BASES MOLECULAIRES DE LA VARIEGATION D'EFFETDE POSITION CHEZ DROSOPHILA MELANOGASTER

Roxane Blattes

▶ **To cite this version:**

Roxane Blattes. BASES MOLECULAIRES DE LA VARIEGATION D'EFFETDE POSITION CHEZ DROSOPHILA MELANOGASTER. Biochimie [q-bio.BM]. Université Paul Sabatier - Toulouse III, 2006. Français. NNT : . tel-00107128

HAL Id: tel-00107128

https://theses.hal.science/tel-00107128

Submitted on 17 Oct 2006

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

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

N° d'ordre

THESE

présentée devant

L'UNIVERSITE TOULOUSE III - PAUL SABATIER
118 route de Narbonne 31062 TOULOUSE CEDEX

en vue de l'obtention du grade de

DOCTEUR DE L'UNIVERSITE TOULOUSE III

spécialité : biologie moléculaire et cellulaire

par

Roxane BLATTES

BASES MOLECULAIRES DE LA VARIEGATION D'EFFET DE POSITION CHEZ *DROSOPHILA MELANOGASTER*

soutenue le vendredi 06 octobre 2006 à 14h30

salle de conférence de l'IEFG, Bâtiment IBCG,
118 route de Narbonne, 31062 Toulouse CEDEX 09

devant de jury de thèse composé de :

M. P.-E. GLEIZES, Professeur, UPS LBME, Toulouse	Président
M. G. CAVALLI, Directeur de Recherche, IGH, Montpellier	Rapporteur
M. P. SPIERER, Professeur, Université de Genève, Genève	Rapporteur
Mme I. LEGER-SILVESTRE, Maître de Conférences, UPS LBME, Toulouse	Examinateur
M. O.CUVIER, Chargé de Recherche, IGH, Montpellier	Examinateur
M. E. KÄS, Directeur de Recherche, LBME, Toulouse	Directeur de Thèse

REMERCIEMENTS

Je tiens tout d'abord à remercier les membres du jury pour avoir accepté d'évaluer ce travail : **Pierre-Emmanuel GLEIZES**, pour avoir accepté d'être président du jury, **Pierre SPIERER** et **Giacomo CAVALLI**, pour avoir accepté d'être rapporteurs de cette thèse et pour leurs remarques constructives, **Isabelle LEGER-SILVESTRE** et **Olivier CUVIER**, pour avoir accepté d'être examinateurs, j'espère que cette première expérience en tant que membre de jury en amènera beaucoup d'autres…

Je remercie aussi **Emmanuel** pour tout ce que tu m'as apporté au cours de ces quatre années tant sur le plan scientifique qu'humain. Je voulais aussi te remercier pour la grande confiance et la liberté que tu m'as accordée. Tu n'as jamais bridé les idées délirantes de manips de la petite étudiante que j'ai été …. Merci aussi à **Caroline**, pour ton optimisme et ta patience, thank you for everything, and I promise: I'll not forget to put the « s » at the third person singular! ; **Guillaume**, merci pour ta présence et les nombreuses discussions ; **Cécile**, je ne te remercierais jamais assez pour ta gaieté et les discussions que nous avons pu avoir sur tous les sujets, plus ou moins sérieux !!! ; **Magali**, qui je l'espère va reprendre le flambeau ; **Sonia**, merci pour ton soutien, Merci aussi à toutes les anciennes X-women du labo **Dorine** et **Vanessa** pour leur bonne humeur permanente… mais à petit **Yan** venu nous amener un peu de son soleil tahitien pendant quelques mois….

Je tenais également à remercier tous les membres du LBME :

Les ARNistes du deuxième étage : pendant quelques années nous nous sommes côtoyés et cette expérience a été pour moi très enrichissante, les portes de vos bureaux ont toujours été ouvertes à mes questions de novices sur le joyeux monde des ARN !

- les Gleizes…. Pendant longtemps j'ai largement abusé de vos moyens techniques et de vos connaissances scientifiques (encore merci pour les fours). Merci à **Nicole**, qui durant mes années sur les bancs de la fac m'a fait découvrir la biologie cellulaire ; **Pierre-Emmanuel** et **Isabelle** grâce à qui j'ai pu faire mes premières armes dans l'enseignement, merci de votre confiance ; **Jacqueline** pour sa gentillesse et aussi pour les cafés du matin ; **Valérie** pour son savoir faire, sa disponibilité et sa gentillesse, **Guillaume**, **Marlène**, **Almass-Houd, Nacer** et **Mme Dupont**. Enfin un grand merci à **Coralie C.** (hic !, elle vient de boire un café) pour sa gentillesse, son humour et sa passion de la coinche !!

- les Ferrer: merci à **Michou** pour ton soutien et tes conseils, **Yves** pour ta disponibilité et aussi ta présence rassurante les WE au labo, **Régine** toujours pleine d'entrain et souriante bref un rayon de soleil quotidien, **Jean-Paul, Anthony**, **Simon**, **Christophe** et **Manu**. Un merci tout particulier à **Coralie HA**…. je te souhaite plein de bonnes choses, ne change surtout rien, et surtout pas tes remontées de bretelles qui m'ont été très utiles.

- les Cavaillé: **Jérôme**, merci pour les discussions sur les petits ARN toujours très enrichissantes et empreintes de passions, **Marie-Hélène** et **Nicole** pour votre aide lors de mes débuts sur les petits ARNs, **Marie-Line**, **Patrice**, **Elène** et **Hervé.**

- les Kiss : merci à **Tamàs**, **Béata**, pour ta gentillesse, **Patricia**, pour les concours de talons dans les couloirs du deuxième et aussi pour avoir été ma coéquipière au volley, bon courage à NY, **Isabelle G**., **Elo**, **Elé**, **grand Sylvain**, et **Arnold**.

Les chromatiniens du RDC : merci pour tous les échanges tant scientifiques qu'humains.

- les D.T. : **Didier**, **Cathy C.**, **Martine**, **Marie**, **Gaëlle**, **Fabrice**, **Sandrine** et **Julie**.

- les H.R.F. : **Hélène**, **Kerstin**, **Françoise**, **Lolotte** pour sa bonne humeur qui traversait souvent le couloir, **'ti Sylvain**, **Lola** et les anciennes **Mathilde** et **Isa K.**, merci pour tous les moments que nous avons passé ensemble depuis le DEUG, bon courage pour la suite.

- les M.G. : **Micha**, pour les « Pas de nouvelle ? bonne nouvelle… », **Kamila**, **Christophe**, et **Nathalie**.

Un grand merci également à ceux sans qui ça marcherait beaucoup moins bien s'ils n'étaient pas là :

- **Safi** et **Cathy R.** merci pour votre disponibilité et pour votre compétence…. et les cafés….

- **David** pour m'avoir expliqué comment on colore « comme il faut » des images sous photoshop…. Euh, désolée j'ai déjà oublié.

- Merci aux membres de l'atelier (pour toutes les réparations et aussi le « 20 minutes » obligatoire le matin au café pour connaître l'horoscope) et du service informatique qui m'ont souvent sortie d'un mauvais pas à cause de mon maudit PC.

Je tenais aussi à remercier le LMGM, nos voisins de l'IBCG, et plus particulièrement **Cathy G.** la motarde, et **Noëlle**….. c'est bon maintenant j'ai plus la pression !!! Merci pour tout ce que tu as fait….

Merci aussi aux squasheurs : les déjà citées Cathy G., Isa K., Noëlle et **Olivier** … pour les parties au cours desquelles j'ai pu démontrer que la variégation était aussi un phénomène humain !

Un grand merci à ceux qui me supportent en dehors du labo …

Lulu… merci pour ta présence, ta gentillesse et ta bonne humeur permanente ! Merci aussi pour les nombreuses parties de belote contre contre **Stéphan** et **Emir** (alias « Mr Patate », pardon Dr. Potatoes) que je remercie d'ailleurs pour leur fair-play et leur absence totale de machisme lors de leurs défaites répétées. **Caro**, alias « blue eyes », pour tout ce qu'on a vécu toutes les deux, le meilleur comme le pire…. Merci pour avoir été là tout au long de ces années. **Marie**, alias « poupoute », merci de m'avoir initié à la dégustation de vin… sauf que je ne sais toujours pas cracher ! merci de ton soutien. **Sèb**, merci pour les soirées œnolo-cinématographiques… ça a toujours été un plaisir même si parfois, le film n'était pas au rendez-vous !!!! Merci aussi à **Riri,** ton comparse de choc pour les sorties ….

Merci à **Jacques** pour ton soutien et ta confiance qui n'a jamais fait défaut….. Merci pour tout ce que tu m'as apporté et m'apporte encore.

Merci à ma famille, **mes parents**, qui m'ont toujours fait confiance et m'ont toujours laissé faire ce que je voulais faire ; sans leur amour et leur soutien je n'en serais pas là, **Tonton Parrain Gâteau**, **Marraine**, **Tatie So**, **mes p'tits cousins** futur Maître Pierrot et Chef Jéré**, mes p'tis frères** Docteur Ben et Manip Jèr, **mes grands-parents** sans oublier mon grand-père Charles a qui je dédie cette thèse …

SOMMAIRE

A. RESUME **12**

B. ABSTRACT **16**

C. ABREVIATIONS **19**

D. INTRODUCTION **22**

I. THEME DE RECHERCHE 23

II. INTRODUCTION 25

1. Historique de la génétique et de la biologie moléculaire 25

a. L'émergence de la génétique formelle 25

b. La nucléine 25

c. Chromosomes, gènes et liaison factorielle 26

d. L'ADN, support de l'information génétique 26

2. La génétique de la drosophile 27

a. Les premières études dans le « laboratoire des mouches » 27

b. Le concept d'hétérochromatine 29

c. Les premières mutations artificielles 32

d. Les effets de position 32

e. La PEV du gène *white* dans l'allèle *white-mottled 4* 33

III. STRUCTURE DE LA CHROMATINE 36

1. Premier niveau de compaction : la fibre de 11 nm 36

a. Les histones 38

b. Le nucléosome 40

2. Deuxième niveau de compaction : la fibre de 30 nm 42

3. Troisième niveau de compaction : la fibre de 300 nm et les chromosomes mitotiques 44

a. Les séquences MAR/SAR 44

b. Le squelette protéique des chromosomes 47

c. Rôle des histones de cœur dans la condensation 51

d. Attachement des SAR/MAR à la matrice nucléaire 52

4. Structures d'ordres supérieurs 53

a. Bandes chromosomiques 54

b. Modèle de repliement des chromosomes 59

IV. L'HETEROCHROMATINE 61
1. Hétérochromatine constitutive et hétérochromatine facultative 61
2. Composition nucléique de l'hétérochromatine 63
a. Les séquences satellites péricentromériques 63
b. Les séquences télomériques 64
c. Les gènes hétérochromatiniens 67
d. Les éléments frontières 68
3. Ligands de l'hétérochromatine 71
a. Les protéines à motifs AT Hook 73
b. Les mimétiques du motif AT Hook 80
c. Les protéines à motifs « chromo-domain » 82
d. Les autres protéines ligands de l'hétérochromatine 90
4. Mécanisme d'assemblage de la structure hétérochromatinienne 95
a. Modèle théorique de l'assemblage 95
b. Initiation de l'assemblage de la structure hétérochromatinienne 97
c. Les complexes de déposition des histones 102
d. Les complexes de remodelage de la chromatine ATP-dépendants 104
e. La méthylation de l'ADN 107
f. Modifications, variants et code des histones 110
5. Hétérochromatine et régulation de l'expression génique 121
a. Régulation par les protéines Polycomb chez la drosophile 123
b. Variégation d'effet de position de l'allèle *white-mottled* w^{m4h} chez la drosophile 124
c. Effet de position chez la levure 125
d. Locus HML/HMR chez la levure 125
e. Compensation de dose et inactivation du chromosome X 127
f. Le locus des gènes ribosomiques 131

E. RESULTATS ET DISCUSSIONS 137

I. DE L'UTILISATION DE NOUVEAUX OUTILS POUR L'ETUDE DE L'HETEROCHROMATINE : « TARGET PRACTICE : AIMING AT SATELLITES REPEATS WITH DNA MINOR GROOVE BINDERS » 139
1. Introduction 139

a. Les oligopyrroles 140
b. Les diamidines 141
2. Objectif de l'étude *142*
3. Résultats *142*
4. Conclusions et discussions *155*
II. DU ROLE DE PROTEINES D1, HP1 ET TOPOÏSOMERASE II DANS LA STRUCTURE HETEROCHROMATINIENNE : « DISPLACEMENT OF D1, HP1 AND TOPOISOMERASE II FROM SATELLITE HETEROCHROMATIN BY A SPECIFIC POLYAMIDE » 158
1. Introduction *158*
a. Les polyamides P9 et P31 158
b. La protéine chromosomale non-histone D1 159
2. Objectifs de l'étude *160*
3. Résultats *161*
4. Conclusions et discussions *174*
III. DU ROLE DE LA PROTEINE TOPOÏSOMERASE II DANS L'ASSEMBLAGE DE L'HETEROCHROMATINE : « TOPOISOMERASE II AND MINOR GROOVE BINDERS ACT COORDINATELY TO FORM SPECIALIZED CHROMATIN STRUCTURES » 178
1. Introduction *178*
2. Objectifs de l'étude *179*
3. Résultats *180*
4. Conclusions et discussions *219*
IV. DE L'IMPORTANCE DES SEQUENCES SATELLITES ET DES ELEMENTS FRONTIERES DANS LA VARIEGATION D'EFFET DE POSITION « A MOLECULAR BASIS FOR VARIEGATING POSITION EFFECTS » 225
1. Introduction *225*
2. Objectifs de l'étude *226*
3. Résultats *228*
a. Effet du traitement par le polyamide P9 sur la localisation des protéines D1 et HP1 dans la lignée *white-mottled* 228
b. Localisation de séquences « marqueurs » à proximité des points de cassures dans la lignée *white-mottled* 232
c. L'élément frontière BE28 est déplacé par l'inversion chromosomique *white-mottled* 234

d. Effet de la perte de l'élément frontière sur la transcription des gènes ribosomiques 236
e. Rôles des petits ARN issus des séquences SAT III 238
4. Conclusions et discussions 242

F. CONCLUSIONS ET PERSPECTIVES 248

I. OBJECTIFS DE CETTE ETUDE 249
II. LES LIGANDS DU PETIT SILLON DE L'ADN : UN NOUVEL OUTIL D'ETUDE DE L'HETEROCHROMATINE 249
III. LA PROTEINE D1 : PLATEFORME DE RECRUTEMENT DE PROTEINES DE L'HETEROCHROMATINE 250
IV. LES SEQUENCES SATELLITES ALPHOÏDES : REGULATEUR DE LA TRANSCRIPTION DES GENES RIBOSOMIQUES ? 252
V. PERSPECTIVES 254

G. MATERIEL ET METHODES 256

I. CULTURE DE LIGNEES DE DROSOPHILES 257
1. Conditions de culture 257
2. Lignée sauvage 257
3. Lignées mutantes 257
II. CULTURE DE CELLULES DE DROSOPHILE 259
III. TECHNIQUES DE BIOLOGIE MOLECULAIRE 259
1. Gel retard 259
2. Footprint à la DNAse I 261
3. Préparation de sondes pour hybridation in situ de fluorescence (FISH) contre l'ADN 262
4. Préparation de sondes pour hybridation in situ de fluorescence (FISH) contre l'ARN 263
5. Northern blot 264
6. Northern blot contre les petits ARN 267
IV. TECHNIQUES DE BIOLOGIE CELLULAIRE 269
***1.** Traitement in vivo de larves par des drogues : « expériences de feeding » 269*
2. Traitement ex vivo des cerveaux perméabilisés par des drogues 270
3. Immunolocalisations sur des cerveaux de larves 270

4. Marquage rapide par les polyamides fluorescents 272
5. Hybridation in situ de fluorescence (FISH) contre l'ADN 272
6. Hybridation in situ de fluorescence (FISH) contre l'ARN 274

H. REFERENCES BIBLIOGRAPHIQUES 275

I. ANNEXES 308

A. RESUME

Les séquences satellites, riches en paires de bases dA•dT, sont les principales séquences d'ADN constituant l'hétérochromatine de la drosophile. La fonction de ces séquences satellites, si elle existe, n'a jamais été clairement établie. Pourtant, des évidences telles que la conservation de leur structure, leur compartimentalisation ou leur présence chez tous les eucaryotes supérieurs, suggèrent une fonction essentielle pour ces séquences. Il en est de même pour les protéines qui les reconnaissent spécifiquement.

Notre modèle d'étude est la mouche du vinaigre *Drosophila melanogaster* et plus précisément son chromosome X. L'hétérochromatine constitue plus de 30% de ce chromosome et est elle-même composée à près de 70% par des séquences satellites de type alphoïde, les séquences satellites de type III (SAT III), similaires aux séquences satellites péricentriques des mammifères. Les protéines de la famille Multi AT Hook, telles que la protéine D1 de drosophile et la protéine HMG-A des mammifères, interagissent avec ces séquences par une reconnaissance sélective du petit sillon de l'ADN. Afin d'étudier à la fois le rôle des séquences satellites et des protéines qui leurs sont spécifiquement associées, nous avons développé de nouvelles approches basées sur l'utilisation de ligands synthétiques du petit sillon de l'ADN, les techniques classiques de biologie moléculaire ou de génétique n'étant pas adaptées à l'étude du rôle biologique de ces interactions qui impliquent des séquences hautement répétées et non codantes.

Ces molécules ont été conçues afin de reconnaître de façon sélective et avec une grande affinité le petit sillon de l'ADN riche en paires de bases dA•dT. L'utilisation de ces composés, des mimétiques du motif AT Hook, nous a permis de mettre en évidence le rôle de la protéine D1 dans le recrutement de protéines impliquées dans la mise en place de la structure hétérochromatinienne du chromosome X. Parmi elles, la protéine HP1 (Heterochromatin Protein 1) est impliquée dans la propagation de modifications épigénétiques caractéristiques de l'hétérochromatine telles que la méthylation de la lysine 9 de l'histone H3. De même, nous avons pu établir un lien entre la localisation de D1 sur les séquences SAT III et le recrutement de la topoïsomérase II, un composant structural et fonctionnel majeur des chromosomes eucaryotes.

Ces études ont été menées en parallèle sur des lignées de drosophiles de types sauvage et *white-mottled* (w^{m4h}). Cette dernière constitue un système modèle classique pour l'étude de l'hétérochromatine en ce que l'expression du gène *white*, requise pour la coloration normale

des yeux de la mouche, y est soumise à une variégation d'effets de position (PEV), un phénomène résultant d'une inversion du bras euchromatinien du chromosome X. Cette inversion re-localise le gène *white*, normalement situé dans la partie centromère-distale du bras euchromatinien, à proximité immédiate de l'hétérochromatine péricentromérique du chromosome. Il est habituellement supposé que la PEV de *white* résulte de l'inactivation stochastique du gène *white* dans ce nouveau contexte génomique.

Nous avons mis en évidence l'existence d'un élément frontière, appelé BE28, qui sépare normalement l'hétérochromatine de l'euchromatine sur le chromosome X. Nos travaux ont permis de le localiser avec précision et de démontrer sa perte par inversion à l'autre extrémité du chromosome X dans la lignée *white-mottled*. Nous proposons que la perte de BE28 résulte en une invasion de l'euchromatine avoisinante —dont le gène *white*— par une structure de type hétérochromatinienne, rendant ainsi compte de l'inactivation stochastique de l'allèle w^{m4h} du gène *white*. Cette découverte démontre donc l'existence d'une barrière physique dont le rôle est de confiner l'hétérochromatine à des régions aux bornes précisément définies, une hypothèse attirante qui n'avait cependant jamais été directement démontrée jusqu'à présent dans le cas de l'hétérochromatine péricentromérique.

Enfin, nos travaux tendent à impliquer les séquences SAT III alphoïdes du chromosome X de la drosophile dans la régulation de la transcription des gènes ribosomiques. Chez la drosophile, comme chez la plupart des eucaryotes supérieurs, les loci des gènes ribosomiques (le rDNA) correspondent à la répétition en tandem de l'unité transcriptionnelle codant pour les ARN ribosomiques (rRNA) et sont le plus souvent retrouvés à proximité de l'hétérochromatine péricentromérique et de ses séquences satellites associées. Nous avons démontré que D1 et HP1 entraient en compétition pour la liaison à ces séquences. Nous proposons que la protéine HP1, dont nous avons montré qu'elle était normalement localisée sur le rDNA, joue un rôle direct dans la régulation transcriptionnelle des gènes codant pour le rRNA.

Nos résultats suggèrent l'existence d'un équilibre dynamique d'association/dissociation des protéines HP1 et D1 entre le locus des gènes ribosomiques (rDNA) et les séquences SAT III adjacentes. Un tel équilibre de compétition pourrait constituer la base moléculaire de l'activation/inactivation stochastique d'un sous-ensemble des unités de transcription codant pour les ARN ribosomiques observée chez tous les

eucaryotes. Sur la base de ces observations, nous proposons enfin que la PEV du gène *white* dans la lignée *white-mottled*, un phénomène décrit il y a plus de 75 ans et qui a permis d'identifier les acteurs impliqués dans les grandes voies de régulation épigénétique, n'est rien d'autre que le reflet d'un mécanisme global de mise en silence facultative et réversible sur lequel repose la régulation fine de l'expression des ARN ribosomiques au cours de la croissance et de la différenciation cellulaire.

B. ABSTRACT

dA•dT-rich satellite sequences constitute the bulk of heterochromatin in *Drosophila melanogaster*. The function of these sequences has never been established. However, their conserved structure, their compartmentalization and their presence in all higher eukaryotes, suggest an essential function for these sequences and for the proteins that preferentially bind them. Our model is the fruit fly *Drosophila melanogaster* and, more specifically, its X chromosome. More than 30% of the X consists of heterochromatin, of which alpha-like satellite III (SAT III) sequences, similar to those found in mammals, represent about seventy percent. Multi-AT Hook proteins, like the *Drosophila* D1 protein and the mammalian HMG-A proteins, interact with these sequences by recognizing the minor groove of dA•dT-rich DNA. In order to study the role of satellite repeats and of the proteins that recognize them, we developed new approaches using synthetic ligands specific for the DNA minor groove, as classical molecular biology and genetic approaches are not applicable to the analysis of these highly repeated non-coding sequences.

These molecules were designed to recognize selectively and with a remarkably high affinity the minor groove of dA•dT-rich DNA. Using these compounds, which mimic the structure and properties of the AT Hook motif, we showed that the D1 protein recruits proteins implicated in heterochromatin formation on the X chromosome. Among those, the HP1 protein (heterochromatin protein 1) has been implicated in the expansion of epigenetic marks characteristic of heterochromatin, such as histone H3 lysine 9 methylation. Furthermore, we established a link between the SAT III localization of D1 and the recruitment of topoisomerase II, a major structural and functional component of eukaryotic chromosomes. These studies were performed both in wild-type and *white-mottled* (w^{m4h}) strains. The latter is a model system for the study of heterochromatin because expression of the *white* gene, necessary for the normal pigmentation of fly eyes, is subject to Position-Effect Variegation (PEV) as a result of a chromosomal inversion on the X. In the w^{m4h} strain, *white* is inverted from its normal position on the centromere-distal euchromatic arm to X-chromosome pericentromeric heterochromatin. PEV is thought to result from the stochastic inactivation of *white* in this new genomic context.

We demonstrate the existence of a boundary element, called BE28 that normally separates heterochromatin from euchromatin on the X chromosome. Our studies localize BE28 precisely and show that it is inverted to the other side of the X chromosome in *white-*

mottled flies. We propose that loss of BE28 results in the invasion of neighbouring euchromatin —including the *white* gene— by a heterochromatin-like structure, thus explaining the stochastic inactivation of the w^{m4h} allele of the *white* gene. This discovery confirms the existence of a long-suspected physical boundary, whose role is to confine heterochromatin to discrete regions, an appealing hypothesis that, however, had not been previously demonstrated for pericentromeric heterochromatin.

Our studies finally implicate X-chromosome alpha-like SAT III sequences in the transcriptional regulation of ribosomal DNA arrays. In *Drosophila*, as in most eukaryotes, ribosomal DNA (rDNA) arrays correspond to tandemly repeated copies of the transcriptional unit encoding ribosomal RNAs (rRNA) and are usually found next to pericentromeric heterochromatin and its associated satellite repeats. We demonstrate that D1 and HP1 compete for binding to these sequences. We propose that HP1, which we have shown to localize within rDNA repeats, is required for the transcriptional regulation of rRNA genes.

Our results suggest a dynamic equilibrium of association/dissociation of D1 and HP1 between the adjacent rDNA and SAT III arrays. This competitive equilibrium may constitute the molecular basis for the stochastic activation/inactivation of a subset of rRNA transcriptional units observed in all eukaryotes. Based on these observations, we propose that *white-mottled* PEV, a phenomenon described more than 75 years ago and that has been used to identify the factors implicated in major epigenetic regulatory pathways, simply reflects the global mechanism of reversible and facultative silencing that provides for the finely tuned regulation of ribosomal RNAs as a function of cell growth and differentiation.

C. ABREVIATIONS

ADN	acide désoxyribonucléique
APC	Anaphase Promoting Complex
ARN	acide ribonucléique
ATP	adénosine triphosphate
BD	Bromo-Domain
BE28	Boundary Element 28
BEAF32	Boundary Element Associated Factor 32
BX-C	Bithorax Complex
C-terminal	extrémité carboxy-terminale
CD	Chromo-domain
CPN	Complexe de Pore Nucléaire
CSD	Chromo-Shadow-Domain
D1ΔE	forme de la protéine D1 délétée du domaine carboxy-terminal
DCC	Dosage Compensation Complex
E(Var)	Enhancer of Variegation
E(Z)	Enhancer of Zeste
FISH	Fluorescence *In Situ* Hybridization
HAT	Histone Acétyl-Transférase
HDAC	Histone Désacétylase
HMG	High Mobility Group protein
HML	HoMothallic Left
HMR	HoMothallic Right
HMT	Histone Methyl Transférase
HP1	Heterochromatin Protein 1
LINE	Long Interspersed Repeated Sequence
MATH	Multi AT Hook
MAR	Matrix Associated Region

N-terminal	extrémité amino-terminale
Pc-G	Polycomb Group
PEV	Position-Effect Variegation
Pol I	ARN Polymérase I
Pol II	ARN Polymérase II
Pol III	ARN Polymérase III
rasiRNA	repeat associated small interfering RNA
rDNA	locus des gènes ribosomiques
RISC	RNA Induced Silencing Complex
RITS	RNA-induced Initiation of Transcriptional Silencing
RNAi	machinerie d'interférence à l'ARN
rRNA	ARN ribosomiques
SAR	Scaffold Associated Region
SAT I	séquences satellites de type I
SAT III	séquences satellites de type III
SAT V	séquences satellites de type V
shRNA	small heterochromatic RNA
SINE	Short Interspersed Repeated Sequence
SIR	Silent Information Regulator
siRNA	small interfering RNA
SMC	Structural Maintenance of Chromosomes proteins
snoARN	Small Nucleolar RNA
SSDNA	Salmon Sperm DNA
SU(Hw)	Suppressor of Hairy Wing
SU(Var)	Suppressor of Variegation
SWI/SNF	Switch/sucrose non fermenting

D. INTRODUCTION

I. Thème de recherche

L'objectif de ce travail de thèse était double. Dans un premier temps, il a consisté à valider l'utilisation de molécules capables de lier le petit sillon de l'ADN en tant qu'outils d'études de l'hétérochromatine. La plupart des facteurs régulant la transcription interagissent avec l'ADN en fixant le grand sillon de la double hélice, plus large et moins profond que le petit sillon. La validation de ces techniques pourrait déboucher sur des utilisations à la fois en recherche fondamentale et en recherche appliquée.

Outre le fait de représenter une avancée considérable dans la recherche sur la structure et la fonction de l'hétérochromatine, ces nouvelles techniques pourraient permettre la mise en place de nouvelles générations de thérapies, à la fois dans le domaine de la cancérologie mais aussi dans la parasitologie. La majorité des molécules synthétiques utilisées dans cette étude sont des dérivés structuraux d'antibiotiques naturels, dont les organismes producteurs se servent pour attaquer leurs prédateurs. De nombreuses études ont pu démontrer que ces molécules, naturelles ou synthétiques, interféraient majoritairement avec la topoïsomérase II. Actuellement, cette protéine représente une cible majeure des thérapies anticancéreuses et antiparasitaires. La caractérisation des mécanismes d'actions de ces composés revêt donc un intérêt tout particulier.

Dans un deuxième temps, nous nous sommes intéressés au mécanisme d'assemblage de la structure hétérochromatinienne sur des séquences particulières : les séquences satellites de type III (SAT III) de *Drosophila melanogaster*, la mouche du vinaigre. Cet organisme représente depuis longtemps un modèle de choix pour l'étude de l'hétérochromatine. Les premières études cytologiques identifiant la composante hétérochromatinienne ont été menées chez la drosophile et de même, les premières études sur la structure de cette partie de la chromatine ont été rendues possibles grâce à la genèse de lignées mutantes. Les travaux antérieurs du groupe avaient identifié un candidat potentiel, la protéine D1 (Aulner et al., 2002; Monod et al., 2002), pour le rôle de ciblage de la machinerie de formation de l'hétérochromatine au niveau des séquences SAT III.

Après validation des techniques permettant l'utilisation de molécules ligands du petit sillon de l'ADN, je me suis donc intéressée au rôle de la protéine D1 dans l'initiation de l'assemblage d'une structure hétérochromatinienne.

Dans cette introduction, je m'attacherai donc dans un premier temps à évoquer l'historique de l'hétérochromatine, des fondements de la génétique à la caractérisation de cette structure. Dans un deuxième temps, je décrirai les mécanismes de repliement mis en jeu dans le passage de la simple double hélice d'ADN à la structure de plus haut ordre : le chromosome mitotique. Enfin, je terminerai par l'analyse des différentes composantes, nucléiques et protéiques, impliquées dans la formation de l'hétérochromatine.

II. Introduction

1. Historique de la génétique et de la biologie moléculaire

a. L'émergence de la génétique formelle

La génétique a pris naissance avec les travaux du moine Gregor Mendel. C'est dans une communication de 1865 qu'il met en évidence la transmission héréditaire des caractères grâce à l'utilisation du pois Pisum (Mendel, 1865). Jusqu'à cette date, les scientifiques pensaient que seuls les spermatozoïdes ou les grains de pollens pouvaient générer des embryons. De ces expériences, G. Mendel tirera trois lois qui seront publiées en 1866 (Mendel, 1866).

Suite à ces communications, les travaux de G. Mendel retombèrent dans l'oubli. Pourtant, il ne cessa pas ses recherches, en étendant celles-ci à de nouvelles espèces végétales. Les théories de G. Mendel réapparurent en 1900 grâce aux travaux de Hugo De Vries, le père de la théorie des mutations (De Vries, 1900). C'est pour la première fois, en 1905, que le mot « génétique » est utilisé par William Bateson (Bateson et al., 1905).

La génétique est donc née, quarante ans après les travaux de G. Mendel, mais il faudra attendre la découverte de l'ADN pour que toutes ses théories soient validées.

b. La nucléine

Voilà plus d'un siècle que la chromatine a été découverte, grâce aux expériences de Friedrich Miescher, biochimiste suisse. En 1869, F. Miescher met en évidence ce qu'il appellera la « nucléine » suite au traitement de cellules lymphocytaires par la pepsine, une protéase stomacale (Miescher, 1871). Des études complémentaires lui permettront de mettre en évidence deux composantes principales à cette nucléine : une partie basique de type protéique, et une partie acide comportant un sucre, le désoxyribose, et du phosphate. Il nommera cette dernière « l'acide désoxyribonucléique » ou ADN.

La composante protéique, caractérisée en 1884 par Albrecht Kossel, sera nommée « histone » (Kossel, 1884).

Plus tard, il sera mis en évidence que l'ADN est un assemblage de quatre constituants distincts : quatre bases azotées appelées adénine, guanine, thymine et cytosine. Ces découvertes ne sucitèrent pas l'attention, aucun rôle ne pouvant être attribué à l'ADN au sein d'une cellule.

c. Chromosomes, gènes et liaison factorielle

Depuis 1880, il était admis que le noyau de la cellule était le siège de l'hérédité. Les travaux de Wilhelm Waldeyer, furent les premiers à décrire les « chromosomes », ou littéralement « corps colorés » (du grec « khroma » couleur et « sôma » corps), structures correspondant à des sortes de bâtonnets présents dans les noyaux et visibles au moment de la division cellulaire (Waldeyer, 1888). Par la suite, les expériences de Walter Sutton fonderont la théorie chromosomique de l'hérédité (Sutton, 1903). Il attirera l'attention sur le lien existant entre la réduction du nombre de chromosomes lors de la méiose et la séparation des caractères alternatifs, ou « allèles », qui doit être réalisée lors de la production des gamètes, ainsi que le postulait G. Mendel. Ainsi, il propose que les chromosomes soient porteurs des gènes.

En étudiant pour la première fois la mouche du vinaigre, ou *Drosophila melanogaster*, en tant que modèle génétique, Thomas H. Morgan met en évidence que les gènes sont organisés en série linéaire le long du chromosome (Morgan et al., 1915). De même, il constate que la liaison entre ces gènes est incomplète et il découvre ainsi le phénomène de recombinaison qu'il associe rapidement à l'appariement des chromosomes observé lors de la méiose. C'est en l'honneur de T. Morgan que l'on donnera son nom à l'unité mesurant la distance entre loci : le centiMorgan.

d. L'ADN, support de l'information génétique

En 1944, alors que la communauté scientifique restait persuadée que le caractère héréditaire de l'information génétique était transmis par l'intermédiaire des protéines, Oswald Avery démontre, chez les bactéries, que le stockage et la transmission de cette information se fait grâce à un support chimique, déjà décrit quelques années auparavant par F. Miescher : l'ADN (Avery et al., 1944). Il confirme ainsi les résultats similaires obtenus en 1928 par Fred

Griffith qui postulait l'existence d'un facteur transformant capable de conférer de manière héréditaire de nouvelles propriétés génétiques (Griffith, 1928).

Une grande question subsistait alors : quelle est la structure exacte de l'ADN ? Grâce à l'utilisation de nouvelles techniques, comme par exemple la chromatographie, Erwin Chargaff démontre que l'ADN est constitué à proportions égales de thymine (T) et adénine (A), et de guanine (G), cytosine (C) (Chargaff, 1950). Il découlera de cette découverte que l'ADN n'est pas un mélange aléatoire des quatre bases A, T, G et C.

Une autre technique, la cristallographie par diffraction aux rayons X, sera utilisée dans le même but, à la fois par Rosalind Franklin en Angleterre, et par Linus Pauling aux Etats-Unis. Les résultats présentés par ces deux équipes comportent des défauts et sont parfois en contradiction (Franklin and Gosling, 1953; Pauling and Corey, 1953). Ce sont finalement James Watson, Francis Crick et Maurice Wilkins qui établiront la structure tridimensionnelle de l'ADN (Watson and Crick, 1953) : c'est une double hélice formée de deux brins chacun composé de groupements phosphates et de sucres, les bases azotées A, T, G et C sont liées sur les sucres des deux brins, une cytosine faisant toujours face à une guanine, et pareillement pour les adénines et les thymines. Cette découverte confirme donc les résultats de E. Chargaff et leur apporte une explication moléculaire.

2. *La génétique de la drosophile*

a. Les premières études dans le « laboratoire des mouches »

Intéressé par les travaux de H. De Vries sur les mutations, T. H. Morgan se met en quête de mutations de grande ampleur permettant d'expliquer l'évolution et l'apparition de nouvelles espèces. C'est dans cet objectif qu'il choisit pour modèle d'étude *Drosophila melanogaster*. En effet, la drosophile est un petit insecte qui est facile à élever et à observer. Son cycle de reproduction est très court, environ 2 semaines à une température de 25°C, et la descendance est nombreuse à chaque génération (figure 1).

La première mutation qu'il observe est une mouche mâle présentant les yeux blancs, ceux-ci étant normalement rouges (mutation *white*). T. H. Morgan, et son étudiant Calvin Bridges, remarquent que ce caractère n'est présent que chez des mâles, ce qui lui permet de proposer que ce caractère est porté par un chromosome sexuel. Cette assignation entre un

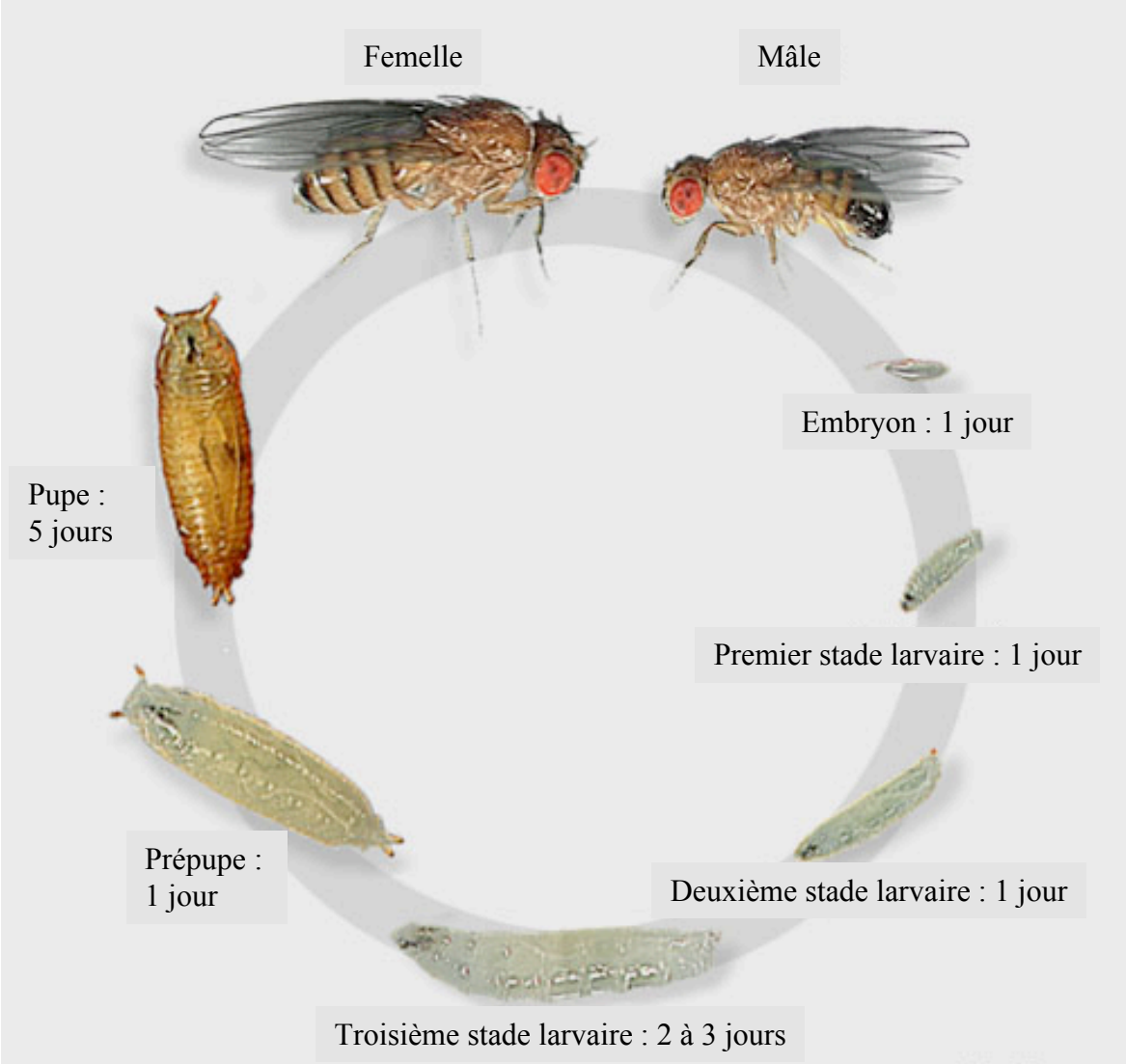


Figure 1: cycle de vie de la drosophile.

Le cycle de vie de *Drosophila melanogaster* à 25 °C dure environ deux semaines. Les femelles pondent environ 400 œufs (embryons). Les œufs, qui ont une longueur d'environ 0,5 millimètres, éclosent après 24 h. La larve à laquelle ils aboutissent croît durant cinq jours en muant deux fois, ces mues définissant les trois stades larvaires. Ensuite, les larves s'encapsulent dans la pupe (phase de prépupaison) et subissent une métamorphose qui dure cinq jours, avant l'éclosion de l'adulte (image tirée du site flymove.uni-muenster.de).

caractère et un chromosome est la première jamais démontrée, et qui plus est par T. H. Morgan, ancien détracteur des théories de G. Mendel (Morgan, 1910).

D'autres observations, menées avec Alfred Sturtevant permettent de mettre en évidence le lien entre le comportement des chromosomes au cours de la reproduction sexuelle et celui des déterminants ou gènes dont les modes de transmissions ont été énoncés par G. Mendel (Morgan, 1911a; Morgan, 1911b).

T. H. Morgan et A. Sturtevant ont étudié la ségrégation de mutations et ainsi ont été à l'origine des toutes premières cartes génétiques de la drosophile, plus précisément celle du chromosome X sur lequel ils localisèrent les gènes *yellow* (corps jaune), *white* (yeux blancs) et *m* (ailes vestigiales) (Sturtevant, 1913). Ces cartes sont complètement superposables aux chromosomes car toutes se rangent en quatre groupes de liaisons correspondant aux quatre chromosomes de la drosophile. Les chromosomes sont donc les supports physiques des gènes.

Les travaux de T. H. Morgan ont ainsi établi la théorie chromosomique de l'hérédité : les chromosomes sont responsables du stockage et de la transmission du patrimoine héréditaire et portent des gènes, à des emplacements fixes.

Suite à ces études, T. H. Morgan et A. Sturtevant ont l'idée de corréler la fréquence de tels évènements à la distance qui sépare les gènes et de proposer que la fréquence de recombinaison exprime une distance entre gènes. L'unité de distances adoptée est le centiMorgan (cM) : un centiMorgan correspond à une fréquence de recombinaison de 1% entre deux marqueurs, soit un évènement de recombinaison pour cent méioses.

b. Le concept d'hétérochromatine

Le terme d'hétérochromatine a été pour la première fois utilisé par Emil Heitz afin de décrire les différences de structure de la chromatine qu'il observait sur des chromosomes mitotiques de plantes (Heitz, 1928). Le terme hétérochromatine, littéralement « l'autre chromatine », se distingue de l'euchromatine, la « vraie chromatine », par une coloration à la Carmine plus marquée. Il avait déjà été décrit que la chromatine était la structure qui permettait de former les chromosomes au cours de la mitose (Boveri, 1904). Des expériences ont permis alors de caractériser des différences morphogénétiques au sein de cette même entité chromatinienne, lors de sa condensation en chromosomes. En effet, l'hétérochromatine semble ne pas participer à la réorganisation qui à lieu pendant la télophase (Heitz, 1935).

Après avoir étudié cette nouvelle structure chez les plantes, E. Heitz s'est intéressé à la drosophile après la publication d'une traduction autorisée du livre de T. H. Morgan sur « les bases physiques de l'hérédité » (Morgan, 1919).

Son travail portait sur cinq espèces de drosophiles qui présentaient la caractéristique d'être toutes reconnaissables en microscopie grâce à des différences aux niveaux de la taille ou du nombre de leurs chromosomes (figure 2). Ses premières observations sur *Drosophila funebris* lui permirent de décrire les chromosomes sexuels comme étant de taille équivalente mais différents en structure chromatinienne : le chromosome Y était totalement « hétéropycnotique » (hétérochromatique) alors que le chromosome X présentait une moitié euchromatique et une autre section centromérique hétérochromatinienne (Heitz, 1933). En montrant qu'il existait une section d'hétérochromatine sur le chromosome X, E. Heitz caractérisa un nouveau type structural pour les chromosomes sexuels, en contraste avec la théorie dite de « type quantitative », selon laquelle les chromosomes X et Y devaient avoir des tailles différentes. Par la suite, il refit les mêmes observations pour *D. melanogaster* et *D. virilis* (Heitz, 1934a), en les complétant par une autre caractéristique commune à ces espèces : les autosomes possèdent aussi une partie hétérochromatinienne, localisée en partie proximale du centromère.

Les observations de E. Heitz étaient en accord avec les théories de l'époque selon lesquelles les chromosomes représentaient le « substrat matériel » des gènes, qui étaient arrangés de manière linéaire sur ces derniers, selon T. H. Morgan.

E. Heitz émet alors la théorie selon laquelle l'euchromatine est génétiquement active, alors que l'hétérochromatine est génétiquement passive. Les chromosomes hétérochromatiniens, ou les parties hétérochromatiniennes, ne contiennent pas de gènes ou alors des gènes « passifs » (Heitz, 1929; Heitz, 1932). Les résultats de T. H. Morgan l'amenèrent à revoir ses hypothèses sur les gènes de l'hétérochromatine. Ainsi, il postula que « la densité de gènes sur un chromosome était reliée à la différentiation longitudinale entre euchromatine et hétérochromatine, et que des gènes présents dans l'hétérochromatine pouvaient jouer un rôle dans le développement » (Heitz, 1934a). Il fut donc un des premiers à évoquer la possibilité de gènes hétérochromatiniens actifs, comme les gènes *rolled* et *light*.

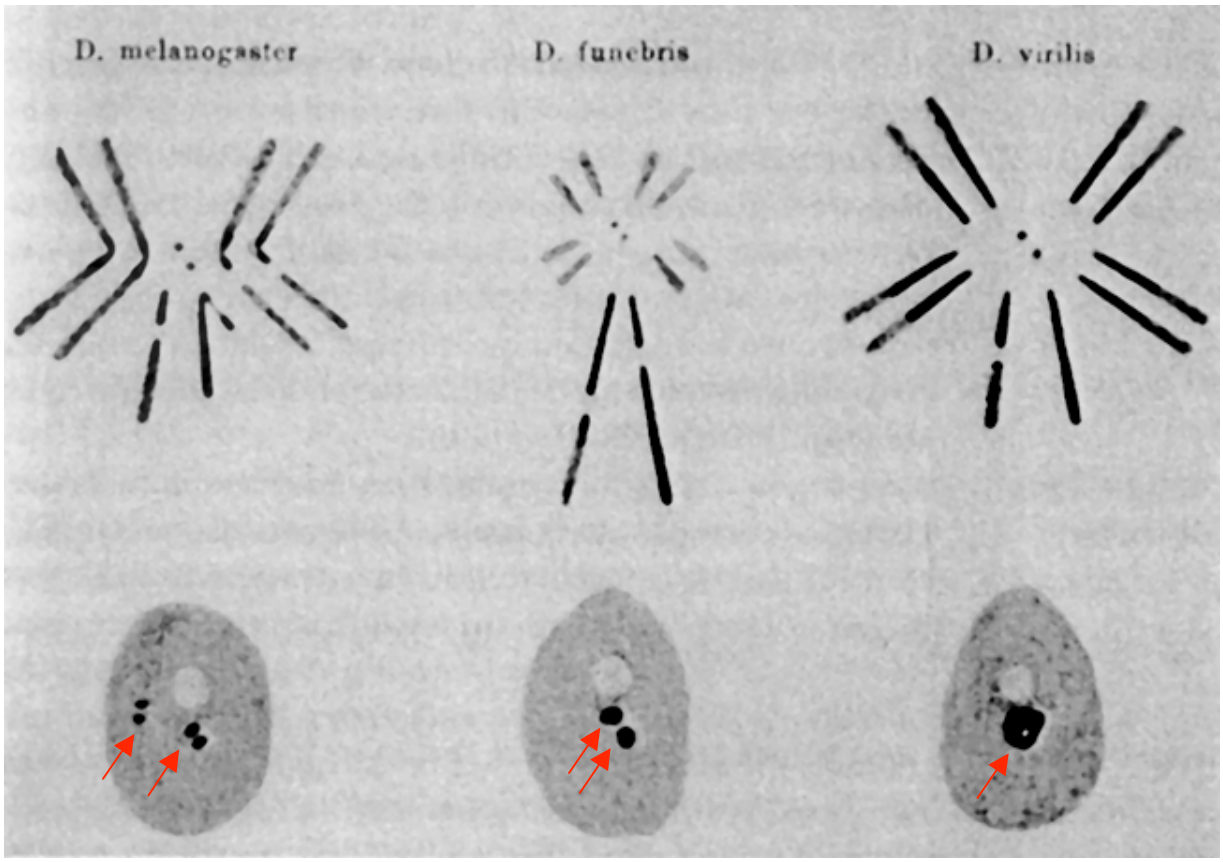


Figure 2: différents types de chromosomes de drosophile observés par Emil HEITZ.

Cette représentation montre trois des cinq espèces étudiées par Emil HEITZ (*Heitz, 1928*). Sur chaque plaque métaphasique, il est possible d'observer que le nombre de chromosomes est différent entre ces trois espèces. De même, sur les noyaux interphasiques correspondants, l'hétérochromatine (flèche rouge) forme une structure différente, qui semble spécifique de l'espèce. Il définit l'hétérochromatine comme la fraction « allocyclique » de chromosome qui ne se décondense pas en interphase.

c. Les premières mutations artificielles

C'est à l'arrivée de Hermann Muller dans le laboratoire de T. H. Morgan que les premières expériences de mutations artificielles furent mises en œuvre grâce à l'utilisation des rayons X, dès 1927.

Il démontre que l'irradiation aux rayons X augmente d'un facteur 100 le nombre de mutations chez la drosophile (Muller, 1927). Cette technique de mutation n'est pas spécifique puisqu'il n'est pas possible de cibler un gène choisi à l'avance. L'augmentation de la fréquence de mutations va cependant s'avérer d'un intérêt considérable puisque le nombre de mutations décrites – tout du moins celles produisant un phénotype aisément détectable – va croître, permettant ainsi l'établissement de cartes génétiques de plus en plus précises. Les techniques de H. Muller seront reprises par Max Delbrück afin de corréler l'intensité de rayonnement X à la fréquence de mutation.

Grâce à ces observations, H. Muller avance une première estimation du nombre de gènes chez la drosophile dès 1929 : il prédit qu'il y aurait de 1 400 à 1 800 gènes, une estimation qui reste toutefois bien inférieure à celle des dernières données issues du séquençage qui suggèrent un nombre de gènes proche de 13 400.

C'est à lui que l'on doit une des mutations de drosophile les plus étudiées : l'inversion chromosomique *white-mottled 4* aussi notée In(1)wm4 ou w^{m4} (Muller, 1930), qui présente une variégation d'effet de position affectant la régulation du gène *white*.

d. Les effets de position

Un effet de position désigne une dérégulation de l'expression d'un gène sous l'influence d'éléments régulateurs du génome.

L'effet de position a été décrit pour la première fois dans une souche de drosophile portant une mutation appelée *Bar eye* (May, 1917). Cette mutation génère un phénotype correspondant à une réduction de la taille des yeux par rapport à la normale. Elle a été étudiée principalement par Charles Zeleny et ses étudiants dès 1919 (Zeleny, 1919). Ils ont pu mettre en évidence une liaison avec le sexe, seules les femelles portaient la mutation, mais aussi une autre forme de phénotype plus aggravé qu'ils ont appelé *double-Bar*. A. Sturtevant et T. H. Morgan étudièrent cette mutation et mirent en évidence un défaut dans les recombinaisons

méiotiques (Sturtevant and Morgan, 1923). De nombreuses études cytologiques, menées par la suite, montrèrent l'implication d'une duplication et d'une délétion d'une section de chromosome dans les phénotypes associés, respectivement *double-Bar* et *Bar* (Bridges, 1936).

D'autres types d'effets de position furent décrits et caractérisés par la suite et leurs analyses amenèrent rapidement l'hypothèse de l'existence de deux types d'effet de position essentiellement différents (Lewis, 1950). Ces deux types sont appelés effet de position de « type S », c'est-à-dire stable, et un dit « type V », c'est-à-dire bigarré (Position Effect Variegation ou PEV). Les effets de position de type S sont dépendants d'un remaniement chromosomique, généralement consécutif à une recombinaison défectueuse, générant la juxtaposition d'un gène à des séquences régulatrices d'un autre gène. La mutation *Bar* fait partie des effets de position de type S.

Le second type d'effet de position, la PEV, est fondamentalement différent de l'effet de position de type S. En effet, les principaux exemples de PEV proviennent de réarrangements chromosomiques induits par des traitements aux rayons X, comme c'est le cas pour l'allèle w^{m4}, décrit par H. Muller (Muller, 1930). Une inactivation a généralement lieu quand des gènes de l'euchromatine sont relocalisés à proximité de l'hétérochromatine, comme dans l'allèle w^{m4}. Moins souvent, la PEV peut être observée lorsque des gènes de l'hétérochromatine, ou placés à proximité de l'hétérochromatine, sont relocalisés à proximité de l'euchromatine, comme dans l'exemple du gène *cubitus interruptus* (Dubinin and Sidorov, 1934). Des études menées sur ces réarrangements chromosomiques montrent qu'il existe un effet d'« étalement » de la structure nouvellement adjacente au gène déplacé. Les effets de position clairement caractérisés supposent alors l'existence de séquences capables d'assurer l'initiation, la propagation et la terminaison de tels phénomènes de mise en silence (Schultz, 1936).

e. La PEV du gène *white* dans l'allèle *white-mottled 4*

Cette inversion, caractérisée dès 1930 par H. Muller (Muller, 1930), déplace le gène *white*, nécessaire à la couleur rouge normale des yeux de drosophile, à proximité de l'hétérochromatine péricentromérique du chromosome X (Schultz, 1936), alors que ce gène est normalement situé à l'extrémité distale de ce même chromosome (figure 3). Les effets générés sur le gène *white* représentent l'exemple classique de PEV (Muller, 1930). En effet, le

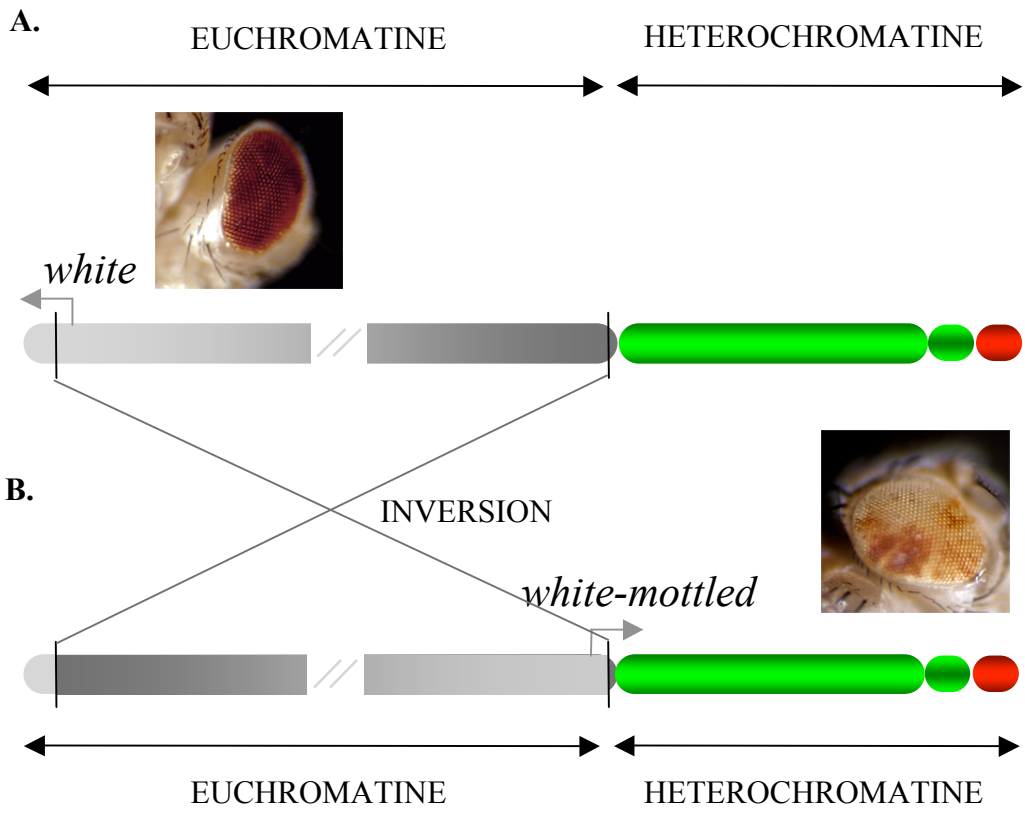


Figure 3: l'inversion white-mottled 4.

A. Schéma d'un chromosome X et photo d'un œil d'une mouche de lignée sauvage (de type Oregon).

B. Schéma d'un chromosome X et photo d'un œil d'une mouche de la lignée *white-mottled 4 h*.

Dans la lignée sauvage (A), le gène *white* qui code un transporteur de pigment essentiel à la coloration rouge normale des yeux (photo A) est localisé à l'extrémité du bras long euchromatinien du chromosome X (en grisé). Dans la lignée *In(1)wm4*, l'inversion chromosomique juxtapose l'allèle de *white* à proximité de l'hétérochromatine péricentrique (en vert). L'inactivation stochastique de *white* dans certaines cellules ou groupes de cellules par l'hétérochromatine conduit à un phénotype bigarré (photo B). L'allèle de *white* ainsi relocalisé dans la lignée *In(1)wm4* a été appelé *white-mottled*, du fait de ce phénotype caractéristique. La lignée wm^{4h} que nous utilisons au laboratoire est une des inversions *white-mottled* de ce type dont le phénotype de PEV est le plus prononcé.

phénotype est directement observable au niveau des yeux qui deviennent, dans les mouches portant l'allèle w^{m4}, bigarrés, c'est-à-dire que certaines ommatidies sont rouges et d'autres plus pâles. Les ommatidies rouges ont un gène *white* actif alors que celles qui sont les plus pâles ont un gène *white* plus ou moins exprimé. Cette répression est héritée de manière clonale, elle traduit une régulation négative stochastique du gène *white* lors du développement et de la formation de l'œil de la drosophile.

Des études complémentaires menées sur cet allèle w^{m4} ont montré que le gène *notch*, très proche du gène *white* sur le chromosome X, était lui aussi soumis à cette PEV (Schultz, 1936). Ces observations vont dans le sens d'une expansion de la structure hétérochromatinienne vers les gènes inversés *white* et *notch*. Elles sont à rapprocher des hypothèses émises par J. Schultz sur l'existence de séquences capables de stopper l'extension d'une telle structure, dans des contextes sauvages. L'inversion w^{m4} semble responsable du déplacement d'une séquence permettant en temps normal de stopper l'expansion de la structure hétérochromatinienne, protégeant ainsi l'euchromatine proximale de toute inactivation (Schultz, 1939).

L'avènement des nouvelles techniques de biochimie, et plus particulièrement les techniques de coloration différentielle des chromosomes présentées dans la partie suivante, ont permis une nouvelle avancée dans l'étude de la structure de la chromatine. Paradoxalement, la compréhension des mécanismes à l'origine de la formation de ces bandes colorées reste incomplète et encore du domaine de l'hypothèse.

D'autre part, les progrès en microscopie électronique et photonique ont permis l'obtention d'images de plus en plus précises autorisant ainsi l'émergence d'hypothèses quant à l'organisation architecturale, topologique et fonctionnelle de la chromatine.

III. Structure de la chromatine

L'évolution des génomes eucaryotes s'est accompagnée de deux événements majeurs : l'augmentation de la quantité d'ADN par cellule entraînant l'émergence de nouvelles fonctions cellulaires, mais aussi le confinement de cet imposant matériel génétique dans un espace restreint : le noyau. La réponse à cette dernière exigence est apportée par la capacité de la molécule d'ADN à se replier sur elle-même, via des mécanismes plus ou moins complexes, tout en autorisant l'accès des séquences nucléotidiques aux machineries moléculaires capables de s'y lier.

Ainsi, chaque molécule d'ADN, correspondant à chaque chromosome, est enroulée et compactée en plusieurs étapes, grâce à l'association de différentes protéines histones et non-histones (Gasser and Laemmli, 1987a) (figure 4).

1. Premier niveau de compaction : la fibre de 11 nm

Le premier niveau de compaction correspond à un arrangement linéaire de la molécule d'ADN autour d'un octamère d'histones, appelé le nucléosome (Germont et al., 1976). Ce premier niveau de compaction a été historiquement mis en évidence par des expériences de biochimie et de microscopie électronique (figure 5).

Des expériences de digestion partielle de la molécule d'ADN avec une nucléase micrococcale (MNase) suivies de migration électrophorétique, ont montré des images en « échelle » avec une périodicité variable selon les espèces et s'échelonnant de 140 paires de bases (pb) à 200 pb (Williamson, 1970).

Par la suite, la régularité de cette structure a été confirmée par des expériences de microscopie électronique. Les observations révélèrent des particules, les nucléosomes, régulièrement espacées sur la molécule d'ADN, donnant ainsi un aspect en « collier de perle » à la chromatine (Olins and Olins, 1974; Oudet et al., 1975).

Ces différents résultats réfutent les hypothèses ayant cours jusqu'alors et qui supposaient que les histones, protéines basiques, recouvraient la molécule d'ADN en formant une sorte de gaine continue.

La première étape de compaction comprend donc l'assemblage de la molécule d'ADN autour d'un octamère d'histones, suivie d'une étape de maturation. Cette étape de maturation

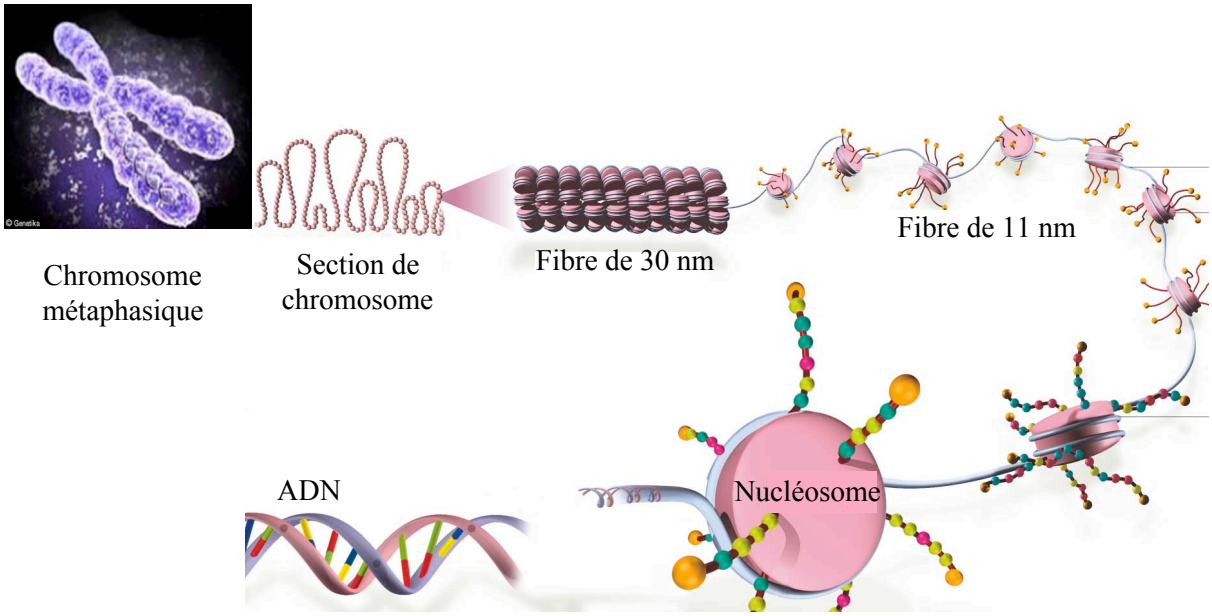
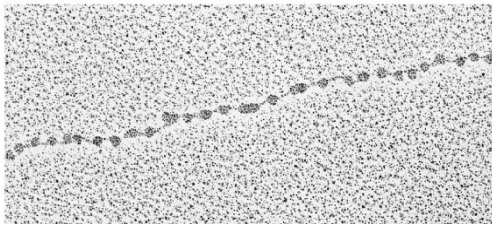


Figure 4: les différents niveaux de compaction de la chromatine.

L'ADN s'enroule autour d'un octamère d'histones pour former le nucléosome, particule fondamentale de la chromatine dont les premiers niveaux de compaction sont la fibre de 11 nm aussi appelée « collier de perles » et la fibre de 30 nm. (Images tirées de Genetika®)

A.



B.

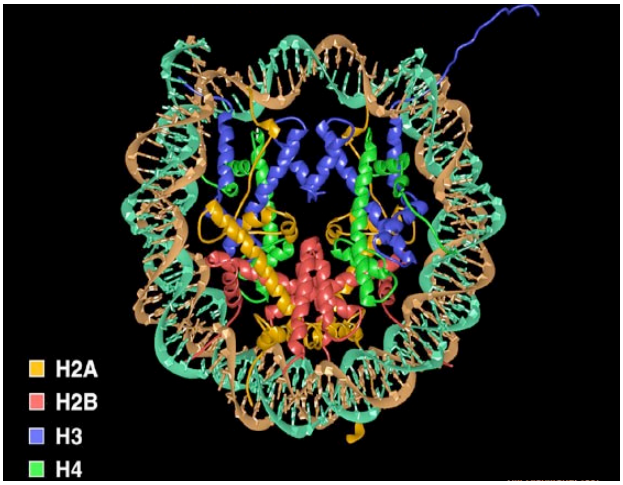


Figure 5: premier niveau de compaction de la chromatine.

A. Fibre de chromatine de 11 nm de diamètre en collier de perle, au microscope électronique (image tirée du site *edoc.hu-berlin.de*).

B. Structure du nucléosome à partir de la diffraction aux rayons X à 2.8 Å : H2A est en jaune, H2B en rouge, H3 en bleu et H4 en vert (d'après Luger *et al*., 1997).

nécessite de l'énergie, fournie par hydrolyse de l'ATP, qui va permettre un espacement régulier des nucléosomes sur l'ADN, formant ainsi le nucléofilament.

a. Les histones

Les histones sont de petites protéines globulaires basiques, très conservées au cours de l'évolution. Deux sortes d'histones peuvent être distinguées : les histones de cœur, qui font partie de l'octamère d'histone du nucléosome, et les histones de liaison.

Les histones de cœur sont au nombre de quatre : les histones H2A, H2B, H3 et H4. Toutes ces protéines sont essentielles pour les organismes eucaryotes. Ces histones ont des caractéristiques structurales communes puisqu'elles présentent toutes quatre une queue amino-terminale (N-terminale), et une partie globulaire centrale hautement structurée et appelée le « histone fold domain ». Celui-ci se caractérise par trois hélices alpha reliées entre elles par deux boucles (Arents and Moudrianakis, 1995). Cette structure particulière favorise les interactions entre les histones et la molécule d'ADN (figure 6).

La partie N-terminale est quant à elle libre de toute structuration remarquable. De plus, elle présente de nombreux acides aminés basiques, accessibles à des complexes protéiques de modifications post-traductionnelles. Ces modifications des histones jouent un rôle prépondérant dans de nombreux processus cellulaires, comme par exemple dans la compaction de la chromatine (Peterson and Laniel, 2004).

Cependant, on peut distinguer des différences dans la composition en acides aminés de ces protéines : les histones H2A et H2B sont plutôt riches en lysines, alors que les histones H3 et H4, sont plutôt composées d'arginines.

Il existe d'autres protéines histones, appelées « variants d'histone », dont quelques acides aminés diffèrent de ceux de la protéine histone parente dite « canonique ». Les variants d'histone sont souvent impliqués dans des processus cellulaires particuliers comme par exemple l'activation transcriptionnelle, la mitose, ou la ségrégation des chromosomes.

Les histones de liaison, dont l'histone H1 est la principale représentante, jouent un rôle dans la formation du chromatosome et la fibre chromatinienne de 30 nm, pour revue

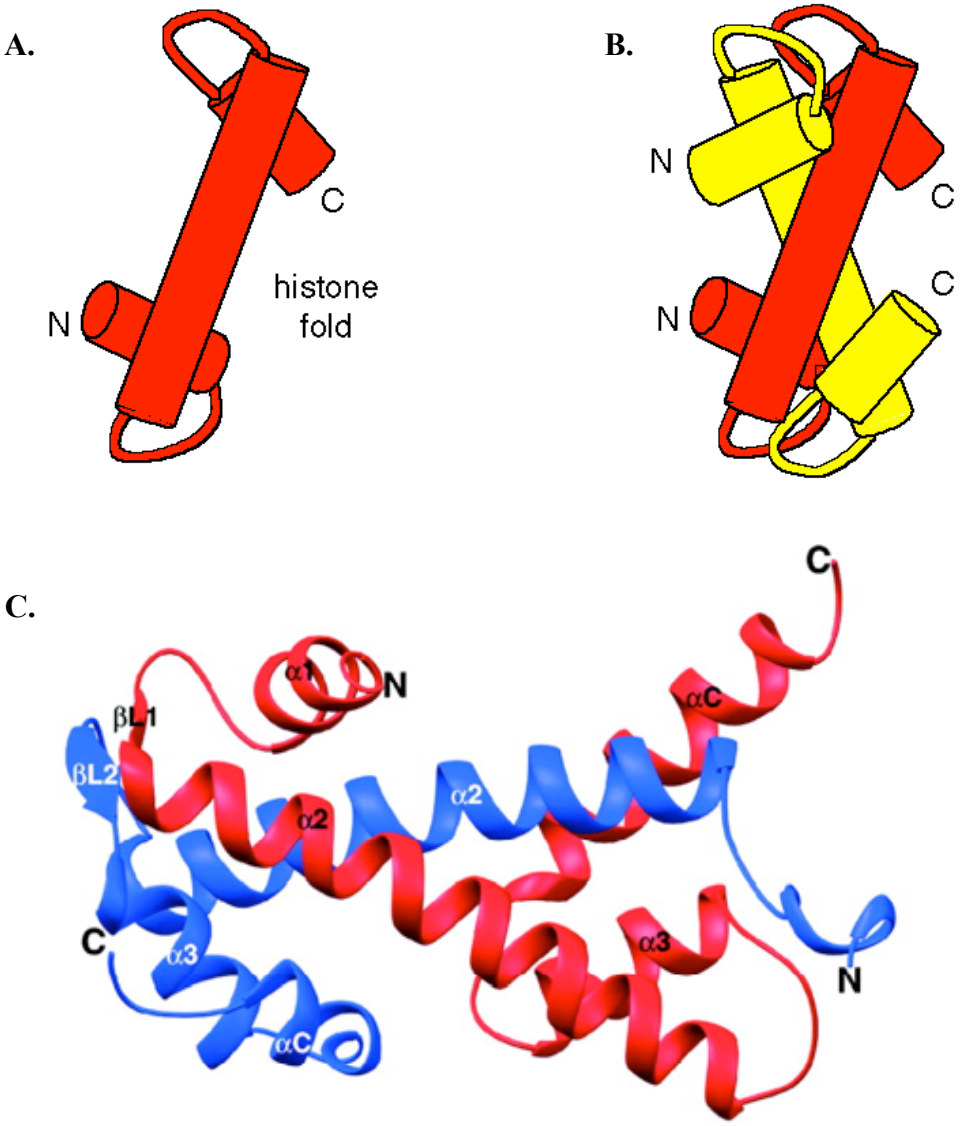


Figure 6 : le domaine histone-fold.

A. Schéma représentant le repliement d'une histone de cœur en domaine *histone-fold*.

B. Schéma représentant l'interaction en « poignée de main » (*handshake*) entre deux histones.

C. Structure cristallographique de deux domaines histone-fold impliqués dans une interaction en poignée de main (d'après Bottomley, 2004).

Ce type de repliement en domaine histone-fold est adopté par la partie structurée des histones de la particule nucléosomale et de l'histone de liaison H1. L'interaction en poignée de main permet l'association des multimères d'histones au sein du nucléosome et est donc impliquée dans la formation de l'octamère d'histone.

(Khochbin, 2001). Elles sont formées de trois domaines protéiques : une queue N-terminale riche en lysines, un domaine central appelé « histone-fold », et une très longue queue carboxy-terminale (C-terminale).

Chez Tetrahymena, il semble que l'histone H1 ne soit pas essentielle puisque sa perte ne perturbe que certains gènes et la division cellulaire n'est pas affectée (Shen and Gorovsky, 1996). Cependant chez l'Homme, il existe de nombreux variants de l'histone H1 dont les séquences primaires sont assez différentes de l'histone modèle. Pour autant, on ne sait pas s'ils ont une fonction particulière dans la chromatine, outre leur fonction de liaison, ou si leur incorporation dépend uniquement de leur lieu et/ou moment de synthèse (Doenecke et al., 1997). Seule la drosophile ne semble pas présenter de variants pour l'histone H1 (Nagel and Grossbach, 2000).

b. Le nucléosome

Le nucléosome est considéré comme l'unité de base de la chromatine (Kornberg and Thomas, 1974). La molécule d'ADN qui s'entoure autour de l'octamère d'histones, appelé la « particule de cœur », prend le nom d'ADN nucléosomal alors que l'ADN reliant un octamère à l'autre est appelé ADN internucléosomal.

L'ADN nucléosomal a une longueur d'environ 147 pb et fait environ 1,7 tours autour de la particule de cœur. Par contre, la longueur de l'ADN internucléosomal est variable, et semble refléter le degré de complexité des organismes.

L'assemblage de la particule de cœur débute par la formation d'un tétramère central comportant les histones H3 et H4 (Kornberg and Thomas, 1974; Roark et al., 1974), qui après fixation sur la molécule d'ADN forme la particule subnucléosomale (figure 7). Viennent ensuite s'adjoindre, de part et d'autre, deux hétérodimères constitués par les histones H2A et H2B (Kelley, 1973). Cette structure octamérique particulière a pu être caractérisée de manière très précise par cristallographie au rayons X ((Luger et al., 1997) et figure 5).

Lors de l'incorporation d'histones nouvellement synthétisées, celles-ci seront spécifiquement modifiées post-traductionnellement sur leur queue N-terminale. Les modifications les plus conservées sont celles de l'histone H4 dont les lysines 5 et 12 sont acétylées (Sobel et al., 1995).

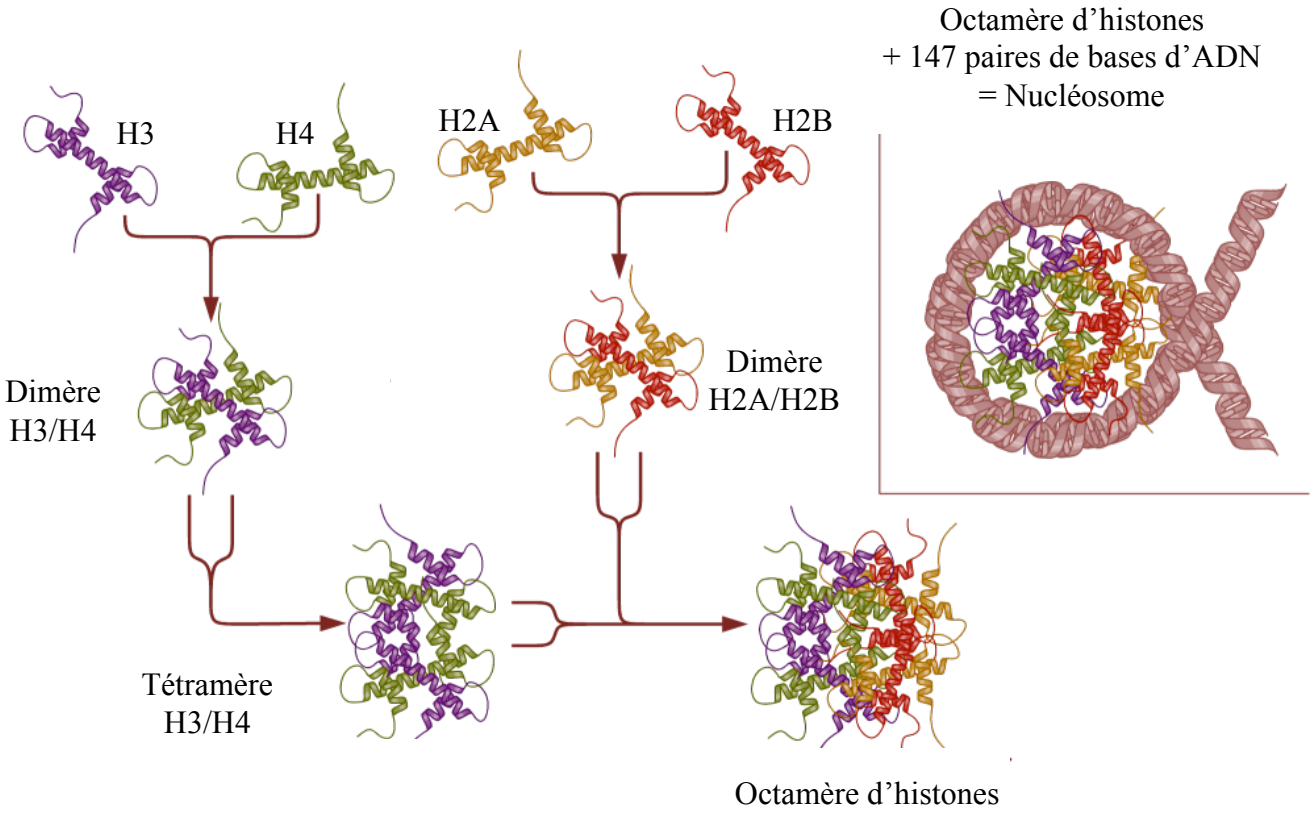


Figure 7 : formation d'un nucléosome.

L'assemblage de la structure nucléosomale se fait en plusieurs étapes. Dans un premier temps, un tétramère d'histone H3 et H4, comportant deux histones H3 et deux histones H4, s'assemble. Ce tétramère est ensuite encadré par deux dimères d'histones H2A et H2AB, formant un octamère d'histones appelé particule de cœur. L'ADN va s'enrouler autour de cette particule et ainsi donner naissance à ce qui est appelé le nucléosome. Seuls les domaines histone-fold des histones sont impliqués dans la formation du nucléosome, les extrémités amino-terminales émanent de cette structure et restent accessibles aux protéines qui les reconnaissent (image tirée du site *www.answers.com/topic/histone***)**.

Les interactions existantes entre les histones, protéines basiques, et les groupements phosphates de l'ADN, chargés négativement, sont de type électrostatiques, ou liaisons hydrogènes. On peut aussi observer des contacts non polaires avec le désoxyribose des bases. L'absence de contact « direct » entre les bases nucléotidiques de l'ADN et les histones expliquerait pourquoi la déposition des nucléosomes est indépendante de la séquence. Cependant, ce postulat est fortement remis en question suite à la mise en évidence de l'implication de séquences d'ADN dans le positionnement des nucléosomes (Young and Carroll, 1983).

Même si dans la plupart des cas les nucléosomes semblent être déposés de manière aléatoire, leur localisation est parfois contrainte par des éléments structuraux : protéines qui lient une séquence nucléotidique particulière, topologie de l'ADN…. Ainsi, les nucléosomes sont souvent positionnés au niveau de promoteurs ou d'éléments régulateurs car ils facilitent l'interaction entre les séquences d'ADN, contrainte par leur engagement dans le nucléosome, et les protéines qui reconnaissent ces séquences (Lu et al., 1995).

2. *Deuxième niveau de compaction : la fibre de 30 nm*

Avant ce repliement, les histones nouvellement incorporées sont désacétylées. La fibre chromatinienne se compacte et forme alors une super hélice de 30 nm de diamètre grâce à l'intervention des histones de liaison H1. Ces dernières peuvent interagir d'une part avec l'ADN internucléosomal et d'autre part avec le cœur du nucléosome pour former le « chromatosome » (figure 8).

A l'heure actuelle, deux principaux modèles de repliement existent : le modèle de type « zig-zag » et le modèle de type « solénoïde » (Van Holde and Zlatanova, 1995; Woodcock et al., 1993).

Le modèle de repliement de type « solénoïde » postule que l'ADN internucléosomal garde la même courbure que celle infligée par son engagement dans le nucléosome. La chaîne de nucléosomes formerait une structure en hélice, l'axe des nucléosomes étant perpendiculaire à l'axe du solénoïde. Le rôle de l'ADN internucléosomal est donc de se replier de telle façon à relier les nucléosomes adjacents dans le solénoïde, selon le modèle simple tour ou double tour, comprenant 6 ou 8 nucléosomes par tour. Cependant, l'aspect régulier de ce solénoïde

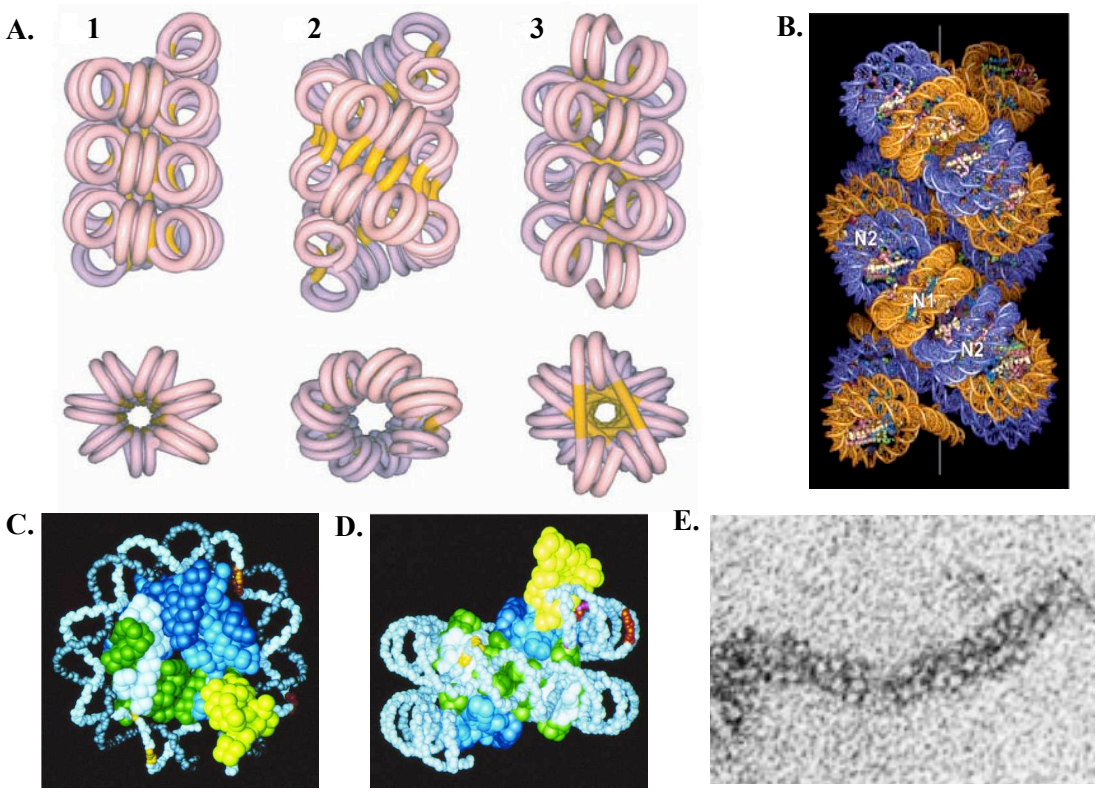


Figure 8: deuxième niveau de compaction de la chromatine.

A. Schéma représentant trois modèles de repliement de l'ADN dans la fibre de chromatine de 30 nm en vue latérale (en haut, l'axe de la fibre est vertical) et selon l'axe de la fibre (en bas).

1 : modèle du simple tour.

2 : modèle du double tour.

3 : modèle du double tour « twisted ».

Le modèle du solénoïde (1), longtemps privilégié ne semble plus rendre compte des plus récentes observations et a été abandonné au profit d'un modèle en double hélice (2 et 3) qui seraient soit surenroulées (2) soit entrelacées (3) selon que l'ADN linker (en jaune) traverse (3) ou non (2) le centre de la fibre (Dorigo *et al.* 2004). L'ADN nucléosomal est représenté en rose.

B. Modèle de repliement de la fibre de 30 nm basé sur les données cristallographiques de la structure d'un tétranucléosome à une résolution de 9Å. (Schalch *et al.* 2005).

C. et D. Structure du chromatosome, proposée par Crane-Robinson *et al.*, 1989. L'histone de liaison H1 est figurée en jaune, les dimères H2A-H2B en bleu clair et bleu foncé, et le tétramère H3-H4 en vert et blanc.

E. Fibre de chromatine de 30 nm en microscopie électronique (image tirée du site *www.csu.edu.au*).

est controversé et il se pourrait que seules des petites régions soient compactées de façon régulière.

Des études complémentaires ont pu déterminer que le modèle le plus probable était celui d'un double tour, qui en fait a donné la prédominance à la théorie de repliement dite en « zig-zag » (Dorigo et al., 2004; Richmond and Davey, 2003). Ce modèle, aussi appelé modèle d'entrecroisement, postule que l'ADN internucléosomal croise la fibre chromatinienne (Woodcock and Dimitrov, 2001). Dans ce modèle, l'ADN internucléosomal adopte une conformation rectiligne entre deux nucléosomes situés sur des cotés opposés de la fibre chromatinienne.

Finalement, plusieurs repliements successifs conduisent à des niveaux d'organisation supérieurs, mais aussi à la formation de domaines spécifiques dans le noyau.

3. *Troisième niveau de compaction : la fibre de 300 nm et les chromosomes mitotiques*

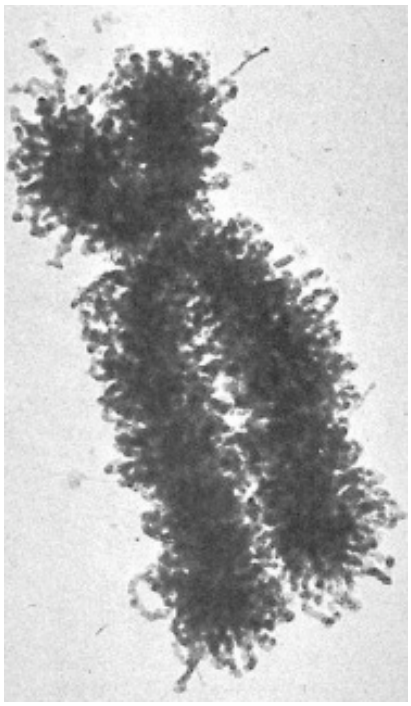
La fibre de 30 nm va former des boucles et des structures secondaires qui vont ainsi générer la fibre de 300 nm (figure 9). Il est supposé que des boucles de chromatine se forment à partir de certaines régions d'ADN qui s'attachent à l'enveloppe nucléaire, ou à un squelette protéique, qui confèrent leur morphologie aux chromosomes métaphasiques. Ces séquences très conservées au cours de l'évolution et riches en paires de bases dA•dT sont appelées séquences MAR (Matrix Associated Region) ou séquences SAR (Scaffold Associated Region).

La condensation des chromosomes nécessite l'action de nombreux acteurs, qui sont à la fois les séquences ADN MAR/SAR mais aussi de très nombreuses protéines.

a. Les séquences MAR/SAR

Les séquences MAR/SAR ont tout d'abord été identifiées chez la drosophile (Mirkovitch et al., 1984), suite à des expériences de préparations de matrices ou échafaudages nucléaires dans des conditions mieux contrôlées que celles utilisées lors d'extractions classiques en présence de 2 M de chlorure de sodium. Dans ces expériences, ces séquences

A.



B.

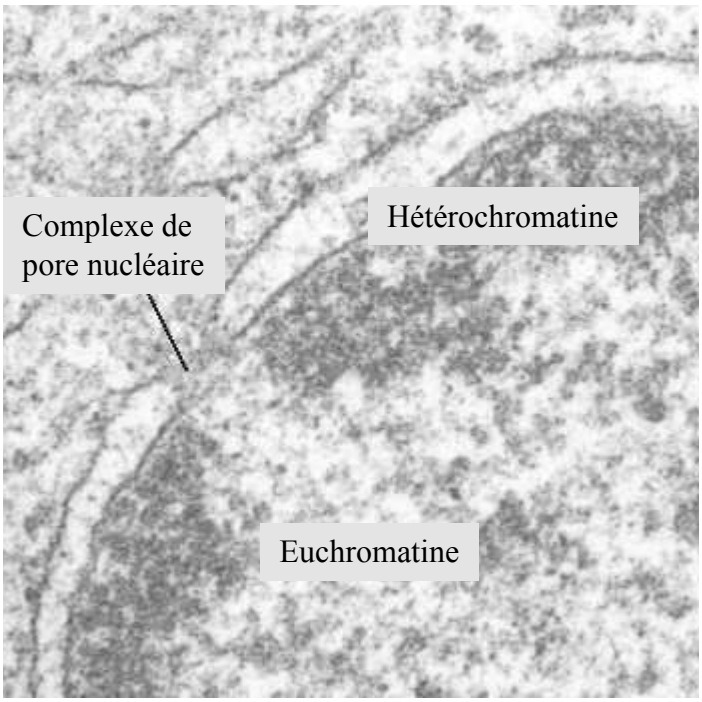


Figure 9: troisième niveau de compaction de la chromatine.

A. Chromosome métaphasique en microscopie électronique (image tirée du site *http://www.roycastle.org*)

B. Domaines nucléaires en microscopie électronique (image tirée du site *http://cellbio.utmb.edu/cellbio/nucleus2.htm*).

L'euchromatine se distingue de l'hétérochromatine par son faible contraste, l'hétérochromatine apparaissant plus dense aux électrons.

étaient retenues dans une fraction insoluble car elles restaient liées, après digestion par des enzymes de restrictions, à la matrice nucléaire. Ces séquences ont ensuite été caractérisées dans d'autres organismes comme l'Homme, le hamster et la levure (Cockerill and Garrard, 1986; Gasser and Laemmli, 1987b; Käs and Chasin, 1987).

Les séquences MAR/SAR sont des séquences de 500 à 1 000 pb majoritairement constituées de paires de bases dA•dT (environ 65%). Dans ces enchaînements, il est à noter la présence de nombreux « A-tracts », c'est-à-dire des enchaînements homopolymériques dA ou dT, espacés de manière irrégulière.

Ces séquences sont à l'origine de la formation de boucles de chromatine dont la taille moyenne est de 85 kb, et pouvant s'échelonner de 5 à 200 kb (Benyajati and Worcel, 1976). En moyenne, chaque boucle de chromatine comprendrait de 180 à 300 nucléosomes (Marsden and Laemmli, 1979).

Une étude exhaustive menée chez la drosophile a permis de mettre en évidence que les séquences MAR/SAR se localiseraient à proximité d'éléments régulateurs, ou à la frontière de domaines de fonctions (Gasser et al., 1986). Pour autant, il est difficile d'appréhender la fonction de ces éléments *in vivo*, même si de nombreux travaux les définissent comme des éléments clés de la structuration des génomes (Hart and Laemmli, 1998; Laemmli et al., 1992).

D'autres études ont aussi permis de démontrer que ce n'était pas à proprement parler la séquence nucléotidique des MAR/SAR qui était importante pour la liaison des facteurs qui la reconnaissent mais plutôt la structure locale de l'ADN. En effet, les « A-tracts » confèrent des particularités à la double hélice : le petit sillon est plus profond et plus étroit, et ces séquences présentent une courbure plus accentuée que la normale.

Des oligopeptides cytotoxiques, comme la distamycine, ont la capacité de se lier à une telle structure (Adachi et al., 1989; Izaurralde et al., 1989; Käs et al., 1989). De même, l'étroitesse et la profondeur du petit sillon sont reconnues spécifiquement par un groupe de protéines possédant des motifs particuliers appelés « AT Hook » : les protéines HMG-A (Zhao et al., 1993). D'autres protéines peuvent se lier aux séquences MAR/SAR comme, par exemple l'histone de liaison H1, pour revue (Bode et al., 2003), qui rentre en compétition avec HMG-A pour la liaison à ces séquences, ou la Topoïsomérase II alpha, qui est la

composante non histone majeure du squelette protéique de l'échafaudage métaphasique des chromosomes.

b. Le squelette protéique des chromosomes

Malgré les efforts mis en œuvre, l'architecture basique des chromosomes mitotiques, de même que le niveau hiérarchique de la structure chromatinienne qu'ils représentent sont peu connus (figure 10 et (Paulson and Laemmli, 1977)). Cette structure est formée de différentes protéines comme le complexe *Condensin*, la Topoïsomérase II, ou les histones dont les queues N-terminales jouent un rôle dans l'assemblage chromosomique.

En lui-même, le squelette protéique des chromosomes est composé de plusieurs protéines, ayant différents rôles au sein de ce complexe. Parmi ces protéines se trouvent les protéines du groupe des SMC (Structural Maintenance of Chromosomes). Ce groupe de protéines participe à la dynamique des chromosomes et est associé à des protéines comme la Topoïsomérase II (Gasser et al., 1986), Barren (Bhat et al., 1996b) ou la Titine (Machado and Andrew, 2000) permettant de réguler leur activité.

Un des complexes de protéines SMC, appelé complexe *Cohesin*, joue un rôle dans la cohésion entre les chromatides sœurs pendant la phase S et ce jusqu'à la transition métaphase/anaphase (Losada and Hirano, 2001). Ce complexe comprend les protéines SMC 1 et SMC 3, associées à des protéines non SMC. Il est particulièrement présent au niveau de l'hétérochromatine centromérique, dont la structure intrinsèque permet une concentration de ce complexe. Ainsi, la cohésion centromérique est suffisante pour prévenir la séparation des chromatides sœurs (Hagstrom and Meyer, 2003).

Cette cohésion centromérique est augmentée grâce au remplacement de l'histone H3 par le variant d'histone centromérique CENP-A. Ce variant d'histone facilite le recrutement du complexe *Cohesin* mais favorise aussi l'attachement aux kinétochores (Van Hooser et al., 2001). De même, le complexe *Cohesin* associé à la kinase mitotique Aurora-B joue un rôle prépondérant dans la bi-orientation des kinétochores et dans le contrôle de la tension exercée sur les deux chromatides sœurs. La protéine kinase Aurora-B, qui fait partie d'un complexe comportant les protéines INCENP et *Survivin*, est ainsi impliquée dans la ségrégation des chromatides sœurs (Tanaka et al., 2002).

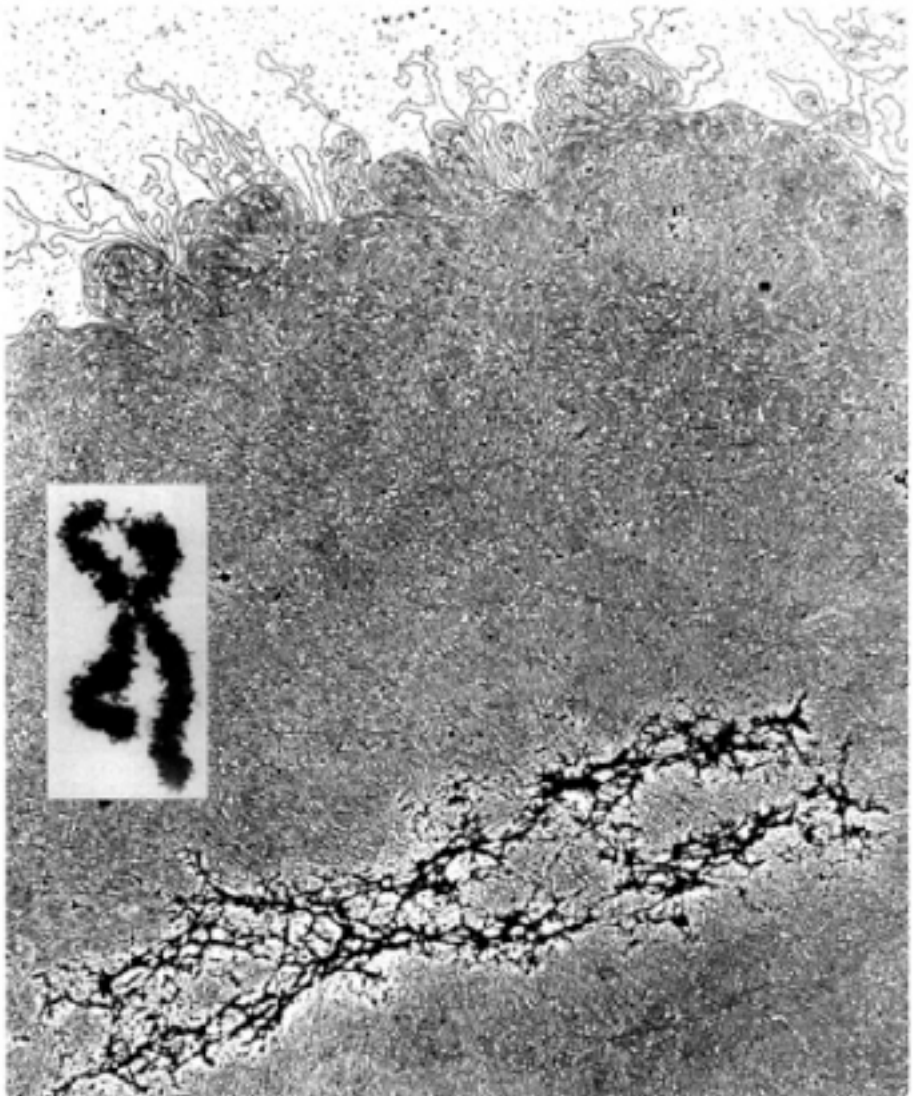


Figure 10 : échafaudage protéique des chromosomes (d'après Paulson et al., 1977).

Image de microscopie électronique montrant la libération de l'ADN après extraction des histones après traitement à 2M NaCl. L'ADN est retenu sous forme de boucles attachées à leur base à une structure fixe, la matrice ou échafaudage métaphasique. La base des boucles est constituée de séquences spécifiques (séquences SAR/MAR), très hautement conservées au cours de l 'évolution. En encart, un chromosome métaphasique natif en microscopie électronique.

Seule l'action de la *Separase*, une protéine qui clive le complexe *Cohesin*, permettra de séparer les chromatides sœurs en anaphase (Nasmyth, 2002).

La *Separase* est inhibée par la *Securin*, qui sera détruite suite à une ubiquitinylation par le complexe APC (Anaphase Promoting Complex), qui joue un rôle essentiel dans le passage en anaphase (Peters, 2002).

D'autres protéines SMC, les SMC 2 et SMC 4, forment le complexe *Condensin* qui peut s'associer à la protéine Topoïsomérase II. Ces protéines se distribuent de manière axiale en une structure hélicoïdale en « échelle de barbier » au sein des chromosomes (Coelho et al., 2003; Maeshima and Laemmli, 2003). Cette distribution apparaît en prophase tardive, et pourrait être stabilisée par une protéine « glue » axiale en fin de mitose (Kireeva et al., 2004). Le complexe *Condensin* participe à la condensation et à la ségrégation des chromosomes mitotiques en induisant des super tours positifs de manière dépendante de l'ATP (Kimura and Hirano, 1997; Kimura et al., 1999). Pour autant, le mécanisme par lequel le complexe *Condensin* agit au niveau de la chromatine reste incertain.

L'interaction du complexe *Condensin*, essentiel à la viabilité cellulaire, avec la Topoïsomérase II, joue un rôle central dans la condensation et la ségrégation des chromatides sœurs. La protéine Barren, aussi associée à ce complexe, serait alors un régulateur pour l'activité catalytique de la topoïsomérase II (Coelho et al., 2003).

En effet, outre son rôle de protéine de structure, la Topoïsomérase II a une activité de clivage spécifique des séquences MAR/SAR, mais aussi au niveau des séquences satellites de type III (SAT III) du chromosome X de la drosophile et des régions actives en transcription (Poljak and Käs, 1995). Ces différentes séquences ont la particularité de présenter des sites de clivage *in vivo* riche en paires de bases dG•dC contrairement à ce qui est décrit *in vitro*, où les sites consensus de clivages sont majoritairement riches en paires de bases dA•dT (Käs and Laemmli, 1992).

La Topoïsomérase II est une enzyme qui forme un homodimère capable de relâcher les super-tours de l'ADN, de manière dépendante de l'ATP (figure 11). Son cycle catalytique lui permet également de prendre en charge de nombreuses transactions topologiques qui dépendent du substrat. Elle joue un rôle essentiel dans la ségrégation des chromatides sœurs et dans leur condensation, principalement en anaphase (Cortes et al., 2003). Le rôle de la Topoïsomérase II dans la condensation a été clairement établi grâce à des expériences d'immunodéplétion de cette protéine suivies de complémentation dans des extraits mitotiques d'œufs de Xénope (Adachi et al., 1991). Il a été montré que la Topoïsomérase II se trouve

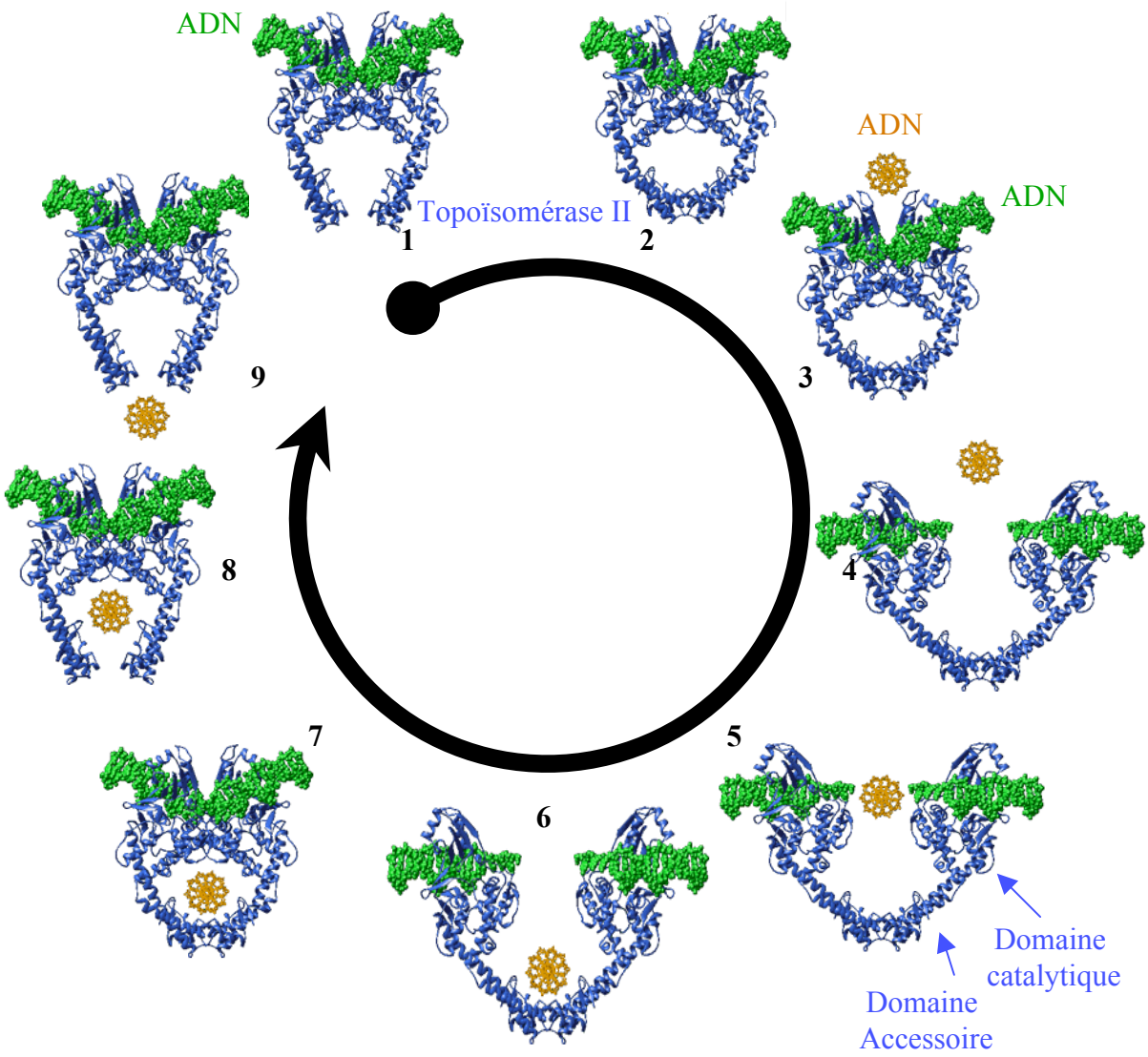


Figure 11 : Cycle catalytique de clivage par la topoïsomérase II.

1 : le cycle s'amorce lorsque un segment d'ADN (en vert) s'associe à un homodimère de topoïsomérase II (en bleu, le domaine ATPase n'est pas représenté).

2 et 3 : lors du rapprochement d'un deuxième segment d'ADN, le dimère de topoïsomérase II induit la coupure de chacun des deux brins du premier segment et leur séparation.

4, 5 et 6 : le deuxième segment passe à travers du premier fragment ainsi ouvert.

7 : une étape de réversion permet la fermeture du premier segment.

8 : suite à un nouveau changement conformationnel, la topoïsomérase II provoque l'ouverture de la pince formée par les domaines accessoires ce qui permet la libération du deuxième segment d'ADN.

9 : le recyclage de la topoïsomérase II s'achève par la dimérisation des domaines accessoires qui restaurent la structure initiale de l'enzyme et la rendent disponible pour un nouveau cycle catalytique. (images tirées du site *www.virtuallaboratory.net*)

localisée sous sa forme catalytiquement active sur l'axe longitudinal des chromosomes (Agostinho et al., 2004) mais aussi au niveau de régions hétérochromatiniennes, comme les séquences SAT III de l'hétérochromatine du chromosome X de la drosophile (Käs and Laemmli, 1992), qui représenteraient ses sites majoritaires de dépôts.

Enfin, un troisième complexe de protéine SMC, comportant les SMC 5 et SMC 6 (aussi appelée Rad18), est impliqué dans la réparation de l'ADN (Fujioka et al., 2002). Ce complexe joue aussi un rôle sur l'organisation des chromosomes puisque des mutations affectant à la fois Rad 18 (SMC 6) et la Topoïsomérase II sont synthétiques létales (Verkade et al., 1999). Dans ces expériences, des souches de levure à fission exprimant de manière thermosensible une Topoïsomérase II mutée ont été utilisées à température semi-permissive afin de diminuer la mortalité liée à la perte de la protéine.

Une autre protéine interviendrait dans la mise en place de la structure des chromosomes : la Titine. Cette protéine semble conférer une certaine élasticité à la structure des chromosomes mitotiques (Machado and Andrew, 2000). Des expériences réalisées *in vitro* mais aussi sur des chromosomes natifs montrent des propriétés d'élasticité identiques, suggérant une structure sous-jacente similaire, qui pourrait être sensiblement différente des modèles actuellement discutés : le modèle de l'échafaudage en boucle ou du repliement en hélice hiérarchisé.

c. Rôle des histones de cœur dans la condensation

Les queues N-terminales des histones de cœur sont particulièrement impliquées dans la condensation de la chromatine. Plus particulièrement, la phosphorylation de la sérine 10 de l'histone H3 est importante en mitose, tout comme la phosphorylation de l'histone H2A (Guo et al., 1995).

La phosphorylation de la sérine 10 de l'histone H3, probablement par une kinase de la famille Aurora, marque le début de la phase de condensation des chromosomes mais n'est pourtant pas requise pour le maintien de la condensation des chromosomes (Van Hooser et al., 1998). Dans le cas de l'expression d'une histone mutante non-phosphorylable sur la sérine 10, les chromosomes sont incapables de se condenser et d'avoir une ségrégation correcte. Un

autre résidu, la sérine 28 de la même histone H3 pourrait aussi être impliquée dans la condensation correcte des chromosomes mitotiques (Goto et al., 1999).

Par contre, l'intervention de l'histone de liaison H1 dans ce mécanisme reste sujet à controverse. Du fait de son importante phosphorylation en début de mitose et sa déphosphorylation en fin d'anaphase, elle pourrait n'être impliquée que de manière accessoire dans la condensation des chromosomes.

d. Attachement des SAR/MAR à la matrice nucléaire

Des expériences menées par différents groupes ont permis de mettre en évidence l'existence d'une structure protéique qui retenait les boucles de chromatine à l'intérieur du noyau (Lebkowski and Laemmli, 1982a; Lebkowski and Laemmli, 1982b). Ce « squelette nucléaire », lié par les séquences MAR/SAR, serait constitué à la fois d'un réseau protéique interne et de la lamina nucléaire en périphérie.

La composition biochimique du réseau protéique interne reste à l'heure actuelle peu caractérisée. Pour autant, il a été montré que certaines protéines pouvaient faire partie de ce réseau, comme par exemple les ARN polymérases I (Pol I) et II (Pol II) (Keppel, 1986; Lewis et al., 1984).

La lamina nucléaire périphérique est impliquée dans l'organisation interphasique de la chromatine. Elle est composée de trois protéines apparentées : les Lamines A, B et C. Outre son rôle dans l'attachement des boucles de chromatines, elle permet aussi la liaison des filaments intermédiaires (Aebi et al., 1986).

Cette lamina consiste en un réseau protéique continu qui est intimement en relation avec la membrane nucléaire interne dont il suit les contours (Gerace et al., 1984). Les seules interruptions qu'il présente se trouvent au niveau des complexes de pores nucléaires (CPN) (Aaronson and Blobel, 1975).

Le rôle de la lamina dans la formation de domaines chromatiniens interphasiques passerait par l'interaction, directe ou indirecte, entre un de ses composants et une protéine présente en périphérie des chromosomes mitotiques. Des résultats récents mettent en évidence une interaction entre la lamine B et la protéine HP1 (Heterochromatin Protein 1), qui peut

passer ou non par l'intervention du récepteur de la Lamine B, (Baricheva et al., 1996; Eissenberg and Elgin, 2000). Il est à noter que la Lamine B est aussi responsable de la liaison entre l'enveloppe nucléaire et les deux autres Lamines, A et C, qui ont quant à elles des propriétés de liaison à l'ADN (Lebkowski and Laemmli, 1982b).

Ces données d'attachement de la chromatine à la matrice nucléaire ainsi que les résultats actuels semblent aller en faveur de l'existence d'une compartimentalisation nucléaire des chromosomes, ceux-ci étant eux-mêmes arrangés en fonction de l'activité transcriptionnelle de certains gènes. Les gènes actifs seraient en effet principalement localisés à proximité des CPN ou y seraient conduit pour leur activation (Blobel, 1985). Cette hypothèse a été récemment confirmée grâce à des études chez la levure *Saccharomyces cerevisiae* (Schmid et al., 2006).

4. Structures d'ordres supérieurs

Les boucles de chromatine reliées à une armature s'enroulent en une spirale plus ou moins resserrée selon le stade du cycle cellulaire. En interphase, cette spirale est relâchée et les chromosomes ne peuvent plus être distingués les uns des autres car ils sont « débobinés ». Néanmoins, les chromosomes ne s'emmêlent pas et semblent au contraire occuper des domaines bien définis dans le volume nucléaire (Lichter et al., 1988; Manuelidis, 1985).

Au cours de la division cellulaire, la spirale de la chromatine se condense encore afin d'atteindre, en métaphase, une structure de 700 nm de diamètre qui reflète le degré de compaction maximale de la fibre chromatinienne.

Selon les techniques de colorations et les traitements que subissent les chromosomes, les chromosomes deviennent visibles et apparaissent composés de deux entités distinctes : les deux chromatides sœurs. Ces techniques permettent aussi de visualiser artificiellement deux types de structures présentant deux niveaux de compaction différents : l'euchromatine, moins condensée, et l'hétérochromatine, plus compacte.

a. Bandes chromosomiques

La coloration des chromosomes en métaphase génère le long de ces structures, une alternance de bandes transversales de coloration distinctes. Les bandes ainsi obtenues prennent le nom du type de traitement qui les génère, pour revue (Craig and Bickmore, 1993).

Le traitement au Giemsa fait apparaître les bandes dites G. Cette coloration est appliquée après dégradation partielle des protéines par action de la Trypsine. Ainsi, certaines données suggèrent que les bandes G correspondent à des zones plus condensées, alors que d'autres concluent à une variation sur grande échelle de la composition en paires de bases. En effet, lors de l'utilisation de fluorochromes spécifiques des paires de bases dA•dT, tels que la quinacrine ou la daunomocyne, ceux-ci font apparaître des bandes appelées bandes Q, qui se superposent parfaitement aux bandes G. Les séquences qui forment les bandes G/Q sont de manière générale pauvres en gènes et en cytosines méthylées, présentent des séquences de type LINE (Long Interspersed Repeated Sequence) et une réplication tardive en phase S (figure 12).

Le R-banding, qui est une autre technique de marquage utilisant un tampon phosphate chauffé et du Giemsa, révèle des bandes R qui sont complémentaires des bandes G/Q. Des études ont montré que les bandes R et G/Q correspondent respectivement à l'euchromatine et à l'hétérochromatine facultative. Il a été montré que les bandes R étaient, par opposition aux bandes G/Q, riches en paires de bases dG•dC, en gènes, en cytosines déméthylées et en séquences de types SINE (Short Interspersed Repeated Sequence). De plus, les bandes R ont la propriété de se répliquer de manière précoce durant la phase S.

Une troisième technique, appelée C-banding et qui consiste à traiter des chromosomes avec des acides et des bases puis à les colorer avec du Giemsa, révèle les bandes C qui représentent l'hétérochromatine constitutive. Elles se retrouvent majoritairement localisées au niveau des centromères et des constrictions secondaires. Ces bandes sont principalement composées de séquences d'ADN dites satellites, des séquences hautement répétées, dont la composition biaisée en paires de bases leur confère une densité caractéristique.

Bandes G	Bandes R (+T)
Bandes G+ colorées	Bandes G- claires
Bandes Q+ fluorescentes	Bandes Q- non fluorescentes
Bandes R- et T- claires	Bandes R+ et T+ colorées
Régions chromomériques au stade pachytène de la méïose	Régions interchromomériques
Réplication tardive au cours de la phase S	Réplication précoce
Riches en paires de bases dA et dT	Riches en paires de bases dG et dC
Pauvres en cytosines méthylées	Riches en cytosines méthylées
Pauvres en gènes	Majorité des gènes de ménage et de spécificité tissulaire
Séquences LINE	Séquences SINE

Figure 12 : tableau regroupant les principales caractéristiques biochimiques et fonctionnelles des bandes chromosomiques.

Un dernier type de bande peut également être visible grâce à la technique du T-banding qui utilise une étape de dénaturation thermique : ce sont les bandes T, qui correspondent à un sous-ensemble des bandes R. Ces bandes T se retrouvent au niveau des séquences riches en dG•dC, elles sont localisées principalement aux télomères et sont plus résistantes que les bandes R à la dénaturation thermique.

D'après toutes ces données, les bandes chromosomiques peuvent être considérées soit comme des segments d'ADN avec un contenu spécifique en paire de bases, en gènes et en séquences répétées dispersées, soit comme des unités fonctionnelles de réplication et de condensation. Ceci indique que le génome des mammifères est une structure organisée dont les caractéristiques fonctionnelles ne se distribuent pas de manière aléatoire.

Jusqu'aux années 80, il était admis que les bandes reflétaient une différenciation longitudinale de la structure nucléoprotéique des chromosomes, mais la nature de cette différenciation n'était pas connue. Différents modèles découpant le génome en sous-ensembles d'ADN avec des caractéristiques variables et une répartition différente selon les bandes G ou R et T ont été proposés. Ils permettent d'aller un peu plus loin dans la compréhension de l'organisation du génome et des mécanismes à l'origine des bandes chromosomiques.

Un des premiers modèles est celui dit des « isochores » (Bernardi, 1995; Bernardi et al., 1985). Ce terme désigne les deux compartiments majeurs traduisant une organisation fondamentale du génome apparue au cours de l'évolution chez les vertébrés à sang chaud. Ce modèle subdivise le génome en cinq classes d'isochores selon leur composition en bases : L1 et L2 (riches en dA•dT, donc plutôt dans les bandes G), H1, H2 et H3 (de plus en plus riches en dG•dC, H1 et H2 se trouvent plutôt dans les bandes R et H3 dans les bandes T) (figure 13).

En 1988, un modèle basé sur la répartition des séquences répétées dispersées le long des chromosomes est formulé (Korenberg and Rykowski, 1988). Cette hypothèse se base plus particulièrement sur la répartition de deux types de séquences : les séquences SINE et les séquences LINE.

Les séquences SINE sont des séquences de taille inférieure à 1 kb, dont 300 000 à 900 000 copies sont réparties au travers du génome. Les LINE sont des séquences de plus

Bandes	T	R	G
Isochores	~ H3	~ H1 + H2	~ L1 + L2
% en paires de bases dG et dC	~ 53	~ 46	~ 39
% du génome	~ 3	~ 31	~ 62
% de gènes	39	38	34
ADN en Mb	90	190	1860
Densité de gènes	1 / 3,2 kb	1 / 25 kb	1 / 50 kb

Figure 13 : tableau présentant la distribution globale des isochores et des gènes dans les chromosomes humains selon les bandes T, R et G.

grande taille, de 6 à 7 kb, mais qui sont beaucoup moins représentées que les SINE dans le génome : on en retrouve seulement de 10 000 à 100 000 copies.

