

50376  
1994  
67

cc0 gen 2015/10/03

50376  
1994  
67

UNIVERSITE DES SCIENCES ET TECHNOLOGIES DE LILLE

U F R de BIOLOGIE

Année 1994

N° d'ordre 1280

## THÈSE

PRÉSENTÉE POUR L'OBTENTION DU GRADE DE  
DOCTEUR DE L'UNIVERSITÉ EN  
SCIENCES DE LA VIE ET DE LA SANTÉ



Anne LELIEVRE-CHOTTEAU

### **Identification de deux gènes de la famille des proto- oncogènes *ets* chez une Annélide Polychète, *Nereis diversicolor***

Présentée le 22 Février 1994 devant la Commission d'examen :

Président                      Professeur A. DHAINAUT

Rapporteurs                    Professeur J. HOURDRY  
Docteur D. LEPRINCE

Examineurs                    Professeur J. KREMBEL  
Professeur M. PORCHET  
Docteur F. FONTAINE

A Eric, à mes parents, à ma famille, à Céline et à mes amis,  
je dédie ce mémoire.

## REMERCIEMENTS

Ce mémoire est pour moi l'occasion de remercier Messieurs Maurice PORCHET et André DHAINAUT, successivement directeurs du Laboratoire de Phylogénie Moléculaire des Annélides, qui ont eu confiance en Ets et m'ont permis de développer ce thème.

Je remercie également M. DHAINAUT d'avoir bien voulu présider mon jury de thèse, et M. PORCHET d'avoir accepté d'être examinateur.

Je remercie Messieurs Jacques HOURDRY et Jean KREMBEL de me faire l'honneur d'être rapporteur et examinateur du travail présenté.

Un grand merci à Dominique LEPRINCE pour de nombreuses raisons et notamment pour sa disponibilité, sa sympathie et pour avoir toujours considéré ce travail avec beaucoup d'attention. Je lui suis très reconnaissante et le remercie aussi d'avoir accepté d'être rapporteur de cette thèse.

A François FONTAINE, avec qui j'ai découvert la biologie moléculaire, pour ces années pleines d'embuches mais aussi toujours pleines d'espoir, merci pour la confiance et le soutien.

Un autre grand merci à Agnès BÈGUE qui m'a toujours été d'un grand secours avec beaucoup de gentillesse et qui est toujours disponible pour me faire profiter de ses connaissances en bio mol. Je lui suis très reconnaissante de tout ce qu'elle a fait pour moi.

Merci beaucoup à Xavier DESBIENS pour avoir toujours été présent aux moments opportuns, pour sa sympathie et son dévouement.

Je remercie Vincent LAUDET pour avoir effectué le traitement informatique des séquences présentées.

Un merci admiratif à Françoise BONET, pour ses qualités (entre autres) de Word et Superpaint woman, pour sa rigueur, sa bonne humeur et ses crocodiles...

Je remercie également toutes les personnes du laboratoire qui ont participé à leur façon à ce que ces années et surtout ces derniers mois se déroulent pour le mieux : Marie-Christine SLOMIANNY, Françoise BONET (la même), Lucette MEUNIER, Annie DESMONS, Yvette HIMPENS, Gérard MONTAGNE et Daniel LAZARECKI.

Merci aussi à toutes les personnes qui ont bien voulu lire et critiquer ce mémoire et notamment Jean-Luc BAERT, Xavier DESBIENS et Agnès WATTEZ.

## SOMMAIRE

	Page
ABBREVIATIONS	8
RÉSUMÉ	10
PRÉSENTATION DU MÉMOIRE	11
INTRODUCTION	14
I - Origine de la famille ETS	15
A - Le rétrovirus E26	15
1 - Structure du rétrovirus	15
2 - Genèse du virus E26	15
3 - Mécanismes de transduction	16
4 - Organisation de la P135 <i>gag-myb-ets</i>	16
5 - Les gènes cibles de la P135 <i>gag-myb-ets</i>	16
6 - Les cellules cibles du rétrovirus E26	17
7 - Implications respectives de <i>v-myb</i> et <i>v-ets</i>	18
B - Identification des homologues cellulaires de l'oncogène viral <i>v-ets</i>	18
1 - Le proto-oncogène <i>c-ets-1</i> aviaire	18
2 - Les homologues de <i>c-ets-1</i> chez l'homme, le poulet et la souris	20
II - Modèle d'étude de la famille ETS : le proto-oncogène <i>c-ets-1</i>	21
A - Fonction des produits de <i>c-ets-1</i>	21
1 - Les protéines Ets-1 sont des facteurs de transcription	21
a) Le domaine de liaison à l'ADN de <i>c-ets-1</i>	21
b) Séquences de fixation des protéines Ets (EBS)	22
c) Reconnaissance moléculaire	24
2 - Les gènes cibles des protéines Ets-1	24
3 - Coopération des protéines Ets-1 avec d'autres facteurs de transcription	26
B - Les domaines fonctionnels des protéines Ets-1	27
1 - Les domaines de régulation de la transcription	27
2 - Répression intramoléculaire de la fixation à l'ADN	28
C - Expression de <i>c-ets-1</i>	30
D - Régulation de la fonction des produits du gène <i>c-ets-1</i>	32
1 - Epissage alternatif	32
2 - Maturation post-traductionnelle des protéines Ets-1	34
3 - Localisation nucléaire	36
4 - Stabilité des protéines Ets-1	36

E - Régulation de <i>c-ets-1</i>	36
1 - Caractérisation du promoteur	36
2 - Régulation de l'expression	37
F - Les protéines Ets-1 et la transformation	40
1 - Transformation de cellules <i>in vitro</i>	40
2 - Translocation chromosomique	40
3 - Ets-1 dans les tumeurs humaines	40
III - La famille ETS	41
A - Critères de classification des membres de la famille ETS	41
1 - Les domaines A, B, C	41
2 - Le domaine ETS	42
3 - Classement des protéines Ets	43
B - Modèle d'évolution	43
1 - Classement des domaines ETS	43
2 - Phylogénie du domaine ETS	44
IV - Les protéines de la famille Ets chez les Vertébrés	45
1 - Fonction des protéines Ets	45
2 - Le site de fixation à l'ADN des protéines Ets	45
3 - Le domaine inhibiteur de la fixation à l'ADN	46
4 - Les domaines transactivateurs	47
5 - Interactions avec diverses protéines régulatrices	47
a) La reconnaissance de deux sous-unités protéiques permet la constitution d'un complexe transcriptionnel efficace	48
b) La fixation d'un facteur dépend de son interaction avec une protéine de la famille ETS	49
c) La fixation à l'ADN d'une protéine Ets nécessite la fixation préalable d'un autre facteur protéique	49
6 - Régulation de l'activité des protéines Ets	50
V - Rôle physiologique des protéines Ets	52
1 - Les protéines Ets dans le système immunitaire	52
2 - Les protéines Ets dans le développement embryonnaire	54
3 - L'activation des protéines Ets	55
a) Activation par insertion rétrovirale	55
b) Activation par réarrangement génomique	56
VI - La famille ETS chez les Invertébrés	56
A - Les protéines E74 de la drosophile	57

1 - Le modèle d'Ashburner	57
2 - Le gène <i>E74</i> et ses produits	57
3 - Mode d'action de l'ecdysone	59
4 - Régulation transcriptionnelle des gènes précoces	59
B - Les autres gènes de la famille Ets chez la drosophile	60
C - La famille ETS chez d'autres Invertébrés	62
VII - Conclusion	63
<b>MATÉRIEL BIOLOGIQUE</b>	65
I - Situation phylogénique des Annélides Polychètes	66
II - Organisation générale de la néréis	67
III - Les cellules cœlomiques	68
IV - La cicatrisation et la régénération	69
A - La cicatrisation	70
B - La formation du pygidium	70
C - La mise en place des métamères	70
<b>RÉSULTATS</b>	71
I - Mise en évidence d'homologues de la séquence <i>v-ets</i> chez la néréis	72
II - Recherche des gènes de la famille ETS dans des banques d'ADN génomique	74
A - Construction et criblage de la banque $\lambda$ EMBL4	74
B - Construction et criblage de la banque $\lambda$ NM1149	77
III - Recherche des membres de la famille ETS par réaction de polymérisation en chaîne (PCR)	78
IV - Recherche du gène correspondant à la sonde Nd <i>erg</i> dans une banque d'ADN génomique construite dans le vecteur $\lambda$ DASHII	80
A - Criblage de la banque $\lambda$ DASHII	80
B - Cartographie partielle des fragments insérés dans les phages	82
C - Clonage des fragments <i>EcoRI</i> possédant la séquence homologue de la sonde Nd <i>erg</i>	82
D - Séquençage	84
V - Analyse des produits de l'amplification de l'ADN génomique par les oligonucléotides D-OL1 et D-OL2	85

VI - Recherche de fragments d'ADNc codant des membres de la famille ETS	87
A - Recherche par la technique de rétrotranscription associée à la PCR (RT PCR)	87
B - Clonage des fragments d'ADNc et séquençage	90
1 - Clonage des produits correspondant à l'amplification des ADNc avec les oligonucléotides D-OL1 et D-OL2	90
2 - Clonage des produits correspondant à l'amplification des ADNc avec les oligonucléotides D-OL1 et oligo(dT) <sub>18</sub>	94
Mise en évidence de l'expression du gène Nd <i>ets</i>	96
VII - Recherche de l'expression des gènes correspondant aux sondes Nd <i>erg</i> et Nd <i>ets</i>	98
A - Recherche par "Northern blot"	98
B - Recherche par RT PCR	98
C - Recherche par hybridation <i>in situ</i>	99
VIII - Recherche des transcrits des gènes correspondant aux sondes Nd <i>erg</i> et Nd <i>ets</i> dans une banque d'ADN complémentaire	99
A - Criblage de la banque avec les sondes nucléotidiques	99
B - Criblage de la banque par PCR avec les oligonucléotides D-OL1 et D-OL2	101
C - Criblage de la banque par PCR avec des oligonucléotides spécifiques du gène Nd <i>erg</i>	101
IX - Recherche par RT PCR de transcrits correspondant au gène Nd <i>ets</i>	103
DISCUSSION	104
I - Identification du gène Nd <i>erg</i> chez <i>Nereis diversicolor</i>	105
A - Domaine ETS de Nd Erg	106
B - Domaine R de Nd Erg	107
C - Extrémité 5' de la séquence Nd <i>erg</i>	108
II - Recherche d'autres gènes de la famille ETS	109
III - Recherche de l'expression des gènes de la famille ETS et identification du gène Nd <i>ets</i>	110
A - Recherche de l'expression par RT PCR	111
B - Recherche de l'expression par "Northern blots" et hybridation <i>in situ</i>	113

IV - Recherche des transcrits des gènes Nd <i>erg</i> et Nd <i>ets</i>	115
A - Criblage d'une banque d'ADNc avec les sondes Nd <i>erg</i> et Nd <i>ets</i>	115
B - Criblage d'une banque d'ADNc par PCR	115
V - Notion d'évolution et situation phylogénique de Nd Erg et Nd Ets	116
A - Evolution phylogénique des domaines ETS	116
1 - Analyse phylogénique selon Lautenberger <i>et al.</i> (1992)	116
2 - Analyse phylogénique selon Laudet <i>et al.</i> (1993)	117
3 - Contribution de Degnan <i>et al.</i> (1993) à la phylogénie des domaines ETS	118
B - Situation phylogénique des gènes <i>erg</i> et <i>ets</i> de la néréis	120
CONCLUSION	124
MÉTHODES	128
I - Extraction des acides nucléiques	129
A - Extraction de l'ADN génomique	129
B - Extraction des ARN	130
1 - Précautions à prendre	130
2 - Extraction des ARN totaux	130
3 - Extraction des ARN messagers	131
4 - Traduction des ARN en lysat de réticulocytes de lapin	132
II - Préparation des sondes d'ADN	132
A - Amplification bactérienne	132
B - Préparation de l'ADN plasmidique	133
1 - Méthode au lysozyme	133
2 - Méthode de lyse alcaline	133
C - Préparation de l'insert	133
1 - Fenêtre en gel d'agarose à 0,3 %	133
2 - Membrane de DEAE cellulose (Schleicher et Schuell)	134
3 - Electroélution	134
III - Marquage des sondes d'ADN	134
A - Marquage avec un radioisotope	134
1 - Marquage de sondes double brin	134
2 - Marquage d'oligonucléotides	135
B - Marquage sans radioisotope	136

IV - Méthode d'analyse des acides nucléiques	136
A - Analyse de l'ADN	136
1 - Techniques électrophorétiques	136
2 - Transfert-hybridation selon Southern (1975)	136
B - Analyse des ARN	139
1 - Gel dénaturant	139
2 - "Northern blot"	140
3 - Hybridation	140
4 - Lavages	141
V - Construction de banques d'ADN génomique	141
A - Différentes banques construites	141
B - Préparation des bras du bactériophage pour les phages λ EMBL4 et λ DASHIII	141
C - Préparation de l'ADN génomique	142
D - Encapsidation <i>in vitro</i>	142
VI - Construction d'une banque d'ADN complémentaire	143
A - Préparation des ARNm	143
B - Synthèse des brins d'ADN complémentaire (Kit Time Saver cDNA synthesis de Pharmacia)	143
C - Ligation des ADNc dans le vecteur λ gt11D	144
D - Encapsidation <i>in vitro</i>	144
VII - Criblage des banques d'ADN génomique et d'ADN complémentaire	144
A - Titre de la banque	144
B - Etalement et repérage des clones	144
C - Hybridation	144
D - Isolement des clones	144
E - Extraction de l'ADN phagique	145
VIII - Amplification enzymatique par la réaction de polymérisation en chaîne (PCR)	145
A - Conditions d'utilisation	145
1 - Amorces	145
2 - Conditions d'amplification	146
3 - Amplification sur colonie bactérienne	147
B - RT PCR	147
1 - Synthèse des brins codant des ADN complémentaires	147

2 - Synthèse d'une queue poly(A) au niveau de l'extrémité 3' du brin d'ADN complémentaire	147
3 - Amplification enzymatique par PCR	148
4 - 5' RACE "Rapid Amplification of cDNA 5' Ends"	148
IX - Techniques de clonage	148
A - Clonage et sous clonage d'insert	148
1 - Préparation des inserts	148
2 - Préparation des vecteurs	149
3 - Ligation	150
4 - Transformation bactérienne	151
5 - Sélection des colonies et vérification du clonage	153
B - Clonage des produits de PCR	154
X - Séquençage selon Sanger <i>et al.</i> (1977)	155
A - Séquençage d'une matrice simple brin	155
1 - Préparation de matrices simple brin d'ADN de phage M13	155
2 - Séquençage d'une matrice simple brin M13	156
B - Séquençage d'une matrice double brin	156
C - Electrophorèse	156
XI - Hybridation moléculaire <i>in situ</i>	157
A - Préparation du matériel histologique	157
B - Marquage radioactif de sondes ARN sens et antisens	157
C - Purification des sondes	159
D - Hybridation-lavages	159
1 - Déparaffinage et réhydratation des coupes	159
2 - Elimination du paraformaldéhyde	159
3 - Perméation des tissus	159
4 - Post-fixation	159
5 - Acétylation	160
6 - Déshydratation	160
7 - Hybridation	160
8 - Lavages	160
9 - Déshydratation	161
E - Autoradiographie	161
RÉFÉRENCES BIBLIOGRAPHIQUES	162

## ABBREVIATIONS

ADNc	Acide Désoxyribonucléique complémentaire
AMPc	Adénosine monophosphate cyclique
AMV	"Avian Myeloblastosis Virus"
AP-1	"Activator protein-1"
AP-2	"Activator protein-2"
ARNm	ARN messenger
EBS	"Ets Binding Site"
F-Mu LV	"Friend Murine Leukemia Virus"
HIV-1	"Human Immunodeficiency Virus-1"
HLH	"Helix-Loop-Helix"
HMG	"High Mobility Group"
HSV	"Herpes Simplex Virus"
HTLV-1	"Human Lymphotropic Virus-type 1"
IL 2	Interleukine 2
IP 3	Inositol tri-Phosphate
kDa	kiloDalton
kpb	kilopaire de bases
LTR	"Long Terminal Repeat"
MAP kinase	"Mitogen Activating Protein" kinase
MMTV	"Mouse Mammary Tumor Virus"
MSV	"Moloney Sarcoma Virus"
ORU	"Oncogene Response Unit"
pb	paire de bases
PCR	"Polymerase Chain Reaction"
PEA3	"Polyome Enhancer 3"
PEST	Région riche en résidus proline, glutamine, sérine et thréonine
PKC	"Protein Kinase C"
PMA	"Phorbol Myristate Acetate"
RAV	"Rous Associated Virus"
RT PCR	Rétrotranscription couplée à la PCR
SAAB	"Selected And Amplified Binding"
SFFV	"Spleen Focus Forming Virus"
SRE	"Serum Response Element"
SRF	"Serum Response Factor"

SV 40	"Simian Virus 40"
TCF	"Ternary Complex Factor"
TCF-1	"T Cell Factor-1"
TCR	"T Cell Receptor"
TPA	"Tumor Promoting Agent" (12-O-tétra-décanoyl-phorbol-13-acétate)
uPA	"urokinase Plasminogen Activator"

Ck	Poulet
D	Drosophile
Mo	Souris
Hu	Homme
Su	Oursin
X	Xénope

Pour clarifier la compréhension du texte, nous avons convenu de la présentation suivante : les gènes sont mis en caractères *italiques* et les produits des gènes en caractères **Droits**. Les protéines Ets sont classées en **GROUPES** et **SOUS-GROUPES** écrits en majuscules.

## RÉSUMÉ

Les protéines de la famille ETS sont des facteurs de transcription constituant une nouvelle famille de protéines eucaryotiques se liant à l'ADN. Ces protéines ont au moins en commun le domaine de liaison à l'ADN, appelé domaine ETS, qui constitue un nouveau motif structural de reconnaissance ADN-protéines et qui permet de les classer en différents groupes. Elles reconnaissent une cible nucléotidique riche en purine (GGA) et plusieurs protéines sont capables de reconnaître le même motif. La spécificité et la sélectivité d'action sur la régulation de gènes particuliers, en dépit de la reconnaissance de séquences nucléotidiques proches, constituent l'originalité de cette famille de facteurs de transcription. Ainsi, ces protéines interagissent avec d'autres protéines de régulation pour former des complexes stables nécessaires à l'activation de la transcription de gènes uniques, parfois dans un type cellulaire spécifique.

Dans le cadre d'une étude phylogénique, nous avons recherché la présence de membres de la famille ETS chez l'Annélide Polychète, *Nereis diversicolor*.

Nous avons identifié deux gènes appartenant aux groupes ETS et ERG de cette famille.

La partie identifiée du gène *ets* de néréis, traduite en acides aminés, présente une homologie importante avec les séquences des protéines Ets de l'homme, de la drosophile et de l'oursin.

Nous avons une connaissance plus complète de l'*erg* de néréis. Le fragment d'ADN séquencé montre qu'il contient dans le même exon la région codant la totalité du domaine ETS et une partie de la région codant le domaine R caractéristique des protéines Erg. Le domaine ETS présente une forte homologie avec les domaines ETS des protéines Erg de Vertébrés et d'Invertébrés. Le domaine R présente de l'homologie avec celui des protéines Erg et Fli de l'homme, de la souris et de l'oursin, seuls organismes où ce domaine est actuellement connu. L'homologie est cependant plus importante entre les domaines R d'Invertébrés d'une part et de Vertébrés d'autre part.

La famille des proto-oncogènes *ets* est une famille ancienne déjà diversifiée dès les Métazoaires primitifs, dont le domaine de liaison à l'ADN est bien conservé jusqu'aux Vertébrés et aux Arthropodes.

La diversification des gènes *ets* s'est effectuée par mouvement d'exons, en même temps que par duplication de gènes, puis divergence. La situation phylogénique particulière des Annélides Polychètes, à proximité des Cœlomates primitifs, en font un matériel de choix pour la compréhension de l'origine des différents groupes de protéines de la famille ETS et du rôle des différents domaines fonctionnels de ces protéines.

# **PRÉSENTATION DU MÉMOIRE**

Les rétrovirus transformants sont des vecteurs naturels portant des séquences cellulaires dérivées de gènes normaux appelés proto-oncogènes. Les rétrovirus de type C possèdent, en plus des gènes de structure *gag*, et/ou *pol* et *env*, un ou deux gènes responsables de la transformation des cellules infectées. Ces oncogènes viraux (*v-onc*) ont été intégrés dans le génome viral par recombinaison de séquences virales avec des séquences cellulaires.

Un grand nombre de proto-oncogènes (*c-onc*), homologues des *v-onc*, ont été caractérisés chez de nombreux organismes et sont impliqués dans les processus de transformation cellulaire non viro-induite. L'activation des proto-oncogènes peut s'effectuer soit par modification qualitative du gène (mutation, délétion, insertion), soit par modification quantitative de l'expression du produit normal par amplification génique ou translocation chromosomique.

Les produits de certains proto-oncogènes sont des facteurs de transcription intervenant dans la régulation génique; l'homologie de certains produits des proto-oncogènes avec des facteurs de croissance ou des récepteurs de facteurs de croissance, ainsi que leur implication dans les différentes étapes de la transmission du signal mitotique, attestent en faveur du rôle fondamental qu'ils jouent dans la vie de la cellule au niveau de la croissance, de la prolifération et de la différenciation cellulaire. Au cours du développement embryonnaire, le passage de l'œuf à un organisme différencié implique la mise en œuvre de mécanismes d'une grande diversité, se traduisant par une modulation de l'expression des gènes. Certains gènes vont être induits, d'autres réprimés. Cette régulation se fait à la fois dans le temps et dans l'espace. La compréhension de ces mécanismes qui aboutissent à la formation d'un organisme complexe est essentielle. La régulation des gènes impliqués nécessite la mise en œuvre de protéines appelées facteurs de transcription qui se lient à proximité ou dans la zone promotrice de ces gènes. Le nombre restreint de facteurs de transcription existant par rapport au nombre de gènes exprimés dans une cellule implique une coopération entre ces facteurs de transcription et d'autres molécules.

La diversité et la complexité du contrôle transcriptionnel sont permises par la combinaison de divers événements. Les gènes possèdent généralement plusieurs séquences *cis* régulatrices permettant la fixation de facteurs de transcription différents donnant ainsi une combinatoire importante. D'autre part, une même protéine peut exister sous différentes formes alternatives pouvant entrer en compétition et avoir des effets antagonistes ; il en est de même pour différents facteurs de transcription appartenant à une même famille de gènes. Les

modifications post-traductionnelles, telle que la phosphorylation, sont aussi des éléments intervenant dans la diversité du contrôle transcriptionnel.

De nombreux proto-oncogènes, identifiés chez des Vertébrés, ont été décrits chez divers Invertébrés.

Une remarquable conservation des séquences peptidiques, voire nucléotidiques, est observée au cours de l'évolution, depuis les Eucaryotes inférieurs jusqu'aux Vertébrés supérieurs. C'est pourquoi, afin de mieux comprendre l'organisation, le rôle et le mode d'action des proto-oncogènes, des études ont été réalisées sur des modèles Invertébrés.

La néréis, ver marin de l'Embranchement des Annélides Polychètes, a fait l'objet de nombreux travaux au sein de notre laboratoire. Les Polychètes, par suite de leur situation phylogénique (ce sont les plus archaïques des Cœlomates) occupent une position clé dans l'évolution animale. Une simplicité de leur organisation, des fonctions métaboliques bien circonscrites dans le temps, un système immunitaire mettant en jeu une coopération cellulaire et un mécanisme humoral ainsi qu'un grand pouvoir de régénération de leur extrémité postérieure, sont quelques aspects qui en font un modèle d'un grand intérêt biologique.

La famille des protéines Ets est une famille de plus en plus importante de facteurs de transcription dont la plupart des membres ont été récemment identifiés.

Dans le cadre d'une étude phylogénique, nous avons recherché la présence de membres de la famille ETS, chez la néréis.

La première partie de ce mémoire présente la famille des facteurs de transcription ETS. La fonction et le rôle de cette famille de gènes sont illustrés par le gène *c-ets-1* du poulet, le premier identifié et le plus étudié de la famille ETS. Les membres de la famille ETS, ainsi que leurs fonctions biologiques connues, sont ensuite présentés.

Les résultats obtenus chez *Nereis diversicolor* sont exposés et discutés dans une deuxième partie. A partir des connaissances accumulées à ce jour, et notamment à partir des résultats fragmentaires récemment publiés pour les Métazoaires primitifs, nous discutons de la position phylogénique des protéines Ets de la néréis et de l'hypothèse de l'évolution de cette famille de gènes.

# **INTRODUCTION**

## **I - Origine de la famille ETS**

### **A - Le rétrovirus E26**

#### **1 - Structure du rétrovirus**

Le rétrovirus E26 appartient au groupe des virus défectifs leucémogènes aviaires (DLV). Son génome est représenté par un ARN défectif pour la propagation à cause de l'absence de séquences virales remplacées par des séquences d'origine cellulaire. Il a donc besoin d'un virus auxiliaire pour se propager. Ce virus auxiliaire code par l'intermédiaire des gènes *gag*, *pol* et *env*, les protéines nécessaires à la structure du virus (les protéines de la capsidie et de l'enveloppe) et à sa réplication (la transcriptase réverse).

L'ADN proviral du virus E26 contient deux oncogènes distincts dérivés de séquences cellulaires : *v-myb* et *v-ets* (Roussel *et al.*, 1979 ; Leprince *et al.*, 1983 ; Nunn *et al.*, 1983) en plus des gènes *gag* et *env* tronqués. Ces séquences sont flanquées de deux LTR ("Long Terminal Repeat") qui sont les séquences promotrices de la transcription virale (Figure 1). E26 code une protéine de fusion nucléaire de 135 kDa, la P135*gag-myb-ets* (Beug *et al.*, 1982 ; Bister *et al.*, 1982 ; Nunn *et al.*, 1983 ; Boyle *et al.*, 1984), identifiée dans les cellules transformées par le rétrovirus (Klempnauer *et al.*, 1984). Le rétrovirus E26 est jusqu'alors le seul virus ayant transduit des déterminants *ets*.

#### **2 - Genèse du virus E26**

Les séquences *v-myb* et *v-ets* transduites par le rétrovirus proviennent des gènes cellulaires *c-myb* et *c-ets-1*. Dans le génome du virus E26, l'oncogène est homologue aux six premiers exons codants du gène cellulaire *c-myb*, le premier exon étant amputé de son extrémité 5'. Le cadre de lecture est cependant conservé. La jonction *myb/ets* a lieu au milieu de l'exon 6 et au début du premier exon  $\alpha$  de *c-ets-1* (Figure 1). La fusion en bonne phase des séquences *v-myb* et *v-ets* résulte donc probablement d'un épissage illégitime ayant accolé un site cryptique donneur de l'exon 6 de *c-myb* et un site accepteur normal de l'exon  $\alpha$  de *c-ets-1* (Leprince *et al.*, 1988). Les séquences sont homologues aux régions codantes du gène *c-ets-1* excepté au niveau carboxy-terminal et pour deux mutations ponctuelles. D'autre part, au niveau de l'extrémité amino-terminale, les onze nucléotides 5' non codants de *c-ets-1* sont traduits dans *v-Ets*.

Au niveau de l'extrémité carboxy-terminale, les treize derniers acides aminés codés par le gène *c-ets-1* sont remplacés par seize acides aminés



n'ayant aucune relation dans v-Ets. Ces acides aminés sont en fait codés par le brin complémentaire de celui codant les treize derniers acides aminés de la protéine Ets-1. L'homologie se termine au niveau d'un codon stop potentiel, qui se trouve 16 codons après le point de divergence entre *c-ets-1* et *v-ets* (Lautenberger et Papas, 1993 ; Leprince *et al.*, 1993) (Figure 2).

### 3 - Mécanismes de transduction

Plusieurs hypothèses ont été formulées sur les mécanismes aboutissant à la transduction des oncogènes.

L'une de ces hypothèses propose que lors de la recombinaison d'un provirus à côté d'un proto-oncogène, sa transcription se poursuit au delà du site de polyadénylation virale. Le proto-oncogène est donc transcrit dans la foulée, suivi ou non par un épissage anormal entre les gènes viraux et cellulaire. Ces ARN sont donc encapsidés avec des ARN viraux normaux. Au cycle viral suivant, de nouveaux virus sont créés par des recombinaisons dues à des changements de matrice d'un brin d'ARN sur l'autre (Swain et Coffin, 1992 ; Zhang et Temin, 1993a).

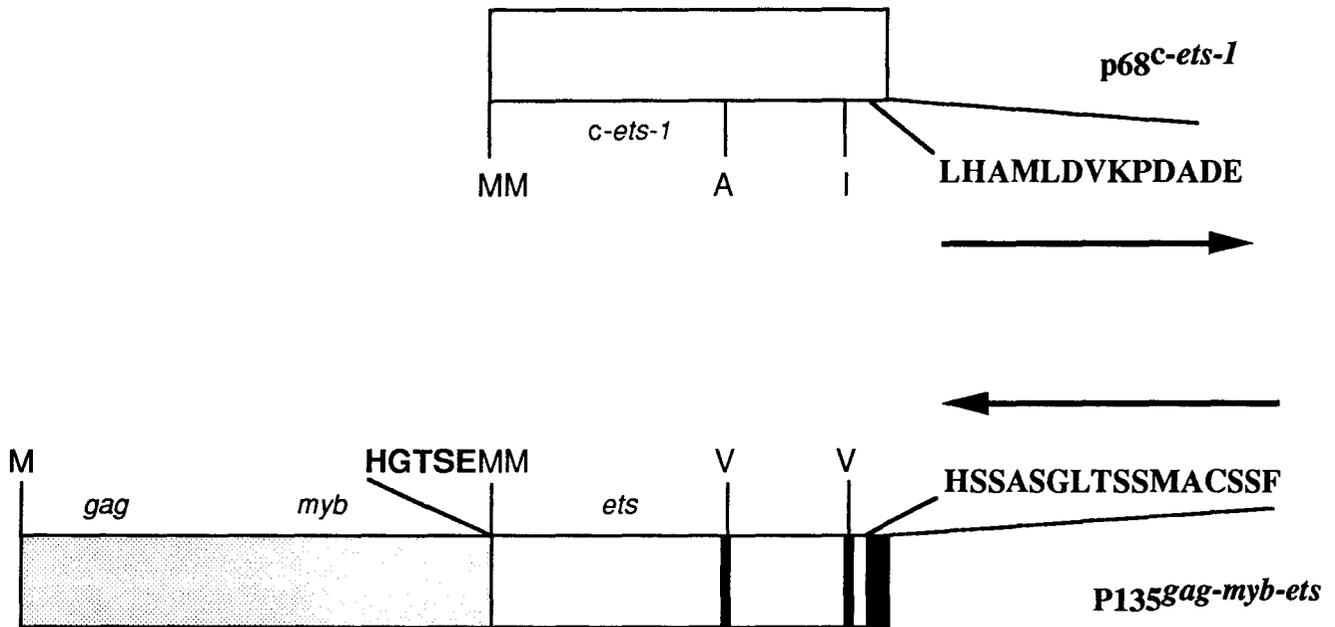
Le rétrovirus E26 fait partie du groupe 2 selon la classification de Zhang et Temin (1993b), car la jonction entre les séquences virales et cellulaires est séparée par une insertion. L'analyse informatique de la séquence de cette insertion (Lautenberger et Papas, 1993 ; Leprince *et al.*, 1993) a révélé une grande homologie avec le gène *env* du rétrovirus RAV-2 (Bova *et al.*, 1988), qui se poursuit avec les derniers codons de *v-ets*. Cette étude suggère qu'un rétrovirus du type RAV-2 est intervenu dans la genèse du rétrovirus E26.

### 4 - Organisation de la P135gag-myb-ets

Plusieurs unités fonctionnelles ont pu être définies dans la P135gag-myb-ets par analogie avec la structure des proto-oncogènes *c-myb* et *c-ets-1*. Les domaines de fixation à l'ADN de *v-myb* et *v-ets* sont situés respectivement à l'extrémité amino-terminale et à l'extrémité carboxy-terminale de la protéine. Deux domaines de transactivation (un pour v-Myb, un pour c-Ets) sont localisés au centre de la protéine, plus un autre domaine transactivateur situé dans la partie C-terminale de v-Ets (Klempnauer et Sippel, 1987 ; Weston et Bishop, 1989 ; Nye *et al.*, 1992 ; Schneikert *et al.*, 1992) (Figure 3).

### 5 - Les gènes cibles de la P135gag-myb-ets

L'expression de la P135gag-myb-ets par le virus E26 altère le déroulement normal de la différenciation, vraisemblablement par la régulation de



**Figure 2 :** Représentation schématique des divergences existant entre la protéine virale et la protéine Ets-1 aviaire, d'après Leprince *et al.*, 1993.

gènes essentiels à la prolifération et à la différenciation hématopoïétique.

La P135*gag-myb-ets* peut se fixer sur le promoteur du gène *mim-1* (Myb induced myeloid protein 1) qui s'exprime dans les cellules myéloïdes, et transactiver l'expression de ce gène par l'intermédiaire du domaine responsable de la fixation à l'ADN de v-Myb (Ness *et al.*, 1989).

Un autre gène cible de la P135*gag-myb-ets* a été découvert après construction d'un virus murin contenant les séquences v-*ets* et v-*myb* équivalent au rétrovirus E26 (ME26) et codant une protéine nucléaire de 135 kDa (Yuan *et al.*, 1989). Ce virus provoque des érythroleucémies chez le souriceau. Les cellules transformées sont des précurseurs très immatures nécessitant la présence d'érythropoïétine. La protéine virale est capable de transactiver le promoteur du gène codant le facteur GATA-1 qui est un effecteur de la régulation de la transcription du gène codant le récepteur de l'érythropoïétine. D'autre part, la P135*gag-myb-ets* coopère avec la protéine GATA-1 pour augmenter la transactivation du gène *GATA-1* (Aurigemma *et al.*, 1992 ; Ruscetti *et al.*, 1992).

## 6 - Les cellules cibles du rétrovirus E26

Le rétrovirus E26 a été isolé en Bulgarie en 1962 dans un poulet de basse-cour (Ivanov *et al.*, 1962). Il induit chez le poulet une leucémie mixte érythroïde et myéloïde (Moscovici *et al.*, 1981 ; Radke *et al.*, 1982) et transforme les myéloblastes et les érythroblastes de tissus en culture (Moscovici *et al.*, 1983). Trois types de colonies transformées sont obtenus (érythroïdes, myéloïdes et des colonies mixtes composées de cellules exprimant soit des marqueurs érythroïdes soit des marqueurs myéloïdes) (Moscovici *et al.*, 1983 ; Graf *et al.*, 1992). Le rétrovirus aurait donc pour cible des cellules progénitrices des lignées érythroïdes et myéloïdes et des cellules progénitrices bipotentes, communes aux deux lignées (Graf *et al.*, 1992). D'autres expériences tendent à prouver que toutes les voies de la différenciation hématopoïétique peuvent être transformées par le rétrovirus E26 à l'exception des voies lymphoïdes (Domenget *et al.*, 1992).

Le rétrovirus E26 est aussi capable d'infecter des fibroblastes et de modifier leurs caractéristiques morphologiques et physiologiques. Il ne peut pas transformer des fibroblastes embryonnaires (Moscovici, 1975 ; Graf *et al.*, 1979 ; Bister *et al.*, 1982), mais il a un effet mitogène sur ces fibroblastes, même avec une expression faible de la P135*gag-myb-ets* (Jurdic *et al.*, 1987).

La délétion des différents domaines de la P135*gag-myb-ets* a montré que les oncogènes v-*myb* et v-*ets* coopèrent pour faire acquérir aux fibroblastes de nouvelles capacités de prolifération (Ravel-Chapuis *et al.*, 1991). Le même phénomène est observé pour les cellules de neurorétine embryonnaires



infectées par le rétrovirus : la P135*gag-myb-ets* ne peut transformer ces cellules mais elle peut induire leur prolifération (Amouyel *et al.*, 1989).

## 7 - Implications respectives de *v-myb* et *v-ets*

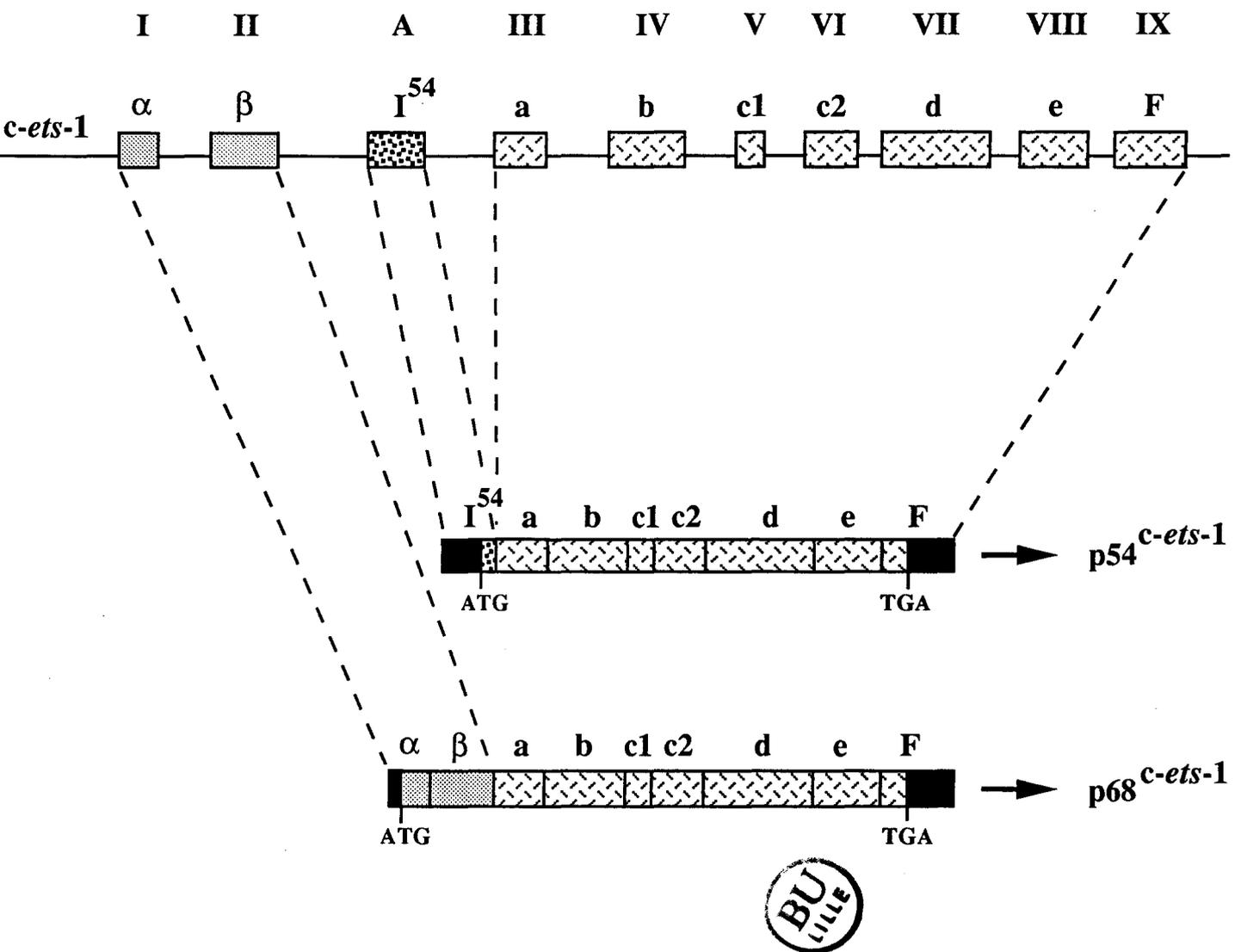
Le rétrovirus de la myéloblastose aviaire, AMV, qui ne possède dans son génome que l'oncogène *v-myb*, ne transforme que les cellules myéloïdes *in vivo* et *in vitro* (Moscovici *et al.*, 1975 ; Durban et Boettiger, 1981 ; Introna *et al.*, 1990). Des mutants de E26 délétés de *v-ets* sont incapables de transformer des cellules érythroïdes *in vivo* et *in vitro*, mais peuvent toujours transformer des cellules myéloïdes. La délétion du domaine de fixation à l'ADN de *v-Ets* inhibe la capacité de transformation des cellules érythroïdes (Golay *et al.*, 1988 ; Nunn et Hunter, 1989 ; Domenget *et al.*, 1992). De même, une mutation ponctuelle dans le domaine de fixation à l'ADN de *v-Myb* n'a pas d'effet sur la transformation érythroïde (Beug *et al.*, 1984) mais entraîne la perte des capacités à transformer des cellules myéloïdes, à 42°C (Frykberg *et al.*, 1988). D'après ces observations, la capacité de transformation des cellules myéloïdes est attribuée à la partie *Myb* de la protéine virale alors que la partie *Ets* permet la transformation des cellules érythroïdes.

Néanmoins, des travaux plus récents basés sur l'étude de virus recombinants exprimant soit *Myb* ou *Ets* seules, soit *Myb* et *Ets* mais séparées dans un même virus, soit *Myb* et *Ets* fusionnées dans la P135*gag-myb-ets*, semblent démontrer une coopération entre *v-Myb* et *v-Ets* au sein de la protéine de fusion (Metz et Graf, 1991a et b). L'expression de *v-myb* et *v-ets* sous la forme d'une protéine de fusion est indispensable à l'activité leucémogène des virus. Les virus qui expriment les protéines *Myb* et *Ets* séparément ne causent pas de leucémies, bien que ces deux protéines soient capables de coopérer dans les tests de transformation *in vitro*. Les domaines de fixation à l'ADN de *v-Myb* et de *v-Ets* de la protéine virale, ainsi que le domaine transactivateur de *v-Ets* peuvent être délétés sans altérer l'activité transformatrice (Metz et Graf, 1991b). Cependant, il semble que le domaine transactivateur de *v-Myb* et celui de *v-Ets* soient interchangeables car la délétion de ce domaine dans *v-Myb* ne perturbe pas la transformation myéloïde lorsque *v-Ets* est intact (Domenget *et al.*, 1992).

## **B - Identification des homologues cellulaires de l'oncogène viral *v-ets***

### **1 - Le proto-oncogène *c-ets-1* aviaire**

Les sondes synthétisées à partir de la séquence virale *v-ets* ont



$\alpha$  à F selon Leprince *et al.*, 1988  
 I à IX selon Watson *et al.*, 1988  
 ■ régions non codantes

Figure 4 : Le locus *c-ets-1* du poulet et ses produits.

permis le clonage de son homologue cellulaire chez le poulet : le proto-oncogène *c-ets-1*. Le locus Ck *c-ets-1* est réparti sur au moins 80 kpb d'ADN génomique (Figure 4). Les régions homologues à *v-ets* sont séparées en deux domaines situés de part et d'autre d'une séquence de 40 kpb non apparentée au gène viral. Le premier domaine est homologue aux 223 nucléotides de la partie 5' du *v-ets*. Il est composé de deux séquences exoniques nommées  $\alpha$  et  $\beta$ . Ces séquences ne sont pas retrouvées dans l'ARN messager majoritaire de 7,5 kpb exprimé particulièrement dans les cellules lymphoïdes (Gégonne *et al.*, 1987a ; Watson *et al.*, 1988a). Le deuxième domaine comprend 7 régions, a, b, c<sub>1</sub>, c<sub>2</sub>, d, e, F, homologues aux 1250 pb restantes de séquences virales.

La protéine majeure de poids moléculaire 54 kDa (*p54<sup>c-ets-1</sup>*) codée par le gène *c-ets-1* n'est pas immunoprécipitée par un anticorps dirigé contre la partie amino-terminale de *v-ets* alors que la *P135<sup>gag-myb-ets</sup>* est, elle, immunoprécipitée. Par contre, des anticorps dirigés contre la partie C-terminale de *v-ets* sont capables de reconnaître ces deux protéines (Ghysdael *et al.*, 1986a ; Gégonne *et al.*, 1987b). Le clonage de l'ADN complémentaire correspondant à la *p54<sup>c-ets-1</sup>* a permis de démontrer que *c-ets 1* est identique à *v-ets* excepté au niveau des extrémités 5' et 3' et de deux mutations ponctuelles.

Les 27 premiers acides aminés de la *p54<sup>c-ets-1</sup>* n'ont pas d'équivalent dans la *P135<sup>gag-myb-ets</sup>* (Duterque-Coquillaud *et al.*, 1988 ; Watson *et al.*, 1988a). Ces acides aminés N-terminaux sont en fait codés par un exon, appelé I<sup>54</sup>, localisé dans les 40 kpb séparant les exons  $\beta$  et a (Duterque-Coquillaud *et al.*, 1988). Dans la partie 3', l'exon F de *c-ets-1* est tronqué : les 13 derniers codons de cet exon sont remplacés par 16 codons d'origine inconnue dans *v-ets* (exon noté f pour *v-ets*).

Une autre protéine contenant les acides aminés codés par les exons  $\alpha$  et  $\beta$  a ensuite été mise en évidence (Leprince *et al.*, 1988). Cette protéine qui a un poids moléculaire de 68 kDa (*p68<sup>c-ets-1</sup>*) est le résultat de la traduction d'un ARNm de 7,5 kpb, uniquement détecté dans la rate.

Le locus Ck *c-ets-1* est donc transcrit en deux ARNm majeurs de 7,5 kpb résultant d'un processus d'initiation de la transcription et d'épissages alternatifs (Leprince *et al.*, 1988 ; Crépieux *et al.*, 1993) et en deux ARNm mineurs de 1,5 et 2 kpb (Leprince *et al.*, 1983 ; Chen, 1985).

L'ARNm correspondant à la *p54<sup>c-ets-1</sup>* est initié au niveau d'un promoteur en amont de l'exon I<sup>54</sup> : il contient l'exon I<sup>54</sup> accolé aux exons a à F identiques à *v-ets* (Duterque-Coquillaud *et al.*, 1988 ; Chen, 1988).

L'ARNm correspondant à la *p68<sup>c-ets-1</sup>* est, lui, initié au niveau d'un promoteur situé en amont d'un exon nommé Y (Crépieux *et al.*, 1993). Par

un processus d'épissage alternatif, l'exon I<sup>54</sup> est éliminé et l'ARNm généré comprend les exons  $\alpha$  et  $\beta$  fusionnés aux exons a à F. Cet ARNm est le progéniteur cellulaire de *v-ets*.

## 2 - Les homologues de *c-ets-1* chez l'homme, le poulet et la souris

Le proto-oncogène *c-ets-1* a ensuite été identifié chez d'autres vertébrés, ainsi qu'un autre gène hautement homologue, *c-ets-2*.

La plus grande homologie avec *c-ets-1* a été découverte en recherchant des séquences homologues de *v-ets* chez l'homme.

Chez l'homme, le locus *c-ets-1* ou Hu *c-ets-1* est localisé sur le bras long du chromosome 11 (De Taisne *et al.*, 1984). Il comprend 8 exons (A, III, IV, V, VI, VII, VIII, IX) qui sont distribués sur plus de 60 kpb. Les exons sont identiques en taille à ceux trouvés chez le poulet avec une identité de 95 % au niveau de la séquence traduite en acides aminés (Jorcyk *et al.*, 1991). Hu *c-ets-1* est transcrit en un ARNm majeur de 6,8 kpb (Watson *et al.*, 1985) et code quatre protéines nucléaires majeures de 39, 42, 48 et 51 kDa, ainsi que deux mineures de poids moléculaire 49 et 52 kDa. Ces différentes protéines sont le résultat d'épissages alternatifs, d'événements de phosphorylation et de modifications de groupements thiol.

Hu *c-ets-2* est situé sur le chromosome 21. Il est transcrit en trois ARNm de 2,7 ; 3,2 et 4,7 kpb (Watson *et al.*, 1985) et code une protéine de 54 kDa qui est rapidement phosphorylée (pp56).

Le locus *c-ets-2* est présent aussi chez le poulet (Boulukos *et al.*, 1988). Il est transcrit en un ARNm de 4 kpb et code des protéines de poids moléculaire de 58, 60, 62 et 64 kDa résultant d'une maturation post-traductionnelle différente (Boulukos *et al.*, 1988 ; Ghysdael *et al.*, 1986b). Les séquences *c-ets-1* et *c-ets-2* sont situées sur le même chromosome en des loci contigus.

Chez la souris, le locus *c-ets-1* est transcrit en un ARNm majeur de 5,3 kpb et en des ARNm mineurs de 2,2; 2,5 et 4 kpb. Il est fortement exprimé dans le thymus et code les protéines p63<sup>*c-ets-1*</sup> et p52<sup>*c-ets-1*</sup> résultant d'un épissage alternatif (Pognonec *et al.*, 1990). Le locus *c-ets-2*, transcrit en un ARNm majeur de 3,5 kpb et traduit en une protéine de 56 kDa, est exprimé dans la plupart des tissus.

Le domaine codé par l'exon I<sup>54</sup> est largement conservé chez les Vertébrés alors que les exons  $\alpha$  et  $\beta$  n'ont été détectés que chez les Oiseaux et les Reptiles (Albagli *et al.*, 1992). Chez les Mammifères, la p63<sup>*c-ets-1*</sup> murine et la

p55<sup>c-ets-1</sup> humaine sont des homologues de la p54<sup>c-ets-1</sup> aviaire.

Les gènes *c-ets-1* et *c-ets-2* ont ensuite été mis en évidence chez le xénope (Stiegler *et al.*, 1990 ; Wolff *et al.*, 1990), chez la drosophile (Pribyl *et al.*, 1988) et chez l'oursin (Chen *et al.*, 1988). D'autres gènes présentant une homologie avec les gènes *c-ets* ont été ensuite isolés dans de nombreuses espèces, donnant naissance à une grande famille de gènes, la famille ETS.

Afin de décrire les différents aspects de la famille des gènes *ets*, nous avons choisi d'approfondir l'étude du proto-oncogène *c-ets-1*. Le proto-oncogène *c-ets-1*, homologue cellulaire de l'oncogène viral *v-ets*, a été découvert chez le poulet, puis décrit dans d'autres espèces animales. Il est à l'origine de la découverte de la famille des gènes *ets* et est à ce jour le mieux étudié. Nous pourrions ainsi illustrer les différentes caractéristiques de cette famille de gènes avant de décrire, dans un prochain chapitre, les autres membres de cette famille ainsi que leurs relations respectives.

## **II - Modèle d'étude de la famille ETS : le proto-oncogène *c-ets-1***

### **A - Fonction des produits de *c-ets-1***

#### **1 - Les protéines Ets-1 sont des facteurs de transcription**

Les protéines Ets ont été définies comme étant des facteurs de transcription car elles reconnaissent un motif nucléotidique particulier et se lient à l'ADN de façon spécifique pour réguler l'expression de certains gènes.

##### **a) Le domaine de liaison à l'ADN de *c-ets-1***

Chez l'homme, la souris et le poulet, le gène *c-ets-1* code des protéines localisées dans le noyau des cellules, associées à la chromatine et possédant une capacité de liaison à l'ADN (Boulukos *et al.*, 1989 ; Pognonec *et al.*, 1989 ; Leprince *et al.*, 1990). Le domaine responsable de ces activités est situé à l'extrémité carboxy-terminale des protéines (Boulukos *et al.*, 1989 ; Reddy et Rao, 1990 ; Gégonne *et al.*, 1992 ; Lim *et al.*, 1992 ; Nye *et al.*, 1992 ; Wang *et al.*, 1992 ; Wasylyk *et al.*, 1992). Ce domaine très conservé au cours de l'évolution et qui a permis de définir les différents membres de la famille ETS a été appelé le domaine ETS (Karim *et al.*, 1990 ; MacLeod *et al.*, 1992).

Le domaine ETS, qui comprend environ 85 acides aminés hautement conservés, constitue un nouveau motif structural de reconnaissance ADN-protéines. Ce domaine est essentiel pour la liaison spécifique des protéines Ets à l'ADN. Les protéines contenant cette séquence conservée constituent donc une nouvelle famille de protéines eucaryotiques se liant à l'ADN (Karim *et al.*, 1990).

#### b) Séquences de fixation des protéines Ets (EBS)

Les premières expériences montrant que les protéines Ets reconnaissent une séquence spécifique ont été réalisées par criblage d'une banque d'expression d'ADN complémentaire de thymus de souris avec une sonde oligonucléotidique représentant 20 pb du LTR du virus MSV (Moloney murine Sarcoma Virus) (Gunther *et al.*, 1990).

La protéine de 30 kDa ainsi produite se lie spécifiquement sur cet oligonucléotide. La séquence déduite en acides aminés de l'ADN complémentaire ainsi isolé montre qu'il code les 272 acides aminés carboxy-terminaux de la p63<sup>c-ets-1</sup> murine.

Wasylyk *et al.* (1990) ont également montré que les protéines p54<sup>c-ets-1</sup> et p68<sup>c-ets-1</sup> (ainsi d'ailleurs que les p58-64<sup>c-ets-2</sup>) transactivent l'"enhancer" du virus du polyome dans des fibroblastes LMTK<sup>-</sup>. Cet "enhancer" comprend le domaine de réponse à l'oncogène composé de deux sites juxtaposés AP-1 ("Activator Protein-1") et PEA3 ("Polyome Enhancer Activity 3") qui coopèrent en *cis* pour médier l'activation transcriptionnelle à partir de différents promoteurs par les facteurs de croissance du sérum, le promoteur de tumeurs TPA et plusieurs oncogènes (Wasylyk *et al.*, 1989, 1990). La p68<sup>c-ets-1</sup> se lie spécifiquement, mais avec une affinité modérée, à l'élément PEA3 de l'"enhancer" (5' CGAGGAAGTG 3') : une mutation de 4 pb dans cette séquence prévient la transactivation par la p68<sup>c-ets-1</sup> alors que les mutations dans le site AP-1 n'ont aucun effet (Wasylyk *et al.*, 1990). La séquence PEA3 ne représente cependant pas le site de fixation optimal pour les protéines Ets-1 et Ets-2.

Des séquences de fixation des protéines Ets, appelées EBS pour "Ets Binding Site" ont été identifiées dans des séquences de régulation de la transcription de différents gènes viraux et cellulaires incluant le promoteur du MSV (Gunther *et al.*, 1990), des récepteurs  $\alpha$  et  $\beta$  des cellules T (TCR- $\alpha$  et  $\beta$ ) (Ho *et al.*, 1990 ; Prosser *et al.*, 1992), le motif PEA3 de l'"enhancer" du polyome (Wasylyk *et al.*, 1990), le LTR du virus HTLV-1 (Human T Lymphotropic Virus type 1) (Bosselut *et al.*, 1990 ; Gitlin *et al.*, 1991), le promoteur de la stromélysine 1 (Wasylyk *et al.*, 1991), de *c-ets-1* et *c-ets-2*

Protéine	Séquence de fixation à l'ADN	Promoteur/Enhancer
Ets-1 (poulet)	GCA <b>GGAA</b> GCATT	Stromélysine
Ets-1 (poulet)	TCGA <b>GGAA</b> GTGAC	Virus du polyome
Ets-1 (homme)	GAAA <b>GGAA</b> GGGCG	Promoteur <i>Ets-1</i>
Ets-1 (homme)	GGA <b>GGAA</b> ATGG	LTR du HTLV-1 (ERE-A)
Ets-1 (homme)	CCG <b>GGAA</b> GCCA	LTR du HTLV-1 (ERE-B)
Ets-1 (homme)	CGC <b>GGAA</b> ATGT	LTR du HTLV-1 (ERR-2)
Ets-1 (homme)	CAGA <b>GGAT</b> GTGG	TCR- $\alpha$
Ets-1 (homme)	GCA <b>GGAA</b> GT	Promoteur <i>Ick</i> type I
Ets-1 (homme)	CGA <b>GGAA</b> GG	Promoteur <i>GATA-1</i>
Ets-1 (souris)	GAGC <b>GGAA</b> GCGC	LTR du MSV

**Tableau 1 : Quelques séquences de fixation des protéines Ets-1.**

(Mavrothalassitis et Papas, 1991 ; Oka *et al.*, 1991 ; Ascione *et al.*, 1992 ; Majérus *et al.*, 1992 ; Crépieux *et al.*, 1993), de la collagénase 1 (Nerlov *et al.*, 1992), de l'activateur du plasminogène de type urokinase (uPA) (Gutman et Wasylyk, 1990 ; Rorth *et al.*, 1990) ainsi que dans le promoteur du gène *mb-1* (Hagman et Grosschedl, 1992). Plus récemment, de tels sites ont été mis en évidence dans le promoteur du gène *Ick* de type I qui code une tyrosine kinase lymphocytaire (Leung *et al.*, 1993) et dans le promoteur du gène du facteur GATA-1 (Seth *et al.*, 1993) (Tableau 1).

La comparaison de différentes séquences de fixation des protéines Ets a permis de définir le noyau GGAA, riche en purine, qui est retrouvé dans tous les sites de fixation des protéines Ets (Karim *et al.*, 1990).

Une mutation au niveau d'un G (GGAA → AGAA) abolit complètement la fixation spécifique à l'ADN de la p68<sup>c-ets-1</sup> (Wasylyk *et al.*, 1990). D'autre part, une différence d'affinité entre les protéines Ets-1 et Ets-2 en fonction des promoteurs reconnus a été mise en évidence, indiquant l'importance des nucléotides proximaux de la séquence GGAA/T (Ascione *et al.*, 1992).

A partir du fait que les protéines Ets-1 reconnaissent des séquences possédant le noyau GGAA/T, des expériences ont été tentées dans le but de pouvoir définir des séquences consensus de fixation d'une protéine donnée. Les séquences de fixation des protéines Ets, ont été déterminées par les techniques de retard sur gel quantitatif associées à la sélection d'oligonucléotides de séquence aléatoire par PCR (technique SAAB : "Selected And Amplified Binding"), soit avec le domaine ETS, soit avec la protéine Ets-1 purifiée.

De cette façon, des séquences consensus pour les protéines Ets-1 aviaire (Woods *et al.*, 1992), murine (Nye *et al.*, 1992) et humaine (Fisher *et al.*, 1991) ont été définies, *in vitro* :

Ets-1 (poulet)	A C C C	GGAA T	A T T G
Ets-1 (souris)	A C C G	GGAA T	G T C
Ets-1 (homme)	G G C A C A	GGAA	G T C

Les nucléotides flanquant le noyau GGAA/T sont donc nécessaires à la reconnaissance par les protéines Ets et à la sélectivité d'action dans la transactivation. En dehors de ce motif, il n'existe pas de séquence consensus identifiée.

### **Proto-oncogènes**

Ets-1  
Ets-2  
GM-CSF  
Fos

### **Récepteurs**

Chaîne  $\alpha$  récepteur cellules T  
Chaîne  $\beta$  récepteur cellules T  
Récepteur transferrine

### **Virus**

LTR du MSV  
Enhancer du polyome  
SV 40  
LTR du HTLV-1  
Antigène cytoplasmique précoce EBV

### **Facteurs de croissance**

Sous-unité  $\alpha$  du NGF Souris  
Interleukine 2 (IL-2)  
Interleukine 3 (IL-3)  
Interleukine 4 (IL-4)  
Interleukine 6 (IL-6)

### **Autres**

Facteur de transcription, GATA-1  
Urokinase  
Stromelysine 1  
ARN nucléaire U 3  
Collagénase 1  
Synthétase humaine 5A  
Interféron humain inductible IFI-54K  
Facteur IX humain  
LTE 12 de levure  
MRS 3 de levure  
Alcool deshydrogénase de drosophile  
Gène H2-K MCH classe I souris  
Gène I-A $\beta$  MCH classe II souris

**Tableau 2 : Liste de gènes dont les promoteurs contiennent des motifs Ets de liaison à l'ADN (AGGAAAG), d'après Seth *et al.*, 1992.**

### c) Reconnaissance moléculaire

Le domaine de fixation à l'ADN de Ets-1, qui se lie à la séquence GGAA/T, est localisé dans le grand sillon de l'ADN. La zone de contact s'étend au niveau des petits sillons adjacents et recouvre une vingtaine de bases. Ce mode de fixation n'aurait encore jamais été observé pour d'autres facteurs de transcription et serait particulier au domaine ETS (Nye *et al.*, 1992).

## 2 - Les gènes cibles des protéines Ets-1

L'identification d'EBS dans différents promoteurs/"enhancers" a permis l'identification de gènes pouvant être régulés par des protéines Ets. L'expression de ces gènes est souvent corrélée à l'expression des protéines Ets. La liste précédemment citée n'est pas exhaustive et pour l'instant, le motif EBS a été mis en évidence dans au moins 30 promoteurs de gènes uniques (Tableau 2). La plupart de ces gènes semblent être activés dans les cellules hématopoïétiques.

Un nombre significatif de gènes spécifiques des cellules lymphoïdes contiennent des EBS dans leurs séquences de régulation. Les EBS de l'"enhancer" du TCR- $\beta$  médient l'expression du gène codant le TCR- $\beta$ , qu'elle soit basale ou induite par l'activation des cellules T par le TPA (Prosser *et al.*, 1992). L'expression de Ets-1 entraîne la répression de l'activité basale ou induite de l'"enhancer" du TCR- $\beta$  (alors que l'expression de Ets-2 n'a aucun effet). L'effet de dérèglement de la transcription du TCR- $\beta$  pourrait être médié par l'inactivation de l'activité de liaison à l'ADN de Ets-1.

L'induction simultanée de l'expression de *c-ets-1* et du TCR- $\alpha$  dans les thymocytes au 18ème jour du développement de l'embryon de souris, ajoutée au fait qu'une mutation de l'EBS dans l'"enhancer" du TCR- $\alpha$  empêche l'activité de transcription (Ho *et al.*, 1990), suggère que Ets-1 intervient dans la régulation transcriptionnelle de ce gène dans les lymphocytes T (Bhat *et al.*, 1989). D'autre part, les protéines Ets-1 (et Ets-2) peuvent activer le promoteur du gène *CD4 in vitro*, qui est activé dans certaines sous-populations de lymphocytes T (Salmon *et al.*, 1993).

Les protéines Ets-1 (et Ets-2) transactivent le LTR du virus HTLV-1 dans les cellules NIH3T3 et HeLa (Bosselut *et al.*, 1990). Il existe au moins deux régions de liaison pour Ets-1 dans le LTR du HTLV-1, ERR-1 et ERR-2, pouvant médier la transactivation. Des analyses de retard sur gel compétitif ont montré que l'interaction de Ets-1 avec ERR-1 est plus stable qu'avec ERR-2. ERR-1 contient deux EBS : ERE-A et ERE-B et l'affinité de Ets-1 pour ERE-A est supérieure à l'affinité pour ERE-B (Gitlin *et al.*, 1991).

Le promoteur du gène codant le facteur de transcription

GATA-1, spécifique des cellules érythroïdes, chez le poulet, contient trois EBS dont deux forment un palindrome. La protéine Ets-1 ne se lie qu'aux EBS palindromiques (Seth *et al.*, 1993).

Des EBS ont aussi été identifiés au niveau de promoteurs de gènes impliqués dans la dégradation de la matrice extracellulaire comme la stromélysine 1 (Wasylyk *et al.*, 1991), la collagénase 1 (Nerlov *et al.*, 1992), ainsi que l'activateur du plasminogène de type urokinase (uPA) (Gutman et Wasylyk, 1990 ; Rorth *et al.*, 1990). Les protéines Ets-1 activent l'expression de la stromélysine 1 du rat en test de cotransfection. Les deux EBS de ce promoteur sont aussi efficaces que la séquence PEA3 du virus du polyome présente en quatre exemplaires (Wasylyk *et al.*, 1991).

Ets-1 peut aussi transactiver son propre promoteur (Oka *et al.*, 1991 ; Ascione *et al.*, 1992 ; Majérus *et al.*, 1992) suggérant que les protéines Ets-1 ou d'autres membres de la famille ETS peuvent réguler leur expression et ainsi moduler leur niveau à l'intérieur d'un type cellulaire spécifique. A cet égard, il a été montré qu'une surexpression de Ets-1 dans les cellules NIH3T3 induit l'expression du gène endogène murin *ets-1*, indiquant que *ets-1* peut réguler sa propre expression (Seth et Papas, 1990).

Dans la plupart des cas, pour qu'un promoteur soit inductible (*in vitro*) par les protéines Ets, il faut la présence dans ce promoteur de plusieurs EBS. Le nombre d'EBS est important pour avoir une activation maximale de la transcription. Pour les promoteurs ne possédant qu'un seul EBS ("enhancer" TCR- $\alpha$ , LTR du MSV, promoteurs de la collagénase et de l'uPA), la transactivation pourrait être régulée par une association entre les protéines Ets et d'autres facteurs activateurs comme cela a été montré pour Ets-1 et AP-1 (Gutman et Wasylyk, 1990 ; Wasylyk *et al.*, 1990).

Les protéines Ets-1 se fixent sous forme monomérique sur leurs séquences (Wasylyk *et al.*, 1990 ; Fischer *et al.*, 1991 ; Hagman et Grosschedl, 1992 ; Nye *et al.*, 1992) mais pourraient peut-être se fixer sous forme dimérique dans certains cas (Gitlin *et al.*, 1991).

En résumé, Ets-1 peut agir en tant qu'activateur ou répresseur (TCR- $\beta$ ). Son activité dépend du contexte du promoteur/"enhancer" et du type cellulaire. L'analyse de la transactivation de différents gènes cibles suggère un rôle des protéines Ets-1 à la fois dans la différenciation de la lignée lymphoïde et dans le contrôle de la dégradation de la matrice extracellulaire.

