



Etude structurale et fonctionnelle du complexe ZEBRA / ADN méthylé

Priscilla Pagniez

► To cite this version:

Priscilla Pagniez. Etude structurale et fonctionnelle du complexe ZEBRA / ADN méthylé. Biochimie [q-bio.BM]. Université Joseph-Fourier - Grenoble I, 2008. Français. NNT: . tel-00356802

HAL Id: tel-00356802

<https://theses.hal.science/tel-00356802>

Submitted on 28 Jan 2009

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 JOSEPH FOURIER - GRENOBLE 1
ECOLE DOCTORALE CHIMIE ET SCIENCES DU VIVANT

Année 2008

N° attribué par la bibliothèque

____/____/____/____/____/____/____/____/____/____

THÈSE

pour l'obtention du grade de Docteur de l'Université Joseph Fourier
Discipline : Biologie Structurale et Nanobiologies

Sujet de la thèse :

Etude structurale et fonctionnelle du complexe

ZEBRA / ADN méthylé

présentée par

Priscilla PAGNIEZ

Thèse soutenue publiquement le 24 octobre 2008

devant le jury composé de :

Mr Patrice MORAND
Mme Irène JOAB
Mr Henri GRUFFAT
Mme Andrea DESSEN
Mr Carlo PETOSA

Directeur de thèse
Rapporteur
Rapporteur
Examineur
Invité

Remerciements...

Je tiens d'abord à remercier Stephen CUSACK, Rob RUIGROK et Christoph MULLER pour m'avoir accueillie dans leurs groupes de recherche tout d'abord à l'EMBL puis à l'UVHCI.

Mes remerciements vont aussi aux membres de mon jury de thèse, Irène JOAB, Andrea DESSEN et Henri GRUFFAT. Merci de l'intérêt qu'ils ont porté à ce travail ainsi que pour leurs commentaires.

Merci à mon directeur de thèse, Patrice MORAND, pour les nombreux encouragements et conseils prodigués durant ces trois années de thèse.

J'exprime ma profonde reconnaissance à Carlo PETOSA. Son intérêt dans ce projet et ses encouragements ont su replacer mes doutes dans un contexte plus optimiste. Merci pour la confiance qu'il m'a accordée et pour son ouverture d'esprit. Merci pour son calme, sa patience et sa fidélité tout au long de ces trois années de thèse. Enfin, merci pour sa disponibilité, ses conseils précieux lors de la rédaction de mon manuscrit et pour les heures qu'il a passé à relire cette thèse et à la corriger.

J'exprime toute ma reconnaissance à Monique PERISSIN qui a guidé mes premiers pas en cristallogénèse. Son départ en retraite est survenu trop tôt à mon goût mais cette thèse aurait été bien différente sans elle.

Ce travail a été effectué en collaboration avec François PARCY et Cécile HAMES de l'IRTSV du CEA. Je les remercie sincèrement pour leur disponibilité et leur enthousiasme à l'égard de notre collaboration. Merci pour le temps qu'ils m'ont accordé, leurs conseils avisés et leurs connaissances.

Je remercie particulièrement Marc JAMIN pour l'intérêt et l'aide précieuse qu'il a apporté à mon travail sur les mesures d'anisotropie de fluorescence.

J'exprime ma plus profonde reconnaissance à l'ensemble du personnel de l'UVHCI et de l'antenne grenobloise de l'EMBL, notamment Virginie, pour sa bonne humeur et sa patience, et les heures de discussion passées ensemble...(!!), Benjamin, pour avoir répondu à tous mes appels au secours suite à mes (trop nombreuses) gaffes..., et tous les autres (la liste serait trop longue) pour tous les petits (et grands) services quotidiens rendus (qui, mis bout à bout, représentent plusieurs mois de travail économisés !!),

Comment ne pas introduire ici ceux qui ont ponctué cette année à coup d'onomatopées, d'éclats de rire et de moments forts : Francine, Cécile, Majida, Julien P.,

Julien L., Pierre-Alain, Alex, Lucy, Euripidès ...À tous ceux-là, je dis merci pour ces agréables moments d'aide et de soutien bienvenus et leur bonne humeur communicative.

Durant une thèse la vie continue, donc merci à tous ceux qui m'ont soutenu en dehors du laboratoire, et particulièrement mes parents : leur soutien, leurs encouragements, leur écoute, leur foi en moi m'ont permis d'aller toujours plus loin, sans jamais me décourager.

*Il est souvent nécessaire d'entreprendre pour espérer
et de persévérer pour réussir.*

Gilbert Cesbron

Table des matières

CHAPITRE I : INTRODUCTION..... 12

I. GENERALITES SUR LE VIRUS D'EPSTEIN-BARR	13
1. <i>L'EBV est un herpes-virus</i>	13
2. <i>Structure du virion</i>	14
3. <i>Génome de l'EBV</i>	15
4. <i>Cycle viral</i>	16
5. <i>Transmission du virus et pathologies associées</i>	18
II. LA PROTEINE ZEBRA :	24
1. <i>Historique de sa découverte</i>	24
2. <i>Les variants de ZEBRA</i>	28
III. ZEBRA ET SES MULTIPLES FONCTIONS... ..	30
1. <i>Les protéines bZIPs</i>	30
2. <i>La régulation de l'expression de ZEBRA</i>	31
3. <i>Les fonctions de trans-activation de ZEBRA</i>	33
4. <i>Les fonctions de réplication de ZEBRA</i>	36
5. <i>ZEBRA et le contrôle du cycle cellulaire</i>	36
6. <i>Les interactions ZEBRA-protéines cellulaires</i>	37
IV. STRUCTURE CRISTALLOGRAPHIQUE DE ZEBRA EN COMPLEXE AVEC LE SITE AP1	41
1. <i>Organisation en domaines de ZEBRA</i>	41
2. <i>Analyse du complexe ZEBRA/AP1 et études de mutagenèse de ZEBRA</i>	42
V. ACTIVATION DES PROMOTEURS METHYLES PAR ZEBRA	49
1. <i>Conséquences fonctionnelles de la méthylation de l'ADN</i>	49
2. <i>Reconnaissance préférentielle de promoteurs méthylés par ZEBRA</i>	52

OBJECTIFS DU TRAVAIL DE THESE **ERREUR ! SIGNET NON DEFINI.5**

CHAPITRE II : PREPARATION ET CRISTALLISATION DE ZEBRA EN COMPLEXE AVEC UN SITE ZRE2 METHYLE..... 56

I. STRATEGIE DE CRISTALLISATION DU COMPLEXE ZEBRA/ZRE2 METHYLE.....	57
1. <i>Choix du clone de ZEBRA</i>	57
2. <i>Choix des oligonucléotides contenant le site ZRE2</i>	58
II. PURIFICATION DU COMPLEXE ZEBRA/ZRE2 METHYLE	59
1. <i>Purification de ZEBRA</i>	59

2. Purification des oligonucléotides contenant le site ZRE2	62
III. CRISTALLISATION DU COMPLEXE ZEBRA/ZRE2 METHYLE	63

CHAPITRE III : RESOLUTION DE LA STRUCTURE DU COMPLEXE

ZEBRA/ZRE2 METHYLE 68

I. COLLECTE DES DONNEES DE DIFFRACTION.....	70
II. TRAITEMENT DES DONNEES DE DIFFRACTION	72
III. AFFINEMENT ET VALIDATION	80

CHAPITRE IV : ANALYSE DE LA STRUCTURE DU COMPLEXE ZEBRA/ZRE2

METHYLE 82

I. LA PROTEINE ZEBRA.....	83
II. LA DOUBLE HELICE D'ADN	86
III. ANALYSE DE L'INTERACTION ENTRE ZEBRA ET LE SITE ZRE2 METHYLE	93

CHAPITRE V : ANALYSES DU MECANISME DE RECONNAISSANCE

PREFERENTIELLE DE L'ADN METHYLE PAR ZEBRA..... 95

I. PRINCIPE DE LA TECHNIQUE D'ANISOTROPIE DE FLUORESCENCE	97
II. MISE AU POINT DES EXPERIENCES D'ANISOTROPIE DE FLUORESCENCE	98
1. Choix des clones de ZEBRA et cJun et purification des protéines	98
2. Choix des oligonucléotides.....	101
III. AFFINITE DE ZEBRA POUR LES MUTANTS DU SITE ZRE2 METHYLE	103
1. Reconnaissance différentielle du groupement méthyle.....	103
2. Reconnaissance différentielle de la paire de bases -3/3' flanquant le site CpG méthylé.....	107
3. Reconnaissance indépendante du site meZRE2 par la sérine 186 et l'asparagine 182.....	109
4. Discussion.....	111
IV. IDENTIFICATION DES RESIDUS DE ZEBRA SUFFISANTS A UNE FIXATION PREFERENTIELLE DU SITE ZRE2 METHYLE.....	116
1. cJun ne fait aucune distinction entre les sites meZRE2 et ZRE2	117
2. Le domaine de liaison à l'ADN de ZEBRA permet à cJun de reconnaître préférentiellement le site meZRE2	118
3. La sérine 186 de ZEBRA est suffisante pour que cJun distingue les sites meZRE2 et ZRE2.....	119
4. Discussion.....	120
V. HYPOTHESES SUR LE MECANISME DE RECONNAISSANCE DE L'ADN METHYLE PAR ZEBRA.....	121

CHAPITRE VI : CRIBLAGE CHIMIQUE A HAUT DEBIT D'INHIBITEURS DE ZEBRA	122
I. ZEBRA : UNE CIBLE IDEALE	124
II. PRINCIPE DE CRIBLAGE CHIMIQUE A HAUT DEBIT	124
III. RESULTATS PRELIMINAIRES	124
IV. CONCLUSIONS	130
 CHAPITRE VII : CONCLUSION GENERALE ET PERSPECTIVES	 131
 CHAPITRE VIII : MATERIELS ET METHODES	 136
I. MATERIELS	137
1. <i>Clonage</i>	137
2. <i>Expression des protéines</i>	137
II. METHODES	137
1. <i>Construction des plasmides d'expression en bactéries</i>	137
1.1. Mutagénèse dirigée : clones ZEBRA C189S ¹⁷⁵⁻²³⁶ – pET28a, cJun A275S/C278S ²⁶⁴⁻³²⁴ – pETM40	137
1.2. Clones ZEBRA C189S ¹⁷⁵⁻²³⁶ – pETM40 et cJun C278S ²⁶⁴⁻³²⁴ – pETM40.....	139
1.3. Clone Chimère ZEBRA-cJun – pETM40.....	140
2. <i>Expression et purification des protéines</i>	140
2.1. Expression des protéines chez <i>E.coli</i>	140
2.2. Protocole de purification de la protéine ZEBRA	141
2.3. Protocole de purification des protéines de fusion-MBP (clones ZEBRA C189S ¹⁷⁵⁻²³⁶ – pETM40, cJun C278S ²⁶⁴⁻³²⁴ – pETM40, cJunA275S/C278S ²⁶⁴⁻³²⁴ – pETM40, Chimère ZEBRA-cJun – pETM40)	144
3. <i>Purification et hybridation des oligonucléotides simple brin</i>	145
4. <i>Cristallogénèse</i>	147
4.1. Principe de la cristallogénèse.....	147
4.2. Mise en œuvre	148
4.3. Cristallogénèse du complexe ZEBRA/meZRE2	149
5. <i>Résolution de la structure du complexe ZEBRA/meZRE2 par cristallographie aux rayons X</i>	150
5.1. Les rayons X et les cristaux de protéines.....	150
5.2. Acquisition et traitement des données	150
6. <i>Mesures d'affinité de l'interaction ZEBRA/ADN par anisotropie de fluorescence (AF)</i>	155
6.1. La technique d'anisotropie de fluorescence.....	155
6.2. Méthodes de détermination des constantes de dissociation	156
6.3. Protocole de mesure d'anisotropie de fluorescence	158
 ANNEXES.....	 159

REFERENCES BIBLIOGRAPHIQUES	180
--	------------

ABBREVIATIONS

A : adénine

Å : Ångstroms (10^{-10} m)

aa : acide aminé

ADN : acide désoxyribonucléique

AF ou A_f : anisotropie de fluorescence

AP1 : « *Activator protein 1* »

APS : Ammonium persulfate

BJAB : cellules humaines EBV-négatives

BL : lymphome de Burkitt

bp : « *base pairs* »

BR : région basique

C : cytosine

CT : C-terminal

DEL : diode électroluminescente

Dnmt : ADN méthyltransférases

DO : densité optique

DTT : dithiothréitol

E.coli : *Escherichia coli*

EA : « *early antigen* »

EBV : Epstein-Barr virus

EMBL : « *European Molecular Biology Laboratory* »

EMSA : « *electrophoretic mobility shift assay* » (ou gel retard)

G : guanine

HDAC : histone-désacétylase

het DNA : « *heterogeneous DNA* »

IPTG : isopropyl-β-D-thio-galactopyranoside

kDa : kilo Dalton

LB : Luria-Bertiani

M : molaire (1M = 1 mole/L)

MBD : « *Methyl-Binding-Domain Protein* »

MBP : « *maltose binding protein* »

meZRE2 : ZRE2 méthylé

MNI : mononucléose infectieuse

ORF : « *open reading frame* »

PCR : « *Polymerase Chain Reaction* »

PEG : polyéthylène glycol

Pfu : *Pyrococcus furiosus* (organisme qui synthétise cette polymérase)

PMSF : Phénylméthylsulfonyl fluoride

rmsd : « *Root Mean Square Deviation* »

rpm : rotation par minute

SDS-PAGE : « *Sodium dodécyl sulfate polyacrylamide gel electrophoresis* »

T : thymine

TAD : « *transcriptional activation domain* »

TAMRA : « *tetramethylrhodamine molecule* »

TEMED : Tétraméthylethylènediamine

T_m : « *melting temperature* »

TPA : 12-O-tétradécanoylphorbol-13-acétate

Tris : Tris-(hydrométhyl)aminométhane

u.a. : unité arbitraire

WT : « *wild type* »

ZEBRA : « *BamHI Z Epstein-Barr Replication Activator* »

ZIP : région de dimérisation *leucine-zipper*

ZRE : « *ZEBRA response element* »

Liste des figures

Figure I.1 : Morphologie et organisation structurale des herpes-virus	Page 15
Figure I.2. : Physiopathologie de l'infection EBV	Page 20
Figure I.3 : Enfant atteint par un lymphome de Burkitt	Page 21
Figure I.4 : Immunoblot des polypeptides antigènes de l'EBV induits après transfection du fragment BamHI-WZ <i>het</i>	Page 25
Figure I.5 : Carte génomique du fragment BamHI-WZ <i>het</i>	Page 26
Figure I.6 : Mobilité électrophorétique de la protéine ZEBRA (<i>BamHI-Z Epstein-Barr replication activator</i>) du virus d'Epstein-Barr	Page 27
Figure I.7 : Représentation schématique de la structure en hélice d'une protéine à leucine zipper	Page 30
Figure I.8: Organisation des sites de fixation en <i>cis</i> des facteurs de transcription cellulaires sur le promoteur Zp du gène BZLF1	Page 31
Figure I.9 : Schéma du promoteur viral du gène BRLF1	Page 34
Figure I.10 : Représentation schématique de l'organisation en domaines de ZEBRA	Page 41
Figure I.11 : Vues orthogonales du complexe ZEBRA/AP1	Page 42
Figure I.12 : Structure du complexe ZEBRA/AP1 : comparaison avec l'hétérodimère bZIP Fos-Jun par alignement structural	Page 43
Figure I.13 : Alignement de séquences et structure secondaire	Page 44
Figure I.14 : Structure du complexe ZEBRA/ADN : mise en évidence des résidus d'acides aminés présents à l'interface des deux monomères	Page 45
Figure I.15 : Résidus hydrophobes à l'interface dimérique de ZEBRA	Page 46
Figure I.16 : Contacts protéine/ADN du complexe ZEBRA/AP1 au niveau du sillon majeur d'ADN	Page 47
Figure I.17 : Mécanisme par lequel la méthylation de l'ADN réprimerait l'expression des gènes	Page 50
Figure I.18 : Réaction de transfert d'un groupement méthyle sur l'ADN	Page 51
Figure I.19 : ZEBRA fixe préférentiellement le site ZRE2 dans son état méthylé	Page 52
Figure I.20 : Prédiction de structure du complexe ZEBRA/meZRE2	Page 54
Figure II.1: Séquence en acides aminés du clone ZEBRA	Page 57
Figure II.2 : Purification de ZEBRA	Page 60
Figure II.3 : Dernière étape de purification de ZEBRA par chromatographie d'exclusion de taille sur colonne S75	Page 61
Figure II.4 : Purification des oligonucléotides simple brin	Page 63
Figure II.5 : Analyse du complexe ZEBRA/ADN par gel retard	Page 65
Figure II.6 : Images des cristaux obtenus	Page 66

Figure III.1 : Clichés de diffraction à $\Phi=0^\circ$ et $\Phi=90^\circ$	Page 70
Figure III.2 : Cartes de Patterson	Page 74
Figure III.3 : Analyse de l'empilement cristallin du complexe ZEBRA/meZRE2	Page 77
Figure III.4 : Vues orthogonales de la carte de densité électronique 2Fo-Fc du complexe ZEBRA/meZRE2	Page 78
Figure III.5 : Carte Fo-Fc de la 1ère solution proposée par PHASER calculée à partir du modèle réduit utilisé pour le remplacement moléculaire	Page 79
Figure III.6 : Alignement des structures des 2 molécules du complexe ZEBRA/meZRE2 présentes dans l'unité asymétrique	Page 81
Figure IV.1 : Séquences oligonucléotidiques des heptamères AP1 et meZRE2	Page 83
Figure IV.2 : Structures des bases cytosine, cytosine méthylée et thymine	Page 83
Figure IV.3 : Superposition des complexes ZEBRA/AP1 et ZEBRA/meZRE2 : alignement global	Page 84
Figure IV.4 : Superposition des complexes ZEBRA/AP1 et ZEBRA/meZRE2 : alignement selon les doubles hélices d'ADN	Page 85
Figure IV.5 : Superposition des complexes ZEBRA/AP1 et ZEBRA/meZRE2 : alignement selon les doubles hélices d'ADN	Page 87
Figure IV.6 : Comparaison des largeurs des sillons majeur et mineur des doubles hélices d'ADN des sites AP1 et meZRE2	Page 88
Figure IV.7 : Comparaison des largeurs des sillons majeur et mineur de l'oligonucléotide X non méthylé ou méthylé sur une cytosine	Page 89
Figure IV.8 : Contacts ZEBRA/ADN au niveau du sillon majeur	Page 91
Figure IV.9 : Schématisation des interactions ZEBRA/meZRE2 au niveau des résidus N182, S186 et R190 des hélices gauche et droite du dimère de ZEBRA	Page 93
Figure V.1 : Schéma représentant le principe d'anisotropie de fluorescence	Page 98
Figure V.2 : Purification de MBP-ZEBRA par chromatographie d'affinité sur résine d'amylose	Page 99
Figure V.3 : Dernière étape de purification de MBP-ZEBRA par chromatographie d'exclusion de taille sur colonne S200	Page 100
Figure V.4 : Courbes de fixation de ZEBRA sur les oligonucléotides meZRE2 et ZRE2	Page 104
Figure V.5 : Courbes de fixation de ZEBRA sur les oligonucléotides hemi1-ZRE2 et hemi2-ZRE2	Page 105
Figure V.6 : Courbes de fixation de ZEBRA sur les oligonucléotides meZRE2, ZRE2, hemi1-ZRE2 et hemi2-ZRE2	Page 106
Figure V.7 : Courbes de fixation de ZEBRA sur les oligonucléotides AP1, AP1(mut3), meZRE2 et meZRE2(mut3)	Page 108
Figure V.8 : (A) Courbes de fixation de ZEBRA sur les oligonucléotides meZRE2, AP1/meZRE2 et meZRE2/AP1	Page 110
Figure V.9 : Vues en stereo (« cross-eyed stereo ») de l'interaction entre la sérine 186 et les cytosines méthylées C-2 et C1' du site meZRE2	Page 113

Figure V.10 : Agrandissement de la structure du complexe ZEBRA/AP1 représentant les interactions protéine/ADN au niveau du sillon majeur de l'ADN	Page 114
Figure V.11 : Agrandissement de la structure du complexe ZEBRA/meZRE2 représentant les interactions protéine/ADN au niveau du sillon majeur de l'ADN	Page 115
Figure V.12 : Séquence en acides aminés des protéines ZEBRA, cJun et de la chimère ZEBRA-Jun	Page 116
Figure V.13 : Courbes de fixation de cJun sur les oligonucléotides meZRE2 et ZRE2	Page 117
Figure V.14 : Courbes de fixation du clone chimère ZEBRA-cJun sur les oligonucléotides meZRE2 et ZRE2	Page 118
Figure V.15 : Courbes de fixation de cJun A275S sur les oligonucléotides meZRE2 et ZRE2	Page 119
Figure VI.1 : Résultats préliminaires de l'inhibition de liaison à l'ADN de ZEBRA	Page 126
Figure VI.2 : Caractéristiques d'inhibition des 3 composés n°102320 (A), 100025 (B) et 100035 (C)	Page 130
Figure VIII.1 : Schéma résumant les différentes étapes de purification de la protéine ZEBRA	Page 141
Figure VIII.2 : Schéma résumant les différentes étapes de purification des protéines de fusion MBP	Page 144
Figure VIII.3 : Schéma résumant les étapes successives de purification des oligonucléotides utilisés en cristallogénèse	Page 146
Figure VIII.4 : Diagramme de phases décrivant la cristallisation d'un complexe protéique par diffusion de vapeur	Page 147
Figure VIII.5 : Mise en œuvre de la technique de cristallisation par diffusion de vapeur en goutte suspendue	Page 149

LISTE DES TABLES

Table I.1 : Liste des différents types d'herpes-virus	Page 13
Table I.2. : Principales caractéristiques du cycle viral, latent et lytique (in vitro et in vivo)	Page 18
Table I.3 : Représentation des séquences nucléotidiques des 3 heptamères ZREs du promoteur Rp ainsi que du site AP1	Page 35
Table I.4 : Résumé des interactions ZEBRA-protéines cellulaires	Page 37
Table II.1 : récapitulatif des différents oligonucléotides testés pour la cristallisation du complexe ZEBRA/ADN	Page 58
Table II.2 : Résumé des conditions de cristallisation des complexes ZEBRA/meZRE2(a) et ZEBRA/ZRE2(a)	Page 64
Table III.1 : Résumé des statistiques de traitement des données de diffraction et de l'affinement du modèle	Page 71
Table III.2 : Comparaison des paramètres de maille des complexes ZEBRA/ADN	Page 72
Table III.3 : Table résumant les statistiques de remplacement moléculaire des 4 premières solutions proposées par le programme PHASER	Page 75
Tableau V.1: Récapitulatif des oligonucléotides utilisés lors des expériences d'anisotropie de fluorescence	Page 102
Tableau V.2 : Récapitulatif des constantes de dissociation (Kd) des complexes ZEBRA/ADN en fonction de l'oligonucléotide testé	Page 112
Tableau V.3 : Récapitulatif des valeurs de Kd des mutants de cJun pour les sites meZRE2 et ZRE2	Page 120

CHAPITRE I :

INTRODUCTION

I. Généralités sur le virus d'Epstein-Barr

1. L'EBV est un herpes-virus

Les herpes-virus constituent une grande famille de plus de 100 virus qui infectent les vertébrés, en général les mammifères et les oiseaux, mais également les amphibiens, les reptiles et les poissons (Strauss and Strauss, 2002). Chaque herpes-virus n'infecte qu'une espèce de vertébrés dans la nature. Les herpes-virus sont classés en trois groupes : alpha, beta et gamma, en fonction de leurs propriétés biologiques ainsi que des informations disponibles relatives à leurs séquences d'ADN (table I.1.) :

Herpes-virus	Non commun	Sous-famille	Taille du génome (Kbp)	Siège de la latence
HHV-1	Herpes simplex virus 1	Alpha	152	Neurones
HHV-2	Herpes simplex virus 2	Alpha	152	Neurones
HHV-3	Varicella-zoster virus	Alpha	125	Neurones
HHV-4	Epstein-Barr virus	Gamma	172	Lymphocytes B
HHV-5	Cytomégalovirus	Bêta	248	Cellules hématopoïétiques
HHV-6A	Roseolovirus	Bêta	159	Monocytes
HHV-6B	Roseolovirus	Bêta	162	Monocytes
HHV-7		Bêta	145	Lymphocytes T (CD4 ⁺)
HHV-8	KSHV	Gamma	170	Cellules épithéliales, lymphocytes

Table I.1. : Liste des différents types d'herpes-virus.

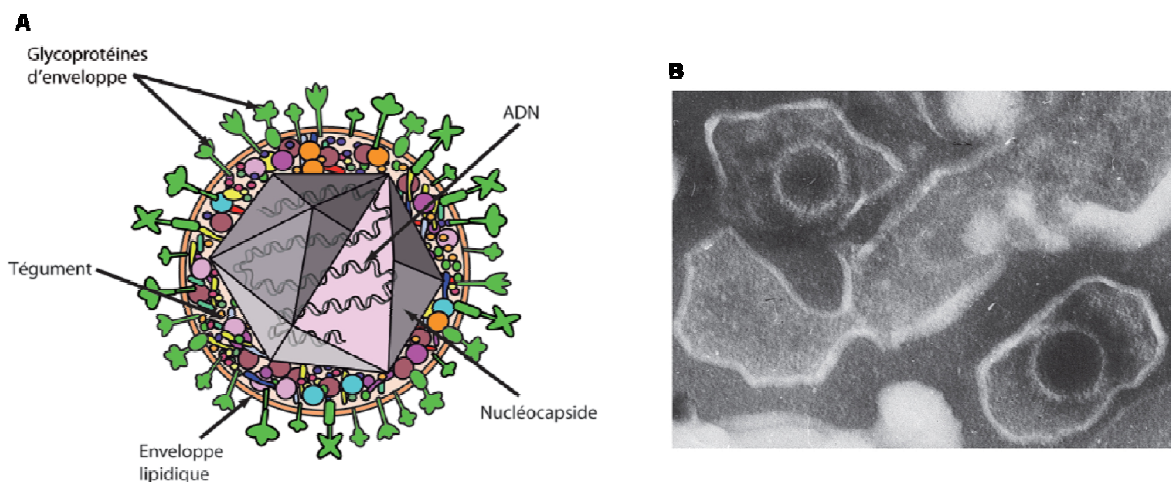
Les alpha herpes-virus infectent un grand nombre d'animaux et de types cellulaires. En culture, ils se propagent rapidement et lysent efficacement les cellules infectées. Chez leurs hôtes naturels, ils établissent une infection latente des neurones sensoriels et l'infection lytique a lieu, en général au niveau des cellules de l'épiderme. Chez l'Homme, cette famille est représentée par le virus Herpes simplex et par Varicella zoster (Strauss and Strauss, 2002).

Les beta herpes-virus infectent un nombre restreint d'hôtes et leur cycle d'infection en culture est long. Chez son hôte naturel, le virus persiste de façon latente, notamment dans les cellules mononuclées sanguines. Le cytomégalovirus est un des représentants humains de cette famille (Strauss and Strauss, 2002).

Les gamma herpes-virus sont les herpes-virus qui possèdent le nombre d'hôtes le plus restreint. Ils se répliquent dans les cellules lymphoïdes et certains peuvent infecter de façon lytique les cellules épithéliales ou endothéliales. Le virus d'Epstein-Barr fait partie de cette sous-famille des *Gammaherpesvirinae* et du genre *Lymphocryptovirus*. Il a été mis en évidence par microscopie électronique en 1964 par Epstein, Achong et Barr (Epstein et al., 1964), dans une lignée cellulaire obtenue à partir d'un lymphome de Burkitt, ce qui fait de l'EBV le premier virus humain à avoir été associé à un cancer.

2. Structure du virion

De l'intérieur vers l'extérieur, l'EBV est constitué d'un nucléoïde, d'une capside, d'un tégument et d'une enveloppe. Le nucléoïde comprend un ADN bicaténaire, linéaire enroulé autour d'une structure protéique. La capside icosaédrique recouvre le nucléoïde et est elle-même entourée d'un ensemble fibrillaire asymétrique : le tégument. Enfin, l'enveloppe est issue des membranes nucléaires et intracytoplasmiques : elle porte en surface une dizaine de glycoprotéines virales sous forme de spicules.



3. Génome de l'EBV

Le génome du virus d'Epstein-Barr se présente sous la forme d'un ADN double brin d'environ 172 Kbp. Il est linéaire dans la particule virale mais c'est sous une forme épisomale que le virus persiste de façon latente dans la cellule. Durant la phase de latence, l'EBV se réplique de façon synchrone avec la cellule infectée lors de la phase S du cycle cellulaire, en utilisant l'origine de réplication latente : Ori-P. Deux autres régions, appelées DR et DL, constituent les deux origines de réplication lytique (*OryLyt*) de l'EBV.

L'analyse de la séquence génomique de la souche B95-8 a identifié une centaine de cadres de lecture (ou ORF). Ces ORFs sont répertoriés selon les fragments de restriction obtenus après digestion du génome par l'enzyme BamHI. Les gènes EBV sont nommés selon le fragment BamHI dans lequel ils se trouvent (de A à Z), selon l'orientation de leur transcription (L pour « *leftward* » ou R pour « *rightward* ») et le numéro de phase de lecture (ou « *frame* »). Par exemple, le gène « BZLF1 » codant pour la protéine ZEBRA correspond au fragment BamHI-Z, d'orientation « *leftward* », et appartient à la phase de lecture (« *frame* ») n°1 de cette orientation.

On distingue les gènes codant pour les protéines de latence et les gènes codant pour les protéines du cycle lytique. Les premières sont exprimées lors de l'infection dite latente, qui permet la persistance du virus dans la cellule, en absence de synthèse de nouvelles particules

virales. Les protéines du cycle lytique sont, pour la plupart, des enzymes de transcription et de réplication virales ainsi que des protéines de structure.

4. Cycle viral

Le virus établit une infection majoritairement latente et l'activation spontanée du cycle lytique ne survient que dans 10 % des cellules au maximum.

4.1. Infection latente

Elle est caractérisée par l'expression d'un nombre restreint des antigènes viraux et par le maintien du génome sous forme épisomale. Chaque copie de l'épisome est répliquée une fois, durant le cycle cellulaire, et ségrège dans les cellules filles. L'origine de réplication de l'épisome est appelée Ori-P. C'est l'ADN polymérase de la cellule hôte qui assure la réplication de l'épisome viral.

Il existe quatre types de latence (table I.2.). Durant la latence de type I, seuls sont exprimés la protéine EBNA1 et certains ARNs non codants (EBERs et BARTs). La latence de type II est caractérisée par l'expression supplémentaire des protéines LMP1 et LMP2. Enfin, dans la latence de type III, tous les gènes de latence s'expriment : les 6 antigènes nucléaires (EBNAs 1, 2, 3A, 3B, 3C et LP), les 3 antigènes membranaires (LMPs 1, 2A et 2B) et les 2 ARNs nucléaires non codants EBERs 1 et 2.

4.2. Cycle lytique

Le cycle lytique, aussi appelé cycle réplcatif ou productif, peut être induit *in vitro* par différents composés tels que les esters de phorbol (TPA) ou le butyrate. Le génome de l'EBV, linéaire durant le cycle lytique, utilise la machinerie enzymatique virale ainsi que celle de la cellule infectée pour exprimer les protéines du cycle lytique, se répliquer et produire de nouveaux virions.

Les gènes du cycle lytique s'expriment séquentiellement : les produits des gènes très précoces activent l'expression des gènes précoces, dont les produits activent à leur tour la réplication du génome viral nécessaire à la transcription des gènes tardifs (dont les produits permettront l'assemblage et la maturation des particules virales).

Les protéines très précoces sont les protéines ZEBRA (« *Z EBV Replication Activator* ») et Rta codées respectivement par les gènes BZLF1 et BRLF1. Ce sont des facteurs de transcription activant plusieurs gènes précoces. ZEBRA suffit à elle seule pour lever la latence du virus et le faire entrer en phase lytique, notamment dans les lymphocytes B (plasmocytes) et les cellules épithéliales.

In vivo, la nature des cellules responsables de la réplication du virus et les signaux conduisant une cellule infectée par l'EBV en phase latente à entrer en réplication lytique sont inconnus. Il a été proposé que les cellules B infectées par le virus latent recirculent vers l'épithélium oropharyngé où elles se réactivent périodiquement et produisent de nouveaux virions libérés dans la salive (Laichalk et al., 2002). Les signaux initiant la réplication virale *in vivo* sont inconnus. Cependant, ils ne semblent pas provenir du virus étant donné que les protéines virales ZEBRA et Rta ne sont pas exprimées dans les cellules B mémoires infectées par le virus en phase latente. Le mécanisme d'activation des cellules B mémoires au niveau de l'épithélium oropharyngé entraîne leur différenciation en plasmocytes producteurs d'anticorps (Liu et al., 1997). L'équipe de David Thorley-Lawson a démontré que la différenciation des cellules B mémoire infectées en plasmocytes constitue le signal initiant la réplication virale étant donné que la différenciation cellulaire doit avoir lieu avant que le promoteur du gène BZLF1 codant pour le facteur de transcription ZEBRA soit activé (Laichalk et al., 2005). La première étape de différenciation cellulaire implique l'expression de la protéine régulatrice de la différenciation en plasmocytes BLIMP-1 (« B-lymphocyte-induced maturation protein 1 ») (Shapiro-Shelef et al., 2003). BLIMP-1 réprime l'expression de Pax-5, un régulateur négatif d'un autre facteur de différenciation en plasmocytes : XBP-1 (« X-box binding protein-1 »). Donc BLIMP-1 induit indirectement la transcription de XBP-1, qui est essentielle à la différenciation cellulaire en plasmocytes. Or, XBP-1 induit la transcription du facteur de transcription viral ZEBRA après la différenciation des cellules B mémoires en plasmocytes. En effet, le promoteur Zp du gène BZLF1 possède un site de liaison (le site ZII) du facteur XBP-1 (Sun et al., 2007, Bhende et al., 2007). Cette observation met en évidence un possible mécanisme moléculaire concernant l'utilisation par le virus EBV de la différenciation cellulaire en plasmocytes pour activer son cycle lytique.

Programmes de transcription virale		Principaux gènes exprimés	Conséquence sur lymphocytes B (individu sain)	Mise en évidence <i>in vivo</i>
Infection Latente	<u>Programme de croissance (Latence III)</u>	Les 6 EBNA Les 3 LMPs (+ARNs non codant EBER et BART)	Prolifération des Lymphocytes B néo-infectés	Primo-infection et MNI : amygdales et organes lymphoïdes <u>Porteur sain :</u> amygdales lors des réactivations. <u>Cancers associés à EBV :</u> lymphomes B immunodéprimés (30 à 100%) leyomyosarcome chez immunodéprimés (100% ?)
	<u>Programme par défaut (Latence II)</u>	EBNA1 et 3 LMPs ARNs EBER	Première étape de la différenciation des cellules B activées en lymphocytes B mémoires	<u>Primo-infection et MNI</u> principalement dans les amygdales (centre germinatif) Porteur sain : ? <u>Cancers associés à EBV :</u> .Maladie de Hodgkin (30 à 100%) .Carcinome. Indifférencié cavum (100%) .Certains lymphome T
	<u>Latence 0</u>	Aucun gène transcrit ou seulement LMP2A?	inconnue	<u>Primo-infection et MNI</u> à la fin de la primo-infection <u>Porteur sain :</u> lymphocytes B mémoires
	<u>Programme EBNA1 (Latence I)</u>	EBNA1 ARN EBER	maintien de l'épisode lors de la division des cellules mémoires	<u>Porteur sain :</u> lymphocytes B mémoires <u>Cancers associés à EBV :</u> lymphome de Burkitt endémique (100%)
Infection lytique		<u>Gènes Précoces immédiats</u>	Transactivation des gènes du cycle lytique	<u>Primo-infection et MNI :</u> amygdales et autres organes lymphoïdes <u>Porteur sain :</u> amygdales lors des réactivations. <u>Cancers associés à EBV :</u> (expression partielle ou complète du cycle lytique) Lymphomes B immunodéprimés Carcinome indifférencié du cavum <u>Autre pathologie non cancéreuse</u> Leucoplasie orale chevelue chez immunodéprimés
		<u>Gènes Précoces</u>	Synthèse ADN vir.	
		<u>Gènes Tardifs</u>	Synthèse des protéines de structure et des enzymes nécessaires à la formation de la particule virale infectieuse	

Table I.2. : Principales caractéristiques du cycle viral, latent et lytique (in vitro et in vivo).

5. Transmission du virus et pathologies associées

Les études épidémiologiques ont montré que 95 % de la population mondiale est infectée par l'EBV, dès le début de l'âge adulte. Cette prévalence importante et précoce est due à la transmission salivaire. En effet, il est bien établi que la transmission de l'EBV se fait principalement par voie salivaire. Les autres modes de contamination sont moins connus.

De nombreuses pathologies sont causées et associées à l'EBV. Ces maladies résultent de la particularité de ce virus qui persiste de façon latente tout au long de la vie de l'individu infecté. En général, l'infection latente reste asymptomatique. Néanmoins, dans certains contextes physiologiques, le sujet infecté développe des pathologies qu'il convient de distinguer selon que le sujet est immuno-compétent ou immuno-déprimé. Voici un bref résumé des pathologies associées à l'EBV, sachant que pour la plupart, elles sont la conséquence de déséquilibre entre les cycles latents et lytiques du virus et les réactions immunitaires du patient (Figure I.2.).

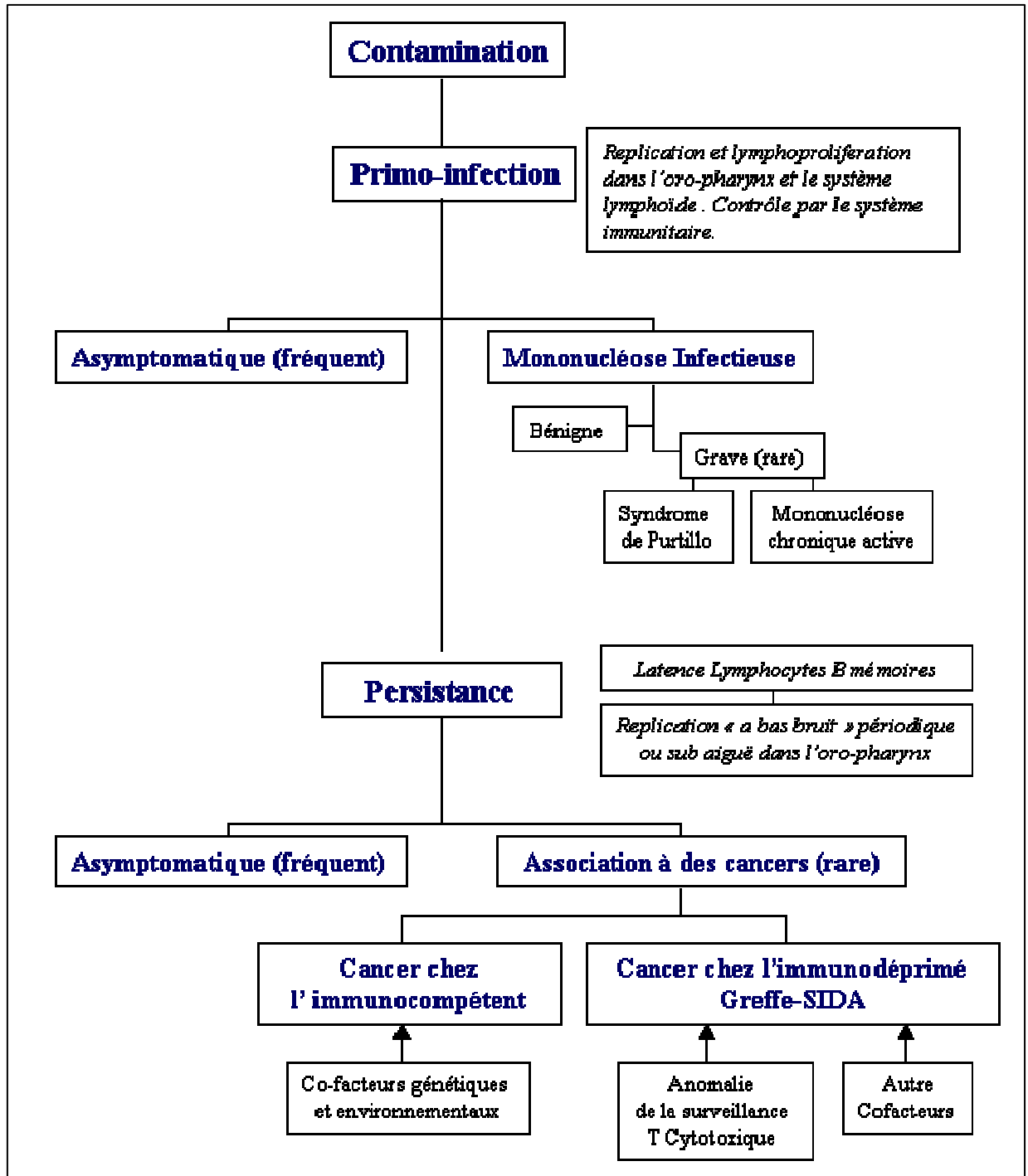


Figure I.2. : Physiopathologie de l'infection EBV

La mononucléose infectieuse (ou MNI)

La primo-infection à l'EBV est acquise dans la petite enfance et reste le plus souvent asymptomatique. Cependant, si le premier contact avec le virus survient chez l'adolescent ou le jeune adulte, l'infection se manifeste sous forme symptomatique (Mononucléose Infectieuse) dans 50% des cas. Les symptômes comprennent de la fièvre, une angine et des polyadénopathies (Schooley, 1988). Les patients atteints de MNI souffrent d'une grande fatigue, les symptômes pouvant durer plus de 6 mois.

L'EBV peut infecter de façon latente les lymphocytes B et stimuler leur prolifération. Ainsi, le virus d'Epstein-Barr est associé à de nombreuses tumeurs développées au dépens des lymphocytes (particulièrement de lymphocytes B) mais également dans certains cas, de cellules épithéliales.

Le lymphome de Burkitt

Historiquement, le lymphome de Burkitt (BL) fut le premier cancer associé à l'EBV. Le virus est retrouvé dans 95 % des cas de lymphomes de Burkitt africains. Des formes non endémiques et sporadiques du BL, observées dans les autres parties du monde, sont associées à l'EBV dans seulement 25 % des cas (30 % pour les immunodéprimés). Les patients souffrant d'un BL développent des tumeurs au niveau des mâchoires et des ganglions lymphatiques (figure I.3.).



Figure I.3. : Enfant atteint par un lymphome de Burkitt.

La latence observée dans le lymphome de Burkitt est de type I : seuls la protéine EBNA1 et les EBERs sont exprimés. Les cellules cancéreuses sont caractérisées par de nombreuses mutations affectant des gènes codant pour des protéines impliquées dans le contrôle de la prolifération cellulaire telles que c-myc (Kelly et al., 2006). De plus, la protéine virale EBNA1 inhibe les protéines cellulaires de contrôle de la prolifération telle que le facteur p53 (Kennedy et al., 2003). Enfin, de nombreux marqueurs antigéniques importants pour l'identification et l'élimination des cellules anormales par les lymphocytes T cytotoxiques (CTL) sont dérégulés dans les cellules cancéreuses du lymphome de Burkitt, ce qui rend le contrôle par le système immunitaire peu efficace (Staeger et al., 2002).

Le lymphome de Hodgkin

La maladie de Hodgkin est une autre forme de tumeur associée dans 40% des cas à l'EBV. Elle est définie comme la prolifération, au sein des ganglions, d'une forme atypique de cellules lymphoïdes B : les cellules malignes de Reed-Sternberg (Murray et al., 1992). Cette maladie se déclenche le plus souvent chez le jeune adulte et présente un second pic d'incidence autour de 45 ans (Flavell et al., 2000). D'autre part, le risque de développer une maladie de Hodgkin est deux à cinq fois plus élevé chez les individus ayant présenté une MNI. En général, la latence impliquée dans le lymphome de Hodgkin est de type II : trois protéines sont exprimées (EBNA1, LMP1 et LMP2). On suppose que l'EBV est un facteur important mais pas suffisant au déclenchement de la maladie, d'autres facteurs doivent entrer en jeu tels que l'origine géographique, le sexe, l'ethnie, l'âge... (Flavell and Murray, 2000).

Lymphomes T

Bien que le tropisme d'EBV soit principalement dirigé vers les lymphocytes B, certains cancers des lymphocytes T, dénommés "lymphome des cellules T", sont associés à l'EBV. Néanmoins les recherches ne sont pas très avancées dans ce domaine (Dupuis et al., 2006).

Cancers épithéliaux

- Le carcinome indifférencié du Rhinopharynx (NPC) :

Le carcinome du Rhinopharynx (NPC), est un carcinome peu ou pas différencié associé à l'EBV dans 100% des cas. Il atteint l'adulte avec deux pics : l'un autour de 20 ans et l'autre aux alentours de 50 ans. Il représente 20 % des cancers en Chine. Il est également présent en Afrique du Nord et plus rarement en Europe et aux Etats-Unis (Wei and Sham, 2005). Les données actuelles suggèrent une combinaison entre la présence du virus EBV et certains facteurs environnementaux et génétiques dans le déclenchement de la maladie. Le NPC est un carcinome qui résulte de la transformation des cellules épithéliales du Rhinopharynx. Comme la plupart des cancers associés à l'EBV, les événements clés dans le déclenchement de la maladie ne sont pas ou peu connus. Il semblerait qu'une infection latente par l'EBV soit nécessaire mais pas suffisante (Strauss & Strauss, 2002).

- Le cancer du sein :

L'association de l'EBV au cancer du sein a été évoquée (Bonnet et al., 1999). Des études suggèrent que les antécédents de MNI constituent un facteur de risque indépendant dans la survenue ultérieure du cancer du sein (Yasui et al., 2001). L'implication réelle de l'EBV reste à élucider et demeure discutée.

II. La protéine ZEBRA

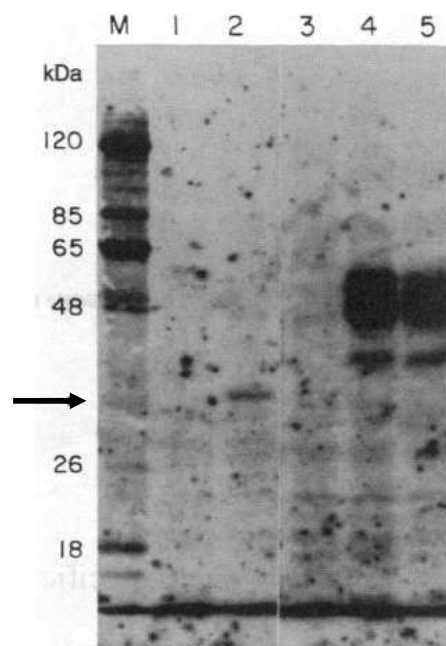
1. Historique de sa découverte

La souche P3-HR-1 de l'EBV est issue de la lignée cellulaire Jijoye Burkitt Lymphoma (Hinuma et al., 1967). Contrairement aux autres souches de l'EBV, P3-HR-1 ne peut immortaliser une lignée de lymphocytes (Menezes et al., 1975; Miller et al., 1974) mais est capable d'activer l'expression d'antigènes précoces (EAs pour « *early antigens* ») d'un génome latent de l'EBV après surinfection de cellules Raji (cellules EBV-positives) (Henle et al., 1970). En 1978, une étude mit en évidence l'hétérogénéité du génome viral de la souche P3-HR-1 de l'EBV (Delius and Bornkamm, 1978). En effet, une sous population de virus P3-HR-1 possède un ADN viral hétérogène (nommé *het* DNA) qui diffère de l'ADN viral classique de la souche P3-HR-1 par des réarrangements et des délétions importantes (Cho et al., 1984; Heston et al., 1982). George Miller et ses collaborateurs montrèrent que la capacité particulière de la souche P3-HR-1 à activer un génome latent de l'EBV était entièrement due à cette sous population de virus P3-HR-1 contenant un *het* DNA (Miller et al., 1984). Par la suite, cette même équipe identifia la région génomique de l'*het* DNA responsable de ses fonctions d'activation: il s'agit d'un fragment de 2,8 kpb qu'ils appelèrent le fragment BamHI-WZ *het* (ou BamHI-*het* 2.7) (Countryman and Miller, 1985). Ce fragment ne contient qu'une seule ORF (*open reading frame*) intacte correspondant à la séquence BZLF1¹ du génome standard de la souche P3-HR-1, et des portions de trois autres ORFs : BWRF1², BRLF1³ et BZLF2¹. Ils se demandèrent alors si ce fragment d'ADN BamHI-WZ *het* seul suffisait à l'induction de génomes latents de l'EBV, ou si le produit de transcription de ce fragment était responsable des effets observés. Leurs études leur ont permis de conclure que le polypeptide de 33 kDa codé par le fragment BamHI-WZ *het* était responsable de l'induction des génomes latents (figure I.4.).

BZLF1¹, BZLF2¹ : « Bam H1-Z fragment – leftward – frame 1 ». C'est-à-dire le cadre de lecture n°1 (ou n°2 pour BZLF2¹) en direction de la gauche du fragment de digestion BamHI-Z du génome de l'EBV.

BWRF1² : « Bam H1-W fragment – rightward – frame 1 ». C'est-à-dire le cadre de lecture n°1 en direction de la droite, du fragment de digestion BamHI-W du génome de l'EBV.

BRLF1³ : « Bam H1-R fragment – leftward – frame 1 ».



(Countryman J. et al., 1985)

Figure I.4. : Immunoblot des polypeptides antigènes de l'EBV induits après transfection du fragment BamHI-WZ *het*.

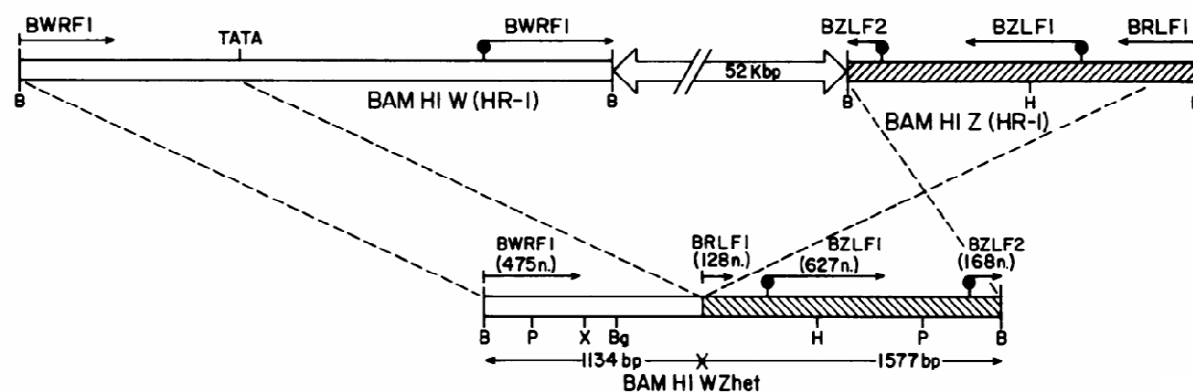
Pistes 1 et 2 : transfection de cellules LTK- par un vecteur vide pSV2neo (piste 1) ou par le vecteur pSV2neo contenant le fragment BamHI-WZ *het* (piste 2). Pistes 3 et 4 : transfection de cellules D98/HR-1 par le vecteur vide pSV2neo (piste 3) ou par le vecteur pSV2neo contenant le fragment BamHI-WZ *het* (piste 4). Piste 5 : cellules D98/HR-1 traitées au PMA (phorbol 12-myristate 13-acétate). La flèche à gauche indique la bande correspondant au polypeptide de 33 kDa issu de la traduction du fragment BamHI-WZ *het*. Piste "M" : marqueurs de poids moléculaires.

La détection des antigènes a été réalisée par un sérum humain polyvalent W.C., et suivie par un traitement à la protéine A staphylococcale marquée à l' ^{251}I .

(Les cellules D98/HR-1 sont des hybrides entre les cellules (adénine phosphoribosyltransférase)⁻ et la lignée du lymphome de Burkitt HR-1 (Glaser and Rapp, 1972). Elles expriment l'antigène nucléaire EBNA de l'EBV, un produit de la phase latente, mais une très faible proportion de cellules produit les antigènes du cycle de réplication virale. Les cellules LTK- correspondent à des cellules (thymidine kinase)⁻ de souris).

Ces observations furent confirmées par Kenzo Takada en 1986 qui montra l'importance de la séquence BamHI Z seule (correspondant à l'ORF BZLF1 du génome standard de l'EBV) sur l'activation de génomes latents de l'EBV de cellules Raji (Takada et

al., 1986). Jill Countryman et ses collègues comparèrent l'effet de la transfection du fragment BamHI-WZ *het* et de la séquence BamHI Z seule sur l'activation du génome latent de l'EBV de cellules D98/HR-1 (Countryman et al., 1987). Ils montrèrent ainsi que la transfection de la séquence BamHI Z seule permettait effectivement d'induire la réplication du virus mais de manière beaucoup moins efficace que suite à la transfection du fragment BamHI-WZ *het*. Ce résultat intrigant fut éclairci par Cliona Rooney de l'équipe de George Miller l'année suivante : ils démontrèrent que les réarrangements génomiques particuliers du fragment BamHI-WZ *het* permettait l'expression du gène BZLF1 qui est normalement silencieux dans les génomes non réarrangés de l'EBV (Rooney et al., 1988). En effet, la fusion des séquences BamHI W et BamHI Z a entraîné le positionnement de régions promotrices contenues dans la séquence BamHI W en amont de l'ORF BZL1, permettant ainsi le recrutement de facteurs de transcription cellulaires activant l'expression du gène BZLF1 (figure I.5.).



(Rooney et al., 1988)

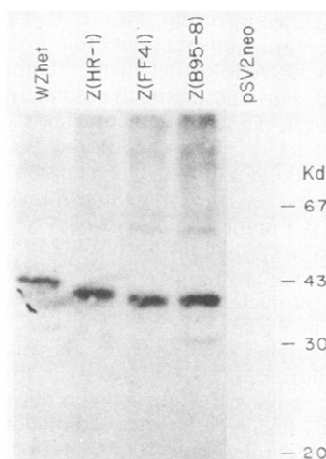
Figure I.5. : Carte génomique du fragment BamHI-WZ *het*.

Ces réarrangements génomiques de l'ADN *het* des virus P3-HR-1 résulteraient d'un mécanisme de dispersion utilisé par le virus de l'EBV pour rompre la latence et entrer en phase de réplication, tout en échappant aux voies de contrôle des cellules hôtes.

