



HAL
open science

Ciblage de l'ADN par de molécules antitumorales et modulation de l'activité des partenaires protéiques

Paul Peixoto

► **To cite this version:**

Paul Peixoto. Ciblage de l'ADN par de molécules antitumorales et modulation de l'activité des partenaires protéiques. Biochimie [q-bio.BM]. Université du Droit et de la Santé - Lille II, 2008. Français. NNT: . tel-00322954

HAL Id: tel-00322954

<https://theses.hal.science/tel-00322954>

Submitted on 19 Sep 2008

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.

Université du Droit et de la Santé – Lille II
Ecole Doctorale Biologie-Santé

N° d'ordre :

THESE

Présentée pour l'obtention du grade de :

DOCTEUR DE L'UNIVERSITE DE LILLE II
Discipline : Biologie moléculaire et cellulaire

Par

Paul PEIXOTO

CIBLAGE DE L'ADN PAR DES MOLECULES
ANTITUMORALES ET MODULATION DE L'ACTIVITE DES
PARTENAIRES PROTEIQUES

Soutenue le 18 juillet 2008 devant le jury :

Rapporteurs :	Dr Emmanuel KAS Dr Yves COLLETTE
Examineurs :	Dr Marie Hélène DAVID Pr Pierre FORMSTECHER Dr Sylvain ROUTIER
Membre invité :	Dr Christian BAILLY
Directeur de thèse :	Dr Amélie LANSIAUX

A Florence,

A mes parents,

A mon frère.

Je tiens à exprimer mes plus vifs remerciements aux docteurs Emmanuel Käs et Yves Collette d'avoir accepté d'être rapporteurs de cette thèse. Je remercie également les docteurs Marie Hélène David, Sylvain Routier, Christian Bailly ainsi que le professeur Pierre Formstecher, de me faire l'honneur de siéger dans ce jury. J'exprime également toute ma reconnaissance à mon tuteur de thèse le Dr Amélie Lansiaux.

Ce travail de thèse a été réalisé dans un premier temps au sein de l'unité INSERM 524 puis au sein de l'équipe 4 du centre Jean Pierre Aubert INSERM unité 837 dirigé respectivement par le Dr Jean Pierre Kerckaert et le Pr Pierre Formstecher. Je les remercie de m'avoir accueilli au sein de leur unité et de m'avoir permis de développer ce projet.

Je remercie très chaleureusement le Dr Amélie Lansiaux pour m'avoir accordé sa confiance et de m'avoir dirigé tout particulièrement sur mes travaux de distribution cellulaire. Je lui adresse toute ma gratitude pour m'avoir toujours aidé quand j'en ai eu besoin et notamment pendant la rédaction de cette thèse. Amélie je vous remercie pour votre gentillesse et votre écoute et vos nombreux conseils prodigués durant ces 5 années passées au laboratoire.

Je tiens à adresser mes plus sincères remerciements au Dr Marie Hélène David pour m'avoir guidé durant ce travail de thèse. Son dynamisme, ses idées et ses nombreux conseils techniques m'ont permis de « grandir » durant ces années passées au laboratoire. C'est aussi l'occasion pour moi de la remercier pour m'avoir permis de participer à de nombreux congrès. *Mary* un grand merci pour ta grande disponibilité et ta gentillesse.

Merci donc à Amélie et à Marie pour tout le travail accompli.

J'adresse également toute ma gratitude au Dr Christian Bailly pour m'avoir permis de travailler dans cette formidable équipe et pour ses précieux conseils.

J'adresse ma plus grande gratitude à Marie-Paule Hildebrand pour l'aide qu'elle m'a apporté durant ces travaux. « *Le grand machin* » vous remercie pour votre bonne humeur, votre amitié, votre sincérité et votre soutien.

Je tiens aussi à associer à ce travail Sabine Depauw pour son dévouement et son aide dans l'étude des produits DB. Merci Sabine de m'avoir supporté ! Je te remercie également de m'avoir donné les bases essentielles de la culture Ch'ti et surtout prend bien soin de la mimine.

Je remercie également Magali Rebucci pour son amitié, son écoute et ses encouragements tout au long de cette rédaction. Un grand merci à Gâelle Lenglet (toujours prête à vous aider) pour sa grande générosité.

Un grand merci aux drôles de dames Brigitte Baldeyrou, Christine Bal et Nicole Lemahieu pour l'aide qu'elles m'ont apporté à un moment ou un autre durant ce travail de thèse. Je n'oublie pas non plus Nathalie Dias et Christelle Tardy qui m'ont elles aussi aidé à mon arrivée au laboratoire.

J'aimerais également dire merci à mes quatre amis *masculins* Jérôme, Romain, Sébastien et William. Jéjé je te remettrai mon trophée si durement gagné, merci pour tous les conseils que tu continues de m'apporter ; Main-Main merci de ne jamais avoir craqué (et encore désolé pour ton oeil) ; Deb, merci pour ta grande « tinglitude » ; et Will je te remercie de m'avoir formé à des techniques variées comme le foot printing (à l'ancienne, avec le scotch), les marquages radioactifs, la bière Belge, le Whisky écossais et le poker...

Merci à toutes les personnes qui ont contribué de près ou de loin à la rédaction de ce mémoire.

J'adresse mes remerciements les plus sincères ainsi que ma plus profonde gratitude au Pr Jean Krembel, directeur de l'Institut de Recherches sur le Cancer de Lille (IRCL), à tous les membres du Conseil d'Administration de l'IRCL et au conseil régional Nord/Pas de Calais pour le financement de mes trois premières années de thèse. Je remercie également l'Association pour la Recherche sur le Cancer de m'avoir accordé un financement pour ces 9 mois supplémentaires.

Je tiens à adresser ma gratitude la plus sincère au Dr Maud Collyn, directeur administratif de l'IMPRT IFR 113 pour la mise à disposition d'équipements de haute qualité et pour sa bonne humeur. Je tiens à associer à ces remerciements Harold Fauvel et Nathalie Jouy pour leur aide.

Je remercie également les Pr David Wilson et David Boykin pour cette excellente collaboration sans laquelle cette étude n'aurait pas été possible.

J'associe aussi à ce travail Christèle Dassi, Yoann Laitem et tous les stagiaires de plus courte durée qui ont contribué à ce travail. Merci à Elodie Monceau et Raja Nhili pour leur travaux en cours sur les DB. Et bonne chance à Raja, qui poursuivra l'étude, alors prends bien soin des DB, je te les confie.

Merci aux membres de l'équipe les plus fraîchement arrivés Marie-Adeline, Julien, Amélie et Minh Hien de contribuer à l'animation de l'équipe.

Je tiens à exprimer ma reconnaissance aux membres du personnel de l'IRCL et tout particulièrement à Micheline Magdelon pour son dévouement et pour m'avoir aidé dans mes démarches administratives. Merci également à Matthias Gérard pour sa bonne humeur légendaire.

Florence, je te remercie pour ton soutien depuis maintenant plus de sept ans. Merci également de me supporter, ce travail est aussi le tien, merci d'être là.

Un grand MERCI à Michelle et Roger de leur grande gentillesse, leur aide et de m'avoir accepté dans leur famille.

Pour terminer, je remercie tout particulièrement mes parents et mon frère pour tout ce qu'ils m'apportent. Merci grand frère pour ton soutien et pour l'exemple. Merci à mes parents, pour leur gentillesse, leur précieux réconfort durant toutes ces années d'études.

LISTE DES ABREVIATIONS

Liste des abréviations

ADN	Acide désoxyribonucléique
AMV-RT	Avian Myeloblastosis Virus Reverse Transcriptase
Antp	Antennapedia
ARN	Acide ribonucléique
ARNg	Acide ribonucléique guide
ATP	Adénosine triphosphate
Bcd	Bicoid
BCR/ABL	Breakpoint Cluster Region/Abelson
BRCA1	Breast Cancer 1
Brn	Brain
C/EBP	CCAAT/Enhancer Binding Protein
CBF	CCAAT-Binding Factor
CBP	CREB Binding Protein
CMV	Cytomégalo Virus
c-Myc	cellular Myelocytomatosis
CRE	cAMP-Regulated Enhancer
CREB	cAMP-Regulated Enhancer Binding
Cybb	Cytochrome b-245 beta polypeptide
DiOC ₆	3,3'-dihexyloxacarboyanine iodide
DMSO	Dimethylsulfoxyde
DO	Densité Optique
DTT	Dithiothréitol
EBS	ETS binding site
EDTA	Ethylene diamine tetraacetic acid
Erg	Ets related gene
ERK	Extracellular Signal-regulated Kinases
ETS	E Twenty-six
Fli-1	Friend leukemia integration 1
Flt3	fms-like tyrosine receptor-3
GCN4	general control nondepressible 4
G-CSF	granulocyte colony-stimulating factor
GH	Growth hormone
GIST	GastroIntestinal Stromal Tumor
GM-CSF	Granulocyte-Macrophage Colony-Stimulating Factor
h	Heure
HAT	Histone-acétyl-transférases
HBP	Hexapeptide Binding Pocket
HDAC	Histones desacétylases
HIV	Human Immunodeficiency virus 1
HLH	Helix-Loop-Helix
HMG	High Mobility Group
Hox	Homeobox
HPV-16	Human Papilloma Virus de type 16
HRP	Horseradish
Hsp70	Heat Shock protein 70
HTH	Hélice-Tour-Hélice
IP	Iodure de Propidium
JNK	c-Jun N-terminal Kinase

LAL-T	Leucémies Aiguës Lymphoïdes-T
LAM	Leucémies Aiguës Myéloïdes
LB	Luria Bertoni
LEF-1	Lymphoid Enhancer Bining Factor
LMC	Leucémies Myéloïdes Chroniques
MAPK	Mitogen-Activated Protein Kinase
MBq	Méga Becquerel
MeCP2	Methyl-CpG-binding Protein 2
Meis-1	Myeloid ecotropic viral integration site 1 homolog
min	Minute
mL	Millilitre
MLL	Mixed-Lineage Leukemia
mM	Millimolaire
MMP1	Matrix MetalloProtéase-1
NES	Nuclear Export Signal
NF-Y	Nuclear Factor-Y
NF-κB	Nuclear Factor-kappa B
ng	Nanogramme
NK	Natural Killer
NLS	Nuclear Localisation Signal
nM	Nanomolaire
NP 40	Nonidet P-40
Nup98	Nucléoporine 98
Oct-1	Octamer 1
OMS	Organisation Mondiale de la Santé
PBS	Phosphate Buffer Saline
Pbx	Pre-B cell leukemia homeobox
PEA3	Polyoma Enhancer Activator 3
PEG	Polyéthylène Glycol
PFA	Paraformaldéhyde
PIM	Pbx Interaction Motif
Pit-1	Pituary-specific positive transcription factor 1
PKC	Proteine Kinase C
PNT	Pointed
<i>PRL</i>	Prolactine
RAR	Retinoic Acid Receptor
RARE	Retinoic Acid Receptor Element
RelA	Reticuloendotheliosis viral oncogene homolog A
RXR	Retinoic X Receptor
SAPK	Strees Activated Protein Kinase
SCF	Stem Cell Factor
SDS	Sodium Dodécyl Sulfate
SMD	Syndrômes Myélodysplasiques
SV40	Simian Vacuolating virus 40
SVF	Sérum de Veau Fœtal
SWI/SNF	SWItch/Sucrose NON Fermentable
TA	Température Ambiante
TAFs	TBP-Associated Factors
TALE	Three Amino acid Loop Extension
TBE	Tris/Borate/EDTA

TBP	TATA Box Protein
TBS	Tris Buffered Saline
TBS _T	Tris Buffered Saline Tween 20
TCF	Ternary Complex Factor
TE	Tris EDTA
TEL	Translocation Ets leukemia
TGF β	Transforming Growth Factor β
TMB	3,3',5,5' -TetraMethylBenzidène
TPA	12-O-TetradecanoylPhorbol-13-Acetate
TPO	Thrombopoïétine
Tris	Tris(hydroxymethyl)aminomethane
TSH β	Thyroid-Stimulating Hormone β
UPE	Upstream Promoter Element
VDRE	Vitamine D Response Element
virus E26	Avian Erythroblastosis Virus
W	Watt
$\mu\text{g/L}$	Microgramme par litre
μM	Micromolaire

TABLE DES MATIERES

Table des matières

INTRODUCTION	1
I. Cancer et dérégulation de l'expression des gènes	1
II. Initiation de la transcription des gènes et régulation de l'expression	2
A. La chromatine et le contrôle de l'expression des gènes	2
1. Organisation de la chromatine	2
2. Altération de la structure de la chromatine	3
a. Altération structurale	3
b. Modifications chimiques des histones	4
B. Région promotrice et partenaires protéiques impliqués dans l'initiation de la transcription	5
1. La région promotrice	5
2. Le complexe de base	6
3. Les facteurs de transcription spécifiques	9
C. La régulation de l'expression des gènes	11
III. Structure de l'ADN et reconnaissance spécifique par les facteurs de transcription	12
A. Structure de l'ADN	12
B. Motifs de reconnaissance de l'ADN des facteurs de transcription	14
1. Le motif hélice-tour-hélice	14
2. Le motif à doigts de zinc	15
3. Les motifs à domaine basique	17
a. Le motif leucine zipper	17
b. Le motif hélice-boucle-hélice	18
4. Autres domaines	19
IV. Inhibition de l'expression des gènes par reconnaissance spécifique de l'ADN	21
A. Les lexitropsines, dérivés de la distamycine et de la nétropsine	21
1. La distamycine et la nétropsine	21
a. Structure	21
b. Interaction avec l'ADN	22
c. Reconnaissance des paires de bases G•C et C•G	22
d. Discrimination entre les appariements A•T et T•A	23
e. Augmentation de l'affinité et de la sélectivité de séquences	23
2. Effets biologiques	25
a. Distribution cellulaire	25
b. Inhibition de l'expression des gènes	25
B. Les dérivés DB	27
1. Relation structure/affinité, structure/sélectivité	27
a. Structure du bérénil et du DB75	27
b. Dérivés diphenyl-furanes	29
c. Les dérivés phényl-furane-benzimidazoles	30
2. Activité biologique des composés DB	36
a. Activité antiparasitaire	36
b. Activité cellulaire	39

V. Modèles d'études et leur implication dans le cancer	42
A. HoxA9 et les leucémies	42
1. Les gènes <i>hox</i> : organisation et fonction	42
a. Rôle dans le développement	42
b. Rôle des gènes <i>hox</i> chez l'adulte	43
(1) Maintien de la plasticité	43
(2) L'hématopoïèse	44
2. Régulation de l'expression des gènes <i>hox</i>	45
3. L'homéodomaine	46
a. Structure et reconnaissance spécifique	46
b. L'homéodomaine et la formation d'hétérodimère	47
4. Rôle dans la leucémogénèse	48
a. Le monomère HoxA9	49
b. <i>HoxA9</i> et ses cofacteurs	50
(1) Meis-1	50
(2) Pbx	51
c. La translocation t(7;11) et la protéine Nup98/HoxA9	51
5. Régulation de la protéine HoxA9	53
a. Le TGF β (transforming growth factor β)	53
b. Les phosphorylations	53
c. Localisation subcellulaire de HoxA9	53
d. Epissage alternatif de HoxA9	54
B. Les facteurs de transcription à POU domaine	54
1. Classification et fonctions	55
2. Le POU domaine: structure et fonctions	55
3. Implication dans le cancer	58
a. Le facteur de transcription Pit-1	58
b. Les facteurs de transcription Brn-3	59
(1) La protéine Brn-3a	59
(2) La protéine Brn-3b	59
C. Les facteurs de transcription de la famille Ets	60
1. Structure des facteurs de transcription ETS	60
a. Le domaine Ets	61
b. Le domaine PNT	62
2. Fonctions des protéines ETS	62
a. Fonctions cellulaires	62
b. Différenciation cellulaire	63
c. L'angiogénèse et la vascularisation	63
3. Régulation des protéines ETS	64
a. Régulation par interactions protéines-protéines	64
b. Régulation post-transcriptionnelle	64
(1) La phosphorylation	64
(2) L'acétylation	65
(3) La sumoylation	65
(4) L'auto-inhibition	65
4. Implication dans la cancérogenèse	66
a. Modifications géniques de ETS	66
(1) Amplification et surexpression des protéines ETS	66

(2) Les mutations	66
(3) Les translocations	67
b. Expression des gènes Ets et cancers	67
c. Rôle de la famille ETS dans l'invasion et l'angiogenèse	68
OBJECTIFS	69
RESULTATS	71
I. Nouveaux dérivés DB et sélectivité de reconnaissance de l'ADN	71
A. Nouveaux composés : Modifications du corps polycyclique	71
1. Les composés diphenyl-furanes	71
a. Sélectivité de séquence	71
b. Affinité relative des composés pour l'ADN	75
c. Relations structure/sélectivité et structure/affinité	77
2. Les composés phényl-furane-benzimidazoles	78
a. Sélectivité d'interaction	78
b. Affinité relative des composés pour l'ADN	80
c. Relations structure/sélectivité et structure/affinité	82
(1) Modification du cycle benzimidazole	82
(2) Modification du cycle phényle	82
(3). Modification du cycle furane	82
B. Nouveaux composés : Nouveaux sites de reconnaissance	82
1. DB1242 et séquence GCTCG	83
2. DB1255 et séquence CCATG	85
3. RT-29 et site ATTC	87
C. Nouveaux composés : Distribution et toxicité cellulaires	89
1. Influence des substituants portés par le corps de la molécule sur la distribution cellulaire	89
2. Influence du mode de fixation des cellules dans la localisation subcellulaire des composés	90
D. Effets cytotoxiques des composés DB	96
II. Modulation de la fixation des facteurs de transcription à l'ADN par les composés DB	98
A. Modulation d'un ensemble de facteurs de transcription par le DB293	98
1. TranSignal™ Protein/DNA Array I	98
2. Modulation de l'activité des facteurs de transcription par le DB293	102
3. Interaction du DB293 sur les sites consensus Pit-1, Brn-3 et IRF-1	105
a. Localisation du DB293 sur les sites Pit-1, Brn-3 et IRF-1	106
b. Stœchiométrie d'interaction	108
c. Localisation du DB293 sur le site consensus Pit-1	110
4. Identification des sites essentiels à la formation du complexe protéine/ADN	115
5. Importance du site ATGA dans l'inhibition des protéines Pit-1 et Brn-3	117
CONFIDENTIEL	
DISCUSSIONS ET PERSPECTIVES	137

CONCLUSION GENERALE	163
MATERIEL ET METHODES	165
I. Clonage moléculaire	165
A. Vecteurs de clonage	165
B. Culture bactérienne	165
1. Souche bactérienne	165
2. Culture	165
C. Transformation bactérienne	165
D. Amplification des plasmides	167
E. Construction plasmidique	167
II. Etude moléculaire	169
A. Etude de l'interaction ADN-ligand	169
1. Marquage radioactif sur l'extrémité 3' de l'ADN	169
a. Marquage à l'AMV-RT	169
b. Marquage à la Klenow	169
2. Marquage radioactif sur l'extrémité 5' de l'ADN	170
3. Empreinte à la DNase I	171
4. Test de dénaturation thermique de l'ADN	172
5. Résonance plasmonique de surface	173
B. Modulation de la fixation des facteurs de transcription à l'ADN	173
1. TranSignal TM /protein/DNA arrays (Panomics)	173
2. Expression des protéines en lysat de réticulocytes	175
3. Retard en gel	175
4. Screening de molécules : technique ELISA	176
5. Western Blotting	177
III. Etudes en cellules eucaryotes	178
A. Culture cellulaire	178
1. Cellules adhérentes	178
a. Cellules HT-29	178
b. Cellules B-16 GM-CSF	178
2. Cellules en suspension	179
a. Cellules MigA9	179
b. Cellules THP-1, Molt-3, K562	179
B. Microscopie de fluorescence	179
1. Distribution cellulaire	179
2. Fixation différentielle	180
C. Tests de cytotoxicité	180
D. Tests clonogéniques	181
1. Cellules HT-29	181
2. Cellules MigA9	181
E. Préparation d'extraits nucléaires	182
F. Extraction des protéines totales	182
G. Activité luciférase	183
REFERENCES BIBLIOGRAPHIQUES	184
ANNEXES	203

INTRODUCTION

I. Cancer et dérégulation de l'expression des gènes

Selon les sources de l'OMS, le cancer est responsable de 7,6 millions de décès dans le monde en 2005 et on estime qu'il en causera 9 millions en 2015 et 11,4 millions en 2030 (<http://www.who.int/fr>).

Le cancer est un désordre prolifératif qui survient quand l'équilibre entre la prolifération et la différenciation cellulaire est rompu. La prédominance de la prolifération sur la différenciation n'est pas le simple résultat d'une activation de gènes nécessaires à la division cellulaire mais est aussi due à l'inhibition de l'expression d'autres gènes limitant la prolifération (Kirmizis *et al.*, 2004). L'acquisition des caractéristiques de promoteur de prolifération et d'inhibiteur de mort cellulaire requière de multiples changements génétiques et épigénétiques dans la cellule. Parmi l'éventail thérapeutique proposé en oncologie, la chimiothérapie antitumorale fait appel bien souvent à des molécules anciennes comme le 5 fluoro-uracile, le cisplatine, le méthotrexate, les anthracyclines et les alcaloïdes de la pervenche. Ces molécules ciblent principalement l'ADN et agissent au moment de l'étape de synthèse de l'ADN où le fuseau mitotique et bloque la mitose. Cependant, ces molécules à large spectre d'action possèdent un index thérapeutique limité lié tout particulièrement à leur faible sélectivité d'action. Depuis la fin des années 1990 et l'avènement du glivec® dans le traitement des leucémies myéloïdes chroniques et des tumeurs gastro-intestinales stromales (GIST) (Wong *et al.*, 2004), la chimiothérapie antitumorale commence à faire place à de nouvelles molécules plus finement adaptées à la nature même de la tumeur et de son environnement.

L'expression inappropriée d'activateur et/ou de répresseur de transcription représente souvent un stade précoce du développement tumoral. Les facteurs de transcription régulent de nombreux phénomènes cellulaires comme la croissance, la différenciation ou encore l'apoptose. Certains oncogènes ont été identifiés comme étant des facteurs de transcription capables d'induire des transformations cellulaires responsables de phénomènes néoplasiques. De nombreuses études montrent l'intérêt du ciblage des facteurs de transcription dans le traitement du cancer (Boyd *et al.*, 1999 ; Pandolfi, 2001 ; Darnell, 2002). Les facteurs de transcription représentent donc une cible intéressante dans le développement de nouvelles molécules antitumorales. Ainsi la modulation de l'expression des gènes par des petites molécules antitumorales interagissant de façon séquence spécifique avec l'ADN représente une stratégie ambitieuse et novatrice pour le traitement des cancers.

II. Initiation de la transcription des gènes et régulation de leur expression

L'initiation de la transcription est l'un des points de contrôle majeur de l'expression des gènes eucaryotes. Depuis environ trente ans, l'utilisation de petites molécules, de nature et de mode d'interaction à l'ADN différents permet de moduler cette initiation en l'activant ou en l'inhibant. Pour bien comprendre le mode d'action de ces petites molécules, il est important de faire le point sur les mécanismes complexes qui régissent le recrutement de l'ARN polymérase et ainsi, l'initiation de la transcription proprement dite. Les ARN polymérases permettent la polymérisation des ribonucléotides dans la molécule d'ARN en fonction de la séquence d'ADN. Il existe 3 classes d'ARN polymérases dépendantes des gènes transcrits : (i) les ARN polymérases de classe I qui permettent la synthèse des ARN ribosomiaux (45S, 28S, 18S et 5.8S) (Sommerville *et al.*, 1984) ; (ii) les ARN polymérases de classe II qui contrôlent la genèse des ARN messagers de toutes les protéines ; (iii) les ARN polymérases de classe III responsables de la synthèse des ARN de transfert ainsi que de l'ARN ribosomal 5S (Ciliberto *et al.*, 1983). Nous nous intéresserons essentiellement à l'initiation de la transcription des gènes de type II.

A. La chromatine et le contrôle de l'expression des gènes

1. Organisation de la chromatine

Dans le noyau de chaque cellule humaine, l'ADN est fractionné et condensé sous forme de chromatine répartie en 23 paires de chromosomes. Ainsi une molécule d'ADN de longueur linéaire d'environ 10 cm sera condensée en un chromosome mesurant à peine 10 micromètres de long. De nombreux complexes protéiques sont impliqués dans cette condensation, les plus connus étant les histones. Ces protéines contiennent de nombreux résidus lysine et arginine dont les charges positives interagissent avec le squelette phosphodiester de la double hélice d'ADN.

L'ADN présente différents niveaux de compaction. L'unité fondamentale de la structure condensée est le nucléosome où 146 pb s'enroulent autour d'un octamère d'histones (deux molécules H2A, H2B, H3 et H4). Dans un second temps, les nucléosomes se compactent et forment une hélice au rythme de 6 nucléosomes par tour produisant une fibre de 30 nm

d'épaisseur. Les histones H1 permettent de maintenir en place la fibre de 30 nm (Horton *et al.*, 1993). Cette fibre est finalement condensée en euchromatine (peu condensée) ou en hétérochromatine (fortement condensée) constituant le troisième niveau d'organisation (Figure 1).

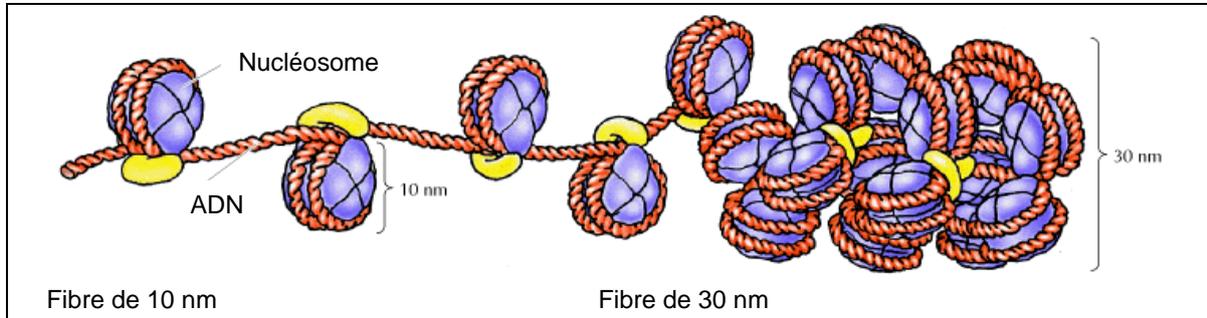


Figure 1 : Structure et assemblage des nucléosomes.

L'ADN (rouge) est enroulé autour d'un octamère d'histone (bleu) formant le nucléosome. Cette structure est verrouillée par l'histone H1 (jaune). Le repliement de ces histones forme la fibre de 30 nm (image provenant de : http://148.216.10.83/CELULA/4,2_chromosomes_and_chromatin.htm).

Il semble que seuls les gènes localisés dans l'euchromatine soient potentiellement transcrits. De nombreuses études ont montré l'importance de la structure de la chromatine et notamment du nucléosome dans la régulation de l'expression des gènes. Cette structure organisationnelle de l'ADN n'est pas uniquement un mode de compaction de l'ADN mais elle possède d'autres rôles importants notamment dans les mécanismes de répression de la transcription. Dans les régions promotrices ou dans les régions activatrices des gènes, ces structures doivent être éliminées ou modifiées pour faciliter l'interaction des facteurs de transcription sur l'ADN (Adams et Workman, 1993).

2. Altération de la structure de la chromatine

Certains mécanismes permettent d'altérer la structure de la chromatine et facilitent l'initiation de la transcription.

a) Altération structurale

Les facteurs de remodelage de la chromatine sont capables d'éliminer les nucléosomes par hydrolyse de l'ATP. C'est le cas des protéines SWI/SNF (*SWItch/Sucrose NON Fermentable*) (Pazin et Kadonaga, 1997). Ce complexe altère la structure du nucléosome et permet ainsi de libérer le site d'interaction des facteurs de transcription.

b) Modifications chimiques des histones

L'acétylation des groupements amines libres portés par des résidus lysine des histones diminuent leurs charges positives nettes et donc leur affinité pour l'ADN. La forme acétylée des histones est préférentiellement retrouvée au niveau de la chromatine active ou potentiellement active (Grunstein, 1997). Ainsi, la transcription peut être activée par des histone-acétyl-transférases (HAT) et inhibée par des histones-desacétylases (HDAC) (Figure 2).

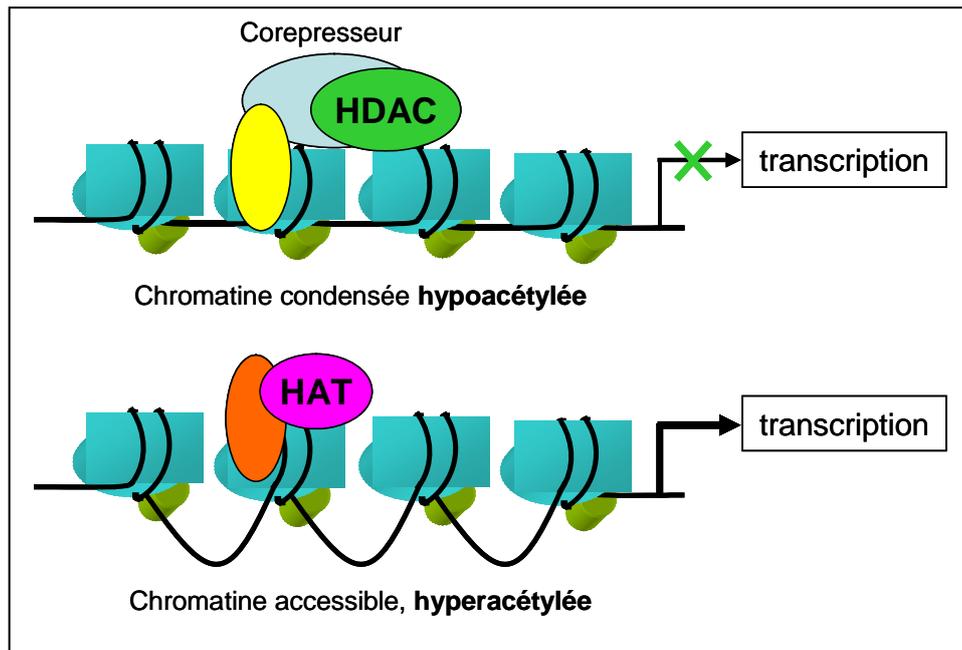


Figure 2 : Activation de la chromatine par acétylation des histones.

La déacétylation des histones par les HDAC permet le maintien de l'ADN sous forme condensé alors que l'acétylation des histones par les HAT diminue l'affinité des protéines pour l'ADN. Cette réduction de l'affinité passe par une réduction du nombre de liaisons électrostatiques entre l'ADN et les protéines.

Certaines protéines non histones comme les facteurs de transcription peuvent aussi être acétylés. L'acétylation de résidus lysine situés dans la partie C-terminale de la protéine p53 par la protéine p300 permet d'activer d'un facteur 30 la fixation du facteur à son site consensus. Cette activation corrèle avec l'effet antiprolifératif de la protéine p53 (Gu et Roeder, 1997). Le facteur de transcription NF- κ B peut lui aussi être acétylé au niveau de sa sous unité RelA (*Reticuloendotheliosis viral oncogene homolog A*) par les protéines p300 et CBP (*CREB Binding protein*). L'acétylation de NF- κ B (*Nuclear factor-kappa B*) diminue l'affinité de I κ B α pour le facteur de transcription. Ceci permet l'exportation du facteur de transcription dans le noyau par exposition du signal de localisation nucléaire. La protéine RelA peut être désacétylée par la protéine HDAC3 (Chen *et al.*, 2001).

Actuellement, de nombreux composés inhibiteurs d'histones déacétylases sont en développement clinique. Ils exercent leur activité cytotoxique par le biais d'un remodelage de la chromatine mais aussi par l'activation de protéines comme p53. Ces modulations directes et indirectes de la transcription impliquent des modifications de la croissance, de la survie et de la mort cellulaire (Marks *et al.*, 2001, Peixoto et Lansiaux, 2006).

Il existe d'autres modifications des histones comme la phosphorylation, l'ubiquitination ou encore la méthylation. A l'inverse de l'acétylation, ces modifications ont souvent lieu sur des sites spécifiques (Li *et al.*, 2007). Récemment, un lien a pu être établi entre méthylation de l'ADN et inhibition de la transcription. La méthylation de sites CpG d'un gène sur le résidu cytosine par une CpG méthyl-transférase diminue l'efficacité de transcription (Razin, 1998 ; Schübeler *et al.*, 2000). A l'inverse, des gènes inactifs peuvent produire des transcrits suite à la déméthylation des CpG par la 5-azacytidine (Chiurazzi *et al.*, 1998). La méthylation pourrait réduire l'affinité des facteurs de transcription pour leurs sites de reconnaissances spécifiques ou au contraire, elle pourrait augmenter la capacité de liaison de répresseurs sur l'ADN. L'un de ces répresseurs, MeCP2 (*methyl-CpG-binding protein 2*), interagit également avec des histones désacétylases établissant ainsi un lien moléculaire direct entre méthylation et déacétylation (Nan *et al.*, 1998).

B. Région promotrice et partenaires protéiques impliqués dans l'initiation de la transcription

1. La région promotrice

Il est difficile d'établir un modèle de région promotrice des gènes tant celle-ci est variable. Cependant, certains motifs basiques sont retrouvés dans un grand nombre de gènes comme les boîtes TATA, CAAT et GC qui permettent le recrutement de l'ARN polymérase II et un taux basal de transcription. A ces éléments s'ajoutent des éléments régulateurs spécifiques qui permettent une régulation intrinsèque plus fine dans l'espace et dans le temps, ou extrinsèque en réponse à des stimuli extérieurs.

Dans la majorité des zones promotrices, on trouve une région riche en paires de bases AT appelée boîte TATA (*TATA-box*) ou « promoteur proximal » de séquence

TATA(A/T)A(A/T). Elle se situe entre les paires de bases -30 et +30 du site d'initiation. Ce site, aussi appelé promoteur minimal, joue un rôle crucial dans l'initiation de la transcription (Latchman *et al.*, 1998) par sa capacité à induire un niveau basal de transcription. C'est à ce niveau que sont recrutés les facteurs de transcription généraux et l'ARN polymérase de type II (Sentenac, 1985).

En amont de la *TATA-box*, on peut trouver des éléments proximaux ou UPE pour (*Upstream Promoter Element*) capables d'augmenter le niveau basal de transcription. Ces éléments sont retrouvés dans une large variété de gènes et permettent la stimulation de l'activité constitutive du promoteur. On trouve notamment les boîtes riches en paires de bases GC (connues sous le nom de GC-box ou encore *Sp -box*) mais aussi la *CCAAT-box* (Dynam et Tjian, 1985).

D'autres éléments sont identifiés dans un petit nombre de gènes, on les appelle les éléments régulateurs. Ces séquences permettent l'expression de gènes en réponse à un stimulus comme pour le *heat-shock element* qui permet l'expression du gène codant pour la protéine Hsp70 (*Heat Shock protein*) en réponse à une augmentation de la température (Pelham, 1982). Les éléments régulateurs sont aussi responsables de l'expression de gènes spécifiques à certains tissus, c'est notamment le cas de la séquence ATGAATAA reconnue par le facteur de transcription Pit-1 (pituitary-specific positive transcription factor 1). Ces sites sont retrouvés dans des zones promotrices de gènes exprimés spécifiquement dans la glande pituitaire antérieure et permettent notamment l'expression de la prolactine (Andersen et Rosenfeld., 1994).

Une des caractéristiques des gènes eucaryotes est l'existence d'éléments régulateurs qui peuvent être localisés à une grande distance du site d'initiation de la transcription et qui influent sur le niveau d'expression des gènes. Ces éléments peuvent se situer en amont, en aval et même directement sur le site de transcription. Ils peuvent augmenter l'activité transcriptionnelle (*enhancer*) ou la diminuer (*silencer*). Certains modulateurs sont actifs dans tous les tissus et augmentent l'activité des promoteurs alors que d'autres sont spécifiques à certains tissus. Les modulateurs interagissent avec les protéines impliquées dans le complexe d'initiation de la transcription par formation d'une boucle sur l'ADN.

2. Le complexe de base

Le complexe de base est composé des facteurs de transcription généraux et de l'ARN polymérase. Ces facteurs de transcription sont des complexes protéiques nécessaires au recrutement de l'ARN polymérase II.

Le premier complexe à intervenir est TF_{II}D dont la sous unité appelée TBP (*TATA box protein*) se lie à l'ADN. Cette protéine interagit directement sur la *TATA-box* et est la seule composante du complexe basal à interagir de façon spécifique avec l'ADN. Les autres composants de TF_{II}D, les TAFs (*TBP-associated factors*) n'interagissent pas directement avec la *TATA-box* mais permettent une réponse aux activateurs de transcription (Tansey et Herr, 1997). Certaines protéines du complexe TF_{II}D permettent la modification de la structure de la chromatine comme la protéine TAF_{II}250. Cette acétyl-transférase est capable d'acétyler les histones et permet la décompaction de la chromatine (Mizzen *et al.*, 1996 ; Wu, 1997). La TBP joue un rôle crucial dans la transcription impliquant l'ARN polymérase de type II de part son interaction avec la *TATA-box* (Burley *et al.*, 1997). Elle permet la première étape de l'assemblage du complexe d'initiation. La structure de cette protéine a été identifiée par cristallographie aux rayons X et ressemble à une selle de cheval. La protéine interagit avec l'ADN et est capable de modifier localement la structure de la double hélice en lui conférant une forme concave (Figure 3).

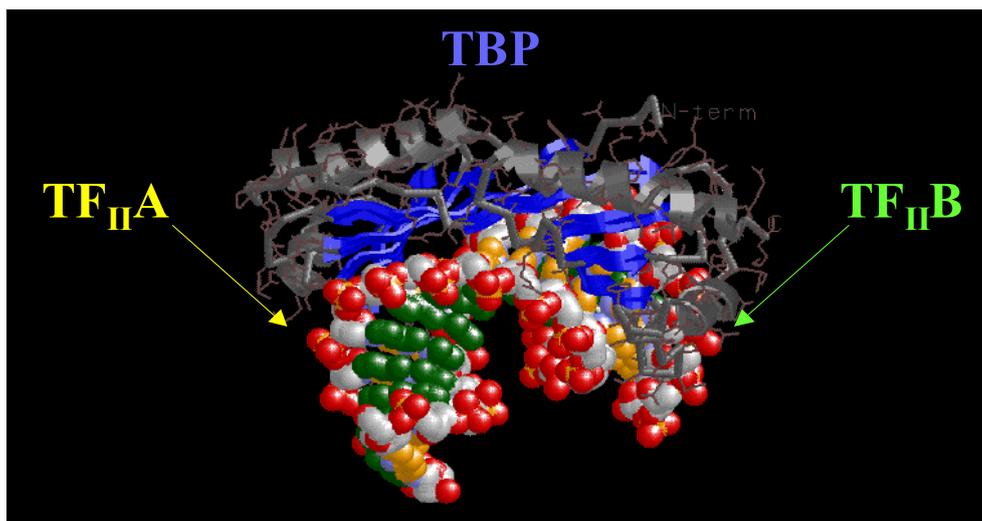


Figure 3 : Interaction entre la TBP et la TATA box.

L'interaction entre la TBP (par ces feuillet β en bleu) et la *TATA box* induit une torsion importante au niveau de l'ADN permettant le recrutement des facteurs TF_{II}A et TF_{II}B.

La courbure de l'ADN induite par la TBP permet l'assemblage du complexe de transcription basal. La première protéine recrutée est TF_{II}A. Ce facteur stabilise le complexe, il interagit avec des acides aminés situés sur la partie N-terminale de la TBP et sur des régions situées en amont de la *TATA-box* de l'ADN. De plus TF_{II}A joue un rôle dans la réponse médiée par les activateurs de transcription.

TF_{II}B, puis le complexe TF_{II}F-ARN polymérase sont recrutés au niveau du site d'initiation de la transcription (Roeder, 1996). TF_{II}B interagit du côté C-terminal de TBP et permet de faire un lien entre la TBP et l'ARN polymérase II.

L'ARN polymérase de type II est composée de 12 sous unités de tailles variables (de 10 à 220 kDa) qui interagissent entre elles pour former un complexe multimérique (Sentenac, 1985). Une des sous unités (la seconde par la taille) contient le site enzymatique responsable de la réaction de polymérisation. La sous-unité la plus volumineuse du complexe possède une extrémité carboxy-terminale unique comportant plusieurs répétitions de l'heptapeptide de séquence consensus YSPTSPS. Ce polypeptide se localise à l'extérieur du complexe de l'ARN polymérase et est accessible à la phosphorylation. Ce domaine carboxy-terminal doit être phosphorylé pour permettre l'initiation de la transcription. Cette phosphorylation est un moyen de régulation de la transcription, des stimuli extérieurs comme les facteurs de croissance permettent, par une cascade signalétique une augmentation de la phosphorylation et donc de la transcription (Dubois *et al.*, 1994). D'autres études ont montré que des protéines impliquées dans les modifications post-transcriptionnelles peuvent s'associer avec cette région qui aurait aussi un rôle dans les phénomènes d'épissage (Corden et Patturajan, 1997).

TF_{II}F, composé de deux sous-unités, est le seul facteur qui reste associé à l'ARN polymérase pendant la phase d'élongation. Après le recrutement de la polymérase, trois autres facteurs se fixent au complexe : TF_{II}E, TF_{II}J et TF_{II}H. TF_{II}H possède une activité hélicase qui permet de décondenser l'ADN et possède une activité kinase capable de phosphoryler le domaine C terminal YSPTSPS de l'ARN polymérase II. Cette forme phosphorylée est active et nécessaire à l'élongation.

Ce complexe formé de 7 facteurs de transcription généraux et de l'ARN polymérase est appelé complexe basal de transcription ou complexe stable (Figure 4). Après l'initiation de la transcription, seuls TF_{II}D et TF_{II}A restent liés à l'ADN et permettent le recrutement des autres membres du complexe basal pour recommencer un cycle de transcription.

Il existe un second modèle d'initiation de la transcription appelé « holoenzyme ». Ce modèle est basé sur la mise en évidence de l'association en solution de l'ARN polymérase avec TF_{II}B, TF_{II}F, TF_{II}E, TF_{II}J et TF_{II}H en absence d'ADN. Ce complexe pourrait se lier directement à la TBP pour débiter la transcription (Pugh, 1996).

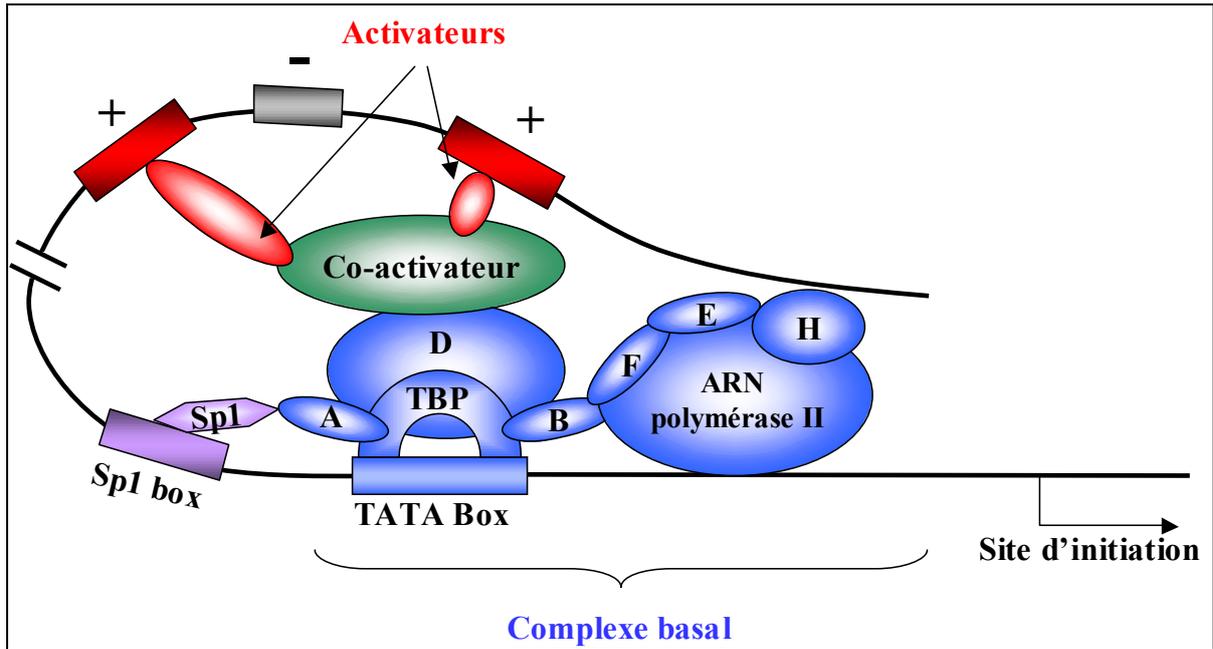


Figure 4 : Complexe d'initiation de la transcription.

Interaction classique des différents complexes protéiques sur une zone promotrice. L'interaction entre le facteur de transcription Sp1 (en violet) et la *SP1-box*, ainsi que le recrutement d'activateur (en rouge) et de co-activateur (en vert) permettent la mise en place du complexe basal de transcription (en bleu) et de l'initiation de la transcription.

La TBP permet également l'initiation de la transcription de gènes qui ne contiennent pas de *TATA-box*. Dans ce cas, elle est recrutée par une protéine qui se fixe à l'ADN et chevauche le site d'initiation de la transcription. La TBP peut alors recruter le facteur $TF_{II}B$ et l'ARN polymérase de type II pour initier la transcription.

3. Les facteurs de transcription spécifiques

Le contrôle de la transcription est en grande partie géré par les facteurs de transcription. Ce sont des protéines qui se lient à une séquence spécifique de l'ADN et induisent, facilitent ou inhibent la transcription des gènes. Ils peuvent notamment influencer, par le recrutement de protéines associées, sur le niveau de condensation de la chromatine et sur l'initiation de la transcription. Ces facteurs sont généralement composés de deux domaines distincts : un domaine de liaison à l'ADN et un domaine nécessaire à l'activité transcriptionnelle. Alors que le domaine de liaison permet au facteur de transcription de reconnaître un site spécifique de l'ADN sur un promoteur précis, le(s) domaine(s) qui porte(nt) l'activité transcriptionnelle

permette(nt) le recrutement de la machinerie basale de transcription de façon directe ou indirecte via des co-activateurs (Figure 4).

Les facteurs de transcription peuvent être classés selon leur site de fixation à l'ADN. De nombreux domaines de liaisons à l'ADN ont été identifiés. On les classe en fonction de l'homologie de leurs domaines de liaison à l'ADN en « *super-classes* » (ex : facteurs de transcription à hélice-tour-hélice) divisée en « *familles* » (ex : la famille à POU domaine) et même en « *sous-familles* » (ex : Pit-1).

Fonctionnellement, on peut également distinguer 2 types de facteurs spécifiques :

i) ***Les facteurs de transcription ubiquitaires.*** Ces facteurs augmentent l'efficacité de la transcription et ont des séquences nucléotidiques cibles présentes dans un grand nombre de promoteurs de gènes dans des tissus très variés. C'est le cas du facteur de transcription Sp1 qui reconnaît une boîte spécifique riche en paire de base GC : la *Sp1-box* (GGGCGG) (Lania *et al.*, 1997). Ce facteur est capable d'activer le promoteur du virus SV40 (Simian vacuolating virus 40), où la *Sp-box* est présente en tandem de 6 copies. Ce facteur de transcription appartient à la super-classe des facteurs de transcription à doigt de zinc. Il possède dans son domaine carboxy-terminal deux régions riches en glutamine qui interagissent avec les TAFs associées à la TBP. Comme le motif *Sp1*, la *CCAAT-box* est présente dans de nombreux promoteurs de gènes et joue un rôle essentiel dans l'activation de la transcription. La *CCAAT-box* peut interagir avec différents facteurs de transcription exprimés dans de nombreux tissus. Le premier facteur de transcription identifié comme interagissant avec la *CCAAT-box* a été la protéine CTF (*CCAAT-box transcription factor*). D'autres facteurs peuvent interagir avec la *CCAAT-box* comme C/EBP (*CCAAT/enhancer binding protein*) ou encore NF-Y (*Nuclear factor-Y*) aussi appelé CBF (*CCAAT-binding factor*). Ce dernier est impliqué dans la transactivation de nombreux gènes comme le gène codant pour l'ADN topoisomérase II α humaine (Maity et de Crombrughe, 1998).

ii) ***Les facteurs de transcription synthétisés ou activés de façon spécifique dans des tissus précis et à différents stades du développement.*** Ces facteurs peuvent être inductibles, répondant à un stimulus extérieur. Les séquences qu'ils reconnaissent sont appelées les éléments de réponse. Leur dérégulation peut être responsable de différentes pathologies. Différents exemples sont développés partie V.

C. La régulation de l'expression des gènes

L'activation ou l'inhibition d'un gène passe par des stimuli externes ou internes entraînant la liaison d'un certain nombre de facteurs de transcription sur leur site consensus de l'ADN. Certains activateurs de transcription vont recruter des complexes protéiques capables de désorganiser les nucléosomes empêchant la formation du complexe basal sur la *TATA-box*, alors que d'autres, appelés répresseurs, stabilisent les nucléosomes (Kingston *et al.*, 1996 ; Felsenfeld *et al.*, 1996). Par exemple, le facteur de transcription CREB (*cAMP-regulated enhancer binding*) interagit avec l'élément de réponse CRE (*cAMP-regulated enhancer*) et il est activé par phosphorylation par l'intermédiaire de la protéine kinase A en réponse à une augmentation de la concentration en AMP cyclique dans la cellule (Mellon *et al.*, 1989). CREB recrute alors la protéine co-activatrice CBP (Chrivia *et al.*, 1993). Deux mécanismes d'activation de la transcription entrent alors en jeu. Tout d'abord CBP interagit directement avec TF_{II}B et aide au recrutement de l'ARN polymérase (Nakajima *et al.*, 1997). Ensuite, CBP recrute d'autres co-activateurs comme la protéine p300 ; ces deux protéines possèdent une activité histone-acétylase rendant la structure de la chromatine compatible avec la transcription (Ogryzko *et al.*, 1996). Le complexe basal de transcription se met alors en place et la transcription peut démarrer.

III. Structure de l'ADN et reconnaissance spécifique par les facteurs de transcription

A. Structure de l'ADN

Pour comprendre comment les protéines ou les petites molécules interagissent avec l'ADN, il est nécessaire de revenir sur quelques règles simples qui gouvernent la structure secondaire de l'ADN.

La structure la plus courante de l'ADN, la forme B, se présente sous forme d'hélice droite qui contient 10 paires de bases par tour. L'angle formé entre deux paires de bases consécutives est de $35,9^\circ$ avec une distance entre deux paires de bases consécutives est de $3,4 \text{ \AA}$. L'ADN est formé de deux sillons de taille différente. Le premier, large est appelé grand sillon et le second, plus étroit, est appelé petit sillon (Figure 5a). La taille des sillons dépend des paires de bases qui les composent. Il est important de remarquer que chaque paire de base présente une partie de leur structure dans le grand sillon et une autre dans le petit sillon (Figure 5b). Il a été établi par le groupe de Rich en 1976 que la reconnaissance spécifique entre l'ADN et les protéines s'effectue selon un modèle simple de donneur et d'accepteur de liaisons hydrogène au niveau des sillons de l'ADN (Figure 5c) (Seeman *et al.*, 1976). Selon les paires de bases et le sillon ciblés, les réseaux de liaisons hydrogène potentiellement réalisables entre l'ADN et les protéines seront différents. Le code de reconnaissance au niveau du grand sillon fait intervenir 4 liaisons hydrogène alors qu'au niveau du petit sillon ce code est simplifié avec seulement 3 liaisons hydrogène. Il existe cependant d'autres types d'interactions faibles moins spécifiques entre l'ADN et les protéines, comme les interactions électrostatiques ou hydrophobes entre les résidus de la protéine et le squelette phosphodiester de l'ADN ou encore les forces de Van der Waals. La prédominance d'interaction au niveau d'un sillon plutôt que l'autre dépend en grande partie des caractéristiques énergétiques des deux sillons. L'association d'une protéine au niveau du grand sillon est favorisée par un processus enthalpique. Le petit sillon est caractérisé par une enthalpie défavorable qui peut être contrebalancée par des caractéristiques entropiques favorables. Ces différences énergétiques sont en grande partie dues à la différence des propriétés d'hydratations des deux sillons, notamment au niveau des sites riches en paires de bases AT. Sur ces séquences, le petit sillon de l'ADN présente un haut niveau d'hydratation et d'organisation qui apporte une

contribution entropique positive lorsqu'il est désorganisé par une molécule ou une protéine (Haq, 2002 ; Privalov *et al.*, 2007).

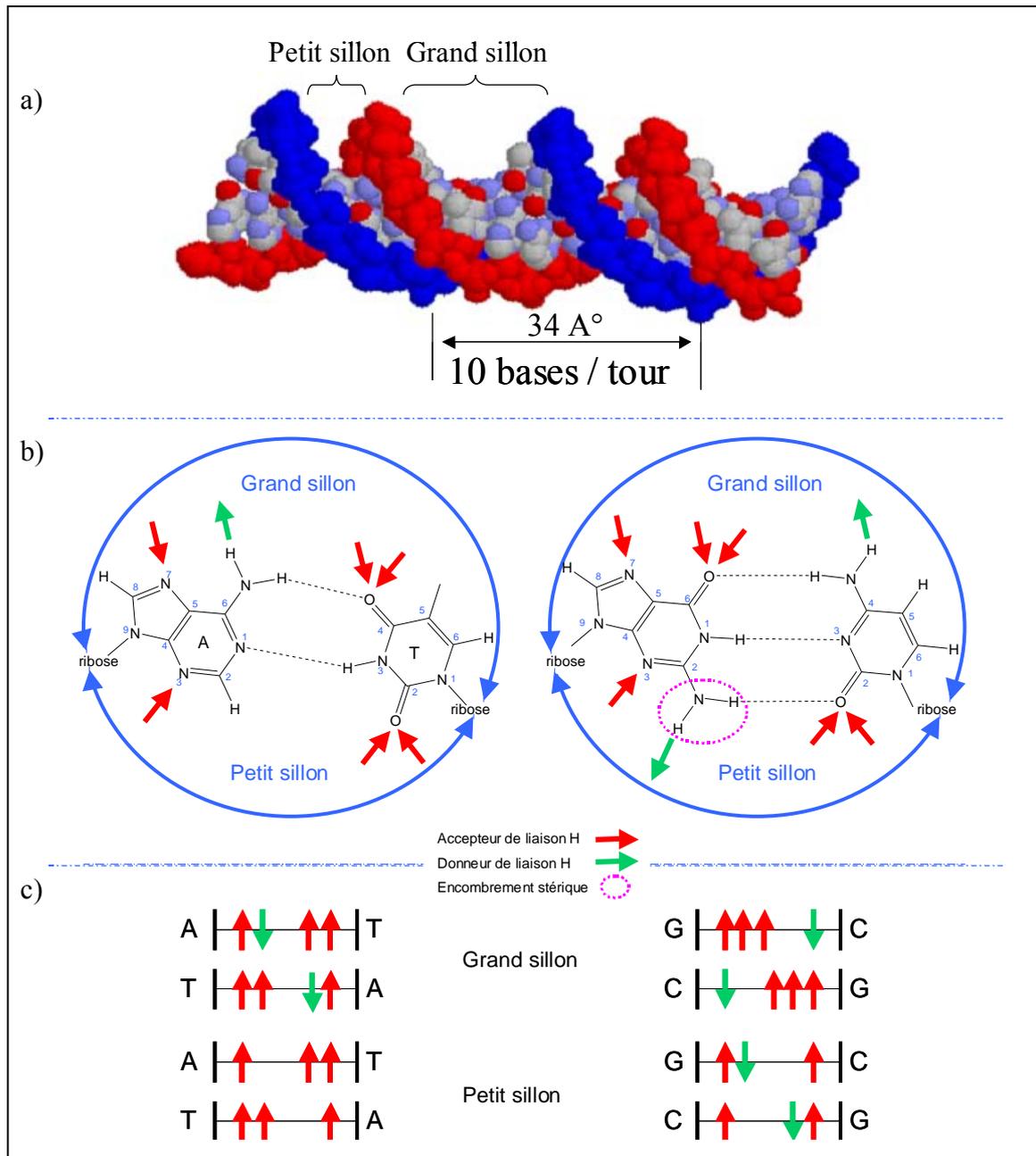


Figure 5 : Structure de l'ADN de forme B.

a) Structure de l'ADN. Les chaînes phosphodiesteres des deux brins d'ADN sont représentées en bleu et en rouge. Les deux sillons de l'ADN présentés sont de profondeur presque identique mais de largeur différente. b) Orientation des paires bases au niveau des deux sillons de l'ADN c) Code de reconnaissance des différentes paires de bases A-T/T-A et G-C/C-G au niveau des deux sillons de l'ADN.

B. Motifs de reconnaissance de l'ADN des facteurs de transcription

Afin de bien positionner les acides aminés essentiels à l'établissement des liaisons hydrogène spécifiques entre les protéines et l'ADN, les facteurs de transcription sont dotés d'un motif de reconnaissance de l'ADN.

Ces trente dernières années, les connaissances sur les domaines de liaison à l'ADN des facteurs de transcription ont connu un essor considérable. L'amélioration des techniques (Résonance magnétique de surface (RMN), cristallographie aux rayons X, modélisation moléculaire) permet de mieux appréhender ces interactions spécifiques qui représentent la base même de la modulation de l'expression des gènes. Les différents motifs présentés ici sont ceux qui sont le plus souvent retrouvés dans les facteurs de transcription. La plupart d'entre eux reconnaissent le grand sillon de l'ADN où ils forment des liaisons faibles et spécifiques avec les paires de bases. Ces structures permettent d'orienter certains acides aminés au niveau de paires de bases précises de la double hélice. Ces modèles permettent de comprendre comment des modifications locales de la structure de la double hélice (induites par exemple par des petites molécules) peuvent être dramatiques pour cette reconnaissance ADN/facteurs de transcription.

1. Le motif hélice-tour-hélice

Le motif hélice-tour-hélice (HTH) est retrouvé dans la plupart des facteurs de transcription procaryotes et eucaryotes. On le retrouve notamment dans la famille des facteurs de transcription à homéodomaine. Le HTH est généralement défini comme un segment de 20 résidus s'organisant en deux hélices α liées par un coude (tour), formant un angle de 120° entre les deux hélices. Ce motif peut être complété par une ou plusieurs hélice(s) du côté C-terminal et un bras de six acides aminés du côté N-terminal. La ou les hélice(s) supplémentaire(s) permette(nt) d'orienter le motif HTH au niveau du grand sillon de l'ADN alors que le bras du côté N-terminal interagit avec le petit sillon de l'ADN.

La seconde hélice α du motif HTH est l'hélice de reconnaissance qui se lie au grand sillon de l'ADN de manière séquence-spécifique (Whartonn *et al.*, 1984). Le positionnement correct de l'hélice dans le grand sillon de l'ADN est primordial car il permet l'alignement des acides aminés au niveau du site consensus de l'ADN et dépend des structures adjacentes à ce domaine minimal (Figure 6a). La reconnaissance de sites consensus s'effectue par des

interactions spécifiques entre les paires de bases et les acides aminés présents au niveau de l'hélice de reconnaissance. De façon intéressante, la seule mutation du résidu lysine en position 9 de l'hélice de reconnaissance de la protéine Bcd (*Bicoid*) en une glutamine confère à cette dernière les propriétés de reconnaissance de la protéine Antp (*Antennapedia*) (Figure 6b). En effet, cette lysine effectue des liaisons hydrogène avec l'oxygène 6 et l'azote 7 du résidu guanine du site consensus Bcd, alors que le groupe amide de la glutamine dans la même position reconnaît une adénine par formation de liaisons hydrogène avec les azotes N6 et N7.

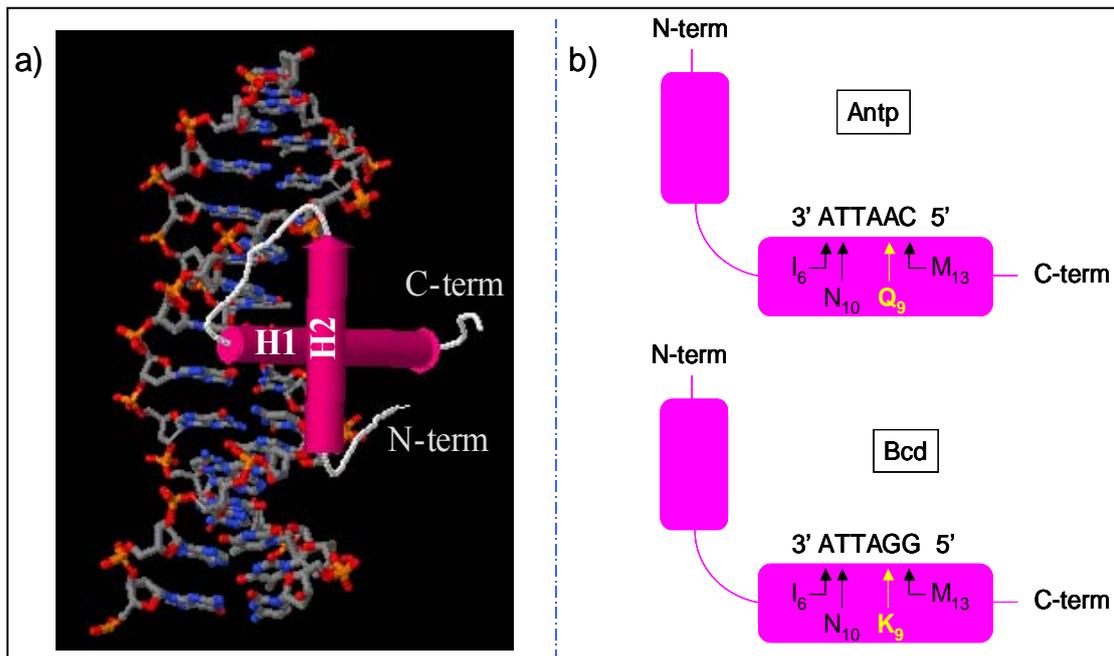


Figure 6 : Motif hélice-tour-hélice.

a) L'hélice de reconnaissance du motif HTH se positionne au niveau du grand sillon de l'ADN où elle effectue des liaisons spécifiques. (<http://arapaho.nsuok.edu/~biology/Tutorials/DNAbinding.htm>). b) Le résidu en position 9 de l'hélice de reconnaissance des protéines Antp et Bcd permet de discriminer la reconnaissance des sites consensus 3'-ATTAAC et 3'-ATTAGG par formation de liaisons hydrogène spécifiques.

2. Le motif à doigts de zinc

Ce second type de domaine d'interaction à l'ADN contient un atome de zinc comme élément structurant. Il existe 3 familles de protéines à doigts de zinc. Tout d'abord le motif à doigts de zinc proprement dit, constitué d'environ 30 acides aminés dont la structure est maintenue par un atome de zinc lié à deux résidus cystéine et deux résidus histidine. Il constitue notamment le domaine d'interaction de la protéine TF_{III}A. La seconde famille contient environ 70 résidus

et est retrouvée dans les récepteurs stéroïdiens où deux atomes de zinc sont liés chacun par quatre résidus cystéine.

La séquence en acides aminés retrouvée dans les facteurs de transcription possédant un motif à doigts de zinc deux cystéines, deux histidines est le suivant : Tyr/Phe-X-Cys-X-Cys-X₂₋₄-Cys-X₃-Phe-X₅-Leu-X₂-His-X₃₋₄-His-X₅, où X représente un résidu à caractère variable. Cette structure contient deux cystéines (en rouge) et deux histidines (en bleu) invariables qui interagissent avec l'atome de zinc (Figure 7a). Entre la seconde cystéine et la première leucine, on retrouve 12 acides aminés. Ces derniers sont composés de résidus leucine, phénylalanine ainsi que des acides aminés basiques très conservés indispensables à l'établissement de liaisons spécifiques avec l'ADN. Un unique motif à doigts de zinc ne permet pas de reconnaître à lui seul une séquence spécifique; il est alors indispensable qu'une protéine possède plusieurs motifs à doigts de zinc pour reconnaître une séquence spécifique. C'est pour cela que cette séquence est répétée successivement trois fois ou plus au sein du domaine de liaison à l'ADN (Figure 7c), on retrouve notamment neuf motifs répétés dans le facteur de transcription TF_{III}A.

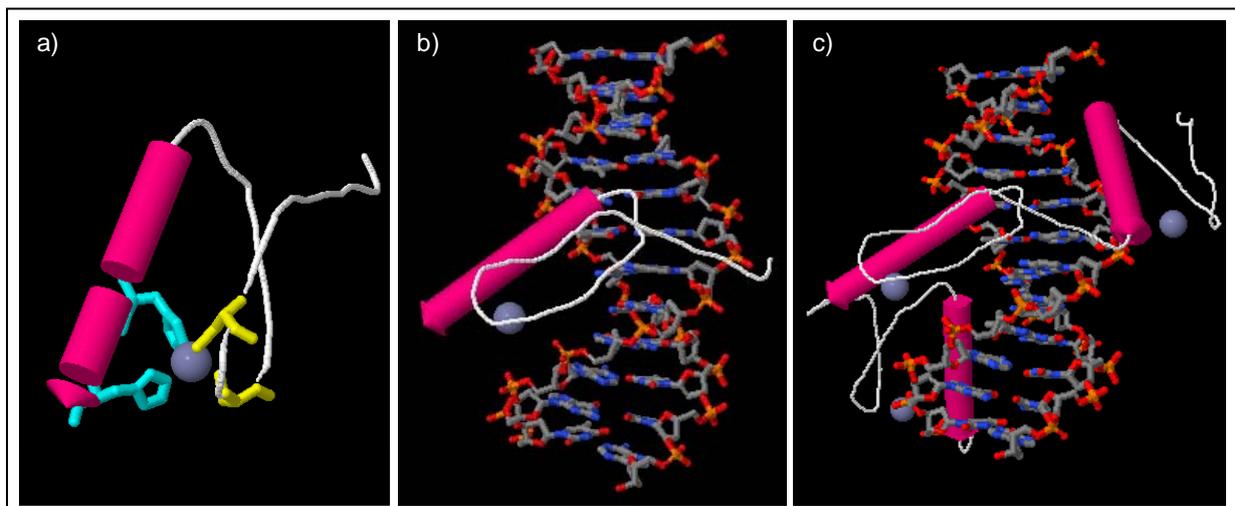


Figure 7 : Structure du motif à doigts de zinc.

a) Les deux histidines (en bleu) et les deux cystéines (en jaune) interagissent avec l'atome de zinc, servant de base à la formation de l'hélice α (en rose) et au feuillet β (en gris). b) Dans l'interaction entre le motif à doigts de zinc et l'ADN, l'hélice α effectue des liaisons spécifiques avec le grand sillon de l'ADN alors que les feuillets β se lient au squelette phosphodiester de l'ADN. c) Les facteurs de transcription à doigts de zinc possèdent plusieurs motifs consécutifs responsables de la reconnaissance spécifique de l'ADN (modélisation réalisée à partir du site <http://arapaho.nsuok.edu/~biology/Tutorials/DNAbinding.htm>).

Dans cette structure, l'atome de zinc engage 4 liaisons avec les cystéines et les histidines et sert d'échafaudage à l'établissement des structures secondaires et tertiaires du motif. Ce domaine forme deux feuillets β anti-parallèles et une hélice α adjacente (Figure 7b) (Lee *et*

al., 1989). L'hélice α effectue des contacts spécifiques avec le grand sillon de l'ADN pendant que les feuillets β effectuent des contacts avec le squelette phosphodiester de l'ADN. L'hélice α , et plus particulièrement les résidus situés entre les deux histidines, possède donc un rôle critique dans la reconnaissance spécifique de l'ADN (Nardelli *et al.*, 1991 ; Choo et Klug, 1994).

