La gravité différente de ces 2 phénotypes corrèle avec l'étendue de la déplétion de H2Aub dans ces souris KO. L'ubiquitinylation globale de H2A est dramatiquement réduite dans les cellules souches embryonnaires déficientes en RNF2 par rapport à RING1 (de Napoles M, 2004, Dev Cell). Un travail biochimique supporte les fonctions différentes observées dans l'étude sur les souris. Des protéines recombinantes pleine longueur de RNF2 mais pas RING1 possèdent une activité ubiquitine-ligase pour H2A, et l'association de RNF2 avec RING1 augmente cette activité (Wang H, 2004, Nature / Cao R, 2005, Mol Cell). Cependant, les situations *in vivo* sont probablement plus nuancées que ce que les résultats obtenus *in vitro* suggèrent. Bien que la modification H2Aub globale soit dramatiquement réduite dans les cellules déficientes en RNF2, le marquage de H2Aub est maintenu au niveau du chromosome X inactif (de Napoles M, 2004, Dev Cell), seules les cellules doublement KO RING1/RNF2 perdent H2Aub au niveau de cette structure, révélant la redondance fonctionnelle de RING1 et RNF2 dans certains contextes.

3.4 Les homologues de Ph

Chez la drosophile, le locus *polyhomeotic* code pour 2 protéines : Phd (distal) et Php (proximal), qui possèdent des profils de localisation exactement superposables avec celui de Pc sur les chromosomes polytènes (Franke A, 1992, EMBO J). Environ 80 sites de fixation pour la protéine Php ont pu être identifiés par immuno-localisation (DeCamillis M, 1992, Genes Dev). Il existe 3 homologues chez les mammifères : PHC1, PHC2 et PHC3 qui ont été purifiées dans un même complexe PRC1 (Levine SS, 2002, Mol Cell Biol). Toutes ces protéines possèdent 2 domaines remarquables. D'une part un domaine SAM carboxy-terminal (Sterile α Motif) de 70 acides aminés, responsable d'interactions protéine-protéine, en particulier d'homo- ou d'hétéro-oligomérisations, par exemple avec le SAM de Scm (Kim CA, 2005, JBC). De plus, le domaine SAM posséderait la capacité de se lier à l'ARN (Kim CA, 2003, Trends Biochem Sci). D'autre part il existe un motif FCS qui correspond à un domaine à doigt de zinc susceptible de se fixer à l'ARN (Zhang H, 2004, Mol Cell) ou l'ADN. Les souris déficientes pour PHC2 montrent des transformations homéotiques et leurs fibroblastes embryonnaires rentrent prématurément en sénescence (Isono K, 2005, Mol Cell Biol). PHC3 interagit avec le facteur de transcription E2F6 en G₀, cette interaction est perdue dans des cellules en prolifération et dans des cellules d'ostéosarcomes, suggérant que la perte d'expression de cette protéine pourrait favoriser la tumorigenèse (Deshpande AM, 2007, Oncogene).

Drososophile	Domaine	Activité	Orthologues humains	Complexe
Pc	chromodomaine, Pc box	Fixation sur H3K27me3	CBX2, CBX4, CBX6, CBX7, CBX8	PRC1
Php, Phd	SAM, FCS	Fixation ARN ?	PHC1, PHC2, PHC3	PRC1
Psc	RING		BMI1, Mel-18, PCGF-1-3-5 et 6	PRC1
Ring	RING	E3 ligase H2AK119ub1	RING1, RNF2	PRC1
Scm	2 x MBT, SAM, FCS		SCMH1, SCML2 (pas de FCS)	PRC1 ?
E(Z)	SET, SANT	Méthylation de H3K27	EZH1, EZH2	PRC2
Esc	5 x WD40	Cofacteur de E(Z)	EED	PRC2
Escl	5 x WD40	Cofacteur de E(Z)	EED	PRC2
Su(z)12	Zn-finger C2H2, VEFS box		SUZ12	PRC2
Pcl	2 x PHD, TUDOR		PHF19 (Pcl3), MTF2 (Pcl2), PHF1 (Pcl1)	PRC2
Pho	4 x Zn-finger C2H2	Liaison à l'ADN	YY1, YY2	PhoRC
Phol	4 x Zn-finger C2H2	Liaison à l'ADN	YY1, YY2	PhoRC ?
Sfmbt	4 x MBT, SAM, FCS	Fixation sur H3K9 et H4K20me1 et me2	L3MBTL2, MBTD1 (pas de SAM) SFMBT1, SFMBT2 (pas de FCS)	PhoRC

Tableau 2 - Protéines Polycomb et complexes associés chez la drosophile et l'Homme

4 Les gènes cibles des PcG

Pendant longtemps, les PcG ont été suspectés de cibler d'autres gènes en plus des clusters de gènes Hox. Les mécanismes PcG sont indispensables à de nombreux processus, de la spermatogenèse (Chen X, 2005, Science) au renouvellement des cellules souches neurales et hématopoïétiques (Molofsky AV, 2003, Nature / Lessard J, 2003 Exp Hematol / Park IK, 2003, Nature) et la progression du cycle cellulaire (Martinez AM, 2006, Genes Dev). La combinaison de l'immunoprécipitation de la

chromatine associée aux puces à ADN (ChIP on chip) ou au pyroséquençage a permis d'apporter une image à l'échelle du génome. Ces études montrent que, chez les mammifères et les drosophiles, un grand nombre de gènes sont ciblés par les PcG. Une grande proportion de ces gènes code pour des régulateurs transcriptionnels, des morphogènes, des récepteurs et des protéines de signalisation qui sont impliqués dans toutes les principales voies du développement.

De plus, bien que les mutants PcG de drosophile et les souris déficientes pour des membres individuels du PRC1 partagent des effets homéotiques similaires, chacun possède un phénotype unique, suggérant que différents complexes PRC1 existent et possèdent certainement quelques gènes cibles non redondants.

4.1 Gènes Hox et régulateurs du développement

- Drosophile

Une étude *in silico* a permis de prédire 167 séquences PRE dans le génome de la drosophile, et certaines ont été validées expérimentalement (Ringrose L, 2003, Dev Cell). Parmi toutes ces cibles, un grand nombre correspond à des gènes agissant précocement au cours du développement comme les gènes responsables : de la segmentation, du profil des disques imaginaux, de l'oogenèse, de l'organisation des chromosomes et du contrôle du cycle cellulaire. Une autre étude, présentant des résultats de ChIP on chip, détaille les sites de fixation des protéines Pc et Ph sur les chromosomes 2 et X (Nègre N, 2006, PLoS Biol). Pc et Ph co-localisent au niveau de larges régions chromosomiques, plus de 70 séquences communes d'ADN ont été identifiées, représentant 37 gènes uniques. D'une manière intéressante, les profils de liaison de ces protéines ont un comportement dynamique pendant le développement embryonnaire. Par exemple, de nouveaux sites de fixation apparaissent à des stades plus tardifs ou certains sites présents à un stade précoce disparaissent au cours du temps, suggérant que la répression par les PcG est un processus dynamique. A l'aide du même type d'approche, 2 autres études ont été publiées sur des cellules drosophiles en culture (Schwartz YB, 2006, Nat Genet / Tolhuis B, 2006, Nat Genet). Elles ont confirmé que les complexes PcG co-localisaient et ciblaient de nombreux régulateurs du développement et des facteurs de transcription. De plus, les régions ciblées par Pc sont enrichies en H3K27me3.

- Mammifères :

Une étude à l'échelle du génome dans des fibroblastes embryonnaires humains a mis en évidence que de nombreux gènes ciblés par les complexes PRC1 et PRC2 correspondaient aux orthologues des gènes prédits par l'analyse *in silico* chez la drosophile, en particulier les gènes homéotiques et les gènes de la segmentation (Bracken AP, 2006, Genes Dev). A ceux-là s'ajoutent plus d'un millier d'autres cibles, des gènes impliqués dans différents processus de différenciation : musculaire, neuronale, osseuse, sexuelle, épidermique et l'hématopoïèse. Des analyses de ChIP on chip avec des membres du PRC1 et du PRC2 dans des cellules souches humaines et murines ont présenté des résultats superposables. Les gènes ciblés sont importants pour le développement, la morphogenèse, l'organogenèse, la neurogenèse, la signalisation cellulaire et la régulation transcriptionnelle (Boyer LA,

2006, Nature). Durant la différenciation, l'occupation des sites décrits précédemment par les complexes PcG est drastiquement réduite, démontrant que les gènes réprimés par les PcG dans les cellules souches embryonnaires sont potentiellement activés lorsque les cellules perdent leur état totipotent. Le concept *trivial* qui émerge de ces résultats suggère que pour maintenir la totipotence, la cellule utilise les complexes PcG afin de maintenir la répression des régulateurs du développement, et que pour se différencier, elle les dé-réprime.

4.2 Cellules souches et domaines bivalents

Il a été proposé que dans les cellules souches embryonnaires, la totipotence est corrélée à des états « bivalents » de la chromatine. Dans ces cellules, les gènes ciblés par les PcG possèdent simultanément la marque répressive H3K27me3 et la marque activatrice H3K4me3 (Bernstein BE, 2006, Cell / Azuara V, 2006, Nat Cell Biol). Les mécanismes résultant de la mise en place de ces 2 marques sont encore inconnus. Pendant la différenciation, ces domaines bivalents possèdent exclusivement une des 2 modifications activatrices ou répressives, indiquant que cette bivalence est une fonction importante des gènes ciblés dans les cellules souches embryonnaires. De plus, une analyse des localisations génomiques des complexes PRC1 et PRC2 conclut que les domaines bivalents dans les cellules souches murines et humaines peuvent être classés en 2 catégories, d'une part les régions qui contiennent les deux complexes PRC, d'autre part les régions enrichies en PRC2 exclusivement (Ku M, 2008, PLoS Genet). Les îlots CpG présents au niveau des domaines bivalents, de par leur longueur, leur état de méthylation et leur proximité les uns des autres, définiraient les domaines spécifiques de fixation des PRC à l'image des PRE chez la drosophile. Les conclusions de ces analyses suggèrent que les PcG pourraient maintenir la totipotence des cellules souches, cependant une autre étude a réfuté cette hypothèse car les mutants EED maintiennent leur totipotence à travers plusieurs cycles cellulaires (Chamberlain SJ, 2008, Stem cells).

5 Marques épigénétiques associées aux PRC1

5.1 Les « corps Polycomb » et la sumoylation

Les complexes PcG forment des structures nucléaires associées de manière stable aux régions péricentromériques des chromosomes autosomes humains (Saurin AJ, 1998, J Cell Biol). En particulier, les corps Polycomb ont été retrouvés au niveau de l'hétérochromatine péricentromérique du chromosome 1 de plusieurs lignées cellulaires. Des structures similaires ont été décrites chez la drosophile (Buchenau P, 1998, J Cell Biol) et le nématode (Zhang T, 2006, Development). Les corps Polycomb semblent représenter des centres d'extinction ou des sous-domaines nucléaires qui contiennent de la chromatine silencieuse. Ils sont considérés comme des sites de sumoylation depuis que l'activité E3 sumo-ligase de CBX4 a été découverte (Kagey MH, 2003, Cell / Kagey MH, 2005, EMBO J). En effet, CBX4 recrute CtBP et l'enzyme de conjugaison Ubc9 au niveau des corps polycomb, où la sumoylation de CtBP est effectuée. D'autres protéines sont probablement le substrat de cette réaction au niveau de ces régions spécialisées.

5.2 Ubiquitinylation de H2A

La protéine Ring de drosophile et H2AK119ub1 co-localisent au niveau du promoteur du gène *Ubx* et la déplétion de Ring dans des cellules SL2 en culture entraîne la perte de cette modification et la perte de la répression de *Ubx* (Wang H, 2004, Nature). Chez les mammifères, la sous-unité catalytique responsable de l'ubiquitinylation de H2A semble être RNF2, bien que RING1 et BMI1 semblent également toutes les deux participer à l'activité E3 ligase *in vivo* (Cao R, 2005, Mol Cell). L'étude de souris knock-out pour *Bmi1* a démontré que l'absence de BMI1 affectait l'ubiquitinylation de H2A au niveau du promoteur du gène homéotique *HoxC*, entraînant une sur-expression de ce dernier et que la méthylation de H3K27 par EZH2 précédait cet événement car cette modification était toujours présente. Le rôle exact de H2AK119ub dans le processus de répression n'a pas été clairement défini, il pourrait servir, comme les méthylations, de site d'ancrage pour d'autres complexes répresseurs ou alors pourrait directement bloquer le remodelage de la chromatine ou les facteurs de transcription. Pour finir, H2AK119ub est enrichi au niveau du chromosome X inactif. L'ubiquitinylation par RING semble jouer un rôle dans l'initiation de l'inactivation du X (de Napoles M, 2004, Dev Cell / Fang J, 2004, JBC / Hernandez-Munoz I, 2005, PNAS).

5.3 Machinerie ARNi

La répression transcriptionnelle des transgènes présents en multiples copies chez la drosophile, dont la répression nécessite la présence de la machinerie ARNi, semble être dépendante des PcG (Pal-Bhadra M, 1997, Cell). L'orthologue murin de Ph, PHC1, est capable de se lier à l'ARN (Zhang H, 2004, Mol Cell). Le domaine responsable de la liaison à l'ARN est une région en doigt de zinc appelée domaine FCS, qui est également présent au niveau des homologues Ph, Scm et Sfmtb (Tableau 2). Chez les mammifères, très peu de données expliquent comment les PRC2 sont recrutés. Des études récentes ont impliqué la machinerie ARN (Rinn JL, 2007, Cell), cependant cela ne doit pas être un mécanisme général puisqu'un traitement à la ribonucléase de cellules U2OS (Human Bone Osteosarcoma Epithelial Cells) n'affecte pas la localisation globale de EZH2 ou SUZ12 (Aoto T, 2008, JBC). De plus, à quelques exceptions près, le recrutement des PcG au niveau des PRE n'est pas affecté suite des mutations de la machinerie ARNi chez la drosophile (Grimaud C, 2006, Cell).

5.4 Déméthylases de H3K27me3

Les déméthylases des groupes JARID (H3K4) et UTX/JMJD3 (H3K27) jouent un rôle au niveau des de la différenciation des cellules (Agger K, 2007, Nature / Christensen J, 2007, Cell / De Santa F, 2007, Cell / Lan F, 2007, Nature / Lee MG, 2007, Science). En réponse à la différenciation des cellules NT2 par l'acide rétinoïque, ces enzymes sont recrutées ou déplacées des promoteurs des gènes homéotiques. Elles apparaissent indispensables pour leur régulation, révélant une relation directe entre les déméthylation de H3K4 et H3K27 et le contrôle de l'activité transcriptionnelle pendant la différenciation. Ces données confirment que les complexes PcG sont dynamiques, en particulier au cours du développement.

6 Conclusions

Le nombre de gènes Hox s'est agrandi au cours de l'évolution des vertébrés. Ces gènes sont typiquement organisés en « clusters » génomiques, la drosophile possède un cluster unique, tandis que les vertébrés possèdent de multiples clusters répartis sur des chromosomes différents (Finnerty JR, 1998, Curr Opin Genet Dev). Depuis que les gènes Hox sont considérés comme des cibles conservées des protéines PcG et comme les clefs du déterminisme de la spécification du plan du corps des organismes bilatéraux, il est possible d'émettre l'hypothèse qu'après la duplication génomique des gènes Hox et PcG, ils se soient diversifiés et aient co-évolués afin d'assurer de nouveaux rôles dans le développement des vertébrés. En plus du développement, les membres du PcG sont connus pour réguler la mémoire cellulaire épigénétique, la pluripotence et le renouvellement des cellules souches. L'expansion des gènes PcG et leur diversification contribuent significativement à la complexité des mécanismes hérités de répression des gènes dans les organismes multicellulaires. Chez les vertébrés, les protéines PcG répriment les gènes *HOX*, de plus elles sont régulatrices du cycle cellulaire en se fixant sur le locus *INK4A-ARF* et ciblent un grand nombre de gènes contrôlant le développement (la croissance cellulaire et la prolifération). Très peu d'éléments sont connus sur la régulation des gènes PcG ou à propos des modifications post-traductionnelles qui pourraient influencer l'activité de leurs protéines. Les connaissances sur les modifications post-traductionnelles de ces protéines sont également limitées. BMI1 est ubiquitinylé au niveau de l'X inactif (Hernandez-Munoz I, 2005, PNAS), son association à la chromatine semble être dépendante de son état de phosphorylation (Voncken JW, 1999, J Cell Sci / Voncken JW, 2005, JBC). Si on acquiert la connaissance nécessaire pour comprendre comment les PcG programment les gènes du développement pour induire la différenciation des cellules souches embryonnaires, alors les opportunités de reprogrammation de cellules différenciées afin de ré-acquérir la pluripotence ou alors d'entrer dans une voie de différenciation alternative deviendraient envisageables. Enfin, comprendre le rôle de chaque paralogue permettrait de mieux appréhender la fonction des multiples complexes PRC qui semblent intervenir pour le contrôle de divers processus cellulaires.

Les méthodes d'étude des interactions protéine-protéine et purification de complexes

Au fur et à mesure que le séquençage du génome a avancé, les chercheurs ont revu à la baisse le nombre de gènes que possédaient les humains, passant de 120000 en 2000 (Liang F, 2000, Nature Genetics) à environ 20-25000 depuis 2004 (International Human Genome Sequencing Consortium, 2004, Nature). Les épissages alternatifs et les modifications post-traductionnelles remplissant à merveille leur rôle dans la multiplicité des formes protéiques exprimées à partir d'un gène unique. C'est ainsi qu'après l'ère génomique, est apparue depuis quelques années l'ère de la protéomique. Tandis que le génome de toutes les cellules d'un organisme est le même, le protéome (ensemble des protéines exprimées au sein de la cellule) est quand à lui bien spécifique à chaque type cellulaire. C'est ainsi que des études de protéomes ont vu le jour avec l'apparition de la chromatographie bidimensionnelle conjuguée à la spectrométrie de masse.

Dans la cellule, les protéines agissent rarement seules, mais plutôt sous forme de complexes dans lesquels chaque protéine aura un rôle bien précis : enzymatique, structural, adressage...

Les interactions protéine-protéine existent sous deux formes :

- permanente : elle est très stable et les protéines sont purifiées uniquement sous leur forme associée
- transitoire : les 2 protéines existent sous forme dissociée ou et/ou sous forme complexée.

Plusieurs techniques ont été développées pour l'analyse des interactions protéine-protéine. La plus utilisée est la technologie dite du double-hybride.

1 Le double-hybride (DH)

Introduite en 1989 (Fields S, 1989, Nature) pour identifier des interactions de type protéine-protéine dans la levure, cette technique a ensuite été étendue à l'analyse des interactions de type protéine-ADN et protéine-protéine dans la bactérie (Joung J, 2000, PNAS).

1.1 Principe de la technologie DH

Le principe de base est le suivant (Figure 9) : il faut disposer d'une souche de levure contenant dans son génome un gène rapporteur (LacZ, histidine) dont la séquence promotrice n'active pas la transcription, en fusionnant d'une part une protéine A (appât) à un domaine de liaison à l'ADN (GAL4) qui va reconnaître le promoteur du gène rapporteur (UAS) et d'autre part une protéine B (proie) à un domaine activateur (DA) de la transcription. En co-exprimant les 2 protéines de fusion dans la levure, la protéine A va se fixer au niveau du promoteur du gène rapporteur. Si une interaction physique existe entre les protéines A et B, alors le gène rapporteur sera exprimé grâce au domaine activateur de la transcription fusionné à B. Dès qu'une interaction est identifiée (coloration bleue ou croissance sur milieu en absence d'histidine), elle est validée par réciprocité (protéine B fusionnée au domaine de liaison à l'ADN et protéine A fusionnée à l'activateur de la transcription).

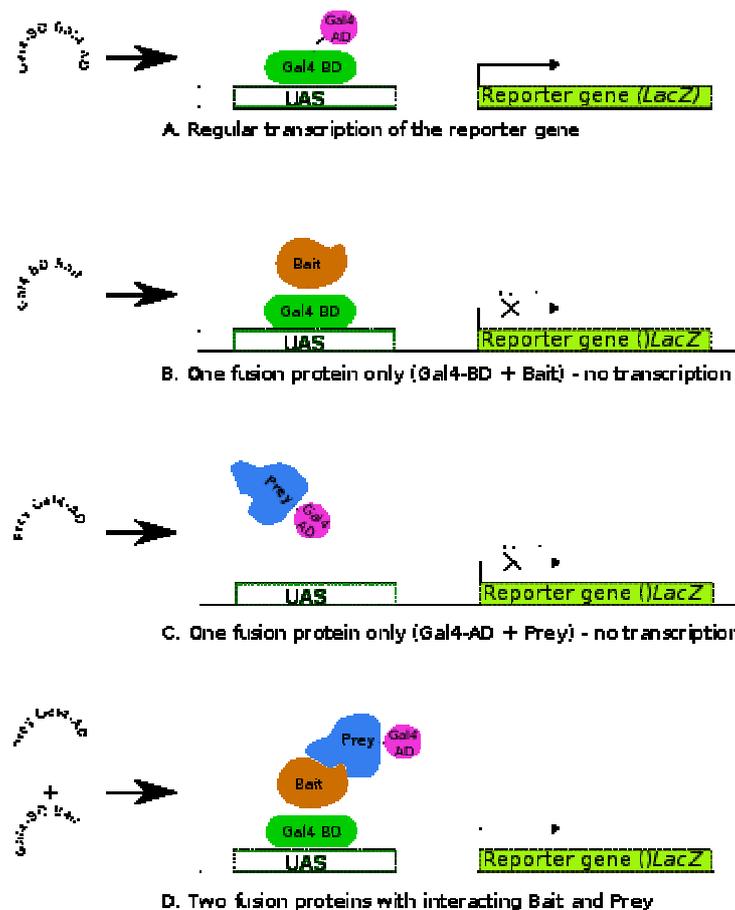


Figure 10 – Principe de la technique dite du « double hybride » (d'après wikipedia). A - En présence du domaine Gal4-BD, la protéine se fixe sur la séquence UAS, la présence de Gal4-AD permet la transcription du gène rapporteur. B – En absence de Gal4-AD, pas de transcription du gène rapporteur. C – En absence de Gal4-BD, la protéine ne se fixe pas sur UAS, pas de transcription du gène rapporteur. D – En présence d'un appât (bait) fusionné à Gal4-BD qui interagit avec une cible (prey) fusionnée à Gal4-AD, la transcription du gène rapporteur est détectable. AD = domaine d'activation, BD = domaine de fixation.

1.2 Limitations

Cette technologie permet d'identifier aussi bien des interactions permanentes que transitoires, cependant il existe des limitations car certains appâts sont auto-activateurs. Aussi, certaines interactions validées en DH n'existent pas physiologiquement (faux positifs). Enfin, d'autres limitations ont été décelées : les repliements structuraux, la stabilité, l'éventuelle toxicité, la mauvaise localisation et les modifications post-traductionnelles des protéines fusion peuvent affecter les interactions potentielles qui ne sont alors pas détectées (faux négatifs).

Pour pallier à ces problèmes, les cribles double-hybride utilisent désormais des domaines protéiques plutôt que des protéines entières. C'est ainsi que les protéines de nombreuses espèces ont ainsi pu être analysées à l'aide d'études à grande échelle : *E. coli* (Butland G, 2005, Nature), *S. cerevisiae* (Fromont-Racine M, 1997, Nature Genet), *D. melanogaster* (Giot L, 2003, Science), *C. elegans* (Li S, 2004, Science) et *H. sapiens* (Stelzl U, 2005, Cell) générant les éléments de base nécessaires à la compréhension du fonctionnement des complexes protéiques au sein des cellules. Certains cribles sur

la même espèce possèdent peu de recouvrement, ce sont les domaines protéiques choisis qui peuvent justifier ces divergences.

Bien que les informations apportées par le double-hybride ne soient pas négligeables, elles ne permettent pas de statuer sur la composition exacte des complexes multi-protéiques présents dans les cellules. D'autres techniques d'analyse ont ainsi été élaborées, en particulier la purification d'affinité en tandem.

2 La purification d'affinité en tandem (TAP) : nouvel outil pour l'exploration des protéomes

Mise au point par une équipe de l'EMBL (European Molecular Biology Laboratory) dirigée par Bertrand Séraphin en 1999 (Rigaut G, 1999, Nat Biotechnol), la technologie TAP est actuellement la technique de référence pour la purification des complexes en condition native, même si des purifications d'affinité en une étape (Flag, biotinylation, anticorps) ou « tag-less » sont également utilisées.

2.1 Principe de la technologie TAP

Cette méthode simple et élégante a été décrite chez la levure, elle consiste à fusionner une protéine d'intérêt à une étiquette constituée de 3 régions importantes (Figure 10A) :

- un domaine de liaison à la calmoduline (affinité forte à la calmoduline en présence d'ions Ca^{2+})
- un site de clivage par la protéase TEV (Tobacco Etch Virus) isolée du tabac : dont la séquence consensus est Glu-X-X-Tyr-X-Gln-Ser (Dougherty WG, 1989, Virology) (X peut correspondre à différents acides-aminés mais pas n'importe lequel).
- une séquence contenant 2 domaines de la protéine A de *Staphylococcus aureus*, qui possèdent une forte affinité pour les immunoglobulines G (IgG).

L'étiquette est constituée de moins de 200 acides aminés, lui conférant une masse moléculaire de 20kDa environ. Pour rappel l'étiquette de la protéine fluorescente verte GFP (Green Fluorescent Protein) qui permet de localiser *in vivo* les protéines à l'aide d'un microscope à fluorescence, possède une masse de 27kDa. Comme toutes les étiquettes, la TAP peut-être fusionnée à l'extrémité amino- ou carboxy-terminale de la protéine d'intérêt.

La séquence nucléotidique de la protéine fusion est insérée dans le génome de l'hôte sous le contrôle d'un promoteur proche du promoteur endogène de la protéine d'intérêt. La protéine « TAP » est donc exprimée dans la cellule et est susceptible de se lier aux protéines avec lesquelles la protéine sauvage est associée naturellement. Une extraction des protéines totales dans des conditions natives permet de récupérer toutes les protéines complexées ou non qui sont exprimées chez l'hôte. Il faut s'assurer que la protéine TAP est bien dans la fraction soluble avant de poursuivre la purification.

La purification en tandem est schématisée en Figure 10B : l'extrait protéique est mis au contact de billes couplées à des IgG de lapin pour une première purification d'affinité, après plusieurs lavages les complexes retenus sont élués en ajoutant la protéase TEV dans la première colonne d'affinité. Ainsi la fixation et l'éluion sont basées sur 2 principes biochimiques différents, éliminant les contaminants qui se seraient fixés de manière aspécifique sur les billes IgG. L'éluat, additionné de calcium, est transféré

dans une seconde colonne de purification contenant des billes associées à la calmoduline, cette seconde étape permet de retenir les complexes associés à la protéine TAP et d'éliminer la protéase et d'autres contaminants. Après plusieurs lavages avec un tampon contenant des ions Ca^{2+} , les complexes sont élués par ajout d'acide éthylène-diamine/glycol-tétraacétique (EDTA/EGTA) dans la colonne. L'EDTA/EGTA est un chélateur d'ions bivalents qui va complexer les ions calcium, l'étiquette TAP va donc perdre son affinité pour les billes de calmoduline et les complexes se retrouveront dans la fraction mobile.

Les protéines éluées sont séparées par électrophorèse sur gel monodimensionnel de polyacrylamide en présence de dodécylsulfate de sodium (SDS-PAGE). Le gel est ensuite coloré au bleu de coomassie ou au nitrate d'argent afin de visualiser les protéines purifiées. L'identification de ces dernières est réalisée par spectrométrie de masse après une digestion protéolytique (trypsine) (approche « bottom-up »).

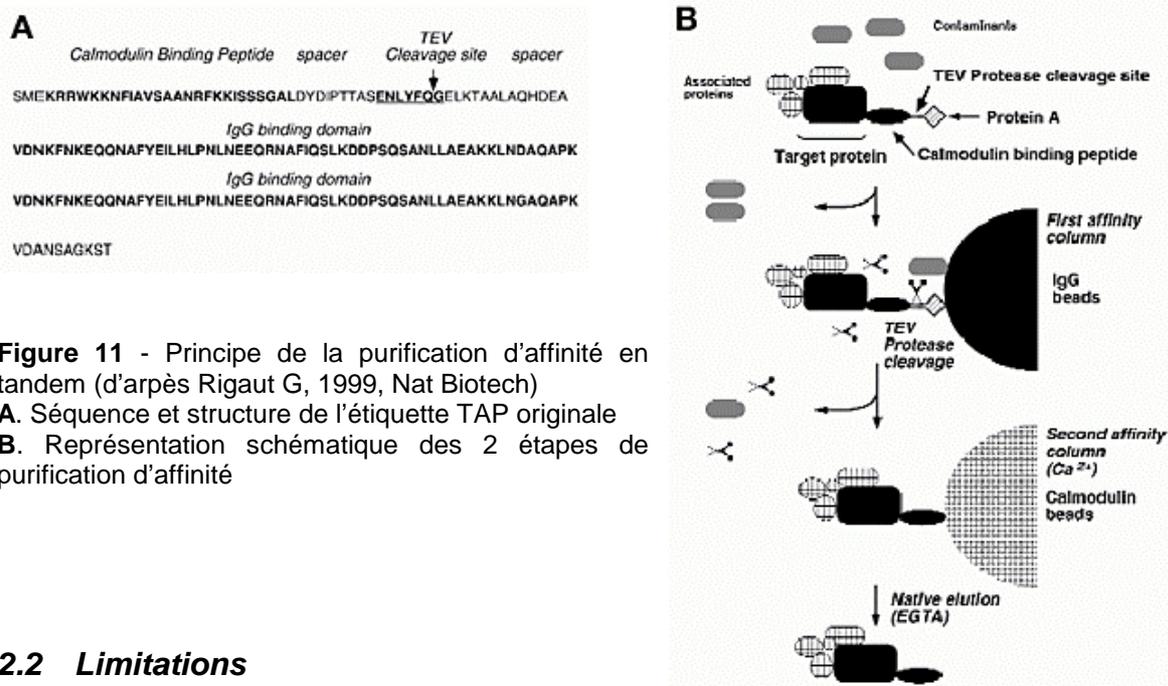


Figure 11 - Principe de la purification d'affinité en tandem (d'après Rigaut G, 1999, Nat Biotech)
A. Séquence et structure de l'étiquette TAP originale
B. Représentation schématique des 2 étapes de purification d'affinité

2.2 Limitations

La purification TAP permet de purifier les protéines associées par des interactions permanentes, dans des conditions proches des conditions physiologiques. Cependant certaines protéines majoritairement exprimées dans la cellule peuvent être identifiées en tant que faux positifs (contaminants), il en est de même pour les protéines qui possèdent une affinité pour l'étiquette TAP. Cette technologie permet de comparer la composition des complexes : en fonction du cycle cellulaire, après un stimulus (physique, chimique) et d'avoir une idée sur la stœchiométrie des sous-unités des complexes.

Cependant certaines limitations peuvent affecter l'efficacité de cette technique. La conformation de la protéine de fusion n'est parfois pas identique à la protéine endogène, ce qui limite, voire abolit, l'interaction avec ses partenaires protéiques, mais également influence sa localisation. Aussi, les repliements ou interactions peuvent affecter l'étiquette qui ne sera plus reconnue lors des 2 étapes de purification. Il est alors judicieux de placer l'étiquette à l'autre extrémité de la protéine. Les protéines

non solubles posent un problème au moment de l'extraction des protéines totales. Bien que le site de clivage par la protéase TEV soit rare, si une des protéines du complexe contient ce dernier, elle sera clivée lors de la première élution. Après les deux étapes de chromatographie d'affinité, des études *in vitro* de l'activité des complexes purifiés peuvent être réalisées mais l'EGTA peut potentiellement jouer un rôle dans la stabilité des complexes, cependant il n'a aucun effet sur les identifications par spectrométrie de masse.

2.3 Les étiquettes TAP pour la purification de complexes à partir de cellules de mammifères

Comme l'étiquette TAP a été décrite pour l'étude de complexes chez la levure (Gavin AC, 2002, Nature), et bien qu'elle soit également efficace pour la purification de complexes dans d'autres organismes ou types cellulaires : la bactérie *E. coli* (Gully D, 2003, FEBS Letters), le nématode *C elegans* (Polanowska J. et al, 2004, Biotechniques), le tabac *Nicotiana benthamiana* (Rohila JS, 2004, Plant J), le riz *Oryza sativa* (Rohila JS, 2006, Plant J), l'arabette *Arabidopsis thaliana* (Van Leene J, 2007, MCP), la drosophile *Drosophila melanogaster* (Veraksa A, 2005, Developmental dynamics) et les cellules de mammifères : souris (Bertwistle D, 2004, Mol Cell Biol) et humaines (Bouwmeester T, 2004, Nature Cell Biology), des équipes de recherche ont élaboré des variants de l'étiquette originale pour pallier à certaines limitations du TAP conventionnel.

- le rendement

Le **GS-TAP** (Bürckstümmer T, 2006, Nat Methods), développé pour la purification de complexes dans les cellules humaines, est constitué de 2 domaines de liaison aux protéines G, d'un site de clivage par la TEV et d'un peptide de liaison à la streptavidine (SBP). Cette étiquette a permis d'isoler la même quantité de complexes lié à IKK γ par rapport à l'étiquette TAP conventionnelle, en partant de 10 fois moins de cellules (HEK293). Cette optimisation du rendement de la purification permet de purifier plus efficacement les complexes présents en faible quantité dans les cellules.

- la stabilité de la protéine d'intérêt et l'encombrement stérique de l'étiquette

Comme l'étiquette TAP peut influencer le repliement de la protéine d'intérêt et/ou son activité, mettant en péril la purification et la véracité des partenaires identifiés, une étiquette de taille réduite a été mise au point. Le **SF-TAP** (Gloeckner CJ, 2007, Proteomics) ne fait que 4,6kDa mais il permet pourtant une purification d'affinité en tandem grâce à ses deux « petits » domaines Streptavidine et FLAG. Les éluions ne se font pas par ajout d'une protéase ou d'un chélateur mais d'un métabolite : la desthiobiotine pour la première dimension et d'un octapeptide FLAG pour la deuxième. Cette technique a permis de purifier les protéines associées à la protéine Raf (MEK1 et 14-3-3) dans des cellules humaines (HEK293). Il est à noter que le niveau d'expression de la protéine Raf fusionnée à l'étiquette conventionnelle dans les mêmes cellules était sensiblement moins important que Raf-SF-TAP, suggérant une instabilité de cette protéine quand elle est associée à une étiquette trop grande.

- TAP *in vivo*

La première souris knockin TAP date de 2004, mais les résultats de la purification n'ont pas été publiés dans l'article original, ni ailleurs (Zhou D, 2004, NAR). Jusqu'en 2006, toutes les purifications TAP effectuées dans des cellules de mammifères l'ont été à partir de cellules en culture, à l'exception d'une étude *in vivo* à partir d'une souris transgénique (Angrand PO, 2006, MCP). La protéine 14-3-3ζ a été fusionnée à l'étiquette TAP conventionnelle, et son expression contrôlée par le promoteur humain de l'ubiquitine C, permettant d'obtenir une expression comparable à celle de l'endogène et de purifier plus d'une trentaine de nouveaux partenaires protéiques par rapport à la même purification dans des cellules HEK293 (Brajenovic M, 2004, JBC), mettant en évidence la tissu-spécificité de certains complexes.

- les protéines non solubles (membranaires)

Les récepteurs de la mélatonine font partie de la superfamille des récepteurs couplés aux protéines G (GPCR), ils possèdent 7 domaines transmembranaires. Ils sont localisés au niveau de la membrane plasmique des cellules, ils sont donc insolubles avec le protocole classique de purification TAP. Une mise au point des conditions de solubilisation à partir des complexes associés à la membrane, notamment à l'aide de faibles quantités de détergents comme la digitonine, a permis de purifier les protéines associées aux récepteurs MT1 et MT2 dans des cellules HEK293 (Daulat AM, 2007, MCP). Ce fut la première fois que des protéines associées aux GPCR furent purifiées dans des conditions natives et dans des quantités compatibles avec l'identification par spectrométrie de masse.

- expression proche de l'endogène et haut débit

Pour les purifications d'affinité en tandem dans les cellules de mammifères, la séquence de la protéine d'intérêt fusionnée à l'étiquette TAP est sous le contrôle d'un promoteur artificiel qui a été ajouté en amont, et son expression dépend également du lieu où le rétrotransposon s'est inséré dans le génome. La protéine TAP est donc parfois largement surexprimée par rapport à l'endogène, ce qui peut altérer son rôle au sein de la cellule et entraîner un déséquilibre dans les complexes. L'utilisation de chromosomes bactériens artificiels (BAC) comprenant les régions promotrices et les éléments régulateurs des gènes d'intérêt permet de contrecarrer cet inconvénient. A l'aide d'une technique baptisée « recombineering », il est aisé d'introduire l'étiquette LAP (Purification d'Affinité et Localisation) dans les transgènes BAC (Poser I, 2008, Nat Methods) qui contiennent la séquence de la protéine d'intérêt et ses régions régulatrices naturelles. L'étiquette LAP est composée de 4 domaines : la protéine GFP, un site de clivage par la protéase PreScission, un peptide-S et un site de clivage par la TEV, permettant une élution finale dans des conditions natives. A l'aide d'une transfection stable dans des cellules de mammifères, cette technologie permet une étude complète de la protéine d'intérêt : sa localisation (GFP), ses partenaires protéiques (LAP) et ses gènes cibles en effectuant une immunoprécipitation de la chromatine avec un anticorps dirigé contre la GFP. Le protocole décrit est applicable sur des cellules souches embryonnaires de souris afin d'obtenir des animaux transgéniques (permettant, entre autres, d'étudier les protéines impliquées dans le développement). Toutes les étapes de cette technologie ont été automatisées et une étude à l'échelle du protéome entier est en cours.

3 Bilan et choix de la méthode

Bien que les 2 technologies détaillées précédemment présentent des avantages et des inconvénients, il est évident qu'elles apparaissent comme étant complémentaires pour la description des complexes. Cependant, le challenge de purifier des complexes à partir de cellules humaines en culture en utilisant la technologie TAP a été considéré comme étant un modèle plus « naturel » compte tenu de l'hôte utilisé pour les cribles double-hybride. Aussi, les complexes associés aux protéines Polycomb et HP1 semblent stables donc la méthode TAP est applicable. Enfin, compte tenu de la divergence entre le nombre de paralogues pour chaque famille de protéines entre l'Homme et la levure, une étude par DH n'apparaît pas comme étant un choix judicieux si une de ces protéines est nécessaire comme intermédiaire entre l'appât et la proie.

La technologie TAP est une méthode de purification générique qui semble tout à fait adaptée pour la comparaison entre membres d'une même famille protéique car le protocole peut être facilement standardisé. Alors il est relativement simple de générer une liste des contaminants, qui sont retirés de la liste des protéines identifiées. Cependant, un contaminant pour la plupart des protéines peut représenter un réel partenaire protéique pour une autre, et des analyses plus approfondies doivent permettre de statuer.

Nous avons opté pour l'étiquette TAP traditionnelle car les premières purifications ont été réalisées avant la publication du GS-TAP qui semble être devenue l'étiquette de référence pour les purifications TAP en cellules humaines. Notons que cette étiquette a également démontré son efficacité dans des cellules végétales (Van Leene J, 2008, Trends in Plant Sci) et chez la drosophile (Kyriakakis P, 2008, Fly).

OBJECTIFS DE LA THÈSE

Arrivé au laboratoire en octobre 2005 alors que le Docteur Pierre-Olivier Angrand venait d'être sélectionné pour installer une nouvelle équipe à l'Institut de Recherches Interdisciplinaires. J'ai intégré une équipe émergente qui a comptabilisé de 3 à 5 personnes au cours de mes 3 années de doctorat. La thématique de l'équipe concerne l'épigénétique d'une manière générale, les protéines de la superfamille CBX (HP1 et Polycomb) en particulier. Compte tenu de l'absence de « passé » de notre équipe, le premier objectif de ma thèse a été de participer à la mise en place d'une grande partie des techniques utilisées aujourd'hui en routine au laboratoire, en particulier la purification TAP et les études biochimiques pour la validation des partenaires identifiés.

En ce qui concerne la partie scientifique, mon sujet de thèse s'est inscrit dans le projet de l'équipe, tous les membres de l'équipe ont été associés à l'ensemble des projets initiés pendant ces 3 années, donc les objectifs scientifiques concernaient davantage un travail d'équipe plutôt qu'un travail individuel. Le point de départ scientifique de notre équipe de recherche était **l'analyse protéomique des complexes associés aux protéines CBX dans des cellules humaines**, afin de pouvoir statuer sur l'hypothèse suggérant qu'il existe différents complexes PRC1 dans les cellules (Otte AP, 2003, Curr Opin Genet Dev) et d'autre part d'identifier des partenaires spécifiques à chaque isoforme HP1.

Dans un premier temps, nous avons étudié la capacité des protéines CBX à réprimer la transcription d'un gène rapporteur. Puis nous avons mis au point un système inductible de répression d'un transgène intégré au génome de cellules humaines. Ensuite, nous avons tiré profit de l'étiquette TAP couplée à un système d'expression rétroviral afin d'isoler différents complexes associés aux protéines CBX à partir de cellules humaines en culture. Les lignées cellulaires exprimant de manière stable les différentes protéines de fusion ont été caractérisées et leurs localisations nucléaires ont été confirmées. Des expériences de chromatographie d'exclusion ont permis de décrire que ces protéines participent à la formation de complexes de haut poids moléculaires. Enfin, les purifications TAP ont été réalisées et les identifications par spectrométrie de masse ont pu mettre en évidence que plusieurs PRC1, de composition proche, étaient présents dans des cellules humaines et que certains partenaires protéiques étaient spécifiques à une isoforme CBX particulière, suggérant un rôle biologique bien particulier pour chacune. L'identification de différentes protéines constituant ces complexes nous a amené à valider les interactions à l'aide d'autres méthodes biochimiques comme le GST-pulldown, la co-immunoprécipitation et /ou la co-localisation.

À la suite des matériels et méthodes, je présenterai dans un premier chapitre les résultats obtenus avec les complexes PRC1 purifiés à partir des CBX Polycomb. Dans un second chapitre, je discuterai les résultats préliminaires des purifications TAP à partir des CBX HP1. Pour finir, je dresserai un bilan sur les contaminants omniprésents dans les purifications TAP réalisées à partir de cellules HeLa, dans nos conditions expérimentales.

MATÉRIELS ET MÉTHODES

1 Clonage GATEWAY (recombinaison *in vitro*)

De façon à pouvoir manipuler plusieurs séquences codantes à la fois, nous avons opté pour la technologie GATEWAY (Invitrogen), qui permet une standardisation des clonages. Celle-ci met en jeu des transferts d'ADN entre différents vecteurs grâce à des recombinaisons sites spécifiques. Elle s'inspire du processus naturel d'intégration/excision du phage λ dans le génome d'*E. coli*. Les sites naturels de recombinaisons attB, attL, attR et attP ont été modifiés de façon à augmenter l'efficacité de recombinaison et à permettre un clonage directionnel. Le clonage GATEWAY est simple, rapide et efficace :

1^{ère} étape : clonage de la séquence codante (ORF) en ajoutant à chacune de ses extrémités deux sites de recombinaison attB : attB1 et attB2. Ces deux sites sont légèrement différents l'un de l'autre, ce qui permet d'assurer tout au long du processus un clonage directionnel. Cette opération est réalisée à l'aide d'une réaction d'amplification en chaîne (PCR) utilisant un couple d'amorces dont les extrémités 5' contiennent les séquences attB1 et attB2. Les ORF proviennent d'ADN complémentaires IMAGE commandés au RZPD (Deutsches Ressourcenzentrum für Genomforschung). Les amorces ont été commandées chez Invitrogen. La liste des ORF et des amorces est disponible en Tableau 3.

Proteine	ORF	Gateway 5'	Gateway 3'
CBX2	BC004252.1	<i>ggggacaagttgtacaaaaagcaggctcc</i> ATGGAGGAGCTGAGCAGC	<i>ggggaccactttgtacaagaagctgggt</i> TCAGTAATGCCTCAGGTT
CBX4	BC014967.1	<i>ggggacaagttgtacaaaaagcaggctcc</i> ATGGAGCTGCCAGTGTT	<i>ggggaccactttgtacaagaagctgggt</i> CTACACCGTCACGTA
CBX6	BC012111.1	<i>ggggacaagttgtacaaaaagcaggctcc</i> ATGGAGCTGTCTGCAGTG	<i>ggggaccactttgtacaagaagctgggt</i> TCACTTGTCTGCCCAAT
CBX7	pcDNA3.Gal4.CBX7	<i>ggggacaagttgtacaaaaagcaggctcc</i> ATGGAGCTGTACGCCATC	<i>ggggaccactttgtacaagaagctgggt</i> TCAGAACTTCCCCTGCG
CBX8	BC008937.2	<i>ggggacaagttgtacaaaaagcaggctcc</i> ATGGAGCTTTCAGCGGTG	<i>ggggaccactttgtacaagaagctgggt</i> TCATCTTTTCTTTTAA
CBX1	BC002609.1	<i>ggggacaagttgtacaaaaagcaggctcc</i> ATGGGAAAAAACAAC	<i>ggggaccactttgtacaagaagctgggt</i> TTAGTTCTTGTCTTTT
CBX3	BC000954.1	<i>ggggacaagttgtacaaaaagcaggctcc</i> ATGGCCTCCAACAAACT	<i>ggggaccactttgtacaagaagctgggt</i> TTATTGAGCTTCTATCTT
CBX5	BC006821.1	<i>ggggacaagttgtacaaaaagcaggctcc</i> ATGGAAAGAAAACAAG	<i>ggggaccactttgtacaagaagctgggt</i> TTAGCTCTTGTCTTTT
PHC1	BC073964.1	<i>ggggacaagttgtacaaaaagcaggctcc</i> ATGGAGACTGAGAGCGAG	<i>ggggaccactttgtacaagaagctgggt</i> TTAGGTCTCTTGAGGAC
PHC2	BC130630.1	<i>ggggacaagttgtacaaaaagcaggctcc</i> ATGGAGAATGAGCTGCCAG	<i>ggggaccactttgtacaagaagctgggt</i> CTAGGAGTCCCTGAGCATG
PHC3	BC131772.1	<i>ggggacaagttgtacaaaaagcaggctcc</i> ATGGATACTGAACCAAAC	<i>ggggaccactttgtacaagaagctgggt</i> TTAAGATTCCCTCAGAGAG
RING1	BC051866.2	<i>ggggacaagttgtacaaaaagcaggctcc</i> ATGACGACGCCCGCAATG	<i>ggggaccactttgtacaagaagctgggt</i> TCACTTTGGATCCTTGGTG
RNF2	BC012583.1	<i>ggggacaagttgtacaaaaagcaggctcc</i> ATGTCTCAGGCTGTGCAG	<i>ggggaccactttgtacaagaagctgggt</i> TCATTTGTCTCTTTGT
PCGF1	BC004952.1	<i>ggggacaagttgtacaaaaagcaggctcc</i> ATGAGGCTTCGGAACCAG	<i>ggggaccactttgtacaagaagctgggt</i> CTACCTCCTTCTCTTT
MEL18	BC004858.2	<i>ggggacaagttgtacaaaaagcaggctcc</i> ATGCATCGGACTACCGG	<i>ggggaccactttgtacaagaagctgggt</i> TCAAGTTAAGGGGGCAC
PCGF3	BC107061.2	<i>ggggacaagttgtacaaaaagcaggctcc</i> ATGTTGACCAGGAAGATC	<i>ggggaccactttgtacaagaagctgggt</i> TCACAGCAAGTCCATCTTG
BMI1	BC011652.2	<i>ggggacaagttgtacaaaaagcaggctcc</i> <i>agactgcc</i> ATGCATCGAACAACGAGA	<i>ggggaccactttgtacaagaagctgggt</i> ACCAGAAGAAGTTGCTGA
PCGF5	BC051845.1	<i>ggggacaagttgtacaaaaagcaggctcc</i> ATGGCTACCCAAAGGAAC	<i>ggggaccactttgtacaagaagctgggt</i> CTAACCGAAATCAATCTTG
PCGF6	BC007602.1	<i>ggggacaagttgtacaaaaagcaggctcc</i> ATGGAGGGGGTCCGGGTG	<i>ggggaccactttgtacaagaagctgggt</i> TCAAGTTATCTTCAGAGG
14-3-3 ϵ	BC000179.2	<i>ggggacaagttgtacaaaaagcaggctcc</i> ATGGATGATCGAGAGGATC	<i>ggggaccactttgtacaagaagctgggt</i> TCAGTATTTTCGTCTTCC
14-3-3 ζ	BC003623.2	<i>ggggacaagttgtacaaaaagcaggctcc</i> ATGGATAAAATGAGCTGG	<i>ggggaccactttgtacaagaagctgggt</i> TTAATTTTCCCCTCTTC

Tableau 3 - Tableau des différents oligonucléotides utilisés. En rouge les séquences GATEWAY, en italique la séquence Kozak, en majuscule les 18 premiers ou derniers nucléotides correspondant à la séquence d'intérêt. pcDNA3.Gal4.CBX7 nous a été généreusement fournis par David Bernard (CLB, Lyon).

2^{ème} étape : sous-clonage du produit PCR obtenu dans un vecteur d'entrée universel pDONR201 (Invitrogen) qui contient les sites attP1 et attP2 capables de recombiner avec attB1 et attB2 au cours de la réaction BP (Intégrase + Integration Host Factor). On obtient alors un « **clone d'entrée** » dans lequel l'ORF est maintenant flanquée de deux sites attL : attL1 et attL2. Les clones d'entrée sont résistants à la kanamycine.

3^{ème} étape : incubation du clone d'entrée avec le vecteur de destination. Ce dernier contient les sites attR1 et attR2 qui recombinent spécifiquement avec attL1 et attL2 au cours de la réaction LR (excisionase), permettant l'obtention d'un « **clone d'expression** ». Le produit de cette réaction est transformé dans des bactéries compétentes TOP10 (Invitrogen) pour l'amplification et la purification du clone d'expression. Les clones d'expression sont résistants à l'ampicilline.

2 Amplification PCR

Pour le clonage des EST, les réactions d'amplification en chaîne sont réalisées avec les polymérases KOD HiFi ou KOD XL (Invitrogen) en fonction de la taille des fragments à amplifier, en suivant les instructions du fournisseur. Les cycles d'amplification sont réalisés dans un thermocycleur PTC-200 (MJ Research).

3 Vecteurs rétroviraux

Les vecteurs rétroviraux de la leucémie murine de Moloney (MoMLV) ont été construits à partir du PLNCX2 (Clontech) dans lequel la cassette IRES-pac issue de pIRESpuro (Clontech) a été introduite, afin de lui conférer une résistance à la puromycine.

Alors, des vecteurs compatibles pour une recombinaison GATEWAY ont été générés pour l'expression d'une protéine étiquetée TAP (pRP-NTAP-GW) ou GFP (pRP-NGF-GW) en insérant les cassettes GATEWAY et : TAP (à partir de pZome-1-N, Cellzome AG) ou GFP (à partir de pEGFP-N1, Clontech) respectivement.

4 Vecteurs d'expression

Pour la production de protéines recombinantes étiquetées GST (glutathione-S-transférase), une cassette GATEWAY a été introduite dans le vecteur pGEX-4T1 (Amersham). Pour les étiquettes HA, Flag, GAL4 et GAL-EBD, nous avons utilisé des plasmides qui contiennent ces séquences et les motifs de recombinaison GATEWAY, un promoteur de mammifère, une résistance antibiotique et une origine de répllication bactérienne. Ils ont été généreusement fournis par Cellzome AG.

Pour tester les activités répressives nous avons utilisé les vecteurs plasmidiques suivants :

- rapporteurs :

E1b (Cellzome AG), bax (promoteur entier du gène murin PM bax PF -2673/-1) et p21, généreusement fournis par Didier Monté (IBL, Lille)

- normalisateur :

pRL-TK (Promega) contenant le gène codant la luciférase Renilla sous le contrôle du promoteur minimal de la Thymidine kinase du virus de l'herpès simplex.

5 Culture cellulaire

Nous avons utilisé plusieurs lignées cellulaires :

- HEK293gp (ATCC CRL-1573): production de rétrovirus
- HEK293-CI7 : lignée cellulaire de base HEK293 contenant une copie unique (clone n°7) d'un transgène composé de 4 sites UAS_{GAL4} et d'un promoteur de la thymidine kinase en amont de la séquence codant pour la luciférase Firefly (lignée généreusement fournie par D. Schultz)
- HeLa (Human cervix epithelioid carcinoma – HPA cultures),
- HEK293 (Human Embryonic Kidney, HPA cultures)

Les cellules sont cultivées dans du milieu de culture DMEM (Dulbecco's Modified Eagle Medium) complété par 10% de SVF (Sérum de veau fœtal, Gibco BRL, Life Technologies, Invitrogen), 100 unités/mL de pénicilline (Invitrogen) et 100µg/mL de streptomycine (Invitrogen), à 37°C sous une atmosphère humide à 5% de dioxyde de carbone. Lorsque les cellules sont à confluence, elles sont lavées avec du PBS (Phosphate Buffer Saline, Gibco, Invitrogen) puis décollées du support et dissociées à l'aide d'une solution de 0,05% trypsine-EDTA (Gibco, Invitrogen).

6 Infection rétrovirale pour l'obtention de lignée stable

Les cellules productrices de rétrovirus recombinants (HEK 293gp), dérivées de cellules HEK293, sontensemencées dans une plaque 6 puits à raison de $3 \cdot 10^5$ cellules par puits, puis transfectées 24h plus tard avec 0,8µg de plasmide rétroviral à l'aide de l'Exgen 500 (Euromedex) en suivant les instructions du fournisseur. Après 48h, le milieu de culture contenant les virus est filtré à travers un filtre 0,45µm (Millipore).

Les cellules HeLa ou 293-CI.7 sontensemencées dans une plaque 6 puits à raison de 10^5 cellules par puits. Trois mL de surnageant filtré en présence de 8µg/mL de Polybrene sont utilisés pour réaliser une infection pendant 24h. Les cellules sont ensuite incubées pendant 24h avec du milieu de culture normal. La population polyclonale est ensuite sélectionnée avec 1µg/mL de puromycine. Les cellules résistantes en prolifération sont testées pour l'expression de la protéine étiquetée par western blot et/ou immunofluorescence et/ou activité sur le gène rapporteur luciférase.

7 Transfection des cellules

Les transfections sont réalisées en plaque de 6 (co-immunoprécipitations, infection virale), 12 ou 24 puits (activité luciférase). Les cellules sontensemencées à J-1 pour obtenir une confluence de 60% le jour J (1,2 millions de cellules / plaque). Les cellules sont rincées une fois avec de l'OPTI-MEM puis mises en culture dans 3 mL d'OPTI-MEM par puits d'une plaque 6 puits. L'agent de transfection utilisé est un polymère d'éthylènimine (PEI, Exgen 500, Euromedex). Le mélange de transfection contient : 500ng d'ADN final et 2 pmoles de PEI pour 1 puits d'une plaque 6 puits. L'ADN et le PEI sont dilués dans 200µL d'eau physiologique et incubés pendant 30 minutes à température ambiante avant d'être déposés sur les cellules. Après 6h, le milieu est aspiré et remplacé par du milieu complet normal.

Quarante huit heures après la transfection, les surnageants sont collectés (infection virale) ou les cellules sont lavées une fois au PBS puis lysées dans 300µL de tampon de lyse cellulaire (Tris 250

mM pH8, glycérol 25% (v/v), MgCl₂ 7,5 mM, NaCl 500 mM, NaF 125 mM et Na₃VO₄ 5 mM, NP40 0,2 % (v/v), cocktail d'inhibiteurs de protéases (cOmplete, EDTA-free, Roche)).

8 Mesure de l'activité luciférase

- Transfection transitoire : luciférase firefly (rapporteur) / luciférase renilla (normalisateur)

Vingt µL d'extrait cellulaire par échantillon sont déposés dans une plaque 96 puits. La plaque est déposée dans le luminomètre Centro LB 960 (Berthold). La luciférine est injectée à raison de 30µL de luciférine par puits. En présence de luciférase Firefly, cette dernière est dégradée et émet de la lumière. Suite à l'injection, la lumière est mesurée pendant 5 secondes. La réaction est stoppée par l'injection de 30µL du second réactif « stop n glo », qui permet le dosage de l'activité luciférase Renilla. La lumière émise est mesurée pendant 5 secondes. Ainsi un rapport entre les 2 lumières mesurées permet d'obtenir une valeur normalisée.

- Lignées stables

Les cellules sont traitées avec 1/1000^{ème} d'éthanol ou 1µM d'hydroxytamixofen (4-OHT) pendant la durée déterminée, puis les cellules sont lavées au PBS et lysées avec le tampon de lyse cellulaire.

La procédure est similaire à la précédente, mais seulement la première injection est réalisée, la normalisation est effectuée par la mesure de la concentration en protéines totales des échantillons par Bradford (BioRad Protein Assay).

9 Western blot

- Préparation des échantillons

Les cellules en culture subconfluentes sont lavées une fois avec du PBS, puis lysées pendant 15 minutes à 4°C dans un tampon de lyse cellulaire. L'extrait cellulaire est centrifugé à 16000 g pendant 10 minutes à 4°C. Le surnageant est conservé et la concentration en protéines est mesurée à l'aide du réactif Bradford.

- Migration

Vingt à 40µg de protéines sont dénaturés dans une solution de Laemmli (pour du 5x : 5% SDS, 1% β-mercaptoéthanol, 0,25M Tris HCl pH6,5, 5% glycérol, 0,05% bleu de bromophénol) et ébouillantés à 99°C pendant 5 minutes avant d'être déposés dans un puits d'un mini- ou midi-gel NuPAGE Novex 4-12% Bis-Tris (Invitrogen). La migration électrophorétique à 200V est réalisée dans un tampon de migration dénaturant NuPAGE MOPS SDS Running Buffer (Invitrogen) agrémenté d'un antioxydant (Invitrogen).

- Transfert

Après la migration, un transfert semi-sec du gel d'acrylamide est effectué sur une membrane de nitrocellulose (Nitrocellulose transfert membrane, Whatman) à l'aide d'un tampon de transfert Towbin (25mM Tris, 192mM Glycine, 20% v/v méthanol, pH 8.3, 0,1% SDS) pendant 2h à 100mA constants.

L'efficacité du transfert est vérifiée par coloration de la membrane au Rouge Ponceau (0.1% de ponceau S dans 5% d'acide acétique, Sigma Aldrich).

- Révélation immunologique

La membrane est incubée pendant une heure dans une solution de saturation (100mg/mL lait, 1% Tween 20 dans PBS) sous agitation. L'anticorps primaire utilisé à la dilution adéquate dans la solution de saturation est incubé avec la membrane pendant 1h à température ambiante. Après 3 lavages de 5 minutes au PBS, la membrane est incubée avec l'anticorps secondaire couplé à la peroxydase, dilué au 1/10000^{ème}. Après 3 lavages de 5 minutes dans du PBS, les protéines immunodétectées sont révélées par chimioluminescence (GE Healthcare).

10 Immunodétection sur cellules fixées

Les lignées cellulaires qui expriment de manière stable la protéine d'intérêt fusionnée à la GFP ou à l'étiquette TAP sontensemencées sur des lamelles rondes qui ont été préalablement traitées avec de la poly-D-lysine. Quand les cellules atteignent une confluence de 70-80%, elles sont rincées une fois avec du PBS puis fixées avec 4% de paraformaldéhyde pendant 20 minutes. Elles sont ensuite rincées 2 fois au PBS puis perméabilisées dans 0,1% Triton X-100 pendant 5 minutes. Après 2 lavages au PBS, les cellules sont saturées avec 1% de sérum de veau fœtal dans du PBS pendant 1h. Ensuite les incubations avec les anticorps primaires et secondaires (Alexa Fluor 594) adéquats sont réalisées. Enfin, une coloration de l'ADN des noyaux est effectuée à l'aide du colorant Hoechst 33342 (Sigma) 10µg/mL dans du PBS pendant 1 minute. Après 5 lavages au PBS, la lamelle est montée sur lame dans du MOWIOL (Sigma).

TAP : Ire = polyclonal anti-protéine A de lapin (sigma) à 0,5µg/mL, Ilre = chèvre anti-lapin-Alexa594 (Invitrogen) au 1/1000^{ème}.

GFP : Ire = monoclonal anti-GFP de souris (Clontech JL8) à 0,5µg/mL, Ilre = chèvre anti-souris-Alexa594 (Invitrogen) au 1/1000^{ème}.

11 Immunoprécipitation

Quarante huit heures après la transfection transitoire, les cellules en plaques 6 puits sont lavées avec du PBS froid puis lysées dans 300µl de tampon de lyse cellulaire. Les lysats sont centrifugés à 16000 g, 4°C pendant 30 minutes. Vingt µL de billes anti- HA ou anti-Flag (sigma) préalablement rincées avec du tampon de lyse cellulaire sont déposées dans 580µL de surnageant (20µL sont aliquotés en tant qu'input). Les échantillons sont incubés sur une roue à 4°C sur la nuit. Après 4 lavages avec le tampon de lyse cellulaire, les protéines immunoprécipitées sont éluées à l'aide de 40µl de la solution de Laemmli.

12 GST-pulldown

Les protéines fusionnées à la GST sont exprimées dans des bactéries *E. coli* BL21 (DE3) et purifiées avec des billes de sépharose 4B couplée à la glutathione (GE Healthcare) en suivant les instructions du fournisseur. Le tampon des protéines purifiées est ensuite changé par 100 mM Tris-HCl pH8 et 10

% sucrose à l'aide de colonnes Amicon (Millipore) et d'une centrifugation. Les protéines purifiées sont à nouveau immobilisées sur les billes de séphaore 4B-glutathione et maintenues dans un tampon de stockage (10 mM Tris-HCl pH8, 150 mM NaCl, 1 mM EDTA et inhibiteurs de protéases [Roche]). Environ 2mg de protéines issues d'une extraction de protéines nucléaires à partir de cellules adéquates en culture sont incubées avec les billes+GST sur une roue à 4°C pendant une nuit. Les billes sont ensuite lavées 4 fois avec le tampon E1A (50 mM Hepes, pH 7,9, 1 mM EGTA, 250 mM NaCl, 1 mM DTT, 1 mM EDTA et des inhibiteurs de protéases). Les protéines retenues sont éluées avec le tampon d'éluion (10 mM glutathione in 50 mM Tris-HCl pH 8.0). Elles sont ensuite analysées par SDS-PAGE puis par western blot.

Les extraits nucléaires pour le GST-pulldown sont préparés à partir de cellules HEK293 transfectées de manière transitoire avec un plasmide exprimant une protéine TAP. Quarante huit heures après la transfection, les cellules sont lysées dans un tampon pauvre en sels (10 mM Tris-HCl pH 7.4, 25 mM NaCl, 2 mM MgOAc, 1 mM DTT, 1mM EDTA, 0.05 % NP-40 et des inhibiteurs de protéases [Roche]) pendant 15 minutes à 4°C. Les noyaux sont culotés par centrifugation à 800g pendant 10 minutes à 4°C, puis resuspendus dans le tampon E1A pendant 2h à 4°C sur une roue. Les protéines nucléaires sont récupérées par centrifugation à 16000 g pendant 10 minutes à 4°C.

13 Chromatographie d'exclusion

Deux mg de protéines nucléaires fraîchement préparées sont fractionnés dans une colonne Superose 6 HR 10/300 GL (GE Healthcare) préalablement équilibrée avec la phase mobile (50 mM Tris-HCl pH 7,4, 100 mM KCl). La chromatographie d'exclusion par FPLP (fast protein liquid chromatography) est réalisée à l'aide d'un ÄKTA purifier (GE Healthcare). Les expériences sont réalisées à température ambiante à un débit de 0,5 mL/min. En sortie de colonne, les éluions sont fractionnées par aliquots de 500µL (une trentaine de fractions collectées par expérience). Vingt cinq µL de chaque éluion sont utilisés pour une étude par western blot.

Pour l'obtention de la courbe de calibration, les protéines témoins (sigma-aldrich) suivantes ont été analysées : bleu dextran (volume vide), thyroglobuline (669 kDa), apoferritine (443 kDa), beta-amylase (200 kDa), alcool deshydrogénase (150 kDa), albumine de sérum bovin (66 kDa) et anhydrase carbonique (29 kDa).

14 Purification TAP

Pour réaliser la purification TAP, les cellules exprimant la protéine fusionnée à l'étiquette TAP sont mises en culture dans 40 boîtes de 15 cm de diamètre. A la confluence, les cellules sont collectées avec un grattoir, lavées dans du PBS froid, resuspendues dans du tampon de lyse (50 mM Tris-HCl pH 7,5, 125 mM NaCl, 5% glycérol, 0,2% NP-40, 1.5 mM MgCl₂, 25mM NaF, 1 mM Na₃VO₄, 1 mM dithiothréitol [DTT] et inhibiteurs de protéases [Roche]) et homogénéisées mécaniquement par 20 aller-retours dans un pilon de type-B. Les extraits sont ensuite incubés pendant une heure sur une roue à 4°C pour extraire les protéines nucléaires. Les lysats sont ensuite clarifiés par ultracentrifugation à 100000g.

Les surnageants sont incubés avec des billes couplées à des Immunoglobulines G (Sigma) sur une roue pendant 2h à 4°C. Les protéines immobilisées sont lavées avec un excès de tampon de lyse puis avec le tampon de digestion par la TEV (10 mM Tris-HCl pH 7,5, 100 mM NaCl and 0.2% NP-40) avant d'être éluées par addition de 30mg de protéase TEV (Invitrogen) pendant 2h à 16°C. L'éluat est ensuite incubé avec des billes de sépharose couplées à la calmoduline (Amersham) en présence de 2mM de CaCl₂ pendant 30 minutes à 4°C sous rotation. Après plusieurs lavages dans le tampon de lavage (100 mM Tris-HCl pH8, 100 mM NaCl, 0,5 mM EDTA, 2 mM CaCl₂), les protéines fixées à la calmoduline sont éluées par ébullition dans une solution de Laemmli.

15 Analyse de spectrométrie de masse

Les protéines éluées après purification TAP sont séparées dans un gel dénaturant NuPAGE Novex 4-12% (Invitrogen) puis sont colorées avec du Bleu de coomassie colloïdal (Imperial Protein Stain, Pierce). La piste est découpée à l'aide d'un scalpel sur toute sa hauteur en bandes d'environ 2mm de hauteur. Les bandes obtenues sont décolorées dans une solution contenant 50 mM NH₄HCO₃ et 50% acétonitrile. Les protéines sont ensuite réduites dans du DTT puis alkylées avec de l'iodoacétamide. Les bandes sont déshydratées avec de l'acétonitrile et rehydratées dans le tampon de lyse (25 mM NH₄HCO₃) contenant 12,5ng/μL de trypsine pendant 30 minutes à 4°C, l'excès de trypsine est retiré et les bandes sont recouvertes par le tampon de lyse sans trypsine. Les échantillons sont ensuite incubés sur la nuit à 37°C. Les peptides tryptiques sont élués dans 3 tampons successifs, chaque élution est réalisée sous agitation pendant 10 minutes à température ambiante : deux bains dans une solution 45% acétonitrile, 10% acide formique puis une dernière élution dans 95% ACN, 5% acide formique. Toutes les éluions sont poolées puis évaporées dans un speed-vac. Les échantillons, repris dans 3μL d'acide formique 0,1%, sont analysés par chromatographie liquide en phase inverse (C18 Pepmap, Dionex) couplée par un électrospray à un spectromètre de masse en tandem (LCQ Deca XP⁺, Thermo Electron Corp). L'acquisition permet de réaliser successivement un full scan dans l'échelle 500-2000 m/z, et un scan MS/MS de l'ion le plus intense. Les données brutes sont transformées en fichiers .dta puis soumises à la base de données publique Mascot (www.matrixscience.com) de manière à être confrontées aux séquences de la banque Swiss-Prot. Les paramètres pour l'interrogation de la base de données sont les suivants : 3 clivages oubliés, oxydation des méthionines et carbamidométhylation des cystéines en modifications variables, 2 Da de tolérance sur le peptide et 0,8 Da de tolérance pour la MS2. Les résultats sont annotés en fonction du score de probabilité Mowse (-10 x log(p), avec p la probabilité que la séquence identifiée corresponde à un événement aléatoire). La plupart des protéines ont été identifiées avec plusieurs peptides indépendants. Les protéines dont le score Mascot est inférieur à 38 ou le seuil de significativité est supérieur à 5% ont été rejetées.

RESULTATS ET DISCUSSION

Partie 1 : Les protéines CBX Polycomb (CBX2, 4, 6, 7 et 8)

1 Etude de l'activité répressive des protéines CBX Polycomb

1.1 *Système rapporteur exogène et sur-expression d'une protéine fusionnée à GAL4.*

Nous avons entrepris de caractériser l'activité répressive de tous les membres de la famille CBX Polycomb sur un plasmide exogène codant pour la luciférase.

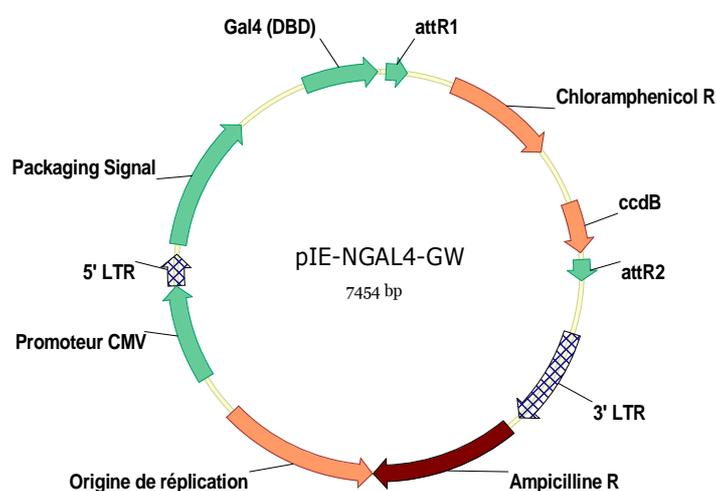


Figure 12 - Carte du vecteur de destination pIE-NGAL4-GW utilisé pour obtenir un effecteur fusionné à GAL4 (DBD)

Pour ce faire, nous avons construit un plasmide de destination (Figure 11) contenant :

- une cassette GATEWAY contenant les sites attR1 et attR2 afin de pouvoir cloner rapidement les séquences codantes des CBX insérées dans le plasmide donneur (pDONR201)
- un domaine GAL4 (DBD) : domaine de liaison à l'ADN qui code pour les acides aminés 1 à 147 de la protéine GAL4 issue de la levure *Saccharomyces cerevisiae*. Ce domaine est capable de reconnaître et de se fixer spécifiquement sur les séquences UAS_{GAL} (Upstream Activation Sequence) présentes dans l'ADN.
- un promoteur fort CMV (cytomégalovirus) pour obtenir de hauts niveaux d'expression de la protéine de fusion dans les cellules humaines
- une résistance à l'ampicilline et une origine de répliation bactérienne (afin de pouvoir sélectionner les clones contenant le plasmide et de l'amplifier)

Les 5 séquences codantes de CBX ont été clonées dans le plasmide de destination pIE-N-Gal4-GW par une réaction LR. Nous avons obtenus 5 clones d'expression qui codent pour des protéines de fusion GAL4-CBX, avec GAL4 situé à l'extrémité amino-terminale des protéines d'intérêt.

Nous disposons de 3 plasmides rapporteurs : G5-E1b-Fluc, G5-bax-Fluc et G5-p21-Fluc contenant tous 5 sites UAS_{GAL}, en amont d'un promoteur et de la séquence codante du gène de la luciférase de luciole (Firefly) : Fluc. Les promoteurs des 3 plasmides sont respectivement :

E1b (130nt) : le promoteur du gène codant pour le polypeptide IX de l'adénovirus

bax (337nt) : le promoteur du gène codant pour la protéine pro-apoptotique Bcl-2-associated X

et p21 (2278nt) : le promoteur du gène CDKN1A codant pour la protéine Cyclin-dependent kinase inhibiteur 1A.

Bax et p21 sont parmi les cibles favorites du facteur de transcription suppresseur de tumeur p53.

Ces 3 séquences correspondent à des promoteurs « forts » qui permettent d'obtenir une activité transcriptionnelle élevée de Fluc quand ils sont transfectés dans des cellules de mammifères.

Un seul et unique plasmide normalisateur a été utilisé dans cette première série d'expériences : ubi-Rluc exprimant la luciférase de méduse (Renilla) sous le contrôle du promoteur fort de l'ubiquitine C humaine.

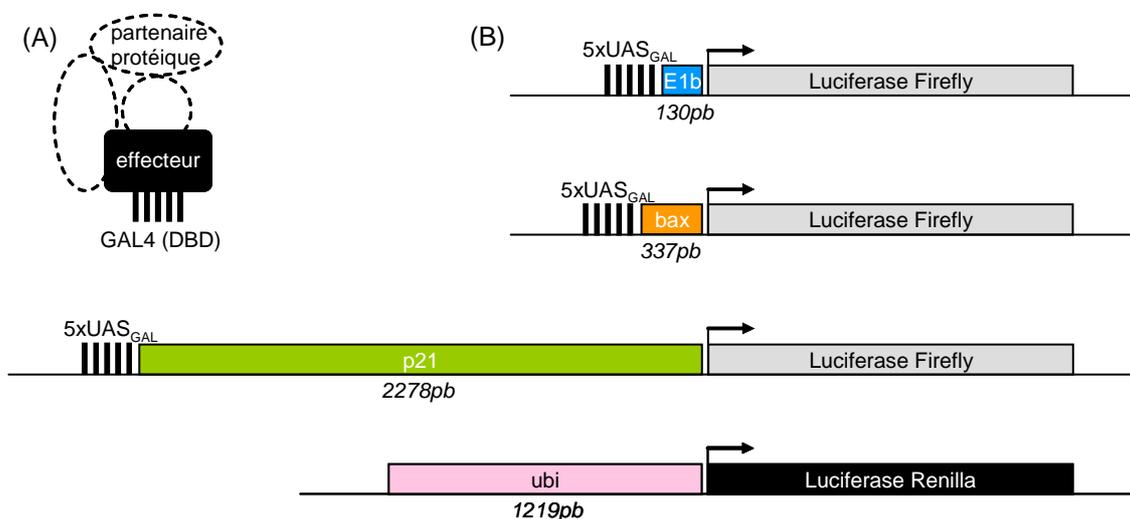


Figure 13 - Système rapporteur luciférase exogène

A - Représentation schématique de la protéine de fusion entre l'effecteur et le domaine de liaison à l'ADN GAL4 (DBD). **B** - Représentation graphique des 3 plasmides rapporteurs et du plasmide normalisateur utilisés

La protéine de fusion exprimée à partir du vecteur d'expression est susceptible d'interagir avec ses partenaires protéiques naturels qui sont présents de manière endogène dans la cellule. Le domaine GAL(DBD) permet l'adressage de la protéine (et de ses partenaires) en amont du promoteur de la luciférase sur le plasmide rapporteur (Figure 12). L'activité répressive ou activatrice de l'effecteur (ou du complexe) sur la transcription est évaluée en mesurant la quantité de luciférase présente dans les cellules transfectées.

Nous avons réalisé des expériences de co-transfection dans des cellules HEK293, pour chaque valeur, nous avons réalisé au moins 3 points, en triplicat. Un clone d'expression contenant la séquence de la GFP fusionnée à GAL4 nous a servi de contrôle visuel de transfection et de référence d'expression de luciférase normalisée.

Des études préliminaires nous ont permis d'estimer les quantités de plasmide à utiliser, pour 1 puits d'une plaque 6 puits : plasmide normalisateur : 5ng, plasmide rapporteur : 100ng pour bax et p21, 500ng pour E1B-F-luc et vecteur d'expression : 100ng.

Les valeurs obtenues représentent le rapport des intensités de lumière mesurées entre la luciférase Firefly et la luciférase Renilla. Les résultats obtenus avec la GFP, ajustés à 100%, ont servi de référence.

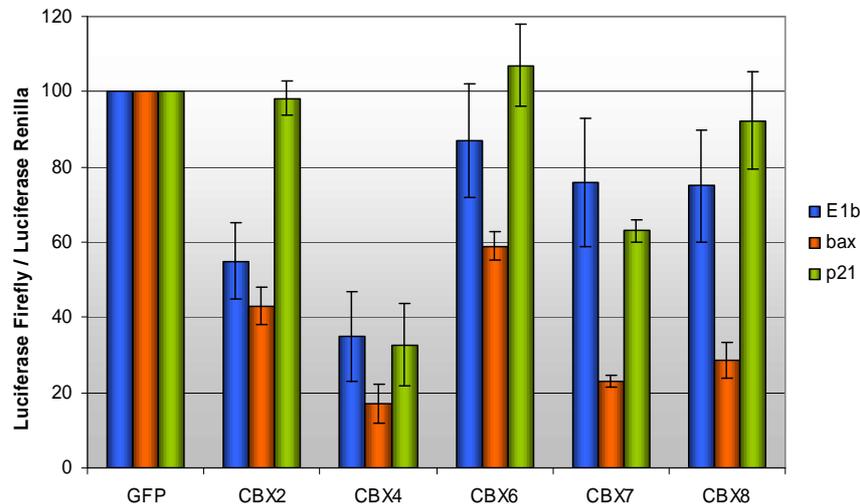


Figure 14 - Activité répressive des protéines CBX Polycomb sur 3 promoteurs différents. L'intensité des luciférase a été mesurée 48h après transfection dans des cellules HEK293

D'une manière générale, nous remarquons que CBX4 est la protéine CBX Polycomb qui réprime le plus dans nos conditions expérimentales. A l'inverse, CBX6 est la protéine ayant le moins d'effet sur la transcription de la luciférase. Nous pouvons résumer les activités constatées de la manière suivante :

Sur E1b : CBX4 > CBX2 > CBX7 = CBX8 > CBX6

Sur bax : CBX4 > CBX7 > CBX8 > CBX2 > CBX6

Sur p21 : CBX4 > CBX7 > CBX8 > CBX2 = CBX6

A l'exception de CBX2 sur E1b, les activités répressives générales observées sont conservées entre les différents promoteurs testés. Le promoteur p21 apparaît comme un promoteur puissant difficile à réprimer contrairement au promoteur bax.

Compte tenu de ces dernières observations, nous avons décidé de tester l'activité répressive d'autres protéines associées au PRC1, que nous avons cloné dans le vecteur d'expression pIE-NGAL4-GW, sur les promoteurs bax et p21.

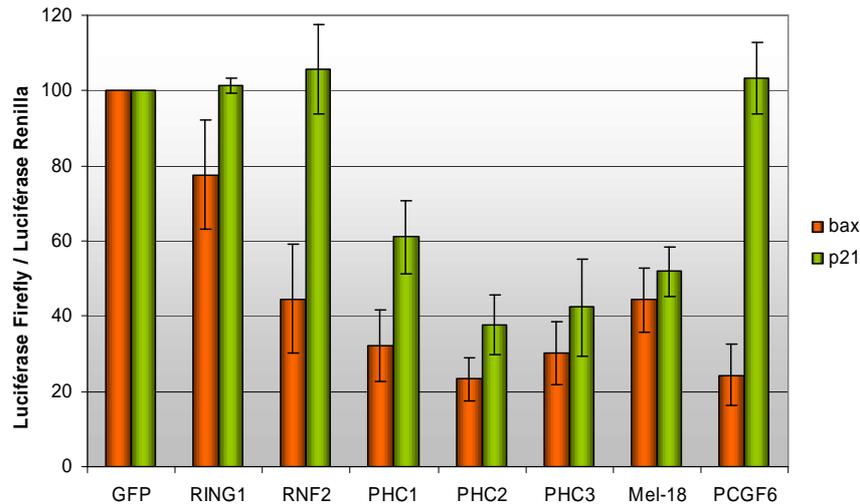


Figure 15 - Activité répressive de protéines décrites comme appartenant au complexe PRC1 sur deux promoteurs différents. L'intensité des luciférases a été mesurée 48h après transfection dans des cellules HEK293

Nous constatons que les protéines RING1 et RNF2 n'ont pas d'effet sur p21 tandis qu'elles répriment respectivement de 20 et 55% la transcription de la luciférase sous le contrôle du promoteur bax. En ce qui concerne les protéines polyhoméotiques (PHC1, PHC2 et PHC3), elles semblent se comporter d'une manière similaire sur les 2 promoteurs testés, réprimant à plus de 70% l'activité de bax et de 40 à 60% l'activité de p21. Enfin, les 2 protéines de la famille PCGF que nous avons étudiées se comportent différemment. D'un côté Mel-18 est capable de réprimer de moitié les deux promoteurs bax et p21, d'un autre PCGF6 ne réprime que bax, mais à presque 80%.

Discussion

Dans cette première série d'expériences, nous avons évalué indirectement l'activité répressive de protéines effectrices fusionnées à GAL4 sur 3 plasmides rapporteurs contenant des promoteurs qui leur étaient spécifiques. Nos résultats confirment que les protéines CBX sont des répresseurs transcriptionnels. Même si en quantité absolue, le promoteur bax permet d'obtenir de grandes quantités de luciférase dans les cellules par rapport aux deux autres plasmides testés, c'est celui qui est le plus facilement réprimé par les protéines CBX. Compte tenu de la structure de ces promoteurs, nous remarquons que les promoteurs E1b et bax sont relativement courts, 130 et 337 paires de bases respectivement, par rapport au promoteur p21 qui ne compte pas moins de 1219 paires de bases. Ainsi, même si le plasmide est circulaire, le site d'ancrage des protéines de fusion est physiquement plus éloigné du site d'initiation de la transcription pour p21 par rapport aux 2 autres promoteurs. Cependant il est admis qu'au sein du génome, les protéines Polycomb sont localisées dans de gros domaines (plusieurs kilobases), couvrant parfois plusieurs gènes comme les gènes homéotiques. Compte tenu que dans nos conditions, p21 est le promoteur le plus « difficile » à réprimer, et que bax et p21 sont tous les deux activés par le même facteur de transcription p53, ce n'est pas la mécanistique d'activation qui change d'un promoteur à l'autre donc le site de fixation de l'effecteur joue un rôle sur son activité.

Les CBX Polycomb ne répriment pas de la même manière un promoteur commun, cela suggère qu'elles possèdent des activités répressives distinctes, suggérant qu'elles recrutent des complexes de composition différentes pour réaliser la répression. Nous avons pu vérifier que l'activité répressive était causée par leurs protéines partenaires en générant une forme tronquée de CBX7 (1-219) dans laquelle la PC box est manquante (Figure 4). Ce mutant CBX7 ne possède plus aucune activité répressive, comme la Pc box est indispensable pour l'interaction avec RING (Satijn DP, 1997, Mol Cell Biol / Bardos JI, 2000, JBC), nous pouvons penser que la Pc box et les protéines qu'elle permet de recruter sont importantes pour la répression médiée par les CBX. Mais comme nous l'avons vu, les protéines RING à elles seules, ne sont pas en mesure de réprimer le promoteur p21, d'autres protéines sont probablement requises pour avoir un complexe PRC1 fonctionnel.

Au vu des résultats de la Figure 14, les protéines PHC, décrites comme pouvant jouer un rôle dans la diffusion des PRC1, sont en mesure de réprimer fortement et de manière assez similaire les 2 promoteurs testés, ce qui suggère qu'elles pourraient faire partie ou recruter un complexe identique. Enfin, les 2 membres de la famille PCGF qui ont été testés révèlent des comportements divergents, supposant qu'ils ne participent pas au même complexe dans la cellule.

1.2 *Système rapporteur endogène et expression inductible d'une protéine fusionnée à GAL4.*

1.2.1 *Système rapporteur chromatinien, inductible par un analogue de l'œstrogène*

Compte tenu du caractère « artificiel » des tests préliminaires, nous avons souhaité mettre au point un système qui allait nous permettre de suivre l'activité des CBX sur le promoteur d'un gène rapporteur intégré de manière stable dans le génome. Pour ce faire, nous avons utilisé la lignée cellulaire HEK293-CI.7 fournie par David C. Schultz (The Wistar Institute, Philadelphie).

Ce dernier a mis au point une lignée cellulaire dont le génome contient une copie unique d'un transgène rapporteur 5xGAL4-TK-luciférase, on parle de rapporteur « chromatinien ». Le transgène contient la séquence codante pour la luciférase Firefly sous le contrôle du promoteur minimal de la thymidine kinase du virus de l'herpès simplex. En amont de ce promoteur, 5 sites UAS_{GAL} sont également présents (Sripathy SP, 2006, Mol Cell Biol). A l'image des expériences transitoires, l'expression d'un effecteur contenant une étiquette GAL4 (DBD) permet son adressage auprès du promoteur TK et la mesure de l'activité luciférase permet d'évaluer l'activité de l'effecteur sur la transcription.

Dans le but de pouvoir suivre la répression du gène rapporteur chromatinien par les protéines CBX, nous avons généré des lignées cellulaires dont l'activité des effecteurs serait inductible. Nous avons opté pour un système dépendant de la présence d'une hormone. Nous avons tiré profit du domaine de liaison aux hormones du récepteur aux œstrogènes (DLO). Une région spécifique du récepteur aux œstrogènes, dont la glycine 521 a été remplacée par une sérine, n'est quasiment plus sensible aux œstrogènes naturelles présentes dans le sérum, alors qu'elle l'est 1000 fois plus à l'anti-hormone 4-hydroxitamoxifen (4-OHT) (Littlewood TD, 1995, NAR). La génération d'une protéine de fusion avec le

domaine DLO (G521R) permet de contrôler post-traductionnellement l'activité de l'effecteur en ajoutant ou non du 4-OHT dans le milieu de culture.

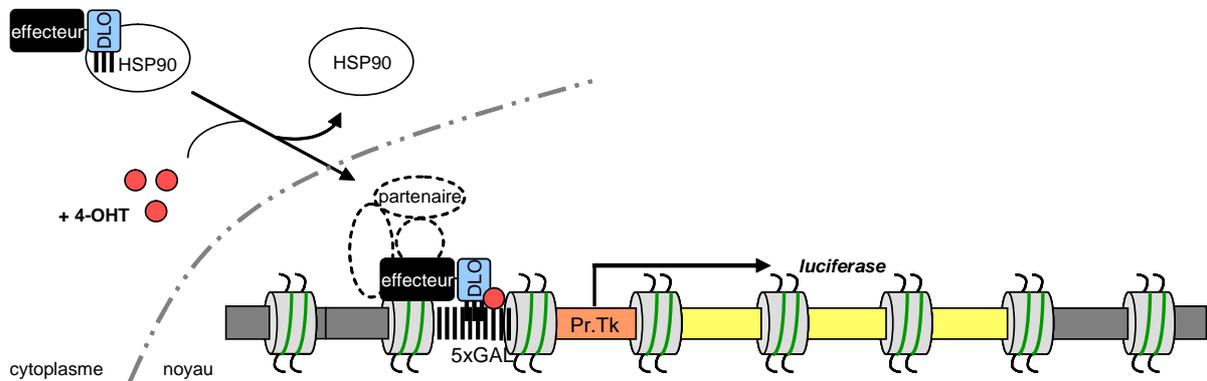


Figure 16 - Représentation schématique du système rapporteur luciférase « chromatinien » dont la répression ou l'activation transcriptionnelle est inducible par une hormone (4-OHT)

En effet, la protéine contenant le DLO est séquestrée dans le cytoplasme des cellules, car en absence de ligand (4-OHT), le DLO interagit avec le complexe associé aux protéines HSP90, empêchant son entrée dans le noyau. Par ajout de 4-OHT dans le milieu de culture, le complexe HSP90 est remplacé par l'anti-hormone et permet à l'effecteur de pénétrer à l'intérieur du noyau (Figure 15).

Nous avons donc construit un plasmide de destination pour générer des lignées stables inducibles par infection virale à partir de la lignée HEK293-CI.7.

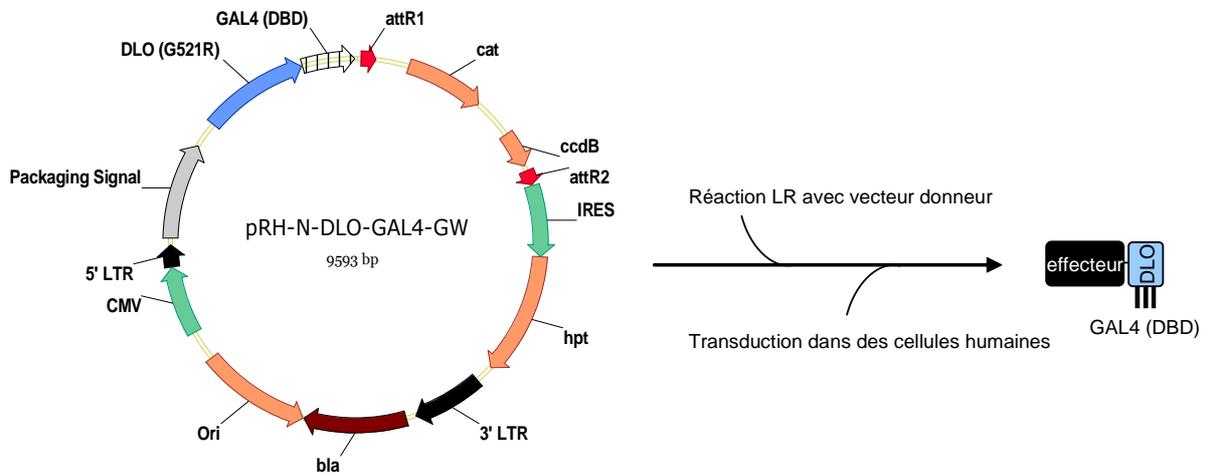


Figure 17 - Carte du vecteur de destination pRH-N-DLO-GAL4-GW utilisé pour obtenir un effecteur fusionné à un domaine de liaison à l'œstrogène et un site GAL4 de fixation à l'UAS_{GAL} à son extrémité amino-terminale.

Le plasmide pRH-N-DLO-GAL4-GW contient les séquences indispensables :

- pour la fabrication du virus : le signal d'emballage et les séquences LTR (Long Terminal repeat) 5' et 3' et un promoteur CMV
- pour l'amplification du plasmide : une origine de répllication bactérienne (Ori) et deux résistances à l'ampicilline (*bla* = gène codant pour la β -lactamase) et au chloramphénicol (*cat*: gène codant pour la chloramphénicol acétyl-transférase)

- pour le clonage de l'effecteur : une cassette GW avec les séquences de recombinaison attR1 et attR2.
- pour sélectionner les cellules qui ont intégré le transgène : un IRES (internal ribosome entry site) qui permet l'initiation de la traduction en amont d'un gène conférant une résistance à l'hygromycine (*hpt* = gène codant pour l'hygromycine phospho-transférase). Notons que la lignée HEK293-CI.7 est déjà résistante à la puromycine.
- pour obtenir l'étiquette d'intérêt en N-terminal : les séquences codantes pour DLO (G521R) et le domaine de liaison à l'ADN GAL4.

Nous avons cloné les différentes protéines CBX Polycomb dans le plasmide pRH-N-DLO-GAL4-GW par une réaction LR, puis nous avons réalisé une infection virale dans la lignée HEK293-CI.7. La concentration optimale en 4-OHT (dilué dans l'éthanol) pour l'induction a été évaluée à 500nM, parallèlement, des cellules de la même lignée sont traitées avec de l'éthanol, les quantités de luciférase mesurées sont normalisées par la quantité totale de protéines dans l'extrait.

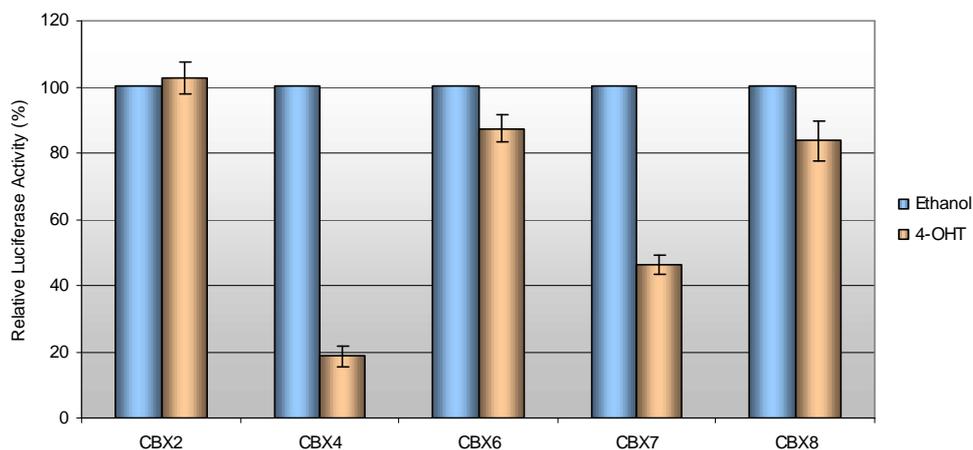


Figure 18- Activité répressive des protéines CBX Polycomb sur le rapporteur chromatinien. L'intensité de luciférase a été mesurée 48h après ajout de 4-OHT. Les valeurs ont été normalisées par la quantité de protéines totales

Les protéines CBX fusionnées à DLO-GAL4 possèdent des activités répressives différentes sur le gène rapporteur chromatinien. Dans ces conditions, les modifications post-traductionnelles des histones et le contexte génomique devaient nous permettre de suivre l'activité répressive de tous les membres de cette famille. Nous constatons que CBX2 ne possède aucune activité répressive dans notre essai, que CBX6 et CBX8 ont des activités proches, mais ne sont capables de réprimer la transcription de la luciférase que de 15-20% environ. Enfin, CBX4 et CBX7 apparaissent comme des répresseurs forts de la transcription, en diminuant respectivement de 80 et 50% la quantité de luciférase mesurée après traitement au 4-OHT pendant 48 heures par rapport au contrôle. Le mutant CBX7 qui ne possède plus la Pc box perd également sa capacité répressive dans ce test. Nous pouvons sommairement résumer nos résultats de la manière suivante :

CBX4 > CBX7 > CBX8 = CBX6 > CBX2

Discussion

Nous avons réussi à mettre au point des lignées stables inductibles pour suivre l'activité des protéines CBX sur un gène rapporteur chromatinien. Nous avons travaillé sur des populations hétérogènes suite à l'infection virale et une sélection par un traitement à l'hygromycine. Parallèlement, nous avons sélectionné une dizaine de clone pour chaque lignée, et les valeurs de luciférase mesurées étaient en moyenne similaires à celles de la population hétérogène.

Notons que ce modèle de répression inductible sur un gène chromatinien a été adopté par d'autres équipes de recherche et qu'il a permis, entre autres, d'établir un modèle pour la transmission de la marque H3K27me3 au cours des divisions cellulaires (Hansen KH, 2008, Nat Cell Biol). Nous avons tenté de réaliser le même type d'expérience en retirant du milieu le 4-OHT après 48h de répression, cependant, la répression diminuait en même temps que les cellules se divisaient, signifiant que la répression induite par les PRC1 est perdue au cours de la division des cellules HeLa.

Les résultats obtenus dans cette deuxième série d'expériences confirment les résultats précédents. Les « forces » de répression sont corrélées entre les deux systèmes utilisés. Nous pouvons conclure que les protéines CBX possèdent des activités répressives différentes dans le modèle choisi, ce qui suggère qu'elles appartiennent à des complexes distincts. De plus, compte tenu de la perte d'activité avec le mutant CBX7 [1-219], nous pensons que ce sont les protéines qui interagissent avec les CBX qui sont responsables de la répression. C'est pourquoi nous avons envisagé d'étudier les complexes qui sont associés aux CBX, afin de mieux comprendre pourquoi elles avaient des activités si différentes quand elles étaient adressées en amont d'un gène activement transcrit, dans un contexte génomique.

1.2.2 Système rapporteur chromatinien, inductible par un analogue de la progestérone

Afin de pouvoir étudier le recrutement séquentiel des différentes protéines associées au PRC1, nous avons envisagé de mettre au point un second système inductible basé sur les mêmes propriétés que le précédent. Il existe une version tronquée du récepteur de la progestérone (42 acides aminés manquant en région C-terminale), qui lui confère une augmentation de son affinité pour le ligand RU486 (ou mifépristone, un analogue chimique de la progestérone) et une forte diminution de la sensibilité à l'hormone endogène. Nous avons cloné cette version tronquée du DLP (domaine de liaison à la progestérone) dans un vecteur rétroviral similaire au plasmide pRH-N-DLO-GAL4-GW, mais portant un gène de résistance à la néomycine à la place de celui codant pour l'hygromycine. La carte du plasmide pRN-N-DLP-GAL4-GW est présentée en Figure 18.

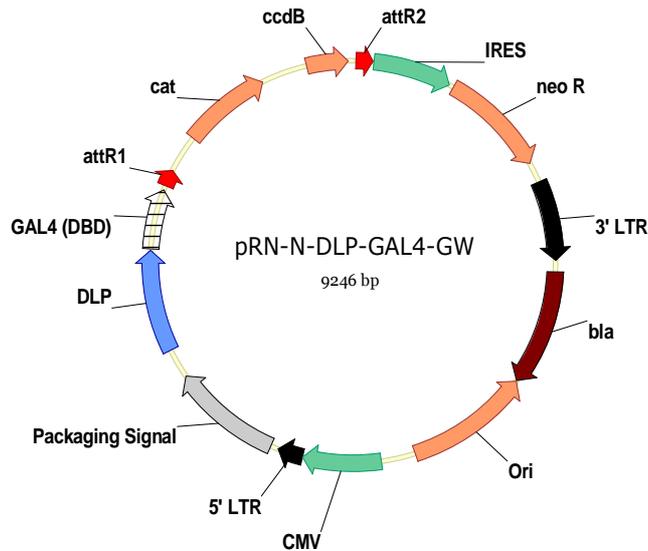


Figure 19 - Carte du vecteur de destination pRN-N-DLP-GAL4-GW utilisé pour obtenir un effecteur fusionné à un domaine de liaison à la progestérone et un site GAL4 de fixation à l'UAS_{GAL} à son extrémité amino-terminale.

Malheureusement, même si l'expression des protéines de fusion a pu être détectée (contrôle avec une protéine de fusion fluorescente rouge), l'induction par le RU486 n'a jamais été fonctionnelle, tant sur des activateurs (domaine transactivateur de la protéine virale VP16) que sur des répresseurs de la transcription (HDAC1). Différentes doses et durées de traitement en RU486 ont été testées mais aucune n'a permis d'obtenir des résultats satisfaisants.

2 Génération de lignées cellulaires exprimant de manière stable les effecteurs-TAP

Nous avons décidé d'utiliser des rétrovirus recombinants pour générer des lignées cellulaires exprimant de manière stable les effecteurs-TAP. Les rétrovirus possèdent une très grande capacité à s'intégrer dans le génome de la cellule hôte, et ils permettent donc d'obtenir une expression stable du transgène sur du long terme.

2.1 Principe d'infection avec des vecteurs viraux

La construction de vecteurs rétroviraux recombinants nécessite deux étapes successives. Au cours de la première, il faut transfecter le plasmide contenant la séquence du gène d'intérêt encadrée par les LTR dans une lignée cellulaire d'empaquetage (HEK293gp). Ces cellules expriment de manière constitutive les gènes *gag*, *pol* et *env*, qui codent respectivement pour la capsid, la transcriptase inverse et l'enveloppe. Cependant, aucun signal d'encapsidation (Ψ) n'est présent dans ces cellules, en revanche, il est retrouvé dans le clone d'expression TAP. Lorsque les cellules 293gp sont transfectées, suite à la transcription du fragment compris entre les régions régulatrices virales LTR, l'ARN est encapsidé puis enveloppé, formant un rétrovirus recombinant qui va être libéré dans le milieu de culture.

Cellule d'encapsulation HEK293gp

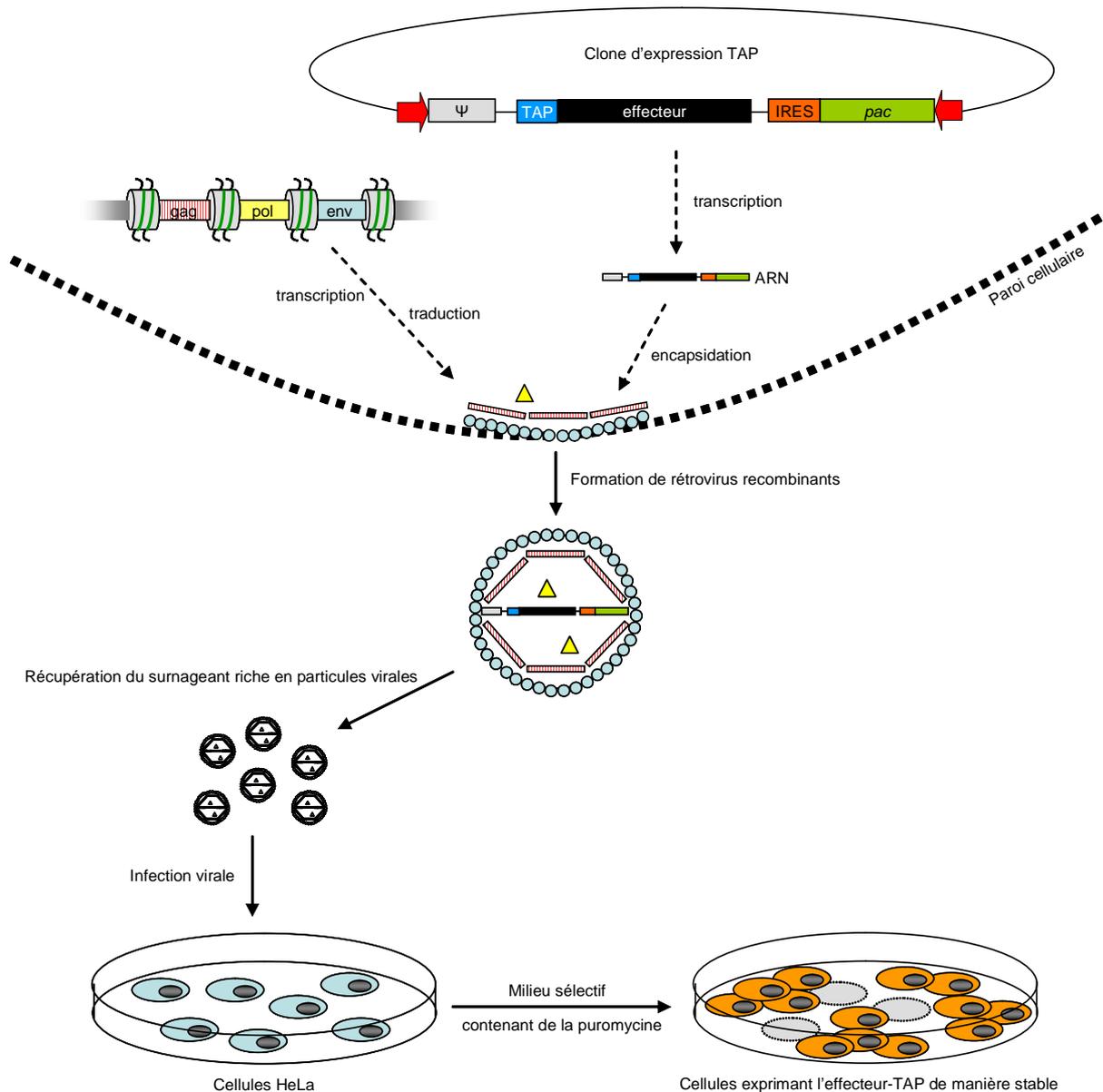


Figure 20 - Principe de l'infection virale pour l'obtention de lignées stables TAP

Les cellules d'emballage permettent une forte production de particules virales, leur milieu de culture est collecté et filtré (pour éviter d'ensemencer des cellules 293gp dans les cellules HeLa et éliminer les débris cellulaires).

Au cours de la seconde étape appelée « infection », le milieu de culture contenant les rétrovirus recombinants est déposé sur des cellules HeLa en culture, pendant 24h. Ensuite le milieu de culture est remplacé par du milieu contenant l'antibiotique puromycine afin de sélectionner les cellules qui ont intégré le transgène dans leur génome. En effet, le transgène contient le gène *pac* qui code pour la puromycine N-acétyltransférase, qui confère à la cellule une résistance. Remarque : comme les particules virales ne contiennent pas les gènes *gag*, *pol* et *env*, les cellules HeLa transduites ne peuvent pas produire de rétrovirus recombinants.

Avantages : les rétrovirus ont la capacité d'encapsider des fragments d'ADN d'une taille importante (7kpb), l'efficacité d'infection est généralement élevée par rapport aux autres méthodes de transfert de gènes, enfin la stabilité d'expression permet une expression du transgène sur du long terme.

Inconvénients : l'utilisation des rétrovirus nécessite de suivre des précautions particulières au cours de leur manipulation (en laboratoire L3), le nombre de particules virales qui pénètrent dans les cellules n'est pas contrôlé, donc certaines cellules hôtes peuvent posséder plusieurs transgènes dans leur génome, ce qui peut modifier l'expression de gènes importants. Enfin les niveaux d'expression de l'effecteur-TAP ne sont pas contrôlés puisque les niveaux de transcription dépendent du contexte génomique du site d'intégration et du nombre de transgènes incorporés.

2.2 Choix du modèle cellulaire

Comme nous souhaitons travailler avec des cellules humaines en culture et qu'il allait nous falloir des quantités importantes de complexes pour réaliser les identifications par spectrométrie de masse, un grand nombre de cellules était requis. Compte tenu de leur vitesse de prolifération et de leur protéome relativement bien décrit, nous avons donc opté pour des cellules HeLa. Cette lignée cellulaire immortalisée provient d'un prélèvement de métastase effectué sur une patiente (Henrietta Lacks) atteinte d'un cancer du col de l'utérus.

Inconvénient : ces cellules possèdent des chromosomes surnuméraires (4 copies du chromosome 12, 2 des chromosomes 6, 8 et 17), donc l'organisation du génome, en particulier les régions eu- et hétéro-chromatiques ne sont probablement pas strictement identiques à celles de lignées primaires ou de cellules différenciées.

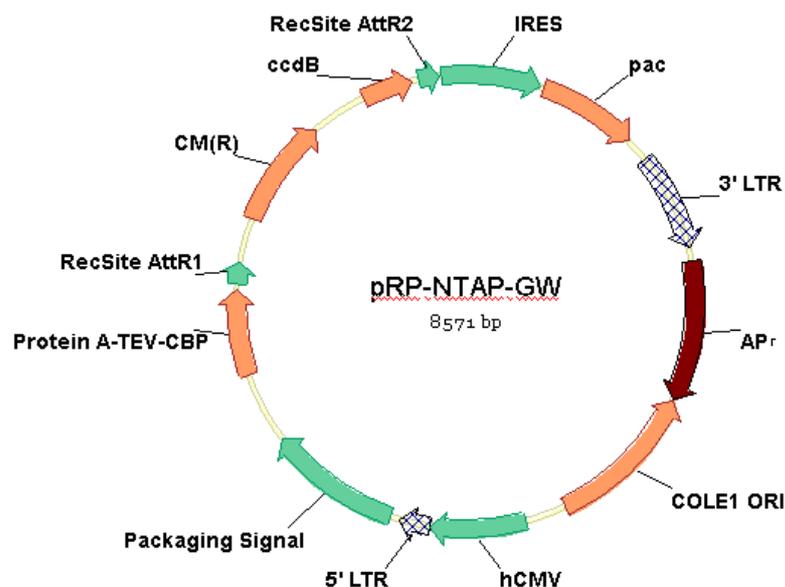


Figure 21 - Carte du vecteur de destination pRP-NTAP-GW

Nous avons généré un plasmide d'expression contenant tous les éléments nécessaires à la production de particules virales par les cellules 293gp (Figure 20) :

- LTR
- Séquence TAP contenant la protéine A, le site de clivage par la TEV et le domaine de liaison à la calmoduline
- la cassette GATEWAY
- IRES + gène *pac*
- promoteur eucaryote fort du cytomégalo virus (CMV)

2.3 Validation des lignées stables et localisation des protéines CBX-TAP

Afin de valider la production de protéines CBX-TAP par les lignées HeLa infectées, nous avons effectué des expériences d'immunomarquage fluorescent à l'aide de l'anticorps de lapin PAP (peroxydase anti peroxydase) dont la chaîne lourde a une forte affinité pour la protéine A de l'étiquette TAP. Un anticorps secondaire dirigé contre le lapin et couplé à l'alexa fluor 594 permet de localiser les protéines immunomarquées, car en l'excitant à 590nm, ce fluorophore émet à 619nm (rouge). Postérieurement, un traitement de ces mêmes cellules avec du Hoechst permet de marquer l'ADN des cellules, délimitant ainsi le noyau du cytoplasme. L'agent Hoechst 33342 est excité à 350nm (ultraviolet) et émet une fluorescence bleue à environ 461nm.

Les observations réalisées avec les lignées TAP à l'aide d'un microscope à fluorescence sont résumées en Figure 21. Nous avons choisi de présenter le grossissement 20x qui permet d'avoir un champ couvrant plusieurs cellules mais permettant de visualiser les structures particulières du noyau. D'une manière générale, nous constatons que les taux d'infection sont élevés, estimés visuellement à 90%, seulement quelques cellules ne présentent pas de marquage rouge. Ces dernières sont probablement infectées puisqu'elles présentent une résistance à la puromycine, mais produisent des quantités trop faibles d'effecteur-TAP pour qu'on puisse les détecter avec l'anticorps PAP. Aussi, nous remarquons que les protéines TAP sont très majoritairement localisées au niveau du noyau, conformément à ce qui est décrit dans la littérature (Ren X, 2008, Mol Cell Biol). L'étiquette TAP ne modifie donc pas la localisation cellulaire des protéines CBX. Ces premiers résultats valident nos lignées TAP, nous avons ensuite entrepris d'autres types d'expériences à partir de ce matériel de départ.

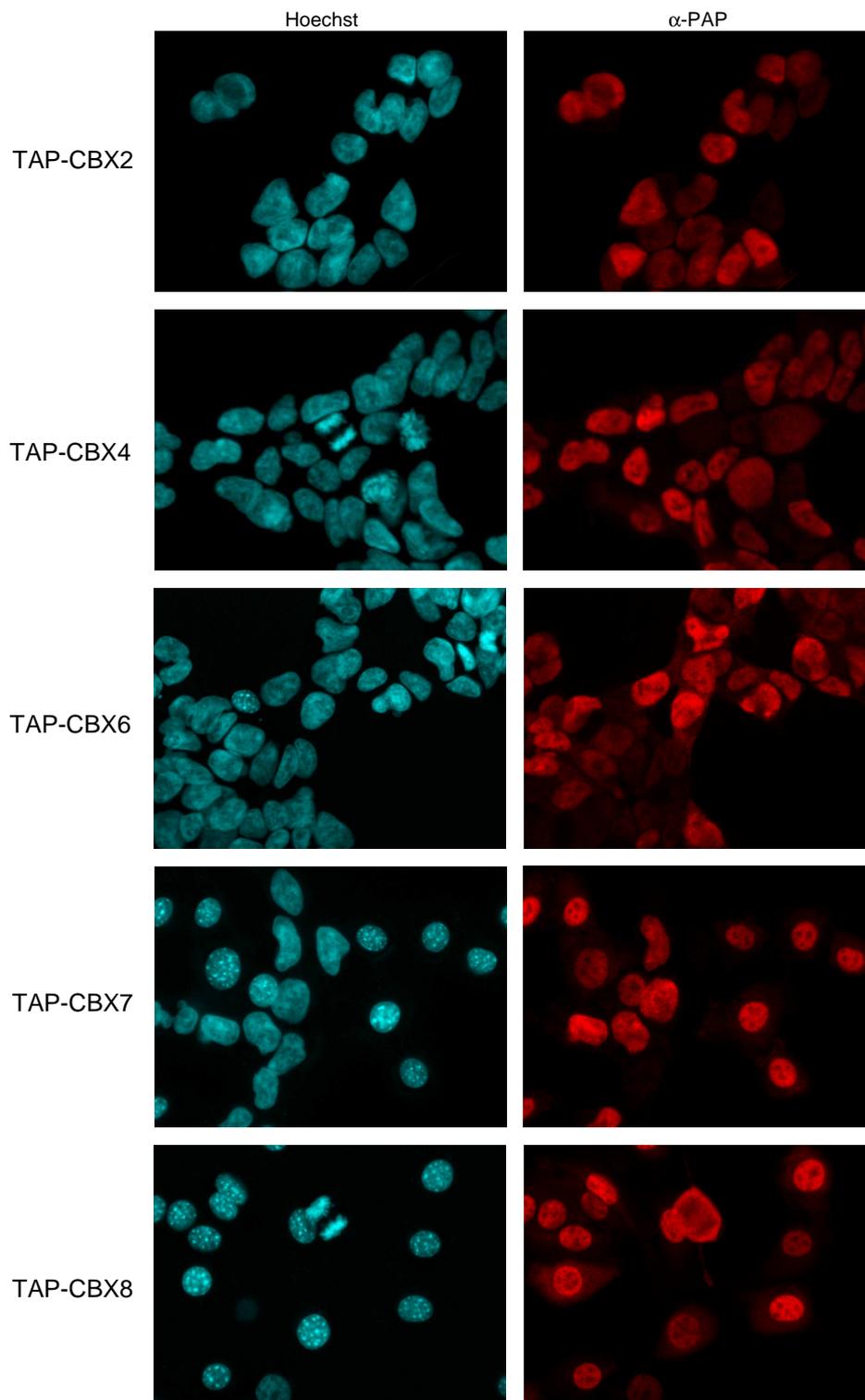


Figure 22 - Immunomarquage des protéines CBX dans les lignées stables HeLa-TAP. Hoechst: marquage de l'ADN, PAP: marquage de la protéine TAP. Grossissement x 200

2.4 Discussion

Chez les mammifères, les membres du PRC1 ne sont pas uniquement localisés dans des régions d'euchromatine, ils sont également accumulés au niveau de foyers spécifiques dans

l'hétérochromatine péricentrique, appelés « corps PcG » (Saurin AJ, 1998, J Cell Biol / Voncken JW, 1999, J Cell Sci). La signification fonctionnelle de ces domaines nucléaires reste méconnue, mais l'intégrité de l'hétérochromatine péricentrique d'une manière générale est essentielle pour assurer la ségrégation des chromosomes, la stabilité génomique et la répression transcriptionnelle (Peters AH, 2001, Cell / Taddei A, 2001, Nat Cell Biol). Les corps PcG apparaissent comme étant des domaines stables qui peuvent être observés dans de nombreuses lignées transformées et dans quelques cellules primaires (Saurin AJ, 1998, J Cell Biol / Voncken JW, 1999, J Cell Sci). Des expériences de retour de fluorescence après photoblanchiment (FRAP : fluorescence recovery after photobleaching) avec une protéine de fusion BMI1-GFP (Green fluorescent protein) ont mis en évidence que cette protéine était beaucoup moins mobile en phase G1 (croissance, préparation de la réplication) par rapport à G2 (croissance, préparation de la mitose) et que le recrutement de cette protéine au niveau des corps PcG était dépendante du complexe PRC2 et de Dnmt1 (Hernandez-Munoz I, 2005, Mol Cell Biol).

Plusieurs protéines composant le PRC1 (BMI1, RING1, PHC1, CBX4) ont été associées avec des séquences spécifiques de régions hétérochromatiques à proximité du centromère du chromosome 1 et des séquences similaires de plusieurs autres chromosomes, pour former les corps PcG (Satiijn DP, 1997, Mol Cell Biol / Saurin AJ, 1998, J Cell Biol / Voncken JW, 1999, J Cell Sci). Cependant, la chromatine péricentrique des corps PcG est à distinguer de la chromatine péricentrique constitutive contenant les HP1. En effet, les régions « décorées » par les corps PcG sont enrichies en H3K27me3, tandis que les régions réprimées par les HP1 possèdent des méthylations sur H3K9, aussi bien chez la drosophile (Tolhuis B, 2006, Nat Genet) que dans des cellules de mammifères (Ku M, 2008, PLoS Genet).

Nous n'avons pas retrouvé ces structures particulières dans les cellules HeLa infectées, bien que certaines régions soient plus « sombres » à l'intérieur du noyau, nous n'avons pas pu distinguer de points lumineux intenses synonymes d'éventuels corps PcG. En revanche le marquage de l'ADN par le Hoechst révèle la présence de foyers d'hétérochromatine dans presque toutes les cellules CBX8-TAP, dans la moitié des cellules CBX7-TAP et dans 2% des cellules CBX6-TAP. Ces foyers ne co-localisent pas avec un enrichissement des protéines TAP, on pourrait même constater une exclusion, signifiant que leur formation n'était pas structurellement formée par ces protéines. Ces régions font penser à des foyers d'hétérochromatine induits par la sénescence : les SAHF (senescence associated heterochromatic foci) (Zhang R, 2007, Cell Cycle). Cependant, les protéines CBX7 et CBX8 (Gil J, 2004, Nat Cell Biol / Dietrich N, 2007, EMBO J) sont capables d'induire un échappement à la sénescence lorsqu'elles sont sur-exprimées dans des cellules primaires normales, donc nous supposons qu'il s'agit plutôt de foyers d'hétérochromatisation qui ne sont pas des SAHF, d'autant plus que nos lignées TAP ont des vitesses de prolifération similaires à la lignée HeLa et que leur morphologie reste inchangée (pas de phénotype de cellules sénescents). Un immunomarquage avec un anticorps dirigé contre les HP1 ou H3K9me3 permettrait de statuer, le chapitre de ce manuscrit consacré aux HP1 permet de conclure qu'il s'agit d'hétérochromatine constitutive (ou assimilée).

3 Evaluation des masses moléculaires des complexes associés aux protéines CBX Polycomb

3.1 Principe de la chromatographie d'exclusion stérique

La chromatographie d'exclusion stérique (ou gel filtration ou tamisage moléculaire) est basée sur la filtration en phase liquide de molécules de tailles différentes à travers un gel constitué de microbilles poreuses, et permet leur séparation physique en fonction de leur taille et de leur forme, c'est-à-dire en fonction de leur encombrement stérique (Figure 22). Les grosses molécules, dont la taille est supérieure à celle des pores sont exclues du gel et donc éluées les premières. Les petites et moyennes molécules sont retenues plus ou moins dans le gel et sont éluées plus tardivement. Il existe une relation linéaire entre le volume d'éluion et le logarithme de la masse moléculaire. Les molécules sont éluées dans l'ordre inverse de leur masse moléculaire.

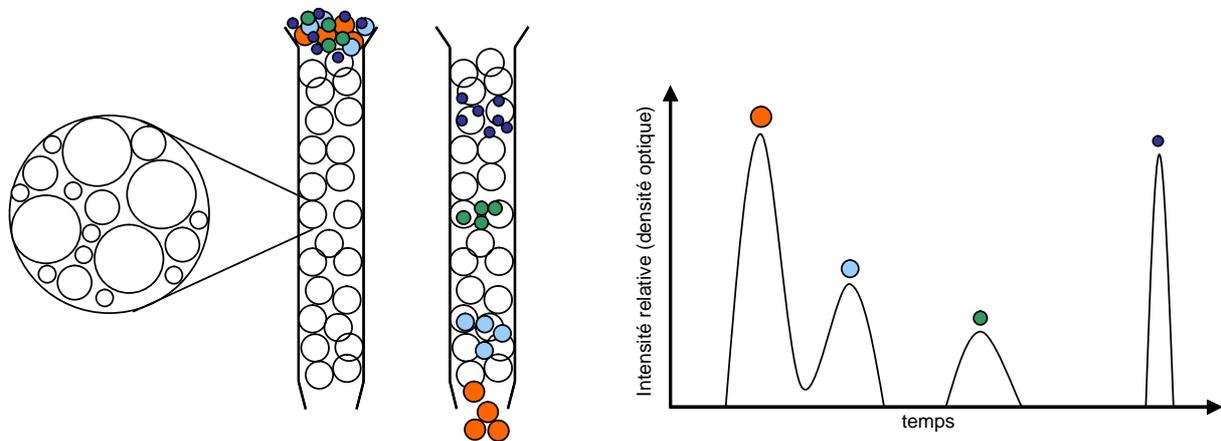


Figure 23 - Principe de la chromatographie d'exclusion, l'ordre d'éluion est inversement proportionnel à la taille des complexes.