Chez d'autres virus à ADN, l'initiation du cycle lytique repose sur l'expression de facteurs de *trans*-activation précoces (« *immediate early proteins* ») comme la phosphoprotéine E1a de l'adénovirus (Berk et al., 1979; Jones and Shenk, 1979) ou la protéine ICP4 du virus HSV (Dixon and Schaffer, 1980; Preston et al., 1978; Watson and Clements, 1978). Les travaux de l'équipe de George Miller avaient mis en évidence l'implication du polypeptide de 33kDa résultant de l'expression de la séquence BamHI Z du

fragment *het* DNA dans le mécanisme de rupture de la latence. Mais c'est l'équipe d'Alain Sergeant. en 1986 qui démontra que ce polypeptide était un facteur de *trans*-activation viral précoce qu'ils nommèrent EB1 (Chevallier-Greco et al., 1986). Leurs travaux leur permirent également d'identifier un deuxième facteur de *trans*-activation viral précoce du virus EBV : le facteur EB2 (encore appelé SM, BMLF1 ou Mta, produit de l'ORF *BSLF2-BMLF1*). En effet, en transfectant la séquence BamHI Z non seulement dans des cellules EBV-positives telles que les cellules Raji mais également dans des cellules EBV-négatives telles que les cellules BJAB, ils prouvèrent la nécessité d'une coopération entre les facteurs EB1 et EB2 pour *trans*-activer le génome de l'EBV.

D'autres noms furent par la suite attribués au produit d'expression du gène BZLF1 : en 1987, l'équipe de George Miller (Grogan et al., 1987) l'appelèrent « ZEBRA » pour « *BamHI-Z Epstein-Barr replication activator* » ou tout simplement « Z » ou « Zta ». Et cette même équipe présenta en 1987 les premiers immunoblots de la protéine ZEBRA et mit en évidence des différences de taille et de migration électrophorétique de ZEBRA suivant que le gène BZLF1 provenait d'un génome de souche B95-8, FF41, P3-HR1 ou P3-HR-1 *het* (figure I.6.) (Countryman et al., 1987). L'existence de différents variants de ZEBRA est discutée dans la partie II.2.



(Countryman et al., 1987)

Figure I.6. : Mobilité électrophorétique de la protéine ZEBRA (*BamHI-Z Epstein-Barr replication activator*) du virus d'Epstein-Barr.

Immunoblot de protéines extraites de cellules COS-1 transfectées avec différents plasmides pSV2neo contenant l'insert BamHI-WZ *het* (piste1) ou l'insert BamHI-Z provenant des souches HR-1 (piste2), FF41 (piste 3) ou B95-8 (piste 4). Le vecteur pSV2neo vide a été chargé piste 5. The immunoblot was reacted with a 1:250 dilution of a polyvalent human antiserum, SC.

Enfin, un éclaircissement fut apporté en 1988 avec la détermination de la séquence nucléotidique du fragment BamHI-WZ l'équipe de George Miller : il est formé de la fusion des séquences BamHI W et BamHI Z du génome standard de la souche P3-HR-1, séquences qui se trouvent normalement à plus de 55 kpb l'une de l'autre dans le génome classique de l'EBV (figure I.5.) (Jenson and Miller, 1988). Cette équipe confirma la présence de régions promotrices dans la séquence BamHI W placées en amont de l'ORF BZL1 grâce au réarrangement particulier de l'*het* DNA. Comme aucune homologie significative n'apparaît entre les séquences BamHI W et BamHI Z, leur fusion résulterait d'une recombinaison non-homologue.

La protéine ZEBRA, qui apparaît déjà comme un facteur clé du déclenchement du cycle lytique de l'EBV, va être intensivement étudiée et révèle de nombreuses fonctions tant au niveau du développement du virus que du contrôle de l'expression des gènes cellulaires.

2. Les variants de ZEBRA

Certains variants de ZEBRA et leur relation à des pathologies associées à l'EBV ont été mis en évidence. Le gène BZLF1 codant pour la protéine ZEBRA est composé de trois exons et deux introns. L'équipe d'Irène Joab a étudié les séquences de transcrits BZLF1 issus de biopsies de tumeurs NPC (Cochet et al., 1993). En les comparant aux transcrits BZLF1 de la souche B95-8, deux différences ont été détectées : la première concerne le 1^{er} exon et résulte en une mutation de l'acide aminé leucine 61 en isoleucine. La seconde différence concerne une délétion de deux bases dans le site donneur d'épissage du 2^{ème} intron. Cette mutation jouerait un rôle dans la vitesse d'épissage des ARN messagers BZLF1. En effet, les analyses PCR réalisées lors de cette étude ont montré que la majorité des transcrits BZLF1 ne sont pas du tout ou incomplètement épissés, ce qui expliquerait la très faible proportion de cellules tumorales NPC exprimant la protéine ZEBRA.

Une autre mutation a été mise en évidence par cette même équipe et concerne la mutation de l'acide aminé 206 en sérine : variant « Zser206 » (Grunewald et al., 1998). Cette mutation est plus fréquemment observée dans les virus de lignée B95-8 de tumeurs NPC de patients provenant d'Afrique du Nord et d'Europe (72 % des transcrits issus de tissus NPC provenant de ces zones géographiques présentent ce variant Zser206 alors qu'il n'est présent que dans 13 à 18 % des cas dans les cellules périphériques du sang de sujets sains issus des mêmes zones géographiques). Une mutation de l'acide aminé alanine 205 en sérine a également été mise en évidence dans les virus EBV issus des lignées cellulaires P3HR1 et

Akata, ainsi que de virus B95-8 isolés à partir de lésions OHL (Lau et al., 1992). Les résidus 205 et 206 appartiennent au domaine de dimérisation de ZEBRA. Bien qu'une diminution significative de la capacité à dimériser ait été observée pour le variant Zser206 (Martel-Renoir et al., 2000), la mutation de ce résidu ne semble pas affecter les fonctions de transactivation de ZEBRA. Cependant, les résidus 205 et 206 pourraient contribuer, via leur localisation sur les hélices α du dimère de ZEBRA, aux interactions de ZEBRA avec des protéines cellulaires étant donné que la double mutation A205R/A206D annule l'interaction entre ZEBRA et les récepteurs aux acides rétinoïques RAR α and RXR α (Sista et al., 1995). Plus récemment, 5 nouveaux variants BZLF1 ont été identifiés dans une région sud vde la Chine, endémique au NPC. 11 acides aminés diffèrent de la souche B95.8 (V01555) (Ji et al., 2009).

III. ZEBRA et ses multiples fonctions...

1. Les protéines bZIPs

Les protéines à motif leucine zipper, y compris les protéines bZIP, contiennent dans leur séquence primaire, une région riche en leucines dans laquelle un résidu leucine est présent de façon répétitive tous les 7 acides aminés.

Dans tous les cas, la région riche en leucine adopte une structure en hélice α dans laquelle les leucines apparaissent tous les deux tours d'hélices et sur le même côté de l'hélice. Cette configuration leur permet ainsi d'interagir, via leur chaîne latérale, avec les résidus leucine d'un deuxième monomère, formant un motif en « leucine zipper » et résultant en la dimérisation des deux monomères.

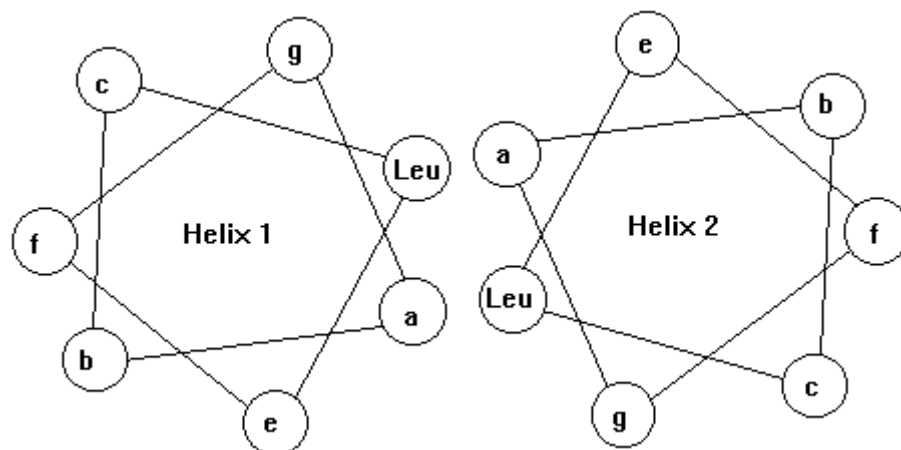


Figure I.7. : Représentation schématique de la structure en hélice d'une protéine à leucine zipper présentant une séquence de 7 résidus **abcdefg**, où la position « **d** » est toujours occupée par une leucine. La position « a » étant généralement occupée par un résidu hydrophobe, cela permet de former des contacts hydrophobes à l'interface hélice-hélice des deux monomères.

Ces résidus leucines présents tous les 7 acides aminés (figure I.7.) sont extrêmement conservés chez toutes les protéines bZIPs. Des substitutions de ces résidus leucine en isoleucine ou valine chez les protéines bZIPs telles que C/EBP α ou Fos et Jun empêchent leur dimérisation avec un monomère intact, indiquant le rôle critique de ce domaine de

dimérisation. Les régions riches en leucines de deux monomères s'associent de manière parallèle, avec leurs hélices orientées dans la même direction comme illustré grâce aux études structurales de l'hétérodimère Fos/Jun lié à l'ADN (Glover and Harrison, 1995) : chaque monomère forme une hélice α et la dimérisation résulte de l'association de deux hélices parallèles qui s'enroulent l'une autour de l'autre pour former un motif structural appelé « *coiled-coil* » (ou super-enroulement d'hélices), similaire à celui adopté par les protéines fibreuses comme la kératine ou la myosine.

Contrairement aux protéines « à doigts de zinc » ou aux protéines « hélice-boucle-hélice », le motif leucine zipper ne contacte pas directement l'ADN. Ce rôle est assuré par une région adjacente très riche en résidus basiques, permettant à cette région d'interagir directement avec l'ADN acide. Le motif leucine zipper jouerait cependant un rôle structural indirect dans la fixation à l'ADN : en facilitant la dimérisation, il permettrait ainsi de positionner correctement les régions basiques des deux monomères sur la double hélice d'ADN. De façon intéressante, la région basique d'une protéine bZIP n'est pas complètement structurée avant sa liaison à l'ADN : elle ne se structure complètement en hélices α qu'une fois liée à sa séquence ADN.

2. La régulation de l'expression de ZEBRA

En plus d'une auto-activation par la protéine ZEBRA elle-même, le promoteur Zp contrôlant l'expression du gène BZLF1 est régulé par de nombreux facteurs de transcription cellulaires. La figure I.8. présente l'organisation des sites de fixation de ces facteurs de régulation cellulaires dans le promoteur Zp :

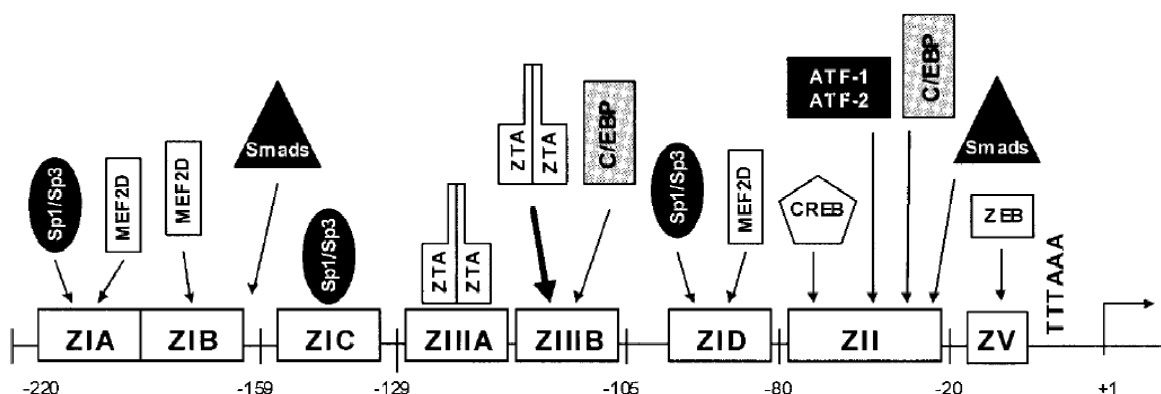


Figure I.8. : Organisation des sites de fixation en *cis* des facteurs de transcription cellulaires sur le promoteur Zp du gène BZLF1. Les éléments ZI (ZI A, B et D) sont liés par les facteurs

MEF2D et Sp1/Sp3. Le site ZII est lié par les facteurs CREB, ATF1/2, cJun et C/EBP α (« *CAAT enhancer binding protein α* »). Les deux sites ZIIIA et ZIIIB sont reconnus par la protéine ZEBRA pour une auto-régulation positive. Cependant, ZIIIB peut également être lié par la protéine C/EBP α . Les séquences ZIV et HBox sont des éléments de régulation négative respectivement reconnus par les facteurs YY1 et E2-2. Enfin, près de la TATA box, le site ZV est un site important de régulation négative lié par la protéine ZEB. Des sites potentiels de fixation du facteur SMAD, responsable de l'activation du facteur TGF- β , sont indiqués.

A l'origine, les séquences de régulation du promoteur Zp ont été classées selon 4 groupes : ZI, ZII, ZIII et ZIV. La région -233 à +12 du promoteur Zp rassemble les principales séquences de régulation associées à l'induction du cycle lytique. Les 4 membres du groupe « ZI » (ZI A à ZI D) sont des motifs riches en paires de bases A-T régulant négativement le promoteur Zp en absence de stimuli du cycle lytique. Le facteur cellulaire MEF2D, fixé aux sites ZI A, B et D (Liu et al., 1997), recrute les complexes HDACs (« *histone deacetylase complexes* ») réduisant ainsi l'activité du promoteur Zp (Gruffat et al., 2002). Cependant, ces sites ZI deviennent des éléments de régulation positive en présence de certains stimuli d'induction du cycle lytique, comme les esters du phorbol, les ionophores calcium ou les agents de chimiothérapie (Flemington and Speck, 1990a).

Le motif ZII est reconnu par de multiples facteurs cellulaires comme le facteur CREB, ATF-1 et l'hétérodimère ATF-2/cJun, et l'hétérodimère cJun/cFos, aussi connu sous le nom de AP1 pour « *activator protein 1* » (Adamson et al., 2000; Flamand and Menezes, 1996; Liu et al., 1998; Wang et al., 1998). Ces facteurs de transcription cellulaires sont communément exprimés dans différents types cellulaires mais leur capacité d'activation est souvent dépendante de leur état de phosphorylation (notamment par les kinases p38 et JNK (« *cJun N-terminal kinase* »)). Il a également été démontré que le motif ZII, en conjonction avec le site Smad, est important pour l'activation du promoteur Zp par le facteur TGF- β dans certaines lignées cellulaires du lymphome de Burkitt.

Les deux sites ZIII A et ZIII B sont des éléments « ZREs » (pour « *Z-responsive elements* ») qui sont reconnus directement par la protéine ZEBRA. Cependant, la capacité de ZEBRA à activer son propre promoteur dans le contexte d'un génome viral latent reste controversé (Le Roux et al., 1996; Yin et al., 2004; Zalani et al., 1996).

De façon intéressante, les sites ZII et ZIII B du promoteur Zp sont également reconnus par le facteur de transcription cellulaire C/EBP α (« *CAAT enhancer binding protein α* »). Il a été suggéré qu'une coopération existerait entre ZEBRA et C/EBP α pour activer le promoteur Zp au niveau de ces sites (Wu et al., 2004).

Le promoteur Zp contient aussi différents sites de régulation négative. Parmi eux, le motif ZV, situé juste à côté de la TATA box, apparaît comme le plus important. ZV est reconnu par le facteur ZEB, une protéine à doigts de zinc (Kraus et al., 2003). Le facteur YY-1, lié aux sites ZIV, régule également de façon négative le promoteur Zp (Montalvo et al., 1985; Montalvo et al., 1991). Enfin, un certain nombre de loci « HBox », reconnus par le facteur de transcription E2-2, régulent négativement le promoteur Zp en cellules B alors qu'ils l'activent en cellules épithéliales (Thomas et al., 2003).

3. Les fonctions de *trans*-activation de ZEBRA

ZEBRA est membre de la famille des protéine bZIPs (ou protéines à « leucine-zipper »), incluant entre autres les facteurs de transcription cellulaires cJun et cFos (Farrell et al., 1989; Kouzarides et al., 1991). ZEBRA reconnaît des sites similaires aux sites AP1 appelés sites ZREs (« *Z-responsive elements* ») (Chang et al., 1990; Lieberman and Berk, 1990). Les motifs ZREs sont présents dans de nombreux promoteurs viraux précoces. La présence de deux sites ZREs ou plus conduisent à une *trans*-activation synergique de ZEBRA (Carey et al., 1992). ZEBRA interagit directement avec de nombreux facteurs de transcription cellulaires, incluant TFIID et TFIIA (Chi and Carey, 1993; Lieberman and Berk, 1994). De plus, ZEBRA interagit directement avec les complexes HDACs (notamment les facteurs CBP et p300) ce qui résulte en l'acétylation de la chromatine et ainsi l'activation de la transcription (Adamson and Kenney, 1999; Chen et al., 2001; Deng et al., 2003). Bien que ZEBRA ne puisse former d'hétérodimères stables avec les facteurs cJun ou cFos, elle semble s'associer à un autre facteur de transcription de la famille des bZIPs, le facteur C/EBP α , pour *trans*-activer son propre promoteur Zp via des sites de fixation C/EBP α (Wu et al., 2004).

Au niveau du génome viral, ZEBRA active en tout premier lieu le promoteur Rp du gène BRLF1 (figure I.9.), afin de rompre la phase de latence (Adamson and Kenney, 1998; Francis et al., 1997).

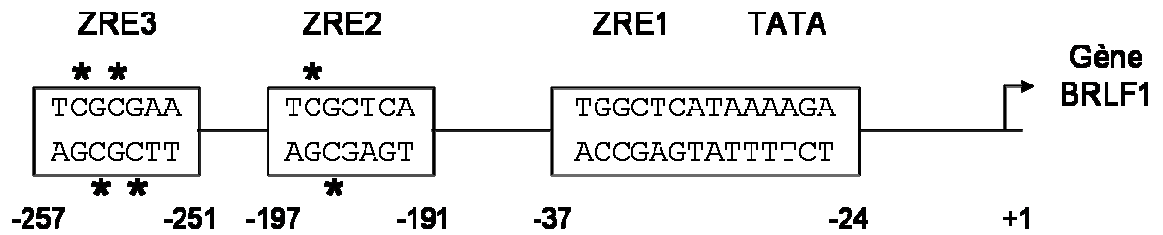


Figure I.9. : Schéma du promoteur viral du gène BRLF1. Les séquences des trois sites ZREs (ZRE1, ZRE2 et ZRE3) du promoteur Rp sont représentées. Les îlots CpG méthylables des sites ZRE2 et ZRE3 sont indiqués par un astérisque. Le site ZRE1 ne contient pas de cytosine méthylable.

Le mécanisme d'activation du promoteur Rp par ZEBRA semble différent de celui qu'utilise ZEBRA pour activer d'autres promoteurs viraux précoces. Par exemple, un mutant de ZEBRA où la sérine 186 a été mutée en alanine (résidu normalement présent chez tous les autres bZIPs connus), empêche ZEBRA de lier et d'activer les sites ZREs du promoteur Rp, mais n'affecte pas ses capacités de liaison à d'autres sites ZREs de promoteurs précoces (Adamson and Kenney, 1998; Francis et al., 1997). Le promoteur Rp présente trois sites ZREs atypiques dont deux contiennent des îlots CpG (ZRE2 et ZRE3) qui peuvent ainsi être méthylés (table I.3.) :

Heptamère	Séquence
ZRE1	TGGCTCA ACCGAGT
ZRE2	* TCGCTCA AGCGAGT *
ZRE3	* * TCGCGAA AGCGCTT * *
AP1	TGACTCA ACTGAGT

Table I.3. : Représentation des séquences nucléotidiques des 3 heptamères ZREs du promoteur Rp ainsi que du site AP1. Les astérisques indiquent les cytosines méthylables des îlots CpG des sites ZRE2 et ZRE3.

En effet, le promoteur Rp est intensivement méthylé durant la phase de latence (Bhende et al., 2004). De façon surprenante, ZEBRA fixe préférentiellement le site ZRE2 dans son état méthylé (elle lie également le site ZRE3 méthylé) (Bhende et al., 2004). De plus, la structure cristallographique du complexe ZEBRA/AP1 résolue par Carlo Petosa et ses collaborateurs en 2006 (Petosa et al., 2006), suggère un rôle clé de la sérine 186 de ZEBRA, qui interagirait directement avec une cytosine méthylée du site ZRE2. Cette particularité unique à ZEBRA à activer préférentiellement des promoteurs méthylés, suggère que le virus de l'EBV a trouvé un échappatoire aux effets inhibiteurs de la méthylation du génome viral.

Au niveau des gènes cellulaires, ZEBRA *trans*-active via des sites AP1 ^{et/ou} des sites ZREs, de nombreux gènes tel que le facteur de transcription cFos (Flemington and Speck, 1990a), l'interleukine IL-10 (Mahot et al., 2003), le facteur de croissance TGF- β (Cayrol and Flemington, 1995), la tyrosine-kinase TKT (Lu et al., 2000) et les métalloprotéases de matrice 1 et 9 (Lu et al., 2003; Yoshizaki et al., 1999). L'activation des cytokines humaines IL-10 et TGF- β par ZEBRA aurait pour but de supprimer la réponse immunitaire cellulaire vis à vis du virus de l'EBV. Le rôle de l'activation par ZEBRA des autres gènes cellulaires cités ici est moins clair.

4. Les fonctions de réplication de ZEBRA

En plus du rôle essentiel de facteur de transcription viral, la protéine ZEBRA joue un rôle prédominant dans le mécanisme de réplication virale. L'origine de réplication de l'EBV, *oriLyt*, contient une série de sites ZREs, et la fixation de ZEBRA sur ces sites est requise pour déclencher la réplication virale (Fixman et al., 1992; Hammerschmidt and Sugden, 1988). La mutation des résidus 12 et 13 de ZEBRA ne lui permet plus d'assurer ces fonctions de réplication mais n'affecte pas ses fonctions de *trans*-activation (Sarisky et al., 1996). Enfin, plusieurs équipes ont démontré que ZEBRA interagissait directement avec différentes protéines du complexe de réplication virale (Gao et al., 1998; Takagi et al., 1991; Zhang et al., 1996). Ces résultats suggèrent que la fixation de ZEBRA sur l'origine de réplication *oriLyt* est essentielle à la réplication de l'EBV et que cette fixation engendrerait la formation du complexe d'initiation de la réplication.

5. ZEBRA et le contrôle du cycle cellulaire

Les Herpesvirus en général interviennent dans les mécanismes de contrôle du cycle cellulaire pour assurer la disponibilité en substrats adéquats à la réplication de l'ADN viral et également, pour avoir le temps de se répliquer. ZEBRA exerce différents effets sur le cycle cellulaire qui sont dépendants du type cellulaire concerné. Chez certains types cellulaires, incluant les fibroblastes, ZEBRA active l'expression de la protéine p21 tandis qu'elle inhibe l'expression de la cycline A et du facteur c-myc. Ces fonctions résultent en un arrêt brutal du cycle cellulaire en phase G1/S (Cayrol and Flemington, 1996; Mauser et al., 2002; Rodriguez et al., 1999). Dans les cellules HeLa, ZEBRA diminue le taux de cycline B, résultant en un arrêt du cycle en phase G2 et en une inhibition de la condensation de la chromatine (Mauser et al., 2002). Dans les keratinocytes modèles illustrant le mieux les cellules cibles que l'EBV infecte *in vivo* dans l'oropharynx, ZEBRA active l'expression des protéines associées au déclenchement de la phase S : E2F61 et les cyclines E et A. Enfin, beaucoup moins d'études ont été réalisées concernant le contrôle du cycle cellulaire par ZEBRA dans les lymphocytes. Ces résultats suggèrent qu'un environnement cellulaire bloqué en phase G1/S, dans lequel la réplication de l'ADN cellulaire est impossible mais où certaines protéines impliquées dans la phase G1/S sont exprimées, serait un environnement idéal pour la réplication de l'ADN viral.

L'inhibition de la réplication de l'ADN cellulaire diminuerait la compétition pour les substrats nécessaires au mécanisme de réplication entre le virus et la cellule hôte, tandis que l'expression de protéines cellulaires impliquées dans la phase G1/S, comme E2F-1, semblent nécessaires à la réplication virale.

6. Les interactions ZEBRA-protéines cellulaires

Les interactions ZEBRA-protéines cellulaires connues à ce jour sont listées dans la table I.4. Pour la plupart de ces interactions, les études qui ont été entreprises pour évaluer la pertinence de l'interaction ont été réalisées *in vitro* après sur-expression d'au moins un des partenaires. Il n'est donc pas encore prouvé que ces interactions se produisent *in vivo* et le défi actuel est de comprendre la pertinence de ces interactions vis-à-vis de la biologie du virus d'Epstein-Barr.

Protéine cellulaire interagissant avec ZEBRA	Références
p53	Zhang Q. <i>et al.</i> , 1994 Mauser A. <i>et al.</i> , 2002
CBP	Adamson A.L. et Kenney S., 1999 Zerby D. <i>et al.</i> , 1999
C/EBP α	Wu F.Y. <i>et al.</i> , 2003
TFII A-D	Lieberman P.M. et Berk A.J., 1991, 1994 Chi T. et Carey M., 1993, 1996 Lieberman P.M. <i>et al.</i> , 1994, 1997 Chi T. <i>et al.</i> , 1995 Ozer J. <i>et al.</i> , 1996 Berk A.J. <i>et al.</i> , 1998 Deng Z. <i>et al.</i> , 2001
RXR α , RAR α	Sista N.D. <i>et al.</i> , 1993, 1995 Pfitzner <i>et al.</i> , 1995
NF κ B	Gutsch <i>et al.</i> , 1994 Hong <i>et al.</i> , 1997
Ubinucléine	Aho <i>et al.</i> , 2000
Rack1	Baumann <i>et al.</i> , 2000

Table I.4. : Résumé des interactions ZEBRA-protéines cellulaires (Sinclair, 2003).

▪ p53

Zta interagit avec p53 *in vivo* et *in vitro* et régule négativement la fonction de *trans*-activation de p53 (Mauser et al., 2002; Zhang et al., 1994). La région bZIP de ZEBRA interagit avec la région C-terminale de p53. Cependant, Mauser A. *et al.* ont récemment montré que certains des effets de régulation négative de ZEBRA sur les mécanisme de transcription p53-dépendante nécessiterait l'intervention de la protéine « *TATA binding protein* » (TBP)(Mauser et al., 2002).

▪ La protéine CBP (“CREB-binding protein”)

L'interaction entre ZEBRA et la protéine CBP a été mise en évidence dans des systèmes de sur-expression et dans des cellules infectées par l'EBV en cycle lytique (Adamson and Kenney, 1999; Zerby et al., 1999). La région de CBP impliquée dans l'interaction avec ZEBRA correspond au motif CH/1 de la moitié N-terminale et à la région CH/3 du domaine C-terminal. La région CH/1 stimule l'expression du gène BZLF1 dans un système gène-reporteur alors que la région CH/3 réprime les fonctions de *trans*-activation de ZEBRA. Cependant, dans son contexte physiologique, la protéine CBP coopère clairement avec ZEBRA dans ses fonctions de *trans*-activation et d'induction du cycle lytique. L'identification de la région de ZEBRA impliquée dans l'interaction avec CBP est moins connue. Des études de mutagenèse suggèrent que la région de dimérisation de ZEBRA est nécessaire ainsi que certains résidus du domaine de *trans*-activation.

Le fait que CBP régule positivement les fonctions de *trans*-activation de ZEBRA serait du au fait que ZEBRA stimule l'activité « histone-acétyl transférase de CBP (Chen et al., 2001), ce qui entraîne la réorganisation de la chromatine et améliore l'accessibilité des sites ZREs de l'ADN.

▪ C/EBP α

L'interaction directe entre ZEBRA et C/EBP α semble activer la fonction de *trans*-activation de C/EBP α , conduisant à la stimulation de l'expression de C/EBP α et d'une cible en aval : p21(CIP-1), entraînant ainsi l'arrêt du cycle cellulaire (Wu et al., 2003; Wu et al., 2004). La capacité à activer l'expression des protéines C/EBP α et p21(CIP-1) dépend du

domaine C-terminal de ZEBRA (Wu et al., 2003). La fonction de cette interaction semble être principalement en rapport avec le contrôle du cycle cellulaire par ZEBRA.

▪ **Les protéines de la machinerie de transcription**

La première indication d'une interaction entre ZEBRA et certaines protéines spécifiques de la machinerie transcriptionnelle est venue des travaux de Lieberman P.M. et Berk A.J. qui ont démontré que non seulement ZEBRA et TBP interagissent directement, mais que cette interaction stabilise l'association de TFIID avec la « *TATA box* » (Lieberman and Berk, 1991). D'autres travaux ont par la suite décrit la formation de complexes stables d'initiation de la transcription contenant ZEBRA (Chi and Carey, 1996; Chi et al., 1995; Deng et al., 2001; Lieberman and Berk, 1994; Lieberman et al., 1997; Ozer et al., 1996).

▪ **Autres protéines cellulaires**

ZEBRA interagit avec les récepteurs aux acides rétinoïques RAR α et RXR α (Pfitzner et al., 1995; Sista et al., 1995). La pertinence fonctionnelle de cette interaction vient du fait que les acides rétinoïques sont capables d'empêcher la réactivation de l'EBV en phase de latence (Sista et al., 1993). Les régions de dimérisation et de *trans*-activation de ZEBRA semblent être impliquées dans cette interaction (Pfitzner et al., 1995).

La région de dimérisation de ZEBRA est également impliquée dans l'interaction avec la sous-unité p65 du facteur de transcription NF κ B (Gutsch et al., 1994; Hong et al., 1997). La sur-expression de p65 en cellules inhibe les fonctions de *trans*-activation de ZEBRA. Cependant, on ne sait pas encore si cette inhibition s'exerce de façon réciproque ni si l'interaction ZEBRA-p65 se produit dans des conditions physiologiques.

Une autre protéine cellulaire interagissant avec ZEBRA, appelée ubinucléine, a été identifiée en 2000 (Aho et al., 2000). Une interaction similaire existe entre l'ubinucléine et le facteur de transcription cellulaire c-jun, mais la signification fonctionnelle de ces interactions est pour l'instant obscure.

La protéine RACK1 a été identifiée comme partenaire cellulaire de ZEBRA grâce à un système double hybride utilisant le domaine de *trans*-activation de ZEBRA comme appât (Baumann et al., 2000). Malgré les relations entre la protéine RACK1 et la kinase C, et malgré le fait que la kinase C soit impliquée dans le mécanisme de rupture de la latence de

l'EBV par les esters de phorbol, RACK1 n'influence ni l'état de phosphorylation de ZEBRA ni ou les fonctions de *trans*-activation de Zta *in vitro* ou *in vivo* (Baumann et al., 2000). Cela suggère que l'interaction RACK1-ZEBRA n'intervient pas dans la transduction du signal des esters de phorbol à ZEBRA.

ZEBRA peut enfin interagir directement avec le facteur CREB lui permettant de réguler négativement les fonctions de *trans*-activation de ce dernier (Adamson and Kenney, 1999). Il est intéressant de remarquer que ZEBRA dispose de deux voies de régulation négative du facteur CREB : soit par une interaction directe avec celui-ci soit en interagissant avec la protéine CBP (« *CREB-binding protein* »). Il a été suggéré que CREB jouait un rôle dans la régulation du promoteur du gène BZLF1 et que la capacité de ZEBRA à réprimer les fonctions de CREB lui permettrait d'exercer une auto-régulation de la transcription du gène BZLF1.

IV. Structure cristallographique de ZEBRA en complexe avec le site AP1

1. Organisation en domaines de ZEBRA

ZEBRA est une protéine de 245 acides aminés organisée en deux domaines principaux : un domaine de *trans*-activation N-terminal et un domaine de dimérisation et de liaison à l'ADN (domaine DBD) (figure I.10.). Ce dernier est un domaine de type bZIP (« *Basic region leucine zipper* ») composé de trois régions : une région basique (BR), une région de dimérisation « *leucine zipper* » atypique vis-à-vis des séquences consensus (ZIP), et une queue C-terminale (CT). Celle-ci comprend les 24 derniers résidus de ZEBRA et est essentielle aux fonctions d'homodimérisation et de fixation à l'ADN de ZEBRA (Hicks et al., 2003).

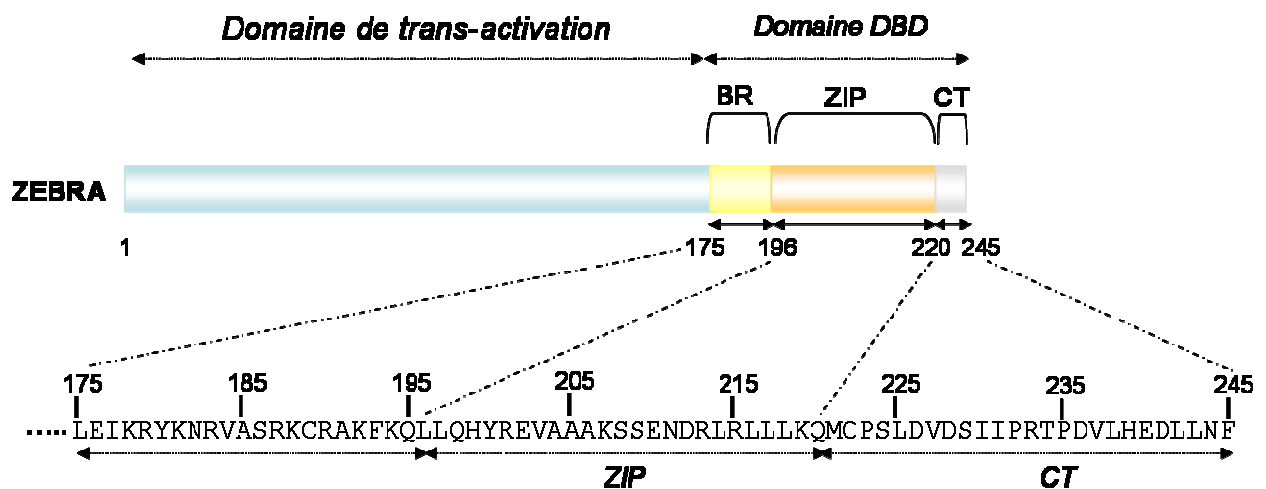


Figure I.10. : Représentation schématique de l'organisation en domaines de ZEBRA.

Le domaine de *trans*-activation N-terminal est représenté en bleu. Les régions BR (région basique), ZIP (région de dimérisation « *leucine zipper* ») et CT (queue C-terminale) du domaine DBD sont représentées respectivement en jaune, orange et gris. Les séquences en acides aminés de ces régions sont indiquées sous le schéma.

2. Analyse du complexe ZEBRA/AP1 et études de mutagenèse de ZEBRA

• STRUCTURE DU COMPLEXE ZEBRA/AP1 :

Le fragment C-terminal de ZEBRA (résidus 175 à 236) a été cristallisé par Carlo Petosa et ses collaborateurs en complexe avec un site AP1 provenant du promoteur du gène lytique précoce de l'EBV BSLF2/BMLF1 (Petosa et al., 2006). Comme les autres protéines bZIP, l'homodimère de ZEBRA saisit l'ADN via ses deux longues hélices α N caractéristiques des bZIPs, la région basique de chaque hélice α formant des contacts avec le sillon majeur de la double hélice d'ADN et la région leucine zipper formant un « *coiled-coil* » (autrement dit, un super enroulement d'hélices α) (figure I.11.).

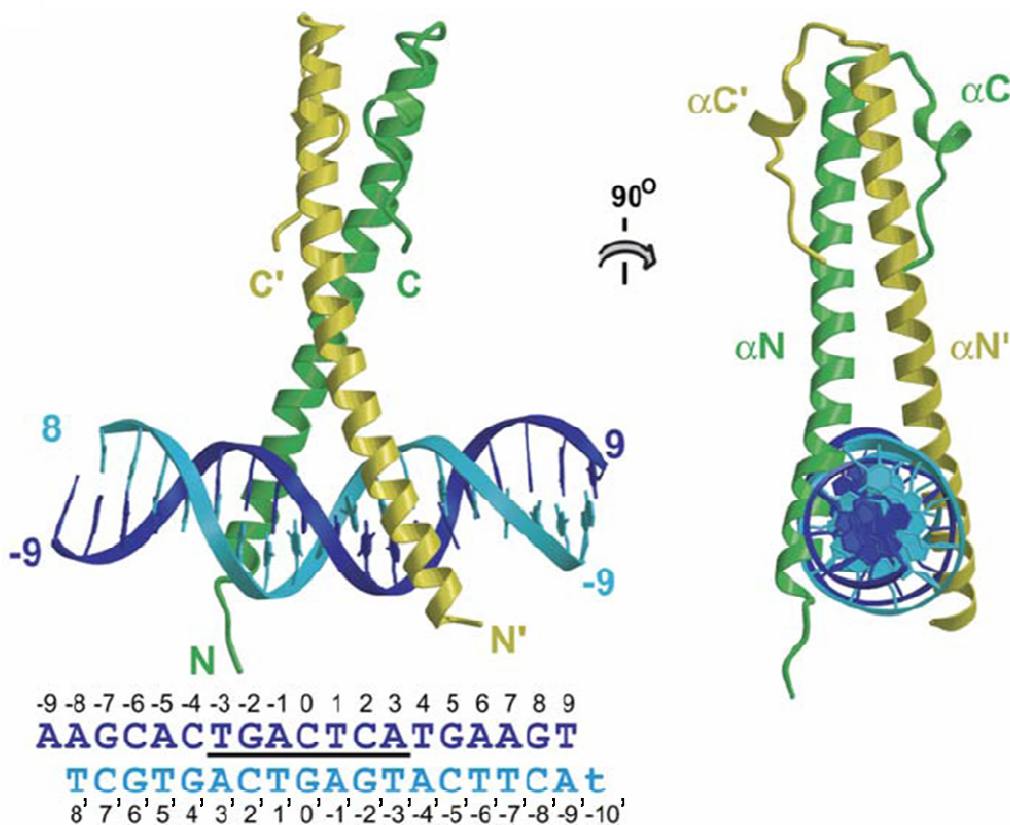


Figure I.11. : Vues orthogonales du complexe ZEBRA/AP1. Le *coiled-coil* est légèrement courbé : l'une des hélices du bZIP (en jaune) est courbée et l'autre hélice (en vert) est essentiellement droite. La séquence nucléotidique ainsi que le code numérique qui leur est attribué est indiqué ; l'heptamère AP1 est souligné.

De façon surprenante cependant, les résidus de la région C-terminale de ZEBRA (queue CT) forment un motif structural additionnel (figure I.10.). L'hélice α N de chaque monomère se coude au niveau du résidu Pro223 pour redescendre de façon anti-parallèle au *coiled-coil* et former une hélice à un tour appelée « hélice α C ». La queue CT forme ainsi de nombreuses interactions avec les hélices α N du *coiled-coil* permettant ainsi de le stabiliser et d'augmenter considérablement l'interface de dimérisation. Ceci explique la stabilité de l'homodimère de ZEBRA en dépit de l'absence d'un motif leucine zipper conventionnel, et également l'incapacité à lier l'ADN entraînée par des délétions affectant l'intégrité de cette queue CT (Hicks et al., 2003).

Un alignement structural entre ZEBRA et l'hétérodimère Fos-Jun (Glover and Harrison, 1995) montre que le *coiled-coil* de Fos-Jun forme 4 tours d'hélices supplémentaires à partir du résidu où la chaîne polypeptidique de ZEBRA change de direction. (figure I.12.).

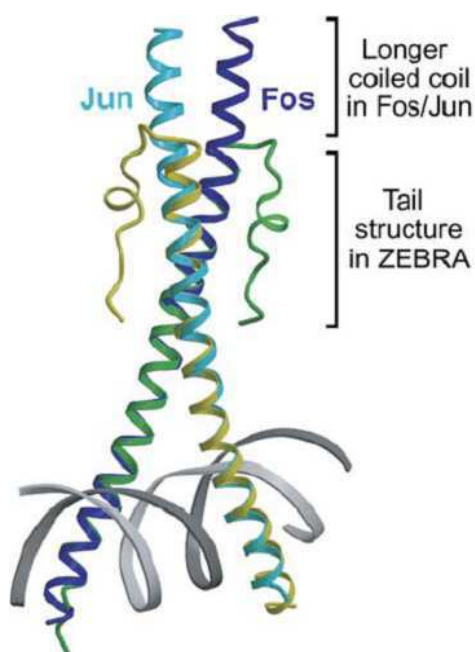


Figure I.12. : Structure du complexe ZEBRA/AP1 : comparaison avec l'hétérodimère bZIP Fos-Jun par alignement structural ($\text{rmsd}_{85\text{Ca}}=0.91\text{\AA}$).

Ainsi, ZEBRA a du évoluer à partir d'un ancêtre commun à Fos à travers des mutations (ex : l'introduction du résidu Proline en position 223 qui entraîne une cassure de l'hélice α N) qui ont contraint la queue CT de l'hélice bZIP de ZEBRA à redescendre le long du *coiled-coil* et à former de nouveaux contacts entre monomères. Cette structure ne ressemble à aucune autre protéine bZIP cellulaire d'après les banques de données, suggérant ainsi qu'elle serait uniquement présente dans la lignée virale.

• INTERFACE DIMERIQUE ET SELECTIVITE DE L'HOMODIMERE DE ZEBRA

Chez les protéines bZIP usuelles, les contacts entre hélices des deux monomères sont assurés par des résidus hydrophobes (a) et des résidus leucines (d) dans le motif répété d'acides aminés (abcdefg)_n (figure I.13.).

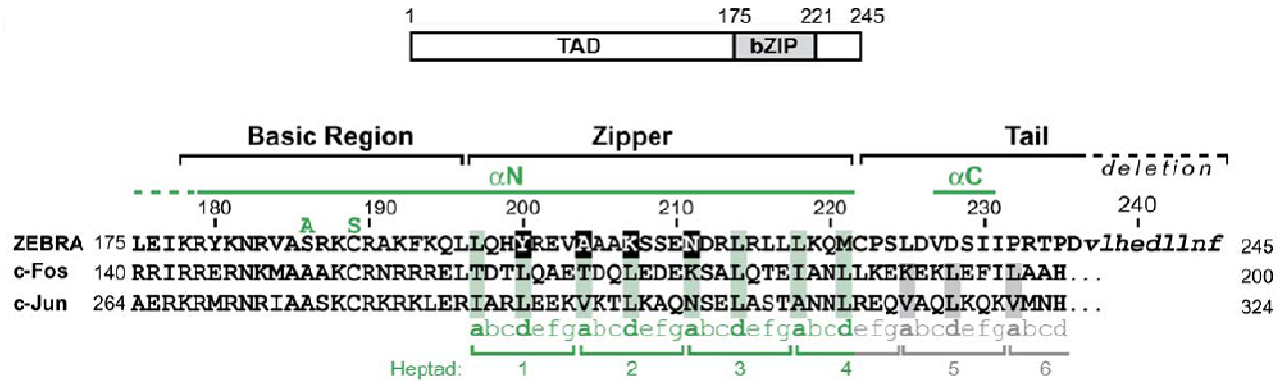


Figure I.13. : Alignement de séquences et structure secondaire. Les 4 résidus consécutifs « a » et « d » de ZEBRA déviant du motif consensus sont surlignés en noir. Les lignes vertes délimitent les structures en hélice ; l'hélice bZIP du monomère jaune de ZEBRA forme un tour d'hélice supplémentaire en N terminal (ligne verte en pointillés). Les mutations S186A et C189S sont indiquées. (TAD, « transcriptional activation domain »).

ZEBRA diffère de ce schéma classique au niveau de 4 positions (a) et (d) consécutives. Deux résidus en position (d), la Tyrosine 200 et la lysine 207, dirigent leur groupement hydrophobe vers l'extérieur de l'axe de l'hélice αN (voir figure I.14. ci-dessous), ce qui confère à la lysine 200 son exposition au solvant et à la lysine 207 de former un pont salin inter-monomères avec la queue CT de ZEBRA (voir figure I.14. ci-dessous).

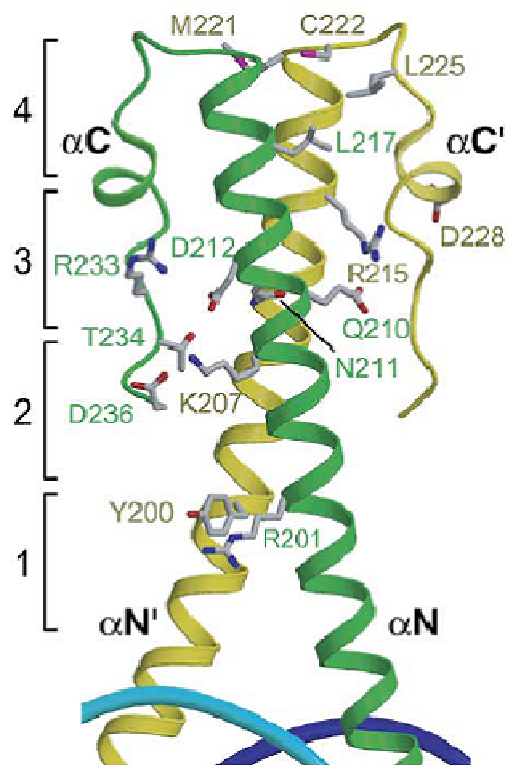


Figure I.14. : Structure du complexe ZEBRA/ADN : mise en évidence des résidus d'acides aminés présents à l'interface des deux monomères. Les 4 répétitions du motif « abcdefg » caractéristiques des protéines bZIP sont indiquées.

La tyrosine 200 est impliquée dans des interactions avec de nombreux facteurs de transcription. Au sommet du *coiled-coil*, les deux cystéines 222 positionnent leurs groupements thiols à 3,9 Å l'un de l'autre, ne leur permettant pas de former un pont disulfure. Cependant, ces deux cystéines 222 contribuent partiellement à la sensibilité redox de l'activité de liaison à l'ADN de ZEBRA (Wang et al., 2005).

L'homodimère de ZEBRA est stabilisé par une large interface hydrophobe impliquant les hélices αN et αC (figure I.15.).

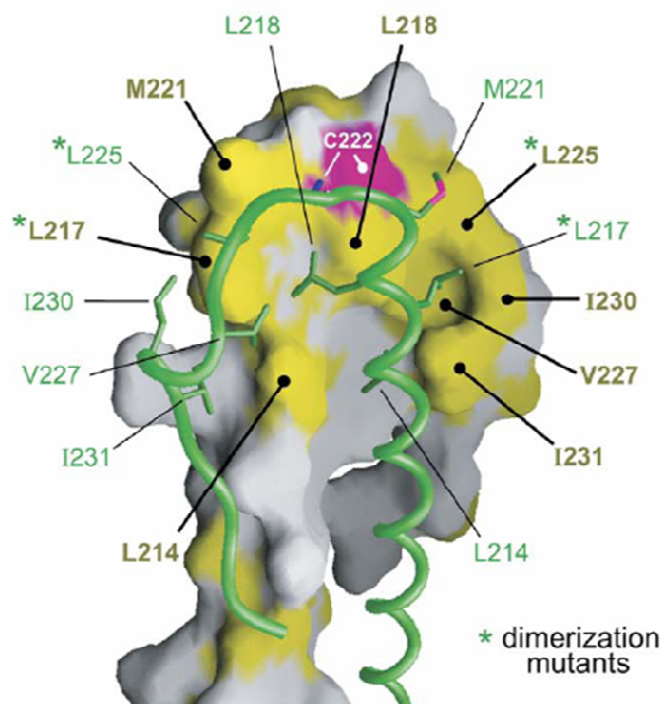


Figure I.15. : Résidus hydrophobes à l'interface dimérique de ZEBRA. La surface occupée par les chaînes latérales des résidus hydrophobes est indiquée en jaune et celle occupée par la cystéine 222, en magenta.

Cinq résidus hydrophobes (Leu218, Leu225, Val227, Ile230 et Ile231) forment une poche à la surface de chaque monomère. Le résidu leucine 217 du deuxième monomère s'insère profondément dans la poche hydrophobe du premier monomère alors que les résidus hydrophobes adjacents (Leu214, Leu218 et Met221) interagissent avec la surface de cette poche. Cette reconnaissance très spécifique de la leucine 217 fait de ce résidu un épitope clé de l'interface dimérique.

ZEBRA ne peut pas former d'hétérodimères avec les protéines bZIPs cellulaires cFos, cJun, C/EBP α et CREB (Chang et al., 1990; Kouzarides et al., 1991; Wu et al., 2004), permettant ainsi à ZEBRA d'exploiter les voies cellulaires d'activation de la transcription indépendamment de ces protéines. L'instabilité de tels hétérodimères est partiellement due à des contacts inter-hélices insuffisants dans le *coiled-coil* ainsi formé et à l'exposition au solvant de résidus hydrophobes du domaine C-terminal du monomère bZIP cellulaire (la région de dimérisation de ZEBRA étant plus courte). Enfin, de façon surprenante, ce résidu leucine 217 de ZEBRA correspond à un résidu polaire ou chargé chez presque toutes les

autres protéines bZIPs usuelles, suggérant que la reconnaissance très spécifique de ce résidu hydrophobe unique à ZEBRA au niveau de l'interface dimérique favoriserait la formation d'homodimères au détriment d'hétérodimères.

● SPECIFICITES DE ZEBRA POUR SES SITES ADN CIBLES

ZEBRA a une plus grande variété de sites ADN cibles que les autres protéines bZIPs, lui permettant ainsi d'activer de nombreux gènes cellulaires et viraux. La sous-famille de protéines bZIP Fos/Jun reconnaît l'heptamère AP1 « TGA^G/C TCA » alors que la sous-famille de protéines bZIP C/EBP reconnaît l'octamère « TTGCGCAA » ; ZEBRA peut se fixer sur les deux sites AP1 et C/EBP, ainsi que sur d'autres sites ZREs dérivant de ceux-ci (Farrell et al., 1989; Kouzarides et al., 1991). Le site ZRE utilisé par Carlo Petosa dans la structure de ZEBRA/AP1 contient le site heptamère AP1. Ce dernier est reconnu par ZEBRA en formant les mêmes contacts « bases-spécifiques » que Fos/Jun : l'asparagine 182 interagit avec la cytosine C^{2'} et la thymine T³, et l'arginine 190 avec la base centrale G⁰ (figure I.16. A et B).

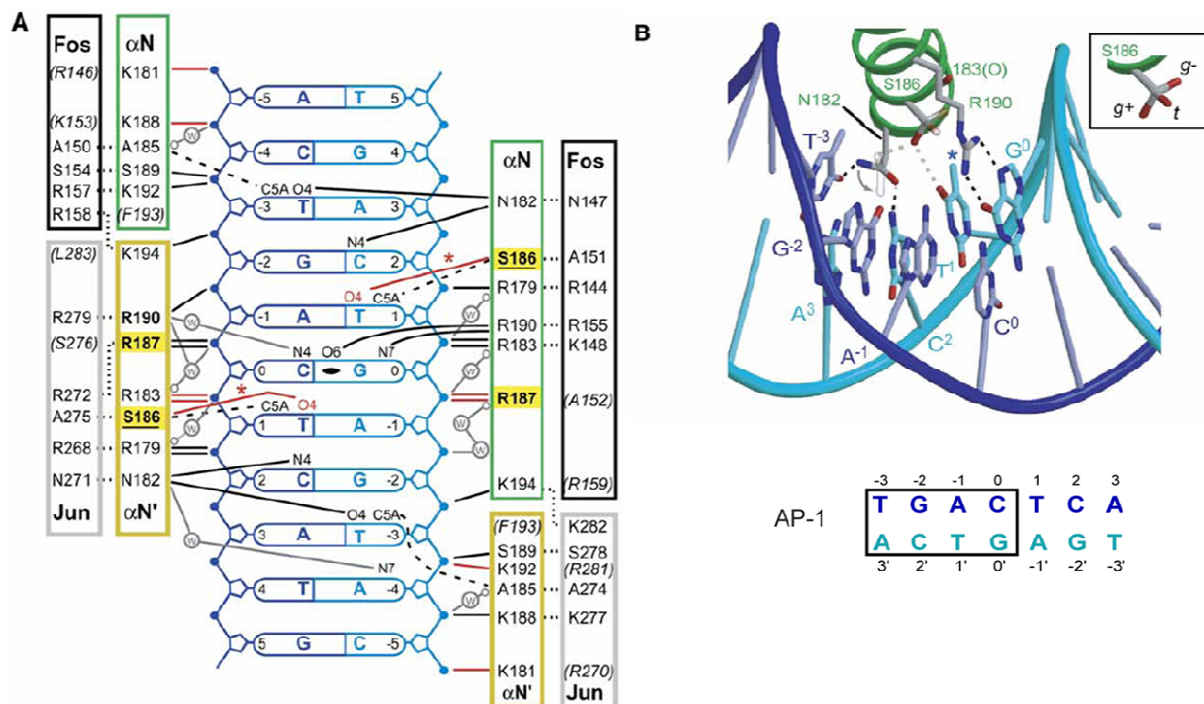


Figure I.16. : Contacts protéine/ADN du complexe ZEBRA/AP1 au niveau du sillon majeur d'ADN.

(A) Résumé des contacts Protéine/ADN. Les résidus de ZEBRA (encadrés en jaune et en vert) sont listés à côté de leurs homologues chez Fos/Jun (encadrés en noir et en gris). Les résidus effectuant les mêmes interactions sont reliés entre eux en pointillés. Les résidus ne

formant aucune interaction sont en parenthèses. Les lignes continues représentent les liaisons hydrogènes et les discontinues, les contacts hydrophobes. Les molécules d'eau sont signalées par un « w ». Les interactions protéine-ADN uniques à ZEBRA sont en rouge. La liaison hydrogène Ser186/T^{1'} (astérisque rouge) est hypothétique, dépendant de la conformation *gauche*⁺ ou *trans* de la sérine 186 (B).

(B) Interactions de ZEBRA avec le site AP1 dans le sillon majeur de l'ADN. La chaîne latérale de la sérine 186 a été représentée selon ses trois rotamères préférentiels (t : *trans*, g : *gauche*). La capacité de la sérine 186 à diriger son groupement hydroxyle vers ou à l'opposé de la base en position 1' et celle de l'asparagine 182 à fournir un donneur ou un accepteur de lien hydrogène à la base en position 2' pourrait expliquer la tolérance de ZEBRA pour les paires de bases 1/1' et 2/2'.