3. Les motifs à domaine basique

a) Le motif leucine zipper

Cette structure est présente dans de nombreux facteurs de transcription comme C/EBP (*CCAAT/enhancer binding protein*), GCN4 (*general control nondepressible 4*) ou encore les protéines oncogènes Myc, Fos et Jun. Ce motif consiste en une région riche en leucine où cet acide aminé y est retrouvé tout les sept résidus (Hurst, 1996). Ce domaine a la capacité de se dimériser avec un segment d'hélice α pour former le domaine leucine *zipper*. La substitution d'un résidu leucine par une valine abolit totalement la capacité de ces protéines à former des dimères ce qui, par la même occasion, inhibe l'interaction entre les protéines et l'ADN. Les hélices permettant l'interaction entre les deux protéines sont orientées parallèlement et dans la même direction (Figure 8a). Les motifs leucine *zipper* des deux protéines (semblables aux deux composantes d'une fermeture éclair) sont capables d'interagir entre elles comme une glissière pour former un dimère stable (Figure 8a). Par opposition aux motifs HTH et doigts de zinc, le motif leucine *zipper* n'est pas en soit le domaine d'interaction avec l'ADN et n'effectue pas de contact direct avec l'ADN. Cependant, il facilite le contact entre une région adjacente riche en résidus basiques et le grand sillon de l'ADN (Figure 8b,c). En accord avec ce modèle, des mutations sur les domaines basiques inhibent l'interaction ADN/protéines mais n'inhibent pas la formation du dimère. Ainsi la spécificité de reconnaissance des facteurs de transcription à leucine *zipper* est déterminée par la séquence en acides aminés du domaine basique.

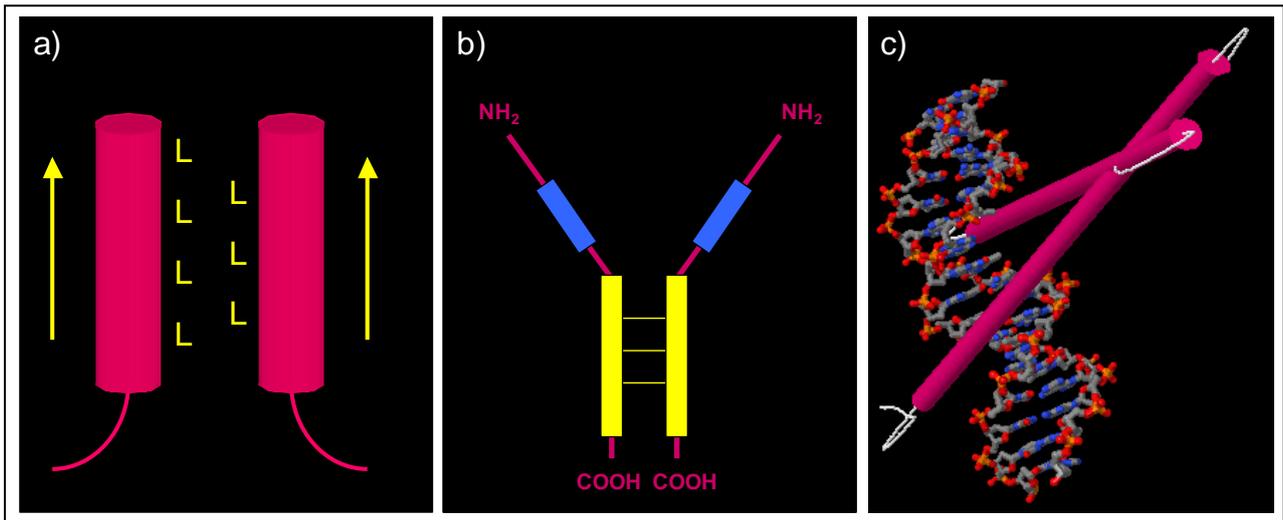


Figure 8: Motif leucine zipper.

a) Les domaines riches en leucines des deux protéines interagissent de façon parallèle avec la même orientation. b) L'interaction entre les domaines leucine zipper (en jaune) des deux protéines permettent le bon positionnement des régions riches en résidus basiques (en bleu). c) Interaction entre le dimère de protéines à leucine zipper et le grand sillon de l'ADN (modélisation réalisée à partir du site <http://arapaho.nsuok.edu/~biology/Tutorials/DNAbinding.htm>).

b) Le motif hélice-boucle-hélice

Le domaine HLH (pour *Helix-Loop-Helix*) est constitué de 40 à 50 acides aminés formant deux hélices α séparées par une boucle (Figure 9) (Littlewood et Evan, 1995). Comme précédemment, ce motif ne se lie pas directement à l'ADN, il est complété par des régions basiques capables d'interagir de façon spécifique avec l'ADN. Ces protéines forment des dimères via l'interaction entre des résidus hydrophobes situés sur les hélices α (Figure 9a). Dans la plupart des protéines à HLH, le domaine basique est adjacent au domaine HLH et effectue des interactions spécifiques avec l'ADN (Figure 9a-b). Le domaine HLH joue donc un rôle similaire au motif leucine zipper permettant la dimérisation de deux protéines, ce qui facilite les interactions entre l'ADN et le complexe protéique.

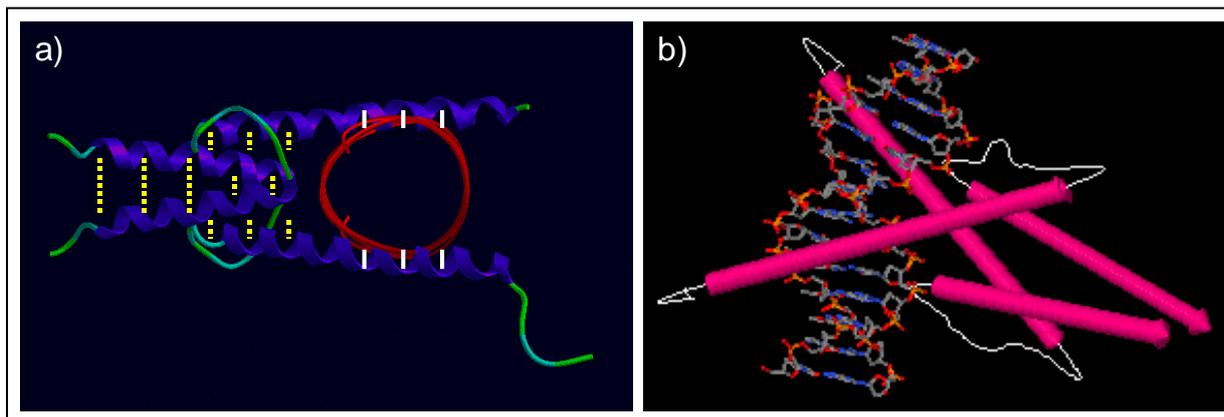


Figure 9 : Motif Hélice-boucle-Hélice.

a) Les deux hélices (en bleu) du domaine HLH effectuent des liaisons hydrophobes entre-elles (traits pointillés jaunes). L'hélice α (généralement la plus longue) de chaque monomère du complexe effectue des interactions spécifiques électrostatiques et hydrogène (traits pleins blancs) avec l'ADN (en rouge). b) Les régions basiques des protéines interagissent au niveau du grand sillon de l'ADN. (modélisation réalisée à partir du site <http://arapaho.nsuok.edu/~biology/Tutorials/DNAbinding.htm>).

4. Autres domaines

La grande majorité des facteurs de transcription identifiée dérive de ces motifs de liaison à l'ADN. Il existe néanmoins d'autres motifs de reconnaissance moins fréquents. Sept familles de protéines, facteurs de transcription ou protéines impliquées dans la structure de la chromatine, possèdent des hélices α orientées de manière différente de celles décrites précédemment. C'est notamment le cas des protéines HMG (High Mobility Group) ou encore des histones (Figure 10). L'octamère d'histone adopte une structure cylindrique autour de laquelle l'ADN vient s'enrouler. Les hélices α effectuent de nombreux contacts avec le squelette phosphodiester pour stabiliser la distortion de l'ADN. Aucune de ces hélices ne s'incèrent dans les sillons de l'ADN, il existe donc très peu de contacts entre les histones et les paires de bases. C'est pour cela que l'interaction entre ces protéines et l'ADN est non spécifique (Luger *et al.*, 1997). Pour les protéines HMG, chaque sous unité comprend 3 hélices α qui forment un L. Les hélices 1 et 2 interagissent avec les paires de bases et le squelette phosphodiester au niveau du petit sillon de l'ADN ce qui cause de sévères distorsions de la structure de l'ADN (Murphy *et al.*, 1999). D'autres protéines, comme par la protéine TBP, possèdent des motifs totalement différents et interagissent avec l'ADN par l'intermédiaire de feuillets β (Figure 10).

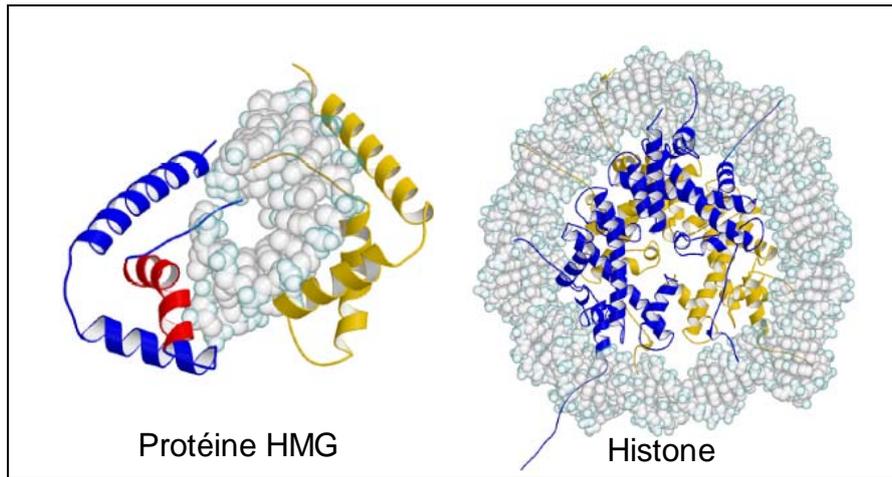


Figure 10 : Autres motifs d'interaction à l'ADN.

Les protéines HMG et histones possèdent des hélices α impliquées dans la reconnaissance plus ou moins spécifique de l'ADN (figure tirée de Luscombe *et al.*, 2000). L'interaction entre ces protéines et l'ADN provoque de fortes modifications structurales de la double hélice.

IV. Inhibition de l'expression des gènes par reconnaissance spécifique de l'ADN

De nombreuses équipes développent des agents qui modulent l'expression de certains gènes par reconnaissance spécifique de la double hélice. La capacité d'inhiber la fixation d'un facteur de transcription oncogène à sa séquence cible ; ou au contraire, d'activer l'interaction d'un facteur anti-oncogène à son site consensus représente une approche thérapeutique très prometteuse.

Certaines règles régissent l'interaction entre ces composés et l'ADN. La plupart d'entre elles rejoignent celles énoncées pour l'interaction entre l'ADN et les protéines. Cependant, elles nécessitent d'être complétées à cause des différences d'orientation et de structure entre les molécules et les protéines. Les molécules qui interagissent au niveau du petit sillon de l'ADN sont ancrées dans celui-ci, il est donc important que la molécule adopte une courbure proche de celle du petit sillon de l'ADN pour se positionner correctement. La taille ainsi que l'angle formés par la molécule doivent respecter les caractéristiques de flexibilité de la double hélice (Goodsell et Dickerson, 1986 ; Zhurkin *et al.*, 1991). Ces caractéristiques sont dépendantes de la nature des bases qui compose la séquence d'ADN. Comme pour les protéines, le complexe molécule/ADN doit être stabilisé par des liaisons faibles de différents types : i) liaisons hydrogène, par l'intermédiaire de groupements accepteurs et donneurs de liaison hydrogène (Zakrzewska *et al.*, 1987) ; ii) liaisons électrostatiques, c'est pour cela que la plupart de ces composés sont des molécules cationiques pouvant interagir avec le squelette phosphate électronégatif de l'ADN ; iii) liaisons hydrogène indirectes par l'intermédiaire de molécules d'eau pour assurer la stabilité entre l'ADN et son ligand (Brown *et al.*, 1990).

A. Les lexitropsines, dérivés de la distamycine et de la nétropsine

1. La distamycine et la nétropsine

a) Structure

La distamycine et la nétropsine sont des composés naturels (Figure 11) formées de 2 ou 3 cycles pyrroles liés les uns aux autres par des groupements amides capables d'effectuer des liaisons hydrogène avec les paires de bases de l'ADN. La partie C-terminale (définie par homologie avec les séquences peptidiques) de ces deux molécules possède un groupement

propylamidine. La partie N-terminale de la distamycine porte un groupement formyle alors que la nétropsine porte un groupement guanidine (revue : Wemmer, 2000 ; Dervan *et al.*, 2005).

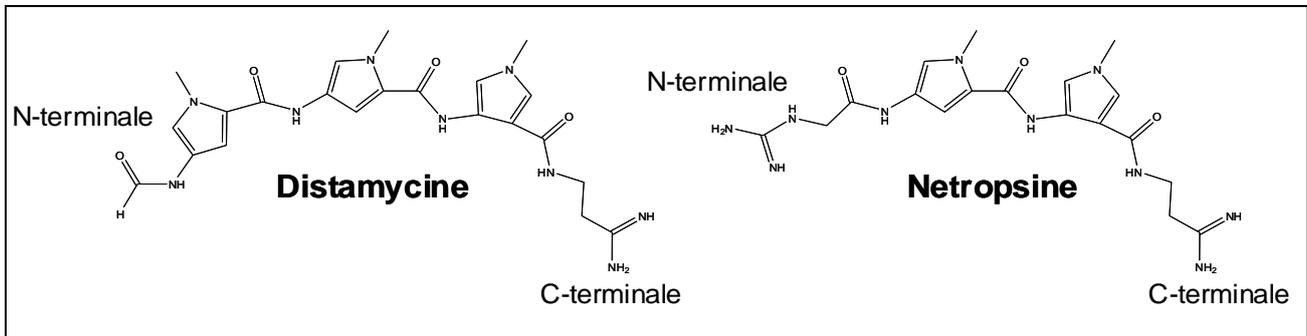


Figure 11 : Structures chimiques de la distamycine et de la nétropsine.

b) Interaction avec l'ADN

Ces dérivés reconnaissent des séquences riches en paires de bases AT au niveau du petit sillon de l'ADN. (Patel, 1982 ; Kopka *et al.*, 1985 ; Coll *et al.*, 1989 ; Pelton et Wemmer, 1989). Le composé disymétrique distamycine s'oriente dans la double hélice, le côté N-terminal de la molécule interagit avec une région 5' riche en adénine (Van Dyke *et al.*, 1982). Pour la nétropsine les deux orientations sont équivalentes quelque soit l'enchaînement des paires de bases AT. La distamycine a la capacité de former des dimères au niveau des séquences riches en paires de bases AT (Pelton *et al.*, 1989) alors que la nétropsine se fixe uniquement sous forme de monomère. Cette différence s'explique par le nombre de charges portée par les molécules. Les deux charges positives de la nétropsine empêchent la formation de dimère (répulsion des charges).

c) Reconnaissance des paires de bases G•C et C•G

Pour comprendre les mécanismes de reconnaissance de l'ADN par des petites molécules, il est indispensable de bien connaître la nature et l'organisation spatiale des substituants portés par les différentes paires de bases au niveau du petit sillon. La différence importante entre les appariements A•T et G•C provient du groupement amine situé sur l'atome de carbone 2 de la guanine (Figure 5c). Ce dernier crée un encombrement stérique important qui empêche la reconnaissance du petit sillon par les petites molécules.

Il est important de pouvoir créer des molécules capables de reconnaître les paires de bases GC afin de moduler la fixation de facteurs de transcription qui reconnaissent des sites consensus contenant ces paires de bases. Pour créer des molécules capables d'interagir avec les paires de

bases GC, il est nécessaire de diminuer l'encombrement stérique et d'établir de nouvelles liaisons avec le groupement amine de la guanine. Dans cette hypothèse, les équipes de Dickerson et Lown (Lown *et al.*, 1986) ont remplacé un cycle pyrrole de la nétropsine par un cycle imidazole. Malgré cette modification, les nouvelles molécules n'ont pas montré de préférence significative pour les séquences contenant des paires de bases GC par rapport aux sites riches en paires de bases AT (Lee *et al.*, 1988). Des composés pyrrole-imidazole-pyrrole (PyImPy) dérivés de la distamycine ont également été synthétisés. Ces molécules, du type lexitropsine se fixent préférentiellement sur la séquence AAGTT par rapport à la séquence AAATT (Dwyer *et al.*, 1992). La reconnaissance entre la molécule et la séquence s'effectue alors sous forme de dimère antiparallèle. La structure du petit sillon au niveau des séquences contenant des paires de bases GC étant moins profond mais plus large que celui des sites riches en paires de bases AT (Heinemann *et al.*, 1992), l'interaction de deux molécules permet un meilleur ancrage au sein de la double hélice. De plus, le dimère permet d'augmenter la spécificité de fixation à l'ADN par l'établissement d'un plus grand nombre de liaisons hydrogène. Néanmoins, ces liaisons n'augmentent pas l'affinité de la molécule pour l'ADN à cause de la proximité entre la cytosine du brin complémentaire et le cycle imidazole porté par la seconde molécule. Pour une meilleure interaction avec les paires de bases GC, des hétérodimères PyImPy et PyPyPy sont utilisés. Ceci permet de discriminer la reconnaissance des appariements G•C (par le couple de molécule Im-Py) par rapport aux appariements C•G (par le couple Py-Im) (Figure 12a) (Geierstanger *et al.*, 1994).

d) Discrimination entre les appariements A•T et T•A

Ce code de reconnaissance établi par l'agencement de cycles pyrroles et imidazoles possède une dégénérescence au niveau de la reconnaissance des appariements A•T et T•A, l'appariement des cycles Py-Py ne discriminant pas ces deux sites. Le remplacement d'un pyrrole par un cycle hydroxypyrrole (Hp) permet d'introduire un nouveau groupement fonctionnel capable d'effectuer des liaisons hydrogène préférentielles avec le carbonyle de la thymine (Figure 12b) (White *et al.*, 1998). Ainsi le couple Hp-Py reconnaît préférentiellement l'appariement T•A et le couple Py-Hp interagit préférentiellement avec l'appariement A•T.

e) Augmentation de l'affinité et de la sélectivité de séquences

Pour éviter que les deux molécules du dimère n'interagissent sous forme de monomères indépendants, les deux molécules composant le dimère ont été liées de façon covalente par

différents *linkers* intramoléculaire (Figure 12c) (Mrksich *et al.*, 1993 ; 1994, Parks, *et al.*, 1996).

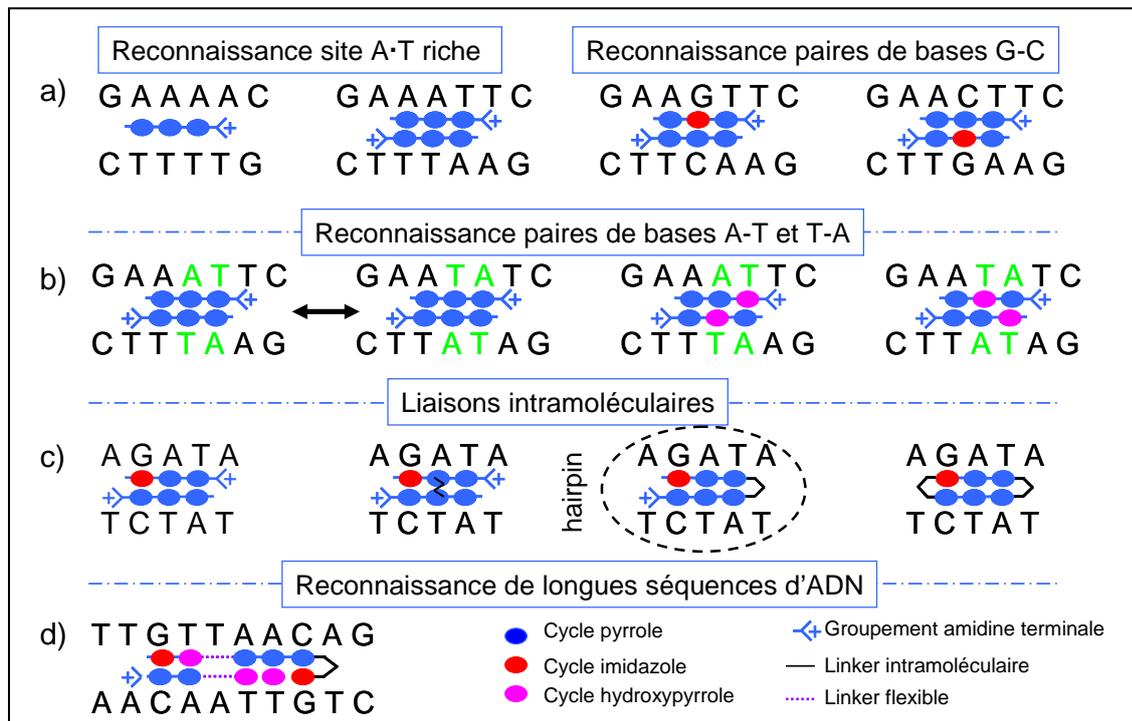


Figure 12 : Reconnaissance polyamines/ADN.

a) Le groupement amidine terminal interagit avec les paires de bases AT. La reconnaissance entre les paires de bases G•C et C•G s'effectue par le remplacement d'un cycle pyrrole par un imidazole (accepteur de liaisons hydrogène). b) Discrimination entre les paires de bases A•T et T•A par le cycle hydroxy-pyrrole (donneur de liaison hydrogène). c) Liaisons covalentes possibles entre les deux molécules du dimère. Les molécules les plus utilisées étant les *hairpins*. d) Flexibilité apportée à la molécule par des linkers du type glycine ou β -alanine. Ces linkers sont reconnus par les paires de bases AT.

En résumé, le *design* à façon de molécules reconnaissant des séquences spécifiques de l'ADN a été rendu possible par une meilleure compréhension des interactions entre les paires de bases. L'affinité entre les polyamines et l'ADN est optimale pour des molécules qui possèdent une structure composée de 5 cycles consécutifs au maximum. Les composés plus longs sont moins affins à cause des différences entre la courbure du petit sillon et la courbure des molécules qui les éloignent l'un de l'autre (Kielkopf *et al.*, 1997). Pour palier à ce problème des linkers flexibles (glycine et β -alanine) ont été ajoutés à l'intérieur de la molécule. Cette flexibilité facilite le positionnement optimal des différents cycles dans le petit sillon (Figure 12d).

2. Effets biologiques

a) *Distribution cellulaire*

S'il est possible d'établir des règles précises de la reconnaissance séquence-spécifique entre les pyrrole-imidazoles-polyamides et l'ADN, la distribution cellulaire de ces composés est régie par des règles plus complexes. De plus, ces molécules ne sont pas fluorescentes, ce qui complique l'étude de leur distribution cellulaire.

Ces composés pénètrent dans la cellule par un processus de diffusion. Pour rendre ces molécules fluorescentes, toute une série de dérivés possédant un fluorophore a été synthétisée. Mais l'ajout d'un fluorophore peut cependant modifier la distribution de ces composés. En effet, même si ces molécules gardent, pour la plupart d'entre elles, une spécificité équivalente à la molécule sans le fluorophore, leur affinité pour l'ADN est diminuée d'un facteur 10.

La distribution nucléaire des polyamides dépend tout d'abord du type cellulaire (Belisky *et al.*, 2002). Des études de relation structure/distribution ont montré que la présence sur la partie C-terminale d'un résidu β -alanine est défavorable à la pénétration du composé dans le noyau. La distribution cellulaire de ces composés dépend aussi de l'enchaînement des cycles qui compose la molécule. La masse moléculaire du composé n'influe pas sur la distribution du produit. (Best *et al.*, 2003 ; Edelson *et al.*, 2004). Il est donc assez difficile de prédire la localisation cellulaire de ces molécules au regard de leur structure.

b) *Inhibition de l'expression des gènes*

L'inhibition sélective du gène codant pour l'ARNr 5S par le blocage stérique du facteur de transcription TF_{III}A sur sa séquence cible a été le premier modèle d'inhibition de l'activité d'un facteur de transcription par les composés lexitropsines. Ces molécules interagissent avec l'ADN sur la séquence 5'-TGGTTA et bloquent l'accès au petit sillon du quatrième domaine à doigts de zinc du facteur de transcription (Neely *et al.*, 1997; Gottesfeld *et al.*, 1997), ce qui a pour conséquence d'empêcher le recrutement de l'ARN polymérase III.

Par la suite, de nombreux exemples de modulation de la fixation de facteurs de transcription par les lexitropsines ont été décrits. Certains composés inhibent l'interaction du complexe ternaire TF_{II}D/TF_{II}A/ADN sur la TATA box (Ehley *et al.*, 2002) alors que d'autres inhibent l'interaction entre la CCAAT box et le complexe NF-Y sur le promoteur du gène Topo II α . (Hochhauser *et al.*, 2007).

L'interaction entre de multiples facteurs de transcription et l'ADN est nécessaire à la synthèse du virus HIV-1 (*Human Immunodeficiency virus 1*). La combinaison de deux polyamines

interagissant sur le promoteur de HIV-1 permet d'inhiber l'activité de liaison à l'ADN de trois facteurs de transcription TFIIID, Ets-1 et LEF-1 (*Lymphoid enhancer binding factor 1*) bloquant ainsi la réplication du virus dans des lymphocytes isolés. (Dickinson *et al.*, 1998).

D'autres molécules inhibent l'interaction à l'ADN du facteur de transcription NFκB. L'interaction des polyamines au niveau du petit sillon modifie la structure de l'ADN ce qui empêcherait la protéine p60 d'interagir avec le grand sillon de la double hélice (Wurtz *et al.*, 2002).

Cependant, l'inhibition par les lexitropsines de facteurs de transcription qui interagissent au niveau du grand sillon de l'ADN n'est pas toujours possible. Les modifications de la structure de l'ADN induites par les polyamines ne sont pas toujours suffisantes pour bloquer l'accès au grand sillon, c'est notamment le cas du facteur de transcription GCN4. Ainsi, pour réussir à inhiber la fixation de ce facteur à l'ADN, deux types de molécules bifonctionnelles ont été synthétisées. Tout d'abord des composés possédant un motif Arg-Pro-Arg. Le peptide cationique porté par la molécule interagit avec les phosphates électro-négatifs de l'ADN et empêche le facteur GCN4 de reconnaître le squelette phosphate de la double hélice (Bremer *et al.*, 1998). La seconde stratégie utilisée pour bloquer l'interaction entre GCN4 et l'ADN a été d'ajouter un groupement intercalant dérivé de l'acridine au composé polyamine. L'intercalant se fixe alors au niveau du site de reconnaissance du facteur de transcription GCN4 et induit des changements considérables de la structure du grand sillon. Ceci a pour conséquence d'abolir l'interaction entre le facteur de transcription et l'ADN (Fechter *et al.*, 2003).

Il est également possible d'empêcher le facteur de transcription ESX de se fixer sur le site *ETS binding site* (EBS) dans la zone promotrice du gène codant pour le récepteur HER2/neu. Cette inhibition bloque l'expression du récepteur HER2 dans des modèles de transcription *in vitro* utilisant des extraits nucléaires. Néanmoins, la faible distribution nucléaire du composé n'a pas permis de valider ce résultat dans la cellule (Belitski *et al.*, 2002).

Ces composés représentent donc un modèle d'étude intéressant. De plus, les connaissances accumulées lors du développement de ces molécules ont servi de base à l'élaboration de nouveaux composés aux structures plus originales.

B. Les dérivés DB

Comme pour les lexitropsines, des études qui ont été elles menées au laboratoire en collaboration entre le Dr Christian Bailly et les équipes d'Atlanta ont permis de définir les spécificités d'interaction des dérivés DB avec l'ADN au niveau du petit sillon de la double hélice.

1. Relation structure/affinité, structure/sélectivité

a) Structure du bérénil et du DB75

Les dérivés DB sont des analogues des composés bérénil et de la pentamidine (Figure 13). Ces composés sont classiquement utilisés dans le traitement d'infections parasitaires. Le bérénil interagit au niveau du petit sillon de l'ADN où il reconnaît trois paires de bases AT consécutives. Le DB75 est un dérivé diphényl-furane (figure 13a). Le remplacement du groupement central triazène (bérénil) par un cycle furane (DB75) permet d'élargir l'angle formé par le composé (de 121 à 125° (Figure 13b)). Cet élargissement de quatre degrés augmente la distance entre les carbones portés par les groupements amidines terminales (11.9 Å pour le bérénil à 12.5 Å pour le DB75) et l'affinité de la molécule pour l'ADN car il permet un meilleur ancrage dans le petit sillon. Même si cette différence paraît minime (0,6 Å) elle est suffisante pour accroître la reconnaissance par la molécule pour l'ADN. En effet, le DB75 se fixe sur 4 paires de bases AT consécutives (Laughton et al., 1996).

Le mode d'interaction entre l'ADN et chacune des molécules est différent : pour le bérénil, une des extrémités amidine terminale interagit directement avec l'ADN et l'autre extrémité indirectement par l'intermédiaire d'une molécule d'eau (Figure 13d). Pour le complexe bérénil/ADN, l'atome d'azote N_1 forme une liaison hydrogène avec l'atome d'azote N_3 de l'amine portée par l'adénine 18 ; à l'autre extrémité, l'atome d'hydrogène porté par l'atome d'azote N_{20} effectue des liaisons hydrogène avec l'atome d'oxygène O_4 de l'adénosine 6 et de l'atome d'azote N_3 de l'amine portée par l'adénine 5 par l'intermédiaire d'une molécule d'eau.

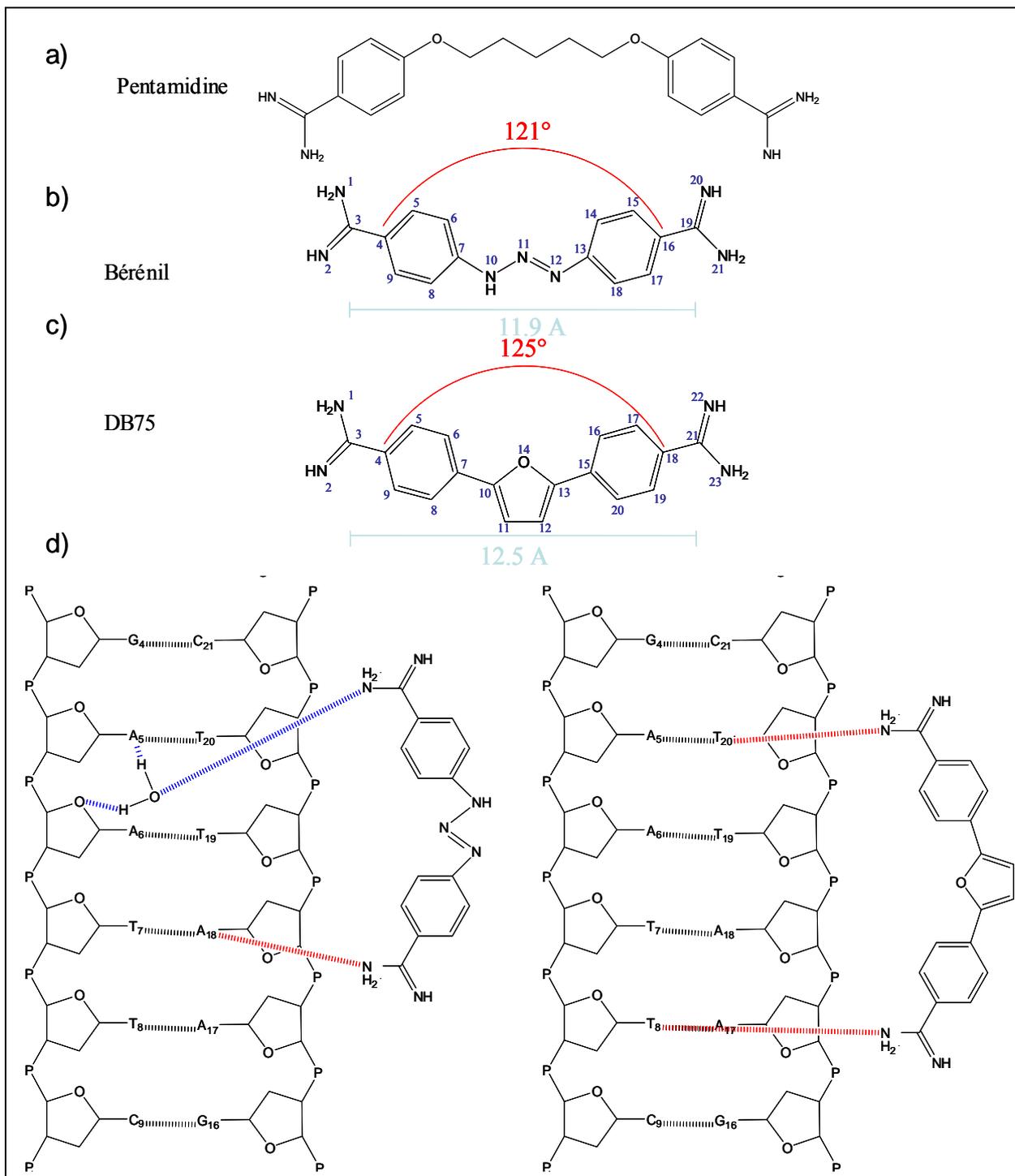


Figure 13 : Structures et interactions du bérénil et du DB75 avec l'ADN.

a) Structure de la pentamidine b) c) Structure, taille et angle formé par les composés bérénil et DB75. c) Liaisons hydrogène formées par le bérénil (à gauche) et le DB75 (à droite) avec l'ADN. Les traits pointillés rouges représentent les liaisons hydrogène directes et les traits pointillés bleus représentent les liaisons hydrogène établies par l'intermédiaire d'une molécule d'eau.

Pour le DB75, les deux groupements amidines interagissent directement avec l'ADN. Un des atomes d'hydrogène porté par les atomes d'azotes en position N₁ et N₂₃ effectue des liaisons

hydrogène avec l'atome d'oxygène O₂ des thymines 8 et 20 (figure 13c). Des interactions de van der Waals stabilisent le complexe entre chacune de ces molécules et l'ADN en particulier entre les carbones C₁₁ et C₁₂ des cycles furanes et les sucres des thymidines T8 et T20. Enfin, des liaisons électrostatiques s'établissent entre les amidines terminales chargées positivement à pH physiologique et le squelette phosphate électronégatif de l'ADN.

b) Dérivés diphényl-furanes

Comme nous venons de l'évoquer, la courbure de la molécule joue un rôle important dans son affinité et son mode d'interaction à l'ADN entre le bérénil et le DB75. Dans cette optique des nouveaux composés pour lesquels les groupements amidines présentes différentes positions ont été synthétisés. La modification de l'angle de la molécule formé par le déplacement des groupements amidines sur les cycles phényles en position para-para (DB75), méta-para (DB555) ou méta-méta (DB359) (Figure 14a) change totalement le mode de fixation, l'affinité, la sélectivité et la cinétique d'interaction de la molécule à l'ADN.

En effet, Le DB555 interagit faiblement sur les séquences riches en paires de bases AT au niveau du petit sillon de l'ADN et plus fortement sur des séquences riches en paires de bases GC sous forme d'intercalation. Le DB359 interagit quant à lui avec l'ADN uniquement sous forme d'intercalation sur des séquences riches en paires de base GC (Nguyen et al., 2002).

Comme nous l'avons vu, le DB75 se fixe sur 4 paires de bases AT successives. Ce sont les amidines terminales qui permettent la reconnaissance spécifique des dérivés pour les séquences riches en paires de bases AT. Des molécules aux amidines terminales diversement substituées ont été synthétisées (Figure14b). Il existe un degré de liberté important au niveau des substitutions possibles apportées sur ces groupements, même si certaines d'entre elles sont plutôt défavorables. L'ajout de groupements cyclopropyles (DB181) et surtout isopropyles (DB180) et cyclopentyles (DB244) (figure 21c) augmente l'affinité des molécules pour l'ADN. Cette meilleure affinité s'explique par l'augmentation de la surface de contact entre le groupement amidine substitué et l'ADN, qui permet de nouvelles liaisons non covalentes par l'intermédiaire de molécules d'eau. Il n'est cependant pas possible d'affirmer que les composés possédant des amidines substituées sont tous plus affins pour l'ADN que le DB75 non substitué, le composé 3-pentyle (DB226) étant moins affiné pour l'ADN que le DB75 (Trent et al., 1996 ; Wilson et al., 1998).

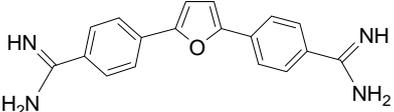
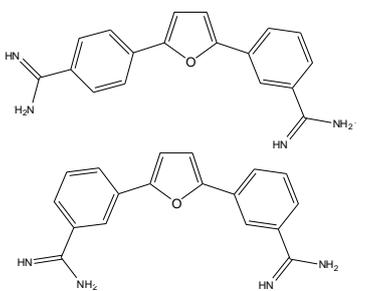
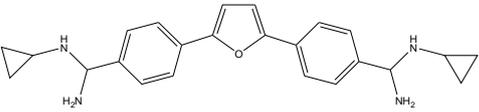
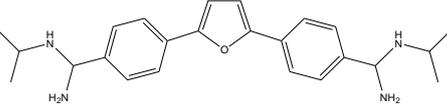
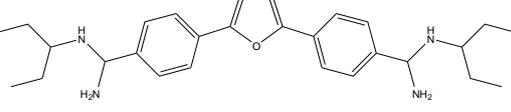
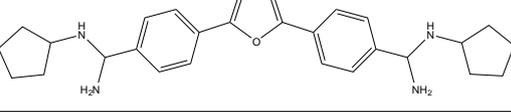
a)	Structure	c) Affinité	
		AT riches	GC riches
DB75		++	+/-
DB359		-	++
b)			
DB180		++	+/-
DB181		+++	+/-
DB226		+	+/-
DB244		+++	+/-

Figure 14 : Structure et affinités relatives des composés diphenyl-furanes.

a) Structure des composés DB75, DB555 et DB359 b) Structures des dérivés DB substitués au niveau des groupements amidines c) Affinité relative des dérivés DB pour les séquences riches en paires de bases AT (AT riches) ou GC (GC riches) de l'ADN.

c) Les dérivés phényl-furane-benzimidazoles

Dans le but d'augmenter la spécificité de séquence d'interaction des dérivés DB à l'ADN, de nouveaux composés ont encore été synthétisés dans lesquels un ou les deux cycle(s) phényle(s) ont été substitués par un ou deux cycle(s) benzimidazole(s). Ces modifications ont permis d'obtenir deux nouvelles molécules : l'une asymétrique aboutissant au composé phényl-furane-benzimidazole (DB293) et l'autre symétrique correspondant au dérivé dibenzimidazole-furane DB270 (figure 15a).

Le DB293 se fixe sous forme de dimère antiparallèle dans le petit sillon de l'ADN sur des séquences spécifiques contenant des paires de bases AT et GC. Cette molécule garde néanmoins la capacité d'interagir sous forme de monomère sur des séquences riches en paires de bases AT de l'ADN (Figure 15c).

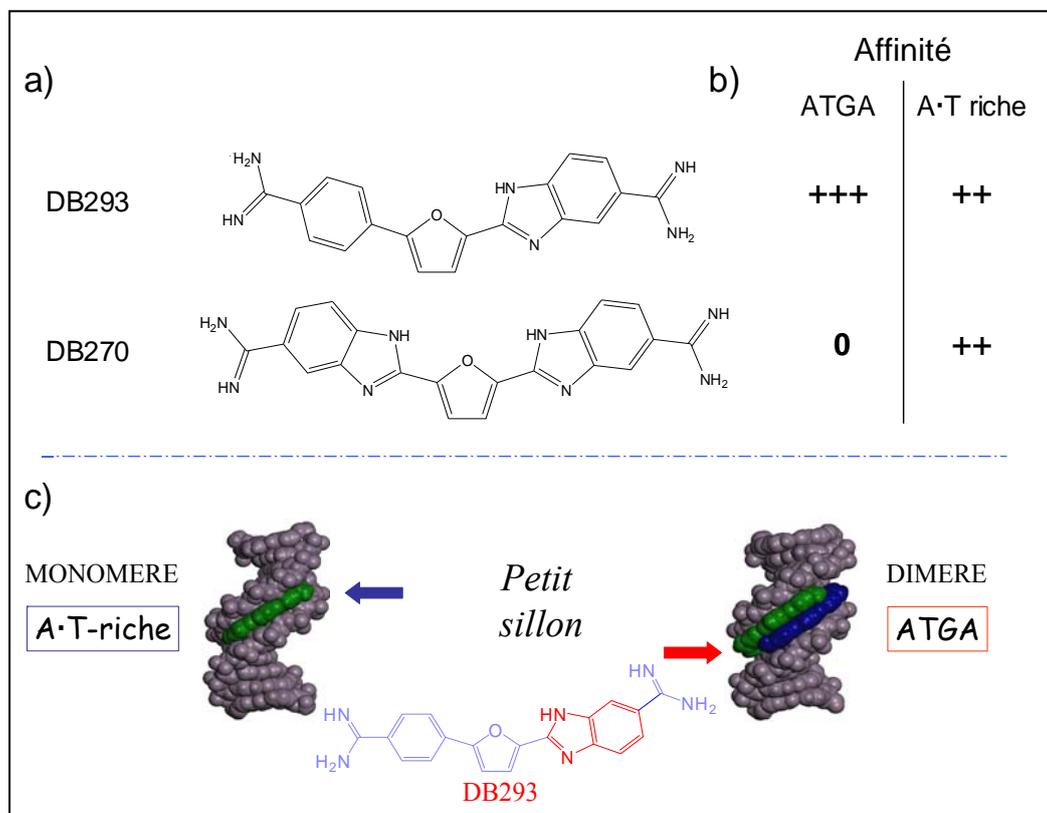


Figure 15 : Structures et affinités relatives pour l'ADN du DB293 et du DB270.

a) Structure des composés symétrique (DB293) et asymétrique (DB270). b) Capacité d'interaction relative des dérivés DB pour les séquences riches en paires de bases AT et 5'ATGA. c) Le DB293 interagit sur les sites riches en paires de bases AT sous forme de monomère. La reconnaissance au site ATGA s'effectue sous forme de dimère antiparallèle au niveau du petit sillon de l'ADN.

Le dérivé symétrique DB270 se fixe uniquement sous forme de monomère aux séquences riches en paires de bases AT au niveau du petit sillon. Jusqu'alors, aucun dérivé dicationique n'avait montré la capacité de former des dimères de manière coopérative du fait de la répulsion des charges positives et de l'encombrement stéréochimique des molécules. Des expériences d'empreintes des sites de reconnaissance utilisant différents agents de clivage de l'ADN (DNase I, radicaux libres, photosensibilité aux anthraquinone-2-sulfonate) ont montré que le DB293 possède une grande spécificité d'interaction avec la séquence 5'-ATGA et son complémentaire 5'-TCAT. L'orientation de la séquence d'ADN est importante dans cette interaction spécifique, les séquences 3'-ATGA et 3'-TCAT n'étant pas reconnues par le DB293 (Figure 15c) (Wang et al., 2000 ; Bailly et al., 2001).

D'un point de vue interaction ADN/molécule, la guanine du site 5'-ATGA forme des liaisons hydrogène avec le benzimidazole d'une première molécule dont le groupement NH pointe vers l'extérieur du petit sillon. L'adénine du brin complémentaire 5'-TCAT est quant à elle reconnue par le benzimidazole de la seconde molécule avec son groupement NH pointé vers l'intérieur du petit sillon. Les deux paires de bases AT aux extrémités de la séquence ATGA et de son complémentaire forment des liaisons hydrogène spécifiques avec les amidines terminales des deux dérivés DB293. Chaque paire de bases de la séquence de reconnaissance forme des liaisons hydrogène avec au moins une des deux molécules du dimère. Cette reconnaissance sous forme de dimère implique des modifications locales de la structure de l'ADN mises en évidence par l'augmentation du clivage par différents agents sur des sites proches de la séquence de reconnaissance du composé (Bailly et al., 2001).

L'interaction entre le DB293 et différentes séquences portant un ou deux point(s) de mutation sur le site 5'-ATGA a été étudiée. Les paires de bases centrales TG sont primordiales pour cette reconnaissance : une mutation sur ces bases annule la reconnaissance du DB293 pour l'ADN. Le DB293 reconnaît préférentiellement la séquence 5'-ATGA, néanmoins certains autres sites sont reconnus par le dimère. L'affinité relative de la molécule pour ces différents sites est présentée sur la Figure 16c (Taniou et al., 2003).

De nouveaux dérivés de la famille du DB293 avec des substitutions au niveau des amidines terminales ou au niveau du cycle benzimidazole (Figure 16) ont été synthétisés. Ces molécules ont aussi permis d'affiner les connaissances sur le ciblage des séquences nucléotidiques. Il s'est avéré que la reconnaissance spécifique sous forme de dimère suit des règles très strictes. La plupart des modifications apportées (modification du cycle benzimidazole, modification des amidines terminales) abolissent la reconnaissance du site ATGA (Figure 16a-b). Ces molécules gardent tout de même une bonne affinité pour les sites riches en paires de bases AT où ils interagissent sous forme de monomère (Wang et al., 2001).

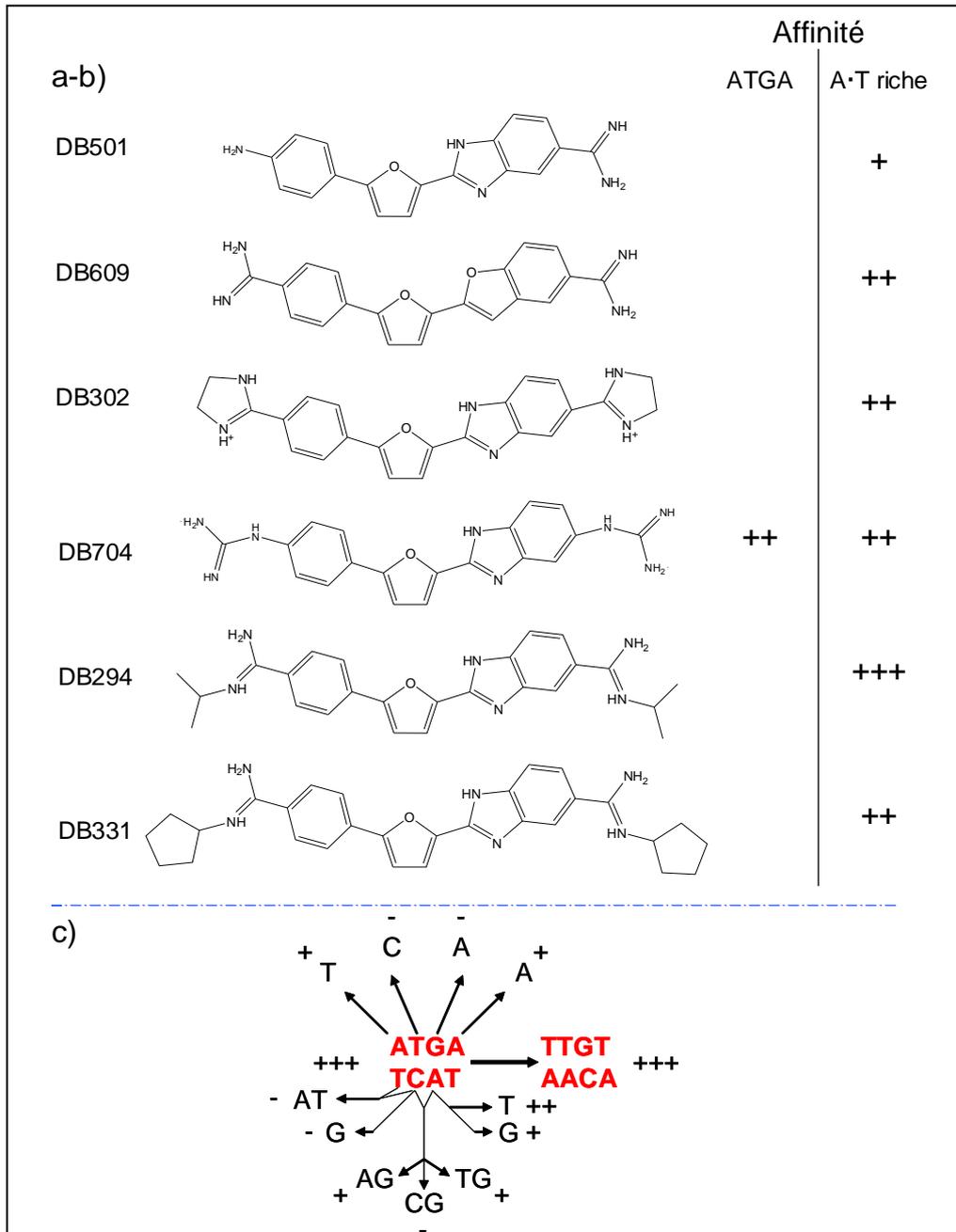


Figure 16 : Structures, affinités relatives et spécificité d'interaction de composés phényl-furane-benzimidazoles.

a-b) Structures et affinité relative des composés du type phényl-furane-benzimidazole. b) Capacité d'interaction relative des dérivés DB pour les séquences 5'ATGA sous forme de dimère et les séquences riches en paire de base AT sous forme de monomère. Séquences reconnues par le composé DB293 sous forme de dimère. Les symboles (+++), (++) et (+) correspondent respectivement aux sites de forte, moyenne et faible affinité. Le symbole (-) correspond aux sites de perte d'interaction entre le composé DB293 et l'ADN.

En résumé, le cycle benzimidazole joue un rôle primordial dans la reconnaissance de séquences ATGA tout comme les groupements amidines terminaux qui forment des liaisons π

avec le cycle phényle de l'autre membre du dimère indispensable à la stabilisation du dimère. Seul le remplacement des amidines terminales par des groupements guanidinium (DB704) permet de garder la fixation coopérative sous forme de dimère. Néanmoins, cette modification implique une diminution de l'affinité de la molécule pour l'ADN.

Comme pour le DB75 diphényl-furane, l'angle formé par la molécule est important pour la reconnaissance à l'ADN. Le DB818 diffère du DB293 par le remplacement du cycle furane par un cycle thiofène (remplacement de l'oxygène par un soufre) qui a pour conséquence d'augmenter l'angle formé par la molécule (Figure 17). Cette variation de quelques degrés abolit la fixation au site ATGA mais confère à la molécule une grande affinité pour les séquences riches en paires de bases AT, 10 fois supérieure à celle du DB293 pour cette même séquence. Ce changement d'angle modifie le positionnement des substituants portés par la molécule (amidines terminales et le cycle benzimidazole) qui peut alors former de nouvelles liaisons hydrogène avec l'ADN (Mallena et al., 2004).

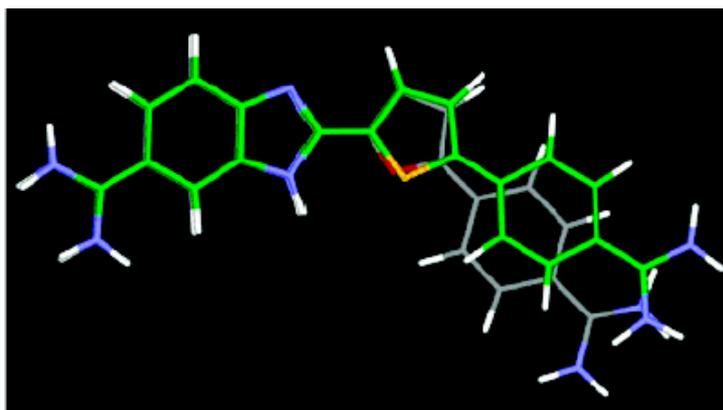


Figure 17 : Modélisation plane comparative entre le DB293 et le DB818.

Le DB818 (en vert) adopte une forme moins concave que le DB293 (en gris). La distance entre les deux atomes d'azotes portés par les amidines est de 14.4 Å pour le DB818, alors qu'elle est de 12.9 Å pour le DB293 (Mallena et al., 2004).

Pour mieux caractériser encore les interactions des DB au niveau du petit sillon, de nouveaux composés diphényl-benzimidazoles ont été synthétisés comme (i) le DB911, dont les substituants phényle et benzimidazole en position méta lui confèrent une courbure proche des autres composés DB et (ii) le composé linéaire DB921, substitué en para, qui possède une structure presque linéaire (Figure 18a). La structure du DB921 pouvait paraître non compatible avec une interaction classique au niveau du petit sillon. Il s'avère pourtant que ce produit linéaire possède une affinité proche de celle observée pour le composé courbe DB818 sur les sites riches en paires de bases AT, alors que le composé courbe DB911 possède une

affinité semblable à celle observée pour les autres composés DB. Des études de cristallographie aux rayons X et de modélisation moléculaire ont permis d'élucider le mode de fixation du DB921 à l'ADN. Le groupement amidine terminal porté par le cycle benzimidazole effectue des liaisons hydrogène avec l'atome d'oxygène de la thymine 8 (Figure 18b) directement dans le petit sillon. Le groupement NH du cycle benzimidazole effectue des liaisons hydrogène avec l'atome d'oxygène O₂ de la thymine 7. L'amidine à l'autre extrémité se situe à l'extérieur du petit sillon et ne peut pas former de liaisons hydrogène directes avec l'ADN. Par conséquent, ce groupement effectue un réseau de liaisons hydrogène indirectes par l'intermédiaire de molécules d'eau. Cette structure énergétiquement favorable confère à cette molécule une très grande affinité pour l'ADN (Miao et al., 2005).

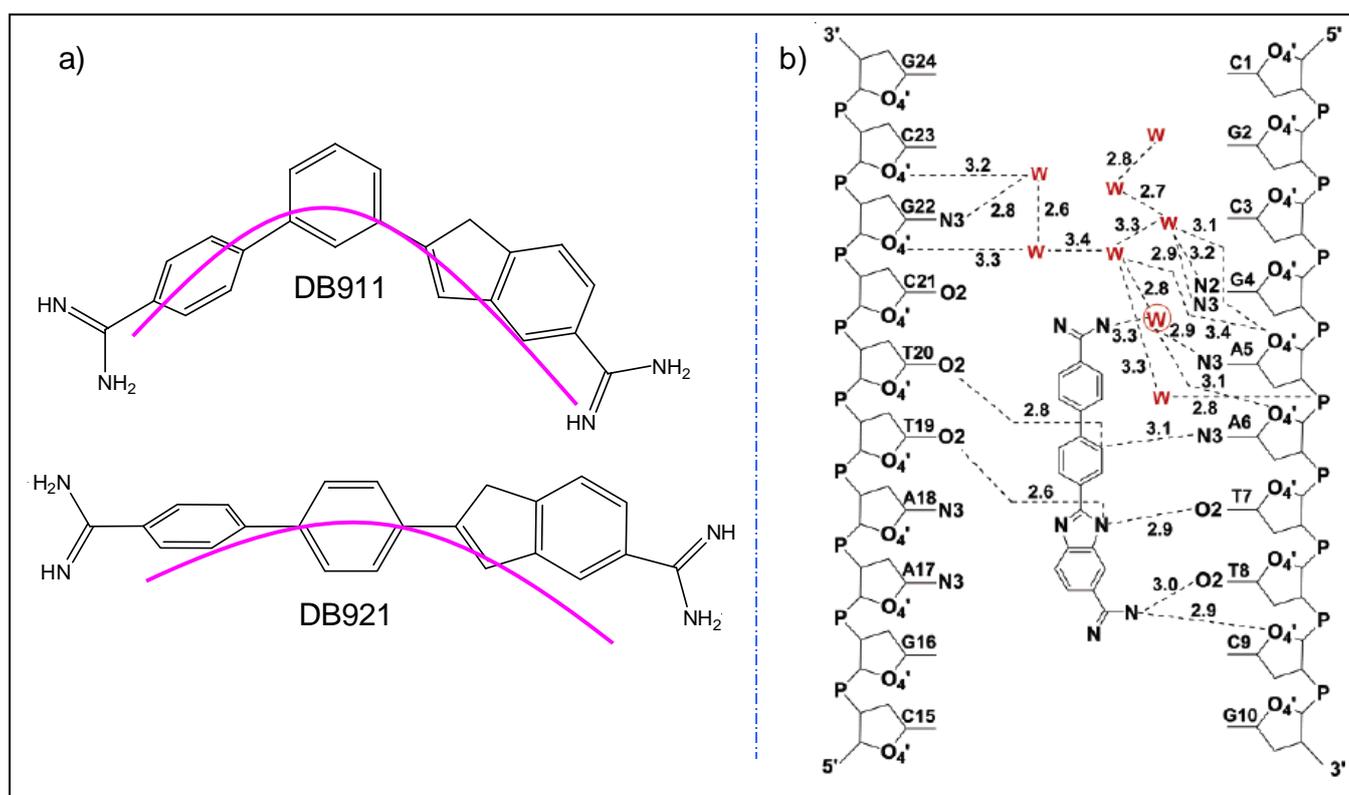


Figure 18 : Structures, angles formés par le DB911 et le DB921 et interactions à l'ADN.

a) Structure et ellipticité des composés DB911 et DB921. La courbure des molécules est représentée en rose b) Représentation schématique des liaisons hydrogène directs et indirectes (par l'intermédiaire de molécules d'eau (W)) formées par l'interaction du composé DB921 et la séquence d'ADN d(CGCGAATTCGCG)₂ (Miao et al., 2005).

De manière générale, les relations structure/affinité de ces composés phényl-furane-benzimidazoles ont permis d'établir plusieurs conclusions :

(i) Pour un grand nombre de substitutions, la reconnaissance des séquences riches en paires de bases AT semble constante.

(ii) Le développement de ces dérivés a fait évoluer les connaissances sur la reconnaissance spécifique au niveau du petit sillon de l'ADN et a permis d'établir de nouvelles règles :

- Des composés dicationiques, comme le DB293, peuvent former des dimères pour interagir sur des séquences contenant les paires de bases GC, ce qui d'un point de vue énergétique paraissait défavorable.

- De plus, même si l'angle formé par la molécule est primordial pour une bonne reconnaissance des amidines terminales à l'ADN, certaines molécules planes peuvent interagir fortement avec l'ADN par l'établissement de liaisons hydrogène par l'intermédiaire de molécules d'eau.

Ces études ont donc permis de développer des molécules (DB293) originales capables de reconnaître des sites qui contiennent les paires des bases GC et des composés qui, comme le dérivé diphenyl-furane DB75, reconnaissent des sites riches en paires de bases AT.

2. Activité biologique des composés DB

a) Activité antiparasitaire

Les pentamidines, comme le bérénil, sont classiquement utilisés dans le traitement des stades précoces d'infections parasitaires comme *Pneumocystis carinii* (pneumocystose), *Plasmodium falciparum* (paludisme), *Trypanosoma brucei rhodesiense* (maladie du sommeil) ou encore *Giardia lamblia* (giardiase) (Bouteille et al., 2003 ; Bray et al., 2003). Les composés DB, analogues des pentamidines possèdent une forte activité antiparasitaire (De Souza et al., 2004). Il existe une importante corrélation entre l'activité antiparasitaire des composés DB et leur affinité pour les séquences riches en paires de bases AT de l'ADN.

Les molécules dicationiques pénètrent majoritairement dans le parasite par transport actif par des transporteurs comme la P2 adénosine/aminopurine (Carter et al., 1995). La perte de fonction de ce transporteur est impliquée dans des mécanismes de résistance à ces agents (Matovu et al., 2003). Des expériences de microscopie de fluorescence montrent une accumulation rapide des composés diamidines dans le kinétoplaste (Wilson et al., 2005). Cet organelle sert d'unique mitochondrie du microorganisme et joue un rôle crucial dans sa survie (Figure 19a).

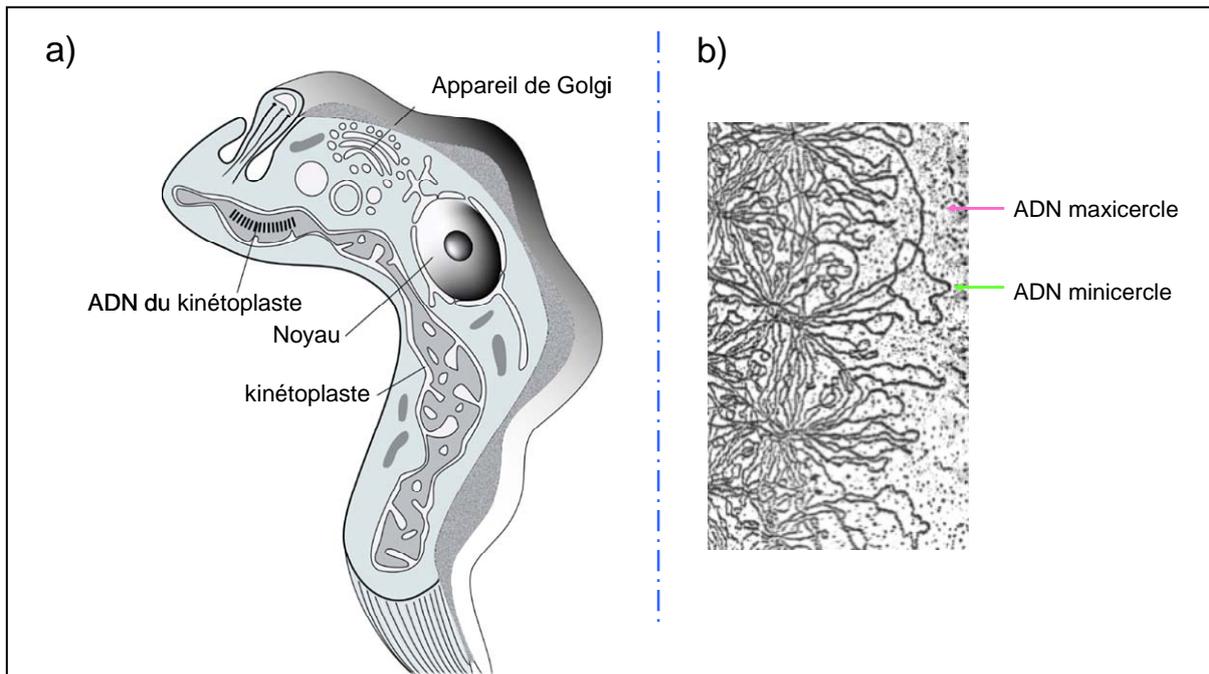


Figure 19 : Parasite à kinétoplaste.

a) Schéma d'un parasite à kinétoplaste (www.itg.be/.../03_African_trypanosomiasisp4.htm). b) Photographie de microscopie électronique montrant l'enchevêtrement des ADN maxicercles et minicercles du kinétoplaste (www.bio.umass.edu/micro/faculty/klingbeil.html).

L'ADN du kinétoplaste est constitué d'un réseau de centaines de molécules d'ADN circulaires topologiquement liées de deux types (Figure 19b) : un petit nombre de maxicercles de 23 kb de séquences homologues aux ADN mitochondriaux des eucaryotes et un grand nombre de minicercles (1kb) de séquences hétérogènes. Les maxicercles codent pour les ARN ribosomiaux et pour les protéines mitochondriales nécessaires à la synthèse d'énergie (Simpson, 1987). Les minicercles codent pour des ARN guides (ARNg) indispensables à l'édition des ARN messagers issus des maxicercles (Blanc et al., 2002). Les ADN minicercles possèdent de nombreuses régions non codantes riches en paires de bases AT qui induisent une certaine topologie à l'ADN. Les molécules dicationiques ciblent spécifiquement ces séquences au niveau du petit sillon. L'interaction des composés DB sur ces ADN circulaires modifie leurs caractéristiques topologiques. Ces contraintes vont bloquer l'accès à certaines protéines : topoisomérases, polymérases et nucléases, indispensables au maintien de l'organite et à la survie du parasite (Chen et al., 1993 ; Beerman et al., 1992).

Le mécanisme d'action des dérivés dicationiques passe par la destruction du kinétoplaste qui intervient après 24 h de traitement. Une localisation secondaire du composé est observée dans le noyau et dans des vésicules acides appelés acidocalcisomes, impliqués dans la séquestration de molécules cationiques et phosphatées (Figure 20).

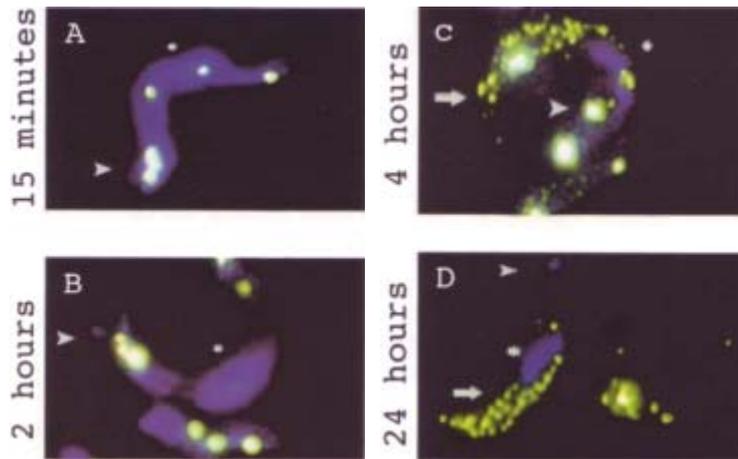


Figure 20 : Microscopie de fluorescence du trypanosome *b. brucei* chez une souris infectée après injection de 7.5 µmol/Kg de dérivé DB75.

Les différents organelles sont anotés comme suit : noyaux (*) kinétoplaste (►) acidocalcisomes (→). On note une accumulation rapide du composé dans le kinétoplaste. Après 4H de traitement, une fluorescence secondaire apparaît dans les vésicules acides. L'intensité de fluorescence dans le kinétoplaste disparaît au bout de 24 h de traitement, ce qui correspond à la destruction du kinétoplaste (figure tirée de Wilson et al., 2005).

Les dérivés pentamidines sont utilisés dans le traitement des stades précoces de la parasitémie. Le DB289, prodrogue du DB75 est actuellement en essai clinique de phase 2 dans le traitement des à *Pneumocystis carinii*. Cette prodrogue possède une bonne biodisponibilité due à l'apport de substituant O-méthoxyl sur les amidines terminales. Ces substituants permettent de masquer les charges positives du DB75. Le DB289 est métabolisé dans la cellule par les cytochromes P450 (permettant la O-deméthoxylation) et cytochromes b5 (permettant une déhydroxylation réductrice) (Figure 21) (Wilson et al., 2005).

