

Université de Montréal

**Étude du promoteur de ZNF74 et des séquences d'ADN reconnues
par ce répresseur transcriptionnel**

Par

Imene BENSMINA

Département de Biochimie

Faculté de médecine

Mémoire présenté à la faculté des études supérieures
en vue de l'obtention du grade de Maîtrise
en biochimie

Janvier 2006

© Imene BENSMINA, Janvier 2006



W

H

US8

2006

V-139

AVIS

L'auteur a autorisé l'Université de Montréal à reproduire et diffuser, en totalité ou en partie, par quelque moyen que ce soit et sur quelque support que ce soit, et exclusivement à des fins non lucratives d'enseignement et de recherche, des copies de ce mémoire ou de cette thèse.

L'auteur et les coauteurs le cas échéant conservent la propriété du droit d'auteur et des droits moraux qui protègent ce document. Ni la thèse ou le mémoire, ni des extraits substantiels de ce document, ne doivent être imprimés ou autrement reproduits sans l'autorisation de l'auteur.

Afin de se conformer à la Loi canadienne sur la protection des renseignements personnels, quelques formulaires secondaires, coordonnées ou signatures intégrées au texte ont pu être enlevés de ce document. Bien que cela ait pu affecter la pagination, il n'y a aucun contenu manquant.

NOTICE

The author of this thesis or dissertation has granted a nonexclusive license allowing Université de Montréal to reproduce and publish the document, in part or in whole, and in any format, solely for noncommercial educational and research purposes.

The author and co-authors if applicable retain copyright ownership and moral rights in this document. Neither the whole thesis or dissertation, nor substantial extracts from it, may be printed or otherwise reproduced without the author's permission.

In compliance with the Canadian Privacy Act some supporting forms, contact information or signatures may have been removed from the document. While this may affect the document page count, it does not represent any loss of content from the document.

Université de Montréal
Faculté des études supérieures

Ce mémoire intitulé :

**Étude du promoteur de ZNF74 et des séquences d'ADN reconnues
par ce répresseur transcriptionnel**

Présenté par :
Imene BENSMINA

A été évalué par un jury composé des personnes suivantes :

Pascal CHARTRAND
Président – rapporteur

Muriel AUBRY
Directrice de recherche

Luc DESGROSEILLES
Membre du jury

Résumé

Notre laboratoire s'intéresse à la famille des protéines à motif KRAB et multidoigt de zinc qui constitue une famille élargie de facteurs transcriptionnels potentiels. Nous étudions plus particulièrement la protéine KRAB multidoigt de zinc ZNF74. Le gène *ZNF74* qui encode cette protéine est un candidat pour un désordre du développement connu sous le nom du syndrome de DiGeorge et est exprimé de façon restreinte au niveau embryonnaire dans des structures affectées chez les patients. Quoique l'activité répressive de ZNF74 ait pu être démontrée en fusionnant son domaine KRAB à un domaine bien caractérisé de liaison à l'ADN, les études de ZNF74 sont limitées par le fait que l'on ne connaît pas les séquences spécifiques d'ADN que cette protéine reconnaît ni ses gènes cibles. Dans cette étude, la recherche des séquences d'ADN spécifiquement liées par ZNF74 a été effectuée par une méthode de sélection cyclique qui combine une purification par South-Western et une amplification par PCR. Les séquences obtenues après trois cycles ont été analysées par le logiciel bioinformatique Gibbs qui a permis la caractérisation d'un motif conservé **APuATTC/G**. Quoique la spécificité de liaison de ZNF74 à des séquences contenant ce motif ait pu être confirmée par retardement sur gel, des essais de gène rapporteur ont révélé qu'elles n'ont pas la capacité d'agir comme éléments de réponse fonctionnels. En effet, ces séquences placées en amont du site d'initiation de la transcription ne permettent pas à ZNF74 d'être recrutée efficacement au promoteur pour réprimer la transcription. Par la suite, nous avons cherché à déterminer si comme plusieurs facteurs transcriptionnels, ZNF74 utilisait des éléments *cis* de son propre promoteur pour réprimer la transcription via un mécanisme d'autorégulation transcriptionnelle. La caractérisation du promoteur par des essais de gène rapporteur a permis de déterminer que la région -399 à +97 relativement au site d'initiation de la transcription correspond au promoteur minimal et qu'un élément *cis* important pour une transcription maximale est situé entre +50 et +97. Par contre, nous n'avons pas pu mettre en évidence l'existence d'un mécanisme d'autorégulation du promoteur (région -1429 à +97 testée) et donc identifier d'élément *cis* reconnu par ZNF74. Cette étude témoigne de la difficulté, déjà évoquée dans la littérature, qui est rencontrée pour l'identification des éléments de réponses spécifiques à des facteurs transcriptionnels à motif multidoigt de zinc. Actuellement, l'identification de ces éléments par immunoprécipitation de la chromatine apparaît comme la meilleure approche possible pour ce type de protéine.

Mots clés :

Syndrome de DiGeorge, protéine KRAB multidoigt de zinc, ZNF74, liaison à l'ADN, éléments de réponse, sélection par South-Western, caractérisation du promoteur, autorégulation transcriptionnelle.

Summary

Our laboratory is interested in the family of KRAB zinc finger proteins, which constitutes a larger family of potential transcriptional factors. This study is focusing more specifically on the KRAB zinc finger protein ZNF74. The gene, which encodes this protein, *ZNF74*, is a candidate for DiGeorge developmental disorder and is embryonically expressed in structures affected in patients. Although ZNF74 was shown to repress transcription when fused with GAL4 DNA binding domain, the studies of ZNF74 are limited by the fact that neither its specific DNA sequences nor its target genes are known. In this study, a search for the DNA sequences specifically recognized by ZNF74 was carried out by a cyclic selection method, which combines purification by South-Western and amplification by PCR. The sequences obtained after three cycles of purification were analyzed by the bioinformatic software Gibbs. This allowed the identification of the conserved motif **APuATTTC/G**. Although the binding specificity of ZNF74 to sequences containing this motif could be confirmed by gel shift assays, gene reporter assays revealed that sequences corresponding to this consensus are not functioning as efficient transcriptional response elements and no repression was observed in the presence of ZNF74. Thereafter, we determine whether ZNF74 might as several transcriptional factors use elements of its promoter to repress its own expression via an auto-regulatory mechanism of transcriptional control. First, a characterization of ZNF74 promoter was performed using a gene reporter assay. This allowed us to determine that the area -399 to +97 relative to the transcriptional start site correspond to the minimal promoter and that a *cis* element important for maximal transcription is located between +50 and +97. Second, co-transfection of promoter fragments and of a ZNF74 expressing construct revealed that ZNF74 is not able to repress its own expression (area of the promoter tested -1429 to +97). This study points to the difficulty inherent to the identification of response element for transcription factors with multifinger as already suggested in few publications. Currently, it seems that the best alternative approach for identifying response elements for multifinger proteins would be chromatin immunoprecipitation and cloning of the chromatin bound sequences.

Key words: DiGeorge syndrome, KRAB zinc finger protein, ZNF74, DNA binding, response element, South-Western selection, characterization of promoter, transcriptional autoregulation.

Table des matières

Chapitre I. Introduction	1
1.1 - La famille des KRAB à multidoigt de zinc	2
1.1.1 - La structure des protéines KRAB à multidoigt de zinc	3
1.1.2 - Le modèle de répression	14
1.1.3 - Le rôle des protéines KRAB à multidoigt de zinc	17
1.2 - ZNF74, une protéine KRAB multidoigt de zinc.	19
1.2.1 - Le gène <i>ZNF74</i>	19
1.2.2 - La protéine ZNF74	20
1.3 - ZNF74, et le syndrome de DiGeorge	19
1.3.1 - Le syndrome de DiGeorge	22
1.3.2 - ZNF74, candidat pour le syndrome de DiGeorge	24
1.4 - Objectifs de recherche	25
Chapitre II. Matériel et méthodes	26
2.1 – Vecteurs	26
2.1.1 - Vecteurs d'expression protéique	27
2.1.2 - Vecteurs de gène rapporteur Luciférase.....	28
2.2- Expression protéique	32
2.2.1 - Protéines d'origine bactérienne	32
2.2.2 - Expression protéique dans les cellules de mammifères	35
2.3 - Expérience de purification par South-Western	37
2.3.1- Préparation des oligonucléotides dégénérés double brin	38
2.3.2- Détermination de l'activité spécifique des oligonucléotides marqués et rendement de marquage	39
2.3.3- Purification des oligonucléotides marqués sur colonne Sephadex	40
2.3.4- Expérience de purification par South-Western	40
2.3.5- Amplification des oligonucléotides élus	41
2.4 - Caractérisation des oligonucléotides sélectionnés	44
2.4.1- Clonage des oligonucléotides sélectionnés	44
2.4.2- Séquençage et analyse des clones	44
2.5- Étude par retardement sur gel de la liaison de ZNF74 aux oligonucléotides sélectionnés	45

2.5.1 - Préparation des oligonucléotides	45
2.5.2 - Expérience de retardement sur gel	45
2.6 - Essais transcriptionnels de gène rapporteur	46
2.6.1 - Détermination de l'activité transcriptionnelle du gène <i>luciférase</i>	46
2.6.2 - Vérification de l'expression des protéines	47
Chapitre III. Résultats	48
3.1 - Recherche de l'élément de réponse de ZNF74.	49
3.1.1 - Sélection d'oligonucléotides dégénérés via purification par South-Western combinée à une amplification par PCR	49
3.1.2 - Analyse des séquences des oligonucléotides sélectionnés	55
3.1.3 - Test de la spécificité de liaison de ZNF74 aux oligonucléotides sélectionnés par retardement sur gel	58
3.1.4 - Test de la fonction du motif identifié comme élément de réponse pour ZNF74 par essai transcriptionnel	63
3.2 - Recherche de l'élément de réponse de ZNF74 dans son propre promoteur .70	
3.2.1 - Analyse de la séquence du promoteur de ZNF74	70
3.2.2 – Analyse fonctionnelle du promoteur	73
3.2.3 – Recherche de séquences cibles dans le promoteur de ZNF74	77
Chapitre IV. Discussion	83
4.1 - À la recherche d'éléments de réponse de ZNF74	85
4.1.1 - Choix de la stratégie de recherche	85
4.1.2 - Résultats de la recherche	86
4.1.3 - Les causes probables des difficultés d'identification de l'élément de réponse pour ZNF74	88
4.1.4 – Recherche d'éléments de réponse de ZNF74 dans la séquence de son propre promoteur	90
4.1.5 – L'immunoprécipitation de la chromatine, l'autre alternative	92
4.2 – Conclusion	93
Bibliographie	95

Liste des tableaux

Tableau I. Résumé de toutes les protéines KRAB multidoigt de zinc pour lesquelles les éléments de réponse ont été caractérisés.	8
Tableau II. Exemple de protéines multidoigt de zinc sans motif KRAB pour lesquelles les éléments de réponse ont été caractérisés.	11
Tableau III. Les constructions des trois oligonucléotides purifiés par South-Western dans le vecteur pGL3 <i>contrôle</i>	29
Tableau IV. Stratégie de clonage des différents fragments du promoteur de ZNF74.....	31

Liste des figures

Figure 1. Structure du domaine doigt de zinc C_2H_2	4
Figure 2. Principe de base de la technique «Target detection assay» (TDA)	6
Figure 3. Organisation structurale des protéines à motifs KRAB et multidoigt de zinc	13
Figure 4. Représentation schématique de la structure de la protéine KAP-1	15
Figure 5. Mécanisme de répression transcriptionnelle par les protéines KRAB à multidoigt de zinc et KAP-1	16
Figure 6. Organisation du gène ZNF74	20
Figure 7. Représentation schématique de la structure primaire des deux isoformes de la protéine ZNF74	22
Figure 8. Représentation schématique de la région chromosomique 22q11.2	24
Figure 9. Représentation schématique des différentes constructions préparées à partir du promoteur de ZNF74	31
Figure 10. Schéma général de l'expérience de purification par South-Western.	39
Figure 11. Représentation schématique du protocole expérimental de l'expérience de purification par South-Western	43
Figure 12. Analyse des protéines de fusion bactériennes MBP, MBP-Zn et MBP-ZNF74-II (Δ 1-176) par électrophorèse sur gel de SDS-polyacrylamide	51
Figure 13. Dosage des protéines de fusion bactériennes MBP, MBP-Zn et MBP-ZNF74-II (Δ 1-176)	52
Figure 14. Les protéines MBP-Zn et MBP-ZNF74-II (Δ 1-176) lient les oligonucléotides doubles brins dégénérés en présence du zinc dans l'expérience de purification par South-Western	54
Figure 15. Analyse des séquences des 35 clones purifiés par le logiciel Gibbs	56
Figure 16. Dosage des protéines bactériennes MBP, MBP-Zn et MBP-ZNF74-II présentes dans les extraits utilisés dans les expériences de retardement sur gel	60
Figure 17. Essai de retardement sur gel de la protéine MBP-ZNF74-II (Δ 1-176) avec les oligonucléotides sélectionnés (98, 76 et 67)	61

- Figure 18.** Essai de retardement sur gel de la protéine MBP-ZNF74-II (Δ 1-176) avec les oligonucléotides sélectionnés (72, 99,73 et 31)62
- Figure 19.** L'effet de pCGN-ZNF74-II sur la transcription médiée par la séquence 16N du clone 76 multimérisée 4 fois en amont du promoteur de SV4066
- Figure 20.** L'effet de pCGN-ZNF74-II sur la transcription médiée par la séquence 16N du clone 31 ou du clone 73 clonée en amont du promoteur de SV4068
- Figure 21.** Analyse de la séquence du promoteur de *ZNF74* par le logiciel Match72
- Figure 22.** Analyse de l'activité transcriptionnelle du promoteur de *ZNF74*75
- Figure 23.** L'effet de pCGN-ZNF74-II sur la transcription médiée par la région -1429/+49 de son propre promoteur79
- Figure 24.** L'effet de pCGN-ZNF74-II sur la transcription médiée par les régions -399/+49 et -399/+97 de son propre promoteur81

Listes des abréviations

ADN	Acide désoxyribonucléique
ARN	Acide ribonucléique
BSA	«Bovine serum albumin»
CAFS	«Conotruncal anomaly face syndrome»
CPM	Comptes par minute
del22q11	Le syndrome de délétion 22q11.2
DMEM	«Delbecco's modified eagle's medium»
dNTP	Desoxy nucleotide triphosphate
EDTA	Acide éthylène diamine tétra acétique
FBS	Sérum foetal bovin
GAL4DB	GAL4 DNA binding domain
H3K9	La lysine 9 de l'histone
HA	Hémagglutinine
HPIBD	«HP1 binding domain»
HRP	«Horsradish peroxydase»
IPTG	L'isopropyl- β -D-thiogalactopyranoside
KAP1	«KRAB/multifinger associated protein 1»
Kb	Kilobases
KLF4	«Krüppel-like factor 4»
KRAB	«Krüppel associated box»
KRIP1	«KRAB interacting protein 1»
KS1	«KRAB suppressor of transformation 1»
LB	«Luria Broth»

LeR	«Leucin rich»
MBP	«Maltose binding protein»
MZF1	«Myeloid zinc finger protein1»
NRIF	«Neurotrophin receptor interacting factor»
NuRD	«Nucleosome remodelling and desacetylation complex»
OSE2	«Osteoblast-specific «AACCACA» enhancer element 2»
PBS	«Phosphate buffered saline»
PHD	«Plant-like homeodomain»
PCR	«Polymerase chain reaction»
RBCC	«RING finger B boxes coiled-coil»
RACE	«Rapid amplification of cDNA ends»
SAAB	«Selected and amplified binding sites imprint assay»
SDS	«Sodium dodecyl sulfate»
TDA	«Target detection assay»
TFIIIA	Le facteur transcriptionnel III A
TIF1 β	«Transcriptional intermediary protein 1 β »
VCFS	Le syndrome vélocardio-facial
WT1	«Wilms tumor suppressor 1»
ZNF	«Zinc finger»

À mon père, mon premier maître, celui qui m'a inculqué les valeurs du savoir et le respect de ceux qui les transmettent, m'a guidé et m'a encouragé tout au long de mes études.

À mon époux qui m'était d'une aide inestimable.

À mes filles, **Meriem et Hajar**, lumière de ma vie.

Remerciements :

Au nom d'Allah clément et miséricordieux

« Gloire à toi, dirent-ils, nous n'avons de connaissance que celle que tu nous a enseignée. Car en vérité, tu es le connaissant et le sage » Coran II, 32.

Je tiens à exprimer mon remerciement envers Muriel AUBRY la personne pour son soutien, sa générosité et sa compréhension. En tant que directrice de recherche, je lui dois toute ma reconnaissance de m'avoir accueillie dans son laboratoire et de m'avoir initiée au monde de la recherche avec ses discussions, ses consignes et ses précieux conseils. Merci beaucoup Muriel.

Je remercie tous les membres de notre laboratoire pour leur sympathie.

CHAPITRE I
INTRODUCTION

1.1 - La famille des protéines KRAB à multidoigt de zinc :

L'ensemble des cellules spécialisées qui constituent un organisme multicellulaire possèdent un bagage génétique identique. La régulation de l'expression génétique est à la base de cette différenciation cellulaire. Cette régulation consiste, selon le type cellulaire, à activer l'expression de certains gènes et à en réprimer d'autres. Elle est médiée en partie par des facteurs transcriptionnels qui lient l'ADN de façon spécifique et qui modifient par la suite le niveau transcriptionnel des gènes cibles en les activant (activateurs) ou en les réprimant (répresseurs).

Les facteurs transcriptionnels ont été subdivisés en plusieurs familles selon leur motif de liaison à l'ADN (Pabo *et al.*, 1992) dont l'une d'elles est la famille des protéines à motif doigt de zinc de type C₂H₂ («Krüppel-related»). C'est une famille élargie de facteurs transcriptionnels potentiels (Bellefroid *et al.*, 1989) dont les membres contiennent deux ou plusieurs doigts de zinc de type C₂H₂ en tandem. Ce motif a été identifié pour la première fois chez *Xenopus* au niveau du Facteur Transcriptionnel III A (TFIIIA) (Miller *et al.*, 1985). D'autres protéines ont été par la suite caractérisées chez plusieurs espèces incluant les vertébrés et les invertébrés. Chez la drosophile, le gène de segmentation *Krüppel* (Schuh *et al.*, 1986) qui encode des motifs doigts de zinc est à l'origine de la désignation des protéines à motifs C₂H₂ comme étant de type Krüppel.

Dans le génome humain, on dénombre plus de 700 gènes encodant des protéines à motif doigt de zinc de type TFIIIA/Krüppel. Le tiers de ces protéines possède en plus du domaine C₂H₂ un autre domaine conservé nommé KRAB (**K**rüppel **A**ssociated **B**ox) (Bellefroid *et al.*, 1991; Crossley *et al.*, 1991). Ce motif est spécifique aux vertébrés.

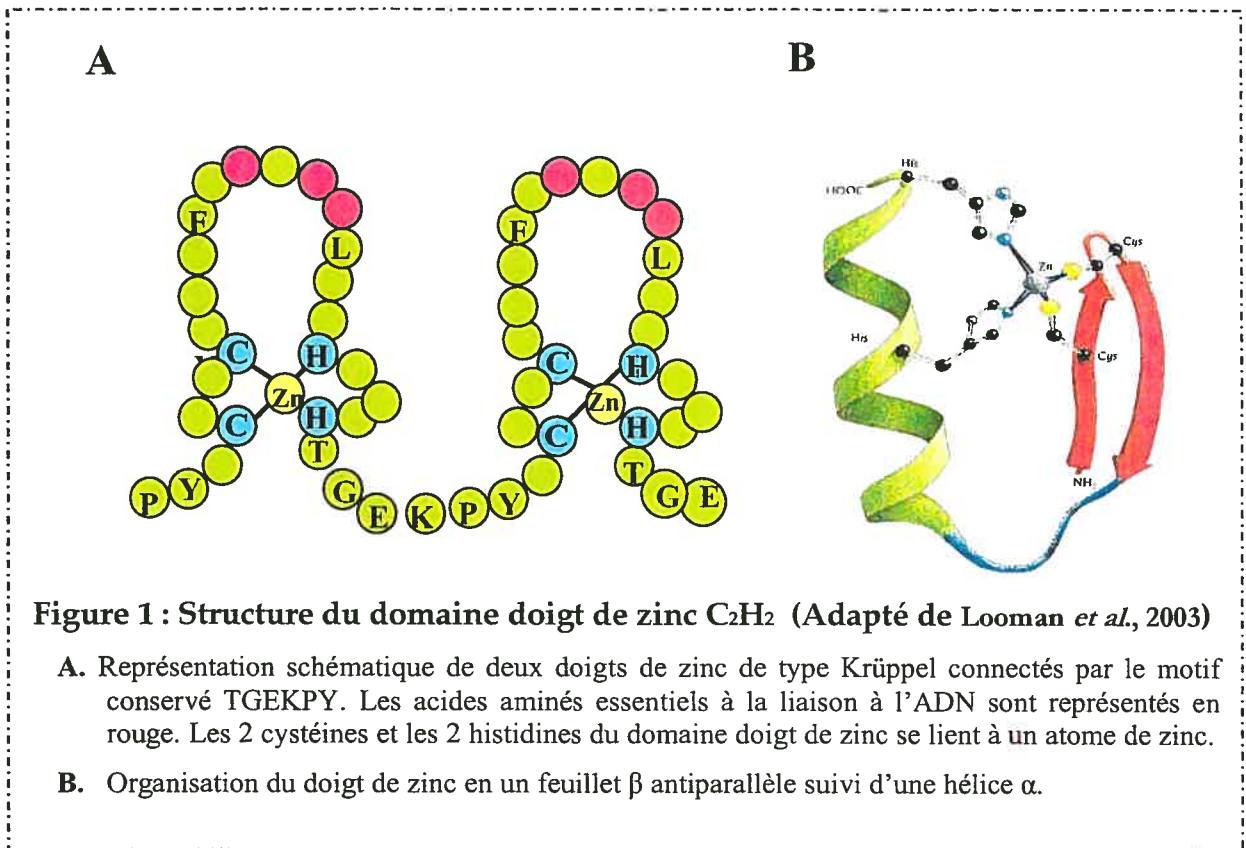
Dans la grande majorité des cas, la présence d'un motif KRAB est associée à plus de cinq motifs C_2H_2 positionnés en tandem. Les protéines KRAB à multidoigt de zinc fonctionnent comme des répresseurs transcriptionnels. Elles se lient à l'ADN par leur domaine multidoigt de zinc tandis que la répression est assurée par le domaine KRAB qui recrute d'autres complexes protéiques incluant des histones desacétylases, des histones méthylases et des protéines de l'hétérochromatine (Schultz *et al.*, 2001; Schultz *et al.*, 2002; Ryan *et al.*, 1999; Lechner *et al.*, 2000)

1.1.1 – La structure des protéines KRAB à multidoigt de zinc :

Tel que mentionné ci-dessus, les protéines KRAB à multidoigt de zinc possèdent une structure conservée avec deux domaines fonctionnels, le domaine doigt de zinc et le domaine KRAB.

1.1.1.1- Domaine doigt de zinc :

Le domaine doigt de zinc de type C_2H_2 est une séquence de 21-23 acides aminés, soit **CX₂₋₄CX₃FX₅LX₂HX₃₋₄H**, séparée par un motif conservé de 7 acides aminés **TGEKPYX** (X désigne un acide aminé non conservé). En présence de zinc, les 2 Cystéines et les 2 Histidines de chaque motif doigt de zinc complexent un atome de zinc (**Figure 1**). Les études cristallographiques des protéines à motif C_2H_2 telles Zif268/Egr1 (Pavletich *et al* 1991) et GLI-1 (Pavletich *et al* 1993) indiquent que le doigt de zinc est organisé en un feuillet β antiparallèle suivi d'une hélice α . Les acides aminés exposés à la surface de l'hélice α participent à la liaison avec l'ADN. Chaque doigt de zinc reconnaît potentiellement une séquence de 3 nucléotides au sein du sillon majeur de l'hélice d'ADN.



En plus de sa propriété commune de liaison à l'ADN, certains domaines doigt de zinc peuvent lier l'ARN, comme c'est l'exemple pour TFIIIA (Theunissen *et al.*, 1992), WT1 (Wilms Tumor suppressor 1) (Caricasole *et al.*, 1996) et ZNF74 (Grondin *et al.*, 1996). La liaison aux protéines par le biais de ce motif doigt de zinc a été aussi rapportée pour les protéines KRAB à multidoigt de zinc ZBRK1 (Zheng *et al.*, 2000) et ZNF74 (Grondin *et al.*, 1997).

Malgré l'abondance des protéines à motif KRAB et multidoigt de zinc, les systèmes biologiques qu'elles régulent restent peu connus en partie à cause de la difficulté d'identifier les séquences d'ADN reconnues par des protéines ayant des doigts de zinc multiples. Il y a encore peu de protéines KRAB à multidoigt de zinc pour lesquelles les séquences spécifiques d'ADN ont été identifiées (**Tableau I**).

De plus, à l'heure actuelle, l'identité d'un gène cible *in vivo* n'a été confirmée que pour la protéine NT2 (Tanaka *et al.*, 2002)

La majorité des études qui ont permis d'identifier des éléments de réponse ont utilisé un même principe de base, celui de la technique décrite en 1990 par Thiesen et Bach qu'ils ont nommée TDA (**T**arget **D**etection **A**ssay) (Thiesen et Bach., 1990). Cet essai consiste à effectuer plusieurs cycles de sélection et amplification par PCR d'oligonucléotides doubles brins dégénérés en N positions après leur liaison à la protéine d'intérêt. À chaque extrémité, ces oligonucléotides sont délimités par des séquences spécifiques qui permettent leur amplification par PCR. Chaque cycle de cet essai comporte quatre étapes principales (**Figure 2**): 1- L'incubation de la protéine d'intérêt avec l'ADN; 2- L'isolation du complexe ADN-protéine; 3- L'éluion de l'ADN lié ; 4- L'amplification par PCR de l'ADN élué.

Durant les premiers cycles de l'essai, l'isolation du complexe ADN-protéine peut être réalisée par liaison à un filtre de nitrocellulose. Les oligonucléotides sélectionnés à la fin de ces cycles peuvent être clonés directement comme ils peuvent servir à réaliser d'autres cycles durant lesquels les complexes ADN-protéine pourront être isolés par retardement sur gel. À la fin de l'essai, les oligonucléotides sont clonés, séquencés et analysés afin d'identifier une séquence consensus d'ADN reconnue spécifiquement par la protéine d'intérêt.

«Target Detection Assay» (TDA)

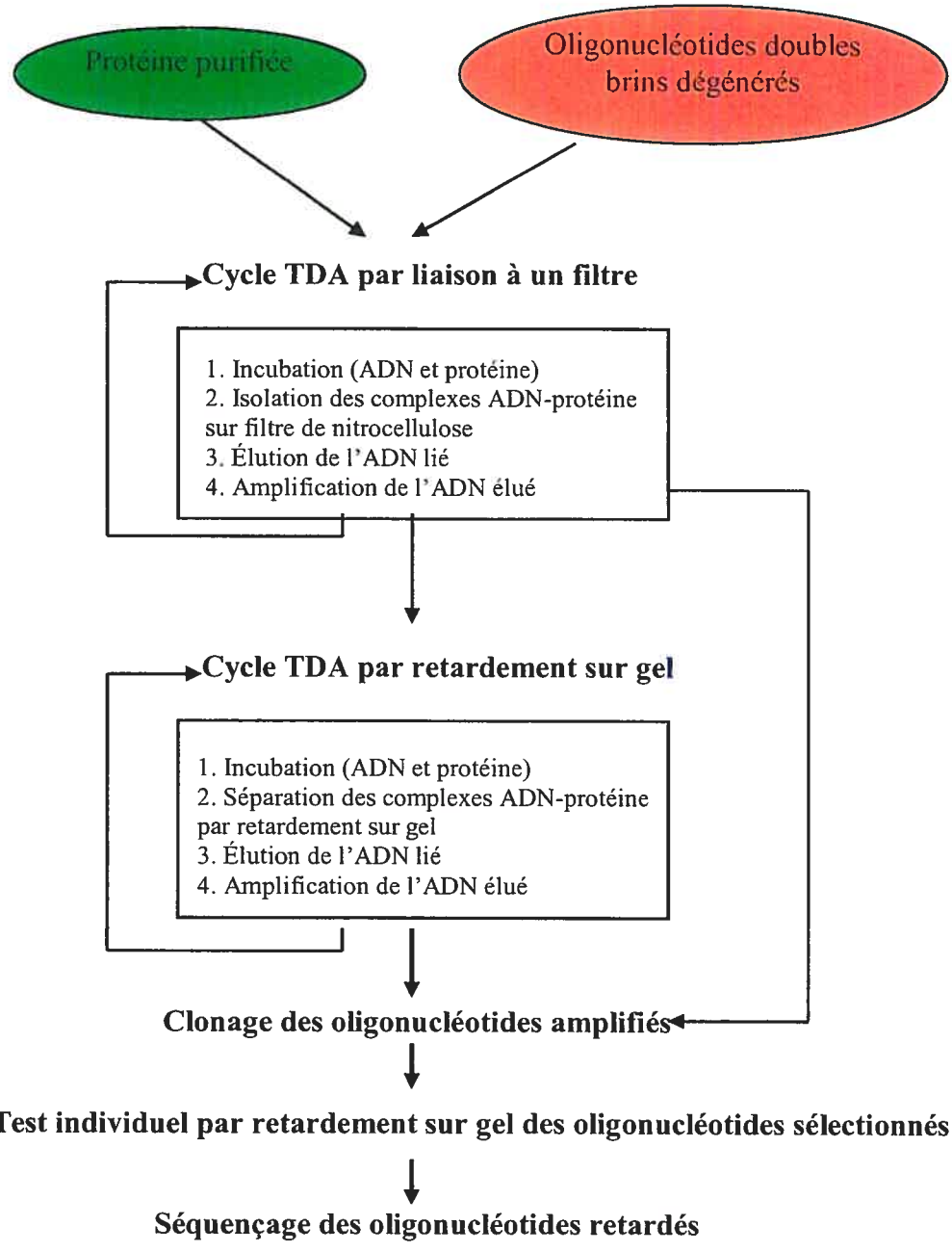


Figure 2: Principe de base de la technique «Target Detection Assay» (TDA) (Adapté de Thiesen et Bach., 1990)

Dans la même année, un autre groupe de chercheurs Blackwell et Weintraub décrivent une méthode similaire qu'ils ont dénommée «SAAB imprint assay» (Selected And Amplified Binding sites imprint assay) (Blackwell et Weintraub., 1990). Ils utilisent également des oligonucléotides dégénérés mais l'étape de sélection se fait uniquement par retardement sur gel. L'essai a été développé initialement dans le but de déterminer les séquences d'ADN reconnues par les protéines bHLH (basic Helix-Loop-Helix) qui entourent leur site consensus traditionnel CANNTG.