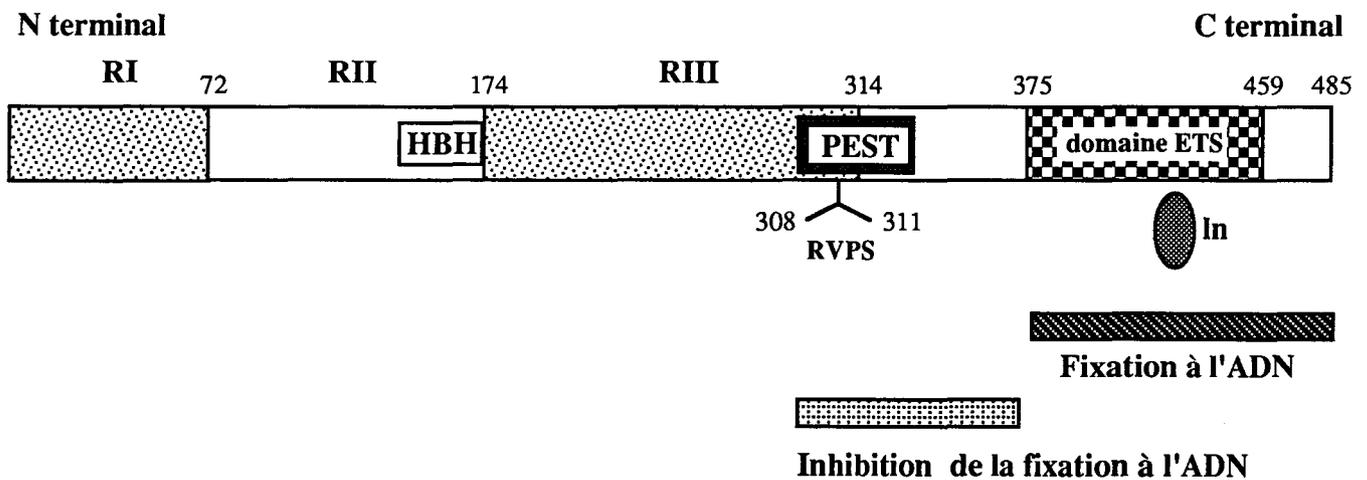
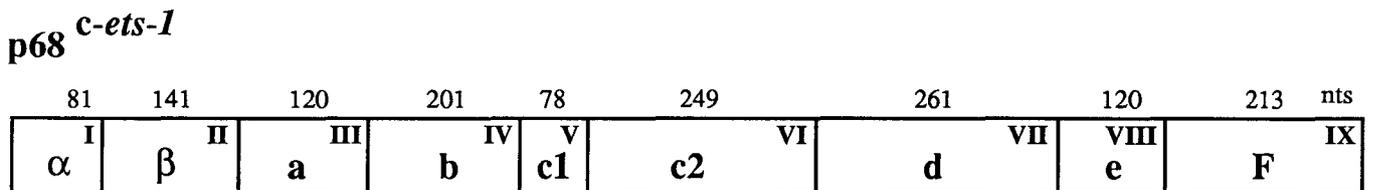
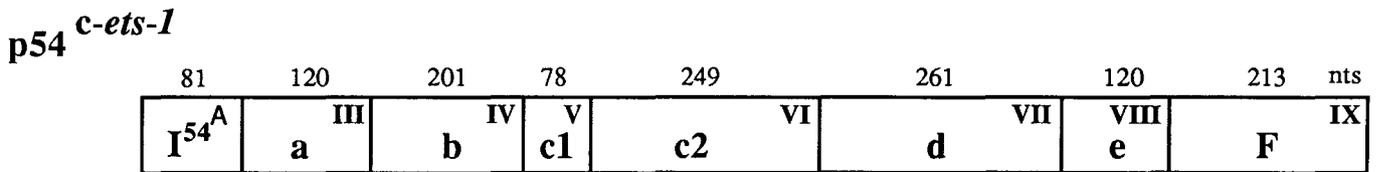
### 3 - Coopération des protéines Ets-1 avec d'autres facteurs de transcription

La p68c-*ets-1* est capable de coopérer avec le complexe AP-1 (Fos-Jun) sur l'élément de réponse à l'oncogène de l'"enhancer" des gènes précoces du virus du polyome qui contient un EBS et un site AP-1 juxtaposés (Wasylyk *et al.*, 1990). Ces auteurs suggèrent que les sites de liaison de Ets-1 et de AP-1 sont requis pour qu'il y ait un effet coopératif, car il n'y a pas de complémentation si l'un des deux sites est absent de l'"enhancer".

Cependant, la coopération entre AP-1 et la p68c-*ets-1* n'est pas médiée au niveau de la liaison à l'ADN puisque la diminution de l'affinité pour la liaison à l'ADN de l'EBS ou du site AP-1 n'entraîne pas la capacité pour ces facteurs de coopérer pour la transcription. Une surexpression de la p68c-*ets-1* entraîne une inhibition de la transactivation par AP-1 de l'unité de réponse aux oncogènes mutée dans son site PEA3. Ces observations suggèrent que la coopération entre AP-1 et la p68c-*ets-1* est médiée par des interactions protéines-protéines (Wasylyk *et al.*, 1990). La coexpression des protéines Ets-1 et du complexe AP-1 entraîne aussi la stimulation du promoteur de l'urokinase (Nerlov *et al.*, 1992). Lorsque le promoteur est stimulé par le TPA, des protéines intermédiaires interviennent pour permettre la coopérativité entre un élément Ets/AP-1 et un site AP-1, situés à grande distance (Nerlov *et al.*, 1992).

Un phénomène de coopération a été démontré pour la protéine p54c-*ets-1* et le facteur de transcription Sp-1 *in vitro* (Gégonne *et al.*, 1993). Ces deux facteurs coopèrent pour la transactivation et la liaison à l'ADN au niveau de la région ERR-1 du LTR du virus HTLV-1. Sp-1 et Ets-1 forment un complexe ternaire avec l'élément Sp-1/ERE-A. Ets-1 et Sp-1 agissent de façon synergique pour activer la transcription à partir des éléments ERR-1 ou Sp-1/ERR-A et pour cela, la liaison de Sp-1 et de Ets-1 au niveau de leur site de reconnaissance est requise. Les éléments Sp-1/ERE-A sont trouvés dans la région promotrice/"enhancer" de plusieurs gènes cellulaires, semblant montrer que la synergie entre Ets-1 et Sp-1 n'est pas restreinte à la région LTR de l'"enhancer" du virus HTLV-1 (Gégonne *et al.*, 1993). Ces résultats suggèrent que les protéines Ets fonctionneraient comme des composants de complexes de transcription plus importants pour réguler l'activité d'une variété de gènes viraux et cellulaires.

Un autre exemple de coopération a été mis en évidence concernant la protéine p54c-*ets-1* et le facteur GATA-1 exprimé dans la lignée érythrocytaire, dans les mégacaryocytes et dans les cellules de moëlle osseuse, de façon spécifique. Le promoteur du gène de la glycoprotéine IIB contient un EBS et un site de liaison pour le facteur GATA-1 situés à proximité. La protéine Ets-1



RI, RIII, domaines de transactivation,  
 RII, domaine de régulation,  
 HBH, région en hélice-boucle-hélice,  
 PEST, région riche en Pro, Glu, Ser et Thr,  
 RVPS, séquence consensus de phosphorylation par la calcium/calmoduline kinase II,  
 In, séquence de localisation nucléaire.

**Figure 5 : Structure et domaines fonctionnels des protéines c-Ets-1 du poulet.**

et le facteur GATA-1 coopèrent de façon additive pour transactiver le promoteur GPIIB et assurer l'expression spécifique de la glycoprotéine dans les mégacaryocytes (Lemarchandel *et al.*, 1993).

## **B - Les domaines fonctionnels des protéines Ets-1**

### **1 - Les domaines de régulation de la transcription**

Des premiers essais de fusion de la protéine Ets-1 de poulet avec le domaine de liaison à l'ADN GAL4 de levure suggéraient que la protéine Ets-1 possédait un ou plusieurs domaines de transactivation intrinsèques (Seneca *et al.*, 1991).

Des expériences utilisant la fusion de mutants de délétion de la protéine p68<sup>c-ets-1</sup> et un domaine de fixation à l'ADN hétérologue ont permis d'identifier deux domaines de transactivation RI et RIII et un domaine de régulation RII, localisés dans la partie N-terminale de cette protéine (Schneikert *et al.*, 1992) (Figure 5).

Les domaines RI et RIII sont capables d'activer la transcription de façon indépendante. Le domaine RI correspond aux 71 acides aminés N-terminaux codés par les exons  $\alpha$  et  $\beta$ . Le domaine RIII est localisé entre les acides aminés 174 et 314 correspondant aux exons V, VI et une partie de l'exon VII. Un domaine correspondant à RIII a été identifié dans la p55<sup>c-ets-1</sup> humaine entre les acides aminés 131 et 242 (Gégonne *et al.*, 1992). L'exon VII qui n'est pas présent dans la p41<sup>c-ets-1</sup> chez l'homme n'est pas nécessaire à l'activité transactivatrice (Gégonne *et al.*, 1992).

Les séquences de ces deux domaines sont différentes et ne présentent pas beaucoup d'homologie dans leur composition en acides aminés : le domaine RI est riche en résidus cystéine (10 %) et proline (14 %) et pauvre en acides aminés basiques ; RIII est lui riche en résidus sérine et présente un caractère acide.

RI possède une succession de cytéines pouvant former un doigt de zinc impliqué dans des interactions protéines-protéines, qui ressemble au domaine transactivateur de la protéine Tat du virus HIV-1 (Human Immunodeficiency Virus-1) (Kuppuswamy *et al.*, 1989). Aucune séquence similaire n'a cependant pu être détectée dans d'autres domaines des protéines Ets ni dans des domaines d'activation de facteurs de transcription connus.

Les 71 acides aminés qui spécifient la partie N-terminale de la p68<sup>c-ets-1</sup> apparaissent globalement très hydrophobes (Leprince *et al.*, 1990) et pourraient former une structure en boucle  $\omega$ . Cette boucle peut être exposée à la

surface de la protéine et pourrait ainsi interagir avec d'autres protéines par des interactions hydrophobes.

Les 17 acides aminés N-terminaux de la p54<sup>c-ets-1</sup> codés par l'exon I<sup>54</sup> ne sont pas capables d'activer la transcription lorsqu'ils sont fusionnés à un domaine de fixation hétérologue (Schneikert *et al.*, 1992 ; Gégonne *et al.*, 1992). La p54<sup>c-ets-1</sup> ne possède donc que les domaines RII et RIII.

La région RII située entre les deux régions transactivatrices RI et RIII a une fonction modulatrice de la transactivation. Elle ne peut activer la transcription seule, mais combinée avec les domaines d'activation, RII a un effet de régulation négative sur l'activité de RI et un effet positif sur l'activité de RIII.

RII qui contient un motif hélice-boucle-hélice est conservée entre Ets-1 et Ets-2 (70 % d'homologie), suggérant que cette région puisse médier un contrôle commun des interactions de Ets-1 et Ets-2. L'hélice-boucle-hélice peut être impliquée dans des interactions protéines-protéines avec d'autres facteurs pouvant affecter l'activité de RII. Des expériences de compétition entre Ets-1 et Ets-2 ont montré que ces protéines interagissent avec un coactivateur commun présent en quantité limitante dans les cellules. Ces cofacteurs ou adaptateurs serviraient de liaison entre les facteurs de transcription et la machinerie transcriptionnelle (Schneikert *et al.*, 1992). La différence d'activité des protéines résiderait dans les mécanismes par lesquels les facteurs se lient à l'ADN interagissent avec d'autres activateurs ou coactivateurs de la transcription.

## **2 - Répression intramoléculaire de la fixation à l'ADN**

Différentes séries d'expériences ont suggéré l'existence dans les protéines Ets-1 humaines, aviaires et murines, de régions régulant de manière négative l'activité de liaison à l'ADN (Lim *et al.*, 1992 ; Nye *et al.*, 1992 ; Hagman et Grosschedl, 1992 ; Wasylyk *et al.*, 1992 ; Wasylyk et Wasylyk, 1993).

La protéine v-Ets, ainsi que le peptide représentant le domaine ETS pris isolément, se lient à des EBS avec une affinité supérieure par rapport aux protéines Ets-1 de pleine longueur (Nye *et al.*, 1992 ; Lim *et al.*, 1992). L'élimination d'une partie des acides aminés carboxy-terminaux de v-Ets n'affecte pas sa capacité de liaison alors que la même délétion dans la protéine c-Ets-1 entraîne une augmentation considérable de son activité de liaison. Ces résultats sembleraient donc démontrer que les acides aminés carboxy-terminaux de c-Ets-1 répriment la capacité de liaison à l'ADN de la protéine alors que les acides aminés C-terminaux de v-Ets ne sont pas requis pour la liaison à l'ADN (Hagman et Grosschedl, 1992 ; Lim *et al.*, 1992). Ces auteurs préconisent une interaction de cette région avec les acides aminés interagissant avec l'ADN, ou avec une autre

région de la protéine. L'intégrité de l'extrémité carboxy-terminale de c-Ets serait donc nécessaire pour la suppression de la liaison à l'ADN mais pas essentielle pour la liaison à l'ADN.

Une deuxième région est impliquée dans la répression de la liaison à l'ADN de la p68<sup>c-ets-1</sup> (Lim *et al.*, 1992 ; Wasylyk *et al.*, 1992). Cette région est située au centre de la molécule (Lim *et al.*, 1992) (Figure 5). Des délétions successives ont permis de localiser précisément cette région au niveau des acides aminés 291 à 377 correspondant à l'exon VII (Wasylyk *et al.*, 1992). Le domaine ETS isolé se lie avec la même spécificité à l'ADN lorsqu'il est associé au domaine répresseur, mais avec une affinité moins grande (Lim *et al.*, 1992). L'association entre le domaine ETS et le domaine de répression empêche la reconnaissance du domaine de fixation par un sérum dirigé contre lui (Wasylyk *et al.*, 1992). Le domaine répresseur pourrait agir en masquant le site de fixation de l'ADN à la protéine. La répression de la liaison à l'ADN ne serait donc pas due à l'interaction avec des cofacteurs et la déphosphorylation des protéines ne restaure pas la capacité de liaison à l'ADN (Lim *et al.*, 1992).

Selon Lim *et al.* (1992) les deux domaines répresseurs sont nécessaires pour réprimer la fixation à l'ADN. Ces auteurs ont formulé l'hypothèse suivante : l'interaction moléculaire entre ces deux domaines de c-Ets maintiendrait la protéine dans une conformation fermée incapable de se lier à l'ADN. Un changement conformationnel résultant en une configuration ouverte de la protéine capable de lier l'ADN serait dû à l'élimination d'une des deux régions répressives ou à un autre mécanisme cellulaire inconnu (Lim *et al.*, 1992). D'autre part, aux limites du domaine de fixation à l'ADN, il existe des régions riches en résidus proline et alanine qui sont flexibles, suggérant la possibilité de former un coude créant ainsi un repliement permettant l'interaction de domaines éloignés (Lim *et al.*, 1992 ; Wasylyk *et al.*, 1992).

D'autres auteurs ne sont pas d'accord sur le rôle inhibiteur de la partie C-terminale. Pour ces auteurs, la région carboxy-terminale de Ets-1 serait directement impliquée dans la fixation à l'ADN et/ou modulerait son activité par un effet structural indirect, alors que la région carboxy-terminale de v-Ets, qui est différente serait, elle, inhibitrice (Leprince *et al.*, 1992 ; Wasylyk et Wasylyk, 1992).

D'autres travaux récents vont dans le sens de cette théorie : le domaine ETS serait suffisant pour la liaison à l'ADN. Le domaine carboxy-terminal situé au delà du domaine ETS altérerait la sensibilité aux conditions de liaison. Une modification dans ce domaine induirait une conformation ouverte, exposant ainsi le domaine ETS aux anticorps et aux enzymes. Cette conformation

ouverte pourrait conférer une sensibilité à la régulation par phosphorylation ou par oxydoréduction (Wasylyk et Wasylyk, 1993). L'activité de liaison à l'ADN de v-Ets est particulièrement sensible aux conditions d'incubation (salinité, présence d'ions, phosphorylation, potentiel redox...). L'augmentation de sensibilité serait due à la perte de la fonction protectrice des séquences C-terminales de c-Ets-1. Les auteurs suggèrent que le domaine de liaison à l'ADN de v-Ets est plus accessible aux interactions protéines-protéines et aux mécanismes de régulation facilitant ainsi la transformation (Wasylyk et Wasylyk, 1993).

En résumé, il n'y aurait qu'une région impliquée dans l'inhibition de la liaison à l'ADN, la région correspondant à l'exon VII. La région C-terminale aurait, elle, un rôle dans la sensibilité aux conditions de liaison.

### C - Expression de c-ets-1

Chez le poulet, la p54<sup>c-ets-1</sup> qui représente le principal produit de traduction est préférentiellement exprimée dans les cellules lymphoïdes (B et T), du thymus et de la bourse de Fabricius (Ghysdael *et al.*, 1986a ; Pognonec *et al.*, 1988) et de la rate (Chen, 1985). Par contre, la p54<sup>c-ets-1</sup> est faiblement exprimée dans les lymphocytes circulants et dans les autres tissus testés (embryons, moëlle osseuse, fibroblastes, macrophages et érythroblastes) (Ghysdael *et al.*, 1986a).

Chez l'homme, des résultats similaires montrent une expression importante de *ets-1* dans les cellules lymphoïdes et hématopoïétiques ainsi que dans les astrocytes du cortex (Amouyel *et al.*, 1988).

De même, chez la souris, l'expression d'*ets-1* est très importante dans les cellules hématopoïétiques et dans le thymus (Bhat *et al.*, 1987, 1989). Dans le thymus de souris, l'expression de *c-ets-1* augmente au cours de l'ontogenèse pour se stabiliser après la naissance (Pognonec *et al.*, 1988 ; Bhat *et al.*, 1987, 1989). Le thymus est le site majeur de développement des cellules T, de leur différenciation et de leur maturation fonctionnelle. Il contient des populations cellulaires à des stades différents de prolifération et de maturation. L'expression du gène *ets-1* a été examinée dans des thymocytes fœtaux à différents stades du développement, dans des sous-populations de thymocytes adultes et dans des lymphocytes T périphériques. L'expression d'*ets-1* est initialement détectée dans des thymocytes fœtaux de 18 jours, correspondant à la première apparition des thymocytes auxiliaires différenciés. Le maximum d'expression est trouvé dans le thymus, un à deux jours après la naissance.

Dans le thymus d'adulte, les niveaux d'expression les plus importants sont trouvés dans les thymocytes CD4<sup>+</sup>, un des sous-types cellulaires les plus

matures, alors que le niveau est beaucoup plus faible dans les lymphocytes T auxiliaires ou cytotoxiques.

Selon d'autres auteurs, qui utilisent une méthode de fractionnement des sous-types cellulaires différente, la protéine Ets-1 murine est détectée à un même niveau dans les différentes populations lymphocytaires (Pognonec *et al.*, 1988). L'expression du gène *ets-1* semble donc corrélée à la maturation fonctionnelle des thymocytes plutôt qu'à leur prolifération cellulaire (Bhat *et al.*, 1989), puisque des thymocytes proliférants n'expriment pas plus *ets-1* que des thymocytes différenciés.

Le niveau d'expression d'*ets-1* dans les cellules T périphériques quiescentes est cependant plus faible que dans les thymocytes. Si ces lymphocytes quiescents sont stimulés par un anticorps anti-CD3 mimant l'interaction TCR-antigène, par un mitogène ou un ionophore calcique, l'expression de *c-ets-1* diminue de façon considérable (Bhat *et al.*, 1990). Ces résultats suggèrent que les produits de *c-ets-1* sont nécessaires pour maintenir les cellules T à un stade quiescent et que l'expression de *c-ets-1* doit être réprimée lors de l'activation lymphocytaire.

La recherche de l'expression des gènes *ets-1* (et *ets-2*) a été effectuée au cours du développement embryonnaire de la souris (Kola *et al.*, 1993). Ces auteurs montrent que l'expression du gène *ets-1* est beaucoup plus importante pendant le développement embryonnaire, comparée au taux à la naissance et au stade adulte. L'expression du gène *ets-1* augmente dans l'embryon après l'implantation et pendant l'organogenèse. Une expression importante est mise en évidence dans tous les organes des embryons de 15 jours examinés (intestin, estomac, rate, rein, langue, foie, cœur, cerveau et thymus).

A des stades fœtaux plus tardifs, l'expression d'*ets-1* est prédominante dans les tissus lymphoïdes, le cerveau et dans les organes qui sont remodelés ou qui subissent des ramifications lors de l'organogenèse (rein, langue et glandes salivaires), mais est réduite de façon considérable dans les autres organes comme l'estomac et l'intestin. Après la naissance, *ets-1* n'est exprimé que dans les organes lymphoïdes et le cerveau alors que chez l'adulte, le gène est uniquement exprimé dans les organes lymphoïdes.

Des expériences d'hybridation *in situ* ont permis de montrer que le gène *ets-1* intervient dans les cellules du mésenchyme des organes en développement, dans le système nerveux et dans la formation des os.

L'expression du gène *ets-1* pendant le développement embryonnaire semble donc associée à la morphogenèse et/ou au modelage des tissus. Ces processus morphogénétiques impliquent la dégradation et le remodelage de la

matrice extracellulaire. Or, les protéines Ets-1 sont connues pour réguler la transcription de métalloprotéinases dégradant la matrice comme la stromélysine et l'uPA. La fonction d'Ets-1 au cours de ces processus morphogénétiques pourrait être de réguler la transcription des gènes impliqués dans la dégradation et le remodelage de la matrice extracellulaire.

Des travaux effectués sur des ovaires de souris adultes ont montré une expression importante d'ARNm du gène *ets-1* (et *ets-2*) (Rowe et Propst, 1992). Par hybridation *in situ*, ces auteurs montrent que l'expression est localisée au niveau des cellules de la thèque et du corps jaune. D'autre part, des extraits d'ovaires de souris adultes présentent des facteurs pouvant se lier à des EBS, semblant ainsi prouver la présence de protéines de la famille ETS fonctionnellement actives. Les cellules de la thèque qui produisent les androgènes et les progestatifs interviennent dans le remodelage du tissu ovarien, en association avec le développement des follicules. Les gènes *ets* pourraient contrôler la régulation transcriptionnelle de gènes impliqués dans la réorganisation de ces tissus.

Des expériences d'hybridation *in situ* chez le poulet ont montré qu'au cours de l'embryogenèse, *ets-1* est exprimé dans toutes les lignées d'origine mésodermique, dont certaines sont à l'origine du système sanguin. L'expression la plus forte a été mise en évidence dans les cellules endothéliales formant la paroi interne des vaisseaux sanguins. En dehors du système vasculaire, *c-ets-1* est exprimé dans des groupes de cellules mésodermiques engagées dans des processus morphogénétiques mais ne l'est pas dans les épithéliums (Vandenbunder *et al.*, 1989). L'expression transitoire de *c-ets-1* au début de la formation des vaisseaux sanguins a aussi été montrée au cours de l'embryogenèse humaine ainsi que dans la cicatrisation des plaies et la vascularisation des tumeurs (Wernert *et al.*, 1992).

L'expression de *c-ets-1* a été bien étudiée lors de l'édification du germe plumaire, lors de processus de migration cellulaire des cellules dermiques vers le site où s'édifie la future plume (Desbiens *et al.*, 1991). *c-ets-1* est aussi transcrit dans les cellules de crêtes neurales céphaliques et troncales en migration (Vandenbunder *et al.*, 1989 ; Queva *et al.*, 1993).

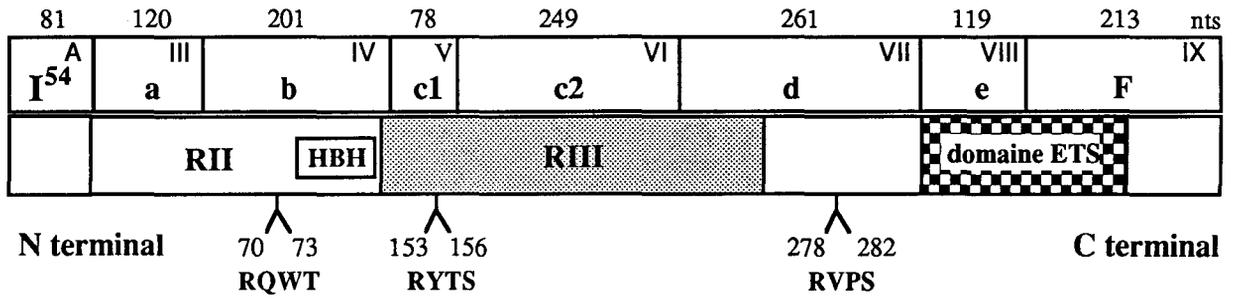
## **D - Régulation de la fonction des produits du gène *c-ets-1***

### **1 - Epissage alternatif**

L'épissage alternatif des messagers du locus *c-ets-1* a été mis en évidence chez l'homme, le poulet et la souris (Leprince *et al.*, 1988 ; Reddy et

A

p51 *c-ets-1*



RIII, domaine de transactivation,  
 RII, domaine de régulation,  
 RQWT, RYTS et RVPS, séquences consensus de phosphorylation,  
 par la calcium/calmoduline kinase II,  
 HBH, région en hélice-boucle-hélice.

B

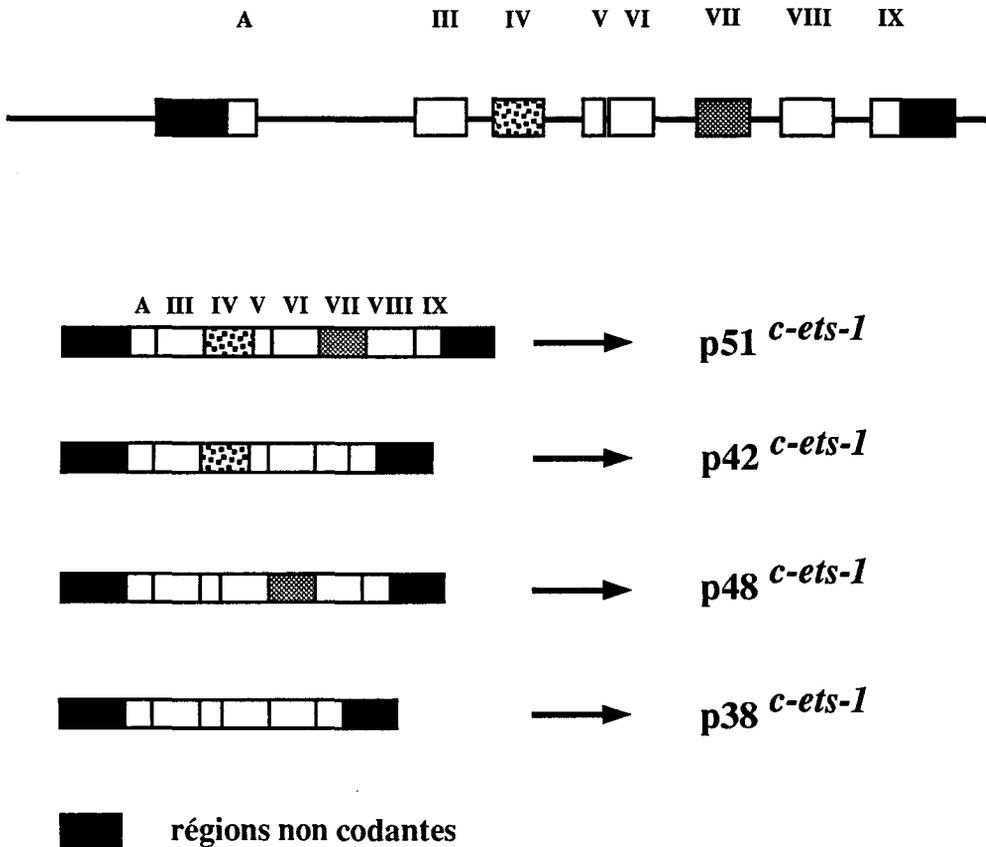


Figure 6 : A - Structure et domaines fonctionnels de la p51 *c-ets-1* humaine.  
 B - Les transcrits du locus *c-ets-1* humain et leur produit.

Rao, 1988 ; Koizumi *et al.*, 1990 ; Jorcyk *et al.*, 1991 ; Crépieux *et al.*, 1993).

Différents transcrits alternatifs ont été identifiés, chez l'homme, dans des lignées de lymphocytes (Figure 6). Deux isoformes des protéines Ets-1 sont traduites, la p51 et la p42<sup>c-ets-1</sup>. L'épissage alternatif éliminant l'exon VII donne naissance au transcrit codant la p42<sup>c-ets-1</sup> (Reddy et Rao, 1988 ; Koizumi *et al.*, 1990 ; Jorcyk *et al.*, 1991). Le messager contenant l'exon VII est cependant le plus abondant dans la lignée cellulaire testée. Il code la p51<sup>c-ets-1</sup>. L'élimination de l'exon VII entraîne la perte du site majeur de phosphorylation. Une protéine variante équivalente à la protéine humaine p42 a été isolée chez la souris (p52<sup>c-ets-1</sup>) (Pognonec *et al.*, 1990). Cette protéine n'est pas affectée par la phosphorylation dans sa capacité de se lier à l'ADN. La protéine p42<sup>c-ets-1</sup> humaine doit vraisemblablement posséder les mêmes propriétés. D'autre part, le domaine codé par l'exon VII est impliqué dans l'inhibition de la fixation à l'ADN et donc la protéine totale se fixe moins bien à l'ADN que la protéine amputée de ce domaine (Wasylyk *et al.*, 1992).

D'autres transcrits, plus rares, ont été détectés par RT PCR dans des lymphocytes sanguins (Jorcyk *et al.*, 1991) (Figure 6). Un épissage alternatif de l'exon IV a été mis en évidence en plus de l'épissage de l'exon VII, ainsi qu'une combinaison de ces deux types d'épissage alternatif, mais aucune fonction n'a pu être déterminée. L'élimination de l'exon IV entraîne la formation d'une protéine de poids moléculaire 48 kDa (p48<sup>c-ets-1</sup>) qui est dépourvue d'une partie de la région codant le domaine de régulation de la transactivation correspondant chez le poulet. Ce domaine possède une structure supposée en hélice-boucle-hélice, structure généralement impliquée dans des interactions protéines-protéines.

Récemment, dans certaines lignées cellulaires, ont été décrits des ARNm du locus *ets-1* résultant d'épissages anormaux : l'ordre normal des exons dans le transcrit est modifié mais le cadre de lecture est conservé (Cocquerelle *et al.*, 1992). Dans ces ARN, les exons VI ou VII peuvent être placés avant l'exon III. Ces ARN ne sont pas polyadénylés et sont circulaires (Cocquerelle *et al.*, 1993). La fonction de ces ARN est encore inconnue.

Chez le poulet, l'épissage alternatif des transcrits du locus *c-ets-1* génère deux protéines, la p54<sup>c-ets-1</sup> et la p68<sup>c-ets-1</sup> (Leprince *et al.*, 1988) (Figure 4). La p68<sup>c-ets-1</sup> diffère de la p54<sup>c-ets-1</sup> par la présence d'un domaine transactivateur supplémentaire localisé dans la partie N-terminale de la p68<sup>c-ets-1</sup>. Les différents épissages alternatifs identifiés chez l'homme n'ont pas d'équivalent chez le poulet. Dans les études *in vitro* menées jusqu'alors, l'efficacité de la transactivation n'apparaît guère différente excepté sur le promoteur de la stromélysine de rat. En cotransfection dans des cellules humaines, la p68<sup>c-ets-1</sup> a

une activité transactivatrice supérieure à celle de la p54c-*ets-1* (Wasylyk *et al.*, 1991). De ce fait, l'existence d'un domaine transactivateur supplémentaire doit permettre à la p68c-*ets-1* d'interagir avec d'autres facteurs et donc d'avoir une activité différente.

## 2 - Maturation post-traductionnelle des protéines Ets-1

Les seules modifications post-traductionnelles des protéines Ets-1 mises en évidence jusqu'à présent sont la phosphorylation et la modification des groupements thiols. Les protéines Ets-1 humaines, murines et aviaires migrent en SDS-PAGE sous différentes formes. Cette hétérogénéité est en grande partie due à des événements de phosphorylation ayant lieu principalement sur des résidus sérine et à un niveau plus faible sur des résidus thréonine. Les protéines phosphorylées ont ainsi une structure tridimensionnelle qui ralentit leur migration (Pognonec *et al.*, 1988 ; Koizumi *et al.*, 1990).

Chez le poulet, après marquage métabolique au  $^{32}\text{P}$ -orthophosphate, trois protéines sont immunoprécipitées et ont des poids moléculaires apparents de 62, 66 et 72 kDa. Ces protéines sont le résultat de la phosphorylation de la p54c-*ets-1*, suggérant l'existence de plusieurs sites de phosphorylation (Pognonec *et al.*, 1988).

Dans les thymocytes, la proportion des formes phosphorylées est faible, indiquant que ces modifications post-traductionnelles auraient une fonction de régulation. La phosphorylation de la p54c-*ets-1* ne dépend ni de l'activation des protéines kinases C, ni de l'activation des protéines kinases AMP cyclique dépendantes. Par contre, le traitement des thymocytes avec l'ionophore calcique A23178 entraîne une augmentation considérable et rapide des formes phosphorylées et la disparition de la p54c-*ets-1*. Des résultats similaires sont obtenus avec les protéines Ets-1 humaines et murines dont les formes phosphorylées ont des poids moléculaires de 60 et 73 kDa respectivement. L'effet de stimulation de l'ionophore A23178 est complètement inhibé par un prétraitement avec du cobalt ou de l'EGTA qui chélatent le calcium. Les événements de phosphorylation  $\text{Ca}^{2+}$  dépendants requièrent donc l'entrée de  $\text{Ca}^{2+}$  extracellulaire, le relargage du  $\text{Ca}^{2+}$  intracellulaire étant insuffisant à lui seul.

La stimulation mitogène des thymocytes par les lectines comme la Concanavaleine A ou par des anticorps monoclonaux dirigés contre des composants du récepteur des cellules T entraîne une phosphorylation de Ets-1 dépendante de la concentration intracellulaire de  $\text{Ca}^{2+}$ , de la même façon qu'avec l'ionophore calcique (Pognonec *et al.*, 1988, 1990).

Les différentes isoformes des protéines Ets-1 humaines ont été

identifiées par marquage au  $^{32}\text{P}$ -orthophosphate dans la lignée cellulaire CEM (cellules T) et aussi dans des lignées de cellules lymphoïdes. Les protéines p52 et p49 sont les versions phosphorylées des p51 et p48 respectivement (Koizumi *et al.*, 1990). L'analyse de la séquence en acides aminés de la protéine p51c-*ets-1* humaine a révélé la présence de sites consensus de la calcium/calmoduline kinase II dans l'exon IV (acides aminés 70-73), dans l'exon V (acides aminés 153-156) et dans l'exon VII (acides aminés 279-282) (Figure 6). Le site situé dans l'exon VII est conservé dans la protéine Ets-2. D'autre part, la p42c-*ets-1* qui résulte de l'épissage alternatif de cet exon, n'est pas phosphorylée après stimulation mitogénique. Le site consensus de l'exon VII semble donc être le site principal de phosphorylation des protéines Ets-1 humaines provoquée par l'augmentation de calcium intracellulaire (Koizumi *et al.*, 1990). Deux sites consensus de fixation de la kinase II calcium/calmoduline dépendante existent aussi dans la p54c-*ets-1*, dont un est situé dans le domaine de répression de la fixation à l'ADN (exon VII - acides aminés 308-311). Les formes non phosphorylées ont une localisation nucléaire et peuvent se lier à l'ADN (Pognonec *et al.*, 1989 ; Koizumi *et al.*, 1990) contrairement aux formes phosphorylées qui perdent leur capacité de liaison à l'ADN et leur localisation nucléaire (Pognonec *et al.*, 1989). La phosphorylation des protéines Ets-1 est transitoire et intervient après stimulation lymphocytaire qui entraînerait l'inactivation des protéines et leur détachement des régions promotrices des gènes qu'elles régulent. Récemment, une forme hyperphosphorylée de la protéine Ets-1 a été mise en évidence au cours de la mitose dans des lignées de lymphocytes T (Fleischman *et al.*, 1993). Cette forme est présente au début du cycle mitotique, avant ou aux alentours de la transition métaphase-anaphase. La forme Ets-1 hyperphosphorylée serait due à de multiples événements de phosphorylation sur des résidus sérine situés au niveau du domaine codé par l'exon VII par activation de la calcium/calmoduline kinase II, et sur un autre site proche du premier. L'activité de liaison spécifique à l'ADN de cette forme hyperphosphorylée ne semble pas altérée. Les changements dans l'activité des phosphatases cellulaires sont importants pour la régulation de Ets-1 en fonction du cycle cellulaire et donc pour la régulation de ses gènes cibles.

Les protéines Ets-1 humaines peuvent subir des modifications au niveau de groupements thiols résultant de l'activité d'inhibiteurs de protéases (Fischer *et al.*, 1992). Ces modifications peuvent avoir lieu sur les résidus cystéine 99, 112 et 169 situés en dehors du domaine de liaison à l'ADN et n'affectent pas la capacité de liaison à l'ADN.

### 3 - Localisation nucléaire

Les protéines p54<sup>c-ets-1</sup> et p68<sup>c-ets-1</sup> exprimées dans des cellules COS à partir d'un vecteur d'expression sont localisées dans le noyau (Boulukos *et al.*, 1989 ; Pognonec *et al.*, 1989 ; Leprince *et al.*, 1990). Le domaine responsable de cette localisation est une séquence riche en acides aminés basiques (376 - GKRKNKPK pour la p54<sup>c-ets-1</sup>) ; cependant, les acides aminés situés en amont par rapport au signal de localisation nucléaire sont aussi impliqués dans cette activité. L'intégrité des 78 acides aminés carboxy-terminaux est requise pour la localisation nucléaire de la p54<sup>c-ets-1</sup> (Boulukos *et al.*, 1989).

La séquence basique GKRKNKPK est similaire à la séquence minimale conférant la localisation nucléaire de l'antigène T du virus SV 40 ainsi que plusieurs autres protéines nucléaires (Richardson *et al.*, 1986 ; Dang et Lee, 1988). L'analyse informatique de cette région prédit qu'elle ferait partie d'une boucle s'étendant à la surface de la protéine, permettant ainsi l'interaction avec un transporteur la guidant jusqu'au noyau (Pognonec *et al.*, 1989).

### 4 - Stabilité des protéines Ets-1

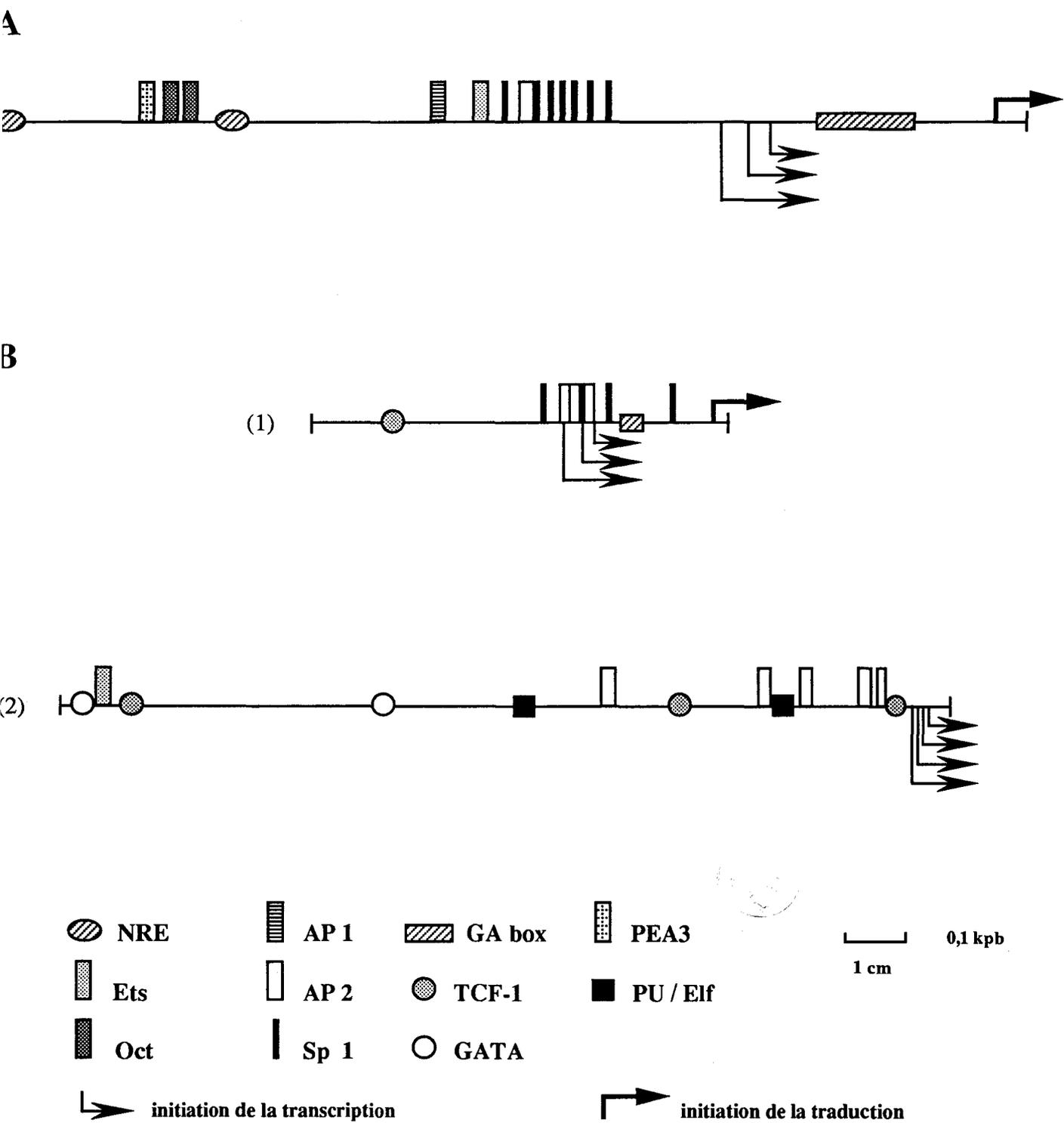
La demi-vie des protéines Ets-1 est relativement courte, de 45 à 85 minutes (Pognonec *et al.*, 1988 ; Koizumi *et al.*, 1990). La stabilité est augmentée pour les protéines Ets-1 résultant de l'épissage alternatif de l'exon VII qui présente le site majeur de phosphorylation. Pour ces protéines (p48 et p39<sup>c-ets-1</sup> humaines), la demi-vie est de 135 minutes (Koizumi *et al.*, 1990). Les séquences PEST riches en résidus proline, glutamine, sérine et thréonine semblent jouer un rôle dans le renouvellement des protéines Ets (Rogers *et al.*, 1986). Des séquences de ce type sont retrouvées entre la région RIII et le domaine ETS. Elles seraient responsables de la protéolyse de la protéine, générant ainsi deux fragments pouvant interférer avec l'activité des protéines totales (pour revue, MacLeod *et al.*, 1992). En effet, le clivage crée ainsi un domaine de liaison à l'ADN exempt des restrictions de régulation imposées par les domaines N-terminaux, toujours capable de reconnaître des EBS.

## E - Régulation de c-ets-1

### 1 - Caractérisation du promoteur

Les séquences promotrices situées en amont de l'exon I<sup>54</sup> dans le locus *c-ets-1* ont été clonées chez l'homme et le poulet et commencent à être bien caractérisées sur le plan fonctionnel (Figure 7).

La transcription du gène humain *ets-1* s'initie en de multiples



**Figure 7 : Promoteurs des gènes *c-ets-1* humain et aviaires**

**A - promoteur humain**

**B - promoteur aviaire dirigeant la transcription des ARNm codant la p54<sup>c-ets-1</sup> (1) et la p68<sup>c-ets-1</sup> (2)**

sites de démarrage. Il existe 14 sites d'initiation couvrant une région de 137 pb, le site situé le plus en 3' étant localisé 305 pb en amont du codon ATG d'initiation de la traduction (Jorcyk *et al.*, 1991). Le promoteur du gène *ets-1* humain ne contient pas les éléments TATA et CAAT classiquement reconnus par l'ARN polymérase II pour la plupart des gènes eucaryotes. Par contre, il contient un grand pourcentage en bases G et C.

L'absence d'éléments TATA-CAAT, le pourcentage élevé de GC ainsi que les multiples sites de démarrage de la transcription sont retrouvés dans des promoteurs d'autres gènes humains comme les "house keeping genes", des gènes en relation avec la croissance, et des oncogènes tel Hu *c-ets-2* (Oka *et al.*, 1991 ; Jorcyk *et al.*, 1991).

Une région de 706 pb incluant les sites multiples d'initiation de la transcription contient l'activité promotrice maximale de l'*ets-1* humain (Jorcyk *et al.*, 1991) (Figure 7A). Cette région présente des sites putatifs de fixation pour les facteurs de transcription Sp-1, AP-1, AP-2 et Ets, ainsi qu'un élément similaire à l'élément de réponse au sérum de *c-fos*, qui s'est ensuite avéré non fonctionnel (Majérus *et al.*, 1992). Le promoteur *ets-1* humain est aussi caractérisé par de longues séquences polypuriques-polypyrimidiques (G/A). Ces séquences semblent importantes pour l'activité promotrice car elles peuvent former une structure en triple hélice de l'ADN, sensible à la nucléase S<sub>1</sub>, sur laquelle peuvent se fixer des protéines nucléaires qui pourraient ainsi réguler l'inductibilité du promoteur (Mavrothalassitis *et al.*, 1990 ; Jorcyk *et al.*, 1991). Ces structures homopolymériques pourraient servir d'éléments fonctionnels du promoteur.

Un certain nombre de tests ont ensuite été effectués pour vérifier si ces sites étaient fonctionnels.

## 2 - Régulation de l'expression

Des expériences de protection contre la nucléase S<sub>1</sub> ont permis de montrer que trois des sites consensus Sp-1 localisés près du site "cap" sont protégés de la digestion par des extraits nucléaires de cellules épithéliales de souris. Une telle protection est aussi observée dans une zone où l'on trouve un motif PEA3 chevauchant des motifs AP-1 ressemblant à un ORU (unité de réponse à l'oncogène). L'expression de *c-ets-1* est induite par le sérum, mais lentement. L'activité du promoteur est augmentée par le sérum (Majérus *et al.*, 1992). De même, des tests de cotransfection de *c-jun* + *c-fos* et de *c-ets-1* montrent une stimulation de l'activité promotrice, impliquant une région contenant des motifs putatifs AP-1 et PEA3. Ces expériences tendent à prouver que la transactivation du promoteur par c-Fos + c-Jun implique un site consensus pour AP-1. De façon

contradictoire, c-Fos + c-Jun ne coopèrent pas dans la transactivation par un ORU de *ets-1* contrairement au PyORU de l'"enhancer" du virus du polyome (Oka *et al.*, 1991 ; Majérus *et al.*, 1992). Les tests de cotransfection ont été réalisés dans différentes lignées cellulaires : lignée cellulaire lymphoïde Daudi, lignée cellulaire érythroïde K562 et lignée cellulaire promyélocytaire HL60 (Oka *et al.*, 1991) ainsi que dans des cellules épithéliales de souris (Majérus *et al.*, 1992). Le promoteur d'*ets-1* fonctionne bien dans les cellules lymphoïdes, les cellules épithéliales mais aussi de façon étonnante dans les cellules érythromyéloïdes K562, car les produits de transcription et de traduction du gène *ets-1* n'ont pas été détectés dans cette lignée cellulaire (Chen, 1985 ; Bhat *et al.*, 1987). Il existerait dans ces cellules un facteur fort, spécifique et de régulation négative, qui se fixerait dans une région du promoteur antérieure à celle utilisée pour les tests de cotransfection (Oka *et al.*, 1991).

En plus de la présence de sites pour la fixation des facteurs de transcription AP-1, AP-2 et Sp-1, le promoteur d'*ets-1* contient un site de liaison pour les protéines Ets-1, qui s'est avéré fonctionnel après un test de transfection dans des cellules Daudi (Majérus *et al.*, 1992). Ces résultats suggèrent que la transcription du gène *ets-1* peut être régulée de façon positive par son propre produit, comme cela a aussi été rapporté par Seth et Papas (1990). La présence d'un facteur de régulation négative en amont de la séquence promotrice étudiée est envisagée pour contrebalancer l'action de la protéine Ets-1 sur son propre promoteur (Oka *et al.*, 1991).

Ces éléments de régulation négative préalablement envisagés viennent d'être identifiés (Chen *et al.*, 1993a). Par une série de délétions, deux éléments de régulation négative ont pu être localisés dans la région promotrice du gène *c-ets-1* humain : NRE1 et NRE2. La présence de ces deux éléments est nécessaire pour exercer un effet répressif sur le promoteur *ets-1*. D'autre part, les auteurs ont également montré que ces NRE sont aussi fonctionnels sur un promoteur hétérologue (Chen *et al.*, 1993a). Des extraits nucléaires de cellules lymphoïdes B et T (Daudi et HPB) et érythromyéloïdes (K562) incubées avec les NRE1 et NRE2 forment des complexes, alors qu'aucun complexe n'est formé dans les cellules K562. Il existe donc des facteurs de transcription capables de se lier sur ces éléments, ces facteurs n'étant présents que dans les cellules où le promoteur *ets-1* humain est actif (Chen *et al.*, 1993a). L'étude plus approfondie du promoteur a aussi permis de mettre en évidence des motifs de liaison pour les facteurs de transcription PEA3 et Oct (un site pour PEA3, deux sites pour Oct) situés entre les NRE. Ces sites représenteraient d'autres éléments de régulation positive du promoteur *ets-1*, ces éléments n'étant pas aussi forts que AP-1, AP-2

et Ets.

Des travaux très récents ont permis de mettre en évidence l'activité de régulation des facteurs de transcription PEA3, Oct-1 et Oct-2 sur le promoteur du gène *ets-1* humain (Chen *et al.*, 1993b). Ces trois facteurs régulent de façon positive le promoteur du gène *ets-1*. Le facteur PEA3 est le stimulateur le plus fort, et il a plus d'affinité pour les motifs PEA3 et Ets que la protéine Ets-1, peut-être parce que la protéine PEA3 ne possède pas de domaine inhibiteur de la fixation à l'ADN (Xin *et al.*, 1992). AP-2 est le régulateur le plus fort du promoteur *ets-1* après PEA3. Son activité est aussi plus importante que celle de la protéine Ets-1. La présence de facteurs se liant au NRE2 empêche les facteurs Oct-1 et Oct-2 d'exercer leur activité de régulation positive. Les facteurs se liant au NRE2 et le facteur Oct-2 interagiraient de façon allostérique. Les auteurs montrent que la combinaison des facteurs AP-2, Ets-1, Oct-1, Oct-2 et PEA3 augmente l'expression du promoteur *ets-1* de façon additive plutôt que synergique.

En résumé, AP-1, AP-2 et Ets sont des facteurs de transcription intervenant dans la régulation positive du gène *ets-1*. Les sites de liaison de ces facteurs sont localisés dans une région du promoteur. Dans une autre région se trouvent des sites de liaison pour les facteurs PEA3, Oct-1 et Oct-2 qui régulent aussi de façon positive. Pour maintenir une expression normale, deux régions contenant des NRE peuvent fixer des facteurs de régulation négative empêchant ainsi une surexpression du gène.

Chez le poulet, deux régions promotrices ont été caractérisées (Crépieux *et al.*, 1993) (Figure 7B) ; la première se trouve en amont de l'exon I<sup>54</sup> et permet la synthèse des ARNm codant la p54<sup>c-ets-1</sup>. Cette région promotrice est très homologue à celle caractérisée chez l'homme. Le deuxième promoteur qui fonctionne dans des fibroblastes embryonnaires est localisé en 5' d'un exon Y non codant situé en amont des exons  $\alpha$  et  $\beta$ . Il initie la synthèse des messagers codant la p68<sup>c-ets-1</sup> pour laquelle le domaine codé par l'exon I<sup>54</sup> est éliminé.

Tout comme pour les promoteurs caractérisés chez l'homme, les deux promoteurs aviaires ne présentent pas d'élément TATA ou CAAT et la transcription des ARNm codant les protéines p54<sup>c-ets-1</sup> et p68<sup>c-ets-1</sup> s'initie en de multiples sites.

Le promoteur initiant la synthèse de la p68<sup>c-ets-1</sup> n'est pas spécialement riche en bases G et C, même s'il ne comporte pas de boîte TATA ou CAAT, mais la transcription est initiée dans une zone contenant une dizaine de bases C ou T. Des sites potentiels de fixation de différents facteurs de transcription spécifiques des cellules lymphoïdes ont été identifiés tels que Elf-1, PU.1, Ets-1,

TCF-1 et GATA-3 en plus de sites AP-2. L'ARNm codant la p54<sup>c-ets-1</sup> s'initie, lui dans une région riche en G et C. L'organisation de ce promoteur est similaire à celle décrite pour le promoteur humain, comprenant plusieurs sites AP-2 et Sp-1 ainsi qu'une longue répétition de G/A située en aval des sites d'initiation de la transcription.

## **F - Les protéines Ets-1 et la transformation**

### **1 - Transformation de cellules *in vitro***

La transfection de fibroblastes de souris NIH3T3 par un vecteur plasmidique exprimant la p54<sup>c-ets-1</sup> (ou la p55<sup>c-ets-2</sup>, Seth *et al.*, 1989) entraîne une modification de leurs caractéristiques (Seth et Papas, 1990). Les lignées cellulaires dérivées des foyers de cellules transformées peuvent pousser sans ancrage en milieu appauvri en sérum et former des colonies en milieu semi-solide et sont tumorigènes chez la souris athymique. Toutefois, le pouvoir transformant de *c-ets-1* est faible et la capacité à induire des foyers est beaucoup plus importante si les cellules NIH3T3 sont transfectées par le virus murin ME26 qui exprime la P135<sup>gag-myb-ets</sup> (Yuan *et al.*, 1989). Les cellules infectées par le virus ME26 n'induisent cependant pas de tumeurs chez la souris athymique ; les cellules transformées par *c-ets-1* ont donc subi des modifications qui ont augmenté leur pouvoir oncogénique.

### **2 - Translocation chromosomique**

Le locus *c-ets-1* a été impliqué dans des translocations chromosomiques entre le chromosome 11 et les chromosomes 1, 4 et 9. Ces translocations ont été mises en évidence chez des malades atteints de leucémies myéloïdes, mais le point de cassure dans le locus *c-ets-1* n'a pas été localisé (Diaz *et al.*, 1986 ; Sacchi *et al.*, 1986). Un réarrangement de ce locus a été montré dans un cas de leucémie myélomonocytaire et dans un cas de lymphome (Rovigatti *et al.*, 1986). Le rôle de *c-ets-1* dans l'apparition ou la maintenance de ces leucémies reste encore à démontrer.

### **3 - Ets-1 dans les tumeurs humaines**

Par la technique d'hybridation *in situ*, l'expression du gène *ets-1* a été mise en évidence dans des cellules de tumeurs humaines d'origine neuroectodermique du système nerveux périphérique (Sacchi *et al.*, 1991). De la même façon, l'expression du gène *ets-1* a été démontrée dans des sarcomes d'Ewing. Cependant, rien n'indique ici dans quelle mesure le gène *ets-1* peut être

Espèce	Gène	Produit (en acides aminés)	Références
Homme	<i>ets-1</i>	441	Watson <i>et al.</i> (1988b)
		354	Reddy et Rao (1988)
	<i>ets-2</i>	469	Watson <i>et al.</i> (1988b)
	<i>erg</i>	462	Reddy <i>et al.</i> (1987)
		363	Rao <i>et al.</i> (1987)
	<i>fli-1/erg B</i>	452	Prasad <i>et al.</i> (1992)
			Watson <i>et al.</i> (1992)
	<i>elk-1</i>	428	Rao <i>et al.</i> (1989)
	<i>SAP-1</i>	453	Dalton et Treisman (1992)
	<i>elf</i>	619	Thompson <i>et al.</i> (1992)
<i>PEA3</i>	476	Higashino <i>et al.</i> (1993)	
<i>ERM</i>	510	Monté <i>et al.</i> (soumis)	
Souris	<i>ets-1</i>	440	Chen (1990)
	<i>ets-2</i>	468	Watson <i>et al.</i> (1988b)
	<i>spi-1/PU.1</i>	272	Moreau-Gachelin <i>et al.</i> (1988)
			Klemsz <i>et al.</i> (1990)
			Moreau-Gachelin <i>et al.</i> (1989)
	<i>fli-1</i>	452	Ben-David <i>et al.</i> (1991)
	<i>GABP<math>\alpha</math></i>	454	La Marco <i>et al.</i> (1991)
	<i>ER71</i>	335	Brown et McKnight (1992)
	<i>ER81</i>	477	Brown et McKnight (1992)
	<i>PEA3</i>	555	Xin <i>et al.</i> (1992)
Poulet	<i>ets-1</i>	441	Watson <i>et al.</i> (1988a)
		485	Duterque-Coquillaud <i>et al.</i> (1988)
	<i>ets-2</i>	479	Leprince <i>et al.</i> (1988) Boulukos <i>et al.</i> (1988)
Xénope	<i>ets-1a</i>	438	Stiegler <i>et al.</i> (1990)
	<i>ets-1b</i>	> 268	Stiegler <i>et al.</i> (1990)
	<i>ets-2a</i>	472	Wolff <i>et al.</i> (1991)
	<i>ets-2b</i>	472	Burdett <i>et al.</i> (1992)
Oursin	<i>ets</i>	ND	Chen <i>et al.</i> (1988)
	<i>erg</i>	ND	Qi <i>et al.</i> (soumis)
Drosophile	<i>ets-2</i>	ND	Pribyl <i>et al.</i> (1988)
		135	Pribyl <i>et al.</i> (1991)
		883	Janknecht <i>et al.</i> (1989) Burtis <i>et al.</i> (1990)
	<i>ets-3</i>	ND	Chen <i>et al.</i> (1992)
	<i>ets-4</i>	ND	
	<i>ets-6</i>	ND	
	<i>pokkuri/yan</i>	761	Tei <i>et al.</i> (1992)
			Lai et Rubin (1992)

Tableau 3 : Les gènes de la famille ETS.

impliqué dans l'apparition ou le développement de ces tumeurs.

Les transcrits du gène *c-ets-1* ont été détectés dans l'endothélium des vaisseaux sanguins néoformés qui irriguent la tumeur (carcinomes du sein et du poumon) et dans les fibroblastes constituant le stroma de tumeurs humaines, mais il n'ont pas été détectés dans les cellules cancéreuses (Wernert *et al.*, 1992).

Ici, comme au cours de l'embryogenèse aviaire, *c-ets-1* est exprimé dans les vaisseaux sanguins en voie de formation (Vandenbunder *et al.*, 1989).

### **III - La famille ETS**

La famille ETS est une famille de facteurs de transcription constituée de nombreux membres présentant entre eux une ou plusieurs régions d'homologie. L'ensemble des gènes constituant cette famille est décrit dans le tableau 3.

#### **A - Critères de classification des membres de la famille ETS**

##### **1 - Les domaines A, B, C**

Selon le degré d'homologie avec la protéine Ets-1 humaine, on distingue trois régions, A, B et C (Watson *et al.*, 1988b) (Figure 8) :

— La région A ne se rencontre que dans quelques membres de la famille ETS, où elle présente une homologie moyenne (66 % d'homologie entre la protéine Ets-2 de xénope et la protéine Ets-1 humaine).

— La région B est celle qui présente le moins d'homologie pour les membres de la famille.

— La région C est la seule région très bien conservée dans tous les membres de la famille ETS : elle présente plus de 90 % d'homologie entre la protéine Ets-2 de drosophile et la protéine Ets-1 humaine. C'est cette région C qui détermine l'appartenance à la famille ETS.

Ces trois régions sont des domaines fonctionnels :

— La région A comprend 60 acides aminés, présentant une certaine homologie avec le domaine HLH des protéines HLH ("helix loop helix") (Anthony-Cahill *et al.*, 1992 ; Vison et Garcia, 1992).

Le domaine HLH est une région caractéristique de nombreux facteurs de transcription, exigée pour les interactions protéine-protéine.

— La région B, spécifique des protéines Ets-1, comprend un



domaine transactivateur (Gégonne *et al.*, 1992), ainsi qu'un domaine de répression de la fixation à l'ADN (Lim *et al.*, 1992 ; Nye *et al.*, 1992 ; Wasylyk *et al.*, 1992).

— La région C est constituée de 110 acides aminés. C'est une région de liaison à l'ADN riche en résidus basiques (arginine et lysine). Ce domaine est différent de ceux impliqués jusqu'ici dans les interactions protéines-acides nucléiques, tels que les motifs "leucine zipper" (Landschultz *et al.*, 1988), doigts de zinc (Evans et Hollenberg, 1988), hélice-boucle-hélice (Murre *et al.*, 1989) et homéodomaine (Levine et Hoey, 1988). Un trait caractéristique du domaine C est la présence d'un triplet de résidus tryptophane, les résidus étant séparés par 17 à 18 acides aminés. Une exception cependant, dans la protéine PU.1, le troisième tryptophane est remplacé par une tyrosine.

## 2 - Le domaine ETS

Dans le domaine C une séquence de 85 acides aminés est particulièrement bien conservée des Invertébrés à l'homme. L'alignement de 27 de ces domaines (Laudet *et al.*, 1993) (Figure 9) montre une forte homologie en deux endroits de ces séquences : 14 acides aminés situés en position N-terminale (structure prédite en hélice  $\alpha$ ), et environ 50 acides aminés situés en position C-terminale (le domaine basique). La région charnière de 23 acides aminés connectant ces deux régions est beaucoup moins bien conservée parmi les membres de la famille ETS. Ainsi, en prenant les membres les plus éloignés de cette famille, Ets-1 et Fli-1, on observe 57 % d'homologie dans l'hélice  $\alpha$  et 35 % d'homologie dans le domaine basique, mais seulement 14 % d'homologie dans la région charnière. Dans cette région charnière, seulement deux acides aminés, un tryptophane et une phénylalanine, sont communs à tous les membres de la famille ETS ; pourtant, Wang *et al.* (1992) ont montré que le remplacement de ce tryptophane par un acide aminé hydrophobe ne modifie pas la capacité de fixation à l'ADN de la protéine c-Ets-1.

Le domaine ETS peut se situer en position carboxy-terminale (protéines Ets-1, Ets-2, Erg, Fli-1, GABP $\alpha$ , Spi-1/PU.1, PEA3 et E74), amino-terminale (protéines Elk-1, SAP-1), ou même médiane (protéines Elf-1 et Pok) de la protéine (Figure 10).

L'analyse des séquences des protéines Ets montrent qu'elles renferment un triplet de résidus tryptophane, espacés de 18 ou 19 acides aminés. Pour certains auteurs, le domaine ETS est à rapprocher de celui des domaines de liaison à l'ADN des protéines Myb (Anton et Frampton, 1988 ; Karim *et al.*, 1990). Cependant, dans les protéines Myb, ce motif de trois résidus tryptophane

Hélice  $\alpha$

Charnière

Domaine basique

	10	20	30	40	50	60	70	80	90	100	
HETS1	GSG---PI	QLWQFLLELLTDKS	--CQSFISWTG-DGWEFKLS	--DPDEVA	RRWGKRR-NKPKMNYEKL	SRGLRYYYDKNI	IHKTA-GKRYVYRFVCDL	QSLLG	93		
CETS1	GSG---PI	QLWQFLLELLTDKS	--CQSFISWTG-DGWEFKLS	--DPDEVA	RRWGKRR-NKPKMNYEKL	SRGLRYYYDKNI	IHKTA-GKRYVYRFVCDL	QSLLG	93		
VETS	GSG---PI	QLWQFLLELLTDKS	--CQSFISWTG-DGWEFKLS	--DPDEVA	RRWGKRR-NKPKMNYEKL	SRGLRYYYDKNI	IHKTA-GKRYVYRFVCDL	QSLLG	93		
XETS1A	GSG---PI	QLWQFLLELLTDKS	--CQSFISWTG-DGWEFKLS	--DPDEVA	RRWGKRR-NKPKMNYEKL	SRGLRYYYDKNI	IHKTA-GKRYVYRFVCDL	QSLLG	93		
XETS1B	GSG---PI	QLWQFLLELLTDKS	--CQSFISWTG-DGWEFKLS	--DPDEVA	RRWGKRR-NKPKMNYEKL	SRGLRYYYDKNI	IHKTA-GKRYVYRFVCDL	QSLLG	93		
HETS2	GSG---PI	QLWQFLLELLSDKS	--CQSFISWTG-DGWEFKLA	--DPDEVA	RRWGKRR-NKPKMNYEKL	SRGLRYYYDKNI	IHKTS-GKRYVYRFVCDL	QNLLG	93		
CETS2	GSG---PI	QLWQFLLELLTDKS	--CQSFISWTG-DGWEFKLA	--DPDEVA	RRWGRRK-NKPKMNYEKL	SRGLRYYYDKNI	IHKTS-GKRYVYRFVCDL	QNLLG	93		
XETS2A	GSG---PI	QLWQFLLELLTDKS	--CQSFISWTG-DGWEFKLT	--DPDEVA	RRWGKRR-NKPKMNYEKL	SRGLRYYYDKNI	IHKTS-GKRYVYRFVCDL	HNLLG	93		
XETS2B	GSG---PI	QLWQFLLELLTDKS	--CQSFISWTG-DGWEFKLA	--DPDEVA	RRWGKRR-NKPKMNYEKL	SRGLRYYYDKNI	IHKTS-GKRYVYRFVCDL	HNLLG	93		
SUETS2	GSG---PI	QLWQFLLELLTDKT	--CQHIISWTG-DGWEFKLS	--DPDEVA	RRWGKRR-NKPKMNYEKL	SRGLRYYYDKNI	IHKTA-GKRYVYRFVCDL	QSLLG	93		
DETS2	GSG---PI	QLWQFLLELLDKT	--CQSFISWTG-DGWEFKLT	--DPDEVA	RRWGIRK-NKPKMNYEKL	SRGLRYYYDKNI	IHKTA-GKRYVYRFVCDL	QNLVG	93		
ERG	GSG---QI	QLWQFLLELLSDSS	--NSSCITWEG-TNGEFKMT	--DPDEVA	RRWGERK-SKPNMNYDKLS	RALRYYYDKNIMTKVH	-GKRYAYKF--DF	HGIAQ	91		
FLI	GSG---QI	QLWQFLLELLSDSA	--NASCITWEG-INGEFKMT	--DPDEVA	RRWGERK-SKPNMNYDKLS	RALRYYYDKNIMTKVH	-GKRYAYKF--DF	HGIAQ	91		
SUERG	?SG---QI	QLWQFLLELLSDSS	--NASCITWEG-TNGEFKMT	--DPDEVA	RRWGERK-SKPNMNYDKLS	RALRYYYDKNIMTKVH	-GKRYAYKF--DF	AGLAQ	90		
DETS3	GSG---QI	QLWQFLLELLSDSN	--NASCITWEG-TNGEFKLT	--DPDEVA	RRWGERK-SKPNMNYDKLS	RALR????????????????????	????????????????	?????	64		
DETS4	GG---SHI	HLWQFLKELLASPQ	VNGTA-IRWIDRSKGIFKIE	--DSVRVA	KLWGRRK-NRPAMNYDKLS	RSLRYYYDKNIMTKVH	-GKRYAYKF--DF	HGLMA	92		
DETS6	GSG---GQI	QLWQFLLELLADSS	--NANAISWEG-QSGEFRLI	--DPDEVA	RRWGERK-AKPNMNYDKLS	RALRYYYDKNIMTKVH	-GKRYAYKF--DF	HGLMA	92		
GABP	G---NNGQI	QLWQFLLELLTDKD	--ARDCISWVG-DEGEFKLN	--QPELVA	QKWGQRK-NKPTMNYEKL	SRALRYYYDGDMI	CKVQ-GKRFVYKFVCDL	KTLIG	94		
DELG	GSGNNGQV	QLWQFLLEYLTDCE	--HTDVIEWVG-TEGEFKLT	--DPDRVA	RLWGEKK-NKPNMNYEKL	SRALRYYYDGDMI	SKVS-GKRFAYKFDL	KLLIG	96		
POK	NTN---GRL	-LWDFLQQLLNDRN	QKYSDLIAWKCRDTGVFKIV	--DPAGLA	KLWGIQK-NHLSMNYDKMS	RALRYYYRVNI	LRKVQ-GERHCYQFLRNP	TELKN	96		
SAP1	---MDSAI	TLWQFLQLLQKQPQ	--NKHMICWTSMDG-QFKLL	--QAEVA	RLWGIRK-NKPNMNYDKLS	RALRYYYVKN	IKKVN-GQKFVYKFVS-Y	PEILN	92		
ELK	---MDPSV	TLWQFLQLLREQG	--NGHIISWTSRDGGEFKLV	--DAEEVA	RLWGLRK-NKTNMNYDKLS	RALRYYYDKNI	IRKVS-GQKFVYKFVS-Y	PEVAG	93		
PEA3	PYQRRGAL	QLWQFLVALLDDPT	--NAHFIAWTG-RGMEFKLI	--EPEEVA	RLWGIQK-NRPAMNYDKLS	RSLRYYYEKGIMQVA	-GERVYKFVCEP	EALFS	93		
ELF1	K-G---NTI	YLWEFLLALLQDKA	T-CPKYIKWTQREKGFKLV	--DSKAVS	RLWGKHK-NKPDNMYETM	GRALRYYYQGI	LAKVE-GGRLVYQFKEMP	KDLIY	95		
E74	REG---STT	YLWEFLLKLLQDRE	Y-CPRFIKWTNREKGVFKLV	--DSKAVS	RLWGMHK-NKPDNMYETM	GRALRYYYQGI	LAKVD-GQRLVYQFVDVP	KDIEE	96		
SP11	GSKKK--I	RLYQFLDLLRSKD	MK--DSIWWVDKDKGTFQFSKHKKEALA		HRWGIQKGNRKKMTYQK	MARALRNYGKTGEVKKVKK	--KLTYQFSGEV	--LGR	95		
SP12	GTRKK--L	RLYQFLGLLTRGD	MR--ECVWVVEPGAGVQFSKHKKELLA		RRWGQOKGNRKRMTYQK	LARALRNYAKTGEIRKVKR	--KLTYQFDSAL	--LPA	95		
CONS		L**FL *L	* W	F*	*	WG K	M*Y * R *R*Y*	K	* Y F		



\* acides aminés conservés dans trois protéines sur quatre

Figure 9 : Alignement du domaine ETS de 27 membres de la famille ETS, d'après Laudet *et al.*, 1993. La ligne consensus désigne les acides aminés invariants.

est lui-même répété trois fois, constituant trois hélices  $\alpha$  : Myb  $\alpha$ 1, Myb  $\alpha$ 2 et Myb  $\alpha$ 3 (Frampton *et al.*, 1989). Une exception cependant : la répétition des motifs Myb ne se fait que deux fois dans la protéine C1 de la plante *Zea mays* (Paz-Ares *et al.*, 1987) considérée comme un facteur de transcription de la famille MYB.