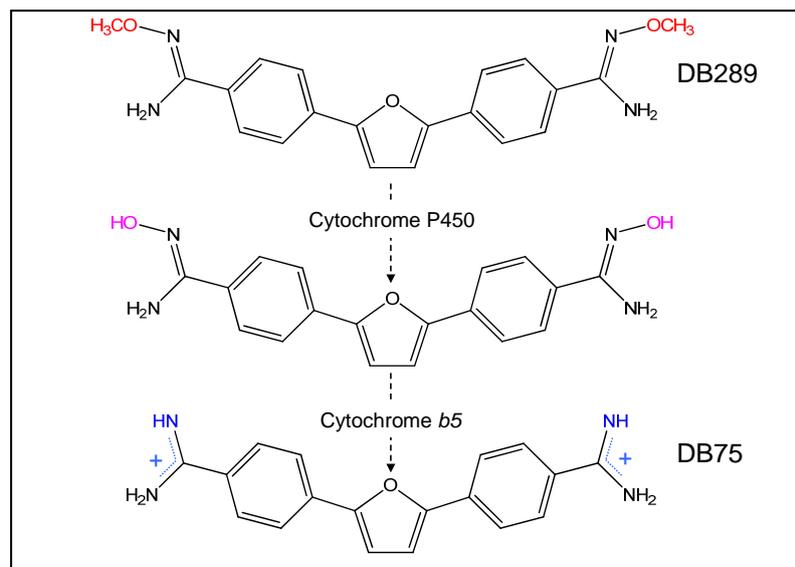


Figure 21 : Métabolisme du DB289 dans la cellule.

b) Activité cellulaire

Des tests de cytotoxicité ont mis en évidence l'activité cytotoxique de certains composés DB de type diphényl-furane sur des lignées cellulaires tumorales. Cette toxicité serait liée à leur affinité pour les séquences riches en paires de bases AT de l'ADN (Neidle et al., 1997). Mais, pour interagir avec l'ADN dans la cellule, les composés doivent traverser la membrane cytoplasmique et pénétrer dans le noyau. L'étude de la distribution cellulaire des composés DB a été effectuée par microscopies de fluorescence et confocale. Contrairement aux lexitropsines, les dérivés DB possèdent une fluorescence intrinsèque avec une longueur d'onde d'excitation située en général entre 360 et 380 nm pour une *longueur d'onde* d'émission située entre 450 et 500 nm. Même si le mécanisme par lequel ces molécules très hydrophiles pénètrent dans la cellule est encore mal connu, il a été mis en évidence une accumulation rapide des dérivés DB dans la cellule, mais le mécanisme d'internalisation n'est pas encore élucidé. Un simple phénomène de diffusion est à exclure, il existe très certainement des transporteurs spécifiques comme chez le parasite pour le transport des pentamidines (transporteur à adénosine P2) (de Koning et al., 2001). De plus, des transports actifs permettent à certains composés de type polyamines (putrescine, spermidine, spermine) de pénétrer dans la cellule (Satriano et al., 2001).

A l'inverse des polyamines, il est apparu que des règles précises gouvernaient la localisation nucléaire ou cytoplasmique des composés DB. La distribution cellulaire des composés DB est variable : (i) les composés DB purement nucléaires, (ii) les composés DB strictement cytoplasmiques et (iii) les composés DB qui se localisent à la fois dans le noyau et dans le cytoplasme (Figure 22).

Les groupements amidines terminaux sont ainsi des éléments cruciaux pour la distribution des composés DB dans la cellule comme ils jouent aussi un rôle dans la reconnaissance de séquence spécifique de l'ADN. En effet, dans la série diphényl-furanes, l'ajout de groupements aromatiques sur ces amidines terminales confère une localisation cytoplasmique (Figure 22a) à ces composés, à l'instar des composés DB569, DB613 et DB667 (Lansiaux et al., 2002a).

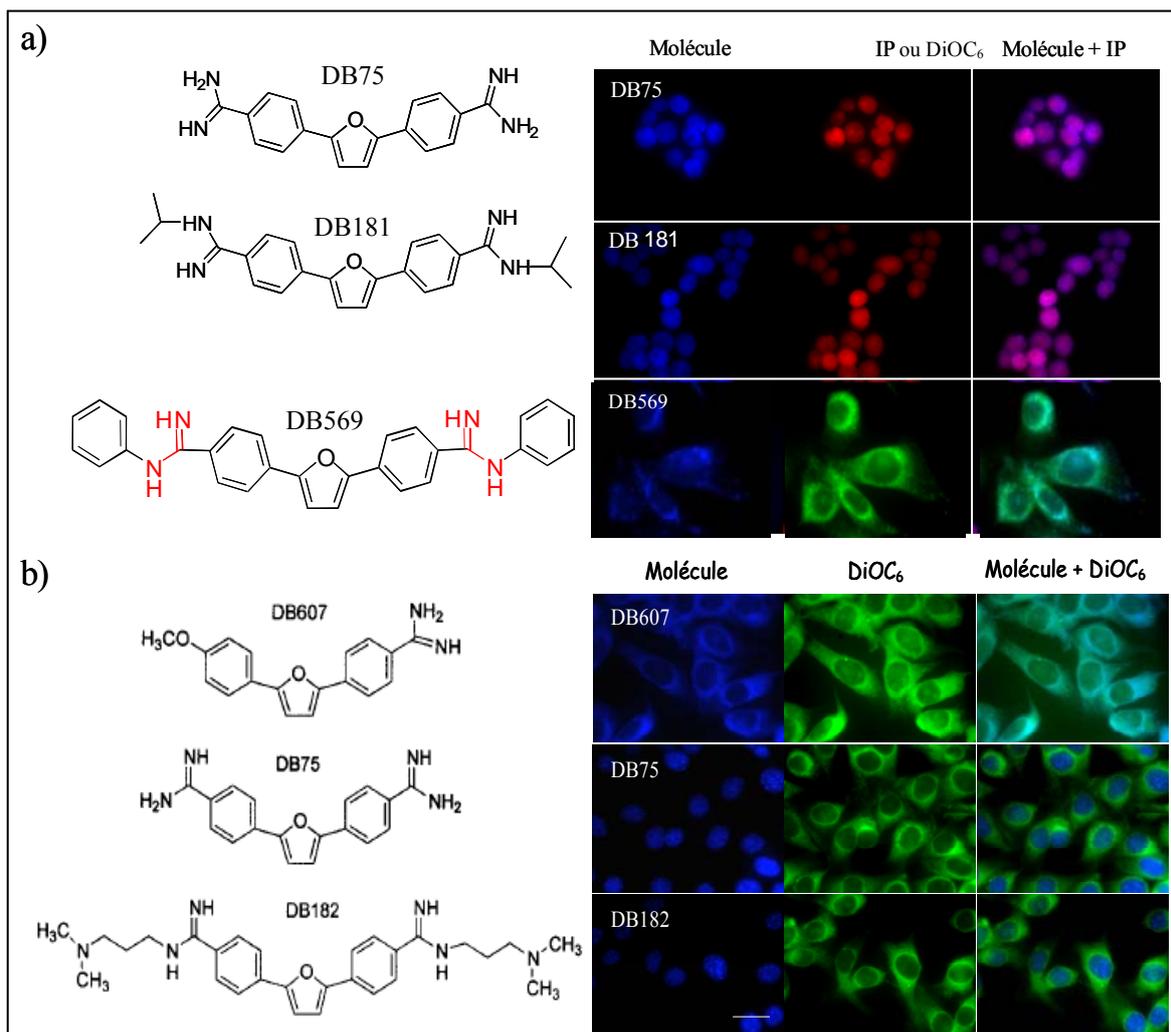


Figure 22 : Distribution intracellulaire de dérivés diphenyl-furanes.

Influence des a) substituants et b) du nombre de charges portés par les amidines sur la distribution des composés diphenyl-furanes. A gauche : a) structure chimique des composés qui portent des groupements amidines aromatiques ou non b) structure chimique des composés mono- (DB607), di- (DB75) ou tétracationiques (DB182). A droite : Localisation intracellulaire des composés visualisée en microscopie de fluorescence. Image de microscopie confocale (x63) sur cellules de mélanome B16. En bleu : fluorescence intrinsèque des composés DB. En vert : fluorescence du DiOC₆, marqueur cytoplasmique. Image de droite: superposition (Lansiaux et al., 2002a ; 2002b).

De plus, le nombre de charge porté par ces groupements influent aussi sur la distribution de ces composés. Ainsi, les composés monocationiques (DB607) pénètrent moins efficacement dans la cellule et possèdent une localisation cytoplasmique. Quant aux composés polycationiques, ils pénètrent efficacement dans la cellule et se localisent dans le noyau (Figure 22b) (Lansiaux et al., 2002b).

La localisation nucléaire des dérivés DB est dépendante de leur affinité pour l'ADN. La pré-incubation des cellules avec l'agent intercalant daunomycine a mis en évidence une

délocalisation du dérivé DB75 dans le cytoplasme. En effet, la daunomycine interagit avec l'ADN et empêche l'interaction du complexe DB75/ADN qui va alors se relocaliser dans le cytoplasme (Figure 23). Le phénomène responsable de la localisation cytoplasmique de certains dérivés DB est encore mal connu. Cependant des expériences de microscopie confocale ont montré une co-localisation des dérivés DB cytoplasmiques avec des fluorophores spécifiques de la mitochondrie. Il existe très certainement une compétition entre l'affinité des composés DB pour l'ADN nucléaire et l'électronégativité relative de la mitochondrie.

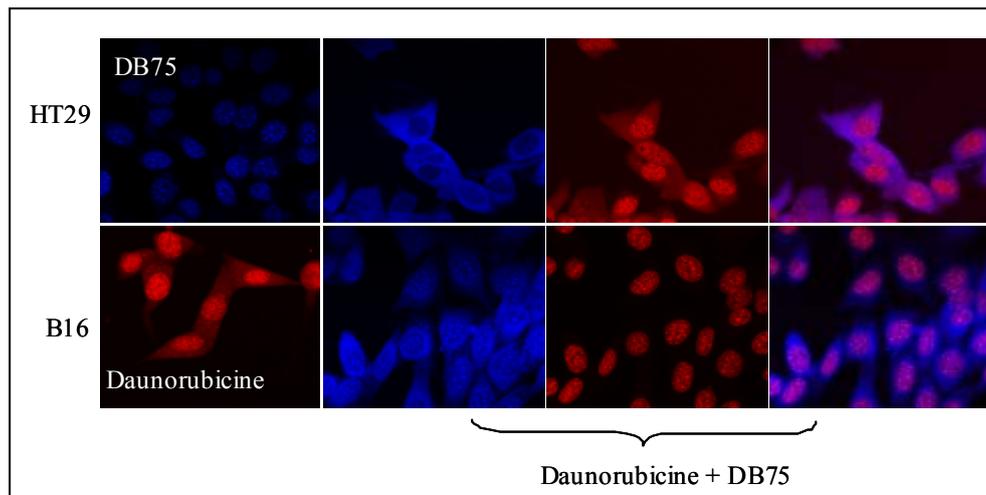


Figure 23 : Compétition entre la daunorubicine et le DB75.

Visualisation en microscopie de fluorescence de la compétition nucléaire entre le DB75 et la daunorubicine sur lignées HT29 et B16 (d'après Lansiaux et al., 2002 b).

Si, à ce jour, la fixation de ces molécules à l'ADN est relativement bien comprise et documentée par des approches biochimiques, peu d'informations sont disponibles quant aux effets de ces molécules sur la machinerie protéique nucléaire qui régule l'utilisation du matériel génétique. Connaître ces effets permettrait d'affiner la spécificité d'action de ces molécules et de développer des molécules plus sélectives.

V. Modèles d'études et leur implication dans le cancer

Dès les années 1970 et la découverte des premiers oncogènes, les facteurs de transcription ont été identifiés comme étant des pièces essentielles dans les processus de cancérogenèse. Six des premiers oncogènes décrits étaient des facteurs de transcription. Leur dysfonctionnement est à l'origine de nombreuses pathologies. C'est notamment le cas de la protéine c-Myc (cellular myelocytomatosis), dont la surexpression est retrouvée dans de nombreuses cellules transformées (Cole et McMahon, 1999). Les facteurs de transcription Fos et Jun sont eux aussi des oncogènes. Ils sont capables d'interagir en homodimère (Jun/Jun) et en hétérodimère (Fos/Jun). La surexpression de ces protéines aboutit à une croissance cellulaire incontrôlée (Jonhson et Nakamura, 2007). Vu le nombre important de facteurs de transcription impliqués dans de tels phénomènes, seuls nos modèles d'études sont présentés par la suite.

A. HoxA9 et les leucémies

1. Les gènes *hox* : organisation et fonction

a) Rôle dans le développement

Les gènes *hox* (*homeobox*) sont très conservés au cours de l'évolution et proviennent de la multiplication d'un cluster ancestral (Duboule, 2007). Ces gènes sont impliqués dans le développement embryonnaire et dans l'identification des structures le long de l'axe antéro-postérieur de l'embryon (Daftary et Taylor, 2006). Cette fonction cruciale dans la morphogenèse de l'embryon est intimement liée à la répartition de ces gènes le long des chromosomes. Ainsi, chez la drosophile les gènes situés côté 3'- sont exprimés dans des stades précoces du développement, alors que les gènes situés en 5'- sont exprimés dans des phases tardives (Mlodzik *et al.*, 1988). Ces gènes permettent le développement et la mise en place de nombreux segments du corps de la drosophile. L'identité des segments du corps est déterminée par la combinaison de l'expression des gènes *hox*. Chez les vertébrés, ces clusters montrent un arrangement spatial et temporel identique. Il existe au moins 39 protéines appartenant à la famille Hox divisées chez l'homme en 4 clusters A, B, C et D respectivement localisés sur les chromosomes 7, 17 12 et 2. Chaque cluster est composé de 9 à 11 gènes (Figure 24). Les gènes qui portent un numéro identique dans les différents clusters sont appelés gènes paralogues et sont exprimés en parallèle. Chez l'homme, la perte de fonction d'un gène peut être compensée par ces paralogues. De plus il existe une certaine redondance fonctionnelle entre les gènes adjacents d'un même cluster. Cependant, une mutation double

ou triple de gènes paralogues résulte en une transformation sévère. Ces redondances de fonctions permettent une compensation entre les gènes *hox* et montrent leur importance dans le développement. Les gènes *hox* interviennent aussi dans la mise en place du système reproductif (Hsieh-Li *et al.*, 1995). Il est assez difficile de déterminer le rôle de chaque protéine Hox car elles interagissent souvent en combinaison, à des concentrations différentes sur des séquences légèrement différentes et avec des affinités pour l'ADN variables. Toutes ces caractéristiques composent le *code Hox* (Pellerin *et al.*, 1994).

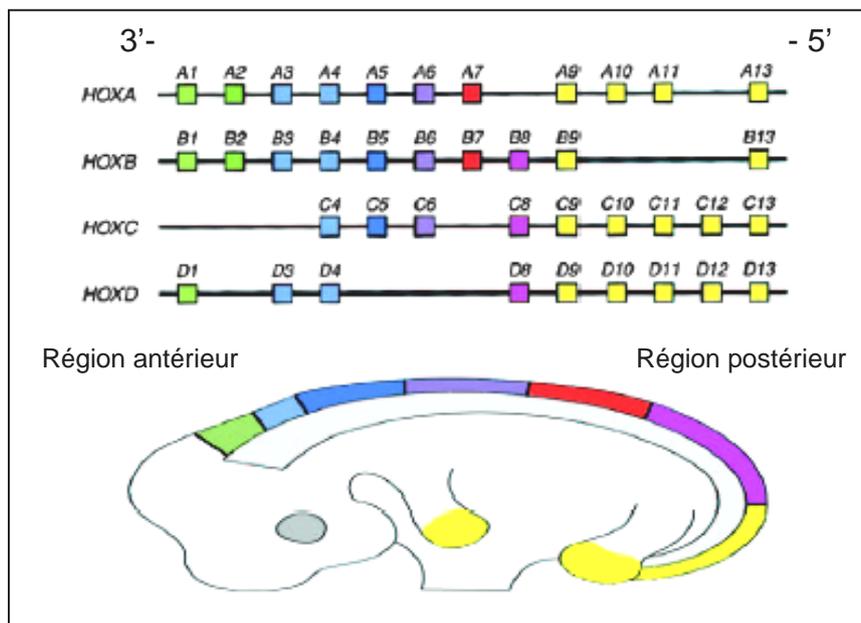


Figure 24: Représentation schématique de l'organisation génomique des clusters des gènes *hox* chez les mammifères et paterne d'expression embryogénique.

Les gènes *hox* situés coté 3' sont exprimés dans les régions antérieures de l'axe du corps et les gènes situés en 5' sont exprimés dans des régions plus postérieures. Les domaines colorés représentent les limites d'expression des gènes (tiré de Daftary *et al.*, 2006).

b) Rôle des gènes *hox* chez l'adulte

(1) Maintien de la plasticité

Chez l'adulte, l'expression des gènes *hox* fait encore l'objet de multiples investigations. Ces gènes permettraient le maintien de la plasticité de certains tissus ou organes qui continuent d'évoluer comme l'appareil reproducteur féminin (Daftary *et al.*, 2006).

(2) L'hématopoïèse

Les gènes *hox* jouent un rôle fondamental dans l'hématopoïèse. De nombreux gènes *hox* sont exprimés dans les cellules hématopoïétiques (Figure 25). Les différents types cellulaires qui constituent le système hématopoïétique subissent un renouvellement perpétuel. Les cellules souches de la moelle osseuse se multiplient et se différencient pour donner les cellules retrouvées dans le sang. Les gènes *hox* présentent un profil d'expression important dans ce système. Ils sont exprimés en fonction du type cellulaire. Ainsi, les cellules CD34+ expriment les gènes des clusters A et B, alors que les cellules CD34- n'expriment pas ces clusters. (Sauvageau *et al.*, 1994). Certaines protéines Hox se retrouvent uniquement dans un type spécifique comme HoxC4 dans les lymphocytes. D'autres se retrouvent dans tous les types cellulaires comme HoxC8, HoxB9 et HoxA7. Les gènes appartenant au cluster D ne sont pas exprimés dans les cellules hématopoïétiques.

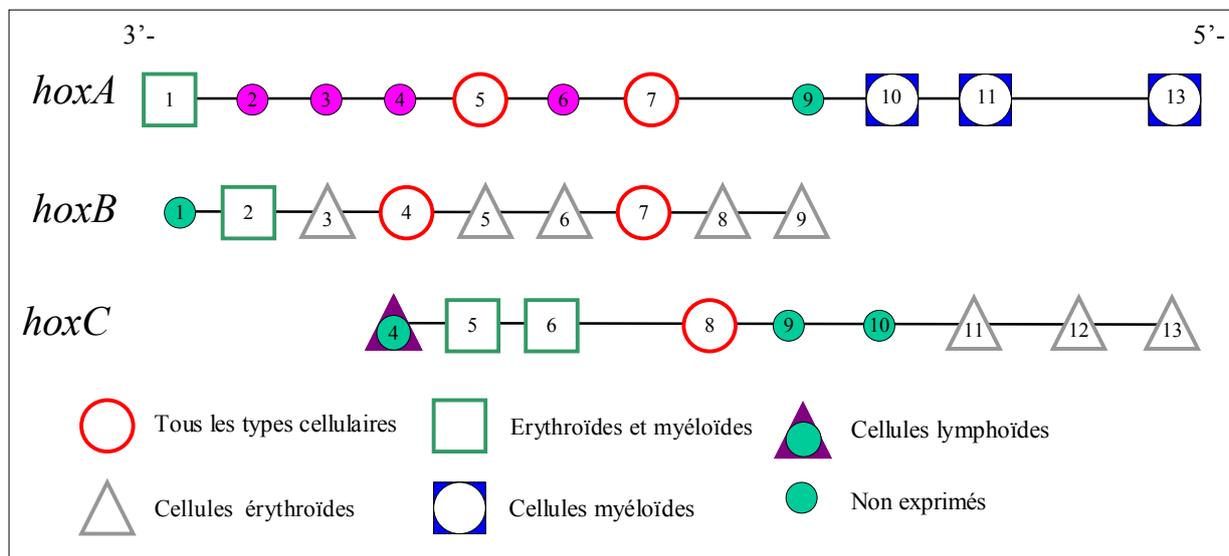


Figure 25 : Expression des gènes *hox* dans les cellules hématopoïétiques.

Les clusters A, B et C contiennent tous les trois des protéines retrouvées dans les différents types cellulaires hématopoïétiques, à l'inverse du cluster D. Chaque type cellulaire est représenté par les symboles différents précisés en dessous (d'après Lawrence et Largman, 1992).

L'expression des gènes *hox* n'est pas uniquement corrélée au type cellulaire mais aussi au stade de différenciation des cellules sanguines. Une expression temporelle liée à la topographie génomique est retrouvée. Ainsi, les gènes situés du côté 3' sont activés dans des stades plus précoces du développement hématopoïétique que les gènes situés en 5'. Par exemple, HoxB2 est exprimé dans les cellules pluripotentes progénitrices alors que HoxB7 ne l'est pas (Lawrence *et al.*, 1992). Il est assez difficile d'établir le rôle des protéines Hox dans la différenciation des cellules hématopoïétiques normales à cause de leur faible taux

d'expression. Ceci est néanmoins possible pour certaines protéines, ainsi il a été montré que la répression de HoxB6 permet la maturation des cellules érythroïdes (Mathews *et al.*, 1991). Certaines protéines Hox ont des effets antagonistes entre elles ou de manière dépendante des co-facteurs. Ainsi, HoxA10 réprime l'expression du gène *cybb* codant pour la glycoprotéine gp91^{phox} alors que HoxA9 l'active en collaboration avec Pbx (*Pre-B cell leukemia homeobox*) ou la réprime en collaboration avec Meis-1 (*Myeloid ecotropic viral integration site 1 homolog*) dans des cellules différenciées. Le gène *cybb* (*cytochrome b-245 beta polypeptide*) est un acteur important dans la différenciation des cellules myéloïdes. Le contrôle de ces gènes par des protéines Hox témoigne de leur implication dans la myélogenèse (Bei *et al.*, 2005). La surexpression de la protéine HoxA9 est d'ailleurs associée à un phénomène d'expansion des cellules progénitrices.

2. Régulation de l'expression des gènes hox

Chez l'embryon comme chez l'adulte, les gènes *hox* sont régulés par de nombreuses hormones qui agissent par l'intermédiaire de leurs récepteurs. Cette régulation est contrôlée par les acides rétinoïques, les oestrogènes, la progestérone et la 1,25-dihydroxycholecalciferol (métabolite de la vitamine D).

Ainsi, la séquence RARE (*retinoic acid receptor element*), reconnue par les facteurs de transcription RAR (*retinoic acid receptor*) et RXR (*retinoic X receptor*), activée par les acides rétinoïques se retrouve dans la zone promotrice des gènes *hoxa1*, *hoxb1* et *hoxb4* (Langston *et al.*, 1992 ; Marshall *et al.*, 1994). La régulation de l'expression des gènes par les acides rétinoïques reste toutefois limitée aux gènes situés sur la partie 3' des chromosomes et donc précocement exprimés dans l'embryogenèse.

Les oestrogènes régulent l'expression de gènes situés côté 5' comme *hoxa9*, *hoxa10* et *hoxa11* exprimés dans les domaines postérieurs de l'axe du corps (Block *et al.*, 2000).

La vitamine D permet la régulation de certains gènes via l'élément de réponse VDRE (*Vitamine D Response Element*) présent dans la zone promotrice de *hoxa10* dans les cellules hématopoïétiques et de l'endomètre (Du *et al.*, 2005).

3. L'homéodomaine

a) Structure et reconnaissance spécifique

Au niveau structural, toutes les protéines Hox possèdent un domaine très conservé de 183 paires de bases connu sous le nom d'*homéobox*. Cette région code pour un domaine de 61 acides aminés appelé homéodomaine. Il adopte une structure HTH qui interagit avec la séquence d'ADN minimale 5'-TANT-3'. La troisième paire de bases est dépendante de la protéine Hox (Laughon *et al.*, 1991). Des expériences de retard en gel ont mis en évidence l'importance des bases flanquantes 3' - dans la reconnaissance du site consensus (Catron *et al.*, 1993). L'homéodomaine est composé de trois hélices α dont les hélices 2 et 3 composent le motif HTH et d'un bras N-terminal. Des études de RMN (Otting *et al.*, 1990) et de cristallographie aux rayons X (Kissinger *et al.*, 1990) montrent que la structure de l'homéodomaine permet une bonne orientation de l'hélice de reconnaissance (hélice 3) dans le grand sillon de l'ADN. Le bras amino-terminal interagit avec le petit sillon adjacent en effectuant des liaisons hydrogène avec les bases et les sucres (Figure 26). Parmi les protéines Hox, la région la plus conservée de l'homéodomaine est l'hélice de reconnaissance. Pourtant, les protéines Hox ne reconnaissent pas toutes des séquences identiques. L'élément déterminant discriminant la reconnaissance entre les différentes protéines fut révélé par des expériences d'empreintes à la DNaseI couplées à des expériences d'interférence de méthylation. La méthylation de l'adénine en position N3 modifie la base au niveau du petit sillon de l'ADN, alors que la méthylation en N7 des guanines modifie la base au niveau du grand sillon. Les méthylations au niveau du petit sillon de l'ADN inhibent l'interaction entre la protéine à homéodomaine et l'ADN alors que les méthylations au niveau du grand sillon (N7 de la guanine) n'empêchent pas l'interaction (Draganescu *et al.*, 1995). L'orientation du bras amino-terminal de l'homéodomaine est donc un acteur prépondérant dans la reconnaissance spécifique des différentes protéines Hox avec l'ADN.

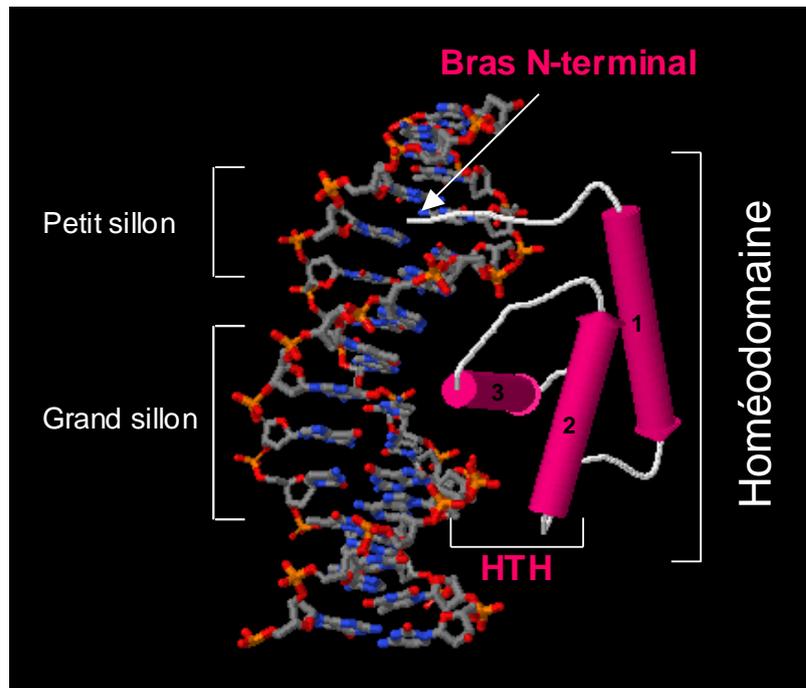


Figure 26 : Interaction entre l'homéodomaine de *Hoxa9* et l'ADN.

L'hélice de reconnaissance 3 du motif HTH interagit avec le grand sillon de la double hélice d'ADN. Cette structure est stabilisée par la seconde hélice du motif HTH ainsi que par l'hélice 1 de l'homéodomaine. Côté N-terminal de l'homéodomaine, le bras interagit avec le petit sillon de l'ADN. (modélisation effectuée à partir du site : <http://arapaho.nsuok.edu/~biology/Tutorials/DNAbinding.htm>).

L'homéodomaine et la formation d'hétérodimère

Comme les protéines Hox se lient à l'ADN avec une faible spécificité, il a très vite été proposé qu'elles agissaient avec des cofacteurs et que seuls les hétéro-dimères seraient dotés d'une reconnaissance suffisante pour rendre compte de leur fonction *in vitro*. Il est aujourd'hui clairement démontré que les protéines Hox se lient à l'ADN de façon coopérative avec des cofacteurs de la famille TALE (*three amino acid loop extension*) (Chang *et al.*, 1995). Les protéines TALE possèdent un homéodomaine particulier qui présente trois acides aminés supplémentaires entre les résidus 23 et 24 de la boucle séparant les hélices 1 et 2. Ces résidus permettent l'interaction protéine/protéine avec les facteurs de transcription Hox. Ainsi, le complexe HoxA9/Pbx se lie au site 5'-ATGATTTA-3' et possède une spécificité de liaison accrue comparé au monomère Hox.

La protéine HoxA9 reconnaît la protéine Pbx1 par un hexapeptide situé du côté N-terminal de l'homéodomaine. Ce dernier aussi appelé PIM (*Pbx Interaction Motif*) possède une structure en coude β capable d'interagir avec une poche appelée *Hexapeptide Binding Pocket* (HBP) de la protéine Pbx (Chang *et al.*, 1995 ; LaRonde-LeBlanc *et al.*, 2003). Cette poche est formée par trois acides aminés situés entre l'hélice 1 et 2 de la protéine Pbx (constituée des acides aminés 23a, sérine 23b et asparagine 23c) et de la partie C-terminale de l'hélice 3 (Figure 27).

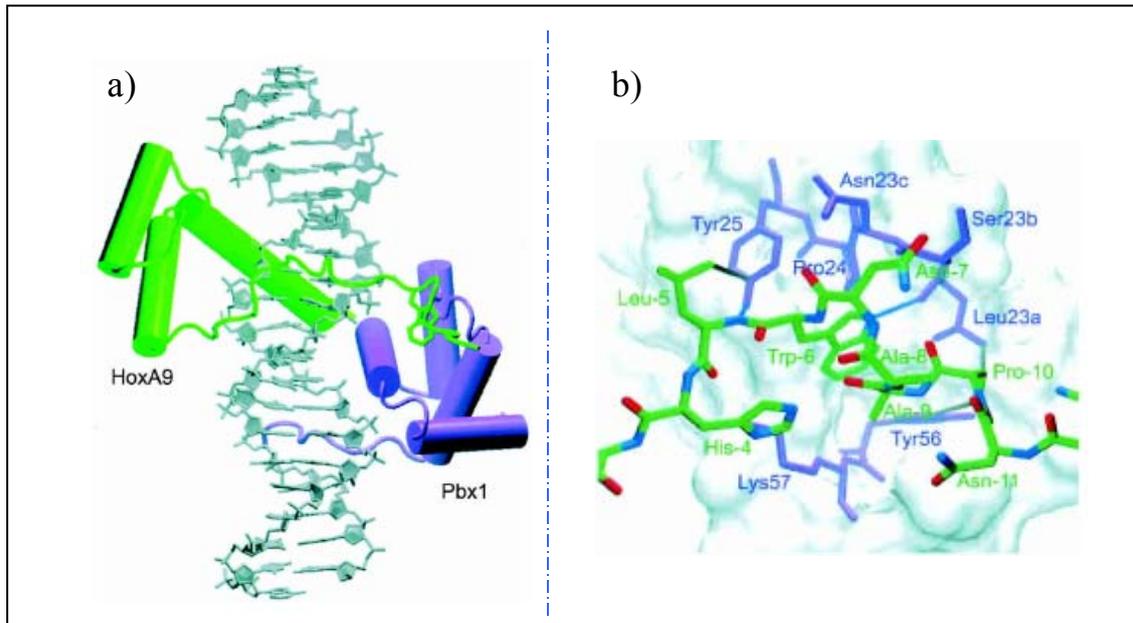


Figure 27 : Structure du complexe ternaire HoxA9/PBX/ADN sur la séquence ATGATTTA.

a) Les deux protéines à homéodomaine HoxA9 et Pbx1 interagissent ensemble pour reconnaître la séquence ATGATTTAcgac. b) Les interactions entre l'hexapeptide de HoxA9 (en vert) et la poche HBP de Pbx1 (en bleu) mettent en jeu des liaisons faibles (d'après LaRonde-Leblanc *et al.*, 2003).

Les protéines paralogues Hox9 à 13 ont la capacité de former des dimères avec les protéines Meis qui appartiennent aussi à la famille TALE. Cependant, l'interaction entre HoxA9 et Meis1 ne fait pas intervenir le domaine PIM de la protéine HoxA9. Des expériences de mutagenèse ont mis en évidence que l'interaction entre HoxA9 et les protéines Meis ne fait pas intervenir l'homéodomaine. Les contacts entre les deux protéines s'effectuent via des interactions entre la partie N-terminale de la protéine HoxA9 et la région C-terminale de l'homéodomaine des protéines Meis (Shen *et al.*, 1999).

4. Rôle dans la leucémogénèse

Le facteur de transcription HoxA9 joue un rôle central dans différents types de leucémies comme les leucémies aiguës myéloïdes (LAM), les leucémies aiguës lymphoïdes-T (LAL-T), les syndrômes myélodysplasiques (SMD). C'est pour cela qu'il représente une cible antitumorale très intéressante. Sa surexpression est souvent retrouvée chez les patients en échec thérapeutique et elle est corrélée à un mauvais pronostic (Golub *et al.*, 1999). Le rôle précis du facteur de transcription HoxA9 dans la leucémogénèse n'est pas clairement défini. Ceci est notamment dû aux phénomènes de compensation de fonctions qui existent entre les

différentes protéines de la famille Hox. Ainsi, la modulation de l'expression d'une de ces protéines peut passer inaperçue par l'activation ou la surexpression d'autres membres de la famille Hox. La multiplication des modèles d'étude engendre parfois des résultats et des conclusions différentes voir opposées.

Le facteur de transcription HoxA9 est normalement exprimé dans les stades intermédiaires et assez tardifs de la différenciation en cellules myéloïdes (Daftary et al., 2006). HoxA9 est exprimé dans les cellules CD34+ et il est sous-exprimé pendant la différenciation des autres types cellulaires.

Pour déterminer le rôle de HoxA9 dans la leucémogénèse, de nombreuses études *in vitro* et *in vivo* sont fréquemment menées. Dans la plupart des cas, la protéine HoxA9 seule ou avec ses cofacteurs est surexprimée dans différents types cellulaires. Par la suite, différents types d'études ont été mises en œuvre : i) l'identification des gènes cibles de HoxA9 seul ou avec ses cofacteurs a été étudiée par des approches de microarrays ; ii) le potentiel de HoxA9 à immortaliser et/ou à permettre la prolifération des cellules a été évalué par des tests clonogéniques ; iii) le potentiel leucémogène de HoxA9 a été déterminé par implantation de ces cellules chez la souris. Cependant, la diversité des modèles cellulaires induit parfois des résultats aux conclusions divergentes.

a) *Le monomère HoxA9*

Des études de microarrays sur lignées myéloïdes (U937), érythroïdes (K562) et lymphoïdes (Jurkat) dans lesquelles la protéine HoxA9 a été surexprimée par transfection, montrent des profils d'expression qui présentent des changements plus importants dans les cellules myéloïdes par rapport aux cellules lymphoïdes (Dorsam *et al.*, 2004). Ce résultat suggère que les effets de HoxA9 dépendent du contexte cellulaire. Ainsi, la surexpression de HoxA9 dans les cellules myéloïdes réprime l'expression d'anti-oncogènes comme *jun-b* ou *lyn* alors qu'elle permet la surexpression d'oncogènes comme *c-myb* et des membres de la famille ras (*eps8* et *rap1B*). La surexpression de ces gènes permet la transformation de cellules en culture (Lipsick *et al.*, 1999 ; Altschuler *et al.*, 1998). D'autres protéines cibles de HoxA9 sont associées à la leucémogénèse comme HoxB2 et Pim1.

De nombreuses études montrent que la surexpression de HoxA9 induit des LAM après une longue période de latence. La protéine HoxA9 est capable de bloquer la différenciation des cellules progénitrices murines en culture stimulées au GM-CSF (*granulocyte-macrophage colony-stimulating factor*) ; par contre, ces mêmes cellules ne prolifèrent pas en présence de SCF (*stem cell factor*) et se différencient en neutrophiles quand le GM-CSF est remplacé par

du G-CSF (*granulocyte colony-stimulating factor*) (Calvo *et al.*, 2000). Cette même étude a été réalisée avec le cofacteur Meis-1 (Calvo *et al.*, 2001). Tandis que Meis-1 seul est incapable d'induire des leucémies sous l'action des cytokines, HoxA9 et Meis-1 sont capables de supprimer la différenciation de ces cellules induite par le G-CSF et permettent la prolifération des cellules en culture avec du SCF. La régulation des phénomènes de différenciation et de prolifération induite par HoxA9 est donc cytokine-dépendante. Cependant, les mécanismes responsables de ces phénomènes sont encore mal connus.

b) HoxA9 et ses cofacteurs

(1) Meis-1

Meis-1 est un cofacteur de HoxA9 dans l'induction rapide de leucémies. Le caractère leucémogène reste en grande partie tributaire de HoxA9 puisque des cellules progénitrices de moelle osseuse transformées par Meis-1 seule n'induisent pas de LAM après réimplantation chez la souris. Les cellules transformées avec HoxA9 seul, sont capables d'induire des LAM au bout de 185 +/-17 jours alors que les progéniteurs transformés avec HoxA9 et Meis-1 induisent des LAM dès 54 +/-9 jours (Thorsteinsdottir *et al.*, 2001 ; Kroon *et al.*, 1998). Meis-1 est alors considéré comme un « accélérateur de leucémogénèse ».

Les gènes HoxA9 et Meis-1 seraient en partie responsables du caractère leucémogène des protéines MLL (*mixed-lineage leukemia*). Le gène MLL (11q23) est impliqué dans de nombreuses translocations et perd sa partie régulatrice carboxy-terminale, ce qui lui confère son activité transformante. Des embryons de souris qui sont déficients en protéine MLL ne sont pas viables et présentent un défaut d'expression des gènes *hox* qui se traduit par des aberrations dans l'identité des segments du corps (Yu *et al.*, 1995). Ces protéines chimères sont retrouvées dans certains cas de leucémies myéloïdes mais surtout dans les lymphoïdes (Armstrong *et al.*, 2001). Les leucémies qui présentent cette translocation possèdent un mauvais pronostic car les thérapies classiquement utilisées dans le traitement des leucémies lymphoïdes s'avèrent peu efficaces. Des expériences utilisant un gène inductible MLL montrent la surexpression de HoxA9 et de Meis-1. Le caractère transformant de MLL serait en partie dû à la surexpression de ces deux gènes par action directe sur le promoteur. Les cellules différenciées (c-Kit-, Gr1+ et Mac1+) qui surexpriment les protéines HoxA9 et Meis-1 ont la capacité de se dédifférencier (c-Kit+, Gr1- et Mac1-) (Zeisig *et al.*, 2004). Dans ce modèle MLL, la surexpression du récepteur tyrosine kinase Flt3 (*fms-like tyrosine receptor-3*) a aussi été montrée. Le groupe de Wang a d'ailleurs proposé que l'acquisition du caractère

leucémogène de HoxA9 et Meis-1 passerait par la surexpression de la protéine Flt3 (Wang *et al.*, 2005), protéine intervenant dans l'expansion des cellules leucémiques (Griffin *et al.*, 2004). Cette hypothèse a été contredite récemment par le groupe de Lavau, qui a montré que des souris Flt3^{-/-} qui sur-expriment les protéines HoxA9 et Meis-1 gardent un caractère leucémogène identique aux souris Flt3^{-/+} (Morgado *et al.*, 2007).

Les mécanismes qui gouvernent l'effet synergique des facteurs de transcription HoxA9 et Meis1 dans la leucémogénèse restent cependant mal connus.

(2) Pbx

Certains membres de la famille Pbx agissent aussi comme cofacteur de HoxA9 et Meis1 dans la cancérogénèse. Ainsi, des mutations de HoxA9 sur le site d'interaction du cofacteur Pbx2 diminuent le potentiel oncogénique de HoxA9/Meis1a dans des progéniteurs myéloïdes (Schnabel *et al.*, 2000).

La localisation nucléaire de Pbx est régulée par les protéines la famille Meis (Berthelsen *et al.*, 1999) telles que Meis-11 et PREP1. Bien que PREP1 possède une certaine homologie avec Meis-1, elle n'interagit pas avec HoxA9 mais garde la capacité de relocaliser Pbx dans le noyau. La co-expression de PREP1 et HoxA9 ne transforme pas plus rapidement des cellules de moelle osseuse de souris par rapport aux cellules qui surexpriment HoxA9 seul (Thorsteinsdottir *et al.*, 2001).

c) La translocation t(7;11) et la protéine Nup98/HoxA9

La protéine de fusion Nup98/HoxA9 a été identifiée en 1996 dans les leucémies myéloïdes. Elle implique la translocation t(7;11)(p15;p15) (Nakamura *et al.*, 1996) et est présente de façon rare mais récurrente dans les LAM et les SMD (syndromes myélo-prolifératif). Le gène *Nup98*, localisé en position 11p15, code pour une nucléoporine de 98 kDa. Elle est un constituant des pores nucléaires qui régulent le trafic nucléo-cytoplasmique des ARNs et des protéines. Nup98/HoxA9 est une protéine chimère composée du domaine amino-terminal de Nup98, conservant le domaine d'interaction protéine-protéine au niveau du pore nucléaire fusionné au domaine carboxy-terminal de HoxA9 contenant son domaine d'interaction à l'ADN.

Le mode d'action de la protéine chimérique semble complexe. Différentes hypothèses existent quant au caractère leucémogène de cette translocation. Il pourrait s'agir d'une simple augmentation de l'expression de HoxA9 dont la fonction resterait inchangée. La protéine chimère pourrait aussi réguler négativement la fixation de la protéine HoxA9. Un rôle

spécifique de Nup98 peut être envisagé, d'autant qu'un autre gène codant pour une nucléoporine (Nup214) a déjà été impliqué dans les fusions de gènes au cours de la leucémogénèse (t(6;9)(p23;q34)) (Kraemer *et al.*, 1994). D'autre part, la présence d'une protéine chimère peut rendre compte d'une localisation aberrante d'une protéine à homéodomaine dans le pore nucléaire, ce qui pourrait interférer avec la fonction normale de Nup98. Ceci serait susceptible d'aboutir à des anomalies de répartition intracellulaire de certaines protéines. Par ailleurs, la cassure du chromosome 7 qui a lieu lors de la translocation t(7;11) prive les gènes HoxA10, HoxA11 et HoxA13 de séquences de régulation en particulier d'un élément de réponse RARE qui pourrait avoir des répercussions dans la leucémie (Soulier, 1996).

La portion amino-terminale de HoxA9 tronquée lors de la translocation contient un site de répression de la transcription. Cependant, cette seule délétion ne permet pas d'expliquer le caractère leucémogène de Nup98/HoxA9 (Kroon *et al.*, 2001). Il existe deux cassures possibles au niveau du gène de la nucléoporine Nup98 au niveau d'un site qui contient des répétitions FG (Phe-Gly). Ce site est indispensable pour le contact des protéines qui subissent le transport nucléo-cytoplasmique même si sa fonction précise n'est pas définie. Des mutations sur ce site au niveau de la protéine chimère ont mis en évidence l'importance de ces sites dans l'activation de la transcription (Kasper *et al.*, 1999). Cette région répétée a la capacité d'interagir avec la protéine CBP qui recrute p300, co-activateur de la transcription (Chapitre II).

La capacité de Nup98/HoxA9 à induire des SMD pouvant aboutir à des LAM a été souvent décrite (Kroon *et al.*, 2001 ; Calvo *et al.*, 2002 ; Ghannam *et al.*, 2004 ; Takeda *et al.*, 2006). Ainsi, dans les cellules souches hématopoïétiques, l'expression de Nup98/HoxA9 entraîne un SMD qui évolue en LAM chez la souris après une période de latence de quatre mois. Cette période peut être accélérée par la co-expression de Meis1. On sait pourtant que Meis-1 se lie à HoxA9 par son domaine amino-terminal absent dans la protéine de fusion. Le phénomène impliquant Meis-1 est encore mal connu. La comparaison entre les cellules souches hématopoïétiques exprimant Nup98/HoxA9 ou HoxA9 seul suggère que l'effet prolifératif est supérieur pour les cellules Nup98/HoxA9, mais la capacité à induire des LAM est supérieure pour les cellules exprimant HoxA9. Ce résultat suggère que les caractères prolifératifs et transformants sont indépendants.

Des expériences de co-transfection entre Nup98/HoxA9 et BCR/ABL ont mis en évidence une synergie entre ces deux protéines pour induire des LAM chez la souris. Sept à 9 jours sont suffisants pour induire une LAM chez les souris exprimant les deux protéines, alors que la

protéine Nup98/HoxA9 seule induit une LAM au bout de 52+/-4 jours et que BCR/ABL seul est incapable d'induire une LAM, mais induit un SMD de cellules matures (Mayotte *et al.*, 2002). BCR/ABL (*Breakpoint cluster region/Abelson*) serait impliqué dans la prolifération et la survie sans affecter la différenciation alors que Nup98/HoxA9 conférerait des avantages de croissance hématopoïétique mais surtout bloquerait en premier lieu la différenciation (Dash *et al.*, 2001). Le remplacement de BCR/ABL par une autre protéine chimère impliquant une tyrosine kinase comme TEL/PDGFR montre les mêmes effets.

5. Régulation de la protéine HoxA9

a) Le TGF β (*transforming growth factor β*)

Même si tous les mécanismes de régulation de la protéine HoxA9 sont loin d'être connus, certains mécanismes ont été décrits. Le TGF β régule de nombreuses fonctions durant l'hématopoïèse incluant la différenciation, la prolifération et l'apoptose. Il a été montré que l'activation de Smad4 par le TGF β inhibe la capacité transformante de HoxA9. Smad4 interagit directement avec l'homéodomaine de HoxA9 ou de Nup98/HoxA9 et bloque l'activité de liaison à l'ADN. La protéine HoxA9 comme d'autres protéines Hox permet le maintien des cellules souches hématopoïétiques dans un état non différencié (Wang *et al.*, 2006).

b) Les phosphorylations

HoxA9 possède différents sites de phosphorylation. Dans l'homéodomaine, il existe un site STRK contenant la sérine 204 spécifiquement phosphorylable par la protéine kinase C (PKC). La phosphorylation de ce site empêche la fixation de HoxA9 à l'ADN. Ainsi, les cellules immortalisées avec HoxA9 et traitées au TPA (12-O-tétradécanoylphorbol-13-acétate, activateur de l'activité tyrosine kinase de la PKC) sont capables de se différencier. Au contraire, une inhibition de l'activité kinase par le bisindolylmaleamide 1 (inhibiteur spécifique de la PKC) empêche la différenciation des cellules préalablement traitées au TPA (Vijapurkar *et al.*, 2004).

c) Localisation subcellulaire de HoxA9

Des expériences de western blot sur extrait nucléaire de cellules hématopoïétiques immatures montrent une augmentation rapide du taux de HoxA9 dans le noyau après traitement par la TPO (thrombopoïétine) alors que le niveau global d'expression de HoxA9 reste identique. La

TPO active l'expression de Meis-1 et donc la formation du complexe HoxA9/Meis1. Des mutations sur le site d'interaction de Meis-1 sur la protéine HoxA9 montrent que ce site est crucial pour la localisation nucléaire de HoxA9. Il a été montré que Meis-1 bloque le signal NES (*nuclear export signal*) de Pbx chez la drosophile. Les gènes Hox contiennent deux signaux de localisation nucléaire (NLS) et des signaux NES qui peuvent être inhibés par phosphorylation (Maizel *et al.*, 1999). Meis-1 pourrait aussi réprimer ce signal en interagissant avec HoxA9 (Kirito *et al.*, 2004).

d) *Epissage alternatif de HoxA9*

Les gènes *hox* sont connus pour produire des transcrits alternatifs codant pour différentes isoformes dont le rôle physiologique est mal connu (Popovic *et al.*, 2007). L'épissage alternatif de HoxA9 peut aboutir à la protéine HoxA9T qui ne possède pas d'homéodomaine. Cette protéine tronquée est présente dans de nombreuses espèces comme chez l'homme, la poule et la souris. HoxA9T garde la capacité d'interagir avec le facteur de transcription CBP, activateur de transcription (Dintilhac *et al.*, 2004). Il pourrait donc exister une compétition entre HoxA9 et HoxA9T dans le recrutement de CBP comme protéine co-activatrice de la transcription. Cette protéine pourrait être un moyen de contrôle de l'activité de HoxA9 par la compétition dans le recrutement de la protéine CBP.

B. Les facteurs de transcription à POU domaine

Le nom de POU domaine dérive des initiales des trois premiers facteurs de transcription découverts possédant cette structure : Pit-1 (Bodner *et al.*, 1988), Oct-1 (Muller *et al.*, 1988) et Unc-86 (Finney *et al.*, 1990). Cette région bipartite unique est composée de deux motifs HTH indépendants liés par un *linker* peu conservé. Ces motifs sont appelés POU homéodomaine (POU_H) et POU spécifique (POU_S). Les gènes qui codent pour les facteurs de transcription à POU domaine sont répartis sur des chromosomes différents et ne sont pas organisés sous forme de cluster comme c'est le cas pour les gènes *hox* (Xia *et al.*, 1993).

1. Classification et fonctions

Ces gènes possèdent des fonctions importantes dans le développement et la différenciation cellulaire. A l'heure actuelle, 15 membres de cette famille ont été découverts chez les mammifères. Ils se classent en six sous-familles basées sur l'homologie de séquence du POU domaine. La classe I (Pit-1) intervient chez les mammifères dans la genèse et le maintien de trois types cellulaires endocrines différents dans la glande pituitaire antérieure (Li *et al.*, 1990). Pit-1 est responsable de l'expression des gènes de la prolactine (*PRL*), de l'hormone de croissance (*GH*, pour *Growth Hormone*) et de l'hormone stimulant la thyroïde (*TSH β* , pour *thyroid-stimulating hormone*) (Bodner *et al.*, 1988 ; Ingraham *et al.*, 1988 ; Steinfelder *et al.*, 1991). Ces gènes interviennent dans le développement et le maintien des cellules somatotropes, lactotropes et thyrotropes dans la glande pituitaire. Oct-1 (*octamer*) et Oct-2 appartiennent à la classe II et possèdent un rôle dans le développement neuroendocrinien (Andersen *et al.*, 1993). La classe III est représentée par les facteurs de transcription Brn-1, -2 et -4 (*Brain*). Brn-2 intervient dans la formation du noyau hypothalamique et de la glande pituitaire postérieure (Nakai *et al.*, 1995 ; Schonemann *et al.*, 1995). Les autres membres de la sous-famille sont importants pour le fonctionnement neuroendocrinien et le développement. Les membres du groupe IV comme Brn-3 sont importants pour le développement sensoriel, mais ne sont pas impliqués dans le système neuroendocrinien (Mc Evilly *et al.*, 1996). Les membres de la classe V comme Oct-3 et 4 sont exprimés dans les cellules souches embryonnaires et dans les phases précoces de la neurogenèse (Nichols *et al.*, 1998). Le facteur de transcription Brn-5, membre de la classe VI, est fortement exprimé dans la région hypothalamique (Andersen *et al.*, 1993) alors que RPF-1, autre membre de la sous-famille, est exprimé dans la rétine et dans diverses populations de neurones au sein de l'hypothalamus dorsal (Zhou *et al.*, 1996).

2. Le POU domaine: structure et fonctions

Parmi les facteurs de transcription, les protéines à POU domaine sont uniques de par la présence d'éléments structurellement indépendants POU_H et POU_S. Ces régions fonctionnent de façon coopérative, ce qui confère aux facteurs de transcription à POU domaine une étonnante flexibilité. Ce principe est démontré par l'interaction de Oct-1 sur différents sites consensus et notamment l'octamère 5' - ATGCAAAT pour lequel le domaine POU_S reconnaît le site ATGC et la région POU_H le site AAAT (Figure 28). L'orientation et l'espacement entre

ces deux domaines peut varier et l'habilité des protéines à POU domaine à reconnaître un site consensus dépend en partie de la taille et de la flexibilité du linker qui sépare les deux motifs HTH (Li *et al.*, 1993). La diversité des sites reconnus par les facteurs de transcription à POU domaine dépend aussi de la flexibilité importante des liaisons établies entre l'ADN et l'hélice de reconnaissance. Cette flexibilité est plus importante que celle retrouvée chez les protéines à homéo-domaines (Cleary *et al.*, 1995). Le facteur de transcription Pit-1 reconnaît de multiples sites retrouvés sur des régions promotrices de gènes exprimés dans la glande pituitaire comme le promoteur codant pour la prolactine (Bodner *et al.*, 1988) ou encore son propre promoteur qu'il auto-régule (Mc Cormick *et al.*, 1990). Les sites reconnus par Pit-1 sont variables et peuvent être du type 5'-ATGA/T ATA(A/T)(A/T)(A/T) retrouvés dans de nombreuses régions promotrices qu'il régule. Pit-1 peut interagir sous forme de monomère, d'homodimère ou d'hétérodimère, notamment avec Oct-1 (Voss *et al.*, 1991).

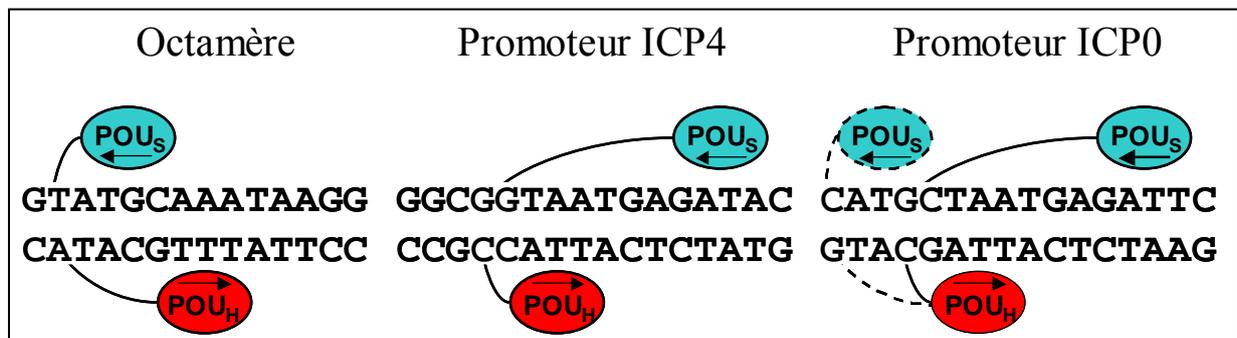


Figure 28 : Flexibilité dans la reconnaissance sur différents sites consensus par les facteurs de transcription à POU domaine : exemple de Oct-1.

Le facteur de transcription Oct-1 est capable de reconnaître différents promoteurs grâce à la grande flexibilité de reconnaissance du POU domaine. Les flèches représentent l'orientation des domaines POU_H et POU_S. Les pointillés montrent la seconde orientation possible de Oct-1 sur la séquence promotrice ICP0 (Phillips *et al.*, 2000).

Au niveau fonctionnel, certaines caractéristiques sont communes aux protéines à POU domaine (Andersen et Rosenfeld, 2001). Ainsi, les membres d'une même sous-classe tendent à reconnaître des sites consensus de séquences très homologues. De même, des membres de classes différentes peuvent reconnaître des séquences identiques avec la même affinité. De plus, toutes les protéines à POU domaine reconnaissent, avec des affinités différentes, l'octamère 5'-ATGCAAAT initialement identifié comme séquence cible de Oct-1. Le site reconnu sur l'ADN est primordial à la fixation, il détermine le mode d'interaction (monomère, homo ou hétérodimère) et le recrutement des protéines associées au facteur de transcription. Ainsi, l'interaction entre Oct-1 et OCA-B sur le site TATGARAT inhibe la transcription alors

que l'interaction entre Oct-2 et OCA-B sur ce même site l'active. Les séquences qui juxtaposent le site consensus jouent elles aussi un rôle dans la reconnaissance spécifique du facteur de transcription.

La structure du POU domaine des protéines Pit-1 et Oct-1 a été mise en évidence par cristallographie et RMN. Les similarités de structure permettent d'établir des bases quant à l'interaction des protéines à POU domaine et l'ADN. Ainsi, le domaine POU_S est composé de 74 à 82 résidus alors que le domaine POU_H contient 60 acides aminés. La taille du linker est très variable (entre 15 et 56 résidus) ce qui confère à la protéine une plus ou moins grande flexibilité d'interaction à l'ADN (Jacobson *et al.*, 1997). Individuellement, les deux domaines HTH peuvent interagir avec l'ADN avec une faible affinité (Verrijzer *et al.*, 1990). Pour avoir une interaction de forte affinité, il est indispensable que chacun des deux domaines se lie à l'ADN. Le domaine POU_S présente quatre hélices α alors que le domaine POU_H est constitué de trois hélices. Pour ces deux régions, le motif HTH est composé par les hélices 2 et 3, avec l'hélice 3 comme hélice de reconnaissance (Figure 29) (Klemm et Pabo, 1996). Les deux domaines HTH n'effectuent aucune liaison entre eux, ce qui participe à augmenter la flexibilité de la protéine.

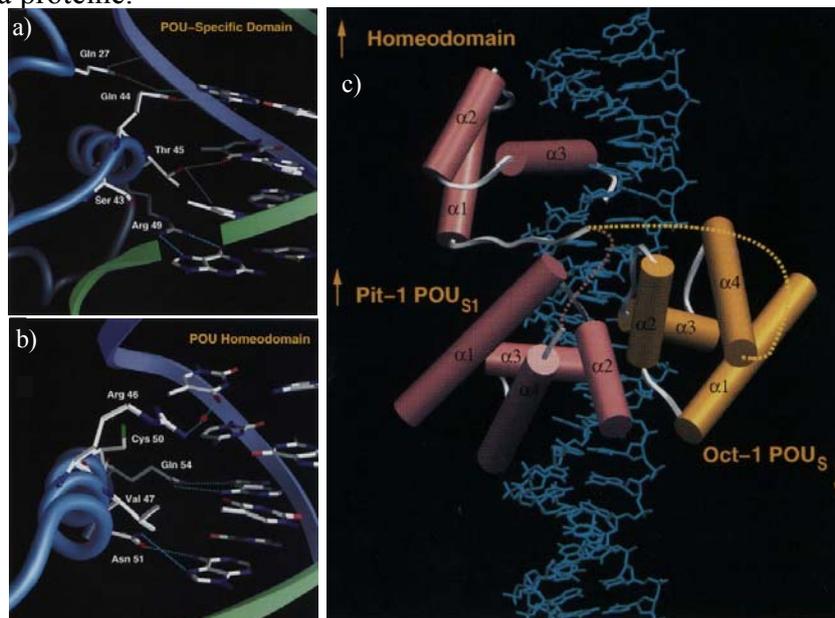


Figure 29 : Structure du POU domaine lié à l'ADN.

Interactions entre les hélices de reconnaissance de POU_S (a) ou POU_H (b) et l'ADN. La reconnaissance s'effectue via des liaisons hydrogène et électrostatiques entre certains résidus des hélices alpha et l'ADN. Les hélices de reconnaissance sont positionnées dans le grand sillon. c) Structure tridimensionnelle de Pit-1 et de Oct-1 sur le site octamère. Pour Pit-1 et Oct-1, le domaine POU_H est orienté de façon similaire (en violet) alors que les domaines POU_S (en violet pour Pit-1 et en orange pour Oct-1) sont orientés sur les côtés opposés de l'ADN (Jacobson *et al.*, 1997).

Les protéines à POU domaine modulent la transcription en interagissant avec d'autres protéines. Ces interactions protéine/protéine s'effectuent via le POU domaine. Le mode d'interaction entre les deux partenaires dépend du facteur de transcription à POU domaine et de la protéine reconnue.

L'activation de la transcription s'effectue par des contacts entre les régions C-terminales ou N-terminales du POU domaine et des co-activateurs ou co-répresseurs de la transcription. Ces séquences amino- et carboxy-terminales sont très variables entre les différentes protéines POU (Tanaka et Herr, 1990). Le facteur de transcription Pit-1 peut être associé au complexe protéique CBP/p300 par son domaine POU_H afin d'activer la transcription (Xu *et al.*, 1998).

3. Implication dans le cancer

Dans la plupart des cas, les facteurs de transcription à POU domaine jouent un rôle prépondérant dans la différenciation cellulaire. Les dérégulations de l'expression de certains de ces facteurs de transcription sont impliquées dans des phénomènes de tumorigenèse.

a) Le facteur de transcription Pit-1

Il existe une controverse concernant le rôle de Pit-1 dans les tumeurs pituitaires. Une étude effectuée sur quinze tumeurs de la glande pituitaire a montré une expression normale du facteur de transcription Pit-1 et des gènes qu'il module (Pellegrini *et al.*, 1994), alors que des études menées auparavant avaient révélé qu'une surexpression de Pit-1 pouvait être retrouvée dans ces types de tumeurs (Delhasse *et al.*, 1993 ; Asa *et al.*, 1993). De plus, l'utilisation d'oligonucléotides *anti-sens* de Pit-1 avait mis en évidence, non seulement un arrêt de la transcription des gènes cibles *GH* et *PRL*, mais aussi l'inhibition de la prolifération des cellules somatotrophes et lactotrophes. Cette inhibition a été montrée par des études d'incorporation de thymidine [³H] (Castrillo *et al.*, 1991).

Il a aussi été montré que Pit-1 est exprimé dans d'autres tissus et types cellulaires comme dans le placenta (Bamberger *et al.*, 1995), certaines cellules leucémiques (HL-60 et Raji) (Delhasse *et al.*, 1993) et même dans les glandes mammaires (Gil-Puig *et al.*, 2002). La protéine Pit-1 est surexprimée dans les cancers du sein où elle pourrait avoir un rôle dans les processus de cancérisation. Des expériences de transfection de la protéine Pit-1 dans des cellules MCF-7 (cancer du sein) montrent une augmentation de la prolifération dose-dépendante (Gil-Puig *et al.*, 2005). L'augmentation de la prolifération serait donc due à la surexpression de Pit-1 et non l'inverse. De plus, l'augmentation de l'expression de GH dans

les glandes mammaires est associée à des désordres prolifératifs incluant le carcinome invasif ductal. Pour la prolactine, il a été rapporté qu'au moins 44% des patientes atteintes d'un cancer du sein possèdent une hyper-prolactinémie durant certains stades du développement de la maladie (Holtkamp *et al.*, 1984).

b) Les facteurs de transcription Brn-3

La sous famille IV est constituée de trois protéines Brn-3a, Brn-3b et Brn-3c aussi appelées Pou4f1, Pou4f2 et Pou4f3. Ces facteurs de transcription sont codés par trois gènes différents. L'implication de Brn-3a et Brn-3b dans les processus de cancérogenèse de tissus ou ces protéines sont normalement peu exprimées est bien établie.

(1) La protéine Brn-3a

La protéine Brn-3a voit son expression augmentée d'un facteur 300 dans les cancers du col de l'utérus de type 3 (CIN3 cervical intra-epithelial neoplasia) où elle active notamment l'expression des protéines transformantes E6 et E7 du papilloma virus de type 16 humain (HPV-16) (Ndisang *et al.*, 1998 ; Schiffman *et al.*, 1993). L'utilisation d'ARN *anti-sens* a permis de mettre en évidence le rôle de Brn-3a dans l'expression du HPV : une diminution du taux de Brn-3a est corrélée à une diminution de l'expression de HPV (Ndisang *et al.*, 1999). La diminution du taux d'expression de Brn-3a réduit la capacité des cellules SiHa (cellules du cancer du col de l'utérus) à former des tumeurs chez la souris et montre une diminution de l'expression de la protéine HPV E6 et de l'oncogène cellulaire Bcl-2 (Ndisang *et al.*, 2001). Le facteur de transcription Brn-3a est aussi fortement exprimé dans les sarcomes d'Ewing et dans les tumeurs neuro-épithéliales. Il a été proposé que Brn-3a jouerait un rôle dans ces pathologies, tout comme dans les tumeurs neuroendocrines où Brn-3a pourrait exercer un rôle oncogénique. Dans ces tumeurs, un taux élevé en protéine Brn-3a présage d'un mauvais pronostic (Leblond-Francillard *et al.*, 1997).

(2) La protéine Brn-3b

Le facteur de transcription Brn-3b joue un rôle clef dans la croissance et la prolifération cellulaire des cancers du sein (Dennis *et al.*, 2001). La surexpression de Brn-3b dans des lignées de cancer du sein permet d'inhiber l'expression de la protéine anti-oncogène BRCA-1 (*Breast cancer 1*) (Budhram-Mahadeo *et al.*, 1999). De plus, des expériences de microarrays ont montré que cette surexpression active l'expression de la protéine CDK4 (*Cyclin-*

dependant Kinases). Cette protéine est importante dans la régulation de la croissance cellulaire (Samady *et al.*, 2004).

Brn3-b possède aussi un rôle significatif dans la croissance et la prolifération des cellules de neuroblastome *in vitro* et *in vivo*. Des tests de croissance cellulaire montrent une augmentation d'un facteur 5 de la croissance des cellules de neuroblastomes où le facteur de transcription Brn-3b est surexprimé. Ainsi, l'injection chez la souris de cellules de neuroblastome surexprimant Brn-3b permet une augmentation rapide du volume tumoral par rapport à des cellules ne sur-exprimant pas la protéine (Irshad *et al.*, 2004). De plus, des tests d'invasivité suggèrent que Brn-3b contribue au potentiel invasif de ces cellules.

C. Les facteurs de transcription de la famille Ets

La première protéine identifiée possédant un domaine ETS provient du virus E26 (avian erythroblastosis virus) et porte le nom de v-ets (E Twenty-six). Basé sur l'homologie de séquence du domaine d'interaction avec l'ADN, de nombreuses protéines Ets ont été identifiées dans de nombreux organismes. Chez l'homme, on en dénombre actuellement 27 (Sharrocks, 2001). Ces protéines sont impliquées dans de nombreux processus au cours du développement mais aussi chez l'adulte. La dérégulation des protéines Ets par une expression inappropriée ou une fusion avec une autre protéine peut entraîner des processus de cancérogenèse.

1. Structure des facteurs de transcription ETS

Les membres de la famille ETS sont divisés en sous-familles en fonction de leur élément de reconnaissance à l'ADN. La plupart d'entre eux ont leur domaine Ets du côté C-terminal, cependant certains comme la sous famille des TCF (Ternary Complex Factor) possèdent leur domaine Ets du côté N-terminal. En plus de ce domaine, certains facteurs possèdent en N-terminal un autre domaine très conservé appelé PNT (pointed) de type HLH permettant les interactions protéines-protéines (Figure 30a).