Comme le complexe PRC1 humain a été estimé à environ 500kDa (Levine SS, 2002, Mol Cell Biol), nous avons choisi une colonne dont les caractéristiques permettent la séparation de complexes allant de 5 à 5000 kDa (Superose 6 10/300 GL, Amersham). Nous avons établi une courbe de calibration (Figure 23) à partir des volumes d'éluion de protéines commerciales de masse connue : Thyroglobuline (669 kDa), Apoferritine (443 kDa), β -amylase (200 kDa), alcool déshydrogénase (150 kDa), albumine de sérum bovin (66 kDa) et anhydrase carbonique (29 kDa). Par extrapolation, nous pouvons estimer la masse moléculaire apparente des complexes analysés par cette colonne en fonction de leur volume d'éluion.

Inconvénient :

Il n'existe aujourd'hui aucune standardisation sur les tampons qui sont utilisés pour réaliser ce type de chromatographie, d'un côté le fournisseur déconseille l'utilisation de certains composants comme le glycérol, de l'autre les scientifiques cherchent à s'approcher le plus possible de la viscosité et des conditions cellulaires pour conserver les complexes en conformation native. Les résines de filtration possédant une limite d'exclusion élevée sont très fragiles car elles ont une grande porosité. Les

vitesse de débit et les pressions doivent donc être soigneusement contrôlées pour éviter d'endommager les billes du gel. Pour finir, la séparation des molécules est réalisée en fonction du rayon de Stokes et non en fonction de la masse moléculaire seule, donc les valeurs obtenues par extrapolation grâce à la courbe de calibration, bien que fiables d'une manière générale, ne sont pas très précises.

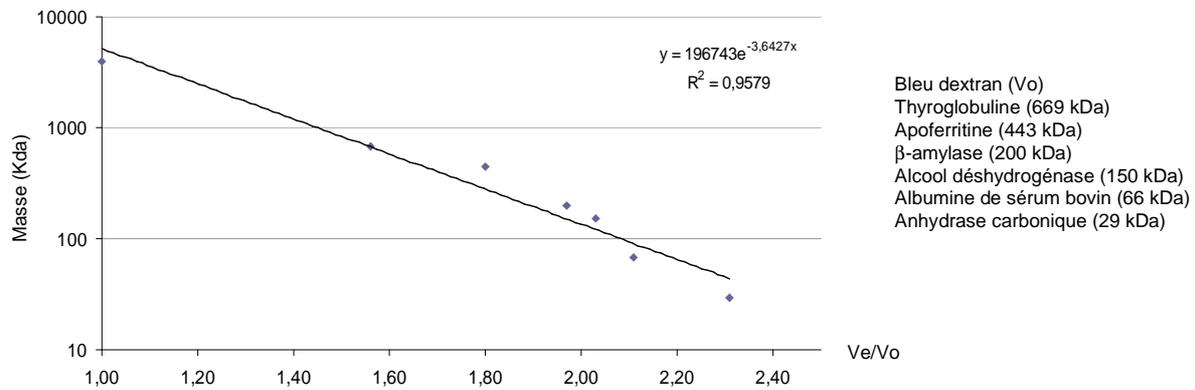


Figure 24 - Courbe de calibration de la colonne Superose 6 10/300 GL représentée sur une échelle semi-logarithmique. V_e/V_o = Volume d'éluion / Volume vide de la colonne

3.2 Résultats

Afin d'étudier les complexes associés aux protéines CBX, nous avons extrait les protéines nucléaires des différentes lignées TAP dans un tampon non dénaturant permettant de conserver les interactions protéiques, donc les complexes. Après avoir équilibré la colonne avec un tampon de composition identique à celui utilisé pour l'extraction des protéines, nous avons analysé un à un les 5 extraits nucléaires, puis nous avons caractérisé les différentes fractions collectées par western blot. Les résultats obtenus sont présentés en Figure 24.

D'une manière identique, nous avons analysé les protéines RING1-TAP et RNF2-TAP avec l'anticorps PAP, et nous avons également fractionné un échantillon de protéines nucléaires provenant de la lignée HeLa et effectué une révélation de la protéine RING1 avec un anticorps dirigé contre la protéine endogène.

Toutes les protéines CBX-TAP sont majoritairement associées à des complexes de haut poids moléculaire (> 1000 kDa) et minoritairement présentes dans des complexes de taille plus faible (entre 2000 et 500). En ce qui concerne les 2 protéines de la famille RING, elles sont éluées principalement dans les fractions correspondant à des complexes de masse apparente comprise entre 200 et 500 kDa. Les protéines RING1-TAP et RNF2-TAP sont également détectées dans la première fraction tandis que la protéine RING1 endogène n'est pas présente dans cette fraction.

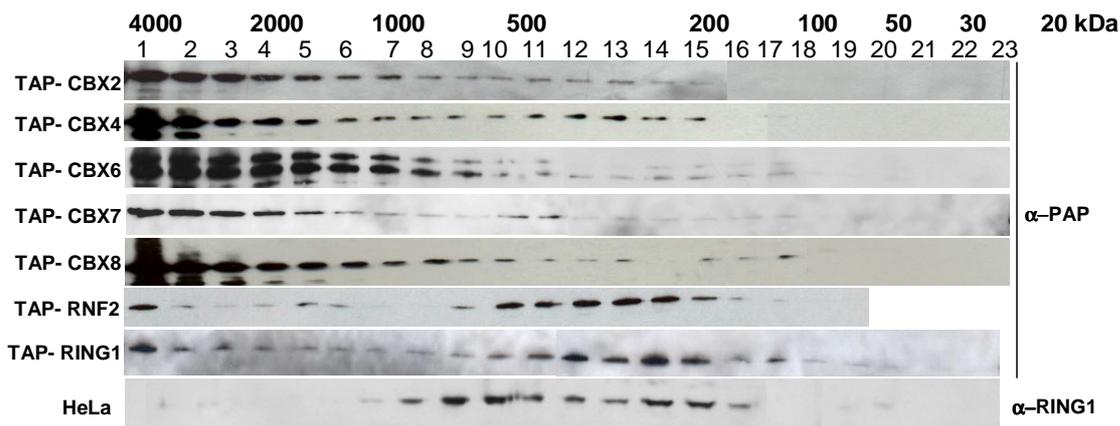


Figure 25 - Analyse par western blot (anticorps PAP ou RING1) des fractions obtenues après chromatographie d'exclusion stérique de 2mg d'extraits nucléaires (cellules HeLa-TAP ou HeLa). Poids moléculaire approximatif indiqué en kilodaltons, fractions 1 à 25.

3.3 Discussion

L'absence de détection de RING1 endogène dans les extraits HeLa au niveau de la première fraction (dernière ligne, Figure 24) prouve que la présence de RING1-TAP et RNF2-TAP dans cette même fraction correspond à un artéfact d'expérience, en effet, nous pensons que la première fraction collectée, qui coïncide avec le volume mort de la colonne, contiendrait d'éventuels agrégats dus à la surexpression de la protéine étudiée ou à l'étiquette TAP. La superposition des profils dans les fractions suivantes entre RING1-TAP et RING1 endogène valide la technologie utilisée et démontre que l'étiquette TAP n'interfère pas avec la formation de complexes dans la cellule.

Les protéines RING1 et RNF2 participent à la formation de complexes de même masse moléculaire, correspondant à la taille estimée du PRC1 dans des cellules humaines en culture (Levine SS, 2002, Mol Cell Biol).

Les protéines CBX participent toutes à la formation de complexes de haut poids moléculaire de taille supérieure à 1000 kDa, même si les signaux de la première fraction peuvent être considérés comme artéfactuels, les intensités des fractions suivantes prouvent que ces protéines sont bien présentes dans les premières fractions d'éluion. La présence de ces protéines est également retrouvée, en plus faible intensité, au niveau des masses moléculaires allant de 200 à 500 kDa. Nous remarquons que parmi les protéines analysées, aucune n'est retrouvée seule dans les extraits nucléaires puisqu'aucun signal n'est détecté au niveau de leur poids moléculaire. Compte tenu de la limite de détection par l'anticorps, cela signifie simplement que la majorité des protéines TAP sont associées à un complexe dans les lignées HeLa, leur forme libre étant minoritaire.

En conclusion, nous proposons que les complexes PRC1 existent sous plusieurs formes dans les cellules humaines, d'une part un complexe de 200 à 500 kDa, d'autre part des complexes d'encombrement stérique plus important de masse comprise entre 1000 et 3000kDa.

4 Purification TAP des protéines CBX Polycomb

4.1 Principe

Le principe de la purification TAP a été présenté dans la 4^{ème} partie de l'introduction, nous avons également discuté les avantages et les inconvénients de cette technique.

4.2 Résultats

Nous avons réalisé chaque purification en duplicata (Figures 25) et nous avons suivi les différentes étapes de la purification en réalisant des western blots pour suivre la présence de la protéine d'intérêt au cours des deux étapes du processus de purification (Figure 26). Nous remarquons que les profils obtenus par coloration au bleu de coomassie colloïdal des gels d'acrylamide sont très proches entre 2 purifications réalisées indépendamment, confirmant que la technologie est maîtrisée et que sa reproductibilité est bonne. Les colorations des gels pour CBX6 et CBX8 sont les plus intenses alors que pour CBX2 et CBX7, peu de protéines sont révélées.

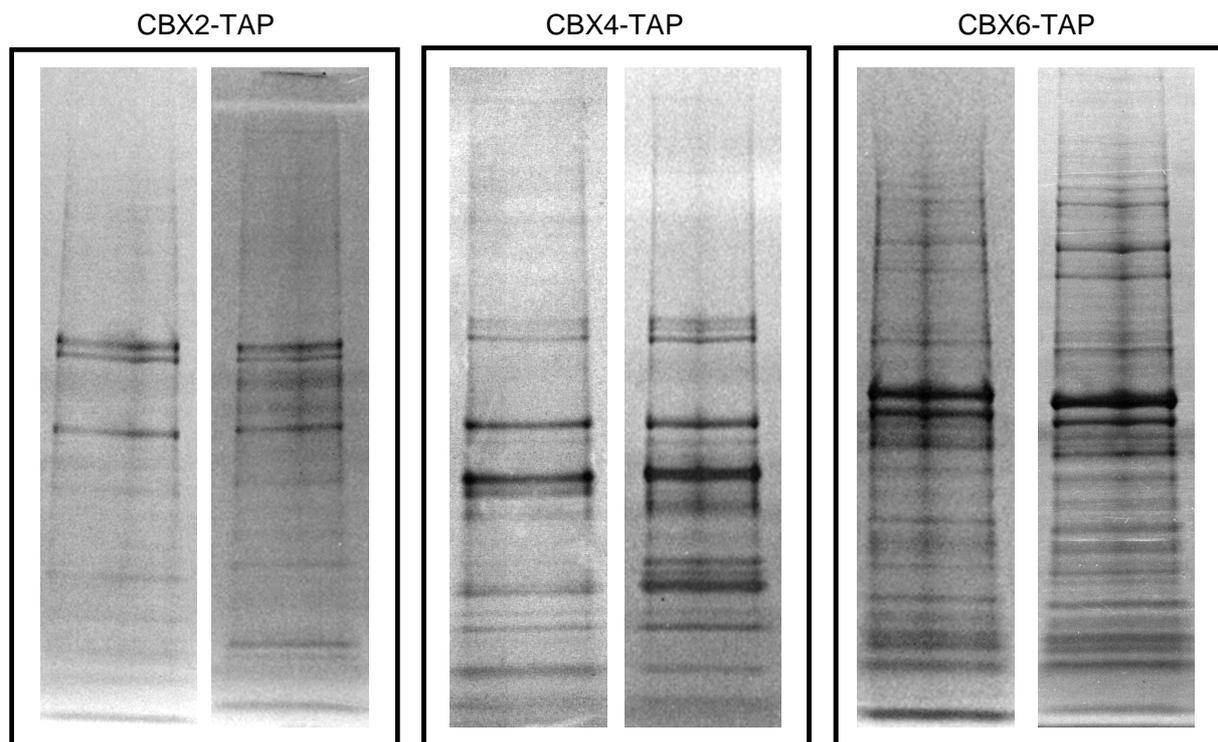


Figure 26.1 - Gels SDS-PAGE TAP des protéines CBX Polycomb colorés au bleu de coomassie colloïdal

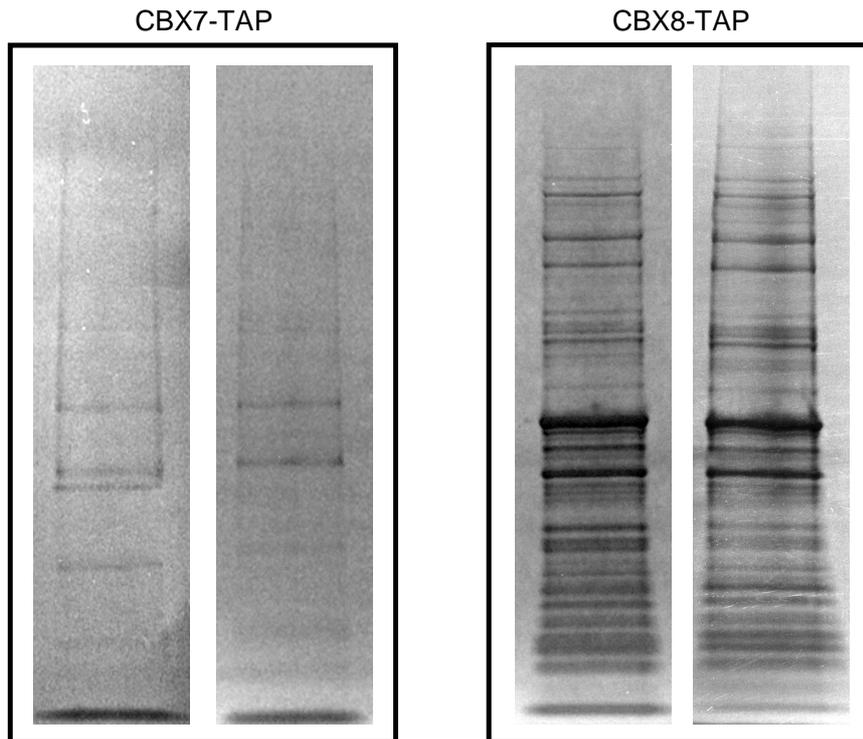


Figure 25.2 - Gels SDS-PAGE TAP des protéines CBX Polycomb colorés au bleu de coomassie colloïdal

D'une manière générale nous constatons que la protéine d'intérêt est enrichie au cours des différentes étapes de purification (Figure 26). En colorant au rouge ponceau la membrane qui a servi pour le western blot, nous avons constaté une perte d'intensité du marquage de la piste de migration au fur et à mesure des étapes de purification, signifiant que nous avons purifié les partenaires protéiques spécifiques des protéines TAP et que les autres protéines nucléaires ont été éliminées.

Nous constatons qu'une partie des protéines TAP n'est pas éluée des billes car un signal est révélé dans la piste billes CBP, cependant le signal est faible en comparaison avec l'éluat (piste élution CBP).

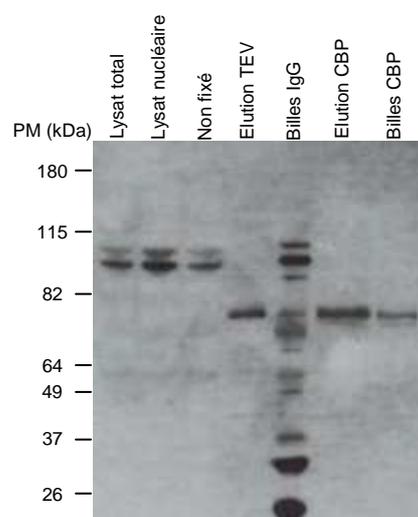


Figure 27 - Suivi de la purification de la protéine TAP-CBX6 par western blot anti-CBP

5 Identification des partenaires protéiques purifiés par TAP

5.1 Principe

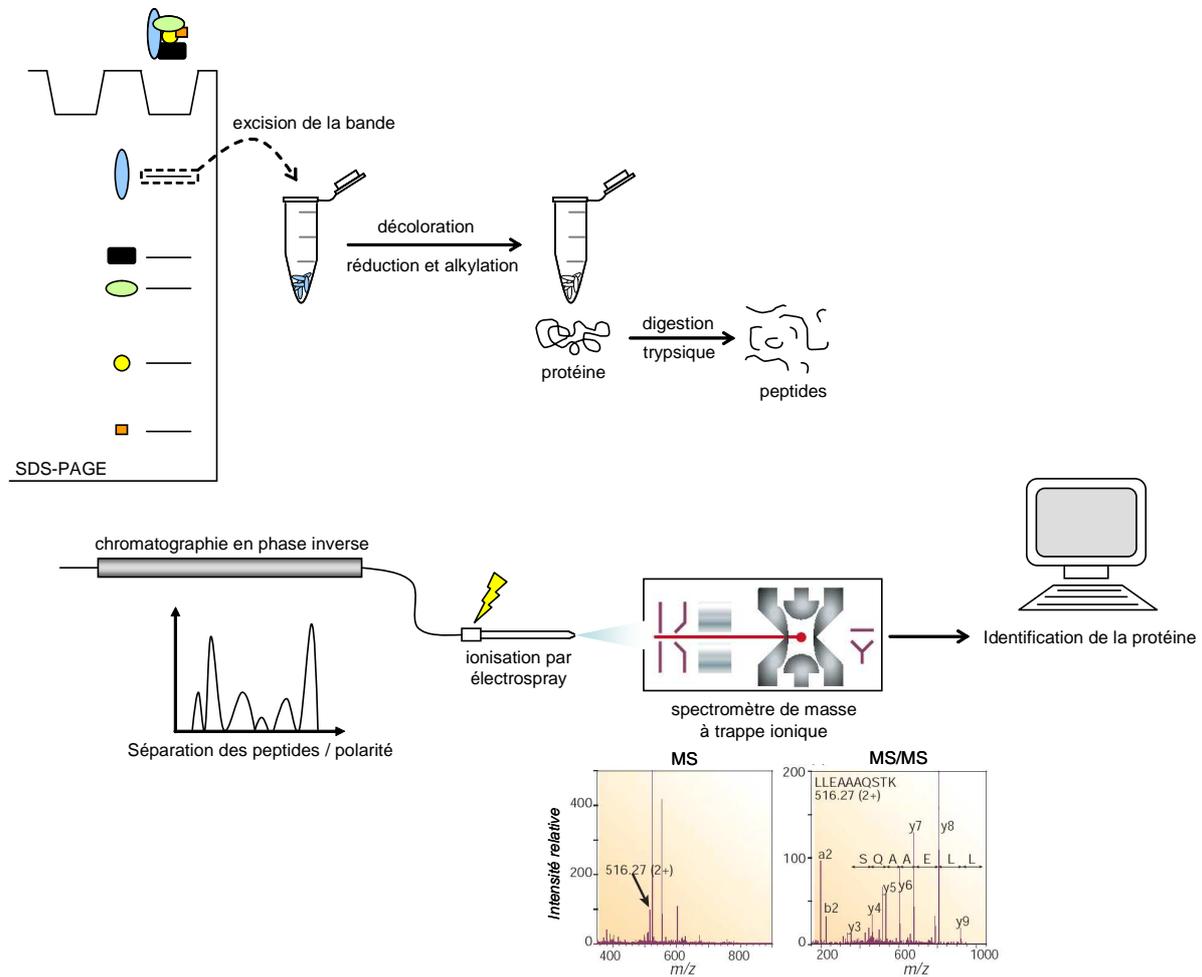


Figure 28 - Etapes successives pour l'identification par spectrométrie de masse des protéines présentes dans un gel de polyacrylamide (= approche bottom-up)

Migration et digestion trypsique des protéines co-purifiées

La purification d'affinité en tandem permet d'obtenir un échantillon contenant les protéines d'intérêt associées à leurs partenaires protéiques. Les échantillons sont dénaturés en les mélangeant avec du tampon Laemmli et par chauffage, puis chargés dans un gel d'acrylamide afin de séparer les protéines co-purifiées en fonction de leur poids moléculaire dans un tampon de migration dénaturant, sous l'action d'un champ électrique. Après migration, les protéines sont colorées à l'aide d'un bleu de coomassie colloïdal. La piste correspondant à l'échantillon TAP est découpée de manière systématique en une trentaine de fragments de 2mm de hauteur environ, en essayant de séparer 2 protéines « voisines » révélées par la coloration. Les fragments sont isolés dans des tubes 1,5mL, échantillonnés en plusieurs petits morceaux, puis décolorés. Suite à la décoloration, les protéines sont alkylées et réduites pour éliminer les ponts disulfures, enfin l'ajout de trypsine permet de digérer les protéines en peptides qui vont pouvoir être élués du gel d'acrylamide et analysés par spectrométrie de masse. La trypsine coupe spécifiquement après les résidus basiques (arginine et lysine) et permet

d'obtenir en moyenne des peptides dans une gamme de 500 à 3500 Da. Les peptides élués du gel sont ensuite analysés par chromatographie liquide couplée à une spectrométrie de masse en tandem (LC-MS/MS).

Nano-chromatographie en phase inversée

Dans un premier temps les peptides sont séparés en fonction de leur polarité par nano-chromatographie liquide en phase inversée. Tandis que la phase stationnaire se distingue par son apolarité (silices apolaires avec 18 atomes de carbone), la phase mobile (peptides) est polaire et hydrophile. La séparation est fondée sur des interactions hydrophobes entre les molécules à séparer et la phase stationnaire. Plus le peptide sera apolaire, plus il sera retenu dans la colonne, à l'inverse plus un peptide sera polaire, plus il sera rapidement élué, entraîné par la phase mobile liquide. La composition en acides aminés (hydrophile ou hydrophobes) permet donc de séparer les différents peptides.

Ionisation des peptides par électrospray

En sortie de colonne, les peptides en solution sont transférés en phase gazeuse grâce à l'ionisation par électrospray. L'électrospray est produit par application d'un champ électrique élevé et positif sur le liquide traversant un fin capillaire. C'est ainsi que des gouttelettes chargées positivement sont produites par le jet du cône de nébulisation se déplaçant vers le spectromètre de masse, qui rencontre un gaz sec. L'évaporation du solvant provoque une réduction de la taille des gouttelettes pendant leur déplacement, jusqu'à leur fission qui produit des ions non fragmentés possédant plusieurs protons (peptides multichargés positivement). Ces derniers sont analysés dans un spectromètre de masse (à trappe ionique dans le cadre de nos études).

Spectromètre de masse à trappe ionique (analyseur à piégeage d'ions)

L'analyseur permet de mesurer le rapport masse/charge (m/z) des ions (peptides chargés). La trappe ionique peut piéger des ions (gamme de 50 à 2000 m/z), mesurer leur masse (1^{ère} MS), isoler un ion (parent ou précurseur), le fragmenter et mesurer les m/z des fragments (MS/MS). Les ions produits par la source électrospray sont focalisés et transmis à la trappe ionique par un quadripôle. La trappe est constituée de 3 électrodes (une annulaire centrale, une d'entrée et une de sortie). L'électrode annulaire permet de piéger les ions et de les retenir. En augmentant la tension de cette électrode, les ions sont éjectés progressivement de la trappe en fonction de leur m/z et sont détectés en sortant de la trappe. C'est ainsi qu'un premier spectre MS permet d'obtenir les m/z de tous les peptides présents dans l'échantillon à un instant t .

Avant la fragmentation, l'ion d'intérêt est isolé dans la trappe. Tous les autres ions sont éjectés de la trappe en utilisant d'une part le processus d'éjection résonnante pour les ions plus lourds que l'ion d'intérêt, d'autre part par l'application sur les électrodes d'entrée et de sortie des fréquences de résonance des ions à éjecter (= différentes de celle de l'ion d'intérêt). L'ion isolé est fragmenté en percutant de l'hélium (gaz neutre) présent dans la trappe, avec une énergie suffisante pour induire sa

fragmentation. Les fragments générés n'ont pas la même fréquence de résonance que l'ion parent, ils sont analysés suivant le même principe que celui décrit précédemment.

Identification des protéines par analyse bio-informatique

Le traitement des signaux et la visualisation des spectres sont gérés par un système informatique, les informations sur les peptides sont extraites et converties dans un format qui permet leur confrontation à une base de données (Swissprot : la plus annotée, la moins redondante) regroupant les séquences du protéome humain. L'identification est réalisée en ligne à l'aide du serveur Mascot (matrix science). Elle doit répondre à des critères d'acceptation spécifiques (valeur du score, nombre de peptides identifiés/protéine, couverture de séquence, adéquation avec la masse moléculaire estimée sur le gel...).

Avantage : la fragmentation dans la trappe ionique est particulièrement efficace grâce à l'ajustement automatique de l'énergie de collision en fonction de la masse de l'ion parent. La trappe permet de réaliser des MS^n , c'est-à-dire de réaliser des MS/MS successives à partir d'un unique ion parent. L'analyse d'ions multichargés permet d'étudier une molécule de plus haut poids moléculaire que la limite de l'analyseur (un peptide de 6000 Da avec 3 charges positives a un m/z de 2000 et peut être analysé par la trappe).

Inconvénient : Le temps d'analyse car l'étape de chromatographie est souvent longue (45-60 minutes / découpe de gel). En ce qui concerne la trappe, la répétabilité des spectres d'un même échantillon est assez moyenne et la gamme m/z limitée. Même si les ions multichargés permettent l'analyse de peptides de plus haut poids moléculaire, cela complique l'analyse des spectres. Lors de l'identification, certains peptides ne sont pas attribués à une protéine (protéine absente de la base de données, modifications post-traductionnelles, clivage trypsique non spécifique...), donc certaines informations sont perdues.

5.2 Protéines co-éluées avec les CBX-TAP

Les données brutes (fichiers de type raw) obtenues après analyse par le spectromètre de masse à trappe ionique sont converties dans un format .dta qui nous permet de confronter les résultats obtenus à la banque de données swissprot sur le site internet Mascot de Matrix Science. Nous avons utilisé la version publique de ce site, et les identifications présentées dans ce manuscrit proviennent de l'interrogation de la même version de la base de données.

Dans un premier temps, les identifications ont été obtenues avec les critères de sélection correspondant au spectromètre de masse utilisé, de plus nous avons exclu les identifications dont le score Mascot était inférieur à 38 (en deçà, l'identification est considérée comme ayant une probabilité inférieure ou égale à l'identification d'un peptide aléatoire). Après avoir compilé l'ensemble des identifications de chaque fragment de gel dans un même tableau, nous avons étudié la localisation cellulaire (quand elle était disponible) de chacune des protéines identifiées. Compte tenu de la localisation des protéines CBX et de notre intérêt d'étudier les complexes PRC1, nous avons choisi de

retirer de la liste toutes les protéines décrites comme n'étant pas localisées au niveau du noyau cellulaire. Les protéines qui ont franchi avec succès les critères de sélection que nous nous sommes imposés pour éviter les faux positifs sont présentées en Tableau 4. Nous remarquons que les protéines CBX2 et CBX7, qui avaient une coloration faible sur gel, comptent le moins de protéines partenaires potentielles (respectivement 25 et 15) tandis que CBX6 et CBX8, qui avaient les colorations les plus intenses, dénombrent une soixantaine d'identifications. CBX4, avec ses 41 protéines co-éluées présente un nombre intermédiaire. L'ensemble des peptides identifiés sont consultables dans les Tableau S1 à S5 en Annexe 1. Le Tableau 8 regroupe l'ensemble des protéines considérées comme contaminants dans le cadre de notre étude, son contenu sera discuté dans la troisième partie.

	purif	accession	description	daltons	score	peptides
CBX2	A	ANR26_HUMAN	Ankyrin repeat domain-containing protein 26	196202	39	2
CBX2	B	BMI1_HUMAN	Polycomb complex protein BMI-1	36925	201	4
CBX2	A / B	CBX2_HUMAN	Chromobox protein homolog 2	56046	583 / 379	43 / 26
CBX2	B	CENPE_HUMAN	Centromeric protein E	316219	62	6
CBX2	A / B	CSK21_HUMAN	Casein kinase II subunit alpha	45115	136 / 66	2 / 2
CBX2	B	CSK22_HUMAN	Casein kinase II subunit alpha'	41187	94	1
CBX2	B	CSK2B_HUMAN	Casein kinase II subunit beta	24926	136	3
CBX2	B	DI3L2_HUMAN	DIS3-like exonuclease 2	99147	43	2
CBX2	B	GTF2I_HUMAN	General transcription factor II-I	112346	211	4
CBX2	B	H2A1B_HUMAN	Histone H2A type 1-B	14127	80	6
CBX2	B	H2B1C_HUMAN	Histone H2B type 1-C/E/F/G/I	13811	98	4
CBX2	B	H4_HUMAN	Histone H4	11360	57	1
CBX2	A	IQWD1_HUMAN	Nuclear receptor interaction protein	96292	48	2
CBX2	A	JHD2A_HUMAN	JmjC domain-containing histone demethylation protein 2A	147234	49	3
CBX2	A	LBA1_HUMAN	Lupus brain antigen 1 homolog	336007	64	3
CBX2	A	MYCB2_HUMAN	Probable E3 ubiquitin-protein ligase MYCBP2	509759	39	5
CBX2	A	NPM_HUMAN	Nucleophosmin	32555	51	1
CBX2	A	PAR12_HUMAN	Poly [ADP-ribose] polymerase 12	79013	44	2
CBX2	B	PARP1_HUMAN	Poly [ADP-ribose] polymerase 1	113012	113	2
CBX2	B	PHC2_HUMAN	Polyhomeotic-like protein 2	90657	350	7
CBX2	A / B	PSME3_HUMAN	Proteasome activator complex subunit 3	29488	280 / 195	5 / 8
CBX2	A / B	RING2_HUMAN	E3 ubiquitin-protein ligase RING2	37632	281 / 429	7 / 8
CBX2	A	RN213_HUMAN	RING finger protein 213	373742	57	5
CBX2	A	YD007_HUMAN	Coiled-coil domain-containing protein FLJ25770	127062	41	6
CBX2	A	ZN167_HUMAN	Zinc finger protein 167	84979	48	2

	purif	accession	description	masse	score	peptides
CBX4	A / B	1433B_HUMAN	14-3-3 protein beta/alpha	28065	154 / 329	8 / 21
CBX4	A / B	1433E_HUMAN	14-3-3 protein epsilon	29155	566 / 1151	25 / 97
CBX4	A / B	1433F_HUMAN	14-3-3 protein eta	28201	79 / 670	2 / 42
CBX4	A / B	1433G_HUMAN	14-3-3 protein gamma	28285	188 / 467	8 / 36
CBX4	B	1433S_HUMAN	14-3-3 protein sigma	27757	151	7
CBX4	A / B	1433T_HUMAN	14-3-3 protein theta	27747	182 / 600	9 / 20
CBX4	A / B	1433Z_HUMAN	14-3-3 protein zeta/delta	27728	254 / 572	10 / 44
CBX4	A	ADT2_HUMAN	ADP/ATP translocase 2	32874	143	3
CBX4	A / B	BMI1_HUMAN	Polycomb complex protein BMI-1	36925	212 / 337	7 / 18
CBX4	A / B	CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	266 / 641	66 / 137
CBX4	B	CENPF_HUMAN	Centromere protein F	367537	46	4
CBX4	B	CHD7_HUMAN	Chromodomain-helicase-DNA-binding protein 7	335717	42	4
CBX4	A	CSK21_HUMAN	Casein kinase II subunit alpha	45115	169	3
CBX4	A	CSK22_HUMAN	Casein kinase II subunit alpha'	41187	46	1
CBX4	A / B	CSK2B_HUMAN	Casein kinase II subunit beta	24926	59 / 107	1 / 2

CBX4	A	DDEF1_HUMAN	130 kDa phosphatidylinositol 4,5-biphosphate-dpdt ARF1	125393	49	2
CBX4	B	ERCC6_HUMAN	DNA excision repair protein ERCC-6	168416	43	3
CBX4	B	FER_HUMAN	Proto-oncogene tyrosine-protein kinase FER	94638	64	2
CBX4	A	FHAD1_HUMAN	Forkhead-associated domain-containing protein 1	161804	40	1
CBX4	A	H11_HUMAN	Histone H1.1	21829	41	4
CBX4	B	KHDR2_HUMAN	KH domain-containing, RNA-binding,	38927	46	2
CBX4	B	KI67_HUMAN	Antigen KI-67	358474	60	4
CBX4	B	MDM1_HUMAN	Nuclear protein MDM1	80686	41	2
CBX4	A	MOV10_HUMAN	Putative helicase MOV-10	113599	44	1
CBX4	A / B	PCGF2_HUMAN	Polycomb group RING finger protein 2	37764	104 / 143	3 / 12
CBX4	A / B	PCGF6_HUMAN	Polycomb group RING finger protein 6	39023	63 / 78	1 / 1
CBX4	A	PHC1_HUMAN	Polyhomeotic-like protein 1	105353	40	1
CBX4	A / B	PHC2_HUMAN	Polyhomeotic-like protein 2	90657	141 / 628	4 / 19
CBX4	A	PHC3_HUMAN	Polyhomeotic-like protein 3	106096	103	4
CBX4	A / B	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	345 / 379	12 / 8
CBX4	A / B	RING2_HUMAN	E3 ubiquitin-protein ligase RING2	37632	284 / 598	7 / 19
CBX4	B	RNF17_HUMAN	RING finger protein 17	184643	43	2
CBX4	B	RUVB2_HUMAN	RuvB-like 2	51125	84	2
CBX4	A	SMC3_HUMAN	Structural maintenance of chromosomes protein 3	141454	61	5
CBX4	A	SMG1_HUMAN	Serine/threonine-protein kinase SMG1	410261	46	3
CBX4	B	SSX10_HUMAN	Protein SSX10	17575	42	2
CBX4	A	STALP_HUMAN	AMSH-like protease	49783	38	2
CBX4	B	WDHD1_HUMAN	WD repeat and HMG-box DNA-binding protein 1	125888	41	4
CBX4	B	YD007_HUMAN	Coiled-coil domain-containing protein FLJ25770	127062	38	3
CBX4	A	YR010_HUMAN	Putative uncharacterized protein ENSP00000344348	46232	42	3
CBX4	B	ZN546_HUMAN	Zinc finger protein 546	98369	40	2

	purif	accession	description	masse	score	peptides
CBX6	B	AHDC1_HUMAN	A.T hook DNA-binding motif-containing protein 1	168245	41	3
CBX6	B	BCAS3_HUMAN	Breast carcinoma-amplified sequence 3	101116	42	2
CBX6	B	BPTF_HUMAN	Nucleosome-remodeling factor subunit BPTF	324019	49	3
CBX6	B	BRWD1_HUMAN	Bromodomain and WD repeat-containing protein 1	262756	40	3
CBX6	B	CAF1A_HUMAN	Chromatin assembly factor 1 subunit A	105223	47	2
CBX6	B	CASC5_HUMAN	Protein CASC5	265193	43	5
CBX6	A	CBX2_HUMAN	Chromobox protein homolog 2	56046	47	2
CBX6	A / B	CBX6_HUMAN	Chromobox protein homolog 2	43871	375 / 572	72 / 96
CBX6	B	CD014_HUMAN	Uncharacterized protein C4orf14	78409	64	1
CBX6	B	CE350_HUMAN	Centrosome-associated protein 350	350716	46	3
CBX6	A	CP135_HUMAN	Centrosomal protein of 135 kDa	133422	43	2
CBX6	B	DACT1_HUMAN	Dapper homolog 1	90119	48	2
CBX6	B	DBPA_HUMAN	DNA-binding protein A	40066	88	1
CBX6	B	DNMT1_HUMAN	DNA (cytosine-5)-methyltransferase 1	183050	45	3
CBX6	B	ELP4_HUMAN	Elongator complex protein 4	46588	59	3
CBX6	B	ESF1_HUMAN	ESF1 homolog	98735	69	1
CBX6	B	GNL3_HUMAN	Guanine nucleotide-binding protein-like 3	61958	107	3
CBX6	B	H12_HUMAN	Histone H1.2	21352	45	1
CBX6	B	H14_HUMAN	Histone H1.4	21852	44	2
CBX6	B	H1X_HUMAN	Histone H1x	22474	50	1
CBX6	B	H2A1B_HUMAN	Histone H2A type 1-B	14127	43	1
CBX6	B	H2B1B_HUMAN	Histone H2B type 1-B	13942	118	2
CBX6	B	H31T_HUMAN	Histone H3.1t	15499	47	3
CBX6	A / B	H4_HUMAN	Histone H4	11360	68 / 54	1 / 1
CBX6	B	IF2B1_HUMAN	Insulin-like growth factor 2 mRNA-binding protein 1	63417	576	15
CBX6	B	IF2B2_HUMAN	Insulin-like growth factor 2 mRNA-binding protein 1	61804	89	3
CBX6	A / B	IF2B3_HUMAN	Insulin-like growth factor 2 mRNA-binding protein 1	63681	322 / 39	8 / 1
CBX6	B	ILF2_HUMAN	Interleukin enhancer-binding factor 2	43035	284	4
CBX6	B	ILF3_HUMAN	Interleukin enhancer-binding factor 3	95279	162	4
CBX6	B	KHDR1_HUMAN	KH domain-containing, RNA-binding	48197	49	3
CBX6	B	KRI1_HUMAN	Protein KRI1 homolog	83201	224	4

CBX6	B	LA_HUMAN	Lupus La protein	46808	88	1
CBX6	B	LAS1L_HUMAN	LAS1-like protein	83013	126	2
CBX6	B	LTV1_HUMAN	Protein LTV1 homolog	54821	93	3
CBX6	B	MOV10_HUMAN	Putative helicase MOV-10	113599	254	6
CBX6	B	MPPH1_HUMAN	M-phase phosphoprotein 1	210553	49	7
CBX6	B	NCOR2_HUMAN	Nuclear receptor corepressor 2	274034	49	4
CBX6	B	NUMA1_HUMAN	Nuclear mitotic apparatus protein 1	238260	45	2
CBX6	B	PABP1_HUMAN	Polyadenylate-binding protein 1	70626	155	3
CBX6	B	PHC2_HUMAN	Polyhomeotic-like protein 2	90657	61	1
CBX6	B	PPR3A_HUMAN	Protein phosphatase 1 regulatory subunit 3A	125759	52	3
CBX6	B	RAD17_HUMAN	Cell cycle checkpoint protein RAD17	77006	41	2
CBX6	B	RAD26_HUMAN	Putative DNA repair and recombination protein RAD26-like	81094	42	2
CBX6	B	RGAG1_HUMAN	Retrotransposon gag domain-containing protein 1	144280	58	2
CBX6	A	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	73	1
CBX6	B	RING2_HUMAN	E3 ubiquitin-protein ligase RING2	37632	91	2
CBX6	B	SCML2_HUMAN	Sex comb on midleg-like protein 2	77257	40	3
CBX6	B	THOC4_HUMAN	THO complex subunit 4	26872	64	1
CBX6	B	UBE2O_HUMAN	Ubiquitin-conjugating enzyme E2 O	141265	221	5
CBX6	B	ULK2_HUMAN	Serine/threonine-protein kinase ULK2	112654	41	3
CBX6	B	WD40C_HUMAN	WD repeat-containing protein 40C	50803	45	4
CBX6	B	XYLB_HUMAN	Xylulose kinase	57374	50	1
CBX6	A / B	YBOX1_HUMAN	Nuclease-sensitive element-binding protein 1	35903	237 / 127	8 / 3
CBX6	B	YC009_HUMAN	Uncharacterized protein FLJ36157	190890	52	3
CBX6	B	ZMYM4_HUMAN	Zinc finger MYM-type protein 4	172677	43	3
CBX6	A / B	ZN33A_HUMAN	Zinc finger protein 33A	94384	45 / 42	3 / 3
CBX6	B	ZN33B_HUMAN	Zinc finger protein 33B	90683	60	3

	purif	accession	description	masse	score	peptides
CBX7	A / B	BMI1_HUMAN	Polycomb complex protein BMI-1	36925	199 / 100	4 / 3
CBX7	A / B	CBX7_HUMAN	Chromobox protein homolog 7	28323	232 / 55	10 / 1
CBX7	B	CSK21_HUMAN	Casein kinase II subunit alpha	45115	255	3
CBX7	B	CSK2B_HUMAN	Casein kinase II subunit beta	24926	103	2
CBX7	A / B	PCGF2_HUMAN	Polycomb group RING finger protein 2	37764	75 / 47	4 / 2
CBX7	A / B	PCGF6_HUMAN	Polycomb group RING finger protein 6	39023	95 / 56	2 / 3
CBX7	A / B	PHC2_HUMAN	Polyhomeotic-like protein 2	90657	111 / 66	3 / 2
CBX7	A	RAD17_HUMAN	Cell cycle checkpoint protein RAD17	77055	45	2
CBX7	A / B	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	384 / 334	8 / 7
CBX7	A / B	RING2_HUMAN	E3 ubiquitin-protein ligase RING2	37632	627 / 571	15 / 12
CBX7	A	SPEC1_HUMAN	Sperm antigen with calponin homology and coiled-coil dom	118512	48	3
CBX7	B	SRBD1_HUMAN	S1 RNA-binding domain-containing protein 1	111705	39	3
CBX7	A	TPR_HUMAN	Nucleoprotein TPR	267293	45	3
CBX7	B	UBE4A_HUMAN	Ubiquitin conjugation factor E4 A	122561	41	2
CBX7	B	UBP20_HUMAN	Ubiquitin carboxyl-terminal hydrolase 20	101938	46	1

	purif	accession	description	masse	score	peptides
CBX8	B	BCLF1_HUMAN	Bcl-2-associated transcription factor 1	106059	84	2
CBX8	A / B	BMI1_HUMAN	Polycomb complex protein BMI-1	36925	154 / 122	9 / 13
CBX8	A / B	CBX8_HUMAN	Chromobox protein homolog 8	43369	866 / 803	217 / 143
CBX8	A / B	CSK21_HUMAN	Casein kinase II subunit alpha	45115	63 / 236	1 / 6
CBX8	A / B	CSK2B_HUMAN	Casein kinase II subunit beta	24926	76 / 111	2 / 2
CBX8	A / B	DBPA_HUMAN	DNA-binding protein A	40066	213 / 218	8 / 4
CBX8	B	DGKZ_HUMAN	Diacylglycerol kinase zeta	124044	67	5
CBX8	A	DIMT1_HUMAN	Probable dimethyladenosine transferase	35214	253	4
CBX8	B	DNL3_HUMAN	DNA ligase 3	102625	44	6
CBX8	B	EP300_HUMAN	Histone acetyltransferase p300	263973	43	2
CBX8	B	FBRL_HUMAN	rRNA 2'-O-methyltransferase fibrillar	33763	94	2
CBX8	A	GBLP_HUMAN	Guanine nucleotide-binding protein subunit beta-2-like 1	35055	110	2
CBX8	A / B	H2A1B_HUMAN	Histone H2A type 1-B	14127	62 / 42	2 / 1
CBX8	B	H2B1A_HUMAN	Histone H2B type 1-A	14159	40	2

CBX8	A / B	H2B1B_HUMAN	Histone H2B type 1-B	13942	116 / 85	2 / 2
CBX8	A / B	H4_HUMAN	Histone H4	11360	143 / 66	6 / 4
CBX8	A / B	HLF_HUMAN	Hepatic leukemia factor	33199	59 / 40	3 / 3
CBX8	A / B	IF2B1_HUMAN	Insulin-like growth factor 2 mRNA-binding protein 1	63417	365 / 169	6 / 6
CBX8	A / B	IF2B2_HUMAN	Insulin-like growth factor 2 mRNA-binding protein 2	61804	65 / 46	2 / 1
CBX8	B	IF2B3_HUMAN	Insulin-like growth factor 2 mRNA-binding protein 3	63681	49	2
CBX8	A / B	ILF2_HUMAN	Interleukin enhancer-binding factor 2	43035	381 / 65	7 / 2
CBX8	A / B	ILF3_HUMAN	Interleukin enhancer-binding factor 3	95279	146 / 76	2 / 2
CBX8	B	JIP4_HUMAN	C-jun-amino-terminal kinase-interacting protein 4	146115	48	2
CBX8	B	K0020_HUMAN	Pumilio domain-containing protein KIAA0020	73538	47	1
CBX8	A / B	KRI1_HUMAN	Protein KRI1 homolog	83201	241 / 100	4 / 1
CBX8	B	KRR1_HUMAN	KRR1 small subunit processome component homolog	43638	91	2
CBX8	B	LTV1_HUMAN	Protein LTV1 homolog	54821	127	3
CBX8	A	MK67I_HUMAN	MKI67 FHA domain-interacting nucleolar phosphoprotein	34201	67	1
CBX8	A / B	MOV10_HUMAN	Putative helicase MOV-10	113599	364 / 268	9 / 6
CBX8	B	MPPH1_HUMAN	M-phase phosphoprotein 1	210553	42	3
CBX8	B	NBN_HUMAN	Nibrin	84906	48	2
CBX8	B	NCOAT_HUMAN	Bifunctional protein NCOAT	102849	59	4
CBX8	B	NFIL3_HUMAN	Nuclear factor interleukin-3-regulated protein	51472	53	2
CBX8	A	NOG1_HUMAN	Nucleolar GTP-binding protein 1	73918	71	2
CBX8	A / B	NOL1_HUMAN	Putative RNA methyltransferase NOL1	89247	105 / 62	2 / 1
CBX8	B	NPM_HUMAN	Nucleophosmin	32555	165	4
CBX8	A	NSBP1_HUMAN	Nucleosome-binding protein 1	31525	47	2
CBX8	B	PCGF1_HUMAN	Polycomb group RING finger protein 1	29202	54	3
CBX8	A / B	PCGF2_HUMAN	Polycomb group RING finger protein 2	37764	110 / 59	4 / 2
CBX8	B	PCGF3_HUMAN	Polycomb group RING finger protein 3	28097	58	1
CBX8	B	PCGF5_HUMAN	Polycomb group RING finger protein 5	29694	50	2
CBX8	A / B	PCGF6_HUMAN	Polycomb group RING finger protein 6	39023	136 / 42	3 / 1
CBX8	A / B	PHC2_HUMAN	Polyhomeotic-like protein 2	90657	192 / 53	3 / 1
CBX8	A	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	124	3
CBX8	A / B	RING2_HUMAN	E3 ubiquitin-protein ligase RING2	37632	395 / 346	11 / 11
CBX8	A / B	RRP5_HUMAN	Protein RRP5 homolog	208570	272 / 61	11 / 1
CBX8	A	SSF1_HUMAN	Suppressor of SWI4 1 homolog	53161	66	2
CBX8	A	STAU1_HUMAN	Double-stranded RNA-binding protein Staufen homolog 1	63143	113	2
CBX8	A / B	THOC4_HUMAN	THO complex subunit 4	26872	98 / 65	2 / 1
CBX8	A / B	TSR1_HUMAN	Pre-rRNA-processing protein TSR1 homolog	91752	100 / 92	2 / 2
CBX8	B	UBR1_HUMAN	E3 ubiquitin-protein ligase UBR1	200080	62	3
CBX8	A / B	YBOX1_HUMAN	Nuclease-sensitive element-binding protein 1	35903	339 / 204	11 / 7
CBX8	A	YTDC2_HUMAN	YTH domain-containing protein 2	74853	98	2
CBX8	B	ZCHC3_HUMAN	Zinc finger CCHC domain-containing protein 3	43591	86	2
CBX8	A / B	ZFR_HUMAN	Zinc finger RNA-binding protein	116939	115 / 39	2 / 1
CBX8	B	ZN717_HUMAN	Zinc finger protein 717	105183	46	3
CBX8	A	ZN828_HUMAN	Zinc finger protein 828	89099	51	3

Tableau 4 – Protéines co-éluées avec les CBX Polycomb-TAP. Orange = appât, bleu = PRC1, rose = retrouvé avec 4 appâts différents, vert = retrouvés avec 2 appâts différents (3 pour MOV10), jaune = protéine exclusive de l'appât, identifiée avec une forte confiance (dans les 2 purifications et/ou score Mascot élevé)

5.2.1 Les protéines du PRC1

Les protéines décrites comme appartenant au PRC1 qui ont été identifiées sont résumées dans le Tableau 5, les couvertures de séquence sont présentées en Annexe 1 dans les Figure S1 à S5. Nous remarquons que CBX6 et CBX2 sont les 2 protéines TAP qui interagissent avec le moins de membre du PRC1 dans notre étude. Cependant, toutes les protéines permettent la purification de RNF2 et PHC2, qui semblent être des protéines possédant une forte affinité pour les 5 CBX. Dans les cellules de mammifères, la protéine PHC2 est exprimée sous une forme longue (90kDa) et une forme courte (36kDa) (Yamaki M, 2002, Gene). Dans nos purifications elle a exclusivement été identifiée dans le

gel à hauteur de 40kDa environ, l'isoforme courte de PHC2 est donc majoritairement (ou exclusivement) présent dans les PRC1. Malgré l'intensité de la coloration sur gel, la purification avec CBX6-TAP ne permet pas l'identification de membres du PRC1 avec une grande confiance, en effet, au maximum 3 peptides sont identifiés pour SCML2, une protéine Polycomb non retrouvée avec les autres CBX-TAP. En comparant ces données avec les identifications de CBX2 dont la coloration était également faible, nous observons la présence d'au moins 4 peptides par protéine identifiée. Les identifications associées à CBX6 sont relativement atypiques par rapport aux autres CBX, la protéine SCML2 n'est pas retrouvée dans les autres purifications, aussi, 2 peptides correspondant à CBX2 sont identifiés. En analysant plus en détails ces peptides, nous constatons qu'un des deux : LEYLVK, correspond également à la séquence de CBX6. De plus, aucun peptide correspondant à CBX2 n'est identifié dans les deux purifications à partir de CBX6. Aucune autre CBX que celle qui a servi d'effecteur n'est identifiée dans les autres purifications. La protéine BMI1 est également retrouvée avec toutes les CBX, à l'exception de CBX6. RING1, quant à elle, est présente partout sauf avec CBX2. Les protéines CBX4, CBX7 et CBX8 sont les protéines avec lesquelles le plus grand nombre de protéines PRC1 sont co-éluées : BMI1, PCGF2 (MEL-18), PCGF6, PHC2, RING1 et RNF2. Certaines protéines ne sont identifiées que dans une seule des 10 purifications TAP réalisées, c'est le cas de PHC1 et PHC3 avec CBX4, PCGF1, PCGF3 et PCGF5 avec CBX8.

	CBX2-A	CBX2-B	CBX4-A	CBX4-B	CBX6-A	CBX6-B	CBX7-A	CBX7-B	CBX8-A	CBX8-B
appât										
	43/583	26/379	65/266	137/641	72/375	96/572	10/232	1/55	217/866	143/803
membres du PRC1										
PCGF1										3/54
PCGF2			12/143	3/104			2/47	4/75	4/110	2/59
PCGF3										1/58
BMI1		4/201	18/337	7/212			4/199	3/100	9/154	13/122
PCGF5										2/50
PCGF6			1/63	1/78			2/95	3/56	3/136	1/42
PHC1			1/40							
PHC2		7/350	19/628	4/141		1/61	2/66	3/111	3/192	1/53
PHC3			4/103							
RING1			12/345	8/379	1/73		8/384	7/334	3/124	
RNF2	7/281	8/429	19/598	7/284		2/91	15/627	12/571	11/346	11/395
SCML2						3/40				
CBX2					2/47					
Nouveau partenaire ?										
CSK21	2/136	2/66	3/169					3/255	1/63	6/236
CSK22		1/94	1/46							
CSK2B		3/136	2/107	1/59				2/103	2/76	2/111

Tableau 5 – Résumé des membres du PRC1 identifiés dans les purifications TAP (nombre de peptides/score Mascot)

5.2.2 La caséine kinase 2 est-elle un nouveau membre du PRC1 ?

La question mérite d'être posée compte tenu de sa présence avec toutes les CBX en dehors de CBX6. Le nombre de peptides et le score élevé des deux sous-unités de cette kinase suggèrent que cette dernière pourrait jouer un rôle avec les protéines associées au PRC1.

Même si certaines protéines ont été retrouvées avec 2 (YD007, YBOX1, NPM, KRI1...), ou 3 (MOV10) protéines CBX différentes (protéines surlignées en vert dans le Tableau 4) : seules les

peptides correspondant à la caséine kinase 2 sont omniprésents dans la quasi-totalité des purifications.

5.2.3 Partenaires spécifiques à chaque protéine CBX

La comparaison des listes des protéines identifiées avec chacune des protéines CBX nous a permis de suggérer la présence de partenaires spécifiques : PSME3 avec CBX2, 14-3-3 avec CBX4 et RPP5 + ZFR avec CBX8. Compte tenu de leurs présences dans les 2 purifications indépendantes par CBX, du nombre de peptides et du score des ions, nous pouvons conclure que ces protéines sont fortement associées à ces protéines CBX spécifiques. Aucune protéine ne correspond aux critères fixés pour CBX6 et CBX7, même si UBE2O a été identifié avec une forte confiance dans une des deux purifications CBX6. Pour finir, dans la purification CBX4 qui a été réalisée avec l'isoforme courte, nous avons identifié 3 peptides qui appartiennent à l'isoforme longue (Figure 28). Nous en concluons que la protéine CBX4 homodimérise dans les cellules HeLa.

MELPAVGEHVFAVESIEKKRIRKGRVEYLVKWRGWSPKYNTWEPEENILDPRLLIAFONRERQEQLM
GYRKRGPKPKPLVVQVPTFARRSNVLTGLQDSSTDNRAKLDLGAQGKGQGHQYELNSKKHHQYQPHS
 KEGKPRPPGKSGKYYYQLNSKHHHPYQDPDKMYDLQYQGGHKEAPSPTCPDLGAKSHPPDKWAQGAG
 AKGYLGAVKPLAGAAGAPGKGSEKGPPNGMMPAPKEAVTGNGIGGKMKIVKNKNKNGRIVIVMSKYM
ENGMQAVKIKSGEVAEGEARSPSHKKRAADERHPPADRTFKKAAGAEKKVEAPPKRREEEVSGVSD
 PQQDAGSRKLSPTKEAFGEQPLQLTTKPDLLAWDPARNTHTPPSHHPHPHHHHHHHHHHHHAVGL
 NLSHVRKRCLSETHGEREPCKKRLTARSISTPTCLGGSPAERPADLPPAAALRQPEVILLDSDLDE
 PIDLRSVKSRSEAGEPPSSLQVKPETPASAAVAVAAAAAPTPTTAEKPPAEAQDEPAESLSEFKPFFG
 NIIITDVTANCLTVTFKEYVTV

Figure 29 - Peptides de CBX4 identifiés par MS-MS au cours de la purification TAP à partir de l'isoforme courte de CBX4. En gris la séquence manquante dans CBX4-TAP (isoforme courte). En bleu les 3 peptides identifiés appartenant à l'isoforme longue de CBX4.

Nous pouvons mentionner les présences de PARP1 et JHD2A (KDM3A) qui peuvent respectivement poly-ADP ribosyler les histones H2B/H4 ou déméthyliser les lysines 9 mono- et di-méthylées de l'histone H3. Nous avons vu en introduction que ces activités conduisaient à une dérégulation transcriptionnelle, une activité surprenante pour des protéines supposées être associées à CBX2, mais qui pourrait expliquer pourquoi cette protéine n'a pas ou peu d'activité répressive sur les gènes rapporteurs testés.

5.3 Discussion

D'une manière générale, les protéines CBX semblent avoir une forte affinité pour les protéines BMI1 (sauf CBX6), PCGF2 et PCGF6 (sauf CBX2 et CBX6), PHC2, RING1 (sauf CBX2) et RNF2 (la sous-unité catalytique du complexe PRC1). Dans une moindre mesure, les autres isoformes de PHC et PCGF sont retrouvées ponctuellement. C'est l'isoforme courte de PHC2 qui est préférentiellement recrutée dans les complexes PRC1. Compte tenu que les isoformes courte et longue possèdent toutes les deux le domaine SAM carboxy-terminal, ce dernier peut potentiellement être impliqué pour l'interaction avec le PRC1. En revanche, nous pouvons exclure ce rôle pour toute la région aminotermine de l'isoforme longue qui est absente dans la courte. Des études complémentaires permettraient de mettre en évidence le rôle joué par les PHC (et le domaine SAM en particulier) au sein des PRC1.

Le coeur des complexes PRC1 associés aux protéines CBX est donc conservé entre ces différentes protéines, en revanche les protéines CBX semblent être mutuellement exclusives car les autres isoformes ne sont pas identifiées avec une CBX donnée. Ces résultats renforcent l'idée que plusieurs PRC1, de coeur similaire, mais pouvant varier en composition, co-existent dans les cellules humaines. Cependant comme des peptides correspondants à l'isoforme longue de CBX4 ont été identifiés dans la purification à partir de l'isoforme courte, nous suggérons que CBX4 homodimérise dans la cellule. Dans notre étude, la caséine kinase apparaît comme étant un nouveau partenaire des protéines associées au PRC1, en effet, comme cela avait été suggéré par la purification de CBX8 fusionnée à l'étiquette FLAG à partir de cellules HEK293 (Dietrich N, 2007, EMBO J), nous identifions avec une forte confiance les différentes sous-unités de cette protéine tandis que la purification avec FLAG n'avait identifié que la sous-unité CSK21. La caséine kinase 2 (CK2) est un tétramère constitué de deux sous-unités catalytiques α ou α' et de deux sous-unités régulatrices β (Litchfield DW, 2003, Biochem J). Cette enzyme est ubiquitaire chez les eucaryotes, elle interagit et phosphoryle une multitude de protéines différentes. Elle est impliquée dans la transformation cellulaire et le développement des tumeurs (Olsten ME, 2004, Biochem Cell Biol). Elle reconnaît une séquence consensus contenant une sérine ou une thréonine dans un environnement acide, tel que S/T-X-X-X-D/E. La CK2 est très conservée, impliquée dans la prolifération cellulaire, la différenciation et l'apoptose, de plus elle est indispensable à la survie des levures (Glover CV 3rd, 1998, Prog Nucleic Acid Research Mol Biol / Faust RA, 2000, Head Neck / Wang H, 2001, Mol Cell Biochem). Des études sur des souris dont les gènes CK2 ont été invalidés démontrent une létalité embryonnaire (sous-unité β), ou des souris viables dont la spermatogenèse est altérée (sous-unité α') (Xu X, 1999, Nat Genet / Buchou T, 2003, Mol Cell Biol / Escalier D, 2003, Mol Reprod Dev), aucune publication concernant la sous-unité α n'a pour le moment été publiée. Ces informations peuvent suggérer un lien entre le PRC1 et les protéines composant la CK2. Enfin, chez la levure, CK2 est responsable de la phosphorylation de la sérine 1 de l'histone H4 lors de dommages à l'ADN (Cheung WL, 2005, Curr Biol).

L'activité du PRC1 peut-être régulée suite à la phosphorylation d'une de ses sous-unités. La phosphorylation de BMI1 par la MAPKAP kinase 3pk entraîne la dissociation du complexe PRC1 de la chromatine (Voncken JW, 2005, JBC). La phosphorylation de PCGF2 (MEL-18) augmente l'activité catalytique du PRC1 en stimulant la reconnaissance du nucléosome et l'ubiquitinylation de la lysine 119 de l'histone H2A (Elderkin S, 2007, Mol Cell). La ou les kinases responsables de la multiple phosphorylation de MEL-18 n'ont pas été identifiées, mais certains résidus (S110, S132 et S254) sont situés dans un contexte favorable à une phosphorylation par CK2. Cette dernière pourrait donc réguler positivement l'activité du PRC1. Pour finir, CK2 a été identifiée comme étant responsable de la phosphorylation de ESC (sous-unité du PRC2), favorisant la dimérisation de cette dernière (Tie F, 2005, BBRC).

En ce qui concerne les protéines qui ne semblent pas avoir un rôle au niveau du PRC1, qui ont à nos yeux un peu moins d'importance que CK2 par exemple : la protéine PSME3 appartient à une famille de sous-unités d'activateurs du protéasome, bien qu'elle ait été localisée au niveau du noyau et du cytoplasme, aucun lien direct avec la chromatine n'a été rapporté. En ce qui concerne MOV-10, elle

est putativement une hélicase d'ARN qui interagit avec DICER et les protéines ribosomales 60S, donc peut-être un contaminant car identifiée dans les 2 purifications qui ont fourni le plus grand nombre de partenaires protéiques (CBX6 et CBX8). RRP5 a également été décrite comme pouvant jouer un rôle dans la synthèse des ARN. Les protéines ZFR et KRI-1 ont été peu étudiées et aucune fonction claire ne leur a été attribuée.

6 Validation des interactions avec les membres du PRC1

6.1 Co-immunoprécipitations (CoIP) : interactions indirectes

6.1.1 Principe

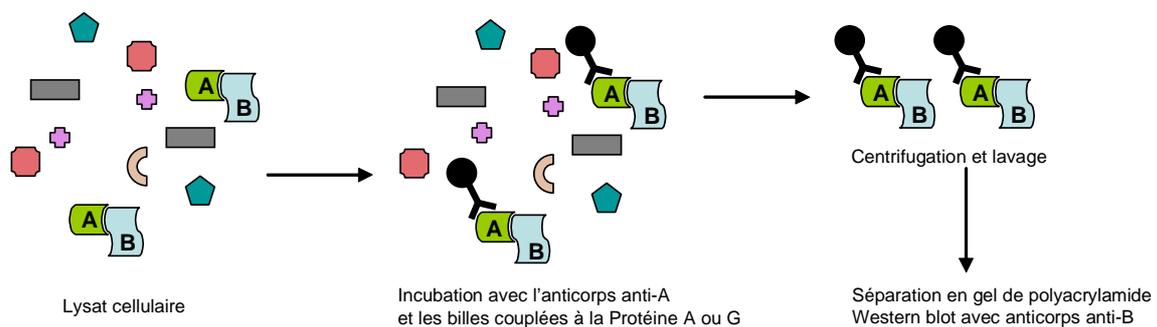


Figure 30 - Principe de la co-immunoprécipitation (co-IP)

Le principe de la co-IP permet de mettre en évidence l'interaction entre une protéine A et une protéine B, sans pour autant donner d'informations sur la nature du lien : direct ou indirect (intervention d'autres protéines). A partir d'un extrait protéique dans lesquels A et B sont exprimées, il s'agit d'immunoprécipiter la protéine A en incubant l'extrait avec un anticorps dirigé contre A (ou une étiquette si A est une protéine de fusion). Des billes de sépharose couplées aux protéines A et G permettent d'immobiliser l'anticorps et les protéines immunoprécipitées, après plusieurs lavages, les billes sont éluées dans un tampon dénaturant. L'éluat est analysé par SDS-PAGE et western blot avec un anticorps dirigé contre la protéine B (ou une étiquette si B est une protéine de fusion). La révélation d'une bande à la taille de B indique que A et B sont bien partenaires dans la cellule, B a été co-immunoprécipitée avec A.

Inconvénients

Certaines protéines se fixent de manière non spécifique sur les billes A/G, ce qui abolit leur analyse par co-IP. Certains épitopes sont « cachés » lors d'une interaction, la protéine A sous forme de complexe ne peut donc pas être immunoprécipitée.

6.1.2 Résultats

Nous avons effectué un nombre restreint de co-immunoprécipitations compte tenu de la pertinence des identifications obtenues avec les membres du PRC1 dans la partie I.

Nous avons cloné les protéines RNF2, PCGF6 et BMI dans des plasmides contenant l'étiquette GFP d'une part, et la protéine CBX8 dans un plasmide contenant l'étiquette HA d'autre part. Nous avons

immunoprécipité la protéine CBX8 exogène à l'aide d'un anticorps dirigé contre l'étiquette HA, puis nous avons révélé avec un anticorps reconnaissant la GFP (Figure 30).

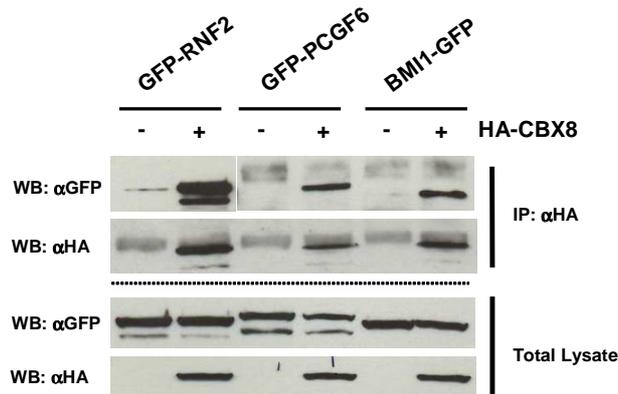


Figure 31 - Co-IP des protéines RNF2, PCGF6 et BMI1 avec CBX8. Les protéines étiquetées avec GFP ou HA sont co-exprimées dans des cellules HeLa. A partir de l'extrait cellulaire, immunoprécipitation avec un anticorps anti-HA. Les protéines retenues sont analysées par western blot avec un anticorps anti-GFP

Les 3 protéines testées ont été co-immunoprécipitées avec CBX8 alors qu'elles ne sont pas retenues sur les billes anti-HA quand elles sont sur-exprimées en absence de CBX8-HA, validant à la fois la méthode de co-IP et les interactions identifiées par la purification TAP.

Nous avons souhaité valider l'interaction de RNF2 avec toutes les protéines CBX, nous avons réalisé l'expérience de co-IP avec les mêmes constructions (GFP et HA). Les résultats obtenus (Figure 31) démontrent que RNF2 est co-précipitée avec toutes les CBX. Cependant, la présence de CBX2 dans l'échantillon immunoprécipité n'a pas pu être démontrée compte tenu de la révélation des immunoglobulines avec l'anticorps secondaire, qui ont le même poids moléculaire que CBX2-HA.

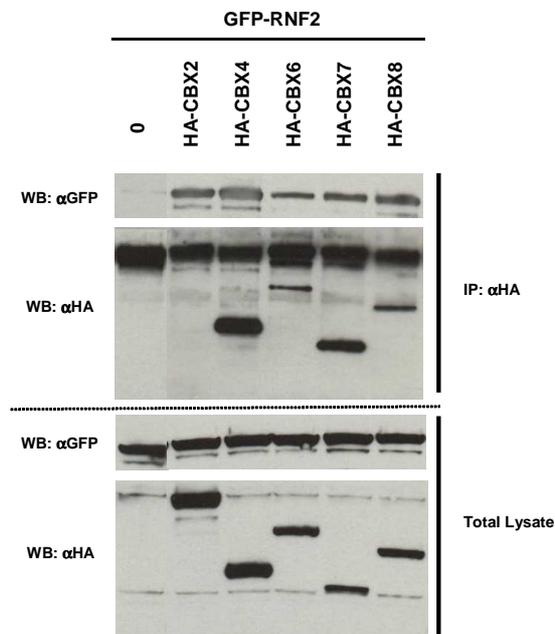


Figure 32 - Co-IP de la protéine RNF2 avec toutes les protéines CBX Polycomb. Les protéines étiquetées avec GFP ou HA sont co-exprimées dans des cellules HeLa. A partir de l'extrait cellulaire, immunoprécipitation avec un anticorps anti-HA. Les protéines retenues sont analysées par western blot avec un anticorps anti-GFP

Nous avons alors réalisé des constructions pour exprimer les protéines CBX fusionnées à l'étiquette flag et répété la même expérience (Figure 32), cette dernière nous a permis de valider l'interaction spécifique de RNF2 avec toutes les isoformes CBX Polycomb car dans un extrait dans lequel aucune protéine CBX-flag n'est exprimée, elle n'est pas retenue quand elle est mise en présence de billes couplées à l'anticorps anti-flag.

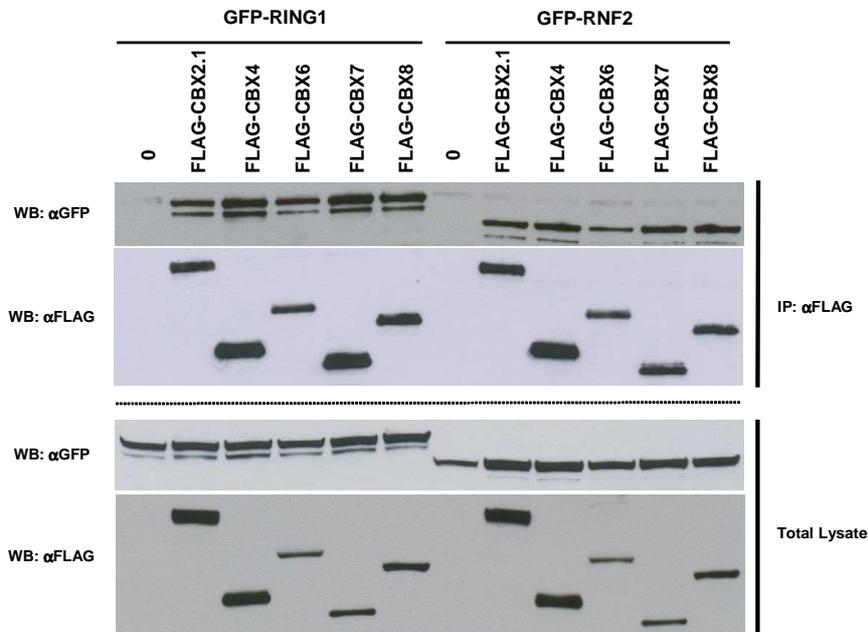


Figure 33 - Co-IP des protéines RING1 et RNF2 avec toutes les CBX Polycomb

6.2 GST-pulldown : interactions directes

6.2.1 Principe

Il existe une approche biochimique lorsqu'on veut tester la liaison entre la protéine X d'intérêt et une autre protéine Y, appelée GST-pulldown. La protéine X est fusionnée à l'étiquette GST (glutathione S transférase), la protéine Y est fusionnée à une autre étiquette pour faciliter sa détection par western blot, ou traduite *in vitro* en présence d'acides aminés marqués radioactivement (révélation par autoradiographie) ou non taguée si on dispose d'un anticorps efficace qui reconnaît cette protéine.

Le principe est de produire en grande quantité la protéine X-GST dans un système procaryote (ou eucaryote), de purifier cette protéine recombinante en l'immobilisant sur des billes de sépharose couplées à la glutathione (pour laquelle la GST a une forte affinité). Après plusieurs lavages pour éliminer d'éventuels contaminants, la protéine Y est ajoutée sur les billes. Après une période d'incubation, les billes sont lavées pour retirer les partenaires non spécifiques, puis les protéines immobilisées sur la résine sont éluées par ajout de glutathione réduite dans la colonne (qui va entrer en compétition avec la glutathione des billes et libérer les protéines GST). L'éluat est déposé dans un gel d'acrylamide puis analysé par la méthode adéquate pour la révélation de la protéine Y.

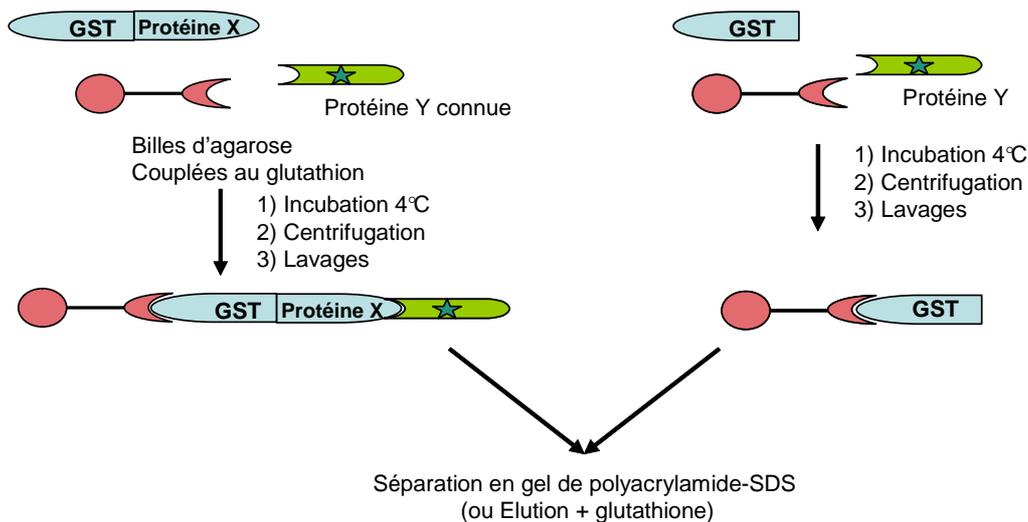


Figure 34 – Principe du GST-pulldown pour valider une interaction *in vitro*

Avantage : simple et rapide, permet de tester des interactions directes

Inconvénients : la taille de l'étiquette GST peut entraîner un encombrement stérique qui peut gêner la fixation des partenaires. Cette technique révèle de nombreux faux positifs.

6.2.2 Validation d'interaction directe entre les protéines PCGF et PHC avec les CBX

Afin d'étudier les interactions directes entre les protéines CBX avec les protéines des familles PHC et PCGF, nous avons entrepris des expériences de GST-pulldown. Dans un premier temps nous avons produit des protéines recombinantes correspondant aux protéines CBX dans un système procaryote (*E. coli* BL21) (Figure 34), après avoir purifié les protéines recombinantes, nous les avons immobilisées sur des billes GST. D'autre part nous avons produit *in vitro* les protéines PCGF et PHC fusionnées à l'étiquette TAP à l'aide de lysats de réticulocytes de lapin.

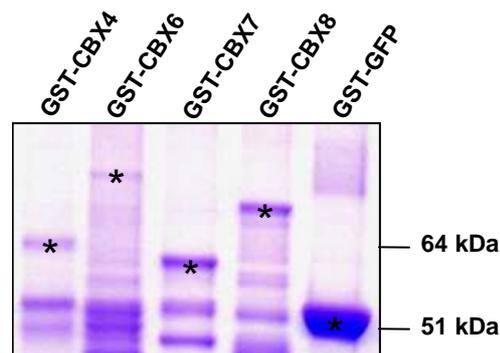


Figure 35 - Production des protéines recombinantes en système procaryote. L'étoile indique la protéine d'intérêt

Nous avons mis en présence chacune des combinaisons CBX/PCGF et CBX/PHC, après plusieurs lavages nous avons élué les protéines retenues avec de la glutathione réduite puis nous avons réalisé un western blot sur l'éluat. Une révélation avec l'anticorps PAP permet de valider ou non l'interaction potentielle directe entre les 2 protéines testées. Une partie des résultats obtenus est présentée en Figure 35. Nous n'avons pas réussi à produire une protéine complète CBX2 dans notre système de

production *in vitro*, malgré plusieurs tests (variation de la température de la culture, concentration d'IPTG, durée d'induction...), cependant les résultats obtenus avec les 4 autres CBX et le contrôle GST nous permettent de dresser quelques tendances.

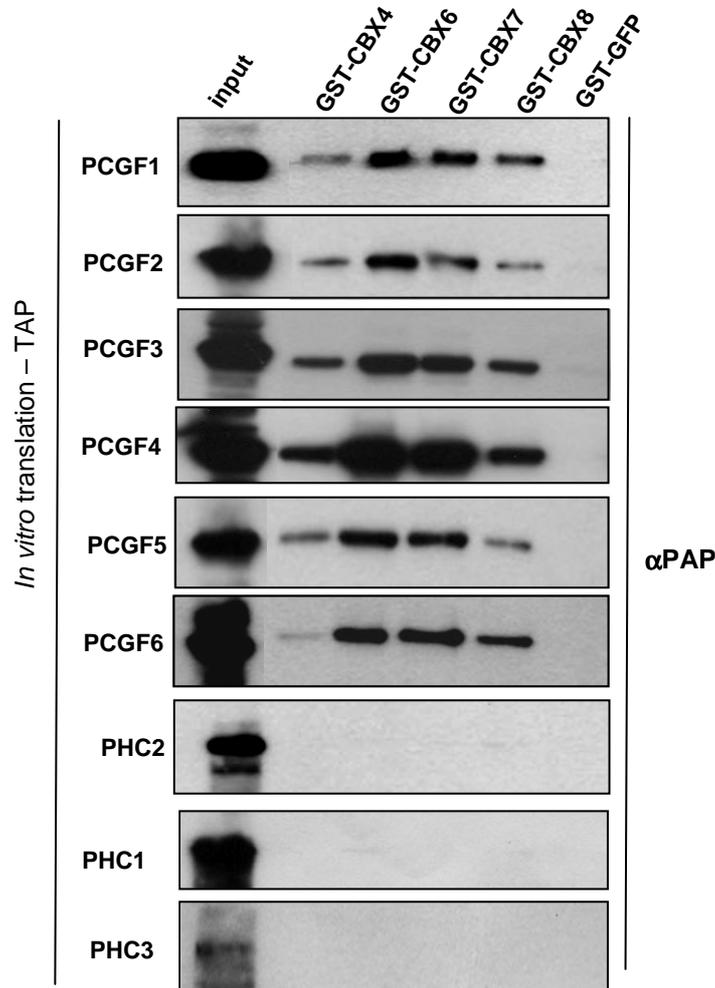


Figure 36 - GST-pulldown entre les CBX et les membres des familles PCGF et PHC

D'une manière générale nous observons que l'ensemble des protéines CBX (hormis CBX2 qui n'a pas été testée) sont en mesure d'interagir directement *in vitro* avec les 6 membres de la famille PCGF. En revanche, elles ne sont pas capables d'interagir directement avec les protéines de la famille PHC. Le contrôle négatif est représenté par la protéine GFP fusionnée à l'étiquette GST, pour laquelle aucune interaction n'est identifiable, démontrant la spécificité d'interaction directe entre les CBX et les PCGF.

6.2.3 Discussion

Des expériences simples comme la co-immunoprécipitation ou le GST-pulldown nous ont permis de valider par d'autres approches biochimiques les identifications obtenues par purification TAP. La co-immunoprécipitation révèle des interactions directes et indirectes, tandis que le pulldown permet de vérifier si une interaction est directe ou non, à condition qu'elle ne soit pas dépendante d'une modification post-traductionnelle particulière. Nous avons réussi à démontrer que les protéines RING étaient des partenaires des CBX dans la cellule, aussi, nous avons pu mettre en évidence que toutes

les protéines de la famille Psc pouvaient interagir directement avec des protéines CBX recombinantes, ainsi la spécificité des identifications obtenues par la purification TAP démontre qu'au sein de la cellule, il existe des régulations qui permettent à ces protéines de se lier à une CBX particulière plutôt qu'aux autres (par exemple PCGF1 avec CBX8), alors que d'autres (PCGF2, 4 et 6) sont retrouvées dans plusieurs purifications avec CBX4, 7 et 8. D'une manière intéressante, certaines protéines capables de s'associer dans le test *in vitro* ne semblent pas l'être *in vivo* (par rapport aux résultats obtenus par TAP). Cela suggère que dans la cellule, une partie des PCGF est préférentiellement associée avec certaines CBX plutôt qu'avec les autres. Nous n'avons pas réalisé de tests de compétition pour vérifier si c'est l'affinité entre ces protéines qui est différente ou si c'est le contexte physico-chimique au sein du complexe qui contrôle l'association ou non avec les différentes CBX.

7 CBX4 et les 14-3-3

Compte tenu de l'abondance, ainsi que la spécificité des protéines 14-3-3 retrouvées exclusivement dans la purification TAP avec CBX4 en comparaison avec les autres membres de la famille, nous avons souhaité caractériser cette potentielle interaction.

Les protéines 14-3-3

Elles constituent une famille de petites protéines acides (28-33kDa) ubiquitaires dans les tous les organismes eucaryotes. Il existe 7 isoformes chez les mammifères qui sont capables d'homo- ou d'hétéro-dimériser, suggérant que des combinaisons différentes puissent avoir des rôles spécifiques (Jones DH, 1995, FEBS Lett / Chaudhri M, 2003, BBRC). Comme nous l'avons brièvement abordé dans le premier chapitre de ce manuscrit, ces protéines sont capables de reconnaître et de se fixer sur une sérine ou une thréonine phosphorylée.

Ces protéines sont impliquées dans la signalisation cellulaire, la régulation de la progression du cycle cellulaire, le trafic intracellulaire, la structure du cytosquelette et la transcription. En général les protéines qui interagissent avec les 14-3-3 possèdent une préférence pour une isoforme particulière. La phosphorylation de la protéine cible peut influencer l'interaction, mais la phosphorylation des 14-3-3 elles-mêmes peut également jouer ce rôle. Plus de 200 protéines ont été rapportées comme partenaires des 14-3-3 et la liste s'agrandit de jour en jour (Pozuelo Rubio M, 2004, Biochem J / Jin J, 2004, Curr Biol / Meek SE, 2004, JBC / Benzinger A, 2005, MCP), aussi elles peuvent se fixer sur de l'ADN cruciforme (Todd A, 1998, Biochemistry). Le motif d'interaction a été classé en deux sous-types majoritaires, le mode 1 : RSXpSXP et le mode 2 : RXY/FXpSXP ou pS correspond à une phosphosérine (Yaffe MB, 1997, Cell). Cependant, certaines protéines ne possédant pas ces séquences sont capables d'interagir directement avec les 14-3-3, suggérant que d'autres séquences ou modes d'interaction existent. Ayant identifié toutes les isoformes, nous pensons que CBX4 n'interagit pas exclusivement avec une 14-3-3 particulière même si on retrouve majoritairement la forme ϵ .

7.1 Localisation d'un site de fixation mode 1 dans la séquence de CBX4

Une analyse bioinformatique des séquences en acides aminés des protéines CBX nous a permis d'identifier un motif putatif dans CBX4. L'alignement des séquences des CBX au niveau du site potentiel d'interaction avec 14-3-3 prouve que seule la protéine CBX4 possède un tel motif. De plus, le motif « RSISTP » est conservé dans les variants courts et longs de cette protéine (Figure 36). La sérine potentiellement phosphorylée correspond aux acides aminés 432 et 164 pour les isoformes respectives longue et courte. Ce motif est situé à l'extérieur du chromodomaine et de la Pc box, c'est pourquoi il n'est pas retrouvé dans les autres protéines CBX. Nous pensons donc que la protéine CBX4 est phosphorylée *in vivo* au niveau de la sérine 432 ou 164 pour interagir avec les 14-3-3. Nous avons réalisé une série d'expériences avec l'isoforme courte pour valider cette hypothèse.

CBX4_iso2 :

```
MELPAVGEHVF AVESIEKKRIRKGRVEYLVKWRGWSPKYNTWEPEENILDPRLLI AFQNRERQEQLMGYRKRGP KP  
KPLVVQVPTFARRSNVLTGLQDSSTDNR AKLDLGAQGKGQGHQYELNSKKHHHHAVGLNLSHVRKRCLSETHGERE  
PCKKRLTARRSISTPTC LGGSPA AERP ADLPPAAALPQPEVILLDSLDLDEPIDLRCVKTRSEAGEPPSSLQVKPETP  
ASAAVAVAAAAAPT TTA EKPPAEAQDEPAESLSEFKPFFGNI IITDVTANCLTVTFKEYVTV
```

CBX4_iso1 :

```
MELPAVGEHVF AVESIEKKRIRKGRVEYLVKWRGWSPKYNTWEPEENILDPRLLI AFQNRERQEQLMGYRKRGP KP  
KPLVVQVPTFARRSNVLTGLQDSSTDNR AKLDLGAQGKGQGHQYELNSKKHQYQPHSKEGKPRPPGKSGKYYYQLN  
SKKHHHPYQDPDKMYDLQYQGGHKEAPSPTCPDLGAKSHPPDKWAQGAGAKGYLGAVKPLAGAAGAPGKGSEKGP  
GMMPAPKEAVTGNIGGKMKIVKNKNNGRIVIVMSKYMENGMQAVKIKSGEVAEGEARSPSHKKRAADERHPPADR  
TFKKAAGAEKKVEAPPKRREEEVSGVSDPQPQDAGSRKLSPTKEAFGEQPLQLTTKPDLLAWDPARNTHPPSHHP  
HHHHHHHHHHHHHHAVGLNLSHVRKRCLSETHGEREPCKKRLTARRSISTPTC LGGSPA AERP ADLPPAAALRQPE  
VILLDSLDLDEPIDLRSVKSRSRSEAGEPPSSLQVKPETPASAAVAVAAAAAPT TTA EKPPAEAQDEPELSEFKPFFGN  
I IITDVTANCLTVTFKEYVTV
```

Figure 37 - Identification d'un site putatif mode1 de fixation des protéines 14-3-3 dans la séquence de CBX4

7.2 Etude de l'interaction de CBX4 avec les 14-3-3 et de la phosphorylation de CBX4

7.2.1 Identification d'un site de phosphorylation *in vivo*

La spectrométrie de masse est un outil puissant qui peut mettre en évidence la présence d'une phosphorylation au sein d'un peptide. Les peptides correspondant à CBX4 identifiés par la trappe ionique couvrent 42% de l'isoforme courte et 30% de l'isoforme longue. Bien que ces valeurs soient élevées, aucun peptide ne couvrant la séquence RSISTP n'a été identifié. En effet, le spectromètre de masse utilisé ne peut analyser que des ions (peptides) dont la masse/charge est comprise entre 500 et 2500 Da. Une étude théorique des peptides générés par digestion trypsique de CBX4 permet de constater que le peptide contenant le motif d'intérêt a une masse de 4620 Da, c'est pourquoi il n'a pas été détecté lors de l'analyse (Figure 37).

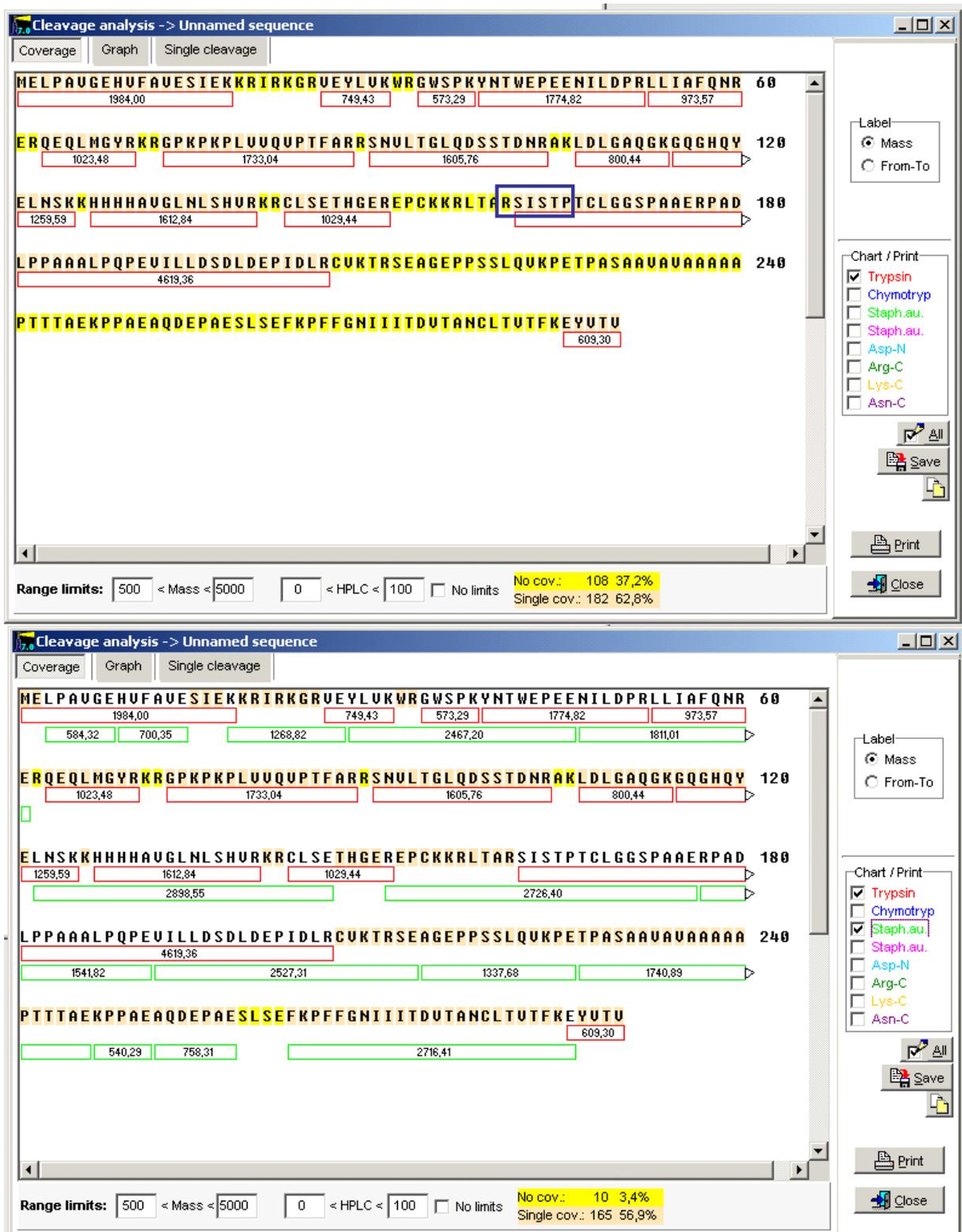


Figure 38 - Digestions théoriques de la protéine CBX4 (iso2) avec les protéases trypsine ou Glu-C (*Staph aureus*) (d'après le logiciel gpmaw)

Afin de remédier à ce problème, nous avons étudié les clivages théoriques de CBX4 par 7 protéases différentes (trypsine, chymotrypsine, Asp-N, Arg-C, Lys-C, Asn-C et Glu-C) qui possèdent des particularités de coupure spécifiques. L'endoprotéinase Glu-C (*Staph aur* en vert sur la Figure 37) est capable de cliver spécifiquement les protéines après les acides glutamiques et aspartiques. Si le

tampon de réaction ne contient pas de phosphate, elle ne digère qu'après les acides glutamiques. Une digestion de CBX4 par Glu-C permet d'obtenir un fragment contenant les acides aminés 151 à 176 (2726 Da) et une double digestion trypsine + Glu-C un peptide encore plus petit : aa162-176 (1525 Da).

Nous avons réalisé ces 2 digestions sur un fragment de gel de polyacrylamide contenant la protéine CBX4 purifiée par TAP. Les peptides générés ont été fournis au centre commun de mesures de spectrométrie de masse qui a réalisé l'analyse des échantillons. Malheureusement, aucune identification correspondant au peptide d'intérêt n'a pu être validée, ne permettant pas de statuer sur la présence ou non d'une phosphorylation sur la sérine 164 *in vivo*. Parallèlement, nous avons également testé un enrichissement en phosphopeptides sur une colonne hydroxyde de titanium (TiO₂) pour enrichir notre échantillon après digestion simple ou double, mais aucune identification par spectrométrie de masse n'a été possible sur ces échantillons.

Nous pouvons expliquer cette absence de résultats par, peut-être la digestion incomplète des échantillons avec Glu-C seule ou combinée à la trypsine. Aussi, le rendement des cônes TiO₂ est parfois faible, et la perte d'échantillon peut expliquer pourquoi nous n'avons pas été en mesure d'identifier des peptides après l'enrichissement en phosphopeptides. Enfin, comme la quantité d'échantillon à la fin de la purification TAP est limitée, nous n'avons pas eu l'opportunité de pouvoir répéter cette série d'expériences en privilégiant d'autres purifications TAP.

7.2.2 Validation de l'interaction directe de 14-3-3 avec CBX4

Compte tenu de l'identification d'un site de fixation de 14-3-3 dans CBX4, nous avons remplacé, par mutagenèse dirigée, la sérine 164 de l'isoforme courte de CBX4 par une alanine. Le choix de cette dernière est guidé par le fait que cet acide aminé possède une masse voisine de celle de la sérine (89,09 contre 105,09g/mol), mais est dépourvu de fonction hydroxyle (-OH). Elle ne peut donc pas être phosphorylée. Nous avons sur-exprimé la protéine CBX4-S164A ou CBX4 fusionnée à l'étiquette HA et l'isoforme ϵ de 14-3-3 fusionnée à la GFP dans des cellules HeLa en culture et réalisé une immunoprécipitation de ces protéines à l'aide d'un anticorps dirigé contre l'étiquette HA. La présence de 14-3-3 dans les protéines co-précipitées est analysée par western blot avec un anticorps anti-GFP (Figure 38). Tandis que la protéine non mutée est en mesure d'interagir avec 14-3-3, cette dernière n'est pas retenue dans l'échantillon avec CBX4 S164A. Nous en concluons que l'interaction entre CBX4 et 14-3-3 est dépendante de la sérine 164 de l'isoforme courte. Des résultats identiques ont été obtenus avec d'autres isoformes 14-3-3.

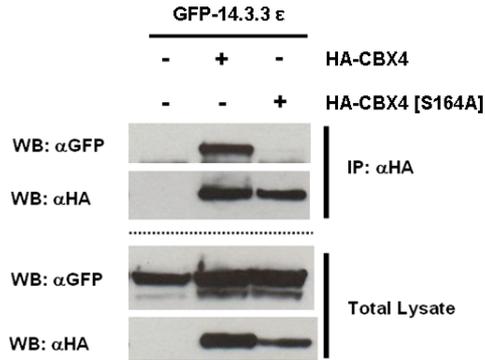


Figure 39 – Co-immunoprécipitations de CBX4 et CBX4-S164A avec 14-3-3 ε

Pour tenter de valider la nécessité d'une phosphorylation de cette sérine pour l'interaction, nous avons produit un certain nombre de protéines recombinantes fusionnées à la GST dans la bactérie *E. coli*. En théorie, les protéines produites dans les procaryotes ne sont pas ou peu modifiées post-traductionnellement. Les protéines CBX4, CBX4-S164A, CBX8, 14-3-3 (contrôle positif) et GFP (contrôle négatif) sont ainsi produites et purifiées sur des colonnes contenant du glutathion immobilisé sur des billes de silice. Les protéines recombinantes couplées aux billes GST permettent de tester les interactions. Les protéines produites et purifiées sont déposées dans un gel de polyacrylamide et révélées à l'aide du bleu de coomassie (Figure 39).

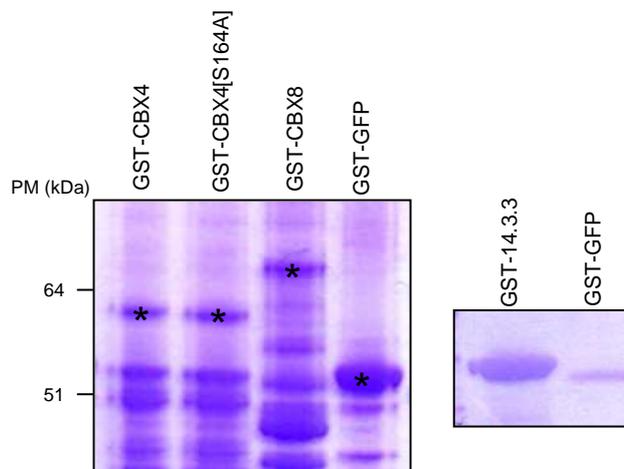


Figure 40 - Coloration au bleu de coomassie des protéines recombinantes produites en système procaryote (pour réaliser les GST pulldown)

Une quantification relative des protéines d'intérêt nous permet d'utiliser approximativement la même quantité de chacune de ces protéines pour réaliser des expériences d'interaction avec des extraits nucléaires issus de cellules HeLa. Après plusieurs lavages, les protéines retenues sur les billes sont éluées à l'aide de glutathion réduit et déposées sur un gel d'acrylamide pour réaliser un western blot avec un anticorps anti-14-3-3 (polyclonal, abcam ab9063) (Figure 40).

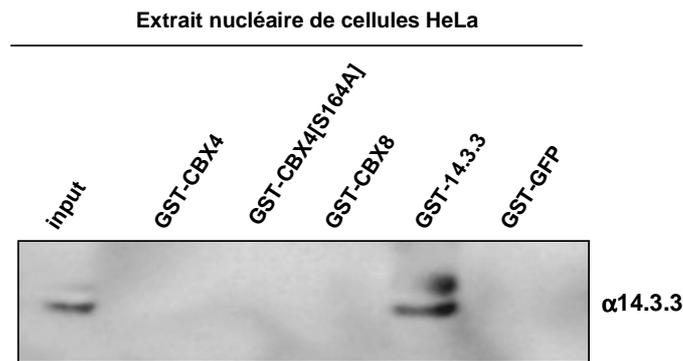


Figure 41 - GST pulldown pour tester l'interaction entre 14-3-3 et CBX4 non phosphorylé

Nous observons que seule la protéine 14-3-3 recombinante est capable d'interagir avec les protéines 14-3-3 contenues dans l'extrait nucléaire. Nous remarquons plusieurs bandes correspondant probablement aux différentes isoformes 14-3-3. Aucun signal n'est détecté dans les autres échantillons, indiquant que les 14-3-3 n'interagissent pas avec les protéines recombinantes CBX4, CBX4-S164A, CBX8 et GFP. Nous supposons que l'absence de phosphorylation dans la protéine recombinante CBX4 peut expliquer l'absence d'interaction avec 14-3-3.

Nous avons réalisé l'expérience dans l'ordre inverse, à savoir la production de protéines recombinantes 14-3-3 et GFP puis nous avons déposé des extraits nucléaires issus des lignes cellulaires CBX4-TAP et CBX8-TAP (Figure 41). Un western blot avec l'anticorps PAP (reconnaissant la protéine A de l'étiquette TAP) sur les extraits retenus permet de mettre en évidence l'interaction spécifique de CBX4-TAP avec 14-3-3-GST recombinant.

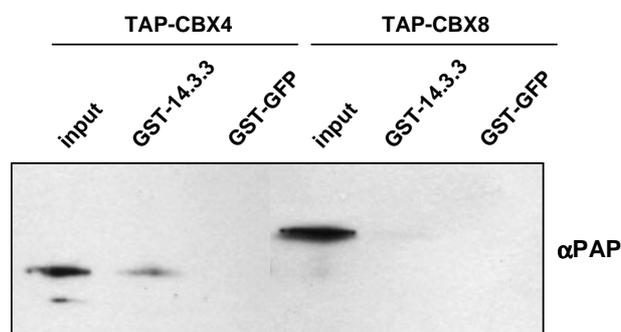


Figure 42 - GST pulldown pour tester l'interaction entre 14-3-3 recombinant et des extraits de protéines nucléaires des lignées TAP CBX4 et CBX8

Ces deux expériences démontrent que CBX4-GST ne retient pas 14-3-3 alors que 14-3-3-GST retient le CBX4-TAP d'un extrait nucléaire. Nous assumons que CBX4-TAP est potentiellement phosphorylé alors que CBX4-GST n'est pas modifié, ainsi nous en concluons que l'interaction entre CBX4 et 14-3-3 est dépendante de la sérine 164 phosphorylée.

Notons que nous avons tenté de répéter les co-immunoprécipitations en ajoutant une étape avec un traitement à l'antartique phosphatase afin de potentiellement retirer la phosphorylation de CBX4, mais nous n'avons pas perdu l'interaction avec 14-3-3. Nous en avons conclu que dans nos conditions d'expérience la phosphatase n'est pas efficace (nous n'avons pas de contrôle pour suivre le bon

déroulement de la réaction) ou alors l'interaction entre 14-3-3 et CBX4 ne permet pas à l'enzyme d'accéder à la sérine. Pour finir des études de co-localisation ont été envisagées, cependant les protéines 14-3-3 sont très largement exprimées aussi bien dans le noyau que dans le cytoplasme donc aucune information supplémentaire n'a pu être obtenue par microscopie à fluorescence. Aussi, il serait intéressant de répéter l'expérience avec une protéine TAP-CBX4-S164A, mais au moment où ce mémoire a été rédigé la lignée stable TAP n'avait pas été générée.

7.2.3 Importance de la phosphorylation de CBX4 sur la composition du PRC1 associé

Après avoir mis en évidence l'interaction spécifique entre CBX4 et 14-3-3, nous avons étudié la composition du PRC1 associé à CBX4-S164A (Figure 42). Pour cela, nous avons réalisé des expériences de co-immunoprécipitation avec RNF2 et PCGF6. Tandis que la protéine RNF2 est identiquement co-précipitée avec CBX4 et CBX4-S164A, la protéine PCGF6 est uniquement retenue avec CBX4 et non CBX4-S164A. L'interaction entre la partie carboxy-terminale de RNF2 et la Pc box des CBX a été récemment caractérisée par RMN (Neira JL, 2009, Biochemistry). La phosphorylation de la sérine 164 semble être responsable de l'association directe entre CBX4 et PCGF6, ou de son maintien au sein du complexe.

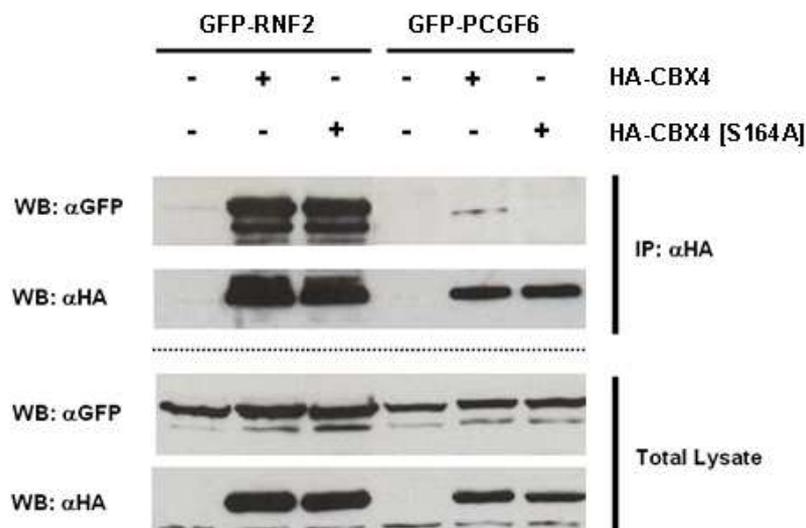


Figure 43 – Co-immunoprécipitation de RNF2 et PCGF6 avec CBX4 sauvage et mutée

7.2.4 Activité répressive de CBX4-S164A

Comme nous avons démontré que la phosphorylation de CBX4 était indispensable pour son interaction avec 14-3-3 et PCGF6, nous avons étudié l'activité répressive de CBX4-S164A dans le système rapporteur exogène décrit précédemment (Figure 43).

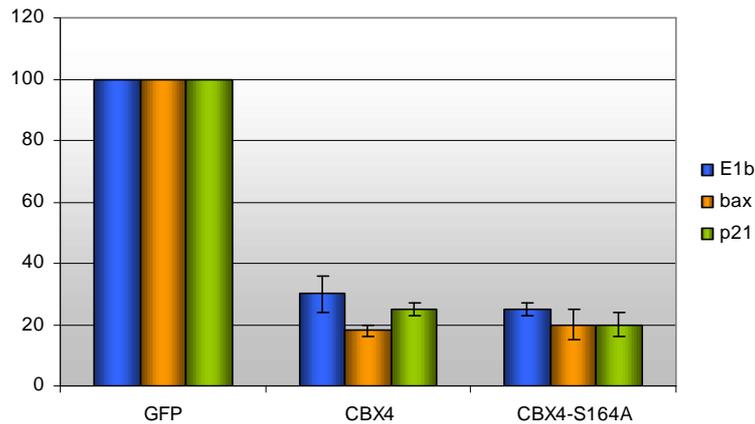


Figure 44- Activité répressive de CBX4-S164A sur 3 promoteurs différents dans un système exogène

Nous observons que la diminution d'expression du gène rapporteur exprimant la luciférase est identique entre CBX4 et CBX4-S164A. Nous en concluons que 14-3-3 et PCGF6 ne sont pas indispensables pour la répression transcriptionnelle du gène rapporteur par CBX4. D'autres protéines, comme RNF2 (la sous-unité catalytique du PRC1), sont toujours en mesure d'interagir avec CBX4-S164A, pouvant expliquer pourquoi la protéine mutée possède toujours une activité répressive.

7.3 Discussion

Nous rapportons pour la première fois une interaction entre la protéine CBX4 et les 14-3-3 dans la cellule, et ceci dépendamment de la phosphorylation de la deuxième sérine du motif « RSISTP » car un remplacement de cette dernière par une alanine abolit l'interaction. Nous ignorons si les 14-3-3 sont nécessaires pour la formation du PRC1-CBX4 ou si elles jouent un rôle de régulateur pour la protéine CBX4 non complexée, car d'une part la protéine mutée possède la même activité répressive que la protéine sauvage mais d'autre part la composition du PRC1 est incomplète avec le mutant. Nous rappelons que la phosphorylation de PCGF2 joue un rôle important dans l'activité du PRC1 (Elderkin S, 2007, Mol Cell). Nous proposons que d'autres membres du PRC1 (comme CBX4) sont également susceptibles d'être phosphorylés et que ces modifications participent à la régulation de la composition du PRC1, mais également à son activité. De telles modifications ont par ailleurs été rapportées comme par exemple lors de dommages à l'ADN, la protéine HIPK2 phosphoryle la thréonine 495 de l'isoforme longue de CBX4, qui à son tour sumoyle HIPK2, la rendant plus efficace pour la répression transcriptionnelle (Roscic A, 2006, Mol Cell). Les modifications post-traductionnelles jouent un rôle important dans la régulation des activités de certaines protéines, c'est le cas des phosphorylations dans la voie des ubiquitine E3 ligases (Gao M, 2005, Mol Cell). Des études de phosphorylation à grande échelle ont permis de mettre en évidence la phosphorylation de CBX4 sur plusieurs résidus (sérine, thréonine) (Olsen JV, 2006, Cell), mais la sérine en question n'a pas été identifiée. D'autres bases de données telle que phosida (www.phosida.de), référencant un grand nombre de phosphorylations sur des protéomes complets ou partiels de la bactérie à l'Homme

en passant par la souris, n'ont pas non plus permis d'identifier la phosphorylation de la sérine 432 ou 164 de CBX4. Pour terminer, nous avons tenté de déterminer quelle pourrait être la potentielle kinase responsable de cette phosphorylation. Une analyse informatique sur plusieurs serveurs dédiés à ce type d'analyses (<http://kinasephos.mbc.nctu.edu.tw> et <http://bioinformatics.lcd-ustc.org>), a été effectuée. Malheureusement plusieurs enzymes sont susceptibles de reconnaître le motif en question, cependant cette étude permet d'exclure la caséine kinase 2. Aussi, la Ser/Thr kinase SMG1 a été co-éluee avec CBX4 (Tableau 4), cependant sa séquence consensus est [ST]-Q donc nous avons invalidé l'hypothétique activité de cette kinase sur CBX4.

8 Purification TAP de protéines non-CBX présentes dans le PRC1

Comme nous l'avons constaté dans la première partie des résultats, les protéines CBX sont mutuellement exclusives au sein des PRC1. Nous avons réalisé les purifications TAP de deux autres membres du PRC1 : PCGF2 et RING1. Ces protéines ont été choisies car des purifications similaires ont déjà été effectuées avec les protéines RNF2 (Sanchez C, 2007, MCP) et BMI1/PCGF4 (Levine SS, 2002, MCB). De plus, RING1 est associée à presque toutes les CBX à l'exception de l'isoforme 2. Une purification des complexes associés à PCGF2 dans des cellules HEK293 a cependant été publiée après que nous avons réalisé les nôtres (Elderkin S, 2007, Mol Cell).

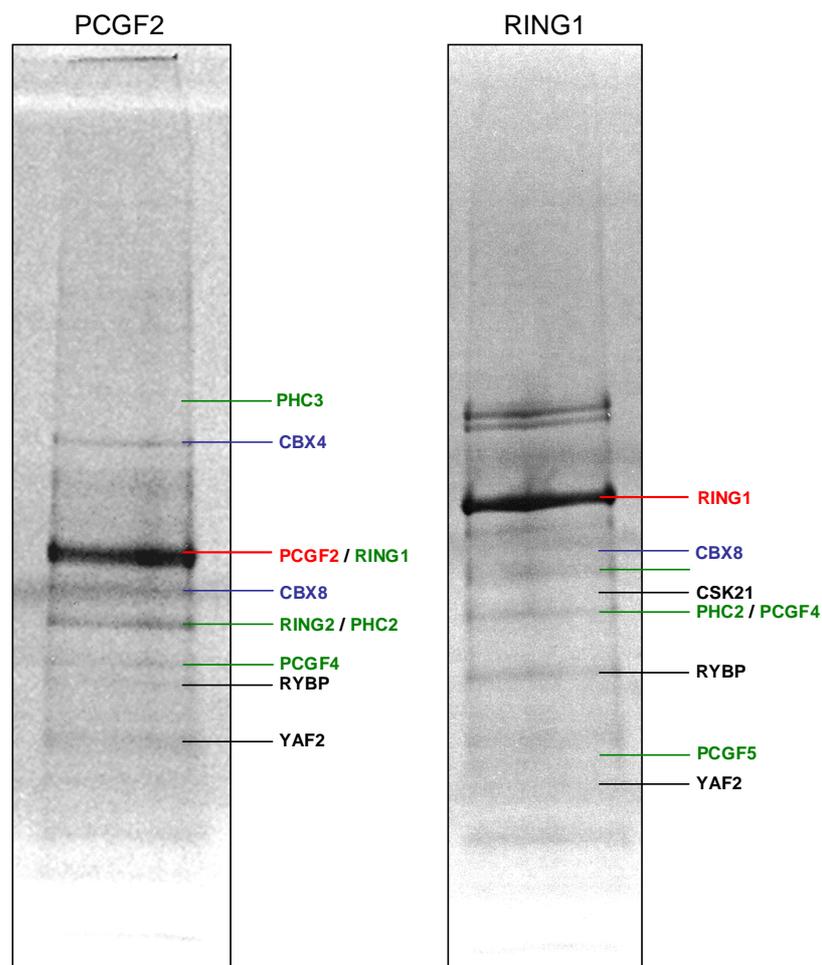


Figure 45 - Purification TAP avec les protéines PCGF2 et RING1. SDS-PAGE colorés au bleu de coomassie colloïdal

8.1 Résultats

Nous avons généré des lignes cellulaires qui expriment de manière stable les protéines PCGF2 et RNF2 fusionnées à l'étiquette TAP. Après avoir amplifié ces cellules, nous avons réalisé les purifications TAP (Figure 44) et identifié les protéines présentes dans le gel par spectrométrie de masse. De manière identique à ce qui a été décrit précédemment, des critères de sélection ont été appliqués sur les résultats bruts obtenus et nous permettent de résumer les identifications dans les Tableaux 6, les peptides sont référencés en Annexe 1 Tableaux S6 et S7.

	accession	description	masse	score	peptide
RING1	BMI1_HUMAN	Polycomb complex protein BMI-1	36925	268	12
RING1	CBX8_HUMAN	Chromobox protein homolog 8	43369	187	4
RING1	CSK21_HUMAN	Casein kinase II subunit alpha	45115	295	8
RING1	NUMA1_HUMAN	Nuclear mitotic apparatus protein 1	238115	44	3
RING1	PCGF2_HUMAN	Polycomb group RING finger protein 2	37764	60	3
RING1	PCGF5_HUMAN	Polycomb group RING finger protein 5	29694	41	1
RING1	PCGF6_HUMAN	Polycomb group RING finger protein 6	39023	97	2
RING1	PHC2_HUMAN	Polyhomeotic-like protein 2	90657	110	3
RING1	PSD12_HUMAN	26S proteasome non-ATPase regulat subunit 12	52871	42	2
RING1	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	530	94
RING1	RYBP_HUMAN	RING1 and YY1-binding protein	24806	120	6
RING1	SKP1_HUMAN	S-phase kinase-associated protein 1	18646	65	2
RING1	UBE2O_HUMAN	Ubiquitin-conjugating enzyme E2 O	141265	42	1
RING1	YAF2_HUMAN	YY1-associated factor 2	19872	139	5

Tableau 6.1 – Protéines co-éluées avec RING1-TAP

	accession	description	masse	score	pept
PCGF2	BMI1_HUMAN	Polycomb complex protein BMI-1	36925	110	5
PCGF2	BRCA2_HUMAN	Breast cancer type 2 susceptibility protein	383986	43	2
PCGF2	CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	64	2
PCGF2	CBX8_HUMAN	Chromobox protein homolog 8	43369	341	10
PCGF2	CENPE_HUMAN	Centromeric protein E	316219	41	2
PCGF2	DPLL_HUMAN	DNA polymerase lambda	63443	42	3
PCGF2	INCE_HUMAN	Inner centromere protein	106049	40	3
PCGF2	NEST_HUMAN	Nestin	177332	42	2
PCGF2	PCGF2_HUMAN	Polycomb group RING finger protein 2	37764	410	30
PCGF2	PHC2_HUMAN	Polyhomeotic-like protein 2	90657	207	5
PCGF2	PHC3_HUMAN	Polyhomeotic-like protein 3	106096	153	6
PCGF2	RECQ1_HUMAN	ATP-dependent DNA helicase Q1	73409	42	2
PCGF2	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	492	33
PCGF2	RING2_HUMAN	E3 ubiquitin-protein ligase RING2	37632	252	5
PCGF2	RYBP_HUMAN	RING1 and YY1-binding protein	24806	58	1
PCGF2	TPR_HUMAN	Nucleoprotein TPR	265441	40	4
PCGF2	YAF2_HUMAN	YY1-associated factor 2	19872	68	5

Tableau 6.2 – Protéines co-éluées avec PCGF2-TAP. Orange = appât, bleu = PRC1, rose = retrouvé avec 4 appâts CBX différents, vert = retrouvés avec 2 appâts différents

Au moment où cette thèse a été rédigée, les purifications n'ont été réalisées qu'une seule fois, contrairement aux purifications avec les CBX présentées avec les CBX Polycomb. Cependant, nous pouvons extraire quelques informations importantes avec ces résultats préliminaires.

8.1.1 peu de CBX

D'une manière assez surprenante, seule CBX8 est présente dans la purification avec RING1. Comme attendu par rapport aux résultats des purifications avec les CBX, les protéines CBX4 et CBX8 sont retrouvées dans la purification avec PCGF2, en revanche aucun peptide de CBX7 n'a été identifié. L'identification de CBX4 et CBX8 avec PCGF2 démontre que ces protéines sont présentes en quantité suffisante de manière endogène pour pouvoir être co-purifiées et identifiées. Peut-être que CBX7 est moins exprimée et que sa détection n'est pas possible dans nos conditions expérimentales. Cependant, comme aucune autre CBX en dehors de l'isoforme 8 n'est co-purifiée avec RING1 (alors que l'inverse est vrai) nous pouvons suggérer que la majeure partie des protéines RING1 dans les cellules HeLa ne sont pas associées avec les protéines CBX, ce qui confirme les résultats obtenus par gel filtration (Figure 24).

8.1.2 Le cœur du PRC1 est conservé

Les protéines BMI1, PHC2, PCGF2 et RING1 sont retrouvées dans les 2 purifications, cela suggère que le cœur du PRC1, également retrouvé avec la plupart des CBX, est bien composé de ces quelques protéines. Cependant la stœchiométrie de chacune de ces protéines au sein du complexe ne peut pas être analysée dans les conditions dans lesquelles nous avons réalisé nos purifications. Les intensités des protéines associées au PRC1 varient entre les 2 purifications, par exemple avec RING1 nous retrouvons PCGF2, PCGF5 et PCGF6 alors qu'avec PCGF2, nous constatons la présence de peptides correspondant à l'effecteur uniquement, suggérant que les complexes qui contiennent PCGF2 ne contiennent pas les autres isoformes, à l'image des CBX qui sont mutuellement exclusives.

8.1.3 RYBP et YAF2

L'identification de YAF2 et RYBP avec RING1 et PCGF2, et leur absence totale dans toutes les purifications faites avec les protéines CBX suggèrent que ces 2 protéines interagissent avec des complexes PRC1 en absence des CBX. YAF2 et RYBP appartiennent à la même famille de cofacteurs transcriptionnels, elles possèdent des structures similaires avec un doigt de zinc de type RANBP2 et une masse moléculaire de 20kDa environ.

RYBP (RING1 et YY1 binding protein) joue un rôle essentiel lors du développement embryonnaire (Pirity MK, 2005, Mol Cell Biol). Elle interagit avec des facteurs de transcription et des protéines Polycomb (Garcia E, 1999, EMBO J) (sans pour autant avoir été associée avec le PRC1) et est capable de se fixer sur l'histone H2A ubiquitinyllée (Arrigoni R, 2006, FEBS Lett). YAF2 (YY1-associated factor 2) a été un peu moins étudiée que RYBP mais son implication dans les stades précoces du développement du poisson zèbre ne fait aucun doute (Stanton SE, 2006, JBC). Ces 2 protéines, bien qu'associées aux membres du PRC1, ne sont pas responsables de transformations homéotiques lorsque leur expression est inhibée c'est pourquoi elles ne peuvent pas être considérées comme des protéines Polycomb.

Comme ces protéines sont capables de se lier à des facteurs de transcription, peut-être qu'elles peuvent recruter les PRC1 au niveau des promoteurs des gènes ciblés par YY1 et ainsi assurer la

répression transcriptionnelle. La présence des protéines RING dans le complexe permettrait de catalyser l'ubiquitylation des histones H2A environnantes.

8.1.4 La caséine kinase 2

Même si elle n'est identifiée que dans la purification de RING1, nous constatons une fois de plus la présence de la caséine kinase 2 dans la purification. Cette protéine est identifiée avec 8 peptides, représentant un score Mascot élevé donc la confiance de l'identification est relativement bonne. Cette protéine pourrait donc être associée aux protéines du PRC1 indépendamment de la présence des CBX.

8.1.5 pas de RNF2 avec RING1 ?

Le serveur en ligne Mascot est très utile pour l'identification des protéines à partir des spectres issus de l'analyse par le spectromètre de masse. A partir de nos données brutes, il nous renseigne sur la présence de RING1 exclusivement, et non RNF2, dans la purification à partir de RING1-TAP. En réalité, aucun peptide spécifique de RNF2 n'a pu être identifié (raison pour laquelle elle n'a pas été listée), en revanche 2 peptides communs aux 2 protéines RING ont été identifiés à plusieurs reprises : SLHSELMCPICLMLK et SLRPDPNFDALISK. C'est ainsi que nous ne pouvons pas exclure la présence de RNF2 dans la purification RING1 bien que mascot suggère le contraire.

8.2 Discussion

La purification TAP de 2 protéines non-CBX appartenant au complexe PRC1 nous a permis d'avoir une vision plus élargie sur la composition des complexes PRC1 dans les cellules HeLa. En effet, nos résultats suggèrent que des complexes PRC1 dépourvus de protéines CBX et contenant les protéines YAF2 et RYBP sont présents dans la cellule. Ces résultats suggèrent un recrutement possible de ces PRC1 au niveau des gènes ciblés par les facteurs de transcription qui interagissent avec ces 2 co-facteurs, indépendamment de la marque H3K27me3.

Ces deux purifications doivent être répétées pour valider ou infirmer les conclusions possibles que nous avons déduites avec nos premières identifications. De plus, il semble intéressant de pouvoir réaliser une purification TAP avec la protéine PHC2 qui est présente dans l'ensemble des purifications TAP réalisées avec un membre du PRC1. Cela permettrait de valider la présence endogène des protéines CBX2, CBX6 et CBX7 dans les cellules HeLa. Remarquons que la liste des protéines identifiées dans notre étude couvre la totalité des protéines identifiées dans l'étude de PCGF2 publiée en 2007 et que nous observons la présence de partenaires additionnels, en particulier BMI1 (PCGF4). Les auteurs ont suggéré à l'époque que PCGF2 et BMI1 étaient mutuellement exclusives dans les complexes PRC1. Avec les résultats que nous avons obtenus dans des cellules HeLa, nous remettons en question cette hypothèse, à moins que cela soit vrai dans des HEK293, soulignant une fois de plus des complexes PRC1 divergents en composition entre différents types cellulaires.

Bien entendu, ces résultats prometteurs nécessitent une caractérisation plus approfondie des interactions que nous avons identifiées. Mais si le modèle de recrutement du PRC1 indépendamment de la marque H3K27me3 était vérifié, il expliquerait pourquoi on peut retrouver des protéines

Polycomb au niveau de certains promoteurs qui ne portent pas cette marque (Bracken AP, 2006, Genes Dev).

Pour finir, la protéine UBE2O, qui avait été suggérée en tant que nouveau partenaire pour CBX6 a également été retrouvée dans la purification réalisée à partir de PCGF2-TAP. Comme CBX6 n'a pas été identifiée dans les protéines présentes dans le gel, nous pouvons exclure la possibilité d'une interaction entre UBE2O, CBX6 et PCGF2, d'autant plus que cette dernière n'a pas été trouvée dans la purification avec CBX6-TAP. Des études complémentaires permettraient de statuer.

RESULTATS ET DISCUSSION

Partie 2 : Les protéines CBX HP1 (CBX1, 3 et 5)

Après avoir caractérisé les complexes CBX associés à l'hétérochromatine facultative, nous nous sommes intéressés aux 3 protéines CBX retrouvées majoritairement au niveau de l'hétérochromatine : HP1 α , HP1 β et HP1 γ . Comme nous avons pu le constater dans l'introduction, plusieurs études ont mis en lumière les différents rôles effectués par les HP1 dans la cellule : la formation de l'hétérochromatine et l'extinction des gènes, la protection des télomères et le contrôle de l'expression des gènes dans certaines régions euchromatiques. Ces fonctions sont réalisées à l'aide d'autres protéines, partenaires stables ou transitoires, qui catalysent des réactions enzymatiques ou changent la conformation tridimensionnelle de la chromatine. Il nous est apparu intéressant de caractériser par purification TAP les trois protéines HP1 afin de pouvoir statuer sur la présence éventuelle de partenaires spécifiques à une isoforme particulière, qui pourraient favoriser une des fonctions sus-citées et comprendre pourquoi plusieurs isoformes de la même protéine sont exprimées chez les eucaryotes. Compte tenu de la mise au point des différentes techniques utilisées pour la caractérisation des CBX Polycomb, nous avons adopté la même stratégie pour les HP1.