Chez l'Homme, plus de 90 % de ces séquences correspondent, pour les SINE, aux séquences Alu de 300 pb, ainsi nommées car elles possèdent généralement un site de restriction pour l'enzyme *Alu*I, et pour les LINE aux séquences L1 ou *Kpn*I, ces dernières possédant un site de restriction reconnu par l'enzyme *Kpn*I.

Il a été montré par hybridation *in situ* que ces séquences ne se répartissent pas au hasard le long des chromosomes humains, mais suivent une distribution caractéristique avec une prédominance des séquences Alu, environ 18 % de l'ADN, dans les bandes R alors que les séquences L1 sont principalement présentes dans les bandes G, environ 14 % de l'ADN (Chen and Manuelidis, 1989). Ceci suggère qu'elles sont impliquées au moins dans certaines propriétés opposées des deux fractions de l'euchromatine.

Dans le modèle dit « des saveurs », le rapport entre composition en bases et richesse en séquence Alu est déterminé, ce qui divise la chromatine des bandes R du chromosome métaphasique en quatre classes, ou « saveurs » (Holmquist, 1992). Ces quatre classes se répartissent ainsi : les séquences riches en dA•dT et riches en séquences Alu, les séquences riches en dA•dT mais pauvres en séquences Alu, les séquences riches en dG•dC et en séquences Alu et enfin les séquences riches en dG•dC et pauvres en séquences Alu. La densité des gènes varierait alors en suivant ces quatre « saveurs ».

Une autre théorie suppose que l'ADN est fractionné en sous-ensembles selon la distance entre îlots CpG. Selon leur état de méthylation, les îlots CpG peuvent être coupés ou non par une enzyme sensible à la méthylation. On trouve particulièrement ces îlots CpG dans la partie 5' des gènes de ménage, d'où leur qualité d'indicateur de la densité en gènes de ménage. Quatre classes ont ainsi été définie selon la distance entre les îlots CpG : si la distance est supérieure à 1 000 kb, cette partie est pauvre en gènes et se trouve principalement dans les bandes G, si la distance est comprise entre 500 et 1 000 kb, alors il s'agit de bandes R. Enfin, si la distance est comprise entre 15 et 100 kb, ou entre 100 et 500 kb, ce sont des séquences riches en gènes situées plutôt dans les bandes T (Craig and Bickmore, 1994).

b. Modèle de repliement des chromosomes

Ces modèles proposent des théories de différenciation structurale des chromosomes. Le premier modèle se base sur la composition en paires de bases différente le long de la fibre de chromatine, une variabilité se traduisant par une compaction différente entre les régions riches en dA•dT, plus condensées, et les régions riches en dG•dC, moins condensées. Ce modèle se réfère aux différentes expériences de coloration permettant de visualiser les bandes chromosomiques. Il fait donc la liaison entre la compaction faible de l'euchromatine, riche en dG•dC, et la forte compaction de l'hétérochromatine, riche en dA•dT.

L'autre modèle proposé se base sur une différence de densité de la chromatine selon les bandes, ce qui permettrait de relier l'organisation de la fibre de la chromatine et l'alternance des bandes le long des chromosomes en métaphase. Cette théorie suppose que les séquences SAR, très riches en paires de bases dA•dT, seraient organisées de manière non aléatoire sur le génome et formeraient ainsi une structure centrale le long de l'axe des chromosomes appelée alignement AT (Saitoh and Laemmli, 1994). Cet alignement AT présenterait un arrangement et une compaction différente le long des chromosomes générant ainsi les bandes Q et les bandes R (figure 14). Dans les bandes Q, la compaction de l'alignement AT serait importante et il serait enroulé en spirale, formant des boucles d'ADN courtes et orientées parallèlement à l'axe du chromosome. Dans les bandes R, les caractéristiques de l'alignement AT sont opposées, à savoir que celui-ci serait arrangé de manière plus linéaire le long de l'axe, les boucles d'ADN, plus longues que dans les bandes G, seraient perpendiculaires à la structure centrale des chromosomes.

Dans ce modèle, la différenciation des bandes ne serait pas due à une différence des compositions en paires de bases mais plutôt à une différence de repliement de la fibre chromatinienne. Les bandes R seraient donc peu condensées alors que les bandes G seraient fortement compactées.

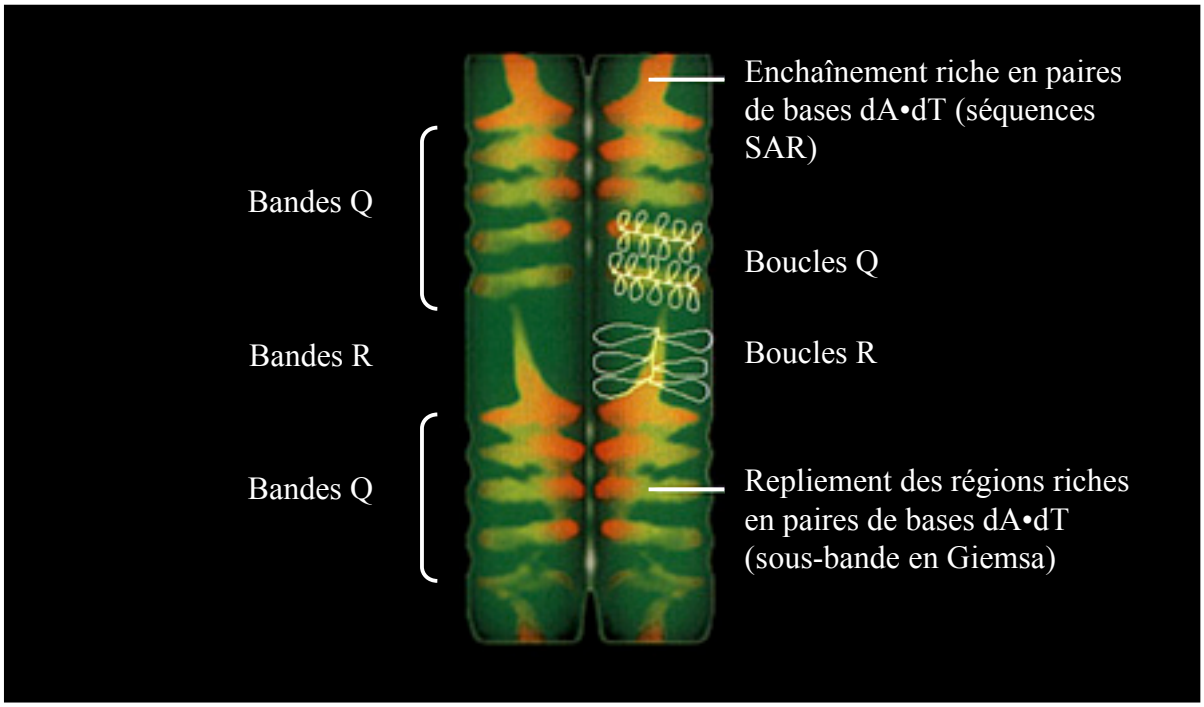


Figure 14 : modèle de repliement des chromosomes de Saitoh (d'après Saitoh and Laemmli, 1994).

Les bandes des chromosomes métaphasiques reflètent le repliement différentiel de l'axe central des chromosomes à travers les régions condensées et décondensées de la chromatine. L'échafaudage chromosomique suit un chemin hélicoïdal à travers l'axe longitudinal du chromosome. Ce modèle prédit un repliement différentiel des boucles de chromatines dans les bandes R et Q. Les bandes R correspondent à une compaction modérée de l'enchaînement riche en paires de bases dA•dT, la chromatine y forme des boucles longues et le long de l'axe. Ces caractéristiques sont opposées dans les bandes Q. Les axes de l'échafaudage des chromatides sœurs sont symétriques.

IV. L'hétérochromatine

On distingue communément deux types de chromatine, définis initialement sur la base d'observations cytologiques menées sur différentes espèces de drosophile par Emil Heitz en 1928 (Heitz, 1928). Ces études sont restées relativement confidentielles à l'époque et ce n'est qu'au cours de ces dernières décennies que ses travaux ont été largement reconnus.

L'euchromatine correspond à la moins dense des deux phases retrouvées dans le noyau, elle représente la portion du génome enrichie en gènes et, du fait de sa décondensation en interphase, elle est plus accessible aux machineries transcriptionnelles. Sur le plan moléculaire, elle est caractérisée par un fort enrichissement en histones acétylées sur leurs extrémités N-terminales.

L'hétérochromatine, quant à elle, fut tout d'abord décrite comme la portion du génome qui ne se décondense pas en interphase. Elle se distingue de l'euchromatine par la nature même des séquences qui la constituent. Pauvre en gènes, présentant des histones hypoacétylées et globalement méthylées sur leur queue N-terminale, elle est majoritairement composée de séquences répétées de nature et de complexité variables. Sur le plan moléculaire, ce sont la relative hypoacétylation des histones, la méthylation des lysines 9 et 27 de l'histone H3 ainsi que son fort enrichissement en protéine HP1 qui reconnaît spécifiquement la méthylation de la lysine 9 de l'histone H3, qui représentent les principales marques épigénétiques de l'hétérochromatine (Dillon, 2004; Elgin and Grewal, 2003). Chez les plantes et les vertébrés, la méthylation des cytosines des îlots CpG est une marque supplémentaire de ce domaine chromatinien.

1. Hétérochromatine constitutive et hétérochromatine facultative

Au sein même de l'hétérochromatine, deux types sont généralement distinguables : l'hétérochromatine constitutive et facultative. Cette dernière correspond plus précisément aux régions génomiques portant les gènes régulés au cours du cycle cellulaire (Gilbert et al., 2003). Un exemple d'hétérochromatine facultative est celui des gènes régulés par le système de E2F/Rb, dans lequel la répression de la transcription se fait par la mise en place d'une

structure de type hétérochromatinienne (Nielsen et al., 2001b; Vandel et al., 2001). Cependant, il n'y a aucune évidence concernant l'intervention de la protéine HP1, comme c'est le cas dans l'hétérochromatine constitutive. La formation d'une structure d'hétérochromatine facultative intervient aussi dans les mécanismes de compensation de dose chez les mammifères. L'inactivation du chromosome X chez la femelle fait intervenir la méthylation de la lysine 9 de l'histone H3 (Avner and Heard, 2001; Heard et al., 2001). Cet état de mise en silence est réversible contrairement à l'hétérochromatine constitutive.

L'hétérochromatine constitutive, qui sera l'objet de ce développement, se retrouve quant à elle localisée majoritairement au niveau des centromères et des télomères de chaque chromosome (Grewal and Elgin, 2002; Grewal and Moazed, 2003). Elle est essentielle pour la ségrégation des chromosomes et pour la stabilité du génome (Wallrath, 1998).

Chez la drosophile, l'hétérochromatine constitutive et facultative prennent respectivement le nom d'hétérochromatine alpha (α) et bêta (β) (Heitz, 1934b). L'hétérochromatine α correspond à la partie qui forme le chromocentre des chromosomes polytènes. Ces chromosomes géants, présents dans les cellules des glandes salivaires, ont une partie euchromatinienne endorépliquée, amplifiée environ un millier de fois. Certaines régions euchromatiniennes échappent cependant à cette endoréplication, dont les gènes codant les histones. De même, l'hétérochromatine β est endorépliquée alors que l'hétérochromatine α ne l'est pas (Hilliker et al., 1980). L'hétérochromatine α, considérée comme l'hétérochromatine mitotique, est riche en séquences répétées. Par contre, l'hétérochromatine β n'est pas distinguable de l'hétérochromatine α sur les chromosomes mitotiques (Miklos and Cotsell, 1990). Cette distinction entre les deux types d'hétérochromatine chez la drosophile n'est utilisée que dans les études cytologiques menées sur les chromosomes polytènes.

L'hétérochromatine constitutive est caractérisée par de nombreux points : elle est constituée de séquences particulières, comme par exemple les séquences satellites, mais aussi de protéines spécifiques. Ces caractéristiques vont être développées dans la partie suivante.

2. *Composition nucléique de l'hétérochromatine*

a. Les séquences satellites péricentromériques

Les séquences satellites ont d'abord été décrites dans les expériences de centrifugation sur gradient de densité en Chlorure de Césium. La composition en paire de bases des génomes est biaisée : les séquences riches en gènes sont plutôt riches en paires de bases dG•dC alors que les séquences satellites sont majoritairement riches en dA•dT. Cette différence de composition se traduit par une sédimentation différente dans des expériences d'ultracentrifugation isopycnique, c'est-à-dire une séparation de molécules selon leur densité par centrifugation sur un gradient continu : ainsi les séquences satellites ne sédimentent pas au même niveau que le reste de l'ADN génomique, particularité qui leur a donné leur dénomination de « satellite ».

Les séquences satellites sont constituées de séquences courtes et répétées en tandem un très grand nombre de fois. Il en existe de plusieurs types qui comprennent de longues ou de courtes répétitions. Ces séquences satellites courtes, appelées microsatellites, présentent des répétitions pouvant aller de 2 à 20 paires de bases. Sur un chromosome eucaryote, ces microsatellites sont distribués uniformément dans l'euchromatine (Csink and Henikoff, 1998). Les longues séquences d'ADN répétés, plus ou moins complexes, sont généralement non transcrites et ce sont elles qui forment l'hétérochromatine constitutive.

Chez les primates et chez l'Homme (Waye and Willard, 1987), les régions centromériques de tous les chromosomes sont riches en ADN répété de type satellite alpha ou alphoïdes, composé de séries de monomères de 171 paires de bases répétées en tandem sur plusieurs centaines de kilobases (Devilee et al., 1986). Ces séquences peuvent représenter de 3 à 5 % de l'ADN centromérique de chaque chromosome (Willard and Waye, 1987). Il existe d'autres séquences d'ADN satellites situés dans les régions adjacentes, appelées péricentromériques, de certains chromosomes : les satellites 1, 2, 3 et bêta (Waye and Willard, 1989). Chez d'autres mammifères comme la souris, le mouton et la vache, l'ADN centromérique est également constitué de séquences répétées en tandem qui semblent spécifiques de l'espèce, mais leur organisation n'est pas aussi bien connue (Burkin et al., 1996; Singer, 1982).

Chez la drosophile, il existe plusieurs classes de séquences satellites dont les principales sont les séquences satellites de type I, III et V (Gall et al., 1971; Lohe et al., 1993) (figure 15). Les séquences satellites de type I (SAT I) et de type V (SAT V) sont de courtes séquences répétées, respectivement $(AATAT)_n$ et $(GAGAA)_n$ (Lohe et al., 1993). Seules les séquences SAT III sont de type alphoïdes et sont composées de la répétition d'un motif de 359 paires de bases riches en dA•dT (Hilliker et al., 1980). Elles représentent la quasi-totalité de l'hétérochromatine péricentromérique du chromosome X de *Drosophila melanogaster*. Les chromosomes X, Y, 3 et 4 possèdent des séquences SAT I et SAT V, alors que le chromosome 2 ne possède que des séquences SAT V (figure 15 et (Peacock et al., 1974)).

Malgré la conservation et la caractérisation de ces séquences, tant chez les mammifères que chez la drosophile, aucun rôle n'a pu être attribué aux séquences satellites de type alphoïde. Des hypothèses tendent à les impliquer dans la cohésion des chromosomes pendant la mitose. Par exemple, les séquences SAT III de la drosophile sont retrouvées à la fois sur les chromosomes X et Y, en proportion relativement moindre sur le chromosome Y, ce qui pourrait permettre leur appariement en mitose, mais aussi en méiose (Hsieh and Brutlag, 1979). Mais cette hypothèse a été réfutée depuis la mise en évidence du rôle du locus des gènes ribosomiques (rDNA) dans l'appariement des chromosomes X et Y de la drosophile en méiose (Park and Yamamoto, 1995). La question du rôle de ces séquences SAT III reste donc en suspend.

Pourtant, des protéines sont connues pour interagir avec ces séquences satellites, ce qui relance la spéculation sur leur rôle hypothétique. Des expériences ont aussi caractérisé des petits ARN issus de la transcription des séquences SAT III chez la drosophile (Aravin et al., 2003). Il en est de même chez d'autres eucaryotes, de la grenouille taureau à l'Homme (Valgardsdottir et al., 2005; Wu et al., 1986). La fonction des ces transcrits issus des séquences satellites reste à élucider, mais des hypothèses tendent à les impliquer dans la formation de la structure hétérochromatinienne (Allshire, 2002).

b. Les séquences télomériques

Les télomères sont des structures particulières présentes aux extrémités des chromosomes eucaryotes. Les télomères protègent les chromosomes et empêchent leur dégradation ou leur raccourcissement jusqu'à une longueur critique qui conduirait à

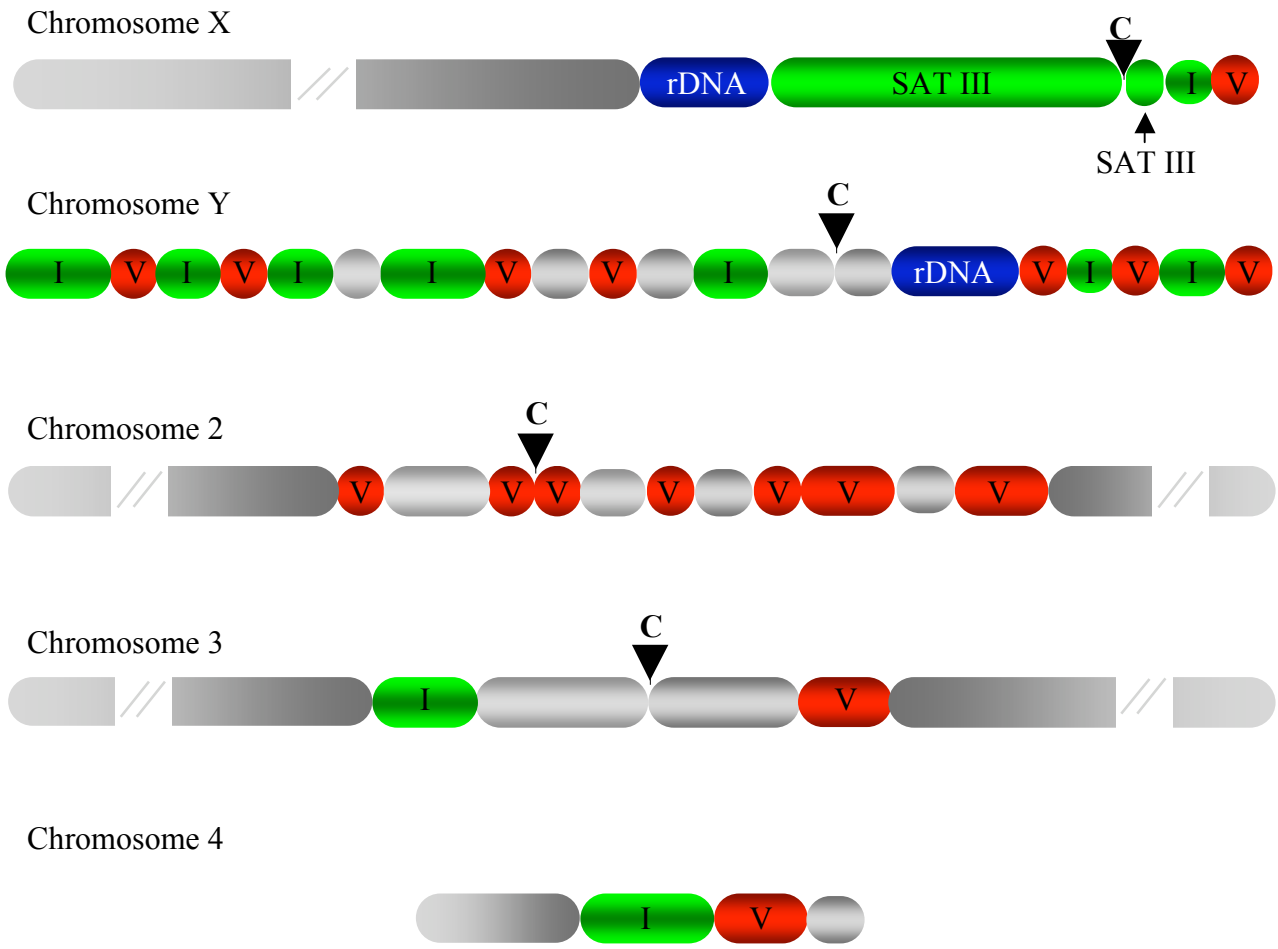


Figure 15 : cartographie des principales séquences satellites de l'hétérochromatine de* Drosophila melanogaster *(d'après Lohe et al. 1993).

Les principales séquences satellites sont représentées sur cette carte.

Locus des gènes ribosomiques (rDNA), présent sur les chromosomes X et Y uniquement.

Séquences satellite de type III (SAT III) : répétition d'un motif de 359 pb riche en paires de bases dA•dT

Séquences satellite de type I (SAT I): répétition d'un motif $(AATAT)_n$

Séquences satellite de type I (SAT I): répétition d'un motif $(AAGAG)_n$

l'altération de leur intégrité et à l'entrée de la cellule en sénescence. Ces structures protectrices sont constituées par l'association d'ADN télomérique et de protéines spécifiques.

Les télomères sont des éléments essentiels préservant la stabilité du génome. Dans les cellules humaines, un raccourcissement des chromosomes dans la plupart des cellules somatiques a été observé au fil des divisions cellulaires. L'hypothèse selon laquelle les télomères joueraient un rôle déterminant dans la sénescence cellulaire a donc été proposée. Une enzyme capable de prévenir ce raccourcissement naturel, la télomérase, a été identifiée dans un premier temps chez *Tetrahymena* (Greider and Blackburn, 1985).

Le séquençage de l'ADN télomérique chez différentes espèces montre que la séquence de l'ADN présent aux extrémités des chromosomes est relativement bien conservée ce qui est en faveur de son importance fonctionnelle. Dans la plupart des cas, l'ADN télomérique est constitué de courtes séquences répétées en tandem (Wellinger and Sen, 1997). Chez la plupart des vertébrés, cette séquence correspond à la répétition d'un motif de six paires de bases : $(TTAGGG)_n$ (Moyzis et al., 1988).

La drosophile, qui a servi à l'élaboration même du concept de télomère, est paradoxalement une des espèces qui font exception à ces règles générales de constitution des télomères. En effet, elle possède aux extrémités de ses chromosomes un nombre multiple de copies de rétrotransposons spécifiques des télomères : Het-A (Danilevskaya et al., 1994), TART (Abad et al., 2004a), TAHRE (Abad et al., 2004b). Il a été démontré que la protéine HP1 jouait un rôle fonctionnel dans le maintien de la fonction des télomères chez la drosophile (Fanti et al., 1998). De plus, la protéine HP1 est impliquée dans le phénomène de variégation d'effet de position télomérique (Cryderman et al., 1999).

Cette localisation télomérique de la protéine HP1 suppose la présence d'histone H3 méthylée sur la lysine 9, modification créée au niveau des télomères par la protéine Enhancer of Zeste, E(Z) (Kuzmichev et al., 2002). Cependant, la présence de la protéine HP1 au niveau télomérique pourrait aussi faire intervenir une interaction directe entre cette protéine et l'ADN télomérique (Perrini et al., 2004). La protéine E(Z), tout comme la protéine HP1, est impliquée dans la mise en silence de gènes médiée par l'hétérochromatine télomérique à la fois chez la drosophile, mais aussi chez la levure (Laible et al., 1997). Cependant, chez la levure, il n'y a pas à proprement parler de structure hétérochromatinienne, sauf au niveau des télomères (Dujon, 1996). La formation de cette structure particulière de mise en silence de la

transcription par des régions télomériques nécessite la présence de protéines du groupe des protéines SIR (Silent Information Regulator, (Gartenberg, 2000)).

Il est à noter que les télomères de la plupart des organismes dériveraient des mêmes télomères ancestraux. Les séquences télomériques « atypiques » de la drosophile résulteraient de modifications des séquences télomériques ancestrales plutôt que de la création *de novo* d'une nouvelle structure télomérique.

c. Les gènes hétérochromatiniens

Dans les premiers temps, l'hétérochromatine était définie comme la partie de la chromatine qui était pauvre en gènes. Pourtant, chez la drosophile une cinquantaine de gènes, parmi lesquels *rolled* (Berghella and Dimitri, 1996), *light* (Devlin et al., 1990b) ou *Nipped-B* (Rollins et al., 1999), est localisée dans l'hétérochromatine. Ils y sont normalement transcrits (Corradini et al., 2003; Hilliker, 1976) et leur structure nucléosomale est proche de celle des gènes euchromatiniens (Dillon and Festenstein, 2002) malgré un environnement chromatinien de type hétérochromatinien dans lequel l'arrangement nucléosomal est régulier (Sun et al., 2001).

Des expériences ont démontré que la localisation de ces gènes dans des structures hétérochromatiniennes était essentielle à leur expression normale (Eberl et al., 1993). Ces données décrivent donc l'hétérochromatine comme un compartiment spécialisé, adapté à l'expression des gènes hétérochromatiniens. En effet, leur déplacement dans un contexte euchromatinien entraîne leur inactivation. Cette activation transcriptionnelle des gènes hétérochromatiniens mettrait en jeu la protéine HP1, marqueur de l'hétérochromatine initialement décrite pour son rôle de mise en silence de la transcription (Lu et al., 2000).

Contrairement aux principales définitions, l'hétérochromatine n'est donc pas seulement une structure de mise en silence des gènes. Elle peut aussi représenter une structure nécessaire à l'activité transcriptionnelle de certains gènes (Eissenberg and Hilliker, 2000). Pour autant, certains auteurs supposent que la localisation de ces gènes était initialement euchromatinienne et que ces loci auraient été envahis et submergés par des séquences hétérochromatiniennes. Après évolution, le patron d'expression de ces gènes ce serait adapté à cette localisation (Miklos and Cotsell, 1990).

d. Les éléments frontières

La chromatine est également sous-divisée en zones ou « domaines chromatiniens » qui sont définis de manière opérationnelle par des éléments d'ADN spécifiques. Ces sites appelés éléments frontières, ou isolateurs, définissent ainsi les limites de domaines chromatiniens et restreignent les interactions activateurs/promoteurs aux seules zones où ils sont conjointement présents (Vazquez et al., 1993). Ces éléments frontières fonctionnent généralement par paires et ils interviendraient dans la constitution de l'ultra-structure chromatinienne et l'organisation nucléaire (figure 16).

Ce type d'éléments frontières peut être identifié *in vivo* par deux méthodes :

- la première technique, dénommée « position-independent expression » consiste à vérifier l'activité de protection contre les effets de position (Kellum and Schedl, 1991).
- la deuxième technique, appelée « enhancer blocking assay » (Kellum and Schedl, 1992), permet de contrôler la capacité de cet élément à bloquer la relation d'activation entre une séquence amont et le promoteur correspondant.

Il est important de noter que ces domaines n'influent pas intrinsèquement sur la transcription : ils n'inactivent pas les promoteurs, ou les éléments activateurs, quand ils sont interposés entre ces deux séquences. Ils ne font que limiter l'action des facteurs à la seule zône qu'ils définissent et ils bloquent la communication entre un élément activateur et le promoteur lorsqu'ils sont interposés entre ces deux séquences. Ainsi, le fonctionnement de tels éléments est totalement différent de la répression bidirectionnelle de certains éléments répresseurs.

Chez la drosophile, trois domaines frontières ont été particulièrement bien caractérisés : les éléments scs et scs', l'isolateur du rétrotransposon *gypsy* et les éléments frontières du complexe bithorax.

Les éléments scs et scs' forment les frontières proximales et distales du locus 87A7, contenant deux gènes divergents *hsp70* (Udvardy et al., 1985). Ils permettraient de limiter la

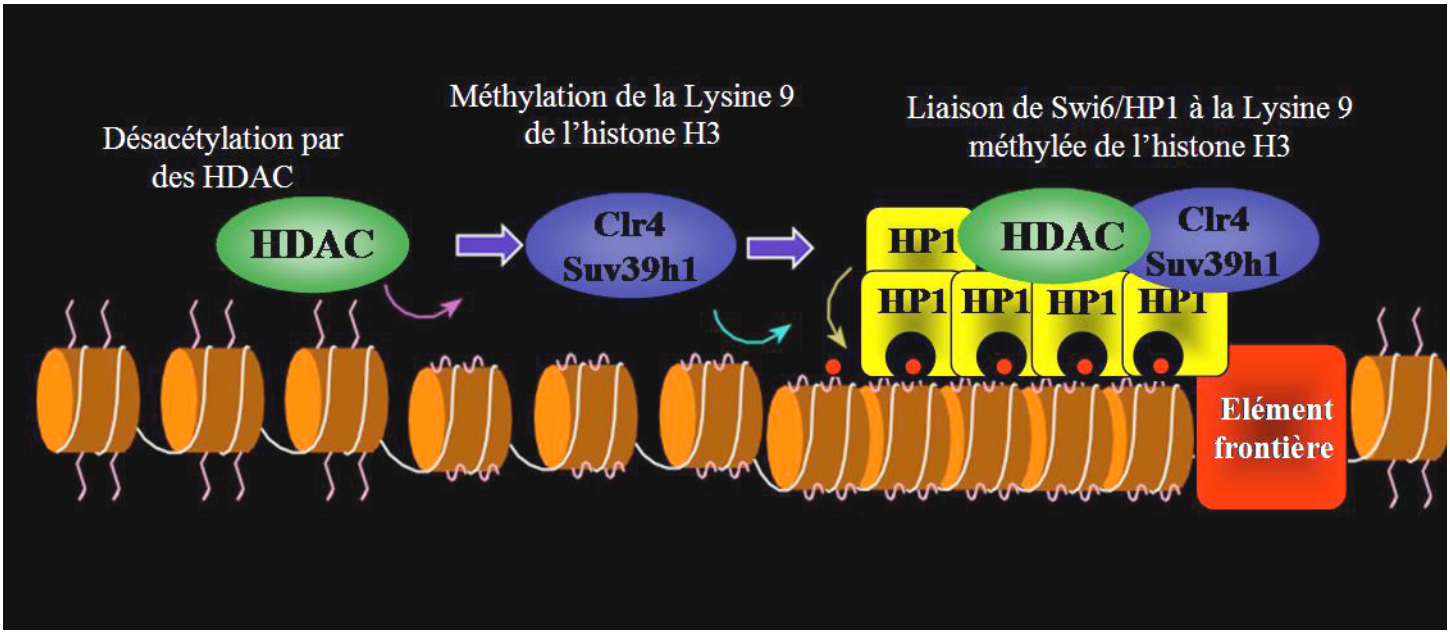


Figure 16 : rôle des éléments frontières dans l'arrêt de l'expansion de la structure hétérochromatinienne (d'après Elgin et al. 2003).

Les éléments frontières sont des séquences d'ADN interposées entre deux domaines chromatiniens de structure différente. De manière générale, les histones de ces séquences sont hyper-acétylées, permettant de rentrer en compétition avec la méthylation.

réponse au choc thermique à la seule région qu'ils englobent. Ils ont été caractérisés comme possédant une activité de blocage des effets de position lorsqu'ils entourent un transgène rapporteur, mais sont également actifs dans le « enhancer blocking assay » (Vazquez et al., 1993).

Le rétrotransposon *gypsy*, et plus particulièrement un fragment de 340 paires de bases, présente une activité de blocage (Cai and Levine, 1995; Gdula et al., 1996; Geyer and Corces, 1992). Ce fragment contient, par ailleurs, une séquence SAR/MAR qui pourrait elle aussi jouer un rôle dans ce phénomène en changeant la localisation de cette séquence dans le volume nucléaire (Nabirochkin et al., 1998). L'activité de blocage de cet élément met en jeu la reconnaissance de cette séquence par la protéine à doigt de zinc Suppressor of Hairy Wing (SU(Hw)) (Roseman et al., 1995a; Roseman et al., 1995b).

La protéine SU(Hw) présente trois domaines fonctionnels : un domaine en doigt de zinc lui permettant d'interagir avec l'ADN (Kim et al., 1996), un domaine de blocage de l'effet activateur (Harrison et al., 1993), et un domaine d'interaction avec le produit du gène *mod(mdg4)* (Gdula and Corces, 1997). L'interaction entre la protéine SU(Hw) et le produit du gène *mod(mdg4)* est essentielle à l'activité de frontière de la séquence (Gerasimova et al., 1995).

Le complexe bithorax (BX-C) de la drosophile comporte trois gènes homéotiques : les gènes *ultrabithorax* (*Ubx*), *Abdominal A* (*abd-A*) et *Abdominal B* (*abd-B*). Les produits de ces gènes sont responsables de l'identité des parasegments qui forment la moitié postérieure du thorax et l'abdomen de la mouche (Casanova et al., 1985). De nombreuses données suggèrent que chaque gène ferait partie d'une unité fonctionnelle indépendante grâce à la présence d'éléments frontières (Galloni et al., 1993).

La plupart des séquences décrites précédemment sont spécifiquement reconnues par des protéines qui jouent elles aussi un rôle important dans la formation, la structure et la fonction de l'hétérochromatine. C'est le cas de l'élément frontière BE28, et de sa protéine transactivatrice BEAF32, auquel je me suis intéressée au cours de ma thèse.

L'élément frontière BE28 correspond à une séquence de 269 pb, incluse dans une région de 1,2 kb répétée environ 150 fois sur le chromosome X. Elle est localisée dans la région séparant l'euchromatine de l'hétérochromatine de ce chromosome (Cuvier et al., 2002). La séquence BE28 possède les propriétés d'insulation mais à l'heure actuelle, aucune

séquence « partenaire », les éléments fontières agissant généralement par paire, n'a pu être identifiée.

La protéine de drosophile BEAF32 (Boundary Element Associated Factor of 32 kDa), a tout d'abord caractérisée comme interagissant avec l'élément frontière scs' qui, avec son partenaire scs, encadre le locus 87A7 contenant deux gènes divergents codant *hsp70* (Zhao et al., 1995). La protéine BEAF32 reconnaît des motifs palindromiques de type CGATA TATCG et des études complémentaires ont montré qu'elle était responsable de l'activité d'isolation de l'élément frontière BE28 (Cuvier et al., 1998).

La reconnaissance de l'ADN par la protéine BEAF32 s'effectue grâce à son domaine N-terminal, les domaines central et C-terminal de la protéine BEAF32 étant impliqués dans l'auto-association de la protéine. Deux isoformes de la protéine BEAF32 existent : les protéines BEAF32A et BEAF32B (Hart et al., 1997; Zhao et al., 1995). Ces protéines s'associent pour former un trimère comportant une protéine BEAF32A et deux protéines BEAF32B.

La protéine BEAF32 peut aussi interagir, toujours grâce aux domaines central et C-terminal, avec le domaine C-terminal de la protéine chromosomale non-histone D1 (Cuvier et al., 2002). Cette interaction pourrait réguler l'activité de l'élément frontière BE28 (figure 17). Actuellement, aucun partenaire du BE28 n'a été caractérisé. Cependant, la mise en évidence d'une interaction entre les protéines BEAF32 et Zw5, une autre protéine impliquée dans le mécanisme d'isolation, suppose l'existence d'une séquence partenaire.

La protéine de drosophile Zw5, précédemment dénommée SBP, est le produit du gène *zeste-white 5* (Gaszner et al., 1999). Elle joue un rôle dans l'activité isolatrice de l'élément scs en permettant la formation d'une boucle de chromatine, réunissant les éléments frontières scs et scs'. Cette boucle, qui formerait un mini-domaine contenant les deux gènes *hsp70* divergents du locus 87A7, passe par l'interaction entre les protéines Zw5 et BEAF32, qui reconnaissent respectivement les éléments scs et scs', qui encadrent ce même locus (Blanton et al., 2003).

3. *Ligands de l'hétérochromatine*

De nombreuses protéines ont été décrites comme interagissant spécifiquement avec les régions hétérochromatiniennes. Le rôle de certaines protéines dans la formation ou le maintien

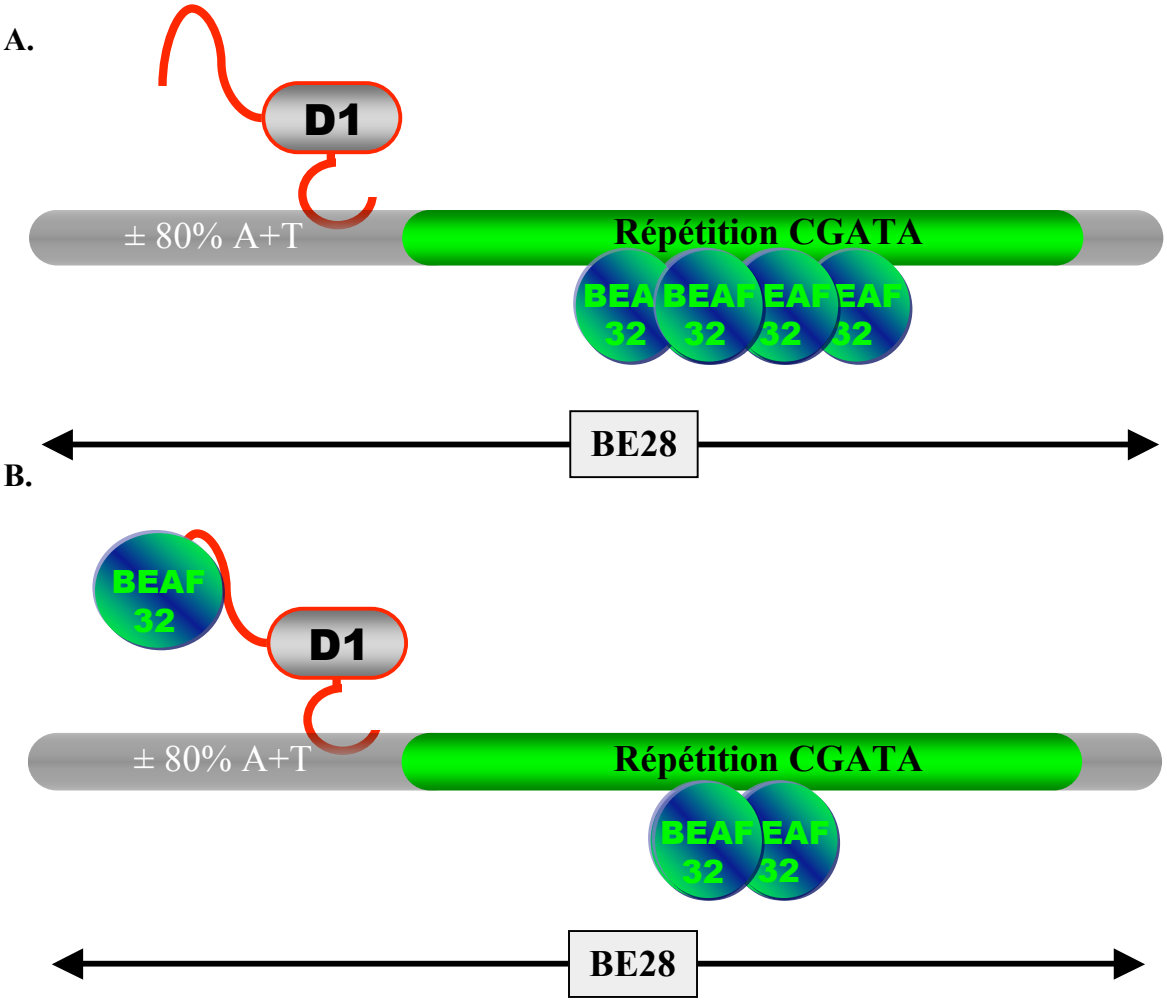


Figure 17 : la protéine BEAF32 et l'élément frontière BE28.

Régulation de l'activité de barrière de l'élément BE28 par les protéines BEAF32 et D1. La protéine BEAF32 est nécessaire à l'activité isolatrice de l'élément frontière BE28. Celui-ci possède des séquences palindromiques CGATA•TATCG reconnus par la protéine BEAF32. Sa liaison à l'élément BE28 est modulée par la protéine D1, capable de se fixer aux séquences riches en dA•dT adjacentes à l'élément BE28 (**A.**). L'interaction entre la protéine D1 et la protéine BEAF32 passe par le domaine C-terminal acide de la protéine D1, qui séquestrerait la protéine BEAF32 et l'empêcherait d'agir au niveau des séquences BE28 (**B.**).

de la structure hétérochromatinienne a été mis en évidence grâce à leur capacité à modifier le phénotype de variégation de position (PEV) de la lignée *white-mottled* de drosophile (Reuter et al., 1982; Reuter and Wolff, 1981). Des mutations générées dans cette lignée ont permi de définir deux classes : les gènes dont la mutation réduit le phénotype de variégation (gènes suppresseurs de variégation ou *Su(var)*) ou l'accentue (gènes enhancer de variégation ou *E(var)* (figure 18).

La plupart des gènes codant les protéines impliquées dans la formation de l'hétérochromatine présentent un phénotype d'haplo-suppressor/triplo-enhancer de PEV (Wustmann et al., 1989). Celui-ci correspond à une diminution de la PEV lorsque une copie du gène est éteinte, et à une augmentation de la PEV lorsque la transcription du gène est augmentée, par insertion d'un allèle supplémentaire. Certaines de ces protéines exercent une activité enzymatique au sein de l'hétérochromatine, d'autres n'y ont pas encore de fonction définie.

a. Les protéines à motifs AT Hook

Le motif AT Hook

Le motif AT Hook a été identifié pour la première fois dans la protéine chromosomique non histone HMG-A (anciennement dénommée HMG-I/Y) (Reeves and Nissen, 1990). Ce motif est retrouvée dans une centaine de protéines (Aravind and Landsman, 1998). Il peut constituer les seuls motifs remarquables d'une protéine, comme c'est le cas pour les protéines de la même famille que la protéine HMG-A. Le motif AT Hook peut aussi être associé à d'autres domaines fonctionnels, comme d'autres domaines de liaison à l'ADN (Bourachot et al., 1999) ou des domaines d'interaction entre protéines.

Le motif AT Hook est un motif de liaison à l'ADN constitué d'un cœur conservé formé des acides aminés Arginine-Glycine-Arginine-Proline (R-G-R-P). Ce motif consensus est entouré d'acides aminés basiques, essentiellement des lysines, ou des arginines. La résolution de la structure tertiaire de ce motif à été effectuée par spectroscopie de résonance magnétique nucléaire (RMN) (Huth et al., 1997). Les résultats indiquent que le squelette peptidique est en forme de croissant plan lorsqu'il est associé à l'ADN. Cette structure est similaire à celle de certaines drogues comme la distamycine A ou la nétropsine (Radic et al., 1992). Ces drogues interagissent spécifiquement avec le petit sillon de l'ADN riche en dA•dT

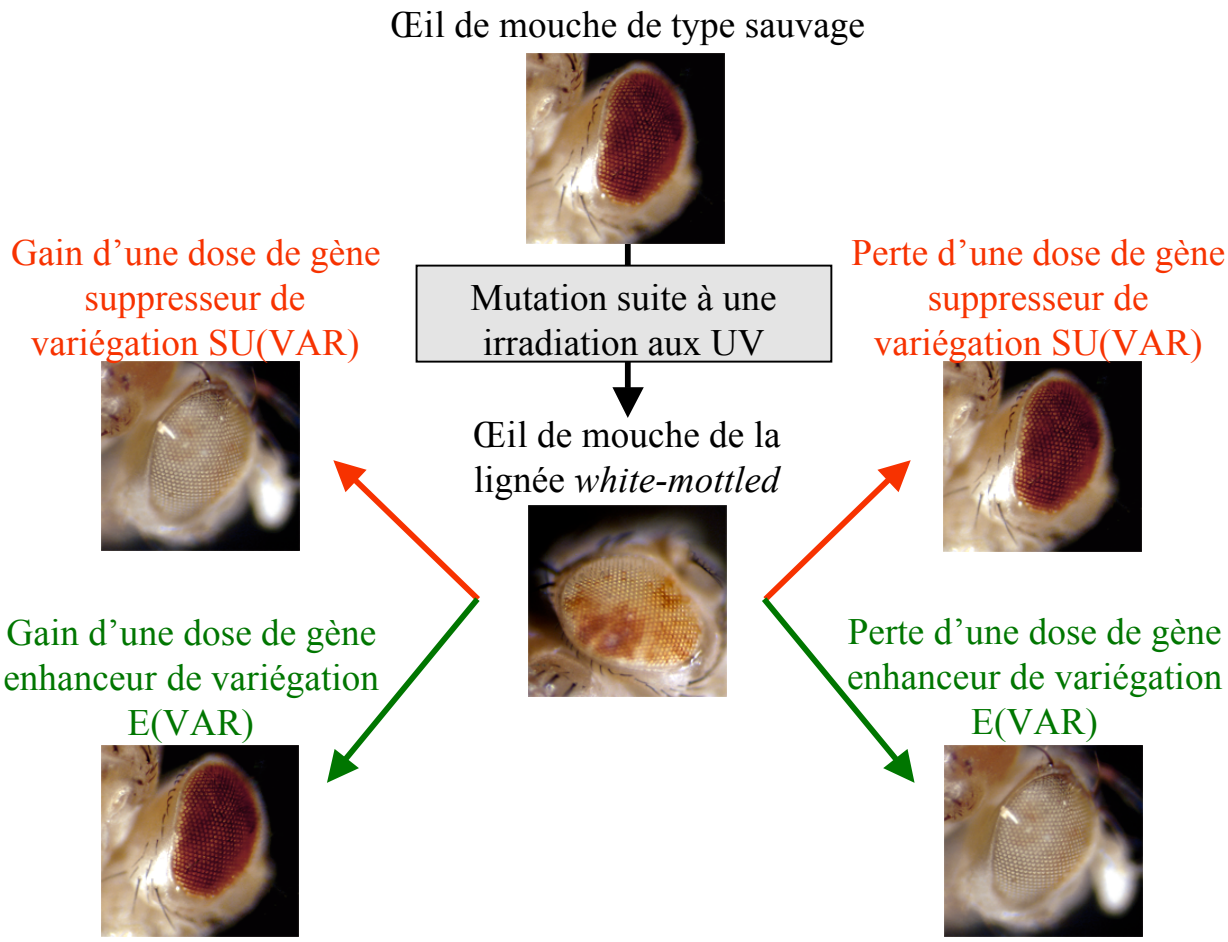


Figure 18 : les gènes modificateurs de la PEV.

Ce sont le plus souvent des gènes identifiés chez la drosophile et dont la mutation induit une modification de la variégation d'effet de position (PEV), par exemple de l'allèle *white-mottled* dans la lignée *In(1)wm4*.

La mutation d'un gène suppresseur de PEV (*Su(Var)*) conduisant à une haploinsuffisance (perte d'une dose) de son produit (SU(VAR)) induit une suppression de la PEV, correspondant à une restauration de l'expression de *white-mottled* et de la coloration rouge de l'œil (en haut, à droite). A l'inverse l'expression d'une copie surnuméraire de ce gène (gain d'une dose) renforce la mise en silence de *white-mottled* par l'hétérochromatine se traduisant par un œil majoritairement blanc (en bas, à droite).

A l'inverse, la mutation d'un gène enhancer de PEV augmente le *silencing* transcriptionnel alors que sa surexpression rétablit l'expression de *white-mottled.* Ceci suggère que les E(VAR) sont des protéines intervenant dans une régulation naturelle des propriétés fonctionnelles de l'hétérochromatine et pourraient par exemple être impliquée dans le contrôle de l'expression de gènes hétérochromatiniens.

(Geierstanger et al., 1994), particulièrement étroit et profond. Le motif AT Hook interagit avec le petit sillon sur une longueur de cinq à six paires de bases dA•dT consécutives (Huth et al., 1997).

L'interaction entre le motif AT Hook et l'ADN implique la présence du cœur peptidique canonique R-G-R-P (figure 19) (Reeves and Nissen, 1990). Les résidus basiques qui entourent ce motif conservé semblent intervenir dans cette reconnaissance en augmentant l'affinité et/ou la stabilité du complexe ADN/protéine (Geierstanger et al., 1994). De même, le nombre de motifs AT Hook présent sur un polypeptide influence aussi l'affinité de la liaison entre cette protéine et son substrat (Huth et al., 1997; Reeves and Nissen, 1990; Strick and Laemmli, 1995).

La protéine HMG-A

La protéine HMG-A appartient à la famille des protéines HMG (High Mobility Group). Cette famille regroupe un ensemble de protéines chromosomiques non histones (Bustin, 1999). Elles sont dénommées ainsi à cause de leur mobilité électrophorétique accentuée sur gel de polyacrylamide.

Chez les mammifères, les protéines HMG sont regroupées en trois familles : les familles HMG-B, HMG-N et HMG-A.

La famille HMG-B, anciennement dénommée HMG-1/2, est constituée des protéines HMG qui possèdent un motif fonctionnel appelé « boîte HMG » (Baxevanis and Landsman, 1995). Les protéines HMG-B lient le petit sillon de l'ADN et lui imposent une courbure accentuée (Love et al., 1995).

La famille HMG-N, aussi connue sous le terme HMG-14/17, regroupe les protéines HMG comportant un motif de liaison au nucléosome (Davie and Delcuve, 1991). Elles ont la particularité de reconnaître spécifiquement la structure nucléosomale (Bustin and Reeves, 1996).

Les protéines de la famille HMG-A, appelée auparavant HMG-I/Y, possèdent toutes le motif AT Hook dans leur structure (Huth et al., 1997). Elles ont la propriété d'intervenir sur la structure de l'ADN en prévenant ou en inversant ces distorsions internes (Falvo et al., 1995).

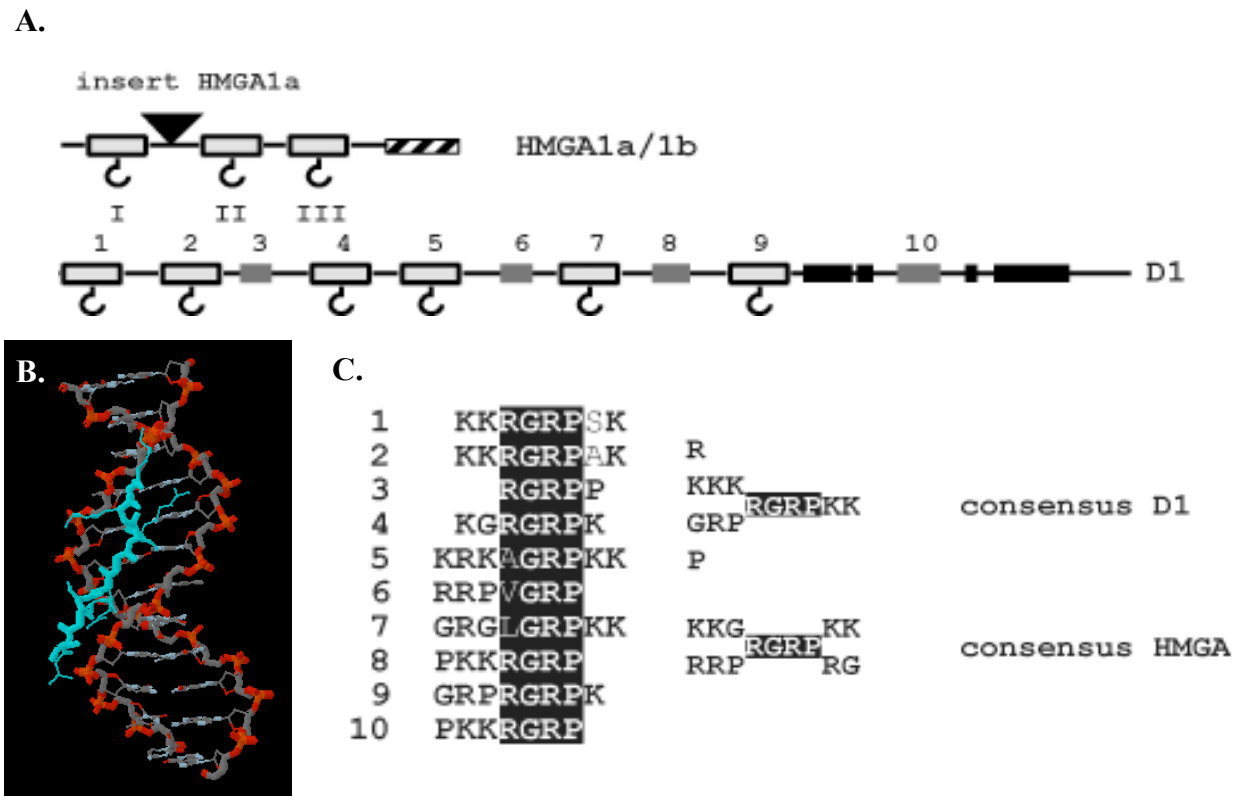


Figure 19 : le motif AT Hook et les protéines MATH.

Les protéines Multi-AT Hook (MATH) se caractérisent par la répétition du motif AT Hook capable d'interagir sélectivement avec les séquences riches en dA•dT. De manière générale, les protéines MATH ne possèdent pas d'autre motif peptidique remarquable si ce n'est un domaine carboxy-terminal (C-terminal) acide.

A. La protéine HMG-A humaine, archétype de la famille MATH, possède 3 motifs AT Hook (I, II, III) et un domaine C-terminal acide impliqué dans des interactions protéine/protéine. La protéine D1, qui pourrait être l'homologue fonctionnel de HMG-A chez la drosophile, possède 10 motifs AT Hook et un domaine C-terminal acide (d'après Aulner *et al*., 2002).

B. Représentation de l'interaction d'un motif AT Hook unique (en turquoise) avec le petit sillon d'une double hélice d'ADN riche en dA•dT (en rouge et gris).

C. Alignement des 10 motifs AT Hook de la protéine D1. Le consensus du motif, un peptide fortement basique, est constitué d'un cœur RGRP (Arginine, Glycine, Arginine, Proline) conservé à la fois dans les protéines HMG-A et D1. Ce motif permet la

Les trois motifs AT Hook des protéines HMG-A leur permettent de reconnaître préférentiellement l'ADN riche en dA•dT aussi bien *in vitro* qu'*in vivo* (Radic et al., 1992). Les séquences SAR/MAR, du fait de leur richesse en dA•dT, constituent une cible préférentielle de ces protéines aussi bien *in vitro* (Zhao et al., 1993), que *in vivo* (Saitoh and Laemmli, 1994).

Des études suggèrent que les protéines de cette famille reconnaissaient plus la structure de l'ADN que sa séquence propre (Reeves and Nissen, 1990). Elles sont ainsi capables de se lier à l'ADN enroulé autour de la particule nucléosomale (Reeves and Nissen, 1993). La liaison des protéines HMG-A induit alors des changements dans la conformation de l'ADN qu'elles reconnaissent (Slama-Schwok et al., 2000).

La liaison des protéines HMG-A avec l'ADN est modulée par la phosphorylation des motifs AT Hook (Nissen et al., 1991; Reeves and Nissen, 1995). Cette modification post-traductionnelle, réalisée par la kinase CDC2, a pour conséquence de diminuer l'affinité de ces protéines pour leur substrat (Reeves et al., 1991).

Les rôles biologiques des protéines HMG-A sont encore peu clairs. Pourtant, trois hypothèses sont émises. Elles ont été évoquées d'une part d'après des caractéristiques d'expression des protéines HMG-A, et d'autre part d'après les propriétés de liaison de ces protéines au sillon mineur de l'ADN riche en dA•dT et leur capacité à induire des changements structuraux.

Les protéines HMG-A sont fréquemment associées à la régulation, positive ou négative, de l'expression génique. Elles sont peu ou pas exprimées dans les cellules différenciées (Johnson et al., 1990). Par contre, leur niveau d'expression est élevé dans les cellules transformées (Ram et al., 1993). Pour autant, les gènes codant ces protéines ne se comportent pas comme des oncogènes classiques. En effet, leur surexpression n'induit pas la transformation cellulaire (Berlingieri et al., 1995).

Les protéines HMG-A interagissent avec les régions riches en dA•dT des séquences promotrices de certains gènes. Au niveau de ces séquences, les protéines HMG-A influencent l'activité d'autres facteurs de transcription (Thanos et al., 1993). Il semblerait donc que les protéines HMG-A agissent comme des facteurs architecturaux de transcription (Wolffe, 1994). Leur liaison, en modifiant la structure locale de l'ADN créerait les conditions favorables au recrutement des activateurs de transcription.

La modulation de l'accessibilité de la fibre de chromatine pourrait être une autre fonction des protéines HMG-A. Ce rôle s'appuierait essentiellement sur l'antagonisme qui

existe entre les protéines HMG-A et l'histone H1 pour la liaison au petit sillon des séquences riches en dA•dT. Il a été démontré *in vitro* que les protéines HMG-A se comportent comme des compétiteurs puissants de la liaison de l'histone H1 aux séquences SAR/MAR (Käs et al., 1989; Käs et al., 1993; Zhao et al., 1993). Ces résultats ont permis de postuler un modèle d'ouverture de la chromatine dans lequel les protéines HMG-A pourraient intervenir dans une étape précoce des mécanismes d'activation transcriptionnelle des gènes. Dans ce modèle, les séquences SAR/MAR se comportent comme des sites de nucléation de l'assemblage coopératif de l'histone H1 sur l'ADN (Izaurralde et al., 1989). Le domaine demeure ainsi transcriptionnellement inactif jusqu'à ce que la protéine HMG-A, en se fixant aux séquences SAR/MAR, chasse l'histone H1 de ces régions. Ce déplacement s'étend sur toute la boucle de chromatine en raison de la coopérativité des interactions entre les histones H1.

Les protéines HMG-A sont les archétypes d'une famille de protéine appelée « multi AT Hook », ou MATH. Cette famille comporte plusieurs protéines dont la protéine de drosophile D1, orthologue des protéines HMG-A.

La protéine D1

La découverte de la protéine D1 est conséquente à des expériences d'isolation d'histones à partir d'embryons de drosophile (Alfageme et al., 1974). Pour autant, les auteurs distinguent cette petite protéine de 37 kDa des autres protéines histones du fait de sa composition différente en acides aminés.

La protéine non-histone chromosomale D1 possède deux domaines remarquables : un enchaînement de 10 motifs AT Hook et un domaine C-terminal acide (figure 19) (Ashley et al., 1989; Reeves and Nissen, 1990). Ce dernier domaine pourrait intervenir dans les interactions entre la protéine D1 et des partenaires protéiques.

La protéine D1 présente des propriétés de liaison au petit sillon de l'ADN riche en dA•dT (Levinger and Varshavsky, 1982). La reconnaissance de ces séquences riches en dA•dT met en jeu ses 10 motifs AT Hook. Il a été démontrée qu'elle se localisait au niveau de l'hétérochromatine péricentromérique riche en dA•dT (Levinger and Varshavsky, 1982), et qu'elle reconnaissait plus particulièrement les séquences SAT III alphoïdes (Rodriguez Alfageme et al., 1980). La protéine D1 est alors considérée comme le premier exemple de protéine qui pourrait être impliquée dans la formation de la structure compacte des régions hautement répétées constituant l'hétérochromatine (Levinger, 1985). Cette interaction avec les

séquences riches en dA•dT pourraient être modulée par la phosphorylation de la protéine D1, qui présente six sites potentiels de modification dans sa partie C-terminale.

L'hypothèse selon laquelle la protéine D1 serait impliquée dans la modification de l'effet de position de l'allèle *white-mottled* a été émise suite à la mise en évidence d'un tel effet dans des expériences utilisant la protéine artificielle MATH 20 (Girard et al., 1998). Cette dernière est constituée de 20 motifs AT Hook dérivés de ceux de la protéine HMG-A et ce, afin de reconnaître spécifiquement les séquences SAR/MAR. Les premiers résultats obtenus grâce à l'utilisation de la protéine MATH 20 démontrent qu'elle a de profonds effets sur la structure et la dynamique des chromosomes (Strick and Laemmli, 1995), mais aussi sur l'expression génique par modification de la structure chromatinienne (Satou et al., 2004; Suzuki et al., 2001). Lorsqu'elle est surexprimée, la protéine MATH 20 est capable de supprimer la PEV du gène *white* dans la lignée *white-mottled* (Girard et al., 1998). Ces résultats supposaient alors un rôle de MATH 20 au niveau des séquences SAT III, sans pour autant identifier les protéines cibles dont l'activité serait affectée par la présence de cette protéine.

Les premiers résultats de l'équipe concernant la protéine D1 avaient dans un premier temps permis de démontrer que celle-ci était essentielle et exprimée de manière ubiquitaire tout au long du développement de la drosophile (Aulner et al., 2002). Sa capacité à interagir avec les séquences SAT III alphoïdes et les résultats démontrant l'effet de suppresseur de la variégation de la lignée *white-mottled* par la protéine MATH 20 permettaient d'envisager un rôle de la protéine D1 dans ce même phénomène.

Ainsi, il a été montré que la surexpression de la protéine D1 augmente la PEV de la lignée *white-mottled* (Aulner et al., 2002), alors que la surexpression de la protéine synthétique MATH 20 supprimait la PEV (Girard et al., 1998). Afin d'expliquer ces effets antagonistes, il a été proposé que la protéine D1 serait une cible de la protéine MATH 20, et que ces deux protéines rentreraient en compétition pour la liaison aux séquences SAT III. Des études complémentaires ont confirmé cette hypothèse et il a donc été proposé que la protéine D1 augmente, alors que la protéine MATH 20 diminue l'effet de mise en silence médié par la proximité des séquences SAT III dans l'inversion (Monod et al., 2002).

Des études complémentaires ont été menées en utilisant la protéine D1ΔE. Celle-ci correspond à la protéine D1 délétée de son domaine C-terminal acide. La surexpression de la protéine D1ΔE supprime aussi la PEV de l'allèle *white-mottled*, tout comme lors de

l'expression de la protéine artificielle MATH20, exerçant un effet dominant négatif sur la protéine D1 endogène. Il a donc été proposé que le domaine C-terminal de la protéine D1 puisse servir à recruter les protéines requises pour la mise en place de la structure hétérochromatinienne au niveau des séquences SAT III (Aulner et al., 2002).

b. Les mimétiques du motif AT Hook

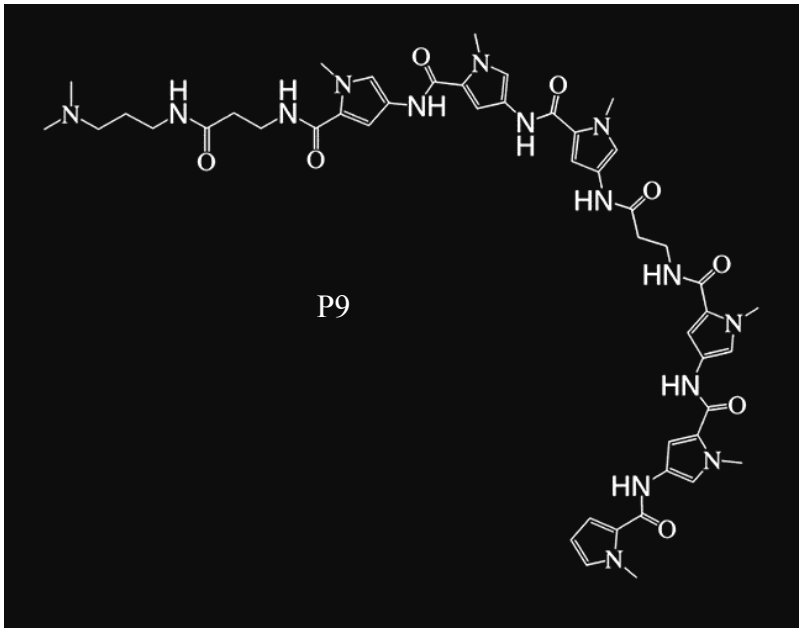
Les mimétiques du motif protéique AT Hook sont de deux catégories. Ce sont soit des protéines synthétiques (voir la protéine synthétique MATH 20 dans la partie précédente), soit des drogues qui se lient spécifiquement au petit sillon de l'ADN riche en dA•dT. La plupart de ces drogues ont été synthétisées par analogie avec la distamycine, un oligopeptide naturel produit par *Streptomyces distallicus* (Hopwood, 1988).

Cet antibiotique est capable de s'insérer dans le petit sillon des séquences riches en dA•dT, comme par exemple les séquences SAR/MAR. En se fixant sur ces séquences, il est capable d'en dissocier l'histone H1 (Käs et al., 1989) et la Topoïsomérase II (Adachi et al., 1989). Des composés synthétiques ont alors été générés dans le but d'imiter les propriétés de liaison à l'ADN de cette molécule et ce afin de mieux caractériser les mécanismes d'action de cette molécule (figure 20). Les propriétés avérées de liaison du petit sillon de l'ADN riche en dA•dT par ces molécules en ont ensuite fait des outils puissants pour l'étude de l'hétérochromatine.

Les polyamides

Cette série de drogues, les « polyamides », a été synthétisée dans un premier temps dans le laboratoire de Peter Dervan, pour revue (Wemmer and Dervan, 1997), qui a mis au point un code d'enchaînement de structures chimiques, des noyaux pyrroles et imidazoles, permettant de reconnaître spécifiquement le petit sillon de séquences d'ADN respectivement riches en dA•dT ou dG•dC. Leurs propriétés de liaison à l'hétérochromatine ont été étudiées principalement par le groupe d'Ulrich K. Laemmli. Les drogues synthétisées et étudiées par ce groupe, le polyamide P9 et le polyamide P31, avaient pour cible les séquences SAT I, SAT III, pour le P9, et SAT V, pour le P31, de la drosophile (Janssen et al., 2000a; Janssen et al., 2000b).

A.



B.

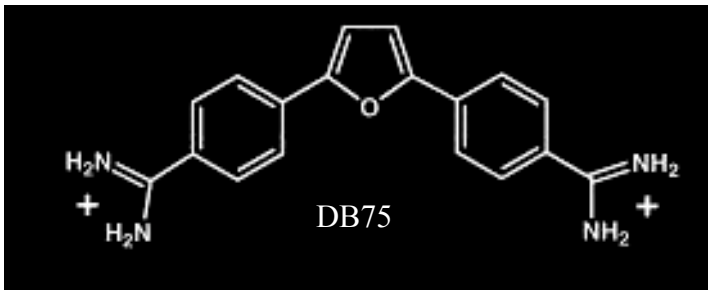


Figure 20 : les mimétiques du motif AT Hook.

A. Le polyamide P9 est constitué d'un enchaînement de noyaux pyrroles. Sa structure forme un croissant plan, tout comme le motif AT Hook. Il est capable de se lier au petit sillon de l'ADN riche en dA•dT et ainsi de rentrer en compétition avec les protéines s'y liant grâce à des motif AT Hook (d'après Janssen *et al*., 2000a).

B. La diamidine DB75 est une molécule de type diphénylfuran. Sa structure lui confère, tout comme les polyamides, une forme en croissant plan. Il a été aussi démontré que ces molécules étaient capables d'interagir avec le petit sillon de l'ADN riche en dA•dT (d'après Lansiaux *et al*., 2002b).

Les diamidines

Les diamidines sont des molécules synthétisées dans le laboratoire de David W. Boykin et de W. David Wilson. Elles possèdent pour la plupart des activités antiparasitaires avérées contre *Plasmodium falciparum, Pneumocystis carinii* et *Trypanosoma brucei*. Ces agents infectieux sont responsables respectivement du paludisme, de la pneumonie à PCP (une forme de pneumonie opportuniste touchant principalement les patients atteints du SIDA) et de la maladie du sommeil (Bouteille et al., 2003; Boykin et al., 1996; Fairlamb, 2003; Tidwell et al., 2003).

Ces drogues, à l'instar du polyamide P9, sont affines pour les séquences d'ADN riches en dA•dT ce qui explique leur efficacité contre ces parasites (Susbielle et al., 2005). En effet, les génomes de *P. falciparum* et *P. carinii* sont riches en paires de bases dA•dT respectivement à 70% et 80%. De plus, les kinétoplastes des trypanosomes ou de Leishmania, structures d'ADN caractéristiques de ces organismes codant pour les protéines ribosomiques et celles de la chaîne respiratoire, ont également une composition nucléotidique fortement biaisée en faveur des paires de bases dA•dT (Bailly et al., 1999).

Tout comme les polyamides, les diamidines sont capables d'interagir spécifiquement, mais pas exclusivement, avec le petit sillon des séquences riches en paires de bases dA•dT, dont les séquences satellites représentent un exemple extrême. Elles représentent aussi un outil cytologique puissant grâce à leurs propriétés d'autofluorescence lorsqu'elles sont liées à la séquence d'ADN cible (Wilson et al., 2005).

c. Les protéines à motifs « chromo-domain »

Les motifs « chromo-domain » et « chromo-shadow-domain »

Le motif « chromo-domain » (CD), pour « CHRomatin Organization MOdifier », est retrouvé dans certaines protéines interagissant avec la chromatine (Kellum, 2003b). Il a été découvert grâce à une homologie de 44 acides aminés fortement enrichi en cystéines entre les domaines C-terminaux des protéines Pc (Polycomb) et HP1 (Paro and Hogness, 1991). Sa structure tridimensionnelle a été résolue par RMN (Ball et al., 1997) (figure 21) : elle se compose de trois feuillets bêta empaquetés contre une hélice alpha (Bottomley, 2004). La

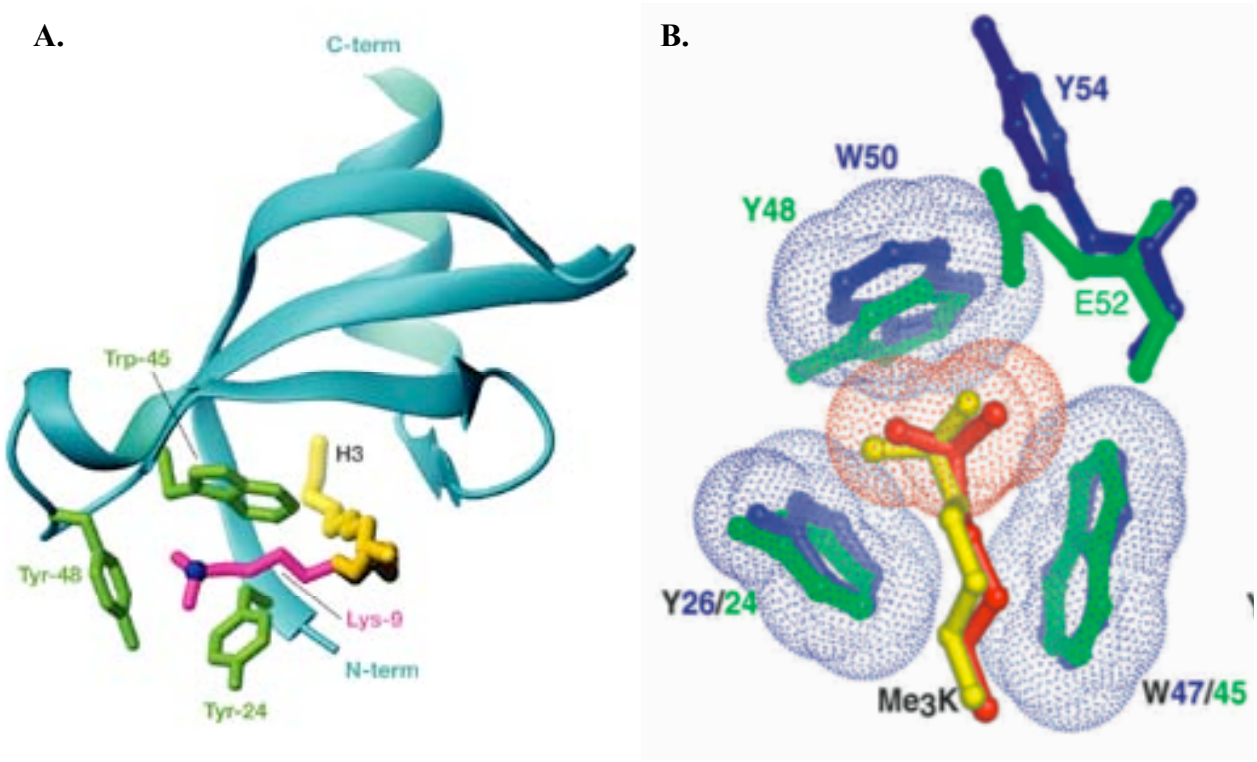


Figure 21 : le « chromo-domain ».

Le « chromo-domain » (CD) est un court motif peptidique enrichi en cystéines permettant la reconnaissance spécifique des résidus lysines méthylées.

A. Représentation schématique de la structure du CD d'HP1β (cyan) lié à un peptide correspondant aux résidus 5 à 10 de l'extrémité amino-terminale de l'histone H3 (squelette couleur or, chaîne latérale de la lysine 9 méthylée de l'histone H3 en magenta). L'atome d'azote en bleu occupe une place centrale au sein de la cage aromatique formée par trois résidus tyrosines qui permet la reconnaissance des lysines di- ou triméthylées (d'après Bottomley *et al*., 2004).

B. Superposition des cages aromatiques des CD d'HP1 et de Polycomb (Pc) (respectivement en vert et en bleu) liées aux lysines 9 et 27 méthylées de l'histone H3 (en jaune et rouge respectivement). Les rayons de van der Waals des anneaux aromatiques de Polycomb et du groupement methylammonium de la lysine 27 de l'histone H3 sont représentés par les trames bleue et rouge respectivement (d'après Jacobs *et al*., 2002).

forte charge négative du CD suppose qu'il intervient plutôt dans les interactions avec d'autres protéines, qu'avec l'ADN (Eissenberg and Elgin, 2000).

La plupart des protéines possédant des CD reconnaissent la structure chromatinienne, et plus précisément les histones modifiées (Koonin et al., 1995), et sont aussi impliquées dans la régulation transcriptionnelle négative de gènes (Eissenberg et al., 1992). Cette régulation est exercée à la fois au niveau de l'hétérochromatine, par les protéines à domaine CD HP1 et SU(VAR)3-9, et de l'euchromatine, via les protéines du groupe Polycomb (Pc-G). Il est à noter que les protéines de l'hétérochromatine ont souvent un CD associé à un « chromo-shadow-domain » (CSD).

Le motif CSD, structurellement proche du CD, est impliqué dans les interactions entre protéines, et plus particulièrement dans la dimérisation de protéines possédant chacune un CSD, comme HP1 (Aasland and Stewart, 1995). La présence de ce domaine permet la localisation des protéines le portant au niveau de l'hétérochromatine (Aasland and Stewart, 1995).

La protéine HP1

Les protéines de la famille HP1 sont retrouvées dans différents organismes, de la levure *S. pombe* à l'homme. Ce sont de petites protéines de 15 kiloDalton (kDa) à 35 kDa (Eissenberg and Elgin, 2000). Elles prennent le nom de Swi6, chez *Schizosaccharomyces pombe*, HP1 chez la souris, l'homme et la drosophile (Kellum, 2003b). Dans cet organisme, la protéine HP1 est parfois appelée SU(VAR)2-5, en référence au gène qui la code (Sinclair et al., 1983). Les protéines de la famille HP1 ont d'abord été mises en évidence chez la drosophile grâce à leur capacité de supprimer la PEV du gène *white* (Eissenberg et al., 1992), probablement en recrutant des partenaires protéiques (Seum et al., 2001). Des expériences de mutations du gène *su(var)2-5* ont démontré que la protéine HP1 était essentielle chez la drosophile, la létalité survenant relativement tardivement, à la fin du troisième stade larvaire (Lu et al., 2000).

Les protéines HP1 ont pour caractéristiques principales de posséder un CD N-terminal et un CSD C-terminal (Eissenberg and Elgin, 2000). Le CD reconnaît spécifiquement la lysine 9 méthylée de l'histone H3 (Bannister et al., 2001; Jacobs and Khorasanizadeh, 2002;

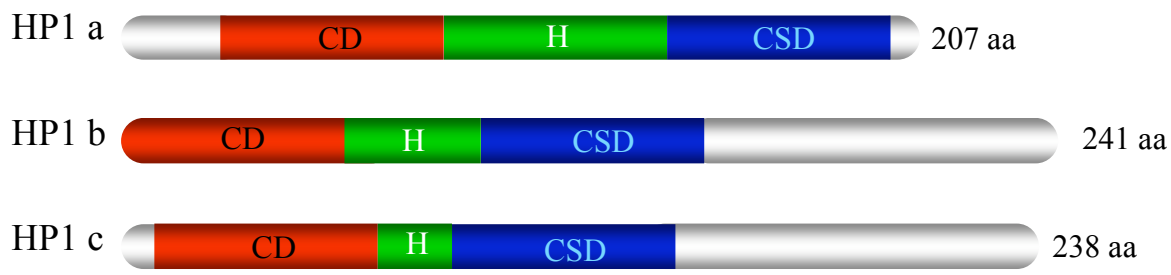
Lachner et al., 2001). Cette modification est une marque de l'hétérochromatine, au même titre que la protéine HP1, selon la théorie du « code des histones » (Strahl and Allis, 2000).

Le CSD serait impliqué dans la dimérisation des protéines HP1 (Brasher et al., 2000), grâce à la reconnaissance d'un enchaînement de cinq acides aminés PXVHH (Smothers and Henikoff, 2000). Cette homodimérisation permet aux protéines HP1 de reconnaître deux histones H3 méthylées sur leurs lysines 9 adjacentes. Le CSD interviendrait aussi dans les interactions avec les protéines partenaires de la protéine HP1, comme le récepteur de la Lamine B (Ye et al., 1997), l'histone méthyltransférase SU(VAR)3-9 (Aagaard et al., 1999; Greil et al., 2003), et la protéine à doigt de zinc SU(VAR)3-7 (Cléard et al., 1997).

Outre les motifs CD et CSD, les protéines HP1 possèdent un pont peptidique, appelé « hinge », qui relie ces deux domaines. Ce domaine hinge aurait la capacité de reconnaître l'ARN (Muchardt et al., 2002), l'ADN (Nielsen et al., 2001a; Smothers and Henikoff, 2001), mais aussi des protéines comme la protéine centromérique INCENP (Ainsztein et al., 1998). Les propriétés de liaison du hinge suggèrent l'existence de plusieurs parties dans ce domaine. Du fait de sa variabilité, il serait responsable de la différence de localisation existant entre les isoformes de la protéine HP1. En effet, la protéine HP1 est présente sous trois isoformes chez la drosophile et chez les mammifères appelées respectivement HP1a, HP1b, HP1c, et HP1α, HP1β, HP1γ. Chez ces derniers, les isoformes HP1α et HP1β sont localisées au niveau de l'hétérochromatine péricentromérique (Minc et al., 2001), tout comme l'isoforme HP1a de la drosophile. L'isoforme HP1γ est quant à elle principalement associée à l'euchromatine (Minc et al., 1999; Minc et al., 2000), tout comme les isoformes HP1b et HP1c de la drosophile (figure 22).

Les fonctions des protéines HP1 ne sont pas encore clairement établies. La présence de ces protéines est nécessaire à la formation et au maintien de la structure hétérochromatinienne (Maison and Almouzni, 2004). Dans certaines cas, elles pourraient cependant ne pas s'avérer nécessaires pour la mise en place d'une telle structure (Gilbert et al., 2003). Les protéines HP1 seraient impliquées dans la régulation de la transcription de certains gènes. Ainsi, la protéine HP1α jouerait un rôle à la fois dans la répression transcriptionnelle (Cheutin et al., 2003), mais aussi dans l'activation transcriptionnelle (Piacentini et al., 2003). L'isoforme HP1γ interviendrait dans l'activation transcriptionnelle au niveau de l'euchromatine (Vakoc et al., 2005).

A.



B.

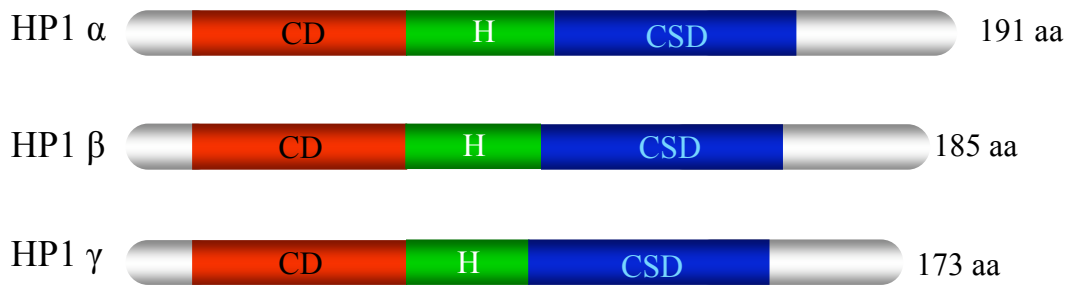


Figure 22 : les isoformes de la protéine HP1.

De manière générale, les protéines HP1 sont de petites protéines de moins de 35 kDa. Elles comportent toutes un CD, et un « chromo-shadow-domain » (CSD), séparés par un domaine « hinge ». Ce dernier pourrait intervenir dans la spécificité d'interaction entre chaque isoforme et sa cible.

A. Représentation schématique de la structure des différentes isoformes de la protéine HP1 de drosophile.

B. Représentation schématique de la structure des différentes isoformes de la protéine HP1 humaine.

« Chromo-Domain » (CD)

« Hinge » (H)

« Chromo-Shadow-Domain » (CSD)

Les données récentes montrent que les protéines HP1 lieraient l'ARN, interaction qui interviendrait dans la mise en place d'une structure de type hétérochromatinienne (Pal-Bhadra et al., 2004). Cette initiation de la formation de la structure hétérochromatinienne impliquerait des petits ARN de la famille des siRNA (small interfering RNA) (Jenuwein, 2002). Ces siRNA seraient capable de générer localement une méthylation de la lysine 9 de l'histone H3 via un complexe protéique appelé RISC (RNA Induced Silencing Complex (Hammond et al., 2000)). Cette formation locale d'hétérochromatine a déjà été caractérisée au niveau des séquences répétées des régions centromériques (Allshire, 2002) et des loci sexuels chez la levure à fission (Hall et al., 2002; Volpe et al., 2002).

Pour autant, les données actuelles envisagent l'implication d'une autre protéine à CD dans la reconnaissance des ces petits ARN : l'histone méthyltransférase Clr4, orthologue de la protéine SU(VAR)3-9 chez la levure à fission *Schizosaccharomyces pombe* (Jenuwein, 2002). Cette protéine serait capable de lier les petits ARN et de modifier directement les lysines 9 des histones H3, permettant ainsi de créer un site de fixation pour la protéine Swi6 et de favoriser ainsi l'extension/propagation d'une structure hétérochromatinienne.

La protéine SU(VAR)3-9

La famille des histones méthyltransférases SU(VAR)3-9 est présente dans tous les organismes, sauf chez la levure *Saccharomyces cerevisae*. Ces protéines prennent les noms de Clr4 chez la levure *Schizosaccharomyces pombe* (Ekwall and Ruusala, 1994), SU(VAR)3-9 chez la drosophile (Tschiersch et al., 1994), suv39h1 et suv39h2 chez la souris (Aagaard et al., 1999), et Suv39H1 et Suv39H2 chez l'homme (Aagaard et al., 1999) (figure 23). La première protéine de la famille a avoir été identifiée est la protéine de drosophile SU(VAR)3-9. Elle a été caractérisées grâce à sa capacité de suppression de la PEV du gène *white* dans la lignée *white-mottled* (Tschiersch et al., 1994).

Toutes ces protéines ont une caractéristique commune : elles possèdent à la fois un CD leur permettant d'interagir avec la lysine 9 méthylée de l'histone H3, et un domaine catalytique appelé « SET domain » (Jenuwein et al., 1998), responsable de la méthylation de ce même résidu lysine (Rea et al., 2000). Il est retrouvé à la fois dans les protéines SU(VAR)3-9, E(Z) et Trithorax (trx), d'où son nom (Jones and Gelbart, 1993; Stassen et al., 1995; Tschiersch et al., 1994).

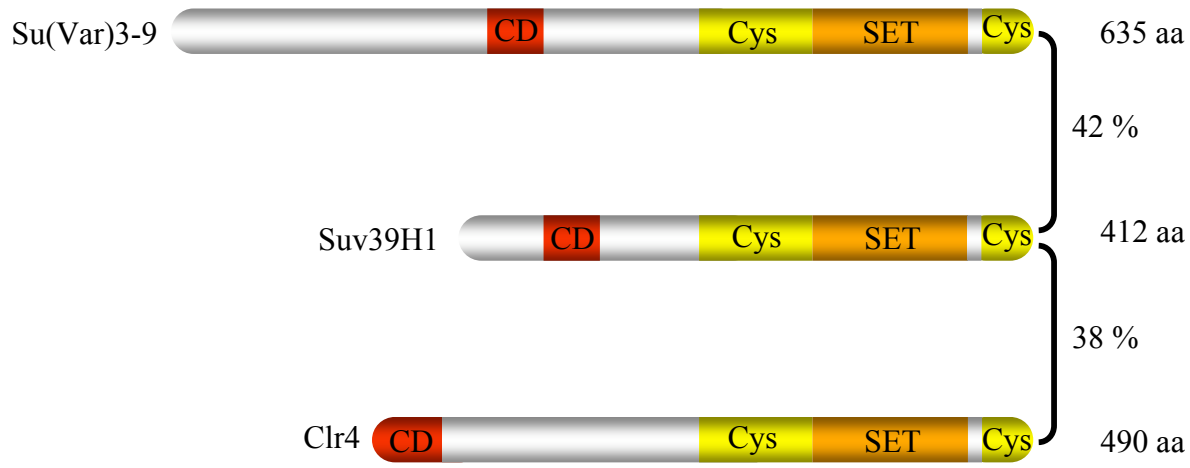


Figure 23 : les homologues de la protéine Su(Var)3-9.

Les protéines de la famille de Su(Var)3-9 possèdent toutes un domaine catalytique appelé le domaine SET, associé à un domaine riche en cystéines. Ce domaine SET est souvent associé à un « chromo-domain », permettant les interactions avec les lysines 9 de l'histone H3 qu'il méthyle. Le « chromo-domain » pourrait aussi intervenir dans les interactions entre les protéines de la famille Su(Var)3-9 et les protéines HP1. Les pourcentages d'identité entre ces différentes protéines sont indiqués.

« Chromo-domain » (CD)

Domaine riche en cystéine associé au domaine SET (Cys)

Domaine SET

L'étude de la protéine Clr4 a mis en évidence l'importance du domaine SET et du CD pour le maintien du silence transcriptionnel au niveau des centromères et des loci sexuels (Ivanova et al., 1998). Une étude de la relation structure/fonction de la protéine Suv39H1 montre que cette protéine est localisée au sein de l'hétérochromatine péricentrique et qu'elle colocalise avec la protéine HP1 (Melcher et al., 2000).

De par leur activité de méthylation de la lysine 9 de l'histone H3, les protéines de la famille SU(VAR)3-9 sont impliquées dans la formation de la structure hétérochromatinienne. En effet, la méthylation de la lysine 9 de l'histone H3 est spécifiquement reconnue par la protéine HP1 via son CD (Bannister et al., 2001; Nielsen et al., 2001b). Des interactions directes entre les protéines Suv39H1 et HP1 peuvent avoir lieu par l'intermédiaire de son CSD. Ces résultats semblent donc démontrer que, la modification mise en place, la méthylation de la lysine 9 de l'histone H3 est capable de recruter l'enzyme qui en est responsable, suggérant un rôle majeur de la protéine Suv39H1 dans la propagation de l'hétérochromatine.

Chez la levure à fission *Schizosaccharomyces pombe*, les orthologues des protéines Suv39H1 et HP1, respectivement les protéines Clr4 et Swi6, ont également été impliquées dans l'établissement et la propagation de l'hétérochromatine. Des centres initiateurs, appelés répétitions cenH, seraient capables de recruter les protéines Swi6 et Clr4 et ainsi d'entraîner localement la méthylation des lysines 9 des histones H3. Cette modification se propagerait grâce à la coopération entre les deux protéines (Hall et al., 2002).

Cependant, les protéines de la famille SU(VAR)3-9 ne sont pas les seules enzymes capables de méthyler la lysine 9 de l'histone H3. Dans les souris dont les gènes codant Suv39h1 et Suv39h2 ont été inactivés, la lysine 9 de l'histone H3 est toujours présente sous forme méthylée. Ces modifications sont mises en place par les protéines G9a (Tachibana et al., 2001; Tachibana et al., 2002), SETDB1 (Schultz et al., 2002) ou ESET (Blackburn et al., 2003; Dodge et al., 2004).

La particularité de la protéine Suv39H1 est qu'elle triméthyle la lysine 9 de l'histone H3, modification caractéristique de l'hétérochromatine péricentrique, contrairement à la protéine G9a qui mono- ou diméthyle ce même résidu (Rice et al., 2003). Dans les cellules qui n'expriment plus la protéine Suv39H1, la triméthylation de la lysine 9 est perdue au profit

d'une monométhylation, alors que dans les cellules n'exprimant plus la protéine G9a, la triméthylation de la lysine 9 n'est pas affectée (Peters et al., 2003).

La triméthylation de la lysine 9 de l'histone H3 semble également jouer un rôle important au niveau des télomères chez les mammifères. Chez les souris qui n'expriment plus la protéine Suv39H1, les télomères sont plus longs que dans les cellules normales. Des études ont démontré que ces télomères allongés n'arboraient plus une triméthylation de la lysine 9 de l'histone H3 mais un monométhylation (Garcia-Cao et al., 2004).

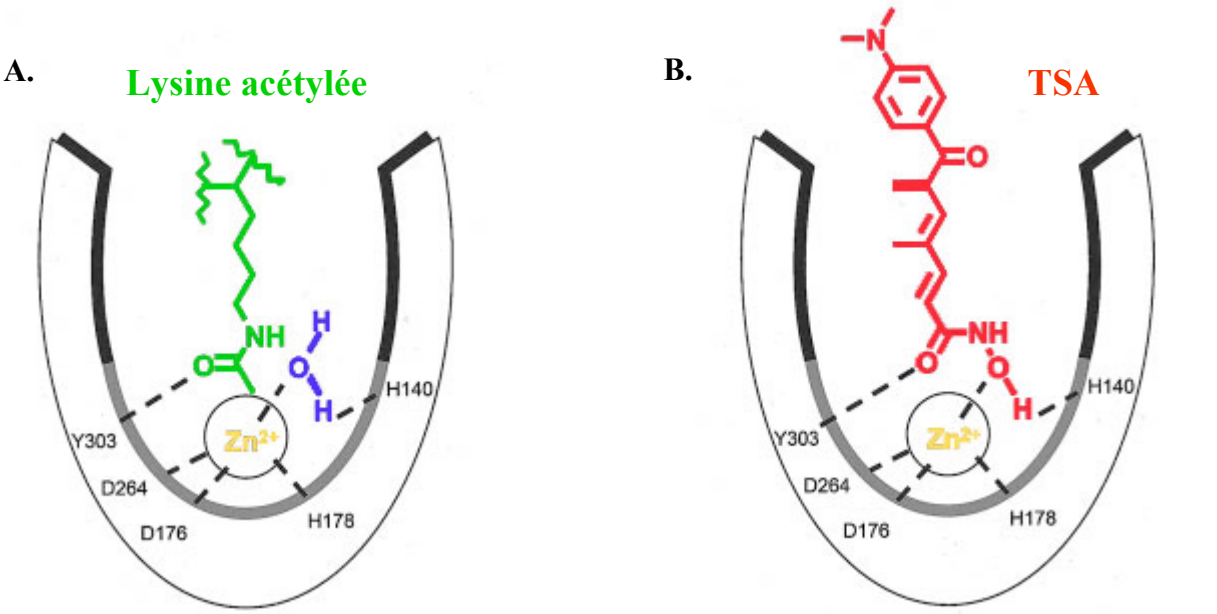
Un lien entre méthylation des histones et méthylation de l'ADN a été également mis en évidence chez les mammifères. En effet, la méthylation de la lysine 9 de l'histone H3 par Suv39H1 est un pré-requis pour la méthylation de l'ADN au sein de l'hétérochromatine péricentrique et des séquences satellites (Lehnertz et al., 2003). De même, la protéine G9a joue un rôle dans la mise en place de l'empreinte génomique parentale par méthylation de l'ADN au niveau du locus du centre de l'empreinte du syndrome de Prader-Willis (Xin et al., 2003).

d. Les autres protéines ligands de l'hétérochromatine

De nombreuses protéines interagissent de manière spécifique ou non avec la structure hétérochromatinienne. Nous nous concentrerons sur les protéines qui sont impliquées dans la mise en place de la structure hétérochromatinienne, ou dans la structuration des chromosomes.

La protéine HDAC1

La protéine HDAC1 fait partie de la famille des Histones DésACétylases (HDAC) dont il existe trois classes (Peterson, 2002c). Les HDAC des classes I et II sont dites classiques et comportent 11 membres. Elles ont la particularité de présenter un site catalytique conservé et un domaine appelé « bromo-domain » (Haynes et al., 1992; Tamkun et al., 1992), capable de reconnaître les queues N-terminales acétylées des histones (Ladurner et al., 2003) (figure 24). Les HDAC de classe III ont été les dernières découvertes et sont de ce fait moins caractérisées.



« Bromo-domain » de la protéine HDAC1 « Bromo-domain » de la protéine HDAC1

Figure 24 : le « bromo-domain » des HDAC.

Ces schémas représentent le site catalytique de l'histone désacétylase HDAC1. En gris, la zone riche en acides aspartiques et histidines qui sert de relais de charge à l'ion Zn^{2+}, nécessaire à la catalyse de la désacétylation des lysines.

A. Interaction entre une lysine acétylée (en vert), une molécule d'eau (en bleu) et l'ion Zn^{2+} (jaune) du site catalytique de l'histone désacétylase.

B. Interaction entre l'ion Zn^{2+} (en jaune) du site catalytique et la Trichostatine A (TSA, en rouge), inhibiteur des histones désacétylases.

Les HDAC de type I sont homologues à la protéine Rpd3, un régulateur transcriptionnel de la levure *Saccharomyces cerevisae*. Cette classe comprend quatre membres dont les protéines HDAC1, HDAC2 et HDAC3.

Les protéines HDAC1 et HDAC2 sont très similaires. Elles présentent un domaine catalytique dans leur partie N-terminale, et une séquence de localisation nucléaire (NLS) (de Ruijter et al., 2003). *In vivo*, elles sont associées aux protéines RbAp46 et RbAp48, avec qui elles forment le complexe « core deacetylase », fréquemment retrouvé dans les complexes de répression transcriptionnelle tels que Sin3, NuRD (Nucleosome Remodelling and Deacetylases) ou Co-REST (Heinzel et al., 1997; Zhang et al., 1999).

Ces complexes protéiques peuvent être recrutés par des facteurs de transcription. Par exemple, le complexe Sin3 reconnaît la protéine E2F qui est liée au niveau des promoteurs à réprimer (Rayman et al., 2002). Ce recrutement va entraîner une désacétylation des queues N-terminales des histones, et ainsi une répression de la transcription (Ahringer, 2000). Il en est de même pour le complexe NuRD, qui est impliqué dans la répression stable de certains loci faisant intervenir les protéines du Pc-G (Ahringer, 2000).

La protéine HDAC3 est la seule protéine de la famille à ne pas être strictement nucléaire (Emiliani et al., 1998). De même, elle est la seule capable à s'oligomériser avec les autres HDAC (Yang et al., 2002).

Les HDAC de type II sont homologues à la protéine de levure Hda1. Cette classe comprend sept membres qui sont aussi impliqués dans la répression transcriptionnelle, tout comme les HDAC de type I.

Les HDAC de type III sont homologues à la protéine Sir2 qui appartient à la famille des protéines SIR découvertes chez *Saccharomyces cerevisae* (Gartenberg, 2000). Elles sont requises pour la répression transcriptionnelle des séquences d'ADN répétées comme les loci de typage sexuel HML/HMR (Rine and Herskowitz, 1987), les télomères (Ng et al., 2002) et le rDNA (Bryk et al., 1997; Smith and Boeke, 1997).

De manière intéressante, la protéine HDAC1 intervient dans la mise en place de la structure hétérochromatinienne, (Taddei et al., 2001), en promouvant la désacétylation de la lysine 9 acétylée de l'histone H3, étape préliminaire à la méthylation de ce résidu par la protéine SU(VAR)3-9. Tout comme le gène *Su(var)3-9*, le gène codant la protéine HDAC1 fait partie du groupe des suppresseurs de PEV du gène *white* de l'allèle *white-mottled 4*

(Mottus et al., 2000). De plus, il a été montré que les protéines HDAC1 et SU(VAR)3-9 interagissaient de manière physique et fonctionnelle (Czermin et al., 2001; Vaute et al., 2002).

Outre son rôle dans la formation de structures répressives à grande échelle, la protéine HDAC1 est aussi impliquée dans la répression de la transcription du rDNA (Zhou et al., 2002). De même, la protéine HDAC1 jouerait un rôle dans le contrôle de la prolifération des cellules de mammifères en réprimant des gènes codant pour des inhibiteurs du cycle cellulaire (Lagger et al., 2002).

La protéine SU(VAR)3-7

La fonction précise de cette protéine de drosophile est à l'heure actuelle incomprise. Le gène *Su(var)3-7* code une protéine volumineuse de 175 kDa (Reuter et al., 1990), qui est associée à l'hétérochromatine péricentromérique, aux télomères et à quelques sites euchromatiniens sur des chromosomes polytènes interphasiques (Cléard et al., 1997). La protéine SU(VAR)3-7 interagit avec l'ADN grâce à la présence de sept motifs en « doigt de Zinc » dans sa partie amino-terminale (Cléard et al., 1997; Cléard et al., 1995; Reuter et al., 1990). Les cibles principales de cette protéine semblent être principalement les séquences SAT I, formés de répétitions $(AATAT)_n$, de l'hétérochromatine péricentromérique de la drosophile (Cléard and Spierer, 2001). Dans une moindre mesure, la protéine SU(VAR)3-7 pourrait interagir avec les séquences SAT III, constituées par la répétition d'un motif de 359 pb majoritairement riche en dA•dT. Des expériences réalisées *in vivo* ont confirmé que la protéine SU(VAR)3-7 pouvait lier directement l'ADN (Cléard and Spierer, 2001; Perrini et al., 2004).

La partie C-terminale de la protéine SU(VAR)3-7 serait impliquée dans son auto-association mais aussi dans les interactions avec l'hétérochromatine péricentromérique (Jaquet et al., 2002). Des expériences de double-hybride menées chez la levure ont démontré que la protéine SU(VAR)3-7 interagissait *in vitro* (Schotta et al., 2002) et *in vivo* (Delattre et al., 2004) avec la protéine SU(VAR)3-9. De même, il a été montré que la protéine SU(VAR)3-7 pouvait interagir génétiquement et physiquement avec la protéine HP1 (Cléard et al., 1997; Delattre et al., 2000). Pour autant, le rôle de la protéine SU(VAR)3-7 dans la formation de la structure hétérochromatinienne reste indéterminé.

Afin d'étudier la protéine SU(VAR)3-7, des mutants ont été générés par recombinaison homologue. L'étude de ces mutants a montré que cette protéine était essentielle mais que cependant, la contribution maternelle était suffisante pour assurer la viabilité (Seum et al., 2002). La cause de la létalité reste inconnue et il apparaît que les mâles sont plus sensibles que les femelles à la perte de la protéine SU(VAR)3-7 (Seum et al., 2002).

Des expériences complémentaires ont démontré que les phénotypes observés sur les chromosomes polytènes dans le cas d'une délétion de la protéine SU(VAR)3-7 étaient similaires à ceux observés dans la cas d'une mutation de la protéine HP1. Ces phénotypes affectent préférentiellement les chromosomes X des mâles ce qui suggère un rôle de la machinerie de compensation de dose (Spierer et al., 2005).

La Topoïsomérase II

L'activité enzymatique de la Topoïsomérase II permet de relâcher les super tours de l'ADN. Cette protéine est principalement connue pour son activité de décaténtation mais aussi pour son intervention dans les mécanismes de condensation des chromosomes mitotiques (Adachi et al., 1991). Outre cette activité, elle a été démontrée comment pouvant s'accumuler au niveau des séquences SAT III de la drosophile (Käs and Laemmli, 1992; Marshall et al., 1997). Cette région particulière représenterait un lieu de dépôt et/ou de stockage pour la Topoïsomérase II.

De plus, la Topoïsomérase II est aussi massivement localisée au niveau du rDNA chez la drosophile. Encore une fois, cette localisation particulière reste sans fonctionnalité précise, même si il a été décrit que la Topoïsomérase II intervenait dans la régulation de la transcription des gènes ribosomiques (Cavalli et al., 1996). Cette fonction est cependant à rapprocher de l'effet de la Topoïsomérase II sur la recombinaison puisqu'il a été montré que chez la levure qu'elle intervenait dans la répression des mécanismes de recombinaison du rDNA (Kim and Wang, 1989), et ce, entre autre, grâce à un partenaire : la protéine Pat1p (Wang et al., 1999).

4. *Mécanisme d'assemblage de la structure hétérochromatinienne*

a. Modèle théorique de l'assemblage

De nombreux travaux ont permis d'élucider la nature des mécanismes mis en œuvre au cours de l'assemblage d'un domaine hétérochromatinien et un modèle a ainsi pu être proposé (Grewal and Moazed, 2003). La nature des modifications post-traductionnelles portées par la lysine 9 de l'histone H3 est un marqueur relativement fiable du type de chromatine. Sous forme acétylée, ce résidu corrèle généralement avec une chromatine dite « ouverte » dans laquelle la transcription est facilitée. Par contre, la méthylation de la lysine 9 de l'histone H3 est une marque épigénétique de la chromatine dite « fermée » ou hétérochromatine. La transition d'une structure euchromatinienne vers une structure de type hétérochromatinienne de mise en silence impliquerait donc une transition entre une forme acétylée et une forme méthylée de la lysine 9 de l'histone H3 (figure 25).

Ce modèle est soutenu par différentes évidences. Par exemple, les protéines de la famille SU(VAR)3-9, qui sont les histone-méthyltransférases spécifiques de la Lysine 9 de l'Histone H3, sont impliquées dans la répression transcriptionnelle (Stewart et al., 2005). Pour autant, une certaine modération s'impose à la vue de travaux récents démontrant l'implication de la méthylation de la lysine 9 de l'histone H3 dans les processus d'élongation des transcrits ARN (Vakoc et al., 2005). Au niveau des régions péricentromériques du chromosome X de la drosophile, les différentes étapes du processus d'assemblage de l'hétérochromatine sont désormais mieux définies (Blattes et al., 2006), même si des incertitudes subsistent quant aux mécanismes responsables des étapes d'initiation.

Il est généralement admis que la mise en place d'un domaine hétérochromatinien fait appel à plusieurs étapes. La première étape de ce processus correspond à la transition de la lysine 9 de l'histone H3 d'un état acétylé vers un état méthylé. Cette transition nécessite l'action conjuguée de deux activités catalytiques. La première activité consiste en l'enlèvement du motif acétyl de la lysine 9. Cette modification est assurée par la protéine HDAC1 (Czermin et al., 2001). Consécutivement à cette désacétylation, le résidu lysine est méthylé par la protéine SU(VAR)3-9 (Schotta et al., 2003; Vaute et al., 2002).

Sous sa forme méthylée, la lysine 9 de l'histone H3 permet le recrutement de la protéine HP1 via son CD, ce qui constitue la deuxième étape du mécanisme (Bannister et al.,

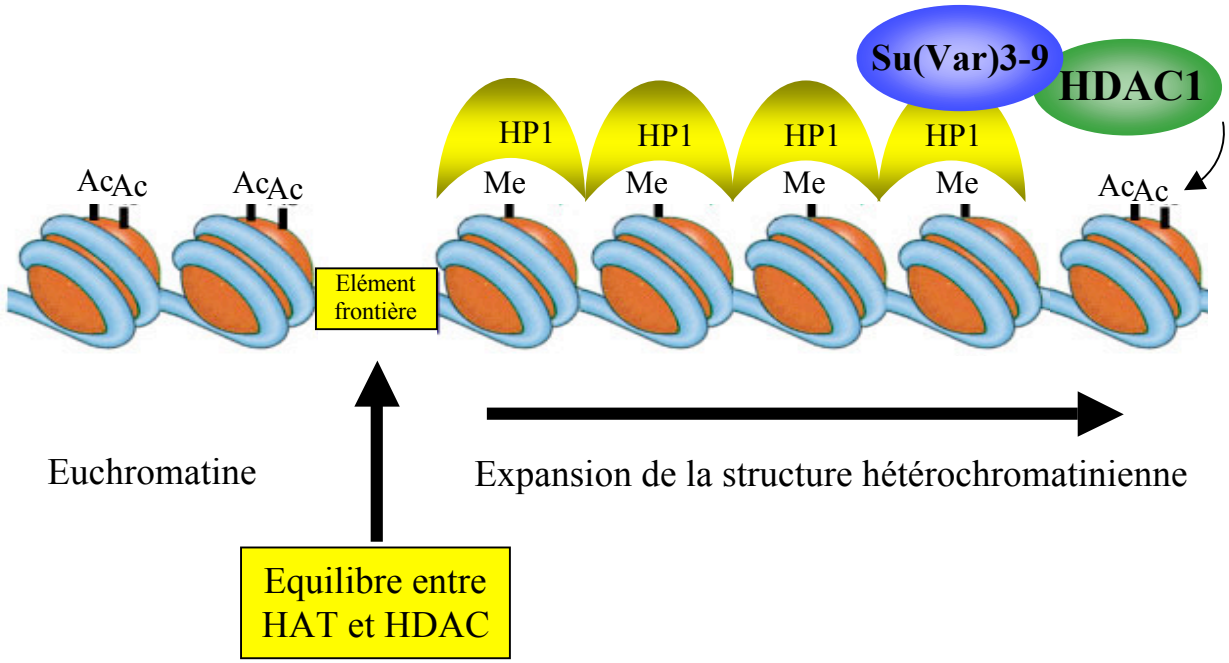


Figure 25 : modèle théorique de l'assemblage de la structure hétérochromatinienne (d'après Bannister et al., 2001).

La mise en place de la structure hétérochromatinienne est effectuée par modification des histones de proches en proches. A partir d'un point de départ, une lysine 9 de l'histone H3 sera méthylée par l'action coordonnée d'une histone désacétylase, qui va dans un premier temps enlever le groupement acétyl présent sur ce résidu, puis une histone méthyltransférase va venir méthyler la lysine 9 de l'histone H3. La lysine 9 méthylée va alors servir de site de fixation pour la protéine HP1, qui à son tour va recruter les protéines de modifications des histones et ainsi de proche en proche, la structure hétérochromatinienne de mise en silence va se mettre en place.

Cette structure ne sera arrêtée que grâce à des séquences de type élément frontière où un équilibre entre acétylation et désacétylation est mis en place. Il est important de noter que cette expansion peut aussi être stoppée à cause de l'épuisement d'une des protéines impliquées dans la mise en place d'une telle structure, en l'absence de tout élément frontière. La terminaison se produirait alors de manière plus graduelle.

2001; Lachner et al., 2001). La protéine HP1 va à son tour recruter la protéine SU(VAR)3-9 et participe ainsi, de proche en proche, à la propagation de la structure hétérochromatinienne par la modification des nucléosomes voisins. Le recrutement de la protéine SUV4-20, une histone-méthyltransférase qui triméthyle spécifiquement la lysine 20 de l'histone H4, va assurer de manière concomitante la mise en place de l'empreinte épigénétique de ces régions (Kourmouli et al., 2004; Schotta et al., 2004). La protéine HP2 a été décrite comme compagnon de la protéine HP1 (Shaffer et al., 2002) mais son rôle dans ce mécanisme est encore mal connu. De plus, même si la mutation du gène codant la protéine HP2 est associée à une suppression de la PEV du gène *white* dans l'allèle *white-mottled*, et malgré ses deux motifs AT Hook, la protéine HP2 ne semble pas posséder de propriété de liaison directe à l'ADN (Stephens et al., 2005). Enfin, un traitement avec l'oligopyrrole P9 n'affecte pas la localisation de cette protéine du chromocentre des chromosomes polytènes (S. C. Elgin, communication personnelle).

b. Initiation de l'assemblage de la structure hétérochromatinienne

Une question majeure quant au processus d'assemblage de l'hétérochromatine péricentrique subsiste : quel est le déterminant spécifique de l'assemblage de l'hétérochromatine sur les séquences satellites riches en dA•dT qui constituent les régions péricentriques du chromosome X chez la drosophile ? A l'heure actuelle, trois hypothèses sont émises sur les mécanismes responsables de l'initiation de la formation de la structure hétérochromatinienne. L'initiation pourrait faire intervenir les petits ARN de la famille des siARN, l'intégration de variants d'histones ou des protéines reconnaissant à la fois une séquence d'ADN spécifique mais aussi les protéines de la machinerie d'assemblage.

Les petits ARN interférents

L'initiation de l'assemblage de la structure hétérochromatinienne pourrait impliquer la machinerie d'interférence à l'ARN (RNAi, (Montgomery, 2004)). Sous ce terme est désigné un mécanisme par lequel des molécules d'ARN double brin conduisent à la dégradation spécifique des molécules d'ARN homologues dans une cellule (figure 26). Décrit initialement chez le vers *Cænorhabditis elegans*, ce mécanisme est conservé chez tous les eucaryotes (Zamore, 2001). Son implication dans les processus d'assemblage de la structure

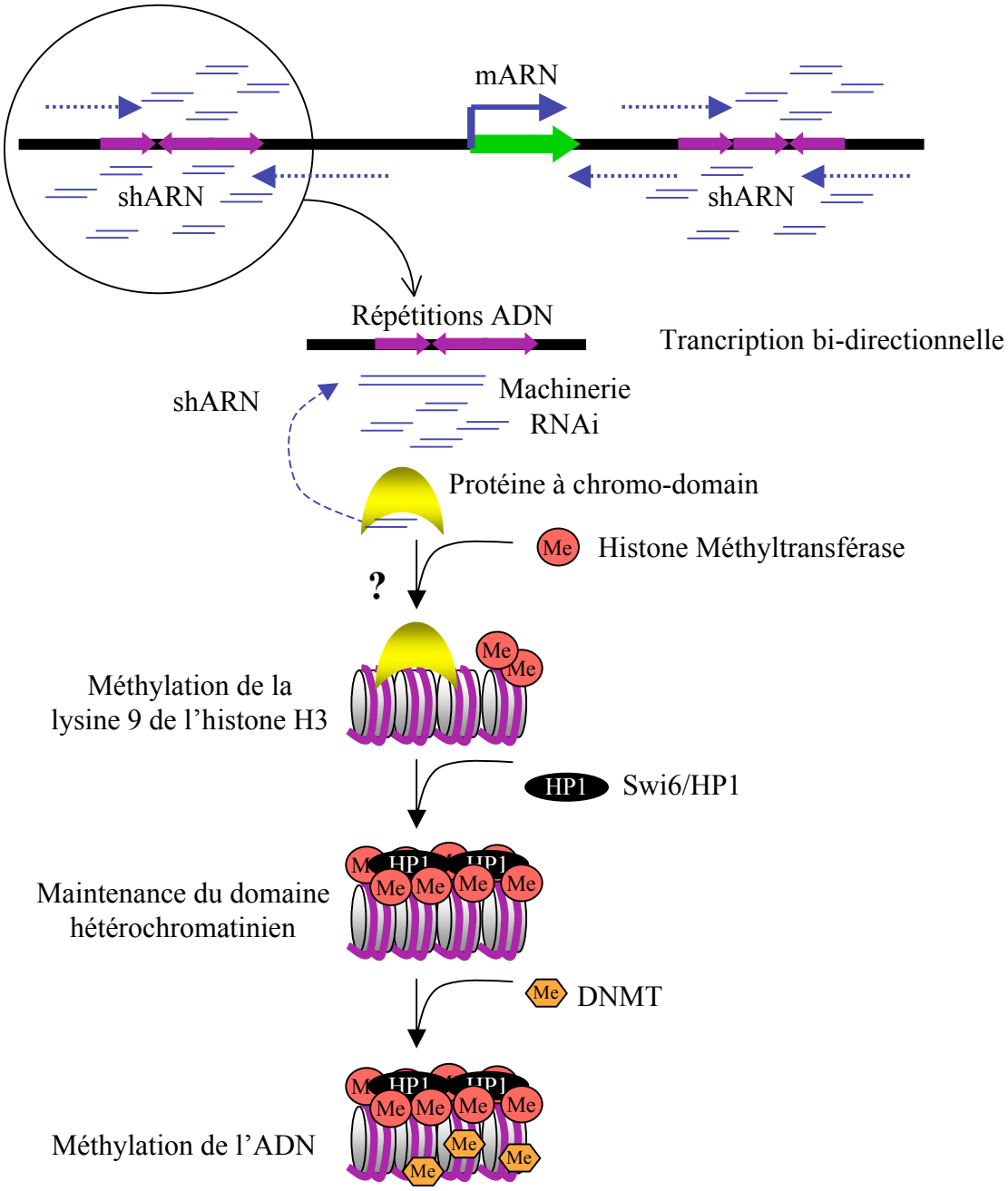


Figure 26 : modèle d'initiation de l'assemblage de la structure hétérochromatinienne par la machinerie d'interférence à ARN chez la levure à fission (d'après Jenuwein et al., 2002).

Des répétitions d'ADN de la levure *Schizosaccharomyces pombe* pourraient générer des transcrits ARN double brins (shRNAs, short heterochromatic RNA) suite à une transcription bidirectionnelle. Ces shRNAs seraient alors pris en charge par la machinerie de RNAi pour être maturés et participeraient au ciblage des facteurs responsables de la formation d'un domaine hétérochromatinien. La question des déterminants responsables de la transcription bidirectionnelle des séquences répétées reste cependant ouverte.

hétérochromatinienne a d'abord été démontrée chez la levure à fission *Schizosaccharomyces pombe* (Volpe et al., 2002). En effet, chez *S. pombe* la délétion des gènes codant les facteurs intervenants dans la voie RNAi s'accompagne d'une expression aberrante de transcrits issus des répétitions de l'hétérochromatine péricentromérique et d'une levée de la répression de transgènes insérés dans l'hétérochromatine (Volpe et al., 2002).

Parmi les facteurs intervenant dans la machinerie RNAi, on trouve les protéines DICER, une RNAseIII hélicase responsable du clivage des ARN double brins en petits ARN de 21 à 24 nucléotides (Bernstein et al., 2001), ARGONAUTE, un des composants du complexe RISC (RNA Interference Silencing Complex, (Hammond et al., 2001)), et Rdrp, une RNA polymérase RNA dépendante qui a pour rôle d'amplifier ce phénomène (Sijen et al., 2001).

La dérépression des transgènes situés dans l'hétérochromatine en absence de la machinerie RNAi est concomitante à une diminution de la méthylation de la lysine 9 de l'histone H3. Par ailleurs, l'existence de transcrits correspondants aux répétitions péricentriques a été mise en évidence chez *S. pombe* (Reinhart and Bartel, 2002), mais aussi chez la drosophile (Aravin et al., 2003). Ces petits ARN, appelés shRNA (small heterochromatic RNA), (Jenuwein, 2002; Vaistij et al., 2002)), ou rasiRNA (repeat associated small interfering RNA), (Djikeng et al., 2001)), seraient capables de guider la méthylation de la lysine 9 de l'histone H3, au niveau des séquences satellites qui sont elles mêmes à l'origine de ces shRNA (Volpe et al., 2002). La mise en place d'une structure hétérochromatinienne via ces petits ARN ferait intervenir les complexes RITS (RNA-induced Initiation of Transcriptional Silencing) et RISC (RNA Induced Silencing Complex), impliqués dans le mécanisme de mise en silence par les siRNA (small interfering RNA) (Hammond et al., 2000; Pal-Bhadra et al., 2004; Verdel et al., 2004).

L'étude des mutants pour la machinerie RNAi chez la drosophile (Pal-Bhadra et al., 2004), et chez les vertébrés (Fukagawa et al., 2004), a confirmé le lien existant entre la machinerie RNAi et la mise en place de la structure hétérochromatinienne. Ce mécanisme d'initiation aurait donc été conservé au cours de l'évolution.

Pour autant, même si l'existence de petits ARN correspondant aux répétitions SAT III a été mise en évidence chez la drosophile, l'origine exacte de ces transcrits n'a pas été

formellement établie (Aravin et al., 2003). De telles séquences sont dispersées dans le génome, à la fois dans des localisations hétérochromatiques et euchromatiques, aussi bien sur le chromosome X que sur les autosomes (Losada and Villasante, 1996). De plus, les petits ARN impliqués dans l'assemblage d'hétérochromatine agissant en *cis*, il n'est pas évident que les transcrits issus des séquences SAT III proviennent d'une transcription bidirectionnelle des copies de ces séquences présentes au niveau des régions péricentriques, ni que ces ARN soient responsables de l'initiation du mécanisme d'assemblage d'un compartiment d'hétérochromatine. Il semble en effet peu probable que la présence de promoteurs cryptiques au niveau de ces régions suffise à rendre compte de l'existence de ces transcrits.