Par une méthode d'analyse permettant de détecter des relations structurales entre des protéines, même avec une faible homologie de séquence, Laget *et al.* (1993) ont montré que les domaines de liaison à l'ADN des protéines des familles ETS, MYB et HMG ("High Mobility Group") présentent des caractéristiques structurales communes. L'analyse des domaines MYB et HMG décrit la présence d'un motif en hélice-boucle-hélice; les auteurs concluent que le domaine ETS doit présenter le même motif.

### 3 - Classement des protéines Ets

Les 3 régions, A, B et C, permettent de grouper les protéines Ets dans les trois classes I, II et III (Watson *et al.*, 1988b) :

— Dans la classe I, les protéines Ets possèdent les trois régions A, B et C homologues. Cette classe est représentée par les protéines Ets-1 et v-Ets.

— Dans la classe II, les protéines présentent les deux régions A et C homologues. Cette classe est représentée par les protéines Ets-2 de poulet, de souris, d'homme ou de xénope ; Erg et Fli-1 humaines; Fli-1 et GABP- $\alpha$  de souris et Pok et Pointed (D Ets-2) de drosophile.

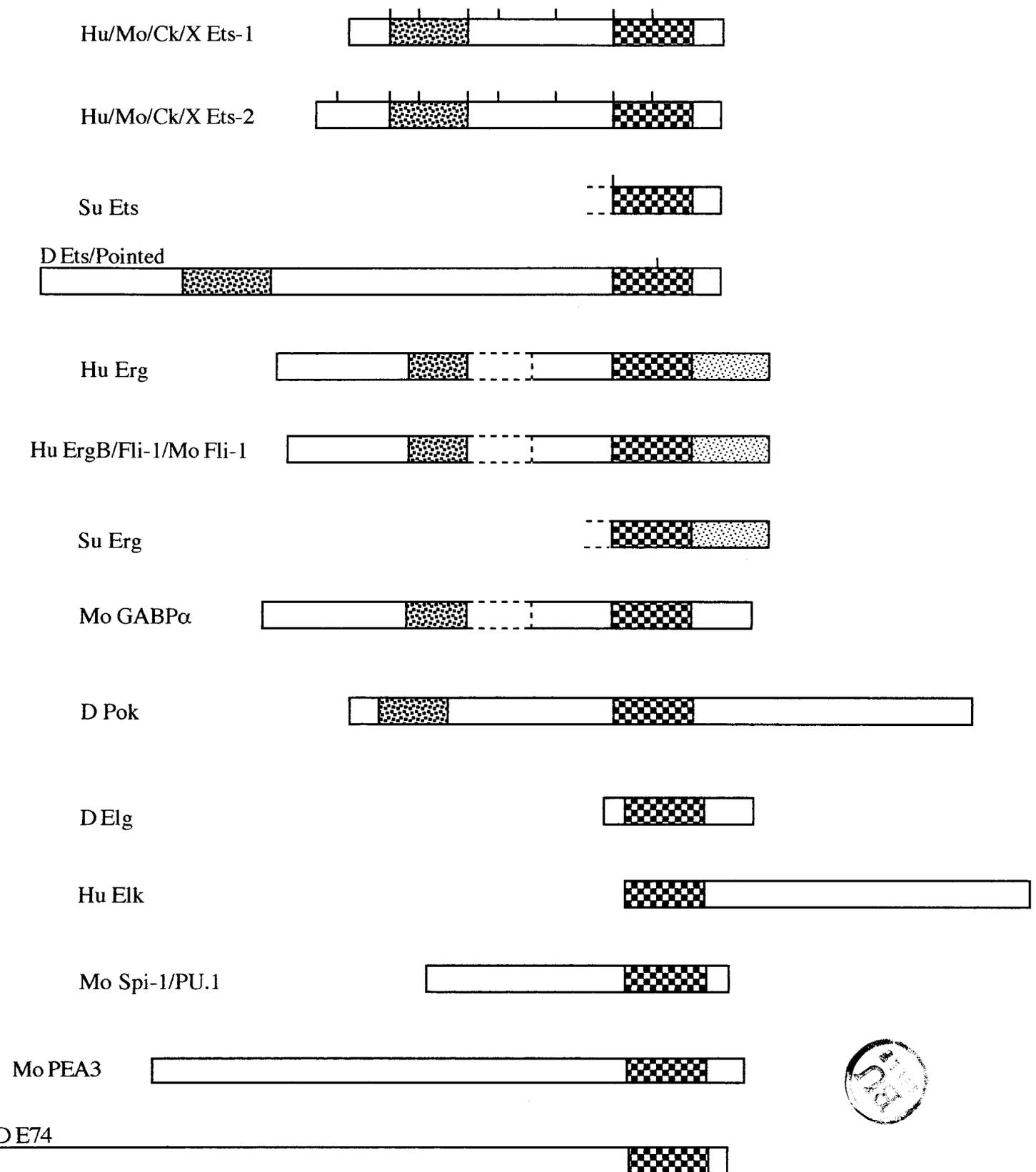
— Dans la classe III, les protéines n'ont pour région homologue que le domaine C. Cette classe est représentée par les protéines Elk humaine, PU.1 et PEA3 de souris et E74A de drosophile.

## B - Modèle d'évolution

### 1 - Classement des domaines ETS

L'analyse des taux de divergence entre les différents domaines ETS (Lautenberger *et al.*, 1992 ; Laudet *et al.*, 1993) indique que la famille ETS constitue un groupe très ancien de gènes, qui peuvent être classés en neuf groupes (Laudet *et al.*, 1993) : ETS, ERG, ELG, PEA3, ELK, ELF, D ETS-4, POK et SPI ; ces neuf groupes peuvent être rassemblés en deux sous-familles : la sous-famille ETS, la plus importante, et la sous-famille SPI (Figure 11).

Dans l'état actuel de nos connaissances, les groupes ERG et surtout ETS sont les plus diversifiés. Le groupe ETS est lui-même subdivisé en sous-groupes ETS-1 et ETS-2. ERG, ETS-1 et ETS-2 sont représentés dans des



 Domaine amino-terminal conservé pour les protéines Ets-1, Ets-2, Erg, Fli, GABP $\alpha$  et Pok.

 Domaine ETS.

 Domaine R spécifique des protéines Erg et Fli.

Les traits verticaux indiquent la position des sites connus d'épissage.

**Figure 10 : Comparaison des structures des protéines de la famille ETS.**

phyla éloignés du point de vue phylogénique : ERG a des représentants chez l'homme, la souris, la drosophile et l'oursin ; ETS-1 est représenté chez l'homme, le poulet, le xénope et l'oursin ; ETS-2 est représenté chez l'homme, le poulet, le xénope et la drosophile.

## 2 - Phylogénie du domaine ETS

Plusieurs gènes *ets* de Vertébrés ont un homologue chez la drosophile :

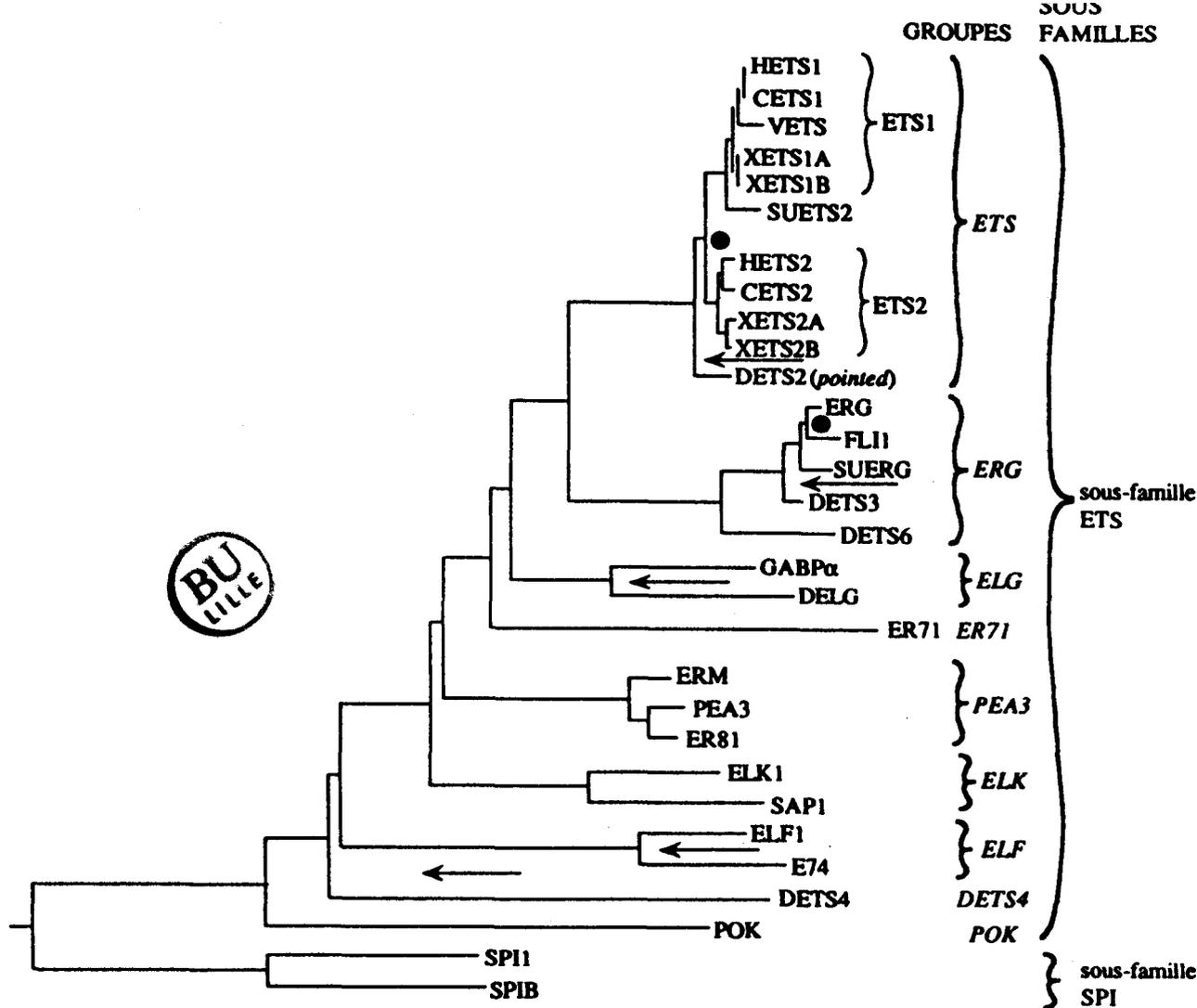
Gènes de Vertébrés	Gènes de drosophile
<i>elf-1</i>	<i>E74</i>
GABP $\alpha$	D <i>elg</i>
<i>erg-1</i>	D <i>ets-3</i>
<i>fli-1</i>	D <i>ets-6</i>
<i>ets-1</i>	D <i>ets-2 /pointed</i>
<i>ets-2</i>	

Dans le phylum des Deutérostomiens, un seul membre du groupe ETS (Su *Ets-2*) et un seul membre du groupe ERG (Su *Erg*) ont jusqu'ici été mis en évidence dans l'embranchement des Echinodermes. Chez les Vertébrés, il existe par contre deux types de gènes du groupe ETS (*ets-1* et *ets-2*), et deux types de gènes du groupe ERG (*erg* et *fli*).

La divergence du gène *ets* se serait produite avant l'apparition des Amphibiens, il y a au moins 400 millions d'années, puisque le xénope possède déjà des exemplaires des gènes *c-ets-1* et *c-ets-2* séparés. Pour les gènes *erg* et *fli*, l'existence de deux gènes séparés a été montrée chez les Mammifères (l'homme et la souris) : la divergence *erg/fli* pourrait s'être produite, de la même façon, avec l'apparition des Vertébrés (Lautenberger *et al.*, 1992 ; Laudet *et al.*, 1993).

Chez l'homme, les gènes *ets-1* et *fli-1* sont localisés sur le chromosome 11 et sont distants de 240 kpb. Les gènes *ets-2* et *erg* sont eux localisés sur le chromosome 21 et sont distants de 260 kpb. La disposition en tandem sur des chromosomes différents des gènes *ets-1/fli-1* et *ets-2/erg* évoque la possibilité d'une duplication du gène ancestral *ets/erg* sur un chromosome, événement qui aurait secondairement été suivi d'une translocation d'un locus porteur d'un exemplaire de chaque gène dupliqué (Figure 12).

Selon les modèles proposés par Lautenberger *et al.* (1992) et par Laudet *et al.* (1993), l'évolution de la famille ETS jusqu'à l'espèce humaine se serait faite ainsi :



- duplications dans les groupes ETS et ERG donnant respectivement d'une part les sous-groupes ETS-1 et ETS-2 et d'autre part ERG et FLI

Figure 11 : Arbre phylogénétique de la famille ETS d'après Laudet *et al*, 1993. Les flèches indiquent la correspondance entre les gènes de drosophile et les gènes de Vertébrés.

- 1ère étape : duplication primitive donnant les ancêtres de *ets* et de *erg*.
- 2ème étape : duplications, d'une part du gène *ets* pour donner les gènes *ets-1* et *ets-2* chez le xénope, les Oiseaux et les Mammifères (donc avant l'apparition des Amphibiens) ; d'autre part du gène *erg* pour donner les gènes *erg* et *fli-1* (avec l'apparition des Vertébrés).
- 3ème étape : remaniement du locus contenant ces gènes, *ets-1* et *fli-1* se retrouvant alors sur le chromosome 11 et *ets-2* et *erg* se retrouvant sur le chromosome 21, chez l'homme.

Les *ets* et *erg* d'oursin seraient ainsi les ancêtres n'ayant pas encore subi la deuxième duplication.

Pour la famille des gènes *Hox* (Murtha *et al.*, 1991) et la superfamille des récepteurs nucléaires (Laudet *et al.*, 1992), une évolution semblable est observée : une première duplication suivie d'une seconde duplication lors de l'émergence des Vertébrés.

Pour compléter cet aspect phylogénique de la famille ETS, il nous faut nous interroger sur le problème de la localisation du domaine ETS. En effet, la plupart des membres de la famille possèdent ce domaine en position C-terminale, mais quelques-uns l'ont en position N-terminale : il s'agit de Elk-1 et SAP-1 et d'autres en position centrale (Elf-1 et Pok). Cette disposition serait due à un mouvement d'exon dans un gène ("exon shuffling"), qui serait en partie responsable de l'évolution des gènes de la famille ETS.

#### **IV - Les protéines de la famille ETS chez les Vertébrés**

##### **1 - Fonction des protéines Ets**

Un rôle de facteur de transcription a pu être initialement attribué aux protéines c-Ets-1 murines, puis aviaires et humaines ( Bosselut *et al.*, 1990 ; Gunther *et al.*, 1990 ; Ho *et al.*, 1990; Wasylyk *et al.*, 1990 ). D'autres protéines présentant un domaine ETS ont aussi montré leur aptitude à reconnaître et à se fixer sur un motif particulier du promoteur de certains gènes. Ces protéines ont toutes les chances d'être douées de propriétés transactivatrices et donc d'intervenir dans la régulation de l'expression de ces gènes. Cette définition passera par la connaissance de leurs gènes cibles.

##### **2 - Le site de fixation à l'ADN des protéines Ets**

La recherche du site de reconnaissance spécifique du domaine ETS a d'abord été faite sur des "enhancers" viraux, avec les protéines Ets-1



aviaires : le site nucléotidique de reconnaissance est centré autour du motif GGA (Bosselut *et al.*, 1990 ; Gunther *et al.*, 1990 ; Wasylyk *et al.*, 1990). L'environnement nucléotidique de la séquence de reconnaissance GGA est important et semble intervenir (sûrement avec d'autres facteurs) dans la spécificité de reconnaissance de ces protéines vis-à-vis des gènes cibles qu'elles sont sensées réguler. Cependant, il n'est pas possible à l'heure actuelle d'attribuer un site de reconnaissance privilégié à chaque membre de la famille ETS. Des études *in vitro* ont montré l'importance de l'environnement nucléotidique :

- Les protéines Ets-1 et Elf-1 ne sont pas interchangeables vis-à-vis de leurs cibles naturelles (Wang *et al.*, 1992).
- Les protéines Ets-1 et Ets-2 ont des affinités différentes selon les promoteurs, en fonction des nucléotides proches du site de reconnaissance (Ascione *et al.*, 1992);
- Ets-1 est moins exigeante que la protéine PU.1 quant à l'importance du motif de liaison à l'ADN (Wasylyk *et al.*, 1992).

La technique de sélection d'oligonucléotides de séquence aléatoire (SAAB) a été utilisée pour tenter de définir une séquence consensus idéale pour plusieurs membres de la famille ETS (Erg, E74, GABP $\alpha$ , ER71 et ER81). Les résultats obtenus sont indiqués dans le tableau 4. Ces travaux montrent que dans tous les cas la séquence possède le motif GGA, qui est donc nécessaire à leur fixation, mais ces résultats n'indiquent pas la spécificité d'action des protéines Ets, dont les sites de reconnaissance sont proches.

La spécificité de reconnaissance et d'action de chaque membre de la famille ETS sur leurs gènes cibles doit donc être assurée, en dehors de la reconnaissance d'une séquence particulière, par l'interaction avec d'autres molécules et par la spécificité d'expression tissulaire de certains membres.

### **3 - Le domaine inhibiteur de la fixation à l'ADN**

La spécificité de liaison à l'ADN implique la présence de séquences en acides aminés particulières. Des expériences de délétions ont été effectuées sur les protéines Ets-1 et Ets-2 (Reddy et Rao, 1990 ; Hagman et Grossechedl, 1992 ; Lim *et al.*, 1992 ; Wasylyk *et al.*, 1992). Les protéines non modifiées se lient moins efficacement à l'ADN que les protéines amputées d'une région amino-terminale, région homologue à 55% pour les protéines Ets-1 et Ets-2, située juste en amont du domaine ETS. La protéine Elk-1 présente aussi un domaine inhibiteur de sa fixation à l'ADN (Rao et Reddy, 1992b), mais ce domaine n'existe pas dans la protéine PU.1 (Wasylyk *et al.*, 1992).

Protéine Ets	Site reconnu	Auteurs
Hu Erg	C/G C/A GGAA G/A T	Murakami <i>et al.</i> , 1993
Mo GABP $\alpha$	G/A C/G C/A GGAA/T G/A T/C	Brown et Mc Knight, 1992
Mo ER71	G/A C/G C/A GGAA/T G/A G/A T/C C	Brown et Mc Knight, 1992
Mo ER81	G/C C/G C/A GGAA/T G/A T/C N	Brown et Mc Knight, 1992
D E74	C/T C A GGAA G T	Urness et Thummel, 1990

**Tableau 4 : Séquences de fixation de protéines de la famille ETS, optimisées par sélection d'oligonucléotides aléatoires.**

#### 4 - Les domaines transactivateurs

L'ensemble des protéines Ets semble posséder une activité de transactivation. Pour les différents membres de la famille ETS, les domaines transactivateurs n'ont pas la même localisation dans la molécule et il n'existe pas vraiment de structure particulière pour ces domaines (Figure 13).

La protéine p68<sup>c-ets-1</sup> de poulet, comme nous avons pu l'indiquer précédemment, possède deux domaines d'activation de la transcription, à l'extrémité amino-terminale, alors que la p54<sup>c-ets-1</sup> n'en a qu'un seul. La protéine p58<sup>c-ets-2</sup> a une organisation analogue à la p68<sup>c-ets-1</sup> (Schneikert *et al.*, 1992 ; Chumakov *et al.*, 1993).

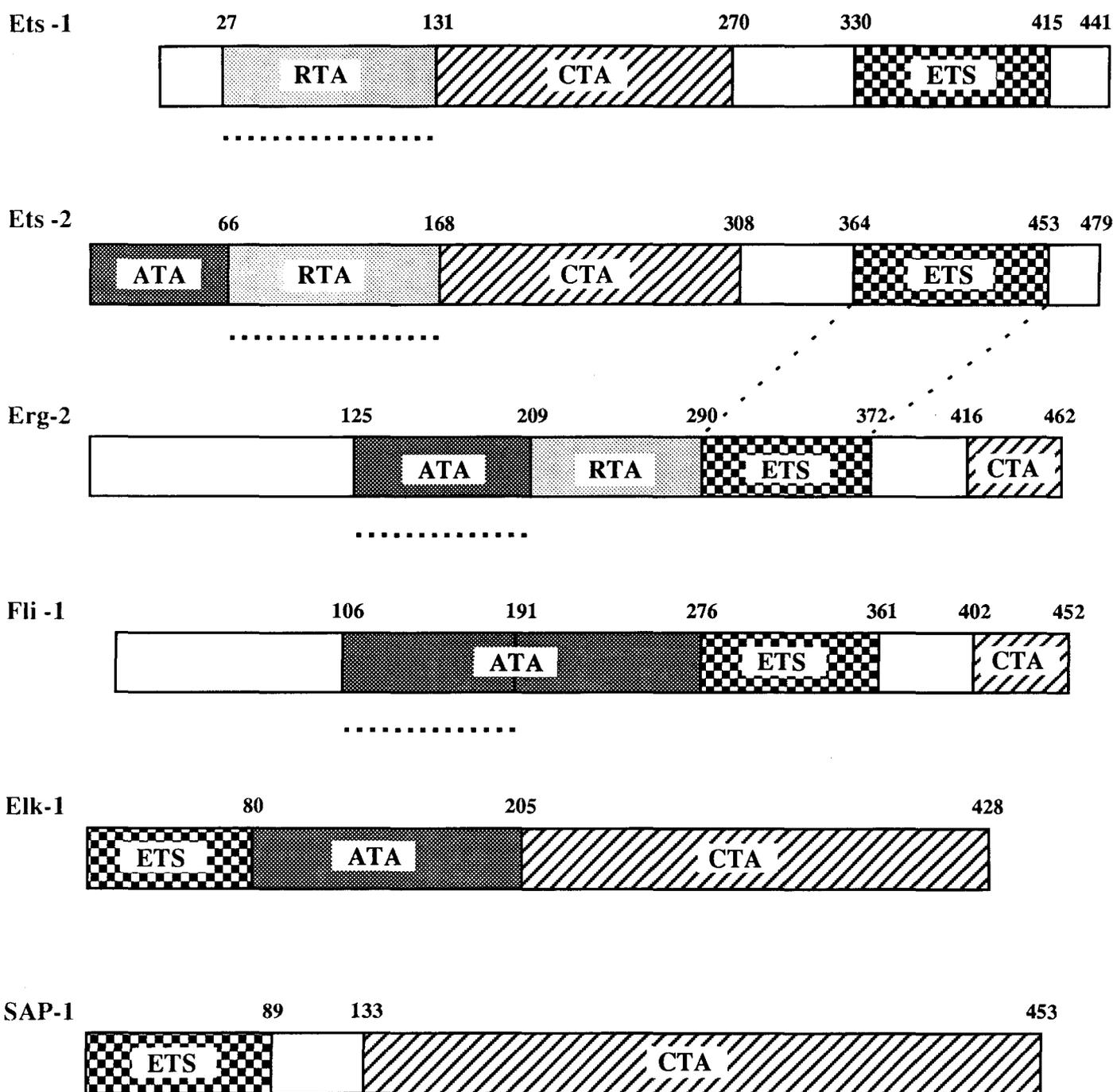
Les protéines Erg et Fli-1 humaines possèdent deux domaines transactivateurs, mais l'un est à l'extrémité N-terminale (ATA) comme pour les protéines Ets-1 et Ets-2, alors que l'autre est à l'extrémité C-terminale (CTA) (Rao *et al.*, 1993 ; Siddique *et al.*, 1993 ; Zhang *et al.*, 1993). Le domaine ATA de ces protéines est homologue par sa composition en acides aminés au domaine régulateur des protéines Ets-1 et Ets-2. Ce domaine correspond au domaine A défini par Watson *et al.* (1988b).

La protéine PU.1 a un domaine transactivateur de 75 acides aminés en position N-terminale (Hagemeier *et al.*, 1993). Le facteur PU.1 peut interagir directement, *in vitro*, avec le facteur général de transcription TFIID et avec la protéine suppressive de tumeur RB (rétinoblastome). Ces interactions sont médiées par le domaine transactivateur de PU.1, suggérant qu'elles sont importantes pour la fonction d'activation de la transcription de ce facteur spécifique des cellules lymphoïdes (Hagemeier *et al.*, 1993).

La protéine Elk-1 possède un domaine transactivateur en position amino-terminale et en position carboxy-terminale (Rao et Reddy, 1992a ; Bhattacharya *et al.*, 1993), alors que la protéine SAP-1, qui lui est très homologue, ne possède un domaine transactivateur qu'en position C-terminale (Bhattacharya *et al.*, 1993).

#### 5 - Interactions avec diverses protéines régulatrices

Comme nous avons pu le voir pour les protéines Ets-1, certaines protéines Ets semblent se fixer à l'ADN sous forme monomérique. Cependant, elles entretiennent aussi des interactions avec d'autres protéines, ces protéines étant de la famille ETS ou appartenant à d'autres familles de protéines. Ces associations permettent d'augmenter les possibilités de régulation des gènes cibles par la mise en œuvre de nombreuses combinaisons possibles. La spécificité de reconnaissance de séquences nucléotidiques peut de cette façon être modulée par la combinaison



ATA = domaine de transactivation amino-terminal  
 CTA = domaine de transactivation carboxy-terminal  
 RTA = domaine de régulation de la transactivation  
 ETS = domaine ETS

.... = région homologue entre les protéines Ets-1, Ets-2, Erg-2 et Fli-1



**Figure 13 : Structure et domaines fonctionnels des protéines**

**Ets-1, Ets-2, Erg-2, Fli-1, Elk-1 et SAP-1 humaines.**

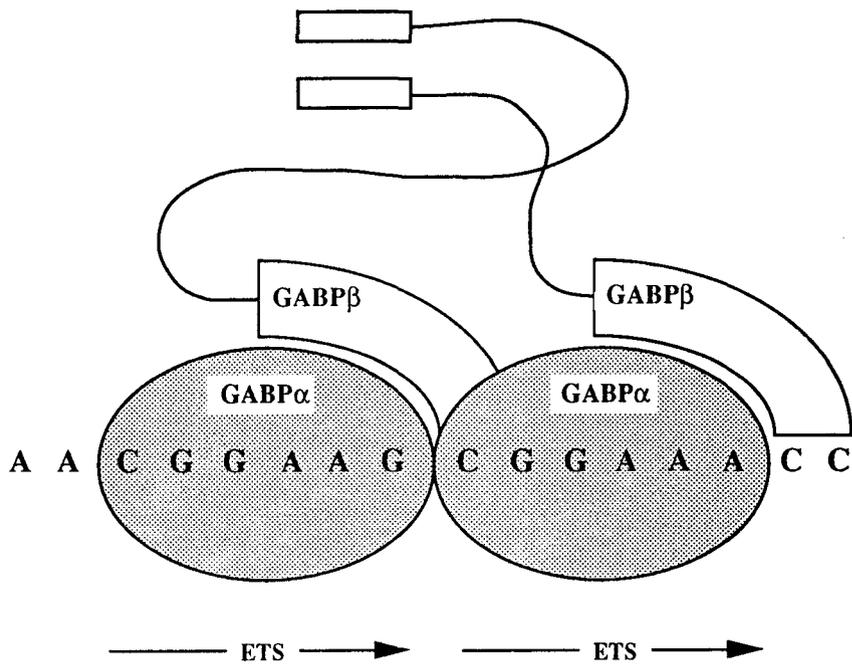
de différents facteurs sous forme de complexes. La formation des complexes aboutissant à l'activation transcriptionnelle des gènes cibles peut résulter d'une fixation indépendante des partenaires, comme nous avons pu le voir pour les protéines Ets-1 associées à d'autres facteurs au niveau de promoteurs de certains gènes. D'autre part, l'interaction de facteurs peut s'effectuer selon un ordre précis, le premier facteur étant indispensable à la fixation du deuxième, permettant ainsi l'activation transcriptionnelle. Nous allons illustrer ce propos par trois exemples.

**a) La reconnaissance de deux sous-unités protéiques permet la constitution d'un complexe transcriptionnel efficace**

La description moléculaire d'un complexe transcriptionnel impliquant une protéine de la famille ETS a été faite pour le promoteur d'un gène précoce IE ("Immediate Early") du virus de l'herpès (HSV-1) (Brown et Mc Knight, 1992). Il contient un élément *cis* régulateur riche en bases puriques (AACGGAAGCGGAAACC) avec deux séquences de fixation de type EBS. Le complexe protéique se liant à ce motif a été appelé GABP ("GA binding protein"). Le complexe GABP est constitué de deux types de sous-unités, GABP $\alpha$  qui possède un domaine ETS, a isolément une affinité faible pour l'ADN, et GABP $\beta$  qui seule n'a pas d'affinité pour l'ADN. GABP $\alpha$  et GABP $\beta$  s'associent pour former le tétramère  $\alpha_2\beta_2$  (Thompson *et al.*, 1991). Cet exemple illustre une caractéristique générale des protéines Ets : seules, elles sont des transactivateurs extrêmement faibles ; en interaction avec d'autres molécules, leur pouvoir transactivateur peut être considérablement augmenté. Dans le cas du complexe GABP, trois interfaces moléculaires hautement spécifiques ont été mises en évidence (Figure 14A) :

- interface GABP $\alpha$ /GABP $\beta$  : du côté GABP $\alpha$  c'est un domaine ETS complété par 37 acides aminés de l'extrémité C-terminale qui est impliqué dans l'interaction avec GABP $\beta$ . Du côté GABP $\beta$  il existe quatre répétitions de 32 à 33 acides aminés situés dans la moitié N-terminale ("ankyrin-type, ANK repeats"). GABP $\beta$  n'interagit avec aucune autre protéine Ets (Brown et Mc Knight, 1992).
- interface entre le complexe GABP et la séquence nucléotidique riche en A/G : elle se fait grâce au domaine ETS de GABP $\alpha$  et aux motifs répétés de GABP $\beta$ .
- dernière interface : le domaine de dimérisation localisé à l'extrémité C-terminale de GABP $\beta$ . GABP $\beta$  forme des homodimères en solution (Thompson *et al.*, 1991), et en présence de GABP $\alpha$ , chaque dimère forme avec

A



B

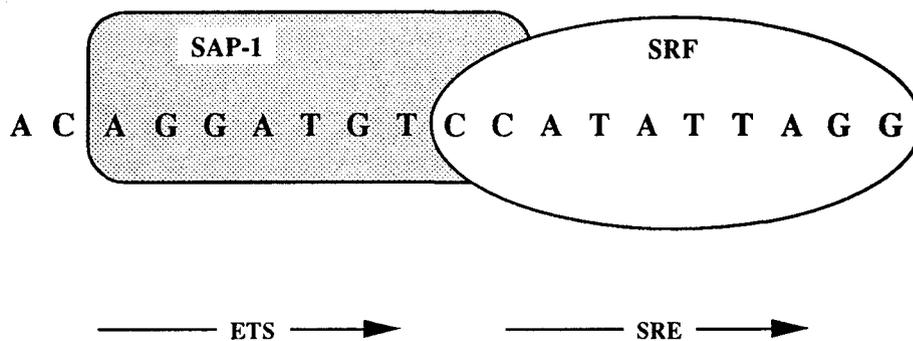


Figure 14 : Représentation schématique du complexe entre le tétramère GABP $\alpha$ 2 $\beta$ 2 et l'"enhancer" d'un gène IE du virus HSV-1 (A) et du complexe SAP-1/SRF lié au promoteur du gène *c-fos* (B), d'après Brown et McKnight, 1992.

deux GABP $\alpha$  un complexe hétéro-tétramérique qui permet la reconnaissance du motif (CGGAAG/A)<sub>2</sub> (Bzik et Preston, 1986 ; Triezenberg *et al.*, 1988 ; Spector *et al.*, 1990).

Ainsi dans l'exemple du complexe GABP, c'est vraisemblablement la reconnaissance initiale des sous-unités  $\alpha$  et  $\beta$  entre elles qui permet secondairement la reconnaissance spécifique de la séquence nucléotidique.

**b) La fixation d'un facteur dépend de son interaction avec une protéine de la famille ETS**

La protéine PU.1 qui est le membre le plus divergent de la famille ETS, s'associe avec le facteur de transcription NF-EM5 exprimé dans les lymphocytes B, sur l'"enhancer" 3' du gène de la chaîne kappa des immunoglobulines (Pongubala *et al.*, 1992). La fixation de NF-EM5 à l'ADN dépend de son association préalable avec la protéine PU.1. Le domaine de PU.1 impliqué dans l'association avec NF-EM5 inclut une séquence PEST riche en résidus proline, glutamine, sérine et thréonine, pouvant être impliquée dans la dégradation enzymatique des protéines (Rogers *et al.*, 1986). La spécificité d'action peut être ici attribuée à l'association d'une protéine de la famille ETS à un facteur s'exprimant dans un type cellulaire.

**c) La fixation à l'ADN d'une protéine Ets nécessite la fixation préalable d'un autre facteur protéique**

Le motif SRE ("Serum Responsive Element") est une composante importante du promoteur du gène *c-fos*. Le facteur p62<sup>TCF</sup> ("Ternary Complex Factor") forme un complexe ternaire avec le SRF ("Serum Response Factor") et la séquence cible de ce dernier, le SRE. p62<sup>TCF</sup> ne se fixe pas de manière autonome à l'ADN mais il est indispensable à la fixation du SRF sur le SRE. La protéine Elk-1 est fonctionnellement équivalente au p62<sup>TCF</sup> (Hipskind *et al.*, 1991). Un autre membre, SAP-1 ("SRF Accessory Protein-1") peut aussi former un complexe ternaire avec le SRF (Dalton et Treisman, 1992). Les deux facteurs de transcription Elk-1 et SAP-1 peuvent établir une liaison avec le motif nucléotidique SRE, cette liaison nécessitant cependant la fixation préalable ou simultanée du SRF sur le site SRE (Rao et Reddy, 1992b) (Figure 14B). Elk-1 et SAP-1 coopèrent individuellement avec le SRF pour activer le promoteur de *c-fos*, en stabilisant sans doute la liaison du SRF au motif SRE (Dalton et Treisman, 1992 ; Hill *et al.*, 1993). Les protéines Elk-1 et SAP-1 ont trois régions homologues : A, B et C (Rao *et al.*, 1989 ; Dalton et Treisman, 1992) (Figure 15). Les deux régions amino-terminales A et B sont indispensables à la fixation à

l'ADN et à la formation du complexe transcriptionnel (Dalton et Treisman, 1992 ; Janknecht et Nordheim, 1992 ; Rao et Reddy, 1992b). Le domaine C de la protéine Elk-1 est un domaine d'activation transcriptionnelle, phosphorylé sur de nombreux sites potentiels par une MAP kinase (Gille *et al.*, 1992 ; Marais *et al.*, 1993). Cette phosphorylation joue sur l'activité transcriptionnelle de Elk-1, mais pas sur sa fixation au SRE.

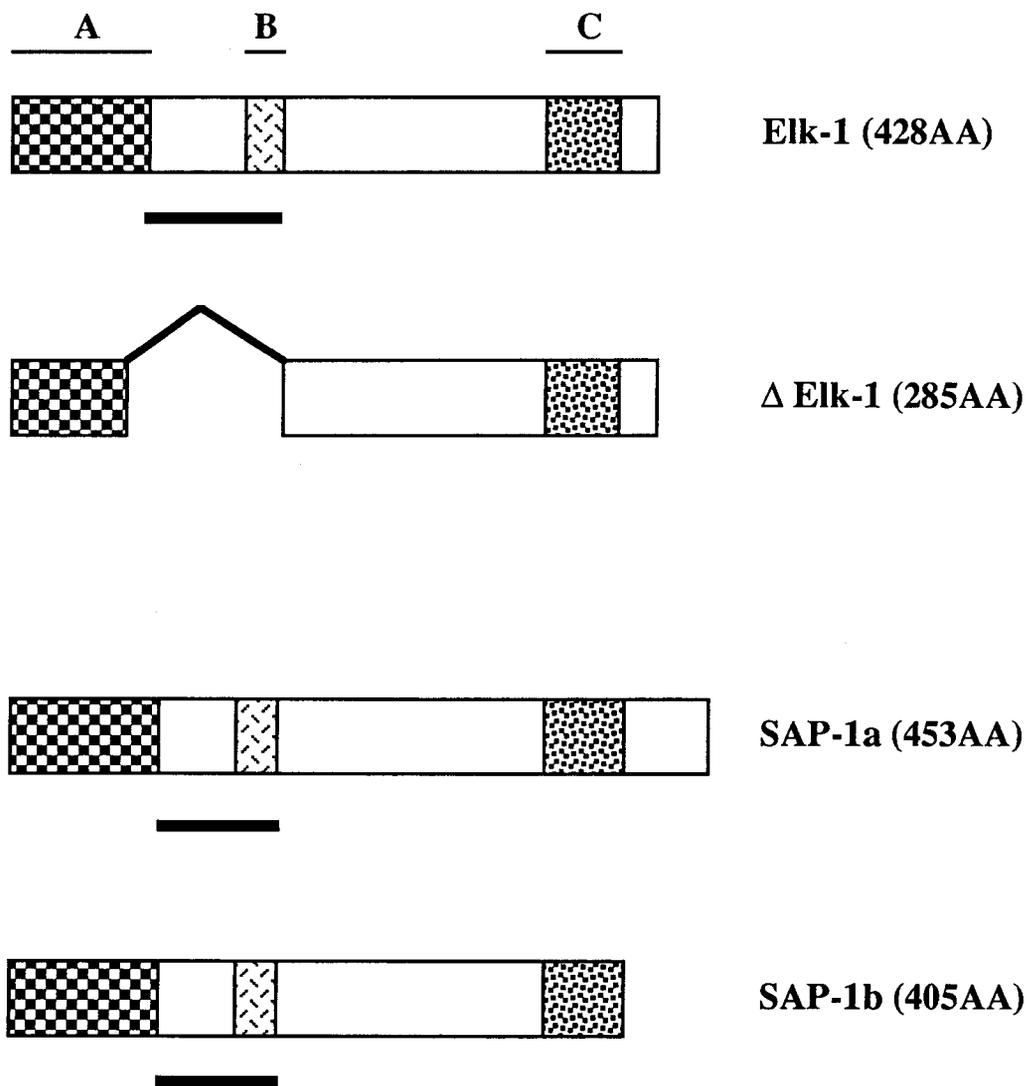
Des travaux récents ont montré que la protéine Elk-1 est phosphorylée *in vivo* par des MAP kinases activées (Janknecht *et al.*, 1993). D'autre part, les auteurs ont mis en évidence que la protéine Elk-1 ne fonctionne pas exclusivement sous forme d'un complexe ternaire avec le SRF et le SRE de *c-fos* mais qu'elle peut être aussi active en se fixant de façon autonome sur un EBS du SRF. Ainsi, cette double fonction d'un TCF pourrait permettre l'activation de différents types de gènes par l'intermédiaire des MAP kinases (Janknecht *et al.*, 1993).

## 6 - Régulation de l'activité des protéines Ets

Certains mécanismes de régulation n'ont été mis en évidence que pour les protéines Ets-1 et ont été développés plus haut : il s'agit de la régulation du transport membranaire et de la modification des groupements thiols. Cependant, des mécanismes de régulation ont aussi été observés pour d'autres membres de la famille ETS.

Les phosphorylations, bien décrites pour les protéines du gène *c-ets-1* pour lesquelles elles diminuent de façon sensible l'affinité de ces protéines pour l'ADN (Pognonec *et al.*, 1989), influencent aussi la liaison à l'ADN de Ets-2 (Boulukos *et al.*, 1990) et de Erg (Murakami *et al.*, 1993). La protéine Erg-2 humaine est phosphorylée et le taux de protéine phosphorylée augmente après 10 minutes de traitement avec le PMA qui est un activateur de la voie aboutissant à l'activation de la protéine kinase C, alors que le traitement avec un ionophore calcique n'a aucun effet, contrairement à ce qui est observé pour les protéines Ets-1 et Ets-2. En accord avec ce résultat, les protéines Ets-1 et Ets-2 possèdent des sites consensus de phosphorylation par la calcium/calmoduline kinase II alors que la protéine Erg-2 possède plusieurs sites potentiels de phosphorylation par la protéine kinase C (Murakami *et al.*, 1993).

La phosphorylation de la protéine PU.1 pourrait jouer un rôle différent (Pongubala *et al.*, 1993) : elle serait indispensable à l'interaction avec le co-facteur NF-EM5, pour former un complexe de transactivation efficace de l'"enhancer" 3' du gène de la chaîne kappa des immunoglobulines. La sérine



**—** domaine intervenant dans la liaison au SRE et dans l'inhibition de la liaison à l'ADN

**▣** domaine de liaison à l'ADN

A, B, C régions d'homologie entre les protéines Elk-1 et SAP-1

**Figure 15 :** Structure des protéines Elk-1 et SAP-1 humaines et présentation de leurs régions homologues.

(résidu 148) serait phosphorylée par la caséine kinase II.

Enfin, l'activité transcriptionnelle de la protéine Elk-1 sur le promoteur du gène *c-fos* dépend de la phosphorylation de la protéine qui s'effectue par une MAP kinase *in vitro* (Gille *et al.*, 1992 ; Marais *et al.*, 1993). Récemment, il a été montré que la protéine Elk-1 est hyperphosphorylée *in vivo*, après activation conjointe de MAP kinases et de la kinase Raf-1, cette dernière servant d'activateur indirect des MAP kinases *in vivo*. Cette phosphorylation qui peut s'effectuer sur plusieurs sites potentiels est accompagnée d'un changement conformationnel et se traduit par une augmentation du potentiel transactivateur de Elk-1. Ainsi, il semble que les MAP kinases contrôlent l'activité de la protéine Elk-1 par phosphorylation directe *in vivo* (Janknecht *et al.*, 1993).

Il existe donc différents chemins de transduction du signal membranaire aboutissant à la phosphorylation des membres de la famille ETS.

L'épissage alternatif est un autre mode de régulation des protéines Ets. Nous avons vu pour les protéines Ets-1 humaines que l'épissage alternatif peut aboutir à la création de facteurs de transcription dont les propriétés de liaison à l'ADN sont modifiées. L'épissage différentiel de l'ARNm de la protéine Elk-1 aboutit à la création de deux protéines à rôles antagonistes. L'association de Elk-1 au motif SRE pour activer la transcription de *c-fos* nécessite la présence du facteur SRF. Il existe une autre forme de Elk-1, la protéine  $\Delta$ Elk-1 (Rao et Reddy, 1993), amputée de la totalité du domaine concerné par l'interaction avec le facteur SRF et par l'inhibition de la fixation à l'ADN (Figure 15). La protéine  $\Delta$ Elk-1 est déletée de 11 acides aminés C-terminaux du domaine ETS ce qui ne l'empêche pas de reconnaître le motif SRE indépendamment de la présence du facteur SRF. Elle pourrait avoir un effet répresseur du complexe Elk-1/SRF par compétition pour la fixation au motif SRE, et donc empêcher l'activation du gène *c-fos*.

La protéine SAP-1 existe sous deux formes SAP-1a et SAP-1b résultant de l'épissage alternatif du gène (Dalton et Treisman, 1992) (Figure 15). La protéine SAP-1b semble interagir de façon plus efficace avec le SRF par rapport à la protéine SAP-1a mais elles activent de la même façon le promoteur du gène *c-fos*.

Des ARNm d'autres membres de la famille ETS, tels que ceux produits par le gène *erg* humain, présentent des épissages alternatifs. Cinq ADNc ont été identifiés (Rao *et al.*, 1987 ; Duterque-Coquillaud *et al.*, 1993). Ces isoformes sont générées par des épissages alternatifs et diffèrent par leur extrémité 5' et par la présence ou l'absence d'exons nommés A81 et A72

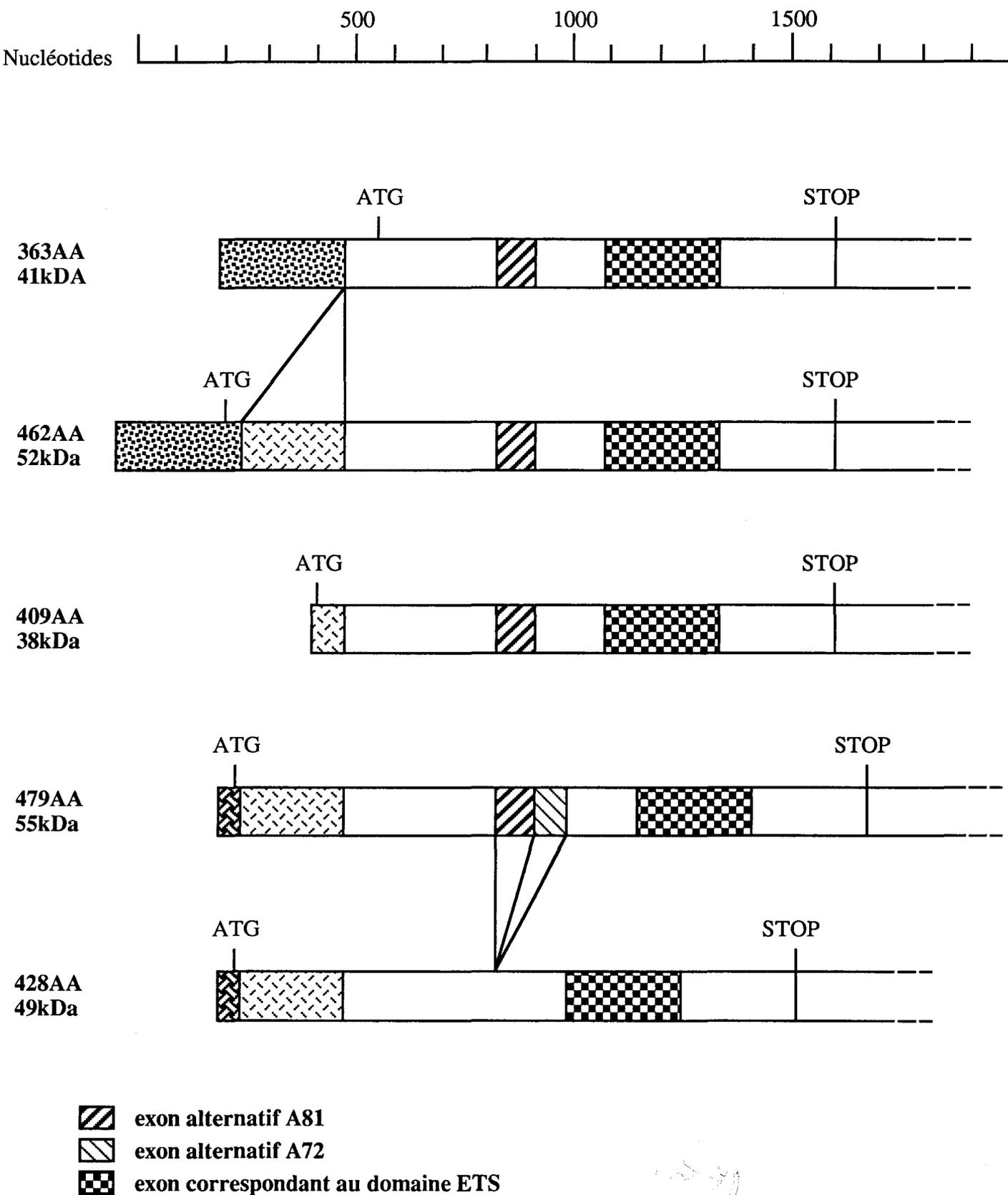


Figure 16 : Représentation schématique des ADNc du gène *erg* humain et des protéines correspondantes.

(Figure 16). Ces différents ADNc ont un cadre de lecture ouvert permettant la synthèse de protéines de 41 kDa (Erg-1), 52 kDa (Erg-2) (Rao *et al.*, 1987 ; Reddy *et al.*, 1987), 38 kDa, 49 kDa et 55 kDa (Duterque-Coquillaud *et al.*, 1993). Cependant, aucune relation n'a encore été établie entre ce mode de régulation et la spécificité de fonction des protéines étudiées.

La **stabilité** des protéines Ets peut aussi être un mode de régulation : comme beaucoup de facteurs de transcription, les protéines de la famille ETS ont une durée de vie relativement courte, de 45 à 85 minutes. La phosphorylation de certaines d'entre elles pourrait allonger leur durée de vie : ainsi, après une phosphorylation induite directement ou indirectement par la protéine kinase C, la durée de vie de la protéine Ets-2 passe de 20 minutes à plus de deux heures (Fujiwara *et al.*, 1988). Certaines protéines ont cependant une durée de vie nettement supérieure : plus de 21 heures pour la demi-vie de la protéine Erg-2 humaine (Murakami *et al.*, 1993).

## **V - Rôle physiologique des protéines Ets**

Afin de déterminer les rôles physiologiques des protéines, l'expression des gènes *ets* a été recherchée par les techniques de "Northern blots", d'hybridation *in situ* ou de RT PCR. Aucun type tissulaire ne semble complètement exempt de l'expression de gènes de la famille ETS. Cependant, le mécanisme exact de la régulation *in vivo* d'un gène par un membre de la famille ETS reste problématique : plusieurs facteurs interviennent dans les mécanismes complexes de régulation d'un gène eucaryote et les modèles de régulation actuellement proposés ont été étudiés *in vitro*.

### **1 - Les protéines Ets dans le système immunitaire**

L'expression des protéines Ets a été principalement étudiée dans le système immunitaire. Il a ainsi été montré que beaucoup d'entre elles s'exprimaient dans les cellules immunocompétentes (Ets-1, Ets-2, Spi-1, PU.1, Elf-1 et Fli-1). D'autre part, ces cellules sont bien typées en fonction de leur stade de différenciation et de leur état d'activation, ce qui en fait un matériel de choix pour l'étude de la régulation de gènes.

Au cours de la maturation des thymocytes, les gènes codant les chaînes ( $\alpha/\beta$  et  $\gamma/\delta$ ) du récepteur des cellules T (TCR) se réarrangent. L'expression de la protéine Ets-2 est concomitante au réarrangement de la chaîne  $\alpha$ . Il a été montré que les protéines Ets doivent se lier à un élément ( $T\alpha 2$ ) de

l'"enhancer" 3' du gène codant la chaîne  $\alpha$  (TCR- $\alpha$ ) en même temps que d'autres facteurs de transcription, pour que ce gène soit activé (Ho *et al.*, 1990). D'autre part, ces auteurs ont montré que la protéine Ets qui se fixe sur le motif T $\alpha$ 2 pourrait interagir avec les autres facteurs se liant aux autres éléments (T $\alpha$ 1, T $\alpha$ 3 et T $\alpha$ 4) de cet "enhancer" et moduler l'activité transcriptionnelle du gène.

Dans les lymphocytes T, un gène est potentiellement la cible des protéines Ets : le gène codant l'enzyme terminale désoxynucléotidyl transférase (TdT) (Ernst *et al.*, 1993). Cette enzyme est exprimée au début du développement lymphocytaire et son taux chute au cours de la maturation des lymphocytes (Hermans *et al.*, 1989). Un élément *cis* régulateur du promoteur du gène TdT est reconnu à la fois par les protéines Ets-1 et Fli-1 *in vitro*. Cependant, le mécanisme de cette régulation reste inconnu.

Un autre phénomène intéressant, bien qu'imparfaitement compris, est la balance existant dans les lymphocytes T humains entre les taux de protéines Ets-1 et Ets-2. En culture, ces lymphocytes expriment assez fortement la protéine Ets-1 à l'état quiescent, mais très peu la protéine Ets-2. L'activation des lymphocytes par un anticorps anti-CD3 inverse les rapports Ets-1/Ets-2, sauf si les synthèses protéiques sont bloquées par la cycloheximide (Bhat *et al.*, 1990). Ainsi, Ets-2 semble jouer un rôle dans la prolifération cellulaire, alors que Ets-1 permettrait de maintenir les lymphocytes T à l'état quiescent.

Au contraire, les études effectuées sur des macrophages montrent que Ets-2 y est exprimée surtout dans les cellules à l'état quiescent et différencié (Ghysdael *et al.*, 1986b ; Boulukos *et al.*, 1988), où son expression est directement liée au MGF ("Macrophage Growth Factor") nécessaire à la survie des macrophages aviaires en culture (Boulukos *et al.*, 1990).

Un autre gène, celui de l'interleukine 2 (IL-2), n'est pas transcrit dans les lymphocytes T quiescents, mais peut être induit. Ce gène possède dans son promoteur des sites de reconnaissance pour les protéines de la famille ETS, permettant la fixation de la protéine Elf-1 (Thompson *et al.*, 1992). Elf-1 est liée à la protéine RB non phosphorylée dans les lymphocytes quiescents, l'ensemble RB/Elf-1 se séparant lors de la phosphorylation de RB, au moment de l'entrée des cellules dans le cycle cellulaire (Wang *et al.*, 1993). La protéine Elf-1 ne pourrait pas à elle seule induire l'expression du gène de l'IL-2, deux autres protéines, Jun B et Fra-1 sont aussi impliquées dans cette induction.

Dans les lymphocytes B, la protéine PU.1 complexée à un autre facteur, NF-EM5, participe à la régulation de l'"enhancer" 3' du gène de la chaîne légère kappa des immunoglobulines (Grant *et al.*, 1992 ; Pongubala *et al.*, 1992, 1993). PU.1 coopère aussi avec Ets-1 pour activer l'"enhancer" de la chaîne mu

non réarrangée des immunoglobulines (Nelsen *et al.*, 1993).

## 2 - Les protéines Ets dans le développement embryonnaire

La recherche de l'expression des protéines Ets dans le développement embryonnaire a été surtout effectuée chez le poulet et la souris. Ces travaux concernent surtout les transcrits du gène *c-ets-1* et ont déjà été exposés.

Chez le xénope, quatre membres de la famille ETS ont jusqu'ici été signalés : X Ets-1 (Stiegler *et al.*, 1990) et X Ets-2 (Wolff *et al.*, 1990 ; Burdett *et al.*, 1992), qui existent sous deux versions, a et b, du fait de la tétraploïdie du génome. Une première étude a permis la mise en évidence des transcrits de 4,4 kpb et 7,5 kpb de X *ets-1* dès les premiers stades de l'ovogénèse jusqu'à la fin de l'embryogénèse. Ces transcrits ne sont pas détectés dans les organes adultes, sauf le transcrit de 7,5 kpb qui a été détecté dans la rate (Stiegler *et al.*, 1990).

D'autres transcrits de 3,2 et 9,5 kpb ont ensuite été détectés au cours de l'ovogénèse et de l'embryogénèse (Stiegler *et al.*, 1993). Les transcrits de 4,4 et 9,5 kbp seraient en fait des messagers en cours de maturation, les autres transcrits pouvant coder deux protéines isoformes. Dans la région non traduite à l'extrémité 3' de l'ARNm, une séquence UUUUUAU, existant en trois exemplaires, et localisée en amont du signal de polyadénylation AAUAAA, serait impliquée dans le contrôle de la polyadénylation spécifique de la maturation ovocytaire.

Les transcrits de 1,7 kpb et 3,2 kpb de X *ets-2* ont été mis en évidence au cours de l'ovogénèse, au début de l'embryogénèse et dans de nombreux tissus adultes (Wolff *et al.*, 1991). Le transcrit de 3,2 kpb majoritairement exprimé serait polyadénylé de façon plus importante au niveau de la maturation ovocytaire. Dans la région non traduite en 3' de l'ARNm de 3,2 kpb, on observe la séquence UUUUUAU mise en évidence dans les transcrits de X *ets-1* et pouvant correspondre au même signal de contrôle de la polyadénylation spécifique de la maturation ovocytaire.

L'ovocyte du xénope est un matériel de choix aisément accessible à l'expérimentation : la microinjection d'oligonucléotides anti-sens de l'ARNm X *ets-2* de 3,2 kpb ne permet pas la dissolution de l'enveloppe nucléaire qui est un indicateur important de la maturation méiotique de l'ovocyte de xénope (Chen *et al.*, 1990). La méiose peut cependant reprendre son cours normal après microinjection d'ADNc *ets-2* humain. Ainsi, d'une part les protéines Ets-2

humaine et de xénope peuvent ici se substituer l'une à l'autre et ont donc les mêmes capacités fonctionnelles, d'autre part la protéine Ets-2 paraît jouer un rôle clé dans la maturation méiotique.

### 3 - L'activation des protéines Ets

L'expression aberrante et l'altération structurale de facteurs de transcription constituent des mécanismes moléculaires fréquents et intervenant précocement dans l'oncogenèse (Lewin, 1991). Chez les Mammifères inférieurs et les espèces aviaires, ces altérations sont souvent médiées par une insertion rétrovirale. Chez l'homme, la dérégulation ou l'altération structurale des facteurs de transcription résultent fréquemment de réarrangements génomiques somatiques (Cleary, 1991).

#### a ) Activation par insertion rétrovirale

En plus de *c-ets-1* (sous sa forme transduite dans le rétrovirus E26), deux gènes de la famille ETS (*PU.1/spi-1* et *fli-1*) sont impliqués dans l'apparition d'érythroleucémies causées par un rétrovirus chez les Mammifères inférieurs, suggérant ainsi l'importance des gènes *ets* dans les érythroleucémies et les lymphomes.

La transformation par certains rétrovirus peut impliquer l'intégration de l'ADN proviral à proximité ou à l'intérieur de gènes cellulaires spécifiques.

Les gènes *fli-1* et *spi-1* sont impliqués dans l'érythroleucémie murine de Friend causée par différentes souches de virus (Moreau-Gachelin *et al.*, 1988 ; Ben-David *et al.*, 1990). Le complexe de Friend est composé d'un virus déficient pour la réplication, le SFFV (Spleen Focus Forming Virus) et d'un virus compétent pour la réplication, le F-MuLV (pour revue, Kabat, 1990). Ce complexe infecte les animaux adultes alors que le F-MuLV seul infecte l'animal à la naissance. Pour certaines souris "prédisposées", le SFFV peut induire une érythroleucémie aiguë. L'analyse d'érythroleucémies induites par deux variants du SFFV (FV-A et FV-P) montre que dans 95% des cas, les clones de cellules érythroleucémiques murines contiennent une insertion provirale à proximité du gène *spi-1* (SFF-proviral integration site 1). Cette intégration active l'expression du gène *spi-1* et contribue au processus tumorigène (Moreau-Gachelin *et al.*, 1988, 1989).

De la même façon, un autre gène de la famille ETS, *fli-1* (Friend leukemia integration-site 1) est activé par une intégration provirale dans 75 % des cas de souris infectées à la naissance par le F-MuLV (Ben David *et al.*,

1991). De façon intéressante, le gène *fli-1* n'est pas réarrangé, ni activé dans les érythroleucémies causées par le SFFV et de façon similaire, le gène *spi-1* n'est pas réarrangé ni activé dans les érythroleucémies causées par le F-MuLV. Les deux virus SFFV et F-MuLV présentent une spécificité stricte d'intégration, respectivement en amont des gènes *spi-1* et *fli-1*. Cette spécificité, qui semble surprenante au premier abord car ces deux gènes appartiennent à la même famille, peut refléter une différence dans les cellules cibles de la transformation par SFFV et F-MuLV.

### **b ) Activation par réarrangement génomique**

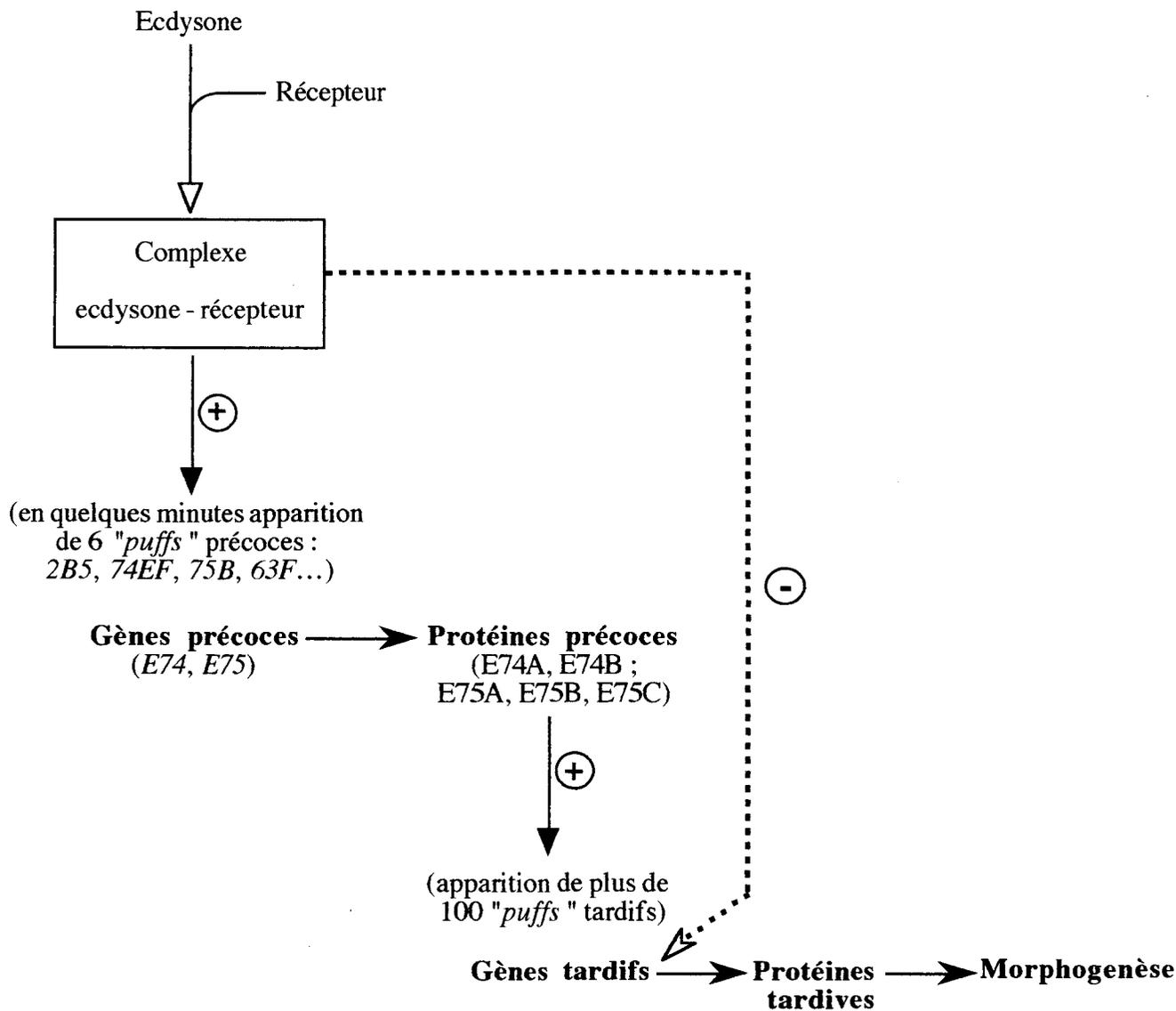
Jusqu'à une époque récente, les protéines Ets n'étaient pas directement impliquées dans les cancers chez l'homme. Cette implication a été mise en évidence dans le cas du sarcome d'Ewing qui est une tumeur d'origine neuroectodermique. Une translocation réciproque des chromosomes 11 et 22 permet à partir du chromosome 22 la synthèse d'une protéine de fusion comprenant un domaine amino-terminal homologue à 40 % de l'extrémité carboxy-terminale de l'ARN polymérase II (dénommé EWS) rabouté au domaine ETS carboxy-terminal de la protéine Fli-1 (Delattre *et al.*, 1992).

La protéine EWS/Fli-1 (contrairement à Fli-1) possède la capacité de transformer des cellules NIH3T3, et les deux domaines EWS et Fli-1 sont nécessaires à cette activité (May *et al.*, 1993a). Cette protéine pourrait fonctionner comme un facteur de transcription aberrant contribuant au pouvoir transformant dans les sarcomes d'Ewing. Des travaux récents ont montré que la protéine EWS/Fli-1 est localisée dans le noyau des cellules, qu'elle se lie à l'ADN sur une séquence spécifique et que la région EWS peut être considérée comme un domaine activateur de la transcription plus actif que le domaine correspondant de la protéine Fli-1 (May *et al.*, 1993b). La protéine EWS/Fli-1 est donc le premier exemple d'une protéine Ets directement impliquée dans la genèse d'une tumeur humaine ne concernant pas les cellules hématopoïétiques.

Récemment il a été montré après analyse de 89 sarcomes d'Ewing et tumeurs comparables, qu'il pouvait y avoir, outre la fusion EWS/Fli-1, la fusion du gène *EWS* au gène *erg*. La position des points de cassure au niveau des chromosomes est restreinte aux introns 7 à 10 du gène *EWS* et est dispersée au niveau des introns 3 à 9 des gènes *fli-1* et *erg* (Zucman *et al.*, 1993).