La forme cristalline contenant la mutation S186A, on ne peut déterminer quel rotamère adopterait la sérine 186 ; cependant, des considérations stéréochimiques favorise une conformation de la sérine 186 dans laquelle elle pourrait former un lien hydrogène avec l'atome d'oxygène O4 de la thymine T^{1'} (figure I.16.B.).

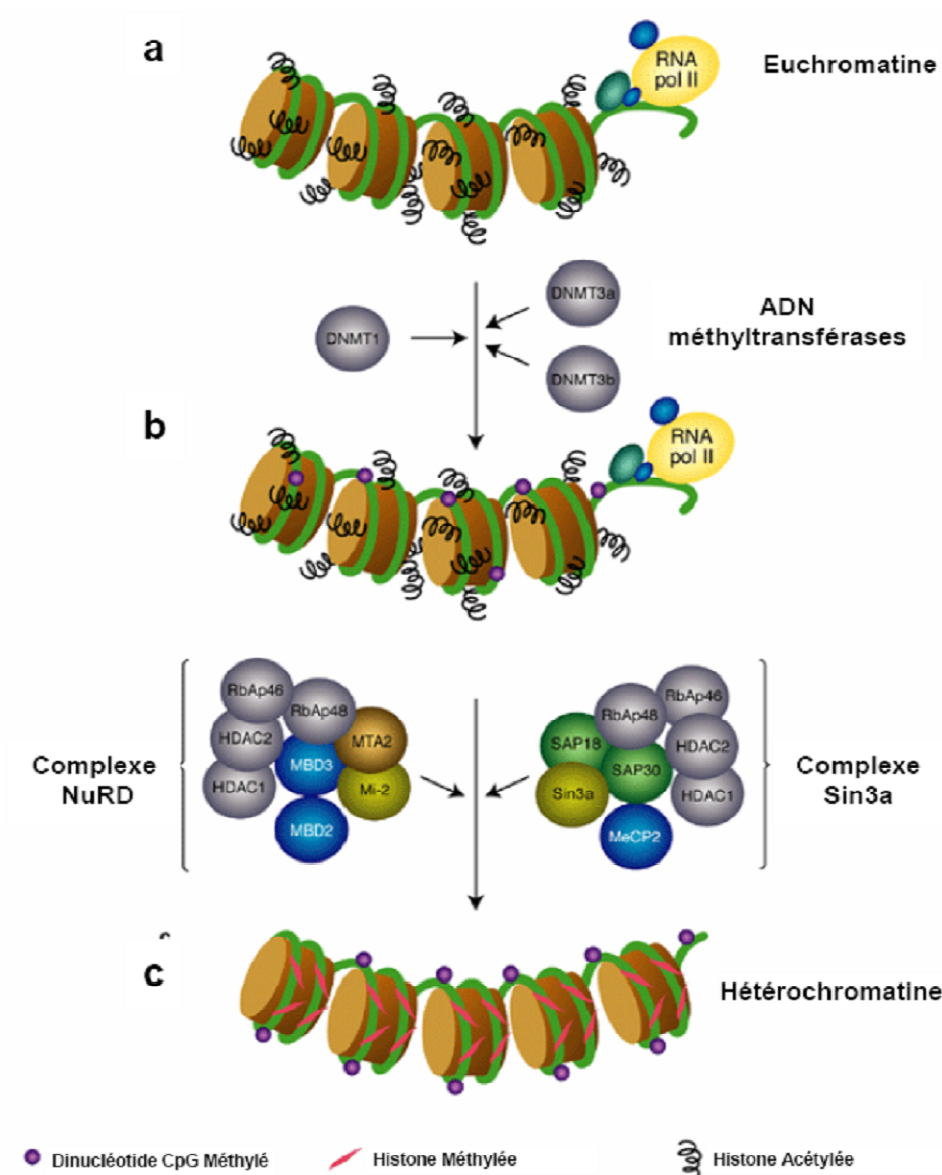
ZEBRA fixe des sites ZREs dont les nucléotides en position 1' et 2' varient considérablement alors que l'hétérodimère Fos/Jun a des exigences strictes concernant ces positions (Kouzarides et al., 1991). Ces nucléotides se trouvent respectivement à proximité des résidus sérine 186 et asparagine 182. L'asparagine 182 est universellement conservée alors que la sérine 186 est unique à ZEBRA : la plupart des protéines bZIPs ont une alanine et aucun n'a à cette position un résidu capable de former une liaison hydrogène. La seule exception notable concerne le résidu glutamine 90 de la protéine Pap-1. Cette glutamine influence la spécificité de fixation à l'ADN de Pap-1 en induisant une conformation alternative de l'asparagine 86 (équivalente à l'asparagine 182 de ZEBRA) (Fujii et al., 2000). La sérine 186 de ZEBRA jouerait un rôle similaire en permettant à l'asparagine 182 de fournir un groupement donneur de liaison hydrogène à la base en position 2'.

V. Activation des promoteurs méthylés par ZEBRA

1. Conséquences fonctionnelles de la méthylation de l'ADN

La méthylation de l'ADN est dite « épigénétique » parce qu'elle module l'activité du génome sans en affecter directement la séquence. Elle concerne principalement les cytosines. La méthylation des cytosines n'existe pas chez tous les organismes. *Saccharomyces cerevisiae* et *Caenorhabditis elegans* en sont, par exemple, totalement dépourvus. Dans le règne animal les cytosines méthylées se trouvent presque invariablement dans le dinucléotide CpG. Chez les vertébrés, ces sites méthylables suivent une distribution non uniforme: il existe des domaines, appelés îlots CpG, où ce dinucléotide est plus fortement représenté. Bien qu'identifiée dès 1948, la méthylation a été perçue pendant de longues années par la communauté scientifique comme un épiphénomène biologique. Cependant, les progrès réalisés au cours des dernières années ont permis au contraire de démontrer le rôle essentiel joué par la méthylation au cours de plusieurs processus biologiques associés au développement, comme l'inactivation du chromosome X chez les mammifères femelles, l'imprinting génomique ou encore l'expression de gènes spécifiques de tissus (Jones and Takai, 2001). En outre, il semble de plus en plus clair qu'une méthylation aberrante de l'ADN joue un rôle primordial dans le processus de cancérogenèse. En particulier, une méthylation altérée peut conduire à l'inactivation de gènes suppresseurs de tumeurs et à un processus de tumorigenèse comparable à celui induit par les mutations génétiques (Baylin and Herman, 2000).

La méthylation exerce une action directe sur l'expression des gènes en empêchant la fixation de protéines qui contrôlent la transcription. Ainsi les activateurs transcriptionnels Sp1, CREB, YY1, et d'autres encore, ne se lient plus à certaines de leurs cibles si elles sont méthylées. En effet, de récents développements suggèrent que le verrouillage de la transcription par méthylation de l'ADN est intimement connecté au remodelage de la chromatine par des enzymes de modification des histones (figure I.17.).



(Expert Reviews in Molecular Medicine©2002 Cambridge University Press)

Figure I.17. : Mécanisme par lequel la méthylation de l'ADN réprimerait l'expression des gènes.

Ces récentes découvertes ont permis l'ébauche d'un modèle de répression de la transcription. Dans un premier temps, les méthyltransférases de l'ADN, Dnmt, utilisent le S- adénosyl-méthionine comme substrat pour catalyser le transfert d'un groupement méthyle sur l'anneau de la cytosine avec formation de la cinquième base, la 5-méthylcytosine (figure I.18.).

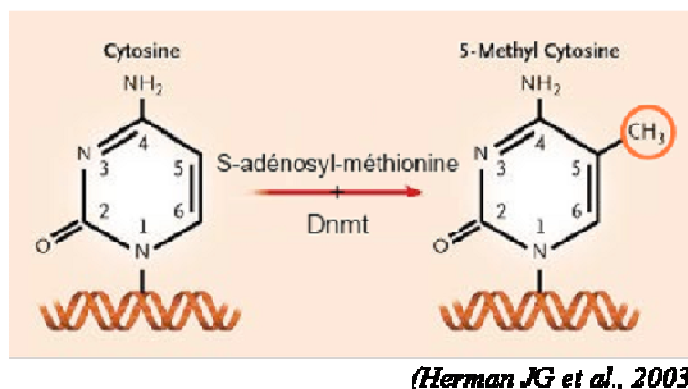


Figure I.18. : Réaction de transfert d'un groupement méthyle sur l'ADN.

Les Dnmt catalysent le transfert du résidu méthyle CH₃ à partir de la S-adénosyl-méthionine sur le cinquième carbone pyrimidique des cytosines, converties ainsi en 5-méthyl-cytosine.

Dans un deuxième temps, les protéines MBD (« Methyl-Binding-Domain Proteins ») se lient spécifiquement aux dinucléotides CpG méthylés, en s'associant à des complexes répresseurs de la transcription (complexes Sin3a ou NuRD) (Fuks, 2003). La famille MBD comprend cinq membres : MeCP2, MBD1, MBD2, MBD3, MBD4 (Hendrich and Bird, 1998). Tous les membres de la famille MBD, via leur domaine conservé MBD, peuvent fixer les dinucléotides CpG méthylés à l'exception de MBD3. Enfin, les histone-désacétylases (HDAC) sont recrutées au niveau des séquences génomiques méthylées, ce qui a pour effet de désacétyler les histones et de maintenir une structure condensée de la chromatine, réprimant ainsi la transcription (Jones et al., 1998; Nan et al., 1998).

Etant donné l'effet répresseur de la méthylation sur l'expression des gènes, on ne comprend pas comment certains gènes méthylés viraux ou cellulaires sont activés sous certaines circonstances. Au moins un virus, le virus Frog 3, code pour une fonction permettant une transcription efficace de gènes méthylés. Il serait avantageux pour les gammaherpès virus qui ont le plus souvent un génome méthylé de façon intensive durant leur phase de latence, de coder pour une ou plusieurs protéines capables d'activer des promoteurs viraux méthylés. Dans le cas de l'EBV, même si le promoteur Rp est hautement méthylé en phase de latence, ZEBRA active efficacement la transcription du gène BRLF1 (Bhende et al., 2004). En effet, Bhende P.M. *et al.* ont démontré que ZEBRA fixait et activait préférentiellement le promoteur Rp lorsque le motif CpG du site ZRE2 de ce promoteur était méthylé.

2. Reconnaissance préférentielle de promoteurs méthylés par ZEBRA

ZEBRA reconnaît et fixe préférentiellement les promoteurs viraux de l'EBV dans leur état méthylé (Bhende et al., 2004). Cette caractéristique unique à ZEBRA est cruciale pour l'activation du cycle lytique car le génome de l'EBV est hautement méthylé durant la phase latente. Plus précisément, la principale cible de ZEBRA est le gène précoce BRLF1, dont le promoteur Rp contient 3 sites ZREs (ZRE1, ZRE2 et ZRE3) dont deux sont méthylables (ZRE2 et ZRE3), car contenant un îlot CpG. Bhende P.M. *et al.* ont montré en 2004 que ZEBRA reconnaissait préférentiellement le site ZRE2 dans son état méthylé (figure I.19.).

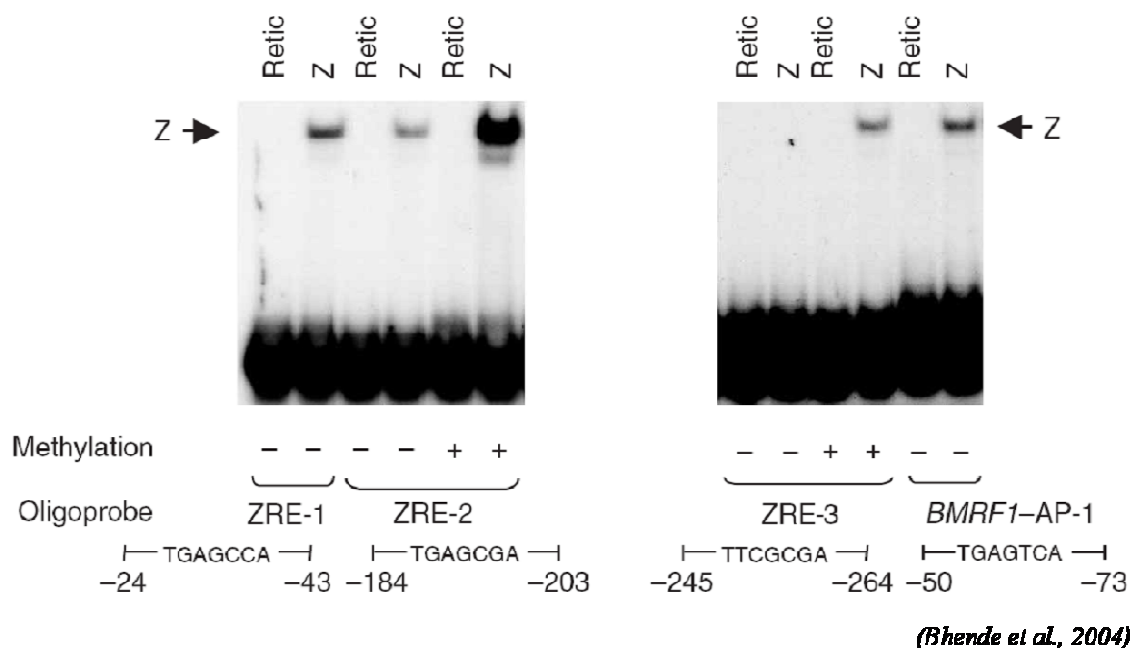


Figure I.19. : ZEBRA fixe préférentiellement le site ZRE2 dans son état méthylé.

Comparaison de la fixation de la protéine ZEBRA (traduite in vitro) sur des oligonucléotides marqués au ^{32}P contenant les sites cibles Rp ZRE1, Rp ZRE2, Rp ZRE3 ou AP1 du promoteur de l'EBV BMRF1. Un lysat de réticulocyte de lapin non traduit sert de contrôle négatif pour chaque oligonucléotide (« Retic »).

Le site ZRE2 contient la séquence TGAGCGGA (souligné : îlot CpG). Bhende P.M. *et al.* ont estimé que ZEBRA fixait ce site ZRE2 méthylé avec dix fois plus d'affinité que l'équivalent non méthylé. Cependant, ZEBRA ne peut se fixer au site mutant ZRE2^{mut} : TGGACGG, que ce soit dans son état méthylé ou non. Ceci indique que ZEBRA ne peut pas reconnaître n'importe quel site contenant un îlot CpG méthylé (Bhende et al., 2004).

Concernant le deuxième site ZRE méthylable du promoteur Rp, ZRE3 (TTCGCGA), il n'est reconnu par ZEBRA que dans son état méthylé (figure I.19.).

Un deuxième article de Bhende P.M. *et al.* en 2005 ont permis d'identifier les résidus de ZEBRA jouant un rôle essentiel dans la reconnaissance préférentielle du site meZRE2 (Bhende et al., 2005). En se basant sur la structure cristallographique de l'hétérodimère Fos-Jun en complexe avec le site AP1, ils ont émis l'hypothèse que la sérine 186 formerait une liaison directe avec la cytosine méthylée du motif meZRE2 (TGAGCGA). Pour vérifier cette hypothèse, ils ont testé la fixation des mutants ZEBRA S186T et ZEBRA S186A sur les sites meZRE2 et ZRE2. Le mutant ZEBRA S186T garde la capacité à fixer le site meZRE2 mais ne reconnaît plus le site ZRE2 non méthylé. Le mutant ZEBRA S186A, quant à lui, ne fixe ni le site meZRE2 ni le site ZRE2 non méthylé. La sérine et la thréonine sont des acides aminés très semblables contenant un groupement hydroxyle. Ceci peut expliquer le fait que le mutant ZEBRA S186T est toujours capable de reconnaître le site meZRE2. Contrairement à cela, la mutation de la sérine en alanine, un acide aminé non polaire, ne permet plus à ZEBRA de fixer le site ZRE, qu'il soit méthylé ou non. Une autre explication serait que différentes conformations de ZEBRA sont requises pour la reconnaissance des sites meZRE2 et ZRE2 non méthylé et que seule ZEBRA WT est capable d'adopter ces différentes conformations.

D'après la structure cristallographique du complexe ZEBRA/AP1, Petosa C. *et al.* ont proposé une hypothèse sur le mécanisme de reconnaissance préférentielle des promoteurs méthylés par ZEBRA. Le site ZRE2 diffère du site AP1 au niveau des positions 1' et 2' (nucléotides soulignés dans la séquence du site ZRE2) qui constituent le motif CpG, méthylé sur la cytosine C^{1'}. Tel que l'évoque la structure du complexe ZEBRA/AP1, la présence de ce motif CpG ne permettrait plus à l'asparagine 182 de former une liaison hydrogène avec la base en position 2' comme observé dans la structure ZEBRA/AP1. Cependant, la méthylation de cet îlot CpG pourrait stabiliser la sérine 186 dans une conformation lui permettant d'interagir avec l'asparagine 182 ce qui, par conséquent permettrait à cette dernière de former un lien hydrogène avec la guanine G^{2'}, augmentant ainsi l'affinité de ZEBRA pour le site meZRE2 (figure I.16.A et B et figure I.20.).

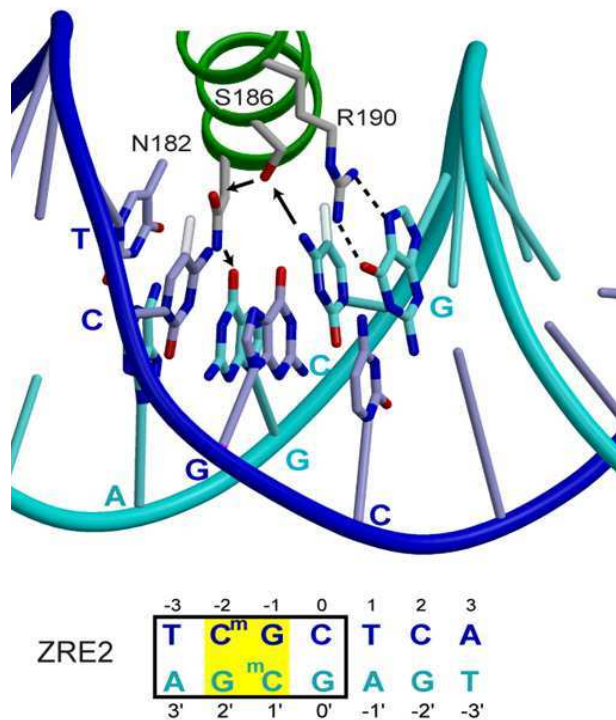


Figure I.20. : Prédiction de structure du complexe ZEBRA/meZRE2. La sérine 186, stabilisée par le groupement méthyle de la cytosine C⁻², interagirait avec l'asparagine 182 qui, à son tour, formerait une liaison hydrogène avec la guanine G^{2'}.

Cette hypothèse est en accord avec l'observation selon laquelle le mutant ZEBRA S186A ne peut fixer aucun des sites ZRE2 méthylé ou non, et que le mutant ZEBRA S186T est capable de reconnaître le site meZRE2 (Bhende et al., 2005). De plus, cette hypothèse expliquerait également la meilleure affinité entre le mutant ZEBRA S186A et le site AP1 (Francis et al., 1999) (figure I.16.B. : légende).

OBJECTIFS DU TRAVAIL DE THESE

Les récentes avancées scientifiques ont démontré que ZEBRA était le seul facteur de transcription connu à ce jour capable de *trans*-activer préférentiellement des séquences d'ADN virales méthylées (Bhende et al., 2004). De plus, la structure cristallographique du domaine bZIP de ZEBRA en complexe avec le site AP1 (Petosa et al., 2006) montre que ZEBRA adopte une structure tridimensionnelle tout à fait atypique au vu des autres membres de la famille des protéines bZIPs tels que l'hétérodimère cFos/cJun.

Or, ZEBRA est le facteur responsable de la transition entre phase latente et phase lytique du virus de l'EBV. Cette fonction primordiale ajoutée à l'unique capacité de ZEBRA à fixer préférentiellement des promoteurs viraux méthylés et à la structure particulière de la protéine ZEBRA, fait de l'élucidation du mécanisme de reconnaissance de l'ADN méthylé par ZEBRA un atout essentiel pour la compréhension de l'activation du cycle lytique du virus de l'EBV.

Mon travail de thèse a tout d'abord consisté à obtenir la structure cristallographique du domaine bZIP de ZEBRA en complexe avec un site ZRE du promoteur viral BRLF1 que ZEBRA fixe préférentiellement dans son état méthylé : le site meZRE2. D'après la structure du complexe ZEBRA/AP1 obtenue précédemment, Petosa C. *et al.* ont proposé une hypothèse quant au mécanisme de reconnaissance de l'ADN méthylé par ZEBRA. Des études de mutagenèse de la protéine ZEBRA et du site meZRE2 ainsi que des expériences de mesures d'affinité de ZEBRA pour différents sites ZREs ont été prévues afin de compléter ou de confirmer cette hypothèse.

Enfin, la structure tridimensionnelle hors norme de la protéine ZEBRA a mis en évidence, sur chaque monomère, la présence d'une poche hydrophobe capable d'accueillir une petite molécule chimique suggérant la drugabilité de cette protéine virale. Nous avons ainsi prévu un criblage de chimiothèque afin d'identifier un inhibiteur spécifique de la protéine ZEBRA qui permettrait de bloquer la transition entre phase de latence et phase lytique du virus EBV.

CHAPITRE II :

PREPARATION ET CRISTALLISATION

DE ZEBRA EN COMPLEXE AVEC UN

SITE ZRE2 METHYLE

I. Stratégie de cristallisation du complexe ZEBRA/ZRE2 méthylé

1. Choix du clone de ZEBRA

La structure cristallographique du bZIP de ZEBRA en complexe avec le site AP1 a été résolue en 2006 (Petosa et al., 2006). Les premiers essais de cristallisation ont été réalisés à partir du domaine C-terminal de ZEBRA, du résidu 175 au résidu 245. Ces premiers essais se révélant infructueux, 9 résidus (VLHEDLLNF) en C terminal de leur clone ont été délétés. En effet, 5 résidus sur les 9 délétés étaient hydrophobes : leur élimination a permis d'accroître la solubilité de la protéine. Le clone ZEBRA^{175-236(mut)} contenait également deux mutations ponctuelles : S186A et C189S. La mutation de la cystéine 189 en sérine protège la protéine contre les phénomènes d'oxydation et permet d'améliorer la qualité des cristaux comme l'ont montré Glover et Harrison (Glover and Harrison, 1995) lors de la cristallisation du complexe Fos-Jun/DNA. La mutation de la sérine 186 en alanine avait pour objectif de rendre la séquence primaire de ZEBRA plus similaire à celle de Fos et Jun (puisque seule ZEBRA possède une sérine à cette position 186, voir chapitre I.IV.2.).

Cependant, il a été démontré que cette sérine jouait un rôle crucial dans la reconnaissance préférentielle du site ZRE2 dans son état méthylé, appelé site meZRE2 (Bhende et al., 2005). Nous avons donc décidé de restituer cette sérine 186 et de ne garder que la mutation ponctuelle C189S. Les essais de cristallisation sont donc réalisés à partir du clone « ZEBRA C189S¹⁷⁵⁻²³⁶ ». Nous ferons référence à ce clone sous le nom simplifié de « ZEBRA ». La séquence en acides aminés de ce clone est présentée ci-dessous (figure II.1.) :

175	186	196
LEIKRYKNRVA	SRKCR	AKFKQ LLQHYREVAA
206	216	226
AKSS	ENDRLR LLLKQ	MCP
		SL DVDSIIPRTPD

Figure II.1: Séquence en acides aminés du clone ZEBRA. En jaune : région basique. En orange : domaine de dimérisation.

2. Choix des oligonucléotides contenant le site ZRE2

Différents duplexes oligonucléotides contenant l'heptamère ZRE2 sous forme méthylée (meZRE2) ou non méthylée (ZRE2) ont été testés (tableau II.1.).

nom	séquence nucléotidique
meZRE2(a)	-3 -2 -1 0 1 2 3 5', T A C T T C A 3', T G A A G T T C^m G C A G C G -3 -2 -1 0 1 2 3 T C A G T G C T A G T C A C G A
meZRE2(b)	-3 -2 -1 0 1 2 3 5', A T A A A A 3', A T T T T T C^m G C A G C G -3 -2 -1 0 1 2 3 T C A T A A G C T A G T A T T C G A
ZRE2(a)	-3 -2 -1 0 1 2 3 5', T A C T T C A 3', T G A A G T T C G C A G C G -3 -2 -1 0 1 2 3 T C A G T G C T A G T C A C G A
ZRE2(b)	-3 -2 -1 0 1 2 3 5', A T A A A A 3', A T T T T T C G C A G C G -3 -2 -1 0 1 2 3 T C A T A A G C T A G T A T T C G A

Table II.1 : récapitulatif des différents oligonucléotides testés pour la cristallisation du complexe ZEBRA/ADN. meZRE2(a) et ZRE2 (a) : oligonucléotides double brin de 19 bp contenant le site ZRE2 sous forme méthylée (meZRE2) ou non (ZRE2) et présentant les séquences flanquantes utilisées pour la résolution de la structure ZEBRA/AP1. meZRE2(b) et ZRE2 (b) : oligonucléotides double brin de 19 bp contenant le site ZRE2 sous forme méthylée (meZRE2) ou non (ZRE2) et présentant les séquences flanquantes naturelles du site ZRE2 dans le promoteur viral du gène *BRLF1* du génome de l'EBV.

Nous avons joué sur deux caractéristiques :

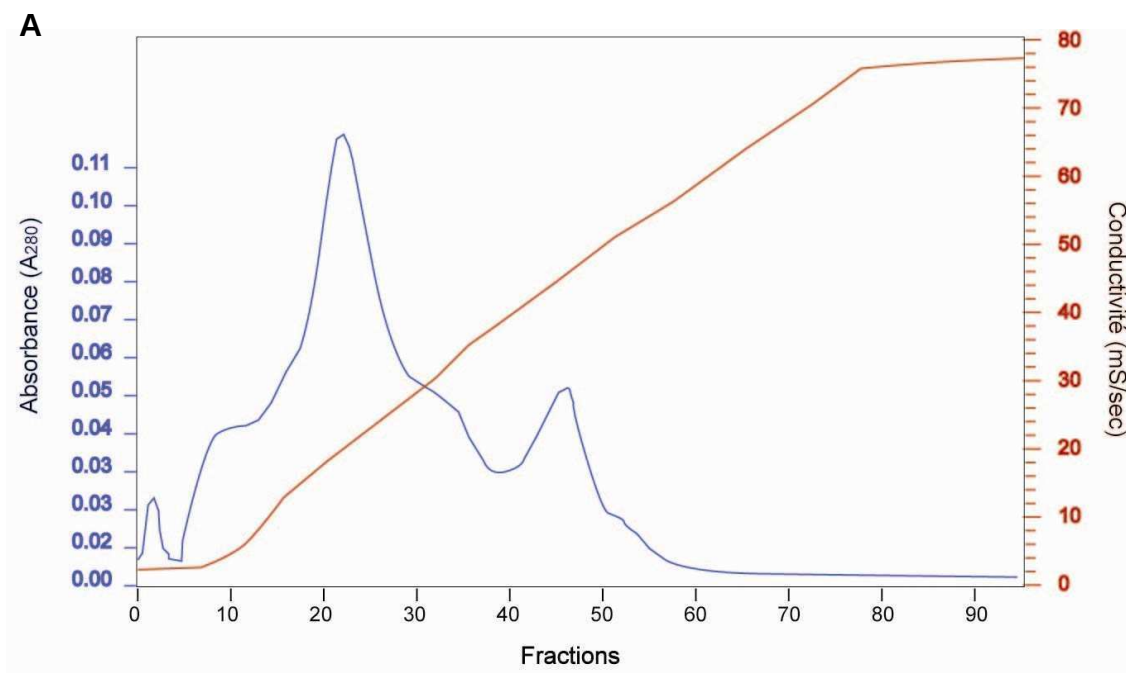
- la méthylation ou non des cytosines en position -2/2' et -1/1',
- la nature des paires de bases flanquant l'heptamère ZRE2. Ces séquences flanquantes représentent dans un cas celles présentes dans le duplexe oligonucléotide de la structure ZEBRA/AP1 (oligonucléotides « meZRE2(a) », et « ZRE2(a) »), et dans l'autre cas représentent les paires de bases originelles flanquant le site ZRE2 dans le promoteur viral du gène *BRLF1* du génome de l'EBV (oligonucléotides « meZRE2(b) », et « ZRE2(b) »).

De plus, tous les duplexes oligonucléotides testés présentent des extrémités cohésives afin de favoriser les contacts cristallins.

II. Purification du complexe ZEBRA/ZRE2 méthylé

1. Purification de ZEBRA

La première étape de purification est une chromatographie échangeuse de cations réalisée sur une résine SP Sépharose (*GE Healthcare*). Après élution à l'aide d'un gradient NaCl, les fractions contenant la protéine ZEBRA (figure II.2.) sont rassemblées et soumises à une précipitation au sulfate d'ammonium.



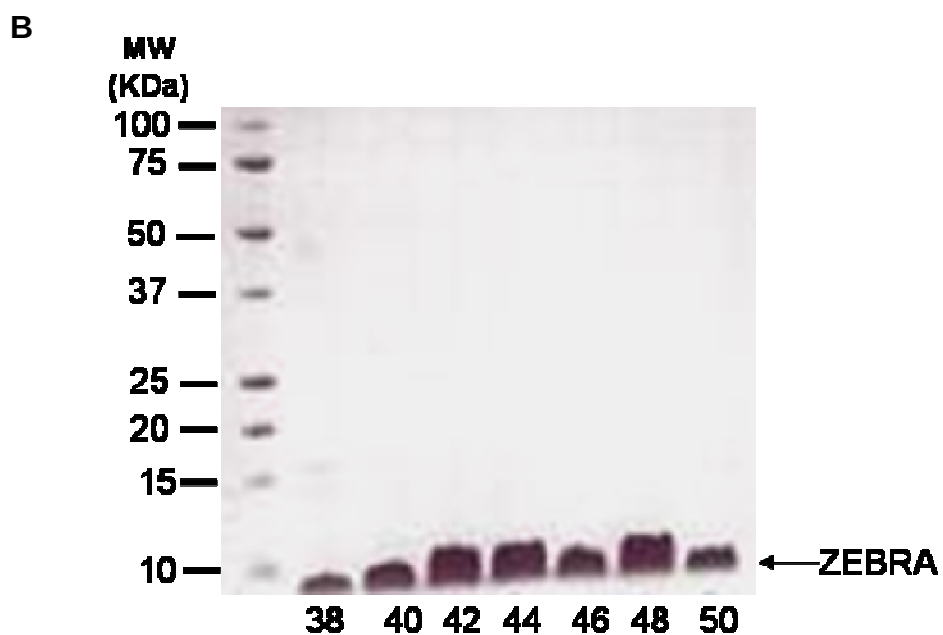


Figure II.2 : Purification de ZEBRA

(A) Profil d'élution de la protéine sur résine échangeuse de cations.

(B) Gel SDS-PAGE des fractions issues de la première étape de purification de ZEBRA par chromatographie échangeuse de cations. Les fractions analysées (38 à 50) sont de 1ml.

Après resuspension du culot de protéines précipitées au sulfate d'ammonium dans le tampon GF (voir Annexe C.2.3.), la protéine ZEBRA subit une ultime étape de purification par chromatographie d'exclusion de taille sur colonne Superdex S75 (*Pharmacia/GE Healthcare*) (figure II.3.) :

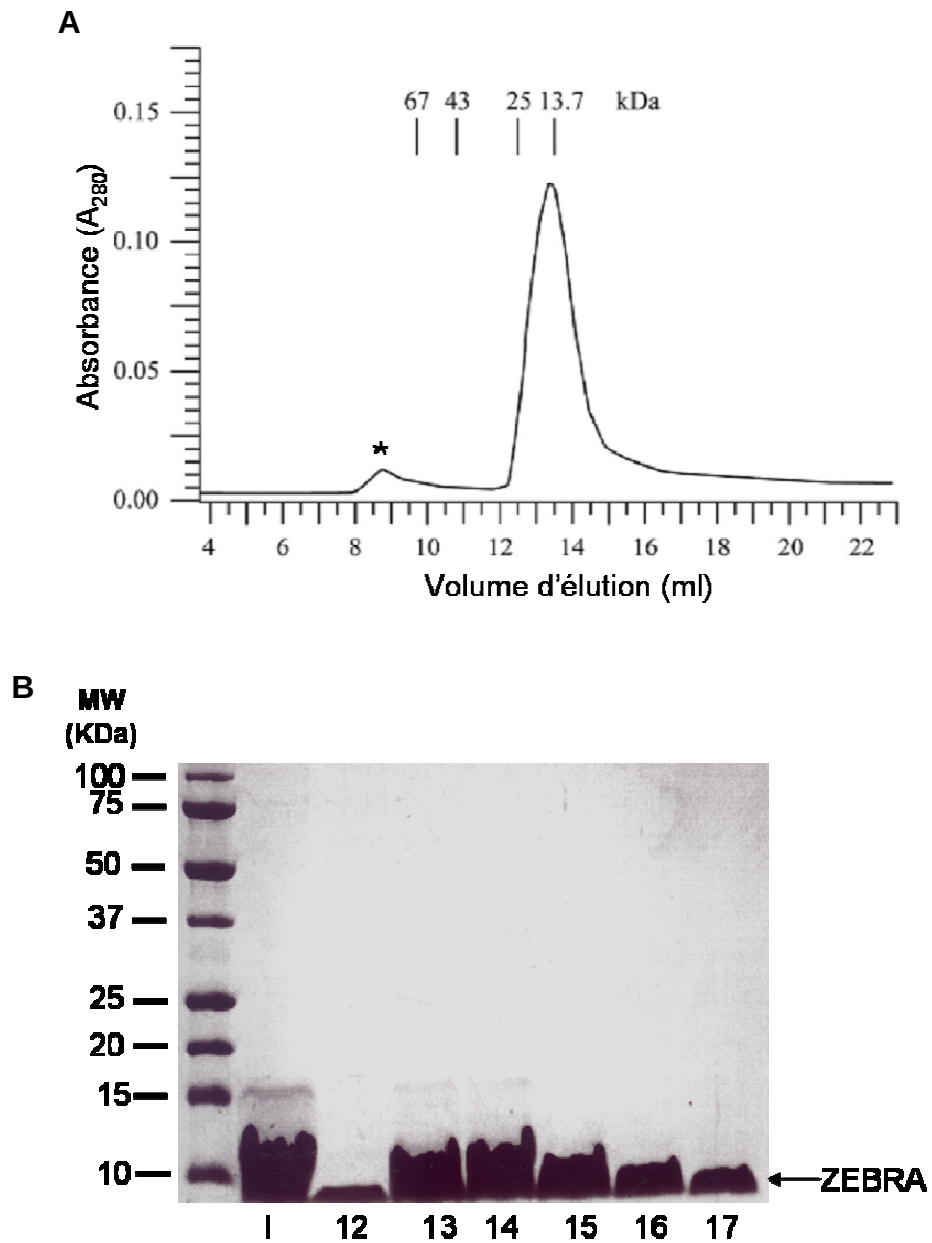


Figure II.3 : Dernière étape de purification de ZEBRA par chromatographie d'exclusion de taille sur colonne S75.

(A) Profil d'éluion de la protéine de ZEBRA. En haut sont indiqués les poids moléculaires de protéines de référence (67, 43, 25 et 13,7 kDa). * : pic correspondant à la protéine ZEBRA en complexe avec de l'ADN contaminant.

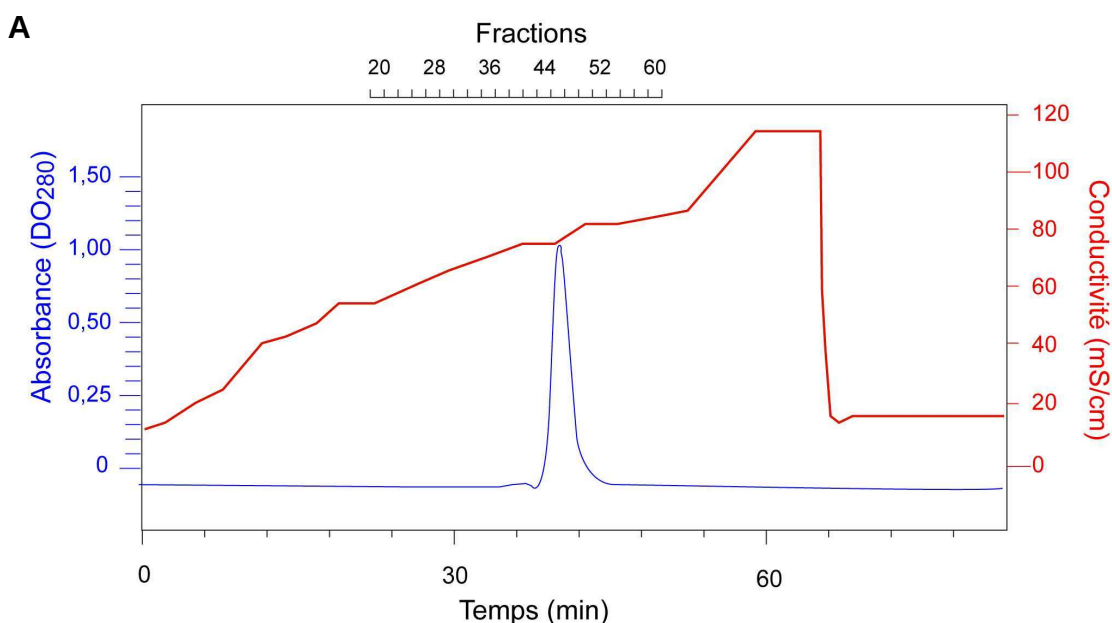
(B) Gel SDS-PAGE des fractions issues de la dernière étape de purification de ZEBRA par chromatographie d'exclusion de taille. I : input.

Le profil d'élution de la protéine ZEBRA se présente sous la forme d'un pic unique correspondant à un poids moléculaire apparent d'environ 15 kDa (figure II.3.A.) soit approximativement deux fois le poids moléculaire calculé pour un monomère de ZEBRA (7300 kDa). Ceci indique que ZEBRA existe en tant qu'homodimère en solution. Toutefois, un petit pic se forme en amont (indiqué par une astérisque aux volumes d'élution 8 à 10 ml). Après analyse des fractions d'élution de ce pic, nous pouvons affirmer qu'il s'agit de la protéine ZEBRA en complexe avec de l'ADN contaminant issu de la lyse des cellules *E.coli* BL21(DE3) ayant servi à exprimer la protéine.

La protéine ZEBRA est très pure après ces deux étapes de purification comme le montre le gel SDS-PAGE des fractions d'élutions issues de la chromatographie d'exclusion de taille (figure II.3.B.).

2. Purification des oligonucléotides contenant le site ZRE2

Les oligonucléotides sont synthétisés sous forme simple brin et concernant l'oligonucléotide meZRE2, nous le recevons déjà méthylé. Chaque oligonucléotide sous forme simple brin est purifié par chromatographie échangeuse d'anions sur colonne MonoQ HR10/10 (*Pharmacia/GE Healthcare*) (figure II.4.) :



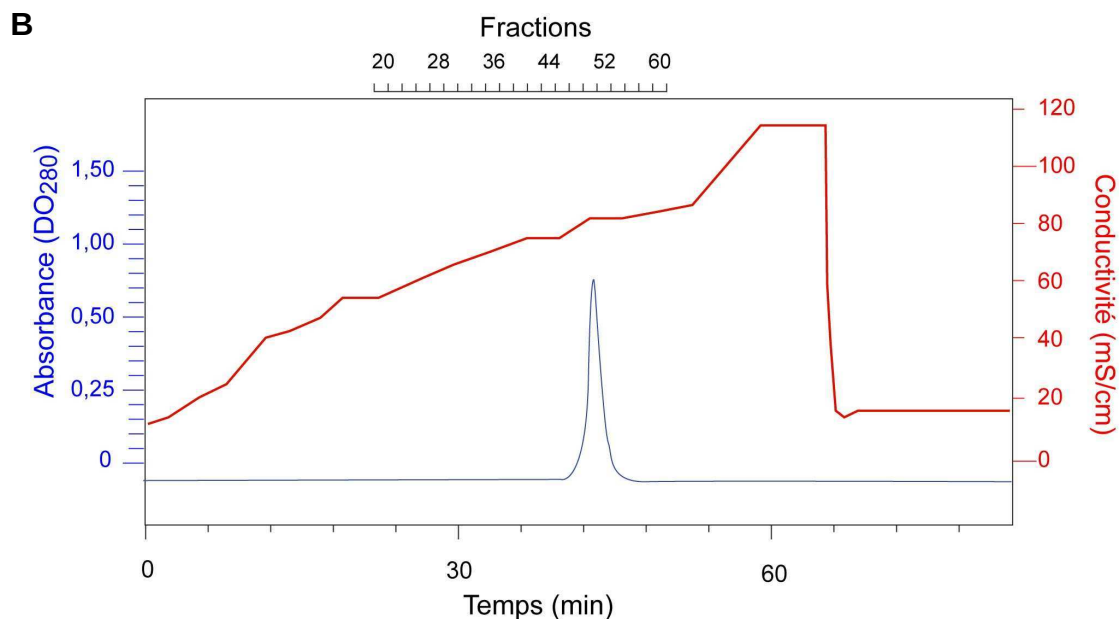


Figure II.4 : Purification des oligonucléotides simple brin.

en bleu : profil d'élution des oligonucléotides meZRE2 simple brin 1(A) et 2 (B) sur colonne MonoQ HR10/10. en rouge : courbe de conductivité du tampon.

Après élution par un gradient NaCl, les fractions collectées sont rassemblées, dialysées et dessalées. Les simples brins d'oligonucléotides complémentaires sont appariés pour former un duplexe d'oligonucléotides.

III. Cristallisation du complexe ZEBRA/ZRE2 méthylé

Avant de procéder aux expériences de cristallogenèse, l'interaction entre ZEBRA et chaque duplexe oligonucléotide est testée par gel retard (ou EMSA).

Seuls les complexes contenant les duplexes oligonucléotides meZRE2(a) et ZRE2(a) (c'est-à-dire contenant les séquences flanquantes issues de la structure ZEBRA/AP1) ont permis d'obtenir des cristaux (figure II.6.). Les conditions de cristallisation de ces derniers ainsi que les différents ratios protéine/ADN testés sont résumés dans le tableau II.2. :

Ratio protéine/ADN	Nom	Conditions de cristallisation
1/1	I	22 % PEG 4000 18 % PEG 400 Tampon AA pH 4,0
	II	12 % PEG 4000 15 % PEG 400 + 20 % isopropanol Tampon AA pH 3,8
1,4/1	III	25 % PEG 4000 10 % PEG 400 Tampon AA pH 4,0
1/1,2	IV	20 % PEG 4000 15 % PEG 400 Tampon AA pH 3,8
1/1	V	20 % PEG 4000 15 % PEG 400 Tampon AA pH 3,8
1/1,2	VI	12 % PEG 4000 15 % PEG 400 Tampon AA pH 4,2

Table II.2 : Résumé des conditions de cristallisation des complexes ZEBRA/meZRE2(a) et ZEBRA/ZRE2(a). AA : tampon acétate de sodium/acide acétique à 50 mM.

Quelque soit l'oligonucléotide testé, le mélange ZEBRA/ADN entraîne l'apparition d'un culot de précipitation. Après centrifugation de ce mélange, le culot et le surnageant ont été analysés par gel retard (figure II.5.) :

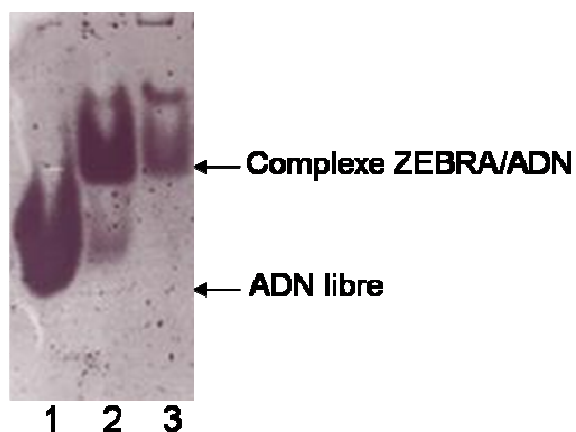


Figure II.5 : Analyse du complexe ZEBRA/ADN par gel retard.

L'oligonucléotide testé ici est le meZRE2(a).

Piste 1 : oligonucléotide meZRE2(a) seul.

Pistes 2 et 3 : surnageant et culot du mélange ZEBRA/meZRE2(a) après centrifugation.

L'ajout de l'oligonucléotide double brin à la solution de ZEBRA entraîne la précipitation d'une partie du complexe ZEBRA/ADN. Seul le surnageant est utilisé pour réaliser les gouttes de cristallogénèse. Nous ne pouvons donc pas savoir exactement quel ratio protéine/ADN est réellement présent dans la goutte.

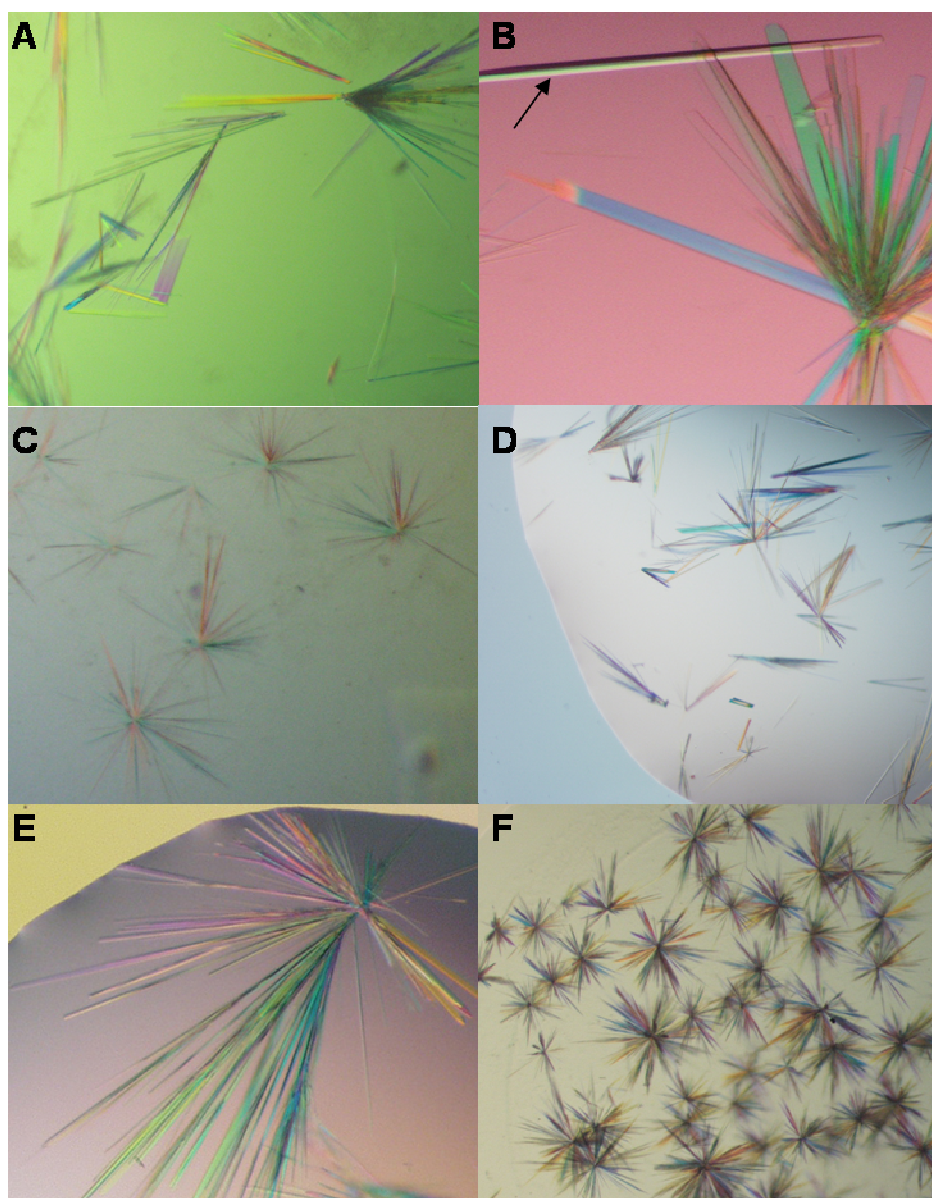


Figure II.6 : Images des cristaux obtenus.

(A) et (B) complexe ZEBRA/meZRE2(a) : conditions de cristallisation I (A) et II (B).

(→ : aiguille de dimensions 800 x 40 μm).

(C à F) complexe ZEBRA/ZRE2(a) : conditions de cristallisation III (C), IV (D), V (E) et VI (F).

Cependant, seuls les cristaux du complexe ZEBRA/meZRE2(a) issus de la condition de cristallisation I ont diffracté à haute résolution (figure II.6.A.). Cette condition était donc prometteuse pour la détermination de la structure du complexe. Mais ces cristaux en forme d'aiguilles étaient très fins et petits et ne permettaient donc pas d'effectuer plusieurs irradiations avant de se dégrader complètement. En les reproduisant, nous avons ajouté 20%

d'isopropanol au milieu de cristallisation (condition de cristallisation II). Les cristaux sont apparus au bout de 48H et en grand nombre (figure II.6.B.). La majorité des cristaux est toujours sous forme d'aiguilles, mais des cristaux en forme de plaques sont apparus. Ces derniers d'une taille approximative de 800 x 40 μM , ont permis d'obtenir un jeu complet de données à 2,6 Å de résolution.

CHAPITRE III :

RESOLUTION DE LA STRUCTURE

DU COMPLEXE ZEBRA/ZRE2

METHYLE

La structure du complexe ZEBRA/meZRE2 a été résolue par remplacement moléculaire en utilisant comme modèle une partie réduite du complexe ZEBRA/AP1 (Petosa et al., 2006). Le complexe ZEBRA/meZRE2 cristallise dans le même groupe d'espace (C2) que le complexe ZEBRA/AP1. Cependant, les paramètres de maille sont différents : notamment, la taille de la maille a doublé dans la direction « *a* ». En effet, après plusieurs recherches, nous avons découvert la présence d'une deuxième molécule du complexe ZEBRA/meZRE2 dans l'unité asymétrique, en relation avec la première molécule par une simple opération de translation. De plus, cette 2^{ème} molécule n'étant pas aussi bien ordonnée que la 1^{ère}, nous avons choisi de nous focaliser sur la 1^{ère} molécule pour la description et l'interprétation de la structure.

I. Collecte des données de diffraction

Les données de diffraction ont été collectées dans des conditions cryoprotectantes sur la ligne ID23-eh2 de l'ESRF à 100K et à une longueur d'onde de 0,873 Å. Cette source de rayons X est une ligne micro-focus : elle génère un faisceau de rayons X très étroit qui permet de viser un endroit très précis du cristal et de se déplacer le long du cristal dans le cas où la zone irradiée se dégraderait.

Cette ligne micro-focus nous a été d'une grande aide étant donné la fragilité et l'étroitesse de nos cristaux en forme d'aiguilles. Grâce à elle, nous avons ainsi pu collecter un jeu complet de données à 2,4 Å (figure III.1.).

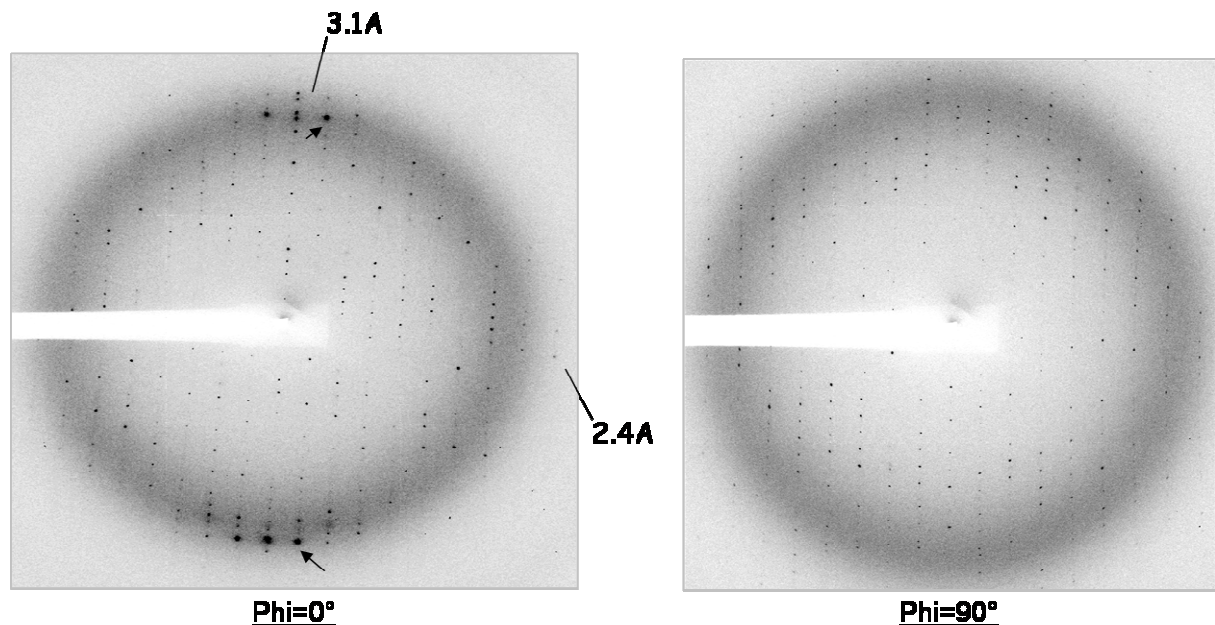


Figure III.1 : Clichés de diffraction à $\Phi=0^\circ$ et $\Phi=90^\circ$. L'axe de rotation est horizontal. L'oscillation est de 1° et le temps d'exposition est de 4 secondes. Le faisceau de rayons X est de 10 μm . La transmission est de 10 %. Les tâches de plus forte intensité (flèches) sur le cliché $\Phi=0^\circ$ sont dues à la présence de l'ADN.