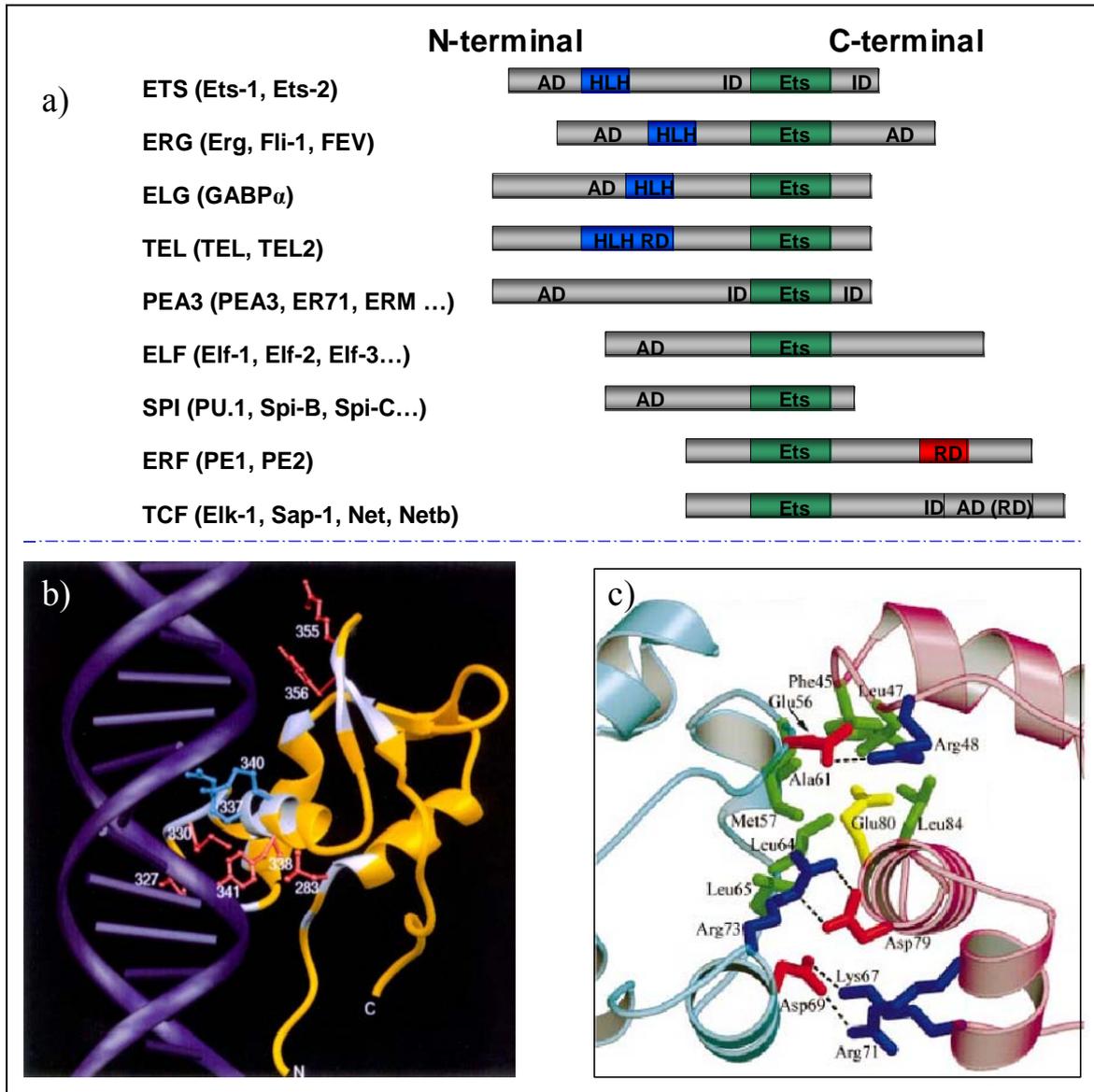


Figure 30 : Membres de la famille ETS et interaction à l'ADN.

a) Structure schématique des membres de la famille ETS. Chaque protéine contient un site Ets (en vert). Le domaine HLH correspond au site PNT (en bleu). AD, domaine d'activation ; ID, domaine d'auto-inhibition ; RD, domaine de répression (Oikawa et al., 2002). b) Organisation structurale de l'interaction entre la protéine Fli-1 et l'ADN. La plupart des contacts avec l'ADN s'effectue via les résidus de l'hélice 3, R355 et Y356 de l'aile formée entre les feuillets β 3 et 4 et K327 et M330 situés sur le tour du motif HTH (Liang et al., 1994). c) Interaction entre deux protéines TEL par leur domaine HLH. Les résidus basiques (en rouge) effectuent des liaisons électrostatiques avec les résidus acides (en bleu) portés par l'une des hélices α de la seconde protéine. Des liaisons hydrophobes entre les hélices (résidus en vert) stabilisent également cette interaction (Kim et al., 2001).

a) Le domaine Ets

Le domaine Ets d'interaction à l'ADN est constitué de 85 acides aminés et reconnaît la séquence minimale GGA(A/T) appelée EBS (Ets Binding Sequence) (Seth et Watson, 2005). Les protéines ETS interagissent avec l'ADN via un motif HTH modifié appelé winged helix-

turn-helix. Ce domaine est englobé dans une structure composée de trois hélices α et quatre feuillets β (Kodandapani et al., 1996) organisés de la façon suivante : $\alpha 1$ - $\beta 1$ - $\beta 2$ - $\alpha 2$ - $\alpha 3$ - $\beta 3$ - $\beta 4$. Les hélices 2 et 3 composent le motif HTH avec l'hélice $\alpha 3$ comme hélice de reconnaissance. La plupart des contacts entre l'ADN et la protéine ont lieu au niveau de l'hélice $\alpha 3$, au niveau de « l'aile » formée entre les feuillets $\beta 3$ et 4 et aussi avec le tour situé entre les hélices $\alpha 2$ et 3 (Figure 30b). Deux résidus arginine très conservés de l'hélice $\alpha 3$ établissent des liaisons spécifiques avec les deux guanines du site EBS. Les sites réellement reconnus par les différentes protéines ETS dépendent des séquences flanquantes du site minimal GGA(A/T). L'interaction entre les protéines ETS et le site consensus est également facilitée par l'interaction avec d'autres facteurs de transcription comme AP-1 pour le promoteur du gène de la métalloprotéinase 1 (mmp-1).

b) Le domaine PNT

Le second domaine conservé retrouvé chez les protéines ETS est le domaine PNT. Cette région composée de 65 à 85 résidus est retrouvée dans 11 des 27 protéines ETS humaines, elle permet l'interaction protéine-protéine et l'oligomérisation. Des analyses structurales ont montré que cette région présente une architecture unique composée d'un motif HLH englobé dans cinq hélices α . Ce motif est capable d'interagir avec d'autres protéines possédant un motif similaire (Figure 30c) (Kim et al., 2001).

2. Fonctions des protéines ETS

Les protéines ETS possèdent des fonctions très variées et sont impliquées dans différents processus biologiques (Sharrocks *et al.*, 2001).

a) Fonctions cellulaires

Des études *in vitro* ont permis d'identifier divers rôles des protéines ETS. Par exemple, les facteurs de transcription appartenant à la sous-famille TCF sont généralement impliqués dans la régulation de la prolifération en réponse à l'activation de la voie ERK (*Extracellular Signal-regulated Kinases*) (Treisman *et al.*, 1994). D'autres protéines ETS, comme Ets-1 et Ets-2 interviennent dans l'apoptose avec un rôle pro- ou anti-apoptotique. Leur rôle dans ce processus est souvent défini par le partenaire protéique (Oikawa et Yamada, 2003). Ets-1 peut notamment former un complexe stable dans les cellules souches embryonnaires avec p53, CBP/p300 et l'ADN pour induire l'apoptose (Xu *et al.*, 2002). Le facteur de transcription Ets-

Il est aussi impliqué dans des phénomènes de sénescence par activation du gène p16 qui induit un blocage du cycle cellulaire et permet l'entrée des cellules en sénescence.

b) Différenciation cellulaire

Un des rôles les plus connus des protéines ETS est celui de la régulation de l'hématopoïèse. Ainsi, Ets-1 possède un rôle central dans la survie des cellules T et pour la production des cellules NK (*Natural Killer*) (Bories *et al.*, 1995 ; Muthusamy *et al.*, 1995). La protéine PU.1 intervient dans la différenciation des cellules lymphoïdes et myéloïdes. Ainsi, une délétion sur le gène *pu.1* aboutit à une réduction du nombre des cellules progénitrices lymphoïdes et myéloïdes (Scott *et al.*, 1997). La décision d'induire la différenciation des cellules myéloïdes et lymphoïdes semble dépendre du niveau d'expression de PU.1. Un taux élevé de cette protéine induit la différenciation des cellules en macrophages, alors qu'un faible taux se traduit par une différenciation en cellules B (DeKoeter et Singh, 2000).

Ces exemples de fonctions dans la différenciation des cellules hématopoïétiques montrent la complexité et le rôle croisé des facteurs de transcription. Ainsi, comme nous l'avons vu précédemment, la surexpression de la protéine HoxA9 permet de maintenir les cellules progénitrices dans un état non différencié. La différenciation est quant à elle dépendante d'autres facteurs de transcription comme certains facteurs de la famille ETS.

c) L'angiogenèse et la vascularisation

Certaines protéines ETS sont mises en jeu dans l'angiogenèse et la vascularisation (Lelièvre *et al.*, 2001). L'expression de nombreux membres de la protéine ETS, comme Ets-1, Erg (*Ets related gene*) et Fli-1 (*friend leukemia integration 1*) corrèle avec l'existence de phénomènes angiogéniques normaux et pathologiques (Mattot *et al.*, 1999). Des embryons de souris ayant des mutations sur le gène codant pour la protéine Fli-1 présentent des hémorragies au niveau du cerveau et dans le tube neural indiquant des défauts possibles dans l'aorte dorsale (Spiropoulos *et al.*, 2000). L'expression d'autres protéines ETS, comme Net, Erg et Ets-1, dans des cellules endothéliales indique un rôle important de ces dernières dans la vascularisation et l'angiogenèse. Des souris invalidées pour la protéine Net possèdent un système lymphatique défectueux (Ayadi *et al.*, 2001).

Enfin, les protéines ETS possèdent bien d'autres fonctions notamment dans la connexion entre les neurones sensoriels et moteurs (Lin *et al.*, 1998), les souris déficientes en protéine ER81 présentent un déficit dans ces connexions.

3. Régulation des protéines ETS

L'activité transcriptionnelle des facteurs de transcription ETS peut être modulée à de multiples niveaux. Cette activité est dépendante des interactions avec d'autres facteurs de transcription et des modifications post-transcriptionnelles (Seth *et al.*, 2005).

a) Régulation par interactions protéines-protéines

La régulation des facteurs de transcription ETS est dépendante de la combinaison de multiples facteurs de transcription. De nombreux facteurs de transcription ont leurs séquences consensus accolées au site EBS (Li et Watson, 2000). Ainsi, de manière contexte-dépendant, l'interaction des protéines ETS avec d'autres facteurs de transcription augmente l'affinité de ces protéines pour l'ADN et résulte en une synergie d'activation ou de répression de la transcription. Une coopérativité a été montrée entre certaines protéines ETS et le complexe AP-1 pour activer la transcription de nombreux gènes comme MMP1 (matrix metalloprotéase-1) (Imler *et al.*, 1988). D'autres interactions impliquent les protéines SRF et Elk-1 qui forment un complexe ternaire avec le motif SRE présent dans la zone promotrice de nombreux gènes comme ceux de c-fos ou erg-1 (Buchwalter *et al.*, 2004).

b) Régulation post-transcriptionnelle

(1) La phosphorylation

La fonction des protéines ETS est régulée par phosphorylation. Ces modifications ont des effets sur l'interaction avec l'ADN, sur les interactions protéine/protéine, sur l'activité transcriptionnelle et sur la localisation cellulaire de ces protéines (Yordy et Muise-Helmericks, 2000). Un des pré-requis à l'induction du cancer est la dérégulation des voies de signalisation induisant une activité constitutive de certaines protéines. Erk, JNK (c-Jun N-terminal kinase) et la p38 MAPK (Mitogen-Activated Protein Kinase) sont des composantes des cascades de phosphorylation. Les protéines Erk sont activées en réponse à un signal mitogène alors que les protéines JNK et p38/SAPK (Stress Activated Protein Kinase) répondent à un stress (Rubinfeld et Seger, 2005). Les facteurs de transcription ETS sont des substrats de ces voies de signalisation, la phosphorylation d'un site adjacent au site PNT de

Ets-1 ou Ets-2 augmente leur activité transcriptionnelle sans pour autant modifier leur interaction à l'ADN. La phosphorylation calcium-dépendante de Ets-1 sur un site adjacent au domaine d'interaction à l'ADN inhibe l'interaction entre la protéine et la double hélice (Pufall et al., 2005). Par opposition à Ets-1, la phosphorylation de la protéine TEL (Translocation Ets leukemia) par la voie MAP kinase résulte en une exportation de la protéine en dehors du noyau et diminue l'affinité de la protéine pour l'ADN (Hanson et al., 2007). Ces processus de phosphorylation sont réversibles et contrôlés par des phosphatases.

(2) L'acétylation

En plus des phosphorylations, les protéines ETS peuvent être régulées par acétylation. Ainsi, l'acétylation de la protéine ER81 augmente sa capacité d'interaction à l'ADN et par la même occasion son activité transcriptionnelle (Goel et Janknecht, 2003). Il a été montré que le facteur de transcription Ets1 peut être acétylé en réponse à une stimulation au TGF β (Czuwara-Ladykowska et al., 2002). Cette acétylation pourrait jouer des rôles multiples comme augmenter la stabilité de la protéine, augmenter les interactions avec d'autres facteurs de transcription, moduler les interactions avec l'ADN, modifier la sélectivité de séquences ou encore intervenir dans la localisation de la protéine.

(3) La sumoylation

La sumoylation altère la fonction de nombreuses protéines incluant des membres de la famille ETS. Elle peut affecter la stabilité, l'activité et la localisation des protéines (Zhao et al., 2007). Par exemple, la sumoylation de Elk-1 facilite le recrutement de l'histone désacétylase-2 (HDAC-2). Ce recrutement permet de diminuer le niveau d'acétylation des histones ce qui a pour conséquence d'altérer la structure de la chromatine. Comme précédemment décrit, cette déacétylation des histones résulte en une compaction de la chromatine et induit une répression des gènes cibles de la protéine Elk-1 (Yang et Sharrocks, 2004). Cet effet peut être annihilé par la phosphorylation de Elk-1 par les MAP kinases. Par opposition, la sumoylation de la protéine TEL inhibe la répression de certains gènes cibles du facteur de transcription par séquestration de la protéine dans le cytoplasme (Chakrabarti et al., 1999).

(4) L'auto-inhibition

De nombreux membres de la famille ETS présentent une auto-inhibition de leur fixation à l'ADN. La délétion de résidus situés dans la région centrale et du côté carboxy-terminal de la protéine Ets-1 permet d'augmenter de 10 à 20 fois l'activité de liaison à l'ADN de la protéine.

Ce résultat suggère que les séquences qui juxtaposent le domaine Ets sont impliquées dans l'auto-inhibition (Lim et al., 1992). Cette auto-inhibition est renforcée par la phosphorylation de Ets-1 par la voie signalétique calcium-dépendante mais non par la voie des MAP kinases. Il est possible que cette interaction intramoléculaire entre les régions inhibitrices déstabilise les interactions ADN/protéines par blocage stérique (Jonsen et al., 1996). Les groupements phosphates, chargés négativement effectuent des interactions électrostatiques avec les modules d'inhibition de Ets-1 (Cowley et Graves, 2000).

L'auto-inhibition peut aussi être annihilée par des interactions avec d'autres protéines. Ainsi, l'activation mutuelle de Ets-1 et AML1/RUNX1 par interaction de leurs domaines d'auto-inhibition a pu être observée dans des cellules leucémiques murines (Goetz et al., 2000). Cet effet est observé avec d'autres membres de la famille ETS comme avec la protéine PEA3 (polyoma enhancer activator 3) (Greenall et al., 2001).

4. Implication dans la cancérogenèse

Dès la fin des années 1980, l'implication des gènes ets dans les processus de cancérisation avait déjà été démontrée. Ainsi, la surexpression des protéines Ets-1 dans des fibroblastes avait permis la mise en évidence du caractère transformant de ces facteurs de transcription (Seth et al., 1989). D'autres phénomènes comme l'amplification, les mutations ou les translocations touchent ces protéines et sont impliqués dans la cancérogenèse.

a) Modifications géniques de ETS

(1) Amplification et surexpression des protéines ETS

Différentes modifications géniques des protéines ETS sont observées dans le cancer. Dans certains cas de LAM, les gènes Ets-1 et Fli-1 sont amplifiés. Des amplifications du gène Ets-2 ont aussi été démontrées chez des patients atteints de leucémies aiguës non lymphoblastiques. Dans les chromosomes 6, 21 et 18 (Santoro *et al.*, 1992). La protéine Erg est surexprimée dans certains cas de LAM et de LLA-T (Baldus *et al.*, 2004) et elle est associée à un mauvais pronostic (Marcucci *et al.*, 2007).

(2) Les mutations

De multiples points de mutations sur le facteur de transcription PU.1 ont été identifiés chez les patients atteints de LAM. La plupart de ces mutations étant localisées sur le domaine

d'interaction à l'ADN. Ces événements génétiques ont pour conséquence de diminuer la capacité de régulation de l'expression de certains gènes cibles et de moduler les interactions entre ce facteur de transcription et d'autres protéines (Mueller *et al.*, 2002).

(3) Les translocations

Le cancer implique des aberrations chromosomiques dont les translocations. Les membres de la famille ETS sont impliqués dans ces phénomènes, les exemples les plus connus étant ceux qui impliquent la protéine TEL. Le facteur de transcription TEL (*Translocation Ets Leukemia*) comme son nom l'indique est impliqué dans de nombreuses translocations retrouvées dans les leucémies. A l'origine, cette protéine a été identifiée dans des leucémies myélomonocytaires chroniques présentant la translocation t(5,12)(q33;p13) partenaire de fusion du récepteur PDGF β (Golub *et al.*, 1994). Cette protéine est réarrangée dans d'autres types de leucémie comme les LAM, les SMD ou encore les LAL. L'un ou l'autre des domaines PNT et/ou Ets sont impliqués dans plus de 20 translocations différentes observées dans les leucémies humaines et plus rarement dans les tumeurs solides (Mavrothalassitis et Ghysdael, 2000). Les protéines de fusions impliquant le domaine PNT aboutissent souvent en une oligomérisation nécessaire pour activer de façon constitutive des protéines kinases comme le récepteur PDGFR- β ou encore JAK2. Ces activations permettent la transformation cellulaire. Alors que la plupart des protéines de fusion impliquant le domaine Ets résultent en une régulation aberrante de l'expression des gènes cibles (Seth *et al.*, 2005). Il existe cependant d'autres translocations qui impliquent d'autres protéines ETS.

b) *Expression des gènes Ets et cancers*

Il existe une corrélation entre le niveau d'expression des gènes Ets et la progression tumorale rencontrée dans certaines tumeurs solides. La surexpression de certains gènes Ets est associée avec les stades précoces des cancers du pancréas et du foie (Ito *et al.*, 2002). Des expériences d'immunohistochimie ont mis en évidence la surexpression de Fli-1 et Elf-1 dans les adénocarcinomes prostatiques (Gavrilov *et al.*, 2001). Dans le cancer du colon, certaines protéines ETS sont retrouvées avec des taux élevés et corrélent avec des stades avancés. Ainsi, dans les tissus sains du colon et dans les polypes hyperplasiques on ne retrouve pas les gènes Ets-1 et Ets-2 alors que leur surexpression est retrouvée dans les grades avancés. La protéine PEA3 est surexprimée dans les cancers du sein invasifs. Sa surexpression est liée à l'augmentation de l'expression du récepteur tyrosine kinase HER2/NEU dont le rôle dans ce type de cancer est bien établi (Benz *et al.*, 1997). Les protéines ETS sont impliquées dans de

nombreux autres cas de cancer comme les cancers de la prostate, du sein, de l'ovaire, des tumeurs d'Ewing (Seth et al., 2005).

c) Rôle de la famille ETS dans l'invasion et l'angiogenèse

Les protéines ETS régulent l'expression de gènes qui codent pour des enzymes capables de dégrader la matrice extracellulaire comme les MMP (Westermarck et Kähäri, 1999) importantes pour l'invasion tumorale. Trois protéines ETS structurellement distinctes Ets-1, Fli-1 et PU.1 permettent l'augmentation de l'expression de la protéine MMP-1 (Gavrilov et al., 2001). De manière similaire, l'expression de Ets-1 et Ets-2 potentialise l'activation des promoteurs des gènes uPA et MMP-9 en réponse au signal EGF (*Epidermal Growth Factor*) (Watabe *et al.*, 1998).

Les protéines Ets-1, Erg et Fli-1 sont aussi impliquées dans les phénomènes d'angiogenèse. La protéine Ets-1 notamment, est capable d'induire l'expression du récepteur au VEGF, indispensable à la mise en place de la néovascularisation (Plate *et al.*, 1993).

OBJECTIFS

Ainsi, le ciblage de la liaison à l'ADN de facteurs de transcription impliqués dans les phénomènes de tumorigenèse permettrait d'en moduler son activité et représente une nouvelle stratégie qui tend à inhiber spécifiquement l'activité de certains gènes. Les composés DB, nouvelle classe de molécules capable de cibler spécifiquement l'ADN, entrent dans cette stratégie.

Il existe déjà des familles de molécules capables d'inhiber l'interaction ADN/Facteurs de transcription comme les polyamines capables de reconnaître des séquences spécifiques ou encore l'ecteinascidine (ET743) pour laquelle des travaux récents ont montré son implication dans le contrôle de la transcription. En effet, cet alkylant du petit sillon de l'ADN inhibe la fixation du facteur de transcription NF- κ B sur les séquences promotrices du gène *mdr*. Ces exemples argumentent l'approche moléculaire envisagée et la nature des cibles retenues comme éléments pertinents de la réponse antitumorale recherchée.

Comme nous venons de le voir, si les règles qui régissent les interactions entre les composés DB et l'ADN sont assez bien connues et argumentées, il reste encore 1) à trouver de nouvelles molécules capables de reconnaître des séquences d'ADN spécifiques et 2) à caractériser leur(s) capacité(s) à moduler la liaison des facteurs de transcription puisque cette propriété n'a pas encore été élucidée pour cette famille de molécules.

Le premier objectif de ce travail est de découvrir de nouveaux dérivés DB capables de reconnaître l'ADN de manière spécifique sur des sites originaux et variés. Nous commencerons par étudier la sélectivité d'interaction et l'affinité pour l'ADN de ces nouveaux composés de la famille des diphényl-furanes et des phényl-furane-benzimidazoles substitués sur le corps polycyclique qui pourraient reconnaître des sites riches en paires de bases AT et ATGA, voire d'autres sites originaux et spécifiques. Les conséquences cellulaires en termes de distribution et de cytotoxicité de ces composés seront alors évaluées. En plus de la découverte de nouvelles molécules capables de cibler spécifiquement l'ADN, ces données permettront de compléter les relations de structure/sélectivité, structure/affinité et structure/activité cellulaire de ces composés.

Dans le second objectif de ce travail, les différentes molécules sélectionnées pourront alors être évaluées en terme de capacité à moduler la liaison de certains facteurs de transcription à l'ADN. Le but de cette partie du travail est d'établir un modèle de régulation des gènes par les

composés DB. Pour cela, nous utiliserons deux stratégies : 1) dans un premier temps, une étude de la modulation de la fixation à l'ADN d'un large panel de 54 facteurs de transcription par les composés DB sera réalisée. Pour cela, nous optimiserons et validerons les conditions d'utilisation des membranes Transignal/Protein/DNA array. Pour les facteurs de transcription dont la liaison à l'ADN sera modulée par un ou plusieurs composés DB, le mode d'interaction entre le composé et le site consensus du complexe ADN/facteur de transcription sera élucidé ; et 2) dans un second temps, un ciblage d'un facteur de transcription associé de manière privilégiée à un cancer donné et se liant au même type de séquence que les composés DB comme le complexe de transcription PBX/HoxA9 sera alors réalisé. Enfin, les conséquences cellulaires de l'inhibition de la protéine HoxA9 seront évaluées en termes de cytotoxicité, de survie et de prolifération cellulaire sur des modèles cellulaires qui expriment ou non la protéine HoxA9.

Dans ce travail, nous espérons établir les premières bases moléculaires de l'interaction de composés DB avec un ou plusieurs complexes ADN/facteur de transcription qui permettront peut être de sélectionner de nouvelles molécules « ciblées » pour leur développement préclinique en oncologie.

RESULTATS

I. Nouveaux dérivés DB et sélectivité de reconnaissance de l'ADN

Cette première partie de l'étude a pour objectif de découvrir de nouveaux composés DB sélectifs dans leur reconnaissance à l'ADN et d'en étudier pour certains les conséquences cellulaires. Dans un premier temps, nous nous sommes proposés évalué le rôle du corps polycyclique de nouveaux composés DB (Figure 31) dans la reconnaissance et l'affinité de ces molécules pour les séquences riches en paires de bases AT et le site particulier ATGA. Pour cela, nous avons décidé d'orienter notre étude sur des composés du type diphényl-furane et phényl-furane-benzimidazole. Ces modifications portent i) sur les cycles phényles, sur le cycle furane ou sur la position des amidines terminales pour les composés diphényl-furanes (Figure 31a) ; ii) sur le cycle phényle, le cycle furane ou le cycle benzimidazole pour les composés du type phényl-furane-benzimidazole (Figure 31b). Ces nouvelles données permettront d'affiner les règles qui régissent l'interaction entre les composés diphényl-furanes ou phényl-furane-benzimidazoles et l'ADN. Dans un second temps, nous avons analysé les interactions moléculaires de nouvelles classes de composés DB ou assimilés potentiellement capables de reconnaître des nouvelles séquences nucléotidiques. Dans un troisième temps, nous avons étudié les conséquences cellulaires pour certains composés.

A. Nouveaux composés : Modifications du corps polycyclique

1. Les composés diphényl-furanes

a) Sélectivité de séquence

Pour déterminer la spécificité de reconnaissance de ces composés DB à l'ADN, des tests d'empreinte à la DNase I ont été réalisés. Dans le cas présent, les dérivés DB diphényl-furanes ont été incubés à des concentrations croissantes avec un fragment d'ADN de 265 pb marqué en position 3' contenant des séquences riches en paires de bases AT (rectangles bleus) ainsi que deux sites spécifiques ATGA consécutifs (rectangles rouges) séparés par deux cytosines (Figure 32a).

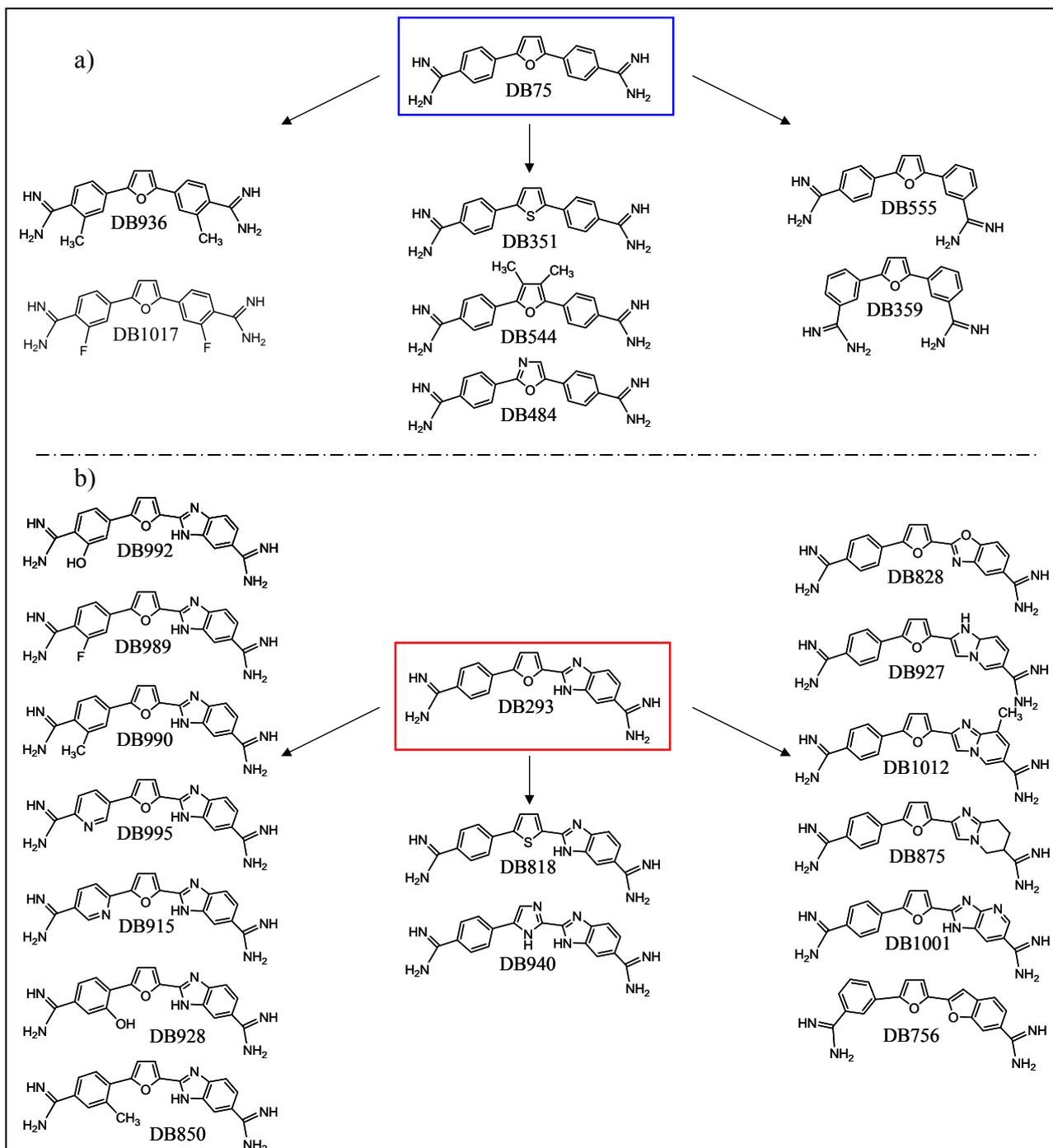


Figure 31 : Structures des composés du type diphenyl- furane et phényl-furane- benzimidazole.

a) Composés du type diphenyl-furane dérivés du DB75. b) Composés du type phényl-furane- benzimidazole dérivés du DB293.

L'analyse du gel (Figure 32a) met en évidence la présence de 2 profils distincts :

Un premier, semblable à celui observé pour le DB75, où l'on observe une protection importante de la digestion sur le site situé entre les paires de bases 70 à 80 (p 70-80). Il faut noter que pour des concentrations fortes (de l'ordre de 5 μ M) en DB351 et DB544, le profil observé est proche de celui du « G-track ». Ce profil révèle une action sur les guanines induisant un site abasique labile au sein du fragment d'ADN. Le second profil, notamment observé pour le DB936, montre un effacement des bandes sur l'ensemble de la piste qui correspond à une fixation non spécifique de la molécule.

L'analyse densitométrique du gel (Figure 32b) permet de comparer des molécules capables d'interagir spécifiquement avec les sites riches en paires de bases AT (DB484) ou non (DB555) en utilisant le DB75 comme molécule de référence. Les composés qui se lient à l'ADN de manière spécifique présentent un clivage différentiel négatif sur la zone de fixation correspondant à une protection du clivage par rapport au contrôle.

L'ensemble de ces résultats montre donc qu'il existe 2 sous-familles de composés :

- Les composés pour lesquels on observe une protection du clivage sur les sites AATT (p 73-79) et 5'-ATTA (p 92-95). Les molécules DB351, DB484, DB544, DB1017 induisent ce type d'empreinte et sont donc capables d'interagir de façon spécifique avec les séquences riches en paires de bases AT de l'ADN comme le composé de référence DB75.
- La seconde classe de composés se lie avec une très faible spécificité aux sites riches en paires de bases AT, avec des empreintes très peu marquées comme le composé DB555. Ces molécules sont classées parmi les composés qui interagissent de façon non spécifique avec l'ADN, c'est le cas des DB936, DB555.

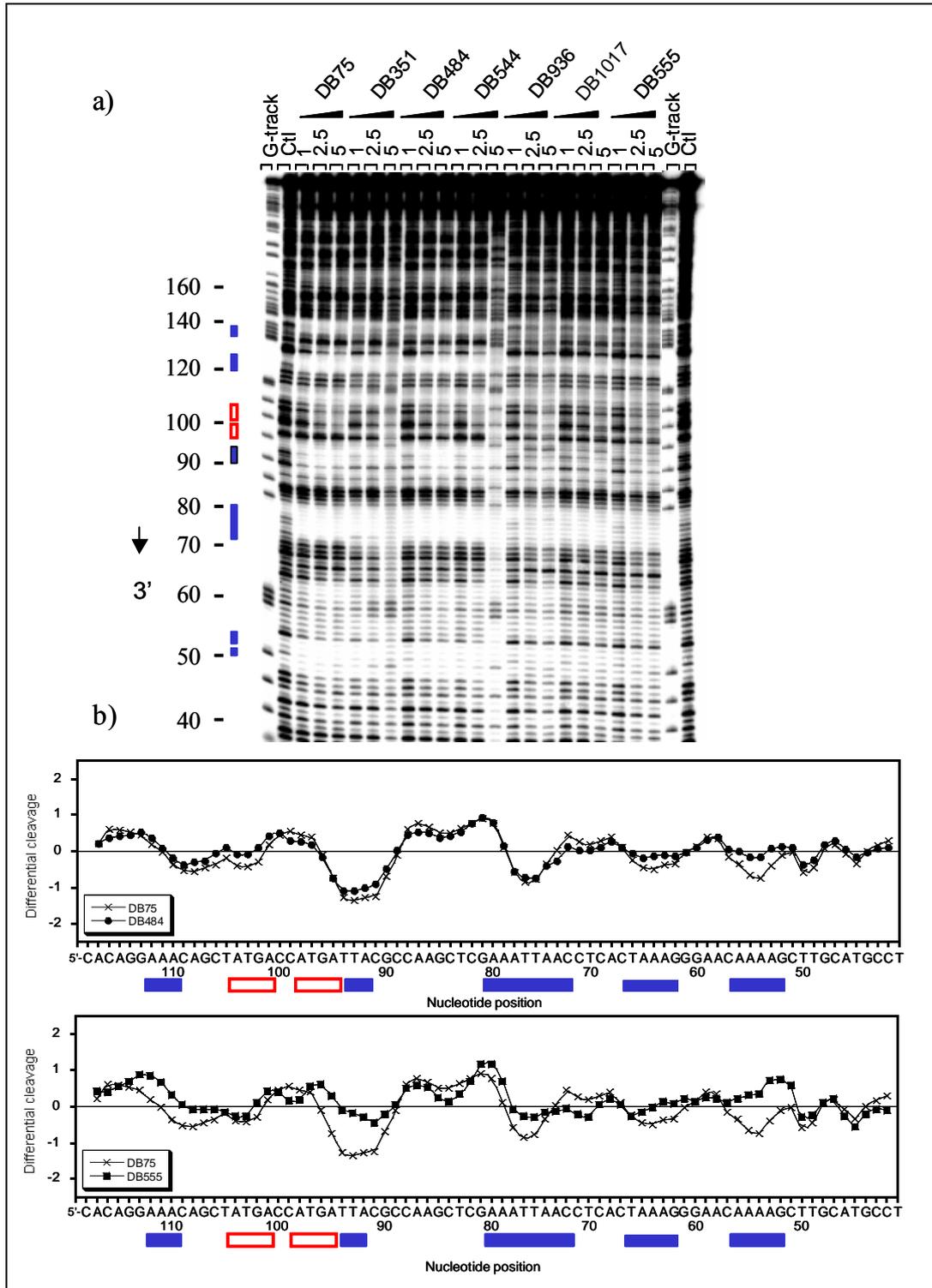


Figure 32 : Empreinte à la DNase I des composés diphényl-furanes.

a) Empreintes à la DNase I des composés du type diphényl-furane. Un fragment d'ADN de 265 pb, obtenu par digestion EcoR I-Pvu II du plasmide pBS marqué en 3', est incubé ou non pendant 15 min à TA avec une gamme de dérivés DB (en μM) puis digéré par la DNase I. Les contrôles (Ctl) ne contiennent pas de composés. Les lignes « G-track » identifient les guanines du fragment d'ADN. b) Analyses densitométriques comparatives des DB75/DB484 ou DB75/DB555. Les points situés au dessous de zéro correspondent aux sites de reconnaissance des composés DB protégés de la digestion. Les rectangles bleus représentent les séquences riches en paires de bases AT et les rectangles rouges les sites ATGA.

b) Affinité relative des composés pour l'ADN

Pour déterminer l'affinité relative de ces composés pour l'ADN, des tests spectroscopiques ont été réalisés sur différents types d'ADN afin de déterminer les variations de température de fusion. Puisque les tests d'empreinte à la DNase I ont permis de sélectionner des molécules capables d'interagir spécifiquement sur des sites riches en paires de bases AT, un fragment d'ADN riche en paires de bases AT (poly(dAdT)₂) a été sélectionné pour déterminer l'affinité relative de ces composés pour ces sites (Figure 33a). Dans un second temps, nous avons déterminé les conséquences de la perte du site riche en paires de bases AT sur l'affinité de la molécule. Pour cela, nous avons utilisé un premier oligonucléotide de référence contenant le site TTAATTT retrouvé dans le fragment de 265 pb (oligonucléotide AT), puis nous avons utilisé un second oligonucléotide TTGGTTT où le site riche en paires de bases AT est muté (oligonucléotide ATm) (Figure 33b).

Les mesures de variation des températures de fusion effectuées sur le fragment poly (dAdT)₂ montrent qu'aucune des molécules testées n'interagit plus fortement avec l'ADN que le composé de référence DB75. Ces résultats démontrent qu'il existe deux familles de composés :

- la première famille possède une forte affinité pour l'ADN (DB75, DB351, DB544, DB484) avec des variations de température de fusion (ΔT_m) supérieures à 20°C
- la seconde famille possède une affinité plus faible pour l'ADN (DB1017, DB555 et DB936).

Ces résultats corrélerent avec les tests d'empreintes à la DNase I car, à l'exception du DB1017, les molécules qui se lient de manière spécifique aux sites riches en paires de bases AT possèdent une meilleure affinité relative pour le poly (dAdT)₂ que les composés non spécifiques.

Les composés DB75, DB351, DB544 et DB484 montrent une affinité relative élevée pour l'ADN avec des ΔT_m d'environ 10°C. Cette affinité relative est moindre pour les composés non spécifiques.

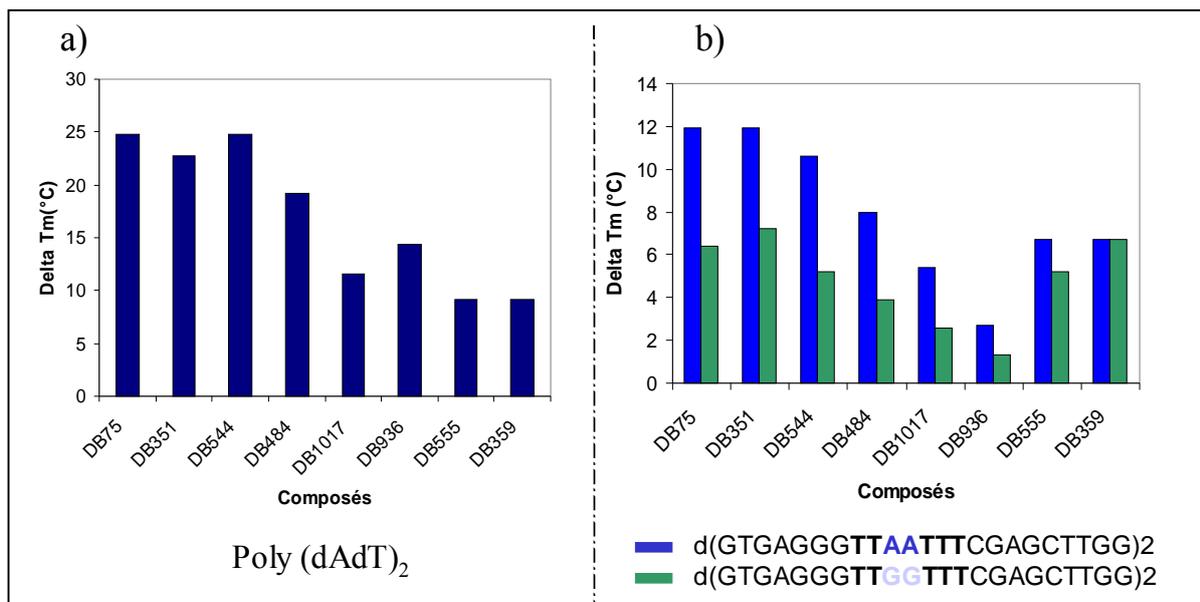


Figure 33 : Diagrammes représentant les variations de T_m des dérivés diphenyl-furanes.

Variation de la température de fusion ($\Delta T_m = T_m^{\text{complexe}} - T_m^{\text{DNA}}$) induite par les composés DB sur différents fragments d'ADN. a) Mesures de ΔT_m effectuées sur un fragment poly (dAdT)₂. Pour cette étude, 20 μM de poly (dAdT)₂ sont incubés avec 1,5 μM de composé DB ; b) Mesures de ΔT_m effectuées sur les oligonucléotides AT (en bleu) et ATm (en vert) à une concentration de 1 μM incubés avec 2 μM de dérivés de la famille du DB75.

L'utilisation de l'oligonucléotide AT valide les résultats obtenus précédemment. L'oligonucléotide ATm entraîne une diminution de la température de fusion pour la plupart des composés DB. Néanmoins, de manière prévisible, on remarque que les composés DB qui interagissent spécifiquement avec l'ADN sont plus affectés que les composés qui interagissent de façon non spécifique. Ainsi, on observe une diminution du ΔT_m de l'ordre de 1,5°C pour les composés non spécifiques, alors que pour les composés qui se lient aux sites riches en paires de bases AT, la diminution moyenne est de 4,1°C. On remarque que le DB359 présente un ΔT_m de 6,6°C identique pour les deux oligonucléotides utilisés.

Une différence de variation de température de fusion de l'ADN observée entre les oligonucléotides AT et ATm conforte la spécificité de reconnaissance des sites riches en paires de bases AT par les composés concernés.

c) Relations structure/sélectivité et structure/affinité

Les composés DB351, DB544 et DB484 dont le cycle furane est respectivement substitué par un cycle thiofène, oxazole et diméthyl-furane garde une forte affinité pour les séquences riches en paires de bases AT. Ces modifications sont les seules qui permettent de garder une forte affinité pour l'ADN et une bonne sélectivité de séquences.

Les modifications apportées sur les deux cycles phényles DB936 et DB1017 (méthyle ou fluor) diminuent l'affinité de la molécule pour l'ADN. En ce qui concerne la spécificité, il est étonnant de remarquer que le DB936 perd sa spécificité de fixation à l'ADN alors que le DB1017 di-fluoré garde sa spécificité de reconnaissance.

L'angle formé par toute molécule joue un rôle important dans leur interaction avec l'ADN. Il a déjà été montré que l'angle formé par les DB555 et DB359 est préjudiciable en termes d'interaction et de spécificité pour l'ADN (Nguyen et al., 2001). Le DB359 possède ses deux fonctions amidines terminales en position méta, alors que le DB75 possède ses deux substituants en position para, et le DB555 possède un groupement en méta et un groupement en position para. Le DB75 interagit au niveau du petit sillon de l'ADN de façon spécifique aux sites riches en paires de bases AT. Le DB555 interagit quant-à lui faiblement au niveau du petit sillon et interagit par intercalation de façon non spécifique à l'ADN. Le DB359 se fixe à l'ADN uniquement par intercalation. Les résultats obtenus dans cette étude valident donc notre hypothèse car on observe une corrélation forte entre l'affinité des composés pour les sites riches en paires de bases AT de l'ADN et leur spécificité de reconnaissance. Ainsi, le composé intercalant DB359 est insensible à la modification de la séquence de l'oligonucléotide ; le DB555, qui possède une faible capacité de se fixer à l'ADN au niveau des séquences riches en paires de bases AT du petit sillon voit son ΔT_m diminuer faiblement ; et le DB75 qui interagit fortement au niveau du petit sillon de l'ADN voit son ΔT_m diminuer fortement. On remarque que la valeur de ΔT_m du DB75 sur l'oligonucléotide AT_m est identique à celui du DB359. Il a été montré que le composé DB75 possède lui aussi la capacité de s'intercaler sur les séquences riches en paires de bases GC avec une faible affinité pour l'ADN en l'absence de séquences riches en paires de bases AT.

1. Les composés phényl-furane-benzimidazoles

a) Sélectivité d'interaction

Comme pour la série des composés de la famille des diphényl-furanes, des tests d'empreintes à la DNase I ont été réalisés pour rechercher la sélectivité de séquence des différents composés phényl-furane-benzimidazoles.

Pour ces composés, une troisième catégorie de molécules apparaît (Figure 34a) pour lesquelles on observe une protection au niveau des sites **ATGACCATGA** (rectangles rouges) en plus de la protection sur les sites riches en paires de bases AT (rectangles bleus), comme pour le DB293.

L'analyse densitométrique du gel (Figure 34b) permet de comparer des molécules capables d'interagir spécifiquement avec les sites riches en paires de bases AT et les séquences ATGA (DB915) avec seulement les séquences riches en paires de bases AT (DB818) en utilisant le DB293 comme molécule de référence.

L'ensemble de ces résultats révèle effectivement trois types de composés du type phényl-furane-benzimidazole :

- les composés pour lesquels sont identifiés trois sites de protections majeurs 5'-AATT (p 73 à 79); (5'-ATTA) (p 92 à 95) et 5'-ATGACCATGA (p 95 à 104). Les composés DB940, DB989, DB915, DB928, DB850 et DB995 reconnaissent donc les séquences riches en paires de bases AT et les sites ATGA.
- les composés pour lesquels on observe une protection du clivage sur les sites 5'-AATTAAA (p 73 à 79) et 5'-ATTA (p 92 à 95). Ces molécules reconnaissent spécifiquement les séquences riches en paires de bases AT. C'est le cas des composés DB992, DB818, DB990, DB927, DB1012 et DB875.
- les composés DB1001, DB828 et DB756 qui se lient à l'ADN de façon non spécifique et pour lesquels la courbe de clivage différentiel reste aux alentours de 0.

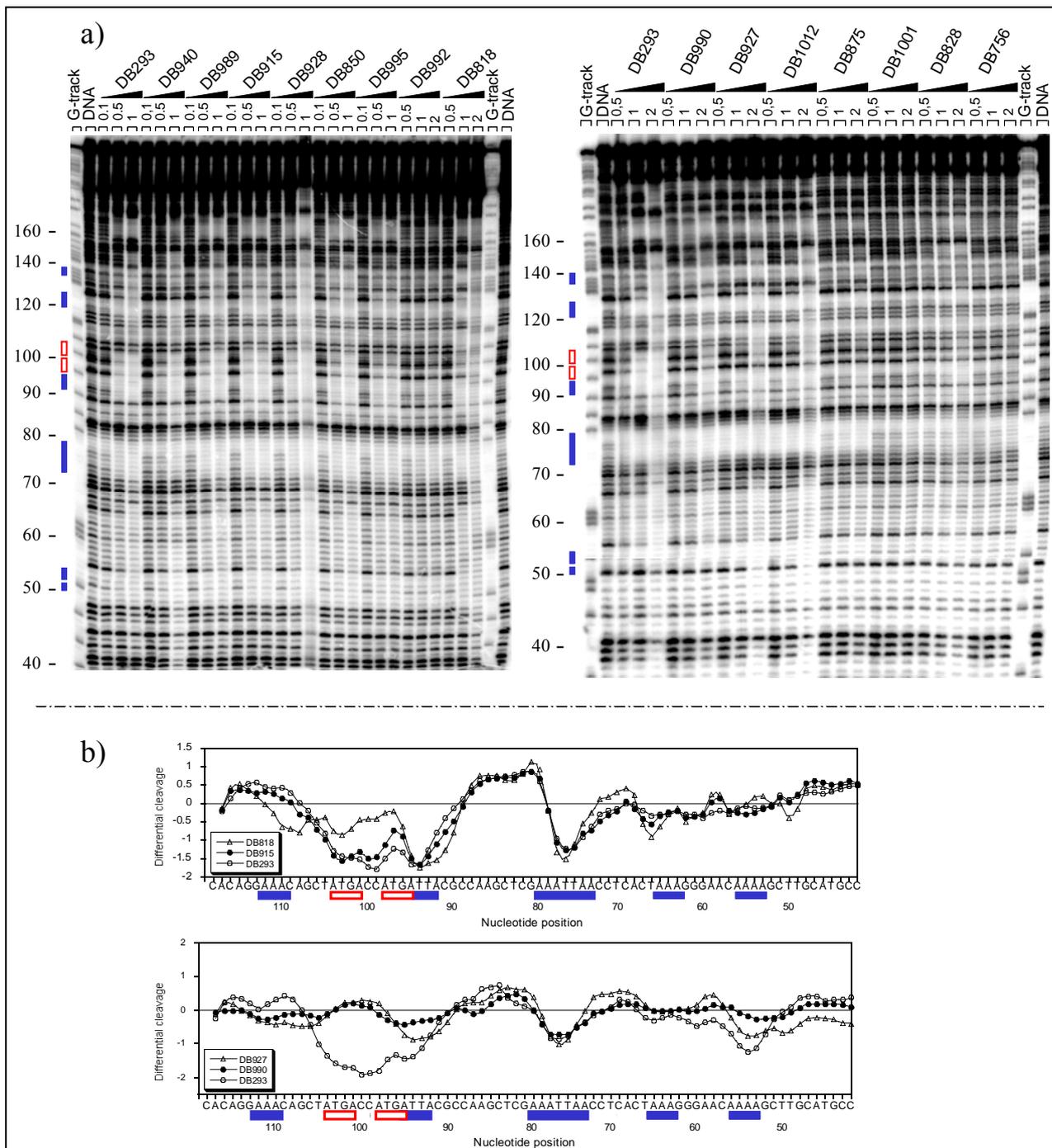


Figure 34 : Empreinte à la DNase I des composés phényl-furane-benzimidazoles.

a) Empreintes à la DNase I des composés du type phényl-furane-benzimidazole. Un fragment d'ADN de 265 pb, obtenu par digestion EcoR I-Pvu II du plasmide pBS marqué en 3', est incubé ou non pendant 15 min à TA avec une gamme de dérivés DB (en μM) puis digéré par la DNase I.

b) Analyse densitométrique des cinq dérivés DB293, DB915, DB818, DB927 et DB990. Les rectangles bleus représentent les séquences riches en paires de bases AT et les rectangles rouges les sites ATGA.

b) Affinité relative des composés pour l'ADN

Pour déterminer l'affinité relative pour l'ADN des composés précédemment testés, des expériences de dénaturation thermique de l'ADN ont été réalisées. L'analyse des tests d'empreinte à la DNase I a permis de sélectionner en plus de l'oligonucléotide AT, l'oligonucléotide ATGA : ACAGCT**ATGACCATGATT**ACGCC. Cette séquence du plasmide pBS située entre les paires de bases 88 et 110 est particulièrement intéressante car elle contient à la fois le site 5'-ATGA et le site riche en paires de bases AT (5'-ATTA) (Figure 35a et b).

La comparaison des tests de variations de la température de fusion réalisés sur ces deux oligonucléotides montre clairement que les molécules qui reconnaissent les séquences ATGA possèdent une meilleure affinité relative pour ce site par rapport aux sites riches en paires de bases AT. La valeur moyenne de ΔT_m est proche de 18°C pour l'oligonucléotide ATGA et de 9,8°C pour l'oligonucléotide AT. On peut noter que pour l'oligonucléotide AT, ces molécules possèdent une affinité relative comparable aux composés qui se lient aux sites riches en paires de bases AT seuls. Leur capacité de fixation aux séquences ATGA ne modifie donc pas leur capacité de se fixer aux séquences AT.

Pour les composés qui se lient aux sites riches en paires de bases AT, on observe et de façon prévisible que le DB818 possède une affinité relative forte pour l'ADN, semblable pour les deux oligonucléotides avec un ΔT_m moyen de 18°C. D'autre part, les DB927 et DB1012 possèdent une affinité relative pour l'ADN également élevée (ΔT_m moyen de l'ordre de 13,5°C). Les composés DB992, DB990 et DB875 qui reconnaissent les sites riches en paires de bases AT sont moins affins pour l'ADN avec des valeurs de ΔT_m situées autour de 6,5°C.

Jusqu'à présent, seul le DB293 était connu pour interagir sur les séquences 5'-ATGA avec une forte affinité pour l'ADN. Cette étude montre que d'autres produits ont la capacité de se fixer à cette séquence avec des affinités comparables à celle du DB293.

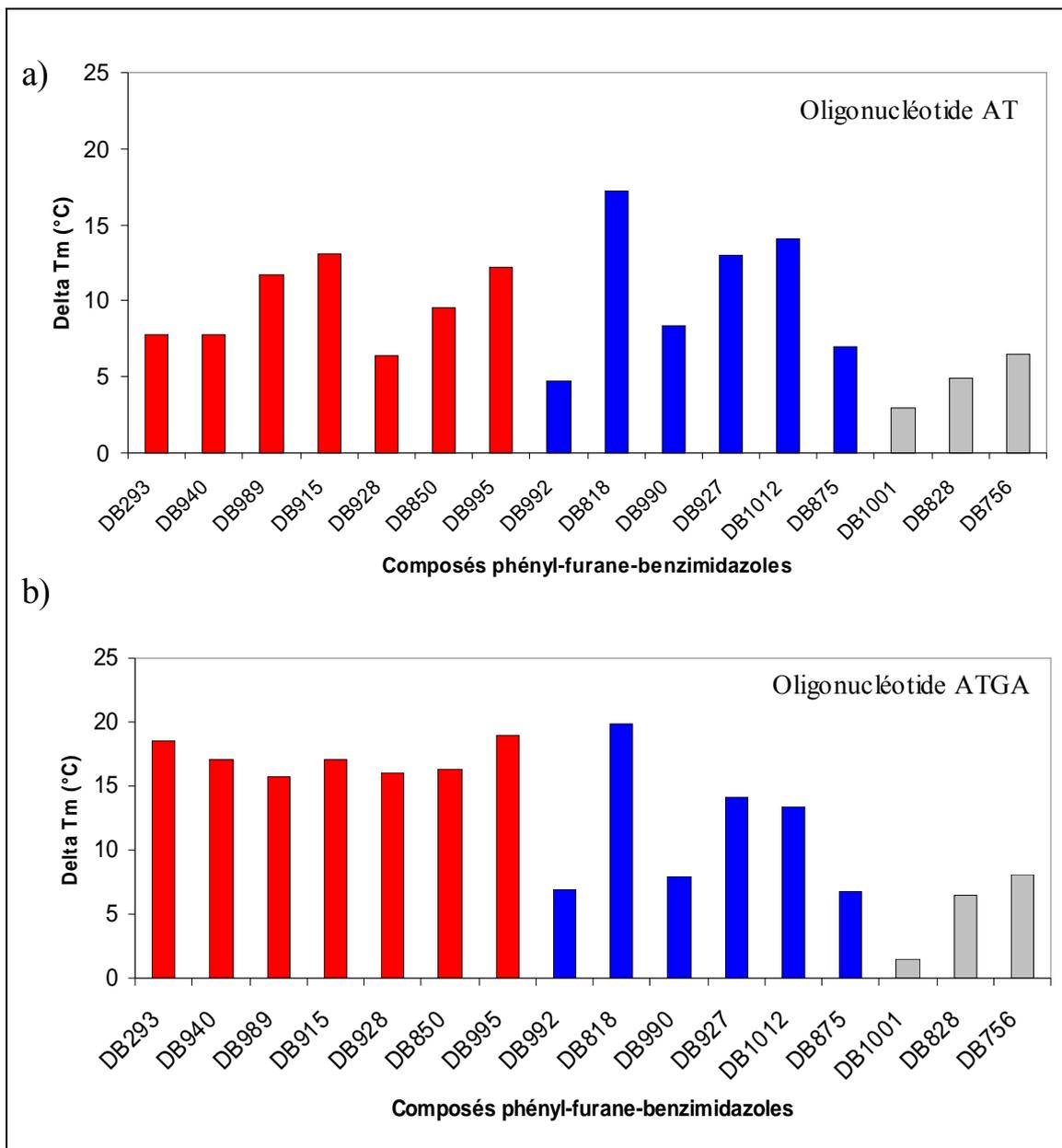


Figure 35 : Diagramme représentant les variations de T_m des dérivés phényl-furane-benzimidazoles.

Les composés DB à la concentration de $2 \mu\text{M}$ sont incubés avec $1 \mu\text{M}$ d'oligonucléotide double brin de séquence a) 5'-GTGAGGGTTAATTCGAGCTTGG-3' et b) 5'-ACAGCTATGACCATGATTACGCC-3'. En rouge sont représentés les composés DB qui reconnaissent les sites ATGA, en bleu les molécules qui se lient uniquement aux sites riches en paires de bases AT et en gris les composés non spécifiques.

c) Relations structure/sélectivité et structure/affinité

(1) Modification du cycle benzimidazole

Ces résultats confortent les résultats obtenus précédemment par notre groupe et celui du Pr David Wilson (Wang et al., 2001). Les substitutions sur le cycle benzimidazole abolissent l'interaction sous forme de dimère ; ceci avait été montré pour le composé DB609. L'utilisation de six autres composés DB : DB1012, DB927, DB875, DB1001, DB828 et DB756 valide cette hypothèse.

(2) Modification du cycle phényle

De façon intéressante, il existe un plus grand degré de liberté en ce qui concerne les modifications sur le cycle phényle et surtout les modifications en position ortho. Ainsi, les composés DB850 et DB928 substitués en position ortho respectivement par un groupement hydroxyle ou méthyle se lient à l'ADN de façon coopérative aux sites 5'-ATGA alors que les composés équivalents DB990 et DB992 substitués en position méta perdent cette capacité. Il existe tout de même une exception, le DB989 substitué par un fluor en position méta interagit sous forme de dimère au site 5'-ATGA.

Il est également possible de remplacer le cycle phényle par un cycle pyridine sans altérer la reconnaissance spécifique du composé.

(3) Modification du cycle furane

Comme il a été démontré précédemment, le DB818 interagit uniquement au niveau des sites riches en paires de bases AT avec une affinité très forte pour l'ADN (Mallena *et al.*, 2004).

Le DB940 garde quant à lui la capacité d'interagir sur les sites 5'-ATGA. Cette molécule, où le cycle furane est remplacé par un cycle imidazole, possède un angle de courbure très proche du DB293 ; en effet, l'atome d'azote possède un encombrement stérique proche de l'atome d'oxygène.

B. Nouveaux composés : Nouveaux sites de reconnaissance

Les résultats précédents ont permis de déterminer deux séries de molécules capables de reconnaître les séquences ATGA et/ou les sites riches en paires de bases AT de l'ADN. Dans le but de découvrir de nouveaux composés DB qui reconnaîtraient de nouvelles séquences d'ADN, diverses séries de dérivés ont été synthétisés par les Prs David Boykin et David Wilson à Atlanta (Figure 36a, 37a et 38a).

1. DB1242 et séquence GCTCG

Pour caractériser sur le plan moléculaire ces nouveaux dérivés DB capables de reconnaître des séquences cibles originales, nous avons également effectué des tests d'empreinte à la DNase I sur des molécules du type triphényle (Figure 36a). Ces molécules, analogues du composé diphényl-benzimidazole DB921 (Figure 18), sont elles aussi totalement planes.

Le gel présenté en Figure 36b montre clairement un site de reconnaissance unique pour le DB1242 (rectangle jaune) qui n'a été retrouvé jusqu'à présent pour aucune autre molécule, alors que les composés DB1111, DB1164 et DB1128 reconnaissent les sites riches en paires de bases AT.

L'analyse densitométrique (Figure 36c) confirme ces résultats et met en évidence que les composés DB1111, DB1164 et DB1228 interagissent avec l'ADN de manière identique au DB75. Ces molécules ciblent les sites AAATTAA (p 74-80), ATTA (p 93 à 96), TTAA (p 121 à 125) et TTAA (p 139-142). Le DB1242 possède quant à lui une empreinte totalement différente de celle obtenue avec le DB75 avec un site d'interaction majeur centré sur la séquence GCTCG (p 81 à 85). Il existe un second site GCTCG lui aussi reconnu par le DB1242 situé en position 155 à 159. Ce site n'apparaît pas sur l'analyse densitométrique du gel. Les bandes, issues de la digestion à la DNase I, sont difficilement dissociables les unes des autres à ce niveau du gel.

Cette séquence reconnue par le DB1242, riche en paire de bases GC, est totalement nouvelle et n'est à ce jour reconnue par aucun autre dérivé DB. Le mode de fixation de ce composé tout à fait original a été élucidé récemment et sera développé ultérieurement.

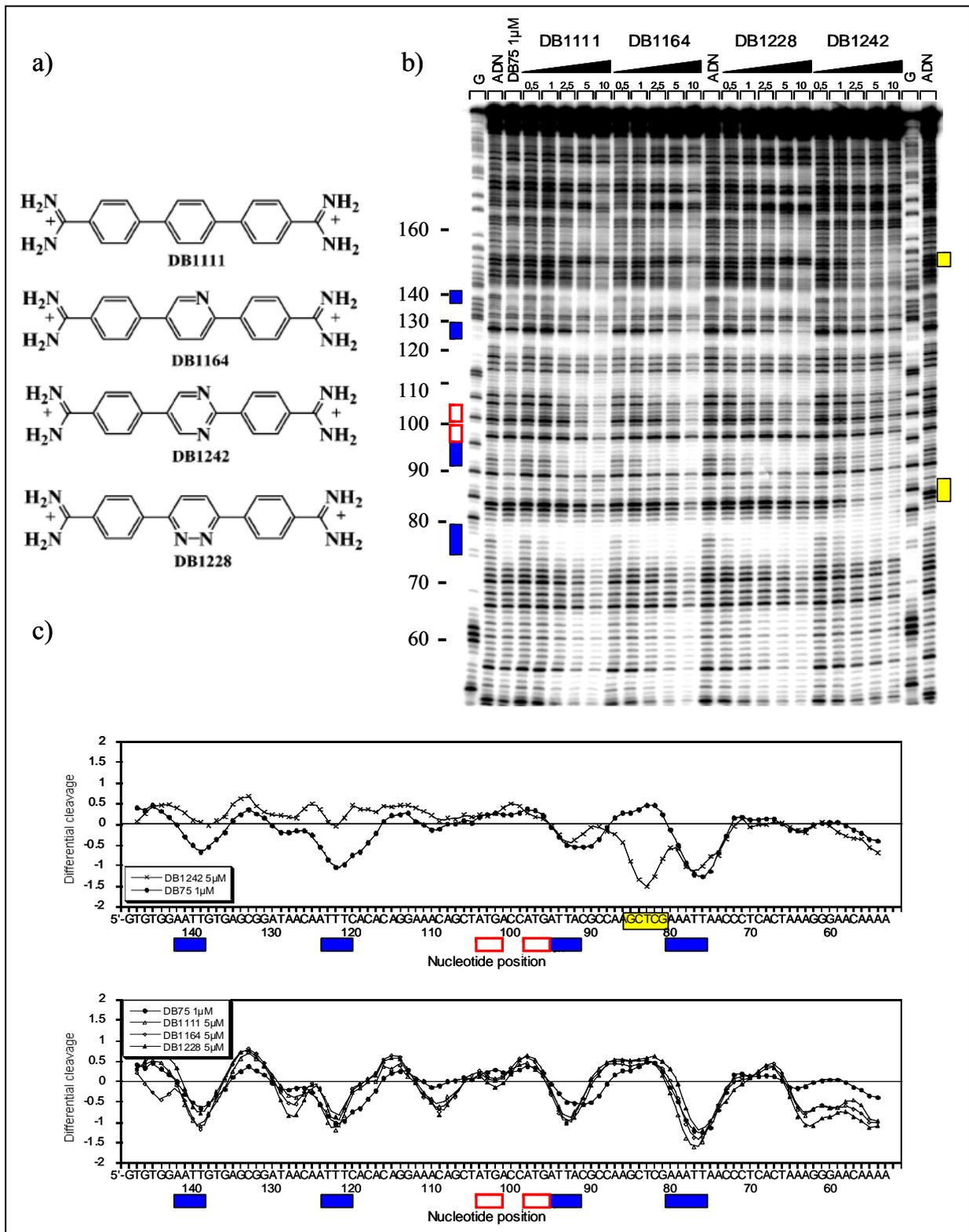


Figure 36 : Test d'empreinte à la DNase I des composés triphényles.

a) Structure des composés plans du type triphényle. b) Tests d'empreinte à la DNase I réalisés sur un fragment de 265 pb, issu de la digestion EcoR I-Pvu II du plasmide pBS. Après incubation du fragment d'ADN avec des concentrations croissantes en composés DB et digestion ménagée à la DNase I, les échantillons sont séparés par électrophorèse dénaturante. c) Analyses densitométriques du gel. Les rectangles bleus représentent les séquences riches en paires de bases AT, les rectangles rouges les sites ATGA et les rectangle jaunes les séquence GCTCG.

2. DB1255 et séquence CCATG

Des composés possédant une structure plus longue caractérisée par la présence de deux ou trois cycles centraux au lieu d'un seul et susceptibles de reconnaître des sites de reconnaissance supérieurs à 4 paires de bases ont été synthétisés (Figure 37a). Les expériences d'empreinte à la DNase I mettent en évidence là encore un nouveau profil de reconnaissance pour le DB1255 (Figure 37b et c). Alors que la plupart des composés testés reconnaissent les sites riches en paires de bases AT ou reconnaissent l'ADN de manière non spécifique, le DB1255 présente une empreinte particulière (rectangle rose). La comparaison réalisée entre les zones de reconnaissance du DB75 (rectangles bleus) et du DB293 (rectangles rouges) montre clairement que le DB1255 reconnaît des sites riches en paires de bases AT et un nouveau site de reconnaissance qui juxtapose les sites ATGA.

Les analyses densitométriques différentielles (Figure 36d et e) révèlent que le DB1255 reconnaît les sites riches en paires de bases AT : AAAA (p 53-56) ; AAATTAA (p 74-80) ; ATTA (p 93 à 96) et un dernier site beaucoup plus original centré sur le site CCATG. (p 96-100). On observe très clairement que le site de fixation de ce composé n'englobe pas les deux sites ATGA typique du DB293 (Figure 36e). Des études complémentaires d'empreintes à la DNase I et de résonance plasmonique de surface sont actuellement en cours pour déterminer le mode d'interaction, ainsi que les sites de reconnaissance précis de ce nouveau composé.

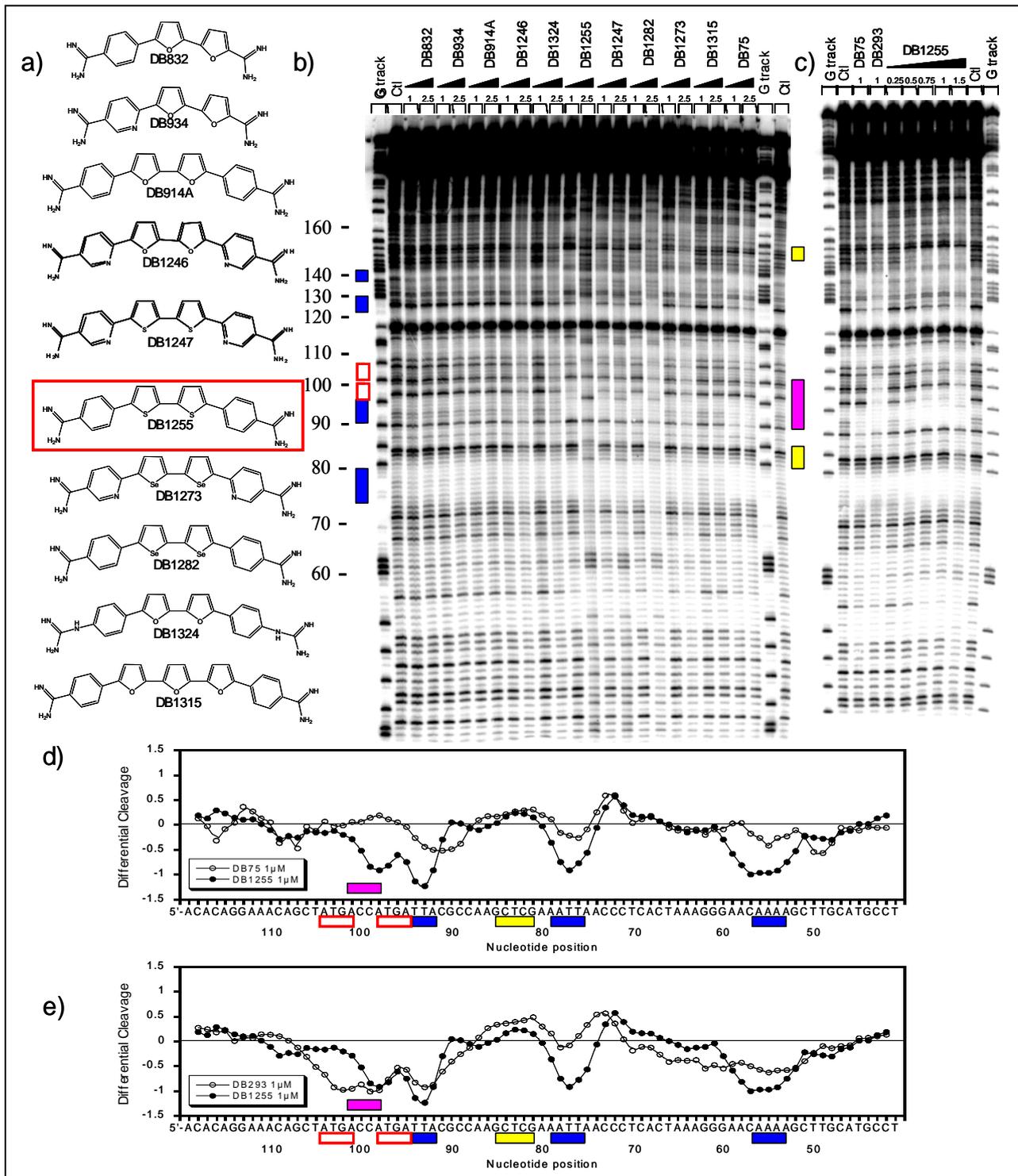


Figure 37 : Test d'empreinte à la DNase I des composés du type diphenyl-difurane.

a) Structure des composés du type diphenyl-difurane. b) Empreinte à la DNase I réalisée sur un fragment de 265 pb, issu de la digestion *EcoRI-PvuII* du plasmide pBS avec les composés diphenyl-difuranes (en μM) c) Empreinte à la DNase I réalisée sur des concentrations croissantes en DB1255 comparées aux DB75 et DB293 à 1 μM . Les rectangles bleus correspondent aux sites riches en paires de bases AT, les rouges aux sites ATGA, les jaunes aux sites GCTCG et en violets le site ACGAT. d) et e) Analyses densitométriques des composés DB75 (d) et (e) DB293 comparés au DB1255.