## **VI - La famille ETS chez les Invertébrés**

La recherche de gènes de la famille ETS a été entreprise chez les



**Figure 17 : Modèle d'Ashburner (1974) pour le contrôle génétique des "puffs" du chromosome polyténique, par l'ecdysone, complété selon les travaux de Burtis *et al.* (1990).**

Invertébrés. Ces travaux, encore fragmentaires, concernent principalement la drosophile. Ils ont été récemment complétés par les études de Degnan *et al.* (1993) concernant des Métazoaires Diploblastiques et Triploblastiques.

## A - Les protéines E74 de la drosophile

### 1 - Le modèle d'Ashburner

A la fin du troisième stade larvaire, chez la drosophile, le passage au stade nymphal est initié par un pic d'une hormone stéroïde, l'ecdysone. Ce pic hormonal induit la différenciation des tissus imaginaux en tissus adultes, et la disparition des tissus larvaires. L'étude de la régulation génique de ces modifications dépendantes de l'ecdysone est facilitée par l'observation des chromosomes polyténiques géants des glandes salivaires de la drosophile. On peut ainsi observer des "puffs" chromosomiques, représentant une décondensation de la chromatine et correspondant à des régions transcriptionnellement actives. Le maintien des glandes salivaires larvaires de drosophile en culture permet de disposer d'un modèle "*in vitro*" pour étudier l'apparition de ces "puffs" chromosomiques sous l'action de l'ecdysone. Ashburner *et al.* (1974) ont pu ainsi proposer un modèle avec une hiérarchie dans le contrôle génétique de la formation des "puffs", sous l'action de l'hormone stéroïde ecdysone (Figure 17). Complété par les études de biologie moléculaire, le modèle d'Ashburner met en jeu différentes protéines intervenant dans le contrôle transcriptionnel, dont certaines présentent un domaine ETS de liaison à l'ADN.

Le "puff" précoce 2B5 a été étudié du point de vue génétique et cytogénétique. Ce "puff" comporte le locus *BR-C* ("Broad Complex"). Chez les larves déficientes en *BR-C* (Thummel, 1990) le "puff" précoce 63F ne se forme pas, les "puffs" précoces 74EF et 75B n'ont qu'un tiers de leur taille normale et il n'y a pas formation de "puffs" tardifs. Donc, dans le locus *BR-C*, se situerait un gène précoce dont le produit faciliterait la formation des "puffs" précoces et serait essentiel à l'induction des "puffs" tardifs. Selon le modèle d'Ashburner, un gène précoce s'exprime sous l'action de l'ecdysone ; sa transcription serait réprimée par les protéines synthétisées, et il coderait des protéines régulatrices se liant à l'ADN et interagissant avec d'autres "puffs" induits par l'ecdysone. Au locus 74EF, le gène *E74* a été particulièrement bien étudié.

### 2 - Le gène *E74* et ses produits

Après l'isolement d'un ADNc (Wolfner, 1980), qui s'hybridait *in situ* en huit sites sur des chromosomes polytènes de drosophile, notamment au

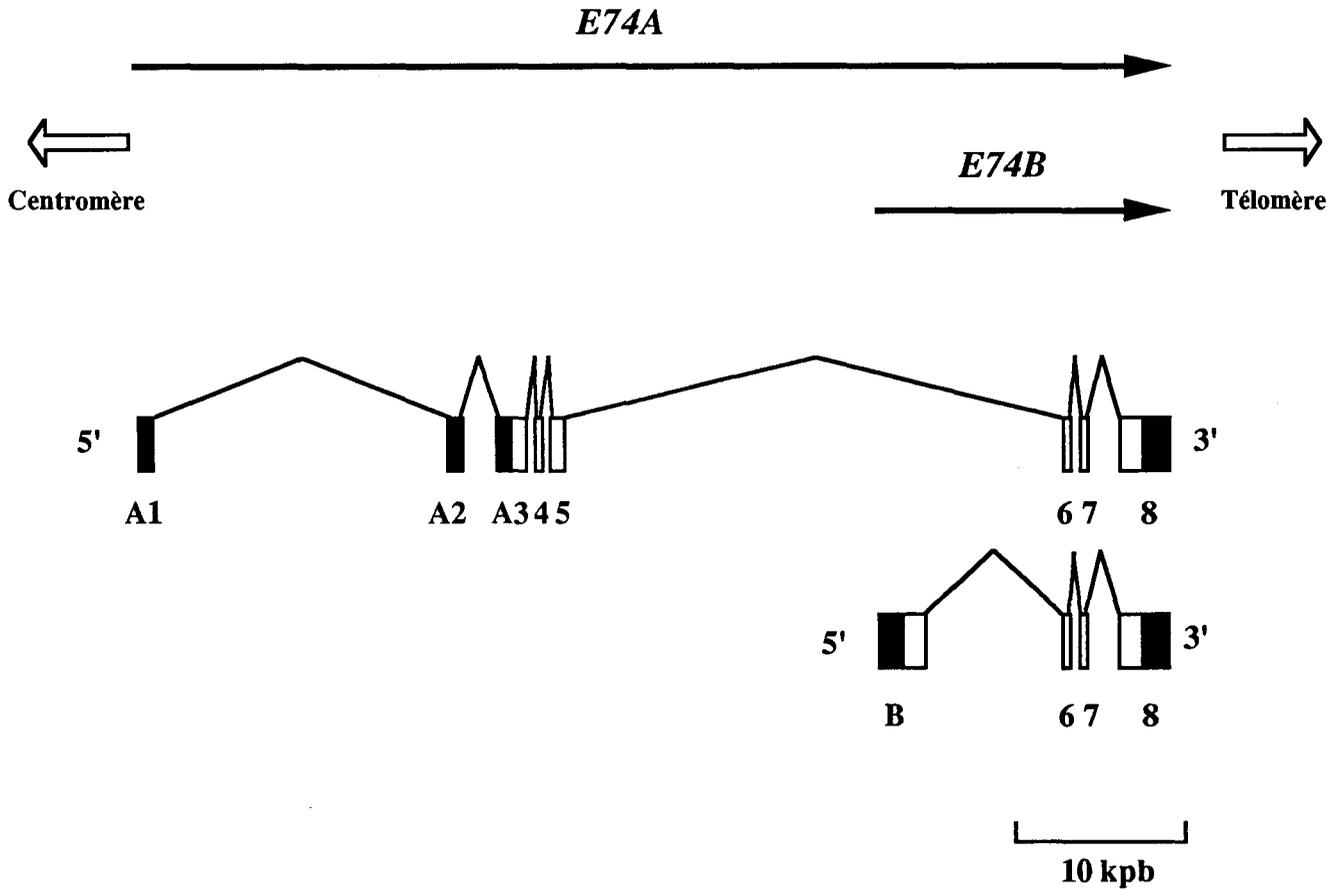


Figure 18 : Organisation du gène *E74* de la drosophile

Les flèches indiquent l'orientation de la transcription de *E74A* et *E74B*, par rapport au centromère et au télomère du bras gauche du chromosome 3 de la drosophile. Les rectangles correspondent aux exons, avec en noir les séquences 5' et 3' non traduites.

niveau du "puff" précoce *74E*, le criblage d'une banque d'ADN génomique puis une marche sur le chromosome entre les loci *74E* et *74F*, ont permis de cloner toute cette région *74EF* (Burtis *et al.*, 1990). Une région de 105 kpb a alors été cartographiée. A partir d'ARNm extraits de tissus stimulés par l'ecdysone deux ARNm ont été mis en évidence : un ARNm de 6 kpb, appelé ARNm *E74A* et un ARNm d'environ 4,9 kpb appelé ARNm *E74B* (Figure 18). Les analyses d'extension d'amorce et de protection contre la nucléase  $S_1$  ont en fait permis d'identifier trois sites d'initiation de la transcription pour l'ARNm *E74B*, donnant des ARNm d'environ 4,8 kpb et 5,1 kpb. Les trois transcrits *E74B1a*, *B1b*, *B2* du gène *E74B* ont des sites d'initiation de la transcription respectivement distants de 8 et 300 pb, mais ne permettent la production que d'une seule protéine.

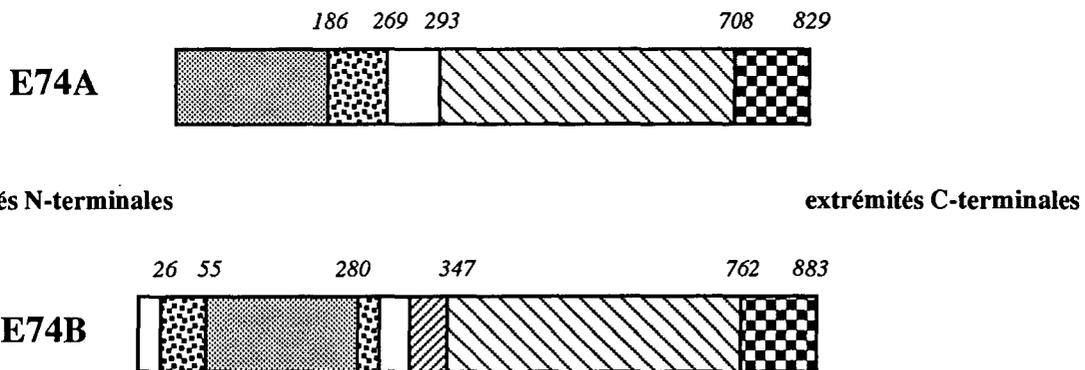
L'unité de transcription *E74A* s'étend sur 60 kpb et contient dans son cinquième exon les sites de départ de la synthèse des transcrits *E74B*. L'ARNm *E74A* de 6 kpb a un aspect inhabituel : il dispose d'une longue séquence de tête en 5' de 1891 nucléotides avec 17 triplets AUG, cette séquence, dont la taille et la complexité sont inhabituelles, pouvant être impliquée dans un mécanisme de régulation post-transcriptionnelle (Boyd *et al.*, 1991).

Les différents ARNm conduisent à la formation des deux protéines apparentées, *E74A* et *E74B*, ayant un domaine C-terminal commun, mais différant par leur extrémité N-terminale (Burtis *et al.*, 1990 ; Urness et Thummel, 1990) (Figure 19).

L'extrémité C-terminale commune contient un motif de 85 acides aminés homologue au domaine ETS. C'est ici le premier cas connu d'une protéine de la famille ETS dont la synthèse est sous contrôle d'une hormone stéroïde.

Le fait que les protéines *E74A* et *B* n'aient pas la même extrémité N-terminale suggère que ces protéines ont des rôles différents dans la régulation. D'ailleurs, la transcription de *E74B* n'est vraisemblablement pas régie par les mêmes mécanismes que celle de *E74A* : les transcrits *E74B* précèdent les transcrits *E74A*, la transcription de *E74B* nécessitant une concentration en ecdysone 25 fois plus faible que pour *E74A* (Karim et Thummel, 1991).

L'étude de la séquence peptidique a permis de classer les protéines *E74* dans un nouveau groupe de molécules dans lequel on a rangé depuis, la protéine Elf-1 ("E74-like factor 1") (Thompson *et al.*, 1992), qui peut se fixer sur des EBS de promoteurs de gènes actifs dans les cellules T humaines. Le domaine ETS de Elf-1 ne présente que 52 % d'homologie avec celui des protéines Ets-1 ou Ets-2 humaines, mais 85 % d'homologie avec celui de *E74*.



— Partie C-terminale commune :

-  région de 415 acides aminés, avec des motifs répétés : répétition d'une Ser et plusieurs Glu, de plusieurs Ala, ou de motifs PRD (une His alternant avec des résidus Pro ou Ser)
-  région riche en acides aminés basiques (Lys et Arg), contenant un domaine ETS

— Partie N-terminale spécifique :

-  domaine riche en acides aminés acides (Glu et Asp)
-  répétition d'acides aminés :
  - E74A : répétition de Ala, Asn et Gly
  - E74B : à l'extrémité N-terminale du domaine acide : répétition de Ser
  - à l'extrémité C-terminale du domaine acide : répétition de Ala
-  région riche en acides aminés basiques dans E74B

Figure 19 : Structure des protéines E74A et E74B de la drosophile.

### 3 - Mode d'action de l'ecdysone

L'ecdysone active le promoteur de *E74A* (Thummel *et al.*, 1990), probablement au niveau d'un élément de réponse situé dans le premier intron (Burtis *et al.*, 1990).

L'ecdysone agit par l'intermédiaire d'un complexe avec un récepteur protéique. Le mode d'action de cette hormone stéroïde peut être illustré par le cas de gènes de choc thermique, qui sont, chez la drosophile, activés par l'ecdysone : dans le promoteur du gène *hsp27*, Riddihough et Pelham (1987) ont montré l'existence d'un élément de réponse de 23 pb, sur lequel peut se lier de façon spécifique un polypeptide de 80 à 90 kDa. Cet élément de réponse à l'hormone ecdysone (EcRE), semblable aux éléments de réponse aux hormones trouvés chez les mammifères (Martinez *et al.*, 1991), serait en fait le site de fixation de deux protéines : l'une est le produit du gène du récepteur de l'ecdysone (*EcR*) (Kœlle *et al.*, 1991) ; l'autre est le produit du gène ultraspiracle *usp* (Oro *et al.*, 1990). Le gène *usp* ne possède pas d'intron, donc pas de plasticité génétique. Le gène *EcR* permet de produire par épissage alternatif trois protéines qui diffèrent par leur extrémité N-terminale et qui sont exprimées à des stades et dans des tissus différents.

La fixation d'une protéine EcR à une séquence EcRE de réponse à l'hormone ecdysone ne peut se faire que s'il y a présence de la protéine Usp. Il y a ainsi formation d'un complexe de deux protéines et de l'hormone qui seul est actif *in vivo* (Yao *et al.*, 1992). Une modulation de la réponse hormonale est possible par la présence de trois produits alternatifs du gène *EcR*, pouvant intervenir dans la formation du complexe.

### 4 - Régulation transcriptionnelle des gènes précoces

L'auto-répression des gènes exprimant les protéines précoces permettrait de déterminer la durée de leur activité en même temps que la quantité de molécules à produire, en réponse à un pic d'ecdysone.

Dans le milieu du gène *E74*, trois sites de liaison de la protéine *E74A* ont été identifiés : ces trois sites possèdent la séquence G/AGGAA, séquence de fixation de domaines ETS (Urness et Thummel, 1990). Mais bien que la protéine *E74A* soit capable de reconnaître *in vitro* ces trois sites, il ne semble pas qu'elle soit l'effecteur de cette répression. C'est apparemment la protéine de réponse précoce *E75* qui en est responsable.

Localisé au niveau d'un autre "puff" précoce, *75B*, le gène *E75* a une unité de transcription de 50 kpb avec un seul site de départ de la transcription. Les transcrits *E75A*, *E75B* et *E75C* permettent la synthèse de trois

protéines apparentées avec la même extrémité C-terminale, mais qui ne présentent pas le domaine ETS observé dans les protéines E74.

Les gènes *E74* et *E75* s'expriment à six moments du développement qui sont caractérisés par un pulse d'ecdysone. Ils sont aussi transcrits à d'autres moments, chez l'embryon de 18 h où aucun pic d'ecdysone n'est détecté, suggérant une méthode alternative de contrôle de la transcription.

### **B - Les autres gènes de la famille ETS chez la drosophile**

Six membres de la famille ETS, autres que *E74*, ont été isolés chez la drosophile.

Le gène D *ets-2* a été partiellement isolé par Pribyl *et al.* (1987, 1988) à partir d'une banque d'ADN génomique. Le domaine ETS se présente sous forme de deux exons séparés par un intron de 63 pb. En 1992, Chen *et al.* (1992) ayant isolé un ADNc, ont montré la présence d'un deuxième intron en aval du domaine ETS. Par la suite, Klämbt (1993) a caractérisé un gène qu'il a appelé *pointed* : ce gène est très vraisemblablement le gène D *ets-2*. Le gène *pointed* possède deux régions promotrices distantes de 50 kpb, permettant la synthèse de deux transcrits : les transcrits *P1* et *P2*. *P1* est constitué des quatre exons 1, 2, 3, 4 ; *P2* est constitué des exons I, II, III, IV, V, 2, 3, 4, les exons 2, 3 et 4 étant communs aux deux transcrits. Le domaine ETS se situe à l'extrémité C-terminale des protéines et est homologue à 95 % de la protéine Ets-1 humaine. Du côté N-terminal du produit correspondant au transcrit *P2*, un domaine présente une forte homologie avec les protéines Ets-1 et Ets-2 sur 70 acides aminés que Klämbt a appelé domaine Pointed. L'expression du gène *pointed* est exigée lors de la constitution du système nerveux central, ainsi que dans la trachée et certains muscles.

Le gène D *ets-3* a été isolé par PCR sur de l'ADN génomique avec des oligonucléotides dégénérés correspondant à des régions bien conservées dans différents domaines ETS alors connus (Chen *et al.*, 1992). Il présente un intron dans le domaine ETS, qui est la seule partie du gène actuellement connue. Traduite en acides aminés, cette séquence se présente comme très proche des protéines Erg et Fli, avec respectivement 95 et 97 % d'homologie. D *ets-3* produit un ARNm de 0,8 kpb, présent au même taux quel que soit le stade de développement.

Utilisant la même technique que pour D *ets-3*, Chen *et al.* (1992) ont aussi isolé une partie des gènes D *ets-4* et D *ets-6*. D *ets-6* présente un intron

dans la région codant le domaine ETS. Ce domaine présente 86 et 87 % d'homologie avec respectivement les protéines Fli et Erg.

Les transcrits du gène D *ets-3* et le transcrit de 0,85 kpb du gène D *ets-6* sont présents à peu près au même taux quel que soit le stade de développement. Les transcrits des gènes D *ets-2* et D *ets-4* sont exprimés à un faible taux durant le stade larvaire. Le transcrit D *ets-2* semble corrélé à la régulation par l'ecdysone et son expression semble complexe, en différents points de l'embryon. Le transcrit de 2,5 kpb du gène D *ets-6* induit au stade pupal, coïncide avec le taux d'ecdysone (Chen *et al.*, 1992).

L'expression des gènes D *ets-3* et D *ets-6* est localisée à certains neurones du système nerveux ventral de l'embryon, mais tandis que celle de D *ets-3*, contrôlée par le *complexe bithorax*, est exclusivement thoracique, celle de D *ets-6* est aussi abdominale (Chen *et al.*, 1992).

Le gène D *elg* (*Drosophila ets like gene*) a d'abord été identifié par Pribyl *et al.* (1991). L'ADNc a été isolé par Chen *et al.* (1992). La séquence en acides aminés prédite du domaine ETS de ce gène est proche de celle de GABP $\alpha$  (73 % d'homologie). L'ARNm de 2 kpb de D *elg* est détecté par hybridation *in situ* dans la plupart des cellules de l'embryon (Pribyl *et al.*, 1991) et l'expression de ce gène ne peut être reliée à un type cellulaire. Au même locus, le brin complémentaire du brin permettant la transcription de l'ARNm de D *elg*, produit un ARNm de 2,8 kpb : ce transcrit est exprimé au même taux et en même temps que le messager de 2 kpb de D *elg* (cependant cet autre gène n'est pas caractérisé).

Une mutation au codon 397 du transcrit du gène D *elg* entraîne la substitution d'une tyrosine, présente dans tous les domaines ETS, par une cystéine : cette mutation *tne* ("tiny eggs") est responsable de la stérilité des femelles (Schulz *et al.*, 1993a). Le gène D *elg* est nécessaire au développement normal de l'embryon de drosophile : des mutations situées en dehors du domaine ETS provoquent des anomalies de la segmentation abdominale (Schulz *et al.*, 1993b).

La déficience du gène *pokkuri* (ou *yan*), découvert de façon concomitante par Lai et Rubin (1992) et Tei *et al.* (1992) provoque une différenciation erronée de certaines cellules nerveuses de l'œil en photorécepteurs de type R7 surnuméraires (Tei *et al.*, 1992). Le domaine ETS de la protéine Pok est homologue à 46 % de la séquence en acides aminés du produit du gène *erg* humain et ce domaine est en position centrale dans la protéine. L'extrémité amino-

terminale de Pok possède une séquence de 80 acides aminés fortement homologue aux extrémités amino-terminales des protéines Ets-1, Ets-2 et Erg humaines.

### **C - La famille ETS chez d'autres Invertébrés**

Chez les Invertébrés autres que la drosophile, le premier membre de la famille ETS a été décrit par Chen *et al.* (1988) chez l'oursin *Lytechinus variegatus*. D'abord mis en évidence par hybridation d'ADN génomique avec un fragment de *v-ets* du virus E26, le fragment *HindIII* cloné dans un phage a été isolé par hybridation avec une sonde D *ets-2*, alors récemment isolée (Pribyl *et al.*, 1987). Après séquençage, Chen *et al.* (1988) ont observé une forte homologie de la séquence traduite en acides aminés avec la protéine Ets-2 humaine, et ont appelé l'Ets d'oursin : Su Ets-2. Les analyses ultérieures ont révélé que Su Ets-2 avait en fait une plus grande homologie avec les protéines Ets-1.

Le gène Su *ets-2* s'exprime sous forme d'un seul transcrit de 6,8 kpb. Chez l'embryon, ce transcrit est trouvé dès les premiers stades du développement, jusqu'au stade blastula. Chen *et al.* (1988) indiquent qu'il n'est présent dans aucun autre tissu adulte. Cependant, nous pensons que ces observations devraient être revues à la lumière des nombreux travaux effectués ces dernières années sur la famille des gènes *ets*.

Chez l'oursin, un autre gène a été identifié. Il s'agit d'un gène proche du groupe ERG : Su *erg*, cité par Lautenberger *et al.* (1992), mais dont la séquence, à notre connaissance, n'a pas encore été publiée et n'a été que partiellement décrite dans les banques de données. Cette séquence traduite comporte 115 acides aminés et est très proche des protéines Erg et Fli de l'homme et de la souris, tant au niveau du domaine ETS, que de la région qui lui est contiguë à l'extrémité 3' et que Lautenberger *et al.* (1992) ont appelé domaine R.

Les membres de la famille ETS chez d'autres Invertébrés ont été mis en évidence par Degnan *et al.* (1993). Ces auteurs ont recherché ces gènes, en criblant les ADN génomique ou complémentaire extraits d'organismes Métazoaires Diploblastiques et Triploblastiques, par PCR. Ils ont ainsi isolé des séquences leur permettant d'aligner 33 acides aminés, dont certaines présentent une forte homologie avec des membres de la famille ETS.

Chez l'ascidie *Styela montereyensis* (Tunicier) et l'étoile de mer *Patiria miniata* (Echinoderme), les séquences restituées sont à 100 % identiques de la protéine Ets-1 humaine. Outre ces deux Deutérostomiens, les travaux de Degnan *et al.* (1993) ont porté sur différents Protostomiens : le crabe *Cancer*

*antennarius* (Crustacé) ; l'ormeau *Haliotis rufescens* (Mollusque Gastéropode) ; la coquille Saint-Jacques *Hinnites giganteus* (Mollusque Bivalve) ; le polychète *Chaetopterus variopedatus* (Annélide) et deux Nématodes appartenant à des espèces non précisées. Chez toutes ces espèces, sauf un Nématode, les auteurs ont obtenu des séquences ne différant que d'un, deux ou trois acides aminés, par rapport à la protéine Ets-1 humaine. Chez une espèce de Nématode, ils ont trouvé une séquence proche de la protéine Erg humaine (seuls deux acides aminés différent).

Les travaux de Degnan *et al.* (1993) se sont aussi étendus aux Diploblastiques : un Cnidaire (l'anémone de mer *Anthropleura elegantissima*), un Cténaire (*Pleurobranchia bachei*) et deux espèces d'éponges (*Haliclona* sp et *Tethya aurantia*). Ils ont ainsi identifié chez l'anémone un membre du groupe ETS et un membre du groupe ELG, chez le Cténaire un membre du groupe D ETS-4 et chez les éponges un membre du groupe ETS pour une espèce, et des membres des groupes ETS et ELK pour l'autre.

Ces résultats indiquent donc que la famille ETS est une famille de gènes bien conservés, qu'elle est fort ancienne, et enfin qu'elle est déjà bien diversifiée dès le stade Diploblastique puisque l'on y trouve déjà représentés, les groupes ETS, ELG, ELK et D ETS-4. Cependant, on peut objecter à cette étude que les séquences traduites ne sont que de 33 acides aminés, et leur situation dans les différents groupes de la famille ETS reste donc à confirmer. D'autre part, il est possible de se demander si ces séquences amplifiées à partir d'ADN génomique, correspondent bien à des gènes fonctionnels.

## **VII- Conclusion**

Les protéines de la famille ETS présentent toutes un domaine de liaison à l'ADN et un ou plusieurs domaines de transactivation. Ces protéines sont des facteurs de transcription constituant une nouvelle famille de protéines eucaryotiques se liant à l'ADN, car leur domaine de liaison à l'ADN constitue un nouveau motif structural de reconnaissance ADN-protéines. Les protéines appartenant à cette famille reconnaissent une cible nucléotidique riche en purine (GGA) et plusieurs protéines sont capables de reconnaître le même motif. La spécificité et la sélectivité d'action pour la régulation de gènes particuliers, en dépit de la reconnaissance de séquences nucléotidiques proches, constituent l'originalité de cette famille de facteurs de transcription. Ainsi, ces protéines interagissent avec d'autres protéines de régulation ou facteurs de transcription

pour former des complexes stables nécessaires à l'activation de la transcription de gènes uniques, parfois dans un type cellulaire spécifique. Les régions divergentes (domaines A et B) de ces protéines Ets peuvent représenter des structures interagissant avec des protéines cellulaires particulières, conférant ainsi une fonction spécifique à chaque membre de la famille ETS. Comme cela a déjà été rapporté, la formation du complexe entre les sous-unités GABP $\alpha$  et GABP $\beta$  ou entre Elk-1 et SRF apparaît être très spécifique, de façon qu'aucun autre membre de la famille ETS ne peut les remplacer dans ces interactions. Enfin, ces interactions permettent une activation synergique de la transcription et compensent ainsi le faible pouvoir transactivateur des protéines Ets. Cette famille de gènes est aussi très ancienne car de nombreux membres identifiés chez les Vertébrés ont été mis en évidence chez des Métazoaires Diploblastiques. Ceci indique l'importance des protéines codées par ces gènes, dont le rôle reste encore en grande partie à découvrir.

# **MATÉRIEL BIOLOGIQUE**

Les *Nereis diversicolor*, Métazoaires Triploblastiques Coelomates métamérisés, appartiennent à l'Embranchement des Annélides (Figure 20). La présence de nombreuses soies portées par des parapodes les placent dans la Classe des Polychètes.

Cette espèce marine ou d'eau saumâtre vit dans la vase noire des estuaires où elle participe à la chaîne alimentaire du littoral marin. Les animaux utilisés dans nos manipulations, ont été récoltés dans l'estuaire de l'Aa, à Grand-Fort-Philippe (Nord). Les vers récoltés sont conservés au laboratoire en eau de mer à la température d'environ 17°C.

### **I - Situation phylogénique des Annélides Polychètes**

Contrairement aux Métazoaires Diploblastiques tels que les Spongiaires et les Cœlentérés, les Métazoaires Triploblastiques Cœlomates sont en principe des animaux libres et mobiles. Ils sont aquatiques, primitivement marins et deviennent dulçaquicoles puis terrestres dans les groupes les plus évolués.

Les Triploblastiques sont caractérisés par un troisième feuillet, le mésoderme, interposé entre l'endoderme et l'ectoderme. Chez les Cœlomates, le mésoderme est généralement creusé de cavités, les cavités cœlomiques.

Il existe deux types de Triploblastiques Cœlomates : l'un a le système nerveux le long de la région ventrale et correspond aux Invertébrés, l'autre, au contraire, étend son système nerveux dans la région dorsale et répond à la structure des Cordés. Selon la terminologie d'A. Lameere, ils sont respectivement appelés les Hyponeuriens et les Epineuriens.

Les Invertébrés Cœlomates Hyponeuriens constituent de loin la partie la plus importante des Cœlomates. Tandis que les Cœlomates Epineuriens ou Cordés comprennent deux Embranchements, les Procordés et les Vertébrés (comptant environ 50 000 espèces), les Cœlomates Hyponeuriens considérablement plus variés, plus nombreux, plus diversifiés, sont représentés par plusieurs Embranchements, eux-mêmes souvent très complexes, les Annélides, les Echinodermes, les Mollusques, les Arthropodes, comptant des centaines de milliers d'espèces animales actuellement vivantes. Ils sont aussi représentés par les formes fossiles les plus anciennes. Les grands Embranchements des Invertébrés se sont différenciés au cours des longues époques précambriennes. En dépit de leur diversité, il est possible de retrouver l'unité de structure du Cœlomate Hyponeurien, conservée lors du développement larvaire, et de suivre les

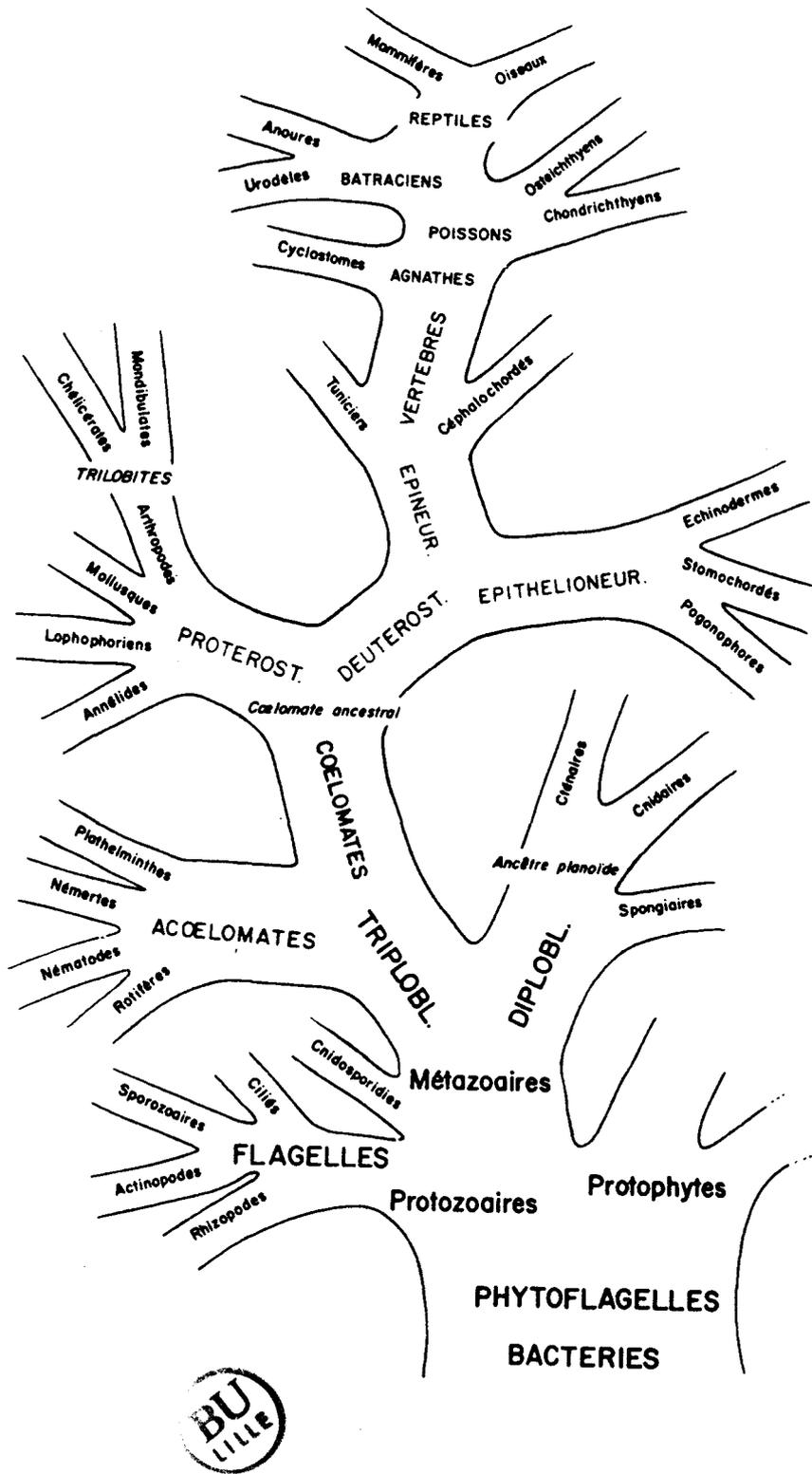


Figure 20 : Arbre généalogique du Règne animal, d'après Beaumont et Cassier, 1973.

modifications subies par cette structure, pour constituer les divers Embranchements.

Ce sont les Annélides, et plus particulièrement les Polychètes, qui constituent, par la conservation à l'état adulte de structures archaïques, le type même du Cœlomate.

## **II - Organisation générale de la néréis**

Le corps des néréis comporte trois régions distinctes (Figure 21) :

- le prostomium, région céphalique pourvue des organes des sens (appendices céphaliques et yeux) et du cerveau.
- le métastomium, métamérisé, c'est-à-dire formé d'une série d'anneaux pratiquement tous semblables contenant des structures et organes identiques. Ils comprennent chacun une paire de cavités cœlomiques. Les enveloppes épithéliales des sacs cœlomiques s'affrontent dans le plan sagittal pour former le mésentère ventral et le mésentère dorsal assurant la suspension des organes axiaux (tube digestif et vaisseaux). En dehors du premier segment (peristomium), chaque métamère présente une paire de parapodes (expansions latérales garnies de nombreuses soies, à fonction locomotrice).
- le pygidium, segment postérieur portant deux cirres et percé par l'anus n'a ni cœlome, ni néphridie.

La coordination de l'organisme est assurée par un système nerveux central et condensé, qui comprend une paire de ganglions cérébroïdes (le cerveau) situés dans le prostomium et une chaîne nerveuse ventrale (type hyponeurien).

L'acquisition de l'état triploblastique s'est traduite par une complexité croissante de l'organisation générale et par la différenciation d'organes nouveaux (appareils digestif, circulatoire et excréteur).

Le système circulatoire est clos : sang et hémocytes circulent dans un compartiment séparé anatomiquement de la cavité cœlomique. Il n'y a pas de gonades différenciées : les gamètes et d'autres cellules cohabitent dans la cavité cœlomique et sont soumis aux messagers chimiques et aux métabolites du liquide cœlomique. Nos connaissances sur les cellules cœlomiques sont encore fragmentaires par rapport à celles d'autres Invertébrés (Annélides Oligochètes, Mollusques ou Insectes).

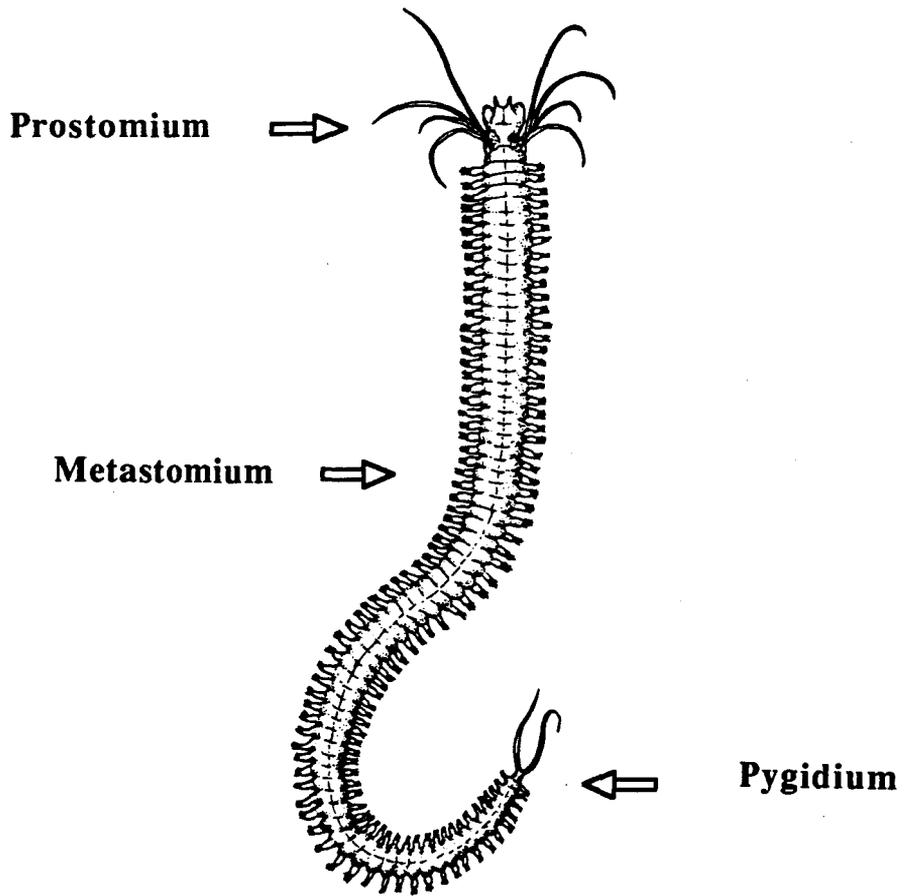


Figure 21 : *Nereis diversicolor*

Les *Nereidae* ont été étudiées dans notre laboratoire selon plusieurs axes et notamment le rôle des cellules cœlomiques et les phénomènes de régénération et de cicatrisation.

### III - Les cellules cœlomiques

Les cellules libres dans le liquide cœlomique, outre les gamètes, sont de différents types :

- les éléocytes, souvent comparés au "corps gras" d'Insecte, synthétisent la vitellogénine qui est absorbée par l'ovocyte en cours de maturation (Baert et Slomianny, 1992). Ils phagocytent aussi les sarcolytes, c'est-à-dire les débris de fibres musculaires en dégénérescence (Dhainaut, 1966a et b) ;
- les granulocytes appartiennent à différents types cellulaires, dont trois sont bien identifiés : les G I (20 % des granulocytes), les G II (54 %) et les G III (24 %).

*N. diversicolor* vit dans des estuaires particulièrement riches en polluants. Ces animaux supportent une pollution en métaux lourds assez élevée, la pollution par le plomb et le cadmium étant mieux supportée que celle par le chrome ou le mercure.

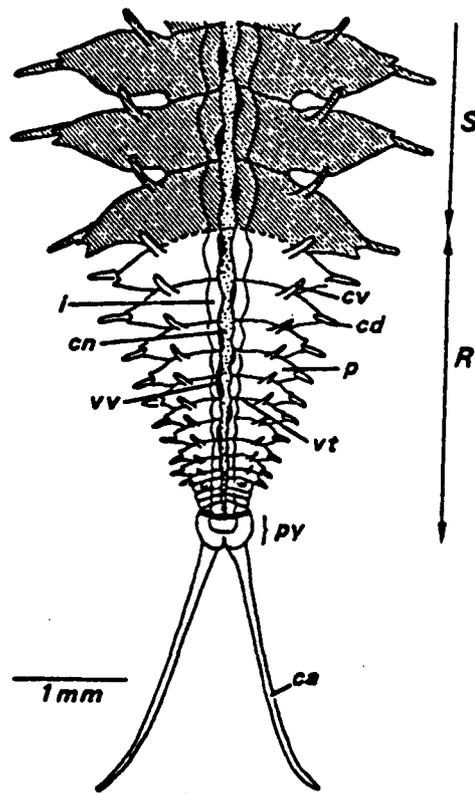
L'invasion du milieu intérieur par des éléments exogènes (bactéries, parasites, implants expérimentaux) ou une intoxication par des substances toxiques, provoquent l'activation cellulaire des granulocytes avec un processus de phagocytose, de dégranulation ou d'encapsulation. L'intégrité du cerveau est indispensable au moins dans les réactions antiparasitaires.

**La dégranulation** est une réaction d'exocytose avec libération du matériel granulaire dans le milieu intérieur.

Les G III dégranulent complètement selon les stimulants (liposaccharides bactériens, métaux lourds...), les G II présentent une dégranulation plus ou moins complète et les G I ne sont pas ou sont peu impliqués dans ces processus.

Le matériel libéré par dégranulation contient une teneur élevée en lysozymes et en  $\beta$ -glucuronidase ainsi qu'une métalloprotéine, la MP II, contenue dans les granulocytes G I (Dhainaut-Courtois *et al.*, 1987) et qui possède une activité antibactérienne (Dhainaut *et al.*, 1989).

La protéine cœlomique MP II, qui peut lier les métaux, se rencontre sous une forme monomérique de 14 kDa (Demuyne *et al.*, 1993). Le site de synthèse dans les granulocytes de type G I a été précisé par immunologie (Dhainaut-Courtois *et al.*, 1987 ; Porchet-Henneré *et al.*, 1987).



ca cirre anal  
 cd cirre dorsal  
 cn chaîne nerveuse  
 cv cirre ventral  
 vt vaisseau sanguin transversal  
 vv vaisseau sanguin ventral

p parapode  
 py pygidium  
 R régénérat  
 S souche

**Figure 22 : Représentation schématique de la régénération caudale chez *Nereis diversicolor* (vue ventrale d'un régénérat, 30 jours après l'amputation), d'après Combaz et Boilly, 1974.**

**L'encapsulation** consiste en l'isolement, grâce à une enveloppe continue de coelomocytes, des corps étrangers dans une capsule qui forme finalement un nodule brun (présence de mélanine) dans lequel le parasite, le tissu altéré ou le corps étranger se nécrosent, si leur nature chimique le permet, et sont mis hors d'état de nuire. Ces mécanismes de défense représentent un modèle remarquable de coopération cellulaire (Porchet-Henneré *et al.*, 1987).

La chronologie des événements (établie par implantation de billes de plastique dans la cavité coelomique) est la suivante :

- reconnaissance par les cellules G III qui émettent de larges filopodes,
- lyse des cellules au contact de l'implant entraînant la formation d'une couche de débris cellulaires : "couche collante" supposée nécessaire et attractive pour les vagues suivantes de cellules réactives,
- les cellules G II très étirées, aplaties, viennent se superposer comme des écailles d'un bulbe d'oignon formant une épaisse couche feuilletée. Ce sont les premières cellules à recouvrir le corps étranger. On observe une intense dégranulation et le produit de dégranulation circule jusqu'à l'implant.
- les cellules G I se disposent en périphérie et forment un réseau d'emballage.

Les granulocytes, cellules circulantes à l'origine, forment dans la capsule un tissu de cellules juxtaposées, avec de nombreux desmosomes et jonctions de type "gap".

#### **IV - La cicatrisation et la régénération**

##### **Régénération caudale des néréis**

Le processus de régénération des néréis comprend trois phases (Combaz et Boilly, 1974) : après la cicatrisation, on observe la mise en place d'un blastème de régénération puis la différenciation de ce blastème qui s'effectue en deux étapes : la première aboutissant à l'édification du pygidium; la seconde permettant la différenciation des segments sétigères (Figure 22). L'évolution de la régénération est présentée dans le tableau 5. Les délais d'apparition des différents stades de la régénération sont sujets à des variations individuelles parfois considérables ; les temps notés correspondent à ceux de la majorité des individus observés à la température de l'expérience.

Etapes de la régénération postérieure	Stades successifs	Date d'apparition des différents stades (Nombre de jours après la section)
Cicatrisation	—	2 - 3
Formation du pygidium	<ul style="list-style-type: none"> <li>• Edification du blastème de régénération</li> <li>• Différenciation et croissance du blastème de régénération <ul style="list-style-type: none"> <li>- Différenciation du segment pygidial <ul style="list-style-type: none"> <li>. Apparition du vaisseau sanguin périanal</li> <li>. Différenciation de la musculature pygidiale</li> </ul> </li> <li>- Différenciation des cirres anaux <ul style="list-style-type: none"> <li>. Apparition des 2 mamelons anaux</li> <li>. Apparition des ébauches des cirres anaux</li> </ul> </li> <li>- Croissance du pygidium (achevée)</li> </ul> </li> </ul>	<p>5 - 6</p> <p>6 - 8 7 - 8</p> <p>6 - 8 7 - 8</p> <p>13 - 15</p>
Mise en place des métamères	<ul style="list-style-type: none"> <li>• Apparition de la zone de prolifération</li> <li>• Différenciation et croissance des métamères <ul style="list-style-type: none"> <li>- Segmentation du régénérat (apparition des dissépiments) <ul style="list-style-type: none"> <li>. Dissépiments 1 et 2</li> <li>. Dissépiments 3, 4 et 5</li> <li>. Dissépiments 6, 7, 8 et 9</li> <li>. Dissépiments 10, 11, 12, 13 et 14</li> </ul> </li> <li>- Différenciation parapodiale <ul style="list-style-type: none"> <li>. Apparition des 1ères ébauches parapodiales</li> <li>. Apparition des 1ers cirres parapodiaux</li> <li>. Apparition d'acicules et de soies sur les 1ers segments néoformés</li> </ul> </li> <li>- Croissance segmentaire (achevée) (longitudinale et transversale)</li> </ul> </li> </ul>	<p>8 - 9</p> <p>9 - 11 12 - 18 18 - 28 30 - 60</p> <p>12 - 15 13 - 16</p> <p>18 - 20</p> <p>60 - 70</p>

**Tableau 5 : Etude chronologique de la régénération postérieure chez *Nereis diversicolor* après amputation simple (selon Combaz et Boilly, 1974).**

### **A - La cicatrisation**

Immédiatement après l'amputation, la paroi intestinale sectionnée vient en contact avec le tégument proche de la section. La soudure entre ces deux parois survient deux à trois jours plus tard.

### **B - La formation du pygidium**

L'édification du blastème de régénération s'effectue en cinq à six jours après l'amputation. Un blastème annulaire non pigmenté, traversé par l'intestin, est édifié au niveau de la jonction des tissus impliqués dans la cicatrisation.

La différenciation du blastème de régénération en un pygidium complet est achevée environ 13 à 15 jours après l'amputation.

### **C - La mise en place des métamères**

La croissance du régénérat s'effectue simultanément à sa différenciation. L'allongement du régénérat se manifeste 8 à 9 jours après l'amputation, avec l'apparition d'une zone de prolifération (zone de croissance). Le régénérat s'allonge et se segmente ; les premiers segments formés sont repoussés vers l'avant au fur et à mesure que de nouveaux segments se différencient, les segments les plus postérieurs étant toujours les plus jeunes.

*N. diversicolor* ne régénère qu'en direction caudale. Il n'a jamais été observé de migration des cellules régénératrices : les cellules à l'origine des structures régénérées proviennent du seul segment contigu au plan d'amputation.

Tous les éléments régénérateurs présentent une phase d'activation au cours de laquelle les cellules synthétisent une quantité importante d'ARN avant de s'engager dans l'histogenèse (Wattez-Combaz, 1985) : mais (c'est une différence avec les Annélides Oligochètes) ce phénomène ne se produit qu'au niveau du front de section et vraisemblablement sous l'influence "trophique" du système nerveux. D'autre part, chaque type de cellule régénératrice assure la reconstitution de tous les tissus correspondant au feuillet embryonnaire dont il provient. Enfin, les blastocytes mésodermiques (provenant des épithéliums cœlomiques) jouent un rôle de premier plan dans la morphogenèse du régénérat : ils sont indispensables à la différenciation, à la croissance et à la métamérisation du régénérat.

La régénération caudale dépend étroitement de la présence de l'intestin ; l'absence de la chaîne nerveuse ne l'empêche pas, bien qu'elle retentisse sur la vitesse de régénération et l'organisation du régénérat (différenciation des cirres anaux et des parapodes).

# RÉSULTATS

La stratégie utilisée pour mettre en évidence des membres de la famille ETS chez la néréis a été la suivante :

#### **Etude à partir de l'ADN génomique**

- après avoir montré l'existence de membres potentiels de la famille ETS par "Southern" avec deux sondes (une virale et une aviaire), ces gènes ont été recherchés dans deux génothèques, mais aucun résultat n'a été obtenu.
- une stratégie utilisant la PCR a alors été envisagée et a permis d'identifier un membre de cette famille. Une partie de ce gène a ensuite été isolé d'une banque d'ADN génomique. De nouveaux membres de la famille ont aussi été recherchés par le séquençage d'autres produits de PCR.

#### **Etude de l'expression**

Plusieurs axes de recherche ont été suivis dans le but de rechercher l'expression de gènes de la famille ETS :

- par rétrotranscription suivie de PCR, qui nous a permis d'isoler un deuxième membre de la famille ETS.
- par "Northern blot" et hybridation *in situ* afin de rechercher l'expression des deux gènes identifiés.
- par criblage d'une banque d'ADN complémentaire.

### **I - Mise en évidence d'homologues de la séquence v-ets chez la néréis**

Deux types de sondes ont été utilisées pour la recherche de séquences homologues des gènes de la famille ETS :

- un fragment *Pst*I de 423 pb (sonde 0,4 kpb-*Pst* I) de v-ets (nt 1504-1927, selon Nunn *et al.*, 1983) qui a permis d'identifier les gènes c-ets-1 et c-ets-2 de poulet (Leprince *et al.*, 1983 ; Boulukos *et al.*, 1988),
- un fragment *Eco*RI de 1203 pb (sonde 1,2 kpb-*Eco*RI) correspondant à l'ADNc de la protéine p54c-ets-1 du poulet (pCkE54A, Duterque-Coquillaud *et al.*, 1988).

Afin de mettre en évidence une séquence homologue des gènes *ets*, l'ADN génomique de *Nereis diversicolor* a été digéré par plusieurs enzymes de restriction (Figure 23). Après digestion par l'enzyme de restriction *Hind*III,

**Figure 23 : Electrophorèse des produits de digestion de l'ADN génomique de *Nereis diversicolor* par les enzymes de restriction *Hind*III (a), *Hind*III/*Pst*I (b), *Eco*RI (c), *Eco*RI/*Hind*III (d) et *Pst*I (e).  
f : Marqueur de poids moléculaire (ADN du phage  $\lambda$  digéré par les enzymes *Eco*RI et *Hind*III.**

**Figure 24 : Autoradiographie après hybridation avec la sonde 0,4 kpb-*Pst*I du transfert de la migration électrophorétique des produits de digestion de l'ADN génomique par les enzymes *Hind*III (a), *Hind*III/*Pst*I (b) et *Eco*RI (c).**

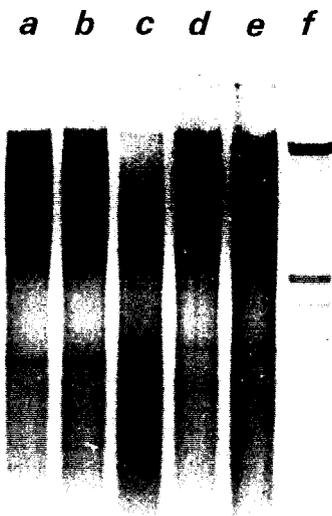


Figure 23

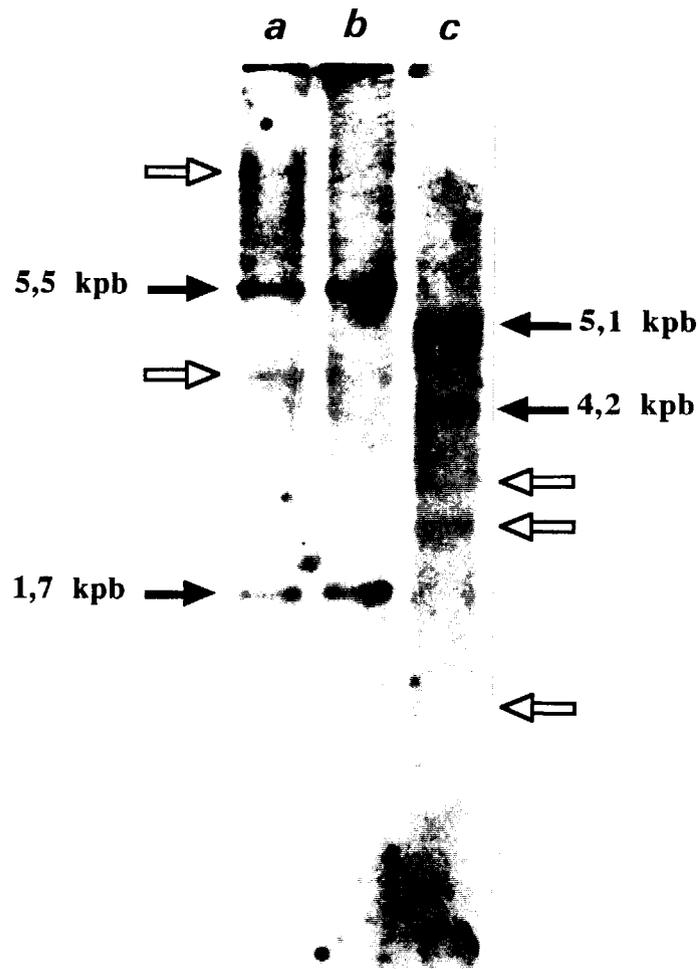


Figure 24

1112

l'hybridation en condition de moyenne stringence avec le fragment 0,4 kpb-*Pst*I a permis la détection de signaux d'environ 1,7 et 5,5 kpb (Figure 24, pistes a et b). Les mêmes hybridations ont été observées avec le fragment 1,2 kpb-*Eco*RI. La digestion par l'enzyme *Eco*RI révèle deux signaux autoradiographiques à 5,1 et 4,2 kpb, mais ces signaux semblent plus faibles (Figure 24, piste c). La double digestion *Bam*HI/*Hind*III a produit le même fragment de 1,7 kpb que précédemment. Les autres produits de digestion testés (*Pst*I, *Bam*HI, *Eco*RI/*Hind*III, *Bam*HI/*Eco*RI, *Bam*HI/*Pst*I et *Eco*RI/*Pst*I) n'ont pas permis de mettre en évidence de signal distinct.

La sonde 0,4 kpb-*Pst*I est identique à l'extrémité 3' de la sonde 1,2 kpb-*Eco*RI sauf pour deux mutations ponctuelles. Les conditions d'hybridation et de lavages ont du être adaptées avant d'obtenir des signaux autoradiographiques. Les signaux les plus intenses ont été obtenus pour des fragments *Hind*III de 1,7 et 5,5 kpb, avec les deux types de sondes. Le domaine ETS est la région la mieux conservée parmi les différents membres de la famille ETS et au cours de l'évolution. Les deux sondes utilisées se recouvrant notamment au niveau du domaine ETS, les hybridations correspondent vraisemblablement à des séquences homologues de ce domaine. D'autres signaux autoradiographiques moins intenses ont aussi été obtenus (Figure 24, flèches blanches) et peuvent correspondre à des séquences moins homologues de nos sondes. Connaissant actuellement la diversité de la famille ETS, on peut penser que ces hybridations correspondent à différents membres de cette famille plus ou moins divergents par rapport aux *c-ets-1*.

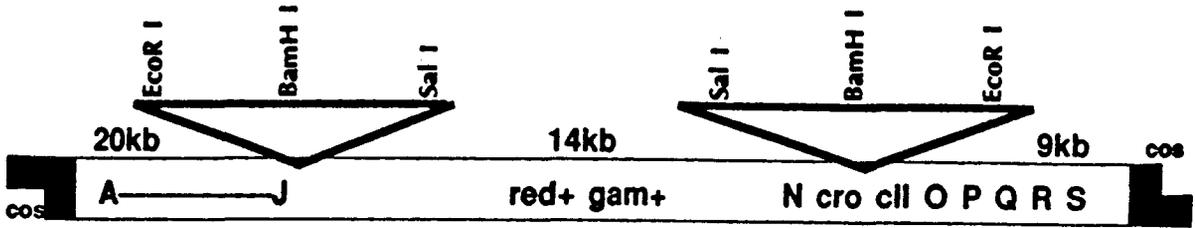
## **II - Recherche des gènes de la famille ETS dans des banques d'ADN génomique**

### **A - Construction et criblage de la banque $\lambda$ EMBL4**

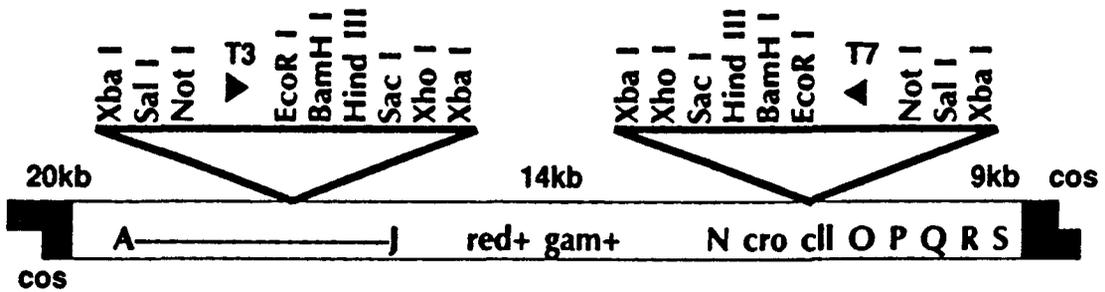
Une première banque d'ADN génomique de *Nereis diversicolor* a été construite en collaboration avec Agnès BÈGUE (Laboratoire d'Oncologie Moléculaire, Institut Pasteur de Lille) afin d'isoler un homologue des gènes de la famille ETS.

Après digestion de l'ADN du phage  $\lambda$  EMBL4 (Figure 25) par l'enzyme de restriction *Bam*HI, un insert de 14 kpb est éliminé et les bras du phage, de 9 et 20 kpb, sont récupérés en gradient de sucrose.

L'ADN génomique natif purifié et testé en électrophorèse migre aux environs de 50 kpb et ne présente aucune dégradation. Une cinétique de digestion par l'enzyme *Sau*3A sur une aliquote permet de définir un temps de digestion et



Phage  $\lambda$  EMBL4



Phage  $\lambda$  DASHII



Figure 25 : Carte physique des phages  $\lambda$  EMBL4 et  $\lambda$  DASHII

une quantité d'enzyme nécessaires à l'obtention du maximum de fragments de tailles comprises entre 10 et 20 kpb, qui seront les seuls utilisés pour la préparation de phages recombinés. Ces fragments ont des extrémités cohésives *Sau3A* compatibles avec celles obtenues après coupure par l'enzyme *BamHI*, ce qui permet leur ligation aux bras du phage.

Ne connaissant pas la taille du génome de la néréis, il est difficile d'évaluer le nombre de clones indépendants à obtenir pour que la banque soit représentative, c'est-à-dire pour qu'il y ait une chance raisonnable de trouver une séquence donnée.

Ce nombre est habituellement donné par la formule :

$$N = \frac{\text{Log}_{10}(1-P)}{\text{Log}_{10}(1-L/G)}$$

P = probabilité

G = taille du génome cloné

L = taille moyenne des fragments clonés

N = nombre de clones indépendants

Après encapsidation, le titre de la banque est de 50 000 phages. Trois génothèques ont été réalisées et nous avons obtenu trois fois environ 50 000 phages. Pour une meilleure représentativité du génome de la néréis, les trois banques ont été rassemblées. Une amplification dans la souche bactérienne DP 50 permet d'augmenter le titre à  $8.10^7$  phages/ml.

Deux séries d'étalements de 600 000 phages ont été réalisées. Les répliques sur des membranes de nitrocellulose hybridées avec le fragment 0,4 kpb-*PstI* n'ont présenté qu'un seul signal autoradiographique après un temps court d'exposition. Ce signal s'est avéré être aspécifique au cours des étapes de purification.

De nouveaux criblages, effectués à partir de l'étalement d'un million de phages selon le titre initial, n'ont rien révélé. Plus aucune confluence n'était obtenue. Après un nouveau titrage de la banque, nous nous sommes aperçus que le titre avait anormalement chuté en quelques semaines, passant de  $8.10^7$  phages/ml à  $4.10^6$  phages/ml.

Cette banque n'a pu être utilisée de nouveau, son titre ne cessait de chuter.

Le faible nombre de clones obtenus peut être expliqué par le fait que l'ADN génomique de néréis possède des propriétés particulières (par exemple qu'il soit fortement méthylé), le rendant difficilement clonable et instable dans ce type de vecteur et dans la souche bactérienne utilisée.

**Figure 26 : Electrophorèse des produits de digestion de l'insert 3,5 kpb-*Hind*III par les enzymes *Sau*3A (a), *Hae*III (b), *Taq*I (c), *Hinf*I (d) et *Xba*I (e).**

**Figure 27 : Electrophorèse des produits de digestion de l'insert 3,5 kpb-*Hind*III par les enzymes de restriction *Eco*RI (a), *Pst*I (b), *Pst*I/*Eco*RI (c), *Msp*I/*Apa*I (d), *Msp*I (e), *Msp*I/*Ava*I (f) et *Sau*3A/*Hae*III(g).**

**Figure 28 : Electrophorèse des produits d'amplification par PCR de l'ADN génomique de *Nereis diversicolor* avec les oligonucléotides D-OL1 et D-OL2 (a et c).  
b : Marqueur de poids moléculaire pBR322 digéré par l'enzyme de restriction *Hinf*I.**

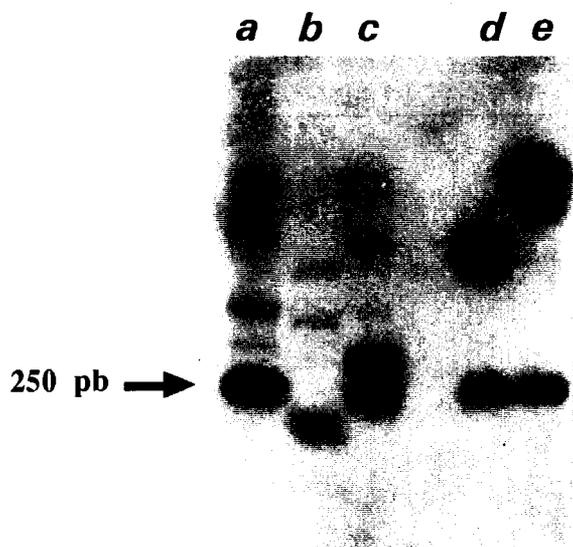


Figure 26

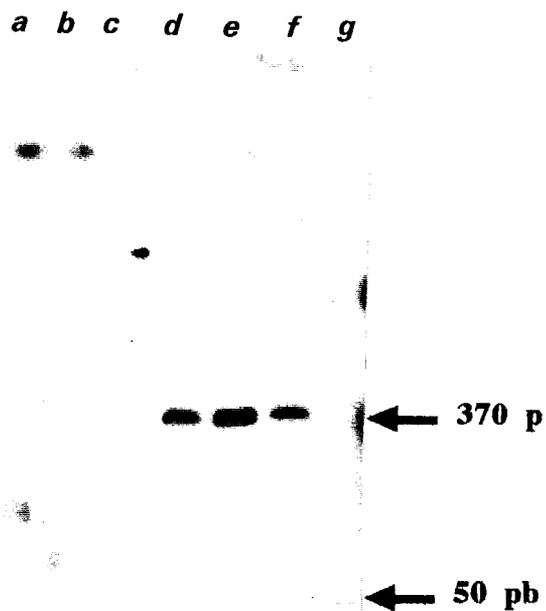


Figure 27

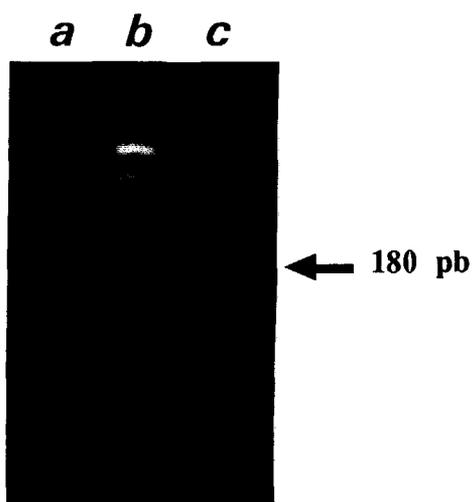


Figure 28



D'autre part, le fait de ne pas avoir trouvé de séquences homologues des gènes *ets* peut être attribué au fait que la banque  $\lambda$  EMBL4 était instable et n'était pas représentative de la totalité du génome de la néréis.

Cette hypothèse a été renforcée par la suite par la construction d'une banque, dans un autre vecteur et une autre souche bactérienne, ayant un titre initial de  $5.10^5$  phages (cf. IV).

### **B - Construction et criblage de la banque $\lambda$ NM1149**

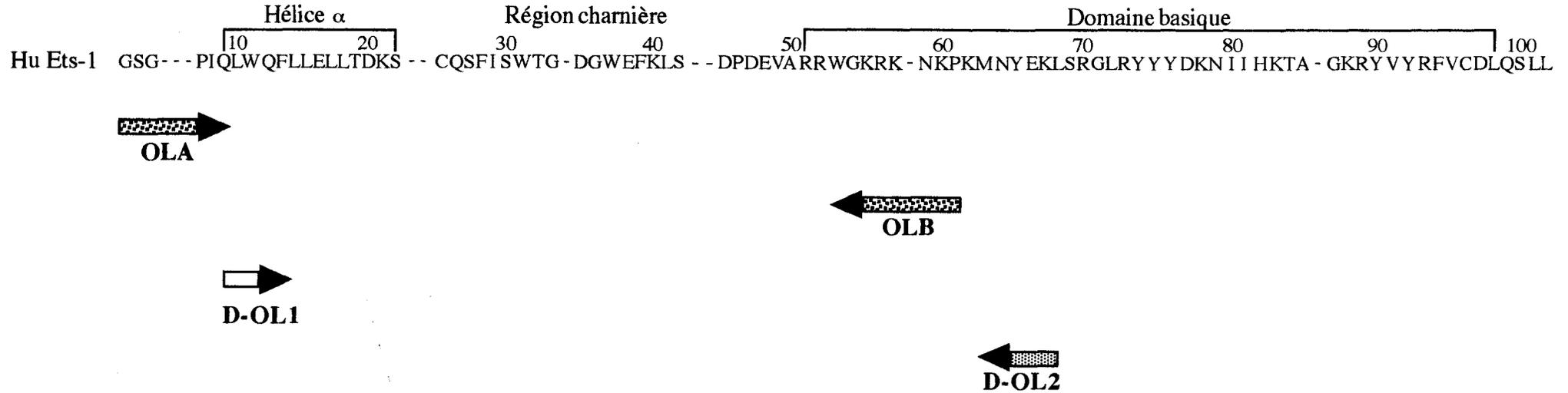
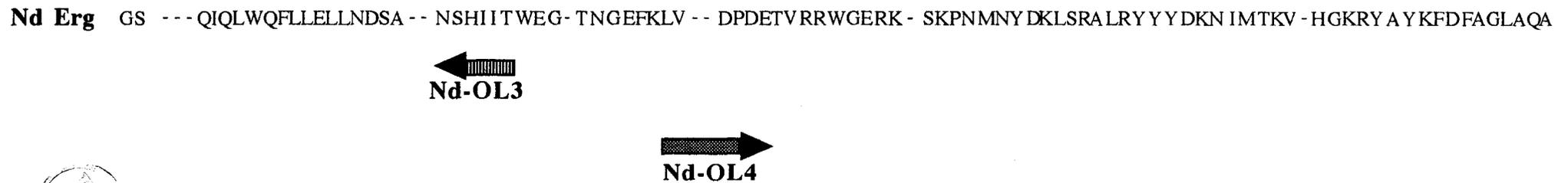
Une banque d'ADN génomique a été construite, après digestion totale de l'ADN par l'enzyme de restriction *HindIII*, dans le vecteur phagique  $\lambda$  NM1149. Ces choix ont été faits, d'une part après observation des signaux d'hybridation en "Southern blot", le vecteur  $\lambda$  NM1149 permettant le clonage de fragments *HindIII*, et d'autre part dans le but de modifier certains paramètres, espérant ainsi augmenter la stabilité de la banque. Le titre initial de la banque est de  $7.10^5$  phages. Après amplification, le titre est de  $8.10^8$  phages/ml.