Par ailleurs, il n'existe aucune démonstration expérimentale de l'implication de la machinerie d'interférence à l'ARN dans l'établissement d'une structure de type hétérochromatinienne dans les régions péricentriques du chromosome X. Si les mutations des gènes codant pour des facteurs intervenant dans la voie RNAi, tels que les protéines Piwi et Argonaute, suppriment la PEV de transgènes *mini-white* en tandem ou d'insertions sur le chromosome 4 hétérochromatinien (Dorer and Henikoff, 1994; Dorer and Henikoff, 1997; Pal-Bhadra et al., 2004), elles sont sans effet sur la lignée *white-mottled*. Par contre, dans les mutants de la protéine Dicer, le phénotype au niveau des ommatidies passe d'une répartition « poivre et sel », les cellules exprimant *white* sont dispersées, à une répartition dite « sectorisée » (communication personnelle, U. Bhadra). Cette dernière observation semble plutôt refléter un impact développemental de la mutation de *Dicer* qu'un effet direct sur la PEV. Cette mutation aurait alors un effet sur le stade développemental au cours duquel l'allèle *white-mottled* serait activé ou réprimé plus que sur la décision elle-même.

Intégration de variants d'histones

Une autre piste a été ouverte chez les levures *S. cerevisiae* et *S. pombe* où de nombreux travaux soulignent le rôle des homologues de la protéine CENP-B (Masumoto et al., 2004) et du variant de l'histone H3 centromérique CENP-A dans ce mécanisme d'intégration de variants dans la chromatine (Tanaka et al., 2005). Chez l'homme, l'interaction entre CENP-B et les boîtes CENP-B des séquences satellites alphoïdes et le recrutement de CENP-A semblent être essentiels pour l'assemblage du kinétochore. Par homologie, leurs orthologues chez la levure assureraient ces fonctions en permettant l'assemblage d'une structure hétérochromatinienne à partir des sites de nucléation constitués

par les séquences riches en dA•dT. Ces protéines sont en effet capables de s'associer spécifiquement aux séquences satellites centromérique de la levure *S. cerevisiae* (Meluh et al., 1998) et *S. pombe* (Meluh et al., 1998; Nakagawa et al., 2002). Elles constituent également des composants essentiels des centromères et sont nécessaires à la désacétylation de la lysine 9 de l'histone H3 et au recrutement de la protéine Swi6, l'homologue de la protéine HP1 chez la levure *S. pombe* (Nakagawa et al., 2002). Jusqu'à il y a peu, l'absence d'homologues de CENP-B chez la drosophile supposait l'existence, chez cet organisme, d'un autre type de facteur assurant cette fonction de nucléation d'un domaine hétérochromatinien au niveau des séquences satellites riches en dA•dT.

L'existence d'un variant d'histone H2A, appelé H2Av, a été démontrée et son rôle dans la mise en silence des gènes homéotiques par les complexes Polycomb mais aussi dans l'établissement de l'hétérochromatine péricentrique ont été caractérisés chez la drosophile (Leach et al., 2000; Swaminathan et al., 2005).

Ce variant d'histone non centromérique serait recruté sur les régions à hétérochromatiniser antérieurement à la méthylation de la lysine 9 de l'histone H3. Ce recrutement provoquerait l'acétylation de la lysine 12 de l'histone H4 qui serait une nouvelle marque épigénétique de l'hétérochromatine (Swaminathan et al., 2005). La mutation de l'histone H2Av entraîne une diminution de la méthylation de la lysine 9 de l'histone H3, et par conséquent une diminution du recrutement de la protéine HP1. De même, une mutation du gène codant ce variant d'histone entraîne une suppression de la PEV de l'allèle *white-mottled* (Swaminathan et al., 2005).

Le variant H2Av semble intervenir dans le recrutement des répresseurs Polycomb par l'histone méthyltransférase E(Z), en aval de la méthylation de la lysine 27 de l'histone H3 (Cao et al., 2002; Czermin et al., 2002). En plus des régions péricentriques et des sites Polycomb, H2Av pourrait être recrutée dans l'euchromatine pour la mise en silence de transgènes portant des copies répétées en tandem d'un gène *mini-white* par la formation de sites d'hétérochromatine ectopiques (Dorer and Henikoff, 1994; Dorer and Henikoff, 1997; Swaminathan et al., 2005). Malgré ces nouvelles indications, l'identité du déterminant de localisation de ce variant H2Av sur les séquences riches en dA•dT demeure une question ouverte.

Rôle de protéines de liaison à l'ADN

Les travaux de l'équipe ont permis de proposer la protéine D1, l'unique représentant de la famille des protéines Multi AT Hook (MATH) chez la drosophile comme candidat pour cette fonction d'initiation.

La protéine D1 a tout d'abord été identifiée du fait de son enrichissement au niveau des séquences SAT III (Rodriguez Alfageme et al., 1980), et dans une moindre mesure sur les séquences SAT I. Son implication dans la formation d'une structure hautement compacte de type hétérochromatinienne sur ces séquences satellites a immédiatement été suggérée (Levinger and Varshavsky, 1982; Levinger, 1985) (figure 27), mais n'a pas été démontrée avant la caractérisation d'un mutant de la protéine D1 en 2002 (Aulner et al., 2002).

Ces caractéristiques et le phénotype d'haplo-suppressor/triplo-enhancer de PEV indiquent que la protéine D1 joue un rôle central dans l'assemblage de la structure hétérochromatinienne (Aulner et al., 2002). Cette hypothèse a constitué la base de mon projet de recherche.

c. Les complexes de déposition des histones

Au cours de la réplication de l'ADN, la moitié des nucléosomes est formée par transfert des nucléosomes déjà existants sur l'un des deux brins d'ADN néosynthétisés. La moitié manquante est formée, quant à elle, par des histones nouvellement produites et hyperacétylées.

Lors de la déposition des nucléosomes nouvellement formés, le tétramère d'histones H3-H4 hyperacétylées vient contacter l'ADN au milieu des 147 paires de bases qui vont l'entourer. Les deux dimères d'histones H2A-H2B sont rajoutés par la suite. Le tétramère d'histone H3-H4 forme un complexe stable avec l'ADN grâce à des facteurs d'assemblage de la chromatine. Ces facteurs sont des protéines chaperonnes qui permettent l'acheminement des histones de leur lieu de synthèse jusqu'à la fourche de réplication et facilitent leur déposition sur l'ADN.

Des mécanismes similaires doivent être mis en place lors de la réparation de lésions de l'ADN. Dans ce cas, les nucléosomes doivent également être enlevés puis repositionnés sur l'ADN. Il existe également des complexes capables de déposer des nucléosomes sur l'ADN déjà empaqueté sous forme de chromatine. En effet, on constate que dans certains cas, il est

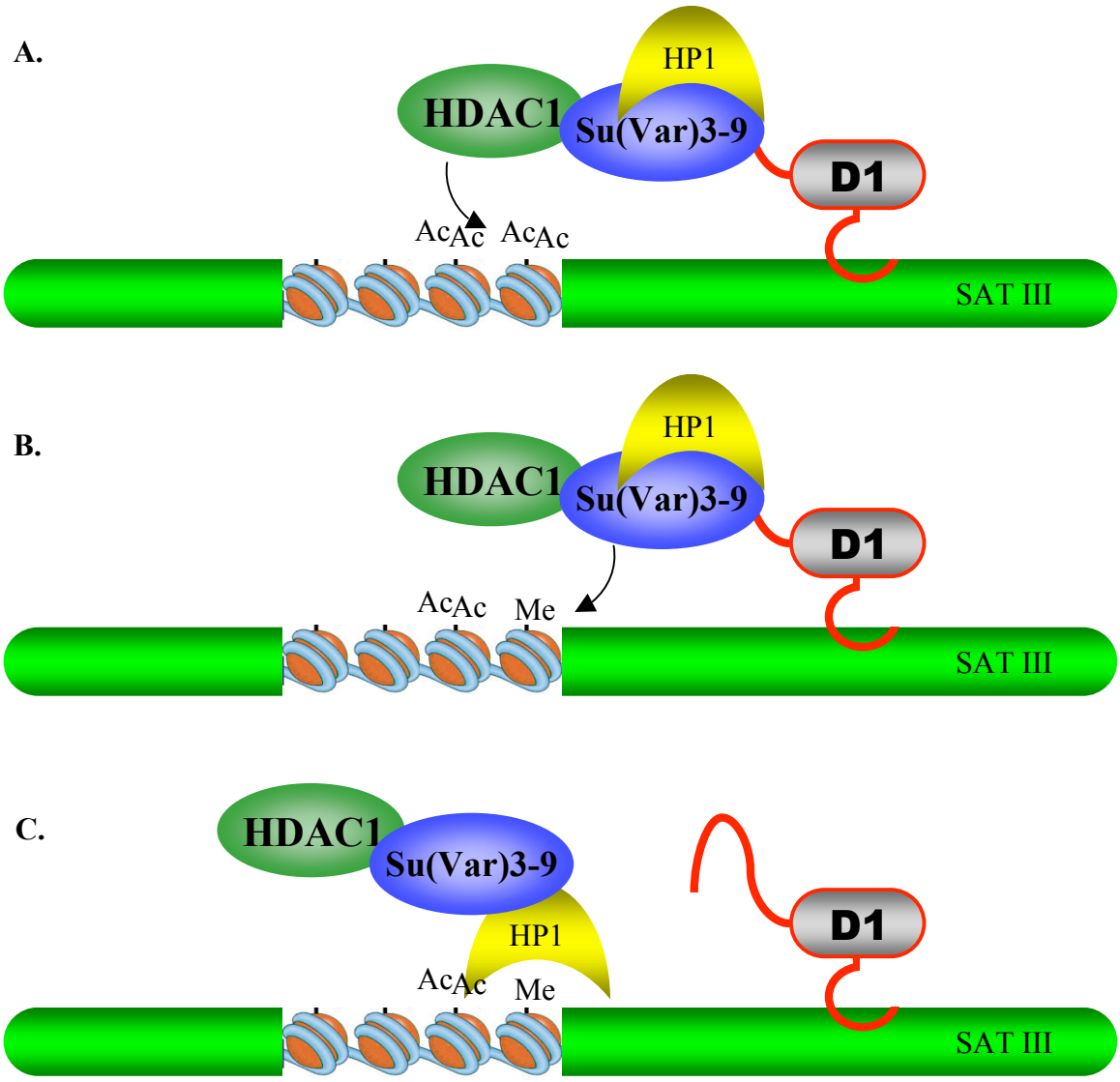


Figure 27 : modèle d'initiation de l'assemblage de la structure hétérochromatinienne par des protéines reconnaissant l'ADN.

Une protéine capable de reconnaître spécifiquement une séquence d'ADN à hétérochromatiniser, comme par exemple les séquences satellites riches en dA•dT, serait capable de recruter les protéines capables de mettre en place une telle structure. Ici, la protéine D1, qui interagit spécifiquement avec les séquences SAT III du chromosome X de la drosophile, recruterait les protéines de modification des histones. Suite à ce recrutement, les histones seraient désacétylées grâce à l'action de la protéine HDAC1 (**A**). L'histone méthyltransférase Su(Var)3-9 méthylerait donc la lysine 9 de l'histone H3 (**B**) créant un site de liaison pour la protéine HP1 (**C**), capable de mettre en place la structure hétérochromatinienne de proche en proche.

nécessaire de remplacer les histones comme, par exemple, dans les zones transcriptionnellement actives, l'histone H3 est remplacée par son variant H3.3 (Ahmad and Henikoff, 2002) et, au niveau des séquences centromériques, le variant d'histone CENP-A prend la place de l'histone H3 (Henikoff et al., 2000). La déposition des histones peut se faire soit d'une manière réplication dépendante, soit de façon réplication indépendante.

Le complexe CAF-1 (Chromatin Assembly Factor 1) est un exemple de complexe de déposition des histones dépendant de la réplication mais aussi à la réparation de l'ADN. CAF-1 a été purifié à partir de cellules humaines en exploitant sa capacité à promouvoir l'assemblage de la chromatine *in vitro* (Smith and Stillman, 1989). Celui-ci joue non seulement un rôle dans le maintien de l'information épigénétique mais peut également générer un nouveau patron de marques épigénétiques dans les cellules filles.

La protéine HIRA (HIstone RegulAtion) est impliquée dans la déposition des histones de manière indépendante de la réplication. Lors de la surexpression de cette protéine, les gènes codant les histones subissent une répression transcriptionnelle (Nelson et al., 2002). De plus, il a été montré qu'elle pouvait interagir avec les histones (Lorain et al., 2001). Une immunodéplétion de la protéine HIRA a démontré que seule la déposition des histones indépendante de la réplication était abolie, démontrant ainsi le rôle de cette protéine dans ce processus (Ray-Gallet et al., 2002). La protéine HIRA est ainsi impliquée dans la déposition du variant d'histone H3.3 (Tagami et al., 2004). En purifiant les complexes protéiques, il a également été montré que des dimères H3-H4, et non pas des tétramères, étaient retrouvés, suggérant un héritage des nucléosomes par dimères (Tagami et al., 2004).

d. Les complexes de remodelage de la chromatine ATP-dépendants

Du fait de la compaction extrême que peut adopter la chromatine et de la nécessité d'accéder au message génétique porté par la molécule d'ADN, les processus tels que la réplication et la transcription requièrent une ouverture de la chromatine et un désassemblage des nucléosomes pour permettre le passage des polymérases.

Les complexes de remodelage de la chromatine ATP-dépendants peuvent modifier la position des nucléosomes sur l'ADN, permettant une meilleure accessibilité. Les complexes

protéiques impliqués dans le remodelage de la chromatine comportent plusieurs sous-unités dont au moins une porte un domaine ATPase similaire à celui de la famille des protéines Swi2/Snf2. La majorité des complexes caractérisés comme étant des complexes de remodelage de la chromatine *stricto sensus* sont impliqués dans des mécanismes tels que la réparation de l'ADN, la réplication ou la transcription.

De manière générale, les complexes de remodelage de la chromatine vont mobiliser les nucléosomes en utilisant l'énergie issue de l'hydrolyse de l'ATP pour déstabiliser les contacts entre l'ADN et l'octamère d'histone (Peterson, 2002a; Peterson, 2002b). Les complexes de remodelage de la chromatine peuvent être recrutés de deux manières différentes au niveau d'un promoteur : soit par les facteurs de transcription séquence-spécifiques (Struhl and Greenwald, 1999), soit par des modifications spécifiques des histones, comme par exemple l'acétylation de la lysine 14 de l'histone H3 (Kasten et al., 2004). Il existe trois grandes familles de complexes de remodelage de la chromatine : la famille SWI/SNF, la famille ISWI et la famille CHD ou Mi2 (figure 28).

La famille SWI/SNF

Le premier membre de la famille SWI/SNF (pour Switch/sucrose non fermenting) a été caractérisé au cours de cribles génétiques chez la levure (Laurent et al., 1991). Dans ce complexe, la sous unité protéique Swi2/Snf2 porte le domaine ATPase (Vignali et al., 2000). Parmi les dix autres sous-unités du complexe, des protéines possédant des domaines d'interaction avec l'ADN ont été caractérisées, de même que des protéines Arp (Actin Related Proteins) pouvant interagir avec la matrice nucléaire (Cairns et al., 1998).

Le complexe SWI/SNF de levure peut jouer un rôle d'activateur ou de répresseur transcriptionnel, selon sa localisation (Burns and Peterson, 1997). Un autre complexe de remodelage de la chromatine, RSC (Remodel the Structure of Chromatin), a été caractérisé chez la levure sur les bases d'une homologie avec le complexe SWI/SNF (Cairns et al., 1996). Dans ce complexe, la protéine Sth1 (Snf Two Homologous) porte l'activité ATPasique (Du et al., 1998).

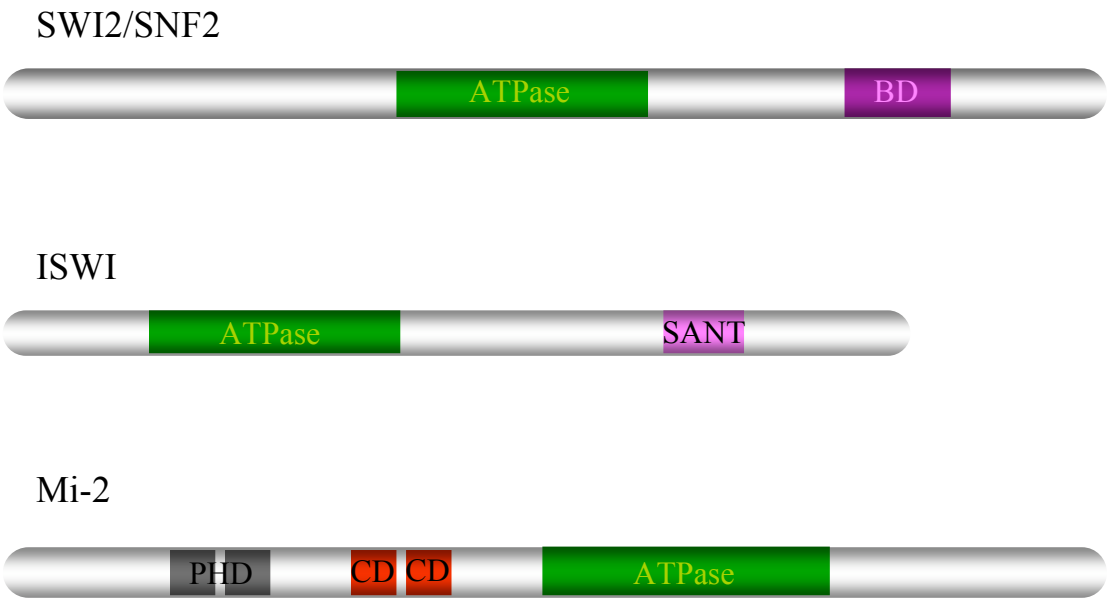


Figure 28 : les facteurs de remodelage de la chromatine ATP-dépendants.

Il existe trois familles de facteurs de remodelage de la chromatine ATP-dépendants : la famille SWI2/SNF2, la famille ISWI et la famille MI-2. Les domaines conservés d'hydrolyse de l'ATP et d'interaction avec la chromatine sont représentés.

Domaine d'hydrolyse de l'ATP

« Bromo-domain » (BD)

Domaine SANT (domaine d'interaction à l'ADN qui reconnaît des séquences spécifiques)

« Chromo-domain » (CD)

PHD finger (domaine homologue au Zinc Finger, permettant la liaison à l'ADN)

Des complexes très similaires à SWI/SNF de levure ont été découverts chez la drosophile, le complexe Brahma (Tamkun et al., 1992), et chez les mammifères, le complexe hSWI/SNF (Kwon et al., 1994). Des protéines homologues à Swi2/Snf2 ou aux Arp y ont de même été caractérisées (Phelan et al., 1999). Chez l'homme, deux complexes sont connus comme étant les orthologues de SWI/SNF : les complexes BRG1 et hBRM.

La famille ISWI

Le domaine ATPase de ISWI (Imitation SWI) est l'un des plus conservés (Tsukiyama et al., 1995). Les études les plus complètes de cette famille ont été menées chez la drosophile. Trois complexes dont la sous-unité catalytique est ISWI ont ainsi été caractérisés : les complexes NURF (NUcleosome Remodeling Factor) (Tsukiyama and Wu, 1995), CHRAC (CHRomatin Accessibility Complex) (Eberharter et al., 2001; Varga-Weisz et al., 1997) et ACF (ATP-dependant Chromatin assembly Factor) (Ito et al., 1997; Ito et al., 1999). Des études menées *in vitro* sur des gènes rapporteurs ont montré que le complexe NURF est capable de faciliter la transcription (Mizuguchi et al., 2001), alors que *in vivo*, le complexe ISWI peut activer ou réprimer la transcription (Badenhorst et al., 2002; Deuring et al., 2000).

La famille CHD ou Mi2

Plusieurs complexes contenant la protéine Mi2, possédant un domaine ATPasique de la famille SNF2, ont été identifiés (Tong et al., 1998; Xue et al., 1998). Le complexe Mi2, ou NuRD (Nucleosome Remodeling and Deacetylases), possède un domaine de remodelage de la chromatine et une activité désacétylase portée par les protéines HDAC1 et HDAC2 (Zhang et al., 1999). L'activité de la protéine Mi2 pourrait donc servir à faciliter l'accès aux histones afin que les protéines HDAC puissent les désacétyler. Ce complexe comporte aussi les protéines MBD2 et MBD3 qui sont capables de lier l'ADN méthylé (Zhang et al., 1999). Des études ont montré que ce complexe était impliqué dans la répression transcriptionnelle.

e. La méthylation de l'ADN

La méthylation de l'ADN est importante tant pour le développement que pour la vie adulte. Elle joue un rôle central dans le contrôle de l'expression génique, l'empreinte

génomique parentale ou l'inactivation du chromosome X. Le patron de méthylation de l'ADN est mis en place de manière très précoce, dès le stade embryonnaire. Cette modification épigénétique peut être retrouvée sur de très grandes zones de chromatine transcriptionnellement silencieuse. Sur l'ADN, les quatre bases peuvent être méthylées mais c'est la cytosine qui l'est le plus fréquemment. La modification est mise en place sur le carbone 5 à la place d'un atome d'hydrogène.Les cytosines méthylées peuvent représenter jusqu'à 7 % des cytosines totales et elles sont toujours retrouvées au niveau d'un enchaînement de dinucléotides CpG.

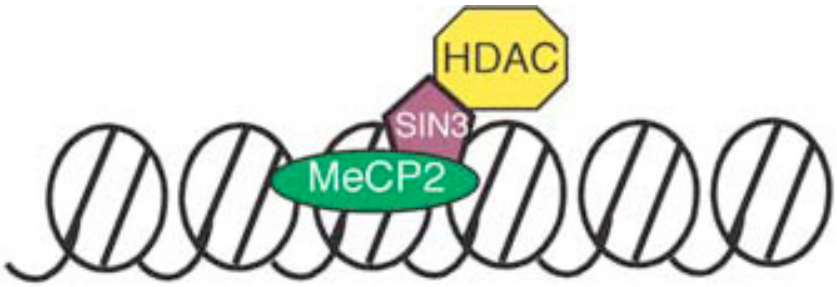
Deux classes d'enzymes capables de méthyler l'ADN ont été caractérisées. La première classe, correspondant à la protéine DNMT1 chez les mammifères (Yen et al., 1992), est impliquée dans la maintenance du patron de méthylation durant la réplication. La deuxième classe, regroupant les protéines de mammifères DNMT3a et DNMT3b (Okano et al., 1999), participe à l'établissement de nouveaux patrons de méthylation au cours du développement embryonnaire précoce.

L'une des caractéristiques importante de la méthylation de l'ADN correspond à la localisation des groupements méthyles dans le grand sillon de l'ADN. Cette modification a non seulement des répercussions importantes sur la structure même de l'ADN, mais aussi sur les interactions entre les protéines et l'ADN. Cette variation dans la structure de l'ADN peut entraîner des modifications de l'affinité de certaines protéines comme les histones ou les facteurs de transcription. Certaines protéines vont lier spécifiquement l'ADN méthylé grâce à un motif appelé MBD (Methyl CpG Binding Domaine, (Nan et al., 1993)).

Les protéines à domaine MBD ont pour modèle la protéine MeCP2 (figure 29). Elle a été caractérisée comme pouvant interagir de manière séquence spécifique avec de l'ADN méthylé (Meehan et al., 1989; von Kries et al., 1991). Ce domaine MBD est souvent associé à un autre domaine protéique permettant d'interagir ou de modifier les histones, comme par exemple l'association d'un domaine MBD avec un domaine SET dans la protéine SETDB1 (Blackburn et al., 2003; Schultz et al., 2002).

Dans certains cas, il a été démontré que la méthylation de l'ADN pouvait agir de manière coopérative avec des protéines impliquées dans la méthylation des histones, et plus particulièrement dans la méthylation la lysine 9 de l'histone H3 (Bowen et al., 2004; Fuks et al., 2003). Ces deux mécanismes agiraient donc de concert afin de réprimer la transcription. Cette coopération est d'ailleurs à la base du phénomène d'inactivation du chromosome X chez la plupart des mammifères (Heard et al., 2001).

A.



B.

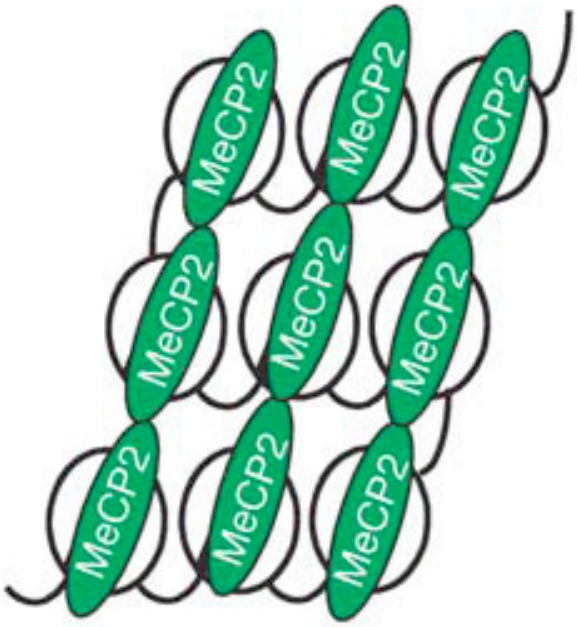


Figure 29 : modèles de condensation de la chromatine initiée par la méthylation de l'ADN.

A. Modèle enzymatique. La protéine MeCP2, capable de se lier spécifiquement aux cytosines méthylées, recrute des enzymes modifiant les histones, comme par exemple l'histone désacétylase HDAC1.

B. Modèle structural. La protéine MeCP2 est capable de s'homopolymériser pour former une structure secondaire de la chromatine (d'après Bowen *et al*., 2004).

De nombreuses protéines à domaine MBD ont été mises en évidence à la fois chez l'homme, la souris ou la drosophile, même si chez cette dernière, la présence d'ADN méthylé n'a pas été clairement été établie et reste très controversée.

f. Modifications, variants et code des histones

Parce que différentes modifications sont corrélées à différents états transcriptionnels, l'hypothèse d'un « code des histones » a été émise (Strahl and Allis, 2000). Selon cette hypothèse, les différentes combinaisons de modifications des histones attribuent une identité à chaque nucléosome et sont lues dans la cellule comme des codes barres moléculaires par des machineries capables de générer des états spécifiques de la chromatine. Une telle théorie d'un code des histones suppose l'existence à la fois de protéines capables d'opérer ces modifications post-traductionnelles des histones mais aussi de machineries capables de déchiffrer ce langage (de la Cruz et al., 2005). Plusieurs types de modifications ont été mises en évidence (figure 30 et (Margueron et al., 2005)). Elles peuvent agir dans le but de favoriser l'activité transcriptionnelle ou de la réprimer.

Acétylation des histones

Les histones peuvent être acétylées et cette modification est souvent corrélée à une activation transcriptionnelle (Allfrey et al., 1964; Hebbes et al., 1988). Il a été démontré que l'acétylation globale des histones permettait de favoriser le désassemblage des nucléosomes grâce à la diminution des interactions électrostatiques entre l'ADN chargé négativement et les histones chargés positivement (Travers and Thompson, 2004). Cette déstabilisation est un processus nécessaire à l'activation transcriptionnelle (Lee et al., 1993).

L'acétylation des histones s'effectue principalement sur des résidus lysines (Brownell and Allis, 1996). Cette modification est mise en place par les protéines appelées HAT (Histone Acetyl Transferases), enlevée par les protéines HDAC et reconnue par les protéines à « Bromo-domain » (Dhalluin et al., 1999; Loyola and Almouzni, 2004).

Les HAT sont classées en trois familles : les HAT de la famille des GNAT (Gcn5 related N-Acetyltransferases, (Georgakopoulos and Thireos, 1992)), celle des MYST (MOZ

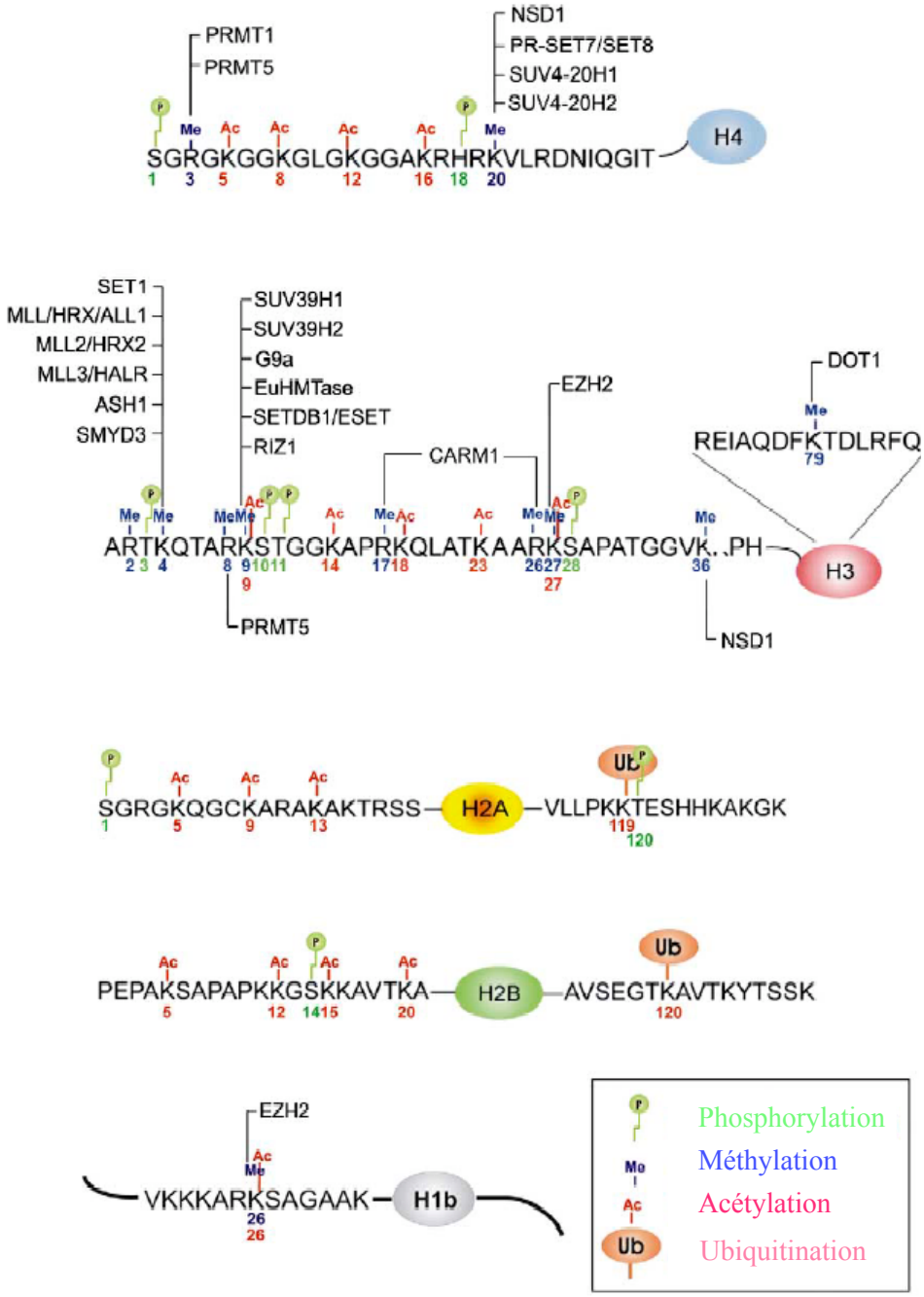


Figure 30 : Principales modifications post-traductionnelles des histones canoniques et les histones méthyltransférases (d'après Margueron et al. 2005).

Les lysines des histones peuvent être la cible d'acétylations, de méthylations (mono-, di- et tri-) et d'ubiquitination. Les résidus arginines peuvent également être méthylés (mono-, diméthylation symétrique ou asymétrique) et enfin certaines sérines et thréonines sont phosphorylées. Parmi de ces modifications, certaines peuvent être mises en place par plusieurs enzymes comme dans le cas de la méthylation des lysines 4 et 9 de l'histone H3.

Tbf2/Sas3 Sas2 Tip60, (Utley and Cote, 2003)) et enfin celle des protéines p300/CBP (Bannister and Kouzarides, 1996). La plupart des HAT fonctionnent au sein de complexes multiprotéiques (Yang et al., 2004), et les facteurs associés à ces complexes semblent contrôler leur spécificité de substrat et augmenter leur activité catalytique (Grant et al., 1997). Ces complexes protéiques sont par ailleurs conservés de la levure à l'homme comme l'illustre la conservation des protéines constituant le complexe SAGA chez la levure et le complexe PCAF chez l'homme (Sterner and Berger, 2000).

Méthylation des histones

Les histones peuvent être méthylées sur des résidus lysines qui peuvent être mono-, di- ou triméthylés (Lachner et al., 2001; Rea et al., 2000). Cette modification peut aussi intervenir sur des résidus arginines, de manière symétrique ou asymétrique (Lee et al., 2005; Zhang and Reinberg, 2001) (figure 31).

Contrairement à l'acétylation, la méthylation n'est pas directement corrélée à un type fonctionnel de la chromatine. L'influence de la méthylation sur le potentiel transcriptionnel d'un domaine chromatinien dépend avant tout du résidu modifié. La méthylation des lysines 9 et 27 de l'histone H3 est une marque de la chromatine silencieuse (Fischle et al., 2003), alors que la triméthylation de la lysine 4 de l'histone H3 marque les régions transcriptionnellement actives (Santos-Rosa et al., 2002). La mise en place de la méthylation des lysines est due à l'activité des protéines à domaine SET.

La même dualité des rôles de la méthylation est retrouvée au niveau des résidus arginines. Dans sa forme asymétrique, la méthylation des arginines par les protéines PRMT1 et CARM1 est impliquée dans l'activation transcriptionnelle, alors que dans sa forme symétrique, la diméthylation des arginines suite à l'action de la protéine PRMT5 est associée à la répression transcriptionnelle (Wysocka et al., 2006).

L'accumulation des études démontrant le caractère dynamique des méthylations des lysines des histones a récemment fait tomber le dogme selon lequel cette marque était irréversible (Dillon and Festenstein, 2002). Ainsi, qu'il s'agisse de la lysine 4 de l'histone H3 (Ma et al., 2001), de la méthylation de la lysine 9 de l'histone H3 (Ghosh and Harter, 2003; Janicki et al., 2004; Nicolas et al., 2003; Saccani and Natoli, 2002) ou de résidus arginines (Metivier et al., 2003), les preuves de l'existence d'une déméthylation active des histones en réponse à des signaux cellulaires s'accumulent.

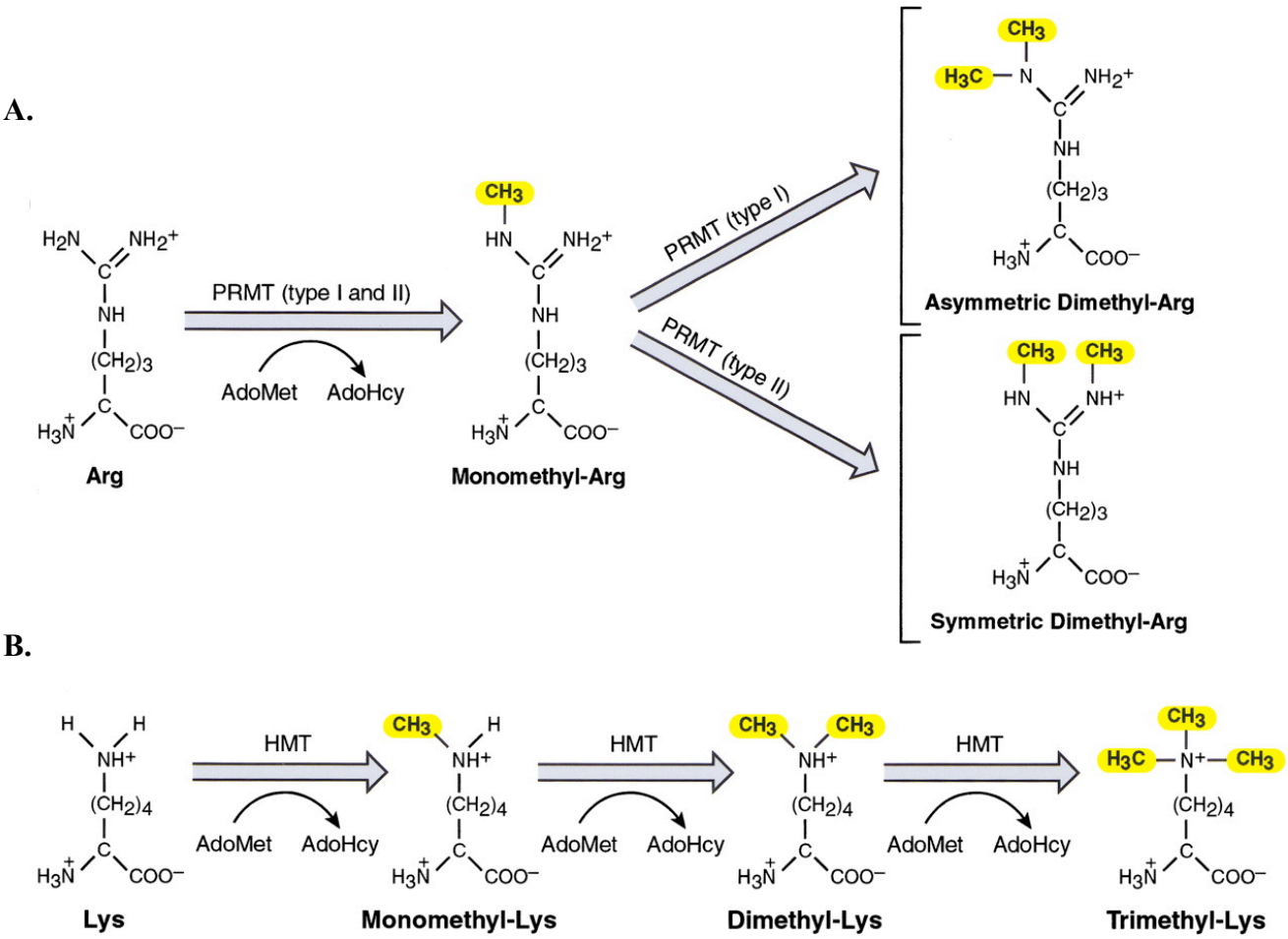


Figure 31 : voies de méthylations des résidus lysines et arginines (d'après Zhang et al., 2001).

A. Voie de méthylation des résidus arginines. En présence de S-Adénosyl Méthionine (AdoMet), les arginines (Arg) sont monométhylées par des méthyltransférases spécifiques des arginines (PRMT). La diméthylation asymétrique et symétrique des résidus monométhylés sont catalysées par les PRMT de type I et II respectivement.

B. Voie de méthylation des résidus lysines. Les lysines (Lys) peuvent être mono-, di- ou triméthylées par des histones méthyltransférases (HMT).

Trois mécanismes sont proposés pour expliquer la déméthylation des histones. Le premier de ces mécanismes repose sur le désassemblage du nucléosome qui porte la modification à enlever et la substitution de l'histone marquée par une nouvelle histone (Ahmad and Henikoff, 2002). Ce mécanisme implique la mise en place de variants d'histones, tel que l'histone H3.3, spécifiquement déposé de manière réplication indépendante au niveau des gènes actifs en transcription.

Un deuxième mécanisme propose l'existence d'une activité enzymatique capable d'assurer la réaction de déméthylation (figure 32). Cette nouvelle famille de protéines a été pour la première fois mise en évidence chez l'homme (Cuthbert et al., 2004; Denman, 2005; Wang et al., 2004). Les protéines arginines déméthylases ont été les premières caractérisées. Leur existence ayant d'abord été supposée (Bannister et al., 2002; Kim et al., 1964; Paik and Kim, 1973), des lysines déméthylases ont été découvertes (Shi et al., 2004; Tsukada et al., 2006). Cette nouvelle famille vient donc s'ajouter aux nombreuses protéines impliquées dans la mise en place et la lecture du code des histones (Forneris et al., 2005a; Forneris et al., 2005b).

Enfin, le troisième mécanisme implique les protéines qui reconnaissent ces modifications. La mise en place de modifications plus facilement réversibles, comme l'acétylation ou la phosphorylation, permettent de moduler les interactions entre les résidus méthylés et les protéines capables de s'y fixer. Par exemple, la phosphorylation de la thréonine 3 de l'histone H3 bloque la reconnaissance de la lysine 4 de l'histone H3 méthylée par la protéine CHD1 (Flanagan et al., 2005). De même, la phosphorylation de la sérine 27 de l'histone H3 empêche la reconnaissance par la protéine HP1 de la lysine 26 de l'histone H3 sous sa forme méthylée (Daujat et al., 2005).

Phosphorylation des histones

La phosphorylation est une modification dynamique dont la régulation repose sur un équilibre entre les activités opposées de kinases, qui catalysent l'ajout d'un groupement phosphate, et de phosphatases, qui enlèvent ces mêmes groupements. Cette modification peut impliquer des résidus thréonines ou sérines.

La phosphorylation de l'histone H1 est très souvent liée à la condensation et à la ségrégation des chromosomes durant la mitose et la méiose (Gurley et al., 1978). Si la fonction principale de la phosphorylation de la sérine 10 de l'histone H3 est la condensation

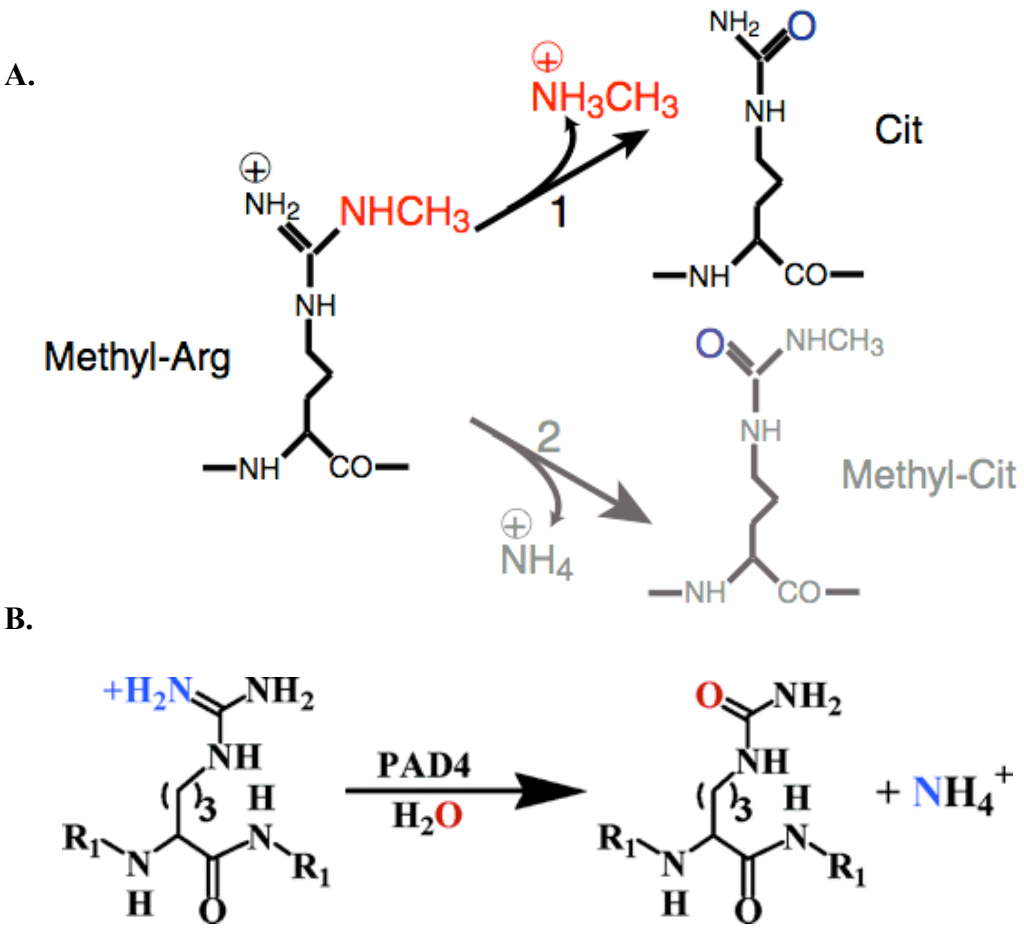


Figure 32 : voie de démethylation par l'enzyme PAD4, spécifique des résidus arginines (d'après Wang et al., 2004).

A. L'arginine déméthylase PAD4 catalyse la déméthylation des arginines méthylées. En fonction du groupement qui sera ôté par l'enzyme, les produits de la réaction sont soit une citrulline et une méthylamine (voie 1, partie haute) soit une méthylcitrulline et un ion ammonium (voie 2, partie basse).

B. L'enzyme PAD4 est également capable de réaliser la conversion d'arginines non méthylées en citrulline, ce qui empêchera une méthylation ultérieure.

des chromosomes, son rôle n'est pas limité à la mitose. Des travaux récents ont montré une implication de cette modification dans l'activation de la transcription par l'inhibition de l'interaction entre la protéine HP1 et la lysine 9 de l'histone H3 sous sa forme méthylée (Eissenberg and Elgin, 2005).

La phosphorylation des histones semble également pouvoir jouer un rôle général dans le contrôle de l'apoptose (Lee et al., 1999a). Plus particulièrement, la phosphorylation de la sérine 32 de l'histone H2B semble être un bon marqueur des cellules apoptotiques (Ajiro, 2000). La réparation des cassures double brin chez la levure semble aussi impliquer la phosphorylation des histones et plus précisément la phosphorylation de la sérine 129 de l'histone H2A (Downs et al., 2000).

Les autres modifications des histones

Acétylation, méthylation et phosphorylation sont les trois modifications les plus étudiées, mais il existe d'autres modifications qui jouent un rôle important dans les processus cellulaires.

La mono-ADP-ribosylation peut être mise en place sur des résidus arginines et glutamates principalement sur les histones H1 et H2B (Burzio et al., 1979; Golderer and Grobner, 1991), mais cette modification semble aussi être présente sur les autres histones. De manière générale, l'ADP-ribolysation est impliquée dans la réparation des dommages de l'ADN (Kreimeyer et al., 1984).

La mono-ubiquitination consiste en l'ajout covalent d'un peptide ubiquitine sur les résidus lysines (Jason et al., 2002). Cette modification repose sur l'activité des enzymes Bre1 et Rad6 chez la levure (Hwang et al., 2003; Wood et al., 2003). Par nature, elle est antagoniste de l'acétylation et la méthylation de ces mêmes résidus. Alors que la poly-ubiquitination conduit à la dégradation de la protéine modifiée, l'ajout d'un seul peptide ubiquitine semble corrélé à l'activation transcriptionnelle (Thorne et al., 1987). Par ailleurs, il a été démontré que l'ubiquitination de l'histone H2B pouvait favoriser la méthylation de la lysine 4 de l'histone H3 chez la levure (Dover et al., 2002; Sun and Allis, 2002), et ainsi contribuer à l'activation transcriptionnelle.

La sumoylation, modification consistant à l'ajout d'un peptide SUMO normalement impliquée dans la dégradation des protéines, a été décrite sur les histones H4. Cette modification interviendrait dans la mise en silence (Shiio and Eisenman, 2003).

La biotinylation est elle aussi associée aux résidus lysines et l'ajout d'un peptide Biotine entrerait en compétition avec la modification de ces mêmes résidus (Camporeale et al., 2004). Elle serait impliquée dans la prolifération cellulaire (Narang et al., 2004; Stanley et al., 2001).

Les variants d'histones

Jusqu'à maintenant, les principaux variants d'histones sont de la famille des histones H2A, H3 et H1, sauf chez la drosophile où cette dernière ne possède pas de variants (figure 33 et (Sarma and Reinberg, 2005)).

Les variants de l'histone H1 sont généralement exprimés en fonction du cycle cellulaire (Khochbin and Wolffe, 1994). Parmi les différents variants de l'histone H1, certains peuvent être exprimés de manière tissu-spécifique comme le variant H1t, spécifique des testicules (Doenecke et al., 1997), ou le variant CSH1 (Cleavage Stage H1) présent dans les oocytes (Mandl et al., 1997).

Des variants de l'histone H1, les histones H1° et H5, sont exprimés de manière dépendante de la différenciation. Par exemple, l'histone H5 n'est retrouvé que dans les érythrocytes des oiseaux (Khochbin and Wolffe, 1994; Zlatanova and Doenecke, 1994).

Même si la question du rôle de ces variants de l'histone H1 reste ouverte, il semblerait que ces protéines puissent réguler la structure de la chromatine et donc moduler l'accessibilité de l'ADN aux protéines qui le reconnaissent (Woodcock et al., 2006).

Il existe quatre isoformes de l'histone H3 : les histones H3.1, qui correspond à la forme canonique de l'histone H3, H3.2, H3.3 et enfin CENP-A, un variant centromérique qui diverge fortement des autres isoformes (Henikoff et al., 2000; Smith, 2002).

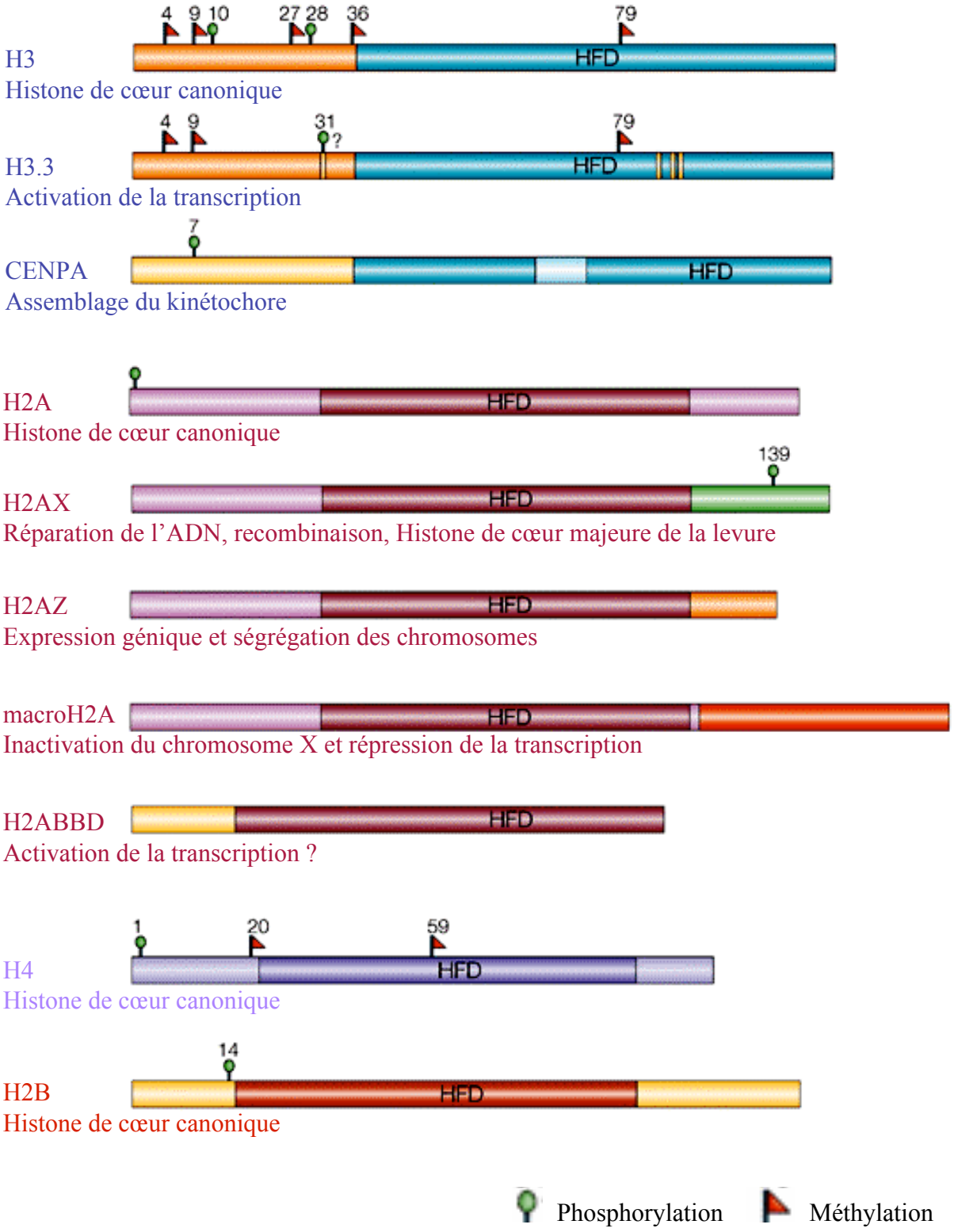


Figure 33 : histones canoniques et variants d'histones (d'après Sarma et al., 2005).

Les histones sont composés d'un domaine central qui adopte un repliement de type « histone fold » ainsi que de domaines amino- et carboxy-terminaux non structurés qui sont la cible de nombreuses modifications post-traductionnelles telles que la méthylation des lysines ou la phosphorylation des sérines.

CENP-A prend la place de l'histone H3 dans les nucléosomes centromériques et joue un rôle dans la ségrégation des chromatides sœurs en mitose (Howman et al., 2000).

Le variant H3.3 est associé à l'activation de la transcription (Daury et al., 2006; Schwartz and Ahmad, 2005). Il ne diffère que d'un acide aminé au niveau de la queue N-terminale et que de trois acides aminés dans le domaine « histone fold » avec l'histone H3.1 et de quatre acides aminés avec le variant H3.2 (figure 34). Le variant H3.3 est la seule histone exprimée dans les cellules en dehors de la phase S avec le variant CENP-A. Elle va donc remplacer progressivement l'histone H3 dans des cellules qui ont quitté le cycle et qui se différencient chez la drosophile (Lennox and Cohen, 1988) et chez le rat (Pina and Suau, 1987). Le variant H3.3 est incorporé de manière dépendante de la transcription, et sa présence pourrait amener un nouveau patron de modification épigénétique et ainsi remplacer l'ancien porté par l'histone H3.1.

Les principaux variants de l'histone H2A sont au nombre de quatre : H2AX, MacroH2A, H2ABBD et H2AZ.

Chez les eucaryotes supérieurs, le variant H2AX joue un rôle dans la recombinaison homologue et dans la réparation des cassures double-brins (Rogakou et al., 1998). Son niveau d'expression corrèle avec le niveau de recombinaison dans une espèce donnée.

Le variant MacroH2A participe chez les vertébrés à l'inactivation d'une des deux copies du chromosome X de la femelle qui formera le corpuscule de Barr. La déposition du variant MacroH2A sur le chromosome X à inactiver est dépendante de la présence de l'ARN Xist (Costanzi and Pehrson, 1998).

Par contre, le variant H2ABBD, pour Barr Body Deficient, est absent du chromosome X inactivé et pourrait jouer un rôle dans l'activation de la transcription (Chadwick and Willard, 2001b).

Le variant H2AZ serait lui aussi impliqué dans l'activation transcriptionnelle et pourrait de plus jouer le rôle de frontière en empêchant l'invasion des régions euchromatiniennes par une structure hétérochromatinienne (Meneghini et al., 2003; Raisner et al., 2005). De manière contradictoire, le variant H2AZ pourrait intervenir dans la répression transcriptionnelle (Dhillon and Kamakaka, 2000). Ce variant est localisé au niveau des régions péricentriques et interviendrait dans la ségrégation des chromosomes (Rangasamy et al., 2003; Rangasamy et al., 2004).

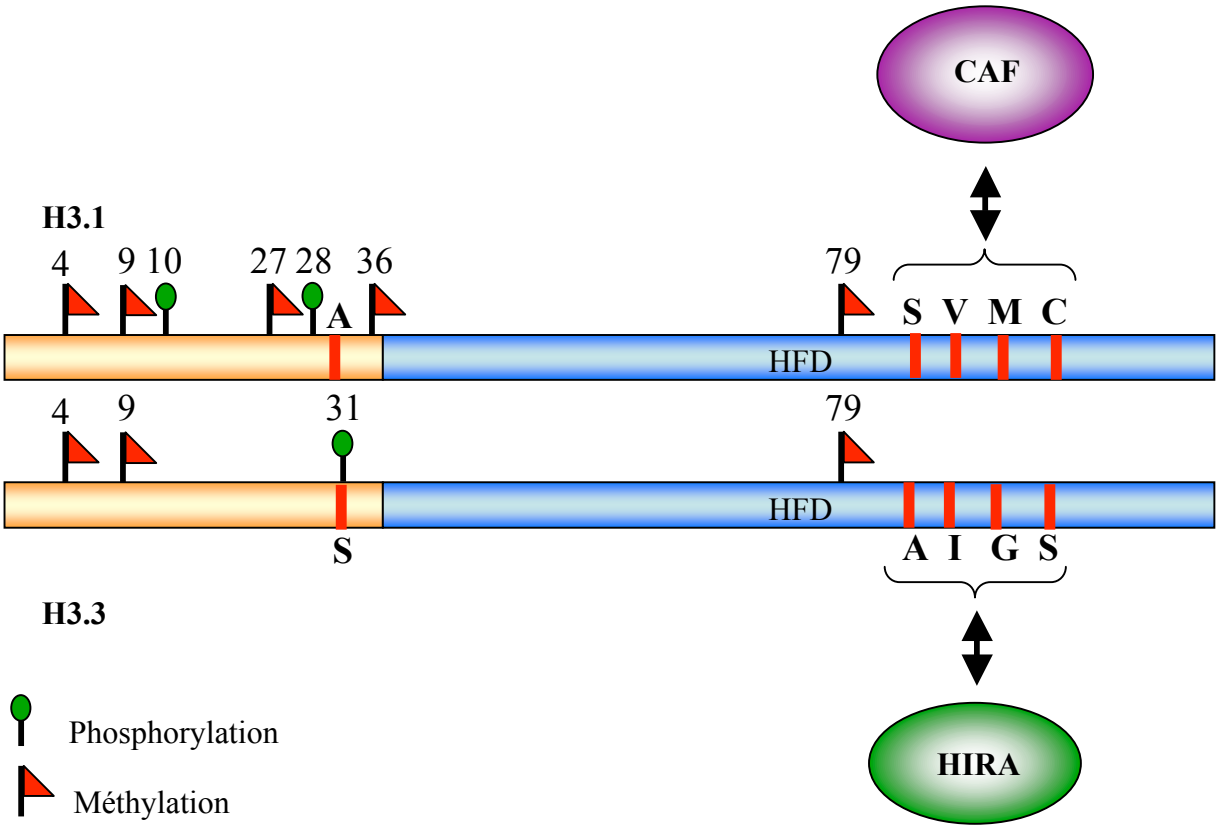


Figure 34 : représentation schématique de l'histone canonique H3.1 et du variant H3.3 (d'après Sarma et al., 2005).

L'histone canonique H3.1 (en haut) dont la déposition par le complexe CAF est couplée à la réplication ne diffère de son variant H3.3 que par cinq acides aminés (indiqués en rouge). Quatre de ces résidus se situent dans le domaine globulaire constitué par le domaine « histone-fold » (HFD en bleu) et trois d'entre eux, les résidus arginine 87, isoleucine 89 et glycine 90 de l'histone H3.3, sont impliqués dans la spécificité de reconnaissance par la chaperonne HIRA et la déposition de l'histone H3.3 indépendante de la réplication. Le dernier acide aminé divergent entre les histones H3.1 et H3.3 se situe à l'extrémité amino-terminale (en orange) de l'histone H3. Dans l'histone H3.3, la sérine 31 qui remplace l'alanine de l'histone H3.1 pourrait être la cible d'une phosphorylation (symbole ovale vert). Les principales lysines cibles de méthylation sont indiquées par un drapeau rouge.

Chez la drosophile, un autre variant a été décrit, le variant H2Av, qui regroupe en fait des caractéristiques structurales de plusieurs autres variants. Il possède la queue C-terminale du variant H2AX et le domaine « histone-fold » d'H2AZ (figure 35). Ce variant serait impliqué dans l'apoptose induite par les dommages à l'ADN (Madigan et al., 2002) et dans les étapes précoces de l'assemblage de structures hétérochromatiniennes (Swaminathan et al., 2005).

5. *Hétérochromatine et régulation de l'expression génique*

L'hétérochromatine est définie historiquement comme étant la partie du génome qui ne se décondense pas en interphase. Elle est riche en séquences répétées, pauvre en gènes et dite inactive en transcription par opposition à l'euchromatine qui est riche en gènes activement transcrits. Une définition plus récente prend en compte les modifications épigénétiques, telle que la méthylation de la lysine 9 de l'histone H3 pour l'hétérochromatine, ou la présence de la protéine HP1 généralement assimilé à une répression transcriptionnelle (Huisinga et al., 2006).

Cependant, cette dichotomie est à modérer. En effet, un certain nombre de gènes ont été localisées dans l'hétérochromatine et semblent nécessiter cet environnement pour être exprimés efficacement. Chez la drosophile, c'est le cas par exemple des gènes *rolled* (Eberl et al., 1993) et *light* (Devlin et al., 1990a; Devlin et al., 1990b), qui requièrent la protéine HP1 pour leur expression (Eissenberg and Hilliker, 2000; Lu et al., 2000). Des études récentes ont démontré l'importance de sa présence pour l'expression de gènes euchromatiniens essentiels comme le gène *rpl15*, codant pour une protéine ribosomique (Schulze et al., 2005), ou comme le gène *cdc2*, une protéine kinase impliquée dans la régulation du cycle cellulaire (Cryderman et al., 2005). Le rôle de la protéine HP1 n'est donc pas limité à l'inactivation transcriptionnelle dans l'hétérochromatine : elle peut également avoir une fonction dans l'activation de la transcription à la fois dans l'hétérochromatine (Lu et al., 2000) et dans l'euchromatine (Kellum, 2003a; Piacentini et al., 2003). Mais ces conclusions doivent être modérées puisqu'il est nécessaire de vérifier quel est l'isoforme de la protéine HP1 recrutée. Ainsi, ce n'est pas l'isoforme HP1α, décrit comme hétérochromatinienne, mais l'isoforme HP1γ qui peut être impliqué dans la transcription de gènes euchromatiniens (Vakoc et al., 2005).

HISTONE	S. p	S. c	T	D. m	M. m	H. s	NOTES
H3.3 (a)	hht3	H3	hv2	His3.3	h3f3	H3.3 H3F3	Activation
CENP A	cnp1 sim2	Cse4	TCEN PA	Cid	Cenpa	CENPA	Assemblage du kinétochore
H3.4						H3.4	Gamétogenèse (b)
H2AX	H2A	H2A	H2A X	H2Av (c)	H2afx	H2A.X H2AF/X	Réparation de l'ADN Recombinaison
H2AZ	pht1	Htz1	hv1		H2afz H2afv	H2A.Z H2AF/Z	Activation/Répression Ségrégation
macroH2A					H2afy	MacroH2A H2AF/Y	Répression Inactivation du X
H2ABBD					H2abbd	H2ABbd H2AF/B	Activation
H2B(1)					TH2B	hTSH2B	Compaction (d)

Figure 35 : tableau de répartition des principaux variants d'histones chez les eucaryotes et leur fonction associée.

S. p : *Schizosaccharomyces pombe*, S. c : *Saccharomyces cerevisiae*, T. : *Tetrahymena*, D. m : *Drosophila melanogaster*, M. m : *Mus musculus*, H. s : *Homo sapiens*.

(a) Chez les plantes, le variant H3.2 joue un rôle dans l'activation de la transcription.
(b) Variant spécifique dans les testicules des mammifères, H3.3A chez la souris.
(c) H2Av chez la drosophile partage les caractéristiques des variants H2AX et H2AZ et est impliqué dans la formation de l'hétérochromatine péricentromérique.
(d) Variant spécifique dans les testicules des mammifères.

Il est aussi important de bien séparer les domaines d'hétérochromatine, riches en séquences répétées, des régions présentant des modifications épigénétiques de type hétérochromatiniennes, comme la méthylation de la lysine 9 de l'histone H3. Ces régions ne peuvent être assimilées à de l'hétérochromatine. Cependant, ce mode de régulation est fréquent au niveau de régions dont l'inactivation transcriptionnelle doit être relativement pérenne. De nombreuses observations mettant en jeu des mécanismes épigénétiques de répression transcriptionnelle ont été faites.

a. Régulation par les protéines Polycomb chez la drosophile

Les protéines du groupe Polycomb (Pc-G) interviennent dans la répression de l'expression des gènes homéotiques, nécessaires à la détermination du destin cellulaire de différentes parties du corps. Cette régulation met en balance les protéines Pc-G un autre groupe de protéines activatrices : le groupe Trithorax (Trx-G).

La fonction de répression par une des protéines du Pc-G, la protéine Polycomb (PC) passe par la présence dans sa partie amino-terminale d'un domaine de 37 acides aminés : le « chromo-domain » (CD) (Muller, 1995). Tout comme la protéine HP1, qui possède aussi un CD, la protéine PC est capable d'induire un état réprimé (Zink and Paro, 1995), maintenu pendant de nombreux cycles cellulaires. Leur surexpression est capable d'induire une augmentation du phénomène de variégation d'effet de position (Pirrotta, 1997).

L'activité de répression transcriptionnelle des protéines Pc-G passe par la formation d'une structure chromatinienne inerte en transcription, aux propriétés proches de l'hétérochromatine (Bienz and Muller, 1995). La mise en place de marques épigénétiques répressives, telle que la méthylation de la lysine 27 de l'histone H3, fait intervenir un complexe appelé Polycomb Repressive Complex 2 (PRC2) (Ng et al., 2002). Ce complexe comporte entre autre les protéines ESC, E(Z) une HMT (Czermin et al., 2002; Kuzmichev et al., 2002) et RPD3, une HDAC (Tie et al., 2001).

Des séquences cibles des protéines (PRE, Polycomb Responsive Element) sont présentes au niveau des promoteurs des gènes homéotiques (Chan et al., 1994). Les protéines Pc-G sont capables de former des complexes multiprotéiques (Muller, 1995), à l'intérieur desquels certaines protéines sont connues pour interagir avec l'ADN (Blastyak et al., 2006). Ainsi, les deux complexes PRC1 et PRC2 agissent de concert afin de permettre la répression transcriptionnelle. Le complexe PRC2 est responsable de la mise en place des marques

épigénétiques répressives et du recrutement du complexe PRC1, qui reconnaît ces marques et empêche l'initiation de la transcription des gènes cibles par séquestration des facteurs nécessaires à cette étape (Dellino et al., 2004).

b. Variégation d'effet de position de l'allèle *white-mottled* w^{m4h} chez la drosophile

Bien que ce phénomène soit connu depuis 1930, ses bases moléculaires restent aujourd'hui encore peu comprises, les niveaux d'expression de l'allèle *white-mottled* variant fortement au sein d'une même population. Dans la variégation d'effet de position de l'allèle *white-mottled* w^{m4}, le gène *white*, normalement localisé dans la partie distale du chromosome X, est relocalisé à proximité du bloc d'hétérochromatine péricentromérique, principalement constitué de séquences SAT III (Tartof et al., 1984).

L'observation du phénotype de la PEV suppose l'inactivation stochastique du gène *white* par une structure hétérochromatinienne répressive (Henikoff, 1990; Henikoff, 1992). Dans cette lignée, le phénotype d'inactivation stochastique du gène *white* est visualisable au niveau des yeux. En effet, le produit du gène *white* est un transporteur transmembranaire des pigments rouges de l'œil (Zhang and Odenwald, 1995).

Ce phénotype particulier a servi de base à la caractérisation de nombreuses protéines impliquées dans l'assemblage de la structure chromatinienne. Des mutations de gènes ont été décrites comme affectant le phénotype de variégation. Deux classes ont ainsi été caractérisées : les mutations engendrant une perte de fonction du gène *white* (appelés enhancer de variégation ou *E(var)*), et des mutations présentant un gain de fonction du gène *white* (appelés suppresseurs de variégation ou *Su(var)*) (Reuter and Wolff, 1981). Parmi ces suppresseurs de variégation, certains ont été impliquées dans la mise en place de la structure hétérochromatinienne, comme les protéines SU(VAR)3-9, HP1 ou HDAC1 déjà décrites précédemment.

L'expansion de la structure hétérochromatinienne est normalement stoppée dans un contexte cellulaire sauvage. Deux hypothèses peuvent être formulées : soit cet arrêt impliquerait un mécanisme passif qui reposerait sur l'épuisement d'un des facteurs nécessaires à sa mise en place (effet de dose caractérisé pour les suppresseurs de variégation), soit l'hétérochromatine est confinée par la présence de séquences capables d'arrêter son expansion, comme des séquences isolantes ou des éléments frontières (Cuvier et al., 2002;

Cuvier et al., 1998; Zhao et al., 1995) où des facteurs activateurs ou répresseurs seraient en compétition (Lam et al., 2005; Suka et al., 2002).

c. Effet de position chez la levure

Chez la levure, un phénomène de mise en silence à proximité des régions télomériques a été décrit (Tartof, 1994). En effet, ces régions ont des propriétés tout à fait similaires à celle de l'hétérochromatine des eucaryotes supérieurs (Gabriel-Robez and Rumpler, 1992; Gotta and Cockell, 1997; Gotta et al., 1996).

Chez la levure *Saccharomyces cerevisae*, les régions de type hétérochromatiniennes sont présentes au niveau des télomères (Gottschling et al., 1990), des loci du typage sexuel et des gènes ribosomiques (Smith and Boeke, 1997) (figure 36). Chez la levure à fission *Schizosaccharomyces pombe*, les centromères sont plus étendus que chez la levure *S. cerevisae* et l'hétérochromatine y est présente (Ekwall et al., 1997), tout comme sur le locus de typage sexuel (Ekwall and Ruusala, 1994).

L'effet répresseur observé au niveau des télomères est la conséquence de la formation d'un complexe multiprotéique composé à la fois de facteurs spécifiques et d'histones (Gartenberg, 2000). La protéine Rap1 se lie aux répétitions télomériques et initie la polymérisation du complexe Sir3/Sir4 sur le domaine chromatinien adjacent, par interaction avec les parties amino-terminales des histones H3 et H4 (Hecht et al., 1995). Une surexpression d'un des éléments du complexe répresseur peut générer un effet, appelé TPE (Telomere Position Effect), similaire à celui de la PEV.

d. Locus HML/HMR chez la levure

Chez la levure *Saccharomyces cerevisae*, les copies des gènes du typage sexuel, appelés a et α, sont localisés dans deux loci silencieux respectivement dénommés HMR (HoMothallic Right) et HML (HoMothallic Left). L'un ou l'autre sera exprimé par transposition au site MAT. La répression transcriptionnelle est réalisée par la condensation très importante de la chromatine sur ces sites, générant une structure dont les propriétés sont proches de l'hétérochromatine (Hecht et al., 1995).

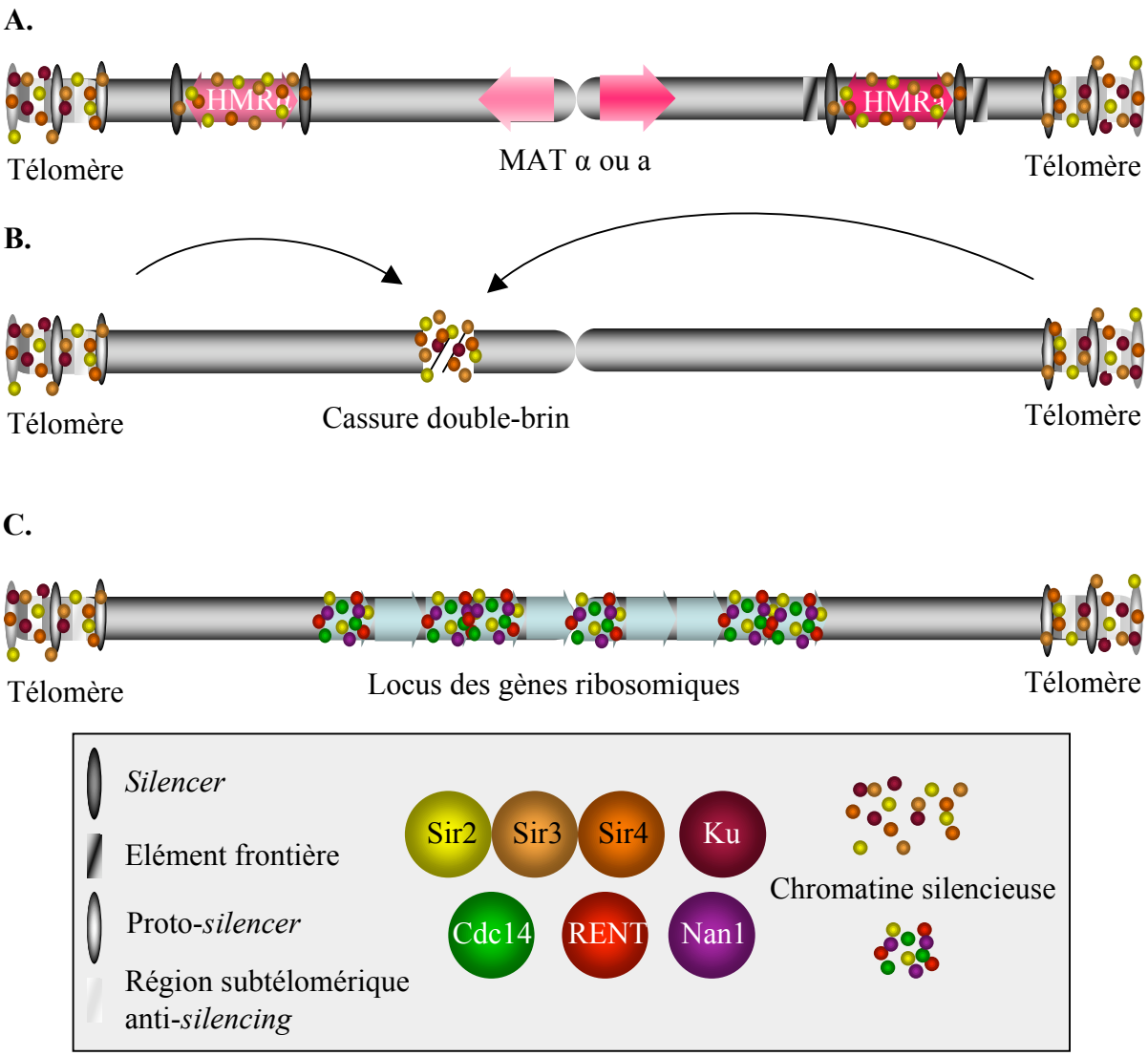


Figure 36 : représentation de la distribution des protéines Sir et de la chromatine silencieuse sur le génome de la levure.

A. Le chromosome III de la levure *Saccharomyces cerevisiae* comporte les loci de conjugaison silencieux HMLa et HMRα, et le locus de conjugaison actif, MAT, qui contient soit une copie du gène de conjugaison a, soit une copie du gène α. Le gène HMRα est encadré par des *silencers* et des éléments frontières, alors que le gène HMLa n'est encadré que par des *silencers*. La protéine Ku ne se lie qu'aux extrémités distales des télomères des chromosomes. Dans les parties sub-télomériques, les proto-silencers et les régions sub-télomériques de mise en silence sont localisés de manière aléatoire.

B. Au niveau d'une cassure double-brin, les protéines Ku et Sir sont déplacées des extrémités des chromosomes vers les points de cassure.

C. Le chromosome XII porte le locus des gènes ribosomiques. Une chromatine silencieuse est arbitrairement assemblée sur certains gènes ribosomiques. La distribution réelle de la chromatine silencieuse au niveau de ce locus reste inconnue.

La répression transcriptionnelle y est induite par la formation sur ces loci d'un complexe multiprotéique contenant principalement les protéines SIR (Lustig, 1998), qui interagissent avec les histones H3 et H4 hypoacétylées (Braunstein et al., 1993). De nombreuses évidences, biochimiques et génétiques, ont permis d'isoler d'autres membres de ce complexe silencieux dont le facteur de régulation Rap1 (Repressor/Activator Protein) (Buck and Shore, 1995), et le complexe ORC (Origin of Replication Complex, (Loo et al., 1995)).

e. Compensation de dose et inactivation du chromosome X

Le phénomène de compensation de dose a été initialement décrit par H. Muller (Muller, 1948; Muller and Kaplan, 1966). Des études moléculaires ont alors été mises en œuvre principalement dans trois organismes modèles : les mammifères, la mouche *Drosophila melanogaster* et le nématode *Cænorhabditis elegans*. Dans ces organismes, la régulation transcriptionnelle permettant un niveau équivalent entre les produits des gènes du chromosome X entre mâles et femelles s'effectue de manières différentes. Trois mécanismes ont été ainsi caractérisés : l'augmentation de la transcription du chromosome X des mâles chez la drosophile, l'inactivation d'un des deux chromosomes X de la femelle chez les mammifères et la réduction de moitié de la transcription des deux chromosomes X de la femelle chez le ver (figure 37).

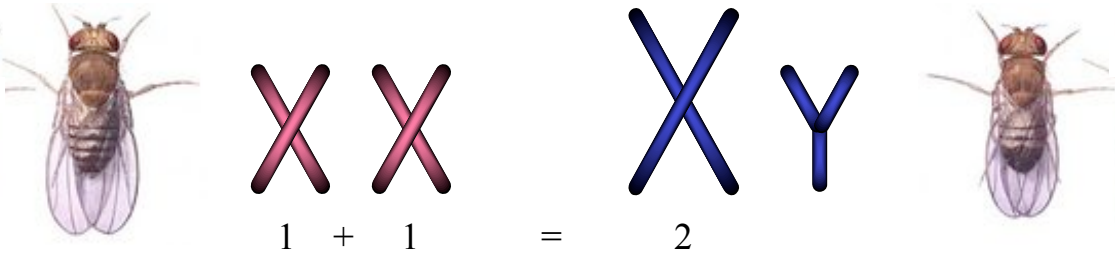
Compensation de dose chez la drosophile

Chez la drosophile, le sexe est déterminé par le calcul du ratio entre le nombre de chromosomes X et le nombre d'autosomes. Un complexe de cinq protéines est impliqué dans le mécanisme de compensation de dose : le complexe MSL (Male Sex Lethal, (Copps et al., 1998)).

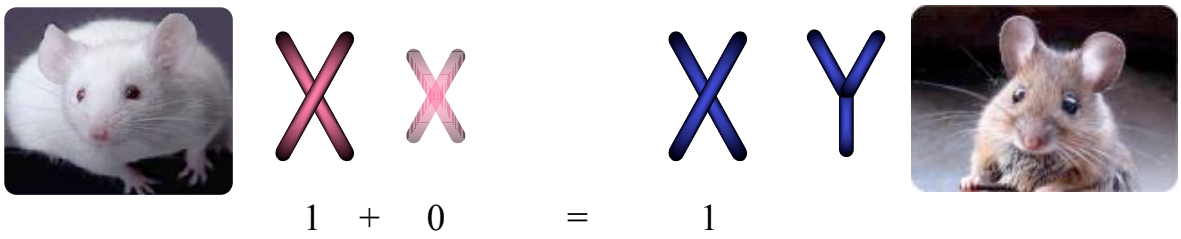
Le complexe MSL comporte entre autre une activité hélicase à ARN dépendante de l'ATP, portée par la sous unité protéique MLE (Lee et al., 1997), et une activité HAT, via la protéine MOF (Akhtar and Becker, 2000). Deux ARN non codants interviennent aussi dans la formation du complexe MSL : les ARN roX1 et roX2 (Franke and Baker, 1999).

Le mécanisme de compensation de dose chez la drosophile est associé à un patron de modification des histones. Le chromosome X des mâles, actif en transcription, présente une

A. Drosophile: hypertranscription du chromosome X des mâles



B. Mammifères: inactivation du chromosome X des femelles



C. Nématode : hypotranscription des chromosomes X des hermaphrodites (femelles)

Figure 37 : les différents mécanismes de compensation de dose.

Le phénomène de compensation de dose a été découvert par H. Müller, et a ensuite été étudié à un niveau moléculaire dans différents organismes tels que la drosophile, les mammifères et le nématode.

A. Chez la drosophile, les gènes liés au chromosome X des mâles présentent un niveau de transcription environ deux fois supérieur à celui des femelles.

B. Chez les mammifères, un des deux chromosomes X de la femelle est inactivé et ce, dans toutes les cellules somatiques.

C. Chez le nématode, les niveaux de transcription des gènes liés au chromosome X sont diminués chez les hermaphrodites de manière à atteindre un niveau équivalent à celui présent chez les mâles.

forte acétylation de la lysine 16 de l'histone H4 (Bone et al., 1994; Suka et al., 2002), modification mise en place par la protéine MOF (Smith et al., 2000). De même, la phosphorylation de la sérine 10 de l'histone H3 semble être impliquée dans ce mécanisme de compensation de dose (Wang et al., 2001).

Des complexes de remodelage de la chromatine interviennent aussi dans le processus de compensation de dose, comme par exemple les complexes ACF, CHRAC et NURF. De même, la protéine SU(VAR)3-7, composant structural de l'hétérochromatine, semble jouer un rôle dans la compaction globale du chromosome X des mâles (Spierer et al., 2005).

Dans cet organisme, les mécanismes épigénétiques sont impliqués dans l'hyperactivation transcriptionnelle du chromosome X des mâles alors que chez les mammifères, les mécanismes épigénétiques interviennent dans la répression transcriptionnelle d'un des chromosomes X de la femelle.

Compensation de dose chez les mammifères

Dans les cellules de mammifères, un des deux chromosomes X est transcriptionnellement mis en silence (Lyon, 1961). Le chromosome X inactivé (Xi), aussi appelé « corpuscule de Barr », est caractérisé par une combinaison particulière de modifications épigénétiques tant au niveau des histones que de l'ADN.

Le Xi est enrichi en histone H3 diméthylée sur la lysine 9, marque de l'hétérochromatine et d'inactivité transcriptionnelle (Heard et al., 2001). Cette modification va permettre, dans les cellules humaines, le recrutement de la protéine HP1 (Peters et al., 2002), contribuant ainsi à la formation d'une structure hétérochromatinienne (Chadwick and Willard, 2003). De même, la triméthylation de la lysine 27 de l'histone H3 fait partie des modifications présentes sur les Xi (Gilbert et al., 2003) et elle permettrait le recrutement du complexe PRC1, responsable de la monoubiquitination de l'histone H2A sur le Xi (de Napoles et al., 2004; Fang et al., 2004). D'autres modifications sont aussi retrouvées spécifiquement sur le Xi : de manière plus générale, les lysines des histones y sont majoritairement hypoacétylées (Jeppesen and Turner, 1993), caractéristique couramment associée à une inactivité transcriptionnelle (Richards and Elgin, 2002).

Les variants de l'histone HA2 jouent aussi un rôle de marqueurs à la fois pour le Xi et pour le X activé (Xa). Les variants macroH2A sont présents au niveau des nucléosomes du Xi (Chadwick and Willard, 2001a; Costanzi and Pehrson, 1998; Costanzi and Pehrson, 2001). Par contre, les variants H2A-Bbd et H2AZ sont moins abondants sur le Xi que sur le Xa et sur les autosomes (Chadwick and Willard, 2001b; Chadwick and Willard, 2003). La méthylation de l'ADN au niveau des dinucléotides CpG semble jouer un rôle important dans le maintien de la mise en silence des gènes du Xi. Ces séquences sont hyperméthylées sur le Xi, et hypométhylées sur le Xa (Bartlett et al., 1991).

Tout comme chez la drosophile, le mécanisme de compensation de dose fait intervenir un ARN, l'ARN Xist, qui recouvre totalement le Xi et participe au maintien de son état inactivé (Brown et al., 1992). Cet ARN est transcrit uniquement à partir du futur chromosome Xi et est régulé par un ARN qui lui est antisens : l'ARN tsiX (Lee et al., 1999b), qui serait quant à lui impliqué dans le choix du chromosome X à inactiver (Sado et al., 2001).

Compensation de dose chez le nématode

Chez le nématode, le sexe est déterminé par un simple ratio entre le nombre de chromosome X et le nombre d'autosomes (X:A). Lorsque ce rapport est inférieur à 0,67, les animaux sont des mâles, si le rapport est supérieur à 0,75, ce sont des animaux hermaphrodites qui se développent et qui sont capables de s'auto-fertiliser (Cline and Meyer, 1996). Un mécanisme de réduction de la transcription est alors mis en place dans les animaux hermaphrodites permettant de diminuer la quantité de transcrits provenant des chromosomes X.

Ce mécanisme de réduction de la transcription impliquerait un complexe de type *Condensin* (Chan et al., 2004) appelé DCC (Dosage Compensation Complex). Ce DCC s'assemble spécifiquement sur le chromosome X de manière dépendante du ratio X:A (Cline and Meyer, 1996). Ce recrutement se fait grâce à la reconnaissance de séquences particulières, spécifiquement présentes sur les chromosomes X (Csankovszki et al., 2004). Récemment, il a été démontré que le complexe DCC comportait une activité HMT spécifique pour la lysine 4 de l'histone H3.

f. Le locus des gènes ribosomiques

Plus de la moitié des ARN transcrits dans une cellule correspond à des ARN ribosomiques (ARNr) (Moss, 2004). Les gènes qui codent pour les ARNr existent dans deux états distincts de chromatine : un état dit « ouvert » et qui correspond à une conformation qui permet la transcription, et un état « fermé » qui est réfractaire à la transcription. Des études récentes ont démontré que la mise en silence des clusters de gènes ribosomiques, arrangés en tandem inversé, faisait intervenir à la fois des facteurs protéiques en *trans*, mais aussi des mécanismes épigénétiques tels que la méthylation de l'ADN et les modifications des histones. Ces mécanismes auraient pour conséquences une altération de la structure chromatinienne et un contrôle du rapport de gènes ribosomiques actifs et inactifs.

Ces clusters, appelés NOR (Nucleolar Organizer Regions), contiennent plus d'une centaine de gènes (Reeder, 1974) et peuvent être localisés sur une seule paire de chromosomes, comme chez la drosophile où ils se trouvent sur les chromosomes sexuels (Ritossa and Spiegelman, 1965) et la levure *S. cerevisiae* sur le chromosome XII, ou sur plusieurs paires de chromosomes, comme chez les mammifères.

La transcription de ces gènes par l'ARN Polymérase I (Pol I) entraîne la synthèse d'un long précurseur ARN, qui après maturation donnera trois des quatre formes matures : les ARNr 5,8S, 18S et 28S chez les mammifères et 5,8S, 18S et 25S chez la levure (figure 38). Ces étapes de synthèse et de maturation sont, entre autre, à l'origine de la formation du nucléole. Un quatrième ARNr mature, l'ARNr 5S transcrit indépendamment par l'ARN Polymérase III (Pol III), s'associera avec ces ARNr ainsi qu'à des protéines ribosomiques afin de créer le ribosome (Moore and Steitz, 2002), ribonucléoparticule responsable de la traduction.

La quantité de gènes ribosomiques transcrits dans une cellule dépend de la demande en ribosomes matures afin de permettre la traduction. Dans les cellules en croissance, la transcription est abondante alors que dans les cellules quiescentes, la transcription est réduite. Deux mécanismes de contrôle de la transcription ont été évoqués : soit la cellule module le nombre de transcrits issus d'un même gène, soit le nombre de gènes actifs varie. Des études de microscopie électronique ont démontré que des gènes actifs pouvaient être adjacents à des gènes inactifs (French et al., 2003), suggérant que seule une sous population de gènes

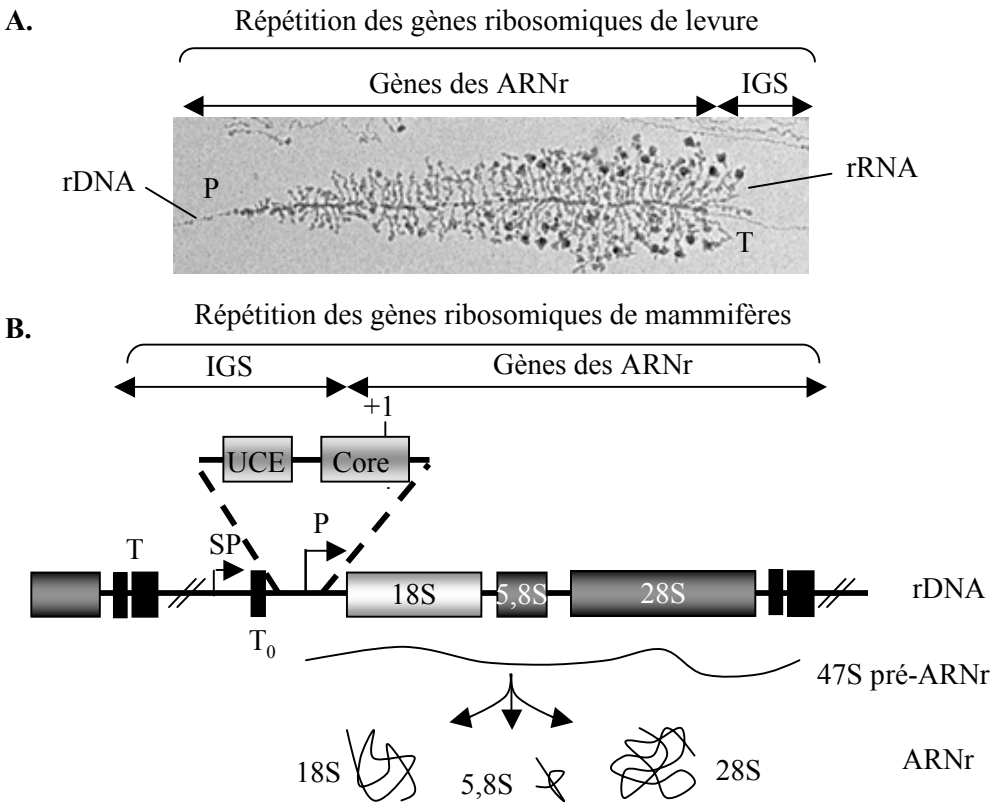


Figure 38 : organisation des gènes codant pour les ARN ribosomiques.

A. La nature répétitive des gènes codant pour les ARN ribosomiques (rDNA) est illustrée par une image de microscopie électronique à partir d'un étalement de chromatine de levure (étalement de Miller ou « arbre de Miller »). Progressivement, des ARN ribosomiques (ARNr) de plus en plus longs émanent des complexes ARN Polymérase I (Pol I) au cours de la transcription qui commence au promoteur (P) et se termine au terminateur (T).

B. Représentation d'une répétition rDNA de mammifères. Chaque répétition de 43 kb comporte un espace intergénique (IGS) de 30 kb contenant les éléments régulateurs de la transcription, et la séquence codante de 13 kb (gènes des ARNr). Les gènes codant les ARNr sont présents sous forme d'unité transcriptionnelle unique, transcrite par la Pol I et générant un précurseur des ARNr de 47S. De manière concomitante à sa transcription, ce précurseur est modifié par méthylation et pseudouridinylation puis clivé afin de générer les ARNr matures de 18S, 5,8S et 28S. La Pol I débute sa transcription au niveau du promoteur (P), qui comporte le « core element » (Core) dans lequel se trouve le point de départ de la transcription (+1) et une élément de contrôle amont (UCE). Un autre promoteur amont est présent (SP) et permet la transcription d'un transcrit de fonction inconnue par la Pol I. Des nombreux éléments de terminaison de la transcription sont situés dans la partie 3' de la région transcrite (T), mais aussi immédiatement en amont du point de départ de la transcription (T0).

ribosomiques était active. Des mécanismes épigénétiques semblent être impliqués dans le passage entre un état activé et un état inactivé.

L'hypothèse de la régulation de la transcription par la structure de la chromatine est basée sur des observations reposant sur l'utilisation du psoralène, un agent qui réagit de manière photoréactive avec les acides nucléiques doubles brins (Toussaint et al., 2005). Cet agent a permis de mettre en évidence deux classes de gènes ribosomiques. La première correspond aux gènes sensibles au psoralène : l'ADN est majoritairement dépourvu de nucléosomes, car ceux-ci sont irrégulièrement espacés, et des transcrits naissants y sont détectés (Conconi et al., 1989). Les gènes de l'autre classe présentent au contraire une insensibilité au psoralène car les nucléosomes sont plus régulièrement espacés donc l'ADN est moins accessible. Cette classe correspond aux gènes inactifs en transcription. Pour autant, ces deux types de gènes rentrent dans la formation des NOR, ce qui élimine une hypothèse d'inactivation des gènes ribosomiques sur une large région (French et al., 2003). De plus, il a été démontré qu'après chaque étape de réplication, les marques épigénétiques sont transitoirement effacées (voir ci-dessous), puis remises en place (Lucchini and Sogo, 1994; Lucchini and Sogo, 1995).

Certaines protéines interagissant avec les régions en amont des promoteurs des gènes ribosomiques jouent un rôle à la fois dans le contrôle de la transcription, mais aussi dans le remodelage chromatinien. La protéine TTF-I (Transcription Terminator Factor I) est impliquée à la fois dans l'initiation de la transcription des ARNr dans leur contexte chromatinien, mais aussi dans le positionnement des nucléosomes au niveau des promoteurs de ces gènes (Langst et al., 1998).

Les modifications des histones, et plus particulièrement de la lysine 9 de l'histone H3 jouent un rôle central dans la mise en place d'une structure chromatinienne inactive au niveau des gènes ribosomiques à mettre en silence (Prior et al., 1983). La méthylation de la lysine 9 de l'histone H3 est présente au niveau des gènes inactifs en transcription (Santoro et al., 2002). Par contre, les nucléosomes présents au niveau des gènes nouvellement activés sont hyperacétylées et comportent le variant d'histone H3.3, traduisant une activité transcriptionnelle suite à une absence de méthylation sur la lysine 9 de ce variant (Ahmad and Henikoff, 2002). Dans ces expériences, l'incorporation du variant d'histone H3.3 fait suite à une reprise de croissance de cellules en culture préalablement maintenues en phase stationnaire. Ce mécanisme de remplacement d'histone pourrait suggérer l'existence d'une

voie d'activation rapide de certaines copies de gènes ribosomiques suite à un stimulus cellulaire et ce afin d'enlever les marques épigénétiques de répression de la transcription.

Tout comme au niveau des séquences satellites, des facteurs permettant de reconnaître spécifiquement l'ADN sont impliquées dans le ciblage des mécanismes de mise en place d'une structure particulière de la chromatine : la protéine UBF (Upstream Binding Factor, (Jantzen et al., 1990)), une protéine à boîte HMG lui permettant de se lier au petit sillon de l'ADN, et la protéine SL1 (Selectivity Factor 1, (Learned et al., 1985)) interagissant avec les protéines TBP et trois TAF (Transcription Associated Factor) (Moss and Stefanovsky, 2002). Ces protéines sont des homologues des protéines de levure UAF (Upstream Activating Factor), associée à la protéine TBP (TATA-Box Binding Protein), et CF (Core Factor). Ces deux facteurs se lient respectivement aux séquences UCE (Upstream Control Element) et CE (Core Element) (Nomura, 2001) afin de permettre l'initiation de la transcription.

La protéine UBF est connue pour être un activateur de la transcription des ARNr (Kuhn and Grummt, 1992; Kuhn et al., 1993) en permettant le recrutement de la Pol I, stabilisant la liaison du facteur SL1 et en rentrant en compétition avec des protéines telles que l'histone H1 (Kermekchiev et al., 1997). Les protéines UBF et SL1 s'assemblent au niveau des promoteurs pour former le PIC (Pre-Initiation Complex) qui sera reconnu par la Pol I (figure 39).

La protéine Tip5 du complexe NoRC (Nucleolar Remodeling Complex) possède des motifs AT Hook lui permettant la liaison au petit sillon de l'ADN riche en dA•dT et intervient aussi dans cette régulation chez les mammifères (Zhou et al., 2002). Le complexe NoRC, un facteur de remodelage de la chromatine dépendant de l'ATP, est impliqué dans la formation d'une structure de type hétérochromatinienne permettant la mise en silence de la transcription des gènes ribosomiques (Santoro et al., 2002). Cette inactivation serait rendue possible par le recrutement d'une activité histone désacétylase au niveau du rDNA (Hirschler-Laszkiewicz et al., 2001; Mayer et al., 2006; Zhou et al., 2002).

La régulation de la transcription implique aussi des protéines comme la Topoïsomérase I qui, associée à Arf une protéine suppresseur de tumeur, peut aussi intervenir dans la biogenèse des formes matures des ARN ribosomiques (Ayrault et al., 2004). La Topoïsomérase II, décrite comme interagissant avec le rDNA chez la drosophile, intervient aussi dans cette régulation (Cavalli et al., 1996).

Enfin, il a été démontré que les nucléosomes avaient une structure particulière au niveau des copies des gènes ribosomiques activement transcrites. Au lieu d'être parfaitement

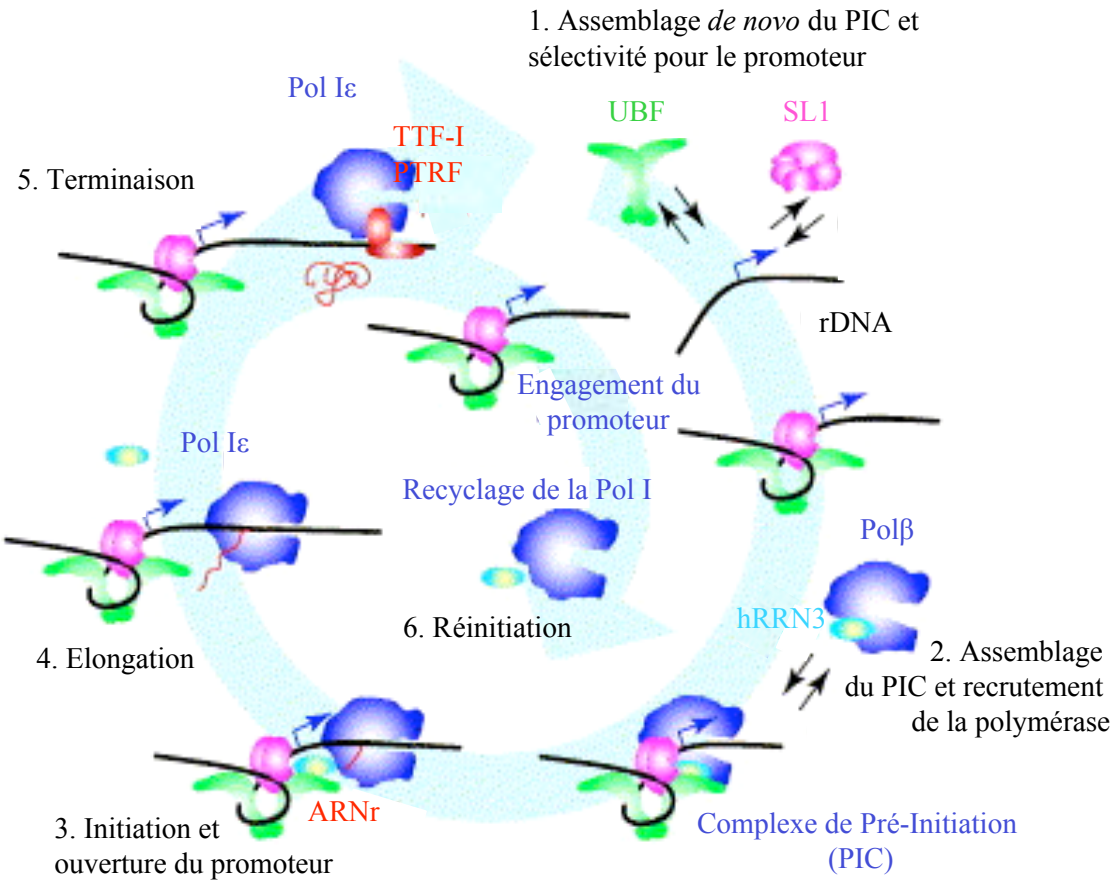


Figure 39 : cycle de transcription par l'ARN polymérase I (Pol I) (d'après Russel and Zomerdjik, 2005).

1. La formation *de novo* du complexe de pré-initiation implique la liaison sélective du facteur de sélectivité SL1 sur le promoteur du gènes codant les ARN ribosomiques (ARNr) et l'incorporation de la protéine activatrice de liaison amont (UBF).

2. L'étape suivante consiste au recrutement de la Pol Iβ par le facteur SL1.

3. La Pol I initie la transcription au niveau du promoteur « ouvert », puis elle est ensuite convertie en une forme processive appelée Pol Iε.

4. La forme processive Pol Iε est responsable de l'élongation du transcrit naissant.

5. La transcription par la Pol I se termine à l'extrémité 3' du gène au niveau d'une séquence spécifique, liée au facteur de terminaison de la transcription TTF-I et au facteur de relargage des transcrits PTRF. De manière concomitante, la Pol I et le précurseur ARNr sont libérés.

6. Les facteurs SL1 et UBF restent liés au promoteur après le passage de la Pol I et forment ainsi un support protéique permettant la réinitiation de la transcription par un autre complexe Pol Iβ. Ce complexe Pol Iβ pourrait provenir du recyclage des complexes Pol I associés à hRRN3. La formation d'un nouveau complexe PIC peut alors initier un autre cycle de transcription.

conformés, les nucléosomes y sont dépliés alors qu'il a été décrit que, lors de la transcription par l'ARN Polymérase II, les nucléosomes étaient désassemblés suite à la perte transitoire des dimères H2A/H2B (Kireeva et al., 2002).

Un autre mécanisme épigénétique pourrait intervenir dans la régulation de la transcription. La méthylation de l'ADN, principalement au niveau des dinucléotides CpG présents dans les séquences régulatrices UCE et CE localisées en amont du promoteur, serait un inhibiteur de la transcription des gènes ribosomiques (Houchins et al., 1997). La seule méthylation de la séquence UCE empêcherait la liaison de la protéine UBF (Santoro and Grummt, 2001), qui reconnaît normalement cette séquence et permet l'initiation de la transcription. L'activité cytosine méthyltransférase serait amenée par le complexe de remodelage de la chromatine NoRC.

Chez la drosophile, un autre mécanisme de régulation de la transcription est mis en place. Dans cet organisme, le rDNA se trouve uniquement sur les chromosomes sexuels et plus précisément à proximité des séquences SAT III du chromosome X (Hilliker and Appels, 1982). Par ailleurs, il représenterait la seule partie homologue entre les chromosomes X et Y, se qui permettrait leur appariement lors de la méiose.

Des rétrotransposons sans LTR (Long Terminal Repeat), appelés R1 et R2, sont insérés dans la partie du gène codant pour l'ARNr de 28S (Jakubczak et al., 1991). Leur présence empêche la synthèse d'un ARNr de 28S fonctionnel (Long et al., 1981).

R1 et R2 ont la capacité de rétrotransposer ce qui assure leur maintien dans ces séquences (Perez-Gonzalez and Eickbush, 2002). Pour autant ce phénomène n'est pas équivalent entre les deux séquences et il semblerait que la rétrotransposition de R1 soit prédominante chez le mâle, alors que pour R2, ce phénomène aurait la même importance dans les deux sexes (Perez-Gonzalez et al., 2003).

De manière inattendue, les évidences selon lesquelles les protéines de la mise en place de l'hétérochromatine sont présentes sur le rDNA de la drosophile m'ont amenée à m'intéresser aux mécanismes épigénétiques pouvant affecter la régulation de la transcription des gènes ribosomiques.

E. RESULTATS ET DISCUSSIONS

Dans un premier temps, l'objectif de mon travail de thèse était d'utiliser des molécules capables de lier le petit sillon de l'ADN en tant qu'outils d'études de l'hétérochromatine.

Ces molécules, les diamidines et les polyamides, nous ont permis de nous intéresser plus précisément aux effets du déplacement de la topoïsomérase II des séquences SAT III. Cette étude, de manière réciproque nous a permis de mieux comprendre les mécanismes d'action des ces molécules au sein des séquences riches en paires de bases dA •dT.

Après validation des techniques permettant l'utilisation de ces molécules, je me suis intéressée au rôle de la protéine D1 dans l'initiation de l'assemblage d'une structure hétérochromatinienne. Ainsi, la protéine D1 recruterait, directement ou indirectement, les protéines de mises en place de la structure hétérochromatinienne, mais aussi la topoïsomérase II sur les séquences SAT III de l'hétérochromatine péricentromérique du chromosome X de la drosophile.

Des résultats obtenus à la fois sur une lignée sauvage de drosophile et sur la lignée *white-mottled 4h* soumise à variégation d'effet de position, nous ont conduit à rechercher la présence d'un élément frontière entre l'euchromatine et l'hétérochromatine péricentromérique du chromosome X. J'ai démontré que cet élément frontière, appelé BE28, sépare dans un contexte sauvage l'euchromatine de l'hétérochromatine, et plus précisément du rDNA. Dans le contexte inversé de la lignée *white-mottled 4h*, cet élément frontière est déplacé à l'autre extrémité du bras euchromatinien du chromosome X.

De manière surprenante, des immunolocalisations réalisées sur ces deux lignées de drosophile montrant un patron de marquage mutuellement exclusif pour les protéine D1 et HP1 au niveau du rDNA, nous ont conduit à nous intéresser à la régulation de la transcription des gènes ribosomiques par des mécanismes épigénétiques.

I. De l'utilisation de nouveaux outils pour l'étude de l'hétérochromatine : « Target practice : aiming at satellites repeats with DNA minor groove binders »

1. Introduction

Les expériences de délétions ou de mutations des séquences satellites, constituant majeur des domaines hétérochromatiniens, ne sont pas possibles. Afin de pallier à ce problème, de nouvelles stratégies utilisant des outils synthétiques ont été élaborées.

Les travaux réalisés dans l'équipe (Aulner et al., 2002; Blattes et al., 2006; Monod et al., 2002) se sont basés sur deux stratégies relativement proches : l'utilisation de la protéine synthétique MATH20 (Strick and Laemmli, 1995) et des polyamides (Janssen et al., 2000a; Janssen et al., 2000b) générés dans le groupe d'U. K. Laemmli.

Ces deux méthodes d'étude des interactions entre les protéines et l'hétérochromatine sont basées sur des travaux antérieurs qui utilisaient un oligopeptide naturel : la distamycine (Käs et al., 1989; Käs et al., 1993). La distamycine, un antibiotique produit par *Streptomyces distallicus* (Hopwood, 1988), est un oligopeptide qui lie préférentiellement le petit sillon de l'ADN riche en paires de bases dA•dT. Cet antibiotique est un moyen de défense pour *S. distallicus*, dont le génome est riche en paires de bases dG•dC, contre des organismes possédant un génome majoritairement riche en paires de bases dA•dT. Il a été démontré que l'utilisation de cet antibiotique pouvait interférer avec de nombreuses interactions entre des séquences d'ADN riches en paires de bases dA•dT et les protéines qui les reconnaissent (Käs et al., 1989; Radic et al., 1992; Suzuki et al., 2001).

Par la suite, un corollaire à cette stratégie de défense a été reproduit grâce à la synthèse chimique de molécules capables de mimer les caractéristiques structurales et possédant le même type d'affinité pour les séquences riches en paires de bases dA•dT. Plusieurs classes de dérivés ont été générées, mais nous nous sommes intéressés plus particulièrement aux oligopyrroles et aux diamidines.

a. Les oligopyrroles

Les oligopyrroles ont été développés dans le groupe de U.K. Laemmli sur la base des règles de reconnaissance de l'ADN établies par P. B. Dervan (Dervan and Edelson, 2003; Wemmer and Dervan, 1997). Ainsi, ils ont généré des polyamides en couplant des noyaux pyrroles et des noyaux imidazoles, et ce afin de cibler spécifiquement les séquences satellites de l'hétérochromatine de *Drosophila melanogaster* (Janssen et al., 2000a; Janssen et al., 2000b). Par la suite, ils ont démontré que l'utilisation de telles molécules pouvait altérer un patron d'expression génique en ciblant spécifiquement des séquences non codantes, telles que les séquences satellites de la drosophile, enrichies au niveau de l'hétérochromatine péricentromérique. De plus, cette approche représente une alternative prometteuse aux thérapies actuelles car elle permet de cibler des séquences non codantes tout en ayant un fort impact sur l'expression de séquences codantes.

Le pyrrole/imidazole P31 a avant tout été conçu en tant que molécule servant de contrôle dans les expériences mettant en jeu l'oligopyrrole P9. L'oligopyrrole P31 reconnaît spécifiquement les séquences SAT V de la drosophile formées de la répétition de motifs simples $(AAGAG)_n$ et $(AAGAGAG)_n$. Les effets observés suite à l'ingestion de cet oligopyrrole par des larves correspondent exactement à ceux observés lors de la mutation du gène codant le facteur GAGA (GAF). Le facteur GAF est un régulateur de la transcription, plus particulièrement des gènes homéotiques (Espinas et al., 2000), et il joue aussi un rôle dans la division cellulaire (Bhat et al., 1996a). Il a donc été proposé que l'oligopyrrole P31 agirait en affectant la localisation du facteur GAF : ce dernier serait délocalisé de ses sites de liaison euchromatiniens vers les séquences SAT V (Janssen et al., 2000b). Ce mécanisme passerait par une « ouverture » de la chromatine permettant le démasquage de site de liaison du facteur au niveau des séquences SAT V. Ces sites lui sont normalement inaccessibles car ils sont impliqués dans une structure hétérochromatinienne « fermée ».

Le polyamide P9 est un oligopyrrole relié structuralement à la distamycine. Il reconnaît spécifiquement les séquences SAT I et SAT III de la drosophile et a un effet remarquable sur la variégation d'effet de position du gène *white* dans la lignée *white-mottled* (Janssen et al., 2000a). Dans cette inversion, le gène *white* est relocalisé à proximité du rDNA, région qui s'étend sur 3,5 Mpb et qui sépare le gène *white* ainsi repositionné et les séquences SAT III riches en paires de bases dA•dT. Suite à l'ingestion de polyamide P9 par

les larves de la lignée *white-mottled*, le phénotype de variégation d'effet de position est supprimé (Janssen et al., 2000a). Les mécanismes par lesquels le P9 supprime la variégation d'effet de position ont été caractérisés dans une autre étude, qui sera détaillée dans la partie suivante (voir chapitre II. et (Blattes et al., 2006)).

b. Les diamidines

Les résultats obtenus avec les polyamides ont ouvert une nouvelle voie d'étude de la structure hétérochromatinienne. Ainsi, nous avons établi une collaboration avec des chimistes américains, David Wilson et David Boykin, spécialistes de la synthèse des diamidines, appelées « DB », qui présentent aussi de fortes similarités structurales et une même spécificité que la distamycine. Ces composés sont des ligands du petit sillon de l'ADN riche en paires de bases dA•dT et ont tout d'abord été développés pour leur activité antiparasitaire, pour revue (Wilson et al., 2005). Ces drogues possèdent des effets antiprolifératifs marqués (Lansiaux et al., 2002a; Lansiaux et al., 2002b) et une activité antiparasitaire avérée contre *Plasmodium falciparum*, *Pneumocystis carinii* et *Trypanosoma brucei* (Bouteille et al., 2003; Fairlamb, 2003). Ces agents infectieux sont responsables respectivement du paludisme, de la pneumonie à PCP (forme opportuniste qui se développe chez les patients atteints du syndrome d'immunodéficience acquise (SIDA)), et de la maladie du sommeil. Ces trois maladies sont endémiques dans les pays en voie de développement, et la recherche de nouveaux traitements représente donc un axe fondamental pour les organismes de santé mondiale. En effet, les traitements utilisant des antibiotiques et anti-malariaux classiques sont confrontés à des problèmes de résistances.

Le projet de développement de ces diamidines fait l'objet de nombreux financements, publics et privés comme ceux fournis par la « Bill and Melinda Gates Fundation », qui ont ainsi autorisé la synthèse de différents composés, parmi lesquels les plus prometteurs sont deux dérivés de pentamidine : le DB293 et le DB75, aussi appelé furamidine (Wilson et al., 2005).

Ces drogues ont une forte activité antiparasitaire contre *P. falciparum* et *P. carinii*. En effet, les génomes de ces parasites sont fortement enrichis en paires de bases dA•dT. La cible de ces diamidines chez *T. brucei* est légèrement différente puisque ce n'est pas à proprement parler le génome qui est visé, mais plutôt les kinétoplastes (Shapiro and Englund, 1990). Ces

structures, des cercles d'ADN concaténés, sont enrichies en paires de bases dA•dT et nécessitent l'activité de la topoïsomérase II pour leur réplication et leur ségrégation (Bailly et al., 1999). Les molécules d'ADN ainsi structurées sont essentielles car elles portent les gènes codant les protéines ribosomiques et les enzymes intervenants dans la chaîne respiratoire.

Dans notre modèle d'étude, les séquences SAT III de la drosophile représentent une cible potentielle pour ces diamidines dont nous avons recherchés les effets de la fixation au niveau des séquences SAT III. Cette étude a été possible de par la capacité des diamidines à lier le petit sillon de l'ADN de ces séquences d'intérêts, mais aussi grâce à leur propriété d'autofluorescence lorsqu'ils sont liés à l'ADN. Les résultats de ces travaux ont également permis de proposer un mécanisme d'action anti-parasitaire pour le DB75 (voir chapitre III de cette même section).

2. *Objectif de l'étude*

L'étude menée avec les diamidines avait pour but de démontrer tout d'abord la spécificité de liaison de ces composés aux séquences SAT III de la drosophile, par des expériences à la fois *in vitro* et *in vivo*. J'ai ensuite recherché quelles pouvaient être les protéines affectées par la fixation des diamidines à ces séquences. Les données bibliographiques ainsi que les expériences déjà réalisées nous ont amenés à envisager la topoïsomérase II comme cible principale des diamidines. Des études *in vitro* tendant à démontrer l'implication de la topoïsomérase II dans la suppression de la variégation ont alors été réalisées en utilisant un des dérivés du DB75, le DB351.

3. *Résultats*

Publication parue dans « Current Medicinal Chemistry – Anti Cancer Agent » en 2005.

Target Practice: Aiming at Satellite Repeats with DNA Minor Groove Binders

Guillaume Susbielle, Roxane Blattes, Vanessa Brevet, Caroline Monod and Emmanuel Käs*

Laboratoire de Biologie Moléculaire Eucaryote, UMR 5099 CNRS/UPS, IFR 109, 118, route de Narbonne, 31062 Toulouse Cedex 4, France

Abstract: Much progress has been made in recent years in developing small molecules that target the minor groove of DNA. Striking advances have led to the design of synthetic molecules that recognize specific DNA sequences with affinities comparable to those of eukaryotic transcription factors. This makes it feasible to modulate or inhibit DNA/protein interactions *in vivo*, a major step towards the development of general strategies of anti-gene therapy. Examples from anti-parasitic drugs also suggest that synthetic molecules can affect a variety of cellular functions crucial to cell viability by more generally targeting vast portions of genomes based on their biased base composition. This provides a rationale for developing approaches based on selective interactions with broad genomic targets such as satellite repeats that are associated with structural or architectural components of chromatin essential for cellular proliferation. Using examples drawn from the *Drosophila melanogaster* model system, we review here the use of synthetic polyamides or diamidines that bind the DNA minor groove and can be used as highly selective agents capable of interfering with specific protein/DNA interactions that occur in A+T-rich repeated sequences that constitute a significant portion of eukaryotic genomes. The satellite localization of cellular proteins that bind the minor groove of DNA *via* domains such as the AT hook motif is highly sensitive to these molecules. A major consequence of the competition between these proteins and their synthetic mimics is an alteration of the nuclear localization and function of proteins such as topoisomerase II, a major target of anti-cancer drugs.

Key Words: DNA minor groove, polyamides, diamidine antiparasitic drugs, AT hook motif, satellite DNA, heterochromatin, silencing, topoisomerase II.

DRUG/DNA *VS* PROTEIN/DNA INTERACTIONS

Two meetings on the theme of "Drugs acting on nucleic acids" were held in 1991 and 1994 as part of the Jacques Monod Conferences organized by the CNRS (the French National Center for Scientific Research). Drawing scientists from a variety of backgrounds, these meetings highlighted exciting new advances, but also a broad chasm that existed between the development and characterization of molecules capable of interacting with nucleic acids with variable affinities or specificities and their possible biological applications. As one conference speaker -a biologist- rather dismissively put it during the last of these two meetings (which have not been held again since), proteins -not drugs- act on nucleic acids *in vivo*.

Ten years later, it is clear that the gap between chemistry and biology has been greatly reduced, if not eliminated. The development of highly specific small molecules that bind DNA has led to the demonstration that they could interfere with protein/DNA interactions *in vitro* and *in vivo* [1]. Some of these molecules mimic naturally occurring DNA-binding motifs and thus act as powerful competitors capable of selectively disrupting the association of their target sequences with the proteins that recognize them. We will focus here on a review of the possible uses of two classes of small synthetic molecules that act on nucleic acids: the mixed pyrrole/imidazole polyamides initially developed by P.B. Dervan and collaborators (reviewed in [2-4]) and antiparasitic diamidines developed by D. W. Boykin and W. D. Wilson (see W. D. Wilson *et al.* in this issue). As the synthesis and applications of some of these molecules and their derivatives are extensively reviewed in this issue, we will address instead novel strategies based on their use as general tools directed against bulk regions of eukaryotic genomes which, while non-coding, serve as major interaction sites for structural/architectural components of chromatin important for cell growth and proliferation.

Such a strategy has been used in the course of evolution time and time again. One example relevant to our discussion is that of distamycin [5, 6]. This crescent-shaped oligopeptide preferentially binds the minor groove of dA•dT sequences and is produced by *Streptomyces distallicus*, whose own genome is highly G+C-rich [7, 8]. This toxic antibiotic potentially interferes with a broad range of protein/DNA interaction involving A+T-rich sequences [9, 10]. One can speculate that the general preference of distamycin for dA•dT sequences has been selected as part of a strategy for competing with other microorganisms whose genome might be A+T-rich, while largely protecting G+C-rich *S. distallicus* from self-destruction by its own potentially lethal product.