Statistiques de collecte des données	
Paramètre	Valeur
Source de rayons X	ID23-eh 2, ESRF
Longueur d'onde (Å)	0,873
Température de collection des données (K)	100
Détecteur	MAR CCD 165mm
Groupe d'espace	C2
Paramètres de la maille (Å)	a = 208, b = 26.6, c = 80.9, β = 103.1°
Résolution (Å)	30–2.4
R_{sym} (%) ^{a,b}	9.6 (61.7)
Nombre total de réflexions	64,010 (7350)
Nombre de réflexions uniques	17,346 (1931)
Complétude (%)	98.5 (98.7)
Redondance	3.7 (3.8)
Mosaïcité (°)	0.24
$I/\sigma(I)$	9.5 (2.4)
Statistiques de traitement des données	
Paramètre	Valeur
$R_{\text{cryst}}/R_{\text{free}}$ (%) ^c	27.0 / 29.9
<u>Nombre d'atomes</u>	
Protéine	1866
ADN	1119
H ₂ O	126
<u>Facteurs de température (Å²)</u>	
Protéine	44.0
ADN	59.4
H ₂ O	44.1
Rmsd longueurs liaison (Å)/angles (°)	0.012/1.9
Résidus présents dans les régions les plus favorables/ autorisées du diagramme de Ramachandran (%)	80.5/18.5
^a Les valeurs en parenthèses font référence à la couche de plus haute résolution (2.4–2.5 Å). ^b $R_{\text{sym}} = \sum I - \langle I \rangle / \sum I$, où I est l'intensité observée pour une intensité donnée. ^c $R_{\text{cryst}} = \sum F_o - F_c / \sum F_o$, où F_o et F_c sont respectivement les facteurs de structure observés et calculés. R_{free} est égal à R_{cryst} pour une portion de 5 % de réflexions sélectionnées au hasard et non utilisées dans le processus d'affinement.	

Table III.1 : Résumé des statistiques de traitement des données de diffraction et de l'affinement du modèle.

II. Traitement des données de diffraction

En premier lieu, nous avons comparé les paramètres de maille avec ceux de la structure de ZEBRA en complexe avec AP1 et calculé le coefficient de Matthews (Matthews, 1968) ainsi que le pourcentage de solvant selon qu'il y ait 1, 2 ou 3 molécules du complexe protéine/ADN par unité asymétrique (table III.2.). Cette analyse nous permet d'obtenir des indications sur le nombre de molécules du complexe présentes dans l'unité asymétrique.

Complexe	ZEBRA/AP1	ZEBRA/meZRE2		
Paramètres de maille (Å)	a = 94.2, b = 26.5, c = 98.1, β = 103.9	a = 208, b = 26.6, c = 80.9, β = 103.1		
Molécules/u.a.	1	1	2	3
Coefficient de Matthews (Å ³ /Da)	2,19	3,99	1,99	1,33
% de solvant	49 %	72,5 %	45 %	17,5 %

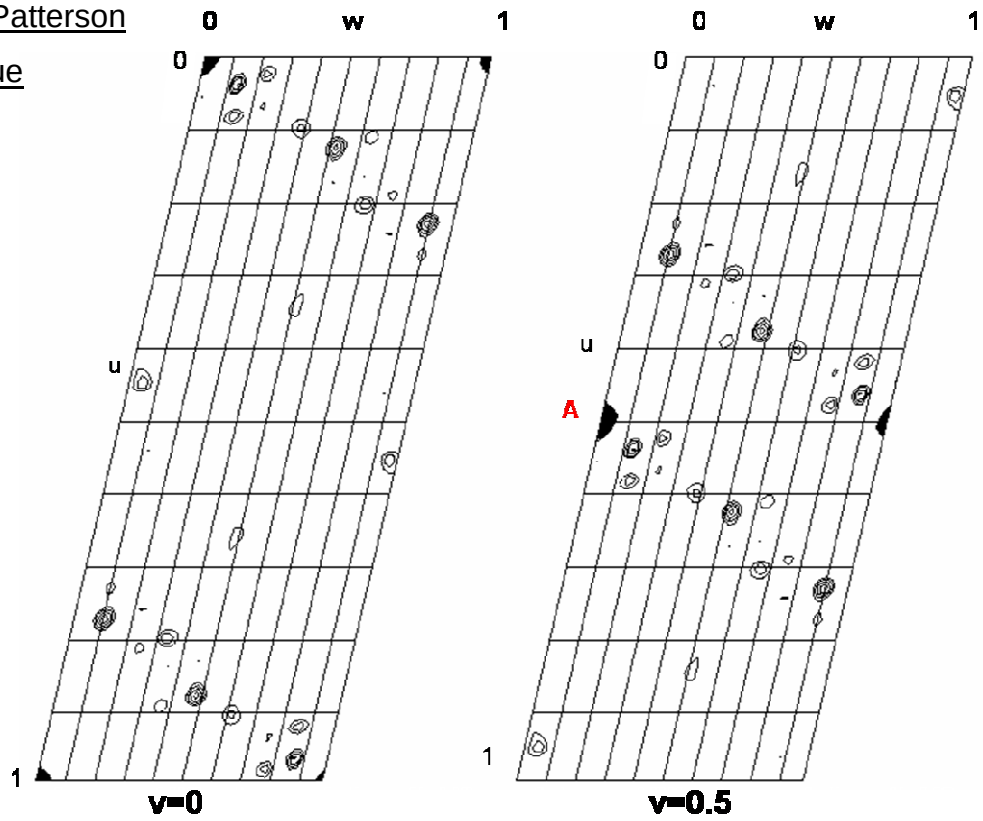
Table III.2 : Comparaison des paramètres de maille des complexes ZEBRA/ADN. (u.a. : unité asymétrique. Da : Dalton).

Cette table indique clairement que notre structure ne peut pas contenir 3 molécules par unité asymétrique car le pourcentage de solvant serait trop faible (17,5 %). En revanche, l'unité asymétrique peut contenir 1 ou 2 molécules du complexe ZEBRA/meZRE2 étant donné que ces deux cas de figures présentent respectivement des pourcentages de solvant cohérents de 72,5 % et 45%. Cependant, un pourcentage de solvant de 45 % est plus cohérent avec la résolution de la diffraction des cristaux. De plus, le coefficient de Matthews correspondant à 2 molécules/u.a. est similaire à celui calculé pour la structure du complexe ZEBRA/AP1. Cela suggère que l'unité asymétrique contient 2 molécules du complexe ZEBRA/meZRE2.

Afin de vérifier la présence de 2 molécules/u.a., nous avons tout d'abord calculé une fonction d'auto-rotation. Cette dernière n'ayant fait apparaître aucun pic correspondant à de la symétrie non cristallographique, cela indique que si une deuxième copie du complexe ZEBRA/meZRE2 est présente dans l'unité asymétrique, elle pourrait se trouver dans une

orientation très similaire à la première copie du complexe. Pour vérifier la présence d'une symétrie non-cristallographique purement translationnelle, une fonction de Patterson a été calculée à partir des données de diffraction (carte de Patterson expérimentale ; figure III.2.B.). Celle-ci a été comparée à une fonction calculée pour le même groupe d'espace à partir d'un modèle hypothétique placé arbitrairement dans la même maille à raison d'une molécule/unité asymétrique (carte de Patterson typique ; figure III.2.A.).

A. Carte de Patterson
typique



B. Carte de Patterson
expérimentale

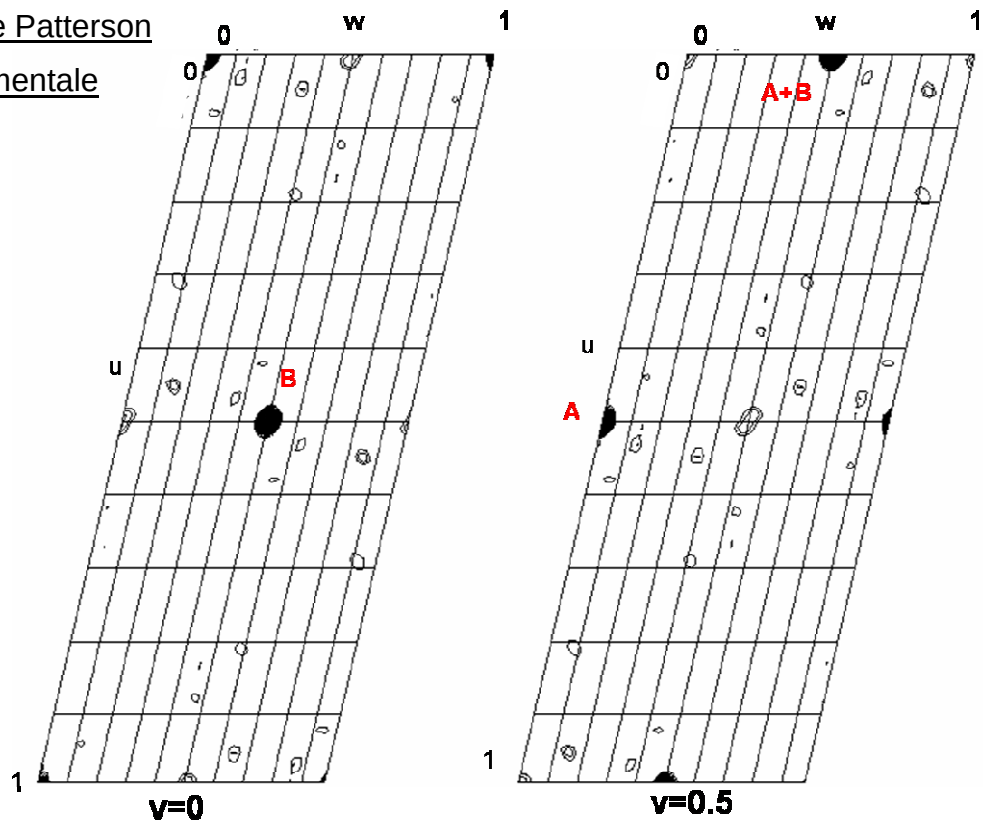


Figure III.2 : Cartes de Patterson.

(A). Carte de Patterson générée pour le groupe d'espace C2 à partir d'un modèle hypothétique placé arbitrairement dans la maille à raison d'une molécule/u.a.(carte de Patterson typique).

(B). Carte de Patterson générée à partir des données de diffraction (carte de Patterson expérimentale).

Le pic « A » présent sur la carte de Patterson typique (figure III.2.A.) correspond à de la symétrie cristallographique propre à ce groupe d'espace. En revanche, un nouveau pic (pic « B ») apparaît sur la carte de Patterson expérimentale (figure III.2.B.). La position de ce pic est cohérent avec la présence de 2 molécules dans l'unité asymétrique, reliées par un facteur de translation (0,0 ; 0,5 ; 0,5).

(Un deuxième pic « A+B » apparaît sur la carte de Patterson expérimentale section $v=0,5$ (figure III.2.B.) : il correspond à la somme des symétries cristallographiques (pic « A ») et non cristallographiques (pic « B »).)

Pour résoudre la structure par remplacement moléculaire, nous avons utilisé le programme PHASER. Le premier modèle choisi correspondant à la structure complète de ZEBRA en complexe avec AP1, n'a pas permis de résoudre la structure. Nous avons donc utilisé un modèle réduit comportant uniquement les résidus 175 à 200 de chaque hélice du bZIP de ZEBRA ainsi que les 12 paires centrales du duplexe oligonucléotides de 19 bp (voir Chapitre VIII : Matériels et Méthodes, section II.5.) Les caractéristiques des quatre premières solutions proposées par PHASER sont résumées dans la table III.3. :

Rang de la solution	Angles d'Euler (°)	Translation (Å)	Z-SCORE		LLG
			RF	TF	
1	182.5 35.6 177.7	-0.321 0.082 -0.774	7.4	8.2	597
2	231.3 106.7 336	0.004 -0.236 1.036	5.2	11.4	558
3	179.5 36.5 179.2	-0.181 0.113 -0.985	7.4	5.8	558
4	75.2 129.3 22.8	0.107 1.209 0.906	5.4	8.8	539

Table III.3 : Table résumant les statistiques de remplacement moléculaire des 4 premières solutions proposées par le programme PHASER.

RF : fonction de rotation.

TF : fonction de translation.

LLG : « log-likelihood-gain ».

La 1^{ère} solution semble correcte étant donné l'écart important entre son score LLG et celui de la 2^{ème} solution.

Nous avons analysé l'empilement cristallin de cette première solution en générant les copies du complexe ZEBRA/meZRE2 obtenues par symétrie du groupe d'espace C2 (figure III.3.) :

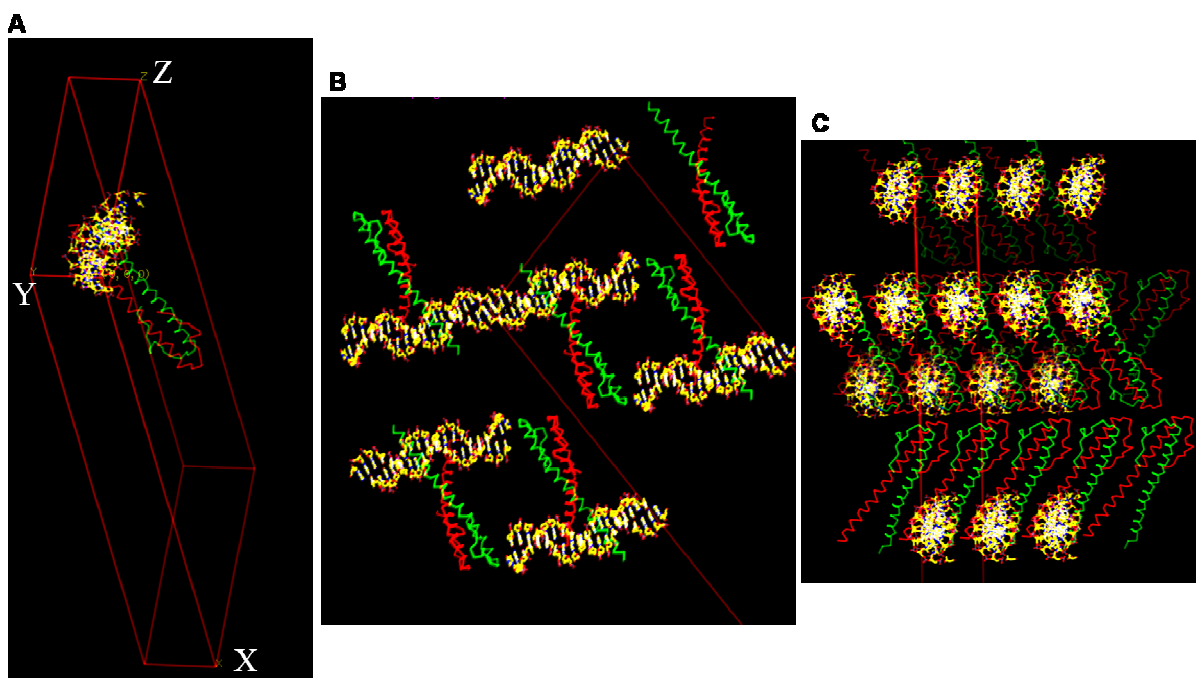


Figure III.3 : Analyse de l'empilement cristallin du complexe ZEBRA/meZRE2.

(A) La maille est représentée avec une molécule du complexe ZEBRA/meZRE2 par unité asymétrique. Les trois directions X, Y et Z de la maille sont indiquées.

(B) (C) Empilement cristallin correspondant à la première solution proposée par le programme PHASER. Observation selon l'axe Y (B) et selon l'axe de l'ADN (C).

Nous n'observons aucun chevauchement entre les différentes copies du complexe ZEBRA/meZRE2, ce qui indique que la première solution proposée par le programme PHASER est convenable. De plus, il reste beaucoup d'espace dans la maille qui se trouve inoccupé et qui pourrait accueillir une 2^{ème} molécule du complexe ZEBRA/meZRE2 (figure III.3.B.).

Une carte 2Fo-Fc calculée à partir de cette 1^{ère} solution révèle de la densité hélicale supplémentaire, en prolongation de l'hélice partielle présente dans le modèle (figure III.4.) :

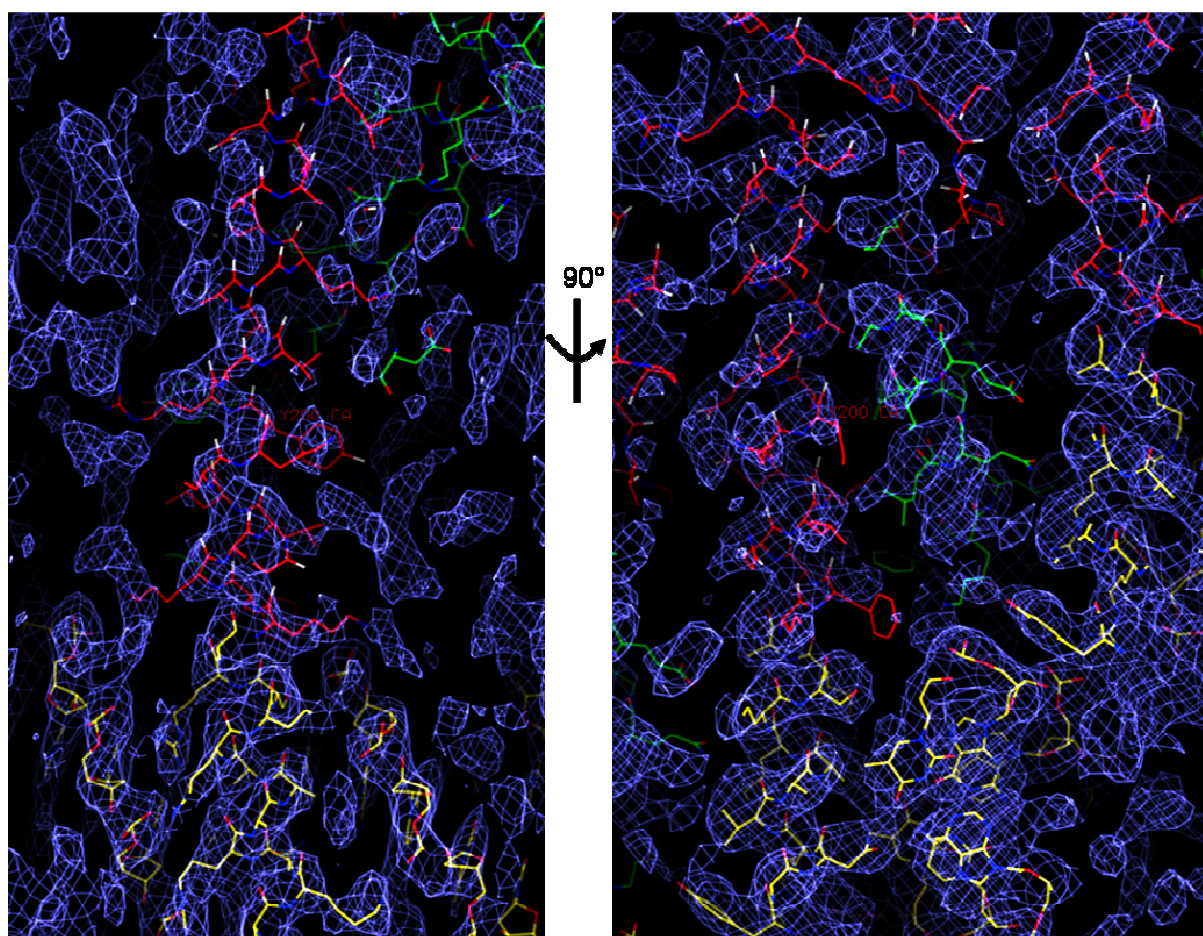


Figure III.4 : Vues orthogonales de la carte de densité électronique 2Fo-Fc du complexe ZEBRA/meZRE2.

En jaune : modèle utilisé pour le remplacement moléculaire.

En rouge et en vert : partie omise de la structure provenant du modèle complet ZEBRA/AP1.

Le fait que le modèle utilisé pour le remplacement moléculaire soit réduit (voir Chapitre VIII : Matériels et Méthodes) et que nous puissions clairement distinguer la densité électronique correspondant aux deux hélices α du bZIP de ZEBRA confirme que la première solution proposée par le programme PHASER est une solution correcte.

L'empilement cristallin a été réexaminé en rajoutant sur l'image la densité électronique d'une carte Fo-Fc calculée à partir de la première solution proposée par PHASER (figure III.5.) :

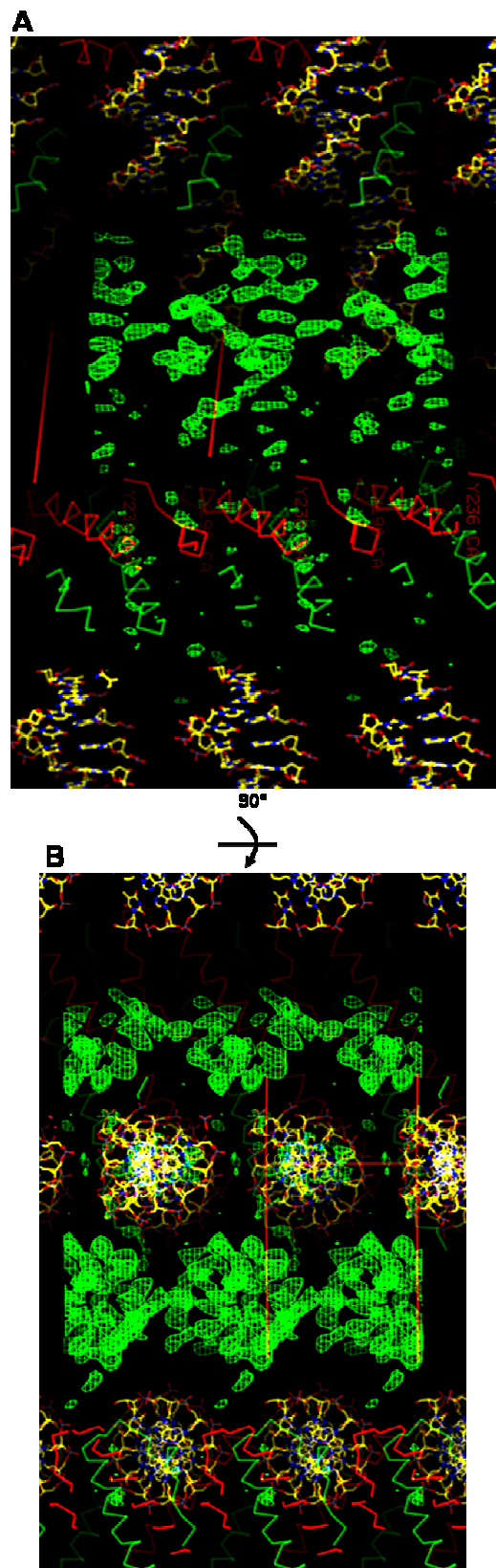


Figure III.5 : Carte Fo-Fc de la 1^{ère} solution proposée par PHASER calculée à partir du modèle réduit utilisé pour le remplacement moléculaire. (A) Observations selon un axe perpendiculaire à l'ADN et (B) selon l'axe de l'ADN. En vert : densité hélicale supplémentaire.

Cette carte montre clairement une densité électronique supplémentaire correspondant à une deuxième molécule du complexe ZEBRA/meZRE2 dans l'unité asymétrique. L'opération de translation reliant ces deux molécules dans l'unité asymétrique est en accord avec celle déterminée par l'analyse des cartes de Patterson.

III. Affinement et validation

La 2^{ème} molécule a été placée dans la densité électronique à l'aide de la partie réduite du modèle. Le programme CNS a été utilisé pour débiter l'affinement. Au début du processus d'affinement, les paramètres R_{cryst} et R_{free} étaient respectivement de 53 et 56 % avec une seule molécule dans l'unité asymétrique puis de 48 et 51,6 % en rajoutant la deuxième molécule. Puis la structure du complexe ZEBRA/meZRE2 a été construite manuellement à l'aide du programme « O » et affinée par le programme CNS jusqu'à construction complète du complexe. La densité de la 2^{ème} molécule est de qualité inférieure à celle de la 1^{ère}, surtout au niveau des extrémités du double brin ADN et de la partie C terminale de la protéine (figure III.6.). Cette 2^{ème} molécule du complexe est plus désordonnée car moins bien tenue en place par des contacts cristallins. Cela explique pourquoi la valeur de $R_{\text{cryst}}/R_{\text{free}}$ est légèrement plus élevée que ce qu'elle devrait être à cette résolution. Nous avons donc choisi de concentrer l'interprétation de la structure du complexe ZEBRA/meZRE2 sur la 1^{ère} molécule présente dans l'unité asymétrique.

Le modèle a été validé grâce au programme PROCHECK en vérifiant que la géométrie de la structure était satisfaisante (voir table III.1. et diagramme de Ramachandran en Annexe E).

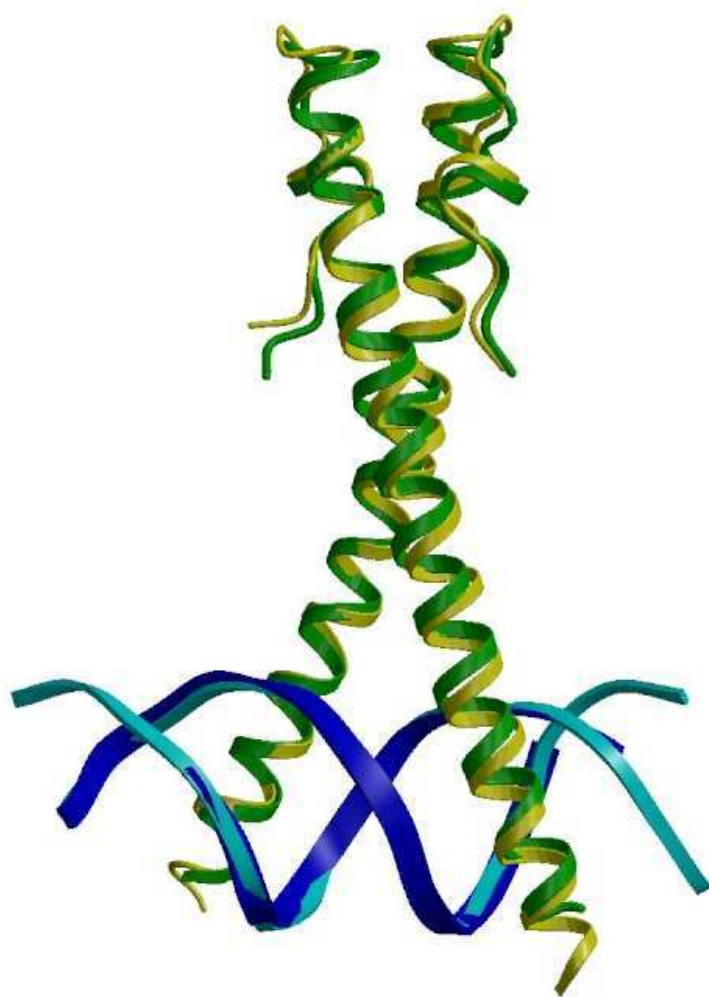


Figure III.6 : Alignement des structures des 2 molécules du complexe ZEBRA/meZRE2 présentes dans l'unité asymétrique. En jaune et bleu : 1^{ère} molécule du complexe ; en vert et cyan : 2^{ème} molécule du complexe. Les paramètres relatifs à cet alignement sont présentés en Annexe F.1.

CHAPITRE IV :

ANALYSE DE LA STRUCTURE DU COMPLEXE ZEBRA/ZRE2 METHYLE

I. La protéine ZEBRA

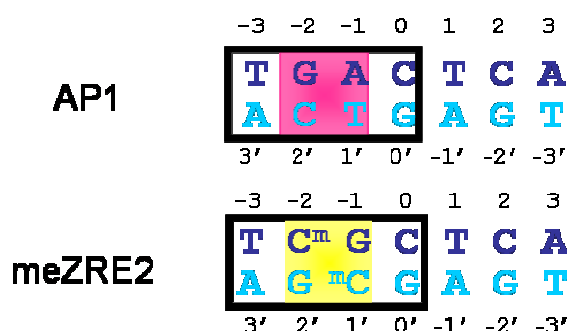


Figure IV.1. : Séquences oligonucléotidiques des heptamères AP1 (en haut) et meZRE2 (en bas). Les paires de bases surlignées en rose sont spécifiques de AP1 et celles surlignées en jaune spécifiques de meZRE2.

Le site meZRE2 diffère du site AP1 uniquement au niveau des paires de bases -2/2' et -1/1' : les bases G⁻² et T^{1'} sont remplacées par deux cytosines méthylées (figure IV.1.). La structure d'une cytosine méthylée se rapproche de celle d'une thymine étant donné qu'elle possède un groupement méthyle additionnel en position 5 de son cycle (figure IV.2.) :

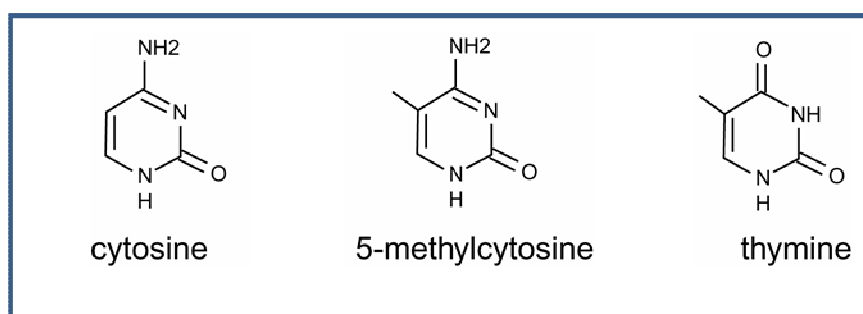


Figure IV.2. : Structures des bases cytosine, cytosine méthylée et thymine.

Les heptamères en général, tel que le site AP1, sont asymétriques. L'heptamère meZRE2 marque encore plus cette asymétrie : il présente un demi-site droit identique à celui d'AP1 et un demi-site gauche contenant un îlot CpG méthylé (figure IV.1.). Chaque demi-site est reconnu par un monomère de ZEBRA : les contacts protéine-ADN dans le sillon majeur ne seront donc pas les mêmes pour les deux hélices du bZIP de ZEBRA. L'arginine 190 du premier monomère de ZEBRA forme deux liaisons hydrogènes avec les atomes O6 et N7 de la base centrale G^{0'} de l'heptamère meZRE2 alors que l'arginine 190 du deuxième monomère

forme des contacts avec le squelette phosphate de l'ADN (figure IV.3.). Nous avons tout d'abord aligné les complexes ZEBRA/AP1 et ZEBRA/meZRE2 en minimisant les rmsd des Ca de ZEBRA et des phosphates de l'ADN, tout en respectant les différences de conformation des deux arginines 190 de chaque complexe (figure IV.3.).

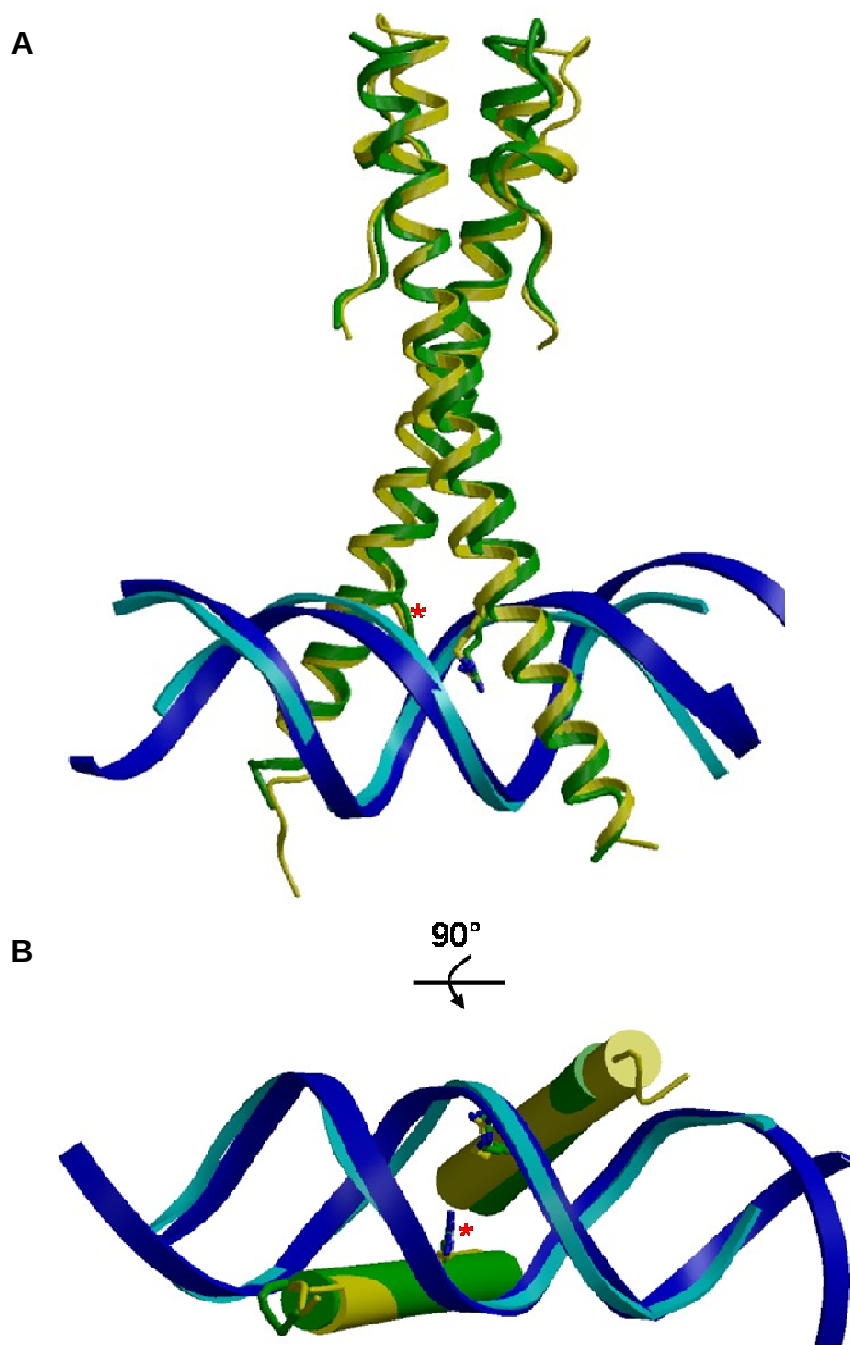


Figure IV.3. : Superposition des complexes ZEBRA/AP1 et ZEBRA/meZRE2 : alignement global.

(A) Vue d'ensemble de l'alignement structural des complexes ZEBRA/AP1 (en jaune et bleu) et ZEBRA/meZRE2 (en vert et cyan).

(B) Vue orthogonale. Seule une partie des hélices bZIP de ZEBRA est représentée sous forme de « tonneaux ».

Alignement réalisé en minimisant les rmsd des C α de ZEBRA et des phosphates de l'ADN (les paramètres associés à cet alignement sont présentés en annexe F.2.). L'alignement conserve l'asymétrie des sites AP1 et meZRE2 : les arginines 190 des monomères de ZEBRA formant des contacts avec le squelette phosphates de l'ADN sont alignées, et celles interagissant avec la base centrale G^{0'} (indiquées par un astérisque) le sont également.

A première vue, les structures de ZEBRA en complexe avec les sites meZRE2 et AP1 semblent très proches l'une de l'autre. Les changements sont minimes au niveau des protéines. Cependant, des variations plus importantes se remarquent au niveau de la double hélice d'ADN. Les deux complexes ont donc été alignés cette fois, uniquement par rapport à la double hélice d'ADN (figure IV.4.).

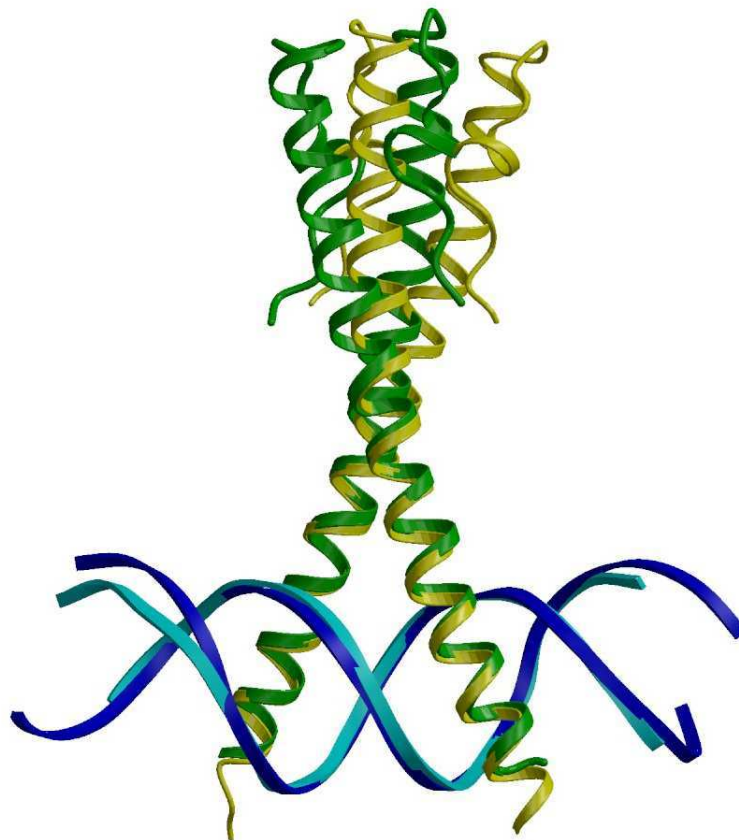


Figure IV.4. : Superposition des complexes ZEBRA/AP1 et ZEBRA/meZRE2 : alignement selon les doubles hélices d'ADN.

Le complexe ZEBRA/AP1 est représenté en jaune et bleu et le complexe ZEBRA/meZRE2 en vert et cyan. L'alignement est réalisé en minimisant uniquement les rmsd des atomes de l'ADN (les paramètres associés à cet alignement sont présentés en annexe F.2.).

L'alignement des complexes vis-à-vis de la double hélice d'ADN met en évidence une courbure importante du dimère ZEBRA associé au site AP1 (en jaune). Au sommet des 2 hélices bZIP, c'est-à-dire au niveau des résidus sérines 224, la distance $C\alpha-C\alpha$ entre les deux complexes est de 10 Å. Cependant, cette courbure ne concerne que la partie C terminale de la protéine, l'orientation des régions basiques étant parfaitement conservée. Cette inflexion n'a donc certainement pas de conséquence sur la fixation à l'ADN de la protéine ZEBRA et doit être due aux contacts cristallins, différents dans les deux complexes.

II. LA DOUBLE HELICE D'ADN

En alignant les complexes vis-à-vis de la double hélice d'ADN, plusieurs changements significatifs sont à présent visibles (figure IV.5.) :

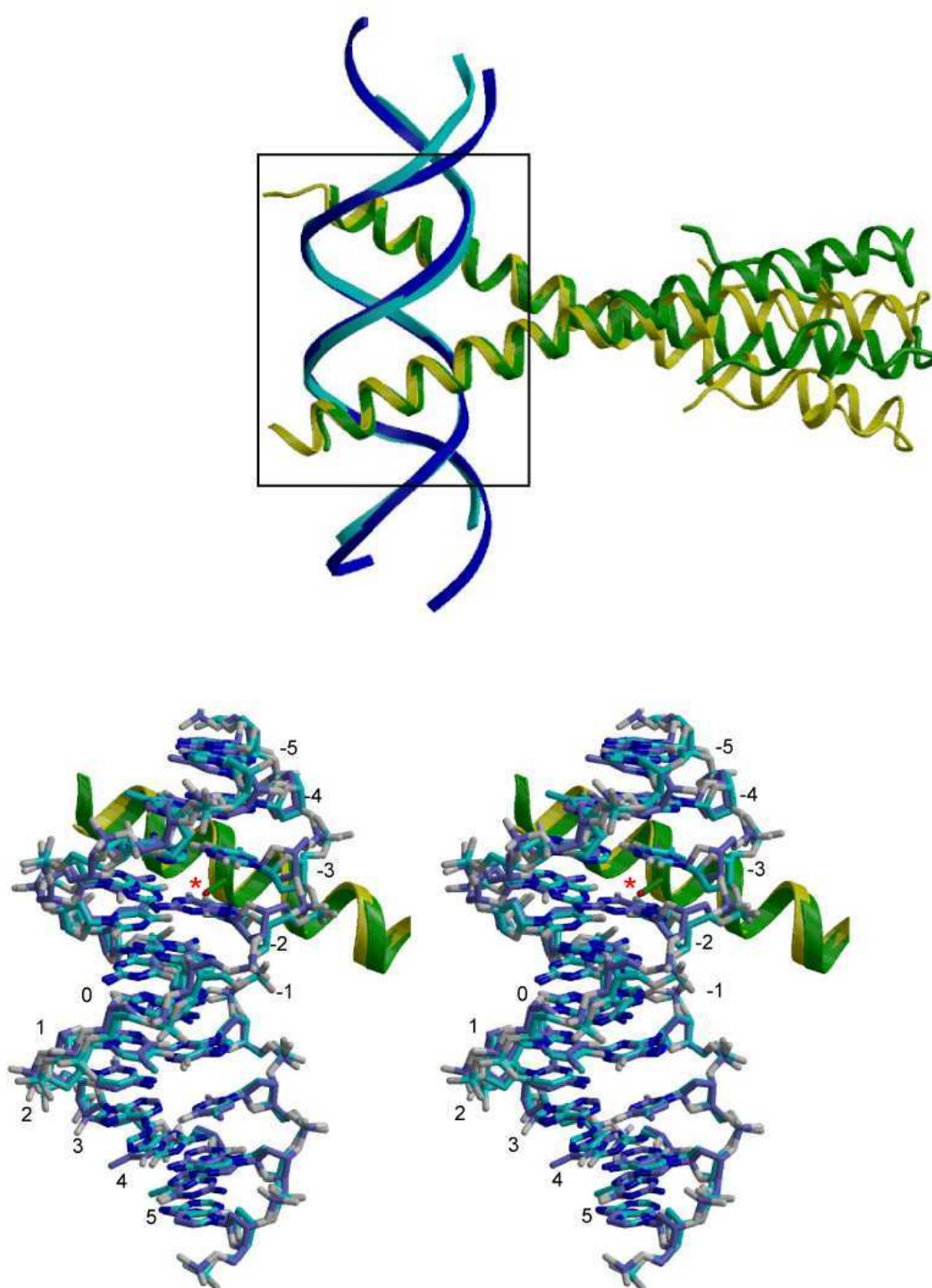
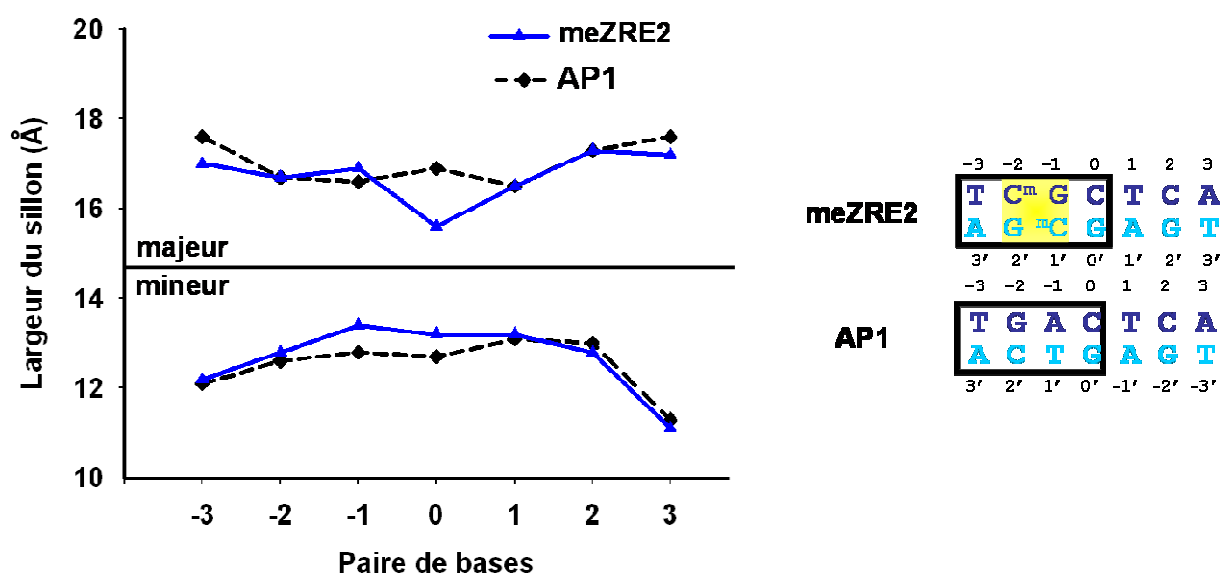


Figure IV.5. : Superposition des complexes ZEBRA/AP1 et ZEBRA/meZRE2 : alignement selon les doubles hélices d'ADN.

(A) Vue d'ensemble de l'alignement structural des complexes ZEBRA/AP1 (en jaune et bleu) et ZEBRA/meZRE2 (en vert et cyan). L'alignement est réalisé en minimisant uniquement les rmsd des atomes de l'ADN (les paramètres associés à cet alignement sont présentés en annexe F.2.).

(B) Vue en stéréo de l'alignement des doubles hélices ADN correspondantes aux sites AP1 et meZRE2. Seules la région comprise entre les paires de base (-5/5') et (5/-5') est représentée. Une partie de l'hélice α d'un monomère de ZEBRA est représentée et la sérine 186 est indiquée par une astérisque (*).

On ne décèle aucun changement significatif au niveau des demi-sites droits des heptamères AP1 et meZRE2 : les paires de bases (0/0' à 5/-5') et les squelettes phosphates sont alignés. Par contre, un décalage apparaît au niveau des phosphates de l'ADN du demi-site gauche (paire de bases -5/5' à -1/1') du site meZRE2, notamment au niveau de la région formant des contacts avec la sérine 186. Afin d'estimer plus précisément ces variations, le programme 3DNA a été utilisé pour mesurer et comparer les largeurs des sillons majeur et mineur de chaque double hélice d'ADN (figure IV.6.) :



Les largeurs des sillons majeurs des sites AP1 et meZRE2 sont pratiquement identiques excepté au niveau de la paire de base centrale 0/0', soit très près de l'îlot CpG du site meZRE2. On observe alors un rétrécissement du sillon majeur du site meZRE2 : sa taille est de 15 Å comparé à 16,3 Å pour le site AP1. Ceci est compensé par une augmentation de taille du sillon mineur au niveau des paires de bases -1/1' et 0/0'. L'introduction d'un îlot

CpG semble entraîner un rétrécissement du sillon majeur du site meZRE2. Mais la méthylation d'îlot CpG joue-t-elle un rôle dans ce changement structural ? Peu d'études ont été réalisées concernant l'influence de la méthylation d'un îlot CpG sur la structure de l'ADN. Heinemann et Hahn ont comparé les structures d'un oligonucléotide de 10 bp méthylé ou non sur une cytosine (nous l'appellerons oligonucléotide X) (Heinemann and Hahn, 1992). Les largeurs des sillons majeur et mineur de cet oligonucléotide sont représentées ci-dessous (figure IV.7.) en fonction de son état de méthylation :

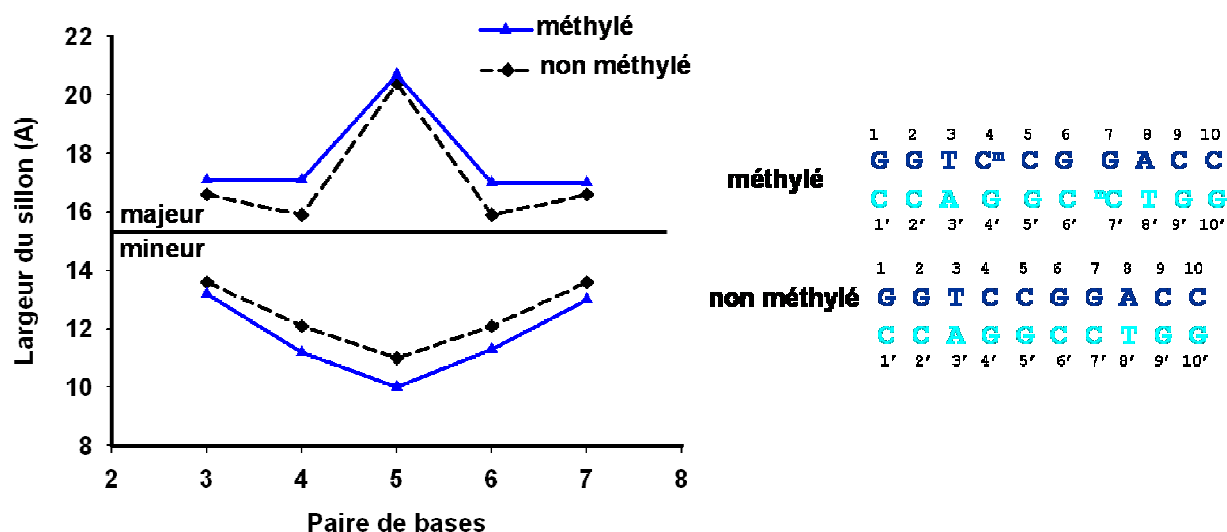


Figure IV.7. : Comparaison des largeurs des sillons majeur et mineur de l'oligonucléotide X non méthylé (en noir) ou méthylé sur une cytosine (en bleu). Seule la région comprise entre les paires de base 2/2' et 8/8' est représentée. Les données ont été obtenues grâce au programme 3DNA et d'après Heinemann et Hahn, 1992. Les n° d'entrée pdb des oligonucléotides X méthylé et non méthylé sont respectivement 122D et 1BD1.

Contrairement au site meZRE2, la méthylation des cytosines C⁴ et C^{7'} semble entraîner un élargissement du sillon majeur de l'oligonucléotide X au niveau des positions 4 et 6 ; cela est compensé par un rétrécissement du sillon mineur. Cependant, les auteurs précisent que la méthylation de ces cytosines ne change que très légèrement les paramètres structuraux de l'oligonucléotide X. Nous avons observé l'effet inverse concernant le site meZRE2 puisque le sillon majeur se rétrécit. Cependant, les cytosines méthylées de l'oligonucléotide X ne font pas partie d'un îlot CpG contrairement au site meZRE2. Cela suggère que les changements structuraux observés entre les sites AP1 et meZRE2 seraient principalement dus à la présence de l'îlot CpG, et non à la méthylation des cytosines C^{1'} et C².

Avant d'analyser les contacts établis entre ZEBRA et le site meZRE2, rappelons ceux décrits par Carlo Petosa concernant le complexe ZEBRA/AP1 (figure IV.8.).

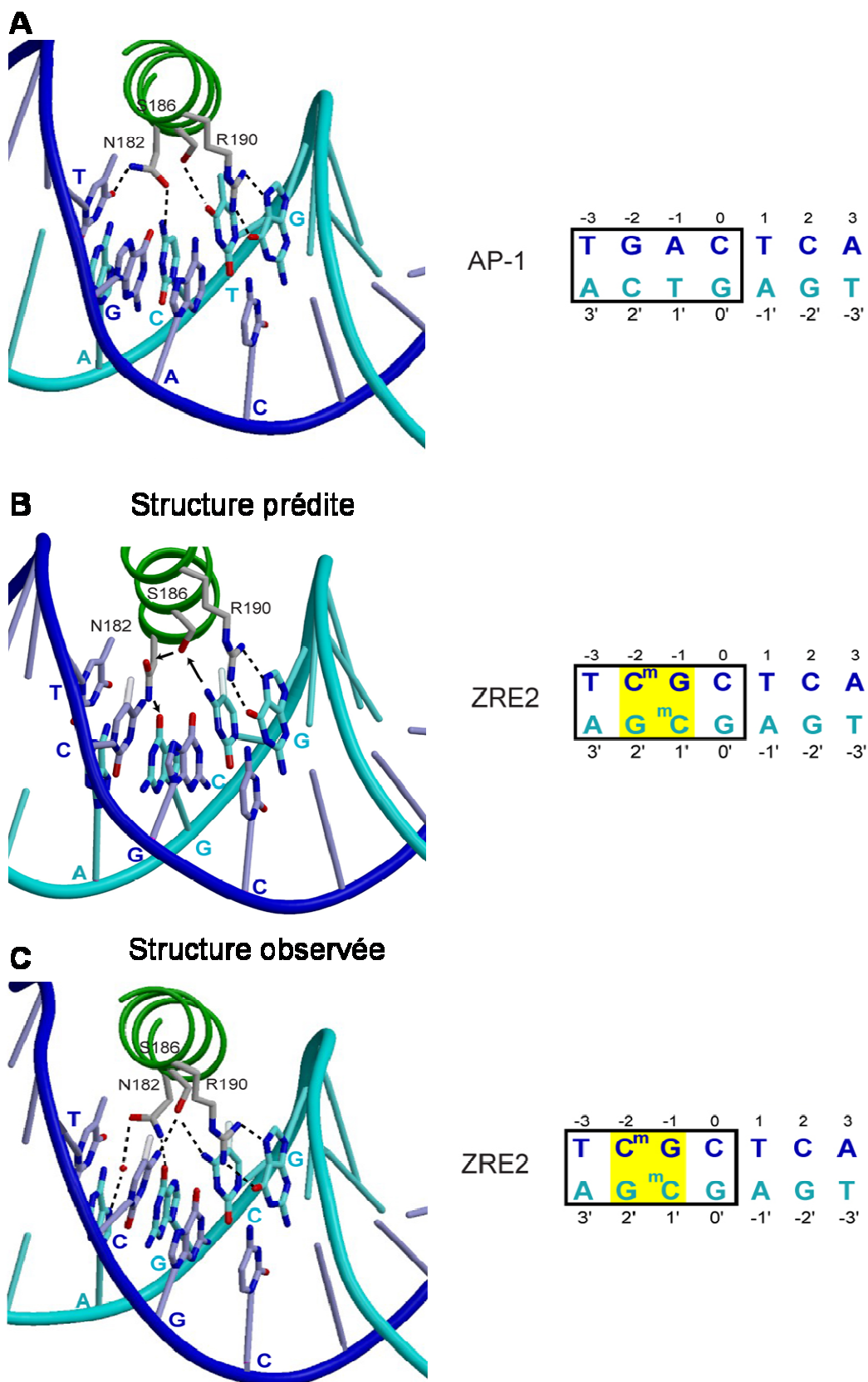


Figure IV.8. : Contacts ZEBRA/ADN au niveau du sillon majeur.

(A) Contacts entre ZEBRA et le site AP1. La sérine 186 a été modélisée étant donné que ce résidu a été muté en alanine dans la structure réelle. (B) Prédiction de structure du complexe ZEBRA/meZRE2. les flèches indiquent les contacts hypothétiques entre les résidus de ZEBRA et le site meZRE2. (C) Contacts protéine/ADN de la structure réelle du complexe ZEBRA/meZRE2. Seul les demi-sites gauches des oligonucléotides sont représentés ainsi que les 3 résidus N182, S186 et R190 de ZEBRA. Les liaisons hydrogènes sont représentées en pointillés.