3. RT-29 et site ATTC

Certaines molécules qui dérivent des composés DB ont été précédemment étudiées au laboratoire. Le RT-29, diphényl-éther-benzimidazole (Figure 38a), possède une structure plus courbe que celle du petit sillon de l'ADN. Pourtant, cette molécule a montré une étonnante capacité de liaison à l'ADN. Les expériences d'empreinte à la DNase I avaient montré une reconnaissance spécifique du site TAAC. Des expériences de résonance plasmonique de surface laissaient présager d'une meilleure reconnaissance du site ATTC par le RT-29. Afin de valider ces résultats, nous avons construit à partir du plasmide pUC19 un fragment d'intérêt contenant le site ATTC. Les tests d'empreinte à la DNase I ont été effectués sur deux fragments d'ADN différents : un premier qui possède de nombreux sites riches en paires de bases AT (rectangles bleus) et le site TAAC (rectangle gris) (Figure 38b) et un second, qui possède le site ATTC (rectangle vert) (Figure 38c). Ces tests révèlent divers sites de fixation pour ce composé. L'utilisation du fragment de 117 pb montre une fixation du composé sur les sites riches en paires de bases AT (Figure 38b et d) situés entre les positions AATTGTAATA (p 20–30), TTTT (p 63-66), TAAC (p 74–77) et ATTAA (p 84 à 88).

L'analyse densitométrique du gel (Figure 38c et e) montre clairement une reconnaissance sur le site GAAT (p 43 à 46) (rectangle vert) (séquence complémentaire du site ATTC).

Cette première partie de l'étude nous a permis de déterminer plusieurs classes de composés capables de reconnaître des séquences différentes variées. Ces molécules pourront être utilisées pour l'étude des conséquences cellulaires et la modulation de l'activité de liaison à l'ADN de divers facteurs de transcription.

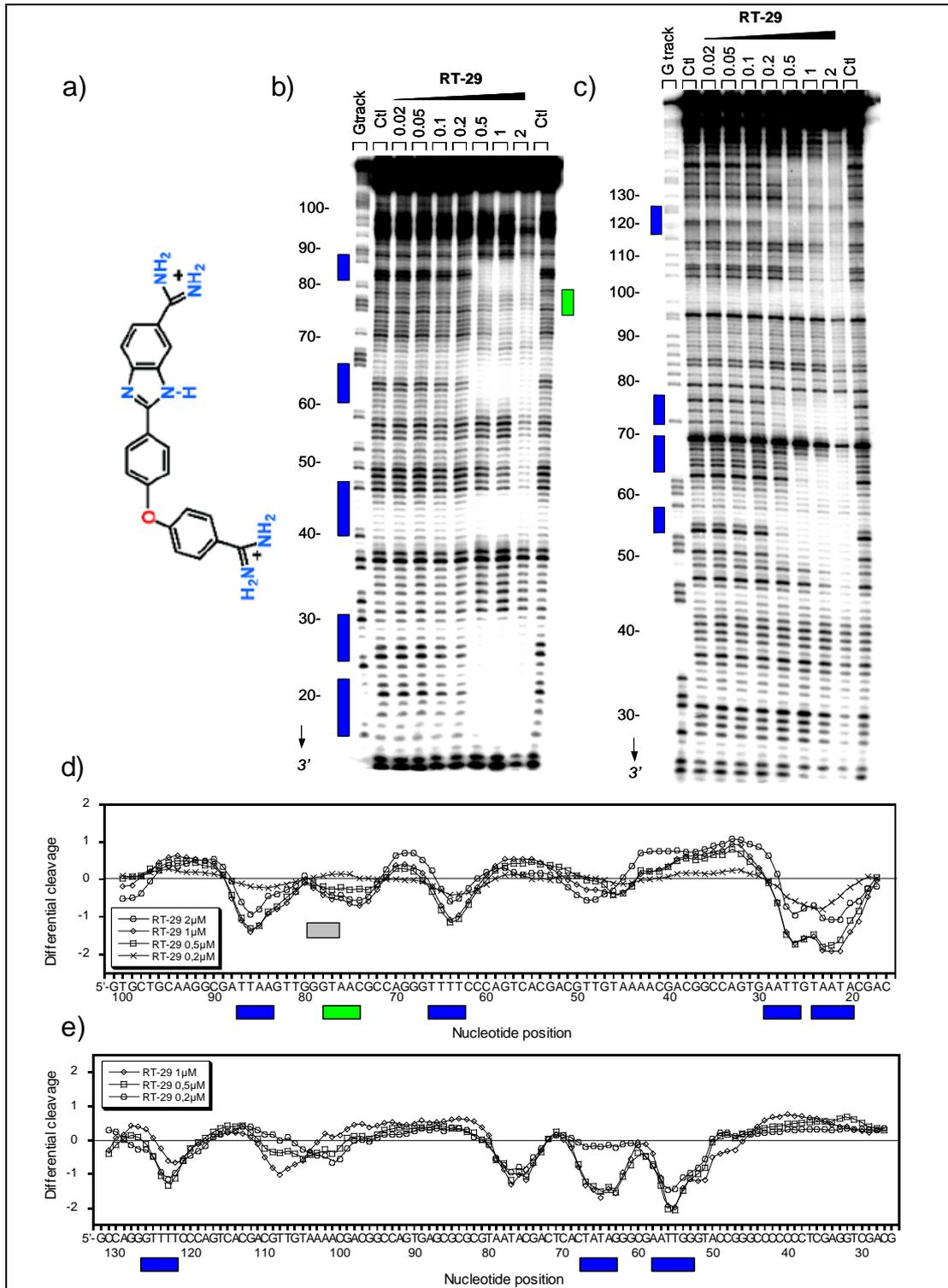


Figure 38 : Test d'empreinte à la DNase I du composé RT-29.

a) Structure du RT-29. b) et c) Tests d'empreintes à la DNase I réalisés sur un fragment de 117 pb, issu de la digestion *EcoR* I-*Pvu* II du plasmide pBS (b) et sur un fragment de 110 pb issu de la digestion *EcoR* I-*Pst* I du plasmide pUC19 ATTC (c). Après incubation du fragment d'ADN avec des concentrations croissantes en composé RT-29 et digestion ménagée à la DNase I, les échantillons sont séparés par électrophorèse dénaturante. d) et e) Analyses densitométriques issues des gels b) et c) respectivement. Les rectangles bleus correspondent aux sites riches en paires de bases AT et le rectangle vert correspond au site TAAC.

C. Nouveaux composés : Distribution et toxicité cellulaires

Pour atteindre leur cible ADN, les composés moléculaires doivent traverser la membrane cytoplasmique et pénétrer dans le noyau. Les études menées auparavant au laboratoire se sont focalisées sur l'influence de la nature des groupements amidines terminaux de ces molécules sur leur distribution cellulaire (études de relation de structure/localisation). En effet, ces études ont montré le rôle déterminant de ces éléments pour la pénétration cellulaire et la localisation des dérivés DB. Grâce aux propriétés de fluorescence intrinsèque des molécules (longueur d'onde d'excitation située entre 360 et 380 nm et une longueur d'onde d'émission située entre 450 et 500 nm), l'équipe a étudié les relations de structure - localisation des analogues de la furamide par microscopie de fluorescence et microscopie confocale. Dans la série diphényl-furane, les auteurs ont clairement démontré que la localisation cellulaire entre les composés mono-(DB607) et tétra-cationique (DB182) était différentielle. Les mêmes résultats sont retrouvés dans la série phényl-furane-benzimidazole (Lansiaux *et al.*, 2002a). Le second article démontre quant à lui qu'en fonction des substitutions des molécules (présence d'un groupement phényle au niveau du groupement amidine), il y a possibilité de cibler les mitochondries des cellules tumorales (Lansiaux *et al.*, 2002b). Comme les chimistes à Atlanta ont synthétisé de nouvelles molécules substituées sur le corps polycyclique pour lesquelles nous venons de caractériser les modes de reconnaissance de l'ADN, nous avons recherché en microscopie de fluorescence si ces nouvelles molécules étaient potentiellement capables d'atteindre leur cible. Nous avons alors analysé l'influence des substituants portés par le corps polycyclique des dérivés diphényl-furanes et phényl-furane-benzimidazoles (Figure 31) sur la distribution des dérivés DB. Dans ce but, la localisation des composés DB au sein de la cellule a été étudiée en utilisant des marqueurs spécifiques de compartiments cellulaires : le DiOC₆ comme marqueur cytoplasmique et l'IP comme marqueur nucléaire.

1. Influence des substituants portés par le corps de la molécule sur la distribution cellulaire

De manière générale, pour les composés DB nucléaires, on observe une fluorescence bleue intrinsèque dans le noyau qui colocalise avec la fluorescence rouge de l'IP, alors que pour les composés DB cytoplasmiques, la fluorescence bleue est colocalisée avec la fluorescence verte du DiOC₆ dans le cytoplasme. Sur la Figure 39a, on observe que les dérivés diphényl-furanes de la famille du DB75 possèdent une localisation nucléaire à l'exception du DB936 (substitué

par deux groupements méthyles en position méta des cycles phényles). Concernant les composés phényl-furane-benzimidazoles du type DB293 (Figure 39b), ils possèdent tous une localisation nucléaire. Aucune des substitutions effectuées sur ce produit n'apporte de modification de localisation.

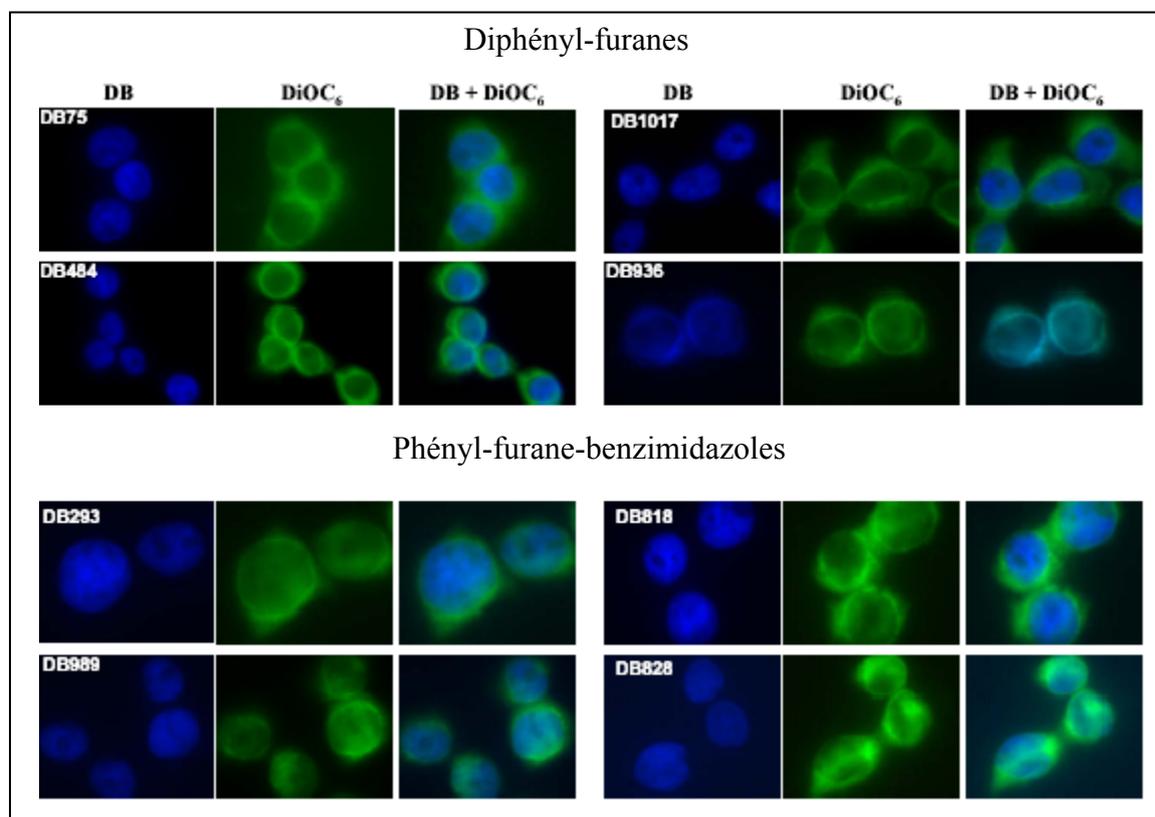


Figure 39 : Distribution cellulaire des dérivés DB.

Microscopie de fluorescence de cellules HT-29 (x 63) traitées 16 h par 5 μM de composés a) diphényl-furanes ou b) phényl-furane-benzimidazoles. Les cellules sont ensuite fixées avec une solution de PFA 2% et incubées avec 75 nM de DiOC₆ (en vert). Les images de droite montrent la superposition du DiOC₆ et du dérivé DB. Seuls six produits sont représentés mais tous hormis le dérivé DB936 possèdent une localisation nucléaire.

Ces résultats montrent que les modifications apportées sur le corps polycyclique ont peu d'importance sur la pénétration et la distribution cellulaire des composés DB à l'exception du DB936 qui possède une localisation cytoplasmique et non nucléaire.

2. Influence du mode de fixation des cellules dans la localisation subcellulaire des composés

Ces expériences de distribution cellulaire sont effectuées sur des cellules fixées au paraformaldéhyde (PFA). Cette étape permet de mieux conserver les échantillons. Cependant, il a été montré pour les polyamines que la distribution cellulaire peut être modifiée par le

mode de fixation des cellules. Ainsi, la fixation au méthanol des cellules SKBR-3 permet aux composés testés N-méthylpyrrole/N-méthylimidazole qui possèdent une localisation cytoplasmique sur des cellules non fixées de se relocaliser dans le noyau (Belitsky *et al.*, 2002).

Nous avons donc vérifié le rôle de la fixation au PFA sur la localisation des composés DB. Dans cette optique, d'autres modes de fixation ont été testés. Pour cela, les cellules HT-29 ont été incubées avec le DB75 (nucléaire) ou le DB936 (cytoplasmique). Après 16h d'incubation, les cellules ont été fixées ou non avec différentes solutions de concentrations croissantes de PFA (Figure 40) ou de méthanol (Figure 41).

Nos résultats montrent clairement que la fixation avec des solutions de 1 à 4 % de PFA aux propriétés de fixation sélectives ne modifie pas la localisation cellulaire du DB75 et du DB936. En effet, le DB75 qui possède une localisation nucléaire sur les cellules non fixées reste dans le noyau même après fixation avec 4% de PFA. De manière identique, le DB936 qui se localise dans le cytoplasme des cellules non fixées reste cytoplasmique après fixation avec 4 % de PFA. Ce mode de fixation ne modifie donc pas la localisation des composés DB.

Le méthanol, aux propriétés de fixation et de perméabilisation, influe sur la localisation des composés. Ainsi, tandis que le composé nucléaire DB75 conserve sa localisation quelque soit la concentration en méthanol, le composé cytoplasmique DB936 se délocalise dans le noyau (colocalisation de la fluorescence bleue du composé DB936 avec la fluorescence rouge de l'IP). Ce phénomène est observé même pour des concentrations faibles en méthanol (à partir de 25 %). Ces résultats concordent avec ceux de la littérature et notamment les résultats obtenus par l'équipe de Dervan sur la distribution cellulaire des polyamines (Belitsky *et al.*, 2002). Ce résultat pourrait s'expliquer par la perméabilisation des membranes induite par le méthanol.

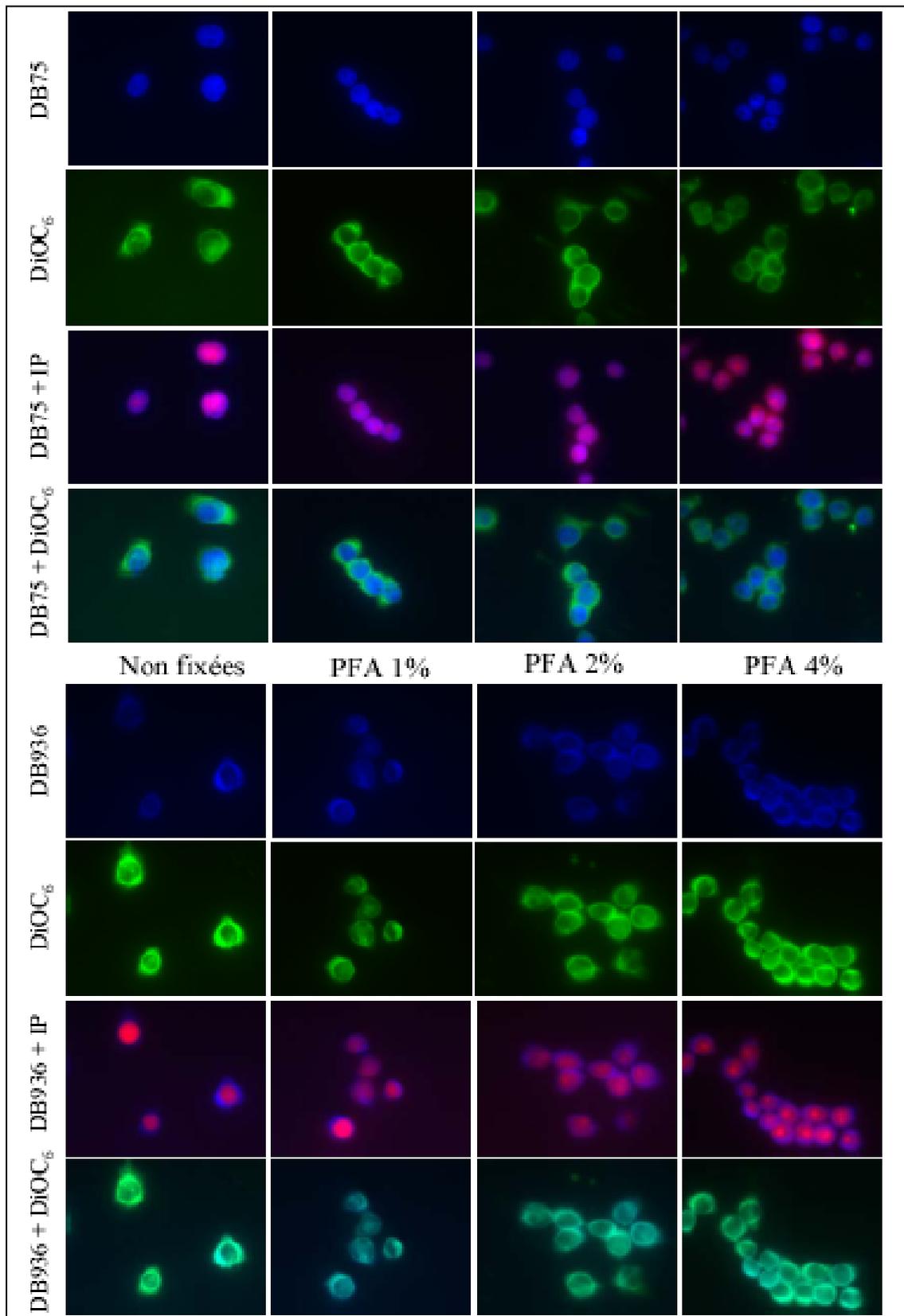


Figure 40 : Distribution cellulaire des cellules HT-29 fixées au PFA.

Microscopie de fluorescence (x63) de cellules HT-29 incubées avec 5 μ M de DB75 ou de DB936 (en bleu) puis fixées ou non avec une solution de PFA de 1 à 4%. Les cellules sont ensuite incubées dans une solution de DiOC₆ à 75 nM (en vert) puis avec 1 μ g/ml de PI (en rouge).

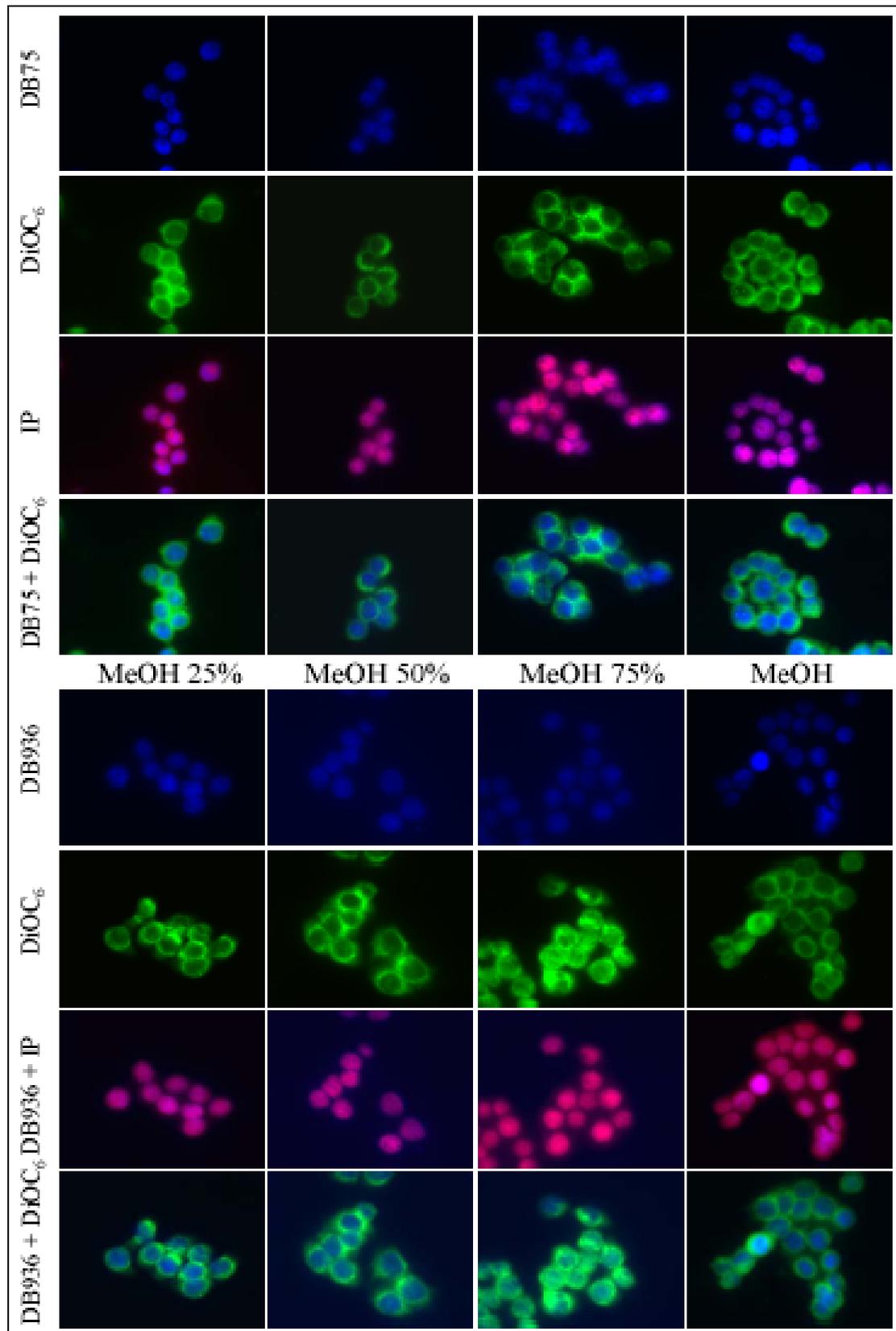


Figure 41 : Distribution cellulaire des cellules HT-29 fixées au Méthanol.

Microscopie de fluorescence (x63) de cellules HT-29 incubées avec 5 μ M de DB75 ou de DB936 (en bleu) puis fixées avec une solution de méthanol de 25 à 100 %. Les cellules sont ensuite incubées dans une solution de DiOC₆ à 75 nM (en vert) puis avec 1 μ g/ml de PI (en rouge).

Pour vérifier cette hypothèse, des expériences complémentaires ont été effectuées. Pour cela, après incubation des cellules avec les DB75 et DB936, les cellules sont fixées avec une solution de NP40 (0,1 à 2 %) dans du PFA à 2 % (Figure 42).

Cette solution est classiquement utilisée en immunohistochimie. Elle permet la fixation des cellules par le PFA et la perméabilisation des membranes par le détergent NP40. Comme dans le cas du méthanol, la fixation au NP40-PFA ne modifie pas la distribution du composé nucléaire, même pour des concentrations importantes en NP40 (2 %). En ce qui concerne le composé cytoplasmique DB936, la fixation au NP40 2 % dans une solution de PFA à 2% modifie sa distribution subcellulaire : le dérivé se relocalise dans le noyau.

En accord avec la littérature, ces résultats montrent que l'analyse de la localisation cellulaire des molécules intrinsèquement fluorescentes après simple fixation des cellules (PFA 2% par exemple) ne modifie en rien la localisation des dérivés DB. Elle a pour avantage d'améliorer la qualité de l'analyse par une meilleure conservation des lames. En revanche et de manière prévisible, il n'est pas recommandé de modifier la perméabilité membranaire (par du méthanol et du NP40) pour analyser la localisation de nouvelles molécules.

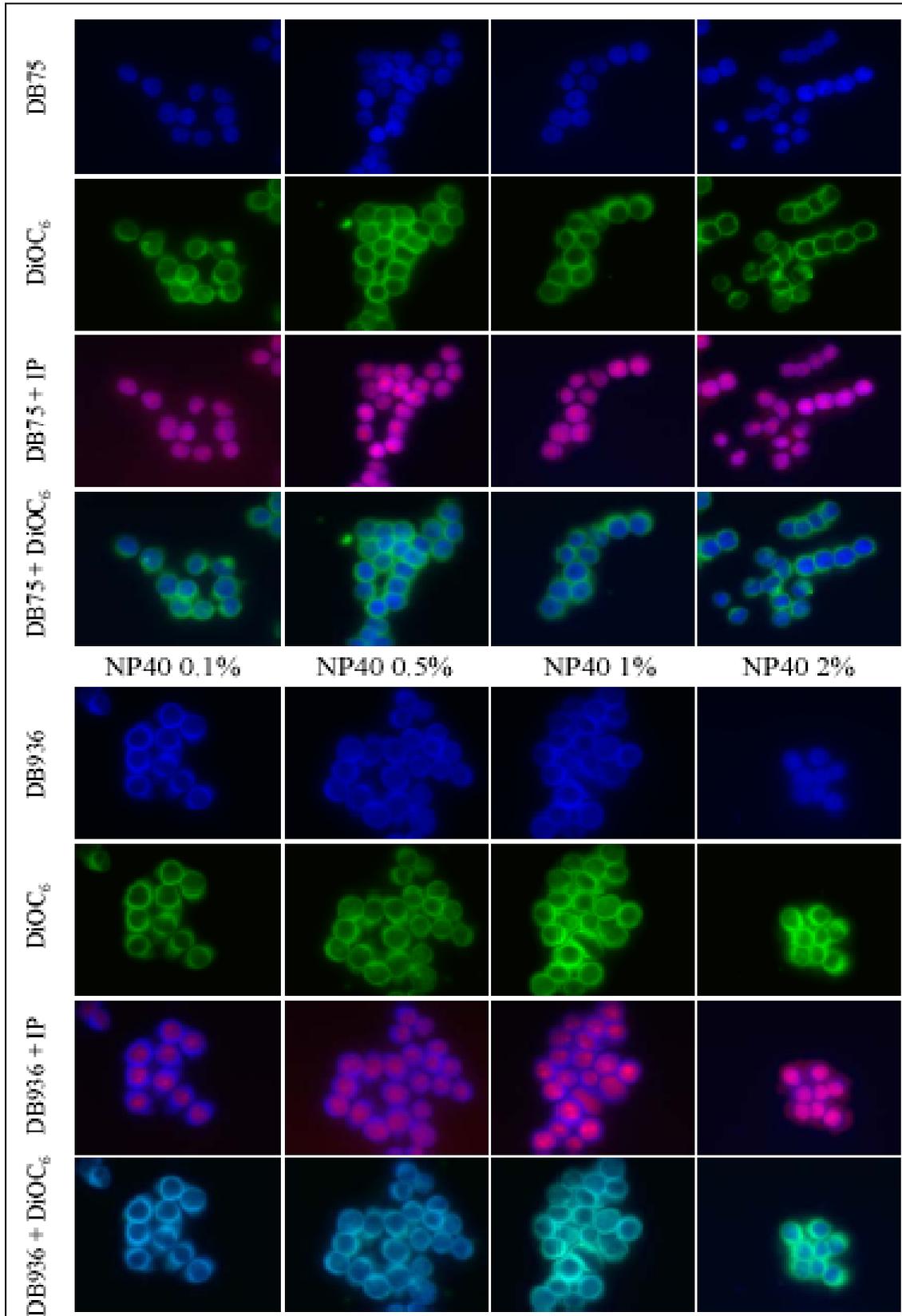


Figure 42 : Distribution cellulaire des cellules HT-29 fixées au NP40 –PFA2%.

Microscopie de fluorescence (x63) de cellules HT-29 incubées avec 5 μ M de DB75 ou de DB936 (en bleu) puis fixées ou non avec une solution de NP40 dans 2 % PFA. Les cellules sont ensuite incubées dans une solution de DiOC₆ à 75 nM (en vert) puis avec 1 μ g/ml de PI (en rouge).

D. Effets cytotoxiques des composés DB

Afin d'évaluer la toxicité des dérivés DB, des tests de cytotoxicité au réactif MTS ont été réalisés sur les cellules HT-29 (Tableau 1). Pour cela un panel de composés appartenant aux différentes séries testées a été utilisé.

Séries	Composés	IC ₅₀ (µM)
Diphényl-furane	DB75	16 ± 1
	DB351	18.5 ± 1.5
	DB484	29 ± 2
	DB544	25.5 ± 2.5
	DB936	3.8 ± 0,85
Phényl-furane-benzimidazole	DB293	>40
	DB995	33.4 ± 0.5
	DB818	35.5 ± 2.5
	DB828	>50
Autres	DB1242	>50
	DB1255	7.7 ± 3.8
	RT-29	47.7 ± 2.4

Tableau 1 : IC₅₀ d'un panel de composés DB.

Tests réalisés sur des cellules HT-29. Les valeurs d'IC₅₀ obtenues à 72 h sont indiquées en µM. Les manipulations sont effectuées en double afin de calculer les écarts types.

Les résultats de cytotoxicité obtenus pour ces différents composés montrent tout d'abord une faible toxicité des composés diphényl-furanes et phényl-furane-benzimidazoles à l'exception du DB936 sur la lignée cellulaire HT-29. On remarque que cette toxicité est supérieure pour les composés de la série diphényl-furane par rapport à la série phényl-furane-benzimidazole. En effet, les DB75 et DB351 possèdent une toxicité de l'ordre de 17 µM alors que leurs analogues respectifs DB293 et DB818 possèdent une toxicité supérieure à 35 µM.

Le composé qui possède une localisation cytoplasmique, le DB936, est plus toxique pour la cellule que les autres composés avec une valeur d'IC₅₀ de 3.8 µM. Ceci a d'ailleurs été montré dans des études antérieures pour des composés de cette famille de dérivés (Lansiaux *et al.*, 2002a ; 2002b). En ce qui concerne les nouvelles molécules, on remarque que seul le composé DB1255 possède une cytotoxicité moyenne de l'ordre de 8 µM.

Ces résultats sont encourageants, puisque les molécules étudiées présentent un effet cytotoxique faible (supérieur à 1 μM).

III. Modulation de la fixation des facteurs de transcription à l'ADN par les composés DB

Dans cette seconde partie du travail, l'ensemble des résultats présentés et publiés permettent de décrire assez précisément les interactions entre l'ADN et les composés DB. Aucune étude, à l'instar des lexitropsines, n'avait mis en évidence la modulation de l'interaction des facteurs de transcription à l'ADN par ces dérivés. C'est pourquoi nous avons posé l'hypothèse que les composés DB pourraient moduler l'activité de liaison des facteurs de transcription au regard de leurs sélectivités de reconnaissance de l'ADN.

Pour cela, nous avons utilisé deux approches :

- Tout d'abord, un screening large de facteurs de transcription à l'aide du kit Transignal protein/DNA arrays. Cette méthode permet d'étudier la modulation de l'activité de liaison à l'ADN de 54 facteurs de transcription exprimés dans une lignée cellulaire donnée selon un mode compétitif.
- Dans un second temps, nous avons réalisé une étude fine sur le facteur de transcription HoxA9. Cette protéine interagit en hétérodimère avec PBX sur le site 5'-ATGATTTA.

A. Modulation d'un ensemble de facteurs de transcription par le DB293

1. TranSignal™ Protein/DNA Array I

Pour cette étude nous avons opté pour un composé capable d'interagir avec la séquence 5'-ATGA. Comme montré précédemment, les composés DB293, DB940, DB989, DB915, DB928, DB850 et DB995 possèdent une affinité relative comparable pour le site 5'-ATGA. Le DB293 reconnaît les séquences 5'-ATGA sous forme de dimère et les sites riches en paires de bases AT sous forme de monomère au niveau du petit sillon de l'ADN. C'est pour cette raison que nous l'avons sélectionné pour cette étude, en tant que chef de file des composés qui se fixent aux séquences 5'-ATGA.

Le kit TranSignal™ Protein/DNA Array I est basé sur le principe des macroarrays et est classiquement utilisé pour comparer le niveau d'expression de différents facteurs de transcription sur des lignées cellulaires différentes ou sur des lignées ayant subi ou non un

traitement (Lam et Li, 2002). Par cette technique nous avons déterminé les effets du DB293 sur la fixation concomitante de 54 facteurs de transcription à leur site consensus.

Ce kit contient 54 sondes biotinylées, reconnues spécifiquement par autant de facteurs de transcription. Ces oligonucléotides peuvent être hybridés sur des membranes où sont greffés les fragments d'ADN complémentaires à ceux utilisés précédemment. Chaque sonde est immobilisée en double exemplaire à une concentration standard et à une concentration au 1/10 pour éviter certains problèmes de saturation du signal (Figure 43).

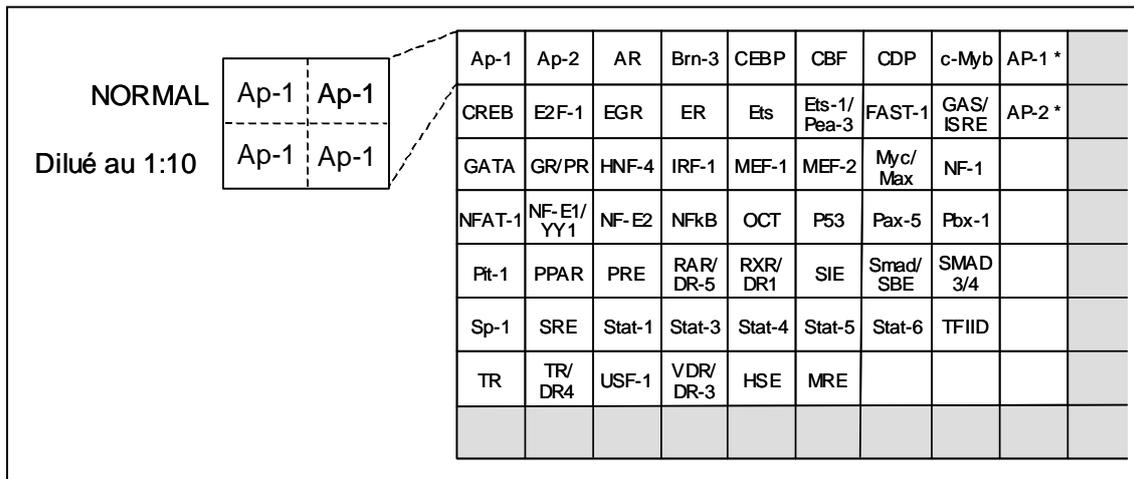


Figure 43 : Organisation des oligonucléotides sur la membrane.

Chaque sonde reconnue est hybridée en 4 exemplaires, diluée (2) ou non au 1/10 (2). En gris sont représentés les contrôles internes qui permettent de comparer les résultats entre les membranes.

Les résultats obtenus sont répertoriés Figure 44. On constate tout d'abord qu'à 5 µM la grande majorité des facteurs de transcription sont inhibés par le DB293 d'un facteur d'environ 1,7. A cette concentration, il peut exister une faible fixation non spécifique du composé qui entraînerait une inhibition globale de la fixation des facteurs de transcription. Ainsi, pour cette concentration nous nous sommes intéressés uniquement aux facteurs de transcription inhibés avec un facteur supérieur à 3.

Certains facteurs de transcription sont inhibés par le DB293 de manière importante. Les facteurs de transcription Pit-1 et Brn-3 appartiennent à la famille des facteurs de transcription à POU domaine et sont inhibés respectivement d'un facteur 5 et 3 à la concentration de 5 µM. Ces deux protéines possèdent dans leur site consensus la séquence ATGA. Par opposition, le facteur de transcription IRF-1 qui possède lui aussi le site ATGA et des sites riches en paires de bases AT n'est pas inhibé par le composé DB293 (inhibition non-spécifique de 1,4 fois à 5 µM).

On observe que le DB293 inhibe la fixation à l'ADN d'autres facteurs de transcription comme le montre la Figure 44d. Ces facteurs de transcription possèdent pour la plupart d'entre eux des sites riches en paires de bases AT (comme stat 4) au niveau de leur site consensus, ou juxtaposés. L'unique exception réside dans l'inhibition des facteurs de transcription de la famille ETS (5'- GGAA) qui ne contient ni site ATGA, ni site riche en paires de bases AT.

On remarque enfin qu'aucun facteur de transcription testé ne voit sa fixation à l'ADN activée par l'interaction du DB293 sur son site consensus.

Cette approche innovante a donc permis de déterminer rapidement l'inhibition de la fixation à l'ADN de certains facteurs de transcription par le DB293. Dans la suite de cette étude nous nous avons plus finement caractérisé l'inhibition des facteurs de transcription Pit-1 et Brn-3 qui reconnaissent une séquence possédant le site ATGA et les facteurs de transcription de ETS n'en possédant pas.

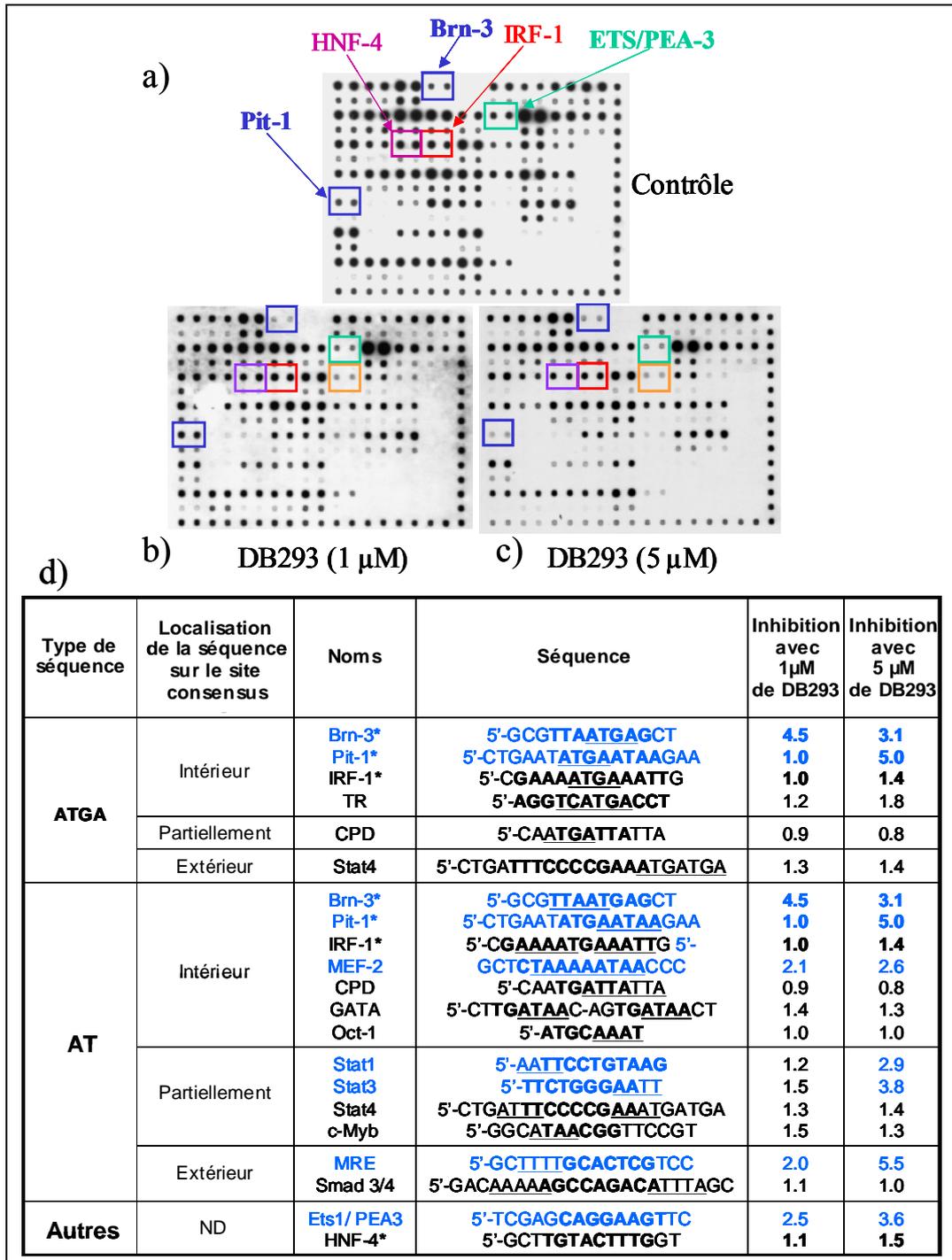


Figure 44 : Modulation de l'activité des facteurs de transcription par le DB293 mise en évidence par le kit TransSignalTM Protein/DNA Array I.

Résultats obtenus par l'utilisation des membranes Protein/DNA Array I en absence (a) ou en présence de 1 µM (b) ou 5 µM (c) de DB293. Les spots représentant la fixation de certains des facteurs de transcription d'intérêt sont encadrés. d) Ratio d'inhibition de la formation du complexe ADN/protéine par le DB293. Ces ratios sont calculés de la manière suivante : membrane contrôle / membrane DB293 (à 1 ou 5 µM). Le tableau indique si la séquence contient ou non le site ATGA, des sites riches en paires de bases AT (AT) ou d'autres séquences (autres). Ces sites se situent à l'intérieur, à l'extérieur ou juxtaposent le site consensus strict de reconnaissance par les facteurs de transcription. En bleu sont représentés les facteurs de transcription inhibés par le DB293.

2. Modulation de l'activité des facteurs de transcription par le DB293

La validation de ces résultats reste cependant primordiale. Pour cela, des expériences de retard en gel ont été effectuées sur les facteurs de transcription Pit-1, Brn-3, IRF-1 et HNF-4. Ce dernier ne possède ni site ATGA, ni site riche en paires de bases AT et n'est pas affecté par la présence du DB293, ce qui justifie son utilisation comme contrôle négatif. Pour réaliser les retards en gel, nous avons utilisé des extraits nucléaires de cellules HT-29 (Figure 46) ou des protéines exprimées en lysat de réticulocytes (Figure 47). En ce qui concerne le site Brn-3, il correspond à la cible de 3 protéines distinctes codées par 3 gènes différents : Brn-3a, Brn-3b et Brn-3c. Lors des expériences de retard en gel nous n'avons pas mis en évidence une fixation à l'ADN de la protéine Brn-3b exprimée en lysat de réticulocytes.

Pour exprimer la protéine en lysat de réticulocytes Pit-1, nous ne disposons pas de vecteur d'expression de la protéine Pit-1 humaine mais d'un vecteur codant pour la protéine Pit-1 de rat. Cependant, comme on peut le constater sur l'alignement de séquences réalisé entre ces 2 protéines Figure 45, les domaines POU_H et POU_S présentent un pourcentage d'identité de 96,6% et un pourcentage d'homologie de 98,3% dans l'homéodomaine POU. Les sites impliqués directement dans la reconnaissance de l'ADN sont identiques dans ces deux espèces, ce qui justifie totalement l'utilisation de cette protéine dans nos expériences de retard en gel.

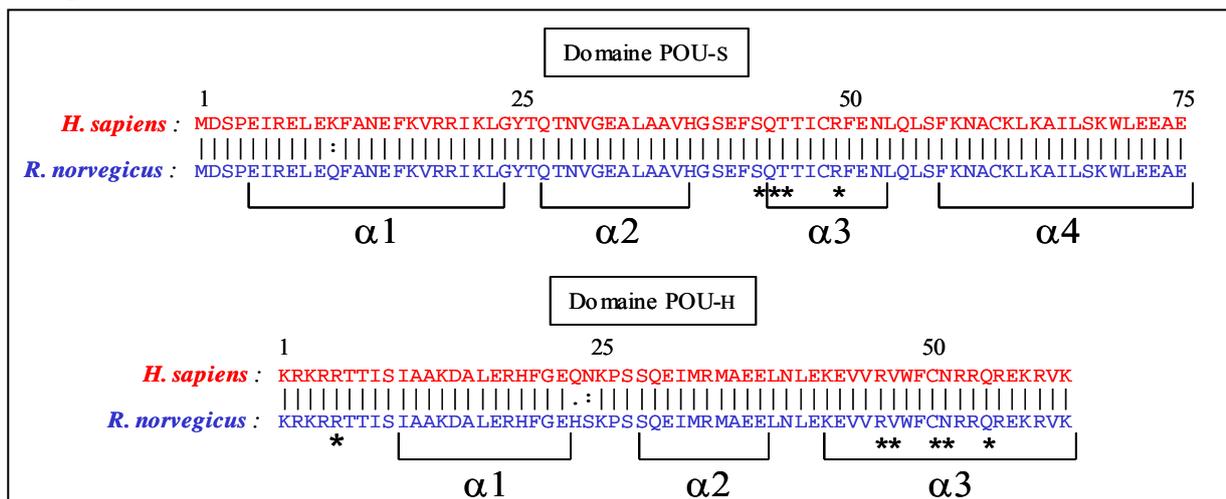


Figure 45 : Alignement de séquences des homéodomaines POU des protéines d'*Homo sapiens* et de *Rattus norvegicus*.

L'alignement de séquence montre un pourcentage d'identité de 96,6% (représenté par un bâton) et un pourcentage d'homologie de 98,3% (représenté par deux points). Les appariements de résidus ni homologues, ni identiques sont représentés par un point. Les étoiles représentent les résidus impliqués dans la reconnaissance de l'ADN.

Les séquences consensus utilisées pour effectuer les gels retard sont identiques à celles présentes sur le kit TranSignal. Pour la protéine Pit-1, l'utilisation d'extrait nucléaire et de sonde froide montre une bande spécifique ayant peu migré (Figure 46a). Etant donné la masse moléculaire de la protéine Pit-1 et la différence de migration obtenue avec la protéine exprimée en lysat de réticulocytes (Figure 47c), il s'agit très probablement d'un homo ou hétérodimère. L'utilisation d'une gamme de concentrations du DB293 (de 0,5 μM à 50 μM) met en évidence une inhibition efficace de la fixation de la protéine sur sa séquence cible. La quantification du gel (Figure 46 d) montre une inhibition de plus de 90 % de la fixation du facteur de transcription Pit-1 par le DB293 à 2,5 μM . Le même type de résultat est retrouvé quand on utilise la protéine Pit-1 exprimée en lysat de réticulocytes (Figure 47c).

En ce qui concerne la protéine Brn-3, l'utilisation d'extraits nucléaires (Figure 46 b) montre des résultats comparables à ceux décrits pour Pit-1. La quantification du gel (Figure 46d) montre là encore une inhibition de plus de 90 % de la fixation du facteur de transcription Brn-3 par le DB293 à 2,5 μM sur sa séquence cible, avec des valeurs d' IC_{50} d'environ 1,5 μM . Les expériences effectuées avec les protéines exprimées en lysat de réticulocytes confirment ces résultats et mettent en évidence l'inhibition de la fixation des protéines Brn-3a et Brn-3c de manière identique (Figure 47a et b).

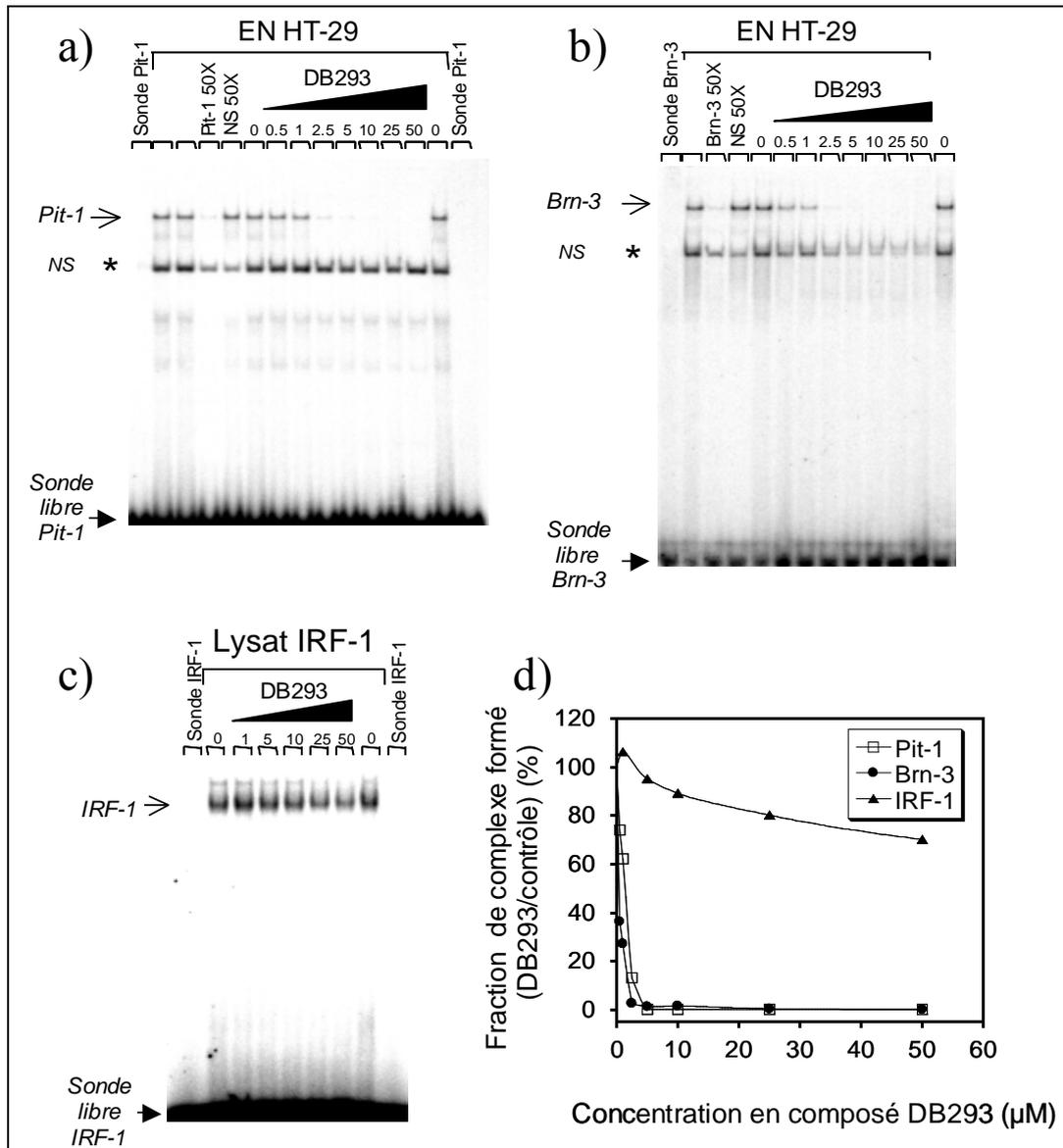


Figure 46 : Inhibition de la formation des complexes ADN/facteurs de transcription par le DB293 visualisée par retard en gel.

Les sondes Pit-1 (a), Brn-3 (b), IRF-1 (c) sont incubées seules (ligne sonde), avec 5 μg d'extrait nucléaire (EN) de cellules HT-29 (EN HT-29) ou avec 2 μL de protéines exprimées en lysat de réticulocytes. Une gamme de concentrations en DB293 (μM) est ajoutée et permet la modulation de la fixation des facteurs de transcription. L'ajout d'un facteur 50 de sonde froide spécifique ou non spécifique (NS) par rapport à la sonde marquée met en évidence la bande d'intérêt. L'étoile montre la fixation non spécifique (NS) de protéines contenues dans les extraits nucléaires. (d) Quantification de l'inhibition de la formation du complexe ADN/facteur de transcription par le DB293.

Les expériences de retard en gel sur les facteurs de transcription IRF-1 (Figure 45c) et HNF-4 (Figure 46d) n'ont montré qu'une faible inhibition de la fixation des facteurs par le DB293. Ces manipulations valident donc les résultats obtenus à l'aide du kit TranSignal™ Protein/DNA Array I.

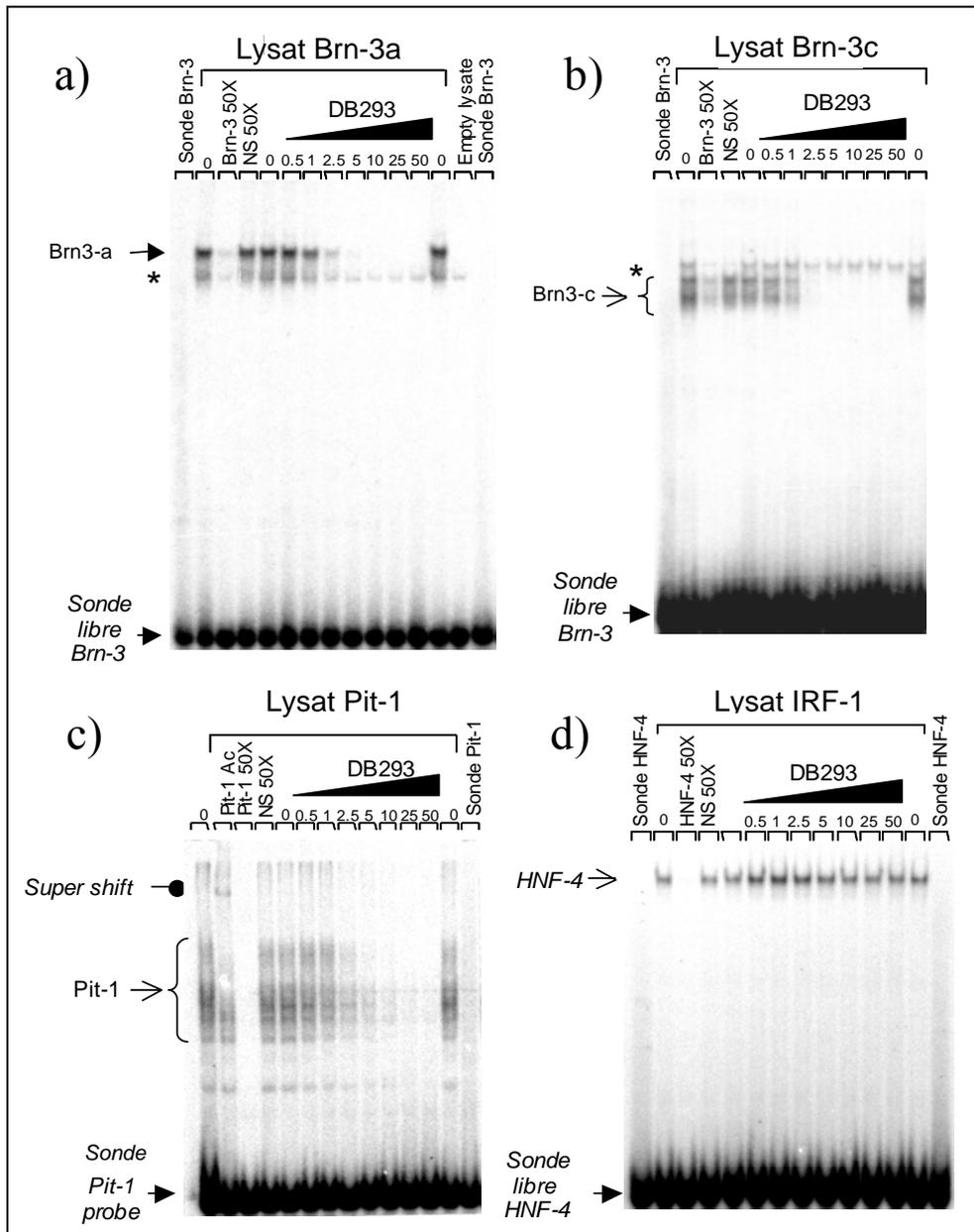


Figure 47 : Inhibition de la formation des complexes ADN/facteurs de transcription exprimés en lysat de réticulocytes par le DB293 visualisée par retard en gel.

La modulation de la fixation des facteurs de transcription Brn-3a (a), Brn-3b (b) et Pit-1 (c) est mise en évidence par retards en gel effectués avec des protéines exprimées en lysat de réticulocytes. La spécificité de la bande Pit-1 est montrée par l'utilisation d'un anticorps anti Pit-1 (supershift). d) Le facteur de transcription HNF-4 est utilisé comme contrôle négatif et reste insensible à l'utilisation du composé DB293.

3. Interaction du DB293 sur les sites consensus Pit-1, Brn-3 et IRF-1

Afin de comprendre pourquoi le DB293 inhibe la formation des complexes ADN / Pit-1 et ADN / Brn-3 mais n'inhibe pas la formation du complexe ADN / IRF-1, nous avons étudié le

mode précis d'interaction du DB293 sur ces trois sites consensus. Si l'on s'intéresse à la séquence, on constate que ces trois sites possèdent un site 5'-ATGA entouré de sites riches en paires de bases AT. Des expériences d'empreinte à la DNase I comparées à des expériences de résonance plasmonique de surface (SPR, Biosensor-surface plasmon resonance) ont permis de répondre à cette question.

a) Localisation du DB293 sur les sites Pit-1, Brn-3 et IRF-1

Dans un premier temps, nous avons effectué des tests d'empreinte à la DNase I (Figure 48). Pour cela, nous avons cloné les séquences consensus Pit-1, Brn-3 et IRF-1 dans le plasmide pUC19 (en pointillé sur la Figure 48). Ces fragments d'ADN radiomarqués sont soumis à une digestion ménagée par la DNase I en l'absence (ligne ADN) ou en présence d'une gamme de concentrations croissantes en DB293. L'insertion du site Pit-1 sur le site de restriction *Bam*H I est inversée par comparaison aux sites Brn-3 et IRF-1, relativement à la lecture du site cible ATGA ou de son complémentaire TCAT.

Pour le site Pit-1 (Figure 48a et d), l'analyse densitométrique du gel montre une empreinte centrée sur la séquence **TTATCATATT** (p 34-43). Ce site contient donc le site 5'-TCAT (ou son complémentaire 5'-ATGA) et les séquences adjacentes riches en paires de bases AT. Néanmoins, il est difficile de localiser précisément le ou les sites de reconnaissance du DB293. Pour la séquence consensus Brn-3 (Figure 48b et e), l'autoradiographie du gel et l'analyse densitométrique révèlent une interaction spécifique sur le site 5'-ATGA et sur le site adjacent 5'- riche en paires de bases AT (**TTAATGA**).

Concernant le fragment IRF-1 (Figure 48c et f), l'analyse du clivage différentiel montre que le DB293 se fixe sur les deux sites riches en paires de bases AT adjacents le site 5'- ATGA mais en aucun cas sur la séquence centrale 5'-ATGA (5'-**AAAATGAAATTGAC**).

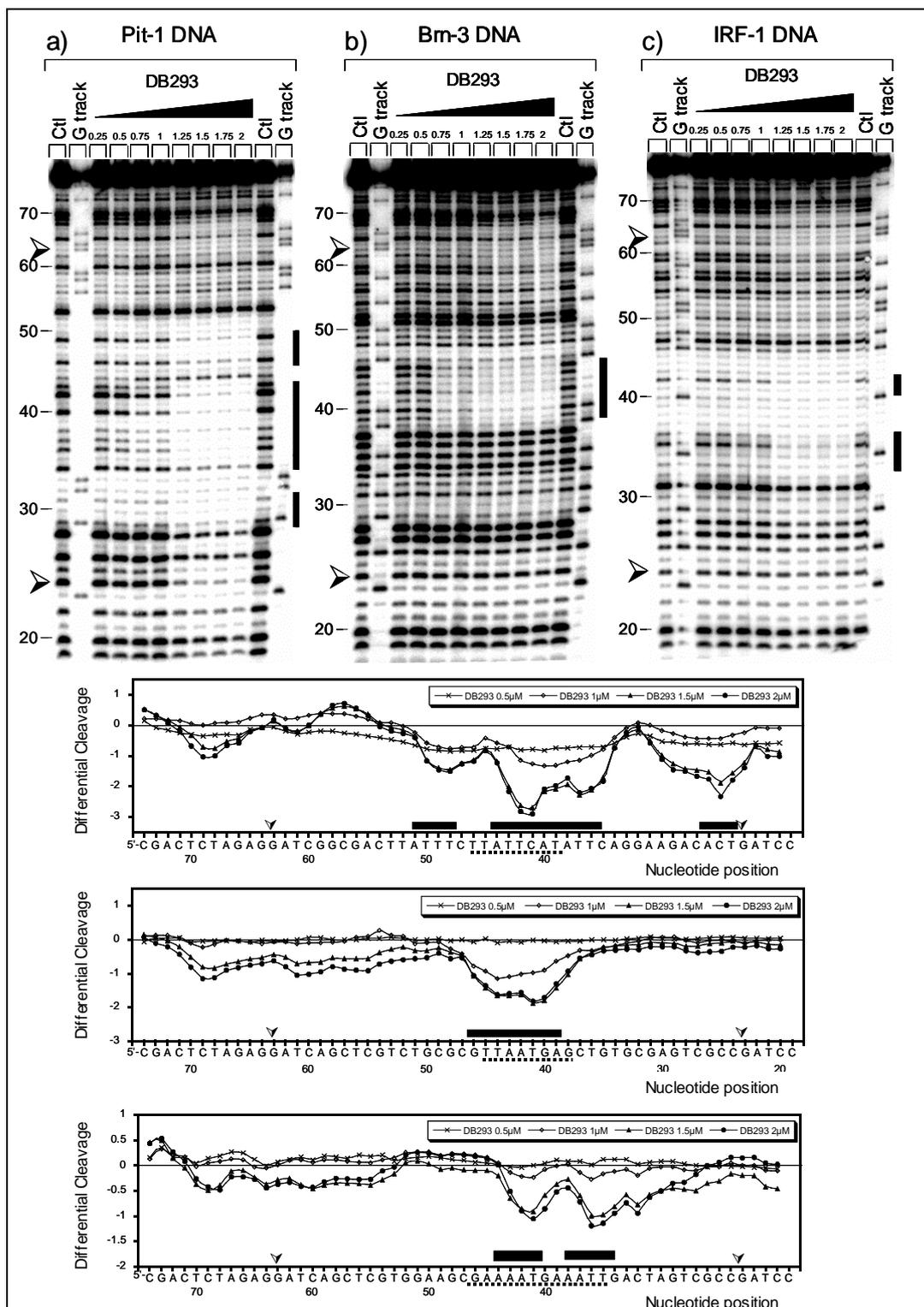


Figure 48 : Empreintes à la DNase I du DB293 sur les sites consensus Pit-1, Brn-3 et IRF-1.

Les fragments issus du plasmide pUC19-Pit-1 (a et d), -Brn-3 (b et e) et -IRF-1 (c et f) sont incubés 20 min avec une gamme de concentration en DB293 (de 0,25 μM à 2 μM) et soumis à une digestion ménagée par la DNase I. La séquence d'ADN des différents sites consensus clonée dans le plasmide pUC19 est située entre les deux flèches (➤). La localisation de l'empreinte est représentée par des rectangles noires sur les gels (a, b et c) et sur les analyses densitométriques respectives (d, e et f).

Ces résultats révèlent que le DB293 reconnaît préférentiellement les sites riches en paires de bases AT par rapport à la séquence centrale 5'-ATGA du site IRF-1. Ceci contraste avec les résultats obtenus pour les séquences Pit-1 et Brn-3 où l'empreinte est centrée sur le site 5'-ATGA central et suggère que la reconnaissance du site 5'-ATGA serait un élément déterminant dans l'inhibition de la formation des complexes ADN/facteurs de transcription.

b) Stœchiométrie d'interaction

Pour déterminer la stœchiométrie et le mode d'interaction (mode coopératif ou non) du DB293 sur les sites consensus des facteurs de transcription Pit-1, Brn-3 et IRF-1, des expériences de SPR ont été effectuées. Ces manipulations sont réalisées par le Docteur Yang Liu au laboratoire du professeur David Wilson (Atlanta, USA). Trois oligonucléotides (en forme d'hairpin) biotinylés en 5' et contenant les sites consensus Pit-1, Brn-3 et IRF-1 (séquences en rouge Figure 49a) sont immobilisés par interaction avec la streptavidine sur 3 des 4 cellules de flux (biosensorchip) que contient la sensorchip. La quatrième cellule sert de contrôle négatif et permet de soustraire le bruit de fond. Les sensorgrammes (RU pour les pistes contenant l'ADN – RU pour le contrôle sans ADN) (Figure 49) sont obtenus par injection de concentrations croissantes en DB293. Les sensorgrammes sont alors analysés selon un modèle d'interaction à deux molécules (Figure 49b), ce qui permet de déterminer les constantes d'équilibre de l'interaction (figure 4 c). La valeur de RU (RU_{max}) maximale pour chacun des ADN est deux fois supérieure à la valeur de RU_{max} pour un composé qui interagit sur un site unique. Cela montre clairement que deux molécules de DB293 interagissent sur chacun des oligonucléotides. Les valeurs relatives des constantes d'équilibre macroscopiques, K_1 et K_2 , reflètent le mode d'interaction de la molécule. Ainsi, lorsque $K_1 = 4 \times K_2$ le mode est non coopératif, lorsque $K_1 < 4 \times K_2$ le mode est coopératif positif et lorsque $K_1 > 4 \times K_2$ le mode est dit coopératif négatif. Le facteur de coopérativité qui permet d'accéder au degré de coopérativité est défini de la manière suivante : $CF = (K_2/K_1) \times 4$. Pour une interaction non coopérative, $CF = 1$, pour une interaction coopérative $CF > 1$ et pour une interaction coopérative selon un mode négatif $CF < 1$. Le tableau de la Figure 49c montre clairement que le DB293 se fixe sur l'oligonucléotide Brn-3 selon un mode fortement coopératif ($CF = 62,78$). Pour l'oligonucléotide IRF-1, il n'existe pas d'effet coopératif ($CF = 1,32$) et pour Pit-1 les résultats montrent une coopérativité de valeur intermédiaire ($CF = 4,55$).