Par la suite, d'autres auteurs ont utilisé des adaptations des méthodes de Thiesen et Bach ou de Blackwell et Weintraub en effectuant la sélection par purification sur colonne d'affinité (exemple : colonne de glutathione) ou par buvardage South-Western (**Tableau I**).

La longueur des séquences consensus identifiées pour ces protéines KRAB à multidoigt de zinc (**Tableau I**) qui varie de façon indépendante du nombre des motifs doigts de zinc suggère que : a) ce ne sont pas tous les doigts qui lient l'ADN et b) ce ne sont pas tous les doigts de zinc qui participent à la spécificité de cette liaison. Il n'est pas exclu non plus que des protéines KRAB à multidoigt de zinc puissent faire appel à des combinaisons spécifiques de doigts de zinc selon le contexte à l'instar de la protéine multidoigt de zinc CTCF qui utilise des combinaisons différentes de ces 11 doigts de zinc pour se lier à ses gènes cibles (Ohlsson *et al.*, 2001) (**Tableau II**). La protéine CTCF est en faite considérée comme le premier facteur nucléaire polyvalent capable de reconnaître des séquences d'ADN variées au sein de plusieurs gènes.




Certains facteurs transcriptionnels en plus de lier des séquences spécifiques au sein de leurs gènes cibles sont capables de lier des séquences situées dans leurs propres promoteurs et d'exercer par la suite une autorégulation de leur transcription. Ce mécanisme de régulation a été observé pour plusieurs protéines. On peut citer comme exemple Barx2,

un facteur transcriptionnel à homéodomaine qui participe au remodelage du follicule pileux en réprimant la transcription de son propre gène (Olson *et al.*, 2005), KLF4 (**K**rüppel-**L**ike **F**actor **4**), un régulateur négatif de la prolifération cellulaire qui active sa transcription en se liant à trois boîtes GC présentes dans son propre promoteur (Mahatan *et al.*, 1999) et Gfi1b, une protéine nucléaire de 37 kDa qui possède 6 doigts de zinc en tandem de type C₂H₂ et qui est régulée par une boucle de rétroaction négative via sa liaison à son propre promoteur (Vassen *et al.*, 2005) (**Tableau II**).




TABEAU I :
Résumé de toutes les protéines KRAB à multidoigt de zinc pour lesquelles les éléments de réponse ont été caractérisés

PROTÉINES MULTIDOIGT DE ZINC À MOTIF KRAB						
Protéine	Caractéristiques structurales	Profil d'expression	Méthode utilisée et Séquence consensus identifiée	Gènes cibles	Références bibliographiques	
KS1	 <p>KRAB <i>Doigts de zinc</i></p>	Ubiquiste	Retard sur gel et PCR (<i>Oligo dégénérés N35, 7 cycles de sélection</i>) Séquence : 5'TCCTACAGTACCAACCCTACAGATAA3'	Aucun gène cible n'a été identifié	Gebelein et urrutia 2001	
SZF1	 <p>KRAB <i>Doigts de zinc</i></p>	Cellules souches CD34*	Retard sur gel et PCR (<i>Oligo dégénérés N14, 3 cycles de sélection</i>) Séquence : 5'CCAGGGTACAGCCG3' (Ressemble à la séquence consensus de ZBRK1 5'GGGXXVCAGXXTTT3')	Aucun gène cible n'a été identifié	Peng <i>et al</i> 2002	
ZNF333	 <p>KRAB KRAB <i>Doigts de zinc</i></p>	Cœur	South-Western et PCR, (<i>Oligo dégénérés N14, 5 cycles de sélection</i>) Séquence : 5'ATAAT3'	EIF3s2, Cox6A1, PPP2R4, BCL2L12 = Gènes cibles potentiels sur la base de l'identification de la séquence ATAAT dans leurs séquences. Aucun test fonctionnel effectué.	Jing <i>et al</i> 2004	
ZBRK1	 <p>KRAB <i>Doigts de zinc</i></p>	Cellules du muscle squelettique	Retard sur gel et PCR ¹ (<i>Oligo dégénérés N23, 6 cycles de sélection</i>) Séquence : 5'GGGXXVCAGXXTTT/G3'	GADD45 (gène suppresseur de tumeur) = Gène cible potentiel induit par BRCA1, un partenaire protéique de ZBRK1. Test fonctionnel : Un intron de GADD45 contenant la séquence 5'TGGXXVCAGXXTTG3' réprime la transcription d'un gène rapporteur en présence de ZBRK1.	Zheng <i>et al</i> 2000	

PROTÉINES MULTIDOIGT DE ZINC À MOTIF KRAB (Suite)

Protéine	Caractéristiques structurales	Profil d'expression	Méthode utilisée et Séquence consensus identifiée	Gènes cibles	Référence bibliographique
ZNF202	<p>SCAN KRAB Doigts de zinc</p> 	Cœur, rein, foie et testicules principalement	<p>Purification par affinité sur colonne de glutathion-agarose et PCR (<i>Oligo dégénérés N36, 5 cycles de sélection</i>)</p> <p>Séquence : 5'GGGGT3'</p>	<p>ApoAIV, ApoCIII, ApoE et HNF4 = Gènes cibles potentiels contenant dans leurs promoteurs plusieurs motifs GGGT</p> <p><u>Tests fonctionnels</u> : Les promoteurs de ces gènes sont réprimés par ZNF202 dans des essais de gènes rapporteurs</p>	Wagner <i>et al.</i> 2000
AJ18	<p>KRAB Doigts de zinc</p> 	Cellules osseuses	<p>South-western et PCR² (<i>Oligo dégénérés N12, 5 cycles de sélection</i>)</p> <p>Séquence : 5'cCACa3'</p>	<p>Ostéocalcine et ostéopontine = Gènes cibles potentiels avec des promoteurs qui contiennent la séquence 5'AACCA3' qui peut lier le transactivateur Runx2.</p> <p><u>Test fonctionnel</u>: AJ18 réprime l'activité de Runx2 dans des essais de gènes rapporteurs sous le contrôle de l'élément 5'AACCA3' répété six fois (compétition possible entre AJ18 et Runx2 pour l'élément de réponse).</p>	Jheon <i>et al.</i> 2001
NT2	<p>SCAN KRAB A Doigts de zinc</p> 	Cerveau, thymus, rate et cartilage principalement	<p>Approche de promoteur candidat : Essai direct sur le promoteur du <i>Coll1a2</i> gène en raison de son profil d'expression qui corrèle inversement avec celui de NT2 dans les chondrocytes.</p>	<p>Gène du collagène Coll1a2 = Gène cible contenant dans son promoteur la séquence régulatrice 5'CAGGAGGAGGAGGAGGCTGCT3'</p> <p><u>Tests fonctionnels</u>: NT2 réprime l'activité du promoteur de <i>Coll1a2</i> dans des essais de gènes rapporteurs.</p>	Tanaka <i>et al.</i> 2002

TABEAU II : Exemples de protéines multidoigt de zinc pour lesquelles les éléments de réponse ont été caractérisés

EXEMPLES DES PROTÉINES MULTIDOIGT DE ZINC SANS MOTIF KRAB (suite)					
Protéine	Caractéristiques structurales	Profil d'expression	Méthode utilisée et Séquence consensus identifiée	Gènes cibles	Références bibliographiques
MZF1 (Myeloid zinc finger protein1)	 <p>SCAN</p> <p>Doigts de zinc</p>	Cellules souches hématopoïétiques	Retard sur gel et PCR (<i>Oligonucleotides dérogés N14 et N20, 2 et 5 cycles de sélection</i>) Doigts de zinc 1 à 4: 5'AGTGGGA3' Doigts de zinc 5 à 13: 5'CGGNGAGGGGAA3' = Séquence riche en GA	<p>N-cadhérine =</p> <p>Gène cible contenant la séquence 5'GGGTGGGGAGGGGG3' riche en GA au niveau de son promoteur</p> <p>Test fonctionnel : dans les ostéoblastes, MZF1 active la transcription du gène rapporteur sous le contrôle du promoteur de la N-cadhérine</p>	Morris <i>et al</i> 1994 Le Mété <i>et al</i> 2005
CTCF	 <p>Doigts de zinc</p>	Ubiquiste		<p>MYC-FpV, FI silen, FII insul, MYC A, MYC B, MYC N, MYCW... etc</p> <p>Gènes cibles : La multiplicité des complexes ADN-CTCF caractérisés peut être expliquée par l'utilisation différentielle des 11 doigts de zinc de CTCF.</p>	Ohlsson <i>et al</i> 2001
Gfi 1b	 <p>Doigts de zinc</p>	Thymocytes CD4 ⁺ et CD8 ⁺	Approche de promoteur candidat : Essai direct sur le promoteur de Gfi1b	<p>Gfi1b =</p> <p>Agit comme son propre gène cible. Gfi1b se lie à son promoteur qui contient 2 répétitions du site de liaison de Gfi1b «AATC core»</p> <p>Test fonctionnel : Gfi1b autorégule son propre gène <i>in vivo</i> (en utilisant des souris transgéniques)</p>	Vassen <i>et al</i> 2005

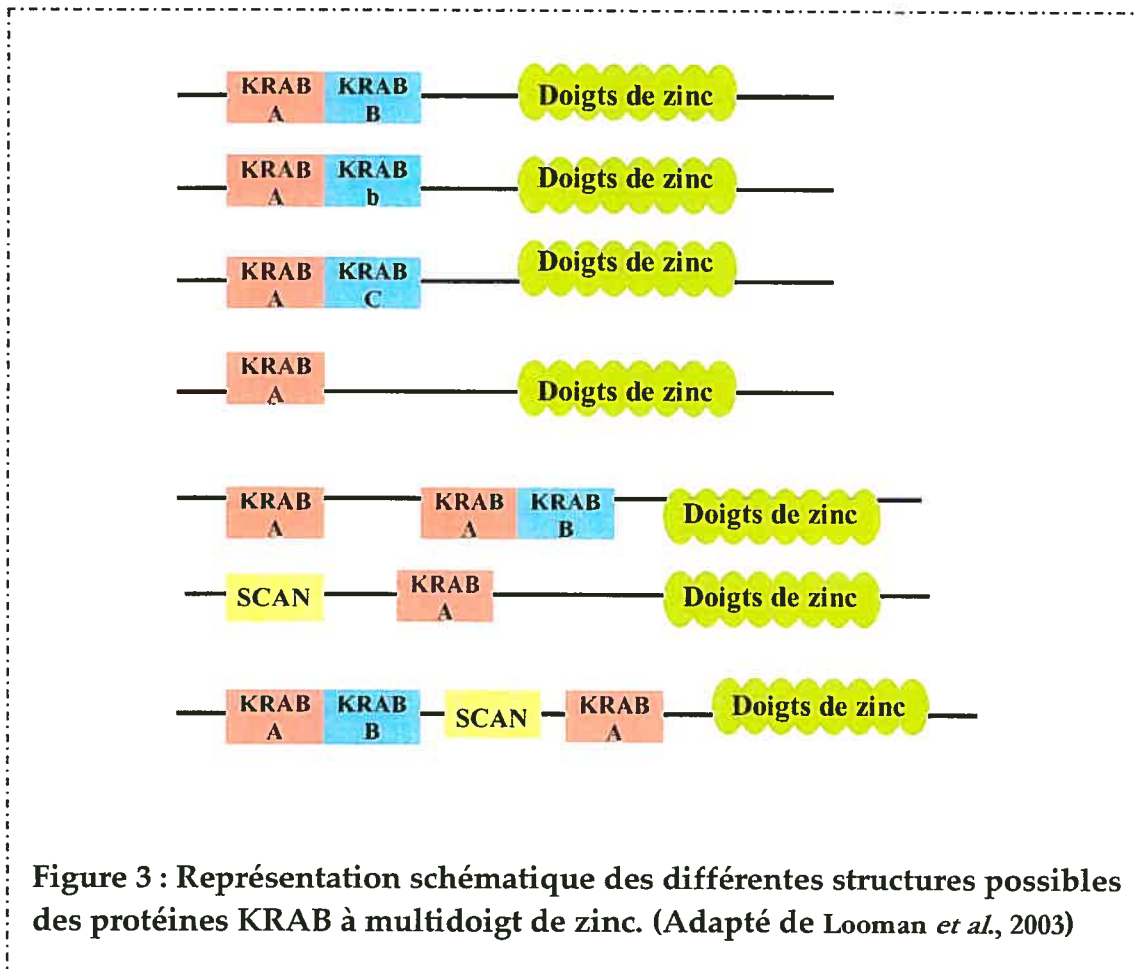
1.1.1.2 - Domaine KRAB :

Ce domaine a été identifié initialement comme un motif conservé de 75 acides aminés présent à l'extrémité N terminale des protéines à motif doigt de zinc Type C₂H₂ (Bellefroid *et al.*, 1991). Quelques années plus tard, il s'est avéré que ce domaine possède un potentiel de répression transcriptionnelle dépendant de la liaison à l'ADN (Margolin *et al.*, 1994; Pengue *et al.*, 1994; Witzgall *et al.*, 1994; Vissing *et al.*, 1995). Quoique le motif KRAB n'ait pas été cristallisé, la structure primaire permet de prédire qu'il est organisé probablement en deux hélices amphipatiques. L'organisation de plusieurs gènes codant pour des protéines KRAB à multidoigt de zinc indique que le domaine KRAB est codé par deux exons différents définissant deux boîtes nommées boîte A et boîte B (Rosati *et al.*, 1991; Witzgall *et al.*, 1994; Bellefroid *et al.*, 1991)

Le module de répression minimal correspond aux 45 acides aminés présents au niveau de la boîte A (Margolin *et al.* 1994). La boîte B seule ne semble pas jouer un rôle primordial dans la répression transcriptionnelle (Margolin *et al.*, 1994) quoiqu'une étude ait démontré qu'elle potentialise le pouvoir répressif de la boîte A (Vissing, *et al.*, 1995)

En plus de l'organisation classique du domaine KRAB en deux boîtes A et B, d'autres arrangements ont été caractérisés (**Figure 3**). On retrouve en effet des protéines ne contenant pas de boîte B et des protéines contenant à la place de la boîte B deux autres types de séquences définissant des boîtes nommées boîte b et boîte C. Ces boîtes sont situées comme la boîte B de façon adjacente et en C-terminal de la boîte A. Certaines protéines KRAB à multidoigt de zinc contiennent en plus des domaines KRAB et multidoigt de zinc, un autre domaine appelé SCAN ou **LeR** (Leucine-rich) qui comme le domaine KRAB est spécifique aux vertébrés. Le domaine SCAN ne semble pas posséder

une activité transcriptionnelle (Pengue *et al.*, 1996; Williams *et al.*, 1995) mais il a été démontré qu'il assurait l'oligomérisation des protéines à motifs SCAN (Williams *et al.*, 1995; Schumacher *et al.*, 2000, Sander *et al.*, 2000)



1.1.2 – Le modèle de répression :

L'activité répressive des protéines KRAB à multidoigt de zinc est attribuée au domaine KRAB. Cette activité est dépendante du recrutement d'une phosphoprotéine nucléaire de 97 kDa caractérisée en 1996 par trois groupes de chercheurs différents et appelée KAP-1 (**K**RAB **A**ssociated **P**rotein-1) (Friedman *et al.*, 1996), TIFI- β (**T**ranscriptional **I**ntermediary **F**actor **I**- β) (Moosman *et al.*, 1996) ou KRIP-I (**K**RAB **I**nteracting **P**rotein-**I**) (Kim *et al.*, 1996). Ce co-répresseur commun de toutes les protéines KRAB à multidoigt de zinc appartient à la famille des TIFI, qui inclue TIFI- α et TIFI- γ . Tous les membres de cette famille répriment la transcription lorsque fusionnés à un domaine hétérologue de liaison à l'ADN. Une seule étude rapporte un effet activateur de KAP-1 mais le mécanisme impliqué reste inconnu (Chang *et al.*, 1998)

Le rôle de KAP-1 comme co-répresseur est maintenant mieux compris. KAP-1 est organisé en plusieurs domaines fonctionnels (**Figure 4**) soit en N-terminal le domaine RBCC (**R**ING finger, **B** Boxes, **C**oiled-Coil) qui est nécessaire et suffisant à sa multimérisation et à sa liaison au domaine KRAB (Peng *et al.*, 2000; Saurin *et al.*, 1996) et en C-terminal le domaine PHD (**P**lant-like **H**omeo**D**omain) et le bromodomaine. Cette région C-terminale forme une unité de répression lorsque fusionnée à un domaine hétérologue de liaison à l'ADN (Friedman *et al.*, 1996; Moosmann *et al.*, 1996) et est caractéristique des protéines agissant au niveau de la chromatine (Aasland *et al.*, 1995; Koken *et al.*, 1995). L'interaction de cette région C-terminale avec plusieurs complexes connus pour réprimer la transcription permet d'expliquer son rôle répresseur. Les complexes recrutés sont d'une part le complexe NuRD (**N**ucleosome **R**emodelling and

Desacetylation complex) (Schultz *et al.*, 2001) contenant les deux histones désacétylases HDAC1 et HDAC2 et d'autre part, SETDB1, une protéine à domaine SET qui a une activité méthyltransférase spécifique de la lysine 9 de l'histone 3 (H3K9) (Schultz *et al.*, 2002). La méthylation de H3K9 constitue un signal de déposition des protéines de l'hétérochromatine HP1 en augmentant la liaison de ces dernières aux histones. Les protéines de l'hétérochromatine HP1 sont des protéines non-histones qui ont un rôle important dans la conversion de la chromatine de la forme active (l'euchromatine) à la forme inactive (l'hétérochromatine). Il est à noter que KAP-1 interagit aussi directement avec HP1. Cette interaction est médiée par un segment de 5 acides aminés PxVxL appelé le HP1BD (**HP1 Binding Domain**) situé dans la partie centrale de KAP-1 (Ryan *et al.*, 1999; Lechner *et al.*, 2000)

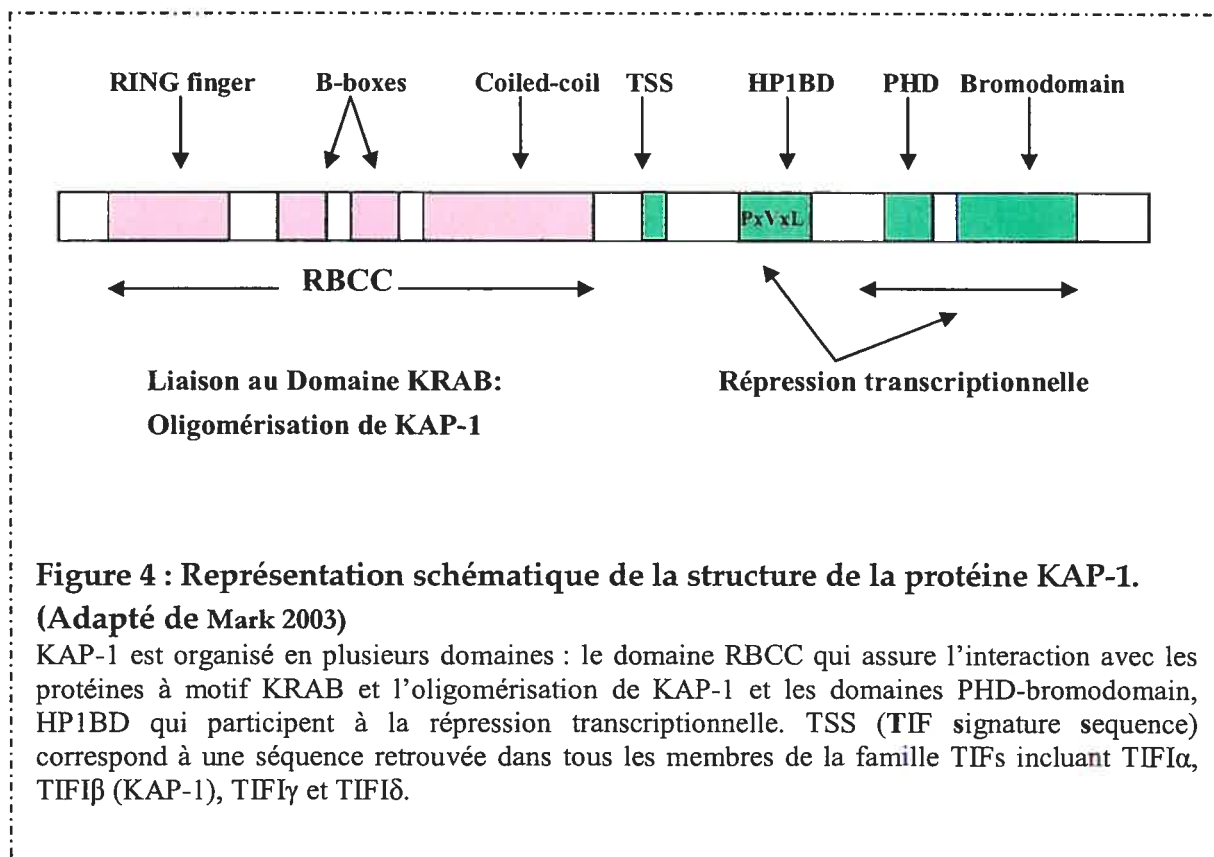
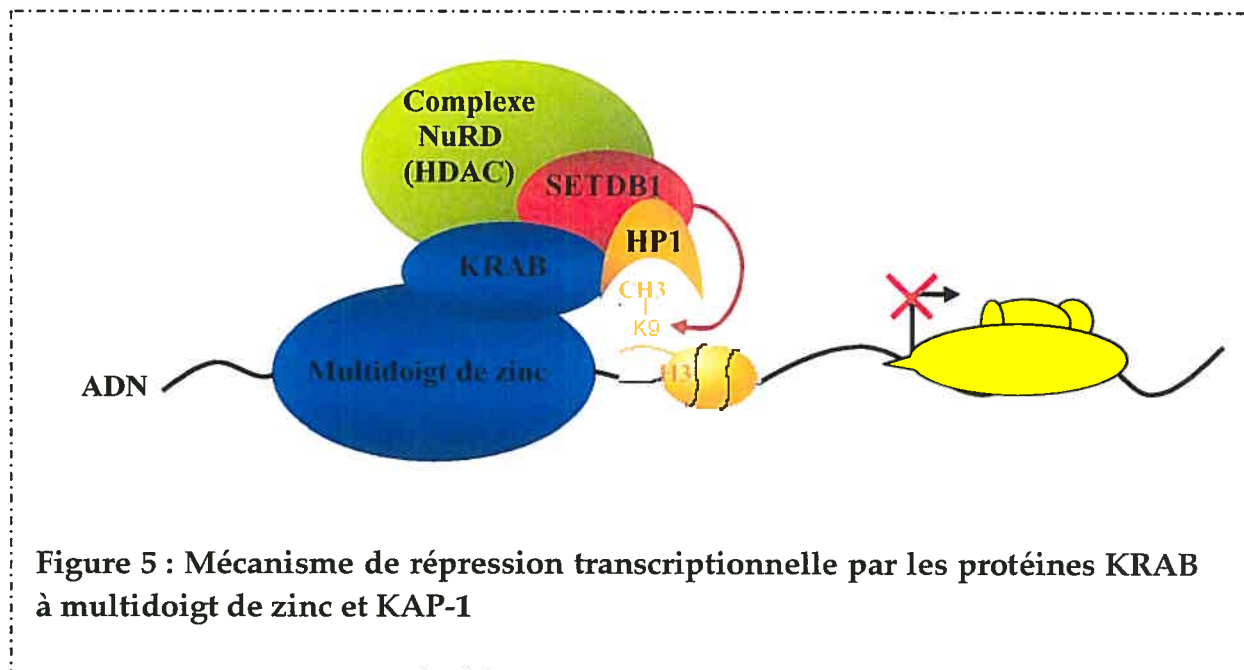


Figure 4 : Représentation schématique de la structure de la protéine KAP-1.

(Adapté de Mark 2003)

KAP-1 est organisé en plusieurs domaines : le domaine RBCC qui assure l'interaction avec les protéines à motif KRAB et l'oligomérisation de KAP-1 et les domaines PHD-bromodomain, HP1BD qui participent à la répression transcriptionnelle. TSS (TIF signature sequence) correspond à une séquence retrouvée dans tous les membres de la famille TIFs incluant TIF1 α , TIF1 β (KAP-1), TIF1 γ et TIF1 δ .

Toutes ces données suggèrent que KAP-1 est recruté à l'ADN via le domaine KRAB des protéines KRAB à multidoigt de zinc et fonctionne comme une plate forme moléculaire qui permet le recrutement des histones méthylases, des histones desacétylases et des protéines de l'hétérochromatine; la formation de ces complexes au niveau de l'ADN conduit à la répression transcriptionnelle (**Figure 5**).



Sur la base du fait que la protéine KRAB à multidoigt de zinc Kox1 réprime la transcription lorsque ciblée à des promoteurs de la polymérase II et III mais pas à des promoteurs de la polymérase I, Moosmann et al (Moosmann *et al.*, 1997) ont suggéré que la répression dépendante de la boîte KRAB et du co-répresseur KAP-1 pourrait aussi impliquer une inhibition directe de composantes de l'ARN pol II et pol III. Ceci reste à prouver.

1.1.3 – Les rôles des protéines KRAB à multidoigt de zinc :

L'ablation génique de KAP-1, le co-répresseur universel des protéines KRAB à multidoigt de zinc est létale au niveau embryonnaire chez la souris (Cammass *et al.*, 2000). Ceci pourrait suggérer un rôle des protéines KRAB à multidoigt de zinc dans le développement. Malgré l'abondance des gènes à motifs KRAB et multidoigt de zinc dans les génomes des vertébrés, leurs rôles physiologiques demeurent cependant très peu étudiés. La plupart des études ont été réalisées dans des modèles cellulaires et leur rôle suggéré dans le développement, la différenciation et la prolifération reste donc à confirmer chez l'animal, par ablation génique en particulier. L'expansion récente de cette famille de gènes observée chez les mammifères, parfois de façon spécifique à l'espèce (Looman *et al.*, 2002) suggère que les protéines KRAB à multidoigt pourraient avoir un rôle dans la spéciation.