En complexe avec le site AP1 (figure IV.8.A.), l'asparagine 182 forme deux liaisons hydrogènes avec les atomes N4 de la cytosine C^{2'} et O4 de la thymine T³. Le résidu 186 a été modélisé en sérine étant donné qu'il a été muté en alanine à l'origine. Cette sérine 186 adopterait une conformation lui permettant ainsi d'interagir avec l'atome O4 de la thymine T^{1'}. Quant à l'arginine 190, elle forme deux liens hydrogènes avec les atomes O6 et N7 de la guanine G^{0'}.

Le site meZRE2 diffère du site AP1 uniquement au niveau des paires de bases -2/2' et -1/1'. Petosa et ses collaborateurs ont ainsi prédit la structure du complexe ZEBRA/meZRE2 et suggéré une hypothèse sur le mécanisme de reconnaissance préférentielle de l'ADN méthylé par ZEBRA (figure IV.8.B.). D'après eux, la présence du motif CpG ne permettrait plus à l'asparagine 182 de former une liaison hydrogène avec la base (C^m) en position 2' comme observé dans le complexe ZEBRA/AP1. Cependant, la méthylation de cet îlot CpG, et plus particulièrement la présence du groupement méthyle de la cytosine C^{1'}, pourrait stabiliser la sérine 186 dans une conformation lui permettant d'interagir avec l'asparagine 182. En effet, la chaîne latérale d'une sérine peut adopter 3 conformations préférentielles : *gauche*⁻, *gauche*⁺ et *trans*. La sérine 186, stabilisée en conformation *gauche*⁺, interagirait avec l'asparagine 182, qui, à son tour pourrait reformer un lien hydrogène avec la guanine G^{2'}, augmentant ainsi l'affinité de ZEBRA pour le site meZRE2 (Petosa et al., 2006).

Analysons à présent la structure réelle de ZEBRA en complexe avec le site méthylé meZRE2 (figures IV.8.C et IV.9.).

III. Analyse de l'interaction entre ZEBRA et le site ZRE2 méthylé

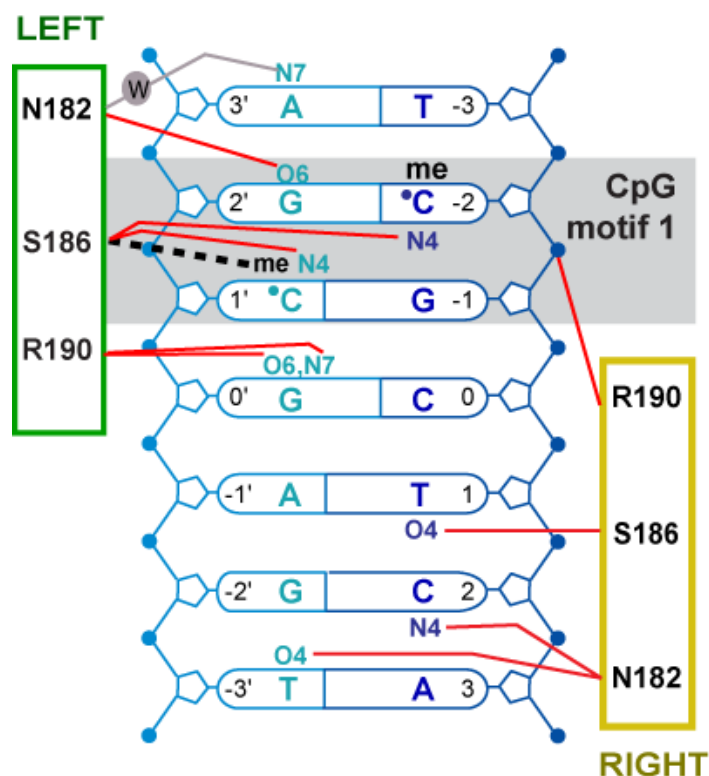


Figure IV.9. : Schématisation des interactions ZEBRA/meZRE2 au niveau des résidus N182, S186 et R190 des hélices gauche et droite du dimère de ZEBRA. L'hélice gauche (en vert) est celle représentée dans les figures A,B et C. Les liaisons hydrogènes sont représentées en rouge et les liaisons hydrophobes en pointillés noirs. La liaison médiée par une molécule d'eau (« W ») est indiquée en gris.

En accord avec la structure prédite, le lien hydrogène entre le groupement hydroxyle de la sérine 186 et l'atome N4 de la cytosine C^{1'} est maintenu (figure IV.8.C.) mais la direction de la liaison est inversée : la cytosine méthylée se retrouve donneur de liaison hydrogène et le groupement OH de la sérine 186, accepteur (alors qu'il était donneur de liaison hydrogène en complexe avec le site AP1). De plus, le groupement méthyle de la cytosine C^{1'} ne se trouvant qu'à 4,25 Å du groupement méthylène de la sérine 186, forme un contact hydrophobe avec cette dernière. Le rotamère *gauche*⁺ de la sérine 186 est stabilisé par ce contact hydrophobe. Enfin, l'arginine 190 conserve les deux liaisons hydrogènes qu'elle formait auparavant avec la base centrale G^{0'}.

Cependant, plusieurs différences avec la structure prédite sont à noter. Tout d'abord, le rétrécissement du sillon majeur d'ADN induit par la présence de l'îlot CpG (voir partie II de ce chapitre) permet à la sérine 186 d'être plus proche de la cytosine C⁻² et ainsi de former une nouvelle liaison hydrogène avec l'atome N4 de cette cytosine. L'asparagine 182, quant à elle, a pivoté de 180°. Elle interagit toujours avec la base en position 2' mais la direction de ce lien est inversée comparé au complexe ZEBRA/AP1 : le groupement amine de l'asparagine 182 est donneur de lien hydrogène à l'atome O6 de la guanine G^{2'}. L'interaction qu'elle formait avec la thymine T⁻³ est remplacée par une interaction indirecte via une molécule d'eau entre son atome Oδ1 et l'atome N7 de l'adénine A^{3'}. Enfin, la sérine 186 ne forme aucune interaction avec l'asparagine 182, comme cela avait été proposé par Carlo Petosa. La structure de ZEBRA en complexe avec le site meZRE2 permet d'apporter une révision à l'hypothèse du mécanisme de reconnaissance préférentielle de l'ADN méthylé par ZEBRA.

Le rétrécissement du sillon majeur du site meZRE2 induit par la présence de l'îlot CpG et la stabilisation de la conformation *gauche*⁺ de la sérine 186 par le groupement méthyle de la cytosine C^{1'}, permet à la sérine 186 de former une nouvelle liaison hydrogène avec la cytosine méthylée C⁻². Ceci explique également le rôle clé joué par la sérine 186 dans le mécanisme de reconnaissance préférentielle de l'ADN méthylé par ZEBRA et mis en évidence par Prasanna Bhende (Bhende et al., 2005).

CHAPITRE V :

ANALYSE DU MECANISME DE

RECONNAISSANCE

PREFERENTIELLE DE L'ADN

METHYLE PAR ZEBRA

Shannon Kenney et ses collaborateurs ont montré que ZEBRA fixe préférentiellement l'ADN méthylé (Bhende et al., 2004). En réalisant des EMSAs, ils ont estimé que ZEBRA lie un site ZRE2 méthylé avec 10 fois plus d'efficacité qu'un site non méthylé. Cependant, aucune mesure précise d'affinité n'a à ce jour été réalisée. La structure cristallographique de ZEBRA en complexe avec un site ZRE2 méthylé résolue au cours de ma thèse, apporte des indices importants pouvant améliorer la compréhension du mécanisme de reconnaissance préférentielle de sites méthylés par ZEBRA. Afin de confirmer et de compléter les informations apportées par cette structure, nous avons mis au point, en collaboration avec l'équipe de François PARCY du CEA, une technique permettant de mesurer avec précision les affinités de ZEBRA pour divers sites ADN. En réalisant une étude de mutagenèse du site ZRE2 méthylé ainsi que de la protéine ZEBRA, nous apportons une révision importante du mécanisme de fixation préférentielle de ZEBRA sur le site ZRE2 méthylé proposé par Petosa et al. en 2006.

I. Principe de la technique d'anisotropie de fluorescence :

La technique d'anisotropie de fluorescence (Lundblad et al., 1996) repose sur l'étude de la mobilité rotationnelle d'une molécule fluorescente (dans notre cas des oligonucléotides marqués à la rhodamine correspondants aux sites de fixation de ZEBRA, voir tableau V.1.). En excitant le groupement fluorescent par une lumière polarisée, la lumière est émise dans des directions qui dépendent de : i) la direction d'observation (par rapport à la direction de lumière incidente) ; ii) l'orientation initiale, la rotation $^{\text{et}}_{\text{ou}}$ la déformation de la molécule pendant le temps qui sépare l'absorption de l'émission. C'est ce dernier aspect qui est le plus intéressant car il donne des informations, notamment, sur la mobilité des molécules sondées. La valeur d'anisotropie de fluorescence (AF) est comprise entre 0 et 1. Elle est calculée comme suit (Canet et al., 2001) :

$$\boxed{AF = (A_v - A_h) / (A_v + 0,5 A_h)}$$

où $A_v = I_{vh} + I_{vv}$ (la somme des intensités de lumière émises après excitation selon l'axe vertical) et $A_h = I_{hh} + I_{hv}$ (la somme des intensités de lumière émises après excitation selon l'axe horizontal). En effet, seule la lumière d'excitation est polarisée. La détection à l'émission se fait sans polariseur (voir Chapitre VIII : Matériels et Méthodes).

Si la molécule marquée est libre, elle a une grande mobilité rotationnelle donc les dipôles émetteurs se désorientent très vite par rapport à la durée de vie de fluorescence (qui est de quelques nanosecondes pour la rhodamine). La valeur d'anisotropie de fluorescence sera ainsi très faible car l'intensité de lumière émise après excitation selon l'axe vertical (A_v) sera pratiquement équivalente à l'intensité de lumière émise après excitation selon l'axe horizontal (A_h). A l'inverse, si cette même molécule est en complexe avec une molécule de plus grande taille (comme ZEBRA), sa mobilité rotationnelle est réduite donc les dipôles émetteurs se réorientent plus lentement par rapport au temps de vie de fluorescence : la valeur d'anisotropie de fluorescence sera donc plus importante (voir figure V.1.).

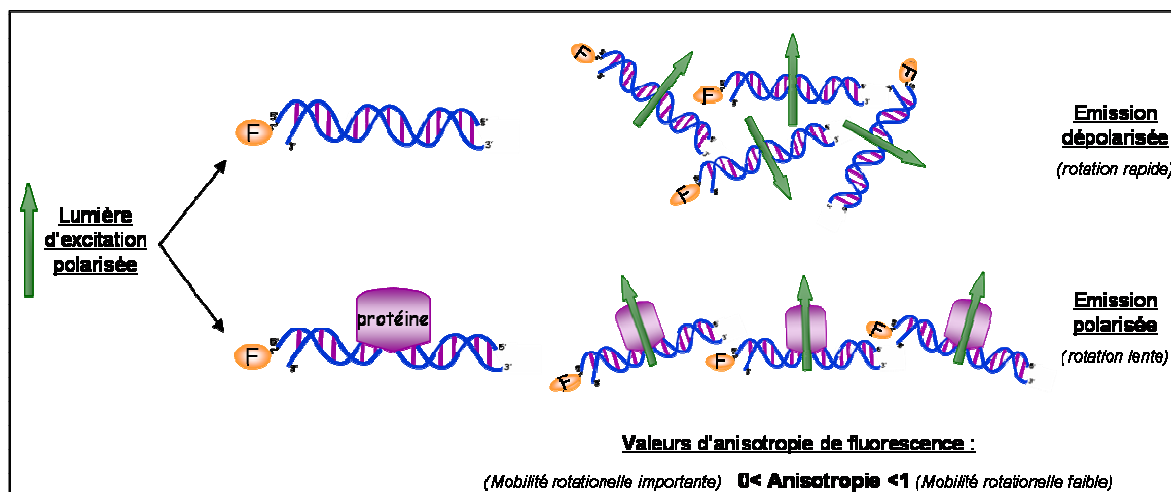


Figure V.1. : Schéma représentant le principe d'anisotropie de fluorescence

L'analyse de la courbe de fixation obtenue permet de déterminer la constante de dissociation à l'équilibre qui informe sur l'affinité et la spécificité des complexes.

II. Mise au point des expériences d'anisotropie de fluorescence

1. Choix des clones de ZEBRA et cJun et purification des protéines

Nous utilisons le même clone de ZEBRA ayant servi à la résolution du complexe ZEBRA/meZRE2. Cependant, bien que ZEBRA fixe l'ADN en tant que dimère, ce clone de ZEBRA de 61 résidus (7298 Da) n'avait pas une taille assez importante pour affecter de façon significative la mobilité rotationnelle de l'oligonucléotide marqué à la rhodamine. Nous avons donc ajouté une étiquette MBP (« *Maltose binding protein* ») du côté N-terminal du bZIP de ZEBRA. A la différence d'une étiquette GST, la protéine MBP ne dimérise pas en solution, elle ne risque donc pas de fausser les capacités de dimérisation ou de liaison à l'ADN de ZEBRA. La construction MBP-ZEBRA forme à présent un polypeptide de plus de 50 KDa qui va permettre d'obtenir un meilleur signal d'anisotropie.

Afin de comparer le comportement de ZEBRA à une autre protéine bZIP classique, nous avons choisi le facteur de transcription cJun. Ce dernier peut former des homodimères mais également des hétérodimères en interagissant avec le facteur de transcription cFos. Pour plus de simplicité, nous avons décidé d'étudier le comportement de cJun en l'absence de

cFos. Etant donné que nous ne nous intéressons qu'au domaine bZIP de cJun, ce dernier a également été exprimé en fusion avec l'étiquette MBP.

Toutes les protéines de fusion-MBP étudiées sont purifiées selon le même protocole (voir Chapitre VIII : Matériels et Méthodes). Les figures illustrant cette partie présentent l'exemple de la protéine MBP-ZEBRA.

La première étape de purification est une chromatographie d'affinité sur résine d'amylose (NEB).

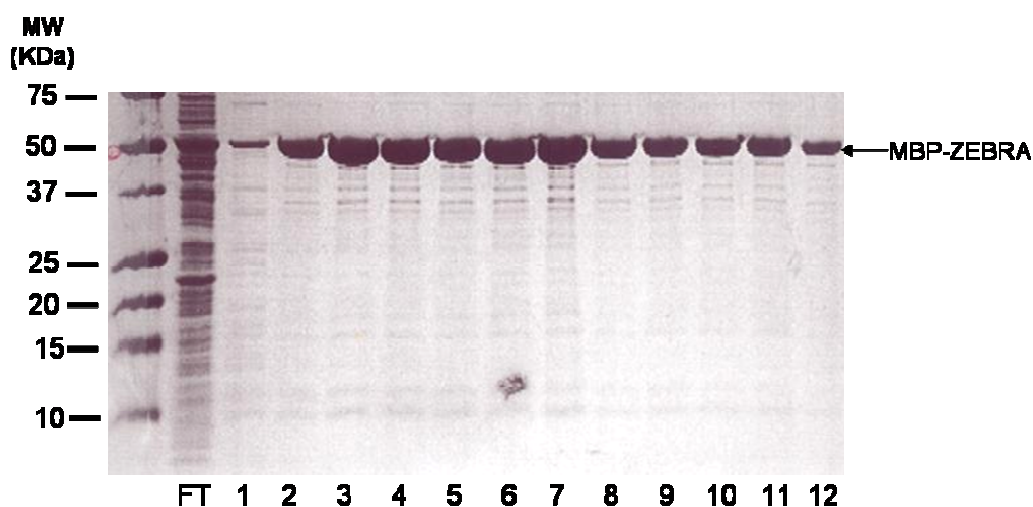


Figure V.2. : Purification de MBP-ZEBRA par chromatographie d'affinité sur résine d'amylose. Gel SDS-PAGE des fractions d'élution. FT : « *flow-through* » ; 1 à 12 : fractions de 1ml contenant la protéine de fusion-MBP.

Après élution par ajout de maltose, les fractions contenant la protéine de fusion-MBP sont analysées sur gel SDS-PAGE (figure V.2.) et concentrées afin de subir une seconde étape de purification par chromatographie d'exclusion de taille sur colonne Superdex 200 10/300 GL (GE Healthcare) (figure V.3.A.).

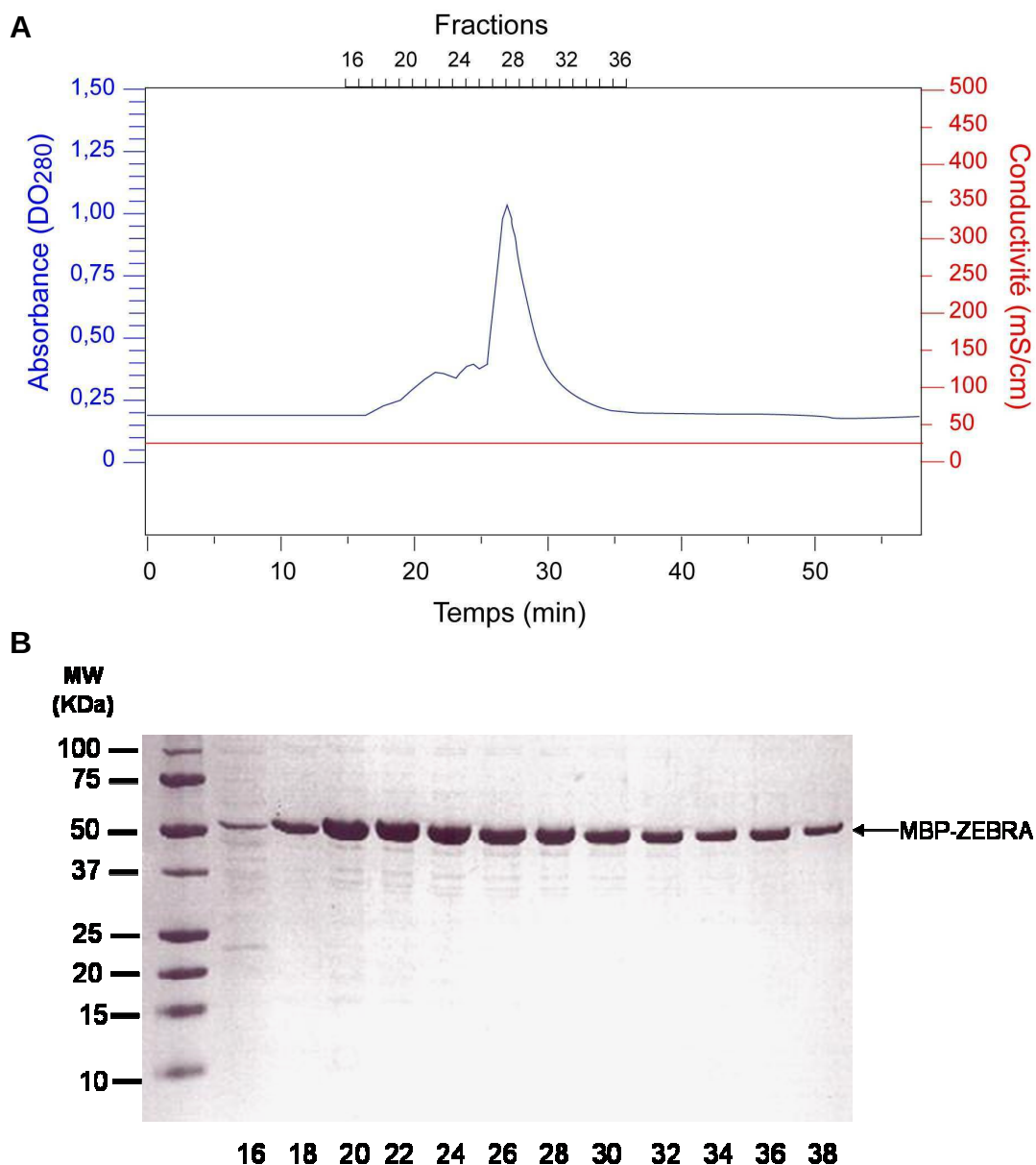


Figure V.3. : Dernière étape de purification de MBP-ZEBRA par chromatographie d'exclusion de taille sur colonne S200.

(A) Profil d'élution de la protéine de MBP-ZEBRA. * : pic correspondant à la protéine MBP-ZEBRA en complexe avec de l'ADN contaminant.

(B) Gel SDS-PAGE des fractions issues de la dernière étape de purification de MBP-ZEBRA par chromatographie d'exclusion de taille. Les numéros des fractions déposées correspondant à ceux des fractions du pic d'élution présenté en (A).

Le profil d'élution se présente sous la forme d'un pic unique correspondant à l'homodimère de MBP-ZEBRA. De la même façon que lors de la purification de la protéine ZEBRA (voir chapitre II.2.1.), un petit pic se forme en amont du pic d'élution correspondant à

la protéine MBP-ZEBRA en complexe avec de l'ADN contaminant issu de la lyse des cellules *E.coli* ayant servi à exprimer la protéine.

2. Choix des oligonucléotides :

Nous avons choisi des oligonucléotides de 19 pb, c'est-à-dire de même longueur que ceux utilisés en cristallogénèse. Ces oligonucléotides sont marqués à la rhodamine en 5' sur un seul brin (voir tableau...). L'heptamère reconnu par ZEBRA est flanqué des séquences présentes naturellement dans le promoteur du gène *BRLF1* de l'EBV (tableau V.1.). Nous avons au préalable vérifié qu'il n'y avait aucune différence d'affinité que l'on utilise des oligonucléotides présentant ces séquences flanquantes ou celles utilisées en cristallographie.

Nom	Séquence oligonucléotidique
AP1	-3 -2 -1 0 1 2 3 5' A A T A A A A T G A C T C A T A A G C^{3'} 3' T A T T T T T A C T G A G T A T T C G A F 3' -2' -1' 0' -1' -2' -3'
AP1(mut3)	-3 -2 -1 0 1 2 3 G G A C T C A C C T G A G T 3' -2' -1' 0' -1' -2' -3'
meZRE2	-3 -2 -1 0 1 2 3 T C^m G C T C A A G^m C G A G T 3' -2' -1' 0' -1' -2' -3'
ZRE2	-3 -2 -1 0 1 2 3 T C G C T C A A C C G A G T 3' -2' -1' 0' -1' -2' -3'
meZRE2(mut3)	-3 -2 -1 0 1 2 3 G C^m G C T C A C G^m C G A G T 3' -2' -1' 0' -1' -2' -3'
AP1/meZRE2	-3 -2 -1 0 1 2 3 T G G C T C A A C^m C G A G T 3' -2' -1' 0' -1' -2' -3'
meZRE2/AP1	-3 -2 -1 0 1 2 3 T C^m A C T C A A G T G A G T 3' -2' -1' 0' -1' -2' -3'
Hemi1-ZRE2	-3 -2 -1 0 1 2 3 T C G C T C A A G^m C G A G T 3' -2' -1' 0' -1' -2' -3'
Hemi2-ZRE2	-3 -2 -1 0 1 2 3 T C^m G C T C A A G C G A G T 3' -2' -1' 0' -1' -2' -3'

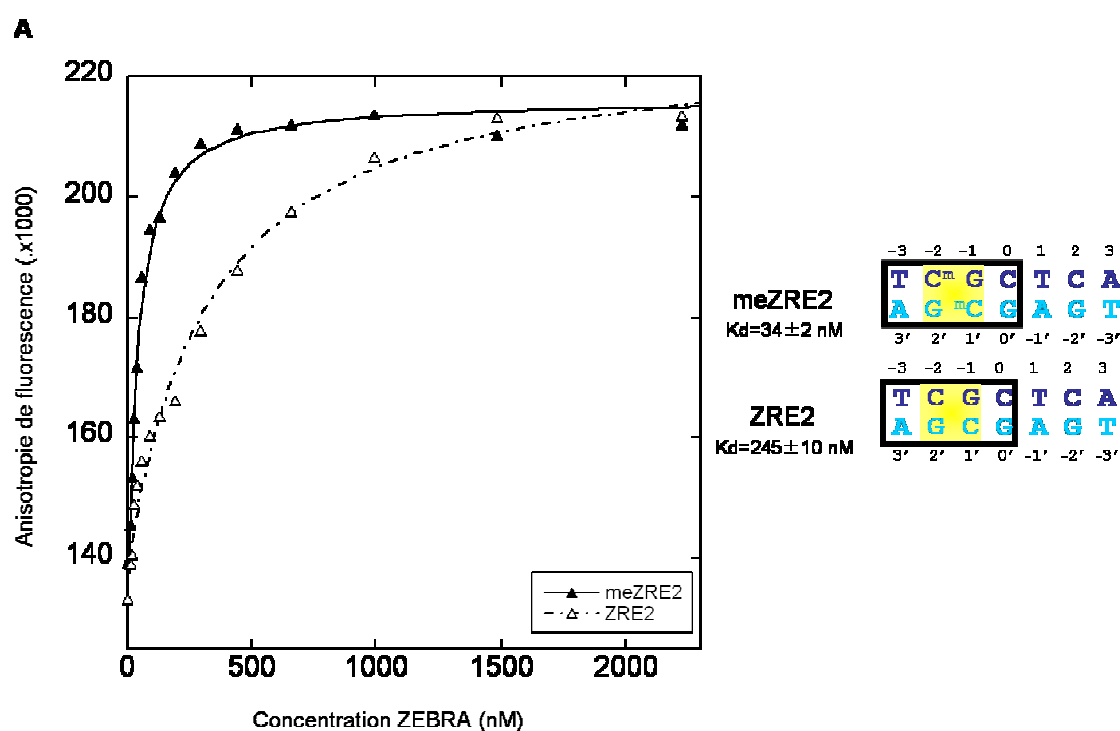
Tableau V.1.: Récapitulatif des oligonucléotides utilisés lors des expériences d'anisotropie de fluorescence. Les paires de bases surlignées en jaune sont caractéristiques de l'heptamère meZRE2, celles surlignées en rose caractéristiques du site AP1 et celles surlignées en gris indiquent la mutation en position -3/3'. Les séquences flanquant l'heptamère sont présentées sur le 1^{er} exemple « AP1 ». Le brin marqué à la rhodamine est indiqué par la lettre « F ». Les oligonucléotides « hémi1-ZRE2 » et « hémi2-ZRE2 » (pour hémi-méthylé) représentent le site meZRE2 où seul un des deux brins est méthylé. Les oligonucléotides « AP1(mut3) » et « meZRE2 (mut3) » ont leur paire de bases en position -3/3' mutée en « G-C ». Enfin, les oligonucléotides « AP1/meZRE2 » et « meZRE2/AP1 » représentent des hybrides des sites AP1 et meZRE2.

III. Affinité de ZEBRA pour les mutants du site ZRE2 méthylé

Les courbes de fixation de ZEBRA sur les différents oligonucléotides ont été analysées selon la méthode décrite dans le chapitre VIII, Matériels et Méthodes. Précisons que le but de cette étude est de comparer entre elles les affinités relatives de ZEBRA pour ces différents sites (et non d'obtenir les valeurs de K_d absolues) afin de déterminer le mécanisme de reconnaissance préférentielle de l'ADN méthylé par ZEBRA.

1. Reconnaissance différentielle du groupement méthyle :

Nous avons tout d'abord mesuré et comparé les affinités de ZEBRA pour le site ZRE2 dans ses états méthylé et non méthylé :



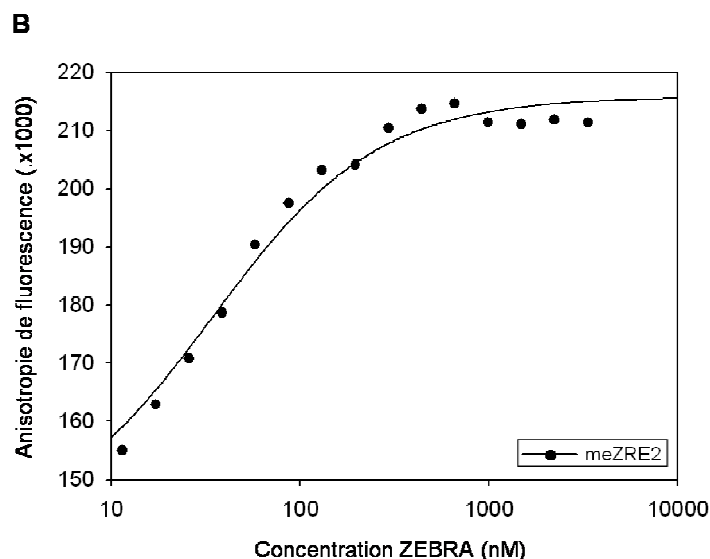


Figure V.4. : (A) Courbes de fixation de ZEBRA sur les oligonucléotides meZRE2 (▲) et ZRE2 (Δ). (B) Représentation logarithmique de la courbe de fixation de ZEBRA sur l'oligonucléotide meZRE2.

La courbe de fixation a une forme hyperbolique (figure V.4.A.). De plus, la représentation logarithmique de cette courbe montre que les valeurs d'anisotropie augmentent de $AF_{\min}$ à $AF_{\max}$ en 2 unités de log. Ces caractéristiques sont celles d'une fixation simple indiquant que ZEBRA se fixe en tant que dimère sur l'heptamère ADN. Les valeurs d'anisotropie mesurées sont typiques de ce genre d'expérience avec un écart d'amplitude d'anisotropie de 80 en moyenne.

Le complexe ZEBRA/meZRE2 a une constante de dissociation (K_d) de 34 ± 2 nM alors que le complexe ZEBRA/ZRE2 a un K_d de 245 ± 10 nM. ZEBRA présente une affinité 7 fois plus élevée pour l'état méthylé du site ZRE2 que pour son état non méthylé (ce qui est en accord avec les données publiées par Prasanna Bhende en 2004).

D'après la structure cristallographique du complexe ZEBRA/meZRE2, les cytosines méthylées C^{-2} et $C^{1'}$ sont toutes deux proches de la sérine 186. Mais jouent-elles toutes les deux un rôle crucial dans la reconnaissance préférentielle du site meZRE2 par ZEBRA ? Pour répondre à cette question, nous avons testé la fixation de ZEBRA sur les sites hemi1-ZRE2 (uniquement méthylé sur la cytosine 1') et hemi2-ZRE2 (uniquement méthylé sur la cytosine - 2) :

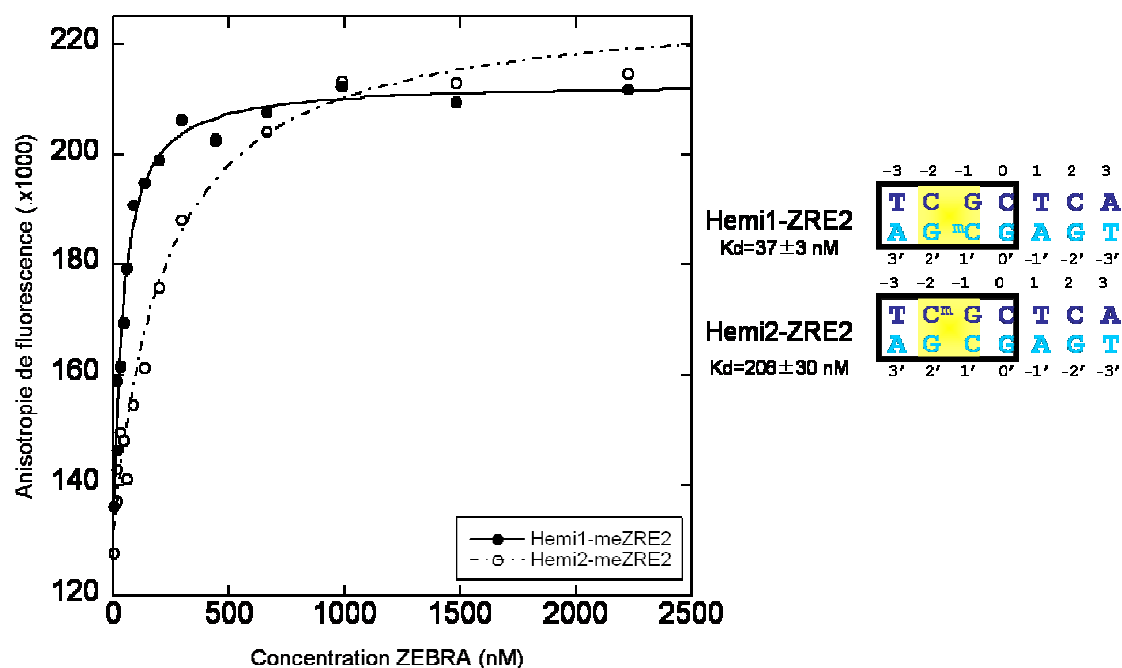


Figure V.5. : Courbes de fixation de ZEBRA sur les oligonucléotides hemi1-ZRE2 (●) et hemi2-ZRE2 (○).

ZEBRA a une affinité pour l'heptamère hemi1-ZRE2 ($K_d = 37 \pm 3 \text{ nM}$) nettement supérieure à celle pour le site hemi2-ZRE2 ($K_d = 206 \pm 30 \text{ nM}$).

En effet, ZEBRA fixe les oligonucléotides meZRE2 et hemi1-ZRE2 avec la même affinité (figure V.6.). De même, les complexes ZEBRA/ZRE2 et ZEBRA/hemi2-ZRE2 présentent sensiblement la même constante de dissociation (figure V.6.).

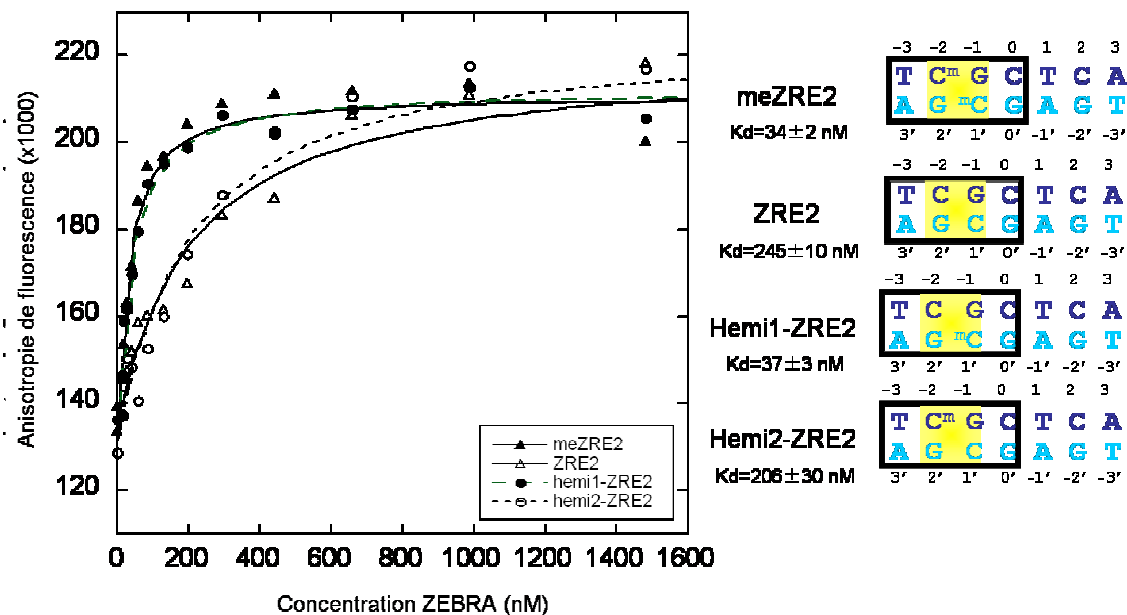


Figure V.6. : Courbes de fixation de ZEBRA sur les oligonucléotides meZRE2 (▲), ZRE2 (△), hemi1-ZRE2 (●) et hemi2-ZRE2 (○).

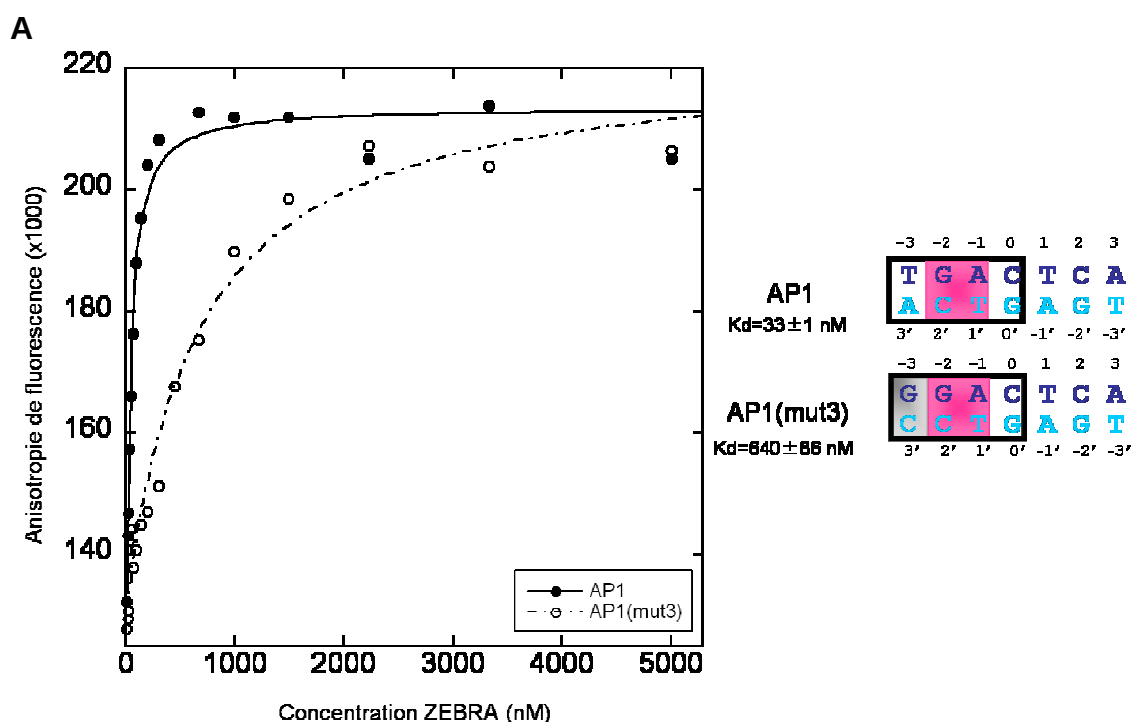
Ces résultats indiquent que seule la cytosine méthylée C^{1'} joue un rôle crucial dans la reconnaissance préférentielle du site meZRE2 par ZEBRA.

Mais est-ce l'identité de cette cytosine C^{1'} en position 1' qui est importante ou seulement le groupement méthyle qu'elle porte ? Afin de répondre à cette question, nous avons muté la paire de bases G-^mC en position -1/1' en paire A-T et obtenu ainsi l'oligonucléotide « meZRE2/AP1 » (voir ci-dessous, partie IV.3.). En effet une base thymine comporte également un groupement méthyle en même position qu'une cytosine méthylée (voir Chapitre IV., figure IV.2.). D'après les expériences d'anisotropie de fluorescence, ZEBRA a la même affinité pour meZRE2 ($K_d = 34 \pm 2$ nM) que pour meZRE2/AP1 ($K_d = 36 \pm 2$ nM). Ce résultat nous permet d'affirmer que l'identité de la pyrimidine en position 1' n'est pas cruciale, pourvu qu'elle porte un groupement méthyle.

2. Reconnaissance différentielle de la paire de bases -3/3' flanquant le site CpG méthylé :

La comparaison des structures de ZEBRA en complexe avec les sites AP1 et meZRE2, suggère un rôle différent de la paire de bases en position -3/3' : en complexe avec AP1, l'asparagine 182 forme une liaison hydrogène directe avec l'oxygène O4 de la thymine T⁻³ (figure V.10.) alors qu'en complexe avec le site meZRE2, elle ne forme plus qu'une liaison indirecte médiée par une molécule d'eau avec l'atome N4 de l'adénine A^{3'} (figure V.11.). Cette observation suggère qu'une mutation de la paire de bases en position -3/3' aurait un effet moins perturbant dans le cas du complexe ZEBRA/meZRE2 que dans le complexe ZEBRA/AP1.

Afin de vérifier cette hypothèse, nous avons comparé les affinités de ZEBRA pour les oligonucléotides AP1(mut3) et meZRE2(mut3).



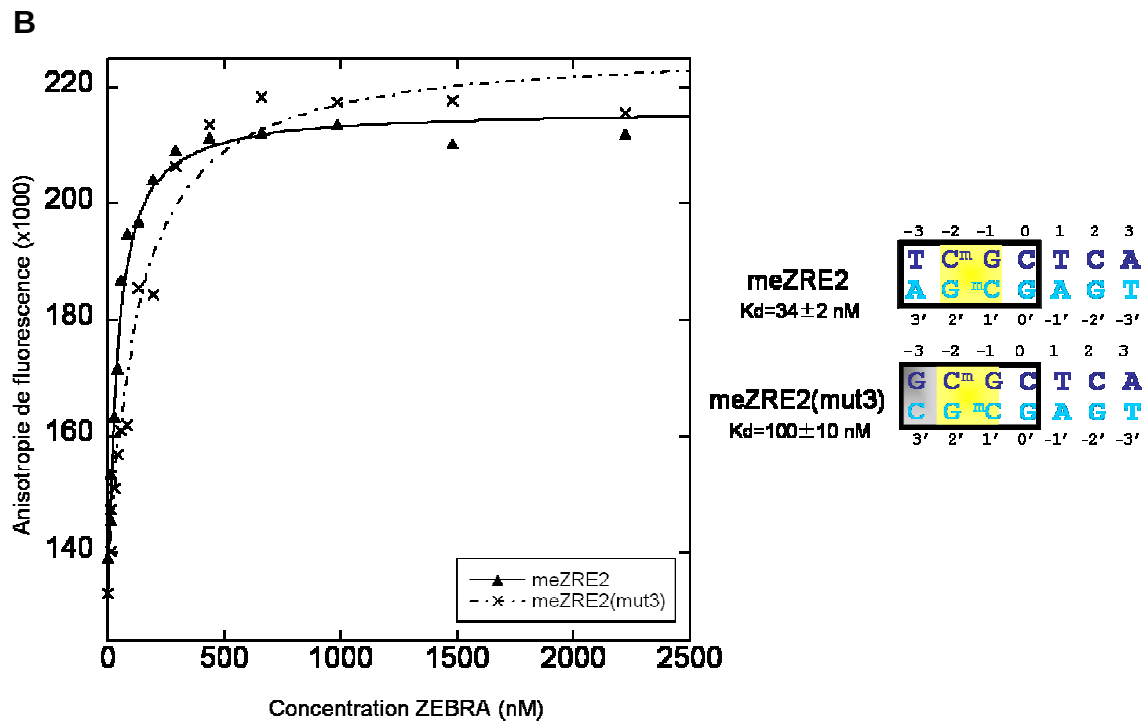


Figure V.7. : (A) Courbes de fixation de ZEBRA sur les oligonucléotides AP1 (●) et AP1(mut3) (○). (B) Courbes de fixation de ZEBRA sur les oligonucléotides meZRE2 (▲) et meZRE2(mut3) (×).

ZEBRA a une affinité pour le site AP1 ($K_d = 33 \pm 1$ nM) similaire à celle pour le site meZRE2. La mutation de la paire de bases en position -3/3' (site AP1(mut3)) entraîne une chute brutale de cette affinité d'un facteur 20 ($K_d = 640 \pm 86$ nM). Cependant, la même mutation dans le site meZRE2 a un effet beaucoup moins drastique puisque ZEBRA fixe l'oligonucléotide meZRE2(mut3) avec un K_d de 100 ± 10 nM, soit avec une affinité seulement 3 fois moins importante que pour le site meZRE2.

L'identité de la paire de bases en position -3/3' a donc un rôle moins crucial pour le complexe ZEBRA/meZRE2 que pour le complexe ZEBRA/AP1, comme décrit par l'analyse des structures de ZEBRA en complexe avec les sites AP1 et meZRE2.

3. Reconnaissance indépendante du site meZRE2 par la sérine 186 et l'asparagine 182 :

D'après la prédiction de structure du complexe ZEBRA/meZRE2, Carlo Petosa et ses collaborateurs ont proposé une hypothèse sur le mécanisme de reconnaissance préférentielle par ZEBRA des promoteurs méthylés (Petosa et al., 2006). La sérine 186, stabilisée par le groupement méthylé de la cytosine C⁻², interagirait avec l'asparagine 182 qui, à son tour, formerait un lien hydrogène avec la guanine G^{2'}, augmentant ainsi l'affinité de ZEBRA pour le site meZRE2 (voir Chapitre I, section V.3.). D'après cette hypothèse, la mutation de la paire de bases G^mC en position -1/1' modifierait la conformation de la sérine 186, ce qui devrait avoir des conséquences sur l'asparagine 182 et donc sur l'affinité de ZEBRA pour le site meZRE2/AP1. De même, en mutant la paire de bases C^m-G en position -2/2' en paire G-C, l'asparagine 182 perdrait les liaisons hydrogènes qu'elle formait avec la guanine G^{2'} et avec la sérine 186, ce qui devrait déstabiliser le complexe ZEBRA/(AP1/meZRE2). En fait, la structure du complexe ZEBRA/meZRE2 ne montre aucune liaison hydrogène entre la sérine 186 et l'asparagine 182. Ce qui suggère, contrairement à ce qui a été prédit, que ces deux résidus reconnaissent le site meZRE2 indépendamment l'un de l'autre. Pour vérifier cela, l'affinité de ZEBRA pour les sites hybrides (AP1/meZRE2) et (meZRE2/AP1) a été mesurée.

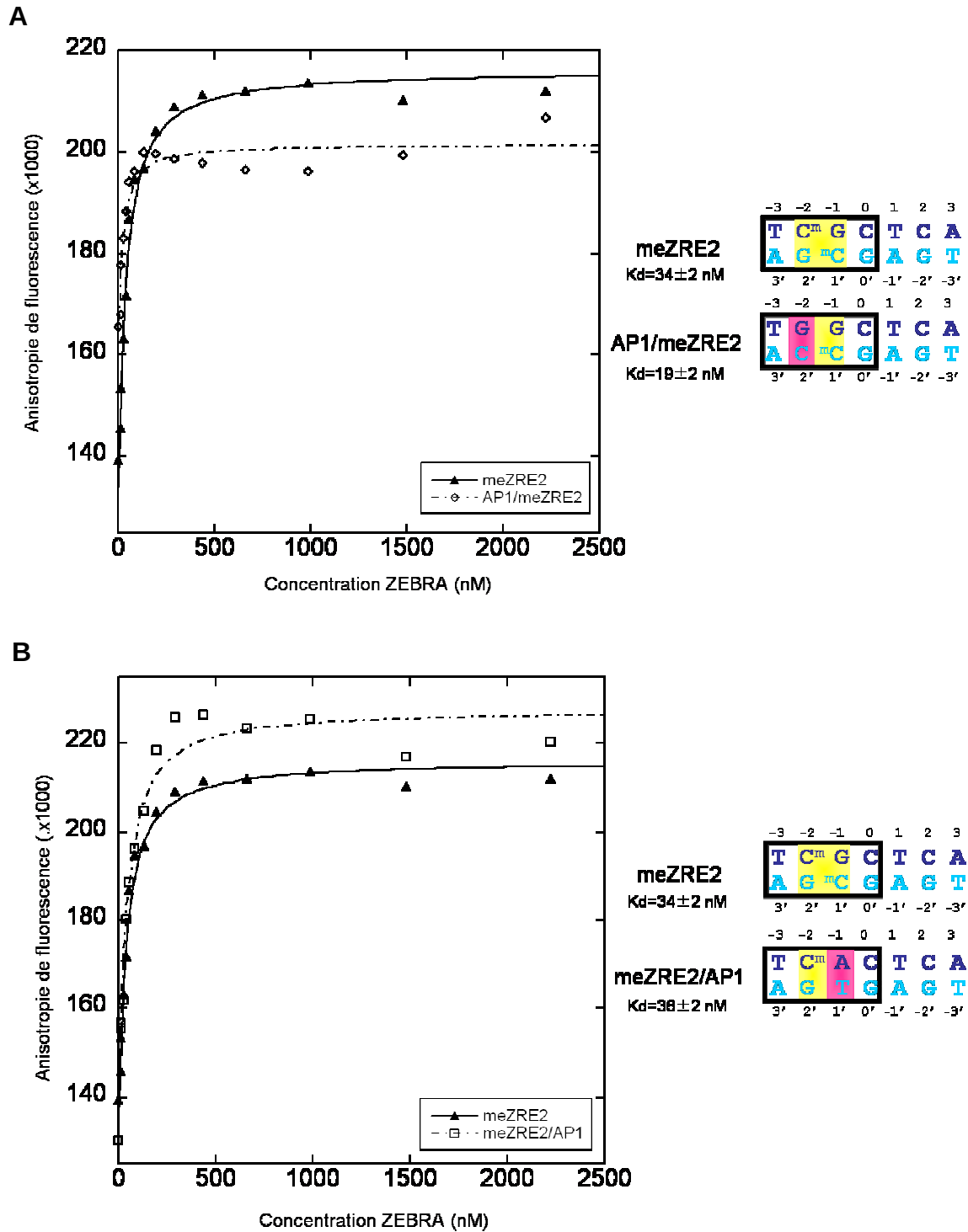


Figure V.8. : (A) Courbes de fixation de ZEBRA sur les oligonucléotides meZRE2 (▲) et AP1/meZRE2 (◊). (B) Courbes de fixation de ZEBRA sur les oligonucléotides meZRE2 (▲) et meZRE2/AP1 (◻). Même si les plateaux des courbes ne sont pas atteints pour les mêmes valeurs, la superposition des courbes à basses concentrations de [ZEBRA] montre que les courbes ont sensiblement le même K_d .

ZEBRA présente les mêmes constantes de dissociation pour les sites meZRE2 ($K_d = 34 \pm 2$ nM) et meZRE2/AP1 ($K_d = 36 \pm 2$ nM) alors que son affinité pour le site AP1/meZRE2 est sensiblement plus importante ($K_d = 19 \pm 2$ nM). Or, selon l'hypothèse proposée, l'affinité de ZEBRA pour ces sites hybrides aurait du chuter considérablement étant donné l'interaction prédite entre la sérine 186 et l'asparagine 182. Donc, le fait que ZEBRA ait la même affinité ou même, une meilleure affinité pour les sites meZRE2/AP1 et AP1/meZRE2 suggère que la mutation de la paire de bases en position -1/1' ou en position -2/2' du site meZRE2 ne compromet pas la fixation de ZEBRA sur le site meZRE2. Contrairement à l'hypothèse de Carlo Petosa, la sérine 186 et l'asparagine 182 reconnaissent le site meZRE2 indépendamment l'une de l'autre, comme observé dans la structure du complexe ZEBRA/meZRE2.

4. Discussion

Les mesures d'affinité de ZEBRA pour les mutants du site meZRE2 et leur corrélation avec la structure cristallographique du complexe ZEBRA/meZRE2 nous permettent de proposer trois conclusions préliminaires (voir tableau V.2.) : les deux groupements méthyles présents sur les cytosines $C^{1'}$ et $C^{2'}$ n'ont pas la même importance dans le mécanisme de reconnaissance préférentielle du site meZRE2 par ZEBRA, la paire (en position -3/3') flanquant l'îlot CpG méthylé du site meZRE2 ne joue pas le même rôle que dans le complexe ZEBRA/AP1 et les interactions de la sérine 186 et de l'asparagine 182 avec le site meZRE2 sont indépendantes l'une de l'autre.

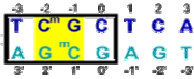
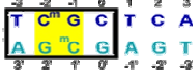
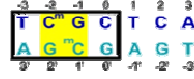
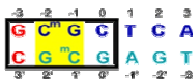
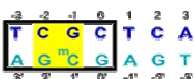
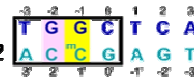
Reconnaissance différentielle du groupement Me	K_d (nM)	Reconnaissance différentielle de la paire de base -3/3'	K_d (nM)	Reconnaissance indépendante des paires de base -3/3' et -2/2'	K_d (nM)
meZRE2 	34 ± 2	meZRE2 	34 ± 2	meZRE2 	34 ± 2
ZRE2 	245 ± 10	meZRE2 (mut3) 	100 ± 10	AP1 	33 ± 1
Hemi1-ZRE2 	37 ± 3	AP1 	33 ± 1	meZRE2 / AP1 	36 ± 2
Hemi2-ZRE2 	208 ± 30	AP1 (mut3) 	840 ± 88	AP1 / meZRE2 	19 ± 2

Tableau V.2. : Récapitulatif des constantes de dissociation (K_d) des complexes ZEBRA/ADN en fonction de l'oligonucléotide testé.

Des deux cytosines méthylées $C^{1'}$ et C^{-2} , seule la cytosine méthylée $C^{1'}$ joue un rôle crucial dans la reconnaissance préférentielle du site meZRE2 par ZEBRA. En effet, la cytosine $C^{1'}$ est engagée dans une liaison hydrogène avec l'hydroxyle de la sérine 186 via son atome N4. De plus, le groupement méthyle de cette cytosine ne se trouve qu'à 4,25 Å du groupement méthylène de la sérine 186 (figure V.9.), lui permettant ainsi de former un contact hydrophobe avec cette dernière. Or, la chaîne latérale d'une sérine peut adopter 3 conformations préférentielles : *gauche*⁻, *gauche*⁺ et *trans*. Le rotamère *gauche*⁺ observé dans la structure cristallographique du complexe ZEBRA/meZRE2 serait stabilisé par ce contact hydrophobe entre le groupement méthyle de la cytosine $C^{1'}$ et le groupement méthylène de la sérine 186.