Six millions de phages ont été étalés, transférés sur des membranes de nitrocellulose qui ont été hybridées avec la sonde 0,4 kpb-*PstI* en condition de moyenne stringence, avant de repérer un signal autoradiographique. Trois étapes de purification ont permis d'isoler le phage contenant une séquence homologue de notre sonde. La digestion de l'ADN phagique par l'enzyme *HindIII* a permis d'isoler un fragment de 3,5 kpb. Ce fragment a ensuite été cloné dans un vecteur plasmidique pUC 18.

Des digestions enzymatiques du fragment de 3,5 kpb ont été effectuées par différentes enzymes de restriction. L'hybridation avec la sonde 0,4 kpb-*PstI* virale a révélé différents fragments (Figures 26 et 27).

Le fragment obtenu après digestion par l'enzyme *Sau3A* a une taille de 250 pb et est le seul possédant une séquence homologue de la sonde virale (Figure 26, piste a).

Cependant, il existe deux fragments *Sau3A* de la même taille d'environ 250 bp, dont un seul porte l'homologie recherchée. Ne pouvant les séparer par électrophorèse, le clonage de ces deux inserts a été réalisé dans un vecteur plasmidique. Environ une centaine de mini-préparations d'ADN plasmidiques recombinés ont été réalisées, pour isoler celles contenant le fragment homologue de notre sonde. Ces préparations ont été digérées pour ressortir l'insert. Après transfert de la migration électrophorétique, les filters ont été hybridés avec la sonde 0,4 kpb-*PstI*. Aucun de ces inserts ne s'est hybridé avec la sonde. Le fait que le bon fragment n'ait jamais pu être sous-cloné semble vraisemblablement dû à une structure secondaire particulière.

**A****B**

**Figure 29 : Localisation des différents oligonucléotides dégénérés par rapport au domaine ETS de la séquence Hu Ets-1 (A) et des oligonucléotides spécifiques par rapport à la séquence Nd Erg (B).**

Un fragment *MspI* de 370 pb portant une partie de l'homologie, l'autre partie se situant sur un fragment de 50 pb (Figure 27, pistes d, e et f), a été cloné à la fois dans un vecteur plasmidique pUC 18 et dans l'ADN double brin du phage M13. Le séquençage de cet insert a été effectué, puis la séquence comparée à celle du *v-ets*. L'homologie montrée n'était pas celle attendue au niveau de la zone correspondant au domaine ETS, qui est la mieux conservée. Cependant, l'analyse de la séquence a montré une homologie en nucléotides de 40 % portant sur 170 pb. Cette homologie semble suffisante pour expliquer l'hybridation en condition de moyenne stringence.

### **III - Recherche des membres de la famille ETS par réaction de polymérisation en chaîne (PCR)**

Ne parvenant pas à isoler un membre de la famille ETS par les techniques classiques de criblage de banque, une autre stratégie, utilisant la PCR, a été envisagée. Karim *et al.* (1990) avaient défini le domaine ETS en alignant les séquences des protéines suivantes : PU.1 et Ets-1 murines ; Ets-2, Erg et Elk humaines ; E74 et Ets-2 de drosophile. Le domaine ETS était défini comme une séquence d'environ 85 acides aminés avec 22 acides aminés identiques dont trois résidus tryptophane.

Les seules séquences connues chez les Invertébrés concernaient l'oursin (Su Ets-2) et la drosophile (D Ets-2). Notre objectif étant alors de rechercher une séquence proche de ces gènes *ets-2* d'Invertébrés, nous avons donc aligné les séquences en acides aminés des domaines ETS des différentes protéines Ets-2 connues : celles de l'homme, du poulet, de la souris, de l'oursin, de la drosophile, ainsi que les deux protéines Ets-2 du xénope. La recherche des acides aminés les mieux conservés et impliquant le moins de dégénérescence du code génétique, nous a permis de définir les séquences nucléotidiques dénommées oligonucléotide A (OLA) et oligonucléotide B (OLB) (Figure 29A). La région pouvant être amplifiée par ces oligonucléotides comprend l'hélice  $\alpha$  et la région charnière du domaine ETS.

Ces oligonucléotides ont été testés sur des préparations d'ADN génomique par PCR. La modification de différents paramètres concernant le cycle d'amplification, la quantité d'ADN et les préparations d'ADN, n'a pas permis de mettre en évidence de produits d'amplification, même après un deuxième cycle d'amplification.

Chen *et al.* (1992), utilisant la PCR, ont isolé cinq membres de la famille ETS chez la drosophile. Les oligonucléotides utilisés avaient été conçus après alignement de tous les domaines ETS alors connus : ceux des protéines Ets-1, Ets-2, Erg et Elk humaines ; des protéines Spi-1/PU.1, Fli-1 et GABP $\alpha$  de souris et de la protéine E74 de drosophile. Les auteurs ont observé que tous les domaines ETS, sauf celui de Spi-1/PU.1, présentaient deux régions hautement conservées : LW(E/Q)F (acides aminés 10 à 13) et MNY(D/E)K (acides aminés 63 à 87). Tous les codons possibles ont été utilisés pour chaque acide aminé à l'exception de la leucine pour laquelle deux codons (CTG et CTC) sur les six possibles ont été utilisés. Les oligonucléotides ainsi définis sont appelés D-OL1 et D-OL2 et permettent d'amplifier une partie de la région en hélice  $\alpha$ , la région charnière et une petite partie du domaine basique du domaine ETS (Figure 29A). Ils ont été synthétisés en leur ajoutant un site de restriction *Xba*I à leur extrémité 5', permettant ainsi un clonage ultérieur. La taille du produit d'amplification attendu est de 180 pb, sauf s'il existe un intron comme pour le D *ets-2* (Pribyl *et al.*, 1988).

Avec ces oligonucléotides, des PCR ont été réalisées sur différentes préparations d'ADN génomique extrait d'animaux entiers. Les premiers essais ont été effectués selon le cycle suivant : hybridation à 52°C et allongement à 72°C, pendant 40 cycles. Ne visualisant rien après migration électrophorétique des produits d'amplification, un deuxième cycle d'amplification a été réalisé selon les mêmes paramètres.

Après migration électrophorétique des produits d'amplification dans un gel d'agarose à 2 %, une bande est obtenue à environ 180 pb (Figure 28). Ces produits de PCR ont ensuite été dialysés sur une membrane de microdialyse puis digérés par l'enzyme de restriction *Xba*I avant d'être clonés dans le vecteur plasmidique pUC 18. Après transformation bactérienne, plusieurs colonies ont été sélectionnées. Une préparation de plasmide a été réalisée dans le but d'établir la séquence du fragment cloné. Une PCR avec les oligonucléotides D-OL1 et D-OL2 sur la préparation d'ADN plasmidique a permis de vérifier le clonage. Le séquençage a été réalisé et la séquence traduite en acides aminés comparée à celles des différents membres de la famille ETS. La recherche effectuée sur la séquence protéique a montré l'homologie la plus grande (minimum 89 %) avec les séquences des protéines Hu Erg, Mo Fli-1, Su Erg et D Ets-3. La séquence nucléotidique obtenue ainsi que la séquence traduite en acides aminés sont les suivantes :

				D-OL1	CTT	CTT	GAA	CTG	CTA	TCC	GAT
					L	L	E	L	L	S	D
TCC	GCC	AAC	TCG	CAC	ATC	ATC	ACC	TGG	GAA	GGA	
S	A	N	S	H	I	I	T	W	E	G	
ACT	AAC	GGG	GAG	TTC	AAA	CTG	GTC	GAC	CCG	GAT	
T	N	G	E	F	K	L	V	D	P	D	
GAA	ACC	GTG	AGA	CGG	TGG	GGA	GAG	CGA	AAA	TCC	
E	T	V	R	R	W	G	E	R	K	S	
AAA	CCC	AAC		D-OL2							
K	P	N									

Cette séquence isolée de l'ADN génomique de *Nereis diversicolor* a été appelée Nd *erg* et a ensuite servi de sonde pour les autres expériences (sonde Nd *erg*)

#### IV - Recherche du gène correspondant à la sonde Nd *erg* dans une banque d'ADN génomique construite dans le vecteur $\lambda$ DASH II

##### A - Criblage de la banque $\lambda$ DASHII

Une nouvelle banque d'ADN génomique a été construite dans le vecteur phagique  $\lambda$  DASHII (Figure 25) après digestion partielle de l'ADN génomique par l'enzyme de restriction *Sau3A*.

Après encapsidation dans les lysats bactériens complémentaires, le titre est de  $5 \cdot 10^5$  phages. L'amplification dans la souche bactérienne SRB donne le titre final de  $6 \cdot 10^8$  phages/ml. Cette banque a été criblée avec la sonde Nd *erg*.

Après étalement d'un million de phages, 11 clones se sont avérés positifs. Afin de contrôler que ces clones possédaient bien la séquence recherchée, un oligonucléotide (Nd-OL3) a été synthétisé. Cet oligonucléotide correspond à une partie de la sonde Nd *erg* dont la séquence est moins bien conservée au cours de l'évolution, à l'intérieur du domaine ETS. La position de cet oligonucléotide par rapport à la séquence Nd *erg* est figurée par un ombrage (cf ci-dessus) et est présentée sur la figure 29B. A l'étape de la dernière purification des phages isolés positifs, une double réplique a été effectuée pour chaque étalement de phages : l'hybridation a été réalisée en condition de forte stringence, d'une part avec la sonde Nd *erg* et, d'autre part, avec l'oligonucléotide Nd-OL3. Les 11 phages

**Figure 30 : Electrophorèse des produits de digestion enzymatique par l'enzyme de restriction *EcoRI* de l'ADN de sept phages isolés de la génothèque  $\lambda$  DASHII, avec la sonde Nd *erg*. La piste a correspond à un type de phage ( $\lambda 2$ ), les pistes b, d, f et h correspondent à un autre type de phage ( $\lambda 1$ ), les pistes c et g correspondent à un troisième type de phage ( $\lambda 3$ ).**

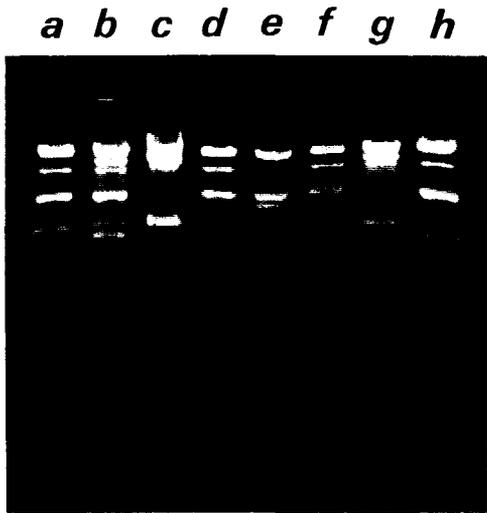
**Figure 31 : Révélation autoradiographique après hybridation avec la sonde Nd *erg* du transfert correspondant à la migration électrophorétique de la figure 30.**

**Figure 32 : Electrophorèse des produits de digestion des ADN phagiques des différents phages ( $\lambda 1$ ,  $\lambda 2$  et  $\lambda 3$ ) par les enzymes de restriction *EcoRI/HindIII*, *EcoRI/BamHI* et *BamHI/HindIII*.**

**a, b et c = phage  $\lambda 1$**

**d, e et f = phage  $\lambda 3$**

**i, j et k  
l, m et n** | = phage  $\lambda 2$



← 4,8 kpb →  
 ← 3,4 kpb →

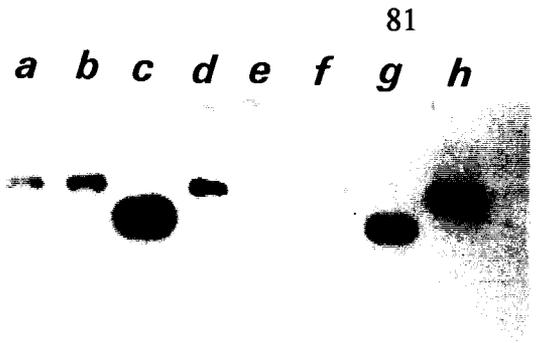


Figure 31

Figure 30

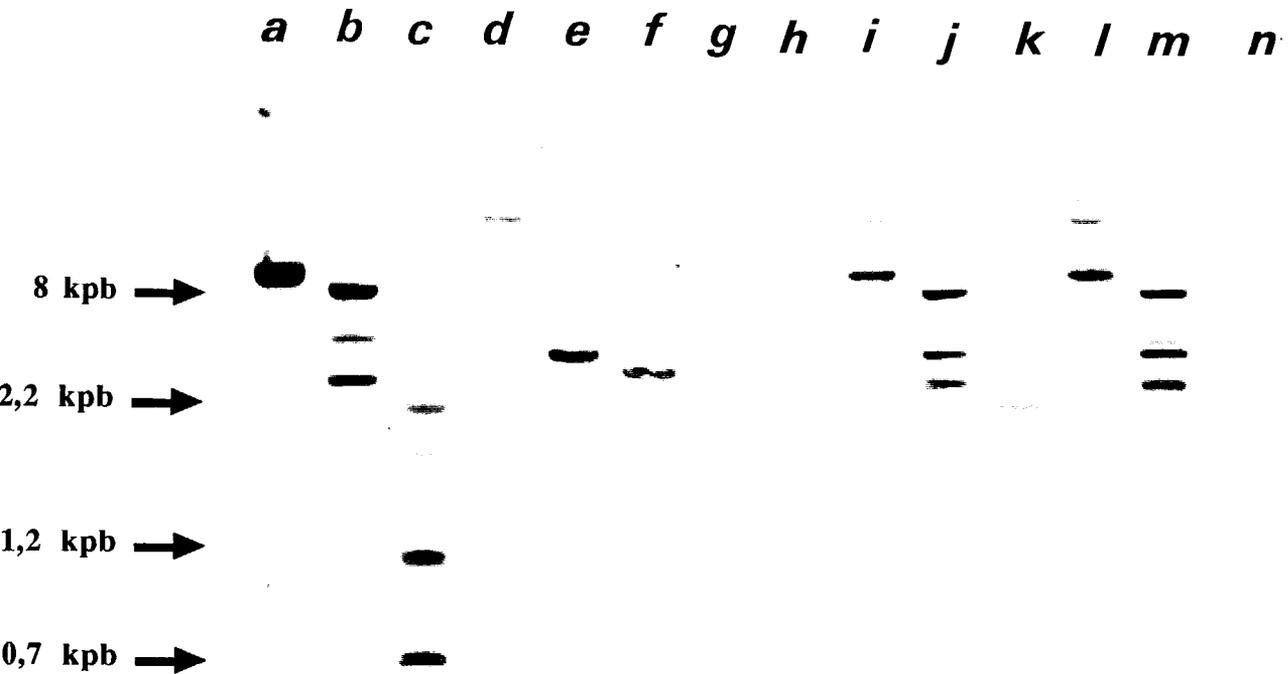
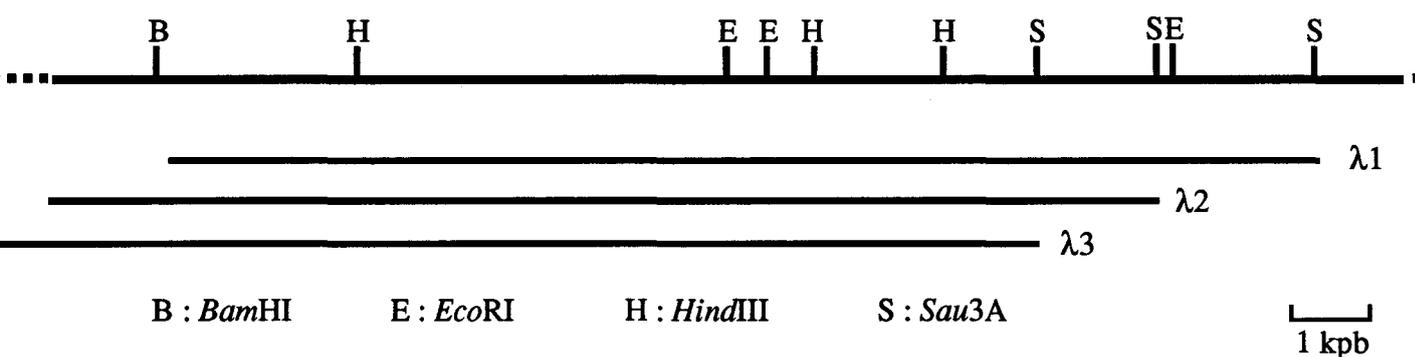


Figure 32

sélectionnés se sont hybridés à la fois avec la sonde oligonucléotidique Nd-OL3 et avec la sonde Nd *erg*. Sept de ces phages ont été isolés et leur ADN recombiné extrait.

### **B - Cartographie partielle des fragments insérés dans les phages**

Les différentes préparations d'ADN phagiques recombinés ont été digérées par l'enzyme de restriction *EcoRI*. En comparant les profils obtenus après migration électrophorétique, plusieurs préparations se sont avérées identiques (Figure 30). En définitive, sur les sept phages isolés, seulement trois possèdent un profil de restriction différent (phages  $\lambda 1$ ,  $\lambda 2$  et  $\lambda 3$ ). La migration électrophorétique de ces digestions enzymatiques a été transférée et la membrane mise en hybridation avec la sonde Nd *erg*. Pour ces phages, les fragments possédant la séquence homologue de notre sonde, sont soit un fragment *EcoRI* de 3,4 kpb, soit un fragment *EcoRI* de 4,8 kpb (Figure 31). Des doubles digestions combinant les enzymes de restriction *EcoRI*, *BamHI* et *HindIII* ont été effectuées pour établir une cartographie partielle des différents phages (Figure 32). L'analyse des résultats du transfert des différentes digestions, suivi de l'hybridation avec la sonde Nd *erg* a permis de proposer la carte suivante :



### **C - Clonage des fragments *EcoRI* possédant la séquence homologue de la sonde Nd *erg***

Les fragments *EcoRI* de 3,4 et 4,8 kpb ont été clonés dans le vecteur pUC 19 et préparés en grande quantité. Ces fragments ont ensuite été digérés par les enzymes de restriction suivantes : *HaeIII*, *HpaII*, *MspI*, *TaqI*, *HinfI*, *BamHI*, *PstI* et *HindIII*, dans le but de sous-cloner un fragment de plus petite taille possédant la séquence homologue de notre sonde. La migration électrophorétique est représentée sur la figure 33. Après transfert et hybridation avec la sonde Nd *erg*, les différents fragments présentant un signal autoradiographique sont

**Figure 33 : Electrophorèse des produits de digestion des fragments *EcoRI* de 3,4 kpb (pistes a à h) et 4,8 kpb (pistes j à p) par les enzymes de restriction *HaeIII* (a, j), *HpaII* (b, k), *MspI* (c, l), *TaqI* (d, m), *HinfI* (e, n), *BamHI* (f, o), *PstI* (g, p) et *HindIII* (h).  
i : Marqueur de taille  $\lambda$  *EcoRI/HindIII*.**

**Figure 34 : Révélation autoradiographique après hybridation du transfert correspondant à l'électrophorèse de la figure 33, avec la sonde Nd *erg*.**



Figure 33

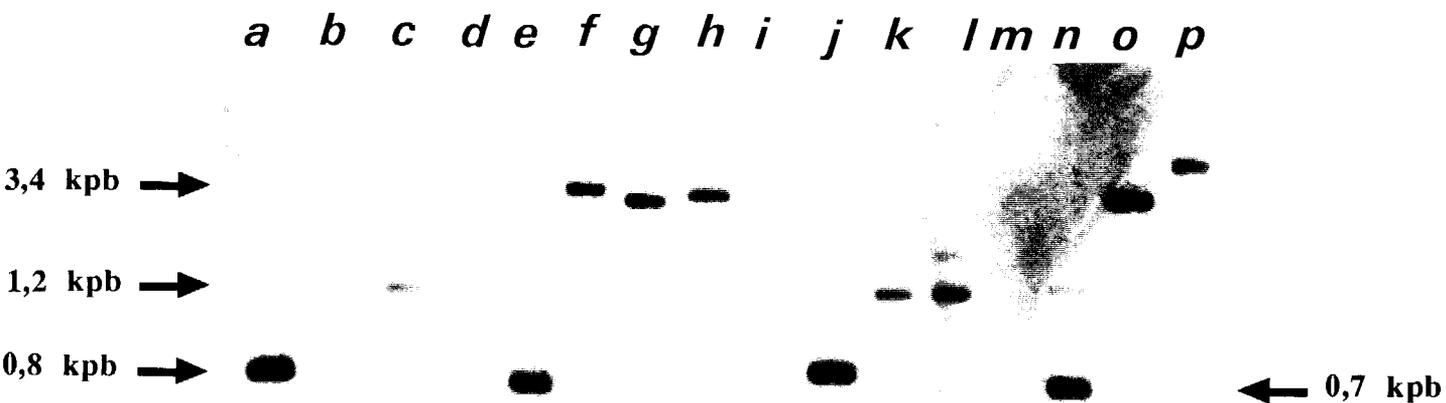


Figure 34

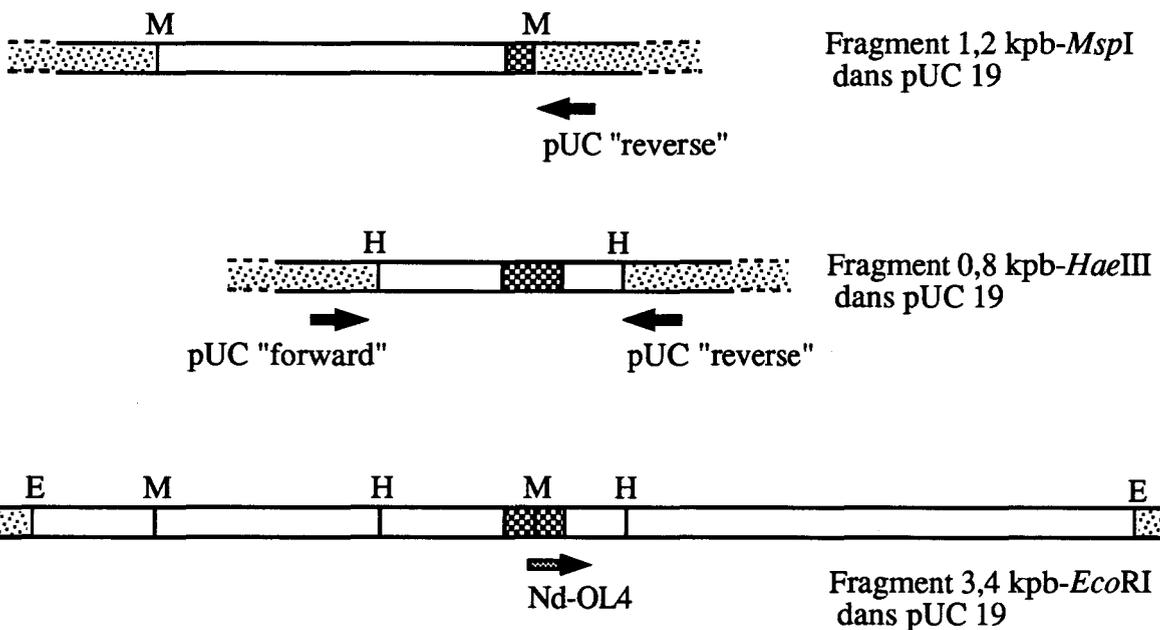
identiques pour les différentes digestions des fragments de 3,4 et 4,8 kpb (Figure 34). Ces deux inserts de 3,4 et 4,8 kpb portés par des phages différents, appartiennent donc au même gène et se recouvrent partiellement.

Dans le but de vérifier que les différents phages isolés avec la sonde Nd *erg* étaient bien spécifiques de cette sonde et donc pour confirmer les résultats précédemment obtenus avec l'oligonucléotide Nd-OL3, le fragment *EcoRI* de 3,4 kpb a été marqué et hybridé avec le transfert précédent. Tous les phages se sont hybridés avec la sonde *EcoRI*-3,4 kpb en condition de forte stringence. Nous sommes donc en présence de phages dont les séquences clonées se recouvrent en partie et appartiennent au même gène.

A partir des fragments de 3,4 et 4,8 kpb, deux inserts ont été sous-clonés dans le vecteur pUC 18 : un fragment de 1,2 kpb obtenu après digestion par l'enzyme *MspI* et un fragment de 0,8 kpb obtenu après digestion par l'enzyme *HaeIII* (Figure 34). Le clonage a été vérifié par hybridation de l'insert ressorti du vecteur avec la sonde Nd *erg* marquée par la technique à la digoxigénine.

#### D - Séquencage

La stratégie de séquencage est résumée dans le schéma ci-dessous :



▤ vecteur pUC 19

▨ sonde Nd *erg*

E = *EcoRI* ; M = *MspI* ; H = *HaeIII*

aaagaaaagaacaacatgctagaataaaaaattctatcatattcactatt

ttatttataatcattcttattttctgatttgca											GGG	AGT	GGA	CAA	ATC	CAG				
											G	S	G	Q	I	Q				
CTG	TGG	CAA	TTT	CTT	CTT	GAA	CTG	CTA	TCC	GAT										
L	W	Q	F	L	L	E	L	L	S	D										
TCC	GCC	AAC	TCG	CAC	ATC	ATC	ACC	TGG	GAA	GGA										
S	A	N	S	H	I	I	T	W	E	G										
ACT	AAC	GGG	GAG	TTC	AAA	CTG	GTC	GAC	CCG	GAT										
T	N	G	E	F	K	L	V	D	P	D										
GAA	AAC	GTG	AGA	CGG	TGG	GGA	GAG	CGA	AAA	TCC										
E	T	V	R	R	W	G	E	R	K	S										
AAA	CCC	AAC	ATG	AAC	TAC	GAC	AAA	CTG	TCC	CGA										
K	P	N	M	N	Y	D	K	L	S	R										
GCG	CTG	CGC	TAC	TAC	TAC	GAC	AAA	AAC	ATT	ATG										
A	L	R	Y	Y	Y	D	K	N	I	M										
ACG	AAG	GTG	CAC	GGC	AAA	CGT	TAC	GCT	TAC	AAG										
T	K	V	H	G	K	R	Y	A	Y	K										
TTT	GAC	TTT	GCT	GGG	TTG	GCC	CAG	GCC	ATG	CAG										
F	D	F	A	G	L	A	Q	A	M	Q										
CCC	TCT	ACC	ACA	GAC	CCA	GCG	GCC	TAC	AAA	TAC										
P	S	T	T	D	P	A	A	Y	K	Y										
CAG	ACA	AGA	CCT	ACT	CAT	GTC	TGG	GTA	CAT	CAC										
Q	T	R	P	T	H	V	W	V	H	H										
ACT	TCA	AAG	TTG	AAC																
T	S	K	L	N																

La séquence ombrée correspond à la sonde Nd *erg*.

└ : site accepteur d'épissage.

Figure 35 : Séquence nucléotidique de Nd *erg* et séquence prédite en acides aminés.

L'extrémité 5' du fragment, par rapport à la sonde Nd *erg*, a été obtenue par le séquençage du fragment *MspI* de 1,2 kpb avec l'amorce "reverse" du pUC.

Le séquençage du fragment *HaeIII* de 0,8 kpb avec les amorces "reverse" et "forward" du pUC, a permis de compléter la séquence Nd *erg* vers l'extrémité 3'.

Un oligonucléotide (Nd-OL4) correspondant à une partie de la séquence Nd *erg* a été synthétisé. Sa localisation et son orientation au sein de la séquence Nd *erg* sont figurées par un encadré (cf. page 80) et présentées sur la figure 29B. Utilisant le fragment 3,4 kpb-*EcoRI* comme matrice, cet oligonucléotide a permis de confirmer et de compléter la séquence de l'extrémité 3'.

La totalité de la séquence nucléotidique obtenue ainsi que la séquence prédite en acides aminés est présentée sur la figure 35.

## **V - Analyse des produits de l'amplification de l'ADN génomique par les oligonucléotides D-OL1 et D-OL2**

Parallèlement au criblage de la banque  $\lambda$  DASHII avec la sonde Nd *erg*, d'autres séquences présentant des homologies avec les membres de la famille ETS ont été recherchées dans les produits de PCR effectuée sur l'ADN génomique. Huit clones ont été repiqués et des préparations d'ADN plasmidiques ont été effectuées. Ces préparations ont servi de matrice à une PCR effectuée avec les oligonucléotides D-OL1 et D-OL2. L'analyse des produits de PCR sur un gel d'agarose à 3 % a montré une légère différence de taille entre les différents produits d'amplification (Figure 36). Les préparations e et i présentent une bande à environ 220 pb alors que les autres préparations présentent toutes une bande de même taille (environ 180 pb). Il n'a pas été obtenu d'amplification avec la préparation h. La piste a correspond à la sonde Nd *erg*.

Après migration électrophorétique, les produits de PCR ont été transférés sur une membrane de nylon chargé qui a été mise en hybridation en condition de forte stringence avec la sonde Nd *erg* marquée à la digoxygénine (Figure 37). La révélation immunologique nécessitée par la technique présente l'avantage de pouvoir définir un ordre d'apparition de la coloration en fonction du temps, permettant ainsi de le corréliser au degré d'homologie avec la sonde. Le but de cette manipulation était d'essayer de mettre en évidence des fragments plus ou moins homologues de la sonde et de les séquencer pour vérifier s'ils pouvaient correspondre à des membres de la famille ETS. L'échantillon i, qui est le témoin

**Figure 36 : Electrophorèse des produits d'amplification par PCR de préparations d'ADN plasmidiques recombinés avec un insert de 180 pb ou 220 pb.  
a = sonde Nd *erg*.**

**Figure 37 : Révélation immunologique après hybridation du transfert correspondant à l'électrophorèse de la figure 36, avec la sonde Nd *erg* marquée à la digoxigénine.  
L'ordre des pistes est inversé par rapport à la figure-36 (a = i; b = h...).**

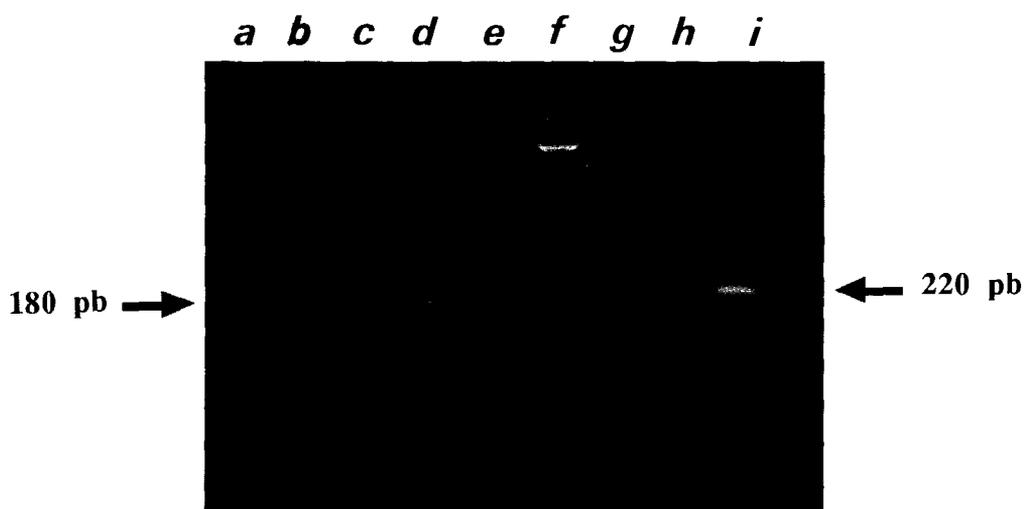


Figure 36

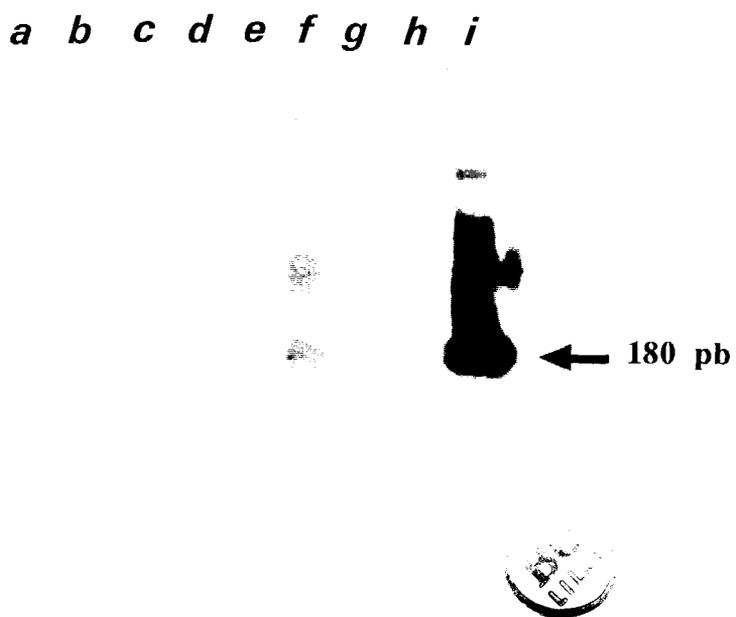


Figure 37

(sonde Nd *erg*) s'est révélé en premier, suivi du f, du a puis du c et au bout d'une nuit du e. Les échantillons b et g ne présentent pas de signal d'hybridation.

Ces différentes préparations ont été séquencées. Aucune des séquences n'a présenté d'homologie avec la sonde Nd *erg*, ni avec aucun des membres de la famille ETS. Par contre, dans chacun des cas, les oligonucléotides D-OL1 et D-OL2 ont été repérés. Les préparations f et c présentent le même insert.

L'amplification avec les oligonucléotides D-OL1 et D-OL2 sur l'ADN génomique produit donc des fragments de taille attendue pour une amplification de membres de la famille ETS, mais les inserts séquencés ne présentent aucune homologie avec les séquences *ets*.

## **VI - Recherche de fragments d'ADNc codant des membres de la famille ETS**

### **A - Recherche par la technique de rétrotranscription associée à la PCR (RT PCR)**

Dans le but d'isoler un ou plusieurs ADNc correspondant à des homologues de la famille ETS, des manipulations de RT PCR ont été entreprises. Ces manipulations ont été effectuées sur des ARN messagers isolés des différentes parties du corps de l'animal :

- le prostomium,
- le métastomium,
- des régénérats de pygidium coupés à des temps différents.

Disposant des oligonucléotides D-OL1 et D-OL2 permettant d'amplifier une partie du domaine ETS qui est le mieux conservé au cours de l'évolution et à travers les différents membres de la famille ETS, plusieurs expériences ont été réalisées dans le but d'isoler à la fois l'extrémité 3' et l'extrémité 5' des ADNc.

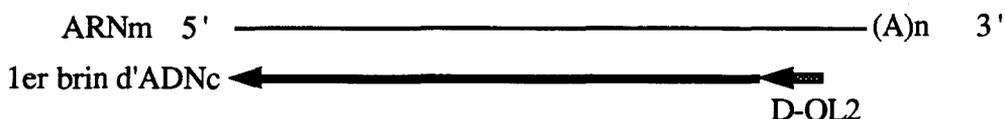
Avant d'entreprendre les manipulations sur les différentes parties du corps de l'animal, une manipulation test a été effectuée sur une préparation d'ARNm de régénérats coupés à différents temps.

Le premier brin d'ADN complémentaire a été synthétisé de trois façons :

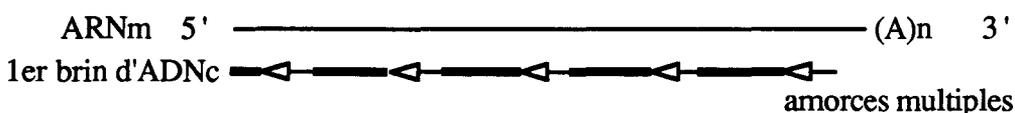
— en utilisant comme amorce un oligo(dT)18,



— en utilisant comme amorce l'oligonucléotide D-OL2,



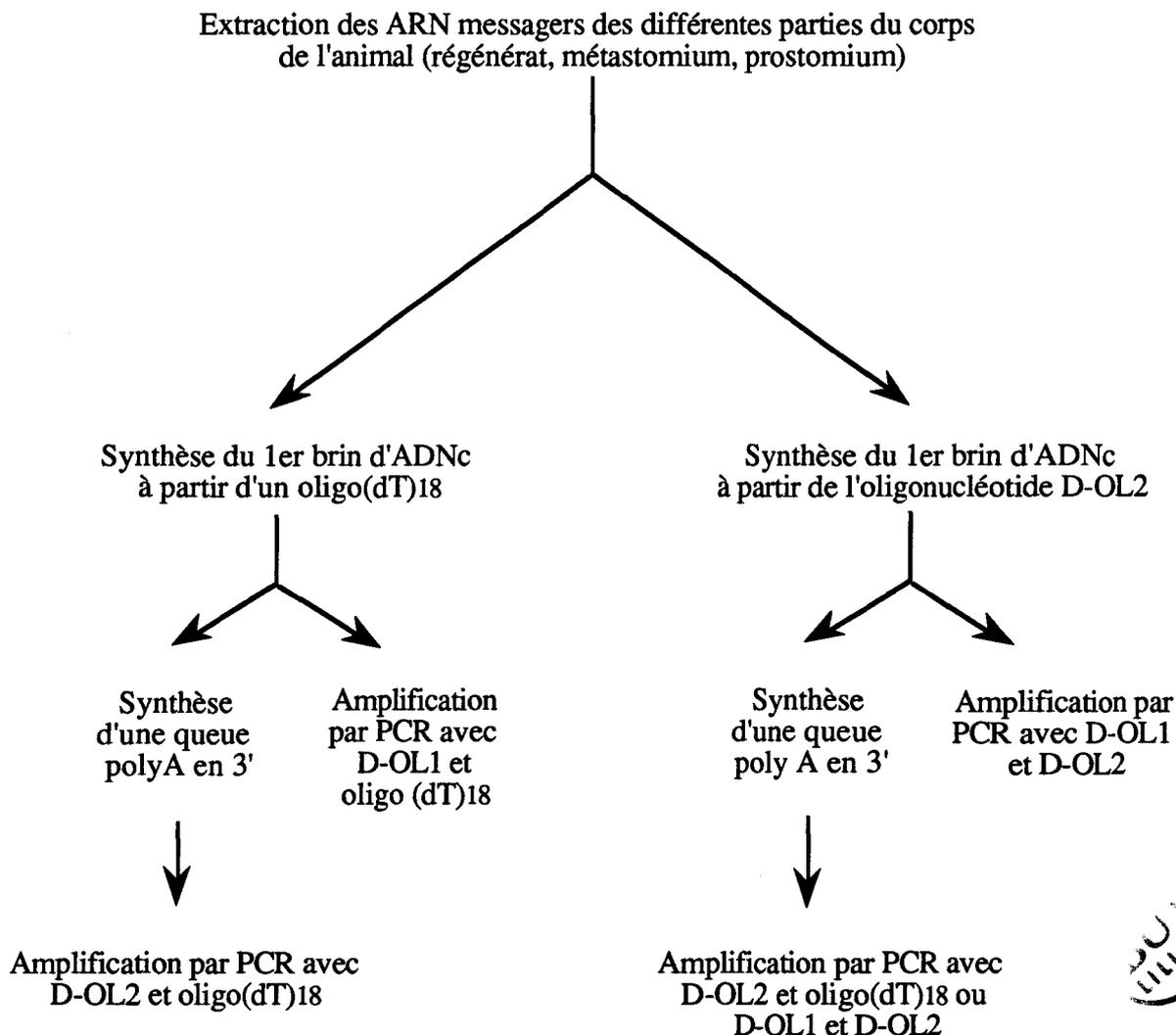
— en utilisant un multiamorçage au hasard,



Une fois la synthèse terminée, la partie ARN de l'hybride ADN/ARN est éliminée par action de l'ARNase H. Pour tester le bon fonctionnement de ces manipulations et afin de vérifier la présence d'ADN complémentaires correspondant aux membres de la famille ETS, une PCR utilisant comme amorces les oligonucléotides D-OL1 et D-OL2 a été effectuée sur les trois types de préparations d'ADNc. Un deuxième cycle d'amplification a été réalisé puis les produits obtenus ont été déposés sur un gel d'agarose à 2 % (Figure 39). L'amplification de fragments de bonne taille (180 pb) a eu lieu pour les préparations d'ADNc initiés avec l'amorce oligo(dT)18 et avec l'oligonucléotide D-OL2 (pistes a et b respectivement). Par contre, pour la préparation d'ADNc initiés par multiamorçage au hasard, aucune amplification n'est obtenue (piste c).

Pour la synthèse des ADNc à partir des ARNm extraits des différentes parties du corps de l'animal, le multiamorçage au hasard a été abandonné.

Les principales étapes suivies pour l'obtention des différentes extrémités des ADNc et leur amplification sont résumées dans le schéma suivant :

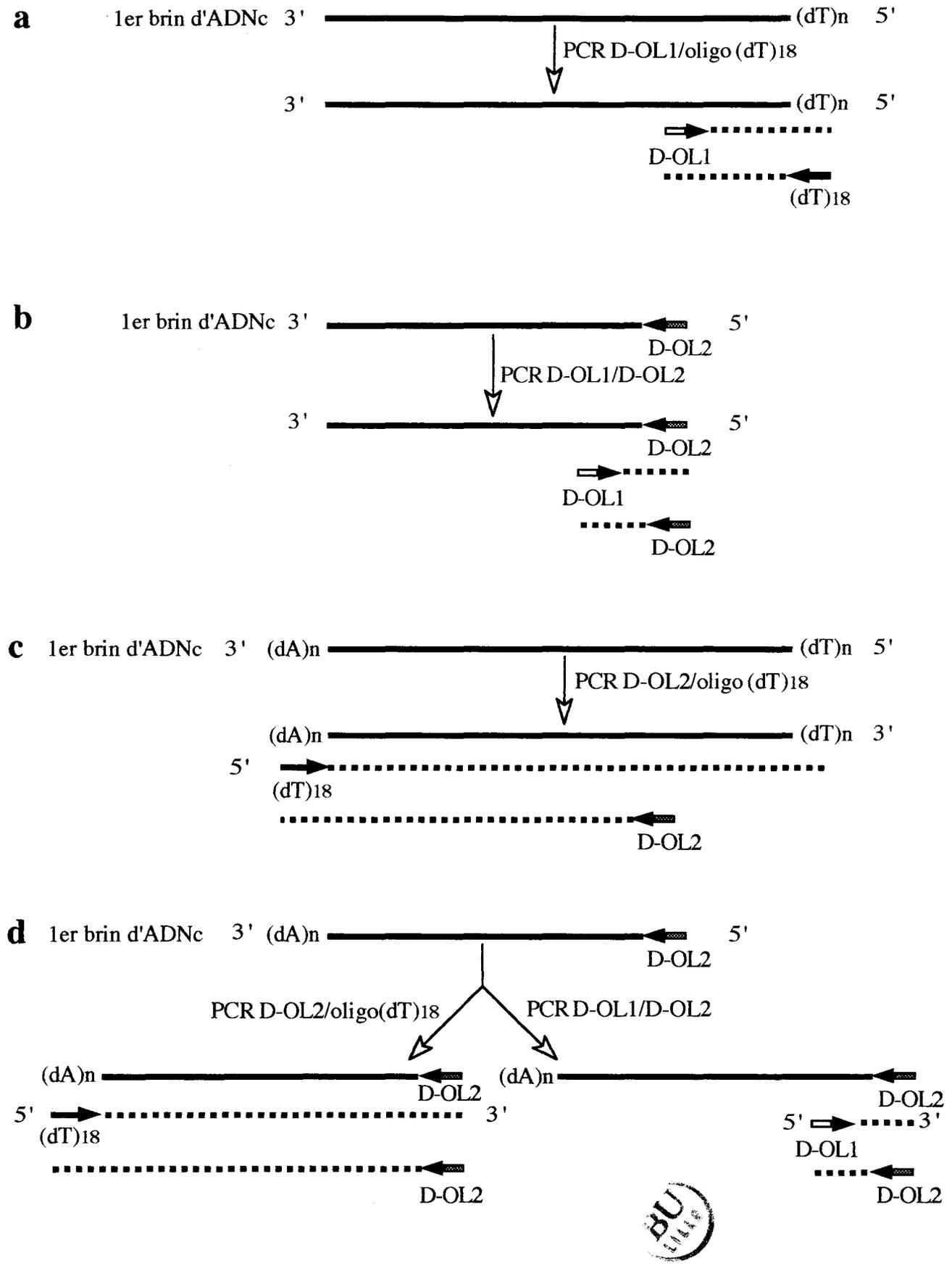


Ne sachant pas si le domaine de liaison à l'ADN, pour les ADNc recherchés, se trouve dans la partie 5' ou dans la partie 3' du gène, différentes manipulations ont été réalisées à partir des ARN messagers isolés des différentes parties du corps de l'animal.

Pour chaque préparation d'ARN messagers (prostomiums, métastomiums, régénérats), le premier brin d'ADN complémentaire est synthétisé, d'une part avec l'oligo(dT)18 et, d'autre part, avec l'oligonucléotide D-OL2.

A partir de là, le matériel a été séparé en deux parties pour chaque type d'ADNc.

— Sur une moitié du matériel, les ADN complémentaires ont servi directement de matrice dans des PCR de deux types, utilisant soit les oligonucléotides D-OL1 et oligo(dT)18, soit les oligonucléotides D-OL1 et D-OL2 (Figure 38, a et b).



**Figure 38 : Représentation schématique des différentes manipulations de RT PCR effectuées sur des préparations d'ARNm extraits des différentes parties du corps de l'animal.**

— Sur l'autre moitié du matériel, une queue poly A est ajoutée à l'extrémité 3' des ADN complémentaires par la terminale transférase. Ces ADNc ont servi de matrice dans des PCR utilisant soit les oligonucléotides D-OL2 et oligo(dT)<sub>18</sub>, soit les oligonucléotides D-OL1 et D-OL2 (Figure 38, c et d).

De toutes ces expériences, seules les PCR effectuées avec les oligonucléotides D-OL1/oligo(dT)<sub>18</sub> et les oligonucléotides D-OL1/D-OL2 ont donné des produits d'amplification. Les PCR utilisant les oligonucléotides D-OL2 et oligo(dT)<sub>18</sub> n'ont jamais donné de résultat d'amplification, même en faisant varier les paramètres des amplifications et la quantité de matériel.

Pour les amplifications utilisant les oligonucléotides D-OL1 et oligo(dT)<sub>18</sub>, visant à isoler l'extrémité 3' des ADNc, plusieurs fragments ayant des tailles de 200 à 600 pb sont visualisés après migration des produits d'amplification dans un gel d'agarose à 2 % (Figure 40).

Pour les amplifications utilisant les oligonucléotides D-OL1 et D-OL2, deux bandes d'environ 180 et 220 pb sont obtenues après migration des produits de PCR dans un gel d'agarose à 3 %, quel que soit le type d'ADNc utilisé comme matrice (Figure 41).

Ces résultats montrent qu'il y a bien une expression de gènes de la famille ETS puisqu'une amplification de fragments de taille correspondant à celle attendue est obtenue avec les oligonucléotides situés dans le domaine ETS et que cette expression semble avoir lieu dans toutes les parties du corps de l'animal. Par contre, le fait de ne pas obtenir d'amplification lors de la réaction avec les oligonucléotides D-OL2 et oligo(dT)<sub>18</sub> reste inexplicable. Peut-être les conditions de PCR ne sont-elles pas adaptées malgré toutes les modifications de cycle qui ont été effectuées, le fragment pouvant être de grande taille.

## **B - Clonage des fragments d'ADNc et séquençage**

### **1 - Clonage des produits correspondant à l'amplification des ADNc avec les oligonucléotides D-OL1 et D-OL2**

Les oligonucléotides D-OL1 et D-OL2 possèdent à leur extrémité 5' un site de restriction pour l'enzyme *Xba*I.

Les deux fragments de 180 et 220 pb correspondant aux amplifications obtenues par RT PCR à partir des différentes parties du corps de l'animal ont été extraits du gel et dialysés sur une membrane de microdialyse

**Figure 39** : Electrophorèse des produits d'amplification par RT PCR d'ARNm extraits de régénérats, avec les oligonucléotides D-OL1 et D-OL2.

**a** : La synthèse du premier brin d'ADNc est initiée avec l'oligonucléotide oligo(dT)<sub>18</sub>.

**b** : La synthèse du premier brin d'ADNc est initiée avec l'oligonucléotide D-OL2.

**c** : La synthèse du premier brin d'ADNc est initiée par multiamorçage au hasard.

**d et f** : Fragments témoins de 180 pb.

**e** : Témoin blanc (la PCR est réalisée sans ADNc).

**Figure 40** : Electrophorèse des produits d'amplification par RT PCR d'ARNm extraits de régénérats (pistes a et d), de métamères (pistes b et e) et de prostomiums (pistes c et f), avec les oligonucléotides D-OL1 et oligo(dT)<sub>18</sub>.

**g et h** : Témoins de 180 pb.

**Figure 41** : Electrophorèse des produits d'amplification par RT PCR d'ARNm extraits de régénérats (piste a), de métamères (piste b) et de prostomiums (piste c), avec les oligonucléotides D-OL1 et D-OL2.

**d et e** : Témoins de 180 pb.

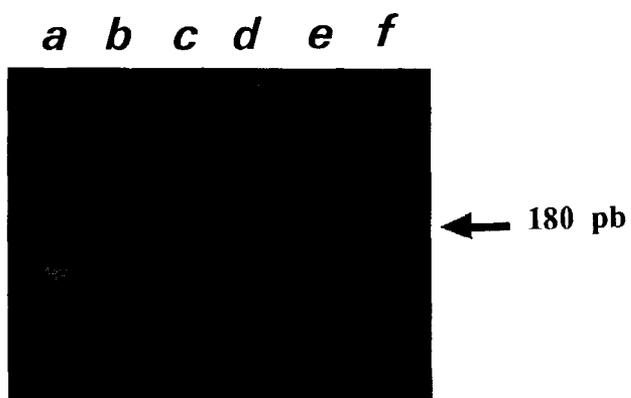


Figure 39

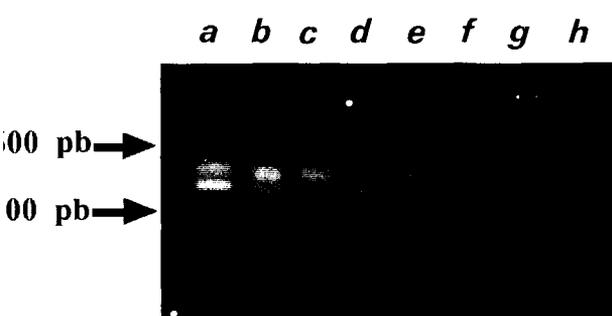


Figure 40

← 180 pb →

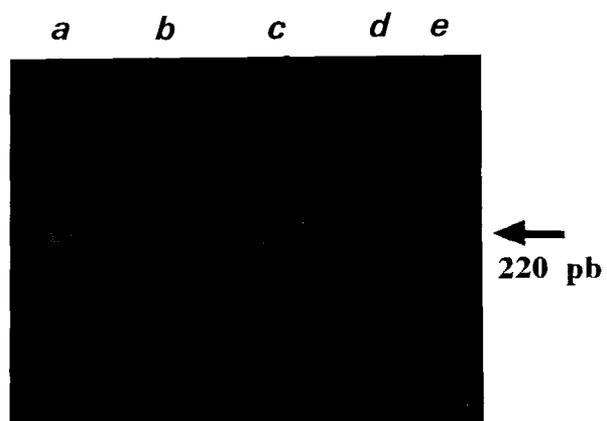


Figure 41



(Millipore). Une digestion enzymatique par l'enzyme *XbaI* a été réalisée, puis pour chaque type de préparation (régénérats, métamères, prostomiums), les fragments ont été clonés dans le site unique *XbaI* du pUC 19. Après transformation dans la souche bactérienne JM 83, de nombreuses colonies ont été obtenues.

Trois extractions de plasmides ont été effectuées pour chacun des deux clonages, soit trois préparations de plasmides contenant un insert de 180 pb et trois préparations contenant un insert de 220 pb pour chaque préparation initiale de régénérats, métamères et prostomiums.

Par PCR utilisant les oligonucléotides "reverse" et "forward" du pUC, la taille des fragments amplifiés a pu être vérifiée après migration électrophorétique (Figure 42).

Un transfert des produits de cette migration a été effectué, suivi d'une hybridation avec l'oligonucléotide Nd-OL3, cet oligonucléotide correspondant à une séquence de 18 nucléotides de la sonde Nd *erg*. Il a été choisi de façon à être spécifique de notre sonde, car cette séquence se situe dans la région charnière qui est moins bien conservée, au sein du domaine ETS. Cette manipulation a été réalisée dans le but de rechercher, par l'intermédiaire de l'oligonucléotide Nd-OL3, des séquences homologues de la sonde Nd *erg* et de montrer ainsi leur expression dans les différentes parties du corps de l'animal, mais aussi dans le but de rechercher d'autres membres de la famille ETS (fragments ne s'hybridant pas forcément voire faiblement avec l'oligonucléotide Nd-OL3).

L'autoradiographie est représentée sur la figure 43. Parmi les clones isolés, certains s'hybrident fortement avec l'oligonucléotide alors que d'autres ne s'hybrident pas, et cela pour les deux types de clonages. Ce résultat semblait conforter l'idée qu'il y a bien, dans les différentes parties du corps de l'animal, une expression de membres de la famille ETS dont certains seraient fortement homologues de la sonde Nd *erg*. Quelques-uns des clones isolés ont été séquencés en utilisant l'amorce universelle du pUC/M13 (pistes marquées par un astérisque) :

- les clones a et c correspondent à des produits d'amplification de 180 pb de régénérats, qui s'hybrident fortement avec l'oligonucléotide Nd-OL3,
- les clones d et e correspondent à des produits d'amplification de 220 pb de régénérats qui ne s'hybrident pas avec l'oligonucléotide Nd-OL3,
- les clones h et i correspondent à des produits d'amplification de 180 pb de métamères. Le clone h s'hybride fortement avec l'oligonucléotide Nd-OL3, le clone i ne s'hybride pas,

**Figure 42 : Electrophorèse des produits d'amplification des préparations de plasmides contenant un insert de 180 pb ou 220 pb.**

**a, b, c : Insert de 180 pb |**  
**d, e, f : Insert de 220 pb |** **provenant de la RT PCR**  
**sur des ARNm extraits de régénérats.**

**g : Témoin de 180 pb.**

**h, i, j : Insert de 180 pb |**  
**k, l, m : Insert de 220 pb |** **provenant de la RT PCR**  
**sur des ARNm extraits de métamères.**

**n, o, p : Insert de 180 pb |**  
**r, s, t : Insert de 220 pb |** **provenant de la RT PCR**  
**sur des ARNm extraits de prostomiums.**

**q : Témoin de 180 pb.**

**La migration électrophorétique correspondant aux pistes h à t a été légèrement décalée vers le haut.**

**Figure 43 : Autoradiographie correspondant au transfert de la migration électrophorétique de la figure 42, et hybridation avec l'oligonucléotide Nd-OL3.**  
**Les pistes marquées par un astérisque correspondent aux fragments qui ont été séquencés.**



Figure 42

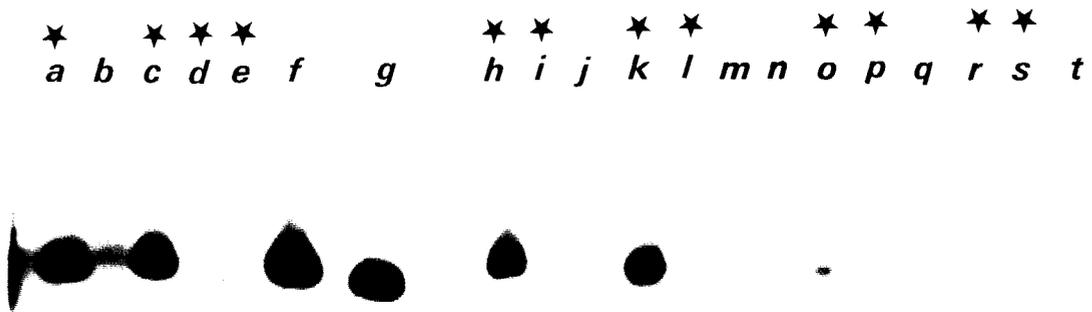


Figure 43



- les clones k et l correspondent à des produits d'amplification de 220 pb de métamères. Le clone k s'hybride fortement avec l'oligonucléotide Nd-OL3, le clone l ne s'hybride pas,
- les clones o et p correspondent à des produits d'amplification de 180 pb de prostomiums. Le clone o s'hybride fortement avec l'oligonucléotide Nd-OL3, le clone p ne s'hybride pas, et
- les clones r et s correspondent à des produits d'amplification de 220 pb de prostomiums, qui ne s'hybrident pas avec l'oligonucléotide Nd-OL3.

Aucun de ces fragments n'a montré d'homologie avec les séquences nucléotidiques des membres de la famille ETS. Pour les clones s'hybridant avec l'oligonucléotide Nd-OL3, des séquences proches de cet oligonucléotide ont été retrouvées, présentant au moins 50 % de nucléotides homologues.

## **2 - Clonage des produits correspondant à l'amplification des ADNc avec les oligonucléotides D-OL1 et oligo(dT)<sub>18</sub>**

L'amplification des ADN complémentaires des préparations de prostomiums, metastomiums et régénérats avec les oligonucléotides D-OL1 et oligo(dT)<sub>18</sub> a produit plusieurs fragments de tailles de 200 à 600 pb. Les produits d'amplification des différentes préparations ont été dialysés sur une membrane de microdialyse puis phosphorylés par la kinase du phage T4 et clonés dans le vecteur pUC 19 préalablement digéré par l'enzyme de restriction *Sma*I et déphosphorylé. Après transformation bactérienne, 15 colonies ont été sélectionnées pour chaque type de clonage (prostomiums, metastomiums, régénérats) et des mini-préparations de plasmides ont été effectuées. Par PCR avec les oligonucléotides "reverse" et "forward" du pUC, la taille des fragments amplifiés a pu être estimée (Figures 44 et 46). Des fragments de tailles très diverses ont été obtenus.

Un transfert de ces différents fragments amplifiés a été effectué puis la membrane mise en hybridation avec la sonde Nd *erg* marquée à la digoxigénine. Les lavages ont été réalisés en condition de forte stringence afin de pouvoir déceler des variations d'hybridation dues à une différence d'homologie. La révélation immunologique a montré le résultat présenté sur les figures 45 et 47.

Plusieurs fragments ont été séquencés (pistes marquées par un astérisque) : six fragments s'hybridant fortement avec la sonde (deux fragments issus de préparations de métamères, deux fragments issus de préparations de

**Figure 44A, 44B et 46C : Electrophorèse des produits d'amplification de préparations de plasmides issus du clonage des produits de RT PCR avec les oligonucléotides D-OL1 et oligo(dT)<sub>18</sub>.  
Différents fragments de tailles comprises entre 200 et 600 pb sont obtenus, pour les préparations de prostomiums, métastomiums et régénérats.**

**Figure 45A, 45B et 47C : Autoradiographie correspondant au transfert des migrations électrophorétiques des figures 44A, 44B et 46C, suivi de l'hybridation avec la sonde Nd *erg*.  
Les pistes marquées par un astérisque correspondent aux fragments qui ont été séquencés.**

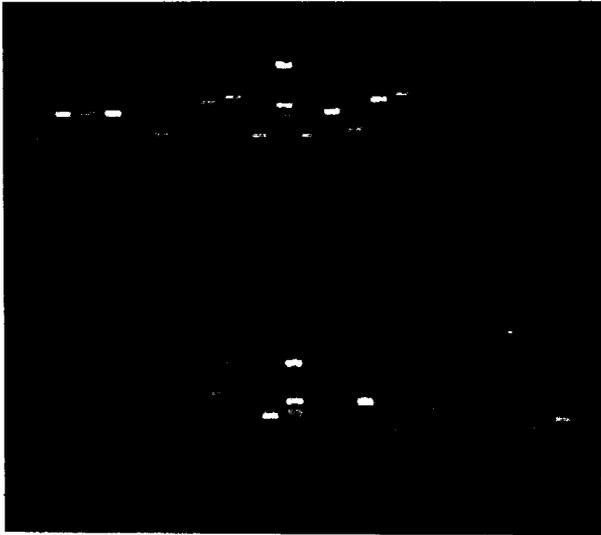


Figure 44



Figure 45

C

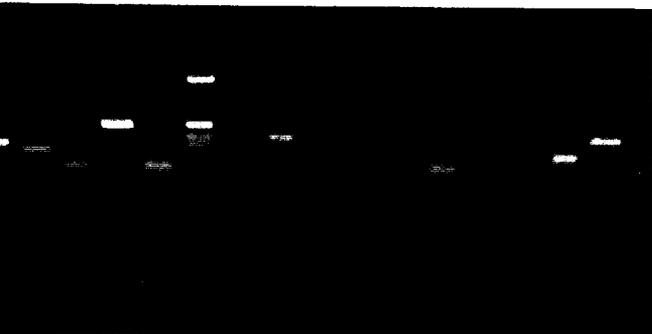


Figure 46



Figure 47



régénérats et deux fragments issus de préparations de prostomiums) et un fragment témoin ne s'hybridant quasiment pas, provenant d'une préparation d'ARNm de régénérats (cette faible intensité étant considérée comme du bruit de fond).

Pour les différents fragments s'hybridant fortement avec la sonde, aucune homologie n'a pu être trouvée avec les différents membres de la famille ETS. Mais, pour chacun des fragments l'oligonucléotide D-OL1 et parfois même l'oligonucléotide D-OL2 ont été retrouvés, ce qui permet d'expliquer les forts signaux d'hybridation. En fait, l'amplification par PCR serait aspécifique.

### Mise en évidence de l'expression du gène Nd *ets*

Par contre, le fragment de 400 pb considéré comme témoin négatif a montré, après comparaison de sa séquence prédite en acides aminés avec les différents domaines ETS, une homologie très importante avec les membres de la famille ETS. Le fragment présentant l'homologie a une taille de 140 pb et il semble s'être accolé, sûrement lors du clonage, avec un autre fragment n'ayant aucun rapport avec les séquences *ets*. L'homologie porte sur 40 acides aminés du domaine ETS. L'homologie la plus importante est trouvée avec les membres des sous groupes ETS-1 et ETS-2 de la famille ETS (85 % d'homologie). Les séquences nucléotidique et protéique de ce fragment sont les suivantes :

5'	CTC	CTG	GAG	CTG	CTG	ACC	GAC	AAG	AAC	TGT	
	L	L	E	L	L	T	D	K	N	C	
	CAG	CAC	TTC	ATC	TCG	TGG	TCC	GGG	GAC	GGC	
	Q	H	F	I	S	W	S	G	D	G	
	TGG	GAA	TTC	AAA	TTA	GCC	GAC	CCG	GAC	GAG	
	W	E	F	K	L	A	D	P	D	E	
	GTG	GCC	CGG	CGG	TGG	GGT	ATC	AGA	AAG	AAG	3'
	V	A	R	R	W	G	I	R	K	K	

Cette séquence obtenue à partir d'ARNm de régénérats de *Nereis diversicolor* a été appelée Nd *ets*.