Certaines protéines KRAB à multidoigt de zinc ont été sélectionnées comme candidates pour des désordres génétiques en se basant sur leur localisation chromosomique. C'est le cas par exemple de ZNF141 qui a été retenu comme candidat pour le syndrome de Wolf-Hirschhorn (Tommerup *et al.*, 1993) et ZNF74 pour le syndrome de DiGeorge (Aubry *et al.*, 1993; Ravassard *et al.*, 1999). D'autre part, le profil d'expression de la majorité des protéines KRAB à multidoigt de zinc étudiées suggère leur implication dans le développement embryonnaire, la différenciation et la transformation cellulaire (Bellefroid *et al.*, 1991; Bellefroid *et al.*, 1993 ; Gebelein *et al.*, 1998; Liu *et al.*, 1999)

Quelques études ont permis d'attribuer un rôle plus spécifique à des protéines KRAB à multidoigt de zinc. Jheon et ses collaborateurs ont rapporté en 2001 le rôle de la protéine AJ18 dans la modulation de la différenciation ostéogénique. AJ18 est activée de

façon transitoire durant la différenciation des ostéoblastes et régule négativement l'activité de Runx 2, un facteur transcriptionnel clé dans le développement osseux, tel que démontré dans des essais de gènes rapporteurs (Jheon *et al.*, 2001). Le facteur transcriptionnel Zfp 60 est un régulateur négatif de la différenciation du cartilage tel que démontré par la prévention de la différenciation chondrogénique de lignées cellulaires après transfection transitoire de Zfp 60 (Ganss et Kobayashi., 2002). De plus, ce rôle est supporté par son profil d'expression spatio-temporel chez la souris. Des études de gènes rapporteurs ont aussi mis en évidence que ZNF 202 est un facteur transcriptionnel qui régule les promoteurs de plusieurs gènes impliqués dans le métabolisme lipidique (Wagner *et al.*, 2000). D'autres chercheurs suggèrent l'implication des protéines KRAB à multidoigt de zinc dans la différenciation des cellules germinales mâles et femelles (Noce *et al.*, 1992) et dans la régulation du développement des cellules immunitaires (Bellefroid *et al.*, 1993; Blanchard *et al.*, 2000).

Une autre membre de la famille NRIF (**N**eurotrophin **R**eceptor **I**nteracting **F**actor) exprimé de façon ubiquiste est important pour la mort cellulaire durant le développement de certains tissus comme la rétine tel que démontré par la diminution de l'apoptose cellulaire observée au niveau rétinien dans des souris NRIF^{-/-} (Casademunt *et al.*, 1999)

L'abondance des protéines de la famille KRAB à multidoigt de zinc ainsi que leur implication possible dans des processus importants tels que le développement, la différenciation cellulaire, la transformation cellulaire et l'apoptose en font une famille de facteurs transcriptionnels très intéressante à étudier. Plusieurs études sont venues éclaircir le mécanisme de répression des membres de cette famille, mais il reste beaucoup à faire pour

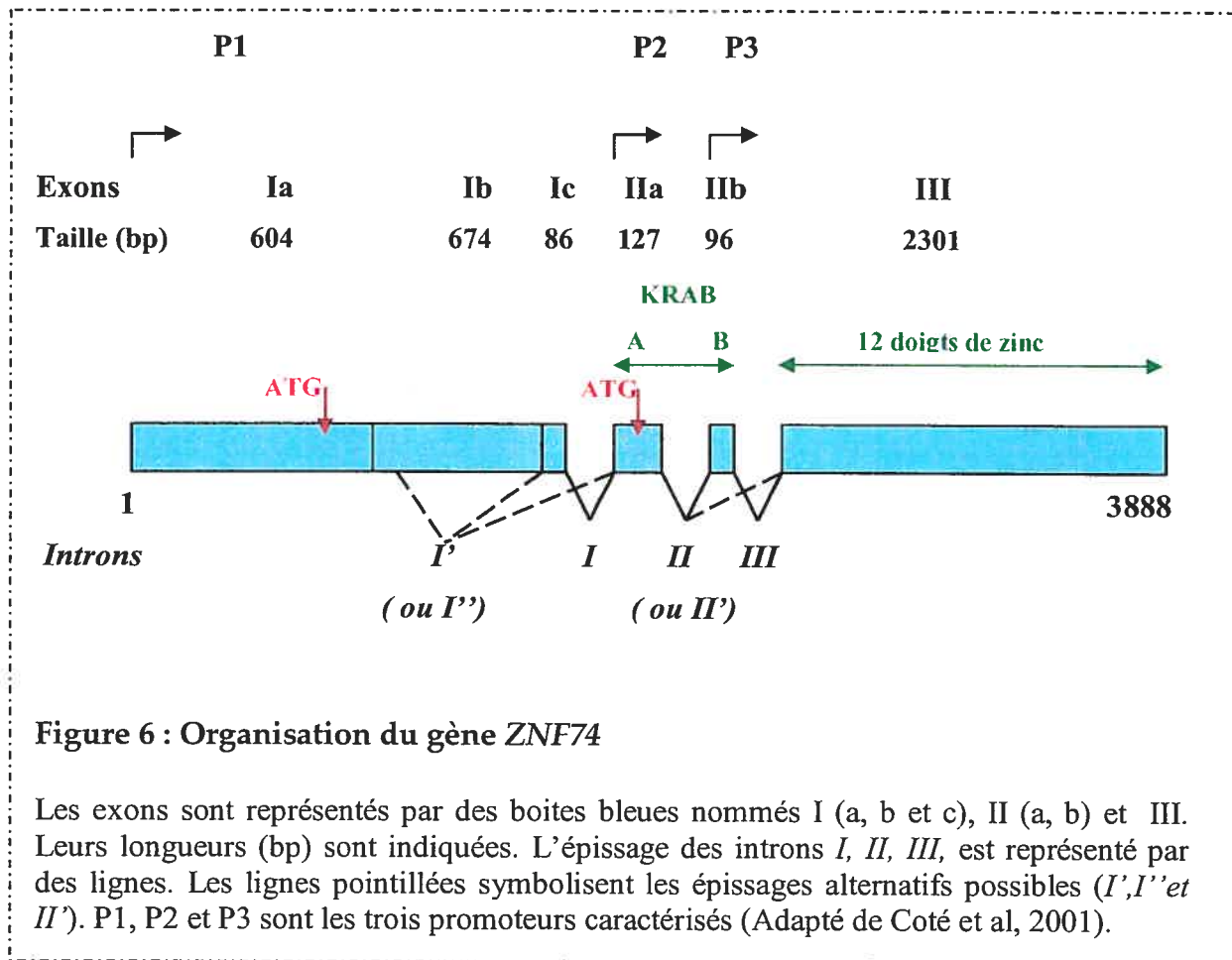
comprendre leurs propriétés de liaison à l'ADN et les différentes voies biologiques qu'ils régulent.

1.2 - ZNF74, une protéine KRAB multidoigt de zinc :

Dans notre laboratoire, nous nous intéressons à la famille de protéines KRAB à multidoigt de zinc en étudiant un prototype de cette famille : la protéine ZNF74.

1.2.1 - Le gène *ZNF74*:

Le gène humain *ZNF74* est situé au niveau du locus 22q11.2. Chez l'embryon humain, son profil d'expression est restreint aux tissus et aux structures qui dérivent de la crête neurale (ganglions de la racine dorsale, ganglions sympathiques, gros vaisseaux du cœur) et aux épithelia des voies respiratoires de l'œsophage (Ravassard *et al.*, 1999). Le gène *ZNF74* est organisé en 6 exons nommés Ia, Ib et Ic, IIa, IIb et III (Coté *et al.*, 2001) (**Figure 6**). Le gène est sujet à un épissage qui élimine les introns *I*, *II* et *III*. Un épissage alternatif éliminant les introns *I'* et *II'* a été aussi caractérisé. Les résultats obtenus par les expériences de RACE (**R**apid **A**mplification of **c**DNA **E**nds) et de RT-PCR suggèrent que la transcription de *ZNF74* peut être initiée au niveau de 3 promoteurs différents appelés P1, P2 et P3. L'exon IIa et IIb codent pour les deux boîtes A et B du domaine KRAB alors que l'exon III code pour le domaine multidoigt de zinc de la protéine KRAB multidoigt de zinc ZNF74.



1.2.2 - La protéine ZNF74 :

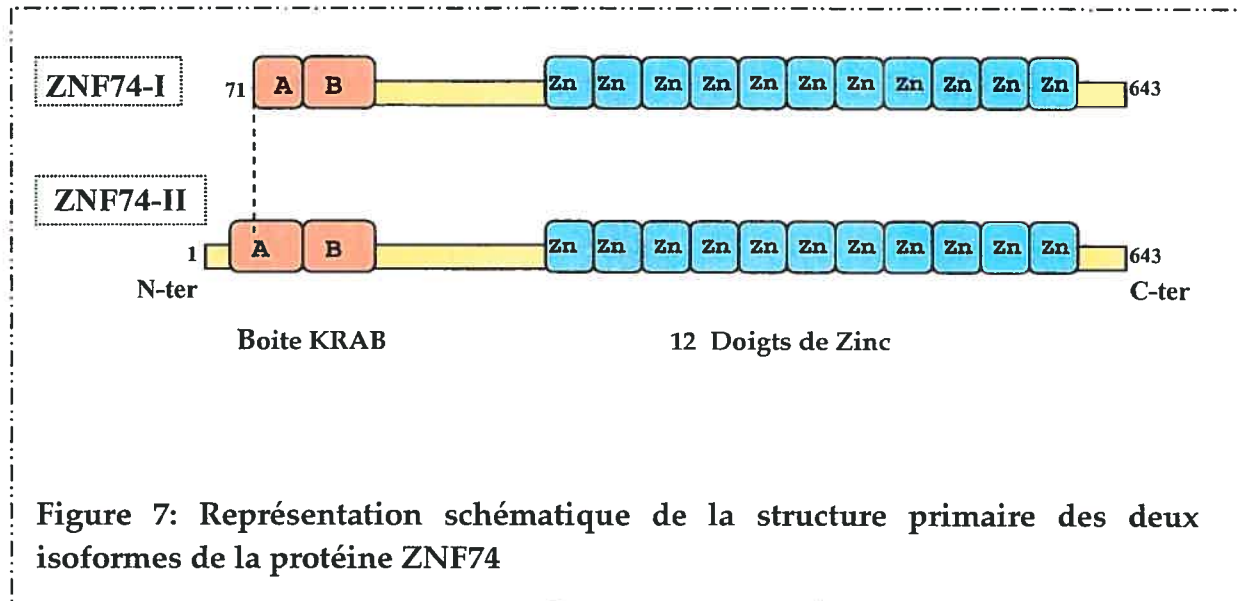
La protéine ZNF74 possède deux domaines fonctionnels : un domaine KRAB classique incluant les deux boîtes A et B et un domaine multidoigt de zinc avec 12 doigts de zinc en tandem.

Le domaine KRAB de la protéine ZNF74 réprime la transcription lorsque fusionné au domaine de liaison à l'ADN de GAL4 (Coté *et al.*, 2001). Le domaine multidoigt de zinc dirige la localisation nucléaire de la protéine et médie sa liaison à la matrice nucléaire (Grondin *et al.*, 1996). D'autre part ZNF74 interagit via son domaine doigt

de zinc avec l'ADN simple brin et double brin tel que démontré par retardement sur gel et Southern Blot (Guévremont, 1999). Guévremont (1999) a tenté d'identifier l'élément de réponse de ZNF74 en utilisant une procédure cyclique de purification sur colonne d'affinité couplée à une amplification par PCR mais aucune séquence spécifique d'ADN n'a pu être identifiée. ZNF74 est aussi capable d'interagir avec l'ARN par le biais du domaine doigt de zinc tel que démontré en utilisant des polymères d'ARN poly (U) et poly (G) (Grondin *et al.*, 1996). Ceci suggère outre un rôle de facteur transcriptionnel, un rôle possible de ZNF74 dans le métabolisme de l'ARN.

Le gène *ZNF74*, par utilisation alternative de promoteurs et par épissage, donne naissance à deux isoformes principales de la protéine ZNF74 (Côté *et al.*, 2001) (**Figure 7**). L'isoforme ZNF74-I (aa 71-643) est traduite à partir de deux variants de l'ARNm qui incluent soit les exons I (a,b,c), II (a,b) et III soit les exons I (a, c), I (a,b) et III. L'isoforme ZNF74-I, initiée au niveau d'une méthionine incluse dans la boîte KRAB A possède un domaine KRAB incomplet et non répresseur. Cette isoforme est localisée majoritairement au niveau de domaines sous-nucléaires de forme irrégulière et ponctuée appelés aussi les «speckles» qui sont connus pour être enrichis en facteurs de maturation des pré-ARNm tels les facteurs d'épissage (Côté *et al.*, 2001). De façon intéressante, il a été montré que ZNF74 interagissait avec la forme hyperphosphorylée de la large sous-unité de l'ARN polymérase II (pol IIo) (Grondin *et al.*, 1997) aussi connue pour interagir avec des facteurs d'épissage et des facteurs impliqués dans l'addition de la coiffe et le clivage de l'ARN suggérant à nouveau un rôle possible de ZNF74 dans la maturation de l'ARN. L'isoforme ZNF74-II (aa 1-643), initiée à une méthionine encodée dans l'exon Ia est la forme répressive avec un domaine KRAB complet possédant les deux boîtes A et B. Cette

isoforme contrairement à la première est localisée de façon plus diffuse dans le noyau sous forme de domaines micro-ponctués ne colocalisant pas spécifiquement avec les facteurs d'épissage (Côté *et al.*, 2001)



1.3 - ZNF74 et le syndrome de DiGeorge:

1.3.1 - Le syndrome de DiGeorge:

Le syndrome de DiGeorge est une maladie congénitale attribuée à une délétion hémizygote d'une région spécifique du chromosome 22 soit la région 22q11.2 (Driscoll *et al.*, 1992). C'est la microdélétion chromosomique la plus commune chez l'humain avec une incidence de 1 sur 4000 à 5000 naissances. Dans la majorité des cas, la délétion est sporadique mais dans 10 à 20 % des cas, elle se transmet selon un mode autosomique dominant.

Les symptômes cliniques associés au syndrome de DiGeorge sont multiples. Parmi les anomalies les plus courantes, on compte l'hypoplasie voir même l'aplasie du

thymus qui se traduit par une immunodéficience. D'autre part une hypoplasie des parathyroïdes est observée entraînant une hypocalcémie. Les anomalies cardiaques liées au syndrome de DiGeorge touchent principalement les gros vaisseaux du cœur incluant la carotide, la valve aortique et la paroi qui sépare l'aorte et l'artère pulmonaire. Ces anomalies affectent sérieusement le pronostic des patients et constituent la principale cause de mortalité. De plus, les patients présentent une dysmorphie faciale légère mais caractéristique.

Au niveau embryologique, les anomalies observées sont probablement la cause d'une défaillance de migration et/ou d'interaction des cellules de la crête neurale avec le mésenchyme des arches pharyngées. Par conséquent, le syndrome de DiGeorge a été décrit comme une neurocristopathie (pathologie liée aux cellules de la crête neurale).

Dans la plupart des cas, la délétion au niveau du locus 22q11.2 élimine une région d'ADN de 3 Mb qui encode approximativement une trentaine de gènes (McDermid *et al* 2002). Dans certains cas, des délétions non superposables à l'intérieur de la région de 3Mb sont observées alors que dans d'autres, le segment délété se situe toujours dans le locus 22q11.2 mais en dehors de la région standard de 3Mb (**Figure 8**).

Le syndrome de DiGeorge apparaît donc comme un syndrome génétique complexe et il a été suggéré que plus d'un gène de la région 22q11.2 pourrait seul ou en combinaison avec d'autres gènes de cette région participer au phénotype. Le syndrome de DiGeorge pourrait être ainsi un syndrome de gènes contigus. En fait, l'ablation génique complète de *Tbx1* ou bien de *Crkl1*, deux gènes de la région 22q11 reproduit le phénotype du syndrome de DiGeorge (Jerome *et al.*, 2001)

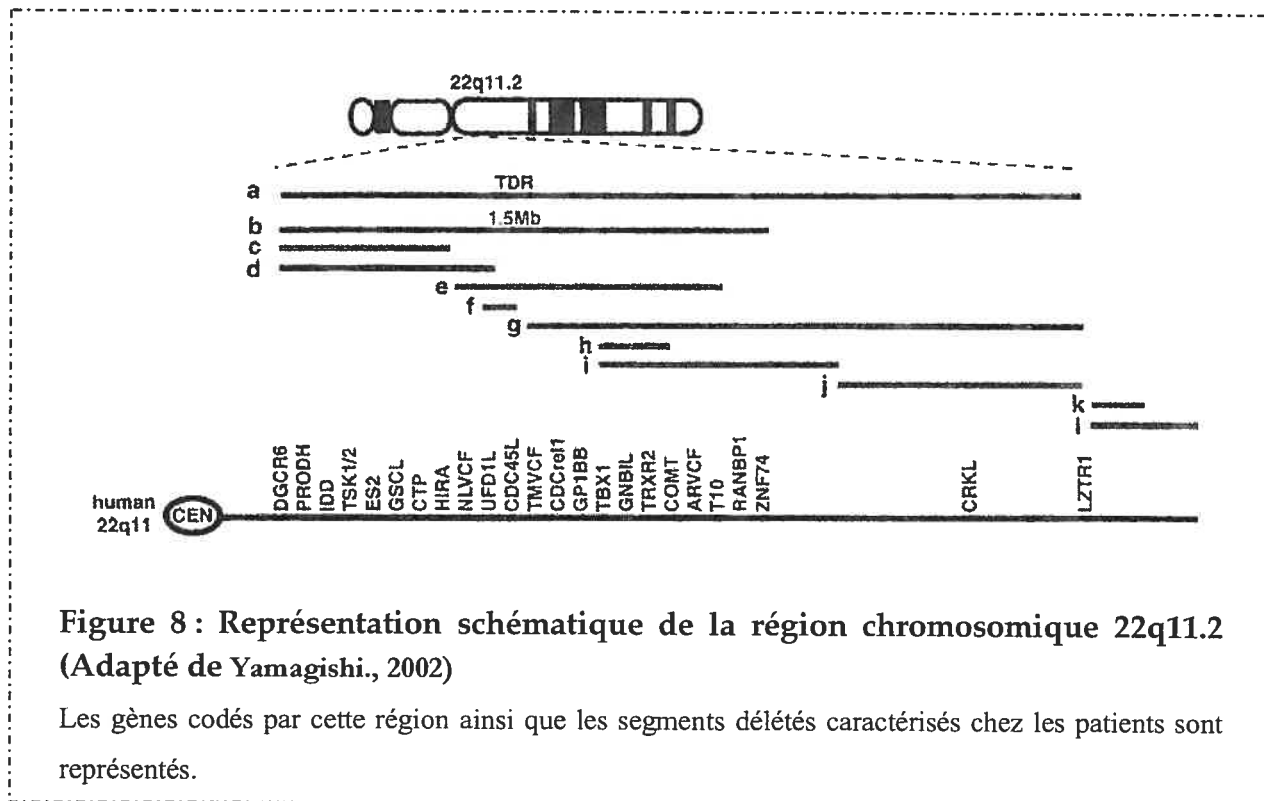


Figure 8 : Représentation schématique de la région chromosomique 22q11.2 (Adapté de Yamagishi., 2002)

Les gènes codés par cette région ainsi que les segments délétés caractérisés chez les patients sont représentés.

1.3.2- ZNF74, candidat pour le syndrome de DiGeorge:

ZNF74 a été isolé initialement par Aubry et al en 1993 (Aubry *et al.*, 1993) comme un gène candidat pouvant être impliqué dans le syndrome de DiGeorge. Plusieurs éléments étaient en faveur de ce rôle proposé de candidat dans le syndrome de DiGeorge :

- 1- Le gène *ZNF74* fait partie de la région de délétion standard de 3 MB perdue chez la grande majorité des patients atteints du syndrome de DiGeorge (Aubry et al 1993)
- 2- Le profil d'expression de *ZNF74* chez l'homme est restreint aux structures qui dérivent des cellules de la crête neurale ce qui corrèle étroitement avec celui des structures affectées par le syndrome de DiGeorge;
- 3- *ZNF74* appartient à une famille de gènes encodant des facteurs transcriptionnels de type KRAB multidoigt de zinc et des

modifications de dosage ou d'activité de facteurs transcriptionnels sont souvent à la base de maladies du développement; 4 - L'exposition à la bisdiamine, un agent chélateur du zinc durant l'embryogenèse chez le rat induit un phénotype similaire au syndrome de DiGeorge ce qui pourrait perturber la structure et l'activité de protéines à doigt de zinc comme ZNF74. (Osler et al 1983)

1.4 – Objectifs de recherche :

L'un des moyens pour mieux comprendre la fonction de la protéine ZNF74 durant le développement serait d'identifier la ou les séquences d'ADN ainsi que les gènes cibles auxquels elle se lie de façon spécifique. L'identification des gènes cibles régulés devrait en effet conduire à déterminer dans quelle(s) voie(s) de signalisation développementale cette protéine intervient.

Les objectifs spécifiques de ce projet de recherche sont les suivants :

- A. Identifier les séquences spécifiques d'ADN reconnues par la protéine ZNF74 en utilisant une technique de sélection qui combine une purification par South-Western et une amplification par PCR.
- B. Confirmer la spécificité de liaison des séquences identifiées *in vitro* par gel de retardement et *in vivo* par essai de gène rapporteur.
- C. Déterminer si l'une des cibles de ZNF74 pourrait être son propre promoteur à l'instar de certains facteurs transcriptionnels.
- D. Rechercher dans les banques de données des gènes humains cibles contenant ces séquences dans leurs promoteurs et tester la fonctionnalité des meilleurs candidats.

CHAPITRE II

MATÉRIEL ET MÉTHODES

2.1 - Vecteurs :

2.1.1 - Vecteurs d'expression protéique :

2.1.1.1 – Vecteur pMAL-c :

Le vecteur pMAL-c (New England Biolabs) encode pour la protéine Maltose Binding Protein (MBP). Ce vecteur est utilisé pour exprimer deux protéines d'intérêt sous forme de protéines de fusion avec la MBP, soient pMAL-c-Zn (P.M calculé = 80 200) qui correspond à la région multidoigt de zinc (doigts de zinc 1 à 12) de la protéine ZNF74 [aa 246-580] décrite précédemment dans (Germain-Desprez *et al.*, 2003) et pMAL-c-ZNF74-II Δ 1-176 (P.M.calculé = 94 300) qui correspond à la région codante complète de ZNF74 moins les 176 premiers acides aminés [aa 177-643]. Cette dernière est préparée par clonage d'un fragment de restriction EcoRI de 1.9 Kilobases (Kb) dans le site de restriction EcoRI du vecteur pMAL-c. Le vecteur pMAL-c non modifié génère la protéine de fusion MBP- β -galactosidase (P.M.calculé = 51 000).

2.1.1.2 – Vecteur pCGN :

Le vecteur d'expression protéique pCGN (Tanaka et Her., 1990) a servi à diriger l'expression de la protéine ZNF74 sous forme de protéine de fusion avec l'épitope hémagglutinine (HA) du virus de l'influenza dans les expériences de transfection transitoire dans les cellules de mammifères. Ce vecteur possède un promoteur CMV en amont du site de multiclonage. L'ADNc de ZNF74-II a été excisé du vecteur SK par les enzymes de restriction XbaI et KpnI puis cloné entre les sites XbaI et KpnI du vecteur pCGN de façon à inclure toute la séquence codante de ZNF74-II (Coté *et al.*, 2001) en aval et en phase avec l'épitope HA.

2.1.1.3 – Vecteur pRSV :

Le vecteur pRSV est utilisé pour exprimer la protéine ZNF74 en fusion avec le domaine de liaison à l'ADN de GAL4 (GAL4 DBD) (**GAL4 DNA Binding Domain**). Ce dernier correspond aux acides aminés 1 à 147 de la protéine GAL4 et est cloné dans le site XbaI du vecteur pRSV. L'ADNc de ZNF74 II est cloné en aval et en phase avec le GAL4DBD dans le site de restriction XbaI du vecteur pRSV.

2.1.2 - Vecteurs de gène rapporteur Luciférase :

Dans les essais de gène rapporteur, trois vecteurs ont été utilisés : Le vecteur pGL3 *basic*, le vecteur pGL3 *enhancer* et le vecteur pGL3 *contrôle* (Promega). Tous les trois sont prévus pour diriger l'expression de la protéine Luciférase de la mouche à feu «firefly Luciférase» .

2.1.2.1 – Vecteur pGL3 *basic* :

Le vecteur pGL3 *basic* (Promega) ne contient ni promoteur ni *enhancer* ce qui rend l'expression de la luciférase dépendante entièrement de l'insertion de régions promotrices en amont du gène *luciférase*.

2.1.2.2 – Vecteur pGL3 *contrôle* :

Le vecteur pGL3 *contrôle* (Promega) exprime la luciférase de la mouche à feu sous le contrôle du promoteur SV40. L'activité de ce promoteur est soutenue par l'*enhancer* SV40 situé en aval du gène *luciférase*.

2.1.2.3 – Vecteur pGL3 *enhancer* :

Le vecteur pGL3 *enhancer* (Clone392) a été généré à partir du vecteur pGL3 contrôle par digestion enzymatique BglII et HindIII afin d'éliminer le promoteur de SV40. Par conséquent, ce vecteur contient l'*enhancer* de SV40 seulement.

2.1.2.4 – Clonage des oligonucléotides sélectionnés via la purification par South-Western dans le vecteur pGL3 *contrôle* :

Des oligonucléotides doubles brins sélectionnés par South-Western ont été clonés dans le vecteur pGL3 *contrôle* sous forme de monomère de 16 nucléotides incluant la séquence consensus identifiée (1 X) ou de multimère (4 X), les oligonucléotides doubles brins phosphorylés ont été clonés dans le site de restriction SmaI du vecteur pGL3 *contrôle* déphosphorylé en amont du promoteur de SV40. Les séquences clonées sont présentées au **Tableau III**.

Tableau III: Les constructions de 3 oligonucléotides purifiés par South-Western dans le vecteur pGL3 *contrôle*.

Nom de la Construction	Séquence des 16 Nucléotides	Noms des clones
pGL3 contrôle 76 (4X)	CATCTA <u>A</u> GATTGTTAC	Clone 399
pGL3 contrôle 31 (1X)	CATA <u>AAATTC</u> GCCAACA	Clone 400
pGL3 contrôle 73 (1X)	TTTCAG <u>A</u> GATTCCGCT	Clone 401

Pour chaque séquence, les oligonucléotides doubles brins devant être utilisés pour le clonage ont été préparés à partir de deux oligonucléotides simple brin 1 µg des deux oligonucléotides simples brins synthétiques phosphorylés qui ont été hybridés dans une

réaction de 50 µl de tampon d'hybridation (Tris HCl 100 mM pH 8, EDTA 50 mM, NaCl 500 mM). Le mélange a été chauffé pendant 5 min à 95 °C puis refroidi lentement dans le bac métallique de chauffage placé à température de la pièce. Alors que pour les constructions pGL3 *contrôle* 31 (1 X) et pGL3 *contrôle* 73 (1 X), on a inséré un monomère de 16 nucléotides, le pGL3 *contrôle* 76 (4 X) contient 4 copies en tandem des 16 nucléotides. Ce multimère a été préparé par ligation de l'oligonucléotide double brin dans 70 µl de tampon de ligation (Tris HCl 50 mM pH 7.6, MgCl₂ 10 mM, ATP 1 mM, DTT 1 mM, polyéthylène glycol-8000 5 %) contenant 48 µl de la réaction d'hybridation, 70 unités de la T4 DNA Ligase et 70 mM d'ATP à 14° C pendant 16h. Les multimères ont été purifiés sur gel d'acrylamide 15 % puis clonés dans le vecteur pGL3 *contrôle*.

2.1.2.5 – Clonage de l'élément de réponse de GAL4 dans le vecteur pGL3 *contrôle* :

L'élément de réponse de la protéine GAL4 AGGGTATATAATG multimérisé 5 fois a été excisé du vecteur SK⁺ par les enzymes de restriction SacI et KpnI et cloné dans les sites SacI et KpnI du vecteur pGL3 *contrôle* en amont du promoteur de SV40.