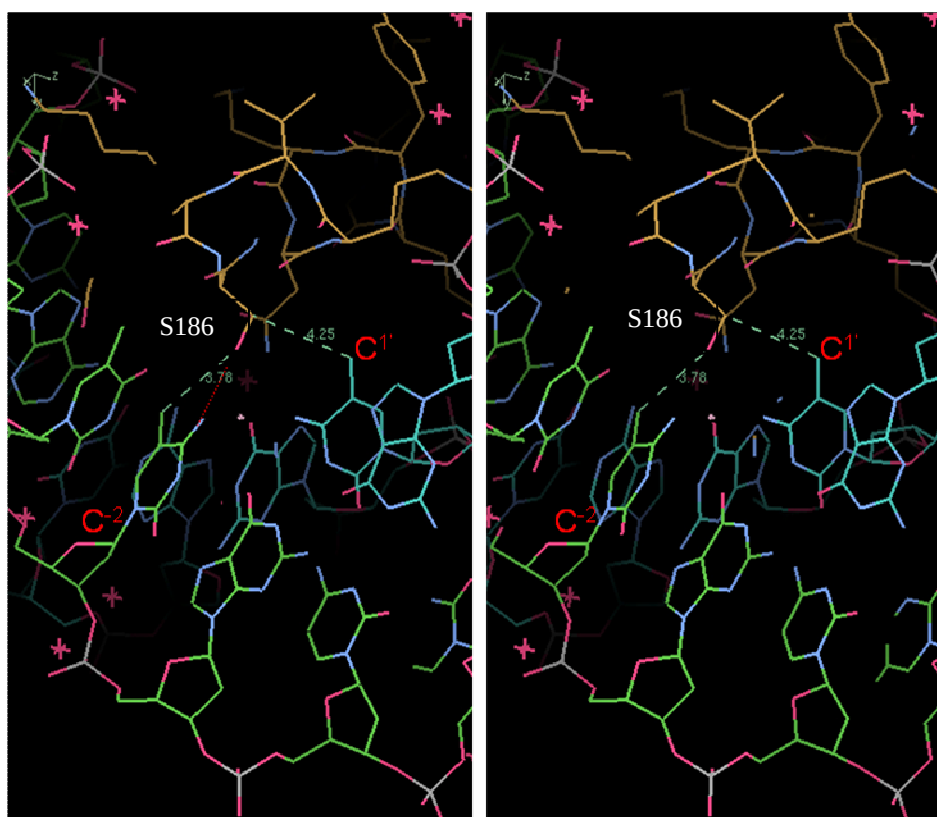


Figure V.9. : Vues en stereo (« *cross-eyed stereo* ») de l'interaction entre la sérine 186 et les cytosines méthylées C⁻² et C^{1'} du site meZRE2.

Dû au rétrécissement du sillon majeur d'ADN induit par la présence de l'îlot CpG (voir Chapitre IV., figure IV.6.), la cytosine C⁻² peut elle aussi engager une liaison hydrogène avec l'hydroxyle de la sérine 186 via son atome N4. De ce fait, son groupement méthyle se retrouve à proximité de l'hydroxyle de la sérine 186 (3,8 Å). Cependant, il n'influence pas la conformation de cette dernière et c'est pourquoi cette cytosine méthylée C⁻² ne jouerait pas un rôle crucial dans la reconnaissance du site meZRE2.

L'identité de la paire de bases en position -3/3' a un rôle moins crucial pour le complexe ZEBRA/meZRE2 que pour le complexe ZEBRA/AP1. Dans le complexe ZEBRA/AP1(mut3), l'asparagine 182 ne peut plus former avec la guanine G⁻³ la liaison hydrogène qu'elle formait auparavant avec la thymine T³ du site AP1 (base encadrée sur la figure V.10.). Cette perte d'une liaison hydrogène directe affecte grandement la stabilité du complexe ZEBRA/AP1(mut3).

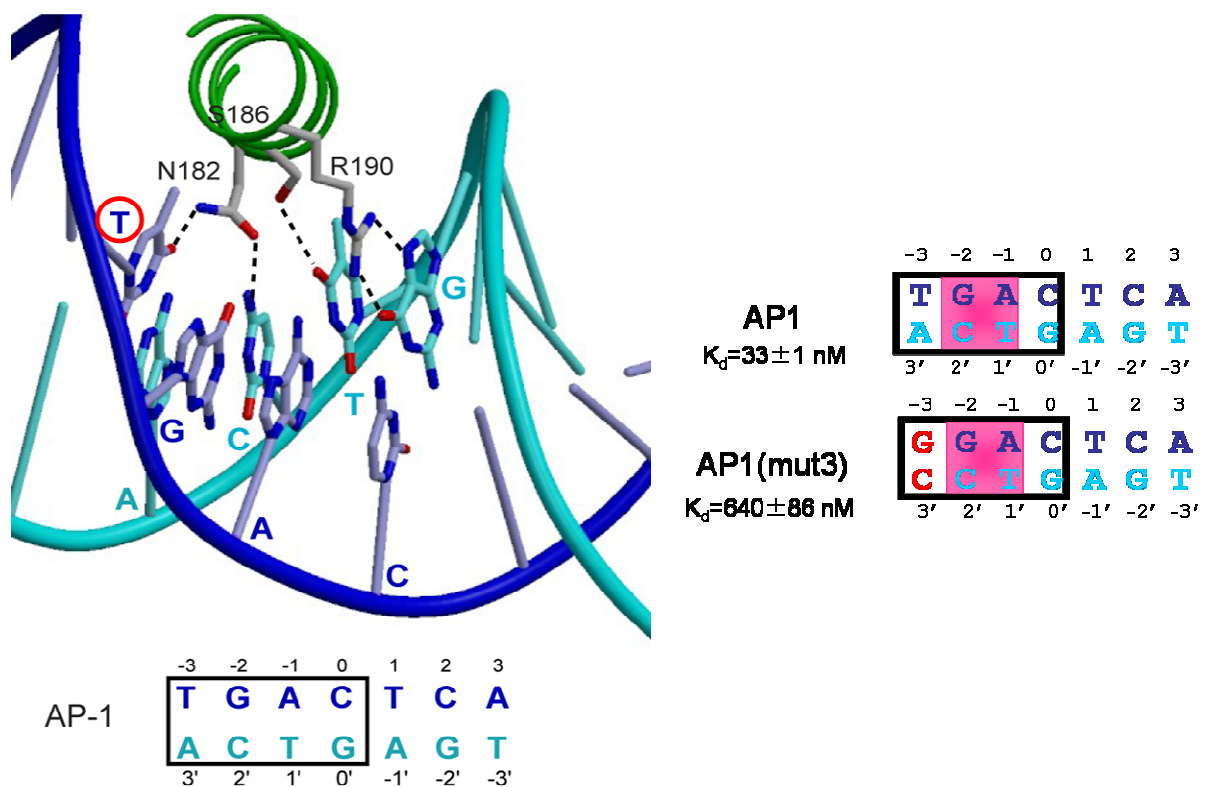


Figure V.10. : Agrandissement de la structure du complexe ZEBRA/AP1 représentant les interactions protéine/ADN au niveau du sillon majeur de l'ADN. Les liaisons hydrogènes sont représentées en pointillés. A droite sont représentées les séquences des oligonucléotides AP1 et AP1(mut3).

En ce qui concerne le complexe ZEBRA/meZRE2(mut3), la mutation de la paire de bases en position -3/3' A-T en G-C n'affecte que la liaison indirecte via la molécule d'eau qu'entretenait l'asparagine 182 avec l'adénine A^{3'} (base encerclée sur la figure V.11.). Les conséquences sur l'affinité de ZEBRA pour ce site meZRE2(mut3) sont donc moins importantes.

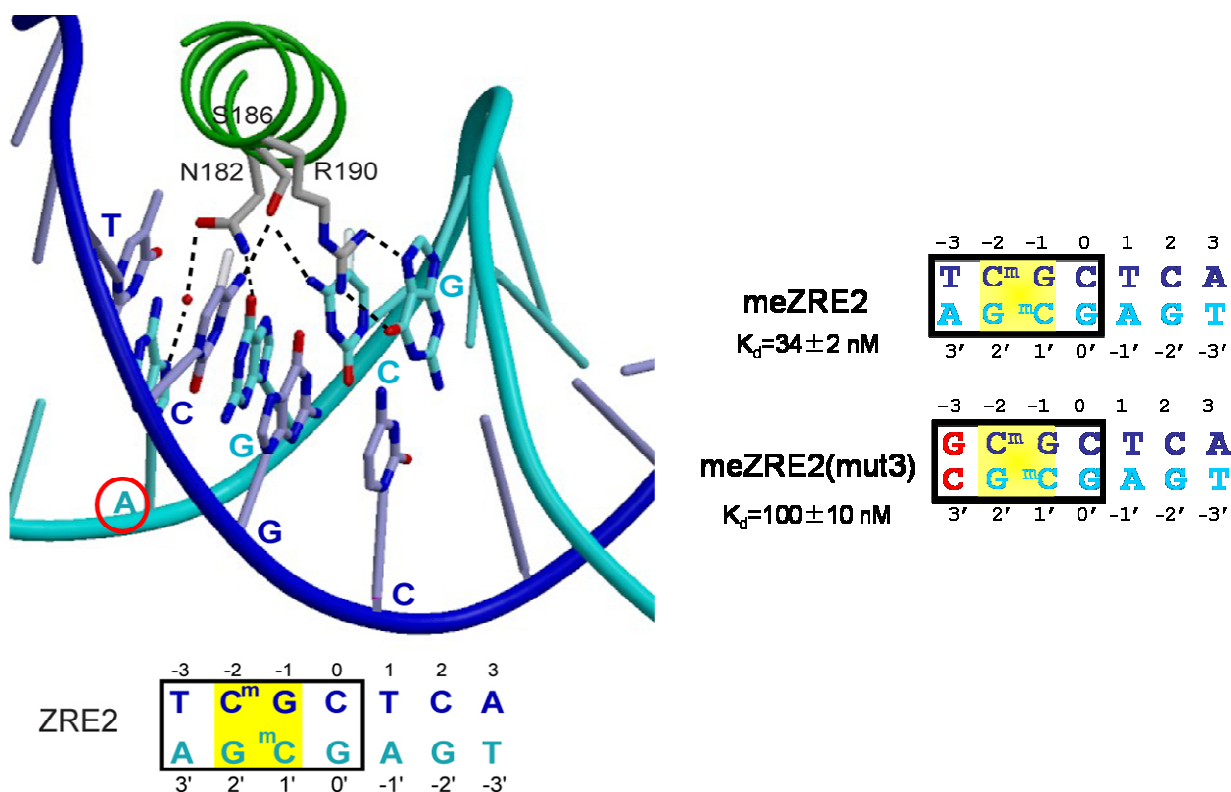


Figure V.11. : Agrandissement de la structure du complexe ZEBRA/meZRE2 représentant les interactions protéine/ADN au niveau du sillon majeur de l'ADN. Les liaisons hydrogènes sont représentées en pointillés. La molécule d'eau est représentée par une sphère rouge. Les séquences des oligonucléotides meZRE2 et meZRE2(mut3) sont indiquées à droite.

Contrairement à la prédiction de structure du complexe ZEBRA/meZRE2, la sérine 186 n'interagit pas avec l'asparagine 182 par un lien hydrogène. Les résultats des expériences d'anisotropie de fluorescence semblent indiquer que ces deux résidus reconnaissent le site meZRE2 indépendamment l'un de l'autre.

D'après la structure cristallographique de ce complexe, la sérine 186 forme des liaisons hydrogènes avec les paires de bases -1/1' et -2/2' tandis que l'asparagine 182 interagit directement avec la paire de bases -2/2' et via une molécule d'eau avec la paire de bases -3/3'. En mutant la paire de bases -2/2' (C^m-G) en G-C (oligonucléotide AP1/meZRE2), le premier quart du site devient identique au site AP1. L'asparagine 182 pourrait alors changer de conformation et retrouver les deux liaisons hydrogènes directes qu'elle formait avec la thymine T⁻³ et la cytosine C^{2'} (figure V.10. ci-dessus) du site AP1. Et d'après les mesures d'affinité de ZEBRA pour le site AP1/meZRE2, ce changement de conformation de l'asparagine 182 n'a aucune incidence sur le comportement de la sérine 186 et donc sur

l'affinité de ZEBRA pour ce site. Ces deux résidus reconnaissent donc le site meZRE2 indépendamment l'un de l'autre. Cette observation est confirmée par les mesures d'affinité de ZEBRA pour le site hybride meZRE2/AP1 : la mutation de la paire de bases -1/1' (G-C^m) en A-T transforme le deuxième quart du site meZRE2 en site AP1. La sérine 186 pourrait de nouveau interagir avec la thymine T^{1'} sans que cela affecte la conformation de l'asparagine 182 puisque les affinités pour les sites meZRE2 et meZRE2/AP1 sont équivalentes.

IV. Identification des résidus de ZEBRA suffisants à une fixation préférentielle du site ZRE2 méthylé

Nous avons identifié les bases nucléotidiques qui jouent un rôle crucial dans la reconnaissance de l'ADN méthylé par ZEBRA. Qu'en est-il concernant les résidus de ZEBRA impliqués dans cette reconnaissance ? Pour répondre à cette question, le comportement de ZEBRA vis-à-vis des sites meZRE2 et ZRE2 a été comparé à cJun, une autre protéine bZIP classique. Un clone chimère (Chimère ZEBRA-cJun) a été construit comportant la région basique de ZEBRA fusionnée au domaine de dimérisation de cJun, afin de déterminer si les résidus responsables de la distinction entre ADN méthylé et non méthylé se trouvaient uniquement dans la région basique de ZEBRA. Enfin, d'après la structure cristallographique du complexe ZEBRA/meZRE2 et les travaux de Prasanna Bhende en 2005 suggérant un rôle clé de la sérine 186 dans la fixation du site meZRE2, nous nous sommes demandé si cette sérine était suffisante au mécanisme de reconnaissance préférentielle ou si d'autres résidus de ZEBRA étaient impliqués de manière cruciale. Nous avons donc construit le mutant cJun A275S possédant l'équivalent de la sérine 186 de ZEBRA. Les séquences de ces différentes protéines sont présentées ci-dessous (figure V.12.) :

	175	186	196	206	216	226
ZEBRA	LEIKRYKNRVA	SRKCRAKFKQ	LLQHYREVAA	AKSSENDRLR	LLLKQMCPSL	DVDSIIPRTPD
	264	275	285	295	305	315
cJun	AERKRMNRRIA	ASKSRKRKLE	RIARLEEKVK	TLKAQNSELA	STANNLREQV	AQLKQKVMNH
	175	186	197/287	295	305	315
Chimère	LEIKRYKNRVA	SRKCRAKFKQ	LIARLEEKVK	TLKAQNSELA	STANNLREQV	AQLKQKVMNH
ZEBRA-Jun						

Figure V.12. : Séquence en acides aminés des protéines ZEBRA, cJun et de la chimère ZEBRA-Jun.

Contrairement au dimère de ZEBRA, l'homodimère de cJun est moins stable. cJun fixe ainsi l'ADN en tant que monomère puis recrute le deuxième monomère dans un deuxième temps (Kohler and Schepartz, 2001). La fixation de cJun sur l'ADN est donc coopérative, représentée par une forme sigmoïdale des courbes de fixation (non montrées). Nous avons décidé de représenter l'axe des abscisses en échelle logarithmique afin de mieux mettre en évidence l'accord entre les courbes ajustées et les données expérimentales. La méthode de détermination des K_d a été ajustée pour prendre en compte l'allure sigmoïdale de ces courbes (voir Chapitre VIII, Matériels et Méthodes).

1. cJun ne fait aucune distinction entre les sites meZRE2 et ZRE2 :

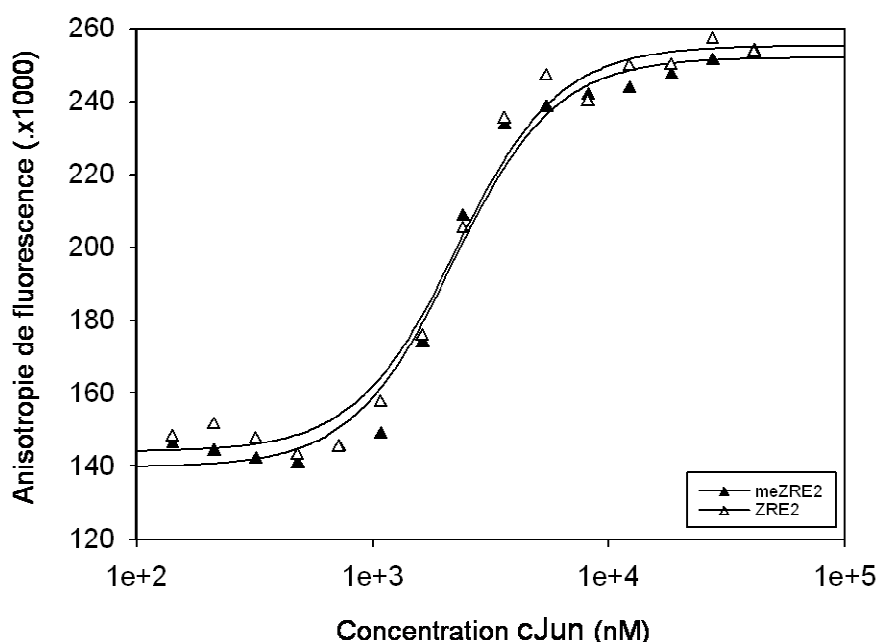


Figure V.13. : Courbes de fixation de cJun sur les oligonucléotides meZRE2 (▲) et ZRE2 (Δ).

Les valeurs d'anisotropie de fluorescence obtenues sont similaires à celles mesurées pour ZEBRA, avec un écart d'amplitude légèrement plus élevé de 120 en moyenne. Les valeurs d'anisotropie augmentent de AF_{min} à AF_{max} en 1 unité de log, ce qui est caractéristique d'une fixation coopérative.

cJun présente des constantes de dissociation similaires pour les sites meZRE2 ($K_d = 2,26 \pm 0,19 \mu\text{M}$) et ZRE2 ($K_d = 2,28 \pm 0,21 \mu\text{M}$) : il ne reconnaît pas préférentiellement un de ces deux sites.

2. Le domaine de liaison à l'ADN de ZEBRA permet à cJun de reconnaître préférentiellement le site meZRE2 :

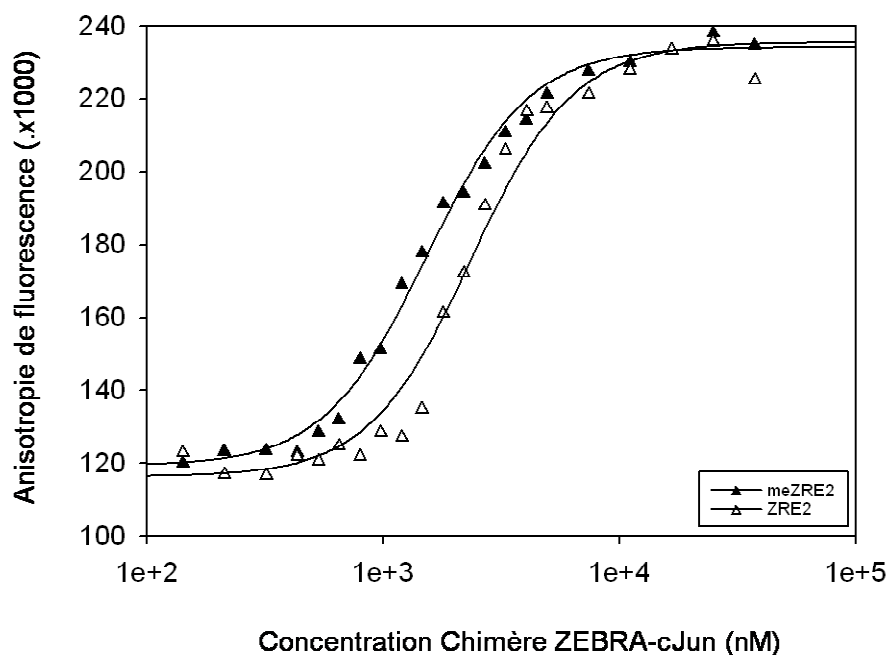


Figure V.14. : Courbes de fixation du clone chimère ZEBRA-cJun sur les oligonucléotides meZRE2 (▲) et ZRE2 (Δ).

La courbe de la protéine chimère en complexe avec le site meZRE2 s'est décalée vers la gauche, soit vers des valeurs de K_d plus faibles. La protéine chimère ZEBRA/cJun reconnaît préférentiellement le site meZRE2 ($K_d = 1,53 \pm 0,07 \mu\text{M}$) soit avec une affinité 1,5 fois plus importante que pour le site ZRE2 ($K_d = 2,35 \pm 0,18 \mu\text{M}$). Cependant, cette différence d'affinité est faible puisqu'elle atteignait un facteur 7 pour la protéine ZEBRA. Ceci semble indiquer que d'autres résidus du domaine de dimérisation de ZEBRA seraient impliqués dans la reconnaissance préférentielle de l'ADN méthylé. Toutefois, nous pouvons conclure que la région basique de ZEBRA est suffisante pour permettre à cJun de différencier un site Adn méthylé d'un site non méthylé.

3. La sérine 186 de ZEBRA est suffisante pour que cJun distingue les sites meZRE2 et ZRE2 :

La sérine 186, nécessaire à la reconnaissance préférentielle de l'ADN méthylé, est unique à ZEBRA (Bhende et al., 2005) : cJun, comme toutes les autres protéines bZIPs connues, possède une alanine à la place (A275). Pour vérifier si cette sérine 186 est suffisante pour distinguer un site ADN méthylé d'un site non méthylé ou si la région basique entière est nécessaire, nous avons construit un clone de cJun possédant cette sérine (cJun A275S).

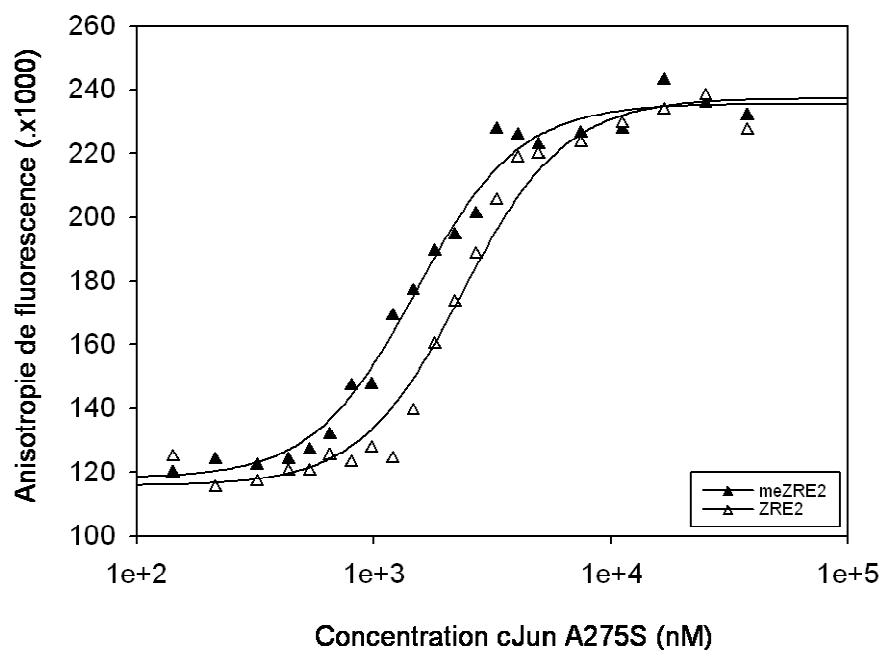


Figure V.15. : Courbes de fixation de cJun A275S sur les oligonucléotides meZRE2 (▲) et ZRE2 (Δ).

cJun A275S a une meilleure affinité pour le site meZRE2 ($K_d = 1,50 \pm 0,08 \mu\text{M}$) que pour le site ZRE2 ($K_d = 2,43 \pm 0,17 \mu\text{M}$). De plus, la différence d'affinité (1,5) observée entre les sites meZRE2 et ZRE2 est identique à celle calculée pour la protéine chimère ZEBRA-cJun : cela prouve que la sérine 186 de ZEBRA (représentée ici par la sérine 275) est suffisante pour distinguer et fixer préférentiellement un site ADN méthylé.

4. Discussion :

Les valeurs de K_d du bZIP de cJun et de ses mutants pour les sites meZRE2 et ZRE2 sont résumées dans le tableau V.3. :

Protéine	K_d (μ M)	
	meZRE2	ZRE2
cJun	$2,26 \pm 0,19$	$2,28 \pm 0,21$
Chimère ZEBRA-cJun	$1,53 \pm 0,07$	$2,35 \pm 0,18$
cJun A275S	$1,50 \pm 0,08$	$2,43 \pm 0,17$

Tableau V.3. : Récapitulatif des valeurs de K_d des mutants de cJun pour les sites meZRE2 et ZRE2.

Remarquons que ces valeurs de K_d sont bien plus élevées que celles obtenues pour ZEBRA. En effet, l'homodimère de cJun a 30 fois moins d'affinité pour ses sites cibles que l'hétérodimère cJun-cFos (Kohler and Schepartz, 2001).

Comme nous nous y attendions, la protéine cJun ne fait aucune distinction entre les sites meZRE2 et ZRE2 : cette capacité à fixer préférentiellement le site ZRE2 dans son état méthylé est unique à ZEBRA. Par ailleurs, le remplacement de la région basique de cJun par celle de ZEBRA permet à cJun de distinguer entre les deux sites ADN. De plus, la mutation de l'alanine 275 de cJun en sérine (équivalente à la sérine 186 de ZEBRA) permet également de fixer préférentiellement le site meZRE2. Or, la différence d'affinité de la protéine chimère ZEBRA-cJun pour les sites méthylé et non méthylé est la même que pour la protéine cJun A275S possédant l'équivalent de la sérine 186. Ceci prouve que la sérine 186 de ZEBRA est suffisante pour distinguer un site ADN méthylé d'un site non méthylé. Cependant, la faible différence d'affinité d'un facteur 1,5 semble indiquer que d'autres résidus du domaine de dimérisation de ZEBRA seraient impliqués dans la reconnaissance préférentielle de l'ADN méthylé.

V. Hypothèses sur le mécanisme de reconnaissance de l'ADN méthylé par ZEBRA

Après avoir étudié les interactions protéine-ADN au niveau du sillon majeur du site meZRE2, nous pouvons émettre une révision de l'hypothèse proposée précédemment par Petosa et ses collaborateurs sur le mécanisme de fixation préférentielle de l'ADN méthylé par ZEBRA.

Au niveau de l'ADN, l'îlot CpG du site meZRE2 entraîne un rétrécissement du sillon majeur de l'ADN, qui se produit probablement indépendamment de l'état de méthylation de cet îlot. Ainsi, le groupement méthyle de la cytosine C^{1'} stabilise la sérine 186 dans une conformation *gauche*⁺. En conséquence, la sérine 186 forme un nouveau lien hydrogène avec l'atome N4 de la cytosine C⁻², ce qui stabilise d'autant plus le complexe ZEBRA/meZRE2. Des deux cytosines méthylées du site meZRE2, nous avons prouvé que seule la cytosine méthylée C^{1'} jouait un rôle crucial dans le mécanisme de reconnaissance préférentielle du site meZRE2 par ZEBRA.

Au niveau des résidus de ZEBRA impliqués dans la reconnaissance du site meZRE2, nous avons démontré que l'asparagine 182 ne jouait pas un rôle crucial dans le mécanisme de fixation préférentielle : elle conserve la liaison hydrogène qu'elle établissait avec la base en position 2' dans le cas du complexe ZEBRA/AP1 mais n'interagit plus avec la base en position -3. Cependant, elle satisfait son potentiel de liaisons hydrogènes en interagissant avec la base en position 3' via une molécule d'eau. De plus, contrairement à l'hypothèse proposée, l'asparagine 182 et la sérine 186 reconnaissent le site meZRE2 indépendamment l'une de l'autre. Enfin, bien que d'autres résidus du domaine de dimérisation de ZEBRA semblent nécessaires pour obtenir une différence d'affinité conséquente entre deux sites méthylé et non méthylé, la sérine 186 joue un rôle clé dans le mécanisme de reconnaissance préférentielle du site meZRE2 puisque sa seule présence dans une protéine bZIP classique tel que cJun, permet à une protéine de distinguer entre un site ADN méthylé ou non.

CHAPITRE VI :

CRIBLAGE CHIMIQUE A HAUT DEBIT

D'INHIBITEURS DE ZEBRA

En parallèle des travaux concernant le sujet principal de cette thèse, nous avons établi une collaboration avec la plate-forme de criblage chimique à haut-débit de l'EMBL à Heidelberg afin d'identifier de petites molécules qui pourraient inhiber la fixation à l'ADN de ZEBRA. ZEBRA est une cible prometteuse pour un traitement chimiothérapeutique étant donné que l'induction du cycle lytique de l'EBV est complètement bloquée quand l'expression de ZEBRA ou son activité de liaison à l'ADN sont inhibés. Or, en 2005, le rôle de ZEBRA dans la lymphomatogénèse a été démontré *in vitro* et chez la souris SCID (13, 14). Dans ces travaux, les souches EBV délétées pour ZEBRA n'entraînent pas de lymphomes chez la souris alors que la restauration de l'expression de ZEBRA re-déclenche les phénomènes lymphoprolifératifs via l'expression de cytokines cellulaires impliquées dans la multiplication cellulaire et l'angiogénèse (IL-6, IL10 et VEGF). De manière intéressante, dans ce modèle, l'inhibition du cycle lytique par des antiviraux agissant sur l'ADN polymérase de l'EBV n'empêche pas la lymphomatogénèse, suggérant que c'est la partie très précoce du cycle lytique qui est impliquée dans ces mécanismes. ZEBRA constitue donc une cible thérapeutique potentielle pour le traitement des lymphoproliférations associées à l'EBV chez l'immunodéprimé. Afin d'identifier d'éventuels inhibiteurs de ZEBRA, nous avons développé un essai à haut débit utilisant la technique de mesure d'anisotropie de fluorescence décrite au chapitre V.

I. ZEBRA : une cible idéale

Les composés capables d'inhiber le cycle lytique de l'EBV ont un intérêt considérable, puisque leur utilisation peut réduire le risque de développer des maladies associées à l'EBV (Gershburg and Pagano, 2005). En effet, l'expression des gènes lytiques de l'EBV stimule la croissance de cellules tumorales EBV-positives (Hong et al., 2005). Les agents pharmacologiques actuels ciblant le cycle lytique de l'EBV inhibent la réplication de l'ADN viral et empêchent l'expression de gènes tardifs de l'EBV. Cependant, ils ne peuvent affecter l'expression de plus de 30 gènes précoces du virus (Amon and Farrell, 2005). ZEBRA est une cible idéale pour un traitement chimiothérapeutique étant donné que l'induction du cycle lytique de l'EBV est complètement bloquée quand l'expression de ZEBRA (Chang et al., 2004) ou son activité de liaison à l'ADN sont inhibés (Bhende et al., 2005).

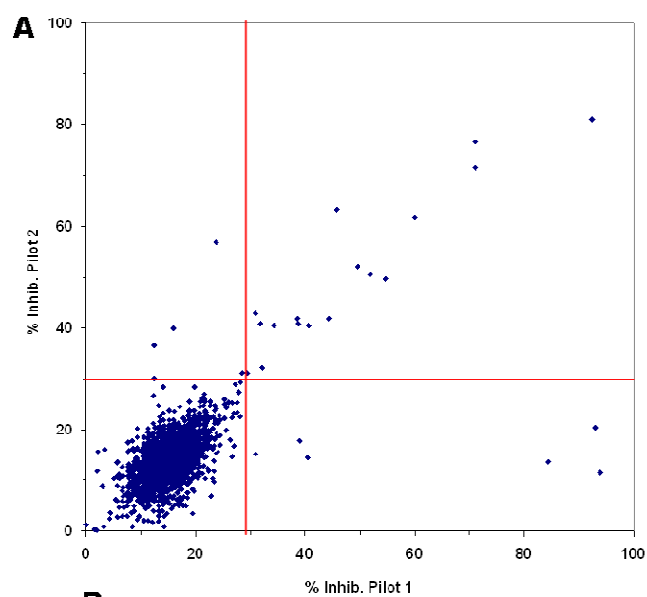
II. Principe de criblage chimique à haut débit

En collaboration avec Joe Lewis de la plate-forme criblage chimique à haut-débit du centre EMBL d'Heidelberg, nous avons développé un essai à haut-débit utilisant la technique de mesure d'anisotropie de fluorescence décrite au chapitre V, afin d'identifier de potentiels inhibiteurs de la fixation à l'ADN de ZEBRA. L'oligonucléotide contenant le site ADN cible AP1 de ZEBRA est marqué à la fluorescéine. Après excitation de cette molécule fluorescente par une lumière polarisée, les valeurs d'anisotropie de fluorescence sont collectées afin d'évaluer la mobilité rotationnelle de l'oligonucléotide : s'il est libre, la valeur d'anisotropie sera très faible ; s'il est complexé à ZEBRA, sa mobilité rotationnelle sera réduite et donc, la valeur d'anisotropie sera plus importante (figure V.1.).

III. Résultats préliminaires

Cet essai est tout d'abord mis au point sur une banque réduite de 3000 composés chimiques. En appliquant la technique d'anisotropie de fluorescence à une recherche haut-débit de composés chimiques, nous avons identifié des inhibiteurs faibles mais spécifiques de ZEBRA (les concentrations correspondant à 50 % d'inhibition (IC50) sont comprises entre 50

et 100 μ M). Les essais sont réalisés en duplicata (« *pilot 1* » et « *pilot 2* ») : le pourcentage d'inhibition est calculé en fonction d'une concentration fixe d'inhibiteur. Les 30 molécules présentant des pourcentages d'inhibition supérieurs à 30 % sont listées dans la table B de la figure VI.1 :



B

EMBL ID	NW	Pilot 1	Pilot 2	mean	SD
100026	500	32.1	32.2	32.2	0.1
100035	470	38.6	41.9	40.3	2.3
100075	503	40.5	14.5	27.5	18.4
100208	410	15.9	40	28.0	17.0
100311	374	30.1	17.0	28.5	16.1
100481	305	40.6	40.5	40.8	0.1
100484	305	38.8	40.9	39.9	1.5
100406	266	44.3	41.9	43.1	1.7
100538	520	31	15.2	23.1	11.2
100682	340	23.8	56.9	40.4	23.4
101002	358	49.6	52.1	50.9	1.8
101164	400	28.5	31.2	29.9	1.9
101283	302	34.3	40.5	37.4	4.4
101316	302	31.9	40.9	36.4	6.4
101308	341	102.6	15.0	59.1	61.5
101488	405	51.9	50.5	51.2	1.0
101531	374	71	71.5	71.3	0.4
101768	435	92.9	20.4	56.7	61.3
101760	512	28.4	31.2	30.3	1.3
101773	388	92.4	81.1	86.8	8.0
101827	404	93.8	11.0	62.7	68.1
102318	290	71	76.8	73.9	4.1
102320	315	84.4	13.7	49.1	50.0
102322	372	45.8	63.4	54.6	12.4
102330	271	12.4	30.2	21.3	12.6
102430	443	54.7	49.8	62.3	3.6
102473	307	31	42.9	37.0	8.4
102640	431	60	61.8	60.9	1.3
102648	486	12.4	36.6	24.5	17.1

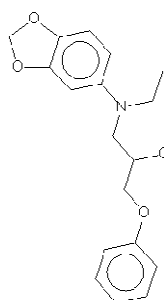
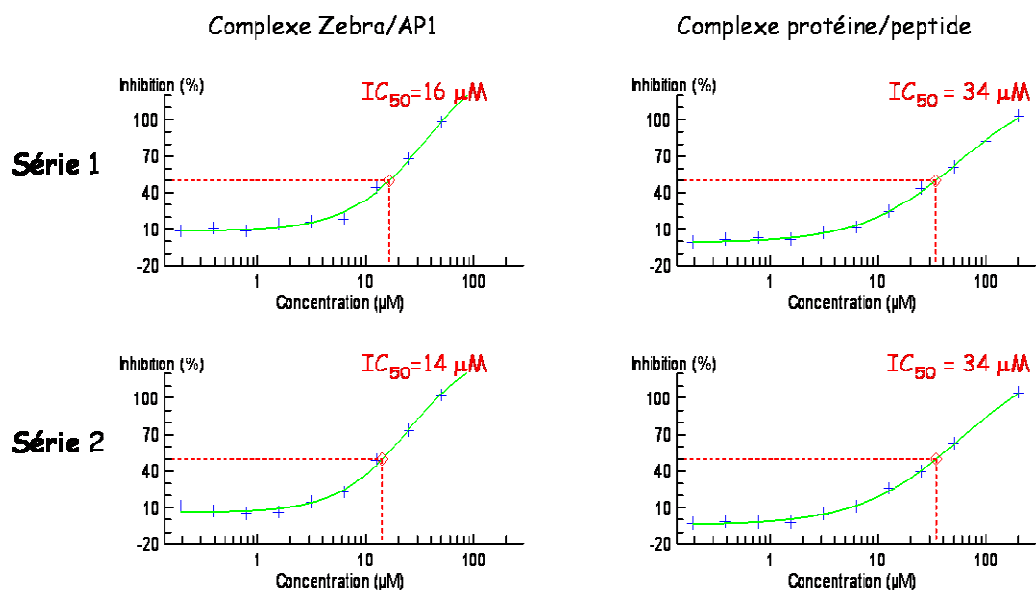
Figure VI.1 : Résultats préliminaires de l'inhibition de liaison à l'ADN de ZEBRA.

(A) Graphique représentant le % d'inhibition issu de l'essai « Pilot 1 » en fonction du % d'inhibition issu de l'essai « Pilot 2 ». Les composés ayant un % d'inhibition inférieur à 30% sont rassemblés dans le carré en bas à gauche. Tous les autres composés présents sur le graphique inhibent plus efficacement et plus spécifiquement ZEBRA : leurs caractéristiques sont présentées dans la table (B). *EMBL ID* : code du composé chimique testé. *MW* : poids moléculaire du composé. *Mean* : moyenne des % d'inhibition issus des 2 essais. *SD* : écart type de la moyenne.

Les capacités d'inhibition de ces 30 molécules ont été étudiées de plus près : leur valeur d'IC₅₀ ont été calculées en testant le % d'inhibition de la fixation de ZEBRA au site AP1 en fonction d'une gamme de concentrations en inhibiteur (figure VI.2 A, B, C). Le pouvoir inhibiteur de ces composés a également été testé sur une protéine non homologue à ZEBRA fixant un peptide marqué à la fluorescéine. Les résultats permettent de définir 3 catégories d'inhibiteurs : 21 des 30 composés de la table VI.1.B. sont en fait des faux-positifs, 2 sont des inhibiteurs non-spécifiques et 1 composé se révèle prometteur. La figure VI.2 présente les déterminations d'IC₅₀ de composés chimiques caractéristiques de chacune de ces catégories.

A

EMBL ID	MW	Pilot 1	Pilot 2	mean	SD
102320	316	84.4	13.7	49.1	50.0

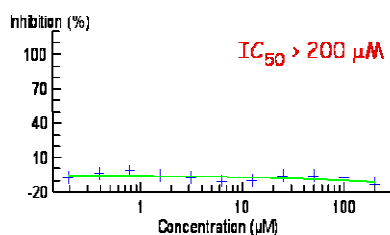


Non spécifique

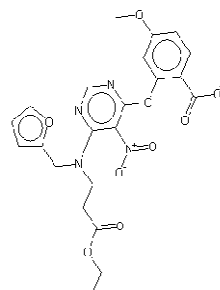
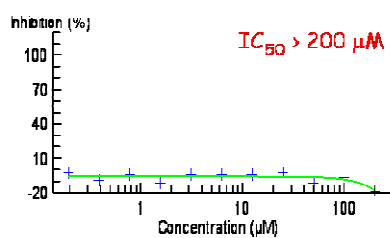
B

EMBL ID	MW	Pilot 1	Pilot 2	mean	SD
100025	600	32.1	32.2	32.2	0.1

Série 1



Série 2

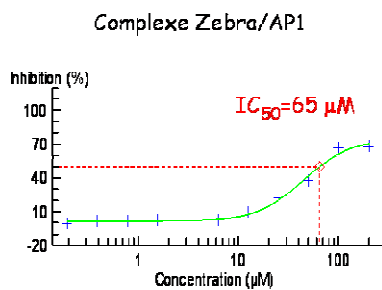


Faux positif

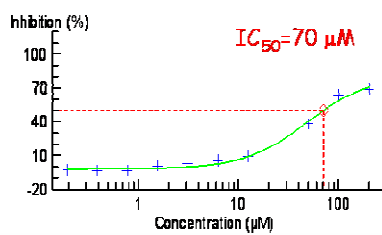
C

EMBL ID	MW	Pilot 1	Pilot 2	mean	SD
100025	478	38.6	41.9	40.3	2.3

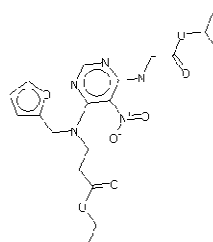
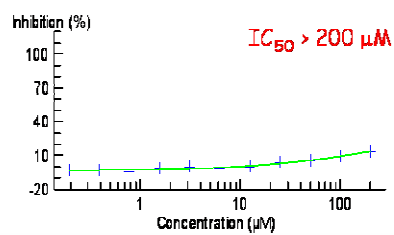
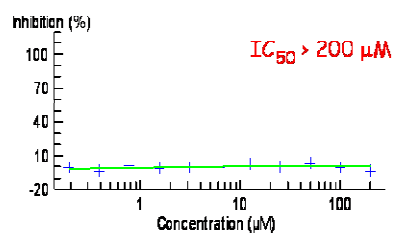
Série 1



Série 2



Complexe protéine/peptide



prometteur

Figure VI.2 : Caractéristiques d'inhibition des 3 composés n°102320 (A), 100025 (B) et 100035 (C). Les graphiques représentent le % d'inhibition en fonction de la concentration en inhibiteur. Les valeurs d'IC50 calculées à partir de ces courbes sont indiquées. Les essais sont réalisés en duplicata (série 1 et 2) sur le complexe ZEBRA/AP1. Un contrôle négatif est réalisé sur une protéine non homologue à ZEBRA en complexe avec un peptide marqué à la fluorescéine. La structure de chacun des 3 composés chimiques est représentée.

Des concentrations croissantes en composé n°102320 entraînent une augmentation du pourcentage d'inhibition de la fixation de ZEBRA sur le site AP1 (figure VI.2.A). Cependant, les valeurs d'IC50 calculées sont identiques concernant le complexe ZEBRA/AP1 et le complexe protéine non homologue/peptide. Ce composé est donc un inhibiteur non spécifique de ZEBRA.

D'après le criblage automatisé réalisé en premier lieu, le composé n°100025 avait un pouvoir inhibiteur moyen de 32 %. Or, les courbes de détermination d'IC50 de ce composé (figure VI.2.B) montrent qu'il n'inhibe en rien la fixation de ZEBRA sur le site AP1 : le pourcentage d'inhibition n'augmente pas avec la concentration en composé chimique. Il s'agit donc d'un faux-positif.

Le composé n°100035 (figure VI.2.C) inhibe la formation du complexe ZEBRA/AP1 de façon concentration-dépendante. Il présente un IC50 moyen de 67,5 μ M. De plus, il n'a aucun effet inhibiteur sur la fixation de la protéine non-homologue à son peptide. Il s'agit donc d'un inhibiteur spécifique de ZEBRA.

IV. Conclusions

Les résultats préliminaires sur cette première banque de 3000 composés chimiques ont permis d'identifier 1 composé prometteur spécifique de ZEBRA (n°100035). Nous testons actuellement des chimiothèques plus importantes afin d'identifier des inhibiteurs tout aussi spécifiques mais plus puissants. Le but à long terme est de faciliter le développement d'agents anti-EBV pour une application thérapeutique.

CHAPITRE VII :

CONCLUSIONS

ET

PERSPECTIVES

Les connaissances actuelles du mécanisme d'induction du cycle lytique de l'EBV par ZEBRA sont insuffisantes à ce jour pour permettre de lutter efficacement contre ce virus. Le facteur de transcription ZEBRA est responsable de la transition entre phase latente et lytique du virus d'Epstein-Barr. Une étape précoce et primordiale de l'induction du cycle lytique est la *trans*-activation par ZEBRA du promoteur viral (Rp) du gène BRLF1 codant pour la protéine Rta. Ainsi, Rta et ZEBRA vont agir en synergie pour déclencher toute une cascade d'induction des gènes lytiques de l'EBV. ZEBRA *trans*-active ses promoteurs cibles en reconnaissant des sites ADN cibles appelés sites ZREs. Le promoteur Rp contient trois sites ZREs dont deux possèdent des îlots CpG qui peuvent donc être méthylés (ZRE2 et ZRE3). ZEBRA est le seul facteur de transcription connu à ce jour qui *trans*-active préférentiellement un site ADN (en l'occurrence, le site ZRE2) dans son état méthylé (Bhende et al., 2004). Cette capacité particulière doit jouer un rôle important dans le mécanisme d'activation du cycle lytique étant donné que le génome de l'EBV est intensivement méthylé durant la phase de latence.

De plus, la structure cristallographique du complexe ZEBRA/AP1 met en évidence une conformation tridimensionnelle tout à fait atypique de la protéine ZEBRA comparé aux autres membres de la famille des protéines bZIPs (Petosa et al., 2006). La modélisation de ZEBRA fixée au site ZRE2 méthylé (meZRE2) a suggéré un mécanisme hypothétique concernant la reconnaissance préférentielle de l'ADN méthylé par ZEBRA. Selon cette hypothèse, la méthylation de la cytosine C^{1'} de l'îlot CpG du site ZRE2 pourrait stabiliser la sérine 186 de ZEBRA dans une conformation lui permettant d'interagir avec l'asparagine 182, ce qui, par conséquent permettrait à cette dernière de former un lien hydrogène avec la guanine G^{2'}, augmentant ainsi l'affinité de ZEBRA pour le site meZRE2.

Afin de vérifier cette hypothèse, nous avons cristallisé le domaine bZIP de ZEBRA en complexe avec le site meZRE2. L'analyse de la structure permet de confirmer une partie de l'hypothèse: le groupement méthyle de la cytosine C^{1'} forme un contact hydrophobe avec la sérine 186, stabilisant ainsi sa chaîne latérale dans une conformation *gauche*⁺. De plus, bien que n'adoptant pas la conformation prédite, l'arginine 182 forme bien un lien hydrogène avec la guanine G^{2'}. Cependant, des différences importantes apparaissent par rapport à la structure prédite. Tout d'abord, la présence de l'îlot CpG du site meZRE2 induit un rétrécissement du sillon majeur d'ADN. Cela permet à la sérine 186 stabilisée en conformation *gauche*⁺ d'être

plus proche de la cytosine C⁻² et ainsi de former une nouvelle liaison hydrogène avec l'atome N4 de cette dernière. L'asparagine 182, quant à elle, a pivoté de 180°. En plus du lien hydrogène qu'elle forme avec la guanine G^{2'}, elle forme une interaction indirecte avec l'adénine A^{3'} via une molécule d'eau. Enfin, la sérine 186 ne forme aucune interaction avec l'asparagine 182, contrairement au mécanisme proposé. Afin de confirmer et de compléter les informations apportées par cette structure, nous avons réalisé des mesures d'affinité de la protéine ZEBRA pour divers sites ADN grâce à la technique d'anisotropie de fluorescence. En réalisant une étude de mutagenèse du site ZRE2 méthylé, nous avons démontré qu'une seule des deux cytosines méthylées (la cytosine C^{1'}) jouait un rôle crucial dans la fixation préférentielle de ZEBRA sur le site meZRE2. Plus précisément, l'importance de ce rôle est due au groupement méthyle de la cytosine C^{1'} puisque la mutation de cette base en thymine ne change en rien l'affinité de ZEBRA pour le site. La paire de bases flanquant l'îlot CpG (position -3/3') du site meZRE2 n'a pas la même importance que dans la reconnaissance du site AP1 par ZEBRA étant donné qu'elle n'engage qu'une liaison indirecte via une molécule d'eau avec l'asparagine 182. Enfin, nous avons confirmé que la sérine 186 et l'asparagine 182 n'interagissaient pas ensemble puisque ces deux résidus reconnaissent le site meZRE2 indépendamment l'un de l'autre.

Après avoir identifié les bases nucléotidiques ayant un rôle crucial dans la reconnaissance de l'ADN méthylé par ZEBRA, nous avons cherché à savoir quels résidus de ZEBRA étaient cruciaux pour cette reconnaissance. Pour cela, nous avons comparé le comportement de ZEBRA au facteur de transcription cJun, membre de la famille des protéines bZIP mais qui ne fait aucune distinction entre les sites meZRE2 et ZRE2. Le remplacement de la région basique de cJun par celle de ZEBRA confère à cJun la capacité de fixer préférentiellement le site meZRE2 avec une affinité 1,5 fois plus importante. La région basique de ZEBRA est donc suffisante pour reconnaître préférentiellement un site ADN méthylé. Cependant, cette préférence est ici moins marquée comparé à ZEBRA qui présentait une affinité 7 fois plus importante pour le site meZRE2. Il semblerait donc que d'autres résidus de la région de dimérisation de ZEBRA soient impliqués dans la reconnaissance préférentielle de l'ADN méthylé. Une étude en 2005 a mis en évidence la nécessité de la sérine 186 dans la fixation préférentielle du site meZRE2 (Bhende et al., 2005). La sérine 186 est unique à ZEBRA : cJun, comme toutes les autres protéines bZIPs connues, possède une alanine à la place (A275). Nous avons introduit cette sérine chez cJun pour vérifier si elle est suffisante pour distinguer un site ADN méthylé d'un site non méthylé ou si la région basique entière de

ZEBRA est nécessaire. La protéine cJun A275S reconnaît préférentiellement le site meZRE2 et la différence d'affinité observée entre les sites méthylé et non méthylé atteint un facteur de 1,5, exactement comme la protéine chimère de cJun possédant la région basique de ZEBRA. Cela prouve que dans toute la région basique de ZEBRA, la sérine 186 seule est suffisante pour distinguer et fixer préférentiellement un site ADN méthylé.

Ces études nous permettent de réviser le mécanisme de reconnaissance préférentielle de l'ADN méthylé par ZEBRA précédemment proposé. Le rétrécissement du sillon majeur du site meZRE2 induit par la présence de l'îlot CpG et la stabilisation de la conformation *gauche*⁺ de la sérine 186 par le groupement méthyle de la cytosine C^{1'}, permet à la sérine 186 de former une nouvelle liaison hydrogène avec la cytosine méthylée C⁻². Ceci explique également le rôle clé joué par la sérine 186 dans le mécanisme de reconnaissance préférentielle de l'ADN méthylé par ZEBRA et mis en évidence par Prasanna Bhende (Bhende et al., 2005).

Ces travaux améliorent considérablement la compréhension du mécanisme d'induction du cycle lytique par ZEBRA. Toutefois, des études complémentaires seront nécessaires pour comprendre pleinement le rôle joué par ZEBRA dans la réactivation du virus d'Epstein-Barr. Tout d'abord, le rôle de la sérine 186 devra être confirmé en testant le comportement d'un mutant de ZEBRA ne possédant pas cette sérine (par exemple, un mutant S186A) sur les sites meZRE2 et ZRE2. De même, les mesures d'affinités vis-à-vis des sites meZRE2 et ZRE2 d'une protéine chimère de cJun possédant des parties de la région de dimérisation et la sérine 186 de ZEBRA permettraient de déterminer si, en plus de la sérine 186, d'autres résidus de ZEBRA sont impliqués dans la reconnaissance préférentielle de l'ADN méthylé. D'autre part, l'équipe d'Alison Sinclair a souligné le rôle de la cystéine 189 de ZEBRA dans la reconnaissance des sites ZREs méthylés (Karlsson et al., 2008). Or, la protéine ZEBRA utilisée lors de nos travaux présente une mutation de cette cystéine en sérine. Il serait donc intéressant de reproduire ces travaux en présence de la cystéine 189. Enfin, nous envisageons de reproduire ces expériences en testant le comportement de ZEBRA et de ces différents mutants en cellules EBV-négatives à l'aide d'un système gène-reporter présentant les sites ZREs du promoteur Rp afin de vérifier la pertinence fonctionnelle de nos résultats *in vivo*.

Parallèlement à ces travaux sur la compréhension du mécanisme d'induction du cycle lytique par ZEBRA, nous avons développé une technique pour effectuer un criblage à haut

débit de composés chimiques pouvant potentiellement inhiber la fixation à l'ADN de ZEBRA. Malgré la taille peu importante de la première chimiothèque testée (composée de 3000 molécules), nous avons obtenu des résultats prometteurs avec l'identification d'un inhibiteur spécifique de ZEBRA. Grâce au développement de cette technique, nous pouvons à présent tester des chimiothèques plus importantes afin d'identifier des inhibiteurs tout aussi spécifiques et plus puissants qui aideront potentiellement au développement de nouveaux agents thérapeutiques contre les pathologies associées à l'EBV.

CHAPITRE VIII :

MATERIELS

ET

METHODES

I. Matériels

1. Clonage

Tous les clonages sont réalisés dans les cellules DH5 α -ETM d'Invitrogen (code n°11319-019 ; génotype : *F⁻ ϕ 80lacZ Δ M15 Δ (lacZYA-argF) U169 recA1 endA1 (rk⁻, mk⁺) gal⁻ phoA supE44 λ^- thi⁻¹ gyrA96 relA1*). Une seule exception concerne les expériences de mutagenèse dirigée pour lesquelles sont utilisées des cellules supercompétentes XL1-Blue de Stratagène (code n° 200236 ; génotype : *recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac [F' proAB lacIqZ Δ M15 Tn10 (Tetr)]*).

2. Expression des protéines

Toutes les protéines sont exprimées en système bactérien *E.coli* grâce aux cellules BL21-CODONPLUS[®] (DE3)-RIPL de Stratagène (code n° 230280 ; génotype : *E. coli B F⁻ ompT hsdS(rB⁻ mB⁻) dcm⁺ Tet^r gal λ (DE3) endA Hte [argU proL Camr] [argU ileY leuW Strep/Spec^r]*). Ces cellules sont supplémentées en gène codant pour des ARN de transfert rares chez *E.coli* (*argU*, *ileY*, *leuW*, *proL*) ce qui permet d'améliorer grandement l'expression de séquences provenant d'autres organismes.

II. Méthodes

1. Construction des plasmides d'expression en bactéries

1.1 Mutagenèse dirigée : clones ZEBRA C189S¹⁷⁵⁻²³⁶ – pET28a

cJun A275S/C278S²⁶⁴⁻³²⁴ – pETM40

Les mutations ponctuelles ont été réalisées grâce au kit de mutagenèse dirigée « Quick Change mutagenesis kit » (Stratagène, code n° 200523). Le milieu réactionnel de 25 μ l contient les réactifs suivants :

- 30 ng de matrice ADN,
- 10 pmoles de chaque amorce nucléotidique,
- 10 mM du mélange des quatre dNTPs,

- 2,5 µl de tampon de réaction 10X,
- 2,5 unités de polymérase Pfu Ultra II (*Stratagène*)

Les séquences des amorces sont fournies en Annexe (Annexe B.1.).

La mutagenèse se déroule dans une machine à PCR (thermocycleur) selon le programme suivant, répété n fois (n=16 pour une modification d'un acide aminé) :

- 95°C pendant 45 secondes,
- 55°C pendant 1 minute,
- 68°C pendant (1 minute x nb de kb du plasmide) minutes.