**Figure 48 : "Northern" d'ARN messagers extraits de régénérats coupés à différents temps et hybridation avec la sonde Nd *erg*.  
Un signal à 2,5 kpb est visible sur la piste a.**

**Figure 49 : Traduction d'ARN messagers en lysat de réticulocytes de lapin.  
a : Traduction d'ARNm de sangsue (témoin).  
b : Traduction d'ARNm de néréis.**

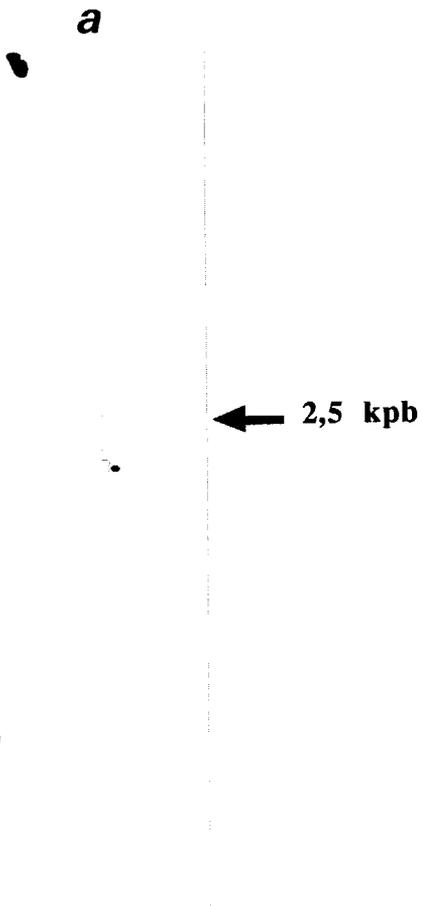


Figure 48

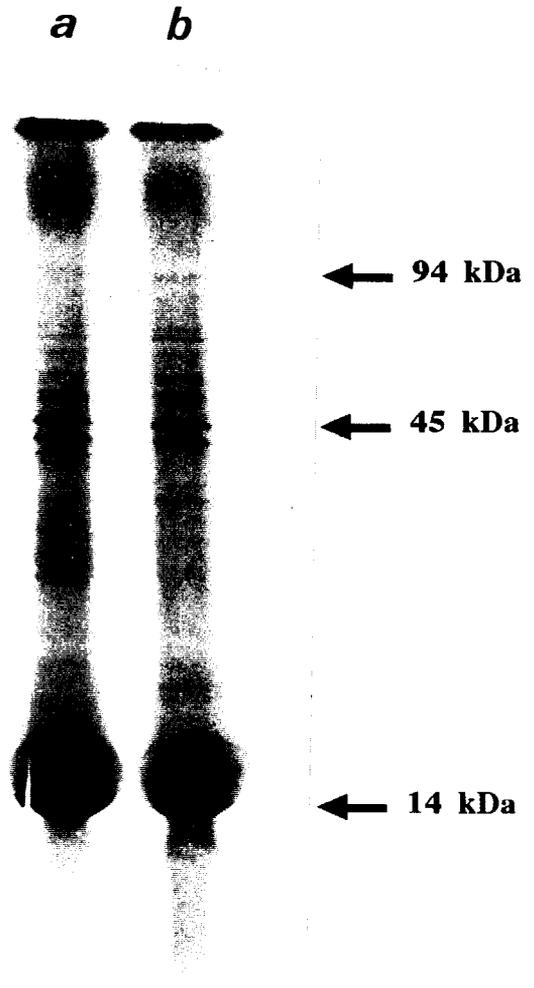
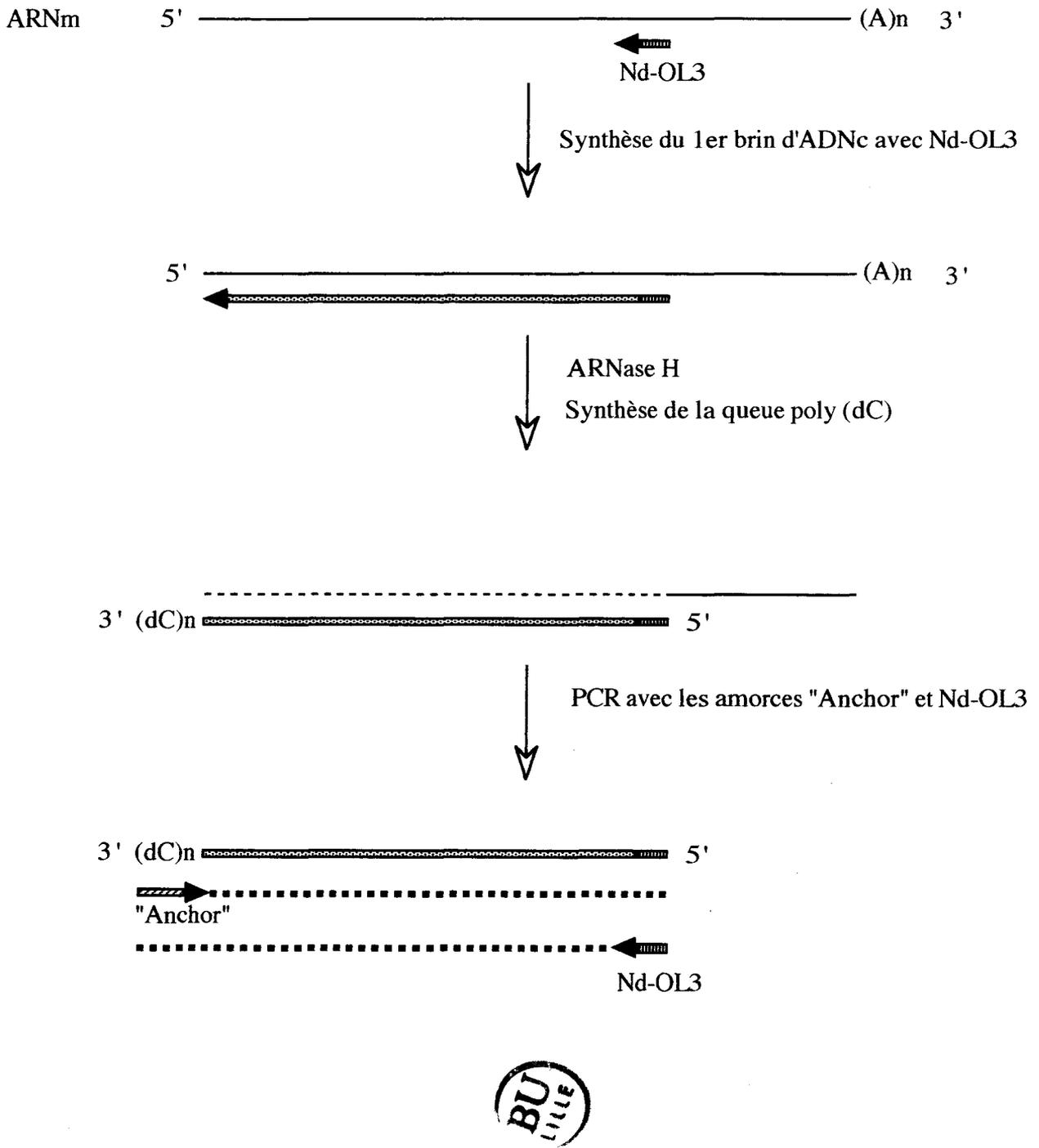


Figure 49





**Figure 50 : Principe d'amplification par la technique "5' RACE System" (Gibco BRL).**

## **VII - Recherche de l'expression des gènes correspondant aux sondes Nd *erg* et Nd *ets***

### **A - Recherche par "Northern blot"**

Les premiers essais de "Northern blots" ont été effectués avec des préparations d'ARN totaux extraits des différentes parties du corps de l'animal (prostomiums, métastomiums et régénérats).

Ne disposant alors que de la sonde Nd *erg*, les hybridations sur les différents "Northerns" d'ARN totaux n'ont été effectuées qu'avec cette sonde. Aucun signal d'hybridation n'a pu être mis en évidence, même après un temps d'exposition très long. Cette manipulation a été répétée plusieurs fois avec différentes préparations d'ARN totaux extraits par différentes techniques mais aucun résultat n'a été obtenu.

Par la suite, des préparations d'ARN messagers ont été réalisées à partir des différentes parties du corps de l'animal, et pour les régénérats les extractions ont été effectuées soit sur un ensemble de régénérats coupés à différents temps, soit sur des régénérats coupés à quatre temps spécifiques marquant la progression du phénomène de régénération (3 jours, 7 jours, 15 jours et 3 semaines).

Avec la sonde Nd *erg*, un signal d'hybridation a été obtenu sur un transfert de 8 µg d'ARN messagers d'un ensemble de régénérats à différents temps. Le signal se situe à environ 2,5 kpb (Figure 48). Pour affiner la manipulation, la même quantité d'ARNm de régénérats prélevés aux temps 3 j, 7 j, 15 j et 3 semaines a été transférée et hybridée avec la sonde. La manipulation a été répétée plusieurs fois sur différentes préparations d'animaux. Le même signal autoradiographique correspondant à un transcrit d'environ 2,5 kpb a de nouveau été obtenu pour chaque type de préparation de régénérats (3 j, 7 j, 15 j et 3 semaines). Cependant, le signal est très faible malgré une exposition longue (1 mois) et n'a pas été obtenu à chaque fois. Sur les autres parties du corps de l'animal, aucun transcrit n'a pu être mis en évidence.

Ces mêmes transferts ont aussi été hybridés avec la sonde Nd *ets* et, pour l'instant, aucun signal d'hybridation n'a été obtenu pour cette sonde.

### **B - Recherche par RT PCR**

Les manipulations de RT PCR ont été effectuées avec un oligonucléotide spécifique de la sonde Nd *erg*, le Nd-OL3 qui dirige la synthèse vers l'extrémité 5', dans le but de montrer une expression du gène correspondant.



Des préparations d'ARN totaux et d'ARN messagers ont été utilisées comme matrice.

Le premier brin d'ADN complémentaire est donc initié avec l'oligonucléotide spécifique. Puis, après dégradation du brin d'ARN de l'hybride ADN/ARN ainsi constitué, une queue poly A est ajoutée par la terminale transférase. Cette expérience a aussi été réalisée avec un "kit" permettant l'amplification rapide des extrémités 5' des ADNc (5' RACE, Gibco BRL) dont le principe est schématisé sur la figure 50. Dans ce cas une séquence spécifique est ajoutée à l'extrémité 3'.

Ensuite, dans les deux cas, le deuxième brin est synthétisé suivi des cycles de PCR. Le principe est le même que celui décrit dans la partie VI. Aucune de ces manipulations n'a donné de résultat même après deux cycles d'amplification.

### **C - Recherche par hybridation *in situ***

Des expériences d'hybridation *in situ* ont été réalisées sur des coupes effectuées à partir des différentes parties du corps de l'animal (prostomium, métastomium et régénérat coupé à différents temps). Les sondes Nd *erg* et Nd *ets* ont été utilisées pour cette étude, mais jusqu'à présent aucun signal distinct n'a pu être détecté.

## **VIII - Recherche des transcrits des gènes correspondant aux sondes Nd *erg* et Nd *ets* dans une banque d'ADN complémentaire**

### **A - Criblage de la banque avec les sondes nucléotidiques**

Une banque d'ADN complémentaire a été construite dans le vecteur phagique  $\lambda$  gt11D commercialisé par la firme Pharmacia. La qualité des ARN messagers utilisés pour la construction de la banque a été vérifiée après traduction des ARN messagers *in vitro* en lysat de réticulocytes de lapin (Figure 49). Le clonage des ADN complémentaires est orienté dans le vecteur : la synthèse du 1er brin d'ADN complémentaire est initiée avec un oligonucléotide oligo(dT) flanqué d'un site de restriction *NotI* et du côté 3' un adaptateur *EcoRI* est ajouté. L'insertion des ADN complémentaires ainsi synthétisés s'effectue dans le vecteur déphosphorylé  $\lambda$  gt11D *EcoRI/NotI* (Figure 51). Après encapsidation, la banque a



**Figure 52 : Electrophorèse des produits de PCR réalisée sur la banque d'ADNc.**

**a et b : Amplification avec les oligonucléotides D-OL2 et Nd-OL4 (1 et 5  $\mu$ l de banque, respectivement).**

**c : Marqueur de taille pBR322 digéré par l'enzyme *HinfI*.**

**d : Témoin de 180 pb.**

**e et f : Amplification avec les oligonucléotides D-OL1 et D-OL2 (1 et 5  $\mu$ l de banque, respectivement).**

**g et h : Réamplification des produits (e) et (f) avec les oligonucléotides D-OL1 et D-OL2.**

**Figure 53 : Electrophorèse des produits de PCR réalisée sur la banque d'ADNc.**

**a, b; f, g et k, l : Amplification avec les oligonucléotides Nd-OL3 et "forward"  $\lambda$  gt11 (1 et 5  $\mu$ l de banque, respectivement).**

**c, d; h, i et m, n : Amplification avec les oligonucléotides Nd-OL4 et "reverse"  $\lambda$  gt11 (1 et 5  $\mu$ l de banque, respectivement).**

**e et j : Marqueur de taille pBR322/*HinfI*.**

**Figure 54 : Electrophorèse des produits de RT PCR réalisée avec les oligonucléotides OL785 et D-OL1.**

**a et c : ARNm de régénérats.**

**b et d : ARNm de métamères.**

**e : Marqueur de taille pBR322/*HinfI*.**

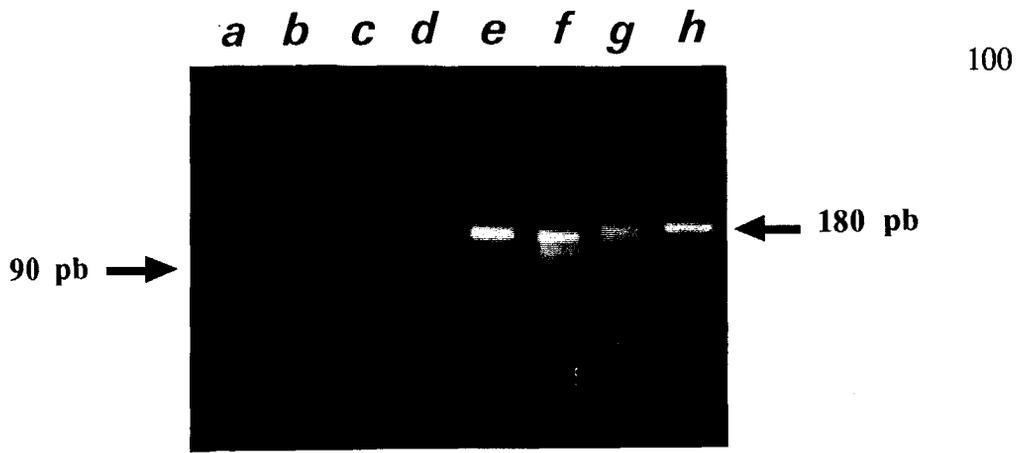


Figure 52



Figure 53

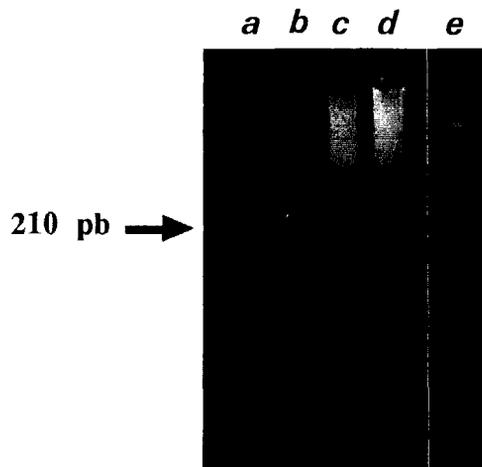


Figure 54

un titre de  $1.10^6$  phages. L'amplification dans la souche bactérienne Y1088 donne un titre final de  $3.10^9$  phages/ml.

Pour la recherche des transcrits codés par les gènes Nd *erg* et Nd *ets*, l'étalement d'un million de phages a été réalisé. Les étalements ont été répliqués sur des filtres de nitrocellulose en deux exemplaires. Chaque série de filtres a été ensuite hybridée, soit avec la sonde Nd *erg*, soit avec la sonde Nd *ets*.

Plusieurs séries d'étalements ont été réalisées, correspondant en tout à quatre millions de phages sans n'avoir jamais pu obtenir le moindre signal pour aucune des deux sondes. Plusieurs indices laissent ici à penser que la non obtention de signaux est plutôt due à un problème d'ordre technique plutôt qu'à un problème de représentativité de la banque (voir ci-après).

### **B - Criblage de la banque par PCR avec les oligonucléotides D-OL1 et D-OL2**

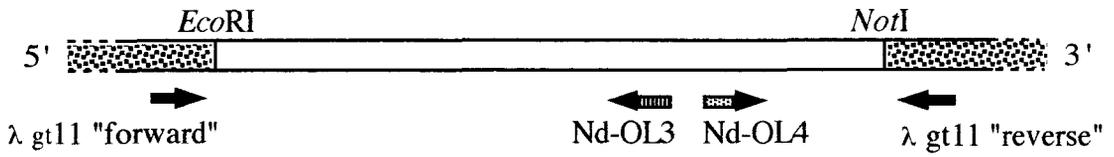
Afin d'essayer de montrer la présence de transcrits de membres de la famille ETS, des essais de PCR ont été réalisés sur des aliquotes de la banque d'ADN complémentaire avec les oligonucléotides D-OL1 et D-OL2. Les résultats présentés sur la figure 52 (pistes e, f, g et h) montrent une amplification de fragments ayant une taille d'environ 180 pb et de fragments plus petits pour les pistes f et g. Cette première manipulation tendait à montrer l'existence de transcrits de gènes de la famille ETS dans la banque, en y mettant des réserves connaissant l'aspécificité potentielle de l'amplification avec les oligonucléotides dégénérés D-OL1 et D-OL2.

Ne pouvant attribuer ces produits d'amplification aux transcrits du gène Nd *erg*, d'autres manipulations ont été effectuées.

### **C - Criblage de la banque par PCR avec des oligonucléotides spécifiques du gène Nd *erg***

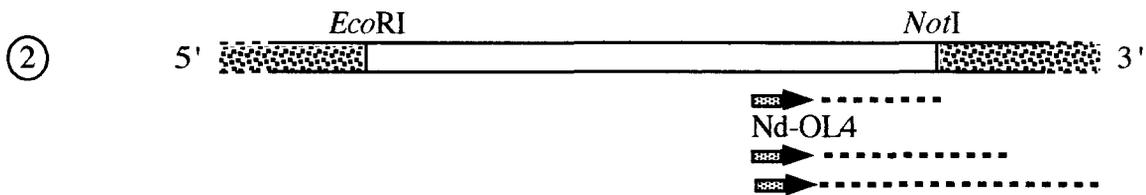
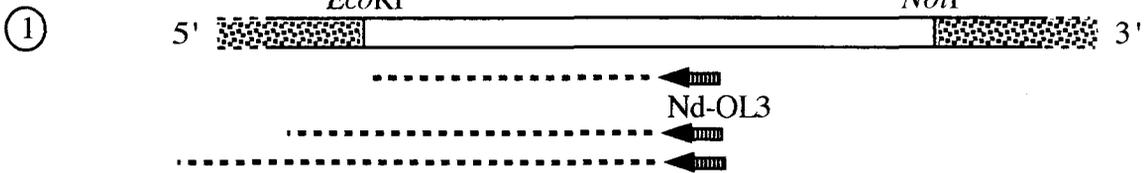
La mise en évidence de transcrits dans la banque d'ADN complémentaire, par PCR, n'a pu être effectuée que pour les transcrits correspondant au gène Nd *erg* car nous ne possédions que des oligonucléotides spécifiques de ce gène.

Le principe de la manipulation est schématisé sur la figure 55 : sur 1 et 5  $\mu$ l de la banque amplifiée, un premier cycle d'amplification par PCR est réalisé en prenant comme amorces, d'une part l'oligonucléotide Nd-OL3 et, d'autre part, l'oligonucléotide Nd-OL4. L'oligonucléotide Nd-OL3 dirige la synthèse du côté 5' alors que l'oligonucléotide Nd-OL4 dirige la synthèse du côté 3'.

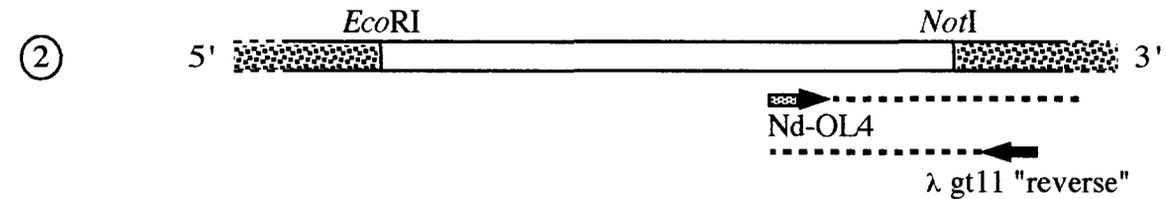
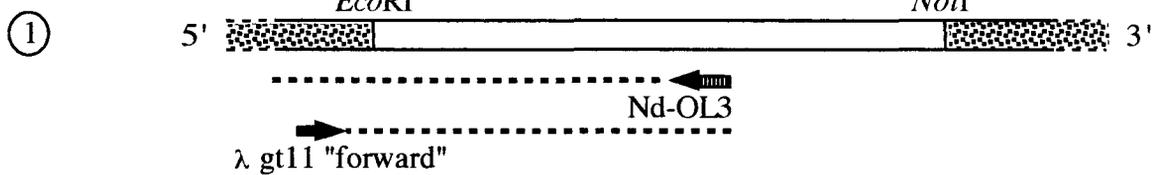


 Vecteur  $\lambda$  gt11D  
 Insert

**1er cycle d'amplification : Nd-OL3  
Nd-OL4**



**2ème cycle d'amplification : Nd-OL3 - "Forward"  
Nd-OL4 - "Reverse"**



**Figure 55 : Principe d'amplification par PCR des ADNc correspondant au gène *Nd erg*, à partir de la banque d'ADNc.**

Pour cette première série d'amplifications, le cycle classique préalablement défini a été utilisé en augmentant le temps d'élongation à 2 minutes 30.

Ensuite, sur 1/5 du produit d'amplification obtenu pour chaque test, un second cycle d'amplification est effectué en utilisant les amorces "reverse" et "forward" du vecteur  $\lambda$  gt11 commercialisées par la firme Promega. Pour les produits obtenus par réaction avec l'oligonucléotide Nd-OL3, le deuxième cycle d'amplification est effectué avec l'oligonucléotide Nd-OL3 et l'amorce "forward" du  $\lambda$  gt11. Pour les produits obtenus avec l'oligonucléotide Nd-OL4, le deuxième cycle d'amplification est réalisé avec l'oligonucléotide Nd-OL4 et l'amorce "reverse" du  $\lambda$  gt11, selon le cycle habituel.

Une aliquote des produits d'amplification de chaque type de préparation a été testée en électrophorèse sur un gel d'agarose à 2 % (Figure 53). La manipulation a été répétée trois fois et a donné à chaque fois le même résultat. Les pistes a, b, f, g, k et l représentent l'amplification avec les oligonucléotides Nd-OL3 et "forward" sur respectivement 1 et 5  $\mu$ l de banque. Les pistes c, d, h, i, m et n représentent, elles, l'amplification de respectivement 1 et 5  $\mu$ l de banque avec les oligonucléotides Nd-OL4 et "reverse".

Pour les amplifications initiées avec l'oligonucléotide Nd-OL3, cinq fragments sont obtenus de tailles de 270, 400, 450, 480 et 500 pb environ. Pour les amplifications initiées avec l'oligonucléotide Nd-OL4, quatre fragments sont obtenus (240, 290, 320 et 400 pb) dont un est plus représenté (400 pb).

Une autre manipulation a aussi été effectuée pour pouvoir confirmer ou infirmer la présence de transcrits codés par le gène Nd *erg*, alors que le criblage traditionnel avec la sonde nucléotidique a échoué.

Deux aliquotes de la banque (1 et 5  $\mu$ l) ont été utilisées dans une PCR avec les amorces oligonucléotidiques Nd-OL4 et D-OL2.

Après deux cycles d'amplification, les produits de la réaction ont été déposés sur un gel d'agarose à 2,5 %. Pour chaque préparation, un fragment de 90 pb est obtenu (Figure 52, pistes a et b). La taille de ce fragment correspond à la taille attendue pour l'amplification.

Ces deux types d'expériences vont dans le sens de l'existence de transcrits codés par le gène Nd *erg* dans la banque d'ADN complémentaire, bien que cela ne sera réellement confirmé qu'après clonage et séquençage des différents fragments obtenus.

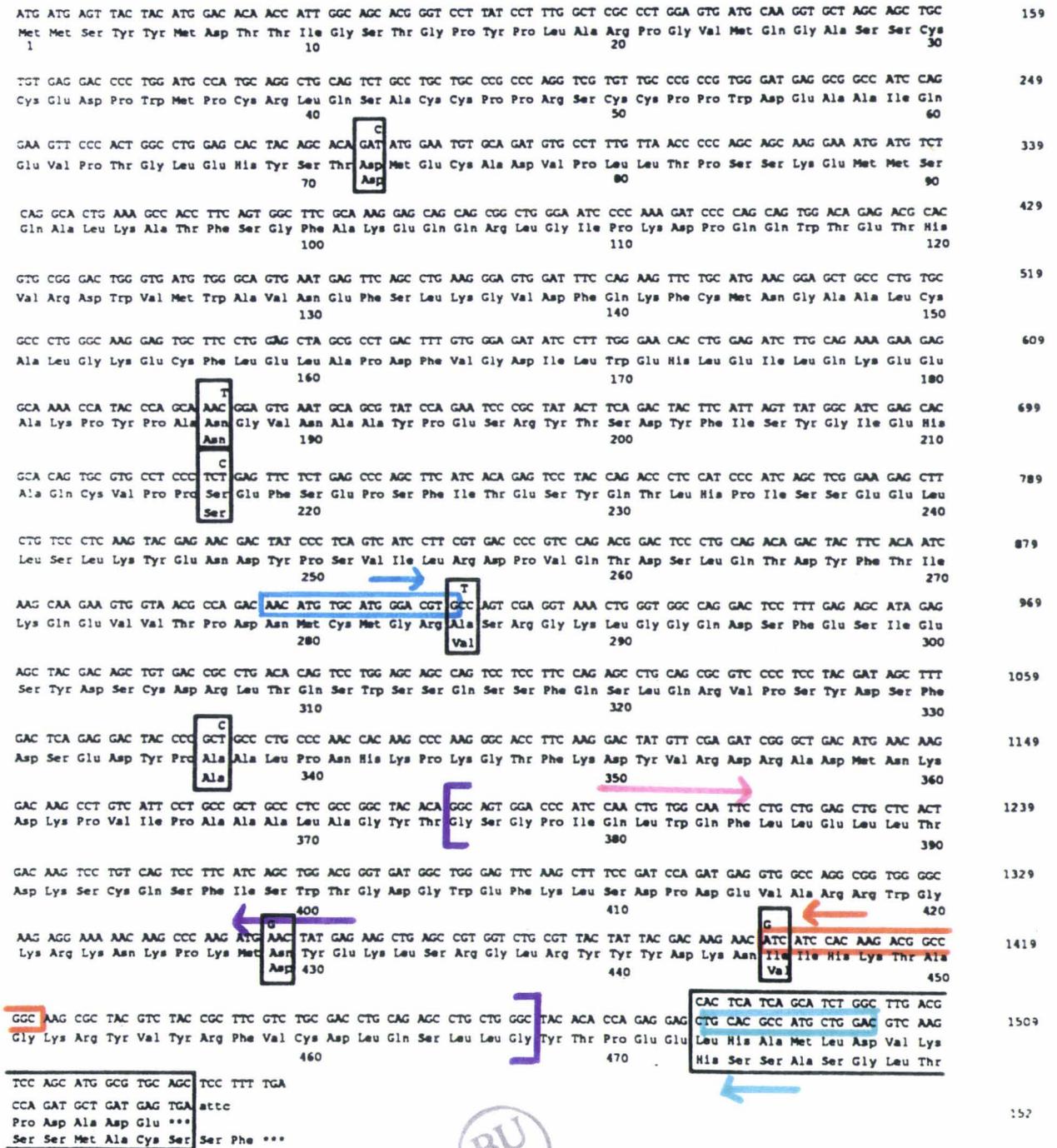


Figure 56 : Position et orientation des oligonucléotides OL 789 (bleu), OL 785 (orange) et OL 793 (vert) par rapport à la séquence de l'ADNc correspondant à la p68<sup>c-ets-1</sup> du poulet (Leprince *et al.*, 1988). Le domaine ETS est délimité par les crochets. La position des oligonucléotides D-OL1 ( flèche rose) et D-OL2 (flèche violette) est représentée à titre indicatif.

## **IX - Recherche par RT PCR de transcrits correspondant au gène Nd *ets***

Le fragment Nd *ets* a été obtenu préalablement par RT PCR. De façon à essayer d'isoler un ou des transcrits plus grands, des manipulations de RT PCR ont été effectuées sur des ARNm extraits de métamères et de régénérats en utilisant des oligonucléotides spécifiques de l'*ets-1* du poulet, fournis par D. LEPRINCE de l'Institut Pasteur de Lille. Les oligonucléotides utilisés sont localisés de la façon suivante sur l'ADNc *ets-1* de 1,5 kpb du poulet (Figure 56).

Les réactions suivantes ont été essayées :

OL 789 - OL 785

OL 789 - OL793

D-OL1 - OL 785

D-OL2 - OL 789

Seule la PCR utilisant les oligonucléotides D-OL1 - OL785 a permis de visualiser l'amplification d'un fragment de 210 pb après migration dans un gel d'agarose à 2 % (Figure 54). Ce produit n'a été obtenu que pour les préparations d'ADNc de régénérats. Ce fragment a une taille correspondant à celle attendue. Cependant, il est localisé dans une trainée, ne permettant pas de l'isoler pour le cloner. Le fragment a été extrait du gel, puis des PCR ont été effectuées en augmentant la température d'hybridation. Cependant, jamais un fragment amplifié bien isolé n'a pu être obtenu.

Ces expériences n'ont donc pas encore permis d'isoler un ADNc de plus grande taille correspondant au gène Nd *ets*.

# **DISCUSSION**

La famille des protéines Ets est une famille de plus en plus importante de facteurs de transcription possédant un domaine de fixation à l'ADN hautement conservé. La famille ETS est largement représentée dans diverses espèces de Vertébrés comme chez des Métazoaires primitifs, montrant ainsi l'ancienneté de ses représentants.

Quand ce travail de recherche a été initié, quelques séquences cellulaires homologues de *v-ets* avaient été identifiées chez certains Vertébrés (poulet, homme et souris). Chez les Invertébrés, la recherche d'homologues n'avait été effectuée que chez la drosophile (*D ets-2*) (Pribyl *et al.*, 1988) et l'oursin (*Su ets-2*) (Chen *et al.*, 1988). La richesse de cette famille de gènes n'avait pas encore été montrée, ni la fonction, ni le rôle et l'ancienneté.

Dans ce cadre, nous avons recherché un homologue des gènes de la famille ETS chez un Invertébré, Métazoaire Cœlomate primitif, le ver marin *Nereis diversicolor*.

### **I - Identification du gène Nd *erg* chez *Nereis diversicolor***

Ne parvenant pas à isoler un membre de la famille ETS par les techniques classiques de criblage de banque, une autre stratégie, utilisant la PCR, a été entreprise.

Les oligonucléotides A et B conçus pour amplifier spécifiquement la région codant une partie du domaine ETS des protéines Ets-2, n'ont pas permis d'obtenir de produits d'amplification de séquences d'ADN génomique, même après modification des conditions expérimentales. Cet échec peut être attribué au fait que les séquences nucléotidiques recherchées chez la néréis sont sensiblement différentes de celles des oligonucléotides utilisés, qui sont peu dégénérés.

Par contre, l'amplification par PCR avec les oligonucléotides D-OL1 et D-OL2 utilisés par Chen *et al.* (1992) pour obtenir différents gènes de la famille ETS chez la drosophile, a permis d'isoler un produit d'amplification d'environ 180 pb.

La séquence prédite en acides aminés de ce fragment a montré une forte homologie avec les domaines ETS des protéines de la famille ETS et principalement avec les domaines des membres du groupe ERG, comprenant les protéines Erg, Fli, D Ets-3 et D Ets-6 (Tableau 6).

La séquence ainsi isolée a servi de sonde pour le criblage d'une banque d'ADN génomique construite dans le vecteur  $\lambda$  DASHII. Cette banque  $\lambda$  DASHII

	Hu Ets-1	Hu Ets-2	Hu Erg-2	Hu ErgB/Fli-1	Mo Fli-1	D Ets-6	Su Erg	Nd Erg
Hu Ets-1	100 %	95,6 %	68,1 %	68,1 %	68,1 %	67,4 %	68,1 %	67 %
Hu Ets-2		100 %	70,3 %	70,3 %	70,3 %	66,3 %	70,3 %	68,1 %
Hu Erg-2			100 %	96,7 %	97,8 %	83,7 %	94,5 %	89 %
Hu ErgB/Fli-1				100 %	98,9 %	82,6 %	93,4 %	89 %
Mo Fli-1					100 %	83,7 %	94,5 %	89 %
D Ets-6						100 %	85,8 %	82,6 %
Su Erg							100 %	91,2 %
Nd Erg								100 %

**Tableau 6 : Pourcentages d'homologie entre les domaines ETS des protéines Hu Ets-1, Hu Ets-2, Hu Erg-2, Hu Erg-B/Fli-1, Mo Fli-1, D Ets-6, Su Erg et Nd Erg.**

a donné un titre initial de  $5.10^5$  phages et s'est avérée être stable. La souche bactérienne utilisée pour la construction de la banque est la souche SRB commercialisée par Stratagene. Elle a été mutée au niveau de certains gènes, permettant ainsi le clonage de fragments difficilement clonables et instables. Notamment, cette souche a été mutée dans des gènes synthétisant les enzymes dégradant l'ADN méthylé (*mcr A*, *mcr CB* et *mrr*).

L'utilisation de la souche SRB a permis d'obtenir une banque stable et qui semble beaucoup plus représentative du génome de la néréis que la première banque construite dans le vecteur  $\lambda$  EMBL4 et la souche bactérienne DP 50.

Un fragment de 3,4 kpb-*EcoRI* a été isolé. Ce fragment contient toute la séquence réputée commune aux protéines du groupe ERG, c'est-à-dire renfermant le domaine ETS et un domaine qui lui est contigu, le domaine R, défini par Lautenberger *et al.* (1992) (Figure 10).

La séquence que nous avons obtenue à partir d'ADN génomique de *Nereis diversicolor* a été appelée Nd *erg*. Cette séquence, traduite en acides aminés (Nd Erg), présente une homologie importante avec les domaines ETS et R des protéines Erg. Ces domaines sont contigus chez la néréis et semblent codés par le même exon.

### **A - Domaine ETS de Nd Erg**

Le domaine ETS est le domaine de liaison à l'ADN caractéristique de toutes les protéines de la famille ETS, leur conférant une activité transactivatrice.

Wang *et al.* (1992) ont montré que ce domaine peut être divisé en trois régions : une hélice  $\alpha$  prédite du côté N-terminal et un domaine riche en acides aminés basiques du côté C-terminal, séparés par une région charnière. L'analyse de la composition en acides aminés de l'hélice  $\alpha$  et du domaine basique a montré que ces deux régions sont bien conservées pour les différents membres de la famille ETS alors que la région charnière est plus divergente.

Comparée à la séquence de la protéine Ets-1 humaine, la séquence Nd Erg présente 78,5 %, 47,8 % et 67,3 % d'homologie avec respectivement les régions en hélice  $\alpha$ , charnière et le domaine basique. La même comparaison effectuée avec le domaine ETS de la protéine Erg-2 humaine montre respectivement pour ces trois régions 92,3 %, 65,2 % et 100 % d'homologie. Ceci montre la bonne conservation de la région en hélice  $\alpha$  et du domaine basique pour des phyla éloignés ainsi que la forte homologie de notre séquence avec la protéine Erg humaine.



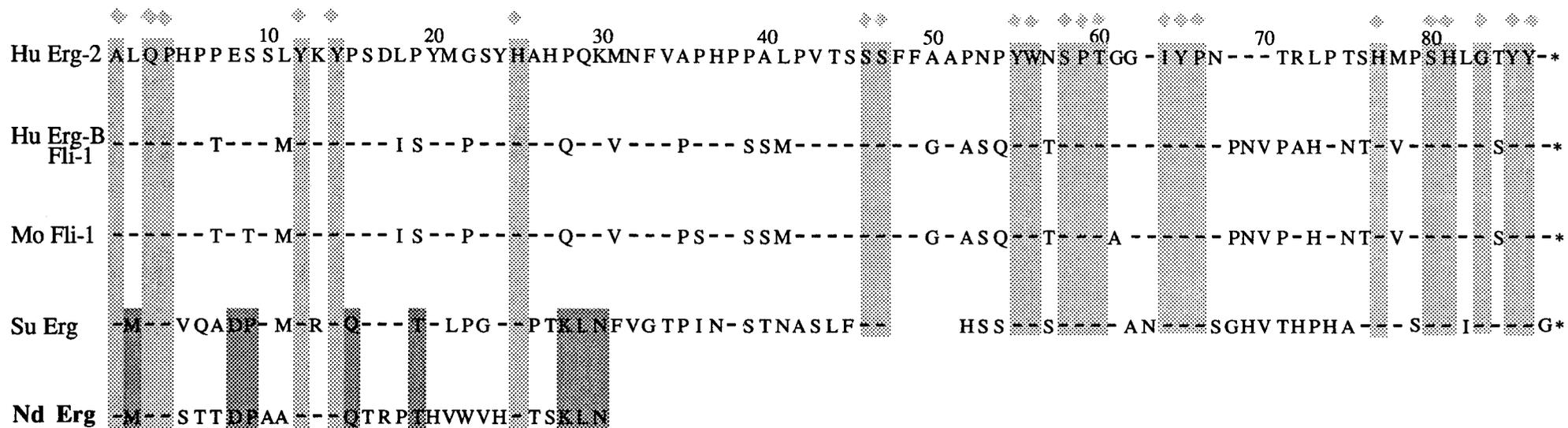
La séquence Nd Erg présente les particularités des domaines ETS des protéines du groupe ERG. On retrouve les acides aminés conservés pour tous les membres de la famille ETS décrits dans la figure 9, et surtout les acides aminés particulièrement bien conservés dans les protéines du groupe ERG (Figure 57). La séquence Nd Erg présente cependant une particularité par rapport aux séquences des protéines Erg jusqu'ici publiées, une valine et une alanine (acides aminés 46 et 47) sont respectivement remplacées par une thréonine et une valine.

Au sein du groupe ERG, les pourcentages d'homologie de notre séquence avec les différents domaines ETS vont de 89 % à 91,2 % alors que l'homologie n'est que de 67 à 68 % vis-à-vis des protéines Ets-1 et Ets-2 humaines (Tableau 6). L'homologie la plus importante est trouvée avec l'Erg d'oursin (91,2 %). Chez d'autres Invertébrés, on ne connaît des séquences appartenant au groupe ERG que chez la drosophile (D Ets-3 et D Ets-6) et chez un ver Nématode. La connaissance du D Ets-3 est incomplète ce qui ne nous permet pas de faire des comparaisons de séquences. Le D Ets-6 est le plus éloigné de tous les membres du groupe ERG. Par contre, chez un ver Nématode, Degnan *et al.* (1993) ont publié une séquence de 33 acides aminés identique à celle de la néréis (excepté les résidus thréonine et valine 46 et 47) avec les deux mêmes acides aminés modifiés (une leucine et une valine). Cette observation est cohérente car les Nématodes constituent un groupe de Métazoaires assez proche des Annélides Polychètes. Cependant, la séquence publiée par Degnan *et al.* (1993) est insuffisante pour pouvoir en faire une étude plus approfondie. Nous pouvons aussi noter que la leucine en position 38 est conservée dans la lignée des Protostomiens chez le ver Nématode, la néréis et chez la drosophile.

Enfin, l'homologie plus grande avec Su Erg peut s'expliquer par la proximité de la néréis et de l'oursin du point de vue évolutif.

## **B - Domaine R de Nd Erg**

Le domaine appelé R par Lautenberger *et al.* (1992) est contigu au domaine ETS et s'étend jusqu'à l'extrémité carboxy-terminale des protéines du groupe ERG. La séquence en acides aminés de ce domaine est assez bien conservée dans le groupe des protéines Erg (Figure 58). Nous avons calculé les pourcentages d'homologie entre les domaines R actuellement connus des différentes protéines du groupe ERG (Tableau 7). Contrairement à ce qui est ici observé pour les protéines Erg, on ne retrouve pas pour les autres protéines Ets d'homologie au niveau de la séquence en acides aminés de la région carboxy-terminale jouxtant le domaine ETS.



◆ : acides aminés conservés au sein du groupe ERG.  
 ■ : acides aminés conservés entre Su Erg et Nd Erg.

**Figure 58 : Alignement du domaine R des différents membres du groupe ERG et de la séquence Nd Erg.**



Une forte homologie (de 68 à 98%) est observée entre les domaines R des protéines Erg et Fli de Vertébrés.

Chez les Invertébrés, seul le domaine R de l'oursin est connu. L'alignement des séquences Nd Erg et Su Erg montre qu'elles sont relativement proches l'une de l'autre mais qu'elles sont, par contre, un peu plus éloignées des domaines R de Vertébrés. En effet, l'homologie entre les domaines R des protéines Erg de Vertébrés et le domaine R de Su Erg ne dépasse pas 40%. Par contre, l'homologie calculée entre la partie connue du domaine R de Nd Erg et la partie correspondante du domaine R de Su Erg est de 50%. D'autre part, cette région amino-terminale est moins homologue entre les protéines de Vertébrés et Su Erg, par rapport à la région carboxy-terminale. La moins bonne homologie entre les domaines R de Vertébrés et d'Invertébrés peut être en relation avec une fonction différente de ce domaine au sein de ces deux grands groupes zoologiques.

Les protéines Erg et Fli humaines possèdent dans leur partie carboxy-terminale un domaine transactivateur, le domaine CTA, avec une homologie de plus de 80 % entre les deux protéines. Ce domaine CTA correspond à la moitié carboxy-terminale du domaine R. L'analyse de la composition en acides aminés du domaine CTA a montré qu'il est particulièrement riche en résidus proline (17,5 % pour Fli-1 et 21 % pour Erg) et sérine (16 % pour Fli-1 et 12 % pour Erg) (Rao *et al.*, 1993 ; Siddique *et al.*, 1993). L'analyse de la structure secondaire de la séquence peptidique du domaine CTA des protéines Erg et Fli prédit une structure de type tour-boucle-tour. Cette structure pourrait être impliquée dans des mécanismes d'activation de la transcription par interactions intermoléculaires avec d'autres protéines.

Le domaine fonctionnel CTA ne correspond qu'à l'extrémité C-terminale du domaine R. Or, l'extrémité N-terminale du domaine R est elle aussi bien conservée chez les Vertébrés d'une part et chez les Invertébrés d'autre part, comme le confirme la séquence Nd Erg. Cette extrémité N-terminale pourrait donc correspondre à un domaine fonctionnel non encore décrit.

### **C - Extrémité 5' de la séquence Nd erg**

A l'extrémité 5' de la séquence Nd *erg*, flanquant la région codant le domaine ETS, existe une séquence particulièrement riche en adénine et thymine (80 % de A-T). Entre cette séquence et la séquence codant le domaine ETS, nous observons un site accepteur d'épissage conforme aux séquences consensus publiées par Mount (1982). Ce site consensus d'épissage délimite le début de l'exon codant le domaine ETS pour tous les membres de la famille ETS.

	Hu Erg-2	Hu Erg B/Fli-1	Mo Fli-1	Su Erg
Hu Erg-2	100 %	70,4 %	68,2 %	40 %
Hu Erg B/Fli-1		100 %	95,4 %	40,9 %
Mo Fli-1			100 %	34,7 %
Su Erg				100 %

**Tableau 7 : Pourcentages d'homologie entre les domaines R des protéines Hu Erg-2, Hu Erg-B/Fli-1, Mo Fli-1 et Su Erg.**

La présence d'un intron très riche en A-T n'est sans doute pas aléatoire mais la littérature ne donne pas d'information à ce sujet, à notre connaissance. Pour les autres gènes *ets*, les introns ne possèdent pas cette caractéristique. Seul l'intron situé à l'intérieur de la séquence codant le domaine ETS du gène D *ets-3* de drosophile est, lui aussi, riche en A-T.

## **II - Recherche d'autres gènes de la famille ETS**

Les oligonucléotides D-OL1 et D-OL2 ont été choisis pour leur haute conservation à l'intérieur du domaine ETS, pour les différents membres de la famille ETS.

Par PCR sur de l'ADN génomique de néréis, ces oligonucléotides nous ont permis d'isoler une séquence (Nd *erg*) appartenant au groupe ERG. Cependant, du fait de leur situation dans une zone hautement conservée, ces oligonucléotides sont susceptibles d'amplifier par PCR d'autres membres de la famille ETS. Plusieurs clones provenant du clonage des produits de 180 pb correspondant à l'amplification de l'ADN génomique par les oligonucléotides D-OL1 et D-OL2, ont été isolés. Huit de ces clones ont été sélectionnés et six ont été séquencés. Sur les six clones séquencés, deux ont une séquence identique, mais aucun n'a montré d'homologie de leur séquence traduite en acides aminés avec des membres de la famille ETS.

Degnan *et al.* (1993) ont mis en évidence par la technique de PCR, des membres de la famille ETS chez de nombreuses espèces de Métazoaires, Triploblastiques et Diploblastiques. Les oligonucléotides utilisés par cette équipe sont des oligonucléotides dégénérés conçus à partir de séquences en acides aminés appartenant aux domaines ETS des protéines humaines *Ets-1* et *Ets-2*.

Sur les 171 clones séquencés à partir des produits d'amplification obtenus pour 12 espèces d'Invertébrés marins, 52 clones comportant 16 séquences *ets* différentes ont été obtenus, les autres clones correspondent à une amplification spécifique.

Le fait de ne pas avoir obtenu d'autre séquence chez la néréis est peut-être dû à un nombre statistiquement insuffisant de clones séquencés. Disposant de nombreux autres clones, leur séquençage pourra être envisagé par la suite.

Chen *et al.* (1992), ont été beaucoup plus chanceux dans leurs expériences de PCR en utilisant les oligonucléotides D-OL1 et D-OL2. En effet, l'amplification d'ADN génomique de *Drosophila melanogaster* avec ces oligonucléotides a donné des produits de 180 pb, 250 pb et des produits de plus

grande taille. Sur cinq clones portant un insert de 250 pb, un seul clone avait une origine inconnue et quatre correspondaient au D *ets-2* (le D *ets-2* possède dans la région codant le domaine ETS un intron de 63 pb expliquant la taille de 250 pb du produit d'amplification). Sur 14 clones portant un fragment de 180 pb, deux clones correspondaient à la séquence de l'*E74*, un clone correspondait au gène *elg* et les 11 clones restant possédaient deux nouvelles séquences apparentées aux gènes *ets* : D *ets-3* et D *ets-4*. De même, des expériences de PCR sur une banque d'ADN complémentaire suivies du clonage des produits obtenus et de leur séquençage leur ont permis de retrouver 19 séquences identiques à celles précédemment citées et 15 séquences représentant un nouveau membre, D *ets-6*, sur les 51 clones séquencés.

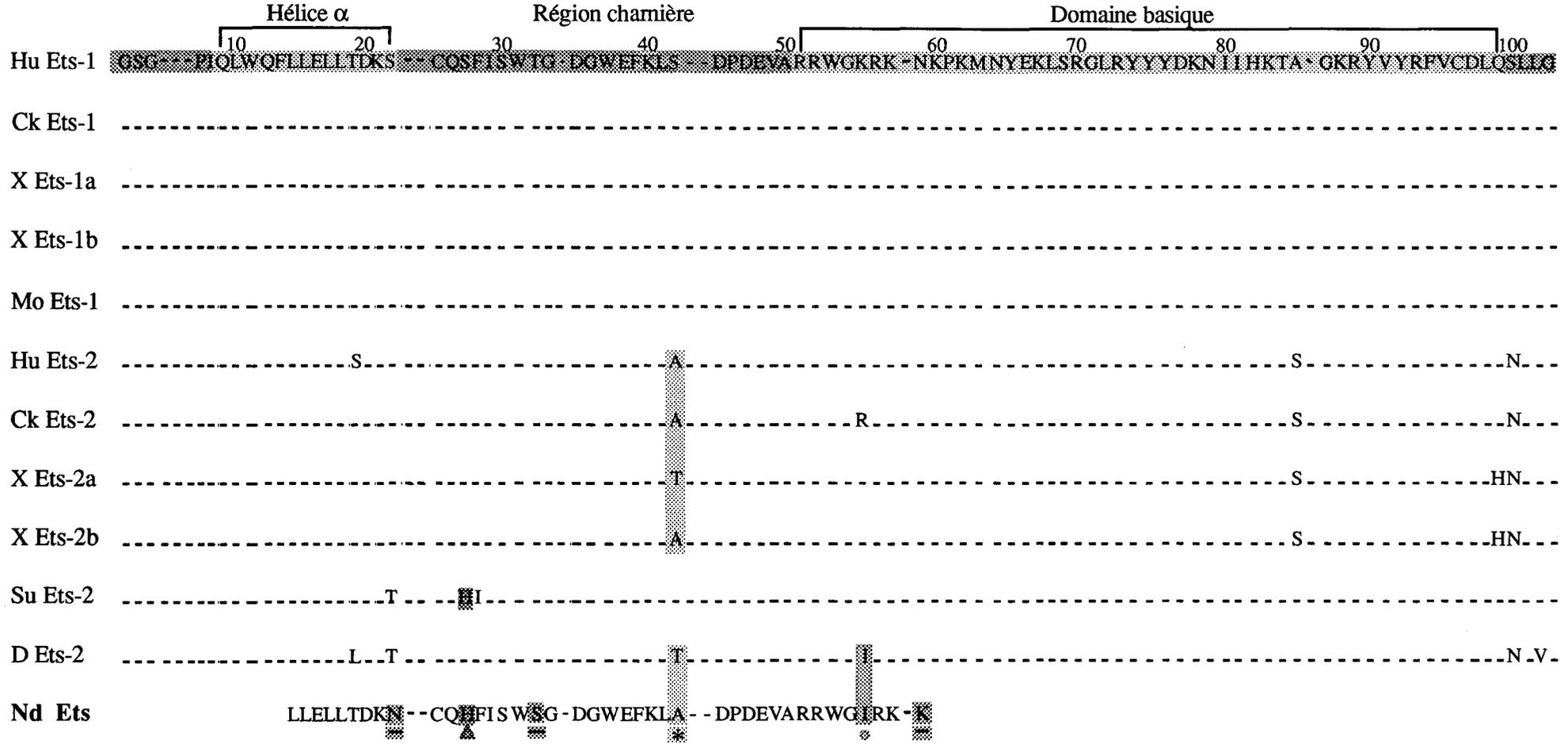
Ces résultats confirment que le nombre de clones séquencés chez la néréis n'est sûrement pas suffisant pour pouvoir trouver de nouveaux membres de la famille ETS. Les paramètres des cycles d'amplification que nous avons effectués sont les mêmes que ceux utilisés par Chen *et al.* (1992) ; le rendement qu'ils ont obtenu est cependant supérieur au nôtre.

### **III - Recherche de l'expression des gènes de la famille ETS et identification du gène Nd *ets***

La recherche d'un proto-oncogène chez une Annélide Polychète n'avait jusqu'ici jamais été entreprise. Ayant mis en évidence l'existence de séquences géniques appartenant à la famille ETS, nous avons recherché leur éventuelle expression dans différents tissus.

L'Annélide Polychète *Nereis diversicolor* est un ver marin caractérisé par une succession de segments identiques, les métamères. Seuls les segments antérieur (prostomium) et postérieur (pygidium) sont morphologiquement différents. L'extrémité pygidiale est précédée d'une zone en constante prolifération, aboutissant à terme à l'édification de métamères. La néréis est capable de régénérer, mais uniquement en direction caudale et sous influence endocrine (Durchon, 1967). Le processus de régénération implique des phénomènes de cicatrisation et de prolifération cellulaire, pour lesquels de nombreux gènes sont susceptibles d'être impliqués et notamment ceux codant des facteurs de transcription.

Nous avons donc recherché l'expression de membres de la famille ETS et notamment au cours du processus de régénération. La régénération caudale comprend trois phases : la cicatrisation, la mise en place d'un blastème de régénération se différenciant ensuite en pygidium, et enfin l'allongement du



- Acides aminés spécifiques de Nd Ets.
- \* Acide aminé existant dans les protéines Ets-2 humaine, du poulet et du xénope et dans le D Ets-2.
- ▲ Acide aminé existant dans Nd Ets et Su Ets-2.
- ◇ Acide aminé existant dans Nd Ets et D Ets-2.

**Figure 59 : Alignement du domaine ETS des différents membres du groupe ETS et de la séquence Nd Ets.**

régénérat et son organisation en segments sétigères. L'ensemble de ces processus peut mettre jusqu'à 60 jours.

Nous avons choisi différents temps de régénération correspondant à des stades où les phénomènes de prolifération sont importants, pensant détecter une expression plus importante des proto-oncogènes de la famille ETS. Ces temps sont de :

3 jours caractérisant la cicatrisation,

7 jours caractérisant l'édification du blastème de régénération,

15 jours caractérisant l'organisation et l'allongement du régénérat.  
3 semaines

Des régénérats de ces différents stades ont été préparés en vue de la confection de coupes pour l'étude de l'expression *in situ* et de préparations d'ARNm pour les études en "Northern blot" et en RT PCR.

Parallèlement, des préparations de métamères et de prostomiums ont été utilisées pour les mêmes études.

Nous avons donc recherché une éventuelle expression des membres de la famille ETS dans ces trois parties du corps de l'animal.

#### **A - Recherche de l'expression par RT PCR**

Des manipulations de RT PCR ont été menées de différentes façons dans le but d'isoler des séquences correspondant à des transcrits des membres de la famille ETS.

Les expériences de RT PCR effectuées avec les oligonucléotides D-OL1 et D-OL2 sur des ARNm extraits des différentes parties du corps de l'animal ont permis d'obtenir l'amplification de fragments de taille attendue (180 pb) et de fragments de 220 pb. Plusieurs clones correspondant aux deux types de fragments ont été sélectionnés et 12 d'entre eux ont été séquencés. L'analyse de la séquence de ces clones, traduite en acides aminés, n'a montré aucune homologie avec des membres de la famille ETS, bien que certains d'entre eux s'hybridaient fortement avec l'oligonucléotide Nd-OL3.

D'autres manipulations de RT PCR ont été effectuées dans le but d'isoler les extrémités 5' et 3' des transcrits des membres de la famille ETS (oligonucléotides D-OL1 et D-OL2) ou l'extrémité 5' du ou des transcrits du gène Nd *erg* (oligonucléotide Nd-OL3).

Une amplification de fragments de 200 à 600 pb a été obtenue, uniquement avec les oligonucléotides D-OL1 et oligo(dT)<sub>18</sub> pour des RT PCR effectuées sur des préparations d'ARNm extraits des différentes parties du corps de l'animal. Après clonage, différentes préparations de fragments ont été réalisées. L'hybridation avec la sonde Nd *erg* a montré des fragments pouvant présenter une

forte homologie. Sept fragments, dont six qui s'hybridaient fortement avec la sonde Nd *erg*, ont été séquencés : seule la séquence du fragment s'hybridant faiblement, traduite en acides aminés, correspond à un nouveau membre de la famille ETS. Ce fragment est issu d'une préparation d'ARN messagers effectuée à partir de régénérats de pygidium.

Cette séquence a été appelée Nd Ets car elle présente une grande homologie en acides aminés avec les domaines ETS des membres du groupe ETS. Elle comporte 40 acides aminés, homologues à plus de 85 % aussi bien avec la protéine Ets-1 humaine qu'avec les séquences en acides aminés des protéines Su Ets-2 et D Ets-2 (Figure 59). Comme pour la séquence D Ets-2 de drosophile, le même acide aminé isoleucine est observé en position 55. Cette isoleucine existe aussi chez quelques autres Invertébrés (Mollusques, Cnidaires) (Degan *et al.*, 1993). De même, la sérine en position 27 est remplacée par une histidine, aussi bien dans la séquence Nd Ets que pour l'Ets-2 d'oursin.

Disposant d'une séquence proche des membres du groupe ETS, nous avons cherché à isoler par RT PCR les séquences en 3' et en 5' du fragment Nd *ets*. Pour cela, nous avons utilisé des oligonucléotides spécifiques de la séquence *c-ets-1* aviaire fournis par D. LEPRINCE (Institut Pasteur de Lille). Les différentes combinaisons possibles de ces oligonucléotides entre eux et avec les oligonucléotides D-OL1 et D-OL2 dégénérés ont été essayées. Seule la réaction utilisant l'oligonucléotide OL-785 et l'oligonucléotide D-OL1, tous deux situés dans le domaine ETS, a donné l'amplification d'un fragment de taille attendue mais uniquement à partir de préparations de régénérats.

En résumé, quelle que soit la partie du corps de l'animal utilisée, les expériences de RT PCR, même après un nouveau cycle d'amplification, n'ont pas toujours donné de produit d'amplification. Avec les oligonucléotides D-OL1 et D-OL2, l'amplification de fragments d'une taille attendue a été obtenue, mais les séquencages effectués suivis de la traduction en acides aminés n'ont pas montré d'homologie avec les domaines ETS.

Les manipulations réalisées pour isoler l'extrémité 5' des transcrits de Nd *erg* (Nd-OL3) et d'autres membres de la famille ETS (D-OL2) n'ont donné aucun produit d'amplification. Cependant, cette expérience nécessite l'addition d'une séquence homopolymérique susceptible de diminuer le rendement de la manipulation. Les expériences visant à isoler l'extrémité 3' des transcrits avec les oligonucléotides D-OL1 et oligo(dT)<sub>18</sub> ont donné des produits d'amplification.

L'ensemble de ces résultats nous amène à faire plusieurs remarques :

- l'utilisation d'oligonucléotides dégénérés a l'avantage d'étendre nos investigations à la plupart des membres de la famille ETS. Cependant, chez la néréis, d'une part les séquences nucléotidiques recherchées peuvent être plus éloignées, d'autre part la dégénérescence des oligonucléotides utilisés diminue le rendement des cycles d'amplification.
- une seule séquence, *Nd ets*, a été obtenue. Cependant, le nombre de séquençages effectués est très faible par rapport à la diversité des types de manipulations réalisées.
- pour les RT PCR qui ont donné des produits d'amplification, plusieurs remarques peuvent être formulées : avec les oligonucléotides D-OL1 et oligo(dT)<sub>18</sub>, les mêmes produits d'amplification sont obtenus quelles que soient les parties du corps de l'animal testées. Cependant, ne connaissant pas la taille attendue des fragments, nous ne pouvons pas conclure de façon précise sur la nature d'une éventuelle expression. Par contre, avec les oligonucléotides D-OL1 et D-OL2 les amplifiats obtenus ont toujours une taille d'environ 180 pb ; il est donc raisonnable de penser que certains de ces fragments correspondent à des membres de la famille ETS.
- les RT PCR montrent qu'il y a une expression des gènes de la famille ETS quelle que soit la partie du corps de l'animal étudiée. Les études de l'expression des protéines de la famille ETS ont montré qu'elles sont en effet représentées dans une grande variété de tissus. Un gène du groupe ETS (*Nd ets*) s'exprime dans des régénérats de pygidium. L'amplification d'un fragment de 210 pb avec les oligonucléotides D-OL1 et OL-785 a été mise en évidence dans les régénérats et nous pouvons penser qu'il appartient au même gène *Nd ets*. Cependant, cette expression n'est pas nécessairement spécifique du phénomène de régénération.

## **B - Recherche de l'expression par "Northern blots" et hybridation *in situ***

Nous avons montré l'existence de gènes de la famille ETS chez la néréis : *Nd erg* obtenu à partir d'ADN génomique et *Nd ets* isolé d'ARN messagers de régénérats de pygidium de l'animal.

Ne connaissant pas le (ou les) type(s) de tissu(s) où ces proto-oncogènes peuvent s'exprimer chez la néréis, nous avons recherché l'expression des gènes *Nd erg* et *Nd ets* dans le prostomium qui contient le cerveau, dans les métamères, et dans l'extrémité caudale en cours de régénération.

Les expériences de "Northern blots" avec la sonde Nd *erg* ont permis de mettre en évidence un ARN messager d'environ 2,5 kpb sur une préparation d'ARNm de régénérats à différents temps. Ce résultat a été confirmé par la suite sur des préparations d'ARNm de régénérats coupés aux quatre temps spécifiques marquant la progression du phénomène de régénération, mais le signal obtenu est de très faible intensité. Par contre, aucune expression n'a pu être montrée avec la sonde Nd *ets*, dans les différentes parties du corps de l'animal.

De même, les expériences d'hybridation *in situ* n'ont pas donné de résultat permettant de repérer un signal distinct.

La néréis ne possède pas d'organes différenciés nous permettant d'extraire des ARN messagers en quantité suffisante et correspondant à une fonction métabolique particulière, c'est pourquoi nos préparations d'ARNm ont été effectuées sur des parties du corps de l'animal.

Pour les régénérats de pygidium, l'extraction d'environ 4 µg d'ARN messagers est obtenue à partir de 100 mg de tissus, qui nécessitent la section et la régénération d'environ 400 animaux. Pour les métamères et le prostomium, le facteur limitant n'est pas tant la quantité de matériel que l'impossibilité d'isoler les cellules exprimant potentiellement le gène étudié.

L'expression généralement faible des proto-oncogènes, notamment ceux de la famille ETS, et la faible quantité d'ARNm obtenue par extraction, nous amènent à penser que l'expression des gènes recherchée est très faible et ne peut être aisément détectée par les techniques utilisées.

La question de la localisation tissulaire de l'expression de gènes de la famille ETS chez la néréis, montrée par "Northern blot" pour *erg* et par RT PCR pour *ets*, reste donc posée. La littérature ne donne que peu d'informations sur la localisation de l'expression des gènes du groupe ERG. A ce jour, les transcrits du gène *erg* humain ont été détectés dans des lignées de cellules lymphoïdes (COLO 320 et Molt 4) (Rao *et al.*, 1987) et n'ont pas été mis en évidence dans les muscles, l'intestin, le cœur, la langue et le cerveau. Des travaux plus récents indiquent que la protéine Erg-2 codée par le gène *erg*, dont l'expression a été recherchée dans différentes lignées myéloïdes, érythroïdes et lymphoïdes, est trouvée dans un nombre limité de lignées cellulaires mais est présente en plus grande quantité dans les cellules myéloïdes que dans les cellules lymphoïdes (Murakami *et al.*, 1993).

Les gènes du groupe ETS ont été particulièrement bien étudiés dans les lignées lymphoïdes de Vertébrés. Les gènes *ets-1* sont exprimés à un haut niveau dans les cellules lymphoïdes alors que les gènes *ets-2* ont une expression ubiquitaire mais qui est faible. Chez les Annélides Polychètes il est possible de

comparer ces types cellulaires à certaines catégories de cellules cœlomiques, les granulocytes.

Les granulocytes appartiennent à différents types cellulaires, dont trois sont bien identifiés. Les granulocytes de type I, que certains comparent aux éosinophiles et aux neutrophiles sanguins des Mammifères, synthétisent notamment une métalloprotéine, la MP II (Dhainaut-Courtois *et al.*, 1987) ayant une activité antibactérienne (Dhainaut *et al.*, 1989). Les granulocytes de type II jouent un rôle dans la cicatrisation et ceux de type III dans la reconnaissance du "non-soi". Les trois types de granulocytes coopèrent au cours des phénomènes d'encapsulation des corps étrangers introduits dans le liquide cœlomique (Porchet-Henneré *et al.*, 1987).

Ainsi, les cœlomocytes de *Nereis diversicolor* sont des cellules diverses, spécialisées, jouant un rôle important chez l'animal, particulièrement dans les phénomènes immunitaires et de cicatrisation.

Connaissant l'implication des gènes *ets* et *erg* dans le système immunitaire des Vertébrés, il sera intéressant de préciser l'importance et le rôle éventuel de ces gènes dans un système immunitaire primitif tel que celui des Annélides Polychètes.

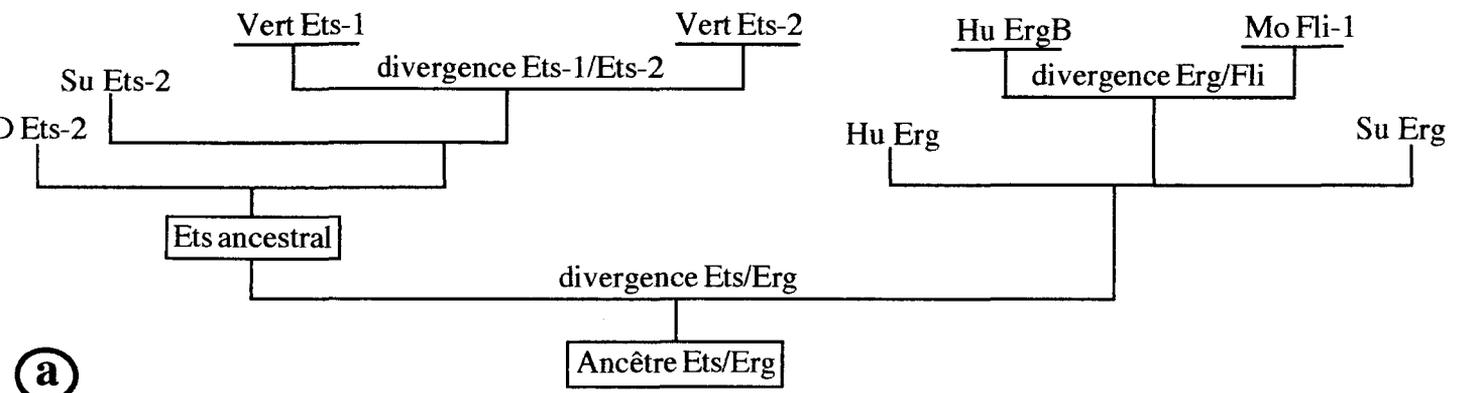
#### **IV - Recherche des transcrits des gènes Nd *erg* et Nd *ets***

##### **A - Criblage d'une banque d'ADNc avec les sondes Nd *erg* et Nd *ets***

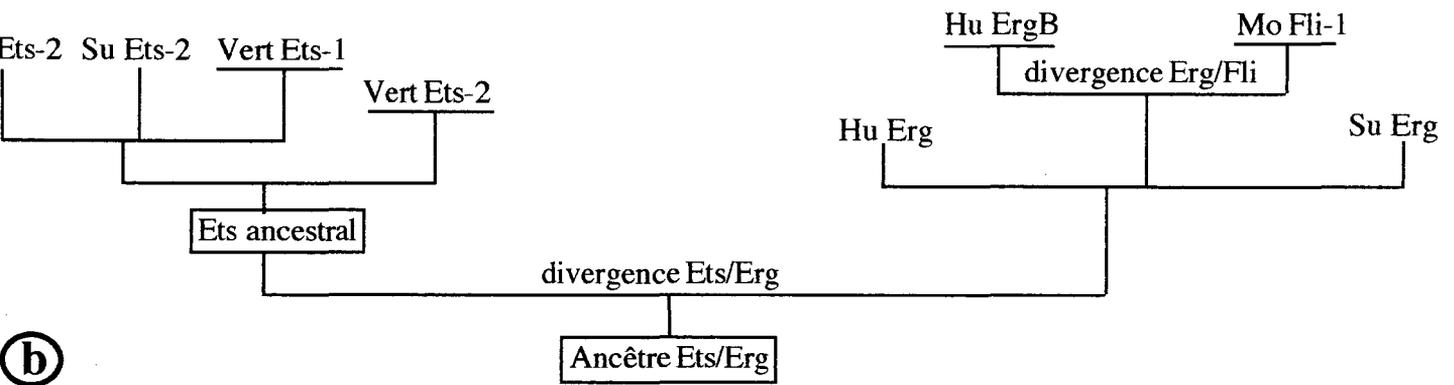
Une banque d'ADNc a été construite dans le vecteur déphosphorylé  $\lambda$  gt11D à partir d'ARNm extraits de régénérats de pygidium et de métamères. La banque a un titre initial de  $1.10^6$  phages, puis de  $3.10^9$  phages/ml après amplification. Plusieurs criblages de 600 000 phages ont été effectués avec les deux sondes. Aucun signal d'hybridation n'a été obtenu. Cet échec peut être attribué soit à un manque de représentativité de notre banque, soit à un problème technique inhérent à la manipulation. Cette deuxième solution semble vraisemblable, vu les résultats obtenus ultérieurement par PCR.

##### **B - Criblage d'une banque d'ADNc par PCR**

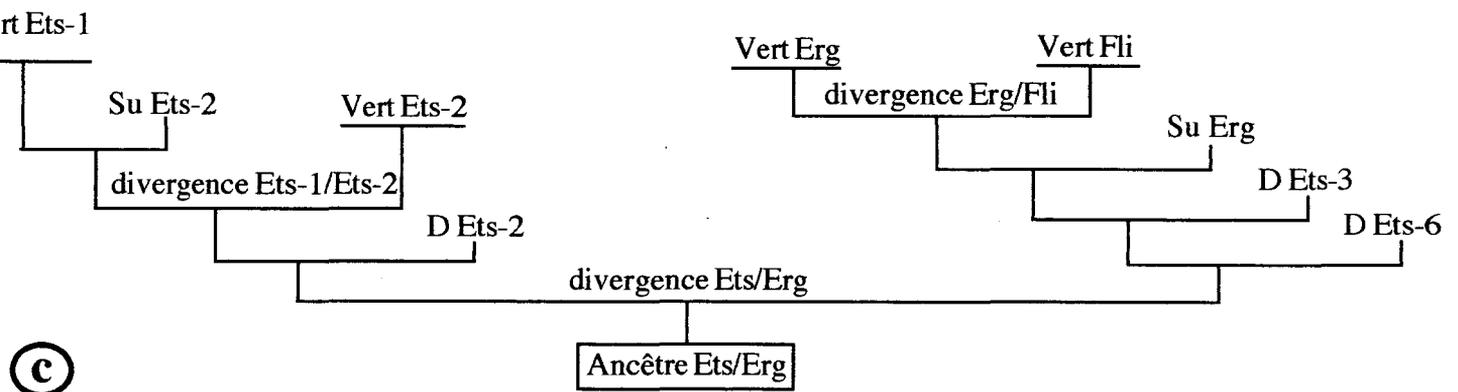
Des amplifications par PCR ont été réalisées sur des aliquotes de banque d'ADNc. Une première série d'amplifications utilisant soit l'oligonucléotide Nd-OL3, soit l'oligonucléotide Nd-OL4 a été réalisée, suivie



(a)



(b)



(c)



Figure 60 : Arbres phylogéniques construits à partir de la séquence en acides aminés des domaines ETS et présentant les divergences Ets-1/Ets-2 et Erg/Fli.