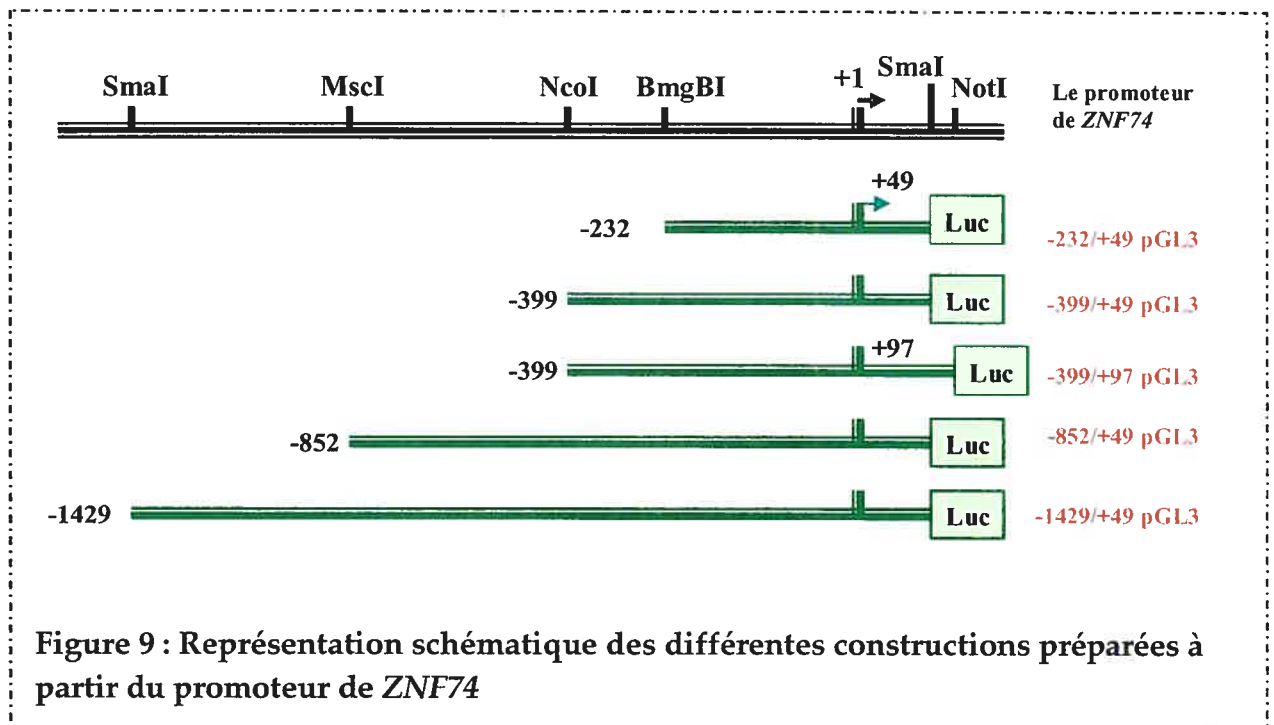
2.1.2.6 – Clonage du promoteur de ZNF74 dans le vecteur pGL3 *enhancer* et pGL3 *basic* :

Afin de caractériser l'activité du promoteur de ZNF74, des constructions contenant différents fragments du promoteur ont été préparées par sous-clonage dans le site SmaI des vecteurs pGL3 *enhancer* et pGL3 *basic* (**Figure 9**) et (**Tableau VI**). Les régions promotrices ont été libérées de la construction (Clone FC5) qui correspond au vecteur SK⁺ contenant le fragment BamHI (2.136Kb) d'ADN génomique de la région de 5' non traduite du gène *ZNF74* (Genbank NT_0111520.10, position 137204 à 139422). Les enzymes utilisées pour libérer les inserts sont mentionnées dans le Tableau II et les constructions générées sont représentées à la **Figure 9**.

Tableau VI: Stratégie de clonage des différents fragments du promoteur de ZNF74

Nom de la Construction	Enzymes utilisées pour libérer l'insert	Taille de l'insert (bp)	Noms des clones dans PGL3 <i>basic</i> et <i>enhancer</i>
-232/+49 pGL3	SmaI - BmgBI	281	Clones 388 et 393
-399/+49pGL3-	NcoI- SmaI	448	Clonc 397*
-399/+97pGL3	NcoI-NotI	496	Clone 391 et 396
-852/+49 pGL3	SmaI-MscI	901	Clones 389 et 394
-1429/+49 pGL3	SmaI	1478	Clones 390 et 395

*Ce fragment a été cloné dans le pGL3 *enhancer* seulement.



2.1.2.7- Vecteur pHRLuc :

Ce vecteur encode la protéine luciférase *Renilla* sous le contrôle du promoteur CMV. Il est co-transfecté dans cellules HEK293T avec les constructions pGL3 utilisant la luciférase de la mouche à feu comme rapporteur de transcription. Il sert en effet à normaliser les essais de gène rapporteur en renseignant sur l'efficacité de la transfection.

2.2 - Expression protéique :

2.2.1 - Protéines d'origine bactérienne :

2.2.1.1 – Préparations des extraits bactériens :

Des bactéries compétentes *Escherichia coli* de la souche DH5 α ont été transformées par les vecteurs suivants (décrits précédemment dans la section 2.1.1.1): p-MAL-c, p-MAL-c-Zn, p-MAL-c-ZNF74-II (Δ 1-176) et génèrent respectivement les protéines suivantes : MBP, MBP-Zn et MBP-ZNF74-II (Δ 1-176).

Initialement, une pré-culture de bactéries transformées est préparée dans 30 ml de milieu Luria Broth (LB) contenant de l'ampicilline à 100 μ g/ml (incubation de 16 h à 37°C). Cette culture est diluée au 1/10 dans un volume de 300 ml de LB et réincubée pendant 1 h à 37°C. L'expression des protéines à partir des vecteurs pMAL-c est induite par l'ajout de l'IsoPropyl- β -D-Thiogalactopyranoside (IPTG) à une concentration finale de 0,3 mM pour une période de 3 h à 37°C.

Le culot bactérien obtenu après une centrifugation de 20 min à 1500 g à 4°C est resuspendu dans 20 ml de tampon de lyse (Tris HCl 200 mM, NaCl 200 mM, β -mercapthoéthanol 10 mM). Pour préparer des extraits bactériens, on a procédé par la suite à

une sonication de 8 cycles de 30 sec chacun (vitesse 3, réglage constant) dans un sonificateur (Sonifier 450 Branson). Une incubation d'une minute sur glace entre les cycles est respectée pour éviter un réchauffement des échantillons. Les extraits bactériens ont été centrifugés pendant 20 min à 5200 g à 4°C et les surnageants obtenus ont été aliquotés et entreposés à - 80°C.

2.2.1.2 – Détermination de la concentration des protéines de fusion dans les extraits bactériens :

Les quantités de protéines de fusion MBP-Zn et MBP-ZNF74-II (Δ 1-176) présentes dans les extraits bactériens sont évaluées par buvardage Western avec un anticorps anti-MBP en utilisant comme standard la protéine MBP préalablement dosée sur gel coloré au bleu de Coomassie en comparaison avec une gamme étalon de BSA (Bovine Serum Albumine).

Dans les extraits bactériens, MBP représente la protéine majoritaire (plus de 80%). Pour estimer la quantité de protéine MBP présente dans des extraits bactériens qui sera utilisée comme standard, des quantités croissantes des extraits bactériens sont séparées sur gel de SDS-polyacrylamide (8 %; 29.2 : 0.8) en parallèle avec une gamme étalon de BSA (50 ng à 1 μ g) et les gels sont colorés au bleu de Coomassie. On estime la quantité de MBP par comparaison visuelle de l'intensité de la coloration obtenue pour la protéine MBP (poids moléculaire apparent : 51 kDa) avec celle des différentes concentrations de BSA (poids moléculaire apparent : 67 kDa).

Pour doser les protéines de fusion MBP-Zn et MBP-ZNF74-II (Δ 1-176) dans les extraits bactériens, des quantités croissantes de chaque extrait sont séparées sur gel SDS-

polyacrylamide en parallèle avec des quantités connues de MBP provenant d'extraits dosés par gel de Coomassie. Après une migration de 2 heures à 105 V, les protéines sont électro-transférées sur une membrane de nitrocellulose. Par la suite, on procède à la détection des protéines de fusion immobilisées sur membrane par Western, en utilisant des anticorps primaires dirigés contre la protéine MBP et des anticorps secondaires couplés à la peroxydase HRP (HorsRadish Peroxydase). La révélation se fait par la trousse «Renaissance» de (NEN) utilisant le luminol comme substrat pour la HRP. L'émission de la lumière à 428 nm résultant de la dégradation du substrat est captée par une caméra grâce à un appareil Multi Imager MAX (Biorad). L'analyse des profils de chimiluminescence est effectuée avec le logiciel Quantity one (Option Fluor S Max). L'intensité de trois échantillons de protéines de fusion est comparée à la gamme étalon de cinq quantités croissantes de MBP.

2.2.1.3 – Purification des protéines de fusion :

Les protéines exprimées sous forme de fusion avec la protéine MBP sont purifiées par chromatographie d'affinité sur colonne d'amylose.

Les extraits bactériens totaux sont incubés avec les billes d'amylose (0.5 ml) préalablement lavées avec le tampon de lyse (Tris HCl 200 mM pH 7.5, NaCl 200 mM, β -mercapthoéthanol 10 mM). L'incubation est faite sous agitation pendant 1 h à 4°C. Le culot obtenu par une centrifugation de 5 min à 2250 g est transféré dans une colonne (Biorad). On a procédé à deux lavages avec 20 ml de tampon de lyse (Tris HCl 200 mM pH 7.5, NaCl 200 mM, β -mercapthoéthanol 10 mM). Le culot final contenant les protéines de fusion purifiées est resuspendu dans du tampon Laemmli (Glycérol 15%, SDS 2%, Tris HCl 80 mM pH 6.8, BRB 0,01 %, β -mercapthoéthanol 10 mM).

2.2.1.4 – Détermination de la concentration des protéines de fusion purifiées :

Les protéines de fusion purifiées MBP-Zn et MBP-ZNF74-II (Δ 1-176) ont été dosées sur gel de SDS-polyacrylamide (8 %; 29.2:0.8) coloré au bleu de Coomassie, en comparaison à une gamme étalon contenant sept quantités croissantes de BSA: 50, 100, 200, 300, 500, 750 et 1000 ng.

2.2.2 - Expression protéique dans les cellules de mammifères :

2.2.2.1 – Transfection transitoire :

Les cellules HEK293T utilisées pour les transfections transitoires ont été étalées à raison de 2×10^5 cellules / puits dans des plaques de 6 puits (diamètre = 35 mm) 18 h avant transfection. Chaque puits contient 3 ml de milieu **Delbecco's Modified Eagle's Medium** (DMEM) incluant 10 % de **Sérum Fœtal Bovin** (FBS), 100 μ g/ml de streptomycine et 100 μ g/ml de pénicilline.

Le Fugene (Roche) (3 μ l) est utilisé pour transférer les cellules HEK293T avec un total de 1 μ g d'ADN provenant de maxi-préparation (Marlingene). L'ADN est tout d'abord mélangé avec le Fugene puis incubé pendant 20 minutes à température ambiante avant d'être rajouter goutte à goutte à la culture cellulaire. Les cellules sont ensuite incubées pendant 46 h à 37 °C dans un incubateur contenant 5 % de CO₂.

2.2.2.2 – Préparation des extraits cellulaires :

Après avoir aspiré le milieu de culture, les cellules sont récoltées avec 1ml de tampon phosphate salin (PBS) (NaCl 137 mM; KCl 2.7 mM; Na₂HPO₄ 10 mM; KH₂HPO₄ 2 mM) contenant 2 mM d'EDTA puis rincées une fois avec 1 ml de PBS contenant 1 mM

de glucose. À chaque fois, les cellules sont centrifugées pendant 15 minutes à 1500 g à température ambiante. Le culot final est resuspendu dans 250 μ l de PBS-Glucose 1 mM.

2.2.2.3 – Dosage des protéines dans les extraits cellulaires :

Le dosage des protéines présentes dans les extraits cellulaires est réalisé selon la méthode de Bradford. Un échantillon de 10 μ l / 250 μ l de l'extrait cellulaire est mélangé à 790 μ l du réactif de dosage (Biorad protein assay). L'absorbance est ensuite mesurée à 595 nm après une incubation de 10 min et est comparée à une gamme étalon de BSA constituée de 5 points : 2 μ g, 5 μ g, 10 μ g, 15 μ g et 20 μ g.

2.3 – Expérience de purification par South-Western :

L'essai de sélection d'oligonucléotides liés par ZNF74 par la technique de South-Western est une adaptation de la méthode initialement décrite par Thiesen et Bach (Thiesen et Bach., 1990). Cet essai est détaillé à la **figure 10**.

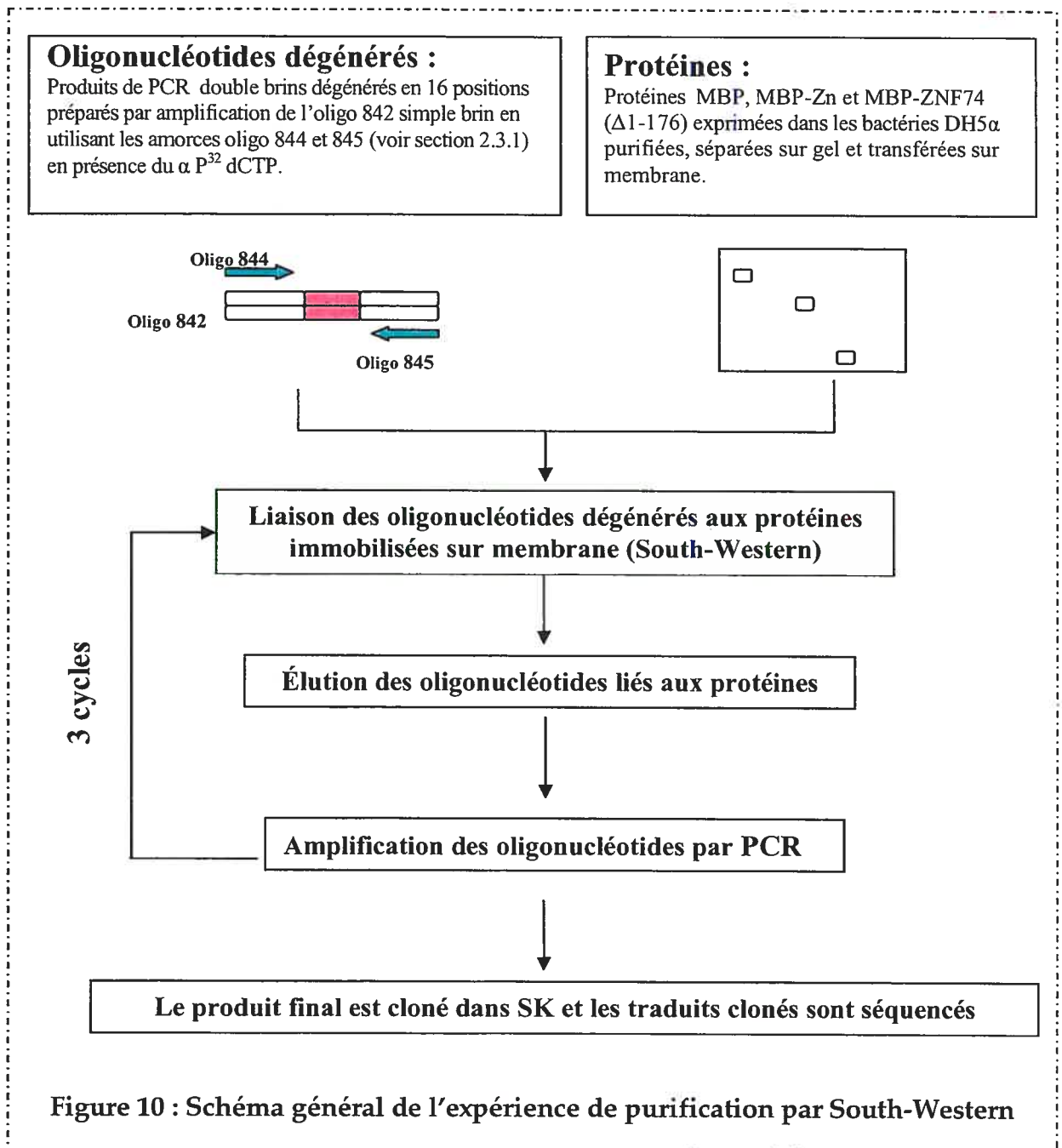
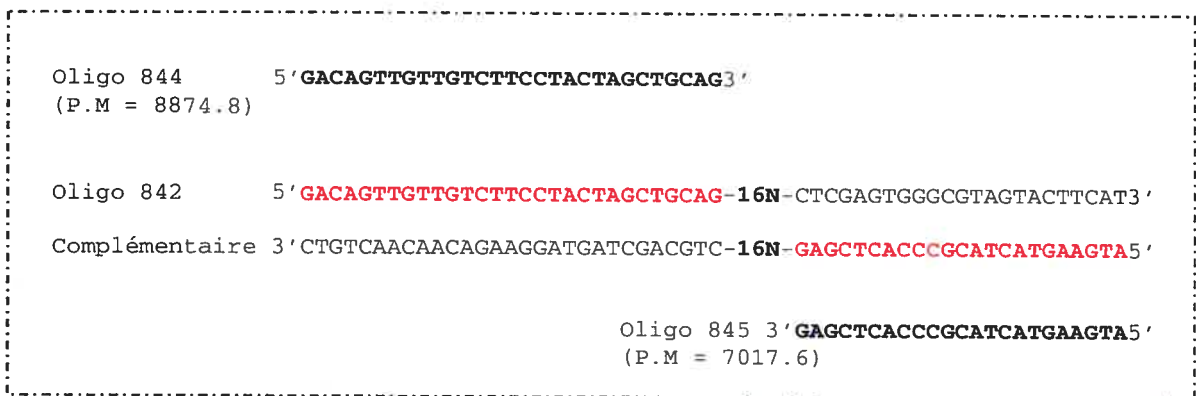


Figure 10 : Schéma général de l'expérience de purification par South-Western

2.3.1-Préparation des oligonucléotides dégénérés doubles brins :

Les oligonucléotides doubles brins dégénérés en 16 positions et marqués au α -P³², sont obtenus par amplification par PCR de l'oligonucléotide simple brin 842 [5' GACAGTTGTTGTCTTCCTACTAGCTGCAG-16N-CTCGAGTGGGCGTAGTACTTCAT3'] en présence de deux amorces : l'oligo 845 qui correspond au fragment complémentaire des 23 derniers nucléotides de l'oligo 842 et l'oligo 844 qui correspond aux 29 premiers nucléotides de l'oligo 842 (voir schéma plus bas).



L'amplification de 10 ng de l'oligo 842 se fait dans une réaction de PCR de 25 μ l de tampon de PCR (Tris HCl 20 Mm pH 8.4, KCl 50 Mm) en présence de 200 ng de chaque amorce, de dNTPs 0.26 mM, de MgCl₂ 1.9 mM, de 2.5 UI de Taq DNA polymérase (Invitrogen) et de 40 μ ci du [α -P³²] dCTP (l'activité spécifique = 3000 ci / mmol) (Perkin Elmer). On a effectué 7 cycles d'amplification incluant 1 minute de dénaturation à 94 °C, 1 minute d'hybridation des amorces avec la matrice à 46 °C et 1 minute d'élongation à 72 °C. Les produits de PCR sont dosés sur gel d'acrylamide 15 % coloré au bromure d'éthidium en comparaison au marqueur ϕ X174 digéré par HaeIII en mesurant la fluorescence sous UV avec un appareil Transluminateur (Biorad). La quantification est effectuée avec le programme Gel doc du logiciel Quantity one.

2.3.2 - Détermination de l'activité spécifique des oligonucléotides marqués et rendement de marquage:

Pour déterminer l'activité spécifique des oligonucléotides marqués radioactivement, on quantifie par essai sur filtre la radioactivité liée aux oligonucléotides. Ainsi, 2 μ l du produit de PCR dilué dans l'eau sont déposés sur filtre DE81 (Whatman). Le filtre est rincé immédiatement 2 fois pendant 2 minutes dans du tampon phosphate 0.5 M (Na₂ HPO₄ 0.5 M, NaHPO₄ 0.5 M) puis 1 fois pendant 1 minute dans de l'eau et enfin 1 fois pendant 1 minute dans de l'éthanol 95 %. Le test est réalisé en duplicata. Les filtres secs sont mis en présence de liquide à scintillation et la radioactivité en cpm est mesurée dans un compteur à scintillation liquide Wallac 1409 de Perkin Elmer. L'activité spécifique correspond à :

$$\text{Activité spécifique de l'oligo (cpm/ ng)} = \frac{\text{Radioactivité liée (cpm)}}{\text{Quantité en ADN (ng)}}$$

Pour estimer le rendement de marquage de l'oligonucléotide, on compare la radioactivité liée aux oligonucléotides à la radioactivité totale utilisée. Pour déterminer la radioactivité totale, les filtres DE81 sur lesquels les échantillons de produits de PCR sont déposés sont directement séchés sans lavage préalable. Le rendement de marquage correspond à :

$$\text{Rendement de marquage} = (\text{Radioactivité liée/ Radioactivité totale}) \times 100$$

2.3.3 - Purification des oligonucléotides marqués sur colonne Sephadex :

Pour éliminer la radioactivité libre présente dans la réaction de PCR, les oligonucléotides marqués et amplifiés par PCR sont purifiés par chromatographie d'exclusion en utilisant une colonne Sephadex G 50. La colonne est préparée en utilisant une seringue de 1 ml stérile bouchée à son extrémité inférieure par une petite quantité de la fibre de verre stérile siliconisée et remplie de billes de Séphadex G 50 (Amersham Pharmacia). Après une centrifugation de 2 min à 1500 g pour tasser le Séphadex G 50 jusqu'à un volume de 1 ml, la colonne est conditionnée par addition de 100 μ l d'eau stérile et une deuxième centrifugation. Le produit de PCR est amené à un volume de 100 μ l avec de l'eau et déposé sur la colonne pré-conditionnée et pré-tassée. Le produit purifié est récupéré dans un tube de 15 ml après une centrifugation de 2 min à 1500 g.

Pour calculer le rendement de la purification des oligonucléotides marqués après passage sur la colonne de Séphadex G 50, les radioactivités liée et libre sont déterminées sur un échantillon de 2 μ l du produit purifié dans les mêmes conditions que celles décrites à la section 2.3.2. Le rendement de récupération de la radioactivité liée varie classiquement entre 30 et 80 % et la radioactivité libre est éliminée à un rendement de plus de 95 %.

2.3.4- Expérience de sélection par South-Western :

Pour sélectionner les oligonucléotides pouvant être liés par les protéines d'intérêt, on utilise la technique de South-Western (**Figure 11**). Les protéines de fusion purifiées MBP, MBP-Zn et MBP-ZNF74-II (Δ 1-176) sont séparées par électrophorèse sur gel de SDS-polyacrylamide (8 %, 29.2 : 0.8) puis transférées sur une membrane de nitrocellulose.

Dans un premier temps, deux membranes identiques sont renaturées pendant 16 h à 4 °C dans 10 ml de tampon de renaturation (Tris HCl 50 mM pH 7.5, KCl 100 mM,

Triton 1 %, Glycérol 10 %) en présence de zinc (+ ZnCl₂ 1 mM) ou en son absence (+EDTA 50 mM + dithiothreitol 10 mM).

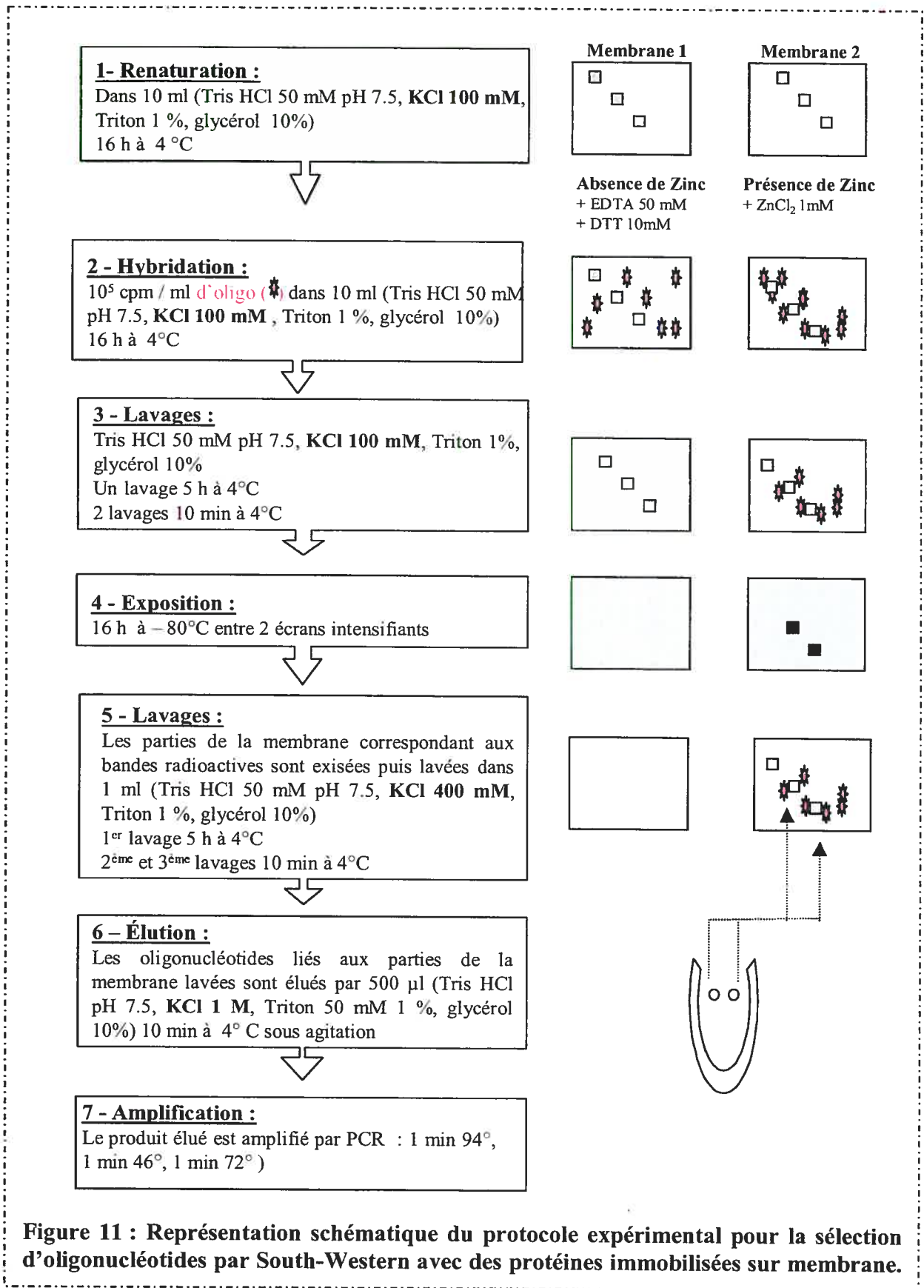
Dans un deuxième temps, chaque membrane est hybridée pendant 16 h à 4°C avec les oligonucléotides dégénérés (10⁵ cpm /ml) dans 10 ml de tampon de renaturation incluant 1 mM de ZnCl₂ ou 50 mM d'EDTA ainsi que 10 mM dithiothreitol. On procède par la suite à trois lavages à 4° C avec le tampon de renaturation. Le premier lavage est de 5 h et les deux derniers de 10 min. La révélation se fait par autoradiographie en appliquant un film Kodak sur la membrane pendant 16 h (exposition faite à - 80°C dans une cassette contenant deux écrans intensifiants) (Fisher Biotech L-Plus). Les parties de la membrane correspondant aux bandes positives peuvent ensuite être découpées. Chaque bande est lavée avec 500 µl de tampon de lavage (Tris HCl 50 mM pH 7.5, KCl 400 mM, Triton 1%, Glycérol 10 %) pendant 10 min à 4°C.

Dans un troisième temps, on procède à l'éluion des oligonucléotides liés aux protéines immobilisées sur la membrane avec 500 µl de tampon d'éluion (Tris HCl 50 mM, KCl 1 M, Triton 1 %, glycérol 10 %). L'éluion est réalisée sous agitation à température ambiante pendant 10 min. Les oligonucléotides sélectionnés contenus dans l'éluat sont récupérés pour une amplification par PCR et servent au cycle de purification suivant.

2.3.5- Amplification des oligonucléotides élués :

Dans un premier temps, un échantillon de 1 µl / 500 µl de l'éluat est amplifié par PCR dans 50 µl de tampon de PCR (Tris HCl 20 mM pH 8.4, KCl 50 mM) en présence de 400 ng de chaque amorce, de dNTPs 0.26 mM, de MgCl₂ 1.9 mM et de Taq DNA polymérase 2.5 UI (Invitrogen) pendant 30 cycles. Chaque cycle correspond à 1 minute de dénaturation à 94 °C, 1 minute d'hybridation des amorces avec la matrice à 46 °C et 1 minute d'élongation à 72°C.

Pour obtenir des oligonucléotides marqués utilisables pour un nouveau cycle de sélection par South-Western, un échantillon de 1 μ l de cette réaction de PCR est amplifié en présence de 40 μ ci du [α -P³²] dCTP (Perkin Elmer) dans une réaction de PCR de 25 μ l. La détermination de l'activité spécifique ainsi que la purification des oligonucléotides marqués est réalisée dans les mêmes conditions que celles décrites précédemment (Voir sections 2.3.2 – 2.3.3)



2.4 – Caractérisation des oligonucléotides sélectionnés :

2.4.1- Clonage des oligonucléotides sélectionnés :

Les oligonucléotides doubles brins obtenus après le troisième cycle de l'expérience de sélection par South-Western sont tout d'abord amplifiés par PCR dans les conditions décrites plus haut (voir section 2.3.5) puis clonés dans le site de restriction EcoRV du vecteur pBluescript SK⁺. La ligation des inserts non phosphorylés (oligonucléotides doubles brins amplifiés) et du vecteur phosphorylé (pBluescript SK⁺ préalablement digéré par EcoRV et purifié à partir d'un gel d'agarose 1%) se fait dans un volume de 10 µl en présence d'une unité de T4 DNA Ligase (InVitrogen), de 5 unités d'EcoRV pour rouvrir le vecteur qui se referme sans insert et de tampon ligation (Tris HCl 5 mM pH 7.6, MgCl₂ 10 mM, ATP 1 mM, DTT 1 mM, polyéthylène glycol-8000 5 % (w/v)) à 14 °C pendant 16 h. Un échantillon de 2 µl de la réaction de ligation est utilisé pour transformer 100 µl de bactéries compétentes DH5 α. L'ADN plasmidique est ensuite préparé à partir des colonies poussant sur pétri d'agar en présence d'ampicilline selon le protocole du Maniatis (Sambrook, 1987). La confirmation de la présence de l'insert dans les clones d'ADN plasmidique est obtenue après digestion enzymatique avec les enzymes de restriction ApaI et SpeI et séparation des produits de digestion sur gel de polyacrylamide non dénaturant à 15 % (taille attendue des inserts = 138 bp)

2.4.2 - Séquençage et analyse des clones :

Plusieurs clones contenant des inserts correspondant à des oligonucléotides sélectionnés par South-Western/PCR ont été séquencés. Les séquences obtenues ont été analysées par le logiciel Gibbs disponible à l'adresse suivante :

<http://bayesweb.wadsworth.org/gibbs/gibbs.html>. Ce logiciel sert à identifier un élément consensus commun par alignement de multiples séquences.