A corollary of using such a strategy is illustrated by the biological activity of diamidines and their derivatives (see W. D. Wilson *et al.*, this issue). These molecules are highly active against several pathogenic parasites [11-15] whose

*Address correspondence to this author at the LBME UMR 5099, 118, route de Narbonne, 31062 Toulouse Cedex 4, France; Tel: +33 (0)5 61 33 59 59; E-mail: kas@ibcg.biotoul.fr

1568-0118/05 $50.00+.00 © 2005 Bentham Science Publishers Ltd.

common characteristic is an extraordinarily A+T-rich genome or portion thereof [16, 17]. As in the case of distamycin, diamidines are also expected to interfere in a potent way with the general metabolism of genomes whose base composition is highly skewed in favor of interactions with the drugs. In this case, the selective toxicity of these synthetic minor groove binders is determined by the biased DNA base composition of their targets, an observation that can be used as a paradigm for targeting specific subsets of genomes based not on DNA sequence *per se*, but on overall base composition.

TARGETING A+T-RICH SEQUENCES: EARLY EXPERIMENTS WITH DISTAMYCIN

One of the early uses of distamycin in biological investigations was for the purpose of "statistical sequencing," a method developed by A. A. Travers and collaborators [18]. Statistical sequencing takes advantage of the fact that distamycin binds consecutive dA•dT base pairs highly selectively and protects these sequences from digestion by DNAse I. This results in footprinted regions that are indicative of highly A+T-rich sequences in a given DNA fragment. Similarly, G+C-rich stretches can be detected as regions protected from DNAse I digestion by chromomycin A3, which selectively binds dG•dC base pairs. Thus, protection from nuclease digestion in the presence of these drugs readily identifies regions of biased base composition. As DNAse I is itself a minor groove binder [19, 20], a corollary of the statistical sequencing method is that a ligand such as distamycin should be able to affect proteins that bind A+T-rich sequences *via* minor groove interactions. This was indeed shown to be the case for the interaction of sequences such as satellite repeats or Scaffold-Associated Regions (SAR or MAR sequences that are believed to form the base of chromatin loops in eukaryotic cells, reviewed in [21]) with proteins such as histone H1 and topoisomerase II, which preferentially bind these AT-rich sequences [22-24]. Later experiments demonstrated secondary effects of AT-binding by distamycin, which was also shown to interfere with the formation of complexes involving the major groove, including homeodomain/DNA complexes, presumably as a result of drug-induced changes in the conformation of DNA [25, 26]. More generally, numerous studies identified distamycin as a potent regulator of DNA/topoisomerase II interactions at the level of both DNA binding and cleavage, suggesting that the activity of this important enzyme, a major target of anticancer drugs, could be regulated by minor groove ligands [27-34].

CELLULAR DISTAMYCIN "ANALOGUES:" IDENTFYING AT-SPECIFIC PROTEINS

Given the ability of distamycin to selectively perturb protein/DNA interactions involving AT-rich sequences, an intriguing possibility is that there might exist cellular analogues of distamycin capable of binding to AT-rich genomic regions [31]. While initially highly speculative, this hypothesis was first validated following the characterization of the AT hook DNA-binding domain (reviewed in [35, 36]), whose planar crescent structure resembles that of distamycin. The AT hook motif was first discovered in the mammalian HMGA proteins (formerly known as HMG-I/Y proteins) [37], which contain three copies of this domain and constitute the archetype for the multi-AT hook (MATH) proteins found in all living organisms but absent in viruses [38]. HMGA was indeed found to share with distamycin the ability to compete with histone H1 for binding to AT-rich DNA such as SAR/MAR sequences [39]. HMGA proteins are highly expressed in proliferating cells but not in differentiated quiescent cells (see [36] for review). However, they are re-expressed in tumor cells and their aberrant expression has been correlated with various tumors. As such, they constitute a potential target for anticancer therapy (reviewed in [40, 41]).

HMGA proteins have been shown to possess a striking range of biological functions, including the assembly of enhancesomes required for the transcriptional activation of certain genes or of specific complexes involved in transcriptional repression (see [35, 36] for review) and even in the integration of the HIV-1 retrovirus [42, 43]. Aside from these diverse activities, the predominant localization of HMGA proteins in AT-rich alpha satellite repeats, at the origin of their identification [44], has eluded a functional characterization. A similar situation is encountered in *Drosophila melanogaster*, where a structural homologue of HMGA, the D1 protein [45, 46], is exclusively associated with the major AT-rich satellites found in this organism, in particular the alpha-like 1.688 g/cm^3 359-bp satellite III (SAT III) repeat and the simple $(AATAT)_n$ satellite I (SAT I) repeats (see Fig. (**1**)) [47, 48]. D1, which contains 10 AT hook motifs [49], thus appears to be the *Drosophila* orthologue of HMGA and the characterization of its biological function(s) was recently made possible using a combination of chemical, genetic, molecular and cell biology approaches [48, 50], which will be discussed below.

IMPROVING ON DISTAMYCIN: SYNTHETIC POLYAMIDES THAT TARGET SATELLITE REPEATS

The specificity of distamycin for dA•dT sequences and its general preference for AT-rich satellites is compromised by its ability to bind dG•dC-containing sequences non specifically at higher concentrations [5]. This relative promiscuity probably explains its lack of suitability as a therapeutic agent against viruses or cancer cells due to its high toxicity. Building on the DNA recognition rules developed by P.B. Dervan and collaborators (reviewed in this issue and in [24]), U. K. Laemmli and collaborators reported on the remarkable activities of synthetic oligopyrrole and mixed pyrrole/imidazole polyamides designed to target satellite repeats of *Drosophila melanogaster* [51, 52]. Their landmark experiments yielded two key findings: such polyamides can be ingested by developing organisms as complex as a fruit fly; once ingested, they will induce loss- or gain-of-function phenotypes consistent with their DNA-binding specificity (see [53] for review). Most importantly, they demonstrate that it is possible to alter overall patterns of gene expression by targeting non-coding satellite repeats highly enriched in pericentric heterochromatin, an intriguing alternative to direct anti-gene therapies.

The P31 pyrrole/imidazole polyamide [51] was designed to bind the $(AAGAG)_n$ and $(AAGAGAG)_n$ repeats of the *Drosophila* satellite V (SAT V, see Fig. (**1**)). In the appro-

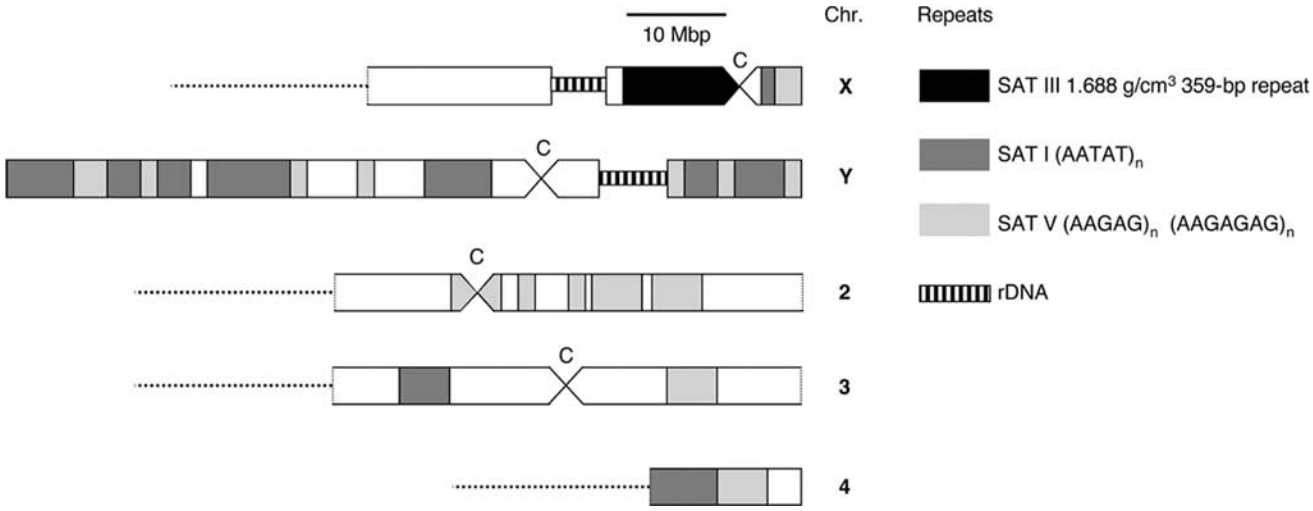


Fig. (1). ***Drosophila melanogaster*** **heterochromatin and major satellite repeats.** The schematized pericentric heterochromatin regions of the *Drosophila* chromosomes show the localization of the major satellite repeats discussed in this review. SAT III, I and V sequences are represented by black, dark gray and light gray boxes, respectively, rDNA repeats located on the X and Y chromosomes by hatched boxes. White boxes indicate sequences distinct from satellite repeats. These maps are adapted from *in situ* hybridization data [47]. Note that the Y chromosome is composed almost exclusively of heterochromatin and is shown here in its entirety.

priate genetic background, feeding P31 to flies resulted in well-defined homeotic transformations that mimic the effects of mutations of the gene encoding GAGA factor (GAF) [52]. The effect of P31 was proposed to be mediated by a redistribution of GAF from its binding sites in euchromatin, where this protein acts as a positive transcription activator, to SAT V. The loss-of-function phenotypes induced by P31 are thus consistent with a reduced availability of GAF for gene expression due to the massive recruitment of GAF to SAT V repeats spanning millions of base pairs. In this case, it was proposed that drug-induced "opening" of SAT V heterochromatin unmasks AAGAG sites recognized by GAF [51].

The P9 polyamide is an oligopyrrole structurally related to distamycin that highly specifically targets the *Drosophila* SAT I and SAT III repeats (Fig. (**1**)) [51]. P9 was shown to exert a profound effect on position-effect variegation (PEV) [52]. PEV arises when a euchromatic gene is juxtaposed to heterochromatin as a result of a chromosomal rearrangement or gene insertion into a host genome. This phenomenon denotes the heterochromatin-mediated silencing of the gene, reflected by its stochastic inactivation and mosaic expression which, once established, is stably propagated to daughter cells (see [54, 55] for review). A classical example of PEV in *Drosophila* is the *white-mottled* inversion, which places the *white* gene, responsible for the red eye color of wild-type flies, close to rDNA repeats in X-chromosome pericentric heterochromatin (Figs. (**1**) and (**3**)), within approximately 3.5 Mbp of the SAT III repeat array [56]. This localization leads to a partial silencing of the gene, resulting in its variegated "mottled" expression, a phenotype that was suppressed by feeding P9 to developing *white-mottled* larvae [52]. P9-induced suppression of PEV is reflected by the recovery of normal or near-normal expression of the *white-mottled* allele, a striking gain-of-function similar to that resulting from mutations in genes encoding proteins implicated in silencing, such as heterochromatin protein 1 (HP1) [55].

These remarkable results rest on the ability to design and synthesize DNA-binding molecules of the appropriate specificity and affinity (polyamides typically have apparent dissociation constants in the nanomolar range or below, comparable to those of eukaryotic transcription factors). The selectivity of polyamides is enhanced by the unexpected accessibility of the DNA minor groove in chromatin, even in its most compact forms -heterochromatin and mitotic chromosomes. This is illustrated in Fig. (**2**), which shows staining of interphase and prophase nuclei and mitotic chromosomes with Texas Red-labeled P31 (P31TR, red) and fluorescein-labeled P9 (P9F, green) (see also [57, 58]). In the case of mitotic chromosomes, the staining pattern observed is fully consistent with the localization of SAT I, SAT III and SAT V repeats derived from high-resolution *in situ* hybridization data [47]. Wide-field deconvolution microscopy followed by 3-D reconstruction (lower photographs) also reveals discrete SAT I, SAT III and SAT V territories, an approach which might be useful to map the localization of active or silenced genes relative to heterochromatin domains. Thus, in addition to their ability to affect gene expression *via* satellite-repeat binding, these molecules can also serve as powerful cytological staining tools, as also illustrated by the development of polyamides capable of selectively recognizing either insect or mammalian telomeric repeats [59].

MOLECULAR CONSEQUENCES OF POLYAMIDE BINDING TO SATELLITE REPEATS

As in the case of the loss-of-function phenotypes induced by P31 as a result of SAT V binding [51, 52], suppression of PEV mediated by P9 was proposed to act through SAT III "chromatin opening" and a subsequent reduction of silencing effects on the relatively nearby *white* gene. Similar effects had been previously observed following expression in *Drosophila* of MATH20, an artificial protein containing 20 tandem repeats of the AT hook motif derived from the

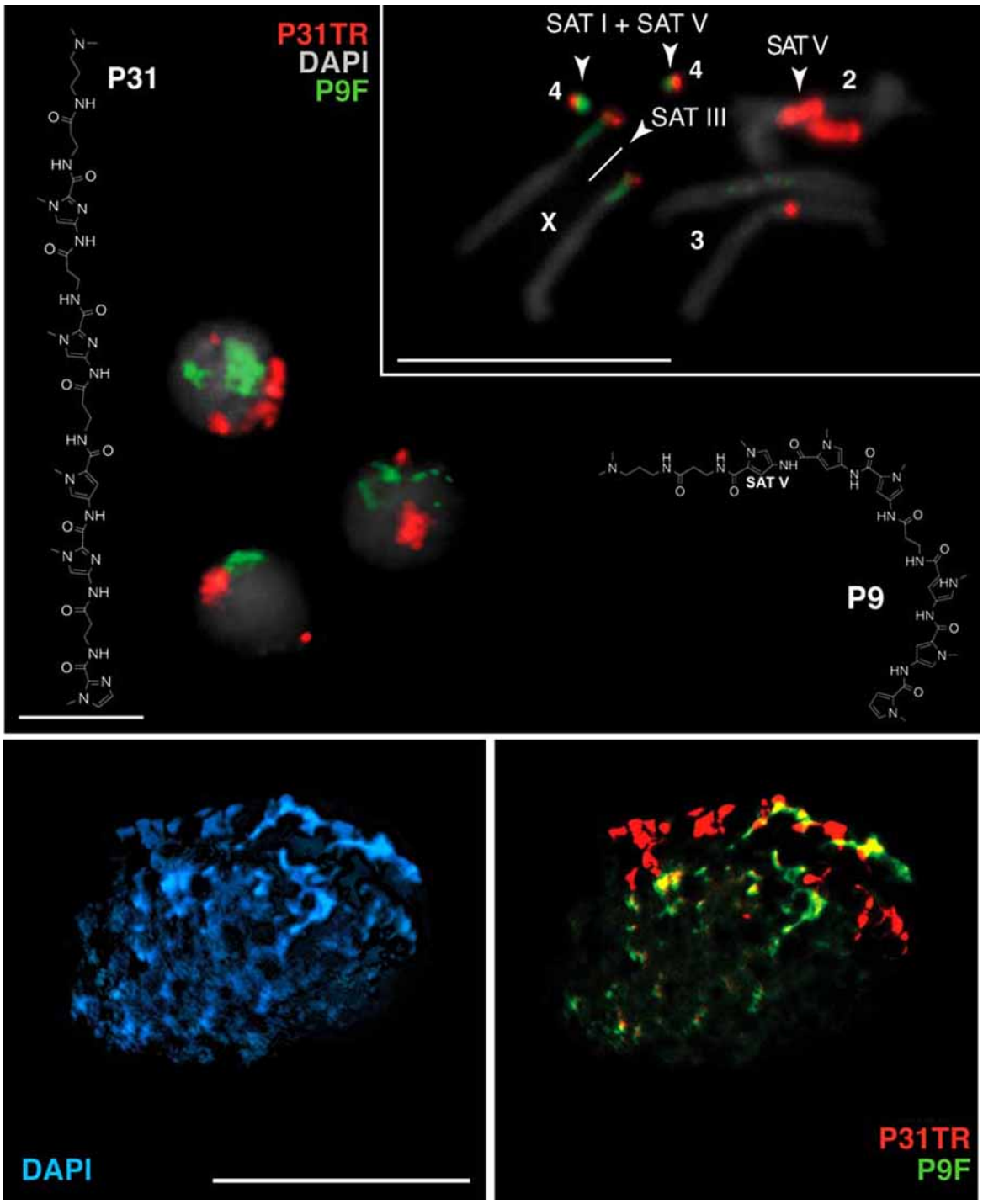


Fig. (2). Localization of SAT repeats in *Drosophila* nuclei and chromosomes. Panel A: dye-conjugated polyamides were used to localize SAT repeats in nuclei and chromosomes from fixed larval brains. The overall structure of P9 is $(Py)_3$-β-$(Py)_3$-β-Dp, that of P31 is Im-β-Im-Py-β-Im-β-Im-β-Dp, where Py denotes a pyrrole, Im an imidazole, β is β-alanine and Dp is dimethylaminopropylamide [51]. P31TR (2 nM), a Texas Red-conjugated P31 derivative (red), detects SAT V $(AAGAG)_n$ or $(AAGAGAG)_n$ repeats located predominantly on chromosome 2 and, to a lesser degree, on chromosomes 4 and 3 and at the very tip of the X. P9F (5 nM), a fluorescein-conjugated P9 oligopyrrole (green), detects the prominent SAT III-repeat array in the pericentric region of the X chromosome and $(AATAT)_n$ SAT I repeats on chromosome 4, at the tip of the X and on both sides of the centromeric regions of chromosome 3. Samples were counterstained with DAPI (gray). Comparison of the staining pattern detected on chromosomes with the map shown in Fig. (**1**) demonstrates the remarkable specificity of these molecules, which can recognize their targets even in highly compact chromatin. Panel B shows the DAPI (blue) and merged P31TR (red) and P9F (green) signals in an early prophase nucleus examined by wide-field deconvolution microscopy and 3-D reconstruction. Discrete SAT I/III and SAT V territories are clearly visible and define largely non-overlapping heterochromatin domains. Scale bars: 5 μm.

HMGA protein [60, 61]. A plausible explanation for these effects is that SAT III chromatin opening by MATH20 and P9 might be due to displacement of an endogenous protein specifically bound to SAT III repeats and which might be a target for direct competition by either molecule [50].

The HMGA-like D1 protein of *Drosophila* is predominantly associated with SAT III repeats *in vivo* [46, 48]. The hypothesis that its interaction with SAT III might be the major target of both MATH20 and P9 is supported by several observations. First, inactivation of the gene encoding D1 also results in suppression of *white-mottled* PEV, the first example of a DNA-binding protein possessing the appropriate specificity that affects heterochromatin-mediated silencing [48]. In addition, a similar effect is obtained following over-expression of D1ΔE, a protein that lacks the C-terminal acidic domain of wild-type D1 and acts as a dominant negative mutant form of the full-length protein. This suggests that the biological function of D1 rests on two separable activities: a DNA-binding function provided by its 10 AT hook motifs and a secondary function implicating its C-terminal domain, most likely required for protein/protein interactions. Finally, it was shown that D1 and MATH20 indeed compete for identical binding sites on SAT III repeats and that suppression of *white-mottled* PEV by MATH20 correlates with the displacement of D1 from SAT III sequences [50]. More recently, results of experiments carried out in our laboratory demonstrated that P9 similarly competes with D1 for binding to SAT III, further suggesting that the D1 protein is indeed the main target of P9 and that competition for binding to similar sites in DNA constitutes the main mechanism whereby this synthetic compound can profoundly affect gene expression.

This work, details of which will be published elsewhere, provides additional information on the mechanism of action of the P9 oligopyrrole. The D1 protein is essential for *Drosophila* development [50] and results summarized above suggest that its function is mediated by DNA binding and recruitment of other proteins important for heterochromatin assembly. Using P9 to displace D1 selectively in an *ex vivo* system of permeabilized *Drosophila* larval brains, we indeed found that competition by P9 for binding to SAT III repeats results in the loss of other heterochromatin-associated proteins, including HP1 and, more unexpectedly, topoisomerase II. While topoisomerase II has long been known to be associated with SAT III repeats [30, 62, 63], the significance of this localization has remained unclear. Our results suggest that the enzyme functions as both a structural and functional component of heterochromatin, classically defined as the portion of the genome which fails to decondense after mitosis. As topoisomerase II is required for chromosome condensation [64] (see also [65] and references therein), it is possible that its association with heterochromatin is required to maintain pericentric regions in a highly compacted state. A role for topoisomerase II in heterochromatin-mediated silencing is also supported by our observation that feeding *white-mottled* larvae VM26, a topoisomerase II poison, at sublethal doses which do not interfere with development, results in a strong suppression of PEV (to be published elsewhere).

These recent results are summarized in Fig. (**3**), which illustrates our current model for heterochromatin assembly in the pericentric region of the *Drosophila* X chromosome. We propose that D1 initiates heterochromatin assembly by binding to SAT III repeats. The acidic C-terminal domain of D1, which resembles that found in HMGA proteins and is required for the biological activity of the *Drosophila* protein [48], most likely serves to recruit other proteins that are key players in determining the functional properties of heterochromatin and include the silencing HP1 protein [54, 55] and topoisomerase II. Competition by MATH20, D1ΔE or P9 results in the displacement of D1 and the subsequent loss of HP1 and topoisomerase II from SAT III repeats. Note that the nearby rDNA repeats, which span approximately 3.5 Mbp, are not affected by this competition and normally remain associated with HP1 and topoisomerase II. Most importantly, the 3' end of the *white-mottled* allele of the *white* gene is located about 20 kbp from the rDNA array [56], but its expression (schematized in the Figure as an OFF/ON state) correlates with the protein composition of SAT III repeats, but not with that of the rDNA array. It is unclear at present how the loss of HP1 and topoisomerase II from SAT III repeats can be transduced to the *white* gene over 3.5 Mbp away. One possible explanation, consistent with spatial models for heterochromatin assembly and propagation [66-68], is that the compartmentalization of the rDNA repeats in the nucleolus places the *white-mottled* allele in close proximity to the SAT III array, making it sensitive to the effects of competitors such as P9 or MATH20.

DRAWING FROM THE EXAMPLE OF ANTIPARASITIC DRUGS: THE BIGGER THE TARGET, THE BETTER

Satellite repeats, because of their highly biased base composition and the huge chromosomal distances they span, can serve as enormous sinks for small molecules of the appropriate specificity. While they are usually located in highly condensed constitutive heterochromatin, they are readily available for interaction with minor groove binders such as polyamides [57, 58], with pronounced effects on global patterns of gene expression [52]. As such, the concept of targeting satellite repeats to develop new therapeutic approaches to the treatment of human diseases is remarkably similar to the strategies evolved by the microorganisms that produce DNA-binding antibiotics that chiefly target base composition rather than specific sequences.

More to the point, the development of diamidines used as antiparasitic drugs rests on a similar overall strategy of targeting genomes that exhibit a highly skewed base composition. The genomes of *Pneumocystis carinii* or *Plasmodium falciparum* are extraordinarily AT-rich (approximately 70% and 80% A+T content, respectively) [69], while the kinetoplast circles of trypanosomes and Leishmania are also highly AT-rich [70]. In this type of biological warfare using small molecules that preferentially -but not exclusively- interact with the minor groove of AT-rich sequences, the potential size of the target is clearly an important consideration. As such, diamidines might similarly exert biological effects on those portions of the genomes of higher eukaryotes that are highly AT-rich, of which satellite repeats constitute one extreme example. Alternatively, they might also serve as useful cytological tools because of their fluorescence properties when bound to DNA (W. D. Wilson *et al.*, this issue, and see below).

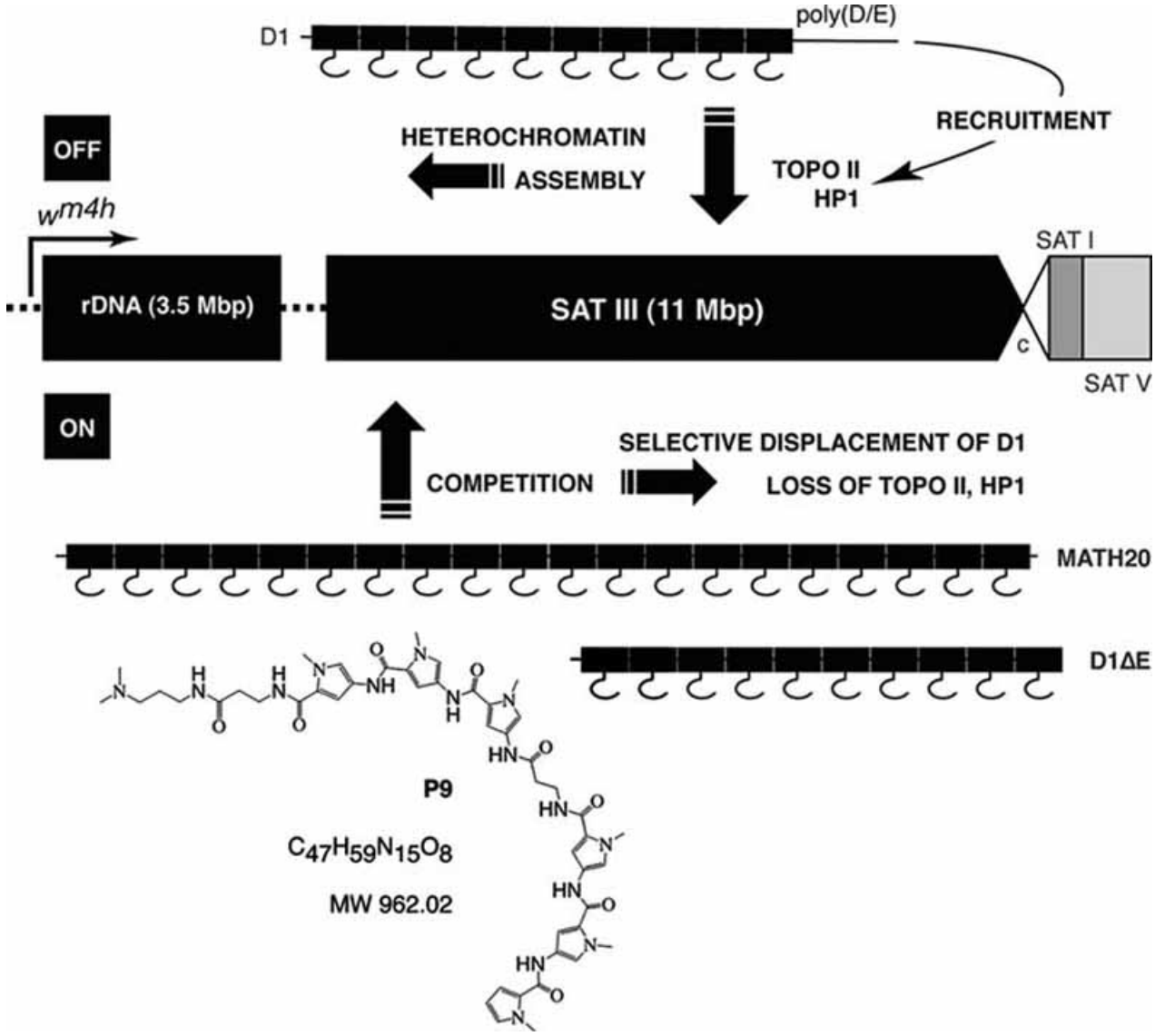


Fig. (3). A model for the initiation of heterochromatin assembly by the D1 AT-hook protein. Results from our laboratory and from previous work support a model of heterochromatin assembly promoted by specific binding of the D1 protein to SAT III repeats in the pericentric region of the *Drosophila* X chromosome. This is followed by recruitment, *via* the C-terminal domain of D1, of heterochromatin proteins such as HP1 and topoisomerase II, resulting in the assembly of a silenced structure, reflected by the variegated expression of the *white-mottled* allele (w^{m4h}), represented as the OFF state. The 3' end of the gene is approximately 20 kbp 5' of the 3.5-Mbp rDNA array in the *white-mottled* X-chromosome inversion [56]. Direct competition for binding to SAT III by the artificial MATH20 protein [50, 61], by the D1ΔE dominant negative mutant of D1 lacking its C-terminal domain [48], or by the synthetic P9 AT-hook mimic [52] results in complete D1 displacement from SAT III and loss of HP1 and topoisomerase II from the 11-Mbp satellite array. This protein redistribution is somehow transduced to the *white-mottled* allele whose expression is returned to near wild-type levels, represented by the ON state. An intriguing possibility is that rDNA repeats located between w^{m4h} and the SAT III array might be subjected to a mode of transcriptional regulation related to that seen for the variegating *white* gene.

For these reasons, we recently started experimenting with diamidines developed by D. W. Boykin and W. D. Wilson (reviewed in this issue) in the *Drosophila* system. Initial experiments focused on furamidine (DB75) and its derivative DB293 (see Fig. (**1**) of Wilson *et al*.,). As shown in Fig. (**4**), these molecules readily interact with *Drosophila* chromatin and give rise to staining patterns that correlate with their known DNA-binding properties. DB75 (blue) preferentially stains AT-rich regions, giving rise to bright spots that can be mapped largely to SAT I and SAT III repeats on mitotic chromosomes, which can be distinguished from SAT V repeats revealed by staining with the P31TR polyamide (red). Note that relatively A+T-poor chromosome arms are only weakly stained (if at all) by DB75. In contrast, DB293 (blue), which can bind more GC-rich sequences to some degree as a stacked dimer [71-73], stains chromosome arms somewhat more strongly, but still displays a strong preference for AT-rich satellite repeats. This is particularly

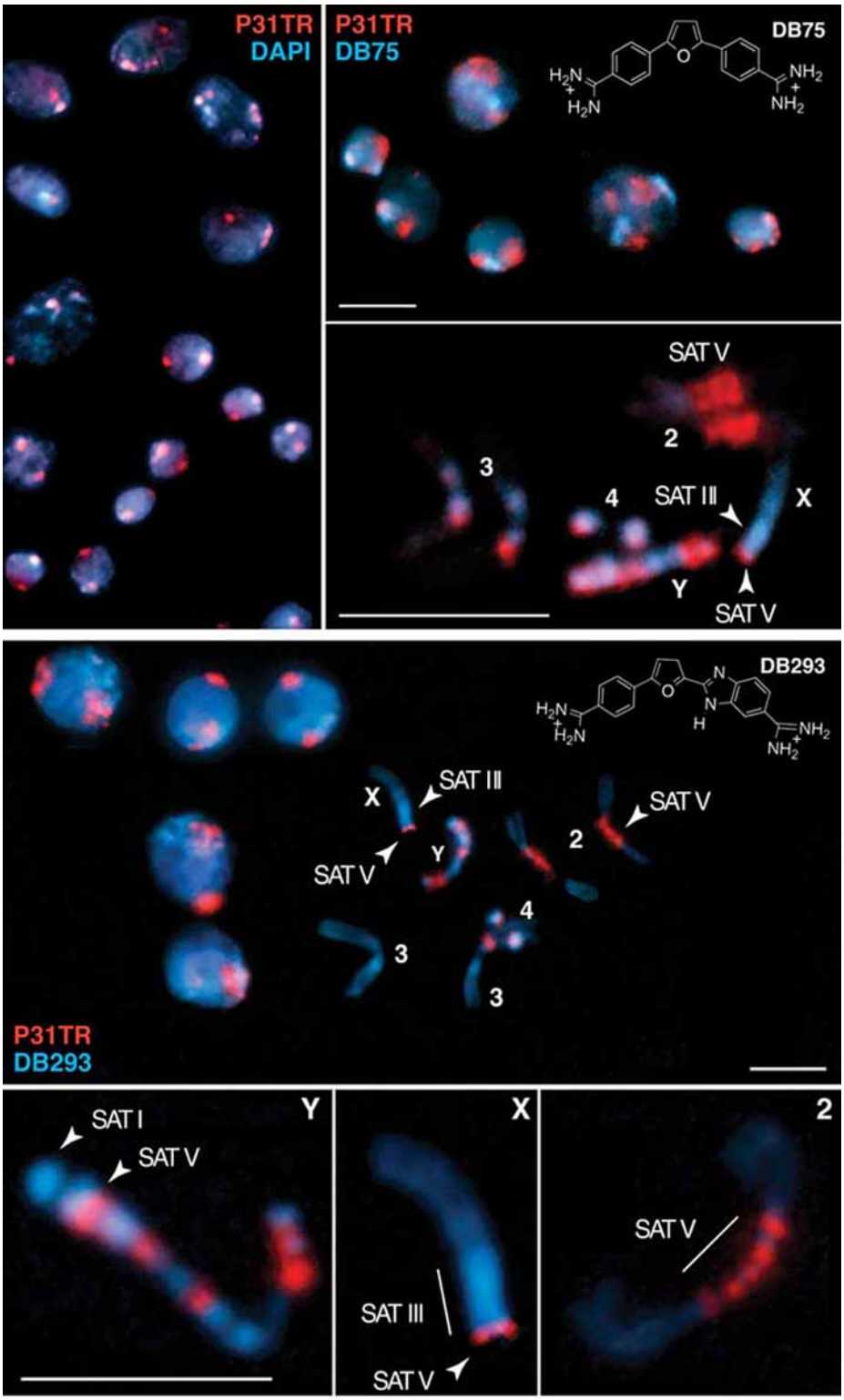


Fig. (4). Binding of diamidines to *Drosophila* nuclei and chromosomes. Fixed larval brains were stained with 2.5 μM DB75 or DB293 (blue) as shown in the photographs and counterstained with P31TR (red). A P31TR/DAPI stain is shown for comparison. Enlarged mitotic chromosomes shown in the inserts demonstrate that these diamidines preferentially recognize AT-rich satellite repeats, albeit at concentrations much higher than those used for detection by P9F (Fig. (**2**)). The SAT III repeats on the X chromosome and SAT V repeats on chromosome 2 and at the tip of the X are indicated. Note the strong staining of SAT III by the diamidines. Arrowheads indicate two of the several alternating blocks of SAT I and SAT V on the Y chromosome stained by DB293 and P31TR, respectively. Scale bars: 5 μm.

obvious in the enlarged chromosomes Y, X and 2 shown: the Y chromosome consists largely of alternating blocks of SAT I and SAT V (see Fig. (**1**)) and consequently shows an alternating red and blue staining pattern. Similarly, SAT III repeats are prominently stained on the X chromosome. Again, chromosome 2 is only weakly stained by DB293 and only shows its distinctive SAT V banded pattern revealed by P31TR staining.

As DB75 and DB293 preferentially stain SAT III repeats in *Drosophila* cells, we investigated these interactions in more detail. A major target of DB75 is trypanosome kinetoplast DNA (see W. D. Wilson *et al.*, this issue and [74]). These sequences exhibit strong intrinsic DNA curvature due to the presence of a "bending locus" of phased homopolymeric dA•dT tracts [75, 76], a characteristic shared by SAT III sequences. Fig. (**5**) shows a gel migration assay of a purified SAT III monomer compared to unrelated DNA fragments of differing sizes and base composition. Note in particular the small difference in size between the SAT III and F2 fragments (359 and 326 bp, respectively). F2 (60% A+T, mostly in the form of alternating dA-dT base pairs) is not curved and migrates normally while the SAT III fragment (which contains large phased homopolymeric dA•dT tracts) is strongly retarded, with an apparent mobility of about 620 bp. Incubation with increasing concentrations of distamycin increases the mobility of the SAT III fragment but has no effect on the migration of F1 (150 bp, 45% A+T), while the migration of F2 becomes slightly retarded as a result of distamycin binding.

We performed a similar experiment with DB75 and DB293, comparing the migration of the SAT III monomer to that of linearized pUC18 vector DNA (2.7 kb). Similar to the effect of distamycin, binding of either drug to SAT III increased its mobility, although DB293 had a stronger effect.

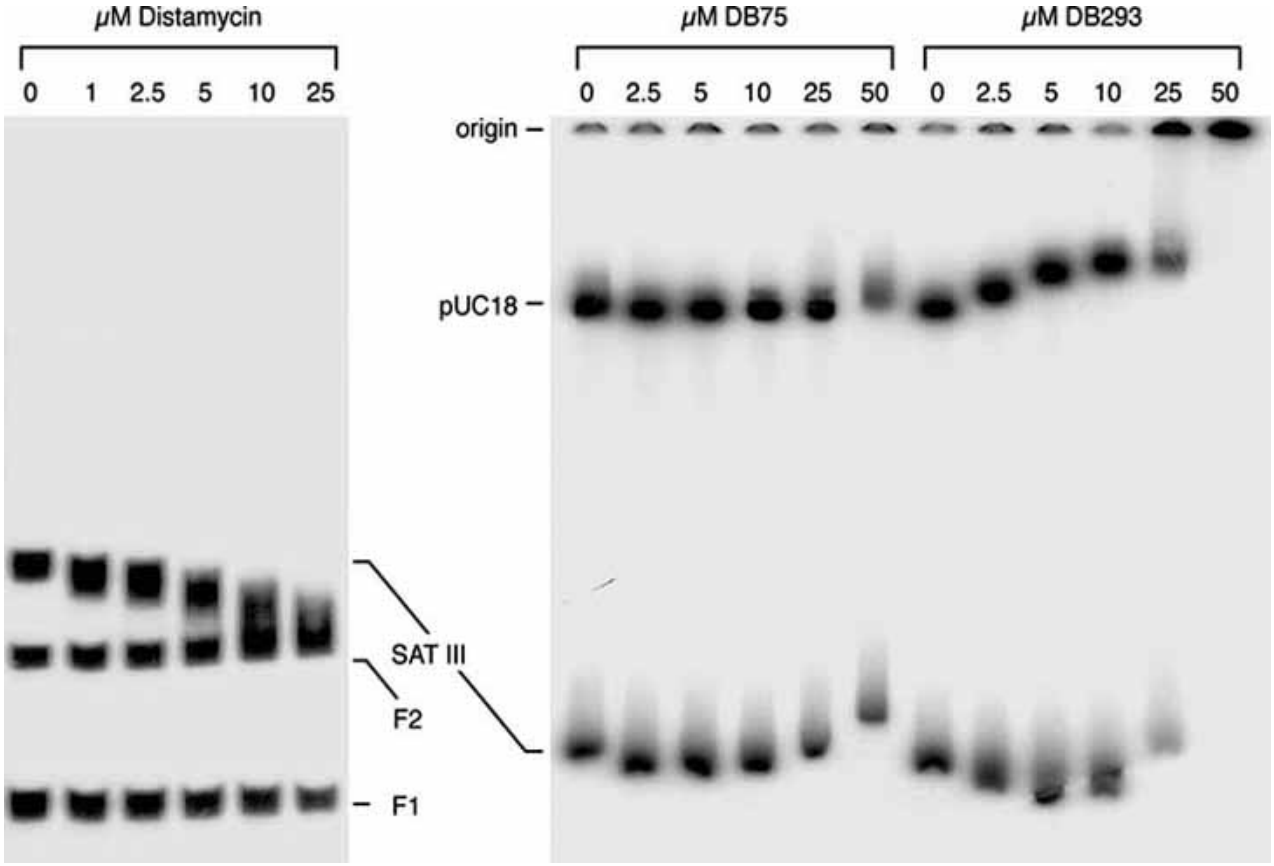


Fig. (5). Distamycin and diamidines induce conformational changes in SAT III DNA. Panel A: radiolabeled 359-bp SAT III monomer (> 65% A+T with large numbers of homopolymeric dA•dT tracts), F1 DNA (a 150-bp 45% A+T fragment) and F2 DNA (a 326-bp 60% A+T fragment containing mostly alternating dA•dT base pairs) were incubated with increasing concentrations of distamycin in the presence of a 250-fold excess of salmon sperm competitor DNA and electrophoresed on a 6% polyacrylamide gel at 4 ºC. Similar to kinetoplast DNA, SAT III DNA is intrinsically curved and migrates abnormally slowly, a behavior reversed in dose-dependent fashion by distamycin. The F1 and F2 fragments are not affected or retarded by drug binding, respectively. Panel B shows a similar experiment performed with DB75 and DB293, using pUC18 plasmid vector DNA as a control. Both diamidines initially increase the mobility of the SAT III fragment, which is preferentially bound by DB75. A more pronounced effect is seen in the presence of DB293, which also binds the vector fragment but retards its migration in dose-dependent fashion. High drug concentrations result in the retardation of both SAT III and vector fragments, or their retention in the well (origin). Drug-induced conformational changes and reduction of DNA curvature might also play a role in the interaction between diamidines and kinetoplast DNA, the presumed targets of the drugs in trypanosomes and Leishmania.

In addition, as expected from its DNA-binding characteristics [71, 73], DB293 bound the pUC18 fragment, as inferred from its increasing retardation, an effect opposite to that on SAT III, except at high drug concentrations where both fragments became retarded or remained in the well. These results suggest that intrinsically curved SAT III and kinetoplast DNA sequences may be similarly subjected to conformational changes induced by diamidines, as has been shown to be the case for distamycin [77]. Results to be published elsewhere show that the P9 oligopyrrole exhibits similar effects on SAT III DNA.

IDENTIFYING PROTEIN TARGETS AFFECTED BY DIAMIDINE BINDING

Binding of DB75 to SAT III heterochromatin (Fig. (**4**)) suggests that, like P9, it may possibly interfere with protein/DNA interactions in the pericentric region of the *Drosophila* X chromosome. We recently tested this hypothesis by immunolocalizing topoisomerase II in nuclei of permeabilized larval brains that had been incubated with DB75. We chose to focus on this protein as it is crucial for the maintenance of kinetoplast DNA in trypanosomes [78] and might be affected by binding of DB75. Panel A of Fig. (**6**) shows the topoisomerase II pattern (red) in untreated *Drosophila* tissues counterstained with DAPI (blue). Most of the topoisomerase II signal is nucleoplasmic and is excluded from a prominent DAPI-bright spot that corresponds to chromosome 4 enriched in SAT I and SAT V repeats (see Fig. (**1**)). Panel B shows an identical sample counterstained with DB75 instead of DAPI after immunolocalization, yielding a similar pattern and exclusion of the protein signal from a DB75-bright spot. We found that incubation of the tissues with DB75 before fixation had a dramatic effect on the nuclear localization of topoisomerase II. Treatment with 2.5 µM DB75 for 90 minutes resulted in a complete reversal of the protein staining pattern (panel C). In this case, topoisomerase II was found to redistribute to DB75-bright foci and was almost completely excluded from the nucleoplasm. Treatment with higher concentrations of the drug (10 µM, panel D) resulted in the apparent expulsion of topoisomerase II from both the nucleoplasm and DB75-bight foci and its retention as a bright ring at the nuclear periphery.

While the molecular mechanisms that underlie these striking effects need to be explored further, it is clear from these experiments that furamidine binding to AT-rich regions of the *Drosophila* genome can directly influence the localization of topoisomerase II in nuclei. Such a massive redistribution is expected to have profound biological effects. It is intriguing to speculate that the consequences of diamidine binding to kinetoplast DNA or, in bulk, to AT-rich genomes of organisms such as *P. carinii* or *P. falciparum*, might similarly affect, *via* mass-action phenomena, the cellular distribution of topoisomerase II and might explain the antiparasitic activity of drugs such as DB75. One possibility is that the effects of diamidines might be readily tested by assaying the response of *white-mottled* flies to these drugs. Because diamidine binding to AT-rich satellites promotes a massive recruitment of topoisomerase II to these sites (Fig. (**6**)), such a relocalization could affect expression of the variegating *white* gene. If it were possible to document such effects, *Drosophila* PEV might then constitute a useful and simple

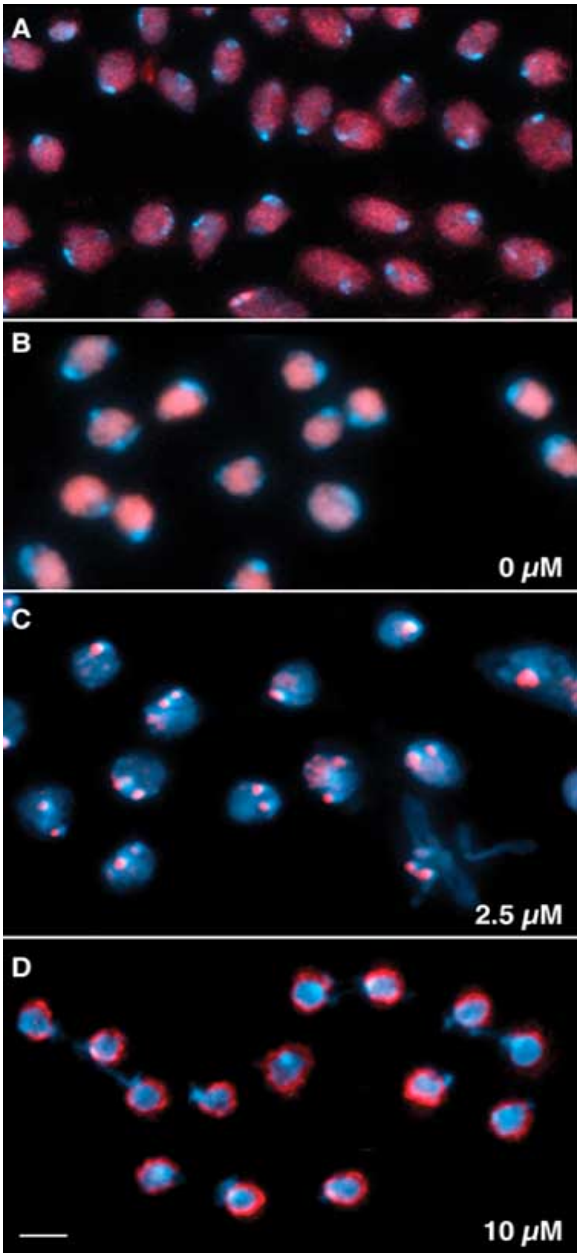


Fig. (6). Redistribution of topoisomerase II by DB75. Permeabilized *Drosophila* larval brains were incubated in the absence or presence of DB75, fixed and processed for topoisomerase II immunolocalization. Panel A shows control nuclei immunostained for topoisomerase II (red) and counterstained with DAPI (blue). Note the exclusion of the protein from DAPI-bright foci. Panel B shows similar control nuclei counterstained with DB75 (2 µM) instead of DAPI, yielding a pattern similar to that seen in panel A. Incubation with increasing concentrations of DB75 before fixation results in a massive redistribution of topoisomerase II. Treatment with 2.5 µM DB75 (panel C) completely reverses the localization of the protein from the nucleoplasm to the DB75-bright foci, detected here by the autofluorescence of the drug used in the incubation. Such a redistribution of the enzyme to drug-bound sequences may underlie the basic mechanism of the antiparasitic activity of DB75. Treatment with 10 µM DB75 results in a dramatic expulsion of topoisomerase II from the nucleoplasm to the nuclear periphery (panel D. Scale bars: 5 µm.

model system to assay the biological effects (and toxicity to the whole developing organism) of therapeutically important molecules that selectively target AT-rich sequences.

To test the feasibility of using such a system, we fed developing *white-mottled* larvae yeast paste containing 1mM DB351, a diamidine closely related to DB75. This choice was determined by the seemingly poor uptake of DB75 by larvae and a lack of observable effects in initial feeding experiments. Fig. (**7**) shows representative photographs of a collection of 30 control hatched *white-mottled* adult flies (top row) and of 30 flies recovered from DB351-fed *white-mottled* larvae (bottom row). Mutant *white*$^-$ and wild-type flies are shown at the far left for comparison. *white-mottled* PEV results in a range of eye-color phenotypes in individual flies, from a pale orange color with a few red spots to darker but mottled colors, of which three representative examples are shown. While this heterogeneity reappears in each succeeding generation, most *white-mottled* flies usually have the lighter eye-color phenotype (20 out of 30 flies in the example shown). In contrast, adult flies hatched from DB351-fed larvae had markedly redder eyes, the vast majority (24/30) with an eye color virtually indistinguishable from that of wild-type flies. Obviously, we cannot determine the actual intracellular drug concentrations that produce such effects but, as in the case of polyamides, it is remarkable that such crude drug feeding experiments yield measurable

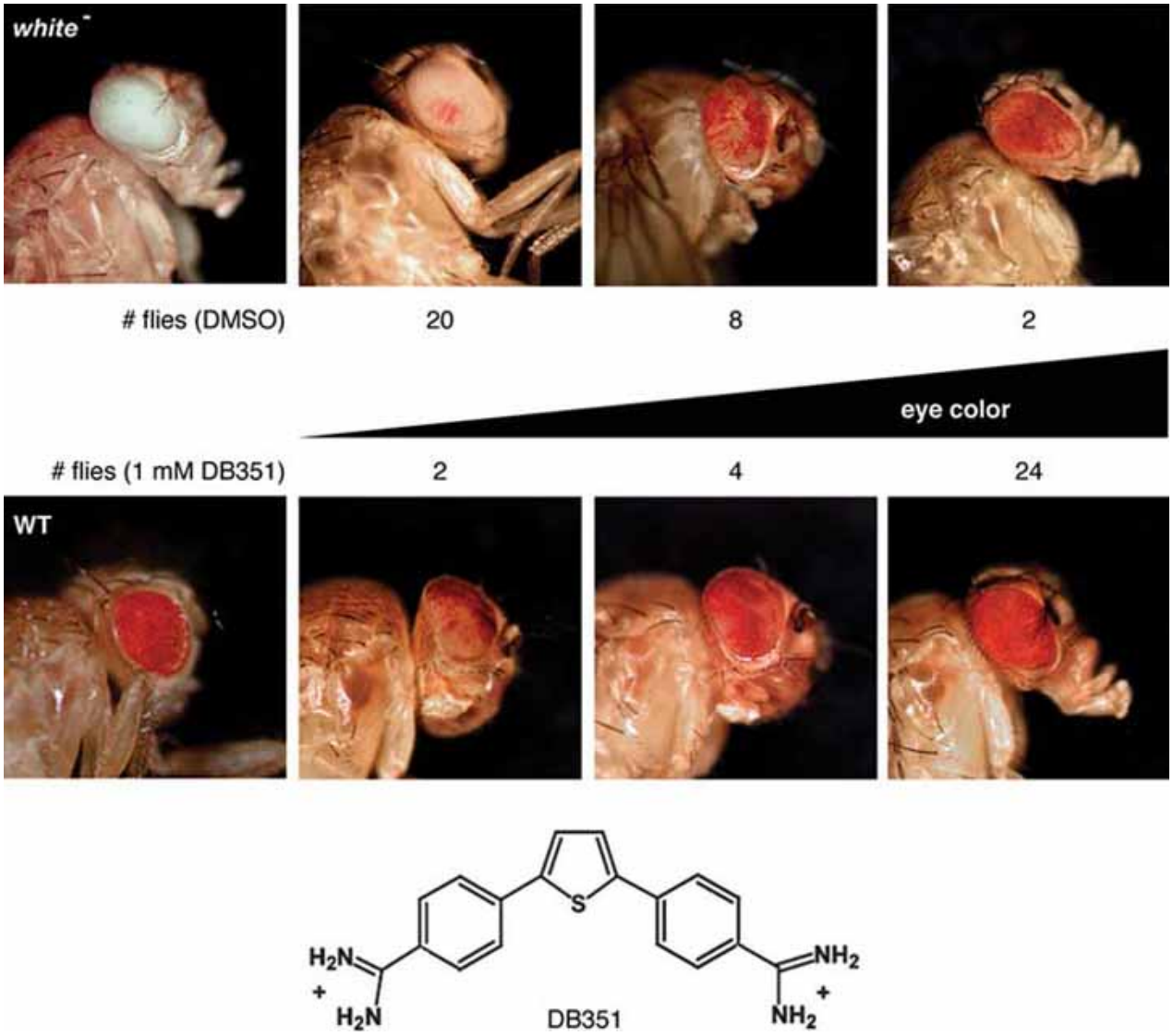


Fig. (7). Suppression of heterochromatin-mediated silencing by a diamidine. *white-mottled* larvae were fed yeast paste containing 1 mM DB351, closely related to DB75, and 1% DMSO. From the results presented in Fig. (**6**), drug feeding may affect topoisomerase II localization and, more generally, heterochromatin structure, during development. 30 adults that hatched from drug-fed larvae were scored for eye color and compared to 30 *white-mottled* adults that had been fed yeast paste supplemented with 1% DMSO alone. Photographs are shown for control and drug-fed larvae and illustrate the variegating phenotypes which can range from lightly colored to a darker eye color. The distribution of flies in each broad phenotypic category is indicated. DB351 induces a strong suppression of *white-mottled* PEV (80% of recovered adults with near-wild type eye color). Wild-type flies (WT) or flies mutated for the *white* gene (*white*$^-$) are shown for comparison in the left-most photographs.

effects at all, and not just dead larvae (drug-induced lethality was less than 10% at 1 mM DB351).

Thus, using a system as complex as that of a developing fly, it is possible to use the DB351 diamidine at concentrations that do not give rise to significant toxicity and lethality and observe a striking gain-of-function effect correlated with the relocalization of topoisomerase II to AT-rich sequences induced by DB75, a closely related molecule. These results further support a functional role for topoisomerase II in heterochromatin assembly and in heterochromatin-mediated silencing. They also suggest a possible mechanism for the anti-parasitic activity of diamidines, providing an intriguing link between AT-binding to curved DNA and topoisomerase II localization. Diamidine binding to trypanosome or Leishmania kinetoplast DNA results in the drug-induced destruction of the kinetoplast circles. The ensuing loss of mitochondrial function, RNA editing and rRNA expression encoded by the kinetoplast maxi- and mini-circles leads to cell death. It has been suggested (W. D. Wilson *et al*., this issue) that topoisomerase II-catalyzed reactions on kinetoplast DNA may be involved in the mechanism of action of these drugs. The massive relocalization of topoisomerase II to AT-rich diamidine-bound regions observed in the *Drosophila* system (Fig. (**6**)) lends support to such a hypothesis. It seems reasonable to suppose that diamidine-induced topoisomerase II redistribution might be accompanied by a significant increase in enzymatic activity at or near drug-bound sites, a hypothesis we are currently testing.

WHAT CAN WE LEARN FROM FLIES

While the biology of fruit flies is not a common topic for review in this journal, we believe that the *Drosophila* system described here can be successfully used to characterize some of the biological effects of small molecules that act on nucleic acids. In addition to polyamides and diamidines discussed here, these molecules also include the triplex-forming oligonucleotides (TFOs) whose use was pioneered by Claude Hélène and Heinz Moser ([79-82] and references therein). As a case in point, the poly-purine AAGAG/AAGAGAG repeat units of the *Drosophila* satellite V (Fig. (**1**)) are potentially a prime target for TFOs, a possibility currently under study in collaboration with C. Hélène's former laboratory. Most importantly, as already discussed, albeit tentatively, during the Jacques Monod conferences of 1991 and 1994, the conjugation of such sequence-specific molecules with DNA-damaging/DNA-cleaving agents [79] or topoisomerase poisons opens the way for anti-gene therapeutic approaches and much progress has been made in this area in recent years [83, 84]. A related example is that of the attachment of an Fe(II)EDTA cleaving moiety to oligopyrrole dimers, yielding molecules that are active and highly specific for DNA affinity cleavage [51].

Results reviewed above, some of which are not yet published, illustrate how synthetic polyamides and diamidines can be used as biological tools to probe the structure and function of non-coding sequences such as satellite repeats. Classical genetic or biological approaches do not lend themselves to a detailed analysis of the role of these highly repeated sequences. The "chemical" approaches discussed here demonstrate that these small molecules not only can serve as highly specific inhibitors of DNA/protein interactions *in vivo* but can also potently affect the localization of a nuclear protein such as topoisomerase II. The resulting biological effects (Fig. (**7**)) demonstrate the feasibility of using *Drosophila* satellite repeat-mediated phenomena as simple model systems to assay drug activity and toxicity. In addition, fluorescent polyamides used for the direct detection of their target satellite sequences are powerful cytological tools. Similarly, the fluorescence properties of diamidines potentially allow for the simultaneous assay of drug binding and biological effects. Taken together, these approaches can be successfully used to alter the protein composition and, presumably, the structure and function of heterochromatin in a highly selective way, with marked effects on transcriptional silencing pathways.

ACKNOWLEDGMENTS

This work was supported by the CNRS and by the Association pour la Recherche sur le Cancer (ARC grant No. 3319 to E. Käs). G. Susbielle and R. Blattes are supported by doctoral fellowships from the Ligue Nationale Contre le Cancer and the French Ministry of Research, respectively. V. Brevet was the recipient of an ARC post-doctoral fellowship. We thank the Centre de Biologie du Développement (CBD, UMR 5547) for *Drosophila* culture facilities and the IEFG (IFR 109) confocal microscopy platform and Brice Ronsin for help with wide-field microscopy and deconvolution. We are grateful to U.K. Laemmli for generous gifts of polyamides synthe-sized in his laboratory. Christian Bailly (Institut de Recherche Pierre Fabre) encouraged us to initiate an exciting collaboration between our laboratory and David Boykin and David Wilson, whom we warmly thank for providing us with an extensive panel of diamidines and for stimulating discussions.

REFERENCES

[1] Gniazdowski, M.; Denny, W.A.; Nelson, S.M.; Czyz, M. *Curr. Med. Chem.*, **2003**, *10*, 909.
[2] Dervan, P.B.; Edelson, B.S. *Curr. Opin. Struct. Biol.*, **2003**, *13*, 284.
[3] Wemmer, D.E.; Dervan, P.B. *Curr. Opin. Struct. Biol.*, **1997**, *7*, 355.
[4] Melander, C.; Burnett, R.; Gottesfeld, J.M. *J. Biotechnol.*, **2004**, *112*, 195.
[5] Zimmer, C. *Prog. Nucl. Acids Res. Mol. Biol.*, **1975**, *15*, 285.
[6] Baraldi, P.G.; Nunez Mdel, C.; Espinosa, A.; Romagnoli, R. *Curr. Top. Med. Chem.*, **2004**, *4*, 231.
[7] Hopwood, D.A. *Proc. R. Soc. Lond. B. Biol. Sci.*, **1988**, *235*, 121.
[8] Bentley, S.D.; Parkhill, J. *Annu. Rev. Genet.*, **2004**, *38*, 771.
[9] Wahnert, U.; Zimmer, O.; Luck, G.; Pitra, O. *Nucleic Acids Res.*, **1975**, *02*, 391.
[10] Simon, H.; Wittig, B.; Zimmer, C. *FEBS Lett.*, **1994**, *353*, 79.
[11] Boykin, D.W.; Kumar, A.; Hall, J.E.; Bender, B.C.; Tidwell, R.R. *Bioorg. Med. Chem. Lett.*, **1996**, *6*, 3017.
[12] Barrett, M.P.; Burchmore, R.J.; Stich, A.; Lazzari, J.O.; Frasch, A.C.; Cazzulo, J.J.; Krishna, S. *Lancet*, **2003**, *362*, 1469.
[13] Bouteille, B.; Oukem, O.; Bisser, S.; Dumas, M. *Fundam. Clin. Pharmacol.*, **2003**, *17*, 171.
[14] Tidwell, R.R.; Boykin, D.W. In *DNA and RNA binders: from small molecules to drugs*; Demeunynck, M.; Bailly, C.Wilson, W.D., Eds.; Wiley-VCH, **2003**; *2*, Chapter 16, pp. 414-460.
[15] Fairlamb, A.H. *Trends Parasitol.*, **2003**, *19*, 488.
[16] Marini, J.C.; Levene, S.D.; Crothers, D.M.; Englund, P.T. *Cold Spring Harb. Symp. Quant. Biol.*, **1983**, *47*, 279.

[17] Marini, J.C.; Effron, P.N.; Goodman, T.C.; Singleton, C.K.; Wells, R.D.; Wartell, R.M.; Englund, P.T. *J. Biol. Chem.*, **1984**, *259*, 8974.
[18] Drew, H.R.; Travers, A.A. *J. Mol. Biol.*, **1985**, *186*, 773.
[19] Drew, H.R. *J. Mol. Biol.*, **1984**, *176*, 535.
[20] Suck, D.; Oefner, C. *Nature*, **1986**, *321*, 620.
[21] Hart, C.M.; Laemmli, U.K. *Curr. Opin. Genet. Dev.*, **1998**, *8*, 519.
[22] Käs, E.; Izaurralde, E.; Laemmli, U.K. *J. Mol. Biol.*, **1989**, *210*, 587.
[23] Izaurralde, E.; Käs, E.; Laemmli, U.K. *J. Mol. Biol.*, **1989**, *210*, 573.
[24] Adachi, Y.; Käs, E.; Laemmli, U.K. *EMBO J.*, **1989**, *8*, 3997.
[25] Broggini, M.; Ponti, M.; Ottolenghi, S.; D'Incalci, M.; Mongelli, N.; Mantovani, R. *Nucleic Acids Res.*, **1989**, *17*, 1051.
[26] Dorn, A.; Affolter, M.; Muller, M.; Gehring, W.J.; Leupin, W. *EMBO J.*, **1992**, *11*, 279.
[27] Woynarowski, J.M.; McHugh, M.; Sigmund, R.D.; Beerman, T.A. *Mol. Pharmacol.*, **1989**, *35*, 177.
[28] Woynarowski, J.M.; Sigmund, R.D.; Beerman, T.A. *Biochemistry*, **1989**, *28*, 3850.
[29] Fesen, M.; Pommier, Y. *J. Biol. Chem.*, **1989**, *264*, 11354.
[30] Käs, E.; Laemmli, U.K. *EMBO J.*, **1992**, *11*, 705.
[31] Käs, E.; Poljak, L.; Adachi, Y.; Laemmli, U.K. *EMBO J.*, **1993**, *12*, 115.
[32] Negri, R.; Costanzo, G.; Buttinelli, M.; Venditti, S.; Di Mauro, E. *Biophys Chem*, **1994**, *50*, 169.
[33] Bell, A.; Kittler, L.; Lober, G.; Zimmer, C. *J. Mol. Recognit.*, **1997**, *10*, 245.
[34] Bartulewicz, D.; Bielawski, K.; Bielawska, A. *Arch. Pharm. (Weinheim)*, **2002**, *335*, 422.
[35] Reeves, R. *Environ. Health Perspect.*, **2000**, *108* Suppl. 5, 803.
[36] Reeves, R. *Gene*, **2001**, *277*, 63.
[37] Reeves, R.; Nissen, M.S. *J. Biol. Chem.*, **1990**, *265*, 8573.
[38] Aravind, L.; Landsman, D. *Nucleic Acids Res.*, **1998**, *26*, 4413.
[39] Zhao, K.; Käs, E.; Gonzalez, E.; Laemmli, U.K. *EMBO J.*, **1993**, *12*, 3237.
[40] Reeves, R.; Beckerbauer, L.M. *Prog. Cell Cycle Res.*, **2003**, *5*, 279.
[41] Evans, A.; Lennard, T.W.; Davies, B.R. *J. Surg. Oncol.*, **2004**, *88*, 86.
[42] Farnet, C.M.; Bushman, F.D. *Cell*, **1997**, *88*, 483.
[43] Hindmarsh, P.; Leis, J. *Microbiol. Mol. Biol. Rev.*, **1999**, *63*, 836.
[44] Strauss, F.; Varshavsky, A. *Cell*, **1984**, *37*, 889.
[45] Levinger, L.; Varshavsky, A. *Cell*, **1982**, *28*, 375.
[46] Levinger, L.; Varshavsky, A. *Proc. Natl. Acad. Sci. USA*, **1982**, *79*, 7152.
[47] Lohe, A.R.; Hilliker, A.J.; Roberts, P.A. *Genetics*, **1993**, *134*, 1149.
[48] Aulner, N.; Monod, C.; Mandicourt, G.; Jullien, D.; Cuvier, O.; Sall, A.; Janssen, S.; Laemmli, U.K.; Käs, E. *Mol. Cell. Biol.*, **2002**, *22*, 1218.
[49] Ashley, C.T.; Pendleton, C.G.; Jennings, W.W.; Saxena, A.; Glover, C.V. *J. Biol. Chem.*, **1989**, *264*, 8394.
[50] Monod, C.; Aulner, N.; Cuvier, O.; Käs, E. *EMBO Rep.*, **2002**, *3*, 747.
[51] Janssen, S.; Durussel, T.; Laemmli, U.K. *Mol. Cell*, **2000**, *6*, 999.
[52] Janssen, S.; Cuvier, O.; Muller, M.; Laemmli, U.K. *Mol. Cell*, **2000**, *6*, 1013.
[53] Henikoff, S.; Vermaak, D. *Cell*, **2000**, *103*, 695.
[54] Eissenberg, J.C.; Elgin, S.C. *Curr. Opin. Genet. Dev.*, **2000**, *10*, 204.
[55] Elgin, S.C.; Grewal, S.I. *Curr. Biol.*, **2003**, *13*, 895.
[56] Tartof, K.D.; Hobbs, C.; Jones, M. *Cell*, **1984**, *37*, 869.
[57] Suto, R.K.; Edayathumangalam, R.S.; White, C.L.; Melander, C.; Gottesfeld, J.M.; Dervan, P.B.; Luger, K. *J. Mol. Biol.*, **2003**, *326*, 371.
[58] Gottesfeld, J.M.; Melander, C.; Suto, R.K.; Raviol, H.; Luger, K.; Dervan, P.B. *J. Mol. Biol.*, **2001**, *309*, 615.
[59] Maeshima, K.; Janssen, S.; Laemmli, U.K. *EMBO J.*, **2001**, *20*, 3218.
[60] Strick, R.; Laemmli, U.K. *Cell*, **1995**, *83*, 1137.
[61] Girard, F.; Bello, B.; Laemmli, U.K.; Gehring, W.J. *EMBO J.*, **1998**, *17*, 2079.
[62] Swedlow, J.R.; Sedat, J.W.; Agard, D.A. *Cell*, **1993**, *73*, 97.
[63] Marshall, W.F.; Straight, A.; Marko, J.F.; Swedlow, J.; Dernburg, A.; Belmont, A.; Murray, A.W.; Agard, D.A.; Sedat, J.W. *Curr. Biol.*, **1997**, *7*, 930.
[64] Adachi, Y.; Luke, M.; Laemmli, U.K. *Cell*, **1991**, *64*, 137.
[65] Maeshima, K.; Laemmli, U.K. *Dev. Cell*, **2003**, *4*, 467.
[66] Csink, A.K.; Henikoff, S. *Nature*, **1996**, *381*, 529.
[67] Dernburg, A.F.; Broman, K.W.; Fung, J.C.; Marshall, W.F.; Philips, J.; Agard, D.A.; Sedat, J.W. *Cell*, **1996**, *85*, 745.
[68] Talbert, P.B.; Henikoff, S. *Genetics*, **2000**, *154*, 259.
[69] Ginsburg, H.; Nissani, E.; Krugliak, M.; Williamson, D.H. *Mol. Biochem. Parasitol.*, **1993**, *58*, 7.
[70] Shlomai, J. *Curr. Mol. Med.*, **2004**, *4*, 623.
[71] Wang, L.; Bailly, C.; Kumar, A.; Ding, D.; Bajic, M.; Boykin, D.W.; Wilson, W.D. *Proc. Natl. Acad. Sci. USA*, **2000**, *97*, 12.
[72] Bailly, C.; Tardy, C.; Wang, L.; Armitage, B.; Hopkins, K.; Kumar, A.; Schuster, G.B.; Boykin, D.W.; Wilson, W.D. *Biochemistry*, **2001**, *40*, 9770.
[73] Wang, L.; Carrasco, C.; Kumar, A.; Stephens, C.E.; Bailly, C.; Boykin, D.W.; Wilson, W.D. *Biochemistry*, **2001**, *40*, 2511.
[74] De Souza, E.M.; Lansiaux, A.; Bailly, C.; Wilson, W.D.; Hu, Q.; Boykin, D.W.; Batista, M.M.; Araujo-Jorge, T.C.; Soeiro, M.N. *Biochem. Pharmacol.*, **2004**, *68*, 593.
[75] Ntambi, J.M.; Marini, J.C.; Bangs, J.D.; Hajduk, S.L.; Jimenez, H.E.; Kitchin, P.A.; Klein, V.A.; Ryan, K.A.; Englund, P.T. *Mol. Biochem. Parasitol.*, **1984**, *12*, 273.
[76] Wu, H.M.; Crothers, D.M. *Nature*, **1984**, *308*, 509.
[77] Radic, M.Z.; Lundgren, K.; Hamkalo, B.A. *Cell*, **1987**, *50*, 1101.
[78] Wang, Z.; Englund, P.T. *EMBO J.*, **2001**, *20*, 4674.
[79] Moser, H.E.; Dervan, P.B. *Science*, **1987**, *238*, 645.
[80] Le Doan, T.; Perrouault, L.; Praseuth, D.; Habhoub, N.; Decout, J.L.; Thuong, N.T.; Lhomme, J.; Hélène, C. *Nucleic Acids Res.*, **1987**, *15*, 7749.
[81] Strobel, S.A.; Dervan, P.B. *Nature*, **1991**, *350*, 172.
[82] Strobel, S.A.; Dervan, P.B. *Science*, **1990**, *249*, 73.
[83] Seidman, M.M.; Glazer, P.M. *J. Clin. Invest.*, **2003**, *112*, 487.
[84] Nagatsugi, F.; Sasaki, S. *Biol. Pharm. Bull.*, **2004**, *27*, 463.

4. *Conclusions et discussions*

Les résultats de cette étude nous ont permis de démontrer que tout comme la distamycine, les diamidines provoquent *in vitro* un changement conformationnel des séquences d'ADN auxquelles elles se lient. La mobilité électrophorétique de ces séquences est affectée par la liaison de ces composés. Les séquences SAT III de la drosophile présentent une courbure naturelle du fait de l'enrichissement en paires de bases dA•dT. La fixation de molécules telles que la distamycine ou les diamidines génère un changement de cette conformation. Ces molécules s'insèrent dans le petit sillon de l'ADN et entraînent une correction de la structure se traduisant alors par une mobilité électrophorétique accélérée. En effet, la fixation de ces molécules à faible poids moléculaire n'entraîne pas un retard de la migration mais une accélération.

Le comportement de deux molécules, les DB75 et DB293, a été testé face à des séquences témoins, le plasmide pUC18, et une séquence correspondant à une répétition de la séquence SAT III. La différence de comportement entre les deux diamidines reflète la spécificité de reconnaissance de ces deux composés. *In vitro*, le DB293 semble montrer une plus grande affinité que le DB75 pour les séquences SAT III. Ces résultats sont à modérer car le DB293 semble avoir une affinité pour l' ADN supérieure au DB75, et ce quelque soit la séquence nucléotidique. Pour autant, les deux diamidines sont capables de se lier à ces séquences et de changer leur courbure. Ce changement topologique augmente la vitesse de migration des complexes ADN-diamidines, la taille de ces composés étant trop faible pour entraîner un retard à de faibles concentrations.

Les études menées *in vivo* on montré qu'au contraire le DB75 avait une très grande affinité pour les séquences riches en paires de bases dA•dT, le DB293 présentant quant à elle une affinité de moindre mesure dans ces conditions.

Cette observation nous a donc tout naturellement amené à rechercher les composants structurels et fonctionnels de l'hétérochromatine affectés par une telle interaction. Une hypothèse a été envisagée puis testée. La cible majeure des diamidines étant à la fois des séquences riches en paires de bases dA•dT mais aussi les kinétoplastes, une des protéines affectée pouvait être la topoïsomérase II.

J'ai donc réalisé des expériences d'immunolocalisation de la topoïsomérase II en présence ou absence de diamidines. En accord avec notre hypothèse, la localisation de la topoïsomérase II est affectée par les traitements : la protéine est relocalisée massivement sur les séquences SAT I et SAT III, mais aussi sur les chromosomes 4 normalement dépourvus de marquage pour la topoïsomérase II. Cette relocalisation pourrait correspondre à un démasquage de sites de fixation pour cette protéine suite à la fixation des diamidines et ouverture de la chromatine. Ces observations recoupent les résultats obtenus avec le prototype des diamidines, la pentamidine actuellement utilisés contre la maladie du sommeil et qui agit en dérégulant l'activité de la topoïsomérase II (Shapiro and Englund, 1990).

Ces mêmes expériences nous aussi montré que la localisation de la topoïsomérase II était affectée en interphase. Le marquage nucléoplasmique est beaucoup moins intense et la topoïsomérase II est relocalisée massivement sur le point brillant de DAPI qui correspond principalement aux séquences SAT III. Ceci conforte donc l'hypothèse d'ouverture de la chromatine par les diamidines.

Des traitements *in vivo* sur des mouches ont ensuite été réalisés et ont démontré le rôle à la fois structural et fonctionnel de la topoïsomérase II au sein des séquences SAT III. La relocalisation massive de la topoïsomérase II au niveau de ces séquences suite à des expériences d'ingestion de DB351 par des larves de la lignée *white-mottled 4h* résulte en un phénotype de suppression de la variégation d'effet de position du gène *white*. Cette observation pourrait traduire une séquestration massive de la topoïsomérase II sur les séquences SAT III ce qui limiterait sa disponibilité envers d'autres séquences. Cette diminution de « dose de protéine disponible » est à relier avec les expériences de suppression de la variégation classiques, dans lesquelles, une copie de gène est enlevée ou inactivée, diminuant ainsi le stock de protéine disponible dans une cellule.

Ces résultats illustrent le fait que des composés synthétiques tels que les diamidines peuvent être utilisés en tant qu'outils biologiques permettant l'étude de la structure et de la fonction de séquences non codantes telles que les séquences satellites. Une approche chimique de cette étude montre que ces composés sont non seulement spécifiques pour la liaison à une séquence mais aussi qu'ils peuvent affecter la localisation de composants protéiques tels que la topoïsomérase II. Ils permettant également d'envisager un mécanisme plausible de leur activité anti-parasitaire, proposé dans (Susbielle et al., 2005), et récemment validé (voir chapitre III de cette même section).

Une combinaison de ces approches pourrait permettre de caractériser le rôle et la fonction de l'hétérochromatine. De même, l'utilisation de ces molécules s'avère relativement prometteuse en terme de traitements anti-cancéreux. La topoïsomérase II est à l'heure actuelle une cible majeure de ces thérapies et la combinaison de poisons directs de la topoïsomérase II et des ces composés ligands du petit sillon de séquences non codantes, en tant qu'adjuvants ou inhibiteurs, pourrait ouvrir une nouvel axe de recherche.

II. ***Du rôle de protéines D1, HP1 et topoïsomérase II dans la structure hétérochromatinienne : « Displacement of D1, HP1 and topoisomerase II from satellite heterochromatin by a specific polyamide »***

1. Introduction

En parallèle à l'étude précédente, j'ai réalisé des expériences utilisant l'oligopyrrole P9. Ces expériences avaient pour but de déterminer le rôle de la protéine D1 dans l'initiation de l'assemblage de la structure hétérochromatinienne au niveau des séquences SAT III de la drosophile.

a. Les polyamides P9 et P31

Le polyamide P9 a tout d'abord été synthétisé afin d'étudier la condensation des chromosomes. Il reconnait spécifiquement et avec une très forte affinité les séquences SAT I et SAT III de la drosophile, riches en paires de bases dA•dT (Janssen et al., 2000b). Le polyamide P31, quant à lui, interagit spécifiquement avec les séquences SAT V, riches en paires de bases dA•dG (Janssen et al., 2000b).

L'étude de la structure de l'hétérochromatine grâce à l'utilisation de petites molécules chimiques présentes un intérêt certain, comme nous l'avons déjà démontré dans les travaux précédents. De plus, ils permettent de pallier au manque d'outils classiques dirigés contre ces séquences satellites, tels que la génétique, puisque ces derniers ne peuvent pas s'appliquer aux séquences fortement répétées et intrinsèquement hétérogènes et variables.

Les premières études avec les polyamides P9 et P31 ont permis de démontrer leur efficacité de fixation mais aussi leur grande spécificité de reconnaissance (Janssen et al., 2000b).

Le polyamide P31 est capable d'induire des transformations homéotiques bien caractérisées dans un fond génétique de type *brown-dominant*. Ces transformations miment la

mutation du gène codant pour le facteur GAGA (GAF) (Janssen et al., 2000a). L'effet du polyamide P31 correspond à une ouverture de la chromatine, permettant le démasquage de sites de liaison pour la protéine GAF au niveau des séquences SAT V. Ces sites, dans le contexte hétérochromatinien classique, ne sont pas accessibles à la protéine et leur reconnaissance par la protéine GAF va entraîner une diminution de la quantité de protéine disponible pour son activité normale.

L'ingestion de P9 par des larves entraîne par contre un phénotype de gain de fonction. Les larves de la lignée *white-mottled* nourries avec ce polyamide présente une forte suppression de la variégation d'effet de position, traduisant un mécanisme de rétraction ou d' « ouverture » de la structure hétérochromatinienne (Janssen et al., 2000a). Ce phénomène permet au gène *white* de s'exprimer de manière quasiment normale.

b. La protéine chromosomale non-histone D1

La protéine D1 de drosophile est essentielle et exprimée de manière ubiquitaire au cours du développement (Aulner et al., 2002). Elle possède dix motifs AT Hook ce qui la classe dans la famille des protéines MATH. Ces protéines, dont l'archétype est la protéine humaine HMG-A, sont capables d'interagir avec le petit sillon de l'ADN riche en paires de bases dA•dT. Il a été démontré que la protéine D1 se liait au petit sillon de l'ADN des séquences SAT I et SAT III de la drosophile (Levinger and Varshavsky, 1982) (Reeves and Nissen, 1990).

Les premières études menées dans le laboratoire ont été faites grâce à la protéine artificielle MATH 20 qui possède vingt motifs AT Hook (Monod et al., 2002). Nos résultats ont démontré que la protéine MATH 20 se liait elle aussi aux séquences satellites riches en paires de bases dA•dT de la drosophile, tout comme la protéine D1. Pour autant, la surexpression de ces deux protéines, D1 et MATH 20, n'a pas le même effet sur la variégation d'effet de position du gène *white*. La surexpression de la protéine D1 entraîne une augmentation de la variégation (Aulner et al., 2002) alors qu'une surexpression de la protéine MATH 20 supprime ce phénotype (Girard et al., 1998). Ces effets contradictoires sont à rapprocher d'un autre résultat démontrant qu'une forme dominante négative de la protéine D1 supprime la variégation d'effet de position. Cette protéine, appelée D1ΔE, correspond à la

protéine D1 pleine taille délétée de son domaine C-terminal acide. Ces différents résultats suggèrent que le domaine acide C-terminal de la protéine D1 est impliqué dans les interactions avec d'autres protéines, elles-mêmes requises pour l'assemblage d'une structure de type hétérochromatinienne.

2. *Objectifs de l'étude*

Il a été proposé que la suppression de la variégation d'effet de position par le polyamide P9 passe par une perturbation de la structure de l'hétérochromatine présente au niveau des séquences SAT III, tant au niveau des protéines qui s'y fixent que de sa structure intrinsèque. J'ai donc réalisé des expériences à la fois *in vivo* et *in vitro* permettant d'appréhender la composition biochimique des SAT III, avant et après traitement par le polyamide P9.

A la vue des résultats précédents obtenus dans l'équipe, j'ai testé l'hypothèse selon laquelle la protéine D1 interviendrait dans le recrutement des protéines impliquées dans la formation de la structure hétérochromatienne, telles que la protéine HP1. De nombreuses études tendent à démontrer que la protéine HP1 reconnaît la structure hétérochromatinienne non pas en se liant à la composante ADN mais plutôt à la lysine 9 de l'histone H3 préalablement méthylée. Cette reconnaissance est à la base du mécanisme d'expansion de la structure hétérochromatinienne.

Pour autant, un mécanisme de reconnaissance de l'ADN par la protéine HP1 n'est pas exclu. La protéine HP1 a donc besoin d'être guidée vers les séquences à hétérochromatiniser et cette étape d'initiation de l'assemblage pourrait impliquer la protéine D1. Cette hypothèse a donc été testée au cours de mes travaux, grâce à l'utilisation du polyamide P9 en tant que compétiteur de la protéine D1 pour la liaison aux séquences satellites. Dans un premier temps, des expériences *in vitro* ont été menées afin de caractériser les sites de liaison de ces deux composants au niveau des séquences SAT III de la drosophile. Dans un deuxième temps, des expériences d'immunolocalisation ont permis de démontrer les effets du polyamide P9 sur la localisation des protéines D1 et HP1. De même, le polyamide P31 a été utilisé afin de valider notre méthode de déplacement sélectif par compétition entre des polyamides et des protéines endogènes pour la reconnaissance et la fixation de séquences cibles.

La protéine D1 pourrait aussi intervenir dans le recrutement de composants structuraux de l'hétérochromatine péricentromérique de la drosophile, tels que la topoïsomérase II. L'accumulation de la topoïsomérase II sur ces séquences a été documentée (Käs and Laemmli, 1992; Swedlow et al., 1993), mais le rôle fonctionnel de cette interaction reste encore à définir. Des résultats obtenus au sein de l'équipe suggèrent que la topoïsomérase II pourrait intervenir dans la formation de la structure hétérochromatinienne. J'ai testé cette hypothèse grâce à des expériences de traitement par l'oligopyrrole P9 suivies d'immunolocalisations.

3. Résultats

Publication parue dans « The EMBO journal » en 2006.

The EMBO Journal (2006), 1–12 | © 2006 European Molecular Biology Organization | All Rights Reserved 0261-4189/06
www.embojournal.org

THE EMBO JOURNAL

Displacement of D1, HP1 and topoisomerase II from satellite heterochromatin by a specific polyamide

Roxane Blattes[1,5], Caroline Monod[1,5], Guillaume Susbielle[1], Olivier Cuvier[2], Jian-hong Wu[3], Tao-shih Hsieh[3], Ulrich K Laemmli[4] and Emmanuel Käs[1,*]

[1]Laboratoire de Biologie Moléculaire Eucaryote, UMR 5099 CNRS-Université Paul Sabatier, Toulouse Cedex, France, [2]Institut de Génétique Humaine, CNRS UPR 1142, Montpellier Cedex, France, [3]Department of Biochemistry, Nanaline H Duke Building, Duke University Medical Center, Durham, NC, USA and [4]Département de Biologie Moléculaire, Université de Genève, Sciences II, Geneva, Switzerland

The functions of DNA satellites of centric heterochromatin are difficult to assess with classical molecular biology tools. Using a chemical approach, we demonstrate that synthetic polyamides that specifically target AT-rich satellite repeats of *Drosophila melanogaster* can be used to study the function of these sequences. The P9 polyamide, which binds the X-chromosome 1.688 g/cm^3 satellite III (SAT III), displaces the D1 protein. This displacement in turn results in a selective loss of HP1 and topoisomerase II from SAT III, while these proteins remain bound to the adjacent rDNA repeats and to other regions not targeted by P9. Conversely, targeting of $(AAGAG)_n$ satellite V repeats by the P31 polyamide results in the displacement of HP1 from these sequences, indicating that HP1 interactions with chromatin are sensitive to DNA-binding ligands. P9 fed to larvae suppresses the position-effect variegation phenotype of *white-mottled* adult flies. We propose that this effect is due to displacement of the heterochromatin proteins D1, HP1 and topoisomerase II from SAT III, hence resulting in stochastic chromatin opening and desilencing of the nearby *white* gene.

The EMBO Journal advance online publication, 4 May 2006; doi:10.1038/sj.emboj.7601125
Subject Categories: chromatin & transcription
Keywords: D1; heterochromatin; HP1; topoisomerase II; 1.688 g/cm^3 satellite

Introduction

Polyamides are short polymers composed of aromatic pyrrole and imidazole amino acids that recognize specific DNA sequences with remarkable affinities (Dervan and Edelson, 2003). The use of polyamides provides an alternative approach to the study of DNA satellites (Janssen *et al*, 2000b), which, unlike regulatory sequences of genes, are difficult to study with classical molecular biology approaches.

It was previously demonstrated that satellite-specific polyamides fed to developing *Drosophila melanogaster* can induce either gain or loss-of-function phenotypes in adult flies. One of these polyamides, P31, binds with great specificity the $(AAGAG)_n$ repeats of the *Drosophila* satellite V (SAT V) (Janssen *et al*, 2000b). P31 was found to induce well-defined homeotic transformations in a *brown-dominant* genetic background that are akin to mutations of the gene encoding GAGA factor (GAF) (Janssen *et al*, 2000a). The effect of P31 was proposed to be mediated by a redistribution of GAF from its euchromatin-binding sites, where this protein acts as a positive transcription activator to SAT V. The homeotic phenotype is then explained by a reduced availability of GAF for gene expression. In contrast, a gain-of-function phenotype was induced by the P9 polyamide. P9 specifically targets the AT-rich satellites SAT I $(AATAT)_n$ and, presumably, SAT III (density 1.688 g/cm^3), which is composed of 359-base pairs (bp) repeats. It was observed that feeding P9 to *white-mottled* (w^{m4}) larvae suppressed the position-effect variegation (PEV) phenotype in adult flies. The w^{m4} fly strain contains a chromosomal inversion that places the *white* gene near the centric heterochromatin of the X chromosome, which harbors the rDNA repeats and the enormous 11-Mbp SAT III array (Lohe *et al*, 1993). PEV is thought to arise through heterochromatin-mediated silencing of the juxtaposed *white* gene (for a review, see Huisinga *et al*, 2006). It was proposed that suppression of PEV by P9 (increased expression of the *white* gene) acts through chromatin opening of SAT III and a subsequent reduction of silencing effects on the nearby *white* gene. This paper aims to dissect the biological effect of P9 by examining the biochemical composition of SAT III using a combination of *in vivo* and *in vitro* experiments.

D1 is an essential protein ubiquitously expressed throughout *Drosophila* development (Aulner *et al*, 2002) containing 10 copies of the AT hook motif, a DNA-binding domain that interacts with the minor groove of AT-rich DNA (Reeves and Nissen, 1990). D1 is associated with SAT I and SAT III repeats. Related to D1 is the artificial MATH20 protein that contains 20 AT hook motifs and, as D1, is expected to bind SAT I and SAT III. Overexpression of D1 enhances w^{m4} PEV (Aulner *et al*, 2002), while overexpression of MATH20 suppresses it (Girard *et al*, 1998). To explain these opposing effects, it was suggested that D1 increases, while MATH20 reduces, silencing effects mediated by SAT III (Monod *et al*, 2002). In contrast, overexpression of a D1 transgene carrying a deletion of its C-terminal domain (D1ΔE) suppresses w^{m4} PEV, an effect similar to that of MATH20. Because D1ΔE behaves as a dominant negative mutation of D1, we proposed that the C-terminal domain of D1 might serve to recruit proteins required for heterochromatin assembly to SAT I and SAT III.

*Corresponding author. LBME, UMR5099, IBCG, 118 route de Narbonne, 31062 Toulouse Cedex 9, France. Tel.: +33 561 335959; Fax: +33 561 335886; E-mail: kas@ibcg.biotoul.fr
[5]These authors contributed equally to this work

Received: 21 October 2005; accepted: 11 April 2006

©2006 European Molecular Biology Organization

Heterochromatin protein 1 (HP1) is a well-characterized structural and functional component of heterochromatin. Binding of HP1 to heterochromatin occurs through the specific recognition of histone H3 methylated at the lysine 9 position by the chromo domain of HP1. Mutations in the gene encoding HP1 or in the genes that encode the histone deacetylase and methyl transferase activities required for the creation of an HP1-binding site suppress w^{m4} PEV (for a review, see Huisinga *et al*, 2006). The interplay between these proteins provides a rationale for the assembly and propagation of heterochromatin, which might involve other proteins as well.

Topoisomerase II (topo II) might conceivably be implicated in w^{m4} PEV as it is prominently associated with centric heterochromatin and SAT III repeats, as demonstrated by *in vivo* cleavage (Käs and Laemmli, 1992) and localization studies (Swedlow *et al*, 1993). Topo II is a DNA topological handler that relaxes positively and negatively supercoiled DNA and mediates DNA strand passage during catenation/decatenation reactions. Previous studies established that topo II is required for the segregation and assembly of mitotic chromosomes (Adachi *et al*, 1991). Hence, this enzyme may perhaps play roles in heterochromatin and chromosome structure, in addition to its enzymatic function.

One possibility is that D1 might help to recruit HP1 and topo II to AT-rich satellites and we explore here the mechanism whereby P9 exerts suppression of the w^{m4} phenotype. Using an immunofluorescence approach, we show that P9 mediates the displacement of D1 from SAT I and SAT III *in vivo* and *in vitro*. Moreover, P9 also specifically displaces HP1 and topo II from SAT III, while these proteins remain bound to other genomic sites that are not targeted by P9. Conversely, targeting SAT V repeats with the P31 polyamide results in the selective displacement of HP1 from these sites, supporting the hypothesis that HP1 association with chromatin involves a direct DNA-binding component. We propose that the suppression of the w^{m4} phenotype by P9 is mediated by displacement of D1, HP1 and topo II from SAT III. To refine our analysis, we studied the effect of topo II inhibitors on w^{m4} flies and we show that feeding VM26, a topo II poison, to w^{m4} flies also results in suppression of *white-mottled* PEV, supporting a structural and/or functional role for this enzyme in heterochromatin-mediated silencing.

Results

Specificity of DNA-binding polyamides for different Drosophila satellites

Previous studies established that the satellite-specific polyamides, P9 and P31, fed to developing flies induce gain or loss-of-function phenotypes, respectively. The binding properties and the biological effects of these polyamides are recapitulated in Supplementary Table I and discussed below.

P31: This compound is known to bind with impressive specificity the $(AAGAG)_n$ repeats of satellite V (SAT V) as determined by footprinting and fluorescence microscopy. P31TR, a derivative labeled with Texas red, was demonstrated to bind to the chromocenter of polytene chromosomes and to the 1.7-Mbp $(AAGAG)_n$ repeats inserted in the *brown-dominant* allele of the *brown* gene (Janssen *et al*, 2000b). This analysis was extended here to high resolution using mitotic chromosomes obtained from larval neuroblasts (Figure 1A).

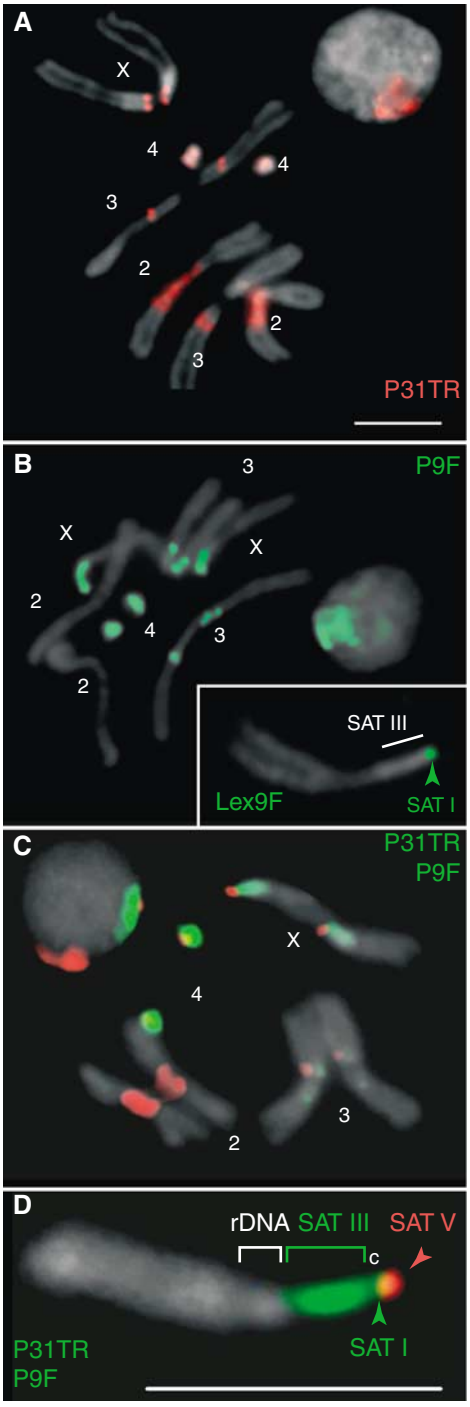


Figure 1 Localization of satellites in diploid cells using fluorescent polyamides. (A–D) Brains dissected from female third-instar larvae were fixed and stained using fluorescent polyamides and counterstained with DAPI (gray). The properties of the drugs are described in Supplementary Table I. P31TR (**A**) detects satellite V, P9F (**B, D**) satellites III and I, Lex9F (inset in B) satellite I. SAT I, III and V maps are diagrammed in Figure 4G below. A sample double-stained with P9F and P31TR is shown in (**C**). The localization of satellite and rDNA repeats is shown for a magnified X chromosome double-stained with P9F and P31TR (D; c: centromere). Scale bars: 5 μm.

This panel localizes SAT V repeats to pericentric regions on chromosomes 2, 3 and 4 and to the very tip of the X chromosome. Figure 1A also includes a nucleus, where the signal is well confined to the SAT V repeats congregated in the chromocenter.

©2006 European Molecular Biology Organization

P9: This oligopyrrole is specific for AT-rich satellites and, when fed to developing w^{m4} flies, suppresses the *white-mottled* PEV phenotype (Janssen *et al*, 2000a). The biological effect of P9 is most likely exerted through binding to the alpha-like satellite III (SAT III), the major satellite juxtaposed to the inverted *white* gene of w^{m4} flies. This notion was examined further here by microscopy of spread mitotic chromosomes stained with the fluorescein-labeled derivative P9F. Figure 1B shows that P9F strongly stains the large (11 Mbp) SAT III abutting the centromere of the X chromosome (see also the enlarged X chromosome in panel D). P9F also detects SAT I $(AATAT)_n$ repeats throughout chromosome 4 and related repeats on either side of the centromere of chromosome 3, as well as a short array of SAT I repeats near SAT V (red) on the X chromosomes (Figure 1D). The order of satellites from the tip of the X chromosome to the rDNA locus is SAT V, SAT I and SAT III, as shown in panel D. These assignments are based on staining with Lex9F (inset Figure 1B) and co-staining with P31TR and P9F (Figure 1C). Lex9F (see Supplementary Table I) is a dimeric oligopyrrole that prefers interaction with long (>W9) AT-rich runs. Hence, Lex9F serves to distinguish between the $(AATAT)_n$ repeats of SAT I and the shorter AT-runs of SAT III, which it does not bind (inset panel B). These results are in excellent agreement with previous satellite mapping data (Lohe *et al*, 1993) and are summarized in Figure 4G below.

The main aim of this report is to elucidate the molecular basis of the gain-of-function phenotype mediated by P9. We proposed that P9 reduces the heterochromatin state of SAT III, presumably as a result of protein displacement. This proposal is explored here, using an experimental approach where polyamides serve two functions: (a) fluorescent derivatives are used to localize the DNA satellites in fixed material (Figure 1) in combination with immunostaining studies and (b) to act as 'drugs' *in vivo* and *in vitro* so as to affect the biochemical composition of heterochromatin.

D1 binds SAT I and SAT III: P9-mediated displacement of D1 in vitro

The D1 protein is associated with SAT I and SAT III repeats and overexpression of D1 enhances PEV of w^{m4} flies. Figure 2A shows that the immunosignal of D1 (red) colocalizes with the pattern of P9F staining (green, panel B and merged in panel C). Hence, D1 binds SAT I and SAT III repeats in interphase and metaphase chromatin.

We explored whether the biological effect of P9 on PEV might arise by displacement of D1. This notion was first studied *in vitro* with gel retardation/competition experiments. Figure 3A shows that binding of purified D1 to a 359-bp SAT III monomer generates a ladder of retarded complexes reflecting increased binding of D1 (Figure 3A, lanes 1–4). In contrast to D1, addition of P9 (MW = 962) resulted in faster migrating complexes (lanes 5–9), due to unbending of the curved SAT III fragment, although retardation was seen at the highest P9 concentration tested (500 nM, lane 10).

Figure 3B establishes that P9 dissociates D1/SAT III complexes *in vitro*. Examination of panel B shows that the D1/SAT III ladder diminishes in a concentration-dependent manner upon addition of P9 (compare lanes 5–7, 8–10 and 11–13 with lanes 2–4). This is accompanied by an increase in the mobility of the free DNA and protein-bound fragments indicative of P9 binding. This displacement reaction was independent of the order in which P9 and D1 were added (data not shown). Taken together, these results demonstrate that the P9 oligopyrrole binds SAT III repeats *in vitro* and displaces the D1 protein bound to these sequences.

DNAse I footprinting experiments demonstrate that the competition between P9 and D1 is the result of nearly identical DNA-binding properties. Figure 3C (lanes 2–5) shows that P9 primarily protects stretches of dA·dT base pairs (W7–W14) of the SAT III repeat from DNAse I digestion. These tracts are also the main interaction sites for D1 (Figure 3D and E). Although both ligands bind similar sequences, some minor differences can be noted, as is manifested by the greater affinity of P9 for longer dA·dT tracts. A similar experiment performed with the artificial MATH20 protein yielded a footprint similar to that of P9 or D1 (lanes 5–7 of panel E). These interaction results are graphically summarized in Figure 3F.

In vivo displacement by P9 of D1 from SAT I and III in embryos

We tested whether P9 displaces D1 from SAT III *in vivo* by exposing whole embryos at the cellular blastoderm stage to this compound. Panels D–F of Figure 2 show that P9 induces a dose-dependent decrease of the D1 immunosignal in the chromocenter of nuclei. In untreated embryos, D1 (red) was found to localize predominantly to two foci in each nucleus (panel D) that correspond to SAT I and SAT III as verified by P9F binding (not shown). Exposure of embryos to 10 μM P9 (panel E), however, led to a loss of D1 from some but not all foci, typically resulting in the detection of only one D1-positive region in each nucleus. This displacement of D1 was essentially complete at the highest drug concentration tested (50 μM, panel F). These results extend our *in vitro* competition experiments and demonstrate that P9 displaces D1 from AT-rich satellite repeats *in vivo*.

D1 displacement from chromosomes of eye disks in vivo

The displacement of D1 was studied in more detail on mitotic chromosomes from eye imaginal disks isolated from P9-fed larvae. As shown above for neuroblast chromosomes (Figure 2A), the immunosignal of D1 is also predominantly confined to SAT III and SAT I in eye disk chromosomes isolated from control larvae (no P9, Figure 2G). In contrast, panels H and I show that D1 is displaced if larvae were fed yeast paste containing 100 μM P9. While displacement by P9 of D1 from SAT III was complete (panel H), we observed a stochastic loss of D1 from SAT I in about 20% of the examined mitotic spreads. Panel I shows such an example, where D1 was completely displaced from both SAT I and III repeats. This experiment shows that P9 displaces D1 from chromosomes isolated from eye disks, where the suppression of w^{m4} PEV is exerted.

The above experiments were carried out with large amounts of P9. As our polyamide supply is very limited, we used permeabilized larval brains in the following experiments, where lower concentrations can be used. Squashes of larval brains also have the advantage that they yield mitotic chromosome spreads of excellent quality (Figure 2A–C). As shown in panels J–M of Figure 2, the displacement of D1 in permeabilized brains occurs at a much lower P9 concentration. Notably, we observed that a P9 concentration of 250 nM suffices to displace D1 from SAT III (panel K, red arrow-

©2006 European Molecular Biology Organization

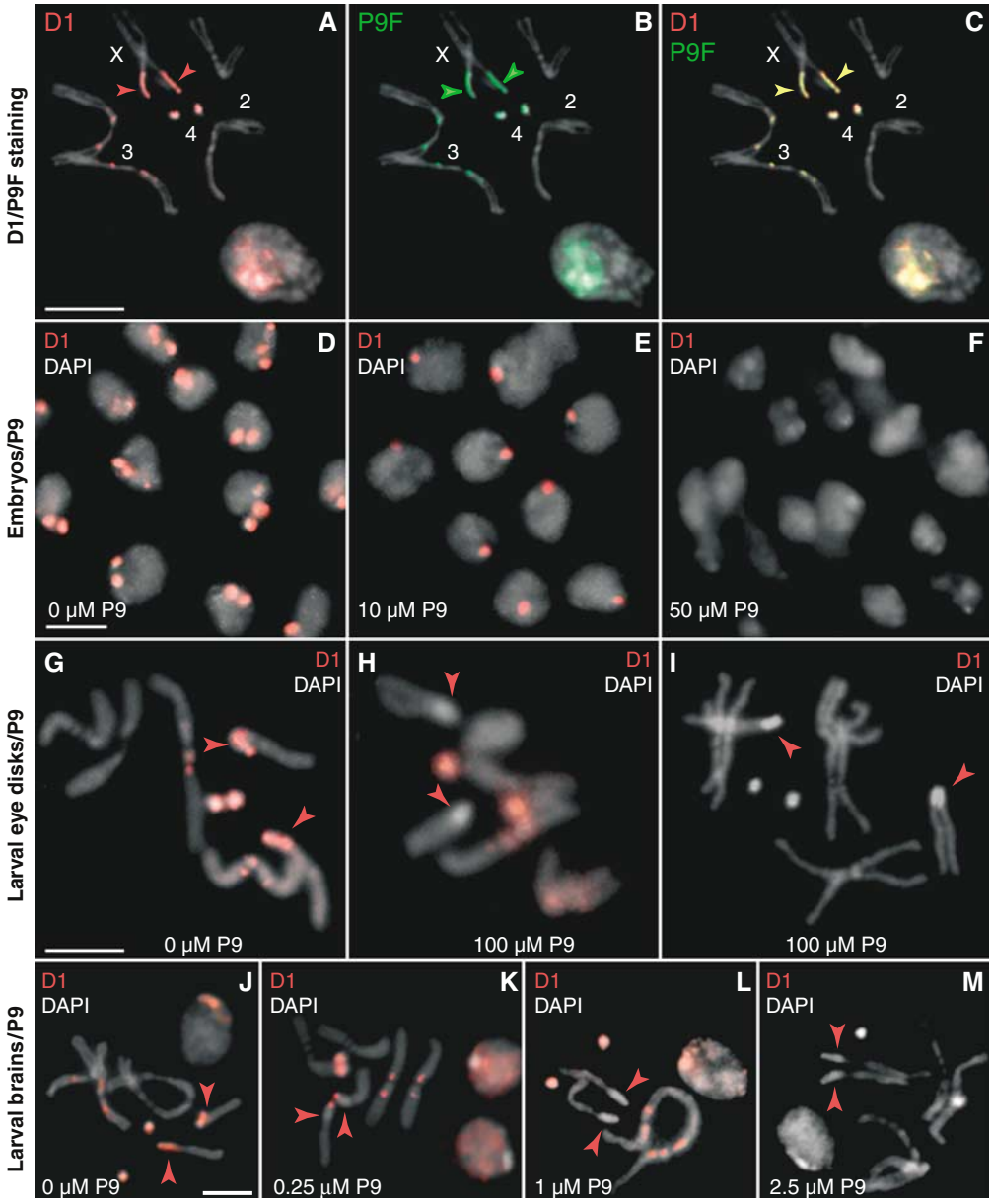


Figure 2 Uptake of oligopyrroles by whole embryos and larvae and displacement of D1. (**A–C**) Larval neuroblast nuclei and mitotic chromosomes stained with an antibody against D1 (red) and P9F (green) and counterstained with DAPI (gray). (**D–F**) Cellular blastoderm embryos were treated with 0, 10 or 50 µM P9, immunostained for D1 (red) and mounted in DAPI (gray). Treatment with 10 µM P9 results in a partial loss of the D1 signal, which is complete at 50 µM P9. (**G–I**) Third-instar larvae were fed colchicine (100 µg/ml) and P9 (0 or 100 µM) to assess displacement of D1 (red signal) from SAT I and SAT III repeats on mitotic chromosomes from eye imaginal disks. Arrowheads indicate the position of SAT III repeats, as determined by P9F staining (not shown). (**J–M**) The D1 signal in permeabilized larval brains treated with 0, 0.25, 1 or 2.5 µM P9, respectively. Merged DAPI (gray) and D1 (red) signals are shown. Arrowheads show the DAPI-bright SAT III array on the X chromosome. Scale bars: 5 µm.

heads), while higher concentrations (1–2.5 µM) are required to evict D1 from SAT I (panels L and M).

P9 displaces HP1 from SAT III but not from rDNA repeats

The possible implication of HP1 in the suppression of w^{m4} PEV by P9 was also studied. Figure 4A shows that HP1 (red) is bound to the pericentric regions of mitotic chromosomes, with a predominant association with SAT III and SAT V repeats of chromosomes X and 2, respectively. In contrast, the HP1 immunosignal is undetectable on the SAT I repeats on chromosomes 4 and Y (not shown) and at the tip of the X chromosome. It is important to note that Lex9F was used to counterstain these panels of Figure 4, as it serves to highlight SAT I (not SAT III) repeats that remain green since their HP1 immunosignal remains undetectable (panel A). This conclusion is also supported by an examination of the staining pattern of interphase nuclei. The inset of panel A shows the intense HP1 staining of SAT V and partial staining of the SAT III chromocenter (yellow), while HP1 is excluded from a region (green) that also harbors SAT I repeats.

Careful inspection of the HP1 immunosignal over the X chromosome at higher magnification (panel B) shows that its signal extends from the rDNA repeats to the adjacent SAT III array, where staining reproducibly appears in the form

©2006 European Molecular Biology Organization

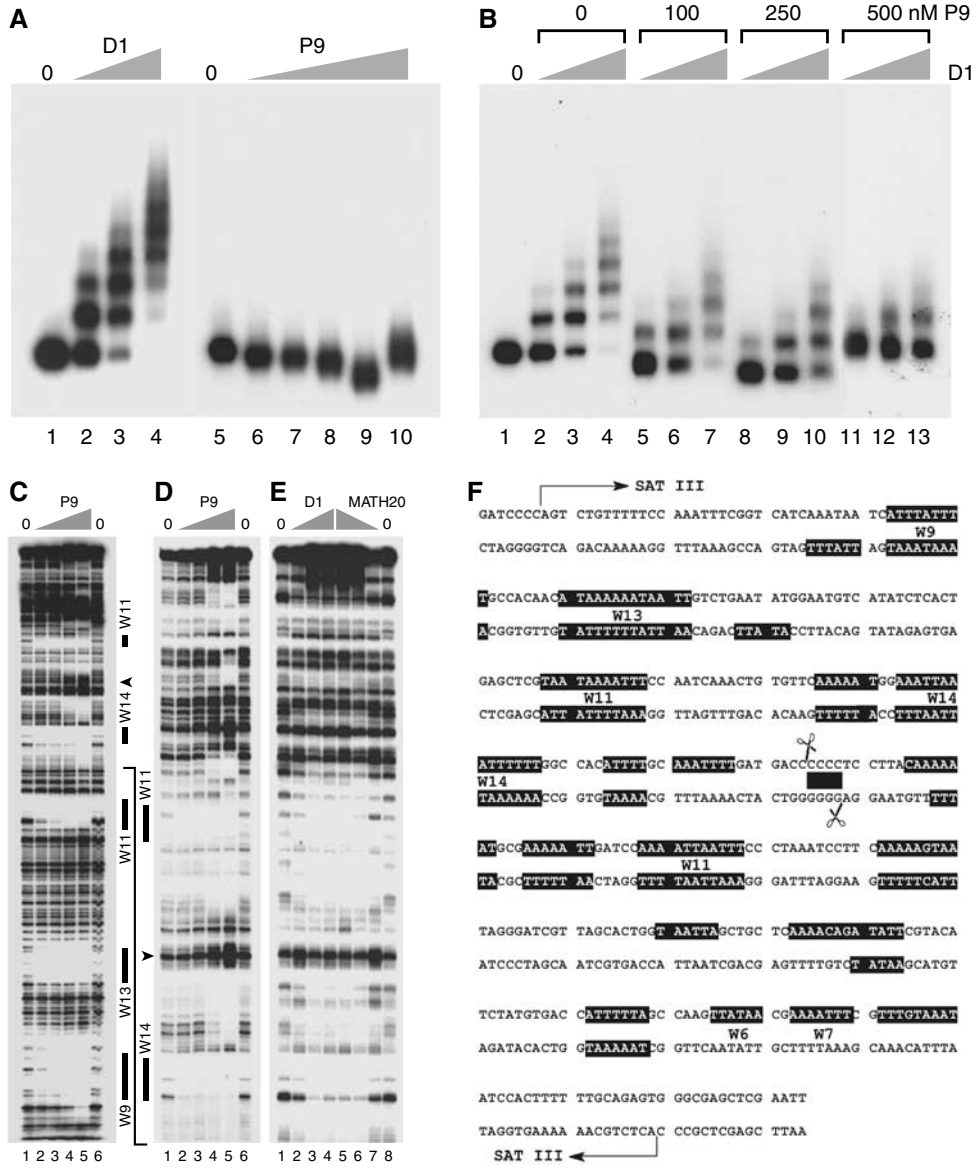


Figure 3 The P9 oligopyrrole competes with D1 for binding to SAT III repeats *in vitro*. (**A**) A SAT III monomer was incubated with 2, 5 or 10 ng D1 (2.7, 6.75 and 13.5 nM, lanes 2–4) or P9 (25, 50, 100, 250 or 500 nM, lanes 6–10). No-protein controls are shown in lanes 1 and 5. (**B**) SAT III was incubated with 1, 5 or 10 ng D1 (1.35, 6.75 or 13.5 nM) in the presence of 0, 100, 250 or 500 nM P9 as shown above the gel. Lane 1 contained no protein. The displacement of D1 by the drug is almost complete at 500 nM P9 (lanes 11–13). (**C**) SAT III monomer DNA 3′-end labeled on the lower strand was incubated with 0 (lanes 1 and 6) or 50, 100, 250, 500 nM P9 (lanes 2–5) and digested with DNAse I. Solid bars denote regions protected from digestion at the lowest P9 concentrations used and correspond to the largest dA·dT tracts (W) of SAT III (W6/W7–W14, where the number denotes the number of dA·dT bp). Note that only W tracts that are clearly resolved on the gels shown are indicated. The arrowhead indicates a P9-induced hypersensitive site. Binding of the same concentrations of P9 to SAT III (**D**) was compared to that of D1 and MATH20 (**E**). Lanes 2–4 of panel (E) correspond to 2, 5 and 10 ng D1 (2.7, 6.75 and 13.5 nM), while lanes 5–7 contained 2.5, 1.5 and 0.5 ng MATH20 (1.5, 0.9 and 0.3 nM). Lanes 1 and 8 are no-protein controls. The solid line starting at the bottom of (D) indicates the SAT III region shown on these gels relative to (C). These results are summarized on the sequence of a SAT III monomer shown in (**F**). Filled boxes indicate the regions protected by binding of D1, MATH20 and P9. Scissors indicate the sequence cleaved by topoisomerase II *in vivo* (Borgnetto *et al*, 1996), the filled box corresponds to the four consecutive dC·dG bp located within the staggered cut. This region spans the P9-induced DNAse I-hypersensitive site and is not bound by D1 or MATH20.

of a gradient that diminishes toward the centromere. In contrast to the binding of HP1 to the SAT III and rDNA repeats, the immunosignal of D1 shown for comparison is restricted to the centromere-proximal region of SAT III and only partially overlaps with HP1 (panel C).

Panels D and E of Figure 4 show that P9 selectively displaces HP1 from SAT III repeats, but not from the rDNA array. This conclusion is best noted on the enlarged X chromosomes shown in panel E. This micrograph shows the green Lex9F stain of SAT I at the chromosomal tip and the red immunosignal of HP1 over the rDNA repeats, while the SAT III region remains unstained (gray). Hence, following treatment with P9, HP1 is delocalized from the SAT III repeats, but remains associated with the rDNA array. Examination of panel D demonstrates that, in contrast, the protein remains prominently associated with SAT V repeats on chromosome 2. The nuclei shown in panel D also confirm this conclusion. In this case, the loss of HP1 from SAT III is

©2006 European Molecular Biology Organization

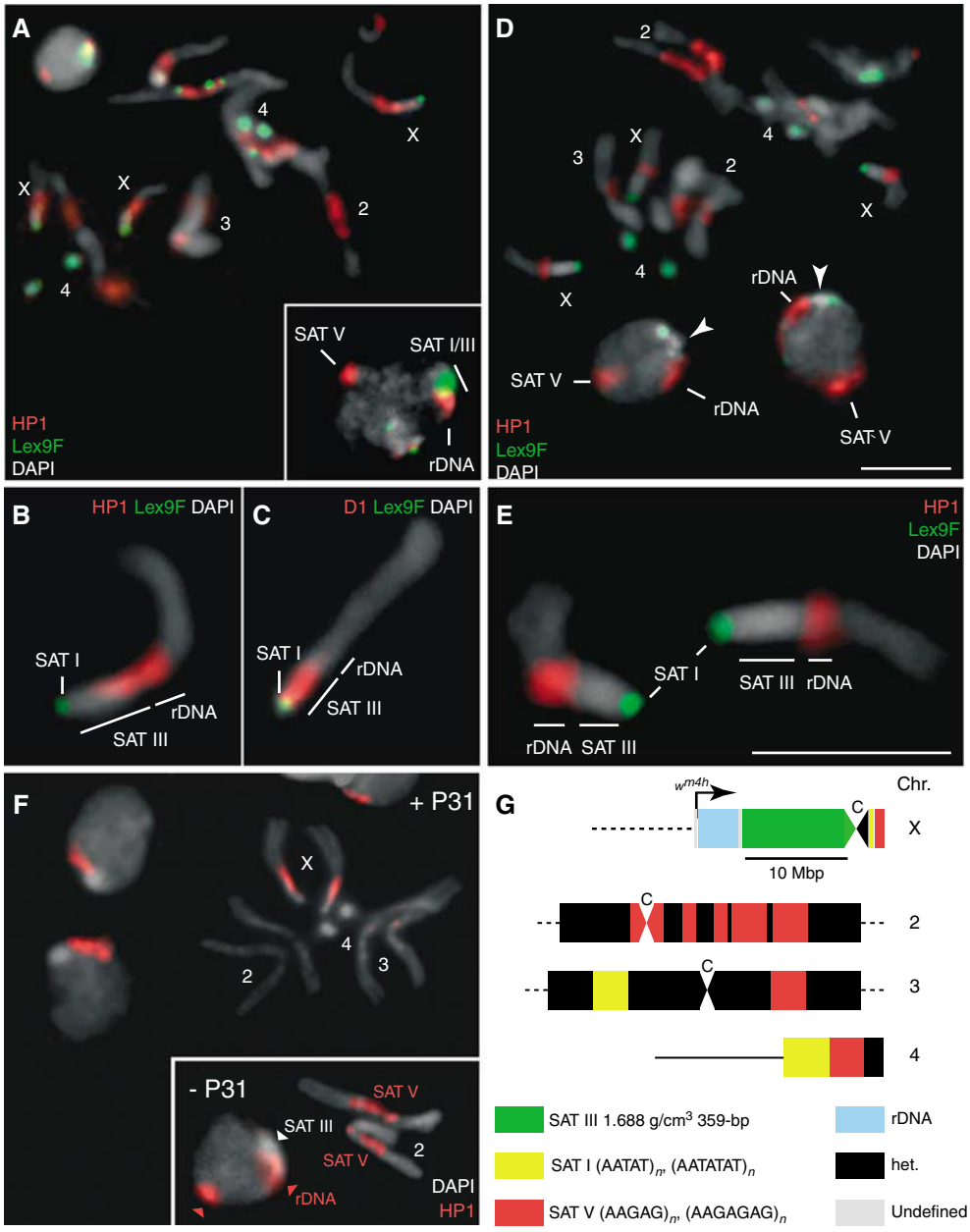


Figure 4 D1 Displacement results in a selective loss of HP1. (**A**) Localization of HP1 (red) in interphase nuclei and mitotic chromosomes of female diploid larval neuroblasts. Samples were counterstained with Lex9F (green) and DAPI (gray). The inset shows a similarly stained interphase nucleus. A region that stains strongly for HP1 alone corresponds to SAT V repeats on chromosome 2, as assessed by P31TR staining (not shown). rDNA repeats are indicated. (**B**, **C**) Magnified X chromosomes after staining for HP1 or D1 (red), respectively, and counterstaining with Lex9F (green) and DAPI (gray). (**D**–**E**) Localization of HP1 (red) after treatment with 2.5 μM P9. Magnified X chromosomes are shown in (E). Samples were counterstained with Lex9F (green) and DAPI (gray). In (D), arrowheads indicate the HP1-depleted DAPI-bright region that corresponds to SAT III repeats in interphase nuclei after treatment with P9; SAT V repeats on chromosome 2 are also indicated. (**F**) The localization of HP1 following treatment with P31. The inset shows untreated chromosomes 2 and a nucleus with the SAT V-associated HP1 signal (arrowhead). A general representation of major satellite blocks is schematized in (**G**), which also shows the approximate location of the *white* gene on a w^{m4} X chromosome, within a type I insertion sequence characteristic of some rDNA repeats. 'C' denotes the centromere. Scale bars: 5 μm.

reflected by the appearance of a green–gray–red pattern over the chromocenter, corresponding to the SAT I, SAT III and rDNA repeats, respectively, while the strong HP1 association site on SAT V repeats is unaffected.

We tested next the sensitivity of HP1 binding to chromatin to treatment with the P31 polyamide. As shown in the inset of Figure 4F, HP1 is prominently associated with SAT V repeats on chromosome 2. Incubation with P31 led to a selective displacement of HP1 from the pericentric region of chromosome 2, without affecting the association of the protein with SAT III and rDNA repeats on the X chromosome. We conclude from these results that HP1 can be targeted by polyamides in a sequence-specific fashion that reflects the selectivity of P9 and P31 for SAT III and SAT V, respectively. This suggests that the association of HP1 with satellites repeats may occur, at least in part, via DNA sequence recognition (Zhao *et al*, 2000; Perrini *et al*, 2004).