1 Etude de l'activité répressive des protéines CBX HP1

1.1 Système rapporteur exogène et sur-expression d'une protéine fusionnée à GAL4

Après avoir cloné les séquences codantes des 3 protéines HP1 dans le plasmide pIE-N-GAL4-GW (Figure 11), nous avons testé leur activité sur les promoteurs bax et p21. Chaque expérience a été répétée au moins 3 fois avec 3 points de mesure à chaque fois. La quantité de luciférase firefly mesurée par le luminomètre a été normalisée par la quantité de luciférase renilla.

Résultats

D'une manière générale, les 3 protéines HP1 répriment très fortement les 2 promoteurs testés dans notre étude (Figure 45). Nous observons une réduction de la quantité relative de luciférase d'environ 90% quel que soit le promoteur testé.

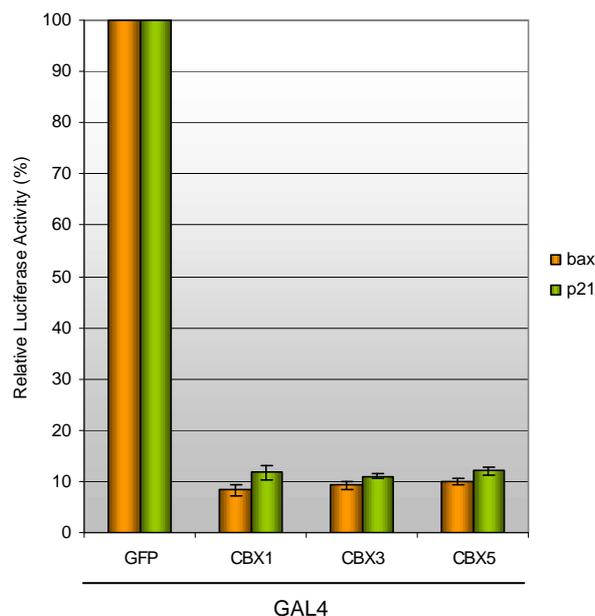


Figure 46 - Activité répressive des protéines HP1 sur 2 promoteurs différents

1.2 *Système rapporteur endogène et expression inductible d'une protéine fusionnée à GAL4*

Système rapporteur chromatinien inductible par un analogue de l'œstrogène.

Comme précédemment, les séquences codantes des protéines CBX1, CBX3 et CBX5 ont été clonées par gateway dans le plasmide pRH-N-DLO-GAL4-GW (Figure 16). Ces plasmides ont permis de produire des virus dans la lignée cellulaire d'empaquetage (HEK293gp) pour infecter les cellules HEK293-CI.7 qui contiennent dans leur génome une copie unique du gène rapporteur et obtenir une lignée stable. Dans cette dernière, l'activation de l'effecteur fusionné à GAL4 est subordonnée à l'ajout de 4-OHT dans le milieu de culture.

Résultats

Nous avons effectué des tests préliminaires pour déterminer la concentration en 4-OHT (50, 100, 200, 500, 1000nM) à utiliser et la durée du traitement (6, 12, 24, 48, 72h). Les cellules ont été collectées 48h après avoir ajouté 500nM de 4-OHT ou de l'éthanol (référence) dans le milieu de culture. Les valeurs de luciférase mesurée au luminomètre ont été normalisées par la quantité totale de protéines (bradford). Les répressions constatées sont de l'ordre de 70 à 80%, les 3 protéines testées se comportent de la même façon une fois qu'elles sont recrutées en amont du promoteur de la thymidine kinase (tk) (Figure 46).

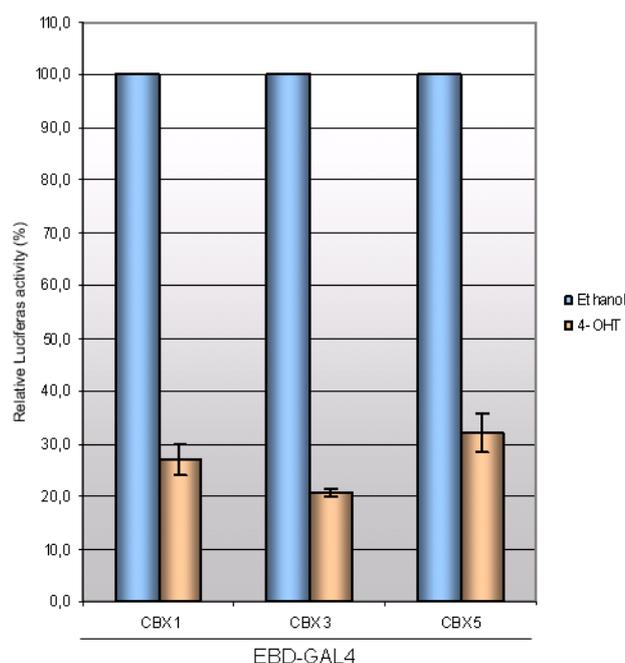


Figure 47 - Activité répressive des protéines HP1 sur un rapporteur chromatinien

1.3 Discussion

Les protéines HP1 sont des répresseurs transcriptionnels puissants capables de réprimer très fortement les promoteurs bax et p21 présents sur un plasmide exogène ou le promoteur tk intégré dans le génome des cellules HEK293. Leur activité répressive a été découverte en 1990 chez la drosophile lorsqu'une mutation dans un des gènes codant pour les HP1 a causé la perte de la variéation par effet de position (Eissenberg JC, 1990, PNAS USA). Nous observons que dans un contexte chromatinien, les répressions transcriptionnelles sont un peu moins importantes que sur un plasmide rapporteur, l'accèsion aux sites UAS_{GAL} peut en partie expliquer ce phénomène. Egalement, en fonction du site d'intégration du transgène, certaines cellules peuvent être également résistantes à l'antibiotique qui permet de faire la sélection, mais exprimer des quantités trop faible de HP1-GAL4 pour pouvoir observer une répression. Ces 3 protéines supposées agir de concert possèdent des activités similaires lorsqu'elles sont ciblées sur des promoteurs transcriptionnellement actifs. N'ayant aucune activité catalytique ou mécanique sur l'ADN ou les histones connue à ce jour, leur rôle de répresseur transcriptionnel est dépendant de leurs partenaires protéiques.

2 Validation des lignées cellulaires et localisation des protéines HP1-TAP

Nous avons généré des lignées HeLa exprimant de manière stable les protéines CBX1, CBX3 et CBX5 fusionnées à l'étiquette TAP à l'aide d'une infection rétrovirale. Après sélection des clones résistants à la puromycine, certaines cellules ont été analysées par immunomarquage avec un anticorps PAP (anti-protéine A). Les noyaux de ces mêmes cellules ont été marqués par le Hoechst 33342.

Résultats

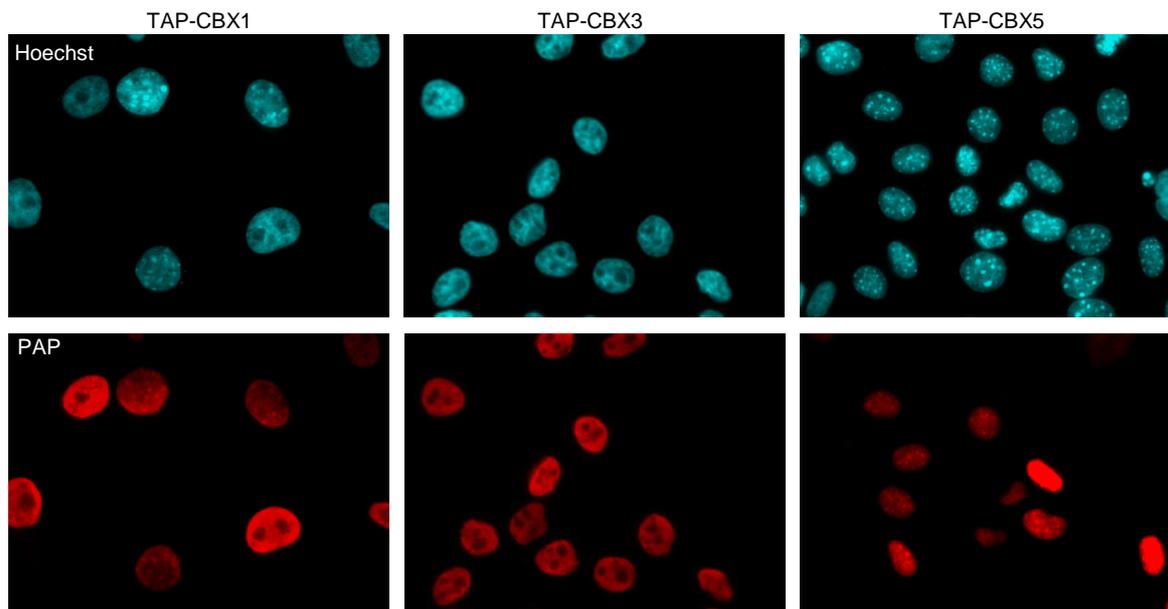


Figure 48 – Immunomarquage des protéines HP1-TAP dans les lignées stables HeLa-TAP

Alors que 100% des cellules TAP CBX1 et CBX3 expriment suffisamment l'effecteur pour qu'il soit détecté par immunomarquage, la protéine CBX5-TAP est visualisable dans moins de 50% des cellules infectées et sélectionnées (Figure 47). Les 3 protéines TAP sont exclusivement nucléaires. L'intensité du marquage pour CBX3 est assez homogène alors que pour CBX1 et CBX5, certaines cellules expriment beaucoup plus fortement que d'autres la protéine effectrice. Enfin, l'organisation de la chromatine révélée par la coloration Hoechst met en évidence des ponctuations intensément colorées dans les cellules CBX1 et CBX5, signifiant que la chromatine est très condensée dans ces régions (foyers d'hétérochromatine). Les profils de localisation des protéines TAP prouvent qu'elles sont localisées au niveau de ces régions (contrairement aux observations réalisées avec CBX7 et CBX8 en partie 1 - 2.3).

Discussion

L'infection rétrovirale de cellules HeLa en culture nous a permis d'obtenir des lignées cellulaires qui expriment de manière stable les protéines HP1 fusionnées à l'étiquette TAP. Les trois isoformes sont localisées exclusivement dans le noyau des cellules. Pour CBX1 et CBX5, la localisation est même majoritairement située au niveau des foyers d'hétérochromatine. Ces données correspondent à ce qui a été rapporté dans la littérature : les localisations de CBX1 et CBX5 préférentiellement au niveau de l'hétérochromatine et la localisation euchromatique de CBX3. Cependant, nous constatons que la sur-expression des ces protéines dans les cellules HeLa n'entraîne pas une augmentation des foyers d'hétérochromatine en comparant les cellules immunomarquées ou non pour CBX5 avec la coloration Hoechst correspondante (Figure 47). Tandis que les lignées TAP CBX1 et CBX3 sont tout a fait satisfaisantes d'un point de vue d'expression, nous remarquons que pour CBX5 seulement la moitié

des cellules exprime de manière satisfaisante la protéine TAP. Cependant, dans le cadre de notre étude nous avons tout de même utilisé cette lignée, en maintenant la pression de sélection, pour réaliser la purification TAP.

2.1 Purification TAP

Comme pour les membres du PRC1, nous avons amplifié les lignées TAP dans 40 boîtes de 15cm de diamètre avant de les collecter. Les protéines extraites à partir de ces cellules sont soumises aux deux étapes de purification d'affinité du protocole TAP. Les protéines co-éluées avec les protéines HP1 sont analysées par migration dans un gel à gradient d'acrylamide 4-12% en conditions dénaturantes. Les images des gels obtenus après coloration au bleu de coomassie colloïdal sont présentées en Figure 48.

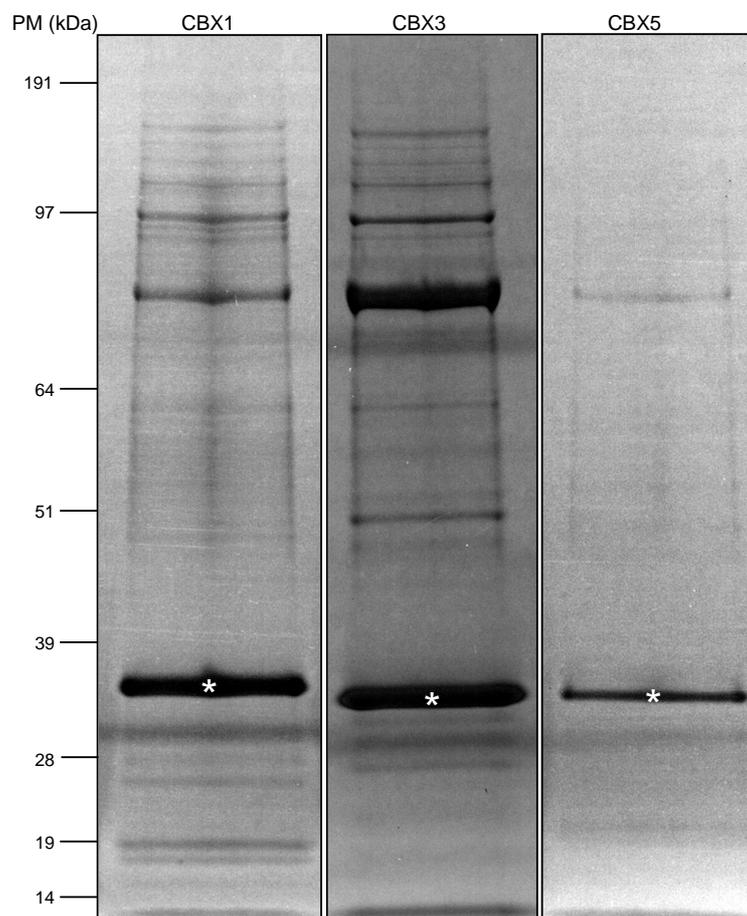


Figure 49 - Coloration au bleu de coomassie colloïdal des gels HP1-TAP. L'étoile indique la protéine TAP. PM = poids moléculaire en kilodaltons

Nous constatons que les intensités sur le gel CBX5 sont très inférieures à celles obtenues avec CBX1 et CBX3. Ce résultat n'est pas surprenant compte tenu des observations faites précédemment dans la caractérisation des lignées TAP. Les pistes ont été découpées sur toute leur hauteur et analysées par spectrométrie de masse pour identifier les protéines présentes dans le gel. Les purifications CBX3 et CBX5 ont été réalisées deux fois indépendamment tandis que pour CBX1, une seule purification a été réalisée.

	purif	accession	description	masse	score	peptides
CBX1	A	ADNP_HUMAN	Activity-dependent neuroprotector	123485	919	57
CBX1	A	ADNP2_HUMAN	ADNP homeobox protein 2	122755	93	4
CBX1	A	CAF1A_HUMAN	Chromatin assembly factor 1 subunit A	105158	100	4
CBX1	A	CAF1B_HUMAN	Chromatin assembly factor 1 subunit B	61454	43	1
CBX1	A	CBX1_HUMAN	Chromobox protein homolog 1 (HP1 beta)	21405	606	65
CBX1	A	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	492	8
CBX1	A	CBX5_HUMAN	Chromobox protein homolog 5 (HP1 alpha)	22211	98	9
CBX1	A	CHD3_HUMAN	Chromodomain-helicase-DNA-binding protein 3	226450	45	2
CBX1	A	CHD4_HUMAN	Chromodomain-helicase-DNA-binding protein 4	217853	388	10
CBX1	A	H2A1B_HUMAN	Histone H2A type 1-B	14127	182	6
CBX1	A	H2A1D_HUMAN	Histone H2A type 1-D	14099	172	3
CBX1	A	H2AV_HUMAN	Histone H2AV	13501	63	2
CBX1	A	H2B1B_HUMAN	Histone H2B type 1-B	13942	168	7
CBX1	A	H2B1C_HUMAN	Histone H2B type 1-C/E/F/G/I	13811	171	10
CBX1	A	H31T_HUMAN	Histone H3.1t	15499	132	4
CBX1	A	H4_HUMAN	Histone H4	11360	201	9
CBX1	A	IMA3_HUMAN	Importin subunit alpha-3	57775	97	2
CBX1	A	IMA4_HUMAN	Importin subunit alpha-4	57851	59	2
CBX1	A	INT1_HUMAN	Integrator complex subunit 1	244143	41	2
CBX1	A	KU70_HUMAN	ATP-dependent DNA helicase 2 subunit 1	69799	58	1
CBX1	A	POGZ_HUMAN	Pogo transposable element with ZNF domain	155245	271	10
CBX1	A	SMCA4_HUMAN	Probable global transcription activator SNF2L4	184530	107	3
CBX1	A	TIF1B_HUMAN	Transcription intermediary factor 1-beta	88493	630	27

Tableau 7.1 – Protéines co-éluées avec CBX1 (HP1β)

	purif	accession	description	masse	score	peptides
CBX3	A / B	ADNP_HUMAN	Activity-dependent neuroprotector	123485	1500 / 805	45 / 33
CBX3	A	CA103_HUMAN	Uncharacterized protein C1orf103	84516	201	3
CBX3	A / B	CBX1_HUMAN	Chromobox protein homolog 1 (HP1 beta)	21405	295 / 45	6 / 2
CBX3	A / B	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	532 / 727	109 / 72
CBX3	A / B	CBX5_HUMAN	Chromobox protein homolog 5 (HP1 alpha)	22211	166 / 146	13 / 9
CBX3	A	CHD4_HUMAN	Chromodomain helicase-DNA-binding protein 4	217853	526	14
CBX3	B	DCK_HUMAN	Deoxycytidine kinase	30499	48	3
CBX3	B	GTF2I_HUMAN	General transcription factor II-I (GTFII-I) (TFII-I)	112346	169	6
CBX3	A	H2A1D_HUMAN	Histone H2A type 1-D (H2A.3)	14099	115	2
CBX3	A	H31T_HUMAN	Histone H3.1t (H3t) (H3/t) (H3/g)	15499	67	2
CBX3	A	H4_HUMAN	Histone H4	11360	69	1
CBX3	A / B	IMA4_HUMAN	Importin alpha-4 subunit (Karyopherin alpha-4)	57851	70 / 44	2 / 1
CBX3	B	K0774_HUMAN	Uncharacterized protein KIAA0774	150032	48	4
CBX3	A / B	K1802_HUMAN	Zinc finger protein KIAA1802	89043	147 / 104	2 / 1
CBX3	A	KU70_HUMAN	ATP-dependent DNA helicase 2 subunit 1	69799	121	2
CBX3	A	POGZ_HUMAN	Pogo transposable element with ZNF domain	155245	712	14
CBX3	A	SMCA4_HUMAN	Probable global transcription activator SNF2L4	184530	82	2
CBX3	A	SUHW4_HUMAN	Suppressor of hairy wing homolog 4	109229	107	2
CBX3	A / B	TIF1B_HUMAN	Transcription intermediary factor 1-beta TIF1-beta	88493	769 / 629	73 / 44
CBX3	A	ZNF8_HUMAN	Zinc finger protein 8 (Zinc finger protein HF.18)	64930	222	3

Tableau 7.2 – Protéines co-éluées avec CBX3 (HP1γ)

	purif	accession	description	masse	score	peptides
CBX5	B	CAF1B_HUMAN	Chromatin assembly factor 1 subunit B	61454	94	3
CBX5	A / B	CBX5_HUMAN	Chromobox protein homolog 5 (HP1 alpha)	22211	659 / 544	75 / 23
CBX5	B	H2A1B_HUMAN	Histone H2A type 1-B	14127	66	1
CBX5	A	H31T_HUMAN	Histone H3.1t	15499	118	4
CBX5	A / B	H4_HUMAN	Histone H4	11360	93 / 42	2 / 1
CBX5	A	IMA2_HUMAN	Importin subunit alpha-2	57826	164	2
CBX5	A / B	IMA3_HUMAN	Importin subunit alpha-3	57775	419 / 84	10 / 2
CBX5	A / B	IMA4_HUMAN	Importin subunit alpha-4	57851	247 / 107	4 / 2
CBX5	A	RBBP4_HUMAN	Histone-binding protein RBBP4	47626	86	2
CBX5	B	SHOC2_HUMAN	Leucine-rich repeat protein SHOC-2	64848	45	1
CBX5	A/B	TIF1B_HUMAN	Transcription intermediary factor 1-beta	88493	451/64	15/2
CBX5	B	UN13D_HUMAN	Protein unc-13 homolog D	123205	44	2

Tableau 7.3 – Protéines co-éluées avec CBX5 (HP1 α)

Orange = appât, bleu = avec les 3, vert = avec 2, jaune = avec une seule

Les Tableaux 7 résumant les noms des protéines qui ont été identifiées avec certitude (les contaminants ont été retirés ainsi que les protéines avec un score Mascot inférieur à 37). Nous avons retenu 22, 19 et 11 partenaires pour CBX1, CBX3 et CBX5. Les couvertures de séquences des protéines TAP sont de 64%, 72% et 66% respectivement pour CBX1, CBX3 et CBX5 (Figure 49). Les peptides identifiés par spectrométrie de masse correspondant à ces protéines sont listés dans l'Annexe 1 Tableaux S8 à S10.

CBX1 – 64% de couverture de séquence

MGKKQNKKKVEEVLEEEEEYVVEKVLDRRVVKGKVEYLLKWKGFSDNTWEPEENLDCPDLIAEFLQSQKTAHETDKSEGGK
RKADSDSEDKGEESKPKKKKEESEKPRGFARGLEPERIIGATDSSGELMFLMKWKNSDEADLVPAKEANVKCPQVVISFYEERL
TWHSYPSEDDDKDDKN

CBX3 – 72% de couverture de séquence

MASNKTTLQKMGKQNGKSKKVEEAEPEEFVVEKVLDRRVVNGKVEYFLKWKGFSDADNTWEPEENLDCPELIEAFILNSQKAGK
EKDGTKRKSLSDSESDSKSKKRRDAADKPRGFARGLDPERIIGATDSSGELMFLMKWKDSDEADLVLAKEANMKCPQIVIAFY
EEERLTWHSCPEDAQ

CBX5 – 66% de couverture de séquence

MGKTKRKTADSSSEDEEEYVVEKVLDRRVVKGQVEYLLKWKGFSEEHNTWEPEKNLDCPELISEFMKYYKMKKEGNNKPREK
SESNKRKSNSFNSADDIKSKKKREQSNDIARGFERGLEPEKIIIGATDSCGDLMLMKWKDTDEADLVLAKEANVKCPQIVIAFY
EEERLTWHAYPEDAENKEKETAKS

Figure 50 - Couverture de séquence des protéines HP1 qui ont servi d'appât pour les purifications TAP. Les peptides soulignés sont retrouvés dans 2 ou 3 isoformes différentes

- protéines communes aux 3 isoformes HP1

Dans chacune des purifications réalisées, nous avons pu observer la présence de peptides correspondant à différentes isoformes d'histones qui composent les nucléosomes (H2A, H2B, H3 et H4), sans prédominance de l'une d'entre elle bien que les HP1 soient connues pour se fixer directement sur les histones H3. Le facteur de transcription TIF1b est également identifié dans toutes les purifications, avec un nombre de peptides important (minimum = 15 avec CBX5). Enfin, différentes sous-unités constituant les importines sont caractérisées avec les 3 protéines HP1.

- HP1 entre elles

Tandis que CBX3 et CBX5 sont identifiées avec CBX1, ainsi que CBX1 et CBX5 avec CBX3, seule CBX5 apparaît dans la liste des identifications réalisées avec la purification CBX5. Cependant, comme pour les CBX Polycomb, ces protéines possèdent des séquences communes et le peptide IIGATDSSGELMFLMK est retrouvé dans les 3 isoformes tandis que CPQIVIAFYEER existe dans les séquences de CBX3 et CBX5. Cela pourrait suggérer que dans la purification CBX5, deux peptides correspondant à CBX3 pourraient avoir été identifiés, mais l'absence d'autres peptides spécifiques pour cette protéine suggère son absence dans les protéines identifiées par spectrométrie de masse.

Les protéines identifiées dans les purifications sont résumées dans la Figure 50.

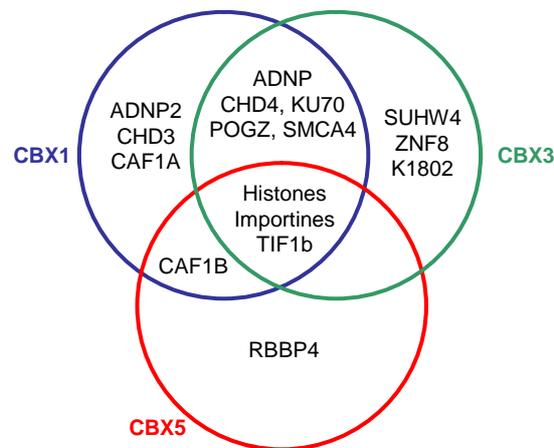


Figure 51 - Résumé des protéines majoritaires identifiées lors des purifications TAP des 3 isoformes HP1

Protéines communes aux purifications CBX1 et CBX3 : ADNP, CHD4, KU70, POGZ, SMCA4

Identifications exclusives :

CBX1 : ADNP2, CHD3 et CAF1A

CBX3 : SUHW4, ZNF8 et K1802

CBX5 : RBBP4

2.2 Discussion

Les protéines HP1 interagissent avec un grand nombre de protéines non histones associées à la chromatine (Lomberck G, 2006, Genome Biol).

Absence de méthyltransférases d'histones et d'ADN

Le paradigme de propagation de l'hétérochromatine a été détaillé dans l'introduction, et l'interaction des HP1 avec les méthyltransférases d'histones (SUV39H1 et SUV420H2) et d'ADN (Dnmt1 et Dnmt3a) (Fuks F, 2003, NAR) a été mise en évidence par différentes équipes de recherche. De manière assez inattendue, nous n'avons retrouvé aucune de ces enzymes dans nos purifications. Plusieurs hypothèses sont envisageables, tout d'abord l'interaction entre les HP1 et ces protéines est peut-être trop faible pour résister aux deux étapes de la purification TAP. De plus, leur association est

peut-être transitoire (voir Annexe 4 : interaction entre HP1 et SUV420H2), ne permettant pas de recueillir suffisamment de matériel pour qu'il puisse être identifié par spectrométrie de masse. Pour finir, ces enzymes ne sont peut-être tout simplement pas les partenaires majoritaires des protéines HP1 dans les cellules HeLa, c'est pourquoi nous ne les avons pas identifiées.

Protéines communes

Parmi les protéines identifiées lors de nos purifications, certaines ont déjà été identifiées en tant que partenaire des HP1, nous allons faire une revue concise de ces macromolécules. Nous pouvons ajouter que leur identification spécifique confirme bien la méthode employée pour la purification de complexes à partir de cellules en culture.

TIF1 β : L'interaction directe entre le répresseur transcriptionnel TIF1 β et les HP1 a d'abord été découverte par double-hybride avec des protéines humaines (Ryan RF, 1999, Mol Cell Biol), puis a été confirmée *in vivo* par immunolocalisation et co-immunoprécipitations sur les protéines endogènes (Nielsen AL, 1999, EMBO J).

CAF1B : Chromatin Assembly Factor 1 (CAF1), comme son nom l'indique, dépose les histones néo-synthétisés H3/H4 au niveau de la chromatine pendant la réplication. D'une manière générale cette protéine est un facteur d'assemblage de la chromatine au niveau des foyers de réplication aussi bien de l'eucaryote que de l'hétéro-chromatine (Krude T, 1995, Exp Cell Res). Son interaction avec 2 des 3 HP1 chez la souris est dépendante d'un domaine baptisé MIR (MOD1 Interaction Region). La mutation de ce dernier abolit la reconnaissance entre CAF1 et HP1 mais ne perturbe pas l'activité de CAF1 (Murzina, N, 1999, Mol Cell).

RBBP4 : La retinoblastoma-binding protein 4, également intitulée CAF1C (p48kDa), n'est autre qu'une des petites sous-unités du complexe CAF précédemment détaillé.

KU70 : Ku est un hétérodimère constitué de 2 sous-unités de 70 et 80kDa, capable de se lier à l'ADN (Rathmell WK, 1994, PNAS), il joue un rôle important dans la régulation et la maintenance des télomères, des levures aux mammifères (Hsu HL, 1999, PNAS). La sous-unité Ku70 est majoritairement exprimée en comparaison de Ku80, ce qui lui permet d'interagir avec d'autres protéines et par conséquent de participer à d'autres fonctions biologiques. Parmi les partenaires protéiques spécifiques de Ku70, on retrouve CBX5 dont l'interaction au niveau du chromoshadow-domaine a été découverte par double-hybride, confirmée par immunolocalisation, GST-pulldown et co-immunoprécipitation sur protéines endogènes (Song K, 2001, JBC).

SMCA4 : Cette protéine, également appelée BRG1 ou SNF2 β , est une homologue de la protéine SWI2/SNF2 de levure (Muchardt C, 1999, J Mol Biol). Elle appartient à un complexe qui possède une activité hélicase sur l'ADN et les nucléosomes de manière ATP-dépendante (Wang W, 1996, EMBO J). Son interaction directe avec les HP1 a d'abord été démontrée par des études *in vitro*, puis dans

des cellules de mammifères par co-immunoprécipitations (Nielsen AL, 2002, EMBO J). De plus, SMCA4 pourrait jouer un rôle dans l'activité répressive des HP1, en modifiant la structure tridimensionnelle de la chromatine.

Nous pouvons mentionner que le facteur de transcription GTF2I qui a été co-élué avec CBX3 (6 peptides, score = 169), l'a également été avec une des purifications CBX2 (4 peptides, score = 201). Le poids moléculaire de cette protéine est de 112kDa, et elle a été identifiée dans les 2 purifications à une hauteur correspondante (premier tiers des découpes). Comme le nombre de peptides est conséquent et le score fiable, mais que cette protéine n'a pas été retrouvée dans les autres purifications, elle est un partenaire potentiel de ces 2 CBX sans toutefois avoir une exclusivité pour l'une d'elle.

Les nouveaux partenaires identifiés

L'hétérochromatine et les transposases

Parmi les protéines majoritaires dans les éluats après les deux étapes de la purification TAP, nous retrouvons *pogZ* (14 peptides avec CBX3, 10 peptides avec CBX1). Sa dénomination complète est la suivante : Pogo transposable element avec domaines ZNF (doigt de zinc). Elle en contient 9 de type C2H2, mais également un domaine DDE et un domaine de liaison à l'ADN de type HTH (helix turn helix) CENPB. Son domaine DDE est similaire à celui des transposases TIGD (Tigger-derived) (Namgoong SY, EMBO J, 1998). Le domaine catalytique DDE, également retrouvé dans les intégrases lentivirales (Dyda F, Science, 1994), permet la modification de l'ADN comme le clivage simple ou double brin ou la ligation. Un criblage double-hybride a permis de mettre en évidence l'interaction entre LEDGF/p75 et *pogZ* (Bartholomeeusen K, 2009, JBC). LEDGF/p75 (Lens Epithelium-derived Growth Factor) est un partenaire cellulaire prépondérant de l'intégrase de HIV-1. Les auteurs ont conclu, entre autres, que *pogZ* était une transposase d'ADN domestiquée à l'image d'autres transposases domestiquées qui ont un rôle dans l'organisation nucléaire (ex : CENPB), la recombinaison (ex : RAG-1) ou la réparation de l'ADN (ex : SETmar) (Volf JN, 2006, Bioessays). Notons que la présence d'une transposase au niveau de sites hétérochromatiques a été rapportée en 2007 (Ikeda R, 2007, Mol Cell Biol), il s'agissait de la transposase Sleeping Beauty (SB). Cette dernière possède une activité importante dans les cellules de mammifères et est utilisée comme outil de biologie moléculaire pour insérer des mutations ou des gènes dans un génome. Cette étude réalisée dans des cellules souches de souris démontre que l'efficacité de l'excision du transposon SB est 100 fois supérieure lorsque cette séquence est située dans une région de chromatine compacte. De plus, la transposase possède une forte affinité pour l'hétérochromatine et s'y fixe de manière très forte. Cependant, la validation d'une interaction entre les HP1 et cette transposase n'a pu être mise en évidence par co-immunoprécipitation.

Depuis les premiers résultats des projets concernant le séquençage de génomes de mammifères, nous nous sommes aperçu que près de la moitié de notre ADN était composé d'éléments transposables (Waterston RH, 2002, Nature). Il est évident que les éléments transposables sont

retrouvés sous forme d'hétérochromatine dans les génomes (Junakovic N, 1998, J Mol Evol), et qu'ils jouent un rôle non négligeable dans l'évolution des génomes. Et ce champ de recherche semble pouvoir tenir une importance relative dans les prochaines années, en effet il apparaît désormais évident que des événements de transposition sont associés à des réarrangements chromosomiques (translocations, les inversions ou les duplications), et qu'ils peuvent être responsables d'une cassure et d'un réarrangement de l'ADN (Zhang J, 2009, Genes Dev). Comprendre la mécanistique de ces translocations permettrait de mieux comprendre l'évolution des génomes et la relation entre *pogZ* et les HP1 pourrait être un point de départ. Le rôle fonctionnel des transposases dans l'organisation des génomes eucaryotes a été récemment démontré de manière élégante à l'aide de RNAi (micro-injection d'ADN double brin) dirigés contre les différentes transposases de *Oxytricha trifallax* (protozoaire cilié) au cours de son développement. Cet eucaryote unicellulaire détruit 95% du génome de sa lignée germinale au cours du développement, et les 3 types de transposases sont essentielles pour cet événement (Nowacki M, Science, 2009).

La protéine *pogZ* a également été baptisée SUHW5 (suppressor of hairy wing 5) à cause de la similarité de séquence de ses domaines répétés en doigt de zinc avec ceux de la protéine de drosophile *su(Hw)* (Parkhurst SM, 1988, Genes Dev), cependant cette dernière ne possède pas les domaines DDE et HTH, tout comme SUHW4 qui a été identifiée dans la purification de CBX3. *su(Hw)* joue un rôle au niveau des insulateurs chromatiniens, en bloquant la communication entre un enhancer et un promoteur. D'une manière plus générale, les insulateurs sont des éléments régulateurs qui établissent des domaines d'activités transcriptionnelles indépendantes au niveau des génomes eucaryotes (formation de boucles au niveau de la chromatine).

ZNF8

La protéine Zinc Finger 8 contient 7 domaines à doigt de zinc et un domaine KRAB, elle a été clonée chez l'Homme en 1990 (Lania L, 1990, Genomics). L'homologue murin est un répresseur transcriptionnel localisé dans le noyau des cellules, qui interagit avec une forme phosphorylée de SMAD1 (Jiao K, 2002, Mol Cell Biol).

ADNP et ADNP2

Découverte chez la souris en 1999, la protéine ADNP (Activity Dependent Neuroprotective Protein) a rapidement été identifiée dans des cellules humaines l'année suivante (Zamostiano R, 2000, JBC). Elle contient 5 domaines à doigt de zinc, une région riche en prolines et un domaine homéobox, suggérant une fonction de facteur de transcription. L'extinction de son expression par des oligonucléotides antisens entraîne l'augmentation de l'expression du suppresseur de tumeur p53 et réduit la viabilité de cellules cancéreuses intestinales de 90%. Cette protéine est donc supposée jouer un rôle dans le maintien de la survie cellulaire. De plus, elle est essentielle pour le développement du cerveau et contient un motif peptidique (NAPVSIPQ) qui favorise la neuroprotection contre une grande variété d'agents cytotoxiques (en particulier les peptides β -amyloïdes de la maladie d'Alzheimer), même à des doses relativement faibles (Steingart RA, 2006, Mol Cell Endocrinol). Les souris knock-out pour ADNP ne sont pas viables : malformation au niveau du tube neural crânien et mortalité

embryonnaire à 8,5 jours (Pinhasov A, 2003, Dev Brain Res). Une approche protéomique a permis d'identifier les partenaires d'ADNP dans des cellules HEK293 (Mandel S, 2007, JBC) : trois composants (parmi une dizaine) du complexe de remodelage de la chromatine SWI/SNF : BAF250a, BRG1 et BAF170 et des expériences complémentaires montrent que l'expression d'ADNP est associée à la morphologie des cellules en modifiant l'organisation des microtubules. Dans cette publication, la protéine CBX5 ne semble pas avoir été identifiée dans la purification alors que la même équipe de recherche avait validé leur interaction *via* le domaine PxVxL par co-immunoprécipitation dans des cellules différenciées et non différenciées (P19, cellules de carcinome embryonnaire) l'année précédente (Mandel S, 2006, Dev Biology). Les connaissances sur le rôle d'ADNP au sein des cellules est de plus en plus détaillé, mais son rôle biologique au niveau de l'interaction avec les HP1 nécessite une étude plus approfondie mais sa fonctionnalité pour le bon développement des cellules ne fait plus aucun doute. La caractérisation moléculaire et fonctionnelle d'ADNP2 a débuté récemment (Kushnir M, 2008, J Neurochem), elle semble avoir un rôle similaire à ADNP (survie cellulaire au niveau du cerveau). D'après nos résultats, nous pouvons d'ores et déjà supposer qu'elle interagit également avec les protéines HP1, en particulier CBX1.

CHD3/CHD4 : Les protéines Mi2- α (CHD3) et Mi2- β (CHD4) sont également des hélicases appartenant à la famille SWI2/SNF2 et qui sont retrouvées au niveau de l'hétérochromatine dans les lymphocytes humains (Kim J, 1999, Immunity). Ces 2 protéines possèdent un double chromo-domaine (comme les HP1) et un double-PHD domaine (Woodage T, 1997, PNAS), ces éléments suggèrent qu'elles sont capables de reconnaître des modifications post-traductionnelles des histones. Elles ont été décrites comme sous-unités du complexe NuRD. Celui-ci est capable de réprimer l'expression des gènes par remodelage de la chromatine *via* l'activité des hélicases mais également par l'activité histone déacétylase des HDAC1 et HDAC2 (Tong JK, 1998, Nature / Xue Y, 1998, Mol Cell / Zhang Y, 1998, Cell). L'interaction entre CHD3/CHD4, ni même le complexe NuRD et les HP1 n'a pas été clairement démontrée à ce jour, mais nos résultats et leurs localisations et fonctions nous permettent de penser que les HP1 (CBX1 et CBX3) peuvent interagir de concert avec CHD3 et CHD4 pour réprimer l'expression des gènes localisés au niveau de l'hétérochromatine.

Les importines alpha

Identifiées dans toutes les purifications avec plusieurs peptides et des scores Mascot satisfaisants, les importines α (ou karyophérines α) apparaissent comme étant de nouveaux partenaires des 3 HP1. Il existe plusieurs voies différentes pour le transport nucléo-cytoplasmique et l'importation classique des protéines vers le noyau suite à leur synthèse dans le cytoplasme est la voie la mieux caractérisée. Le transport des macromolécules vers le noyau ou vers le cytoplasme est régulé par des structures appelées Nuclear Pore Complexes (NPC). Ces pores, localisés à la surface de l'enveloppe nucléaire, permettent la diffusion passive des ions et des petites protéines (<40kDa) mais limitent le passage des molécules plus importantes qui ne possèdent pas un signal d'adressage adapté (Paine PL, 1975, Nature). Le transport actif fait intervenir 2 types de protéines : les karyophérines qui regroupent les importines et les exportines (Radu A, 1995, PNAS) et les GTPases de la famille Ras, les Ran (Quimby

BB, 2003, Curr Opin Cell Biol). Le modèle suivant est admis pour le transport actif vers le noyau : dans le cytoplasme, l'importine α reconnaît et se fixe sur la protéine à importer, permettant la fixation de l'importine β . Cette dernière interagit avec un NPC afin de réaliser la translocation du complexe. Une fois dans le noyau, la fixation de RanGTP sur l'importine β entraîne sa libération du complexe importine α -« protéine à importer » par changement conformationnel (Lee SJ, 2005, Nature). Une fois que la protéine a été correctement adressée à l'intérieur du noyau, l'importine α est recyclée pour un autre cycle d'importation.

Parmi toutes les protéines synthétisées dans le cytoplasme, seules les protéines à vocation nucléaire doivent être importées dans le noyau. Afin de pouvoir être discriminées, elles possèdent toutes un signal de localisation vers le noyau appelé NLS (Nuclear Localisation Signal), qui peut être situé n'importe où dans la séquence de la protéine. Les NLS contiennent une séquence courte (4-8 aa) souvent riche en acides aminés chargés positivement (Lysine, Arginine). Certains NLS sont constitués de 2 blocs séparés par une dizaine de résidus (Robbins J, 1991, Cell), comme celui des HP1. En effet, bien que ces protéines possèdent un poids moléculaire faible (environ 21kDa), elles possèdent un NLS qui a d'abord été décrit pour Swi6 (homologue HP1) chez *S. pombe* (Wang G, 2000, Mol Cell Biol), il se trouve au niveau de la charnière entre le chromo-domaine et le chromoshadow-domaine. Ce NLS est retrouvé dans les HP1a et HP1b de la drosophile et les 3 HP1 humaines, où les séquences KRK sont conservées (Smothers JF, 2001, Mol Cell Biol). D'autre part la phosphorylation de la sérine 83 de HP1 γ par PKA joue un rôle primordial dans sa localisation (exclusivement euchromatique quand modifié), et ce résidu se trouve précisément au niveau du NLS (Lomberk G, 2006, Nat Cell Biol). De manière surprenante, aucune publication n'a rapporté l'interaction directe entre les HP1 et les importines α alors qu'il ne fait aucun doute qu'elle existe compte tenu des informations contenues dans la littérature. Nos purifications confirment cette idée et il serait intéressant d'étudier l'implication des possibles modifications post-traductionnelles des HP1 sur leur interaction avec les karyophérines. Il semblerait cependant que l'équipe de Raul Urrutia ait présenté ce type d'étude lors d'une conférence en 2004 (résumé dans : Lomberk G, 2004, Pancreas).

La protéine K1802, également dénommée ZN828, a été identifiée avec les 2 purifications de CBX3. Aucune information fonctionnelle n'est disponible pour cette protéine de 89kDa, possédant 5 doigts de zinc de type C2H2 répartis aux extrémités amino- et carboxy-terminales. Elle apparaît comme un partenaire spécifique de CBX3 car aucun peptide correspondant n'a pu être identifié dans les autres purifications réalisées.

Les résultats obtenus avec les HP1 confirment la présence d'un grand nombre de partenaires déjà décrits dans la littérature, validant la technologie TAP dans les cellules HeLa. Cependant, bien qu'elles soient parmi les protéines chromatiniennes non-histones les plus étudiées, nous avons eu la possibilité d'identifier de nouveaux partenaires protéiques qui, compte tenu de leur intensité sur gel, du nombre de peptides identifiés et du score Mascot, sont de sérieux candidats qui pourraient être impliqués dans la structure de l'hétérochromatine constitutive : pogZ, CHD3/CHD4 et ADNP2.

RESULTATS ET DISCUSSION

Partie 3 : Les contaminants

n° accession	description	masse
1 ACTA_HUMAN	Actin, aortic smooth muscle (Alpha-actin-2)	41982
2 ACTB_HUMAN	Actin, cytoplasmic 1 (Beta-actin)	41710
3 ACTC_HUMAN	Actin, alpha cardiac muscle 1 (Alpha-cardiac actin)	41992
4 ALBU_HUMAN	Serum albumin precursor	69321
5 ANK3_HUMAN	Ankyrin-3	480403
6 CALM_HUMAN	Calmodulin (CaM)	16827
7 COPB_HUMAN	Coatamer subunit beta	107142
8 CPNS1_HUMAN	Calpain small subunit 1 (CSS1) (Calcium-dependent protease small subunit 1)	28298
9 CPNS2_HUMAN	Calpain small subunit 2 (CSS2) (Calcium-dependent protease small subunit 2)	27643
10 CRUM1_HUMAN	Crumbs homolog 1 precursor	154081
11 DDX18_HUMAN	ATP-dependent RNA helicase DDX18 (EC 3.6.1.-) (DEAD box protein 18)	75359
12 DDX21_HUMAN	Nucleolar RNA helicase 2 (EC 3.6.1.-) (Nucleolar RNA helicase II)	87290
13 DDX24_HUMAN	ATP-dependent RNA helicase DDX24	96271
14 DDX5_HUMAN	Probable ATP-dependent RNA helicase DDX5 (EC 3.6.1.-) (DEAD box protein 5)	69105
15 DHX15_HUMAN	Putative pre-mRNA-splicing factor ATP-dependent RNA helicase DHX15	90875
16 DHX30_HUMAN	Putative ATP-dependent RNA helicase DHX30 (EC 3.6.1.-) (DEAH box protein 30)	133854
17 DHX57_HUMAN	Putative ATP-dependent RNA helicase DHX57 (EC 3.6.1.-) (DEAH box protein 57)	155507
18 DHX9_HUMAN	ATP-dependent RNA helicase A (EC 3.6.1.-) (Nuclear DNA helicase II)	140869
19 DOK7_HUMAN	Protein Dok-7	53064
20 DYH9_HUMAN	Dynein heavy chain 9, axonemal	511605
21 DYL1_HUMAN	Dynein light chain 1, cytoplasmic (Dynein light chain LC8-type 1)	10359
22 EF1A1_HUMAN	Elongation factor 1-alpha 1 (EF-1-alpha-1) (Elongation factor 1 A-1)	50109
23 FBSH_HUMAN	Probable fibrosin-1 long transcript protein	38890
24 FUS_HUMAN	RNA-binding protein FUS	53394
25 GRP75_HUMAN	Stress-70 protein, mitochondrial precursor (75 kDa glucose-regulated protein)	73635
26 GRP78_HUMAN	78 kDa glucose-regulated protein precursor (GRP 78) (Heat shock 70 kDa protein 5)	72288
27 HNRCL_HUMAN	Heterogeneous nuclear ribonucleoprotein C-like 1 (hnRNP core protein C-like 1)	32123
28 HNRL2_HUMAN	Heterogeneous nuclear ribonucleoprotein U-like protein 2	85052
29 HNRPG_HUMAN	Heterogeneous nuclear ribonucleoprotein G (hnRNP G)	42306
30 HNRPM_HUMAN	Heterogeneous nuclear ribonucleoprotein M (hnRNP M)	77464
31 HNRPQ_HUMAN	Heterogeneous nuclear ribonucleoprotein Q (hnRNP Q)	69560
32 HNRPR_HUMAN	Heterogeneous nuclear ribonucleoprotein R (hnRNP R)	70899
33 HNRPU_HUMAN	Heterogeneous nuclear ribonucleoprotein U (hnRNP U)	90457
34 HS71L_HUMAN	Heat shock 70 kDa protein 1L	70331
35 HS90B_HUMAN	Heat shock protein HSP 90-beta (HSP 84) (HSP 90)	83212
36 HSP71_HUMAN	Heat shock 70 kDa protein 1 (HSP70.1) (HSP70-1/HSP70-2)	70009
37 HSP72_HUMAN	Heat shock-related 70 kDa protein 2 (Heat shock 70 kDa protein 2)	69978
38 HSP77_HUMAN	Heat shock 70 kDa protein 7	26890
39 HSP7C_HUMAN	Heat shock cognate 71 kDa protein (Heat shock 70 kDa protein 8)	70854
40 IGKC_HUMAN	Ig kappa chain C region	11609
41 K6PL_HUMAN	6-phosphofructokinase, liver type (EC 2.7.1.11) (Phosphofructokinase 1)	84964
42 KAC_HUMAN	Ig kappa chain C region	11602
43 LRR7_HUMAN	Leucine-rich repeat-containing protein 7	172473
44 LRRK2_HUMAN	Leucine-rich repeat serine/threonine-protein kinase 2	286058
45 LYSC_HUMAN	Lysozyme C precursor	16526
46 MTHR_HUMAN	Methylenetetrahydrofolate reductase (EC 1.5.1.20)	74549
47 MYH10_HUMAN	Myosin-10	228798
48 MYH9_HUMAN	Myosin-9 (Myosin heavy chain 9) (Myosin heavy chain, nonmuscle IIa)	226392
49 MYL6B_HUMAN	Myosin light polypeptide 6B (Smooth muscle and nonmuscle myosin light chain alkali 6B)	22750
50 MYLK2_HUMAN	Myosin light chain kinase 2, skeletal/cardiac muscle (EC 2.7.11.18) (MLCK2)	64644
51 NFM_HUMAN	Neurofilament medium polypeptide	102448
52 NIPBL_HUMAN	Nipped-B-like protein	316051
53 PRKDC_HUMAN	DNA-dependent protein kinase catalytic subunit	469089
54 PTC3_HUMAN	Pentatricopeptide repeat-containing protein 3, mitochondrial	78550
55 RL10_HUMAN	60S ribosomal protein L10 (QM protein) (Tumor suppressor QM)	24561
56 RL10A_HUMAN	60S ribosomal protein L10a (CSA-19)	24816
57 RL11_HUMAN	60S ribosomal protein L11 (CLL-associated antigen KW-12)	20240
58 RL12_HUMAN	60S ribosomal protein L12	17808
59 RL13_HUMAN	60S ribosomal protein L13 (Breast basic conserved protein 1)	24247
60 RL13A_HUMAN	60S ribosomal protein L13a (23 kDa highly basic protein)	23562
61 RL14_HUMAN	60S ribosomal protein L14 (CAG-ISL 7)	23275
62 RL15_HUMAN	60S ribosomal protein L15	24131
63 RL17_HUMAN	60S ribosomal protein L17 (L23)	21383
64 RL18_HUMAN	60S ribosomal protein L18	21621

65 RL18A_HUMAN	60S ribosomal protein L18a	20749
66 RL19_HUMAN	60S ribosomal protein L19	23451
67 RL1D1_HUMAN	Ribosomal L1 domain-containing protein 1 (Cellular senescence-inhibited gene protein)	54939
68 RL21_HUMAN	60S ribosomal protein L21	18553
69 RL22_HUMAN	60S ribosomal protein L22 (Epstein-Barr virus small RNA-associated protein)	14778
70 RL22L_HUMAN	Ribosomal protein L22-like 1	14598
71 RL23_HUMAN	60S ribosomal protein L23 (Ribosomal protein L17)	14856
72 RL23A_HUMAN	60S ribosomal protein L23a	17684
73 RL24_HUMAN	60S ribosomal protein L24 (Ribosomal protein L30)	17768
74 RL26L_HUMAN	60S ribosomal protein L26-like 1	17246
75 RL27_HUMAN	60S ribosomal protein L27	15788
76 RL27A_HUMAN	60S ribosomal protein L27a	16551
77 RL29_HUMAN	60S ribosomal protein L29	17741
78 RL3_HUMAN	60S ribosomal protein L3 (HIV-1 TAR RNA-binding protein B) (TARBP-B)	46080
79 RL30_HUMAN	60S ribosomal protein L30	12776
80 RL31_HUMAN	60S ribosomal protein L31	14454
81 RL32_HUMAN	60S ribosomal protein L32	15850
82 RL35_HUMAN	60S ribosomal protein L35	14543
83 RL35A_HUMAN	60S ribosomal protein L35a	12530
84 RL36_HUMAN	60S ribosomal protein L36	12246
85 RL37A_HUMAN	60S ribosomal protein L37a	10268
86 RL4_HUMAN	60S ribosomal protein L4 (L1)	47667
87 RL5_HUMAN	60S ribosomal protein L5	34341
88 RL6_HUMAN	60S ribosomal protein L6	32708
89 RL7_HUMAN	60S ribosomal protein L7	29207
90 RL7A_HUMAN	60S ribosomal protein L7a (Surfeit locus protein 3) (PLA-X polypeptide)	29977
91 RL8_HUMAN	60S ribosomal protein L8	28007
92 RL9_HUMAN	60S ribosomal protein L9	21850
93 RLA0_HUMAN	60S acidic ribosomal protein P0 (L10E)	34252
94 RLA1_HUMAN	60S acidic ribosomal protein P1	11507
95 RLA2_HUMAN	60S acidic ribosomal protein P2 (Renal carcinoma antigen NY-REN-44)	11658
96 RM12_HUMAN	39S ribosomal protein L12, mitochondrial precursor	21335
97 ROA0_HUMAN	Heterogeneous nuclear ribonucleoprotein A0 (hnRNP A0)	30822
98 ROA1_HUMAN	Heterogeneous nuclear ribonucleoprotein A1 (Helix-destabilizing protein)	38822
99 ROA3_HUMAN	Heterogeneous nuclear ribonucleoprotein A3 (hnRNP A3)	39571
100 RS10_HUMAN	40S ribosomal protein S10	18886
101 RS11_HUMAN	40S ribosomal protein S11	18419
102 RS12_HUMAN	40S ribosomal protein S12	14516
103 RS13_HUMAN	40S ribosomal protein S13	17212
104 RS14_HUMAN	40S ribosomal protein S14	16263
105 RS15_HUMAN	40S ribosomal protein S15	17029
106 RS15A_HUMAN	40S ribosomal protein S15a	14830
107 RS16_HUMAN	40S ribosomal protein S16	16435
108 RS17_HUMAN	40S ribosomal protein S17	15540
109 RS18_HUMAN	40S ribosomal protein S18	17708
110 RS19_HUMAN	40S ribosomal protein S19	16051
111 RS2_HUMAN	40S ribosomal protein S2	31305
112 RS20_HUMAN	40S ribosomal protein S20	13364
113 RS23_HUMAN	40S ribosomal protein S23	15798
114 RS24_HUMAN	40S ribosomal protein S24	15413
115 RS25_HUMAN	40S ribosomal protein S25	13734
116 RS26_HUMAN	40S ribosomal protein S26	13007
117 RS26L_HUMAN	40S ribosomal protein S26-like 1	12994
118 RS27_HUMAN	40S ribosomal protein S27	9455
119 RS27A_HUMAN	40S ribosomal protein S27a	9412
120 RS27L_HUMAN	40S ribosomal protein S27-like protein	9471
121 RS29_HUMAN	40S ribosomal protein S29	6672
122 RS3_HUMAN	40S ribosomal protein S3	26671
123 RS30_HUMAN	40S ribosomal protein S30	6644
124 RS3A_HUMAN	40S ribosomal protein S3a	29926
125 RS4X_HUMAN	40S ribosomal protein S4	29579
126 RS5_HUMAN	40S ribosomal protein S5	22862
127 RS6_HUMAN	40S ribosomal protein S6	28663
128 RS7_HUMAN	40S ribosomal protein S7	22113
129 RS8_HUMAN	40S ribosomal protein S8	24190
130 RS9_HUMAN	40S ribosomal protein S9	22578
131 RT06_HUMAN	Mitochondrial 28S ribosomal protein S6	14218
132 RT07_HUMAN	28S ribosomal protein S7, mitochondrial precursor (S7mt) (MRP-S7) (bMRP27a)	28144
133 RT09_HUMAN	28S ribosomal protein S9, mitochondrial precursor (S9mt) (MRP-S9)	45794
134 RT14_HUMAN	Mitochondrial 28S ribosomal protein S14	15129
135 RT15_HUMAN	28S ribosomal protein S15, mitochondrial precursor	29823
136 RT18B_HUMAN	28S ribosomal protein S18b, mitochondrial precursor	29377

137 RT21_HUMAN	Mitochondrial 28S ribosomal protein S21	10734
138 RT22_HUMAN	Mitochondrial 28S ribosomal protein S22 (S22mt) (MRP-S22)	41254
139 RT23_HUMAN	Mitochondrial ribosomal protein S23 (S23mt) (MRP-S23)	21757
140 RT24_HUMAN	28S ribosomal protein S24, mitochondrial precursor	19003
141 RT25_HUMAN	Mitochondrial 28S ribosomal protein S25	20103
142 RT26_HUMAN	28S ribosomal protein S26, mitochondrial precursor	24197
143 RT27_HUMAN	Mitochondrial 28S ribosomal protein S27	47639
144 RT29_HUMAN	Mitochondrial 28S ribosomal protein S29	45538
145 RT31_HUMAN	28S ribosomal protein S31, mitochondrial precursor	45290
146 RT34_HUMAN	Mitochondrial 28S ribosomal protein S34 (S34mt) (MRP-S34)	25634
147 S100B_HUMAN	Protein S100-B	10706
148 S10A2_HUMAN	Protein S100-A2 (S100 calcium-binding protein A2) (Protein S-100L) (CAN19)	11109
149 S10A8_HUMAN	Protein S100-A8 (S100 calcium-binding protein A8) (Calgranulin-A)	10828
150 SNX25_HUMAN	Sorting nexin-25	97945
151 SPTB2_HUMAN	Spectrin beta chain, brain 1	274439
152 SSH1_HUMAN	Protein phosphatase Slingshot homolog 1	115511
153 STK10_HUMAN	Serine/threonine-protein kinase 10	112065
154 SYTL2_HUMAN	Synaptotagmin-like protein 2	100734
155 TBA1A_HUMAN	Tubulin alpha-1A chain (Tubulin B-alpha-1) (Tubulin alpha-3 chain)	50104
156 TBA1B_HUMAN	Tubulin alpha-1B chain (Tubulin alpha-ubiquitous chain) (Alpha-tubulin ubiquitous)	50120
157 TBA1C_HUMAN	Tubulin alpha-1C chain (Tubulin alpha-6 chain) (Alpha-tubulin 6)	49863
158 TBA4A_HUMAN	Tubulin alpha-4A chain	49892
159 TBA6_HUMAN	Tubulin alpha-6 chain (Alpha-tubulin 6)	49863
160 TBAK_HUMAN	Tubulin alpha-ubiquitous chain (Alpha-tubulin ubiquitous) (Tubulin K-alpha-1)	50120
161 TBB1_HUMAN	Tubulin beta-1 chain	50295
162 TBB2A_HUMAN	Tubulin beta-2A chain	49875
163 TBB2B_HUMAN	Tubulin beta-2B chain	49921
164 TBB2C_HUMAN	Tubulin beta-2C chain (Tubulin beta-2 chain)	49799
165 TBB5_HUMAN	Tubulin beta chain (Tubulin beta-5 chain)	49639
166 TEN2_HUMAN	Teneurin-2 (Ten-2) (Tenascin-M2) (Ten-m2) (Protein Odd Oz/ten-m homolog 2)	307496
167 TITIN_HUMAN	Titin	3816188
168 TRY3_HUMAN	Trypsin-3	32529
169 WDR68_HUMAN	WD repeat-containing protein 68	38926
170 ZCH11_HUMAN	Zinc finger CCHC domain-containing protein 11	185166

Tableau 8 – Liste des protéines considérées comme contaminants après 25 purifications TAP

Pendant la durée de ma thèse, notre équipe a réalisé plus d'une vingtaine de purifications TAP, seules les identifications obtenues avec tous les membres de la famille CBX, RING1, PCGF2 et la méthyltransférase d'histone SUV420H2 (voir Annexe 4) ont été présentées dans ce manuscrit. Les listes des protéines identifiées par spectrométrie de masse nous ont permis de caractériser les potentiels contaminants de la purification TAP à partir des cellules HeLa dans nos conditions expérimentales.

Les contaminants majeurs, et « ennemis publics » des massistes : les kératines. En effet, bien que différentes isoformes de kératines soient exprimées intrinsèquement dans les cellules humaines, dont les cellules cancéreuses HeLa, nous avons dû faire face à la contamination de nos gels par les kératines présentes dans l'environnement au moment des expériences. Si les kératines endogènes migrent à leurs masses moléculaires respectives dans le gel monodimensionnel (60-70kDa) : les cheveux, poils et peaux mortes de l'expérimentateur peuvent contaminer à tout moment n'importe quelle partie du gel. Nous avons donc établi des règles de manipulation très stricte au laboratoire pour éviter au maximum ce genre de contaminations (gants, charlotte, masque, travail sous hotte...), sans pour autant arriver à s'en affranchir totalement. Ces contaminants, célèbres pour tous les scientifiques qui veulent réaliser des identifications par spectrométrie de masse à partir d'échantillons « humains », n'ont pas été listés dans le tableau récapitulatif des protéines considérées comme contaminants dans le cadre de nos purifications.

Un autre contaminant récurrent est la trypsine. Comme chaque fragment de gel est soumis à l'action de cette endo-protéase, si il y a trop peu de matériel à digérer, la trypsine s'auto-clive et des peptides sont identifiés lors de l'analyse par SM. Elle n'est jamais apparue dans les fragments de gel contenant des protéines identifiées avec plusieurs peptides.

Les membres du cytosquelette font partie des protéines contaminantes. Bien que le protocole utilisé pour la purification des complexes nous permette en théorie de réaliser la purification à partir d'un extrait de protéines nucléaires, 6 familles de protéines décrites comme appartenant au cytosquelette ont été retrouvées dans quasiment toutes les purifications TAP : les actines, les tubulines, les myosines, les spectrines, les dynéines et la titine.

Autres protéines intrinsèques au protocole qui ont parfois été identifiées : la calmoduline et les immunoglobulines. Ces dernières permettent de retenir les protéines fusionnées à l'étiquette TAP lors de la première étape de purification, et certaines d'entre elles peuvent potentiellement se décrocher des billes de sépharose, tout en conservant leur affinité pour la protéine A de l'étiquette. Il en est de même pour la seconde étape de purification où une fraction de la calmoduline fixée sur les billes de sépharose peut également être éluée en même temps que les complexes. D'autre part, les petites protéines de la famille S100 qui ont une affinité pour le calcium ont également été identifiées en tant que contaminants, l'utilisation de fortes concentrations en calcium lors de la seconde étape de la purification peut expliquer un enrichissement en cette famille de protéines. Le même raisonnement peut expliquer la présence de spectrine (SPTB2) qui possède une affinité pour la calmoduline en présence de calcium.

D'autres protéines, associées à nos purifications, ne sont pas apparues comme étant des partenaires protéiques spécifiques d'intérêt : les HSP (Heat Shock Protein) et les GRP (Glucose Related Protein). Ces protéines ubiquitaires sont très largement exprimées dans toutes les cellules et apparaissent d'un intérêt moindre dans le cadre de notre étude (complexes PRC1 et HP1).

Les protéines ribosomales (RS, RL, RO et RT), les ribonucléoprotéines nucléaires (hnRNP), les hélicases d'ARN ATP-dépendante (DDH et DHX) ainsi que la nucléoline participent toutes au même complexe (Fujiyama S, 2002, JBC / Czuby A, 2005, BBA) qui interagit avec les transcrits d'ARN pour réaliser leur maturation et leur traduction, notre étiquette étant placée au niveau amino-terminal des protéines, des formes non matures de la protéine d'intérêt (TAP seul ou avec les premiers acides aminés de la protéine appât) peuvent être associées à ce complexe et être retenues lors des 2 étapes de la purification TAP. Toutes ces protéines ont été identifiées dans les purifications des CBX Polycomb et de SUV420H2.

NIPBL (Nipped B like protein / delangin) a été publiée en tant que partenaire protéique de CBX5 *via* son motif PxVxL (Lechner MS, 2005, BBRC). Etant donné que nous avons identifié cette protéine dans les purifications réalisées avec CBX1 et CBX5, mais également avec CBX2 et CBX8. Nos résultats suggèrent que NIPBL n'est pas exclusivement un partenaire protéique des HP1, mais qu'il

semble également capable d'interagir avec les membres du groupe Polycomb, c'est pourquoi elle a été classée dans les contaminants.

D'autres protéines ont également été identifiées avec les 2 types de protéines HP1 et Polycomb : la protéine à doigts de zinc ZCH11 est présente dans les purifications de CBX1, CBX2 et CBX6, le facteur de transcription EF1A1 et la protéase calpaïne (calcium-dépendante) sont présents dans les purifications de CBX3 et de toutes les CBX Polycomb, la teneurine 2 (protéine membranaire) et FUS (RNA-binding protein) dans CBX5 d'une part et CBX2, CBX4 et CBX7 d'autre part. Les protéines ZCH11, EF1A1 et la calpaïne ont également été retrouvées dans la purification réalisée avec SUV420H2. La 6-phosphofructokinase (K6PP et K6PL) est également ubiquitaire dans nos purifications (HP1, Polycomb et SUV420H2).

En ce qui concerne la Kallikréine 15, identifiée seulement avec CBX6 et CBX8, des peptides pour son identification sont retrouvés dans quasiment toutes les découpes de la piste du gel monodimensionnel. Cette petite protéine de 28kDa ne peut pas migrer à des tailles supérieures dans des conditions dénaturantes (présence de sodium dodécylsulfate). Elle a donc été classée dans les contaminants, au même titre que la N-acétyltransférase ESCO2 et la méthyltransférase d'histones SETB2, entre autres.

Enfin, nous tenons à souligner l'absence de peptides correspondant à la protéase TEV dans les identifications, cela souligne l'efficacité de la seconde étape de purification dans nos conditions expérimentales.

Cette base de données constitue un véritable atout pour notre laboratoire qui dispose de la liste des protéines qui co-éluent avec les appâts fusionnés à l'étiquette TAP dans les cellules HeLa, « collent » aux billes ou à l'étiquette au cours de la purification, ou sont tout simplement trop abondantes dans les extraits pour pouvoir être éliminées complètement, malgré les 2 étapes d'immobilisation spécifique.

Nous tenons à souligner que nous n'avons pas réalisé de purification TAP « contrôle » sans protéine étiquetée ou avec une étiquette TAP seule pour plusieurs raisons :

- nous savions que nous réaliserions les purifications d'au moins 8 protéines (CBX) en duplicata, représentant 2 groupes de protéines ayant des activités distinctes et recrutant des complexes différents, donc les protéines communes pourraient être listées en tant que contaminants.
- bien souvent, les expériences contrôles avec un extrait nucléaire ou total appliqué successivement sur les 2 colonnes d'affinité (immunoglobulines et calmoduline) de la purification TAP ne reflètent pas l'ensemble des contaminants possibles.
- l'expression de l'étiquette TAP seule permet de recruter les éventuelles protéines qui pourraient se lier à cette structure d'une vingtaine de kilodaltons, seulement sa localisation, sans NLS (signal de localisation nucléaire), est cytoplasmique, ce qui limite le nombre de partenaires potentiels. De plus, nous avons travaillé avec des protéines exprimées de manière quasi-exclusive dans le noyau.

- la comparaison d'une purification unique avec un contrôle négatif (extrait cellulaire ou étiquette seule) peut induire en erreur car la présence d'une protéine avec l'appât et son absence dans le contrôle négatif ne signifie pas forcément qu'il s'agit d'un partenaire spécifique à la protéine d'intérêt. Même si des validations biochimiques sont nécessaires, toutes les identifications de toutes les purifications n'ont pas été rigoureusement validées. Par conséquent, nous pouvons raisonnablement penser que certaines des protéines décrites en tant que partenaires peuvent être de faux positifs.

Pour finir, il faut garder à l'esprit que les identifications dépendent de la base de données utilisée pour faire la recherche Mascot, les protéines ayant les scores Mascot les plus élevés sont retrouvées de manière identique entre les différentes bases de données, en revanche, certaines protéines ayant des scores plus faibles (comme parfois les contaminants) peuvent varier suivant les interrogations, c'est pour cela que nous sommes restés fidèles à la base de données Swissprot pour toutes nos identifications *via* le serveur Mascot.

Problème des faux négatifs avec WDR68-HAN11

Lors de notre sélection des protéines co-éluées avec la protéine TAP d'intérêt, nous avons éliminé certaines protéines dont la localisation supposée ou avérée était cytoplasmique (d'après Uniprot et Human Protein Reference Database). C'est le cas de la protéine WDR68 (HAN11 ou AN11) qui contient 5 domaines WD40. Ces domaines sont référencés en tant que plateforme pour l'assemblage de complexes protéiques ou médiateurs transitoires. Nous avons identifié cette protéine de 39kDa avec CBX4, CBX7 et CBX8. Elle est absente des purifications RING1, PCGF2 et HP1. Un screen réalisé chez le poisson zèbre pour identifier les gènes responsables de malformations crâno-faciales a permis d'identifier le gène codant pour WDR68. En effet, la protéine codée par ce gène, exprimée dans le noyau des cellules (transfection transitoire dans des cellules HEK293) et qui co-localise avec la kinase Dyrk-1, est nécessaire pour l'expression de l'endothéline-1. La voie de signalisation impliquant cette dernière est essentielle pour la formation de la mâchoire supérieure. Les auteurs suggèrent que le complexe formé par WDR68 et Dyrk-1 est un régulateur négatif de la transcription (Nissen RM, 2006, BMC Dev Biol). D'autre part, une étude protéomique de HAN11 (purification TAP et GST-pulldown) a permis de confirmer son interaction avec Dyrk-1A mais également avec Dia1 un régulateur du cytosquelette et dont la sur-expression entraîne la translocation de HAN11 du cytoplasme vers le noyau. HAN participerait à la régulation transcriptionnelle de GLI-1 (un des facteurs de transcription de la voie de signalisation Hedgehog) (Morita K, 2006, J Dermatol Sci). Le gène codant pour HAN11 appartient à la liste des 7 gènes considérés comme étant impliqués dans la transformation des tissus prostatiques dans les cancers correspondants (Fujita A, 2008, BMC Syst Biol). Enfin, la protéine HAN11 a été co-éluée dans la purification de CBX8 (Dietrich N, 2007, EMBO J) et des études complémentaires ont permis d'obtenir les résultats suivants: l'utilisation d'un ARN interférent dirigé contre HAN11 réduit la capacité répressive de CBX8 sur un gène rapporteur exprimant la luciférase, mais également sur des cibles endogènes telles que SORBS ou INK4A (PCR quantitative), accompagnée d'une augmentation de p16 (western blot). La purification des complexes associés à HAN11 a permis d'identifier plusieurs membres du PRC1 : CBX4, CBX6, CB8, RING1,

RNF2, PCGF3 et PCGF5 mais également des membres du complexes NuRD et des facteurs de transcription. L'interaction avec CBX4, CBX8 a été validée par co-immunoprécipitation, BMI-1 (autre membre du PRC1) retient également HAN11 par IP. Enfin, lors de la différenciation des cellules NT2-D1 par l'acide rétinoïque, les gènes HoxA1, Oct4 et Nanog sont réprimés. Dans des cellules traitées avec des ARN interférents dirigés contre HAN11, la répression des ces 3 gènes est réduite. Le Dr Klaus H. Hansen (BRIC, Copenhague), qui supervise ces travaux, suggère que HAN11 joue un rôle dans la différenciation des cellules (communication personnelle).

Ces différents éléments nous permettent de relativiser les informations suggérées par les outils bioinformatiques gérant des bases de données importantes. Les informations fournies ne sont pas forcément erronées mais ne sont malheureusement parfois pas toujours fiables et une étude bibliographique complémentaire est indispensable pour la validation des résultats.

CONCLUSIONS

Analyse des protéines CBX associées au PRC1

Dans la première partie de ce manuscrit, nous avons étudié les protéines CBX-Polycomb (CBX2, CBX4, CBX6, CBX7 et CBX8) (Tajul-Arifin K, 2003, Genome Res). Dans un premier temps nous avons voulu confirmer qu'elles se comportaient en tant que répresseurs transcriptionnels. Pour ce faire, nous avons utilisé des protéines fusionnées à l'étiquette GAL4 et étudié leur activité sur un gène rapporteur dont la séquence promotrice contenait le site de fixation de la séquence GAL4. Ainsi nous avons pu mettre en évidence que ces protéines étaient capables de réprimer différents promoteurs (fort ou faible, exogène ou endogène) et que leurs activités étaient différentes : CBX4 et CBX6 étant respectivement la plus et la moins répressive. Ces résultats suggèrent qu'elles recrutent des complexes de compositions différentes dans la cellule. Cette étude préliminaire nous a permis de valider les outils cellulaires que nous avons choisis, en particulier le système inductible.

Nous avons ensuite généré des lignées cellulaires, à partir de cellules HeLa, qui exprimaient de manière stable ces différentes protéines fusionnées à l'étiquette TAP (par infection rétrovirale). Une étude de la localisation des protéines TAP nous a permis de confirmer qu'elles étaient exclusivement nucléaires. Cependant, leurs localisations précises ne seraient pas redondantes au sein des noyaux de cellules souches de souris, suggérant qu'elles ciblent chacune des gènes de manière spécifique (Ren X, 2008, Mol Cell Biol). Donc leur recrutement supposé au niveau des tri-méthyl-lysines des histones H3 dépendrait davantage de leurs partenaires protéiques plutôt que des CBX elles-mêmes et leur chromodomaine en particulier.

Dans l'optique de la caractérisation des complexes associés aux CBX Polycomb, nous avons évalué leurs masses moléculaires à l'aide d'une chromatographie d'exclusion stérique adaptée pour l'étude des complexes de haut poids moléculaires (gamme dynamique de 5 à 5000 kDa). Nous avons observés que ces 5 protéines participent à deux types de complexes : un premier compris entre 200 et 500kDa et un second, majoritaire, de plus haut poids moléculaire (supérieur à 1000kDa). Cependant l'interprétation des résultats obtenus est à relativiser car la gel filtration évalue davantage un encombrement stérique plutôt qu'une masse moléculaire précise. D'autre part aucune standardisation des tampons à utiliser n'a été adoptée par la communauté scientifique, et une étude datant de 2002 suggère que le PRC1 correspond a un complexe de 500kDa (Levine SS, 2002, MCB). Nos résultats confirment que les protéines CBX recrutent des complexes d'encombrement stérique important, et que la grande majorité des protéines TAP sont retrouvées dans des complexes car la forme libre n'est pas détectée par western blot.

Nous avons réalisé en duplicata les purifications TAP de ces 5 protéines. Bien que nous soyons parti de la même quantité de cellules au départ, les quantités de protéines co-éluées sont variables, mais également les protéines associées au PRC1 (Schwartz YB, 2007, Nat Rev Genet). Toutes les protéines CBX recrutent des protéines du PRC1, dont les plus abondantes sont : BMI1, PCGF2, PCGF6, RING1, RNF2 et PHC2. D'une part les protéines CBX Polycomb apparaissent comme mutuellement exclusives, il s'agit de la première étude biochimique systématique de toutes les CBX

Polycomb mettant en évidence leur exclusivité au sein des PRC1. D'autre part les complexes PRC1 recrutés sont de composition différente. De plus, bien que la plupart des protéines co-éluées avec certains membres de la famille CBX (CBX4, CBX7 et CBX8) appartiennent au PRC1, les protéines identifiées avec CBX2 et CBX6 ne sont pas particulièrement des membres du PRC1. Certaines protéines, en particulier les sous-unités de la caséine kinase 2, sont retrouvées dans la plupart des purifications, suggérant que cette protéine pourrait être un nouveau partenaire des complexes PRC1. D'autres protéines telles que l'hélicase MOV10, le facteur de transcription YBOX1 sont identifiées avec plusieurs CBX Polycomb. Ces purifications permettent également de mettre en évidence des partenaires spécifiques de chacune des CBX : PSME3 avec CBX2, 14-3-3 et l'isoforme longue CBX4 avec CBX4 et RPP5 ainsi que ZFR avec CBX8. Des études biochimiques complémentaires permettraient de valider ou non ces potentielles interactions.

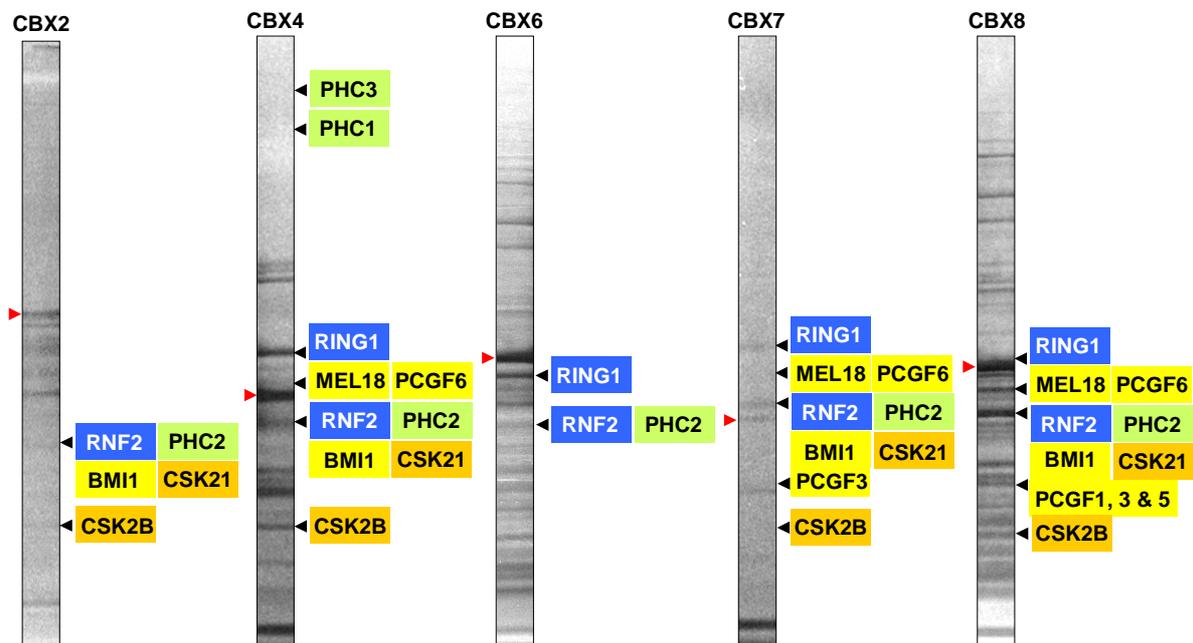


Figure 52 – Bilan des protéines associées au PRC1 identifiées avec les CBX Polycomb
L'appât est indiqué avec une flèche rouge, chaque couleur correspond à une famille différente.

Nous avons entrepris certaines validations à l'aide de co-immunoprécipitations et de GST-pulldowns, qui nous ont permis de mettre en évidence les interactions directes entre les CBX et les membres de la famille PCGF mais aussi que l'interaction entre les CBX et les PHC était vraisemblablement indirecte.

Nous avons ensuite focalisé nos études sur la caractérisation de l'interaction spécifique entre CBX4 et les 14-3-3, compte tenu de leur intensité dans la purification, nous avons choisi d'étudier ces petites protéines régulatrices ubiquitaires. La protéine CBX4 contient un site de fixation de mode1 des 14-3-3, nous n'avons malheureusement pas réussi à démontrer la présence d'une phosphorylation sur la sérine (S164) impliquée dans ce motif. Cependant, nous avons pu valider l'interaction spécifique directe entre CBX4 et 14-3-3 et la perte de cette interaction lors de la mutation de la sérine 164. De plus, la protéine CBX4 mutée semble déstabiliser la composition du complexe PRC1 recruté par CBX4 car, tandis que RNF2 est identiquement retenue par coIP entre la protéine mutée et la sauvage, PCGF6 n'est plus retenue avec la mutée. Il a été rapporté que PCGF2 pouvait être phosphorylé et

que cette modification influençait l'activité du PRC1 (Elderkin S, 2007, Mol Cell), nous suggérons que d'autres membres des PRC1, en particulier CBX4, peuvent subir ce genre de modifications post-traductionnelles.

Analyse des protéines non-CBX associées au PRC1

Parallèlement à l'étude des CBX Polycomb, nous avons étudié quelques membres du PRC1. Nous avons étudié l'activité répressive de RING1, RNF2 et 2 membres de la famille PCGF (2 et 6). Toutes ces protéines possèdent des activités différentes sur les différents promoteurs testés, confirmant l'idée qu'il existe plusieurs complexes PRC1 différents dans les cellules HeLa. Les masses moléculaires apparentes des complexes associés aux protéines RING1-TAP et RNF2-TAP sont majoritairement présents dans une gamme comprise entre 200 et 500kDa, soulignant que ces protéines participent majoritairement à des complexes ayant ces masses, sans pour autant exclure leur participation aux complexes de plus haut poids moléculaires observés avec les protéines CBX. Les purifications TAP de PCGF2 et RNF2 nous permettent de tirer plusieurs enseignements : à l'exception de CBX8, les protéines CBX ne sont pas identifiées avec RING1 alors que l'inverse est vrai. Cela suggère que la grande majorité des protéines RING1 n'est pas associée aux CBX alors que le cœur du PRC1 est composé de : BMI1, PHC2, PCGF2 et RING1, les protéines retrouvées dans toutes les purifications. Nous avons remarqué l'absence d'autres membres de la famille PCGF dans la purification PCGF2, suggérant comme pour les CBX, une exclusivité d'un membre par rapport aux autres dans les PRC1. Aussi, les identifications de YAF2 et RYBP suggèrent que ces protéines sont présentes dans les PRC1 qui ne possèdent pas de CBX car ces protéines n'ont pas été identifiées avec ces dernières. Pour finir, nous retrouvons la caséine kinase 2 qui semble bien faire partie des PRC1 dans les cellules HeLa.

Analyse des protéines HP1

A partir des mêmes outils cellulaires ou biochimiques, développés et mis au point pour les protéines Polycomb, nous avons entrepris la caractérisation des complexes associés aux 3 protéines HP1 humaines (Lomberck G, 2006, Genome Biol). Ces 3 CBX possèdent des activités répressives voisines sur les différents promoteurs testés. L'immunomarquage des lignées TAP réalisées avec les différentes isoformes souligne leur localisation nucléaire, en particulier aux niveaux des régions hétérochromatiques pour CBX1 et CBX3. Les protéines co-éluées avec les HP1 ont pour la plupart déjà été décrites dans la littérature : comme TIF1b, les histones, les autres HP1, ADNP, CHD4, KU70 ou encore SMCA3 et 4. En revanche, nous rapportons pour la première fois l'interaction entre les HP1 et : les importines- α et pogoZ. De plus, certaines protéines identifiées semblent spécifiques à une isoforme particulière : RPPB4 avec CBX5, SUHW4, ZNF8 et K1802 avec CBX3 et pour finir ADNP2 avec CBX1.

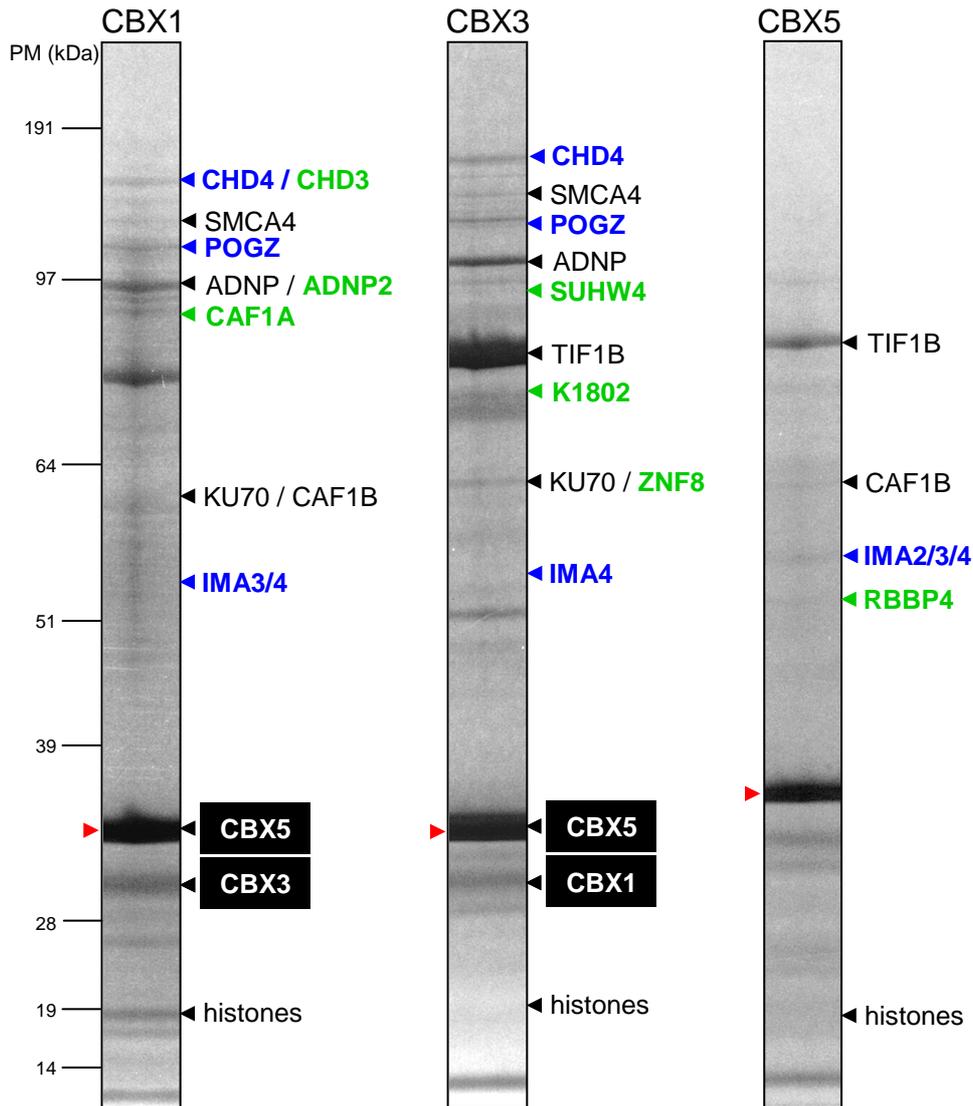


Figure 53 – Bilan des protéines majoritaires co-éluées avec les 3 isoformes HP1 par purification TAP
 Flèche rouge = appât. En noir = partenaires connus. En bleu = nouveaux partenaires retrouvés dans
 avec les 3 isoformes. En vert = partenaires spécifiques à une isoforme.

Bilan des purifications TAP réalisées

Bien qu'un nombre non négligeable de validations biochimiques sont requises, les purifications TAP réalisées avec les 8 membres de la famille CBX dans des cellules HeLa en culture nous ont permis de définir les premiers éléments nécessaires à la caractérisation des complexes PRC1 et HP1 dans les cellules de mammifères. En effet, ces données sont une base solide de travail pour des caractérisations fonctionnelles futures qui permettront de mettre en lumière le rôle joué par les différents potentiels nouveaux partenaires protéiques décrits dans ce manuscrit. Pour les Polycomb, la caséine kinase 2 est une de nos futures cibles d'étude privilégiées tandis que pour les HP1, c'est la transposase pogZ qui a attiré notre attention.

PERSPECTIVES

A court terme :

Validation des nouveaux partenaires protéiques par les techniques biochimiques classiques et usuelles:

- double-hybride (domaines d'interaction)
- co-localisation par immunofluorescence
- co-immunoprécipitation (protéines endogènes si anticorps disponibles ou exogènes étiquetées)
- GST-pulldown / extrait nucléaire ou protéine traduite *in vitro* pour interaction directe.

Validation des interactions par des outils techniquement plus élaborés comme le BiFC (bimolecular fluorescence complementation) ou le FRET en collaboration avec l'équipe de Biophotonique dirigée par Laurent Héliot (Institut de Recherches Interdisciplinaires, Lille).

Identification de la phosphorylation potentielle de CBX4 sur la sérine 164 *in vivo* par spectrométrie de masse à partir d'une protéine CBX4 sur-exprimée dans les cellules et enrichie avec une étape de purification (billes IgG avec étiquette TAP par exemple).

A moyen terme :

Identification de la kinase responsable de la phosphorylation de CBX4

Si la phosphorylation de CBX4 est avérée *in vivo*, il serait intéressant de déterminer l'enzyme qui catalyse cette réaction à l'aide des nombreux kits de réaction de phosphorylation *in vitro* disponibles commercialement.

Rôle fonctionnel - modèle *in vivo*

Compte tenu de l'importance des protéines Polycomb pour le développement des organismes supérieurs, il serait intéressant de pouvoir invalider ou surexprimer les protéines partenaires candidates. Certains organismes modèles comme la drosophile, le poisson zèbre ou le nématode sont particulièrement bien adaptés pour ce type d'études ou les manipulations génétiques sont plus aisées qu'avec la souris. Aussi, comme nous l'avons dans l'introduction les protéines identifiées possèdent des homologues dans les organismes proposés. Le laboratoire envisage de mettre en œuvre une animalerie équipée d'aquariums pour une étude des partenaires protéiques identifiés par morpholinos chez le poisson-zèbre. D'autre part, l'équipe dirigée par Yasushi Saka (Institut de Recherches Interdisciplinaires, Lille), qui étudie la différenciation cellulaire et la prolifération cellulaire chez le Xénope, pourrait nous permettre de faire certaines études chez cet autre organisme modèle.

Optimisation de notre purification TAP

- Purifier les complexes à partir de cellules synchronisées. Les purifications présentées dans ce manuscrit proviennent de populations hétérogènes, et si les partenaires protéiques des complexes varient au cours du cycle cellulaire, il serait intéressant de pouvoir réaliser des purifications sur des cellules bloquées à certaines phases importantes de leur cycle. Mais compte tenu de la prolifération anarchique des cellules cancéreuses (dont les cellules HeLa), une telle expérience n'est peut-être pas évidente à mettre en œuvre. Cependant, certains partenaires spécifiques ayant une interaction transitoire avec la protéine d'intérêt au cours du cycle cellulaire, non identifiés dans notre étude car présents en trop faible quantité, pourraient être révélés.

- Purifier des complexes à partir de cellules différenciables, avant et après différenciation. Il existe différents modèles cellulaires qui peuvent se différencier en cellules épidermales ou neuronales sous l'effet d'un agent physiologique (comme les cellules NT2-D1 traitées avec de l'acide rétinoïque). La purification des complexes avant et après l'induction pourrait permettre de comprendre comment s'organisent les complexes, en particulier les PRC1, qui sont impliqués dans le développement. Aussi, la validation des nouveaux partenaires protéiques dans une autre lignée cellulaire permettrait de valider leur participation au complexe.

- Digestion protéolytique des éluions TAP par une autre enzyme que la trypsine. Bien que cette dernière soit l'enzyme de choix pour les identifications par spectrométrie de masse car elle fournit en moyenne des peptides de tailles compatibles avec cette technologie, il existe d'autres endoprotéases qui coupent avant ou après d'autres résidus que les lysine/arginine comme Glu-C, Arg-C, Asp-N ou la chymotrypsine. Leur utilisation, unique ou combinée avec la trypsine, permettrait d'identifier des peptides supplémentaires à partir des purifications, mais aussi d'augmenter potentiellement la couverture de séquence des partenaires identifiés et de l'appât. Cette approche a été adoptée avec succès pour atteindre des couvertures de séquences de l'ordre de 95% et permettre l'identification des modifications post-traductionnelles des 3 isoformes HP1 étiquetées avec un flag et sur-exprimées dans des cellules HEK293 (LeRoy G, 2009, MCP).

- Utiliser une autre technologie TAP pour les PRC1 : split-TAP avec une CBX et BMI1 ou RING1 car, d'après nos résultats ce sont les protéines les plus stables dans les PRC1 purifiés. Cette technologie permettrait d'enrichir en partenaires spécifiques du PRC1, et notamment de statuer sur la place de 14-3-3 dans le PRC1-CBX4.

Non redondance des CBX Polycomb dans les PRC1... et cibles exclusives ?

- Les gènes ciblés par les complexes PRC1 ont été identifiés grâce à des études de ChIP on chip ou ChIP seq. Ces études ont été réalisées avec une ou plusieurs protéines du complexe, aujourd'hui aucune étude n'a permis de comparer les localisations d'au moins 2 CBX entre elles. Nous supposons

que certaines cibles sont communes comme le locus INK4A-ARF (CBX7 et CBX8) mais que certaines sont spécifiques d'une isoforme particulière. Bien qu'une étude par ChIP seq soit onéreuse, des alternatives moins coûteuses peuvent être envisagées comme des immunoprécipitations de la chromatine avec toutes les CBX Polycomb en ciblant les promoteurs de gènes décrits dans la littérature comme cible du PRC1, en particulier celles de BMI (Kallin EM, 2009, PLoS Genet). Une comparaison des différentes CBX permettrait de statuer sur la présence et/ou l'absence des unes et des autres.

- Les cellules murines L929 présentent la particularité d'avoir des corps Polycomb bien définis. Une étude systématique par immunofluorescence *in vivo* de toutes les CBX Polycomb permettrait de statuer sur leur colocalisation ou non au niveau de ces structures particulières du noyau.

A long terme ou en collaboration :

Prochains défis pour l'analyse protéomique

- Identifier les modifications post-traductionnelles de l'appât et éventuellement des protéines co-éluées. Toutes les protéines sont potentiellement les cibles de modifications post-traductionnelles et leur caractérisation ouvre de nouvelles opportunités pour comprendre la régulation des complexes étudiés. La phosphorylation de MEL-18 augmente l'activité de RING1B sur H2A (Elderkin S, 2007, Mol Cell). RNF2 peut être phosphorylée par p38 MAPK et ERK1/2 (Rao PS, 2009, Proteomics). Les protéines HP1 sont aujourd'hui au cœur de la théorie de l' « histone subcode », leurs ADP-ribosylation (Quénet D, 2008, FASEB J) et phosphorylation (Lomberck G, 2006, Nat Cell Biol) ont déjà été mises en évidence, de même que leurs méthylation, acétylation et formylation (LeRoy G, 2009, MCP).

- Augmenter les quantités de protéines éluées afin d'identifier de potentiels nouveaux partenaires qui auraient été présents en trop faible quantité lors de nos purifications pour pouvoir être listés. Une alternative serait d'étudier les mêmes quantités de matériel avec un spectromètre de masse ayant une sensibilité supérieure à la trappe ionique utilisée : LCQ Deca XP+ (Thermo electron) datant de fin 2002.

- Exprimer une protéine TAP au même niveau que la protéine endogène (comme pour les purifications TAP réalisées dans la levure). Bien que nos résultats soient en accord avec les précédents partenaires protéiques décrits dans la littérature, nous savons que les protéines fusionnées à l'étiquette TAP sont sur-exprimées par rapport à l'endogène. La sélection de clones après l'infection virale permettrait d'obtenir une lignée stable qui exprime la protéine TAP à un niveau proche de celui de l'endogène.

- Protéomique quantitative : un laboratoire japonais a développé une purification TAP associée à des isotopes lourd et léger de Leucine pour évaluer les protéines stables et dynamiques du complexe eIF2

à partir de levures (Kito K, 2008, Proteomics). Cette technologie est applicable à des cellules en culture.

- Développement de la technologie TAP : comme nous l'avons vu dans l'introduction, d'autres étiquettes TAP ont été développées et la dernière génération applicable aux cellules de mammifères s'appelle le GS-TAP (Bürckstümmer T, 2006, Nat Methods). Comme le plasmide qui code pour cette étiquette est disponible, le laboratoire pourrait tester une purification avec un appât déjà étudié et comparer les identifications.

- Analyses structurales qui permettraient de déterminer la stœchiométrie de chacune des sous-unités du complexe (Zhang Y, 2006, PNAS / Aloy P, 2004, Science). La spectrométrie RMN est un outil efficace pour déterminer les structures tridimensionnelles des protéines, et la résolution des protéines étudiées permettrait de modéliser la structure des complexes PRC1 et HP1.

ANNEXES

Annexe 1 – Listes de tous les peptides identifiés dans les purifications TAP

Tableau S1 : Peptides identifiés dans les 2 purifications CBX2-TAP (n°15 et 18)

band	hit	accession	description	mass	score	peptide
15E03.txt	3	JHD2A_HUMAN	JmjC domain-containing histone demethylation protein	147234	49	FLGDQQRVFLSK
15E03.txt	3	JHD2A_HUMAN	JmjC domain-containing histone demethylation protein	147234	49	KRRVIEVYSLLR
15E03.txt	3	JHD2A_HUMAN	JmjC domain-containing histone demethylation protein	147234	49	LQVKNVIYHAVKDAVAMLK
15E03.txt	4	ZN167_HUMAN	Zinc finger protein 167	84979	48	ELEGQTSDEEGSR
15E03.txt	4	ZN167_HUMAN	Zinc finger protein 167	84979	48	GNLGLIPRSTAFQK
15E09.txt	18	MYCB2_HUMAN	Probable E3 ubiquitin-protein ligase MYCBP2	509759	39	SMMCPPGMHK
15E09.txt	18	MYCB2_HUMAN	Probable E3 ubiquitin-protein ligase MYCBP2	509759	39	SMMCPPGMHK
15E09.txt	18	MYCB2_HUMAN	Probable E3 ubiquitin-protein ligase MYCBP2	509759	39	IAGQISASVAQQR
15E09.txt	18	MYCB2_HUMAN	Probable E3 ubiquitin-protein ligase MYCBP2	509759	39	SDGRMPSSRAESPGPGRS
15E09.txt	18	MYCB2_HUMAN	Probable E3 ubiquitin-protein ligase MYCBP2	509759	39	NEEHCLPCLHGCDKATSCLK
15E13.txt	1	CBX2_HUMAN	Chromobox protein homolog 2	56046	265	LLLAFQK
15E13.txt	1	CBX2_HUMAN	Chromobox protein homolog 2	56046	265	ESPTSVGVFFNLR
15E13.txt	1	CBX2_HUMAN	Chromobox protein homolog 2	56046	265	AQILVAKPELKDPIR
15E13.txt	1	CBX2_HUMAN	Chromobox protein homolog 2	56046	265	HNSWEPEENILDPR
15E13.txt	1	CBX2_HUMAN	Chromobox protein homolog 2	56046	265	GTGSLIGASGATMPTDTSK
15E13.txt	1	CBX2_HUMAN	Chromobox protein homolog 2	56046	265	GTGSLIGASGATMPTDTSK
15E13.txt	1	CBX2_HUMAN	Chromobox protein homolog 2	56046	265	VPAGCPGQPAPTQELSLQVLDLQSVK
15E14.txt	1	CBX2_HUMAN	Chromobox protein homolog 2	56046	583	LEVLVK
15E14.txt	1	CBX2_HUMAN	Chromobox protein homolog 2	56046	583	LLLAFQK
15E14.txt	1	CBX2_HUMAN	Chromobox protein homolog 2	56046	583	CGLGLDLK
15E14.txt	1	CBX2_HUMAN	Chromobox protein homolog 2	56046	583	NGMPGVLLAR
15E14.txt	1	CBX2_HUMAN	Chromobox protein homolog 2	56046	583	AQILVAKPELK
15E14.txt	1	CBX2_HUMAN	Chromobox protein homolog 2	56046	583	ESPTSVGVFFNLR
15E14.txt	1	CBX2_HUMAN	Chromobox protein homolog 2	56046	583	LPPPLSAPVAGLAALK
15E14.txt	1	CBX2_HUMAN	Chromobox protein homolog 2	56046	583	AQILVAKPELKDPIR
15E14.txt	1	CBX2_HUMAN	Chromobox protein homolog 2	56046	583	AQILVAKPELKDPIR
15E14.txt	1	CBX2_HUMAN	Chromobox protein homolog 2	56046	583	HNSWEPEENILDPR
15E14.txt	1	CBX2_HUMAN	Chromobox protein homolog 2	56046	583	GGISWQSSIVHYMNR
15E14.txt	1	CBX2_HUMAN	Chromobox protein homolog 2	56046	583	GTGSLIGASGATMPTDTSK
15E14.txt	1	CBX2_HUMAN	Chromobox protein homolog 2	56046	583	AQILVAKPELKDPIR
15E14.txt	1	CBX2_HUMAN	Chromobox protein homolog 2	56046	583	GTGSLIGASGATMPTDTSK
15E14.txt	1	CBX2_HUMAN	Chromobox protein homolog 2	56046	583	VPAGCPGQPAPTQELSLQVLDLQSVK
15E14.txt	1	CBX2_HUMAN	Chromobox protein homolog 2	56046	583	VPAGCPGQPAPTQELSLQVLDLQSVK
15E14.txt	10	PAR12_HUMAN	Poly [ADP-ribose] polymerase 12	79013	44	MGLSADALER
15E14.txt	10	PAR12_HUMAN	Poly [ADP-ribose] polymerase 12	79013	44	GDGPHGSCAFQK
15E15.txt	4	CBX2_HUMAN	Chromobox protein homolog 2	56046	162	AQILVAKPELKDPIR
15E15.txt	4	CBX2_HUMAN	Chromobox protein homolog 2	56046	162	GTGSLIGASGATMPTDTSK
15E15.txt	4	CBX2_HUMAN	Chromobox protein homolog 2	56046	162	AQILVAKPELKDPIR
15E15.txt	4	CBX2_HUMAN	Chromobox protein homolog 2	56046	162	VPAGCPGQPAPTQELSLQVLDLQSVK
15E15.txt	8	YD007_HUMAN	Coiled-coil domain-containing protein FLJ25770	127062	41	NQAMSSMIR
15E15.txt	8	YD007_HUMAN	Coiled-coil domain-containing protein FLJ25770	127062	41	EKELSLEKEQNK
15E15.txt	8	YD007_HUMAN	Coiled-coil domain-containing protein FLJ25770	127062	41	VAMGMQKQITAKR
15E15.txt	8	YD007_HUMAN	Coiled-coil domain-containing protein FLJ25770	127062	41	ERAIEATNAEITKLR
15E16.txt	3	YD007_HUMAN	Coiled-coil domain-containing protein FLJ25770	127062	41	EKELSLEKEQNK
15E16.txt	3	YD007_HUMAN	Coiled-coil domain-containing protein FLJ25770	127062	41	MTLESSERTISDLTSLQEK
15E17.txt	5	CBX2_HUMAN	Chromobox protein homolog 2	56046	100	HNSWEPEENILDPR
15E17.txt	5	CBX2_HUMAN	Chromobox protein homolog 2	56046	100	GTGSLIGASGATMPTDTSK
15E19.txt	21	CBX2_HUMAN	Chromobox protein homolog 2	56046	41	HNSWEPEENILDPR
15E20.txt	6	IQWD1_HUMAN	Nuclear receptor interaction protein	96292	48	MLGTR
15E20.txt	6	IQWD1_HUMAN	Nuclear receptor interaction protein	96292	48	EMEELDTLNIIRPLVK
15E22.txt	4	CSK21_HUMAN	Casein kinase II subunit alpha	45115	136	QLYQTLTDYDIR
15E22.txt	4	CSK21_HUMAN	Casein kinase II subunit alpha	45115	136	TPALVFEHVNNDFDK
15E22.txt	2	RING2_HUMAN	E3 ubiquitin-protein ligase RING2	37632	281	FCADCITLALR
15E22.txt	2	RING2_HUMAN	E3 ubiquitin-protein ligase RING2	37632	281	SLRPDPNFDALISK
15E22.txt	2	RING2_HUMAN	E3 ubiquitin-protein ligase RING2	37632	281	SLRPDPNFDALISK
15E22.txt	2	RING2_HUMAN	E3 ubiquitin-protein ligase RING2	37632	281	IYPSRDEYEAHQER
15E22.txt	2	RING2_HUMAN	E3 ubiquitin-protein ligase RING2	37632	281	TPQEAITDGLEIVSPR
15E23.txt	2	CBX2_HUMAN	Chromobox protein homolog 2	56046	72	HNSWEPEENILDPR
15E23.txt	2	CBX2_HUMAN	Chromobox protein homolog 2	56046	72	GTGSLIGASGATMPTDTSK
15E23.txt	1	RING2_HUMAN	E3 ubiquitin-protein ligase RING2	37632	95	FCADCITLALR
15E23.txt	1	RING2_HUMAN	E3 ubiquitin-protein ligase RING2	37632	95	SLRPDPNFDALISK
15E24.txt	1	CBX2_HUMAN	Chromobox protein homolog 2	56046	108	LLLAFQK
15E24.txt	1	CBX2_HUMAN	Chromobox protein homolog 2	56046	108	AQILVAKPELKDPIR
15E24.txt	1	CBX2_HUMAN	Chromobox protein homolog 2	56046	108	AQILVAKPELKDPIR
15E24.txt	1	CBX2_HUMAN	Chromobox protein homolog 2	56046	108	HNSWEPEENILDPR
15E24.txt	1	CBX2_HUMAN	Chromobox protein homolog 2	56046	108	AQILVAKPELKDPIR
15E25.txt	3	CBX2_HUMAN	Chromobox protein homolog 2	56046	86	LLLAFQK
15E25.txt	3	CBX2_HUMAN	Chromobox protein homolog 2	56046	86	AQILVAKPELKDPIR

15E25.txt	3	CBX2_HUMAN	Chromobox protein homolog 2	56046	86	HNSWEPEENILDP
15E25.txt	1	PSME3_HUMAN	Proteasome activator complex subunit 3	29488	280	YPHVEDYR
15E25.txt	1	PSME3_HUMAN	Proteasome activator complex subunit 3	29488	280	SNQQLVDIIK
15E25.txt	1	PSME3_HUMAN	Proteasome activator complex subunit 3	29488	280	NQYVTLHDMILK
15E25.txt	1	PSME3_HUMAN	Proteasome activator complex subunit 3	29488	280	TVSEEAASYLDQISR
15E25.txt	1	PSME3_HUMAN	Proteasome activator complex subunit 3	29488	280	TVTIDEKEYISLR
15E26.txt	3	CBX2_HUMAN	Chromobox protein homolog 2	56046	95	LLAFQK
15E26.txt	3	CBX2_HUMAN	Chromobox protein homolog 2	56046	95	HNSWEPEENILDP
15E26.txt	3	CBX2_HUMAN	Chromobox protein homolog 2	56046	95	HNSWEPEENILDP
15E26.txt	4	RN213_HUMAN	RING finger protein 213	373742	57	RSLETNGEINLPK
15E26.txt	4	RN213_HUMAN	RING finger protein 213	373742	57	IYGDVPTFLPHLPR
15E26.txt	4	RN213_HUMAN	RING finger protein 213	373742	57	MKMQLNVKNVPLKTIR
15E26.txt	4	RN213_HUMAN	RING finger protein 213	373742	57	MKMQLNVKNVPLKTIR
15E26.txt	4	RN213_HUMAN	RING finger protein 213	373742	57	HKGEMAYIVVQNHMNLSENANNNVPSWK
15E31.txt	17	ANR26_HUMAN	Ankyrin repeat domain-containing protein 26	196202	39	ELCSLR
15E31.txt	17	ANR26_HUMAN	Ankyrin repeat domain-containing protein 26	196202	39	LKHKIDDLTAELETAGSKCLHLDTK
15E33.txt	2	LBA1_HUMAN	Lupus brain antigen 1 homolog	336007	64	LALAHDTALSMKSK
15E33.txt	2	LBA1_HUMAN	Lupus brain antigen 1 homolog	336007	64	ALKRVTNNAEKEMVK
15E33.txt	2	LBA1_HUMAN	Lupus brain antigen 1 homolog	336007	64	ELGDVFSIIQEYKPK
18E14.txt	3	GTF2I_HUMAN	General transcription factor II-I	112346	211	EQVNDLFSR
18E14.txt	3	GTF2I_HUMAN	General transcription factor II-I	112346	211	FAQALGLTEAVK
18E14.txt	3	GTF2I_HUMAN	General transcription factor II-I	112346	211	ANELPQPPVPEPANACK
18E14.txt	3	GTF2I_HUMAN	General transcription factor II-I	112346	211	DQSAVVVQGLPEGVAFK
18E14.txt	5	PARP1_HUMAN	Poly [ADP-ribose] polymerase 1	113012	113	AEPVEVVAPR
18E14.txt	5	PARP1_HUMAN	Poly [ADP-ribose] polymerase 1	113012	113	SDAYYCTGDVTAWTK
18E21.txt	4	CBX2_HUMAN	Chromobox protein homolog 2	56046	159	ESPTSVGFFNLR
18E21.txt	4	CBX2_HUMAN	Chromobox protein homolog 2	56046	159	GTGSLIGASGATMPTDTSK
18E21.txt	4	CBX2_HUMAN	Chromobox protein homolog 2	56046	159	GTGSLIGASGATMPTDTSK
18E22.txt	1	CBX2_HUMAN	Chromobox protein homolog 2	56046	379	NGMPGVGLLAR
18E22.txt	1	CBX2_HUMAN	Chromobox protein homolog 2	56046	379	NGMPGVGLLAR
18E22.txt	1	CBX2_HUMAN	Chromobox protein homolog 2	56046	379	AQILVAKPELK
18E22.txt	1	CBX2_HUMAN	Chromobox protein homolog 2	56046	379	ESPTSVGFFNLR
18E22.txt	1	CBX2_HUMAN	Chromobox protein homolog 2	56046	379	LPPLSAPVAGLAALK
18E22.txt	1	CBX2_HUMAN	Chromobox protein homolog 2	56046	379	GTGSLIGASGATMPTDTSK
18E22.txt	1	CBX2_HUMAN	Chromobox protein homolog 2	56046	379	GTGSLIGASGATMPTDTSK
18E22.txt	1	CBX2_HUMAN	Chromobox protein homolog 2	56046	379	VPAGCPGPQAPTQELSLQVLDLQSVK
18E23.txt	4	CBX2_HUMAN	Chromobox protein homolog 2	56046	205	ESPTSVGFFNLR
18E23.txt	4	CBX2_HUMAN	Chromobox protein homolog 2	56046	205	HNSWEPEENILDP
18E23.txt	4	CBX2_HUMAN	Chromobox protein homolog 2	56046	205	GTGSLIGASGATMPTDTSK
18E23.txt	4	CBX2_HUMAN	Chromobox protein homolog 2	56046	205	VPAGCPGPQAPTQELSLQVLDLQSVK
18E25.txt	12	CENPE_HUMAN	Centromeric protein E	316219	62	LHENLEEMK
18E25.txt	12	CENPE_HUMAN	Centromeric protein E	316219	62	DNLRRVEETLK
18E25.txt	12	CENPE_HUMAN	Centromeric protein E	316219	62	FQESQEEIKSLTK
18E26.txt	13	CENPE_HUMAN	Centromeric protein E	316219	43	SENLEKEK
18E26.txt	13	CENPE_HUMAN	Centromeric protein E	316219	43	LHENLEEMK
18E26.txt	13	CENPE_HUMAN	Centromeric protein E	316219	43	FQESQEEIKSLTK
18E27.txt	19	CBX2_HUMAN	Chromobox protein homolog 2	56046	45	GTGSLIGASGATMPTDTSK
18E32.txt	9	CSK21_HUMAN	Casein kinase II subunit alpha	45115	46	QLYQTLTDYDIR
18E33.txt	4	BMI1_HUMAN	Polycomb complex protein BMI-1	36925	201	TLQDIVYK
18E33.txt	4	BMI1_HUMAN	Polycomb complex protein BMI-1	36925	201	CPAAMTVMHLR
18E33.txt	4	BMI1_HUMAN	Polycomb complex protein BMI-1	36925	201	SDKTLQDIVYK
18E33.txt	4	BMI1_HUMAN	Polycomb complex protein BMI-1	36925	201	DGLTNAGELESDSGSK
18E33.txt	17	CBX2_HUMAN	Chromobox protein homolog 2	56046	41	HNSWEPEENILDP
18E33.txt	17	CBX2_HUMAN	Chromobox protein homolog 2	56046	41	GGISWQSSIVHYMNR
18E33.txt	10	CSK21_HUMAN	Casein kinase II subunit alpha	45115	66	QLYQTLTDYDIR
18E33.txt	2	PHC2_HUMAN	Polyhomeotic-like protein 2	90657	350	AQEIDGQALLLK
18E33.txt	2	PHC2_HUMAN	Polyhomeotic-like protein 2	90657	350	DLVGMGHFLPSEPTK
18E33.txt	2	PHC2_HUMAN	Polyhomeotic-like protein 2	90657	350	RASKASLPPLTK
18E33.txt	2	PHC2_HUMAN	Polyhomeotic-like protein 2	90657	350	SLPGCQIEAEFFR
18E33.txt	2	PHC2_HUMAN	Polyhomeotic-like protein 2	90657	350	SSLLVGNLK
18E33.txt	2	PHC2_HUMAN	Polyhomeotic-like protein 2	90657	350	WNVEDVYEFIR
18E33.txt	2	PHC2_HUMAN	Polyhomeotic-like protein 2	90657	350	YAQGFLEPK
18E33.txt	1	RING2_HUMAN	E3 ubiquitin-protein ligase RING2	37632	429	FCADCIITLR
18E33.txt	1	RING2_HUMAN	E3 ubiquitin-protein ligase RING2	37632	429	TWELSLYELQR
18E33.txt	1	RING2_HUMAN	E3 ubiquitin-protein ligase RING2	37632	429	VNKPMEIYYAPTK
18E33.txt	1	RING2_HUMAN	E3 ubiquitin-protein ligase RING2	37632	429	SLRPDPNFALISK
18E33.txt	1	RING2_HUMAN	E3 ubiquitin-protein ligase RING2	37632	429	SLRPDPNFALISK
18E33.txt	1	RING2_HUMAN	E3 ubiquitin-protein ligase RING2	37632	429	HNNQALSHSIEEGLK
18E33.txt	1	RING2_HUMAN	E3 ubiquitin-protein ligase RING2	37632	429	TPQEAITDGLSIVVSPR
18E33.txt	1	RING2_HUMAN	E3 ubiquitin-protein ligase RING2	37632	429	SLHSELMCPICLDMLKNTMTTK
18E34.txt	1	CSK22_HUMAN	Casein kinase II subunit alpha	41187	94	TPALVFEYINNTDFK
18E35.txt	4	CBX2_HUMAN	Chromobox protein homolog 2	56046	79	AQILVAKPELKDPIR
18E35.txt	4	CBX2_HUMAN	Chromobox protein homolog 2	56046	79	HNSWEPEENILDP
18E35.txt	4	CBX2_HUMAN	Chromobox protein homolog 2	56046	79	HNSWEPEENILDP
18E36.txt	3	CBX2_HUMAN	Chromobox protein homolog 2	56046	77	LLAFQK
18E36.txt	3	CBX2_HUMAN	Chromobox protein homolog 2	56046	77	HNSWEPEENILDP
18E36.txt	1	PSME3_HUMAN	Proteasome activator complex subunit 3	29488	195	SNQQLVDIIK
18E36.txt	1	PSME3_HUMAN	Proteasome activator complex subunit 3	29488	195	TVSEEAASYLDQISR
18E36.txt	1	PSME3_HUMAN	Proteasome activator complex subunit 3	29488	195	TVTIDEKEYISLR
18E37.txt	1	PSME3_HUMAN	Proteasome activator complex subunit 3	29488	180	YPHVEDYR
18E37.txt	1	PSME3_HUMAN	Proteasome activator complex subunit 3	29488	180	NQYVTLHDMILK
18E37.txt	1	PSME3_HUMAN	Proteasome activator complex subunit 3	29488	180	TVSEEAASYLDQISR
18E37.txt	1	PSME3_HUMAN	Proteasome activator complex subunit 3	29488	180	TVSEEAASYLDQISR
18E37.txt	1	PSME3_HUMAN	Proteasome activator complex subunit 3	29488	180	ITSEADLVANFFPK

18E39.txt	5	CBX2_HUMAN	Chromobox protein homolog 2	56046	78	LEYLVK
18E39.txt	5	CBX2_HUMAN	Chromobox protein homolog 2	56046	78	HNSWEPEENILDP
18E39.txt	5	CBX2_HUMAN	Chromobox protein homolog 2	56046	78	HNSWEPEENILDP
18E40.txt	2	CSK2B_HUMAN	Casein kinase II subunit beta	24926	136	RPANQFVPR
18E40.txt	2	CSK2B_HUMAN	Casein kinase II subunit beta	24926	136	YQQGDFGYCPR
18E40.txt	2	CSK2B_HUMAN	Casein kinase II subunit beta	24926	136	FNLTLGNEQVPHYR
18E40.txt	8	H2A1B_HUMAN	Histone H2A type 1-B	14127	59	VTIAGGVLVLPNIQAVLLPK
18E41.txt	12	DI3L2_HUMAN	DIS3-like exonuclease 2	99147	43	SKNRSTRGK
18E41.txt	12	DI3L2_HUMAN	DIS3-like exonuclease 2	99147	43	DETTTCISQDTRALSEK
18E41.txt	7	H2A1B_HUMAN	Histone H2A type 1-B	14127	80	VTIAGGVLVLPNIQAVLLPK
18E43.txt	1	H2B1B_HUMAN	Histone H2A type 1-B	13942	109	LLLPGLAK
18E43.txt	1	H2B1B_HUMAN	Histone H2A type 1-B	13942	109	ESYSIYVYK
18E43.txt	1	H2B1B_HUMAN	Histone H2A type 1-B	13942	109	AMGIMNSFVNDIFER
18E43.txt	1	H2B1B_HUMAN	Histone H2A type 1-B	13942	109	AMGIMNSFVNDIFER
18E43.txt	2	H2B1C_HUMAN	Histone H2B type 1-C/E/F/G/I	13811	98	LLLPGLAK
18E43.txt	2	H2B1C_HUMAN	Histone H2B type 1-C/E/F/G/I	13811	98	ESYSVYVYK
18E43.txt	2	H2B1C_HUMAN	Histone H2B type 1-C/E/F/G/I	13811	98	AMGIMNSFVNDIFER
18E43.txt	2	H2B1C_HUMAN	Histone H2B type 1-C/E/F/G/I	13811	98	AMGIMNSFVNDIFER
18E46.txt	1	H4_HUMAN	Histone H4	11360	57	TVTAMDVVYALK