2.5 - Étude par retardement sur gel de la liaison de ZNF74 aux oligonucléotides sélectionnés:

La spécificité de liaison des oligonucléotides sélectionnés à la protéine bactérienne de fusion MBP-ZNF74-II (Δ 1-176) est testée par la technique de retardement sur gel. Le protocole utilisé est une adaptation de celui décrit par Garner et Revzin (1981).

2.5.1 – Préparation des oligonucléotides :

Les oligonucléotides utilisés pour l'expérience de retardement sur gel sont obtenus par PCR. Une première amplification est réalisée en utilisant un échantillon de la minipréparation d'ADN qui correspond à l'oligonucléotide cloné dans le vecteur SK⁺. Le produit amplifié est mis sur gel d'acrylamide afin d'éliminer le vecteur SK⁺ qui a servi de matrice. Une deuxième amplification est faite en utilisant le produit purifié à partir du gel d'acrylamide en présence de 20 μ ci du [α -P³²] dCTP (Perkin Elmer). La détermination de l'activité spécifique ainsi que la purification des oligonucléotides marqués sont réalisées dans les conditions décrites précédemment (Voir sections 2.3.2 et 2.3.3).

2.5.2- Expérience de retardement sur gel :

Dans un premier temps, des quantités équivalentes à 4 pmoles des protéines de fusion MBP (= 200 ng) et MBP-ZNF74-II (Δ 1-176) (= 377 ng) ont été incubées à 4°C pendant 30 min dans 20 μ l de tampon de liaison (Hepes 13.4 mM pH 8, DTT 2 mM, Nonidet P-40 0.1% (v/v), glycérol 3% (v/v) et 250 μ g de BSA) contenant 1.2 μ g de poly

(dI-dC) et 30 mM de KCl. Dans un deuxième temps, 50 000 cpm d'oligonucléotides marqués (activité spécifique de 2000 à 10000 cpm / ng) sont rajoutés et le mélange est réincubé pendant 30 min à température ambiante pour permettre la liaison des oligonucléotides aux protéines.

La séparation des complexes ADN-protéine est effectuée sur gel de polyacrylamide 5 % (29,5:0.5). La migration se fait à 20 mA à 4° C dans du TBE 0.25 X (Tris-borate 22.5 mM pH 8.3, EDTA 0.5 mM). Le gel est ensuite séché à 80°C et les complexes ADN-protéine sont visualisés à l'aide d'un film pour autoradiographie (Kodak X-Omat Blue XB-1).

2.6- Essais transcriptionnels de gène rapporteur :

Les essais transcriptionnels du gène luciférase ont été réalisés en utilisant le kit «Dual-Glo™ luciferase assay system» de Promega selon les recommandations du fournisseur.

Les cellules HEK293T ont été transfectées par les vecteurs d'intérêt tel qu'indiqué à la section 2.2.2.1. Les extraits cellulaires sont préparés et dosés selon les protocoles décrits aux sections 2.2.2.2 et 2.2.2.3.

2.6.1 - Détermination de l'activité transcriptionnelle du gène rapporteur luciférase :

Un échantillon de 75 µl de l'extrait cellulaire dilué à une concentration de 35µg / 75 µl de protéines totales est utilisé pour mesurer l'activité transcriptionnelle du gène rapporteur luciférase de la mouche à feu. L'essai se fait à température ambiante dans

des plaques de 96 puits. Dans un premier temps, on mesure l'activité de la luciférase de la mouche à feu («firefly luciferase») transcrite à partir du gène rapporteur en rajoutant 75 µl du réactif Dual-Glo™ Luciférase. Dans un deuxième temps, l'activité transcriptionnelle de la luciférase *Renilla* transcrite à partir du vecteur servant à mesurer l'efficacité de la transfection et donc à normaliser l'essai est mesurée après l'ajout de 75 µl du réactif Dual-Glo™ Stop & Glo®. À chaque étape, un temps d'incubation de 10 min est respecté avant d'effectuer la mesure. Le rapport entre l'activité luciférase de la mouche à feu et l'activité luciférase *Renilla* correspond à l'activité transcriptionnelle relative du gène rapporteur luciférase.

2.6.2 - Vérification de l'expression des protéines :

L'expression de la protéine HA-ZNF74-II dans les cellules HEK293T a été vérifiée par buvardage Western. Des échantillons d'extrait cellulaire correspondant à 30 µg de protéines totales ont été séparés sur gel de SDS-polyacrylamide 8 % puis transférés sur membrane de nitrocellulose. La membrane a été soumise à un buvardage Western avec un anticorps anti-HA.

CHAPITRE III
RÉSULTATS

3.1- Recherche de l'élément de réponse de ZNF74 :

3.1.1 –Sélection d'oligonucléotides dégénérés via une purification par South-Western combinée à une amplification par PCR :

En 1999, Guévremont (Guévremont, 1999) a montré par les expériences de retardement sur gel et Southern Blot que le domaine doigt de zinc de la protéine ZNF74 a une affinité pour l'ADN double brin. En 2001, Côté *et al* (Côté *et al.*, 2001) démontrent que le domaine KRAB de ZNF74 réprime la transcription lorsque fusionné au domaine de liaison à l'ADN de GAL4. Ces deux études ensemble indiquent que la protéine ZNF74 peut agir comme un facteur transcriptionnel potentiel. On sait que les facteurs transcriptionnels agissent de façon très spécifique au niveau de l'ADN et que cette spécificité est dictée en grande partie par la reconnaissance de séquences spécifiques au niveau de leurs cibles d'ADN. Afin de mieux caractériser la régulation transcriptionnelle qu'exerce ZNF74, il était essentiel d'identifier l'élément de réponse auquel il se lie en préalable à son action comme répresseur de la transcription.

Une première tentative d'identification de l'élément de réponse a été entamée par Guévremont en 1999 (Guévremont; 1999) par le biais d'une procédure cyclique qui combine une première étape de purification sur colonne d'affinité d'amylose et une deuxième d'amplification par PCR mais aucune séquence spécifique d'ADN n'a pu être identifiée. En 2001, Jheon *et al* (Jheon *et al.*; 2001) publie une étude dans laquelle il leur a été possible d'identifier l'élément de réponse de la protéine KRAB à multidoigt de zinc AJ18 en utilisant une technique reposant sur le même principe cyclique que la méthode décrite par Thiesen et Bach en 1990 (Thiesen et Bach., 1990) mais au lieu de faire une purification sur

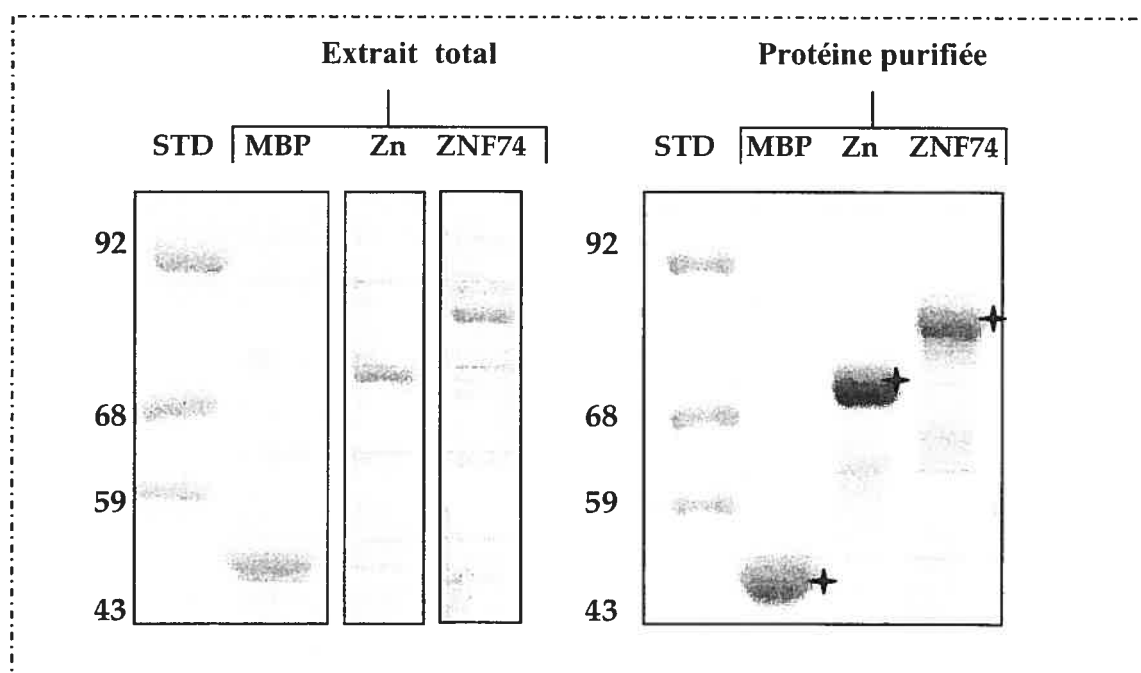
colonne d'affinité comme dans le cas de Guévremont (Guévremont; 1999), ils ont effectué une purification stringente par South-Western.

Cette étude encourageante nous a poussés à tenter d'identifier la ou les séquences consensus d'ADN reconnues par ZNF74 en utilisant une purification par South-Western couplée à une amplification par PCR tel que décrit à la section 2.3.

Pour le South-Western, on a immobilisé sur membrane de nitrocellulose des protéines de fusion entre MBP et la partie doigts de zinc de ZNF74 (MBP-Zn) ou une version plus longue de ZNF74 (MBP-ZNF74-II Δ 1-176) exprimées dans des bactéries et purifiées sur colonne d'amylose. On a choisi de ne pas utiliser la protéine complète ZNF74-II (1-643) car elle est dégradée en large proportion dans les bactéries (résultats non présentés). L'incubation des membranes a été réalisée avec un mélange d'oligonucléotides de 68 bp dégénérés en 16 positions et marqués radioactivement avec du α -P³². L'identité ainsi que l'intégrité des protéines de fusion MBP-Zn, MBP-ZNF74-II (Δ 1-176) et MBP seule (la protéine contrôle) ont pu être confirmées sur gel de SDS-polyacrylamide à la **Figure 12**. Les protéines purifiées ont été dosées par la suite par rapport à une gamme étalon de BSA sur gel de SDS-polyacrylamide 8 % coloré au bleu de Coomassie (**Figure 13**).

Figure 12 : Analyse des protéines de fusion bactériennes MBP, MBP-Zn et MBP-ZNF74-II (177-643) par électrophorèse sur gel de SDS-polyacrylamide.

Les extraits totaux bactériens ainsi que les protéines de fusion bactériennes purifiées MBP, MBP-Zn et MBP-ZNF74-II (Δ 1-176) ont été séparés par électrophorèse sur gel de SDS-polyacrylamide. Les protéines MBP, MBP-Zn et MBP-ZNF74-II (Δ 1-176) ont été purifiées par chromatographie d'affinité sur colonne d'amylose. Elles ont été détectées après coloration du gel au bleu de Coomassie. Les protéines purifiées sont identifiées par une étoile. Un échantillon d'extrait total de bactéries exprimant chaque protéine de fusion et correspondant à 1/15000 de la culture initiale de 300 ml a été chargé sur le gel. Le standard de masse moléculaire (STD) contient les protéines suivantes : phosphorylase B (92 kDa), BSA (68 kDa), catalase (59 kDa), albumine d'œuf de poulet (43 kDa). La masse moléculaire des protéines du standard est indiquée à gauche du gel.



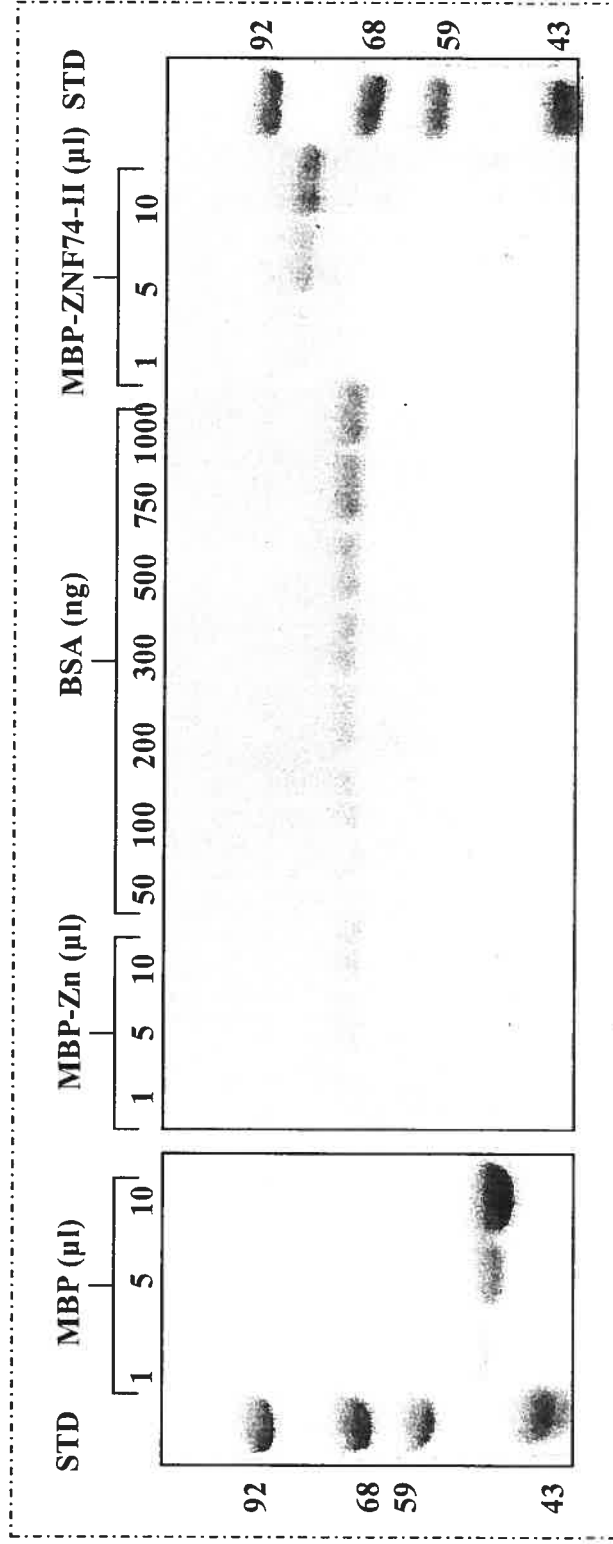
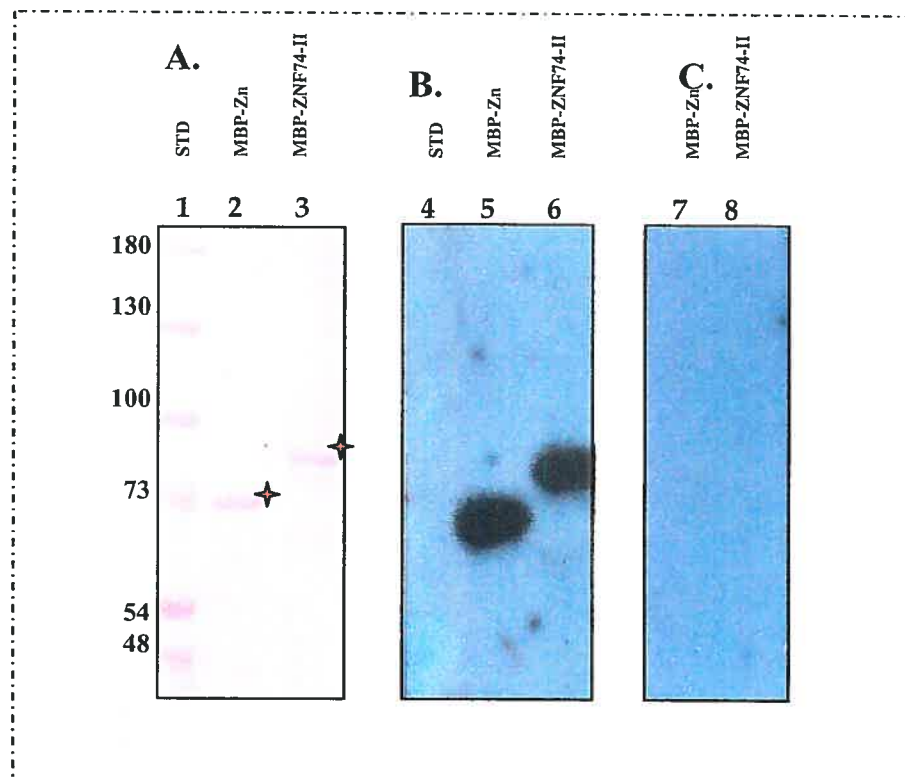


Figure 13: Dosage des protéines de fusion bactériennes purifiées MBP, MBP-Zn et MBP-ZNF74-II (Δ 1-176) par Coomassie. Les protéines bactériennes de fusion MBP, MBP-Zn et MBP-ZNF74-II (Δ 1-176) purifiées sur colonne d'amylose ont été dosées sur gel de SDS-polyacrylamide (8 %; 29.2:0.8) coloré au bleu de Coomassie. Pour chaque protéine de fusion, 3 échantillons correspondant à 1/15000 (1 µl), 1/3000 (5 µl) et 1/1500 (10 µl) de la culture initiale de 300 ml ont été chargés sur gel et ont servi au dosage par rapport à la gamme étalon de BSA qui contient 7 points: 50, 100, 200, 300, 500 750 et 1000 ng. Le standard de masse moléculaire (STD) contient les protéines suivantes: phosphorylase B (92 kDa), BSA (68 kDa), catalase (59 kDa), albumine d'œuf de poulet (43 kDa). La masse moléculaire des protéines du standard est indiquée de chaque côté du gel.

Avant de procéder à l'étape d'hybridation par South-Western, l'intégrité des protéines MBP-Zn et MBP-ZNF74-II (Δ 1-176) (500 ng) transférées sur la membrane de nitrocellulose ainsi que l'efficacité du transfert ont été confirmées par une coloration de la membrane au ponceau (**Figure 14 A**) (Pistes 2 et 3 respectivement). L'hybridation des protéines de fusion immobilisées sur membrane avec le mélange d'oligonucléotides dégénérés et marqués radioactivement a été effectuée dans deux conditions expérimentales soit en présence de zinc (**Figure 14 B**) ou en absence de zinc (grâce à l'ajout d'un chélateur du zinc) (**Figure 14 C**). Les rinçages des membranes ont été réalisés dans des conditions stringentes en présence d'une haute concentration en sel (KCl = 400 mM). Tel que démontré à la **Figure 14**, les oligonucléotides dégénérés et marqués au α -P³² sont liés spécifiquement par les protéines de fusion MBP-Zn (piste 5) et MBP-ZNF74-II (Δ 1-176) (pistes 6) seulement en présence du zinc. L'ajout d'un chélateur de zinc dans des conditions réductrices (EDTA et DTT) au mélange réactionnel prévient complètement la liaison des oligonucléotides aux protéines d'intérêt (pistes 7 et 8) ce qui indique que l'intégrité du domaine doigt de zinc est requise pour assurer la liaison à l'ADN. Dans ces mêmes conditions expérimentales, ni le standard de masse moléculaire (piste 4), ni la protéine MBP seule (résultats non présentés) ne lient les oligonucléotides dégénérés.

Les oligonucléotides purifiés à partir de ce premier cycle de South-Western ont été amplifiés par PCR et ont servi par la suite au deuxième cycle. L'expérience a été répétée pour un total de trois fois et le produit final du troisième cycle a été amplifié puis cloné.

Figure 14: Les protéines MBP-Zn et MBP-ZNF74-II (Δ 1-176) lient les oligonucléotides doubles brins dégénérés en présence du zinc dans l'expérience de purification par South-Western. Des quantités équivalentes à 500 ng des protéines de fusion purifiées MBP-Zn et MBP-ZNF74-II (Δ 1-176) ont été chargées en triplicata sur gel de SDS-polyacrylamide (8 %; 29.2:0.8) puis transférées sur membrane de nitrocellulose. (A) - Le premier 1/3 de la membrane a été coloré au ponceau et confirme l'intégrité des protéines transférées MBP-Zn (piste 2) et MBP-ZNF-II (Δ 1-176) (piste 3) indiquées par des étoiles. La masse moléculaire des protéines du standard est indiquée à gauche du gel. (B) - Le deuxième 1/3 de la membrane a été hybridé avec 10^5 cpm / ml d'oligonucléotides dégénérés marqués au α -P³² (activité spécifique = 7169 cpm / ng) en présence du zinc (+ZnCl₂). (C) - Le dernier 1/3 de la membrane a été hybridé dans les mêmes conditions que (B) mais en absence de zinc (+ EDTA + DTT).



3.1.2 – Analyse des séquences des oligonucléotides sélectionnés:

Les oligonucléotides sélectionnés par trois cycles de South-Western/PCR ont été clonés dans le site EcoRV du vecteur SK⁺. Parmi les clones obtenus ayant intégré un insert, 35 ont été séquencés et présentent des séquences différentes. Dans le but de déterminer un élément commun à l'intérieur de ce groupe de séquences oligonucélotidiques, les séquences des deux brins de ces 35 clones ont été analysées par le logiciel bioinformatique Gibbs. Ce logiciel utilise un algorithme pour localiser des éléments communs au sein d'une collection de séquences d'ADN. Le logiciel calcule la fréquence de la présence d'un nucléotide pour chaque position de la séquence et aligne par la suite toutes les différentes séquences selon le profil optimal identifié. Les oligonucléotides sélectionnés après le troisième cycle de purification ont été analysés par le logiciel Gibbs après séquençage. L'analyse est effectuée plusieurs fois en modifiant à chaque fois la taille du consensus recherché (3 à 15 nucléotides dans cette étude). Le motif identifié ainsi que l'alignement des 35 séquences sont représentés à la **Figure 15**. La séquence de 6 nucléotides identifiée par le logiciel Gibbs correspond au motif **APuATTC/G**.

Figure 15 : Analyse par le logiciel Gibbs des séquences des 35 clones purifiés:

Les 35 séquences obtenues ont été analysées par le logiciel bioinformatique Gibbs. Le motif de 6 nucléotides identifié correspond à la séquence consensus **APuATTC/G**. Pour chaque position du motif, le pourcentage du nucléotide correspondant est indiqué. Les séquences ont été alignées selon le motif identifié. Les nucléotides identiques au motif ont été surlignés en vert. Les oligonucléotides en majuscules correspondent aux 16 nucléotides dégénérés alors que ceux en minuscule correspondent à la région flanquante conservée. (S) sens, (R) complémentaire inverse. Les clones analysés ultérieurement par retardement sur gel et/ou dans des essais de gène rapporteur sont surlignés en jaune.

3.1.3 – Test par retardement sur gel de la spécificité de liaison de ZNF74 aux oligonucléotides sélectionnés:

Ayant pour but de déterminer si le motif identifié par le logiciel Gibbs correspond à un élément de réponse reconnu par la protéine ZNF74, la liaison de quelques séquences contenant ce motif à la protéine ZNF74 présente dans des extraits bactériens a tout d'abord été confirmée par retardement sur gel.

En préalable, les quantités de protéines de fusion présentes dans les extraits bactériens totaux ont été évaluées par buvardage Western en utilisant un anticorps anti-MBP avec comme courbe standard de référence des quantités connues de la protéine MBP (**Figure 16**).

Par la suite, les insertions des clones 98, 76, 67, 72, 99, 73 et 31 (**Figure 15**) qui contiennent ou non le motif **APuATTC/G** dans leurs séquences dégénérées ont été marquées au α -P³² et incubées séparément avec les extraits bactériens totaux contenant 3 pmoles de la protéine MBP-ZNF74-II (Δ 1-176) ou de la protéine contrôle MBP seule. La séparation des complexes ADN-protéine a été faite sur gel de polyacrylamide 5 %. La présence de bandes retardées indique que la séquence 76 (**Figure 17**; pistes 14 et 16) ainsi que les séquences 99, 73 et 31 (**Figure 18**; pistes 7, 10 et 13) lient la protéine ZNF74-II (Δ 1-176). Une bande plus faible est observée avec la séquence 67 (**Figure 17**; piste 18). La spécificité des complexes obtenus est confirmée par l'ajout de chacun des oligonucléotides compétiteurs non marqués qui en excès de 30 fois par rapport à la sonde correspondante marquée prévient complètement l'apparition de la bande retardée (**Figure 17**; pistes 15, 17 et 19). Contrairement aux séquences 31, 67, 73, 76 et 99, la séquence 98 qui dérive du consensus par un nucléotide (**Figure 17**; piste 12) et la séquence 72 qui ne contient pas le consensus (**Figure 18**; piste 5) ne lient pas la protéine MBP-ZNF74-II (Δ 1-176). Par contre, le contrôle positif (842) correspondant au mélange initial d'oligonucléotides

dégénérés forme un complexe retardé avec la protéine ZNF74-II (Δ 1-176) (**Figure 17**; piste 10). D'autre part, il apparaît que la présence des complexes retardés est dépendante de ZNF74 puisqu'aucune bande retardée n'est observée avec la protéine contrôle MBP seule (**Figure 17**; pistes 5, 6, 7, 8 et 9) et (**Figure 18**; pistes 3, 6, 9 et 12).

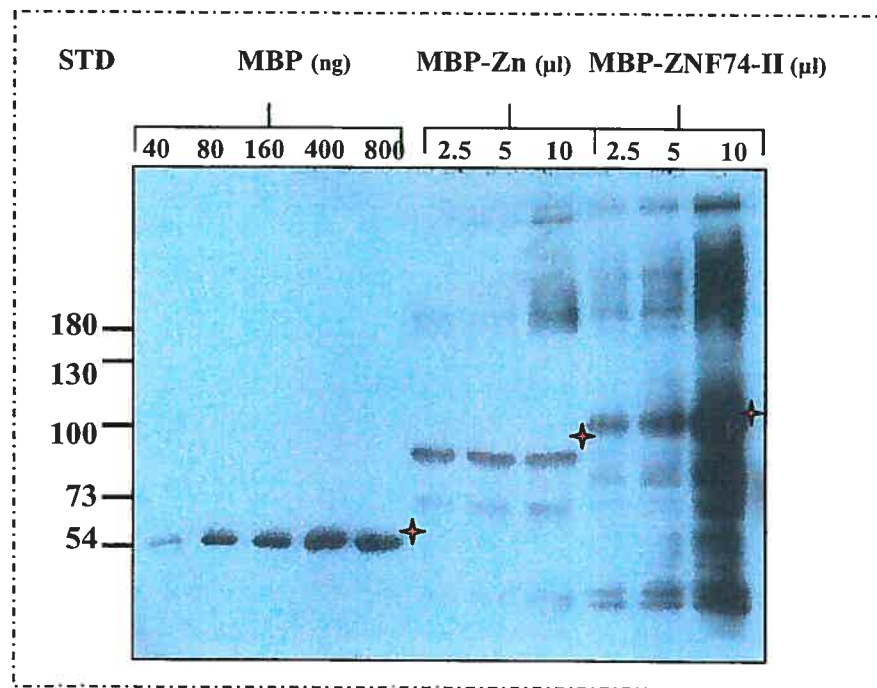


Figure 16 : Dosage des protéines bactériennes MBP, MBP-Zn et MBP-ZNF74-II (Δ 1-176) présentes dans les extraits utilisés dans les expériences de retardement sur gel. La concentration des protéines de fusion, MBP-Zn et MBP-ZNF74-II (Δ 1-176), présentes dans les extraits bactériens totaux a été évaluée par buvardage Western avec un anticorps anti-MBP en comparaison à des extraits bactériens contenant des quantités connues de MBP (40, 80, 160, 400 et 800 ng) (la protéine MBP représente plus que 90 % des protéines des extraits bactériens et est préalablement dosée sur gel SDS-polyacrylamide coloré au bleu de Coomassie par rapport à une gamme étalon de BSA). Pour chaque protéine de fusion, trois échantillons de l'extrait bactérien total correspondant respectivement à 1/8000 (2.5 μ l), 1/4000 (5 μ l) et 1/2000 (10 μ l) de la culture initiale de 300 ml ont été séparés sur gel SDS-polyacrylamide 8% puis transférés sur membrane de nitrocellulose. Après incubation avec les anticorps primaires et secondaires, la chemiluminescence libérée a été évaluée par le logiciel «Quantity One» avec l'option «Fluor S Max» pour une analyse quantitative. L'exposition d'un film «Kodak» à la chemiluminescence est présentée ici. La masse moléculaire des protéines du standard en (kDa) est indiquée à gauche du gel. Les masses moléculaires apparentes des protéines MBP, MBP-Zn et MBP-ZNF74-II (Δ 1-176) sont respectivement 51 kDa, 80 kDa et 94 kDa.