L'ADN parental non muté a été synthétisé chez *E.coli*. Donc, contrairement à l'ADN synthétisé *in vitro* pendant le programme de mutagenèse dirigée, il est méthylé. Afin de sélectionner l'ADN muté, les produits de la réaction sont digérés pendant 1 heure à 37°C par 10 unités de l'enzyme de restriction DpnI (*Stratagène*, code n° 500402) qui clive spécifiquement les sites méthylés).

Après transformation de 50 µl de cellules supercompétentes XL1-Blue (*Stratagène*, code n° 200236) par 5 µl de réaction, les clones obtenus sont tout d'abord sélectionnés par digestion avec les enzymes de restriction correspondant au site de clonage du gène d'intérêt. Les clones positifs à la présence du gène d'intérêt sont alors séquencés afin de vérifier la présence de la mutation.

- clone ZEBRA C189S¹⁷⁵⁻²³⁶ – pET28a

Le domaine bZIP de ZEBRA (résidus 175-236) avait été cloné dans le vecteur pET28a (*Novagen*, voir Annexe D) par Patrice Morand (Morand P. *et al.*, 2006). Dans cette construction, le domaine de liaison à l'ADN de ZEBRA comporte deux mutations : la sérine 186 a été mutée en alanine et la cystéine 189 en sérine. D'après la structure cristallographique du bZIP de ZEBRA en complexe avec le site AP-1 (Petosa C. *et al.*, 2006), la cystéine 189 forme des contacts avec l'ADN. Afin d'éviter toute modification de ces contacts protéine-ADN par une oxydation éventuelle de cette cystéine 189 lors de la cristallogénèse, la mutation C189S sera conservée. Cependant, nous avons restitué la sérine 186 par mutagenèse dirigée.

Le nom de la protéine correspondant à ce clone a été simplifié en « ZEBRA ». Les amorces utilisées sont résumées en Annexe B.1.

- clone cJun A275S/C278S²⁶⁴⁻³²⁴ – pETM40

Nous disposions du clone cJun C278S²⁶⁴⁻³²⁴ – pETM40 où seule la cystéine 278 avait été mutée en sérine (pour les mêmes raisons que le clone ZEBRA C189S175-236 – pET28a, voir plus haut). Pour obtenir le clone cJun A275S/C278S²⁶⁴⁻³²⁴ – pETM40, l'alanine 275 du bZIP de cJun a été mutée en sérine par mutagenèse dirigée afin de se rapprocher de la séquence protéique de ZEBRA. En effet, ZEBRA est le seul bZIP possédant une sérine en position 186 alors que tous les autres bZIPs présentent une alanine à cette position (alanine 275 concernant cJun). Or, cette sérine 186 de ZEBRA semble jouer un rôle primordial dans la reconnaissance préférentielle de l'ADN méthylé. Nous avons donc construit un clone de cJun possédant cette sérine particulière : cJun A275S/C278S²⁶⁴⁻³²⁴ – pETM40. Le nom de la protéine correspondant à ce clone a été simplifié en « cJun A275S ».

1.2. Clones ZEBRA C189S¹⁷⁵⁻²³⁶ – pETM40 et cJun C278S²⁶⁴⁻³²⁴ – pETM40

Nous disposions des clones ZEBRA C189S¹⁷⁵⁻²³⁶ – pET28a et cJun C278S²⁶⁴⁻³²⁴ – pET11. Nous les avons sous-cloné dans le vecteur d'expression pETM40 afin d'exprimer les protéines ZEBRA et cJun en fusion avec une étiquette MBP (l'étiquette MBP étant positionné en N-ter de ces protéines). Les inserts ADN cJun C278S²⁶⁴⁻³²⁴ et ZEBRA C189S¹⁷⁵⁻²³⁶ ont été amplifiés par la méthode de réaction de polymérisation en chaîne (PCR) (les amorces utilisées pour la PCR sont résumées en Annexe B.3.). Les inserts et le vecteur pETM40 sont digérés par les enzymes de restriction NcoI et KpnI (*NEB*), puis purifiés grâce au « PCR purification kit » de *QIAGEN* (code n° 28106). Les ligations « insert-vecteur pETM40 » (ratio vecteur : insert de 1 : 5) sont réalisées par la T4 DNA ligase de *NEB* (code n° m02021) à 30°C pendant 1h, suivies d'une inactivation de la ligase à 65°C pendant 20 minutes. Les produits de ligation sont utilisés pour transformer des cellules *E.coli* DH5 α -ETM d'*Invitrogen* (code n°11319-019) selon le protocole suivant : les bactéries sont mises sur glace pendant 30 minutes en présence d'ADN puis soumises à un choc thermique de 45 secondes à 42°C. Elles sont ensuite mises sur glace 2 minutes puis resuspendues dans 150 μ l de LB. Après une incubation d'1h à 37°C, les cellules sont étalées sur milieu LB-agar contenant l'antibiotique d'intérêt (kanamycine 60 μ g/ml), puis incubées de 12 à 16 h à 37°C. Les colonies sont alors testées par « colonie-

PCR » puis par digestion avec les enzymes de restriction correspondant au site de clonage du gène d'intérêt. Les clones positifs à la présence du gène d'intérêt sont alors séquencés afin de déceler d'éventuelles mutations non désirées.

1.3. Clone Chimère ZEBRA-cJun – pETM40

Ce clone contient la région 175-196 du bZIP de ZEBRA fusionné à la région 287-324 du bZIP de cJun. Il a été construit par Cécile Hamès (collaboration CEA) selon la technique des mégaprimers (Kirsch R.D. and Joly E., 1998). Brièvement, cette technique comporte deux PCR : la première a pour but d'amplifier un fragment d'ADN (en l'occurrence la région 175-196 du bZIP de ZEBRA) à partir du plasmide ZEBRA C189S¹⁷⁵⁻²³⁶ – pETM40 : ce fragment d'ADN double brin jouera le rôle de mégaprimers. La deuxième PCR utilise ces mégaprimers pour insérer la région 175-196 du bZIP de ZEBRA en amont de la région 287-324 du bZIP de cJun (la région 264-286 du bZIP de cJun étant éliminé lors de cette PCR). L'ADN polymérase fidèle Phusion (*Finnzymes*) est utilisée pour la première étape et l'ADN polymérase fidèle Ex Taq (*Takara*) pour la deuxième.

2. Expression et purification des protéines

2.1. Expression des protéines chez *E.coli*

L'expression des protéines se fait dans les bactéries compétentes *E.coli* BL21-CODONPLUS[®] (DE3)-RIPL de *Stratagène* (code n° 230280). Ces bactéries sont transformées par choc thermique comme il a été décrit précédemment (voir partie 1.2 de ce chapitre). Une pré-culture est préparée en inoculant 25 ml de milieu LB contenant 60 µg/ml de kanamycine par les colonies transformées ayant pousser sur milieu LB-agar. La pré-culture est incubée à 37°C pendant 16 heures sous agitation à 200 rpm. 10 ml de cette pré-culture sont utilisés pour inoculer 1 l de milieu LB (contenant 60 µg/ml de kanamycine). La croissance de cette culture est réalisée dans les mêmes conditions que la pré-culture et maintenue jusqu'à ce que la densité optique à 600 nm (DO₆₀₀) du milieu ait atteint 0,7. L'expression des protéines est induite par l'ajout d'IPTG à une concentration de 0,5 mM et la culture est maintenue sous agitation à 37°C pendant 5 h.

Les bactéries sont centrifugées à 8000 rpm pendant 10 minutes à 4°C. Le culot bactérien est suspendu dans le tampon de lyse (voir Annexe C.2.1.) à raison de 20 ml de tampon de lyse par litre de culture. Les bactéries sont lysées par sonication en réalisant 6 cycles de 30 sec, chacun entrecoupés de 45 sec de pause. Le lysat est centrifugé pendant 20 minutes à 20,000 rpm à 4°C. Le surnageant est conservé avant de poursuivre à l'étape de purification.

2.2 Protocole de purification de la protéine ZEBRA

Les différentes étapes de cette purification sont résumées dans le schéma suivant (figure VIII.1.) :

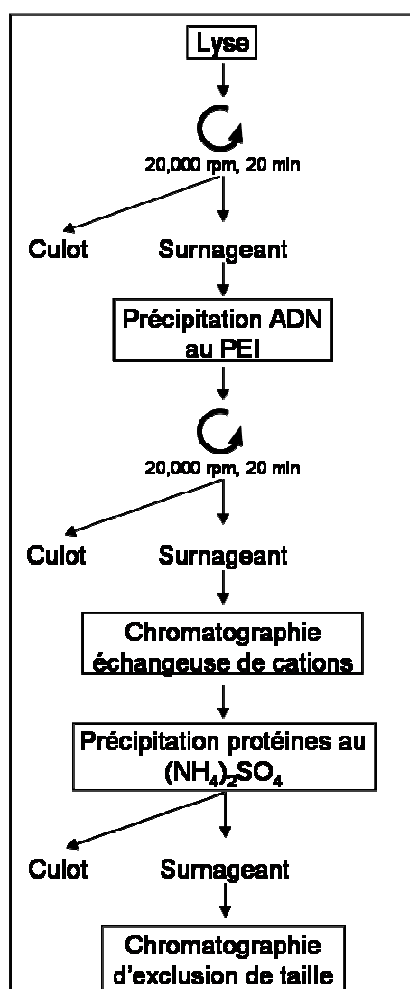


Figure VIII.1. : Schéma résumant les différentes étapes de purification de la protéine ZEBRA.

1^{ère} étape de purification : précipitation de l'ADN contaminant au PEI

Le PEI (poly-éthylène-imine) est un polycation qui forme un énorme polymère en réseau qui emprisonne l'ADN. Il est utilisé pour éliminer la majeure partie de l'ADN contaminant provenant de la lyse des bactéries. Le PEI est donc ajouté au surnageant de lyse à une concentration finale de 0,3 % (v/v). Ce mélange est incubé 30 min à 4°C sous agitation douce avant d'être centrifugé 20 min à 20,000 rpm. Le surnageant est conservé à 4°C avant de procéder à l'étape de purification suivante.

2^{ème} étape de purification : chromatographie échangeuse de cations

La résine échangeuse de cations utilisée est la résine SP Sepharose Fast flow (*GE Healthcare*, code n° 17-0729-10). Ce type de résine est utilisée pour la séparation de macromolécules ionisées (en l'occurrence le domaine bZIP de ZEBRA globalement chargé positivement du à sa région basique de liaison à l'ADN). La matrice polyosidique (composée de cellulose ou de dextrans) porte des groupements chargés sulfopropyles (SP) jouant le rôle de contre-ion échangeable.

La résine est préalablement équilibrée par 5 volumes de tampon A (voir Annexe C.2.2.). Cette chromatographie est réalisée sur un système « *Biologic Low Pressure* » de *Bio-Rad* (code n° 358BR2423). Les protéines sont éluées par un gradient NaCl 0,01-1M à l'aide des tampons A et B (voir Annexe C.2.2.). Les fractions contenant la protéine sur-exprimée sont analysées sur gel SDS-PAGE puis rassemblées et conservées à 4°C avant de procéder à l'étape de purification suivante.

Gel SDS-PAGE Tris-Tricine :

Gel de concentration

- 5% acrylamide/bisacrylamide 37,5 :1 (*Merck*)
- 0,7 M Tris-HCl pH 8,45
- 0,1% SDS
- 0,1% APS
- 0,01% TEMED

Gel de séparation

- 12 % acrylamide/bisacrylamide 37,5 :1 (*Merck*)
- 1 M Tris-HCl pH 8,45
- 0,1% SDS
- 10 % glycérol
- 0,1% persulfate d'ammonium (APS)
- 0,01% TEMED

Les tampons de migration utilisés sont présentés en Annexe C1.3.

3^{ème} étape de purification : précipitations des protéines au sulfate d'ammonium

Cette technique utilise la solubilité différentielle des protéines en fonction de la force ionique de la solution qui les contient. Le sulfate d'ammonium est ajouté à une concentration finale de 30 % (w/v) aux fractions contenant la protéine d'intérêt. Le mélange est mis sur glace sous agitation douce pendant 30 min, puis centrifugé 20 min à 20,000 rpm. Le culot, bien séché, est ensuite conservé à 4°C.

4^{ème} étape de purification : chromatographie d'exclusion de taille

Le culot de sulfate d'ammonium est repris par 1 ml de tampon GF (voir Annexe C.2.3.) et centrifugé 5 min à 10,000 rpm. Le surnageant est alors injecté sur une colonne Superdex S75 HR 10/30 (*Pharmacia/GE Healthcare*, code n° 71-6031-00) montée sur un système « *Biologic FPLC* » de *Bio-Rad*. Le tampon GF utilisé est présenté en Annexe C.2.3. Les fractions contenant la protéine purifiée sont analysées sur gel SDS-PAGE puis rassemblées et concentrées sur système de concentration Amicon (*Millipore*, code n° UFC9 010 08). La concentration de la protéine purifiée est ensuite dosée en triplica par absorbance à 280nm (A_{280}). L'absorbance à 260 nm (A_{260}) est également mesurée afin de contrôler la quantité d'Adn contaminant encore présent dans l'échantillon ; le rapport A_{280}/A_{260} devant être supérieur ou égal à 1,8. La protéine peut ensuite être conservée à 4°C pour une courte période ou congelée à -20°C après ajout de 50 % de glycérol.

2.3. Protocole de purification des protéines de fusion-MBP (clones ZEBRA C189S¹⁷⁵⁻²³⁶ – pETM40, cJun C278S²⁶⁴⁻³²⁴ – pETM40, cJun A275S/C278S²⁶⁴⁻³²⁴ – pETM40, Chimère ZEBRA-cJun – pETM40)

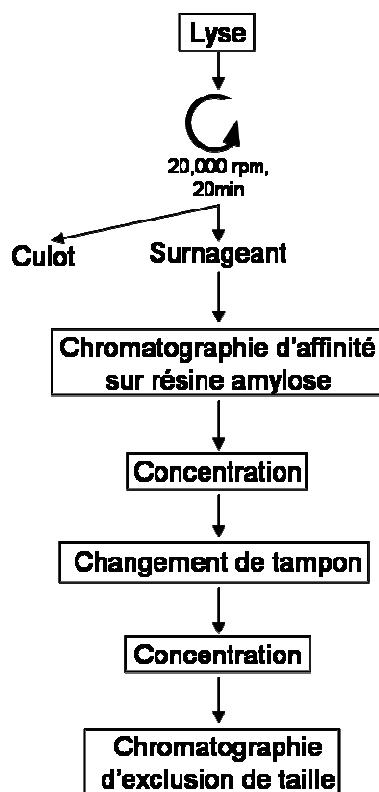


Figure VIII.2. : Schéma résumant les différentes étapes de purification des protéines de fusion MBP.

1^{ère} étape de purification : chromatographie d'affinité sur résine amylose

Cette étape de purification par chromatographie d'affinité utilise une résine d'amylose (NEB, code n° E8021L) sur laquelle vont se fixer les protéines de fusion-MBP. La résine d'amylose est préalablement équilibrée par 3 x 1 volume de tampon de lyse (voir Annexe C.2.1.) puis incubée avec le surnageant de lyse sous agitation douce à 4°C pendant 1 h. L'ensemble est ensuite coulé dans une colonne où la résine va se tasser sous effet de la gravité. Après élimination du « *flow-through* », la résine est lavée par 5 volumes de tampon de lavage 1 M puis 5 volumes de tampon de lavage 1,5 M (Annexe C.2.4.). Enfin, les

protéines de fusion-MBP sont éluées par 1 volume de tampon d'élution (Annexe C.2.4.) dans des fractions de 1 ml. Ces fractions sont ensuite analysées sur gel SDS-PAGE et rassemblées.

L'étape suivante consiste à échanger le tampon dans lequel a été éluée la protéine de fusion-MBP pour un tampon PBS (Annexe C.2.5.) qui sera le tampon de la chromatographie d'exclusion de taille (étape suivante). Pour cela, les fractions éluées de la résine d'amylose sont déposées sur une colonne Econo-Pac 10 DG de *Bio-Rad* (code n° 732-2010). Chaque chargement sur la colonne est de 3 ml et chaque élution est réalisée par 4 ml de PBS. Les fractions éluées sont rassemblées puis concentrées sur système de concentration Amicon (*Millipore*, code n° UFC9 010 08). La protéine est conservée à 4°C jusqu'à l'étape suivante de purification.

2^{ème} étape de purification : chromatographie d'exclusion de taille

Le protocole est décrit dans le paragraphe 2.2.4. Les protéines de fusion-MBP purifiées ont un poids moléculaire avoisinant les 50 kDa (l'étiquette MBP ayant une taille importante) : nous utilisons donc une colonne Superdex 200 10/300 GL (*Pharmacia/GE Healthcare*, code n° 17-5175-01) montée sur un système « Biologic FPLC » de *Bio-Rad*. Le tampon GF2 utilisé est présenté en Annexe C.2.5. Les fractions contenant la protéine purifiée sont analysées sur gel SDS-PAGE puis rassemblées et concentrées. La concentration de la protéine purifiée est ensuite dosée en triplica par absorbance à 280 nm et 260 nm (A_{280}/A_{260}). La protéine est congelée à -20°C après ajout de 50 % de glycérol et de 2,5 mM de β -mercaptoéthanol.

3. Purification et hybridation des oligonucléotides simple brin

Les oligonucléotides sont synthétisés sous forme simple brin. Les oligonucléotides utilisés en cristallogénèse proviennent de BioSpring (le site meZRE2 étant méthylé durant la synthèse). Ceux utilisés dans les expériences d'anisotropie de fluorescence sont synthétisés par MWG : un seul des deux brins complémentaires de chaque paire d'oligonucléotide est marqué à la rhodamine en 5' (pic d'excitation à 546 nm, pic d'émission à 576 nm). Contrairement à ceux utilisés en cristallogénèse, les oligonucléotides d'anisotropie de

fluorescence ne seront pas re-purifiés : leur purification par HPLC opérée par le fournisseur suffit.

Purification des oligonucléotides simple brin utilisés en cristallogénèse

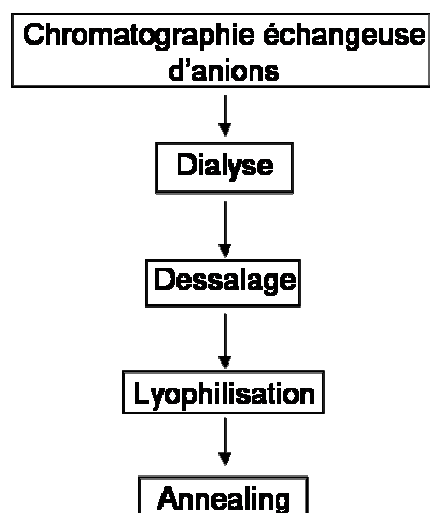


Figure VIII.3. : Schéma résumant les étapes successives de purification des oligonucléotides utilisés en cristallogénèse.

Chaque oligonucléotide est reçu en poudre. Après resuspension dans 1 ml d'un mélange 90 % tampon DNA/A-10% tampon DNA/B (voir Annexe C.3.), chacun est injecté sur une colonne MonoQ HR 10/10 (*Pharmacia/GE Healthcare*, code n°17-0506-01) préalablement équilibrée par le même mélange. L'élution est réalisée par un gradient 0-1 M NaCl à l'aide des tampons DNA/A et DNA/B (voir Annexe C.3.).

Les fractions correspondant au pic d'élution sont rassemblées puis dialysées contre 4 l d'H₂O déminéralisée pendant 14h à 4°C.

Elles sont ensuite déposées sur une colonne Econo-Pac 10 DG (selon le protocole décrit en partie 2.3.1. de ce chapitre). Les oligonucléotides simple brin sont élués par de l'H₂O déminéralisée, congelés à l'azote et lyophilisés jusqu'à séchage complet. Ils sont enfin repris dans de l'H₂O déminéralisée à une concentration finale de 20 mg/ml.

Hybridation des oligonucléotides simple brin de cristallographie/ des oligonucléotides simple brin fluorescents

Chaque paire d'oligonucléotides complémentaires est hybridée à un ratio 1 :1 dans un tampon d'hybridation (tampon d'hybridation-X pour les oligonucléotides de cristallogénèse et tampon d'hybridation-AF pour les oligonucléotides fluorescents) par passage à 95°C pendant 2 à 3 min et refroidissement progressif jusqu'à température ambiante.

4. Cristallogénèse

4.1. Principe de la cristallogénèse

Un complexe protéique cristallise lorsqu'il se trouve en phase de sursaturation. D'après le diagramme de phases suivant, cette phase se situe entre la zone de solubilité et la zone de précipitation (figure VIII.4.) :

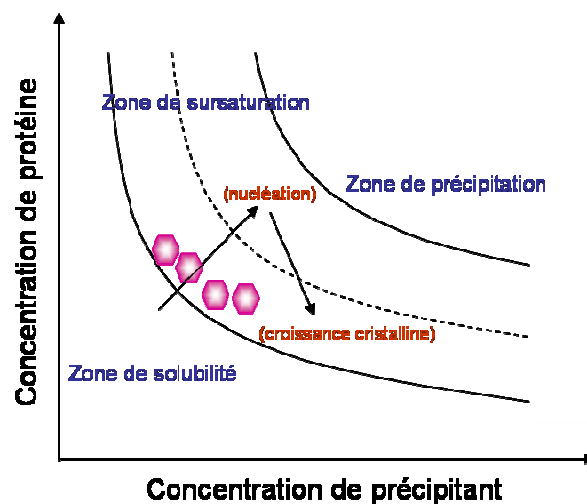


Figure VIII.4. : Diagramme de phases décrivant la cristallisation d'un complexe protéique par diffusion de vapeur.

La solubilité d'un complexe protéique dépend de nombreux paramètres comme sa concentration, le pH, la température, la présence d'agents précipitants (notamment le PEG

et/ou le sels chaotropiques à haute concentration) ou solubilisants (sels à base concentration par exemple). L'aspect cinétique est lui aussi important. Le travail de cristallogénèse consiste à modifier ces différents paramètres afin que le complexe protéique cristallise.

La sursaturation, rapport entre la concentration initiale en complexe protéique et la concentration à l'équilibre, est la force motrice des processus de nucléation et de croissance cristalline.

La première étape de la cristallisation consiste à rendre les complexes protéiques insolubles suite à l'ajout d'agents précipitants afin d'amener les complexes dans la partie métastable de la zone de sursaturation où les précipités cristallins se forment. Suite à cela la sursaturation des complexes dans le reste de la goutte de cristallisation diminue, ce qui permet de repasser dans le domaine de sursaturation propice à la croissance cristalline. Cette dernière se poursuit alors autour des précipités cristallins formés précédemment et la taille des cristaux augmente, jusqu'à ce qu'il n'y ait plus de complexe protéique disponible dans la goutte ou que la croissance cristalline soit interrompue suite à de trop nombreux défauts de cristallisation (irrégularités dans la maille, addition de molécules légèrement différentes).

4.2. Mise en œuvre

Le complexe ZEBRA/ADN a été cristallisé selon la technique de diffusion de vapeur en goutte suspendue (McPherson A. *et al.*, 1995). Elle consiste en la diffusion de vapeur entre une goutte de solution protéique et la solution de cristallisation contenue dans le puits jusqu'à ce que les potentiels chimiques de la goutte et de la solution de cristallisation s'équilibrent. Pour cela, une goutte d'1 μ l de complexe protéine/ADN est déposée sur une lamelle de verre siliconée et 1 μ l de solution de cristallisation lui est ajouté. La lamelle de verre vient fermer un puits dans lequel 500 μ l de solution de réservoir ont été déposés, l'étanchéité étant assurée par de la graisse de silicone déposée entre la lamelle et le dessus du puits (figure VIII.5.).

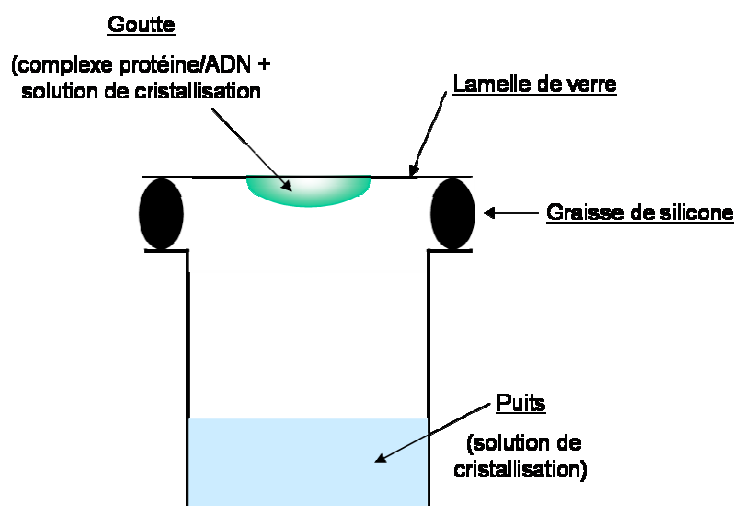


Figure VIII.5. : Mise en œuvre de la technique de cristallisation par diffusion de vapeur en goutte suspendue.

Pendant que les potentiels chimiques s'équilibrent, le volume de la goutte suspendue diminue. Il s'en suit une augmentation de la concentration en complexe protéique et en sels dans la goutte, et selon le diagramme de phases, le complexe protéique est alors dans le domaine de sursaturation ou de précipitation. C'est là que les germes cristallins apparaissent. Puis le processus de croissance cristalline se poursuit, comme décrit précédemment (figure VIII.4.).

4.3. Cristallogenèse du complexe ZEBRA/meZRE2

Le complexe protéine/ADN est formé à un ratio 1 : 1 à partir d'une solution de ZEBRA à 6,7 mg/ml et du duplexe oligonucléotides meZRE2 à 20 mg/ml. Le complexe est alors incubé pendant 30 min sur glace puis centrifugé 5 min à 10,000 rpm. En effet, l'ajout de l'ADN à la solution protéique entraîne une précipitation d'une partie du complexe ZEBRA/meZRE2. Après centrifugation, seul le surnageant est utilisé pour réaliser les gouttes sur lamelles. Pour cela, 1 μ l du complexe ZEBRA/meZRE2 est déposé sur une lamelle de verre comme décrit précédemment, et mélangé à 1 μ l de la solution de cristallisation suivante : 12 % PEG 4000, 15 % PEG 400, 20 % isopropanol et 50 mM de tampon acétate de

sodium/acide acétique pH 3,8. Les cristaux en forme d'aiguilles apparaissent 72 h après et la croissance cristalline semble stagner au bout d'une semaine.

5. Résolution de la structure du complexe ZEBRA/meZRE2 par cristallographie aux rayons X

5.1. Les rayons X et les cristaux de protéines

Pour sonder la structure de la matière, il est nécessaire d'utiliser un rayonnement de longueur d'onde de l'ordre de grandeur des dimensions caractéristiques de l'objet étudié. Or, les distances interatomiques d'une molécule sont de l'ordre de l'Angström (Å, 10^{-10} m). Les rayons X, ondes électromagnétiques de longueur d'onde de l'ordre de l'Angström, sont donc propices à l'étude des protéines par cristallographie.

Cependant, l'intensité de l'onde diffractée par une seule molécule serait trop faible et ne pourrait pas être mesurée. Afin d'augmenter ce signal, celui-ci est mesuré en faisant diffracter le faisceau de rayons X par un cristal de protéines, de sorte que ces dernières soient ordonnées, que les ondes diffractées soient donc cohérentes et que la somme de leurs intensités soit alors observable.

5.2. Acquisition et traitement des données

Source de rayons X

Préalablement à leur congélation, les cristaux ont été cryo-protégés par ajout de 25 % de glycérol afin d'éviter leur destruction par la glace qui se formerait. Les données de diffraction ont été collectées à l'ESRF (Grenoble) sur la ligne micro-focus ID23-eh2 à 100 K à une longueur d'onde de 0,873 Å. Le complexe ZEBRA/meZRE2 appartenant au groupe d'espace C2, toutes les tâches de diffraction sont collectées après rotation du cristal de 180°. L'angle d'oscillation choisi est de 1° et l'exposition est de 4 sec/image.

Traitement des données

Le groupe d'espace des cristaux a été déduit de l'analyse des clichés de diffraction par le programme XDS (Kabsch, 1988 ?????).

Indexation des données : les données sont indexées image par image en affinant un certain nombre de paramètres comme la distance cristal-détecteur, la position du centre du détecteur, l'orientation du cristal, les paramètres de maille, etc. Les profils de diffraction prédits sont ajustés aux taches observées. Chaque tache est alors indexée en h, k et l.

Intégration et correction des intensités : dans le même temps, l'intensité de chaque tache de diffraction est mesurée en appliquant un masque autour de la tache. La couronne interne délimite la surface à intégrer et la couronne externe permet d'estimer le bruit de fond à retrancher de l'intensité intégrée. Chaque tache est donc corrigée du bruit de fond et l'ajustement des masques pour les taches doit être optimal afin de maximiser le rapport signal/bruit.

Mise à l'échelle et réduction du jeu de données : l'ensemble des images intégrées sont mises à l'échelle par le programme XSCALE afin de constituer un jeu de données complet. Les données mises à l'échelle sont finalement réduites à l'unité asymétrique et les intensités « I » sont transformées en amplitudes « F » avec le programme TRUNCATE.

Analyses cristallographiques préliminaires

Analyse du contenu de l'unité asymétrique : détermination du nombre de molécules dans l'unité asymétrique : l'analyse du contenu en solvant du cristal repose sur les travaux de Matthews B.W. (Matthews B.W., 1968). Le nombre de Matthews est défini par :

$$V_M = \frac{V}{M \times n}$$

où V est le volume de la maille, M la masse moléculaire du complexe protéine/ADN et n le nombre de molécules dans la maille.

Pour des cristaux ayant de 30 à 75% de solvant, V_M varie de 1,66 à 4,0. Même s'il existe des exceptions, la valeur moyenne de V_M est de 2,4. Cette valeur permet d'estimer le pourcentage de solvant dans les cristaux en supposant que la densité de la protéine a une valeur standard.

L'unité asymétrique de la maille est le plus petit volume où les symétries cristallographiques du groupe d'espace ne s'appliquent plus. Autrement dit, l'ensemble de la maille est engendré par l'unité asymétrique et l'ensemble des symétries de cette dernière par les opérations de symétrie cristallographique. L'unité asymétrique peut contenir plusieurs molécules qui ne sont alors pas reliées par des opérations de symétrie cristallographique. Les relations au sein de l'unité asymétrique sont donc non-cristallographiques. Le calcul du V_M peut donner des indications sur le nombre de molécules dans la maille et donc dans l'unité asymétrique si l'on connaît le groupe d'espace du cristal.

Symétries non cristallographiques : si le cristal contient plusieurs molécules dans l'unité asymétrique, celles-ci peuvent être reliées par des symétries non cristallographiques. La recherche des symétries non cristallographiques purement translationnelles se fait en établissant une carte de Patterson. La fonction de Patterson se calcule comme une synthèse de Fourier avec les intensités diffractées. Les cartes de Patterson contiennent alors des pics correspondant à tous les vecteurs interatomiques.

Remplacement moléculaire

Principe : la connaissance des structures déjà résolues peut être utile à la résolution de structures inconnues présentant une certaine homologie, évitant ainsi un phasage expérimental souvent assez long. En effet, à partir d'une structure tridimensionnelle connue, un diagramme complet de diffraction incluant amplitudes et phases peut être généré. La connaissance d'une structure homologue à celle à déterminer peut donc fournir des phases initiales, pourvu que les diagrammes de diffraction soient comparables, c'est à dire qu'ils soient ramenés à la même maille et au même groupe d'espace. Il faut donc que la structure modèle et la structure à déterminer occupent la même position dans la maille cristalline. Le remplacement moléculaire consiste à placer la « molécule modèle » dans la maille. Le modèle est d'abord orienté dans la maille par une fonction de rotation puis positionné dans le cristal par une fonction de translation (Lattman E., 1985).

Le remplacement moléculaire est réalisé par le programme PHASER à l'aide d'un modèle réduit de la structure ZEBRA/AP1 comportant les résidus 175 à 200 du bZIP de ZEBRA ainsi que les 12 bp centrales du duplexe ADN de 19 bp.

Construction du modèle et affinement cristallographique

Le modèle a été construit à la main dans la carte de densité électronique expérimentale à l'aide du programme « O ». Il a été affiné par le programme CNS selon la stratégie décrite plus bas.

Principes de l'affinement cristallographique : le modèle construit dans la carte de densité électronique, même s'il est proche de la structure réelle, ne rend pas bien compte des données expérimentales. L'affinement cristallographique vise donc à réduire l'écart entre les facteurs de structure observés et calculés en affinant les paramètres du modèle. Le facteur de structure F_{calc} est calculé par transformée de Fourier à partir du modèle atomique. Une grandeur indicatrice de l'état de l'affinement est le facteur R cristallographique :

$$R_{\text{cryst}} = \frac{\sum_{hkl} ||F_{\text{obs}}| - k|F_{\text{calc}}||}{\sum_{hkl} |F_{\text{obs}}|}$$

Du fait du grand nombre de paramètres à affiner, l'affinement est une étape qui est limitée par le rapport du nombre d'observations sur le nombre de paramètres. L'affinement n'est possible, et ne converge de manière satisfaisante, que si le nombre d'observations est supérieur au nombre de paramètres. Les paramètres sont en général au nombre de 4 par atome (x,y,z, B). En règle générale, ces 4 paramètres peuvent être affinés si les cristaux diffractent au moins jusqu'à 2,5Å.

Afin de contourner le problème du manque d'observations par rapport au nombre de paramètres, une solution est soit d'augmenter le nombre d'observations, soit de diminuer le nombre de paramètres.

Il faut également avoir des critères objectifs pour suivre l'affinement. Le facteur R n'est pas très objectif puisqu'il fait partie de la fonction à minimiser et que la modification des paramètres durant l'affinement va directement influencer sa valeur. En revanche, l'utilisation du facteur R_{free} , calculé sur un jeu de réflexion exclues de l'affinement, permet d'avoir une estimation objective du protocole d'affinement (Brünger A.T., 1992). La diminution concomitante du R et du R_{free} valide le protocole d'affinement. D'autre part, la stéréochimie de la macromolécule sera également un critère validant la qualité de l'affinement.

Stratégie d'affinement : Les 2 molécules sont tout d'abord placées dans l'unité asymétrique par le programme PHASER. Ces deux molécules sont affinées par CNS selon deux corps rigides indépendants puis l'ensemble est affiné selon 4 corps rigides indépendants (les 2 doubles brins d'ADN et les 2 dimères de ZEBRA). Une série de cycles de minimisation visant à affiner les coordonnées des atomes ainsi que les facteurs B est alors effectuée. Les cartes 2Fo-Fc et Fo-Fc sont examinées et les résidus sont ajoutés manuellement dans la densité électronique. D'autres cycles d'affinement sont régulièrement lancés avec le programme CNS jusqu'à construction complète du modèle. Puis les molécules d'eau sont ajoutées et une étape de recuit simulé est effectuée avant de lancer une dernière fois un cycle d'affinement.

Validation des structures

Le programme PROCHECK a été utilisé pour vérifier la géométrie du modèle et le valider. Les diagrammes de Ramachandran (Ramachandran G.N. and Sasisekharan V., 1968) ont également été examinés : ils donnent la répartition des angles dièdres ϕ et ψ . Les couples (ϕ, ψ) sont caractéristiques d'une structure secondaire du squelette C_α des protéines. La répartition de ces angles donne donc un critère supplémentaire de qualité du modèle affiné.

6. Mesures d'affinité de l'interaction ZEBRA/ADN par anisotropie de fluorescence (AF)

6.1. La technique d'anisotropie de fluorescence

Les données d'anisotropie de fluorescence ont été obtenues en utilisant un Tecan Safire² (MTX Lab Systems) par l'analyse d'échantillons de 40µl final en plaques de 384 puits.

Les échantillons sont excités par une diode électroluminescente (DEL) produisant un rayonnement à 530nm. Cette lumière est ensuite polarisée et acheminée au monochromateur final par l'intermédiaire d'un système mettant en jeu trois lentilles optiques : la première lentille collecte et aligne selon un axe vertical la lumière produite par la DEL ; la deuxième focalise cette lumière d'excitation polarisée sur les puits puis collecte la lumière plus ou moins polarisée émise par l'échantillon; la troisième focalise la lumière d'émission sur une fibre optique qui transmet le signal lumineux à un monochromateur centré sur 580nm. Un photomultiplicateur convertit en dernier lieu le signal lumineux en signal électrique amplifié.

Les enregistrements sont soumis à l'opération mathématique suivante afin d'en déduire la valeur d'anisotropie de fluorescence (Canet D. *et al.*, 2001) :

$$AF = (A_v - A_h) / (A_v + 0,5 A_h)$$

où $A_v = I_{vh} + I_{vv}$ c'est-à-dire la somme des intensités de lumière émises après excitation selon l'axe vertical et $A_h = I_{hh} + I_{hv}$ c'est-à-dire la somme des intensités de lumière émises après excitation selon l'axe horizontal. En effet, seule la lumière d'excitation est polarisée. La détection à l'émission se fait sans polariseur.

6.2. Méthodes de détermination des constantes de dissociation :

Constante de dissociation des complexes ZEBRA/ADN :

Les données d'anisotropie de fluorescence permettent de déduire le K_d de la réaction de fixation. En effet :

$$AF = A_f F_f + A_b F_b$$

avec :

AF : anisotropie de fluorescence totale

A_f : anisotropie de fluorescence des oligonucléotides marqués libres

F_f : fraction des oligonucléotides marqués libres

A_b : anisotropie de fluorescence des oligonucléotides marqués liés

F_b : fraction des oligonucléotides marqués liés

$$\text{Sachant } F_f + F_b = 1 \text{ et } F_b = \frac{[ADN/Z]}{[ADN_{total}]}, \text{ alors } AF = A_f + (A_b - A_f) \frac{[ADN-Z]}{[ADN_{total}]} \quad (1)$$

avec : $[ADN/Z]$: concentration du complexe ZEBRA/ADN

$[ADN_{total}]$: concentration totale en ADN

En considérant la définition de la constante de dissociation K_d

$$\text{dans la réaction } [ADN] + [Z] \rightleftharpoons [ADN/Z] : K_d = \frac{[ADN][Z]}{[ADN/Z]}$$

Et la loi de conservation des éléments : $[ADN_{total}] = [ADN] + [ADN/Z]$

$$\text{et } [Z_{total}] = [Z] + [ADN/Z]$$

Nous pouvons écrire :

$$[ADN/Z]^2 + [ADN/Z] (-K_d - [ADN_{total}] - [Z_{total}]) + [ADN_{total}] [Z_{total}] = 0 \quad (2)$$

D'après (1) et par résolution de (2), nous obtenons la relation permettant de déterminer les constantes de dissociation des complexes ZEBRA/ADN :

$$AF = A_f + (A_b - A_f) \frac{(K_d + [ADN_{total}] + [Z_{total}]) - \sqrt{(K_d + [ADN_{total}] + [Z_{total}])^2 - 4[ADN_{total}][Z_{total}]}}{2[ADN_{total}]}$$

Les courbes de fixation de ZEBRA sur les oligonucléotides testés sont analysées par le logiciel *KaleidaGraph*. Chaque essai est réalisé en triplica : les K_d sont mesurés sur la moyenne des points de ces 3 essais.

Remarque : la fixation de ZEBRA sur les sites testés suit un modèle simple : ZEBRA se fixe en tant que dimère sur l'ADN (voir Chapitre V.III.1.). Cependant, nous ne pouvons utiliser l'équation typique correspondant à une fixation simple étant donné la gamme de concentration en protéine utilisée :

$$AF = \frac{AF_{max} \times [ZEBRA]_{libre}}{K_d + [ZEBRA]_{libre}}$$

(avec AF_{max} : valeur d'anisotropie atteinte par le plateau de l'hyperbole et $[ZEBRA]_{libre}$: concentration en protéine ZEBRA libre)

En effet, l'utilisation de cette relation suppose que pour tous les points de la courbe, la concentration en protéine ZEBRA libre soit en large excès par rapport à la concentration en ADN ($[ZEBRA]_{libre} \gg [ADN]$). Or, la concentration en ADN est de 10 nM et la gamme de concentrations en ZEBRA s'étale de 4 nM à 20 μ M (voir partie 6.3. ci-dessous). Donc, la condition $[ZEBRA]_{libre} \gg [ADN]$ ne pourrait être respectée pour tous les points de la courbe.

Constante de dissociation des complexes cJun/ADN, cJun A275S/ADN et Chimère ZEBRA-cJun/ADN :

Le facteur de transcription cJun existe en tant qu'hétérodimère en solution mais également en tant qu'homodimère. Cependant, en présence d'une séquence ADN cible, l'homodimère de cJun se dissocie : en effet, cJun fixe l'ADN en tant que monomère dans un premier temps. Le premier monomère ainsi fixé sur sa séquence cible ADN recrute le

deuxième monomère (Kolher J.J. *et al.*, 2001). Il s'agit d'une fixation coopérative : les courbes de fixation de l'homodimère cJun se présentent comme des sigmoïdes. Nous avons donc analysé les courbes de fixation des complexes cJun/ADN à l'aide de l'équation utilisée par Kohler J.J. *et al.* et dérivée de l'équation de Hill (Hill A.V., 1910) :

$$AF = \frac{((AF_{\max} - AF_{\min}) \times [cJun]^n)}{(K_d^n + [cJun]^n)} + AF_{\min}$$

où $AF_{\max}$ est la valeur d'anisotropie maximale

$AF_{\min}$, la valeur d'anisotropie minimale

$[cJun]$, la concentration en protéine cJun

n , le coefficient de Hill.

Les courbes de fixation ont été analysées à l'aide du logiciel *SigmaPlot*. Tout comme Kohler J.J. *et al.*, nous avons fixé le coefficient de Hill à $n=2$ (après avoir vérifié que cette contrainte ne modifiait ni la constante de dissociation calculée ni les statistiques de fit).

6.3. Protocole de mesure d'anisotropie de fluorescence

L'analyse des échantillons est réalisée en plaques opaques de 384 puits dans un tampon PBS 1X (voir Annexe C.2.5.). Pour chaque essai, 15 puits sont utilisés, contenant chacun des concentrations fixes d'oligonucléotides fluorescents (10 nM) et d'ADN aspécifique (28 nM de fragments d'ADN de poisson) et des concentrations variables de protéines (de 4 nM à 20 μ M concernant ZEBRA et de 100 nM à 37,5 μ M pour les clones de cJun). Le volume total dans chaque puits est de 40 μ l.

Chaque essai est mesuré en triplica à température ambiante et les points sont moyennés avant analyse de la courbe de fixation et détermination du K_d .

Remarque : chaque valeur d'anisotropie mesurée par l'appareil Tecan Safire² est déjà moyennée puisque l'appareil effectue 100 mesures dans chacun des 15 puits d'un même essai.

ANNEXES

Annexe A :

Table des constructions réalisées

Vecteur	Mutant	Description
pET28a	ZEBRA S186A/C189S¹⁷⁵⁻²³⁶ (ou ZEBRA ^{175-236(mut)})	▪ Résidus 175-236 du domaine bZIP de ZEBRA. ▪ Double mutant de ZEBRA où la sérine 186 a été mutée en alanine et la cystéine 189 en sérine.
pET28a	ZEBRA C189S¹⁷⁵⁻²³⁶	▪ Résidus 175-236 du domaine bZIP de ZEBRA. ▪ Mutant de ZEBRA où la cystéine 189 a été mutée en sérine.
pETM-40	ZEBRA C189S¹⁷⁵⁻²³⁶	▪ Résidus 175-236 du domaine bZIP de ZEBRA. ▪ Mutant de ZEBRA où la cystéine 189 a été mutée en sérine.
pETM-40	cJun C278S²⁶⁴⁻³²⁴	▪ Résidus 264-324 du domaine bZIP de cJun. ▪ Mutant de cJun où la cystéine 278 a été mutée en sérine.
pETM-40	cJun A275S/C278S²⁶⁴⁻³²⁴	▪ Résidus 264-324 du domaine bZIP de cJun. ▪ Double mutant de cJun où la cystéine 278 a été mutée en sérine et l'alanine 275 en sérine.
pETM-40	Chimère ZEBRA-cJun	▪ Résidus 175- 196 du domaine bZIP de ZEBRA suivi des résidus 287- 324 du domaine bZIP de cJun. ▪ Mutant où la cystéine 189 de ZEBRA a été mutée en sérine.

Annexe B :

Amorces de clonage et de mutagenèse dirigée

B.1. Mutagenèse dirigée dans le domaine bZIP de ZEBRA :

Clone de départ : ZEBRA^{175-236(mut)} – pET28a (Petosa et al., 2006) contenant les mutations S186A et C189S.

Clone final : ZEBRA C189S¹⁷⁵⁻²³⁶ – pET28a contenant uniquement la mutation C189S.

Amorce sens : 5' tacaagaatcgggtggcttccagaaaatgccgg 3'

T_m = 82.8°

Nombre nt : 33 bp

Amorce anti-sens : 3' atgttcttagcccaccgaaggctctttacggcc 5'

T_m = 79.6°

Nombre nt : 33 bp

B.2. Mutagenèse dirigée dans le domaine bZIP de cJun :

Clone de départ : cJun C278S²⁶⁴⁻³²⁴ contenant la mutation C278S.

Clone final : cJun A275S/C278S²⁶⁴⁻³²⁴ contenant les mutations A275S et C278S.

Amorce sens : 5' atgcgtaaccgcatcgcttctccaagtgccgg 3'

T_m = 85.9°

Nombre nt : 33 bp

Amorce anti-sens : 3' tacgcattggcgtagcgaaggaggttcacggcc 5'

$T_m = 84.7^\circ$

Nombre nt : 33 bp

B.3. Clonage des séquences ZEBRA C189S¹⁷⁵⁻²³⁶ et cJun C278S²⁶⁴⁻³²⁴ dans le vecteur d'expression pETM40 :

- Clone ZEBRA C189S¹⁷⁵⁻²³⁶ – pETM40

Clone de départ : ZEBRA C189S¹⁷⁵⁻²³⁶ – pET28a.

Clone final : ZEBRA C189S¹⁷⁵⁻²³⁶ – pETM40.

Amorce PCR sens : 5' atacactcccatgggcatggaataaagcgatacaagaatcgg 3'

T_m (total) = 83°

Nombre nt (total) : 43 bp

T_m (hybridant) = $65,6^\circ$

Nombre nt (hybridant) : 27 bp

Amorce PCR anti-sens : 3' tccattatcccccgacaccagattaaggtaccatctgat 5'

T_m (total) = $81,3^\circ$

Nombre nt (total) : 40 bp

T_m (hybridant) = $71,3^\circ$

Nombre nt (hybridant) : 27 bp

- Clone cJun C278S²⁶⁴⁻³²⁴ – pETM40

Clone de départ : cJun C278S²⁶⁴⁻³²⁴ – pET11.

Clone final : cJun C278S²⁶⁴⁻³²⁴ – pETM40.

Amorce PCR sens : 5' atacactcccatgggcatgaaagcggaacgtaaacgcatg 3'

T_m (total) = $84,5^\circ$

Nombre nt (total) : 40 bp

Tm (hybridant) = 67,4°

Nombre nt (hybridant) : 24 bp

Amorce PCR anti-sens : 3' gcacagcttaaacagaaagtcatgaaccactaaggtaccatctgat 5'

Tm (total) = 79,8°

Nombre nt (total) : 46 bp

Tm (hybridant) = 68,9°

Nombre nt (hybridant) : 30 bp

Annexe C :

Tampons et solutions : composition

C.1 Milieu de culture

C.1.1 Milieu Luria Broth (LB)

Composants	Quantités par L
Bacto-Tryptone	10 g
Bacto-Yeast extract	5 g
NaCl	10 g
NaOH	à pH=7.0

Le milieu est autoclavé avant utilisation.

C.1.2 Milieu Luria Broth (LB)-agar

Composants	Quantités par L
Bacto-Tryptone	10 g
Bacto-Yeast extract	5 g
NaCl	10 g
Agarose	15 g
NaOH	à pH=7.0

Le milieu est autoclavé avant utilisation. L'agar fondu a été également distribué par agitation douce après stérilisation. Les antibiotiques sont ajoutés quand le milieu est refroidi à 50°C. Environ 30 ml de cette solution sont versés dans chaque boîte de Petri.

C.1.3 Autres solutions

Kanamycin solution stock

Composants	Quantité pour 20 ml
Kanamycin	200 mg

La solution est filtrée à travers une membrane de 0.2 μ m et aliquotée dans des tubes de 1 ml. Le stock est conservé à -20°C.

IPTG (β -isopropyl-thiogalactosidase) solution stock

Composants	Quantités
β -isopropyl-thiogalactosidase	2 g
ddH ₂ O	10 ml

La solution est filtrée à travers une membrane de 0.2 μ m et aliquotée dans des tubes de 1 ml. Le stock est conservé à -20°C.

Tampon 10X de migration “cathode” pour gels SDS-PAGE

Composants	Quantités pour 1 l	Concentration finale
Tris Base	121 g	0,1 M
Tricine	180 g	0,1 M
SDS	10 g	0,1 %

Le tampon est dilué à 1 X avant utilisation et conservé à température ambiante.

Tampon 10X de migration “anode” pour gels SDS-PAGE

Composants	Quantités pour 1 l	Concentration finale
Tris Base	242 g	0,2 M

Le pH du tampon est ensuite ajusté à 8,9.

Le tampon est dilué à 1 X avant utilisation et conservé à température ambiante.

C.2 Tampons de purification protéines

C.2.1 Tampon de lyse

Solution stock de tampon de lyse

Composants	Quantités pour 200 ml	Concentration finale
NaCl	4 ml, 5 M stock	100 mM
Tris/HCl pH 6.8	4 ml, 1 M stock	20 mM

Tampon de lyse

Juste avant utilisation, 70 µL de β-mercaptoethanol (solution stock 14,5 M ; concentration finale 5 mM), 2 ml de PMSF (à 100 mM de solution stock ; 1 mM concentration finale) et deux tablettes de "complete EDTA-free" (cocktail anti-protéases, *Boehringer*) sont ajoutés à 200 ml de solution stock de tampon de lyse. Le tampon de lyse est mélangé jusqu'à la dissolution complète des tablettes.

C.2.2 Chromatographie échangeuse de cations SP/LP

Résine SP Sepharose Fast flow (GE Healthcare, code n° 17-0729-10)

TAMPON A :

Composants	Quantités pour 200 ml	Concentration finale
NaCl	0.4 ml, 5 M stock	10 mM

Tris/HCl pH 6.8	4 ml, 1 M stock	20 mM
-----------------	-----------------	-------

Ce tampon est préparé extemporanément. Une tablette de "complete EDTA-free" (cocktail anti-protéases, *Boehringer*) est ajoutée à 200 ml de solution. Le tampon A est mélangé jusqu'à la dissolution complète de la tablette puis filtré sur membrane de filtration 0,2 μ m et dégazé. Juste avant utilisation sont ajoutés 70 μ l de β -mercaptoethanol (stock 14.5 M, concentration finale 5 mM) et 0,4 ml de PMSF (stock 1 M, concentration finale 2 mM).

TAMPON B :

Composants	Quantités pour 200 ml	Concentration finale
NaCl	80 ml, 5 M stock	2 M
Tris/HCl pH 6.8	4 ml, 1 M stock	20 mM

Ce tampon est préparé extemporanément. Une tablette de "complete EDTA-free" (cocktail anti-protéases, *Boehringer*) est ajoutée à 200 ml de solution. Le tampon B est mélangé jusqu'à la dissolution complète de la tablette puis filtré sur membrane de filtration 0,2 μ m et dégazé. Juste avant utilisation sont ajoutés 70 μ l de β -mercaptoethanol (stock 14.5 M, concentration finale 5 mM) et 0,4 ml de PMSF (stock 1 M, concentration finale 2 mM).

C.2.3 Chromatographie d'exclusion de taille

Colonne Superdex S75 HR 10/30 (*Pharmacia/GE Healthcare*, code n° 71-6031-00)

TAMPON GF :

Composants	Quantités pour 200 ml	Concentration finale
NaCl	6 ml, 5 M stock	150 mM
NH ₄ acétate	30 ml, 1 M stock	150 mM
Tris/HCl pH 7.5	4 ml, 1 M stock	20 mM
DTT	1 ml, 1M stock	5 mM

PMSF	0.4 ml, 1 M stock	2 mM
------	-------------------	------

Ce tampon est préparé extemporanément. Une tablette de "complete EDTA-free" (cocktail anti-protéases, *Boehringer*) est ajoutée à 200 ml de solution. Le tampon GF est mélangé jusqu'à la dissolution complète de la tablette puis filtré sur membrane de filtration 0,2 μ m et dégazé. Juste avant utilisation sont ajoutés 1 ml de DTT (stock 1 M, concentration finale 5 mM) et 0,4 ml de PMSF (stock 1 M, concentration finale 2 mM).

C.2.4 Chromatographie d'affinité pour protéines de fusion MBP

Résine d'amylose (*NEB*, code n° E8021L)

TAMPON de lavage 1 M :

Composants	Quantités pour 200 ml	Concentration finale
NaCl	40 ml, 5 M stock	1 M
Tris/HCl pH 7.2	4 ml, 1 M stock	20 mM

Juste avant utilisation, 70 μ L de β -mercaptoethanol (solution stock 14,5 M ; concentration finale 5 mM), 2 ml de PMSF (à 100 mM de solution stock ; 1 mM concentration finale) et deux tablettes de "complete EDTA-free" (cocktail anti-protéases, *Boehringer*) sont ajoutés à 200 ml de solution stock de tampon de lavage 1 M. tampon de lavage 1 M est mélangé jusqu'à la dissolution complète des tablettes.

TAMPON de lavage 1,5 M :

Composants	Quantités pour 200 ml	Concentration finale
NaCl	60 ml, 5 M stock	1,5 M
Tris/HCl pH 7.2	4 ml, 1 M stock	20 mM

Juste avant utilisation, 70 μ L de β -mercaptoethanol (solution stock 14,5 M ; concentration finale 5 mM), 2 ml de PMSF (à 100 mM de solution stock ; 1 mM concentration finale) et deux tablettes de "complete EDTA-free" (cocktail anti-protéases, *Boehringer*) sont ajoutés à 200 ml de solution stock de tampon de lavage 1,5 M. Le tampon de lavage 1,5 M est mélangé jusqu'à la dissolution complète des tablettes.