Tableau S2 : Peptides identifiés dans les 2 purifications CBX4-TAP (n°1 et 8)

band	hit	accession	description	mass	score	peptide
merge01E02.txt	13	SMC3_HUMAN	Structural maintenance of chromosomes protein 3	141454	58	EMQLSGGQK
merge01E02.txt	13	SMC3_HUMAN	Structural maintenance of chromosomes protein 3	141454	58	LDQDLNEVKAR
merge01E02.txt	13	SMC3_HUMAN	Structural maintenance of chromosomes protein 3	141454	58	LDELSAKRETSGEK
merge01E04.txt	10	SMC3_HUMAN	Structural maintenance of chromosomes protein 3	141454	61	LDQDLNEVKAR
merge01E04.txt	10	SMC3_HUMAN	Structural maintenance of chromosomes protein 3	141454	61	ALEYTIYNQELNETR
merge01E05.txt	17	FHAD1_HUMAN	Forkhead-associated domain-containing protein 1	161804	40	ELKALEEARL
merge01E05.txt	12	PHC2_HUMAN	Polyhomeotic-like protein 2	90657	86	GDGNSVSPGSMGR
merge01E05.txt	12	PHC2_HUMAN	Polyhomeotic-like protein 2	90657	86	AQEIDGQALLLLK
merge01E05.txt	10	PHC3_HUMAN	Polyhomeotic-like protein 3	106096	103	AQEIDGQALLLLK
merge01E05.txt	10	PHC3_HUMAN	Polyhomeotic-like protein 3	106096	103	SPSPDPSHVSVPPPLLLPAATTR
merge01E06.txt	16	MOV10_HUMAN	Putative helicase MOV-10	113599	44	ETGDPGGQLVLAGDPR
merge01E08.txt	17	PHC1_HUMAN	Polyhomeotic-like protein 1	105353	40	KAEADGSGQQNVGMNLTR
merge01E10.txt	8	PHC3_HUMAN	Polyhomeotic-like protein 3	106096	56	KRPLLDNQVINSVVCVPELQNNTK
merge01E10.txt	8	PHC3_HUMAN	Polyhomeotic-like protein 3	106096	56	GEDLTSEHPLELQVELPAVASVSASVIK
merge01E17.txt	9	CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	138	SNVLTGLQDSSTDNR
merge01E17.txt	9	CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	138	YNTWEPEENILDP
merge01E18.txt	5	CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	155	SNVLTGLQDSSTDNR
merge01E18.txt	5	CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	155	YNTWEPEENILDP
merge01E19.txt	16	CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	158	SNVLTGLQDSSTDNR
merge01E19.txt	16	CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	158	RSNVLTGLQDSSTDNR
merge01E19.txt	9	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	345	FCSDCIVTALR
merge01E19.txt	9	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	345	TWELSLYELHR
merge01E19.txt	9	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	345	VSRPLELCYAPTK
merge01E19.txt	9	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	345	SLRPPDPNFDALISK
merge01E19.txt	9	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	345	SLRPPDPNFDALISK
merge01E19.txt	9	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	345	LHNQQALSSSIEEGLR
merge01E19.txt	9	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	345	LHNQQALSSSIEEGLR
merge01E19.txt	9	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	345	IYPSREYEYAHQDR
merge01E19.txt	9	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	345	VSRPLELCYAPTKDPK
merge01E20.txt	7	CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	191	SNVLTGLQDSSTDNR
merge01E20.txt	7	CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	191	RSNVLTGLQDSSTDNR
merge01E20.txt	7	CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	191	YNTWEPEENILDP
merge01E20.txt	8	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	167	FCSDCIVTALR
merge01E20.txt	8	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	167	VSRPLELCYAPTK
merge01E20.txt	8	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	167	LHNQQALSSSIEEGLR
merge01E21.txt	6	CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	218	LLIAFQNR
merge01E21.txt	6	CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	218	SNVLTGLQDSSTDNR
merge01E21.txt	6	CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	218	RSNVLTGLQDSSTDNR
merge01E21.txt	6	CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	218	YNTWEPEENILDP
merge01E21.txt	6	CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	218	GPPNGMMPAPKEAVTGNIGGKMKIVK
merge01E21.txt	10	PCGF2_HUMAN	Polycomb group RING finger protein 2	37764	121	TLQDIVYK
merge01E21.txt	10	PCGF2_HUMAN	Polycomb group RING finger protein 2	37764	121	SDKTLQDIVYK
merge01E21.txt	10	PCGF2_HUMAN	Polycomb group RING finger protein 2	37764	121	YCPMCDVQVHK
merge01E21.txt	14	PCGF6_HUMAN	Polycomb group RING finger protein 6	39023	63	GYLDATTITECLHTFCK
merge01E21.txt	17	SMG1_HUMAN	Serine/threonine-protein kinase SMG1	410261	46	TVGLSLDVSR
merge01E21.txt	17	SMG1_HUMAN	Serine/threonine-protein kinase SMG1	410261	46	HGFDLLTEMK
merge01E21.txt	17	SMG1_HUMAN	Serine/threonine-protein kinase SMG1	410261	46	KAVEHNIQIGKFSQLVMNR
merge01E22.txt	6	BMI1_HUMAN	Polycomb complex protein BMI-1	36925	106	LVPGLFK
merge01E22.txt	6	BMI1_HUMAN	Polycomb complex protein BMI-1	36925	106	TLQDIVYK
merge01E22.txt	6	BMI1_HUMAN	Polycomb complex protein BMI-1	36925	106	SDKTLQDIVYK
merge01E22.txt	6	BMI1_HUMAN	Polycomb complex protein BMI-1	36925	106	DGLTNAGELESDSGSDK
merge01E22.txt	2	CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	234	EYVTV
merge01E22.txt	2	CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	234	LLIAFQNR
merge01E22.txt	2	CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	234	QEQLMGYR
merge01E22.txt	2	CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	234	QEQLMGYR
merge01E22.txt	2	CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	234	SNVLTGLQDSSTDNR
merge01E22.txt	2	CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	234	GPKPKPLVQVPTFAR
merge01E22.txt	2	CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	234	RSNVLTGLQDSSTDNR
merge01E22.txt	2	CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	234	RSNVLTGLQDSSTDNR
merge01E22.txt	4	CSK21_HUMAN	Casein kinase II subunit alpha	45115	169	EAMEHPYFTVVK
merge01E22.txt	4	CSK21_HUMAN	Casein kinase II subunit alpha	45115	169	GGPNIITLADIVKDPVSR

merge01E22.txt	4	CSK21_HUMAN	Casein kinase II subunit alpha	45115	169	VLGTEDLDYIDKYNIELDPR
merge01E22.txt	8	PHC2_HUMAN	Polyhomeotic-like protein 2	90657	92	SSLVGNLK
merge01E22.txt	8	PHC2_HUMAN	Polyhomeotic-like protein 2	90657	92	LPQQDHTTTTDSMEEPYLQESK
merge01E22.txt	8	PHC2_HUMAN	Polyhomeotic-like protein 2	90657	92	LPQQDHTTTTDSMEEPYLQESKEEGAPLK
merge01E22.txt	5	RING2_HUMAN	E3 ubiquitin-protein ligase RING2	37632	124	FCADCIATLR
merge01E22.txt	5	RING2_HUMAN	E3 ubiquitin-protein ligase RING2	37632	124	SLRPPDPNFALISK
merge01E22.txt	13	YR010_HUMAN	Putative uncharacterized protein	46232	42	EAVSSK
merge01E22.txt	13	YR010_HUMAN	Putative uncharacterized protein	46232	42	KIAMISTKLLMEKEWVK
merge01E22.txt	13	YR010_HUMAN	Putative uncharacterized protein	46232	42	DNTASLNMKELTLKDVECK
merge01E23.txt	7	BMI1_HUMAN	Polycomb complex protein BMI-1	36925	337	LVPGLFK
merge01E23.txt	7	BMI1_HUMAN	Polycomb complex protein BMI-1	36925	337	TLQDIVYK
merge01E23.txt	7	BMI1_HUMAN	Polycomb complex protein BMI-1	36925	337	TLQDIVYK
merge01E23.txt	7	BMI1_HUMAN	Polycomb complex protein BMI-1	36925	337	CPAAMTVMHLR
merge01E23.txt	7	BMI1_HUMAN	Polycomb complex protein BMI-1	36925	337	SDKTLQDIVYK
merge01E23.txt	7	BMI1_HUMAN	Polycomb complex protein BMI-1	36925	337	YCPICDVQVHK
merge01E23.txt	7	BMI1_HUMAN	Polycomb complex protein BMI-1	36925	337	DGLTNAGELESDSGSDK
merge01E23.txt	7	BMI1_HUMAN	Polycomb complex protein BMI-1	36925	337	DFYAAHPSADAANGSNEDRGEVADEDKR
merge01E23.txt	9	CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	201	SNVLTGLQDSSTDNR
merge01E23.txt	9	CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	201	SNVLTGLQDSSTDNR
merge01E23.txt	9	CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	201	RSNVLTGLQDSSTDNR
merge01E23.txt	9	CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	201	YNTWEPEENILDPR
merge01E23.txt	11	PCGF2_HUMAN	Polycomb group RING finger protein 2	37764	143	LVPGLFK
merge01E23.txt	11	PCGF2_HUMAN	Polycomb group RING finger protein 2	37764	143	TLQDIVYK
merge01E23.txt	11	PCGF2_HUMAN	Polycomb group RING finger protein 2	37764	143	TLQDIVYK
merge01E23.txt	11	PCGF2_HUMAN	Polycomb group RING finger protein 2	37764	143	SDKTLQDIVYK
merge01E23.txt	11	PCGF2_HUMAN	Polycomb group RING finger protein 2	37764	143	YLETNKYCPMCDVQVHK
merge01E23.txt	2	PHC2_HUMAN	Polyhomeotic-like protein 2	90657	628	VDFAYK
merge01E23.txt	2	PHC2_HUMAN	Polyhomeotic-like protein 2	90657	628	SSLVGNLK
merge01E23.txt	2	PHC2_HUMAN	Polyhomeotic-like protein 2	90657	628	YAQGFLEPK
merge01E23.txt	2	PHC2_HUMAN	Polyhomeotic-like protein 2	90657	628	YAQGFLEPK
merge01E23.txt	2	PHC2_HUMAN	Polyhomeotic-like protein 2	90657	628	SSLVGNLKK
merge01E23.txt	2	PHC2_HUMAN	Polyhomeotic-like protein 2	90657	628	DLELPDMHMR
merge01E23.txt	2	PHC2_HUMAN	Polyhomeotic-like protein 2	90657	628	AQEIDGQALLLLK
merge01E23.txt	2	PHC2_HUMAN	Polyhomeotic-like protein 2	90657	628	WNVEDVYEFIR
merge01E23.txt	2	PHC2_HUMAN	Polyhomeotic-like protein 2	90657	628	SLPGCQEIIEEFR
merge01E23.txt	2	PHC2_HUMAN	Polyhomeotic-like protein 2	90657	628	DLVGMGHFFLPSEPTK
merge01E23.txt	2	PHC2_HUMAN	Polyhomeotic-like protein 2	90657	628	LPQQDHTTTTDSMEEPYLQESK
merge01E23.txt	2	PHC2_HUMAN	Polyhomeotic-like protein 2	90657	628	KQPTGTVPLSVTAALQLTHSQEDSSR
merge01E23.txt	2	PHC2_HUMAN	Polyhomeotic-like protein 2	90657	628	LPQQDHTTTTDSMEEPYLQESKEEGAPLK
merge01E23.txt	2	PHC2_HUMAN	Polyhomeotic-like protein 2	90657	628	LPQQDHTTTTDSMEEPYLQESKEEGAPLK
merge01E23.txt	3	RING2_HUMAN	E3 ubiquitin-protein ligase RING2	37632	598	LALEELR
merge01E23.txt	3	RING2_HUMAN	E3 ubiquitin-protein ligase RING2	37632	598	FCADCIATLR
merge01E23.txt	3	RING2_HUMAN	E3 ubiquitin-protein ligase RING2	37632	598	FCADCIATLR
merge01E23.txt	3	RING2_HUMAN	E3 ubiquitin-protein ligase RING2	37632	598	TWELSLYELQR
merge01E23.txt	3	RING2_HUMAN	E3 ubiquitin-protein ligase RING2	37632	598	GESNQMNLDTASEK
merge01E23.txt	3	RING2_HUMAN	E3 ubiquitin-protein ligase RING2	37632	598	VNKPMELYYAPT
merge01E23.txt	3	RING2_HUMAN	E3 ubiquitin-protein ligase RING2	37632	598	SLRPPDPNFALISK
merge01E23.txt	3	RING2_HUMAN	E3 ubiquitin-protein ligase RING2	37632	598	SLRPPDPNFALISK
merge01E23.txt	3	RING2_HUMAN	E3 ubiquitin-protein ligase RING2	37632	598	IYPSRDEYEAHQER
merge01E23.txt	3	RING2_HUMAN	E3 ubiquitin-protein ligase RING2	37632	598	HNNQQALSHSIEEGLK
merge01E23.txt	3	RING2_HUMAN	E3 ubiquitin-protein ligase RING2	37632	598	TPQEAITDGLIEIVVSPR
merge01E23.txt	3	RING2_HUMAN	E3 ubiquitin-protein ligase RING2	37632	598	SLHSELMCPICLDMLK
merge01E23.txt	3	RING2_HUMAN	E3 ubiquitin-protein ligase RING2	37632	598	SLHSELMCPICLDMLK
merge01E23.txt	3	RING2_HUMAN	E3 ubiquitin-protein ligase RING2	37632	598	INKHNNQQALSHSIEEGLK
merge01E23.txt	3	RING2_HUMAN	E3 ubiquitin-protein ligase RING2	37632	598	INKHNNQQALSHSIEEGLK
merge01E24.txt	4	BMI1_HUMAN	Polycomb complex protein BMI-1	36925	244	LVPGLFK
merge01E24.txt	4	BMI1_HUMAN	Polycomb complex protein BMI-1	36925	244	TLQDIVYK
merge01E24.txt	4	BMI1_HUMAN	Polycomb complex protein BMI-1	36925	244	SDKTLQDIVYK
merge01E24.txt	4	BMI1_HUMAN	Polycomb complex protein BMI-1	36925	244	YCPICDVQVHK
merge01E24.txt	4	BMI1_HUMAN	Polycomb complex protein BMI-1	36925	244	DGLTNAGELESDSGSDK
merge01E24.txt	3	CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	266	LLIAFQNR
merge01E24.txt	3	CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	266	QEQLMGYR
merge01E24.txt	3	CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	266	ERQDELMGYR
merge01E24.txt	3	CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	266	SNVLTGLQDSSTDNR
merge01E24.txt	3	CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	266	GPKPKPLVVQVPTFAR
merge01E24.txt	3	CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	266	YNTWEPEENILDPR
merge01E24.txt	19	CSK22_HUMAN	Casein kinase II subunit alpha'	41187	46	LIDTVKDPVSK
merge01E24.txt	6	PCGF2_HUMAN	Polycomb group RING finger protein 2	37764	111	LVPGLFK
merge01E24.txt	6	PCGF2_HUMAN	Polycomb group RING finger protein 2	37764	111	TLQDIVYK
merge01E24.txt	6	PCGF2_HUMAN	Polycomb group RING finger protein 2	37764	111	SDKTLQDIVYK
merge01E24.txt	6	PCGF2_HUMAN	Polycomb group RING finger protein 2	37764	111	YLETNKYCPMCDVQVHK
merge01E24.txt	5	RING2_HUMAN	E3 ubiquitin-protein ligase RING2	37632	177	SLRPPDPNFALISK
merge01E24.txt	5	RING2_HUMAN	E3 ubiquitin-protein ligase RING2	37632	177	TPQEAITDGLIEIVVSPR
merge01E25.txt	11	BMI1_HUMAN	Polycomb complex protein BMI-1	36925	70	SDKTLQDIVYK
merge01E25.txt	5	CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	210	SNVLTGLQDSSTDNR
merge01E25.txt	5	CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	210	RCLSETHGEREPECK
merge01E25.txt	5	CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	210	RSNVLTGLQDSSTDNR
merge01E25.txt	5	CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	210	YNTWEPEENILDPR
merge01E27.txt	4	1433E_HUMAN	14-3-3 protein epsilon	29155	460	NLLSVAYK
merge01E27.txt	4	1433E_HUMAN	14-3-3 protein epsilon	29155	460	AAAENSLVAYK
merge01E27.txt	4	1433E_HUMAN	14-3-3 protein epsilon	29155	460	AAAENSLVAYK
merge01E27.txt	4	1433E_HUMAN	14-3-3 protein epsilon	29155	460	YLAEFATGNDR
merge01E27.txt	4	1433E_HUMAN	14-3-3 protein epsilon	29155	460	YLAEFATGNDRK
merge01E27.txt	4	1433E_HUMAN	14-3-3 protein epsilon	29155	460	VAGMDVELTVEER
merge01E27.txt	4	1433E_HUMAN	14-3-3 protein epsilon	29155	460	LICCDILDVLDK
merge01E27.txt	4	1433E_HUMAN	14-3-3 protein epsilon	29155	460	GDYHRYLAEFATGNDR
merge01E27.txt	4	1433E_HUMAN	14-3-3 protein epsilon	29155	460	IISSEIQKEENKGGEDKLK
merge01E27.txt	4	1433E_HUMAN	14-3-3 protein epsilon	29155	460	IISSEIQKEENKGGEDKLK

merge01E27.txt	4	1433E_HUMAN	14-3-3 protein epsilon	29155	460	LICCDILDVLDKHLIPAANTGESK
merge01E27.txt	8	ADT2_HUMAN	ADP/ATP translocase 2	32874	143	GLGDCLVK
merge01E27.txt	8	ADT2_HUMAN	ADP/ATP translocase 2	32874	143	EQGVLSFWR
merge01E27.txt	8	ADT2_HUMAN	ADP/ATP translocase 2	32874	143	YFPTQALNFAFK
merge01E27.txt	7	CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	196	LLIAFQNR
merge01E27.txt	7	CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	196	SNVLTGLQDSSTDNR
merge01E27.txt	7	CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	196	RSNVLTGLQDSSTDNR
merge01E27.txt	7	CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	196	YNTWPEENILDPDR
merge01E27.txt	17	H11_HUMAN	Histone H1.1	21829	41	TVKPKKVAKSPAK
merge01E27.txt	17	H11_HUMAN	Histone H1.1	21829	41	SPAKAKAVKPKAAK
merge01E27.txt	17	H11_HUMAN	Histone H1.1	21829	41	ASSVETKPGASKVATK
merge01E27.txt	17	H11_HUMAN	Histone H1.1	21829	41	SETVPPAASAASAEKPLAGK
merge01E27.txt	19	STALP_HUMAN	AMSH-like protease	49783	38	NKYKAEILKK
merge01E27.txt	19	STALP_HUMAN	AMSH-like protease	49783	38	LTNAGMLEVSACKK
merge01E28.txt	11	1433B_HUMAN	14-3-3 protein beta/alpha	28065	154	NLLSVAYK
merge01E28.txt	11	1433B_HUMAN	14-3-3 protein beta/alpha	28065	154	NLLSVAYK
merge01E28.txt	11	1433B_HUMAN	14-3-3 protein beta/alpha	28065	154	DSTLIMQLLR
merge01E28.txt	11	1433B_HUMAN	14-3-3 protein beta/alpha	28065	154	DSTLIMQLLR
merge01E28.txt	11	1433B_HUMAN	14-3-3 protein beta/alpha	28065	154	YLIPNATQPESK
merge01E28.txt	2	1433E_HUMAN	14-3-3 protein epsilon	29155	566	NLLSVAYK
merge01E28.txt	2	1433E_HUMAN	14-3-3 protein epsilon	29155	566	NLLSVAYK
merge01E28.txt	2	1433E_HUMAN	14-3-3 protein epsilon	29155	566	DSTLIMQLLR
merge01E28.txt	2	1433E_HUMAN	14-3-3 protein epsilon	29155	566	EAAENSLVAYK
merge01E28.txt	2	1433E_HUMAN	14-3-3 protein epsilon	29155	566	EAAENSLVAYK
merge01E28.txt	2	1433E_HUMAN	14-3-3 protein epsilon	29155	566	DSTLIMQLLR
merge01E28.txt	2	1433E_HUMAN	14-3-3 protein epsilon	29155	566	YLAEFATGNDR
merge01E28.txt	2	1433E_HUMAN	14-3-3 protein epsilon	29155	566	YLAEFATGNDRK
merge01E28.txt	2	1433E_HUMAN	14-3-3 protein epsilon	29155	566	VAGMDVLTVEER
merge01E28.txt	2	1433E_HUMAN	14-3-3 protein epsilon	29155	566	LICCDILDVLDK
merge01E28.txt	2	1433E_HUMAN	14-3-3 protein epsilon	29155	566	AASDIAMTELPPHPHPIR
merge01E28.txt	2	1433E_HUMAN	14-3-3 protein epsilon	29155	566	GDYHRYLAEFATGNDR
merge01E28.txt	2	1433E_HUMAN	14-3-3 protein epsilon	29155	566	LAEQAERYDEMVMESMKK
merge01E28.txt	2	1433E_HUMAN	14-3-3 protein epsilon	29155	566	LICCDILDVLDKHLIPAANTGESK
merge01E28.txt	7	1433G_HUMAN	14-3-3 protein gamma	28285	188	NLLSVAYK
merge01E28.txt	7	1433G_HUMAN	14-3-3 protein gamma	28285	188	NLLSVAYK
merge01E28.txt	7	1433G_HUMAN	14-3-3 protein gamma	28285	188	YLAEVATGEEK
merge01E28.txt	7	1433G_HUMAN	14-3-3 protein gamma	28285	188	DSTLIMQLLR
merge01E28.txt	7	1433G_HUMAN	14-3-3 protein gamma	28285	188	DSTLIMQLLR
merge01E28.txt	7	1433G_HUMAN	14-3-3 protein gamma	28285	188	NVTELNEPLSNEER
merge01E28.txt	9	1433T_HUMAN	14-3-3 protein theta	27747	161	NLLSVAYK
merge01E28.txt	9	1433T_HUMAN	14-3-3 protein theta	27747	161	NLLSVAYK
merge01E28.txt	9	1433T_HUMAN	14-3-3 protein theta	27747	161	DSTLIMQLLR
merge01E28.txt	9	1433T_HUMAN	14-3-3 protein theta	27747	161	DSTLIMQLLR
merge01E28.txt	9	1433T_HUMAN	14-3-3 protein theta	27747	161	AVTEQGAELSNEER
merge01E28.txt	6	1433Z_HUMAN	14-3-3 protein zeta/delta	27728	254	NLLSVAYK
merge01E28.txt	6	1433Z_HUMAN	14-3-3 protein zeta/delta	27728	254	NLLSVAYK
merge01E28.txt	6	1433Z_HUMAN	14-3-3 protein zeta/delta	27728	254	DSTLIMQLLR
merge01E28.txt	6	1433Z_HUMAN	14-3-3 protein zeta/delta	27728	254	DSTLIMQLLR
merge01E28.txt	6	1433Z_HUMAN	14-3-3 protein zeta/delta	27728	254	DICNDVLSLLEK
merge01E28.txt	6	1433Z_HUMAN	14-3-3 protein zeta/delta	27728	254	SVTEQGAELSNEER
merge01E28.txt	6	1433Z_HUMAN	14-3-3 protein zeta/delta	27728	254	SVTEQGAELSNEER
merge01E28.txt	6	1433Z_HUMAN	14-3-3 protein zeta/delta	27728	254	EKIETELRDICNDVLSLLEK
merge01E28.txt	10	CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	156	SNVLTGLQDSSTDNR
merge01E28.txt	10	CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	156	RSNVLTGLQDSSTDNR
merge01E28.txt	10	CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	156	YNTWPEENILDPDR
merge01E29.txt	6	1433B_HUMAN	14-3-3 protein beta/alpha	28065	91	NLLSVAYK
merge01E29.txt	6	1433B_HUMAN	14-3-3 protein beta/alpha	28065	91	MTMDKSELVQK
merge01E29.txt	6	1433B_HUMAN	14-3-3 protein beta/alpha	28065	91	YLIPNATQPESK
merge01E29.txt	8	1433F_HUMAN	14-3-3 protein eta	28201	79	NLLSVAYK
merge01E29.txt	8	1433F_HUMAN	14-3-3 protein eta	28201	79	NSVVEASEAAYK
merge01E29.txt	7	1433G_HUMAN	14-3-3 protein gamma	28285	88	NLLSVAYK
merge01E29.txt	7	1433G_HUMAN	14-3-3 protein gamma	28285	88	NVTELNEPLSNEER
merge01E29.txt	3	1433T_HUMAN	14-3-3 protein theta	27747	182	NLLSVAYK
merge01E29.txt	3	1433T_HUMAN	14-3-3 protein theta	27747	182	YLAEVACGDDR
merge01E29.txt	3	1433T_HUMAN	14-3-3 protein theta	27747	182	YLIANATNPESK
merge01E29.txt	3	1433T_HUMAN	14-3-3 protein theta	27747	182	AVTEQGAELSNEER
merge01E29.txt	10	1433Z_HUMAN	14-3-3 protein zeta/delta	27728	75	NLLSVAYK
merge01E29.txt	10	1433Z_HUMAN	14-3-3 protein zeta/delta	27728	75	SVTEQGAELSNEER
merge01E29.txt	5	CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	98	RSNVLTGLQDSSTDNR
merge01E29.txt	5	CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	98	RSNVLTGLQDSSTDNR
merge01E29.txt	5	CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	98	YNTWPEENILDPDR
merge01E30.txt	10	CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	255	LLIAFQNR
merge01E30.txt	10	CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	255	ERQEQLMGYR
merge01E30.txt	10	CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	255	SNVLTGLQDSSTDNR
merge01E30.txt	10	CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	255	RSNVLTGLQDSSTDNR
merge01E30.txt	10	CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	255	YNTWPEENILDPDR
merge01E30.txt	10	CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	255	GSEKGGPPNGMMPAPKEAVTNGNGIGGKMK
merge01E30.txt	21	CSK2B_HUMAN	Casein kinase II subunit beta	24926	59	YQQGDFGYGCFR
merge01E31.txt	3	CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	175	LLIAFQNR
merge01E31.txt	3	CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	175	SNVLTGLQDSSTDNR
merge01E31.txt	3	CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	175	SNVLTGLQDSSTDNR
merge01E31.txt	3	CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	175	YNTWPEENILDPDR
merge01E32.txt	7	CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	200	SNVLTGLQDSSTDNR
merge01E32.txt	7	CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	200	RSNVLTGLQDSSTDNR
merge01E32.txt	7	CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	200	YNTWPEENILDPDR
merge01E33.txt	11	CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	96	ERQEQLMGYR
merge01E33.txt	11	CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	96	SPSHKKRAADER
merge01E33.txt	11	CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	96	YNTWPEENILDPDR

merge8E25.txt	1 CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	540 GPKPKPLVVQVPTFAR
merge8E25.txt	1 CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	540 RSNVLTGLQDSSTDNR
merge8E25.txt	1 CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	540 YNTWEPEENILDP
merge8E25.txt	6 PCGF2_HUMAN	Polycomb group RING finger protein 2	37764	104 TLQDIVYK
merge8E25.txt	6 PCGF2_HUMAN	Polycomb group RING finger protein 2	37764	104 SDKTLQDIVYK
merge8E25.txt	6 PCGF2_HUMAN	Polycomb group RING finger protein 2	37764	104 YLETNKYCPMCDVQVHK
merge8E25.txt	2 RING2_HUMAN	E3 ubiquitin-protein ligase RING2	37632	284 YLAVR
merge8E25.txt	2 RING2_HUMAN	E3 ubiquitin-protein ligase RING2	37632	284 LALEELR
merge8E25.txt	2 RING2_HUMAN	E3 ubiquitin-protein ligase RING2	37632	284 TSGNATVDHLSK
merge8E25.txt	2 RING2_HUMAN	E3 ubiquitin-protein ligase RING2	37632	284 FCADCITALR
merge8E25.txt	2 RING2_HUMAN	E3 ubiquitin-protein ligase RING2	37632	284 VNKPMELYYAPTK
merge8E25.txt	2 RING2_HUMAN	E3 ubiquitin-protein ligase RING2	37632	284 VNKPMELYYAPTK
merge8E25.txt	2 RING2_HUMAN	E3 ubiquitin-protein ligase RING2	37632	284 IYPSRDEYEAHQER
merge8E26.txt	1 CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	492 EYTV
merge8E26.txt	1 CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	492 VEYLVK
merge8E26.txt	1 CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	492 VEYLVK
merge8E26.txt	1 CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	492 LDLGAQGK
merge8E26.txt	1 CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	492 LLIAFQNR
merge8E26.txt	1 CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	492 AKLDLGAQGK
merge8E26.txt	1 CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	492 QEQLMGYR
merge8E26.txt	1 CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	492 QEQLMGYR
merge8E26.txt	1 CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	492 ERQEQLMGYR
merge8E26.txt	1 CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	492 ERQEQLMGYR
merge8E26.txt	1 CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	492 SNVLTGLQDSSTDNR
merge8E26.txt	1 CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	492 GPKPKPLVVQVPTFAR
merge8E26.txt	1 CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	492 YNTWEPEENILDP
merge8E26.txt	1 CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	272 VEYLVK
merge8E26.txt	1 CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	272 AKLDLGAQGK
merge8E26.txt	1 CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	272 SNVLTGLQDSSTDNR
merge8E26.txt	1 CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	272 RSNVLTGLQDSSTDNR
merge8E26.txt	1 CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	348 VEYLVK
merge8E26.txt	1 CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	348 LLIAFQNR
merge8E26.txt	1 CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	348 AKLDLGAQGK
merge8E26.txt	1 CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	348 QEQLMGYR
merge8E26.txt	1 CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	348 ERQEQLMGYR
merge8E26.txt	1 CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	348 SNVLTGLQDSSTDNR
merge8E26.txt	1 CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	348 SNVLTGLQDSSTDNR
merge8E26.txt	1 CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	348 RSNVLTGLQDSSTDNR
merge8E26.txt	1 CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	348 YNTWEPEENILDP
merge8E26.txt	1 1433E_HUMAN	14-3-3 protein epsilon	29155	1151 VFYYK
merge8E26.txt	1 1433E_HUMAN	14-3-3 protein epsilon	29155	1151 VFYYK
merge8E26.txt	1 1433E_HUMAN	14-3-3 protein epsilon	29155	1151 NLLSVAYK
merge8E26.txt	1 1433E_HUMAN	14-3-3 protein epsilon	29155	1151 NLLSVAYK
merge8E26.txt	1 1433E_HUMAN	14-3-3 protein epsilon	29155	1151 IISIEQK
merge8E26.txt	1 1433E_HUMAN	14-3-3 protein epsilon	29155	1151 IISIEQK
merge8E26.txt	1 1433E_HUMAN	14-3-3 protein epsilon	29155	1151 QMVETELK
merge8E26.txt	1 1433E_HUMAN	14-3-3 protein epsilon	29155	1151 DSTLIMQLLR
merge8E26.txt	1 1433E_HUMAN	14-3-3 protein epsilon	29155	1151 EAAENSLVAYK
merge8E26.txt	1 1433E_HUMAN	14-3-3 protein epsilon	29155	1151 DSTLIMQLLR
merge8E26.txt	1 1433E_HUMAN	14-3-3 protein epsilon	29155	1151 HLIPAANTGESK
merge8E26.txt	1 1433E_HUMAN	14-3-3 protein epsilon	29155	1151 YLAEFATGNDR
merge8E26.txt	1 1433E_HUMAN	14-3-3 protein epsilon	29155	1151 YLAEFATGNDR
merge8E26.txt	1 1433E_HUMAN	14-3-3 protein epsilon	29155	1151 YDEMVESMK
merge8E26.txt	1 1433E_HUMAN	14-3-3 protein epsilon	29155	1151 YLAEFATGNDR
merge8E26.txt	1 1433E_HUMAN	14-3-3 protein epsilon	29155	1151 YLAEFATGNDR
merge8E26.txt	1 1433E_HUMAN	14-3-3 protein epsilon	29155	1151 IISIEQKEENK
merge8E26.txt	1 1433E_HUMAN	14-3-3 protein epsilon	29155	1151 VAGMDVELTVEER
merge8E26.txt	1 1433E_HUMAN	14-3-3 protein epsilon	29155	1151 VAGMDVELTVEER
merge8E26.txt	1 1433E_HUMAN	14-3-3 protein epsilon	29155	1151 LICCDILDVLDK
merge8E26.txt	1 1433E_HUMAN	14-3-3 protein epsilon	29155	1151 AASDIAMTELPPTHPIR
merge8E26.txt	1 1433E_HUMAN	14-3-3 protein epsilon	29155	1151 AASDIAMTELPPTHPIR
merge8E26.txt	1 1433E_HUMAN	14-3-3 protein epsilon	29155	1151 GDYHRYLAEFATGNDR
merge8E26.txt	1 1433E_HUMAN	14-3-3 protein epsilon	29155	1151 IISIEQKEENKGGEDK
merge8E26.txt	1 1433E_HUMAN	14-3-3 protein epsilon	29155	1151 IISIEQKEENKGGEDK
merge8E26.txt	1 1433E_HUMAN	14-3-3 protein epsilon	29155	1151 IISIEQKEENKGGEDK
merge8E26.txt	1 1433E_HUMAN	14-3-3 protein epsilon	29155	1151 IISIEQKEENKGGEDK
merge8E26.txt	1 1433E_HUMAN	14-3-3 protein epsilon	29155	1151 LICCDILDVLDKHLIPAANTGESK
merge8E26.txt	1 1433E_HUMAN	14-3-3 protein epsilon	29155	1151 AAFDDAIAELDTLSEESYKSTLIMQLLR
merge8E26.txt	1 1433E_HUMAN	14-3-3 protein epsilon	29155	1151 DNLTWLWTSMDQGDGEEQNKEALQDVEDEN Q
merge8E29.txt	5 1433F_HUMAN	14-3-3 protein eta	28201	95 NLLSVAYK
merge8E29.txt	5 1433F_HUMAN	14-3-3 protein eta	28201	95 NLLSVAYK
merge8E29.txt	5 1433F_HUMAN	14-3-3 protein eta	28201	95 DSTLIMQLLR
merge8E29.txt	5 1433F_HUMAN	14-3-3 protein eta	28201	95 DSTLIMQLLR
merge8E29.txt	5 1433F_HUMAN	14-3-3 protein eta	28201	95 VISSIEQKTMADGNEKK
merge8E29.txt	2 CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	234 SNVLTGLQDSSTDNR
merge8E29.txt	2 CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	234 RSNVLTGLQDSSTDNR
merge8E29.txt	2 CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	234 YNTWEPEENILDP
merge8E29.txt	8 RNF17_HUMAN	RING finger protein 17	184643	43 TADQLTTGLER
merge8E29.txt	8 RNF17_HUMAN	RING finger protein 17	184643	43 IKDEFNLNAPEKAICK
merge8E30.txt	1 1433E_HUMAN	14-3-3 protein epsilon	29155	979 VFYYK
merge8E30.txt	1 1433E_HUMAN	14-3-3 protein epsilon	29155	979 VFYYK
merge8E30.txt	1 1433E_HUMAN	14-3-3 protein epsilon	29155	979 NLLSVAYK
merge8E30.txt	1 1433E_HUMAN	14-3-3 protein epsilon	29155	979 NLLSVAYK
merge8E30.txt	1 1433E_HUMAN	14-3-3 protein epsilon	29155	979 IISIEQK
merge8E30.txt	1 1433E_HUMAN	14-3-3 protein epsilon	29155	979 QMVETELK
merge8E30.txt	1 1433E_HUMAN	14-3-3 protein epsilon	29155	979 YDEMVESMK
merge8E30.txt	1 1433E_HUMAN	14-3-3 protein epsilon	29155	979 HLIPAANTGESK

merge8E31.txt	7	1433S_HUMAN	14-3-3 protein sigma	27757	151	VFYLK
merge8E31.txt	7	1433S_HUMAN	14-3-3 protein sigma	27757	151	VLLSIEQK
merge8E31.txt	7	1433S_HUMAN	14-3-3 protein sigma	27757	151	NLLSVAYK
merge8E31.txt	7	1433S_HUMAN	14-3-3 protein sigma	27757	151	NLLSVAYK
merge8E31.txt	7	1433S_HUMAN	14-3-3 protein sigma	27757	151	DSTLIMQLLR
merge8E31.txt	7	1433S_HUMAN	14-3-3 protein sigma	27757	151	DSTLIMQLLR
merge8E31.txt	7	1433S_HUMAN	14-3-3 protein sigma	27757	151	DSTLIMQLLR
merge8E31.txt	2	1433T_HUMAN	14-3-3 protein theta	27747	600	GDYFR
merge8E31.txt	2	1433T_HUMAN	14-3-3 protein theta	27747	600	VFYLK
merge8E31.txt	2	1433T_HUMAN	14-3-3 protein theta	27747	600	VISSIEQK
merge8E31.txt	2	1433T_HUMAN	14-3-3 protein theta	27747	600	NLLSVAYK
merge8E31.txt	2	1433T_HUMAN	14-3-3 protein theta	27747	600	NLLSVAYK
merge8E31.txt	2	1433T_HUMAN	14-3-3 protein theta	27747	600	EMQPTHPIR
merge8E31.txt	2	1433T_HUMAN	14-3-3 protein theta	27747	600	DSTLIMQLLR
merge8E31.txt	2	1433T_HUMAN	14-3-3 protein theta	27747	600	DSTLIMQLLR
merge8E31.txt	2	1433T_HUMAN	14-3-3 protein theta	27747	600	DSTLIMQLLR
merge8E31.txt	2	1433T_HUMAN	14-3-3 protein theta	27747	600	YLAEVACGDDR
merge8E31.txt	2	1433T_HUMAN	14-3-3 protein theta	27747	600	YLIANATNPESK
merge8E31.txt	2	1433T_HUMAN	14-3-3 protein theta	27747	600	SICTTVLELLDK
merge8E31.txt	2	1433T_HUMAN	14-3-3 protein theta	27747	600	YLAEVACGDDR
merge8E31.txt	2	1433T_HUMAN	14-3-3 protein theta	27747	600	AVTEQGAELSNEER
merge8E31.txt	2	1433T_HUMAN	14-3-3 protein theta	27747	600	KQIDNSQGAQYQEAFFDISK
merge8E31.txt	3	1433Z_HUMAN	14-3-3 protein zeta/delta	27728	572	VFYLK
merge8E31.txt	3	1433Z_HUMAN	14-3-3 protein zeta/delta	27728	572	NLLSVAYK
merge8E31.txt	3	1433Z_HUMAN	14-3-3 protein zeta/delta	27728	572	NLLSVAYK
merge8E31.txt	3	1433Z_HUMAN	14-3-3 protein zeta/delta	27728	572	EKIETELR
merge8E31.txt	3	1433Z_HUMAN	14-3-3 protein zeta/delta	27728	572	EMQPTHPIR
merge8E31.txt	3	1433Z_HUMAN	14-3-3 protein zeta/delta	27728	572	YLAEVACGDDR
merge8E31.txt	3	1433Z_HUMAN	14-3-3 protein zeta/delta	27728	572	DSTLIMQLLR
merge8E31.txt	3	1433Z_HUMAN	14-3-3 protein zeta/delta	27728	572	DSTLIMQLLR
merge8E31.txt	3	1433Z_HUMAN	14-3-3 protein zeta/delta	27728	572	DSTLIMQLLR
merge8E31.txt	3	1433Z_HUMAN	14-3-3 protein zeta/delta	27728	572	DSTLIMQLLR
merge8E31.txt	3	1433Z_HUMAN	14-3-3 protein zeta/delta	27728	572	YLAEVACGDDR
merge8E31.txt	3	1433Z_HUMAN	14-3-3 protein zeta/delta	27728	572	FLIPNASQAESK
merge8E31.txt	3	1433Z_HUMAN	14-3-3 protein zeta/delta	27728	572	FLIPNASQAESK
merge8E31.txt	3	1433Z_HUMAN	14-3-3 protein zeta/delta	27728	572	DICNDVLSLLEK
merge8E31.txt	3	1433Z_HUMAN	14-3-3 protein zeta/delta	27728	572	SVTEQGAELSNEER
merge8E31.txt	3	1433Z_HUMAN	14-3-3 protein zeta/delta	27728	572	SVTEQGAELSNEER
merge8E31.txt	3	1433Z_HUMAN	14-3-3 protein zeta/delta	27728	572	IETELRDICNDVLSLLEK
merge8E31.txt	3	1433Z_HUMAN	14-3-3 protein zeta/delta	27728	572	GIVDQSQAYQEAFEISKK
merge8E31.txt	8	CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	104	SNVLTGLDQSDTDNR
merge8E31.txt	16	CENPF_HUMAN	Centromere protein F	367537	46	SCKQLEEEK
merge8E31.txt	16	CENPF_HUMAN	Centromere protein F	367537	46	MESQGMKKNK
merge8E31.txt	16	CENPF_HUMAN	Centromere protein F	367537	46	EQNLSQVECLELEK
merge8E31.txt	16	CENPF_HUMAN	Centromere protein F	367537	46	SLLAFDQQPAMHHSFANIIEQGSMPSEK
merge8E31.txt	23	MDM1_HUMAN	Nuclear protein MDM1	80686	41	SQEAQKDVQTER
merge8E31.txt	23	MDM1_HUMAN	Nuclear protein MDM1	80686	41	SKADKMKEGSDSSVSEK
merge8E31.txt	22	SSX10_HUMAN	Protein SSX10	17575	42	AFDDIAKYFSK
merge8E31.txt	22	SSX10_HUMAN	Protein SSX10	17575	42	MKASEKILYVYMK
merge8E32.txt	8	1433B_HUMAN	14-3-3 protein beta/alpha	28065	158	VFYLK
merge8E32.txt	8	1433B_HUMAN	14-3-3 protein beta/alpha	28065	158	NLLSVAYK
merge8E32.txt	8	1433B_HUMAN	14-3-3 protein beta/alpha	28065	158	NLLSVAYK
merge8E32.txt	8	1433B_HUMAN	14-3-3 protein beta/alpha	28065	158	AVTEQGHLSNEER
merge8E32.txt	1	1433E_HUMAN	14-3-3 protein epsilon	29155	666	VFYK
merge8E32.txt	1	1433E_HUMAN	14-3-3 protein epsilon	29155	666	NLLSVAYK
merge8E32.txt	1	1433E_HUMAN	14-3-3 protein epsilon	29155	666	NLLSVAYK
merge8E32.txt	1	1433E_HUMAN	14-3-3 protein epsilon	29155	666	QMVETELK
merge8E32.txt	1	1433E_HUMAN	14-3-3 protein epsilon	29155	666	EAAENSLVAYK
merge8E32.txt	1	1433E_HUMAN	14-3-3 protein epsilon	29155	666	HLIPAANTGESK
merge8E32.txt	1	1433E_HUMAN	14-3-3 protein epsilon	29155	666	YLAEFATGNDR
merge8E32.txt	1	1433E_HUMAN	14-3-3 protein epsilon	29155	666	IISIEQKEENK
merge8E32.txt	1	1433E_HUMAN	14-3-3 protein epsilon	29155	666	LICCDILDVLDK
merge8E32.txt	1	1433E_HUMAN	14-3-3 protein epsilon	29155	666	AASDIAMTELPPHPIR
merge8E32.txt	1	1433E_HUMAN	14-3-3 protein epsilon	29155	666	GDYHRYLAEFATGNDR
merge8E32.txt	1	1433E_HUMAN	14-3-3 protein epsilon	29155	666	IISIEQKEENKGGEDK
merge8E32.txt	1	1433E_HUMAN	14-3-3 protein epsilon	29155	666	IISIEQKEENKGGEDK
merge8E32.txt	1	1433E_HUMAN	14-3-3 protein epsilon	29155	666	QMVETELKICCDILDVLDK
merge8E32.txt	4	1433F_HUMAN	14-3-3 protein eta	28201	208	VFYLK
merge8E32.txt	4	1433F_HUMAN	14-3-3 protein eta	28201	208	NLLSVAYK
merge8E32.txt	4	1433F_HUMAN	14-3-3 protein eta	28201	208	NLLSVAYK
merge8E32.txt	4	1433F_HUMAN	14-3-3 protein eta	28201	208	YLAEVASGEK
merge8E32.txt	4	1433F_HUMAN	14-3-3 protein eta	28201	208	NSVVEASEAAYK
merge8E32.txt	4	1433F_HUMAN	14-3-3 protein eta	28201	208	AVTELNEPLSNEDR
merge8E32.txt	7	1433G_HUMAN	14-3-3 protein gamma	28285	158	VFYLK
merge8E32.txt	7	1433G_HUMAN	14-3-3 protein gamma	28285	158	NLLSVAYK
merge8E32.txt	7	1433G_HUMAN	14-3-3 protein gamma	28285	158	NLLSVAYK
merge8E32.txt	7	1433G_HUMAN	14-3-3 protein gamma	28285	158	NVTELNEPLSNEER
merge8E32.txt	6	1433T_HUMAN	14-3-3 protein theta	27747	186	VFYLK
merge8E32.txt	6	1433T_HUMAN	14-3-3 protein theta	27747	186	NLLSVAYK
merge8E32.txt	6	1433T_HUMAN	14-3-3 protein theta	27747	186	NLLSVAYK
merge8E32.txt	6	1433T_HUMAN	14-3-3 protein theta	27747	186	YLAEVACGDDR
merge8E32.txt	6	1433T_HUMAN	14-3-3 protein theta	27747	186	AVTEQGAELSNEER
merge8E32.txt	3	1433Z_HUMAN	14-3-3 protein zeta/delta	27728	320	VFYLK
merge8E32.txt	3	1433Z_HUMAN	14-3-3 protein zeta/delta	27728	320	NLLSVAYK
merge8E32.txt	3	1433Z_HUMAN	14-3-3 protein zeta/delta	27728	320	NLLSVAYK
merge8E32.txt	3	1433Z_HUMAN	14-3-3 protein zeta/delta	27728	320	YLAEVACGDDR
merge8E32.txt	3	1433Z_HUMAN	14-3-3 protein zeta/delta	27728	320	DICNDVLSLLEK
merge8E32.txt	3	1433Z_HUMAN	14-3-3 protein zeta/delta	27728	320	SVTEQGAELSNEER
merge8E32.txt	2	CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	376	VEYLK

merge8E35.txt	1	CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	436	ERQEQLMGYR
merge8E35.txt	1	CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	436	SNVLTGLQDSSTDNR
merge8E35.txt	1	CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	436	GPKPKPLVVQVPTFAR
merge8E35.txt	1	CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	436	GPKPKPLVVQVPTFAR
merge8E35.txt	1	CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	436	GPKPKPLVVQVPTFAR
merge8E35.txt	1	CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	436	RSNVLTLGLQDSSTDNR
merge8E35.txt	1	CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	436	GPPNGMMPAPKEAVTNGIGGK
merge8E35.txt	12	WDHD1_HUMAN	WD repeat and HMG-box DNA-binding protein 1	125888	41	SSKKSTALSR
merge8E35.txt	12	WDHD1_HUMAN	WD repeat and HMG-box DNA-binding protein 1	125888	41	SSKKSTALSR
merge8E35.txt	12	WDHD1_HUMAN	WD repeat and HMG-box DNA-binding protein 1	125888	41	GETASEGTEAK
merge8E35.txt	12	WDHD1_HUMAN	WD repeat and HMG-box DNA-binding protein 1	125888	41	LLAIPVEKSVKLYR
merge8E35.txt	13	ZN546_HUMAN	Zinc finger protein 546	98369	40	IYLSQLQTGEK
merge8E35.txt	13	ZN546_HUMAN	Zinc finger protein 546	98369	40	NVCKIYLSQLQTGEK
merge8E36.txt	1	CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	294	VEYLVK
merge8E36.txt	1	CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	294	LLIAFQNR
merge8E36.txt	1	CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	294	ERQEQLMGYR
merge8E36.txt	1	CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	294	SNVLTGLQDSSTDNR
merge8E36.txt	1	CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	294	GPKPKPLVVQVPTFAR
merge8E36.txt	1	CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	294	RSNVLTLGLQDSSTDNR
merge8E37.txt	1	1433E_HUMAN	14-3-3 protein epsilon	29155	346	NLLSVAYK
merge8E37.txt	1	1433E_HUMAN	14-3-3 protein epsilon	29155	346	EAAENSLVAYK
merge8E37.txt	1	1433E_HUMAN	14-3-3 protein epsilon	29155	346	YLAEFATGNDR
merge8E37.txt	1	1433E_HUMAN	14-3-3 protein epsilon	29155	346	VAGMDVELTVEER
merge8E37.txt	1	1433E_HUMAN	14-3-3 protein epsilon	29155	346	GDYHRYLAEFATGNDR
merge8E37.txt	1	1433E_HUMAN	14-3-3 protein epsilon	29155	346	IISSEIQKEENKGGEDK
merge8E37.txt	4	1433F_HUMAN	14-3-3 protein eta	28201	144	VFYLK
merge8E37.txt	4	1433F_HUMAN	14-3-3 protein eta	28201	144	NLLSVAYK
merge8E37.txt	4	1433F_HUMAN	14-3-3 protein eta	28201	144	AVTELNEPLSNEDR
merge8E37.txt	6	1433G_HUMAN	14-3-3 protein gamma	28285	128	VFYLK
merge8E37.txt	6	1433G_HUMAN	14-3-3 protein gamma	28285	128	NLLSVAYK
merge8E37.txt	6	1433G_HUMAN	14-3-3 protein gamma	28285	128	NVTELNEPLSNEER
merge8E37.txt	2	1433Z_HUMAN	14-3-3 protein zeta/delta	27728	232	VFYLK
merge8E37.txt	2	1433Z_HUMAN	14-3-3 protein zeta/delta	27728	232	NLLSVAYK
merge8E37.txt	2	1433Z_HUMAN	14-3-3 protein zeta/delta	27728	232	FLIPNASQAESK
merge8E37.txt	2	1433Z_HUMAN	14-3-3 protein zeta/delta	27728	232	SVTEQGAELSNEER
merge8E37.txt	2	1433Z_HUMAN	14-3-3 protein zeta/delta	27728	232	FLIPNASQAESKVFYLK
merge8E37.txt	3	CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	206	VEYLVK
merge8E37.txt	3	CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	206	LLIAFQNR
merge8E37.txt	3	CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	206	SNVLTGLQDSSTDNR
merge8E37.txt	3	CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	206	GPKPKPLVVQVPTFAR
merge8E37.txt	3	CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	206	YNTWEPEENILDPK

Tableau S3 : Peptides identifiés dans les 2 purifications CBX6-TAP (n°11 et 12)

band	hit	accession	description	mass	score	peptide
11E16.txt	3	IF2B3_HUMAN	Insulin-like growth factor 2 mRNA-binding protein 1	63681	39	FTEEIPLK
11E17.txt	6	CBX6_HUMAN	Chromobox protein homolog 2	43871	57	IILNLK
11E17.txt	6	CBX6_HUMAN	Chromobox protein homolog 2	43871	57	VFAAESIIK
11E18.txt	1	CBX6_HUMAN	Chromobox protein homolog 2	43871	375	IILNLK
11E18.txt	1	CBX6_HUMAN	Chromobox protein homolog 2	43871	375	IILNLK
11E18.txt	1	CBX6_HUMAN	Chromobox protein homolog 2	43871	375	IEYLVK
11E18.txt	1	CBX6_HUMAN	Chromobox protein homolog 2	43871	375	FSESVLR
11E18.txt	1	CBX6_HUMAN	Chromobox protein homolog 2	43871	375	LIAAFEQK
11E18.txt	1	CBX6_HUMAN	Chromobox protein homolog 2	43871	375	VFAAESIIK
11E18.txt	1	CBX6_HUMAN	Chromobox protein homolog 2	43871	375	EFCNPEDFEK
11E18.txt	1	CBX6_HUMAN	Chromobox protein homolog 2	43871	375	VAAGVAGAAGGGGSGISK
11E18.txt	1	CBX6_HUMAN	Chromobox protein homolog 2	43871	375	ISDVHFSVKPSASASSPK
11E18.txt	1	CBX6_HUMAN	Chromobox protein homolog 2	43871	375	ISDVHFSVKPSASASSPK
11E18.txt	6	ZN33A_HUMAN	Zinc finger protein 33A	94384	42	LLLNIK
11E18.txt	6	ZN33A_HUMAN	Zinc finger protein 33A	94384	42	LLLNIK
11E19.txt	1	CBX6_HUMAN	Chromobox protein homolog 2	43871	350	GWAIK
11E19.txt	1	CBX6_HUMAN	Chromobox protein homolog 2	43871	350	TFLLK
11E19.txt	1	CBX6_HUMAN	Chromobox protein homolog 2	43871	350	IILNLK
11E19.txt	1	CBX6_HUMAN	Chromobox protein homolog 2	43871	350	IEYLVK
11E19.txt	1	CBX6_HUMAN	Chromobox protein homolog 2	43871	350	FSESVLR
11E19.txt	1	CBX6_HUMAN	Chromobox protein homolog 2	43871	350	FSESVLR
11E19.txt	1	CBX6_HUMAN	Chromobox protein homolog 2	43871	350	WKGWAIK
11E19.txt	1	CBX6_HUMAN	Chromobox protein homolog 2	43871	350	LIAAFEQK
11E19.txt	1	CBX6_HUMAN	Chromobox protein homolog 2	43871	350	VFAAESIIK
11E19.txt	1	CBX6_HUMAN	Chromobox protein homolog 2	43871	350	LLPETVSPSAPSWR
11E19.txt	1	CBX6_HUMAN	Chromobox protein homolog 2	43871	350	YSTWEPEENILDSR
11E20.txt	1	CBX6_HUMAN	Chromobox protein homolog 2	43871	258	TFLLK
11E20.txt	1	CBX6_HUMAN	Chromobox protein homolog 2	43871	258	IILNLK
11E20.txt	1	CBX6_HUMAN	Chromobox protein homolog 2	43871	258	FSESVLR
11E20.txt	1	CBX6_HUMAN	Chromobox protein homolog 2	43871	258	LIAAFEQK
11E20.txt	1	CBX6_HUMAN	Chromobox protein homolog 2	43871	258	VFAAESIIK
11E20.txt	1	CBX6_HUMAN	Chromobox protein homolog 2	43871	258	VAAGVAGAAGGGGSGISK
11E20.txt	2	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	73	GGAGGSSVGTGGGTGGVGGGAGSEDS GDR
11E20.txt	5	ZN33A_HUMAN	Zinc finger protein 33A	94384	45	LLLNIK
11E21.txt	3	CBX2_HUMAN	Chromobox protein homolog 2	56046	47	LEYLVK
11E21.txt	3	CBX2_HUMAN	Chromobox protein homolog 2	56046	47	MTQSQQAASRLALK
11E21.txt	1	CBX6_HUMAN	Chromobox protein homolog 2	43871	187	TFLLK
11E21.txt	1	CBX6_HUMAN	Chromobox protein homolog 2	43871	187	IILNLK
11E21.txt	1	CBX6_HUMAN	Chromobox protein homolog 2	43871	187	IEYLVK
11E21.txt	1	CBX6_HUMAN	Chromobox protein homolog 2	43871	187	FSESVLR
11E21.txt	1	CBX6_HUMAN	Chromobox protein homolog 2	43871	187	LIAAFEQK

11E21.txt	1 CBX6_HUMAN	Chromobox protein homolog 2	43871	187 VFAAESIIK
11E22.txt	2 CBX6_HUMAN	Chromobox protein homolog 2	43871	107 TFLK
11E22.txt	2 CBX6_HUMAN	Chromobox protein homolog 2	43871	107 IEYLVK
11E22.txt	2 CBX6_HUMAN	Chromobox protein homolog 2	43871	107 FSESVLR
11E22.txt	2 CBX6_HUMAN	Chromobox protein homolog 2	43871	107 VFAAESIIK
11E22.txt	1 YBOX1_HUMAN	Nuclease-sensitive element-binding protein 1	35903	127 NGYGFINR
11E22.txt	1 YBOX1_HUMAN	Nuclease-sensitive element-binding protein 1	35903	127 NGYGFINR
11E22.txt	1 YBOX1_HUMAN	Nuclease-sensitive element-binding protein 1	35903	127 GAEAAANVTGPGGVPVQGSK
11E23.txt	2 CBX6_HUMAN	Chromobox protein homolog 2	43871	76 IILNLK
11E23.txt	2 CBX6_HUMAN	Chromobox protein homolog 2	43871	76 FSESVLR
11E23.txt	2 CBX6_HUMAN	Chromobox protein homolog 2	43871	76 VFAAESIIK
11E25.txt	1 CBX6_HUMAN	Chromobox protein homolog 2	43871	56 IILNLK
11E25.txt	1 CBX6_HUMAN	Chromobox protein homolog 2	43871	56 VFAAESIIK
11E26.txt	3 CBX6_HUMAN	Chromobox protein homolog 2	43871	103 IILNLK
11E26.txt	3 CBX6_HUMAN	Chromobox protein homolog 2	43871	103 IEYLVK
11E26.txt	3 CBX6_HUMAN	Chromobox protein homolog 2	43871	103 FSESVLR
11E26.txt	3 CBX6_HUMAN	Chromobox protein homolog 2	43871	103 VFAAESIIK
11E27.txt	3 CP135_HUMAN	Centrosomal protein of 135 kDa	133422	43 LEHGIK
11E27.txt	3 CP135_HUMAN	Centrosomal protein of 135 kDa	133422	43 LDTFLK
11E28.txt	1 CBX6_HUMAN	Chromobox protein homolog 2	43871	104 IEYLVK
11E28.txt	1 CBX6_HUMAN	Chromobox protein homolog 2	43871	104 LIAAFEQK
11E28.txt	1 CBX6_HUMAN	Chromobox protein homolog 2	43871	104 VFAAESIIK
11E29.txt	2 CBX6_HUMAN	Chromobox protein homolog 2	43871	149 LIAAFEQK
11E29.txt	2 CBX6_HUMAN	Chromobox protein homolog 2	43871	149 VFAAESIIK
11E29.txt	2 CBX6_HUMAN	Chromobox protein homolog 2	43871	149 VAAGVAGAAGGGGSGASK
11E30.txt	2 CBX6_HUMAN	Chromobox protein homolog 2	43871	165 IEYLVK
11E30.txt	2 CBX6_HUMAN	Chromobox protein homolog 2	43871	165 LIAAFEQK
11E30.txt	2 CBX6_HUMAN	Chromobox protein homolog 2	43871	165 VFAAESIIK
11E30.txt	2 CBX6_HUMAN	Chromobox protein homolog 2	43871	165 ISDVHFSVKPSASASSPK
11E31.txt	3 CBX6_HUMAN	Chromobox protein homolog 2	43871	156 IEYLVK
11E31.txt	3 CBX6_HUMAN	Chromobox protein homolog 2	43871	156 LIAAFEQK
11E31.txt	3 CBX6_HUMAN	Chromobox protein homolog 2	43871	156 VFAAESIIK
11E31.txt	3 CBX6_HUMAN	Chromobox protein homolog 2	43871	156 ISDVHFSVKPSASASSPK
11E32.txt	4 CBX6_HUMAN	Chromobox protein homolog 2	43871	118 TFLK
11E32.txt	4 CBX6_HUMAN	Chromobox protein homolog 2	43871	118 IEYLVK
11E32.txt	4 CBX6_HUMAN	Chromobox protein homolog 2	43871	118 LIAAFEQK
11E32.txt	4 CBX6_HUMAN	Chromobox protein homolog 2	43871	118 VFAAESIIK
11E33.txt	3 CBX6_HUMAN	Chromobox protein homolog 2	43871	96 IEYLVK
11E33.txt	3 CBX6_HUMAN	Chromobox protein homolog 2	43871	96 LIAAFEQK
11E33.txt	3 CBX6_HUMAN	Chromobox protein homolog 2	43871	96 VFAAESIIK
11E34.txt	6 CBX6_HUMAN	Chromobox protein homolog 2	43871	119 IILNLKVIDK
11E34.txt	6 CBX6_HUMAN	Chromobox protein homolog 2	43871	119 YSTWPEENILDSR
11E34.txt	6 CBX6_HUMAN	Chromobox protein homolog 2	43871	119 WKGWAIKYSTWPEENILDSR
11E35.txt	7 H4_HUMAN	Histone H4	11360	54 DNIQGITKPAIR
12E04.txt	19 MPPH1_HUMAN	M-phase phosphoprotein 1	210553	38 NSSVKKEQK
12E04.txt	19 MPPH1_HUMAN	M-phase phosphoprotein 1	210553	38 SNAMEEDLVKCNENK
12E04.txt	19 MPPH1_HUMAN	M-phase phosphoprotein 1	210553	38 FFKQQNEMILTAQLTEK
12E04.txt	19 MPPH1_HUMAN	M-phase phosphoprotein 1	210553	38 KHIETMSSSKLSNVEASKENVSQPK
12E04.txt	13 NUMA1_HUMAN	Nuclear mitotic apparatus protein 1	238260	45 EVAEK
12E04.txt	13 NUMA1_HUMAN	Nuclear mitotic apparatus protein 1	238260	45 AEDEWKAQVAR
12E04.txt	14 RAD17_HUMAN	Cell cycle checkpoint protein RAD17	77006	41 GMSLKSDAVLKSKS
12E04.txt	14 RAD17_HUMAN	Cell cycle checkpoint protein RAD17	77006	41 KIILVEDLPNQFYR
12E07.txt	1 UBE2O_HUMAN	Ubiquitin-conjugating enzyme E2 O	141265	221 VQWYPEGVK
12E07.txt	1 UBE2O_HUMAN	Ubiquitin-conjugating enzyme E2 O	141265	221 VVQSMQTLVLR
12E07.txt	1 UBE2O_HUMAN	Ubiquitin-conjugating enzyme E2 O	141265	221 VVVEEVQVVELK
12E07.txt	1 UBE2O_HUMAN	Ubiquitin-conjugating enzyme E2 O	141265	221 SGYPDIFPLPPLSK
12E07.txt	1 UBE2O_HUMAN	Ubiquitin-conjugating enzyme E2 O	141265	221 SFCPCGGTDSVSPPPSVITQENLGR
12E10.txt	2 ESF1_HUMAN	ESF1 homolog	98735	69 YYYAVVDCDSPETASK
12E11.txt	2 MOV10_HUMAN	Putative helicase MOV-10	113599	254 DVPLLPDVK
12E11.txt	2 MOV10_HUMAN	Putative helicase MOV-10	113599	254 VGSVEEFQGER
12E11.txt	2 MOV10_HUMAN	Putative helicase MOV-10	113599	254 HGVDFVEVQGPHEAR
12E11.txt	2 MOV10_HUMAN	Putative helicase MOV-10	113599	254 ETGDPGGQLVLGADPR
12E11.txt	2 MOV10_HUMAN	Putative helicase MOV-10	113599	254 LSPSTSGPHSHDYLPQER
12E11.txt	2 MOV10_HUMAN	Putative helicase MOV-10	113599	254 AHILACAPSNSGADLLCQR
12E12.txt	12 CBX6_HUMAN	Chromobox protein homolog 2	43871	42 LLPETVSPSAPSWR
12E12.txt	3 KRI1_HUMAN	Protein KRI1 homolog	83201	224 FEPPDSASVK
12E12.txt	3 KRI1_HUMAN	Protein KRI1 homolog	83201	224 AQEEADYIEWLK
12E12.txt	3 KRI1_HUMAN	Protein KRI1 homolog	83201	224 LQETSSQSYVEEQK
12E12.txt	3 KRI1_HUMAN	Protein KRI1 homolog	83201	224 TVVPCDFGLSTEEILAADDKELNR
12E13.txt	1 ILF3_HUMAN	Interleukin enhancer-binding factor 3	95279	162 AYAALAALEK
12E13.txt	1 ILF3_HUMAN	Interleukin enhancer-binding factor 3	95279	162 VLGMDFLPSK
12E13.txt	1 ILF3_HUMAN	Interleukin enhancer-binding factor 3	95279	162 VLQDMGLPTGAEGR
12E13.txt	1 ILF3_HUMAN	Interleukin enhancer-binding factor 3	95279	162 VLQDMGLPTGAEGR
12E13.txt	2 LAS1L_HUMAN	LAS1-like protein	83013	126 VECVLAEK
12E13.txt	2 LAS1L_HUMAN	LAS1-like protein	83013	126 SGNELPLAVASTADLIR
12E15.txt	10 BCAS3_HUMAN	Breast carcinoma-amplified sequence 3	101116	42 TLKSGLTMVKG
12E15.txt	10 BCAS3_HUMAN	Breast carcinoma-amplified sequence 3	101116	42 SDPVSMGSSRPVSDRR
12E15.txt	7 CD014_HUMAN	Uncharacterized protein C4orf14	78409	64 EGLGASEAVADIK
12E15.txt	5 PABP1_HUMAN	Polyadenylate-binding protein 1	70626	155 YQGVNLYVK
12E15.txt	5 PABP1_HUMAN	Polyadenylate-binding protein 1	70626	155 FSPAGPILSIR
12E15.txt	5 PABP1_HUMAN	Polyadenylate-binding protein 1	70626	155 ALYDTFSAFGNLSCK
12E16.txt	24 CASC5_HUMAN	Protein CASC5	265193	43 CLSNMREEK
12E16.txt	24 CASC5_HUMAN	Protein CASC5	265193	43 NDKTIVFSENHK
12E16.txt	24 CASC5_HUMAN	Protein CASC5	265193	43 SDIIAKNSLTDWTNK
12E16.txt	24 CASC5_HUMAN	Protein CASC5	265193	43 TVFYSSCNDAMEMTK
12E16.txt	24 CASC5_HUMAN	Protein CASC5	265193	43 NLLLELVQKEQTLAQIDFMQK
12E16.txt	12 CBX6_HUMAN	Chromobox protein homolog 2	43871	78 VAAGVAGAAGGGGSGASK
12E16.txt	8 GNL3_HUMAN	Guanine nucleotide-binding protein-like 3	61958	107 GGIPNVEGAAK

12E16.txt	8	GNL3_HUMAN	Guanine nucleotide-binding protein-like 3	61958	107	VGVIQFPNVGK
12E16.txt	8	GNL3_HUMAN	Guanine nucleotide-binding protein-like 3	61958	107	VIEASDVVLEVLDR
12E16.txt	2	IF2B1_HUMAN	Insulin-like growth factor 2 mRNA-binding protein 1	63417	576	MILEIMHK
12E16.txt	2	IF2B1_HUMAN	Insulin-like growth factor 2 mRNA-binding protein 1	63417	576	IAPPETPDSK
12E16.txt	2	IF2B1_HUMAN	Insulin-like growth factor 2 mRNA-binding protein 1	63417	576	IRDILAQVK
12E16.txt	2	IF2B1_HUMAN	Insulin-like growth factor 2 mRNA-binding protein 1	63417	576	QQQVDIPLR
12E16.txt	2	IF2B1_HUMAN	Insulin-like growth factor 2 mRNA-binding protein 1	63417	576	ISYSGQLVVK
12E16.txt	2	IF2B1_HUMAN	Insulin-like growth factor 2 mRNA-binding protein 1	63417	576	RLEIEHSVPK
12E16.txt	2	IF2B1_HUMAN	Insulin-like growth factor 2 mRNA-binding protein 1	63417	576	LKEENFFGPK
12E16.txt	2	IF2B1_HUMAN	Insulin-like growth factor 2 mRNA-binding protein 1	63417	576	MVIITGPPEAQFK
12E16.txt	2	IF2B1_HUMAN	Insulin-like growth factor 2 mRNA-binding protein 1	63417	576	MVIITGPPEAQFK
12E16.txt	2	IF2B1_HUMAN	Insulin-like growth factor 2 mRNA-binding protein 1	63417	576	LLVPTQYVGAIGK
12E16.txt	2	IF2B1_HUMAN	Insulin-like growth factor 2 mRNA-binding protein 1	63417	576	DQTPDENDQVIVK
12E16.txt	2	IF2B1_HUMAN	Insulin-like growth factor 2 mRNA-binding protein 1	63417	576	LKEENFFGPKKEEVK
12E16.txt	2	IF2B1_HUMAN	Insulin-like growth factor 2 mRNA-binding protein 1	63417	576	TVNELQNLTAEEVVPR
12E16.txt	2	IF2B1_HUMAN	Insulin-like growth factor 2 mRNA-binding protein 1	63417	576	ITISSLQDLTLYNPER
12E16.txt	2	IF2B1_HUMAN	Insulin-like growth factor 2 mRNA-binding protein 1	63417	576	SGYAFVDCPDEHWAMK
12E16.txt	9	IF2B2_HUMAN	Insulin-like growth factor 2 mRNA-binding protein 1	61804	89	MVIITGPPEAQFK
12E16.txt	9	IF2B2_HUMAN	Insulin-like growth factor 2 mRNA-binding protein 1	61804	89	MVIITGPPEAQFK
12E16.txt	9	IF2B2_HUMAN	Insulin-like growth factor 2 mRNA-binding protein 1	61804	89	ILVPTQFVGAIGK
12E16.txt	3	IF2B3_HUMAN	Insulin-like growth factor 2 mRNA-binding protein 1	63681	322	IKEENFVSPK
12E16.txt	3	IF2B3_HUMAN	Insulin-like growth factor 2 mRNA-binding protein 1	63681	322	KIQEILTQVK
12E16.txt	3	IF2B3_HUMAN	Insulin-like growth factor 2 mRNA-binding protein 1	63681	322	MVIITGPPEAQFK
12E16.txt	3	IF2B3_HUMAN	Insulin-like growth factor 2 mRNA-binding protein 1	63681	322	MVIITGPPEAQFK
12E16.txt	3	IF2B3_HUMAN	Insulin-like growth factor 2 mRNA-binding protein 1	63681	322	LLVPTQYVGAIGK
12E16.txt	3	IF2B3_HUMAN	Insulin-like growth factor 2 mRNA-binding protein 1	63681	322	DQTPDENDQVVVK
12E16.txt	3	IF2B3_HUMAN	Insulin-like growth factor 2 mRNA-binding protein 1	63681	322	TVNELQNLSSAEVVPR
12E16.txt	3	IF2B3_HUMAN	Insulin-like growth factor 2 mRNA-binding protein 1	63681	322	TGYAFVDCPDESWALK
12E16.txt	18	RGAG1_HUMAN	Retrotransposon gag domain-containing protein 1	144280	58	SLASGALSKPLMTPK
12E16.txt	18	RGAG1_HUMAN	Retrotransposon gag domain-containing protein 1	144280	58	AWTSETMSTPLMR
12E17.txt	2	CBX6_HUMAN	Chromobox protein homolog 2	43871	318	IILNLK
12E17.txt	2	CBX6_HUMAN	Chromobox protein homolog 2	43871	318	FSESVLR
12E17.txt	2	CBX6_HUMAN	Chromobox protein homolog 2	43871	318	VFAAESIIK
12E17.txt	2	CBX6_HUMAN	Chromobox protein homolog 2	43871	318	EFCNPEDFEK
12E17.txt	2	CBX6_HUMAN	Chromobox protein homolog 2	43871	318	VAAGVAGAAGGGGSGIGASK
12E17.txt	2	CBX6_HUMAN	Chromobox protein homolog 2	43871	318	LLPETVSPSPSWR
12E17.txt	2	CBX6_HUMAN	Chromobox protein homolog 2	43871	318	YSTWEPEENILDSR
12E17.txt	2	CBX6_HUMAN	Chromobox protein homolog 2	43871	318	ISDVHFSVKPSASASSPK
12E17.txt	2	CBX6_HUMAN	Chromobox protein homolog 2	43871	318	ISDVHFSVKPSASASSPK
12E17.txt	2	CBX6_HUMAN	Chromobox protein homolog 2	43871	318	WKGWAIKYSTWEPEENILDSR
12E17.txt	2	CBX6_HUMAN	Chromobox protein homolog 2	43871	318	RPLPRPDQGGSPGLRPPISPFSETVR
12E17.txt	4	LTV1_HUMAN	Protein LTV1 homolog	54821	93	NEQLTLHDER
12E17.txt	4	LTV1_HUMAN	Protein LTV1 homolog	54821	93	LQEVLDNYK
12E17.txt	4	LTV1_HUMAN	Protein LTV1 homolog	54821	93	SQRDPLAADESAPQR
12E18.txt	1	CBX6_HUMAN	Chromobox protein homolog 2	43871	436	IILNLK
12E18.txt	1	CBX6_HUMAN	Chromobox protein homolog 2	43871	436	IEYLVK
12E18.txt	1	CBX6_HUMAN	Chromobox protein homolog 2	43871	436	VFAAESIIK
12E18.txt	1	CBX6_HUMAN	Chromobox protein homolog 2	43871	436	EFCNPEDFEK
12E18.txt	1	CBX6_HUMAN	Chromobox protein homolog 2	43871	436	VAAGVAGAAGGGGSGIGASK
12E18.txt	1	CBX6_HUMAN	Chromobox protein homolog 2	43871	436	LLPETVSPSPSWR
12E18.txt	1	CBX6_HUMAN	Chromobox protein homolog 2	43871	436	YSTWEPEENILDSR
12E18.txt	1	CBX6_HUMAN	Chromobox protein homolog 2	43871	436	ISDVHFSVKPSASASSPK
12E18.txt	1	CBX6_HUMAN	Chromobox protein homolog 2	43871	436	EPEVLDLSLPESAATSK
12E18.txt	1	CBX6_HUMAN	Chromobox protein homolog 2	43871	436	VIDKGAGGGGAGGAGALARPK
12E18.txt	1	CBX6_HUMAN	Chromobox protein homolog 2	43871	436	FGAFALYKPPAPLVLAPSPGK
12E18.txt	1	CBX6_HUMAN	Chromobox protein homolog 2	43871	436	WKGWAIKYSTWEPEENILDSR
12E18.txt	8	ZN33A_HUMAN	Zinc finger protein 33A	94384	40	LLNLK
12E19.txt	1	CBX6_HUMAN	Chromobox protein homolog 2	43871	572	TFLLK
12E19.txt	1	CBX6_HUMAN	Chromobox protein homolog 2	43871	572	IILNLK
12E19.txt	1	CBX6_HUMAN	Chromobox protein homolog 2	43871	572	IEYLVK
12E19.txt	1	CBX6_HUMAN	Chromobox protein homolog 2	43871	572	IEYLVK
12E19.txt	1	CBX6_HUMAN	Chromobox protein homolog 2	43871	572	FSESVLR
12E19.txt	1	CBX6_HUMAN	Chromobox protein homolog 2	43871	572	LIAAFEQK
12E19.txt	1	CBX6_HUMAN	Chromobox protein homolog 2	43871	572	LIAAFEQK
12E19.txt	1	CBX6_HUMAN	Chromobox protein homolog 2	43871	572	EFCNPEDFEK
12E19.txt	1	CBX6_HUMAN	Chromobox protein homolog 2	43871	572	EFCNPEDFEK
12E19.txt	1	CBX6_HUMAN	Chromobox protein homolog 2	43871	572	VAAGVAGAAGGGGSGIGASK
12E19.txt	1	CBX6_HUMAN	Chromobox protein homolog 2	43871	572	VAAGVAGAAGGGGSGIGASK
12E19.txt	1	CBX6_HUMAN	Chromobox protein homolog 2	43871	572	VAAGVAGAAGGGGSGIGASK
12E19.txt	1	CBX6_HUMAN	Chromobox protein homolog 2	43871	572	VAAGVAGAAGGGGSGIGASK
12E19.txt	1	CBX6_HUMAN	Chromobox protein homolog 2	43871	572	VAAGVAGAAGGGGSGIGASK
12E19.txt	1	CBX6_HUMAN	Chromobox protein homolog 2	43871	572	VAAGVAGAAGGGGSGIGASK
12E19.txt	1	CBX6_HUMAN	Chromobox protein homolog 2	43871	572	LLPETVSPSPSWR
12E19.txt	1	CBX6_HUMAN	Chromobox protein homolog 2	43871	572	YSTWEPEENILDSR
12E19.txt	1	CBX6_HUMAN	Chromobox protein homolog 2	43871	572	ISDVHFSVKPSASASSPK
12E19.txt	1	CBX6_HUMAN	Chromobox protein homolog 2	43871	572	EPEVLDLSLPESAATSK
12E19.txt	1	CBX6_HUMAN	Chromobox protein homolog 2	43871	572	VIDKGAGGGGAGGAGALARPK
12E19.txt	1	CBX6_HUMAN	Chromobox protein homolog 2	43871	572	FGAFALYKPPAPLVLAPSPGK
12E19.txt	1	CBX6_HUMAN	Chromobox protein homolog 2	43871	572	RPLPRPDQGGSPGLRPPISPFSETVR
12E19.txt	6	DACT1_HUMAN	Dapper homolog 1	90119	48	ESAQLSGASPKEPSR
12E19.txt	6	DACT1_HUMAN	Dapper homolog 1	90119	48	LGNHASICGGSELDVAK
12E19.txt	7	WD40C_HUMAN	WD repeat-containing protein 40C	50803	45	EVGAR
12E19.txt	7	WD40C_HUMAN	WD repeat-containing protein 40C	50803	45	EVGAR
12E19.txt	7	WD40C_HUMAN	WD repeat-containing protein 40C	50803	45	VFASQWLNAR
12E19.txt	7	WD40C_HUMAN	WD repeat-containing protein 40C	50803	45	QQNIRPLCSR
12E19.txt	5	YC009_HUMAN	Uncharacterized protein FLJ36157	190890	52	MGCEK
12E19.txt	5	YC009_HUMAN	Uncharacterized protein FLJ36157	190890	52	LYINVKMGCEK
12E19.txt	5	YC009_HUMAN	Uncharacterized protein FLJ36157	190890	52	NSVTPVQASSDSFYNNK

12E19.txt	3 ZN33B_HUMAN	Zinc finger protein 33B	90683	60 LLLNIK
12E19.txt	3 ZN33B_HUMAN	Zinc finger protein 33B	90683	60 HYDCGESGNFR
12E19.txt	3 ZN33B_HUMAN	Zinc finger protein 33B	90683	60 SFCVKSGLLIHERK
12E20.txt	9 CBX6_HUMAN	Chromobox protein homolog 2	43871	312 FSESVLR
12E20.txt	9 CBX6_HUMAN	Chromobox protein homolog 2	43871	312 VFAAESIIK
12E20.txt	9 CBX6_HUMAN	Chromobox protein homolog 2	43871	312 VAAGVAGAAGGGGSGASK
12E20.txt	9 CBX6_HUMAN	Chromobox protein homolog 2	43871	312 LLPETVSPSAPSWR
12E20.txt	9 CBX6_HUMAN	Chromobox protein homolog 2	43871	312 YSTWEPEENILDSR
12E20.txt	9 CBX6_HUMAN	Chromobox protein homolog 2	43871	312 ISDVHFSVKPSASASSPK
12E20.txt	9 CBX6_HUMAN	Chromobox protein homolog 2	43871	312 ISDVHFSVKPSASASSPK
12E20.txt	9 CBX6_HUMAN	Chromobox protein homolog 2	43871	312 FGAFALYKPPAPLVPSPGK
12E20.txt	9 CBX6_HUMAN	Chromobox protein homolog 2	43871	312 WKGWAIKYSTWEPEENILDSR
12E21.txt	3 CBX6_HUMAN	Chromobox protein homolog 2	43871	185 IILNLK
12E21.txt	3 CBX6_HUMAN	Chromobox protein homolog 2	43871	185 EFCNPEDFEK
12E21.txt	3 CBX6_HUMAN	Chromobox protein homolog 2	43871	185 LLPETVSPSAPSWR
12E21.txt	3 CBX6_HUMAN	Chromobox protein homolog 2	43871	185 YSTWEPEENILDSR
12E21.txt	3 CBX6_HUMAN	Chromobox protein homolog 2	43871	185 WKGWAIKYSTWEPEENILDSR
12E21.txt	9 LA_HUMAN	Lupus La protein	46808	88 LTDFNVIVEALSK
12E21.txt	17 SCML2_HUMAN	Sex comb on midleg-like protein 2	77257	40 SDVMMKMYGLK
12E21.txt	17 SCML2_HUMAN	Sex comb on midleg-like protein 2	77257	40 LRLRLDGSNDR
12E21.txt	17 SCML2_HUMAN	Sex comb on midleg-like protein 2	77257	40 GQTVNEDSMIVK
12E21.txt	7 YBOX1_HUMAN	Nuclease-sensitive element-binding protein 1	35903	96 GAEANVTGPGGVPVQGSK
12E21.txt	7 YBOX1_HUMAN	Nuclease-sensitive element-binding protein 1	35903	96 NDTKEDVVFHQTAIK
12E21.txt	16 ZN33A_HUMAN	Zinc finger protein 33A	94384	41 LLLNIK
12E22.txt	6 CBX6_HUMAN	Chromobox protein homolog 2	43871	148 IILNLK
12E22.txt	6 CBX6_HUMAN	Chromobox protein homolog 2	43871	148 YSTWEPEENILDSR
12E22.txt	6 CBX6_HUMAN	Chromobox protein homolog 2	43871	148 ISDVHFSVKPSASASSPK
12E22.txt	6 CBX6_HUMAN	Chromobox protein homolog 2	43871	148 FGAFALYKPPAPLVPSPGK
12E22.txt	6 CBX6_HUMAN	Chromobox protein homolog 2	43871	148 WKGWAIKYSTWEPEENILDSR
12E22.txt	1 ILF2_HUMAN	Interleukin enhancer-binding factor 2	43035	284 VLQSALAAIR
12E22.txt	1 ILF2_HUMAN	Interleukin enhancer-binding factor 2	43035	284 QPLALNAVYR
12E22.txt	1 ILF2_HUMAN	Interleukin enhancer-binding factor 2	43035	284 ILPTLEAVALGNK
12E22.txt	1 ILF2_HUMAN	Interleukin enhancer-binding factor 2	43035	284 VKPAPDETSFSEALLK
12E22.txt	3 YBOX1_HUMAN	Nuclease-sensitive element-binding protein 1	35903	237 EDVVFHQTAIK
12E22.txt	3 YBOX1_HUMAN	Nuclease-sensitive element-binding protein 1	35903	237 EDVVFHQTAIKK
12E22.txt	3 YBOX1_HUMAN	Nuclease-sensitive element-binding protein 1	35903	237 GAEANVTGPGGVPVQGSK
12E22.txt	3 YBOX1_HUMAN	Nuclease-sensitive element-binding protein 1	35903	237 NDTKEDVVFHQTAIK
12E22.txt	3 YBOX1_HUMAN	Nuclease-sensitive element-binding protein 1	35903	237 NDTKEDVVFHQTAIK
12E22.txt	3 YBOX1_HUMAN	Nuclease-sensitive element-binding protein 1	35903	237 EDGNEEDKENQGDDETQGGQPPQR
12E23.txt	4 CBX6_HUMAN	Chromobox protein homolog 2	43871	83 YSTWEPEENILDSR
12E23.txt	4 CBX6_HUMAN	Chromobox protein homolog 2	43871	83 ISDVHFSVKPSASASSPK
12E23.txt	4 CBX6_HUMAN	Chromobox protein homolog 2	43871	83 WKGWAIKYSTWEPEENILDSR
12E23.txt	8 MPPH1_HUMAN	M-phase phosphoprotein 1	210553	49 EQLNNQKVEEAIQYER
12E23.txt	8 MPPH1_HUMAN	M-phase phosphoprotein 1	210553	49 FFKQQNEMILTAQLTEK
12E23.txt	8 MPPH1_HUMAN	M-phase phosphoprotein 1	210553	49 DLLKQSRILIFTYGLTNSGK
12E23.txt	7 PHC2_HUMAN	Polyhomeotic-like protein 2	90657	61 AQEIDGQALLLKL
12E23.txt	2 RING2_HUMAN	E3 ubiquitin-protein ligase RING2	37632	91 FCADCITAIR
12E23.txt	2 RING2_HUMAN	E3 ubiquitin-protein ligase RING2	37632	91 IYPSRDEYEAHQER
12E24.txt	3 DBPA_HUMAN	DNA-binding protein A	40066	88 SVGDGETVEFDVVEGEK
12E25.txt	1 CBX6_HUMAN	Chromobox protein homolog 2	43871	72 IILNLK
12E25.txt	1 CBX6_HUMAN	Chromobox protein homolog 2	43871	72 FGAFALYKPPAPLVPSPGK
12E26.txt	3 CBX6_HUMAN	Chromobox protein homolog 2	43871	115 GKPK
12E26.txt	3 CBX6_HUMAN	Chromobox protein homolog 2	43871	115 IILNLK
12E26.txt	3 CBX6_HUMAN	Chromobox protein homolog 2	43871	115 YSTWEPEENILDSR
12E26.txt	3 CBX6_HUMAN	Chromobox protein homolog 2	43871	115 ISDVHFSVKPSASASSPK
12E26.txt	3 CBX6_HUMAN	Chromobox protein homolog 2	43871	115 ISDVHFSVKPSASASSPK
12E26.txt	7 ZN33A_HUMAN	Zinc finger protein 33A	94384	42 LLLNIK
12E27.txt	9 BPTF_HUMAN	Nucleosome-remodeling factor subunit BPTF	324019	49 MIVCNQVMK
12E27.txt	9 BPTF_HUMAN	Nucleosome-remodeling factor subunit BPTF	324019	49 SNGELSESPGAGKGSSTR
12E27.txt	9 BPTF_HUMAN	Nucleosome-remodeling factor subunit BPTF	324019	49 VVSGNVPKVNNINKIIPENDIK
12E27.txt	11 H12_HUMAN	Histone H1.2	21352	45 ASGPPVSELITK
12E27.txt	12 H14_HUMAN	Histone H1.4	21852	44 SPAKAKAVKPK
12E27.txt	12 H14_HUMAN	Histone H1.4	21852	44 ASGPPVSELITK
12E27.txt	8 THOC4_HUMAN	THO complex subunit 4	26872	64 SLGTADVHFER
12E28.txt	6 CAF1A_HUMAN	Chromatin assembly factor 1 subunit A	105223	47 FVKGSTK
12E28.txt	6 CAF1A_HUMAN	Chromatin assembly factor 1 subunit A	105223	47 KFKVKGSTK
12E28.txt	3 CBX6_HUMAN	Chromobox protein homolog 2	43871	132 GKPK
12E28.txt	3 CBX6_HUMAN	Chromobox protein homolog 2	43871	132 RGPKPKTFLLKAR
12E28.txt	3 CBX6_HUMAN	Chromobox protein homolog 2	43871	132 YSTWEPEENILDSR
12E28.txt	3 CBX6_HUMAN	Chromobox protein homolog 2	43871	132 ISDVHFSVKPSASASSPK
12E28.txt	3 CBX6_HUMAN	Chromobox protein homolog 2	43871	132 WKGWAIKYSTWEPEENILDSR
12E28.txt	8 CE350_HUMAN	Centrosome-associated protein 350	350716	46 IEDAFSKEGK
12E28.txt	8 CE350_HUMAN	Centrosome-associated protein 350	350716	46 LKSAGESKLDSDDDDTK
12E28.txt	8 CE350_HUMAN	Centrosome-associated protein 350	350716	46 LEVAPTSTAVCDSDVMDTKSSTSATRK
12E28.txt	5 H1X_HUMAN	Histone H1x	22474	50 YSQLVETIR
12E29.txt	7 NCOR2_HUMAN	Nuclear receptor corepressor 2	274034	49 EITMVEQQISKLK
12E29.txt	7 NCOR2_HUMAN	Nuclear receptor corepressor 2	274034	49 MGSKSPGNTSQPPAFFSK
12E29.txt	7 NCOR2_HUMAN	Nuclear receptor corepressor 2	274034	49 NRLLSPRPLLTPTGDPR
12E29.txt	7 NCOR2_HUMAN	Nuclear receptor corepressor 2	274034	49 KGLLEHGRNWSAIARMVSGK
12E31.txt	10 CBX6_HUMAN	Chromobox protein homolog 2	43871	64 GRIEYLVKWK
12E31.txt	10 CBX6_HUMAN	Chromobox protein homolog 2	43871	64 YSTWEPEENILDSR
12E31.txt	10 CBX6_HUMAN	Chromobox protein homolog 2	43871	64 ISDVHFSVKPSASASSPK
12E31.txt	13 DNMT1_HUMAN	DNA (cytosine-5)-methyltransferase 1	183050	45 FVSNITR
12E31.txt	13 DNMT1_HUMAN	DNA (cytosine-5)-methyltransferase 1	183050	45 FCVSCAR
12E31.txt	13 DNMT1_HUMAN	DNA (cytosine-5)-methyltransferase 1	183050	45 KTTKPEPTEK
12E32.txt	3 CBX6_HUMAN	Chromobox protein homolog 2	43871	148 LIAAFEQK
12E32.txt	3 CBX6_HUMAN	Chromobox protein homolog 2	43871	148 VFAAESIIK
12E32.txt	3 CBX6_HUMAN	Chromobox protein homolog 2	43871	148 YSTWEPEENILDSR

12E32.txt	3	CBX6_HUMAN	Chromobox protein homolog 2	43871	148	ISDVHFSVKPSASASSPK
12E32.txt	3	CBX6_HUMAN	Chromobox protein homolog 2	43871	148	ISDVHFSVKPSASASSPK
12E32.txt	3	CBX6_HUMAN	Chromobox protein homolog 2	43871	148	WKGWAIKYSTWEPEENILDSR
12E32.txt	11	ELP4_HUMAN	Elongator complex protein 4	46588	59	ASVTNDSGPR
12E32.txt	11	ELP4_HUMAN	Elongator complex protein 4	46588	59	MEIGPVSSSR
12E32.txt	11	ELP4_HUMAN	Elongator complex protein 4	46588	59	DLAFKLRK
12E32.txt	18	ZMYM4_HUMAN	Zinc finger MYM-type protein 4	172677	43	SIVAVEPR
12E32.txt	18	ZMYM4_HUMAN	Zinc finger MYM-type protein 4	172677	43	MVTSAGVQVCNSCK
12E32.txt	18	ZMYM4_HUMAN	Zinc finger MYM-type protein 4	172677	43	DFCSQSCLSTYELK
12E34.txt	5	CBX6_HUMAN	Chromobox protein homolog 2	43871	130	IEYLVK
12E34.txt	5	CBX6_HUMAN	Chromobox protein homolog 2	43871	130	IEYLVK
12E34.txt	5	CBX6_HUMAN	Chromobox protein homolog 2	43871	130	LIAAFEQK
12E34.txt	5	CBX6_HUMAN	Chromobox protein homolog 2	43871	130	LLPETVSPSAPSWR
12E34.txt	5	CBX6_HUMAN	Chromobox protein homolog 2	43871	130	YSTWEPEENILDSR
12E34.txt	5	CBX6_HUMAN	Chromobox protein homolog 2	43871	130	WKGWAIKYSTWEPEENILDSR
12E34.txt	6	H2B1B_HUMAN	Histone H2B type 1-B	13942	118	LLLPGLAK
12E34.txt	6	H2B1B_HUMAN	Histone H2B type 1-B	13942	118	AMGIMNSFVNDIFER
12E34.txt	13	H31T_HUMAN	Histone H3.1t	15499	47	YRPGTVLR
12E34.txt	13	H31T_HUMAN	Histone H3.1t	15499	47	LPFQRLMR
12E34.txt	13	H31T_HUMAN	Histone H3.1t	15499	47	KPHRYRPGTVLR
12E34.txt	20	ULK2_HUMAN	Serine/threonine-protein kinase ULK2	112654	41	ECAHCLLVQGSER
12E34.txt	20	ULK2_HUMAN	Serine/threonine-protein kinase ULK2	112654	41	IPKTVASSNLLALVTR
12E34.txt	20	ULK2_HUMAN	Serine/threonine-protein kinase ULK2	112654	41	HLNVMLMFTFECVLDLTAMR
12E35.txt	17	AHDC1_HUMAN	A.T hook DNA-binding motif-containing protein 1	168245	41	CSFSDFEFGIGK
12E35.txt	17	AHDC1_HUMAN	A.T hook DNA-binding motif-containing protein 1	168245	41	NVVVYFPGEMPIILK
12E35.txt	17	AHDC1_HUMAN	A.T hook DNA-binding motif-containing protein 1	168245	41	AAETFPKLVPPPSAMAR
12E35.txt	18	BRWD1_HUMAN	Bromodomain and WD repeat-containing protein 1	262756	40	YLSAGPCRR
12E35.txt	18	BRWD1_HUMAN	Bromodomain and WD repeat-containing protein 1	262756	40	SSRKLPHRNASAAAK
12E35.txt	18	BRWD1_HUMAN	Bromodomain and WD repeat-containing protein 1	262756	40	VIHDSQETAKEVKKR
12E35.txt	13	H2A1B_HUMAN	Histone H2A type 1-B	14127	43	VTIAQGGVLPNIQAVLLPK
12E35.txt	10	PPR3A_HUMAN	Protein phosphatase 1 regulatory subunit 3A	125759	52	ISVAGIGASNR
12E35.txt	10	PPR3A_HUMAN	Protein phosphatase 1 regulatory subunit 3A	125759	52	DTVGVIVYDNDFEK
12E35.txt	10	PPR3A_HUMAN	Protein phosphatase 1 regulatory subunit 3A	125759	52	DTVGVIVYDNDFEK
12E36.txt	8	H4_HUMAN	Histone H4	11360	68	DNIQGITKPAIR
12E36.txt	14	RAD26_HUMAN	Putative DNA repair and recombination protein RAD26-like	81094	42	DVKVLR
12E36.txt	14	RAD26_HUMAN	Putative DNA repair and recombination protein RAD26-like	81094	42	SEERLKIVK
12E36.txt	11	XYLB_HUMAN	Xylulose kinase	57374	50	TKILATGGASHNR
12E38.txt	3	KHDR1_HUMAN	KH domain-containing, RNA-binding, signal transduction-associated protein 1	48197	49	GALVR
12E38.txt	3	KHDR1_HUMAN	KH domain-containing, RNA-binding, signal transduction-associated protein 1	48197	49	NMKLKER
12E38.txt	3	KHDR1_HUMAN	KH domain-containing, RNA-binding, signal transduction-associated protein 1	48197	49	MQRDDPAARMSR

Tableau S4 : Peptides identifiés dans les 2 purifications CBX7-TAP (n°9 et 10)

band	hit	accession	description	mass	score	peptide
10E07.txt	2	UBP20_HUMAN	Ubiquitin carboxyl-terminal hydrolase 20	101938	46	GGGSSQAETELLIPDEAGR
10E10.txt	6	SRBD1_HUMAN	S1 RNA-binding domain-containing protein 1	111705	39	MSECLLKAMLNCK
10E10.txt	6	SRBD1_HUMAN	S1 RNA-binding domain-containing protein 1	111705	39	EMGPLDPNLRSAVSIAR
10E10.txt	6	SRBD1_HUMAN	S1 RNA-binding domain-containing protein 1	111705	39	INSFLEKEGMKIAER
10E15.txt	2	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	334	TTGNATVDHLK
10E15.txt	2	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	334	FCSDCIVTALR
10E15.txt	2	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	334	SLRPDPNFDALISK
10E15.txt	2	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	334	LHNQALSSSIEEGLR
10E15.txt	2	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	334	GGGAGGSSVGTGGGGTGGVGGGAGSEDS
10E15.txt	2	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	334	GDR
10E16.txt	5	PCGF6_HUMAN	Polycomb group RING finger protein 6	39023	56	LVINLEER
10E16.txt	5	PCGF6_HUMAN	Polycomb group RING finger protein 6	39023	56	LQDIVYK
10E17.txt	1	CSK21_HUMAN	Casein kinase II subunit alpha	45115	255	QLYQTLTDYDIR
10E17.txt	1	CSK21_HUMAN	Casein kinase II subunit alpha	45115	255	TPALVFEHVNNTDFK
10E17.txt	1	CSK21_HUMAN	Casein kinase II subunit alpha	45115	255	GGPNIITLADIVKDPVSR
10E17.txt	8	PCGF6_HUMAN	Polycomb group RING finger protein 6	39023	56	GLEVPKPAVPQPVPSSK
10E17.txt	11	PHC2_HUMAN	Polyhomeotic-like protein 2	90657	50	LPQDHTTTTTDSEMEEPYLQESK
10E17.txt	5	RING2_HUMAN	E3 ubiquitin-protein ligase RING2	37632	70	FCADCITLALR
10E18.txt	3	BMI1_HUMAN	Polycomb complex protein BMI-1	36925	100	LVPGLFK
10E18.txt	3	BMI1_HUMAN	Polycomb complex protein BMI-1	36925	100	TLQDIVYK
10E18.txt	3	BMI1_HUMAN	Polycomb complex protein BMI-1	36925	100	YCPICDVQVHK
10E18.txt	5	PCGF2_HUMAN	Polycomb group RING finger protein 2	37764	75	CPAAMTYMHLAKFLRNKMDVPSK
10E18.txt	5	PCGF2_HUMAN	Polycomb group RING finger protein 2	37764	75	LVPGLFK
10E18.txt	5	PCGF2_HUMAN	Polycomb group RING finger protein 2	37764	75	TLQDIVYK
10E18.txt	5	PCGF2_HUMAN	Polycomb group RING finger protein 2	37764	75	YLETNKYCPMCDVQVHK
10E18.txt	2	PHC2_HUMAN	Polyhomeotic-like protein 2	90657	111	YAQGFLPEK
10E18.txt	2	PHC2_HUMAN	Polyhomeotic-like protein 2	90657	111	LPQDHTTTTTDSEMEEPYLQESK
10E18.txt	1	RING2_HUMAN	E3 ubiquitin-protein ligase RING2	37632	571	YLAVR
10E18.txt	1	RING2_HUMAN	E3 ubiquitin-protein ligase RING2	37632	571	LALIELR
10E18.txt	1	RING2_HUMAN	E3 ubiquitin-protein ligase RING2	37632	571	FCADCITLALR
10E18.txt	1	RING2_HUMAN	E3 ubiquitin-protein ligase RING2	37632	571	TWELSLYELQR
10E18.txt	1	RING2_HUMAN	E3 ubiquitin-protein ligase RING2	37632	571	GESNQMLDITASEK
10E18.txt	1	RING2_HUMAN	E3 ubiquitin-protein ligase RING2	37632	571	VNKPMEELYAPTK
10E18.txt	1	RING2_HUMAN	E3 ubiquitin-protein ligase RING2	37632	571	SLRPDPNFDALISK
10E18.txt	1	RING2_HUMAN	E3 ubiquitin-protein ligase RING2	37632	571	RSLRPDPNFDALISK
10E18.txt	1	RING2_HUMAN	E3 ubiquitin-protein ligase RING2	37632	571	IYPSRDEYEAHQER
10E18.txt	1	RING2_HUMAN	E3 ubiquitin-protein ligase RING2	37632	571	IYPSRDEYEAHQER
10E18.txt	1	RING2_HUMAN	E3 ubiquitin-protein ligase RING2	37632	571	HNNQALSHSIEEGLK

10E19.txt	3	CBX7_HUMAN	Chromobox protein homolog 7	28323	55	EAQAAEGFFR
10E20.txt	1	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	49	SLRPDPNFDALISK
10E21.txt	4	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	49	SLRPDPNFDALISK
10E23.txt	6	CSK2B_HUMAN	Casein kinase II subunit beta	24926	103	RPANQFVPR
10E23.txt	6	CSK2B_HUMAN	Casein kinase II subunit beta	24926	103	YQQGDFGYCPR
10E28.txt	11	UBE4A_HUMAN	Ubiquitin conjugation factor E4 A	122561	41	ELNDEERKIK
10E28.txt	11	UBE4A_HUMAN	Ubiquitin conjugation factor E4 A	122561	41	ELNDEERKIKNVHMR
merge9E04.txt	8	TPR_HUMAN	Nucleoprotein TPR	267293	45	IAHLAGVK
merge9E04.txt	8	TPR_HUMAN	Nucleoprotein TPR	267293	45	LSDKVVASVK
merge9E04.txt	8	TPR_HUMAN	Nucleoprotein TPR	267293	45	SQYEGRISRRLER
merge9E09.txt	12	SPEC1_HUMAN	Sperm antigen with calponin homology and coiled-	118512	48	LRSELKYYK
merge9E09.txt	12	SPEC1_HUMAN	Sperm antigen with calponin homology and coiled-	118512	48	VRPLPAASSGMKSSK
merge9E09.txt	12	SPEC1_HUMAN	Sperm antigen with calponin homology and coiled-	118512	48	TLEECRVTLLEGLKMENGLK
merge9E10.txt	16	RAD17_HUMAN	Cell cycle checkpoint protein RAD17	77055	45	IVTIEANKNGGK
merge9E10.txt	16	RAD17_HUMAN	Cell cycle checkpoint protein RAD17	77055	45	KIILVEDLPNQFYR
merge9E10.txt	1	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	384	IALER
merge9E10.txt	1	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	384	TTGNATVDHLSK
merge9E10.txt	1	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	384	FCSDCIVTLR
merge9E10.txt	1	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	384	SLRPDPNFDALISK
merge9E10.txt	1	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	384	LHNQALSSSIEEGLR
merge9E10.txt	1	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	384	VSRPLELCYAPTKDPK
merge9E10.txt	1	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	384	GGGAGGSSVGTGGGGTGGVGGGAGSEDS
merge9E10.txt	1	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	384	GDR
merge9E10.txt	1	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	384	SLHSELMLCPICLDMLKNTMTTK
merge9E11.txt	6	PCGF6_HUMAN	Polycomb group RING finger protein 6	39023	95	VLESVFR
merge9E11.txt	6	PCGF6_HUMAN	Polycomb group RING finger protein 6	39023	95	VSGEATIGHVEK
merge9E12.txt	3	BMI1_HUMAN	Polycomb complex protein BMI-1	36925	199	CPAAMTYMHLR
merge9E12.txt	3	BMI1_HUMAN	Polycomb complex protein BMI-1	36925	199	DGLTNAGELESDSGSDK
merge9E12.txt	3	BMI1_HUMAN	Polycomb complex protein BMI-1	36925	199	TLQDIVYK
merge9E12.txt	3	BMI1_HUMAN	Polycomb complex protein BMI-1	36925	199	YCPICDVQVHK
merge9E12.txt	11	PCGF2_HUMAN	Polycomb group RING finger protein 2	37764	47	TLQDIVYK
merge9E12.txt	11	PCGF2_HUMAN	Polycomb group RING finger protein 2	37764	47	YLETNKYCPMCDVQVHK
merge9E12.txt	9	PHC2_HUMAN	Polyhomeotic-like protein 2	90657	66	LGPALK
merge9E12.txt	9	PHC2_HUMAN	Polyhomeotic-like protein 2	90657	66	YAQGFLEPK
merge9E12.txt	1	RING2_HUMAN	E3 ubiquitin-protein ligase RING2	37632	627	YLAVR
merge9E12.txt	1	RING2_HUMAN	E3 ubiquitin-protein ligase RING2	37632	627	LALEELR
merge9E12.txt	1	RING2_HUMAN	E3 ubiquitin-protein ligase RING2	37632	627	TSGNATVDHLSK
merge9E12.txt	1	RING2_HUMAN	E3 ubiquitin-protein ligase RING2	37632	627	FCADCITLALR
merge9E12.txt	1	RING2_HUMAN	E3 ubiquitin-protein ligase RING2	37632	627	GESNQMNLDTASEK
merge9E12.txt	1	RING2_HUMAN	E3 ubiquitin-protein ligase RING2	37632	627	VNKPMELYAPTK
merge9E12.txt	1	RING2_HUMAN	E3 ubiquitin-protein ligase RING2	37632	627	VNKPMELYAPTK
merge9E12.txt	1	RING2_HUMAN	E3 ubiquitin-protein ligase RING2	37632	627	VNKPMELYAPTK
merge9E12.txt	1	RING2_HUMAN	E3 ubiquitin-protein ligase RING2	37632	627	SLRPDPNFDALISK
merge9E12.txt	1	RING2_HUMAN	E3 ubiquitin-protein ligase RING2	37632	627	SKGESNQMNLDTASEK
merge9E12.txt	1	RING2_HUMAN	E3 ubiquitin-protein ligase RING2	37632	627	IYPSRDEYEAHQER
merge9E12.txt	1	RING2_HUMAN	E3 ubiquitin-protein ligase RING2	37632	627	HNNQALSHSIEEGLK
merge9E12.txt	1	RING2_HUMAN	E3 ubiquitin-protein ligase RING2	37632	627	TPQEAITDGLIIVSPR
merge9E13.txt	2	CBX7_HUMAN	Chromobox protein homolog 7	28323	232	GWPPK
merge9E13.txt	2	CBX7_HUMAN	Chromobox protein homolog 7	28323	232	VEYLVK
merge9E13.txt	2	CBX7_HUMAN	Chromobox protein homolog 7	28323	232	LYSMDLR
merge9E13.txt	2	CBX7_HUMAN	Chromobox protein homolog 7	28323	232	LYSMDLR
merge9E13.txt	2	CBX7_HUMAN	Chromobox protein homolog 7	28323	232	GKVEYLVK
merge9E13.txt	2	CBX7_HUMAN	Chromobox protein homolog 7	28323	232	EAQAAEGFFR
merge9E13.txt	2	CBX7_HUMAN	Chromobox protein homolog 7	28323	232	LVMAYEEKEER
merge9E13.txt	2	CBX7_HUMAN	Chromobox protein homolog 7	28323	232	LVMAYEEKEER
merge9E17.txt	6	RING2_HUMAN	E3 ubiquitin-protein ligase RING2	37632	60	TSGNATVDHLSK
merge9E18.txt	2	CBX7_HUMAN	Chromobox protein homolog 7	28323	91	GKVEYLVK
merge9E18.txt	2	CBX7_HUMAN	Chromobox protein homolog 7	28323	91	LVMAYEEKEER
merge9E18.txt	7	RING2_HUMAN	E3 ubiquitin-protein ligase RING2	37632	45	TSGNATVDHLSK

Tableau S5 : Peptides identifiés dans les 2 purifications CBX8-TAP (n°6 et 13)

band	hit	accession	description	mass	score	peptide
13E04.txt	1	RRP5_HUMAN	Protein RRP5 homolog	208570	272	KQNKGAK
13E04.txt	1	RRP5_HUMAN	Protein RRP5 homolog	208570	272	VFYTGQVVK
13E04.txt	1	RRP5_HUMAN	Protein RRP5 homolog	208570	272	VKDYGCIVK
13E04.txt	1	RRP5_HUMAN	Protein RRP5 homolog	208570	272	TSIIAQYLK
13E04.txt	1	RRP5_HUMAN	Protein RRP5 homolog	208570	272	VGYLNCIVEK
13E04.txt	1	RRP5_HUMAN	Protein RRP5 homolog	208570	272	AIFENTLSTYPK
13E04.txt	1	RRP5_HUMAN	Protein RRP5 homolog	208570	272	KPALVSTVEGGQDPK
13E04.txt	1	RRP5_HUMAN	Protein RRP5 homolog	208570	272	VTPNEGLTVSFPFGK
13E06.txt	6	DNL3_HUMAN	DNA ligase 3	102625	44	LIKHDLKMNISGAK
13E06.txt	6	DNL3_HUMAN	DNA ligase 3	102625	44	SEAHTADGISIRFPR
13E06.txt	6	DNL3_HUMAN	DNA ligase 3	102625	44	SVEYAMKKCPNGMFSEIK
13E07.txt	3	BCLF1_HUMAN	Bcl-2-associated transcription factor 1	106059	84	ETGYVVERPSTTK
13E07.txt	3	BCLF1_HUMAN	Bcl-2-associated transcription factor 1	106059	84	SQEEPDKTFEHDPSIESIDFNK
13E07.txt	8	ZFR_HUMAN	Zinc finger RNA-binding protein	116939	39	AISSASSPQSPGDALR
13E08.txt	2	MOV10_HUMAN	Putative helicase MOV-10	113599	268	HGLGYSLER
13E08.txt	2	MOV10_HUMAN	Putative helicase MOV-10	113599	268	VLITTLITAGR
13E08.txt	2	MOV10_HUMAN	Putative helicase MOV-10	113599	268	KGPDGYDPQFITK
13E08.txt	2	MOV10_HUMAN	Putative helicase MOV-10	113599	268	HGVDEVVQGPHEAR
13E08.txt	2	MOV10_HUMAN	Putative helicase MOV-10	113599	268	ETGDPGGQLVLGADPR
13E08.txt	2	MOV10_HUMAN	Putative helicase MOV-10	113599	268	NGGTQSVTLTHLFLPCR
13E09.txt	1	KRI1_HUMAN	Protein KRI1 homolog	83201	241	FEEDPSASVK

13E09.txt	1	KRI1_HUMAN	Protein KRI1 homolog	83201	241	AQEEADYIEWLK
13E09.txt	1	KRI1_HUMAN	Protein KRI1 homolog	83201	241	LQETSSQSYVEEQK
13E09.txt	1	KRI1_HUMAN	Protein KRI1 homolog	83201	241	SLCREEAETPAEATGKPPQR
13E09.txt	3	NOL1_HUMAN	Putative RNA methyltransferase NOL1	89247	105	FLPAVSDENSKR
13E09.txt	3	NOL1_HUMAN	Putative RNA methyltransferase NOL1	89247	105	VLLDAPCSGTVISK
13E10.txt	6	DGKZ_HUMAN	Diacylglycerol kinase zeta	124044	67	FYRIDR
13E10.txt	6	DGKZ_HUMAN	Diacylglycerol kinase zeta	124044	67	SPTCQKLSPK
13E10.txt	6	DGKZ_HUMAN	Diacylglycerol kinase zeta	124044	67	RASGTTAGTMLPTR
13E10.txt	6	DGKZ_HUMAN	Diacylglycerol kinase zeta	124044	67	RPSSVGLPTGKARRR
13E10.txt	6	DGKZ_HUMAN	Diacylglycerol kinase zeta	124044	67	RASGTTAGTMLPTRVPLSR
13E10.txt	8	NCOAT_HUMAN	Bifunctional protein NCOAT	102849	59	GRSTELIPR
13E10.txt	8	NCOAT_HUMAN	Bifunctional protein NCOAT	102849	59	EIPVESIEEVSK
13E10.txt	8	NCOAT_HUMAN	Bifunctional protein NCOAT	102849	59	EIPVESIEEVSKI
13E10.txt	8	NCOAT_HUMAN	Bifunctional protein NCOAT	102849	59	ISWIPFMQKEYTKPNGDK
13E10.txt	3	TSR1_HUMAN	Pre-rRNA-processing protein TSR1 homolog	91752	100	ENLPQDYAR
13E10.txt	3	TSR1_HUMAN	Pre-rRNA-processing protein TSR1 homolog	91752	100	WTYDYPVPEVPVWLK
13E11.txt	10	EP300_HUMAN	Histone acetyltransferase p300	263973	43	TVLSNNLSPFAMDKK
13E11.txt	10	EP300_HUMAN	Histone acetyltransferase p300	263973	43	RMENLVAVYARKVEGDMYESANNR
13E12.txt	6	ILF3_HUMAN	Interleukin enhancer-binding factor 3	95279	76	EATDAIGHLDR
13E12.txt	6	ILF3_HUMAN	Interleukin enhancer-binding factor 3	95279	76	YELISETGGSHDKR
13E14.txt	9	K0020_HUMAN	Pumilio domain-containing protein KIAA0020	73538	47	EAVVYLAHTHDGAR
13E15.txt	3	IF2B1_HUMAN	Insulin-like growth factor 2 mRNA-binding protein 1	63417	169	IAPPETPDSK
13E15.txt	3	IF2B1_HUMAN	Insulin-like growth factor 2 mRNA-binding protein 1	63417	169	ISYSGQLVK
13E15.txt	3	IF2B1_HUMAN	Insulin-like growth factor 2 mRNA-binding protein 1	63417	169	MVIITGPPEAQFK
13E15.txt	3	IF2B1_HUMAN	Insulin-like growth factor 2 mRNA-binding protein 1	63417	169	DQTPDENQVIVK
13E15.txt	3	IF2B1_HUMAN	Insulin-like growth factor 2 mRNA-binding protein 1	63417	169	AISVHSTPEGSSACK
13E15.txt	3	IF2B1_HUMAN	Insulin-like growth factor 2 mRNA-binding protein 1	63417	169	LKEENFFGPKKEEVK
13E15.txt	3	IF2B1_HUMAN	Insulin-like growth factor 2 mRNA-binding protein 1	63417	169	TVNLQNLTAAEVVPR
13E15.txt	6	IF2B2_HUMAN	Insulin-like growth factor 2 mRNA-binding protein 2	61804	65	MVIITGPPEAQFK
13E15.txt	6	IF2B2_HUMAN	Insulin-like growth factor 2 mRNA-binding protein 2	61804	65	DQTPDENEEVIVR
13E15.txt	7	IF2B3_HUMAN	Insulin-like growth factor 2 mRNA-binding protein 3	63681	49	IKEENFVSPK
13E15.txt	7	IF2B3_HUMAN	Insulin-like growth factor 2 mRNA-binding protein 3	63681	49	MVIITGPPEAQFK
13E16.txt	4	LTV1_HUMAN	Protein LTV1 homolog	54821	127	DPLAADESAPQR
13E16.txt	4	LTV1_HUMAN	Protein LTV1 homolog	54821	127	LQEVLDNYK
13E16.txt	4	LTV1_HUMAN	Protein LTV1 homolog	54821	127	SQRDPLAADESAPQR
13E18.txt	15	SSF1_HUMAN	Suppressor of SWI4 1 homolog	53161	53	VGGSDDEEASGIPSR
13E18.txt	12	CBX8_HUMAN	Chromobox protein homolog 8	43369	57	RQDSDLVQCQGVTPSSAEATGK
13E19.txt	3	CBX8_HUMAN	Chromobox protein homolog 8	43369	306	AAFLEAK
13E19.txt	5	DBPA_HUMAN	DNA-binding protein A	40066	213	NDTKEDVFVHQTAIK
13E19.txt	5	DBPA_HUMAN	DNA-binding protein A	40066	213	NDTKEDVFVHQTAIK
13E19.txt	5	DBPA_HUMAN	DNA-binding protein A	40066	213	GAEANVTGPDGVPVEGSR
13E19.txt	5	DBPA_HUMAN	DNA-binding protein A	40066	213	SVGDGETVEFDVVEGK
13E19.txt	3	CBX8_HUMAN	Chromobox protein homolog 8	43369	306	ELPDPSQRPLGEPAGLGEYLK
13E19.txt	3	CBX8_HUMAN	Chromobox protein homolog 8	43369	306	ESNTDQGF
13E19.txt	3	CBX8_HUMAN	Chromobox protein homolog 8	43369	306	KELPDPSQRPLGEPAGLGEYLK
13E19.txt	3	CBX8_HUMAN	Chromobox protein homolog 8	43369	306	LLAAFEER
13E19.txt	3	CBX8_HUMAN	Chromobox protein homolog 8	43369	306	MEYLVK
13E19.txt	3	CBX8_HUMAN	Chromobox protein homolog 8	43369	306	NMGLSPPASSTSTSSCTR
13E19.txt	3	CBX8_HUMAN	Chromobox protein homolog 8	43369	306	NMGLSPPASSTSTSSCTR
13E19.txt	3	CBX8_HUMAN	Chromobox protein homolog 8	43369	306	RQDSDLVQCQGVTPSSAEATGK
13E19.txt	3	CBX8_HUMAN	Chromobox protein homolog 8	43369	306	RQDSDLVQCQGVTPSSAEATGK
13E19.txt	3	CBX8_HUMAN	Chromobox protein homolog 8	43369	306	YSTWEPEENILDAR
13E20.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	803	ELPDPSQRPLGEPAGLGEYLK
13E20.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	803	EMELYGPK
13E20.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	803	EREMELYGPK
13E20.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	803	FPAGHSVIQLAR
13E20.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	803	KELPDPSQRPLGEPAGLGEYLK
13E20.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	803	KLDDTPSGAGK
13E20.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	803	KLDDTPSGAGKFPAGHSVIQLAR
13E20.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	803	LAVDTFPAR
13E20.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	803	LDDTPSGAGKFPAGHSVIQLAR
13E20.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	803	LLAAFEER
13E20.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	803	LLAAFEERER
13E20.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	803	MEYLVK
13E20.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	803	MEYLVK
13E20.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	803	NMGLSPPASSTSTSSCTR
13E20.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	803	RQDSDLVQCQGVTPSSAEATGK
13E20.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	803	SPQDLASTSR
13E20.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	803	VFAAEALLK
13E20.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	803	VFAAEALLK
13E20.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	803	VFAAEALLKR
13E20.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	803	VVVTDVTSNFLTVTIK
13E20.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	803	VVVTDVTSNFLTVTIK
13E20.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	803	VVVTDVTSNFLTVTIK
13E20.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	803	YSTWEPEENILDAR
13E20.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	803	YSTWEPEENILDAR
13E20.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	803	YSTWEPEENILDAR
13E21.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	610	ELPDPSQRPLGEPAGLGEYLK
13E21.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	610	EMELYGPK
13E21.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	610	ESNTDQGF
13E21.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	610	FPAGHSVIQLAR
13E21.txt	11	ESCO2_HUMAN	N-acetyltransferase ESCO2	68264	58	MAALTPR
13E21.txt	11	ESCO2_HUMAN	N-acetyltransferase ESCO2	68264	58	MAALTPRKR
13E21.txt	10	HLF_HUMAN	Hepatic leukemia factor	33199	59	NNMAAKRSR
13E21.txt	10	HLF_HUMAN	Hepatic leukemia factor	33199	59	NNMAAKRSR
13E21.txt	10	HLF_HUMAN	Hepatic leukemia factor	33199	59	KNNMAAKRSR
13E21.txt	6	KRR1_HUMAN	KRR1 small subunit processome component	43638	91	SVSFEQAVR

13E21.txt	6	KRR1_HUMAN	KRR1 small subunit processome component	43638	91	ILQDDVACDIK
13E21.txt	3	YBOX1_HUMAN	Nuclease-sensitive element-binding protein 1	35903	163	GAEANVTGPGGVPVQGSK
13E21.txt	3	YBOX1_HUMAN	Nuclease-sensitive element-binding protein 1	35903	163	SVGDGETVEFDVVEGEK
13E21.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	610	GPKPK
13E21.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	610	KELPDPSQRPLGEPASAGLGEYLK
13E21.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	610	KLDDTPSGAGKFPAGHSVIQLAR
13E21.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	610	LAVDTFFPAR
13E21.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	610	LLAAFEER
13E21.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	610	MEYLVK
13E21.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	610	NMGLSPPASSTSTSTSCR
13E21.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	610	NMGLSPPASSTSTSTSCR
13E21.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	610	RQDSDLVQCQGVTSPPSAEATGK
13E21.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	610	RQDSDLVQCQGVTSPPSAEATGK
13E21.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	610	SPQDLASTSR
13E21.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	610	VFAAEALLK
13E21.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	610	YSTWEPEENILDAR
13E22.txt	3	DBPA_HUMAN	DNA-binding protein A	40066	157	WFNVR
13E22.txt	3	DBPA_HUMAN	DNA-binding protein A	40066	157	NGYGFNIR
13E22.txt	3	DBPA_HUMAN	DNA-binding protein A	40066	157	NDTKEDVVFHQTAIK
13E22.txt	3	DBPA_HUMAN	DNA-binding protein A	40066	157	GAEANVTGPDGVPVEGSR
13E22.txt	8	ILF2_HUMAN	Interleukin enhancer-binding factor 2	43035	65	VLQSAALAIR
13E22.txt	2	ILF2_HUMAN	Interleukin enhancer-binding factor 2	43035	65	QPLALNVAIR
13E22.txt	2	YBOX1_HUMAN	Nuclease-sensitive element-binding protein 1	35903	204	WFNVR
13E22.txt	2	YBOX1_HUMAN	Nuclease-sensitive element-binding protein 1	35903	204	NGYGFNIR
13E22.txt	2	YBOX1_HUMAN	Nuclease-sensitive element-binding protein 1	35903	204	GAEANVTGPGGVPVQGSK
13E22.txt	2	YBOX1_HUMAN	Nuclease-sensitive element-binding protein 1	35903	204	NDTKEDVVFHQTAIK
13E22.txt	2	YBOX1_HUMAN	Nuclease-sensitive element-binding protein 1	35903	204	EDGNEEDKENQGDDETQGGQPPQR
13E22.txt	5	ZCHC3_HUMAN	Zinc finger CCHC domain-containing protein 3	43591	86	NETVDVDTWTK
13E22.txt	5	ZCHC3_HUMAN	Zinc finger CCHC domain-containing protein 3	43591	86	CGSRTHMSGCTQDR
13E22.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	246	KELPDPSQRPLGEPASAGLGEYLK
13E22.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	246	NMGLSPPASSTSTSTSCR
13E22.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	246	NMGLSPPASSTSTSTSCR
13E22.txt	10	PCGF6_HUMAN	Polycomb group RING finger protein 6	39023	42	GLEVPKPAVPPVPSSK
13E22.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	246	RQDSDLVQCQGVTSPPSAEATGK
13E22.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	246	YSTWEPEENILDAR
13E22.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	246	YSTWEPEENILDAR
13E23.txt	2	CBX8_HUMAN	Chromobox protein homolog 8	43369	158	GIRIPYPGRSPQDLASTSRAR
13E23.txt	1	CSK21_HUMAN	Casein kinase II subunit alpha	45115	236	FYMYEILK
13E23.txt	1	CSK21_HUMAN	Casein kinase II subunit alpha	45115	236	QLYQTLTDYDIR
13E23.txt	1	CSK21_HUMAN	Casein kinase II subunit alpha	45115	236	ALDYCHSMGIMHR
13E23.txt	1	CSK21_HUMAN	Casein kinase II subunit alpha	45115	236	TPALVFEHVNNTDFK
13E23.txt	1	CSK21_HUMAN	Casein kinase II subunit alpha	45115	236	TPALVFEHVNNTDFK
13E23.txt	1	CSK21_HUMAN	Casein kinase II subunit alpha	45115	236	GGPNIITLADIVKDPVSR
13E23.txt	6	DNL3_HUMAN	DNA ligase 3	102625	41	CTANDLK
13E23.txt	6	DNL3_HUMAN	DNA ligase 3	102625	41	DGNMQTAKPSAMKVEGEK
13E23.txt	6	DNL3_HUMAN	DNA ligase 3	102625	41	RVTKALDLADMITRVIQEGLEGLVLK
13E23.txt	4	PHC2_HUMAN	Polyhomeotic-like protein 2	90657	53	YAQGFLEPK
13E23.txt	2	CBX8_HUMAN	Chromobox protein homolog 8	43369	158	KELPDPSQRPLGEPASAGLGEYLK
13E23.txt	2	CBX8_HUMAN	Chromobox protein homolog 8	43369	158	LLAAFEER
13E23.txt	2	CBX8_HUMAN	Chromobox protein homolog 8	43369	158	RQDSDLVQCQGVTSPPSAEATGK
13E23.txt	2	CBX8_HUMAN	Chromobox protein homolog 8	43369	158	YSTWEPEENILDAR
13E24.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	514	EEMELYGPK
13E24.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	514	EREMELYGPK
13E24.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	514	FPAGHSVIQLAR
13E24.txt	3	BMI1_HUMAN	Polycomb complex protein BMI-1	36925	122	DGLTKALDES DSGSDK
13E24.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	514	KELPDPSQRPLGEPASAGLGEYLK
13E24.txt	2	RING2_HUMAN	E3 ubiquitin-protein ligase RING2	37632	395	FCADCIITLR
13E24.txt	2	RING2_HUMAN	E3 ubiquitin-protein ligase RING2	37632	395	FCADCIITLR
13E24.txt	2	RING2_HUMAN	E3 ubiquitin-protein ligase RING2	37632	395	GESNQMNLDTASEK
13E24.txt	2	RING2_HUMAN	E3 ubiquitin-protein ligase RING2	37632	395	HNNQQALSHSIEGLK
13E24.txt	2	RING2_HUMAN	E3 ubiquitin-protein ligase RING2	37632	395	IYPSRDEYEAHQER
13E24.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	514	KLDDTPSGAGK
13E24.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	514	LLAAFEER
13E24.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	514	MEYLVK
13E24.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	514	NMGLSPPASSTSTSTSCR
13E24.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	514	NMGLSPPASSTSTSTSCR
13E24.txt	2	RING2_HUMAN	E3 ubiquitin-protein ligase RING2	37632	395	SLRPPDNFDALISK
13E24.txt	2	RING2_HUMAN	E3 ubiquitin-protein ligase RING2	37632	395	SLRPPDNFDALISK
13E24.txt	3	BMI1_HUMAN	Polycomb complex protein BMI-1	36925	122	LQDQDIVYK
13E24.txt	3	BMI1_HUMAN	Polycomb complex protein BMI-1	36925	122	YCPICDVQVHK
13E24.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	514	SPQDLASTSR
13E24.txt	2	RING2_HUMAN	E3 ubiquitin-protein ligase RING2	37632	395	TPQEAITDGLIIVVSPR
13E24.txt	2	RING2_HUMAN	E3 ubiquitin-protein ligase RING2	37632	395	VNKPMELYYAPT
13E24.txt	2	RING2_HUMAN	E3 ubiquitin-protein ligase RING2	37632	395	VNKPMELYYAPT
13E24.txt	2	RING2_HUMAN	E3 ubiquitin-protein ligase RING2	37632	395	VNKPMELYYAPT
13E24.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	514	VFAAEALLK
13E24.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	514	YSTWEPEENILDAR
13E24.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	514	YSTWEPEENILDAR
13E25.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	564	EEMELYGPK
13E25.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	564	EREMELYGPK
13E25.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	564	FPAGHSVIQLAR
13E25.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	564	KELPDPSQRPLGEPASAGLGEYLK
13E25.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	564	KELPDPSQRPLGEPASAGLGEYLK
13E25.txt	5	BMI1_HUMAN	Polycomb complex protein BMI-1	36925	85	SDKTLQDIVYK
13E25.txt	4	FBRL_HUMAN	rRNA 2'-O-methyltransferase fibrillar	33763	94	VSISEGDDKIEYR
13E25.txt	4	FBRL_HUMAN	rRNA 2'-O-methyltransferase fibrillar	33763	94	DHAVVGVYRPPPK
13E25.txt	2	NPM_HUMAN	Nucleophosmin	32555	165	FINVYK
13E25.txt	2	NPM_HUMAN	Nucleophosmin	32555	165	VDNDENHQLSLR