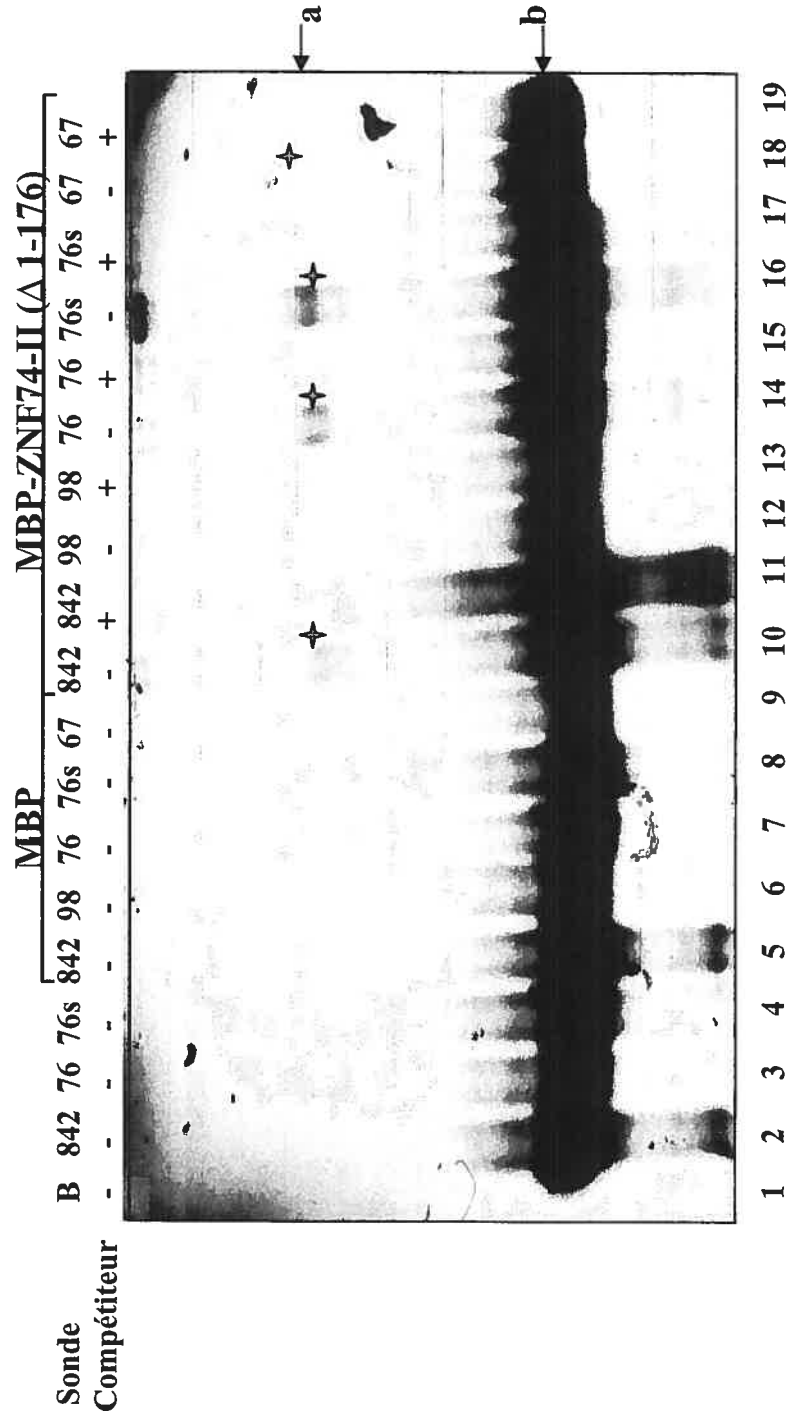


Figure 17 : Essai de retardement sur gel de la protéine MBP-ZNF74-II (Δ 1-176) avec les oligonucléotides sélectionnés (98, 76 et 67) Des quantités équivalentes à 3 pmoles des protéines de fusion d'origine bactérienne MBP (= 150 ng) et MBP-ZNF74-II (Δ 1-176) (= 282 ng) ont été incubées séparément avec les oligonucléotides sélectionnés 76, 98 et 67 amplifiés par PCR et marqués au α -P³² (50000 cpm de chaque oligonucléotide; activité spécifique = 4000 à 10000 cpm / ng). Le 76s est un oligonucléotide d'origine synthétique identique à l'oligonucléotide 76. L'oligonucléotide 842 correspond au mélange initial d'oligonucléotides dégénérés qui a servi au premier cycle de purification par South-Western. Les complexes ADN-protéine ont été analysés sur gel de polyacrylamide 5%. La spécificité de la liaison est évaluée par l'addition de chacun des oligonucléotides non marqués (un excès de 30 fois par rapport à la sonde marquée) (pistes 11, 15, 17 et 19). Les bandes retardées observées sont indiquées par des étoiles. (a) Complexes oligonucléotides marqués-ZNF74-II (Δ 1-176). (b) Oligonucléotides marqués non liés.

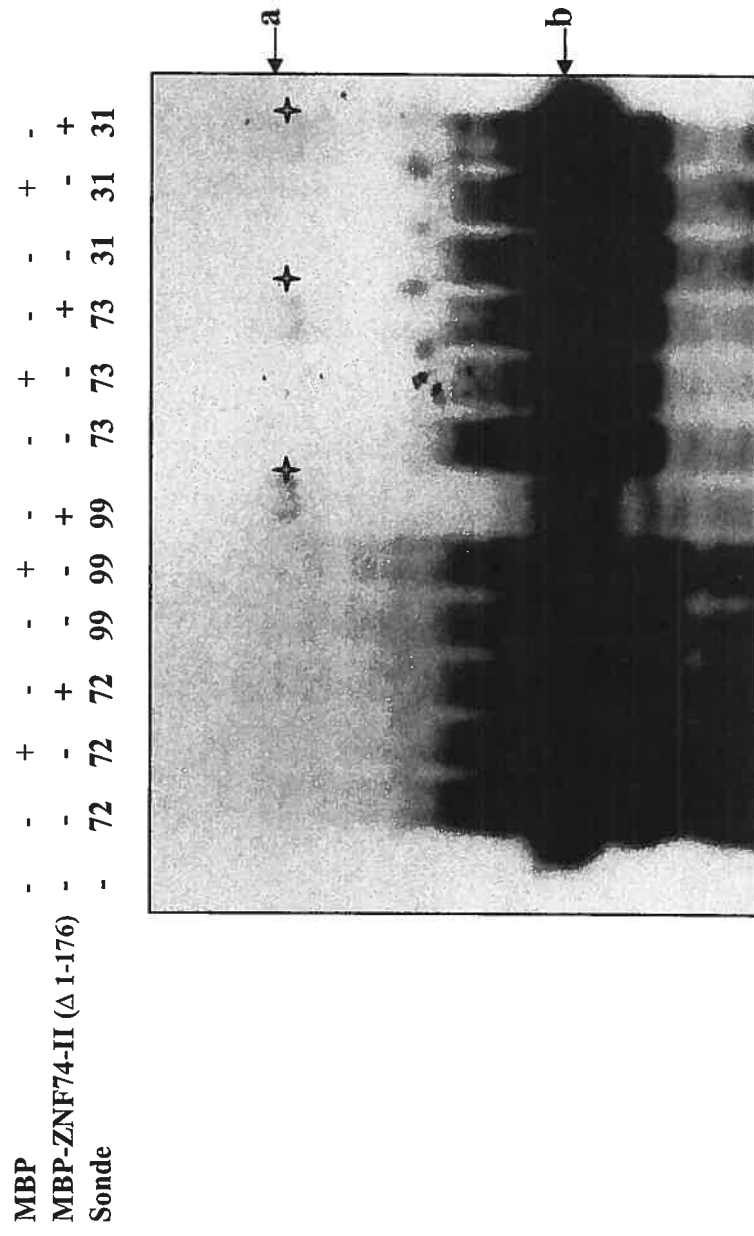


Figure 18: Essai de retardement sur gel de la protéine MBP-ZNF74-II (Δ 1-176) avec les oligonucléotides sélectionnés (72, 99, 73 et 31)

Des quantités équivalentes à 3 pmoles des protéines de fusion d'origine bactérienne MBP (= 150 ng) et MBP-ZNF74-II (Δ 1-176) (= 282 ng) ont été incubées avec les oligonucléotides sélectionnés 72, 99, 73 et 31 amplifiés par PCR et marqués au α -P³² (20000 cpm de chaque oligonucléotide; activité spécifique = 2000 à 4000 cpm / ng). Les complexes ADN-protéine ont été analysés sur gel de polyacrylamide 5 %. Les bandes retardées observées sont indiquées par des étoiles. (a) Complexes oligonucléotides marqués-ZNF74-II (Δ 1-176) (b) Oligonucléotides marqués non liés.

3.1.4 – Test par essai transcriptionnel de la fonction du motif identifié comme élément de réponse pour ZNF74:

Considérant que la protéine ZNF74-II lie des séquences contenant le motif **AA/GATTC/G** identifié par le logiciel Gibbs tel que confirmé par les expériences de retardement sur gel, nous avons testé si ce motif pouvait fonctionner comme un élément de réponse pour la protéine ZNF74. Pour cela, des essais transcriptionnels de gène rapporteur ont été réalisés en utilisant des cellules HEK293T transfectées de façon transitoire avec un gène rapporteur et des quantités croissantes du vecteur d'expression eucaryote pCGN exprimant ZNF74-II (pCGN-ZNF74). Le gène rapporteur correspond au vecteur pGL3 *contrôle* contenant le gène *luciférase* de la mouche à feu sous le contrôle du promoteur fort de SV40 avec en amont le motif identifié comme élément de réponse potentiel. Le vecteur pGL3 *contrôle* contient aussi une séquence «enhancer» de SV40 située en aval du gène de la luciférase et permettant de renforcer l'effet du promoteur.

Ainsi, trois constructions de type gène rapporteur ont été créées en incluant les séquences (76, 73, et 31) qui ont montré une capacité de liaison à la protéine ZNF74-II par retardement sur gel. Les trois constructions correspondent respectivement aux vecteurs *contrôle-76* (4X) qui contient la séquence 76 (CATCTA**AGATTG**TTAC) multimérisée 4 fois, pGL3 *contrôle-31* (1X) qui contient une seule copie de la séquence 31 (CATA**AAATTC**GCCAACA) et enfin pGL3 *contrôle-73* (1X) qui contient une seule copie de la séquence 73 (TTTCAG**AGATTCC**GCT). L'effet de la protéine ZNF74-II sur la transcription médiée par ces trois vecteurs a été analysé.

Tel qu'illustré à la **Figure 19**, la transcription du gène *luciférase* de la mouche à feu sous le contrôle du promoteur de SV40 qui contient en amont le multimère 76 (4X) n'est pas affectée par l'ajout de quantités croissantes du vecteur pCGN-ZNF74-II qui encode la protéine HA-ZNF74-II. La baisse de transcription de 11 à 17 % observée par rapport au contrôle exprimant le gène rapporteur et le vecteur pCGN (vecteur vide) n'est pas significative car une même baisse est observée avec le vecteur pCGN-ZNF74-II anti-sens qui ne code pour aucune protéine (résultats non présentés).

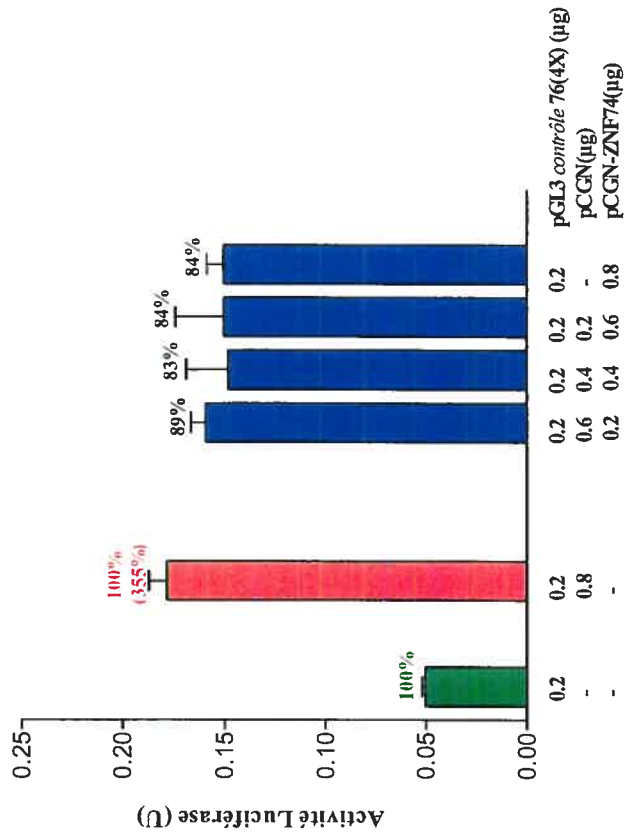
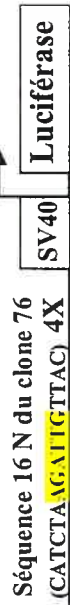
Les résultats obtenus avec les vecteurs pGL3 *contrôle-31* (1X) (**Figure 20; A**) et pGL3 *contrôle-73* (1X) (**Figure 20; B**) ne montrent pas eux non plus une baisse de répression suite à l'addition de quantités croissantes du vecteur d'expression protéique pCGN-ZNF74-II.

En parallèle, dans les mêmes conditions expérimentales le contrôle positif correspondant au vecteur pRSV-GAL4DBD-ZNF74-II qui exprime la protéine ZNF74-II en fusion avec le domaine de liaison à l'ADN de GAL4 réprime fortement la transcription du gène *luciférase* sous le contrôle du promoteur de SV40 qui contient en amont l'élément de réponse de GAL4 (AGGGTATATAATG)⁵ multimérisé 5 fois. Une baisse significative de l'activité transcriptionnelle de 83 % est obtenue (**Figure 19 B**). Cette baisse n'est pas observée lorsqu'on utilise le vecteur pRSV-GAL4DBD vide (résultat non montré). Dans les études de caractérisation de l'activité répressive de ZNF74 (Coté et al., 2001), seul le domaine KRAB de la protéine ZNF74 avait été fusionné avec le GAL4DBD et testé pour son activité répressive. Cette étude démontre que ZNF74 fonctionne aussi dans le contexte de la protéine complète, qui inclut ses deux domaines fonctionnels (KRAB et multidoigt de zinc), comme un puissant répresseur transcriptionnel.

Quoique les trois séquences nommées (76, 73 et 31) aient la capacité de lier la protéine ZNF74 dans des essais de retardement sur gel, l'absence d'un effet fonctionnel de ZNF74 sur l'activité transcriptionnelle des gènes rapporteurs incluant ces séquences suggère qu'elles ne peuvent pas fonctionner comme des éléments de réponse efficaces.

Figure 19 : L'effet de pCGN-ZNF74-II sur la transcription médiée par un multimère de la séquence 76 placé en amont du promoteur de SV40. (A) Les cellules HEK293T ont été transfectées de façon transitoire avec 0.2 μ g du vecteur rapporteur pGL3 *contrôle-76* (4X) qui contient la séquence 76 (CATCTAAGATTGTTAC) multimérisée 4 fois en amont du promoteur de SV40 avec 0.8 μ g de vecteur pCGN vide ou bien avec des quantités croissantes (0.2, 0.4, 0.6, 0.8 μ g) du vecteur d'expression protéique pCGN-ZNF74-II qui exprime la protéine HA-ZNF74-II. Dans chaque cas, on a complété avec le vecteur pCGN vide jusqu'à une quantité totale correspondant à 0.8 μ g. L'activité du gène rapporteur codant pour la luciférase de la mouche à feu a été normalisée en fonction de l'activité de la luciférase *Renilla* (voir Matériel et Méthodes). Les essais ont été réalisés en utilisant une quantité fixe (35 μ g) de protéines totales provenant des différents extraits cellulaires. Les résultats sont présentés en pourcentage par rapport à l'activité luciférase obtenue dans des extraits de cellules transfectées avec 0.2 μ g du vecteur rapporteur ainsi que 0.8 μ g du vecteur pCGN vide. Les pourcentages représentés par rapport aux extraits contenant seulement le vecteur rapporteur sont indiqués entre parenthèses. **(B)** À titre de contrôle, les cellules HEK293T ont été transfectées de façon transitoire avec 0.2 μ g du vecteur rapporteur pGL3 *contrôle GAL4 RE* (5X) qui contient l'élément de réponse de la protéine GAL4 (AGGGTATATAATG)⁵ multimérisé 5 fois en amont du promoteur de SV40 avec ou sans le vecteur pRSV-GAL4DBD-ZNF74-II (0.2 μ g) qui exprime la protéine ZNF74-II en fusion avec le domaine de liaison à l'ADN de GAL4. Chaque point correspond à des transfections effectuées en triplicata dont l'activité luciférase a été déterminée en duplicata. L'activité luciférase présentée correspond à la moyenne des triplicatas. Les barres d'erreur représentent l'erreur standard à la moyenne. Ces résultats représentent une expérience typique et chaque point testé est représentatif d'un minimum de deux expériences indépendantes.

A



B

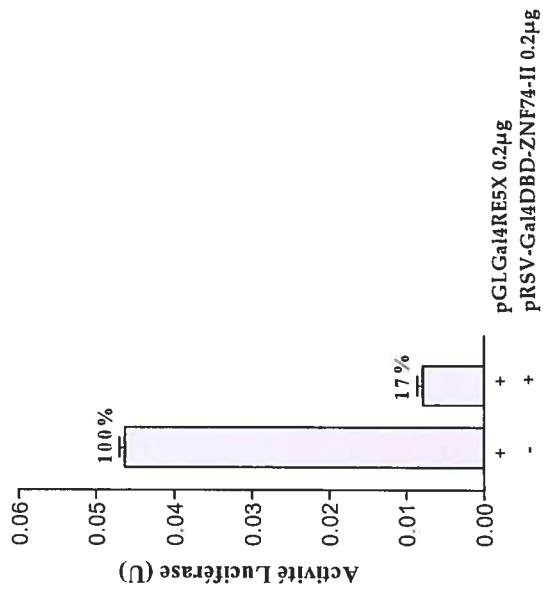
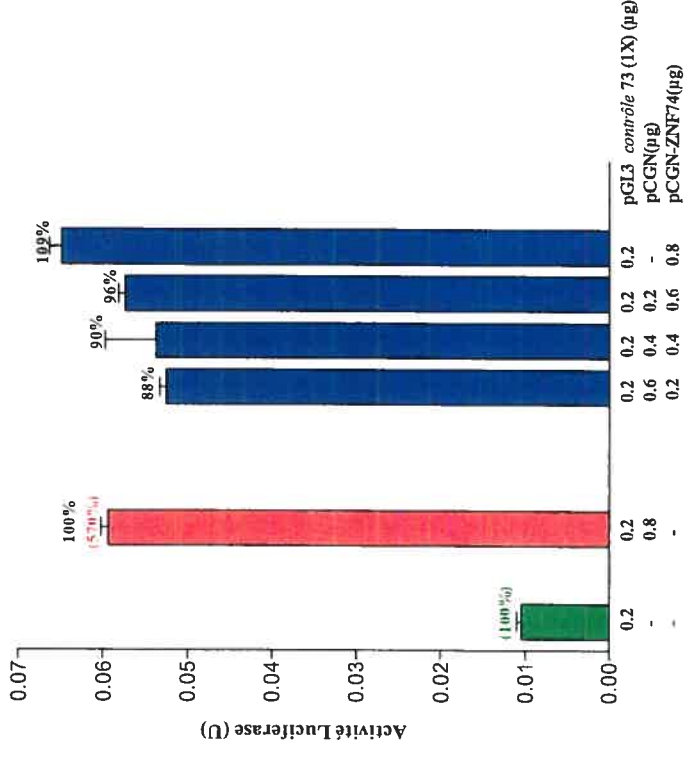
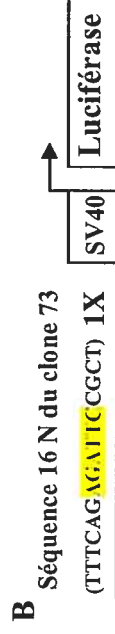
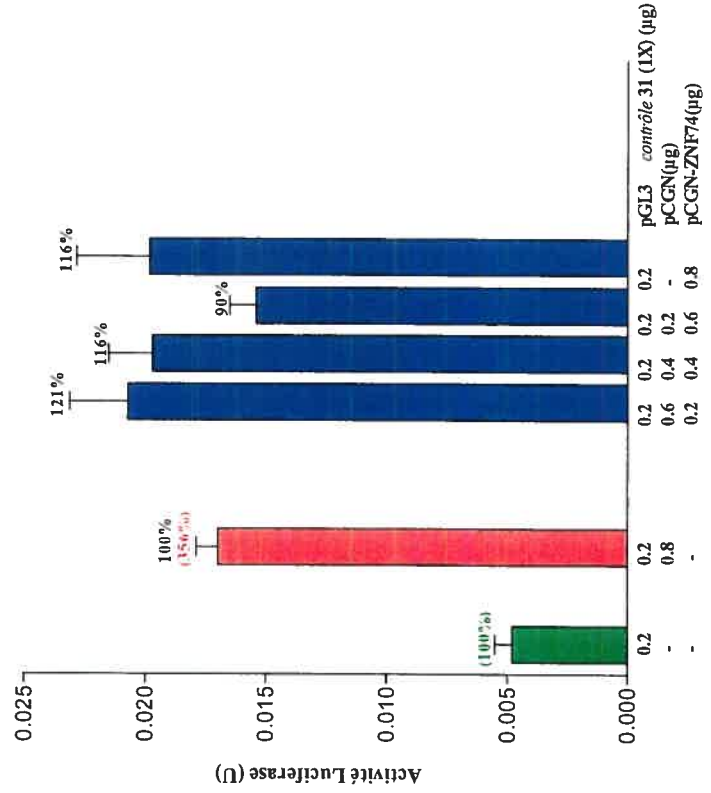
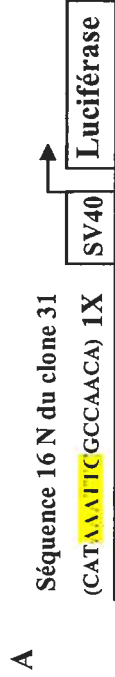


Figure 20 : L'effet de pCGN-ZNF74-II sur la transcription médiée par la séquence 31 ou la séquence 73 clonées en amont du promoteur de SV40.

Les cellules HEK293T ont été transfectées de façon transitoire avec 0.2 µg du vecteur rapporteur pGL3 *contrôle*-31 (1X) qui contient la séquence 31 (CATAAAATTTCGCCAACA) **(A)** ou la séquence 73 (TTTCAGAGATTCCGCT) **(B)** en amont du promoteur de SV40 avec 0.8 µg de vecteur pCGN vide ou bien avec des quantités croissantes (0.2, 0.4, 0.6, 0.8 µg) du vecteur d'expression protéique pCGN-ZNF74-II qui exprime la protéine HA-ZNF74-II. Dans chaque cas, on a complété avec le vecteur pCGN vide jusqu'à une quantité totale correspondante à 0.8 µg. L'activité du gène rapporteur codant pour la luciférase de la mouche à feu a été normalisée en fonction de l'activité de la luciférase *Renilla* (voir Matériel et Méthodes). Les essais ont été réalisés en utilisant une quantité fixe (35 µg) de protéines totales provenant des différents extraits cellulaires. Les résultats sont représentés en pourcentage par rapport à l'activité luciférase obtenue dans des extraits de cellules transfectées avec 0.2 µg du vecteur rapporteur ainsi que 0.8 µg du vecteur pCGN vide. Chaque point correspond à des transfections effectuées en triplicata dont l'activité luciférase a été déterminée en duplicata. L'activité luciférase présentée correspond à la moyenne des triplicatas. Les barres d'erreur représentent l'erreur standard à la moyenne.



3.2 - Recherche de l'élément de réponse de ZNF74 dans son propre promoteur :

L'expérience de purification South-Western/PCR n'ayant pas permis d'identifier l'élément de réponse de ZNF74, nous avons émis l'hypothèse que ZNF74 pourrait réguler son propre promoteur par un mécanisme d'autorégulation tel que déjà démontré pour plusieurs facteurs transcriptionnels incluant des multidoigts de zinc (Olson *et al.*, 2005 ; Mahatan *et al.*, 1999 ; Vassen *et al.*, 2005). Nous avons abordé l'étude du promoteur de ZNF74 dans le but d'y identifier des éléments cibles qui lient potentiellement la protéine ZNF74 et permettent la régulation négative du promoteur.

Avant de procéder à la recherche de séquences cibles dans le promoteur de ZNF74, une analyse de la séquence du promoteur ainsi qu'une étude fonctionnelle de son activité transcriptionnelle ont tout d'abord été effectuées. Ensuite, l'effet de ZNF74 sur son propre promoteur a été testé.

3.2.1 – Analyse de la séquence du promoteur de ZNF74 :

Quoique les éléments *cis* requis pour une expression spatio-temporelle spécifique puissent se situer à plusieurs kilobases du site d'initiation de la transcription, les séquences promotrices sont habituellement situées à moins de 2 kilobases en amont du site d'initiation de la transcription. On a donc limité l'analyse du promoteur de ZNF74 à un fragment d'environ 2 kb. L'analyse de la séquence du promoteur de ZNF74 qui correspond à la région (-1714/+301) par rapport au site d'initiation de la transcription a été réalisée en utilisant le logiciel MatchTM disponible à l'adresse suivante (<http://www.gene-regulation.com/cgi-bin/pub/programs/match/bin/match.cgi>). Ce logiciel permet d'identifier

la présence de sites de liaison potentiels pour des facteurs transcriptionnels au sein d'une séquence d'ADN donnée. Les résultats obtenus sont présentés à la **Figure 21**. On note que le promoteur de *ZNF74* ne contient pas de boite TATA typique proche du site d'initiation de la transcription indiqué par +1. Les six sites potentiels de liaison pour les facteurs transcriptionnels suivants : HNF1 («**H**epatic Nuclear Factor 1»), ARNT («**AhR** Nuclear Translocator homodimers»), USF («**U**pstream Stimulatory Factor»), C/EBP («**C**CAAT/**E**nhancer Binding Protein»), E2F (**E**2 promoter binding Factor) et le proto-oncogène c-Rel ont été identifiés par le logiciel MatchTM tel qu'indiqué à la **Figure 21**.

3.2.2 – Analyse fonctionnelle du promoteur de ZNF74

Dans le but de caractériser l'activité transcriptionnelle du promoteur de *ZNF74*, des essais de gène rapporteur ont été réalisés en utilisant des cellules HEK293T transfectées de façon transitoire avec des constructions contenant des régions de différentes longueurs du promoteur de *ZNF74*. La version la plus longue contient un fragment qui va de -1429 à +49 (-1429/+49) par rapport au site d'initiation de la transcription identifié comme (+1) et la version la plus courte contient le fragment -232/+49. Dans toutes les constructions, la partie située en aval du site d'initiation de la transcription est de 49 bp sauf pour une construction préparée avec 97 bp en aval qui correspond à la région -399/+97. Ces régions du promoteur ont été clonées dans deux types de vecteurs rapporteurs permettant l'expression de la luciférase de la mouche à feu, pGL3 *basic* et pGL3 *enhancer*. Le vecteur pGL3 *basic* ne contient ni promoteur ni élément «enhancer» et le vecteur pGL3 *enhancer* contient seulement un élément «enhancer» en aval du gène *luciférase*. Dans les essais de gène rapporteur, l'expression de la luciférase de la mouche à feu dépend de l'insertion en amont du gène de la luciférase dans le vecteur pGL3 *basic* d'un promoteur fonctionnel et dans la bonne orientation. Le vecteur pGL3 *enhancer* est aussi souvent utilisé pour vérifier la fonctionnalité d'un promoteur car la présence d'un «enhancer» en conjonction avec un promoteur résulte souvent en une activité augmentée du gène de la luciférase.