©2006 European Molecular Biology Organization

HP1 delocalization results in the invasion of rDNA repeats by D1

The displacement of HP1 from SAT III repeats might be due to direct competition with P9, predicting that D1 and HP1 may compete for binding to SAT III. Displacement of HP1 from heterochromatin should then result in additional D1 binding to sites normally occupied by HP1. We tested this possibility by feeding a histone deacetylase inhibitor to growing third-instar larvae. Treatments with trichostatin A (TSA) have been shown to result in the delocalization of HP1 from heterochromatin (Taddei *et al*, 2001). Figure 5A shows the HP1 and D1 localization patterns normally observed in larval neuroblast nuclei. HP1 (red) partially associates with SAT III repeats stained with P9F (green) as well as with SAT V repeats and is largely absent from the nucleoplasm. The localization of D1 is strictly correlated with P9F-positive foci, which correspond to SAT I and SAT III repeats. TSA feeding resulted in a loss of HP1 from heterochromatin and a relocalization of the protein to the nucleoplasm (Figure 5B). Significantly, this displacement was accompanied by an extension of the D1 signal: instead of colocalizing with P9F-positive foci, yielding a characteristic yellow signal (Figure 5A), the D1 antibody also stained an immediately adjacent region, yielding both red and yellow signals. The red-only domain occupied by D1 after TSA treatment corresponds to the rDNA repeats normally associated with HP1 (data not shown and see below).

SAT III protein composition affects the white gene across the rDNA repeats

The delocalization of D1 and HP1 from SAT III repeats by P9 correlates with suppression of w^{m4} PEV. However, as the inverted *white* gene lies near the centromere-distal end of the rDNA repeats (see Figure 4G), this effect must somehow be transduced across the 3.5-Mbp rDNA locus. How satellite protein composition can affect gene expression over such distances is an important question, which we addressed by performing fluorescence *in situ* hybridization (FISH) experiments. Hybridization of rDNA and *white* probes to neuroblast nuclei and chromosomes resulted in well-defined signals. The *Suppressor of forked* gene (*Su(f)*), which maps to the heterochromatin region of the X chromosome and is distal to the

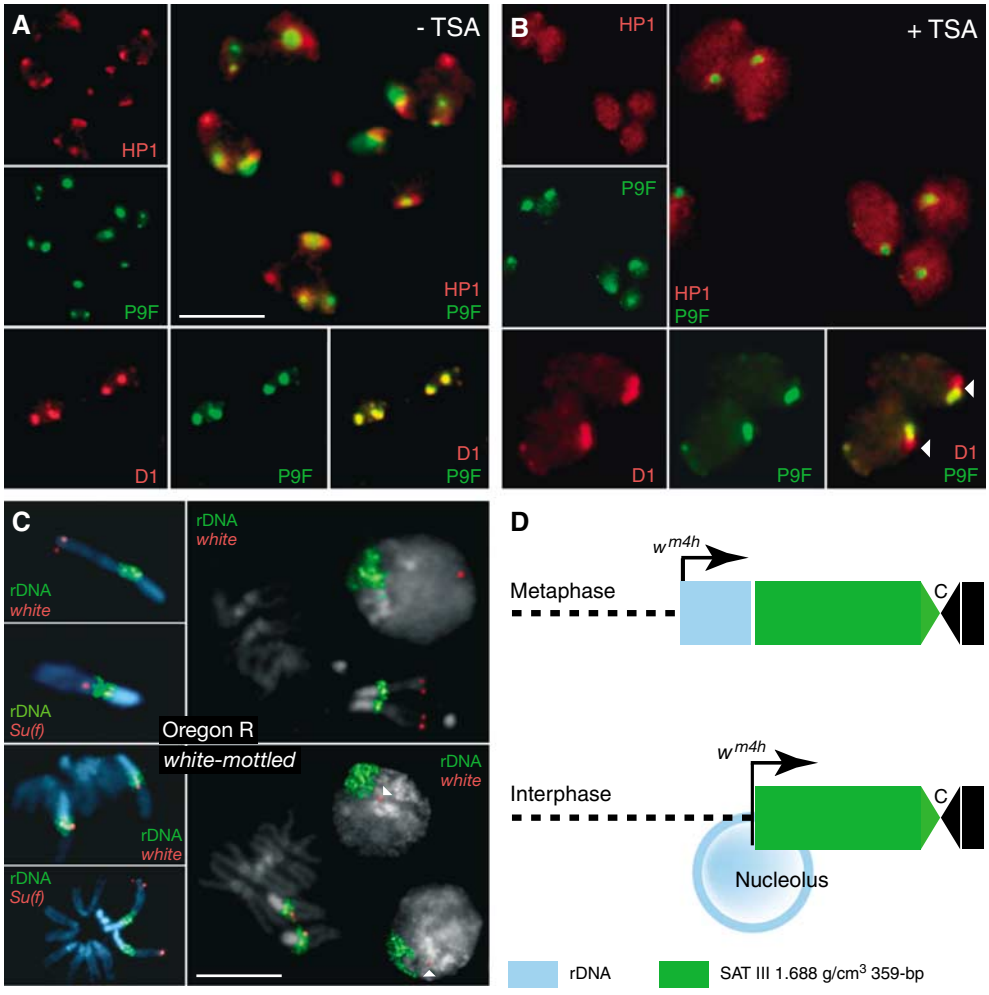


Figure 5 SAT III and rDNA repeats define alternative D1 and HP1 domains. Neuroblasts from control (**A**) or TSA-fed (**B**) third-instar larvae were immunostained for HP1 or D1 (red) and counterstained with P9F (green). Individual and merged signals are as indicated in the photographs. TSA feeding induces a delocalization of HP1 to the nucleoplasm and an extension of the D1 signal from SAT III to an adjacent P9F-negative region that corresponds to the rDNA repeats. (**C**) The results of FISH experiments performed with *white* and *Su(f)* probes (red) or a full-length rDNA probe (green). Photographs at the top and bottom show results obtained from wild-type Oregon R or w^{m4} neuroblasts, respectively. Arrowheads indicate the *white* signal detected in interphase nuclei. Scale bars: 5 μm. (**D**) The juxtaposition of the rDNA-linked *white* gene to SAT III repeats as a result of looping out of rDNA sequences in the nucleolus.

©2006 European Molecular Biology Organization

rDNA repeats, was used as a control. The insets of Figure 4C show hybridization signals for rDNA (green) and *white* or *Su(f)* sequences (red) on X chromosomes from Oregon R (wild-type) or *white-mottled* larvae, clearly showing the inversion of the *Su(f)* and *white* genes on w^{m4} X chromosomes. In this case, the *white* gene could not be resolved from the rDNA repeats, consistent with its localization within type I insertion sequences located at the distal end of the rDNA locus (Schotta *et al*, 2003). In interphase nuclei, the *white* signal in wild-type cells was always localized at the opposite pole of the nucleus relative to rDNA repeats. Significantly, in contrast to the close linkage seen in mitotic chromosomes, the *white* signal was always well resolved from the rDNA repeats in w^{m4} interphase nuclei, localizing instead immediately adjacent to DAPI-bright foci that correspond to SAT III repeats. As diagrammed in Figure 5D, these results support a model whereby looping out of rDNA sequences within the nucleolus places the *white* gene very close to the SAT III repeats where the effects of P9 are exerted.

SAT III-specific displacement of topo II by P9

Topo II is one of the growing numbers of proteins found to be associated with centric heterochromatin. This association is particularly evident in immunofluorescence studies of prophase and mitotic chromosomes, where the topo II signal contains conspicuous centromeric foci in addition to axial, subchromosomal extensions (Maeshima and Laemmli, 2003). While topo II appears clearly enriched in centric heterochromatin of mitotic chromosomes, an important association with heterochromatin is also noted in interphase nuclei. This interaction was previously noted by the prominent cleavage of SAT III repeats by topo II *in vivo* (Käs and Laemmli, 1992). Is topo II associated with SAT III involved in PEV?

Figure 6A shows the localization of topo II in nuclei and mitotic chromosomes of larval neuroblasts immunostained for topo II (red) and counterstained with DAPI (gray) and P9F (green). Examination of interphase nuclei in panel A shows a granular topo II signal that appears excluded from a prominent DAPI-bright (gray) and P9F-bright (green) region thought to correspond to chromosome 4. On chromosome arms, the topo II signal shows a banded pattern that is reminiscent of observations made with mammalian chromosomes. Topo II is also enriched on SAT III and on the adjacent rDNA repeats (see the enlarged X chromosome in the insets of panel A). A comparison of the P9F and protein staining patterns confirms that SAT III repeats constitute the major chromosomal site (yellow merged color in Figure 6A) where satellite sequences and topo II strictly colocalize. Lack of staining of the fourth chromosome, as in the case of HP1, could be due to a reduced accessibility of the antigen in this compact chromosome.

Experiments discussed above establish that D1 and HP1 are displaced from SAT III by P9. This observation is extended to topo II next. Exposure of permeabilized brains to P9 in the range of 250 nM to 1 μM resulted in the selective loss of topo II from SAT III repeats (Figure 6C and D and insets with enlarged X chromosomes), while it remained associated with rDNA repeats and the chromosome arms. Note that topo II is completely evicted at the highest concentration of P9 used (2.5 μM, Panel E). These results show that, similarly to HP1, topo II is selectively displaced from SAT III repeats by P9 (250 nM to 1 μM), while it remains bound to the adjacent rDNA repeats. Such a displacement might be due to direct competition between P9 and topo II for binding to AT-rich satellite sequences, or might be a consequence of the loss of D1 that might serve to stabilize the interaction of topo II with SAT III repeats.

Suppression of PEV by a topo II inhibitor

The colocalization of topo II with two heterochromatin proteins, HP1 and D1, on SAT III repeats suggests that topo II might also be a structural/enzymatic component of heterochromatin. Such a notion is supported by our finding that feeding w^{m4h} larvae sublethal doses of VM26, a topo II poison, resulted in a dose-dependent suppression of PEV (Figure 7). Suppression was complete at the highest drug concentrations used, as judged from the recovery of adult flies with wild-type eye color. A similar observation was made by feeding VM26 to flies harboring a variegating mini-*white* construct (*Tp(3;Y)BL2*) inserted on the Y chromosome (Lu *et al*, 1996). In contrast, m-AMSA, another topo II poison structurally unrelated to VM26, had no effect on w^{m4} PEV (Figure 6): m-AMSA does not induce topo II cleavage in SAT III repeats, although it promotes cleavage at other genomic sites (Borgnetto *et al*, 1996). We conclude from these results that topo II is implicated in *white-mottled* heterochromatin-mediated silencing and that this enzyme might more generally play a functional role in the assembly of heterochromatin, as suggested by the effect of VM26 in the *Tp(3;Y)BL2* line.

Discussion

P9 is a simple polyamide of linked pyrrole amino acids with a subnanomolar binding preference for long dA·dT-runs (Supplementary Table I). It was previously shown that this compound, if fed to developing flies, suppresses the w^{m4} phenotype, an effect manifested by a more homogenous, redder appearance of the eyes. This 'red shift' is due to a generally higher expression of the rearranged *white* gene (Janssen *et al*, 2000a). This observation suggested that P9 could work through chromatin opening, yielding a reduced heterochromatic state of SAT III, the main highly repetitive DNA flanking the *white* gene on the inverted w^{m4} X chromosome. Results reported here establish that, both *in vivo* and *in vitro*, P9 displaces the D1 protein from SAT III. In addition, we show that this compound results in a selective loss of HP1 and topo II from SAT III, while these proteins remain bound to low-affinity targets for P9, such as the rDNA repeats.

The displacement of D1 by P9 is an expected result. Polyamides are deeply buried in the DNA minor groove through specific hydrogen bonding and Van der Waals interactions (Dervan and Edelson, 2003). Similarly, the AT-rich specificity of D1 is mediated through its AT hooks reaching into the DNA minor groove of dA·dT-runs (Slama-Schwok *et al*, 2000). Oligopyrroles such as P9 mimic the DNA minor groove contacts of AT hooks. Hence, the displacement of D1 by P9 is explained by a molecular competition reaction. The salient observation of this paper is the selective displacement of the heterochromatin proteins HP1 and topo II from P9-targeted DNA such as SAT III. This is an unexpected finding since HP1 is thought to bind chromatin largely through protein–protein contacts mediated by its chromo domain,

 ©2006 European Molecular Biology Organization

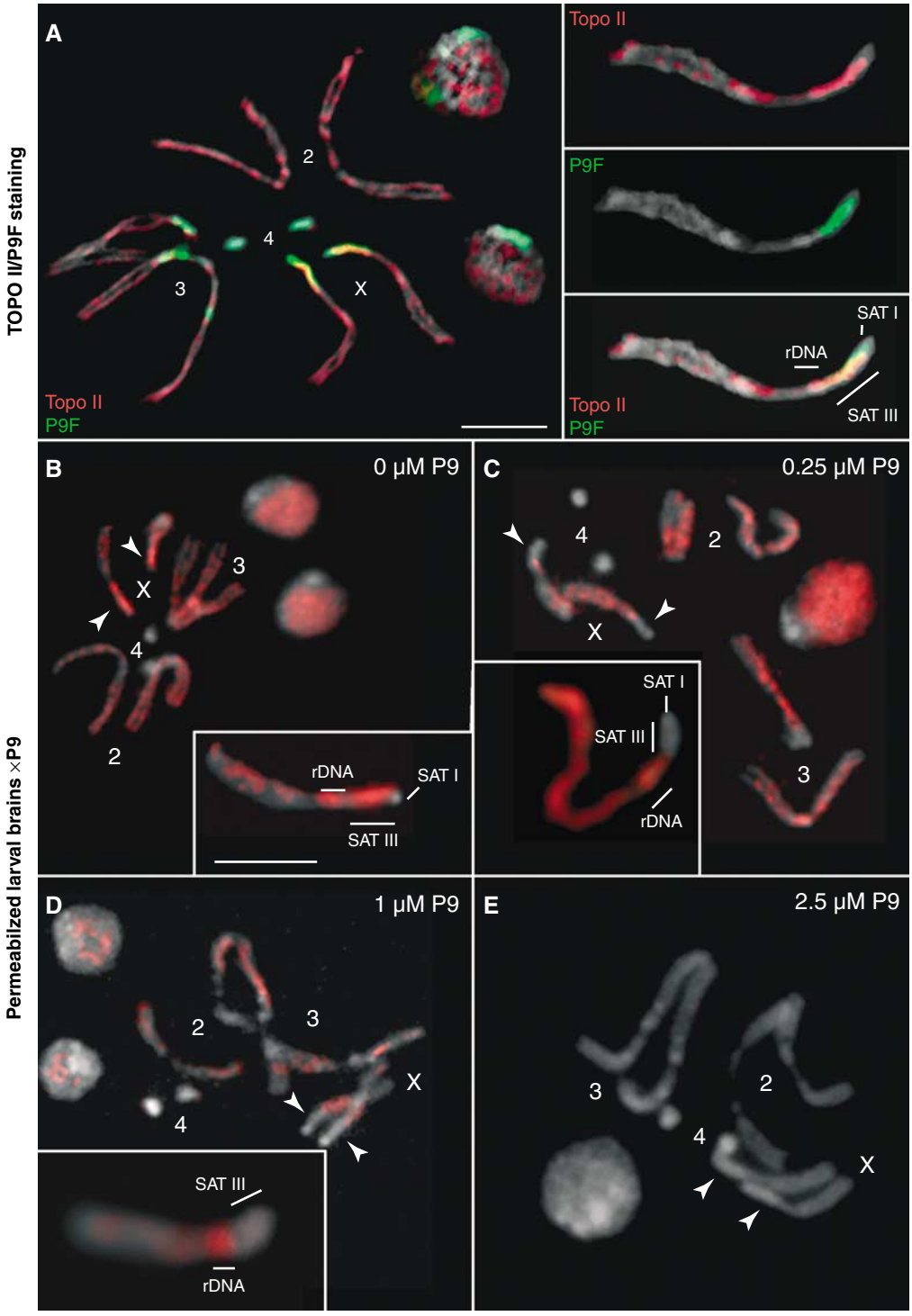


Figure 6 Localization of topoisomerase II and displacement from SAT III by P9. (**A**) Neuroblasts from female third-instar larvae were immunostained for topo II (red) and counterstained with P9F (green) and DAPI (gray). Individual chromosomes are numbered as shown. The localization of SAT I, SAT III and rDNA repeats is indicated in the inset showing a magnified X chromosome. SAT III repeats are associated with a particularly strong topo II signal, the adjacent region spanning the 3.5-Mbp rDNA repeats is also stained. Photographs to the right show the individual topo II, P9F and merged signals. In panels (**B–E**), third-instar female larval brains were permeabilized and treated with 0, 0.25, 1 and 2.5 μM P9, respectively. Drug treatment leads to a delocalization and displacement of topo II that is selective for SAT III repeats at low to intermediate concentrations of the drug (250 nM to 1 μM, C and D), while topo II remains associated with rDNA repeats (see magnified X chromosomes in insets). Topo II is completely displaced from nuclei and mitotic chromosomes at the highest drug concentration tested (2.5 μM, E). Arrowheads indicate the position of SAT III repeats. Samples were counterstained with DAPI (gray). Scale bars = 5 μm.

which recognizes the methylated state of the K9 amino acid of histone H3 (reviewed in Huisinga *et al*, 2006). It is possible that D1 might recruit the enzymatic activities required to create an HP1-binding site. Other sequence-selective satellite-binding proteins such as PROD (Platero *et al*, 1998) and SU(VAR)3–7 (Cléard and Spierer, 2001) might play a similar

©2006 European Molecular Biology Organization

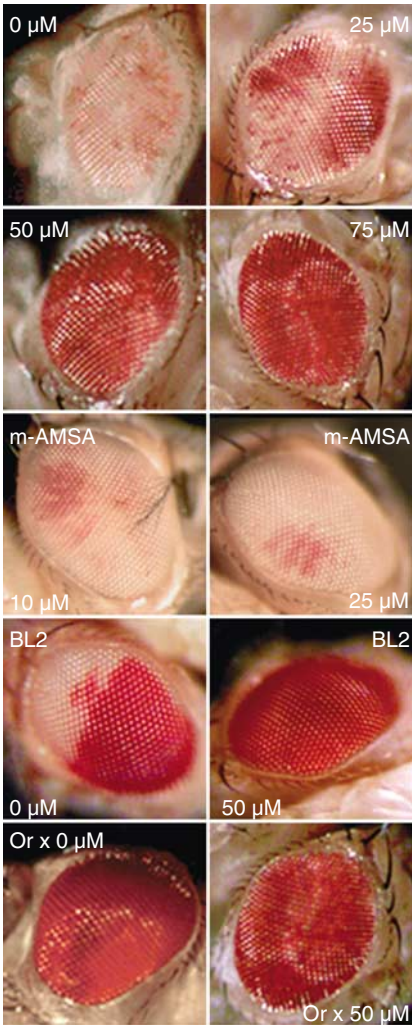


Figure 7 Treatment with VM26, a topoisomerase II poison, suppresses *white-mottled* PEV. Third-instar larvae from the w^{m4h}, *Tp(3;Y)BL2* ('BL2') and Oregon R ('Or') lines were fed VM26 or m-AMSA at the indicated concentrations. Representative eyes of 5-day-old adult flies are shown and correspond to those observed in >80% of hatched flies ($n=30$ for each drug concentration). Eyes from untreated or VM26-treated wild-type Oregon R flies are shown for comparison. Extracted eye pigments from two independent feeding experiments were measured for absorbance at 480 nm and yielded the following values: 0.18 ± 0.03 (w^{m4h}, 0 μM VM26); 0.36 ± 0.05 (w^{m4h}, 25 μM VM26); 0.79 ± 0.05 (w^{m4h}, 50 μM VM26); 0.91 ± 0.03 (w^{m4h}, 75 μM VM26); 0.22 ± 0.06 (w^{m4h}, 10 μM m-AMSA); 0.21 ± 0.03 (w^{m4h}, 25 μM m-AMSA); 0.43 ± 0.05 (BL2, 0 μM VM26); 0.92 ± 0.04 (BL2, 50 μM VM26); 0.92 ± 0.02 (Oregon, 0 μM VM26); 0.89 ± 0.02 (Oregon, 50 μM VM26).

role. Taken together, our data suggest that D1 is directly or indirectly implicated in the association of HP1 and topo II with SAT III repeats and that the effects of P9 are exerted via the selective displacement of these proteins, resulting in a reduction of silencing of the nearby *white* gene in w^{m4} flies.

Our conclusions are based on the results of *in vitro* competition/binding assays and of *in vivo* P9 bathing and feeding experiments performed on embryos and larvae, respectively. Experimentally, we have followed the chromosome association of D1, HP1 and topo II as a function of P9 treatment by immunofluorescence and with the help of fluorescent polyamide derivatives added during the staining procedure. The latter serve as optical highlighters of the targeted genomic regions (Figure 1). Binding experiments show that P9 competes with D1 for binding to similar sites in SAT III repeats that are also bound by the artificial MATH20 protein (Figure 3). This mimicry extends to the ability of P9, following entry in whole embryos or ingestion by larvae, to displace D1 from its SAT III-binding sites *in vivo* (Figure 2), one important consequence of which is the loss of HP1 from SAT III (Figure 4) which demonstrates a disassembly most likely caused by the displacement of D1 by P9.

However, our results also indicate the likely significance of HP1/DNA contacts: displacement of HP1 from SAT III or SAT V repeats is strictly determined by the sequences targeted by the P9 or P31 polyamide used, respectively (Figure 4). This is in line with the report that telomeric sequences are bound directly by HP1 (Perrini *et al*, 2004), an observation which may extend to a more general binding of HP1 to DNA in chromatin (Zhao *et al*, 2000). The displacement of HP1 from SAT III repeats by P9 might be due to direct competition, a consequence of which is that HP1 and D1 would also compete for binding to similar sites. Alternatively, HP2, a companion of HP1 containing two AT-hook motifs (Shaffer *et al*, 2002), might be a target for competition by P9 and D1. The opposite gradient-like distributions of HP1 and D1 on SAT III repeats (Figure 4B and C) support such a possibility, suggesting mutually exclusive interactions that might extend to the rDNA repeats. This hypothesis follows from our observation that TSA feeding results in a displacement of HP1 from heterochromatin and the subsequent invasion of the rDNA repeats by D1 (Figure 5). While not a target for P9 (see Figure 1 and the lack of staining of rDNA repeats by P9F), the rDNA repeats are generally AT-rich and may contain binding sites for D1 that are normally occupied by HP1. A destabilization of HP1/heterochromatin interactions by histone hyperacetylation (Taddei *et al*, 2001) would thus result in an increased accessibility to D1, which is indeed what we observe.

The displacement of topo II from SAT III repeats by P9 (Figure 6) might similarly be due to direct competition with P9 or, alternatively, reflect a secondary delocalization that follows displacement of D1 from that AT-rich sequences. A role for this enzyme in SAT III heterochromatin-mediated silencing, as supported by its localization and the striking suppression of PEV by VM26, a topo II poison, is consistent with the results of recent studies indicating that many modifiers of PEV are involved in several important aspects of chromosome structure and function (Le *et al*, 2004). We note that a link was previously established between chromatin compaction, topo II and the negative control of gene expression: topo II was shown to be implicated in Polycomb-mediated repression (Lupo *et al*, 2001). A similar interplay with HP1 might reflect X chromosome-specific silencing phenomena. Genetic screens did not previously identify topo II as a modifier of PEV and flies heterozygous for the Df(2L)TW158 deficiency, which deletes the gene encoding topo II, do not suppress w^{m4} PEV (Wustmann *et al*, 1989). However, topo II is a very abundant protein and a 50% reduction in protein levels might not have measurable effects on overall levels of enzymatic activity. Our observations and the failure of Df(2L)TW158 to modify w^{m4} PEV are therefore not necessarily contradictory.

Feeding VM26 to developing w^{m4} larvae suppresses PEV, an unexpected result which could be due to a general—but

©2006 European Molecular Biology Organization

reversible—inhibition of topo II, or to the possible accumulation of topo II/DNA complexes in SAT III repeats that might interfere with dynamic interactions with proteins such as HP1. Synthetic diamidines that promote the relocalization of topo II to AT-rich satellites such as SAT III indeed also suppress PEV when fed to w^{m4} larvae (Susbielle *et al*, 2005). We do not rule out alternative mechanisms: trapped complexes might for instance interfere with the bidirectional transcription of SAT III repeats (Aravin *et al*, 2003), whose role in heterochromatin assembly, structure and function remains to be elucidated. The RNAi machinery has been linked to the variegated expression of mini-*white* reporters (Pal-Bhadra *et al*, 2004), but the mutants used were not tested against the w^{m4} inversion. Preliminary results (data not shown) indicate that *piwi* and *argonaute 1* mutants do not modify w^{m4} PEV, suggesting that there exist distinct pathways for the initiation of heterochromatin assembly (Danzer and Wallrath, 2004; Huisinga *et al*, 2006). In addition, SAT III-like sequences are interspersed throughout the *D. melanogaster* genome (Losada and Villasante, 1996) and the genomic origin of putative SAT III siRNAs has not been established. We are developing a combined DNA/RNA FISH method that can be applied to *Drosophila* larval tissues to address this important question in the future.

We favor a model whereby the suppression of w^{m4} PEV by VM26 is a direct result of interfering with topo II activity in SAT III repeats. The suppression of PEV in the BL2 line, in which *white* is inserted in the largely heterochromatic Y chromosome, further indicates that heterochromatin-mediated silencing in regions other than SAT III is also sensitive to topo II activity levels. Such a requirement for catalytically active topo II in heterochromatin assembly is generally supported by results showing active enzyme localized to centromeric regions and to human alpha-satellites that resemble SAT III sequences (Agostinho *et al*, 2004).

The strikingly different behavior of HP1 and topo II localized in the rDNA repeats adjacent to the SAT III array is discussed next. The known preference of topo II for association with AT-rich sequences (Maeshima and Laemmli, 2003) might explain its interaction with both SAT III and rDNA repeats. The lack of sensitivity of HP1 to displacement by P9 suggests that different mechanisms can direct HP1 to these closely juxtaposed chromatin domains. The D1-independent association of HP1 with rDNA correlates instead with the methylation of lysine 9 of histone H3 in this region (Ahmad and Henikoff, 2002). This suggests that packaging of the *Drosophila* rDNA repeats into heterochromatin may serve to regulate their expression. The colocalization of topo II and HP1 in this highly condensed region might reflect a coordinate interaction important for the control of rDNA transcription whose negative regulation involves the recruitment of histone deacetylase and methyl-transferase activities (Santoro *et al*, 2002). In contrast, resumption of proliferation and activation of ribosomal RNA transcription in *Drosophila* tissue culture cells is accompanied by the replication-independent replacement of histone H3 methylated on the lysine 9 position by the histone variant H3.3 in rDNA sequences (Ahmad and Henikoff, 2002). We speculate that the stable association of topo II with rDNA repeats might play a role in regulating the topological constraints that are likely to accompany such a replacement process.

Our results demonstrate that interfering with D1, HP1 and topo II interactions with SAT III correlates with P9-induced w^{m4} gain-of-function. However, the inverted *white* gene has been mapped very close to—or even within—rDNA repeats (Schotta *et al*, 2003): suppression of w^{m4} PEV by P9 acting on SAT III must then somehow be exerted over the distance spanned by the rDNA locus, or approximately 3.5 Mbp. These observations are not consistent with a linear model of heterochromatin spreading and support instead a spatial model of heterochromatin-mediated silencing (Csink and Henikoff (1996); see also Huisinga *et al*, 2006 and references therein). Our FISH results (Figure 5C), which show that the rDNA-linked inverted *white* gene is indeed more closely associated with SAT III repeats in interphase chromatin, provide additional support for such a model, in which looping out of the rDNA into the nucleolus would place *white* in close proximity to SAT III (Figure 5D).

Finally, we note that the stochastic inactivation of the w^{m4} allele might be a direct reflection of developmentally regulated rDNA expression. The association of D1 and HP1 with SAT III and rDNA repeats is sensitive to histone acetylation levels (Figure 5B) and these proteins may compete for binding to similar sites. Our observations raise the intriguing possibility that a dynamic equilibrium of alternative interactions of D1 and HP1 with SAT III and rDNA repeats might serve to regulate the structure of the centromere-proximal region of the X chromosome and participate in the regulation of rDNA expression.

Materials and methods

Fluorescent drug staining and immunostaining

Brains were dissected from Oregon R third-instar larvae in 0.7% saline and incubated in 0.5% sodium citrate for 15 min. Dissected tissues were carefully transferred to a small drop of 0.5% citrate on a poly-lysine-coated microscope slide. Samples were fixed for 10 min in 2% PFA in PBS containing 0.1% Triton X-100 followed by 10 min in 2% PFA, 45% acetic acid. This fixation protocol yields qualitatively better results than the short acid fixation previously used for D1 localization (Aulner *et al*, 2002); avoidance of colchicine yields high-quality mitotic chromosomes. Labeling with rabbit polyclonal antibodies against D1 (1:100 dilution) was performed exactly as described (Monod *et al*, 2002). Samples were stained for 10 min with Lex9F (20 nM), P9F (5 nM) or P31TR (2 nM) in PBS containing 1 mM $MgCl_2$, washed twice in the same buffer, mounted in DAPI and visualized by epifluorescence microscopy. Staining with a polyclonal rabbit antibody against HP1 kindly provided by R Kellum (University of Kentucky) was performed using a 1:5000 dilution and permeabilized larval brains as described in Supplementary data for mock P9 treatments. The mouse C1A9 monoclonal HP1 antibody was also used, yielding identical results. Labeling with rabbit polyclonal antibodies against *Drosophila* topo II (1:1000 dilution) was performed as described above. Photographs were taken on a Leica DMRB microscope equipped with a CoolSnap HQ CCD camera and processed with Adobe Photoshop. *In vivo* and *ex vivo* drug treatments, FISH experiments and DNA binding/footprinting assays were performed as described in Supplementary data.

Supplementary data

Supplementary data are available at *The EMBO Journal* Online.

Acknowledgements

We thank Rebecca Kellum (University of Kentucky) for her gift of antiserum against *Drosophila* HP1 and Lori Wallrath (University of Iowa) and Sally Elgin (Washington University) for samples of the C1A9 antibody. We thank J Eissenberg, P Geyer, E Kuhn-Parnell,

©2006 European Molecular Biology Organization

M Grigoriev and M Weber for comments on the manuscript. The Centre de Biologie du Développement provided *Drosophila* culture facilities and the Institut d'Exploration Fonctionnelle des Génomes microscopy facilities. This work was supported by the Association pour la Recherche sur le Cancer (grant no. 3319), by the ACI Cancéropôle 2004 program 'Genetic instability as a pejorative outcome in cancer' and by the National Institutes of Health (grant GM29006 to T-sH). R Blattes and G Susbielle are supported by doctoral fellowships from the French Ministry of Research and from the Ligue Nationale Contre le Cancer, respectively.

References

Adachi Y, Luke M, Laemmli UK (1991) Chromosome assembly *in vitro*: topoisomerase II is required for condensation. *Cell* **64:** 137–148

Agostinho M, Rino J, Braga J, Ferreira F, Steffensen S, Ferreira J (2004) Human topoisomerase IIalpha: targeting to subchromosomal sites of activity during interphase and mitosis. *Mol Biol Cell* **15:** 2388–2400

Ahmad K, Henikoff S (2002) The histone variant H3*3 marks active chromatin by replication-independent nucleosome assembly. *Mol Cell* **9:** 1191–1200

Aravin AA, Lagos-Quintana M, Yalcin A, Zavolan M, Marks D, Snyder B, Gaasterland T, Meyer J, Tuschl T (2003) The small RNA profile during *Drosophila melanogaster* development. *Dev Cell* **5:** 337–350

Aulner N, Monod C, Mandicourt G, Jullien D, Cuvier O, Sall A, Janssen S, Laemmli UK, Käs E (2002) The AT-hook protein D1 is essential for *Drosophila melanogaster* development and is implicated in position-effect variegation. *Mol Cell Biol* **22:** 1218–1232

Borgnetto ME, Zunino F, Tinelli S, Käs E, Capranico G (1996) Drug-specific sites of topoisomerase II DNA cleavage in *Drosophila* chromatin: heterogeneous localization and reversibility. *Cancer Res* **56:** 1855–1862

Cléard F, Spierer P (2001) Position-effect variegation in *Drosophila*: the modifier Su(var)3–7 is a modular DNA-binding protein. *EMBO Rep* **2:** 1095–1100

Csink AK, Henikoff S (1996) Genetic modification of heterochromatic association and nuclear organization in *Drosophila*. *Nature* **381:** 529–531

Danzer JR, Wallrath LL (2004) Mechanisms of HP1-mediated gene silencing in *Drosophila*. *Development* **131:** 3571–3580

Dervan PB, Edelson BS (2003) Recognition of the DNA minor groove by pyrrole–imidazole polyamides. *Curr Opin Struct Biol* **13:** 284–299

Girard F, Bello B, Laemmli UK, Gehring WJ (1998) *In vivo* analysis of scaffold-associated regions in *Drosophila:* a synthetic high-affinity SAR binding protein suppresses position effect variegation. *EMBO J* **17:** 2079–2085

Huisinga KL, Brower-Toland B, Elgin SC (2006) The contradictory definitions of heterochromatin: transcription and silencing. *Chromosoma* **28:** 28

Janssen S, Cuvier O, Muller M, Laemmli UK (2000a) Specific gain- and loss-of-function phenotypes induced by satellite-specific DNA-binding drugs fed to *Drosophila melanogaster*. *Mol Cell* **6:** 1013–1024

Janssen S, Durussel T, Laemmli UK (2000b) Chromatin opening of DNA satellites by targeted sequence-specific drugs. *Mol Cell* **6:** 999–1011

Käs E, Laemmli UK (1992) *In vivo* topoisomerase II cleavage of the *Drosophila* histone and satellite III repeats: DNA sequence and structural characteristics. *EMBO J* **11:** 705–716

Le HD, Donaldson KM, Cook KR, Karpen GH (2004) A high proportion of genes involved in position effect variegation also affect chromosome inheritance. *Chromosoma* **112:** 269–276

Lohe AR, Hilliker AJ, Roberts PA (1993) Mapping simple repeated DNA sequences in heterochromatin of *Drosophila melanogaster*. *Genetics* **134:** 1149–1174

Losada A, Villasante A (1996) Autosomal location of a new subtype of 1*688 satellite DNA of *Drosophila melanogaster*. *Chromosome Res* **4:** 372–383

Lu BY, Bishop CP, Eissenberg JC (1996) Developmental timing and tissue specificity of heterochromatin-mediated silencing. *EMBO J* **15:** 1323–1332

Lupo R, Breiling A, Bianchi ME, Orlando V (2001) *Drosophila* chromosome condensation proteins Topoisomerase II and Barren colocalize with Polycomb and maintain Fab-7 PRE silencing. *Mol Cell* **7:** 127–136

Maeshima K, Laemmli UK (2003) A two-step scaffolding model for mitotic chromosome assembly. *Dev Cell* **4:** 467–480

Monod C, Aulner N, Cuvier O, Käs E (2002) Modification of position-effect variegation by competition for binding to *Drosophila* satellites. *EMBO Rep* **3:** 747–752

Pal-Bhadra M, Leibovitch BA, Gandhi SG, Rao M, Bhadra U, Birchler JA, Elgin SC (2004) Heterochromatic silencing and HP1 localization in *Drosophila* are dependent on the RNAi machinery. *Science* **303:** 669–672

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:** 467–476

Platero JS, Csink AK, Quintanilla A, Henikoff S (1998) Changes in chromosomal localization of heterochromatin-binding proteins during the cell cycle in *Drosophila*. *J Cell Biol* **140:** 1297–1306

Reeves R, Nissen MS (1990) The A·T-DNA-binding domain of mammalian high mobility group I chromosomal proteins. *J Biol Chem* **265:** 8573–8582

Santoro R, Li J, Grummt I (2002) The nucleolar remodeling complex NoRC mediates heterochromatin formation and silencing of ribosomal gene transcription. *Nat Genet* **32:** 393–396

Schotta G, Ebert A, Dorn R, Reuter G (2003) Position-effect variegation and the genetic dissection of chromatin regulation in *Drosophila*. *Semin Cell Dev Biol* **14:** 67–75

Shaffer CD, Stephens GE, Thompson BA, Funches L, Bernat JA, Craig CA, Elgin SC (2002) Heterochromatin protein 2 (HP2), a partner of HP1 in *Drosophila* heterochromatin. *Proc Natl Acad Sci USA* **99:** 14332–14337

Slama-Schwok A, Zakrzewska K, Leger G, Leroux Y, Takahashi M, Käs E, Debey P (2000) Structural changes induced by binding of the high-mobility group I protein to a mouse satellite DNA sequence. *Biophys J* **78:** 2543–2559

Susbielle G, Blattes R, Brevet V, Monod C, Käs E (2005) Target practice: aiming at satellite repeats with DNA minor groove binders. *Curr Med Chem Anti-canc Agents* **5:** 409–420

Swedlow JR, Sedat JW, Agard DA (1993) Multiple chromosomal populations of topoisomerase II detected *in vivo* by time-lapse, three-dimensional wide-field microscopy. *Cell* **73:** 97–108

Taddei A, Maison C, Roche D, Almouzni G (2001) Reversible disruption of pericentric heterochromatin and centromere function by inhibiting deacetylases. *Nat Cell Biol* **3:** 114–120

Wustmann G, Szidonya J, Taubert H, Reuter G (1989) The genetics of position-effect variegation modifying loci in *Drosophila melanogaster*. *Mol Gen Genet* **217:** 520–527

Zhao T, Heyduk T, Allis CD, Eissenberg JC (2000) Heterochromatin protein 1 binds to nucleosomes and DNA *in vitro*. *J Biol Chem* **275:** 28332–28338

 ©2006 European Molecular Biology Organization

4. *Conclusions et discussions*

Après avoir montré que le polyamide P9, tout comme la protéine D1, se liait spécifiquement aux séquences SAT III, des expériences *in vitro* m'ont permis de caractériser les sites de liaison de ces composés sur ces séquences. J'ai ensuite réalisé des expériences de localisation des séquences satellites de la drosophile sur des chromosomes mitotiques en utilisant des dérivés fluorescents des polyamides P9 et P31, respectivement le P9F couplé à la Fluorescéine et le P31TR couplé au Texas Red.

Des expériences *in vivo* et l'utilisation d'une nouvelle technique de traitement *ex vivo* par les drogues m'a permis de démontrer que le polyamide P9 était capable de déplacer la protéine D1 des séquences SAT III et ce, suite à une compétition directe entre ces deux composants pour la liaison à ces séquences.

Le déplacement de la protéine D1 des séquences SAT III a des conséquences directes sur la localisation de la protéine HP1. Après traitement avec le polyamide P9, la protéine D1 est déplacée des séquences SAT I et SAT III, tout comme la protéine HP1. Cependant, après traitement avec le polyamide P9, la localisation de la protéine HP1 n'est pas affectée au niveau du rDNA, adjacent aux séquences SAT III, et des séquences SAT V. Par contre, lors d'un traitement des cellules par le polyamide P31, la protéine HP1 n'est délocalisée que des séquences SAT V, cibles de ce composé, et reste présente sur les autres séquences satellites non ciblées.

Des expériences de traitement à la Trichostatine A (TSA), un inhibiteur des HDAC, ont permis de confirmer l'interdépendance existant entre la localisation des protéines D1 et HP1. Suite au déplacement de la protéine D1 des séquences SAT III, la protéine HP1 reste accumulée au niveau du rDNA, alors qu'après traitement à la TSA, la protéine HP1 est délocalisée de l'hétérochromatine vers le nucléoplasme et est remplacée par la protéine D1 au niveau des séquences SAT III et du rDNA. Cette invasion du rDNA par la protéine D1 suggère la possibilité d'interactions mutuellement exclusives entre ces séquences et les protéines HP1 ou D1.

Dans le cas de la suppression de la variégation d'effet de position par le polyamide P9, il semblerait donc que la perte de la protéine HP1 des séquences SAT III soit responsable de

ce phénotype. Pourtant, le gène *white* dans la lignée *white-mottled* et les séquences SAT III sont séparés par le rDNA qui s'étend sur plus de 3,5 Mpb. Cette constatation suggère un mécanisme spatial : le rDNA étant engagé dans la formation du nucléole dans les cellules en interphase, les séquences SAT III et le gène *white* peuvent donc être spatialement rapprochés. Ces observations ne sont donc pas en accord avec le modèle d'expansion linéaire de l'hétérochromatine mais plutôt avec un modèle spatial, déjà évoqué dans le cas du mécanisme de mise en silence par les protéines du groupe Polycomb (Bantignies et al., 2003) ou dans la lignée *brown-dominant* (Henikoff et al., 1995). Même si ce modèle séduisant est plausible, il ne résiste pas à une analyse détaillée qui suggère l'existence d'un autre mécanisme de régulation à distance mettant en jeu des petits ARN (partie développée page suivante).

Les mêmes expériences de traitement par le polyamide P9 ont été reproduites mais cette fois, en recherchant les effets de ce traitement sur la topoïsomérase II. J'ai pu démontrer que tout comme la protéine D1, la topoïsomérase II était déplacée des séquences SAT III. Mais cette perte pouvait être soit dépendante de la perte de la protéine D1, soit indépendante. En effet, la protéine D1 pourrait interagir *in vivo* avec la topoïsomérase II mais cette dernière peut aussi interagir spécifiquement avec les séquences SAT III (Käs and Laemmli, 1992).

Ces résultats de déplacement de la topoïsomérase II par le polyamide P9 nous ont amenés à envisager des traitements de mouches de la lignée *white-mottled* avec des inhibiteurs de la topoïsomérase II afin de rechercher un rôle éventuel de cette protéine dans la mise en place de la structure hétérochromatinienne. Un traitement avec le VM26, un poison de la topoïsomérase II, a permis de conclure à l'implication de la topoïsomérase II dans le mécanisme d'assemblage de la structure hétérochromatinienne. Dans la plupart des organismes, la topoïsomérase II est souvent associée à des régions péricentriques hautement compactées suggérant donc un rôle de cette protéine dans la modulation de la structure hétérochromatinienne mais aussi du chromosome lui-même (Maeshima and Laemmli, 2003; Porter and Farr, 2004; Swedlow et al., 1993). De même, il a été montré que la topoïsomérase II associée à la protéine Barren, un membre du complexe *Condensin*, participait à la répression transcriptionnelle médiée par le complexe protéine Polycomb (Lupo et al., 2001).

Tous ces résultats suggèrent que le polyamide P9, tout comme le polyamide P31, agissent en générant une ouverture de la chromatine par déplacement sélectif de protéines qui lui sont associées. Dans le cas du polyamide P31, le déplacement de la protéine HP1 des

séquences SAT V pourrait permettre la fixation de la protéine GAF. En effet, des résultats préliminaires obtenus au laboratoire tendent à démontrer que sur les chromosomes mitotiques, la protéine HP1 est remplacée par la protéine GAF sur les séquences SAT V au fur et à mesure de la progression à travers la mitose.

Grâce à l'utilisation de P9, l'hypothèse de l'implication de la protéine D1 dans l'initiation de l'assemblage de l'hétérochromatine a été fortement appuyée. La protéine D1 serait donc capable de guider les protéines HP1 et topoïsomérase II vers les lieux d'assemblage de l'hétérochromatine.

L'accumulation de la topoïsomérase II au niveau des séquences SAT III avait déjà été décrite (Käs and Laemmli, 1992; Marshall et al., 1997; Swedlow et al., 1993) mais son mode de recrutement ainsi que sa fonction au sein de ces séquences restaient à définir. Les premiers résultats obtenus impliquent la protéine D1 dans le recrutement de la topoïsomérase II sur les séquences SAT III. Pourtant, il a été établi que la topoïsomérase II possédait des sites de liaison spécifiques au sein des séquences SAT III de la drosophile (Poljak and Käs, 1995). Il n'avait pas été décrit que l'activité enzymatique de la topoïsomérase II soit nécessaire au mécanisme d'assemblage de la structure hétérochromatinienne.

Ainsi, la protéine D1 pourrait refléter le rôle des protéines HMG-A au niveau des séquences satellites alphoïdes des eucaryotes supérieurs. Les protéines MATH joueraient dont un rôle de plateforme de recrutement pour les protéines de formation de la structure hétérochromatinienne. Cette hypothèse reste à tester.

D'autre part, l'indépendance de la localisation de la protéine HP1 vis-à-vis de la présence de la protéine D1 au niveau du rDNA suggère l'intervention d'un autre mécanisme de recrutement de cette protéine. La régulation de la transcription des gènes ribosomiques fait intervenir dans de nombreux organismes des mécanismes épigénétiques tels que la méthylation de la lysine 9 de l'histone H3. Il est possible qu'au niveau des ces régions, d'autres protéines soient responsables du recrutement de la protéine HP1, ou que d'autres mécanismes interviennent. Cette localisation pourrait être dépendante de la machinerie d'interférence à ARN, qui dans certains organismes, tels que chez la levure à fission *Schizosaccharomyces pombe*, est couplée à la mise en silence par une structure hétérochromatinienne (Volpe et al., 2002; Zamore, 2001). Cependant, ce mécanisme ne semble pas intervenir chez la drosophile ((Blattes et al., 2006) et résultats personnels), du moins pour le modèle des séquences SAT III, même si des petits ARN provenant de ces

séquences ont été décrits (Aravin et al., 2003). D'autre part, le recrutement de la protéine HP1 sur les séquences SAT III pourrait faire intervenir un mécanisme de liaison directe de cette protéine à la chromatine (Nielsen et al., 2001a). Cette possibilité, même si elle a été déjà décrite, reste cependant à démontrer au niveau des séquences SAT III. Une autre hypothèse est celle d'une interaction directe entre l'ADN et la protéine HP1 (Perrini et al., 2004; Zhao et al., 2000), confortée par nos observations, mais qui reste à démontrer directement.

Enfin, j'ai mis en évidence une interaction mutuellement exclusive des protéines D1 et HP1 avec les séquences SAT III et le rDNA. Cette observation, décrite plus bas, était déjà suggérée par les résultats de traitement par la TSA, montrant que la délocalisation de HP1 du rDNA était accompagnée d'une invasion de ce locus par la protéine D1. J'ai donc étendu une analyse à l'étude de la localisation de ces deux protéines dans le contexte de la lignée *white-mottled*, et avec ou sans traitement avec le polyamide P9. Ces expériences, que j'ai réalisées par la suite (voir ci-dessous), nous ont permis de mieux comprendre le phénomène de la variégation d'effet de position d'un point de vue biochimique, mais aussi le mécanisme mis en jeu dans la suppression de ce phénotype par la molécule P9.

Une autre hypothèse est soulevée par ces résultats. La distinction entre les domaines de localisation de la protéine HP1 dépendante, ou indépendante, de la protéine D1 est nette. Cette observation suggère l'existence d'élément frontières cloisonnant le rDNA en le plaçant ainsi hors du contexte environnant, c'est-à-dire hétérochromatinien du côté des séquences SAT III et euchromatinien de l'autre côté. L'existence d'une telle frontière entre hétérochromatine, regroupant les séquences SAT III et rDNA, et euchromatine avait déjà été prédite il y a plus de 60 ans (Schultz, 1939). Une séquence reconnue pour son activité de frontière, l'élément fontière BE28, avait d'ailleurs été localisée sous forme de cluster dans cette région de jonction euchromatine/hétérochromatine, mais sa localisation fine et précise n'avait pu être confirmée (Cuvier et al., 2002). Dans des travaux postérieurs, je me suis attachée à rechercher la localisation précise de cet élément frontière, non plus sur des chromosomes polytènes mais sur des chromosomes mitotiques (voir chapitre IV de cette même partie). Par ailleurs, un autre candidat est également présent dans cette région, mais ses propriétés de « frontière » n'ont pas été établies (O'Hare et al., 2002).

III. Du rôle de la protéine topoïsomérase II dans l'assemblage de l'hétérochromatine : « Topoisomerase II and minor groove binders act coordinately to form specialized chromatin structures »

1. Introduction

La topoïsomérase II, une protéine essentielle, est impliquée dans le contrôle de la topologie de l'ADN. Elle est capable de relâcher les super-tours d'ADN dans les étapes de caténation/décaténation à la fois de la chromatine (Cortes et al., 2003) ou des kinétoplastes (Bailly et al., 1999). Ces structures, qui sont présentes chez les trypanosomes, correspondent à l'entrecroisement de cercles d'ADN qui portent les gènes codant les composants ribosomiques (ARN ribosomiques et protéines ribosomiques) et des protéines mitochondriales. L'activité de la topoïsomérase II est donc essentielle à la résolution et à la réplication de ces molécules, comme c'est aussi le cas dans les cellules humaines pour les chromatides sœurs.

L'activité de la topoïsomérase II est la cible de nombreuses drogues utilisées dans les thérapies anticancéreuses. Afin d'étudier les mécanismes d'action de certaines molécules, ici les diamidines à visée antiparasitaire, nous les avons utilisées dans notre organisme modèle de la drosophile.

Des études précédentes avaient révélé une forte accumulation de la topoïsomérase II au niveau des séquences SAT III du chromosome X de la drosophile (Blattes et al., 2006). Cette localisation n'est pas sans rappeler l'accumulation de cette protéine sur les séquences satellites alphoïdes péricentriques des mammifères (Maeshima and Laemmli, 2003).

Suite à des traitements avec la diamidine DB351, un dérive du DB75, capable d'agir sur la localisation de la protéine topoïsomérase II, nous avons démontré que la variégation d'effet de position de l'allèle *white-mottled* était supprimée, suggérant un rôle de la topoïsomérase II dans le mécanisme d'assemblage de la structure hétérochromatinienne (Susbielle et al., 2005). Pour la première fois, la topoïsomérase II a été identifiée comme un modificateur possible de la PEV alors qu'auparavant, des approches génétiques visant à déléter une copie du gène codant cette protéine n'avait pas montré pas un tel effet (Wustmann

et al., 1989). L'utilisation du polyamide P9 nous a d'ores et déjà permis de démontrer le rôle de la protéine D1 dans la localisation de la protéine HP1 au niveau des séquences SAT III de l'hétérochromatine péricentromérique du chromosome X (Blattes et al., 2006). La localisation de la protéine topoïsomérase II sur ces mêmes séquences semble elle aussi être dépendante de la protéine D1 (Blattes et al., 2006).

Ce résultat est cependant à modérer car la topoïsomérase II est connue pour interagir directement avec l'ADN riche en paires de bases dA•dT au niveau de séquences appelées MAR/SAR. La topoïsomérase II est capable de cliver ces séquences ainsi que les séquences SAT III au niveau de séquences consensus différentes de celles coupées *in vitro* (Käs and Laemmli, 1992).

En comparant ces données à celles obtenues dans le cadre d'expériences mettant en jeu le polyamide P9, un mécanisme d'action différent pour ces deux ligands du petit sillon semble plausible. De même, il est possible d'envisager un mécanisme de recrutement de la topoïsomérase II sur les séquences SAT III mettant en jeu deux voies différentes, dépendante ou indépendante de la protéine D1, mais cependant complémentaires.

2. Objectifs de l'étude

L'objectif de ce travail a été de définir le rôle de la topoïsomérase II dans l'assemblage de la structure hétérochromatinienne du chromosome X de la drosophile. Nous avons aussi abordé le rôle de la protéine D1 dans le mécanisme de recrutement de la protéine topoïsomérase II sur les séquences SAT III. Grâce à l'utilisation d'un mutant dominant négatif de la protéine D1, le mutant D1ΔE délété du domaine C-terminal, nous avons cherché à déterminer l'importance du rôle de la protéine D1 dans le recrutement de la topoïsomérase II au niveau des séquences SAT III, mais aussi au niveau d'autres séquences où la protéine D1 est localisée.

Nous avons ensuite recherché si la présence de la protéine D1 était nécessaire à l'activité de clivage de la topoïsomérase II sur les séquences SAT III, et si cette activité était requise pour l'assemblage de la structure hétérochromatinienne.

Enfin, nous avons étendu ces études aux effets des diamidines sur la localisation de la topoïsomérase II, afin d'éclaircir le mécanisme d'action de ces drogues lors des traitements antiparasitaires. De même, nous avons vérifié que la relocalisation de la topoïsomérase II sur les séquences SAT I et SAT III n'affectait pas l'activité catalytique de celle-ci.

3. Résultats

Publication soumise à « EMBO Journal » en septembre 2006.

Topoisomerase II and minor-groove binders act coordinately to form specialized chromatin structures

Roxane Blattes[1], Guillaume Susbielle[1], Emily J. Kuhn[2], Nathalie Aulner[1], Pamela K. Geyer[2], David W. Boykin[3], W. David Wilson[3], Amanda M. Mathis[4], James Edwin Hall[4], Emmanuel Käs[1†] and Caroline Monod[1†]

[1]Laboratoire de Biologie Moléculaire Eucaryote, CNRS/Université Paul Sabatier, IFR 109, 118 route de Narbonne, 31062 Toulouse Cedex 9, France

[2]Department of Biochemistry, University of Iowa, 3135 MERF Iowa City, IA 52242-1109, U.S.A.

[3]Department of Chemistry, Georgia State University, Atlanta, GA 30303, U.S.A

[4]Division of Molecular Pharmaceutics, School of Pharmacy, University of North Carolina, Chapel Hill, NC 27599, U.S.A.

[†] Corresponding authors

LBME - UMR 5099, 118 route de Narbonne, 31062 Toulouse Cedex 9, France

phone: +33 (0)5 61 33 59 59

fax: +33 (0)5 61 33 58 86

e-mail: monod@ibcg.biotoul.fr, kas@ibcg.biotoul.fr

54 803 characters

ABSTRACT

AT-rich satellites are Mbp-sized targets for synthetic DNA minor-groove binders whose remarkably specific binding to pericentric heterochromatin alters its functional properties by inhibiting or promoting satellite/protein interactions. The P9 oligopyrrole targets the *Drosophila* D1 protein associated with satellite III (SAT III) repeats in X-chromosome centric heterochromatin, perturbing a D1-dependent pathway that guides topoisomerase II (topo II) to the SAT III array. In contrast, antiparasitic diamidines uncouple the topo II/D1 interaction and cause a massive D1-independent relocalization of topoisomerase II to AT-rich heterochromatin. This striking mobilization results in suppression of heterochromatin-mediated silencing. Diamidines target the highly biased A+T content of pathogenic parasitic microorganisms, or portions thereof as in the kinetoplast DNA of trypanosomes. Their demonstrated effects in the *Drosophila* system provide keys to understanding their mechanism of biological action, which most likely reflects an anti-genome activity that interferes with essential protein/DNA complexes involving AT-rich sequences. We propose that synthetic and natural AT-specific minor-groove binders act coordinately with topo II in the assembly of specialized nucleoprotein structures, of which heterochromatin constitutes an extreme example.

Key words: D1/Topoisomerase II/Polyamides/Diamidines/Minor-groove binding

Running title: Topoisomerase II and heterochromatin assembly

INTRODUCTION

Topoisomerase II (topo II) is an essential enzyme of cell proliferation and an important target for several anti-cancer drugs (see (Wang, 2002) for review). Topo II relaxes positively and negatively supercoiled DNA and mediates DNA strand passage during catenation and decatenation reactions. Topo II is required to resolve sister chromatids (DiNardo et al, 1984), an activity which also plays a role in the decatenation of replicated kinetoplast circles of trypanosomes (Wang & Englund, 2001). A number of observations suggest that this enzyme may play, in addition to its enzymatic function in chromosome dynamics, a role in the structural maintenance of chromosomes (Maeshima & Laemmli, 2003).

In this paper, we study by high-resolution immunofluorescence the localization of topo II in *Drosophila melanogaster* mitotic chromosomes, with an emphasis on its association with centric satellite regions and on the factors that affect this preferential interaction. We previously showed that satellite-specific synthetic polyamides (Janssen et al, 2000a) selectively perturb the protein composition of satellites sequences. The P9 oligopyrrole, which targets AT-rich satellites, selectively displaces topo II from SAT III, the major satellite repeat located on the *Drosophila melanogaster* X chromosome. We established that P9 also displaces D1 and Heterochromatin Protein 1 (HP1) from SAT III repeats (Blattes et al, 2006). The functional significance of these multiple interactions was also assessed genetically using *white-mottled* (w^{m4}) flies as a model system. P9 suppresses the stochastic silencing phenomenon observed in w^{m4} flies, where the *white* gene become juxtaposed to the rDNA and SAT III repeats of the X chromosome (Janssen et al, 2000b). These combined observations suggested that suppression of the silenced phenotype is likely caused by the displacement of the SAT III-associated proteins D1, HP1 and topo II. This gain-of-function phenotype was

shown to implicate topo II activity and/or localization as VM26, a topo II poison, suppressed w^{m4} PEV when fed to developing larvae (Blattes et al, 2006).

Here, we present results demonstrating that the localization of topo II to SAT III repeats requires the presence of the D1 protein and of its C-terminal domain. This interaction can be bypassed by treatment with synthetic antiparasitic diamidines that target AT-rich sequences (Tidwell & Boykin, 2003; Susbielle et al, 2005; Wilson et al, 2005). We show that, like P9, the DB75 diamidine displaces D1 from SAT III repeats. However, in contrast to polyamide-induced effects, DB75 induces a massive relocalization of topo II from euchromatin to AT-rich satellite heterochromatin. This mobilization significantly increases VM26-induced topo II cleavage of SAT III repeats and results in suppression of w^{m4} PEV, demonstrating a structural and/or functional role for topo II in heterochromatin-mediated silencing.

Our results also implicate interactions between topo II and AT-rich genomes or portions thereof in the antiparasitic activity of diamidines. DB75 localizes selectively to the AT-rich mitochondrial kinetoplast DNA (kDNA) of kinetoplastid parasites (Wilson et al, 2005; Mathis et al, 2006). The thousands of catenated minicircle DNAs in the mitochondrial genome have extensive, closely-spaced, phased dA•dT sequences that yield curved DNA structures and serve as preferred target sites for diamidines. Solution studies of diamidine/DNA complexes suggest that targeting of both the AT-rich regions of SAT III and the parasite kDNA must involve similar minor groove complexes. A parasite topo II is critical for decatenation, replication and reinsertion of the minicircles into the kinetoplast DNA matrix (Wang & Englund, 2001). We propose that the mechanism of antiparasitic action of diamidines in these repeated dA•dT sequences is related to their capacity to mobilize topo II to SAT III repeats.

RESULTS

The 11-Mbp SAT III array located in the pericentric region of the *Drosophila* X chromosome is associated with several non-histone proteins, including D1, which specifically binds this satellite sequence via interactions with the 10 AT hooks found in the protein, and HP1 (Blattes et al, 2006). In addition to these proteins, which modify w^{m4} PEV (see (Huisinga et al, 2006) for review), SAT III repeats were also shown several years ago to be associated with topo II (Käs & Laemmli, 1992). Experiments described below aim to characterize the latter interaction in more detail.

Fluorescent polyamides specific for the major *Drosophila* satellites (Janssen et al, 2000a) can be used to map the localization of heterochromatin-associated proteins relative to these highly repeated sequences with a high degree of precision. This is illustrated in Figure 1A, which shows staining of mitotic chromosomes from diploid larval neuroblasts with P9F (green) and P31TR (red). P9F specifically targets the major AT-rich satellites of *Drosophila*, the short satellite I (SAT I) $AATAT_n$ repeats found predominantly on the fourth chromosome as well as on chromosome 3 and at the very tip of the X chromosome, and the 11-Mbp SAT III array at the proximal end of the X chromosome. P31TR targets the $AAGAG_n$ satellite V (SAT V) repeats that localize predominantly to chromosome 2 as well as to chromosome 4 and the tip of the X chromosome.

The focus of the studies presented here is the pericentromeric region of the X chromosome, where SAT III is the largest block of satellite heterochromatin, immediately adjacent to the 4-Mbp rDNA locus. SAT III constitutes the major *in vivo* binding site for the D1 protein (Figure 1B), which also binds SAT I repeats. Experiments using the P9 oligopyrrole have shown that the perturbation of SAT III/D1 interactions correlates with the suppression of w^{m4} PEV (Janssen et al, 2000a; Blattes et al, 2006). Treatment of *Drosophila*

permabilized larval tissues with P9, a synthetic mimic of the D1 AT-hook motifs, results in the displacement of D1 and topo II from SAT III repeats. This is illustrated in panels C and D of Figure 1, which show displacement of D1 from SAT III (but not SAT I) repeats following treatment with a low concentration of P9 (250 mM). A similar effect can be observed for topo II (panels E and F) although, in this case, the protein is not displaced from the adjacent rDNA locus that is not a target for P9. This displacement was shown to correlate with suppression of w^{m4} PEV by P9, an observation which suggests that topo II is implicated in heterochromatin-mediated silencing as further supported by the effects of topo II poisons on the *white-mottled* phenotype (see below).

Localization of topo to SAT III repeats requires the C-terminal domain of D1

Topo II, like D1 and possibly HP1 as well (Blattes et al, 2006), might be delocalized by P9 as a result of direct competition, a possibility supported by the known preferential association of the protein with AT-rich sequences (Saitoh & Laemmli, 1994). Alternatively, the displacement of topo II from SAT III sequences might be caused by the P9-induced loss of D1. D1ΔE is a dominant negative mutant form of D1 which lacks the acidic C-terminal domain of D1 and acts as a suppressor of w^{m4} PEV (Aulner et al, 2002). We proposed that the C-terminal domain of D1 is required for the recruitment of proteins implicated in heterochromatin assembly (Monod et al, 2002). Hence, binding of D1ΔE to SAT III repeats could lead to a reduced association with topo II if D1 serves to direct its association with these sequences.

We tested this possibility by assessing the localization of topo II in larvae expressing D1ΔE under the control of a heat-shock promoter (HS-D1ΔE). Following induction of the transgene, which is expressed at about 3-fold greater levels than the endogenous gene (data not shown), the D1 antibody, which recognizes both D1 and D1ΔE, detected a diffuse staining

pattern in nuclei which most likely reflects displacement of endogenous D1 and/or accumulation of the mutant protein in the nucleoplasm (Figure 2A). The D1 localization pattern observed on mitotic chromosomes was relatively unchanged relative to non heat-shocked flies (see inset). Expression of D1ΔE led to the delocalization of topo II from SAT III repeats, without affecting adjacent regions (panels B and C). Counterstaining with P9F (not shown) confirmed the loss of topo II signal from the SAT III array, but not from adjacent rDNA repeats. These changes were not observed in heat-shocked wild-type larvae or in larvae from a line expressing a heat-inducible D1 transgene (Aulner et al, 2002). In the latter case, over-expression of full-length D1 (also at about 3-fold greater levels than those of the endogenous gene) did not result in a displacement of topo II from SAT III repeats (data not shown). Taken together, these results indicate that topo II requires the presence of full-length D1 for binding to SAT III repeats and we conclude that the selective loss of topo II from this 11-Mbp satellite array in the presence of P9 is a secondary effect of the drug-induced displacement of D1.

Topoisomerase II is a companion of D1 on polytene chromosomes

Results shown above suggest a coordinate interaction between D1 and topo II which we further analyzed under conditions where the localization of D1 is not restricted to AT-rich satellite repeats. In polytene chromosomes, the under-replication of satellite sequences might reveal a more widespread interaction of D1 with endo-replicated euchromatic sites. As shown in Figure 2D, staining of polytene chromosomes for D1 indeed revealed a large number of positive sites localized to DAPI-bright bands, in addition to prominent staining of the chromocenter. Is topo II generally associated with the D1 sites? To answer this question, we performed double-labeling experiments which revealed extensive overlap between the D1 and topo II signals (Figure 2E). A notable exception to this co-localization is the chromocenter,

which stained weakly, if at all, for topo II, in line with the general exclusion of topo II from DAPI-bright regions of diploid nuclei (Figure 1E). The split image of an enlarged chromosome arm (bottom panel of Figure 2E) shows a one-on-one correspondence between many topo II and D1 bands. Note, however, that there are more topo II sites than D1 sites on polytene chromosomes; although we estimate overall that greater than 90% of D1 sites correspond to topo II sites, there exists a sub-population of topo II sites that are not associated with D1. These sites appear in red in the merged image and might represent a population of the enzyme that interacts with chromatin in a D1-independent fashion, as in the case of the rDNA repeats adjacent to SAT III sequences in diploid cells (Figure 1E).

SAT III interacts selectively with catalytically active topoisomerase II

The immunolocalization data shown above demonstrate a selective binding of topo II to the 359-bp SAT III repeats that requires the presence of D1 and might more generally extend to other sites where the two proteins co-localize. Does this ternary interaction implicate catalytically active topo II molecules? Topo II cleavage of SAT III sequences can be induced by treatment of *Drosophila* tissue-culture cells with VM26, a topo II poison (Käs & Laemmli, 1992). VM26 traps a "cleavable complex" in which the transient topo II-induced double-stranded DNA breaks can be revealed by subsequent protein denaturation and degradation. As topo II cleavage of SAT III has only been evidenced in cultured cells, we tested whether treatment of whole developing embryos with VM26 would similarly induce topo II cleavage in SAT III repeats.

Figure 3 shows a time course of VM26 treatment of whole cellular blastoderm embryos. Topo II cleaves the 359-bp sequence once at a specific site located in one of the two nucleosomal linkers accommodated by each repeat, resulting in the appearance of a 359-bp ladder. Topo II cleavage of the SAT III array was detected following a treatment as short as

10 minutes (panel A, lane 2), reaching a maximum at 20 minutes (lane 3) and decreasing thereafter (lanes 4-6). The detection of a monomer cleavage product demonstrates that adjacent repeats are associated with active topo II molecules while the fast kinetics of cleavage suggest the presence of highly active molecules spanning the entire SAT III array, in line with our immunolocalization results. The decrease in the intensity of the cleavage ladder following VM26 treatment times exceeding 20 minutes is consistent with the high degree of reversibility of topo II cleavable complexes stabilized by the drug measured in cultured cells in this region of the *Drosophila* genome (Borgnetto et al, 1996). No VM26-induced cleavage products were detected in $(AATAT)_n$ SAT I repeats (Figure 3B), suggesting that topo II cleavage of SAT III on the X chromosome is a characteristic property of this sequences, the only alpha-like satellite found in *Drosophila*. This is consistent with our immunolocalization results which do not show detectable topo II on SAT I repeats (see Figure 1E and, in particular, the lack of staining of chromosome 4).

D1 and Topoisomerase II cooperate to effect specific cleavage of SAT III repeats

The association of D1 and active topo II with SAT III might reflect a coordinate interaction important for heterochromatin assembly. We show next that D1 strongly influences topo II activity on SAT III repeats and determines the highly specific cleavage observed *in vivo*. Topo II cleaves each SAT III repeat *in vivo* at a short GC-rich sequence embedded in this otherwise AT-rich satellite (Käs & Laemmli, 1992; Borgnetto et al, 1996). However, neither D1 nor P9 bind to this cleavage site *in vitro* and induce instead the formation of a DNAse I-hypersensitive site in this region (Blattes et al, 2006). Such an enhanced accessibility to DNAse I might extend to topo II as well. To test this possibility, we performed *in vitro* cleavage reactions using purified topo II and D1 protein. As shown in Figure 4, topo II cleaves SAT III at numerous sites in naked DNA (lanes 1 and 5), which do

not map to the sequence cleaved *in vivo* (arrowheads). However, addition of D1 to the cleavage reaction had a striking effect on the distribution of cleavage products, precisely guiding cleavage to a single region that corresponds exactly to the *in vivo* site (lanes 2-4). Labeling of the other strand confirmed this effect (data not shown). These results demonstrate that D1 binding to SAT III repeats indeed promotes cleavage by topo II at the sequence recognized by the enzyme *in vivo*. In this case, minor-groove binding creates a highly accessible topo II cleavage site, comparable to the DNAse I hypersensitivity induced by D1 or synthetic ligands discussed next.

Molecules that possess the planar crescent structure characteristic of minor-groove binders such as P9 or AT hook motifs produce identical effects. The distamycin A oligopeptide is the archetype for these ligands and was shown several years ago to enhance topo II cleavage of SAT III repeats in cultured cells (Käs et al, 1993). Distamycin mimics the effect of D1 on topo II cleavage of the SAT III sequence (Figure 4, lanes 6-8), while chromomycin, a GC-specific ligand which strongly inhibits SAT III cleavage *in vivo* (Käs & Laemmli, 1992) has the opposite effect, enhancing the pattern observed in the absence of drug and masking the *in vivo* cleavage site (lanes 9 and 10). These effects are due to drug binding at AT- or GC-rich sites, as demonstrated by the results of DNAse I footprinting experiments performed in the absence or presence of these drugs (lanes 11-13). Binding by distamycin induces protection of many sites in the SAT III sequence, in addition to a strong hypersensitive site mapping to the topo II cleavage site (lane 12), an effect similar to that of P9 and D1 (Blattes et al, 2006). This site indeed corresponds to a sequence protected from DNAse I digestion by chromomycin (lane 13), in line with the general reversal of protected regions relative to distamycin (compare lanes 12 and 13).

Antiparasitic diamidines uncouple the co-localization of D1 and Topoisomerase II and redistribute the enzyme to AT-rich satellites

The striking ability of both natural and synthetic minor-groove binders to affect the localization and activity of topo II prompted us to test additional compounds. Heterocyclic diamidines are an attractive class of DNA targeting agents to use as additional probes for the function of topo II in chromatin. As their structure is quite different from that of polyamide cations such as P9 or distamycin, they may provide very useful additional test compounds for the investigation of topo II effects on chromatin. Diamidines are readily taken up by both human and parasite eukaryotic cells and a prodrug of a diphenylfuran diamidine, DB75, is in phase III clinical trials against sleeping sickness (Lansiaux et al, 2002; Tidwell & Boykin, 2003; Wilson et al, 2005). DB75 (structure in Figure 5) binds in the DNA minor groove at dA•dT sequences of four or more base pairs and the amidine groups form hydrogen bonds with T bases at the floor of the groove (Wilson et al, 2005). DB293 (structure in Figure 5) has one of the phenyl groups of DB75 replaced with a benzimidazole. It binds with similar affinity to AT sequences as DB75 but it can also bind as a dimer to certain dG•dC-containing sequences such as ATGA (Wang et al, 2000) and may thus have different topological effects on DNA.

The antiparasitic properties of diamidines have been difficult to ascertain in the pathogenic microorganisms they target. Appropriate experimental assays are generally lacking and we used the *Drosophila* system instead to understand their possible mechanism of action. This is possible as DNA-bound diamidines have auto-fluorescence properties and their preferential cellular interaction sites can be visualized directly. This is illustrated in Figure 5A, which shows permeabilized larval neuroblasts treated with DB75 (blue) and counterstained with P31TR (red). DB75 strongly fluoresces when bound to AT sequences in DNA and preferential binding to AT-rich DNA is inferred from bright staining of X-

chromosome SAT III repeats and of chromosome 4, while euchromatic chromosomes arms are weakly stained. We performed a similar test on the localization of DB293, which also stained AT-rich sequences, resolved from AG-rich sequences stained by P31TR (Figure 5B). As described above, DB293 also has a dimer binding mode that can selectively recognize some dG•dC-containing sequences. As might be expected from this broader sequence binding ability, DB293 exhibited less specificity in chromatin as judged from the more prominent staining of euchromatic chromosome arms.

DB351 is closely related to DB75, from which it differs by the substitution of an oxygen atom (central furan in DB75) by a sulfur (central thiophene in DB351). DB351 was shown to target topo II interactions with chromatin and drug feeding results in suppression of *white* PEV in adult w^{m4} flies (Susbielle et al, 2005). In these experiments, DB351 was used instead of DB75 because whole *Drosophila* tissues appear relatively impermeable to the latter. Suppression of PEV is most likely mediated by the preferential interaction of DB351 with SAT III repeats, as inferred from the binding ptoperties of DB75 to chromatin of permabilized larval brains (Figure 5A), with SAT III repeats. We tested this hypothesis by examining the effect of DB75 on the localization of D1 and topo II following treatment of permeabilized larval neuroblasts. Treated larval brains were immunostained for D1 (red) and the DB75 signal (blue) was simultaneously detected. This resulted in a selective displacement of D1 from SAT III repeats (Figure 5C), but not from SAT I sequences on chromosomes 4 and 3 and at the very tip of the X chromosome, an effect identical to that seen following treatment with P9 (compare to Figure 1D). In contrast, treatment with DB293, which exhibits a lower selectivity for AT-rich sequences, had no effect on the localization of D1 (Figure 5D). We conclude from these results that the DB75 antiparasitic drug acts similarly to the P9 oligopyrrole on D1 displacement. The lack of displacement by DB293 may be due to its

broader distribution in chromatin and lower concentration in AT sequences and/or to a different topological effect on DNA that is compatible with D1 binding.

Does DB75 affects topo II localization as well? Figure 5E shows results of drug treatments followed by topo II immunostaining. In this case, the effect of DB75 was singularly different from that obtained after treatment with P9. The top inset of panel E shows interphase nuclei from untreated larvae post-stained with DB75. As described above (Figure 1E), the topo II signal is restricted to the nucleoplasm and is excluded from a DB75-bright region that corresponds to the chromocenter that includes the fourth chromosomes. Following treatment with DB75, we observed a complete reversal of the topo II signal from the nucleoplasm to the chromocenter and a quantitative redistibution of the enzyme to SAT III repeats on the X chromosomes and to chromosome 4 which is normally devoid of topo II (see bottom inset in panel E). This result extends those previously reported for DB351 and demonstrates that DB75 has a direct effect on the localization of topo II in chromatin. In contrast, DB293 did not affect the localization of topo II (panel F), clearly showing that diamidine substitutions and structure differentially affect chromatin-bound proteins.

Taken together, these results indicate that diamidines that possess a sufficient affinity for, and possibly induce structural changes in, AT-rich sequences can specifically target SAT III repeats and affect their interactions with proteins such as D1 and topo II. In the case of D1, the displacement may occur via a simple competition mechanism, an effect similar to that of P9 (Blattes et al, 2006), or a more complex topological mechanism may be involved as proposed for kinetoplast DNA (Wilson et al, 2005). Unlike P9, DB75 dramatically affects the intranuclear localization of topo II and quantitatively redistributes it to AT-rich SAT I and SAT III repeats. This uncoupling between the colocalization of D1 and topo II that is normally observed (see Figure 2) is unique to DB75 and most likely reflects a novel activity that may bear on the antiparasitic properties of diamidines.

DB75-mediated potentiation of topoisomerase II activity on SAT III repeats correlates with suppression of PEV

The striking DB75-incuced mobilization of topo II from the nucleoplasm to AT-rich satellites might result in increased enzymatic activity in these regions. We tested this possibility by incubating Schneider tissue-culture cells with increasing concentrations of DB75 before a 10-minute treatment with VM26, a time that corresponds to the onset of detectable VM26-induced topo II cleavage activity in embryos (Figure 3A). We first verified that DB75 alone did not induce topo II cleavage. Genomic DNA purified from cells treated for one hour with 500 nM DB75 and hybridized to a radioactive SAT III probe did not reveal the characteristic ladder of 359-bp topo II cleavage products (Figure 6A, lane 1, compare to Figure 3A). Treatment with VM26 alone led to the appearance of the 359-bp ladder (lane 2), which was strongly enhanced in a dose-dependent manner by prior incubation with DB75 (100-500 nM, lanes 3-5). Hence, DB75-induced relocalization of topo II to satellite sequences is accompanied by a strong increase in cleavage activity. Note that DB75 also relocalizes topo II to the SAT I repeats of chromosome 4 (bottom inset of Figure 5E). However, we could not detect VM26-induced cleavage activity in the $AATAT_n$ SAT I sequences under these conditions (data not shown). This negative result again suggests that topo II cleavage of SAT III is a characteristic feature of this region of the X chromosome. However, as topo II poisons exhibit marked differences in their ability to induce cleavage depending on the sequence of the genomic region studied (Borgnetto et al, 1996), we cannot exclude the possibility that VM26 cannot trap cleavable complexes on SAT I repeats.

Finally, we show that the mobilization of topo II to SAT III repeats by DB75 and the enhancement of topo II cleavage activity by the drug correlates with suppression of w^{m4} PEV. Feeding the closely related DB351 molecule to *white-mottled* larvae led to a dose-dependent suppression of PEV that was maximal at a concentration of 1 mM (Figure 6B). Again, the

control drug, DB293, did not modify the variegation of eye color in *white-mottled* adults recovered from these feeding experiments. We conclude from these results that perturbation of topo II localization and enzymatic activity directly affects heterochromatin-mediated mechanisms that influence w^{m4} PEV.

Diamidine/SAT III interactions as a model for antiparasitic activity

An important corollary of this conclusion is that the antiparasitic activity of diamidines such as DB75 or DB351 is likely determined by selective binding to AT-rich sequences in the genomes of their target microorganisms and by their ability to directly influence the localization and activity of topo II. If so, selective preferential binding of DB75 and DB351, but not DB293, should be observed in AT-rich parasitic genomes or portions thereof, comparable to that demonstrated for SAT I and SAT III repeats. This hypothesis can be readily tested in cultured trypanosomes, whose AT-rich kinetoplast DNA (kDNA) is likely to be a preferred target for DB75 and DB351. These sequences exhibit strong intrinsic DNA curvature due to the presence of a "bending locus" of phased homopolymeric dA•dT tracts (Shapiro & Englund, 1995), a characteristic shared by SAT III sequences (Susbielle et al, 2005).

In agreement with the binding characteristics of DB75 in *Drosophila* tissues, kDNA displayed the brightest fluorescence in trypanosomes treated with the drug for 1, 2 or 4 hours and the distribution and intensity of DB75 and of the closely related DB351 were found to be similar (Figure 7). Nuclear DNA binds these compounds more weakly and is faintly visible with DB75 and DB351 from the early time points. No other regions of significant fluorescence were seen in the trypanosomes. Similar to its binding to *Drosophila* chromosomes, DB293 showed significantly weaker fluorescence and only kDNA could be detected, albeit at a much lower intensity, up to the 4-hour time point.

We conclude from these results that the binding of DB75 and DB351 to trypanosome kDNA follows from our detailed observations in the *Drosophila* system. Future experiments will address the critical question of whether, as on SAT III sequences, selective binding of diamidines to kDNA induces a massive relocalization of parasite topo II, an enzyme crucial for the maintenance of kinetoplast circles (Wang & Englund, 2001), and whose accumulation and, possibly, enhanced activity (see Figure 6), might lead to fragmentation and destruction of the kDNA that carries essential genetic functions.

DISCUSSION

We have used the *Drosophila white-mottled* system to detect perturbations in mass interactions spanning several million DNA base pairs that occur between satellite sequences and proteins such as D1, topo II and HP1 (Blattes et al, 2006). Using PEV as a readily measurable reporter for these perturbations, we have shown that it is possible to follow, in a developing whole organism, the effects of ingested drugs. As demonstrated here, these effects can be addressed at the molecular level by analyzing the localization of potential protein targets, at the biochemical level by measuring drug effects on topo II enzymatic activity and by observing overall changes in the functional properties of heterochromatin as reflected by the extent of silencing of the w^{m4} allele of the *white* gene. Remarkably, these effects can be detected at drug doses, delivered by a simple feeding route, that do not noticeably interfere with development or viability. These properties hence make it possible to address issues of drug toxicity and mechanism of action in a fully integrated, physiologically relevant system, circumventing many of the limitations of tissue culture systems.

Our results show that minor-groove interactions between SAT III sequences and synthetic ligands have profound effects on the protein composition of heterochromatin and its functional properties. Surprisingly, these perturbations can vary considerably depending on the structure of the minor-groove binders used. These include the natural D1 protein, its synthetic P9 mimic and the structurally unrelated DB75 diamidine currently in clinical trials against parasitic diseases.

P9 induces a displacement of D1 from SAT III repeats that is accompanied by a subsequent loss of topo II. This activity was demonstrated in a previous study (Blattes et al, 2006) that could not distinguish between direct and indirect P9-induced effects on topo II localizzation. We have extended these observations here to show that over-expression of

D1ΔE, but not full-length D1, induces a similar displacement of topo II (Figure 2), further supporting the hypothesis that the acidic C-terminal domain absent in D1ΔE is required for the localization of topo II to SAT III sequences. We speculate that, *in vivo*, D1 binding promotes a secondary interaction with topo II that allows for cleavage by the enzyme, as it does *in vitro* (Figure 4). The strikingly different behavior of topo II localized in the region encompassing rDNA repeats adjacent to the SAT III array, which is not affected by P9 treatment, suggests that assembly of topo II on SAT III and rDNA repeats depends on distinct D1-dependent and D1-independent pathways, a hypothesis supported by our observation that a subset of topo II sites on polytene chromosomes is not associated with D1 (Figure 2E).

D1-independent targeting of topo II to the SAT III array is supported by our results showing that AT-specific diamidines can uncouple D1 and topo II binding in this 11-Mbp region. This is an unexpected result as the primary consequence of DB75 binding is to displace D1 from the SAT III array (Figure 5C), an effect similar to that of P9 (Figure 1D). However, DB75 has the opposite effect on the localization of topo II and redistributes the enzyme to SAT III and SAT I sequences (Figure 5E). This unique property, which is reflected by enhancement of VM26-induced cleavage of SAT III repeats and the ability of the closely related DB351 to suppress *white-mottled* PEV, demonstrates that different minor-groove ligands that target AT-rich sequences can selectively affect a subset of DNA/protein interactions that are normally coupled *in vivo*. These results also provide additional support for a functional implication of topo II in heterochromatin assembly.

Topoisomerase II and heterochromatin condensation

Heterochromatin is classically defined as the portion of the genome which fails to decondense after mitosis. Topo II is required for mitotic chromosome condensation (Adachi et al, 1991) and it may play a similar role in the packaging of SAT III heterochromatin in a

condensed state during interphase. *In vivo* topo II cleavage of SAT III repeats (Figure 3A) is likely to play a role in the dynamics of heterochromatin assembly or maintenance, a possibility supported by the suppression of heterochromatin-mediated silencing by VM26, a topo II poison (Blattes et al, 2006). While we have been unable to study the effects of catalytic inhibitors of topo II such as ICRF193, which are lethal during *Drosophila* development, we favor a model whereby the suppression of *white-mottled* PEV by VM26 is a direct result of trapping active topo II in SAT III repeats and, possibly, in adjacent rDNA repeats.

It is noteworthy that topo II cleavage occurs *in vivo* in an internucleosomal region that may be free of D1 protein, as suggested by DNAse I footprinting experiments (Blattes et al, 2006). This region does not bind P9 and indeed becomes hypersensitive to nuclease digestion upon binding of D1 or distamycin to adjacent sequences (Figure 4). We propose that D1 binding to SAT III sequences serves as a guide for topo II, creating an accessible site where the enzyme may be recruited for cleavage. This model is in keeping with the results of our *in vitro* experiments which demonstrate that the *in vivo* specificity of topo II cleavage can be fully reconstituted by complexing SAT III DNA with D1 (Figure 4). This property extends more generally to small minor-groove binders such as the mammalian High Mobility-Group A (HMGA) protein (data not shown). Like D1, HMGA contains repeated copies of the AT hook motif and is associated with alpha-satellites *in vivo* (Reeves, 2001). Based on these similarities, we propose that D1 and HMGA might be functional orthologues. *In vitro* binding of D1 or HMGA restricts topo II cleavage to a single site in each 359-bp SAT III repeat and thus reconstitutes the periodicity of cleavage observed *in vivo* in one of the two positioned SAT III nucleosomal linkers (Käs & Laemmli, 1992), albeit in the absence of nucleosomes. This indicates that the underlying DNA sequence of SAT III repeats and their unexpected accessibility to minor-groove binders play important roles in determining nucleosome

positioning and interactions with a select subset of chromosomal proteins, including topo II. These properties might provide signals for the assembly of heterochromatin on SAT III repeats, which might then spread to adjacent sequences.

Minor grove interactions and modulation of the functional properties of chromatin

Previous reports have identified several minor-groove binders as AT hook mimics or competitors that can selectively displace AT-specific proteins from chromatin. This was originally shown for distamycin (Käs et al, 1993) and MATH20 (Girard et al, 1998), an artificial protein containing 20 AT hooks which selectively binds SAT III repeats and displaces the *Drosophila* D1 protein (Monod et al, 2002). The P9 oligopyrrole, the first synthetic AT hook mimic designed to bind AT-rich satellites with remarkable sub-nanomolar affinity (Janssen et al, 2000a), has the same properties and suppresses w^{m4} PEV when fed to *white-mottled* larvae (Janssen et al, 2000b). The elucidation of the molecular mechanisms that account for these effects has helped identify topo II as a novel target for minor-groove ligands that modulate chromatin structure and function and, more specifically, help determine the assembly of specialized structures such as heterochromatin (Blattes et al, 2006). Taken together, these results suggest that it is possible to induce wide-ranging effects in the functional properties of chromatin by taking advantage of the surprising accessibility of the minor groove of AT-rich sequences, even in highly compact pericentric regions.

An anti-genome mechanism of activity for antiparasitic drugs

We have extended these results by demonstrating that another class of minor groove binders that similarly target AT-rich satellites in the *Drosophila* genome not only share these properties but, in addition, possess unique activities that may account for their biological effects. Diamidines have been developed as novel antiparasitic drugs that, *a priori*, act by

binding to AT-rich DNA (Tidwell & Boykin, 2003; Wilson et al, 2005), a property which might serve to target specific protein/DNA interaction sites. Satellite repeats, because of their highly biased base composition and the huge chromosomal distances they span, can serve as enormous sinks for these small molecules, and are indeed preferentially stained by these autofluorescent drugs (Figure 5). The genomes of parasites such as *Pneumocystis carinii* or *Plasmodium falciparum* have approximately 70% and 80% A+T content, respectively (Ginsburg et al, 1993). Another example of extraordinarily biased sequence content is the kinetoplast circles of trypanosomes and Leishmania (Shlomai, 2004), a likely target of diamidines such as DB75, currently in clinical trials, and which indeed binds these structures preferentially (Mathis et al, 2006).

The striking effects of DB75 and of the closely related DB351 on topo II localization and activity are discussed next. These molecules can cause structural changes in the kinetoplast that cause significant disruptions in kDNA replication with subsequent biological effects that lead to cell death. Topo II is essential for the replication of kinetoplast circles and relocalization of the enzyme to these regions and the accompanying increase in cleavage activity induced by drugs such as DB75 might interfere with the replication of these circular DNA molecules or, alternatively, lead to extensive DNA breakage. Either should prove lethal. Fluorescence microscopic analysis of trypanosomes from infected mice, taken after treatment with DB75 or DB351, clearly show strong initial compound fluorescence in the mitochondrial kinetoplast with much weaker fluorescence in the nucleus (Figure 7). The kinetoplast then disappears from treated cells over a 24-48 hour period and the cells die (Wilson et al, 2005; Mathis et al, 2006). Results obtained from the *Drosophila* system and the demonstrated ability to use this model to study in detail drug/DNA/protein interactions both *in vivo* and *in vitro* suggest a mechanism for diamidine activity based on the targeting of genomic regions such as kinetoplasts based on their highly skewed A+T content and selective interference with

metabolic reactions that implicate these sequences. Other drug-targeted parasites that do not have kinetoplasts, such as *Plasmodium falciparum*, might present an even larger target as their whole genome is remarkably AT-rich. It seems highly likely, based on the results presented here, that selective targeting of DNA/protein complexes in the nuclei of these organisms is involved in the antiparasitic action of diamidines.

An important question is that of the identity of the proteins whose localization and/or activity are perturbed by minor-groove ligands. We have identified topo II as a major target for a broad array of molecules, ranging from the natural D1 protein to the synthetic polyamide and diamidines used in our studies. We do not know at present whether topo II is a target for DB75 in trypanosomes but, taken together, its central role in kDNA maintenance (Wang & Englund, 2001), the selective concentration of DB75 in kDNA (Figure 7) and the ensuing fragmentation and destruction of kDNA (Wilson et al, 2005; Mathis et al, 2006) all lend support for this model. We propose that these phenomena reflect topo II hyperactivity induced at diamidine-bound genome regions, of which kDNA constitutes a special example. The dramatic effect of DB75 on the localization of topo II in *Drosophila* tissues and its strong stimulation of SAT III cleavage likewise support this intriguing hypothesis.

EXPERIMENTAL PROCEDURES

Immunostaining and fluorescence immunolocalization. Labeling with rabbit polyclonal antibodies against *Drosophila* D1 (1:200 dilution) and topo II (1:1,000 dilution), staining with P9F or P31TR and drug treatments of permeabilized larval neuroblasts from third-instar larvae was performed exactly as described (Blattes et al, 2006). Permeabilized brains were incubated for 60 minutes at room temperature with the indicated concentrations of P9, DB75, DB351 or DB293 before fixation and immunostaining. The transgenic HS-D1 and HS-D1ΔE lines used for the experiment shown in Figure 2 have been described (Aulner et al, 2002) and express inducible full-length and truncated D1 sequences, respectively, under the control of a heat-shock promoter. Induction of HS-D1 and HS-D1ΔE transgene expression and recovery were as previously described (Aulner et al, 2002) and dissected larval brains were processed as described above. Samples were counterstained with DAPI or, where indicated, with DB75 or DB293 (50 nM). For visualization of diamidine autofluorescence from drug-treated samples, excitation was performed in the UV range as for DAPI, yielding a blue fluorescent signal. Photographs were taken on a Leica DMRB microscope equipped with a CoolSnap CCD HQ camera (Roper Scientific) and processed for contrast with Adobe Photoshop. *In vivo* drug treatments, topo II cleavage and DNAse I assays, double immunolabeling experiments and trypanosome culture and drug treatments were performed as described in Supplementary Material.

ACKNOWLEDGEMENTS

We thank U. K. Laemmli (University of Geneva) for generous gifts of synthetic polyamides, J.-h. Wu and T.-s. Hsieh (Duke University) for providing us with anti-topo II antibodies, J. Eissenberg, M. Grigoriev and M. Weber for critical reading of the manuscript and for useful suggestions, the Centre de Biologie du Développement (CBD, Toulouse) for use of *Drosophila* culture facilities and the Institut d'Exploration Fonctionnelle des Génomes (IEFG) for microscopy facilities. This work was supported by the ARC grant No. 3319, by a "Genetic instability as a pejorative outcome in cancer" Cancéropôle grant and by the Agence Nationale pour la Recherche (EK), by the National Institutes of Health (NIH grant GM42539 to PKG and grants GM61587 and AI064200 to WDW and DWB), and by a grant from the Bill and Melinda Gates Foundation to JEH, DWB, and WDW. RB and GS were supported by doctoral fellowships from the French Ministry of Research and from the Ligue Nationale Contre le Cancer, respectively, and AMM was supported by a predoctoral fellowship from Amgen.

REFERENCES

Adachi Y, Luke M, Laemmli UK (1991) Chromosome assembly in vitro: topoisomerase II is required for condensation. *Cell* **64**: 137-148

Aulner N, Monod C, Mandicourt G, Jullien D, Cuvier O, Sall A, Janssen S, Laemmli UK, Käs E (2002) The AT-hook protein D1 Is essential for *Drosophila melanogaster* development and is implicated in position-effect variegation. *Mol Cell Biol* **22**: 1218-1232

Blattes R, Monod C, Susbielle G, Cuvier O, Wu JH, Hsieh TS, Laemmli UK, Käs E (2006) Displacement of D1, HP1 and topoisomerase II from satellite heterochromatin by a specific polyamide. *EMBO J* **25**: 2397-2408

Borgnetto ME, Zunino F, Tinelli S, Käs E, Capranico G (1996) Drug-specific sites of topoisomerase II DNA cleavage in Drosophila chromatin: heterogeneous localization and reversibility. *Cancer Res* **56**: 1855-1862

DiNardo S, Voelkel K, Sternglanz R (1984) DNA topoisomerase II mutant of Saccharomyces cerevisiae: topoisomerase II is required for segregation of daughter molecules at the termination of DNA replication. *Proc Natl Acad Sci USA* **81**: 2616-2620

Ginsburg H, Nissani E, Krugliak M, Williamson DH (1993) Selective toxicity to malaria parasites by non-intercalating DNA-binding ligands. *Mol Biochem Parasitol* **58**: 7-15

Girard F, Bello B, Laemmli UK, Gehring WJ (1998) *In vivo* analysis of scaffold-associated regions in *Drosophila:* a synthetic high-affinity SAR binding protein suppresses position effect variegation. *EMBO J* **17**: 2079-2085

Huisinga KL, Brower-Toland B, Elgin SC (2006) The contradictory definitions of heterochromatin: transcription and silencing. *Chromosoma* **28**: 28

Janssen S, Durussel T, Laemmli UK (2000a) Chromatin opening of DNA satellites by targeted sequence-specific drugs. *Mol Cell* **6**: 999-1011

Janssen S, Cuvier O, Muller M, Laemmli UK (2000b) Specific gain- and loss-of-function phenotypes induced by satellite- specific DNA-binding drugs fed to *Drosophila melanogaster*. *Mol Cell* **6**: 1013-1024

Käs E, Laemmli UK (1992) *In vivo* topoisomerase II cleavage of the *Drosophila* histone and satellite III repeats: DNA sequence and structural characteristics. *EMBO J* **11**: 705-716

Käs E, Poljak L, Adachi Y, Laemmli UK (1993) A model for chromatin opening: stimulation of topoisomerase II and restriction enzyme cleavage of chromatin by distamycin. *EMBO J* **12**: 115-126

Lansiaux A, Tanious F, Mishal Z, Dassonneville L, Kumar A, Stephens CE, Hu Q, Wilson WD, Boykin DW, Bailly C (2002) Distribution of furamidine analogues in tumor cells: targeting of the nucleus or mitochondria depending on the amidine substitution. *Cancer Res* **62**: 7219-7229

Maeshima K, Laemmli UK (2003) A two-step scaffolding model for mitotic chromosome assembly. *Dev Cell* **4**: 467-480

Mathis AM, Holman JL, Sturk LM, Ismail MA, Boykin DW, Tidwell RR, Hall JE (2006) Accumulation and intracellular distribution of antitrypanosomal diamidine compounds DB75 and DB820 in African trypanosomes. *Antimicrob Agents Chemother* **50**: 2185-2191

Monod C, Aulner N, Cuvier O, Käs E (2002) Modification of position-effect variegation by competition for binding to Drosophila satellites. *EMBO Rep* **3**: 747-752

Reeves R (2001) Molecular biology of HMGA proteins: hubs of nuclear function. *Gene* **277**: 63-81

Saitoh Y, Laemmli UK (1994) Metaphase chromosome structure: bands arise from a differential folding path of the highly AT-rich scaffold. *Cell* **76**: 609-622

Shapiro TA, Englund PT (1995) The structure and replication of kinetoplast DNA. *Annu Rev Microbiol* **49**: 117-143

Shlomai J (2004) The structure and replication of kinetoplast DNA. *Curr Mol Med* **4**: 623-647

Susbielle G, Blattes R, Brevet V, Monod C, Käs E (2005) Target practice: aiming at satellite repeats with DNA minor groove binders. *Curr Med Chem Anticancer Agents* **5**: 409-420

Tidwell RR, Boykin DW (2003) DNA and RNA binders: from small molecules to drugs. In *Demeunynck, M, Bailly, C, Wilson, WD* pp 414-460

Wang JC (2002) Cellular roles of DNA topoisomerases: a molecular perspective. *Nat Rev Mol Cell Biol* **3**: 430-440

Wang L, Bailly C, Kumar A, Ding D, Bajic M, Boykin DW, Wilson WD (2000) Specific molecular recognition of mixed nucleic acid sequences: an aromatic dication that binds in the DNA minor groove as a dimer. *Proc Natl Acad Sci USA* **97**: 12-16

Wang Z, Englund PT (2001) RNA interference of a trypanosome topoisomerase II causes progressive loss of mitochondrial DNA. *EMBO J* **20**: 4674-4683

Wilson WD, Nguyen B, Tanious FA, Mathis A, Hall JE, Stephens CE, Boykin DW (2005) Dications that target the DNA minor groove: compound design and preparation, DNA interactions, cellular distribution and biological activity. *Curr Med Chem Anticancer Agents* **5**: 389-408

FIGURE LEGENDS

Figure 1. Satellite localization of D1 and topoisomerase II and displacement by P9. A: female neuroblasts stained with P9F (green) and P31TR (red) to detect SAT III and I and SAT V repeats, respectively. The inset shows a 2-fold magnified X chromosome with SAT III repeats prominently stained by P9F. The adjacent rDNA locus is indicated. B: staining with an antibody against D1 (red) and counterstaining with P9F (green) and DAPI (grey). The inset shows a 2-fold enlarged X chromosome. D1 (red) was immunolocalized following treatment with 0 (C) or 250 nM P9 (D) and counterstaining with DAPI (blue). Insets show magnified X chromosomes. E and F: identical experiments using an antibody raised against topo II (red) and counterstaining with DAPI (blue). The localization of SAT I, SAT III and rDNA repeats is indicated in the insets showing 2-fold magnified X chromosomes. Chromosomes are identified as shown; scale bars = 5 μm.

Figure 2. Topoisomerase II is a companion of D1. A: the D1/D1ΔE signal (red) is shown for brains dissected from heat-shocked female larvae expressing the HS-D1ΔE transgene or from non heat-shocked larvae (inset). Samples were counterstained with DAPI (gray). The four chromosomes are numbered as shown. Mitotic chromosomes isolated from this line have a generally puffy appearance, particularly after heat shock. B: the topo II signal (red) is shown for heat-shocked larvae expressing the HS-D1ΔE transgene, arrowheads indicate the SAT III region on the X chromosome. The X chromosomes shown in panel C have been magnified 2-fold; SAT III repeats are indicated by arrowheads. The adjacent rDNA locus is also identified. Scale bars = 5 μm. D: polytene chromosome spreads stained for D1 (red) and counter-stained with DAPI (blue). Arrowheads point to the chromocenter. A broad co-localization of the protein and the dye is shown in greater detail in the split-signal image of a portion of an enlarged chromosome arm in the bottom panel. E: polytene chromosomes double-stained for

D1 (green) and topo II (red). The arrowhead points to the chromocenter. A split-signal image is shown in the bottom photograph for an enlarged portion of a chromosome arm.

Figure 3. Topoisomerase II cleaves SAT III repeats *in vivo*. Southern blots were performed on genomic DNA isolated from whole cellular blastoderm embryos treated with 50 μM VM26 for the times shown above the gels. DNA samples were hybridized to SAT III (panel A) or SAT I (panel B) probes. Lane 1 contained DNA from embryos incubated in the absence of VM26. The band visible near the bottom of the gel in panel A corresponds to the SAT III monomer cleavage product.

Figure 4. D1 guides topoisomerase II cleavage of SAT III repeats. VM26-induced topo II cleavage and DNAse I footprints were performed on a 359-bp SAT III fragment 3' end-labeled on the lower strand. Lanes 1 and 5 show cleavage products formed in the absence of protein or drug, while lanes 2-4 show cleavage products formed in the presence of 2, 5 and 10 ng D1 (2.7, 6.75 and 13.5 nM), distamycin (0.5, 1, 2 μM, lanes 6-8) and chromomycin (0.5, 1 μM, lanes 9-10). DNAse I footprints were performed in the absence of drugs (lane 11, ND) or in the presence of 1 μM distamycin or chromomycin (lanes 12 and 13, respectively). Unnumbered lanes correspond to chemical sequencing reactions of the SAT III fragment. Arrowheads indicate the position of the *in vivo* topo II cleavage site; the filled box to the right of the gel corresponds to the hypersensitive site induced by distamycin, white rectangles correspond to distamycin and D1-bound regions. Chromomycin yields a reversed protection pattern.

Figure 5. Uncoupling of D1/topoisomerase II association with SAT III by diamidines. Neuroblasts were stained with DB75 (blue, panel A) or DB293 (blue, panel B) and P31TR

(red). The structures of DB75, of the closely related DB351 and of DB293 are shown. Female and male neuroblasts were stained in panels A and B, respectively. C and D: permeabilized neuroblasts treated with 2.5 μM DB75 (C) or DB293 (D) and immunostained for D1 (red). The blue signal is that from diamidine autofluorescence. The top inset in C shows interphase nuclei following treatment with a lower concentration (500 nM) of DB75, leading to D1 displacement from DB75-bright regions to the nucleoplasm. Insets at the bottom of each panel show 2-fold enlarged X chromosomes; SAT III repeats are indicated. E and F: permeabilized larval neuroblasts treated with 2.5 μM DB75 (E) or DB293 (F) immunostained for topo II (red). Bottom insets show enlarged X chromosomes and chromosomes 4, normally unstained for topo II (Figure 1E). F: DB293-treated neuroblasts. Individual chromosomes are indicated as shown; scale bars = 5 μm.

Figure 6. Stimulation of topoisomerase II SAT III cleavage by DB75 and suppression of PEV by DB351. A: Schneider S2 cells treated with 0, 100, 250 or 500 nM DB75 for 1 hour (lanes 2-5) before subsequent treatment with 50 μM VM26 for 10 minutes. The sample in lane 1 was from cells treated with 500 nM DB75 alone. Hybridization to a SAT III probe detects VM26-induced topo II cleavage products. M = molecular weight markers. B: *white-mottled* third-instar larvae were fed increasing concentrations of DB351, DB293 or DMSO alone as shown in each photograph. Representative eyes of 5-day old adult flies are shown. An Oregon R adult is shown as a control for wild-type eye color. The arrow in the 500 μM DB351 photograph indicates a small sector of ommatidia in which *white* PEV has not been suppressed.

Figure 7. Preferential binding to kDNA by antiparasitic diamidines. S427 *Trypanosoma brucei brucei* trypanosomes were incubated *in vitro* with 500 nM of DB75 (left), DB351

(center), or DB293 (right) for 1, 2 and 4 hours. Washed trypanosomes were resuspended in mouse blood and viewed as dry smears. For each compound set, phase contrast images are on the left and fluorescence images are on the right side. Both nuclei and kinetoplasts can be seen from the blue fluorescence of all three compounds. Nuclei are found in the center of trypanosomes and the smaller kinetoplasts brightly stained by DB75 and DB351, but not DB293, are found at cell ends.

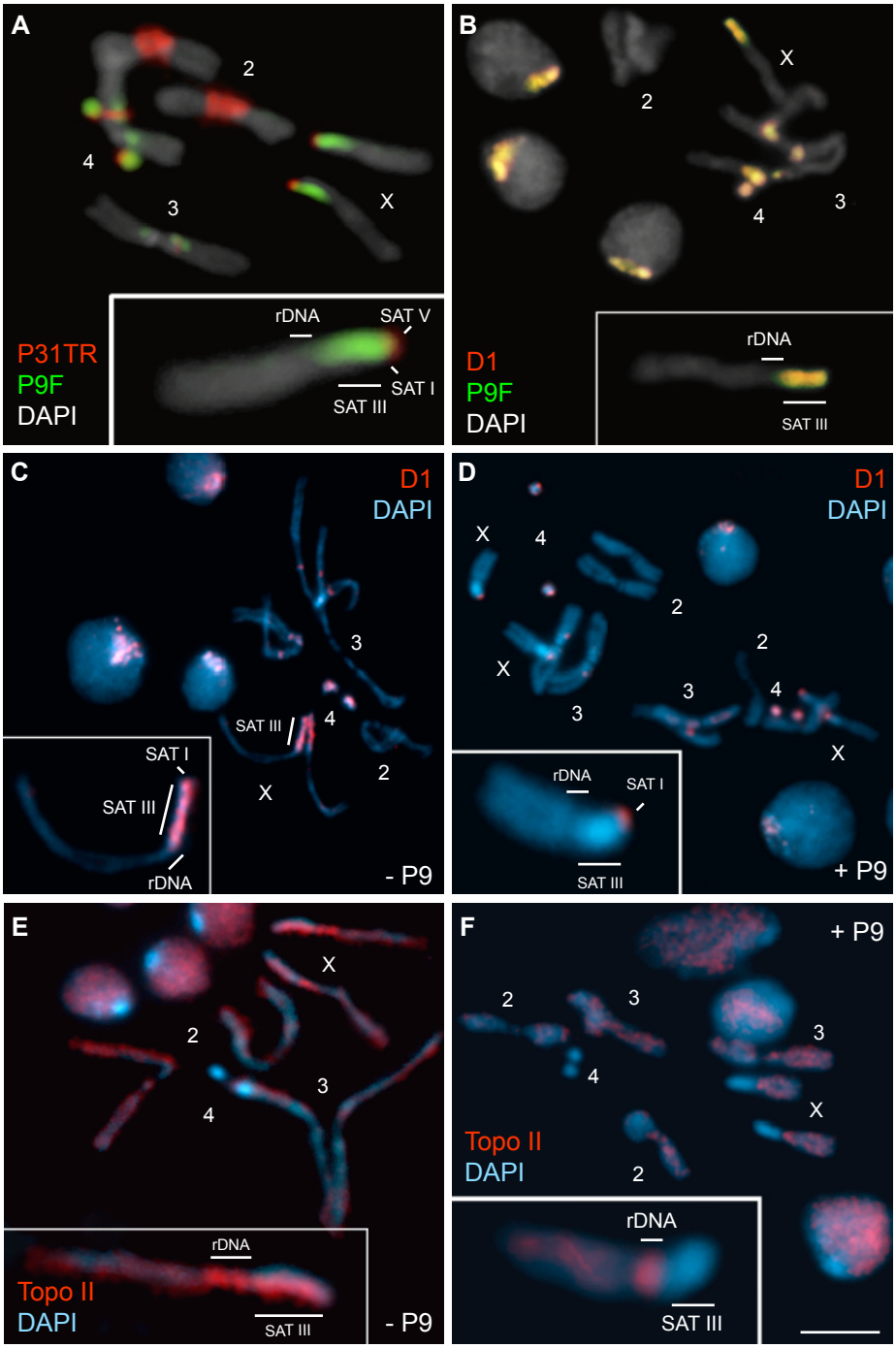

A
2
4
3
X
P31TR
P9F
DAPI
rDNA
SAT V
SAT I
SAT III
B
X
2
4
3
D1
P9F
DAPI
rDNA
SAT III
C
D1
DAPI
3
SAT III
4
2
X
SAT I
SAT III
rDNA
- P9
D
D1
DAPI
X
4
2
2
X
3
4
3
X
rDNA
SAT I
SAT III
+ P9
E
X
2
3
4
rDNA
Topo II
DAPI
SAT III
- P9
F
+ P9
2
3
3
4
X
Topo II
DAPI
2
rDNA
SAT III

Figure 1

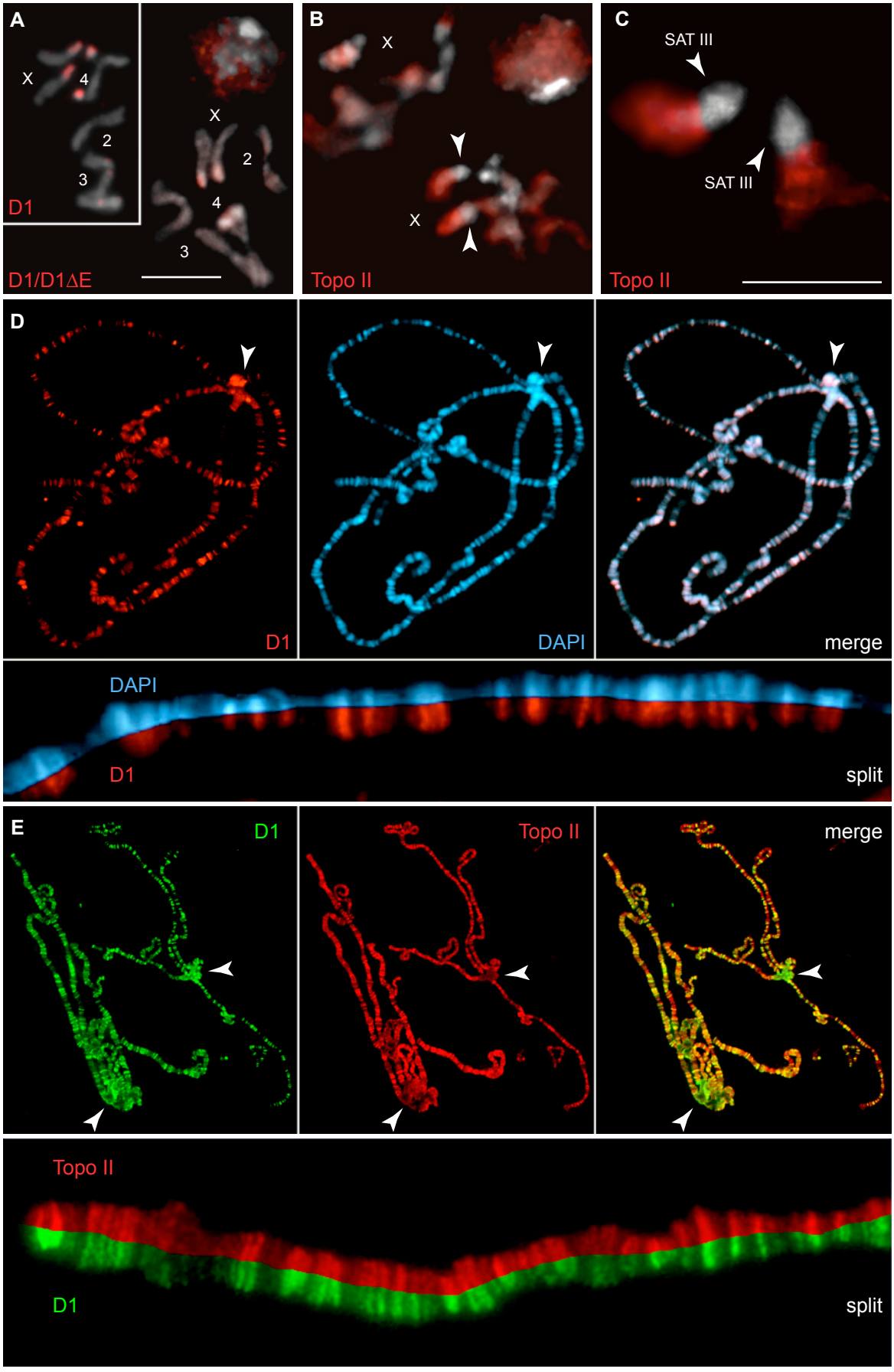

A
X
4
2
3
D1
X
2
4
3
D1/D1ΔE
B
X
X
Topo II
C
SAT III
SAT III
Topo II
D
D1
DAPI
merge
DAPI
D1
split
E
D1
Topo II
merge
Topo II
D1
split

Figure 2

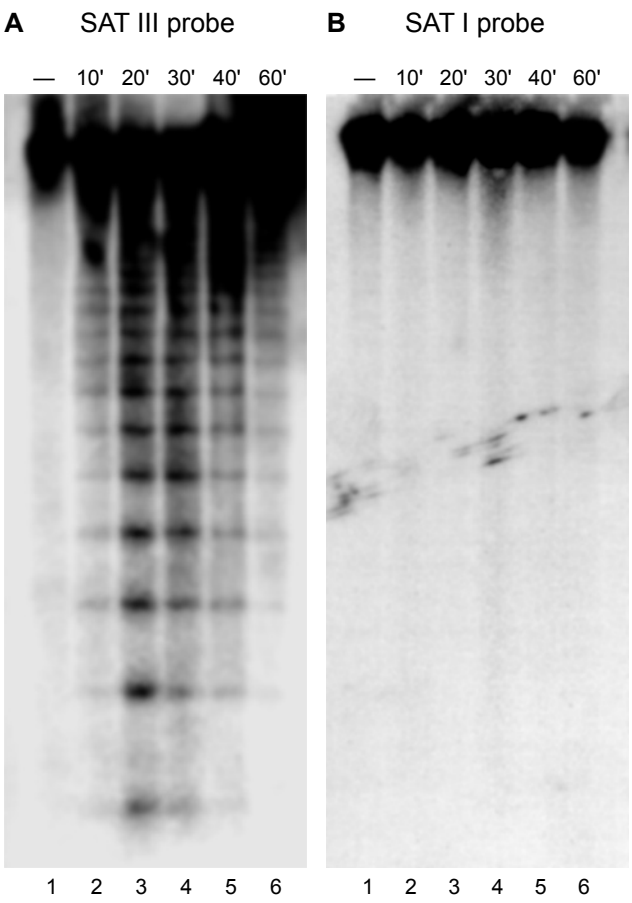

A
SAT III probe
B
SAT I probe
— 10' 20' 30' 40' 60'
— 10' 20' 30' 40' 60'
1 2 3 4 5 6
1 2 3 4 5 6

Figure 3

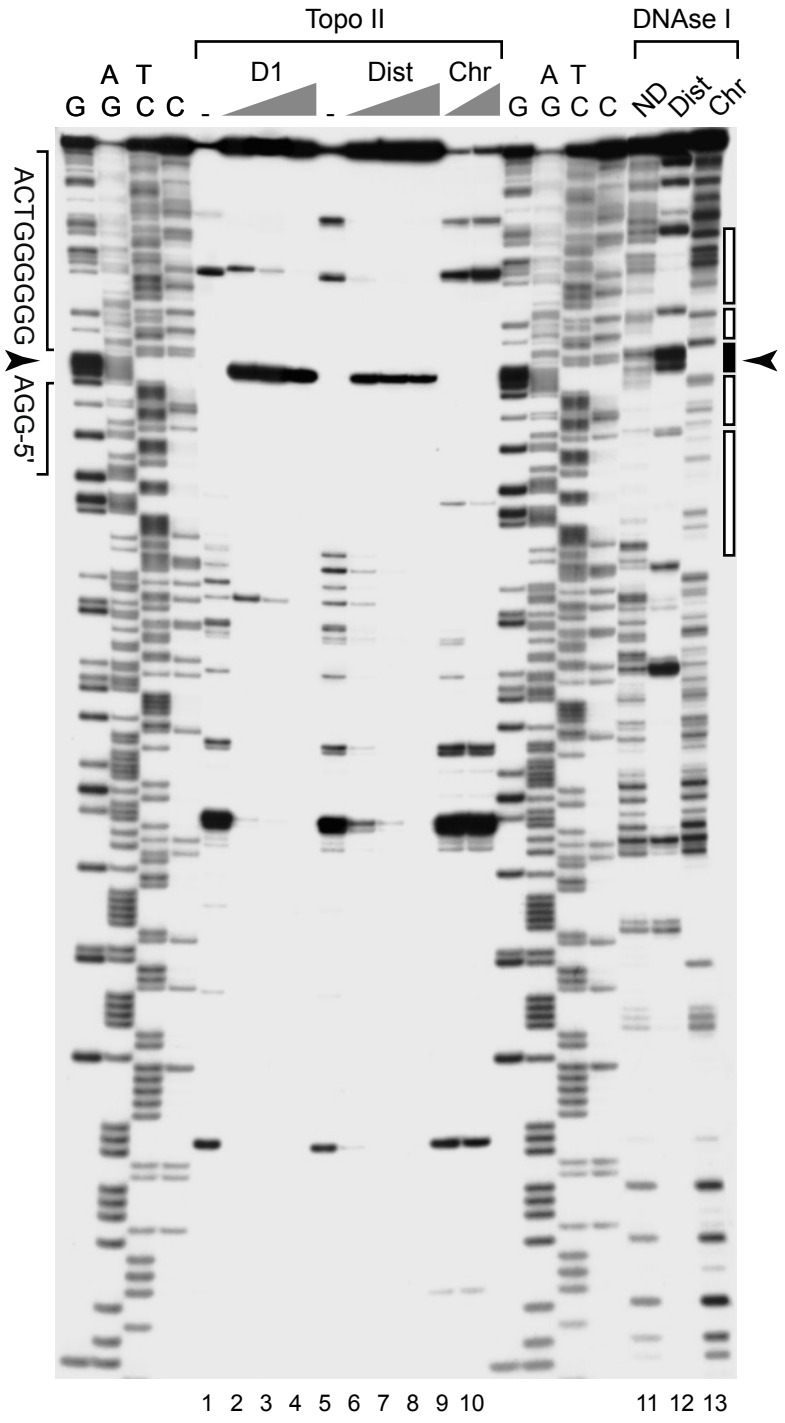

Topo II
DNAse I
D1
Dist
Chr
A T
G G C C
A T
G G C C
ND
Dist
Chr
ACTGGGGGG
AGG-5'
1 2 3 4 5 6 7 8 9 10
11 12 13

Figure 4

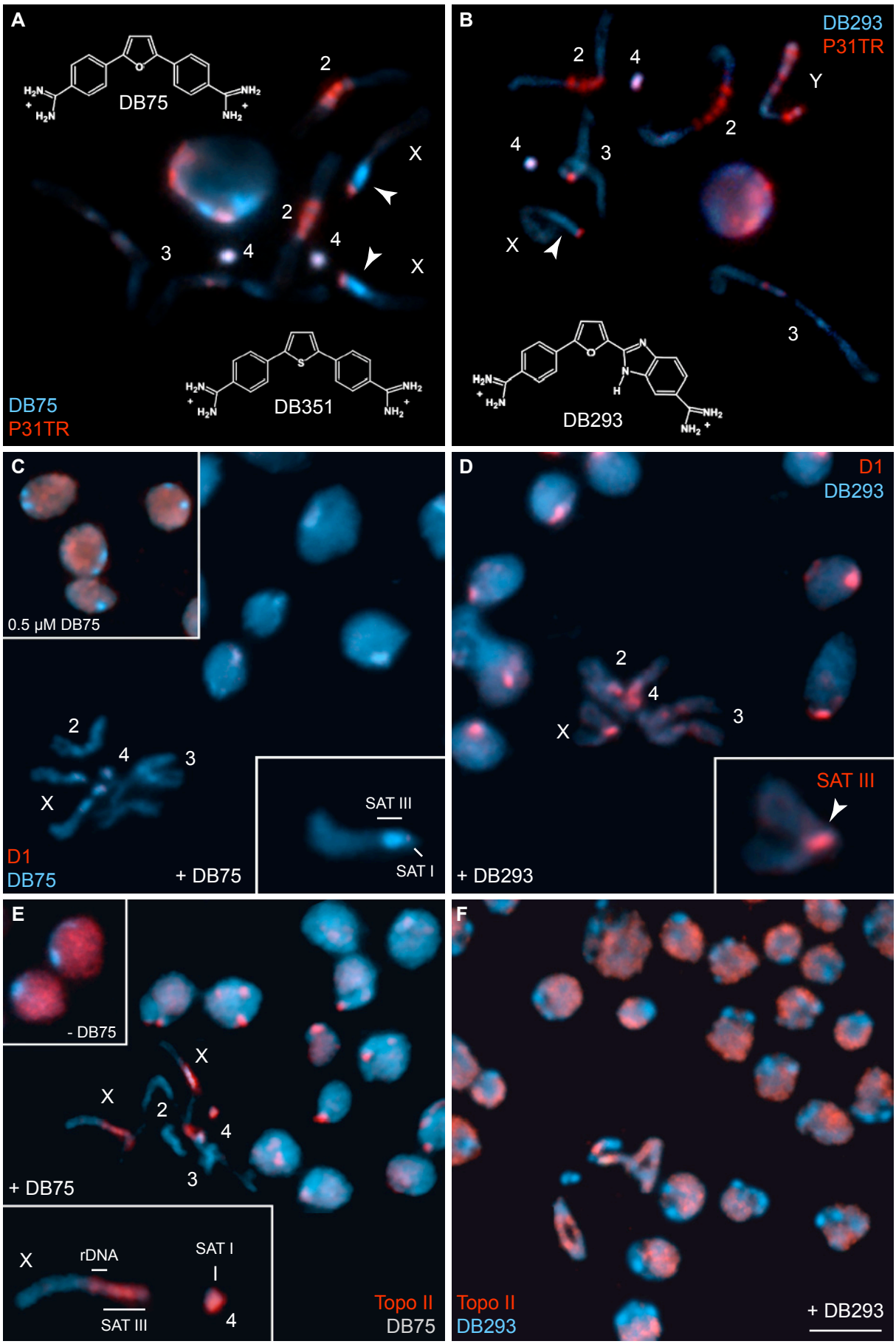

A
DB75
2
X
2
3
4
4
X
DB75
P31TR
DB351
B
DB293
P31TR
2
4
Y
2
4
3
X
3
DB293
C
0.5 μM DB75
2
4
3
X
D1
DB75
+ DB75
SAT III
SAT I
D
D1
DB293
2
4
3
X
SAT III
+ DB293
E
- DB75
X
X
2
4
3
+ DB75
X
rDNA
SAT I
SAT III
4
Topo II
DB75
F
Topo II
DB293
+ DB293

Figure 5

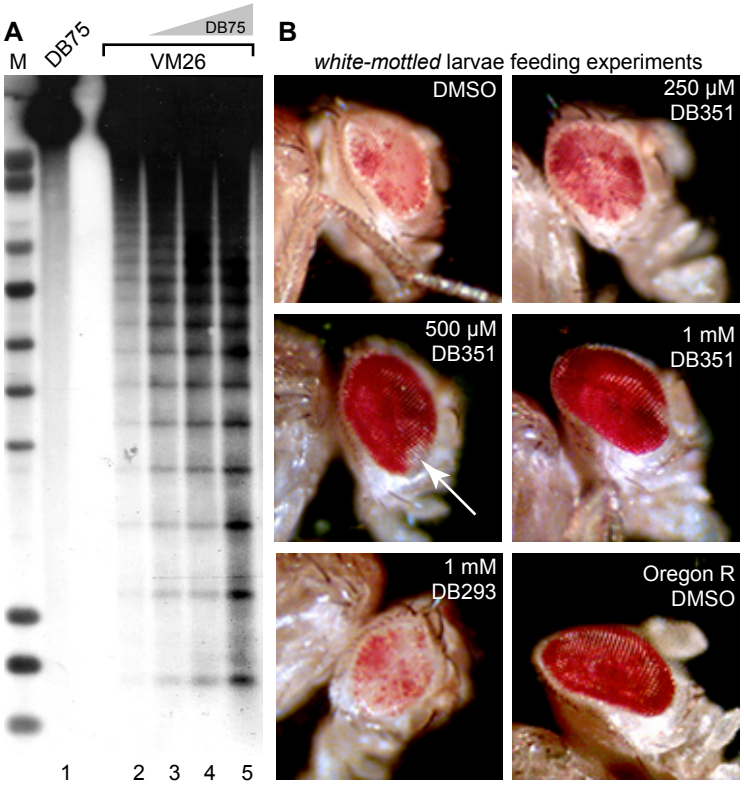

A
M
DB75
DB75
VM26
1
2
3
4
5
B
white-mottled larvae feeding experiments
DMSO
250 µM
DB351
500 µM
DB351
1 mM
DB351
1 mM
DB293
Oregon R
DMSO

Figure 6

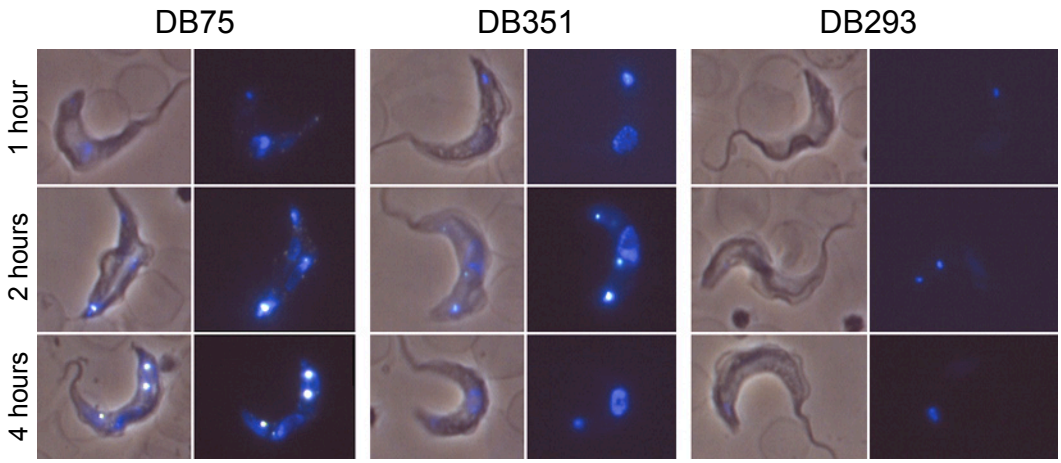


Figure 7

4. *Conclusions et discussions*

Des études précédentes nous ont permis de démontrer qu'au niveau du chromosome X, les patrons de localisation des protéines HP1 et topoïsomérase II sont similaires (Blattes et al., 2006). Cette observation nous avait déjà amené à envisager un rôle de la topoïsomérase II dans l'assemblage de la structure hétérochromatinienne. Cette hypothèse a été confirmée par les résultats des expériences de traitement avec le polyamide P9 dans lesquelles la protéine topoïsomérase II était déplacée des séquences SAT III (Blattes et al., 2006). Mais ces premiers résultats ne permettaient pas de conclure à une relation directe ou indirecte entre la localisation de la protéine topoïsomérase II et celle de la protéine D1.