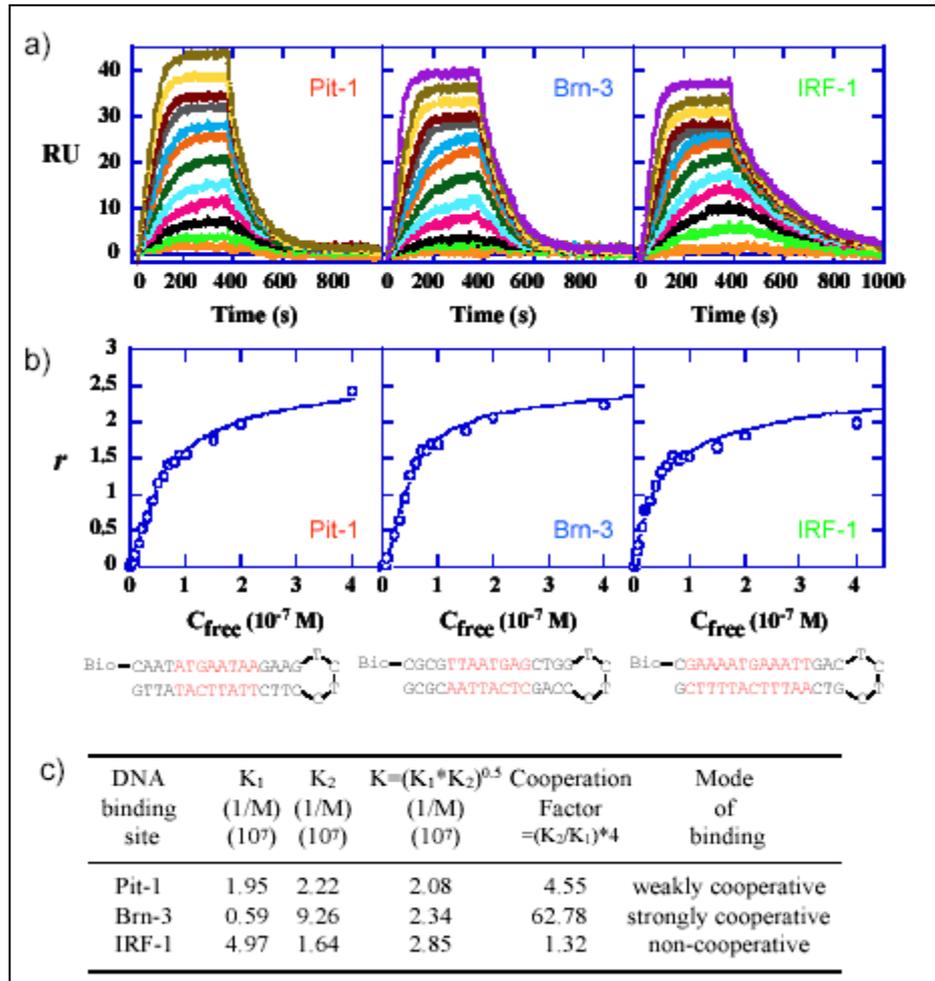


Figure 49 : Analyse en résonance plasmonique de surface de l'interaction entre le DB293 et les sites consensus des facteurs de transcription Pit-1, Brn-3 et IRF-1.

a) Sensorgrammes obtenus en présence de quantité croissante de DB293 (de 1 nM courbe du bas jusque 0,4 μM courbe du haut) sur des oligonucléotides en forme d'hairpin contenant les sites consensus Pit-1, Brn3 et IRF-1 (en rouge). b) Courbe d'interaction dérivée des sensorgrammes. Ces courbes sont utilisées pour déterminer les constantes d'affinité du DB293 pour les sites consensus Pit-1, Brn-3 et IRF-1. La séquence et la structure des *hairpins* sont présentées en bas de la figure. c) Constantes d'équilibre du DB293 pour les différents sites consensus. K_1 et K_2 sont les constantes macroscopiques déduites des courbes d'interaction. Les facteurs de coopérativité sont utilisés pour déterminer le potentiel coopératif du DB293 pour l'ADN.

Le mode d'interaction coopératif entre la molécule et l'ADN prédit une fixation au site 5'-ATGA sous forme de dimère. Ceci suggère que le DB293 interagirait sur le site 5'-ATGA sous forme de dimère au niveau du petit sillon de l'ADN pour les oligonucléotides Brn-3 (et Pit-1 avec dans une moindre mesure). Par opposition, l'absence de coopérativité pour l'oligonucléotide IRF-1 confirme les résultats obtenus précédemment par les tests d'empreintes à la DNase I. Le DB293 interagit sur cette dernière séquence sous forme de deux monomères indépendants au niveau des deux sites riches en paires de bases AT qui entourent le site 5'-ATGA.

Ces conclusions sont confirmées par des études SPR effectuées avec le DB75. Comme précédemment montré, cette molécule interagit uniquement sur un site riche en paires de bases AT. Ainsi, deux molécules de DB75 interagissent fortement de manière non coopérative avec l'oligonucléotide IRF-1, alors qu'elles se fixent faiblement sur la séquence Brn-3 (Figure 50).

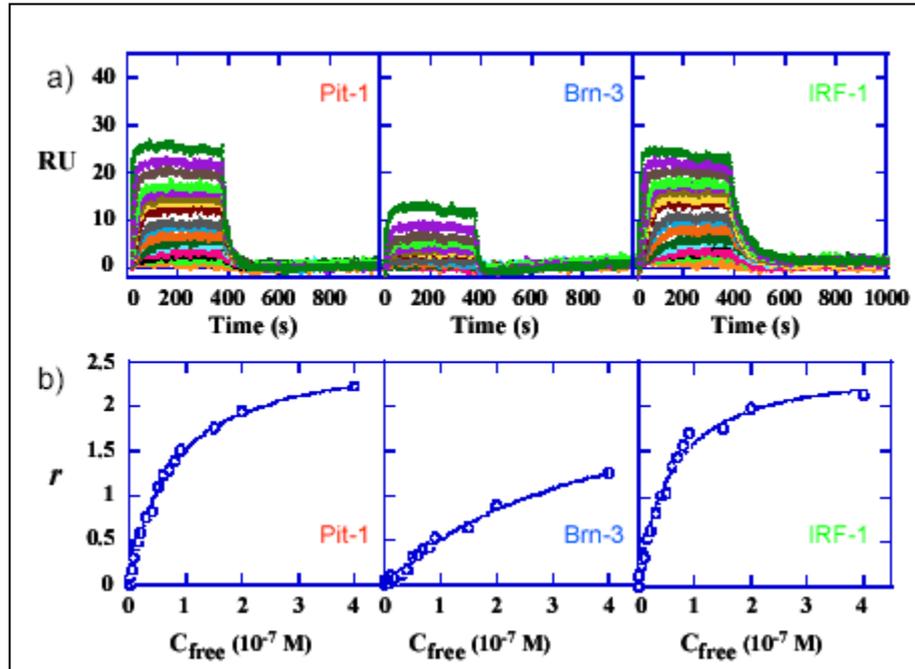


Figure 50 : Analyse en résonance plasmonique de surface du DB75 sur les séquences Pit-1, Brn-3 et IRF-1.

a) Sensorgrammes obtenus en présence de quantité croissante de DB75 (de 1 nM courbe du bas jusque 0,4 µM courbe du haut) sur des oligonucléotides en forme d'hairpin contenant les sites consensus Pit-1, Brn3 et IRF-1 (en rouge). b) Courbes d'interaction dérivées des sensorgrammes.

Les résultats obtenus avec la séquence Pit-1 sont intermédiaires et suggèrent que les sites riches en paires de bases AT entrent en compétition avec le site ATGA pour l'interaction avec le DB293. Pour le DB75, la reconnaissance s'effectue uniquement sur les sites riches en paires de bases AT.

c) Localisation du DB293 sur le site consensus Pit-1

Pour comprendre le mécanisme précis d'interaction entre le site Pit-1 avec le DB293, de nouvelles empreintes à la DNase I ont été effectuées sur une série de séquences Pit-1 mutées (Figures 51 et 52). Ainsi, cinq fragments d'ADN ont été construits :

- le fragment Pit-1 WT (Wild Type) contenant le site consensus strict Pit-1 (TTAATTCATATTCA);

- le fragment Pit-1 M1 qui contient un point de mutation au niveau du site 5'-ATGA / 5'-TCAT (TTAATTCGTATTCA) ;
- le fragment Pit-1 M2 qui contient un point de mutation au niveau du site 5'- riche en paires de bases AT (TCGTTCATATTCA) ;
- le fragment Pit-1 M3 qui contient deux points de mutation au niveau des sites 5'-ATGA et 5'- riche en paires de bases AT (TCGTTCGTATTCA) ;
- le fragment Pit-1 M4 qui contient deux points de mutation au niveau des sites 5'- et 3'- riches en paires de bases AT (TCGTTCATACTCA).

L'empreinte obtenue avec le fragment Pit-1 M1, qui ne touche que le site 5'-TCAT, est différente de celle obtenue avec le fragment Pit-1 WT. En effet, la mutation du site 5'-TCAT par le site 5'- TCGT modifie le clivage observé avec le fragment Pit-1 WT et montre une protection du clivage centrée sur la paire de bases C en position 41. De plus, l'abolition du site 5'-TCAT induit un déplacement de l'empreinte obtenue par la digestion à la DNase I sur le site riche en paires de bases AT, TTATT, localisé en position 5' de la mutation (position 42 à 46). Le fragment Pit-1 M2 est modifié sur les positions 44 et 45 ce qui a pour conséquence de rompre le site riche en paires de bases AT en position 5'- de la séquence TCAT. L'empreinte à la DNase I obtenue avec ce fragment est semblable à celle obtenue avec le fragment Pit-1 WT, avec une reconnaissance importante sur le site TCAT et le site 3'- riche en paires de bases AT. Les mutations du site TCAT et du site 5'- riche en paires de bases AT du fragment Pit-1 M3 abolissent l'interaction entre le DB293 et l'ADN. On observe uniquement une légère empreinte au niveau du site TATT en position 36 à 39. Pour terminer, la mutation des sites riches en paires de bases AT situé en 3' et 5' de la séquence TCAT montre une interaction forte entre le DB293 et le site TCAT. L'empreinte observée est plus marquée sur cette zone par rapport à celle obtenue pour le fragment Pit-1 WT.

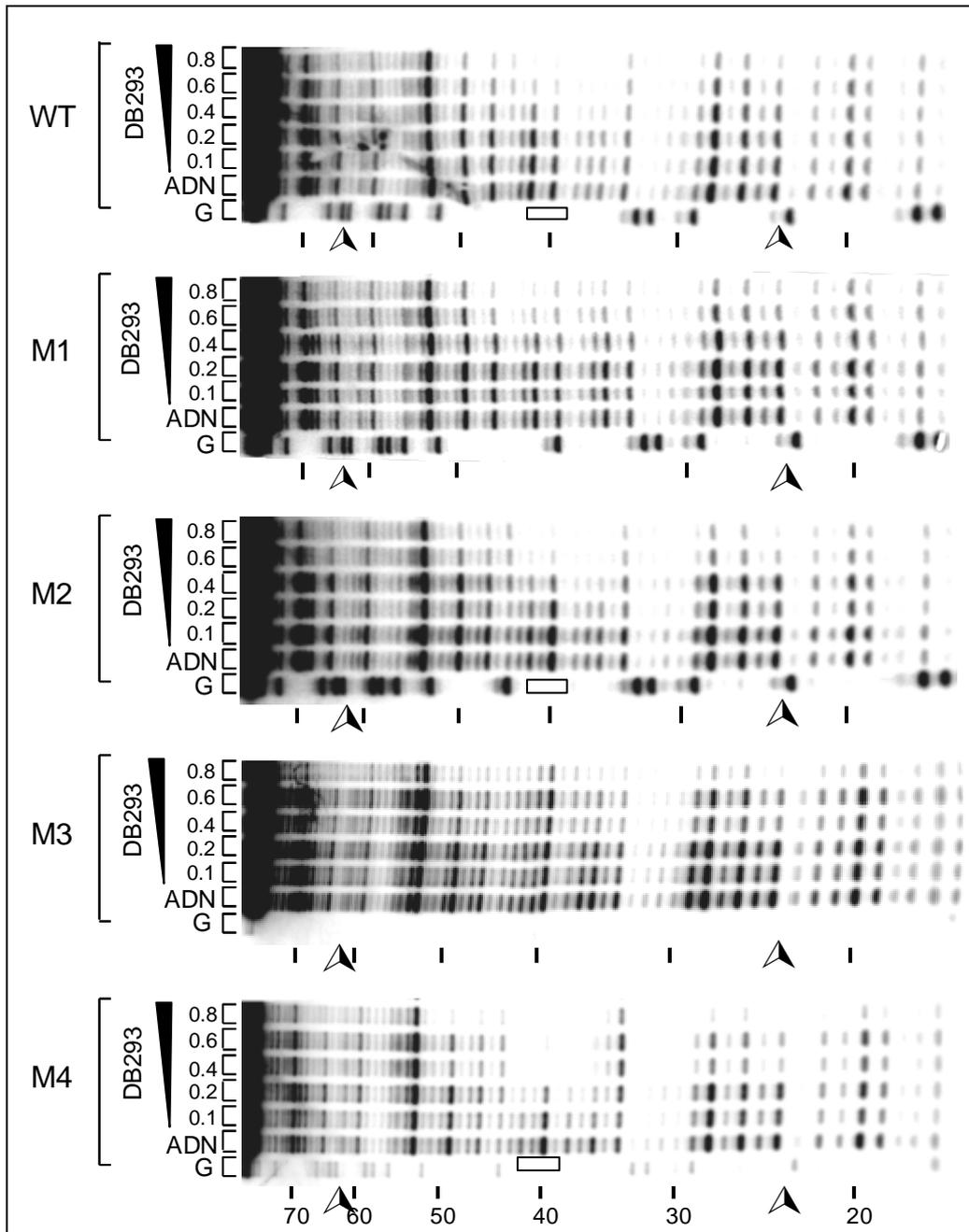


Figure 51 : Empreinte à la DNase I du DB293 sur les fragments d'ADN WT et mutés du site consensus minimal Pit-1.

Les fragments issus de la digestion *Hind* III/*Pst* I des plasmides pUC19 Pit-1 WT, Pit-1 M1, Pit-1 M2, Pit-1 M3 et Pit-1 M4 marqués sur l'extrémité 3' sont incubés avec une gamme de concentrations en DB293 (μ M) avant d'être digérés par la DNase I. La ligne ADN correspond au contrôle sans drogue et la ligne G met en évidence la position des guanines du fragment.

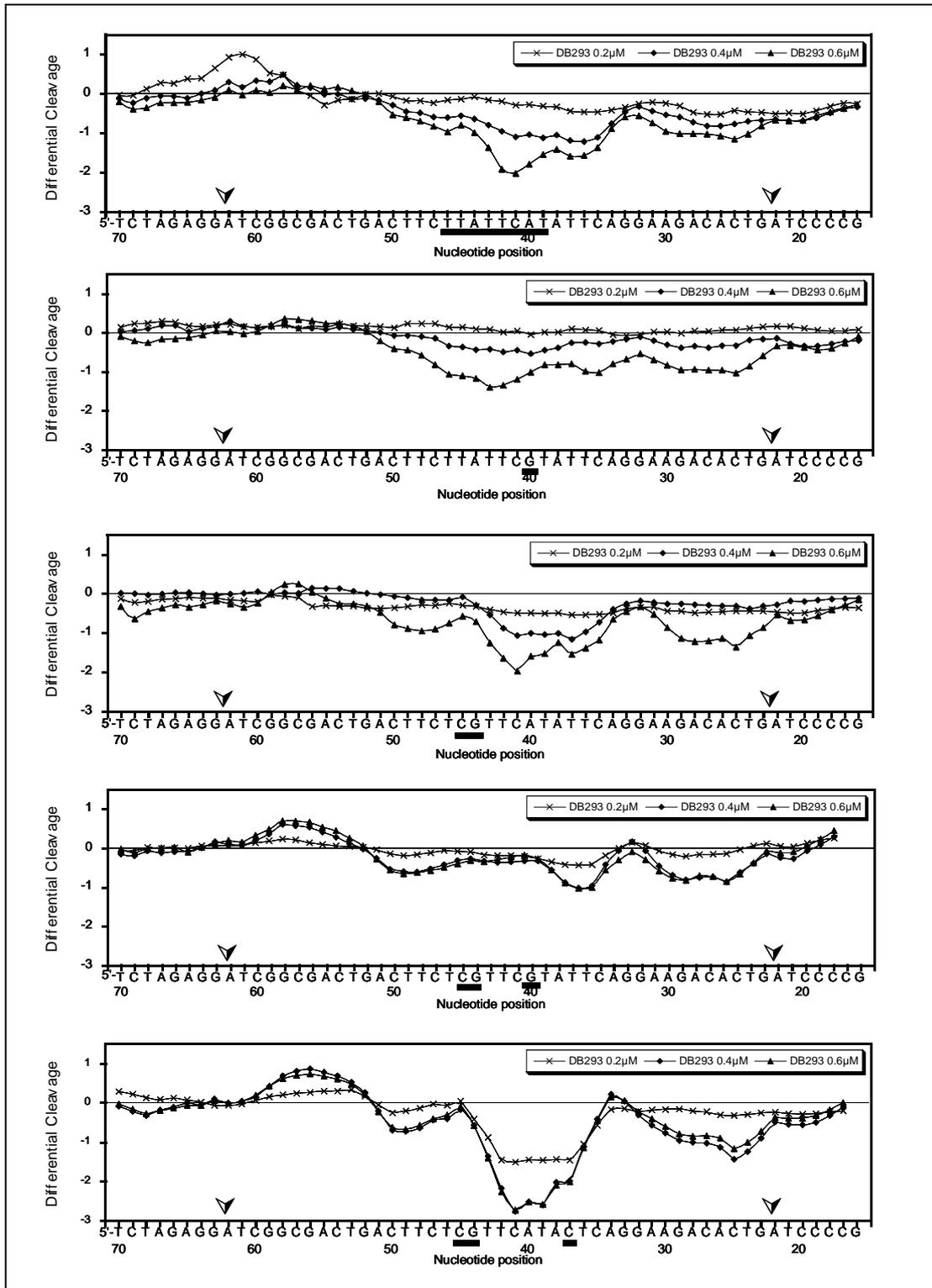


Figure 52 : Analyse densitométrique des empreintes à la DNase I effectuées sur les différents fragments d'ADN présentés en figure 51.

Les différents sites consensus clonés se situent entre les deux flèches. Les boîtes noires représentent les mutations effectuées sur le site Pit-1. Les points situés sous le zéro correspondent aux zones de protection de la digestion et donc aux sites d'interaction du composé DB293.

Ces résultats montrent clairement que l'interaction entre le DB293 et la séquence consensus Pit-1 s'effectue par une fixation au site ATGA/TCAT mais aussi par une fixation sur les sites riches en paires de bases AT adjacents.

Pour définir quantitativement l'interaction entre le composé DB293 et la séquence Pit-1, des expériences de résonance plasmonique de surface ont été réalisées sur les séquences Pit-1 WT, Pit-1 M1 et Pit-1 M2 (Figure 53).

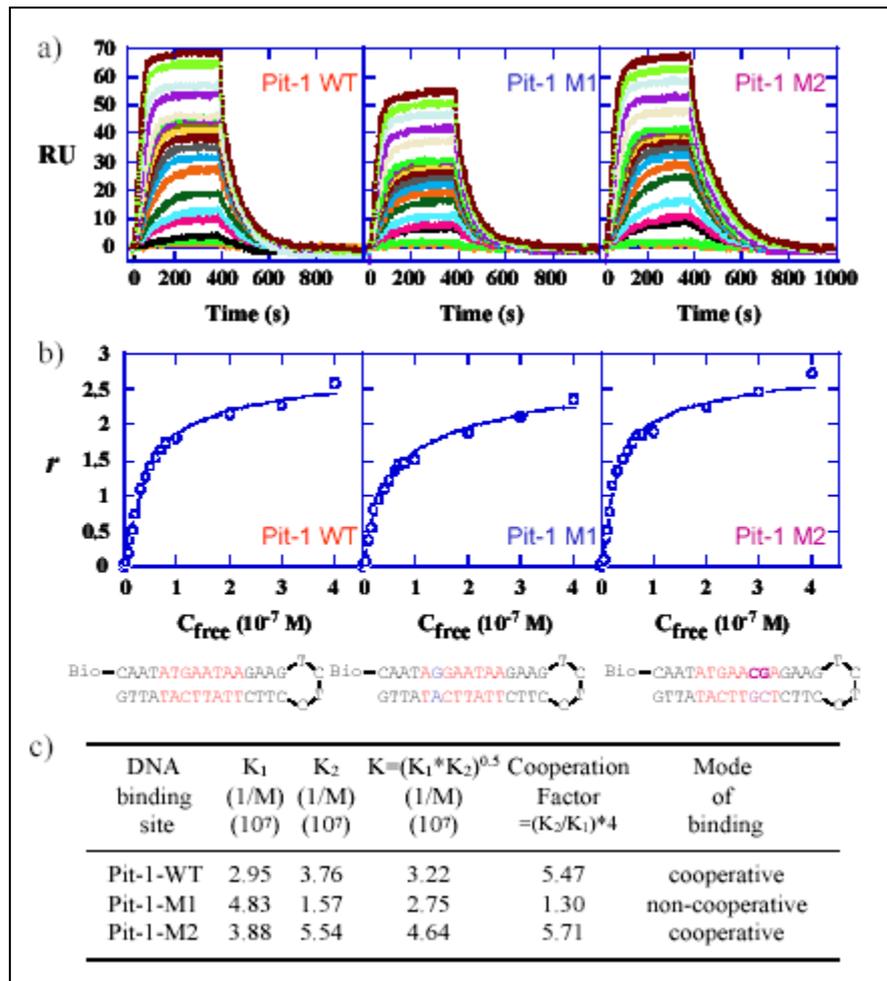


Figure 53 : Analyse en résonance plasmonique de surface de l'interaction du DB293 et les oligonucléotides Pit-1 WT, Pit-1 M1 et Pit-1 M2.

a) Sensorgrammes représentant l'interaction entre le DB293 (de 1 nM à 0,4 μM) et les hairpins d'ADN Pit-1 WT, M1 et M2. b) Courbes d'interaction issues des sensorgrammes SPR. Ces courbes sont utilisées pour calculer les constantes d'affinité de l'interaction. La structure de l'oligonucléotide et le site consensus minimum (en rouge) sont représentés en bas de la figure. Les bases en bleu correspondent aux sites de mutation. c) Constantes d'équilibres de l'interaction du composé DB293 à l'ADN. K_1 et K_2 représentent les constantes macroscopiques déduites des courbes d'interaction. Les facteurs de coopérativité déduits sont utilisés pour déterminer le mode d'interaction.

Les sensorgrammes représentant l'interaction entre ces trois séquences et le DB293 montrent des valeurs de RU_{max} deux fois supérieures à celle obtenues pour la fixation d'une molécule par site. Ceci indique que deux molécules de DB293 se lient à ces séquences (Figure 53 b). Pour les oligonucléotides Pit-1 WT et Pit-1 M2, les CF montrent une interaction de type coopérative, avec des valeurs respectives de 5,47 et 5,71. Pour le fragment M1, où le site ATGA est muté, le résultat est totalement différent avec un CF de 1,3 qui traduit un mode d'interaction non coopératif. Ces résultats sont totalement en accord avec ceux déterminés précédemment et valident la fixation du DB293 au site ATGA/TCAT sur le site consensus Pit-1.

4. Identification des sites essentiels à la formation du complexe protéine/ADN

Pour compléter cette étude des bases moléculaires de l'interaction entre le facteur de transcription Pit-1 et sa séquence consensus, nous avons évalué par retard en gel la capacité de la protéine Pit-1 à interagir avec la séquence Pit-1 WT et les séquences Pit-1 mutées (Figure 54). Ces manipulations ont été réalisées avec des extraits nucléaires de cellules HT-29 et avec la protéine Pit-1 exprimée en lysat de réticulocytes. Pour compléter le panel des mutations, nous avons utilisé deux sites mutés supplémentaires :

- Pit-1 M5 muté uniquement sur le site riche en paires de bases AT situé côté 5'- de la séquence ATGA (TCCTGAGTATGAATAAGAAGT) ;
- Pit-1 M6 qui possède deux points de mutation sur la séquence ATGA et sur le site riche en paires de bases AT situé côté 5'- (TCCTGAGTACGAATAAGAAGT).

Les deux gels présentés Figure 54 révèlent qu'une unique mutation sur un des 3 sites (M1, M2 ou M5) réduit la capacité de la protéine Pit-1 à se lier avec son site consensus mais ne l'abolit pas. L'utilisation d'oligonucléotides doublement mutés (M3) empêche un peu plus fortement la formation du complexe ADN/protéine. Les doubles mutations M6 ou M4 abolissent totalement la fixation de la protéine Pit-1 sur son site consensus. Ces données suggèrent que le site riche en paires de bases AT situé côté 5'- de la séquence ATGA joue un rôle déterminant dans la formation du complexe (Figure 54 c flèche large). Ces résultats sont en accord avec la flexibilité importante des protéines à POU domaine dans la reconnaissance de séquences consensus décrite par l'équipe de Jacobson. (Jacobson et al, 1997).

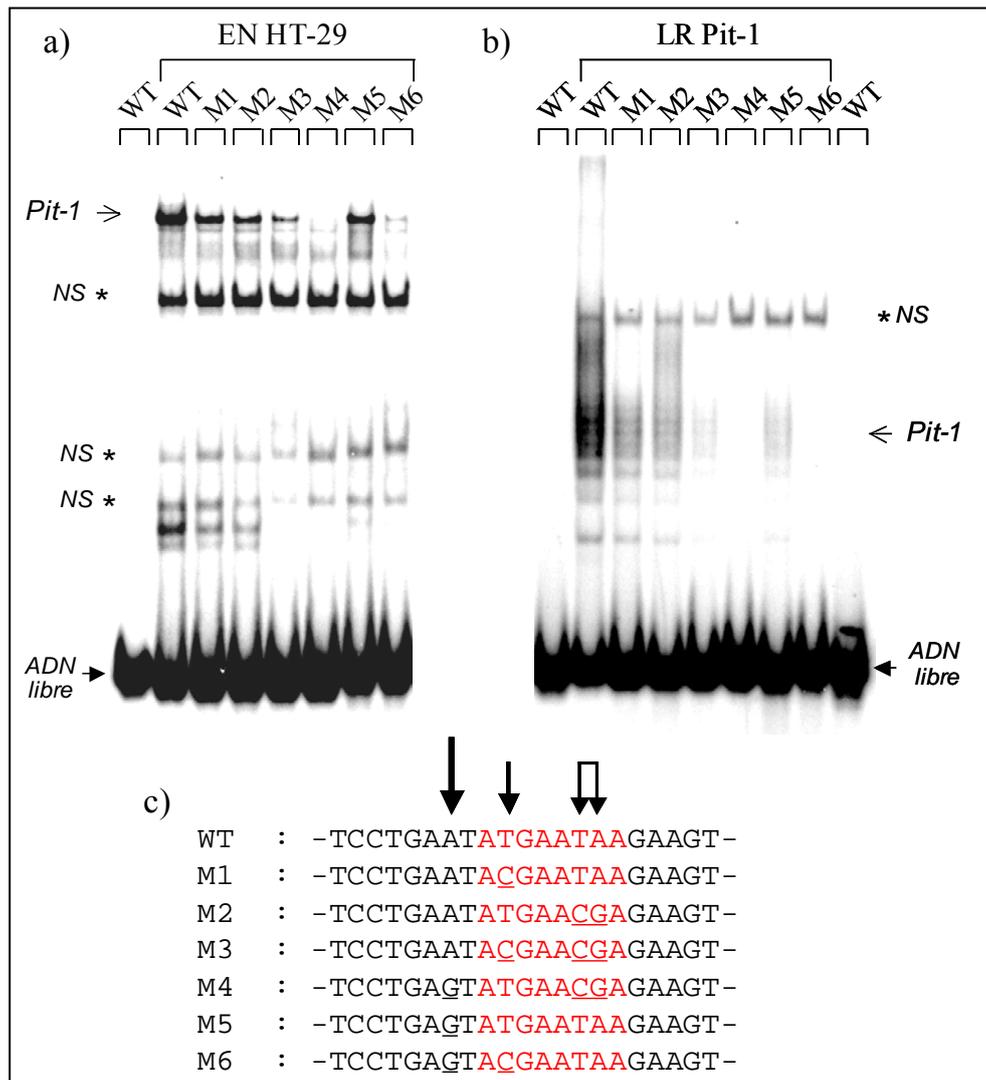


Figure 54 : Mise en évidence des éléments déterminants dans la reconnaissance de la protéine Pit-1 pour sa séquence cible par retard en gel.

a) Les oligonucléotides Pit-1 WT et mutés M1 à M6 sont incubés avec 5 µg d'extraits nucléaires de cellules HT-29 (a) ou 2 µL de protéine Pit-1 exprimée en lysat de réticulocytes (b). La bande Pit-1 correspond à la bande spécifique, les bandes NS, localisées par un astérisque (★) correspondent aux bandes non spécifiques. c) Représentation des différents points de mutations présents sur le site Pit-1 (bases soulignées). Les flèches montrent l'importance de ces sites dans la formation du complexe, leur taille étant fonction de l'importance du site dans la reconnaissance.

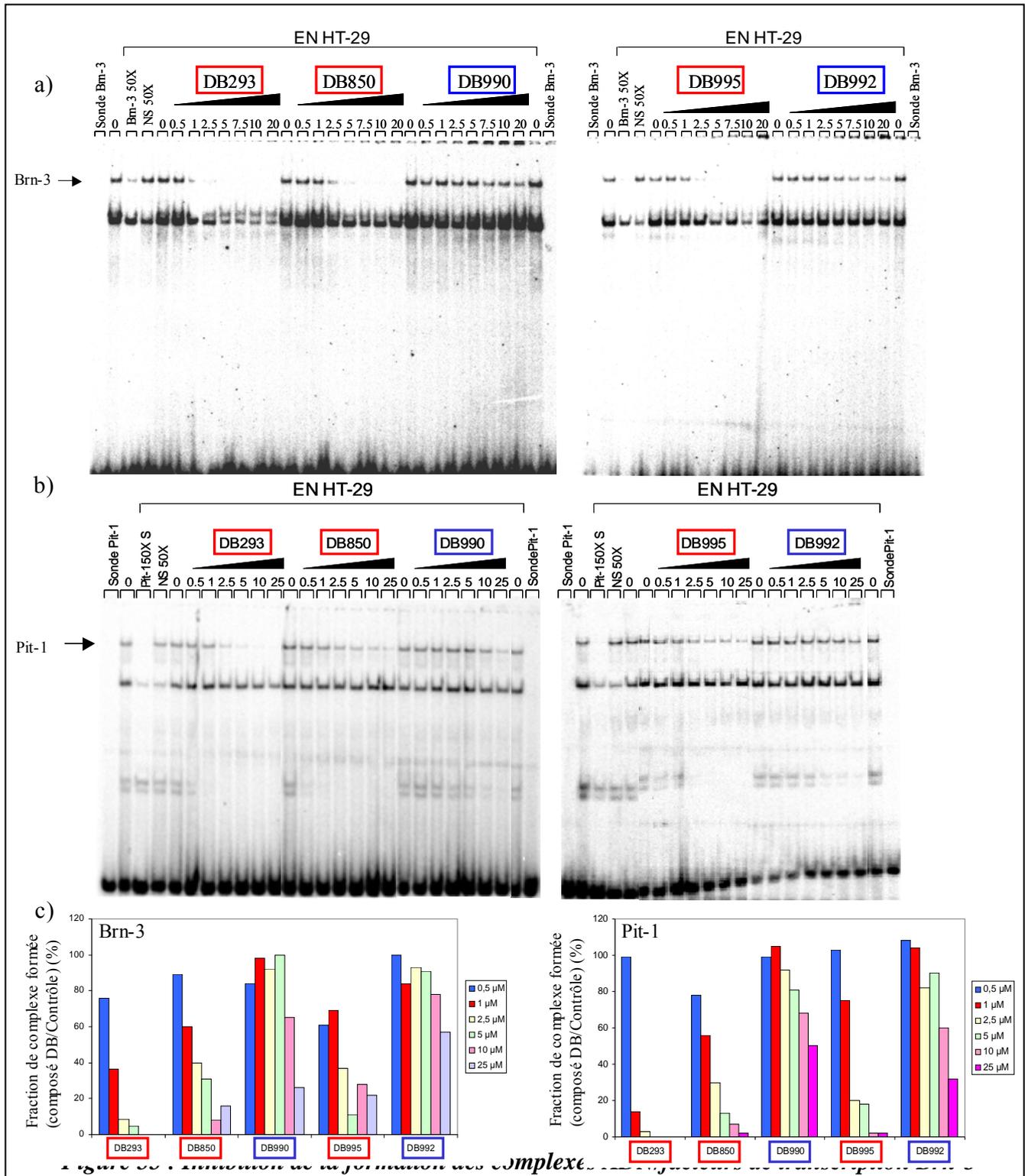
La mutation d'un des trois sites diminue modestement la formation du complexe ADN/facteur de transcription. La protéine possède la capacité de s'adapter à ces modifications de séquences notamment grâce à la flexibilité apportée par le *linker* qui sépare les deux domaines POU_H et POU_S. Par contre, une mutation de deux des trois sites d'interaction potentiels affecte fortement la formation du complexe.

5. Importance du site ATGA dans l'inhibition des protéines Pit-1 et Brn-3

Afin de valider l'importance de la reconnaissance du site ATGA dans l'efficacité d'inhibition des facteurs de transcription Brn-3 et Pit-1, d'autres tests de retard en gel ont été réalisés sur de nouveaux dérivés DB (Figure 55). Pour cela, les produits DB850 et DB995 précédemment décrits (Figure 31), capables de reconnaître des sites 5'-ATGA et deux composés analogues DB990 et DB992, qui interagissent uniquement sur les sites riches en paires de bases AT ont été testés.

Les résultats présentés montrent clairement que les molécules qui forment des dimères spécifiques sur le site ATGA inhibent fortement la formation des complexes Brn-3/ADN. Les mêmes résultats sont observés pour le facteur de transcription Pit-1.

Ainsi, les composés qui se lient aux sites ATGA inhibent beaucoup plus efficacement la formation du complexe ADN/facteur de transcription que les composés qui reconnaissent les sites riches en paires de bases AT. Le facteur d'inhibition de la formation du complexe entre les composés « ATGA » et les composés « AT » est supérieur à 10. La reconnaissance spécifique du site ATGA est donc primordiale pour une inhibition efficace de la fixation de ces facteurs de transcription.



par les composés « ATGA » ou « AT » visualisée par retard en gel.

La sonde Brn-3 a) ou Pit-1 b) est incubée seule (ligne sonde) ou avec 5 μg d'extraits nucléaires de cellules HT-29 (EN HT-29) en présence d'une gamme de DB293, DB850, DB995, DB990 et DB992 (en μM). c) Quantification de la formation du complexe ADN/protéine en présence des différents composés DB. Les produits encadrés en rouge sont capables de se lier au site ATGA et les produits encadrés en bleu se lient uniquement sur les sites riches en paires de bases AT.

CONFIDENTIEL

DISCUSSION ET PERSPECTIVES

Depuis l'avènement de la chimiothérapie antitumorale et l'utilisation des moutardes azotées, l'ADN représente une cible thérapeutique majeure. Ainsi, de nombreux composés cytotoxiques sont utilisés et restent prépondérants dans le traitement de certains cancers. Si l'ADN et ses partenaires protéiques constituent toujours une cible privilégiée en pharmacologie antitumorale, la manière d'appréhender leur modulation par des petites molécules a évolué ces dernières années pour laisser place à des stratégies plus fines, prenant davantage en compte les paramètres dynamiques et la complexité des circuits moléculaires impliqués (Hurley, 2002).

Les dérivés DB analogues des composés bérénil et pentamidine s'intègrent totalement dans cette approche de part leur reconnaissance sélective pour des séquences spécifiques de fixation de l'ADN. En effet, si le bérénil cible le petit sillon de l'ADN et reconnaît 3 paires de bases AT consécutives, le DB75, qui diffère du bérénil par le remplacement du triazène par un cycle furane (Figure 13), présente une meilleure affinité pour l'ADN et reconnaît 4 paires de bases AT consécutives (Laughon *et al.*, 1996). Pour les dérivés diphényl-furanes analogues du DB75, ces propriétés persistent et le choix des amidines terminales jouent un rôle important dans l'affinité à l'ADN (Trent *et al.*, 1996 ; Wilson *et al.*, 1998). Les dérivés phényl-furane-benzimidazoles comme le composé DB293 reconnaissent non seulement les séquences riches en paires de bases AT mais aussi des sites spécifiques ATGA (Bailly *et al.*, 2001 ; Tanious *et al.*, 2003). Pour cette famille de molécules, le choix des amidines terminales joue un rôle important dans la reconnaissance des séquences ATGA (Wang *et al.*, 2001). Différentes études sur leurs activités anti-parasitaires et cellulaires ont été menées. Les variations de localisation cellulaire peuvent dépendre des substitutions au niveau des amidines terminales (Lansiaux *et al.*, 2002a ; 2002b). Les premiers tests de cytotoxicité sur les dérivés DB ont été réalisés il y a une dizaine d'années (Neidle *et al.*, 1997). Aucune donnée n'a par contre été publiée sur leur rôle dans la modulation de facteurs de transcription et de l'expression des gènes.

Ces études de relations structure/sélectivité, structure/affinité et cellulaires que nous envisagions nous ont amené à proposer à l'équipe du Professeur David Boykin la synthèse de nouveaux composés potentiellement plus affins et plus spécifiques. C'est pourquoi nous avons dans ce travail étudié i) l'affinité et la sélectivité dans la reconnaissance de l'ADN de ces nouveaux dérivés ainsi que leurs conséquences cellulaires en termes de localisation cellulaire et de cytotoxicité et ii) la modulation de la fixation des facteurs de transcription et de l'expression des gènes impliqués dans la cancérogenèse par ces nouvelles molécules qui ciblent des séquences nucléotidiques spécifiques de l'ADN.

Relations de structure/activité, structure/affinité et conséquences cellulaires

Structure/activité, structure/affinité : Nous disposons d'un large panel de molécules pour compléter les connaissances sur les relations de structure/affinité et cibler des sites consensus variés avec une meilleure affinité pour l'ADN (Figure 31).

Des expériences d'empreinte à la DNase I et de spectroscopie pour mesurer les variations de température de fusion de l'ADN ont mis en évidence les capacités de nouveaux composés du type diphenyl-furane pour reconnaître les séquences riches en paires de bases AT (DB351, DB484, DB544 et DB1017). Les DB351 et DB544 (Figure 31) où le cycle furane central est remplacé respectivement par des cycles thiofène et oxazole reconnaissent spécifiquement les sites riches en paires de bases AT avec une affinité proche de celle du DB75. Ainsi, des modifications minimales de l'angle formé par la molécule ne modifie ni la spécificité ni l'affinité à l'ADN par rapport au composé de référence le DB75. Si la modification de cet angle est plus importante comme pour la molécule qui porte ses fonctions amidines terminales en position méta-para (DB555), l'interaction et la spécificité de reconnaissance du composé pour l'ADN est bien moindre. D'autre part, l'ajout de 2 groupements méthyles en position méta des cycles phényles (DB936) empêche la reconnaissance spécifique de l'ADN. De façon surprenante, l'ajout d'atomes de fluor, (qui possèdent un encombrement stérique proche du groupement méthyle sur ces mêmes positions (DB1017) n'influe pas sur la sélectivité de reconnaissance des sites riches en paires de bases AT mais diminue l'affinité de la molécule pour l'ADN. Pour ces nouveaux composés capables de reconnaître des séquences riches en paires de bases AT, la mesure des variations de température de fusion d'un oligonucléotide qui présente une mutation sur le site riche en paires de bases AT (Figure 33) montre que les composés spécifiques (DB75, DB351, DB484, DB544 et DB1017) voient leur affinité relative pour l'ADN significativement diminuée (diminution de moitié) alors que les composés peu ou non spécifiques (DB555, DB359) voient leur affinité pour l'ADN peu modifiée. Pour le DB359 qui possède ses deux fonctions amidines terminales en position méta-méta et qui interagit avec l'ADN de façon non spécifique par intercalation, les expériences de dénaturation thermique indiquent un ΔT_m identique pour les deux oligonucléotides. Il existe donc une corrélation entre la spécificité de reconnaissance et l'affinité des molécules pour l'ADN. Les résultats préalables avaient déjà montré le rôle important de la structure et surtout de la courbure de la molécule dans la reconnaissance de l'ADN par les dérivés diphenyl-

furanes (Nguyen *et al.*, 2002). Les résultats de notre étude confirment donc ces premières données.

Dans la série phényl-furane-benzimidazole, seuls les DB293 et, à moindre mesure, DB704 avaient montré une reconnaissance spécifique des sites 5'-ATGA (Wang *et al.*, 2001). Dans cette étude, nous avons mis en évidence de nouvelles molécules du type phényl-furane-benzimidazole capables d'interagir aux sites ATGA (Figure 34). Dans un premier temps, nous avons démontré l'importance du cycle benzimidazole porté par la molécule dans cette reconnaissance spécifique car aucune des 6 molécules substituées au niveau de ce cycle n'a la capacité de se fixer au site 5'-ATGA. Ces résultats convergent avec les observations initiales obtenues avec 2 composés (DB75 et DB609) où ce cycle était modifié (Wang *et al.*, 2001). D'autre part, nous avons démontré que 1) des modifications mineures sur le cycle furane (cycle furane remplacé par un cycle imidazole pour le DB 940) n'altèrent pas la reconnaissance du site ATGA ; alors que 2) des modifications plus importantes sur le cycle furane (cycle furane remplacé par un cycle thiofène pour le DB 818) empêchent la reconnaissance du site ATGA mais pas celle des sites riches en paires de bases AT. Si on regarde l'angle formé par le DB940 par rapport à celui du DB293, les variations sont minimales, et comme l'atome d'azote et l'atome d'oxygène sont proches en terme d'encombrement, ceci peut expliquer la similitude d'interaction avec l'ADN des DB293 et DB940. Il est également possible de remplacer le cycle phényle par un cycle pyridine (DB 995 et 915) sans altérer la reconnaissance spécifique du composé au site 5'-ATGA. Enfin, nous avons démontré l'importance de la position méta du cycle phényle dans la reconnaissance du site ATGA puisque l'ajout d'un groupement hydroxyle ou méthyle respectivement pour le DB928 et le DB850 en position ortho n'altère pas la reconnaissance du site ATGA ; l'ajout de ces mêmes groupements en position méta respectivement pour le DB992 et le DB990, empêche la reconnaissance du site ATGA. Cependant, comme pour la série diphenyl-furane (DB1017), l'ajout d'un atome de fluor en position méta (DB989) ne modifie pas la reconnaissance du site ATGA. Dans cette famille, concernant la reconnaissance des sites riches en paires de bases AT, seules certaines modifications importantes sur le cycle benzimidazole empêchent cette reconnaissance. Là encore des résultats préalables avaient déjà montré le rôle important de la structure et surtout de la courbure de la molécule dans la reconnaissance de l'ADN pour les dérivés phényl-furane-benzimidazoles (Mallena *et al.*, 2004). Les résultats de notre étude confirment également ces données.

Dans cette famille de molécules, même si la reconnaissance du site ATGA joue un rôle majeur dans leur activité, il est important de découvrir de nouveaux composés capables de reconnaître des sites originaux, ce qui est le cas pour les composés DB1242, RT-29 et DB1255. Le DB1242 (Figure 36) possède une structure proche de celle du DB921 (Figure 18). Sa structure plane devrait empêcher son ancrage correct au niveau du petit sillon de l'ADN, mais de manière surprenante, nous avons montré que le DB1242 se fixe sur une séquence d'ADN spécifique riche en paires de bases GC (site GCTCG), à la différence du DB921 qui se fixe sur une séquence riche en paires de bases AT. Cette molécule se lie à l'ADN sous forme de dimère tout-à-fait original (Figure 66), ainsi le dimère formé ne se présente pas dans une orientation « tête-bêche » comme pour le DB293. Il existe en fait des interactions entre un des atomes d'azote du cycle pyrimidine et un des groupements amidines porté par la seconde molécule et inversement. Cette structure dimérique possède alors une courbure proche de celle du petit sillon où elle interagit avec le site GCTCG en établissant des liaisons hydrogène spécifiques sur les différentes paires de bases (Figure 63). L'ensemble de ces résultats sont publiés dans l'article de Mundle (Mundle et al., 2007).

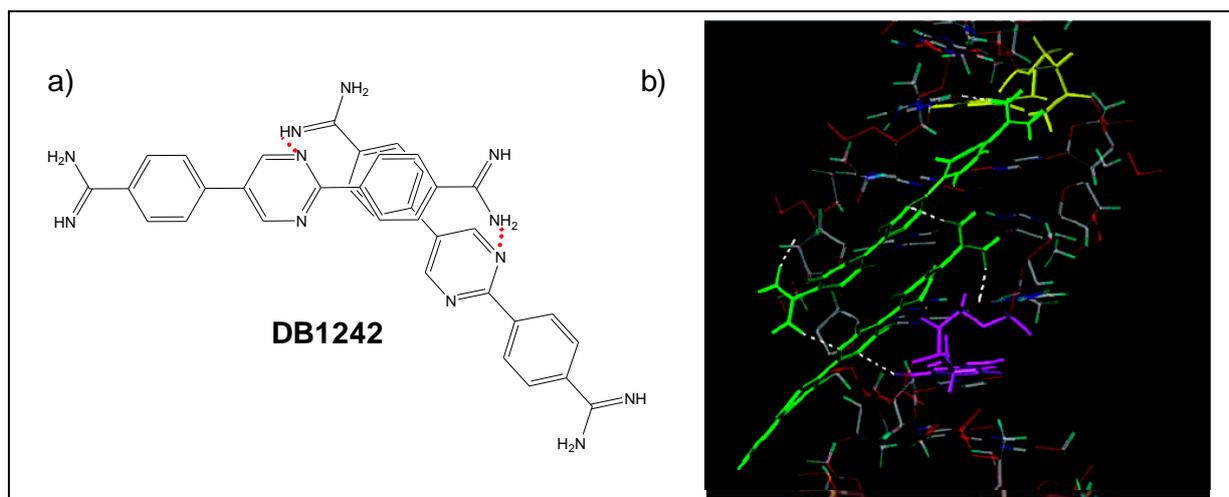


Figure 66 : Structure et mode d'interaction du DB1242 à l'ADN.

a) Structure du dimère formé entre deux molécules de DB1242. Ces deux molécules effectuent des liaisons hydrogène (pointillés rouges) entre les groupements amidines d'une extrémité et un des azotes du cycle pyrimidine. Ce dimère adopte un angle compatible à la courbure de l'ADN et interagit sous forme de dimère à la séquence GCTCG.

A l'inverse du DB1242, le RT-29 possède une structure beaucoup plus courbée que les composés DB de référence (DB75 et DB293). Ce type de structure est là encore généralement défavorable à une bonne interaction avec le petit sillon de l'ADN. Cependant, nous avons montré que cette molécule à forte courbure possède une affinité relative forte pour les sites riches en paires de bases AT et reconnaît également la séquence 5'-ATTC. Pour reconnaître

ces différents sites, le RT-29 effectue des liaisons hydrogène indirectes avec le petit sillon via une molécule d'eau (Figure 67). Ainsi, pour reconnaître le site AATT, le groupe NH du cycle benzimidazole et le groupement amidine accolé effectuent des liaisons hydrogène directes avec le petit sillon de l'ADN. De l'autre côté, le cycle phényl-amidine interagit avec le petit sillon par l'intermédiaire d'une molécule d'eau. Pour la reconnaissance du site ATTC, la position de la molécule d'eau est inversée et effectue des contacts avec l'extrémité benzimidazole-amidine qui est poussée hors de l'axe du petit sillon par le groupement NH₂ de la guanine du brin complémentaire. Ces résultats sont publiés dans l'article de Tanious (Tanious *et al.*, 2007).

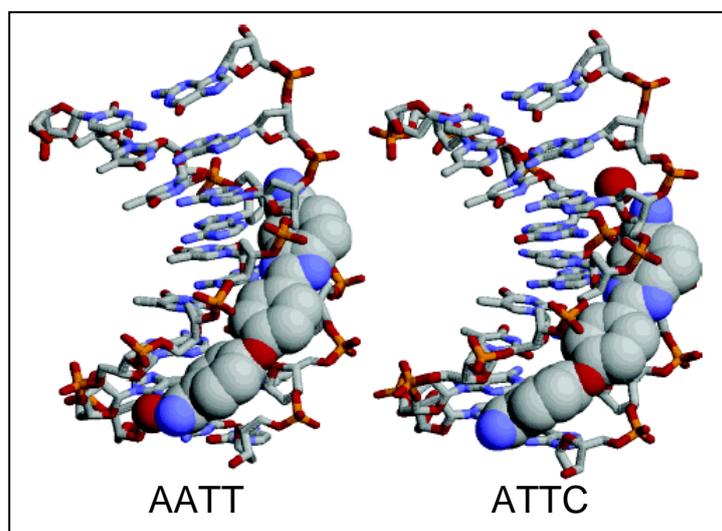


Figure 67 : Structure et mode d'interaction du RT-29 à l'ADN.

Modélisation de l'interaction entre le RT-29 et l'ADN. Le RT-29 possède une courbure plus importante que l'ADN et interagit par l'intermédiaire d'une molécule d'eau (représentée en rouge). La molécule d'eau se place au niveau de la première base sur le site AATT, alors que pour le site ATTC, la molécule se place au niveau de la dernière base sur l'appariement G•C.

Ces deux molécules, qui sont à première vue incompatibles avec les règles qui régissent l'interaction au niveau du petit sillon de l'ADN en terme de respect de la courbure du petit sillon (Kopka *et al.*, 1985 ; Neidle 2001), interagissent pourtant efficacement avec l'ADN. Cette interaction passe par une adaptation de la molécule à la forme du petit sillon par l'intermédiaire de liaisons indirectes avec des molécules d'eau (RT-29) ou par la formation d'un dimère original (DB1242) générant ainsi la courbure nécessaire à l'interaction dans le petit sillon. Il apparaît donc difficile de prédire si une molécule peut interagir ou non avec l'ADN tant est grande la diversité des structures que la molécule peut adopter (seule, en dimère, ou par l'intermédiaire de molécules d'eau).

Pour le DB1255, les tests d'empreinte à la DNase I mettent en évidence un site de reconnaissance très original centré sur la séquence nucléotidique 5'-CCATG en plus d'une reconnaissance des sites riches en paires de bases AT (Figure 37). Les études réalisées sur le site consensus ETS/PEA3 ont révélé une reconnaissance spécifique du site 5'-GTTCG (Figure 57). Il est maintenant important de déterminer le mode d'interaction de ce composé original et son affinité pour ces sites de reconnaissance. Ces études sont actuellement en cours, en collaboration avec le Pr David Wilson.

Ainsi, dans cette première partie du travail, nous avons mis en évidence des dérivés diphenyl-furanes capables d'interagir fortement avec l'ADN (DB75, DB351, DB484, DB544,) et de reconnaître spécifiquement des séquences riches en paires de bases AT (DB75, DB351, DB484, DB544, DB1017) ou non (DB359, DB555). Dans la série phényl-furane-benzimidazole nous avons démontré que d'autres composés étaient capables de se lier fortement aux sites ATGA (DB293, DB940, DB989, DB915, DB928, DB850 et DB995). L'étude de nouvelles séries de composés nous a également permis de mettre en évidence des composés qui présentent des sites de reconnaissance originaux pour l'ADN (séquences riches en paires de bases GC, séquence nucléotidique 5'-CCATG et séquence ATTC respectivement par les composés DB1242, DB1255 et RT-29).

Conséquences cellulaires : devant l'intérêt de ces molécules, nous avons alors étudié les conséquences cellulaires de leur interaction avec l'ADN. Des premiers tests de distribution cellulaire avaient montré que les dérivés DB pouvaient se localiser différemment dans la cellule en fonction des types de substitutions (Lansiaux *et al.*, 2002 b), de la charge des molécules (Lansiaux *et al.*, 2002 a) et que la mitochondrie pouvait jouer un rôle dans leur mécanisme d'action cellulaire. Nous avons mis à profit les capacités de fluorescence des nouveaux dérivés pour essayer d'affiner ces résultats et de corréler les conséquences cellulaires aux mécanismes définis au préalable. Dans un premier temps, nous avons montré que les composés diphenyl-furanes diversement substitués au niveau du cycle furane (Figure 39a) pénètrent efficacement dans la cellule où les molécules les plus affines pour l'ADN se dirigent vers le noyau. En revanche, l'ajout de substituants méthyles sur les cycles phényles (DB936) empêche la molécule de pénétrer dans le noyau, alors que l'ajout d'atomes de fluor (DB1017) ne perturbe pas la localisation nucléaire de la molécule. Le cycle furane n'est donc pas primordial dans la localisation sub-cellulaire des composés. Par contre, la substitution des

cycles phényles joue un rôle beaucoup plus important. Ces substitutions se font aussi au détriment de la spécificité de reconnaissance et de l'affinité du composé pour les sites riches en paires de bases AT (Figure 33).

Ces résultats renforcent ceux déjà obtenus par notre groupe dans la mesure où : i) les amidines terminales apparaissent comme étant les éléments déterminants dans la pénétration des composés DB dans la cellule (Lansiaux *et al.*, 2002a) ; ii) les composés DB les plus affins pour l'ADN possèdent tous une localisation nucléaire (Lansiaux *et al.*, 2002b). Maintenant, nous avons aussi mis en évidence que la substitution du groupe phényle jouerait également un rôle important pour la distribution cellulaire des composés diphényl-furanes.

Dans un second temps, puisque certaines équipes avaient démontré l'influence du mode expérimental de fixation des cellules sur la distribution des polyamines au sein de la cellule (Belitsky *et al.*, 2002), nous avons comparé la distribution cellulaire d'un composé nucléaire (DB75) et d'un composé cytoplasmique (DB936) sur des cellules non fixées, fixées au PFA à différentes concentrations, fixées et perméabilisées par du méthanol et enfin fixées et perméabilisées au PFA/NP40. Les composés N-méthyl-pyrrole/N-méthyl-imidazole qui possèdent normalement une localisation cytoplasmique sur des cellules non fixées se délocalisent dans le noyau après fixation des cellules SKBR-3 au méthanol (Belitsky *et al.*, 2002). Dans notre étude, les mêmes résultats ont été observés pour les dérivés DB, dont la fixation des cellules au méthanol modifie la distribution cellulaire du DB936 qui se délocalise alors dans le noyau (Figure 41). L'utilisation d'une solution de NP40 2% PFA 2% délocalise elle aussi le DB936 dans le noyau (Figure 42). Ce résultat montre que la perméabilisation des membranes influe sur la localisation des composés. Par contre la simple fixation des cellules au PFA, quelque soit la concentration utilisée, ne modifie pas la localisation nucléaire (DB75) ou cytoplasmique (DB936) des composés (Figure 40). Ceci valide d'une part l'utilisation du PFA dans la fixation des cellules pour l'étude de la distribution cellulaire et confirme le rôle important que peuvent jouer les membranes (cytoplasmiques, mitochondriales...) dans la localisation des composés DB.

Après l'étude de la distribution cellulaire des composés DB, les conséquences cellulaires en termes de cytotoxicité ont été évaluées. Ainsi, les résultats présentés montrent tout d'abord qu'à l'exception des composés DB936 et DB1255 qui possèdent une cytotoxicité intermédiaire (de l'ordre de 5 μ M), ces molécules ne peuvent pas être considérées comme des agents cytotoxiques classiques vu leur faible cytotoxicité (Tableau 5). Ceci est en accord avec le mode d'action envisagée. On observe tout de même une graduation de la toxicité entre les

différentes familles de composés. Ainsi, le composé phényl-thiofène-benzimidazole (DB818) est moins cytotoxique que son équivalent diphényl-thiofène (DB351), lui-même moins cytotoxique que son analogue diphényl-dithiofène (DB1255). On remarque de manière générale que les composés diphényl-furanes (DB75, DB351) ont un effet cytotoxique supérieur à leurs analogues dans la série phényl-thiofène-benzimidazole (DB293, DB818). De plus, comme il a déjà été montré par notre équipe, on observe que les composés qui possèdent une localisation cytoplasmique (DB936) possèdent une cytotoxicité supérieure à celle des composés nucléaires dans une même série.

L'ensemble de ces résultats complète les règles de structure/sélectivité, de structure/affinité et de structure/effet cellulaires précédemment définies pour les composés DB. Ils montrent la capacité de liaison à l'ADN et la spécificité de reconnaissance de nouveaux composés DB. Ceci permet là encore d'orienter les synthèses de composés capables de reconnaître l'ADN i) avec une meilleure affinité ; ii) avec une plus grande sélectivité par une reconnaissance de sites supérieurs à 4 pb, comme c'est le cas pour le composé DB1242 (5 pb) ; iii) sur de nouveaux sites afin de pouvoir cibler un panel de facteurs de transcription variés.

Modulation de la fixation des facteurs de transcription

Cette première partie de l'étude a permis de caractériser et de sélectionner des composés DB candidats intéressants dans l'étude de la modulation des facteurs de transcription. Ces molécules, tout comme d'autres familles de composés, interagissent fortement au niveau du petit sillon de l'ADN selon un modèle ancien de donneur et d'accepteur de liaisons hydrogène (Seeman *et al.*, 1976). Cette approche est appréhendée par de nombreuses équipes qui travaillent sur des composés variés. L'exemple des lexitropsines est le plus documenté et représente un modèle intéressant. Pour ce type de molécules, un code simple basé sur le type et l'enchaînement des cycles qui composent la molécule permet de cibler spécifiquement une séquence nucléotidique donnée (Wemmer *et al.*, 2000 ; Dervan *et al.*, 2005). Ainsi, il apparaît que ces composés inhibent la fixation à l'ADN d'un certain nombre de facteurs de transcription comme TFII-D, NFκB ou encore NF-Y (Ehley *et al.*, 2002 ; Wurtz *et al.*, 2002 ; Hochhauser *et al.*, 2007). Cependant, ces composés possèdent une pénétration et une distribution cellulaire difficilement prévisible ce qui aboutit à certains échecs dans l'inhibition de facteurs de transcription au niveau cellulaire (Belistky *et al.*, 2002).

Afin d'étudier la modulation de la fixation des facteurs de transcription à l'ADN par les composés DB, nous avons choisi deux stratégies différentes : une première, totalement originale basée sur **un criblage large de facteurs de transcription** pour un composé donné indépendamment du site consensus de reconnaissance de ces protéines ; une seconde, plus classique, sur un facteur de transcription donné reconnaissant un site consensus potentiellement reconnu par les composé DB, par **un criblage de composés DB**.

Criblage large de facteurs de transcription : pour mettre en place cette première approche originale, nous avons sélectionné le DB293, chef de file des composés phényl-furane–benzimidazoles, pour lequel le mode d'interaction à l'ADN est le mieux détaillé (Wang *et al.*, 2000 ; Bailly *et al.*, 2001 ; Tanius *et al.*, 2003) et parce qu'il possède une affinité relative pour l'ADN comparable à celle déterminée pour les autres composés de ce type.

Pour le DB293, nous avons utilisé les membranes Transignal/Protein DNA array I (Panomics) (Lam et Li ; 2002) pour réaliser un criblage rapide de 54 facteurs de transcription selon un mode compétitif. Nous avons montré que le DB293 inhibait un certain nombre de facteurs de transcription qui possèdent dans leur site consensus (Figure 44) i) la séquence ATGA, comme Pit-1 et Brn-3 ; ii) un site riche en paires de bases AT comme MEF-2, Stat1, Stat3 et Stat 4 ;

iii) une séquence consensus sans site ATGA ni site riche en paires de bases AT pour les protéines ETS. Nous avons alors validé ces résultats par des techniques conventionnelles de retard en gel sur des extraits nucléaires de cellules HT-29 et par l'utilisation de protéines exprimées en lysat de réticulocytes pour les protéines Pit-1, Brn-3 (Figure 46 et 47), IRF-1 (Figure 46), HNF-4 (Figure 47) et ETS (Figure 56). Nous avons alors focalisé notre attention sur les protéines à POU domaine Pit-1 et Brn-3 ainsi qu'au facteur IRF-1 (non inhibé par le DB293), qui possèdent tous les séquences ATGA et les séquences riches en paires de bases AT dans leur site consensus.

Les protéines à POU domaine sont uniques par la présence dans leur domaine de reconnaissance à l'ADN de deux entités structurellement indépendantes, les domaines POU_H et POU_S reliés par un linker variable qui confère une grande flexibilité d'interaction à l'ADN (Herr et Cleary, 1995 ; Jacobson *et al.*, 1997 ; Phillips et Luisi., 2000). Brn-3 et Pit-1 interagissent sur des sites différents qui possèdent une séquence 5'-ATGA dans un contexte riche en paires de bases AT : 5'-TTA**ATGA** pour Brn-3 et 5'-AAT**ATGA**AATAA pour Pit-1. La protéine IRF-1 possède aussi une séquence 5'-ATGA dans un environnement riche en paires de bases AT sur son site consensus mais elle n'est pas inhibée par le DB293 (5'-AAA**ATGA**AATT). Ainsi, la seule présence du site 5'-ATGA sur le site consensus d'un facteur de transcription ne suffit pas à expliquer une inhibition efficace de l'activité de liaison à l'ADN des facteurs de transcription.

Pour comprendre le mode d'inhibition ou non de la fixation de ces 3 protéines, une analyse fine de la fixation du DB293 sur les séquences consensus a été réalisée. Ainsi, le mode d'inhibition des facteurs de transcription à POU domaine et le rôle de la reconnaissance du site 5'-ATGA ont pu être évalués par l'utilisation concomitante d'expériences d'empreinte à la DNase I, de résonance plasmonique de surface et de retard en gel réalisés sur les facteurs de transcription Brn-3, Pit-1 et IRF-1.

Pour Brn-3, l'interaction du DB293 passe par la formation d'un dimère coopératif sur le site ATGA et par une reconnaissance du site riche en paires de bases AT situé côté 5'- de la séquence ATGA (Figure 48b et 49). Ainsi, ce sont 3 molécules DB293, (deux sous forme de dimère et une sous forme monomérique) qui peuvent se lier au site consensus Brn-3.

Pour la reconnaissance du site Pit-1 par le DB293, l'analyse s'est avérée plus complexe par la présence de deux sites riches en paires de bases AT de part et d'autre du site ATGA (Figure 48a et 49). Cette reconnaissance a néanmoins pu être identifiée par l'utilisation de séquences Pit-1 mutées (Figure 51 à 53). Ainsi, cette reconnaissance s'effectue par l'interaction possible de 4 molécules de DB293, 2 sous forme de dimère coopératif sur la séquence ATGA et 2 sous forme de monomères sur les sites riches en paires de bases AT en 3' - et 5' - du site ATGA.

En ce qui concerne IRF-1, le DB293 interagit selon un mode non coopératif sur les deux sites riches en paires de bases AT qui juxtaposent la séquence 5'-ATGA (Figure 48c et 49). Ainsi, par opposition aux deux cas précédents, il s'avère que le site ATGA dans cet environnement n'est pas reconnu par le DB293.

Il a précédemment été montré que les composés DB interagissaient fortement sur les sites qui contiennent un enchaînement d'adénine (A_n) ou d'adénine-thymine (A_nT_n). Par opposition, les composés DB ne reconnaissent que faiblement les enchaînements thymine-adénine (T_nA_n) (Tanious *et al.*, 2003). Pour Brn-3, la reconnaissance du site TTAAT est faible et n'interfère donc pas avec la reconnaissance du site 5'-ATGA. Pour le site Pit-1, le mode de reconnaissance est plus complexe car il existe des courts enchaînements A_nT_n et T_nA_n . Il existe une compétition entre la reconnaissance sous forme de monomère et de dimère. Ceci est d'ailleurs vérifié par les expériences d'empreinte à la DNase I sur les séquences Pit-1 mutées (Figure 51 et 52). En revanche, les deux séquences riches en paires de bases AT du site IRF-1 sont du type A_5T et A_3T_2 . Ces sites sont donc fortement reconnus par le DB293 sous forme de monomère, ce qui empêche la reconnaissance en dimère sur le site 5'-ATGA. Le mode de fixation du DB293 sur ces 3 sites consensus est donc en parfait accord avec les observations faites sur le mode de fixation des composés DB sur les sites riches en paires de bases AT.

D'après ces résultats, nous pouvons émettre l'hypothèse que l'inhibition ou non de la fixation des facteurs de transcription par le DB293 dépendrait du mode de reconnaissance sous forme de dimère du site 5'-ATGA. D'après les données sur : i) l'inhibition ou non de ces facteurs de transcription par le DB293 ; ii) sur le mode d'interaction précis de cette molécule sur ces 3 sites consensus ; iii) et à partir des données bibliographiques qui relatent du mode d'interaction de ces 3 protéines à l'ADN, il est possible d'établir des modèles de la modulation de ces protéines par le DB293.

Modèles de la modulation de Brn3, Pit1, IRF1 par le DB293 : tout d'abord pour Brn-3, à notre connaissance, la structure cristallographique du complexe ADN/Brn-3 n'a pas été déterminée. Il est néanmoins possible que le mode d'interaction de la protéine Brn-3 à l'ADN soit similaire à celui du facteur de transcription à POU domaine Oct-1. La protéine Oct-1 interagit avec une forte affinité sur le site ATGCAAAT sur la zone promotrice du gène codant pour l'histone H2B (Klemm *et al.*, 1994). A partir de ces informations, nous suggérons donc que le domaine POU_H de la protéine Brn-3 interagirait avec le site riche en paires de bases AT de la séquence consensus Brn-3 5'-GTTAATGAG et que le domaine POU_S interagirait avec le site ATGA (Figure 68a). Sur ce site le DB293 interagit sous forme de dimère sur la séquence 5'-ATGA et sous forme de monomère sur le site TTAA (Figure 68b). La fixation de ces molécules au niveau du petit sillon de l'ADN peut empêcher l'interaction des deux domaines de reconnaissance de la protéine Brn-3. Ainsi, l'inhibition de l'interaction de la protéine Brn-3 serait due à l'agrandissement du petit sillon par le dimère de DB293 sur le site ATGA et donc à la réduction du grand sillon sur le côté opposé, ce qui empêcherait le positionnement correct des 2 domaines de liaison de la protéine Brn-3 sur son site consensus.