(a) et (b) d'après Lautenberger *et al.* (1992) selon les méthodes de Fitch-Margoliash (1967) (a), et de Swofford (1984) (b)

(c) d'après Laudet *et al.* (1993) selon les méthodes de Fitch (1981), Saitou et Nei (1987) et UPGMA.

d'une deuxième série utilisant respectivement les oligonucléotides Nd-OL3 et "forward"  $\lambda$  gt11 d'une part, et Nd-OL4 et "reverse"  $\lambda$  gt11 d'autre part.

L'expérience a été répétée trois fois et a toujours mis en évidence les mêmes types de produits d'amplification. Bien que reproductibles ces résultats ne seront confirmés qu'après clonage et séquençage des différents fragments.

Afin de vérifier la qualité de notre banque, d'autres amplifications avec les oligonucléotides D-OL2 et Nd-OL4 ont été effectuées. Une amplification de fragments de la taille attendue de 90 pb a été obtenue. D'autre part, des amplifications ont aussi été réalisées avec les oligonucléotides D-OL1 et D-OL2 : un produit d'amplification d'environ 180 pb a été mis en évidence ; une réamplification avec les oligonucléotides D-OL2 et Nd-OL4 a montré le fragment de 90 pb attendu. L'ensemble de ces résultats laisse à penser que la banque d'ADNc contient bien des transcrits correspondant au gène Nd *erg* et vraisemblablement à d'autres membres de la famille ETS.

## **V - Notion d'évolution et situation phylogénique de Nd Erg et Nd Ets**

Nous avons isolé deux séquences de la famille des gènes *ets*. Après restitution de la séquence en acides aminés, le traitement informatique de la séquence Nd Erg, effectué par V. LAUDET (Institut Pasteur de Lille) montre qu'elle est proche des protéines Erg et Fli de Mammifères. Pour la séquence Nd Ets nous obtenons une forte homologie avec les protéines du groupe ETS.

### **A - Evolution phylogénique des domaines ETS**

Deux publications récentes, celle de Lautenberger *et al.* (1992) et celle de Laudet *et al.* (1993) ont proposé une phylogénie des domaines ETS. Ces études ont été complétées par les travaux de Degnan *et al.* (1993), notamment chez des Métazoaires primitifs.

#### **1 - Analyse phylogénique selon Lautenberger *et al.* (1992)**

Lautenberger *et al.* (1992) ont aligné 21 séquences de domaines ETS appartenant à six espèces (homme, souris, poulet, xénope, oursin et drosophile). Chez les Invertébrés ils n'ont inclus dans leur étude que les protéines D Ets-2, D Elg, E74 de drosophile, et la protéine Su Ets-2 d'oursin.

Ces séquences ont été traitées selon deux techniques :

— la méthode de Fitch et Margoliash (1967) (Figure 60a) ;

— la méthode de Swofford (1984) utilisant le programme PAUP ("Phylogenetic Analysis Using Parsimony") (Figure 60b).

Dans l'analyse de l'évolution des gènes *erg* et *ets*, les deux méthodes s'accordent pour partir d'un ancêtre commun à ces deux gènes qui, après duplication, serait à l'origine de deux lignées, l'une menant au groupe ETS, l'autre au groupe ERG. Cet évènement de duplication se serait produit avant l'apparition des Echinodermes et des Cordés, il y a 500 millions d'années, car les gènes *ets* et *erg* sont distincts chez l'homme et l'oursin.

Au sein du groupe ERG, quelle que soit la technique utilisée, on observe le même évènement de duplication : une divergence entre Hu *erg*, Hu *ergB/Mo fli-1* et Su *erg*, confirmant la parenté entre l'*erg* d'oursin et les *erg* ou *fli* de Vertébrés.

Les études de Lautenberger *et al.* (1992) sont discordantes dans l'analyse de l'évolution du groupe ETS : elles situent différemment la divergence *ets-1/ets-2*. La première méthode utilisée, celle de Fitch-Margoliash, indique d'abord l'existence d'un *ets* ancestral qui aurait pu donner le D *ets-2* et le Su *ets-2*, la divergence *ets-1/ets-2* s'étant produite plus tard, avant l'apparition des Amphibiens, puisque l'on trouve un *ets-1* et un *ets-2* individualisés à la fois chez le xénope, le poulet et l'homme. La deuxième méthode utilisée implique l'existence d'un *ets* ancestral qui aurait donné les *ets-2* de Vertébrés ; la duplication de cet *ets-2*, après divergence aurait donné D *ets-2*, Su *ets-2* et les *ets-1* de vertébrés car, pour ces auteurs, les *ets* d'Invertébrés (oursin ou drosophile) s'apparenteraient aux *ets-1* de Vertébrés. Ainsi, en utilisant le programme PAUP, la divergence *ets-1/ets-2* précéderait la divergence Invertébrés/Vertébrés. Cependant, les observations de Lautenberger *et al.* (1992) qui apparentent les *ets* d'Invertébrés aux *ets-1* sont admises pour l'*ets* d'oursin, qui avait été appelé Su *ets-2* par Chen *et al.* (1988) qui avaient initialement observé une forte homologie avec l'*ets-2* humain, mais sont contestables pour le D *ets-2* qui reste habituellement considéré comme un *ets-2*.

Pour Lautenberger *et al.* (1992), donc, la divergence *ets-1/ets-2* reste en suspens : soit avant la divergence Invertébrés/Vertébrés, soit après, selon la technique d'analyse utilisée.

## 2 - Analyse phylogénique selon Laudet *et al.* (1993)

Les études de Laudet *et al.* (1993) ont porté sur l'alignement de 27 séquences de domaines ETS ; en ce qui concerne les Ets d'Invertébrés, D Ets-3, D Ets-4 et D Ets-6 sont ajoutés aux études précédemment faites par

Lautenberger *et al.* (1992). Ils ont utilisé la technique présentée par Saitou et Nei (1987) ("Neighbor-Joining, NJ method") et la méthode UPGMA.

Comme Lautenberger *et al.* (1992), Laudet *et al.* (1993) montrent l'émergence d'un groupe ETS et d'un groupe ERG, à partir d'un ancêtre commun. L'évolution du groupe ERG n'est pas en contradiction avec le travail de Lautenberger *et al.* (1992), mais il est complété par D Ets-3 et D Ets-6 : ces derniers auraient divergé à partir d'un ancêtre *erg*, qui aurait aussi donné le Su *erg*, ce gène se dupliquant ensuite pour donner les *erg* et les *fli* de Vertébrés.

Dans le groupe ETS, le D *ets-2* apparaît comme un ancêtre de *ets-1/ets-2*, la divergence de ce gène ayant donné le *ets-2* de Vertébré d'une part, et les *ets-1* de Vertébrés et Su *ets-2* d'autre part. Le Su *ets-2* est ainsi très proche des *ets-1* de Vertébrés. La duplication *ets-1/ets-2* se serait ainsi faite après la divergence Arthropodes/Vertébrés, donc la séparation Protostomiens/Deutérostomiens, mais avant l'apparition de la lignée des Amphibiens, puisque chez le xénope les gènes *ets-1* et *ets-2* sont déjà séparés.

Notons que les analyses de Lautenberger *et al.* (1992) et Laudet *et al.* (1993) se réfèrent exclusivement à la phylogénie de la séquence en acides aminés des domaines ETS et non pas à l'étude complète des protéines.

### **3 - Contribution de Degnan *et al.* (1993) à la phylogénie des domaines ETS**

Les travaux de Degnan *et al.* (1993) ont permis, par PCR avec des oligonucléotides dégénérés, d'isoler des séquences *ets* de nombreux organismes Invertébrés. Ces auteurs ont montré que les gènes *ets* constituent une famille de gènes qui existeraient déjà chez les Métazoaires primitifs.

Chez les Deutérostomiens, leurs travaux ont porté sur un Tunicien, l'ascidie *Styela montereyensis*, et un Echinoderme, *Patiria miniata* : chez ces deux organismes ils ont montré la présence d'un membre du groupe ETS, homologue à 100 % de la protéine Ets-1 humaine. Cependant, leurs séquences ne correspondant qu'aux acides aminés 39 à 72 du domaine ETS, il n'est pas possible de savoir s'il s'agit de Ets-1 ou de Ets-2, ou de l'ancêtre de ces deux gènes. Les gènes *ets-1* et *ets-2* se dupliquant et se différenciant avec l'avènement des Vertébrés, il serait intéressant de poursuivre ces travaux afin de savoir si la duplication *ets-1/ets-2* s'est déjà produite chez un Cordé primitif tel que l'ascidie.

Chez les Protostomiens, Degnan *et al.* (1993) ont mis en évidence la présence de membres du groupe ETS chez le Crustacé *Cancer antennarius*, les Mollusques *Haliotis rufescens* et *Hinnites giganteus* et l'Annélide

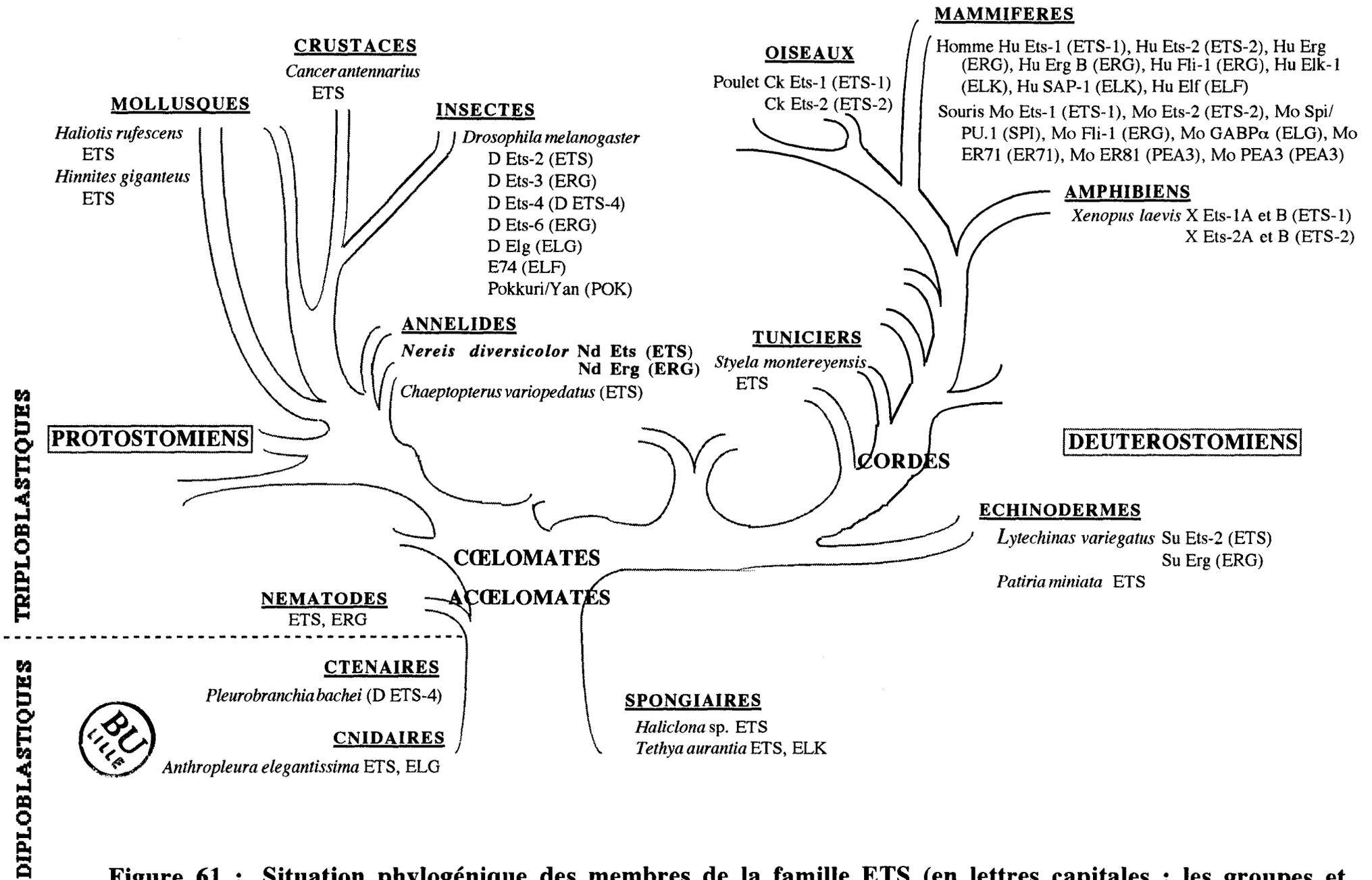


Figure 61 : Situation phylogénique des membres de la famille ETS (en lettres capitales : les groupes et sous groupes d'appartenance à la famille ETS).

*Chaetopterus variopediatus*. Chez deux Nématodes dont les espèces ne sont pas précisées, ils ont noté la présence des groupes ETS et ERG.

Ces auteurs ont enfin montré que la famille ETS était déjà diversifiée chez les Métazoaires Diploblastiques. En effet, chez le Cténaire *Pleurobranchia brachei* ils observent la présence d'une séquence proche de D Ets-4 ; chez le Cnidaire *Anthropleura elegantissima* ils ont isolé des membres des groupes ETS et ELG ; chez les éponges *Haliclona* sp et *Tethya aurantia* ils ont mis en évidence des séquences pouvant appartenir aux groupes ETS et ELK.

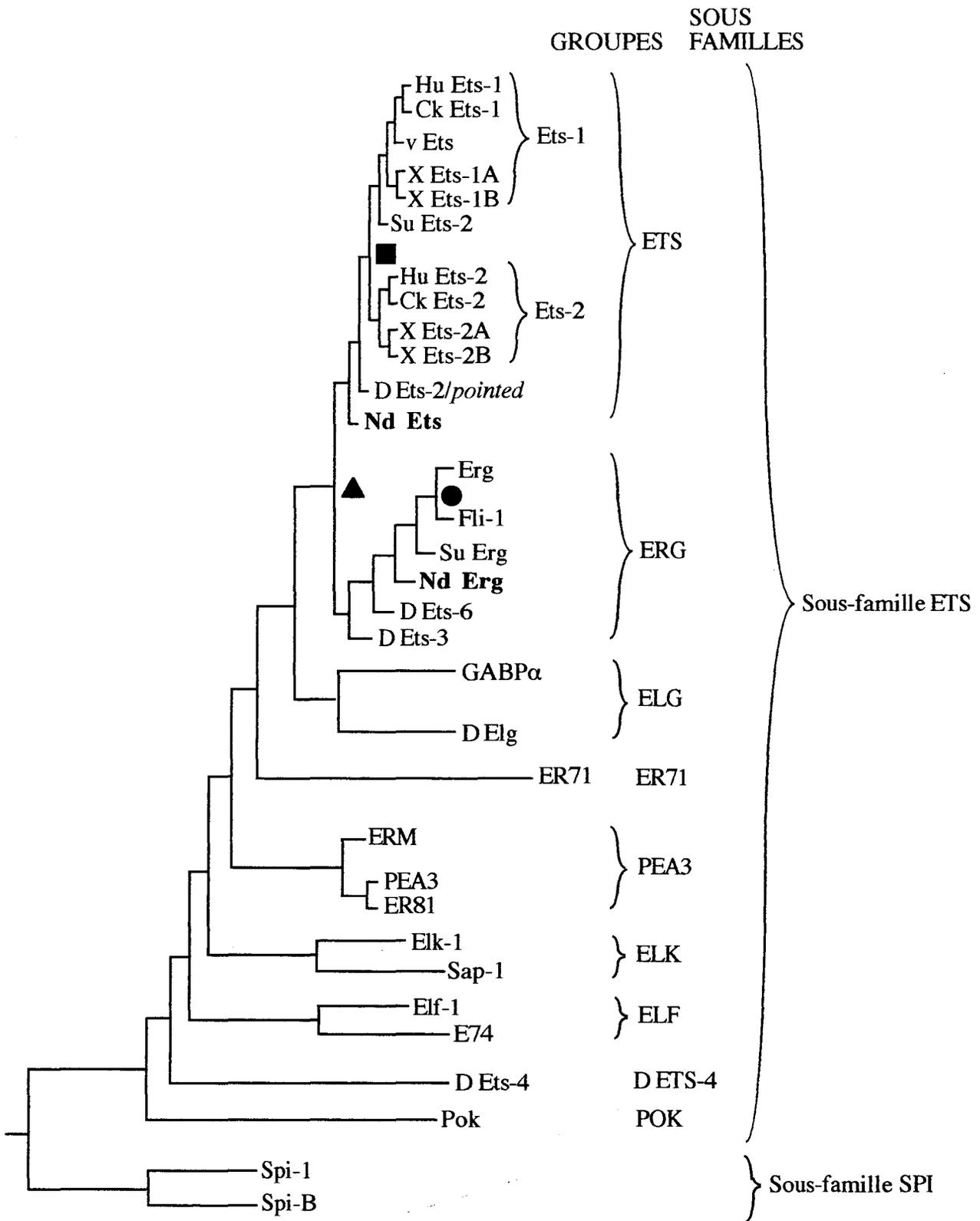
Le groupe D ETS-4 n'était jusqu'ici représenté que chez la drosophile : la présence d'un membre de ce groupe chez un Cténaire peut laisser penser que ce type de gène n'est donc pas apparu avec les Arthropodes, mais pourrait être beaucoup plus primitif.

Cependant, ils n'ont pu détecter de membres de la famille ETS chez des organismes autres que les Métazoaires. Les recherches chez le Protozoaire cilié *Tetrahymena thermophila*, le Dinoflagellé *Alexandrium fundyense*, le champignon *Aspergillus nidulans*, l'amibe *Trichoshaerium*, et une plante, le pois, n'ont permis d'amplifier aucune séquence.

La famille ETS pourrait donc s'être rapidement diversifiée avec l'apparition des Métazoaires (Figure 61). Les protéines Ets seraient ainsi essentielles à l'établissement de l'environnement cellulaire imposant la nécessité de coordonner le taux d'expression génique dans l'espace et le temps, et l'agencement d'un cytosquelette exigé pour la construction d'un organisme multicellulaire.

L'existence de séquences apparentées aux groupes ETS et ELK chez l'éponge *Tethya aurantia* laisse supposer une divergence de ces deux groupes dès l'apparition des premiers Métazoaires Diploblastiques. La présence des groupes ETS et ELG chez l'anémone *Anthropleura elegantissima* permet de faire une hypothèse identique pour ces groupes. L'existence de séquences pouvant être placées dans les groupes ETS, ELK, ELG et D ETS-4, chez les Métazoaires Diploblastiques, montre que ces quatre groupes étaient différenciés avant la divergence Protostomiens/Deutérostomiens. Degnan *et al.* (1993) suggèrent qu'il existerait un gène "D ets-4 like" chez les Vertébrés, et un gène "elk like" chez la drosophile, qui resteraient donc à découvrir.

Ces données nous montrent que seules les divergences *erg/fli* et *ets-1/ets-2* seraient récentes, apparaissant vraisemblablement avec l'émergence des Vertébrés. Si nos observations ne se limitent pas au domaine ETS, on observe que les protéines Erg et Fli sont fortement homologues sur toute leur longueur, alors que les protéines Ets-1 et Ets-2 sont plus divergentes, tout particulièrement au



**Figure 62 : Arbre phylogénique de la famille ETS présentant la situation phylogénique de Nd Ets et Nd Erg, et les trois duplications importantes apparues chez les Métazoaires.**

niveau de leur région centrale. On peut donc penser que si les deux duplications *erg/fli* et *ets-1/ets-2* se situent dans le même Embranchement, la duplication *erg/fli* est vraisemblablement plus tardive dans l'évolution.

Enfin, chez un ver Nématode et chez le ver Plathelminthe *Notoplana acticola*, Degnan *et al.* (1993) ont identifié des séquences qu'ils n'ont pu, après traduction en acides aminés, classer dans aucun des groupes de la famille ETS. On trouve dans ces deux séquences un tryptophane conservé dans les autres membres de la famille ETS et on observe qu'elles ont une homologie de 67 % avec la protéine Ets-1 humaine. Les séquences isolées par ces auteurs sont cependant trop courtes pour nous indiquer s'il s'agit d'un fragment de gène pouvant être effectivement fonctionnel chez ces organismes.

### **B - Situation phylogénique des gènes *erg* et *ets* de la néréis**

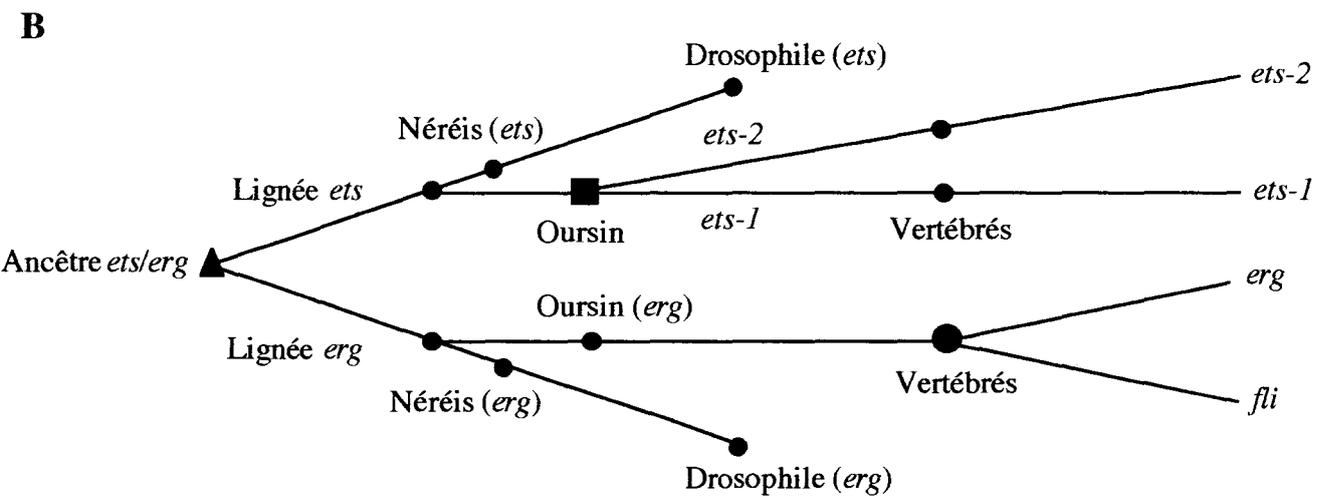
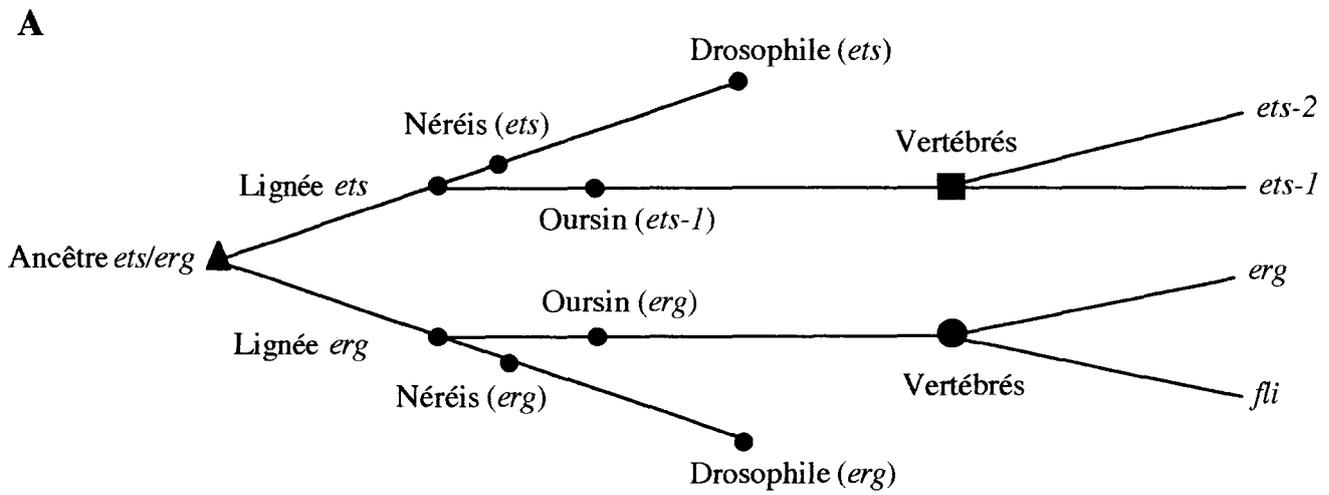
Les séquences Nd Erg et Nd Ets ont été analysées par la technique précédemment utilisée par Laudet *et al.* (1993) (Figure 62).

La séquence Nd Erg apparaît comme très proche des autres séquences Erg d'Invertébrés actuellement connues (Su Erg d'oursin, D Ets-3 et D Ets-6 de drosophile). La néréis étant située à proximité de la divergence entre Protostomiens et Deutérostomiens, l'Erg de néréis est proche de l'Erg ancestral de ces deux lignées évolutives (Figure 62).

L'étude de Nd Ets permet de montrer sa parenté avec l'Ets de drosophile, les deux gènes *ets* de drosophile et de néréis ayant un ancêtre commun. Nous pouvons faire l'hypothèse qu'il y aurait eu une première divergence d'un *ets* ancestral, sans duplication, donnant l'*ets* de la lignée Annélides-Arthropodes, ou l'*ets* de la lignée Echinodermes-Vertébrés, l'*ets* d'oursin étant le précurseur de l'*ets-1* des Vertébrés. Une duplication avec divergence serait à l'origine de l'*ets-2* apparu avec l'émergence des Vertébrés (Figure 63A). Cependant, on ne peut exclure l'existence potentielle d'un *ets-2* chez l'ousin, situant ainsi la duplication *ets-1/ets-2* avec l'émergence des Echinodermes (Figure 63B).

Ces analyses confortent et complètent les travaux antérieurs concernant la phylogénie du domaine ETS et laissent présager qu'il n'existerait qu'un seul gène *erg* et un seul gène *ets* chez les Annélides Polychètes.

Degnan *et al.* (1993) ont mis en évidence la présence d'un gène *ets* chez les Métazoaires Diploblastiques mais n'ont pas montré l'existence d'un gène *erg* chez ces organismes. On ne peut cependant pas exclure l'existence d'un gène *erg* chez les Diploblastiques, le nombre de séquençages effectués par ces auteurs ne prenant pas systématiquement en compte la totalité des membres pouvant être



**Figure 63 : Modèle d'évolution des domaines ETS des sous-groupes ETS-1, ETS-2, ERG et FLI.**

amplifiés par PCR. Dans ce cas, la divergence *ets/erg* à partir d'un ancêtre commun se serait produite avec l'apparition des Diploblastiques, voire antérieurement. Il est aussi possible d'envisager qu'il n'existe pas de gène *erg* chez les Diploblastiques. Dans ce cas, c'est lors de la séparation Protostomiens/Deutérostomiens que se serait produite la divergence *ets/erg*, les gènes *ets* et *erg* de néréis étant les ancêtres de gènes qui se sont diversifiés ensuite.

La diversification de la famille ETS s'est en partie déjà faite chez les Métazoaires Diploblastiques, comme le montrent les travaux de Degnan *et al.* (1993). Chez les Deutérostomiens la diversification la plus tardive s'est dans doute produite lors de l'avènement des Vertébrés, avec l'apparition des *erg* et *fli* mis en évidence chez les Mammifères (homme et souris) et des *ets-1* et *ets-2* sans doute présents chez tous les Vertébrés. Des résultats non publiés indiquent qu'il existerait un gène *fli* chez le xénope (Stiegler *et al.*, 1993). Dans l'état actuel de nos connaissances, ces quatre sous-groupes de gènes *erg*, *fli*, *ets-1* et *ets-2* n'existent donc pas chez les Invertébrés.

Chez les Protostomiens, on peut s'interroger sur la présence de deux membres du groupe ERG chez la drosophile, D Ets-3 et D Ets-6. Etant donné que D Ets-6 est le plus éloigné des membres du groupe ERG, on peut faire deux hypothèses : soit le gène *erg* s'est dupliqué chez les Insectes, ou même dans l'Embranchement des Arthropodes, l'un des deux gènes ayant rapidement divergé ; soit le gène D *ets-6* est en fait un membre ancien de la famille ETS que l'on doit retrouver dans des organismes Protostomiens plus primitifs. Dans ce cas, son homologue doit exister chez la néréis.

Des observations analogues peuvent être faites pour les gènes D *ets-4*, E 74, D *elg* et *pokkuri/yan* de drosophile.

Le gène D *ets-4* existe chez la drosophile et Degnan *et al.* (1993) ont montré son existence chez les Métazoaires Diploblastiques. Il existerait donc un "D *ets-4* like" à découvrir chez la néréis.

Le gène *pokkuri/yan* présent chez la drosophile n'a pas d'homologue dans d'autres organismes. Ce gène peut être, d'après les études phylogéniques, considéré comme un membre primitif de la famille ETS et a aussi vraisemblablement un homologue chez les Annélides.

Les gènes E74 et *elg* de la drosophile possèdent respectivement les gènes *elf* et GABP $\alpha$  comme homologues chez les Vertébrés : leurs ancêtres chez la néréis restent donc à découvrir.

Dans l'étude de la phylogénie de la famille des gènes *ets*, les Annélides Polychètes apparaissent situées à un carrefour important du point de vue évolutif. La recherche des éventuels homologues des gènes *D ets-4*, *D ets-6*, *D elg*, *pokkurilyan*, *E74* et des autres membres de la famille, est un élément dans la compréhension de la diversification de cette famille de facteurs de transcription, dans le phylum des Protostomiens. La connaissance complète des produits des gènes *Nd ets* et *Nd erg* contribuera à la compréhension de l'évolution des différents domaines fonctionnels des protéines *Ets* de Vertébrés.

Ces observations reposent cependant sur l'état actuel de nos connaissances de la famille ETS et restent spéculatives. Ainsi, la diversité de la famille ETS observée chez les Arthropodes peut être en relation avec les exigences de ces organismes, et les gènes découverts chez la drosophile n'ont pas nécessairement un homologue dans des organismes plus archaïques ou appartenant à d'autres lignées évolutives.

L'évolution de la famille des proto-oncogènes *src* montre bien les limites d'une recherche prospective basée sur les analyses phylogéniques. Comme dans la famille ETS, les gènes *src* sont apparus dès les Métazoaires Diploblastiques, puisque Otilie *et al.* (1992) ont trouvé quatre membres de la famille SRC chez l'éponge *Spongilla lacustris*. Or, l'éponge est un organisme plus primitif que le Coelentéré *Hydra* où on ne connaît qu'un membre de la famille SRC, ou que la drosophile qui ne possède que deux gènes de la famille SRC. Les gènes *src* de l'éponge ont donc divergé à l'intérieur même du phylum des éponges, à partir d'un gène *src*, ancêtre des gènes *src* de drosophile et de Vertébrés, et les quatre gènes *src* de *Spongilla lacustris* n'ont pas donné naissance à autant de familles de gènes chez les organismes supérieurs.

Enfin, si la famille ETS est une famille ancienne de facteurs de transcription, déjà diversifiée chez les Métazoaires Diploblastiques, on peut s'interroger sur la date d'apparition de ces gènes. Degnan *et al.* (1993) n'ont pu les mettre en évidence chez les Protozoaires ou chez une plante, le pois. Cependant, le choix des oligonucléotides qu'ils ont utilisés pour rechercher par PCR ces séquences ne permet peut être pas d'isoler des ancêtres lointains de gènes qui ont ensuite fortement divergé et se sont diversifiés. Pour notre part, nous serons donc plus prudents que Shenk et Steele (1993) qui affirment que les résultats négatifs de Degnan *et al.* (1993) sont aussi solides que les résultats positifs. Nous pensons que cette recherche de l'"*ets* ancestral" devrait être approfondie en modifiant les conditions expérimentales.

A l'appui de cette hypothèse, on peut faire référence aux travaux concernant le proto-oncogène *myb* : cette famille de facteurs de transcription existe chez les Protozoaires et les plantes (Paz-Ares *et al.*, 1987). Le domaine de liaison à l'ADN des protéines Ets a une structure apparentée à celui des protéines Myb, consistant en la répétition de motifs renfermant des résidus tryptophane. Les caractéristiques structurales des domaines de liaison à l'ADN des protéines Ets, Myb et HMG, ont été récemment étudiées par Laget *et al.* (1993). Les deux familles de protéines Ets et Myb (et peut être aussi la famille des protéines HMG), pourraient être issues d'un ancêtre commun ne contenant qu'un motif avec des résidus tryptophane, ce motif subissant une duplication ou une triplication pour donner les structures des domaines ETS et MYB. Mais une autre hypothèse permet de considérer ces familles de gènes comme résultant d'une évolution convergente nécessitée par la fonction du domaine de liaison à l'ADN. L'une ou l'autre de ces hypothèses ne pourra être vérifiée que par l'éventuelle mise en évidence de membres de la famille ETS chez les Protozoaires, les champignons et les plantes.

# CONCLUSION

Les produits des proto-oncogènes de la famille ETS sont des protéines faiblement activatrices de la transcription qui agissent généralement en coopération avec d'autres protéines dans différents types cellulaires, chez les Vertébrés et les Invertébrés. Les protéines Ets ont en commun un domaine de liaison à l'ADN, le domaine ETS, particulièrement bien conservé du point de vue de la séquence protéique et qui permet de les classer en différents groupes de protéines (Laudet *et al.*, 1993).

Dans le cadre d'une étude phylogénique nous avons recherché les homologues du proto-oncogène *ets* chez l'Annélide Polychète *Nereis diversicolor*. Les Annélides Polychètes sont des Cœlomates primitifs où des caractères archaïques tels que la métamérie et les cavités cœlomiques persistent dans l'organisme adulte.

Chez *Nereis diversicolor* nous avons identifié deux gènes appartenant aux groupes ETS et ERG de la famille ETS.

L'étude phylogénique montre qu'il n'existerait qu'un seul membre du groupe ETS et qu'un seul membre du groupe ERG chez les Annélides Polychètes, comme chez l'oursin ; ces deux groupes de gènes existent déjà chez les Métazoaires Diploblastiques mais n'ont pas été mis en évidence chez les Protozoaires (Degnan *et al.*, 1993). La diversification de la famille ETS s'est ensuite faite chez les Protostomiens d'une part, et chez les Deutérostomiens d'autre part, la plupart des groupes de la famille ETS étant déjà présents chez les Invertébrés.

La partie identifiée de l'Ets de néréis présente une homologie de plus de 85 % avec les séquences en acides aminés des protéines Ets-1 humaine, Ets-2 de drosophile et Ets-2 d'oursin.

Nous avons une connaissance plus complète de l'*erg* de néréis. Le fragment d'ADN séquencé montre qu'il contient dans le même exon la région codant la totalité du domaine ETS et une partie de la région codant le domaine R caractéristique des protéines Erg. Le domaine ETS présente une forte homologie avec les domaines ETS des protéines Erg et Fli de Vertébrés et d'Invertébrés. L'homologie la plus importante est trouvée avec l'Erg d'oursin (91,2 % d'homologie) et avec une séquence de 33 acides aminés d'un ver Nématode publiée par Degnan *et al.* (1993). Le domaine R présente de l'homologie avec celui des protéines Erg et Fli de l'homme, de la souris et de l'oursin, seuls organismes où ce domaine est actuellement connu. L'homologie est cependant plus

importante entre les domaines R d'Invertébrés d'une part et de Vertébrés d'autre part.

L'expression des gènes *ets* et *erg* de néréis a été mise en évidence dans des régénérats de l'extrémité caudale, mais elle semble très faible et n'est pas nécessairement restreinte à cette partie du corps de l'animal. De fait, l'étude de l'expression de ces deux gènes devra être précisée. L'expression et l'activité des protéines Ets ont surtout été recherchées chez les Vertébrés : les protéines Ets-1 sont exprimées préférentiellement dans les tissus et cellules lymphoïdes et la protéine Erg-2 humaine dans les tissus lymphoïdes et myéloïdes. Les granulocytes des Annélides Polychètes sont habituellement considérés comme l'équivalent des cellules lymphoïdes des Vertébrés. Il sera important de caractériser les protéines Ets et Erg de néréis et de rechercher leur éventuelle expression dans les granulocytes.

La situation phylogénique particulière des Annélides Polychètes, à proximité des Cœlomates primitifs qui ont évolué pour donner les organismes que nous connaissons aujourd'hui, en font un matériel de choix pour la compréhension de l'origine des différents groupes de protéines de la famille ETS, de leurs relations fonctionnelles et du rôle des différents domaines de ces protéines.

Dans les protéines Erg et Fli, chez la néréis comme chez l'oursin et les Vertébrés, le domaine ETS est contigu au domaine R, ces deux domaines étant codés par un même exon, chez la néréis.

Comme l'ont montré des expériences *in vitro* concernant les protéines Erg et Fli humaines (Rao *et al.*, 1993 ; Siddique *et al.*, 1993), la partie carboxy-terminale du domaine R est un domaine transactivateur qui pourrait être impliqué dans les mécanismes d'activation de la transcription par des interactions intermoléculaires avec d'autres protéines. Nous montrons une bonne conservation de la partie amino-terminale du domaine R de l'Erg de néréis, particulièrement avec l'oursin, semblant indiquer que la partie fonctionnelle de ce domaine ne se trouve pas uniquement localisée à la région transactivatrice, et qu'un autre domaine fonctionnel doit être caractérisé à l'extrémité amino-terminale du domaine R.

La plupart des gènes eucaryotes sont constitués de petits exons séparés par de longues séquences introniques. La diversification des protéines résulterait d'un réassortiment des exons, ce qui explique que le domaine ETS puisse être en position amino ou carboxy-terminale, selon les protéines. Parallèlement à ce mouvement d'exons se produisent des duplications de gènes ou d'exons, puis leur

divergence. Dorit *et al.* (1990) estiment que seulement 1000 à 7000 exons seraient ainsi nécessaires à la construction de toutes les protéines. Chaque exon codant partiellement ou totalement les domaines fonctionnels des protéines, comme pour les domaines ETS et R, l'étude phylogénique dans différents organismes permettra de préciser le rôle des domaines fonctionnels et de comprendre leur évolution.

Cette hypothèse nous amène à souligner que jusqu'ici les études phylogéniques de la famille ETS proposées (Lautenberger *et al.*, 1992 ; Laudet *et al.*, 1993) concernent plus précisément l'évolution phylogénique du domaine ETS.

Le mode d'action des proto-oncogènes *ets* reste problématique et la connaissance que nous en avons repose sur des études faites *in vitro*. Des organismes plus simples que les Vertébrés, tels que les Métazoaires Diploblastiques et les Coelomates primitifs peuvent aider à comprendre le mode d'action des proto-oncogènes. Dans cette optique les Annélides Polychètes disposent de fonctions particulièrement intéressantes à étudier : un système immunitaire primitif, qui se développera ensuite chez les Vertébrés où une expression des protéines Ets est bien décrite ; une possibilité de régénération qui sera ensuite perdue au cours de l'évolution, les organismes supérieurs ne conservant que la capacité de cicatrisation. L'étude de l'expression des gènes *ets* et de leurs produits lors de ces phénomènes biologiques contribuera à la compréhension du mode d'action des proto-oncogènes.

De nombreuses molécules sont parfaitement conservées dans différents phyla de Métazoaires (Shenk et Steele, 1993). Les gènes de la famille ETS ont été identifiés chez la néréis en utilisant la PCR avec des oligonucléotides judicieusement choisis d'après les séquences hautement conservées des protéines. Cette même démarche peut être appliquée au criblage de familles de gènes d'organismes variés et a d'ailleurs été récemment utilisée pour montrer la diversification et l'extension de la famille des gènes *Hox* au cours de l'évolution des Cordés (Pendleton *et al.*, 1993). La connaissance de ces gènes pourra permettre de relier leurs caractéristiques structurales à des événements précis de leur histoire évolutive. Ce type d'étude peut avoir la prétention d'améliorer la compréhension des mécanismes moléculaires de l'évolution animale ayant permis la conservation et la diversification des domaines fonctionnels des protéines.

# MÉTHODES

## **I - Extraction des acides nucléiques**

### **A - Extraction de l'ADN génomique**

L'ADN est extrait à partir des vers marins entiers. Les animaux sont lavés plusieurs fois dans du tampon PBS ( $\text{NaH}_2\text{PO}_4$  0,05 M ;  $\text{Na}_2\text{HPO}_4$  0,05 M ;  $\text{NaCl}$  0,96 M ; pH 7,5).

Les animaux sont ensuite plongés dans de l'azote liquide puis conservés à  $-70^\circ\text{C}$  s'ils ne sont pas utilisés immédiatement.

#### **Protocole d'extraction**

Les animaux congelés sont broyés manuellement dans un mortier en présence d'azote liquide. Le broyage s'effectue jusqu'à l'obtention d'une fine poudre blanche en présence de TNE (Tris-HCl 50 mM pH 8,0 ;  $\text{NaCl}$  0,1 M ; EDTA (0,2 M) et de SDS (Sodium Dodecyl Sulfate) à 0,5 % final.

La majorité des protéines et des contaminations est éliminée par une extraction phénolique (phénol-chloroforme 1:1) suivie d'une extraction chloroformique. Après centrifugation, la phase supérieure aqueuse contenant l'ADN est récupérée. L'ADN est précipité par 1/10<sup>e</sup> de volume d'acétate de sodium 3M et deux volumes d'éthanol. Les fibres formées sont enroulées autour d'une pipette Pasteur à bout recourbé puis elles sont dissoutes dans du TE (Tris-HCl 10 mM pH 8,0 ; EDTA 1 mM).

— Cette préparation est purifiée par action d'ARNase (50  $\mu\text{g}/\text{ml}$  final) pendant 2 heures à  $37^\circ\text{C}$  suivie d'une incubation avec de la protéinase K (200  $\mu\text{g}/\text{ml}$  final) pendant 1 heure à  $37^\circ\text{C}$ . Après une nouvelle extraction au phénol puis au chloroforme, l'ADN est récupéré comme précédemment par précipitation alcoolique. L'ADN dissous dans du TE est ensuite dialysé contre du TE 0,1X pour obtenir une meilleure purification.

— Alternativement, la purification de l'ADN est effectuée par une ultracentrifugation en gradient de chlorure de césium (1 mg pour 1 ml de solution) en présence de bromure d'ethidium (BET), pendant une nuit à 214 700 g (VTi 65.1, Beckman). Le BET est ensuite éliminé par plusieurs extractions à l'isopropanol. La préparation est finalement dialysée contre du TE 0,1X pendant une nuit à température ambiante (Sambrook *et al.*, 1989).

L'ADN est testé en électrophorèse avant et après digestion enzymatique. Une cinétique de digestion permet de vérifier la bonne qualité de la préparation.

## B - Extraction des ARN

### 1 - Précautions à prendre

Pour éliminer toutes les traces potentielles d'ARNase, les solutions et la vaisselle subissent un traitement spécial.

— Les solutions sont traitées au diéthylpyrocarbonate (DEPC) à raison de 0,1 % pendant une nuit, puis autoclavées à 120°C pendant 120 minutes.

— Les plastiques sont plongés pendant 12 heures dans une solution d'alcool 100 % contenant 0,1 % de DEPC puis autoclavés 20 minutes à 120°C.

— La verrerie est passée 2 heures à 200°C.

### 2 - Extraction des ARN totaux

Plusieurs techniques ont été utilisées. Les extractions ont été effectuées sur différentes parties du corps de l'animal: le prostomium, le metastomium et des régénérats de pygidium coupés à des temps différents.

— Extraction selon la technique de Chirwin *et al* (1979).

500 mg des différents extraits sont lavés dans du PBS, immédiatement placés dans 5 ml de tampon GT et homogénéisés à l'aide d'un homogénéiseur de type Dounce à température ambiante.

Tampon GT:	Isothiocyanate de guanidium	4 M
	Citrate de sodium pH 7,2	25 mM
	N Lauryl sarcosinate de Na <sup>+</sup>	0,5 %
	β mercaptoéthanol	0,2 M

Ce milieu permet la dénaturation des protéines mais aussi l'inhibition des ARNases libérées lors de la lyse cellulaire.

L'homogénat est centrifugé à 5 000 g (JA 20, Beckman) pendant 5 minutes à température ambiante. La fraction soluble est récupérée et du CsCl est ajouté à raison de 2 g pour 5 ml. L'ensemble est déposé sur un coussin de CsCl.

Coussin de CsCl:	CsCl	5,7 M
	EDTA pH7,5	0,1 M

Une ultracentrifugation est ensuite réalisée pendant 18 heures à 148 500 g (SW 40 Ti, Beckman) à 20°C. Un gradient de densité se forme et à l'équilibre, les ARN sédimentent au fond du tube et sont séparés de l'ADN et des

protéines. Le tube est clampé et le culot d'ARN est lavé plusieurs fois à l'éthanol 70 % puis séché et repris par du tampon TE-SDS 1 %.

Le dosage des ARN s'effectue à 260 nm et le degré de pureté est estimé par le rapport de la DO à 280 nm/260 nm qui doit être supérieur à 2 (Sambrook *et al.*, 1989).

— Extraction selon la technique de Chomczynsky et Sacchi (1987).

Cette méthode combine une déprotéinisation au thiocyanate de guanidium et une extraction au phénol-chloroforme.

Les différents extraits d'animaux (environ 100 mg) sont broyés dans un mortier en présence d'azote liquide et de 1 ml de solution dénaturante.

Solution dénaturante :	Thiocyanate de guanidium	4 M
	Citrate de sodium pH 7,0	25 mM
	N Lauryl sarcosinate de Na <sup>+</sup>	0,5 %
	β mercaptoéthanol	100 mM

Ensuite, 0,1 volume d'acétate de sodium 2 M pH 4,0 est ajouté, suivi d'une extraction au phénol-chloroforme (1 volume de phénol et 0,2 volume de chloroforme/alcool isoamylique). Après une centrifugation à 10000 g pendant 10 minutes à 4°C, la phase supérieure qui contient les ARN est récupérée. Les ARN en solution sont précipités par un volume d'isopropanol à -20°C puis centrifugés selon les mêmes paramètres. Le culot est ensuite repris par 0,8 ml de solution dénaturante. Les ARN sont à nouveau précipités par 1 volume d'isopropanol à -20°C. Après centrifugation, le culot d'ARN, lavé à l'éthanol 75 %, est finalement repris dans 100 µl de tampon TE.

### 3 - Extraction des ARN messagers

L'isolement des ARN messagers poly A+ est réalisé avec le "kit" de micro-extraction "QuickPrep Micro mRNA Purification" (Pharmacia-LKB). Les différentes parties du corps de l'animal sont extraites dans un tampon contenant une forte concentration en thiocyanate de guanidium, permettant une inactivation rapide des ARNases endogènes. L'extrait est ensuite dilué dans un tampon d'élution pour permettre une liaison efficace entre les queues polyA+ des molécules d'ARN et les oligo(dT) couplés à la cellulose tout en conservant une activité inhibitrice vis à vis des ARNases. L'extrait est finalement transféré sur une colonne d'oligo(dT) cellulose et les ARNm polyA+ sont récupérés après

centrifugation. Plusieurs lavages dans des tampons de salinités différentes sont effectués et les ARNm polyA+ sont récupérés après passage sur une colonne de purification (MicroSpin Column).

Chaque purification produit environ 6 µg d'ARN dont au moins 90 % sont polyadénylés.

#### **4 - Traduction des ARN en lysat de réticulocytes de lapin**

Cette technique permet de produire *in vitro* les protéines à partir d'ARN messagers, en présence d'acides aminés radiomarqués.

L'intérêt d'utiliser cette technique est ici de vérifier la qualité des ARN messagers préparés à partir du "kit" de microextraction commercialisé par la firme Pharmacia.

5 µg d'ARN messagers sont repris par 3 µl d'eau milliQ.

50 µCi de méthionine et de cystéine marquées au <sup>35</sup>S sont ajoutés ainsi que 20 µl de lysat de réticulocytes (N90, Amersham). Le lysat représente 80 % du volume total de la préparation.

La réaction est réalisée pendant 1 heure à 30°C. Le 17ème de la préparation est déposé sur un gradient de polyacrylamide 10-25 %. La concentration de l'échantillon est préalablement réalisée dans un gel de concentration contenant 5 % de polyacrylamide. La migration est effectuée pendant une nuit à 3 mA.

Le gel est ensuite coloré dans une solution de bleu de Coomassie (Bleu de Coomassie 0,1 % ; Isopropanol 25 % ; Acide acétique 10 %) pendant 45 minutes puis décoloré dans de l'acide acétique 10 %. Le gel est finalement séché puis mis en autoradiographie avec un film MP-Amersham.

## **II - Préparation des sondes d'ADN**

### **A - Amplification bactérienne**

Les colonies bactériennes contenant les vecteurs sont stockées dans leur milieu de culture en présence de glycérol à -20°C. Une aliquote des cultures sélectionnées est amplifiée sur un plus grand volume et l'addition de chloramphénicol, alors que les bactéries sont en phase de croissance exponentielle, permet d'augmenter la quantité de plasmides recombinés.

## **B - Préparation de l'ADN plasmidique**

Pour préparer l'ADN plasmidique à partir de cultures bactériennes, plusieurs techniques ont été utilisées :

### **1 - Méthode au lysozyme**

Les cultures sont centrifugées et le culot est repris dans une solution de GTE (Glucose 50 mM ; Tris-HCl 25 mM pH 8,0 ; EDTA 20 mM) contenant du lysozyme (1 mg/ml). Du tampon de lyse (NaOH 0,2 N; SDS 1 %) et de l'acétate de potassium 3 M sont ajoutés, le tout est centrifugé et les surnageants filtrés sont précipités par un volume d'isopropanol. Une solution de chlorure de lithium 4 M permet de précipiter une partie des ARN. Les surnageants sont centrifugés après addition de deux volumes d'éthanol et les culots sont repris dans du TE. Après action d'ARNase (50 µg/ml), 30 minutes à 37°C et de protéinase K (200 µg/ml), 2 heures à 37°C, une extraction phénolique puis une extraction chloroformique sont réalisées. L'ADN plasmidique récupéré dans la phase aqueuse est finalement repris dans du TE après précipitation éthanolique.

### **2 - Méthode de lyse alcaline**

La culture est centrifugée et le culot est repris dans une solution de GTE puis du tampon de lyse (NaOH 0,2 N ; SDS 1%) et de l'acétate de potassium sont ajoutés. Le tout est centrifugé et le surnageant est filtré. Un volume d'isopropanol est ajouté et après centrifugation, le culot est repris dans du TE. Les plasmides sont ensuite purifiés par une ultracentrifugation isopycniqne en gradient de chlorure de césium (1 g/ml) en présence de BET pendant 16 heures à 214 700 g (VTi 65-1, Beckman). Après récupération de la bande correspondant aux plasmides, le BET est éliminé par plusieurs extractions à l'isopropanol. L'ADN est finalement dialysé une nuit contre du TE.

## **C - Préparation de l'insert**

Les plasmides recombinés sont digérés par les enzymes de restriction appropriées. Une unité d'enzyme permet de digérer 1 µg d'ADN en 1 heure à 37°C dans le tampon correspondant à l'activité maximale de l'enzyme.

On sépare ensuite l'insert du plasmide par migration en gel d'agarose. L'insert est ensuite récupéré de différentes façons.

### **1 - Fenêtre en gel d'agarose à 0,3 %**

L'insert est récupéré dans une fenêtre d'agarose à bas point de fusion (Biorad) à 0,3 %, et ce morceau de gel est soumis à une succession de

congélations-décongélations afin de casser les mailles du gel. Après filtration sur laine de verre ou sur filtre 0,22  $\mu\text{m}$  (Millex, Millipore) et précipitation éthanolique, le culot d'ADN plasmidique est repris par du TE.

## 2 - Membrane de DEAE cellulose (Schleicher et Schuell)

Après séparation du plasmide et de l'insert, une membrane de DEAE cellulose est placée sous l'insert. La migration électrophorétique est poursuivie et l'ADN adsorbé sur la membrane est ensuite élué par chauffage à 68°C dans un tampon salé, puis précipité à l'éthanol et centrifugé. Le culot est finalement repris dans du TE.

Tampon d'éluion :	Tris-Hcl pH 7,5	20 mM
	EDTA	1 mM
	NaCl	1,5 M

## 3 - Electroélution

Après migration électrophorétique, le gel est découpé au niveau des inserts et les morceaux sont placés dans les enceintes prévues à cet effet au niveau de l'électroéluéur (Biorad). L'appareil est mis sous tension à 100 volts pendant une heure. L'ADN sort du gel et est piégé dans un tampon hautement salé (NaCl 1 M ; Tris-HCl 20 mM pH 7,4 ; EDTA 1 mM). Il est ensuite précipité et récupéré dans du TE.

# III - Marquage des sondes d'ADN

## A - Marquage avec un radioisotope

### 1 - Marquage de sondes double brin

Deux techniques ont été utilisées :

— Multiamorçage au hasard ("Random priming")

20 à 100 ng d'ADN sont dénaturés (10 minutes à 100°C) et additionnés de  $^{32}\text{P}$  dCTP (50  $\mu\text{Ci}$ , 3000 Ci/mmmole), d'oligonucléotides amorces, de dATP, dGTP, dTTP et d'enzyme (T7 DNAPolymerase) (Prime System, Clontech). La réaction de polymérisation incorporant l'isotope radioactif s'effectue pendant 20 minutes à 37°C. Après arrêt de la réaction par addition

d'EDTA 2 mM final, la sonde est purifiée par gel filtration sur colonne de Séphadex G 50 (Pharmacia™).

La sonde est éluée dans du tampon TE et le comptage est effectué sur une aliquote par scintillation liquide. L'activité spécifique de la sonde est de 5 à  $7.10^8$  cpm/ $\mu$ g.

— Coupure-Réparation ("Nick translation")

100 ng de sonde sont utilisés pour le marquage en présence de  $^{32}\text{P}$  dCTP (50  $\mu$ Ci, 3000 Ci/mme), d'ADN polymérase (fragment de Klenow) et d'ADNase I (Nick Translation kit, Amersham). La réaction s'effectue à 16°C pendant 40 minutes et est stoppée par addition de tampon STE (Tris-HCl 10 mM pH 8,0 ; NaCl 100 mM ; EDTA 1 mM pH 8,0) contenant de la protéinase K à 200  $\mu$ g/ml. La sonde est purifiée par passage sur une colonne échangeuse d'ions de DEAE cellulose (Elutip, Schleicher et Schuell) après addition d'ADN entraîneur. L'élution s'effectue par un tampon fortement salé. La sonde ainsi purifiée est dénaturée par de la soude 0,6 N final puis neutralisée par de l'HCl 0,6 N final, le témoin de neutralisation étant le rouge de phénol. L'activité spécifique de la sonde est approximativement de  $3.10^8$  cpm/ $\mu$ g.

## 2 - Marquage d'oligonucléotides

Deux techniques ont été utilisées:

— Marquage par la terminale désoxynucléotidyl transférase

La terminale transférase catalyse l'addition répétitive de désoxynucléotides 5' triphosphates à l'extrémité 3'OH libre d'acides nucléiques, sans nécessiter la présence d'une matrice. Cette réaction produit une queue homopolymérique.

4 pmole d'oligonucléotide sont additionnés à 100  $\mu$ Ci de  $^{32}\text{P}$  à 5000 Ci/mme, de chlorure de cobalt et de 36 unités de terminale transférase, dans un tampon approprié. La réaction s'effectue à 37°C pendant 30 minutes puis l'oligonucléotide marqué est purifié par passage sur une colonne de purification spécifique Nensorb (Kit Oligonucléotide 3' End Labeling System, Dupont/NEN).

L'activité spécifique est calculée après comptage d'une aliquote en scintillation liquide.

— Marquage par la polynucléotide kinase du phage T4

La polynucléotide kinase du phage T4 permet l'addition d'un groupement phosphate sur une extrémité 5'OH. La réaction s'effectue dans un

tampon approprié sur 50 pmole d'oligonucléotide en présence de 250  $\mu\text{Ci}$  de  $\gamma^{32}\text{PdCTP}$  et de 10 unités de kinase pendant 1 heure à 37°C. L'oligonucléotide est ensuite purifié et son activité spécifique calculée comme précédemment.

### **B - Marquage sans radioisotope**

Les sondes d'ADN double brin homologues ont été marquées par multiamorçage au hasard utilisant un nucléotide sur lequel est fixé la digoxigénine (UTP-DIG, Non radioactive Labeling and Detection Kit, Boehringer Mannheim). La détection immunologique se fait à l'aide d'un anticorps anti-digoxigénine couplé à la peroxydase.

## **IV - Méthode d'analyse des acides nucléiques**

### **A - Analyse de l'ADN**

#### **1 - Techniques électrophorétiques**

L'ADN est visualisé sur un gel d'agarose dont la concentration est fonction de la taille des fragments, en présence de bromure d'éthidium (BET) et après migration à voltage constant dans un tampon acétate (TAE 1 X : Tris acétate 40 mM; EDTA 1mM) ou borate (TBE 1 X : Tris borate 90 mM; EDTA 2 mM).

Le dosage de l'ADN s'effectue par spectrophotométrie ou par fluorimétrie. Si la concentration d'ADN est trop faible, une évaluation peut être faite par comparaison avec un marqueur de taille dont on connaît la concentration.

#### **2 - Transfert-hybridation selon Southern (1975)**

##### **a) Transfert**

Les transferts ont été réalisés sur des membranes de nitrocellulose ou de nylon chargé positivement.

— Transfert sur membrane de nitrocellulose (Schleicher et Schuell).

Les fragments d'ADN sont dénaturés dans le gel (NaOH 0,5 M ; NaCl 1,5 M), deux fois quinze minutes puis neutralisés (Tris-HCl 0,5 M pH 7,5 ; NaCl 1,5 M), deux fois quinze minutes.

Pour les transferts d'ADN génomique, un prétraitement du gel dans de l'HCl 0,25 M pendant 10 minutes est nécessaire. Ce traitement permet une dépurination et donc casse les grands fragments qui sont plus facilement transférables.

Une feuille de nitrocellulose imprégnée de SSC 2X (SSC 20 X : NaCl 3 M ; Citrate de sodium 0,3 M pH 7,0) est placée sur le gel et recouverte d'un papier Whatman 3M. Du papier absorbant est déposé sur l'ensemble ainsi qu'un poids. Le transfert s'effectue pendant 12 à 16 heures dans une solution saline de SSC 20X. Le flux de tampon montant par capillarité permet le transfert des fragments d'ADN du gel à la membrane de nitrocellulose. L'ADN est ensuite fixé irréversiblement à la membrane par 1 à 2 heures de cuisson à 80°C ou par irradiation aux ultraviolets pendant 2 minutes. La membrane peut être conservée à 4°C en milieu déshydraté.

— Transfert sur membrane de nylon chargé positivement (Hybond N+, Amersham)

Le gel est directement mis en transfert, sans prétraitement, dans de la soude 0,4 M pendant 16 heures. Après transfert, la membrane est rincée dans du SSC 2X puis séchée et ainsi conservée. Le traitement à la soude ne nécessite pas de fixation ultérieure de l'ADN à la membrane.

#### **b) Hybridation pour sonde radioactive**

Les milieux, la température d'hybridation et la stringence des lavages varient en fonction:

- de la nature homologue ou hétérologue de la sonde
- de la nature du marquage utilisé

#### **1) Sonde double brin hétérologue marquée en "random priming" ou en "nick translation"**

La solution de préhybridation est composée de :

SSC	3 X
Denhardt's	1 X
Tampon phosphate	25 mM pH 7,5
ADN de sperme de saumon	100 µg/ml
Formamide	50 %

Denhardt's 50 X :	Ficoll	1 %
	Polyvinylpyrrolidone	1 %
	Sérumalbumine bovine	1 %

La préhybridation s'effectue pendant 2 heures à 42°C.

La solution d'hybridation diffère de la solution de préhybridation par la concentration en SSC qui est 7X. La sonde dénaturée est ajoutée de façon à avoir  $10^7$  cpm par filtre de 150 cm<sup>2</sup>. L'hybridation s'effectue une nuit à 42°C.

## 2) Sonde double brin homologue marquée en "random priming" ou en "nick translation"

Dans ce cas, la stringence est augmentée en jouant sur la concentration saline.

La solution de préhybridation est identique à celle citée précédemment. Par contre, dans la solution d'hybridation, on retrouve les mêmes composants mais la concentration en SSC est de 4X finale.

La température d'hybridation est calculée en fonction de l'homologie, du pourcentage en bases G et C et du pourcentage de formamide selon la formule suivante:

$$T_m = 79,8^\circ\text{C} + (18,5 * \log \text{Na}^+) + (58,4^\circ\text{C} * \%G+C) - (820/\text{pb en duplex}) - 0,5 * \% \text{ formamide}$$

## 3) Sonde oligonucléotidique dégénérée ou homologue

Le milieu d'hybridation adapté à l'utilisation d'oligonucléotides est le suivant:

SSC	6 X
Denhardt's	5 X
SDS	0,1 %
Tampon phosphate	25 mM pH 7,5
ADN de sperme de saumon	200 µg/ml

Pour un oligonucléotide dégénéré, la température est fixée à 55°C, alors que pour un oligonucléotide homologue, la température est calculée selon la formule suivante:

$$T_m = 81,5^\circ\text{C} - (16,5 * \log \text{Na}^+) + 0,41 (\%G+C) - (820/\text{pb en duplex})$$

### c) Hybridation pour sondes froides

Les solutions de préhybridation et d'hybridation utilisées pour la détection à la digoxygénine sont les suivantes :

SSC	5 X
Blocking Reagent (Boehringer)	2 %
Formamide	5 %
N Laurylsarcosinate de Na <sup>+</sup>	0,1 %
SDS	0,02 %

### d) Lavages

Les conditions de lavage ne varient qu'en fonction de la nature homologue ou hétérologue de la sonde, quel que soit le marquage utilisé. La stringence est augmentée pour les sondes homologues par une diminution de la concentration saline et par une augmentation de la température.

Pour des sondes hétérologues, la membrane est lavée 3 fois 15 minutes dans du SSC 2X ; SDS 0,1 % à 41°C suivi de 2 lavages de 15 minutes dans du SSC 0,1X ; SDS 0,1 % à la même température.

Pour des sondes homologues, la seule modification concerne la température qui est augmentée à 52°C.

### e) Autoradiographie

Les membranes séchées sont mises en autoradiographie sur un Hyperfilm MP commercialisé par la firme Amersham.

### f) Détection immunologique

Pour l'hybridation réalisée avec le système DIG (Boehringer), la révélation s'effectue avec un anticorps reconnaissant la digoxygénine, qui est couplé à la phosphatase alcaline. La coloration apparaît après incubation avec le NBT (nitroblue tetrazolium salt) et le X-phosphate (5-bromo-4-chloro-3-indolyl phosphate), substrat de la phosphatase alcaline.

## B - Analyse des ARN

### 1 - Gel dénaturant

La quantité nécessaire d'ARN est reprise dans le tampon de dénaturation suivant:

Formamide	50 %
Formaldéhyde	18 %
MOPS	1 X
ARN repris dans de l'eau milli Q	

MOPS 10X:	MOPS	0,2 M
	(Acide 4 morpholino propane sulfonique)	
	Acétate de Na	0,05 M pH 7,0
	Na <sub>2</sub> EDTA	0,01 N

L'échantillon est dénaturé pendant 5 minutes à 68°C pour éliminer les structures secondaires, puis placé dans la glace. Du bleu de charge (bleu de charge 6 X : Bleu de Bromophénol 0,25%; Xylène cyanol 0,52%; Glycérol 30%) et du BET sont ajoutés. Le mélange est déposé sur un gel d'agarose dénaturant:

Agarose	1,2 %
Formaldéhyde	17 %
MOPS	1 X

La migration s'effectue dans du MOPS 1X à raison de 5 volts par cm de gel.

## 2 - "Northern blot"

Le transfert est réalisé sur membrane de nitrocellulose ou de nylon chargé positivement dans du tampon SSC 20X pendant 16 heures. Les ARN y sont ensuite fixés par cuisson du filtre pendant 1 heure à 80°C.

## 3 - Hybridation

La solution de préhybridation est composée de:

SSPE		4 X	
Formamide		50 %	
Denhardt's		5 X	
SDS		0,1 %	
ARNt d'E. coli		3 µg/ml	
SSPE 20 X	NaCl	3,6 M	pH 7,4
	NaH <sub>2</sub> PO <sub>4</sub>	200 mM	
	EDTA	0,5 M	

## Souches bactériennes et milieux de culture pour la construction et le criblage des banques d'ADN génomique et complémentaire.

### Milieux de culture

LB (milieu de Luria Bertani)		
Tryptone	10 ‰	pH 7,5
Extrait de levure	5 ‰	
NaCl	5 ‰	

Souche bactérienne	Milieu de culture
DP 50	LB
C 600 <i>Hfl</i>	LB
SRB	LB
Y 10 88	LB + Amp

### Génotypes

DP 50	(F <sup>-</sup> <i>fhu A53 dapD8 lacY1 glnV44 (supE44) Δ (gal-uvrB) 47 tyr T58 (supF58) gyrA29 (Nal<sup>r</sup>) Δ (thyA ) 57hsdS3 (r<sub>k</sub>, m<sub>k</sub><sup>-</sup>)).</i>
<i>mcrA1</i> C600 <i>Hfl</i>	(F <sup>-</sup> <i>hsdR supE44 thr-1 leuB6 thi-1 lacY1 fhuA21 mcrB hflA150 : Tn 10 (Tet<sup>r</sup>).</i>
SRB	<i>e14<sup>-</sup> (mrc), Δ (mcrCB-HSD SMR-mrr) 171, sbcC, recJ, uvrC, umuC : Tn (Kan<sup>r</sup>), supE44, lac, gyrA96, relA1, thi-1, endA1 (F' proAB lacI<sup>q</sup>ZΔM15).</i>
Y1088	F <sup>-</sup> <i>Δ(lac) U169supE supF hsdR (r<sub>k</sub>-m<sub>k</sub><sup>+</sup>) metB trpR fhuA21 proC : Tn5 (pMC9; Tet<sup>r</sup> Amp<sup>r</sup>).</i>

L'hybridation s'effectue dans le même milieu additionné de la sonde à raison de 70 000 cpm/cm<sup>2</sup> de filtre et à la température définie pour la sonde homologue utilisée (cf IV-A-b-2).

#### **4 - Lavages**

Les filtres sont lavés 3 fois 15 minutes dans du SSC 2X, SDS 0,1 % à 60°C puis 2 fois 10 minutes dans du SSC 0,1X, SDS 0,1 % à la même température.