TAMPON d'élution :

Composants	Quantités pour 200 ml	Concentration finale
NaCl	60 ml, 5 M stock	1,5 M
Tris/HCl pH 7.2	4 ml, 1 M stock	20 mM
Maltose	2 ml, 1M stock	10 mM

Juste avant utilisation, 70 μ L de β -mercaptoéthanol (solution stock 14,5 M ; concentration finale 5 mM), 2 ml de PMSF (à 100 mM de solution stock ; 1 mM concentration finale) et deux tablettes de "complete EDTA-free" (cocktail anti-protéases, *Boehringer*) sont ajoutés à 200 ml de solution stock de tampon d'élution. Le tampon d'élution est mélangé jusqu'à la dissolution complète des tablettes.

C.2.5 Chromatographie d'exclusion de taille pour protéines de fusion MBP

Colonne Superdex 200 10/300 GL (*Pharmacia/GE Healthcare*, code n° 17-5175-01)

TAMPON GF2 : PBS 1x

Composants	Quantités pour 200 ml
NaCl	2 g
KCl	50 mg
Na ₂ HPO ₄	360 mg
KH ₂ PO ₄	60 mg
HCl	ajuster à pH 7.4

Le tampon GF2 est filtré sur membrane de filtration 0,2 μ m et dégazé.

C.3 Tampons de purification ADN

Colonne MonoQ HR 10/10 (*Pharmacia/GE Healthcare*, code n°17-0506-01)

TAMPON DNA/A :

Composants	Quantités pour 200 ml	Concentration finale
NaOH	2 ml, 1 M stock	10 mM

Le tampon DNA/A est filtré sur membrane de filtration 0,2 µm et dégazé.

TAMPON DNA/B :

Composants	Quantités pour 200 ml	Concentration finale
NaOH	2 ml, 1 M stock	10 mM
NaCL	40 ml, 5 M stock	1 M

Le tampon DNA/B est filtré sur membrane de filtration 0,2 µm et dégazé.

C.4 Tampons d'hybridation des oligonucléotides simple brin

C.4.1 Tampon d'hybridation des oligonucléotides des expériences de cristallographie

TAMPON d'hybridation-X

Composants	Quantités pour 10 ml	Concentration finale
NaCl	0.2 ml, 5 M stock	0,1 M
Hepes pH 7.5	0.1 ml, 1 M stock	10 mM
MgCl ₂	0.1 ml, 1 M stock	10 mM
DTT	10 µl, 1M stock	1 mM

Le tampon d'hybridation-X est filtré sur membrane de filtration 0,2 µm et conservé à 4°C.

C.4.2 Tampon d'hybridation des oligonucléotides des expériences d'anisotropie de fluorescence

TAMPON d'hybridation-AF

Composants	Quantités pour 10 ml	Concentration finale
NaCl	0.3 ml, 5 M stock	0,15 M
TRIS/HCl pH 7.5	0.1 ml, 1 M stock	10 mM
EDTA	10 µl, 1 M stock	1 mM

Le tampon d'hybridation-AF est filtré sur membrane de filtration 0,2 µm et conservé à 4°C.

Annexe D :

Cartes des vecteurs d'expression en bactéries

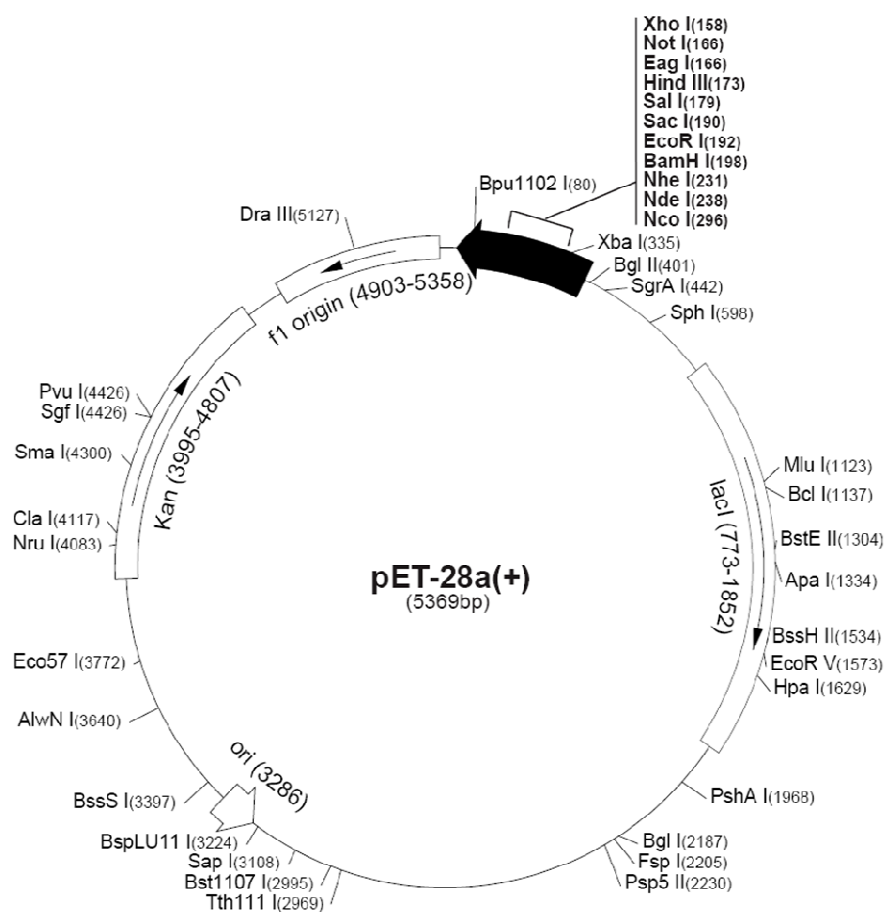
D.1 Vecteur d'expression pET28a :

pET-28a-c(+) Vectors

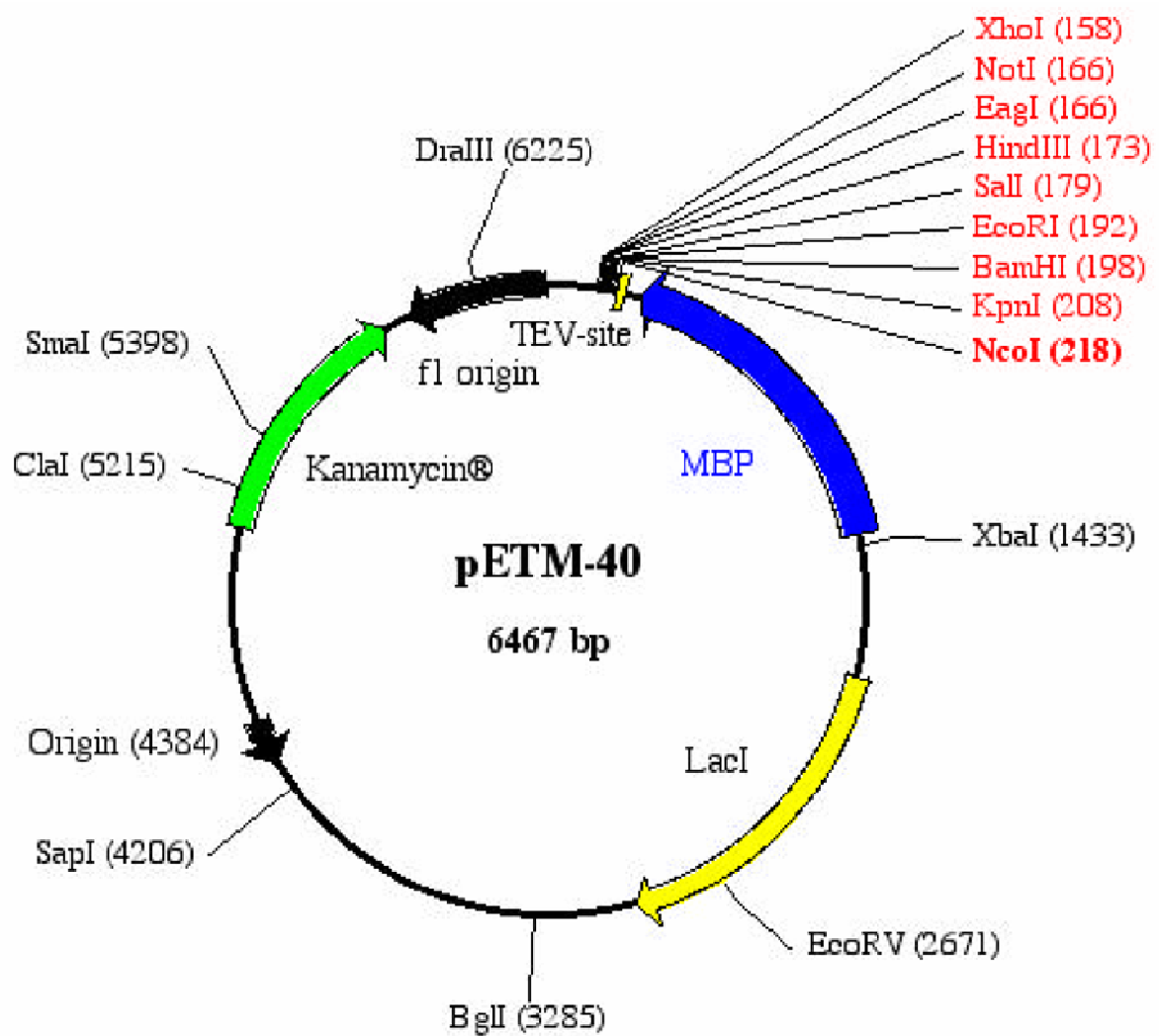
	Cat. No.
pET-28a DNA	69864-3
pET-28b DNA	69865-3
pET-28c DNA	69866-3

pET-28a(+) sequence landmarks

T7 promoter	370-386
T7 transcription start	339
His*Tag coding sequence	270-287
T7*Tag coding sequence	207-239
Multiple cloning sites (<i>Bam</i> HI - <i>Xho</i> I)	158-203
His*Tag coding sequence	140-157
T7 terminator	25-72
<i>lacI</i> coding sequence	773-1852
pBR322 origin	3285
Kan coding sequence	3995-4807
f1 origin	4903-5358

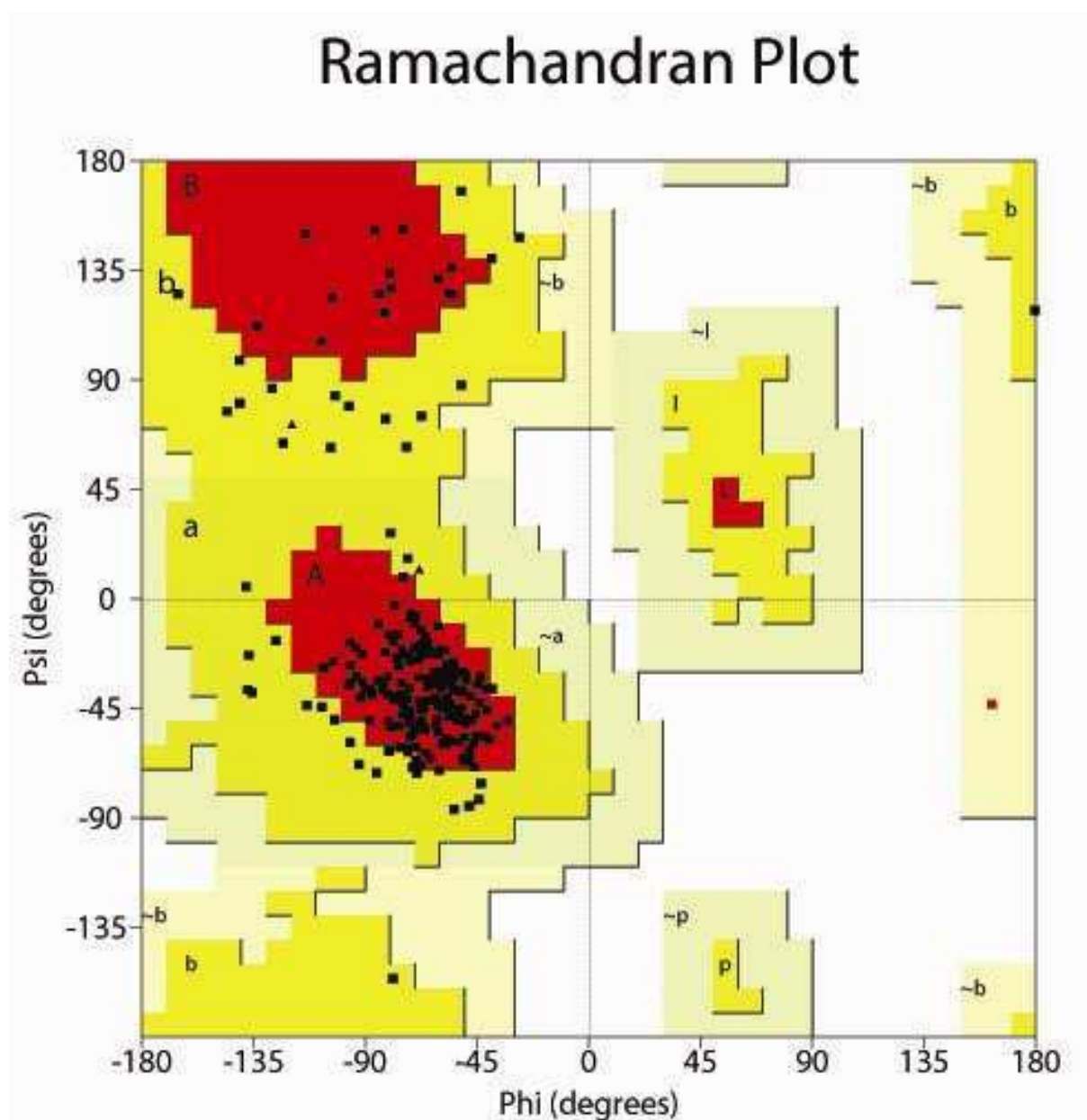


D.2 Vecteur d'expression pETM40 :



Annexe E :

Diagramme de Ramachandran



Annexe F :

Paramètres des alignements structuraux des complexes ZEBRA/ADN

F.1 Alignement structural des 2 molécules du complexe ZEBRA/meZRE2 présentes dans l'unité asymétrique :

F.1.1 Dénomination des hélices α protéines et des doubles hélices ADN

Les chaînes X, Y, W et Z correspondent aux hélices α des dimères de ZEBRA de chaque molécule (molécule I : chaînes X + W et molécule II : chaînes Y + Z).

Les chaînes A, B, C et D correspondent aux brins des doubles hélices ADN de chaque molécule (molécule I : brins D + E et molécule II : brins A + B).

("res" : résidus et "ntd" : nucléotides)

Molécule I du complexe ZEBRA/meZRE2	Chaîne X res 176-236	Chaîne W res 176-235	Chaîne D ntd 3-17	Chaîne E ntd 4-19
Molécule II du complexe ZEBRA/meZRE2	Chaîne Y res 176-236	Chaîne Z res 175-235	Chaîne A ntd 3-15	Chaîne B ntd 6-16

F.1.2 Paramètres des alignements structuraux des molécules I et II du complexe ZEBRA/meZRE2 :

Le site meZRE2 est asymétrique (voir chapitre IV). Les deux hélices α d'un dimère de ZEBRA qui vont fixer cet heptamère ne vont pas former les mêmes interactions protéine/ADN avec les deux demi-sites de l'heptamère meZRE2. Ainsi, deux complexes ZEBRA/meZRE2 peuvent être alignés de 2 façons différentes : soit les hélices α X et W de la molécule I sont respectivement alignées avec leurs hélices α correspondantes de la molécule

II, Y et Z (Alignement 1). Ou bien à l'inverse : les hélices α X et W de la molécule I sont respectivement alignées avec les hélices α Z et Y de la molécule II (Alignement 2).

De plus, l'alignement peut être réalisé de façon « globale », c'est-à-dire en considérant l'ensemble du complexe ZEBRA/meZRE2 (on minimise les rmsd des C α de ZEBRA et des phosphates de l'ADN) ou selon la protéine (minimisation des rmsd des C α de ZEBRA uniquement) ou encore selon la double hélice d'ADN. Dans ce dernier cas, on ne minimise uniquement que les rmsd des atomes de la double hélice d'ADN.

	Alignement 1 (translation) YZAB → XWDE			Alignement 2 (rotation puis translation) YZAB → WXED		
	Nb atomes	rmsd (Å)	rmsd₁₀₀*	Nb atomes	rmsd (Å)	rmsd₁₀₀
Protéine						
Atomes C α	120	1,03	0,94	120	1,87	1,71
Tous atomes	861	1,47	0,71	857	2,27	1,09
ADN						
P atomes	22	0,94	3,86	21	1,12	5,10
Tous atomes	485	1,16	0,65	439	1,36	0,78
Complexe protéine/ADN						
C α & P atomes	142	1,05	0,89	140	3,08	2,64
Tous atomes	1346	1,41	0,61	1274	3,48	0,76

* rmsd normalisé de Carugo and Pongor : $\text{rmsd}_{100} = \text{rmsd} / (1 + \ln(N/100)^{1/2})$

(Carugo O. *et al.*, 2001)

F.2 Alignement structural des complexes ZEBRA/meZRE2 et ZEBRA/AP1 :

F.2.1 Dénomination des hélices α protéines et des doubles hélices ADN

Les chaînes X, Y, W et Z correspondent aux hélices α des dimères de ZEBRA de chaque complexe (complexe ZEBRA/meZRE2 : chaînes X + W et complexe ZEBRA/AP1 : chaînes Y + Z).

Les chaînes A, B, C et D correspondent aux brins des doubles hélices ADN de chaque complexe (complexe ZEBRA/meZRE2 : chaînes D + E et complexe ZEBRA/AP1 : chaînes A + B).

(“res” : résidus et “ntd” : nucléotides)

Complexe ZEBRA/meZRE2	Chaîne X res 176-236	Chaîne W res 176-235	Chaîne D ntd 3-17	Chaîne E ntd 4-19
Complexe ZEBRA/AP1	Chaîne Y res 174-236	Chaîne Z res 175-236	Chaîne A ntd 1-19	Chaîne B ntd 101-119

F.2.2 Paramètres des alignements structuraux des molécules I et II du complexe ZEBRA/meZRE2 :

	Alignement 1 (translation) YZAB → XWDE			Alignement 2 (rotation puis translation) YZAB → WXED		
	Nb atomes	rmsd (Å)	rmsd₁₀₀*	Nb atomes	rmsd (Å)	rmsd₁₀₀
Protéine						
Atomes Ca	120	1,55	1,42	120	1,28	1,17
Tous atomes	976	1,99	0,93	981	1,78	0,83
ADN						
P atomes	29	1,27	3,30	29	1,41	3,68
Tous atomes	577	1,65	0,87	560	1,72	0,92
Complexe protéine/ADN						
Ca & P atomes	149	1,66	1,38	149	3,34	2,78
Tous atomes	1553	2,06	0,87	1541	3,55	1,50

* rmsd normalisé de Carugo and Pongor : $\text{rmsd}_{100} = \text{rmsd} / (1 + \ln(N/100))^{1/2}$

(Carugo and Pongor, 2001)

REFERENCES

BIBLIOGRAPHIQUES

Table bibliographique

Adamson, A. L., and Kenney, S. C. (1998). Rescue of the Epstein-Barr virus BZLF1 mutant, Z(S186A), early gene activation defect by the BRLF1 gene product. *Virology* 251, 187-197.

Adamson, A. L., and Kenney, S. (1999). The Epstein-Barr virus BZLF1 protein interacts physically and functionally with the histone acetylase CREB-binding protein. *J Virol* 73, 6551-6558.

Adamson, A. L., Darr, D., Holley-Guthrie, E., Johnson, R. A., Mauser, A., Swenson, J., and Kenney, S. (2000). Epstein-Barr virus immediate-early proteins BZLF1 and BRLF1 activate the ATF2 transcription factor by increasing the levels of phosphorylated p38 and c-Jun N-terminal kinases. *J Virol* 74, 1224-1233.

Aho, S., Buisson, M., Pajunen, T., Ryoo, Y. W., Giot, J. F., Gruffat, H., Sergeant, A., and Uitto, J. (2000). Ubinuclein, a novel nuclear protein interacting with cellular and viral transcription factors. *J Cell Biol* 148, 1165-1176.

Altmann, M., and Hammerschmidt, W. (2005). Epstein-Barr virus provides a new paradigm: a requirement for the immediate inhibition of apoptosis. *PLoS Biol* 3, e404.

Amon, W., and Farrell, P. J. (2005). Reactivation of Epstein-Barr virus from latency. *Rev Med Virol* 15, 149-156.

Baumann, M., Gires, O., Kolch, W., Mischak, H., Zeidler, R., Pich, D., and Hammerschmidt, W. (2000). The PKC targeting protein RACK1 interacts with the Epstein-Barr virus activator protein BZLF1. *Eur J Biochem* 267, 3891-3901.

Baylin, S. B., and Herman, J. G. (2000). DNA hypermethylation in tumorigenesis: epigenetics joins genetics. *Trends Genet* 16, 168-174.

Berk, A. J., Lee, F., Harrison, T., Williams, J., and Sharp, P. A. (1979). Pre-early adenovirus 5 gene product regulates synthesis of early viral messenger RNAs. *Cell* 17, 935-944.

Bhende, P. M., Seaman, W. T., Delecluse, H. J., and Kenney, S. C. (2004). The EBV lytic switch protein, Z, preferentially binds to and activates the methylated viral genome. *Nat Genet* 36, 1099-1104.

Bhende, P. M., Seaman, W. T., Delecluse, H. J., and Kenney, S. C. (2005). BZLF1 activation of the methylated form of the BRLF1 immediate-early promoter is regulated by BZLF1 residue 186. *J Virol* 79, 7338-7348.

Bhende, P. M., Dickerson, S. J., Sun, X., Feng, W. H. and Kenney, S. C. (2007). X-Box-Binding protein 1 activates lytic Epstein-Barr virus gene expression in combination with protein kinase D. *J Virol* 81, 7763-7770.

- Bonnet, M., Guinebretière, J.M., Kremmer, E., Grunewald, V., Benhamou, E., Contesso, G. and Joab, I. (1999). Detection of Epstein-Barr virus in invasive breast cancers. *J Natl Cancer Inst* 91, 1376-81.
- Canet, D., Doering, K., Dobson, C. M., and Dupont, Y. (2001). High-sensitivity fluorescence anisotropy detection of protein-folding events: application to alpha-lactalbumin. *Biophys J* 80, 1996-2003.
- Carey, M., Kolman, J., Katz, D. A., Gradoville, L., Barberis, L., and Miller, G. (1992). Transcriptional synergy by the Epstein-Barr virus transactivator ZEBRA. *J Virol* 66, 4803-4813.
- Carugo, O., and Pongor, S. (2001). A normalized root-mean-square distance for comparing protein three-dimensional structures. *Protein Sci* 10, 1470-1473.
- Cayrol, C., and Flemington, E. K. (1995). Identification of cellular target genes of the Epstein-Barr virus transactivator Zta: activation of transforming growth factor beta igh3 (TGF-beta igh3) and TGF-beta 1. *J Virol* 69, 4206-4212.
- Cayrol, C., and Flemington, E. K. (1996). The Epstein-Barr virus bZIP transcription factor Zta causes G0/G1 cell cycle arrest through induction of cyclin-dependent kinase inhibitors. *Embo J* 15, 2748-2759.
- Chang, Y. N., Dong, D. L., Hayward, G. S., and Hayward, S. D. (1990). The Epstein-Barr virus Zta transactivator: a member of the bZIP family with unique DNA-binding specificity and a dimerization domain that lacks the characteristic heptad leucine zipper motif. *J Virol* 64, 3358-3369.
- Chang, Y., Chang, S. S., Lee, H. H., Doong, S. L., Takada, K., and Tsai, C. H. (2004). Inhibition of the Epstein-Barr virus lytic cycle by Zta-targeted RNA interference. *J Gen Virol* 85, 1371-1379.
- Chen, C. J., Deng, Z., Kim, A. Y., Blobel, G. A., and Lieberman, P. M. (2001). Stimulation of CREB binding protein nucleosomal histone acetyltransferase activity by a class of transcriptional activators. *Mol Cell Biol* 21, 476-487.
- Chevallier-Greco, A., Manet, E., Chavrier, P., Mosnier, C., Daillie, J., and Sergeant, A. (1986). Both Epstein-Barr virus (EBV)-encoded trans-acting factors, EB1 and EB2, are required to activate transcription from an EBV early promoter. *EMBO J* 5, 3243-3249.
- Chi, T., and Carey, M. (1993). The ZEBRA activation domain: modular organization and mechanism of action. *Mol Cell Biol* 13, 7045-7055.
- Chi, T., Lieberman, P., Ellwood, K., and Carey, M. (1995). A general mechanism for transcriptional synergy by eukaryotic activators. *Nature* 377, 254-257.
- Chi, T., and Carey, M. (1996). Assembly of the isomerized TFIIA--TFIID--TATA ternary complex is necessary and sufficient for gene activation. *Genes Dev* 10, 2540-2550.

Cho, M. S., Bornkamm, G. W., and zur Hausen, H. (1984). Structure of defective DNA molecules in Epstein-Barr virus preparations from P3HR-1 cells. *J Virol* 51, 199-207.

Cochet, C., Martel-Renoir, D., Grunewald, V., Bosq, J., Cochet, G., Schwaab, G., Bernaudin, J. F., and Joab, I. (1993). Expression of the Epstein-Barr virus immediate early gene, BZLF1, in nasopharyngeal carcinoma tumor cells. *Virology* 197, 358-365.

Countryman, J., and Miller, G. (1985). Activation of expression of latent Epstein-Barr herpesvirus after gene transfer with a small cloned subfragment of heterogeneous viral DNA. *Proc Natl Acad Sci U S A* 82, 4085-4089.

Countryman, J., Jenson, H., Seibl, R., Wolf, H., and Miller, G. (1987). Polymorphic proteins encoded within BZLF1 of defective and standard Epstein-Barr viruses disrupt latency. *J Virol* 61, 3672-3679.

Delius, H., and Bornkamm, G. W. (1978). Heterogeneity of Epstein-Barr virus. III. Comparison of a transforming and a nontransforming virus by partial denaturation mapping of their DNAs. *J Virol* 27, 81-89.

Deng, Z., Chen, C. J., Zerby, D., Delecluse, H. J., and Lieberman, P. M. (2001). Identification of acidic and aromatic residues in the Zta activation domain essential for Epstein-Barr virus reactivation. *J Virol* 75, 10334-10347.

Deng, Z., Chen, C. J., Chamberlin, M., Lu, F., Blobel, G. A., Speicher, D., Cirillo, L. A., Zaret, K. S., and Lieberman, P. M. (2003). The CBP bromodomain and nucleosome targeting are required for Zta-directed nucleosome acetylation and transcription activation. *Mol Cell Biol* 23, 2633-2644.

Dixon, R. A., and Schaffer, P. A. (1980). Fine-structure mapping and functional analysis of temperature-sensitive mutants in the gene encoding the herpes simplex virus type 1 immediate early protein VP175. *J Virol* 36, 189-203.

Dupuis, J., Emile, J. F., Mounier, N., Gisselbrecht, C., Martin-Garcia, N., Petrella, T., Bouabdallah, R., Berger, F., Delmer, A., Coiffier, B., *et al.* (2006). Prognostic significance of Epstein-Barr virus in nodal peripheral T-cell lymphoma, unspecified: A Groupe d'Etude des Lymphomes de l'Adulte (GELA) study. *Blood* 108, 4163-4169.

Epstein, M. A., Achong, B. G., and Barr, Y. M. (1964). Virus Particles in Cultured Lymphoblasts from Burkitt's Lymphoma. *Lancet* 1, 702-703.

Farrell, P. J., Rowe, D. T., Rooney, C. M., and Kouzarides, T. (1989). Epstein-Barr virus BZLF1 trans-activator specifically binds to a consensus AP-1 site and is related to c-fos. *EMBO J* 8, 127-132.

Fixman, E. D., Hayward, G. S., and Hayward, S. D. (1992). trans-acting requirements for replication of Epstein-Barr virus ori-Lyt. *J Virol* 66, 5030-5039.

Flamand, L., and Menezes, J. (1996). Cyclic AMP-responsive element-dependent activation of Epstein-Barr virus zebra promoter by human herpesvirus 6. *J Virol* 70, 1784-1791.

Flavell, K. J., Biddulph, J. P., Constandinou, C. M., Lowe, D., Scott, K., Crocker, J., Young, L. S., and Murray, P. G. (2000). Variation in the frequency of Epstein-Barr virus-associated Hodgkin's disease with age. *Leukemia* 14, 748-753.

Flavell, K. J., and Murray, P. G. (2000). Hodgkin's disease and the Epstein-Barr virus. *Mol Pathol* 53, 262-269.

Flemington, E., and Speck, S. H. (1990a). Epstein-Barr virus BZLF1 trans activator induces the promoter of a cellular cognate gene, c-fos. *J Virol* 64, 4549-4552.

Flemington, E., and Speck, S. H. (1990b). Evidence for coiled-coil dimer formation by an Epstein-Barr-virus transactivator that lacks a heptad repeat of leucine residues. *Proc Natl Acad Sci* 87, 9459-9463.

Francis, A. L., Gradoville, L., and Miller, G. (1997). Alteration of a single serine in the basic domain of the Epstein-Barr virus ZEBRA protein separates its functions of transcriptional activation and disruption of latency. *J Virol* 71, 3054-3061.

Francis, A., Ragoczy, T., Gradoville, L., Heston, L., El-Guindy, A., Endo, Y., and Miller, G. (1999). Amino acid substitutions reveal distinct functions of serine 186 of the ZEBRA protein in activation of early lytic cycle genes and synergy with the Epstein-Barr virus R transactivator. *J Virol* 73, 4543-4551.

Fujii, Y., Shimizu, T., Toda, T., Yanagida, M., and Hakoshima, T. (2000). Structural basis for the diversity of DNA recognition by bZIP transcription factors. *Nat Struct Biol* 7, 889-893.

Fuks, F. (2003). [DNA methyltransferases: from chromatin remodeling to cancer]. *Med Sci (Paris)* 19, 477-480.

Gao, Z., Krithivas, A., Finan, J. E., Semmes, O. J., Zhou, S., Wang, Y., and Hayward, S. D. (1998). The Epstein-Barr virus lytic transactivator Zta interacts with the helicase-primase replication proteins. *J Virol* 72, 8559-8567.

Gershburg, E., and Pagano, J. S. (2005). Epstein-Barr virus infections: prospects for treatment. *J Antimicrob Chemother* 56, 277-281.

Glaser, R., and Rapp, F. (1972). Rescue of Epstein-Barr virus from somatic cell hybrids of Burkitt lymphoblastoid cells. *J Virol* 10, 288-296.

Glover, J. N., and Harrison, S. C. (1995). Crystal structure of the heterodimeric bZIP transcription factor c-Fos-c-Jun bound to DNA. *Nature* 373, 257-261.

Grogan, E., Jenson, H., Countryman, J., Heston, L., Gradoville, L., and Miller, G. (1987). Transfection of a rearranged viral DNA fragment, WZhet, stably converts latent Epstein-Barr viral infection to productive infection in lymphoid cells. *Proc Natl Acad Sci U S A* 84, 1332-1336.

Gruffat, H., Manet, E., and Sergeant, A. (2002). MEF2-mediated recruitment of class II HDAC at the EBV immediate early gene BZLF1 links latency and chromatin remodeling. *EMBO Rep* 3, 141-146.

Grunewald, V., Bonnet, M., Boutin, S., Yip, T., Louzir, H., Levrero, M., Seigneurin, J. M., Raphael, M., Touitou, R., Maartel-Renoir, D., Cochet, C., Durandy, A., Andre, P., Lau, W., Zeng, Y., and Joab, I. (1998). Amino-acid change in the Epstein-Barr-virus ZEBRA protein in undifferentiated nasopharyngeal carcinomas from Europe and North Africa. *Int J Cancer* 75, 497-503.

Gutsch, D. E., Holley-Guthrie, E. A., Zhang, Q., Stein, B., Blonar, M. A., Baldwin, A. S., and Kenney, S. C. (1994). The bZIP transactivator of Epstein-Barr virus, BZLF1, functionally and physically interacts with the p65 subunit of NF-kappa B. *Mol Cell Biol* 14, 1939-1948.

Hammerschmidt, W., and Sugden, B. (1988). Identification and characterization of oriLyt, a lytic origin of DNA replication of Epstein-Barr virus. *Cell* 55, 427-433.

Heinemann, U., and Hahn, M. (1992). C-C-A-G-G-C-m5C-T-G-G. Helical fine structure, hydration, and comparison with C-C-A-G-G-C-C-T-G-G. *J Biol Chem* 267, 7332-7341.

Hendrich, B., and Bird, A. (1998). Identification and characterization of a family of mammalian methyl-CpG binding proteins. *Mol Cell Biol* 18, 6538-6547.

Henle, W., Henle, G., Zajac, B. A., Pearson, G., Waubke, R., and Scriba, M. (1970). Differential reactivity of human serums with early antigens induced by Epstein-Barr virus. *Science* 169, 188-190.

Heston, L., Rabson, M., Brown, N., and Miller, G. (1982). New Epstein-Barr virus variants from cellular subclones of P3J-HR-1 Burkitt lymphoma. *Nature* 295, 160-163.

Hicks, M. R., Al-Mehairi, S. S., and Sinclair, A. J. (2003). The zipper region of Epstein-Barr virus bZIP transcription factor Zta is necessary but not sufficient to direct DNA binding. *J Virol* 77, 8173-8177.

Hill, A. V. (1910). A new mathematical treatment of changes of ionic concentration in muscle and nerve under the action of electric currents, with a theory as to their mode of excitation. *J Physiol* 40, 190-224.

Hinuma, Y., Konn, M., Yamaguchi, J., Wudarski, D. J., Blakeslee, J. R., Jr., and Grace, J. T., Jr. (1967). Immunofluorescence and herpes-type virus particles in the P3HR-1 Burkitt lymphoma cell line. *J Virol* 1, 1045-1051.

Hong, Y., Holley-Guthrie, E., and Kenney, S. (1997). The bZip dimerization domain of the Epstein-Barr virus BZLF1 (Z) protein mediates lymphoid-specific negative regulation. *Virology* 229, 36-48.

Hong, G. K., Gulley, M. L., Feng, W. H., Delecluse, H. J., Holley-Guthrie, E., and Kenney, S. C. (2005a). Epstein-Barr virus lytic infection contributes to lymphoproliferative disease in a SCID mouse model. *J Virol* 79, 13993-14003.

Hong, G.K., Kumar, P., Wang, L., Damania, B., Gulley, M.L., Delecluse H.J., Polverini, P.J. and Kenney, S.C. (2005b). Epstein-Barr virus lytic infection is required for efficient

production of the angiogenesis factor vascular endothelial growth factor in lymphoblastoid cell lines. *J Virol* 79, 13984-92.

Jenson, H. B., and Miller, G. (1988). Polymorphisms of the region of the Epstein-Barr virus genome which disrupts latency. *Virology* 165, 549-564.

Ji, K.M., Li, C.L., Meng, G., Han, A.D. and Wu, X.L. (2008). New BZLF1 sequence variations in EBV-associated undifferentiated nasopharyngeal carcinoma in southern China. *Arch Virol* 153, 1949-53.

Jones, N., and Shenk, T. (1979). An adenovirus type 5 early gene function regulates expression of other early viral genes. *Proc Natl Acad Sci U S A* 76, 3665-3669.

Jones, P. L., Veenstra, G. J., Wade, P. A., Vermaak, D., Kass, S. U., Landsberger, N., Strouboulis, J., and Wolffe, A. P. (1998). Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription. *Nat Genet* 19, 187-191.

Jones, P. A., and Takai, D. (2001). The role of DNA methylation in mammalian epigenetics. *Science* 293, 1068-1070.

Karlsson, Q. H., Schelcher, C., Verrall, E., Petosa, C., and Sinclair, A. J. (2008). Methylated DNA recognition during the reversal of epigenetic silencing is regulated by cysteine and serine residues in the Epstein-Barr virus lytic switch protein. *PLoS Pathog* 4, e1000005.

Kelly, G. L., Milner, A. E., Baldwin, G. S., Bell, A. I., and Rickinson, A. B. (2006). Three restricted forms of Epstein-Barr virus latency counteracting apoptosis in c-myc-expressing Burkitt lymphoma cells. *Proc Natl Acad Sci U S A* 103, 14935-14940.

Kennedy, G., Komano, J., and Sugden, B. (2003). Epstein-Barr virus provides a survival factor to Burkitt's lymphomas. *Proc Natl Acad Sci U S A* 100, 14269-14274.

Kohler, J. J., and Schepartz, A. (2001). Kinetic studies of Fos.Jun.DNA complex formation: DNA binding prior to dimerization. *Biochemistry* 40, 130-142.

Kouzarides, T., Packham, G., Cook, A., and Farrell, P. J. (1991). The BZLF1 protein of EBV has a coiled coil dimerisation domain without a heptad leucine repeat but with homology to the C/EBP leucine zipper. *Oncogene* 6, 195-204.

Kraus, R. J., Perrigoue, J. G., and Mertz, J. E. (2003). ZEB negatively regulates the lytic-switch BZLF1 gene promoter of Epstein-Barr virus. *J Virol* 77, 199-207.

Laichalk, L. L., Hochberg, G. J., Babcock, G. J., Freeman, R. B., and Thorley-Lawson, D. A., (2002). The dispersal of mucosal memory B cells : evidence from persistent EBV infection. *Immunity* 16, 745-754.

Laichalk, L. L., and Thorley-Lawson, D. A., (2005). Terminal differentiation into plasma cells initiates the replicative cycle of Epstein-Barr virus in vivo. *J Virol* 79, 1296-1307.

Landschulz, W. H., Johnson, P. F., and McKnight, S. L. (1989). The DNA binding domain of the rat liver nuclear protein C/EBP is bipartite. *Science* 243, 1681-1688.

Lau, R., Packham, G., and Farrell, P. J., (1992). Differential splicing of Epstein-Barr-virus immediate-early RNA. *J Virol* 66, 6233-6236.

Le Roux, F., Sergeant, A., and Corbo, L. (1996). Epstein-Barr virus (EBV) EB1/Zta protein provided in trans and competent for the activation of productive cycle genes does not activate the BZLF1 gene in the EBV genome. *J Gen Virol* 77 (Pt 3), 501-509.

Lieberman, P. M., and Berk, A. J. (1990). In vitro transcriptional activation, dimerization, and DNA-binding specificity of the Epstein-Barr virus Zta protein. *J Virol* 64, 2560-2568.

Lieberman, P. M., and Berk, A. J. (1991). The Zta trans-activator protein stabilizes TFIID association with promoter DNA by direct protein-protein interaction. *Genes Dev* 5, 2441-2454.

Lieberman, P. M., and Berk, A. J. (1994). A mechanism for TAFs in transcriptional activation: activation domain enhancement of TFIID-TFIIA--promoter DNA complex formation. *Genes Dev* 8, 995-1006.

Lieberman, P. M., Ozer, J., and Gursel, D. B. (1997). Requirement for transcription factor IIA (TFIIA)-TFIID recruitment by an activator depends on promoter structure and template competition. *Mol Cell Biol* 17, 6624-6632.

Liu, P., Liu, S., and Speck, S. H. (1998). Identification of a negative cis element within the ZII domain of the Epstein-Barr virus lytic switch BZLF1 gene promoter. *J Virol* 72, 8230-8239.

Liu, S., Liu, P., Borrás, A., Chatila, T., and Speck, S. H. (1997). Cyclosporin A-sensitive induction of the Epstein-Barr virus lytic switch is mediated via a novel pathway involving a MEF2 family member. *Embo J* 16, 143-153.

Liu, Y. J., and Arpin, C. (1997). Germinal center development. *Immunol Rev* 156, 111-126.

Lin, K. I., Angelin-Duclos, C., Kuo, T. C., and Calame, K. (2002). Blimp-1-dependent repression of Pax-5 is required for differentiation of B cells to immunoglobulin M-secreting plasma cells. *Mol Cell Biol* 22, 4771-4780.

Liu, Y. J., and Arpin, C. (1997). Germinal center development. *Immunol Rev* 156, 111-126.

Lu, K., Wang, W., Xie, Z., Wong, B. S., Li, R., Petersen, R. B., Sy, M. S., and Chen, S. G. (2000). Expression and structural characterization of the recombinant human doppel protein. *Biochemistry* 39, 13575-13583.

Lu, J., Chua, H. H., Chen, S. Y., Chen, J. Y., and Tsai, C. H. (2003). Regulation of matrix metalloproteinase-1 by Epstein-Barr virus proteins. *Cancer Res* 63, 256-262.

Lundblad, J. R., Laurance, M., and Goodman, R. H. (1996). Fluorescence polarization analysis of protein-DNA and protein-protein interactions. *Mol Endocrinol* 10, 607-612.

McPherson, A., Malkin, A. J., and Kuznetsov, Y. G. (1995). The science of macromolecular crystallization. *Structure* 3, 759-768.

Mahot, S., Sergeant, A., Drouet, E., and Gruffat, H. (2003). A novel function for the Epstein-Barr virus transcription factor EB1/Zta: induction of transcription of the hIL-10 gene. *J Gen Virol* 84, 965-974.

Martel-Renoir, D., Wesner, M., and Joab, I. (2000). Dimerization of the Epstein-Barr virus ZEBRA protein in the yeast two-hybrid system. Comparison of a ZEBRA variant with the B95-8 form. *Biochimie* 82, 139-145.

Matthews, B. W. (1968). Solvent content of protein crystals. *J Mol Biol* 33, 491-497.

Mausser, A., Holley-Guthrie, E., Simpson, D., Kaufmann, W., and Kenney, S. (2002). The Epstein-Barr virus immediate-early protein BZLF1 induces both a G(2) and a mitotic block. *J Virol* 76, 10030-10037.

Menezes, J., Leibold, W., and Klein, G. (1975). Biological differences between Epstein-Barr virus (EBV) strains with regard to lymphocyte transforming ability, superinfection and antigen induction. *Exp Cell Res* 92, 478-484.

Mettenleiter, T. C., Klupp, B. G., and Granzow, H. (2006). Herpesvirus assembly: a tale of two membranes. *Curr Opin Microbiol* 9, 423-429.

Miller, G., Robinson, J., Heston, L., and Lipman, M. (1974). Differences between laboratory strains of Epstein-Barr virus based on immortalization, abortive infection, and interference. *Proc Natl Acad Sci U S A* 71, 4006-4010.

Miller, G., Rabson, M., and Heston, L. (1984). Epstein-Barr virus with heterogeneous DNA disrupts latency. *J Virol* 50, 174-182.

Montalvo, E. A., Parmley, R. T., and Grose, C. (1985). Structural analysis of the varicella-zoster virus gp98-gp62 complex: posttranslational addition of N-linked and O-linked oligosaccharide moieties. *J Virol* 53, 761-770.

Montalvo, E. A., Shi, Y., Shenk, T. E., and Levine, A. J. (1991). Negative regulation of the BZLF1 promoter of Epstein-Barr virus. *J Virol* 65, 3647-3655.

Morand, P., Budayova-Spano, M., Perrissin, M., Muller, C. W., and Petosa, C. (2006). Expression, purification, crystallization and preliminary X-ray analysis of a C-terminal fragment of the Epstein-Barr virus ZEBRA protein. *Acta Crystallogr Sect F Struct Biol Cryst Commun* 62, 210-214.

Murray, P. G., Young, L. S., Rowe, M., and Crocker, J. (1992). Immunohistochemical demonstration of the Epstein-Barr virus-encoded latent membrane protein in paraffin sections of Hodgkin's disease. *J Pathol* 166, 1-5.

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.

Ozer, J., Bolden, A. H., and Lieberman, P. M. (1996). Transcription factor IIA mutations show activator-specific defects and reveal a IIA function distinct from stimulation of TBP-DNA binding. *J Biol Chem* 271, 11182-11190.

Petosa, C., Morand, P., Baudin, F., Moulin, M., Artero, J. B., and Muller, C. W. (2006). Structural basis of lytic cycle activation by the Epstein-Barr virus ZEBRA protein. *Mol Cell* 21, 565-572.

Pfützner, E., Becker, P., Rolke, A., and Schule, R. (1995). Functional antagonism between the retinoic acid receptor and the viral transactivator BZLF1 is mediated by protein-protein interactions. *Proc Natl Acad Sci U S A* 92, 12265-12269.

Preston, V. G., Davison, A. J., Marsden, H. S., Timbury, M. C., Subak-Sharpe, J. H., and Wilkie, N. M. (1978). Recombinants between herpes simplex virus types 1 and 2: analyses of genome structures and expression of immediate early polypeptides. *J Virol* 28, 499-517.

Rodriguez, A., Armstrong, M., Dwyer, D., and Flemington, E. (1999). Genetic dissection of cell growth arrest functions mediated by the Epstein-Barr virus lytic gene product, Zta. *J Virol* 73, 9029-9038.

Rooney, C., Taylor, N., Countryman, J., Jenson, H., Kolman, J., and Miller, G. (1988). Genome rearrangements activate the Epstein-Barr virus gene whose product disrupts latency. *Proc Natl Acad Sci U S A* 85, 9801-9805.

Sarisky, R. T., Gao, Z., Lieberman, P. M., Fixman, E. D., Hayward, G. S., and Hayward, S. D. (1996). A replication function associated with the activation domain of the Epstein-Barr virus Zta transactivator. *J Virol* 70, 8340-8347.

Schooley, R. T. (1988). Chronic fatigue syndrome: a manifestation of Epstein-Barr virus infection? *Curr Clin Top Infect Dis* 9, 126-146.

Shapiro-Shelef, M., Lin, K. I., McHeyzer-Williams, L. J., Liao, J., McHeyzer-Williams, M. G., and Calame, K. (2003). Blimp-1 is required for the formation of immunoglobulin secreting plasma cells and pre-plasma memory B cells. *Immunity* 19, 607-620.

Sinclair, A. J. (2003). bZIP proteins of human gammaherpesviruses. *J Gen Virol* 84, 1941-1949.

Sista, N. D., Pagano, J. S., Liao, W., and Kenney, S. (1993). Retinoic acid is a negative regulator of the Epstein-Barr virus protein (BZLF1) that mediates disruption of latent infection. *Proc Natl Acad Sci U S A* 90, 3894-3898.

Sista, N. D., Barry, C., Sampson, K., and Pagano, J. (1995). Physical and functional interaction of the Epstein-Barr virus BZLF1 transactivator with the retinoic acid receptors RAR alpha and RXR alpha. *Nucleic Acids Res* 23, 1729-1736.

Staeger, M. S., Lee, S. P., Frisan, T., Mautner, J., Scholz, S., Pajic, A., Rickinson, A. B., Masucci, M. G., Polack, A., and Bornkamm, G. W. (2002). MYC overexpression imposes a nonimmunogenic phenotype on Epstein-Barr virus-infected B cells. *Proc Natl Acad Sci U S A* 99, 4550-4555.

Strauss, J.H. and Strauss, E.G. (2002). *Viruses and human disease*.

Sun, C. C., and Thorley-Lawson D.A. (2007). Plasma cell-specific transcription factor XBP-1s binds to and transactivates the Epstein-Barr virus BZLF1 promoter. *J Virol* 81, 13566-13577.

Takada, K., Shimizu, N., Sakuma, S., and Ono, Y. (1986). Trans activation of the latent Epstein-Barr virus (EBV) genome after transfection of the EBV DNA fragment. *J Virol* 57, 1016-1022.

Takagi, S., Takada, K., and Sairenji, T. (1991). Formation of intranuclear replication compartments of Epstein-Barr virus with redistribution of BZLF1 and BMRF1 gene products. *Virology* 185, 309-315.

Thomas, C., Dankesreiter, A., Wolf, H., and Schwarzmann, F. (2003). The BZLF1 promoter of Epstein-Barr virus is controlled by E box-/HI-motif-binding factors during virus latency. *J Gen Virol* 84, 959-964.

Wang, Z. P., Qing, D. S., Shi, B. G., Chen, Y. R., and Liu, G. D. (1998). Modulating effect of mitomycin or cisplatin on lymphokine-activated killer cell proliferation and antitumor activity to bladder cancer cell lines in vitro. *Zhongguo Yao Li Xue Bao* 19, 369-372.

Wang, P., Day, L., Dheekollu, J., and Lieberman, P. M. (2005). A redox-sensitive cysteine in Zta is required for Epstein-Barr virus lytic cycle DNA replication. *J Virol* 79, 13298-13309.

Watson, R. J., and Clements, J. B. (1978). Virus transcript mapping studies in cells infected with temperature-sensitive mutants of herpes simplex virus type 1. *IARC Sci Publ*, 313-326.

Wei, W. I., and Sham, J. S. (2005). Nasopharyngeal carcinoma. *Lancet* 365, 2041-2054.

Wu, F. Y., Chen, H., Wang, S. E., ApRhys, C. M., Liao, G., Fujimuro, M., Farrell, C. J., Huang, J., Hayward, S. D., and Hayward, G. S. (2003). CCAAT/enhancer binding protein alpha interacts with ZTA and mediates ZTA-induced p21(CIP-1) accumulation and G(1) cell cycle arrest during the Epstein-Barr virus lytic cycle. *J Virol* 77, 1481-1500.

Wu, F. Y., Wang, S. E., Chen, H., Wang, L., Hayward, S. D., and Hayward, G. S. (2004). CCAAT/enhancer binding protein alpha binds to the Epstein-Barr virus (EBV) ZTA protein through oligomeric interactions and contributes to cooperative transcriptional activation of the ZTA promoter through direct binding to the ZII and ZIIIB motifs during induction of the EBV lytic cycle. *J Virol* 78, 4847-4865.

Yasui, Y., Potter, J. D., Stanford, J. L., Rossing, M. A., Winget, M. D., Bronner, M., and Daling, J. (2001). Breast cancer risk and "delayed" primary Epstein-Barr virus infection. *Cancer Epidemiol Biomarkers Prev* 10, 9-16.

Yin, Q., Jupiter, K., and Flemington, E. K. (2004). The Epstein-Barr virus transactivator Zta binds to its own promoter and is required for full promoter activity during anti-Ig and TGF-beta1 mediated reactivation. *Virology* 327, 134-143.

Yoshizaki, T., Sato, H., Murono, S., Pagano, J. S., and Furukawa, M. (1999). Matrix metalloproteinase 9 is induced by the Epstein-Barr virus BZLF1 transactivator. *Clin Exp Metastasis* 17, 431-436.

Zalani, S., Holley-Guthrie, E., and Kenney, S. (1996). Epstein-Barr viral latency is disrupted by the immediate-early BRLF1 protein through a cell-specific mechanism. *Proc Natl Acad Sci U S A* 93, 9194-9199.

Zerby, D., Chen, C. J., Poon, E., Lee, D., Shiekhattar, R., and Lieberman, P. M. (1999). The amino-terminal C/H1 domain of CREB binding protein mediates zta transcriptional activation of latent Epstein-Barr virus. *Mol Cell Biol* 19, 1617-1626.

Zhang, Q., Gutsch, D., and Kenney, S. (1994). Functional and physical interaction between p53 and BZLF1: implications for Epstein-Barr virus latency. *Mol Cell Biol* 14, 1929-1938.

Zhang, Q., Hong, Y., Dorsky, D., Holley-Guthrie, E., Zalani, S., Elshiekh, N. A., Kiehl, A., Le, T., and Kenney, S. (1996). Functional and physical interactions between the Epstein-Barr virus (EBV) proteins BZLF1 and BMRF1: Effects on EBV transcription and lytic replication. *J Virol* 70, 5131-5142.

Le virus Epstein-Barr (EBV) est un γ -Herpesvirus infectant plus de 95 % de la population mondiale. Comme les autres virus de cette famille, il provoque une infection aiguë souvent asymptomatique, puis persiste à l'état latent chez l'individu infecté. Il peut cependant être réactivé et entrer en phase lytique où il provoque la mort des cellules infectées. Le facteur de transcription viral ZEBRA (également appelé BZLF1, Zta, Z ou EB1) est responsable de la transition entre ces deux phases. ZEBRA est une protéine à « leucine-zipper » de la famille des protéines bZIP. Elle active les promoteurs des gènes lytiques de l'EBV en se fixant sur des sites ADN spécifiques appelés sites ZREs. ZEBRA fixe préférentiellement certains sites ZREs lorsqu'ils sont méthylés sur leurs motifs CpG et notamment le site cible ZRE2 du promoteur viral du gène précoce *BRLF1*. Cette capacité particulière est unique à ZEBRA parmi les autres membres de la famille bZIP et s'avère critique pour l'activation du cycle lytique étant donné que le génome de l'EBV est intensivement méthylé durant la phase de latence. Nous avons résolu la structure cristallographique du domaine bZIP de ZEBRA en complexe avec le site ZRE2 méthylé. L'analyse structurale corrélée à une étude de mutagenèse et à la détermination des affinités de ZEBRA pour ses sites ADN cibles, nous permet de proposer une hypothèse quant au mécanisme de fixation préférentielle de l'ADN méthylé par le facteur de transcription ZEBRA. En parallèle, nous avons débuté un criblage à haut débit de composés chimiques pouvant inhiber la fixation à l'ADN de ZEBRA et par conséquent, l'induction du cycle lytique. Les résultats préliminaires ont permis d'identifier des composés prometteurs ayant un pouvoir inhibiteur modéré mais spécifiques de ZEBRA. Ce travail améliore considérablement notre compréhension du mécanisme d'induction du cycle lytique par ZEBRA et aidera potentiellement au développement de nouveaux agents thérapeutiques contre les pathologies associées à l'EBV.

Epstein-Barr virus is a γ -herpesvirus that infects over 95% of the world population. Like other herpesviruses, EBV causes an acute, usually asymptomatic, infection that subsequently persists in the infected individual in a latent state. The virus periodically reactivates to undergo lytic replication, causing death of the infected cells. The viral transcription factor ZEBRA (BZLF1, Zta, Z, EB1) is responsible for switching infection from the latent to the lytic phase. ZEBRA, a member of the basic region leucine zipper (bZIP) protein family, activates the promoters of lytic EBV genes by recognizing specific DNA sites called ZEBRA-responsive elements (ZREs). ZEBRA binds preferentially to certain ZREs when these are methylated on CpG motifs, including a key target site (ZRE2) within the promoter of the immediate-early gene *BRLF1*. This activity is unique to ZEBRA among known transcription factors and is critical for lytic cycle activation, as the EBV genome is highly methylated during latency. We determined the crystal structure of ZEBRA's DNA-binding domain in complex with a methylated ZRE2 site. Structural analysis combined with mutagenesis and binding studies suggest a detailed mechanism by which ZEBRA preferentially recognizes methylated DNA. In parallel, we developed a high throughput assay to identify small molecules that inhibit ZEBRA's DNA-binding activity. This work substantially advances our understanding of ZEBRA-mediated lytic cycle activation and will potentially aid in the development of new therapeutic agents against EBV-associated disease.