Afin de pouvoir répondre à cette question, j'ai utilisé une lignée de drosophile exprimant le mutant D1ΔE. Suite à l'expression de ce mutant, une perte de la topoïsomérase II des séquences SAT III a été observée. Ces résultats suggèrent que la protéine D1 est nécessaire au recrutement de la topoïsomérase II au niveau des séquences SAT III. Plus précisément, c'est la queue acide C-terminale de la protéine D1 qui semble être impliquée dans ce recrutement. L'hypothèse du recrutement indirect de la topoïsomérase II sur des séquences qu'elle clive avait déjà était envisagée suite à des expériences *in vitro* de compétition entre l'histone H1, la protéine HMG-A et la topoïsomérase II (Poljak and Käs, 1995).

Des expériences complémentaires de colocalisation des protéines D1 et topoïsomérase II ont été menées sur des chromosomes polytènes, système dans lequel la localisation de la protéine D1 n'est pas restreinte aux séquences SAT I et SAT III. En effet, l'utilisation des chromosomes polytènes permet de révéler des sites d'association pour la protéine D1 au niveau de séquences endorépliquées euchromatiniennes, ses cibles classiques que sont les séquences satellites riches en paires de bases dA•dT étant sous-répliquées. Les résultats confirment la forte association de ces deux protéines sur de nombreux sites. Cependant, une petite sous-population de topoïsomérase II n'est pas colocalisée avec la protéine D1. Ce résultat est en accord avec la situation présentée au niveau du rDNA où la topoïsomérase II est présente, mais pas la protéine D1. De plus, l'interaction entre les protéines D1 et topoïsomérase II peut être indirecte et impliquer une autre protéine. En effet,

une étude sur les interactions protéines/protéines potentielles n'avait pas identifié la topoïsomérase II comme un interactant de la protéine D1 (Giot et al., 2003).

Ces résultats ont montré que la liaison de la topoïsomérase II aux séquences SAT III nécessite la présence de la forme entière de la protéine D1. Des études complémentaires menées *in vitro* nous ont permis de démontrer que la topoïsomérase II fixée à ces mêmes séquences était catalytiquement active et que cette activité nécessitait la présence de la protéine D1. Cette observation est en corrélation avec les expériences d'ingestion de poisons de la topoïsomérase II par des larves de la lignée *white-mottled*. Suite à ces expériences, nous avions démontré que la topoïsomérase II était probablement un suppresseur de la PEV. Ces résultats sont en accord avec les données montrant que la plupart des suppresseurs de la variégation d'effet de position du gène *white* sont impliqués dans les mécanismes de structuration des chromosomes (Le et al., 2004).

Une des activités de la topoïsomérase II consiste à cliver l'ADN double brin super enroulé pour le relâcher. Cette activité est nécessaire au cours de la réplication (Mondal and Parvin, 2001). Au cours de la progression de la fourche de réplication, chaque côté de la bulle de réplication va former des super tours d'ADN. L'absence d'activité capable de relâcher ces super tours contraindrait trop la molécule et inhiberait l'avancée de la fourche de réplication. La topoïsomérase II joue donc un rôle dans la relaxation de l'ADN, nécessaire à la fois à la réplication mais aussi en début de condensation des chromosomes mitotiques (Cuvier and Hirano, 2003). Il serait donc envisageable que cette enzyme joue aussi un rôle topologique au niveau des séquences SAT III.

L'activité de clivage de la topoïsomérase II pourrait générer un relâchement de la chromatine impliquée dans la formation d'une structure hétérochromatinienne et ainsi permettre l'interaction de certaines protéines avec l'ADN. La topoïsomérase II pourrait donc réguler les interactions protéines/ADN au niveau des séquences SAT III en agissant sur la topologie de la molécule d'ADN, tout comme la protéine D1 modulerait ces interactions par recrutement direct ou indirect d'autres protéines. Les protéines D1 et topoïsomérase II seraient donc des protéines capables de réguler les interactions des protéines de l'hétérochromatine avec les séquences nucléotidiques. Elles seraient donc à la base de la formation et du maintien d'une telle structure.

Cette fonction de la topoïsomérase II semble être affectée par les drogues antiparasitaires de la famille des diamidines. Nous avions déjà mis en évidence la différence d'effet entre les polyamides et les diamidines quant à la localisation de la topoïsomérase II alors que ces deux composés avaient le même effet sur la protéine D1 (ce travail et (Blattes et al., 2006)). Nous avons démontré que lors de l'utilisation de la diamidine DB75, ou de ses dérivés, la topoïsomérase II accumulée sur les séquences SAT III était catalytiquement active. Ainsi, les diamidines agiraient non pas en déplaçant la topoïsomérase II, éliminant ainsi l'activité de cette protéine sur les séquences SAT III, mais au contraire en augmentant la quantité de protéine fixée. Cette accumulation se traduit par une augmentation massive de l'activité de clivage au niveau de ces séquences. Ce résultat est à corréler avec l'effet des diamidines sur les trypanosomes. Suite à un traitement avec ces drogues, les kinétoplastes sont totalement fragmentés et non plus caténés comme dans les organismes non traités (Wilson et al., 2005).

Les résultats obtenus au cours de ce travail suggèrent un rôle de guidage de la topoïsomérase II par la protéine D1 vers les séquences SAT III. En effet, dans un contexte sans drogue, le recrutement de la topoïsomérase II vers ces séquences satellites serait facilité par la protéine D1. Cette hypothèse est d'autant plus appuyée que dans les expériences de footprint à la DNase I réalisées en présence de l'oligopyrrole P9 ou de la distamycine : ces deux molécules ne se fixent pas au niveau du site de clivage de la topoïsomérase II. Ces résultats démontrent que la protéine D1 est absente du site de clivage par la topoïsomérase II. De même, les résultats montrant un découplage entre les protéines D1 et topoïsomérase II dans les expériences de traitements avec les diamidines suggèrent un autre mode de recrutement de cette dernière au niveau des séquences SAT III. La topoïsomérase II possède des propriétés de liaison directes à l'ADN de ce type. Ces propriétés pourraient être altérées lorsque l'ADN est impliqué dans une structure de type hétérochromatinienne.

Les effets différents des oligopyrroles et des diamidines sur cette structure expliqueraient alors la différence de résultats. En présence du polyamide P9, les protéines D1 et topoïsomérase II sont déplacées des séquences SAT III alors qu'en présence de DB75, seule la protéine D1 est déplacée et au contraire, la topoïsomérase II est massivement relocalisée sur ces séquences. Dans un même temps, le marquage nucléoplasmique correspondant aux régions euchromatiniennes en interphase diminue au profit d'un marquage localisé sur les séquences SAT III. Ce résultat est discuté précédemment (voir chapitre I de

cette même section). Cependant, cette perte massive de la topoïsomérase II suppose une liaison de DB75 avec l'euchromatine qui n'est pas clairement démontrée (Susbielle et al., 2005). La relocalisation de la topoïsomérase II sur ces séquences correspondrait à un rééquilibrage rapide de la quantité de protéine fixée par remobilisation de celle-ci fixée sur des sites euchromatiniens. De même, l'hypothèse d'une ouverture massive de la chromatine sous l'action des diamidines permettrait d'expliquer la présence de topoïsomérase II sur les chromosomes 4, localisation non détectée sans traitement.

Une autre hypothèse suppose une ouverture massive de la chromatine par les diamidines entraînant un déplacement quantitatif de la topoïsomérase II des séquences satellites riches en paires de bases dA•dT, effet similaire à celui observé dans le cas d'un traitement avec le polyamide P31 et la localisation du facteur GAF (Janssen et al., 2000b). L'utilisation de la distamycine, archétype des ligands tels que le polyamide P9, favorise également le clivage des séquences SAT III, probablement par un mécanisme d'ouverture de la chromatine, comme cela avait déjà été décrit *in vivo* (Käs et al., 1993).

Les résultats obtenus suite à des traitements par le DB75 tendent à démontrer l'existence d'un découplage entre la localisation de la protéine D1 et de la topoïsomérase II sur les séquences SAT III, conclusion qui va à l'encontre des résultats obtenus avec le polyamide P9. Des expériences en présence de P9 montrait un déplacement de la topoïsomérase II et de la protéine D1 des séquences SAT III (Blattes et al., 2006). Ces observations sont confortées par les expériences de colocalisation de ces deux protéines sur chromosomes polytènes. Les résultats suggèrent l'existence d'une sous-population de topoïsomérase II qui se fixerait à l'ADN sans la présence de la protéine D1. Certes, la présence de la protéine D1 pleine taille sur les séquences SAT III semble jouer un rôle dans la localisation mais aussi l'activité de clivage de la topoïsomérase II. Cependant, les expériences *in vitro* montrent que la relocalisation de la topoïsomérase II causée par les traitements avec le DB75 n'affecte nullement son activité. Par contre, la relocalisation de la topoïsomérase II sur les séquences SAT I du chromosome 4 ne semble pas corrélé avec l'activité de clivage de cette protéine puisque aucun clivage de ces séquences n'a été observé *in vitro*. Il est possible de proposer qu'en se fixant dans le petit sillon de l'ADN riche en paire de bases dA•dT, le polyamide P9 en déplace la protéine D1 alors que les diamidines joueraient un rôle d'ouverture de ce petit sillon dans lequel la protéine D1 ne pourrait plus se fixer, sans pour autant affecter la localisation de la topoïsomérase II.

Des expériences utilisant le DB351, un dérivé de structure très proche au DB75, montrant que ce composé est capable de supprimer la PEV de l'allèle *white-mottled* quand celui-ci est donné à manger à des larves (Susbielle et al., 2005). Des études complémentaires *in vitro* ont été réalisée en présence de VM26, un poison de la topoïsomérase II qui bloque l'activité enzymatique de la protéine après le clivage du premier brin et stabilise le complexe enzyme/ADN de manière réversible. Sans addition de DB75, le clivage des séquences SAT III en présence de VM26 décroît avec le temps, confirmant la réversibilité du complexe de clivage formé par la topoïsomérase II et le VM26 (Borgnetto et al., 1996). En présence de DB75 et de VM26, la formation des produits de clivage est augmentée. Ces expériences confirment que la topoïsomérase II relocalisée sur les séquences SAT III est catalytiquement active et que de plus, le DB75 favorise cette activité de clivage.

L'effet de suppression de la PEV du gène *white* dans la souche w^{m4} par les diamidines passerait donc par une augmentation de l'activité de la topoïsomérase II sur les séquences SAT III. La présence massive de topoïsomérase II sur ces séquences en limiterait alors l'accessibilité pour d'autres protéines telle que la protéine D1. Cette dernière ne se trouvant plus sur les séquences SAT III, elle ne pourrait plus y recruter les protéines de mise en place de la structure hétérochromatinienne (HDAC1, SU(VAR)3-9 et HP1, (Blattes et al., 2006)). Ainsi, le polyamide P9 supprimerait la PEV dans la lignée *white-mottled* en déplaçant la protéine D1 et en perturbant le recrutement des protéines de mise en place de la structure hétérochromatinienne, alors que les diamidines agiraient en saturant les séquences SAT III en topoïsomérase II, entraînant un déplacement indirect de la protéine D1 et de ces partenaires.

De même, il est intéressant de comparer la suppression de la PEV du gène *white* dans le cas de l'utilisation des diamidines (Susbielle et al., 2005) ou du VM26 (Blattes et al., 2006). Le VM26 agit en séquestrant de manière réversible la topoïsomérase II sur les séquences SAT III et cette accumulation pourrait interférer avec la fixation des protéines telles que HP1. L'effet des diamidines, comme il en est discuté précédemment, entraînerait une accumulation de la topoïsomérase II active sur les séquences SAT III, alors que dans le cas du VM26, la protéine est temporairement inactivée. Ces deux suppresseurs de variégation semblent donc agir de manière similaire entraînant une accumulation massive de topoïsomérase II sur les séquences SAT III, créant une gène stérique et empêchant donc la fixation d'autres protéines. Cependant, dans un cas la topoïsomérase II est inactive, avec le VM26, ou active, avec les diamidines. Ceci suggère que toute modification de la fonction de la topoïsomérase II sur les

séquences SAT III se traduit par une suppression de la PEV du gène *white*. Cela place donc l'activité de clivage de la topoïsomérase II au centre de la régulation de la structure de l'hétérochromatine de cette région.

IV. De l'importance des séquences satellites et des éléments frontières dans la variégation d'effet de position « A molecular basis for variegating position effects »

1. Introduction

Les travaux précédents de l'équipe ont permis de mettre en évidence le rôle de la protéine D1 dans la localisation de protéines telles que HP1 ou la topoïsomérase II au niveau des séquences SAT III de la drosophile. Ces études ont été menées grâce à l'utilisation de molécules synthétiques telles que le polyamide P9 dont nous avons caractérisé les propriétés (Blattes et al., 2006). Il avait déjà été montré que le polyamide P9 était un suppresseur de la PEV de l'allèle *white-mottled* (Janssen et al., 2000a).

Nos résultats précédents nous ont permis de démontrer que l'utilisation du polyamide P9 était capable d'affecter la localisation de la protéine D1. Cette délocalisation a pour conséquence la perte de la protéine HP1 des séquences SAT III mais pas du locus rDNA adjacent.

D'autre part, les mécanismes par lesquels la protéine HP1 est recrutée au niveau des séquences SAT III et sur le rDNA sont différents. Nous avons montré qu'au niveau des séquences SAT III, ce recrutement est dépendant de la protéine D1, et requiert probablement, comme dans le cas de la topoïsomérase II, la présence de la queue C-terminale acide de la protéine D1. Par contre, le recrutement de la protéine HP1 sur le rDNA est indépendant de la présence de la protéine D1. Ce mécanisme pourrait refléter un aspect de la régulation de la transcription des gènes ribosomiques. En effet, l'inactivation de certaines copies des gènes ribosomiques passe par la mise en place d'une structure de type hétérochromatinienne (Prior et al., 1983; Santoro et al., 2002). Des résultats obtenus précédemment nous avait permis de démontrer qu'en présence de TSA, un inhibiteur des HDAC, la protéine HP1 était déplacée des sites hétérochromatiniens, et plus précisément des séquences SAT III et du rDNA, et que la protéine D1 envahissait ces régions laissées vacantes (Blattes et al., 2006). Il existerait donc un cloisonnement de ces deux types de séquences à l'intérieur du noyau, hypothèse probable du fait de l'engagement du rDNA dans la formation du nucléole en interphase.

De même, la régulation de la transcription des gènes ribosomiques pourrait faire intervenir des petits ARN, comme cela a déjà été démontré chez l'homme (Mayer et al., 2006). Ces petits ARN transcrits à partir des séquences SAT III ont été mis en évidence à la suite d'une recherche à grande échelle de petits ARN exprimés de manière dépendante du développement (Aravin et al., 2003). De nombreuses études ont pu mettre en évidence un rôle des petits ARN dans la recrutement de protéine impliquées dans la mise en place de structure hétérochromatinienne, via le recrutement de la machinerie d'interférence à ARN par une protéine à CD (voir figure 26 et (Pal-Bhadra et al., 2004)). Cette protéine liant l'ARN et possédant un CD pourrait être HP1. En effet, il a été récemment démontré que la protéine HP1 possédait une telle activité de liaison à l'ARN (Muchardt et al., 2002).

Le cloisonnement des séquences SAT III et rDNA dans deux compartiments différents mais adjacents suggère la présence de séquences capables de séparer les régions hétérochromatiniennes et euchromatiniennes. La présence de telles séquences, appelées éléments frontières ou isolateurs, avait été prédite dès 1939 (Schultz, 1939). Un élément frontière localisé entre la partie hétérochromatine et euchromatine du chromosome X permettrait, dans un contexte normal, de protéger les gènes euchromatiniens à proximité de l'hétérochromatine de la mise en silence médiée par cette structure. De la même manière, un élément frontière placé dans un tel contexte, pourrait protéger le locus du rDNA des effets de position générés par la proximité de l'euchromatine. En effet, il est possible que l'expression des gènes ribosomiques nécessite un contexte hétérochromatinien pour sa régulation normale.

L'hypothèse de la présence d'une telle séquence est d'autant plus fondée que dans le contexte *white-mottled*, l'inversion pourrait déplacer cet élément frontière, et ainsi expliquer l'effet de l'hétérochromatine dans le patron d'expression soumis à PEV du gène *white*. Une séquence possédant les propriétés d'élément frontière, le BE28, avait été décrite comme étant présente sous forme de cluster dans cette région mais sa localisation précise n'était cependant pas connue (Cuvier et al., 2002).

2. Objectifs de l'étude

Les expériences réalisées dans une lignée sauvage suggéraient que l'effet de suppression de la variégation par le polyamide P9 passait par une délocalisation de la protéine

D1 des séquences SAT III. Une délocalisation de cette protéine a pour conséquence une délocalisation de la protéine HP1 de ces mêmes séquences mais pas du rDNA adjacent, mis à l'écart par formation du nucléole. Cet effet expliquerait le mécanisme à l'origine de l'expression relativement normale du gène *white* dans ce contexte particulier. Nous avons donc voulu vérifier cette hypothèse. De même, nous avons voulu vérifier l'effet du polyamide P9 sur la localisation de la protéine HP1 dans les mêmes conditions. La protéine HP1 étant normalement localisée sur les séquences SAT III et sur le rDNA adjacent, un traitement avec le polyamide P9 peut affecter la localisation de la protéine HP1 sur les séquences SAT III mais aussi sur le rDNA suite à au déplacement de la protéine D1.

Afin d'expliquer le phénomène de cloisonnement observé lors de la délocalisation de la protéine HP1 par le polyamide P9, la localisation de l'élément frontière BE28 a été recherchée grâce à des expériences d'hybridation *in situ* en fluorescence (FISH) sur chromosome mitotiques. Cette technique a permis de localiser les régions hétérochromatiniennes avec une précision plus grande que celle obtenue généralement sur les chromosomes polytènes, où ces régions sont sous-répliquées.

La perte de cette séquence dans l'inversion *white-mottled* pourrait expliquer pourquoi le gène *white* est affecté par la structure hétérochromatinienne. Un tel déplacement pourrait également exercer des effets sur la régulation de la transcription des gènes ribosomiques. Enfin, la perte de l'élément frontière BE28 pourrait affecter les gènes euchromatiniens autant que les gènes hétérochromatiniens. Afin de vérifier ces hypothèses, des expériences de Nortern-blot ont été réalisées.

De même, nous avons voulu savoir si les petits ARN dérivés des séquences SAT III (SATIII-si) jouaient un rôle dans la régulation de l'expression des gènes ribosomiques. Dans un premier temps, l'existence de ces SATIII-si a été recherchée par des expériences de Northern-blot. Dans un deuxième temps, le cas échéant, nous avons recherché leur présence à proximité du ou dans le nucléole grâce à des expériences de FISH contre l'ARN. Cette localisation permettrait de confirmer notre hypothèse de régulation des gènes ribosomiques par des petits ARN issus des séquences SAT III.

3. *Résultats*

a. Effet du traitement par le polyamide P9 sur la localisation des protéines D1 et HP1 dans la lignée *white-mottled*

Nos travaux précédents ont mis en évidence la capacité du polyamide P9 à rentrer en compétition avec des protéines endogènes, telles que les protéines D1 et topoïsomérase II, au niveau des séquences SAT III (figure 40, (Blattes et al., 2006) et voir partie précédente).

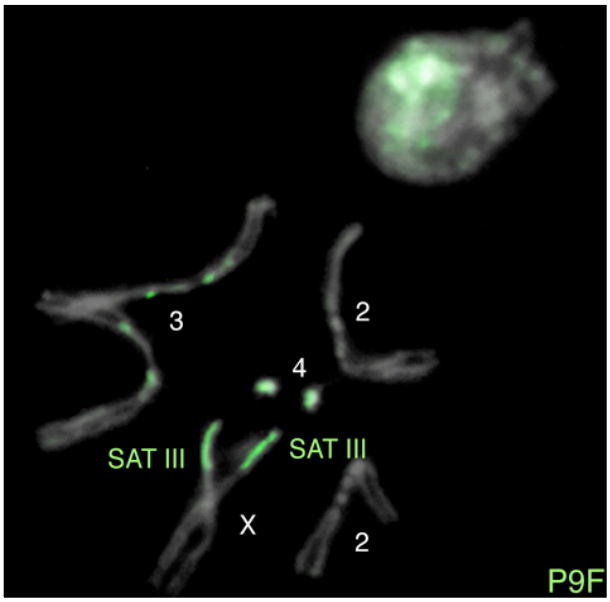
Il a donc été supposé que l'effet de suppresseur de variégation du polyamide P9 impliquait le déplacement de protéines reconnaissant spécifiquement les séquences ciblées. Pour autant, ce mécanisme n'avait jamais été vérifié dans de la lignée *white-mottled*. En effet, outre le phénotype bien connu de variégation au niveau des yeux, ces mouches présentent un cycle de vie plus long que celui des mouches sauvages et de plus, elles semblent se reproduirent beaucoup plus difficilement, ce qui les rends plus difficile à élever.

Afin de confirmer cette hypothèse, j'ai tout d'abord réalisé des immunolocalisations avec des anticorps dirigés contre les protéines D1 et HP1 dans la lignée sauvage Oregon R (voir figure 41). Dans ce contexte, la protéine D1 se localise sur les séquences SAT III. De même, la protéine HP1 est localisée massivement au niveau du rDNA. Un marquage est présent sur les séquences SAT III, mais seulement au niveau de la partie distale de ces régions. Ces immunolocalisations ont été ensuite comparées aux résultats obtenus dans le contexte de la lignée *white-mottled*.

Nous pensions que le polyamide P9 allait également entraîner une délocalisation des protéines D1 et HP1 des séquences SAT III dans un contexte *white-mottled*. En effet, dans un contexte sauvage, un traitement au polyamide P9 entraîne un déplacement de la protéine D1 des séquences SAT III et une délocalisation de la protéine HP1 au niveau de ces séquences, mais pas du rDNA (Blattes et al., 2006).

J'ai donc réalisé des expériences de déplacement de protéines *ex vivo* avec ou sans traitement avec le polyamide P9 à partir de larves *white-mottled*.

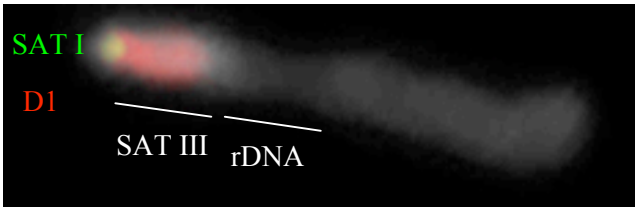
Figure 40 : localisation des séquences satellites sur des chromosomes mitotiques grâce à l'utilisation du polyamide P9.



Des cerveaux disséqués à partir de larves femelles de troisième stade de type sauvage (Oregon R), sont fixés et colorés en utilisant le polyamide fluorescent P9F. Ce polyamide est un dérivé du polyamide P9 directement couplé à la fluorescéine. Il est synthétisé pour reconnaître spécifiquement les séquences SAT I et SAT III.

L'image de microscopie à fluorescence présentée ici est un exemple de marquage obtenu avec le polyamide P9F (en vert), fortement spécifique des séquences SAT I et SAT III. Sur les chromosomes : le marquage du chromosome X correspond majoritairement aux séquences SAT III, seule la pointe de ce chromosome présente des séquences SAT I, comme les autres chromosomes marqués (paires 3 et 4). Sur le noyau interphasique : le marquage se présente sous la forme d'un focus relativement intense, correspondant probablement à la majorité des séquences SAT III et SAT I. Cette localisation se superpose avec le point brillant obtenu avec un marquage au DAPI (en gris) et décrit comme étant une zone hétérochromatinienne.

A.



B.

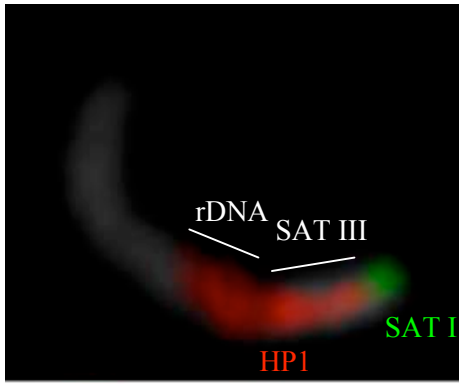


Figure 41 : immunolocalisations des protéines D1 et HP1 sur chromosomes mitotiques de larves sauvages.

Les chromosomes mitotiques issus de neuroblastes larvaires sont marqués avec un anticorps dirigé contre la protéine D1 (en rouge, **A.**) ou la protéine HP1 (en rouge, **B.**), et contremarqués avec le polyamide Lex9F (en vert, **A. et B.**). Ce polyamide reconnaît spécifiquement les séquences SAT I mais très peu les séquences SAT III (Blattes *et al*., 2006). Dans les deux cas, les chromosomes sont contremarqués avec du DAPI (en gris). Les marquages contre les protéines D1 et HP1 semblent former un gradient de sens opposé suggérant une interaction mutuellement exclusive des ces deux protéines avec les séquences SAT III.

Dans un premier temps, j'ai recherché l'effet d'un traitement avec le polyamide P9 sur la localisation de la protéine D1. Sans traitement et dans les conditions des ces expériences, la protéine D1 est localisée sur les séquences SAT III (voir figure 42A). Après un traitement avec 2,5 µM de P9, la protéine D1 est non seulement délocalisée des séquences SAT III mais elle est relocalisée massivement sur le rDNA (voir figure 42B). Contrairement à ce qui était attendu, la protéine D1 n'est donc pas perdue au niveau des chromosomes, mais est relocalisée.

Dans un deuxième temps, les mêmes expériences ont été reproduites, mais en recherchant cette fois les effets du changement de localisation de la protéine D1 sur la protéine HP1. Pour cela, la localisation de la protéine HP1 a été comparée sans (figure 42C) ou suite à un traitement avec 2,5 µM de P9 (figure 42D).

Sans traitement, le marquage obtenu pour la protéine HP1 est semblable à celui observé dans un contexte sauvage sur les séquences SAT III (comparer figure 41B et figure 42C). Pourtant, il apparaît un nouveau point de marquage pour la protéine HP1. Celui-ci se localise sur l'extrémité distale du bras euchromatinien du chromosome X. Ce marquage pourrait correspondre au point distal de cassure de l'inversion *white-mottled*. Cette région serait donc capable de recruter la protéine HP1 et de créer une région ectopique d'hétérochromatine.

Après traitement au P9, la protéine HP1 n'est plus localisée au niveau du rDNA mais est maintenant présente sur les séquences SAT III, à la place de la protéine D1 (voir figure 42D). Par contre, le point de marquage ectopique n'est pas affecté par le traitement avec P9.

Ces résultats aussi surprenant qu'inattendus suggèrent que les protéines D1 et HP1 interagissent de manière mutuellement exclusive avec les séquences SAT III mais aussi avec le rDNA sur des chromosomes portant l'inversion w^{m4}. Par contre, le point de marquage ectopique observé en utilisant un anticorps contre la protéine HP1 suggère la formation d'une structure de type hétérochromatinienne en conséquence de l'inversion *white-mottled*. Ces résultats supposent l'existence de déterminants différents entre les contexte sauvage et *white-mottled*, qui permettraient d'expliquer la différence d'effet du traitement par le polyamide P9 dans ces deux lignées.

Une possibilité intrigante est celle émise par J. Schultz, prédisant l'existence de séquences capables de stopper l'expansion de la structure hétérochromatinienne et de séparer

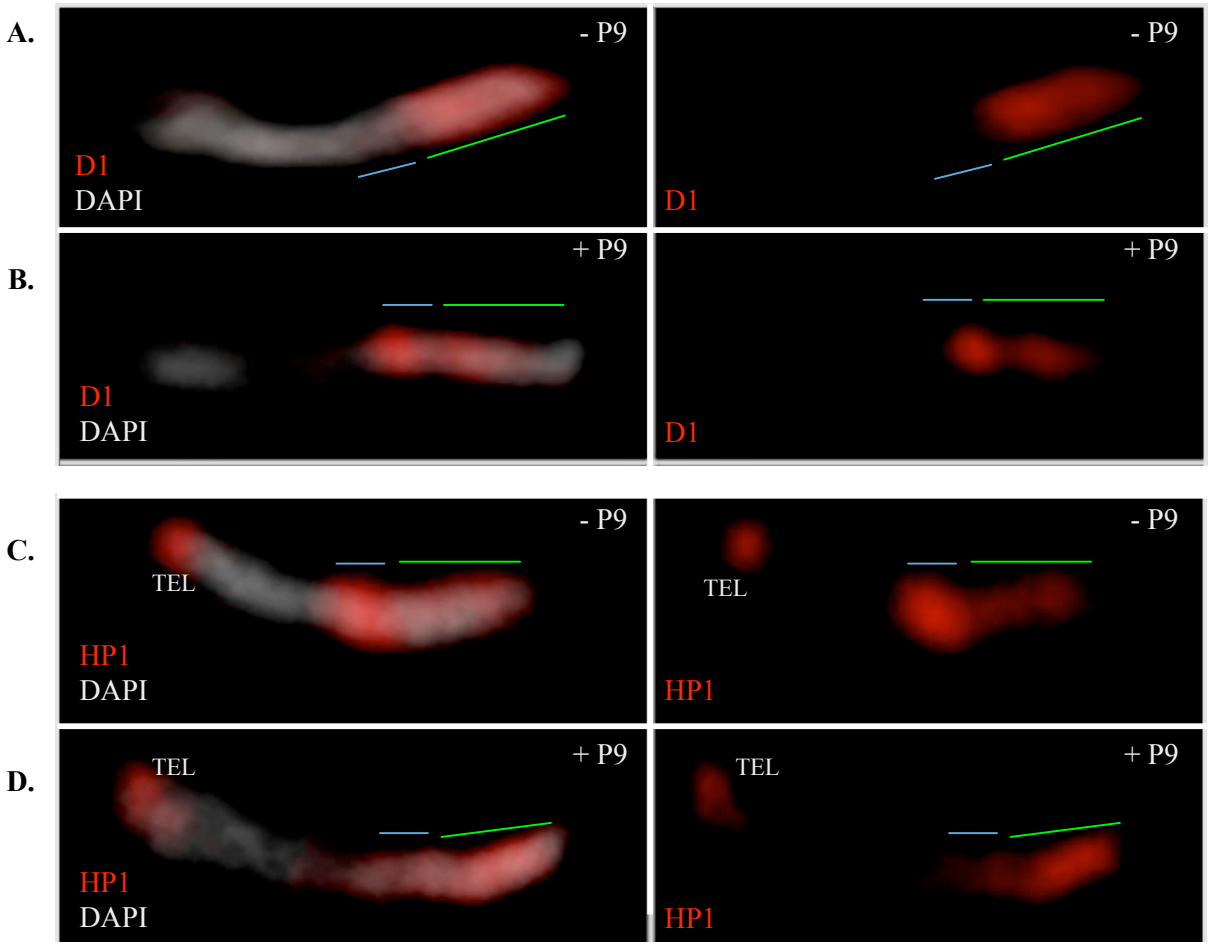


Figure 42 : effet du traitement au P9 sur la localisation des protéines D1 et HP1 dans des larves white-mottled 4h.

Des cerveaux de larves de troisième stade de la lignée *white-mottled* sont perméabilisés, traités (B. et D.) ou non (A. et B.) avec du polyamide P9 à 2,5 µM. Des immunolocalisations contre les protéines D1 (en rouge, A. et B.) et HP1 (en rouge, C. et D.) sont ensuite effectuées, et les chromosomes sont contremarqués avec du DAPI.

Des barres vertes et bleues indiquent respectivement les positions des séquences satellites de type III (en vert) et du locus des gènes ribosomiques (rDNA, en bleu).

A. et B. : sans traitement, la localisation de la protéine D1 correspond à ce qui est décrit dans la lignée sauvage (voir figure 42) alors qu'après traitement, la protéine D1 est déplacée des séquences SAT III et occupe maintenant la place habituelle de la protéine HP1 sur le rDNA.

C. et D. : sans traitement, la localisation de la protéine HP1 correspond à ce qui est décrit dans la lignée sauvage (voir figure 42) excepté au niveau de l'extrémité du chromosome X où un nouveau point de marquage apparaît (noté TEL pour région proche du télomère). Après traitement, la protéine HP1 est déplacée du rDNA et prend maintenant la place de la protéine D1 sur les séquences SAT III. Le nouveau point de marquage n'est pas affecté par ce traitement.

l'hétérochromatine de l'euchromatine (Schultz, 1939). Une telle séquence, si elle existe, pourrait donc être inversée dans la lignée *white-mottled* et sa perte serait à l'origine des différences observées.

b. Localisation de séquences « marqueurs » à proximité des points de cassures dans la lignée *white-mottled*

Afin de mieux décrire les effets de l'inversion au niveau de la localisation des protéines étudiées, il est tout d'abord important de savoir quelles sont les séquences affectées par l'inversion. Pour cela, j'ai choisi de m'intéresser aux séquences proches de chaque extrémité des points de cassures décrits pour l'inversion *white-mottled*.

D'un côté de la cassure, il est connu depuis longtemps que les gènes *white* et *Notch* étaient inversés (Schultz, 1936). La séquence du gène *white* a donc été utilisée en tant que sonde dans des expériences de FISH afin de servir de marqueur pour une des extrémités de l'inversion. De l'autre côté de la cassure, à proximité du rDNA, des gènes ont été caractérisés comme le gène *su(f)* (Yamamoto et al., 1990). J'ai cherché à savoir si ce gène était lui aussi inversé dans la lignée *white-mottled*. J'ai donc réalisé des expériences de FISH en utilisant des sondes contre la séquence génomique des gènes *white*, *su(f)* et d'une répétition du gène codant les ARN ribosomiques (figure 43), dont il ne peut être physiquement résolu. Ces résultats sont en accord avec les résultats décrivant le point de cassure de l'inversion dans le locus des gènes ribosomiques et plus précisément dans un insertion de type I (Ebert et al., 2004).

Sur des chromosomes X de la lignée sauvage, le gène *white* est localisé à l'extrémité du bras euchromatinien (figure 43A). Dans la lignée *white-mottled*, suite à l'inversion le gène *white* est maintenant localisé comme attendu à proximité immédiate du rDNA (figure 43B). Les mêmes expériences ont été réalisées avec une sonde dirigée contre le gène *su(f)*. Dans la lignée sauvage, le gène *su(f)* est localisé à proximité du rDNA (figure 43C). Dans l'inversion, le gène *su(f)* est relocalisé sur la partie distale du bras euchromatinien du chromosome X (figure 43D).Ces résultats montrent donc que dans l'inversion *white-mottled*, la cassure intervient dans une région comprise entre le gène *su(f)* et le rDNA. Le gène *su(f)* est donc un bon marqueur de cette extrémité de l'inversion.

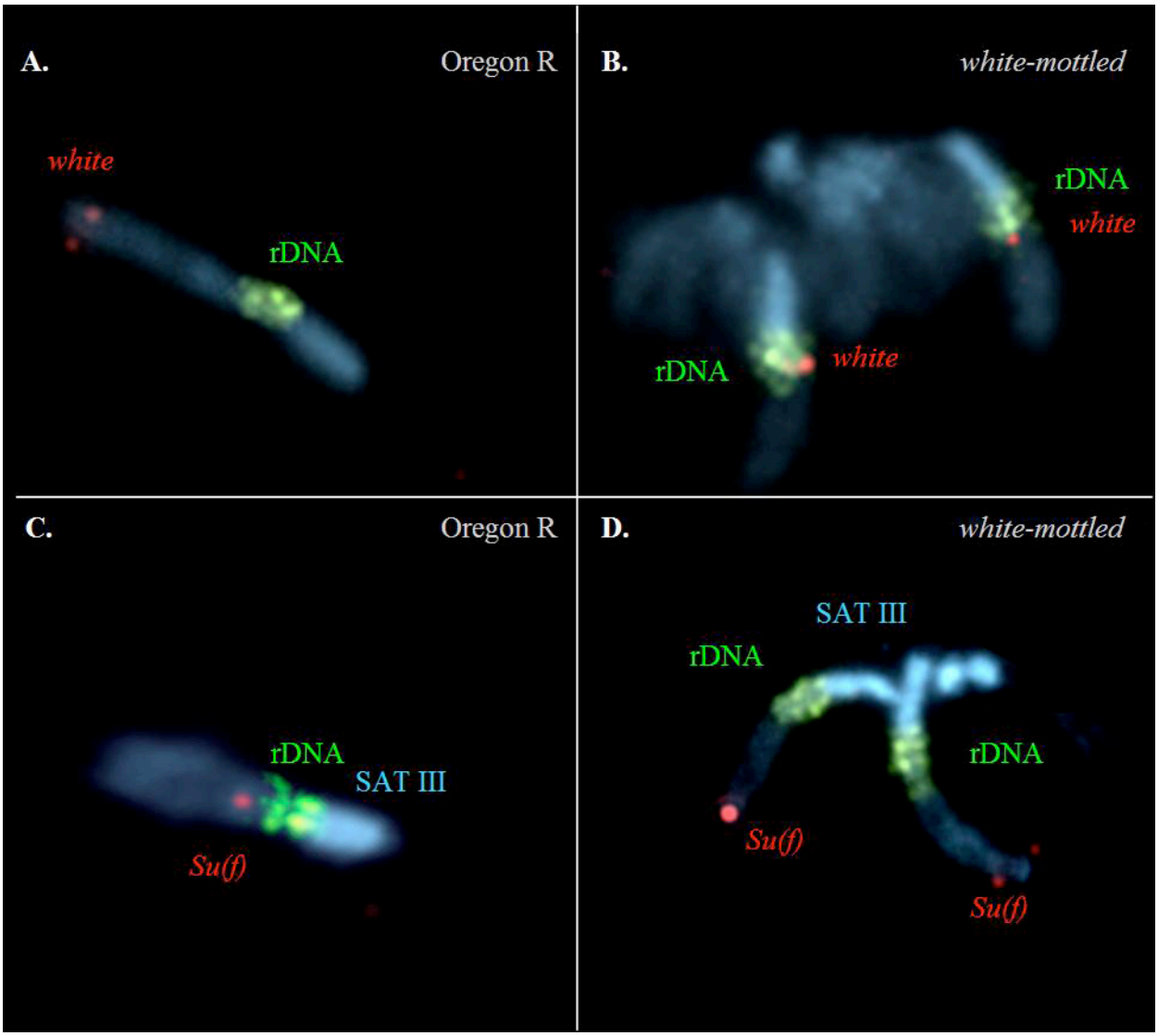


Figure 43 : FISH contre le locus des gènes ribosomiques et contre les gènes white et su(f) sur chromosomes mitotiques.

Dans ces expériences d'hybridation *in situ* de fluorescence (FISH), des sondes contre les séquences génomiques du gène *white* (en rouge, **A. et B.**) ou du gène *su(f)* (en rouge, **C. et D.**) et contre le locus des gènes ribosomiques (vert) sont utilisées. La localisation du gène *white* a été recherchée dans des larves de type sauvage (Oregon R, **A. et C.**) ou de type *white-mottled* (**B. et D.**).

Dans la lignée sauvage, le gène *white* est sur l'extrémité distale du bas euchromatinien du chromosome X (**A.**). Dans l'inversion *white-mottled* (**B.**), le signal correspondant au gène *white* est localisé à proximité du locus des gènes ribosomiques (rDNA), voire même dans son extrémité centromérique-distale.

Inversement, dans la lignée sauvage (**C.**), le gène *su(f)* (en rouge) est localisé à proximité du rDNA (en vert). Dans la lignée *white-mottled* (**D.**), le gène *su(f)* est relocalisé à l'autre extrémité du chromosome X

c. L'élément frontière BE28 est déplacé par l'inversion chromosomique *white-mottled*

L'hypothèse émise par J. Schultz suggère l'existence d'un élément frontière permettant de séparer l'hétérochromatine péricentromérique du chromosome X et la partie euchromatinienne (Schultz, 1939). Un élément frontière, le BE28, a été décrit comme étant localisé dans cette région grâce à des expériences de FISH sur des chromosomes polytènes (Cuvier et al., 2002).

J'ai donc effectué des expériences de FISH sur chromosomes mitotiques afin de localiser plus précisément cette séquence (voir figure 44). L'élément frontière BE28 correspond à une séquence de 269 pb présente au sein d'une séquence plus large de 1,2 kb. Cette dernière est répétée environ une trentaine de fois sur le chromosome X dans la région séparant l'hétérochromatine de l'euchromatine ((Cuvier et al., 2002) et O. Cuvier, communication personnelle).

Je me suis servie de la séquence de 1,2 kb comme sonde dans les expériences de FISH. Dans ces mêmes expériences, j'ai utilisé des sondes dirigées contre les séquences SAT III et contre une copie d'une unité de transcription codant les rRNA.

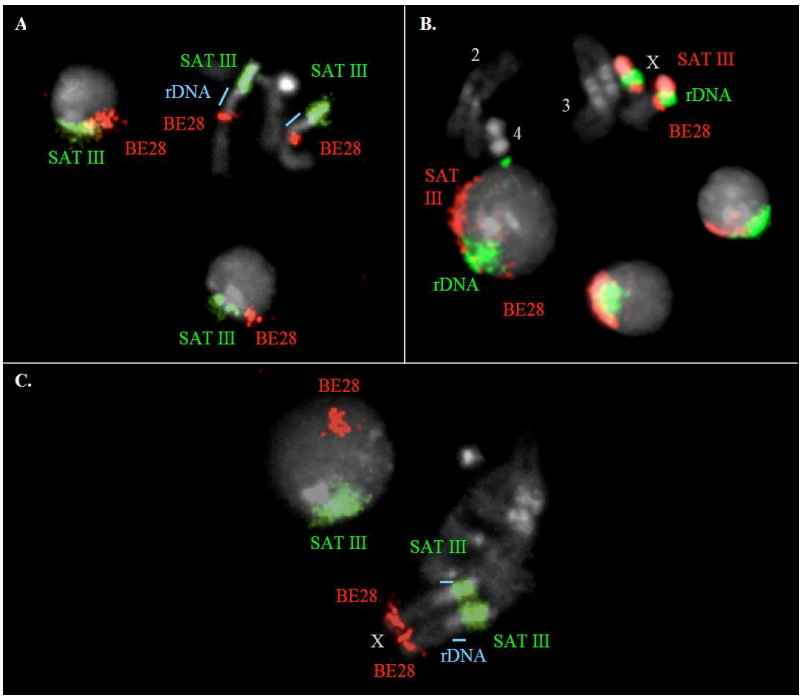
J'ai ainsi pu montrer que l'élément frontière était localisé, sur un chromosome X de lignée sauvage, entre le rDNA et le bras euchromatinien (figure 44A). Les mêmes expériences ont été menées sur des chromosomes X de la lignée *white-mottled* (figure 44B).

Les résultats démontrent que l'élément frontière BE28 est déplacé dans l'inversion chromosomique *white-mottled*. Pour compléter la cartographie des séquences de cette région, j'ai réalisé des expériences de FISH en utilisant des sondes dirigées contre l'élément frontière BE28 et le gène *su(f)* (figure 45).

Des données préalables supposaient que le gène *su(f)* se trouvait localisé à proximité du rDNA, c'est-à-dire dans la partie hétérochromatinienne, et non pas du côté du bras euchromatinien (O'Hare et al., 2002; Yamamoto et al., 1990). Des données plus récentes on montré que le gène *su(f)* était séparé du rDNA par une séquence de 5,9 kb répétée en tandem (O'Hare et al., 2002). Cependant, l'activité de frontière de cet élément n'a pas encore été confirmée.

Dans la lignée sauvage, il apparaît que le gène *su(f)* est effectivement localisé entre l'élément frontière et le rDNA (figure 45A). Conformément à ce qui est attendu, après

Figure 44 : FISH contre les séquences SAT III et l'élément frontière BE28 sur chromosomes mitotiques et noyaux interphasiques.



Dans ces expériences d'hybridation *in situ* de fluorescence (FISH), des sondes contre les séquences SAT III, la séquence de l'élément frontière BE28 et le locus des gènes ribosomiques (rDNA) sont utilisées. Dans la lignée sauvage (**A.**), le rDNA est encadré par l'élément frontière BE28 (en rouge) d'un côté et par les séquences SAT III (en vert) de l'autre.

La même expérience réalisée sur une lignée sauvage (**B.**), mais avec cette fois avec trois sondes confirme cette observation. Ici, les sondes sont dirigées contre les séquences SAT III (en rouge), le rDNA (en vert) et l'élément frontière BE28 (en rouge). Dans l'inversion *white-mottled* (**C.**), le signal correspondant à l'élément frontière (en rouge) est localisé de l'autre côté sur le chromosome X, et se retrouve donc très éloigné du rDNA (signalé par un trait bleu) et des séquences SAT III (en vert).

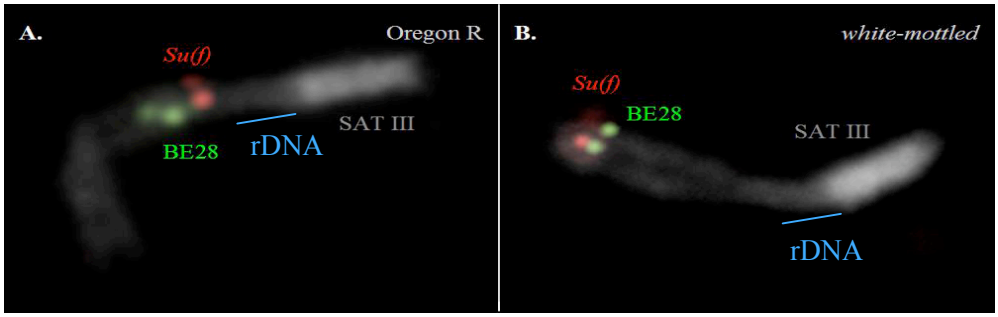


Figure 45 : FISH contre l'élément frontière BE28 et le gène Su(f) sur chromosomes mitotiques.

Dans ces expériences d'hybridation *in situ* de fluorescence (FISH), des sondes contre la séquence de l'élément frontière BE28 (en vert, **A. et B.**), et le gène *su(f)* (en rouge, **A. et B.**) sont utilisées.

Sur la lignée sauvage (**C.**), le gène *su(f)* (en rouge) est localisé entre l'élément frontière BE28 (en vert) et le locus des gènes ribosomiques. Dans l'inversion *white-mottled* (**D.**), les deux séquences sont inversées et le gène *su(f)* est plus près de la région télomérique que l'élément frontière. La localisation du gène *su(f)* mise en évidence ici tend donc à le caractériser comme un gène hétérochromatinien.

inversion le gène *su(f)* et localisé entre l'élément frontière et la région télomérique. Dans les deux contextes, le gène *su(f)* n'est donc pas dans la partie euchromatinienne du chromosome X mais pourrait constituer un nouveau membre de la famille réduite des gènes hétérochromatiniens.

d. Effet de la perte de l'élément frontière sur la transcription des gènes ribosomiques

La régulation de la transcription des gènes ribosomiques fait intervenir, entre autre, des déterminants s'apparentant à la machinerie de mise en place de la structure hétérochromatinienne. Les nucléosomes présents au niveau des copies actives contiennent le variant d'histone H3.3 (Ahmad and Henikoff, 2002; Schwartz and Ahmad, 2005) alors qu'au niveau des copies inactives, la lysine 9 de l'histone H3 est méthylée (Prior et al., 1983; Santoro et al., 2002). Nous avons donc recherché l'effet de certaines mutations suppresseur de variégation sur la transcription des gènes ribosomiques. La perte de l'élément frontière BE28 dans le contexte *white-mottled* pourrait en effet perturber la transcription des gènes ribosomiques.

J'ai donc réalisé des expériences de Northern-blot en utilisant des sondes capables de reconnaître différentes formes des ARN ribosomiques (voir figure 46 et (Fox, 1997; Giordano et al., 1999; Long and Dawid, 1980)). Contrairement à ce qui est décrit dans les autres organismes, la drosophile possède la particularité de présenter un ARN « mature » de 28S scindé en deux parties, le 28Sa et le 28Sb, qui restent cependant associées au sein des ribosomes. L'étape de maturation du 28S en 28Sa et 28Sb enlève une courte séquence d'ARN de 27 nucléotides. Cette particularité permet de détecter tous les ARN précurseurs de l'ARN de 28S mature en utilisant une sonde dirigée contre la partie clivée. Les autres sondes ont été choisies afin de détecter des précurseurs précoces (sonde ETS1), plus tardifs (sonde ITS1) ou la forme mature (sonde 18S).

En comparant les différentes pistes sur chaque membrane, il semble que la transcription des gènes ribosomiques n'est visiblement pas affectée dans la lignée *white-mottled*. Par contre, il semblerait que la transcription des gènes ribosomiques soit diminuée dans les lignées *wm^{4h}/Su(var)2-5* et *wm^{4h}/Su(var)3-9*. Ces lignées correspondent à des lignées

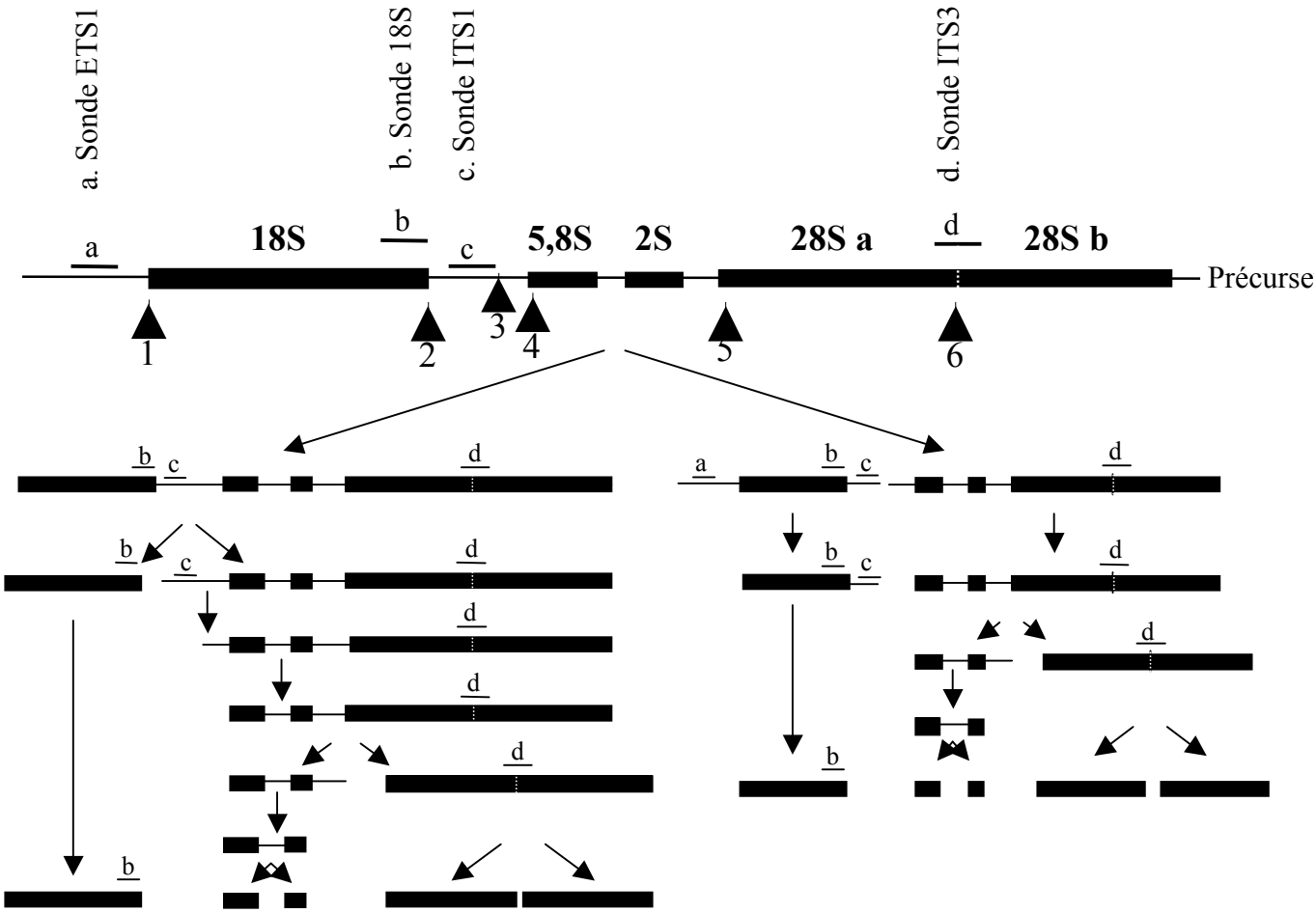


Figure 46 : schéma de maturation de l'ARN ribosomique de la drosophile (d'après Long et al., 1980).

La formation des ARN ribosomiques matures nécessite de nombreux clivages mais aussi des modifications post-transcriptionnelles. Sont représentées ici les deux principales voies de maturation décrites chez la drosophile (Long *et al*, 1980), avec les principaux points de clivage (clivages séquentiels du site 1 au site 6).

Contrairement à la plupart des autres espèces, la voie de maturation de l'ARN de 28S chez la drosophile présente une particularité. L'ARN ribosomique de 28S est clivé en deux sous-parties qui resteront associées dans la particule ribosomale.

Sur ce schéma, sont représentées les sondes utilisées dans les expériences de Northern-blot. Il est à noter que la sonde appelée ITS3 correspond à la totalité de la partie clivée au sein du 28S.

white-mottled (wm^{4h}) respectivement mutées au niveau de gènes suppresseurs de variégation comme le gène *Su(var)2-5*, qui code la protéine HP1, et le gène *Su(var)3-9* codant l'histone méthyl-transférase spécifique de la lysine 9 de l'histone H3. Cette observation est faite en comparant les signaux obtenus en utilisant les sondes dirigées contre les précurseurs des ARN ribosomiques (figure 47A, C et D) et ceux obtenus en utilisant la sonde dirigée contre la forme mature de l'ARN de 18S (figure 47B).

Ces résultats suggèrent donc l'intervention des protéines HP1 et SU(VAR)3-9 dans la régulation positive de la transcription des gènes ribosomiques. Dans un contexte *white-mottled*, la présence de ces protéines n'affecte pas la transcription des gènes ribosomiques. Par contre, ces protéines influencent l'expression du gène *white*, localisé à proximité de ces séquences. Dans un contexte où ces protéines sont moins présentes (effet de dose suite à la mutation d'un des gènes codant la protéine d'intérêt), le gène *white* semble retrouver une expression normale alors qu'au contraire, il semblerait que la transcription des gènes ribosomiques soit affectée.

e. Rôles des petits ARN issus des séquences SAT III

La caractérisation de ces petits ARN soulève la question de l'initiation de la mise en place de la structure hétérochromatinienne grâce à l'intervention de la machinerie RNAi, comme cela a déjà été décrit chez la levure *S. pombe* (Volpe et al., 2002). Cette intervention a été recherchée chez la drosophile, grâce à des expériences mettant en jeu des phénotypes de variégation. Certes, la PEV de certains gènes *mini-white* peut être supprimée dans des lignées mutantes pour des protéines de la machinerie RNAi (Pal-Bhadra et al., 2004). Mais les lignées utilisées dans cette étude ne sont pas des souches de type *white-mottled*. Elles possèdent des insertions de gènes *mini-white* répétés sur les chromosomes 2 ou 4.

Les tests réalisés sur la lignée *white-mottled* dans un contexte mutant pour les protéines de la machinerie RNAi (PIWI et ARGONAUTE1, mutations transhétérozygotes) ne montrent pas de variation dans le phénotype de PEV du gène *white* (communication personnelle, U. Bhadra). Seule la mutation de DICER altère la variégation du gène *white* mais ne la supprime pas (communication personnelle, U. Bhadra). Ces résultats suggèrent donc qu'il existerait une autre voie d'initiation de l'assemblage de l'hétérochromatine, non directement liée à la machinerie RNAi.

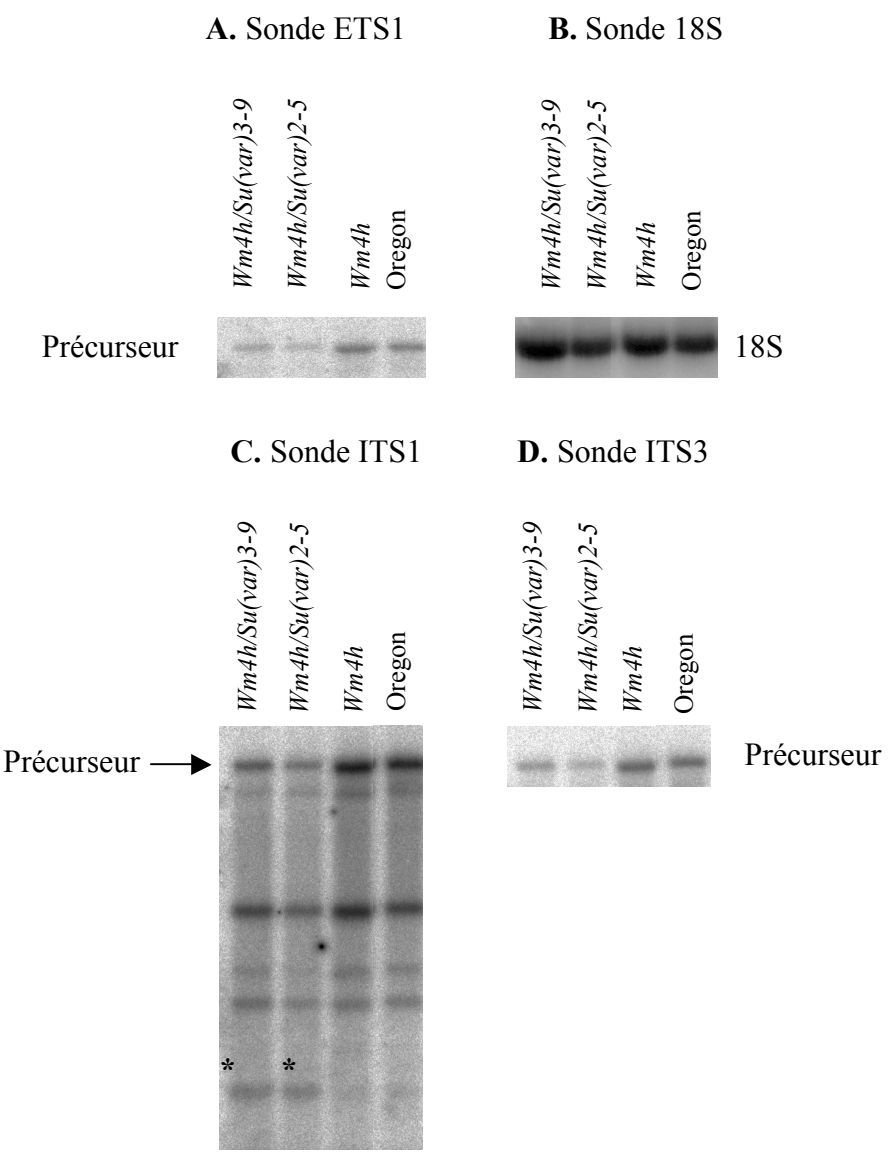


Figure 47 : effet des mutations suppresseurs de variégation sur la transcription des gènes ribosomiques .

Ces figures représentent des membranes de Northern blot hybridées avec différentes sondes (voir figure précédente pour la position des sondes). Quatre types d'ARN totaux ont été séparés sur gel : des ARN extraits à partir de mouches sauvages (Oregon R), de mouches de la lignée *white-mottled* (*wm^{4h}*), et de mouches de la lignée *white-mottled* qui portent des mutations du gène codant la protéine HP1 (*wm^{4h}/Su(var)2-5*) ou la protéine Su(var)3-9 (*wm^{4h}/Su(var)3-9*). En comparant les pistes, il est possible de voir que dans le cas des lignées *wm^{4h}/Su(var)2-5* et *wm^{4h}/Su(var)3-9*, la transcription des gènes ribosomiques semble diminuée. De même, il semblerait qu'une des voies de maturation soit affectées car une espèce, minoritaire dans les autres pistes, est surreprésentée dans les pistes correspondantes (voir l'astérisque dans la partie C.)

Malgré ces résultats, la présence de petits ARN, les SATIII-si, issus probablement de la transcription des séquences SAT III (Aravin et al., 2003) pourrait suggérer l'existence d'un mécanisme impliquant des petits ARN dans la mise en place de la structure hétérochromatinienne. Ces SATIII-si ayant été découvert suite à une approche de clonage, j'ai tout d'abord recherché leur existence dans les cerveaux de larves de drosophile de 3ème stade. Dans un premier temps, je ne me suis intéressée qu'à la plus longue des espèces décrites : une forme de 28 nucléotides que j'appellerai SATIII-si28.

J'ai tout d'abord recherché l'existence du SATIII-si28 grâce à des expériences de Northern-blot. Pour cela, une sonde correspondant à la séquence antisens du SATIII-si28 a été synthétisée puis radiomarquée. Ces expériences ont été réalisées à la fois sur des ARN totaux extraits à partir de larves de type sauvage ou *white-mottled*. Dans les deux cas, la sonde révèle plusieurs espèces qui forment une échelle dont la plus petite bande détectée est de 28 nucléotides (figure 48A). Ces résultats suggèrent l'existence de plusieurs transcrits hétérogènes ou de leurs précurseurs issus de la transcription des séquences SAT III.

Ces résultats préliminaires confirment l'existence *in vivo* de ces SATIII-si28. Même s'il sera difficile d'établir un rôle, s'ils en ont un, pour ces SATIII-si, des données significatives peuvent être obtenues en examinant leur localisation au niveau des noyaux de tissus larvaires. Pour cela, des expériences de FISH contre l'ARN ont été réalisées en utilisant une sonde synthétisée antisens à la séquence décrite pour le SATIII-si28.

En nous basant sur notre hypothèse que les séquences SAT III jouent un rôle dans la régulation des gènes ribosomiques, nous avons supposé que ce SATIII-si28 puisse être impliqué dans ce mécanisme et pourrait donc être accumulé dans le nucléole. Suivant cette hypothèse, nous avons réalisé des co-marquages à la fois contre le SATIII-si28 et contre le snoARN U3, connu pour s'accumuler dans le nucléole. L'utilisation d'un double marquage s'est avéré indispensable du fait de la difficulté à observer la structure nucléolaire par des techniques classiques de microscopie optique (contraste de phase ou Nomarski).

Comme le montre la figure 48B, la sonde dirigée contre le snoARN U3 décore une structure sphérique qui correspond au nucléole. De manière remarquable, la sonde dirigée contre le SATIII-si28 révèle un ou deux foci d'accumulation qui ne sont significativement

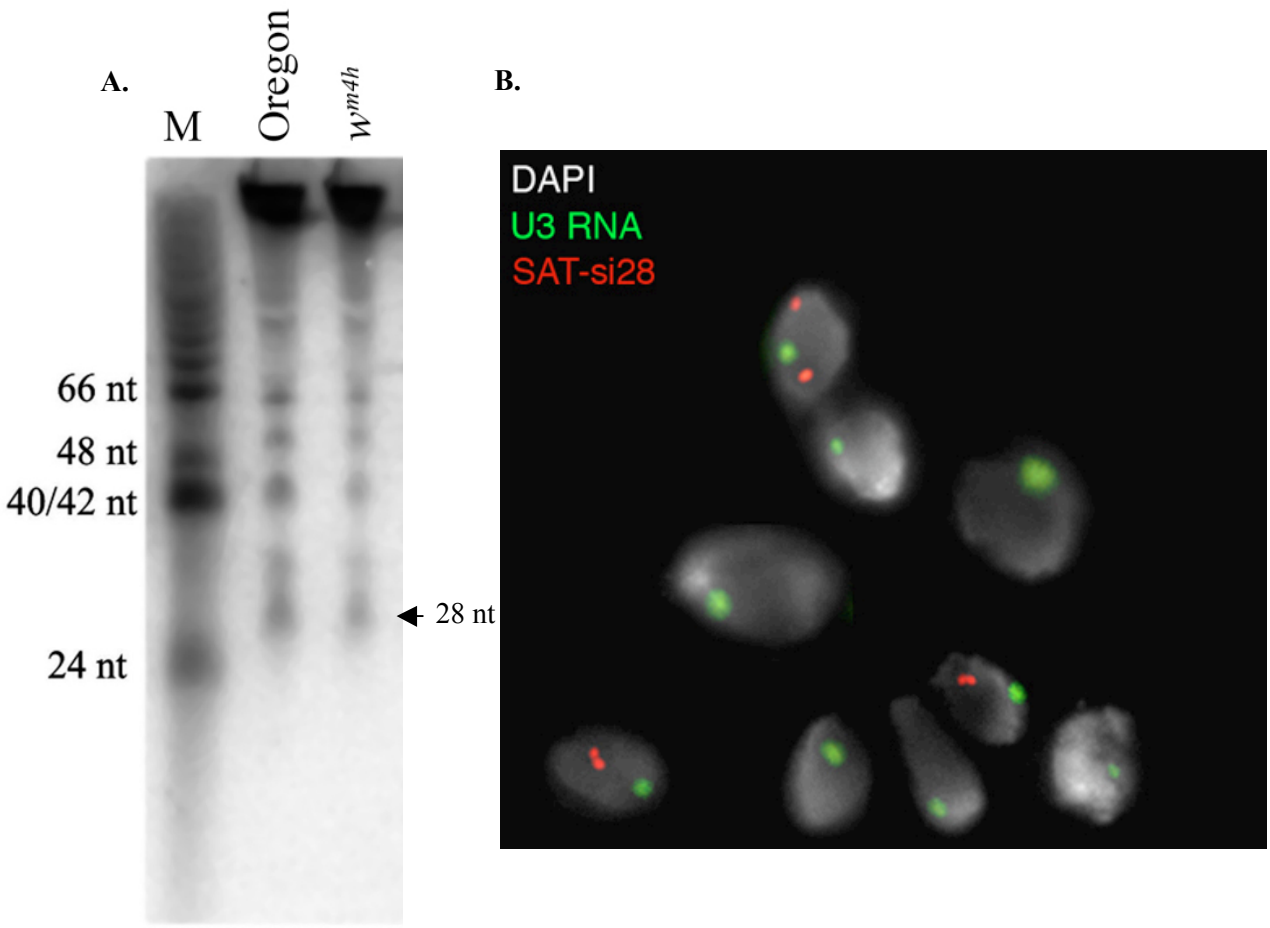


Figure 48 : détection des petits ARN issus des séquences SAT III (SATIII-si) .

A. Northern-blot effectué sur des ARN totaux extraits à partir de larves de troisième stade de la lignée sauvage (Oregon) ou *white-mottled* (*wm*^{4h}).
Dans chaque ligne, une échelle correspondant à des petits ARN apparaît. La bande de plus faible taille correspond à un petit ARN d'environ 28 nucléotides. Le marqueur correspond au plasmide PhiX174 (FERMENTAS) digéré par Hinf I et marqué radioactivement.

B. Hybridation *in situ* de fluorescence (FISH) contre l'ARN.
La sonde contre le SATIII-si28 (en rouge) montre que ces petits ARN forment entre 0 et 2 foci par noyau. Ces foci ne colocalisent pas avec le nucléole révélé par la sonde contre le snoARN U3 (en vert). Ces foci ne présentent pas de localisation définie dans le noyau et peuvent être détectés soit proches, soit loin l'un de l'autre.

Les sondes utilisées dans ces expériences correspondent à la séquence antisens du SATIII-si28 (d'après Aravin *et al.*, 2003), et sont soit marquées radioactivement (**A.**), soit couplées à un fluorochrome rouge (Cy3, **B.**).

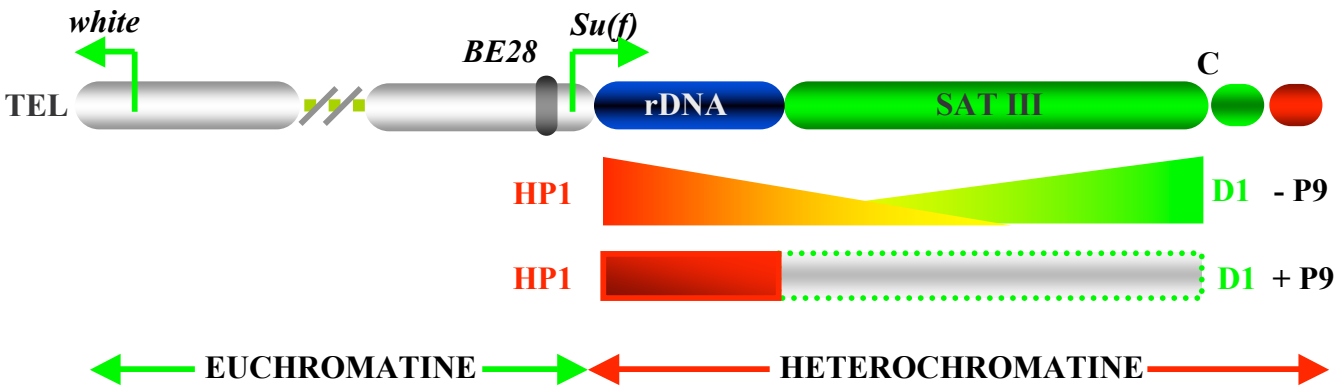
jamais associés à la structure nucléolaire. Ces résultats suggèrent donc que le SATIII-si28 n'est pas impliqué dans la régulation de l'expression des gènes ribosomiques.

4. Conclusions et discussions

Les expériences menées sur la lignée *white-mottled* ont montré que le polyamide P9 n'avait pas le même effets sur les protéines D1 et HP1, par rapport à ce qui avait été décrit sur la lignée sauvage. Dans cette lignée, un traitement au P9 entraîne une délocalisation de la protéine HP1, conséquente à la perte de la protéine D1 des séquences SAT III. Au contraire, dans la lignée *white-mottled*, P9 entraîne un échange de la localisation des deux protéines sur deux domaines adjacents que sont les séquences SAT III et le rDNA. Cette différence pourrait s'expliquer par la perte de l'élément frontière BE28. Cette hypothèse émise en 1939 par J. Schultz a été confirmée par notre étude (voir figure 49). Cependant, le BE28 peut ne pas être le seul élément frontière de cette région. En effet, une séquence de 5,9 kb répétée en tandem a été décrite comme étant localisée entre le gène *su(f)* et le rDNA (O'Hare et al., 2002). Même si l'activité de cette séquence en tant qu'élément frontière n'a pas encore été validée, sa localisation en fait aussi un candidat pour la fonction de frontière entre hétérochromatine et euchromatine. De plus, cette séquence est elle aussi déplacée dans l'inversion *white-mottled*. Cependant, le déplacement du BE28 a plusieurs conséquences qui laissent penser que cette séquence est un bon candidat pour être la frontière entre hétérochromatine et euchromatine.

La relocalisation du BE28 au niveau de l'extrémité du bras euchromatinien va entraîner la formation d'un domaine hétérochromatinien ectopique, caractérisé par la présence de la protéine HP1. Les éléments frontières fonctionnant généralement par paires, la formation d'un tel domaine suppose donc la présence d'un autre élément frontière et permet donc de formuler plusieurs hypothèses. Une possibilité est que l'élément frontière BE28 soit relocalisé « à proximité » d'un autre élément frontière déjà existant dans cette région sub-télomérique et dont le rôle serait de protéger le bras euchromatinien des effets de l'hétérochromatine télomérique. Une telle hypothèse est compatible avec l'apparition d'une nouvelle région très délimitée d'hétérochromatine ectopique dans la partie distale du bras euchromatinien du chromosome X.

A. Oregon



B. *White-mottled*

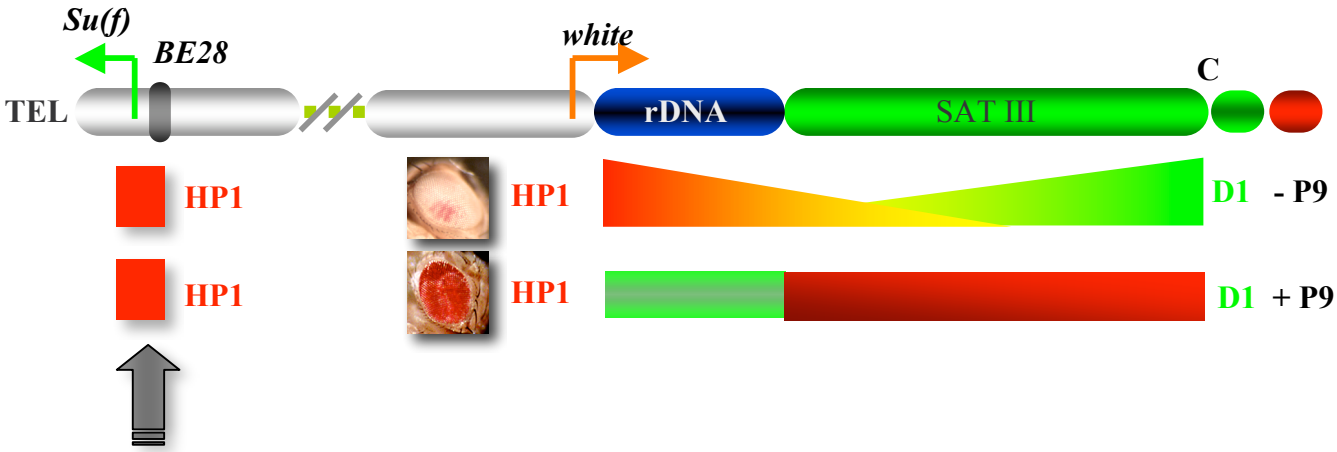


Figure 49 : schémas de chromosomes X de drosophile comparant le positionnement des différentes séquences étudiées et la localisation des protéines HP1 et D1 entre une lignée sauvage et la lignée white-mottled.

A. Lignée sauvage de type Oregon

B. Lignée *white-mottled 4*.

Une autre possibilité est que, lors de la cassure, l'inversion inclue un autre élément frontière qui serait normalement localisé entre le gène *su(f)* et le rDNA. Une séquence qui pourrait avoir une activité de frontière a été décrite dans cette région et son analyse en tant que telle est en cours. Cette séquence est répétée et est riche en paires de bases dA•dT (O'Hare et al., 2002). Une autre évidence qui renforce cette théorie provient de diverses observations réalisées lors de l'analyse des résultats d'expériences de FISH. En utilisant une sonde dirigée contre une répétition des gènes ribosomiques, il est apparu un marquage du côté proximal de l'inversion, à proximité du signal correspondant au gène *su(f)*.

L'une ou l'autre de ces hypothèses est envisageable. Mais en fonction de la taille relative de l'hétérochromatine ectopique formée dans cette région télomérique-proximale, la première hypothèse semble la plus probable. Il est aussi possible que, lors de l'inversion, un élément frontière soit déplacé avec le gène *white*, et que celui-ci ne soit plus actif dans sa nouvelle localisation par perte de son compagnon. Au contraire, l'élément BE28 serait, lui, capable de recréer une frontière active dans sa nouvelle localisation.

La perte de l'élément frontière BE28 qui semble séparer l'hétérochromatine et l'euchromatine pourrait aussi avoir une influence sur la régulation de la transcription des gènes adjacents. Nos résultats semblent déjà montrer que la transcription des gènes ribosomiques n'est visiblement pas affectée dans les lignées *white-mottled*. Ces résultats semblent appuyer l'hypothèse de l'existence d'un autre élément frontière entre le gène *su(f)* et les gènes ribosomiques. Néanmoins, les lignées *white-mottled* sont viables mais leur cycle de vie est plus lent en comparaison à celui de mouches sauvages. Le phénotype de variégation peut être accentué ou allégé en fonction de certains paramètres de culture : température, nombre d'individus par tubes, alimentation…. Par exemple, une augmentation de la température de culture de 18 à 22 °C accentue le phénotype de variégation, les mouches ayant les yeux pâles sont plus nombreuses, et entraîne une diminution du nombre d'individus viables. Il serait donc intéressant de reproduire ces expériences en faisant varier certaines conditions de culture.

Un autre résultat important nous a permis de montrer que la perte des protéines HP1 et SU(VAR)3-9 avait un impact négatif sur la régulation de la transcription des gènes ribosomiques. Les lignées utilisées, wm^{4h}*/Su(var)2-5* et wm^{4h}*/Su(var)3-9*, présentent un contexte chromatinien différent d'une lignée sauvage. Dans ces deux lignées, il semblerait que la présence de la protéine HP1 soit nécessaire à la transcription normale des gènes

ribosomiques. L'effet semble moins marqué dans la lignée *wm^{4h}/Su(var)3-9* car l'absence de l'activité de cette protéine semble pouvoir être compensée par l'activité d'autres histone méthyltransférases telles que la protéine G9a (voir introduction). Cependant, il serait aussi intéressant de reproduire ces expériences en faisant également varier des paramètres de culture, comme pour les lignées *white-mottled*.

La diminution de la transcription des gènes ribosomiques en absence de la protéine HP1 pourrait s'expliquer par un mécanisme impliquant, au contraire, une augmentation de l'accessibilité du locus. L'absence de la protéine HP1 augmenterait le nombre de gènes actifs, et augmenterait donc le nombre de cibles pour les complexes de transcription. Le mécanisme impliquant le facteur UBF indique que ce facteur resterait associé aux promoteurs des gènes actifs. Ceci pourrait conduire à une titration de ces protéines, résultant, contradictoirement, à une perte d'efficacité de ces facteurs de transcription. De même, la différence de structure nucléosomale entre gènes transcrits et non transcrits (Toussaint et al., 2005) peut expliquer l'importance de la présence de la protéine HP1. En effet, si une majorité des gènes présente une structure nucléosomale de type active, la structure globale du nucléole pourrait être altérée et donc générer des défauts de croissance.

Dans ces lignées *wm^{4h}/Su(var)2-5* et *wm^{4h}/Su(var)3-9*, nous avons aussi mis en évidence l'augmentation du signal pouvant correspondre à un précurseur des ARN ribosomiques de 5,8S, 2S et 28S. Cette accumulation pourrait traduire un ralentissement dans la voie de maturation de ces formes, peut être du à l'absence d'un facteur essentiel à cette étape de maturation.

Cette dérégulation de la transcription des gènes ribosomiques génèrerait des petits ARN qui pourraient être à la base du recrutement de la protéine HP1 sur ces séquences. Nous avons mis en évidence que le recrutement de la protéine HP1 sur le rDNA ne dépendait pas de la présence de la protéine D1 mais impliquait un autre mécanisme. La protéine HP1 pourrait donc soit interagir directement avec l'ADN au niveau du rDNA, mais il n'est pas exclu que des formes immatures des ARN ribosomiques soient impliqués dans son recrutement. Cette hypothèse soulève donc la question du rôle essentiel de la protéine HP1 dans la régulation des gènes ribosomiques. La protéine HP1, qui serait donc impliquée directement dans le contrôle de l'expression des gènes ribosomiques, assurerait donc un contrôle fortement régulé du nombre de copies de gènes actifs.

Nous avons vu que la structure hétérochromatinienne, et plus exactement les protéines associées à l'hétérochromatine, pouvait jouer un rôle dans la régulation des gènes ribosomiques. Au cours de ce travail, je me suis aussi intéressée au rôle éventuel des SATIII-si dans ce même mécanisme de régulation. Les résultats que j'ai obtenu ont permis de valider leur existence, et plus particulièrement celle de la forme la plus longue qui avait été décrite : le SATIII-si28.

Les résultats préliminaires obtenus par les expériences de FISH contre l'ARN établissent deux informations importantes. Tout d'abord, le SATIII-si28 s'accumule de manière frappante sous forme de foci, dont le nombre peut varier de aucun à deux par noyaux. Ensuite, contrairement à ce que nous pensions, ces foci ne colocalisent pas avec le nucléole, suggérant donc une fonction distincte de celle de régulation de l'expression des gènes ribosomiques. En effet, les foci d'accumulation des SATIII-si28 sont retrouvés éloignés du nucléole et apparaissent soit sous forme de paires, soit proches, soit distants. Ce patron de marquage est indépendant du sexe, permettant d'écarter l'hypothèse d'une accumulation sur le chromosome X, et de la lignée utilisée pour ces expériences (sauvage ou *white-mottled*). La variation de la distance entre les deux foci de la même paire permet aussi d'écarter la possibilité d'une accumulation sur un autosome. En effet, les chromosomes homologues restent appariés pendant l'interphase chez la drosophile.

Afin de déterminer si ces SATIII-si28 s'accumulent sur sites chromosomiques précis, il sera nécessaire de réaliser le même type d'expériences sur des chromosomes polytènes. Cependant, la sous-réplication de ces séquences dans les glandes salivaires pourrait ne pas permettre la transcription d'une quantité suffisante de SATIII-si28 pour être détectable.

Cependant, il est possible d'envisager que ces SATIII-si28 ne soient pas originaires du bloc de séquences SAT III présent sur le chromosome X. En effet, les séquences SAT III sont retrouvées dispersées dans le génome de la drosophile et il ne faut pas négliger la possibilité que ces SATIII-si proviennent d'une de ces séquences. Nous devons aussi démontrer que ces SATIII-si sont de vrais siARN, plutôt que des petits transcrits d'origine aléatoire. Leur présence devra donc être recherchée dans des lignées mutantes pour des protéines de la machinerie RNAi.

Une autre possibilité serait que les SATIII-si ne soient pas associés avec des sites spécifiques dans la chromatine mais plutôt avec des sous-structures nucléaires. Si SATIII-si28

est effectivement un siARN, on devrait d'attendre à ce que ses sites d'accumulation contiennent une cible ARN, aient une localisation précise et soient présents en nombre variable. Une des sous-structures nucléaires qui remplit ces conditions est le corpuscule de Cajal (Liu et al., 2006).

Nos efforts afin d'identifier des cibles potentielles pour le SATIII-si28 sont biaisées de par le séquençage incomplet des séquences répétées de la drosophile. Cependant, la dernière version du séquençage du génome de la drosophile permet de mettre en évidence certains alignements entre la séquence du SATIII-si28 et des séquences non encore annotées. De même, certains alignements montrent que la séquence du SATIII-si28 possède des cibles potentielles dans la région SAR/MAR de l'isolateur *gypsy*. Il est connu que la protéine SU(Hw) et son compagnon Mod(Mdg4) sont capables de se lier à ces séquences afin de lui conférer son activité d'isolation (Ghosh et al., 2001). De plus, les protéines SU(Hw) et D1 colocalisent sur un sous-ensemble restreint de séquences *gypsy* (communication personnelle, Pamela Geyer). La correspondance entre les séquences du SATIII-si28 et de l'insulateur *gypsy* se trouve au niveau de la partie riche en paires de bases dA•dT de ce dernier, partie qui pourrait potentiellement être reconnue par la protéine D1. Des expériences de localisation de la protéine D1, des SATIII-si28 et des insulateurs *gypsy* pourraient de révéler intéressantes puisque ces séquences peuvent être retrouvées sous forme de clusters dans les noyaux (Byrd and Corces, 2003). Enfin, nous pourrons rechercher si la protéine D1 joue un rôle dans la transcription de ces SATIII-si au niveau des séquences SAT III grâce à des techniques de siARN dont la cible serait l'ARNm de D1.

F. CONCLUSIONS et PERSPECTIVES

I. Objectifs de cette étude

Dans un premier temps, l'objectif de mon projet de thèse était de valider l'utilisation de ligands synthétiques du petit sillon de l'ADN afin d'étudier la structure biochimique de l'hétérochromatine. Après avoir mis au point certaines techniques spécifiques, j'ai étudié les mécanismes d'action de ces drogues, à la fois *in vitro*, *ex vivo* et *in vivo*.

Dans un deuxième temps, l'élaboration de techniques d'utilisation de ces molécules nous a permis d'accéder à de nouveaux outils pour l'étude de la fonction de l'hétérochromatine. Ainsi, nous avons pu mettre en évidence le rôle central de la protéine D1 dans le recrutement des protéines responsables de la mise en place de la structure hétérochromatinienne. La compréhension du rôle de la protéine D1 au sein de l'hétérochromatine nous amène donc à envisager plusieurs hypothèses sur la fonction de l'hétérochromatine elle-même.

II. Les ligands du petit sillon de l'ADN : un nouvel outil d'étude de l'hétérochromatine

Les ligands du petit sillon de l'ADN, comme les oligopyrroles (Janssen et al., 2000a; Janssen et al., 2000b) ou les diamidines (Susbielle et al., 2005), ont tout d'abord été utilisés pour leur capacité à interagir spécifiquement avec des séquences cibles présentant des propriétés structurales particulières. Les séquences d'ADN riches en paires de bases dA•dT possèdent un petit sillon beaucoup plus profond et plus étroit que celui des séquences mixtes (Crothers et al., 1990).

L'hétérochromatine possède une composition en paires de bases fortement biaisée car les séquences satellites qui la constituent sont en majorité riches en paires de bases dA•dT (Lohe et al., 1993). Les enchaînements réguliers de ces paires de bases altèrent la structure générale de la double hélice. De ce fait, le petit sillon étroit et profond de ces séquences représente une cible préférentielle pour ces petites molécules.

Nous avons montré que l'interaction de molécules telles que les oligopyrroles ou les diamidines entraînent un changement de conformation de ces séquences et influencent ainsi les interactions entre l'ADN et les protéines qui reconnaissent ces structures.

Les différentes familles de composés possédant cette capacité de liaison au petit sillon présentent aussi la particularité d'affecter la localisation et/ou l'activité de la topoïsomérase II (Adachi et al., 1989; Bell et al., 1996; Bell et al., 1997; Fesen and Pommier, 1989; Woynarowski et al., 1989a; Woynarowski et al., 1989b), et plus particulièrement sur les séquences SAT III (Käs and Laemmli, 1992; Käs et al., 1993).

Nos études nous ont permis de conforter la plupart de ces résultats. Actuellement, outre le fait de représenter des outils puissants pour l'étude des interactions entre l'hétérochromatine et les protéines qui la reconnaissent, les ligands du petit sillon de l'ADN riche en paires de bases dA•dT pourraient s'avérer être des molécules relativement prometteuses dans des thérapies cellulaires. En effet, sans interagir avec des séquences codantes, ils sont cependant capables de moduler l'activité transcriptionnelle de certains gènes (Dervan and Edelson, 2003), mais aussi d'agir sur la disponibilité d'enzymes essentielles telle que la topoïsomérase II (Blattes et al., 2006). La compréhension des mécanismes par lesquels ils influence la localisation et l'activité de la topoïsomérase II représente donc un axe de recherche majeur à développer.

III. La protéine D1 : plateforme de recrutement de protéines de l'hétérochromatine

La protéine D1 de la drosophile avait été décrite comme pouvant interagir spécifiquement avec les séquences SAT I et SAT III, riches en paires de bases dA•dT (Levinger and Varshavsky, 1982). Mais son rôle au sein de ces séquences restait à déterminer. Grâce à l'utilisation de ligands du petit sillon de l'ADN synthétisés pour reconnaître spécifiquement les séquences SAT I et SAT III de la drosophile (Janssen et al., 2000a), nous avons pu déterminer que la protéine D1 était capable de recruter des protéines essentielles à la formation ou au maintien de la structure hétérochromatinienne sur ces séquences (Blattes et al., 2006). Ces résultats renforcent ceux des approches plus classiques précédemment mises en œuvre (Aulner et al., 2002; Monod et al., 2002).

Ainsi, la protéine D1 est capable d'influer sur la localisation de la protéine HP1 au niveau des séquences satellites à « hétérochromatiniser ». La protéine HP1 est impliquée dans la mise en place des marques épigénétiques caractérisant l'hétérochromatine (Eissenberg and Elgin, 2000). De même, la protéine D1 intervient dans la localisation de la topoïsomérase II sur ces mêmes séquences. Nous avons pu démontrer que cette interaction était essentielle au maintien de l'activité cette enzyme et à sa localisation. La protéine D1 est l'orthologue de la protéine humaine HMG-A, qui est elle-même localisée préférentiellement au niveau des séquences riches en paires de bases dA•dT (Reeves and Nissen, 1990). Il serait intéressant de poursuivre l'étude du rôle des protéines MATH, dont les protéines D1 et HMG-A font parties, dans des cellules humaines grâce à une stratégie utilisant des molécules chimiques spécifiques des séquences cibles de la protéine HMG-A. Nous pourrions ainsi proposer un modèle plus général pour la fonction des protéines MATH au niveau de ces séquences riches en paires de bases dA•dT, fonction qui reste à déterminer dans le cas de la protéine HMG-A.

Nos études de la localisation des protéines D1 et HP1 sur une lignée modèle de drosophile, la lignée *white-mottled*, nous ont aussi permis de valider l'hypothèse selon laquelle un élément frontière était présent dans la région séparant hétérochromatine et euchromatine. Afin d'étudier l'influence de cet élément frontière, le BE28 (Cuvier et al., 2002), des études visant à l'inactiver biochimiquement pourraient conforter ces résultats. Cet élément frontière correspondant à une séquence répétée, ces études pourraient mettre en œuvre l'utilisation de ligands du petit sillon spécifiques de cette séquence, ou la dégradation des protéines qui le reconnaissent, BEAF32 (Zhao et al., 1995), par des techniques d'interférence à l'ARN. L'inactivation de cet élément frontière pourrait avoir des répercussions sur les gènes qui se trouvent à proximité, tels que le rDNA et le gène *su(f)*.

Nos observations sur les localisations des protéines D1 et HP1 nous ont aussi amené à nous intéresser au rDNA. Les gènes ribosomiques sont toujours retrouvés associés à des séquences satellites de type alphoïdes chez les eucaryotes supérieurs comme chez la souris et l'homme (Hirai et al., 1999). Les séquences SAT III sont les seuls représentants de cette famille chez la drosophile. La conservation de cette juxtaposition et nos résultats sur la localisation de la protéine HP1 nous ont amené à proposer un rôle pour ces séquences satellites.

IV. Les séquences satellites alphoïdes : régulateur de la transcription des gènes ribosomiques ?

Les premiers résultats que nous avons obtenus semblent indiquer que les séquences SAT III pourraient servir de site « stockage » pour la protéine HP1, une fonction qui pourrait jouer un rôle important dans la régulation de la transcription des gènes ribosomiques. Les modifications épigénétiques de type méthylation de la lysine 9 de l'histone H3 sont déjà connues pour intervenir dans la régulation de la transcription de certaines copies de gènes ribosomiques (Santoro and Grummt, 2001; Santoro et al., 2002; Zhou et al., 2002). La présence de la protéine HP1 sur des promoteurs inactifs a été confirmée (Santoro et al., 2002). Nous avons montré qu'une perte de dose de protéine HP1 semblait entraîner une diminution du taux de transcription de ces gènes.

La juxtaposition entre les séquences satellites alphoïdes et le rDNA suggère donc un rôle de stockage des protéines capables d'intervenir dans la régulation de la transcription des gènes ribosomiques. Cette hypothèse relance donc la théorie selon laquelle l'hétérochromatine serait un site de stockage de protéines et leur séquestration au niveau de cette structure permettrait d'en réguler la disponibilité. Les évidences montrant que toute perturbation des interactions entre l'hétérochromatine et les protéines qui la reconnaissent influent sur la régulation génique confortent cette hypothèse (Lu et al., 2000).

Nous avons aussi pu démontrer que la perte de l'élément frontière BE28 dans la lignée *white-mottled* ne semblait pas affecter la régulation des gènes ribosomiques. Cette constatation suggère plusieurs hypothèses quant aux mécanismes possibles de compartimentalisation du nucléole. Il est connu de longue date que la séquence codante des gènes ribosomiques sont capables de former le nucléole, quelle que soit leur localisation (Ritossa and Spiegelman, 1965). Mais la question de leur compartimentalisation reste cependant en suspens, et plus particulièrement suite à nos résultats concernant les localisations des protéines D1 et HP1. Nous avons démontré que ces deux protéines avaient la même localisation sur les séquences SAT III et le rDNA, à la fois dans une souche sauvage et dans une souche *white-mottled*. Par contre, nous avons aussi pu établir une différence de localisation de ces deux protéines entre une souche sauvage et une souche *white-mottled* suite à un traitement au P9.

Ces différents résultats supposent un rôle à la fois des protéines HP1 et D1 dans la régulation de la transcription des gènes ribosomiques. Il est aussi possible d'envisager que la perte de l'élément frontière BE28 ne perturbe pas l'équilibre de localisation de ces protéines établit entre les séquences SAT III et le rDNA. Ces données suggèrent donc que le nucléole est un compartiment qui fonctionne de manière autonome et est capable de s'exclure du contexte environnant. Cependant, la proximité des séquences SAT III lui serait nécessaire afin de mobiliser le plus rapidement possible les protéines dont il a besoin pour sa régulation.

Mais qu'en est il des séquences situées à proximité de la jonction euchromatine/hétérochromatine ? Lors de la perte de l'élément frontière BE28, la régulation de la transcription du gène *white* dans la souche *white-mottled* est affectée par la présence d'une structure de type hétérochromatinienne avoisinante. Cette hypothèse suggère une expansion spatiale de l'hétérochromatine, des séquences SAT III vers le gène *white*, sans affecter le rDNA. Dans ce modèle, le compagnon de l'élément frontière BE28 pourrait se situer à plusieurs endroits : soit entre le rDNA et les séquences SAT III, soit entre le rDNA et le gène *su(f)*. Cette hypothèse semble la plus probable du fait de l'apparition d'un point d'hétérochromatine ectopique à l'extrémité télomérique proximale, site du deuxième point de cassure de l'inversion *white-mottled*, le premier point se situant à proximité directe, ou même directement dans la partie distale du rDNA.

Même si la fonction de l'hétérochromatine reste encore à définir, il est envisageable qu'elle évolue au cours du développement. Cette hypothèse peut être émise suite à l'analyse des petits ARN non codants à différents stades du développement de la drosophile (Aravin et al., 2003). Des petits ARN non codants transcrits à partir des séquences SAT III ont été isolés seulement dans des drosophiles au stade embryonnaire ou larvaire. La structure hétérochromatinienne et ses propriétés fonctionnelles pourraient donc évoluer. Une autre hypothèse peut être soulevée quant aux mécanismes temporels agissants pour la mise en place d'une structure hétérochromatinienne. Ces petits ARN pourraient intervenir dans l'assemblage d'une telle structure à certains moments du développement afin de suppléer au mécanisme impliquant la protéine D1. Cependant, les premiers résultats que nous avons obtenus au laboratoire tendent à indiquer un rôle différent de ces petits ARN. En effet, ni leur localisation, ni leur expression ne semble refléter un rôle dans la mise en place ou le maintien d'une structure de type hétérochromatinienne.

De même, il semble maintenant peu plausible qu'ils soient impliqués dans la régulation de la transcription des gènes ribosomiques, comme nous l'avions envisagé. Leur

localisation précise en foci réguliers reste encore élucider. Elle pourrait correspondre aux « corps de Cajal » (Liu et al., 2006). Les corps de Cajal correspondent à une sous région nucléaire et sont le lieu de maturation des petits ARN du spliceosome. Leur présence chez la drosophile a été confirmé, et leur nombre varie entre 0 et 2 par cellule (Richard et al., 2003). Il serait donc intéressant de réaliser des expériences de co-localisation entre les SAT III-si et un petit scaARN (small cajal body ARN), dU85 (Jady and Kiss, 2001), afin de vérifier si les SAT III-si ne se localisent pas au sein de ces structures. .

L'absence d'évidences directes sur la fonctionnalité de l'hétérochromatine ne signifie aucunement qu'il n'en existe pas. Si elle n'avait aucune utilité, pourquoi serait-elle si hautement conservée ? La même question se pose au sujet de la conservation des protéines qui lui sont associées. D'autres évidences pragmatiques suggèrent l'existence d'une fonction essentielle pour cette structure. En effet, la protéine D1 qui reconnaît spécifiquement ces séquences est essentielle (Aulner et al., 2002). Sa délétion est létale à un stade très précoce ce qui confirme l'hypothèse de l'existence d'une fonction majeure de l'hétérochromatine au cours des premières étapes du développement.

V. Perspectives

Outre les pistes développées suite aux résultats obtenus au cours de ma thèse, les futurs travaux du laboratoire s'orientent vers la caractérisation des effets des protéines de l'hétérochromatine sur la régulation de la transcription de certains gènes. En effet, grâce à une technique d'interférence à l'ARN sur des cellules en culture, nous allons analyser les conséquences de la perte des protéines D1, HP1 et SU(VAR)3-9 sur la transcription des gènes ribosomiques. La déplétion des protéines de l'hétérochromatine nous permettra aussi d'étudier la structure du nucléole dans de telles conditions.

Dans ce système *in vitro*, nous allons pouvoir mimer l'effet de la perte de l'élément frontière par extinction de la protéine BEAF32, nécessaire à l'activité de frontière du BE28. Ainsi, nous rechercherons les effets d'une telle perte sur la transcription des gènes ribosomiques mais aussi du gène *su(f)*, choisi du fait de sa proximité avec l'hétérochromatine. De même, la structure du nucléole sera observée dans de telles conditions. Le ciblage de deux protéines pourrait aussi s'avérer intéressant comme par exemple dans le cas de la déplétion

des protéines D1 et BEAF32, ces deux protéines étant toutes deux impliquées dans la régulation de l'élément frontière BE28.

D'autre part, le résultat concernant les SATIII-si ouvrent une nouvelle voie d'investigation mais aussi de nouvelles hypothèses sur le rôle des séquences SAT III. Quelle est l'origine de ces transcrits ? Où sont-ils localisés dans le noyau ? Dans quel mécanisme sont-ils impliqués ? Autant de questions auxquelles nos futures expériences devront répondre.

De même, notre modèle d'étude, la lignée *white-mottled*, se prête maintenant à l'élucidation du rôle de molécules d'intérêt thérapeutique ciblant les séquences satellites alphoïdes. En effet, nos travaux soulignent la possibilité d'étendre à l'avenir l'utilisation de ce modèle à l'évaluation des propriétés pharmacologiques de ces composés. Les larves de drosophile offrent ainsi un nouveau moyen d'étudier, dans un organisme métazoaire vivant et au cours du développement, l'impact de la prise par voie orale de ces molécules. Le test fonctionnel de la modification de la PEV du gène *white* dans la lignée *white-mottled* autorise le criblage rapide de nombreux composés, tout en permettant de s'affranchir dans un premier temps de modèles animaux plus lourd à manipuler et donc moins adaptés aux premières étapes de sélection de larges familles de molécules à intérêt thérapeutique.

.

G. MATERIEL et METHODES

I. Culture de lignées de drosophiles

1. Conditions de culture

Tous les stocks de mouches ont été maintenus dans un milieu standard contenant de la levure, du sucre, de la polenta et de l'agar. Les températures de culture étaient de 22°C ou de 18°C selon la vitesse de développement souhaitée (rapide à 22°C, plus lente à 18°C). Seuls les stocks de la lignée *white-mottled* ont été maintenus à 18°C afin de préserver leur phénotype. L'induction de l'expression de transgène, comme par exemple pour D1ΔE (voir plus bas), par choc thermique a été réalisée par incubation à 37°C pendant 30 min.

2. Lignée sauvage

La lignée sauvage utilisée dans l'étude est de type Oregon R.

3. Lignées mutantes

a. Lignées *white-mottled*

La lignée présentant le phénotype de variégation d'effet de position du gène *white* utilisée au cours de cette étude est la lignée *white-mottled 4h* (wm^{4h}).

D'autres souches dérivées de cette même lignée ont été utilisées :

- la lignée *white-mottled 4h*/*Su(var)2-5* (wm^{4h}; Su(var)2-5^{05}/In2R+In2L) portant une mutation pour le gène codant la protéine Su(Var)2-5 (HP1)
- la lignée *white-mottled 4h*/*Su(var)3-9* (wm^{4h}; Su(var)3-9^{17}/TM3) portant une mutation pour le gène codant la protéine Su(Var)3-9.

b. Lignées mutantes pour la protéine D1

Plusieurs lignées de drosophile présentant des modifications affectant le gène D1 ont été utilisées :

- la lignée EP^{473} présente l'insertion d'un élément transposable (EP) à la position -271, correspondant au promoteur du gène codant la protéine D1. Cette insertion est létale à l'état homozygote et est maintenue sur différents types de chromosomes balancers.
- la lignée *hsD1 1d* possède une insertion d'un élément transposable sur le chromosome X, portant le cDNA du gène codant D1, en aval du promoteur du gène codant la protéine hsp70.
- la lignée *hsD1ΔE 27a* porte un élément transposable inséré sur le chromosome II. Cet élément transposable contient le cDNA du gène codant pour la protéine D1 délétée de son domaine acide C-terminal placé sous le contrôle du promoteur du gène codant la protéine hsp70.

c. Lignées mutantes pour des protéines de la machinerie d'interférence à l'ARN

La machinerie d'interférence à l'ARN fait intervenir différentes protéines dont le gène codant certaines ont été mutées. Les souches utilisées dans cette étude portent des mutations affectant les gènes codant les protéines Argonaute I (AgoI) et Piwi :

- les lignées *AgoI K08121/CyO*, *AgoI 04845/CyO* et *AgoI K00208/CyO* présentent des mutations dans le gène codant la protéine Argonaute I.
- la lignée *Piwi 06845/CyO* est mutée au niveau du gène codant la protéine Piwi.

II. Culture de cellules de drosophile

La lignée utilisée au cours de cette étude est de type « Schneider S2 ». Ce sont des cellules de type épithéliales embryonnaires dérivées de *Drosophila melanogaster*.

Les cellules S2 sont cultivées dans un milieu adapté : Schneider Drosophila Medium (GIBCO/Invitrogen) supplémenté de 10% sérum de veau fœtal (SIGMA) et additionné d'antibiotiques (10 000 unités/ml Pénicilline et 10 000 µg/ml Streptomycine, GIBCO). Les cellules sont maintenues à une température de 23°C sous une atmosphère dépourvue de dioxyde de carbone. Elles sont passées par simple décollement à la pipette tous les 2 ou 3 jours afin de les maintenir en phase exponentielle de croissance (entre 2.10^6 cellules/ml et 3.10^6 cellules/ml).

III. Techniques de biologie moléculaire

1. Gel retard

a. Préparation de la sonde radiomarquée par le grand fragment de l'ADN polymérase I (fragment de Klenow)

La sonde est issue d'un vecteur portant une répétition de la séquence SAT III, insérée entre les sites *Bam*HI et *Eco*RI. Une première digestion (avec l'une ou l'autre des enzymes) permet de dégager une extrémité à bouts cohésifs.

Des nucléotides sont rajoutés au niveau de cette extrémité grâce à l'activité 5' polymérase de la Klenow. Cette réaction se fait dans un milieu contenant 50 à 100 ng de vecteur digéré repris dans 109 µl d'eau et préalablement dénaturé à 100°C pendant 5 minutes, additionné de 4 µl de Random primers à 500 µg/ml (ROCHE), 6 µl de dNTP froids à 10 mM, 5 µl de α^{32}P-dATP, 14 µl de tampon de réaction 10X pour la Klenow et 2 µl de Klenow à 9 U/µl. Le tout est incubé 1 heure à 37°C puis la réaction est stoppée par addition de 8 µl d'EDTA à 0,25 M pH 8, 4 µl de NaCl à 5 M, 2 µl de SDS à 20%, 2 µl d'ADN de sperme de

saumon (SSDNA) soniqué à 10 mg/ml puis de 144 µl de TE. La sonde est ensuite purifiée à partir de ce milieu de réaction par une extraction au phénol/chloroforme/alcool isoamylique (24 : 24 : 1) volume à volume, précipitée dans l'éthanol puis reprise dans 30 µl de TE. La deuxième réaction de digestion est alors effectuée afin de séparer la sonde SAT III du reste du vecteur.

Afin de purifier le fragment SAT III marqué, la réaction de digestion est mise à migrer sur un gel de polyacrylamide à 6% dans du tampon de migration TBE 1X. La migration est effectuée à 25 mA et ce pendant 5 heures environ. Après la migration, un film autoradiographique est mis en contact avec le gel et après 5 minutes d'exposition, le film est révélé. La bande correspondant à la taille désirée (environ 370 pb) est découpée sur le gel en utilisant le film comme repère.

La sonde est alors extraite du gel polyacrylamide en ajoutant 1 ml de tampon d'élution d'ADN (0,5 M Acétate d'Ammonium, 10 mM Acétate de Magnésium tétrahydrate, 1 mM EDTA pH 8, 0,1% SDS sauf si utilisation de la sonde dans des expériences de retard sur gel). La bande écrasée dans un tube contenant ce milieu est mise sous agitation toute la nuit à température ambiante. Le tube est ensuite vortexé, centrifugé quelques secondes afin de récupérer toutes les gouttes, puis le surnageant est prélevé. Ce surnageant est mis dans un nouveau tube et conservé tandis que le culot de gel de polyacrylamide restant dans le premier tube est rincée deux fois avec 500 µl de tampon d'élution, qui sera ensuite récupéré et ajouté au précédent. Le surnageant est enfin filtré dans une seringue au travers d'un filtre dont les pores sont de 0,75 µm. Le surnageant est alors aliquoté puis compté dans un compteur à scintillation.

b. Préparation des échantillons et migration

Des quantités variables (2, 5 ou 10 ng) de protéine D1 (D1) purifiée sont mises en présence de 1 ng de sonde SAT III préalablement radiomarquée additionnée de 250 ng de SSDNA le tout dans un volume de 20 µl contenant 10 mM de Tris-HCl pH 7,5, 50 mM de NaCl, 1 mM d'EDTA pH 8 et 5% de glycérol. Le tout est incubé 30 minutes à 28°C puis la réaction est stoppée par ajout de 3 µl de Ficoll 20 % repris dans du TBE 0,25X.

Les mêmes expériences sont réalisées avec le polyamide P9 (P9) à des concentrations variables (25, 50, 100, 250 ou 500 nM). Pour les expériences de compétition, des complexes préformés entre la sonde SAT III et D1 (1, 5 ou 10 ng) sont ensuite incubés avec des quantités croissantes de P9 (100, 250 ou 500 nM). Cette expérience consiste tout d'abord à incuber la sonde SAT III avec D1 ou P9 pendant les 15 premières minutes puis de rajouter le compétiteur (respectivement P9 ou D1) pendant 15 minutes supplémentaires. Avant de charger les réactions sur le gel de polyacrylamide à 6%, TBE 0,25X, celui-ci est préalablement mis à prémigrer 1 heure 30 minutes à 150 V. La migration dure toute la nuit à 4°C et à 175 V. Après migration, le gel est séché puis un film autoradiographique est mis en contact pour une exposition.

2. *Footprint à la DNAse I*

a. Préparation de la sonde

La sonde est préparée de la même façon que pour le gel retard. L'élution du fragment marqué est effectuée dans le même tampon en l'absence de SDS.

b. Préparation des échantillons et migration

Des échantillons contenant des quantités croissantes de P9 (50, 100, 250 ou 500 nM) sont préparés en ajoutant 3 µl de sonde SAT III radiomarquée à une extrémité, 2 µg SSDNA et 4 µl de tampon 5X d'incubation (50 mM Tris-HCl, 250 mM NaCl, 5 mM EDTA pH 8, 25% glycérol et 0,5% digitonine, ces deux derniers composés ne sont rajoutés que dans le cas de l'utilisation de P9). Les tubes sont incubés à 30°C pendant 30 minutes.

Au bout de 30 minutes, 5 µl d'une solution contenant de la DNAse I est ajoutée. Cette solution est préparée comme suit : 3,75 µl d'une solution stock de DNAse I à 100 ng/µl (10 mM Tris-HCl pH 8, 10 mM NaCl, 1 mM $CaCl_2$, DNAse I) additionnés de 15 µl de tampon 5X d'incubation (décrit ci-dessus, sans glycérol et sans digitonine), 18,75 µl de $MgCl_2$ à 100 mM et de l'eau en quantité suffisante pour compléter à 75 µl.

30 secondes après addition de la DNAse I, 12 µl de l'échantillon sont prélevés et mis dans un tube maintenu au froid et contenant 25 µl d'une solution permettant de stopper la réaction (0,6 M Acétate de Sodium, 50 mM EDTA pH 8, 10 ng/ml SSDNA). 30 secondes après le premier prélèvement, les 13 µl restants sont à leur tour prélevés et mis dans les tubes contenant la solution de blocage et les 12 µl déjà prélevés. Le volume de chaque tube est ramené à 200 µl en ajoutant 130 µl de TE et 20 µl d'Acétate de Sodium à 3 M. Les échantillons sont purifiés par une extraction phénol/chloroforme/alcool isoamylique (24 : 24 : 1), puis précipités en ajoutant 600 µl d'éthanol à 100% et laissés toute la nuit à -80°C. Après centrifugation, lavage à l'éthanol à 70% et séchage, les échantillons sont repris dans 6 µl d'eau avant d'être mis à migrer.

Le gel de séquence correspond à un gel polyacrylamide de 6% additionné d'urée à une concentration finale de 8 M. Avant de charger les échantillons sur le gel, celui-ci est mis à prémigrer 30 minutes à 2000 V. De même, les échantillons sont chauffés 3 minutes à 100°C, couvercles ouverts. Les dépôts se font en deux partie : la moitié du tube est déposée en premier, et la deuxième moitié 1 heure 30 minutes après dans d'autres puits. Après la migration, le gel est fixé pendant 20 minutes dans un bain contenant 10% de méthanol et 10% d'acide acétique, puis mis à sécher et mis à exposer avec un film autoradiographique.