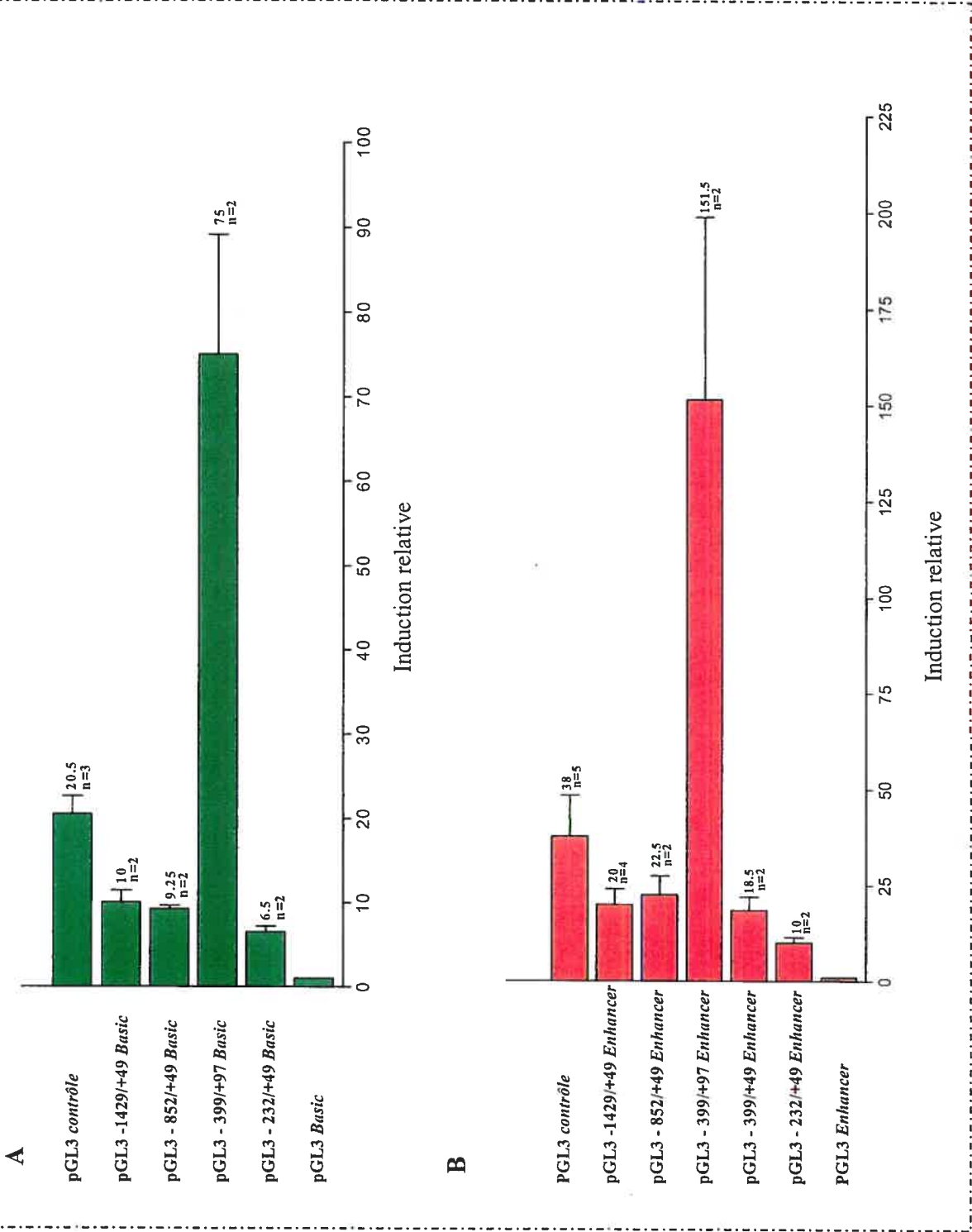
Une première série d'essais de gène rapporteur avec les différentes constructions du promoteur contenues dans le pGL3 *basic* a été effectuée. Les régions -1429/+49 et -852/+49 sont toutes deux capables d'activer la transcription à des niveaux comparables soit environ 10 fois par rapport au pGL3 *basic* vide qui constitue le contrôle négatif sans aucune région promotrice. Il apparaît qu'il n'y a pas d'éléments majeurs influençant la

transcription dans les conditions de l'essai entre -852 et -1429. La région -232/+49 du promoteur permet une induction de l'activité transcriptionnelle du gène *luciférase* légèrement inférieure à celle obtenue avec la construction -852/+49 (65%). Ceci indique que la majorité des éléments importants pour l'activation de la transcription se trouvent en aval de la région -232/+49 bp. De façon notable, des éléments situés après le site d'initiation de la transcription entre +49 et +97 sont requis pour une transcription de 75 fois supérieure au contrôle négatif sans promoteur tel qu'observé avec la construction -399/+97. Ainsi, cette construction fonctionne comme un promoteur fort puisqu'elle active presque 4 fois mieux la transcription que le vecteur pGL3 *contrôle* qui contient le promoteur fort et l'«enhancer» de SV40 (**Figure 22; A**).

Une deuxième série d'essais de gène rapporteur a été réalisée avec les différentes régions du promoteur de *ZNF74* contenues dans le vecteur pGL3 *enhancer* qui contient l'«enhancer» de SV40 en aval du gène luciférase. On n'observe en fait aucun changement dans l'induction relative de la transcription générée par les différents fragments du promoteur de *ZNF74* qu'ils soient clonés dans le vecteur pGL3 *basic* ou le pGL3 *enhancer* par rapport au vecteur pGL3 *contrôle* contenant le promoteur et l'«enhancer». Il est intéressant de constater sur cette figure que la région -399/+97 permet une activation de la transcription 8 fois plus élevée que la construction -399/+49 soulignant l'importance d'éléments *cis* situés entre +49 et +97 pour l'activation du promoteur de *ZNF74*.

Figure 22 : Analyse de l'activité transcriptionnelle du promoteur de ZNF74

L'activité du gène rapporteur *luciférase* de la mouche à feu, normalisée par l'activité du gène *luciférase Renilla* ainsi que par la concentration protéique totale, a été mesurée dans les différents extraits cellulaires. Les cellules HEK293T ont été transfectées de façon transitoire avec 0.5 µg des constructions pGL3 basic (**A**) ou pGL3 *enhancer* (**B**). Dans chaque cas, 0.05 µg du vecteur de normalisation pHRLUC a été transfecté parallèlement. La quantité totale d'ADN transfecté est amenée à 1 µg en complétant avec le vecteur vide pGEM qui n'affecte pas la transcription. L'activité luciférase des différents gènes rapporteurs représente la moyenne de *n* expériences indépendantes tel qu'indiqué sur la figure. Les barres d'erreur représentent l'erreur standard à la moyenne. Dans toutes les expériences, chaque point correspond à des transfections effectuées en triplicata dont l'activité luciférase a été déterminée en duplicata. Les niveaux d'induction présentés correspondent à l'activité luciférase comparée à celle du vecteur pGL3 *basic* vide (**A**) ou à celle du vecteur pGL3 *enhancer* vide (**B**).



3.2.3 – Recherche de séquences cibles de ZNF74 au sein de son

promoteur :

Dans le but d'identifier des séquences cibles potentielles pour la protéine ZNF74 au sein de son propre promoteur, des essais transcriptionnels de gène rapporteur ont été effectués pour tester si la protéine ZNF74 pouvait exercer une régulation de la transcription en agissant sur son propre promoteur.

On a tout d'abord testé la région la plus longue du promoteur $-1429/+49$ que l'on a clonée dans le vecteur pGL3 *enhancer* en amont du gène *luciférase*. Il apparaît cependant que des quantités croissantes de pCGN-ZNF74 codant pour la protéine HA-ZNF74 n'induisent pas une répression significative de la transcription du gène *luciférase* (**Figure 23**). Selon les expériences, une baisse allant jusqu'à 20 % par rapport au contrôle contenant le vecteur vide pCGN a pu être observée. Cependant ce type de variation ne semble pas spécifique à l'expression de ZNF74 à partir de pCGN-ZNF74. En effet, tel qu'observé à la figure 23, des quantités comparables du vecteur pCGN-ZNF74 anti-sens qui n'exprime pas la protéine ZNF74-II et d'autre part, un deuxième contrôle négatif qui consiste en la substitution du vecteur contenant des régions promotrices de *ZNF74* par le vecteur pGL3 *contrôle* qui contient le promoteur de SV40 induit des variations de même amplitude que celles observées avec pCGN-ZNF74 sur son promoteur (**Figure 23**).

Afin de vérifier que l'absence d'effet recherché n'est pas due à un déficit d'expression de la protéine HA-ZNF74-II dans les cellules HEK293T, un buvardage Western a été réalisé avec une partie des extraits cellulaires utilisés pour les essais transcriptionnels. Les résultats présentés à la **Figure 23; B** indiquent que le niveau d'expression de la protéine HA-ZNF74-II corrèle bien avec les quantités de vecteur pCGN-ZNF74-II transfectées dans

chaque cas. Ceci exclue la possibilité qu'une absence d'effet sur la transcription soit liée à une absence d'expression de la protéine ZNF74-II.

Considérant que la région +49 à +97 contient une région très importante pour l'activation de la transcription par le promoteur de ZNF74, nous avons aussi testé l'effet de la protéine ZNF74-II sur la région -399/+97 de son promoteur en comparaison à la région -399/+49. Aucun effet de la protéine.ZNF74 n'a pu être observé sur ces régions du promoteur (**Figure 24**). Les résultats des figures 23 et 24 suggèrent donc que le promoteur de ZNF74 n'est pas une cible pour la protéine ZNF74-II.

Figure 23: L'effet de pCGN-ZNF74-II sur la transcription médiée par la région -1429/+49 de son propre promoteur. (A) L'activité du gène rapporteur *Luciférase* de la mouche à feu, normalisée par l'activité du gène *Luciférase Renilla* ainsi que par la concentration protéique totale, a été mesurée dans les différents extraits cellulaires. (A) Les cellules HEK293T ont été transfectées de façon transitoire avec 0.5 µg des vecteurs rapporteurs pGL3 -1429/+49 *enhancer* qui contient la région -1429/+49 du promoteur de *ZNF74* en amont du gène *Luciférase* ou pGL3 *contrôle* qui contient le promoteur de SV40. On a aussi co-transfecté le vecteur pCGN contenant ou non l'ADNc de *ZNF74-II* dans sa forme sens, qui permet l'expression de la protéine HA-ZNF74-II, ou anti-sens (AS). La quantité du vecteur pCGN est maintenue à 0.45 µg en complétant au besoin pCGN-ZNF74-II avec le vecteur pCGN vide. À titre de contrôle, on a pris soin de vérifier qu'il n'y avait pas de différence significative de l'activité transcriptionnelle obtenue avec 0.45 µg et 0.33 µg de pCGN vide, cette dernière quantité étant équimolaire à 0.45 µg de pCGN-ZNF74-II. La quantité totale d'ADN transfecté est amenée au besoin à 1 µg en complétant avec le vecteur vide pGEM qui n'affecte pas la transcription. Chaque point a été fait en triplicata et l'activité luciférase correspond à la moyenne de ces triplicatas. Les barres d'erreur représentent l'erreur standard à la moyenne. Ces résultats représentent une expérience typique parmi trois expériences indépendantes. (B) Pour évaluer l'expression protéique de la protéine HA-ZNF74-II, des échantillons d'extraits cellulaires totaux correspondant à 30 µg de protéines totales ont été séparés par électrophorèse sur gel de SDS-polyacrylamide (8%; 29,2 :0,8) et soumis à un buvardage Western avec un anticorps anti-HA. La protéine HA-ZNF74-II est indiquée par une flèche. Deux bandes non-spécifiques classiquement observées avec l'anticorps anti-HA sont visibles au dessus de la protéine HA-ZNF74-II. Les masses moléculaires des protéines du standard en kDa sont indiquées à gauche du gel.

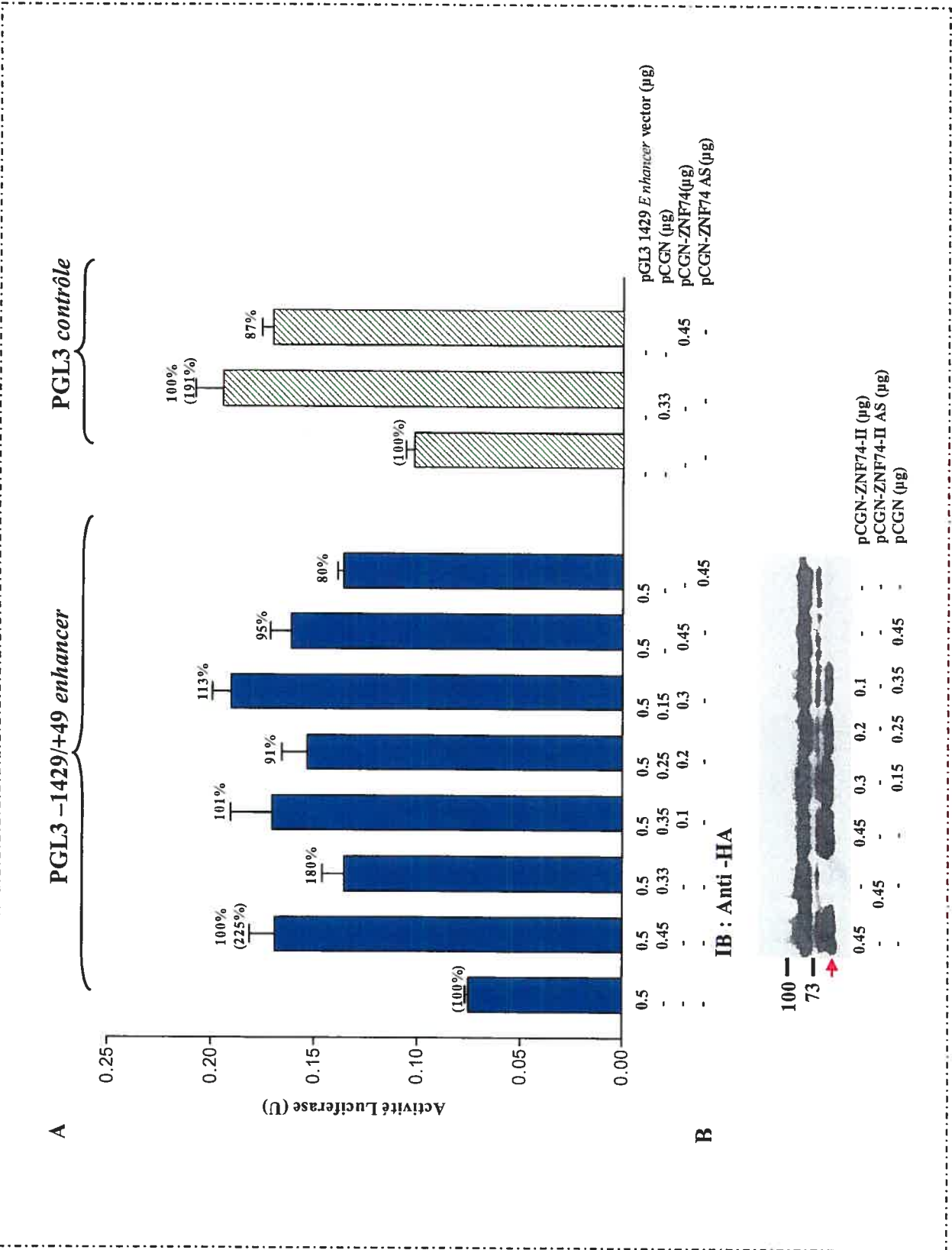
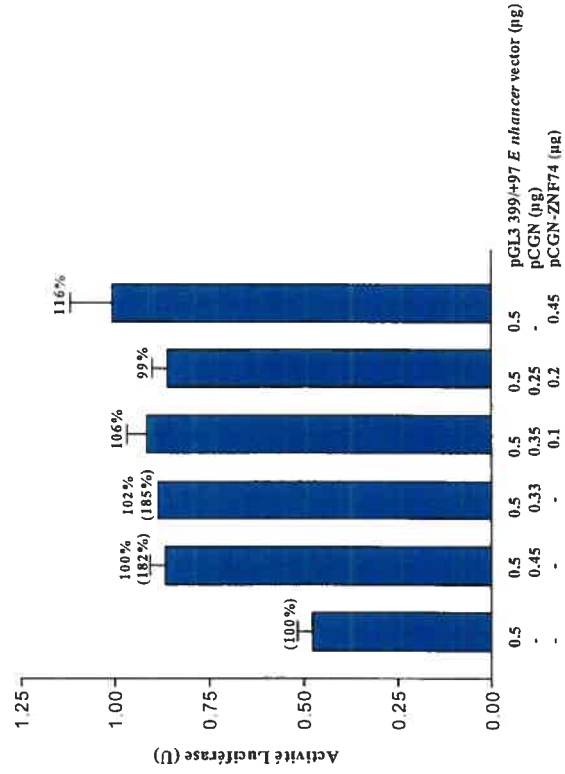


Figure 24 : L'effet de pCGN-ZNF74-II sur la transcription médiée par les régions -399/+49 et -399/+97 de son propre promoteur. (A) L'activité du gène rapporteur *Luciférase* de la mouche à feu, normalisée par l'activité du gène *Luciférase Renilla* ainsi que par la concentration protéique totale, a été mesurée dans les différents extraits cellulaires. Les cellules HEK293T ont été transfectées de façon transitoire avec 0.5 μ g du vecteur rapporteur pGL3 *Enhancer* qui contient la région -399/+97 (A) ou -399/+49 (B) du promoteur de *ZNF74* en amont du gène *Luciférase*. On a aussi co-transfecté le vecteur pCGN contenant ou non l'ADNc de *ZNF74-II* dans sa forme sens qui permet l'expression de la protéine HA-ZNF74-II, ou anti-sens (AS). À titre de contrôle, on a pris soin de vérifier qu'il n'y avait pas de différence significative de l'activité transcriptionnelle obtenue avec 0.45 μ g et 0.33 μ g de pCGN vide, cette dernière quantité étant équimolaire à 0.45 μ g de pCGN-ZNF74-II. La quantité totale d'ADN transfecté est amenée au besoin à 1 μ g en complétant avec le vecteur vide pGEM qui n'affecte pas la transcription. Chaque point a été fait en triplicata et l'activité luciférase correspond à la moyenne de ces triplicatas. Les barres d'erreur représentent l'erreur standard à la moyenne.

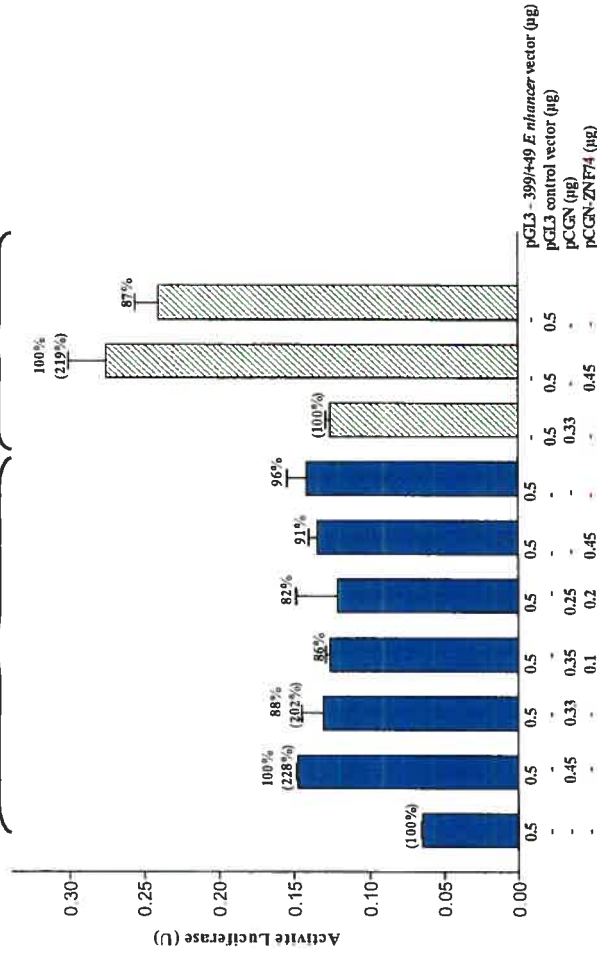
A

pGL3 -399/+97 enhancer



B

pGL3 -399/+49 enhancer pGL3 contrôle



CHAPITRE IV
DISCUSSION

Aubry et ses collaborateurs ont isolé en 1993 le gène codant pour la protéine ZNF74 (Aubry *et al.*, 1993) qui est capable d'une part, de lier l'ADN double brin via son domaine doigts de zinc (Guévremont, 1999) et d'autre part de réprimer la transcription lorsque son domaine KRAB est fusionné à un domaine hétérologue de liaison à l'ADN (Coté *et al.*, 2001). Quoique la capacité de la protéine à réprimer la transcription lorsqu'en fusion avec un domaine connu de liaison spécifique à l'ADN soit établie, les rôles que la protéine peut jouer au niveau physiologique n'ont pu être caractérisés à défaut de connaître la ou les séquences reconnues spécifiquement et les gènes cibles régulés.

Dans ce projet de recherche, le but principal était d'éclaircir les propriétés de répression transcriptionnelle de la protéine KRAB à multidoigt de zinc ZNF74 par le biais d'une caractérisation de la ou des séquences d'ADN spécifiques qui permettent son recrutement à l'ADN et lui permettent donc d'exercer son pouvoir répresseur. L'obtention de cette séquence ou de ces séquences devait nous orienter vers l'identification possible de ses gènes cibles et permettre la modulation de son activité transcriptionnelle dans le contexte de la protéine native plutôt que dans le contexte de fusion avec un domaine hétérologue de liaison à l'ADN. Notre stratégie de recherche était d'identifier initialement un motif consensus reconnu par ZNF74 en utilisant *in vitro* une méthode de sélection South-Western/PCR et puis de chercher ensuite dans les banques d'ADN, les gènes humains cibles qui contiennent l'élément de réponse dans leurs séquences de promoteurs. Une deuxième approche consistait à tenter d'identifier la présence possible de ces éléments de réponse au niveau de son propre promoteur *in vivo* par essais de gène rapporteur.

4.1 - À la recherche d'éléments de réponse de ZNF74 :

4.1.1 – Choix de la stratégie de recherche :

La majorité des études qui ont permis d'identifier des éléments de réponse pour des protéines KRAB à multidoigt de zinc ont utilisé un même principe de base, celui de la technique décrite par Thiesen et Bach (Thiesen et Bach., 1990). Cette méthode consiste en une procédure cyclique de sélection d'oligonucléotides double brin combinant une étape de purification par liaison à un support solide (filtre ou colonne d'affinité) et/ou par retardement sur gel avec une étape d'enrichissement du matériel sélectionné par amplification PCR.

En 1999, Guévremont a tenté d'identifier l'élément de réponse de la protéine ZNF74 en utilisant également une technique se basant sur le principe de Thiesen et Bach. Après plusieurs étapes de purification sur colonne d'affinité d'amylose et d'amplification par PCR, aucun élément consensus n'a pu être identifié parmi les séquences sélectionnées (Guévremont; 1999). En 2001, Jheon et ses collaborateurs ont réussi à identifier l'élément de réponse de la protéine KRAB multidoigt de zinc AJ18 (Jheon *et al.*, 2001) qui contient 11 doigts de zinc en tandem en suivant toujours le même protocole de Thiesen et Bach mais en effectuant la première étape de sélection par South-Western au lieu d'une purification sur colonne d'affinité comme l'a fait Guévremont. La mise en évidence de ces éléments nous a encouragés à opter pour la méthode de sélection South-Western/PCR dans le but d'identifier la séquence consensus d'ADN reconnue par la protéine ZNF74.

4.1.2 - Résultats de la recherche :

L'analyse des 35 séquences isolées après le troisième cycle de sélection a permis de mettre en évidence la présence d'un motif conservé de 6 nucléotides **APuATTC/G**. L'analyse

de la spécificité de liaison de ZNF74 été faite par retardement sur gel pour 6 des clones contenant ce motif. Cinq des six séquences sélectionnées qui contenaient le motif lient ZNF74 tandis qu'aucune liaison n'est détectée avec une septième séquence non reliée. Quoique l'intensité du signal obtenu avec certaines séquences sélectionnées (exemple : clone 76) soit supérieure à celle obtenue avec un mélange hétérogène d'oligonucléotides dégénérés en 16 positions en raison probablement d'une plus grande affinité des séquences pour ZNF74, certaines séquences correspondant au consensus ne semblent avoir qu'une affinité restreinte pour ZNF74 (exemple : clone 67). Considérant le fait que ZNF74 possède douze doigts de zinc et que chaque doigt a théoriquement la possibilité de lier trois nucléotides, il n'est pas impossible que des séquences adjacentes au consensus jouent aussi un rôle pour moduler l'interaction protéine-ADN. On a tenté de tester par retardement sur gel les 35 séquences sélectionnées pour leur capacité de liaison à la protéine ZNF74 afin de refaire une deuxième analyse plus spécifique par le logiciel Gibbs avec seulement les séquences qui manifesteraient des propriétés de liaison à ZNF74. Malheureusement, des difficultés techniques dans l'expérience de retardement sur gel et des contraintes de temps ont restreint nos analyses à 7 séquences sur un total de 35.

Par la suite, nous avons tenté de déterminer si les séquences identifiées *in vitro* qui se conformaient au consensus et étaient retardées sur gel en présence de ZNF74 pouvaient agir comme élément de réponse pour ZNF74 *in vivo*. Ces séquences ont été testées sous forme de monomères (oligo 76 –résultat non montré- et oligo 31 et 73) ou de multimères (oligo 76 multimérisé 4X). La multimérisation des éléments de réponse permet d'augmenter l'effet sur la transcription des facteurs qui s'y lient (Gebelein et Urrutia., 2001; Jheon *et al.*, 2001 ; Peng *et al.*, 2002). Il est à noter que seule la séquence 76 a été testée sous forme de multimère en raison des difficultés rencontrées dans la multimérisation des autres séquences. Dans un essai de gène

rapporteur, aucune séquence monomérique ou multimérique identifiée *in vitro* n'a pu agir comme élément de réponse pour la protéine ZNF74 et réprimer la transcription du gène rapporteur sous le contrôle d'un promoteur fort. Il est à noter qu'il est improbable que la position d'insertion de l'élément par rapport au site d'initiation de la transcription et du promoteur SV40 puisse expliquer la non fonctionnalité des séquences testées comme élément de réponse. En effet, il a été montré précédemment que la répression par les KRAB à multidoigt de zinc est indépendante de la position et de l'orientation de l'élément de réponse dans le promoteur (Senatore *et al.*, 1999)

En conclusion, les résultats obtenus avec les trois séquences testées (76, 73 et 31) indiquent que les séquences isolées par 3 cycles de purification South-Western et qui contiennent un motif consensus de six nucléotides ne peuvent pas fonctionner comme un élément de réponse pour la protéine ZNF74 *in vivo*.

4.1.3 - Les Causes probables des difficultés d'identification de l'élément de réponse pour ZNF74:

Quoique la technique de sélection par South-Western/PCR ait permis de caractériser une séquence consensus reconnue par la protéine KRAB à multidoigt de zinc AJ18 (Jheon *et al.*, 2001), cette même méthode ne nous a pas permis d'obtenir les résultats souhaités dans le cas de ZNF74, soit l'identification d'une séquence consensus agissant comme élément de réponse fonctionnel dans le cadre de la transcription.

Dans notre étude, le nombre de cycles de purification South-Western/PCR (3 cycles) réalisé a été relativement limité en comparaison avec certaines études similaires publiées pour lesquelles le nombre variait généralement de 5 à 7 cycles consécutifs. En fait, ce nombre a été restreint à 3 cycles dans nos expériences en raison de l'absence d'enrichissement des séquences marquées liées par la protéine ZNF74 au niveau du 2^{ème} et 3^{ème} cycle en comparaison avec le 1^{er} cycle. Cette absence d'enrichissement a été mise en évidence par South-Western en comparant la liaison de l'ADN amplifié et marqué radioactivement par PCR après chacun des cycles. L'absence d'enrichissement au cours des cycles pourrait traduire une liaison relativement non spécifique de la protéine ZNF74 à l'ADN dans les conditions *in vitro* utilisées ou être attribuable à la forte stringence des lavages salins permettant une purification efficace dès le premier cycle. De façon compatible avec cette deuxième explication, les études de Peng *et al.* (Peng *et al.*, 2002) qui ont permis d'identifier l'élément de réponse de la protéine KRAB à multidoigt de zinc SZF1 après trois cycles ne rapportaient aucun enrichissement entre le premier et le troisième cycle de retardement sur gel/PCR.