13E25.txt	2	NPM_HUMAN	Nucleophosmin	32555	165	MTDQEAIQDLWQWR
13E25.txt	2	NPM_HUMAN	Nucleophosmin	32555	165	VNDNEHQLSLRVTSLGAGAK
13E25.txt	7	PCGF2_HUMAN	Polycomb group RING finger protein 2	37764	59	SDKTLQDIVYK
13E25.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	564	LLAAFEER
13E25.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	564	MEYLVK
13E25.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	564	NMGLSPPASSTSTSSTCR
13E25.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	564	NMGLSPPASSTSTSSTCR
13E25.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	564	RQDSDLVQCGVTSPSSAEATGK
13E25.txt	5	BMI1_HUMAN	Polycomb complex protein BMI-1	36925	85	YCPICDVQVHK
13E25.txt	7	PCGF2_HUMAN	Polycomb group RING finger protein 2	37764	59	YLETNKYCPMCDVQVHK
13E25.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	564	SPQDLASTSR
13E25.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	564	TYEFR
13E25.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	564	VFAAEALLK
13E25.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	564	YSTWEPEENILDAR
13E26.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	231	KELPDPSQRPLGEPASAGLGEYLK
13E26.txt	7	ZN717_HUMAN	Zinc finger protein 717	105183	46	THTGEKPPFRNCECRK
13E26.txt	7	ZN717_HUMAN	Zinc finger protein 717	105183	46	THTGEKPYECNECGKPFINK
13E26.txt	7	ZN717_HUMAN	Zinc finger protein 717	105183	46	YNEYEKACNNSAVIVQITVQVGQPTCCRK
13E26.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	231	LLAAFEER
13E26.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	231	NMGLSPPASSTSTSSTCR
13E26.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	231	RQDSDLVQCGVTSPSSAEATGK
13E26.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	231	TFLK
13E26.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	231	YSTWEPEENILDAR
13E27.txt	4	CBX8_HUMAN	Chromobox protein homolog 8	43369	152	NMGLSPPASSTSTSSTCR
13E27.txt	4	CBX8_HUMAN	Chromobox protein homolog 8	43369	152	RQDSDLVQCGVTSPSSAEATGK
13E27.txt	4	CBX8_HUMAN	Chromobox protein homolog 8	43369	152	YSTWEPEENILDAR
13E28.txt	8	PCGF3_HUMAN	Polycomb group RING finger protein 3	28097	58	YLEENNTCPTCR
13E28.txt	12	PCGF5_HUMAN	Polycomb group RING finger protein 5	29694	50	TCIVQHFDNSNDCPR
13E28.txt	6	THOC4_HUMAN	THO complex subunit 4	26872	65	SLGTADVHFER
13E28.txt	7	UBR1_HUMAN	E3 ubiquitin-protein ligase UBR1	200080	62	QHNGLK
13E28.txt	7	UBR1_HUMAN	E3 ubiquitin-protein ligase UBR1	200080	62	IMAQMSALQK
13E28.txt	7	UBR1_HUMAN	E3 ubiquitin-protein ligase UBR1	200080	62	GIPQLEGQKDMITWILQMFDTVK
13E28.txt	5	CBX8_HUMAN	Chromobox protein homolog 8	43369	94	NMGLSPPASSTSTSSTCR
13E28.txt	5	CBX8_HUMAN	Chromobox protein homolog 8	43369	94	YSTWEPEENILDAR
13E29.txt	9	JIP4_HUMAN	C-jun-amino-terminal kinase-interacting protein 4	146115	48	QVTNGQGGENKMK
13E29.txt	9	JIP4_HUMAN	C-jun-amino-terminal kinase-interacting protein 4	146115	48	DTSMKLWCAVGNVNLGGK
13E29.txt	7	PCGF1_HUMAN	Polycomb group RING finger protein 1	29202	54	VMQDIVYK
13E29.txt	7	PCGF1_HUMAN	Polycomb group RING finger protein 1	29202	54	LVPGLQDSEEKR
13E29.txt	7	PCGF1_HUMAN	Polycomb group RING finger protein 1	29202	54	YDEQLNLCLER
13E29.txt	11	PCGF5_HUMAN	Polycomb group RING finger protein 5	29694	46	LNTLEEIIFK
13E29.txt	3	CBX8_HUMAN	Chromobox protein homolog 8	43369	166	NMGLSPPASSTSTSSTCR
13E29.txt	3	CBX8_HUMAN	Chromobox protein homolog 8	43369	166	SPQDLASTSR
13E29.txt	3	CBX8_HUMAN	Chromobox protein homolog 8	43369	166	YSTWEPEENILDAR
13E30.txt	5	CSK2B_HUMAN	Casein kinase II subunit beta	24926	111	RPANQFVPR
13E30.txt	5	CSK2B_HUMAN	Casein kinase II subunit beta	24926	111	YQQGDFGYCPR
13E30.txt	8	CBX8_HUMAN	Chromobox protein homolog 8	43369	75	NMGLSPPASSTSTSSTCR
13E30.txt	8	CBX8_HUMAN	Chromobox protein homolog 8	43369	75	YSTWEPEENILDAR
13E31.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	325	EMELYGPK
13E31.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	325	LLAAFEER
13E31.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	325	LLAAFEER
13E31.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	325	MEYLVK
13E31.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	325	MEYLVK
13E31.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	325	MEYLVK
13E31.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	325	NMGLSPPASSTSTSSTCR
13E31.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	325	NMGLSPPASSTSTSSTCR
13E31.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	325	SPQDLASTSR
13E31.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	325	TYEFR
13E31.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	325	VFAAEALLK
13E32.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	321	EMELYGPK
13E32.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	321	LLAAFEER
13E32.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	321	MEYLVK
13E32.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	321	MEYLVK
13E32.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	321	NMGLSPPASSTSTSSTCR
13E32.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	321	SPQDLASTSR
13E32.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	321	VFAAEALLK
13E32.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	321	YSTWEPEENILDAR
13E33.txt	4	CBX8_HUMAN	Chromobox protein homolog 8	43369	158	EREMELYGPK
13E33.txt	4	CBX8_HUMAN	Chromobox protein homolog 8	43369	158	EREMELYGPK
13E33.txt	4	CBX8_HUMAN	Chromobox protein homolog 8	43369	158	GPKPK
13E33.txt	4	CBX8_HUMAN	Chromobox protein homolog 8	43369	158	LLAAFEER
13E33.txt	4	CBX8_HUMAN	Chromobox protein homolog 8	43369	158	MEYLVK
13E33.txt	4	CBX8_HUMAN	Chromobox protein homolog 8	43369	158	VFAAEALLK
13E33.txt	4	CBX8_HUMAN	Chromobox protein homolog 8	43369	158	YSTWEPEENILDAR
13E34.txt	4	CBX8_HUMAN	Chromobox protein homolog 8	43369	162	EREMELYGPK
13E34.txt	4	CBX8_HUMAN	Chromobox protein homolog 8	43369	162	LLAAFEER
13E34.txt	4	CBX8_HUMAN	Chromobox protein homolog 8	43369	162	MEYLVK
13E34.txt	4	CBX8_HUMAN	Chromobox protein homolog 8	43369	162	VFAAEALLK
13E34.txt	4	CBX8_HUMAN	Chromobox protein homolog 8	43369	162	YSTWEPEENILDAR
13E34.txt	4	CBX8_HUMAN	Chromobox protein homolog 8	43369	162	YSTWEPEENILDAR
13E35.txt	5	CBX8_HUMAN	Chromobox protein homolog 8	43369	139	DREDRER
13E35.txt	5	CBX8_HUMAN	Chromobox protein homolog 8	43369	139	LLAAFEER
13E35.txt	26	H2A1B_HUMAN	Histone H2A type 1-B	14127	42	VTAQQGGVLPNIQAVLLPK
13E35.txt	30	H2B1A_HUMAN	Histone H2B type 1-A	14159	40	LLLPGELAK
13E35.txt	30	H2B1A_HUMAN	Histone H2B type 1-A	14159	40	AMSIMNSFVTDIFER
13E35.txt	6	H2B1B_HUMAN	Histone H2B type 1-B	13942	116	LLLPGELAK
13E35.txt	6	H2B1B_HUMAN	Histone H2B type 1-B	13942	116	AMGIMNSFVNDIFER
13E35.txt	18	NBN_HUMAN	Nibrin	84906	48	LQDDSEMLPK
13E35.txt	18	NBN_HUMAN	Nibrin	84906	48	LPHIIGGSDLIAHHARK
13E35.txt	5	CBX8_HUMAN	Chromobox protein homolog 8	43369	139	MEYLVK

13E35.txt	5	CBX8_HUMAN	Chromobox protein homolog 8	43369	139	VFAAEALLK
13E35.txt	5	CBX8_HUMAN	Chromobox protein homolog 8	43369	139	YSTWEPEENILDAR
13E36.txt	6	CBX8_HUMAN	Chromobox protein homolog 8	43369	70	LLAAFEER
13E36.txt	12	MPPH1_HUMAN	M-phase phosphoprotein 1	210553	42	KLLDLIEDLK
13E36.txt	12	MPPH1_HUMAN	M-phase phosphoprotein 1	210553	42	TKGELIKTKEELK
13E36.txt	12	MPPH1_HUMAN	M-phase phosphoprotein 1	210553	42	LRETGNINTSLLTLGKCINVLK
13E36.txt	6	CBX8_HUMAN	Chromobox protein homolog 8	43369	70	VFAAEALLK
13E36.txt	6	CBX8_HUMAN	Chromobox protein homolog 8	43369	70	YSTWEPEENILDAR
13E36.txt	6	CBX8_HUMAN	Chromobox protein homolog 8	43369	70	YSTWEPEENILDAR
13E37.txt	5	H4_HUMAN	Histone H4	11360	66	TLYGFGG
13E37.txt	5	H4_HUMAN	Histone H4	11360	66	DNIQGITKPAIR
13E37.txt	9	NFIL3_HUMAN	Nuclear factor interleukin-3-regulated protein	51472	53	KNNEAAK
13E37.txt	9	NFIL3_HUMAN	Nuclear factor interleukin-3-regulated protein	51472	53	KMQTVKK
13E37.txt	13	RRP5_HUMAN	Protein RRP5 homolog	208570	46	KQNKGAK
13E37.txt	13	RRP5_HUMAN	Protein RRP5 homolog	208570	46	KLMMTLKK
13E37.txt	13	RRP5_HUMAN	Protein RRP5 homolog	208570	46	SYGMLVKVGEQMR
13E38.txt	6	H4_HUMAN	Histone H4	11360	42	GGVKR
13E38.txt	6	H4_HUMAN	Histone H4	11360	42	DNIQGITKPAIR
merge6E05.txt	2	RRP5_HUMAN	Protein RRP5 homolog	208570	61	VGQYLNClVEK
merge6E07.txt	2	YTDC2_HUMAN	YTH domain-containing protein 2	74853	98	DVNTNSENWAVVK
merge6E07.txt	2	YTDC2_HUMAN	YTH domain-containing protein 2	74853	98	APSKPWSQVDEATIR
merge6E09.txt	5	ZFR_HUMAN	Zinc finger RNA-binding protein	116939	115	AISSASSPQSPGDALR
merge6E09.txt	5	ZFR_HUMAN	Zinc finger RNA-binding protein	116939	115	NVNLVLLCSEKPSK
merge6E11.txt	1	MOV10_HUMAN	Putative helicase MOV-10	113599	364	TVTLVEAIK
merge6E11.txt	1	MOV10_HUMAN	Putative helicase MOV-10	113599	364	SVILSTVR
merge6E11.txt	1	MOV10_HUMAN	Putative helicase MOV-10	113599	364	VLITTLITAGR
merge6E11.txt	1	MOV10_HUMAN	Putative helicase MOV-10	113599	364	VGSVEEFQGGQER
merge6E11.txt	1	MOV10_HUMAN	Putative helicase MOV-10	113599	364	SLESNPEQLQAMR
merge6E11.txt	1	MOV10_HUMAN	Putative helicase MOV-10	113599	364	HGVDVEVQGPHEAR
merge6E11.txt	1	MOV10_HUMAN	Putative helicase MOV-10	113599	364	ETGDPGGQLVLGADPR
merge6E11.txt	1	MOV10_HUMAN	Putative helicase MOV-10	113599	364	AHLACAPNSNSGADLLCQR
merge6E11.txt	1	MOV10_HUMAN	Putative helicase MOV-10	113599	364	EGNSPFFNPEEAATVTSYLK
merge6E13.txt	2	KR11_HUMAN	Protein KR11 homolog	83201	100	LQETSSQSYVEEQK
merge6E13.txt	5	NOL1_HUMAN	Putative RNA methyltransferase NOL1	89247	62	LGVTNTIISHYDGR
merge6E14.txt	3	TSR1_HUMAN	Pre-rRNA-processing protein TSR1 homolog	91752	92	WTYDYPVPEVPVWLK
merge6E14.txt	3	TSR1_HUMAN	Pre-rRNA-processing protein TSR1 homolog	91752	92	TSPWDPKLENLPQDYAR
merge6E15.txt	1	ILF3_HUMAN	Interleukin enhancer-binding factor 3	95279	146	EDITQSAQHAR
merge6E15.txt	1	ILF3_HUMAN	Interleukin enhancer-binding factor 3	95279	146	VLAGETLSVNDPPDVLDR
merge6E18.txt	3	NOG1_HUMAN	Nucleolar GTP-binding protein 1	73918	71	MVKKAKTMMIK
merge6E18.txt	3	NOG1_HUMAN	Nucleolar GTP-binding protein 1	73918	71	ADVVDQPYAFTTK
merge6E20.txt	3	IF2B1_HUMAN	Insulin-like growth factor 2 mRNA-binding protein 1	63417	365	QQQVDIPLR
merge6E20.txt	3	IF2B1_HUMAN	Insulin-like growth factor 2 mRNA-binding protein 1	63417	365	ISYGGQLVK
merge6E20.txt	3	IF2B1_HUMAN	Insulin-like growth factor 2 mRNA-binding protein 1	63417	365	LLVPTQYVGAIGK
merge6E20.txt	3	IF2B1_HUMAN	Insulin-like growth factor 2 mRNA-binding protein 1	63417	365	DQTPDENQVIVK
merge6E20.txt	3	IF2B1_HUMAN	Insulin-like growth factor 2 mRNA-binding protein 1	63417	365	AISVHSTPEGCSACK
merge6E20.txt	3	IF2B1_HUMAN	Insulin-like growth factor 2 mRNA-binding protein 1	63417	365	LYIGNLINESVTPADLEK
merge6E20.txt	14	IF2B2_HUMAN	Insulin-like growth factor 2 mRNA-binding protein 2	61804	46	DQTPDENEEVIVR
merge6E23.txt	6	SSF1_HUMAN	Suppressor of SWI4 1 homolog	53161	66	LPGGPTLTFQVK
merge6E23.txt	3	STAU1_HUMAN	Double-stranded RNA-binding protein Staufen	63143	113	NLPVNFVAR
merge6E23.txt	3	STAU1_HUMAN	Double-stranded RNA-binding protein Staufen	63143	113	VSVGEFVGELEGK
merge6E23.txt	9	CBX8_HUMAN	Chromobox protein homolog 8	43369	43	YSTWEPEENILDAR
merge6E24.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	237	ESNTDQGFFK
merge6E24.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	237	KLDDTPSGAGKFPAGHSVIQLAR
merge6E24.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	237	LLAAFEER
merge6E24.txt	2	DBPA_HUMAN	DNA-binding protein A	40066	218	NDTKEDVFVHQTAIK
merge6E24.txt	2	DBPA_HUMAN	DNA-binding protein A	40066	218	SVGDGETVEFDDVVEGK
merge6E24.txt	2	DBPA_HUMAN	DNA-binding protein A	40066	218	NYAGEEEEEEGSGSEGFDPATDR
merge6E24.txt	2	DBPA_HUMAN	DNA-binding protein A	40066	218	IQAGEIGEMKDGVPGEAQLQGPVHR
merge6E24.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	237	NMGLSPPASSTSTSTSCR
merge6E24.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	237	NMGLSPPASSTSTSTSCR
merge6E24.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	237	RQDSDLVQCQGVTPSSSAEATGK
merge6E24.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	237	RQDSDLVQCQGVTPSSSAEATGK
merge6E24.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	237	YSTWEPEENILDAR
merge6E25.txt	2	CBX8_HUMAN	Chromobox protein homolog 8	43369	448	DMGAQGGRPSTLIAR
merge6E25.txt	2	CBX8_HUMAN	Chromobox protein homolog 8	43369	448	DMGAQGGRPSTLIAR
merge6E25.txt	2	CBX8_HUMAN	Chromobox protein homolog 8	43369	448	FPAGHSVIQLAR
merge6E25.txt	2	CBX8_HUMAN	Chromobox protein homolog 8	43369	448	KELPDPSQRPLGEPASAGLGEYLK
merge6E25.txt	2	CBX8_HUMAN	Chromobox protein homolog 8	43369	448	KELPDPSQRPLGEPASAGLGEYLK
merge6E25.txt	2	CBX8_HUMAN	Chromobox protein homolog 8	43369	448	LLAAFEER
merge6E25.txt	2	CBX8_HUMAN	Chromobox protein homolog 8	43369	448	LLAAFEERER
merge6E25.txt	2	CBX8_HUMAN	Chromobox protein homolog 8	43369	448	LLAAFEEREREMELYGPK
merge6E25.txt	4	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	124	FCSDCIVTLR
merge6E25.txt	4	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	124	GGGAGGSSVGTGGGGTGGVGGGAGSEDS
merge6E25.txt	4	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	124	GDR
merge6E25.txt	4	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	124	SLRPDPNFDALISK
merge6E25.txt	2	CBX8_HUMAN	Chromobox protein homolog 8	43369	448	MEYLVK
merge6E25.txt	2	CBX8_HUMAN	Chromobox protein homolog 8	43369	448	NMGLSPPASSTSTSTSCR
merge6E25.txt	2	CBX8_HUMAN	Chromobox protein homolog 8	43369	448	NMGLSPPASSTSTSTSCR
merge6E25.txt	2	CBX8_HUMAN	Chromobox protein homolog 8	43369	448	RQDSDLVQCQGVTPSSSAEATGK
merge6E25.txt	2	CBX8_HUMAN	Chromobox protein homolog 8	43369	448	VVVTDVTSNFLTVTIK
merge6E25.txt	2	CBX8_HUMAN	Chromobox protein homolog 8	43369	448	YSTWEPEENILDAR
merge6E26.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	866	DMGAQGGRPSTLIAR
merge6E26.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	866	EGLRNMGSLSPASSTSTSTSCR
merge6E26.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	866	EGLRNMGSLSPASSTSTSTSCR
merge6E26.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	866	ELPDPSQRPLGEPASAGLGEYLK
merge6E26.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	866	ESNTDQGFFK
merge6E26.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	866	ESNTDQGFFK
merge6E26.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	866	FPAGHSVIQLAR

merge6E26.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	866	ILGDPEEESWSPSLTNLEK
merge6E26.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	866	KELPDPSQRPLGEPGAGLGEYLK
merge6E26.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	866	KELPDPSQRPLGEPGAGLGEYLK
merge6E26.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	866	KLDDTPSGAGKFPAGHSVIQLAR
merge6E26.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	866	KLDDTPSGAGKFPAGHSVIQLAR
merge6E26.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	866	KLDDTPSGAGKFPAGHSVIQLAR
merge6E26.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	866	LAVDTFFPAR
merge6E26.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	866	LDDTPSGAGKFPAGHSVIQLAR
merge6E26.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	866	LLAAFEER
merge6E26.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	866	LLAAFEERER
merge6E26.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	866	LLAAFEEREREMELYGPK
merge6E26.txt	4	HLF_HUMAN	Hepatic leukemia factor	33199	40	NNMAAKRSR
merge6E26.txt	4	HLF_HUMAN	Hepatic leukemia factor	33199	40	NNMAAKRSR
merge6E26.txt	4	HLF_HUMAN	Hepatic leukemia factor	33199	40	KNNMAAKRSR
merge6E26.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	866	NMGLSPPASSTSTSTSCR
merge6E26.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	866	NMGLSPPASSTSTSTSCR
merge6E26.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	866	NMGLSPPASSTSTSTSCR
merge6E26.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	866	RQDSDLVQCQGVTPSSAEATGK
merge6E26.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	866	RQDSDLVQCQGVTPSSAEATGK
merge6E26.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	866	VFAAEALLKR
merge6E26.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	866	VVVTDVTSNFLTVTIK
merge6E26.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	866	WKGWSQKYSTWEPEENILDAR
merge6E26.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	866	YSTWEPEENILDAR
merge6E26.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	866	YSTWEPEENILDAR
merge6E27.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	613	DMGAQGGRPGLIAR
merge6E27.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	613	DMGAQGGRPGLIAR
merge6E27.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	613	ELPDPSQRPLGEPGAGLGEYLK
merge6E27.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	613	EMELYGPK
merge6E27.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	613	FPAGHSVIQLAR
merge6E27.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	613	KELPDPSQRPLGEPGAGLGEYLK
merge6E27.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	613	KELPDPSQRPLGEPGAGLGEYLK
merge6E27.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	613	KLDDTPSGAGKFPAGHSVIQLAR
merge6E27.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	613	KLDDTPSGAGKFPAGHSVIQLAR
merge6E27.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	613	KLDDTPSGAGKFPAGHSVIQLAR
merge6E27.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	613	LLAAFEER
merge6E27.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	613	LLAAFEERER
merge6E27.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	613	LLAAFEEREREMELYGPK
merge6E27.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	613	LLAAFEEREREMELYGPK
merge6E27.txt	9	NSBP1_HUMAN	Nucleosome-binding protein 1	31525	47	LSAMLVPTPEVKPK
merge6E27.txt	9	NSBP1_HUMAN	Nucleosome-binding protein 1	31525	47	LSAMLVPTPEVKPKRTSSSRK
merge6E27.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	613	RQDSDLVQCQGVTPSSAEATGK
merge6E27.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	613	RQDSDLVQCQGVTPSSAEATGK
merge6E27.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	613	RQDSDLVQCQGVTPSSAEATGK
merge6E27.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	613	VFAAEALLKR
merge6E27.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	613	VFAAEALLKR
merge6E27.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	613	YSTWEPEENILDAR
merge6E27.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	613	YSTWEPEENILDAR
merge6E28.txt	2	CBX8_HUMAN	Chromobox protein homolog 8	43369	296	ELPDPSQRPLGEPGAGLGEYLK
merge6E28.txt	4	PCGF6_HUMAN	Polycomb group RING finger protein 6	39023	136	CNIVVHQPVTPEVKPK
merge6E28.txt	2	CBX8_HUMAN	Chromobox protein homolog 8	43369	296	KELPDPSQRPLGEPGAGLGEYLK
merge6E28.txt	2	CBX8_HUMAN	Chromobox protein homolog 8	43369	296	KELPDPSQRPLGEPGAGLGEYLK
merge6E28.txt	2	CBX8_HUMAN	Chromobox protein homolog 8	43369	296	KLDDTPSGAGKFPAGHSVIQLAR
merge6E28.txt	2	CBX8_HUMAN	Chromobox protein homolog 8	43369	296	LLAAFEERER
merge6E28.txt	2	CBX8_HUMAN	Chromobox protein homolog 8	43369	296	NMGLSPPASSTSTSTSCR
merge6E28.txt	2	CBX8_HUMAN	Chromobox protein homolog 8	43369	296	NMGLSPPASSTSTSTSCR
merge6E28.txt	4	PCGF6_HUMAN	Polycomb group RING finger protein 6	39023	136	GLEVPKPAVQPVPSSK
merge6E28.txt	4	PCGF6_HUMAN	Polycomb group RING finger protein 6	39023	136	QLQDIVYK
merge6E28.txt	1	YBOX1_HUMAN	Nuclease-sensitive element-binding protein 1	35903	339	NGYGFINR
merge6E28.txt	1	YBOX1_HUMAN	Nuclease-sensitive element-binding protein 1	35903	339	GAEAAVNTGPGGVVPGQSK
merge6E28.txt	1	YBOX1_HUMAN	Nuclease-sensitive element-binding protein 1	35903	339	NDTKEDVFVHQTAIK
merge6E28.txt	1	YBOX1_HUMAN	Nuclease-sensitive element-binding protein 1	35903	339	NDTKEDVFVHQTAIK
merge6E28.txt	1	YBOX1_HUMAN	Nuclease-sensitive element-binding protein 1	35903	339	NDTKEDVFVHQTAIK
merge6E28.txt	1	YBOX1_HUMAN	Nuclease-sensitive element-binding protein 1	35903	339	SVGDGETVEFDVVEGK
merge6E28.txt	1	YBOX1_HUMAN	Nuclease-sensitive element-binding protein 1	35903	339	NDTKEDVFVHQTAIK
merge6E28.txt	1	YBOX1_HUMAN	Nuclease-sensitive element-binding protein 1	35903	339	AADPPAENSSAPEAQGGAE
merge6E28.txt	1	YBOX1_HUMAN	Nuclease-sensitive element-binding protein 1	35903	339	RPQYSNPPVQGEVMEGADNQGAGEQGRP
merge6E28.txt	1	YBOX1_HUMAN	Nuclease-sensitive element-binding protein 1	35903	339	RPQYSNPPVQGEVMEGADNQGAGEQGRP
merge6E28.txt	2	CBX8_HUMAN	Chromobox protein homolog 8	43369	296	YSTWEPEENILDAR
merge6E28.txt	2	CBX8_HUMAN	Chromobox protein homolog 8	43369	296	YSTWEPEENILDAR
merge6E29.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	599	ESNTDQGGFFK
merge6E29.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	599	FPAGHSVIQLAR
merge6E29.txt	8	RING2_HUMAN	E3 ubiquitin-protein ligase RING2	37632	85	FCADCIITLR
merge6E29.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	599	KELPDPSQRPLGEPGAGLGEYLK
merge6E29.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	599	KLDDTPSGAGKFPAGHSVIQLAR
merge6E29.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	599	LDDTPSGAGKFPAGHSVIQLAR
merge6E29.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	599	LLAAFEER
merge6E29.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	599	LLAAFEERER
merge6E29.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	599	LLAAFEERER
merge6E29.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	599	LLAAFEEREREMELYGPK
merge6E29.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	599	NMGLSPPASSTSTSTSCR
merge6E29.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	599	NMGLSPPASSTSTSTSCR
merge6E29.txt	11	CSK21_HUMAN	Casein kinase II subunit alpha	45115	63	QLYQTLTDYDIR
merge6E29.txt	2	ILF2_HUMAN	Interleukin enhancer-binding factor 2	43035	381	VLQSALAIIR
merge6E29.txt	2	ILF2_HUMAN	Interleukin enhancer-binding factor 2	43035	381	QPLALNAVYR
merge6E29.txt	2	ILF2_HUMAN	Interleukin enhancer-binding factor 2	43035	381	LDPELHLDIK
merge6E29.txt	2	ILF2_HUMAN	Interleukin enhancer-binding factor 2	43035	381	ILITVPPNLR
merge6E29.txt	2	ILF2_HUMAN	Interleukin enhancer-binding factor 2	43035	381	KLDPHLDIK
merge6E29.txt	2	ILF2_HUMAN	Interleukin enhancer-binding factor 2	43035	381	ILPTLEVAALGNK

merge6E29.txt	2	ILF2_HUMAN	Interleukin enhancer-binding factor 2	43035	381	WFEENASQSTVK
merge6E29.txt	3	PHC2_HUMAN	Polyhomeotic-like protein 2	90657	192	YAQGFLEPK
merge6E29.txt	3	PHC2_HUMAN	Polyhomeotic-like protein 2	90657	192	AQEIDGQALLLKK
merge6E29.txt	3	PHC2_HUMAN	Polyhomeotic-like protein 2	90657	192	SLPGQCEIAEFFR
merge6E29.txt	4	YBOX1_HUMAN	Nuclease-sensitive element-binding protein 1	35903	137	GAEANVTGPGGVVQGSK
merge6E29.txt	4	YBOX1_HUMAN	Nuclease-sensitive element-binding protein 1	35903	137	SVGDGETVFEFVVEGEK
merge6E29.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	599	RQDSDLVQCQGVTPSSSAEATGK
merge6E29.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	599	RQDSDLVQCQGVTPSSSAEATGK
merge6E29.txt	8	RING2_HUMAN	E3 ubiquitin-protein ligase RING2	37632	85	VNKPMELYYAPTK
merge6E29.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	599	VFAAEALLK
merge6E29.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	599	VFAAEALLKR
merge6E29.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	599	YSTWEPEENILDAR
merge6E29.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	599	YSTWEPEENILDAR
merge6E30.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	723	EGLRNMGLSPPASSTSTSTSCR
merge6E30.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	723	ELPDPSQRPLGEPASAGLGEYLK
merge6E30.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	723	ELPDPSQRPLGEPASAGLGEYLK
merge6E30.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	723	EMELYGPK
merge6E30.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	723	FPAGHSVIQLAR
merge6E30.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	723	FPAGHSVIQLAR
merge6E30.txt	3	BM11_HUMAN	Polycomb complex protein BMI-1	36925	154	DFYAAHPSADAANGSNEDRGEVADEDKR
merge6E30.txt	3	BM11_HUMAN	Polycomb complex protein BMI-1	36925	154	LVPGLFK
merge6E30.txt	2	RING2_HUMAN	E3 ubiquitin-protein ligase RING2	37632	346	FCADCIITLR
merge6E30.txt	2	RING2_HUMAN	E3 ubiquitin-protein ligase RING2	37632	346	FCADCIITLR
merge6E30.txt	2	RING2_HUMAN	E3 ubiquitin-protein ligase RING2	37632	346	IYPSRDEYEAHQER
merge6E30.txt	2	RING2_HUMAN	E3 ubiquitin-protein ligase RING2	37632	346	IYPSRDEYEAHQER
merge6E30.txt	2	RING2_HUMAN	E3 ubiquitin-protein ligase RING2	37632	346	LAALEELR
merge6E30.txt	3	BM11_HUMAN	Polycomb complex protein BMI-1	36925	154	SDKTLQDIVYK
merge6E30.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	723	KELPDPSQRPLGEPASAGLGEYLK
merge6E30.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	723	KELPDPSQRPLGEPASAGLGEYLK
merge6E30.txt	5	PCGF2_HUMAN	Polycomb group RING finger protein 2	37764	110	LVPGLFK
merge6E30.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	723	KLDDTPSGAGKFPAGHSVIQLAR
merge6E30.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	723	LAVDTFFPAR
merge6E30.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	723	LDDTPSGAGKFPAGHSVIQLAR
merge6E30.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	723	LAAAFEER
merge6E30.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	723	LAAAFEERER
merge6E30.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	723	LAAAFEEREREMELYGPK
merge6E30.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	723	NMGLSPPASSTSTSTSCR
merge6E30.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	723	NMGLSPPASSTSTSTSCR
merge6E30.txt	3	BM11_HUMAN	Polycomb complex protein BMI-1	36925	154	TLQDIVYK
merge6E30.txt	3	BM11_HUMAN	Polycomb complex protein BMI-1	36925	154	YCPCIDVQVHK
merge6E30.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	723	RQDSDLVQCQGVTPSSSAEATGK
merge6E30.txt	5	PCGF2_HUMAN	Polycomb group RING finger protein 2	37764	110	SDKTLQDIVYK
merge6E30.txt	5	PCGF2_HUMAN	Polycomb group RING finger protein 2	37764	110	TLQDIVYK
merge6E30.txt	5	PCGF2_HUMAN	Polycomb group RING finger protein 2	37764	110	YLETNKYCPMCDVQVHK
merge6E30.txt	2	RING2_HUMAN	E3 ubiquitin-protein ligase RING2	37632	346	SLRDPNFDALISK
merge6E30.txt	2	RING2_HUMAN	E3 ubiquitin-protein ligase RING2	37632	346	SLRDPNFDALISK
merge6E30.txt	2	RING2_HUMAN	E3 ubiquitin-protein ligase RING2	37632	346	TPQEAITDGLIIVVSPR
merge6E30.txt	2	RING2_HUMAN	E3 ubiquitin-protein ligase RING2	37632	346	VNKPMELYYAPTK
merge6E30.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	723	VFAAEALLK
merge6E30.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	723	VFAAEALLKR
merge6E30.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	723	YSTWEPEENILDAR
merge6E31.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	747	EGLRNMGLSPPASSTSTSTSCR
merge6E31.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	747	ELPDPSQRPLGEPASAGLGEYLK
merge6E31.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	747	ELPDPSQRPLGEPASAGLGEYLK
merge6E31.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	747	EMELYGPK
merge6E31.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	747	EREMELYGPK
merge6E31.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	747	FPAGHSVIQLAR
merge6E31.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	747	KELPDPSQRPLGEPASAGLGEYLK
merge6E31.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	747	KELPDPSQRPLGEPASAGLGEYLK
merge6E31.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	747	KLDDTPSGAGKFPAGHSVIQLAR
merge6E31.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	747	LAAAFEER
merge6E31.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	747	LAAAFEERER
merge6E31.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	747	LAAAFEEREREMELYGPK
merge6E31.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	747	LAAAFEEREREMELYGPK
merge6E31.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	747	MEYLVK
merge6E31.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	747	NMGLSPPASSTSTSTSCR
merge6E31.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	747	NMGLSPPASSTSTSTSCR
merge6E31.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	747	NMGLSPPASSTSTSTSCR
merge6E31.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	747	RQDSDLVQCQGVTPSSSAEATGK
merge6E31.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	747	RQDSDLVQCQGVTPSSSAEATGK
merge6E31.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	747	TFLK
merge6E31.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	747	VFAAEALLK
merge6E31.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	747	VFAAEALLKR
merge6E31.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	747	YSTWEPEENILDAR
merge6E32.txt	2	CBX8_HUMAN	Chromobox protein homolog 8	43369	312	KELPDPSQRPLGEPASAGLGEYLK
merge6E32.txt	2	CBX8_HUMAN	Chromobox protein homolog 8	43369	312	KELPDPSQRPLGEPASAGLGEYLK
merge6E32.txt	2	CBX8_HUMAN	Chromobox protein homolog 8	43369	312	LAAAFEER
merge6E32.txt	2	CBX8_HUMAN	Chromobox protein homolog 8	43369	312	LAAAFEERER
merge6E32.txt	2	CBX8_HUMAN	Chromobox protein homolog 8	43369	312	LAAAFEEREREMELYGPK
merge6E32.txt	2	CBX8_HUMAN	Chromobox protein homolog 8	43369	312	NMGLSPPASSTSTSTSCR
merge6E32.txt	2	CBX8_HUMAN	Chromobox protein homolog 8	43369	312	NMGLSPPASSTSTSTSCR
merge6E32.txt	2	CBX8_HUMAN	Chromobox protein homolog 8	43369	312	NMGLSPPASSTSTSTSCR
merge6E32.txt	2	CBX8_HUMAN	Chromobox protein homolog 8	43369	312	RQDSDLVQCQGVTPSSSAEATGK
merge6E32.txt	2	CBX8_HUMAN	Chromobox protein homolog 8	43369	312	RQDSDLVQCQGVTPSSSAEATGK
merge6E32.txt	7	MK671_HUMAN	MK167 FHA domain-interacting nucleolar phosphoprotein	34201	67	IVAETMNNYLFGER
merge6E32.txt	2	CBX8_HUMAN	Chromobox protein homolog 8	43369	312	YSTWEPEENILDAR

merge6E33.txt	1	DIMT1_HUMAN	Probable dimethyladenosine transferase	35214	253	LQVLVGDVLK
merge6E33.txt	1	DIMT1_HUMAN	Probable dimethyladenosine transferase	35214	253	SSAVQQLLEK
merge6E33.txt	1	DIMT1_HUMAN	Probable dimethyladenosine transferase	35214	253	LSINTQLLAR
merge6E33.txt	1	DIMT1_HUMAN	Probable dimethyladenosine transferase	35214	253	NPPPPINFQEWDLVLR
merge6E33.txt	3	GBLP_HUMAN	Guanine nucleotide-binding protein subunit beta-2-	35055	110	VWQVTIGTR
merge6E33.txt	3	GBLP_HUMAN	Guanine nucleotide-binding protein subunit beta-2-	35055	110	FSPNSSNPIIVSCGWDK
merge6E33.txt	9	CBX8_HUMAN	Chromobox protein homolog 8	43369	44	YSTWEPEENILDAR
merge6E34.txt	4	CBX8_HUMAN	Chromobox protein homolog 8	43369	210	NMGLSPPASSTSTSTSCR
merge6E34.txt	4	CBX8_HUMAN	Chromobox protein homolog 8	43369	210	RQDSDLVQCQGVTSPPSAEATGK
merge6E34.txt	4	CBX8_HUMAN	Chromobox protein homolog 8	43369	210	RQDSDLVQCQGVTSPPSAEATGK
merge6E34.txt	4	CBX8_HUMAN	Chromobox protein homolog 8	43369	210	YSTWEPEENILDAR
merge6E35.txt	5	CBX8_HUMAN	Chromobox protein homolog 8	43369	147	NMGLSPPASSTSTSTSCR
merge6E35.txt	5	CBX8_HUMAN	Chromobox protein homolog 8	43369	147	NMGLSPPASSTSTSTSCR
merge6E35.txt	7	THOC4_HUMAN	THO complex subunit 4	26872	98	SLGTADVHFER
merge6E35.txt	7	THOC4_HUMAN	THO complex subunit 4	26872	98	QLPDKWQHDLFDLSFGGGAGVETGGK
merge6E35.txt	5	CBX8_HUMAN	Chromobox protein homolog 8	43369	147	YSTWEPEENILDAR
merge6E36.txt	4	CBX8_HUMAN	Chromobox protein homolog 8	43369	218	EGLRNMGLSPPASSTSTSTSCR
merge6E36.txt	4	CBX8_HUMAN	Chromobox protein homolog 8	43369	218	LLAAFEEREREMELYGPK
merge6E36.txt	4	CBX8_HUMAN	Chromobox protein homolog 8	43369	218	NMGLSPPASSTSTSTSCR
merge6E36.txt	4	CBX8_HUMAN	Chromobox protein homolog 8	43369	218	NMGLSPPASSTSTSTSCR
merge6E36.txt	4	CBX8_HUMAN	Chromobox protein homolog 8	43369	218	RQDSDLVQCQGVTSPPSAEATGK
merge6E36.txt	4	CBX8_HUMAN	Chromobox protein homolog 8	43369	218	YSTWEPEENILDAR
merge6E37.txt	3	CBX8_HUMAN	Chromobox protein homolog 8	43369	283	EREMELYGPK
merge6E37.txt	3	CBX8_HUMAN	Chromobox protein homolog 8	43369	283	KELPDPSQRPLGEPASGLGEYLK
merge6E37.txt	3	CBX8_HUMAN	Chromobox protein homolog 8	43369	283	NMGLSPPASSTSTSTSCR
merge6E37.txt	3	CBX8_HUMAN	Chromobox protein homolog 8	43369	283	RQDSDLVQCQGVTSPPSAEATGK
merge6E37.txt	3	CBX8_HUMAN	Chromobox protein homolog 8	43369	283	VFAAEALLKR
merge6E37.txt	8	CSK2B_HUMAN	Casein kinase II subunit beta	24926	76	YQQGDFGYCPR
merge6E37.txt	8	CSK2B_HUMAN	Casein kinase II subunit beta	24926	76	VYCENQPMPLIGLSDIPGEAMVK
merge6E37.txt	3	CBX8_HUMAN	Chromobox protein homolog 8	43369	283	YSTWEPEENILDAR
merge6E38.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	343	EGLRNMGLSPPASSTSTSTSCR
merge6E38.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	343	LLAAFEER
merge6E38.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	343	LLAAFEERER
merge6E38.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	343	LLAAFEEREREMELYGPK
merge6E38.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	343	LLAAFEEREREMELYGPK
merge6E38.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	343	LLAAFEEREREMELYGPK
merge6E38.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	343	LLAAFEEREREMELYGPK
merge6E38.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	343	NMGLSPPASSTSTSTSCR
merge6E38.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	343	NMGLSPPASSTSTSTSCR
merge6E38.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	343	SDSARGIRIPYPGR
merge6E38.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	343	TFLLK
merge6E38.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	343	TFLLK
merge6E38.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	343	VFAAEALLK
merge6E38.txt	11	ZN828_HUMAN	Zinc finger protein 828	89099	51	IPCNSAEPK
merge6E38.txt	11	ZN828_HUMAN	Zinc finger protein 828	89099	51	NVLQFTEEK
merge6E38.txt	11	ZN828_HUMAN	Zinc finger protein 828	89099	51	CDILVQEELLASPKK
merge6E38.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	343	YSTWEPEENILDAR
merge6E38.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	343	YSTWEPEENILDAR
merge6E39.txt	2	CBX8_HUMAN	Chromobox protein homolog 8	43369	186	LLAAFEER
merge6E39.txt	2	CBX8_HUMAN	Chromobox protein homolog 8	43369	186	LLAAFEEREREMELYGPK
merge6E39.txt	2	CBX8_HUMAN	Chromobox protein homolog 8	43369	186	VFAAEALLK
merge6E39.txt	2	CBX8_HUMAN	Chromobox protein homolog 8	43369	186	VFAAEALLKR
merge6E39.txt	2	CBX8_HUMAN	Chromobox protein homolog 8	43369	186	YSTWEPEENILDAR
merge6E40.txt	5	CBX8_HUMAN	Chromobox protein homolog 8	43369	167	LLAAFEER
merge6E40.txt	5	CBX8_HUMAN	Chromobox protein homolog 8	43369	167	LLAAFEEREREMELYGPK
merge6E40.txt	5	CBX8_HUMAN	Chromobox protein homolog 8	43369	167	LLAAFEEREREMELYGPK
merge6E40.txt	5	CBX8_HUMAN	Chromobox protein homolog 8	43369	167	QDSDLVQCQGVTSPPSAEATGK
merge6E40.txt	5	CBX8_HUMAN	Chromobox protein homolog 8	43369	167	VFAAEALLK
merge6E40.txt	5	CBX8_HUMAN	Chromobox protein homolog 8	43369	167	VFAAEALLKR
merge6E40.txt	5	CBX8_HUMAN	Chromobox protein homolog 8	43369	167	YSTWEPEENILDAR
merge6E41.txt	8	CBX8_HUMAN	Chromobox protein homolog 8	43369	135	LLAAFEER
merge6E41.txt	8	CBX8_HUMAN	Chromobox protein homolog 8	43369	135	LLAAFEERER
merge6E41.txt	8	CBX8_HUMAN	Chromobox protein homolog 8	43369	135	LLAAFEEREREMELYGPK
merge6E41.txt	8	CBX8_HUMAN	Chromobox protein homolog 8	43369	135	YSTWEPEENILDAR
merge6E42.txt	4	CBX8_HUMAN	Chromobox protein homolog 8	43369	217	EREMELYGPK
merge6E42.txt	4	CBX8_HUMAN	Chromobox protein homolog 8	43369	217	LLAAFEER
merge6E42.txt	4	CBX8_HUMAN	Chromobox protein homolog 8	43369	217	LLAAFEEREREMELYGPK
merge6E42.txt	4	CBX8_HUMAN	Chromobox protein homolog 8	43369	217	LLAAFEEREREMELYGPK
merge6E42.txt	4	CBX8_HUMAN	Chromobox protein homolog 8	43369	217	VFAAEALLK
merge6E42.txt	4	CBX8_HUMAN	Chromobox protein homolog 8	43369	217	VFAAEALLKR
merge6E42.txt	21	H2A1B_HUMAN	Histone H2A type 1-B	14127	45	VTIAQGGVLPNIQAVLLPK
merge6E42.txt	12	H2B1B_HUMAN	Histone H2B type 1-B	13942	85	LLPGELAK
merge6E42.txt	12	H2B1B_HUMAN	Histone H2B type 1-B	13942	85	AMGIMNSFVNDIFER
merge6E42.txt	4	CBX8_HUMAN	Chromobox protein homolog 8	43369	217	YSTWEPEENILDAR
merge6E43.txt	8	CBX8_HUMAN	Chromobox protein homolog 8	43369	93	VFAAEALLKR
merge6E43.txt	10	H2A1B_HUMAN	Histone H2A type 1-B	14127	62	VTIAQGGVLPNIQAVLLPK
merge6E43.txt	8	CBX8_HUMAN	Chromobox protein homolog 8	43369	93	YSTWEPEENILDAR
merge6E44.txt	2	H4_HUMAN	Histone H4	11360	143	TYLGFEGG
merge6E44.txt	2	H4_HUMAN	Histone H4	11360	143	VFLENVIR
merge6E44.txt	2	H4_HUMAN	Histone H4	11360	143	ISGLIYEETR
merge6E44.txt	2	H4_HUMAN	Histone H4	11360	143	DNIQGITKPAIR
merge6E44.txt	7	CBX8_HUMAN	Chromobox protein homolog 8	43369	42	YSTWEPEENILDAR
merge6E45.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	149	LLAAFEERER
merge6E45.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	149	VFAAEALLKR
merge6E45.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	149	YSTWEPEENILDAR
merge6E45.txt	4	H4_HUMAN	Histone H4	11360	89	TYLGFEGG
merge6E45.txt	4	H4_HUMAN	Histone H4	11360	89	DNIQGITKPAIR

Tableau S6 : Peptides identifiés dans la purification RING1-TAP (n°21)

merge21E17.txt	5 PSD12_HUMAN	26S proteasome non-ATPase regulatory subunit 12	52871	42 EWDLLNENIMLLSKR
merge21E17.txt	5 PSD12_HUMAN	26S proteasome non-ATPase regulatory subunit 12	52871	42 EAASILQELQVETYGSMEKK
merge21E17.txt	3 RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	87 FCSDCIVTALR
merge21E17.txt	3 RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	87 SLRPDPNFDALISK
merge21E18.txt	2 RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	272 TTGNATVDHLSK
merge21E18.txt	2 RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	272 FCSDCIVTALR
merge21E18.txt	2 RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	272 VSRPLELCYPTK
merge21E18.txt	2 RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	272 SLRPDPNFDALISK
merge21E18.txt	2 RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	272 TPQEAIMDGTIAVSPR
merge21E18.txt	7 UBE2O_HUMAN	Ubiquitin-conjugating enzyme E2 O	1E+05	42 FRRTDIVIR
merge21E19.txt	1 RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	389 TTGNATVDHLSK
merge21E19.txt	1 RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	389 FCSDCIVTALR
merge21E19.txt	1 RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	389 TWELSLYELHR
merge21E19.txt	1 RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	389 SLRPDPNFDALISK
merge21E19.txt	1 RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	389 SLRPDPNFDALISK
merge21E19.txt	1 RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	389 LHNQALSSSIEEGLR
merge21E19.txt	1 RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	389 TPQEAIMDGTIAVSPR
merge21E19.txt	1 RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	389 TPQEAIMDGTIAVSPR
merge21E19.txt	1 RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	389 VSRPLELCYPTKDKP
merge21E19.txt	1 RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	389 GGGAGSSVGTGGGGTGGVGGGAGSEDSGDR
merge21E20.txt	1 RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	530 IALER
merge21E20.txt	1 RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	530 YLALR
merge21E20.txt	1 RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	530 TTGNATVDHLSK
merge21E20.txt	1 RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	530 TWELSLYELHR
merge21E20.txt	1 RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	530 TWELSLYELHR
merge21E20.txt	1 RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	530 VSRPLELCYPTK
merge21E20.txt	1 RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	530 VSRPLELCYPTK
merge21E20.txt	1 RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	530 SLRPDPNFDALISK
merge21E20.txt	1 RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	530 SLRPDPNFDALISK
merge21E20.txt	1 RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	530 IYPSREEYEAHQDR
merge21E20.txt	1 RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	530 IYPSREEYEAHQDR
merge21E20.txt	1 RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	530 TPQEAIMDGTIAVSPR
merge21E20.txt	1 RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	530 TPQEAIMDGTIAVSPR
merge21E20.txt	1 RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	530 TPQEAIMDGTIAVSPR
merge21E20.txt	1 RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	530 VSRPLELCYPTKDKP
merge21E20.txt	1 RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	530 SLHSELMCPICLDMLK
merge21E20.txt	1 RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	530 SLHSELMCPICLDMLK
merge21E21.txt	1 RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	357 TTGNATVDHLSK
merge21E21.txt	1 RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	357 FCSDCIVTALR
merge21E21.txt	1 RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	357 TWELSLYELHR
merge21E21.txt	1 RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	357 SLRPDPNFDALISK
merge21E21.txt	1 RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	357 SLRPDPNFDALISK
merge21E21.txt	1 RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	357 LHNQALSSSIEEGLR
merge21E21.txt	1 RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	357 IYPSREEYEAHQDR
merge21E21.txt	1 RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	357 TPQEAIMDGTIAVSPR
merge21E22.txt	3 RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	99 FCSDCIVTALR
merge21E22.txt	3 RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	99 SLRPDPNFDALISK
merge21E22.txt	3 RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	99 SLRPDPNFDALISK
merge21E23.txt	1 RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	185 TTGNATVDHLSK
merge21E23.txt	1 RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	185 FCSDCIVTALR
merge21E23.txt	1 RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	185 VSRPLELCYPTK
merge21E23.txt	1 RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	185 SLRPDPNFDALISK
merge21E23.txt	1 RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	185 SLRPDPNFDALISK
merge21E23.txt	1 RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	185 VSRPLELCYPTKDKP
merge21E24.txt	2 CBX8_HUMAN	Chromobox protein homolog 8	43369	187 SPQDLASTSR
merge21E24.txt	2 CBX8_HUMAN	Chromobox protein homolog 8	43369	187 YSTWEPEENILDAR
merge21E24.txt	2 CBX8_HUMAN	Chromobox protein homolog 8	43369	187 VVVTDVTSNFLTVTIK
merge21E24.txt	2 CBX8_HUMAN	Chromobox protein homolog 8	43369	187 RQDSDLVQCGVTPSSAEATGK
merge21E24.txt	4 PCGF2_HUMAN	Polycomb group RING finger protein 2	37764	60 NGPLPK
merge21E24.txt	4 PCGF2_HUMAN	Polycomb group RING finger protein 2	37764	60 LVPGLFKDEM
merge21E24.txt	4 PCGF2_HUMAN	Polycomb group RING finger protein 2	37764	60 CPAAMTVMHLAK
merge21E24.txt	3 PCGF6_HUMAN	Polycomb group RING finger protein 6	39023	97 GLYLDATTITECLHTFCK
merge21E24.txt	3 PCGF6_HUMAN	Polycomb group RING finger protein 6	39023	97 AIGDAMQDGLLVHLYGLVVSPLK
merge21E24.txt	1 RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	211 IALER
merge21E24.txt	1 RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	211 TTGNATVDHLSK
merge21E24.txt	1 RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	211 FCSDCIVTALR
merge21E24.txt	1 RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	211 SLRPDPNFDALISK
merge21E24.txt	1 RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	211 SLRPDPNFDALISK
merge21E25.txt	1 CSK21_HUMAN	Casein kinase II subunit alpha	45115	295 FYMYEILK
merge21E25.txt	1 CSK21_HUMAN	Casein kinase II subunit alpha	45115	295 QLYQTLTDYDIR
merge21E25.txt	1 CSK21_HUMAN	Casein kinase II subunit alpha	45115	295 EAMEHPYFYTVVK
merge21E25.txt	1 CSK21_HUMAN	Casein kinase II subunit alpha	45115	295 YSEVFEAINITNNEK
merge21E25.txt	1 CSK21_HUMAN	Casein kinase II subunit alpha	45115	295 GGNITLADIVKDPVSR
merge21E25.txt	1 CSK21_HUMAN	Casein kinase II subunit alpha	45115	295 EYWDYESHVVEWGNQDDYQLVR
merge21E25.txt	4 PHC2_HUMAN	Polyhmeotic-like protein 2	90657	89 AQEIDGQALLLLK
merge21E25.txt	2 RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	150 TTGNATVDHLSK
merge21E25.txt	2 RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	150 FCSDCIVTALR
merge21E25.txt	2 RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	150 LHNQALSSSIEEGLR
merge21E25.txt	2 RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	150 VSRPLELCYPTKDKP
merge21E26.txt	1 BMI1_HUMAN	Polycomb complex protein BMI-1	36925	268 LVPGLFK
merge21E26.txt	1 BMI1_HUMAN	Polycomb complex protein BMI-1	36925	268 TLQDIVYK

merge21E26.txt	1	BMI1_HUMAN	Polycomb complex protein BMI-1	36925	268	CPAAMTVMHLR
merge21E26.txt	1	BMI1_HUMAN	Polycomb complex protein BMI-1	36925	268	CPAAMTVMHLR
merge21E26.txt	1	BMI1_HUMAN	Polycomb complex protein BMI-1	36925	268	SDKTLQDIVYK
merge21E26.txt	1	BMI1_HUMAN	Polycomb complex protein BMI-1	36925	268	CPAAMTVMHLR
merge21E26.txt	1	BMI1_HUMAN	Polycomb complex protein BMI-1	36925	268	YCPICDVQVHK
merge21E26.txt	1	BMI1_HUMAN	Polycomb complex protein BMI-1	36925	268	DGLTNAGELESDSGSDK
merge21E26.txt	6	CSK21_HUMAN	Casein kinase II subunit alpha	45115	47	QLYQTLTDYDIR
merge21E26.txt	3	PHC2_HUMAN	Polymethionine-like protein 2	90657	110	AQEIDGQALLLTK
merge21E26.txt	3	PHC2_HUMAN	Polymethionine-like protein 2	90657	110	DLVGMGHFLPSEPTK
merge21E26.txt	2	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	157	FCSDCIVTALR
merge21E26.txt	2	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	157	SLRPDPNFDALISK
merge21E26.txt	2	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	157	LHNQALSSSIEEGLR
merge21E27.txt	5	CSK22_HUMAN	Casein kinase II subunit alpha	41187	45	VYAEVNSLR
merge21E27.txt	4	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	96	FCSDCIVTALR
merge21E27.txt	4	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	96	SLRPDPNFDALISK
merge21E28.txt	1	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	150	TTGNATVDHLSK
merge21E28.txt	1	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	150	FCSDCIVTALR
merge21E28.txt	1	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	150	SLRPDPNFDALISK
merge21E29.txt	2	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	173	TTGNATVDHLSK
merge21E29.txt	2	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	173	FCSDCIVTALR
merge21E29.txt	2	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	173	SLRPDPNFDALISK
merge21E29.txt	2	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	173	SLRPDPNFDALISK
merge21E29.txt	4	RYBP_HUMAN	RING1 and YY1-binding protein	24806	120	DKEISPSVTK
merge21E29.txt	4	RYBP_HUMAN	RING1 and YY1-binding protein	24806	120	INSQLVAQQVAQQYATPPPPK
merge21E29.txt	4	RYBP_HUMAN	RING1 and YY1-binding protein	24806	120	INSQLVAQQVAQQYATPPPPK
merge21E29.txt	4	RYBP_HUMAN	RING1 and YY1-binding protein	24806	120	INSQLVAQQVAQQYATPPPPK
merge21E30.txt	2	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	129	TTGNATVDHLSK
merge21E30.txt	2	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	129	SLRPDPNFDALISK
merge21E30.txt	2	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	129	SLRPDPNFDALISK
merge21E30.txt	2	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	129	IYPSREEYEAHQDR
merge21E32.txt	2	PCGF5_HUMAN	Polycomb group RING finger protein 5	29694	41	GYLIKPTTVTECLHTFCK
merge21E32.txt	3	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	39	SLRPDPNFDALISK
merge21E32.txt	1	YAF2_HUMAN	YY1-associated factor 2	19872	59	SPPASSAASADQHSQSGSSDNTER
merge21E33.txt	3	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	98	FCSDCIVTALR
merge21E33.txt	3	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	98	SLRPDPNFDALISK
merge21E33.txt	3	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	98	SLRPDPNFDALISK
merge21E33.txt	4	RYBP_HUMAN	RING1 and YY1-binding protein	24806	40	SDILKDPPEANSIQSANATTK
merge21E33.txt	1	YAF2_HUMAN	YY1-associated factor 2	19872	139	HAKPSSDEGYWDCSVCTFR
merge21E33.txt	1	YAF2_HUMAN	YY1-associated factor 2	19872	139	SPPASSAASADQHSQSGSSDNTER
merge21E33.txt	1	YAF2_HUMAN	YY1-associated factor 2	19872	139	KPRPVSQLVAQVQVFPPTQSK
merge21E34.txt	1	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	313	FCSDCIVTALR
merge21E34.txt	1	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	313	TWELSLYELHR
merge21E34.txt	1	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	313	LHNQALSSSIEEGLR
merge21E34.txt	1	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	313	LHNQALSSSIEEGLR
merge21E34.txt	1	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	313	IYPSREEYEAHQDR
merge21E34.txt	1	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	313	TPQEAIMDGTIEAVSPR
merge21E34.txt	1	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	313	TPQEAIMDGTIEAVSPR
merge21E34.txt	4	RYBP_HUMAN	RING1 and YY1-binding protein	24806	55	SDILKDPPEANSIQSANATTK
merge21E35.txt	12	NUMA1_HUMAN	Nuclear mitotic apparatus protein 1	2E+05	44	AEDWKAQVAR
merge21E35.txt	12	NUMA1_HUMAN	Nuclear mitotic apparatus protein 1	2E+05	44	AVQAQGGESQEQEQR
merge21E35.txt	12	NUMA1_HUMAN	Nuclear mitotic apparatus protein 1	2E+05	44	LRLQLAETASNSARAAR
merge21E35.txt	8	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	84	FCSDCIVTALR
merge21E35.txt	8	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	84	SLRPDPNFDALISK
merge21E35.txt	8	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	84	SLRPDPNFDALISK
merge21E35.txt	10	SKP1_HUMAN	S-phase kinase-associated protein 1	18646	65	NDFTEEEEAQVR
merge21E35.txt	10	SKP1_HUMAN	S-phase kinase-associated protein 1	18646	65	GLLDVCTKTVANMIK GK
merge21E36.txt	4	YAF2_HUMAN	YY1-associated factor 2	19872	40	SPPASSAASADQHSQSGSSDNTER
merge21E38.txt	1	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	117	SLRPDPNFDALISK
merge21E38.txt	1	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	117	TPQEAIMDGTIEAVSPR
merge21E39.txt	4	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	46	TWELSLYELHR

Tableau S7 : Peptides identifiés dans la purification PCGF2-TAP (n°19)

band	hit	accession	description	mass	score	peptide
19E01.txt	6	BRCA2_HUMAN	Breast cancer type 2 susceptibility protein	383986	43	IVSYSK
19E01.txt	6	BRCA2_HUMAN	Breast cancer type 2 susceptibility protein	383986	43	INDNEIHQFNK
19E01.txt	11	CENPE_HUMAN	Centromeric protein E	316219	41	LHENLEEMK
19E01.txt	11	CENPE_HUMAN	Centromeric protein E	316219	41	FQESQEEIKSLTK
19E01.txt	9	NEST_HUMAN	Nestin	177332	42	TLLEAENSR
19E01.txt	9	NEST_HUMAN	Nestin	177332	42	LELQQLQAERGGLLERR
19E13.txt	8	PHC3_HUMAN	Polymethionine-like protein 3	106096	59	SLLIEQPVK
19E16.txt	2	K6PP_HUMAN	6-phosphofructokinase type C	85542	164	DLQSNVEHLETK
19E16.txt	2	K6PP_HUMAN	6-phosphofructokinase type C	85542	164	AIGVLTSGGDAQGMNAAVR
19E16.txt	4	PHC3_HUMAN	Polymethionine-like protein 3	106096	111	MGYANEFLR
19E16.txt	4	PHC3_HUMAN	Polymethionine-like protein 3	106096	111	SLLIEQPVK
19E18.txt	3	CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	64	YNTWEPEENILDPR
19E18.txt	3	CBX4_HUMAN	E3 SUMO-protein ligase CBX4	61190	64	EYVTV
19E23.txt	13	DPOLL_HUMAN	DNA polymerase lambda	63443	42	AINALK
19E23.txt	13	DPOLL_HUMAN	DNA polymerase lambda	63443	42	YLGVCRLPGPGRRRH
19E23.txt	13	DPOLL_HUMAN	DNA polymerase lambda	63443	42	YLGVCRLPGPGRR
19E23.txt	5	PCGF2_HUMAN	Polycomb group RING finger protein 2	37764	318	NGPLPLK
19E23.txt	5	PCGF2_HUMAN	Polycomb group RING finger protein 2	37764	318	TLQDIVYK
19E23.txt	5	PCGF2_HUMAN	Polycomb group RING finger protein 2	37764	318	KGPLENGDGDK

19E23.txt	5	PCGF2_HUMAN	Polycomb group RING finger protein 2	37764	318	LVPGLFKDEM
19E23.txt	5	PCGF2_HUMAN	Polycomb group RING finger protein 2	37764	318	SDKTLQDIVYK
19E23.txt	5	PCGF2_HUMAN	Polycomb group RING finger protein 2	37764	318	CPAAMTVMHLAK
19E23.txt	5	PCGF2_HUMAN	Polycomb group RING finger protein 2	37764	318	YCPMCDVQVHK
19E23.txt	5	PCGF2_HUMAN	Polycomb group RING finger protein 2	37764	318	LTLATVPTPSEGNTSGASECESVSDK
19E23.txt	6	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	303	TTGNATVDHLSK
19E23.txt	6	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	303	FCSDCIVTALR
19E23.txt	6	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	303	SLRPDPNFDALISK
19E23.txt	6	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	303	SLRPDPNFDALISK
19E23.txt	6	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	303	LHNQQALSSSIEEGLR
19E23.txt	6	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	303	IYPSREEYEAHQDR
19E23.txt	6	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	303	GGGAGGSSVGTGGGTGGVGGGAGEDSGDR
19E24.txt	4	PCGF2_HUMAN	Polycomb group RING finger protein 2	37764	410	NGPLPLK
19E24.txt	4	PCGF2_HUMAN	Polycomb group RING finger protein 2	37764	410	LVPGLFK
19E24.txt	4	PCGF2_HUMAN	Polycomb group RING finger protein 2	37764	410	TLQDIVYK
19E24.txt	4	PCGF2_HUMAN	Polycomb group RING finger protein 2	37764	410	TLQDIVYK
19E24.txt	4	PCGF2_HUMAN	Polycomb group RING finger protein 2	37764	410	LVPGLFKDEM
19E24.txt	4	PCGF2_HUMAN	Polycomb group RING finger protein 2	37764	410	LVPGLFKDEM
19E24.txt	4	PCGF2_HUMAN	Polycomb group RING finger protein 2	37764	410	SDKTLQDIVYK
19E24.txt	4	PCGF2_HUMAN	Polycomb group RING finger protein 2	37764	410	CPAAMTVMHLAK
19E24.txt	4	PCGF2_HUMAN	Polycomb group RING finger protein 2	37764	410	CPAAMTVMHLAK
19E24.txt	4	PCGF2_HUMAN	Polycomb group RING finger protein 2	37764	410	CPAAMTVMHLAK
19E24.txt	4	PCGF2_HUMAN	Polycomb group RING finger protein 2	37764	410	KGPLENGDGDKK
19E24.txt	4	PCGF2_HUMAN	Polycomb group RING finger protein 2	37764	410	YCPMCDVQVHK
19E24.txt	4	PCGF2_HUMAN	Polycomb group RING finger protein 2	37764	410	YCPMCDVQVHK
19E24.txt	4	PCGF2_HUMAN	Polycomb group RING finger protein 2	37764	410	LTLATVPTPSEGNTSGASECESVSDK
19E24.txt	4	PCGF2_HUMAN	Polycomb group RING finger protein 2	37764	410	RLTLATVPTPSEGNTSGASECESVSDK
19E24.txt	11	RECQ1_HUMAN	ATP-dependent DNA helicase Q1	73409	42	EDFPWSGKVK
19E24.txt	11	RECQ1_HUMAN	ATP-dependent DNA helicase Q1	73409	42	MEKKNNGNFQK
19E24.txt	2	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	492	IALER
19E24.txt	2	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	492	TTGNATVDHLSK
19E24.txt	2	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	492	FCSDCIVTALR
19E24.txt	2	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	492	FCSDCIVTALR
19E24.txt	2	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	492	VSRPLELCYAPTK
19E24.txt	2	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	492	SLRPDPNFDALISK
19E24.txt	2	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	492	LHNQQALSSSIEEGLR
19E24.txt	2	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	492	TPQEAIMDGTEIAVSPR
19E24.txt	2	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	492	TPQEAIMDGTEIAVSPR
19E24.txt	2	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	492	VSRPLELCYAPTKDKPK
19E24.txt	13	TPR_HUMAN	Nucleoprotein TPR	265441	40	IVKPGMK
19E24.txt	13	TPR_HUMAN	Nucleoprotein TPR	265441	40	EKAENEKIQNEQLEK
19E24.txt	13	TPR_HUMAN	Nucleoprotein TPR	265441	40	GAILSEEEEAAMSPTAAAVAK
19E24.txt	13	TPR_HUMAN	Nucleoprotein TPR	265441	40	EEEEEDSTIEASDQVSDTVMEMPLPKK
19E25.txt	3	PCGF2_HUMAN	Polycomb group RING finger protein 2	37764	158	TLQDIVYK
19E25.txt	3	PCGF2_HUMAN	Polycomb group RING finger protein 2	37764	158	SDKTLQDIVYK
19E25.txt	3	PCGF2_HUMAN	Polycomb group RING finger protein 2	37764	158	CPAAMTVMHLAK
19E25.txt	3	PCGF2_HUMAN	Polycomb group RING finger protein 2	37764	158	CPAAMTVMHLAK
19E25.txt	2	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	160	VSRPLELCYAPTK
19E25.txt	2	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	160	SLRPDPNFDALISK
19E25.txt	2	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	160	LHNQQALSSSIEEGLR
19E26.txt	5	BMI1_HUMAN	Polycomb complex protein BMI-1	36925	77	TLQDIVYK
19E26.txt	5	BMI1_HUMAN	Polycomb complex protein BMI-1	36925	77	SDKTLQDIVYK
19E26.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	341	TFLK
19E26.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	341	AAFLEAK
19E26.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	341	VFAAEALLK
19E26.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	341	SPQDLASTSR
19E26.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	341	ESNTDQGFFK
19E26.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	341	YSTWEPEENILDAR
19E26.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	341	NMGLSPPASSTSTSSTCR
19E26.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	341	NMGLSPPASSTSTSSTCR
19E26.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	341	RQDSDLVQCQVTSPPSAAEATGK
19E26.txt	1	CBX8_HUMAN	Chromobox protein homolog 8	43369	341	KELPDPSQRPLGEPASAGLGEYLK
19E26.txt	3	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	106	FCSDCIVTALR
19E26.txt	3	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	106	SLRPDPNFDALISK
19E27.txt	8	PCGF2_HUMAN	Polycomb group RING finger protein 2	37764	74	TLQDIVYK
19E27.txt	8	PCGF2_HUMAN	Polycomb group RING finger protein 2	37764	74	SDKTLQDIVYK
19E27.txt	8	PCGF2_HUMAN	Polycomb group RING finger protein 2	37764	74	KMTVNGAPVPLT
19E27.txt	2	PHC2_HUMAN	Polyhomeotic-like protein 2	90657	207	YAQGFLPEK
19E27.txt	2	PHC2_HUMAN	Polyhomeotic-like protein 2	90657	207	AQEIDGQALLLK
19E27.txt	2	PHC2_HUMAN	Polyhomeotic-like protein 2	90657	207	SLPGCQEIIEEFR
19E27.txt	4	PHC3_HUMAN	Polyhomeotic-like protein 3	106096	153	FCTMSCAKR
19E27.txt	4	PHC3_HUMAN	Polyhomeotic-like protein 3	106096	153	SSLIEQPVK
19E27.txt	4	PHC3_HUMAN	Polyhomeotic-like protein 3	106096	153	AQEIDGQALLLK
19E27.txt	5	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	106	SLRPDPNFDALISK

19E27.txt	5	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	106	LHNQQALSSSIEEGLR
19E28.txt	3	PHC2_HUMAN	Polyhomeotic-like protein 2	90657	55	ISMLK
19E28.txt	3	PHC2_HUMAN	Polyhomeotic-like protein 2	90657	55	SLLVGNLK
19E28.txt	1	RING2_HUMAN	E3 ubiquitin-protein ligase RING2	37632	252	FCADCIITLR
19E28.txt	1	RING2_HUMAN	E3 ubiquitin-protein ligase RING2	37632	252	VNKPMELYYPTK
19E28.txt	1	RING2_HUMAN	E3 ubiquitin-protein ligase RING2	37632	252	SLRPDPNFDALISK
19E28.txt	1	RING2_HUMAN	E3 ubiquitin-protein ligase RING2	37632	252	SKGESNQMNLDTASEK
19E28.txt	1	RING2_HUMAN	E3 ubiquitin-protein ligase RING2	37632	252	IYPSRDEYEAHQER
19E30.txt	1	BMI1_HUMAN	Polycomb complex protein BMI-1	36925	110	TLQDIVYK
19E30.txt	1	BMI1_HUMAN	Polycomb complex protein BMI-1	36925	110	LVPGLFKNEMK
19E30.txt	1	BMI1_HUMAN	Polycomb complex protein BMI-1	36925	110	SDKTLQDIVYK
19E33.txt	3	YAF2_HUMAN	YY1-associated factor 2	19872	68	SSSPR
19E33.txt	3	YAF2_HUMAN	YY1-associated factor 2	19872	68	KPRPVSQLVAQVQTQQFVPTQSK
19E33.txt	3	YAF2_HUMAN	YY1-associated factor 2	19872	68	KPRPVSQLVAQVQTQQFVPTQSK
19E34.txt	4	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	46	SLRPDPNFDALISK
19E34.txt	2	RYBP_HUMAN	RING1 and YY1-binding protein	24806	58	SDILKDPPEANSIQSANATTK
19E34.txt	5	YAF2_HUMAN	YY1-associated factor 2	19872	44	KPRPVSQLVAQVQTQQFVPTQSK
19E35.txt	4	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	39	SLRPDPNFDALISK
19E35.txt	5	YAF2_HUMAN	YY1-associated factor 2	19872	38	KPRPVSQLVAQVQTQQFVPTQSK
19E37.txt	2	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	213	FCSDCIVTALR
19E37.txt	2	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	213	SLRPDPNFDALISK
19E37.txt	2	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	213	LHNQQALSSSIEEGLR
19E37.txt	2	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	213	IYPSREEYEAHQDR
19E38.txt	1	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	165	FCSDCIVTALR
19E38.txt	1	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	165	SLRPDPNFDALISK
19E38.txt	1	RING1_HUMAN	E3 ubiquitin-protein ligase RING1	42403	165	IYPSREEYEAHQDR
19E41.txt	12	INCE_HUMAN	Inner centromere protein	106049	40	RHQLLQKK
19E41.txt	12	INCE_HUMAN	Inner centromere protein	106049	40	AAAKMEEVEARR
19E41.txt	12	INCE_HUMAN	Inner centromere protein	106049	40	QNAEQHVHTQLMSTELPR

Tableau S8 : Peptides identifiés dans la purification CBX1-TAP (n°16)

fragment	hit	accession	description	masse	score	peptide
16E07.txt	1	CHD4_HUMAN	Chromodomain-helicase-DNA-binding protein 4 - Homo sapiens (Human)	217853	388	EFSTNNPFK
16E07.txt	1	CHD4_HUMAN	Chromodomain-helicase-DNA-binding protein 4 - Homo sapiens (Human)	217853	388	AFSQFVRPLIAAK
16E07.txt	1	CHD4_HUMAN	Chromodomain-helicase-DNA-binding protein 4 - Homo sapiens (Human)	217853	388	ENEF5FDENAIR
16E07.txt	1	CHD4_HUMAN	Chromodomain-helicase-DNA-binding protein 4 - Homo sapiens (Human)	217853	388	YAILNEPFGKEMNR
16E07.txt	1	CHD4_HUMAN	Chromodomain-helicase-DNA-binding protein 4 - Homo sapiens (Human)	217853	388	KLERPPPTPTVDPTVK
16E07.txt	1	CHD4_HUMAN	Chromodomain-helicase-DNA-binding protein 4 - Homo sapiens (Human)	217853	388	KLERPPPTPTVDPTVK
16E07.txt	1	CHD4_HUMAN	Chromodomain-helicase-DNA-binding protein 4 - Homo sapiens (Human)	217853	388	FNAPGAQFCFLSTR
16E07.txt	1	CHD4_HUMAN	Chromodomain-helicase-DNA-binding protein 4 - Homo sapiens (Human)	217853	388	FAEVECLAESHQHSK
16E07.txt	1	CHD4_HUMAN	Chromodomain-helicase-DNA-binding protein 4 - Homo sapiens (Human)	217853	388	GPFLVSAPLSTIINWER
16E07.txt	1	CHD4_HUMAN	Chromodomain-helicase-DNA-binding protein 4 - Homo sapiens (Human)	217853	388	NQDETETELQGMNEYLSSFK
16E08.txt	2	CHD3_HUMAN	Chromodomain-helicase-DNA-binding protein 3 - Homo sapiens (Human)	226450	45	LLEQLVIEEQLR
16E08.txt	2	CHD3_HUMAN	Chromodomain-helicase-DNA-binding protein 3 - Homo sapiens (Human)	226450	45	GPFLVSAPLSTIINWER
16E09.txt	1	SMCA4_HUMAN	Probable global transcription activator SNF2L4 - Homo sapiens (Human)	184530	107	GLQSYAVAHAVTER
16E09.txt	1	SMCA4_HUMAN	Probable global transcription activator SNF2L4 - Homo sapiens (Human)	184530	107	EVEAQLPEKVEYVIK
16E09.txt	1	SMCA4_HUMAN	Probable global transcription activator SNF2L4 - Homo sapiens (Human)	184530	107	QEVVCMRRDIALETALNAK
16E10.txt	3	POGZ_HUMAN	Pogo transposable element with ZNF domain - Homo sapiens	155245	161	SSLYSEVDVHFR
16E10.txt	3	POGZ_HUMAN	Pogo transposable element with ZNF domain - Homo sapiens	155245	161	EQQLPVNEETLFQK
16E10.txt	3	POGZ_HUMAN	Pogo transposable element with ZNF domain - Homo sapiens	155245	161	VAQLTNFPKVAVSFRCPHCTKR
16E10.txt	3	POGZ_HUMAN	Pogo transposable element with ZNF domain - Homo sapiens	155245	161	QFSTPFQLQCHLENVHSPYESTTK
16E11.txt	1	POGZ_HUMAN	Pogo transposable element with ZNF domain - Homo sapiens	155245	271	EQQLPVNEETLFQK
16E11.txt	1	POGZ_HUMAN	Pogo transposable element with ZNF domain - Homo sapiens	155245	271	VVLFALCCNTEQAAEHFR
16E11.txt	1	POGZ_HUMAN	Pogo transposable element with ZNF domain - Homo sapiens	155245	271	VTEALRGHMICYCCPEMVEYQKK
16E11.txt	1	POGZ_HUMAN	Pogo transposable element with ZNF domain - Homo sapiens	155245	271	QFSTPFQLQCHLENVHSPYESTTK
16E11.txt	1	POGZ_HUMAN	Pogo transposable element with ZNF domain - Homo sapiens	155245	271	EQLSVKLRLVLFALCCNTEQAAEHFR
16E12.txt	1	ADNP_HUMAN	Activity-dependent neuroprotector - Homo sapiens (Human)	123485	919	VPVAVANIYMK
16E12.txt	1	ADNP_HUMAN	Activity-dependent neuroprotector - Homo sapiens (Human)	123485	919	VPVAVANIYMK
16E12.txt	1	ADNP_HUMAN	Activity-dependent neuroprotector - Homo sapiens (Human)	123485	919	GPISDALAHHLR
16E12.txt	1	ADNP_HUMAN	Activity-dependent neuroprotector - Homo sapiens (Human)	123485	919	ILSDIGLEYCK
16E12.txt	1	ADNP_HUMAN	Activity-dependent neuroprotector - Homo sapiens (Human)	123485	919	YKPGVLLGFNMK
16E12.txt	1	ADNP_HUMAN	Activity-dependent neuroprotector - Homo sapiens (Human)	123485	919	EHFQHVAAPIYAK
16E12.txt	1	ADNP_HUMAN	Activity-dependent neuroprotector - Homo sapiens (Human)	123485	919	EHFQHVAAPIYAK
16E12.txt	1	ADNP_HUMAN	Activity-dependent neuroprotector - Homo sapiens (Human)	123485	919	QADSVQAVYYCK
16E12.txt	1	ADNP_HUMAN	Activity-dependent neuroprotector - Homo sapiens (Human)	123485	919	CTYRDPLYEIVR
16E12.txt	1	ADNP_HUMAN	Activity-dependent neuroprotector - Homo sapiens (Human)	123485	919	TKPFCCSACPFSSK
16E12.txt	1	ADNP_HUMAN	Activity-dependent neuroprotector - Homo sapiens (Human)	123485	919	TKPFCCSACPFSSK
16E12.txt	1	ADNP_HUMAN	Activity-dependent neuroprotector - Homo sapiens (Human)	123485	919	KCTYRDPLYEIVR
16E12.txt	1	ADNP_HUMAN	Activity-dependent neuroprotector - Homo sapiens (Human)	123485	919	YKPGVLLGFNMK
16E12.txt	1	ADNP_HUMAN	Activity-dependent neuroprotector - Homo sapiens (Human)	123485	919	EHFQHVAAPIYAK
16E12.txt	1	ADNP_HUMAN	Activity-dependent neuroprotector - Homo sapiens (Human)	123485	919	ILLNCPYCTFNADKK
16E12.txt	1	ADNP_HUMAN	Activity-dependent neuroprotector - Homo sapiens (Human)	123485	919	ILLNCPYCTFNADKK
16E12.txt	1	ADNP_HUMAN	Activity-dependent neuroprotector - Homo sapiens (Human)	123485	919	IFHAPNASAPSSLSSTFK
16E12.txt	1	ADNP_HUMAN	Activity-dependent neuroprotector - Homo sapiens (Human)	123485	919	IFHAPNASAPSSLSSTFK
16E12.txt	1	ADNP_HUMAN	Activity-dependent neuroprotector - Homo sapiens (Human)	123485	919	LGLGGNAPVSIPIQQSQSVK
16E12.txt	1	ADNP_HUMAN	Activity-dependent neuroprotector - Homo sapiens (Human)	123485	919	EHIEDFKQFEPNDFYK
16E12.txt	1	ADNP_HUMAN	Activity-dependent neuroprotector - Homo sapiens (Human)	123485	919	VLGQSSSKPAAAATGPPPGNTSSTQK
16E12.txt	1	ADNP_HUMAN	Activity-dependent neuroprotector - Homo sapiens (Human)	123485	919	ICTICNELFPENVYVHFKE
16E12.txt	1	ADNP_HUMAN	Activity-dependent neuroprotector - Homo sapiens (Human)	123485	919	ICTICNELFPENVYVHFKE
16E12.txt	1	ADNP_HUMAN	Activity-dependent neuroprotector - Homo sapiens (Human)	123485	919	DAPAESVAYHAQNNPPVPPKQPK

16E12.txt	3	INT1_HUMAN	Integrator complex subunit 1 - Homo sapiens (Human)	244143	41	ALENRGADASMACRK
16E12.txt	3	INT1_HUMAN	Integrator complex subunit 1 - Homo sapiens (Human)	244143	41	QTKHEINFQAFCLGLMQR
16E13.txt	1	ADNP_HUMAN	Activity-dependent neuroprotector - Homo sapiens (Human)	123485	393	VPVAVANYIMK
16E13.txt	1	ADNP_HUMAN	Activity-dependent neuroprotector - Homo sapiens (Human)	123485	393	SDIASHFSNKR
16E13.txt	1	ADNP_HUMAN	Activity-dependent neuroprotector - Homo sapiens (Human)	123485	393	HQVIQTVHPVEK
16E13.txt	1	ADNP_HUMAN	Activity-dependent neuroprotector - Homo sapiens (Human)	123485	393	EHFQHVAAPIYAK
16E13.txt	1	ADNP_HUMAN	Activity-dependent neuroprotector - Homo sapiens (Human)	123485	393	NTTWEDVGLWDPSSLTK
16E13.txt	1	ADNP_HUMAN	Activity-dependent neuroprotector - Homo sapiens (Human)	123485	393	VLGQSSSKPAAAATGPPPGNTSSTQK
16E13.txt	1	ADNP_HUMAN	Activity-dependent neuroprotector - Homo sapiens (Human)	123485	393	DAPAESVAYHAQNNPPVPPKQPK
16E13.txt	1	ADNP_HUMAN	Activity-dependent neuroprotector - Homo sapiens (Human)	123485	393	YLPTDTLLNHMLIHLGSLCPYCR
16E13.txt	2	ADNP2_HUMAN	ADNP homeobox protein 2 - Homo sapiens (Human)	122755	93	SLVPVYVVK
16E13.txt	2	ADNP2_HUMAN	ADNP homeobox protein 2 - Homo sapiens (Human)	122755	93	SVISEHIKR
16E13.txt	2	ADNP2_HUMAN	ADNP homeobox protein 2 - Homo sapiens (Human)	122755	93	TCPVCNELFPSNVYQVHMEVAHK
16E13.txt	3	POGZ_HUMAN	Pogo transposable element with ZNF domain - Homo sapiens (Human)	155245	64	EQQLPVNEETLFQK
16E14.txt	5	ADNP_HUMAN	Activity-dependent neuroprotector - Homo sapiens (Human)	123485	47	ILLNCPYCTFNADKK
16E14.txt	5	ADNP_HUMAN	Activity-dependent neuroprotector - Homo sapiens (Human)	123485	47	LGLGGNAPVSIPIQQSQSVK
16E14.txt	4	ADNP2_HUMAN	ADNP homeobox protein 2 - Homo sapiens (Human)	122755	55	VQNYTVNILGETK
16E14.txt	2	CAF1A_HUMAN	Chromatin assembly factor 1 subunit A - Homo sapiens	105158	100	IGCVWAADRDCAGDDLK
16E14.txt	2	CAF1A_HUMAN	Chromatin assembly factor 1 subunit A - Homo sapiens	105158	100	IGCVWAADRDCAGDDLK
16E14.txt	2	CAF1A_HUMAN	Chromatin assembly factor 1 subunit A - Homo sapiens	105158	100	IGCVWAADRDCAGDDLK
16E14.txt	2	CAF1A_HUMAN	Chromatin assembly factor 1 subunit A - Homo sapiens	105158	100	EDSGSVPTSGPSQGTPISLK
16E16.txt	1	ADNP_HUMAN	Activity-dependent neuroprotector - Homo sapiens (Human)	123485	206	VPVAVANYIMK
16E16.txt	1	ADNP_HUMAN	Activity-dependent neuroprotector - Homo sapiens (Human)	123485	206	TLCPLCLFSILK
16E16.txt	1	ADNP_HUMAN	Activity-dependent neuroprotector - Homo sapiens (Human)	123485	206	ADIPVKSSPQAAVPIYK
16E16.txt	1	ADNP_HUMAN	Activity-dependent neuroprotector - Homo sapiens (Human)	123485	206	LGLGGNAPVSIPIQQSQSVK
16E16.txt	1	ADNP_HUMAN	Activity-dependent neuroprotector - Homo sapiens (Human)	123485	206	VLGQSSSKPAAAATGPPPGNTSSTQK
16E16.txt	1	ADNP_HUMAN	Activity-dependent neuroprotector - Homo sapiens (Human)	123485	206	DAPAESVAYHAQNNPPVPPKQPK
16E16.txt	1	ADNP_HUMAN	Activity-dependent neuroprotector - Homo sapiens (Human)	123485	236	SDIASHFSNKR
16E16.txt	1	ADNP_HUMAN	Activity-dependent neuroprotector - Homo sapiens (Human)	123485	236	HQVIQTVHPVEK
16E16.txt	1	ADNP_HUMAN	Activity-dependent neuroprotector - Homo sapiens (Human)	123485	236	STFNDEVKMAAHMR
16E17.txt	1	ADNP_HUMAN	Activity-dependent neuroprotector - Homo sapiens (Human)	123485	236	VLGQSSSKPAAAATGPPPGNTSSTQK
16E17.txt	1	ADNP_HUMAN	Activity-dependent neuroprotector - Homo sapiens (Human)	123485	236	ICTICNELFPENVYSVHFKEK
16E17.txt	1	ADNP_HUMAN	Activity-dependent neuroprotector - Homo sapiens (Human)	123485	236	DAPAESVAYHAQNNPPVPPKQPK
16E17.txt	1	ADNP_HUMAN	Activity-dependent neuroprotector - Homo sapiens (Human)	123485	236	DAPAESVAYHAQNNPPVPPKQPK
16E17.txt	2	TIF1B_HUMAN	Transcription intermediary factor 1-beta - Homo sapiens	88493	165	ADVQSIQLQR
16E17.txt	2	TIF1B_HUMAN	Transcription intermediary factor 1-beta - Homo sapiens	88493	165	MIVDPVEPHGEMK
16E17.txt	2	TIF1B_HUMAN	Transcription intermediary factor 1-beta - Homo sapiens	88493	165	LSPPYSSPQEFQAQDVGR
16E17.txt	2	TIF1B_HUMAN	Transcription intermediary factor 1-beta - Homo sapiens	88493	165	LSPPYSSPQEFQAQDVGRMFKQFNK
16E18.txt	2	ADNP_HUMAN	Activity-dependent neuroprotector - Homo sapiens (Human)	123485	240	SDIASHFSNKR
16E18.txt	2	ADNP_HUMAN	Activity-dependent neuroprotector - Homo sapiens (Human)	123485	240	HQVIQTVHPVEK
16E18.txt	2	ADNP_HUMAN	Activity-dependent neuroprotector - Homo sapiens (Human)	123485	240	LGLGGNAPVSIPIQQSQSVK
16E18.txt	2	ADNP_HUMAN	Activity-dependent neuroprotector - Homo sapiens (Human)	123485	240	VLGQSSSKPAAAATGPPPGNTSSTQK
16E18.txt	2	ADNP_HUMAN	Activity-dependent neuroprotector - Homo sapiens (Human)	123485	240	ICTICNELFPENVYSVHFKEK
16E18.txt	2	ADNP_HUMAN	Activity-dependent neuroprotector - Homo sapiens (Human)	123485	240	DAPAESVAYHAQNNPPVPPKQPK
16E18.txt	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta - Homo sapiens	88493	630	LLASLVK
16E18.txt	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta - Homo sapiens	88493	630	LIYFQLHR
16E18.txt	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta - Homo sapiens	88493	630	SEGEVVSGLMR
16E18.txt	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta - Homo sapiens	88493	630	FQWDLNAWTK
16E18.txt	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta - Homo sapiens	88493	630	MIVDPVEPHGEMK
16E18.txt	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta - Homo sapiens	88493	630	MIVDPVEPHGEMK
16E18.txt	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta - Homo sapiens	88493	630	MIVDPVEPHGEMK
16E18.txt	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta - Homo sapiens	88493	630	DHQQFLEDAVR
16E18.txt	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta - Homo sapiens	88493	630	LDLDTADSQPPVFK
16E18.txt	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta - Homo sapiens	88493	630	LTEDKADVQSIQLQR
16E18.txt	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta - Homo sapiens	88493	630	LSPPYSSPQEFQAQDVGR
16E18.txt	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta - Homo sapiens	88493	630	VFPGSTTEDYNLIVIER
16E18.txt	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta - Homo sapiens	88493	630	VFPGSTTEDYNLIVIER
16E18.txt	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta - Homo sapiens	88493	630	IVAERPGTNSTGPAPMAPPR
16E18.txt	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta - Homo sapiens	88493	630	IVAERPGTNSTGPAPMAPPR
16E18.txt	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta - Homo sapiens	88493	630	LSPPYSSPQEFQAQDVGRMFKQFNK
16E19.txt	2	ADNP_HUMAN	Activity-dependent neuroprotector - Homo sapiens (Human)	123485	130	TLCPLCLFSILK
16E19.txt	2	ADNP_HUMAN	Activity-dependent neuroprotector - Homo sapiens (Human)	123485	130	HQVIQTVHPVEK
16E19.txt	2	ADNP_HUMAN	Activity-dependent neuroprotector - Homo sapiens (Human)	123485	130	DAPAESVAYHAQNNPPVPPKQPK
16E19.txt	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta - Homo sapiens	88493	189	LSPPYSSPQEFQAQDVGR
16E19.txt	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta - Homo sapiens	88493	189	VFPGSTTEDYNLIVIER
16E19.txt	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta - Homo sapiens	88493	189	DCQLNAHKDHQYQFLEDAVR
16E19.txt	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta - Homo sapiens	88493	189	LSPPYSSPQEFQAQDVGRMFKQFNK
16E20.txt	2	ADNP_HUMAN	Activity-dependent neuroprotector - Homo sapiens (Human)	123485	96	SDIASHFSNKR
16E20.txt	2	ADNP_HUMAN	Activity-dependent neuroprotector - Homo sapiens (Human)	123485	96	HQVIQTVHPVEK
16E20.txt	3	TIF1B_HUMAN	Transcription intermediary factor 1-beta - Homo sapiens	88493	84	LSPPYSSPQEFQAQDVGR
16E20.txt	3	TIF1B_HUMAN	Transcription intermediary factor 1-beta - Homo sapiens	88493	84	VFPGSTTEDYNLIVIER
16E20.txt	3	TIF1B_HUMAN	Transcription intermediary factor 1-beta - Homo sapiens	88493	84	LSPPYSSPQEFQAQDVGRMFKQFNK
16E25.txt	5	CAF1B_HUMAN	Chromatin assembly factor 1 subunit B - Homo sapiens	61454	43	AIVEFLSNLAR
16E25.txt	4	KU70_HUMAN	ATP-dependent DNA helicase 2 subunit 1 - Homo sapiens	69799	58	SDSFENPVLQQHFR
16E28.txt	4	IMA3_HUMAN	Importin subunit alpha-3 - Homo sapiens (Human)	57775	97	LLSSDRNPPIDDLK
16E28.txt	4	IMA3_HUMAN	Importin subunit alpha-3 - Homo sapiens (Human)	57775	97	IEQLQNHENEDIYK
16E28.txt	5	IMA4_HUMAN	Importin subunit alpha-4 - Homo sapiens (Human)	57851	59	LLSSDRNPPIDDLK
16E28.txt	5	IMA4_HUMAN	Importin subunit alpha-4 - Homo sapiens (Human)	57851	59	IEQLQNHENEDIYK
16E35.txt	1	CBX1_HUMAN	Chromobox protein homolog 1 - Homo sapiens (Human)	21405	126	CPQVVISFYEEER
16E35.txt	1	CBX1_HUMAN	Chromobox protein homolog 1 - Homo sapiens (Human)	21405	126	KVEEVELEEEEEYVVEK
16E36.txt	4	CBX1_HUMAN	Chromobox protein homolog 1 - Homo sapiens (Human)	21405	59	CPQVVISFYEEER
16E37.txt	1	CBX1_HUMAN	Chromobox protein homolog 1 - Homo sapiens (Human)	21405	409	GKVEYLLK
16E37.txt	1	CBX1_HUMAN	Chromobox protein homolog 1 - Homo sapiens (Human)	21405	409	CPQVVISFYEEER
16E37.txt	1	CBX1_HUMAN	Chromobox protein homolog 1 - Homo sapiens (Human)	21405	409	IIGATDSSGELMFLMK
16E37.txt	1	CBX1_HUMAN	Chromobox protein homolog 1 - Homo sapiens (Human)	21405	409	IIGATDSSGELMFLMK
16E37.txt	1	CBX1_HUMAN	Chromobox protein homolog 1 - Homo sapiens (Human)	21405	409	LTVHWSYPSEDDDKKDDKN
16E37.txt	1	CBX1_HUMAN	Chromobox protein homolog 1 - Homo sapiens (Human)	21405	409	KKVEEVELEEEEEYVVEK
16E37.txt	1	CBX1_HUMAN	Chromobox protein homolog 1 - Homo sapiens (Human)	21405	409	GFSDENTWEPENLDCPDIAEFLQS

16E46.txt	5	H2B1C_HUMAN	Histone H2B type 1-C/E/F/G/I - Homo sapiens (Human)	13811	147	KESYSVYVYK
16E46.txt	5	H2B1C_HUMAN	Histone H2B type 1-C/E/F/G/I - Homo sapiens (Human)	13811	147	AMGIMNSFVNDIFER
16E46.txt	5	H2B1C_HUMAN	Histone H2B type 1-C/E/F/G/I - Homo sapiens (Human)	13811	147	AMGIMNSFVNDIFER
16E46.txt	5	H2B1C_HUMAN	Histone H2B type 1-C/E/F/G/I - Homo sapiens (Human)	13811	147	AMGIMNSFVNDIFER
16E46.txt	6	H31T_HUMAN	Histone H3.1t - Homo sapiens (Human)	15499	132	STELLIR
16E46.txt	6	H31T_HUMAN	Histone H3.1t - Homo sapiens (Human)	15499	132	EIAQDFK
16E46.txt	6	H31T_HUMAN	Histone H3.1t - Homo sapiens (Human)	15499	132	YRPGTVALR
16E46.txt	6	H31T_HUMAN	Histone H3.1t - Homo sapiens (Human)	15499	132	EIAQDFKTLDR
16E47.txt	1	H2A1B_HUMAN	Histone H2A type 1-B - Homo sapiens (Human)	14127	182	AGLQFPVGR
16E47.txt	1	H2A1B_HUMAN	Histone H2A type 1-B - Homo sapiens (Human)	14127	182	NDEELNKLGR
16E47.txt	1	H2A1B_HUMAN	Histone H2A type 1-B - Homo sapiens (Human)	14127	182	VTIAQGGVLPNIQAVLLPK
16E47.txt	2	H2A1D_HUMAN	Histone H2A type 1-D - Homo sapiens (Human)	14099	172	AGLQFPVGR
16E47.txt	2	H2A1D_HUMAN	Histone H2A type 1-D - Homo sapiens (Human)	14099	172	NDEELNKLGR
16E47.txt	2	H2A1D_HUMAN	Histone H2A type 1-D - Homo sapiens (Human)	14099	172	VTIAQGGVLPNIQAVLLPK
16E47.txt	5	H2AV_HUMAN	Histone H2AV - Homo sapiens (Human)	13501	63	AGLQFPVGR
16E47.txt	5	H2AV_HUMAN	Histone H2AV - Homo sapiens (Human)	13501	63	GDEELDSLK
16E47.txt	3	H2B1C_HUMAN	Histone H2B type 1-C/E/F/G/I - Homo sapiens (Human)	13811	171	LLPGELAK
16E47.txt	3	H2B1C_HUMAN	Histone H2B type 1-C/E/F/G/I - Homo sapiens (Human)	13811	171	KESYSVYVYK
16E47.txt	3	H2B1C_HUMAN	Histone H2B type 1-C/E/F/G/I - Homo sapiens (Human)	13811	171	AMGIMNSFVNDIFER
16E47.txt	3	H2B1C_HUMAN	Histone H2B type 1-C/E/F/G/I - Homo sapiens (Human)	13811	171	AMGIMNSFVNDIFER
16E48.txt	3	CBX1_HUMAN	Chromobox protein homolog 1 - Homo sapiens (Human)	21405	54	KVEEVEEEEEEEYVEK
16E48.txt	2	H2A1B_HUMAN	Histone H2A type 1-B - Homo sapiens (Human)	14127	56	VTIAQGGVLPNIQAVLLPK
16E48.txt	1	H4_HUMAN	Histone H4 - Homo sapiens (Human)	11360	142	TLYGFGG
16E48.txt	1	H4_HUMAN	Histone H4 - Homo sapiens (Human)	11360	142	ISGLIYEETR
16E48.txt	1	H4_HUMAN	Histone H4 - Homo sapiens (Human)	11360	142	TVTAMDVVYALK
16E48.txt	1	H4_HUMAN	Histone H4 - Homo sapiens (Human)	11360	142	TVTAMDVVYALK
16E49.txt	3	H2A1B_HUMAN	Histone H2A type 1-B - Homo sapiens (Human)	14127	62	VTIAQGGVLPNIQAVLLPK
16E49.txt	1	H4_HUMAN	Histone H4 - Homo sapiens (Human)	11360	201	TLYGFGG
16E49.txt	1	H4_HUMAN	Histone H4 - Homo sapiens (Human)	11360	201	TLYGFGG
16E49.txt	1	H4_HUMAN	Histone H4 - Homo sapiens (Human)	11360	201	ISGLIYEETR
16E49.txt	1	H4_HUMAN	Histone H4 - Homo sapiens (Human)	11360	201	TVTAMDVVYALK
16E49.txt	1	H4_HUMAN	Histone H4 - Homo sapiens (Human)	11360	201	DNIQGITPAIR

Tableau S9 : Peptides identifiés dans les 2 purifications CBX3-TAP (n°2 et 4)

fragment	hit	accession	description	masse	score	peptide
04E04	1	CHD4_HUMAN	Chromodomain helicase-DNA-binding protein 4	217853	526	GNFLEIK
04E04	1	CHD4_HUMAN	Chromodomain helicase-DNA-binding protein 4	217853	526	TYEIWHR
04E04	1	CHD4_HUMAN	Chromodomain helicase-DNA-binding protein 4	217853	526	WQDIQNDPR
04E04	1	CHD4_HUMAN	Chromodomain helicase-DNA-binding protein 4	217853	526	ENEFSEFNDAIR
04E04	1	CHD4_HUMAN	Chromodomain helicase-DNA-binding protein 4	217853	526	NDKDKPLPPLLAR
04E04	1	CHD4_HUMAN	Chromodomain helicase-DNA-binding protein 4	217853	526	TVQTAFLYSLYK
04E04	1	CHD4_HUMAN	Chromodomain helicase-DNA-binding protein 4	217853	526	FNAPGAQQFCFLSTR
04E04	1	CHD4_HUMAN	Chromodomain helicase-DNA-binding protein 4	217853	526	FAEVECLAESHQHLSK
04E04	1	CHD4_HUMAN	Chromodomain helicase-DNA-binding protein 4	217853	526	HLCEPGADGAETFDGVPR
04E04	1	CHD4_HUMAN	Chromodomain helicase-DNA-binding protein 4	217853	526	EIIKQEEVDPDYWEK
04E04	1	CHD4_HUMAN	Chromodomain helicase-DNA-binding protein 4	217853	526	IEENSLKEEESIEGEKEVK
04E04	1	CHD4_HUMAN	Chromodomain helicase-DNA-binding protein 4	217853	526	NQDETETELQGMNEYLSSFK
04E06	1	CHD4_HUMAN	Chromodomain helicase-DNA-binding protein 4	217853	111	GPFLVSAPLSTIHWER
04E06	1	CHD4_HUMAN	Chromodomain helicase-DNA-binding protein 4	217853	111	HLCEPGADGAETFDGVPR
04E06	3	SMCA4_HUMAN	Probable global transcription activator SNF2L4	184530	82	HMQAKGVLLTDGSEK
04E06	3	SMCA4_HUMAN	Probable global transcription activator SNF2L4	184530	82	GLQSYVAVAHVTER
04E07	1	POGZ_HUMAN	Pogo transposable element with ZNF domain	155245	712	LQFLFAK
04E07	1	POGZ_HUMAN	Pogo transposable element with ZNF domain	155245	712	IQPLDVCIK
04E07	1	POGZ_HUMAN	Pogo transposable element with ZNF domain	155245	712	LAEWWLTQR
04E07	1	POGZ_HUMAN	Pogo transposable element with ZNF domain	155245	712	HLLCPYCLK
04E07	1	POGZ_HUMAN	Pogo transposable element with ZNF domain	155245	712	FQASQGENLEK
04E07	1	POGZ_HUMAN	Pogo transposable element with ZNF domain	155245	712	SSLYSEVDVHFR
04E07	1	POGZ_HUMAN	Pogo transposable element with ZNF domain	155245	712	AYANHMNNHVPR
04E07	1	POGZ_HUMAN	Pogo transposable element with ZNF domain	155245	712	VPEPENVGDAVQTK
04E07	1	POGZ_HUMAN	Pogo transposable element with ZNF domain	155245	712	EQQLPVNEETLFQK
04E07	1	POGZ_HUMAN	Pogo transposable element with ZNF domain	155245	712	LACTSCTFVTSVGDAMAK
04E07	1	POGZ_HUMAN	Pogo transposable element with ZNF domain	155245	712	LACTSCTFVTSVGDAMAK
04E07	1	POGZ_HUMAN	Pogo transposable element with ZNF domain	155245	712	VVLFALCCNTEQAAEHFR
04E07	1	POGZ_HUMAN	Pogo transposable element with ZNF domain	155245	712	DTHKPGEMPYVCVCQYR
04E07	1	POGZ_HUMAN	Pogo transposable element with ZNF domain	155245	712	HHVELDQVQNGEVDGHTICQHCYR
04E09	1	ADNP_HUMAN	Activity-dependent neuroprotector	123485	1500	FFSAYK
04E09	1	ADNP_HUMAN	Activity-dependent neuroprotector	123485	1500	VEGFWSK
04E09	1	ADNP_HUMAN	Activity-dependent neuroprotector	123485	1500	VPAVANYIMK
04E09	1	ADNP_HUMAN	Activity-dependent neuroprotector	123485	1500	VPAVANYIMK
04E09	1	ADNP_HUMAN	Activity-dependent neuroprotector	123485	1500	SLPSQQMVR
04E09	1	ADNP_HUMAN	Activity-dependent neuroprotector	123485	1500	SLNGAVPLGSNAR
04E09	1	ADNP_HUMAN	Activity-dependent neuroprotector	123485	1500	MVHIDEEMGPK
04E09	1	ADNP_HUMAN	Activity-dependent neuroprotector	123485	1500	LNQSPSLAPVKR
04E09	1	ADNP_HUMAN	Activity-dependent neuroprotector	123485	1500	ILSDIGLEYCK
04E09	1	ADNP_HUMAN	Activity-dependent neuroprotector	123485	1500	YFNKQPYPTR
04E09	1	ADNP_HUMAN	Activity-dependent neuroprotector	123485	1500	YKPGVLLGFNMK
04E09	1	ADNP_HUMAN	Activity-dependent neuroprotector	123485	1500	TYEQMEFPLK
04E09	1	ADNP_HUMAN	Activity-dependent neuroprotector	123485	1500	EHFQHVAAPIAK
04E09	1	ADNP_HUMAN	Activity-dependent neuroprotector	123485	1500	EHFQHVAAPIAK
04E09	1	ADNP_HUMAN	Activity-dependent neuroprotector	123485	1500	QADSVEQAVYYCK
04E09	1	ADNP_HUMAN	Activity-dependent neuroprotector	123485	1500	SKPLMLIAPKQDK
04E09	1	ADNP_HUMAN	Activity-dependent neuroprotector	123485	1500	CTYRDPLYEIVR
04E09	1	ADNP_HUMAN	Activity-dependent neuroprotector	123485	1500	TKPFCACSACPFSSK
04E09	1	ADNP_HUMAN	Activity-dependent neuroprotector	123485	1500	QADSVEQAVYYCKK
04E09	1	ADNP_HUMAN	Activity-dependent neuroprotector	123485	1500	SSKPAAKKATMQGDR
04E09	1	ADNP_HUMAN	Activity-dependent neuroprotector	123485	1500	YSLQSANASSLSSGQLK

04E09	1	ADNP_HUMAN	Activity-dependent neuroprotector	123485	1500	ILLNCPYCTFNADKK
04E09	1	ADNP_HUMAN	Activity-dependent neuroprotector	123485	1500	ILLNCPYCTFNADKK
04E09	1	ADNP_HUMAN	Activity-dependent neuroprotector	123485	1500	NTTWEDVGLWDPSTLK
04E09	1	ADNP_HUMAN	Activity-dependent neuroprotector	123485	1500	IFHAPNASAPSSSLSTFK
04E09	1	ADNP_HUMAN	Activity-dependent neuroprotector	123485	1500	IFHAPNASAPSSSLSTFK
04E09	1	ADNP_HUMAN	Activity-dependent neuroprotector	123485	1500	LGLGGNAPVSIPOQSSQSVK
04E09	1	ADNP_HUMAN	Activity-dependent neuroprotector	123485	1500	IGYQVTAMIGHTNVVVPR
04E09	1	ADNP_HUMAN	Activity-dependent neuroprotector	123485	1500	YFNKQPYPTRREIEK
04E09	1	ADNP_HUMAN	Activity-dependent neuroprotector	123485	1500	IGYQVTAMIGHTNVVVPR
04E09	1	ADNP_HUMAN	Activity-dependent neuroprotector	123485	1500	IGYQVTAMIGHTNVVVPR
04E09	1	ADNP_HUMAN	Activity-dependent neuroprotector	123485	1500	EHIEDFKQFEPNDFYLK
04E09	1	ADNP_HUMAN	Activity-dependent neuroprotector	123485	1500	ILSDIGLEYCKEHIEDFK
04E09	1	ADNP_HUMAN	Activity-dependent neuroprotector	123485	1500	VLGQSSSKPAAAAATGPPPGNTSST QK
04E09	1	ADNP_HUMAN	Activity-dependent neuroprotector	123485	1500	VLGQSSSKPAAAAATGPPPGNTSST QK
04E09	1	ADNP_HUMAN	Activity-dependent neuroprotector	123485	1500	DAPAESVAYHAQNNPPVPPKQPK
04E09	1	ADNP_HUMAN	Activity-dependent neuroprotector	123485	1500	YLPTDTRLNHMLIHGLSCPYCR
04E09	1	ADNP_HUMAN	Activity-dependent neuroprotector	123485	1500	LSIPKPNLNSTGVNMMSSVHLQQNN YGVK
04E09	1	ADNP_HUMAN	Activity-dependent neuroprotector	123485	1500	LSIPKPNLNSTGVNMMSSVHLQQNN YGVK
04E09	1	ADNP_HUMAN	Activity-dependent neuroprotector	123485	1500	ILSDIGLEYCKEHIEDFKQFEPNDFYL K
04E11	4	ADNP_HUMAN	Activity-dependent neuroprotector	123485	46	LGLGGNAPVSIPOQSSQSVK
04E11	2	SUHW4_HUMAN	Suppressor of hairy wing homolog 4	109229	107	NIVSCDSNIGADKVEK
04E11	2	SUHW4_HUMAN	Suppressor of hairy wing homolog 4	109229	107	SEESITVSDKENETCLADQETGSK
04E13	3	K1802_HUMAN	Zinc finger protein KIAA1802	89043	147	DNQESSDAELSSSEYIK
04E13	3	K1802_HUMAN	Zinc finger protein KIAA1802	89043	147	NVLQFTTEEKEAFISEEIAK
04E13	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	183	VFPGSTTDEDYNLIVIER
04E13	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	183	IVAERPGTNSTGPAPMAPPR
04E13	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	183	DCQLNAHKDHQYQFLEDAVR
04E14	4	ADNP_HUMAN	Activity-dependent neuroprotector	123485	59	DAPAESVAYHAQNNPPVPPKQPK
04E14	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	692	LLASLVK
04E14	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	692	KLLASLVK
04E14	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	692	MAILQIMK
04E14	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	692	MAILQIMK
04E14	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	692	SGEGEVSGLMR
04E14	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	692	ADVQSIIGLQR
04E14	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	692	KLIYFQLHR
04E14	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	692	FQWDLNAWTK
04E14	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	692	MIVDPVEPHGEMK
04E14	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	692	MIVDPVEPHGEMK
04E14	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	692	LDLDTADSOPPVPFK
04E14	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	692	LTEDKADVQSIIGLQR
04E14	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	692	VLVNDQAKVTEGQQR
04E14	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	692	LSPPYSSPQFAQDVGR
04E14	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	692	VFPGSTTDEDYNLIVIER
04E14	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	692	IVAERPGTNSTGPAPMAPPR
04E14	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	692	IVAERPGTNSTGPAPMAPPR
04E14	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	692	IVAERPGTNSTGPAPMAPPR
04E14	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	692	IVAERPGTNSTGPAPMAPPRAPGPL SK
04E14	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	692	LSPPYSSPQFAQDVGRMFKQFNK
04E15	2	ADNP_HUMAN	Activity-dependent neuroprotector	123485	52	VLGQSSSKPAAAAATGPPPGNTSST QK
04E15	2	ADNP_HUMAN	Activity-dependent neuroprotector	123485	52	DAPAESVAYHAQNNPPVPPKQPK
04E15	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	769	SGEGEVSGLMR
04E15	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	769	KLIYFQLHR
04E15	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	769	FQWDLNAWTK
04E15	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	769	MIVDPVEPHGEMK
04E15	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	769	MIVDPVEPHGEMK
04E15	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	769	MIVDPVEPHGEMK
04E15	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	769	DHQYQFLEDAVR
04E15	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	769	LDLDTADSOPPVPFK
04E15	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	769	LTEDKADVQSIIGLQR
04E15	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	769	VLVNDQAKVTEGQQR
04E15	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	769	LSPPYSSPQFAQDVGR
04E15	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	769	VFPGSTTDEDYNLIVIER
04E15	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	769	IVAERPGTNSTGPAPMAPPR
04E15	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	769	IVAERPGTNSTGPAPMAPPR
04E15	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	769	IVAERPGTNSTGPAPMAPPR
04E15	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	769	HEPLVLFCECSDTLTCR
04E15	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	769	DCQLNAHKDHQYQFLEDAVR
04E15	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	769	DCQLNAHKDHQYQFLEDAVR
04E16	5	ADNP_HUMAN	Activity-dependent neuroprotector	123485	47	DAPAESVAYHAQNNPPVPPKQPK
04E16	3	CA103_HUMAN	Uncharacterized protein C1orf103	84516	201	VTSVGTGNFSSSVSK
04E16	3	CA103_HUMAN	Uncharacterized protein C1orf103	84516	201	DTLQTVSSSPVTEISR
04E16	3	CA103_HUMAN	Uncharacterized protein C1orf103	84516	201	GTLNSNAAYEQSHFFNK
04E16	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	617	ADVQSIIGLQR
04E16	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	617	MIVDPVEPHGEMK
04E16	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	617	MIVDPVEPHGEMK
04E16	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	617	MIVDPVEPHGEMK
04E16	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	617	DHQYQFLEDAVR
04E16	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	617	LDLDTADSOPPVPFK
04E16	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	617	LDLDTADSOPPVPFK
04E16	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	617	LTEDKADVQSIIGLQR
04E16	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	617	LSPPYSSPQFAQDVGR

04E16	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	617	VFPGSTTEDYNLIVIER
04E16	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	617	IVAERPSTNSTGPAPMAPPR
04E16	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	617	IVAERPSTNSTGPAPMAPPR
04E16	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	617	IVAERPSTNSTGPAPMAPPR
04E16	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	617	DCQLNAHKDHQYQFLEDAVR
04E16	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	617	DCQLNAHKDHQYQFLEDAVR
04E16	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	617	DCQLNAHKDHQYQFLEDAVR
04E16	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	617	DCQLNAHKDHQYQFLEDAVR
04E16	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	617	DCQLNAHKDHQYQFLEDAVR
04E17	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	268	MIVDPVEPHGEMK
04E17	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	268	VFPGSTTEDYNLIVIER
04E17	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	268	IVAERPSTNSTGPAPMAPPR
04E17	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	268	DCQLNAHKDHQYQFLEDAVR
04E18	3	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	70	DCQLNAHKDHQYQFLEDAVR
04E18	3	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	70	DCQLNAHKDHQYQFLEDAVR
04E18	3	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	70	DCQLNAHKDHQYQFLEDAVR
04E19	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	81	STGPAK
04E19	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	81	VFPGSTTEDYNLIVIER
04E21	7	KU70_HUMAN	ATP-dependent DNA helicase 2 subunit 1	69799	121	SDSFENPVLQQHFR
04E21	7	KU70_HUMAN	ATP-dependent DNA helicase 2 subunit 1	69799	121	TEGDEEAEEQENLEASGDYK
04E21	4	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	144	VFPGSTTEDYNLIVIER
04E21	4	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	144	DCQLNAHKDHQYQFLEDAVR
04E21	2	ZNF8_HUMAN	Zinc finger protein 8 (Zinc finger protein HF.18)	64930	222	AFSQNSSLDHQR
04E21	2	ZNF8_HUMAN	Zinc finger protein 8 (Zinc finger protein HF.18)	64930	222	EGFPTDAPYPTTLGK
04E21	2	ZNF8_HUMAN	Zinc finger protein 8 (Zinc finger protein HF.18)	64930	222	KEEGPEEESHVTGR
04E22	4	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	53	VFPGSTTEDYNLIVIER
04E25	3	IMA4_HUMAN	Importin alpha-4 subunit	57851	70	NFKNKGDRLETMR
04E25	3	IMA4_HUMAN	Importin alpha-4 subunit	57851	70	IEQLQNHENEDIYK
04E25	4	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	50	DCQLNAHKDHQYQFLEDAVR
04E26	4	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	64	MIVDPVEPHGEMK
04E26	4	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	64	IVAERPSTNSTGPAPMAPPR
04E30	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	79	VEEAPEEFVVEK
04E30	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	79	KVEEAPEEFVVEKVLDRR
04E31	2	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	241	LTWHSCPEDEAQ
04E31	2	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	241	WKDSDEADLVLAKE
04E31	2	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	241	VEEAPEEFVVEK
04E31	2	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	241	KVEEAPEEFVVEK
04E32	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	306	VVNGKVEYFLK
04E32	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	306	LTWHSCPEDEAQ
04E32	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	306	WKDSDEADLVLAKE
04E32	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	306	VEEAPEEFVVEK
04E32	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	306	KVEEAPEEFVVEK
04E32	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	306	KVEEAPEEFVVEKVLDRR
04E33	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	283	VEYFLK
04E33	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	283	VVNGKVEYFLK
04E33	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	283	LTWHSCPEDEAQ
04E33	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	283	WKDSDEADLVLAKE
04E33	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	283	VEEAPEEFVVEK
04E33	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	283	KVEEAPEEFVVEK
04E34	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	532	VEYFLK
04E34	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	532	DSDEADLVLAKE
04E34	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	532	LTWHSCPEDEAQ
04E34	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	532	WKDSDEADLVLAKE
04E34	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	532	VEEAPEEFVVEK
04E34	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	532	KVEEAPEEFVVEK
04E34	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	532	IIGATDSSGELMFLMK
04E34	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	532	IIGATDSSGELMFLMK
04E34	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	532	SKKVEEAPEEFVVEK
04E34	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	532	GFTDADNTWEPEENLDCPELIEAFLNSQK
04E35	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	435	VEYFLK
04E35	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	435	DSDEADLVLAKE
04E35	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	435	VVNGKVEYFLK
04E35	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	435	VVNGKVEYFLK
04E35	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	435	WKDSDEADLVLAKE
04E35	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	435	VEEAPEEFVVEK
04E35	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	435	VEEAPEEFVVEK
04E35	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	435	KVEEAPEEFVVEK
04E35	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	435	IIGATDSSGELMFLMK
04E35	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	435	IIGATDSSGELMFLMK
04E35	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	435	IIGATDSSGELMFLMK
04E35	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	435	GFTDADNTWEPEENLDCPELIEAFLNSQK
04E35	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	435	DSDEADLVLAKE
04E36	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	455	VVNGKVEYFLK
04E36	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	455	VVNGKVEYFLK
04E36	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	455	RVVNGKVEYFLK
04E36	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	455	LTWHSCPEDEAQ
04E36	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	455	WKDSDEADLVLAKE
04E36	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	455	CPQIVIAFYEER
04E36	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	455	VEEAPEEFVVEK
04E36	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	455	MASKNTTLQKMGKK
04E36	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	455	KVEEAPEEFVVEK
04E36	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	455	IIGATDSSGELMFLMK
04E36	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	455	IIGATDSSGELMFLMK
04E36	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	455	IIGATDSSGELMFLMK
04E36	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	455	IIGATDSSGELMFLMK
04E36	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	455	SKKVEEAPEEFVVEK
04E36	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	455	GFTDADNTWEPEENLDCPELIEAFLNSQK
04E36	3	CBX5_HUMAN	Chromobox protein homolog 5 (HP1 alpha)	22211	98	WKDTDEADLVLAKE
04E36	3	CBX5_HUMAN	Chromobox protein homolog 5 (HP1 alpha)	22211	98	WKDTDEADLVLAKE
04E36	3	CBX5_HUMAN	Chromobox protein homolog 5 (HP1 alpha)	22211	98	CPQIVIAFYEER