3. *Préparation de sondes pour hybridation in situ de fluorescence (FISH) contre l'ADN*

Les sondes utilisées pour les FISH contre l'ADN correspondent au cDNA ou au gène de la séquence ciblée. Les plasmides portant ces séquences proviennent de différentes sources :

- le cDNA correspondant au gène *white* a été fourni par Benoît Augé (CBD, Toulouse)
- le cDNA correspondant au gène *Notch* a été fourni par François Schweisguth (ENS, Paris)
- le cDNA et la séquence totale du gène *su(f)* ont été fournis par Kevin O'Hare (Division of Cell and Molecular Biology, Imperial College, Londres, UK)
- l'unité transcriptionnelle codant les ARN ribosomiques (rDNA) provient du laboratoire de Igor B. Dawid (NCI, NIH, Bethesda, USA)

- les sondes dirigées contre SAT III et l'élément frontière BE28 (BE28) proviennent de nos stocks.

Dans un premier temps, les plasmides portant ces séquences d'intérêt ont été amplifiés par transformation bactérienne suivie d'une extraction d'ADN plasmidique. Les plasmides sont ensuite digérés par les enzymes de restriction adaptées puis les produits de digestion sont séparés sur un gel d'agarose. Les bandes migrant aux tailles attendues sont découpées du gel puis l'ADN en est extrait.

Les séquences d'ADN ainsi purifiées sont alors dosées par spectrophotométrie. Les sondes sont préparées grâce à l'utilisation d'un kit de couplage à la Digoxigénine ou à la Biotine (DIG ou BIOTIN-Nick Translation Mix, ROCHE). Le protocole de marquage permet de coupler 1 µg de sonde et d'obtenir des fragments de 200 à 500 pb, selon le temps d'incubation. Après vérification des réactions, le temps d'incubation qui a été choisi est de 1 heure à 15°C. Cette durée permet d'obtenir des sondes dont la taille, après vérification sur gel agarose, est en moyenne de 500 pb. Les sondes sont reprises à une concentration finale de 50 ng/µl.

4. *Préparation de sondes pour hybridation in situ de fluorescence (FISH) contre l'ARN*

Les sondes utilisées dans les expériences de FISH contre l'ARN sont en général de courts oligonucléotides dont la taille varie entre 24 et 28 nucléotides. Elles sont synthétisées et couplées à un fluorochrome, Cy3 ou Fluoresceïne, directement par le fournisseur (SIGMA GENOSYS).

Séquences des sondes utilisées :

- SATIII-si28 : 5' Cy3-ATAGGGATCGTTAGCGCTGGTAATTAGT 3'
- snoARN U3 : 5' GGTGCACTTGGCTTTCAGCTAGTG-Fluoresceïne 3'

5. *Northern blot*

a. Préparation des échantillons

Les ARN totaux sont extraits à partir de 50 mg de mouches adultes (soit 50 femelles et 80 mâles). Les mouches sont homogénéisées dans 1 ml de Trizol (INVITROGEN). Les tubes sont laissés entre 5 minutes et 3 heures à température ambiante puis centrifugés 10 minutes à 13 000 g et à 4°C. Le surnageant est alors transféré dans de nouveaux tubes, laissé 5 minutes à température ambiante et additionné de 200 µl de chloroforme. Les tubes sont vortexés pendant 15 secondes, incubés 3 minutes à température ambiante avant d'être centrifugés 15 minutes, à 4°C et à 13 000 g. La phase supérieure aqueuse est transférée dans un nouveau tube dans lequel sont ajoutés 500 µl d'isopropanol.

Après homogénéisation, les tubes sont incubés 15 minutes à température ambiante avant d'être centrifugés 5 minutes, à 4°C et à 13 000 g. Le culot est ensuite rincé avec 1 ml d'éthanol à 75%. Après 5 minutes d'incubation à température ambiante, les tubes sont centrifugés 5 minutes, à 4°C et à 8 000 g. Le surnageant est alors enlevé et les culots sont ensuite repris dans 100 µl d'eau. Pour faciliter la remise en solution des culots, les tubes sont incubés 10 minutes à 65°C.

Les ARN totaux ainsi extraits sont dosés par spectrophotométrie et vérifiés sur gel agarose. La migration sur gel agarose nécessite l'addition d'un tampon de migration 5X contenant du TBE 5X, formamide 65% et formaldéhyde 9,25%, à une concentration finale de 1X.

b. Préparation du gel et migration

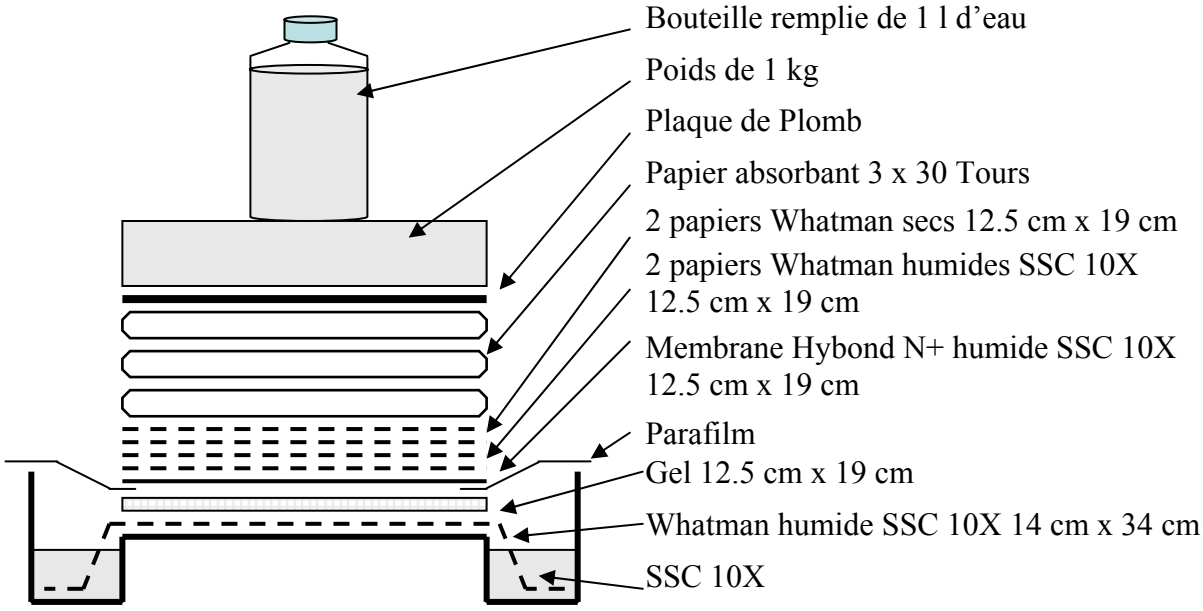
Un gel agarose 1% dénaturant est préparé. L'agarose, 3 g, est mis à fondre dans 216 ml d'eau stérile. Après ébullition, la température de la solution est ramenée à 55°C puis sont ajoutés 30 ml de MOPS 10X (0,2 M MOPS pH 7, 0,1 M Acétate de Sodium, 0,01 M EDTA pH 8) et 54 ml de formaldéhyde déionisée à 37%. Le gel est alors coulé dans une cuve horizontale et mis à sécher sous une hotte.

Les échantillons d'ARN, de 5 à 10 µg dans un volume final maximal de 2,5 µl, sont préparés en ajoutant 7,5 µl de tampon de migration (100 µl MOPS 10X, 165 µl de formaldéhyde à 37%, 500 µl de formamide, 13 µl de BEt à 10 mg/ml et 222 µl d'eau). Les échantillons sont alors chauffés 5 min à 65°C puis refroidis dans la glace. Avant d'être mis à migrer, 2 µl de tampon de charge 6X sont rajoutés aux échantillons.

Le gel est mis à migrer 5 à 6 heures, dans du tampon de migration MOPS 1X et formaldéhyde 6%. Après la migration, le gel subit différents traitements :

- 2 lavages de 1 minute dans 500 ml d'eau stérile
- 1 lavage de 20 minute dans 500 ml de soude à 75 mM (3,75 ml NaOH 10 N dans 500 ml d'eau)
- 2 lavages de 1 minute dans 500 ml d'eau stérile
- 2 lavages de 15 minutes dans du Tris-HCl pH 8 à 0,5 M, NaCl 1,5 M (dans 1,5 L d'eau dissoudre 131,7 g de NaCl et 90,8 g de Tris et ajuster le pH à 8)
- 1 lavage de 1 minute dans 500 ml d'eau
- 1 lavage de 10 minutes dans 500 ml de SSC 10X (175,3 grammes de NaCl, 88,2 g de citrate de sodium dans 800 ml, ajuster le pH avec de l'HCl puis ramener le volume à 1 litre)

Les ARN contenus dans le gel sont alors transférés de manière passive sur un membrane en nylon de type Hybond N+ (Amersham). Le schéma du montage pour ce type de transfert est représenté dans la figure suivante :



Ce montage est laissé toute la nuit et après le transfert, la membrane est récupérée et les ARN y sont pontés par deux cycles automatiques successifs dans le Stratalinker (Stratagène) soit deux fois 1500 J/cm².

c. Préparation des sondes

Les sondes utilisées pour hybrider les ARN ribosomiques cibles sont de courtes séquences et sont synthétisées par SIGMA GENOSYS. Elles ciblent différentes parties des précurseurs des ARN ribosomiques (voir figure 46).

Les séquences des sondes sont les suivantes :

- sonde ETS1 : GAACT GAGAA ACAAA TTTGT ACGGA TAATA ATATC
- sonde 18S : CGCAA GGCGT CACAG TGATC ACGTC CGG
- sonde ITS1 : GTTGC GAAAT GTCTT AGTTT CATAT AAGC
- sonde ITS3 : GTATT ATAAT TCGCT TTGTT TATAT AG

Les sondes sont radiomarquées à leur extrémité 5' avec du γ-P^{32} par action de la T4 polynucleotide kinase (T4 PNK, Promega). La sonde à marquer est tout d'abord diluée à 100 ng/µl puis 100 ng sont mis en présence de 1,5 µl de tampon 10X pour la T4 PNK (Promega), 2,5 µl de γ-P^{32}-ATP, 1 µl de T4 PNK (10 unités/µl), le tout dans un volume final de 15 µl. La réaction est laissée 45 minutes à 37°C. Les sondes sont ensuite séparées des nucléotides non incorporés par passage sur une colonne de chromatographie (Micro Bio-Spin Chromatography Column, BIORAD), puis comptées dans un compteur à scintillation.

d. Hybridation de la sonde sur la membrane

La membrane est mise à préhybrider à 37°C pendant 1 heure dans un four à hybridation, en présence de tampon de préhybridation (rajouté afin respecter une proportion de 0,1 ml de tampon pour 0,1 cm² de membrane). Le tampon de préhybridation contient du SSC 6X, du Denhart's 5X, du SDS 0,5% et du SSDNA à 25 µg/ml.

Après préhybridation de la membrane, la sonde radioactive est rajoutée à raison d'au moins 1 10^6 cpm par membrane et l'hybridation se poursuit toute la nuit à 37°C. Après hybridation, la membrane subit des bains de lavages successifs :

- 2 lavages de 10 minutes dans 250 ml de SSC 2X, SDS 0,1% à 37°C
- 2 lavages de 10 minutes dans 250 ml de SSC 1X, SDS 0,1% à 37°C.

La membrane est alors exposée sur un écran de PhosphorImager durant des temps variables qui dépendent de l'intensité du signal, estimé à l'aide d'un compteur Geiger.

6. *Northern blot contre les petits ARN*

a. Préparation des échantillons

Les ARN totaux sont extraits à partir de larves (50 larves) comme décrit précédemment pour des mouches. Après extraction, les ARN totaux sont dosés par spectrophotométrie et vérifiés sur gel d'agarose.

b. Préparation du gel et migration

Un gel dénaturant à 15 % acrylamide et 7 M urée est préparé. L'urée, 16,82 grammes, est mise à fondre dans 9 ml d'eau stérile additionné de 4 ml de TBE 10X (1X final). Après dissolution, 15 ml d'acrylamide (19 : 1) sont rajoutés. Le mélange est mis à dégazer et après ébullition, 350 µl de TEMED et 17,5 µl d'APS frais sont rajoutés. Le gel est alors rapidement coulé dans une cuve verticale.

Après polymérisation du gel, celui-ci est préchauffé avant la migration pendant 30 minutes à 2 000 V. La migration s'effectuera dans un tampon TBE 1X.

Les échantillons d'ARN, 30 µg dans un volume final maximal de 30 µl, sont préparés en ajoutant un tampon de migration 5X (TBE 5X, formamide 65% et formaldéhyde 9,25%), à une concentration finale de 1X, et un tampon de charge classique 6X (15 % Ficoll type 400, 0,25 % bleu de bromophénol, 0,25 % xylène cyanol) à une concentration finale de 1X. Les échantillons sont alors chauffés 5 min à 65°C puis refroidis dans la glace.

Les échantillons sont alors chargés sur le gel, en ayant auparavant pris le soin d'évacuer l'urée accumulée au fond des puits. Le gel est mis à migrer 1 à 2 heures, dans du tampon de migration 1X TBE.

Après migration, les ARN contenus dans le gel sont alors transférés sur une membrane de type Brightstar Plus (AMBION) par électrotransfert dans un tampon 0,5X TBE, pendant 2 heures à 1 A et à 4°C. Les ARN sont ensuite pontés à la membrane par deux cycles automatiques successifs dans le Stratalinker (Stratagène) soit deux fois 1500 J/cm².

c. Préparation des sondes

Les sondes utilisées pour hybrider les ARN SATIII-si28 cibles sont de courtes séquences synthétisées par SIGMA GENOSYS (voir la séquence dans la partie III. 4. de cette section).

Les sondes sont radiomarquées à leur extrémité 5' par action de la T4 polynucleotide kinase (T4 PNK, Promega) comme décrit précédemment.

d. Hybridation de la sonde sur la membrane

La membrane est mise à préhybrider à 37°C pendant 1 heure dans un four à hybridation, en présence de tampon de préhybridation (rajouté afin respecter une proportion de 0,1 ml de tampon pour 0,1 cm² de membrane). Le tampon de préhybridation contient du SSC 6X, du Denhart's 5X, du SDS 0,5% et du SSDNA à 25 µg/ml.

Après préhybridation de la membrane, la sonde radioactivité est rajoutée à raison d'au moins 1. 10^6 cpm par membrane et l'hybridation se poursuit toute la nuit à 37°C. Après hybridation, la membrane subit des bains de lavages successifs :

- 2 lavages de 10 minutes dans 250 ml de SSC 2X, SDS 0,1% à 37°C
- 2 lavages de 10 minutes dans 250 ml de SSC 1X, SDS 0,1% à 37°C.

La membrane est alors exposée sur un écran de PhosphorImager durant des temps variables qui dépendent de l'intensité du signal, estimé à l'aide d'un compteur Geiger.

IV. Techniques de biologie cellulaire

1. Traitement in vivo de larves par des drogues : « expériences de feeding »

Les larves sont sélectionnées en fonction de leur stade développement. Des larves de 3ème stade de la lignée Oregon R sont choisies afin de réaliser ces expériences d'ingestion de drogues car elles sont dans une phase où elles se nourrissent beaucoup avant de pouvoir rentrer en pupaison.

Des pondoirs (boîtes de Pétri de 5,1 cm de diamètre dont le fond est recouvert par une couche d'agar d'un volume de 6 ml) qui serviront de chambre d'incubation pour les larves sont préalablement recouverts de 60 µl de colchicine à 25 mM (SIGMA). La colchicine permet de bloquer le cycle cellulaire en métaphase et ainsi d'augmenter la probabilité d'observer des chromosomes. Elle est utilisée ici à une concentration finale de 250 µM, dans l'agar, ce qui correspond à une dose sublétale. Afin que les larves puissent continuer à se nourrir, une préparation à base de levure et de colchicine (1 ml de solution saline à 0,7% NaCl, 0,3 grammes de levure déshydratée, 20 µl de colchicine à 10 mM) est rajoutée. Les larves, entre 50 et 100, sont alors rajoutées sur les pondoirs et laissées toute la nuit. Cette technique n'a été utilisée que dans les expériences publiées dans (Blattes et al., 2006), elle a été délaissée car nous avons remarqué que la structure des chromosomes mitotiques était très affectée suite à ce traitement.

Le traitement avec le polyamide P9 est effectué de la même manière en rajoutant du P9 à la fois sur l'agar et dans la nourriture à une concentration finale de 100 µM. Il en est de même pour le traitement à la Trichostatine A (TSA), utilisée à une concentration finale de 10 µM.

Après le traitement, les cerveaux de ces larves seront utilisés dans des expériences d'immunolocalisations décrites dans la partie suivante.

2. *Traitement ex vivo des cerveaux perméabilisés par des drogues*

Les cerveaux de larves de 3ème stade sont disséqués dans du tampon salin (0,7% NaCl), incubés 15 minutes dans du Na-Citrate à 0,5% puis transférés dans une goutte de solution HENMT (10 mM HEPES-NaOH pH 7,4, 1 mM EDTA pH 8, 50 mM NaCl, 5 mM $MgCl_2$, 0,1% Triton X-100). Les cerveaux sont alors laissés 2 fois 1 heure dans cette solution en présence ou non de drogues.

Après incubation, les cerveaux sont rincés 2 fois 5 minutes dans la solution HENMT sans drogue puis sont transférés sur une lame poly-L-Lysine. Les cerveaux sont alors fixés pendant 10 minutes dans une solution HENMT/PFA (paraformaldéhyde) 2%, puis 10 minutes dans une solution contenant PFA 2%/acide acétique 45%.

Une lamelle siliconisée (10 minutes dans la solution de siliconisation puis séchage à l'air libre) préalablement rincée à l'éthanol est placée sur les cerveaux, puis les tissus sont écrasés entre lame et lamelle grâce à une forte pression exercée avec le pouce pendant 10 secondes. Les lames sont alors rapidement plongées dans l'azote liquide, la lamelle est retirée et les lames sont réhydratées dans un tampon de blocage avant que l'expérience d'immunolocalisation ne soit poursuivie.

3. *Immunolocalisations sur des cerveaux de larves*

Les cerveaux sont disséqués à partir de larves de 3ème stade ayant subi ou non un traitement *in vivo*. Les larves choisies sont généralement celles qui ne se nourrissent plus et commencent à sortir de la nourriture pour aller vers un endroit sec afin de réaliser leur pupaison.

La dissection se fait dans du tampon salin 0,7% NaCl. Les cerveaux sont ensuite incubés 15 minutes dans du Na-Citrate 0,5% avant d'être transférés sur une lame Poly-L-Lysine et fixés 10 minutes dans une solution PT (PBS 1X, 0,1% Triton X-100, 10 mM $MgCl_2$) contenant de la PFA à 2%, puis 10 minutes dans une solution contenant PFA 2% /acide acétique 45%.

Les cerveaux sont ensuite écrasés entre lame et lamelle comme décrit précédemment, puis les lames sont réhydratées dans une solution de blocage composée de solution PLT

(solution PT additionnée de 5% de lait en poudre écrémé). Les lames sont incubées au minimum 45 minutes, à température ambiante, ou toute la nuit à 4°C, dans cette solution.

La solution de blocage est enlevée de la lame puis l'anticorps primaire, dilué dans 100 µl de solution PLT à 0,5% de lait, est rajouté. Le temps d'incubation avec cet anticorps primaire est d'environ 2 heures. Les lames sont ensuite lavées, cinq lavages de 5 minutes, avec de la solution PLT à 0,5% de lait, puis l'anticorps secondaire couplé à un fluorochrome et dilué dans 100 µl de solution PLT à 0,5% de lait est rajouté et l'incubation se poursuit pendant 1 heure 30 minutes.

Les différents anticorps primaires et secondaires utilisés, ainsi que leur dilution, sont les suivant :

- anti-D1 (lapin), utilisé au 1/1.000^{ème}, anti-lapin TRITC au 1/2.000^{ème} (SIGMA)
- anti-Topoïsomérase II (lapin) 1/500ème, anti-lapin TRITC au 1/10.000^{ème} (SIGMA)
- anti-HP1 (lapin) 1/10.000^{ème}, anti-lapin TRITC au 1/20.000^{ème} (SIGMA)

Après incubation avec l'anticorps secondaire, les lames sont rincées trois fois 5 minutes dans la solution PLT à 0,5% de lait, puis deux fois 5 minutes dans la solution PT. A ce stade, le milieu de montage contenant du DAPI peut être mis sur les lames avant observation.

Les lames peuvent aussi être incubées avec les polyamides fluorescents afin de contremarquer les séquences satellites. Dans ce cas, le polyamide fluorescent est dilué dans 100 µl de solution PT (P9F 5 nM, P31TR 2 nM, Lex9F 20 nM, P49TR 5 nM) puis incubé 15 minutes sur la lame. Après l'incubation, la lame est rincée cinq fois 5 minutes avec la solution PT avant ajout du milieu de montage additionné de DAPI.

Le milieu de montage contient 100 mM NaCl, 2 mM KCl, 1 mM $MgSO_4$, 2 mM $CaCl_2$, 0,1 mM K-EGTA, 5 mM K-HEPES pH 7,8, 1 mg/ml de P-Phenylenediamine, 78% glycérol. Au dernier moment, 5 µl de DAPI à 10 µg/ml sont rajoutés à 500 µl de milieu de montage, puis 10 µl de ce milieu sont déposés sur lame. Le tout est monté entre lame et lamelle puis observé sur un microscope LEICA DMRB équipé d'une caméra CCD CoolSnap HQ. Les photographies sont prises grâce au logiciel d'acquisition MetaView et reprises dans le logiciel Adobe Photoshop (ajout des couleurs, réglage des contrastes et de la luminosité).

4. *Marquage rapide par les polyamides fluorescents*

Pour des marquages rapides utilisant seulement les polyamides, les cerveaux peuvent être traités plus rapidement que pour des immunolocalisations. Les cerveaux sont disséqués, fixés et écrasés comme décrit précédemment. Les lames sont alors réhydratées non plus avec une solution de blocage mais avec la solution PT. Après 5 min, la solution PT est enlevée et 100 µl de solution PT dans laquelle les polyamides ont été dilués sont rajoutés pendant 15 minutes. Les lames sont ensuite rincées et préparées pour l'observation comme décrit précédemment.

5. *Hybridation in situ de fluorescence (FISH) contre l'ADN*

a. Préparation des tissus

Les expériences de FISH sont adaptées du protocole de M. L. Pardue (Pardue, 1986). Les cerveaux de larves de 3ème stade sont disséqués comme décrit précédemment puis fixés pendant 10 minutes dans 10 µl d'une solution contenant 75% d'éthanol et 25% d'acide acétique. Les cerveaux sont alors transférés sur une lame poly-L-Lysine préalablement chauffée à 40°C et 10 µl d'une solution d'acide acétique à 60% chauffée à 40°C sont rajoutés. Les cerveaux sont étalés sur la lame par des mouvements d'inclinaison. Les lames sont ensuite laissées à l'air libre afin d'être séchées.

b. Dénaturation des tissus

Les lames sont incubées 30 minutes à 70°C dans une solution de SSPE 2X (SSPE 20X pH 7,4 : 3,6 M NaCl, 0,2 M NaH_2PO_4, 0,02 M Na_2-EDTA pH 7,7) puis transférées dans un bain d'alcool à 70% à température ambiante pendant 10 minutes. Les lames sont déshydratées dans des bains d'éthanol à 70% pendant 10 minutes à température ambiante, puis à 95% pendant 5 minutes à température ambiante. Les lames sont ensuite séchées à l'air libre et à température ambiante.

La dénaturation de l'ADN est effectuée par incubation des lames portant les cerveaux dans une solution de soude à 0,07 M pendant 3 minutes. La dénaturation est stoppée par immersion des lames dans une solution d'éthanol à 70% à température ambiante et pendant 10 minutes. Les lames sont déshydratées par deux lavages successifs de 10 minutes et à température ambiante dans un bain d'éthanol à 70%, puis par deux lavages successifs de 5 minutes dans un bain d'éthanol à 95% à température ambiante. Les lames sont ensuite séchées à l'air libre et à température ambiante.

c. Dénaturation des sondes

Les sondes préparées comme décrit précédemment sont mises en présence de 1 µl de SSDNA et le volume est ramené à 25 µl avec de l'eau. La quantité de sonde par lame est d'environ 100 ng. La sonde est alors dénaturée 10 minutes à 100°C puis le tout est refroidi dans la glace. Le milieu d'hybridation (Dextran Sulfate 8%, SSPE 2X, Urée 6M) est alors rajouté de manière à ramener le volume à 50 µl.

Le milieu d'hybridation contenant la sonde est alors réparti sur deux lames (25 µl par lame) puis le tout est placé dans une chambre humide et incubé toute la nuit à 42°C.

Après incubation les lames sont rincées comme décrit ci-dessous :

- 2 lavages de 5 minutes à température ambiante dans une solution de SSPE 2X
- 2 lavages de 5 minutes à 55°C dans une solution de SSPE 2X
- 2 lavages de 5 minutes à 55°C dans une solution de SSPE 1X.

d. Révélation des sondes

Après les lavages, les lames sont incubées 30 minutes, à température ambiante dans une solution de blocage (SSPE 4X, lait écrémé 5% dans solution PT). Les lames sont ensuite incubées pendant 1 heure en présence d'anticorps couplés à un fluorochrome dilués dans une solution contenant du SSPE 0,4X, 0,5% de lait dans la solution PT (anti-BIOTIN-Cy3 dilué à 1/400ème et anti-DIGOXIGENIN-FITC dilué au 1/100ème selon les instructions du fournisseur, ROCHE).

Après incubation, les lames sont lavées deux fois 5 minutes dans la solution SSPE 0,4X, 0,5% de lait dans la solution PT, puis deux fois 5 minutes dans du PBS 1X. Les lames sont ensuite préparées pour l'observation comme décrit précédemment.

6. *Hybridation in situ de fluorescence (FISH) contre l'ARN*

Les cerveaux sont disséqués comme décrit précédemment, transférés sur une lame poly-L-lysine puis fixés 20 minutes dans 10 µl d'une solution contenant 4 % PFA et 1X PBS. Après incubation, les cerveaux sont écrasés entre lame et lamelles, puis trempés dans l'azote liquide et la lamelle est retirée comme décrit précédemment. Les lames sont réhydratées dans la solution PT, puis incubées toute la nuit à 4°C dans de l'éthanol à 70 %.

Avant l'hybridation, les lames sont rincées à température ambiante deux fois 5 minutes dans une solution contenant 2X SSC et formamide 10 %. Les sondes fluorescentes (50 ng par lame) sont ajoutées à une solution contenant 10 % formamide, 2X SSC, 20 µg/µl tRNA, le tout dans un volume final ramené à 14,5 µl. La sonde est alors dénaturée 10 minutes à 100°C avant d'être refroidie sur glace. 10,5 µl d'une solution contenant 10 % de dextran sulfate, 5 µg de BSA et 10 mM de vanadium ribonucleoside sont rajoutés. Les 25 µl du tampon d'hybridation sont alors rajoutés sur la lame et le tout est incubé au minimum 3 heures à 37°C.

Après incubation, les lames sont rincées deux fois 30 minutes à température ambiance dans une solution contenant 2X SSC et 10 % formamide, puis une fois 10 minutes à température ambiante dans une solution de PBS 1X. Les lames sont alors préparées pour l'observation comme décrit précédemment.

H. REFERENCES BIBLIOGRAPHIQUES

1. Aagaard, L., Laible, G., Selenko, P., Schmid, M., Dorn, R., Schotta, G., Kuhfittig, S., Wolf, A., Lebersorger, A., Singh, P.B., Reuter, G. and Jenuwein, T. (1999) Fuctional mammalian homologues of the *Drosophila* PEV-modifier Su(var)3-9 encode centromere-associated proteins which complex with the heterochromatin component M31. *EMBO J*, **18**, 1923-1938.
2. Aaronson, R.P. and Blobel, G. (1975) Isolation of nuclear pore complexes in association with a lamina. *Proc Natl Acad Sci U S A*, **72**, 1007-1011.
3. Aasland, R. and Stewart, A.F. (1995) The chromo shadow domain, a second chromo domain in heterochromatin-binding protein 1, HP1. *Nucleic Acids Res*, **23**, 3168-3173.
4. Abad, J.P., De Pablos, B., Osoegawa, K., De Jong, P.J., Martin-Gallardo, A. and Villasante, A. (2004a) Genomic analysis of Drosophila melanogaster telomeres: full-length copies of HeT-A and TART elements at telomeres. *Mol Biol Evol*, **21**, 1613-1619.
5. Abad, J.P., De Pablos, B., Osoegawa, K., De Jong, P.J., Martin-Gallardo, A. and Villasante, A. (2004b) TAHRE, a novel telomeric retrotransposon from Drosophila melanogaster, reveals the origin of Drosophila telomeres. *Mol Biol Evol*, **21**, 1620-1624.
6. Adachi, Y., Käs, E. and Laemmli, U.K. (1989) Preferential cooperative binding of DNA topoisomerase II to scaffold-associated regions. *EMBO J.*, **8**, 3997-4006.
7. Adachi, Y., Luke, M. and Laemmli, U.K. (1991) Chromosome assembly in vitro: topoisomerase II is required for condensation. *Cell*, **64**, 137-148.
8. Aebi, U., Cohn, J., Buhle, L. and Gerace, L. (1986) The nuclear lamina is a meshwork of intermediate-type filaments. *Nature*, **323**, 560-564.
9. Agostinho, M., Rino, J., Braga, J., Ferreira, F., Steffensen, S. and Ferreira, J. (2004) Human topoisomerase IIα: targeting subchromosomal sites of activity during interphase and mitosis. *Mol. Biol. Cell*, **15**, 2388-2400.
10. Ahmad, K. and Henikoff, S. (2002) The histone variant H3.3 marks active chromatin by replication-independent nucleosome assembly. *Mol. Cell*, **9**, 1191-1200.
11. Ahringer, J. (2000) NuRD and SIN3 histone deacetylase complexes in development. *Trends Genet*, **16**, 351-356.
12. Ainsztein, A.M., Kandels-Lewis, S.E., Mackay, A.M. and Earnshaw, W.C. (1998) INCENP centromere and spindle targeting: identification of essential conserved motifs and involvement of heterochromatin protein HP1. *J Cell Biol*, **143**, 1763-1774.
13. Ajiro, K. (2000) Histone H2B phosphorylation in mammalian apoptotic cells. An association with DNA fragmentation. *J Biol Chem*, **275**, 439-443.
14. Akhtar, A. and Becker, P.B. (2000) Activation of transcription through histone H4 acetylation by MOF, an acetyltransferase essential for dosage compensation in Drosophila. *Mol Cell*, **5**, 367-375.
15. Alfageme, C.R., Zweidler, A., Mahowald, A. and Cohen, L.H. (1974) Histones of Drosophila embryos. Electrophoretic isolation and structural studies. *J Biol Chem*, **249**, 3729-3736.
16. Allfrey, V.G., Faulkner, R. and 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**, 786-794.
17. Allshire, R.C. (2002) RNAi and heterochromatin: a hushed-up affair. *Science*, **297**, 1818-1819.
18. Aravin, A.A., Lagos-Quintana, M., Yalcin, A., Zavolan, M., Marks, D., Snyder, B., Gaasterland, T., Meyer, J. and Tuschl, T. (2003) The small RNA profile during *Drosophila melanogaster* development. *Dev. Cell*, **5**, 337-350.

19. Aravind, L. and Landsman, D. (1998) AT-hook motifs identified in a wide variety of DNA-binding proteins. *Nucleic Acids Res*, **26**, 4413-4421.
20. Arents, G. and Moudrianakis, E.N. (1995) The histone fold: a ubiquitous architectural motif utilized in DNA compaction and protein dimerization. *Proc Natl Acad Sci U S A*, **92**, 11170-11174.
21. Ashley, C.T., Pendleton, C.G., Jennings, W.W., Saxena, A. and Glover, C.V. (1989) Isolation and sequencing of cDNA clones encoding Drosophila chromosomal protein D1. A repeating motif in proteins which recognize at DNA. *J Biol Chem*, **264**, 8394-8401.
22. Aulner, N., Monod, C., Mandicourt, G., Jullien, D., Cuvier, O., Sall, A., Janssen, S., Laemmli, U.K. and Käs, E. (2002) The AT-hook protein D1 is essential for *Drosophila melanogaster* development and is implicated in position-effect variegation. *Mol Cell Biol*, **22**, 1218-1232.
23. Avery, O.T., Mac Leod, C.M. and Mc Carty, M. (1944) Studies on the chimical nature of the substance inducing transformation of pneumococcal types. *J. Exptl. Med.*, **79**, 137-158.
24. Avner, P. and Heard, E. (2001) X-chromosome inactivation: counting, choice and initiation. *Nat Rev Genet*, **2**, 59-67.
25. Ayrault, O., Andrique, L., Larsen, C.J. and Seite, P. (2004) Human Arf tumor suppressor specifically interacts with chromatin containing the promoter of rRNA genes. *Oncogene*, **23**, 8097-8104.
26. Badenhorst, P., Voas, M., Rebay, I. and Wu, C. (2002) Biological functions of the ISWI chromatin remodeling complex NURF. *Genes Dev*, **16**, 3186-3198.
27. Bailly, C., Dassonneville, L., Carrasco, C., Lucas, D., Kumar, A., Boykin, D.W. and Wilson, W.D. (1999) Relationships between topoisomerase II inhibition, sequence-specificity and DNA binding mode of dicationic diphenylfuran derivatives. *Anti-Cancer Drug Design*, **14**, 47-60.
28. Ball, L.J., Murzina, N.V., Broadhurst, R.W., Raine, A.R., Archer, S.J., Stott, F.J., Murzin, A.G., Singh, P.B., Domaille, P.J. and Laue, E.D. (1997) Structure of the chromatin binding (chromo) domain from mouse modifier protein 1. *Embo J*, **16**, 2473-2481.
29. Bannister, A.J. and Kouzarides, T. (1996) The CBP co-activator is a histone acetyltransferase. *Nature*, **384**, 641-643.
30. Bannister, A.J., Schneider, R. and Kouzarides, T. (2002) Histone methylation: dynamic or static? *Cell*, **109**, 801-806.
31. Bannister, A.J., Zegerman, P., Partridge, J.F., Miska, E.A., Thomas, J.O., Allshire, R.C. and Kouzarides, T. (2001) Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. *Nature*, **410**, 120-124.
32. Bantignies, F., Grimaud, C., Lavrov, S., Gabut, M. and Cavalli, G. (2003) Inheritance of Polycomb-dependent chromosomal interactions in Drosophila. *Genes Dev*, **17**, 2406-2420.
33. Baricheva, E.A., Berrios, M., Bogachev, S.S., Borisevich, I.V., Lapik, E.R., Sharakhov, I.V., Stuurman, N. and Fisher, P.A. (1996) DNA from Drosophila melanogaster beta-heterochromatin binds specifically to nuclear lamins in vitro and the nuclear envelope in situ. *Gene*, **171**, 171-176.
34. Bartlett, M.H., Adra, C.N., Park, J., Chapman, V.M. and McBurney, M.W. (1991) DNA methylation of two X chromosome genes in female somatic and embryonal carcinoma cells. *Somat Cell Mol Genet*, **17**, 35-47.
35. Bateson, W., Saunders, E.R. and Punnett, R.C. (1905) Experimental studies in the physiology of heredity. *Reports to Evol. Comm. Royal Soc.*, **2**, 1-131.

36. Baxevanis, A.D. and Landsman, D. (1995) The HMG-1 box protein family: classification and functional relationships. *Nucleic Acids Res*, **23**, 1604-1613.
37. Bell, A., Kittler, L., Lober, G. and Zimmer, C. (1996) Influence of minor groove binders on the eukaryotic topoisomerase II cleavage reaction with 41 base pair model oligonucleotides. *Invest New Drugs*, **13**, 271-284.
38. Bell, A., Kittler, L., Lober, G. and Zimmer, C. (1997) DNA binding properties of minor groove binders and their influence on the topoisomerase II cleavage reaction. *J Mol Recognit*, **10**, 245-255.
39. Benyajati, C.W. and Worcel, A. (1976) Isolation, characterization and structure of the folded interphase genome of Drosophila melanogaster. *Cell*, **9**, 393-407.
40. Berghella, L. and Dimitri, P. (1996) The heterochromatic rolled gene of Drosophila melanogaster is extensively polytenized and transcriptionally active in the salivary gland chromocenter. *Genetics*, **144**, 117-125.
41. Berlingieri, M.T., Manfioletti, G., Santoro, M., Bandiera, A., Visconti, R., Giancotti, V. and Fusco, A. (1995) Inhibition of HMGI-C protein synthesis suppresses retrovirally induced neoplastic transformation of rat thyroid cells. *Mol Cell Biol*, **15**, 1545-1553.
42. Bernardi, G. (1995) The human genome: organization and evolutionary history. *Annu Rev Genet*, **29**, 445-476.
43. Bernardi, G., Olofsson, B., Filipski, J., Zerial, M., Salinas, J., Cuny, G., Meunier-Rotival, M. and Rodier, F. (1985) The mosaic genome of warm-blooded vertebrates. *Science*, **228**, 953-958.
44. Bernstein, E., Caudy, A.A., Hammond, S.M. and Hannon, G.J. (2001) Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature*, **409**, 363-366.
45. Bhat, K.M., Farkas, G., Karch, F., Gyurkovics, H., Gausz, J. and Schedl, P. (1996a) The GAGA factor is required in the early *Drosophila* embryo not only for transcriptionnal regulation but also for nuclear division. *Development*, **122**, 1113-1124.
46. Bhat, M.A., Philp, A.V., Glover, D.M. and Bellen, H.J. (1996b) Chromatid segregation at anaphase requires the barren product, a novel chromosome-associated protein taht interacts with topoisomerase II. *Cell*, **87**, 1103-1114.
47. Bienz, M. and Muller, J. (1995) Transcriptional silencing of homeotic genes in Drosophila. *Bioessays*, **17**, 775-784.
48. Blackburn, M.L., Chansky, H.A., Zielinska-Kwiatkowska, A., Matsui, Y. and Yang, L. (2003) Genomic structure and expression of the mouse ESET gene encoding an ERG-associated histone methyltransferase with a SET domain. *Biochim Biophys Acta*, **1629**, 8-14.
49. Blanton, J., Gaszner, M. and Schedl, P. (2003) Protein:protein interactions and the pairing of boundary elements in vivo. *Genes Dev*, **17**, 664-675.
50. Blastyak, A., Mishra, R.K., Karch, F. and Gyurkovics, H. (2006) Efficient and specific targeting of Polycomb group proteins requires cooperative interaction between Grainyhead and Pleiohomeotic. *Mol Cell Biol*, **26**, 1434-1444.
51. Blattes, R., Monod, C., Susbielle, G., Cuvier, O., Wu, J., Hsieh, T., Laemmli, U.K. and Käs, E. (2006) Displacement of D1, HP1 and topoisomerase II from satellite heterochromatin by a specific polyamide. *Embo J*, **25**, 2397-2408.
52. Blobel, G. (1985) Gene gating: a hypothesis. *Proc Natl Acad Sci U S A*, **82**, 8527-8529.

53. Bode, J., Goetze, S., Heng, H., Krawetz, S.A. and Benham, C. (2003) From DNA structure to gene expression: mediators of nuclear compartmentalization and dynamics. *Chromosome Res*, **11**, 435-445.
54. Bone, J.R., Lavender, J., Richman, R., Palmer, M.J., Turner, B.M. and Kuroda, M.I. (1994) Acetylated histone H4 on the male X chromosome is associated with dosage compensation in Drosophila. *Genes Dev*, **8**, 96-104.
55. Borgnetto, M.E., Zunino, F., Tinelli, S., Käs, E. and Capranico, G. (1996) Drug-specific sites of topoisomerase II DNA cleavage in Drosophila chromatin: heterogeneous localization and reversibility. *Cancer Res*, **56**, 1855-1862.
56. Bottomley, M.J. (2004) Structures of protein domains that create or recognize histone modifications. *EMBO Rep*, **5**, 464-469.
57. Bourachot, B., Yaniv, M. and Muchardt, C. (1999) The activity of mammalian brm/SNF2alpha is dependent on a high-mobility-group protein I/Y-like DNA binding domain. *Mol Cell Biol*, **19**, 3931-3939.
58. Bouteille, B., Oukem, O., Bisser, S. and Dumas, M. (2003) Treatment perspectives for human African trypanosomiasis. *Fundam Clin Pharmacol*, **17**, 171-181.
59. Boveri, T. (1904) Ergebnisse uber die Konstitution der chromatischen Substanz des Zellkerne. *Fischer, Jena*.
60. Bowen, N.J., Palmer, M.B. and Wade, P.A. (2004) Chromosomal regulation by MeCP2: structural and enzymatic considerations. *Cell Mol Life Sci*, **61**, 2163-2167.
61. Boykin, D.W., Kumar, A., Hall, J.E., Bender, B.C. and Tidwell, R.R. (1996) Anti-pneumocystis activity of bis-amidoximes and bis-o-alkylamidoximes prodrugs. *Bioorg. Med. Chem. Lett.*, **6**, 3017-3020.
62. 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. and 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**, 1587-1597.
63. Braunstein, M., Rose, A.B., Holmes, S.G., Allis, C.D. and Broach, J.R. (1993) Transcriptional silencing in yeast is associated with reduced nucleosome acetylation. *Genes Dev*, **7**, 592-604.
64. Bridges, C. (1936) The bar gene duplication. *Science*, **83**, 210-211.
65. Brown, C.J., Hendrich, B.D., Rupert, J.L., Lafreniere, R.G., Xing, Y., Lawrence, J. and Willard, H.F. (1992) The human XIST gene: analysis of a 17 kb inactive X-specific RNA that contains conserved repeats and is highly localized within the nucleus. *Cell*, **71**, 527-542.
66. Brownell, J.E. and Allis, C.D. (1996) Special HATs for special occasions: linking histone acetylation to chromatin assembly and gene activation. *Curr Opin Genet Dev*, **6**, 176-184.
67. Bryk, M., Banerjee, M., Murphy, M., Knudsen, K.E., Garfinkel, D.J. and Curcio, M.J. (1997) Transcriptional silencing of Ty1 elements in the RDN1 locus of yeast. *Genes Dev*, **11**, 255-269.
68. Buck, S.W. and Shore, D. (1995) Action of a RAP1 carboxy-terminal silencing domain reveals an underlying competition between HMR and telomeres in yeast. *Genes Dev*, **9**, 370-384.
69. Burkin, D.J., Broad, T.E. and Jones, C. (1996) The chromosomal distribution and organization of sheep satellite I and II centromeric DNA using characterized sheep-hamster somatic cell hybrids. *Chromosome Res*, **4**, 49-55.
70. Burns, L.G. and Peterson, C.L. (1997) The yeast SWI-SNF complex facilitates binding of a transcriptional activator to nucleosomal sites in vivo. *Mol Cell Biol*, **17**, 4811-4819.

71. Burzio, L.O., Riquelme, P.T. and Koide, S.S. (1979) ADP ribosylation of rat liver nucleosomal core histones. *J Biol Chem*, **254**, 3029-3037.
72. Bustin, M. (1999) Regulation of DNA-dependent activities by the functional motifs of the high-mobility-group chromosomal proteins. *Mol Cell Biol*, **19**, 5237-5246.
73. Bustin, M. and Reeves, R. (1996) High-mobility-group chromosomal proteins: architectural components that facilitate chromatin function. *Prog Nucleic Acid Res Mol Biol*, **54**, 35-100.
74. Byrd, K. and Corces, V.G. (2003) Visualization of chromatin domains created by *gypsy* insulator of *Drosophila*. *J. Cell Biol.*, **162**, 565-574.
75. Cai, H. and Levine, M. (1995) Modulation of enhancer-promoter interactions by insulators in the Drosophila embryo. *Nature*, **376**, 533-536.
76. Cairns, B.R., Erdjument-Bromage, H., Tempst, P., Winston, F. and Kornberg, R.D. (1998) Two actin-related proteins are shared functional components of the chromatin-remodeling complexes RSC and SWI/SNF. *Mol Cell*, **2**, 639-651.
77. Cairns, B.R., Lorch, Y., Li, Y., Zhang, M., Lacomis, L., Erdjument-Bromage, H., Tempst, P., Du, J., Laurent, B. and Kornberg, R.D. (1996) RSC, an essential, abundant chromatin-remodeling complex. *Cell*, **87**, 1249-1260.
78. Camporeale, G., Shubert, E.E., Sarath, G., Cerny, R. and Zempleni, J. (2004) K8 and K12 are biotinylated in human histone H4. *Eur J Biochem*, **271**, 2257-2263.
79. Cao, R., Wang, L., Wang, H., Xia, L., Erdjument-Bromage, H., Tempst, P., Jones, R.S. and Zhang, Y. (2002) Role of histone H3 lysine 27 methylation in Polycomb-group silencing. *Science*, **298**, 1039-1043.
80. Casanova, J., Sanchez-Herrero, E. and Morata, G. (1985) Contrabithorax and the control of spatial expression of the bithorax complex genes of Drosophila. *J Embryol Exp Morphol*, **90**, 179-196.
81. Cavalli, G., Bachmann, D. and Thoma, F. (1996) Inactivation of topoisomerases affects transcription-dependent chromatin transitions in rDNA but not in a gene transcribed by RNA polymerase II. *Embo J*, **15**, 590-597.
82. Chadwick, B.P. and Willard, H.F. (2001a) Histone H2A variants and the inactive X chromosome: identification of a second macroH2A variant. *Hum Mol Genet*, **10**, 1101-1113.
83. Chadwick, B.P. and Willard, H.F. (2001b) A novel chromatin protein, distantly related to histone H2A, is largely excluded from the inactive X chromosome. *J Cell Biol*, **152**, 375-384.
84. Chadwick, B.P. and Willard, H.F. (2003) Chromatin of the Barr body: histone and non-histone proteins associated with or excluded from the inactive X chromosome. *Hum Mol Genet*, **12**, 2167-2178.
85. Chan, C.S., Rastelli, L. and Pirrotta, V. (1994) A Polycomb response element in the Ubx gene that determines an epigenetically inherited state of repression. *Embo J*, **13**, 2553-2564.
86. Chan, R.C., Severson, A.F. and Meyer, B.J. (2004) Condensin restructures chromosomes in preparation for meiotic divisions. *J Cell Biol*, **167**, 613-625.
87. Chargaff, E. (1950) Chemical specificity of nucleic acids and mechanism of their enzymatic degradation. *Experientia*, **6**, 201.
88. Chen, T.L. and Manuelidis, L. (1989) SINEs and LINEs cluster in distinct DNA fragments of Giemsa band size. *Chromosoma*, **98**, 309-316.
89. Cheutin, T., McNairn, A.J., Jenuwein, T., Gilbert, D.M., Singh, P.B. and Misteli, T. (2003) Maintenance of stable heterochromatin domains by dynamic HP1 binding. *Science*, **299**, 721-725.

90. Cléard, F., Delattre, M. and Spierer, P. (1997) SU(VAR)3-7, a *Drosophila* heterochromatin-associated protein and companion of HP1 in the genomic silencing of position-effect variegation. *EMBO J*, **16**, 5280-5288.
91. Cléard, F., Matsarskaia, M. and Spierer, P. (1995) The modifier of position-effect variegation Suvar(3)7 of Drosophila: there are two alternative transcripts and seven scattered zinc fingers, each preceded by a tryptophan box. *Nucleic Acids Res*, **23**, 796-802.
92. Cléard, F. and Spierer, P. (2001) Position-effect variegation in *Drosophila*: the modifier Su(var)3-7 is a modular DNA-binding protein. *EMBO Rep*, **2**, 1095-1100.
93. Cline, T.W. and Meyer, B.J. (1996) Vive la difference: males vs females in flies vs worms. *Annu Rev Genet*, **30**, 637-702.
94. Cockerill, P.N. and Garrard, W.T. (1986) Chromosomal loop anchorage of the kappa immunoglobulin gene occurs next to the enhancer in a region containing topoïsomerase II sites. *Cell*, **44**, 273-282.
95. Coelho, P.A., Queiroz-Machado, J. and Sunkel, C.E. (2003) Condensin-dependent localisation of topoisomerase II to an axial chromosomal structure is required for sister chromatid resolution during mitosis. *J. Cell Science*, **116**, 4763-4776.
96. Conconi, A., Widmer, R.M., Koller, T. and Sogo, J.M. (1989) Two different chromatin structures coexist in ribosomal RNA genes throughout the cell cycle. *Cell*, **57**, 753-761.
97. Copps, K., Richman, R., Lyman, L.M., Chang, K.A., Rampersad-Ammons, J. and Kuroda, M.I. (1998) Complex formation by the Drosophila MSL proteins: role of the MSL2 RING finger in protein complex assembly. *Embo J*, **17**, 5409-5417.
98. Corradini, N., Rossi, F., Verni, F. and Dimitri, P. (2003) FISH analysis of Drosophila melanogaster heterochromatin using BACs and P
elements. *Chromosoma*, **112**, 26-37.
99. Cortes, F., Pastor, N., Mateos, S. and Dominguez, I. (2003) Roles of DNA topoisomerases in chromosome segregation and mitosis. *Mutat. Res.*, **543**, 59-66.
100. Costanzi, C. and Pehrson, J.R. (1998) Histone macroH2A1 is concentrated in the inactive X chromosome of female mammals. *Nature*, **393**, 599-601.
101. Costanzi, C. and Pehrson, J.R. (2001) MACROH2A2, a new member of the MACROH2A core histone family. *J Biol Chem*, **276**, 21776-21784.
102. Craig, J.M. and Bickmore, W.A. (1993) Chromosome bands--flavours to savour. *Bioessays*, **15**, 349-354.
103. Craig, J.M. and Bickmore, W.A. (1994) The distribution of CpG islands in mammalian chromosomes. *Nat Genet*, **7**, 376-382.
104. Crothers, D.M., Haran, T.E. and Nadeau, J.G. (1990) Intrinsically bent DNA. *J Biol Chem*, **265**, 7093-7096.
105. Cryderman, D.E., Grade, S.K., Li, Y., Fanti, L., Pimpinelli, S. and Wallrath, L.L. (2005) Role of Drosophila HP1 in euchromatic gene expression. *Dev Dyn*, **232**, 767-774.
106. Cryderman, D.E., Morris, E.J., Biessmann, H., Elgin, S.C. and Wallrath, L.L. (1999) Silencing at *Drosophila* telomeres: nuclear organization and chromatin structure play critical roles. *EMBO J*, **18**, 3724-3735.
107. Csankovszki, G., McDonel, P. and Meyer, B.J. (2004) Recruitment and spreading of the C. elegans dosage compensation complex along X chromosomes. *Science*, **303**, 1182-1185.
108. Csink, A.K. and Henikoff, S. (1998) Something from nothing: the evolution and utility of satellite repeats. *Trends Genet*, **14**, 200-204.

109. 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. and Kouzarides, T. (2004) Histone deimination antagonizes arginine methylation. *Cell*, **118**, 545-553.
110. Cuvier, O., Hart, C.M., Käs, E. and Laemmli, U.K. (2002) Identification of a multicopy chromatin boundary element at the borders of silenced chromosomal domains. *Chromosoma*, **110**, 519-531.
111. Cuvier, O., Hart, C.M. and Laemmli, U.K. (1998) Identification of a class of chromatin boundary elements. *Mol Cell Biol*, **18**, 7478-7486.
112. Cuvier, O. and Hirano, T. (2003) A role of topoisomerase II in linking DNA replication to chromosome condensation. *J Cell Biol*, **160**, 645-655.
113. Czermin, B., Melfi, R., McCabe, D., Seitz, V., Imhof, A. and Pirrotta, V. (2002) *Drosophila* enhancer of Zeste/ESC complexes have a histone H3 methyltransferase activity that marks chromosomal Polycomb sites. *Cell*, **111**, 185-196.
114. Czermin, B., Schotta, G., Hulsmann, B.B., Brehm, A., Becker, P.B., Reuter, G. and Imhof, A. (2001) Physical and functional association of SU(VAR)3-9 and HDAC1 in *Drosophila*. *EMBO Rep*, **2**, 915-919.
115. Danilevskaya, O., Slot, F., Pavlova, M. and Pardue, M.L. (1994) Structure of the Drosophila HeT-A transposon: a retrotransposon-like element forming telomeres. *Chromosoma*, **103**, 215-224.
116. Daujat, S., Zeissler, U., Waldmann, T., Happel, N. and Schneider, R. (2005) HP1 binds specifically to Lys26-methylated histone H1.4, whereas simultaneous Ser27 phosphorylation blocks HP1 binding. *J Biol Chem*, **280**, 38090-38095.
117. Daury, L., Chailleux, C., Bonvallet, J. and Trouche, D. (2006) Histone H3.3 deposition at E2F-regulated genes is linked to transcription. *EMBO Rep*, **7**, 66-71.
118. Davie, J.R. and Delcuve, G.P. (1991) Characterization and chromatin distribution of the H1 histones and high-mobility-group non-histone chromosomal proteins of trout liver and hepatocellular carcinoma. *Biochem J*, **280 (Pt 2)**, 491-497.
119. de la Cruz, X., Lois, S., Sanchez-Molina, S. and Martinez-Balbas, M.A. (2005) Do protein motifs read the histone code? *Bioessays*, **27**, 164-175.
120. 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. and Brockdorff, N. (2004) Polycomb group proteins Ring1A/B link ubiquitylation of histone H2A to heritable gene silencing and X inactivation. *Dev Cell*, **7**, 663-676.
121. de Ruijter, A.J., van Gennip, A.H., Caron, H.N., Kemp, S. and van Kuilenburg, A.B. (2003) Histone deacetylases (HDACs): characterization of the classical HDAC family. *Biochem J*, **370**, 737-749.
122. De Vries, H. (1900) Sur la loi de disjonction des hybrides. *Comptes Rendus de L'Académie des Sciences de Paris.*, **130**, 845-847.
123. Delattre, M., Spierer, A., Jaquet, Y. and Spierer, P. (2004) Increased expression of Drosophila Su(var)3-7 triggers Su(var)3-9-dependent heterochromatin formation. *J Cell Sci*, **117**, 6239-6247.
124. Delattre, M., Spierer, A., Tonka, C.H. and Spierer, P. (2000) The genomic silencing of position-effect variegation in Drosophila melanogaster: interaction between the heterochromatin-associated proteins Su(var)3-7 and HP1. *J Cell Sci*, **113 Pt 23**, 4253-4261.
125. Dellino, G.I., Schwartz, Y.B., Farkas, G., McCabe, D., Elgin, S.C. and Pirrotta, V. (2004) Polycomb silencing blocks transcription initiation. *Mol Cell*, **13**, 887-893.
126. Denman, R.B. (2005) PAD: the smoking gun behind arginine methylation signaling? *Bioessays*, **27**, 242-246.

127. Dervan, P.B. and Edelson, B.S. (2003) Recognition of the DNA minor groove by pyrrole-imidazole polyamides. *Curr Opin Struct Biol*, **13**, 284-299.
128. Deuring, R., Fanti, L., Armstrong, J.A., Sarte, M., Papoulas, O., Prestel, M., Daubresse, G., Verardo, M., Moseley, S.L., Berloco, M., Tsukiyama, T., Wu, C., Pimpinelli, S. and Tamkun, J.W. (2000) The ISWI chromatin-remodeling protein is required for gene expression and the maintenance of higher order chromatin structure in vivo. *Mol Cell*, **5**, 355-365.
129. Devilee, P., Slagboom, P., Cornelisse, C.J. and Pearson, P.L. (1986) Sequence heterogeneity within the human alphoid repetitive DNA family. *Nucleic Acids Res*, **14**, 2059-2073.
130. Devlin, R.H., Bingham, B. and Wakimoto, B.T. (1990a) The organization and expression of the light gene, a heterochromatic gene of Drosophila melanogaster. *Genetics*, **125**, 129-140.
131. Devlin, R.H., Holm, D.G., Morin, K.R. and Honda, B.M. (1990b) Identifying a single-copy DNA sequence associated with the expression of a heterochromatic gene, the light locus of Drosophila melanogaster. *Genome*, **33**, 405-415.
132. Dhalluin, C., Carlson, J.E., Zeng, L., He, C., Aggarwal, A.K. and Zhou, M.M. (1999) Structure and ligand of a histone acetyltransferase bromodomain. *Nature*, **399**, 491-496.
133. Dhillon, N. and Kamakaka, R.T. (2000) A histone variant, Htz1p, and a Sir1p-like protein, Esc2p, mediate silencing at HMR. *Mol Cell*, **6**, 769-780.
134. Dillon, N. (2004) Heterochromatin structure and function. *Biol Cell*, **96**, 631-637.
135. Dillon, N. and Festenstein, R. (2002) Unravelling heterochromatin: competition between positive and negative factors regulates accessibility. *Trends Genet*, **18**, 252-258.
136. Djikeng, A., Shi, H., Tschudi, C. and Ullu, E. (2001) RNA interference in Trypanosoma brucei: cloning of small interfering RNAs provides evidence for retroposon-derived 24-26-nucleotide RNAs. *Rna*, **7**, 1522-1530.
137. Dodge, J.E., Kang, Y.K., Beppu, H., Lei, H. and Li, E. (2004) Histone H3-K9 methyltransferase ESET is essential for early development. *Mol. Cell. Biol.*, **24**, 2478-2486.
138. Doenecke, D., Albig, W., Bode, C., Drabent, B., Franke, K., Gavenis, K. and Witt, O. (1997) Histones: genetic diversity and tissue-specific gene expression. *Histochem Cell Biol*, **107**, 1-10.
139. Dorer, D.R. and Henikoff, S. (1994) Expansions of transgene repeats cause heterochromatin formation and gene silencing in Drosophila. *Cell*, **77**, 993-1002.
140. Dorer, D.R. and Henikoff, S. (1997) Transgene repeat arrays interact with distant heterochromatin and cause silencing in cis and trans. *Genetics*, **147**, 1181-1190.
141. Dorigo, B., Schalch, T., Kulangara, A., Duda, S., Schroeder, R.R. and Richmond, T.J. (2004) Nucleosome arrays reveal the two-start organization of the chromatin fiber. *Science*, **306**, 1571-1573.
142. Dover, J., Schneider, J., Tawiah-Boateng, M.A., Wood, A., Dean, K., Johnston, M. and Shilatifard, A. (2002) Methylation of histone H3 by COMPASS requires ubiquitination of histone H2B by Rad6. *J Biol Chem*, **277**, 28368-28371.
143. Downs, J.A., Lowndes, N.F. and Jackson, S.P. (2000) A role for Saccharomyces cerevisiae histone H2A in DNA repair. *Nature*, **408**, 1001-1004.
144. Du, J., Nasir, I., Benton, B.K., Kladde, M.P. and Laurent, B.C. (1998) Sth1p, a Saccharomyces cerevisiae Snf2p/Swi2p homolog, is an essential ATPase in RSC and differs from Snf/Swi in its interactions with histones and chromatin-associated proteins. *Genetics*, **150**, 987-1005.

145. Dubinin, N.P. and Sidorov, B.N. (1934) Relationship between the effect of a gene and its position in the system. *Am. Nat.*, **68**, 377-381.
146. Dujon, B. (1996) The yeast genome project: what did we learn? *Trends Genet*, **12**, 263-270.
147. Eberharter, A., Ferrari, S., Langst, G., Straub, T., Imhof, A., Varga-Weisz, P., Wilm, M. and Becker, P.B. (2001) Acf1, the largest subunit of CHRAC, regulates ISWI-induced nucleosome remodelling. *Embo J*, **20**, 3781-3788.
148. Eberl, D.F., Duyf, B.J. and Hilliker, A.J. (1993) The role of heterochromatin in the expression of a heterochromatic gene, the rolled locus of Drosophila melanogaster. *Genetics*, **134**, 277-292.
149. Ebert, A., Schotta, G., Lein, S., Kubicek, S., Krauss, V., Jenuwein, T. and Reuter, G. (2004) Su(var) genes regulate the balance between euchromatin and heterochromatin in Drosophila. *Genes Dev*, **18**, 2973-2983.
150. Eissenberg, J.C. and Elgin, S.C. (2000) The HP1 protein family: getting a grip on chromatin. *Curr Opin Genet Dev*, **10**, 204-210.
151. Eissenberg, J.C. and Elgin, S.C. (2005) Molecular biology: antagonizing the neighbours. *Nature*, **438**, 1090-1091.
152. Eissenberg, J.C. and Hilliker, A.J. (2000) Versatility of conviction: heterochromatin as both a repressor and an activator of transcription. *Genetica*, **109**, 19-24.
153. Eissenberg, J.C., Morris, G.D., Reuter, G. and Hartnett, T. (1992) The heterochromatin-associated protein HP-1 is an essential protein in Drosophila with dosage-dependent effects on position-effect variegation. *Genetics*, **131**, 345-352.
154. Ekwall, K., Olsson, T., Turner, B.M., Cranston, G. and Allshire, R.C. (1997) Transient inhibition of histone deacetylation alters the structural and functional imprint at fission yeast centromeres. *Cell*, **91**, 1021-1032.
155. Ekwall, K. and Ruusala, T. (1994) Mutations in rik1, clr2, clr3 and clr4 genes asymmetrically derepress the silent mating-type loci in fission yeast. *Genetics*, **136**, 53-64.
156. Elgin, S.C. and Grewal, S.I.S. (2003) Heterochromatin: silence is golden. *Curr. Biol.*, **13**, R895-898.
157. Emiliani, S., Fischle, W., Van Lint, C., Al-Abed, Y. and Verdin, E. (1998) Characterization of a human RPD3 ortholog, HDAC3. *Proc Natl Acad Sci U S A*, **95**, 2795-2800.
158. Espinas, M.L., Canudas, S., Fanti, L., Pimpinelli, S., Casanova, J. and Azorin, F. (2000) The GAGA factor of *drosophila* interacts with SAP18, a sin3-associated polypeptide. *EMBO Rep*, **1**, 253-259.
159. Fairlamb, A.H. (2003) Chemotherapy of human African trypanosomiasis: current and future prospects. *Trends Parasitol*, **19**, 488-494.
160. Falvo, J.V., Thanos, D. and Maniatis, T. (1995) Reversal of intrinsic DNA bends in the IFN beta gene enhancer by transcription factors and the architectural protein HMG I(Y). *Cell*, **83**, 1101-1111.
161. Fang, J., Chen, T., Chadwick, B., Li, E. and Zhang, Y. (2004) Ring1b-mediated H2A ubiquitination associates with inactive X chromosomes and is involved in initiation of X inactivation. *J Biol Chem*, **279**, 52812-52815.
162. Fanti, L., Giovinazzo, G., Berloco, M. and Pimpinelli, S. (1998) The heterochromatin protein 1 prevents telomere fusions in Drosophila. *Mol Cell*, **2**, 527-538.
163. Fesen, M. and Pommier, Y. (1989) Mammalian topoisomerase II activity is modulated by the DNA minor groove binder distamycin in simian virus 40 DNA. *J Biol Chem*, **264**, 11354-11359.

164. Fischle, W., Wang, Y. and Allis, C.D. (2003) Histone and chromatin cross-talk. *Curr Opin Cell Biol*, **15**, 172-183.
165. Flanagan, J.F., Mi, L.Z., Chruszcz, M., Cymborowski, M., Clines, K.L., Kim, Y., Minor, W., Rastinejad, F. and Khorasanizadeh, S. (2005) Double chromodomains cooperate to recognize the methylated histone H3 tail. *Nature*, **438**, 1181-1185.
166. Forneris, F., Binda, C., Vanoni, M.A., Battaglioli, E. and Mattevi, A. (2005a) Human histone demethylase LSD1 reads the histone code. *J Biol Chem*, **280**, 41360-41365.
167. Forneris, F., Binda, C., Vanoni, M.A., Mattevi, A. and Battaglioli, E. (2005b) Histone demethylation catalysed by LSD1 is a flavin-dependent oxidative process. *FEBS Lett*, **579**, 2203-2207.
168. Fox, M.G. (1997) Identification of a new ribosomal RNA in Drosophila melanogaster. *Gene*, **194**, 291-295.
169. Franke, A. and Baker, B.S. (1999) The rox1 and rox2 RNAs are essential components of the compensasome, which mediates dosage compensation in Drosophila. *Mol Cell*, **4**, 117-122.
170. Franklin, R.E. and Gosling, R.G. (1953) Evidence for 2-chain helix in crystalline structure of sodium deoxyribonucleate. *Nature*, **172**, 156-157.
171. French, S.L., Osheim, Y.N., Cioci, F., Nomura, M. and Beyer, A.L. (2003) In exponentially growing Saccharomyces cerevisiae cells, rRNA synthesis is determined by the summed RNA polymerase I loading rate rather than by the number of active genes. *Mol Cell Biol*, **23**, 1558-1568.
172. Fujioka, Y., Kimata, Y., Nomaguchi, K., Watanabe, K. and Kohno, K. (2002) Identification of a novel non-structural maintenance of chromosomes (SMC) component of the SMC5-SMC6 complex involved in DNA repair. *J. Biol. Chem.*, **277**, 21585-21591.
173. Fukagawa, T., Nogami, M., Yoshikawa, M., Ikeno, M., Okazaki, T., Takami, Y., Nakayama, T. and Oshimura, M. (2004) Dicer is essential for formation of the heterochromatin structure in vertebrate cells. *Nat Cell Biol*, **6**, 784-791.
174. Fuks, F., Hurd, P.J., Wolf, D., Nan, X., Bird, A.P. and Kouzarides, T. (2003) The methyl-CpG-binding protein MeCP2 links DNA methylation to histone methylation. *J Biol Chem*, **278**, 4035-4040.
175. Gabriel-Robez, O. and Rumpler, Y. (1992) Heterochromatin: role of intercalated heterochromatin in gonosome-autosome translocations. *Eur J Histochem*, **36**, 91-97.
176. Gall, J.G., Cohen, E.H. and Polan, M.L. (1971) Reptitive DNA sequences in drosophila. *Chromosoma*, **33**, 319-344.
177. Galloni, M., Gyurkovics, H., Schedl, P. and Karch, F. (1993) The bluetail transposon: evidence for independent cis-regulatory domains and domain boundaries in the bithorax complex. *Embo J*, **12**, 1087-1097.
178. Garcia-Cao, M., O'Sullivan, R., Peters, A.H., Jenuwein, T. and Blasco, M.A. (2004) Epigenetic regulation of telomere length in mammalian cells by the Suv39h1 and Suv39h2 histone methyltransferases. *Nat Genet*, **36**, 94-99.
179. Gartenberg, M.R. (2000) The Sir proteins of *Saccharomyces cerevisiae*: mediators of trancriptional silencing and much more. *Curr Opin Microbiol*, **3**, 132-137.
180. Gasser, S.M. and Laemmli, U.K. (1987a) A glimpse at chromosomal order. *Trends in Genetics*, **3**, 16-22.
181. Gasser, S.M. and Laemmli, U.K. (1987b) Improved methods for the isolation of individual and clustered mitotic chromosomes. *Exp. Cell Res.*, **173**, 85-98.
182. Gasser, S.M., Laroche, T., Falquet, J., Boy de la Tour, E. and Laemmli, U.K. (1986) Metaphase chromosome structure. Involvment of topoisomerase II. *J. Mol. Biol.*, **188**, 613-629.

183. Gaszner, M., Vazquez, J. and Schedl, P. (1999) The Zw5 protein, a component of the scs chromatin domain boundary, is able to block enhancer-promoter interaction. *Genes Dev*, **13**, 2098-2107.
184. Gdula, D.A. and Corces, V.G. (1997) Characterization of functional domains of the su(Hw) protein that mediate the silencing effect of mod(mdg4) mutations. *Genetics*, **145**, 153-161.
185. Gdula, D.A., Gerasimova, T.I. and G., C.V. (1996) Genetic and molecular analysis of the *gypsy* chromatin insulator of *Drosophila*. *Proc. Natl. Acad. Sci. USA*, **93**, 9378-9383.
186. Geierstanger, B.H., Volkman, B.F., Kremer, W. and Wemmer, D.E. (1994) Short peptide fragments derived from HMG-I/Y proteins bind specifically to the minor groove of DNA. *Biochemistry*, **33**, 5347-5355.
187. Georgakopoulos, T. and Thireos, G. (1992) Two distinct yeast transcriptional activators require the function of the GCN5 protein to promote normal levels of transcription. *Embo J*, **11**, 4145-4152.
188. Gerace, L., Comeau, C. and Benson, M. (1984) Organization and modulation of nuclear lamina structure. *J Cell Sci Suppl*, **1**, 137-160.
189. Gerasimova, T.I., Gdula, D.A., Gerasimov, D.V., Simonova, O. and Corces, V.G. (1995) A Drosophila protein that imparts directionality on a chromatin insulator is an enhancer of position-effect variegation. *Cell*, **82**, 587-597.
190. Germont, J.E., Bellard, M., Oudet, P. and Chambon, P. (1976) Stability of nucleosomes in native and reconstituted chromatins. *Nucleic Acids Res.*, **3**, 3173-3192.
191. Geyer, P.K. and Corces, V.G. (1992) DNA position-specific repression of transcription by a Drosophila zinc finger protein. *Genes Dev*, **6**, 1865-1873.
192. Ghosh, D., Gerasimova, T.I. and Corces, V.G. (2001) Interactions between the Su(Hw) and Mod(mdg4) proteins required for gypsy insulator function. *Embo J*, **20**, 2518-2527.
193. Ghosh, M.K. and Harter, M.L. (2003) A viral mechanism for remodeling chromatin structure in G0 cells. *Mol Cell*, **12**, 255-260.
194. Gilbert, N., Boyle, S., Sutherland, H., De Las Heras, J., Allan, J., Jenuwein, T. and Bickmore, W.A. (2003) Formation of facultative heterochromatin in the absence of HP1. *EMBO J.*, **22**, 5540-5550.
195. Giordano, E., Peluso, I., Senger, S. and Furia, M. (1999) minifly, a Drosophila gene required for ribosome biogenesis. *J Cell Biol*, **144**, 1123-1133.
196. Giot, L., Bader, J.S., Brouwer, C., Chaudhuri, A., Kuang, B., Li, Y., Hao, Y.L., Ooi, C.E., Godwin, B., Vitols, E., Vijayadamodar, 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. and Rothberg, J.M. (2003) A protein interaction map of Drosophila melanogaster. *Science*, **302**, 1727-1736.
197. Girard, F., Bello, B., Laemmli, U.K. and Gehring, W.J. (1998) In vivo analysis of scaffold-associated regions in *Drosophila*: a synthetic high-affinity SAR binding protein suppresses position effect variegation. *EMBO J*, **17**, 2079-2085.
198. Golderer, G. and Grobner, P. (1991) ADP-ribosylation of core histones and their acetylated subspecies. *Biochem J*, **277 (Pt 3)**, 607-610.
199. Goto, H., Tomono, Y., Ajiro, K., Kosako, H., Fujita, M., Sakurai, M., Okawa, K., Iwamatsu, A., Okigaki, T., Takahashi, T. and Inagaki, M. (1999) Identification of a

novel phosphorylation site on histone H3 coupled with mitotic chromosome condensation. *J Biol Chem*, **274**, 25543-25549.
200. Gotta, M. and Cockell, M. (1997) Telomeres, not the end of the story. *Bioessays*, **19**, 367-370.
201. Gotta, M., Laroche, T., Formenton, A., Maillet, L., Scherthan, H. and Gasser, S.M. (1996) The clustering of telomeres and colocalization with Rap1, Sir3, and Sir4 proteins in wild-type Saccharomyces cerevisiae. *J Cell Biol*, **134**, 1349-1363.
202. Gottschling, D.E., Aparicio, O.M., Billington, B.L. and Zakian, V.A. (1990) Position effect at S. cerevisiae telomeres: reversible repression of Pol II transcription. *Cell*, **63**, 751-762.
203. Grant, P.A., Duggan, L., Cote, J., Roberts, S.M., Brownell, J.E., Candau, R., Ohba, R., Owen-Hughes, T., Allis, C.D., Winston, F., Berger, S.L. and Workman, J.L. (1997) Yeast Gcn5 functions in two multisubunit complexes to acetylate nucleosomal histones: characterization of an Ada complex and the SAGA (Spt/Ada) complex. *Genes Dev*, **11**, 1640-1650.
204. Greider, C.W. and Blackburn, E.H. (1985) Identification of a specific telomere terminal transferase activity in Tetrahymena extracts. *Cell*, **43**, 405-413.
205. Greil, F., van der Kraan, I., Delrow, J., Smothers, J.F., de Wit, E., Bussemaker, H.J., van Driel, R., Henikoff, S. and 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**, 2825-2838.
206. Grewal, S.I. and Elgin, S.C. (2002) Heterochromatin: new possibilities for the inheritance of structure. *Curr Opin Genet Dev*, **12**, 178-187.
207. Grewal, S.I.S. and Moazed, D. (2003) Heterochromatin and epigenetic control of gene expression. *Science*, **301**, 798-802.
208. Griffith, F. (1928) The significance of pneumococcal types. *J. Hyg.*, **27**, 113-159.
209. Guo, X.W., Th'ng, J.P., Swank, R.A., Anderson, H.J., Tudan, C., Bradbury, E.M. and Roberge, M. (1995) Chromosome condensation induced by fostriecin does not require p34cdc2 kinase activity and histone H1 hyperphosphorylation, but is associated with enhanced histone H2A and H3 phosphorylation. *Embo J*, **14**, 976-985.
210. Gurley, L.R., D'Anna, J.A., Barham, S.S., Deaven, L.L. and Tobey, R.A. (1978) Histone phosphorylation and chromatin structure during mitosis in Chinese hamster cells. *Eur J Biochem*, **84**, 1-15.
211. Hagstrom, K.A. and Meyer, B.J. (2003) Condensin and cohesin: more than chromosome compactor and glue. *Nature*, **4**.
212. Hall, I.M., Shankaranarayana, G.D., Noma, K.-I., Ayoub, N., Cohen, A. and Grewal, S.I.S. (2002) Establishment and maintenance of a heterochromatin domain. *Science*, **297**, 2232-2237.
213. Hammond, S.M., Bernstein, E., Beach, D. and Hannon, G.J. (2000) An RNA-directed nuclease mediates post-transcriptional gene silencing in Drosophila cells. *Nature*, **404**, 293-296.
214. Hammond, S.M., Boettcher, S., Caudy, A.A., Kobayashi, R. and Hannon, G.J. (2001) Argonaute2, a link between genetic and biochemical analyses of RNAi. *Science*, **293**, 1146-1150.
215. Harrison, D.A., Gdula, D.A., Coyne, R.S. and Corces, V.G. (1993) A leucine zipper domain of the suppressor of Hairy-wing protein mediates its repressive effect on enhancer function. *Genes Dev*, **7**, 1966-1978.
216. Hart, C.M. and Laemmli, U.K. (1998) Facilitation of chromatin dynamics by SARs. *Curr Opin Genet Dev*, **8**, 519-525.

217. Hart, C.M., Zhao, K. and Laemmli, U.K. (1997) The scs' boundary element: characterization of boundary element-associated factors. *Mol Cell Biol*, **17**, 999-1009.
218. Haynes, S.R., Dollard, C., Winston, F., Beck, S., Trowsdale, J. and Dawid, I.B. (1992) The bromodomain: a conserved sequence found in human, Drosophila and yeast proteins. *Nucleic Acids Res*, **20**, 2603.
219. Heard, E., Rougeulle, C., Arnaud, D., Avner, P., Allis, C.D. and Spector, D.L. (2001) Methylation of histone H3 at Lys-9 is an early mark on the X chromosome during X inactivation. *Cell*, **107**, 727-738.
220. Hebbes, T.R., Thorne, A.W. and Crane-Robinson, C. (1988) A direct link between core histone acetylation and transcriptionally active chromatin. *Embo J*, **7**, 1395-1402.
221. Hecht, A., Laroche, T., Strahl-Bolsinger, S., Gasser, S.M. and Grunstein, M. (1995) Histone H3 and H4 N-termini interact with SIR3 and SIR4 proteins: a molecular model for the formation of heterochromatin in yeast. *Cell*, **80**, 583-592.
222. Heinzel, T., Lavinsky, R.M., Mullen, T.M., Soderstrom, M., Laherty, C.D., Torchia, J., Yang, W.M., Brard, G., Ngo, S.D., Davie, J.R., Seto, E., Eisenman, R.N., Rose, D.W., Glass, C.K. and Rosenfeld, M.G. (1997) A complex containing N-CoR, mSin3 and histone deacetylase mediates transcriptional repression. *Nature*, **387**, 43-48.
223. Heitz, E. (1928) Das heterochromatin der moose. *I. Jarhb. Wiss. Botanik.*, **69**, 762-818.
224. Heitz, E. (1929) Heterochromatin, Chromocentren, Chromomeren. (Vorlaufige Mitteilung). *Ber. Dtsch. Bot. Ges.*, **47**, 274-284.
225. Heitz, E. (1932) Geschlechtschromosomen bei einem Laubmoos. (Vorlaufige Mitteilung). *Ber. Dtsch. Bot. Ges.*, **50**, 204-206.
226. Heitz, E. (1933) Uber totale und partielle somatische Heteropyknose, sowie strukturelle Geschlechtschromosomen bei *Drosophila funebris*. (Cytologische Untersuchungen an Dipteren, II). *Z. Zellforsch. mikrosk. Anat.*, **19**, 720-742.
227. Heitz, E. (1934a) Die somatische Heteropyknose bei *Drosophila melanogaster* und ihre genetische Bedeutung. (Cytologische Untersuchungen an Dipteren, III). *Z. Zellforsch. mikrosk. Anat.*, **20**, 237-287.
228. Heitz, E. (1934b) Uber alpha-heterochromatin sowie konstanz und bau der chromomeren bei Drosophila. *Biol. Zentralbl.*, **45**, 588-609.
229. Heitz, E. (1935) Chromosomenstruktur und Gene. *Z. indukt. Abstammungs-Verebungsl.*, **70**, 402-447.
230. Henikoff, S. (1990) Position-effect variegation after 60 years. *Trends Genet*, **6**, 422-426.
231. Henikoff, S. (1992) Position effect and related phenomena. *Curr Opin Genet Dev*, **2**, 907-912.
232. Henikoff, S., Ahmad, K., Platero, J.S. and van Steensel, B. (2000) Heterochromatic deposition of centromeric histone H3-like proteins. *Proc Natl Acad Sci U S A*, **97**, 716-721.
233. Henikoff, S., Jackson, J.M. and Talbert, P.B. (1995) Distance and pairing effects on the brownDominant heterochromatic element in Drosophila. *Genetics*, **140**, 1007-1017.
234. Hilliker, A.J. (1976) Genetic analysis of the centromeric heterochromatin of chromosome 2 of Drosophila melanogaster: deficiency mapping of EMS-induced lethal complementation groups. *Genetics*, **83**, 765-782.
235. Hilliker, A.J. and Appels, R. (1982) Pleiotropic effects associated with the deletion of heterochromatin surrounding rDNA on the X chromosome of *Drosophila*. *Chromosoma*, **86**, 469-490.

236. Hilliker, A.J., Appels, R. and Schalet, A. (1980) The genetic analysis of D. melanogaster heterochromatin. *Cell*, **21**, 607-619.
237. Hirai, H., Taguchi, T. and Godwin, A.K. (1999) Genomic differentiation of 18S ribosomal DNA and beta-satellite DNA in the hominoid and its evolutionary aspects. *Chromosome Res*, **7**, 531-540.
238. Hirschler-Laszkiewicz, I., Cavanaugh, A., Hu, Q., Catania, J., Avantaggiati, M.L. and Rothblum, L.I. (2001) The role of acetylation in rDNA transcription. *Nucleic Acids Res*, **29**, 4114-4124.
239. Holmquist, G.P. (1992) Chromosome bands, their chromatin flavors, and their functional features. *Am J Hum Genet*, **51**, 17-37.
240. Hopwood, D.A. (1988) The Leeuwenhoek lecture, 1987. Towards an understanding of gene switching in Streptomyces, the basis of sporulation and antibiotic production. *Proc R Soc Lond B Biol Sci*, **235**, 121-138.
241. Houchins, K., O'Dell, M., Flavell, R.B. and Gustafson, J.P. (1997) Cytosine methylation and nucleolar dominance in cereal hybrids. *Mol Gen Genet*, **255**, 294-301.
242. Howman, E.V., Fowler, K.J., Newson, A.J., Redward, S., MacDonald, A.C., Kalitsis, P. and Choo, K.H. (2000) Early disruption of centromeric chromatin organization in centromere protein A (Cenpa) null mice. *Proc Natl Acad Sci U S A*, **97**, 1148-1153.
243. Hsieh, T. and Brutlag, D. (1979) Sequence and sequence variation within the 1.688 g/cm3 satellite DNA of Drosophila melanogaster. *J Mol Biol*, **135**, 465-481.
244. Huisinga, K.L., Brower-Toland, B. and Elgin, S.C. (2006) The contradictory definitions of heterochromatin: transcription and silencing. *Chromosoma*, **115**, 110-122.
245. Huth, J.R., Bewley, C.A., Nissen, M.S., Evans, J.N., Reeves, R., Gronenborn, A.M. and Clore, G.M. (1997) The solution structure of an HMG-I(Y)-DNA complex defines a new architectural minor groove binding motif. *Nat Struct Biol*, **4**, 657-665.
246. Hwang, W.W., Venkatasubrahmanyam, S., Ianculescu, A.G., Tong, A., Boone, C. and Madhani, H.D. (2003) A conserved RING finger protein required for histone H2B monoubiquitination and cell size control. *Mol Cell*, **11**, 261-266.
247. Ito, T., Bulger, M., Pazin, M.J., Kobayashi, R. and Kadonaga, J.T. (1997) ACF, an ISWI-containing and ATP-utilizing chromatin assembly and remodeling factor. *Cell*, **90**, 145-155.
248. Ito, T., Levenstein, M.E., Fyodorov, D.V., Kutach, A.K., Kobayashi, R. and Kadonaga, J.T. (1999) ACF consists of two subunits, Acf1 and ISWI, that function cooperatively in the ATP-dependent catalysis of chromatin assembly. *Genes Dev*, **13**, 1529-1539.
249. Ivanova, A.V., Bonaduce, M.J., Ivanov, S.V. and Klar, A.J. (1998) The chromo and SET domains of the Clr4 protein are essential for silencing in fission yeast. *Nat Genet*, **19**, 192-195.
250. Izaurralde, E., Käs, E. and Laemmli, U.K. (1989) Highly preferential nucleation of histone H1 assembly on scaffold-associated regions. *J. Mol. Biol.*, **210**, 573-585.
251. Jacobs, S.A. and Khorasanizadeh, S. (2002) Structure of HP1 chromodomain bound to a lysine 9-methylated histone H3 tail. *Science*, **295**, 2080-2083.
252. Jady, B.E. and Kiss, T. (2001) A small nucleolar guide RNA functions both in 2'-O-ribose methylation and pseudouridylation of the U5 spliceosomal RNA. *Embo J*, **20**, 541-551.
253. Jakubczak, J.L., Burke, W.D. and Eickbush, T.H. (1991) Retrotransposable elements R1 and R2 interrupt the rRNA genes of most insects. *Proc Natl Acad Sci U S A*, **88**, 3295-3299.

254. 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. and Spector, D.L. (2004) From silencing to gene expression: real-time analysis in single cells. *Cell*, **116**, 683-698.
255. Janssen, S., Cuvier, O., Muller, M. and Laemmli, U.K. (2000a) Specific gain- and loss-of-function phenotypes induced by satellite-specific DNA-binding drugs fed to *Drosophila melanogaster*. *Mol Cell*, **6**, 1013-1024.
256. Janssen, S., Durussel, T. and Laemmli, U.K. (2000b) Chromatin opening of DNA satellites by targeted sequence-specific drugs. *Mol Cell*, **6**, 999-1011.
257. Jantzen, H.M., Admon, A., Bell, S.P. and Tjian, R. (1990) Nucleolar transcription factor hUBF contains a DNA-binding motif with homology to HMG proteins. *Nature*, **344**, 830-836.
258. Jaquet, Y., Delattre, M., Spierer, A. and Spierer, P. (2002) Functional dissection of the *Drosophila* modifier of variegation SU(VAR)3-7. *Development*, **129**, 3975-3982.
259. Jason, L.J., Moore, S.C., Lewis, J.D., Lindsey, G. and Ausio, J. (2002) Histone ubiquitination: a tagging tail unfolds? *Bioessays*, **24**, 166-174.
260. Jenuwein, T. (2002) An RNA-guided pathway for the epigenome. *Science*, **297**, 2215-2218.
261. Jenuwein, T., Laible, G., Dorn, R. and Reuter, G. (1998) SET domain proteins modulate chromatin domains in eu- and heterochromatin. *Cell Mol Life Sci*, **54**, 80-93.
262. Jeppesen, P. and Turner, B.M. (1993) The inactive X chromosome in female mammals is distinguished by a lack of histone H4 acetylation, a cytogenetic marker for gene expression. *Cell*, **74**, 281-289.
263. Johnson, K.R., Disney, J.E., Wyatt, C.R. and Reeves, R. (1990) Expression of mRNAs encoding mammalian chromosomal proteins HMG-I and HMG-Y during cellular proliferation. *Exp Cell Res*, **187**, 69-76.
264. Jones, R.S. and Gelbart, W.M. (1993) The Drosophila Polycomb-group gene Enhancer of zeste contains a region with sequence similarity to trithorax. *Mol Cell Biol*, **13**, 6357-6366.
265. Käs, E. and Chasin, L.A. (1987) Anchorage of the Chinese hamster dihydrofolate reductase gene to the nuclear scaffold occurs in an intragenic region. *J Mol Biol*, **198**, 677-692.
266. Käs, E., Izaurralde, E. and Laemmli, U.K. (1989) Specific inhibition of DNA binding to nuclear scaffolds and histone H1 by distamycin. The role of oligo(dA).oligo(dT) tracts. *J. Mol. Biol.*, **210**, 587-599.
267. Käs, E. and Laemmli, U.K. (1992) In vivo topoisomerase II cleavage of the *Drosophila* histone and satellite III repeats: DNA sequence and structural characteristics. *EMBO J*, **11**, 705-716.
268. Käs, E., Poljak, L., Adachi, Y. and Laemmli, U.K. (1993) A model for chromatin opening: stimulation of topoisomerase II and restriction enzyme cleavage of chromatin by distamycin. *EMBO J*, **12**, 115-126.
269. Kasten, M., Szerlong, H., Erdjument-Bromage, H., Tempst, P., Werner, M. and Cairns, B.R. (2004) Tandem bromodomains in the chromatin remodeler RSC recognize acetylated histone H3 Lys14. *Embo J*, **23**, 1348-1359.
270. Kelley, R.I. (1973) Isolation of a histone IIb1-IIb2 complex. *Biochem. Biophys. Res. Commun.*, **54**, 1588-1594.
271. Kellum, R. (2003a) HP1 complexes and heterochromatin assembly. *Curr Top Microbiol Immunol*, **274**, 53-77.
272. Kellum, R. (2003b) Proteins that modify chromatin. *Curr. Topics. Microbiol. Imm.*, **274**, 53-77.

273. Kellum, R. and Schedl, P. (1991) A position-effect assay for boundaries of higher order chromosomal domains. *Cell*, **64**, 941-950.
274. Kellum, R. and Schedl, P. (1992) A group of scs elements function as domain boundaries in an enhancer-blocking assay. *Mol Cell Biol*, **12**, 2424-2431.
275. Keppel, F. (1986) Transcribed human ribosomal RNA genes are attached to the nuclear matrix. *J Mol Biol*, **187**, 15-21.
276. Kermekchiev, M., Workman, J.L. and Pikaard, C.S. (1997) Nucleosome binding by the polymerase I transactivator upstream binding factor displaces linker histone H1. *Mol Cell Biol*, **17**, 5833-5842.
277. Khochbin, S. (2001) Histone H1 diversity: bridging regulatory signals to linker histone function. *Gene*, **271**, 1-12.
278. Khochbin, S. and Wolffe, A.P. (1994) Developmentally regulated expression of linker-histone variants in vertebrates. *Eur J Biochem*, **225**, 501-510.
279. Kim, J., Shen, B., Rosen, C. and Dorsett, D. (1996) The DNA-binding and enhancer-blocking domains of the Drosophila suppressor of Hairy-wing protein. *Mol Cell Biol*, **16**, 3381-3392.
280. Kim, R.A. and Wang, J.C. (1989) A subthreshold level of DNA topoisomerases leads to the excision of yeast rDNA as extrachromosomal rings. *Cell*, **57**, 975-985.
281. Kim, S., Benoiton, L. and Paik, W.K. (1964) Epsilon-Alkyllysinase. Purification And Properties Of The Enzyme. *J Biol Chem*, **239**, 3790-3796.
282. Kimura, K. and Hirano, T. (1997) ATP-dependent positive supercoiling of DNA by 13S condensin: a biochemical implication for chromosome condensation. *Cell*, **90**, 625-634.
283. Kimura, K., Rybenkov, V.V., Crisona, N.J., Hirano, T. and Cozzarelli, N.R. (1999) 13S condensin actively reconfigures DNA by introducing global positive writhe: implications for chromosome condensation. *Cell*, **98**, 239-248.
284. Kireeva, M.L., Walter, W., Tchernajenko, V., Bondarenko, V., Kashlev, M. and Studitsky, V.M. (2002) Nucleosome remodeling induced by RNA polymerase II: loss of the H2A/H2B dimer during transcription. *Mol Cell*, **9**, 541-552.
285. Kireeva, N., Lakonishok, M., Kireev, I., Hirano, T. and Belmont, A.S. (2004) Visualization of early chromosome condensation: a hierarchical folding, axial glue model of chromosome structure. *J Cell Biol*, **166**, 775-785.
286. Koonin, E.V., Zhou, S. and 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**, 4229-4233.
287. Korenberg, J.R. and Rykowski, M.C. (1988) Human genome organization: Alu, lines, and the molecular structure of metaphase chromosome bands. *Cell*, **53**, 391-400.
288. Kornberg, R.D. and Thomas, J.O. (1974) Chromatin structure; oligomers of the histones. *Science*, **184**, 865-868.
289. Kossel, A. (1884) *Z.Physiol.Chem*, **8**.
290. Kourmouli, N., Jeppesen, P., Mahadevhaiah, S., Burgoyne, P., Wu, R., Gilbert, D.M., Bongiorni, S., Prantera, G., Fanti, L., Pimpinelli, S., Shi, W., Fundele, R. and Singh, P.B. (2004) Heterochromatin and tri-methylated lysine 20 of histone H4 in animals. *J Cell Sci*, **117**, 2491-2501.
291. Kreimeyer, A., Wielckens, K., Adamietz, P. and Hilz, H. (1984) DNA repair-associated ADP-ribosylation in vivo. Modification of histone H1 differs from that of the principal acceptor proteins. *J Biol Chem*, **259**, 890-896.
292. Kuhn, A. and Grummt, I. (1992) Dual role of the nucleolar transcription factor UBF: trans-activator and antirepressor. *Proc Natl Acad Sci U S A*, **89**, 7340-7344.

293. Kuhn, A., Stefanovsky, V. and Grummt, I. (1993) The nucleolar transcription activator UBF relieves Ku antigen-mediated repression of mouse ribosomal gene transcription. *Nucleic Acids Res*, **21**, 2057-2063.
294. Kuzmichev, A., Nishioka, K., Erdjument-Bromage, H., Tempst, P. and Reinberg, D. (2002) Histone methyltransferase activity associated with a human multiprotein complex containing the Enhancer of Zeste protein. *Gene Dev*, **16**, 2893-2905.
295. Kwon, H., Imbalzano, A.N., Khavari, P.A., Kingston, R.E. and Green, M.R. (1994) Nucleosome disruption and enhancement of activator binding by a human SW1/SNF complex. *Nature*, **370**, 477-481.
296. Lachner, M., O'Carroll, D., Rea, S., Mechtler, K. and Jenuwein, T. (2001) Methylation of Histone 3 lysine 9 creates a binding site for HP1 proteins. *Nature*, **410**, 116-120.
297. Ladurner, A.G., Inouye, C., Jain, R. and Tjian, R. (2003) Bromodomains mediate an acetyl-histone encoded antisilencing function at heterochromatin boundaries. *Mol Cell*, **11**, 365-376.
298. Laemmli, U.K., Käs, E., Poljak, L. and Adachi, Y. (1992) Scaffold-associated regions: cis-acting determinants of chromatin structural loops and functional domains. *Curr Opin Genet Dev*, **2**, 275-285.
299. Lagger, G., O'Carroll, D., Rembold, M., khier, H., Tischler, J., Weitzer, G., Schuettengruber, B., Hauser, C., Brunmeir, R., Jenuwein, T. and Seiser, C. (2002) Essentiel function of histone deacetylase 1 in proliferation control and CDK inhibitor repression. *EMBO J*, **21**, 2672-2681.
300. Laible, G., Wolf, A., Dorn, R., Reuter, G., Nislow, C., Lebersorger, A., Popkin, D., Pillus, L. and Jenuwein, T. (1997) Mammalian homologues of the polycomb-group gene Enhancer of zeste mediate gene silencing in *Drosophila* heterochromatin and at S. cerevisiae telomeres. *EMBO J*, **16**, 3219-3232.
301. Lam, A.L., Pazin, D.E. and Sullivan, B.A. (2005) Control of gene expression and assembly of chromosomal subdomains by chromatin regulators with antagonistic functions. *Chromosoma*, **114**, 242-251.
302. Langst, G., Becker, P.B. and Grummt, I. (1998) TTF-I determines the chromatin architecture of the active rDNA promoter. *EMBO J.*, **17**, 3135-3145.
303. Lansiaux, A., Dassonneville, L., Facompre, M., Kumar, A., Stephens, C.E., Bajic, M., Tanious, F., Wilson, W.D., Boykin, D.W. and Bailly, C. (2002a) Distribution of furamidine analogues in tumor cells: influence of the number of positive charges. *J. Med. Chem.*, **45**, 1994-2002.
304. Lansiaux, A., Tanious, F., Mishal, Z., Dassonneville, L., Kumar, A., Stephens, C.E., Hu, Q., Wilson, W.D., Boykin, D.W. and Bailly, C. (2002b) Distribution of furamidine analogues in tumor cells: targeting of the nucleus or mitochondria depending on the amidine substitution. *Cancer Res.*, **62**, 7219-7229.
305. Laurent, B.C., Treitel, M.A. and Carlson, M. (1991) Functional interdependence of the yeast SNF2, SNF5, and SNF6 proteins in transcriptional activation. *Proc Natl Acad Sci U S A*, **88**, 2687-2691.
306. Le, H.D., Donaldson, K.M., Cook, K.R. and Karpen, G.H. (2004) A high proportion of genes involved in position effect variegation also affect chromosome inheritance. *Chromosoma*, **112**, 269-276.
307. Leach, T.J., Mazzeo, M., Chotkowski, H.L., Madigan, J.P., Wotring, M.G. and Glaser, R.L. (2000) Histone H2A.Z is widely but nonrandomly distributed in chromosomes of Drosophila melanogaster. *J Biol Chem*, **275**, 23267-23272.
308. Learned, R.M., Cordes, S. and Tjian, R. (1985) Purification and characterization of a transcription factor that confers promoter specificity to human RNA polymerase I. *Mol Cell Biol*, **5**, 1358-1369.

309. Lebkowski, J.S. and Laemmli, U.K. (1982a) Evidence for two levels of DNA folding in histone-depleted HeLa interphase nuclei. *J Mol Biol*, **156**, 309-324.
310. Lebkowski, J.S. and Laemmli, U.K. (1982b) Non-histone proteins and long-range organization of HeLa interphase DNA. *J Mol Biol*, **156**, 325-344.
311. Lee, C.G., Chang, K.A., Kuroda, M.I. and Hurwitz, J. (1997) The NTPase/helicase activities of Drosophila maleless, an essential factor in dosage compensation. *Embo J*, **16**, 2671-2681.
312. Lee, D.Y., Hayes, J.J., Pruss, D. and Wolffe, A.P. (1993) A positive role for histone acetylation in transcription factor access to nucleosomal DNA. *Cell*, **72**, 73-84.
313. Lee, E., Nakatsuma, A., Hiraoka, R., Ishikawa, E., Enomoto, R. and Yamauchi, A. (1999a) Involvement of histone phosphorylation in thymocyte apoptosis by protein phosphatase inhibitors. *IUBMB Life*, **48**, 79-83.
314. Lee, J.H., Cook, J.R., Yang, Z.H., Mirochnitchenko, O., Gunderson, S.I., Felix, A.M., Herth, N., Hoffmann, R. and Pestka, S. (2005) PRMT7, a new protein arginine methyltransferase that synthesizes symmetric dimethylarginine. *J Biol Chem*, **280**, 3656-3664.
315. Lee, J.T., Davidow, L.S. and Warshawsky, D. (1999b) Tsix, a gene antisense to Xist at the X-inactivation centre. *Nat Genet*, **21**, 400-404.
316. Lehnertz, B., Ueda, Y., Derijck, A.A., Braunschweig, U., Perez-Burgos, L., Kubicek, S., Chen, T., Li, E., Jenuwein, T. and Peters, A.H. (2003) Suv39h-mediated histone H3 lysine 9 methylation directs DNA methylation to
major satellite repeats at pericentric heterochromatin. *Curr. Biol.*, **13**, 1192-1200.
317. Lennox, R.W. and Cohen, L.H. (1988) The production of tissue-specific histone complements during development. *Biochem Cell Biol*, **66**, 636-649.
318. Levinger, L. and Varshavsky, A. (1982) Protein D1 preferentially binds A + T-rich DNA in vitro and is a component of Drosophila melanogaster nucleosomes containing A + T-rich satellite DNA. *Proc Natl Acad Sci U S A*, **79**, 7152-7156.
319. Levinger, L.F. (1985) D1 protein of Drosophila melanogaster. Purification and AT-DNA binding properties. *J Biol Chem*, **260**, 14311-14318.
320. Lewis, C.D., Lebkowski, J.S., Daly, A.K. and Laemmli, U.K. (1984) Interphase nuclear matrix and metaphase scaffolding structures. *J Cell Sci Suppl*, **1**, 103-122.
321. Lewis, E.B. (1950) The phenomenon of position effect. *Advances in Genetics*, **3**, 73-115.
322. Lichter, P., Cremer, T., Borden, J., Manuelidis, L. and Ward, D.C. (1988) Delineation of individual human chromosomes in metaphase and interphase cells by in situ suppression hybridization using recombinant DNA libraries. *Human Genetics*, **80**, 224-234.
323. Liu, J.L., Murphy, C., Buszczak, M., Clatterbuck, S., Goodman, R. and Gall, J.G. (2006) The Drosophila melanogaster Cajal body. *J Cell Biol*, **172**, 875-884.
324. Lohe, A.R., Hilliker, A.J. and Roberts, P.A. (1993) Mapping simple repeated DNA sequences in heterochromatin of Drosophila melanogaster. *Genetics*, **134**, 1149-1174.
325. Long, E.O., Collins, M., Kiefer, B.I. and Dawid, I.B. (1981) Expression of the ribosomal DNA insertions in bobbed mutants of Drosophila melanogaster. *Mol Gen Genet*, **182**, 377-384.
326. Long, E.O. and Dawid, I.B. (1980) Alternative pathways in the processing of ribosomal RNA precursor in Drosophila melanogaster. *J Mol Biol*, **138**, 873-878.
327. Loo, S., Laurenson, P., Foss, M., Dillin, A. and Rine, J. (1995) Roles of ABF1, NPL3, and YCL54 in silencing in Saccharomyces cerevisiae. *Genetics*, **141**, 889-902.
328. Lorain, S., Lecluse, Y., Scamps, C., Mattei, M.G. and Lipinski, M. (2001) Identification of human and mouse HIRA-interacting protein-5 (HIRIP5), two

mammalian representatives in a family of phylogenetically conserved proteins with a role in the biogenesis of Fe/S proteins. *Biochim Biophys Acta*, **1517**, 376-383.
329. Losada, A. and Hirano, T. (2001) Shaping the metaphase chromosome: Coordination of cohesion and condensation. *Bioessays*, **23**, 924-935.
330. Losada, A. and Villasante, A. (1996) Autosomal location of a new subtype of 1.688 satellite DNA of Drosophila melanogaster. *Chromosome Res*, **4**, 372-383.
331. Love, J.J., Li, X., Case, D.A., Giese, K., Grosschedl, R. and Wright, P.E. (1995) Structural basis for DNA bending by the architectural transcription factor LEF-1. *Nature*, **376**, 791-795.
332. Loyola, A. and Almouzni, G. (2004) Bromodomains in living cells participate in deciphering the histone code. *Trends Cell Biol*, **14**, 279-281.
333. Lu, B.Y., Emtage, P.C., Duyf, B.J., Hilliker, A.J. and Eissenberg, J.C. (2000) Heterochromatin protein 1 is required for the normal expression of two heterochromatin genes in Drosophila. *Genetics*, **155**, 699-708.
334. Lu, Q., Wallrath, L.L. and Elgin, S.C. (1995) The role of a positioned nucleosome at the Drosophila melanogaster hsp26 promoter. *Embo J*, **14**, 4738-4746.
335. Lucchini, R. and Sogo, J.M. (1994) Chromatin structure and transcriptional activity around the replication forks arrested at the 3' end of the yeast rRNA genes. *Mol Cell Biol*, **14**, 318-326.
336. Lucchini, R. and Sogo, J.M. (1995) Replication of transcriptionally active chromatin. *Nature*, **374**, 276-280.
337. Luger, K., Mader, A.W., Richmond, R.K., Sargent, D.F. and Richmond, T.J. (1997) Crystal structure of the nucleosome core particle at 2.8 A resolution. *Nature*, **389**, 251-260.
338. Lupo, R., Breiling, A., Bianchi, M.E. and Orlando, V. (2001) *Drosophila* chromosome condensation proteins topoisomerase II and barren colocalize with polycomb and maintain Fab-7 PRE silencing. *Mol Cell*, **7**, 127-136.
339. Lustig, A.J. (1998) Mechanisms of silencing in Saccharomyces cerevisiae. *Curr Opin Genet Dev*, **8**, 233-239.
340. Lyon, M.F. (1961) Gene action in the X-chromosome of the mouse (Mus musculus L.). *Nature*, **190**, 372-373.
341. Ma, H., Baumann, C.T., Li, H., Strahl, B.D., Rice, R., Jelinek, M.A., Aswad, D.W., Allis, C.D., Hager, G.L. and Stallcup, M.R. (2001) Hormone-dependent, CARM1-directed, arginine-specific methylation of histone H3 on a steroid-regulated promoter. *Curr Biol*, **11**, 1981-1985.
342. Machado, C. and Andrew, D.J. (2000) D-TITIN: a giant protein with dual roles in chromosomes and muscles. *J. Cell. Biol.*, **151**, 639-651.
343. Madigan, J.P., Chotkowski, H.L. and Glaser, R.L. (2002) DNA double-strand break-induced phosphorylation of Drosophila histone variant H2Av helps prevent radiation-induced apoptosis. *Nucleic Acids Res*, **30**, 3698-3705.
344. Maeshima, K. and Laemmli, U.K. (2003) A two-step scaffolding model for mitotic chromosome assembly. *Dev Cell*, **4**, 467-480.
345. Maison, C. and Almouzni, G. (2004) HP1 and the dynamics of heterochromatin maintenance. *Nature*, **5**, 296-304.
346. Mandl, B., Brandt, W.F., Superti-Furga, G., Graninger, P.G., Birnstiel, M.L. and Busslinger, M. (1997) The five cleavage-stage (CS) histones of the sea urchin are encoded by a maternally expressed family of replacement histone genes: functional equivalence of the CS H1 and frog H1M (B4) proteins. *Mol Cell Biol*, **17**, 1189-1200.
347. Manuelidis, L. (1985) Individual interphase chromosome domains revealed by *in situ hybridization*. *Human Genetics*, **71**, 288-293.

348. Margueron, R., Trojer, P. and Reinberg, D. (2005) The key to development: interpreting the histone code? *Curr Opin Genet Dev*, **15**, 163-176.
349. Marsden, M.P. and Laemmli, U.K. (1979) Metaphase chromosome structure: evidence for a radial loop model. *Cell*, **17**, 849-858.
350. Marshall, W.F., Straight, A., Marko, J.F., Swedlow, J., Dernburg, A., Belmont, A., Murray, A.W., Agard, D.A. and Sedat, J.W. (1997) Interphase chromosomes undergo constrained diffusional motion in living cells. *Curr. Biol.*, **7**, 930-939.
351. Masumoto, H., Nakano, M. and Ohzeki, J. (2004) The role of CENP-B and alpha-satellite DNA: de novo assembly and epigenetic maintenance of human centromeres. *Chromosome Res*, **12**, 543-556.
352. May, H.G. (1917) The appearence of reverse mutations in the Bar-eyed race of Drosophila under experimental control. *Proc. Natl. Acad. Sci. USA*, **3**, 544-545.
353. Mayer, C., Schmitz, K.M., Li, J., Grummt, I. and Santoro, R. (2006) Intergenic transcripts regulate the epigenetic state of rRNA genes. *Mol Cell*, **22**, 351-361.
354. Meehan, R.R., Lewis, J.D., McKay, S., Kleiner, E.L. and Bird, A.P. (1989) Identification of a mammalian protein that binds specifically to DNA containing methylated CpGs. *Cell*, **58**, 499-507.
355. Melcher, M., Schmid, M., Aagaard, L., Selenko, P., Laible, G. and Jenuwein, T. (2000) Structure-function analysis of SUV39H1 reveals a dominant role in heterochromatin organization, chromosome segregation, and mitotic progression. *Mol Cell Biol*, **20**, 3728-3741.
356. Meluh, P.B., Yang, P., Glowczewski, L., Koshland, D. and Smith, M.M. (1998) Cse4p is a component of the core centromere of Saccharomyces cerevisiae. *Cell*, **94**, 607-613.
357. Mendel, G. (1865) Versuche über Pflanzen-Hybriden. *Vorgelegt in den Sitzungen vom 8. Februar und 8. März 1865.*
358. Mendel, G. (1866) Versuche über Pflanzen-Hybriden. *Verh. Naturforsch. Ver. Brünn*, **4**, 3–47.
359. Meneghini, M.D., Wu, M. and Madhani, H.D. (2003) Conserved variant H2A.Z protects euchromatin from the ectopic spread of silent heterochromatin. *Cell*, **112**, 725-736.
360. Metivier, R., Penot, G., Hubner, M.R., Reid, G., Brand, H., Kos, M. and Gannon, F. (2003) Estrogen receptor-alpha directs ordered, cyclical, and combinatorial recruitment of cofactors on a natural target promoter. *Cell*, **115**, 751-763.
361. Miescher, F. (1871) Über die chemische Zusammensetzung der Eiterzellen. *Med-Chem Untersuch*, **4**, 441-460.
362. Miklos, G.L. and Cotsell, J.N. (1990) Chromosome structure at interfaces between major chromatin types: alpha- and beta-heterochromatin. *Bioessays*, **12**, 1-6.
363. Minc, E., Allory, Y., Courvalin, J.C. and Buendia, B. (2001) Immunolocalization of HP1 proteins in metaphasic mammalian chromosomes. *Methods Cell Sci*, **23**, 171-174.
364. Minc, E., Allory, Y., Worman, H.J., Courvalin, J.C. and Buendia, B. (1999) Localization and phosphorylation of HP1 proteins during the cell cycle in mammalian cells. *Chromosoma*, **108**, 220-234.
365. Minc, E., Courvalin, J.C. and Buendia, B. (2000) HP1gamma associates with euchromatin and heterochromatin in mammalian nuclei and chromosomes. *Cytogenet Cell Genet*, **90**, 279-284.
366. Mirkovitch, J., Mirault, M.E. and Laemmli, U.K. (1984) Organization of the higher-order chromatin loop: specific DNA attachment sites on nuclear scaffold. *Cell*, **39**, 223-232.

367. Mizuguchi, G., Vassilev, A., Tsukiyama, T., Nakatani, Y. and Wu, C. (2001) ATP-dependent nucleosome remodeling and histone hyperacetylation synergistically facilitate transcription of chromatin. *J Biol Chem*, **276**, 14773-14783.
368. Mondal, N. and Parvin, J.D. (2001) DNA topoisomerase IIα is required for RNA polymerase II transcription on chromatin templates. *Nature*, **413**, 435-438.
369. Monod, C., Aulner, N., Cuvier, O. and Käs, E. (2002) Modification of Position Effect Variegation by competition for binding to *Drosophila* satellites. *EMBO Rep*, **3**, 747-752.
370. Montgomery, M.K. (2004) RNA interference: historical overview and significance. *Methods Mol Biol*, **265**, 3-21.
371. Moore, P.B. and Steitz, T.A. (2002) The involvement of RNA in ribosome function. *Nature*, **418**, 229-235.
372. Morgan, T.H. (1910) Sex limited inheritance in Drosophila. *Science*, **32**, 120.
373. Morgan, T.H. (1911a) The origin of five mutations in eye color in Drosophila and their mode of inheritance. *Science*, **33**, 534.
374. Morgan, T.H. (1911b) The origin of nine wings mutations in Drosophila. *Science*, **33**, 496.
375. Morgan, T.H. (1919) *The physical basis of heredity*, Philadelphia, London.
376. Morgan, T.H., Sturtevant, A.H., Muller, H.J. and Bridges, C. (1915) The mechanism of mendelian heredity. *New-York*.
377. Moss, T. (2004) At the crossroads of growth control; making ribosomal RNA. *Curr Opin Genet Dev*, **14**, 210-217.
378. Moss, T. and Stefanovsky, V.Y. (2002) At the center of eukaryotic life. *Cell*, **109**, 545-548.
379. Mottus, R., Sobel, R.E. and Grigliatti, T.A. (2000) Mutational analysis of a histone deacetylase in Drosophila melanogaster: missense mutations suppress gene silencing associated with position effect variegation. *Genetics*, **154**, 657-668.
380. Moyzis, R.K., Buckingham, J.M., Cram, L.S., Dani, M., Deaven, L.L., Jones, M.D., Meyne, J., Ratliff, R.L. and Wu, J.R. (1988) A highly conserved repetitive DNA sequence, (TTAGGG)n, present at the telomeres of human chromosomes. *Proc Natl Acad Sci U S A*, **85**, 6622-6626.
381. Muchardt, C., Guilleme, M., Seeler, J.S., Trouche, D., Dejean, A. and Yaniv, M. (2002) Coordinated methyl and RNA binding is required for heterochromatin localization
of mammalian HP1alpha. *EMBO Rep.*
382. Muller, H.J. (1927) Artificial transmutation of the gene. *Science*, **66**, 84-87.
383. Muller, H.J. (1930) Types of viable variations induced by X-rays in Drosophila. *J. Genet.*, **22**, 299-333.
384. Muller, H.J. (1948) Evidence on the precision of genetic adaptation. *Harvey Lect.*, **43**, 165-229.
385. Muller, H.J. and Kaplan, W.D. (1966) The dosage compensation of Drosophila and mammals as showing the accuracy of the normal type. *Genet Res*, **8**, 41-59.
386. Muller, J. (1995) Transcriptional silencing by the Polycomb protein in Drosophila embryos. *Embo J*, **14**, 1209-1220.
387. Nabirochkin, S., Ossokina, M. and Heidmann, T. (1998) A nuclear matrix/scaffold attachment region co-localizes with the Gypsy retrotransposon insulator sequence. *J. Biol. Chem.*, **273**, 2473-2479.
388. Nagel, S. and Grossbach, U. (2000) Histone H1 genes and histone gene clusters in the genus Drosophila. *J Mol Evol*, **51**, 286-298.

389. Nakagawa, H., Lee, J.K., Hurwitz, J., Allshire, R.C., Nakayama, J., Grewal, S.I.S., Tanaka, K. and Y., M. (2002) Fission yeast CENP-B homologs nucleate centromeric heterochromatin by promoting heterechromatin-specific histone tail modifications. *Gene Dev*, **16**, 1766-1778.
390. Nan, X., Meehan, R.R. and Bird, A. (1993) Dissection of the methyl-CpG binding domain from the chromosomal protein MeCP2. *Nucleic Acids Res*, **21**, 4886-4892.
391. Narang, M.A., Dumas, R., Ayer, L.M. and Gravel, R.A. (2004) Reduced histone biotinylation in multiple carboxylase deficiency patients: a nuclear role for holocarboxylase synthetase. *Hum Mol Genet*, **13**, 15-23.
392. Nasmyth, K. (2002) Segregating sister genomes: the molecular biology of chromosome separation. *Science*, **297**, 559-565.
393. Nelson, D.M., Ye, X., Hall, C., Santos, H., Ma, T., Kao, G.D., Yen, T.J., Harper, J.W. and Adams, P.D. (2002) Coupling of DNA synthesis and histone synthesis in S phase independent of cyclin/cdk2 activity. *Mol Cell Biol*, **22**, 7459-7472.
394. Ng, H.H., Feng, Q., Wang, H., Erdjument-Bromage, H., Tempst, P., Zhang, Y. and Struhl, K. (2002) Lysine methylation within the globular domain of histone H3 by Dot1 is important for telomeric silencing and Sir protein association. *Gene Dev*, **16**, 1518-1527.
395. Nicolas, E., Roumillac, C. and Trouche, D. (2003) Balance between acetylation and methylation of histone H3 lysine 9 on the E2F-responsive dihydrofolate reductase promoter. *Mol Cell Biol*, **23**, 1614-1622.
396. Nielsen, A.L., Oulad-Abdelghani, M., Ortiz, J.A., Remboutsika, E., Chambon, P. and Losson, R. (2001a) Heterochromatin formation in mammalian cells: interaction between histones and HP1 proteins. *Mol Cell*, **7**, 729-739.
397. Nielsen, S.J., Schneider, R., Bauer, U.-M., Bannister, A.J., Morrison, A., O'Carroll, D., Firensten, R., Cleary, M., Jenuwein, T., Herrera, R.E. and Kouzarides, T. (2001b) Rb targets histone H3 methylation and HP1 to promoters. *Nature*, **412**, 561-565.
398. Nissen, M.S., Langan, T.A. and Reeves, R. (1991) Phosphorylation by cdc2 kinase modulates DNA binding activity of high mobility group I nonhistone chromatin protein. *J Biol Chem*, **266**, 19945-19952.
399. Nomura, M. (2001) Ribosomal RNA genes, RNA polymerases, nucleolar structures, and synthesis of rRNA in the yeast Saccharomyces cerevisiae. *Cold Spring Harb Symp Quant Biol*, **66**, 555-565.
400. O'Hare, K., Chadwick, B.P., Constantinou, A., Davis, A.J., Mitchelson, A. and Tudor, M. (2002) A 5.9-kb tandem repeat at the euchromatin-heterochromatin boundary of the X chromosome of Drosophila melanogaster. *Mol Genet Genomics*, **267**, 647-655.
401. Okano, M., Bell, D.W., Haber, D.A. and Li, E. (1999) DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell*, **99**, 247-257.
402. Olins, A.L. and Olins, D.E. (1974) Spheroid chromatin units (v bodies). *Science*, **183**, 330-332.
403. Oudet, P., Gross-Bellard, M. and Chambon, P. (1975) Electron microscopic and biochemical evidence that chromatin structure is a repeating unit. *Cell*, **4**, 281-300.
404. Paik, W.K. and Kim, S. (1973) Enzymatic demethylation of calf thymus histones. *Biochem Biophys Res Commun*, **51**, 781-788.
405. Pal-Bhadra, M., Leibovitch, B.A., Gandhi, S.G., Rao, M., Bhadra, U., Birchler, J.A. and Elgin, S.C. (2004) Heterochromatic silencing and HP1 localization in Drosophila are dependent on the RNAi machinery. *Science*, **303**, 669-672.
406. Pardue, M.L. (1986) *In situ hybridization* to DNA of chromosomes and nuclei. *Drosophila: a practical approach.*, 111-137.