Les études des dernières années ont révélé la difficulté à identifier des cibles d'ADN *in vitro* pour les protéines à multidoigt de zinc. Quoique la liste ne soit pas exhaustive car plusieurs de ces tentatives n'ont pas été publiées, deux études au moins font le constat de

l'insuccès rencontré lors de l'isolation des séquences d'ADN spécifiques reconnues par des protéines à multidoigt de zinc par des méthodes de sélection *in vitro*, soit la protéine ZASC1 ou la protéine Kid1 (Bogaert *et al.*, 2005; Elser *et al.*, 1997). Plusieurs raisons possibles peuvent être invoquées. Tout d'abord, il est à noter que chaque motif doigt de zinc a structuralement la possibilité de se lier à trois nucléotides. Ainsi, une protéine comme ZNF74 qui a douze doigts de zinc pourrait reconnaître potentiellement une séquence couvrant 36 nucléotides ou plus. L'identification d'un consensus de liaison de façon visuelle ou avec l'utilisation d'un programme de recherche de consensus est compliquée par le fait que certains doigts peuvent ne pas être impliqués dans la liaison à l'ADN et que donc certains nucléotides inclus dans le consensus peuvent être dégénérés. Ceci est le cas pour les protéines ZBRK1 et SZF1 qui reconnaissent la même séquence de 15 nucléotides (GGGXXXCAGXXXTTT/G) (Peng *et al.*, 2002; Zheng *et al.* 2000). D'autre part, un élément technique limitant tient au fait que pour rechercher des consensus sur de longues séquences la complexité des bibliothèques d'oligonucléotides requises devient en pratique problématique (Exemple : 4^{36} oligonucléotides différents pour une bibliothèque dégénérée en 36 positions). Aussi, la plupart des études ont plutôt utilisé des bibliothèques de nucléotides dégénérés en 16 ou maximum 24 positions. Une deuxième limitation des méthodes de sélection *in vitro* réside dans le fait que les interactions identifiées peuvent être très différentes *in vitro* et *in vivo*. En effet, *in vivo* les interactions peuvent être stabilisées ou modifiées par la présence de co-facteurs interagissant avec le facteur transcriptionnel qui se lie à l'ADN. D'autre part, il ne peut être exclu en se basant sur le cas de la protéine à 11 doigts de zinc CTCF (Ohlson *et al.*, 2001) que ZNF74 pourrait reconnaître des séquences très divergentes via l'utilisation de combinaisons différentes de ses douze doigts de zinc et ce selon ses gènes cibles.

Enfin, quoique cela ne soit pas l'hypothèse que nous favorisons, on ne peut exclure la possibilité que certaines protéines KRAB multidoigt de zinc puissent fonctionner uniquement comme des co-régulateurs de la transcription lorsque recrutés indirectement à l'ADN par un facteur de transcription plutôt que comme facteur reconnaissant des séquences précises d'ADN.

A l'heure actuelle, on connaît seulement sept protéines KRAB à multidoigt de zinc pour lesquelles la ou les séquences spécifiques d'ADN reconnues ont été identifiées *in vitro* (Tableau I). D'autre par, l'identité d'un gène cible n'a été confirmée *in vivo* que pour la protéine KRAB multidoigt NT2 (Tanaka *et al.*, 2002)

4.1.4 – Recherche d'éléments de réponse de ZNF74 dans la séquence de son propre promoteur :

Certains facteurs transcriptionnels sont capables d'exercer une autorégulation de leur transcription en se liant à des séquences spécifiques au niveau de leurs propres promoteurs (Olson *et al.*, 2005 ; Mahatan *et al.*, 1999 ; Vassen *et al.*, 2005). Nous avons donc cherché à déterminer si ZNF74 pouvait avoir son propre promoteur comme cible et ce en vue de pouvoir le cas échéant identifier l'élément impliqué.

Au préalable, nous avons effectué une analyse de la séquence de 2 kb bordant le site d'initiation de la transcription et comprenant selon toute probabilité le promoteur puis nous avons procédé à une étude fonctionnelle de son activité transcriptionnelle.

L'inspection visuelle de la région -1714/+207 (par rapport au site d'initiation de la transcription) du promoteur de *ZNF74* indique que ce promoteur est dépourvu de TATA box proche du site d'initiation de la transcription tandis que l'analyse réalisée avec le logiciel MatchTM a aboutit à l'identification de 6 sites de liaisons pour les facteurs HNF1 (Hepatic

Nuclear Factor 1) (Tronche F, Yaniv M., 1992), **ARNT (AhR Nuclear Translocator homodimers)** (Hoffman *et al.*, 1991), **USF (Upstream Stimulatory Factor)** (Gregor *et al.*, 1990; Sirito *et al.*, 1992), **C/EBP (CCAAT/Enhancer Binding Protein)** (Ramji et Foka., 2002), **C-Rel** (Lioli et Hsia, 2003) et **E2F (E2 promoter binding Factor)** (Dimova et Dyson., 2005). Des études spécifiques (exemple : gel de retardement, mutation des sites, profils d'expression et essais de gène rapporteur) seraient requises pour déterminer si ces facteurs peuvent être requis pour réguler l'activité de ce promoteur.

L'analyse fonctionnelle du promoteur par essai de gène rapporteur, à l'aide de fragments de différentes longueurs compris dans la région -1429 à +97, nous a permis de mettre en évidence une région promotrice minimale correspondant à la région -399/+97. Cette région permet une très forte activation de la transcription. A titre d'indication de la force de ce promoteur minimal, on peut rappeler qu'il est 4 fois plus actif que le promoteur fort de SV40 et respectivement 75 fois et 150 fois plus actif que le vecteur dépourvu de promoteur sans ou avec «enhancer». Il faut noter que des éléments *cis* situés entre +50 et +97 semblent requis pour une activation maximale puisque la délétion dans la construction -399/+97 de cette région (47 bp) réduit de huit fois le niveau de transcription. La recherche d'éléments de réponse potentiels effectuée à l'aide du logiciel MatchTM, indique l'existence d'un site de liaison potentiel pour le facteur E2F (**E2 promoter binding Factor**) situé à +76. Une mutation de ce site aurait pu permettre de déterminer la fonctionnalité de ce site de liaison pour E2F. La présence d'un site de liaison de E2F fonctionnel en aval du site d'initiation de la transcription a déjà été mis en évidence pour d'autres gènes comme pour le récepteur FGF dont la transcription peut être fortement induite par E2F via la liaison à un site situé à +57 (Tashiro et al., 2003).

Après avoir confirmé que la région -1429 à +97 du promoteur de *ZNF74* est capable d'activer la transcription et identifié une région activatrice majeure entre +50 et +97, nous avons déterminé si cette région promotrice pouvait servir de cible pour la protéine ZNF74. Les résultats des essais transcriptionnels de gène rapporteur réalisés en présence de la protéine ZNF74 avec les fragments -1429/+49 et -399/+97 du promoteur ont révélé que ZNF74 n'exerce pas une autorégulation de sa transcription.

4.1.5 – L'immunoprécipitation de la chromatine, l'autre alternative :

Une autre alternative pour l'identification des gènes cibles régulés par la protéine ZNF74 serait l'immunoprécipitation de la chromatine (ChIP) couplée au clonage. La ChIP standard a été décrite initialement pour étudier la liaison *in vivo* de la protéine d'intérêt à une séquence génomique cible préalablement caractérisée par d'autres méthodes. Cette méthode ne permet pas la caractérisation de nouveaux gènes cibles. Quelques années plus tard, une variante de la méthode standard ChIP a été décrite par Weinmann et Farnham (Weinmann et Farnham, 2002). Elle combine la ChIP au Clonage et sert à identifier les gènes cibles liés par la protéine d'intérêt *in vivo*. Elle a permis d'identifier de nouveaux gènes cibles pour le facteur E2F (Weinmann *et al.*, 2001). Dans la méthode de ChIP couplée au clonage, on utilise des cellules qui expriment de façon endogène la protéine d'intérêt et l'on stabilise chimiquement les complexes protéines-ADN par liaison covalente avec la formaldéhyde. Après lyse des cellules, la chromatine traitée est soniquée dans le but de produire des complexes incluant des fragments ADN d'une taille approximative de 1-2 kb. Enfin, une étape d'immunoprécipitation des complexes ADN-protéine utilisant un anticorps qui reconnaît spécifiquement la protéine d'intérêt permet de récupérer la protéine ainsi que les fragments d'ADN associés. Ces fragments d'ADN sont ensuite clonés et séquencés et/ou hybridés à des micropuces d'ADN contenant soit des séquences promotrices (exemple : îlot CpG) ou des gènes spécifiques. La

spécificité des clones sélectionnés doit être confirmée par la méthode standard de ChIP, c'est à dire par couplage de la chromatine avec la formaldéhyde, immunoprécipitation de la chromatine et amplification par PCR avec des amorces reconnaissant les séquences des clones identifiés comme des cibles de la protéine d'intérêt.

4.2 – Conclusion :

L'étude présentée ici a permis de confirmer que si le domaine KRAB de la protéine ZNF74 fonctionnait comme un répresseur transcriptionnel tel que montré par Côté *et al* en 2001 (Côté *et al.*, 2001), cette activité répressive était conservée dans le contexte de la protéine complète. D'autre part, cette étude nous a permis d'initier une caractérisation de la région promotrice du gène *ZNF74*. Nous avons mis en évidence que la région du promoteur de *ZNF74* allant de -1429 à +97 relativement au site d'initiation de la transcription identifié précédemment (Côté *et al.*, 2001) a une activité promotrice. Le promoteur minimal identifié correspond à la région -399 à +49 tandis que des éléments de réponse situés en aval du site d'initiation de la transcription entre +49 et +97 entraîne une très forte augmentation de l'activité transcriptionnelle du promoteur. Notre recherche d'éléments de réponse de ZNF74 au niveau du promoteur de ZNF74 ou par une méthode de sélection *in vitro* s'est révélée infructueuse. Ce constat ainsi qu'une revue de la littérature révèle que la recherche d'éléments de réponse et de gènes cibles pour les protéines à motifs KRAB et multidoigt de zinc devrait probablement privilégier des approches *in vivo* comme la méthode de la chromatine immunoprécipitation (ChIP) couplée au clonage et/ou éventuellement à l'hybridation à des micropuces plutôt que les méthodes *in vitro* classiquement utilisées pour d'autres type de facteurs transcriptionnels.

Bibliographie :

- Aasland R, Gibson TJ, Stewart AF, 1995 **The PHD finger: implications for chromatin-mediated transcriptional regulation.** Trends Biochem Sci. 20:56-9.
- Aubry, M., Demczuk, S., Desmaze, C., Aikem, M., Aurias, A., Julien, J. P., Rouleau, G. A. 1993 **Isolation of a zinc finger gene consistently deleted in DiGeorge syndrome.** Hum.Mol.Genet. 2: 1583-1587.
- Bellefroid E. J., D. A. Poncelet, P. J. Lecocq, O. Revelant, J. A. Martial, 1991 **The evolutionarily conserved Kruppel-associated box domain defines a subfamily of eukaryotic multifingered proteins.** Proc. Natl. Acad. Sci. USA 88:3608-3612.
- Bellefroid EJ, Lecocq PJ, Benhida A, Poncelet DA, Belayew A, Martial JA, 1989 **The human genome contains hundreds of genes coding for finger proteins of the Kruppel type.** DNA. 8:377-87.
- Bellefroid EJ, Marine JC, Ried T, Lecocq PJ, Riviere M, Amemiya C, Poncelet DA, Coulie PG, de Jong P, Szpirer C, et al 1993 **Clustered organization of homologous KRAB zinc-finger genes with enhanced expression in human T lymphoid cells.** EMBO J 12:1363.
- Blackwell TK, Weintraub H, 1990 **Differences and similarities in DNA-binding preferences of MyoD and E2A protein complexes revealed by binding site selection.** Science.250:1104-10.
- Blanchard AD, Page KR, Watkin H, Hayward P, Wong T, Bartholomew M, Quint DJ, Daly M, Garcia-Lopez J, Champion BR, 2000 **Identification and characterization of SKAT-2, a novel Th2-specific zinc finger gene.** Eur J Immunol.30:3100.
- Bogaerts S, Vanlandschoot A, van Hengel J, van Roy F, 2005 **Nuclear translocation of alpha N-catenin by the novel zinc finger transcriptional repressor ZASC1.** Exp Cell Res. 311:1-13.
- Cammas F, Mark M, Dolle P, Dierich A, Chambon P, Losson R, 2000 **Mice lacking the transcriptional corepressor TIF1beta are defective in early postimplantation development.** Development. 127:2955.
- Caricasole A, Duarte A, Larsson SH, Hastie ND, Little M, Holmes G, Todorov I, Ward A, 1996 **RNA binding by the Wilms tumor suppressor zinc finger proteins.** Proc Natl Acad Sci U S A. 93:7562-6.
- Casademunt E, Carter BD, Benzel I, Frade JM, Dechant G, Barde YA, 1999 **The zinc finger protein NRIF interacts with the neurotrophin receptor p75^{NTR} and participates in programmed cell death.** EMBO J. 18:6050.

- Chang CJ, Chen YL, Lee SC, 1998 **Coactivator TIF1beta interacts with transcription factor C/EBPbeta and glucocorticoid receptor to induce alpha1-acid glycoprotein gene expression.** *Mol Cell Biol.*18:5880-7.
- Côté, F., Boisvert, F. M., Grondin, B., Bazinet, M., Goodyer, C. G., Bazett-Jones, D. P., Aubry, M. 2001 **Alternative promoter usage and splicing of ZNF74 multifinger gene produce protein isoforms with a different repressor activity and nuclear partitioning.** *DNA Cell Biol.* 20:159-173.
- Crossley PH, Little PF, 1991 **A cluster of related zinc finger protein genes is deleted in the mouse embryonic lethal mutation tw18.** *Proc Natl Acad Sci U S A.* 88:7923-7.
- Dimova DK, Dyson NJ, 2005 **The E2F transcriptional network: old acquaintances with new faces** *Oncogene.*24:2810-26.
- Driscoll DA, Budarf ML, Emanuel BS.,1992 **A genetic etiology for DiGeorge syndrome: consistent deletions and microdeletions of 22q11.** *Am J Hum Genet.* 50:924-33.
- Elser B, Kriz W, Bonventre JV, Englert C, Witzgall R, 1997 **The Kruppel-associated box (KRAB)-zinc finger protein Kid-1 and the Wilms' tumor protein WT1, two transcriptional repressor proteins, bind to heteroduplex DNA.** 272:27908-12.
- Friedman J. R., W. J. Fredericks, D. E. Jensen, D. W. Speicher, X. P. Huang, E. G. Neilson, F. J. R. Rauscher, 1996 **KAP-1, a novel corepressor for the highly conserved KRAB repression domain** *Genes Dev* 10:2067-2078.
- Ganss, B., and Kobayashi, H., 2002 **The zinc finger transcription factor Zfp60 is a negative regulator of cartilage differentiation.** *J Bone Miner Res* 17:2151.
- Garner, M. M., Revzin, A. 1981 **A gel electrophoresis method for quantifying the binding of proteins to specific DNA regions: application to components of the Escherichia coli lactose operon regulatory system.** *Nucleic Acids Res.* 9, 3047-3060.
- Gebelein B, Fernandez-Zapico M, Imoto M, Urrutia R, 1998 **KRAB-independent suppression of neoplastic cell growth by the novel zinc finger transcription factor KS1.** *J Clin Invest.* 102:1911-9.
- Gebelein B, Urrutia R, 2001 **Sequence-specific transcriptional repression by KS1, a multiple-zinc-finger-Kruppel-associated box protein.** *Mol Cell Biol.*21:928-39.
- Germain-Desprez, D., Bazinet, M., Bouvier, M., Aubry, M. 2003 **Oligomerization of TIF1 transcriptional regulators and interaction with ZNF74 nuclear matrix protein revealed by BRET in living cells.** *J.Biol.Chem.*25:22367-73.
- Gregor PD, Sawadogo M, Roeder RG,1990 **The adenovirus major late transcription factor USF is a member of the helix-loop-helix group of regulatory proteins and binds to DNA as a dimmer** *Genes Dev.* 4:1730-40.

- Grondin, B., Bazinet, M., Aubry, M. 1996 **The KRAB zinc finger gene ZNF74 encodes an RNA-binding protein tightly associated with the nuclear matrix.** J.Biol.Chem. 271:15458-15467.
- Grondin, B., Côté, F., Bazinet, M., Vincent, M., Aubry, M. 1997 **Direct interaction of the KRAB/Cys2-His2 zinc finger protein ZNF74 with a hyperphosphorylated form of the RNA polymerase II largest subunit.** J.Biol.Chem. 272:27877-27885
- Guévremont, M. 1999 **Interactions de la protéine à multidoigt de zinc ZNF74 avec les acides nucléiques.** Mémoire de Maîtrise, Université de Montréal.
- Hoffman EC, Reyes H, Chu FF, Sander F, Conley LH, Brooks BA, Hankinson O.1991 **Cloning of a factor required for activity of the Ah (dioxin) receptor** Science 252:954-958.
- Jerome LA, Papaioannou VE, 2001 **DiGeorge syndrome phenotype in mice mutant for the T-Box gene, Tbx1.** Nat Genet 27 :286-91.
- Jheon AH, Ganss B, Cheifetz S, Sodek J, 2001 **Characterization of a novel KRAB/C2H2 zinc finger transcription factor involved in bone development.** J Biol Chem.276:18282-9.
- Jing Z, Liu Y, Dong M, Hu S, Huang S, 2004 **Identification of the DNA binding element of the human ZNF333 protein.** J Biochem Mol Biol.37:663-70.
- Kim SS, Chen YM, O'Leary E, Witzgall R, Vidal M, Bonventre JV, 1996 **A novel member of the RING finger family, KRIP-1, associates with the KRAB-A transcriptional repressor domain of zinc finger proteins.**Proc Natl Acad Sci U S A. 93:15299-304.
- Koken MH, Saib A, de The H, 1995 **A C₄HC₃ zinc finger motif.** C R Acad Sci III. 318:733-9.
- Le Mee S, Fromiguet O, Marie PJ, 2005 **Sp1/Sp3 and the myeloid zinc finger gene MZF1 regulate the human N-cadherin promoter in osteoblasts** Exp Cell Res. 302:129-42.
- Lechner MS, Begg GE, Speicher DW, Rauscher FJ 3rd, 2000 **Molecular determinants for targeting heterochromatin protein 1-mediated gene silencing: direct chromoshadow domain-KAP-1 corepressor interaction is essential.**Mol Cell Biol. 20:6449-65.
- Liou HC, Hsia CY 2003 **Distinctions between c-Rel and other NF-kappa B proteins in immunity and disease.** Bioessays 25:767-80.
- Liu C, Levenstein M, Chen J, Tsifrina E, Yonescu R, Griffin C, Civin CI, Small D, 1999 **SZF1: a novel KRAB-zinc finger gene expressed in CD34+ stem/progenitor cells.** Exp Hematol. 27:313-25
- Looman C, Abrink M, Mark C, Hellman L, 2002 **KRAB zinc finger proteins: an analysis of the molecular mechanisms governing their increase in numbers and complexity during evolution.** Mol Biol Evol. 19:2118-30.

- Looman Camilla, 2003 **The ABC of KRAB zinc finger proteins**. Comprehensive summaries of Uppsala dissertations from the faculty of science and technology 863, ACTA universitatis upsaliensis
- Mahatan CS, Kaestner KH, Geiman DE, Yang VW, 1999 **Characterization of the structure and regulation of the murine gene encoding gut-enriched Kruppel-like factor (Kruppel-like factor 4)** Nucleic Acids Res. 27:4562-9.
- Margolin J. F., J. R. Friedman, W. K. Meyer, H. Vissing, H. J. Thiesen, F. J. R. Rauscher, 1994 **Kruppel-associated boxes are potent transcriptional repression domains**. Proc. Natl. Acad. Sci. USA 91:4509-4513.
- Mark Charlotta, 2003 **Three subfamilies of KRAB zinc finger proteins A structural, functional and evolutionary analysis**. Comprehensive summaries of Uppsala dissertations from the faculty of science and technology 864, ACTA universitatis upsaliensis.
- McDermid HE, Morrow BE, 2002 **Genomic disorders on 22q11**. Am J Hum Genet. 70:1077-88.
- Miller J., A. D. McLachlan, A. Klug, 1985 **Repetitive zinc-binding domains in the protein transcription factor IIIA from Xenopus oocytes**. EMBO J 4:1609-1614.
- Moosmann P, Georgiev O, Le Douarin B, Bourquin JP, Schaffner W 1996 **Transcriptional repression by RING finger protein TIF1 beta that interacts with the KRAB repressor domain of KOX1**. Nucleic Acids Res. 24:4859-67.
- Moosmann P, Georgiev O, Thiesen HJ, Hagmann M, Schaffner W, 1997 **Silencing of RNA polymerases II and III-dependent transcription by the KRAB protein domain of KOX1, a Kruppel-type zinc finger factor**. Biol Chem. 378:669-77.
- Morris JF, Hromas R, Rauscher FJ 3rd, 1994 **Characterization of the DNA-binding properties of the myeloid zinc finger protein MZF1: two independent DNA-binding domains recognize two DNA consensus sequences with a common G-rich core**. Mol Cell Biol. 14:1786-95.
- Noce T, Fujiwara Y, Sezaki M, Fujimoto H, Higashinakagawa T, 1992 **Expression of a mouse zinc finger protein gene in both spermatocytes and oocytes during meiosis**. Dev Biol 153:356.
- Ohlsson R, Renkawitz R, Lobanenkov V, 2001 **CTCF is a uniquely versatile transcription regulator linked to epigenetics and disease**. Trends Genet. 17:520-7.
- Olson LE, Zhang J, Taylor H, Rose DW, Rosenfeld MG, 2005 **Barx2 functions through distinct corepressor classes to regulate hair follicle remodeling**. Proc Natl Acad Sci. 102:3708-13.
- Pabo CO, Sauer RT. 1992 **Transcription factors: structural families and principles of DNA recognition**. Annu Rev Biochem 61:1053-95.
- Pavletich N. P., Pabo C. O, 1991 **Zinc finger-DNA recognition: crystal structure of a Zif268-DNA complex at 2.1 Å**. Science 252:809-817.

- Pavletich N. P., Pabo C. O, 1993 **Crystal structure of a five-finger GLI-DNA complex: new perspectives on zinc fingers.** *Science* 261:1701-1707.
- Peng H, Begg GE, Schultz DC, Friedman JR, Jensen DE, Speicher DW, Rauscher FJ 3rd, 2000 **Reconstitution of the KRAB-KAP-1 repressor complex: a model system for defining the molecular anatomy of RING-B box-coiled-coil domain-mediated protein-protein interactions.** *J Mol Biol.* 295:1139-62.
- Peng H, Zheng L, Lee WH, Rux JJ, Rauscher FJ 3rd, 2002 **A common DNA-binding site for SZF1 and the BRCA1-associated zinc finger protein, ZBRK1.** *Cancer Res.* 62:3773-81.
- Pengue G, Calabro V, Bartoli PC, Pagliuca A, Lania L, 1994 **Repression of transcriptional activity at a distance by the evolutionarily conserved KRAB domain present in a subfamily of zinc finger proteins.** *Nucleic Acids Res.* 22:2908-14.
- Ramji DP, Foka P, 2002 **CCAAT/enhancer-binding proteins: structure, function and regulation.** *Biochem. J* 561-575.
- Ravassard, P., Côté, F., Grondin, B., Bazinet, M., Mallet, J., Aubry, M. 1999 **ZNF74, a gene deleted in DeGeorge syndrome is expressed in human neural crest-derived tissues and foregut endoderm epithelia.** *Genomics* 62, 82-85.
- Rosati M., M. Marino, A. Franze, A. Tramontano, G. Grimaldi, 1991 **Members of the zinc finger protein gene family sharing a conserved N-terminal module.** *Nucleic Acids Res* 19:5661-5667.
- Ryan RF, Schultz DC, Ayyanathan K, Singh PB, Friedman JR, Fredericks WJ, Rauscher FJ 3rd, 1999 **KAP-1 corepressor protein interacts and colocalizes with heterochromatic and euchromatic HP1 proteins: a potential role for Kruppel-associated box-zinc finger proteins in heterochromatin-mediated gene silencing.** *Mol Cell Biol.* 19:4366-78.
- Sambrook J, Fritsch E.F, Maniatis T. 1989 **Molecular cloning a laboratory manual.** Cold spring harbour laboratory press.
- Sander TL, Haas AL, Peterson MJ, Morris JF, 2000 **Identification of a novel SCAN box-related protein that interacts with MZF1B. The leucine-rich SCAN box mediates hetero- and homoprotein associations.** *J Biol Chem.* 275:12857-67.
- Saurin AJ, Borden KL, Boddy MN, Freemont PS. 1996 **Does this have a familiar RING?.** *Trends Biochem Sci.* 21:208-14.
- Schuh R, Aicher W, Gaul U, Cote S, Preiss A, Maier D, Seifert E, Nauber U, Schroder C, Kemler R, et al, 1986 **A conserved family of nuclear proteins containing structural elements of the finger protein encoded by Kruppel, a Drosophila segmentation gene.** *Cell* 47:1025-32.

- Schultz DC, Ayyanathan K, Negorev D, Maul GG, Rauscher FJ 3rd, 2002. **SETDB1: a novel KAP-1-associated histone H3, lysine 9-specific methyltransferase that contributes to HP1-mediated silencing of euchromatic genes by KRAB zinc-finger proteins.** *Genes Dev* 16:919-32.
- Schultz DC, Friedman JR, Rauscher FJ 3rd, 2001 **Targeting histone deacetylase complexes via KRAB-zinc finger proteins: the PHD and bromodomains of KAP-1 form a cooperative unit that recruits a novel isoform of the Mi-2alpha subunit of NuRD.** *Genes Dev.*15:428-43.
- Schumacher C, Wang H, Honer C, Ding W, Koehn J, Lawrence Q, Coulis CM, Wang L, Ballinger D, Bowen BR, Wagner S, 2000 **The SCAN domain mediates selective oligomerization.** *J Biol Chem.* 275:17173-9.
- Senatore B, Cafieri A, Di Marino I, Rosati M, Di Nocera PP, Grimaldi G, 1999 **A variety of RNA polymerases II and III-dependent promoter classes is repressed by factors containing the Kruppel-associated/finger preceding box of zinc finger proteins.** *Gene.* 234:381-94.
- Sirito M, Walker S, Lin Q, Kozlowski MT, Klein WH, Sawadogo M 1992 **Members of the USF family of helix-loop-helix proteins bind DNA as homo- as well as heterodimers** *Gene Expr.* 2:231-40.
- Sirito M, Lin Q, Maity T, Sawadogo M. 1994 **Ubiquitous expression of the 43- and 44-kDa forms of transcription factor USF in mammalian cells.** *Nucleic Acids Res.* 22:427-33.
- Tanaka K, Tsumaki N, Kozak CA, Matsumoto Y, Nakatani F, Iwamoto Y, Yamada Y, 2002 **A Kruppel-associated box-zinc finger protein, NT2, represses cell-type-specific promoter activity of the alpha 2(XI) collagen gene** *Mol Cell Biol.* 22:4256-67.
- Tanaka, M., Herr, W. 1990 **Differential transcriptional activation by Oct-1 and Oct-2: interdependent activation domains induce Oct-2 phosphorylation.** *Cell* 60, 375-386.
- Tashiro E, Minato Y, Maruki H, Asagiri M, Imoto M, 2003 **Regulation of FGF receptor-2 expression by transcription factor E2F-1.** *Oncogene.* 22:5630-5.
- Theunissen O, Rudt F, Guddat U, Mentzel H, Pieler T, 1992 **RNA and DNA binding zinc fingers in Xenopus TFIIIA.** *Cell* 71:679-90.
- Thiesen HJ, Bach C, 1990 **Target Detection Assay (TDA): a versatile procedure to determine DNA binding sites as demonstrated on SP1 protein.** *Nucleic Acids Res.*18:3203-9.
- Tommerup N, Aagaard L, Lund CL, Boel E, Baxendale S, Bates GP, Lehrach H, Vissing H, 1993 **A zinc-finger gene ZNF141 mapping at 4p16.3/D4S90 is a candidate gene for the Wolf-Hirschhorn (4p-) syndrome.** *Hum Mol Genet.* 2:1571-5.
- Tronche F, Yaniv M 1992 **HNF1, a homeoprotein member of the hepatic transcription regulatory network.** *Bioessays.*14:579-87.

- Vassen L, Fiolka K, Mahlmann S, Moroy T, 2005 **Direct transcriptional repression of the genes encoding the zinc-finger proteins Gfi1b and Gfi1 by Gfi1b.** Nucleic Acids Res. 33(3):987-98.
- Vissing H., W. K. Meyer, L. Aagaard, N. Tommerup, H. J. Thiesen, 1995 **Repression of transcriptional activity by heterologous KRAB domains present in zinc finger proteins** FEBS. Lett 369:153-157.
- Wagner S, Hess MA, Ormonde-Hanson P, Malandro J, Hu H, Chen M, Kehrer R, Frodsham M, Schumacher C, Beluch M, Honer C, Skolnick M, Ballinger D, Bowen BR, 2000 **A broad role for the zinc finger protein ZNF202 in human lipid metabolism.** J Biol Chem. 275:15685.
- Williams AJ, Khachigian LM, Shows T, Collins T. 1995 **Isolation and characterization of a novel zinc-finger protein with transcription repressor activity.** J Biol Chem. 270:22143-52.
- Witzgall R, O'Leary E, Leaf A, Onaldi D, Bonventre JV, 1994 **The Kruppel-associated box-A (KRAB-A) domain of zinc finger proteins mediates transcriptional repression** Proc Natl Acad Sci. 91:4514-8.
- Zheng L, Pan H, Li S, Flesken-Nikitin A, Chen PL, Boyer TG, Lee WH, 2000 **Sequence-specific transcriptional corepressor function for BRCA1 through a novel zinc finger protein, ZBRK1** Mol Cell. 6:757-68
- Yamagishi H, 2002 **The 22q11.2 deletion syndrome.** Keio J Med. 51:77-88.