04E36	3	CBX5_HUMAN	Chromobox protein homolog 5 (HP1 alpha)	22211	98	IIGATDSCGDLMFLMK
04E36	3	CBX5_HUMAN	Chromobox protein homolog 5 (HP1 alpha)	22211	98	IIGATDSCGDLMFLMK
04E36	3	CBX5_HUMAN	Chromobox protein homolog 5 (HP1 alpha)	22211	98	IIGATDSCGDLMFLMK
04E36	3	CBX5_HUMAN	Chromobox protein homolog 5 (HP1 alpha)	22211	98	IIGATDSCGDLMFLMK
04E37	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	482	VEYFLK
04E37	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	482	DSDEADLVLAQ
04E37	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	482	LTWHSCPEDEAQ
04E37	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	482	IIGATDSSGDLVLAQ
04E37	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	482	CPQIVIAFYEER
04E37	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	482	VEEAPEPEEFVVEK
04E37	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	482	KVEEAPEPEEFVVEK
04E37	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	482	IIGATDSSGELMFLMK
04E38	2	CBX1_HUMAN	Chromobox protein homolog 1 (HP1 beta)	21405	295	GKVEYLLK
04E38	2	CBX1_HUMAN	Chromobox protein homolog 1 (HP1 beta)	21405	295	CPQVVISFYEER
04E38	2	CBX1_HUMAN	Chromobox protein homolog 1 (HP1 beta)	21405	295	IIGATDSSGELMFLMK
04E38	2	CBX1_HUMAN	Chromobox protein homolog 1 (HP1 beta)	21405	295	IIGATDSSGELMFLMK
04E38	2	CBX1_HUMAN	Chromobox protein homolog 1 (HP1 beta)	21405	295	KVEEVLEEEEEVYVEK
04E38	2	CBX1_HUMAN	Chromobox protein homolog 1 (HP1 beta)	21405	295	KKVEEVLEEEEEVYVEK
04E38	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	430	VEYFLK
04E38	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	430	LTWHSCPEDEAQ
04E38	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	430	WKDSDEADLVLAQ
04E38	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	430	VEEAPEPEEFVVEK
04E38	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	430	VEEAPEPEEFVVEK
04E38	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	430	KVEEAPEPEEFVVEK
04E38	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	430	IIGATDSSGELMFLMK
04E38	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	430	IIGATDSSGELMFLMK
04E38	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	430	SKKVEEAPEPEEFVVEK
04E38	4	CBX5_HUMAN	Chromobox protein homolog 5 (HP1 alpha)	22211	166	GQVEYLLK
04E38	4	CBX5_HUMAN	Chromobox protein homolog 5 (HP1 alpha)	22211	166	WKDTDEADLVLAQ
04E38	4	CBX5_HUMAN	Chromobox protein homolog 5 (HP1 alpha)	22211	166	WKDTDEADLVLAQ
04E38	4	CBX5_HUMAN	Chromobox protein homolog 5 (HP1 alpha)	22211	166	IIGATDSCGDLMFLMK
04E38	4	CBX5_HUMAN	Chromobox protein homolog 5 (HP1 alpha)	22211	166	IIGATDSCGDLMFLMK
04E38	4	CBX5_HUMAN	Chromobox protein homolog 5 (HP1 alpha)	22211	166	LTWHAYPEDAENKEK
04E39	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	451	VEYFLK
04E39	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	451	DSDEADLVLAQ
04E39	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	451	VVNGKVEYFLK
04E39	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	451	LTWHSCPEDEAQ
04E39	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	451	WKDSDEADLVLAQ
04E39	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	451	VEEAPEPEEFVVEK
04E39	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	451	KVEEAPEPEEFVVEK
04E39	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	451	KVEEAPEPEEFVVEK
04E39	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	451	IIGATDSSGELMFLMK
04E39	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	451	GFTDADNTWEPEENLDCPELIEAFLNSQK
04E40	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	455	VEYFLK
04E40	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	455	VVNGKVEYFLK
04E40	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	455	LTWHSCPEDEAQ
04E40	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	455	WKDSDEADLVLAQ
04E40	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	455	VEEAPEPEEFVVEK
04E40	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	455	KVEEAPEPEEFVVEK
04E40	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	455	IIGATDSSGELMFLMK
04E40	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	455	SKKVEEAPEPEEFVVEK
04E40	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	455	KVEEAPEPEEFVVEKVLDRR
04E41	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	226	VEYFLK
04E41	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	226	VVNGKVEYFLK
04E41	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	226	LTWHSCPEDEAQ
04E41	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	226	WKDSDEADLVLAQ
04E41	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	226	VEEAPEPEEFVVEK
04E41	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	226	KVEEAPEPEEFVVEKVLDRR
04E41	3	H31T_HUMAN	Histone H3.1t (H3t)	15499	67	STELLIR
04E41	3	H31T_HUMAN	Histone H3.1t (H3t)	15499	67	EIAQDFKTDLR
04E42	2	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	146	VVNGKVEYFLK
04E42	2	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	146	LTWHSCPEDEAQ
04E42	2	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	146	VEEAPEPEEFVVEK
04E43	2	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	254	LTWHSCPEDEAQ
04E43	2	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	254	WKDSDEADLVLAQ
04E43	2	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	254	VEEAPEPEEFVVEK
04E43	2	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	254	KVEEAPEPEEFVVEK
04E43	2	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	254	KVEEAPEPEEFVVEKVLDRR
04E43	3	H2A1D_HUMAN	Histone H2A type 1-D	14099	115	NDEELNKLKLGK
04E43	3	H2A1D_HUMAN	Histone H2A type 1-D	14099	115	VVNGKVEYFLK
04E44	2	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	101	LTWHSCPEDEAQ
04E44	2	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	101	KVEEAPEPEEFVVEK
04E44	4	H4_HUMAN	Histone H4	11360	69	DNIQGITKPAIR
04E46	4	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	63	LTWHSCPEDEAQ
merge02E04.txt	2	ADNP_HUMAN	Activity-dependent neuroprotector	123485	805	FFSAYK
merge02E04.txt	2	ADNP_HUMAN	Activity-dependent neuroprotector	123485	805	VEGFWSK
merge02E04.txt	2	ADNP_HUMAN	Activity-dependent neuroprotector	123485	805	SLPSQQMVNR
merge02E04.txt	2	ADNP_HUMAN	Activity-dependent neuroprotector	123485	805	SLNGAVPLGSNAR
merge02E04.txt	2	ADNP_HUMAN	Activity-dependent neuroprotector	123485	805	LNQSPSLAPVKR
merge02E04.txt	2	ADNP_HUMAN	Activity-dependent neuroprotector	123485	805	TLCPCLCFSLK
merge02E04.txt	2	ADNP_HUMAN	Activity-dependent neuroprotector	123485	805	YKPGVLLGFNMK
merge02E04.txt	2	ADNP_HUMAN	Activity-dependent neuroprotector	123485	805	ISNDNPEEHVLK
merge02E04.txt	2	ADNP_HUMAN	Activity-dependent neuroprotector	123485	805	QADSVEQAVYYCK
merge02E04.txt	2	ADNP_HUMAN	Activity-dependent neuroprotector	123485	805	SKPLMLIAPKQDK
merge02E04.txt	2	ADNP_HUMAN	Activity-dependent neuroprotector	123485	805	TKPFCCSACPFSSK
merge02E04.txt	2	ADNP_HUMAN	Activity-dependent neuroprotector	123485	805	YSLQSANASSLSSGQLK

merge02E04.txt	2	ADNP_HUMAN	Activity-dependent neuroprotector	123485	805	VIPDASESEEKLDQK
merge02E04.txt	2	ADNP_HUMAN	Activity-dependent neuroprotector	123485	805	NTTWEDVGLWDPSTLK
merge02E04.txt	2	ADNP_HUMAN	Activity-dependent neuroprotector	123485	805	IFHAPNASAPSSSLSTFK
merge02E04.txt	2	ADNP_HUMAN	Activity-dependent neuroprotector	123485	805	IFHAPNASAPSSSLSTFK
merge02E04.txt	2	ADNP_HUMAN	Activity-dependent neuroprotector	123485	805	LGLGGNAPVSIPOQSSQSVK
merge02E04.txt	2	ADNP_HUMAN	Activity-dependent neuroprotector	123485	805	IFHAPNASAPSSSLSTFKDK
merge02E04.txt	2	ADNP_HUMAN	Activity-dependent neuroprotector	123485	805	VLGQSSSKPAAAAATGPPPGNTSST QK
merge02E04.txt	2	ADNP_HUMAN	Activity-dependent neuroprotector	123485	805	VLGQSSSKPAAAAATGPPPGNTSST QK
merge02E04.txt	2	ADNP_HUMAN	Activity-dependent neuroprotector	123485	805	DAPAESVAYHAQNPPVPPKPQPK
merge02E05.txt	4	ADNP_HUMAN	Activity-dependent neuroprotector	123485	301	ISNDNPEEHVVK
merge02E05.txt	4	ADNP_HUMAN	Activity-dependent neuroprotector	123485	301	YSLQSANASSLSGQLK
merge02E05.txt	4	ADNP_HUMAN	Activity-dependent neuroprotector	123485	301	LGLGGNAPVSIPOQSSQSVK
merge02E05.txt	4	ADNP_HUMAN	Activity-dependent neuroprotector	123485	301	VLGQSSSKPAAAAATGPPPGNTSST QK
merge02E05.txt	4	ADNP_HUMAN	Activity-dependent neuroprotector	123485	301	DAPAESVAYHAQNPPVPPKPQPK
merge02E05.txt	7	GTF2L_HUMAN	General transcription factor II-I	112346	169	RPSTFGIPR
merge02E05.txt	7	GTF2L_HUMAN	General transcription factor II-I	112346	169	RILDSAEIFK
merge02E05.txt	7	GTF2L_HUMAN	General transcription factor II-I	112346	169	EDLQDQKASGVK
merge02E05.txt	7	GTF2L_HUMAN	General transcription factor II-I	112346	169	RPELLTHSTTEVTQPR
merge02E05.txt	7	GTF2L_HUMAN	General transcription factor II-I	112346	169	VPFALFESFPEDFYVEGLPEGVVFR
merge02E05.txt	7	GTF2L_HUMAN	General transcription factor II-I	112346	169	VPPYVFESNPEFLYVEGLPEGIPFR
merge02E05.txt	9	K1802_HUMAN	Zinc finger protein KIAA1802	89043	104	DNQESSDAELSSSEYIK
merge02E05.txt	8	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	130	VFPSTTDEYNLIVIER
merge02E05.txt	8	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	130	IVAERPSTNSTGPAPMAPPR
merge02E06.txt	5	ADNP_HUMAN	Activity-dependent neuroprotector	123485	75	ISNDNPEEHVVK
merge02E06.txt	5	ADNP_HUMAN	Activity-dependent neuroprotector	123485	75	VLGQSSSKPAAAAATGPPPGNTSST QK
merge02E06.txt	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	629	LLASLVK
merge02E06.txt	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	629	SGEGEVSLMR
merge02E06.txt	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	629	KLYFGLHR
merge02E06.txt	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	629	FQWDLNAWTK
merge02E06.txt	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	629	MIVDPVEPHGEMK
merge02E06.txt	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	629	MIVDPVEPHGEMK
merge02E06.txt	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	629	DHQYQFLEDAVR
merge02E06.txt	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	629	LDLDTADSQPPVFK
merge02E06.txt	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	629	LTEDKADVQSIIGLQR
merge02E06.txt	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	629	LTEDKADVQSIIGLQR
merge02E06.txt	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	629	EEDGSLDLGADSTGVVAK
merge02E06.txt	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	629	VFPSTTDEYNLIVIER
merge02E06.txt	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	629	IVAERPSTNSTGPAPMAPPR
merge02E06.txt	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	629	IVAERPSTNSTGPAPMAPPR
merge02E06.txt	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	629	IVAERPSTNSTGPAPMAPPR
merge02E07.txt	6	ADNP_HUMAN	Activity-dependent neuroprotector	123485	129	ISNDNPEEHVVK
merge02E07.txt	6	ADNP_HUMAN	Activity-dependent neuroprotector	123485	129	VLGQSSSKPAAAAATGPPPGNTSST QK
merge02E07.txt	6	ADNP_HUMAN	Activity-dependent neuroprotector	123485	129	DAPAESVAYHAQNPPVPPKPQPK
merge02E07.txt	3	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	481	ADVQSIIGLQR
merge02E07.txt	3	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	481	FQWDLNAWTK
merge02E07.txt	3	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	481	LDLDTADSQPPVFK
merge02E07.txt	3	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	481	LTEDKADVQSIIGLQR
merge02E07.txt	3	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	481	LTEDKADVQSIIGLQR
merge02E07.txt	3	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	481	EEDGSLDLGADSTGVVAK
merge02E07.txt	3	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	481	VFPSTTDEYNLIVIER
merge02E07.txt	3	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	481	IVAERPSTNSTGPAPMAPPR
merge02E07.txt	3	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	481	DCQLNAHKDHQYQFLEDAVR
merge02E08.txt	8	ADNP_HUMAN	Activity-dependent neuroprotector	123485	80	VLGQSSSKPAAAAATGPPPGNTSST QK
merge02E08.txt	8	ADNP_HUMAN	Activity-dependent neuroprotector	123485	80	DAPAESVAYHAQNPPVPPKPQPK
merge02E08.txt	3	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	410	LDLDTADSQPPVFK
merge02E08.txt	3	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	410	LTEDKADVQSIIGLQR
merge02E08.txt	3	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	410	EEDGSLDLGADSTGVVAK
merge02E08.txt	3	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	410	VFPSTTDEYNLIVIER
merge02E08.txt	3	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	410	IVAERPSTNSTGPAPMAPPR
merge02E08.txt	3	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	410	DCQLNAHKDHQYQFLEDAVR
merge02E09.txt	11	IMA4_HUMAN	Importin alpha-4 subunit	57851	44	IEQLQNHENEDIYK
merge02E09.txt	9	K0774_HUMAN	Uncharacterized protein KIAA0774	150032	48	VAESTTK
merge02E09.txt	9	K0774_HUMAN	Uncharacterized protein KIAA0774	150032	48	RDAEIPR
merge02E09.txt	9	K0774_HUMAN	Uncharacterized protein KIAA0774	150032	48	HVPKDKLAK
merge02E09.txt	9	K0774_HUMAN	Uncharacterized protein KIAA0774	150032	48	SGLRPPGYSRPAAK
merge02E09.txt	8	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	82	IVAERPSTNSTGPAPMAPPR
merge02E09.txt	8	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	82	DCQLNAHKDHQYQFLEDAVR
merge02E10.txt	8	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	76	KVEEAPEEFVVEK
merge02E11.txt	6	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	239	VEYFLK
merge02E11.txt	6	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	239	LTWHSCPEDEAQ
merge02E11.txt	6	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	239	WKDSDEADLVLAQ
merge02E11.txt	6	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	239	KVEEAPEEFVVEK
merge02E11.txt	3	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	337	SGEGEVSLMR
merge02E11.txt	3	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	337	ADVQSIIGLQR
merge02E11.txt	3	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	337	LSPANQRKCR
merge02E11.txt	3	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	337	LDLDTADSQPPVFK
merge02E11.txt	3	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	337	LTEDKADVQSIIGLQR
merge02E11.txt	3	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	337	EEDGSLDLGADSTGVVAK
merge02E12.txt	7	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	215	LTWHSCPEDEAQ
merge02E12.txt	7	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	215	WKDSDEADLVLAQ
merge02E12.txt	7	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	215	KVEEAPEEFVVEK
merge02E12.txt	7	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	215	KVEEAPEEFVVEK
merge02E13.txt	4	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	300	VEYFLK
merge02E13.txt	4	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	300	LTWHSCPEDEAQ

merge02E13.txt	4	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	300	WKDSDEADLVLAKE
merge02E13.txt	4	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	300	VEEAPEPEEFVVEK
merge02E13.txt	4	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	300	KVEEAPEPEEFVVEK
merge02E13.txt	4	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	300	KVEEAPEPEEFVVEK
merge02E13.txt	10	DCK_HUMAN	Deoxycytidine kinase	30499	48	EFLSTL
merge02E13.txt	10	DCK_HUMAN	Deoxycytidine kinase	30499	48	DKYESLVEK
merge02E13.txt	10	DCK_HUMAN	Deoxycytidine kinase	30499	48	LKDAEKPLVFFERSVYSDR
merge02E13.txt	6	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	96	LTEDKADVQSIIGLQR
merge02E13.txt	6	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	96	EEDGSLSLDGADSTGVVAK
merge02E14.txt	2	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	486	VEYFLK
merge02E14.txt	2	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	486	LTWHSCPEDEAQ
merge02E14.txt	2	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	486	WKDSDEADLVLAKE
merge02E14.txt	2	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	486	VEEAPEPEEFVVEK
merge02E14.txt	2	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	486	KVEEAPEPEEFVVEK
merge02E14.txt	2	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	486	KVEEAPEPEEFVVEK
merge02E14.txt	2	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	486	IIGATDSSGELMFLMK
merge02E14.txt	2	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	486	IIGATDSSGELMFLMK
merge02E14.txt	2	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	486	KVEEAPEPEEFVVEKVLDRR
merge02E14.txt	2	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	486	GFTDADNTWPEENLDCPELIEAFLNSQK
merge02E14.txt	6	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	133	ADVQSIIGLQR
merge02E14.txt	6	TIF1B_HUMAN	Transcription intermediary factor 1-beta (TIF1-beta)	88493	133	LTEDKADVQSIIGLQR
merge02E15.txt	3	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	546	VEYFLK
merge02E15.txt	3	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	546	DSDEADLVLAKE
merge02E15.txt	3	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	546	RVVNGKVEYFLK
merge02E15.txt	3	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	546	LTWHSCPEDEAQ
merge02E15.txt	3	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	546	WKDSDEADLVLAKE
merge02E15.txt	3	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	546	WKDSDEADLVLAKE
merge02E15.txt	3	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	546	CPQIVIAFYEER
merge02E15.txt	3	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	546	VEEAPEPEEFVVEK
merge02E15.txt	3	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	546	KVEEAPEPEEFVVEK
merge02E15.txt	3	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	546	KVEEAPEPEEFVVEK
merge02E15.txt	3	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	546	IIGATDSSGELMFLMK
merge02E15.txt	3	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	546	IIGATDSSGELMFLMK
merge02E15.txt	3	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	546	KVEEAPEPEEFVVEKVLDRR
merge02E16.txt	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	727	VEYFLK
merge02E16.txt	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	727	DSDEADLVLAKE
merge02E16.txt	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	727	DSDEADLVLAKE
merge02E16.txt	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	727	VVNGKVEYFLK
merge02E16.txt	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	727	MASKNTTLQKMGK
merge02E16.txt	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	727	LTWHSCPEDEAQ
merge02E16.txt	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	727	LTWHSCPEDEAQ
merge02E16.txt	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	727	WKDSDEADLVLAKE
merge02E16.txt	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	727	VEEAPEPEEFVVEK
merge02E16.txt	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	727	VEEAPEPEEFVVEK
merge02E16.txt	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	727	IIGATDSSGELMFLMK
merge02E16.txt	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	727	IIGATDSSGELMFLMK
merge02E16.txt	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	727	DSDEADLVLAKEANMK
merge02E16.txt	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	727	SKKVEEAPEPEEFVVEK
merge02E16.txt	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	727	SKKVEEAPEPEEFVVEK
merge02E16.txt	1	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	727	GFTDADNTWPEENLDCPELIEAFLNSQK
merge02E16.txt	8	CBX5_HUMAN	Chromobox protein homolog 5 (HP1 alpha)	22211	63	WKDTDEADLVLAKE
merge02E16.txt	8	CBX5_HUMAN	Chromobox protein homolog 5 (HP1 alpha)	22211	63	WKDTDEADLVLAKE
merge02E16.txt	8	CBX5_HUMAN	Chromobox protein homolog 5 (HP1 alpha)	22211	63	IIGATDSCGDLMFLMK
merge02E16.txt	8	CBX5_HUMAN	Chromobox protein homolog 5 (HP1 alpha)	22211	63	IIGATDSCGDLMFLMK
merge02E16.txt	8	CBX5_HUMAN	Chromobox protein homolog 5 (HP1 alpha)	22211	63	IIGATDSCGDLMFLMK
merge02E17.txt	4	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	242	VEYFLK
merge02E17.txt	4	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	242	LTWHSCPEDEAQ
merge02E17.txt	4	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	242	VEEAPEPEEFVVEK
merge02E17.txt	4	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	242	KVEEAPEPEEFVVEK
merge02E17.txt	4	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	242	KVEEAPEPEEFVVEK
merge02E18.txt	12	CBX1_HUMAN	Chromobox protein homolog 1 (HP1 beta)	21405	45	GKVEYLLK
merge02E18.txt	12	CBX1_HUMAN	Chromobox protein homolog 1 (HP1 beta)	21405	45	NSDEADLVPAK
merge02E18.txt	5	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	201	LTWHSCPEDEAQ
merge02E18.txt	5	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	201	VEEAPEPEEFVVEK
merge02E18.txt	5	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	201	KVEEAPEPEEFVVEK
merge02E18.txt	6	CBX5_HUMAN	Chromobox protein homolog 5 (HP1 alpha)	22211	146	GQVEYLLK
merge02E18.txt	6	CBX5_HUMAN	Chromobox protein homolog 5 (HP1 alpha)	22211	146	WKDTDEADLVLAKE
merge02E18.txt	6	CBX5_HUMAN	Chromobox protein homolog 5 (HP1 alpha)	22211	146	LTWHAYPEDAENKEK
merge02E18.txt	6	CBX5_HUMAN	Chromobox protein homolog 5 (HP1 alpha)	22211	146	LTWHAYPEDAENKEKETAKS
merge02E19.txt	3	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	379	VEYFLK
merge02E19.txt	3	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	379	WKDSDEADLVLAKE
merge02E19.txt	3	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	379	VEEAPEPEEFVVEK
merge02E19.txt	3	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	379	KVEEAPEPEEFVVEK
merge02E19.txt	3	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	379	KVEEAPEPEEFVVEK
merge02E19.txt	3	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	379	IIGATDSSGELMFLMK
merge02E19.txt	3	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	379	KVEEAPEPEEFVVEKVLDRR
merge02E22.txt	7	CBX3_HUMAN	Chromobox protein homolog 3 (HP1 gamma)	20798	88	KVEEAPEPEEFVVEK

merge22E30.txt	1	CBX5_HUMAN	Chromobox protein homolog 5 - Homo sapiens (Human)	22211	659	IIGATDSCGDLMLFLMK
merge22E30.txt	1	CBX5_HUMAN	Chromobox protein homolog 5 - Homo sapiens (Human)	22211	659	IIGATDSCGDLMLFLMK
merge22E30.txt	1	CBX5_HUMAN	Chromobox protein homolog 5 - Homo sapiens (Human)	22211	659	IIGATDSCGDLMLFLMK
merge22E30.txt	1	CBX5_HUMAN	Chromobox protein homolog 5 - Homo sapiens (Human)	22211	659	LTWHAYPEDAENKEK
merge22E30.txt	1	CBX5_HUMAN	Chromobox protein homolog 5 - Homo sapiens (Human)	22211	659	LTWHAYPEDAENKEK
merge22E37.txt	3	H31T_HUMAN	Histone H3.1t - Homo sapiens (Human)	15499	118	STELLIR
merge22E37.txt	3	H31T_HUMAN	Histone H3.1t - Homo sapiens (Human)	15499	118	EIAQDFK
merge22E37.txt	3	H31T_HUMAN	Histone H3.1t - Homo sapiens (Human)	15499	118	YRPGTVALR
merge22E37.txt	3	H31T_HUMAN	Histone H3.1t - Homo sapiens (Human)	15499	118	KPHRYRPGTVALR
merge22E39.txt	1	H4_HUMAN	Histone H4 - Homo sapiens (Human)	11360	93	ISGLIYEETR
merge22E39.txt	1	H4_HUMAN	Histone H4 - Homo sapiens (Human)	11360	93	DNIQGITKPAIR
merge22E21.txt	3	IMA2_HUMAN	Importin subunit alpha-2 - Homo sapiens (Human)	57826	164	NVSSFPDDATSPLQENR
merge22E21.txt	3	IMA2_HUMAN	Importin subunit alpha-2 - Homo sapiens (Human)	57826	164	GINSNNVENLQATQAAR
merge22E21.txt	1	IMA3_HUMAN	Importin subunit alpha-3 - Homo sapiens (Human)	57775	419	SGILPILVK
merge22E21.txt	1	IMA3_HUMAN	Importin subunit alpha-3 - Homo sapiens (Human)	57775	419	NPPIDDLIK
merge22E21.txt	1	IMA3_HUMAN	Importin subunit alpha-3 - Homo sapiens (Human)	57775	419	DSQVVQVVLDDLK
merge22E21.txt	1	IMA3_HUMAN	Importin subunit alpha-3 - Homo sapiens (Human)	57775	419	EAAWAINLTISSGR
merge22E21.txt	1	IMA3_HUMAN	Importin subunit alpha-3 - Homo sapiens (Human)	57775	419	IEVLQKHENEDIYK
merge22E21.txt	1	IMA3_HUMAN	Importin subunit alpha-3 - Homo sapiens (Human)	57775	419	NILIMAGDEASTIAEIEECGGLEK
merge22E21.txt	1	IMA3_HUMAN	Importin subunit alpha-3 - Homo sapiens (Human)	57775	419	NILIMAGDEASTIAEIEECGGLEK
merge22E21.txt	1	IMA3_HUMAN	Importin subunit alpha-3 - Homo sapiens (Human)	57775	419	KDQVEYLVQQNVIPPCNLLSVK
merge22E21.txt	1	IMA3_HUMAN	Importin subunit alpha-3 - Homo sapiens (Human)	57775	419	SPHQNVCEQAVWALGNIIGDGPQC
merge22E21.txt	1	IMA3_HUMAN	Importin subunit alpha-3 - Homo sapiens (Human)	57775	419	AQNVLTLEAILQNATSDNPVQLSAV
merge22E21.txt	2	IMA4_HUMAN	Importin subunit alpha-4 - Homo sapiens (Human)	57851	247	NPPIDDLIK
merge22E21.txt	2	IMA4_HUMAN	Importin subunit alpha-4 - Homo sapiens (Human)	57851	247	EAAWAINLTISSGR
merge22E21.txt	2	IMA4_HUMAN	Importin subunit alpha-4 - Homo sapiens (Human)	57851	247	IEQLQNHENEDIYK
merge22E21.txt	2	IMA4_HUMAN	Importin subunit alpha-4 - Homo sapiens (Human)	57851	247	DAQVVQVVLDDLNLK
merge22E24.txt	1	RBBP4_HUMAN	Histone-binding protein RBBP4 - Homo sapiens (Human)	47626	86	GEFGGFGSVSGK
merge22E24.txt	1	RBBP4_HUMAN	Histone-binding protein RBBP4 - Homo sapiens (Human)	47626	86	HPSKPDPSGECNPDLR
merge22E11.txt	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta - Homo sapiens	88493	62	LSPPYSSPQFAQDVGR
merge22E13.txt	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta - Homo sapiens	88493	66	LSPPYSSPQFAQDVGR
merge22E13.txt	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta - Homo sapiens	88493	66	LSPPYSSPQFAQDVGRMFKQFNK
merge22E12.txt	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta - Homo sapiens	88493	451	SGEGEVSLMR
merge22E12.txt	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta - Homo sapiens	88493	451	SGEGEVSLMR
merge22E12.txt	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta - Homo sapiens	88493	451	ADVQSIILQGR
merge22E12.txt	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta - Homo sapiens	88493	451	FQWDLNAWTK
merge22E12.txt	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta - Homo sapiens	88493	451	DHQYQFLEDAVR
merge22E12.txt	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta - Homo sapiens	88493	451	LDLTLTADSQPPVFK
merge22E12.txt	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta - Homo sapiens	88493	451	LSPPYSSPQFAQDVGR
merge22E12.txt	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta - Homo sapiens	88493	451	VFPSTTDEYNLIVIER
merge22E12.txt	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta - Homo sapiens	88493	451	IVAERPNTSTGPAPMAPPR
merge22E12.txt	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta - Homo sapiens	88493	451	IVAERPNTSTGPAPMAPPR
merge22E12.txt	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta - Homo sapiens	88493	451	DCQLNAHKDQYQFLEDAVR
merge22E12.txt	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta - Homo sapiens	88493	451	LSPPYSSPQFAQDVGRMFKQFNK
merge23E18.txt	1	CAF1B_HUMAN	Chromatin assembly factor 1 subunit B - Homo sapiens	61454	94	AIVEFLSNLAR
merge23E18.txt	1	CAF1B_HUMAN	Chromatin assembly factor 1 subunit B - Homo sapiens	61454	94	VTCIEIAWHNK
merge23E18.txt	1	CAF1B_HUMAN	Chromatin assembly factor 1 subunit B - Homo sapiens	61454	94	RVTLNTLQAWSK
merge23E30.txt	1	CBX5_HUMAN	Chromobox protein homolog 5 - Homo sapiens (Human)	22211	44	GFSEEHNTWEPEK
merge23E26.txt	1	CBX5_HUMAN	Chromobox protein homolog 5 - Homo sapiens (Human)	22211	85	WKDTEADLVLAKEK
merge23E26.txt	1	CBX5_HUMAN	Chromobox protein homolog 5 - Homo sapiens (Human)	22211	85	GFSEEHNTWEPEK
merge23E28.txt	1	CBX5_HUMAN	Chromobox protein homolog 5 - Homo sapiens (Human)	22211	142	WKDTEADLVLAKEK
merge23E28.txt	1	CBX5_HUMAN	Chromobox protein homolog 5 - Homo sapiens (Human)	22211	142	CPQIVIAFYEER
merge23E28.txt	1	CBX5_HUMAN	Chromobox protein homolog 5 - Homo sapiens (Human)	22211	142	NLDCEPILSEFMK
merge23E28.txt	1	CBX5_HUMAN	Chromobox protein homolog 5 - Homo sapiens (Human)	22211	142	NLDCEPILSEFMK
merge23E29.txt	1	CBX5_HUMAN	Chromobox protein homolog 5 - Homo sapiens (Human)	22211	544	GQVEYLLK
merge23E29.txt	1	CBX5_HUMAN	Chromobox protein homolog 5 - Homo sapiens (Human)	22211	544	DTDEADLVLAKEK
merge23E29.txt	1	CBX5_HUMAN	Chromobox protein homolog 5 - Homo sapiens (Human)	22211	544	KSNFNSNSADDIK
merge23E29.txt	1	CBX5_HUMAN	Chromobox protein homolog 5 - Homo sapiens (Human)	22211	544	WKDTEADLVLAKEK
merge23E29.txt	1	CBX5_HUMAN	Chromobox protein homolog 5 - Homo sapiens (Human)	22211	544	LTWHAYPEDAENKEK
merge23E29.txt	1	CBX5_HUMAN	Chromobox protein homolog 5 - Homo sapiens (Human)	22211	544	GFSEEHNTWEPEK
merge23E29.txt	1	CBX5_HUMAN	Chromobox protein homolog 5 - Homo sapiens (Human)	22211	544	NLDCEPILSEFMK
merge23E29.txt	1	CBX5_HUMAN	Chromobox protein homolog 5 - Homo sapiens (Human)	22211	544	NLDCEPILSEFMK
merge23E29.txt	1	CBX5_HUMAN	Chromobox protein homolog 5 - Homo sapiens (Human)	22211	544	NLDCEPILSEFMK
merge23E29.txt	1	CBX5_HUMAN	Chromobox protein homolog 5 - Homo sapiens (Human)	22211	544	IIGATDSCGDLMLFLMK
merge23E29.txt	1	CBX5_HUMAN	Chromobox protein homolog 5 - Homo sapiens (Human)	22211	544	IIGATDSCGDLMLFLMK
merge23E29.txt	1	CBX5_HUMAN	Chromobox protein homolog 5 - Homo sapiens (Human)	22211	544	IIGATDSCGDLMLFLMK
merge23E29.txt	1	CBX5_HUMAN	Chromobox protein homolog 5 - Homo sapiens (Human)	22211	544	LTWHAYPEDAENKEK
merge23E29.txt	1	CBX5_HUMAN	Chromobox protein homolog 5 - Homo sapiens (Human)	22211	544	TADSSSEDEEEYVVEK
merge23E29.txt	1	CBX5_HUMAN	Chromobox protein homolog 5 - Homo sapiens (Human)	22211	544	WKGFSEEHNTWEPEK
merge23E35.txt	1	H2A1B_HUMAN	Histone H2A type 1-B - Homo sapiens (Human)	14127	66	VTAQQGVLPLNIQAVLLPK
merge23E36.txt	1	H4_HUMAN	Histone H4 - Homo sapiens (Human)	11360	42	ISGLIYEETR
merge23E21.txt	2	IMA3_HUMAN	Importin subunit alpha-3 - Homo sapiens (Human)	57775	84	DSQVVQVVLDDLK

merge23E21.txt	2	IMA3_HUMAN	Importin subunit alpha-3 - Homo sapiens (Human)	57775	84	DYVISLGVVKPLLSFISPSIPITFLR
merge23E21.txt	1	IMA4_HUMAN	Importin subunit alpha-4 - Homo sapiens (Human)	57851	107	DAQVVQVLDGLSNILK
merge23E21.txt	1	IMA4_HUMAN	Importin subunit alpha-4 - Homo sapiens (Human)	57851	107	DYVISLGVVKPLLSFISPSIPITFLR
merge23E34.txt	1	SHOC2_HUMAN	Leucine-rich repeat protein SHOC-2 - Homo sapiens	64848	45	CREENSMRLDLSKR
merge23E13.txt	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta - Homo sapiens	88493	64	LSPPYSSPQEFAQDVGR
merge23E13.txt	1	TIF1B_HUMAN	Transcription intermediary factor 1-beta - Homo sapiens	88493	64	LSPPYSSPQEFAQDVGRMFKQFNK
merge23E29.txt	3	UN13D_HUMAN	Protein unc-13 homolog D - Homo sapiens (Human)	1E+05	44	SSSLASNRLK
merge23E29.txt	3	UN13D_HUMAN	Protein unc-13 homolog D - Homo sapiens (Human)	1E+05	44	IHLFSMAFRELQWLVAK

Annexe 2 : Peptides identifiés appartenant aux protéines Polycomb

Figure S1 : Peptides identifiés pour les protéines Polycomb à partir de CBX2-TAP

CBX2 (A, B)

MEELSSVGEQVFAAECILSKRLRKGKLEYLVKWRGWSSKHNSWEPEENILDPRLLLAFQK
KEHEKEVQNRKRGRPRGRPRKLTAMSSCSRRLKEPDAPSKSKSSSSSSSSSTSSSSSS
DEEDDSDLDAKRGPRGRETHPVPQKKAQILVAKPELKDP~~IR~~KKRGRKPLPPEQKATRRPV
SLAKVLKTARKDLGAPASKLPPPLSAPVAGLAALKAHAKEACGGPSAMATPENLASLMKG
MASSPGRGGISWQSSIVHYMNRMTQSQQAASRLALKAQATNKCGLGDLKVRTQKGEIG
MSPPGSKI PKAPSGGAVEQKVGNTGGPPHTHGASRVPAGCPGPQPAPTQELSLQVLDLQS
VKNGMPGVLLARHATATKGV PATNPAPGKGTGSGLIGASGATMPTDTSKSEKLASRAVA
PPTPASKRDCVKGSATPSGQESRTAPGEARKAATLPEMSAGEESSSSSDSDPDSASPPSTG
QNPSVSVQTSQDWKPTRSLIEHVFVTDVTANLITVTVKESPTSVGFFNLRHY
CHROMO

PHC2

MENELPVPHTSSSACATSSSTSGASSSSGCNNSSSSGGSGRPTGPQISVYSGIPDRQTVQVI
QQALHRQPSTAAQYLQQMYYAAQQOHLMLQTAALQQOHLSSAQLQSLAAVQQASLVSNRQG
STSGSNVSAQAPAQSSINLAASPAQAQLLNRAQSVNSAAASGIAQQAVLLGNTSSPALT
ASQAQMYLRAQMLIFTPTATVATVQPELGTGSPARPPTPAQVQNLTLRTQQTPAAAASGP
TPTQPVLPSLALKPTPGGSQPLPTPAQSRNTAQASAPAGAKPGIADSVMEPHKKGDGNSV
PGSMEGRAGLSRTVPAVAHPLIAPAYAQLQPHQLLPQPSKHLQPFVVIQQQPQPQQQ
PPPQQSRPVLQAEPHPQLASVSPSVALQPSSEAHAMPLGPVTPALPLQCPTANLHKPGGS
QQCHPPTPDTGPPQNGHPEGVPHTPQRRFQHTSAVILQLQPASPPQQCVDDWKEVAPGEK
SVPETRSGSPHQAIVTAMPGGLPVPTSPNIQPSPAHETGQGIVHALTDLSSPGMTSGN
GNSASSIAGTAPQNGENKPPQAIKVPQILTHVIEGFVIEGAEFPVGRSSLLVGNLKKK
YAQGFLEKLPQQDHTTTTDSMEMEPEYLQESKEEGAPLKLKCELCGRVDFAYKFKRSKRF
CSMACAKRYNVGCTKRVGLFHSRSLKQKAGAATHNRRRASKASLPPLTKDTKKQPTGTV
PLSVTAAQLTHSQEDSSRCSNDSYEEPLSPIASASSSTRRRQQRDLELPMHMRDLV
GMGHFLPSEPTKWNVEDVYEFIRSLPGCQEI~~AEEFRAQEIDQALLLLK~~EDHLM~~SAMNI~~
KLGPAIKIYARISMLKDS
SAM

RNF2

MSQAVQTNGTQPLSKTWELSLYELQRTPQEAITDGLEIVVSPRSLHSELMCPICLDMLKN
TMTTKECLHRFCADCIITALRSGNKECPTCRK~~KL~~VSKRSLRDPNFDALISKIYPSRDEY
EAHQERVLARINKHNNQALSHSIEEGLKI~~Q~~AMNRLQRGKKQ~~Q~~IENGSGAEDNGDSSHCS
NASTHSNQEAGPSNKR~~T~~KTSDDSGLELDNNAAMAIDPVMGASEIELVFRPHPTLMEKD
DSAQTRYIKTSGNATVDHLSKYLA~~VRLALEELRSKGESNQMNLDTASEKQYTIYIATASG~~
QFTVLNGSFSLELVSEKYWKVNKPMELLYAPTKEHK
RING

BMI1

MHR~~T~~TRIKITELNPHLMCVLCGGYFIDATTIIECLHSFCKTCIVRYLET~~SKY~~CPICDVQV
HKTRPLL~~NIR~~SDKTLQDIVYKLVPLGFKNEMKRRRDFYAAHPSADAANGSNEDRGEVADE
DKRIITDDEIISLSIEFFDQNR~~L~~DRKVNKDKEKSKEEVNDKRYLRCPAAMTVMHLRKFLR
SKMDIPNTFQIDVMYEEEPKDY~~Y~~TLMDIAYIYT~~WRR~~NGPLPLKYRVRPTCKRMKISHQR
DGLTNAGELES~~DS~~SGDKANS~~P~~AGGIPSTSSCLPSPSTPVQSPHPQ~~F~~PHISS~~T~~MNGTSNSP
SGNHQSSFANRPRKSSVNGSSATSSG
RING

Figure S2 : Peptides identifiés pour les protéines Polycomb à partir de CBX4-TAP

CBX4 (A, B)

MELPAVGEHVFAV **ESIEKKRIRKGRVEYLVKWRGWSPKYNTWEPEENILDPRLLLIAFQNR**
ERQEQLMGYRKR**GPKPKPLVVQVPTFARRSNVLTGLQDSSTDNR****AKLDDLGAQKGQGHQY**
ELNSKKHHQYQPHSKEGKPRPPGKSGKYYYQLNSKHHHPYQDPKMYDLQYQGGHKEAPS
PTCPDLGAKSHPPDKWAQGAGAKGYLGAVKPLAGAAGAPGKSEK **GPPNGMMPAPKEAVT**
GNGIGGKMKIVKNKNKNGR **IVIVMSKYMENGMQAVKIK**SGEVAEGEAR **SPSHKKRAADER**
HPPADRTFKKAAGAEKKEVEAPPKRREEEVSQVSDPQPQDAGSRKLSPTKEAFGEQPLQL
TTKPDLLAWDPARNTHTPPSHHPHPHHHHHHHHHHHHAVGLNLSHVRK **RCLSETHGERE**
PCKKRLTARSISTPTCLGGSPAERPADLPAAALRQPEVILLDSLDLDEPIDLRSVKSRSE
EAGEPPSSLQVKPETPASAAVAVAAAAAPT TTAEKPPAEAQDEPAESLSEFKPFFGNI I I
TDVTANCLTVTFK **EYVTV**
CHROMO Missing in TAP-CBX4

PHC1

METESEQNSNSTNGSSSSGGSSRPQIAQMSLYERQAVQALQALQRPNAAQYFHQFMLQQ
QLSNAQLHSLAAVQATIAASRQASSPNTSTTQQQT TTTTQASINLATTSAAQ LISRSQSV
SSPSATTLTQSVLLGN TTSPLNQSQAQMYLRPQLGNLLQVNRTPGSNVPLASQLILMPN
GAVAAVQQEVPSAQSPGVHADADQVQNLAVRNQOASAQGPQMGGSTQKAI PP GASPVSSL
SQASSQALAVAQASSGATNQLNLSQAGGGSGNSIPGSMGPGGGQAHGGLGQLPSSGMG
GGSCPRKGTGVVQPLPAAQTVTVS QGSQTEAESAAAK **KA EADGSGQNVGMNLTR**TATPA
PSQTLISSATYTIQPHSLIQQQQIHLQOKQVVIQQQIAIHHQQQFQHRQS QLLHTATH
LQLAQQQQQQQQQQQQQQPPQATTLTAPQPPQVPPTQQVPPSQSQQAQTLVVQPMLQSS
PLSLPPDAAPKPIPIQSKPPVAPIKPPQLGAAKMSAAQPPPHIPVQVVGTRQPGTAQA
QALGLAQLAAVPTSRGMPGTVQSGQAHLASSPPSSQAPGALQECPPTLAPGMTLAPVQG
TAHVVKGGATTSSPVVAQVPAAFYMQSVHLPGKPQTLAVKRKADSEERDDVSTLGSMLP
AKASPVAESPKVMDEKSSLGEKAESVANVNANAPSELVALTPAPSVPPP TLAMVSRQMG
DSKPPQAIKVPQILTHIEGFVIEGAEPFPVVGCSQLLKESEKPLQTLPTGLTENQSGG
PLGVDSPSAELDKKANLLKCEYCGKYAPAEQFRGSKRFC SMTCAKRYNVCSHQFRLKRK
KMKEFQEANYARVRRRGRPRSSSDIARAKIQGCHRQEDSSRGSNDSSYDEALSPTSPG
PLSVRAGHGERDLGNPNTAPPTPELHG INPVFLSSNPSR **WSVEEVYEF IASLQGCQEIAE**
EFRSQEIDGQALLLLK EEHLSAMNIKLGPAKIKAKINVLKET
SAM

PHC2

MENELPVPHTSSSACATSSSTSGASSSSGCNNSSSGGSGRPTGPQISVYSGIPDRQTVQVI
QQALHRQPSTAAQYLQOMYAAQQOHLMLQTAALQQOHLSSAQLQSLAAVQASLVSNRQG
STSGSNVSAQAPAQSSSINLAASPAQAQLLNRAQSVNSAAAASGIAQQAVLLGNTSSPALT
ASQAQMYLRAQMLIFTPTATVATVQPELGTGSPARPTPAQVQNLTLRTQQTPAAAAAGP
TPTQPVLPSLALKPTPGGSQPLPTPAQSRNTAQASPAGAKPGIADSVMEPHKKGDGNSV
PGSMGRAGLSRTVPAVAHPLIAPAYAQLQPHQLLPQPSSKHLQPQFVIQQQPQPQQQ
PPPQQSRPVLQAEHPHQLASVSPSVALQPSSEAHAMPLGPVTPALPLQCP TANLHKPGGS
QQCHPPTPDTGPNQNGHPEGVPHTPQRRFQHTSAVILQLQPASPPQCV PDDWKEVAPGEK
SVPETRSGSPHQQAIVTAMPGLPVPTSPNIQPSPAHETGQIVHALTDLSSP GMTSGN
GNSASSIAGTAPQNGENKPPQAIKVPQILTHVIEGFVIEGAEPFPVGR **SLLVGNLKKK**
YAQGFLEKLPQQDHTTTT DSEMEEPYLQESKEEGAPLKLKCELCGRVDFAYKFKRSKRF
CSMACAKRYNVGCTKRVGLFHSDRSKLQKAGAATHNRRRASKASLPPLTKDTK **KQPTGTV**
PLSVTAALQLTHSQEDSSRCSDNSSYEPLSPISASSSTSRRRQGR **DLELPDMHMRDLV**
GMGHFLPSEPTKWNVEDVYEFIRSLPGCQEIAEEFRAQEIDGQALLLLKEDHLSAMNI
KLGPALKIYARISMLKDS
SAM

PHC3

MDTEPNPGTSSVSTTTSTTTTITTTSSSRMQQPQISVYSGSDRHAVQVIQQALHRPPSS
AAQYLQOMYAAQQOHLMLHTAALQQOHLSSSQLSLAAVQASLSSGRPSTSP TGSVTQOS
SMSQTSINLSTSPTPAQLISRSQASSSTSGSITQOTMLLGSTSP TLTASQAQMYLRAQML
IFTATTVAAVQSDIPVSSSSSSSSCQSAATQVQNLTLRSQKLGVLSSSQNGPPKSTSQ
QSLTICHNKQITVTSSKISQRDPSPE SNKKGSPSLESRSTAVTRTSSIHQLIAPASYSP I
QPHSLIKHQIPLHSPPSKRVSHHQLILQEQQQQIQPITLQNSTQDPPPSQHCIPLQNHGL
PPAPSNAQSQHCSPIQSHPSPLTVSPNQSSAQSVVVSPPPHSPSQSPTII IHPQALI
QPHPLVSSALQPGPNLQOSTANQVQATAQLNLP SHLPLPASPVVHIGPVQQSALVSPGQQ

IVSPSHQOYSSLQSSPIPIASPPQMSTSPPAQIPPLPLQSMQSLQVQPEILSQGQVLVQN
ALVSEEEELPAAEALVQLPFQTLPPPQTVAVNLQVQPPAPVDPVQVEDVCEEEMPEES
DECVRMDRTPPPPTLSPAAITVGRGEDLTSEHPLLEQVELPAVASVSASVIKSPSDPSHV
SVPPPLLLPAATTRSNSTSMHSSIPS IENKPPQAIVKPQILTHVIEGFV IQEGLPEFPV
SRSSLLIEQPVKKRPLLDNQVINSVCVQPELQNNTKHADNSSDTEMEDMIAEETLEEMDS
ELLKCEFCGKMGYANEFRLSKRFCTMSCAKRYNVSCSKKFALSRWNRKPDNQSLGHRGR
PSGPDGAAREHILRQLPITYPSAEEDLASHEDSVPSAMTTRLRRQSERERERELRDVIR
KMPENSDLLPVAQTEPSIWTVDVWAFIHSLPGCQDIADEFRAQEQIDGQALLLLKEDHLM
SAMNIKLGPAKICARINSLKES
SAM

RING1

MTTPANAQNASKTWELSLYELHRTPQEAIMDGTEIAVSPRSLHSELMCPICLDMLKNTMT
TKECLHRFCSDCIVTALRSGNKECPTCRKCLVSKRSLRDPNFDALISKIYPSREEYEAH
QDRVIRLSRLLHNQOALSSSIEEGLRMQAMHRAQRVRRPIPQSDQTTTMSGGEGEPGE
GDGEDVSSDSAPDSAPGPAPKRPRGGGAGGSSVGTGGGGTGGVGGGAGSEDSGDRGGTLG
GGTLGPPSPPPGAPSPPEPGEIELVFRPHPLLVEKGEYCQTRYVKTTGNATVDHLSKYLA
LRIALERRQQEAGEPGGPGGASDTGGPDGCGGEGGGAGGGDGPEEPALPSLEGVSEKQ
YTIYIAPGGGAFTTLNGLTLELVNEKFWKVSRPLELCYAPTKDPK
RING

RNF2

MSQAVQTNGTQPLSKTWELSLYELQRTTPQEAITDGLEIIVSPRSLHSELMCPICLDMLKN
TMTTKECLHRFCADCIITALRSGNKECPTCRKCLVSKRSLRDPNFDALISKIYPSRDEY
EAHQERVLARINKHNNQOALSSHSIEEGLKIQAMNRLQRGKKQOIENSGAEDNGDSSHCS
NASTHSNQEAGPSNKRKTSDDSGLELDNNAAMAIDPVMGASEIELVFRPHPTLMEKD
DSAQTRYIKTSGNATVDHLSKYLAVRLALEELRSKGESNQMNLDTASEKQYTIYIATASG
QFTVLNGSFSLELVSEKYWKVNKPMELYYAPTEHK
RING

PCGF1

MRLRNQLQSVYKMDPLRNEEEVRVKIKDLNEHIVCCLCAGYFVDATTITECLHTFCKSCI
VKYLQTSKYCPMCNIKIHETQPLLNKLDKDRVMQDIVYKLVPLQDSEEKRIREFYQSRGL
DRVTQPTGEEPALSNLGLPFSSFDHKAHYRYDEQLNLCLERLSSGKDKNKSVLQNKYV
RCSVRAEVRHLRRVLCHRLMLNPQHVQLLFDNEVLPDHMTMKQIWLRSRWF GKPSPLLLQY
SVKEKRR
RING

PCGF2

MHRTRIKITELNPHLMCALCGGYFIDATTIVECLHSFCKTCIVRYLETNKYCPMCDVQV
HKTRPLLSIRSDKTLQDIVYKLVPLGFKDEMRRRDFYAAYPTEVPNGSNEDRGEVLEQ
EKGALSDDEIVSLSIEFYEGARDRDEKKGLENGDGKEKTGVRFLRCPAAMTMHLAKF
LRNKMDVP SKYKVEVLYEDELKEYTLMIDAIYIPWRRNGPLPLKYRVQPACKRLTLAT
VPTPSEGTNTSGASECESVSDKAPSPATLPATSSSLPATPSHGSPSSHGPPATHPTSP
TPPSTASGATTAANGGSLNCLQTPSSSTRGRKMTVNGAPVPPLT
RING

BMI1

MHRTRIKITELNPHLMCVLCGGYFIDATTIECLHSFCKTCIVRYLETNKYCPICDVQV
HKTRPLLNIRSDKTLQDIVYKLVPLGFKDEMRRRDFYAAHPSADAANGSNEDRGEVADE
DKRIITDDEIISLSIEFFDQNLDRKVNKDKEKSKEEVNDKRYLRCPAAMTMHLRKFLR
SKMDIPNTFQIDVMYEEEPKDYTLMDIAYIYTWRNGPLPLKYRVQPACKRLTLAT
DGLTNAGELESSDSGSKANSPAGGIPSTSSCLPSPSTPVQSPHPQPHISSTMNGTNSNP
SGNHQSSFANRPRKSSVNGSSATSSG
RING

PCGF6

MEGVAVVTAGSVGAakteGAAALPPPPVSPPALTPAPAAGEEGPAPLSETGAPGCSGSR
PPELEPERSLGRFRGRFEDEDELEEEEEEEEEEEEEEDMSHFSLRLEGGRQDSEDEEE
RLINLSELTPYILCSICKGYLIDATTITECLHTFCKSCIVRHFYYSNRCPKCNIVVHQQTQ
PLYNIRLDRQLQDIVYKLVINLEEREKKQMDFYKERGLEVPKPAVPQVPVSSKGRSCKV
LESVFRIPPELDMSLLLEF IGANEGTGHFKPLEKKFVRVSGEATIGHVEKFLRRKMGLDP
ACQVDIICGDHLLQYQTLREIRRAIGDAAMQDGLLVLHYGLVVSPLKIT

RING

SCMH1

MLVCYSVLACEILWDLPCSIMGSPGLGHFTWDKYLKETCSVPAPVHCFKQSYTPPSNEFKI
SMKLEAQDPRNTTSTCIATVVLGTGARLRRLRLDGSNDKNDFWRLVDSAEIQPIGNCEKNG
GMLQPPGLGFRNLNASSWPMFLKTLNGAEMAPIRIFHKEPPSPSHNFFKMGMKLEAVDRKN
PHFICPATIGEVRGSEVLVTFDGRGAFDYWCRFDSRDI F PVGWCSLTGDNLQPHGTKVV
IPKNPYPASDVNTEKPSIHSSTKTVLEHQPGQRGRKPGKKRGRTPKTLISHPISAPSKTA
EPLKFPKKRGPKPGSKRKPRTLLNPPASPTTSTPEPDTSTVPQDAATIPSSAMQAPTVC
IYLNKNGSTGPHLDKKKVQQLPDHFGPARASVVLQQAQACIDCAYHQKTVFSFLKQGHG
GEVISAVFDREQHTLNLPAVNSITYVLRFLLEKLCHNLRSNDLFGNQPFQTHTLSLTAIEY
SHSHDRYLPGETFVLGNLARSLEPHSDSMSASNPNTLVSTSQRHRPLLSSCGLPPSTA
SAVRRCLSRGVLKGSNERRDMESFWKLNRSFGSDRYLESRDASRLSGRDPSSWTVEDVMQ
FVREADPQLGPHADLFRKHEIDGKALLLLRSDMMMKYMGKLGPKLKLKLSYHIDRLKQKGF

MBT (2 domaines)

SAM

Figure S3 : Peptides identifiés pour les protéines Polycomb à partir de CBX6-TAP

CBX6 (A, B)

MELSAVGERVFAAESIIKRRIRKGRIEYLVKWKGWAIKYSTWEPEENILDSRLIAAFEQK
ERERELYGPKKRGPKPKTFLLKARAQAEALRISDVHFSVKPSASASSPKLHSSAAVHRLK
KDIRRCHRMSRPLPRPDPQGGSPGLRPPISPFSETVRI INRKVKPREPKRNR I IILNLKV
IDKGAGGGGAGQAGALARPKVPSRNRVIGKSKKFSSEVLR TQIRHMKFGAFALYKPPPA
PLVAPSPGKAEASAPGPGLLLAAPAAPYDARSSGSSGCPSPPTQSSDPDDTTPPKLLPETV
SPSAPSWREPEVLDLSLPPESAATSKRAPPEVTAAGPAPPTAPEPAGASSEPEAGDWRP
EMSPCCSNVVVTDVTSNLLTVTIK EFCNPEDFEKVAAGVAGAAGGGGSI GASK

PHC2

MENELPVPHTSSSACATSSSTSGASSSSGCNNSSSGGSGRPTGPQISVYSGIPDRQTVQVI
QQALHRQPSTAAQYLQOMYAAQQOHLMLQTAALQQOHLSSAQLQSLAAVQQASLVSNRQG
STSGSNVSAQAPAQSSSINLAASPAQAQLLNRAQSVNSAAAASGIAQQAVLLGNTSSPALT
ASQAQMYLRAQMLIFTPTATVATVQPELGTGSPARPPTPAQVQNLTLRTQQTPAAAASGP
TPTQPVLPSLALKPTPGGSQPLPTPAQSRNTAQASPAKPGIADSVMEPHKKGDGNSV
PGSMEGRAGLSRTVPAVAHPLIAPAYAQLQPHQLLPQPSSKHLQPQFVIQQQPQPQQQ
PPPQQSRPVLQAEHPQLASVSPSVALQPSSEAHAMPLGPVTPALPLQCPTANLHKPGGS
QQCHPPTPDTGPNQGHPEGVPHTPQRRFQHTSAVILQLQPASPPQQCVDDWKEVAPGEK
SVPETRSGPSHPHQAIVTAMPGGPLPVPTSPNIQPSPAHETGQGIVHALTDLSSPGMTSGN
GNSASSIAGTAPQNGENKPPQAIIVKQILTHVIEGFVIQEGAEFPVGRSSLLVGNLKKK
YAQGFLEKLPQQDHTTTTDSMEMEPLYQESKEEGAPLKLKCELCGRVDFAYKFKRSKRF
CSMACAKRYNVGCTKRVGLFHSRDRSKLQKAGAATHNRRRASKASLPPLTKDTKKQPTGTV
PLSVTAALQLTHSQEDSSRCSNDSSSYEPLSPISASSSTSRRRQQRDLLELPDMHMRDLV
GMGHFLPSEPTKWNVEDVYEFIRSLPGCQIEAEFRAQEIDGQALLLLKEDHLSAMNI
KLGPKLIYARISMLKDS

SAM

RING1

MTTPANAQNASKTWELSLYELHRTPOEAIMDGTEIAVSPRSLHSELMCPICLDMLKNMT
TKECLHRFCSDCIVTALRSGNKECPTCRKKLVSKRSLRDPNFDALISKIYPSREEYEAH
QDRVLIRLSRLHNQQALSSSIEEGLRMQAMHRAQRVRRPIPGSDQTTTMSGGEGEPGEGE
GDGEDVSSDSAPDSAPGPAPKRPRGGGAGGSSVGTGGGGTGGVGGGAGSEDSGDRGGTLG
GGTLGPPSPPGAPSPPEPGEIELVFRPHPLLVEKGEYCQTRYVKTGTGNATVDHLSKYLA
LRIALERRQQQEAGEPGGPGGGASDTGGPDGCGGEGGGAGGGDGPEEPALPSLEGVSEKQ
YTIYIAPGGGAFTTLNGLSLTELVNEKFWKVSRLPLELCYAPTCDPK

RING

RNF2

MSQAVQTNGTQPLSKTWELSLYELQRTPOEAITDGLEIVVSPRSLHSELMCPICLDMLKN
TMTTKECLHRFCADCIITALRSGNKECPTCRKKLVSKRSLRDPNFDALISKIYPSRDEY
EAHQERVRLARINKHNNQQALSHSIEEGLKIQAMNRLQRGKKQOIENGSGAEDNGDSSHCS
NASTHSNQEAGPSNKRKTSDDSGLELDNNAAMAIDPVMGASEIELVFRPHPTLMEKD
DSAQTRYIKTSGNATVDHLSKYLA VRLALEELRSKGESNQMNLDTASEKQYTIYIATASG

QFTVLNGSFSLELVSEKYWKVNKPMELYYAPTKEHK

RING

SCML2

LRLRLDGS DNR

GQTVNEDSMDVKK

MGQTVNEDSMDVK

Figure S4 : Peptides identifiés pour les protéines Polycomb à partir de CBX7-TAP

CBX7 (A, B)

MELSAIGEQVFAVESIRKKRVRK**GKVEYLVKWKGWPPK**YSTWEPEEHILDPRILVMAYEEK
EERDRASGYRKRGPVKRLLQL**RLYSMDLR**SSHKAKGKEKLCFSLTCPLGSGSPGCVVKA
GAPELVDKGLVPTLPFPLRKAHKYLRLSRKKFPPRGNLESHSHRRELFLQEPAP
DVLQAAGEWEPAAQPPEEEADADLAEGPPPWPALPSSEVTVDITANSITVTFR**EAQAA**
EGFFRDRSGKF

PHC2

MENELPVPHTSSSACATSSSTSGASSSSGCNNSSSGGSGRPTGPQISVYSGIPDRQTVQVI
QQALHRQPSTAAQYLQOMYAAQQOHLMLQTAALQQOHLSSAQLQSLAAVQQASLVSNRQG
STSGSNVSAQAPAQSSSINLAASPAQAQLLNRAQSVNSAAAASGIAQQAVLLGNTSSPALT
ASQAQMYLRAQMLIFTPTATVATVQPELGTGSPARPPTPAQVQNLTLRTQQTPAAAASGP
TPTQPVLPSLALKPTPGGSQPLPTPAQSRNTAQASPAGAKPGIADSVMEPHKKGDGNSV
PGSMEGRAGLSRTVPAVAHPLIAPAYAQLQPHQLLPQPSSKHLQPQFVIQQQPQPQQQQ
PPPQQSRPVLQAEHPHQLASVSPSVALQPSSEAHAMPLGPVTPALPLQCPATANLHKPGGS
QQCHPPTPDTGPNQNGHPEGVPHTPQRRFQHTSAVILQLQPASPPQQCVDDWKEVAPGEK
SVPETRSGPSPHQQAIVTAMPGGLPVPTSPNIQPSPAHETGQIVHALTDLSSPGMTSGN
GNSASSIAGTAPQNGENKPPQAIVKPQILTHVIEGFVIEGAEFPVGRSSLLVGNLKKK
YAQGFLEKLPQDHTTTT DSEMEEPYLQESKEEGAPLKLKCELCGRVDFAYKFKRSKRF
CSMACAKRYNVGCTKRVGLFHSRSLKQKAGAATHNRRRASKASLPPLTKDTKKQPTGTV
PLSVTAALQLTHSQEDSSRCSNNSYEEPLSPIASSSTSRRRQQRDLELDPDMHRDLV
GMGHHFLPSEPTK**WNVEDVYEFIRSLPGCQEI**AAEFRAQEI**DGQALLLLKEDHLSAMNI**
KLGPALKIYARISMLKDS
SAM

RING1

MTTPANAQNASKTWELSLYELHRTPOEAIMDGTEIAVSPR**SLHSELMCPICLDMLKNTMT**
TKECLHRFCSDCITALRSGNKECPTCRKCLVSKR**SLRPDPNFDALISKI**YPSREYEYEAH
QDRVLIIRLSR**LHNQALSSSIEEGLR**MQAMHRAQRVRRIPGSDQTTTMSGGEGEPGE
GDGEDVSSDSAPDSAPGPAPKRPR**GGGAGGSSVGTGGGGTGGVGGGAGSEDSGDR**GGTLG
GGTLGPPSPPGAPSPPEPGGEIELVFRPHPLLVEKGEYCQTRYVK**TTGNATVDHLSK**YLA
LR**IALERR**QQQEAGEPGGPGGASDTGGPDGCGGEGGGAGGGDGPPEPALPSLEGVSEKQ
YTIYIAPGGGAFTTLNGSLTLELVNEKFWK**VSRPLELCYAPTKDPK**
RING

RNF2

MSQAVQTNGTQPLSK**TWELSLYELQRTPOEAITDGLEIVVSPRSLHSELMCPICLDMLKN**
TMTTKECLHRFCADCITALRSGNKECPTCRKCLVSKR**SLRPDPNFDALISKI**YPSRDEY
EAHQERVLARINK**HNNQALSHSIEEGLK**IQAMNRLQRGKKQIENGSGAEDNGDSSHCS
NASTHSNQEAGPSNKRTKTSDDSGLELDNNAAMAIDPVMGASEIELVFRPHPTLMEKD
DSAQTRYIK**TSGNATVDHLSKYLA**VRLALEELR**SKGESNQMNLDTASEK**QYTIYIATASG
QFTVLNGSFSLELVSEKYWKVN**KPMELYYAPTKEHK**
RING

PCGF2

MHRTRRIKITELNPHLM**CALCGGYFIDATTIVECLHSFCKTCIVRYLETNKYCPMCDVQV**
HKTRPLLSIRSDK**TLQDIVYKLVPLGFK**DEMRRRRDFYAAAYPLTEVPNGSNEDRGEVLEQ
EKGALSDDEIVSLSIEFYEGARDRDEKKGPLENGDGDKEKTGVRFLRCPAAMTMHLAKF
LRNKMDVPSKYKVEVLYEDEPLKEYYTLMDIAYIYPWRRNGPLPLKYRVQPACKRLTLAT
VPTPSEGTNTSGASECESVSDKAPSPATLPATSSSLPSPATPSHGSPSSHGPPATHPTSP
TPPSTASGATTAANGGSLNCLQTPSSSTRGRKMTVNGAPVPLT
RING

PCGF3

MLTRKIKLWDINAHIT CRLCSGYLIDATTVTECLHTFCRSCLVKYLEENNTCPTCR IVIH
QSHPLQYIGHDRMTQDIVYKLVPGLEAEMRKQREFYHKLGMVPGDIKGETCSAQHLD
SHRNETKADSSNKEAAEEKPEEDNDYHRSDEQVSI CLECNSSKLRGLKRKWRCSAQA
TVLHLKKFIAKKLNLSSFNELDILCNEEILGKDHTLK FVVVTR WRFKKAPLLLHYRPKMD
LL

RING

BMI1

MHRTTRIKITELNPHLMCVLCCGGYFIDATTIIECLHSFCKTCIVRYLETSKY CPICDVQV
HKTRPLLNIIRSDK TLQDIVYKLVPGLFK NEMKRRRDFYAAHP SADAANGSNEDRGEVADE
DKRIITDDEIISLSIEFFDQNRLLDRKVNKDKEKSKEEVNDKRYLR CPAAMTVMHLR KFLR
SKMDIPNTFQIDVMYEEEEPLKDYITLMDIAYIYTWRNGPLPLKYRVRPTCKRMKISHQR
DGLTNAGELES SDSGSK ANSPAGGIPSTSSCLPSPSTPVQSPHPQFPHISSTMNGTNSNP
SGNHQSSFANRPRKSSVNGSSATSSG

RING

PCGF6

MEGVAVVTAGSVGAakteGAAALPPPPPVSPALTPAPAAGEEGPAPLSETGAPGCSGSR
PPELEPERSLGRFRGRFEDEDELEEEEELEEEEEEDMSHFSLRLEGGRQDSEDEEE
RLINLSELTPYIL CSICKGYLIDATTITECLHTFCKSCIVRHFYYSNRCPKCN IIVHQIQ
PLYNIRLDR QLQDIVYKLVINLEERE KKQM HDFYKER GLEVPKPAVPQVPSSK GRSKKY
LESVFR IPPELDMSLLLEF IGANEGTGHFKPLEKKFVVR VSGEATIGHVEK FLRRKMGLDP
ACQVDIICGDHLLLEQYQTLREIRRAIGDAAMQDGLLVLHYGLVVSPLKIT

RING

Figure S5 : Peptides identifiés pour les protéines Polycomb à partir de CBX8-TAP

CBX8 (A, B)

MELSAVGER VFAAEALLKRR IRKGR MEYLVKWKGSQKYSTWEPEENILDARLLAAFEER
EREMELYGPKKRGP KPKTFLLKAQAKAKAKTYEFR SDSARGIRIPYPGRSPQDLASTSRA
REGLRNMGLSPASSTSTSTCTR AEAPRDRDRDRDR DRERDRERERERERERERERERER
GTSRVDDKSSPGDSSKRGPKPR KELPDPSQRPLGEP SAGLGEYLKGRKLDDTPSGAGK
FPAGHSVIQLARRQSDLVQCGVTSPSSAEATGKLAVDTFPAR VIKHRAAFLEAKGQGal
DPNGTRVRHGSGPPSSGGGLYR DMGAQGGRP SLIAR IPVAR ILGDPEEESWSPSLTNLEK
VVTDVTSNFLVTITIKESNTDQGFFK EKR

PHC2

MENELPVPHTSSSACATSSTSGASSSSGCNNSSSGGSGRPTGPQISVYSGIPDRQTVQVI
QQALHRQPSTAAQYLQQMYYAAQQOHLMLQTAALQQOHLSSAQLQSLAAVQQASLVSNRQG
STSGSNVSAQAPAQSSINLAASPAQAQLLNRAQSVNSAAAASGIAQQAVLLGNTSSPALT
ASQAQMYLRAQMLIFTPTATVATVQPELGTGSPARPPTPAQVQNLTLRTQQTPAAAAAGP
TPTQPVLPSLALKPTPGGSQPLPTPAQSRNTAQASPAGAKPGIADSVMEPHKKGDGNSV
PGSMEGRAGLSRTVPAVAHPLIAPAYAQLQPHQLLPQSSKHLQPQFVIQQQPQPQQQ
PPPQQSRPVLQAEPHPQLASVSPSVALQPSSEAHAMPLGPVTPALPLQCPTANLHKPGGS
QQCHPPTPDTGPPQNGHPEGVPHTPQRRFQHTSAVILQLQPASPPQCVPDDWKEVAPGEK
SVPETRSGSPHQAIIVTAMPGGLPVPTSPNIQPSPAHETGQGIVHALTDLSSPGMTSGN
GNSASSIAGTAPQNGENKPPQAIIVKPKQILTHVIEGFVIQEGAEPFPVGRSSLLVGNLKKK
YAQGFLEKLP QQDHTTTT DSEMEEPYLQESKEEGAPLKLKCELCGRVDFAYKFKRSKRF
CSMACAKRYNVGCTKRVGLFHSDRSKLQKAGAATHNRRRASKASLPPLTKDTKKQPTGTV
PLSVTAAALQLTHSQEDSSRCSDNSSYEEPLSPIASASSSTRRRQQRDLELPMHMRDLV
GMGHFLPSEPTK WNVEDVYEFIRSLPGCQEI AEFFRAQ EIDGQALLLLKEDH LMSAMNI
KLGPALKIYARISMLKDS

SAM

RING1

MTTPANAQNASKTWELSLYELHRTPOEAIMDGTEIAVSPRSLHSELM CPICLDMLKNTMT
TKECLHRFCSDCIVTALR SGNKECPTCRKLLVSKR SLRDPNFDALISKIYPSREYEYEAH
QDRV LIRLSRLHNQALSSSIEEGLRMQAMHRAQRVRRPIPGSDQTTTMSGGEGEPGE
GDGEDVSSDSAPDSAPGPAPKRPR GGGAGGSSVGTGGGGTGGVGGGAGSEDSGDR GGTLG
GGTLGPPSPPGAPSPPEPGEIELVFRPHPLLVEKGEYCQTRYVKTGTGNATVDHLSKYLA

LRIALERRQQEAGEPGGGGGASDTGGPDGCGGEGGGAGGGDGPEEPALPSLEGVSEKQ
YTIYIAPGGGAFTTLNGSLTLELVNEKFWKVSRLPLELCYAPTKDPK
RING

RNF2

MSQAVQNTGTQPLSKTWELSLYELQRTPQEAITDGLEIVVSPRSLHSELMCPICLDMLKN
TMTTKECLHRFCADCIITALRSGNKECPTCRKKLVSKRSLRPDPNFDALISKIYPSRDEY
EAHQERVLARINKHNNQALSHSIEEGLKIQAMNRLQRGKKQIENGSGAEDNGDSSHCS
NASTHSNQEAGPSNKRKTSDDSGLELDNNAAMAIDPVMGASEIELVFRPHPTLMEKD
DSAQTRYIKTSGNATVDHLSKYLAVALALEELRSKGESNQMNLDTASEKQYTIYIATASG
QFTVLNGSFSLELVSEKYWKVNKPMELYYAPTK~~EHK~~
RING

PCGF1

MRLRNQLQSVYKMDPLRNEEEVRVKIKDLNEHIVCCLCAGYFVDATTITECLHTFCKSCI
VKYLQTSKYCPMCN~~IKIHETQPLLNKLD~~VMQDIVYKLVPGLODSEEKRIREFYQSRGL
DRVTQPTGEEPALSNLGLPFSFDHSAHYR~~YDEQLNLCLE~~RLSSGKDKNSVLQNKYV
RCSVRAEVRHLRRVLC~~HRMLMNPQH~~VLLFDNEVLPDHMTMKQIWL~~SRWFGKPS~~PLLLQY
SVKEKRR
RING

PCGF2

MHRTRRIKITE~~LNPHLMCALCGGYFIDATTIVE~~CLHSFCKTCIVRYLETNKYCPMCDVQV
HKTRPLLSIRSDKTLQDIVYKLVPGLFKDEM~~KRRRDFYAA~~YPLTEVPNGSNEDRGEVLEQ
EKGALSDEIVSLSIEFYEGARDRDEKKGPLENGDGKEKTGVRFLRCPAAMTVMHLAKF
LRNKMDVPSKYKVEVLYE~~DEPLKEYYTLMDIAYI~~YPWRRNGPLPLKYRVQPACKRLTLAT
VPTPSEGTNTSGASECESVSDKAPSPATLPATSSSLPSPATPSHGSPSSHGPPATHPTSP
TPPSTASGATTAANGGSLNCLQTPSSTSRGRKMTVNGAPVPPLT
RING

PCGF3

MLTRKIKLWDINAHITCRLCSGYLIDATTVTECLHTFCRSCLVKYLEENNTCPTCR~~IVIH~~
QSHPLQYIGHDRMQDIVYKLVPGLOEAEMRKQREFYHKLGMVPGDIKGETCSAQHLD
SHRNGETKADSSNKEAAEKEPEEDNDYHRSDEQVSI~~CLECNSSKLRGLKRK~~WIRCSAQA
TVLHLK~~KFI~~AKKLNLSFNELDILCNEEILGKDHTLKFVVVTRWRFKKAPLLHYPKMD
LL
RING

BMI1

MHRTRRIKITE~~LNPHLMCVLCGGYFIDATTIE~~CLHSFCKTCIVRYLETSKYCPICDVQV
HKTRPLLNIRSDKTLQDIVYKLVPGLFKNEM~~KRRRDFYAA~~HPSADAANGSNEDRGEVADE
DKRIITDDEIISLSIEFFDQNR~~LDRKVNKDKEKS~~KEEVNDKRYLRCPAAMTVMHLRKF~~FLR~~
SKMDIPNTFQIDVMYEEEEPLKDY~~YTLMDIAYI~~YTWRNGPLPLKYRV~~RPTCKRMK~~ISHQR
DGLTNAGELES~~DSGSDKAN~~SPAGGIPSTSSCLPSPSTPVQSPHPQPHISSTMNGT~~SN~~SP
SGNHQSSFANRPRKSSVNGSSATSSG
RING

PCGF5

MATQRKHLVKDFNPYITCYICKGYLIKPTTVTECLHTFCKTCIVQHFE~~DSNDCPR~~CGNQV
HETNPLEMLRLDNTLEEIIIFKLVPG~~LEQE~~LERESEFWKKNKPQENGQDDTSKADKPKVD
EEGDENEDDKDYHRSDPQIAICLDCLRNNGQSGDNVVKGLMKK~~FIRCSTRVTVGT~~IKKFL
SLK~~LKLPSS~~YELDVLCNGEIMGKDHTMEFIYMTRWRLRGENFRCLNCSASQVCSQDGP~~LY~~
QSYPMVLQYRPRIDFG
RING

PCGF6

MEGVAVVTAGSVGAAKTEGAAALPPPPVSPALTPAPAAGEEGPAPLSETGAPGCSGSR
PPELEPERSLGRFRGRFEDEDEELEE~~EEEEEEEE~~EDMSHFSLRLEGGRQDSEDEEE
RLINLSELTPYILCSICKGYLIDATTITECLHTFCKSCIVRHFYYSNRCPKCN~~IVVHQ~~TQ
PLYNIRLDRQLQDIVYKLVINLEEREKQMHDFYKERGLEVPKPAVPQVPSSKGRS~~SKV~~
LESVFRIPPELDMSLLLEF~~IGANEGTGHFK~~PLEKKFV~~RVSGEAT~~IGHVEKFLRRKMGLDP
ACQVDIICGDHLLQYQTLREIRRAIGDAAMQDGLLVLHYGLVVSPLKIT
RING

SCML2
RIVQACVDCALETK
LRLRLDGS DNR

Figure S6 : Peptides identifiés pour les protéines Polycomb à partir de PCGF2-TAP

PCGF2

MHRTRRIKITELNPHLMCALCGGYFIDATTIVECLHSFCKTCIVRYLETNKYCPMCDVQV
HKTRPLLSIRSDKTLQDIVYKLVPGFLFKDEMKR RRRDFYAAAYPLTEVPNGSNEDRGEVLEQ
EKGALSDDEIVLSLSIEFYEGARDRDEKKGPLENGDGDEKKTGVRFLRCPAAMTVMHLAKF
LRNKMDVPSKYKVEVLYEDEPLKEYYTLMDIAYIYPWRRNGPLPLKYRVQPACKRLTLAT
VPTPSEGTNTSGASECESVSDKAPSPATLPATSSSLPSPATPSHGSPSSHGPPATHPTSP
TPPSTASGATTAANGGSLNCLQTPSSSTRGRKMTVNGAPVPPLT
RING

PHC2

MENELPVPHTSSSACATSSSTSGASSSSGCNNSSSGGSGRPTGPQISVYSGIPDRQTVQVI
QQALHRQPSTAAQYLQQMYAAQQQHLMLQTAALQQQHLSSAQLQSLAAVQQASLVSNRQG
STSGSNVSAQAPAQSSSINLAASPAQAQLLNRAQSVNSAAAASGIAQQAVLLGNTSSPALT
ASQAQMYLRAQMLIFTPTATVATVQPELGTGSPARPPTPAQVQNLTLRTQQTAAAAASGP
TPTQPVLPSLALKPTPGGSQPLPTPAQSRNTAQASPAKPGIADSVMEPHKKGDGNSV
PGSMEGRAGLSRTVPAVAHPLIAPAYAQLQPHQLLPQPSSKHLQPQFVIQQQPQPQQQ
PPPQQSRPVLQAEPHPQLASVSPSVALQPSSEAHAMPLGPVTPALPLQCPTANLHKPGGS
QQCHPPTPDTGPNQHGPEGPVHTPQRRFQHTSAVILQLQPASPPQQCVDDWKEVAPGEK
SVPETRSGSPHQAIVTAMPGGLPVPTSPNIQPSPAHETGQIVHALTDLSSPGMTSGN
GNSASSIAGTAPQNGENKPPQAIVKPQILTHVIEGFVIQEGAEFPVGRS SLLVGNLKKK
YAQGFLEKLPQQDHTTTTTDSEMEEPYLQESKEEGAPLKLKCELCGRVDFAYKFKRSKRF
CSMACAKRYNVGCTKRVGLFHSRSLKQKAGAATHNRRRASKASLPPLTKDKKQPTGTV
PLSVTAALQLTHSQEDSSRCSDNSSYEEPLSPIASASSSTRRRQQRDLELPMHMRDLV
GMGHFLPSEPTKWNVEDVYEFIRSLPGCQEI AEEFRAQ EIDGQALLLLKEDHLM SAMNI
KLGPAIKIYARISMLKDS
SAM

PHC3

MDTEPNPGTSSVSTTTSTTTTTITSSSRMQQPQISVYSGSDRHAVQVIQQALHRPPSS
AAQYLQQMYAAQQQHLMLHTAALQQQHLSSSQLQSLAAVQASLSSGRPSTSPSTGVTQQS
SMSQTSINLSTSPTPAQLISRSQASSSTSGSITQQTMLLGSTSPTLTASQAQMYLRAQML
IFTPTATTVAAVQSDIPVVSSSSSSSCQSAATQVQNLTLRSQKLGVLSSSQNGPPKSTSTQ
QSLTICHNKTTVTSSKISQRDPSPESSKKGESPSLESSTAVTRTSSIHQLIAPASYSPI
QPHSLIKHQIPLHSPSKVSHHQLILQQQQQIQPITLQNSTQDPPPSQHCIPLQNHGL
PPAPSNAQSQHCSPIQSHPSPLTVSPNQSSQAQSVVSPPPPHSPQSPTI IHPQALI
QPHPLVSSALQPGPNLQOSTANQVQATAQLNLP SHLPLPASPVVHIGPVQQSALVSPGQQ
IVSPSHQQYSSLQSSPIPIASPPQMSTSPPAQIPPLPLQSMQSLQVQPEILSQGVLVQN
ALVSEEEELPAAEALVQLPFQTLPPPQTVAVNLQVQPPAPVDPPVYQVEDVCEEEMPEES
DECVRMDRTPPPPTLSPAAITVGRGEDLTSEHPLLEQVELPAVASVSASVIKSPSDPSHV
SVPPPPLLLPAATTRSNSTSMHSSIPS IENKPPQAIVKPQILTHVIEGFVIQEGLEFPV
SR SLLIEQPVK RPLLDNQVINSVCVQPELQNNTKHADNSSDTEMEDMIAEETLEEMDS
ELLKCEFCGKMGYANEF LRSKRFTMSCAKRYNVSCSKKFALSRWNRKPDNQLGHRGR
PSGPDGAAREHILRQLPITYPSAEEDLASHEDSVPSAMTTRLRQSERERERELRDVIR
KMPENSDLLPVAQTEPSIWTVDVWAFIHS LPGCQDI ADEFRAQ EIDGQALLLLKEDHLM
SAMNIKLGPAIKICARINSLKES
SAM

RING1

MTPPANAQNASKTWELSLYELHRTPQEAIMDGTEIAVSPRSLHSELMCPICLDMLKNMT
TKECLHRFCSDCIVTALRSGNKECPTCRK KLVSKRSLRDPNFDALISKIYPSREEYEAH
QDRVLI RLSRLHNQ QALSSSIEEGLRMQAMHRAQRVRRIPGSDQTTTMSGGEGEPGE
GDGEDVSSDSAPDSAPGPAPKRPRGGGAGGSSVGTGGGGTGGVGGGAGSEDSGDRGGTLG
GGTLGPPSPGAPSPPEPGGEIELVFRPHPLLVEKGEYCQTRYVKT TGNATVDHLSK YLA
LR IALERRQQQEAGEPGGPGGASDTGGPDGCGGEGGGAGGGDGPEEPALPSLEGVSEKQ
YTIYIAPGGGAFTTLNGLTLELVNEKFWK VSRPLELCYAPTKDPK

RING

RNF2

MSQAVQTNGTQPLSKTWELSLYELQRTPEAITDGLEIVVSPRSLHSELM**CPICLDMLKN**
TMTTKECLHRFCADCIITALRSGNKECPTCRKKLVSKR**SLRDPNFDALISKIYPSRDEY**
EAHQERVLARINKHNNQALSHSIEEGLKIQAMNRLQRGKKQOIENSGAEDNGDSSHCS
NASTHSNQEAGPSNKRTKTSDDSGLELDNNAAMAIDPVMGASEIELVFRPHPTLMEKD
DSAQTRYIKTSGNATVDHLSKYLAURLALEELR**SKGESNQMNLDTASEK**QYTIYIATASG
QFTVLNGSFSLELVSEKYWK**VNKPMELYYAPTKEHK**

RING

BMI1

MHRTRRIKITEINPHLM**CVLCGGYFIDATTIIECLHSFCKTCIVRYLETSKYCPICD**VQV
HKTRPLLNR**SDKTLQDIVYKLVPLGFLKNEK**RRRDFYAAHPSADAANGSNEDRGEVADE
DKRIITDDEIISLSIEFFDQNRDRKVNKDKESKEEVNDKRYLRCPAAMTMHLRKFRLR
SKMDIPNTFQIDVMYEEEPKDYTYTLMIDYIYTWRRNGPLPLKYRVRPTCKRMKISHQR
DGLTNAGELESDSGSKANSPPAGGIPSTSSCLPSPSTPVQSPHPQFPHISSTMNGTNSNP
SGNHQSSFANRPRKSSVNGSSATSSG

RING

CBX4

MELPAVGEHVFAVESIEKKRIRKGRVEYLVKWRGWSPK**YNTWEPEENILDPR**LLIAFQNR
ERQEQLMGYRKRGPKPPLVVQVPTFARRSNVLTGLQDSSTDNRAKLDLGAQGGKQGHQY
ELNSKKHHQYQPHSKEGKPRPPGKSGKYYYQLNSKKHHYQDPDKMYDLQYQGGHKEAPS
PTCPDLGAKSHPPDKWAQGAGAKGYLGAVKPLAGAAGAPGKGSEKPPNGMMPAPKEAVT
GNGIGGKMKIVKNKKNKGRIVIVMSKYMENGMQAVKIKSGEVAEGEARSPSHKKRAADER
HPPADRTFKKAAGAEKKVEAPPKRREEEVSQVSDPQDAGSRKLSPTKEAFGEQPLQL
TTKPDLLAWDPARNTHTPPSHHPHPHHHHHHHHHHHHAVGLNLSHVRKRCLSETHGERE
PCKKRLTARISITPTCLGGSPAERPADLPPAAALRQPEVILLDSLDLDEPIDLRVKSRS
EAGEPPSSLQVKPETPASA AVAAAAAPT TTA EKPPAEA QDEPAESLSEFKPFFGNII
TDVTANCLTVTFK**EYVTV**

CBX8 MELSAVGER**VFAAEALLKRR**IRKGRMEYLVKWKGSQ**KYSTWEPEENILDAR**LLA
AFEEREREMELYGPKKRGPKPK**TFL**LKAQAKAKAKTYEFRSDSARGIRIPYPGR**SPQDLA**
STSRAREGLRNMGLSPPASSTSTSSTCRAEAPRDRDRDRDRDRERDRERERERERERE
RERERGTSRVDDKPSPPGDSKRRGPKPR**KELPDP**SQRPLGEP**SAGLGEYLK**GRKLLDTP
SGAGKFPAGHSVIQLAR**RQSD**DLV**QCGVT**SP**SSAEATGK**LAVDTFPARVIKHR**AAFLEAK**
GQALDPNGTRVRHGSPPSSGGGLYRDMGAQGGRP**SLIARIPVARILGDPEEESWSPSL**
TNLEKVVVTDVTSNFLTVTIK**ESNTDQGF**FEKR

Figure S7 : Peptides identifiés pour les protéines Polycomb à partir de RING1-TAP

RING1

MTTPANAQNASK**TWELSLYELHRTPEAIMDGTEIAVSPRSLHSELM****CPICLDMLKN****TMT**
TKECLHRFCSDCIVTALRSGNKECPTCRKKLVSKR**SLRDPNFDALISKIYPSREEYEAH**
QDRVLIRLSRLHN**QALSSSIEEGLRM**QAMHRAQRVRRIPGSDQTTTMSGGEGEPGEGE
GDGEDVSSDSAPDSAPGPAPKRPR**GGGAGGS**SVGT**GGGGTGGVGGGAGS**EDSGDRGGTLG
GGTLGPPSPPPGAPSPPEPGGEIELVFRPHPLLVKEGEYCQTRYVK**TTGNATVDHLSKYLA**
LRIALERRQQQEAGEPGGPGGGASDTGGPDGCGGEGGGAGGGDGPEEPALPSLEGVSEKQ
YTIYIAPGGGAFTTLNGSLTLELVNEKFWK**VSRPLELCYAPT**KDPK

RING

RNF2

MSQAVQTNGTQPLSKTWELSLYELQRTPEAITDGLEIVVSPR**SLHSELM****CPICLDMLKN**
TMTTKECLHRFCADCIITALRSGNKECPTCRKKLVSKR**SLRDPNFDALISKIYPSRDEY**
EAHQERVLARINKHNNQALSHSIEEGLKIQAMNRLQRGKKQOIENSGAEDNGDSSHCS
NASTHSNQEAGPSNKRTKTSDDSGLELDNNAAMAIDPVMGASEIELVFRPHPTLMEKD
DSAQTRYIKTSGNATVDHLSKYLAURLALEELR**SKGESNQMNLDTASEK**QYTIYIATASG
QFTVLNGSFSLELVSEKYWK**VNKPMELYYAPTKEHK**

RING

PHC2

Annexe 3: Article médecine/sciences

Vandamme J, Angrand PO.

La méthylation des histones n'est plus ce qu'elle était.

Med Sci (Paris). 2006 Apr;22(4):361-3

Avant propos

Pendant ma première année de thèse, la première histone déméthylase à domaine Jumonji a été découverte par l'équipe de Yi Zhang (Tsukada Y, 2006, Nature). Ce domaine est capable d'hydroxyler les groupements méthyles des lysines mono-, di- et/ou tri-méthylées. La découverte de toute une famille d'enzymes possédant ce domaine, conservé de la levure à l'Homme, remettait en question le dogme de l'irréversibilité de la méthylation des histones admis jusqu'à cette époque.

Compte tenu de l'avancée importante qu'à permis cette publication dans le domaine de la régulation transcriptionnelle et plus généralement dans le monde de l'épigénétique, nous avons pensé que l'exercice de rédaction d'un premier article en français pourrait me permettre, à la fois de faire le bilan des connaissances sur la déméthylation des histones à cette date, mais aussi de me confronter à la difficulté de l'écriture d'un article scientifique.

C'est donc avec l'aide de mon directeur de thèse que nous avons pu soumettre une « nouvelle » dans la revue scientifique francophone « Médecine/Science », qui a été publiée en avril 2006.



entrée de Ca^{2+} qui va induire une exocytose de glutamate. Cette sécrétion paracrine propage le signal dans d'autres populations cellulaires néocorticales via des récepteurs glutamatergiques NMDA ou AMPA (Figure 1). Le tandem canal Na^+ -échangeur $\text{Na}^+/\text{Ca}^{2+}$, présent dans une sous-population neuronale de la PP, va donc permettre la mise en place d'une exocytose en l'absence de canaux Ca^{2+} activés par dépolarisation [6]. Or, le glutamate sécrété est un agent important de contrôle de la prolifération des progéniteurs [2, 3] et de migration neuronale [4], ce qui suggère un impact physiologique d'envergure pour cette voie de signalisation immature. Une communication paracrine similaire, impliquant le GABA, a été décrite dans l'hippocampe [10]. Mais notre étude identifie pour la première fois l'existence d'une communication paracrine dans le néocortex à un stade aussi précoce et impliquant, cette fois, le glutamate. Le GABA ne semble pas intervenir dans cette voie de signalisation néocorticale [8].

Avant d'être les producteurs de potentiels d'action membranaires, les canaux Na^+ assurent donc des fonctions inhabituelles au cours du développement cortical. Leur expression restreinte à certaines cellules pionnières de la PP, dont les cellules de Cajal-Retzius [7], fait penser à une implication majeure dans le développement cortical. Cette implication semble d'autant plus pertinente que l'analyse préliminaire de l'activité calcique spontanée montre l'existence de réseaux neuronaux organisés et synchrones au cours du temps. La compréhension moléculaire de cette signalisation calcique précoce constitue donc une étape primordiale dans la compréhension du développement cortical embryonnaire. ♦

Neuronal activity before synaptogenesis: Na^+ channels, Ca^{2+} signalling and glutamatergic secretion.

Or « how to play the part when some famous actors are missing in the scene » ?

RÉFÉRENCES

- Benitez-Diaz P, Miranda-Contreras L, Mendoza-Briceno RV, et al. Prenatal and postnatal contents of amino acid neurotransmitters in mouse parietal cortex. *Dev Neurosci* 2003; 25: 366-74.
- Haydar TF, Wang F, Schwartz ML, et al. Differential modulation of proliferation in the neocortical ventricular and subventricular zones. *J Neurosci* 2000; 20: 5764-74.
- LoTurco JJ, Owens DF, Heath MJ, et al. GABA and glutamate depolarize cortical progenitor cells and inhibit DNA synthesis. *Neuron* 1995; 15: 1287-98.
- Bchar TN, Scott CA, Greene CL, et al. Glutamate acting at NMDA receptors stimulates embryonic cortical neuronal migration. *J Neurosci* 1999; 19: 4449-61.
- Moody WJ. The development of voltage-gated ion channels and its relation to activity dependent developmental events. *Curr Top Dev Biol* 1998; 39: 139-85.
- Picken-Bahney HL, Mandy WJ. Voltage-gated currents, dye and electrical coupling in the embryonic mouse neocortex. *Cereb Cortex* 2003; 13: 239-51.
- Albrieux M, Platel JC, Dupuis A, et al. Early expression of sodium channel transcripts and sodium current by Cajal-Retzius cells in the preplate of the embryonic mouse neocortex. *J Neurosci* 2004; 24: 1719-25.
- Platel JC, Boisseau S, Dupuis A, et al. Na^+ channel-mediated Ca^{2+} entry leads to glutamate secretion in mouse neocortical preplate. *Proc Natl Acad Sci USA* 2005; 102: 19174-9.
- Molnar Z, Blakemore C. How do thalamic axons find their way to the cortex? *Trends Neurosci* 1995; 18: 389-97.
- Demarque M, Represa A, Beq H, et al. Paracrine intercellular communication by a Ca^{2+} - and SNARE-independent release of GABA and glutamate prior to synapse formation. *Neuron* 2002; 36: 1051-61.

NOUVELLE

La méthylation des histones n'est plus ce qu'elle était

Julien Vandamme, Pierre-Olivier Angrand

► Dans le noyau, l'ADN eucaryote s'associe avec les histones (H2A, H2B, H3 et H4) pour former l'unité fondamentale de la chromatine : le nucléosome. L'architecture précise de la chromatine détermine si celle-ci est permissive ou non à la transcription et à d'autres événements qui dépendent de son organisation, tels que la réplication, la réparation de l'ADN ou la recombinaison. Alors que la grande majorité des nucléosomes sont constitués des mêmes histones, une immense diversité de nucléosomes

différents découle des modifications post-traductionnelles qui touchent les histones. Celles-ci peuvent être acétylées, méthylées, phosphorylées, ubiquitinylées ou sumoylées. Il est proposé que ces différentes modifications agissent en combinaison pour former un « code histone » [1-3]. Ce code est lu par des protéines non-histones qui se lient aux résidus modifiés des histones, et influencent alors l'organisation de la chromatine, la transcription ou la réplication. Certaines modifications comme

Institut de Recherche Interdisciplinaire,
Cnrs FRE 2963,
IRI @ Institut de Biologie de Lille,
1, rue du Professeur Calmette,
59021 Lille Cedex, France.
pierre-olivier.angrand@ibl.fr

l'acétylation ou la phosphorylation sont réversibles et dynamiques, et souvent associées à la régulation inductible de gènes individuels. En revanche, pendant de nombreuses années, la méthylation des histones était considérée comme une marque épigénétique stable et irréversible qui rendait compte du caractère héritable des états ouverts (euchroma-

tine) ou fermé (hétérochromatine) de la chromatine. Cette vision découlait de l'incapacité à mettre en évidence une activité histone déméthylase et des études qui suggéraient que la demi-vie de la méthylation des histones est similaire à celle des histones elles-mêmes [4]. Cependant, le dogme de l'irréversibilité de la méthylation des histones est maintenant remis en question. Des études ont montré que la méthylation pouvait même être extrêmement dynamique. Par exemple, la méthylation de la lysine 9 de l'histone H3 des nucléosomes localisés au niveau de gènes qui répondent à l'inflammation disparaît lors de leur activation, puis est restaurée

lors de la répression transcriptionnelle post-induction [5]. Depuis fin 2004, un nouveau pas a été franchi avec la découverte des protéines et des mécanismes directement impliqués dans la déméthylation des histones (Figure 1).

Les agents de déméthylation des histones

Les arginines sont sujettes à une conversion enzymatique qui engendre un acide aminé non conventionnel. La peptidylarginine désiminase 4 humaine (PADI4/PAD4) convertit les arginines monométhylées des histones H3 et H4 en citrulline par déméthylimination [6, 7]. Le fait que PADI4 transforme également

les arginines non méthylées en citrullines, qui ne peuvent plus être méthylées, suggère que le rôle de PADI4 serait d'éliminer les histones en tant que substrats des arginine méthyltransférases, et non seulement de déméthyliser les histones.

LSD1 (*lysine-specific demethylase 1*) est une amine oxydase nucléaire dépendante de l'adénine flavine dinucléotide (FAD) capable de déméthyliser la lysine 4 de l'histone H3 [8]. LSD1 catalyse une réaction oxydative de l'amine via le clivage oxydatif de la liaison α -CH du substrat pour former une imine en réduisant un cofacteur, la flavine. L'imine intermédiaire est ensuite hydrolysée spontanément pour produire du carbinolamine, groupe instable qui se transforme en formaldéhyde, régénérant ainsi une lysine déméthylée. Seules les lysines mono- et di-méthylées sont des substrats de LSD1. En effet, la formation de l'imine intermédiaire nécessite une lysine protonée, et les lysines tri-méthylées ne peuvent donc pas être substrats des amine oxydases. L'activité et la spécificité de déméthylation des lysines 4 de l'histone H3 dépend à la fois des autres modifications post-traductionnelles de l'histone H3, comme son état de phosphorylation en sérine 10 [9] et de la présence de cofacteurs. En effet, quand LSD1 est associé au cofacteur CoREST, LSD1 déméthyle la lysine 4 de l'histone H3 [10] (une marque de l'euchromatine), alors qu'associée avec le récepteur des androgènes, LSD1 déméthyle la lysine 9 de l'histone H3 [11] (une marque de l'hétérochromatine). Ainsi, la protéine LSD1 agit à la fois comme un co-activateur ou un co-répresseur transcriptionnel. Plus récemment, il a été proposé [12, 13] que la déméthylation des lysines pouvait se faire par hydroxylation des groupes méthyles. Cette réaction est catalysée par le domaine JmjC possédant une activité di-oxygénase dépendante de l' α -cétoglutarate et du fer divalent. La réaction utilise le Fe(II) pour activer une molécule de di-oxygène et former une molécule haute-

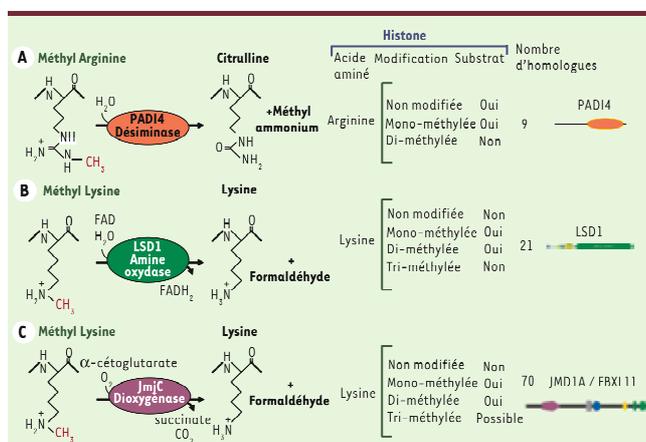


Figure 1. Mécanismes enzymatiques de la déméthylation et spécificité de substrat des histone déméthylases. A. La méthylation des arginines est supprimée par les arginine désiminases (PADI) pour engendrer une citrulline et du méthylammonium. Les enzymes PADI sont peu spécifiques et sont actives sur différentes arginines monométhylées des histones. Elles peuvent même convertir des arginines non-méthylées en citrullines. B. Deux mécanismes rendent compte de la déméthylation des méthyl-lysines. Les amine oxydases (de type LSD1) régénèrent une lysine et du formaldéhyde. La formation d'une imine intermédiaire par transfert de deux atomes d'hydrogène sur le cofacteur FAD nécessite un azote protoné, ce qui limite la réaction aux lysines mono- et diméthylées. C. Les di-oxygénases à domaine JmjC engendrent également une lysine et du formaldéhyde mais peuvent théoriquement déméthyliser les tri-méthyl-lysines. À la différence des histone déméthylases spécifiques des arginines, les histone déméthylases spécifiques des lysines ont une très grande spécificité de substrat: LSD1 pour les lysines 4 ou 9 de l'histone H3, selon le cofacteur associé, et FBXL11 pour la lysine 36 de l'histone H3. Le nombre de protéines humaines qui possèdent des domaines désiminases, amine oxydases ou JmjC similaires à ceux de PADI4, LSD1 ou FBXL11 est indiqué (<http://www.sanger.ac.uk/Software/Pfam/>). Seule la déméthylation des monométhyl-lysine est montrée (adapté de [14]).



ment réactive d'oxoferryl [Fe(IV)O], qui peut hydroxyler la méthyl-lysine ; la déméthylation produit alors du formaldéhyde. Les premières histone déméthylases de ce type caractérisées sont JHMD1A (FBXL11) et JHMD1B (FBXL10) qui déméthylent spécifiquement la lysine 36 de l'histone H3 sous sa forme mono- ou di-méthylée, mais pas sous sa forme tri-méthylée [13].

Alors que la découverte des premières histone déméthylases représente une avancée majeure dans la biologie de la chromatine et de l'épigénétique, elle soulève également de nombreuses questions. La méthylation des arginines peut être le résultat d'une induction hormonale, mais si leur déméthylation est associée à l'apparition d'une citrulline, comment une nouvelle induction peut-elle avoir lieu ? La citrulline constitue-t-elle une nouvelle marque du code histone ? Les amine oxydases (environ 20 membres) et les protéines à domaine

JmJc (environ 70 membres) répondent en termes de nombre à la complexité de la méthylation des lysines contrôlée par plus de 70 protéines à motif méthyl-transférase spécifique (le domaine SET). Toutefois, aucune protéine capable de déméthyliser les tri-méthyl-lysines n'a encore été identifiée, bien que la réaction d'hydroxylation des groupes méthyles en soit théoriquement capable. Et si la méthylation des histones est aussi dynamique, quels sont les mécanismes responsables de l'héritabilité de l'organisation de la chromatine et de son maintien au cours des mitoses successives ? ♦

Histone demethylation unravelled ?

RÉFÉRENCES

1. Strahl BD, Allis CD. The language of covalent histone modifications. *Nature* 2000 ; 403 : 41-5.
2. Jenewein T, Allis CD. Translating the histone code. *Science* 2001 ; 293 : 1074-80.
3. Ray-Gallet D, Gérard A, Polo S, Almouzni G. Variations sur le thème du « code histone ». *Med Sci (Paris)* 2005 ; 21 : 384-9.
4. Szyoet P, Shepherd GR, Hardin JM, Noland RJ. The distribution and turnover of labelled methyl groups in histone fractions of cultured mammalian cells. *Arch Biochem Biophys* 1972 ; 148 : 558-67.
5. Sacconi S, Natoli G. Dynamic changes in histone H3 Lys 9 methylation occurring at tightly regulated inducible inflammatory genes. *Genes Dev* 2002 ; 16 : 2219-24.
6. Cuthbert GL, Daujat S, Snowden AW, et al. Histone deimination antagonizes arginine methylation. *Cell* 2004 ; 118 : 545-55.
7. Wang Y, Wysocka J, Sayegh J, et al. Human PAD4 regulates histone arginine methylation levels via demethyliminination. *Science* 2004 ; 306 : 279-83.
8. Shi Y, Lan F, Matson C, et al. Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. *Cell* 2004 ; 119 : 941-53.
9. Fomenis F, Binda C, Vanoni MA, et al. Human histone demethylase LSD1 reads the histone code. *J Biol Chem* 2005 ; 280 : 41360-5.
10. Shi Y, Matson C, Lan F, et al. Regulation of LSD1 histone demethylase activity by its associated factors. *Mol Cell* 2005 ; 19 : 867-64.
11. Metzger E, Wissmann M, Yin N, et al. LSD1 demethylates repressive histone marks to promote androgen-receptor-dependent transcription. *Nature* 2005 ; 437 : 436-9.
12. Treweek SC, McLaughlin PJ, Allshire RC. Methylation: lost in hydroxylation? *EMBO Rep* 2005 ; 6 : 315-20.
13. Tsukada Y, Fang J, Erdjument-Bromage H, et al. Histone demethylation by a family of JmJc domain-containing proteins. *Nature* 2006 ; 439 : 811-6.
14. Kubicek S, Jenewein T. A crack in histone lysine methylation. *Cell* 2004 ; 119 : 903-6.

NOUVELLE

Pour maigrir, faisons de la bile

Pascal Ferré

Inserm U671,
Université Pierre et Marie Curie, Paris 6,
Centre Biomédical des Cordeliers,
15, rue de l'École de Médecine, 75270 Paris Cedex 06, France.
pferre@bhdj.jussieu.fr

► L'obésité est devenu un problème majeur de santé publique lié entre autres à la multitude des complications qu'elle engendre (par exemple, le diabète de type 2). Son développement est la conséquence de variations rapides et récentes de l'environnement nutritionnel et social. D'un point de vue énergétique, les causes de l'obésité sont simples. Si l'on a stocké des calories sous forme d'acides gras dans le tissu adipeux, c'est que, pendant une période plus ou moins longue, les apports caloriques (alimentation) ont dépassé les dépenses. Les dépenses peuvent être schématiquement divisées en une partie incompressible (métabolisme de base)

lié au fonctionnement obligatoire de nos cellules, et une partie variable, dépendante de paramètres comme l'absorption de nourriture (thermogenèse post-prandiale), l'activité physique et l'adaptation aux conditions climatiques. Si l'on veut maigrir, il faut donc jouer sur l'un ou l'autre des plateaux de la balance énergétique, les entrées ou les sorties. Augmenter l'activité physique se révèle pour la plupart des obèses très difficile. Peut-on alors envisager d'augmenter « artificiellement » la dépense énergétique (maigrir en regardant son émission préférée...) ? Dans la cel-

lule, la production d'énergie chimique sous forme d'ATP s'effectue dans la chaîne respiratoire des mitochondries. La chaîne respiratoire oxyde les coenzymes réduits provenant de l'utilisation des substrats énergétiques (Figure 1) et pompe des protons à l'extérieur de la matrice mitochondriale. L'énergie créée par le gradient de protons sert ensuite à synthétiser de l'ATP (la monnaie énergétique cellulaire) à partir d'ADP, grâce à l'ATP synthase qui peut être considérée comme un dissipateur du gradient de protons puisqu'elle permet

Annexe 4 : article BMC Cell Biology

Souza PP, Völkel P, Trinel D, Vandamme J, Rosnoblet C, Héliot L, Angrand PO.

The histone methyltransferase SUV420H2 and Heterochromatin Proteins HP1 interact but show different dynamic behaviours.

BMC Cell Biol. 2009 Jun 1;10:41.

Avant propos

J'ai eu l'opportunité d'apporter ma contribution au projet scientifique du Dr Patricia Pellegrino da Souza, qui a réalisé un stage post-doctoral de 2 ans dans notre équipe pendant ma thèse. Initialement axé sur la dynamique des protéines HP1 et des méthyl-transférases d'histones SUV420H1/2 dans les cellules de mammifères en culture, ce projet s'est élargi à une analyse biochimique des protéines associées à SUV420H2. C'est ainsi que nous avons réalisé la purification TAP de cette enzyme à partir de cellules HeLa. Compte tenu de sa localisation au niveau de l'hétérochromatine, et du faible rendement de la purification, nous avons conclu que la majorité des protéines TAP restait fixée à la chromatine et que les protéines co-éluées représentaient les partenaires de la fraction soluble de SUV420H2 dans la cellule. Parmi les protéines identifiées, les plus abondantes étaient les protéines HP1, en particulier CBX3.

Nous avons validé ces résultats par GST-pulldown, la région d'interaction entre SUV420H2 et les 3 isoformes HP1 a été réduite à 88 acides aminés (aa 347-435), qui sont également nécessaires et suffisants pour la fixation de SUV420H2 au niveau de l'hétérochromatine.

Des études de recouvrement de fluorescence après photoblanchiment (FRAP) ont permis de démontrer qu'à l'inverse des protéines HP1 qui ont une dynamique élevée au niveau de l'hétérochromatine (même si elle varie en fonction du type cellulaire) : 70 à 80% de recouvrement de fluorescence après 30 secondes, les protéines SUV420H2 restent fixées de manière beaucoup plus stable à la chromatine (10% après 30 secondes). D'après nos résultats, nous suggérons que SUV420H2 présentes de manière soluble dans le noyau sont associées aux HP1, et que ce complexe est recruté au niveau des lysines 9 méthylées de l'histone H3, en particulier au niveau de l'hétérochromatine constitutive. Ensuite, les protéines SUV420H2 s'associent de manière stable sur la chromatine, à l'inverse des HP1 qui ont une dynamique relativement importante par rapport à ce qui était admis il y a encore quelques années.

Research article

Open Access

The histone methyltransferase SUV420H2 and Heterochromatin Proteins HPI interact but show different dynamic behaviours

Patricia P Souza^{1,3}, Pamela Völkel¹, Dave Trinel², Julien Vandamme^{1,4}, Claire Rosnoblet¹, Laurent Héliot² and Pierre-Olivier Angrand*¹

Address: ¹Chromatinomics, Interdisciplinary Research Institute, Université des Sciences et Technologies de Lille/CNRS USR 3078, Parc Scientifique de la Haute Borne, 50 Avenue Halley, F-59658 Villeneuve d'Ascq, France, ²Biophotonique Cellulaire Fonctionnelle, Interdisciplinary Research Institute, Université des Sciences et Technologies de Lille/CNRS USR 3078, Parc Scientifique de la Haute Borne, 50 Avenue Halley, F-59658 Villeneuve d'Ascq, France, ³New York University Langone Medical Center, Pathology Department, 550, First Avenue, New York, NY, 10016, USA and ⁴Biotech Research & Innovation Centre, University of Copenhagen, Ole Maaloes Vej 5, DK-2200 Copenhagen, Denmark

Email: Patricia P Souza - patricia.souza@nyumc.org; Pamela Völkel - pamelavoelkel@iri.univ-lille1.fr; Dave Trinel - dave.trinel@iri.univ-lille1.fr; Julien Vandamme - julien.vandamme@bric.ku.dk; Claire Rosnoblet - claire.rosnoblet@iri.univ-lille1.fr; Laurent Héliot - laurent.heliot@iri.univ-lille1.fr; Pierre-Olivier Angrand* - pierre-olivier.angrand@iri.univ-lille1.fr

* Corresponding author

Published: 1 June 2009

Received: 9 January 2009

BMC Cell Biology 2009, 10:41 doi:10.1186/1471-2121-10-41

Accepted: 1 June 2009

This article is available from: <http://www.biomedcentral.com/1471-2121/10/41>

© 2009 Souza et al; licensee BioMed Central Ltd.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

Background: Histone lysine methylation plays a fundamental role in chromatin organization and marks distinct chromatin regions. In particular, trimethylation at lysine 9 of histone H3 (H3K9) and at lysine 20 of histone H4 (H4K20) governed by the histone methyltransferases SUV39H1/2 and SUV420H1/2 respectively, have emerged as a hallmark of pericentric heterochromatin. Controlled chromatin organization is crucial for gene expression regulation and genome stability. Therefore, it is essential to analyze mechanisms responsible for high order chromatin packing and in particular the interplay between enzymes involved in histone modifications, such as histone methyltransferases and proteins that recognize these epigenetic marks.

Results: To gain insights into the mechanisms of SUV420H2 recruitment at heterochromatin, we applied a tandem affinity purification approach coupled to mass spectrometry. We identified heterochromatin proteins HPI as main interacting partners. The regions responsible for the binding were mapped to the heterochromatic targeting module of SUV420H2 and HPI chromoshadow domain. We studied the dynamic properties of SUV420H2 and the HPI in living cells using fluorescence recovery after photobleaching. Our results showed that HPI proteins are highly mobile with different dynamics during the cell cycle, whereas SUV420H2 remains strongly bound to pericentric heterochromatin. An 88 amino-acids region of SUV420H2, the heterochromatic targeting module, recapitulates both, HPI binding and strong association to heterochromatin.

Conclusion: FRAP experiments reveal that in contrast to HPI, SUV420H2 is strongly associated to pericentric heterochromatin. Then, the fraction of SUV420H2 captured and characterized by TAP/MS is a soluble fraction which may be in a stable association with HPI. Consequently, SUV420H2 may be recruited to heterochromatin in association with HPI, and stably maintained at its heterochromatin sites in an HPI-independent fashion.

Background

Eukaryotic DNA is packaged within the nucleus through its association with histone proteins forming the fundamental repeating unit of chromatin, the nucleosome. The nucleosome consists of 146 bp of DNA wrapped around a histone core octamer composed of two each of histones H2A, H2B, H3, and H4 [1]. Histone C- and N-terminal tails are flexible, protrude from the nucleosome octamer structure, and are subjected to post-translational modifications, including acetylation, methylation, phosphorylation, ubiquitination or sumoylation. Among these modifications, histone lysine methylation patterns have been associated with distinct chromatin states and are proposed to be major epigenetic marks that could extend the information potential of the genetic code by fixing the chromatin organization in a heritable manner (for a review [2]).

In particular, constitutive heterochromatin, considered as the part of the genome that is gene poor, transcriptionally silent and highly condensed in interphase cells, is characterized to harbour nucleosomes rich in trimethylation at lysine 9 of histone H3 (H3K9me3), trimethylation at lysine 20 of histone H4 (H4K20me3) and monomethylation at lysine 27 of histone H3 (H3K27me1) [3-5]. The histone methyltransferases SUV39H1 and SUV39H2 play a crucial role in the initial steps of heterochromatin formation in mammals by selective trimethylation of H3K9 [3,6,7]. Indeed, mice that are deficient for SUV39H activities were shown to display impaired H3K9 trimethylation at pericentric heterochromatin and were subjected to chromosomal instability [8]. The molecular mechanisms by which SUV39H1 and SUV39H2 are recruited at heterochromatin are still unknown but were suggested to be mediated by direct or indirect association with components of a RNA interference pathway [9]. According to current models, H3K9me3 marks placed by SUV39H activities stabilize heterochromatin protein 1 (HP1) binding at heterochromatin [10,11], and HP1 proteins would then recruit the histone methyltransferases SUV420H2 and SUV420H1 which in turn, trimethylate H4K20 [5,12,13]. At present, it is unclear whether SUV420H histone methyltransferases interact only temporally with chromatin to methylate H4K20 or participate in a more stable multiprotein complex together with HP1 or other chromatin proteins to support a stable heterochromatin structure.

Interestingly, maintenance of stable heterochromatin domains in living cells involves the transient binding and dynamic exchange of HP1 from chromatin [14-16] indicating that heterochromatin is not a static and inaccessible higher order conformation but is a dynamic domain of chromatin. In contrast to HP1, SUV39H1 has a significantly slower exchange rate and a substantial fraction

immobile at heterochromatin [17], and nothing is known about the dynamics of the SUV420H class of histone methyltransferases.

To gain further insights into the role that SUV420H2 plays in heterochromatin, we first used a directed proteomic analysis of SUV420H2-binding proteins in cells using the tandem affinity purification (TAP)-mass spectrometry (MS) methodology [18]. We showed that HP1 proteins are the main SUV420H2 interacting partners. We then investigated the *in vivo* kinetics of SUV420H2 in mouse L929 cells and compared them to those of HP1 proteins. Using fluorescence recovery after photobleaching (FRAP) analysis we showed that in contrast to HP1 proteins, SUV420H2 has a slow exchange rate and is largely strongly bound at pericentric heterochromatin. Furthermore, we mapped the domain responsible for stable association to chromatin to a region of 88 amino-acids of SUV420H2, previously defined as the heterochromatic targeting module by Schotta et al. [5].

Results and discussion

SUV420H2 interacts with HP1 proteins

Pericentric regions, as well as telomeres, are part of the constitutive heterochromatin, which is characterized by hypermethylation of DNA, hypoacetylation and hypermethylation of histones. In particular, trimethylation of H3K9 (H3K9me3) and trimethylation of H4K20 (H4K20me3) by the SUV39H and SUV420H histone methyltransferases respectively, mark pericentric chromatin [3-5,19]. In order to gain insight into the mechanisms involved in the recruitment of SUV420H enzymes at heterochromatin and identify their interacting partners, we applied the tandem affinity purification (TAP) technology coupled to tandem mass spectrometry (MS/MS), since this approach has proven to efficiently allow the characterization of protein complexes from different cells in culture or organisms [20]. TAP/MS analysis of protein associations around SUV420H2 was performed from HeLa cells, because we have assembled a database of more than 30 TAP/MS experiments from proteins involved in chromatin regulation in this cell line. This dataset allows reliable assessment of a given interaction (data not shown).

A retrovirus gene transfer strategy was used to generate a cell pool expressing TAP-tagged SUV420H2 fusion protein. Expanded cell pools were subjected to tandem affinity purification, a procedure consisting of two specific binding and two specific elution steps under mild conditions, which preserve the integrity of non-transient protein-protein interactions [18]. The affinity purified complexes were resolved on SDS-PAGE and Coomassie stained (Fig. 1A). Then, proteins were identified by peptide sequence determination using tandem mass spec-

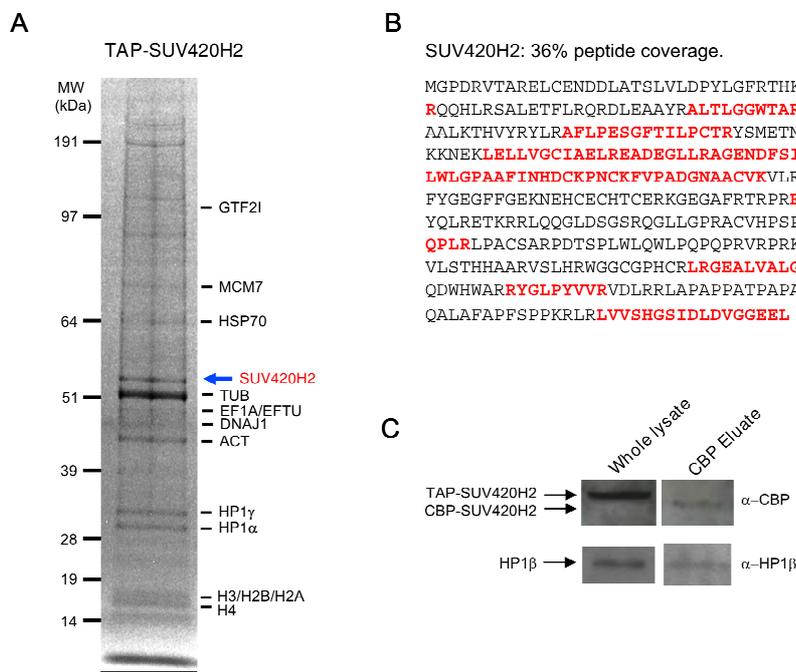


Figure 1

Identification of SUV420H2 associated proteins using TAP/MS. (A) TAP-SUV420H2 purification from HeLa cells. TAP-SUV420H2 protein complexes were purified from HeLa cells, separated by SDS-PAGE, and stained with colloidal Coomassie. Some of the co-purified proteins identified by LC-MS/MS are indicated. **(B)** SUV420H2 peptide coverage is 36%. The identified peptides are indicated in red bold on the SUV420H2 protein sequence. Peptides are dispersed over the full sequence of the protein. **(C)** Western blotting showing the presence of HP1 β in the TAP-SUV420H2 purification. Total protein extracted from HeLa cells expressing TAP-SUV420H2 (Whole lysate; 50 μ g of proteins) and TAP purified sample (CBP Eluate; 20% of the material used for MS identification) were separated by SDS-PAGE and proteins detected by Western blotting using specific antibodies. Tagged SUV420H2 was revealed using an anti-CBP antibody (α -CBP), while HP1 β was revealed by an anti-HP1 β antibody (α -HP1 β). Note that the TEV protease-mediated cleavage of the Protein A moiety of the TAP tag during the purification procedure increases the gel mobility of the tagged SUV420H2 protein in the CBP eluate as compared to whole extract.

trometry (LC-MS/MS). Most proteins were characterized by the identification of several peptides [see Additional files 1 and 2]. Details on protein identification procedures are given in the Methods section. The tagged SUV420H2 protein represents a prominent band on the gel and was identified by peptides covering 36% of its total sequence (Fig. 1B). However, SUV420H2 was not the most abundant protein recovered by TAP/MS since tubulin, which is considered as a TAP/MS contaminant, remains the most prominent identified protein. This relative low recovery of

SUV420H2 by TAP is probably due to its strong association to chromatin (see below). Sonication or addition of intercalating agents such as ethidium bromide [21], prior to protein extraction, did not improve the recovery rate of SUV420H2 by TAP/MS (data not shown). In spite of a weak recovery of the bait, heterochromatin proteins (HP) 1 were identified as specific SUV420H2 interactors (Fig. 1A) by identified peptides covering 53% (HP1 γ - CBX3) and 15% (HP1 α - CBX5) of their total sequence, respectively. The third member of the HP1 protein family, HP1 β

(CBX1) was not formally identified in the purification. However, the peptide IIGATDSSGELMFLMK identified in the SUV420H2 purification, is common to HP1 γ and HP1 β sequences. Moreover, HP1 β was identified by Western blotting using specific anti-HP1 β antibodies in the TAP-SUV420H2 purified samples (Fig. 1C). Consequently, an interaction between SUV420H2 and HP1 β might occur as well. None of the other proteins identified by TAP/MS appeared to be specific to the SUV420H2 purification when compared to our dataset. Thus, we conclude that SUV420H2 mainly interacts with the HP1 isoforms *in vivo*; HP1 α , HP1 γ and HP1 β as well.

Shotta et al [5] mapped a HP1-binding domain of murine Suv420h2 in the AA349–441 protein region referred as the heterochromatic targeting module, once this domain is also responsible for Suv420h2 targeting at heterochromatin. To confirm that such a functional module is also conserved in the human ortholog, the corresponding region (AA347–435) from human SUV420H2 was expressed in *E. coli* as a recombinant GST fusion protein and bound to glutathione-Sepharose beads. Full-length TAP-tagged HP1 α , HP1 β and HP1 γ were expressed in HEK293 cells and incubated with GST-SUV420H2 [347–435] fragments. After extensive washings, bound proteins were eluted with glutathione, separated on SDS-PAGE and Western blots were probed with peroxidase coupled anti-peroxidase antibody for detection of TAP-tagged HP1 isoforms. Figure 2A shows that the SUV420H2 [347–435] fragment interacts with all three HP1 proteins. In contrast, recombinant GST-GFP fusion protein does not interact with HP1, excluding unspecific binding to the GST-tag. Thus, the heterochromatic targeting module from human SUV420H2, as previously described for its mouse counterpart [5], interacts with all three HP1 isoforms, HP1 α , HP1 β and HP1 γ . Interestingly, this domain of SUV420H2 contains a PYVRV sequence, which is similar to the pentameric sequences found through a phage display analysis with the *Drosophila melanogaster* HP1 chromoshadow domain [22] and present in a number of proteins interacting to mammalian HP1 chromoshadow domains [23].

To define the HP1 regions involved in the binding to SUV420H2, TAP-tagged mutants of HP1 γ were transiently expressed in HEK293 cells. Nuclear protein extracts were incubated with the recombinant GST-SUV420H2 [347–435] heterochromatic targeting module (Fig. 2B). Deletion of HP1 γ chromoshadow domain (CSD) abolishes HP1 γ dimerization as expected [24–26], but also the interaction with the SUV420H2 [347–435] fragment, whereas deletion of lysine 105 to lysine 107 within the HP1 γ hinge region does not affect binding to the SUV420H2 heterochromatic targeting module. Thus, the HP1 CSD is required for SUV420H2-HP1 interaction. In order to test whether the HP1 CSD is sufficient for such an interaction,

a GST-CSD fusion protein was expressed in *E. coli*, bound to glutathione-Sepharose and incubated with *in vitro* translated TAP-SUV420H2 [347–435] or TAP-GFP proteins. Figure 2C shows that the SUV420H2 [347–435] fragment specifically interacts with GST-CSD, but not GST-GFP. Taken all together, these results indicate that the interaction between SUV420H2 and HP1 proteins is mediated by the binding of the heterochromatic targeting module to the chromoshadow domain of the respective proteins.

HP1 is a dynamic component of heterochromatin and its dynamics depend on the cell cycle

To assay for protein dynamics at heterochromatin, fluorescence recovery after photobleaching (FRAP) [27] experiments were performed in live murine L929 cells. These cells were used in our imaging analyses, because mouse fibroblastic cells present a defined distribution of heterochromatic domains revealed by Hoechst staining patterns as bright fluorescent nuclear regions (Fig. 3A).

GFP-HP1 expressing cells were generated using retroviral gene transduction. Fusion proteins of GFP with HP1 α , HP1 β and HP1 γ were shown to localize to these heterochromatin foci (Fig. 3A). The endogenous HP1 counterparts were shown to be localized at the same nuclear domains, as revealed by immunocytochemistry using specific anti-HP1 antibodies ([28] and data not shown). Expression of GFP-HP1 fusion proteins was analyzed by Western blotting using specific antibodies for each HP1 isoform (Fig. 3B). GFP-HP1 fusion proteins were expressed as full-length proteins at levels significantly higher (GFP-HP1 β) or similar (GFP-HP1 γ and GFP-HP1 α) to those of their endogenous counterparts.