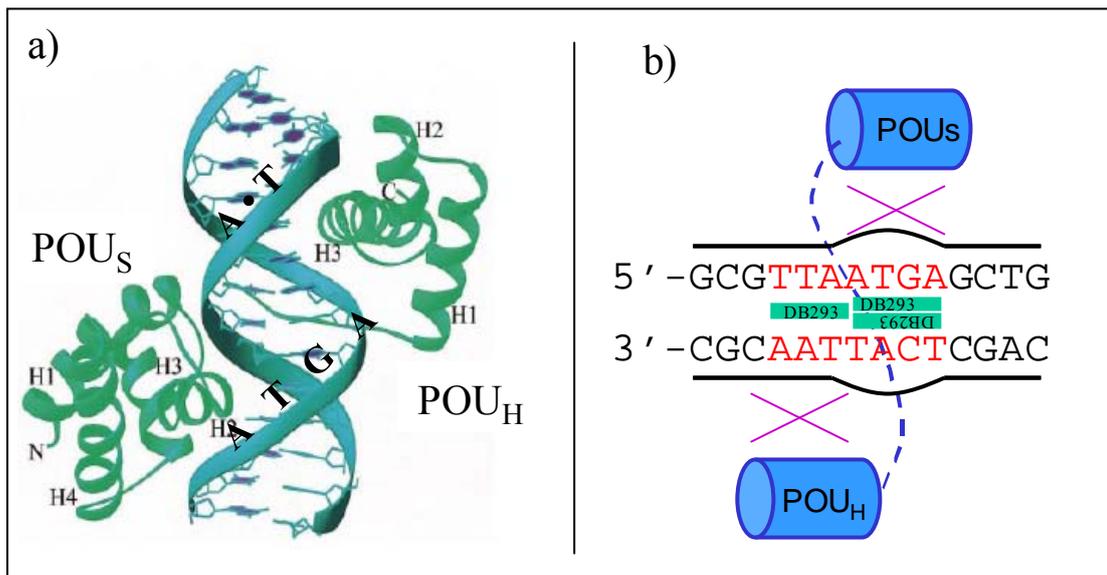


Figure 68 : Modèle de l'inhibition du complexe ADN / Brn-3 par le DB293

a) Modèle du complexe ADN/Brn-3 réalisé à partir du complexe ADN/Oct-1 (Klemm *et al.*, 1994). Le domaine POU_H interagirait avec le site riche en paires de bases AT alors que le domaine POU_S interagirait avec le site ATGA. b) Inhibition du complexe ADN/Brn-3 par le composé DB293. La formation d'un dimère de composé DB293 induirait un changement structurel des sillons ayant pour conséquence l'inhibition de la fixation des domaines POU_H et POU_S sur le site consensus Brn-3.

Pour Pit-1, la structure cristallographique de différents complexes ADN/Pit-1 a été déterminée sur le promoteur du gène codant pour l'hormone de croissance GH1 et sur le promoteur du gène codant pour la prolactine (Jacobson *et al.*, 1997 ; Scully *et al.*, 2000). D'après ces données, nous avons proposé que le domaine POU_H se lierait via les résidus Arg 46, Val 47, Cys 50, Asn 51 et Gln 54 de l'hélice α_3 (hélice de reconnaissance) au site 5'- ATGA au niveau du grand sillon de l'ADN. De plus, il existerait des contacts entre le résidu Arg 5 du bras N-terminal du domaine POU_H et certaines paires de bases du site consensus Pit-1 au niveau petit sillon de l'ADN. Le domaine POU_S effectuerait des contacts au niveau du site riche en paires de bases AT du grand sillon de l'ADN par l'intermédiaire des résidus Ser 43, Gln 44, Thr 45 et Arg 49 de l'hélice α_3 (Figure 69). En accord avec nos résultats, le DB293 interagirait sous forme de dimère sur le site ATGA et sous forme de monomère sur les deux sites riches en paires de bases AT. L'interaction potentielle de ces 4 molécules sur le site Pit-1 inhibe efficacement la fixation du facteur de transcription Pit-1.

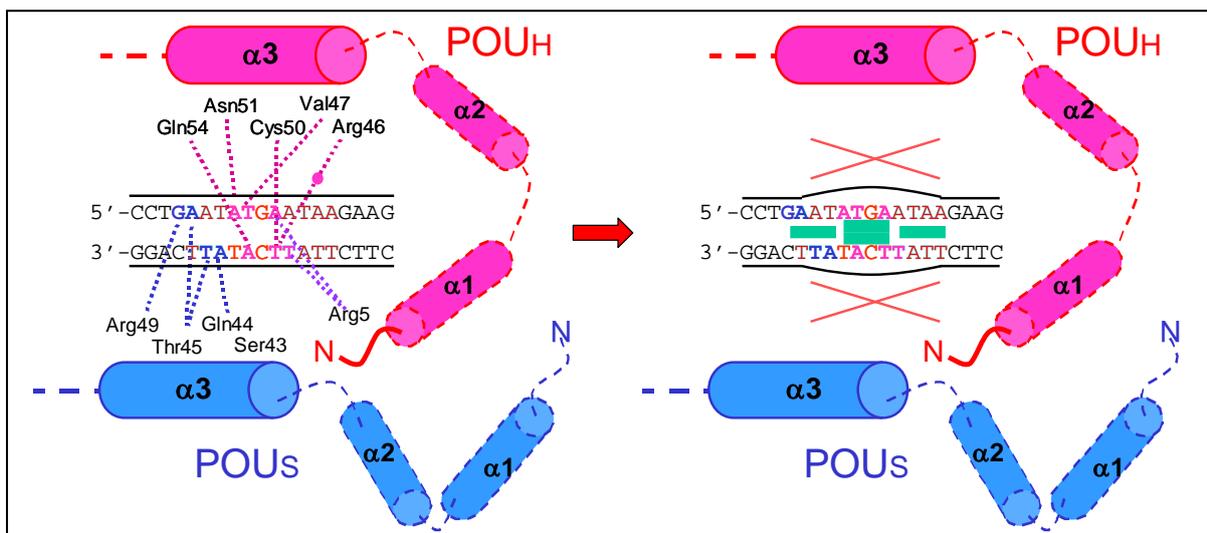


Figure 69 : Inhibition de l'interaction de la protéine Pit-1 sur son site consensus par le DB293.

Les hélices α_3 des domaines POU_H et POU_S effectuent de multiples liaisons hydrogène au niveau du grand sillon de l'ADN (traits pointillés roses pour POU_H et traits pointillés bleus pour POU_S). Le bras N-terminal reconnaît l'ADN et effectue des liaisons hydrogène au niveau du petit sillon (traits pointillés violets). La fixation d'un dimère de composé DB293 sur le site ATGA et de deux monomères sur les séquences riches en paires de bases AT empêche la reconnaissance et le positionnement correct du facteur de transcription Pit-1 sur sa séquence cible.

Nous avons alors étudié l'influence d'autres composés DB structurellement proches du DB293 et reconnaissant exclusivement les sites riches en paires de bases AT (DB990 et DB992) ou reconnaissant des sites ATGA (DB850 et DB995) sur la fixation des protéines Brn-3, Pit-1, et IRF1. Cette analyse a permis de démontrer le rôle primordial de la reconnaissance du site ATGA dans l'inhibition des facteurs de transcription à POU domaine

(Figure 55). Sur cette base, nous avons démontré que les DB850 et DB995, capables d'interagir au niveau des sites ATGA inhibaient plus efficacement l'interaction entre les protéines Pit-1, Brn-3 et l'ADN que les composés DB990 et DB992 qui se lient aux sites riches en paires de bases AT. Ainsi, les modifications structurales induites par la formation du dimère au niveau du petit sillon de l'ADN représentent un élément déterminant pour une inhibition efficace de la fixation des facteurs de transcription.

En ce qui concerne la protéine IRF-1, d'un point de vue structural, le domaine de reconnaissance de la protéine est composé d'une succession d'hélices α , de feuillets β et de linker ($\alpha 1$ - $\beta 1$ - $\beta 2$ -L1- $\alpha 2$ -L2- $\alpha 3$ - $\beta 3$ -L3- $\beta 4$) (Escalante *et al.*, 1998). La topologie des hélices $\alpha 2$ et $\alpha 3$ ressemble à celle du domaine de type HTH. L'hélice $\alpha 3$ interagit avec le grand sillon de l'ADN et joue le rôle de l'hélice de reconnaissance. Cependant, son orientation est différente de celle observée pour le motif HTH. Dans le motif HTH, l'hélice de reconnaissance traverse le grand sillon de l'ADN perpendiculairement à l'axe de l'ADN, alors que pour IRF-1 elle est tangente à l'axe et parallèle au squelette phosphodiester (Figure 70a). Les contacts entre l'ADN et la protéine s'effectuent entre les paires de bases du site GAAA et les résidus Arg 82, Cys 83, Asn 86 et Ser 87 situés sur l'hélice $\alpha 3$ (Figure 70b). Il existe d'autres contacts entre l'ADN et la protéine qui stabilisent la structure. Ainsi, les Trp 11 (de l'hélice $\alpha 1$) et le Trp 58 (de l'hélice $\alpha 2$) effectuent des interactions hydrophobes et des liaisons hydrogène avec le squelette phosphodiester de l'ADN. La chaîne phosphodiester effectue aussi des liaisons hydrogène avec divers résidus situés sur les linker L1, L2 et L3. Nous avons vu précédemment que sur le site consensus IRF-1, deux molécules de DB293 interagissaient de part et d'autre des paires de bases centrales GA (5'-AAAATGAAATTT). L'interaction du DB293 sur le site consensus IRF-1 n'empêche donc pas l'interaction entre l'ADN et le facteur de transcription (Figure 70c). De plus, comme le DB293 n'interagit pas sous forme de dimère sur ce site ATGA, mais sous forme de deux monomères sur les sites AAAT et AATTT, les modifications structurales des sillons dues à la fixation des composés sont insuffisantes pour bloquer le positionnement correct de l'hélice $\alpha 3$ au niveau du grand sillon de l'ADN.

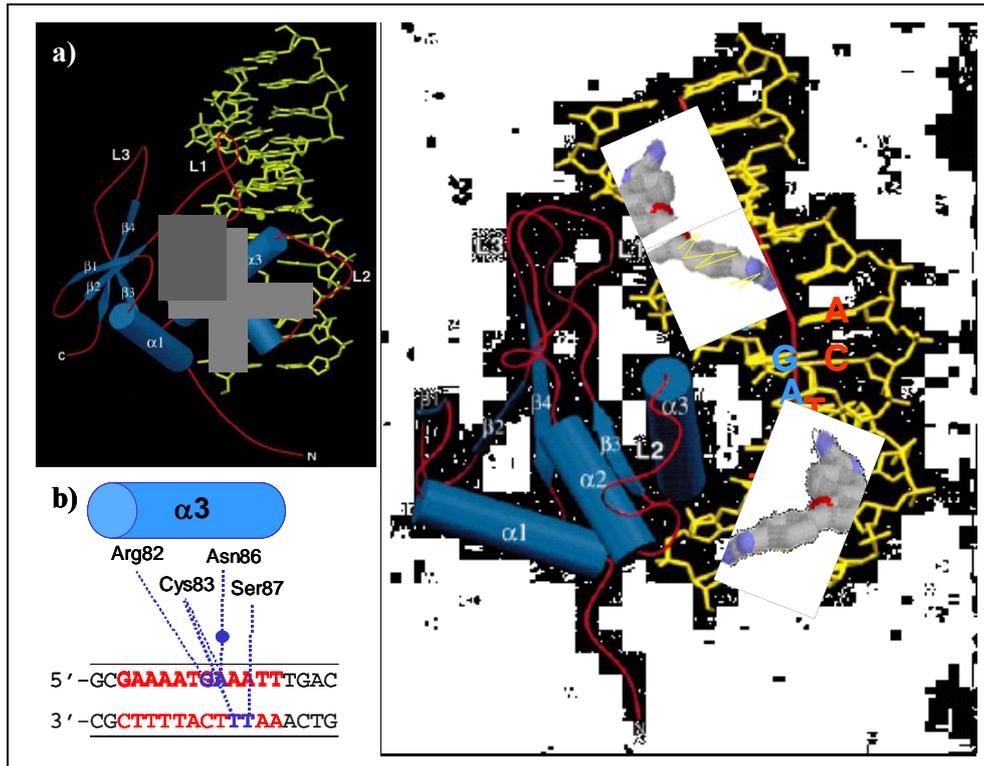


Figure 70 : Modèles d'interaction des complexes IRF-1/ADN et IRF-1/ADN/DB293

a) Complexe IRF-1/ADN. La protéine IRF-1 est composée d'une succession d'hélices α , de feuillets β et de linker. Les hélices $\alpha 2$ et $\alpha 3$ forment un pseudo motif HTH tangent à l'axe de la double hélice. En gris est représentée l'orientation normale du motif HTH perpendiculaire à l'axe de l'ADN. b) Interactions entre l'hélice $\alpha 3$ et l'ADN. Les traits pointillés bleus représentent les interactions hydrogènes qui s'établissent entre les paires de bases de l'ADN et les différents résidus. c) Complexe ADN/IRF-1/composés DB293. Les interactions entre l'ADN et la protéine IRF-1 entraîne une courbure de l'axe de l'ADN (trait rouge). La fixation des deux molécules de DB293 n'empêche pas la fixation du facteur de transcription sur son site consensus. (Modélisation tirée de Escalante *et al.*, 1998).

En perspective, l'inhibition de la fixation des facteurs de transcription Pit-1 et Brn-3 par le DB293 sera validée sur des modèles cellulaires. Pour cela, différentes approches sont envisagées comme l'étude de l'inhibition d'expression d'un gène rapporteur ou encore par des expériences de chromatine immunoprécipitation (ChIP). Tout d'abord, nous déterminerons si le DB293 a la capacité d'inhiber l'expression d'un gène rapporteur placé sous la dépendance des sites consensus Pit-1 et Brn-3. Ces expériences seront réalisées sur différents sites consensus, ces protéines possédant une flexibilité de reconnaissance importante (Cleary *et al.*, 1995). Cette flexibilité a d'ailleurs été vérifiée pour la protéine Pit-1 par des tests de retard en gel réalisés sur un panel d'oligonucléotides mutés sur le site consensus de Pit-1. Ainsi, il s'avère nécessaire de muter au moins 2 des 3 sites d'intérêts (sites 5'-AT, ATGA ou 3'-AT) pour inhiber la fixation de la protéine, la mutation d'un seul site n'altérant pas l'interaction

entre la protéine et sa séquence cible (Figure 54). Pour réaliser les tests d'activité luciférase, nous avons d'ores et déjà sélectionné un panel de sites consensus reconnus par les protéines Pit-1 et Brn-3. Pour Pit-1, on retrouve de nombreux sites consensus qui contiennent ou non le site ATGA avec des séquences environnantes variées (Andersen et Rosenfeld, 1994 ; Matre *et al.*, 1999) (Tableau 3). L'étude de la modulation de l'activité luciférase sur ces différents promoteurs validerait l'effet cellulaire de l'inhibition de la fixation du facteur de transcription Pit-1 à sa séquence cible. Nous commencerons l'étude par le promoteur du gène prolactine 3 (Prl-3P), site utilisé lors de notre étude.

Sites	Gènes
CATGA A ATAATTTG	Pit-1 b
CATGA A TATATATA	Prl-1P
TATGA A ATAAGAAA	Prl-3P
GATGA A TTTTCAAT	TSHbeta-2
GATGA A TGATTAAT	SS1R fp2
CAAATG A AAAAAACAT	GH-N LCR3

Tableau 3 : Exemples de sites consensus reconnus par le facteur de transcription Pit-1

Dans la colonne de gauche sont répertoriés quelques sites consensus reconnus par le facteur de transcription Pit-1. Les séquences ATGA sont représentées en gras et les séquences riches en paires de bases AT sont soulignées. Les gènes où sont retrouvés ces éléments de réponse sont inscrits dans la colonne de droite.

Ensuite, des expériences de chromatine immunoprécipitation (ChIP) sur ces promoteurs seront également réalisées en absence ou en présence de DB293 afin de valider ces résultats. L'étude de ces deux protéines s'avère intéressante car il a été montré que ces facteurs de transcription étaient surexprimés dans différents processus de cancérisation. Ainsi, la surexpression de la protéine Pit-1 serait impliquée dans des phénomènes de prolifération dans des modèles de cancer du sein (Gil-Puig *et al.*, 2005 ; Holtkamp *et al.*, 1984). Il serait alors intéressant de regarder l'expression de protéines directement dépendantes de l'activité du facteur de transcription Pit-1 comme la prolactine ou encore de l'hormone de croissance dont la dérégulation est associée à ces phénomènes de tumorigenèse. En ce qui concerne les protéines Brn-3, nous testerons l'effet du DB293 dans des modèles de cancer du col de l'utérus où l'expression de la protéine Brn-3a peut être augmentée d'un facteur 300 dans des biopsies réalisées sur des patientes atteintes de ce type de cancer (Ndisang *et al.*, 1998). Ces tests pourront aussi être réalisés sur certains modèles cellulaires de cancer du sein où la surexpression de cette protéine induit une inhibition de l'expression de la protéine anti-

oncogène BRCA-1 (Budhram-Mahadeo *et al.*, 1999). Nous pourrions alors déterminer le taux d'expression de cette protéine par RT-PCR et par Western blot après traitement des cellules avec le DB293.

En plus des facteurs à POU domaines, les membranes Transignal/Protein DNA array montrent l'inhibition par le DB293 d'autres classes de facteurs de transcription comme la famille ETS qui ne possède ni site riche en paires de bases AT, ni séquence ATGA dans son site consensus. Cette famille de facteurs de transcription est impliquée dans de nombreux processus cellulaires et est caractérisée par son domaine de liaison à l'ADN appelé EBS (Seth *et al.*, 2005).

D'autre part, l'analyse des résultats obtenus sur les membranes Transignal/Protein DNA array n'est pas encore terminée. Il s'avère que certains facteurs de transcription de la famille STAT (*Signal Transducer and Activator of the Transcription*) (Stat1, Stat3 et Stat4) seraient inhibés par le DB293. Le facteur de transcription Stat4 se lie au site 5'-CTGATTTCCCGAAATGATGA, cette séquence possède donc en 5'- un site riche en paires de bases AT et, chose intéressante, elle contient du côté 3'- deux sites ATGA consécutifs. Ainsi, l'interaction du DB293 sur ces différents sites juxtaposant le site consensus strict de reconnaissance de la protéine Stat4 pourrait empêcher sa fixation à l'ADN. Les protéines Stat1 et Stat3 reconnaissent respectivement les sites consensus 5'-AATTCCTGTAAG et CTTCTGGGAATTC, qui contiennent la séquence AATT (Yu, C.L. *et al.*, 1995 ; Chaturvedi *et al.*, 1997). Ce site est le site préférentiel de fixation des composés DB sur les séquences riches en paires de bases AT (Wang *et al.*, 2000 ; Bailly *et al.*, 2001). Il apparaît donc possible que la fixation du composé DB sur ces séquences inhiberait la fixation à l'ADN des protéines Stat1 et Stat3. Il sera intéressant alors d'évaluer les capacités d'inhibition de formation des complexes Stat/ADN par des composés DB de la famille des diphenyl-furanes n'interagissant qu'avec les sites riches en paires de bases AT. Ces facteurs de transcription Stat sont activés par les protéines kinases JAKs (*Janus tyrosine kinases*). Ils sont impliqués dans des processus variés comme la différenciation cellulaire, la prolifération, l'apoptose et l'angiogenèse (Horvath, 2000). Les protéines Stats représentent un modèle d'étude intéressant car certaines d'entre elles sont activées de manière constitutive dans des tumeurs primaires incluant les leucémies et les tumeurs solides des voies aéro-digestives supérieures, du sein, du cerveau et du pancréas (Calò *et al.*, 2003).

D'autres facteurs de transcription pourraient aussi être analysés par les membranes Transignal/Protein DNA array. A l'heure actuelle, il existe 5 membranes Transignals qui permettent de cribler 350 facteurs de transcription. Une même étude avec le DB293 sur ces nouvelles membranes pourrait donc être entreprise.

De plus, cette étude devrait être élargie sur l'ensemble de ces membranes Transignal/Protein DNA array à de nouveaux composés qui reconnaissent d'autres séquences de fixation comme les molécules DB1242, DB1255 et RT-29.

CONCLUSION GENERALE

L'étude de la reconnaissance spécifique des composés DB à l'ADN ainsi que leurs conséquences cellulaires en termes de cytotoxicité et de distribution cellulaire sont menées depuis une quinzaine d'années en collaboration avec les équipes des Prs David Wilson et David Boykin (Atlanta). Cette étroite collaboration a permis, à partir de molécules simples comme la pentamidine et le bérénil, de développer des molécules avec une reconnaissance et une affinité accrue pour l'ADN. Ces molécules capables de reconnaître des séquences riches en paires de bases AT et/ou ATGA possèdent des propriétés de pénétration et de distribution cellulaire étonnantes d'efficacité. Ces premières années d'étude se sont surtout orientées sur les fonctions amidines terminales et leur rôle majeur dans l'affinité et la sélectivité des composés pour l'ADN, mais aussi dans la distribution cellulaire de ces molécules. Notre étude menée sur les composés diphenyl-furanes et phényl-furane-benzimidazoles s'est dans un premier temps orientée sur les conséquences des modifications des corps polycycliques des molécules sur l'affinité et la sélectivité de reconnaissance des bases de l'ADN mais aussi sur les conséquences cellulaires de ces modifications. Nous avons alors pu compléter les règles de structure/activité et structure/affinité qui gouvernent la reconnaissance spécifique de ces composés. L'élargissement de l'étude à d'autres séries de composés DB qui possèdent des modifications plus importantes du corps polycyclique ont montré une capacité originale de reconnaissance de nouveaux sites spécifiques comme pour les composés DB1242 (GCTCG) et DB1255 (CCATG). Pour cette dernière molécule originale, l'étude actuelle son mode d'interaction avec l'ADN et de son affinité pour ces sites de reconnaissance CCATG devrait ouvrir de nouveaux horizons de synthèse et de développement.

Puisque les dérivés DB reconnaissent des séquences de plus en plus spécifiques, nous avons émis l'hypothèse qu'ils pourraient moduler la fixation de facteurs de transcription afin de réguler l'expression de certains gènes impliqués dans la cancérogenèse. Les premiers résultats présentés dans ce mémoire laissent à penser que cette hypothèse est pertinente. En effet, nous avons mis en évidence pour la première fois, par des techniques variées (membranes Transignal/Protein DNA array I, retard en gel et ELISA), l'inhibition sélective de facteurs de transcription par différents composés DB et par des modes d'action différents :

- Le DB293 inhibe efficacement la liaison à l'ADN des facteurs de transcription à POU domaine Brn-3 et Pit-1 par une reconnaissance du site ATGA situé sur la séquence consensus minimale de fixation de ces protéines. Cette reconnaissance a pour conséquence majeure de modifier la structure des sillons de l'ADN et d'empêcher la reconnaissance des facteurs de transcription.

Ainsi, depuis 15 ans et les premiers travaux menés sur les composés analogues du bérénil, nous avons montré que les composés DB représentaient aujourd'hui une classe de molécules très prometteuse dans le développement de ligands qui inhibent la liaison à l'ADN de facteurs de transcription régulant ainsi l'expression de certains gènes notamment ceux impliqués dans la cancérogenèse. Ces premiers résultats ouvrent des voies de recherches encourageantes pour le traitement en cancérologie.

MATERIEL ET METHODES

I. Clonage moléculaire

A. Vecteurs de clonage

Les vecteurs utilisés et leurs applications sont répertoriés dans le tableau 1. Les vecteurs d'expression nous ont été généreusement fournis par diverses équipes que je tiens à remercier. Les vecteurs pCMV-IRF-1, pCMV-Pit-1 et pcDNA3-HNF-4 nous ont été respectivement fournis par les Dr Angelina Battistini (Rome, Italie), Dr Clara Álvarez Villamarín (Santiago de Compostela, Espagne) et Dr Bernard Laine (INSERM U837, Lille, France). Les plasmides pRK5-Brn-3a, -3b et -3c ont été obtenus auprès du Dr Menping Xiang (Picataway, USA). Les vecteurs d'intérêts pCMV-HoxA9 et pSP64-PBX nous ont été fournis par le Dr Corey Largman (San Fransisco, USA). Les plasmides pcDNA3-HoxA9 et pcDNA3-Nup98/HoxA9 nous ont été fournis par le Dr Nabell R. Yaseen (Chicago, USA). Enfin, les plasmides pSG5-Erg, -Fli, -Ets-1, -Ets-2 et les plasmides pSV6-PEA-3, -ERM et -ETV-1 nous ont été fournis respectivement par les Dr Duterque Martine et Dr Jean-Luc Baert (IBL, Lille, France).

B. Culture bactérienne

1. Souche bactérienne

La souche d'*Escherichia coli* *XL1-Blue strain* possède le génotype suivant : *recA1*, *endA1*, *gyrA96*, *thi-1*, *hsdR17*, *supE44*, *relA1*, *lac* [F', *proAB*, *lac*^qZΔM15, Tn10, (Tet^r)]. Cette souche est communément utilisée pour amplifier des plasmides d'intérêts.

2. Culture

Cette souche bactérienne est cultivée en solution dans du milieu LB (Luria Bertoni : peptone 10 g/L ; extrait sec de levure 5g/L, NaCl 86 mM ; pH 7,5) ou étalée sur boîtes de pétri de LB gelosé (LB contenant de l'agar à 15g/L).

C. Transformation bactérienne

Les plasmides utilisés ainsi que leurs fonctions sont répertoriés tableau 4. L'amplification des plasmides s'effectue dans la bactérie XL1. Un clone bactérien est placé dans 10 mL de milieu

LB à 37°C sous agitation. Après 16 h, les bactéries sont diluées au 1/10 et placées 4h à 37°C sous agitation ; cette incubation permet aux bactéries de reprendre leur phase de croissance exponentielle. Les amplifications bactériennes en milieu liquide s'effectuent toujours à 37°C sous agitation alors que les milieux gélosés sont placés à 37°C sans agitation. Les bactéries sont alors centrifugées à 5000g pendant 8 minutes et le culot est repris dans 1mL d'une solution de TSS (10 mL LB, 1 g PEG, MgCl₂ 90 mM, DMSO 45 mM). 100µL de bactéries sont placées dans la glace en présence de 300 ng de plasmide à amplifier. Après 30 minutes d'incubation, 900 µL d'une solution de TSS glucose (TSS, glucose 200 mM) sont ajoutés. Après 1h d'incubation à 37°C, 60 µL de bactéries sont étalées sur une boîte de pétri contenant le LB gélosé additionné des antibiotiques de sélection tels que l'ampicilline (25 µg/mL), la kanamycine (125 µg/mL) ou la tétracycline (15µg/mL) (en fonction du plasmide).

plasmides	Type de vecteur	Description
pBS _{II} SK (Stratagen)	Amplification	Contient le fragment de 265 pb utilisé pour classer les composés DB en fonction de leur site de fixation par des tests à la DNase I.
pUC19 (Fermentas)	Amplification	Il permet après insertion des sites consensus des différents facteurs de transcription d'effectuer les tests à la DNaseI avec le composé DB293.
pCMV (Invitrogen)	Expression	Ce vecteur d'expression eucaryote contient le promoteur du CMV et est utilisé pour exprimer les protéines HoxA9, IRF-1 et Pit-1 en lysat de réticulocytes.
pcDNA3 (Invitrogen)	Expression	Ce vecteur d'expression eucaryote contient le promoteur du CMV et est utilisé pour exprimer les protéines HoxA9 et Nup98/HoxA9 en lysat de réticulocytes.
pSG5 (Stratagen)	Expression	Ce vecteur d'expression eucaryote contient le promoteur du virus SV40 et est utilisé pour exprimer les protéines Erg, Fli, Ets-1 et Ets-2 en lysat de réticulocytes.
pSP64 (Promega)	Expression	Ce vecteur d'expression eucaryote contient le promoteur du virus SP6 et est utilisé pour exprimer la protéine Pbx en lysat de réticulocytes.
pSV6 (Promega)	Expression	Ce vecteur d'expression contient le promoteur du virus SV40 et est utilisé pour exprimer les protéines ERM, PEA-3, ETV-1 en lysat de réticulocytes.
pGL3 basic (Promega)	Rapporteur	Ce vecteur contient un site de clonage multiple où a été inséré le site consensus de la protéine HoxA9 en amont du gène codant pour la protéine luciférase.
pGL3 promoteur (Promega)	Rapporteur	Ce vecteur contient en plus du vecteur précédant le promoteur minimal du virus SV40 situé entre le site de clonage multiple et le gène luciférase.

Tableau 4 : Descriptif des plasmides utilisés.

Le nom, le type de plasmide et le descriptif de chaque plasmide utilisé sont répertoriés dans ce tableau. Les plasmides dit « d'amplification » sont les vecteurs dans lesquels sont amplifiés les fragments d'ADN utilisés pour effectuer les tests d'empreinte à la DNase I. Les plasmides d'expression sont utilisés pour exprimer les protéines d'intérêt en lysat de réticulocytes ou en cellules eucaryotes. Les plasmides « rapporteurs » sont utilisés pour effectuer des tests d'inhibition d'activité de la luciférase par les composés DB sur certains sites consensus.

D. Amplification des plasmides

Après incubation 16 h à 37°C dans une étuve, les clones bactériens transformés avec un plasmide d'intérêt sur boîte de pétri, sont cultivés dans 10 mL d'une solution de LB additionnée d'antibiotique de sélection à la concentration appropriée. Après incubation à 37°C pendant 8 h, 1 mL de bactéries est prélevé et placé dans 250 mL de milieu LB + antibiotique. Cette solution bactérienne est incubée pendant 16 h à 37°C.

Les plasmides ainsi amplifiés sont récupérés et purifiés à l'aide du kit Qiagen (Maxi prep). Brièvement, les bactéries sont récupérées par centrifugation à 5000 g pendant 8 min puis remises en suspension dans 10 mL de tampon P1. Les bactéries sont ensuite lysées dans 10 mL de tampon P2 pendant 5 min, et la réaction est stoppée par 10 mL de tampon P3. Les lysats sont filtrés à travers une seringue munie d'un filtre pour éliminer les débris bactériens. La solution ainsi obtenue est passée sur une colonne où l'ADN plasmidique est retenu. Ces ADN sont lavés 2 fois par 30 mL de tampon QBT. Les plasmides sont ensuite élués dans 15 mL de tampon QF puis précipités dans 15 mL d'isopropanol puis dans 10 mL d'éthanol à 70 %. Les culots d'ADN plasmidique sont séchés et repris dans l'eau avant d'être dosés au spectrophotomètre.

E. Construction plasmidique

Les oligonucléotides et les enzymes utilisés pour effectuer les différents clonages sont référencés Tableau 5. Deux µg de plasmide d'intérêt sont digérés par 2 unités d'enzyme de restriction correspondantes (Tableau 5). Le vecteur clivé est séparé de l'insert à éliminer par migration sur gel d'agarose 1 % pendant 1 h à 120 V. La bande d'intérêt est excisée du gel sous UV et l'ADN plasmidique est récupéré avec le kit NucleoSpin (Macherey-Nagel). Brièvement, l'agarose est dissous à 50°C pendant 5 min dans 300 µL de tampon NT. La solution contenant l'ADN est placée sur une colonne munie d'un filtre capable de retenir le vecteur, centrifugée 1 min à 11000 g puis lavée deux fois par 600 µL d'une solution NT3. Pour terminer, le vecteur est élué par 30 µL d'eau et récupéré dans un tube 1,5 mL par centrifugation pendant 2 min à 12000 g. La quantité d'insert (oligonucléotide double brin) à utiliser pour la ligation est déterminée par l'équation suivante :

$$\text{ng d'insert} = \frac{\text{ng de vecteur} \times \text{taille de l'insert (Kb)}}{\text{taille du vecteur (Kb)}} \times \text{ration molaire} \frac{\text{insert}}{\text{vecteur}}$$

Le ratio molaire classiquement utilisés pour une efficacité maximale est de 3.

Plasmides	Sites d'insertion	Séquence de l'insert	Sites de digestion et marquage	Sites consensus
pUC19	<i>Bam</i> H I	5'- gatcagctcgtctgcgcg TTAATGAG ctgtgcgagtcgcc	<i>Eco</i> R I/ <i>Pst</i> I	Brn-3
pUC19	<i>Bam</i> H I	5'- gatcagtgcttctctg AATATGAATAA gaaataagtcgcc	<i>Eco</i> R I/ <i>Pst</i> I	Pit-1
pUC19	<i>Bam</i> H I	5'- gatcagtgcttctctg AATATGAATAA gaagtcagtcgcc	<i>Eco</i> R I/ <i>Pst</i> I	Pit-1 WT
pUC19	<i>Bam</i> H I	5'- gatcagtgcttctctg AATACGAATAA gaagtcagtcgcc	<i>Eco</i> R I/ <i>Pst</i> I	Pit-1 m1
pUC19	<i>Bam</i> H I	5'- gatcagtgcttctctg AATATGAACGA gaagtcagtcgcc	<i>Eco</i> R I/ <i>Pst</i> I	Pit-1 m2
pUC19	<i>Bam</i> H I	5'- gatcagtgcttctctg AATACGAACGA gaagtcagtcgcc	<i>Eco</i> R I/ <i>Pst</i> I	Pit-1 m3
pUC19	<i>Bam</i> H I	5'- gatcagtgcttctctg AGTATGAACGA gaagtcagtcgcc	<i>Eco</i> R I/ <i>Pst</i> I	Pit-1 m4
pUC19	<i>Bam</i> H I	5'- gatcagctcgtggaagc GAAAATGAAATT gactagtcgcc	<i>Eco</i> R I/ <i>Pst</i> I	IRF-1

Tableau 5 : Constructions plasmidiques.

La colonne **plasmide** correspond au vecteur originel utilisé pour effectuer la construction. L'insert est intégré dans le plasmide au niveau du site multiple de clonage. Les enzymes de restriction impliqués dans la construction plasmidique sont représentés dans la colonne **Site d'insertion**. La séquence des inserts est indiquée dans la colonne **Séquence de l'insert**. En bleu majuscule sont répertoriés les sites d'intérêt et en rouge les points de mutations éventuels. La colonne Sites de digestion et de Marquage correspond aux sites utilisés pour digérer le plasmide construit afin de récupérer un fragment d'intérêt pour réaliser des tests d'empreinte à la DNase I. Les sites présentés en rose correspondent aux sites où sont réalisés les marquages en 3'. Le sigle Ø indique que les plasmides n'ont pas été utilisés pour effectuer des tests d'empreinte à la DNaseI. La colonne **Sites consensus** représente les différents sites d'intérêts intégrés dans les plasmides.

Les ligations sont effectuées en utilisant 200 ng de vecteur. Les inserts utilisés sont répertoriés dans le Tableau 5. La quantité d'insert à utiliser est déterminée selon la formule ci-dessus et est dépendante de la taille du vecteur et de l'insert. Pour la ligation, le mélange vecteur / insert est incubé avec 4 unités de T4 DNA ligase (New England BioLabs) pendant 16 h à 16°C dans un tampon 50 mM Tris-NaCl, pH 7,5 ; 10 mM MgCl₂ ; 1mM ATP ; 10 mM DTT. L'amplification des plasmides d'intérêts s'effectue comme décrit précédemment.

II. Etude moléculaire

A. Etude de l'interaction ADN-ligand

1. Marquage radioactif sur l'extrémité 3' de l'ADN

a) Marquage à l'AMV-RT

Le fragment d'ADN de 265 pb provenant du plasmide pBS est obtenu par digestion enzymatique *Eco*R I-*Pvu* II. Le marquage est effectué par incorporation d' α -[³²P]-dATP (1,85

MBq; GE Healthcare) au niveau du site EcoR I et en 3' du fragment d'ADN par 40 unités d'enzyme AMV-RT (Avian Myeloblastosis Virus-Reverse Transcriptase, Finzyme, 20U/ μ L). Après 2 h d'incubation à 37°C, 12 μ L d'une solution de dATP non marqué à 100 mM sont ajoutés. Les produits de digestion sont déposés sur gel de polyacrylamide à 6 % non dénaturant (rapport acrylamide:bisacrylamide 19:1) et séparés par une migration de 2 h à 170 V dans une solution de TBE 1X (89 mM Tris, 89 mM Borate, 20 mM EDTA, pH 8.3). Après autoradiographie du gel, les bandes d'intérêt sont excisées, broyées puis éluées dans un tampon TE-NaCl (Tris 10 mM, EDTA 1 mM, NaCl 100 mM, pH 8) pendant une nuit à 4°C. L'ADN radiomarqué en solution est alors séparé du gel d'acrylamide par centrifugation au travers d'un filtre millipore de 0,45 μ m puis précipité dans l'éthanol absolu par centrifugation 30 min à 13800 g. L'éthanol est éliminé et le culot d'ADN est remis en suspension dans 50 μ L d'eau MQ.

b) Marquage à la Klenow

Les fragments d'ADN marqués à l'enzyme Klenow proviennent des différents clonages effectués. Ces plasmides sont référencés Tableau 5. Tous les fragments proviennent de clonages effectués dans le plasmide pUC19 à l'exception d'un seul, nommé pGL3 promoteur HoxA9, qui provient du plasmide pGL3 promoteur. Les sites de digestion des plasmides nécessaires pour le marquage sont répertoriés Tableau 5. Les marquages sont effectués par incorporation sur l'extrémité 3' d' α -[³²P]-dATP (1,85 MBq; GE Healthcare) par 10 unités d'enzyme Klenow (BioLabs) pendant 30 min à 37°C. Pour les fragments issus du plasmide pUC19, après 30 min d'incubation sont ajoutés 12 μ L d'une solution de dATP à 100 mM. Pour le marquage du fragment pGL3-HoxA9, 10 mM de dGTP sont ajoutés pour permettre l'incorporation du dATP marqué en position 2 sur l'extrémité protubérante du site Bgl II.

Les produits de digestion sont déposés sur gel de polyacrylamide à 6 % non dénaturant (rapport acrylamide:bisacrylamide 19:1) et séparés par une migration de 2 h à 170 V dans une solution de TBE 1X. Après autoradiographie du gel, les bandes d'intérêts sont excisées et l'ADN récupéré comme précédemment décrit dans le paragraphe II.A.1.a.

2. Marquage radioactif sur l'extrémité 5' de l'ADN

Les différents oligonucléotides utilisés sont répertoriés dans le Tableau 6. Quatre cents ng d'oligonucléotides simples brins contenant les séquences consensus des facteurs de transcription d'intérêt sont marqués par l'incorporation sur l'extrémité 5' de γ -[³²P]-ATP

(1,85 MBq; GE Healthcare). Le marquage s'effectue par l'ajout de 20 U d'enzyme T4 polynucléotide kinase (BioLabs). Après 1 h 30 d'incubation à 37°C, l'enzyme est inactivé pendant 5 min à 90°C et 600 ng d'oligonucléotides complémentaires sont ajoutés. La réaction d'hybridation des deux brins d'oligonucléotides s'effectue en plaçant la solution 5 min à 90°C puis en laissant redescendre lentement la température à 20°C (environ 2 h). L'oligonucléotide double brin est ensuite purifié par migration sur gel de polyacrylamide et récupéré comme précédemment (les séquences des oligonucléotides utilisés sont répertoriées tableau 3).

Noms	Séquence
Brn-3	5'- GCGCGT <u>T</u> AATGAGCTGTG
Pit-1	5'- GTCTTCCTGA <u>A</u> ATATGAATAAGAAATAA
IRF-1	5'- GGAAGCGAA <u>A</u> ATGAAATTGACT
HNF-4	5'- TCAGCTTGTACTTTGGTACA <u>A</u> CTA
Pit-1 WT	5'- GATCAGTGTCTTCCTGAAT <u>A</u> TGAATAAGAAGTCAGTCGCC
Pit-1 M1	5'- GATCAGTGTCTTCCTGAAT <u>A</u> CGAATAAGAAGTCAGTCGCC
Pit-1 M2	5'- GATCAGTGTCTTCCTGAAT <u>A</u> TGAACGAGAAGTCAGTCGCC
Pit-1 M3	5'- GATCAGTGTCTTCCTGAAT <u>A</u> CGAACGAGAAGTCAGTCGCC
Pit-1 M4	5'- GATCAGTGTCTTCCTGAGT <u>A</u> TGAACGAGAAGTCAGTCGCC
Pit-1 M5	5'- GATCAGTGTCTTCCTGAGT <u>A</u> TGAATAAGAAGTCAGTCGCC
Pit-1 M6	5'- GATCAGTGTCTTCCTGAGT <u>A</u> CGAATAAGAAGTCAGTCGCC
HoxA9	5'-CTGCGATGATTTACGACCGC

Tableau 6 : Séquence des oligonucléotides marqués à l'extrémité 5'.

Oligonucléotides simple brin utilisés pour le marquage en 5'. Après le marquage, 600 ng d'oligonucléotide complémentaire sont ajoutés et hybridés. Les séquences d'intérêt sont soulignées en noir. Ne sont représentés que les oligonucléotides marqués.

3. Empreinte à la DNase I

Cette technique permet de localiser les sites de fixation à l'ADN d'un produit par digestion ménagée à la DNase I des séquences d'ADN non protégées par les produits (Graves *et al.*, 1986 ; Bailly *et al.*, 2005). Dix ng d'ADN marqué en 3' sont incubés 15 min à température ambiante en présence de ligand à la concentration désirée dans un tampon constitué de 10 mM Tris pH 7.0 et 10 mM NaCl. La digestion est initiée par l'addition de 2 µL d'une solution de DNase I diluée à 0,01 unité/mL. Après 3 minutes, la réaction est stoppée par congélation à -80°C dans la carboglace. Les échantillons sont alors lyophilisés sous vide. Les fragments d'ADN digérés sont redissous dans 5 µL de tampon de charge dénaturant (formamide 80%,

xylène cyanol 0.1%, bleu de bromophénol 0.1%), dénaturés par chauffage 5 min à 90°C, suivi d'un brusque refroidissement dans la glace avant d'être déposés sur gel de polyacrylamide dénaturant (Urée 8M) à 8%. Un contrôle de migration (G-track) est lui aussi déposé, il permet de montrer la localisation des guanines dans le brin d'ADN testé. La migration s'effectue pendant 1 h 45 à 65 W dans un tampon TBE 1X. Le gel est fixé dans une solution d'acide acétique à 10 % puis séché sous vide et révélé au PhosphoImager. Les analyses densitométriques sont réalisées par intégration des densités des différents fragments d'ADN issus de la digestion par utilisation du logiciel ImageQuant version 3.3. Chaque bande est identifiée par comparaison avec sa position relative au standard (G-track).

Le G-track permet de visualiser la localisation des guanines dans le brin d'ADN testé. Les guanines sont révélées par digestion chimique à la pipéridine. 200 ng d'ADN marqués sont incubés 2 min à température ambiante avec 200 mM de DMS. Le DMS transfère un groupement méthyle sur l'azote en position N7 des guanines de l'ADN. La réaction est stoppée par l'addition de 200 µL d'éthanol qui permet de précipiter l'ADN. Après centrifugation à 13800 g, l'éthanol est éliminé et l'ADN méthylé est séché 15 min à l'air libre. Le culot est repris dans 25 µL de pipéridine 100 mM et incubé 12 min à 90°C afin de permettre la digestion de l'ADN au niveau des guanines méthylées. Le mélange est séché sous vide, lavé avec 100 µL d'eau et séché sous vide une nouvelle fois avant d'être repris dans un tampon de charge dénaturant.

4. Test de dénaturation thermique de l'ADN

Cette technique consiste à enregistrer des variations des propriétés d'absorbance reflétant un changement conformationnel de la molécule d'ADN (effet hyperchrome). Pour cela, la densité optique (DO) à 260 nm d'une solution d'ADN soumise à une augmentation de température progressive de 20 à 100°C (augmentation de 1°C/min) est mesurée toutes les 90 secondes. Les courbes de dénaturation thermique de l'ADN sont enregistrées avec un spectrophotomètre Uvikon 943 couplé à un régulateur thermique cryostat Neslab RTE 111. La température de fusion d'un fragment d'ADN, ou T_m (melting temperature), correspond à la température où 50 % de l'ADN se retrouve sous forme dénaturée et 50 % sous forme double

brin. La température est mesurée dans une cuve de référence à l'aide d'une sonde de température. Pour chaque série de mesures, 10 échantillons et 2 contrôles sont déposés dans des cuves en quartz (10 mm) et placés dans une cellule chauffée par circulation d'eau contrôlée par le thermostat. Les échantillons d'ADN sont placés dans un tampon BPE (6 mM Na₂HPO₄, 2 mM NaH₂PO₄, 1 mM EDTA, pH 7). Pour cette étude, différentes séquences d'ADN sont utilisées (Tableau 7). En ce qui concerne les oligonucléotides, la concentration en ADN est de 0,5 μM d'oligonucléotide alors que pour le poly(dAdT)₂ la concentration en ADN est de 20 μM. Les ligands sont toujours ajoutés à la concentration de 1 μM.

<i>Noms</i>	<i>Séquences</i>
<i>AT riche</i>	5'- GTGAGGGTTAATTCGAGCTTGG
<i>AT riche m</i>	5'- GTGAGGGTTGGTTTCGAGCTTGG
<i>ATGA</i>	5' - ACAGCTATGACCATGATTACGCC
<i>HoxA9</i>	5'- CTGCGATGATTACGACCGC
<i>HoxA9 AT m</i>	5'- CTGCGATGAGTCACGACCGC
<i>HoxA9 ATGA m</i>	5'- CTGCGTAGCTTACGACCGC
<i>HoxA9 m</i>	5'- CTGCGTAGCGTCACGACCGC

Tableau 7 : Séquences des oligonucléotides utilisés pour réaliser les mesures de dénaturation thermique de l'ADN.

Un des deux oligonucléotides complémentaire est représenté. Les séquences représentées en bleu correspondent aux sites d'intérêts, les bases représentées en rouge représentent les sites de mutations.

5. Résonance plasmonique de surface

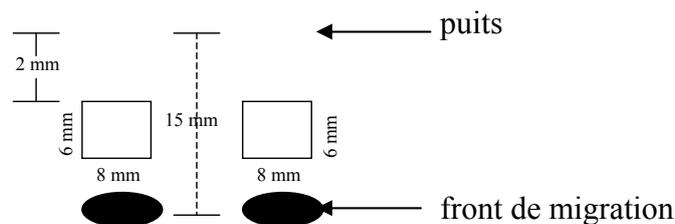
Ces expériences sont réalisées par l'équipe du Pr David Wilson (Atlanta, USA). Pour effectuer ces tests les oligonucléotides suivant (en forme d'hairpin, biotinylés sur l'extrémité 5'-) ont été utilisés : d(Biotine-CAATATGAATAAGAAGTCTCCTTCTTATTCATATTG) (Pit-1), d(Biotine-CGCGTTAATGAGCTGGTCTCCCAGCTCATTAACGCG) (Brn-3), d(Biotine-CGAAAATGAAATTGACTCTCGTCAATTTTCATTTTCG) (IRF-1), ainsi que deux séquences Pit-1 mutées sur le site ATGA et sur le site riche en paires de bases AT situé en position 3' du site ATGA, d(Biotine-CAATACGAATAAGAAGTCTCCTTCTTATTCGTATTG) (Pit-1 M1) et d(Biotine-CAATATGAACGAGAAGT

CTCCTTCTCGTTCATATTG) (Pit-1 M2). Les détails techniques sont détaillés dans Peixoto et al., 2008.

B. Modulation de la fixation des facteurs de transcription à l'ADN

1. TranSignal™/protein/DNA arrays (Panomics)

Cette méthode, basée sur le principe des macroarrays, permet d'effectuer un criblage de 54 facteurs de transcription en une seule manipulation dans un mode compétitif (Figure 74). Des oligonucléotides biotinylés, spécifiquement reconnus par des facteurs de transcription sont incubés à 37°C avec 1 ou 5 µM de DB293. 15 µg d'extraits nucléaires, contenant le mélange de facteurs de transcription, sont ajoutés. Les complexes (oligonucléotides - facteurs de transcription) formés sont isolés des oligonucléotides libres par migration sur gel d'agarose à 2 % dans du TBE 1X pendant 15 min à 120 V. Après la migration, le gel est excisé selon le schéma suivant recommandé :



Les oligonucléotides complexés aux facteurs de transcription sont extraits avec le kit *Get pure DNA-kit Agarose* (Dojundo). L'agarose est tout d'abord dissout dans 300 µL de tampon *gel lysis buffer* à 60°C pendant 10 min puis précipité dans 300 µL de tampon *precipitation solution* et éliminé après centrifugation à 12000 rpm pendant 5 min. Les oligonucléotides contenus dans le surnageant sont ensuite purifiés ; pour cela 2 µl de la solution de *co-precipitation* et 800 µL d'éthanol froid sont ajoutés et vortexés avant d'être centrifugés à 12000 rpm pendant 3 min. Un lavage par une solution d'éthanol -20°C à 70 % suivi d'une centrifugation à 12000 rpm sont effectués. Les oligonucléotides sont repris dans 100 µL d'eau, avant d'être dénaturés et hybridés aux membranes TranSignal dans le tampon d'hybridation fourni, à une température de 42°C pendant 16 h. Sur ces membranes sont greffés chacun des oligonucléotides simple brin correspondant aux séquences complémentaires de 54 facteurs de transcription connus (kit TranSignal™/protein/DNA arrays I). Les membranes sont ensuite lavées et une solution de HRP (péroxydase) couplée à la streptavidine est ajoutée. Pour terminer, les membranes sont incubées dans une solution contenant le substrat de la HRP et sont ensuite révélées par autoradiographie (réaction de

chimioluminescence). L'analyse se fait à l'aide du logiciel ImageQuant qui détermine l'intensité de chaque spot. Le rapport d'intensité est calculé en faisant le ratio intensité du contrôle sans drogue sur l'intensité du test avec drogue. Comme établi par le fournisseur, la modulation de l'interaction ADN/facteur de transcription est significative lorsque le ratio est supérieur à 1,5.

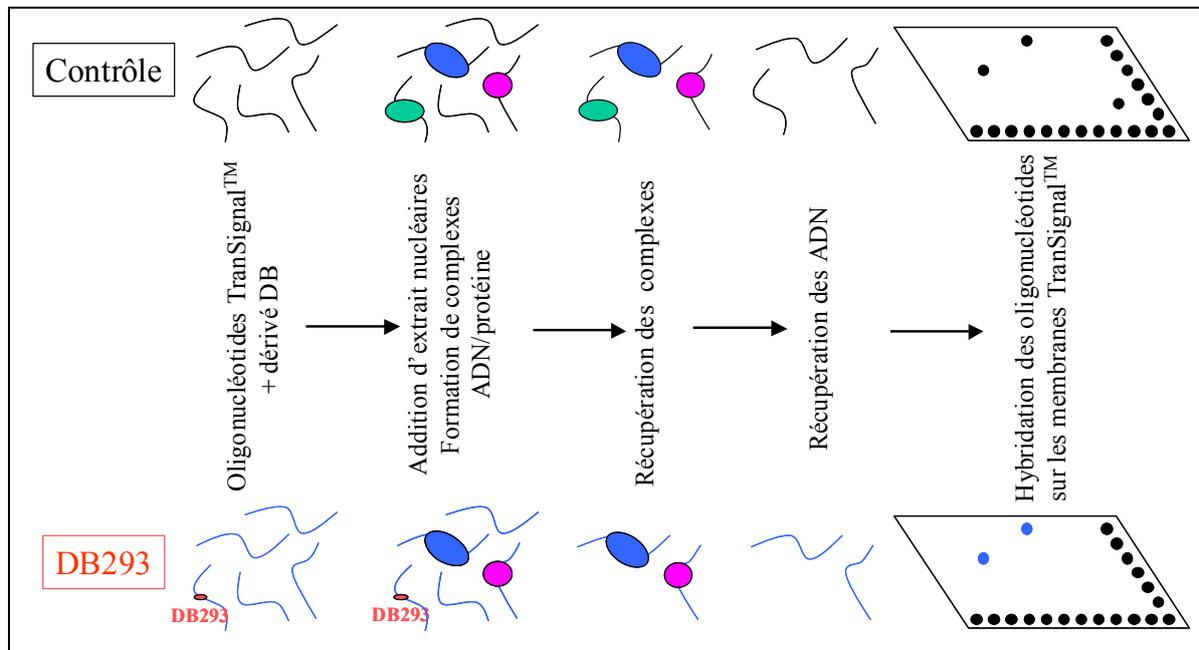


Figure 74 : TransSignal™/protein/DNA arrays (Panomics).

Les oligonucléotides biotinylés sont incubés ou non avec le DB293 puis des extraits nucléaires de cellules HT-29 sont ajoutés. Les complexes facteurs de transcription/ADN ainsi formés sont isolés par gel d'agarose. Les oligonucléotides engagés dans ces complexes sont récupérés et hybridés sur les membranes Transignal™. Sur ces membranes sont immobilisés les oligonucléotides complémentaires à ceux utilisés précédemment. Des billes de streptavidine couplées à la protéine HRP sont ajoutées. Une analyse densitométrique permet de déterminer les rapports d'inhibition ou d'activation de la formation des complexes ADN/facteurs de transcription par le DB293.

2. Expression des protéines en lysat de réticulocytes

L'expression de protéines en lysat de réticulocytes s'effectue par le kit T7 ou SP6 TNT® Coupled Reticulocyte Lysate System (Promega). 25 µL de lysat sont mélangés à 2 µL de tampon TNT® reaction Buffer, 1µL de TNT® T7 RNA polymérase ou SP6 TNT® T7 RNA polymérase (en fonction du plasmide d'expression), 1 µL d'acides aminés à 1 mM, 2 µL d'inhibiteur de ribonucléase et 1 µg de plasmide d'intérêt (vecteurs exprimant Brn-3-a, -b, -c Pit-1, IRF-1, HNF-4 ou HoxA9). Le mélange est complété avec de l'eau sans nucléase pour un volume final de 50 µL et incubé à 30°C pendant 1 h 30. Afin de vérifier l'expression des protéines d'intérêt, l'acide aminé méthionine est remplacé par la [³⁵S] méthionine (1,1 MBq ; GE Healthcare) dans une manipulation contrôle. Les lysats contenant les protéines marquées

sont alors déposés sur un gel de polyacrylamide (rapport acrylamide/bisacrylamide 37,5/1) à 10 % (gel de concentration : 0,125 M Tris, pH 6,8 ; 0,1% SDS ; gel de séparation : 0,375 M Tris ; pH 8,8 ; 0,1 % SDS) et la migration s'effectue contre un tampon Tris 25 mM ; 192 mM glycine ; SDS 0,1 % pendant 4 h à 30 mA. Le gel est alors séché et analysé au phosphoImager.

3. Retard en gel

Cette technique permet de visualiser l'interaction entre une protéine et une sonde d'acide nucléique marquée, en se basant sur la différence de mobilité sur gel de polyacrylamide d'un complexe protéine-ADN par rapport à celle de la sonde ADN libre. 2 µL de composé DB à la concentration requise sont incubés avec 6 ng d'oligonucléotides double brin marqués en 5' et complétés à 10 µL de tampon d'interaction (15 mM Tris-HCl, pH 7.5, 50 mM KCl, 1.5 mM EDTA, 10% glycerol, 1.5 mM DTT, 10 µg BSA). Afin de vérifier la spécificité des bandes obtenues, 2 stratégies peuvent être adoptées : soit ajouter au mélange un excès de sonde froide (50 fois concentré, soit 300 ng) spécifique ou non spécifique, soit ajouter 1 µL d'anticorps spécifique ou non spécifique de la protéine d'intérêt pour effectuer un super-retard ou *supershift*. Après 20 min d'incubation à température ambiante, 5 µg d'extraits nucléaires de cellules HT-29 ou 1 à 2 µl de protéines exprimées en lysat de réticulocytes sont ajoutés et incubés 30 min à 4°C. Les échantillons sont alors déposés sur un gel de polyacrylamide non dénaturant à 6 % (rapport acrylamide:bisacrylamide 37,5:1, TBE 0,5X). L'électrophorèse s'effectue à 300 V pendant 3 h contre un tampon TBE 0,5X. Le gel est ensuite séché et analysé au phosphoImager.

5. Western Blotting

Les lysats protéiques sont repris dans un tampon de charge 2X (50 mM Tris HCl, pH 6,8 ; 15 % Sucrose, 2 mM EDTA , pH 7,5 ; 3% SDS, bleu de bromophénol 0,01%) avant d'être déposés à la concentration désirée sur un gel d'acrylamide 10% (rapport acrylamide/bisacrylamide 37,5/1 ; gel de concentration : 0,125 M Tris, pH 6,8 ; 0,1% SDS ; gel de séparation : 0,375 M Tris ; pH 8,8 ; 0,1% SDS) et la migration s'effectue contre un tampon Tris 25 mM ; 192 mM glycine ; SDS 0,1% pendant 1h30 à 100 mA. Après migration, le gel est transféré sur une membrane PVDF (Amersham biosciences) pendant 90 min à 90 V contre un tampon de transfert (192 mM glycine ; 25 mM Tris, pH 8,3 ; SDS 0,1% ; méthanol 20%). La membrane est alors incubée pendant 1h avec une solution de TBS_T (Tween 20) 0,5

% lait 1 % afin de limiter les interactions non spécifiques de l'anticorps sur la membrane. L'anticorps primaire est alors ajouté à la concentration désirée (dilué au 1/500^{ème} pour HoxA9) et incubé pendant 1 h à température ambiante. Trois lavages successifs au TBS_T sont effectués, puis l'anticorps secondaire (dirigé contre les anticorps produits chez la chèvre pour l'anticorps anti-HoxA9) couplé à la protéine HRP dilué au 1/20000 est ajouté et incubé 1h à température ambiante. Trois lavages successifs au TBS_T sont à nouveau réalisés puis un lavage au TBS. Les protéines d'intérêts fixées sur les membranes sont révélées par chimioluminescence avec le kit ECL (GE Healthcare) sur des films photographiques (Kodak, GE Healthcare). Les temps d'exposition sont fonction des anticorps utilisés.

III. Etudes en cellules eucaryotes

A. Culture cellulaire

1. Cellules adhérentes

a) Cellules HT-29

Les cellules d'adénocarcinome de colon HT-29 (ATCC : HTB-38) sont cultivées dans un milieu DMEM-glutaMAX auquel sont classiquement ajoutés 10% de SVF (sérum de veau fœtal) et 1% de mélange pénicilline-streptomycine (pénicilline G : 10000 unités/mL, streptomycine 10 µg/mL). Les cellules HT-29, comme toutes les cellules utilisées, sont maintenues pour cette étude en culture à 37°C avec 5% de CO₂ dans une étuve à atmosphère humide (90% d'humidité). Le repiquage des cellules s'effectue dans une flasque T150 (150 cm²) dans laquelle on ajoute 25 mL de cellules diluées dans du milieu de culture à la concentration de 200 000 cellules/mL.

B. Microscopie de fluorescence

1. Distribution cellulaire

Les cellules HT-29 sontensemencées dans 1 mL de milieu de culture sur des lames compartimentées en 8 puits à la densité de 20 000 cellules/mL. Après 24 h d'incubation, 5 µM de composés fluorescents sont ajoutés.

Après 18 heures d'incubation, les cellules sont lavées par 1 mL de PBS (NaCl 68 mM, Na₂HPO₄ 85 mM, NaH₂PO₄ 17 mM ; pH 7,4) puis fixées par une solution de PFA 2%

pendant 15 min à 4°C. Entre chaque étape, 3 lavages de 5 min par 1 mL de PBS sont effectués. Elles sont ensuite incubées en présence de 200 µL d'une solution de DiOC₆ (3,3'-dihexyloxycarbocyanine iodide) à 75 nM (utilisé comme marqueur cytoplasmique) pendant 15 min à 37°C puis avec 200 µL d'une solution d'IP (Iodure de propidium) à 1 µg/mL (utilisé comme marqueur nucléaire). Enfin, la préparation est scellée entre lame et lamelle dans 8 µL d'une solution de « Vectashield » (VECTOR). Les lames sont observées au microscope droit à fluorescence LEICA[®] avec différents lasers permettant de visualiser la localisation des différents produits utilisés. Les images sont captées par le logiciel Qfluor.

2. Fixation différentielle

Les cellules HT-29 sont placées dans des lames compartimentées en 8 puits à la densité de 20000 cellules/mL comme précédemment. Après 18 h de traitement avec les composés à étudier, les cellules sont lavées par 1 mL de PBS froid (5 min), puis fixées avec différents agents fixant : i) PFA en concentrations croissantes de 1 à 4% en solution dans du PBS pendant 20 min à 4°C ; ii) NP40 en concentrations croissantes dans du PBS + PFA à 2% pendant 20 min à 4°C (NP40 de 0, 1 à 2 %) ; iii) méthanol en concentrations croissantes de 25 % à 100 % dans du PBS pendant 10 min à -20°C. Après la fixation, les cellules sont lavées trois fois au PBS et analysées comme précédemment.

C. Tests de cytotoxicité

Les tests de cytotoxicité au réactif MTS permettent de déterminer la concentration pour laquelle 50 % des cellules sont tuées par les composés (IC₅₀).

80 µl de cellules à une concentration dépendante du type cellulaire (tableau 8) sont placées dans des plaques 96 puits. Après 2 h (pour des cellules en suspension) ou 24 h (pour des cellules adhérentes) de pré-culture, les cellules sont incubées en présence de concentrations croissantes en composé DB allant de 0,001 µM à 50 µM. Après 72 h de traitement, le réactif MTS (CellTiter 96[®] Aqueous) est ajouté à raison de 20 µL par puits et incubé dans l'étuve de 1h à 2h en fonction du type cellulaire (voir tableau 8). Le réactif MTS contient un dérivé tétrazolium (3-(4,5-diméthylthiazol-2-yl)-2-(4-sulfophényl)-2H-tétrazolium) réduit en formazan par l'enzyme mitochondriale NADP déshydrogénase. Cette réduction entraîne un changement de couleur de la molécule qui passe du jaune au pourpre. Le pourcentage de cellules vivantes est déterminé par lecture de l'absorbance à 492 nm des différents échantillons traités ou non avec le composé d'intérêt. L'absorbance à cette longueur d'onde est maximale lorsque le dérivé tétrazolium est réduit (couleur pourpre).

E. Préparation d'extraits nucléaires

Les extraits nucléaires sont préparés à l'aide du kit Nuclear extraction kit (Panomics). Les cellules HT-29 sont mises en culture dans 20 boîtes de pétri 100-mm dans un milieu DMEM auquel sont ajoutés 10% de SVF et 1% de mélange pénicilline-streptomycine (pénicilline G : 10 000 unités/mL, streptomycine 10 µg/mL) jusqu'à obtention d'un tapis cellulaire à 80% de confluence (environ 1×10^6 cellules par mL). Les cellules sont alors lavées deux fois dans 10 mL de PBS. 1 mL de Solution A (1 mL tampon A, 10 µL DTT, 10 µL d'inhibiteur de protéases, 10 µL d'inhibiteur de phosphatase I, 10 µL d'inhibiteur de phosphatase II) sont ajoutés et les cellules sont placées sous agitation sur un plateau rotatif à 200 rpm. Après 10 min d'agitation, les cellules sont raclées, récupérées dans des tubes de 2 mL et vortexées pendant 10 secondes dans le but de rompre les membranes plasmiques. Après centrifugation à 14 000 g pendant 3 min à 4°C et élimination du surnageant contenant la fraction cytoplasmique, les noyaux sont repris dans 150 µL de Solution B (1 mL de tampon B, 10 µL DTT, 10 µL d'inhibiteur de protéases, 10 µL d'inhibiteur de phosphatase I, 10 µL d'inhibiteur de phosphatase II). L'incubation des noyaux pendant 2 h en agitation à 150 rpm dans cette solution très saline permet l'exclusion des protéines nucléaires hors du noyau. Le surnageant contenant les protéines est alors récupéré par centrifugation à 14 000 g pendant 5 min à 4°C et dialysé à 4°C dans 500 mL de Solution C (20 mM Tris pH 7.9, 75 mM KCl, 0.2 mM EDTA, 20 % glycérol, 1 mM DTT et 0.5 mM PMSF) pendant 16 h. Les extraits nucléaires sont récupérés et dosés selon la méthode de Bradford avant d'être aliquotés et stockés à -80°C.

F. Extraction des protéines totales

Les extraits protéiques totaux sont utilisés pour réaliser les expériences de western blot. Pour réaliser ces extractions 5 millions de cellules sont utilisées. Après comptage, les cellules sont lavées par 5 mL de PBS puis reprises dans 500 µL de tampon de lyse (SDS 1% ; 1mM Na-vanadate; 10 mM Tris, pH 7,4 ; 0,1 mM PMSF, pepstatine 2,5 µg/mL ; aprotinine 10 µg/mL ; leupeptine 5 µg/mL) à 4°C. Après 10 min les échantillons sont passés 10 fois à travers une seringue munie d'une aiguille 21G pour obtenir un lysat cellulaire non visqueux. Les protéines sont ensuite dosées selon la méthode de Bradford.

REFERENCES BIBLIOGRAPHIQUES

- Adams, C.C. and Workman, J.L. (1993) Nucleosome displacement in transcription. *Cell*, **72**, 305-308.
- Altschuler, D.L. and Ribeiro-Neto, F. (1998) Mitogenic and oncogenic properties of the small G protein Rap1b. *Proc. Natl. Acad. Sci. U S A.*, **95**, 7475-7479.
- Andersen, B. and Rosenfeld, M.G. (1991) POU domain factors in the neuroendocrine system: lessons from developmental biology provide insights into human disease. *Endocr. Rev.*, **22**, 2-35.
- Andersen, B. and Rosenfeld, M.G. (1994) Pit-1 determines cell types during development of the anterior pituitary gland. A model for transcriptional regulation of cell phenotypes in mammalian organogenesis. *J. Biol. Chem.*, **269**, 29335-29338.
- Andersen, B., Schonemann, M.D., Flynn, S.E., Pearse, R.V. 2nd, Singh, H. and Rosenfeld, M.G. (1993) Skn-1a and Skn-1i: two functionally distinct Oct-2-related factors expressed in epidermis. *Science*, **260**, 78-82. Erratum (1993), *Science*, **262**, 1499.
- Andersen, B., Schonemann, M.D., Pearse, R.V. 2nd, Jenne, K., Sugarman, J. and Rosenfeld, M.G. (1993) Brn-5 is a divergent POU domain factor highly expressed in layer IV of the neocortex. *J. Biol. Chem.*, **268**, 23390-23398.
- Arai, Y., Kyo, T., Miwa, H., Arai, K., Kamada, N., Kita, K. and Ohki, M. (2000) Heterogenous fusion transcripts involving the NUP98 gene and HOXD13 gene activation in a case of acute myeloid leukemia with the t(2;11)(q31;p15) translocation. *Leukemia*, **14**, 1621-1629.
- Armstrong, S.C., Staunton, J.E., Silverman, L.B., Pieters, R., den Boer, M.L., Minden, M.D., Sallan, S.E., Lander, E.S., Golub, T.R. and Korsmeyer, S.J. (2001) MLL translocations specify a distinct gene expression profile that distinguishes a unique leukemia. *Nat. genet.*, **30**, 41-47.
- Asa, S.L., Puy, L.A., Lew, A.M., Sundmark, V.C. and Elsholtz, H.P. (1993) Cell type-specific expression of the pituitary transcription activator pit-1 in the human pituitary and pituitary adenomas. *J. Clin. Endocrinol. Metab.*, **77**, 1275-1280.
- Ayadi, A., Zheng, H., Sobieszczuk, P., Buchwalter, G., Moerman, P., Alitalo, K. and Wasylyk, B. (2001) Net-targeted mutant mice develop a vascular phenotype and up-regulate egr-1. *EMBO J.*, **20**, 5139-5152.
- Bailly, C., Tardy, C., Wang, L., Armitage, B., Hopkins, K., Kumar, A., Schuster, G.B., Boykin, D.W., Wilson, W.D. (2001) Recognition of ATGA sequences by the unfused aromatic dication DB293 forming stacked dimers in the DNA minor groove. *Biochemistry*, **40**, 9770-9779.
- Baldus, C.D., Martus, P., Burmeister, T., Schwartz, S., Gökbüget, N., Bloomfield, C.D., Hoelzer, D., Thiel, E. and Hofmann, W.K. (2007) Low ERG and BAALC expression identifies a new subgroup of adult acute T-lymphoblastic leukemia with a highly favorable outcome. *J. Clin. Oncol*, **25** 3739-37450.
- Bamberger, A.M., Bamberger, C.M., Pu, L.P., Puy, L.A., Loh, Y.P. and Asa, S.L. (1995) Expression of pit-1 messenger ribonucleic acid and protein in the human placenta. *J. Clin. Endocrinol. Metab.*, **80**, 2021-2026.
- Beck, K., Hunter, I. and Engel, J. (1990) Structure and function of laminin: anatomy of a multidomain glycoprotein. *FASEB J.*, **4**, 148-160.
- Beerman, T.A., McHugh, M.M., Sigmund, R., Lown, J.W., Rao, K.E. and Bathini Y. (1992) Effects of analogs of the DNA minor groove binder Hoechst 33258 on topoisomerase II and I mediated activities. *Biochim. Biophys. Acta.*, **1131**, 53-61.