407. Park, H.S. and Yamamoto, M.T. (1995) The centric region of the X chromosome rDNA functions in male meiotic pairing in Drosophila melanogaster. *Chromosoma*, **103**, 700-707.
408. Paro, R. and 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**, 263-267.
409. Pauling, L. and Corey, R.B. (1953) Structure of the nucleic acids. *Nature*, **171**, 346.
410. Paulson, J.R. and Laemmli, U.K. (1977) The structure of histone-depleted metaphase chromosomes. *Cell*, **12**, 817-828.
411. Peacock, W.J., Brutlag, D., Goldring, E., Appels, R., Hinton, C.W. and Lindsley, D.L. (1974) The organization of highly repeated DNA sequences in Drosophila melanogaster chromosomes. *Cold Spring Harb Symp Quant Biol*, **38**, 405-416.
412. Perez-Gonzalez, C.E., Burke, W.D. and Eickbush, T.H. (2003) R1 and R2 retrotransposition and deletion in the rDNA loci on the X and Y chromosomes of *Drosophila melanogaster*. *Genetics*, **165**, 675-685.
413. Perez-Gonzalez, C.E. and Eickbush, T.H. (2002) Rates of R1 and R2 retrotransposition and elimination from the rDNA locus of *Drosophila melanogaster*. *Genetics*, **162**, 799-811.
414. Perrini, B., Piacentini, L., Fanti, L., Altieri, F., Chichiarelli, S., Berloco, M., Turano, C., Ferraro, A. and Pimpinelli, S. (2004) HP1 controls telomere capping, telomere elongation, and telomere silencing by two different mechanisms in Drosophila. *Mol Cell*, **15**, 467-476.
415. Peters, A.H., Mermoud, J.E., O'Carroll, D., Pagani, M., Schweizer, D., Brockdorff, N. and Jenuwein, T. (2002) Histone H3 lysine 9 methylation is an epigenetic imprint of facultative heterochromatin. *Nat Genet*, **30**, 77-80.
416. Peters, A.H.F.M., Kubicek, S., Mechtler, K., O'Sullivan, R.J., Derijck, A.A.H.A., Perez-Burgos, L., Kohlmaier, A., Opravil, S., Tachibana, M., Shinkai, Y., Martens, J.H.A. and Jenuwein, T. (2003) Partitioning and plasticity or repressive histone methylation states in mammalian chromatin. *Mol. Cell*, **12**, 1577-1589.
417. Peters, J.M. (2002) The anaphase-promoting complex: proteolysis in mitosis and beyond. *Mol. Cell*, **9**, 931-943.
418. Peterson, C.L. (2002a) Chromatin remodeling enzymes: taming the machines. Third in review series on chromatin dynamics. *EMBO Rep*, **3**, 319-322.
419. Peterson, C.L. (2002b) Chromatin remodeling: nucleosomes bulging at the seams. *Curr Biol*, **12**, R245-247.
420. Peterson, C.L. (2002c) HDAC's at Work: Everyone Doing Their Part. *Mol Cell*, **9**, 921-929.
421. Peterson, C.L. and Laniel, M.A. (2004) Histones and histone modifications. *Curr Biol*, **14**, R546-551.
422. Phelan, M.L., Sif, S., Narlikar, G.J. and Kingston, R.E. (1999) Reconstitution of a core chromatin remodeling complex from SWI/SNF subunits. *Mol Cell*, **3**, 247-253.
423. Piacentini, L., Fanti, L., Bercolo, M., Perrini, B. and Pimpinelli, S. (2003) Heterochromatin protein 1 (HP1) is associated with induced gene expression in *Drosophila* euchromatin. *J Cell Biol*, **161**, 707-714.
424. Pina, B. and Suau, P. (1987) Changes in histones H2A and H3 variant composition in differentiating and mature rat brain cortical neurons. *Dev Biol*, **123**, 51-58.
425. Pirrotta, V. (1997) Chromatin-silencing mechanisms in Drosophila maintain patterns of gene expression. *Trends Genet*, **13**, 314-318.
426. Poljak, L. and Käs, E. (1995) Resolving the role of topoisomerase II in chromatin structure and function. *Trends Cell Biol.*, **5**, 348-354.

427. Porter, A.C. and Farr, C.J. (2004) Topoisomerase II: untangling its contribution at the centromere. *Chromosome Res*, **12**, 569-583.
428. Prior, C.P., Cantor, C.R., Johnson, E.M., Littau, V.C. and Allfrey, V.G. (1983) Reversible changes in nucleosome structure and histone H3 accessibility in transcriptionally active and inactive states of rDNA chromatin. *Cell*, **34**, 1033-1042.
429. Radic, M.Z., Saghbini, M., Elton, T.S., Reeves, R. and Hamkalo, B.A. (1992) Hoechst 33258, distamycin A, and high mobility group protein I (HMG-I) compete for binding to mouse satellite DNA. *Chromosoma*, **101**, 602-608.
430. Raisner, R.M., Hartley, P.D., Meneghini, M.D., Bao, M.Z., Liu, C.L., Schreiber, S.L., Rando, O.J. and Madhani, H.D. (2005) Histone variant H2A.Z marks the 5' ends of both active and inactive genes in euchromatin. *Cell*, **123**, 233-248.
431. Ram, T.G., Reeves, R. and Hosick, H.L. (1993) Elevated high mobility group-I(Y) gene expression is associated with progressive transformation of mouse mammary epithelial cells. *Cancer Res*, **53**, 2655-2660.
432. Rangasamy, D., Berven, L., Ridgway, P. and Tremethick, D.J. (2003) Pericentric heterochromatin becomes enriched with H2A.Z during early mammalian development. *Embo J*, **22**, 1599-1607.
433. Rangasamy, D., Greaves, I. and Tremethick, D.J. (2004) RNA interference demonstrates a novel role for H2A.Z in chromosome segregation. *Nat Struct Mol Biol*, **11**, 650-655.
434. Ray-Gallet, D., Quivy, J.P., Scamps, C., Martini, E.M., Lipinski, M. and Almouzni, G. (2002) HIRA is critical for a nucleosome assembly pathway independent of DNA synthesis. *Mol Cell*, **9**, 1091-1100.
435. Rayman, J.B., Takahashi, Y., Indjeian, V.B., Dannenberg, J.H., Catchpole, S., Watson, R.J., te Riele, H. and Dynlacht, B.D. (2002) E2F mediates cell cycle-dependent transcriptional repression in vivo by recruitment of an HDAC1/mSin3B corepressor complex. *Genes Dev*, **16**, 933-947.
436. Rea, S., Eisenhaber, F., O'Carroll, D., Strahl, B.D., Sun, Z., Schmid, M., Opravil, S., Mechtler, K., Ponting, C.P., Allis, C.D. and Jenuwein, T. (2000) Regulation of chromatin structure by site specific histone H3 methyltransferases. *Nature*, **406**, 593-599.
437. Reeder, R.H. (1974) In Nomura, M. (ed.), *Ribosomes*. Cold Spring Harbor Laboratory Press, New York, pp. 489-519.
438. Reeves, R., Langan, T.A. and Nissen, M.S. (1991) Phosphorylation of the DNA-binding domain of nonhistone high-mobility group I protein by cdc2 kinase: reduction of binding affinity. *Proc Natl Acad Sci U S A*, **88**, 1671-1675.
439. Reeves, R. and Nissen, M.S. (1990) The A.T-DNA-binding domain of mammalian high mobility group I chromosomal proteins. A novel peptide motif for recognizing DNA structure. *J Biol Chem*, **265**, 8573-8582.
440. Reeves, R. and Nissen, M.S. (1993) Interaction of high mobility group-I (Y) nonhistone proteins with nucleosome core particles. *J Biol Chem*, **268**, 21137-21146.
441. Reeves, R. and Nissen, M.S. (1995) Cell cycle regulation and functions of HMG-I(Y). *Prog Cell Cycle Res*, **1**, 339-349.
442. Reinhart, B.J. and Bartel, D.P. (2002) Small RNAs correspond to centromere heterochromatic repeats. *Science*, **297**, 1831.
443. Reuter, G., Giarre, M., Farah, J., Gausz, J., Spierer, A. and Spierer, P. (1990) Dependence of position-effect variegation in Drosophila on dose of a gene encoding an unusual zinc-finger protein. *Nature*, **344**, 219-223.
444. Reuter, G., Werner, W. and Hoffmann, H.J. (1982) Mutants affecting position-effect heterochromatinization in Drosophila melanogaster. *Chromosoma*, **85**, 539-551.

445. Reuter, G. and Wolff, I. (1981) Isolation of dominant suppressor mutations for position-effect variegation in Drosophila melanogaster. *Mol Gen Genet*, **182**, 516-519.
446. Rice, J.C., Briggs, S.D., Ueberheide, B., Barber, C.M., Shabanowitz, J., Hunt, D.F., Shinkai, Y. and Allis, C.D. (2003) Histone methyltransferases direct different degrees of methylation to define distinct chromatin domains. *Mol. Cell*, **12**, 1591-1598.
447. Richard, P., Darzacq, X., Bertrand, E., Jady, B.E., Verheggen, C. and Kiss, T. (2003) A common sequence motif determines the Cajal body-specific localization of box H/ACA scaRNAs. *Embo J*, **22**, 4283-4293.
448. Richards, E.J. and Elgin, S.C. (2002) Epigenetic codes for heterochromatin formation and silencing: rounding up the usual suspects. *Cell*, **108**, 489-500.
449. Richmond, T.J. and Davey, C.A. (2003) The structure of DNA in the nucleosome core. *Nature*, **423**, 145-150.
450. Rine, J. and Herskowitz, I. (1987) Four genes responsible for a position effect on expression from HML and HMR in Saccharomyces cerevisiae. *Genetics*, **116**, 9-22.
451. Ritossa, F.M. and Spiegelman, S. (1965) Localization of DNA Complementary to Ribosomal Rna in the Nucleolus Organizer Region of Drosophila Melanogaster. *Proc Natl Acad Sci U S A*, **53**, 737-745.
452. Roark, D.E., Geoghegan, T.E. and Keller, G.H. (1974) A two-subunit histone complex from calf thymus. *Biochem. Biophys. Res. Commun.*, **59**, 542-547.
453. Rodriguez Alfageme, C., Rudkin, G.T. and Cohen, L. (1980) Isolation, properties and cellular distribution of D1, a chromosomal protein of *Drosophila*. *Chromosoma*, **78**, 1-31.
454. Rogakou, E.P., Pilch, D.R., Orr, A.H., Ivanova, V.S. and Bonner, W.M. (1998) DNA double-stranded breaks induce histone H2AX phosphorylation on serine 139. *J Biol Chem*, **273**, 5858-5868.
455. Rollins, R.A., Morcillo, P. and Dorsett, D. (1999) Nipped-B, a Drosophila homologue of chromosomal adherins, participates in activation by remote enhancers in the cut and Ultrabithorax genes. *Genetics*, **152**, 577-593.
456. Roseman, R.R., Johnson, E.A., Rodesch, C.K., Bjerke, M., Nagoshi, R.N. and Geyer, P.K. (1995a) A P element containing suppressor of hairy-wing binding regions has novel properties for mutagenesis in Drosophila melanogaster. *Genetics*, **141**, 1061-1074.
457. Roseman, R.R., Swan, J.M. and Geyer, P.K. (1995b) A Drosophila insulator protein facilitates dosage compensation of the X chromosome min-white gene located at autosomal insertion sites. *Development*, **121**, 3573-3582.
458. Saccani, S. and Natoli, G. (2002) Dynamic changes in histone H3 Lys 9 methylation occurring at tightly regulated inducible inflammatory genes. *Genes Dev*, **16**, 2219-2224.
459. Sado, T., Wang, Z., Sasaki, H. and Li, E. (2001) Regulation of imprinted X-chromosome inactivation in mice by Tsix. *Development*, **128**, 1275-1286.
460. Saitoh, Y. and Laemmli, U.K. (1994) Metaphase chromosome structure: bands arise from a differential folding path of the highly AT-rich scaffold. *Cell*, **76**, 609-622.
461. Santoro, R. and Grummt, I. (2001) Molecular mechanisms mediating methylation-dependent silencing of ribosomal gene transcription. *Mol Cell*, **8**, 719-725.
462. Santoro, R., Li, J. and Grummt, I. (2002) The nucleolar remodeling complex NoRC mediates heterochromatin formation and silencing of ribosomal gene transcription. *Nat Genet*, **32**, 393-396.
463. Santos-Rosa, H., Schneider, R., Bannister, A.J., Sherriff, J., Bernstein, B.E., Emre, N.C., Schreiber, S.L., Mellor, J. and Kouzarides, T. (2002) Active genes are tri-methylated at K4 of histone H3. *Nature*, **419**, 407-411.

464. Sarma, K. and Reinberg, D. (2005) Histone variants meet their match. *Nat Rev Mol Cell Biol*, **6**, 139-149.
465. Satou, W., Suzuki, T., Noguchi, T., Ogino, H., Fujii, M. and Ayusawa, D. (2004) AT-hook proteins stimulate induction of senescence markers triggered by 5-bromodeoxyuridine in mammalian cells. *Exp Gerontol*, **39**, 173-179.
466. Schmid, M., Arib, G., Laemmli, C., Nishikawa, J., Durussel, T. and Laemmli, U.K. (2006) Nup-PI: the nucleopore-promoter interaction of genes in yeast. *Mol Cell*, **21**, 379-391.
467. Schotta, G., Ebert, A., Krauss, V., Fischer, A., Hoffmann, J., Rea, S., Jenuwein, T., Dorn, R. and Reuter, G. (2002) Central role of Drosophila SU(VAR)3-9 in histone H3-K9 methylation and heterochromatic gene silencing. *Embo J*, **21**, 1121-1131.
468. Schotta, G., Ebert, A. and Reuter, G. (2003) SU(VAR)3-9 is a conserved key function in heterochromatic gene silencing. *Genetica*, **117**, 149-158.
469. Schotta, G., Lachner, M., Sarma, K., Ebert, A., Sengupta, R., Reuter, G., Reinberg, D. and Jenuwein, T. (2004) A silencing pathway to induce H3-K9 and H4-K20 trimethylation at constitutive heterochromatin. *Genes & Dev.*, **18**, 1251-1262.
470. Schultz, D.C., Ayyanathan, K., Negorev, D., Maul, G.G. and 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**, 919-932.
471. Schultz, J. (1936) Variegation in Drosophila and the inert chromosome regions. *Proc. Natl. Acad. Sci. USA*, **22**, 27-33.
472. Schultz, J. (1939) The function of heterochromatin. *Proc. VII Int. Congr. Genet.*, pp. 257-262.
473. Schulze, S.R., Sinclair, D.A., Fitzpatrick, K.A. and Honda, B.M. (2005) A genetic and molecular characterization of two proximal heterochromatic genes on chromosome 3 of Drosophila melanogaster. *Genetics*, **169**, 2165-2177.
474. Schwartz, B.E. and Ahmad, K. (2005) Transcriptional activation triggers deposition and removal of the histone variant H3.3. *Genes Dev*, **19**, 804-814.
475. Seum, C., Delattre, M., Spierer, A. and Spierer, P. (2001) Ectopic HP1 promotes chromosome loops and variegated silencing in *Drosophila*. *EMBO J*, **20**, 812-818.
476. Seum, C., Pauli, D., Delattre, M., Jaquet, Y., Spierer, A. and Spierer, P. (2002) Isolation of Su(var)3-7 mutations by homologous recombination in Drosophila melanogaster. *Genetics*, **161**, 1125-1136.
477. Shaffer, C.D., Stephens, G.E., Thompson, B.A., Funches, L., Bernat, J.A., Craig, C.A. and Elgin, S.C. (2002) Heterochromatin protein 2 (HP2), a partner of HP1 in *Drosophila* heterochromatin. *PNAS*, **99**, 14332-14337.
478. Shapiro, T.A. and Englund, P.T. (1990) Selective cleavage of kinetoplast DNA minicircles promoted by antitrypanosomal drugs. *Proc Natl Acad Sci U S A*, **87**, 950-954.
479. Shen, X. and Gorovsky, M.A. (1996) Linker histone H1 regulates specific gene expression but not global transcription in vivo. *Cell*, **86**, 475-483.
480. Shi, Y., Lan, F., Matson, C., Mulligan, P., Whetstine, J.R., Cole, P.A., Casero, R.A. and Shi, Y. (2004) Histone Demethylation Mediated by the Nuclear Amine Oxidase Homolog LSD1. *Cell*, **119**, 941-953.
481. Shiio, Y. and Eisenman, R.N. (2003) Histone sumoylation is associated with transcriptional repression. *Proc Natl Acad Sci U S A*, **100**, 13225-13230.
482. Sijen, T., Fleenor, J., Simmer, F., Thijssen, K.L., Parrish, S., Timmons, L., Plasterk, R.H. and Fire, A. (2001) On the role of RNA amplification in dsRNA-triggered gene silencing. *Cell*, **107**, 465-476.

483. Sinclair, D.A.R., Mottus, R.C. and Grigliatti, T.A. (1983) Genes which suppress position-effect variegation in *Drosophila melanogaster* are clustered. *Mol Gen Genet*, **191**, 326-333.
484. Singer, M.F. (1982) Highly repeated sequences in mammalian genomes. *Int Rev Cytol*, **76**, 67-112.
485. Slama-Schwok, A., Zakrzewska, K., Leger, G., Leroux, Y., Takahashi, M., Käs, E. and Debey, P. (2000) Structural changes induced by binding of the High Mobility Group I protein to a mouse satellite DNA sequence. *Biophys J*, **78**, 2543-2559.
486. Smith, E.R., Pannuti, A., Gu, W., Steurnagel, A., Cook, R.G., Allis, C.D. and Lucchesi, J.C. (2000) The drosophila MSL complex acetylates histone H4 at lysine 16, a chromatin modification linked to dosage compensation. *Mol Cell Biol*, **20**, 312-318.
487. Smith, J.S. and Boeke, J.D. (1997) An unusual form of transcriptional silencing in yeast ribosomal DNA. *Genes Dev*, **11**, 241-254.
488. Smith, M.M. (2002) Centromeres and variant histones: what, where, when and why? *Curr Opin Cell Biol*, **14**, 279-285.
489. Smith, S. and Stillman, B. (1989) Purification and characterization of CAF-I, a human cell factor required for chromatin assembly during DNA replication in vitro. *Cell*, **58**, 15-25.
490. Smothers, J.F. and Henikoff, S. (2000) The HP1 chromo shadow domain binds a consensus peptide pentamer. *Curr. Biol.*, **10**, 27-30.
491. Smothers, J.F. and Henikoff, S. (2001) The hinge and chromo shadow domain impart distinct targeting of HP1-like proteins. *Mol. Cell. Biol.*, **21**, 2555-2569.
492. Sobel, R.E., Cook, R.G., Perry, C.A., Annunziato, A.T. and Allis, C.D. (1995) Conservation of deposition-related acetylation sites in newly synthetized histones H3 and H4. *Proc. Natl. Acad. Sci. USA*, **92**, 1237-1241.
493. Spierer, A., Seum, C., Delattre, M. and Spierer, P. (2005) Loss of the modifiers of variegation Su(var)3-7 or HP1 impacts male X polytene chromosome morphology and dosage compensation. *J Cell Sci*, **118**, 5047-5057.
494. Stanley, J.S., Griffin, J.B. and Zempleni, J. (2001) Biotinylation of histones in human cells. Effects of cell proliferation. *Eur J Biochem*, **268**, 5424-5429.
495. Stassen, M.J., Bailey, D., Nelson, S., Chinwalla, V. and Harte, P.J. (1995) The Drosophila trithorax proteins contain a novel variant of the nuclear receptor type DNA binding domain and an ancient conserved motif found in other chromosomal proteins. *Mech Dev*, **52**, 209-223.
496. Stephens, G.E., Slawson, E.E., Craig, C.A. and Elgin, S.C. (2005) Interaction of heterochromatin protein 2 with HP1 defines a novel HP1-binding domain. *Biochemistry*, **44**, 13394-13403.
497. Sterner, D.E. and Berger, S.L. (2000) Acetylation of histones and transcription-related factors. *Microbiol Mol Biol Rev*, **64**, 435-459.
498. Stewart, M.D., Li, J. and Wong, J. (2005) Relationship between histone H3 lysine 9 methylation, transcription repression, and heterochromatin protein 1 recruitment. *Mol Cell Biol*, **25**, 2525-2538.
499. Strahl, B.D. and Allis, C.D. (2000) The language of covalent histone modifications. *Nature*, **403**, 41-45.
500. Strick, R. and Laemmli, U.K. (1995) SARs are cis DNA elements of chromosome dynamics: synthesis of a SAR repressor protein. *Cell*, **83**, 1137-1148.
501. Struhl, G. and Greenwald, I. (1999) Presenilin is required for activity and nuclear access of Notch in Drosophila. *Nature*, **398**, 522-525.
502. Sturtevant, A.H. (1913) The linear arrangement of six sex-linked factors in Drosophila, as shown by their mode of association. *J. Exp. Zool.*, **14**, 43-59.

503. Sturtevant, A.H. and Morgan, T.H. (1923) Reverse mutation of the Bar gene correlated with crossing over. *Science*, **57**, 746-747.
504. Suka, N., Luo, K. and Grunstein, M. (2002) Sir2p and Sas2p opposingly regulate acetylation of yeast histone H4 lysine16 and spreading of heterochromatin. *Nat Genet*, **32**, 378-383.
505. Sun, F.L., Cuaycong, M.H. and Elgin, S.C. (2001) Long-range nucleosome ordering is associated with gene silencing in Drosophila melanogaster pericentric heterochromatin. *Mol Cell Biol*, **21**, 2867-2879.
506. Sun, Z.W. and Allis, C.D. (2002) Ubiquitination of histone H2B regulates H3 methylation and gene silencing in yeast. *Nature*, **418**, 104-108.
507. Susbielle, G., Blattes, R., Brevet, V., Monod, C. and Käs, E. (2005) Target practice: aiming at satellite repeats with DNA minor groove binders. *Curr. Med. Chem. - Anti Cancer Agents.*, **5**, 409-420.
508. Sutton, W.S. (1903) The chromosomes in heredity. *Biological Bulletin*, **4**, 231-251.
509. Suzuki, T., Yaginuma, M., Oishi, T., Michishita, E., Ogino, H., Fujii, M. and Ayusawa, D. (2001) 5-Bromodeoxyuridine suppresses position effect variegation of transgenes in HeLa cells. *Exp Cell Res*, **266**, 53-63.
510. Swaminathan, J., Baxter, E.M. and Corces, V.G. (2005) The role of histone H2Av variant replacement and histone H4 acetylation in the establishment of Drosophila heterochromatin. *Genes Dev*, **19**, 65-76.
511. Swedlow, J.R., Sedat, J.W. and Agard, D.A. (1993) Multiple chromosomal populations of topoisomerase II detected in vivo by time-lapse, three-dimensional wide-field microscopy. *Cell*, **73**, 97-108.
512. Tachibana, M., Sugimoto, K., Fukushima, T. and Shinkai, Y. (2001) Set domain-containing protein, G9a, is a novel lysine-preferring mammalian histone methyltransferase with hyperactivity and specific selectivity to lysines 9 and 27 of histone H3. *J Biol Chem*, **276**, 25309-25317.
513. Tachibana, M., Sugimoto, K., Nozaki, M., Ueda, J., Ohta, T., Ohki, M., Fukuda, M., Takeda, N., Niida, H., Kato, H. and Shinkai, Y. (2002) G9a histone methyltransferase plays a dominant role in euchromatic histone H3 lysine 9 methylation and is essential for early embryogenesis. *Gene Dev*, **16**, 1779-1791.
514. Taddei, A., Maison, C., Roche, D. and Almouzni, G. (2001) Reversible disruption of pericentric heterochromatin and centromere function by inhibiting deacetylases. *Nat Cell Biol*, **3**, 114-120.
515. Tagami, H., Ray-Gallet, D., Almouzni, G. and Nakatani, Y. (2004) Histone H3.1 and H3.3 complexes mediate nucleosome assembly pathways dependent or independent of DNA synthesis. *Cell*, **116**, 51-61.
516. Tamkun, J.W., Deuring, R., Scott, M.P., Kissinger, M., Pattatucci, A.M., Kaufman, T.C. and Kennison, J.A. (1992) brahma: a regulator of Drosophila homeotic genes structurally related to the yeast transcriptional activator SNF2/SWI2. *Cell*, **68**, 561-572.
517. Tanaka, T.U., Rachidi, N., Janke, C., Pereira, G., Galova, M., Schiebel, E., Stark, M.J. and Nasmyth, K. (2002) Evidence that the Ipl1-Sli15 (Aurora kinase-INCENP) complex promotes chromosome bi-orientation by altering kinetochore-spindle pole connections. *Cell*, **108**, 317-329.
518. Tanaka, Y., Tachiwana, H., Yoda, K., Masumoto, H., Okazaki, T., Kurumizaka, H. and Yokoyama, S. (2005) Human centromere protein B induces translational positioning of nucleosomes on alpha-satellite sequences. *J Biol Chem*, **280**, 41609-41618.
519. Tartof, K.D. (1994) Position effect variegation in yeast. *Bioessays*, **16**, 713-714.

520. Tartof, K.D., Hobbs, C. and Jones, M. (1984) A structural basis for variegating position effects. *Cell*, **37**, 869-878.
521. Thanos, D., Du, W. and Maniatis, T. (1993) The high mobility group protein HMG I(Y) is an essential structural component of a virus-inducible enhancer complex. *Cold Spring Harb Symp Quant Biol*, **58**, 73-81.
522. Thorne, A.W., Sautiere, P., Briand, G. and Crane-Robinson, C. (1987) The structure of ubiquitinated histone H2B. *Embo J*, **6**, 1005-1010.
523. Tidwell, R.R., Boykin, D.W., Demeunynck, M., Bailly, C. and Wilson, W.D. (2003) In DNA and RNA binders/ from small molecules to drugs. *Wiley VCH*, **2**, 414-460.
524. Tie, F., Furuyama, T., Prasad-Sinha, J., Jane, E. and Harte, P.J. (2001) The Drosophila Polycomb Group proteins ESC and E(Z) are present in a complex containing the histone-binding protein p55 and the histone deacetylase RPD3. *Development*, **128**, 275-286.
525. Tong, J.K., Hassig, C.A., Schnitzler, G.R., Kingston, R.E. and Schreiber, S.L. (1998) Chromatin deacetylation by an ATP-dependent nucleosome remodelling complex. *Nature*, **395**, 917-921.
526. Toussaint, M., Levasseur, G., Tremblay, M., Paquette, M. and Conconi, A. (2005) Psoralen photocrosslinking, a tool to study the chromatin structure of RNA polymerase I - transcribed ribosomal genes. *Biochem Cell Biol*, **83**, 449-459.
527. Travers, A.A. and Thompson, J.M. (2004) An introduction to the mechanics of DNA. *Philos Transact A Math Phys Eng Sci*, **362**, 1265-1279.
528. Tschiersch, B., Hofmann, A., Krauss, V., Dorn, R., Korge, G. and Reuter, G. (1994) The protein encoded by the Drosophila position-effect variegation suppressor gene Su(var)3-9 combines domains of antagonistic regulators of homeotic gene complexes. *Embo J*, **13**, 3822-3831.
529. Tsukada, Y., Fang, J., Erdjument-Bromage, H., Warren, M.E., Borchers, C.H., Tempst, P. and Zhang, Y. (2006) Histone demethylation by a family of JmjC domain-containing proteins. *Nature*, **439**, 811-816.
530. Tsukiyama, T., Daniel, C., Tamkun, J. and Wu, C. (1995) ISWI, a member of the SWI2/SNF2 ATPase family, encodes the 140 kDa subunit of the nucleosome remodeling factor. *Cell*, **83**, 1021-1026.
531. Tsukiyama, T. and Wu, C. (1995) Purification and properties of an ATP-dependent nucleosome remodeling factor. *Cell*, **83**, 1011-1020.
532. Udvardy, A., Maine, E. and Schedl, P. (1985) The 87A7 chromomere. Identification of novel chromatin structures flanking the heat shock locus that may define the boundaries of higher order domains. *J Mol Biol*, **185**, 341-358.
533. Utley, R.T. and Cote, J. (2003) The MYST family of histone acetyltransferases. *Curr Top Microbiol Immunol*, **274**, 203-236.
534. Vaistij, F.E., Jones, L. and Baulcombe, D.C. (2002) Spreading of RNA targeting and DNA methylation in RNA silencing requires transcription of the target gene and a putative RNA-dependent RNA polymerase. *Plant Cell*, **14**, 857-867.
535. Vakoc, C.R., Mandat, S.A., Olenchock, B.A. and Blobel, G.A. (2005) Histone H3 lysine 9 methylation and HP1gamma are associated with transcription elongation through mammalian chromatin. *Mol Cell*, **19**, 381-391.
536. Valgardsdottir, R., Chiodi, I., Giordano, M., Cobianchi, F., Riva, S. and Biamonti, G. (2005) Structural and functional characterization of noncoding repetitive RNAs transcribed in stressed human cells. *Mol Biol Cell*, **16**, 2597-2604.
537. Van Holde, K. and Zlatanova, J. (1995) Chromatin higher order structure: chasing a mirage? *J. Biol. Chem.*, **270**, 8373-8376.

538. Van Hooser, A., Goodrich, D.W., Allis, C.D., Brinkley, B.R. and Mancini, M.A. (1998) Histone H3 phosphorylation is required for the initiation, but not maintenance, of mammalian chromosome condensation. *J Cell Sci*, **111 (Pt 23)**, 3497-3506.
539. Van Hooser, A.A., Ouspenski, II, Gregson, H.C., Starr, D.A., Yen, T.J., Goldberg, M.L., Yokomori, K., Earnshaw, W.C., Sullivan, K.F. and Brinkley, B.R. (2001) Specification of kinetochore-forming chromatin by the histone H3 variant CENP-A. *J Cell Sci*, **114**, 3529-3542.
540. Vandel, L., Nicolas, E., Vaute, O., Ferreira, R., Ait-Si-Ali, S. and Trouche, D. (2001) Transcriptional repression by the retinoblastoma protein through the recruitment of a histone methyltransferase. *Mol Cell Biol*, **21**, 6484-6494.
541. Varga-Weisz, P.D., Wilm, M., Bonte, E., Dumas, K., Mann, M. and Becker, P.B. (1997) Chromatin-remodelling factor CHRAC contains the ATPases ISWI and topoisomerase II. *Nature*, **388**, 598-602.
542. Vaute, O., Nicolas, E., Vandel, L. and Trouche, D. (2002) Functional and physical interaction between the histone methyl transferase Suv39H1 and histone deacetylases. *Nucleic Acids Res*, **30**, 475-481.
543. Vazquez, J., Pauli, D. and Tissieres, A. (1993) Transcriptional regulation in Drosophila during heat shock: a nuclear run-on analysis. *Chromosoma*, **102**, 233-248.
544. Verdel, A., Jia, S., Gerber, S., Sugiyama, T., Gygi, S., Grewal, S.I.S. and Moazed, D. (2004) RNAi-mediated targeting of heterochromatin by the RITS complex. *Science*, **303**, 672-676.
545. Verkade, H.M., Bugg, S.J., Lindsay, H.D., Carr, A.M. and O'Connell, M.J. (1999) Rad18 is required for DNA repair and checkpoint responses in fission yeast. *Mol. Biol. Cell*, **10**, 2905-2918.
546. Vignali, M., Hassan, A.H., Neely, K.E. and Workman, J.L. (2000) ATP-dependent chromatin-remodeling complexes. *Mol Cell Biol*, **20**, 1899-1910.
547. Volpe, T.A., Kidner, C., Hall, I.M., Teng, G., Grewal, S.I.S. and Martienssen, M. (2002) Regulation of heterochromatic silencing and histone H3 lysine-9 methylation by RNAi. *Science*, **297**, 1833-1837.
548. von Kries, J.P., Buhrmester, H. and Stratling, W.H. (1991) A matrix/scaffold attachment region binding protein: identification, purification, and mode of binding. *Cell*, **64**, 123-135.
549. Waldeyer, W. (1888) Ueber Karyokinese und ihre Beziehungen zu den Befruchtungsvorgängen. *Arch Mikr Anat.*, **32**, 1-122.
550. Wallrath, L.L. (1998) Unfolding the mysteries of heterochromatin. *Curr Opin Genet Dev*, **8**, 147-153.
551. Wang, X., Watt, P.M., Borts, R.H., Louis, E.J. and Hickson, I.D. (1999) The topoisomerase II-associated protein, Pat1p, is required for maintenance of rDNA locus stability in Saccharomyces cerevisiae. *Mol Gen Genet*, **261**, 831-840.
552. Wang, Y., Wysocka, J., Sayegh, J., L., Y.-H., Perlin, J.R., Lauriebeth, L., Sonbuchner, L.S., McDonald, C.H., Cook, R.G., Dou, Y., Roeder, R.G., Clarke, S., Stallcup, M.R., Allis, C.D. and Coonrod, S.A. (2004) Human PAD4 regulates histone arginine methylation levels via demethylimination. *Science*, **306**, 279-283.
553. Wang, Y., Zhang, W., Jin, Y., Johansen, J. and Johansen, K.M. (2001) The JIL-1 tandem kinase mediates histone H3 phosphorylation and is required for maintenance of chromatin structure in Drosophila. *Cell*, **105**, 433-443.
554. Watson, J.D. and Crick, F.H. (1953) Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature*, **171**, 737-738.

555. Waye, J.S. and Willard, H.F. (1987) Nucleotide sequence heterogeneity of alpha satellite repetitive DNA: a survey of alphoid sequences from different human chromosomes. *Nucleic Acids Res*, **15**, 7549-7569.
556. Waye, J.S. and Willard, H.F. (1989) Human beta satellite DNA: genomic organization and sequence definition of a class of highly repetitive tandem DNA. *Proc Natl Acad Sci U S A*, **86**, 6250-6254.
557. Wellinger, R.J. and Sen, D. (1997) The DNA structures at the ends of eukaryotic chromosomes. *Eur J Cancer*, **33**, 735-749.
558. Wemmer, D.E. and Dervan, P.B. (1997) Targeting the minor groove of DNA. *Curr Opin Struct Biol*, **7**, 355-361.
559. Willard, H.F. and Waye, J.S. (1987) Chromosome-specific subsets of human alpha satellite DNA: analysis of sequence divergence within and between chromosomal subsets and evidence for an ancestral pentameric repeat. *J Mol Evol*, **25**, 207-214.
560. Williamson, R. (1970) Properties of rapidly labelled deoxyribonucleic acid fragments isolated from the cytoplasm of primary cultures of embryonic mouse liver cells. *J. Mol. Biol.*, **51**, 157-168.
561. Wilson, W.D., Nguyen, B., Tanious, F.A., Mathis, A., Hall, J.E., Stephens, C.E. and Boykin, D.W. (2005) Dications that target the DNA minor groove: compound design and preparation, DNA interactions, cellular distribution and biological activity. *Curr Med Chem Anticancer Agents*, **5**, 389-408.
562. Wolffe, A.P. (1994) Architectural transcription factors. *Science*, **264**, 1100-1101.
563. Wood, A., Krogan, N.J., Dover, J., Schneider, J., Heidt, J., Boateng, M.A., Dean, K., Golshani, A., Zhang, Y., Greenblatt, J.F., Johnston, M. and Shilatifard, A. (2003) Bre1, an E3 ubiquitin ligase required for recruitment and substrate selection of Rad6 at a promoter. *Mol Cell*, **11**, 267-274.
564. Woodcock, C.L. and Dimitrov, S. (2001) Higher-order structure of chromatin and chromosomes. *Curr Opin Genet Dev*, **11**, 130-135.
565. Woodcock, C.L., Grigoryev, S.A., Horowitz, R.A. and Whitaker, N. (1993) A chromatin folding model that incorporates linker variability generates fibers resembling the native structures. *Proc. Natl. Acad. Sci. USA*, **90**, 9021-9025.
566. Woodcock, C.L., Skoultchi, A.I. and Fan, Y. (2006) Role of linker histone in chromatin structure and function: H1 stoichiometry and nucleosome repeat length. *Chromosome Res*, **14**, 17-25.
567. Woynarowski, J.M., McHugh, M., Sigmund, R.D. and Beerman, T.A. (1989a) Modulation of topoisomerase II catalytic activity by DNA minor groove binding agents distamycin, Hoechst 33258, and 4',6-diamidine-2-phenylindole. *Mol Pharmacol*, **35**, 177-182.
568. Woynarowski, J.M., Sigmund, R.D. and Beerman, T.A. (1989b) DNA minor groove binding agents interfere with topoisomerase II mediated lesions induced by epipodophyllotoxin derivative VM-26 and acridine derivative m-AMSA in nuclei from L1210 cells. *Biochemistry*, **28**, 3850-3855.
569. Wu, Z.G., Murphy, C. and Gall, J.G. (1986) A transcribed satellite DNA from the bullfrog Rana catesbeiana. *Chromosoma*, **93**, 291-297.
570. Wustmann, G., Szidonya, J., Taubert, H. and Reuter, G. (1989) The genetics of position-effect variegation modifying loci in *Drosophila melanogaster*. *Mol Gen Genet*, **217**, 520-527.
571. Wysocka, J., Allis, C.D. and Coonrod, S. (2006) Histone arginine methylation and its dynamic regulation. *Front Biosci*, **11**, 344-355.

572. Xin, Z., Tachibana, M., Guggiari, M., Heard, E., Shinkai, Y. and Wagstaff, J. (2003) Role of histone methyltransferase G9a in CpG methylation of the Prader-Willi syndrome imprinting center. *J Biol Chem*, **278**, 14996-15000.
573. Xue, Y., Wong, J., Moreno, G.T., Young, M.K., Cote, J. and Wang, W. (1998) NURD, a novel complex with both ATP-dependent chromatin-remodeling and histone deacetylase activities. *Mol Cell*, **2**, 851-861.
574. Yamamoto, M.T., Mitchelson, A., Tudor, M., O'Hare, K., Davies, J.A. and Miklos, G.L. (1990) Molecular and cytogenetic analysis of the heterochromatin-euchromatin junction region of the Drosophila melanogaster X chromosome using cloned DNA sequences. *Genetics*, **125**, 821-832.
575. Yang, L., Yan, D., Bruggeman, M., Du, H. and Yan, C. (2004) Mutation of a lysine residue in a homeodomain generates dominant negative thyroid transcription factor 1. *Biochemistry*, **43**, 12489-12497.
576. Yang, W.M., Tsai, S.C., Wen, Y.D., Fejer, G. and Seto, E. (2002) Functional domains of histone deacetylase-3. *J Biol Chem*, **277**, 9447-9454.
577. Ye, Q., Callebaut, I., Pezhman, A., Courvalin, J.C. and Worman, H.J. (1997) Domain-specific interactions of human HP1-type chromodomain proteins and inner nuclear membrane protein LBR. *J Biol Chem*, **272**, 14983-14989.
578. Yen, R.W., Vertino, P.M., Nelkin, B.D., Yu, J.J., el-Deiry, W., Cumaraswamy, A., Lennon, G.G., Trask, B.J., Celano, P. and Baylin, S.B. (1992) Isolation and characterization of the cDNA encoding human DNA methyltransferase. *Nucleic Acids Res*, **20**, 2287-2291.
579. Young, D. and Carroll, D. (1983) Regular arrangement of nucleosomes on 5S rRNA genes in Xenopus laevis. *Mol Cell Biol*, **3**, 720-730.
580. Zamore, P.D. (2001) RNA interference: listening to the sound of silence. *Nat Struct Biol*, **8**, 746-750.
581. Zeleny, C. (1919) A change in the Bar gene of drosophila involving further decrease in facet number and increase in dominance. *J. Gen. Physiol.*, **2**, 69-71.
582. Zhang, S.D. and Odenwald, W.F. (1995) Misexpression of the white (w) gene triggers male-male courtship in Drosophila. *Proc Natl Acad Sci U S A*, **92**, 5525-5529.
583. Zhang, Y., Ng, H.H., Erdjument-Bromage, H., Tempst, P., Bird, A. and Reinberg, D. (1999) Analysis of the NuRD subunits reveals a histone deacetylase core complex and a connection with DNA methylation. *Genes Dev*, **13**, 1924-1935.
584. Zhang, Y. and Reinberg, D. (2001) Transcription regulation by histone methylation: interplay between different covalent modifications of the core histone tails. *Genes Dev*, **15**, 2343-2360.
585. Zhao, K., Hart, C.M. and Laemmli, U.K. (1995) Visualization of chromosomal domains with boundary element-associated factor BEAF-32. *Cell*, **81**, 879-889.
586. Zhao, K., Käs, E., Gonzalez, E. and Laemmli, U.K. (1993) SAR-dependent mobilization of histone H1 by HMG-I/Y in vitro: HMG-I/Y is enriched in H1-depleted chromatin. *EMBO J*, **12**, 3237-3247.
587. Zhao, T., Heyduk, T., Allis, D. and Eissenberg, J.C. (2000) Heterochromatin protein 1 binds to nucleosome and DNA *in vitro*. *J. Biol. Chem.*, **275**, 28332-28338.
588. Zhou, Y., Santoro, R. and Grummt, I. (2002) The chromatin remodeling complex NoRC targets HDAC1 to the ribosomal gene promoter and represses RNA polymerase I transcription. *Embo J*, **21**, 4632-4640.
589. Zink, D. and Paro, R. (1995) Drosophila Polycomb-group regulated chromatin inhibits the accessibility of a trans-activator to its target DNA. *Embo J*, **14**, 5660-5671.
590. Zlatanova, J. and Doenecke, D. (1994) Histone H1 zero: a major player in cell differentiation? *Faseb J*, **8**, 1260-1268.

I. ANNEXES

Dual-core signature of Boundary Element-Associated Factor binding sites highlights their regulatory role for hundreds of genes through epigenetic regulation of chromatin domains

Eldon Emberly*[1] , Roxane Blattes[2], Emmanuel Käs[2] and Olivier Cuvier*[3,4]

[1]Physics Department, Simon Fraser University, Burnaby, BC, Canada, V5A 1S6

[2] LBME, UMR5099 CNRS/UPS, IFR109, 31062 Toulouse, France

[3] Cold Spring Harbor Laboratory, Cold Spring Harbor, NY 11724, USA

[4]Present address: Genome Dynamics and Development, Institute for Human Genetics, CNRS, 34396 Montpelier, France

*co-corresponding authors: eemberly@sfu.ca; cuvier@igh.cnrs.fr

*these authors contributed equally to the work

ABSTRACT

Boundary Element-Associated Factor (BEAF) blocks enhancer-promoter communication and heterochromatin spreading and binds to hundreds of unknown sites in the *Drosophila* genome. Here, we used a genome-wide approach combining bioinformatic and *in vivo* analyses and identified 2094 BEAF elements sharing a highly conserved *dual-core* signature. These elements further cluster within 227 *BEAF domains* ~ 50-kbp. Transcriptome analysis supports a role for dual-cores as chromatin regulators of transcription within BEAF domains. siRNA-mediated depletion demonstrates that BEAF restricts heterochromatin spreading and regulates the heterochromatin/euchromatin ratio of epigenetic marks. Our *in situ* data identify *BEAF dual-cores* as key modules that functionally regulate higher levels of chromatin organization.

KEY WORDS

Drosophila insulators; acetylated H4; H3MeK9; 30-nm fiber; long-range communication

INTRODUCTION

A key determinant for the regulation of transcription is the accessibility of a promoter within chromatin, which is epigenetically regulated by histone modifications (Jenuwein and Allis 2001). Of interest is the identification of the *cis*-acting elements that drive the deposition of specific epigenetic marks within chromosome territories/domains to direct developmentally regulated or constitutive gene expression. Boundary elements/insulators (BE, Bell and Felsenfeld 1999; Capelson and Corces 2004) may sub-divide the genome into functional domains through epigenetic regulation (West et al. 2002). Although several groups of BEs have been identified (Cuvier et al. 1998; Ramos et al. 2006; Parnell et al. 2006), few factors are known that possess insulating activity and bind to hundreds of genomic sites, as required for their proposed function. These include CTCF in vertebrates (Bell et al. 1999), Boundary Element-Associated Factor (BEAF; Zhao et al. 1995) and *suppressor of hairy-wing* (Corces 1995) in *Drosophila* and TFIIIC in *S. cerevisiae* (Noma et al. 2006).

The *beaf* gene encodes the BEAF-32A and –32B proteins that form a BEAF complex (Zhao et al. 1995; Hart et al. 1997). BEAF blocks promiscuous interactions between enhancers and promoters (Zhao et al. 1995; Cuvier et al. 2002; Gilbert et al. 2006) as well as heterochromatin spreading in a heterologous context (Ishii and Laemmli 2003). Although BEAF was proposed to sub-divide the *Drosophila* genome into chromatin domains (Zhao et al. 1995), very few of its hundreds of genomic binding sites have been characterized (Cuvier et al. 1998). Here, we identify highly conserved BEAF elements bracketing hundreds of genes within BEAF domains. We demonstrate that BEAF has a global positive effect on transcription, is required to regulate epigenetic marks of chromatin and to restrict heterochromatin spreading, supporting its role as a regulator of chromatin compartmentalization *in vivo*.

RESULTS AND DISCUSSION

BEAF dual-cores: a signature for higher-order organization of BEAF elements

The DNA-binding activity of BEAF-32B (referred to as BEAF hereafter) has been well characterized *in vitro* and *in vivo* (Hart et al. 1997; Cuvier et al. 1998). Each BEAF subunit targets a CGATA consensus motif. In the known boundary elements (BE) of the BEAF family (*scs'*, BE-28 -51 and -76; Cuvier et al. 1998), the CGATA motif is present at least 3 times, forming clusters that are targeted by BEAF or by BEAF complexes purified from nuclear extracts. These and six additional genomic sites co-precipitated by BEAF antibodies all contain at least 3 CGATA motifs within a DNA sequence window of 100-200 bp (see "known elements" in our database). In contrast, various DNA fragments containing fewer or unclustered CGATAs were more poorly bound by BEAF *in vitro* and *in vivo* (see below). Point mutations within CGATA abolish BEAF binding and insulating activity (Zhao et al. 1995; Cuvier et al. 1998).

Based on these data, we scanned the *Drosophila melanogaster* genome for all unique 200-bp sequences that contain at least three CGATA BEAF consensus sites. This Gbrowse generated a set of 7045 predictions (Figure 1A), which includes all previously isolated BEAF binding sequences. Comparison of these predictions with the immunolocalization of BEAF on polytene chromosomes was inconclusive as predicted elements were present in all regions of the genome, including those where no staining was observed. Closer inspection showed that predicted BEAF elements were often juxtaposed (Figure 1B). This higher level of organization, with two consecutive clusters of CGATAs separated by a few hundred base pairs, had been previously reported in the *scs'* element, where BEAF binds cooperatively to each of the two clusters (Hart et al. 1997). This raised the possibility that a similarly conserved organization may be found among BEAF elements. Accordingly, we defined a

BEAF dual-core element that consists of two predicted BEAF elements separated by less than 800 bp. We scored 494 dual-core elements, with an impressive p-value of 1e-21.

We tested next the relevance of our BEAF dual-core database by analyzing their overlap with the BEAF immunostaining map of polytene chromosomes. A significant improvement was found compared to single BEAF-element predictions. Remarkably, all major BEAF signals which had been previously mapped on chromosomes 2 and 4 fell into regions harboring a dual-core element (Figure 1C-E; 2A). 4/5 dual-cores matched within the strong BEAF signals on chromosome 4 (dual-cores 2-3, 4, 5; Figure 1E) and 11/13 for chromosome 2 (Figure 2A). For chromosome 4, only one region weakly stained with anti-BEAF had no dual-core (see *, Figure 1E). Hence, BEAF dual-cores are strong predictors of BEAF binding.

As a control, we also mapped the signals obtained for DREF (Figure 1E), a factor required for transcriptional activation of genes involved in DNA replication and cell proliferation (Hirose et al. 1993) and which was shown to compete with BEAF for binding to DNA (Hart et al. 1999). DREF signals on chromosome 4 fell into two BEAF dual-cores harboring a TATCGATA DREF consensus site (Hirose et al. 1993). The overlap with the BEAF consensus indicates that DREF competition for BEAF binding may also occur there. Two additional anti-DREF signals (see *, Figure 1E) mapped to regions where numerous elements unique to DREF are present (see Gbrowser database).

These results suggest that BEAF dual-cores possess a distinct organization, which we analyzed in more detail. In the case of *scs'*, an AT-rich spacer of approximately 200 bp separates two distantly located BEAF binding sites (Udvardy et al. 1985; Zhao et al. 1995). Strikingly, plotting the distribution of 200 CGATA motifs from randomly selected dual-cores revealed that hundreds of these elements harbor an AT-rich spacer (≥70% A+T) averaging 200-300 bp in length (Figure 1F). Further analysis of known BEAF elements showed a

similar organization: a 200-bp AT-rich sequence separates BE76 from a second cluster including 2 CGATAs and repeated BE28-51 elements are organized in tandem with AT-rich spacers (see **Figure S1**). Therefore, all known BEs recognized by BEAF are highly similar to BEAF dual-cores. Together with our immunostaining results, these data show that the dual-core organization defines a clear conserved signature for BEAF elements, which may be important for the activity of BEAF.

Clustering of BEAF elements around dual-cores define BEAF chromatin domains

Lower-resolution inspection of genomic regions shows that BEAF elements tend to cluster into regions of 25-75 kb (Figure 2B; **Figure S2**). We quantified this distribution by making predictions for *BEAF chromatin domains* defined as 50-kb windows containing 7 or more single BEAF elements (p-value = 0.025). 227 BEAF chromatin domains were thus scored (52 expected by chance; Figure 2B and **Figure S2**). As a control, we found that DREF clusters are not similarly distributed within large domains.

All BEAF chromatin domains were positively stained by anti-BEAF antibodies on polytene chromosomes (Figure 2A-B; **Figure S2**, shadowed regions). Significantly, staining was absent outside BEAF domains, where only single BEAF elements are scattered (Figure 1A; **Figure S2**, light-shadowed regions). Nearly all BEAF domains contained a dual-core surrounded by additional single elements (Figure 1B). The assignment of BEAF domains significantly improves the correlation with BEAF immunostaining data over that of dual-cores alone. As the size of BEAF domains and intervening regions can be resolved on polytene chromosomes, the co-localization of BEAF domains and anti-BEAF staining is not fortuitous. Moreover, BEAF domains included the few regions lacking dual-cores which stain positive for BEAF (region marked s*, Figure 1E; region 22A6-7, Figure 2A). Finally, BEAF domains matched extremely well with BEAF signals which occasionally cover a more

extended region (e.g. 21B1-3; 22E-F; 24A1-C9, Figure 2B; **Figure S2**). These results validate our database of BEAF chromatin domains that harbor clustered BEAF elements surrounding dual-cores.

Our findings suggest that BEAF binding may not be restricted to the unique dual-core element present, but may also extend to surrounding single BEAF elements. Since the boundary activity of BEAF requires that it binds to two separate sites bracketing an insulated gene (Cuvier et al. 1998; Ishii et al. 2002), it was important to test whether surrounding single BEAF elements are also bound by BEAF *in vivo*. Chromatin immunoprecipitations (ChIP) were performed in the 9E2-3 region of BE76 (locus 9E3; Figure 2C). The dual-core present in this region is indeed bound by BEAF *in vivo* (Figure 2D). While BE76 is located 6 kbp away (Figure 2C), we confirmed by ChIP that it is also bound by BEAF *in vivo*, whereas we find that BEAF does not bind to the intervening region (Figure 2D). Thus, a strict dual-core organization (with 3 CGATAs in each cluster) is not absolutely required for BEAF binding and we found that two other single BEAF elements (*Jan*/*Sry* and *Rga*/*Atu*) resembling dual cores were also co-precipitated in anti-BEAF chromatin pellets (Figure 2D).

We next asked whether many single elements similarly resemble dual-cores. We found that a statistically significant proportion of 1700 single BEAF elements have a second cluster of 2 CGATAs separated by an intervening sequence of approximately 200-300 bp (**Figure S3**) and share a >70% AT-rich spacer. These results again demonstrate the highly conserved nature of the BEAF dual-core signature. This signature is also quite specific as the region 5' to the actin 5C promoter was not enriched in anti-BEAF ChIP pellets (Figure 2D). This region contains 6 CGATA motifs within 1 kbp (**Figure S3**); as these motifs are not clustered, no BEAF element had been scored for this region in our database.

In summary, the presence of dual-cores together with numerous additional single elements resembling dual-cores is, genome-wide, the feature most relevant to BEAF binding

within ~ 50-kbp chromatin domains. As the boundary activity of BEAF relies on its binding to two elements bracketing a gene (Ishii and Laemmli 2003), our database of multiple BEAF elements within BEAF domains shows that this function of BEAF is reflected by a genomic distribution of hundreds of binding sites.

BEAF dual-cores are enriched near promoters and for specific sub-classes of genes

The insulating function of BEAF has been evidenced using reporter constructs (Kellum and Schedl 1991) where a BE insulates a gene from position effects (Cuvier et al. 1998), from enhancers (Zhao et al. 1995; Cuvier et al. 2002; Gilbert et al. 2006) or from heterochromatin spreading (Ishii et al. 2002; Ishii and Laemmli 2003). Are these activities correlated with a non-random distribution of BEAF dual-cores with respect to genes? A search of our database revealed a statistically significant enrichment of these elements within +/- 500 bp of promoter start sites (Figure 3A): 2521 BEAF elements were found near promoters (748 expected) with a strong enrichment for 347 dual-cores juxtaposed to promoters (107 expected). Moreover, we also found that BEAF elements were strikingly enriched between divergent genes (Figure 1B): of 2652 such gene pairs with an intergenic space < 5 kb, 912 contained a BEAF element. We note that this situation is also shared by TFIIIC binding sites (Noma et al. 2006). The *scs'* element and the dual-core-like BE76, *jan*/*Sry* and *Rga*/*Atu* elements all fall within this class of promoter-associated BEAF elements. Within the context of larger BEAF domains, BEAF elements often bracket a gene, or a group of genes, indicating that BEAF could play an important role in insulating them from neighboring domains.

As a positive control, 1311 DREF elements were found near promoters (290 expected), an unsurprising result for a transcription factor. In addition, based on the previously reported consensus of *zeste* (Benson and Pirrotta 1988), we found that putative

zeste elements are not enriched near promoters (see our Gbrowser), but are found instead in gene-coding regions or gene-poor regions, opposite to what we find for BEAF and DREF.

Interestingly, BEAF elements were significantly enriched within several specific gene-class ontology (GO) groups that affect mitosis, organelle generation, chromosome organization/segregation, the cell cycle, DNA packaging, microtubule-based processes, actin organization and cell-shape (**Table S1**). These essential cellular processes require constitutive activation. In contrast, genes associated with non-constitutive processes such as sensing and behavior were not enriched for BEAF elements. BEAF might also be involved in X chromosome-specific processes as BEAF elements are enriched on the X chromosome (1677 predictions, 1227 expected, p-value ~ 1e-42) while their frequency is less than expected on autosomes. This might reflect the potential role of BEAF in dosage compensation (Gilbert et al. 2006), as also suggested for DREF sites which often overlap with BEAF elements and are enriched in male-specific lethal (MSL)-related genes (Legube et al. 2006). However, for genes harboring a DREF promoter element, the enriched GO groups showed marked differences with BEAF target genes. For instance, DREF target genes are enriched for functions involved in biosynthesis, metabolism and cell death, the latter being in agreement with the known effects of *dref* mutants on apoptosis (Hirose et al. 2001). We conclude from these observations that BEAF elements are enriched within specific gene classes, particularly those involved in the structure and/or dynamics of chromosomes and of the cytoskeleton.

BEAF as a global positive activator of transcription

Given that BEAF elements are juxtaposed to hundreds of promoters, we next asked if they might have any effect on transcription. Analysis of micro-array expression data for *Drosophila* embryos showed that many genes bracketed by BEAF elements had a high expression profile in early embryogenesis which significantly decreased at later stages

(**Figure S4**). This resembles the profile of expression of *beaf* (which contains a BEAF element in its own promoter), suggesting that BEAF could act as a global positive regulator of transcription. Strikingly, genes that share a BEAF-like expression profile are also enriched for promoter-linked BEAF elements genome-wide (Fig. 3B(i)). A correlation coefficient greater than the statistically optimal cutoff of 0.43 was found for 2674 genes; of these, 387 contain a BEAF element (p-value ~ 1e-15). As a control, we found that DREF target genes are enriched for a subpopulation that is co-expressed with DREF (Fig. 3B(ii)). Interestingly, the correlation for BEAF extends over a much broader region than for DREF, indicating that BEAF may act as a more global positive regulator of transcription. Moreover, genes with anti-correlated expression were clearly under-represented in BEAF elements (Fig. 3B(i)), further illustrating the genome-wide impact of BEAF elements on transcription.

In line with this interpretation, we show next that numerous BEAF foci (often found at the nuclear periphery) co-localize with acetylated histone H4 (H4Ac; Figure 3B), an epigenetic mark associated with active or potentially active genes (Schubeler et al. 2004; Pokholok et al. 2005). The frequent association of BEAF with active chromatin is in full agreement with our conclusion that BEAF acts as a global positive regulator of genes bracketed by clustered BEAF elements within BEAF chromatin domains. Significantly, the availability of genome-wide mapping data for H4Ac (Schubeler et al. 2004; Pokholok et al. 2005) makes it possible to determine that BEAF domains are often enriched for this epigenetic mark (Figure 3D), in contrast to sites located outside BEAF domains. Together with our gene expression analyses, these data show that BEAF elements may participate in the homeostasis linking the regulation of a specific subset of genes with the regulation of chromatin structure within BEAF domains. This involves a regulated deposition of chromatin epigenetic marks, as demonstrated next by BEAF depletion experiments.

BEAF globally hinders the spread of heterochromatin

Although gene-poor regions are generally devoid of BEAF elements, a classification of BEAF domains reveals that those with the highest density of BEAF elements are found at the heterochromatin border of all chromosomes, including 2L, 2R, 3L, 3R, X and 4 hetero-/euchromatin junctions (Figure 4A; **Figure S5A**). Accordingly, we detected strong BEAF, but not DREF, signals in these regions (Figure 4B). As BEAF can insulate a gene against heterochromatin repression (Ishii and Laemmli 2003; Gilbert et al. 2006), these sites likely define a sub-category of BEAF elements that protect against spreading of constitutive heterochromatin.

To assess this possibility, we depleted BEAF proteins from Schneider (SL2) cells by siRNA treatment (Figure 4) followed by detection of several chromatin proteins, including histone H3 methylated on lysine 9 (H3MeK9), a characteristic epigenetic mark of heterochromatin (Bannister et al. 2001). BEAF staining was reduced to background levels after five days of siRNA treatment (**Figure S5**), showing efficient depletion of this protein. BEAF depletion led to a striking increase in overall H3MeK9 staining levels (Figure 4C); the same nuclei showed only a slight decrease of the RNA polymerase II signal and no changes in the staining of unmodified histone H3 (data not shown and see below). An increase in H3MeK9 may result from the spreading of heterochromatin into euchromatin regions. Closer inspection showed that approximately 35 +/-10 more weakly stained H3MeK9 foci could be detected in control cells (3x panels, Figure 4C), in addition to the 3-5 strong foci corresponding to constitutive heterochromatin (Blattes et al. 2006). BEAF depletion did not significantly change the total number of H3MeK9 foci detected (to 40 +/-10), but instead led the formation of much broader foci (Figure 4C, 3x enlarged panels). Taken together, these results demonstrate that one important function of BEAF is to restrict heterochromatin spreading to euchromatin regions. Our data further indicate that the blocking activity of

BEAF does not solely apply to constitutive heterochromatin.

As discussed above, H4Ac marks have been implicated in the positive epigenetic regulation of chromatin structure. If BEAF serves to protect active domains against invasion by repressive marks, its depletion should have an opposite effect on modifications such as H4Ac. Indeed, the strong spreading of H3MeK9 levels induced by BEAF depletion was accompanied by marked decrease in H4Ac levels (Figure 4D). Although differences were found among siRNA-treated or control cells, H4Ac staining generally varied according to BEAF levels remaining after 3 days of siRNA treatment (not shown), while the overall chromatin staining of controls such as replication protein A (RPA) and RNA polymerase II, was largely unchanged (**Figure S5B-C**). We conclude that BEAF depletion specifically increases the nuclear ratio of heterochromatin to euchromatin. Therefore, epigenetic regulation of chromatin structure accounts for the role of BEAF as a global activator of gene expression for genes bracketed by BEAF elements in BEAF domains.

BEAF dual-cores partition the genome into BEAF domains

Our *in silico* analysis identifies the BEAF dual-core as a highly conserved signature for BEAF binding sites. A typical BEAF dual-core typically harbors two clustered sets of ≥3 CGATA sequences separated by a >200-bp AT-rich sequence and is juxtaposed to one or two promoters. While they do not appear to play a direct role in boundary activity (Hart and Laemmli 1998), these AT-rich sequences may promote interactions with other proteins (see below) or, as in the case of the *gypsy* insulator, tether chromatin loops to the nuclear matrix (Gerasimova et al. 2000).

In the case of *scs'*, BEAF binds cooperatively to the two CGATA clusters across a 200-bp spacer (Hart et al. 1997). This dual interaction creates a DNA loop which may form a specialized nucleoprotein structure, which could be stabilized through the binding of

additional unknown proteins or by a positioned nucleosome (Zhao et al. 1995). Our results provide a sound basis for understanding how BEs function and for identifying additional factors required to stabilize/regulate the activity of BEAF dual-cores. Indeed, in many dual-cores, our data identify a new AGCAA/TTCGT consensus sequence which may constitute a binding site for BEAF-32A (see Figure S1), in line with the ability of this isoform to form a hetero-complex with BEAF-32B (Hart et al. 1997).

Our genome-wide analysis of gene expression profiles highlights a global positive effect of BEAF dual-cores on the regulation of hundreds of genes, including many that regulate key functions of chromosome or cytoskeletal structure. These conclusions are consistent with the proposed role of BEAF in chromatin structure and with the finding that dominant negative BEAF mutants impair the establishment of BE activity and are lethal to flies (Gilbert et al. 2006). Moreover, we find that BEAF depletion doubles the number of cells with chromosome segregation defects (data not shown). BEAF appears to act as a positive regulator of transcription within large BEAF chromatin domain and to control the deposition of epigenetic marks important for the regulation of chromatin structure. Many BEAF foci co-localize with acetylated histone H4, which was recently shown to impair the higher-order organization of the 30-nm chromatin fiber (Shogren-Knaak et al. 2006). In addition, BEAF domains are largely enriched in H4Ac levels. Therefore, one function of BEAF dual-cores may be to participate in chromatin opening, in line with the localization of strong nucleases hypersensitive sites within many characterized BEs, including the two BEAF-binding sites in *scs'* and BE76 (Udvardy et al. 1985; Cuvier et al. 1998). Dual-cores may also limit the spreading of heterochromatin H3MeK9 epigenetic marks into active regions, as H3MeK9 is not restricted to constitutive heterochromatin but to numerous additional sites. As shown for pericentric heterochromatin (Taddei et al. 2001), the interplay between H4Ac and H3MeK9 epigenetic marks may also apply to euchromatic sites,

particularly those where Heterochromatin Protein 1 (HP1) binds (de Wit et al. 2005).

In pericentric heterochromatin regions, we propose that the high density of BEAF dual-cores is necessary to overcome high concentrations of repressing factors. These may include HP1 and the satellite-associated D1 protein (Aulner et al. 2002). D1 has recently been implicated in heterochromatin dynamics through a competitive interactions with HP1 for binding to satellite repeats in X-chromosome heterochromatin (Blattes et al. 2006). Intriguingly, the AT-rich domain of the BE28 dual-core element located nearby is bound by D1 *in vivo* (Cuvier et al. 2002) and, through its physical interaction with BEAF, may participate in the determination of heterochromatin and euchromatin domain boundaries.

The CTCF boundary factor mediates long-range communication between distant loci (Ling et al. 2006). Several lines of evidence suggest that BEAF acts similarly. One of the BEAF elements identified here has been implicated in mediating *trans*-repression by bringing a gene close to the satellite sequence array inserted in the $brown^{D}$ gene (Sage et al. 2005) and the distant *scs* and *scs'* BEs have been shown to interact (Blanton et al. 2003). The number of BEAF foci detected per nucleus is 3-4 times lower than its number of distinct binding sites on polytene chromosomes (Zhao et al. 1995; Hart et al. 1997), suggesting that they cluster in diploid cells. Most importantly, the boundary activity of BEAF involves communication between two elements bracketing an insulated gene and its binding to these two sites (Ishii and Laemmli 2003). A number of dual-cores identified here also participate in long-range communication (Martin Muller, personal communication; O.C., unpublished results), further supporting a role for BEAF in establishing cross-talk between distant sites. We propose that local disruption of the 30-nm fiber, resulting from BEAF binding to two separate dual-cores clusters to target epigenetic marks such as H4Ac and to generate hypersensitive sites, favors the invasion of, or interaction with, additional chromatin sites, including other BEs.

MATERIALS AND METHODS

Complete bioinformatic analysis results are available for download at: http://www.sfu.ca/~eemberly/insulator/. Additional details are available in the Supplemental data section. siRNA-mediated BEAF depletion, immunostaining, ChIP procedures and polytene chromosome staining are described in Materials and Methods of Supplemental data.

ACKNOWLEDGMENTS:

EE would like to acknowledge the support of NSERC and CIHR in carrying out this work. OC would like to acknowledge the support of Human Frontier Science Program and Fondation pour la Recherche Médicale. Work in EK's lab was supported by ARC, ACI Cancéropôle GSO and ANR grants. RB is supported by a French Research Ministry doctoral fellowship.

REFERENCES:

Aulner, N., C. Monod, G. Mandicourt, D. Jullien, O. Cuvier, A. Sall, S. Janssen, U.K. Laemmli, and E. Kas. 2002. The AT-hook protein D1 is essential for Drosophila melanogaster development and is implicated in position-effect variegation. *Mol Cell Biol* **22**: 1218-32.

Bannister, A.J., R. Schneider, and T. Kouzarides. 2002. Histone methylation: dynamic or static? *Cell* **109**: 801-6.

Bannister, A.J., P. Zegerman, J.F. Partridge, E.A. Miska, J.O. Thomas, R.C. Allshire, and T. Kouzarides. 2001. Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. *Nature* **410**: 120-4.

Bell, A.C. and G. Felsenfeld. 1999. Stopped at the border: boundaries and insulators. *Curr Opin Genet Dev* **9**: 191-8.

Bell, A.C., A.G. West, and G. Felsenfeld. 1999. The protein CTCF is required for the enhancer blocking activity of vertebrate insulators. *Cell* **98**: 387-96.

Benson, M. and V. Pirrotta. 1988. The Drosophila zeste protein binds cooperatively to sites in many gene regulatory regions: implications for transvection and gene regulation. *Embo J* **7**: 3907-15.

Blanton, J., M. Gaszner, and P. Schedl. 2003. Protein:protein interactions and the pairing of boundary elements in vivo. *Genes Dev* **17**: 664-75.

Blattes, R., C. Monod, G. Susbielle, O. Cuvier, J.H. Wu, T.S. Hsieh, U.K. Laemmli, and E. Kas. 2006. Displacement of D1, HP1 and topoisomerase II from satellite heterochromatin by a specific polyamide. *Embo J* **25**: 2397-408.

Capelson, M. and V.G. Corces. 2004. Boundary elements and nuclear organization. *Biol Cell* **96**: 617-29.

Corces, V.G. 1995. Chromatin insulators. Keeping enhancers under control. *Nature* **376**: 462-3.

Cuvier, O., C.M. Hart, E. Kas, and U.K. Laemmli. 2002. Identification of a multicopy chromatin boundary element at the borders of silenced chromosomal domains. *Chromosoma* **110**: 519-31.

Cuvier, O., C.M. Hart, and U.K. Laemmli. 1998. Identification of a class of chromatin boundary elements. *Mol Cell Biol* **18**: 7478-86.

de Wit, E., F. Greil, and B. van Steensel. 2005. Genome-wide HP1 binding in Drosophila: developmental plasticity and genomic targeting signals. *Genome Res* **15**: 1265-73.

Gerasimova, T.I., K. Byrd, and V.G. Corces. 2000. A chromatin insulator determines the nuclear localization of DNA. *Mol Cell* **6**: 1025-35.

Gilbert, M.K., Y.Y. Tan, and C.M. Hart. 2006. The Drosophila Boundary Element-Associated Factors BEAF-32A and 32B Affect Chromatin Structure. *Genetics*.

Hart, C.M., O. Cuvier, and U.K. Laemmli. 1999. Evidence for an antagonistic relationship between the boundary element-associated factor BEAF and the transcription factor DREF. *Chromosoma* **108**: 375-83.

Hart, C.M. and U.K. Laemmli. 1998. Facilitation of chromatin dynamics by SARs. *Curr. Opin. Genet. Dev.* **8**: 519-525.

Hart, C.M., K. Zhao, and U.K. Laemmli. 1997. The scs' boundary element: characterization of boundary element-associated factors. *Mol Cell Biol* **17**: 999-1009.

Hirose, F., N. Ohshima, M. Shiraki, Y.H. Inoue, O. Taguchi, Y. Nishi, A. Matsukage, and M. Yamaguchi. 2001. Ectopic expression of DREF induces DNA synthesis, apoptosis, and unusual morphogenesis in the Drosophila eye imaginal disc: possible interaction with Polycomb and trithorax group proteins. *Mol Cell Biol* **21**: 7231-42.

Hirose, F., M. Yamaguchi, H. Handa, Y. Inomata, and A. Matsukage. 1993. Novel 8-base

pair sequence (Drosophila DNA replication-related element) and specific binding factor involved in the expression of Drosophila genes for DNA polymerase alpha and proliferating cell nuclear antigen. *J Biol Chem* **268**: 2092-9.

Ishii, K., G. Arib, C. Lin, G. Van Houwe, and U.K. Laemmli. 2002. Chromatin boundaries in budding yeast: the nuclear pore connection. *Cell* **109**: 551-62.

Ishii, K. and U.K. Laemmli. 2003. Structural and dynamic functions establish chromatin domains. *Mol Cell* **11**: 237-48.

Jenuwein, T. and C.D. Allis. 2001. Translating the histone code. *Science* **293**: 1074-80.

Kellum, R. and P. Schedl. 1991. A position-effect assay for boundaries of higher order chromosomal domains. *Cell* **64**: 941-50.

Legube, G., S.K. McWeeney, M.J. Lercher, and A. Akhtar. 2006. X-chromosome-wide profiling of MSL-1 distribution and dosage compensation in Drosophila. *Genes Dev* **20**: 871-83.

Ling, J.Q., T. Li, J.F. Hu, T.H. Vu, H.L. Chen, X.W. Qiu, A.M. Cherry, and A.R. Hoffman. 2006. CTCF mediates interchromosomal colocalization between Igf2/H19 and Wsb1/Nf1. *Science* **312**: 269-72.

Noma, K., H.P. Cam, R.J. Maraia, and S.I. Grewal. 2006. A role for TFIIIC transcription factor complex in genome organization. *Cell* **125**: 859-72.

Parnell, T.J., E.J. Kuhn, B.L. Gilmore, C. Helou, M.S. Wold, and P.K. Geyer. 2006. Identification of Genomic Sites That Bind the Drosophila Suppressor of Hairy-wing Insulator Protein. *Mol Cell Biol* **26**: 5983-93.

Pokholok, D.K., C.T. Harbison, S. Levine, M. Cole, N.M. Hannett, T.I. Lee, G.W. Bell, K. Walker, P.A. Rolfe, E. Herbolsheimer, J. Zeitlinger, F. Lewitter, D.K. Gifford, and R.A. Young. 2005. Genome-wide map of nucleosome acetylation and methylation in yeast. *Cell* **122**: 517-27.

Ramos, E., D. Ghosh, E. Baxter, and V.G. Corces. 2006. Genomic organization of gypsy chromatin insulators in Drosophila melanogaster. *Genetics* **172**: 2337-49.

Sage, B.T., J.L. Jones, A.L. Holmes, M.D. Wu, and A.K. Csink. 2005. Sequence elements in cis influence heterochromatic silencing in trans. *Mol Cell Biol* **25**: 377-88.

Schubeler, D., D.M. MacAlpine, D. Scalzo, C. Wirbelauer, C. Kooperberg, F. van Leeuwen, D.E. Gottschling, L.P. O'Neill, B.M. Turner, J. Delrow, S.P. Bell, and M. Groudine. 2004. The histone modification pattern of active genes revealed through genome-wide chromatin analysis of a higher eukaryote. *Genes Dev* **18**: 1263-71.

Shogren-Knaak, M., H. Ishii, J.M. Sun, M.J. Pazin, J.R. Davie, and C.L. Peterson. 2006. Histone H4-K16 acetylation controls chromatin structure and protein interactions. *Science* **311**: 844-7.

Taddei, A., C. Maison, D. Roche, and G. Almouzni. 2001. Reversible disruption of pericentric heterochromatin and centromere function by inhibiting deacetylases. *Nat Cell Biol* **3**: 114-20.

Udvardy, A., E. Maine, and P. Schedl. 1985. The 87A7 chromomere. Identification of novel chromatin structures flanking the heat shock locus that may define the boundaries of higher order domains. *J Mol Biol* **185**: 341-58.

West, A.G., M. Gaszner, and G. Felsenfeld. 2002. Insulators: many functions, many mechanisms. *Genes Dev* **16**: 271-88.

Zhao, K., C.M. Hart, and U.K. Laemmli. 1995. Visualization of chromosomal domains with boundary element-associated factor BEAF-32. *Cell* **81**: 879-89.

FIGURE LEGENDS

Figure 1. *BEAF dual-cores* define a highly conserved organization of BEAF elements

(A) Examples of BEAF element predictions downloaded from our BEAF Gbrowse. Each element contains ≥ 3 BEAF CGATA/TATCG consensus sites within a 200-bp sequence. "D" indicate BEAF elements containing a consensus for DREF (TATCGATA). Bar, 5 kbp.

(B) Example of two BEAF elements organized into a *dual-core* element that resides between divergent promoters. Bar, 5kbp.

(C) A chromosome 4 download showing predictions for genes (red), cytology (blue), BEAF dual-cores (blue boxes) and single BEAF elements (shadowed if clustered).

(D) Mapping of BEAF and DREF signals on polytene chromosome 4 (see E). Solid or dotted-lines indicate strong or weak signals, respectively.

(E) Immunostaining with anti-BEAF and anti-DREF antibodies. Dual-cores are indicated (1-5). Arrow, strong D1 band of chromosome 4. S* indicates a region stained with anti-BEAF outside a dual-core region. c, chromocenter.

(F) The BEAF dual-core signature. The graph shows the distances between 200 CGATAs randomly selected out of our dual-core database. Note the high preference for an intervening sequence of 200-300 bp between the two CGATA clusters (middle gap). AT-richness is >70% A+T. See our dual-core database for details.

Figure 2. **Highly clustered single BEAF elements surrounding dual-cores define ~50kbp *BEAF domains***

(A) Alignment of all chromosome 2 dual-cores (shadowed boxes) with anti-BEAF staining. Dual-cores 3-13 fit within strongly stained regions and dual-cores 1-2 within more weakly stained sites. Only one region (*22A6-7) clearly stained with anti-BEAF does not harbor a

dual-core.

(B) Example of BEAF domains: many (>7) single elements are found within 25-kbp regions surrounding BEAF dual-cores. All BEAF domains are stained with anti-BEAF while all intervening regions (25-400 kbp) are not (see **Figure S2** for a full map of all domains).

(C) Details of a BEAF domain in the BE76 region. BE76 is located 6 kbp from a dual-core element bracketing the *ras* gene. Bars 1-2-3 indicate the probes used for ChIP analyses.

(D) ChIP analyses: the top three panels show results for the BE76 region. BEAF single elements resembling dual-cores (BE76 and *Jan*/*sry* and *Atu*/*rga* regions) were analyzed using the probes indicated by bars. The actin 5C 5' region was further analyzed as a negative control (see text).

Figure 3. BEAF elements positively regulate hundreds of genes

(A) Graphical representation of the distance between BEAF elements and the nearest promoter. Other consensus sequences including DREF elements were used as controls (see our database for further details). Dual-cores are also highly enriched in the category of elements very close to two promoters, most prominently between divergent genes (see classification of dual-cores according to the sum distances to two promoters in our database).

(B) Distributions of correlation coefficients between the expression profile of genes with (red) or without (black) BEAF (i) or DREF (ii) elements. Correlation coefficients were calculated by comparison to the expression profile of the *beaf* (i) or *dref* (ii) genes. "+" signs above the graphs indicate statistical enrichment for co-regulated genes (correlation coefficient >0.43 for BEAF, and 0.80 for DREF (green bars). "-" signs denote anti-correlated gene profiles significantly enriched for genes lacking nearby BEAF sites.

(C) Double immunostaining using anti-BEAF (green) and anti-acetylated histone H4 (H4Ac; red) antibodies in Schneider (SL2) cells. The arrowheads mark sites where strong BEAF foci

are observed at the nuclear periphery and co-localize with H4Ac. The white arrow points to a DAPI-bright DNA region where reproducibly reduced staining of both BEAF and acetylated H4 is observed.

(D) BEAF domains are enriched in acetylated H4. In the example shown, two putative BEAF domains are present (shadowed regions). Position of genes are indicated together with their H4Ac levels measured by ChIP (Schubeler et al. 2004; Pokholok et al. 2005).

Figure 4. BEAF restricts the spread of heterochromatin

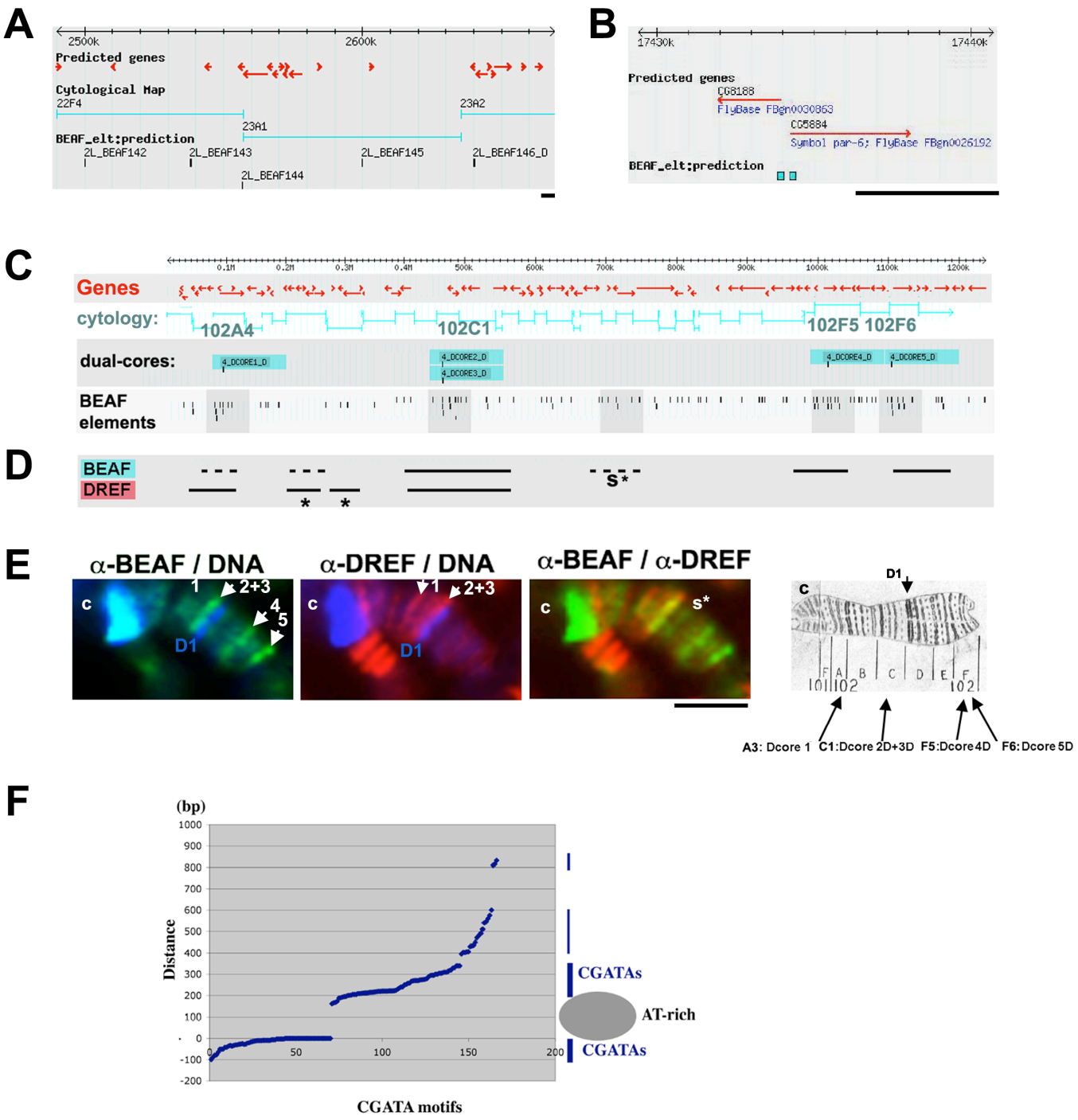
(A) Graphical representation showing that the most highly enriched BEAF domains in dual-cores are located at the heterochromatin borders of all *Drosophila* chromosomes. CEN, centromere.

(B) BEAF immunostaining surrounding the chromocenter (c).

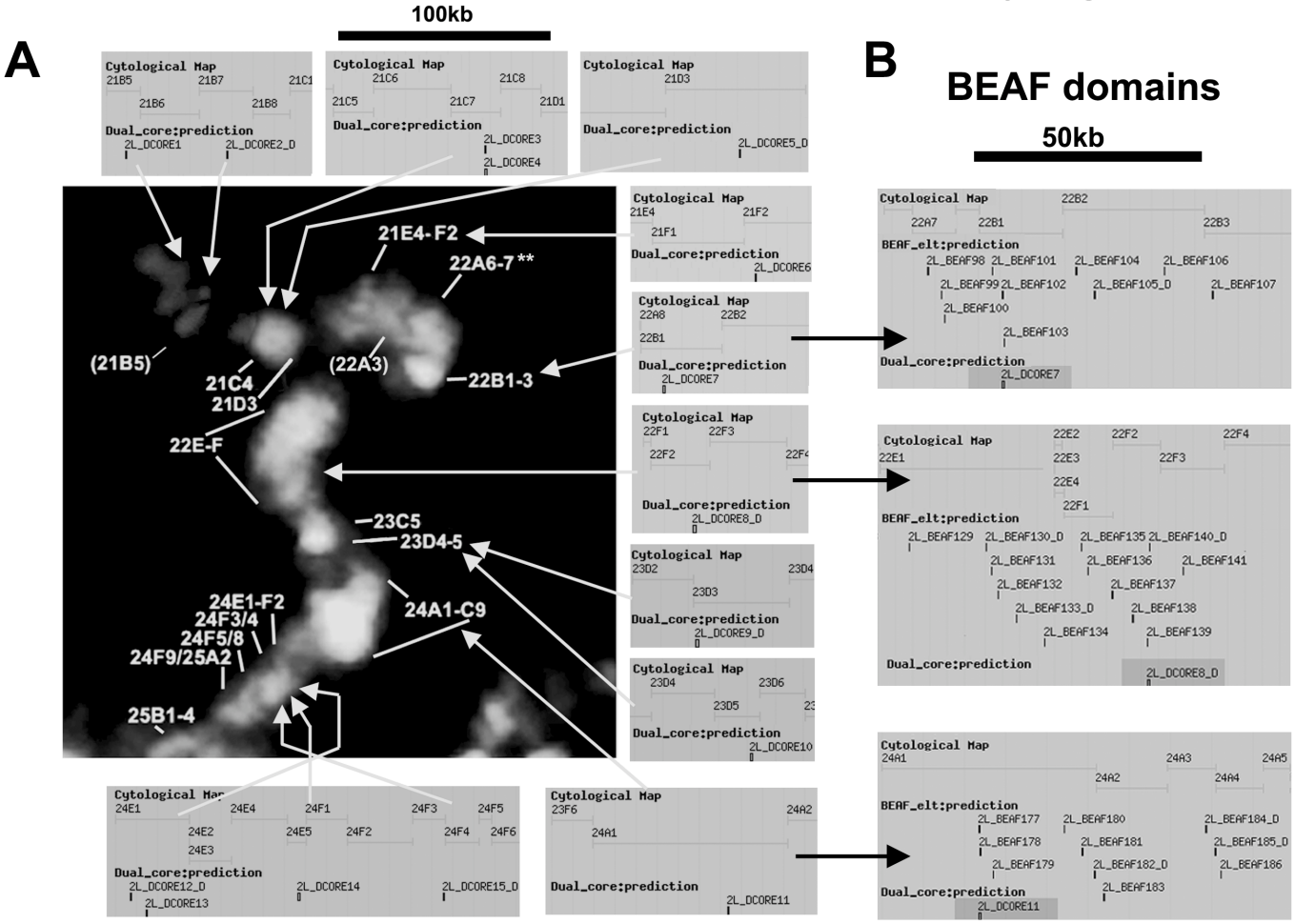
(C) H3MeK9 immunostaining after siRNA treatment of Schneider cells against *beaf* (left panels) or of control cells (right panels). 3x enlargements of confocal images are shown below.

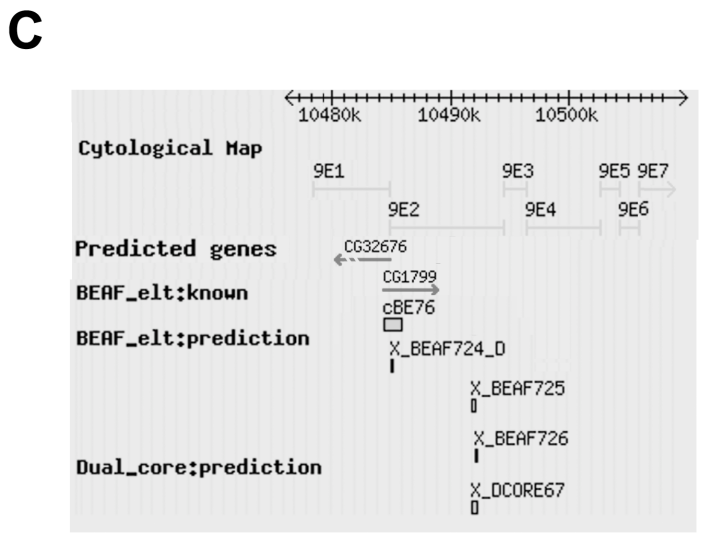
(D) Same as in (C) except nuclei were stained with anti-BEAF (green) and H4Ac (red). Merged images are also shown. See **Figure S6** for further details.

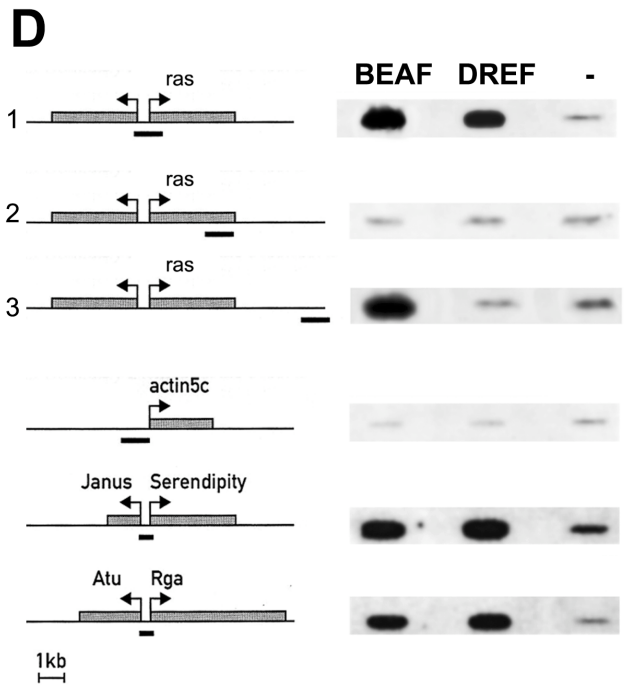
Emberly, Figure 1



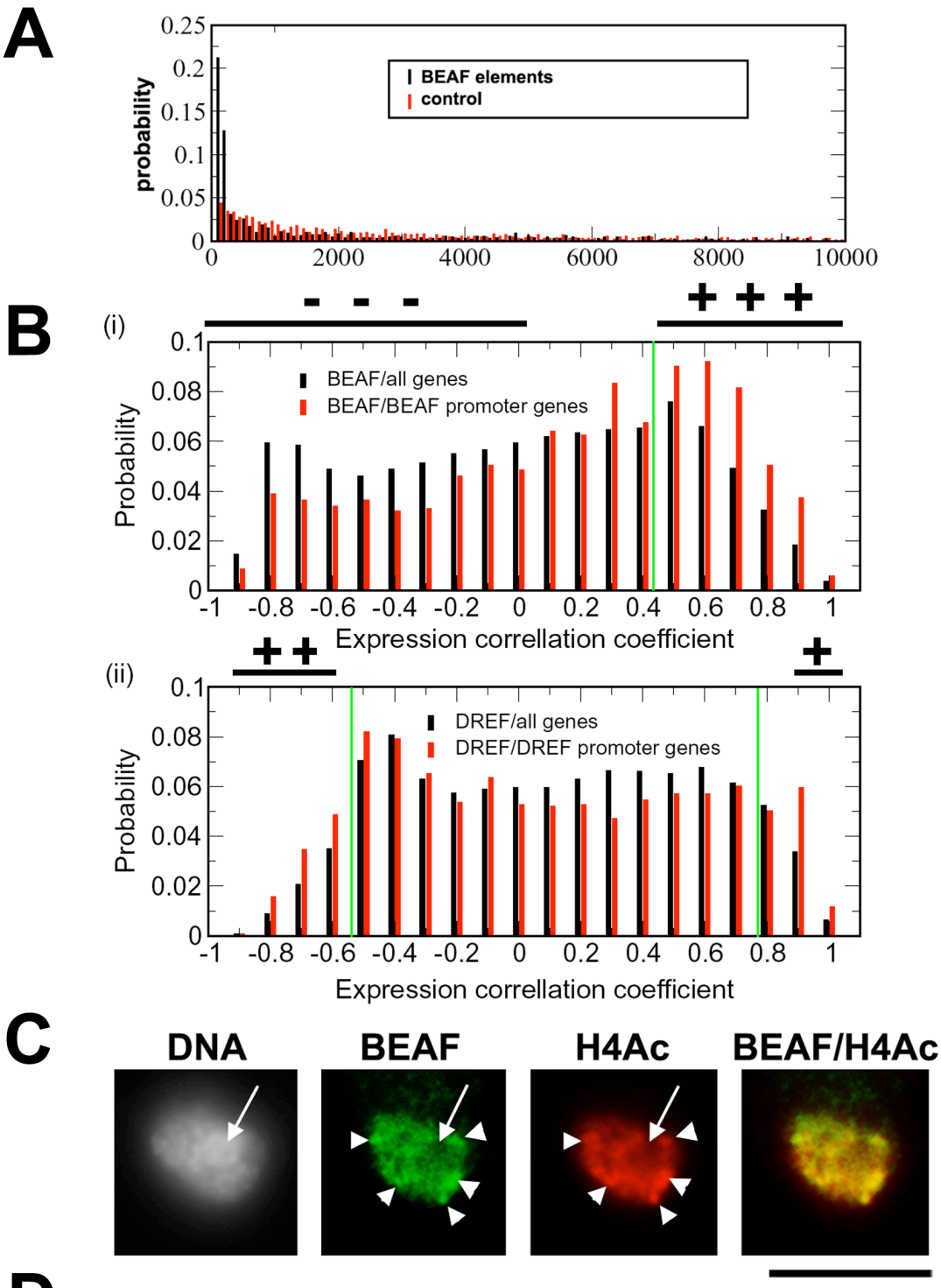
Emberly, Figure 2

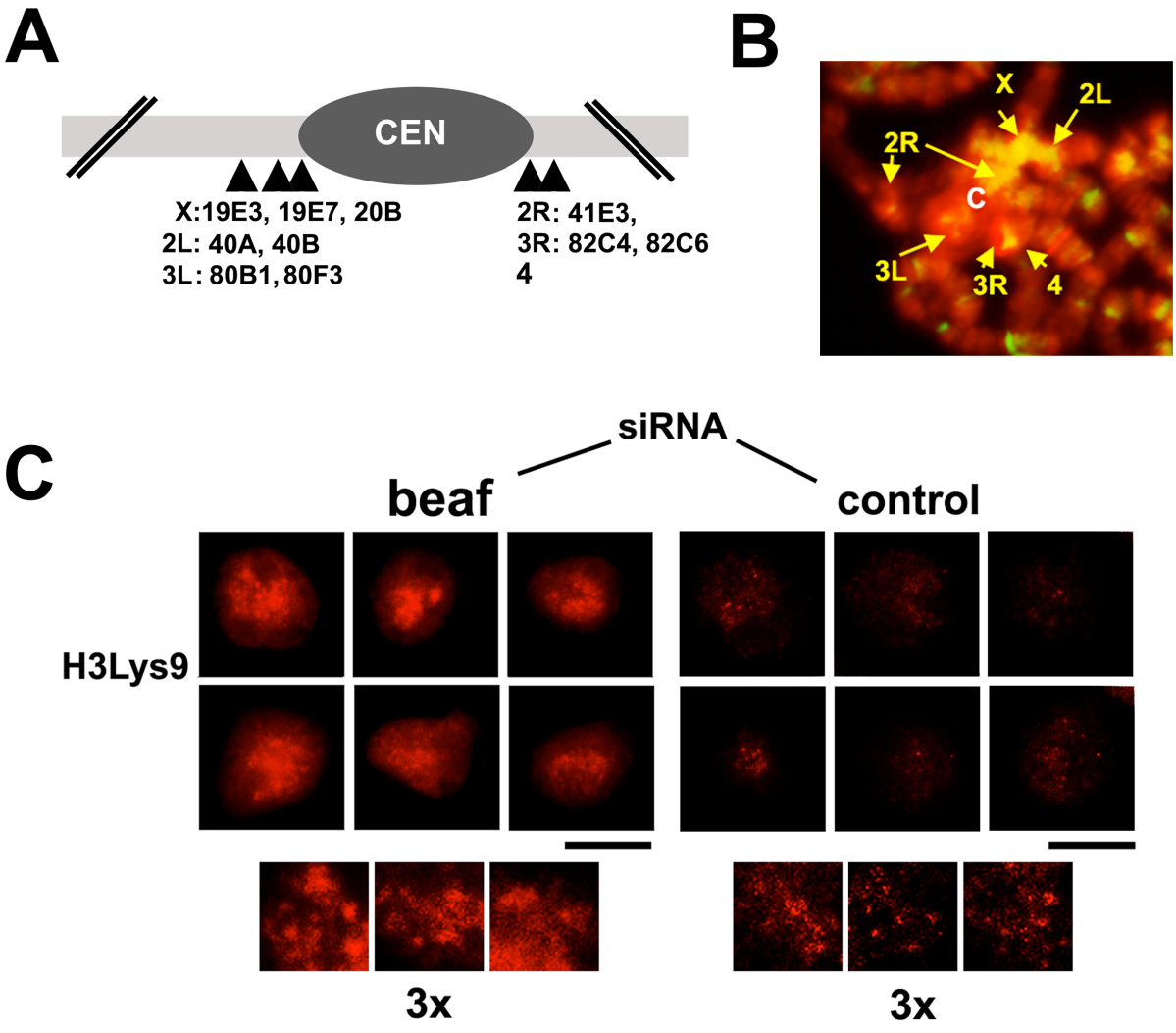






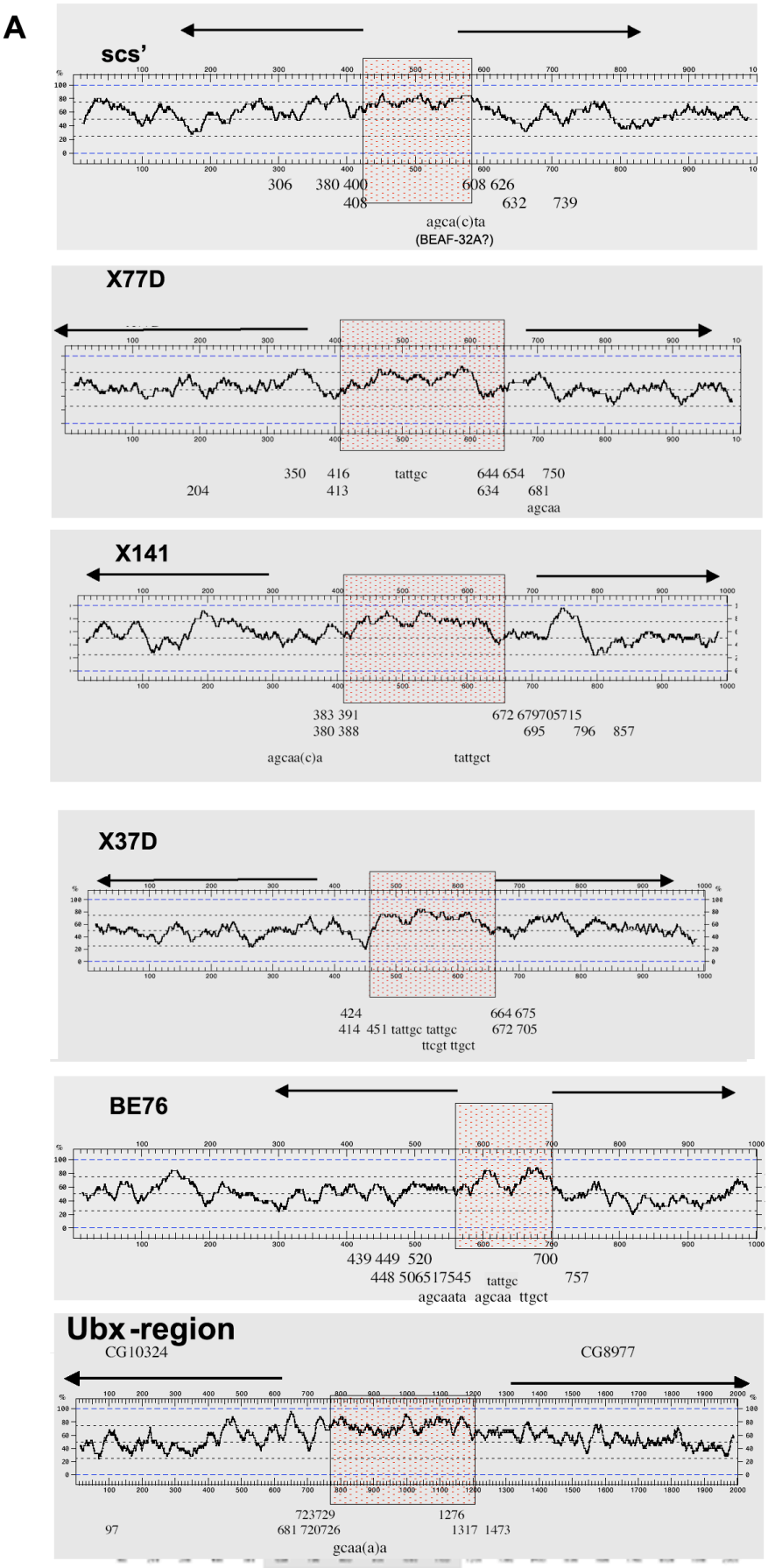
Emberly, Figure 3



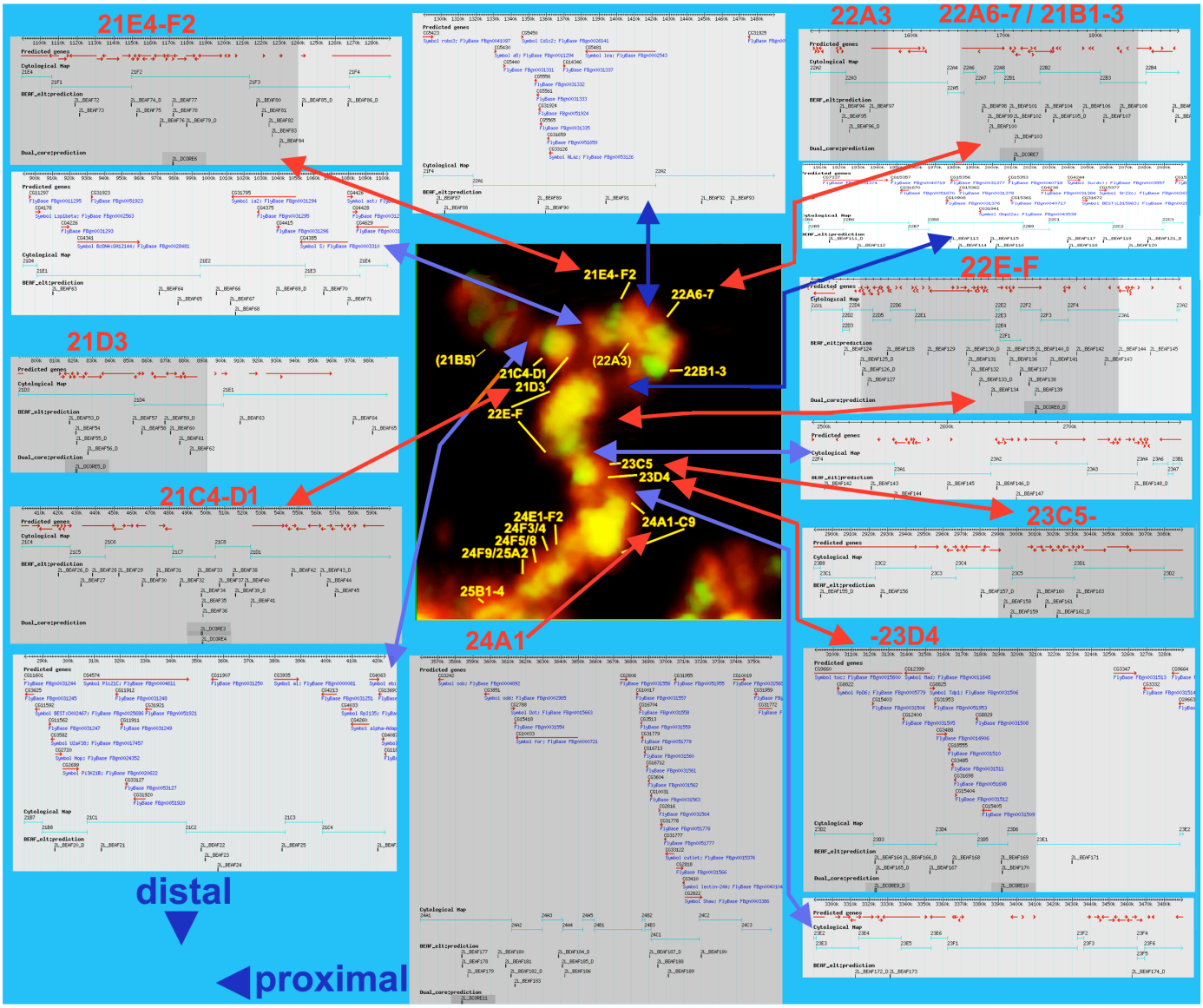

A
CEN
X:19E3, 19E7, 20B
2L: 40A, 40B
3L: 80B1, 80F3
2R: 41E3,
3R: 82C4, 82C6
4
B
X
2L
2R
C
3L
3R
4
C
siRNA
beaf
control
H3Lys9
3x
3x
D
siRNA
beaf
control
DNA
BEAF
DNA
BEAF
H4Ac
BEAF/H4ac
H4Ac
BEAF/H4ac

A

Emberly, Figure S1

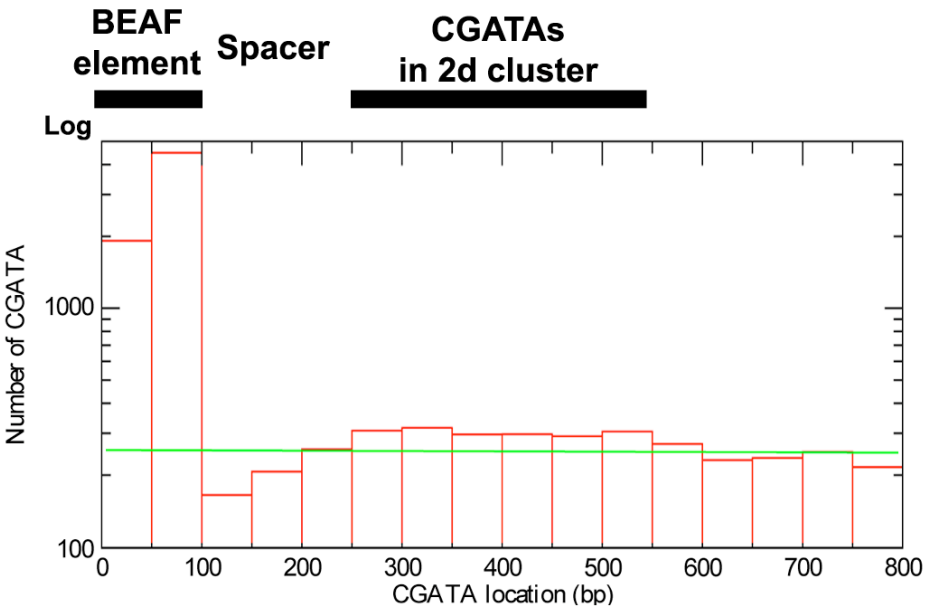


Emberly, Figure S2



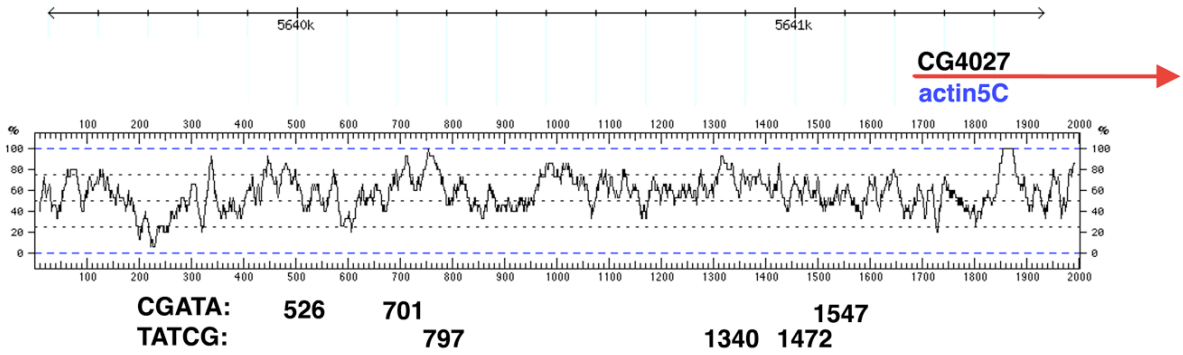
Emberly, Figure S3

A



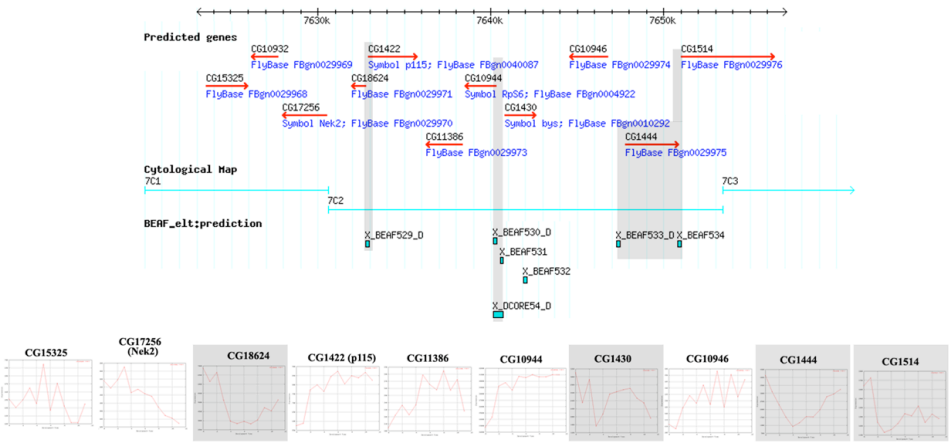
/1700 BEAF single elements

B

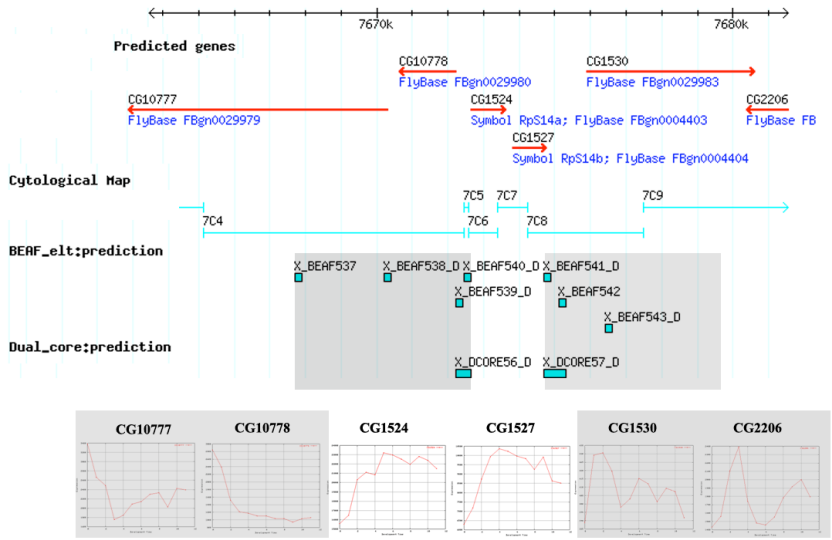


Emberly, Figure S4

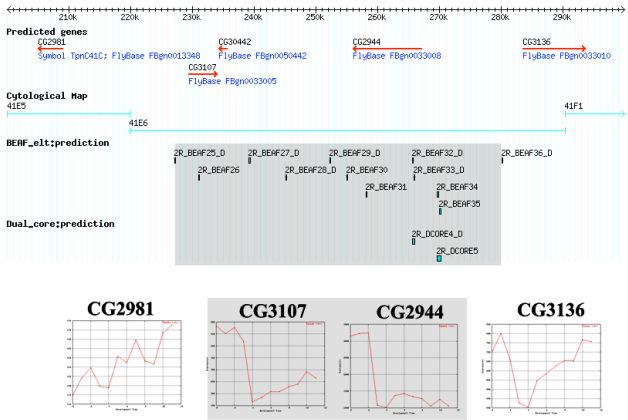
7C2



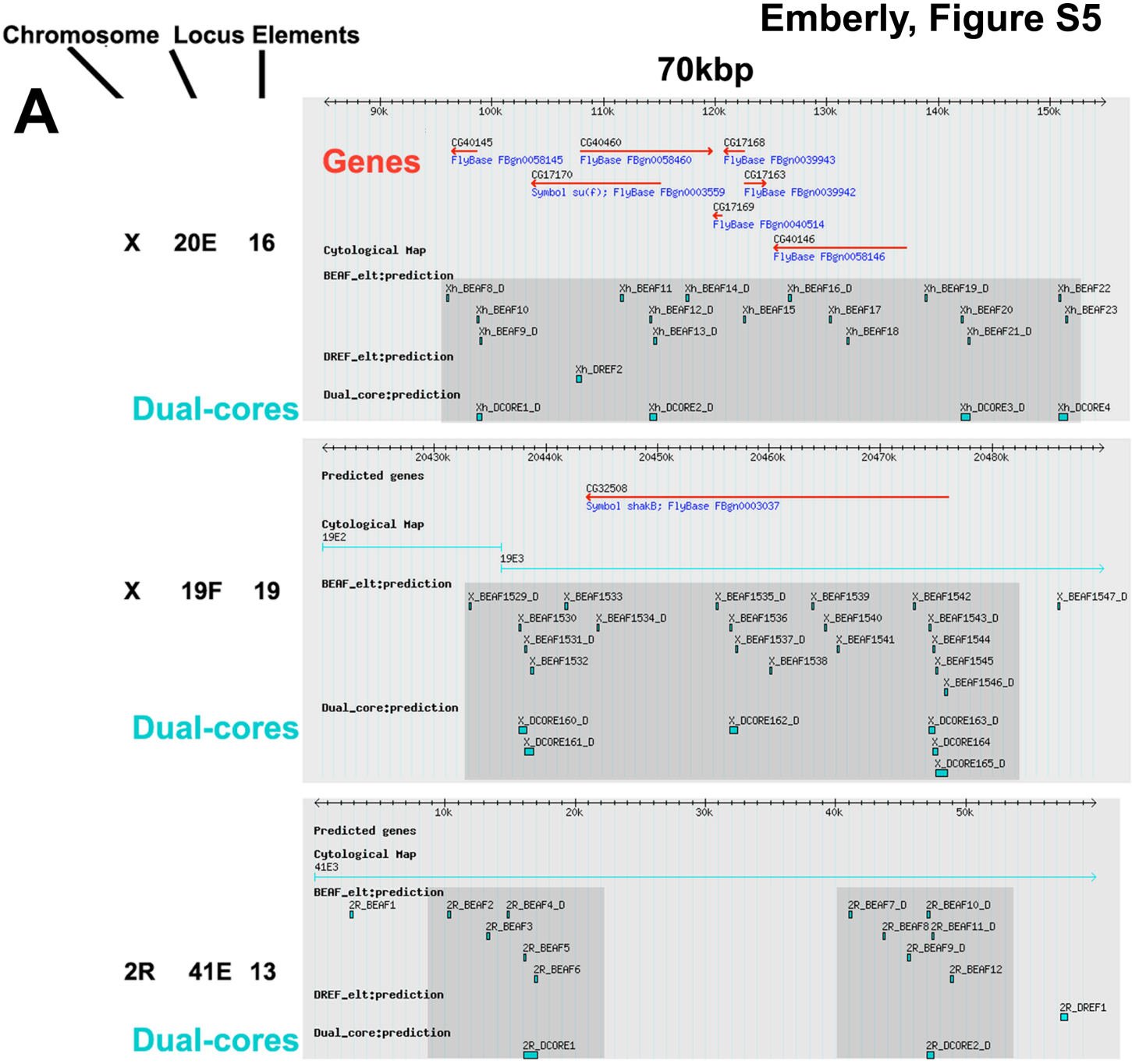
7C4

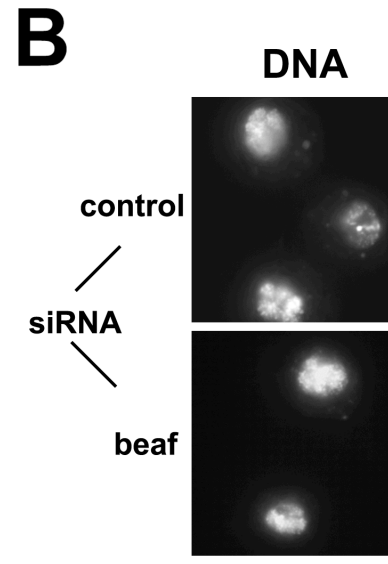


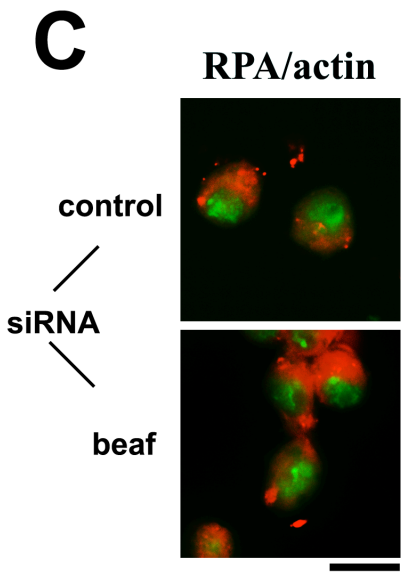
41E



Emberly, Figure S5







MOLECULAR BASIS OF POSITION EFFECT VARIEGATION IN DROSOPHILA MELANOGASTER

ABSTRACT :

Satellite III sequences (SATIII) represent the bulk of drosophila X chromosome heterochromatin and play a key role in position effect variegation (PEV) of the *white* gene in the *white-mottled* strain. Their structural characteristics, based on their biased dA•dT rich nucleotidic composition, make them good targets for artificial DNA minor groove binders. Such new tools allowed us to improve the SATIII sequences' function in heterochromatic structure formation. Our work allows us to shed light on the functional role of the D1 protein, which specifically binds SAT III sequences, as well as topoisomerase II, in the initiation of the heterochromatin assembly. Our results suggest that SATIII sequences represent a reservoir for proteins regulating ribosomal DNA arrays (rDNA) transcription located next to SATIII sequences. We also show involvement of a boundary element in heterochromatin partitioning, which plays a role in regulation of rDNA expression.

AUTEUR : Roxane BLATTES

TITRE : Bases moléculaires de la variégation d'effet de position chez *Drosophila melanogaster*.

DIRECTEUR DE THESE : Emmanuel KÄS

LIEU ET DATE DE SOUTENANCE : salle de conférence de l'IEFG, bâtiment IBCG, 118 route de Narbonne, 31062 Toulouse cedex 09 ; soutenue le vendredi 06 octobre 2006 à 14h30.

RESUME :

Les séquences satellites III (SATIII) représentent la majeure partie de l'hétérochromatine du chromosome X de drosophile et jouent un rôle dans la variégation d'effet de position du gène *white* dans l'inversion *white-mottled*. Leurs caractéristiques structurales dues à leur richesse en dA•dT sont ciblées par des ligands synthétiques du petit sillon de l'ADN, outils qui ont permis de caractériser la fonction des SATIII dans l'assemblage de l'hétérochromatine. Nous avons mis en évidence un rôle de D1, qui reconnaît spécifiquement les SATIII, mais aussi de la topoïsomérase II, dans l'initiation de l'assemblage de l'hétérochromatine. Nos résultats suggèrent que les SATIII pourraient constituer un réservoir de protéines servant à la régulation de la transcription du locus des gènes ribosomiques (rDNA) adjacent aux SATIII. Enfin, nous avons pu identifier un élément frontière qui contrôle le cloisonnement de l'hétérochromatine et jouerait un rôle dans la régulation de l'expression du rDNA.

MOTS-CLES : hétérochromatine, séquences satellites, ligands du sillon mineur, D1, topoïsomérase II, variégation d'effet de position

DISCIPLINE ADMINISTRATIVE : BIOLOGIE MOLECULAIRE ET CELLULAIRE

ADRESSE DU LABORATOIRE :

LABORATOIRE DE BIOLOGIE MOLECULAIRE EUCARYOTE
Université Paul SABATIER, bâtiment IBCG
CNRS UMR 5099
31062 Toulouse cedex 09