Defined areas of $\sim 1.3 \mu\text{m}$ of diameter were irreversibly photobleached by five excitation pulses of 336 ms. Fluorescence recovery in the same areas was imaged at regular time intervals (Fig. 3C and see Methods section). The relative fluorescence intensity values for each experiment were collected and plotted in a FRAP curve after correction for photobleaching. The $t_{1/2}$ value of fluorescence recovery, which is a measure for the speed of replacement of molecules in the bleached area by molecules from the environment was determined, as well as the intensity of fluorescence recovery after 30 s, which is a measure of the mobile fraction of molecules (Fig. 3D and 3E). As previously reported by others in other cell types [14–16,29], Figure 3C–D shows that HP1 proteins are highly dynamic components of heterochromatin in L929 cells. The high mobility of ectopically expressed GFP-HP1 proteins might not be a consequence of protein overexpression since similar observations have been obtained in different experimental settings [14–16,29]. For instances, Schmiedeberg et al [16] determined similar $t_{1/2}$ values from Hep-2 cells

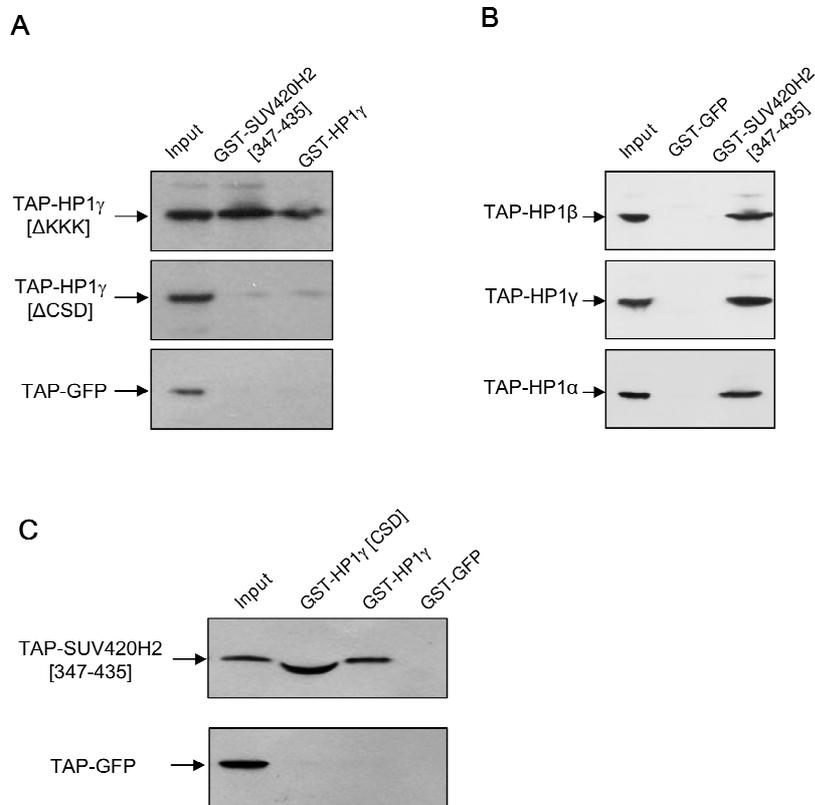


Figure 2
Identification of SUV420H2 and HP1 interacting domains by in vitro GST-pull down. (A) The heterochromatic targeting module of SUV420H2 interacts with HP1 proteins. GST-tagged truncation SUV420H2 [347-435], as well as control GST-GFP fusion were bound to glutathione-Sepharose and incubated with nuclear extracts from HEK293 cells expressing TAP-tagged HP1 α , HP1 β , or HP1 γ as indicated. After extensive washings, proteins were separated on SDS-PAGE and Western blot probed to reveal TAP-tagged proteins. (B) The chromoshadow domain of HP1 γ is required for its interaction with SUV420H2. GST-SUV420H2 [347-435], as well as control GST-HP1 γ fusions were bound to glutathione-Sepharose and incubated with nuclear extracts from HEK293 cells expressing TAP-HP1 γ truncations or TAP-GFP control. HP1 γ [ΔKKK] harbours a deletion from lysine 105 to lysine 107 within the hinge region, whereas HP1 γ [ΔCSD] contains the 114 N-terminus amino-acids of HP1 γ but not the chromoshadow domain. Loss of the chromoshadow domain abolishes the interaction with the heterochromatic targeting module of SUV420H2 and HP1 γ oligomerization. (C) The chromoshadow domain (CSD) of HP1 γ interacts with the SUV420H2 [347-435] region. GST-HP1 γ [CSD] consisting of the HP1 γ chromoshadow domain (AA 115-183), as well as control GST-HP1 γ and GST-GFP fusions were bound to glutathione-Sepharose and incubated with in vitro translated TAP-SUV420H2 [347-435] fusion protein or TAP-GFP control. Bound TAP-tagged proteins are revealed by Western blotting using the peroxidase-anti-peroxidase (PAP) antibody (Sigma) which recognizes the protein A moiety of the TAP tag.

expressing GFP-HP1 proteins at levels lower than those of their endogenous counterparts. It indicates that HP1 dynamic parameters do not depend on the GFP-HP1 expression levels. However, we observed that for each of the three GFP-HP1 isotypes, $t_{1/2}$ values from one photobleached cell to another show a significantly high variability (Fig. 3E).

In order to test whether this large distribution of $t_{1/2}$ values could reflect different states of the cells, we performed FRAP experiments after cell synchronization (Fig. 4). GFP-HP1 expressing cells were blocked in G0 by serum starvation for 72 hours, or in G2/M transition using a nocodazole treatment for 24 hours [30]. Cell synchronization was verified by fluorescent-activated cell sorter (FACS) analysis in presence of propidium iodide [31], and FRAP performed. Figure 4A–B shows that, in all cases, cell synchronization reduces the variability of the $t_{1/2}$ values. After serum deprivation, the $t_{1/2}$ was ~ 11.3 s, ~ 9.4 s, and 11.2 s for HP1 α , HP1 β and HP1 γ , respectively, but only ~ 7.4 s, ~ 6.8 s, and ~ 7.1 s after the nocodazole treatment. Both serum starvation and nocodazole treatment did not affect GFP-HP1 protein localization (Fig. 4A–B) or protein expression (Fig. 4C). This suggests that HP1 proteins at heterochromatin have faster dynamics in G2 rather than in G0 phase of the cell cycle.

In order to demonstrate that these differences in HP1 dynamics were due to intrinsic properties of HP1 but not a consequence of the treatments, similar experiments were performed with NSD3S, another protein associated to heterochromatin. Based on alternative splicing events, the NSD3 gene encodes two different proteins [32]: a long isoform (NSD3L) related to the histone methyltransferase Nsd1, and a short isoform (NSD3S) which does not contain the SET domain and the PHD fingers, but still harbouring a PWWP motif. NSD3S was stably expressed as a GFP fusion protein in L929 cells using retroviral transduction. Figure 5A shows that NSD3S is targeted at pericentric heterochromatin. FRAP experiments revealed that NSD3S is as dynamic as HP1 at heterochromatin (Fig. 5B–C). GFP-NSD3S fluorescence recovery reached $\sim 80\%$ after 30 s with a $t_{1/2}$ value ~ 8.4 s. These dynamic parameters are similar to those of HP1 proteins (Fig. 3D), indicating that HP1 proteins are not the only mobile components associated to heterochromatin with such a highly dynamic exchange rate. Next, FRAP experiments after cell synchronization were performed on GFP-NSD3S expressing L929 cells. Figure 5D–E shows that $t_{1/2}$ values are relatively similar after serum starvation and nocodazole treatment. This result indicates that the treatments used to induce cell synchronization do not modify significantly NSD3S dynamics at heterochromatin. Consequently, the differences of mobility in HP1 protein dynamics during the cell cycle are indeed due to their intrinsic properties, but not to culture conditions.

SUV420H2 is a stable component of heterochromatin

To gain insights into the dynamic properties of SUV420H2 and its association to heterochromatin, FRAP experiments were performed on L929 stably expressing a GFP-SUV420H2 fusion protein (Fig. 6C–D). In contrast to HP1 proteins, GFP-SUV420H2 fluorescence recovery reached only 10% as a maximum after 60 s, suggesting that $\sim 90\%$ of the SUV420H2 fraction is stably bound within heterochromatic domains. Even after several minutes, the intensity of fluorescence recovery remains weak (see below, Figure 7). Together, these results indicate that SUV420H2 is more stably associated to heterochromatin than HP1 proteins. This suggests that in addition to its function as a histone methyltransferase, SUV420H2 might have a structural role in chromatin organization.

To identify the protein domains involved in stable association of SUV420H2 at heterochromatin, we stably expressed various protein deletions as GFP fusions in L929 cells. Figure 6A shows a schematic representation of the protein truncations used. Consistent with previous data [5], the N-terminus of SUV420H2 comprising the catalytically active SET domain is dispersed within the nucleus, whereas the C-terminal part of the protein is associated with heterochromatin that corresponds to Hoechst dense staining (Fig. 6B). Moreover, the 88 amino-acids region comprising the heterochromatic targeting module of SUV420H2 is sufficient to recapitulate full length protein localization. Next, FRAP studies were performed in L929 cells expressing SUV420H2 C-terminal-GFP fusion proteins still located at heterochromatin foci. The C-terminal truncation as well as the heterochromatic targeting module alone, show kinetic properties similar to those of SUV420H2 full length (Fig. 6C–D). At the opposite, the SUV420H2 [1–280] protein region comprising the SET domain, is highly mobile. Thus, in contrast to SUV39H1 for which the SET domain has been shown to contribute to heterochromatin binding [17], all the dynamic characteristics of SUV420H2 are restricted to the 88 amino-acids heterochromatic targeting module.

So far, the mechanism by which the heterochromatic targeting module is tightly bound to chromatin remains elusive. Using a GST-pull down-mass spectrometry strategy, the only proteins found to interact *in vitro* with this domain were the HP1 proteins (data not shown), but other approaches are ongoing to unravel the molecular mechanisms involved in the linkage of SUV420H2 to heterochromatin.

Next, we compared the SUV420H2 kinetics with the one of H2AFY. H2AFY (macroH2A1) is a variant of histone H2A, enriched at the inactive X chromosome in female mammals and found at other chromosomal locations [33]. Since H2AFY is more resistant to salt extraction than canonical H2A [34], it is expected that H2AFY might be in

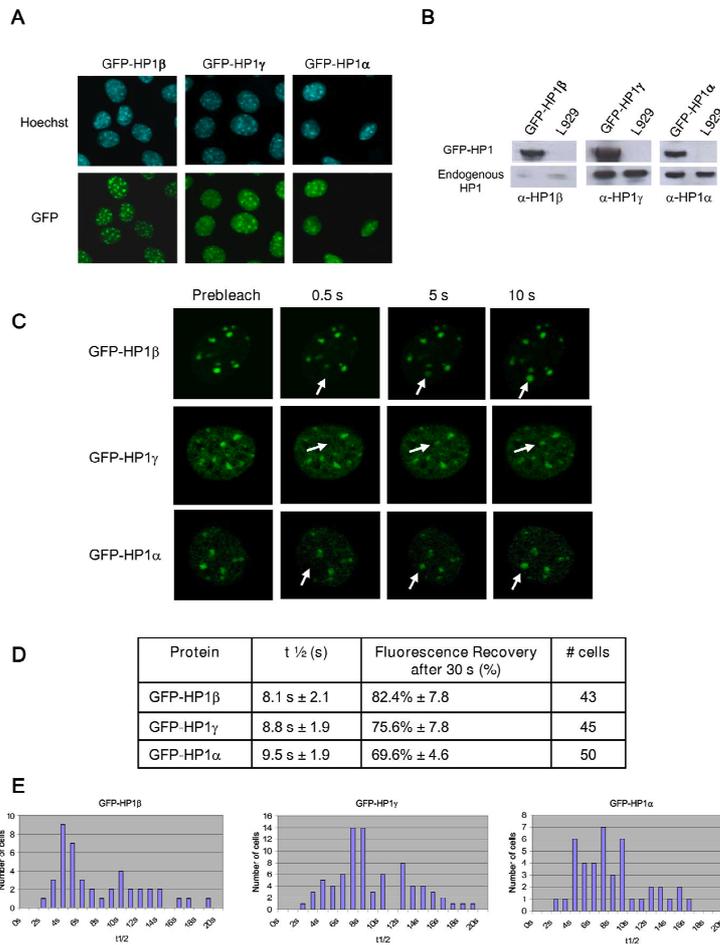


Figure 3
FRAP analysis of GFP-HP1 shows that HP1 proteins are dynamic components of heterochromatin. (A) Nuclear localization of GFP-tagged proteins. GFP-HP1 expressing L929 cells were fixed with paraformaldehyde and localization of fusion protein was visualized with fluorescence microscopy. Dense foci seen by Hoechst staining represent pericentric heterochromatin regions. (B) Detection of GFP-HP1 and endogenous HP1 isoforms in stable cell lines and in L929 cells by Western blotting using specific antibodies for each isoform (α-HP1β, α-HP1γ and α-HP1α). (C) Time-lapse series of confocal images of living cells expressing GFP-HP1β, GFP-HP1γ or GFP-HP1α, as indicated. Heterochromatic foci were selected and photobleached. Images were recorded before (Prebleach) and at different time intervals after the bleach. Arrows indicate the photobleached areas. (D) Parameters of HP1 protein dynamics from FRAP curves. The half-time of fluorescence recovery (t1/2), the percentage of fluorescence recovery after 30 s, and the number of FRAP curves used are indicated. (E) Representation of the number of cells characterized by a given t1/2 value. A single area was photobleached per cell.

more stable association with H3-H4 and/or DNA than H2A. H2AFY was stably expressed as a GFP fusion protein in L929 cells using retroviral transduction. Figure 7A shows that H2AFY is targeted at foci identified by Hoechst staining as pericentric heterochromatin regions, like SUV420H2. FRAP experiments performed over more than 30 min showed that fluorescence recovery after 20 min is ~22% and ~7% for GFP-SUV420H2 and GFP-H2AFY, respectively (Fig. 7B-C). Thus, SUV420H2 is stably bound to heterochromatin, but not as strongly as the nucleosome component, H2AFY.

Conclusion

The SUV420H family of histone methyltransferases catalyzes trimethylation of histone H4K20 which is characteristic to pericentric heterochromatin. So far, little is known about the mechanisms by which this protein class is recruited to and interacts with chromatin.

Here, we applied the TAP/MS approach in order to identify proteins associated to the histone methyltransferase SUV420H2. In our purification, a relatively weak amount of SUV420H2 was recovered, presumably because the protein is tightly bound to chromatin. We assume that the SUV420H2 captured by TAP mainly correspond to the nucleoplasm soluble fraction. Most of the proteins identified by TAP/MS appeared to be nonspecific and recovered in a number of unrelated purifications (data not shown). However, members of the HP1 family were found as specific SUV420H2-binding partners. Since only part of the SUV420H2 population was purified, we cannot exclude that a number of SUV420H2 interacting partners, and in particular the proteins associated to the SUV420H2 fraction the most tightly bound to chromatin, were not identified.

We used *in vitro* GST-pull down assays, to map the domains of interaction between SUV420H2 and HP1 proteins. As previously described for the mouse Suv420h2 ortholog [5], an 88 amino-acids region defined as the heterochromatic targeting module of SUV420H2 interacts with HP1 α , HP1 β and HP1 γ proteins. Using deletion mutants, the HP1 chromoshadow domain was shown to be sufficient for binding to the SUV420H2 heterochromatic targeting module.

To investigate the mechanisms by which SUV420H2 could be recruited at heterochromatin and the possible role of HP1 proteins in this recruitment, protein dynamics were studied using fluorescence recovery after photobleaching (FRAP) in living cells [27]. For a long time, heterochromatin was considered as a stable structure maintained in an inaccessible conformation that would exclude transcriptional activation. This concept has changed when HP1 was shown by FRAP analysis to be

highly mobile [14,15]. In our study, we show that NSD3S is another protein transiently bound to pericentric heterochromatin, reinforcing the idea that the dynamic nature of heterochromatin is essential to its function. Along our study, we also showed that HP1 exchange kinetics is modified during the cell cycle, indicating that heterochromatin organization is modulated during the cell cycle progression. However, not all the proteins bound to pericentric heterochromatin are in continuous flux with the nucleoplasm since Krouels et al. showed that a substantial fraction of the histone methyltransferase SUV39H1 is stably associated with chromatin [17]. In this study, we demonstrated that the major part of SUV420H2 is also tightly bound at pericentric heterochromatin, indicating that in addition to its histone methyltransferase activity, SUV420H2 might have a structural role in chromatin organization. We showed that the heterochromatic targeting module is sufficient to fully recapitulate SUV420H2 dynamic properties. Then, in contrast to SUV39H1 [17], the SET domain of SUV420H2 is not required for stable binding of the protein to pericentric heterochromatin.

Genetic studies indicated that H3K9 trimethylation would precede H4K20 trimethylation, and that HP1 which recognizes H3K9 trimethylation could target SUV420H2 class of histone methyltransferases to heterochromatin [5]. Our data confirm such an interaction *in vitro* and *in vivo*. However, FRAP experiments reveal that in contrast to HP1, SUV420H2 is strongly associated to pericentric heterochromatin. Thus, the fraction of SUV420H2 captured and characterized by TAP/MS is the soluble fraction which may be in a stable association with HP1. Consequently, we speculate that SUV420H2 is recruited to heterochromatin in association with HP1, but stably maintained at its sites in an HP1-independent fashion.

Methods

Constructs and vectors engineering

Full-length open reading frames of all cDNAs, as well as truncated and mutant cDNAs were PCR amplified from IMAGE cDNA clones purchased from Deutsches Ressourcenzentrum für Genomforschung GmbH (RZPD) and cloned into the GATEWAY entry vector pDONR201 (Invitrogen).

Moloney murine leukemia virus-based vectors were generated from pLNCX2 (Clontech) in which an IRES-pac cassette from pIRESpuro (Clontech) and conferring resistance to puromycin was introduced. Then, GATEWAY compatible target vectors for expression of TAP-tagged (pRP-NTAP-GW) or GFP-tagged (pRP-NGFP-GW) proteins were engineered by inserting GATEWAY and TAP (from pZome-1-N, Cellzome) or GFP (from pEGFP-N1, Clontech) cassettes, respectively. For bacterial expression of the GST-tagged proteins, a GATEWAY cassette was

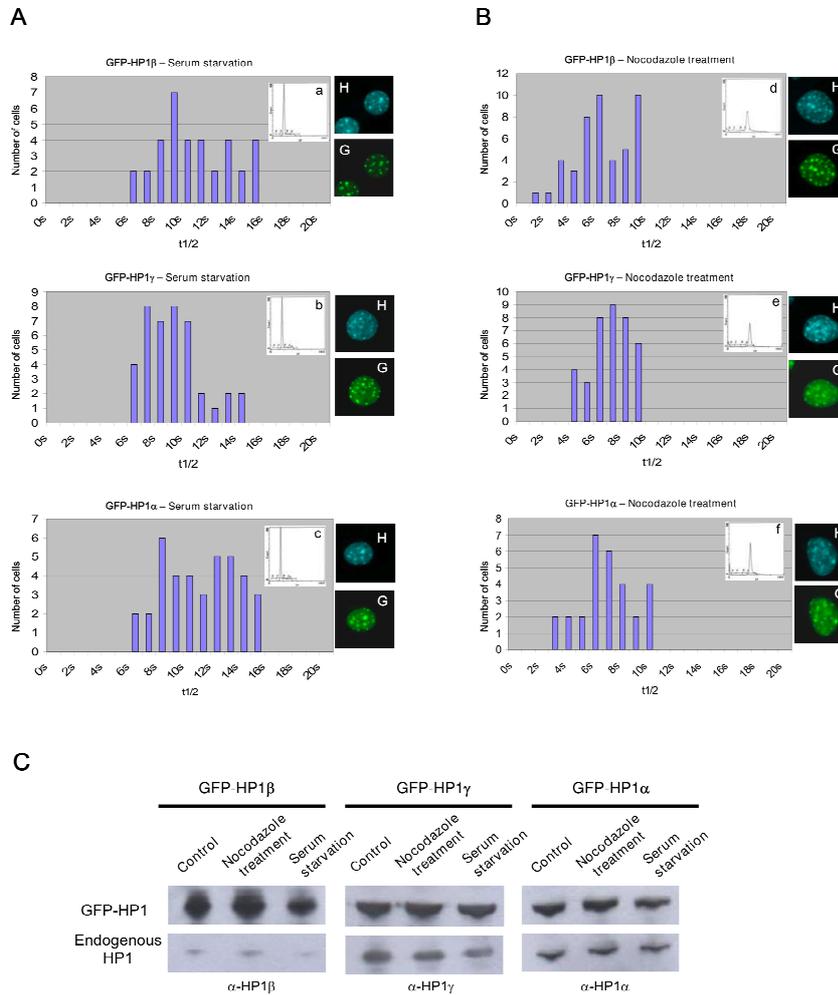


Figure 4
Dynamics of HP1 varies during cell cycle. Distribution of $t_{1/2}$ values deduced from FRAP experiments for GFP-HP1 β (upper panels), GFP-HP1 γ (middle panels), and GFP-HP1 α (lower panels) after serum starvation (A) or nocodazole treatment (B). The bars indicate the number of cells presenting a given $t_{1/2}$ value. The effect of the treatments on cell synchronization is measured by FACS analysis (a-f). Nuclear localization of GFP-tagged HP1 proteins (G) and heterochromatin foci revealed by Hoechst staining (H) are shown. (C) Detection of GFP-HP1 and endogenous HP1 isoforms by Western blotting using specific antibodies for each isoform (α -HP1 β , α -HP1 γ and α -HP1 α) in absence of treatment (Control), in presence of Nocodazole (50 ng/ml) 24 hours or after 72 hours serum starvation.

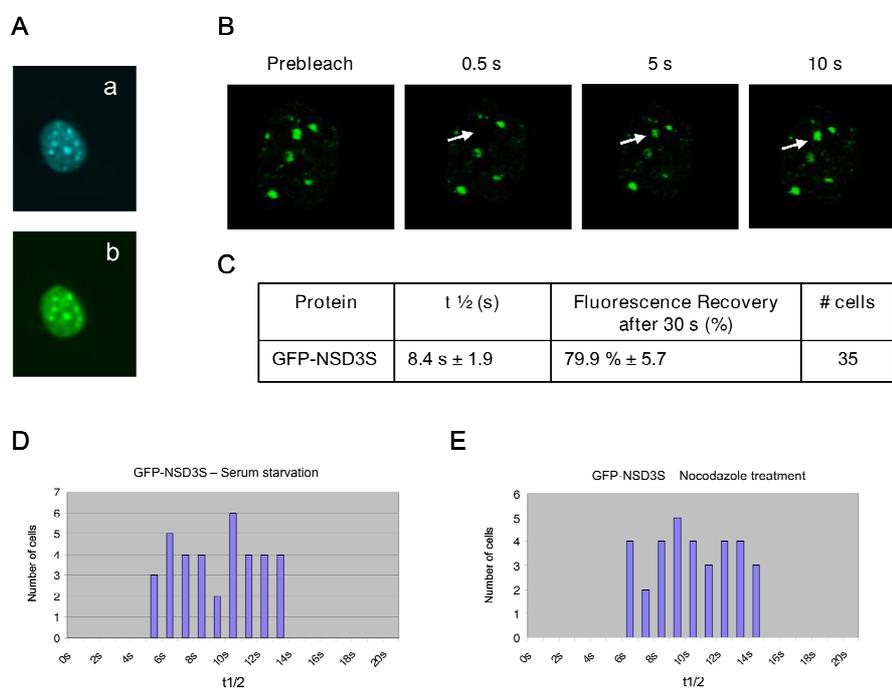


Figure 5
FRAP analysis of GFP-NSD3S shows that NSD3S is a dynamic component of heterochromatin. (A) Cellular localization of NSD3S. GFP-NSD3S expressing L929 cells were fixed with paraformaldehyde and localization of fusion protein was visualized with fluorescence microscopy. Dense foci seen by Hoechst staining represent pericentric heterochromatin regions (a). GFP-NSD3S labeling co-localizes at these sites (b). (B) Time-lapse series of confocal images of living cells expressing GFP-NSD3S during a FRAP experiment. Images were recorded before (Prebleach) and at different time intervals after the bleach. Arrows indicate the photobleached areas. (C) Parameters of NSD3S protein dynamics derived from FRAP curves. The half-time of fluorescence recovery ($t_{1/2}$), the percentage of fluorescence recovery after 30 s, and the number of FRAP curves used are indicated. (D) Representation of the number of cells characterized by a given $t_{1/2}$ value after cell synchronization by serum starvation. (E) Representation of the number of cells characterized by a given $t_{1/2}$ value after cell synchronization by nocodazole treatment.

introduced into pGEX-4T1 (Amersham). Entry clones were recombined into suitable target vectors by GATEWAY LR reactions. A list of clones used, primer and vector sequences is available on request.

Cell culture and stable cell lines

HeLa (Human cervix epitheloid carcinoma - ECACC), L929 (Mouse cell line C3H/An connective tissue - ECACC) and HEK293 (ECACC) cell lines were maintained in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal bovine serum (Invitrogen),

100 U/ml penicillin (Invitrogen), and 100 $\mu\text{g/ml}$ streptomycin (Invitrogen) at 37°C in humidity-saturated 5% CO_2 atmosphere.

Retroviral stable cell lines were generated according to the following procedure. Phoenix amphotropic packaging cells (3×10^5 cells/well) were seeded in a 6-well plate and transfected 24 hours later with 0.8 μg of retroviral plasmid using Exgen 500 (Euromedex) following the instructions of the manufacturer. After 48 hours virus-containing supernatant was filtered through a 0.45- μm -pore-size fil-

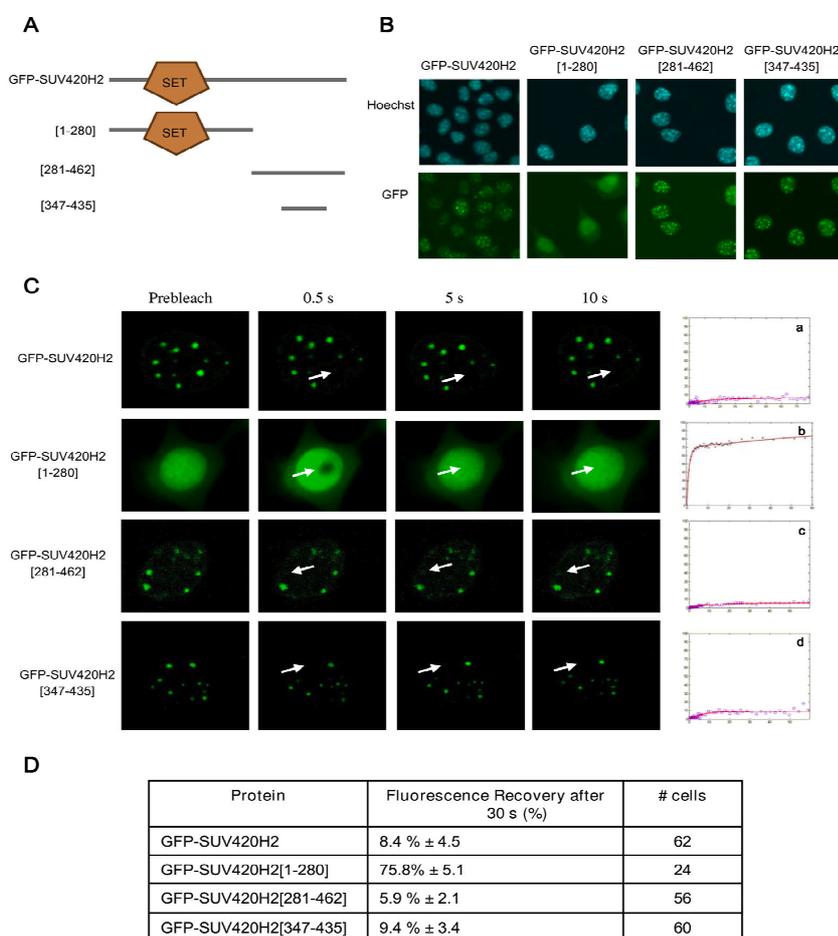


Figure 6
FRAP analysis of GFP-SUV420H2 shows that SUV420H2 is strongly associated to heterochromatin. (A) Schematic representation of various SUV420H2 truncations used. The catalytically active SET domain is indicated. (B) Cellular localization of SUV420H2 and the truncations expressed as GFP fusion proteins in L929 cells. Cells were fixed with paraformaldehyde and localization of fusion proteins was visualized with fluorescence microscopy. Dense foci seen by Hoechst staining represent pericentric heterochromatin regions (upper panels). GFP labelling is shown (lower panels). (C) Time-lapse series of confocal images of living cells expressing GFP-SUV420H2, GFP-SUV420H2 [1-280], GFP-SUV420H2 [281-462], and GFP-SUV420H2 [347-435] as indicated, during FRAP experiments. Images were recorded before (Prebleach) and at different time intervals after the bleach. Arrows indicate the photobleached areas. Relative fluorescence intensities are displayed in recovery curves (a-d). (D) Protein and truncations dynamics derived from FRAP curves. The percentage of fluorescence recovery after 30 s, and the number of FRAP curves used are indicated.

ter. HeLa or L929 cells (10^5 each) were seeded in a six-well plate and transduced with 3 ml filtered virus supernatant in the presence of 8 $\mu\text{g/ml}$ of Polybrene for an infectious round of 24 hours. Cells were then incubated for 24 hours in normal medium. The polyclonal population of cells was then selected with 1 $\mu\text{g/ml}$ of puromycin. Growing cells were then tested for recombinant protein expression using immunocytochemistry for TAP-tagged protein expression or immunofluorescence for GFP-tagged protein expression.

For cell cycle synchronization, cells were seeded at a concentration of 2×10^4 cells/cm² either in 6-well plates or in 10-cm dishes and grown to 50–70% confluence to obtain cultures in the logarithmic growth phase. Synchronization in G0 was achieved by serum deprivation; cultures were washed 3 times with PBS, once with DMEM and then cultured for 72 hours in DMEM without serum, whereas synchronization at G2/M transition was achieved using a nocodazole treatment [30]; cultures were grown in DMEM with 10% fetal bovine serum and 50 ng/ml nocodazole for 24 hours. All synchronization experiments based on serum deprivation or nocodazole treatment were performed in parallel to ensure accurate comparisons.

Antibodies and Western blotting

The following primary antibodies were used for Western blotting: mouse anti-HP1 α antibody (2HP-2G9, Euromedex; used at a dilution of 1:2 000), mouse anti-HP1 β antibody (1MOD-1A9, Euromedex; used at a dilution of 1:2 000), mouse anti-HP1 γ antibody (2MOD-1G6, Euromedex; used at a dilution of 1:2 000), goat anti-CBP antibody (sc-9456, Santa Cruz Biotechnology; used at a dilution of 1:300). Secondary antibodies were: horseradish peroxidase-conjugated goat anti-mouse antibody (115-035-003, Jackson ImmunoResearch; used at a dilution of 1:10 000), horseradish peroxidase-conjugated donkey anti-goat antibody (705-035-003, Jackson ImmunoResearch; used at a dilution of 1:5 000). The protein A moiety of the TAP tag was revealed with rabbit peroxidase anti-peroxidase antibody (P1291, Sigma; used at a dilution of 1:10 000).

For Western blotting, protein samples (50 μg) in SDS loading buffer were electrophoresed on 4–12% Bis-Tris gels (Invitrogen) and transferred to nitrocellulose membranes (Schleicher & Schuell). The membranes were blocked in 10% milk powder in PBS-T (1 \times PBS with 0.1% Tween20) for 1 hour at room temperature, incubated for same time with the primary antibody in PBS-T, and washed three times 10 min in PBS-T. The membranes were then incubated with the peroxidase-conjugated secondary antibody in PBS-T for 1 hour and afterward washed three times 10 min in PBS-T. Signal was detected

using chemiluminescence reagent (ECL, Amersham) on imaging film (GE Healthcare).

Flow cytometry

Cell cycle analysis was performed by flow cytometry as described by Evans et al [31]. Single-cell suspensions were obtained by trypsinization, washed twice with PBS and incubated in 80% ethanol at room temperature for 16–24 hours. Cells were rinsed with PBS, resuspended in 200 $\mu\text{g/ml}$ Propidium Iodide (PI), incubated with 100 $\mu\text{g/ml}$ RNase at 37°C for 30 min, and stored at least one hour at 4°C. Stained cells were analyzed with a Coulter Epics XL flow cytometer equipped with a 488-nm argon laser (Beckman Coulter, Fullerton, CA) and a 530 nm band-pass filter allowing detection of PI fluorescence. A minimum of 10,000 events were collected. GFP or PI fluorescence was expressed as a ratio of the mean channel value of the GFP or PI histogram to the mean channel value of the isotype control histogram.

Tandem affinity purification

For Tandem affinity purification (TAP), cells expressing a TAP-tagged protein were expanded into forty 15-cm dishes. At confluence (about 100 mg total protein), cells were harvested, washed with PBS, resuspended in cell lysis buffer (10 mM Tris-HCl pH 7.4, 1.5 mM MgCl₂, 10 mM KCl, 25 mM NaF, 1 mM Na₃VO₄, 1 mM dithiothreitol [DTT] and complete protease inhibitors [Roche]) and homogenized by 20 strokes with a type B pestle. Nuclei were recovered by centrifugation 10 min at 2000 g, resuspended in nuclear lysis buffer (50 mM Tris-HCl pH 7.4, 1.5 mM MgCl₂, 420 mM NaCl, 20% glycerol, 25 mM NaF, 1 mM Na₃VO₄, 1 mM dithiothreitol [DTT] and complete protease inhibitors [Roche]), and then incubated for 1 hour in a rotation wheel at 4°C to extract nuclear proteins. Lysates were subsequently clarified by ultracentrifugation at 100,000 g, 1 hour.

Nuclear lysate was incubated with IgG agarose beads (Sigma) for 2 hours at 4°C in a rotation wheel. Bound proteins were washed with an excess of lysis buffer and then with a TEV-protease cleavage buffer (10 mM Tris-HCl pH 7.5, 100 mM NaCl and 0.2% NP-40) and eluted by addition of 30 mg TEV protease (Invitrogen) for 2 hours at 16°C. The TEV-protease cleavage product was incubated with calmodulin sepharose (Amersham) in the presence of 2 mM CaCl₂ for 30 min at 4°C in a rotation wheel. After extensive washes in 100 mM Tris-HCl pH8, 100 mM NaCl, 0.5 mM EDTA, 2 mM CaCl₂, calmodulin-bound proteins were eluted by boiling in SDS loading buffer.

Mass spectrometric analysis

Protein eluate was separated on a 4–12% NuPAGE Novex gel (Invitrogen) and stained with Imperial Protein Stain

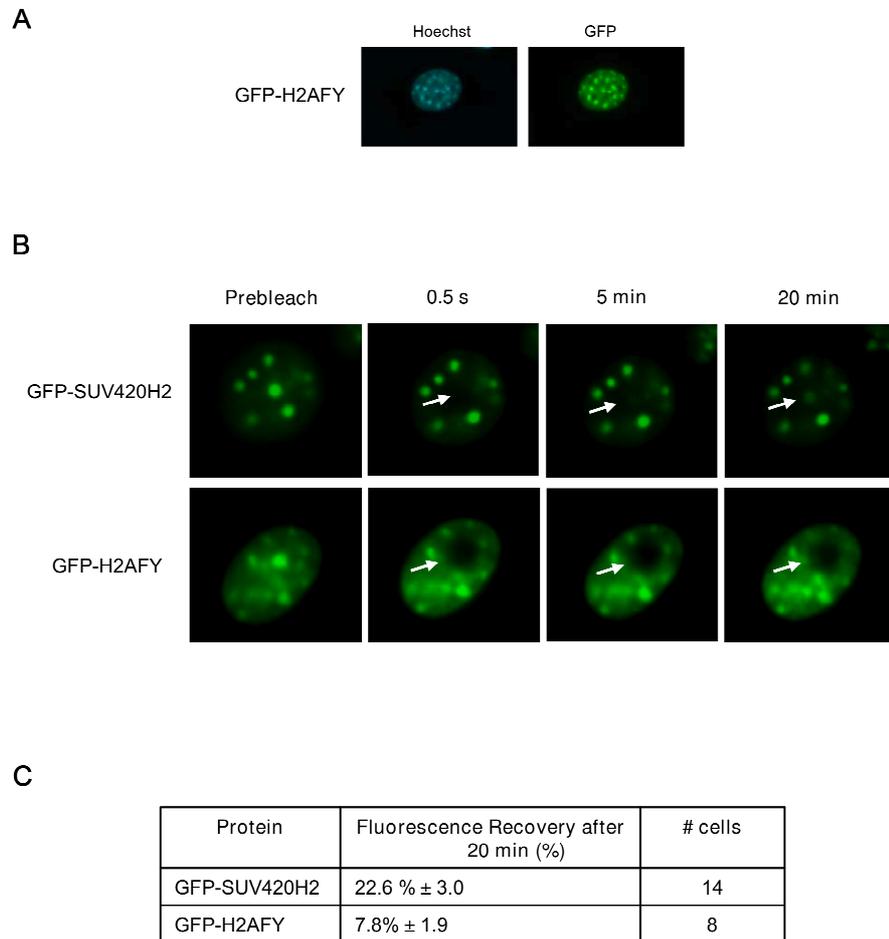


Figure 7
Comparison of SUV420H2 and H2AFY dynamics by FRAP. (A) Cellular localization of GFP-H2AFY fusion proteins stably expressed in L929 cells. Cells were fixed with paraformaldehyde and localization of fusion proteins was visualized with fluorescence microscopy. GFP-H2AFY fusion proteins colocalize with dense foci identified by Hoechst staining as pericentric heterochromatin regions. (B) Time-lapse series of confocal images of living cells expressing GFP-SUV420H2 and GFP-H2AFY as indicated, during FRAP experiments. Images were recorded before (Prebleach) and at different time intervals after the bleach. Arrows indicate the photobleached areas. (C) Protein dynamics derived from FRAP curves. The percentage of fluorescence recovery after 20 min, and the number of FRAP curves used are indicated.

(Pierce). Gel was sliced into 37 bands across the entire separation range of the lane. Cut bands were reduced, alkylated with iodoacetamide, and in-gel digested with trypsin (Promega) as described previously [35]. In brief, gel bands were destained overnight at 4°C in a solution containing 50 mM NH₄HCO₃ and 50% acetonitrile, dehydrated in acetonitrile, and dried in a vacuum centrifuge. Gel pieces were then rehydrated at 4°C for 45 min in a digestion buffer (25 mM NH₄HCO₃ and 12.5 ng/μl trypsin). The supernatant was replaced by 50 μl of 25 mM NH₄HCO₃, and the samples were incubated overnight at 37°C. The tryptic peptides were recovered by 10-min incubations, twice in 45% acetonitrile, 10% HCOOH and once in 95% acetonitrile, 5% HCOOH. All supernatants were pooled and dried in a vacuum centrifuge.

Each tryptic digest sample was subjected to nano-LC-nano-ESI-MS/MS analysis on an ion trap mass spectrometer (LCQ Deca XP⁺, Thermo Electron Corp.), equipped with a nano-electrospray ion source, coupled with a nano-high pressure liquid chromatography system (LC Packings Dionex). Samples were resuspended in 3 μl of 0.1% HCOOH, and 1.4 μl were injected into the mass spectrometer using a Famos autosampler (LC Packings Dionex). The samples were first desalted and then concentrated on a reverse phase precolumn of 5 mm × 0.3 mm inner diameter (Dionex) by solvent A (95% H₂O, 5% acetonitrile, 0.1% HCOOH) delivered by the Switchos pumping device (LC Packings Dionex) at a flow rate of 10 μl/min for 3 min. Peptides were separated on a 15 cm × 75 μm-inner diameter C18 PepMap column (Dionex). The flow rate was set at 200 nl/min. Peptides were eluted using a 5–70% linear gradient of solvent B (20% H₂O, 80% acetonitrile, 0.08% HCOOH) in 45 min. Coated nano-electrospray needles (360 μm outer diameter, 20 μm inner diameter, 10 μm tip inner diameter, standard coating) were obtained from New Objective (Woburn, MA). Spray voltage was set at 1.5 kV, and capillary temperature was set at 170°C. The mass spectrometer was operated in positive ionization mode.

Data acquisition was performed in a data-dependent mode consisting of, alternatively in a single run, a full-scan MS over the range m/z 500–2,000 and a full MS/MS of the ion selected in an exclusion dynamic mode (the most intense ion is selected and excluded for further selection for a duration of 3 min). MS/MS data were acquired using a 2 m/z unit ion isolation window and 35% relative collision energy. MS/MS .raw data files were transformed to .dta files with the Bioworks 3.1 software (Thermo Electron Corp.). The .dta files generated were next concatenated with merge.bat (a DOS batch file for Windows) to be uploaded in Mascot public interface version 2.2.03 <http://www.matrixscience.com> for database searches in

Swiss-Prot 55.1 (359,942 sequences; 129,199,355 residues).

Search parameters in human sequences were: three allowed missed cleavages, methionine oxidation and cysteine carbamidomethylation as variable modifications, 2 Da for peptide tolerance, and 0.8 Da for MS/MS tolerance. Results were scored using the probability-based Mowse score [the protein score is $-10 \times \log(p)$, where p is the probability that the observed match is a random event]. Most proteins were unambiguously identified by the sequencing of several independent peptides. Identifications with Mascot individual ion score < 38 or with the significance threshold p > 5% (indicate identity or extensive homology) were categorically rejected. In addition, because a shared sequence may represent a problem, for single peptide identifications, all sequences obtained by MS/MS analysis were checked using the Basic Local Alignment Search Tool (BLAST) public interface (version 2.2.18) to exclude that sequence sharing with other proteins could interfere with the reliability of the identification.

In vitro GST protein binding assays

GST fusion proteins, were expressed in E. coli BL21 (DE3) and purified on glutathione-Sepharose 4B (GE Healthcare) according to the manufacturer's instructions. GST-proteins were then fixed on glutathione-Sepharose 4B and stored in STE buffer (10 mM Tris-HCl pH8, 150 mM NaCl, 1 mM EDTA and complete protease inhibitors [Roche]). After preclearing with empty beads, nuclear extracts (~500 μg proteins) were incubated with immobilized GST-fusion proteins overnight at 4°C. Beads were washed four times with E1A Buffer (50 mM Hepes, pH 7.9, 1 mM EGTA, 250 mM NaCl, 1 mM DTT, 1 mM EDTA and complete protease inhibitors [Roche]) and bound proteins were recovered with the Elution Buffer (10 mM glutathione in 50 mM Tris-HCl pH 8.0), resolved by 4–12% gradient SDS-PAGE (Invitrogen) and visualized by Western blotting using the peroxidase-anti-peroxidase (PAP) antibody (Sigma) which recognizes the protein A moiety of the TAP tag. Input material corresponds to 2% of the material used in the binding assays.

Nuclear extracts used for GST-pull downs were prepared from HEK293 cells transiently expressing TAP-tagged proteins. Forty eight hours after transfection, cells were lysed in low salt buffer (10 mM Tris-HCl pH 7.4, 25 mM NaCl, 2 mM MgOAc, 1 mM DTT, 1 mM EDTA, 0.05% NP-40 and complete protease inhibitors [Roche]) for 15 min at 4°C. Nuclei were pelleted by centrifugation at 4°C for 10 min at 800 g, and lysed in E1A buffer during 2 hours by gently shaking at 4°C. Nuclear proteins were recovered by centrifugation at 2500 g, 10 min at 4°C.

In vitro translated proteins used for GST-pull downs were produced with the TnT7 Quick Coupled Transcription/Translation System (Promega). Immobilized GST-fusion proteins were incubated with in vitro translated proteins overnight at 4 °C in IP buffer (50 mM Tris-HCl pH7.5, 500 mM NaCl, 1 mM EDTA, 0.5% NP-40, 10% glycerol and complete protease inhibitors [Roche]). The beads were washed four times with IP buffer and resuspended in loading buffer. Bound proteins were resolved by SDS-PAGE and visualized by Western blotting.

Fluorescence recovery after photobleaching

FRAP experiments were performed in living cells and dynamic parameters determined according to McNally [36]. Briefly, GFP-expressing L929 cells were plated on coverslip surfaces, grown in L15 medium (Invitrogen), and maintained at 37 °C in a Leica live-cell chamber. A Leica SP2 confocal microscope equipped with an oil 100× NA 1.4 plan Apo lens objective, a 488 nm mono-ray laser line, and a linear regime detector in intensity measurements were used for FRAP experiments [37]. Five pre-bleach acquisitions were performed to measure background, fluorescence fading, and pre-bleach fluorescence intensity. The laser power was calibrated using an acousto-optical tunable filter (AOTF). Pre-bleach and post-bleach imaging were performed with an AOTF setting at 5% (5 mW), whereas photobleaching was at 100% AOTF (100 mW). The bleached area was chosen circular for simplicity of the equations available to analyze fluorescence recovery for such a shape [38]. The bleached regions had a size of 1 to 1.3 μm of diameter, and were subjected to five excitation pulses of 336 ms each at high laser power. Post-bleach images were collected at 20 time intervals of 336 ms, 15 time intervals of 1,336 s and 15 time intervals of 2,336 s.

Relative fluorescence intensities within the bleached area were plotted as a function of time, yielding the RAW FRAP curve. The recovery curves were corrected for background, fluorescence fading, and decrease in fluorescence during photobleaching. The $t_{1/2}$ value was defined as the time required for reaching half-maximum recovery and was calculated from the corrected recovery curves obtained using an in-house modified MATLAB fitting tool (Mathworks) using fit curves where the intensity recovered in time follows the exponential law,

$$I(t) = 1 - \sum_i A_i \cdot e^{-k_i t},$$

according to McNally and colleagues [36,39].

Abbreviations

TAP: Tandem Affinity Purification; MS: Mass Spectrometry; FRAP: Fluorescence Recovery After Photobleaching;

AA: amino-acid; CSD: Chromoshadow domain; TEV protease: Tobacco etch virus protease

Authors' contributions

PPS performed FRAP experiments and analyses. PV carried out molecular clonings, stable cell line generation, and participated to TAP experiments. DT participated to FRAP curve analyses. JV participated to TAP and performed MS analyses. CR carried out GST-pull down and FRAP experiments. LH participated in the coordination of the FRAP experiments. POA conceived, designed and coordinated the study, and wrote the manuscript. All authors read and approved the final manuscript.

Additional material

Additional file 1

List of proteins identified with SUV420H2. Proteins identified by LC-MS/MS in the SUV420H2 tandem affinity purification are listed.

Click here for file

[http://www.biomedcentral.com/content/supplementary/1471-2121-10-41-S1.pdf]

Additional file 2

List of peptides identified by TAP with SUV420H2. For each protein, the sequences of the peptides identified by LC-MS/MS in the SUV420H2 tandem affinity purification are listed.

Click here for file

[http://www.biomedcentral.com/content/supplementary/1471-2121-10-41-S2.pdf]

Acknowledgements

We are particularly grateful to Christian Rolando, Adeline Page and the Centre Commun de Mesures de Spectrométrie de Masse de l'Université des Sciences et Technologies de Lille for technical expertise in mass spectrometry. We thank Julia Bertout for her help in FACS analysis and Nicolas Fourré for advices in FRAP.

This work was supported by the CNRS, l'Université des Sciences et Technologies de Lille, l'Université de Lille 2, the Région Nord-Pas de Calais, the European Regional Development Funds, and the Agence Nationale de la Recherche (ANR 2007 PFTV 01 I 01). JV was supported by a fellowship from the French Ministère de la Recherche et de l'Enseignement Supérieur.

References

1. Luger K, Mäder AW, Richmond RF, Sargent DF, Richmond TJ: Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* 1997, **389**:251-260.
2. Völkel P, Angrand PO: The control of histone lysine methylation in epigenetic regulation. *Biochimie* 2007, **89**:1-20.
3. Peairs AH, Kubicek S, Mechtler K, O'Sullivan RJ, Denjck AA, Perez-Burgos L, Kohlmaier A, Opravil S, Tachibana M, Shinkai Y, Martens JH, Jenuwein T: Partitioning and plasticity of repressive histone methylation sites in mammalian chromatin. *Mol Cell* 2003, **12**:1577-1589.
4. Rice JC, Briggs SD, Ueberheide B, Barber CM, Shabanowitz J, Hunt DF, Shinkai Y, Allis CD: Histone methyltransferases direct different degrees of methylation to define distinct chromatin domains. *Mol Cell* 2003, **12**:1591-1598.

5. Schotta G, Lachner M, Sarma K, Ebert A, Sengupta R, Reuter G, Reinberg D, Jenuwein T: A silenced pathway to induce H3-K9 and H4-K20 trimethylation at constitutive heterochromatin. *Genes Dev* 2004, **18**:1251-1262.
6. Rea S, Eisenhaber F, O'Carroll D, Strahl BD, Sun ZY, Schmid M, Opravil S, Mechtler K, Poryling CP, Allis CD, Jenuwein T: Regulation of chromatin structure by site-specific histone H3 methyltransferases. *Nature* 2000, **406**:593-599.
7. Garcia-Cao M, O'Sullivan R, Peters AH, Jenuwein T, Blasco MA: Epigenetic regulation of telomere length in mammalian cells by the Suv39h1 and Suv39h2 histone methyltransferases. *Nat Genet* 2004, **36**:94-99.
8. Peters AH, O'Carroll D, Scherthan H, Mechtler K, Sauer S, Schofer C, Weipoltshammer K, Pagani M, Lachner M, Kholmaier A, Opravil S, Doyle M, Sibila M, Jenuwein T: Loss of Suv39h histone methyltransferases impairs mammalian heterochromatin and genome stability. *Cell* 2001, **107**:323-337.
9. Maison C, Bailly D, Peters AH, Quivy JP, Roche D, Taddel A, Lachner M, Jenuwein T, Almouzni G: High-order structure in pericentric heterochromatin involves a distinct pattern of histone modification and an RNA component. *Nat Genet* 2002, **30**:329-334.
10. Bannister AJ, Zegerman P, Partridge JF, Miska EA, Thomas JO, Allshire RC, Kouzarides T: Selective recognition of methylated lysine 9 on histone H3 by the HPI chromo domain. *Nature* 2001, **410**:120-124.
11. Lachner M, O'Carroll D, Rea S, Mechtler K, Jenuwein T: Methylation of histone H3 lysine 9 creates a binding site for HPI proteins. *Nature* 2001, **410**:116-120.
12. Kourmouli N, Sun Y, Sar S van der, Singh PB, Brown JP: Epigenetic regulation of mammalian pericentric heterochromatin in vivo by HPI1. *Biochem Biophys Res Commun* 2005, **337**:901-907.
13. Benetti R, Gonzalo S, Jaco I, Schotta G, Klatt P, Jenuwein T, Blasco MA: Suv4-20h deficiency results in telomere elongation and depression of telomere recombination. *J Cell Biol* 2007, **178**:925-936.
14. Festenstein R, Pagakis SN, Hiragami K, Lyon D, Verrault A, Sekkali B, Kloussis D: Modulation of heterochromatin protein 1 dynamics in primary mammalian cells. *Science* 2003, **299**:719-721.
15. Cheudin T, McNairn AJ, Jenuwein T, Gilbert DM, Singh PB, Mistell T: Maintenance of stable heterochromatin domains by dynamic HPI1 binding. *Science* 2003, **299**:721-725.
16. Schmedeberg L, Welshart K, Diekmann S, Meyer zu Hoersta G, Hemmerich P: High- and low-mobility populations of HPI1 in heterochromatin of mammalian cells. *Mol Biol Cell* 2004, **15**:2819-2833.
17. Krouwels IM, Wiesmuller K, Abraham TE, Molenaar C, Verwoerd NP, Tanke HJ, Dirks RW: A glue for heterochromatin maintenance: stable SUV39H1 binding to heterochromatin is reinforced by the SET domain. *J Cell Biol* 2005, **170**:537-549.
18. Bouwmeester T, Bauch A, Ruffner H, Angrand PO, Bergamini G, Croughton K, Eberhart D, Gagneur J, Ghidella S, Hopf C, Hulse B, Mangano R, Michon AM, Schirle M, Schlegl J, Schwab M, Stein M, Bauer A, Casari G, Drewes G, Gavin AC, Jackson DB, Joberty G, Neubauer G, Rick JM, Kuster B, Superti-Furga G: A physical and functional map of the human TNF α /NF- κ B signal transduction pathway. *Nat Cell Biol* 2004, **6**:97-105.
19. Kourmouli N, Jeppesen P, Mahadevalak S, Burgoyne P, Gilbert DM, Bongiorno S, Prantara G, Fanti L, Pimpinelli S, Shi W, Fundele R, Singh PB: Heterochromatin and tri-methylated lysine 20 of histone 4 in mammals. *J Cell Sci* 2004, **117**:2491-2501.
20. Gingras AC, Aebersold R, Rought B: Advances in protein complex analysis using mass spectrometry. *J Physiol* 2005, **563**:11-21.
21. Schuster H, Maier G, Pfestling H, Nordheim A: DNA intercalators induce specific release of HMG 14, HMG 17 and other DNA-binding proteins from chicken erythrocyte chromatin. *EMBO J* 1985, **4**:3867-3872.
22. Smolchert JF, Henikoff S: The HPI1 chromo shadow domain binds a consensus peptide pentamer. *Curr Biol* 2000, **10**:27-30.
23. Lachner MS, Schultz DC, Negorev D, Maul GG, Rauscher FJ III: The mammalian heterochromatin protein 1 binds diverse nuclear proteins through a common motif that targets the chromoshadow domain. *Biochem Biophys Res Commun* 2005, **331**:929-937.
24. Brasher SV, Smith BO, Fogh RH, Nieldsbach D, Thiru A, Nielsen PR, Broadhurst RW, Ball LJ, Murzina NV, Laue ED: The structure of mouse HPI1 suggests a unique mode of single peptide recognition by the shadow chromo domain dimmer. *EMBO J* 2000, **19**:1587-1597.
25. Cowleson NP, Partridge JF, Allshire RC, McLaughlin PJ: Dimerization of a chromo shadow domain and distinctions from the chromodomain as revealed by structural analysis. *Curr Biol* 2000, **10**:517-525.
26. Lachner MS, Begg GE, Speicher DW, Rauscher FJ: Molecular determinants for targeting heterochromatin protein 1-mediated gene silencing: direct chromoshadow domain-KAP-1 corepressor interaction is essential. *Mol Cell Biol* 2000, **20**:6449-6465.
27. Sprague BL, McNally JG: FRAP analysis of binding: proper and fitting. *Trends Cell Biol* 2005, **15**:84-91.
28. Nielsen AL, Oulad-Abdelghani M, Ortiz JA, Remboutsika E, Chambon P, Losson R: Heterochromatin formation in mammalian cells: interaction between histones and HPI1 proteins. *Mol Cell* 2001, **7**:729-739.
29. Djalmas GK, Terjung S, Brown JP, Aucott RL, Baron-Luhr B, Singh PB, Georgatos SD: Plasticity of HPI1 proteins in mammalian cells. *J Cell Sci* 2007, **120**:3415-3424.
30. Chen Z, Indjejan VB, McManus M, Wang L, Dynlacht BD: CP110, a cell cycle-dependent CDK substrate, regulates centrosome duplication in human cells. *Dev Cell* 2002, **3**:339-350.
31. Evans DL, Bishop GR, Jaso-Friedmann L: Methods for cell cycle analysis and detection of apoptosis of teleost cells. *Methods Cell Sci* 2000, **22**:225-231.
32. Angrand PO, Apicou F, Stewart AF, Dutrillaux B, Losson R, Chambon P: NSD3, a new SET domain-containing gene, maps to 8p12 and is amplified in human breast cancer cell lines. *Genomics* 2001, **74**:79-88.
33. Costanzi C, Pehrson JR: Histone macroH2A1A is concentrated in the inactive X chromosome of female mammals. *Nature* 1998, **393**:599-601.
34. Wade Abbott D, Laszczak M, Lewis JD, Su H, Moore SC, Hills M, Dimitrov S, Auzio J: Structural characterization of macroH2A containing chromatin. *Biochemistry* 2004, **43**:1352-1359.
35. Angrand PO, Segura I, Völkel P, Ghidella S, Terry T, Brajenovic M, Vintersten K, Klein R, Superf-Furga G, Drewes G, Kuster B, Bouwmeester T, Acker-Palmer A: Transgenic mouse proteomics identifies new 14-3-3 associated proteins involved in cytoskeletal rearrangements and cell signalling. *Mol Cell Proteomics* 2006, **5**:2211-2227.
36. McNally JG: Quantitative FRAP in analysis of molecular binding dynamics in vivo. *Methods Cell Biol* 2008, **85**:329-351.
37. Cho EH, Lockett SJ: Calibration and standardization of the emission light path of confocal microscopes. *J Microsc* 2006, **223**:15-25.
38. Carrero G, Crawford E, Hendzel MJ, de Vries G: Characterizing fluorescence recovery curves for nuclear proteins undergoing binding event. *Bull Math Biol* 2004, **66**:1515-1545.
39. Sprague BL, Pego RL, Stavreva DA, McNally JG: Analysis of binding reactions by fluorescence recovery after photobleaching. *Biophys J* 2004, **86**:3473-3495.

Publish with **BioMed Central** and every scientist can read your work free of charge

"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."

Sir Paul Nurse, Cancer Research UK

Your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:
http://www.biomedcentral.com/info/publishing_adv.asp



List of proteins identified with SUV420H2

TAP bait is indicated in bold

Accession	Description	Protein Score	Number of Peptides
SH17R_HUMAN	5-hydroxytryptamine receptor 7	36	3
A28CA_HUMAN	ANKRD28-like family C member 1A - Homo sapiens (Human)	155	4
ACT8_HUMAN	Actin, cytoplasmic 1 - Homo sapiens (Human)	534	17
ACT8L_HUMAN	Protein beta-actin-like - Homo sapiens (Human)	138	7
ACTC_HUMAN	Actin, alpha cardiac muscle 1 - Homo sapiens (Human)	328	16
ADT2_HUMAN	ADP/ATP translocase 2 - Homo sapiens (Human)	46	1
CALM_HUMAN	Calmodulin - Homo sapiens (Human)	101	3
CBX3_HUMAN	Chromobox protein homolog 3 - Homo sapiens (Human)	422	20
CBX6_HUMAN	Chromobox protein homolog 6 - Homo sapiens (Human)	46	3
CH80_HUMAN	80 kDa heat shock protein, mitochondrial precursor - Homo sapiens (Human)	132	3
CPNS1_HUMAN	Calpain small subunit 1 - Homo sapiens (Human)	180	8
DHX9_HUMAN	ATP-dependent RNA helicase A - Homo sapiens (Human)	42	1
DNJA1_HUMAN	DnaJ homolog subfamily A member 1 - Homo sapiens (Human)	132	2
DYH12_HUMAN	Dynein heavy chain 12, axonemal	38	2
EF1A1_HUMAN	Elongation factor 1-alpha 1 - Homo sapiens (Human)	180	4
EFTU_HUMAN	Elongation factor Tu, mitochondrial precursor - Homo sapiens (Human)	278	4
GRF78_HUMAN	78 kDa glucose-regulated protein precursor - Homo sapiens (Human)	398	8
GTF2I_HUMAN	General transcription factor II-I - Homo sapiens (Human)	85	1
H2A1D_HUMAN	Histone H2A type 1-D - Homo sapiens (Human)	128	4
H2B1C_HUMAN	Histone H2B type 1-C/EP/IGH - Homo sapiens (Human)	123	3
H31T_HUMAN	Histone H3.1 - Homo sapiens (Human)	134	8
H4_HUMAN	Histone H4 - Homo sapiens (Human)	162	7
HNR22_HUMAN	Heterogeneous nuclear ribonucleoprotein U-like protein 2 - Homo sapiens (Human)	87	1
HNRP1_HUMAN	Heterogeneous nuclear ribonucleoprotein U - Homo sapiens (Human)	138	2
HSP90A_HUMAN	Heat shock protein HSP 90-alpha - Homo sapiens (Human)	95	1
HSP90B_HUMAN	Heat shock protein HSP 90-beta - Homo sapiens (Human)	88	1
HSP70_HUMAN	Heat shock cognate 71 kDa protein - Homo sapiens (Human)	408	8
IQGAP1_HUMAN	Ras GTPase-activating-like protein IQGAP1 - Homo sapiens (Human)	70	1
K19B7_HUMAN	Protein KIAA1987 - Homo sapiens (Human)	83	2
KBPP_HUMAN	8-phosphotudokinease type C - Homo sapiens (Human)	220	4
KU70_HUMAN	ATP-dependent DNA helicase 2 subunit 1 - Homo sapiens (Human)	51	1
KUB8_HUMAN	ATP-dependent DNA helicase 2 subunit 2 - Homo sapiens (Human)	71	1
LHX1_HUMAN	LIM/homeobox protein Lhx1 - Homo sapiens (Human)	41	2
MATR3_HUMAN	Matrin-3 - Homo sapiens (Human)	120	4
MCM7_HUMAN	DNA replication licensing factor MCM7 - Homo sapiens (Human)	145	3
MFAF4_HUMAN	Microfibril-associated glycoprotein 4 precursor - Homo sapiens (Human)	75	3
MPCP_HUMAN	Phosphate carrier protein, mitochondrial precursor - Homo sapiens (Human)	44	1
MYH14_HUMAN	Myosin-14	68	1
MYH8_HUMAN	Myosin-8 - Homo sapiens (Human)	589	10
MYLK2_HUMAN	Myosin light chain kinase 2, skeletal/cardiac muscle - Homo sapiens (Human)	70	8
ORD1_HUMAN	Oral-facial-digital syndrome 1 protein	40	2
PP1A_HUMAN	Serine/threonine-protein phosphatase PP1-alpha catalytic subunit - Homo sapiens (Human)	48	1
PFKFB1_HUMAN	60S ribosomal protein L11 - Homo sapiens (Human)	232	4
RL11_HUMAN	60S ribosomal protein L11 - Homo sapiens (Human)	82	1
RL12_HUMAN	60S ribosomal protein L12 - Homo sapiens (Human)	85	1
RL23_HUMAN	60S ribosomal protein L23 - Homo sapiens (Human)	84	1
RL23A_HUMAN	60S ribosomal protein L23a - Homo sapiens (Human)	80	1
RL27A_HUMAN	60S ribosomal protein L27a - Homo sapiens (Human)	78	2
RL30_HUMAN	60S ribosomal protein L30 - Homo sapiens (Human)	56	1
RLA0_HUMAN	60S acidic ribosomal protein P0 - Homo sapiens (Human)	104	3
RLA1_HUMAN	60S acidic ribosomal protein P1 - Homo sapiens (Human)	56	1
RS11_HUMAN	40S ribosomal protein S11 - Homo sapiens (Human)	83	2
RS13_HUMAN	40S ribosomal protein S13 - Homo sapiens (Human)	75	2
RS15A_HUMAN	40S ribosomal protein S15a - Homo sapiens (Human)	48	2
RS17_HUMAN	40S ribosomal protein S17 - Homo sapiens (Human)	85	1
RS18_HUMAN	40S ribosomal protein S18 - Homo sapiens (Human)	78	2
RS25_HUMAN	40S ribosomal protein S25 - Homo sapiens (Human)	60	2
RS3_HUMAN	40S ribosomal protein S3 - Homo sapiens (Human)	333	7
RS3A_HUMAN	40S ribosomal protein S3a - Homo sapiens (Human)	49	1
RS4Y1_HUMAN	40S ribosomal protein S4, Y isoform 1 - Homo sapiens (Human)	68	1
RS5_HUMAN	40S ribosomal protein S5 - Homo sapiens (Human)	82	1
S10A2_HUMAN	Protein S100-A2 - Homo sapiens (Human)	88	3
SFPQ_HUMAN	Splicing factor, proline- and glutamine-rich - Homo sapiens (Human)	70	1
SPT16I_HUMAN	FACT complex subunit SPT16 - Homo sapiens (Human)	65	1
SPTG7_HUMAN	Paraplegin - Homo sapiens (Human)	41	2
SPTA2_HUMAN	Spectrin alpha chain, brain - Homo sapiens (Human)	96	2
SBB_HUMAN	Single-stranded DNA-binding protein, mitochondrial precursor - Homo sapiens (Human)	66	1
SLN42_HUMAN	Histone-lysine N-methyltransferase SUV420H2 - Homo sapiens (Human)	587	28
TBA1A_HUMAN	Tubulin alpha-1A chain - Homo sapiens (Human)	81	2
TBA1B_HUMAN	Tubulin alpha-1B chain - Homo sapiens (Human)	879	16
TBA1C_HUMAN	Tubulin alpha-1C chain - Homo sapiens (Human)	512	23
TBA4A_HUMAN	Tubulin alpha-4A chain - Homo sapiens (Human)	686	16
TBA8_HUMAN	Tubulin alpha-8 chain - Homo sapiens (Human)	306	8
TBB1_HUMAN	Tubulin beta-1 chain - Homo sapiens (Human)	211	11
TBB2A_HUMAN	Tubulin beta-2A chain - Homo sapiens (Human)	896	26
TBB2C_HUMAN	Tubulin beta-2C chain - Homo sapiens (Human)	1114	31
TBB3_HUMAN	Tubulin beta-3 chain - Homo sapiens (Human)	560	23
TBB4_HUMAN	Tubulin beta-4 chain - Homo sapiens (Human)	855	23
TBB6_HUMAN	Tubulin beta chain - Homo sapiens (Human)	1082	43
TBB6_HUMAN	Tubulin beta-6 chain - Homo sapiens (Human)	281	11
TNKS1_HUMAN	Tankyrase-1	40	2
WBP4_HUMAN	WW domain-binding protein 4 - Homo sapiens (Human)	41	3

List of peptides identified by TAP with SUV420H2

Data represents a total of 462 identified peptides of which 376 are non redundant.
Average mass accuracy is 52ppm; Standard deviation is 52ppm.
Average peptide size is 1686 Da; Standard deviation is 465 Da; Peptide range is 700-3500 Da.

Protein accession	Description	Observed m/z	Peptide Mass (exp. in Da)	Peptide Mass (theo. in Da)	Delta Mass	Peptide Score	Peptide Misses	Peptide Sequence	Modifications	Peptide count
BTTR_HUMAN	5-hydroxytryptamine receptor 7	656.4361	1270.6703	1270.6703	0.1914	14.51	1	KLSMAGMEALK	Oxidation (M)	1
BTTR_HUMAN	5-hydroxytryptamine receptor 7	789.4719	2250.2253	2251.2278	-1.0026	7.04	3	LQKVEKCANLSPILKHER	Carbamidomethyl (C)	1
BTTR_HUMAN	5-hydroxytryptamine receptor 7	717.9717	1494.9133	1494.9133	0.0000	16.49	3	EDKATKPKMKNRPPKRS	Oxidation (M)	1
A2SC_HUMAN	ANKRD28-like family C member 1A - Homo sapiens (Human)	1271.2587	2542.1136	2542.1136	-0.0000	75.25	0	LCVWLDLDFCEEMATAASSSSLEK	Carbamidomethyl (C)	1
A2SC_HUMAN	ANKRD28-like family C member 1A - Homo sapiens (Human)	1284.2663	2566.5132	2566.5132	0.0000	75.25	0	LCVWLDLDFCEEMATAASSSSLEK	Carbamidomethyl (C)	1
A2SC_HUMAN	ANKRD28-like family C member 1A - Homo sapiens (Human)	669.0048	1789.5086	1789.5086	0.0000	49.35	0	SVFLPQGVITKNER	Carbamidomethyl (C); Oxidation (M)	1
ACTB_HUMAN	Actin, cytoplasmic 1 - Homo sapiens (Human)	1192.9381	1799.9381	1797.9382	0.0000	81.92	0	AVFSSVGRPR		1
ACTB_HUMAN	Actin, cytoplasmic 1 - Homo sapiens (Human)	909.4705	907.4709	907.4709	-0.0168	37.85	0	DLTDYLMK	Oxidation (M)	1
ACTB_HUMAN	Actin, cytoplasmic 1 - Homo sapiens (Human)	1014.4732	1013.4729	1013.4729	-0.0001	43.49	0	DLTDYLMK	Oxidation (M)	2
ACTB_HUMAN	Actin, cytoplasmic 1 - Homo sapiens (Human)	589.6892	1177.3219	1178.6008	0.7159	52.37	0	ETALAPSTMK	Oxidation (M)	1
ACTB_HUMAN	Actin, cytoplasmic 1 - Homo sapiens (Human)	589.6705	1131.7265	1131.5197	0.2068	45.24	0	GYSTFTFAER	Oxidation (M)	2
ACTB_HUMAN	Actin, cytoplasmic 1 - Homo sapiens (Human)	708.2246	1616.4397	1614.7419	1.6949	61.09	0	WHITTYNELR	Oxidation (M)	2
ACTB_HUMAN	Actin, cytoplasmic 1 - Homo sapiens (Human)	1180.6997	2269.1199	2269.1020	0.0042	63.92	1	KDLYANTVALSQKTTWTPQADLR	Oxidation (M)	1
ACTB_HUMAN	Actin, cytoplasmic 1 - Homo sapiens (Human)	1284.1577	2569.3006	2566.1914	1.1363	105.25	0	LCVWLDLDFCEEMATAASSSSLEK	Carbamidomethyl (C); Oxidation (M)	2
ACTB_HUMAN	Actin, cytoplasmic 1 - Homo sapiens (Human)	1276.9777	2549.9399	2546.1993	0.7704	76.99	0	LCVWLDLDFCEEMATAASSSSLEK	Carbamidomethyl (C)	1
ACTB_HUMAN	Actin, cytoplasmic 1 - Homo sapiens (Human)	605.5878	1616.6816	1616.6816	0.0000	48.12	0	QEVDESQFPMHR	Carbamidomethyl (C)	1
ACTB_HUMAN	Actin, cytoplasmic 1 - Homo sapiens (Human)	589.13	1789.2455	1789.2455	0.0000	86.19	0	SVFLPQGVITKNER		2
ACTB_HUMAN	Actin, cytoplasmic 1 - Homo sapiens (Human)	977.7028	1853.3812	1853.0671	0.3341	87.79	0	VAPFEPHLLI TEAPLNPK		1
ACTBL_HUMAN	Protein beta-actin-like - Homo sapiens (Human)	866.4706	967.4779	967.4779	-0.0168	37.85	0	DLTDYLMK	Oxidation (M)	1
ACTBL_HUMAN	Protein beta-actin-like - Homo sapiens (Human)	1014.4752	1013.4879	1013.4879	-0.0001	43.49	0	DLTDYLMK	Oxidation (M)	2
ACTBL_HUMAN	Protein beta-actin-like - Homo sapiens (Human)	980.9648	1199.9751	1199.9751	0.0000	8.46	0	GYNFTTFAER		2
ACTBL_HUMAN	Protein beta-actin-like - Homo sapiens (Human)	686.13	1789.2455	1789.2455	0.0000	86.19	0	SVFLPQGVITKNER		2
ACTBL_HUMAN	Protein beta-actin-like - Homo sapiens (Human)	877.7028	1853.3812	1853.0571	0.3241	81.02	0	VAPFEPHLLI TEAPLNPK		1
ACTC_HUMAN	Actin, alpha cardiac muscle 1 - Homo sapiens (Human)	900.1638	1199.2931	1197.9983	0.6948	61.02	0	AVFSSVGRPR		1
ACTC_HUMAN	Actin, alpha cardiac muscle 1 - Homo sapiens (Human)	866.4705	967.4532	967.4779	-0.0168	37.85	0	DLTDYLMK	Oxidation (M)	2
ACTC_HUMAN	Actin, alpha cardiac muscle 1 - Homo sapiens (Human)	869.9692	1013.4879	1013.4879	-0.0001	43.49	0	DLTDYLMK	Oxidation (M)	1
ACTC_HUMAN	Actin, alpha cardiac muscle 1 - Homo sapiens (Human)	1405.7277	2808.4499	2806.2633	0.7164	4.19	1	ELCYWLDLDFENEMATAASSSSLEK	Carbamidomethyl (C); Oxidation (M)	1
ACTC_HUMAN	Actin, alpha cardiac muscle 1 - Homo sapiens (Human)	759.2256	1519.6597	1514.7418	1.8946	51.09	0	WHITTYNELR	Oxidation (M)	2
ACTC_HUMAN	Actin, alpha cardiac muscle 1 - Homo sapiens (Human)	760.1767	2046.9396	2051.1458	-1.2069	16.73	0	LCVWLDLDFENEMATAASSSSLEK		1
ACTC_HUMAN	Actin, alpha cardiac muscle 1 - Homo sapiens (Human)	677.7028	1496.9196	1496.9196	-0.0000	86.19	0	SVFLPQGVITKNER		2
ACTC_HUMAN	Actin, alpha cardiac muscle 1 - Homo sapiens (Human)	669.9878	1493.5912	1493.5912	0.0000	86.19	0	SVFLPQGVITKNER		2
AD2T_HUMAN	ADP3ATP translocase 2 - Homo sapiens (Human)	709.2955	1470.4416	1470.4416	0.0000	20.81	0	YREHHTVNDQMEK	Oxidation (M)	2
CALM_HUMAN	Calmodulin - Homo sapiens (Human)	681.6997	1121.1249	1120.9866	0.6589	44.76	1	DTDDEEERAFR		1
CALM_HUMAN	Calmodulin - Homo sapiens (Human)	686.216	1558.5044	1558.5044	0.0000	33.36	1	DTDDEEERAFR		1
CBX3_HUMAN	Chromobox protein homolog 3 - Homo sapiens (Human)	1881.8923	2622.6692	2621.9825	1.6867	42.44	0	EADKDDKQWYEEFVQVMNTAK	2 Oxidation (M)	1
CBX3_HUMAN	Chromobox protein homolog 3 - Homo sapiens (Human)	867.4803	1712.9481	1709.8536	1.7695	24.87	1	VFDKDKQWYBAEELR	Carbamidomethyl (C)	1
CBX3_HUMAN	Chromobox protein homolog 3 - Homo sapiens (Human)	867.4811	1729.9481	1711.9623	1.0396	48.12	0	IGKATDSSKQELMFLMK	Oxidation (M)	1
CBX3_HUMAN	Chromobox protein homolog 3 - Homo sapiens (Human)	831.8347	1681.8546	1727.8474	0.4362	99.76	0	IKGATDSSKQELMFLMK	Oxidation (M)	7
CBX3_HUMAN	Chromobox protein homolog 3 - Homo sapiens (Human)	737.0092	1471.6038	1471.6038	0.0000	100.82	1	KVEAEPEEFVYK	Oxidation (M)	1
CBX3_HUMAN	Chromobox protein homolog 3 - Homo sapiens (Human)	826.5398	1676.3974	1676.3974	0.0000	31.56	0	0.4		2
CBX3_HUMAN	Chromobox protein homolog 3 - Homo sapiens (Human)	787.8008	1532.1871	1532.1871	0.0000	82.85	0	VKEAEPEEFVYK	Carbamidomethyl (C)	4
CBX3_HUMAN	Chromobox protein homolog 3 - Homo sapiens (Human)	706.0042	1405.7277	1405.7277	0.0000	20.96	0	VEYFLK		1

KEP8_HUMAN	Keap1-like domain type 8 - Homo sapiens (Human)	1413_2235	1441_5943	6292	6182	0	FLRSNHEHLETK	Carbamidomethyl (C), Oxidation (M)	1
KEP9_HUMAN	Keap1-like domain type 9 - Homo sapiens (Human)	3310_2046	3330_5850	6561	6441	0	FLMSKAVNLDVQVAMDERPDDAVR	Carbamidomethyl (C), Oxidation (M)	1
KEP10_HUMAN	Keap1-like domain type 10 - Homo sapiens (Human)	1339_4884	1359_5950	6850	7050	0	IKSSEKAVNLDVQVAMDERPDDAVR	Carbamidomethyl (C), Oxidation (M)	1
KLTO_HUMAN	ATP-dependent DNA helicase 2 subunit 1 - Homo sapiens (Homo sapiens)	1703_3986	1703_3986	6862	6862	0	KKSSKSNKLVQGFHR	Oxidation (M)	1
KLUS8_HUMAN	ATP-dependent DNA helicase 2 subunit 2 - Homo sapiens (Homo sapiens)	1022_3383	1022_3383	6242	7106	0	IKKSNKLVQGFHR	Oxidation (M)	1
LHK1_HUMAN	LIM-kinase domain protein 1 - Homo sapiens (Human)	1246_3632	1246_3632	6869	3242	0	2PKMLKALGAR	Oxidation (M)	1
LHK2_HUMAN	LIM-kinase domain protein 2 - Homo sapiens (Human)	551_1989	551_1989	6869	3242	0	2PKMLKALGAR	Oxidation (M)	1
MATR3_HUMAN	Matriin-3 - Homo sapiens (Human)	1703_0702	1703_0702	1386	1325	0	0VFLHNGFTSMCHK	Oxidation (M)	1
MATR33_HUMAN	Matriin-33 - Homo sapiens (Human)	1324_1938	1324_1938	1174	7416	0	0GADDAANILASFLGAR	Oxidation (M)	1
MATR333_HUMAN	Matriin-333 - Homo sapiens (Human)	1324_1938	1324_1938	1174	7416	0	0GADDAANILASFLGAR	Oxidation (M)	1
MCM7_HUMAN	DNA replication licensing factor MCM7 - Homo sapiens (Human)	2037_2785	2037_2785	6862	1043	0	0ENLGGANGNLGSPR	Oxidation (M)	2
MCM7L_HUMAN	DNA replication licensing factor MCM7L - Homo sapiens (Human)	1026_5844	1026_5844	6862	3810	0	0VHLNLIHSSYSQSAV.LK	Oxidation (M)	2
MCPK4_HUMAN	Microtubule-associated glycoprotein 4 precursor - Homo sapiens (Homo sapiens)	1473_4308	1473_4308	6862	57363	0	0CAEELTLNAR	Oxidation (M)	1
MCPK4L_HUMAN	Microtubule-associated glycoprotein 4 precursor - Homo sapiens (Homo sapiens)	1051722	1051722	6862	4684	0	0ADGEEIDFYEK	Oxidation (M)	1
MCPK9_HUMAN	Microtubule-associated glycoprotein 9 precursor - Homo sapiens (Homo sapiens)	878_3061	878_3061	6862	2854	0	0TORPADVFAIVR	Oxidation (M)	1
MCPK9L_HUMAN	Microtubule-associated glycoprotein 9 precursor - Homo sapiens (Homo sapiens)	517_3959	517_3959	6862	2854	0	0ADGEEIDFYEK	Oxidation (M)	1
MPCP_HUMAN	Phosphate carrier protein, mitochondrial precursor - Homo sapiens (Homo sapiens)	1381_141	1472_81	6208	1784	0	1GFYYS.LKR	Oxidation (M)	1
MYH6_HUMAN	Myosin-6 - Homo sapiens (Human)	1304_1663	1304_1663	6862	4386	0	0KIQPGYVAVTLR	Oxidation (M)	1
MYH6L_HUMAN	Myosin-6L - Homo sapiens (Human)	1192_8508	1192_8508	6862	8787	0	0EDADFALEALAK	Oxidation (M)	1
MYH8_HUMAN	Myosin-8 - Homo sapiens (Human)	1587_0432	1587_0432	6862	5207	0	2DQRELDTRASR	Oxidation (M)	1
MYH8L_HUMAN	Myosin-8L - Homo sapiens (Human)	1304_1663	1304_1663	6862	8787	0	0EDADFALEALAK	Oxidation (M)	1
MYH9_HUMAN	Myosin-9 - Homo sapiens (Human)	1633_3378	1633_3378	6862	10085	0	0KEFTNLTTEEEK	Oxidation (M)	1
MYH9L_HUMAN	Myosin-9L - Homo sapiens (Human)	1633_3378	1633_3378	6862	10085	0	0KEFTNLTTEEEK	Oxidation (M)	1
MYH10_HUMAN	Myosin-10 - Homo sapiens (Human)	1530_462	1530_462	6862	7403	0	0MLEDQDNETK	Oxidation (M)	2
MYH10L_HUMAN	Myosin-10L - Homo sapiens (Human)	1946_4415	1946_4415	6862	8643	0	0MLEDQDNETK	Oxidation (M)	2
MYH11_HUMAN	Myosin-11 - Homo sapiens (Human)	1673_777	1673_777	6862	8506	0	0LQVELNYTLISQSSK	Oxidation (M)	1
MYH12_HUMAN	Myosin-12 - Homo sapiens (Human)	603_3558	603_3558	6862	7377	0	0FNINPLAQDAVAAK	Oxidation (M)	1
MYH12L_HUMAN	Myosin-12L - Homo sapiens (Human)	1141_1022	1141_1022	6862	6544	0	0TDLLEPRNK	Oxidation (M)	1
MYH13_HUMAN	Myosin-13 - Homo sapiens (Human)	9420_2869	9420_2869	6862	6716	0	2WATERGAVELGRINFTDKAPKSTTQERPLM	Oxidation (M)	1
MYH14_HUMAN	Myosin-14 - Homo sapiens (Human)	131_1263	131_1263	6862	8515	0	0MLESYLRK	Oxidation (M)	1
MYH15_HUMAN	Myosin-15 - Homo sapiens (Human)	1231_3245	1231_3245	6862	8515	0	0MLESYLRK	Oxidation (M)	1
MYH16_HUMAN	Myosin-16 - Homo sapiens (Human)	1130_8336	1130_8336	6862	1734	0	0MLESYLRK	Oxidation (M)	1
MYH17_HUMAN	Myosin-17 - Homo sapiens (Human)	1639_8199	1639_8199	6862	5814	0	0MLESYLRK	Oxidation (M)	1
MYH18_HUMAN	Myosin-18 - Homo sapiens (Human)	820_917	820_917	6862	5814	0	0MLESYLRK	Oxidation (M)	1
MYH19_HUMAN	Myosin-19 - Homo sapiens (Human)	612_9771	612_9771	6862	8017	0	0MLESYLRK	Oxidation (M)	1
MYH20_HUMAN	Myosin-20 - Homo sapiens (Human)	1001_8983	1001_8983	6862	3794	0	0MLESYLRK	Oxidation (M)	1
MYH21_HUMAN	Myosin-21 - Homo sapiens (Human)	1205_0387	1205_0387	6862	1636	0	0MLESYLRK	Oxidation (M)	1
MYH22_HUMAN	Myosin-22 - Homo sapiens (Human)	1536_2291	1536_2291	6862	10023	0	0MLESYLRK	Oxidation (M)	1
MYH23_HUMAN	Myosin-23 - Homo sapiens (Human)	1646_4268	1646_4268	6862	8161	0	0MLESYLRK	Oxidation (M)	1
MYH24_HUMAN	Myosin-24 - Homo sapiens (Human)	1674_7702	1674_7702	6862	8639	0	0MLESYLRK	Oxidation (M)	1
MYH25_HUMAN	Myosin-25 - Homo sapiens (Human)	1490_8653	1490_8653	6862	8426	0	0MLESYLRK	Oxidation (M)	1
MYH26_HUMAN	Myosin-26 - Homo sapiens (Human)	1404_1122	1404_1122	6862	873	0	0MLESYLRK	Oxidation (M)	1
MYH27_HUMAN	Myosin-27 - Homo sapiens (Human)	931_3945	931_3945	6862	873	0	0MLESYLRK	Oxidation (M)	1
MYH28_HUMAN	Myosin-28 - Homo sapiens (Human)	1110_8307	1110_8307	6862	8975	0	0MLESYLRK	Oxidation (M)	1
MYH29_HUMAN	Myosin-29 - Homo sapiens (Human)	1756_3029	1756_3029	6862	6616	0	0MLESYLRK	Oxidation (M)	1
MYH30_HUMAN	Myosin-30 - Homo sapiens (Human)	1285_9288	1285_9288	6862	8612	0	0MLESYLRK	Oxidation (M)	1
MYH31_HUMAN	Myosin-31 - Homo sapiens (Human)	1221_5287	1221_5287	6862	3225	0	0MLESYLRK	Oxidation (M)	1
MYH32_HUMAN	Myosin-32 - Homo sapiens (Human)	1676_1784	1676_1784	6862	513	0	0MLESYLRK	Oxidation (M)	1
MYH33_HUMAN	Myosin-33 - Homo sapiens (Human)	1703_8425	1703_8425	6862	8632	0	0MLESYLRK	Oxidation (M)	1
MYH34_HUMAN	Myosin-34 - Homo sapiens (Human)	1149_0868	1149_0868	6862	5036	0	0MLESYLRK	Oxidation (M)	1
MYH35_HUMAN	Myosin-35 - Homo sapiens (Human)	1965_8104	1965_8104	6862	3217	0	0MLESYLRK	Oxidation (M)	1
MYH36_HUMAN	Myosin-36 - Homo sapiens (Human)	1246_8792	1246_8792	6862	1848	0	0MLESYLRK	Oxidation (M)	1
MYH37_HUMAN	Myosin-37 - Homo sapiens (Human)	1437_3751	1437_3751	6862	8585	0	0MLESYLRK	Oxidation (M)	1
MYH38_HUMAN	Myosin-38 - Homo sapiens (Human)	742_3056	742_3056	6862	1285	0	0MLESYLRK	Oxidation (M)	1
MYH39_HUMAN	Myosin-39 - Homo sapiens (Human)	1126_7951	1126_7951	6862	2088	0	0MLESYLRK	Oxidation (M)	1
MYH40_HUMAN	Myosin-40 - Homo sapiens (Human)	1287_8650	1287_8650	6862	8514	0	0MLESYLRK	Oxidation (M)	1
MYH41_HUMAN	Myosin-41 - Homo sapiens (Human)	1236_2273	1236_2273	6862	447	0	0MLESYLRK	Oxidation (M)	1
MYH42_HUMAN	Myosin-42 - Homo sapiens (Human)	1247_3218	1247_3218	6862	447	0	0MLESYLRK	Oxidation (M)	1
MYH43_HUMAN	Myosin-43 - Homo sapiens (Human)	1060_1619	1060_1619	6862	3706	0	0MLESYLRK	Oxidation (M)	1

R53_HUMAN	46S ribosomal protein S3 - Homo sapiens (Human)	1030,0007	1091,5811	0,4398	44,72	0	DELMEFLTR	1
R53_HUMAN	46S ribosomal protein S3 - Homo sapiens (Human)	1470,1004	1480,1261	0,2361	65,46	0	DELFTTSSGDK	1
R53_HUMAN	46S ribosomal protein S3 - Homo sapiens (Human)	1423,2078	1422,7813	0,2451	62,76	0	ELAEKQVSEK	1
R53_HUMAN	46S ribosomal protein S3 - Homo sapiens (Human)	1573,6083	1572,7937	0,6113	46,46	0	GGKGRPAKAPDPTA	1
R53_HUMAN	46S ribosomal protein S3 - Homo sapiens (Human)	1593,6892	1594,1619	0,6028	41,3	0	GGKGRPAKAPDPTA	1
R53_HUMAN	46S ribosomal protein S3 - Homo sapiens (Human)	1268,4628	1267,6816	0,6201	67,62	0	ELCALQKESLR	1
R53_HUMAN	46S ribosomal protein S3a - Homo sapiens (Human)	1026,7511	1026,823	0,1281	88,99	1	LILPLAIVR	1
R53_HUMAN	46S ribosomal protein S3a - Homo sapiens (Human)	1201,9511	1206,7285	0,2016	40,36	1	TEIPLSISGIDSEK	1
R53_HUMAN	46S ribosomal protein S5 - Homo sapiens (Human)	601,4828	1214,852	0,2981	59,02	0	GPIHLVTHIDAR	1
R53_HUMAN	Protein S100-A2 - Homo sapiens (Human)	1004,5179	1030,2238	0,7767	62,05	0	TAECLADELINAAK	1
R53_HUMAN	Protein S100-A2 - Homo sapiens (Human)	1902,9938	0,2239	1004,752	38,39	2	ELRPSVGEK	1
R53_HUMAN	Splicing factor, proline- and glutamine-rich - Homo sapiens (Human)	627,0319	0,3976	1903,392	47,96	2	ELRPSVGEKVEEGLKK	2
R53_HUMAN	FACT complex subunit SPT16 - Homo sapiens (Human)	1232,0482	1231,5135	0,4356	74,47	0	YGEPSVGVFNK	1
R53_HUMAN	Paraphagin - Homo sapiens (Human)	1472,4294	1424,9838	0,0751	67,62	0	YTESVQVSLNMTK	1
R53_HUMAN	Paraphagin - Homo sapiens (Human)	731,2205	1470,7678	1,0660	40,19	1	LKRAEDELNEAK	1
R53_HUMAN	Speckin alpha chain, brain - Homo sapiens (Human)	1104,8126	3312,8013	-1,1848	3,74	3	MCMAALGRASALSNFNEVTSAGQDURKIVTR	1
R53_HUMAN	Speckin alpha chain, brain - Homo sapiens (Human)	897,6912	1393,3678	0,6688	88,19	0	LGESQTLQDFSR	1
R53_HUMAN	Single-stranded DNA-binding protein, mitochondrial precursor - Homo sapiens (Human)	804,9005	1072,6913	0,2678	34,14	0	SLOQLAEEK	1
R53_HUMAN	Histone-lysine N-methyltransferase SUV420H2 - Homo sapiens (Human)	844,1819	1610,7424	1,0441	67,63	0	SGDGEVYQLGDVSKK	1
R53_HUMAN	Histone-lysine N-methyltransferase SUV420H2 - Homo sapiens (Human)	748,0821	1907,8864	0,4828	67,27	0	AFPLPESGFTLPQTR	1
R53_HUMAN	Histone-lysine N-methyltransferase SUV420H2 - Homo sapiens (Human)	754,2528	1488,8867	0,4368	80,9	0	ADINDFRIMVYTR	1
R53_HUMAN	Histone-lysine N-methyltransferase SUV420H2 - Homo sapiens (Human)	831,4985	1505,9458	0,4954	86,37	0	ADINDFRIMVYTR	1
R53_HUMAN	Histone-lysine N-methyltransferase SUV420H2 - Homo sapiens (Human)	683,0967	1044,8126	0,2397	48,44	0	ALTLGGWIKAR	1
R53_HUMAN	Histone-lysine N-methyltransferase SUV420H2 - Homo sapiens (Human)	942,4431	1333,0907	0,414	65,68	0	DPTCAACQPLR	1
R53_HUMAN	Histone-lysine N-methyltransferase SUV420H2 - Homo sapiens (Human)	617,1685	961,4605	-0,5147	40,36	0	EAEDELRLR	1
R53_HUMAN	Histone-lysine N-methyltransferase SUV420H2 - Homo sapiens (Human)	825,0719	1247,8824	0,5989	32,33	0	EPALPPLPLK	1
R53_HUMAN	Histone-lysine N-methyltransferase SUV420H2 - Homo sapiens (Human)	647,9311	1247,8824	0,4158	26,81	0	PYPALDSIAGACQK	1
R53_HUMAN	Histone-lysine N-methyltransferase SUV420H2 - Homo sapiens (Human)	849,4919	1346,7746	0,4728	61,97	0	LELLVGSALER	1
R53_HUMAN	Histone-lysine N-methyltransferase SUV420H2 - Homo sapiens (Human)	633,35	1893,982	-0,9868	69,24	1	LKAGVSLDQVPHAS	1
R53_HUMAN	Histone-lysine N-methyltransferase SUV420H2 - Homo sapiens (Human)	633,35	1793,1871	-0,7816	69,24	1	LKAGVSLDQVDEEL	1
R53_HUMAN	Histone-lysine N-methyltransferase SUV420H2 - Homo sapiens (Human)	591,9517	1244,7916	0,6721	23,42	1	EVMLPQVPLR	1
R53_HUMAN	Histone-lysine N-methyltransferase SUV420H2 - Homo sapiens (Human)	610,9572	1124,8889	0,2543	23,42	1	EVMLPQVPLR	1
R53_HUMAN	Histone-lysine N-methyltransferase SUV420H2 - Homo sapiens (Human)	894,3923	2423,1767	1,0174	27,84	2	SALMLGSAAFNHDGKRNICK	2
R53_HUMAN	Tubulin alpha-1A chain - Homo sapiens (Human)	890,2925	965,9334	0,0246	11,62	0	YQLPQVYR	1
R53_HUMAN	Tubulin alpha-1B chain - Homo sapiens (Human)	833,4362	1717,8747	0,1678	80,84	0	NLDERPPTTLNLR	2
R53_HUMAN	Tubulin alpha-1B chain - Homo sapiens (Human)	861,6945	1863,8871	0,9607	63,74	0	AVCALSNITVAEAWAR	1
R53_HUMAN	Tubulin alpha-1B chain - Homo sapiens (Human)	917,9508	1700,9385	0,276	48,41	0	APVDALEPTVIDEYR	1
R53_HUMAN	Tubulin alpha-1B chain - Homo sapiens (Human)	823,0939	2749,284	1,2461	32,74	0	AVYEQLSVALETNACFEFANQMK	1
R53_HUMAN	Tubulin alpha-1B chain - Homo sapiens (Human)	603,6311	2785,2789	0,8811	30,74	0	AVYEQLSVALETNACFEFANQMK	1
R53_HUMAN	Tubulin alpha-1B chain - Homo sapiens (Human)	1618,1160	1014,0476	0,1768	67,36	0	DVMAAATIK	1
R53_HUMAN	Tubulin alpha-1B chain - Homo sapiens (Human)	689,0207	3245,3094	0,9198	85,43	1	EDMAALEQYEEVGVDRVGESEEEGEEY	1
R53_HUMAN	Tubulin alpha-1B chain - Homo sapiens (Human)	744,7039	1755,9559	0,6517	60,28	0	HFPLATYAPVSAEK	1
R53_HUMAN	Tubulin alpha-1B chain - Homo sapiens (Human)	800,1944	1385,8967	0,2611	62,01	1	LKHKFDLIMYK	1
R53_HUMAN	Tubulin alpha-1B chain - Homo sapiens (Human)	833,3579	1486,8719	0,5211	56,85	0	LISQVSRITVALNR	1
R53_HUMAN	Tubulin alpha-1B chain - Homo sapiens (Human)	783,2030	1717,8747	0,4965	30,84	0	NLDERPPTTLNLR	1
R53_HUMAN	Tubulin alpha-1B chain - Homo sapiens (Human)	623,2119	1674,7013	0,7255	28,79	0	SIQFYDQKGFQK	2
R53_HUMAN	Tubulin alpha-1B chain - Homo sapiens (Human)	633,4175	1685,7443	0,7084	88,85	0	VGVNYQPTVYFQDGLAK	1
R53_HUMAN	Tubulin alpha-1B chain - Homo sapiens (Human)	624,475	1246,5463	0,7024	28,28	0	YMAQGLLYR	1
R53_HUMAN	Tubulin alpha-1C chain - Homo sapiens (Human)	924,1951	1294,5403	0,2901	60,39	0	AVCALSNITVAEAWAR	1
R53_HUMAN	Tubulin alpha-1C chain - Homo sapiens (Human)	833,4362	1846,8815	0,4142	81,62	2	AVCALSNITVAEAWAR	2
R53_HUMAN	Tubulin alpha-1C chain - Homo sapiens (Human)	861,6945	1865,8764	-1,0168	7,5	0	APVDALEPTVIDEYR	1
R53_HUMAN	Tubulin alpha-1C chain - Homo sapiens (Human)	303,6311	1014,0476	0,4168	67,63	0	DVMAAATIK	2
R53_HUMAN	Tubulin alpha-1C chain - Homo sapiens (Human)	872,9791	1755,9559	0,5817	60,28	0	HFPLATYAPVSAEK	1
R53_HUMAN	Tubulin alpha-1C chain - Homo sapiens (Human)	689,0207	1386,0468	0,2611	62,01	1	LKHKFDLIMYK	1

REFERENCES BIBLIOGRAPHIQUES

- (2004) Finishing the euchromatic sequence of the human genome. *Nature*, 431(7011), 931-45.
- (2008) Moving ahead with an international human epigenome project. *Nature*, 454(7205), 711-5.
- Aagaard, L., Schmid, M., Warburton, P. & Jenuwein, T. (2000) Mitotic phosphorylation of suv39h1, a novel component of active centromeres, coincides with transient accumulation at mammalian centromeres. *J Cell Sci*, 113 (Pt 5)(817-29).
- Aasland, R. & Stewart, A. F. (1995) The chromo shadow domain, a second chromo domain in heterochromatin-binding protein 1, hp1. *Nucleic Acids Res*, 23(16), 3168-73.
- Adams, J. M., Harris, A. W., Pinkert, C. A., Corcoran, L. M., Alexander, W. S., Cory, S., Palmiter, R. D. & Brinster, R. L. (1985) The c-myc oncogene driven by immunoglobulin enhancers induces lymphoid malignancy in transgenic mice. *Nature*, 318(6046), 533-8.
- Agger, K., Cloos, P. A., Christensen, J., Pasini, D., Rose, S., Rappsilber, J., Issaeva, I., Canaani, E., Salcini, A. E. & Helin, K. (2007) Utx and jmjd3 are histone h3k27 demethylases involved in hox gene regulation and development. *Nature*, 449(7163), 731-4.
- Ait-Si-Ali, S., Guasconi, V., Fritsch, L., Yahi, H., Sekhri, R., Naguibneva, I., Robin, P., Cabon, F., Polesskaya, A. & Harel-Bellan, A. (2004) A suv39h-dependent mechanism for silencing s-phase genes in differentiating but not in cycling cells. *EMBO J*, 23(3), 605-15.
- Akasaka, T., Kanno, M., Balling, R., Mieza, M. A., Taniguchi, M. & Koseki, H. (1996) A role for mel-18, a polycomb group-related vertebrate gene, during theanteroposterior specification of the axial skeleton. *Development*, 122(5), 1513-22.
- Allan, J., Mitchell, T., Harborne, N., Bohm, L. & Crane-Robinson, C. (1986) Roles of h1 domains in determining higher order chromatin structure and h1 location. *J Mol Biol*, 187(4), 591-601.
- Allfrey, V. G., Faulkner, R. & Mirsky, A. E. (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).
- Aloy, P., Bottcher, B., Ceulemans, H., Leutwein, C., Mellwig, C., Fischer, S., Gavin, A. C., Bork, P., Superti-Furga, G., Serrano, L. & Russell, R. B. (2004) Structure-based assembly of protein complexes in yeast. *Science*, 303(5666), 2026-9.
- Altenburg, B. C., Via, D. P. & Steiner, S. H. (1976) Modification of the phenotype of murine sarcoma virus-transformed cells by sodium butyrate. Effects on morphology and cytoskeletal elements. *Exp Cell Res*, 102(2), 223-31.
- Alvarez-Venegas, R., Pien, S., Sadler, M., Witmer, X., Grossniklaus, U. & Avramova, Z. (2003) Atx-1, an arabidopsis homolog of trithorax, activates flower homeotic genes. *Curr Biol*, 13(8), 627-37.
- Ame, J. C., Spenlehauer, C. & De Murcia, G. (2004) The parp superfamily. *Bioessays*, 26(8), 882-93.
- Andersen, E. C. & Horvitz, H. R. (2007) Two c. Elegans histone methyltransferases repress lin-3 egf transcription to inhibit vulval development. *Development*, 134(16), 2991-9.
- Angrand, P. O., Segura, I., Volkel, P., Ghidelli, S., Terry, R., Brajenovic, M., Vintersten, K., Klein, R., Superti-Furga, G., Drewes, G., Kuster, B., Bouwmeester, T. & Acker-Palmer, A. (2006) Transgenic mouse proteomics identifies new 14-3-3-associated proteins involved in cytoskeletal rearrangements and cell signaling. *Mol Cell Proteomics*, 5(12), 2211-27.
- Aoto, T., Saitoh, N., Sakamoto, Y., Watanabe, S. & Nakao, M. (2008) Polycomb group protein-associated chromatin is reproduced in post-mitotic g1 phase and is required for s phase progression. *J Biol Chem*, 283(27), 18905-15.
- Arents, G., Burlingame, R. W., Wang, B. C., Love, W. E. & Moudrianakis, E. N. (1991) The nucleosomal core histone octamer at 3.1 a resolution: A tripartite protein assembly and a left-handed superhelix. *Proc Natl Acad Sci U S A*, 88(22), 10148-52.
- Arrighoni, R., Alam, S. L., Wamstad, J. A., Bardwell, V. J., Sundquist, W. I. & Schreiber-Agus, N. (2006) The polycomb-associated protein rybp is a ubiquitin binding protein. *FEBS Lett*, 580(26), 6233-41.
- Auth, T., Kunkel, E. & Grummt, F. (2006) Interaction between hp1alpha and replication proteins in mammalian cells. *Exp Cell Res*, 312(17), 3349-59.
- Ayoub, N., Jeyasekharan, A. D., Bernal, J. A. & Venkitaraman, A. R. (2008) Hp1-beta mobilization promotes chromatin changes that initiate the DNA damage response. *Nature*, 453(7195), 682-6.
- Ayyanathan, K., Lechner, M. S., Bell, P., Maul, G. G., Schultz, D. C., Yamada, Y., Tanaka, K., Torigoe, K. & Rauscher, F. J., 3rd (2003) Regulated recruitment of hp1 to a euchromatic gene induces mitotically heritable, epigenetic gene silencing: A mammalian cell culture model of gene variegation. *Genes Dev*, 17(15), 1855-69.
- Azuara, V., Perry, P., Sauer, S., Spivakov, M., Jorgensen, H. F., John, R. M., Gouti, M., Casanova, M., Warnes, G., Merkenschlager, M. & Fisher, A. G. (2006) Chromatin signatures of pluripotent cell lines. *Nat Cell Biol*, 8(5), 532-8.
- Badugu, R., Yoo, Y., Singh, P. B. & Kellum, R. (2005) Mutations in the heterochromatin protein 1 (hp1) hinge domain affect hp1 protein interactions and chromosomal distribution. *Chromosoma*, 113(7), 370-84.
- Ball, P. (2003) Portrait of a molecule. *Nature*, 421(6921), 421-2.
- Bannister, A. J., Zegerman, P., Partridge, J. F., Miska, E. A., Thomas, J. O., Allshire, R. C. & Kouzarides, T. (2001) Selective recognition of methylated lysine 9 on histone h3 by the hp1 chromo domain. *Nature*,

- 410(6824), 120-4.
- Bardos, J. I., Saurin, A. J., Tissot, C., Duprez, E. & Freemont, P. S. (2000) Hpc3 is a new human polycomb orthologue that interacts and associates with ring1 and bmi1 and has transcriptional repression properties. *J Biol Chem*, 275(37), 28785-92.
- Barski, A., Cuddapah, S., Cui, K., Roh, T. Y., Schones, D. E., Wang, Z., Wei, G., Chepelev, I. & Zhao, K. (2007) High-resolution profiling of histone methylations in the human genome. *Cell*, 129(4), 823-37.
- Bartholomeeusen, K., Christ, F., Hendrix, J., Rain, J. C., Emiliani, S., Benarous, R., Debyser, Z., Gijsbers, R. & De Rijck, J. (2009) Lens epithelium-derived growth factor/p75 interacts with the transposase-derived dde domain of pogz. *J Biol Chem*, 284(17), 11467-77.
- Baulcombe, D. (2004) Rna silencing in plants. *Nature*, 431(7006), 356-63.
- Benzinger, A., Muster, N., Koch, H. B., Yates, J. R., 3rd & Hermeking, H. (2005) Targeted proteomic analysis of 14-3-3 sigma, a p53 effector commonly silenced in cancer. *Mol Cell Proteomics*, 4(6), 785-95.
- Bernstein, B. E., Mikkelsen, T. S., Xie, X., Kamal, M., Huebert, D. J., Cuff, J., Fry, B., Meissner, A., Wernig, M., Plath, K., Jaenisch, R., Wagschal, A., Feil, R., Schreiber, S. L. & Lander, E. S. (2006) A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell*, 125(2), 315-26.
- Bernstein, E. & Allis, C. D. (2005) Rna meets chromatin. *Genes Dev*, 19(14), 1635-55.
- Bernstein, E. & Hake, S. B. (2006) The nucleosome: A little variation goes a long way. *Biochem Cell Biol*, 84(4), 505-17.
- Bertwistle, D., Sugimoto, M. & Sherr, C. J. (2004) Physical and functional interactions of the arf tumor suppressor protein with nucleophosmin/b23. *Mol Cell Biol*, 24(3), 985-96.
- Bird, A. (2002) DNA methylation patterns and epigenetic memory. *Genes Dev*, 16(1), 6-21.
- Blower, M. D. & Karpen, G. H. (2001) The role of drosophila cid in kinetochore formation, cell-cycle progression and heterochromatin interactions. *Nat Cell Biol*, 3(8), 730-9.
- Bocconi, P., Macgrogan, D., Scandura, J. M. & Nimer, S. D. (2003) The human I(3)mbt polycomb group protein is a transcriptional repressor and interacts physically and functionally with tel (etv6). *J Biol Chem*, 278(17), 15412-20.
- Boffa, L. C., Vidali, G., Mann, R. S. & Allfrey, V. G. (1978) Suppression of histone deacetylation in vivo and in vitro by sodium butyrate. *J Biol Chem*, 253(10), 3364-6.
- Botuyan, M. V., Lee, J., Ward, I. M., Kim, J. E., Thompson, J. R., Chen, J. & Mer, G. (2006) Structural basis for the methylation state-specific recognition of histone h4-k20 by 53bp1 and crb2 in DNA repair. *Cell*, 127(7), 1361-73.
- Bourc'his, D., Xu, G. L., Lin, C. S., Bollman, B. & Bestor, T. H. (2001) Dnmt3l and the establishment of maternal genomic imprints. *Science*, 294(5551), 2536-9.
- Bouwmeester, T., Bauch, A., Ruffner, H., Angrand, P. O., Bergamini, G., Croughton, K., Cruciat, C., Eberhard, D., Gagneur, J., Ghidelli, S., Hopf, C., Huhse, B., Mangano, R., Michon, A. M., Schirle, M., Schlegl, J., Schwab, M., Stein, M. A., Bauer, A., Casari, G., Drewes, G., Gavin, A. C., Jackson, D. B., Joberty, G., Neubauer, G., Rick, J., Kuster, B. & Superti-Furga, G. (2004) A physical and functional map of the human tnfr-alpha/nf-kappa b signal transduction pathway. *Nat Cell Biol*, 6(2), 97-105.
- Boyer, L. A., Plath, K., Zeitlinger, J., Brambrink, T., Medeiros, L. A., Lee, T. I., Levine, S. S., Wernig, M., Tajonar, A., Ray, M. K., Bell, G. W., Otte, A. P., Vidal, M., Gifford, D. K., Young, R. A. & Jaenisch, R. (2006) Polycomb complexes repress developmental regulators in murine embryonic stem cells. *Nature*, 441(7091), 349-53.
- Bracken, A. P., Dietrich, N., Pasini, D., Hansen, K. H. & Helin, K. (2006) Genome-wide mapping of polycomb target genes unravels their roles in cell fate transitions. *Genes Dev*, 20(9), 1123-36.
- Brajenovic, M., Joberty, G., Kuster, B., Bouwmeester, T. & Drewes, G. (2004) Comprehensive proteomic analysis of human par protein complexes reveals an interconnected protein network. *J Biol Chem*, 279(13), 12804-11.
- Brasher, S. V., Smith, B. O., Fogh, R. H., Nietlispach, D., Thiru, A., Nielsen, P. R., Broadhurst, R. W., Ball, L. J., Murzina, N. V. & Laue, E. D. (2000) The structure of mouse hp1 suggests a unique mode of single peptide recognition by the shadow chromo domain dimer. *EMBO J*, 19(7), 1587-97.
- Brickner, D. G., Cajigas, I., Fondufe-Mittendorf, Y., Ahmed, S., Lee, P. C., Widom, J. & Brickner, J. H. (2007) H2a.Z-mediated localization of genes at the nuclear periphery confers epigenetic memory of previous transcriptional state. *PLoS Biol*, 5(4), e81.
- Brown, J. L., Fritsch, C., Mueller, J. & Kassis, J. A. (2003) The drosophila pho-like gene encodes a yy1-related DNA binding protein that is redundant with pleiohomeotic in homeotic gene silencing. *Development*, 130(2), 285-94.
- Brown, J. L., Mucci, D., Whiteley, M., Dirksen, M. L. & Kassis, J. A. (1998) The drosophila polycomb group gene pleiohomeotic encodes a DNA binding protein with homology to the transcription factor yy1. *Mol Cell*, 1(7), 1057-64.
- Buchenau, P., Hodgson, J., Strutt, H. & Arndt-Jovin, D. J. (1998) The distribution of polycomb-group proteins during cell division and development in drosophila embryos: Impact on models for silencing. *J Cell Biol*, 141(2), 469-81.
- Buchou, T., Vernet, M., Blond, O., Jensen, H. H., Pointu, H., Olsen, B. B., Cochet, C., Issinger, O. G. & Boldyreff, B. (2003) Disruption of the regulatory beta subunit of protein kinase ck2 in mice leads to a cell-autonomous defect and early embryonic lethality. *Mol Cell Biol*, 23(3), 908-15.
- Buchou, T., Vernet, M., Blond, O., Jensen, H. H., Pointu, H., Olsen, B. B., Cochet, C., Issinger, O. G. & Boldyreff, B. (2003) Disruption of the regulatory beta subunit of protein kinase ck2 in mice leads to a cell-

- autonomous defect and early embryonic lethality. *Mol Cell Biol*, 23(3), 908-15.
- Buchwald, G., Van Der Stoop, P., Weichenrieder, O., Perrakis, A., Van Lohuizen, M. & Sixma, T. K. (2006) Structure and e3-ligase activity of the ring-ring complex of polycomb proteins bmi1 and ring1b. *EMBO J*, 25(11), 2465-74.
- Burckstummer, T., Bennett, K. L., Preradovic, A., Schutze, G., Hantschel, O., Superti-Furga, G. & Bauch, A. (2006) An efficient tandem affinity purification procedure for interaction proteomics in mammalian cells. *Nat Methods*, 3(12), 1013-9.
- Burma, S., Chen, B. P., Murphy, M., Kurimasa, A. & Chen, D. J. (2001) Atm phosphorylates histone h2ax in response to DNA double-strand breaks. *J Biol Chem*, 276(45), 42462-7.
- Butland, G., Peregrin-Alvarez, J. M., Li, J., Yang, W., Yang, X., Canadien, V., Starostine, A., Richards, D., Beattie, B., Krogan, N., Davey, M., Parkinson, J., Greenblatt, J. & Emili, A. (2005) Interaction network containing conserved and essential protein complexes in *escherichia coli*. *Nature*, 433(7025), 531-7.
- Byvoet, P., Shepherd, G. R., Hardin, J. M. & Noland, B. J. (1972) The distribution and turnover of labeled methyl groups in histone fractions of cultured mammalian cells. *Arch Biochem Biophys*, 148(2), 558-67.
- Cam, H. P., Sugiyama, T., Chen, E. S., Chen, X., Fitzgerald, P. C. & Grewal, S. I. (2005) Comprehensive analysis of heterochromatin- and rna-mediated epigenetic control of the fission yeast genome. *Nat Genet*, 37(8), 809-19.
- Cameron, E. E., Bachman, K. E., Myohanen, S., Herman, J. G. & Baylin, S. B. (1999) Synergy of demethylation and histone deacetylase inhibition in the re-expression of genes silenced in cancer. *Nat Genet*, 21(1), 103-7.
- Cammas, F., Oulad-Abdelghani, M., Vonesch, J. L., Huss-Garcia, Y., Chambon, P. & Losson, R. (2002) Cell differentiation induces tif1beta association with centromeric heterochromatin via an hp1 interaction. *J Cell Sci*, 115(Pt 17), 3439-48.
- Cao, R., Tsukada, Y. & Zhang, Y. (2005) Role of bmi-1 and ring1a in h2a ubiquitylation and hox gene silencing. *Mol Cell*, 20(6), 845-54.
- Cao, R., Wang, H., He, J., Erdjument-Bromage, H., Tempst, P. & Zhang, Y. (2008) Role of hphf1 in h3k27 methylation and hox gene silencing. *Mol Cell Biol*, 28(5), 1862-72.
- Cao, R., Wang, L., Wang, H., Xia, L., Erdjument-Bromage, H., Tempst, P., Jones, R. S. & Zhang, Y. (2002) Role of histone h3 lysine 27 methylation in polycomb-group silencing. *Science*, 298(5595), 1039-43.
- Cao, R. & Zhang, Y. (2004) The functions of e(z)/ezh2-mediated methylation of lysine 27 in histone h3. *Curr Opin Genet Dev*, 14(2), 155-64.
- Caretti, G., Di Padova, M., Micales, B., Lyons, G. E. & Sartorelli, V. (2004) The polycomb ezh2 methyltransferase regulates muscle gene expression and skeletal muscle differentiation. *Genes Dev*, 18(21), 2627-38.
- Chadwick, B. P. & Willard, H. F. (2001) A novel chromatin protein, distantly related to histone h2a, is largely excluded from the inactive x chromosome. *J Cell Biol*, 152(2), 375-84.
- Chamberlain, S. J., Yee, D. & Magnuson, T. (2008) Polycomb repressive complex 2 is dispensable for maintenance of embryonic stem cell pluripotency. *Stem Cells*, 26(6), 1496-505.
- Chang, B., Chen, Y., Zhao, Y. & Bruck, R. K. (2007) Jmjd6 is a histone arginine demethylase. *Science*, 318(5849), 444-7.
- Chaudhri, M., Scarabel, M. & Aitken, A. (2003) Mammalian and yeast 14-3-3 isoforms form distinct patterns of dimers in vivo. *Biochem Biophys Res Commun*, 300(3), 679-85.
- Chen, X., Hiller, M., Sancak, Y. & Fuller, M. T. (2005) Tissue-specific tafs counteract polycomb to turn on terminal differentiation. *Science*, 310(5749), 869-72.
- Cheung, P., Allis, C. D. & Sassone-Corsi, P. (2000) Signaling to chromatin through histone modifications. *Cell*, 103(2), 263-71.
- Cheung, W. L., Ajiro, K., Samejima, K., Kloc, M., Cheung, P., Mizzen, C. A., Beeser, A., Etkin, L. D., Chernoff, J., Earnshaw, W. C. & Allis, C. D. (2003) Apoptotic phosphorylation of histone h2b is mediated by mammalian sterile twenty kinase. *Cell*, 113(4), 507-17.
- Cheung, W. L., Turner, F. B., Krishnamoorthy, T., Wolner, B., Ahn, S. H., Foley, M., Dorsey, J. A., Peterson, C. L., Berger, S. L. & Allis, C. D. (2005) Phosphorylation of histone h4 serine 1 during DNA damage requires casein kinase ii in *s. Cerevisiae*. *Curr Biol*, 15(7), 656-60.
- Cheutin, T., Mcnairn, A. J., Jenuwein, T., Gilbert, D. M., Singh, P. B. & Misteli, T. (2003) Maintenance of stable heterochromatin domains by dynamic hp1 binding. *Science*, 299(5607), 721-5.
- Choo, J. H., Kim, J. D., Chung, J. H., Stubbs, L. & Kim, J. (2006) Allele-specific deposition of macroh2a1 in imprinting control regions. *Hum Mol Genet*, 15(5), 717-24.
- Chowdhury, D., Xu, X., Zhong, X., Ahmed, F., Zhong, J., Liao, J., Dykxhoorn, D. M., Weinstock, D. M., Pfeifer, G. P. & Lieberman, J. (2008) A pp4-phosphatase complex dephosphorylates gamma-h2ax generated during DNA replication. *Mol Cell*, 31(1), 33-46.
- Christensen, J., Agger, K., Cloos, P. A., Pasini, D., Rose, S., Sennels, L., Rappsilber, J., Hansen, K. H., Salcini, A. E. & Helin, K. (2007) Rbp2 belongs to a family of demethylases, specific for tri- and dimethylated lysine 4 on histone 3. *Cell*, 128(6), 1063-76.
- Chu, F., Nusinow, D. A., Chalkley, R. J., Plath, K., Panning, B. & Burlingame, A. L. (2006) Mapping post-translational modifications of the histone variant macroh2a1 using tandem mass spectrometry. *Mol Cell Proteomics*, 5(1), 194-203.
- Churikov, D., Siino, J., Svetlova, M., Zhang, K., Gineitis, A., Morton Bradbury, E. & Zalensky, A. (2004) Novel human testis-specific histone h2b encoded by the interrupted gene on the x chromosome. *Genomics*, 84(4), 745-56.