- Bei, L., Lu, Y., Eklund, E.A. (2005) HOXA9 activates transcription of the gene encoding gp91Phox during myeloid differentiation. *J. Biol. Chem.*, **280**, 12359-12370.
- Belitsky, J.M., Leslie, S.J., Arora, P.S., Beerman, T.A. and Dervan, P.B. (2002) Cellular uptake of N-methylpyrrole/N-methylimidazole polyamide-dye conjugates. *Bioorg. Med. Chem.*, **10**, 3313-3318.
- Benz, C.C., O'Hagan, R.C., Richter, B., Scott, G.K., Chang, C.H., Xiong, X., Chew, K., Ljung, B.M., Edgerton, S., Thor, A. and Hassell, J.A. (1997) HER2/Neu and the Ets transcription activator PEA3 are coordinately upregulated in human breast cancer. *Oncogene*, **15**, 1513-1525.
- Berthelsen, J., Kilstrup-Nielsen, C., Blasi, F., Mavilio, F. and Zappavigna, V. (1999) The subcellular localization of PBX1 and EXD proteins depends on nuclear import and export signals and is modulated by association with PREP1 and HTH. *Genes Dev.*, **13**, 946-953.
- Best, T.P., Edelson, B.S., Nickols, N.G. and Dervan, P.B. (2003) Nuclear localization of pyrrole-imidazole polyamide-fluorescein conjugates in cell culture. *Proc. Natl. Acad. Sci U S A.*, **100**, 12063-12068.
- Blanc, V., Farré, J.C., Litvak, S., Araya, A. (2002) Réécriture du matériel génétique : fonctions et mécanismes de l'édition de l'ARN. *Médecine/Sciences*, **18**, 181-192.
- Block, K., Kardana, A., Igarashi, P. and Taylor, H.S. (2000). In utero diethylstilbestrol (DES) exposure alters Hox gene expression in the developing müllerian system. *FASEB J.*, **14**, 1101-1108.
- Bodner, M., Castrillo, J.L., Theill, L.E., Deerinck, T., Ellisman, M. and Karin, M. (1988) The pituitary-specific transcription factor GHF-1 is a homeobox-containing protein. *Cell*, **55**, 505-518.
- Bories, J.C., Willerford, D.M., Grévin, D., Davidson, L., Camus, A., Martin, P., Stéhelin, D. and Alt, F.W. (1995) Increased T-cell apoptosis and terminal B-cell differentiation induced by inactivation of the Ets-1 proto-oncogene. *Nature*, **377**, 635-638.
- Bouteille, B., Oukem, O., Bisser, S. and Dumas, M. (2003) Treatment perspectives for human African trypanosomiasis. *Fundam Clin Pharmacol.*, **17**, 171-181.
- Boyd, K.E. and Farnham, P.J. (1999) Identification of target genes of oncogenic transcription factors. *Proc. Soc. Exp. Biol. Med.*, **222**, 9-28.
- Bray, P.G., Barrett, M.P., Ward, S.A. and de Koning, H.P. (2003) Pentamidine uptake and resistance in pathogenic protozoa: past, present and future. *Trends Parasitol.*, **19**, 232-239.
- Bremer, R.E., Baird, E.E. and Dervan, P.B. (1998) Inhibition of major-groove-binding proteins by pyrrole-imidazole polyamides with an Arg-Pro-Arg positive patch. *Chem. Biol.*, **5**, 119-133.
- Brown, D.G., Sanderson, M.R., Skelly, J.V., Jenkins, T.C., Brown, T., Garman, E., Stuart, D.I. and Neidle, S. (1990) Crystal structure of a berenil-dodecanucleotide complex: the role of water in sequence-specific ligand binding. *EMBO J.*, **9**, 1329-1334.
- Buchwalter, G., Gross, C. and Wasyluk, B. (2004) Ets ternary complex transcription factors. *Gene*, **324**, 1-14.
- Budhram-Mahadeo, V., Ndisang, D., Ward, T., Weber, B.L. and Latchman, D.S. (1999) The Brn-3b POU family transcription factor represses expression of the BRCA-1 anti-oncogene in breast cancer cells. *Oncogene*, **18**, 6684-6691.

- Burley, S.K. and Roeder, R.G. (1996) Biochemistry and structural biology of transcription factor IID (TFIID). *Annu. Rev. Biochem.*, **65**, 769-799.
- Calò, V., Migliavacca, M., Bazan, V., Macaluso, M., Buscemi, M., Gebbia, N. and Russo, A. (2003) STAT proteins: from normal control of cellular events to tumorigenesis. *J. Cell. Physiol.*, **197**, 157-168.
- Calvo, K.R., Knoepfler, P.S., Sykes, D.B., Pasillas, M.P. and Kamps, M.P. (2001) Meis1a suppresses differentiation by G-CSF and promotes proliferation by SCF: potential mechanisms of cooperativity with Hoxa9 in myeloid leukemia. *Proc Natl. Acad. Sci. U S A.*, **98**, 13120-13125.
- Calvo, K.R., Sykes, D.B., Pasillas, M. and Kamps, M.P. (2000) Hoxa9 immortalizes a granulocyte-macrophage colony-stimulating factor-dependent promyelocyte capable of biphenotypic differentiation to neutrophils or macrophages, independent of enforced meis expression. *Mol. Cell. Biol.*, **20**, 3274-3285.
- Calvo, K.R., Sykes, D.B., Pasillas, M.P., Kamps, M.P. (2002) Nup98-HoxA9 immortalizes myeloid progenitors, enforces expression of Hoxa9, Hoxa7 and Meis1, and alters cytokine-specific responses in a manner similar to that induced by retroviral co-expression of Hoxa9 and Meis1. *Oncogene*, **21**, 4247-4256.
- Carter, N.S., Berger, B.J. and Fairlamb, A.H. (1995) Uptake of diamidine drugs by the P2 nucleoside transporter in melarsen-sensitive and -resistant *Trypanosoma brucei brucei*. *J. Mol. Biol. Chem.*, **270**, 28153-28157.
- Castrillo, J.L., Theill, L.E. and Karin, M. (1991) Function of the homeodomain protein GHF1 in pituitary cell proliferation. *Science*, **253**, 197-199.
- Catron, K.M., Iler, N. and Abate C. (1993) Nucleotides flanking a conserved TAAT core dictate the DNA binding specificity of three murine homeodomain proteins. *Mol. Cell. Biol.*, **13**, 2354-2365.
- Celetti, A., Barba, P. And Cillo, C. (1993) Characteristic patterns of HOX gene expression in different types of human leukemia. *Int. J. Cancer*, **53**, 237-244.
- Chakrabarti, S.R., Sood, R., Ganguly, S., Bohlander, S., Shen, Z. and Nucifora, G. (1999) Modulation of TEL transcription activity by interaction with the ubiquitin-conjugating enzyme UBC9. *Proc. Natl. Acad. Sci. U S A.*, **96**, 7467-7472.
- Chang, C.P., Shen, W.F., Rozenfeld, S., Lawrence, H.J., Largman, C. and Cleary, M.L.(1995) Pbx proteins display hexapeptide-dependent cooperative DNA binding with a subset of Hox proteins. *Genes Dev.*, **9**, 663-674.
- Chaturvedi, P., Sharma, S. and Reddy, E.P. (1997) Abrogation of interleukin-3 dependence of myeloid cells by the v-src oncogene requires SH2 and SH3 domains which specify activation of STATs. *Mol. Cell. Biol.*, **17**, 3295-3304.
- Chen, A.Y., Yu, C., Gatto, B., Liu, L.F. (1993) DNA minor groove-binding ligands: a different class of mammalian DNA topoisomerase I inhibitors. *Proc. Natl. Acad. Sci. U S A.* **90**, 8131-8135.
- Chen, L., Fischle, W., Verdin, E. and Greene, C. (2001) Duration of nuclear NF- κ B action regulated by reversible acetylation, *Science*, **293**, 1653-1657.
- Chiang, S.Y., Burli, R.W., Benz, C.C., Gawron, L., Scott, G.K., Dervan, P.B. and Berman TA. (2000) Targeting the ets binding site of the HER2/neu promoter with pyrrole-imidazole polyamides. *J. Biol. Chem.*, **275**, 24246-24254.

- Chiurazzi, P., Pomponi, M.G., Willemsen, R., Oostra, B.A. and Neri, G. (1998) In vitro reactivation of the FMR1 gene involved in fragile X syndrome. *Hum. Mol. Genet.*, **7**, 109-113.
- Choi, H.W., Shin, M.G., Sawyer, J.R., Cho, D., Kee, S.J., Baek, H.J., Kook, H., Kim, H.J., Shin, J.H., Suh, S.P., Hwang, T.J. and Ryang, D.W. (2006) Unusual type of TLS/FUS-ERG chimeric transcript in a pediatric acute myelocytic leukemia with 47,XX,+10,t(16;21)(p11;q22). *Cancer Genet. Cytogenet.*, **167**, 172-176.
- Choo, Y. and Klug, A. (1994) Toward a code for the interactions of zinc fingers with DNA: selection of randomized fingers displayed on phage. *Proc. Natl. Acad. Sci. U S A.*, **23**, 11163-11167.
- Chrivia, J.C., Kwok, R.P., Lamb, N., Hagiwara, M., Montminy, M.R. and Goodman, R.H.. (1993) Phosphorylated CREB binds specifically to the nuclear protein CBP. *Nature*, **365**, 855-859.
- Ciliberto, G., Castagnoli, L. and Cortese, R. (1983) Transcription by RNA polymerase III. *Curr. Top. Dev. Biol.*, **18**, 59-88.
- Cleary, M.A. and Herr, W. (1995) Mechanisms for flexibility in DNA sequence recognition and VP16-induced complex formation by the Oct-1 POU domain. *Mol. Cell. Biol.*, **15**, 2090-2100.
- Cole, M.D. and McMahon, S.B. (1999) The Myc oncoprotein: a critical evaluation of transactivation and target gene regulation. *Oncogene*, **18**, 2916-2924.
- Coll, M., Aymami, J., van der Marel, G.A., van Boom, J.H., Rich, A. and Wang, A.H. (1989) Molecular structure of the netropsin-d(CGCGATATCGCG) complex: DNA conformation in an alternating AT segment. *Biochemistry*, **28**, 310-320.
- Corden, J.L. and Patturajan, M.A. (1997) CTD function linking transcription to splicing. *Trends Biochem. Sci.*, **22**, 413-416.
- Cowley, D.O. and Graves, B.J. (2000) Phosphorylation represses Ets-1 DNA binding by reinforcing autoinhibition. *Genes Dev.*, **14**, 366-376.
- Czuwara-Ladykowska, J., Sementchenko, V.I., Watson, D.K. and Trojanowska, M. (2002) Ets1 is an effector of the transforming growth factor beta (TGF-beta) signaling pathway and an antagonist of the profibrotic effects of TGF-beta. *J. Biol. Chem.*, **277**, 20399-20408.
- Daftary, G.S. and Taylor, H.S. (2006) Endocrine regulation of HOX genes. *Endocr. Rev.*, **27**, 331-355.
- Darnell, J.E. Jr. (2002) Transcription factors as targets for cancer therapy. *Nat. Rev. Cancer*, **2**, 740-749.
- Dash, A.B., Williams, I.R., Kutok, J.L., Tomasson, M.H., Anastasiadou, E., Lindahl, K., Li, S., Van Etten, R.A., Borrow, J., Housman, D., Druker, B. and Gilliland, .D.G. (2001) A murine model of CML blast crisis induced by cooperation between BCR/ABL and NUP98/HOXA9. *Proc. Natl. Acad. Sci. U S A.*, **99**, 7622-7627.
- De Koning, H.P. and Jarvis, S.M. (2001) Uptake of pentamidine in *Trypanosoma brucei brucei* is mediated by the P2 adenosine transporter and at least one novel, unrelated transporter. *Acta. Trop.*, **80**, 245-250.
- De Souza, E.M., Lansiaux, A., Bailly, C., Wilson, W.D., Hu, Q., Boykin, D.W., Batista, M.M., Araújo-Jorge, T.C. and Soeiro, M.N. (2004) Phenyl substitution of furamidine markedly potentiates its anti-parasitic activity against *Trypanosoma cruzi* and *Leishmania amazonensis*. *Biochem. Pharmacol.*, **68**, 593-600.

- DeKoter, R.P. and Singh, H. (2000) Regulation of B lymphocyte and macrophage development by graded expression of PU.1. *Science*, **288**, 1439-1441.
- Delhase, M., Vergani, P., Malur, A., Velkeniers, B., Teugels, E., Trouillas, J. and Hooghe-Peters, E.L. (1993) Pit-1/GHF-1 expression in pituitary adenomas: further analogy between human adenomas and rat SMtTW tumours. *J. Mol. Endocrinol.*, **11**, 129-139.
- Dennis, J.H., Budhram-Mahadeo, V. and Latchman, D.S. (2001) The Brn-3b POU family transcription factor regulates the cellular growth, proliferation, and anchorage dependence of MCF7 human breast cancer cells. *Oncogene*, **20**, 4961-4971.
- Dervan, P.B., Doss, R.M. and Marques, M.A. (2005) Programmable DNA binding oligomers for control of transcription. *Curr. Med. Chem. Anticancer Agents*, **5**, 373-387.
- Dickinson, L.A., Gulizia, R.J., Trauger, J.W., Baird, E.E., Mosier, D.E., Gottesfeld, J.M. and Dervan, P.B. (1998) Inhibition of RNA polymerase II transcription in human cells by synthetic DNA-binding ligands. *Proc. Natl. Acad. Sci. U S A.*, **95**, 12890-12895.
- Dintilhac, A., Bihan, R., Guerrier, D., Deschamps, S. and Pellerin, I. (2004) A conserved non-homeodomain Hoxa9 isoform interacting with CBP is co-expressed with the 'typical' Hoxa9 protein during embryogenesis. *Gene Expr. Patterns*, **4**, 215-222.
- Dorsam, S.T., Ferrell, C.M., Dorsam, G.P., Derynck, M.K., Vijapurkar, U., Khodabakhsh, D., Pau, B., Bernstein, H., Haqq, C.M., Largman, C. and Lawrence, H.J. (2004) The transcriptome of the leukemogenic homeoprotein HOXA9 in human hematopoietic cells. *Blood*, **103**, 1676-1684.
- Draganescu, A., Levin, J.R., Tullius, T.D. (1995) Homeodomain proteins: what governs their ability to recognize specific DNA sequences? *J. Mol. Biol.*, **250**, 595-608.
- Du, H., Daftary, G.S., Lalwani, S.I. and Taylor H.S. (2005) Direct regulation of HOXA10 by 1,25-(OH)₂D₃ in human myelomonocytic cells and human endometrial stromal cells. *Mol Endocrinol.*, **19**, 2222-2233.
- Dubois, M.F., Nguyen, V.T., Dahmus, M.E., Pagès, G., Pouyssegur, J. and Bensaude, O. (1994) Enhanced phosphorylation of the C-terminal domain of RNA polymerase II upon serum stimulation of quiescent cells: possible involvement of MAP kinases. *EMBO J.*, **13**, 4787-4797.
- Duboule D. (2007) The rise and fall of Hox gene clusters. *Development*, **134**, 2549-2560.
- Dwyer, T.J., Geierstanger, B.H., Bathini, Y., Lown, J.W. and Wemmer, D.E. (1992) Design and binding of a distamycin A analog to d(CGCAAGTTGGC).d(GCCAACTTGCG) : synthesis, NMR studies, and implications for the design of sequence-specific minor groove binding oligopeptides. *J. Am. Chem. Soc.*, **114**, 5911-5919.
- Dynan, W.S. and Tjian, R. (1985) Control of eukaryotic messenger RNA synthesis by sequence-specific DNA-binding proteins. *Nature*, **316**, 774-778.
- Edelson, B.S., Best, T.P., Olenyuk, B., Nickols, N.G., Doss, R.M., Foister, S., Heckel, A. and Dervan, P.B. (2004) Influence of structural variation on nuclear localization of DNA-binding polyamide-fluorophore conjugates. *Nucleic Acids Res.*, **32**, 2802-2818.
- Ehley, J.A., Melander, C., Herman, D., Baird, E.E., Ferguson, H.A., Goodrich, J.A., Dervan, P.B. and Gottesfeld, J.M. (2002) Promoter scanning for transcription inhibition with DNA-binding polyamides. *Mol Cell Biol.*, **22**, 1723-1733.

- Eklund, E.A. (2007) The role of HOX genes in malignant myeloid disease. *Curr. Opin. Hematol.*, **14**, 85-89.
- Escalante, C.R., Yie, J., Thanos, D. and Aggarwal, A.K. (1998) Structure of IRF-1 with bound DNA reveals determinants of interferon regulation. *Nature*, **391**, 103-106.
- Fechter, E.J., Dervan, P.B. (2003) Allosteric inhibition of protein-DNA complexes by polyamide-intercalator conjugates. *J. Am. Chem. Soc.*, **125**, 8476-8485.
- Felsenfeld, G., Boyes, J., Chung, J., Clark, D. and Studitsky V. (1996) Chromatin structure and gene expression. *Proc. Natl. Acad. Sci. U S A.*, **93**, 9384-9388.
- Finney, M. and Ruvkun, G. (1990) The unc-86 gene product couples cell lineage and cell identity in *C. elegans*. *Cell*, **63**, 895-905.
- Gavrilov, D., Kenzior, O., Evans, M., Calaluce, R. and Folk, W.R. (2001) Expression of urokinase plasminogen activator and receptor in conjunction with the ets family and AP-1 complex transcription factors in high grade prostate cancers. *Eur. J. Cancer*, **37**, 1033-1040.
- Geierstanger, B.H., Jacobsen, J.P., Mrksich, M., Dervan, P.B. and Wemmer, D.E. (1994) Structural and dynamic characterization of the heterodimeric and homodimeric complexes of distamycin and 1-methylimidazole-2-carboxamide-netropsin bound to the minor groove of DNA. *Biochemistry*, **33**, 3055-3062.
- Ghannam, G., Takeda, A., Camarata, T., Moore, M.A., Viale A. and Yaseen, N.R. (2004) The oncogene Nup98-HOXA9 induces gene transcription in myeloid cells. *J. Biol. Chem.*, **279**, 866-875.
- Gil-Puig, C., Blanco, M., García-Caballero, T., Segura, C. and Pérez-Fernández, R. (2002) Pit-1/GHF-1 and GH expression in the MCF-7 human breast adenocarcinoma cell line. *J. Endocrinol.*, **173**, 161-167.
- Gil-Puig, C., Seoane, S., Blanco, M., Macia, M., Garcia-Caballero, T., Segura, C. and Perez-Fernandez, R. (2005) Pit-1 is expressed in normal and tumorous human breast and regulates GH secretion and cell proliferation. *Eur. J. Endocrinol.*, **153**, 335-344.
- Goel, A. and Janknecht, R. (2003) Acetylation-mediated transcriptional activation of the ETS protein ER81 by p300, P/CAF, and HER2/Neu. *Mol. Cell. Biol.*, **23**, 6243-6254.
- Goetz, T.L., Gu, T.L., Speck, N.A. and Graves, B.J. (2000) Auto-inhibition of Ets-1 is counteracted by DNA binding cooperativity with core-binding factor alpha2. *Mol. Cell. Biol.*, **20**, 81-90.
- Golub, T.R., Barker, G.F., Lovett, M. and Gilliland, D.G. (1994) Fusion of PDGF receptor beta to a novel ets-like gene, tel, in chronic myelomonocytic leukemia with t(5;12) chromosomal translocation. *Cell*, **77**, 307-16.
- Golub, T.R., Slonim, D.K., Tamayo, P., Huard, C., Gaasenbeek, M., Mesirov, J.P., Coller, H., Loh, M.L., Downing, J.R., Caligiuri, M.A., Bloomfield, C.D. and Lander, E.S. (1999) Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science*, **286**, 531-537.
- Goodsell, D. and Dickerson, R.E. (1986) Isohelical analysis of DNA groove-binding drugs. *J. Med. Chem.*, **29**, 727-733.

- Gottesfeld, J.M., Neely, L., Trauger, J.W., Baird, E.E. and Dervan, P.B. (1997) Regulation of gene expression by small molecules. *Nature*, **387**, 202-205.
- Greenall, A., Willingham, N., Cheung, E., Boam, D.S. and Sharrocks, A.D. (2000) DNA binding by the ETS-domain transcription factor PEA3 is regulated by intramolecular and intermolecular protein-protein interactions. *J. Biol. Chem.*, **276**, 16207-16215.
- Griffin, J.D. (2004) FLT3 tyrosine kinase as a target in acute leukemias. *Hematol J.*, **5**, 188-190.
- Grunstein, M. (1997) Histone acetylation in chromatin structure and transcription. *Nature*, **389**, 349-352.
- Gu, W. and Roeder, R.G. (1997) Activation of p53 sequence-specific DNA binding by acetylation of the p53 C-terminal domain. *Cell*, **90**, 595-606.
- Hanson, C.A., Wood, L.D. and Hiebert, S.W. (2007) Cellular stress triggers TEL nuclear export via two genetically separable pathways. *J. Cell. Biochem.*, Epub ahead of print.
- Haq, I. (2002) Thermodynamics of drug-DNA interactions. *Arch. Biochem. Biophys.*, **403**, 1-15.
- Heinemann, U., Alings, C. and Bansal, M. (1992) Double helix conformation, groove dimensions and ligand binding potential of a G/C stretch in B-DNA. *EMBO J.*, **11**, 1931-1939.
- Herr, W. and Cleary, M.A. (1995) The POU domain: versatility in transcriptional regulation by a flexible two-in-one DNA-binding domain. *Genes Dev.*, **9**, 1679-1693
- Hochhauser, D., Kotecha, M., O'hare, C., Morris, P.J., Hartley, J.M., Taherbhai, Z., Harris, D., Forni, C., Mantovani, R., Lee, M. and Hartley, J.A. (2007) Modulation of topoisomerase IIalpha expression by a DNA sequence-specific polyamide. *Mol. Cancer. Ther.*, **6**, 346-354.
- Holtkamp, W., Nagel, G.A., Wander, H.E., Rauschecker, H.F. and von Heyden, D. (1984) Hyperprolactinemia is an indicator of progressive disease and poor prognosis in advanced breast cancer. *Int. J. Cancer.*, **34**, 323-328.
- Horton, H., Moran, L., Ochs, R., Rawn, J. and Scrimgeour, K. (1993) Principle of biochemistry, DeBoeck-Wesmael.
- Horvath, C.M. (2000) STAT proteins and transcriptional responses to extracellular signals. *Trends Biochem. Sci.*, **25**, 496-502.
- Hsieh-Li, H.M., Witte, D.P., Weinstein, M., Branford, W., Li, H., Small, K. and Potter S.S. (1995) Hoxa 11 structure, extensive antisense transcription, and function in male and female fertility. *Development*, **121**, 1373-1385.
- Hurley, L.H. (2002) DNA and its associated processes as targets for cancer therapy. *Nat. Rev. Cancer*, **2**, 188-200.
- Hurst, H.C. (1994) Transcription factors. 1: bZIP proteins. *Protein Profile*, **1**, 123-168.
- Ichikawa, H., Shimizu, K., Hayashi, Y. and Ohki, M. (1994) An RNA-binding protein gene, TLS/FUS, is fused to ERG in human myeloid leukemia with t(16;21) chromosomal translocation. *Cancer Res.*, **54**, 2865-2868.
- Imler, J.L., Schatz, C., Wasylyk, C., Chatton, B. and Wasylyk, B. (1988) A Harvey-ras responsive transcription element is also responsive to a tumour-promoter and to serum. *Nature*, **332**, 275-278.

- Ingraham, H.A., Chen, R.P., Mangalam, H.J., Elsholtz, H.P., Flynn, S.E., Lin, C.R., Simmons, D.M., Swanson, L. and Rosenfeld, M.G. (1988) A tissue-specific transcription factor containing a homeodomain specifies a pituitary phenotype. *Cell*, **55**, 519-529.
- Irshad, S., Pedley, R.B., Anderson, J., Latchman, D.S. and Budhram-Mahadeo, V. (2004) The Brn-3b transcription factor regulates the growth, behavior, and invasiveness of human neuroblastoma cells in vitro and in vivo. *J. Biol. Chem.*, **279**, 21617-2127.
- Ito, Y., Miyoshi, E., Takeda, T. and Sakon, M., Ihara, S., Tsujimoto, M. and Matsuura, N. (2002) Ets-2 overexpression contributes to progression of pancreatic adenocarcinoma. *Oncol. Rep.*, **9**, 853-857.
- Jacobson, E.M., Li, P., Leon-del-Rio, A., Rosenfeld, M.G. and Aggarwal, A.K. (1997) Structure of Pit-1 POU domain bound to DNA as a dimer: unexpected arrangement and flexibility. *Genes Dev.*, **11**, 198-212.
- Johnson, G.L. and Nakamura, K. (2007) The c-jun kinase/stress-activated pathway: regulation, function and role in human disease. *Biochim. Biophys. Acta.*, **1773**, 1341-1348.
- Jonsen, M.D., Petersen, J.M., Xu, Q.P. and Graves, B.J. (1996) Characterization of the cooperative function of inhibitory sequences in Ets-1. *Mol. Cell. Biol.*, **16**, 2065-2073.
- Kasper, L.H., Brindle, P.K., Schnabel, C.A., Pritchard, C.E., Cleary, M.L. and van Deursen J.M. (1999) CREB binding protein interacts with nucleoporin-specific FG repeats that activate transcription and mediate NUP98-HOXA9 oncogenicity. *Mol. Cell. Biol.*, **19**, 764-776.
- Kaushansky, K. (2006) Hematopoietic growth factors, signaling and the chronic myeloproliferative disorders. *Cytokine Growth Factor Rev.*, **17**, 423-430.
- Kielkopf, C.L., Baird, E.E., Dervan, P.B. and Rees, D.C. (1998) Structural basis for G.C recognition in the DNA minor groove. *Nat. Struct. Biol.*, **5**, 104-109.
- Kim, C.A., Phillips, M.L., Kim, W., Gingery, M., Tran, H.H., Robinson, M.A., Faham, S. and Bowie, J.U. (2001) Polymerization of the SAM domain of TEL in leukemogenesis and transcriptional repression. *EMBO J.*, **20**, 4173-4182.
- Kingston, R.E., Bunker, C.A. and Imbalzano, A.N. (1996) Repression and activation by multiprotein complexes that alter chromatin structure. *Genes Dev.*, **10**, 905-920.
- Kirito, K., Fox, N. and Kaushansky, K. (2004) Thrombopoietin induces HOXA9 nuclear transport in immature hematopoietic cells: potential mechanism by which the hormone favorably affects hematopoietic stem cells. *Mol. Cell. Biol.*, **24**, 6751-6762.
- Kirmizis, A. and Farnham, P.J. (2004) Genomic approaches that aid in the identification of transcription factor target genes. *Exp. Biol. Med.*, **229**, 705-721.
- Kissinger, C.R., Liu, B.S., Martin-Blanco, E., Kornberg, T.B. and Pabo, C.O. (1990) Crystal structure of an engrailed homeodomain-DNA complex at 2.8 Å resolution: a framework for understanding homeodomain-DNA interactions. *Cell*, **63**, 579-590.
- Klemm, J.D. and Pabo, C.O. (1996) Oct-1 POU domain-DNA interactions: cooperative binding of isolated subdomains and effects of covalent linkage. *Genes Dev.*, **10**, 27-36.

- Klemm, J.D., Rould, M.A., Aurora, R., Herr, W. and Pabo, C.O. (1994) Crystal structure of the Oct-1 POU domain bound to an octamer site: DNA recognition with tethered DNA-binding modules. *Cell*, **77**, 21–32.
- Kodandapani, R., Pio, F., Ni, C.Z., Piccialli, G., Klemsz, M., McKercher, S., Maki, R.A. and Ely, K.R. (1996) A new pattern for helix-turn-helix recognition revealed by the PU.1 ETS-domain-DNA complex. *Nature*, **380**, 456-460. Erratum in: *Nature* (1998), 392, 630.
- Kopka, M.L., Yoon, C., Goodsell, D., Pjura, P. and Dickerson, R.E. (1985) The molecular origin of DNA-drug specificity in netropsin and distamycin. *Proc. Natl. Acad. Sci. U S A.*, **82**, 1376-1380.
- Kraemer D, Wozniak, R.W., Blobel, G. and Radu A. (1994) The human CAN protein, a putative oncogene product associated with myeloid leukemogenesis, is a nuclear pore complex protein that faces the cytoplasm. *Proc. Natl. Acad. Sci. U S A.*, **91**, 1519-1523.
- Kroon, E., Kros, J., Thorsteinsdottir, U., Baban, S., Buchberg, A.M. and Sauvageau G. (1998) Hoxa9 transforms primary bone marrow cells through specific collaboration with Meis1a but not Pbx1b. *EMBO J.*, **17**, 3714-3725.
- Kroon, E., Thorsteinsdottir, U., Mayotte, N., Nakamura, T. and Sauvageau, G. (2001) NUP98-HOXA9 expression in hemopoietic stem cells induces chronic and acute myeloid leukemias in mice. *EMBO J.*, **20**, 350-361.
- Lam, R. and Li, X. (2002) An array-based method for specifically profiling multiple transcription factor activity. *Am. Biotechnol. Lab.* 8:22–26.
- Langston, A.W. and Gudas, L.J. (1992) Identification of a retinoic acid responsive enhancer 3' of the murine homeobox gene Hox-1.6. *Mech. Dev.*, **38**, 217-227.
- Lania, L., Majello, B. and De Luca, P. (1997) Transcriptional regulation by the Sp family proteins. *Int. J. Biochem. Cell. Biol.*, **29**, 1313-1323.
- Lansiaux, A., Dassonneville, L., Facompré, M., Kumar, A., Stephens, C.E., Bajic, M., Tanious, F., Wilson, W.D., Boykin, D.W. and Bailly, C. (2002a) Distribution of furamide analogues in tumor cells: influence of the number of positive charges. *J Med Chem.*, **45**, 1994-2002.
- 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 furamide analogues in tumor cells: targeting of the nucleus or mitochondria depending on the amidine substitution. *Cancer Res.*, **62**, 7219-7229.
- LaRonde-LeBlanc, N.A. and Wolberger, C. Structure of HoxA9 and Pbx1 bound to DNA: Hox hexapeptide and DNA recognition anterior to posterior. *Genes Dev.*, **17**, 2060-2072.
- Latchman, D. S. (1998) Eukaryotic transcription factors. Third edition. Academic press.
- Laudet, V., Hänni, C., Stéhelin, D. And Duterque-Coquillaud, M. (1999) Molecular phylogeny of ETS gene family. *Oncogene*, **18**, 1351-1359.
- Laughon, A. (1991) DNA binding specificity of homeodomains. *Biochemistry*, **30**, 11357-11367.
- Laughton, C.A., Tanious, F., Nunn, C.M., Boykin, D.W., Wilson, W.D. and Neidle, S. (1996) A crystallographic and spectroscopic study of the complex between d(CGCGAATTCGCG)₂ and 2,5-bis(4-guanylphenyl)furan, an analogue of berenil. Structural origins of enhanced DNA-binding affinity. *Biochemistry*, **35**, 5655-5661.

- Lawrence, H.J. and Largman, C. (1992) Homeobox genes in normal hematopoiesis and leukemia. *Blood*, **80**, 2445-2453.
- Leblond-Francillard, M., Picon, A., Bertagna, X. and de Keyzer, Y. (1997) High expression of the POU factor Brn3a in aggressive neuroendocrine tumors. *J. Clin. Endocrinol. Metab.*, **82**, 89-94.
- Lee, M., Hartley, J.A., Pon, R.T., Krowicki, K. and Lown, J.W. (1988) Sequence specific molecular recognition by a monocationic lexitropin of the decaoxyribonucleotide d-[CATGGCCATG]₂: structural and dynamic aspects deduced from high field 1H-NMR studies. *Nucleic Acids Res.*, **16**, 665-684.
- Lee, M.S., Gippert, G.P., Soman, K.V., Case, D.A. and Wright, P.E. (1989) Three-dimensional solution structure of a single zinc finger DNA-binding domain. *Science*, **245**, 635-637.
- Lelièvre, E., Lionneton, F., Soncin, F. and Vandebunder, B. (2001) The Ets family contains transcriptional activators and repressors involved in angiogenesis. *Int. J. Biochem. Cell. Biol.*, **33**, 391-407.
- Li, S., Crenshaw, E.B. 3rd, Rawson, E.J., Simmons, D.M., Swanson, L.W. and Rosenfeld, M.G. (1990) Dwarf locus mutants lacking three pituitary cell types result from mutations in the POU-domain gene pit-1. *Nature*, **347**, 528-533.
- Li, P., He, X., Gerrero, M.R., Mok, M., Aggarwal, A. and Rosenfeld, M.G. (1993) Spacing and orientation of bipartite DNA-binding motifs as potential functional determinants for POU domain factors. *Genes Dev.*, **7**, 2483-2496.
- Li, R., Pei, H. and Watson, D.K. (2000) Regulation of Ets function by protein - protein interactions. *Oncogene*, **19**, 6514-6523.
- Li, B., Carey, M. and Workman, J.L. (2007) The role of chromatin during transcription. *Cell*, **128**, 707-719.
- Liang, H., Mao, X., Olejniczak, E.T., Nettesheim, D.G., Yu, L., Meadows, R.P., Thompson, C.B. and Fesik, S.W. (1994) Solution structure of the ets domain of Fli-1 when bound to DNA. *Nat Struct Biol.*, **1**, 871-875.
- Lim, F., Kraut, N., Frampton, J. and Graf, T. (1992) DNA binding by c-Ets-1, but not v-Ets, is repressed by an intramolecular mechanism. *EMBO J.*, **11**, 643-652.
- Lin, J.H., Saito, T., Anderson, D.J., Lance-Jones, C., Jessell, T.M. and Arber, S. (1998) Functionally related motor neuron pool and muscle sensory afferent subtypes defined by coordinate ETS gene expression. *Cell*, **95**, 393-407.
- Lipsick, J.S. and Wang, D.M. (1999) Transformation by v-Myb. *Oncogene*. **18**, 3047-3055.
- Littlewood, T.D. and Evan, G.I. (1995) Transcription factors 2: helix-loop-helix. *Protein Profile*, **2**, 621-702.
- Lown, J.W., Krowicki, K., Bhat, U.G., Skorobogaty, A., Ward, B. and Dabrowiak, J.C. (1986) Molecular recognition between oligopeptides and nucleic acids: novel imidazole-containing oligopeptides related to netropsin that exhibit altered DNA sequence specificity. *Biochemistry*, **25**, 7408-7416.
- Luger, K., Mäder, A.W., Richmond, R.K., Sargent, D.F. and Richmond, T.J. (1997) Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature*, **389**, 231-233.

- Luscombe, N.M., Austin, S.E., Berman, H.M. and Thornton, J.M. (2000) An overview of the structures of protein-DNA complexes. *Genome Biol.*, **1**, 1-37.
- Maity, S.N and de Crombrughe, B. (1998) Role of the CCAAT-binding protein CBF/NF-Y in transcription. *Trends Biochem.*, **23**, 174-178.
- Maizel, A., Bensaude, O., Prochiantz, A. and Joliot, A. (1999) A short region of its homeodomain is necessary for engrailed nuclear export and secretion. *Development*, **126**, 3183-3190.
- Mallena, S., Lee, M.P., Bailly, C., Neidle, S., Kumar, A., Boykin, D.W. and Wilson, W.D. (2004) Thiophene-based diamidine forms a "super" at binding minor groove agent. *J. Am. Chem. Soc.*, **126**, 13659-13669.
- Marcucci, G., Maharry, K., Whitman, S.P., Vukosavljevic, T., Paschka, P., Langer, C., Mrózek, K., Baldus, C.D., Carroll, A.J., Powell, B.L., Kolitz, J.E., Larson, R.A. and Bloomfield, C.D. (2007) Cancer and Leukemia Group B Study. High expression levels of the ETS-related gene, ERG, predict adverse outcome and improve molecular risk-based classification of cytogenetically normal acute myeloid leukemia : a Cancer and Leukemia Group B Study. *J. Clin. Oncol.*, **25**, 3337-3343.
- Marks, A. P., Rifkind, R. A., Richon, V. M., Breslow, R., Miller, T. And Kelly, K. K. (2001) Histone deacetylase and cancer : causes and therapies. *Nat .Rev. Cancer*, **1**, 194-1202.
- Marshall, H., Studer, M., Pöpperl, H., Aparicio, S., Kuroiwa, A., Brenner, S. and Krumlauf R. (1994) A conserved retinoic acid response element required for early expression of the homeobox gene Hoxb-1. *Nature*, **370**, 567-571.
- Mathews, C.H., Detmer, K., Boncinelli, E., Lawrence, H.J. and Largman, C. .Erythroid-restricted expression of homeobox genes of the human HOX 2 locus. *Blood*, **78**, 2248-2252.
- Matovu, E., Stewart, M.L., Geiser, F., Brun, R., Mäser, P., Wallace, L.J., Burchmore, R.J., Enyaru, J.C., Barrett, M.P., Kaminsky, R., Seebeck, T. and de Koning, H.P. (2003) Mechanisms of arsenical and diamidine uptake and resistance in *Trypanosoma brucei*. *Eukaryot. Cell*, **2**, 1003-1008.
- Matre, V., Høvring, P.I., Orstavik, S., Frengen, E., Rian, E., Velickovic, Z., Murray-McIntosh, R.P. and Gautvik, K.M. (1999) Structural and functional organization of the gene encoding the human thyrotropin-releasing hormone receptor. *J. Neurochem.*, **72**, 40-50.
- Mattot, V., Vercamer, C., Soncin, F., Fafeur, V. and Vandenbunder, B. (1999) Transcription factors of the Ets family and morphogenesis of the vascular tree. *J. Soc. Biol.*, **193**, 147-153.
- Mavrothalassitis, G. and Ghysdael, J. (2000) Proteins of the ETS family with transcriptional repressor activity. *Oncogene*, **19**, 6524-6532.
- Mayotte, N., Roy, D.C., Yao, J., Kroon, E. and Sauvageau G. (2002) Oncogenic interaction between BCR-ABL and NUP98-HOXA9 demonstrated by the use of an in vitro purging culture system. *Blood*, **100**, 4177-4184.
- McCormick, A., Brady, H., Theill, L.E. and Karin, M. (1990) Regulation of the pituitary-specific homeobox gene GHF1 by cell-autonomous and environmental cues. *Nature*, **345**, 829-832.
- McEvelly, R.J., Erkman, L., Luo, L., Sawchenko, P.E., Ryan, A.F. and Rosenfeld, M.G. (1996) Requirement for Brn-3.0 in differentiation and survival of sensory and motor neurons. *Nature*, **384**, 574-577.

- Mellon, P.L., Clegg, C.H., Correll, L.A. and McKnight, G.S. (1989) Regulation of transcription by cyclic AMP-dependent protein kinase. *Proc. Natl. Acad. Sci. U S A.*, **86**, 4887-4891.
- Miao, Y., Lee, M.P., Parkinson, G.N., Batista-Parra, A., Ismail, M.A., Neidle, S., Boykin, D.W. and Wilson, W.D. (2005) Out-of-shape DNA minor groove binders: induced fit interactions of heterocyclic dications with the DNA minor groove. *Biochemistry*, **44**, 14701-14708.
- Mizzen, C.A., Yang, X.J., Kokubo, T., Brownell, J.E., Bannister, A.J., Owen-Hughes, T., Workman, J., Wang, L., Berger, S.L., Kouzarides, T., Nakatani, Y. and Allis, C.D. (1996) The TAF(II)250 subunit of TFIID has histone acetyltransferase activity. *Cell*, **87**, 1261-1270.
- Mlodzik, M., Fjose, A. and Gehring, W.J. (1988) Molecular structure and spatial expression of a homeobox gene from the labial region of the Antennapedia-complex. *EMBO J.*, **7**, 2569-2578.
- Moore, S.D., Offor, O., Ferry, J.A., Amrein, P.C., Morton, C.C. and Dal Cin, P. (2006) ELF4 is fused to ERG in a case of acute myeloid leukemia with a t(X;21)(q25-26;q22). *Leuk. Res.*, **30**, 1037-1042.
- Morgado, E., Albouhair, S. and Lavau, C. (2007) Flt3 is dispensable to the Hoxa9/Meis1 leukemogenic cooperation. *Blood*, **109**, 4020-4022.
- Mrksich, M. and Dervan, P.B. (1994) Design of a covalent peptide heterodimer for sequence-specific recognition in the minor groove of double-helical DNA. *J. Am. Chem. Soc.*, **116**, 3663-3664.
- Mrksich, M., Wade, W.S., Dwyer, T.J., Geierstanger, B.H., Wemmer, D.E. and Dervan, P.B. (1992) Antiparallel side-by-side dimeric motif for sequence-specific recognition in the minor groove of DNA by the designed peptide 1-methylimidazole-2-carboxamide netropsin. *Proc. Natl. Acad. Sci. U S A.*, **89**, 7586-7590.
- Mueller, B.U., Pabst, T., Osato, M., Asou, N., Johansen, L.M., Minden, M.D., Behre, G., Hiddemann, W., Ito, Y. and Tenen, D.G. (2002) Heterozygous PU.1 mutations are associated with acute myeloid leukemia. *Blood*, **100**, 998-1007.
- Müller, M.M., Ruppert, S., Schaffner, W. and Matthias, P. (1988) A cloned octamer transcription factor stimulates transcription from lymphoid-specific promoters in non-B cells. *Nature*, **336**, 544-551.
- Mundle, M., Ismail, M.A., Arafa, R., Peixoto, P., Collar, C.J., Liu, Y., Hu, L., David-Cordonnier, M.H., Lansiaux, A., Bailly, C., Boykin, D.W. and Wilson, W.D. (2007) Design of DNA minor groove binding diamidines that recognize GC base pair sequences: a dimeric-hinge interaction motif. *J. Am. Chem. Soc.*, **129**, 13732-13743.
- Murphy, F.V. 4th, Sweet, R.M. and Churchill, M.E. (1999) The structure of a chromosomal high mobility group protein-DNA complex reveals sequence-neutral mechanisms important for non-sequence-specific DNA recognition. *EMBO J.*, **18**, 6610-6618.
- Muthusamy, N., Barton, K. and Leiden, J.M. (1995) Defective activation and survival of T cells lacking the Ets-1 transcription factor. *Nature*, **377**, 639-642.
- Nakai, S., Kawano, H., Yodate, T., Nishi, M., Kuno, J., Nagata, A., Jishage, K., Hamada, H., Fujii, H. and Kawamura, K. (1995) The POU domain transcription factor Brn-2 is required for the determination of specific neuronal lineages in the hypothalamus of the mouse. *Genes Dev.*, **9**, 3109-3121.
- Nakajima, T., Uchida, C., Anderson, S.F., Parvin, J.D. and Montminy, M. (1997) Analysis of a cAMP-responsive activator reveals a two-component mechanism for transcriptional induction via signal-dependent factors. *Genes Dev.*, **11**, 738-747.

- Nakamura, T., Largaespada, D.A., Lee, M.P., Johnson, L.A., Ohyashiki, K., Toyama, K., Chen, S.J., Willman, C.L., Chen, I.M., Feinberg, A.P., Jenkins, N.A., Copeland, N.G. and Shaughnessy, J.D. Jr. (1996) Fusion of the nucleoporin gene NUP98 to HOXA9 by the chromosome translocation t(7;11)(p15;p15) in human myeloid leukaemia. *Nat. Genet.*, **12**, 154-158.
- Nan, X., Ng, H.H., Johnson, C.A., Laherty, C.D., Turner, B.M., Eisenman, R.N. and Bird, A. (1998) Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. *Nature*, **393**, 386-389.
- Nardelli, J., Gibson, T.J., Vesque, C. and Charnay, P. (1991) Base sequence discrimination by zinc-finger DNA-binding domains. *Nature*, **349**, 175-178.
- Ndisang, D., Budhram-Mahadeo, V. and Latchman, D.S. (1999) The Brn-3a transcription factor plays a critical role in regulating human papilloma virus gene expression and determining the growth characteristics of cervical cancer cells. *J. Biol. Chem.*, **274**, 28521-28527.
- Ndisang, D., Budhram-Mahadeo, V., Pedley, B. and Latchman, D.S. (2001) The Brn-3a transcription factor plays a key role in regulating the growth of cervical cancer cells in vivo. *Oncogene*, **20**, 4899-4903.
- Ndisang, D., Morris, P.J., Chapman, C., Ho, L., Singer, A. and Latchman, D.S. (1998) The HPV-activating cellular transcription factor Brn-3a is overexpressed in CIN3 cervical lesions. *J. Clin. Invest.*, **101**, 1687-1692.
- Neely, L., Trauger, J.W., Baird, E.E., Dervan, P.B. and Gottesfeld, J.M. (1997) Importance of minor groove binding zinc fingers within the transcription factor IIIA-DNA complex. *J. Mol. Biol.*, **274**, 439-445.
- Neidle, S. (2001) DNA minor-groove recognition by small molecules. *Nat. Prod. Rep.*, **18**, 291-309.
- Neidle, S., Kelland, L. R., Trent, J. O., Simpson, I. J., Boykin, D. W., Kumar, A. and Wilson, W. D. (1997) Cytotoxicity of Bis(phenylamidine)furan Alkyl Derivatives in Human Tumor Cell Lines: Relation to DNA Minor Groove Binding. *Bioorg. Med. Chem. Lett.*, **7**, 1403-1408.
- Nguyen, B., Tardy, C., Bailly, C., Colson, P., Houssier, C., Kumar, A., Boykin, D.W. and Wilson, W.D. (2002) Influence of compound structure on affinity, sequence selectivity, and mode of binding to DNA for unfused aromatic dications related to furamide. *Biopolymers*, **63**, 281-297.
- Nichols, J., Zevnik, B., Anastassiadis, K., Niwa, H., Klewe-Nebenius, D., Chambers, I., Schöler, H. and Smith, A. (1998) Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4. *Cell*, **95**, 379-391.
- Ogryzko, V.V., Schiltz, R.L., Russanova, V., Howard B.H. and Nakatani, Y. (1996) The transcriptional coactivators p300 and CBP are histone acetyltransferases. *Cell*, **87**, 953-959.
- Oikawa, T. and Yamada, T. (2003) Molecular biology of the Ets family of transcription factors. *Gene*, **303**, 11-34.
- Otting, G., Qian, Y.Q., Billeter, M., Müller, M., Affolter, M., Gehring, W.J. and Wüthrich K. (1990) Protein-DNA contacts in the structure of a homeodomain-DNA complex determined by nuclear magnetic resonance spectroscopy in solution. *EMBO J.*, **9**, 3085-3092.
- Pandolfi, P.P. (2001) Transcription therapy for cancer. *Oncogene*, **20**, 3116-3127.

- Parks, M.E., Baird, E.E. and Dervan, P.B. (1996) Optimization of the hairpin polyamide design for recognition of the minor groove of DNA. *J. Am. Chem. Soc.*, **118**, 6147-6152.
- Patel, D.J. (1982) Antibiotic-DNA interactions: intermolecular nuclear Overhauser effects in the netropsin-d(C-G-C-G-A-A-T-T-C-G-C-G) complex in solution. *Proc. Natl. Acad. Sci. U S A.*, **79**, 6424-6428.
- Pazin, M. J. and Kadonaga, J. T. (1997) SWI2/SNF2 and related proteins : ATP-drive motoes that disrupt protein-DNA interaction? *Cell*, **88**, 737-740.
- Peixoto, P. and Lansiaux, A. (2006) Histone-deacetylase inhibitors : from TSA to SAHA. *Bull. Cancer*, **93**, 27-36.
- Peixoto, P., Liu, Y., Depauw, D., Hildebrand, M.P, Boykin, D.W. Bailly, C., Wilson W.D. and David, M.H. (2008) Direct inhibition of the DNA-binding activity of POU transcription factors Pit-1 and Brn-3 by selective binding of a phenyl-furan-benzimidazole dication. *Nucleic Acids Res.*, 3341-3353.
- Pelham, H.R. (1982) A regulatory upstream promoter element in the Drosophila hsp 70 heat-shock gene. *Cell*, **30**, 517-528.
- Pellegrini-Bouiller, I., Morange-Ramos, I., Barlier, A., Gunz, G., Enjalbert, A. and Jaquet P. (1997) Pit-1 gene expression in human pituitary adenomas. *Horm. Res.*, **47**, 251-258.
- Pellerin, I., Schnabel, C., Catron, K.M. and Abate, C. (1994) Hox proteins have different affinities for a consensus DNA site that correlate with the positions of their genes on the hox cluster. *Mol. Cell. Biol.*, **14**, 4532-4545.
- Pelton, J.G. and Wemmer, D.E. (1989) Structural characterization of a 2:1 distamycin A. d(CGCAAATTGGC) complex by two-dimensional NMR. *Proc. Natl. Acad. Sci. U S A.*, **86**, 5723-5727.
- Phillips, K. and Luisi, B. (1990) The virtuoso of versatility: POU proteins that flex to fit. *J. Mol. Biol.*, **302**, 1023-1039.
- Plate, K.H., Breier, G., Millauer, B., Ullrich, A. and Risau, W. (1993) Up-regulation of vascular endothelial growth factor and its cognate receptors in a rat glioma model of tumor angiogenesis. *Cancer Res.*, **53**, 5822-5827.
- Popovic, R., Erfurth, F. and Zeleznik-Le, N. (2008) Transcriptional complexity of the HOXA9 locus. *Blood. Cells. Mol. Dis.*, **40**, 156-159.
- Privalov, P.L., Dragan, A.I., Crane-Robinson, C., Breslauer, K.J., Remeta, D.P. and Minetti, C.A. (2006) What drives proteins into the major or minor grooves of DNA? *J. Mol. Biol.*, **365**, 1-9.
- Pufall, M.A., Lee, G.M., Nelson, M.L., Kang, H.S., Velyvis, A., Kay, L.E., McIntosh, L.P. and Graves, B.J. (2005) Variable control of Ets-1 DNA binding by multiple phosphates in an unstructured region. *Science*, **309**, 142-145.
- Pugh, B.F. (1996) Mechanisms of transcription complex assembly. *Curr. Opin. Cell. Biol.*, **8**, 303-311.
- Razin, A. (1998) CpG methylation, chromatin structure and gene silencing-a three-way connection. *EMBO J.*, **17**, 4905-4908.

- Roeder, R.G. (1996) The role of general initiation factors in transcription by RNA polymerase II. *Trends Biochem. Sci.*, **21**,327-335.
- Rubinfeld, H. and Seger, R. (2005) The ERK cascade: a prototype of MAPK signaling. *Mol. Biotechnol.*, **31**,151-174.
- Samady, L., Dennis, J., Budhram-Mahadeo, V. and Latchman, D.S. (2004) Activation of CDK4 gene expression in human breast cancer cells by the Brn-3b POU family transcription factor. *Cancer Biol. Ther.*, **3**, 317-323.
- Santoro, A., Maggio, A., Carbone, P., Mirto, S., Caronia, F. and Acuto, S. (1992) Amplification of ETS2 oncogene in acute nonlymphoblastic leukemia with t(6;21;18). *Cancer Genet. Cytogenet.*, **58**, 71-75.
- Satriano, J., Isome, M., Casero, R.A. Jr., Thomson, S.C., Blantz, R.C. (2001) Polyamine transport system mediates agmatine transport in mammalian cells. *Am. J. Physiol. Cell. Physiol.*, **281**, C329-334.
- Sauvageau, G., Lansdorp, P.M., Eaves, C.J., Hogge, D.E., Dragowska, W.H., Reid, D.S., Largman, C., Lawrence, H.J. and Humphries, R.K. (1994) Differential expression of homeobox genes in functionally distinct CD34+ subpopulations of human bone marrow cells. *Proc. Natl. Acad. Sci. U S A.*, **91**, 12223-12227.
- Schiffman, M.H., Bauer, H.M., Hoover, R.N., Glass, A.G., Cadell, D.M., Rush, B.B., Scott, D.R., Sherman, M.E., Kurman, R.J. and Wacholder, S. (1993) Epidemiologic evidence showing that human papillomavirus infection causes most cervical intraepithelial neoplasia. *J. Natl. Cancer Inst.*, **85**, 958-964.
- Schnabel, C.A., Jacobs, Y. and Cleary, M.L. (2000) HoxA9-mediated immortalization of myeloid progenitors requires functional interactions with TALE cofactors Pbx and Meis. *Oncogene*, **19**, 608-616.
- Schonemann, M.D., Ryan, A.K., McEvelly, R.J., O'Connell, S.M., Arias, C.A., Kalla, K.A., Li, P., Sawchenko, P.E. and Rosenfeld, M.G. (1995) Development and survival of the endocrine hypothalamus and posterior pituitary gland requires the neuronal POU domain factor Brn-2. *Genes Dev.*, **9**, 3122-3135.
- Schübeler, D., Lorincz, M.C., Cimbora, D.M., Telling, A., Feng, Y.Q., Bouhassira, E.E. and Groudine, M. (2000) Genomic targeting of methylated DNA : influence of methylation on transcription, replication, chromatin structure, and histone acetylation. *Mol. Cell. Biol.*, **20**, 9103-9112.
- Scott, E.W., Fisher, R.C., Olson, M.C., Kehrli, E.W., Simon, M.C. and Singh, H. (1997) PU.1 functions in a cell-autonomous manner to control the differentiation of multipotential lymphoid-myeloid progenitors. *Immunity*, **6**, 437-647.
- Scully, K.M., Jacobson, E.M., Jepsen, K., Lunyak, V., Viadiu, H., Carrière, C., Rose, D.W., Hooshmand, F., Aggarwal, A.K. and Rosenfeld, M.G. (2000) Allosteric effects of Pit-1 DNA sites on long-term repression in cell type specification. *Science*, **290**, 1127-1131.
- Seeman, N. C., Rosenberg, J. M. and Rich, A. (1976) Sequence-specific recognition of double helical nucleic acids by proteins. *Proc. Natl. Acad. Sci. USA.*, **73**, 804–808.
- Sentenac, A. (1985) Eukaryotic RNA polymerases. *CRC Crit. Rev. Biochem.*,**18**, 31-90.

- Seth, A., Watson, D.K. (2005) ETS transcription factors and their emerging roles in human cancer. *Eur. J. Cancer*, **41**, 2462-2478.
- Seth, A., Watson, D.K., Blair, D.G. and Papas, T.S. (1989) c-ets-2 protooncogene has mitogenic and oncogenic activity. *Proc. Natl. Acad. Sci. U S A.*, **86**, 7833-7837.
- Sharrocks, A.D. (2001) The ETS-domain transcription factor family. *Nat. Rev. Mol. Cell Biol.*, **2**, 827-837.
- Shen, W.F., Rozenfeld, S., Kwong, A., Kömüves, L.G., Lawrence, H.J. and Largman, C. (1999) HOXA9 forms triple complexes with PBX2 and MEIS1 in myeloid cells. *Mol. Cell. Biol.*, **19**, 3051-3061.
- Simpson, L. (1987) The mitochondrial genome of kinetoplastid protozoa: genomic organization, transcription, replication, and evolution. *Annu. Rev. Microbiol.*, **41**, 363-382.
- Skyldberg, B., Salo, S., Eriksson, E., Aspenblad, U., Moberger, B., Tryggvason, K. and Auer G. (1999) Laminin-5 as a marker of invasiveness in cervical lesions. *J. Natl. Cancer Inst.*, **91**, 1882-1887.
- Soini, Y., Määttä, M., Salo, S., Tryggvason, K. and Autio-Harminen, H. (1996) Expression of the laminin gamma 2 chain in pancreatic adenocarcinoma. *J. Pathol.*, **180**, 290-294.
- Sommerville, J. (1984) RNA polymerase I promoters and transcription factors. *Nature*, **310**, 189-190.
- Sordat, I., Bosman, F.T., Dorta, G., Rousselle, P., Aberdam, D., Blum, A.L. and Sordat, B. (1998) Differential expression of laminin-5 subunits and integrin receptors in human colorectal neoplasia. *J. Pathol.*, **185**, 44-52.
- Soulier, J. (1996) Fusion entre une protéine HOX et une nucléoporine dans les leucémie myéloïdes. *M.S. Méd. Sci.*, **12**, 1017-1018.
- Steinfeldt, H.J., Hauser, P., Nakayama, Y., Radovick, S., McClaskey, J.H., Taylor, T., Weintraub, B.D. and Wondisford, F.E. (1991) Thyrotropin-releasing hormone regulation of human TSHB expression: role of a pituitary-specific transcription factor (Pit-1/GHF-1) and potential interaction with a thyroid hormone-inhibitory element. *Proc. Natl. Acad. Sci. U S A.*, **88**, 3130-3134.
- Takeda, A., Goolsby, C. and Yaseen, N.R. (2006) NUP98-HOXA9 induces long-term proliferation and blocks differentiation of primary human CD34+ hematopoietic cells. *Cancer Res.* **66**, 6628-6637.
- Taketani, T., Taki, T., Ono, R., Kobayashi, Y., Ida, K. and Hayashi, Y. (2002a) The chromosome translocation t(7;11)(p15;p15) in acute myeloid leukemia results in fusion of the NUP98 gene with a HOXA cluster gene, HOXA13, but not HOXA9. *Genes Chromosomes Cancer*, **34**, 437-443.
- Taketani, T., Taki, T., Shibuya, N., Ito, E., Kitazawa, J., Terui, K. and Hayashi, Y. (2002b) The HOXD11 gene is fused to the NUP98 gene in acute myeloid leukemia with t(2;11)(q31;p15). *Cancer Res.*, **62**, 33-37.
- Tanaka, M. and Herr W. (1990) Differential transcriptional activation by Oct-1 and Oct-2: interdependent activation domains induce Oct-2 phosphorylation. *Cell*, **60**, 375-386.
- Tanious, F., Wilson, W.D., Wang, L., Kumar, A., Boykin, D.W., Marty, C., Baldeyrou, B. and Bailly, C. (2003) Cooperative dimerization of a heterocyclic diamidine determines sequence-specific DNA recognition. *Biochemistry*, **42**, 13576-13586.

- Tanious, F.A., Laine, W., Peixoto, P., Bailly, C., Goodwin, K.D., Lewis, M.A., Long, E.C., Georgiadis, M.M., Tidwell, R.R. and Wilson, W.D. (2007) Unusually strong binding to the DNA minor groove by a highly twisted benzimidazole diphenylether: induced fit and bound water. *Biochemistry*, **46**, 6944-6956.
- Tansey, W.P. and Herr, W. (1997) TAFs: guilt by association ? *Cell*, **88**, 729-732.
- Thorsteinsdottir, U., Kroon, E., Jerome, L., Blasi, F. and Sauvageau, G. (2001) Defining roles for HOX and MEIS1 genes in induction of acute myeloid leukemia. *Mol. Cell. Biol.*, **21**, 224-234.
- Timpl, R. (1989) Structure and biological activity of basement membrane proteins. *Eur. J. Biochem.*, **180**, 487-502.
- Tomlins, S.A., Laxman, B., Varambally, S., Cao, X., Yu, J., Helgeson, B.E., Cao, Q., Prensner, J.R., Rubin, M.A., Shah, R.B., Mehra, R. and Chinnaiyan, A.M. (2008) Role of the TMPRSS2-ERG gene fusion in prostate cancer. *Neoplasia*, **10**, 177-188.
- Treisman, R. (1994) Ternary complex factors: growth factor regulated transcriptional activators. *Curr. Opin. Genet. Dev.*, **4**, 96-101.
- Trent, J.O., Clark, G.R., Kumar, A., Wilson, W.D., Boykin, D.W., Hall, J.E., Tidwell, R.R., Blagburn, B.L., Neidle, S. (1996) Targeting the minor groove of DNA: crystal structures of two complexes between furan derivatives of berenil and the DNA dodecamer d(CGCGAATTCGCG)₂. *J. Med. Chem.*, **39**, 4554-4562.
- Trivedi, C.M., Patel, R.C. and Patel, C.V. (2007) Homeobox gene HOXA9 inhibits nuclear factor-kappa B dependent activation of endothelium. *Atherosclerosis*, **195**, 50-60.
- Van Dyke, M.W., Hertzberg, R.P. and Dervan, P.B. (1982). Map of distamycin, netropsin, and actinomycin binding sites on heterogeneous DNA: DNA cleavage-inhibition patterns with methidiumpropyl-EDTA.Fe(II). *Proc. Natl. Acad. Sci. U S A.*, **79**, 5470-5474.
- Verrijzer, C.P., Kal, A.J. and van der Vliet, P.C. (1990) The oct-1 homeo domain contacts only part of the octamer sequence and full oct-1 DNA-binding activity requires the POU-specific domain. *Genes Dev.*, **4**, 1964-1974.
- Vijapurkar, U., Fischbach, N., Shen, W., Brandts, C., Stokoe, D., Lawrence, H.J. and Largman, C. (2004) Protein kinase C-mediated phosphorylation of the leukemia-associated HOXA9 protein impairs its DNA binding ability and induces myeloid differentiation. *Mol. Cell. Biol.*, **24**, 3827-3837.
- Voss, J.W., Wilson, L. and Rosenfeld, M.G. (1991) POU-domain proteins Pit-1 and Oct-1 interact to form a heteromeric complex and can cooperate to induce expression of the prolactin promoter. *Genes Dev.*, **5**, 1309-1320.
- Wang, G.G., Pasillas, M.P. and Kamps, M.P. (2005) Meis1 programs transcription of FLT3 and cancer stem cell character, using a mechanism that requires interaction with Pbx and a novel function of the Meis1 C-terminus. *Blood*, **106**, 254-264.
- Wang, L., Bailly, C., Kumar, A., Ding, D., Bajic, M., Boykin, D.W. and Wilson, W.D. (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. U S A.*, **97**, 12-16.

- Wang, L., Carrasco, C., Kumar, A., Stephens, C.E., Bailly, C., Boykin, D.W. and Wilson, W.D. (2001) Evaluation of the influence of compound structure on stacked-dimer formation in the DNA minor groove. *Biochemistry*, **40**, 2511-2521.
- Wang, N., Kim, H.G., Cotta, C.V., Wan, M., Tang, Y., Klug, C.A. and Cao, X (2006) TGFbeta/BMP inhibits the bone marrow transformation capability of Hoxa9 by repressing its DNA-binding ability. *EMBO J.*, **25**, 1469-1480.
- Watabe, T., Yoshida, K., Shindoh, M., Kaya, M., Fujikawa, K., Sato, H., Seiki, M., Ishii, S., Fujinaga, K. (1998) The Ets-1 and Ets-2 transcription factors activate the promoters for invasion-associated urokinase and collagenase genes in response to epidermal growth factor. *Int. J. Cancer*, **77**, 128-137.
- Wemmer, D.E. (2000) Designed sequence-specific minor groove ligands. *Ann. Rev. Biomol. Struct.*, **29**, 439-461.
- Westermarck, J. and Kähäri, V.M. (1999) Regulation of matrix metalloproteinase expression in tumor invasion. *FASEB J.*, **13**, 781-792.
- Whartonn R.P., Brown E.L. and Ptashnen M. (1984) Substituting an alpha-helix switches the sequence-specific DNA interactions of a repressor. *Cell*, **38**, 361-369.
- White, S., Szewczyk, J.W., Turner, J.M., Baird E.E. and Dervan P.B. (1998) Recognition of the four Watson–Crick base pairs in the DNA minor groove by synthetic ligands. *Nature*, **391**, 468-471.
- 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.
- Wilson, W.D., Tanious, F.A., Ding, D., Kumar, A., Boykin, D.W., Colson, P., Houssier, C. and Bailly, C. (1998) Nucleic Acid Interactions of Unfused Aromatic Cations: Evaluation of Proposed Minor-Groove, Major-Groove, and Intercalation Binding Modes. *J. Am. Chem. Soc.*, **120**, 10310-10321.
- Wong, S. and Witte, O.N. (2004) The BCR-ABL story: bench to bedside and back. *Annu. Rev. Immunol.*, **22**, 247-306.
- Wu, C., (1997) Chromatin remodelling and the control of gene expression. *J. Biol. Chem.*, **272**, 28171-28174.
- Wurtz, N.R., Pomerantz, J.L., Baltimore, D. and Dervan, P.B. (2002) Inhibition of DNA binding by NF-kappa B with pyrrole-imidazole polyamides. *Biochemistry*, **41**, 7604-7609.
- Xia, Y.R., Andersen, B., Mehrabian, M., Diep, A.T., Warden, C.H., Mohandas, T., McEvelly, R.J., Rosenfeld, M.G. and Lusic, A.J. (1993) Chromosomal organization of mammalian POU domain factors. *Genomics*, **18**, 126-130.
- Xu, D., Wilson, T.J., Chan, D., De Luca, E., Zhou, J., Hertzog, P.J. and Kola, I. (2002) Ets1 is required for p53 transcriptional activity in UV-induced apoptosis in embryonic stem cells. *EMBO J.*, **21**, 4081-4093.
- Xu, L., Lavinsky, R.M., Dasen, J.S., Flynn, S.E., McInerney, E.M., Mullen, T.M., Heinzl, T., Szeto, D., Korzus, E., Kurokawa, R., Aggarwal, A.K., Rose, D.W., Glass, C.K. and Rosenfeld, M.G. (1998) Signal-specific co-activator domain requirements for Pit-1 activation. *Nature*, **395**, 301-306.

- Yang, S.H. and Sharrocks, A.D. (2004) SUMO Promotes HDAC-Mediated Transcriptional Repression. *Mol. Cell.*, **13**, 611-617.
- Yordy, J.S. and Muise-Helmericks, R.C. (2000) Signal transduction and the Ets family of transcription factors. *Oncogene*, **19**, 6503-6513.
- Yu, B.D., Hess, J.L., Horning, S.E., Brown, G.A. and Korsmeyer, S.J. (1995) Altered Hox expression and segmental identity in Mll-mutant mice. *Nature* **378**, 505–508.
- Yu, C.L., Meyer, D.J., Campbell, G.S., Lerner, A.C., Carter-Su, C., Schwartz, J. and Jove, R. (1995) Enhanced DNA-binding activity of a Stat3-related protein in cells transformed by the Src oncoprotein. *Science*, **269**, 81-8.
- Zakrzewska, K., Lavery, R. and Pullman, B.A. (1987) Theoretical study of the sequence specificity in binding of lexitropsins to B-DNA. *J. Biomol. Struct. Dyn.*, **4**, 833-843.
- Zeisig, B.B., Milne, T., García-Cuéllar, M.P., Schreiner, S., Martin, M.E., Fuchs, U., Borkhardt, A., Chanda, S.K., Walker, J., Soden, R., Hess, J.L. and Slany, R.K. (2004) Hoxa9 and Meis1 are key targets for MLL-ENL-mediated cellular immortalization. *Mol. Cell. Biol.*, **24**, 617-628.
- Zhao, J. (2007) Sumoylation regulates diverse biological processes. *Cell Mol Life Sci.*, **64**, 3017-3033.
- Zhou, H., Yoshioka, T. and Nathans, J. (1996) Retina-derived POU-domain factor-1: a complex POU-domain gene implicated in the development of retinal ganglion and amacrine cells. *J. Neurosci.*, **16**, 2261-2274.
- Zhurkin, V.B., Ulyanov, N.B., Gorin, A.A. and Jernigan, R.L. (1991) Static and statistical bending of DNA evaluated by Monte Carlo simulations. *Proc. Natl. Acad. Sci. U S A.*, **88**, 7046-7050.