- Clegg, N. J., Honda, B. M., Whitehead, I. P., Grigliatti, T. A., Wakimoto, B., Brock, H. W., Lloyd, V. K. & Sinclair, D. A. (1998) Suppressors of position-effect variegation in *Drosophila melanogaster* affect expression of the heterochromatic gene *light* in the absence of a chromosome rearrangement. *Genome*, 41(4), 495-503.
- Cloos, P. A., Christensen, J., Agger, K., Maiolica, A., Rappsilber, J., Antal, T., Hansen, K. H. & Helin, K. (2006) The putative oncogene *gasc1* demethylates tri- and dimethylated lysine 9 on histone H3. *Nature*, 442(7100), 307-11.
- Core, N., Joly, F., Boned, A. & Djabali, M. (2004) Disruption of *e2f* signaling suppresses the *ink4a*-induced proliferative defect in *m33*-deficient mice. *Oncogene*, 23(46), 7660-8.
- Cosgrove, M. S., Boeke, J. D. & Wolberger, C. (2004) Regulated nucleosome mobility and the histone code. *Nat Struct Mol Biol*, 11(11), 1037-43.
- Costanzi, C. & Pehrson, J. R. (1998) Histone macroH2A1 is concentrated in the inactive X chromosome of female mammals. *Nature*, 393(6685), 599-601.
- Costello, J. F., Fruhwald, M. C., Smiraglia, D. J., Rush, L. J., Robertson, G. P., Gao, X., Wright, F. A., Feramisco, J. D., Peltomaki, P., Lang, J. C., Schuller, D. E., Yu, L., Bloomfield, C. D., Caligiuri, M. A., Yates, A., Nishikawa, R., Su Huang, H., Petrelli, N. J., Zhang, X., O'dorisio, M. S., Held, W. A., Cavenee, W. K. & Plass, C. (2000) Aberrant CpG-island methylation has non-random and tumour-type-specific patterns. *Nat Genet*, 24(2), 132-8.
- Couture, J. F., Collazo, E. & Trievel, R. C. (2006) Molecular recognition of histone H3 by the WD40 protein Wdr5. *Nat Struct Mol Biol*, 13(8), 698-703.
- Cowell, I. G., Aucott, R., Mahadevaiah, S. K., Burgoyne, P. S., Huskisson, N., Bongiorno, S., Prantera, G., Fanti, L., Pimpinelli, S., Wu, R., Gilbert, D. M., Shi, W., Fundele, R., Morrison, H., Jeppesen, P. & Singh, P. B. (2002) Heterochromatin, hp1 and methylation at lysine 9 of histone H3 in animals. *Chromosoma*, 111(1), 22-36.
- Cowieson, N. P., Partridge, J. F., Allshire, R. C. & McLaughlin, P. J. (2000) Dimerisation of a chromatin shadow domain and distinctions from the chromodomain as revealed by structural analysis. *Curr Biol*, 10(9), 517-25.
- Cryderman, D. E., Grade, S. K., Li, Y., Fanti, L., Pimpinelli, S. & Wallrath, L. L. (2005) Role of *Drosophila* hp1 in euchromatic gene expression. *Dev Dyn*, 232(3), 767-74.
- Cuthbert, G. L., Daujat, S., Snowden, A. W., Erdjument-Bromage, H., Hagiwara, T., Yamada, M., Schneider, R., Gregory, P. D., Tempst, P., Bannister, A. J. & Kouzarides, T. (2004) Histone deimination antagonizes arginine methylation. *Cell*, 118(5), 545-53.
- Czernin, B., Melfi, R., McCabe, D., Seitz, V., Imhof, A. & Pirrotta, V. (2002) *Drosophila* enhancer of *zeste/esc* complexes have a histone H3 methyltransferase activity that marks chromosomal polycomb sites. *Cell*, 111(2), 185-96.
- Czubaty, A., Girstun, A., Kowalska-Loth, B., Trzcinska, A. M., Purta, E., Winczura, A., Grajkowski, W. & Staron, K. (2005) Proteomic analysis of complexes formed by human topoisomerase I. *Biochim Biophys Acta*, 1749(1), 133-41.
- Dai, J., Sultan, S., Taylor, S. S. & Higgins, J. M. (2005) The kinase Haspin is required for mitotic histone H3 Thr 3 phosphorylation and normal metaphase chromosome alignment. *Genes Dev*, 19(4), 472-88.
- Daulat, A. M., Maurice, P., Froment, C., Guillaume, J. L., Broussard, C., Monsarrat, B., Delagrèze, P. & Jockers, R. (2007) Purification and identification of G protein-coupled receptor protein complexes under native conditions. *Mol Cell Proteomics*, 6(5), 835-44.
- Davie, J. R. & Chadee, D. N. (1998) Regulation and regulatory parameters of histone modifications. *J Cell Biochem Suppl*, 30-31(203-13).
- Davie, J. R. & Spencer, V. A. (2001) Signal transduction pathways and the modification of chromatin structure. *Prog Nucleic Acid Res Mol Biol*, 65(299-340).
- De Lucia, F., Ni, J. Q., Vaillant, C. & Sun, F. L. (2005) Hp1 modulates the transcription of cell-cycle regulators in *Drosophila melanogaster*. *Nucleic Acids Res*, 33(9), 2852-8.
- De Napoles, M., Mermoud, J. E., Wakao, R., Tang, Y. A., Endoh, M., Appanah, R., Nesterova, T. B., Silva, J., Otte, A. P., Vidal, M., Koseki, H. & Brockdorff, N. (2004) Polycomb group proteins ring1a/b link ubiquitylation of histone H2A to heritable gene silencing and X inactivation. *Dev Cell*, 7(5), 663-76.
- De Santa, F., Totaro, M. G., Prosperini, E., Notarbartolo, S., Testa, G. & Natoli, G. (2007) The histone H3 lysine-27 demethylase Jmjd3 links inflammation to inhibition of polycomb-mediated gene silencing. *Cell*, 130(6), 1083-94.
- De Wit, E., Greil, F. & Van Steensel, B. (2007) High-resolution mapping reveals links of hp1 with active and inactive chromatin components. *PLoS Genet*, 3(3), e38.
- Decamillis, M., Cheng, N. S., Pierre, D. & Brock, H. W. (1992) The polyhomeotic gene of *Drosophila* encodes a chromatin protein that shares polytene chromosome-binding sites with polycomb. *Genes Dev*, 6(2), 223-32.
- Del Mar Lorente, M., Marcos-Gutierrez, C., Perez, C., Schoorlemmer, J., Ramirez, A., Magin, T. & Vidal, M. (2000) Loss- and gain-of-function mutations show a polycomb group function for ring1a in mice. *Development*, 127(23), 5093-100.
- Deshpande, A. M., Akunowicz, J. D., Reveles, X. T., Patel, B. B., Saria, E. A., Gorlick, R. G., Naylor, S. L., Leach, R. J. & Hansen, M. F. (2007) Phc3, a component of the hprc-h complex, associates with e2f6 during G0 and is lost in osteosarcoma tumors. *Oncogene*, 26(12), 1714-22.
- Dhalluin, C., Carlson, J. E., Zeng, L., He, C., Aggarwal, A. K. & Zhou, M. M. (1999) Structure and ligand of a

- histone acetyltransferase bromodomain. *Nature*, 399(6735), 491-6.
- Dialynas, G. K., Terjung, S., Brown, J. P., Aucott, R. L., Baron-Luhr, B., Singh, P. B. & Georgatos, S. D. (2007) Plasticity of hp1 proteins in mammalian cells. *J Cell Sci*, 120(Pt 19), 3415-24.
- Dietrich, N., Bracken, A. P., Trinh, E., Schjerling, C. K., Koseki, H., Rappsilber, J., Helin, K. & Hansen, K. H. (2007) Bypass of senescence by the polycomb group protein cbx8 through direct binding to the ink4a-arf locus. *EMBO J*, 26(6), 1637-48.
- Dion, M. F., Altschuler, S. J., Wu, L. F. & Rando, O. J. (2005) Genomic characterization reveals a simple histone h4 acetylation code. *Proc Natl Acad Sci U S A*, 102(15), 5501-6.
- Dodd, I. B., Micheelsen, M. A., Sneppen, K. & Thon, G. (2007) Theoretical analysis of epigenetic cell memory by nucleosome modification. *Cell*, 129(4), 813-22.
- Dougherty, W. G., Parks, T. D., Cary, S. M., Bazan, J. F. & Fletterick, R. J. (1989) Characterization of the catalytic residues of the tobacco etch virus 49-kda proteinase. *Virology*, 172(1), 302-10.
- Dryhurst, D., Thambirajah, A. A. & Ausio, J. (2004) New twists on h2a.Z: A histone variant with a controversial structural and functional past. *Biochem Cell Biol*, 82(4), 490-7.
- Dura, J. M., Randsholt, N. B., Deatrick, J., Erk, I., Santamaria, P., Freeman, J. D., Freeman, S. J., Weddell, D. & Brock, H. W. (1987) A complex genetic locus, polyhomeotic, is required for segmental specification and epidermal development in d. *Melanogaster*. *Cell*, 51(5), 829-39.
- Dyda, F., Hickman, A. B., Jenkins, T. M., Engelman, A., Craigie, R. & Davies, D. R. (1994) Crystal structure of the catalytic domain of hiv-1 integrase: Similarity to other polynucleotidyl transferases. *Science*, 266(5193), 1981-6.
- Ebert, A., Schotta, G., Lein, S., Kubicek, S., Krauss, V., Jenuwein, T. & Reuter, G. (2004) Su(var) genes regulate the balance between euchromatin and heterochromatin in drosophila. *Genes Dev*, 18(23), 2973-83.
- Eissenberg, J. C., Ge, Y. W. & Hartnett, T. (1994) Increased phosphorylation of hp1, a heterochromatin-associated protein of drosophila, is correlated with heterochromatin assembly. *J Biol Chem*, 269(33), 21315-21.
- Eissenberg, J. C., James, T. C., Foster-Hartnett, D. M., Hartnett, T., Ngan, V. & Elgin, S. C. (1990) Mutation in a heterochromatin-specific chromosomal protein is associated with suppression of position-effect variegation in drosophila melanogaster. *Proc Natl Acad Sci U S A*, 87(24), 9923-7.
- Elderkin, S., Maertens, G. N., Endoh, M., Mallery, D. L., Morrice, N., Koseki, H., Peters, G., Brockdorff, N. & Hiom, K. (2007) A phosphorylated form of mel-18 targets the ring1b histone h2a ubiquitin ligase to chromatin. *Mol Cell*, 28(1), 107-20.
- Elgin, S. C. & Grewal, S. I. (2003) Heterochromatin: Silence is golden. *Curr Biol*, 13(23), R895-8.
- Esteller, M., Silva, J. M., Dominguez, G., Bonilla, F., Matias-Guiu, X., Lerma, E., Bussaglia, E., Prat, J., Harkes, I. C., Repasky, E. A., Gabrielson, E., Schutte, M., Baylin, S. B. & Herman, J. G. (2000) Promoter hypermethylation and brca1 inactivation in sporadic breast and ovarian tumors. *J Natl Cancer Inst*, 92(7), 564-9.
- Fang, J., Chen, T., Chadwick, B., Li, E. & Zhang, Y. (2004) Ring1b-mediated h2a ubiquitination associates with inactive x chromosomes and is involved in initiation of x inactivation. *J Biol Chem*, 279(51), 52812-5.
- Fanti, L., Berloco, M., Piacentini, L. & Pimpinelli, S. (2003) Chromosomal distribution of heterochromatin protein 1 (hp1) in drosophila: A cytological map of euchromatic hp1 binding sites. *Genetica*, 117(2-3), 135-47.
- Fanti, L., Giovannazzo, G., Berloco, M. & Pimpinelli, S. (1998) The heterochromatin protein 1 prevents telomere fusions in drosophila. *Mol Cell*, 2(5), 527-38.
- Faust, R. A., Tawfic, S., Davis, A. T., Bubash, L. A. & Ahmed, K. (2000) Antisense oligonucleotides against protein kinase ck2-alpha inhibit growth of squamous cell carcinoma of the head and neck in vitro. *Head Neck*, 22(4), 341-6.
- Feinberg, A. P. & Vogelstein, B. (1983) Hypomethylation distinguishes genes of some human cancers from their normal counterparts. *Nature*, 301(5895), 89-92.
- Feng, Q., Wang, H., Ng, H. H., Erdjument-Bromage, H., Tempst, P., Struhl, K. & Zhang, Y. (2002) Methylation of h3-lysine 79 is mediated by a new family of hmtases without a set domain. *Curr Biol*, 12(12), 1052-8.
- Festenstein, R., Pagakis, S. N., Hiragami, K., Lyon, D., Verreault, A., Sekkali, B. & Kioussis, D. (2003) Modulation of heterochromatin protein 1 dynamics in primary mammalian cells. *Science*, 299(5607), 719-21.
- Fields, S. & Song, O. (1989) A novel genetic system to detect protein-protein interactions. *Nature*, 340(6230), 245-6.
- Finnerty, J. R. & Martindale, M. Q. (1998) The evolution of the hox cluster: Insights from outgroups. *Curr Opin Genet Dev*, 8(6), 681-7.
- Fischle, W., Tseng, B. S., Dormann, H. L., Ueberheide, B. M., Garcia, B. A., Shabanowitz, J., Hunt, D. F., Funabiki, H. & Allis, C. D. (2005) Regulation of hp1-chromatin binding by histone h3 methylation and phosphorylation. *Nature*, 438(7071), 1116-22.
- Fischle, W., Wang, Y., Jacobs, S. A., Kim, Y., Allis, C. D. & Khorasanizadeh, S. (2003) Molecular basis for the discrimination of repressive methyl-lysine marks in histone h3 by polycomb and hp1 chromodomains. *Genes Dev*, 17(15), 1870-81.
- Flanagan, J. F., Mi, L. Z., Chruszcz, M., Cymborowski, M., Clines, K. L., Kim, Y., Minor, W., Rastinejad, F. & Khorasanizadeh, S. (2005) Double chromodomains cooperate to recognize the methylated histone h3 tail. *Nature*, 438(7071), 1181-5.
- Font-Burgada, J., Rossell, D., Auer, H. & Azorin, F. (2008) Drosophila hp1c isoform interacts with the zinc-finger proteins woc and relative-of-woc to regulate gene expression. *Genes Dev*, 22(21), 3007-23.
- Fraga, M. F., Herranz, M., Espada, J., Ballestar, E., Paz, M. F., Ropero, S., Erkek, E., Bozdogan, O., Peinado, H.,

- Niveleau, A., Mao, J. H., Balmain, A., Cano, A. & Esteller, M. (2004) A mouse skin multistage carcinogenesis model reflects the aberrant DNA methylation patterns of human tumors. *Cancer Res*, 64(16), 5527-34.
- Francis, N. J., Kingston, R. E. & Woodcock, C. L. (2004) Chromatin compaction by a polycomb group protein complex. *Science*, 306(5701), 1574-7.
- Franke, A., Decamillis, M., Zink, D., Cheng, N., Brock, H. W. & Paro, R. (1992) Polycomb and polyhomeotic are constituents of a multimeric protein complex in chromatin of drosophila melanogaster. *EMBO J*, 11(8), 2941-50.
- Friedman, J. R., Fredericks, W. J., Jensen, D. E., Speicher, D. W., Huang, X. P., Neilson, E. G. & Rauscher, F. J., 3rd (1996) Kap-1, a novel corepressor for the highly conserved krab repression domain. *Genes Dev*, 10(16), 2067-78.
- Fritsch, C., Brown, J. L., Kassis, J. A. & Muller, J. (1999) The DNA-binding polycomb group protein pleiohomeotic mediates silencing of a drosophila homeotic gene. *Development*, 126(17), 3905-13.
- Frodin, M., Jensen, C. J., Merienne, K. & Gammeltoft, S. (2000) A phosphoserine-regulated docking site in the protein kinase rsk2 that recruits and activates pdk1. *EMBO J*, 19(12), 2924-34.
- Fromont-Racine, M., Rain, J. C. & Legrain, P. (1997) Toward a functional analysis of the yeast genome through exhaustive two-hybrid screens. *Nat Genet*, 16(3), 277-82.
- Fujita, A., Gomes, L. R., Sato, J. R., Yamaguchi, R., Thomaz, C. E., Sogayar, M. C. & Miyano, S. (2008) Multivariate gene expression analysis reveals functional connectivity changes between normal/tumoral prostates. *BMC Syst Biol*, 2(106).
- Fujiyama, S., Yanagida, M., Hayano, T., Miura, Y., Isobe, T., Fujimori, F., Uchida, T. & Takahashi, N. (2002) Isolation and proteomic characterization of human parvulin-associating preribosomal ribonucleoprotein complexes. *J Biol Chem*, 277(26), 23773-80.
- Fukagawa, T., Nogami, M., Yoshikawa, M., Ikeno, M., Okazaki, T., Takami, Y., Nakayama, T. & Oshimura, M. (2004) Dicer is essential for formation of the heterochromatin structure in vertebrate cells. *Nat Cell Biol*, 6(8), 784-91.
- Fuks, F., Hurd, P. J., Deplus, R. & Kouzarides, T. (2003) The DNA methyltransferases associate with hp1 and the suv39h1 histone methyltransferase. *Nucleic Acids Res*, 31(9), 2305-12.
- Furuta, K., Chan, E. K., Kiyosawa, K., Reimer, G., Luderschmidt, C. & Tan, E. M. (1997) Heterochromatin protein hp1hsbeta (p25beta) and its localization with centromeres in mitosis. *Chromosoma*, 106(1), 11-9.
- Furuyama, T., Banerjee, R., Breen, T. R. & Harte, P. J. (2004) Sir2 is required for polycomb silencing and is associated with an e(z) histone methyltransferase complex. *Curr Biol*, 14(20), 1812-21.
- Gagne, J. P., Isabelle, M., Lo, K. S., Bourassa, S., Hendzel, M. J., Dawson, V. L., Dawson, T. M. & Poirier, G. G. (2008) Proteome-wide identification of poly(adp-ribose) binding proteins and poly(adp-ribose)-associated protein complexes. *Nucleic Acids Res*, 36(22), 6959-76.
- Gaillard, P. H., Martini, E. M., Kaufman, P. D., Stillman, B., Moustacchi, E. & Almouzni, G. (1996) Chromatin assembly coupled to DNA repair: A new role for chromatin assembly factor i. *Cell*, 86(6), 887-96.
- Gao, M. & Karin, M. (2005) Regulating the regulators: Control of protein ubiquitination and ubiquitin-like modifications by extracellular stimuli. *Mol Cell*, 19(5), 581-93.
- Garcia, E., Marcos-Gutierrez, C., Del Mar Lorente, M., Moreno, J. C. & Vidal, M. (1999) Rybp, a new repressor protein that interacts with components of the mammalian polycomb complex, and with the transcription factor yy1. *EMBO J*, 18(12), 3404-18.
- Garcia-Cao, M., O'sullivan, R., Peters, A. H., Jenuwein, T. & Blasco, M. A. (2004) Epigenetic regulation of telomere length in mammalian cells by the suv39h1 and suv39h2 histone methyltransferases. *Nat Genet*, 36(1), 94-9.
- Gavin, A. C., Bosche, M., Krause, R., Grandi, P., Marzioch, M., Bauer, A., Schultz, J., Rick, J. M., Michon, A. M., Cruciat, C. M., Remor, M., Hofert, C., Schelder, M., Brajenovic, M., Ruffner, H., Merino, A., Klein, K., Hudak, M., Dickson, D., Rudi, T., Gnau, V., Bauch, A., Bastuck, S., Huhse, B., Leutwein, C., Heurtier, M. A., Copley, R. R., Edelmann, A., Querfurth, E., Rybin, V., Drewes, G., Raida, M., Bouwmeester, T., Bork, P., Seraphin, B., Kuster, B., Neubauer, G. & Superti-Furga, G. (2002) Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature*, 415(6868), 141-7.
- Gil, J., Bernard, D., Martinez, D. & Beach, D. (2004) Polycomb cbx7 has a unifying role in cellular lifespan. *Nat Cell Biol*, 6(1), 67-72.
- Gineitis, A. A., Zalenskaya, I. A., Yau, P. M., Bradbury, E. M. & Zalensky, A. O. (2000) Human sperm telomere-binding complex involves histone h2b and secures telomere membrane attachment. *J Cell Biol*, 151(7), 1591-8.
- Ginsburg, E., Salomon, D., Sreevalsan, T. & Freese, E. (1973) Growth inhibition and morphological changes caused by lipophilic acids in mammalian cells. *Proc Natl Acad Sci U S A*, 70(8), 2457-61.
- Giot, L., Bader, J. S., Brouwer, C., Chaudhuri, A., Kuang, B., Li, Y., Hao, Y. L., Ooi, C. E., Godwin, B., Vitols, E., Vijayadamar, G., Pochart, P., Machineni, H., Welsh, M., Kong, Y., Zerhusen, B., Malcolm, R., Varrone, Z., Collis, A., Minto, M., Burgess, S., Mcdaniel, L., Stimpson, E., Spriggs, F., Williams, J., Neurath, K., Ioime, N., Agee, M., Voss, E., Furtak, K., Renzulli, R., Aanensen, N., Carrolla, S., Bickelhaupt, E., Lazovatsky, Y., Dasilva, A., Zhong, J., Stanyon, C. A., Finley, R. L., Jr., White, K. P., Braverman, M., Jarvie, T., Gold, S., Leach, M., Knight, J., Shimkets, R. A., Mckenna, M. P., Chant, J. & Rothberg, J. M. (2003) A protein interaction map of drosophila melanogaster. *Science*, 302(5651), 1727-36.
- Gloeckner, C. J., Boldt, K., Schumacher, A., Roepman, R. & Ueffing, M. (2007) A novel tandem affinity purification strategy for the efficient isolation and characterisation of native protein complexes. *Proteomics*, 7(23),

4228-34.

- Glover, C. V., 3rd (1998) On the physiological role of casein kinase ii in *saccharomyces cerevisiae*. *Prog Nucleic Acid Res Mol Biol*, 59(95-133).
- Golderer, G. & Grobner, P. (1991) Adp-ribosylation of core histones and their acetylated subspecies. *Biochem J*, 277 (Pt 3)(607-10).
- Goll, M. G., Kirpekar, F., Maggert, K. A., Yoder, J. A., Hsieh, C. L., Zhang, X., Golig, K. G., Jacobsen, S. E. & Bestor, T. H. (2006) Methylation of trnaasp by the DNA methyltransferase homolog dnmt2. *Science*, 311(5759), 395-8.
- Goto, H., Tomono, Y., Ajiro, K., Kosako, H., Fujita, M., Sakurai, M., Okawa, K., Iwamatsu, A., Okigaki, T., Takahashi, T. & Inagaki, M. (1999) Identification of a novel phosphorylation site on histone h3 coupled with mitotic chromosome condensation. *J Biol Chem*, 274(36), 25543-9.
- Goto, H., Yasui, Y., Nigg, E. A. & Inagaki, M. (2002) Aurora-b phosphorylates histone h3 at serine28 with regard to the mitotic chromosome condensation. *Genes Cells*, 7(1), 11-7.
- Greger, V., Passarge, E., Hopping, W., Messmer, E. & Horsthemke, B. (1989) Epigenetic changes may contribute to the formation and spontaneous regression of retinoblastoma. *Hum Genet*, 83(2), 155-8.
- Gregoret, I. V., Lee, Y. M. & Goodson, H. V. (2004) Molecular evolution of the histone deacetylase family: Functional implications of phylogenetic analysis. *J Mol Biol*, 338(1), 17-31.
- Greil, F., Van Der Kraan, I., Delrow, J., Smothers, J. F., De Wit, E., Bussemaker, H. J., Van Driel, R., Henikoff, S. & Van Steensel, B. (2003) Distinct hp1 and su(var)3-9 complexes bind to sets of developmentally coexpressed genes depending on chromosomal location. *Genes Dev*, 17(22), 2825-38.
- Grewal, S. I. & Moazed, D. (2003) Heterochromatin and epigenetic control of gene expression. *Science*, 301(5634), 798-802.
- Grimaud, C., Bantignies, F., Pal-Bhadra, M., Ghana, P., Bhadra, U. & Cavalli, G. (2006) Rnai components are required for nuclear clustering of polycomb group response elements. *Cell*, 124(5), 957-71.
- Grishok, A., Sinskey, J. L. & Sharp, P. A. (2005) Transcriptional silencing of a transgene by rnai in the soma of *c. Elegans*. *Genes Dev*, 19(6), 683-96.
- Guenatri, M., Bailly, D., Maison, C. & Almouzni, G. (2004) Mouse centric and pericentric satellite repeats form distinct functional heterochromatin. *J Cell Biol*, 166(4), 493-505.
- Gully, D., Moinier, D., Loiseau, L. & Bouveret, E. (2003) New partners of acyl carrier protein detected in *escherichia coli* by tandem affinity purification. *FEBS Lett*, 548(1-3), 90-6.
- Guo, W. J., Zeng, M. S., Yadav, A., Song, L. B., Guo, B. H., Band, V. & Dimri, G. P. (2007) Mel-18 acts as a tumor suppressor by repressing bmi-1 expression and down-regulating akt activity in breast cancer cells. *Cancer Res*, 67(11), 5083-9.
- Gurley, L. R., D'anna, J. A., Barham, S. S., Deaven, L. L. & Tobey, R. A. (1978) Histone phosphorylation and chromatin structure during mitosis in chinese hamster cells. *Eur J Biochem*, 84(1), 1-15.
- Hake, S. B. & Allis, C. D. (2006) Histone h3 variants and their potential role in indexing mammalian genomes: The "H3 barcode hypothesis". *Proc Natl Acad Sci U S A*, 103(17), 6428-35.
- Hake, S. B., Garcia, B. A., Duncan, E. M., Kauer, M., Dellaire, G., Shabanowitz, J., Bazett-Jones, D. P., Allis, C. D. & Hunt, D. F. (2006) Expression patterns and post-translational modifications associated with mammalian histone h3 variants. *J Biol Chem*, 281(1), 559-68.
- Hake, S. B., Garcia, B. A., Kauer, M., Baker, S. P., Shabanowitz, J., Hunt, D. F. & Allis, C. D. (2005) Serine 31 phosphorylation of histone variant h3.3 is specific to regions bordering centromeres in metaphase chromosomes. *Proc Natl Acad Sci U S A*, 102(18), 6344-9.
- Hall, I. M., Shankaranarayana, G. D., Noma, K., Ayoub, N., Cohen, A. & Grewal, S. I. (2002) Establishment and maintenance of a heterochromatin domain. *Science*, 297(5590), 2232-7.
- Han, Z., Guo, L., Wang, H., Shen, Y., Deng, X. W. & Chai, J. (2006) Structural basis for the specific recognition of methylated histone h3 lysine 4 by the wd-40 protein wdr5. *Mol Cell*, 22(1), 137-44.
- Hansen, K. H., Bracken, A. P., Pasini, D., Dietrich, N., Gehani, S. S., Monrad, A., Rappsilber, J., Lerdrup, M. & Helin, K. (2008) A model for transmission of the h3k27me3 epigenetic mark. *Nat Cell Biol*, 10(11), 1291-300.
- Haupt, Y., Alexander, W. S., Barri, G., Klinken, S. P. & Adams, J. M. (1991) Novel zinc finger gene implicated as myc collaborator by retrovirally accelerated lymphomagenesis in e mu-myc transgenic mice. *Cell*, 65(5), 753-63.
- Hayakawa, T., Haraguchi, T., Masumoto, H. & Hiraoka, Y. (2003) Cell cycle behavior of human hp1 subtypes: Distinct molecular domains of hp1 are required for their centromeric localization during interphase and metaphase. *J Cell Sci*, 116(Pt 16), 3327-38.
- Hearn, M. G., Hedrick, A., Grigliatti, T. A. & Wakimoto, B. T. (1991) The effect of modifiers of position-effect variegation on the variegation of heterochromatic genes of *drosophila melanogaster*. *Genetics*, 128(4), 785-97.
- Hecker, C. M., Rabiller, M., Haglund, K., Bayer, P. & Dikic, I. (2006) Specification of sumo1- and sumo2-interacting motifs. *J Biol Chem*, 281(23), 16117-27.
- Hernandez-Munoz, I., Lund, A. H., Van Der Stoop, P., Boutsma, E., Muijers, I., Verhoeven, E., Nusinow, D. A., Panning, B., Marahrens, Y. & Van Lohuizen, M. (2005) Stable x chromosome inactivation involves the prc1 polycomb complex and requires histone macroh2a1 and the cullin3/spo ubiquitin e3 ligase. *Proc Natl Acad Sci U S A*, 102(21), 7635-40.
- Hiragami, K. & Festenstein, R. (2005) Heterochromatin protein 1: A pervasive controlling influence. *Cell Mol Life Sci*, 62(23), 2711-26.

- Hirota, T., Lipp, J. J., Toh, B. H. & Peters, J. M. (2005) Histone h3 serine 10 phosphorylation by aurora b causes hp1 dissociation from heterochromatin. *Nature*, 438(7071), 1176-80.
- Hook, S. S., Orian, A., Cowley, S. M. & Eisenman, R. N. (2002) Histone deacetylase 6 binds polyubiquitin through its zinc finger (paz domain) and copurifies with deubiquitinating enzymes. *Proc Natl Acad Sci U S A*, 99(21), 13425-30.
- Hsu, H. L., Gilley, D., Blackburn, E. H. & Chen, D. J. (1999) Ku is associated with the telomere in mammals. *Proc Natl Acad Sci U S A*, 96(22), 12454-8.
- Huisinga, K. L. & Elgin, S. C. (2009) Small rna-directed heterochromatin formation in the context of development: What flies might learn from fission yeast. *Biochim Biophys Acta*, 1789(1), 3-16.
- Huletsky, A., De Murcia, G., Muller, S., Hengartner, M., Menard, L., Lamarre, D. & Poirier, G. G. (1989) The effect of poly(adp-ribosylation) on native and h1-depleted chromatin. A role of poly(adp-ribosylation) on core nucleosome structure. *J Biol Chem*, 264(15), 8878-86.
- Hurley, J. H., Lee, S. & Prag, G. (2006) Ubiquitin-binding domains. *Biochem J*, 399(3), 361-72.
- Huyen, Y., Zgheib, O., Ditullio, R. A., Jr., Gorgoulis, V. G., Zacharatos, P., Petty, T. J., Sheston, E. A., Mellert, H. S., Stavridi, E. S. & Halazonetis, T. D. (2004) Methylated lysine 79 of histone h3 targets 53bp1 to DNA double-strand breaks. *Nature*, 432(7015), 406-11.
- Ikeda, R., Kokubu, C., Yusa, K., Keng, V. W., Horie, K. & Takeda, J. (2007) Sleeping beauty transposase has an affinity for heterochromatin conformation. *Mol Cell Biol*, 27(5), 1665-76.
- Isono, K., Fujimura, Y., Shinga, J., Yamaki, M., J, O. W., Takihara, Y., Murahashi, Y., Takada, Y., Mizutani-Koseki, Y. & Koseki, H. (2005) Mammalian polyhomeotic homologues phc2 and phc1 act in synergy to mediate polycomb repression of hox genes. *Mol Cell Biol*, 25(15), 6694-706.
- Jacobs, S. A. & Khorasanizadeh, S. (2002) Structure of hp1 chromodomain bound to a lysine 9-methylated histone h3 tail. *Science*, 295(5562), 2080-3.
- Jacobson, R. H., Ladurner, A. G., King, D. S. & Tjian, R. (2000) Structure and function of a human tafii250 double bromodomain module. *Science*, 288(5470), 1422-5.
- James, T. C., Eissenberg, J. C., Craig, C., Dietrich, V., Hobson, A. & Elgin, S. C. (1989) Distribution patterns of hp1, a heterochromatin-associated nonhistone chromosomal protein of drosophila. *Eur J Cell Biol*, 50(1), 170-80.
- James, T. C. & Elgin, S. C. (1986) Identification of a nonhistone chromosomal protein associated with heterochromatin in drosophila melanogaster and its gene. *Mol Cell Biol*, 6(11), 3862-72.
- Janicki, S. M., Tsukamoto, T., Salghetti, S. E., Tansey, W. P., Sachidanandam, R., Prasanth, K. V., Ried, T., Shav-Tal, Y., Bertrand, E., Singer, R. H. & Spector, D. L. (2004) From silencing to gene expression: Real-time analysis in single cells. *Cell*, 116(5), 683-98.
- Jiao, K., Zhou, Y. & Hogan, B. L. (2002) Identification of mznf8, a mouse kruppel-like transcriptional repressor, as a novel nuclear interaction partner of smad1. *Mol Cell Biol*, 22(21), 7633-44.
- Jin, J., Smith, F. D., Stark, C., Wells, C. D., Fawcett, J. P., Kulkarni, S., Metalnikov, P., O'donnell, P., Taylor, P., Taylor, L., Zougman, A., Woodgett, J. R., Langeberg, L. K., Scott, J. D. & Pawson, T. (2004) Proteomic, functional, and domain-based analysis of in vivo 14-3-3 binding proteins involved in cytoskeletal regulation and cellular organization. *Curr Biol*, 14(16), 1436-50.
- Johansson, A. M., Stenberg, P., Pettersson, F. & Larsson, J. (2007) Pof and hp1 bind expressed exons, suggesting a balancing mechanism for gene regulation. *PLoS Genet*, 3(11), e209.
- Jones, B., Su, H., Bhat, A., Lei, H., Bajko, J., Hevi, S., Baltus, G. A., Kadam, S., Zhai, H., Valdez, R., Gonzalo, S., Zhang, Y., Li, E. & Chen, T. (2008) The histone h3k79 methyltransferase dot1l is essential for mammalian development and heterochromatin structure. *PLoS Genet*, 4(9), e1000190.
- Jones, D. H., Ley, S. & Aitken, A. (1995) Isoforms of 14-3-3 protein can form homo- and heterodimers in vivo and in vitro: Implications for function as adapter proteins. *FEBS Lett*, 368(1), 55-8.
- Jones, D. O., Cowell, I. G. & Singh, P. B. (2000) Mammalian chromodomain proteins: Their role in genome organization and expression. *Bioessays*, 22(2), 124-37.
- Joo, H. Y., Zhai, L., Yang, C., Nie, S., Erdjument-Bromage, H., Tempst, P., Chang, C. & Wang, H. (2007) Regulation of cell cycle progression and gene expression by h2a deubiquitination. *Nature*, 449(7165), 1068-72.
- Joung, J. K., Ramm, E. I. & Pabo, C. O. (2000) A bacterial two-hybrid selection system for studying protein-DNA and protein-protein interactions. *Proc Natl Acad Sci U S A*, 97(13), 7382-7.
- Junakovic, N., Terrinoni, A., Di Franco, C., Vieira, C. & Loevenbruck, C. (1998) Accumulation of transposable elements in the heterochromatin and on the y chromosome of drosophila simulans and drosophila melanogaster. *J Mol Evol*, 46(6), 661-8.
- Kagey, M. H., Melhuish, T. A., Powers, S. E. & Wotton, D. (2005) Multiple activities contribute to pc2 e3 function. *EMBO J*, 24(1), 108-19.
- Kagey, M. H., Melhuish, T. A. & Wotton, D. (2003) The polycomb protein pc2 is a sumo e3. *Cell*, 113(1), 127-37.
- Kalakonda, N., Fischle, W., Boccuni, P., Gurvich, N., Hoya-Arias, R., Zhao, X., Miyata, Y., Macgrogan, D., Zhang, J., Sims, J. K., Rice, J. C. & Nimer, S. D. (2008) Histone h4 lysine 20 monomethylation promotes transcriptional repression by l3mbtl1. *Oncogene*, 27(31), 4293-304.
- Kallin, E. M., Cao, R., Jothi, R., Xia, K., Cui, K., Zhao, K. & Zhang, Y. (2009) Genome-wide uh2a localization analysis highlights bmi1-dependent deposition of the mark at repressed genes. *PLoS Genet*, 5(6), e1000506.
- Kanellopoulou, C., Muljo, S. A., Kung, A. L., Ganesan, S., Drapkin, R., Jenuwein, T., Livingston, D. M. & Rajewsky, K. (2005) Dicer-deficient mouse embryonic stem cells are defective in differentiation and

- centromeric silencing. *Genes Dev*, 19(4), 489-501.
- Kangaspeska, S., Stride, B., Metivier, R., Polycarpou-Schwarz, M., Ibberson, D., Carmouche, R. P., Benes, V., Gannon, F. & Reid, G. (2008) Transient cyclical methylation of promoter DNA. *Nature*, 452(7183), 112-5.
- Kanno, M., Hasegawa, M., Ishida, A., Isono, K. & Taniguchi, M. (1995) Mel-18, a polycomb group-related mammalian gene, encodes a transcriptional negative regulator with tumor suppressive activity. *EMBO J*, 14(22), 5672-8.
- Kato, Y., Kaneda, M., Hata, K., Kumaki, K., Hisano, M., Kohara, Y., Okano, M., Li, E., Nozaki, M. & Sasaki, H. (2007) Role of the dnmt3 family in de novo methylation of imprinted and repetitive sequences during male germ cell development in the mouse. *Hum Mol Genet*, 16(19), 2272-80.
- Kaufman, P. D., Kobayashi, R., Kessler, N. & Stillman, B. (1995) The p150 and p60 subunits of chromatin assembly factor I: A molecular link between newly synthesized histones and DNA replication. *Cell*, 81(7), 1105-14.
- Ketel, C. S., Andersen, E. F., Vargas, M. L., Suh, J., Strome, S. & Simon, J. A. (2005) Subunit contributions to histone methyltransferase activities of fly and worm polycomb group complexes. *Mol Cell Biol*, 25(16), 6857-68.
- Khetchoumian, K., Teletin, M., Marck, M., Lerouge, T., Cervino, M., Oulad-Abdelghani, M., Chambon, P. & Losson, R. (2004) TIF1 δ , a Novel HP1-interacting Member of the Transcriptional Intermediary Factor 1 (TIF1) Family Expressed by Elongating Spermatids. *J Biol Chem*, 279(46), 48329-41.
- Kim, C. A. & Bowie, J. U. (2003) Sam domains: Uniform structure, diversity of function. *Trends Biochem Sci*, 28(12), 625-8.
- Kim, C. A., Gingery, M., Pilpa, R. M. & Bowie, J. U. (2002) The sam domain of polyhomeotic forms a helical polymer. *Nat Struct Biol*, 9(6), 453-7.
- Kim, C. A., Sawaya, M. R., Cascio, D., Kim, W. & Bowie, J. U. (2005) Structural organization of a sex-comb-on-midleg/polyhomeotic copolymer. *J Biol Chem*, 280(30), 27769-75.
- Kim, J., Daniel, J., Espejo, A., Lake, A., Krishna, M., Xia, L., Zhang, Y. & Bedford, M. T. (2006) Tudor, mbt and chromo domains gauge the degree of lysine methylation. *EMBO Rep*, 7(4), 397-403.
- Kim, J., Hake, S. B. & Roeder, R. G. (2005) The human homolog of yeast bre1 functions as a transcriptional coactivator through direct activator interactions. *Mol Cell*, 20(5), 759-70.
- Kim, J., Sif, S., Jones, B., Jackson, A., Koipally, J., Heller, E., Winandy, S., Viel, A., Sawyer, A., Ikeda, T., Kingston, R. & Georgopoulos, K. (1999) Ikaros DNA-binding proteins direct formation of chromatin remodeling complexes in lymphocytes. *Immunity*, 10(3), 345-55.
- Kimmins, S. & Sassone-Corsi, P. (2005) Chromatin remodelling and epigenetic features of germ cells. *Nature*, 434(7033), 583-9.
- Kinney, C. M., Chandrasekharan, U. M., Mavrakis, L. & Dicorleto, P. E. (2008) Vegf and thrombin induce mkp-1 through distinct signaling pathways: Role for mkp-1 in endothelial cell migration. *Am J Physiol Cell Physiol*, 294(1), C241-50.
- Kito, K., Kawaguchi, N., Okada, S. & Ito, T. (2008) Discrimination between stable and dynamic components of protein complexes by means of quantitative proteomics. *Proteomics*, 8(12), 2366-70.
- Kleff, S., Andrulis, E. D., Anderson, C. W. & Sternglanz, R. (1995) Identification of a gene encoding a yeast histone h4 acetyltransferase. *J Biol Chem*, 270(42), 24674-7.
- Klenova, E. & Ohlsson, R. (2005) Poly(adp-ribosylation) and epigenetics. Is ctf part of the plot? *Cell Cycle*, 4(1), 96-101.
- Klose, R. J., Yamane, K., Bae, Y., Zhang, D., Erdjument-Bromage, H., Tempst, P., Wong, J. & Zhang, Y. (2006) The transcriptional repressor jhdm3a demethylates trimethyl histone h3 lysine 9 and lysine 36. *Nature*, 442(7100), 312-6.
- Klymenko, T. & Muller, J. (2004) The histone methyltransferases trithorax and ash1 prevent transcriptional silencing by polycomb group proteins. *EMBO Rep*, 5(4), 373-7.
- Klymenko, T., Papp, B., Fischle, W., Kocher, T., Schelder, M., Fritsch, C., Wild, B., Wilm, M. & Muller, J. (2006) A polycomb group protein complex with sequence-specific DNA-binding and selective methyl-lysine-binding activities. *Genes Dev*, 20(9), 1110-22.
- Koering, C. E., Pollice, A., Zibella, M. P., Bauwens, S., Puisieux, A., Brunori, M., Brun, C., Martins, L., Sabatier, L., Pulitzer, J. F. & Gilson, E. (2002) Human telomeric position effect is determined by chromosomal context and telomeric chromatin integrity. *EMBO Rep*, 3(11), 1055-61.
- Koike, N., Maita, H., Taira, T., Ariga, H. & Iguchi-Ariga, S. M. (2000) Identification of heterochromatin protein 1 (hp1) as a phosphorylation target by pim-1 kinase and the effect of phosphorylation on the transcriptional repression function of hp1(1). *FEBS Lett*, 467(1), 17-21.
- Koonin, E. V., Zhou, S. & Lucchesi, J. C. (1995) The chromo superfamily: New members, duplication of the chromo domain and possible role in delivering transcription regulators to chromatin. *Nucleic Acids Res*, 23(21), 4229-33.
- Kourmouli, N., Theodoropoulos, P. A., Dialynas, G., Bakou, A., Politou, A. S., Cowell, I. G., Singh, P. B. & Georgatos, S. D. (2000) Dynamic associations of heterochromatin protein 1 with the nuclear envelope. *EMBO J*, 19(23), 6558-68.
- Kouzarides, T. (2002) Histone methylation in transcriptional control. *Curr Opin Genet Dev*, 12(2), 198-209.
- Kraus, W. L. & Lis, J. T. (2003) Parp goes transcription. *Cell*, 113(6), 677-83.
- Krude, T. (1995) Chromatin assembly factor 1 (caf-1) colocalizes with replication foci in hela cell nuclei. *Exp Cell Res*, 220(2), 304-11.
- Ku, M., Koche, R. P., Rheinbay, E., Mendenhall, E. M., Endoh, M., Mikkelsen, T. S., Presser, A., Nusbaum, C.,

- Xie, X., Chi, A. S., Adli, M., Kasif, S., Ptaszek, L. M., Cowan, C. A., Lander, E. S., Koseki, H. & Bernstein, B. E. (2008) Genomewide analysis of prc1 and prc2 occupancy identifies two classes of bivalent domains. *PLoS Genet*, 4(10), e1000242.
- Kushnir, M., Dresner, E., Mandel, S. & Gozes, I. (2008) Silencing of the adnp-family member, adnp2, results in changes in cellular viability under oxidative stress. *J Neurochem*, 105(2), 537-45.
- Kustatscher, G., Hothorn, M., Pugieux, C., Scheffzek, K. & Ladurner, A. G. (2005) Splicing regulates nad metabolite binding to histone macroh2a. *Nat Struct Mol Biol*, 12(7), 624-5.
- Kuzmichev, A., Jenuwein, T., Tempst, P. & Reinberg, D. (2004) Different ezh2-containing complexes target methylation of histone h1 or nucleosomal histone h3. *Mol Cell*, 14(2), 183-93.
- Kuzmichev, A., Margueron, R., Vaquero, A., Preissner, T. S., Scher, M., Kirmizis, A., Ouyang, X., Brockdorff, N., Abate-Shen, C., Farnham, P. & Reinberg, D. (2005) Composition and histone substrates of polycomb repressive group complexes change during cellular differentiation. *Proc Natl Acad Sci U S A*, 102(6), 1859-64.
- Kuzmichev, A., Nishioka, K., Erdjument-Bromage, H., Tempst, P. & Reinberg, D. (2002) Histone methyltransferase activity associated with a human multiprotein complex containing the enhancer of zeste protein. *Genes Dev*, 16(22), 2893-905.
- Kyriakakis, P., Tipping, M., Abed, L. & Veraksa, A. (2008) Tandem affinity purification in drosophila: The advantages of the gs-tap system. *Fly (Austin)*, 2(4).
- Lachner, M., O'carroll, D., Rea, S., Mechtler, K. & Jenuwein, T. (2001) Methylation of histone h3 lysine 9 creates a binding site for hp1 proteins. *Nature*, 410(6824), 116-20.
- Lan, F., Bayliss, P. E., Rinn, J. L., Whetstone, J. R., Wang, J. K., Chen, S., Iwase, S., Alpatov, R., Issaeva, I., Canaani, E., Roberts, T. M., Chang, H. Y. & Shi, Y. (2007) A histone h3 lysine 27 demethylase regulates animal posterior development. *Nature*, 449(7163), 689-94.
- Lan, F., Collins, R. E., De Cegli, R., Alpatov, R., Horton, J. R., Shi, X., Gozani, O., Cheng, X. & Shi, Y. (2007) Recognition of unmethylated histone h3 lysine 4 links bhc80 to lsd1-mediated gene repression. *Nature*, 448(7154), 718-22.
- Lania, L., Donti, E., Pannuti, A., Pascucci, A., Pengue, G., Feliciello, I., La Mantia, G., Lanfrancone, L. & Pelicci, P. G. (1990) Cdna isolation, expression analysis, and chromosomal localization of two human zinc finger genes. *Genomics*, 6(2), 333-40.
- Le Douarin, B., Nielsen, A. L., Garnier, J. M., Ichinose, H., Jeanmougin, F., Losson, R. & Chambon, P. (1996) A possible involvement of tif1 alpha and tif1 beta in the epigenetic control of transcription by nuclear receptors. *EMBO J*, 15(23), 6701-15.
- Lechner, M. S., Begg, G. E., Speicher, D. W. & Rauscher, F. J., 3rd (2000) Molecular determinants for targeting heterochromatin protein 1-mediated gene silencing: Direct chromoshadow domain-kap-1 corepressor interaction is essential. *Mol Cell Biol*, 20(17), 6449-65.
- Lechner, M. S., Schultz, D. C., Negorev, D., Maul, G. G. & Rauscher, F. J., 3rd (2005) The mammalian heterochromatin protein 1 binds diverse nuclear proteins through a common motif that targets the chromoshadow domain. *Biochem Biophys Res Commun*, 331(4), 929-37.
- Lee, D. Y., Teyssier, C., Strahl, B. D. & Stallcup, M. R. (2005) Role of protein methylation in regulation of transcription. *Endocr Rev*, 26(2), 147-70.
- Lee, M. G., Villa, R., Trojer, P., Norman, J., Yan, K. P., Reinberg, D., Di Croce, L. & Shiekhattar, R. (2007) Demethylation of h3k27 regulates polycomb recruitment and h2a ubiquitination. *Science*, 318(5849), 447-50.
- Lee, S. J., Matsuura, Y., Liu, S. M. & Stewart, M. (2005) Structural basis for nuclear import complex dissociation by rangtp. *Nature*, 435(7042), 693-6.
- Lehnertz, B., Ueda, Y., Derijck, A. A., Braunschweig, U., Perez-Burgos, L., Kubicek, S., Chen, T., Li, E., Jenuwein, T. & Peters, A. H. (2003) Suv39h-mediated histone h3 lysine 9 methylation directs DNA methylation to major satellite repeats at pericentric heterochromatin. *Curr Biol*, 13(14), 1192-200.
- Leroy, G., Weston, J. T., Zee, B. M., Young, N. L., Plazas-Mayorca, M. D. & Garcia, B. A. (2009) Heterochromatin protein 1 is extensively decorated with histone code-like post-translational modifications. *Mol Cell Proteomics*.
- Lessard, J. & Sauvageau, G. (2003) Polycomb group genes as epigenetic regulators of normal and leukemic hemopoiesis. *Exp Hematol*, 31(7), 567-85.
- Levine, S. S., King, I. F. & Kingston, R. E. (2004) Division of labor in polycomb group repression. *Trends Biochem Sci*, 29(9), 478-85.
- Levine, S. S., Weiss, A., Erdjument-Bromage, H., Shao, Z., Tempst, P. & Kingston, R. E. (2002) The core of the polycomb repressive complex is compositionally and functionally conserved in flies and humans. *Mol Cell Biol*, 22(17), 6070-8.
- Lewis, E. B. (1978) A gene complex controlling segmentation in drosophila. *Nature*, 276(5688), 565-70.
- Li, H., Ilin, S., Wang, W., Duncan, E. M., Wysocka, J., Allis, C. D. & Patel, D. J. (2006) Molecular basis for site-specific read-out of histone h3k4me3 by the bptf phd finger of nurf. *Nature*, 442(7098), 91-5.
- Li, S., Armstrong, C. M., Bertin, N., Ge, H., Milstein, S., Boxem, M., Vidalain, P. O., Han, J. D., Chesneau, A., Hao, T., Goldberg, D. S., Li, N., Martinez, M., Rual, J. F., Lamesch, P., Xu, L., Tewari, M., Wong, S. L., Zhang, L. V., Berriz, G. F., Jacotot, L., Vaglio, P., Reboul, J., Hirozane-Kishikawa, T., Li, Q., Gabel, H. W., Elewa, A., Baumgartner, B., Rose, D. J., Yu, H., Bosak, S., Sequerra, R., Fraser, A., Mango, S. E., Saxton, W. M., Strome, S., Van Den Heuvel, S., Piano, F., Vandenhaute, J., Sardet, C., Gerstein, M., Doucette-Stamm, L., Gunsalus, K. C., Harper, J. W., Cusick, M. E., Roth, F. P., Hill, D. E. & Vidal, M.

- (2004) A map of the interactome network of the metazoan *C. elegans*. *Science*, 303(5657), 540-3.
- Li, Y., Kirschmann, D. A. & Wallrath, L. L. (2002) Does heterochromatin protein 1 always follow code? *Proc Natl Acad Sci U S A*, 99 Suppl 4(16462-9).
- Li, Z., Cao, R., Wang, M., Myers, M. P., Zhang, Y. & Xu, R. M. (2006) Structure of a bmi-1-ring1b polycomb group ubiquitin ligase complex. *J Biol Chem*, 281(29), 20643-9.
- Liang, F., Holt, I., Perteau, G., Karamycheva, S., Salzberg, S. L. & Quackenbush, J. (2000) Gene index analysis of the human genome estimates approximately 120,000 genes. *Nat Genet*, 25(2), 239-40.
- Litchfield, D. W. (2003) Protein kinase ck2: Structure, regulation and role in cellular decisions of life and death. *Biochem J*, 369(Pt 1), 1-15.
- Littlewood, T. D., Hancock, D. C., Danielian, P. S., Parker, M. G. & Evan, G. I. (1995) A modified oestrogen receptor ligand-binding domain as an improved switch for the regulation of heterologous proteins. *Nucleic Acids Res*, 23(10), 1686-90.
- Liu, C. L., Kaplan, T., Kim, M., Buratowski, S., Schreiber, S. L., Friedman, N. & Rando, O. J. (2005) Single-nucleosome mapping of histone modifications in *S. cerevisiae*. *PLoS Biol*, 3(10), e328.
- Liu, H., Schmidt-Supprian, M., Shi, Y., Hobeika, E., Barteneva, N., Jumaa, H., Pelanda, R., Reth, M., Skok, J. & Rajewsky, K. (2007) Yin yang 1 is a critical regulator of B-cell development. *Genes Dev*, 21(10), 1179-89.
- Lo, W. S., Trievel, R. C., Rojas, J. R., Duggan, L., Hsu, J. Y., Allis, C. D., Marmorstein, R. & Berger, S. L. (2000) Phosphorylation of serine 10 in histone H3 is functionally linked in vitro and in vivo to GCN5-mediated acetylation at lysine 14. *Mol Cell*, 5(6), 917-26.
- Lomberk, G. (2004) Hp1 gamma, a regulator of epigenetic gene silencing by DNA methylation, is regulated by changes in nuclear localization in response to phosphorylation. *Pancreas*, 29(4), 356.
- Lomberk, G., Bensi, D., Fernandez-Zapico, M. E. & Urrutia, R. (2006) Evidence for the existence of an hp1-mediated subcode within the histone code. *Nat Cell Biol*, 8(4), 407-15.
- Lomberk, G., Wallrath, L. & Urrutia, R. (2006) The heterochromatin protein 1 family. *Genome Biol*, 7(7), 228.
- Lu, B. Y., Emtage, P. C., Duyf, B. J., Hilliker, A. J. & Eisenberg, J. C. (2000) Heterochromatin protein 1 is required for the normal expression of two heterochromatin genes in *Drosophila*. *Genetics*, 155(2), 699-708.
- Luger, K. (2003) Structure and dynamic behavior of nucleosomes. *Curr Opin Genet Dev*, 13(2), 127-35.
- Luger, K., Mader, A. W., Richmond, R. K., Sargent, D. F. & Richmond, T. J. (1997) Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature*, 389(6648), 251-60.
- Luo, J., Li, M., Tang, Y., Laszkowska, M., Roeder, R. G. & Gu, W. (2004) Acetylation of p53 augments its site-specific DNA binding both in vitro and in vivo. *Proc Natl Acad Sci U S A*, 101(8), 2259-64.
- Lusser, A., Urwin, D. L. & Kadonaga, J. T. (2005) Distinct activities of chd1 and acf in ATP-dependent chromatin assembly. *Nat Struct Mol Biol*, 12(2), 160-6.
- Macdonald, N., Welburn, J. P., Noble, M. E., Nguyen, A., Yaffe, M. B., Clynes, D., Moggs, J. G., Orphanides, G., Thomson, S., Edmunds, J. W., Clayton, A. L., Endicott, J. A. & Mahadevan, L. C. (2005) Molecular basis for the recognition of phosphorylated and phosphoacetylated histone H3 by 14-3-3. *Mol Cell*, 20(2), 199-211.
- Maison, C., Bailly, D., Peters, A. H., Quivy, J. P., Roche, D., Taddei, A., Lachner, M., Jenuwein, T. & Almouzni, G. (2002) Higher-order structure in pericentric heterochromatin involves a distinct pattern of histone modification and an RNA component. *Nat Genet*, 30(3), 329-34.
- Mandel, S. & Gozes, I. (2007) Activity-dependent neuroprotective protein constitutes a novel element in the SWI/SNF chromatin remodeling complex. *J Biol Chem*, 282(47), 34448-56.
- Mandel, S., Rechavi, G. & Gozes, I. (2007) Activity-dependent neuroprotective protein (ADNP) differentially interacts with chromatin to regulate genes essential for embryogenesis. *Dev Biol*, 303(2), 814-24.
- Margueron, R., Li, G., Sarma, K., Blais, A., Zavadil, J., Woodcock, C. L., Dynlacht, B. D. & Reinberg, D. (2008) Ezh1 and Ezh2 maintain repressive chromatin through different mechanisms. *Mol Cell*, 32(4), 503-18.
- Martinez, A. M., Colomb, S., Dejardin, J., Bantignies, F. & Cavalli, G. (2006) Polycomb group-dependent cyclin repression in *Drosophila*. *Genes Dev*, 20(4), 501-13.
- Mateescu, B., England, P., Halgand, F., Yaniv, M. & Muchardt, C. (2004) Tethering of hp1 proteins to chromatin is relieved by phosphoacetylation of histone H3. *EMBO Rep*, 5(5), 490-6.
- Maurer-Stroh, S., Dickens, N. J., Hughes-Davies, L., Kouzarides, T., Eisenhaber, F. & Ponting, C. P. (2003) The tudor domain 'royal family': Tudor, plant agmatin, chromo, PWWP and MBT domains. *Trends Biochem Sci*, 28(2), 69-74.
- Meek, S. E., Lane, W. S. & Piwnicka-Worms, H. (2004) Comprehensive proteomic analysis of interphase and mitotic 14-3-3-binding proteins. *J Biol Chem*, 279(31), 32046-54.
- Merlo, A., Herman, J. G., Mao, L., Lee, D. J., Gabrielson, E., Burger, P. C., Baylin, S. B. & Sidransky, D. (1995) 5' CpG island methylation is associated with transcriptional silencing of the tumour suppressor p16/CDKN2/MTS1 in human cancers. *Nat Med*, 1(7), 686-92.
- Metivier, R., Gallais, R., Tiffocche, C., Le Peron, C., Jurkowska, R. Z., Carmouche, R. P., Ibberson, D., Barath, P., Demay, F., Reid, G., Benes, V., Jeltsch, A., Gannon, F. & Salbert, G. (2008) Cyclical DNA methylation of a transcriptionally active promoter. *Nature*, 452(7183), 45-50.
- Metzger, E., Wissmann, M., Yin, N., Muller, J. M., Schneider, R., Peters, A. H., Gunther, T., Buettner, R. & Schule, R. (2005) LSD1 demethylates repressive histone marks to promote androgen-receptor-dependent transcription. *Nature*, 437(7057), 436-9.
- Mikkelsen, T. S., Ku, M., Jaffe, D. B., Issac, B., Lieberman, E., Giannoukos, G., Alvarez, P., Brockman, W., Kim,

- T. K., Koche, R. P., Lee, W., Mendenhall, E., O'donovan, A., Presser, A., Russ, C., Xie, X., Meissner, A., Wernig, M., Jaenisch, R., Nusbaum, C., Lander, E. S. & Bernstein, B. E. (2007) Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature*, 448(7153), 553-60.
- Min, J., Zhang, Y. & Xu, R. M. (2003) Structural basis for specific binding of polycomb chromodomain to histone h3 methylated at lys 27. *Genes Dev*, 17(15), 1823-8.
- Minc, E., Allory, Y., Worman, H. J., Courvalin, J. C. & Buendia, B. (1999) Localization and phosphorylation of hp1 proteins during the cell cycle in mammalian cells. *Chromosoma*, 108(4), 220-34.
- Minc, E., Courvalin, J. C. & Buendia, B. (2000) Hp1gamma associates with euchromatin and heterochromatin in mammalian nuclei and chromosomes. *Cytogenet Cell Genet*, 90(3-4), 279-84.
- Moazed, D. (2001) Common themes in mechanisms of gene silencing. *Mol Cell*, 8(3), 489-98.
- Molofsky, A. V., Pardal, R., Iwashita, T., Park, I. K., Clarke, M. F. & Morrison, S. J. (2003) Bmi-1 dependence distinguishes neural stem cell self-renewal from progenitor proliferation. *Nature*, 425(6961), 962-7.
- Montgomery, N. D., Yee, D., Chen, A., Kalantry, S., Chamberlain, S. J., Otte, A. P. & Magnuson, T. (2005) The murine polycomb group protein eed is required for global histone h3 lysine-27 methylation. *Curr Biol*, 15(10), 942-7.
- Moosmann, P., Georgiev, O., Le Douarin, B., Bourquin, J. P. & Schaffner, W. (1996) Transcriptional repression by ring finger protein ttf1 beta that interacts with the krab repressor domain of kox1. *Nucleic Acids Res*, 24(24), 4859-67.
- Morita, K., Lo Celso, C., Spencer-Dene, B., Zouboulis, C. C. & Watt, F. M. (2006) Han11 binds mdia1 and controls gli1 transcriptional activity. *J Dermatol Sci*, 44(1), 11-20.
- Motamedi, M. R., Verdel, A., Colmenares, S. U., Gerber, S. A., Gygi, S. P. & Moazed, D. (2004) Two rna complexes, rits and rdrc, physically interact and localize to noncoding centromeric rnas. *Cell*, 119(6), 789-802.
- Muchardt, C., Guilleme, M., Seeler, J. S., Trouche, D., Dejean, A. & Yaniv, M. (2002) Coordinated methyl and rna binding is required for heterochromatin localization of mammalian hp1alpha. *EMBO Rep*, 3(10), 975-81.
- Muchardt, C. & Yaniv, M. (1999) Atp-dependent chromatin remodelling: Swi/snf and co. Are on the job. *J Mol Biol*, 293(2), 187-98.
- Mueller-Dieckmann, C., Kernstock, S., Lisurek, M., Von Kries, J. P., Haag, F., Weiss, M. S. & Koch-Nolte, F. (2006) The structure of human adp-ribosylhydrolase 3 (arh3) provides insights into the reversibility of protein adp-ribosylation. *Proc Natl Acad Sci U S A*, 103(41), 15026-31.
- Mukhopadhyay, D. & Dasso, M. (2007) Modification in reverse: The sumo proteases. *Trends Biochem Sci*, 32(6), 286-95.
- Muller, H. J. (1930) Types of visible variations induced by x-rays in drosophila. *J Genet*, 22), 299-334.
- Muller, J. (1995) Transcriptional silencing by the polycomb protein in drosophila embryos. *EMBO J*, 14(6), 1209-20.
- Muller, J., Hart, C. M., Francis, N. J., Vargas, M. L., Sengupta, A., Wild, B., Miller, E. L., O'connor, M. B., Kingston, R. E. & Simon, J. A. (2002) Histone methyltransferase activity of a drosophila polycomb group repressor complex. *Cell*, 111(2), 197-208.
- Murchison, E. P., Partridge, J. F., Tam, O. H., Cheloufi, S. & Hannon, G. J. (2005) Characterization of dicer-deficient murine embryonic stem cells. *Proc Natl Acad Sci U S A*, 102(34), 12135-40.
- Murzina, N., Verreault, A., Laue, E. & Stillman, B. (1999) Heterochromatin dynamics in mouse cells: Interaction between chromatin assembly factor 1 and hp1 proteins. *Mol Cell*, 4(4), 529-40.
- Nakada, S., Chen, G. I., Gingras, A. C. & Durocher, D. (2008) Pp4 is a gamma h2ax phosphatase required for recovery from the DNA damage checkpoint. *EMBO Rep*, 9(10), 1019-26.
- Nakagawa, T., Kajitani, T., Togo, S., Masuko, N., Ohdan, H., Hishikawa, Y., Koji, T., Matsuyama, T., Ikura, T., Muramatsu, M. & Ito, T. (2008) Deubiquitylation of histone h2a activates transcriptional initiation via trans-histone cross-talk with h3k4 di- and trimethylation. *Genes Dev*, 22(1), 37-49.
- Namgoong, S. Y. & Harshey, R. M. (1998) The same two monomers within a mua tetramer provide the dde domains for the strand cleavage and strand transfer steps of transposition. *EMBO J*, 17(13), 3775-85.
- Nathan, D., Ingvarsdottir, K., Sterner, D. E., Bylebyl, G. R., Dokmanovic, M., Dorsey, J. A., Whelan, K. A., Krsmanovic, M., Lane, W. S., Meluh, P. B., Johnson, E. S. & Berger, S. L. (2006) Histone sumoylation is a negative regulator in saccharomyces cerevisiae and shows dynamic interplay with positive-acting histone modifications. *Genes Dev*, 20(8), 966-76.
- Negre, N., Hennetin, J., Sun, L. V., Lavrov, S., Bellis, M., White, K. P. & Cavalli, G. (2006) Chromosomal distribution of pcg proteins during drosophila development. *PLoS Biol*, 4(6), e170.
- Neira, J. L., Roman-Trufero, M., Contreras, L. M., Prieto, J., Singh, G., Barrera, F. N., Renart, M. L. & Vidal, M. (2009) The transcriptional repressor rybp is a natively unfolded protein which folds upon binding to DNA. *Biochemistry*, 48(6), 1348-60.
- Nekrasov, M., Wild, B. & Muller, J. (2005) Nucleosome binding and histone methyltransferase activity of drosophila prc2. *EMBO Rep*, 6(4), 348-53.
- Nelson, C. J., Santos-Rosa, H. & Kouzarides, T. (2006) Proline isomerization of histone h3 regulates lysine methylation and gene expression. *Cell*, 126(5), 905-16.
- Ng, J., Hart, C. M., Morgan, K. & Simon, J. A. (2000) A drosophila esc-e(z) protein complex is distinct from other polycomb group complexes and contains covalently modified esc. *Mol Cell Biol*, 20(9), 3069-78.
- Nicassio, F., Corrado, N., Vissers, J. H., Areces, L. B., Bergink, S., Martejijn, J. A., Geverts, B., Houtsmuller, A. B., Vermeulen, W., Di Fiore, P. P. & Citterio, E. (2007) Human usp3 is a chromatin modifier required for s phase progression and genome stability. *Curr Biol*, 17(22), 1972-7.

- Nielsen, A. L., Ortiz, J. A., You, J., Oulad-Abdelghani, M., Khechumian, R., Gansmuller, A., Chambon, P. & Losson, R. (1999) Interaction with members of the heterochromatin protein 1 (hp1) family and histone deacetylation are differentially involved in transcriptional silencing by members of the tif1 family. *EMBO J*, 18(22), 6385-95.
- Nielsen, A. L., Oulad-Abdelghani, M., Ortiz, J. A., Remboutsika, E., Chambon, P. & Losson, R. (2001) Heterochromatin formation in mammalian cells: Interaction between histones and hp1 proteins. *Mol Cell*, 7(4), 729-39.
- Nielsen, A. L., Sanchez, C., Ichinose, H., Cervino, M., Lerouge, T., Chambon, P. & Losson, R. (2002) Selective interaction between the chromatin-remodeling factor brg1 and the heterochromatin-associated protein hp1alpha. *EMBO J*, 21(21), 5797-806.
- Nielsen, P. R., Nietlispach, D., Mott, H. R., Callaghan, J., Bannister, A., Kouzarides, T., Murzin, A. G., Murzina, N. V. & Laue, E. D. (2002) Structure of the hp1 chromodomain bound to histone h3 methylated at lysine 9. *Nature*, 416(6876), 103-7.
- Nielsen, S. J., Schneider, R., Bauer, U. M., Bannister, A. J., Morrison, A., O'carroll, D., Firestein, R., Cleary, M., Jenuwein, T., Herrera, R. E. & Kouzarides, T. (2001) Rb targets histone h3 methylation and hp1 to promoters. *Nature*, 412(6846), 561-5.
- Nishizuka, Y., Ueda, K., Honjo, T. & Hayaishi, O. (1968) Enzymic adenosine diphosphate ribosylation of histone and poly adenosine diphosphate ribose synthesis in rat liver nuclei. *J Biol Chem*, 243(13), 3765-7.
- Nissen, R. M., Amsterdam, A. & Hopkins, N. (2006) A zebrafish screen for craniofacial mutants identifies wdr68 as a highly conserved gene required for endothelin-1 expression. *BMC Dev Biol*, 6(28).
- Norwood, L. E., Grade, S. K., Cryderman, D. E., Hines, K. A., Furiasse, N., Toro, R., Li, Y., Dhasarathy, A., Kladd, M. P., Hendrix, M. J., Kirschmann, D. A. & Wallrath, L. L. (2004) Conserved properties of hp1(hsalpha). *Gene*, 336(1), 37-46.
- Nowacki, M., Higgins, B. P., Maquilan, G. M., Swart, E. C., Doak, T. G. & Landweber, L. F. (2009) A functional role for transposases in a large eukaryotic genome. *Science*, 324(5929), 935-8.
- Nowak, S. J., Pai, C. Y. & Corces, V. G. (2003) Protein phosphatase 2a activity affects histone h3 phosphorylation and transcription in drosophila melanogaster. *Mol Cell Biol*, 23(17), 6129-38.
- O'carroll, D., Erhardt, S., Pagani, M., Barton, S. C., Surani, M. A. & Jenuwein, T. (2001) The polycomb-group gene ezh2 is required for early mouse development. *Mol Cell Biol*, 21(13), 4330-6.
- Ogawa, H., Ishiguro, K., Gaubatz, S., Livingston, D. M. & Nakatani, Y. (2002) A complex with chromatin modifiers that occupies e2f- and myc-responsive genes in g0 cells. *Science*, 296(5570), 1132-6.
- Okada, Y., Feng, Q., Lin, Y., Jiang, Q., Li, Y., Coffield, V. M., Su, L., Xu, G. & Zhang, Y. (2005) Hdot1l links histone methylation to leukemogenesis. *Cell*, 121(2), 167-78.
- Okano, M., Bell, D. W., Haber, D. A. & Li, E. (1999) DNA methyltransferases dnmt3a and dnmt3b are essential for de novo methylation and mammalian development. *Cell*, 99(3), 247-57.
- Oktaba, K., Gutierrez, L., Gagneur, J., Girardot, C., Sengupta, A. K., Furlong, E. E. & Muller, J. (2008) Dynamic regulation by polycomb group protein complexes controls pattern formation and the cell cycle in drosophila. *Dev Cell*, 15(6), 877-89.
- Olsen, J. V., Blagoev, B., Gnadt, F., Macek, B., Kumar, C., Mortensen, P. & Mann, M. (2006) Global, in vivo, and site-specific phosphorylation dynamics in signaling networks. *Cell*, 127(3), 635-48.
- Olsten, M. E. & Litchfield, D. W. (2004) Order or chaos? An evaluation of the regulation of protein kinase ck2. *Biochem Cell Biol*, 82(6), 681-93.
- Ooi, S. K., Qiu, C., Bernstein, E., Li, K., Jia, D., Yang, Z., Erdjument-Bromage, H., Tempst, P., Lin, S. P., Allis, C. D., Cheng, X. & Bestor, T. H. (2007) Dnmt3l connects unmethylated lysine 4 of histone h3 to de novo methylation of DNA. *Nature*, 448(7154), 714-7.
- Org, T., Chignola, F., Hetenyi, C., Gaetani, M., Rebane, A., Liiv, I., Maran, U., Mollica, L., Bottomley, M. J., Musco, G. & Peterson, P. (2008) The autoimmune regulator phd finger binds to non-methylated histone h3k4 to activate gene expression. *EMBO Rep*, 9(4), 370-6.
- Osley, M. A. (2006) Regulation of histone h2a and h2b ubiquitylation. *Brief Funct Genomic Proteomic*, 5(3), 179-89.
- Otte, A. P. & Kwaks, T. H. (2003) Gene repression by polycomb group protein complexes: A distinct complex for every occasion? *Curr Opin Genet Dev*, 13(5), 448-54.
- Owen, D. J., Ornaghi, P., Yang, J. C., Lowe, N., Evans, P. R., Ballario, P., Neuhaus, D., Filetici, P. & Travers, A. A. (2000) The structural basis for the recognition of acetylated histone h4 by the bromodomain of histone acetyltransferase gcn5p. *EMBO J*, 19(22), 6141-9.
- Ozdog, H., Teschendorff, A. E., Ahmed, A. A., Hyland, S. J., Blenkinsop, C., Bobrow, L., Veerakumarasivam, A., Burtt, G., Subkhankulova, T., Arends, M. J., Collins, V. P., Bowtell, D., Kouzarides, T., Brenton, J. D. & Caldas, C. (2006) Differential expression of selected histone modifier genes in human solid cancers. *BMC Genomics*, 7(90).
- Paine, P. L., Moore, L. C. & Horowitz, S. B. (1975) Nuclear envelope permeability. *Nature*, 254(5496), 109-14.
- Pak, D. T., Pflumm, M., Chesnokov, I., Huang, D. W., Kellum, R., Marr, J., Romanowski, P. & Botchan, M. R. (1997) Association of the origin recognition complex with heterochromatin and hp1 in higher eukaryotes. *Cell*, 91(3), 311-23.
- Pal, S., Vishwanath, S. N., Erdjument-Bromage, H., Tempst, P. & Sif, S. (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.
- Pal-Bhadra, M., Bhadra, U. & Birchler, J. A. (1997) Cosuppression in drosophila: Gene silencing of alcohol

- dehydrogenase by white-adh transgenes is polycomb dependent. *Cell*, 90(3), 479-90.
- Pal-Bhadra, M., Leibovitch, B. A., Gandhi, S. G., Rao, M., Bhadra, U., Birchler, J. A. & Elgin, S. C. (2004) Heterochromatic silencing and hp1 localization in drosophila are dependent on the rna1 machinery. *Science*, 303(5658), 669-72.
- Papp, B. & Muller, J. (2006) Histone trimethylation and the maintenance of transcriptional on and off states by trrx and pcg proteins. *Genes Dev*, 20(15), 2041-54.
- Park, I. K., Qian, D., Kiel, M., Becker, M. W., Pihalja, M., Weissman, I. L., Morrison, S. J. & Clarke, M. F. (2003) Bmi-1 is required for maintenance of adult self-renewing haematopoietic stem cells. *Nature*, 423(6937), 302-5.
- Parkhurst, S. M., Harrison, D. A., Remington, M. P., Spana, C., Kelley, R. L., Coyne, R. S. & Corces, V. G. (1988) The drosophila su(hw) gene, which controls the phenotypic effect of the gypsy transposable element, encodes a putative DNA-binding protein. *Genes Dev*, 2(10), 1205-15.
- Paro, R. & Hogness, D. S. (1991) The polycomb protein shares a homologous domain with a heterochromatin-associated protein of drosophila. *Proc Natl Acad Sci U S A*, 88(1), 263-7.
- Parthun, M. R., Widom, J. & Gottschling, D. E. (1996) The major cytoplasmic histone acetyltransferase in yeast: Links to chromatin replication and histone metabolism. *Cell*, 87(1), 85-94.
- Partridge, J. F., Scott, K. S., Bannister, A. J., Kouzarides, T. & Allshire, R. C. (2002) Cis-acting DNA from fission yeast centromeres mediates histone h3 methylation and recruitment of silencing factors and cohesin to an ectopic site. *Curr Biol*, 12(19), 1652-60.
- Pasini, D., Bracken, A. P., Jensen, M. R., Lazzarini Denchi, E. & Helin, K. (2004) Suz12 is essential for mouse development and for ezh2 histone methyltransferase activity. *EMBO J*, 23(20), 4061-71.
- Pena, P. V., Davrazou, F., Shi, X., Walter, K. L., Verkhusha, V. V., Gozani, O., Zhao, R. & Kutateladze, T. G. (2006) Molecular mechanism of histone h3k4me3 recognition by plant homeodomain of ing2. *Nature*, 442(7098), 100-3.
- Perrini, B., Piacentini, L., Fanti, L., Altieri, F., Chichiarelli, S., Berloco, M., Turano, C., Ferraro, A. & Pimpinelli, S. (2004) Hp1 controls telomere capping, telomere elongation, and telomere silencing by two different mechanisms in drosophila. *Mol Cell*, 15(3), 467-76.
- Peters, A. H., O'carroll, D., Scherthan, H., Mechtler, K., Sauer, S., Schofer, C., Weipoltshammer, K., Pagani, M., Lachner, M., Kohlmaier, A., Opravil, S., Doyle, M., Sibilia, M. & Jenuwein, T. (2001) Loss of the suv39h histone methyltransferases impairs mammalian heterochromatin and genome stability. *Cell*, 107(3), 323-37.
- Piacentini, L., Fanti, L., Berloco, M., Perrini, B. & Pimpinelli, S. (2003) Heterochromatin protein 1 (hp1) is associated with induced gene expression in drosophila euchromatin. *J Cell Biol*, 161(4), 707-14.
- Pindyurin, A. V., Boldyreva, L. V., Shloma, V. V., Kolesnikova, T. D., Pokholkova, G. V., Andreyeva, E. N., Kozhevnikova, E. N., Ivanoschuk, I. G., Zarutskaya, E. A., Demakov, S. A., Gorchakov, A. A., Belyaeva, E. S. & Zhimulev, I. F. (2008) Interaction between the drosophila heterochromatin proteins suur and hp1. *J Cell Sci*, 121(Pt 10), 1693-703.
- Pinhasov, A., Mandel, S., Torchinsky, A., Giladi, E., Pittel, Z., Goldsweig, A. M., Servoss, S. J., Brenneman, D. E. & Gozes, I. (2003) Activity-dependent neuroprotective protein: A novel gene essential for brain formation. *Brain Res Dev Brain Res*, 144(1), 83-90.
- Pirity, M. K., Locker, J. & Schreiber-Agus, N. (2005) Rybp/dedaf is required for early postimplantation and for central nervous system development. *Mol Cell Biol*, 25(16), 7193-202.
- Platero, J. S., Hartnett, T. & Eisenberg, J. C. (1995) Functional analysis of the chromo domain of hp1. *EMBO J*, 14(16), 3977-86.
- Plath, K., Fang, J., Mlynarczyk-Evans, S. K., Cao, R., Worringer, K. A., Wang, H., De La Cruz, C. C., Otte, A. P., Panning, B. & Zhang, Y. (2003) Role of histone h3 lysine 27 methylation in x inactivation. *Science*, 300(5616), 131-5.
- Pokholok, D. K., Harbison, C. T., Levine, S., Cole, M., Hannett, N. M., Lee, T. I., Bell, G. W., Walker, K., Rolfe, P. A., Herbolzheimer, E., Zeitlinger, J., Lewitter, F., Gifford, D. K. & Young, R. A. (2005) Genome-wide map of nucleosome acetylation and methylation in yeast. *Cell*, 122(4), 517-27.
- Polanowska, J., Martin, J. S., Fisher, R., Scopa, T., Rae, I. & Boulton, S. J. (2004) Tandem immunoaffinity purification of protein complexes from caenorhabditis elegans. *Biotechniques*, 36(5), 778-80, 782.
- Polioudaki, H., Kourmouli, N., Drosou, V., Bakou, A., Theodoropoulos, P. A., Singh, P. B., Giannakouros, T. & Georgatos, S. D. (2001) Histones h3/h4 form a tight complex with the inner nuclear membrane protein lbr and heterochromatin protein 1. *EMBO Rep*, 2(10), 920-5.
- Polo, S. E. & Almouzni, G. (2006) Chromatin assembly: A basic recipe with various flavours. *Curr Opin Genet Dev*, 16(2), 104-11.
- Poser, I., Sarov, M., Hutchins, J. R., Heriche, J. K., Toyoda, Y., Pozniakovsky, A., Weigl, D., Nitzsche, A., Hegemann, B., Bird, A. W., Pelletier, L., Kittler, R., Hua, S., Naumann, R., Augsburg, M., Sykora, M. M., Hofemeister, H., Zhang, Y., Nasmyth, K., White, K. P., Dietzel, S., Mechtler, K., Durbin, R., Stewart, A. F., Peters, J. M., Buchholz, F. & Hyman, A. A. (2008) Bac transgenomics: A high-throughput method for exploration of protein function in mammals. *Nat Methods*, 5(5), 409-15.
- Poux, S., Horard, B., Sigrist, C. J. & Pirrotta, V. (2002) The drosophila trithorax protein is a coactivator required to prevent re-establishment of polycomb silencing. *Development*, 129(10), 2483-93.
- Poux, S., Melfi, R. & Pirrotta, V. (2001) Establishment of polycomb silencing requires a transient interaction between pc and esc. *Genes Dev*, 15(19), 2509-14.
- Pozuelo Rubio, M., Geraghty, K. M., Wong, B. H., Wood, N. T., Campbell, D. G., Morrice, N. & Mackintosh, C.

- (2004) 14-3-3-affinity purification of over 200 human phosphoproteins reveals new links to regulation of cellular metabolism, proliferation and trafficking. *Biochem J*, 379(Pt 2), 395-408.
- Pradhan, S. & Esteve, P. O. (2003) Mammalian DNA (cytosine-5) methyltransferases and their expression. *Clin Immunol*, 109(1), 6-16.
- Preuss, U., Landsberg, G. & Scheidtmann, K. H. (2003) Novel mitosis-specific phosphorylation of histone h3 at thr11 mediated by dlk/zip kinase. *Nucleic Acids Res*, 31(3), 878-85.
- Qin, S. & Parthun, M. R. (2006) Recruitment of the type b histone acetyltransferase hat1p to chromatin is linked to DNA double-strand breaks. *Mol Cell Biol*, 26(9), 3649-58.
- Quenet, D., Gasser, V., Fouillen, L., Cammas, F., Sanglier-Cianferani, S., Losson, R. & Dantzer, F. (2008) The histone subcode: Poly(ADP-ribose) polymerase-1 (parp-1) and parp-2 control cell differentiation by regulating the transcriptional intermediary factor tif1beta and the heterochromatin protein hp1alpha. *FASEB J*, 22(11), 3853-65.
- Quimby, B. B. & Dasso, M. (2003) The small gtpase ran: Interpreting the signs. *Curr Opin Cell Biol*, 15(3), 338-44.
- Quivy, J. P., Roche, D., Kirschner, D., Tagami, H., Nakatani, Y. & Almouzni, G. (2004) A caf-1 dependent pool of hp1 during heterochromatin duplication. *EMBO J*, 23(17), 3516-26.
- Radu, A., Blobel, G. & Moore, M. S. (1995) Identification of a protein complex that is required for nuclear protein import and mediates docking of import substrate to distinct nucleoporins. *Proc Natl Acad Sci U S A*, 92(5), 1769-73.
- Rangasamy, D., Berven, L., Ridgway, P. & Tremethick, D. J. (2003) Pericentric heterochromatin becomes enriched with h2a.Z during early mammalian development. *EMBO J*, 22(7), 1599-607.
- Rangasamy, D., Greaves, I. & Tremethick, D. J. (2004) RNA interference demonstrates a novel role for h2a.Z in chromosome segregation. *Nat Struct Mol Biol*, 11(7), 650-5.
- Rao, P. S., Satelli, A., Zhang, S., Srivastava, S. K., Srivenugopal, K. S. & Rao, U. S. (2009) Rnf2 is the target for phosphorylation by the p38 mapk and erk signaling pathways. *Proteomics*, 9(10), 2776-87.
- Rathmell, W. K. & Chu, G. (1994) Involvement of the ku autoantigen in the cellular response to DNA double-strand breaks. *Proc Natl Acad Sci U S A*, 91(16), 7623-7.
- Ren, X., Vincenz, C. & Kerppola, T. K. (2008) Changes in the distributions and dynamics of polycomb repressive complexes during embryonic stem cell differentiation. *Mol Cell Biol*, 28(9), 2884-95.
- Richon, V. M., Sandhoff, T. W., Rifkind, R. A. & Marks, P. A. (2000) Histone deacetylase inhibitor selectively induces p21waf1 expression and gene-associated histone acetylation. *Proc Natl Acad Sci U S A*, 97(18), 10014-9.
- Rigaut, G., Shevchenko, A., Rutz, B., Wilm, M., Mann, M. & Seraphin, B. (1999) A generic protein purification method for protein complex characterization and proteome exploration. *Nat Biotechnol*, 17(10), 1030-2.
- Ringrose, L., Rehmsmeier, M., Dura, J. M. & Paro, R. (2003) Genome-wide prediction of polycomb/trithorax response elements in *Drosophila melanogaster*. *Dev Cell*, 5(5), 759-71.
- Rinn, J. L., Kertesz, M., Wang, J. K., Squazzo, S. L., Xu, X., Bruggmann, S. A., Goodnough, L. H., Helms, J. A., Farnham, P. J., Segal, E. & Chang, H. Y. (2007) Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. *Cell*, 129(7), 1311-23.
- Robbins, J., Dilworth, S. M., Laskey, R. A. & Dingwall, C. (1991) Two interdependent basic domains in nucleoplasmic nuclear targeting sequence: Identification of a class of bipartite nuclear targeting sequence. *Cell*, 64(3), 615-23.
- Rohila, J. S., Chen, M., Cerny, R. & Fromm, M. E. (2004) Improved tandem affinity purification tag and methods for isolation of protein heterocomplexes from plants. *Plant J*, 38(1), 172-81.
- Roscic, A., Moller, A., Calzado, M. A., Renner, F., Wimmer, V. C., Gresko, E., Ludi, K. S. & Schmitz, M. L. (2006) Phosphorylation-dependent control of pc2 sumo e3 ligase activity by its substrate protein hpk2. *Mol Cell*, 24(1), 77-89.
- Ross, J. M. & Zarkower, D. (2003) Polycomb group regulation of HOX gene expression in *C. elegans*. *Dev Cell*, 4(6), 891-901.
- Ruthenburg, A. J., Allis, C. D. & Wysocka, J. (2007) Methylation of lysine 4 on histone H3: intricacy of writing and reading a single epigenetic mark. *Mol Cell*, 25(1), 15-30.
- Ruthenburg, A. J., Wang, W., Graybosch, D. M., Li, H., Allis, C. D., Patel, D. J. & Verdine, G. L. (2006) Histone H3 recognition and presentation by the WDR5 module of the MLL1 complex. *Nat Struct Mol Biol*, 13(8), 704-12.
- Ryan, R. F., Schultz, D. C., Ayyanathan, K., Singh, P. B., Friedman, J. R., Fredericks, W. J. & Rauscher, F. J., 3rd (1999) Kap-1 corepressor protein interacts and colocalizes with heterochromatic and euchromatic HP1 proteins: A potential role for Kruppel-associated box-zinc finger proteins in heterochromatin-mediated gene silencing. *Mol Cell Biol*, 19(6), 4366-78.
- Saccani, S. & Natoli, G. (2002) Dynamic changes in histone H3 lysine 9 methylation occurring at tightly regulated inducible inflammatory genes. *Genes Dev*, 16(17), 2219-24.
- Salvador, L. M., Park, Y., Cottom, J., Maizels, E. T., Jones, J. C., Schillace, R. V., Carr, D. W., Cheung, P., Allis, C. D., Jameson, J. L. & Hunzicker-Dunn, M. (2001) Follicle-stimulating hormone stimulates protein kinase A-mediated histone H3 phosphorylation and acetylation leading to select gene activation in ovarian granulosa cells. *J Biol Chem*, 276(43), 40146-55.
- Sanchez, C., Sanchez, I., Demmers, J. A., Rodriguez, P., Strouboulis, J. & Vidal, M. (2007) Proteomics analysis of RING1B/RNF2 interactors identifies a novel complex with the FBX10/JHD1B histone demethylase and the BCL6 interacting corepressor. *Mol Cell Proteomics*, 6(5), 820-34.
- Santiveri, C. M., Lechtenberg, B. C., Allen, M. D., Sathyamurthy, A., Jaulent, A. M., Freund, S. M. & Bycroft, M. (2008) The malignant brain tumor repeats of human SCML2 bind to peptides containing monomethylated

- lysine. *J Mol Biol*, 382(5), 1107-12.
- Sarcinella, E., Zuzarte, P. C., Lau, P. N., Draker, R. & Cheung, P. (2007) Monoubiquitylation of h2a.Z distinguishes its association with euchromatin or facultative heterochromatin. *Mol Cell Biol*, 27(18), 6457-68.
- Sarma, K., Margueron, R., Ivanov, A., Pirrotta, V. & Reinberg, D. (2008) Ezh2 requires phf1 to efficiently catalyze h3 lysine 27 trimethylation in vivo. *Mol Cell Biol*, 28(8), 2718-31.
- Sassone-Corsi, P., Mizzen, C. A., Cheung, P., Crosio, C., Monaco, L., Jacquot, S., Hanauer, A. & Allis, C. D. (1999) Requirement of rsk-2 for epidermal growth factor-activated phosphorylation of histone h3. *Science*, 285(5429), 886-91.
- Sathyamurthy, A., Allen, M. D., Murzin, A. G. & Bycroft, M. (2003) Crystal structure of the malignant brain tumor (mbt) repeats in sex comb on midleg-like 2 (scml2). *J Biol Chem*, 278(47), 46968-73.
- Satijn, D. P., Gunster, M. J., Van Der Vlag, J., Hamer, K. M., Schul, W., Alkema, M. J., Saurin, A. J., Freemont, P. S., Van Driel, R. & Otte, A. P. (1997) Ring1 is associated with the polycomb group protein complex and acts as a transcriptional repressor. *Mol Cell Biol*, 17(7), 4105-13.
- Satijn, D. P., Hamer, K. M., Den Blaauwen, J. & Otte, A. P. (2001) The polycomb group protein eed interacts with yy1, and both proteins induce neural tissue in xenopus embryos. *Mol Cell Biol*, 21(4), 1360-9.
- Saurin, A. J., Shao, Z., Erdjument-Bromage, H., Tempst, P. & Kingston, R. E. (2001) A drosophila polycomb group complex includes zeste and dtafi proteins. *Nature*, 412(6847), 655-60.
- Saurin, A. J., Shiels, C., Williamson, J., Satijn, D. P., Otte, A. P., Sheer, D. & Freemont, P. S. (1998) The human polycomb group complex associates with pericentromeric heterochromatin to form a novel nuclear domain. *J Cell Biol*, 142(4), 887-98.
- Savitsky, M., Kravchuk, O., Melnikova, L. & Georgiev, P. (2002) Heterochromatin protein 1 is involved in control of telomere elongation in drosophila melanogaster. *Mol Cell Biol*, 22(9), 3204-18.
- Saxena, A., Yik, J. H. & Weigel, P. H. (2002) H2, the minor subunit of the human asialoglycoprotein receptor, trafficks intracellularly and forms homo-oligomers, but does not bind asialo-orosomucoid. *J Biol Chem*, 277(38), 35297-304.
- Schneider, R., Bannister, A. J. & Kouzarides, T. (2002) Unsafe sets: Histone lysine methyltransferases and cancer. *Trends Biochem Sci*, 27(8), 396-402.
- Schoeftner, S., Sengupta, A. K., Kubicek, S., Mechtler, K., Spahn, L., Koseki, H., Jenuwein, T. & Wutz, A. (2006) Recruitment of prc1 function at the initiation of x inactivation independent of prc2 and silencing. *EMBO J*, 25(13), 3110-22.
- Schoorlemmer, J., Marcos-Gutierrez, C., Were, F., Martinez, R., Garcia, E., Satijn, D. P., Otte, A. P. & Vidal, M. (1997) Ring1a is a transcriptional repressor that interacts with the polycomb-m33 protein and is expressed at rhombomere boundaries in the mouse hindbrain. *EMBO J*, 16(19), 5930-42.
- Schott, S., Coustham, V., Simonet, T., Bedet, C. & Palladino, F. (2006) Unique and redundant functions of c. Elegans hp1 proteins in post-embryonic development. *Dev Biol*, 298(1), 176-87.
- Schotta, G., Lachner, M., Sarma, K., Ebert, A., Sengupta, R., Reuter, G., Reinberg, D. & Jenuwein, T. (2004) A silencing pathway to induce h3-k9 and h4-k20 trimethylation at constitutive heterochromatin. *Genes Dev*, 18(11), 1251-62.
- Schubeler, D., Macalpine, D. M., Scalzo, D., Wirbelauer, C., Kooperberg, C., Van Leeuwen, F., Gottschling, D. E., O'Neill, L. P., Turner, B. M., Delrow, J., Bell, S. P. & Groudine, M. (2004) The histone modification pattern of active genes revealed through genome-wide chromatin analysis of a higher eukaryote. *Genes Dev*, 18(11), 1263-71.
- Schuetz, A., Allali-Hassani, A., Martin, F., Loppnau, P., Vedadi, M., Bochkarev, A., Plotnikov, A. N., Arrowsmith, C. H. & Min, J. (2006) Structural basis for molecular recognition and presentation of histone h3 by wdr5. *EMBO J*, 25(18), 4245-52.
- Schultz, D. C., Ayyanathan, K., Negorev, D., Maul, G. G. & Rauscher, F. J., 3rd (2002) Setdb1: A novel kap-1-associated histone h3, lysine 9-specific methyltransferase that contributes to hp1-mediated silencing of euchromatic genes by krab zinc-finger proteins. *Genes Dev*, 16(8), 919-32.
- Schultz, D. C., Friedman, J. R. & Rauscher, F. J., 3rd (2001) Targeting histone deacetylase complexes via krab-zinc finger proteins: The phd and bromodomains of kap-1 form a cooperative unit that recruits a novel isoform of the mi-2alpha subunit of nurd. *Genes Dev*, 15(4), 428-43.
- Schumacher, A. & Magnuson, T. (1997) Murine polycomb- and trithorax-group genes regulate homeotic pathways and beyond. *Trends Genet*, 13(5), 167-70.
- Schwartz, Y. B., Kahn, T. G., Nix, D. A., Li, X. Y., Bourgon, R., Biggin, M. & Pirrotta, V. (2006) Genome-wide analysis of polycomb targets in drosophila melanogaster. *Nat Genet*, 38(6), 700-5.
- Schwartz, Y. B. & Pirrotta, V. (2007) Polycomb silencing mechanisms and the management of genomic programmes. *Nat Rev Genet*, 8(1), 9-22.
- Seet, B. T., Dikic, I., Zhou, M. M. & Pawson, T. (2006) Reading protein modifications with interaction domains. *Nat Rev Mol Cell Biol*, 7(7), 473-83.
- Seum, C., Reo, E., Peng, H., Rauscher, F. J., 3rd, Spierer, P. & Bontron, S. (2007) Drosophila setdb1 is required for chromosome 4 silencing. *PLoS Genet*, 3(5), e76.
- Sewalt, R. G., Gunster, M. J., Van Der Vlag, J., Satijn, D. P. & Otte, A. P. (1999) C-terminal binding protein is a transcriptional repressor that interacts with a specific class of vertebrate polycomb proteins. *Mol Cell Biol*, 19(1), 777-87.
- Shao, Z., Raible, F., Mollaaghababa, R., Guyon, J. R., Wu, C. T., Bender, W. & Kingston, R. E. (1999) Stabilization of chromatin structure by prc1, a polycomb complex. *Cell*, 98(1), 37-46.

- Shareef, M. M., King, C., Damaj, M., Badagu, R., Huang, D. W. & Kellum, R. (2001) Drosophila heterochromatin protein 1 (hp1)/origin recognition complex (orc) protein is associated with hp1 and orc and functions in heterochromatin-induced silencing. *Mol Biol Cell*, 12(6), 1671-85.
- Shen, X., Liu, Y., Hsu, Y. J., Fujiwara, Y., Kim, J., Mao, X., Yuan, G. C. & Orkin, S. H. (2008) Ezh1 mediates methylation on histone h3 lysine 27 and complements ezh2 in maintaining stem cell identity and executing pluripotency. *Mol Cell*, 32(4), 491-502.
- Shi, Y., Lan, F., Matson, C., Mulligan, P., Whetstone, J. R., Cole, P. A. & Casero, R. A. (2004) Histone demethylation mediated by the nuclear amine oxidase homolog lsd1. *Cell*, 119(7), 941-53.
- Shio, Y. & Eisenman, R. N. (2003) Histone sumoylation is associated with transcriptional repression. *Proc Natl Acad Sci U S A*, 100(23), 13225-30.
- Shimada, M., Niida, H., Zineldeen, D. H., Tagami, H., Tanaka, M., Saito, H. & Nakanishi, M. (2008) Chk1 is a histone h3 threonine 11 kinase that regulates DNA damage-induced transcriptional repression. *Cell*, 132(2), 221-32.
- Shoemaker, C. B. & Chalkley, R. (1980) H3-specific nucleohistone kinase of bovine thymus chromatin. Purification, characterization, and specificity for threonine residue 3. *J Biol Chem*, 255(22), 11048-55.
- Shogren-Knaak, M., Ishii, H., Sun, J. M., Pazin, M. J., Davie, J. R. & Peterson, C. L. (2006) Histone h4-k16 acetylation controls chromatin structure and protein interactions. *Science*, 311(5762), 844-7.
- Sijen, T. & Plasterk, R. H. (2003) Transposon silencing in the caenorhabditis elegans germ line by natural rna. *Nature*, 426(6964), 310-4.
- Sims, R. J., 3rd & Reinberg, D. (2008) Is there a code embedded in proteins that is based on post-translational modifications? *Nat Rev Mol Cell Biol*, 9(10), 815-20.
- Singh, P. B., Miller, J. R., Pearce, J., Kothary, R., Burton, R. D., Paro, R., James, T. C. & Gaunt, S. J. (1991) A sequence motif found in a drosophila heterochromatin protein is conserved in animals and plants. *Nucleic Acids Res*, 19(4), 789-94.
- Smothers, J. F. & Henikoff, S. (2001) The hinge and chromo shadow domain impart distinct targeting of hp1-like proteins. *Mol Cell Biol*, 21(7), 2555-69.
- Soloaga, A., Thomson, S., Wiggan, G. R., Rampersaud, N., Dyson, M. H., Hazzalin, C. A., Mahadevan, L. C. & Arthur, J. S. (2003) Msk2 and msk1 mediate the mitogen- and stress-induced phosphorylation of histone h3 and hmg-14. *EMBO J*, 22(11), 2788-97.
- Song, J. J., Garlick, J. D. & Kingston, R. E. (2008) Structural basis of histone h4 recognition by p55. *Genes Dev*, 22(10), 1313-8.
- Song, K., Jung, Y., Jung, D. & Lee, I. (2001) Human ku70 interacts with heterochromatin protein 1alpha. *J Biol Chem*, 276(11), 8321-7.
- Sparmann, A. & Van Lohuizen, M. (2006) Polycomb silencers control cell fate, development and cancer. *Nat Rev Cancer*, 6(11), 846-56.
- Spencer, V. A. & Davie, J. R. (1999) Role of covalent modifications of histones in regulating gene expression. *Gene*, 240(1), 1-12.
- Sripathy, S. P., Stevens, J. & Schultz, D. C. (2006) The kap1 corepressor functions to coordinate the assembly of de novo hp1-demarcated microenvironments of heterochromatin required for krab zinc finger protein-mediated transcriptional repression. *Mol Cell Biol*, 26(22), 8623-38.
- Stabell, M., Eskeland, R., Bjorkmo, M., Larsson, J., Aalen, R. B., Imhof, A. & Lambertsson, A. (2006) The drosophila g9a gene encodes a multi-catalytic histone methyltransferase required for normal development. *Nucleic Acids Res*, 34(16), 4609-21.
- Stanton, S. E., McReynolds, L. J., Evans, T. & Schreiber-Agus, N. (2006) Yaf2 inhibits caspase 8-mediated apoptosis and regulates cell survival during zebrafish embryogenesis. *J Biol Chem*, 281(39), 28782-93.
- Steger, D. J., Lefterova, M. I., Ying, L., Stonestrom, A. J., Schupp, M., Zhuo, D., Vakoc, A. L., Kim, J. E., Chen, J., Lazar, M. A., Blobel, G. A. & Vakoc, C. R. (2008) Dot1/kmt4 recruitment and h3k79 methylation are ubiquitously coupled with gene transcription in mammalian cells. *Mol Cell Biol*, 28(8), 2825-39.
- Steingart, R. A. & Gozes, I. (2006) Recombinant activity-dependent neuroprotective protein protects cells against oxidative stress. *Mol Cell Endocrinol*, 252(1-2), 148-53.
- Stelzl, U., Worm, U., Lalowski, M., Haenig, C., Brembeck, F. H., Goehler, H., Stroedicke, M., Zenkner, M., Schoenherr, A., Koeppen, S., Timm, J., Mintzlaff, S., Abraham, C., Bock, N., Kietzmann, S., Goedde, A., Toksoz, E., Droege, A., Krobitsch, S., Korn, B., Birchmeier, W., Lehrach, H. & Wanker, E. E. (2005) A human protein-protein interaction network: A resource for annotating the proteome. *Cell*, 122(6), 957-68.
- Serner, D. E. & Berger, S. L. (2000) Acetylation of histones and transcription-related factors. *Microbiol Mol Biol Rev*, 64(2), 435-59.
- Stewart, M. D., Li, J. & Wong, J. (2005) Relationship between histone h3 lysine 9 methylation, transcription repression, and heterochromatin protein 1 recruitment. *Mol Cell Biol*, 25(7), 2525-38.
- Stock, J. K., Giadrossi, S., Casanova, M., Brookes, E., Vidal, M., Koseki, H., Brockdorff, N., Fisher, A. G. & Pombo, A. (2007) Ring1-mediated ubiquitination of h2a restrains poised rna polymerase ii at bivalent genes in mouse es cells. *Nat Cell Biol*, 9(12), 1428-35.
- Strahl, B. D. & Allis, C. D. (2000) The language of covalent histone modifications. *Nature*, 403(6765), 41-5.
- Stucki, M., Clapperton, J. A., Mohammad, D., Yaffe, M. B., Smerdon, S. J. & Jackson, S. P. (2005) Mdc1 directly binds phosphorylated histone h2ax to regulate cellular responses to DNA double-strand breaks. *Cell*, 123(7), 1213-26.
- Suetake, I., Shinozaki, F., Miyagawa, J., Takeshima, H. & Tajima, S. (2004) Dnmt3l stimulates the DNA methylation activity of dnmt3a and dnmt3b through a direct interaction. *J Biol Chem*, 279(26), 27816-23.

- Sullivan, K. F. (2001) A solid foundation: Functional specialization of centromeric chromatin. *Curr Opin Genet Dev*, 11(2), 182-8.
- Taddei, A., Maison, C., Roche, D. & Almouzni, G. (2001) Reversible disruption of pericentric heterochromatin and centromere function by inhibiting deacetylases. *Nat Cell Biol*, 3(2), 114-20.
- Taddei, A., Roche, D., Bickmore, W. A. & Almouzni, G. (2005) The effects of histone deacetylase inhibitors on heterochromatin: Implications for anticancer therapy? *EMBO Rep*, 6(6), 520-4.
- Tajul-Arifin, K., Teasdale, R., Ravasi, T., Hume, D. A. & Mattick, J. S. (2003) Identification and analysis of chromodomain-containing proteins encoded in the mouse transcriptome. *Genome Res*, 13(6B), 1416-29.
- Tamaru, H. & Selker, E. U. (2001) A histone h3 methyltransferase controls DNA methylation in *Neurospora crassa*. *Nature*, 414(6861), 277-83.
- Tan, C. P. & Nakielnny, S. (2006) Control of the DNA methylation system component mbd2 by protein arginine methylation. *Mol Cell Biol*, 26(19), 7224-35.
- Taylor-Harding, B., Binne, U. K., Korenjak, M., Brehm, A. & Dyson, N. J. (2004) P55, the drosophila ortholog of Rbap46/Rbap48, is required for the repression of *de2f2/rbf*-regulated genes. *Mol Cell Biol*, 24(20), 9124-36.
- Tetsu, O., Ishihara, H., Kanno, R., Kamiyasu, M., Inoue, H., Tokuhisa, T., Taniguchi, M. & Kanno, M. (1998) Mel-18 negatively regulates cell cycle progression upon B cell antigen receptor stimulation through a cascade leading to c-myc/cdc25. *Immunity*, 9(4), 439-48.
- Thompson, P. R. & Fast, W. (2006) Histone citrullination by protein arginine deiminase: Is arginine methylation a green light or a roadblock? *ACS Chem Biol*, 1(7), 433-41.
- Tie, F., Prasad-Sinha, J., Birve, A., Rasmuson-Lestander, A. & Harte, P. J. (2003) A 1-megadalton esc/e(z) complex from drosophila that contains polycomblike and rpd3. *Mol Cell Biol*, 23(9), 3352-62.
- Tie, F., Siebold, A. P. & Harte, P. J. (2005) The N-terminus of drosophila esc mediates its phosphorylation and dimerization. *Biochem Biophys Res Commun*, 332(2), 622-32.
- Tie, F., Stratton, C. A., Kurzhals, R. L. & Harte, P. J. (2007) The N terminus of drosophila esc binds directly to histone H3 and is required for e(z)-dependent trimethylation of H3 lysine 27. *Mol Cell Biol*, 27(6), 2014-26.
- Todd, A., Cossons, N., Aitken, A., Price, G. B. & Zannis-Hadjopoulos, M. (1998) Human cruciform binding protein belongs to the 14-3-3 family. *Biochemistry*, 37(40), 14317-25.
- Tolhuis, B., De Wit, E., Muijters, I., Teunissen, H., Talhout, W., Van Steensel, B. & Van Lohuizen, M. (2006) Genome-wide profiling of PRC1 and PRC2 polycomb chromatin binding in drosophila melanogaster. *Nat Genet*, 38(6), 694-9.
- Tong, J. K., Hassig, C. A., Schnitzler, G. R., Kingston, R. E. & Schreiber, S. L. (1998) Chromatin deacetylation by an ATP-dependent nucleosome remodeling complex. *Nature*, 395(6705), 917-21.
- Trojer, P., Li, G., Sims, R. J., 3rd, Vaquero, A., Kalakonda, N., Boccuni, P., Lee, D., Erdjument-Bromage, H., Tempst, P., Nimer, S. D., Wang, Y. H. & Reinberg, D. (2007) L3mbtl1, a histone-methylation-dependent chromatin lock. *Cell*, 129(5), 915-28.
- Tsukada, Y., Fang, J., Erdjument-Bromage, H., Warren, M. E., Borchers, C. H., Tempst, P. & Zhang, Y. (2006) Histone demethylation by a family of JmjC domain-containing proteins. *Nature*, 439(7078), 811-6.
- Turner, B. M. (2005) Reading signals on the nucleosome with a new nomenclature for modified histones. *Nat Struct Mol Biol*, 12(2), 110-2.
- Vakoc, C. R., Mandat, S. A., Olenchok, B. A. & Blobel, G. A. (2005) Histone H3 lysine 9 methylation and H3K9me3 are associated with transcription elongation through mammalian chromatin. *Mol Cell*, 19(3), 381-91.
- Van Der Lugt, N. M., Domen, J., Linders, K., Van Roon, M., Robanus-Maandag, E., Te Riele, H., Van Der Valk, M., Deschamps, J., Sofroniew, M., Van Lohuizen, M. & Et Al. (1994) Posterior transformation, neurological abnormalities, and severe hematopoietic defects in mice with a targeted deletion of the *bmi-1* proto-oncogene. *Genes Dev*, 8(7), 757-69.
- Van Der Stoop, P., Boutsma, E. A., Hulsman, D., Noback, S., Heimerikx, M., Kerkhoven, R. M., Voncken, J. W., Wessels, L. F. & Van Lohuizen, M. (2008) Ubiquitin E3 ligase ring1b/rnf2 of polycomb repressive complex 1 contributes to stable maintenance of mouse embryonic stem cells. *PLoS One*, 3(5), e2235.
- Van Leene, J., Stals, H., Eeckhout, D., Persiau, G., Van De Slijke, E., Van Isterdael, G., De Clercq, A., Bonnet, E., Laukens, K., Remmerie, N., Henderickx, K., De Vijlder, T., Abdelkrim, A., Pharezyn, A., Van Onckelen, H., Inze, D., Witters, E. & De Jaeger, G. (2007) A tandem affinity purification-based technology platform to study the cell cycle interactome in Arabidopsis thaliana. *Mol Cell Proteomics*, 6(7), 1226-38.
- Van Leene, J., Witters, E., Inze, D. & De Jaeger, G. (2008) Boosting tandem affinity purification of plant protein complexes. *Trends Plant Sci*, 13(10), 517-20.
- Vandel, L., Nicolas, E., Vaute, O., Ferreira, R., Ait-Si-Ali, S. & Trouche, D. (2001) Transcriptional repression by the retinoblastoma protein through the recruitment of a histone methyltransferase. *Mol Cell Biol*, 21(19), 6484-94.
- Varambally, S., Dhanasekaran, S. M., Zhou, M., Barrette, T. R., Kumar-Sinha, C., Sanda, M. G., Ghosh, D., Pienta, K. J., Sewalt, R. G., Ote, A. P., Rubin, M. A. & Chinnaiyan, A. M. (2002) The polycomb group protein *ezh2* is involved in progression of prostate cancer. *Nature*, 419(6907), 624-9.
- Veraksa, A., Bauer, A. & Artavanis-Tsakonas, S. (2005) Analyzing protein complexes in drosophila with tandem affinity purification-mass spectrometry. *Dev Dyn*, 232(3), 827-34.
- Verdel, A., Jia, S., Gerber, S., Sugiyama, T., Gygi, S., Grewal, S. I. & Moazed, D. (2004) RNAi-mediated targeting

- of heterochromatin by the rits complex. *Science*, 303(5658), 672-6.
- Vermaak, D., Henikoff, S. & Malik, H. S. (2005) Positive selection drives the evolution of rhino, a member of the heterochromatin protein 1 family in drosophila. *PLoS Genet*, 1(1), 96-108.
- Vermeulen, M., Mulder, K. W., Denissov, S., Pijnappel, W. W., Van Schaik, F. M., Varier, R. A., Baltissen, M. P., Stunnenberg, H. G., Mann, M. & Timmers, H. T. (2007) Selective anchoring of tfiid to nucleosomes by trimethylation of histone h3 lysine 4. *Cell*, 131(1), 58-69.
- Vincenz, C. & Kerppola, T. K. (2008) Different polycomb group cbx family proteins associate with distinct regions of chromatin using nonhomologous protein sequences. *Proc Natl Acad Sci U S A*, 105(43), 16572-7.
- Volff, J. N. (2006) Turning junk into gold: Domestication of transposable elements and the creation of new genes in eukaryotes. *Bioessays*, 28(9), 913-22.
- Volpe, T. A., Kidner, C., Hall, I. M., Teng, G., Grewal, S. I. & Martienssen, R. A. (2002) Regulation of heterochromatic silencing and histone h3 lysine-9 methylation by rnaï. *Science*, 297(5588), 1833-7.
- Voncken, J. W., Niessen, H., Neufeld, B., Rennefahrt, U., Dahlmans, V., Kubben, N., Holzer, B., Ludwig, S. & Rapp, U. R. (2005) Mapkap kinase 3pk phosphorylates and regulates chromatin association of the polycomb group protein bmi1. *J Biol Chem*, 280(7), 5178-87.
- Voncken, J. W., Roelen, B. A., Roefs, M., De Vries, S., Verhoeven, E., Marino, S., Deschamps, J. & Van Lohuizen, M. (2003) Rnf2 (ring1b) deficiency causes gastrulation arrest and cell cycle inhibition. *Proc Natl Acad Sci U S A*, 100(5), 2468-73.
- Voncken, J. W., Schweizer, D., Aagaard, L., Sattler, L., Jantsch, M. F. & Van Lohuizen, M. (1999) Chromatin-association of the polycomb group protein bmi1 is cell cycle-regulated and correlates with its phosphorylation status. *J Cell Sci*, 112 (Pt 24)(4627-39.
- Waddington, C. H. (1942) The epigenotype. *Endeavour*, 1(18-20.
- Wakimoto, B. T. & Hearn, M. G. (1990) The effects of chromosome rearrangements on the expression of heterochromatic genes in chromosome 2l of drosophila melanogaster. *Genetics*, 125(1), 141-54.
- Wang, G., Ma, A., Chow, C. M., Horsley, D., Brown, N. R., Cowell, I. G. & Singh, P. B. (2000) Conservation of heterochromatin protein 1 function. *Mol Cell Biol*, 20(18), 6970-83.
- Wang, G. G., Allis, C. D. & Chi, P. (2007) Chromatin remodeling and cancer, part ii: Atp-dependent chromatin remodeling. *Trends Mol Med*, 13(9), 373-80.
- Wang, H., Davis, A., Yu, S. & Ahmed, K. (2001) Response of cancer cells to molecular interruption of the ck2 signal. *Mol Cell Biochem*, 227(1-2), 167-74.
- Wang, H., Wang, L., Erdjument-Bromage, H., Vidal, M., Tempst, P., Jones, R. S. & Zhang, Y. (2004) Role of histone h2a ubiquitination in polycomb silencing. *Nature*, 431(7010), 873-8.
- Wang, H., Wang, M., Bocker, W. & Iliakis, G. (2005) Complex h2ax phosphorylation patterns by multiple kinases including atm and DNA-pk in human cells exposed to ionizing radiation and treated with kinase inhibitors. *J Cell Physiol*, 202(2), 492-502.
- Wang, H., Zhai, L., Xu, J., Joo, H. Y., Jackson, S., Erdjument-Bromage, H., Tempst, P., Xiong, Y. & Zhang, Y. (2006) Histone h3 and h4 ubiquitylation by the cul4-ddb-roc1 ubiquitin ligase facilitates cellular response to DNA damage. *Mol Cell*, 22(3), 383-94.
- Wang, L., Brown, J. L., Cao, R., Zhang, Y., Kassis, J. A. & Jones, R. S. (2004) Hierarchical recruitment of polycomb group silencing complexes. *Mol Cell*, 14(5), 637-46.
- Wang, W., Cote, J., Xue, Y., Zhou, S., Khavari, P. A., Biggar, S. R., Muchardt, C., Kalpana, G. V., Goff, S. P., Yaniv, M., Workman, J. L. & Crabtree, G. R. (1996) Purification and biochemical heterogeneity of the mammalian swi-snf complex. *EMBO J*, 15(19), 5370-82.
- Wang, Y., Wysocka, J., Sayegh, J., Lee, Y. H., Perlin, J. R., Leonelli, L., Sonbuchner, L. S., McDonald, C. H., Cook, R. G., Dou, Y., Roeder, R. G., Clarke, S., Stallcup, M. R., Allis, C. D. & Coonrod, S. A. (2004) Human pad4 regulates histone arginine methylation levels via demethylination. *Science*, 306(5694), 279-83.
- Waterston, R. H., Lindblad-Toh, K., Birney, E., Rogers, J., Abril, J. F., Agarwal, P., Agarwala, R., Ainscough, R., Alexandersson, M., An, P., Antonarakis, S. E., Attwood, J., Baertsch, R., Bailey, J., Barlow, K., Beck, S., Berry, E., Birren, B., Bloom, T., Bork, P., Botcherby, M., Bray, N., Brent, M. R., Brown, D. G., Brown, S. D., Bult, C., Burton, J., Butler, J., Campbell, R. D., Carninci, P., Cawley, S., Chiaromonte, F., Chinwalla, A. T., Church, D. M., Clamp, M., Clee, C., Collins, F. S., Cook, L. L., Copley, R. R., Coulson, A., Couronne, O., Cuff, J., Curwen, V., Cutts, T., Daly, M., David, R., Davies, J., Delehaunty, K. D., Deri, J., Dermitzakis, E. T., Dewey, C., Dickens, N. J., Diekhans, M., Dodge, S., Dubchak, I., Dunn, D. M., Eddy, S. R., Elnitski, L., Emes, R. D., Eswara, P., Eyas, E., Felsenfeld, A., Fewell, G. A., Flicek, P., Foley, K., Frankel, W. N., Fulton, L. A., Fulton, R. S., Furey, T. S., Gage, D., Gibbs, R. A., Glusman, G., Gnerre, S., Goldman, N., Goodstadt, L., Grafham, D., Graves, T. A., Green, E. D., Gregory, S., Guigo, R., Guyer, M., Hardison, R. C., Haussler, D., Hayashizaki, Y., Hillier, L. W., Hinrichs, A., Hlavina, W., Holzer, T., Hsu, F., Hua, A., Hubbard, T., Hunt, A., Jackson, I., Jaffe, D. B., Johnson, L. S., Jones, M., Jones, T. A., Joy, A., Kamal, M., Karlsson, E. K., et al. (2002) Initial sequencing and comparative analysis of the mouse genome. *Nature*, 420(6915), 520-62.
- Watson, J. D. & Crick, F. H. (1953) Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature*, 171(4356), 737-8.
- Wiederschain, D., Chen, L., Johnson, B., Bettano, K., Jackson, D., Taraszka, J., Wang, Y. K., Jones, M. D., Morrissey, M., Deeds, J., Mosher, R., Fordjour, P., Lengauer, C. & Benson, J. D. (2007) Contribution of polycomb homologues bmi-1 and mel-18 to medulloblastoma pathogenesis. *Mol Cell Biol*, 27(13), 4968-79.

- Wisniewski, J. R., Zougman, A. & Mann, M. (2008) Nepsilon-formylation of lysine is a widespread post-translational modification of nuclear proteins occurring at residues involved in regulation of chromatin function. *Nucleic Acids Res*, 36(2), 570-7.
- Wolffe, A. P. & Hayes, J. J. (1999) Chromatin disruption and modification. *Nucleic Acids Res*, 27(3), 711-20.
- Woodage, T., Basrai, M. A., Baxeivanis, A. D., Hieter, P. & Collins, F. S. (1997) Characterization of the chd family of proteins. *Proc Natl Acad Sci U S A*, 94(21), 11472-7.
- Wysocka, J., Swigut, T., Milne, T. A., Dou, Y., Zhang, X., Burlingame, A. L., Roeder, R. G., Brivanlou, A. H. & Allis, C. D. (2005) Wdr5 associates with histone h3 methylated at k4 and is essential for h3 k4 methylation and vertebrate development. *Cell*, 121(6), 859-72.
- Wysocka, J., Swigut, T., Xiao, H., Milne, T. A., Kwon, S. Y., Landry, J., Kauer, M., Tackett, A. J., Chait, B. T., Badenhorst, P., Wu, C. & Allis, C. D. (2006) A phd finger of nurf couples histone h3 lysine 4 trimethylation with chromatin remodelling. *Nature*, 442(7098), 86-90.
- Xu, W., Chen, H., Du, K., Asahara, H., Tini, M., Emerson, B. M., Montminy, M. & Evans, R. M. (2001) A transcriptional switch mediated by cofactor methylation. *Science*, 294(5551), 2507-11.
- Xu, X., Toselli, P. A., Russell, L. D. & Seldin, D. C. (1999) Globozoospermia in mice lacking the casein kinase ii alpha' catalytic subunit. *Nat Genet*, 23(1), 118-21.
- Xue, Y., Wong, J., Moreno, G. T., Young, M. K., Cote, J. & Wang, W. (1998) Nurd, a novel complex with both atp-dependent chromatin-remodeling and histone deacetylase activities. *Mol Cell*, 2(6), 851-61.
- Yaffe, M. B., Rittinger, K., Volinia, S., Caron, P. R., Aitken, A., Leffers, H., Gambli, S. J., Smerdon, S. J. & Cantley, L. C. (1997) The structural basis for 14-3-3:Phosphopeptide binding specificity. *Cell*, 91(7), 961-71.
- Yamaki, M., Isono, K., Takada, Y., Abe, K., Akasaka, T., Tanzawa, H. & Koseki, H. (2002) The mouse *edr2* (*mph2*) gene has two forms of mrna encoding 90- and 36-kda polypeptides. *Gene*, 288(1-2), 103-10.
- Yamamoto, Y., Verma, U. N., Prajapati, S., Kwak, Y. T. & Gaynor, R. B. (2003) Histone h3 phosphorylation by *ikk-alpha* is critical for cytokine-induced gene expression. *Nature*, 423(6940), 655-9.
- Yamane, K., Toumazou, C., Tsukada, Y., Erdjument-Bromage, H., Tempst, P., Wong, J. & Zhang, Y. (2006) *Jhdm2a*, a *jmjc*-containing h3k9 demethylase, facilitates transcription activation by androgen receptor. *Cell*, 125(3), 483-95.
- Ye, Q., Callebaut, I., Pezhman, A., Courvalin, J. C. & Worman, H. J. (1997) Domain-specific interactions of human *hp1*-type chromodomain proteins and inner nuclear membrane protein *lbr*. *J Biol Chem*, 272(23), 14983-9.
- Ye, Q. & Worman, H. J. (1996) Interaction between an integral protein of the nuclear envelope inner membrane and human chromodomain proteins homologous to *drosophila hp1*. *J Biol Chem*, 271(25), 14653-6.
- Young, A. P. & Longmore, G. D. (2004) Differences in stability of repressor complexes at promoters underlie distinct roles for *rb* family members. *Oncogene*, 23(3), 814-23.
- Zalensky, A. O., Siino, J. S., Gineitis, A. A., Zalenskaya, I. A., Tomilin, N. V., Yau, P. & Bradbury, E. M. (2002) Human testis/sperm-specific histone h2b (*htsh2b*). Molecular cloning and characterization. *J Biol Chem*, 277(45), 43474-80.
- Zamostiano, R., Pinhasov, A., Gelber, E., Steingart, R. A., Seroussi, E., Giladi, E., Bassan, M., Wollman, Y., Eyre, H. J., Mulley, J. C., Brenneman, D. E. & Gozes, I. (2001) Cloning and characterization of the human activity-dependent neuroprotective protein. *J Biol Chem*, 276(1), 708-14.
- Zeitlin, S. G., Shelby, R. D. & Sullivan, K. F. (2001) *Cenp-a* is phosphorylated by *aurora b* kinase and plays an unexpected role in completion of cytokinesis. *J Cell Biol*, 155(7), 1147-57.
- Zeng, L. & Zhou, M. M. (2002) Bromodomain: An acetyl-lysine binding domain. *FEBS Lett*, 513(1), 124-8.
- Zhang, H., Christoforou, A., Aravind, L., Emmons, S. W., Van Den Heuvel, S. & Haber, D. A. (2004) The *c. Elegans polycomb* gene *sop-2* encodes an rna binding protein. *Mol Cell*, 14(6), 841-7.
- Zhang, J., Yu, C., Pulletikurti, V., Lamb, J., Danilova, T., Weber, D. F., Birchler, J. & Peterson, T. (2009) Alternative *ac/ds* transposition induces major chromosomal rearrangements in maize. *Genes Dev*, 23(6), 755-65.
- Zhang, R. & Adams, P. D. (2007) Heterochromatin and its relationship to cell senescence and cancer therapy. *Cell Cycle*, 6(7), 784-9.
- Zhang, R., Poustovoitov, M. V., Ye, X., Santos, H. A., Chen, W., Daganzo, S. M., Erzberger, J. P., Serebriiskii, I. G., Canutescu, A. A., Dunbrack, R. L., Pehrson, J. R., Berger, J. M., Kaufman, P. D. & Adams, P. D. (2005) Formation of *macroh2a*-containing senescence-associated heterochromatin foci and senescence driven by *asf1a* and *hira*. *Dev Cell*, 8(1), 19-30.
- Zhang, T., Sun, Y., Tian, E., Deng, H., Zhang, Y., Luo, X., Cai, Q., Wang, H., Chai, J. & Zhang, H. (2006) Rna-binding proteins *sop-2* and *sor-1* form a novel *pcg*-like complex in *c. Elegans*. *Development*, 133(6), 1023-33.
- Zhang, W., Deng, H., Bao, X., Lerach, S., Girton, J., Johansen, J. & Johansen, K. M. (2006) The *jil-1* histone h3s10 kinase regulates dimethyl h3k9 modifications and heterochromatic spreading in *drosophila*. *Development*, 133(2), 229-35.
- Zhang, X. Y., Pfeiffer, H. K., Thorne, A. W. & McMahon, S. B. (2008) *Usp22*, an *hsaga* subunit and potential cancer stem cell marker, reverses the polycomb-catalyzed ubiquitylation of histone h2a. *Cell Cycle*, 7(11), 1522-4.
- Zhang, X. Y., Varthi, M., Sykes, S. M., Phillips, C., Warzecha, C., Zhu, W., Wyce, A., Thorne, A. W., Berger, S. L. & McMahon, S. B. (2008) The putative cancer stem cell marker *usp22* is a subunit of the human *saga* complex required for activated transcription and cell-cycle progression. *Mol Cell*, 29(1), 102-11.

- Zhang, Y., Hubner, I. A., Arakaki, A. K., Shakhnovich, E. & Skolnick, J. (2006) On the origin and highly likely completeness of single-domain protein structures. *Proc Natl Acad Sci U S A*, 103(8), 2605-10.
- Zhang, Y., Leroy, G., Seelig, H. P., Lane, W. S. & Reinberg, D. (1998) The dermatomyositis-specific autoantigen mi2 is a component of a complex containing histone deacetylase and nucleosome remodeling activities. *Cell*, 95(2), 279-89.
- Zhao, T. & Eisenberg, J. C. (1999) Phosphorylation of heterochromatin protein 1 by casein kinase ii is required for efficient heterochromatin binding in drosophila. *J Biol Chem*, 274(21), 15095-100.
- Zhao, T., Heyduk, T., Allis, C. D. & Eisenberg, J. C. (2000) Heterochromatin protein 1 binds to nucleosomes and DNA in vitro. *J Biol Chem*, 275(36), 28332-8.
- Zhao, T., Heyduk, T. & Eisenberg, J. C. (2001) Phosphorylation site mutations in heterochromatin protein 1 (hp1) reduce or eliminate silencing activity. *J Biol Chem*, 276(12), 9512-8.
- Zhou, D., Ren, J. X., Ryan, T. M., Higgins, N. P. & Townes, T. M. (2004) Rapid tagging of endogenous mouse genes by recombineering and es cell complementation of tetraploid blastocysts. *Nucleic Acids Res*, 32(16), e128.
- Zhou, W., Zhu, P., Wang, J., Pascual, G., Ohgi, K. A., Lozach, J., Glass, C. K. & Rosenfeld, M. G. (2008) Histone h2a monoubiquitination represses transcription by inhibiting rna polymerase ii transcriptional elongation. *Mol Cell*, 29(1), 69-80.
- Zhu, P., Zhou, W., Wang, J., Puc, J., Ohgi, K. A., Erdjument-Bromage, H., Tempst, P., Glass, C. K. & Rosenfeld, M. G. (2007) A histone h2a deubiquitinase complex coordinating histone acetylation and h1 dissociation in transcriptional regulation. *Mol Cell*, 27(4), 609-21.