### **V - Construction de banques d'ADN génomique**

#### **A - Différentes banques construites**

Trois banques d'ADN génomique de Nereis ont été construites en collaboration avec Agnès BÈGUE (Laboratoire d'Oncologie Moléculaire-Institut Pasteur de Lille):

- une première banque a été construite par digestion partielle de l'ADN par l'enzyme de restriction *Sau3A* dans le vecteur phagique  $\lambda$  EMBL4. Les souches bactériennes utilisées pour la propagation des phages sont la DP50 lysogène pour le bactériophage P2 et la DP50.

- la deuxième banque a été construite par digestion totale de l'ADN par l'enzyme de restriction *HindIII* dans le vecteur phagique  $\lambda$  NM1149. La souche bactérienne utilisée est ici la C600Hfl.

- la dernière banque a été construite par digestion partielle de l'ADN par l'enzyme de restriction *Sau3A* dans le vecteur phagique  $\lambda$  DASH II. Les souches bactériennes utilisées sont la SRB lysogène pour le bactériophage P2 et la SRB.

#### **B - Préparation des bras du bactériophage pour les phages $\lambda$ EMBL4 et $\lambda$ DASHII**

Quelques microgrammes d'ADN du phage ( $\lambda$  EMBL4 ou  $\lambda$  DASHII) sont digérés par l'enzyme de restriction *BamHI* dans le tampon salé approprié. Une ultracentrifugation en gradient de sucrose permet de séparer la région centrale du génome qui contient des gènes non essentiels à la multiplication du phage, des bras droit et gauche qui possèdent les gènes essentiels. Les bras droit et gauche sont récupérés en différentes fractions par une minipompe et sont

précipités à l'éthanol. L'ADN des bras est repris dans du TE et dialysé contre de l'eau.

Tampons pour gradient de sucrose:

Tris-HCl	20 mM pH 7,5
NaCl	1 M
EDTA	5 mM
Sucrose	10 %

Tris-HCl	20 mM pH 8,0
NaCl	1 M
EDTA	5 mM
Sucrose	40 %

### **C - Préparation de l'ADN génomique**

Pour deux des banques, une aliquote d'ADN génomique de néréis digéré par l'enzyme *Sau3A* est soumise à une cinétique de digestion.

A partir de 75 µg d'ADN purifié, une digestion partielle par l'enzyme *Sau3A* procure des fragments d'assez grande taille, qui sont séparés sur un gradient de sucrose par ultracentrifugation pendant 17 à 18 heures.

Les différents fragments sont ensuite récupérés à l'aide d'une minipompe par fractions de 0,5 ml et testés en électrophorèse pour n'utiliser que les fractions contenant les fragments de 10 à 20 kpb.

Pour la deuxième banque, l'ADN a été digéré totalement par l'enzyme *HindIII* puis purifié par extraction phénolique et dialysé.

### **D - Encapsidation *in vitro***

Les fractions correspondant aux fragments de 10 à 20 kpb ou l'ADN digéré totalement sont précipités à l'éthanol et l'ADN est mis en ligation avec les bras des bactériophages ( $\lambda$  EMBL4 ou  $\lambda$  DASHII) ou avec l'ADN du bactériophage  $\lambda$  NM1149 digéré par l'enzyme de restriction *HindIII*. Environ 200 ng d'ADN sont mis en présence de 800 ng de bras, de 3 unités de ligase T4 et du tampon approprié à 37°C pendant 2 heures. L'ensemble est ensuite précipité à l'éthanol et mis en encapsidation (Gigapack Gold, Stratagene). La technique consiste en l'encapsidation de l'ADN par deux lysats complémentaires de bactéries, chacun étant incapable d'encapsider à l'état isolé. La complémentation est effectuée en présence de l'ADN du phage.

L'encapsidation s'effectue pendant 2 heures à température ambiante à l'obscurité. Le mélange est complété à 1 ml par du TMG (Tris-HCl 20 mM pH 7,5 ; NaCl 100 mM ; MgSO<sub>4</sub> 10 mM ; Gélatine 0,01 %). Les

bactériophages recombinants sont mis en adsorption avec des bactéries DP50 (ou SRB) lysogènes pour P<sub>2</sub> ou des bactéries C 600*Hfl* (2-5.10<sup>8</sup> bactéries) par fractions d'environ 50 000 phages à 42°C pendant 10 minutes. Du milieu de culture contenant de l'agar (4 ‰) est ajouté dans chacun des tubes et le tout est étalé sur huit boîtes de Pétri contenant du milieu de culture et de l'agar (12 ‰), les exigences de la souche (acide diamino pimélique 50 µg/ml final, thymidine 50 µg/ml final pour la DP50) et du MgSO<sub>4</sub> 10 mM final. Les boîtes sont incubées, une nuit à 37°C.

Le lendemain, les couches supérieures molles contenant les plages de lyse sont récupérées dans un tube et malaxées. 1/5e de volume de chloroforme est ajouté et les bactériophages sont récupérés par centrifugation.

La génothèque ainsi constituée est conservée à 4°C. Une partie est aliquotée et stockée à -20°C dans du glycérol 20 % final.

## **VI - Construction d'une banque d'ADN complémentaire**

### **A - Préparation des ARNm**

Les ARNm ont été extraits de régénérats et de métamères avec le "kit" de microextraction (cf.I, B, 3) (Pharmacia). Les ARNm sont ici purifiés par un deuxième passage sur une colonne d'oligo(dT) cellulose.

### **B - Synthèse des brins d'ADN complémentaire (Kit Time Saver cDNA synthesis de Pharmacia)**

Le premier brin d'ADN complémentaire est synthétisé avec un oligonucléotide oligo(dT)<sub>12-18</sub> possédant un site de restriction *NotI* à son extrémité 5', en présence des dNTP, de BSA (Sérum Albumine Bovine) et de transcriptase réverse (M-MLV RT, Moloney-Murine Leukemia Virus Reverse Transcriptase).

Après dégradation du brin ARN de l'hybride ADNc/ARN par l'ARNase H, le deuxième brin d'ADN est synthétisé par l'ADN polymérase I en présence des dNTP puis une purification est effectuée par passage sur une colonne. Des adaptateurs *EcoRI* sont ensuite ajoutés aux extrémités des ADNc suivi d'une digestion par l'enzyme de restriction *NotI*, générant ainsi des fragments d'ADNc possédant une extrémité *NotI* et une extrémité *EcoRI*. Une deuxième purification par passage sur colonne est effectuée.

### **C - Ligation des ADNc dans le vecteur $\lambda$ gt11D**

Deux  $\mu\text{g}$  de vecteur  $\lambda$  gt11D (Pharmacia) sont mis en ligation avec une partie des ADNc. Le vecteur est déphosphorylé et permet le clonage des fragments *Not I/EcoRI* qui sont ainsi orientés.

### **D - Encapsidation *in vitro***

L'encapsidation de la moitié des produits de ligation est réalisée par l'intermédiaire de lysats complémentaires commercialisés par la firme Stratagene. L'étalement et l'amplification de la banque sont réalisés de la même façon que celle décrite au V-D, mais la souche bactérienne utilisée est ici *E. coli* Y 1088.

## **VII - Criblage des banques d'ADN génomique et d'ADN complémentaire**

### **A - Titre de la banque**

La banque est préalablement titrée pour connaître la quantité de bactériophages par ml. Plusieurs dilutions sont effectuées et des aliquotes étalées après adsorption dans la souche bactérienne appropriée.

### **B - Etalement et repérage des clones**

Le criblage de la banque s'effectue d'abord grossièrement par plusieurs étalements de différentes dilutions sur des boîtes contenant le milieu solide adapté. Une membrane de nitrocellulose 0,45  $\mu\text{m}$  (Schleicher et Schuell) est déposée sur chaque boîte et les phages se fixent à la membrane en quelques minutes. La membrane est ensuite dénaturée dans une solution de NaOH 0,2 N ; NaCl 1,5 M pendant 5 minutes puis neutralisée dans une solution de NaCl 1,5 M ; Tris-HCl 0,5 M, pH 7,5. La cuisson de la membrane est réalisée pendant 1 heure à 80°C.

### **C - Hybridation**

Les mêmes conditions d'hybridation que celles citées en IV-A-2 sont utilisées, selon que l'on travaille avec une sonde homologue ou hétérologue.

### **D - Isolement des clones**

Les clones potentiels répondant par un spot autoradiographique sont purifiés: la zone correspondante est carottée avec une pipette Pasteur puis placée dans du TMG. Des dilutions sont effectuées et plusieurs étalements sont réalisés.

Les boîtes les plus représentatives sont transférées, hybridées et mises en autoradiographie. Deux étapes de purification par dilution sont souvent nécessaires avant d'obtenir des plages de lyse isolées qui, après transfert et hybridation, se révèlent positives.

### **E - Extraction de l'ADN phagique**

Les plages de lyse positives sont carottées avec une pipette Pasteur et mises en adsorption à 42°C avec des bactéries (à raison de  $1.10^9$  bactéries pour  $1.10^7$  phages). Après adsorption, du milieu LB est ajouté et le tout est mis sous agitation à 37°C. 4 à 5 heures après, la lyse des bactéries par les phages est amorcée et le petit stock constitué sert à effectuer un stock sur un plus grand volume. La lyse s'effectue après environ 5 heures à 37°C. Quelques gouttes de chloroforme sont ensuite ajoutées ainsi que du polyéthylèneglycol 10 % final et du NaCl 0,5 M final (qui permettent la précipitation des phages). Après une nuit à 4°C, la culture est centrifugée et le culot repris dans du TMN. Après extraction chloroformique, la phase supérieure contenant les phages est additionnée de chlorure de césium (à raison de 1 g pour 1 ml). Une ultracentrifugation de 5 heures à 189 400 g (SW 50.1, Beckman) permet de récupérer la bande correspondant aux particules phagiques (la bande est visible à partir de  $1.10^{12}$  particules phagiques). Deux dialyses d'une demi-heure dans du TNE sont effectuées. L'ADN est extrait après action de la protéinase K (200 µg/ml ; 5 minutes à 37°C) et de sarcosinate de sodium 1 % (15 minutes à 37°C), suivie d'une extraction phénolique et enfin précipité par trois volumes d'éthanol. Pour une meilleure purification, une dialyse contre de l'eau Milli Q est réalisée pendant une nuit. L'ADN est ensuite digéré par différentes enzymes de restriction.

## **VIII - Amplification enzymatique par la réaction de polymérisation en chaîne (PCR)**

### **A - Conditions d'utilisation**

#### **1 - Amorces**

Plusieurs amorces ont été synthétisées pour les différentes manipulations :

— des oligonucléotides dégénérés

OLA 5' GG(A/C) AGT GGA CCC AT(C/T) CAA CTG 3'

OLB 5' TGG GGC AAG AAG AAA AAC AAG CCG 3'  
 D-OL1 5' G CT CTA G ACT G (G/C)TGG (C/G) A(A/G) TT (C/T) 3'  
 D-OL2 5' G CT CTA G ACT G A (T/C) TT (T/C/A/G) TC (A/G/G) TA  
 (A/G) TTCAT 3'

— des oligonucléotides spécifiques

Nd-OL3 5' AAC TCG CAC ATC ATC ACC 3'

Nd-OL4 5' CTG GTC GAC CCG GAT GAA 3'

## 2 - Conditions d'amplification

— Avec des amorces dégénérées

Les premières PCR ont été réalisées sur de l'ADN génomique et sur des aliquotes de la banque d'ADN génomique  $\lambda$  DASHIII, dans des conditions peu stringentes.

Le milieu d'amplification contient:

ADN	500 ng
MgSO <sub>4</sub>	4 mM
dATP,dTTP,dGTP,dCTP	12,5 mM
Amorces	20 ng
Taq polymérase (Vent (exo-), Biolabs)	1 unité
Glycérol	5-10 %

L'échantillon est d'abord dénaturé dans le milieu 10 minutes à 94°C puis 30 cycles sont effectués:

Dénaturation	1 minute à 94°C
Hybridation	1 minute à 52°C
Elongation	2 minutes à 72°C

Les 30 cycles sont poursuivis par une élongation finale des brins non terminés de 10 minutes à 72°C.

Pour les différentes PCR effectuées, les conditions d'amplification ont parfois été modifiées pour augmenter la spécificité et selon les oligonucléotides utilisés. Ces modifications concernent la concentration en Mg<sup>2+</sup>, la température d'hybridation et le temps d'élongation.

— Avec des amorces homologues

Des amorces "reverse" et "forward" (Promega) complémentaires des extrémités des sites de clonage multiples des plasmides pUC

sont utilisées pour vérifier un clonage ou pour isoler un fragment à l'aide d'un autre oligonucléotide spécifique interne.

Dans ce cas, le même cycle est effectué mais la température d'hybridation est augmentée à 55°C pour diminuer les réactions aspécifiques.

### **3 - Amplification sur colonie bactérienne**

Une colonie bactérienne est reprise dans 500 µl d'eau et chauffée à 100°C pendant 5 minutes. La réaction de PCR s'effectue dans le tampon précédent sur 20 µl de la solution.

## **B - RT PCR**

La PCR est utilisée pour amplifier l'ADN complémentaire synthétisé à partir des ARN messagers.

### **1 - Synthèse des brins codant des ADN complémentaires**

Les premiers brins d'ADN complémentaire sont synthétisés à partir d'une amorce oligo(dT)<sub>18</sub> d'un oligonucléotide dégénéré (D-OL2) ou d'un oligonucléotide spécifique (Nd-OL3).

La réaction s'effectue en présence d'1 µg d'ARN messagers préalablement chauffés 2 minutes à 70°C, de l'amorce (2 µM), de dNTP 0,5 mM chacun et de 20 unités d'inhibiteur d'ARNase dans un tampon adapté (Tris-HCl 50 mM pH 8,3, KCl 75 mM, MgCl<sub>2</sub> 3 mM). 200 unités de transcriptase réverse (M-MLV Reverse Transcriptase, Moloney-Murine Leukemia Virus) (Clontech) sont ajoutées et l'incubation dure 1 heure à 42°C. La réaction est arrêtée par chauffage à 94°C pendant 5 minutes.

L'ARN des hybrides ADN/ARN est ensuite hydrolysé par incubation avec deux unités d'ARNase H (Gibco BRL) pendant 20 minutes à 37°C. L'enzyme est inhibée par chauffage à 100°C pendant 10 minutes.

### **2 - Synthèse d'une queue poly(A) au niveau de l'extrémité 3' du brin d'ADN complémentaire**

La moitié du mélange d'ADN complémentaire est utilisé pour la synthèse de la queue poly(A). La réaction s'effectue en présence d'ATP 5 µM, de MgCl<sub>2</sub> 10 mM et de 15 unités de terminale desoxynucléotidyl transférase (Stratagene) pendant 1 heure à 37°C dans le tampon suivant :

## Souches bactériennes et milieux de culture pour le clonage.

### Milieux de culture

#### LB (milieu de Luria Bertani)

Tryptone	10 ‰		pH 7,5
Extrait de levure	5 ‰		
NaCl	5 ‰		

#### SOB

Tryptone	20 ‰		pH 7,0
Extrait de levure	5 ‰		
NaCl	0,5 ‰		

#### Milieu minimum M 9

Na <sub>2</sub> HPO <sub>4</sub>	12,8 ‰
KH <sub>2</sub> PO <sub>4</sub>	3 ‰
NaCl	0,5 ‰
NH <sub>4</sub> Cl	1 ‰
Glucose	0,4 ‰

	Souche bactérienne	Milieu de culture
Plasmides	JM 83	LB
	DH 5 $\alpha$	LB
Phages M13	JM 109	SOB
	TG 1	SOB
	(milieu de sélection M 9)	

Cacodylate de K <sup>+</sup>	100 mM pH 7,2
CaCl <sub>2</sub>	2 mM
DTT	0,2 mM

La réaction est arrêtée en ajoutant de l'EDTA 40 mM pH 8,0 suivie d'une extraction chloroformique et d'une précipitation éthanolique. L'ADNc est finalement repris dans 50 µl de TE.

### **3 - Amplification enzymatique par PCR**

La PCR est réalisée selon le cycle décrit en VIII-A sur le quart de la préparation d'ADNc, en utilisant comme amorces l'oligo(dT)<sub>18</sub> et un de nos oligonucléotides selon les cas envisagés.

### **4 - 5' RACE "Rapid Amplification of cDNA 5' Ends"**

Le système RACE génère des ADNc en utilisant la PCR pour amplifier des copies de la région désirée entre une séquence dans le transcrit et l'extrémité 3' ou 5'.

Pour l'extrémité 5', la transcription inverse est effectuée en utilisant une amorce spécifique du gène. Un homopolymère est ensuite ajouté aux extrémités des produits de la réaction de synthèse du premier brin. Finalement, la réaction d'amplification est réalisée en utilisant l'amorce hybride préalablement décrite et une seconde amorce spécifique du gène située en amont de la première.

## **IX - Techniques de clonage.**

### **A - Clonage et sous clonage d'insert**

Le clonage d'insert consiste à intégrer dans un vecteur plasmidique un fragment d'ADN au niveau d'un site de restriction unique et à sélectionner, après transformation bactérienne, les vecteurs recombinés.

#### **1 - Préparation des inserts**

Les inserts à cloner proviennent de fragments plus importants digérés par des enzymes de restriction. En fonction de la nature des extrémités du fragment, différentes stratégies sont utilisées.

Si les fragments portent des extrémités cohésives compatibles avec celles retrouvées dans les sites multiples de clonage du plasmide, les fragments sont clonés tels quels dans le vecteur préparé à cet effet (voir ci-après).

## Génotypes

- JM 83            F- *ara*Δ(*lac-proAB*) *rpsL* (Str<sup>r</sup>) (80dΔ (*lacZ*) M15)).
- DH 5α            *supE44* Δ*lac* U169 (80 *lacZ*ΔM15) *hsdR17* *recA1* *endA1*  
*gyrA96* *thi-1* *relA1*.
- JM 109           F' *traD36* *lacIq*Δ(*lacZ*)M15 *proAB/recA1* *endA1* *gyrA96*  
(Nal<sup>r</sup>) *thi* *hsdR17* (*r<sub>k</sub>-m<sub>k</sub>*+) *supE44* *relA1* Δ(*lac-proAB*).
- TG 1             *supE* *hsd*Δ5 *thi* Δ(*lac-proAB*) F' (*traD36* *proAB*<sup>+</sup> *lacIq*  
*lacZ*ΔM15).

Si les fragments portent des extrémités franches, ils sont clonés dans le polylinker au niveau d'un site à coupure franche comme *SmaI* ou *HincII*.

Enfin, si le fragment ne possède pas d'extrémités compatibles avec les sites uniques de restriction des sites multiples de clonage, celui-ci est préparé comme suit :

- lorsque les enzymes utilisées produisent une extrémité 5' protubérante, le large fragment de Klenow en présence de dNTP crée, par remplissage, des extrémités franches. La réaction s'effectue sur 1 µg d'ADN en présence de 2 unités du large fragment de Klenow et de dNTP à 0,5mM chacun pendant 30 minutes à 37°C.

- dans le cas d'enzymes produisant des extrémités 5', la nucléase S1 est alors utilisée pour digérer les portions d'ADN simple brin et fournir des extrémités franches.

Après ces divers traitements, l'insert est purifié par une électrophorèse sur un gel d'agarose à bas point de fusion et élué comme décrit en II-C-1.

## 2 - Préparation des vecteurs

### a) Digestions enzymatiques

Les vecteurs utilisés pour la majeure partie des clonages sont les pUC 18 et 19, le pGEM 3 Z(-) notamment pour la synthèse de sondes ARN et l'ADN RF (forme répliquative) du phage M13.

Les plasmides comme l'ADN du phage M13 sont digérés par la ou les enzymes appropriées à raison de 1 unité par microgramme de vecteur pendant 1 heure à 37°C dans le tampon adapté.

Si le vecteur doit être digéré par deux enzymes de restriction, une extraction chloroformique est effectuée après action de la première enzyme. La deuxième enzyme est alors ajoutée dans son tampon. Le vecteur est finalement purifié par électrophorèse sur un gel d'agarose (cf. II-C-1).

Après digestion enzymatique, l'enzyme est éliminée par une extraction chloroformique suivie d'une précipitation éthanolique ou inhibée par chauffage à 65°C pendant 10 minutes.

### b) Déphosphorylation

Les plasmides linéarisés par une enzyme de restriction sont déphosphorylés par la phosphatase alcaline bactérienne (BAP). Cela permet d'augmenter la probabilité de ligation de l'insert dans le vecteur en interdisant la

recircularisation du vecteur sur lui-même. La BAP élimine le groupement 5' phosphate des ADN linéaires. 250 unités de BAP sont ajoutés à 0,5 µg d'ADN plasmidique dans le tampon suivant (Tris-HCl 10 mM pH 7,5 ; MgCl<sub>2</sub> 10 mM). La réaction s'effectue pendant 1 heure à 60°C.

Du SDS 0,5 % et de l'EDTA 5 mM permettent de stopper la réaction. De la protéinase K (200 µg/ml) est ajoutée et le mélange est incubé pendant 30 minutes à 37°C. Après extraction chloroformique et précipitation éthanolique, l'ADN plasmidique est repris dans du TE de façon à avoir une concentration finale de 50 µg/ml.

### 3 - Ligation

#### a) Ligation d'extrémités cohésives

Pour une efficacité maximale de cette réaction, il faut utiliser un rapport de 1 mole de vecteur pour 4 à 5 moles d'insert.

Le volume réactionnel adéquat est de 10 à 20 µl et comprend:

de la T4 ADN Ligase	1 unité Weiss
du vecteur	30 à 60 ng
de l'insert	
du tampon de ligation	1 X

Tampon de ligation 10 X :	Tris-HCl	0,5 M pH 7,6
	MgCl <sub>2</sub>	100 mM
	DTT	100 mM
	BSA	500 µg/ml

L'incubation s'effectue à 16°C pendant 3 à 4 heures. L'enzyme est inhibée par chauffage à 65°C pendant 10 minutes.

#### b) Ligation d'extrémités franches

La ligation d'extrémités franches est plus délicate et requiert des conditions particulières (faible concentration en ATP, absence de polyamines, forte concentration en ligase et en extrémités franches).

Le milieu réactionnel est donc adapté: de l'héxamine cobalt est ajouté à raison de 1µM (cette molécule augmente considérablement la capacité de ligation des extrémités franches) et la quantité de ADN ligase du phage T4 est augmentée à 5 unités.

#### 4 - Transformation bactérienne

Les sondes, pour être utilisées, doivent être amplifiées. Pour cela, les plasmides recombinés sont intégrés dans des bactéries rendues compétentes.

##### a) Bactéries compétentes

###### Technique au chlorure de calcium

1 colonie est incubée une nuit sous agitation dans son milieu de culture à 37°C. 5 ml de cette culture sont ajoutés à 50 ml de milieu de culture et incubés environ 2 heures à 37°C jusqu'à l'obtention d'une DO de 0,5 à 600 nm (soit environ 10<sup>8</sup> cellules/ml).

La culture est centrifugée à 2 500 g (GP centrifuge, Beckman) pendant 10 minutes à 4°C. Le culot est repris dans 40 ml de CaCl<sub>2</sub> 50 mM froid et laissé 20 minutes dans la glace. Après une centrifugation selon les mêmes paramètres, le culot est repris par 0,4 ml de CaCl<sub>2</sub> 50 mM. Les bactéries ainsi rendues compétentes sont utilisables pendant 48 heures par aliquotes de 300 µl pour chaque transformation.

Alternativement, le culot peut être repris par 0,4 ml de CaCl<sub>2</sub> 50mM contenant 20 % de glycérol. Les aliquotes de bactéries peuvent alors être conservées plusieurs mois à -70°C.

###### Technique au chlorure de rubidium

Une colonie est reprise dans 5 ml de milieu de croissance et incubée à 37°C jusqu'à l'obtention d'une DO de 0,3 à 550 nm.

Milieu de croissance:	Bactotryptone	2 %
	Extrait de levure	0,5 %
	NaCl	10 mM
	KCl	2,5 mM
	MgSO <sub>4</sub>	0,4 %

100 ml de milieu inoculés avec 1 ml de la préculture sont incubés à 37°C jusqu'à l'obtention d'une DO de 0,48 à 550 nm. Après une centrifugation de 5 minutes à 200 g (JA 14, Beckman) à 4°C, le culot est repris dans 30 ml de tampon Tfb1 et placé dans la glace pendant 20 minutes.

Tampon TfB1:	RbCl <sub>2</sub>	100 mM	pH 5,8
	MnCl <sub>2</sub>	50 mM	
	CaCl <sub>2</sub>	10 mM	
	KAc	10 mM	
	Glycérol	15 %	

Les cellules sont à nouveau centrifugées selon les mêmes paramètres, puis elles sont reprises dans 4ml de tampon TfB2.

Tampon TfB2:	MOPS	10 mM	pH 7,0
	RbCl <sub>2</sub>	10 mM	
	CaCl <sub>2</sub>	75 mM	
	Glycérol	15 %	

Les bactéries rendues compétentes sont stockées à -70°C par aliquotes de 200 µl.

Par cette technique, la fréquence de transformation peut être supérieure à  $5 \cdot 10^8$  colonies transformées par microgramme d'ADN plasmidique superenroulé.

### **Technique utilisée pour les bactéries JM109 ou TG1 en vue d'une transformation par l'ADN RF du phage M13**

Quelques colonies bactériennes sont inoculées dans 30 ml de milieu SOB contenant 20 mM de MgSO<sub>4</sub> à 37°C jusqu'à l'obtention d'une DO de 0,8 à 600 nm.

La culture est centrifugée à 2 500 g (GP centrifuge, Beckman) pendant 10 minutes à 4°C et le culot est repris dans 20 ml de milieu FSB.

Milieu FSB:	AcK	10 mM	pH 7,5
	MnCl <sub>2</sub>	45 mM	
	CaCl <sub>2</sub>	50 mM	
	KCl	100 mM	
	Hexamine CoCl	3 mM	
	Glycérol	10 %	

Le mélange est laissé 10 minutes dans la glace puis est centrifugé selon les mêmes paramètres. Le culot est repris dans 4 ml de milieu FSB auxquels sont ajoutés 140 µl de DMSO (Diméthyl Sulfoxyde). Après 15 minutes dans la glace, 140 µl de β Mercaptoéthanol 2,25 M sont ajoutés, et la suspension est placée 5 minutes dans la glace. Après avoir ajouté à nouveau

140 µl de DMSO, les bactéries ainsi rendues compétentes sont aliquotées par fractions de 50 µl et conservées à -70°C jusqu'à utilisation.

#### **b) Transformation bactérienne**

L'ADN plasmidique est introduit dans les bactéries compétentes par choc thermique (perméation temporaire de la paroi des bactéries).

##### **Transformation par un plasmide**

La moitié du mélange de ligation est ajoutée à une aliquote de bactéries compétentes. Le tout est laissé 20 minutes dans la glace. Le choc thermique est effectué en plaçant ensuite le mélange 90 secondes à 42°C puis immédiatement 2 minutes dans la glace.

1 ml de milieu de culture supplémenté de 16 µl de glucose 10 % est ajouté. L'incubation dure 1 heure à 37°C. Différents volumes sont ensuite étalés sur un milieu gélosé contenant les facteurs de sélection (Ampicilline 50 mg/ml, X-Gal 0,005 % et IPTG 0,05 M).

Les boîtes sont incubées retournées pendant une nuit à 37°C.

##### **Transformation par l'ADN RF du phage M13**

Des bactéries TG1 ou JM109 sont rendues compétentes par la technique de Hanahan (1983).

La moitié du mélange de ligation est utilisé pour la transformation. Après le choc thermique effectué comme cité ci-dessus, 200 µl de bactéries indicatrices (culture d'une nuit), sont ajoutés. Différentes dilutions sont étalées en présence de X-Gal et d'IPTG sur milieu du gélosé LB.

## **5 - Sélection des colonies et vérification du clonage**

#### **a) Plasmides**

Les colonies transformées par un plasmide recombiné apparaissent blanches, car l'insertion modifie le cadre de lecture du gène *lacZ* et empêche la synthèse de l'enzyme  $\beta$  galactosidase. Les vecteurs non recombinés permettent la synthèse de cette enzyme qui peut ainsi dégrader son substrat (X-Gal) entraînant l'apparition de la couleur bleue.

Quelques clones mis en culture sont testés par la méthode dite de Boiling (mini-préparation) (Sambrook *et al.*, 1989) qui permet un isolement

rapide de l'ADN plasmidique pour s'assurer de la présence des plasmides recombinés.

Les plasmides sont alors testés par des digestions enzymatiques appropriées puis visualisés sur un gel d'agarose.

La présence du plasmide recombiné peut aussi être testée par PCR, directement sur colonie bactérienne en utilisant les oligonucléotides "reverse" et "forward" du pUC (cf VIII-A-2).

Si le clonage s'avère bon, une préparation plus importante et plus pure des plasmides est alors effectuée (cf II-B).

#### **b) ADN RE du phage M13**

Une minipréparation de la forme double brin du phage M13 va permettre de vérifier que l'insert est bien cloné.

Une plage repiquée dans 2 ml de milieu de culture est incubée pendant 4 heures à 37°C en présence de 50 µl de bactéries (culture de 8 heures) sous agitation. La culture est centrifugée à 15 000 g (Microfuge, Beckman) pendant 5 minutes à température ambiante. Les bactéries infectées sont culottées alors que les particules phagiques se trouvent dans le surnageant.

Le culot cellulaire est repris dans 100 µl de solution 1 (Glucose 50 mM ; EDTA 10 mM pH 8,0 ; Tris-HCl 25 mM pH 8,0) et 100 µl de solution 2 (NaOH 0,2 N ; SDS 1 %). Après une incubation de 5 minutes à température ambiante, 150 µl d'acétate de potassium 3 M sont ajoutés et le mélange est laissé 5 minutes dans la glace. Une centrifugation de 5 minutes à 15 000 g (Microfuge, Beckman) à 4°C est réalisée. Le surnageant est extrait par un volume de phénol/chloroforme (1/1) puis précipité par 2 volumes d'éthanol.

Après une nouvelle centrifugation, le culot est lavé à l'éthanol 75 % puis repris dans 50 µl de TE. De l'ARNase T1 est ajoutée à raison de 20 µg/ml et des digestions enzymatiques sont effectuées afin de ressortir le fragment cloné.

### **B - Clonage des produits de PCR**

Pour les amplifications enzymatiques réalisées avec des oligonucléotides synthétisés avec un site de restriction au niveau de l'extrémité 5', les produits de PCR sont dialysés sur une membrane 0,022 µm (Millipore) contre du TE pendant 2 heures. Une digestion enzymatique par l'enzyme de restriction (*Xba*I) est effectuée et le produit de PCR est finalement cloné dans un vecteur pUC digéré par la même enzyme et déphosphorylé (cf. IX-A-2). La transformation bactérienne est ensuite réalisée comme décrit en IX-A-4.

Pour les amplifications enzymatiques réalisées avec des oligonucléotides ne possédant pas de site de restriction, les produits de PCR sont dialysés puis un groupement phosphate est ajouté au niveau de l'extrémité 5'. La réaction s'effectue sur 500 ng de produits de PCR dans un tampon supplémenté en ATP 1 mM pendant 45 minutes à 37°C avec 20 unités de polynucléotide kinase du phage T4. La réaction est arrêtée en chauffant 10 minutes à 60°C. Une extraction chloroformique puis une précipitation éthanolique sont effectuées et l'ADN est repris dans du TE. Le clonage dans un vecteur digéré par l'enzyme *SmaI* (extrémités franches) et la transformation bactérienne sont réalisés comme cité au paragraphe IX-A.

## **X - Séquençage selon Sanger *et al.* (1977)**

Cette méthode de séquençage est basée sur la terminaison base spécifique de réactions d'extension d'amorces catalysées par l'ADN polymérase du phage T7. Quatre réactions sont menées parallèlement, chacune contenant l'amorce, la matrice et les quatre désoxynucléotides, et un didéoxynucléotide différent qui termine spécifiquement l'extension.

### **A - Séquençage d'une matrice simple brin**

#### **1 - Préparation de matrices simple brin d'ADN de phage M13**

Les fragments à séquencer sont clonés au niveau d'un site unique de clonage de la forme double brin de l'ADN du bactériophage M13. Ces ADN de phages recombinés sont ensuite introduits dans des bactéries comme décrit en IX-A.

#### **Culture des phages M13 recombinés**

Une colonie bactérienne est relancée dans le milieu de culture SOB jusqu'à une DO de 1 à 600 nm. Cette culture est diluée au cinquantième dans un volume de 12,5 ml et inoculée avec une plaque de M13. L'incubation dure environ 5 heures à 37°C sous agitation. Les bactéries sont ensuite culotées par deux centrifugations successives de 10 minutes à 5 000 g (JA 20, Beckman) à 4°C.

### **Isolement de la matrice simple brin**

Aux 10 ml de surnageant sont ajoutés 2 ml de PEG 20 % et NaCl 2,5 M pour précipiter les phages. La solution est placée 30 minutes dans la glace. Les phages précipités sont collectés par une centrifugation de 15 minutes à 5 000 g (JA 20, Beckman). Après élimination du surnageant, le culot est séché puis repris par 500 µl de TE. Deux extractions au phénol/chloroforme suivies de deux extractions au chloroforme sont réalisées, puis l'ADN est précipité par 0,1 volume d'acétate de sodium et deux volumes d'éthanol pendant une demi-heure à -20°C. Après une centrifugation de 30 minutes à 15 000 g (Microfuge ; Beckman,), le culot est séché et repris par 50 µl de TE. Cette méthode produit environ 80 µg d'ADN.

### **2 - Séquençage d'une matrice simple brin M13**

1,5 à 2 µg d'ADN simple brin sont nécessaires pour la réaction de séquence. L'amorce utilisée pour la réaction peut être soit le Primer Universel, soit un oligonucléotide spécifique du fragment à séquencer.

L'amorce est hybridée à la matrice et à cette solution est ajouté un mélange de marquage, le nucléotide marqué au <sup>35</sup>S et l'ADN polymérase du phage T7. Le tout est réparti dans les quatre tubes contenant les dNTP et respectivement un didéoxynucléotide.

Une aliquote de chacun des échantillons est déposée sur un gel de polyacrylamide dénaturant.

### **B - Séquençage d'une matrice double brin**

Un fragment cloné dans un vecteur plasmidique tel que le pUC peut être séquencé directement en utilisant les amorces "reverse" et "forward" spécifiques du pUC ou un oligonucléotide spécifique du fragment à séquencer.

Préalablement à la réaction d'hybridation de l'amorce, le vecteur double brin doit être dénaturé par de la soude à 0,2 N finale pendant 10 minutes à température ambiante. 1,5 à 2 µg d'ADN sont ici aussi nécessaires pour la réaction de séquence qui se poursuit comme décrit en IX-A

### **C - Electrophorèse**

La migration des échantillons s'effectue sur un gel de polyacrylamide 6 %-urée 7 M à 1400 volts. Le gel est ensuite séché puis directement mis en autoradiographie pendant 1 ou 2 jours.

## **XI - Hybridation moléculaire *in situ***

L'hybridation *in situ* permet de détecter et de localiser des séquences nucléotidiques (ADN et ARNm) au niveau d'une coupe de tissus, d'étalement cellulaires ou de préparation chromosomiques. Elle repose sur la spécificité de l'appariement entre deux séquences nucléotidiques complémentaires, celle à détecter et celle de la sonde (ADN ou ARN).

Les précautions d'utilisation définies pour le travail des ARN sont à prendre pour la fabrication et la manipulation des coupes ainsi que pour la préparation des sondes et l'hybridation.

Les lames sont traitées par une solution de TESPA (3 Amino-Propyl-Tri-Ethoxy-Silane) à 2 % dans de l'acétone et les lamelles siliconées.

### **A - Préparation du matériel histologique**

Les différentes parties du corps de l'animal utilisées pour la préparation de coupes sont:

- des tronçons de métamères
- des régénérats de pygidium à 3, 7, 12 et 21 jours
- des prostomiums

Les différentes pièces sont lavées plusieurs fois dans du PBS puis fixées dans du PAF (paraformaldéhyde à 4 % dans du tampon PBS) pendant 16 heures à 4°C. Elles sont ensuite lavées plusieurs fois dans du PBS puis déshydratées par une série de bains dans l'éthanol à 70 %, 96 % et 100 %, suivie d'un bain dans du butanol.

Les pièces sont finalement incluses dans du Paraplast (paraffine et DMSO). Les blocs ainsi constitués peuvent être conservés plusieurs mois à température ambiante.

Les coupes sont effectuées au microtome avec une épaisseur de 7 µm et récupérées sur les lames passées au TESPA. Elles sont ainsi conservées jusqu'à utilisation à 4°C.

### **B - Marquage radioactif de sondes ARN sens et antisens**

La sonde est clonée dans le vecteur pGEM 3Zf(-) qui contient les promoteurs des ARN polymerases des bactériophages SP6 et T7 dans des orientations opposées et flanquant la région des multiples sites uniques de clonage. Ce vecteur possède le gène *lacZ* qui permet une sélection des clones recombinés par la coloration blanc-bleu.

Ce vecteur permet la synthèse sélective d'ARN à partir de l'un ou l'autre des deux brins d'ADN en fonction de la polymérase utilisée.

Le marquage de la sonde se faisant avec un précurseur radioactif, l'ATP <sup>35</sup>S, la réaction est effectuée dans des tubes siliconnés et en présence de dithiothréitol 10 mM (DTT).

Le plasmide recombiné est linéarisé par une enzyme de restriction appropriée, avant d'être marqué.

Le mélange réactionnel est le suivant:

Tampon de transcription (Riboprobe Gemini System II, Promega)	1 X
DTT	30 mM
ATP	0,4 mM
GTP	0,4 mM
UTP	0,4 mM
RNase Block II (Stratagene)	1 U
ARNt d' <i>E. coli</i>	200 µg/ml
Plasmide linéarisé	2 µg
ATP <sup>35</sup> S (1000Ci/mmmole)	200 µCi
ARN polymérase SP6 (Stratagene)	8 U
ARN polymérase T7 (Stratagene)	2,5 U

Le mélange réactionnel est incubé 1 heure à 37°C. 8 unités d'ARN polymérase SP6 (Stratagene) ou 2,5 unités d'ARN polymérase T7 sont de nouveau ajoutées, puis une incubation de 1 heure à 37°C est réalisée.

Le plasmide et le fragment d'ADN sont ensuite digérés par une ADNase (RQ1 DNase, Promega) dans du tampon RQ1 (Tris-HCl 10 mM pH 7,9 ; NaCl 10 mM ; MgCl 6 mM) pendant 30 minutes à 37°C.

Une extraction au phénol-chloroforme est effectuée, puis la sonde est précipitée pendant 10 minutes à -70°C dans le mélange suivant:

Acétate de sodium	0,2 M
DTT	0,1 M
ARNt d' <i>E. coli</i>	100 ng
Ethanol	2 volumes

Après une centrifugation de 15 minutes à 15 000 g (Microfuge, Beckman), la sonde est reprise par 70 µl d'eau Milli Q, 20 µl ARNt d'*E. coli* (5 µg/ml) et 40 µl de DTT 1 M.

### **C - Purification des sondes**

Les sondes sont purifiées par passage sur une colonne de purification des ARN radiomarqués (Sephadex fine G-50 ; Quick Spin Column, Boehringer).

La longueur de la sonde est caractérisée sur un gel d'agarose dénaturant (formamide/formaldéhyde) (voir "Northern blot" en IV-B-2).

Le dosage de la sonde est effectué sur une aliquote de 2  $\mu$ l par scintillation liquide.

### **D - Hybridation-Lavages**

#### **1 - Déparaffinage et réhydratation des coupes**

Les coupes sont plongées dans un solvant de la paraffine, le xylène, 2 fois 10 minutes à température ambiante, suivi d'une série de bains d'éthanol (96 %, 70 % puis 30 %) et d'eau MilliQ pour permettre la réhydratation.

#### **2 - Elimination du paraformaldéhyde**

Cette étape permet d'éliminer le paraformaldéhyde qui peut inhiber la protéinase K.

Les lames sont plongées 10 minutes à température ambiante dans une solution de Glycine 0,1 M ; Tris-Hcl 0,2 M pH 7,4.

#### **3 - Perméation des tissus**

Les lames sont rincées dans du tampon PK pendant 5 minutes à 37°C.

Tampon PK:	Tris-Hcl	100 mM	pH 8,0
	EDTA	50 mM	

Elles sont ensuite incubées avec de la protéinase K à raison de 1 mg/ml dans le même tampon, 15 minutes à 37°C. Les lames sont rincées brièvement dans de l'eau Milli Q.

#### **4 - Post-fixation**

Cette étape est réalisée dans du PAF pendant 15 minutes à température ambiante, suivie d'un lavage dans du PBS de 5 minutes. Elle permet un meilleur attachement du matériel sur les lames.

## 5 - Acétylation

L'acétylation est effectuée dans le tampon suivant:

Triéthanolamine	0,1 M pH 8,0
Anhydride acétique	0,25 %

pendant 10 minutes à température ambiante. Cela permet la neutralisation des charges négatives et évite l'apparition du bruit de fond.

Les lames sont lavées brièvement dans du SSC 2X puis 2 fois 5 minutes dans de l'eau Milli Q.

## 6 - Deshydratation

Les coupes sont déshydratées par des passages successifs de 5 minutes chacun dans des bains d'éthanol à 30 %, 70 %, 96 % et 100 %.

## 7 - Hybridation

Le mélange d'hybridation est le suivant:

NaCl	300 mM
Tris-Hcl pH8,0	20 mM
EDTA	5 mM
Sulfate de Dextran	10 %
Denhart's	1 X
DTT	100 mM
Formamide	50 %
Sonde	1 à 20.10 <sup>3</sup> cpm/ $\mu$ l

Le milieu d'hybridation est déposé sur la lame à raison de 30  $\mu$ l par demi lame et recouvert d'une lamelle siliconée.

L'hybridation est réalisée à 55°C dans une boîte humidifiée par des mouchoirs absorbants saturés par une solution de 50 % de formamide dans du SSC 4X.

## 8 - Lavages

Les lavages permettent de décrocher la sonde fixée de manière aspécifique et donc de réduire au maximum le bruit de fond.

En fonction de l'homologie de la sonde avec la séquence recherchée, deux facteurs interviennent, la température et la salinité.

Les lames sont plongées dans la solution de lavage 1 pendant 30 minutes puis 1 heure à température ambiante sous agitation puis pendant 30 minutes à 60°C dans la solution 2.

Solution 1:	SSC	4 X
	DTT	10 mM
Solution 2:	NaCl	0,15 M
	Tris-HCl	20 mM pH 8,0
	EDTA	5 mM
	DTT	100 mM
	Formamide	50 %

Une digestion par l'ARNase A est ensuite effectuée. Elle permet d'éliminer spécifiquement les portions simple brin des ARN mal appariés.

Les lames sont plongées dans du tampon ARNase A contenant 20mg/ml d'ARNase A pendant 1 heure à 37°C, puis rincées dans le tampon ARNase A pendant 5 minutes à 37°C.

Tampon ARNase A :	NaCl	0,4 M
	Tris-Hcl pH 7,5	10 mM
	EDTA	0,05 M

Les lavages sont terminés par un bain de 15 minutes à 60°C dans du SSC 2X suivi d'un autre bain de 15 minutes à 60°C dans du SSC 0,1X.

## 9 - Deshydratation

Les lames sont déshydratées par des passages successifs de 5 minutes dans de l'éthanol à 30%, 70%, 96% et 100%.

## E - Autoradiographie

L'émulsion (LM1, Amersham) est fluidifiée à 45°C.

Les lames sont plongées verticalement dans l'émulsion puis séchées et stockées dans une boîte contenant un agent dessiccateur (Silicagel) à l'abri de la lumière pendant une période allant de 30 jours minimum à 3 mois.

Les lames sont développées dans du D19b (Kodak) à 18-20°C pendant 2 minutes et demie, rincées rapidement à l'eau puis fixées par du thiosulfate de sodium pendant 5 minutes. Elles sont lavées 10 minutes à l'eau courante puis 5 minutes dans de l'eau distillée. Les lames sont finalement colorées au Bleu Azur II (0,1 %) et montées dans du XAM (Gurr, BDH).

# **RÉFÉRENCES BIBLIOGRAPHIQUES**

- Albagli, O., Flourens, A., Crépieux, P., Bègue, A., Stéhelin, D. and Leprince, D. (1992).  
Phylogeny of the p68<sup>c-ets-1</sup> amino-terminal transactivating domain reveals some highly conserved structural features. *Oncogene* **7**, 1435-1439.
- Amouyel, P., Gégonne, A., Delacourte, A., Défossez, A. and Stéhelin, D. (1988).  
Expression of *ets* proto-oncogenes in astrocytes in human cortex. *Brain Res.* **447**, 149-153.
- Amouyel, P., Laudet, V., Martin, P., Li, R., Quatannens, B., Stéhelin, D. and Saule, S. (1989).  
Two nuclear oncogenic proteins, P135<sup>gag-myb-ets</sup> and p61/63<sup>myc</sup>, cooperate to induce transformation of chicken neuroretina cells. *J. Virol.* **63**, 3382-3388.
- Anthony-Cahill, S.J., Benfield, P.A., Fairman, R. *et al* (1992).  
Molecular characterization of helix-loop-helix peptides. *Science* **255**, 979-983.
- Anton, I.A. and Frampton, J. (1988).  
Tryptophans in Myb proteins. *Nature* **336**, 719.
- Ascione, R., Thompson, D.M., Thomas, R., Panayiotakis, A., Ramsay, R., Tymms, M., Kola, I. and Seth, A. (1992).  
Influence of nucleotides flanking the -GGAA- core sequence on *ets-1* and *ets-2* DNA-binding activity and the mechanism of *ets-1* autoregulation. *Intern. J. of Onc.* **1**, 631-637.
- Ashburner, M., Chihara, C., Meltzer, P. and Richards, G. (1974).  
Temporal control of puffing activity in polytene chromosomes. *Cold Spring Harbor Symp. Quant. Biol.* **38**, 665-662.
- Aurigemma, R.E., Blair, D.G. and Ruscetti S.K. (1992).  
Transactivation of Erythroid transcription factor GATA-1 by a *myb-ets*-containing retrovirus. *J. Virol.* **66**, 3056-3061.
- Baert, J.-L. and Slomianny, M.-C. (1992).  
Vitellin accumulation and vitellogenin synthesis in relation to oogenesis in *Perinereis cultrifera* (Polychaete Annelida). *Invert. Reprod. Develop.* **21**, 121-128.
- Beaumont, A. and Cassier, P. (1973).  
Biologie animale. Des Protozoaires aux Métaozaires Epithélioneuriens. Tome I, Ed. Bordas.

- Ben-David, Y., Giddens, E.B. and Bernstein, A. (1990).  
Identification and mapping of a common proviral integration site Fli-1 in erythroleukemia cells induced by Friend murine leukemia virus. *Proc. Natl. Acad. Sci. USA* **87**, 1332-1336.
- Ben-David, Y., Giddens, E.B., Letwin K. and Bernstein, A. (1991).  
Erythroleukemia induction by Friend murine leukemia virus : insertional activation of a new member of the ets gene family, Fli-1, closely linked to c-ets-1. *Genes Dev.* **5**, 908-918.
- Beug, H., Hayman, M.J. and Graf, T. (1982).  
Myeloblasts transformed by the avian leukemia virus E26 are hormone-dependent for growth and for the expression of a putative *myb*-containing protein p135 E26. *The EMBO J.* **1**, 1069-1073.
- Beug, H., Leutz, A., Kahn, P. and Graf, T. (1984).  
Ts mutant of E26 leukemia virus allow transformed myeloblasts, but not erythroblasts or fibroblasts, to differentiate at the non permissive temperature. *Cell* **39**, 579-588.
- Bhat, N.K., Fisher, T.J., Fujiwara S., Ascione, R. and Papas, T.S. (1987).  
Temporal and tissue-specific expression of mouse *ets* genes. *Proc. Natl. Acad. Sci. USA* **84**, 3163-3165.
- Bhat, N.K., Komschlies, K.L., Fujawara S., Fisher, R.J., Mathieson, B.J., Gegorio T.A., Yound, H.A., Kasik, J.W., Ozato, K. and Papas T.S. (1989).  
Expression of *ets* genes in mouse thymocyte subsets and T cells. *J. Immun.*, **142**, 672-678.
- Bhat, J.K., Thompson, C.B., Lindsten, T., June, C.H., Fujiawara, S., Koizumi, S., Fisher, R.J. and Papas, T.S. (1990).  
Reciprocal expression of human *ets1* and *ets2* genes during T-cell activation : regulatory role for the protooncogene *ets1*. *Proc. Natl. Acad. Sci. USA* **87**, 3723-3727.
- Bhattacharya, G., Lee, L., Reddy, E.S.P. and Rao, V.N. (1993).  
Transcriptional activation domains of *elk-1*,  $\Delta$ *elk-1* and SAP-1 proteins. *Oncogene* **8**, 3459-3464
- Bister, K., Nunn, M., Moscovici, C., Perbal, B., Baluda, M.A. and Duesberg, P.H. (1982).  
Acute leukemia viruses E26 and avian myeloblastosis virus have related transformation-specific RNA sequences, but different genetic structures, gene products and oncogenic properties. *Proc. Natl. Acad. Sci. USA* **79**, 3677-3681.

- Bosselut, R., Duvall, J.F., Gégonne, A., Bailly, M., Hémar, A., Brady, J. and Ghysdael, J. (1990).  
The product of the *c-ets-1* proto-oncogene and the related *Ets2* protein act as transcriptional activators of the long terminal repeat of human T cell leukemia virus HTLV-1. *EMBO J.* **9**, 3315-3322.
- Boulukos, K.E., Pognonec, P., Bègue, A., Galibert, F., Gesquière, J.C., Stéhelin, D. and Ghysdael, J. (1988).  
Identification in chickens of an evolutionary conserved cellular *ets-2* gene (*c-ets-2*) encoding nuclear proteins related to the products of the *c-ets* proto-oncogene. *EMBO J.* **7**, 697-705.
- Boulukos, K.E., Pognonec, P., Rabault, B., Bègue, A. and Ghysdael, J. (1989).  
Definition of an *Ets1* protein domain required for nuclear localization in cells and DNA-binding activity *in vitro*. *Mol. Cell. Biol.* **9**, 5718-1721.
- Boulukos, K.E., Pognonec, P., Sariban, E., Bailly, M., Lagrou, C. and Ghysdael, J. (1990).  
Rapid and transient expression of *Ets-2* in mature macrophages following stimulation with cMGF, LPS and PKC activators. *Genes Dev.* **4**, 401-409.
- Bova, C.A., Olsen, J.C. and Swanstrom, R. (1988).  
The avian retrovirus env gene family : molecular analysis of host range and antigenic variants. *J. Virol.* **62**, 75-83.
- Boyd, L., O'Toole, E. and Thummel, C.S. (1991).  
Patterns of *E74A* RNA and protein expression at the onset of metamorphosis in *Drosophila*. *Development* **112**, 981-995.
- Boyle, W.J., Lampert, M.A., Lipsick, J.S. and Baluda, M.A. (1984).  
Avian myeloblastosis virus and E26 virus oncogene products are nuclear proteins. *Proc. Natl. Acad. Sci. USA* **81**, 4265-4269.
- Brown, T.A. and McKnight, S.L. (1992).  
Specificities of protein-protein and protein-DNA interaction of GABP $\alpha$  and two newly defined *ets*-related proteins. *Genes Dev.* **6**, 2502-2512.
- Burdett, L.A., Qi S., Chen, Z.Q., Lautenberger, J.A. and Papas, T.S. (1992).  
Characterization of the cDNA sequences of two *Xenopus ets-2* proto-oncogenes. *Nucl.Acids Res.* **20**, 371.
- Burtis, K.C., Thummel, C.S., Weldon Jones, C., Karim, F.D. and Hogness, D.S. (1990).  
The *Drosophila* 74EF early puff contains E74, a complex ecdysone-inducible gene that encodes two *ets*-related proteins. *Cell* **61**, 85-99.

- Bzik, D.J. and Preston, C.M. (1986).  
Analysis of DNA sequences which regulate the transcription of herpes simplex virus immediate early gene 3 : DNA sequences required for enhancer-like activity and response to trans-activation by a virion polypeptide. *Nucleic Acids Res.* **14**, 929-943.
- Chen, J.H. (1985).  
The proto-oncogene *c-ets-1* is preferentially expressed in lymphoid cells. *Mol. Cell. Biol.* **5**, 2993-3000.
- Chen, J.H. (1988).  
Complementary DNA clones of chicken proto-oncogene *c-ets* : sequence divergence from the viral oncogene *v-ets*. *Onc. Res.* **2**, 371-384.
- Chen, J.H. (1990).  
Cloning, sequencing, and expression of mouse *c-ets-1* cDNA in baculovirus expression system. *Onc. Res.* **5**, 277-285.
- Chen, J.H., Jeha, S. and Oka, T. (1993a).  
Negative regulatory elements in the human ETS1 gene promoter. *Oncogene* **8**, 133-139.
- Chen, J.H. and Wright, C.D. (1993b).  
PEA2, oct 1 and oct 2 positively regulate the human ETS 1 promoter. *Oncogene* **8**, 3375-3383.
- Chen, T., Bunting, M., Karim, F.D. and Thummel, C.S. (1992).  
Isolation and characterization of five Drosophila genes that encode an ets-related DNA binding domain. *Dev. Biol.* **151**, 176-191.
- Chen, Z.Q., Kan, N.C., Pribyl, L., Lautenberger, J.A., Moudrianakis E. and Papas, T.S. (1988).  
Molecular cloning of the *ets* proto-oncogene of the sea urchin and analysis of its developmental expression. *Dev. Biol.* **125**, 432-440.
- Chen, Z.Q., Burdett L.A., Seth A.K., Lautenberger, J.A. and Papas, T.S. (1990).  
Requirement of *ets-2* expression for xenopus oocyte maturation. *Science* **250**, 1416-1418.
- Chirgwin, J.M., Przybyla, A.E., Mac Donald, R.J. and Rutter, W.J. (1979).  
Isolation fo biologically active ribonucleic acid from sources enriched in ribonuclease. *Biochemistry* **18**, 5294- 5299.
- Chomczynski, P. and Sacchi, N. (1987).  
Single-step method of RNA isolation by acid guanidium thiocyanate-phenol-chloroform extraction. *Analytical Biochem.* **162**, 156-159.

- Chumakov, A.M., Chen, D.L., Chumakova E.A. and Koeffler, H.P. (1993).  
Localization of the c-ets-2 transactivation domain. *J. Virol.* **67**, 2421-2425.
- Clearly, M.L. (1991).  
Oncogenic conversion of transcription factors by chromosomal translocations. *Cell* **58**, 619-622.
- Cocquerelle, C., Daubersies, P., Majérus, M.A., Kerckaert, J.P. and Bailleul, B. (1992).  
Splicing with inverted order of exons occurs proximal to large introns. *EMBO J.* **11**, 1095-1098.
- Cocquerelle, C., Mascrez, B., Hetuin, D. and Bailleul, B. (1993).  
Missplicing yields circular RNA molecules. *FASEB J.* **7**, 155-160.
- Combaz, A. and Boilly, B. (1974).  
Etude expérimentale et histologique de la régénération caudale en l'absence de chaîne nerveuse chez les *Nereidae* (Annélides Polychètes). *Ann. Embryol. Morphogen.* **VII**, 171-197.
- Crepieux, P., Leprince, D., Flourens, A., Albagli, O., Ferreira, E. and Stéhelin, D. (1993).  
The two functionally distinct amino termini of chicken *c-ets-1* products arise from alternative promoter usage. *Gene expression* **3**, 215-225.
- Dalton, S. and Treisman, R. (1992).  
Characterization of SAP-1, a protein recruited by serum response factor to the c-fos serum response element. *Cell* **68**, 597-612.
- Dang, C.V. and Lee, W.M.F. (1988).  
Identification of the human c-myc protein nuclear translocation signal. *Mol. Cell. Biol.* **8**, 4048-4054.
- Degnan, B.M., Degnan, S.M., Naganuma, T. and Morse, D.E. (1993).  
The *ets* multigene family is conserved throughout the Metazoa. *Nucleic Acids Res.* **21**, 3479-3484.
- Delattre O., Zucman, J., Plougastel, B., Desmaze, C., Melot, T., Peter, M., Kovar, H., Joubert, I., de Jong, P., Rouleau, G., Aurias, A. and Thomas, G. (1992).  
Gene fusion with an ETS DNA-binding domain caused by chromosome translocation in human tumours. *Nature* **359**, 162-165.
- Demuyneck, S., Li, K.W., Van der Schors, R. and Dhainaut-Courtois, N. (1993).  
Amino acid sequence of the small cadmium-binding protein (MP II) from *Nereis diversicolor* (annelida, polychaeta). Evidence for a myohemerythrin structure. *Eur. J. Biochem* **217**, 151-156.

- Desbiens, X., Quéva, C., Jaffredo, T., Stéhelin, D. and Vandebunder, B. (1991).  
The relationship between cell proliferation and the transcription of the nuclear oncogenes *c-myc*, *c-myb* and *c-ets-1* during feather morphogenesis in the chick embryo. *Development* **111**, 699-713.
- De Taisne, C., Gégonne, A., Stéhelin, D., Bernheim, A. and Berger, R. (1984).  
Chromosomal localization of the human proto-oncogene *c-ets*. *Nature* **310**, 581-583.
- Dhainaut, A. (1966a).  
Etude ultrastructurale de l'évolution cytoplasmique au cours des premiers stades de l'ovogenèse chez *N. pelagica*. *C. R. Acad. Sci. Paris* **262**, 2616-2619.
- Dhainaut, A. (1966b).  
Etude ultrastructurale de l'évolution cytoplasmique chez *Nereis pelagica* L. (Annélide Polychète) à l'approche de la maturité sexuelle. *C. R. Acad. Sci. Paris* **262**, 2740-2743.
- Dhainaut, A., Raveillon, B., M'Béri, M., Porchet-Henneré, E. and Demuynck, S. (1989).  
Purification of an antibacterial protein in the coelomic fluid of *Nereis diversicolor* (Annelida, Polychaeta). Similitude with a cadmium-binding protein. *Comp. Biochem. Physiol.* **94C**, 555-560.
- Dhainaut-Courtois, N., Nejmeddine, A., Baert, J.L. and Dhainaut, A. (1987).  
Localisation immunocytochimique en microscopie électronique d'une protéine complexant le cadmium chez un ver marin (*Nereis diversicolor* O.F. Müller). *C. R. Acad. Sci. Paris* **305**, Sér. III, 237-241.
- Diaz, M.O., Le Beau, M.M., Pitha, P. and Rowley, J.D. (1986).  
Interferon and *c-ets-1* genes in the translocation (9 ; 11) (p22 ; q23) in human acute monocytic leukemia. *Science* **231**, 265-267.
- Domenget, C., Leprince, D., Pain, B., Peyrol, S., Li, R.P., Stéhelin, D., Samarut, J. and Jurdic, P. (1992).  
The various domain of *v-myb* and *v-ets* oncogenes of E26 retrovirus contribute differently, but cooperatively, in transformation of hematopoietic lineages. *Oncogene*, **7**, 2231-2241.
- Dorit, R.L., Schoenbach, L. and Gilbert, W. (1990).  
How big is the universe of exons ? *Science* **250**, 1377-1382.
- Durban, E.M. and Boettiger, D. (1981).  
Differential effects of transforming avian RNA tumor viruses on avian macrophages. *Proc. Natl. Acad. Sci. USA* **78**, 3600-3604.

- Durchon, M. (1967).  
Rôle du système nerveux dans la régénération chez les Annélides. *Bull. Soc. Zool. Fr.* **92**, 319-331.
- Duterque-Coquillaud, M., Leprince, D., Flourens A., Henry, C., Ghysdael, J., Debuire, B. and Stéhelin, D. (1988).  
Cloning and expression of chicken p54<sup>c-ets</sup> cDNAs : the first p54<sup>c-ets</sup> coding exon is located into the 40,0 kbp genomic domain unrelated to *v-ets*. *oncogene Res.* **2**, 335-344.
- Duterque-Coquillaud, M., Niel, C., Plaza, S. and Stéhelin, D. (1993).  
New human erg isoforms generated by alternative splicing are transcriptional activators. *Oncogene* **8**, 1865-1873.
- Ernst P., Kahm, K. and Smale, S.T. (1993).  
Both Lyf-1 and Ets protein interact with a critical promoter element in the murine terminal transferase gene. *Mol. Cell. Biol* **13**, 2982-2992.
- Evans, M.E. and Hollenberg, S.M. (1988).  
Zinc fingers : gilt by association. *Cell* **52**, 1-3.
- Fisher R.J., Mavrothalassitis G., Kondoh A. and Papas, T.S. (1991).  
High-affinity DNA-protein interactions of the cellular ETS1 protein : the determination of the ETS binding motif. *Oncogene* **6**, 2249-2254.
- Fisher R.J., Koizumi S., Kondoh A., Mariano J.M., Mavrothalassitis G., Bhat N.K. and Papas, T.S. (1992).  
Human ETS1 oncoprotein : purification, isoforms, -SH modification and DNA sequence-specific binding. *J. Biol. Chem* **267**, 17957-17965.
- Fitch, W.M. and Margoliash, E. (1967).  
Construction of phylogenetic trees. *Science* **155**, 279-284.
- Fleishman L.F., Pilaro A.M., Murakami K., Kondoh A., Fisher R.J. and Papas, T.S. (1993).  
*c-ets-1* protein is hyperphosphorylated during mitosis. *Oncogene* **8**, 771-780.
- Frampton, J., Leutz, A., Gibson, T. and Graf, T. (1989).  
DNA-binding domain ancestry. *Nature* **342**, 134.
- Frykberg, L., Metz, T., Brady, G., Introna, M., Beug, H., Vennstrom, B. and Graf, T. (1988).  
A point mutation in the DNA-binding domain of the *v-myb* oncogene of E26 virus confers temperature sensitivity for transformation of myelomonocytic cells. *Onc. Res.* **3**, 313-322.

- Fujiwara S., Fisher, R.J., Seth, A., Bhat, N.K., Showalter, S.D., Zweig, M. and Papas, T.S. (1988).  
Characterization and localization of the products of the human homologs of the *v-ets* oncogene. *Oncogene* **2**, 99-103.
- Gégonne, A., Leprince, D., Duterque-Coquillaud, M., Vandebunder, B., Flourens, A., Ghysdael, J., Debuire, B. and Stéhelin, D. (1987a).  
Multiple domains for the chicken cellular sequences homologous to the *v-ets* oncogene of the E26 retrovirus. *Mol. Cell. Biol.* **7**, 806-812.
- Gégonne, A., Leprince, D., Pognonec, P., Dernis, D., Faes, M.B., Stéhelin, D. and Ghysdael, J. (1987b).  
The 5' extremity of the *v-ets* oncogene of avian leukemia virus E26 encodes amino acid sequences not derived from the major *c-ets*-encoded cellular proteins. *Virology* **156**, 177-180.
- Gégonne, A., Punhyammalee, B., Rabault, B., Bosselut, R., Seneca, S., Crabeel, M. and Ghysdael, J. (1992).  
Analysis of the DNA binding and transcriptional activation properties of the Ets1 oncoprotein. *New Biol.* **4**, 512-519.
- Gégonne, A., Bosselut, R., Bailly, R.A and Ghysdael, J. (1993).  
Synergistic activation of the HTLV1 LTR Ets-responsive region by transcription factors Ets1 and Sp1. *EMBO J.* **12**, 1169-1178.
- Ghysdael, J., Gégonne, A., Pognonec, P., Dernis, D., Leprince, D. and Stéhelin, D. (1986a).  
Identification and preferential expression in thymic and bursal lymphocytes of a *c-ets*-oncogene-encoded Mr54,000 cytoplasmic protein. *Proc. Natl. Acad. Sci. USA* **83**, 1714-1718.
- Ghysdael, J., Gégonne, A., Pognonec, P., Boulukos, K., Leprince, D., Dernis, D., Lagrou C. and Stéhelin, D. (1986b).  
Identification in chicken macrophages of a set of proteins related to, but distinct from, the chicken cellular *c-ets*-encoded protein p54<sup>c-ets</sup>. *EMBO J.* **5**, 2251-2256.
- Gille H., Sharrocks, A.D. and Shaw, P. (1992).  
Phosphorylation of p62<sup>TCF</sup> by MAP kinase stimulates ternary complex formation at *c-fos* promoter. *Nature* **358**, 414-417.
- Gitlin, S.D., Bosselut, R., Gégonne, A., Ghysdael, J. and Brady, J.N. (1991).  
Sequence-specific interaction of the Ets1 protein with the long terminal repeat of the human T-lymphotropic virus type I. *J. Virol.* **65**, 5513-5523.
- Golay, J., Introna, M. and Graf, T. (1988).  
A single point mutation in the *v-ets* oncogene affects both erythroid and myelomonocytic cell differentiation. *Cell* **55**, 1147-1158.

- Graf, T., Oker-Blom, N., Todorov, T.G. and Beug, H. (1979).  
Transforming capacities and defectiveness of avian leukemia virus OK10 and E26. *Virology* **99**, 431-436.
- Graf, T., McNagny, K., Brady, G. and Frampton, J. (1992).  
Chicken "erythroid" cells transformed by the gag-myb-ets-encoding E26 leukemia virus are multipotent. *Cell* **70**, 201-213.
- Grant, P.A., Arulampalam, V., Ahrlund-Richter L. and Pettersson, S. (1992).  
Identification of Ets-like lymphoid specific element within the immunoglobulin heavy chain 3' enhancer. *Nucleic Acids Res.* **20**, 4401-4408.
- Gunther, C.V., Nye, J.A., Bryner, R.S. and Graves, B.J. (1990).  
Sequence-specific DNA binding of the proto-oncoprotein ets-1 defines a transcriptional activator sequence within the long terminal repeat of the Moloney murine sarcoma virus. *Genes Dev.* **4**, 667-679.
- Gutman, A. and Wasylyk, B. (1990).  
The collagenase gene promoter contains a TPA and oncogene-responsive unit encompassing the PEA3 and AP-1 binding sites. *EMBO J.* **9**, 2241-2246.
- Hagemeier, C., Bannister, A.J., Cook A. and Kouzarides, T. (1993).  
The activation domain of transcription factor PU.1 binds the retinoblastoma (RB) protein and the transcription factor TFIID *in vitro*: RB shows sequence similarity to TFIID and TFIIB. *Proc. Natl. Acad. Sci. USA* **90**, 1580-1584.
- Hagman, J. and Grosschedl, R. (1992).  
An inhibitory carboxyl-terminal domain in Ets-1 and Ets-2 mediates differential binding of ETS family factors to promoter sequences of the *mb-1* genes. *Proc. Natl. Acad. Sci. USA* **89**, 8889-8893.
- Hanahan, D. (1983).  
Studies on transformation of *Escherichia coli* with plasmids. *J. Mol. Biol.* **166**, 557.
- Hermans, M.H.A., Hartsuiker, H. and Opstelten, D. (1989).  
An *in situ* study of B-lymphopoiesis in rat bone marrow: topographical arrangement of nucleotidyl transferase-positive cells and pre-B cells. *J. Immunol.* **142**, 67-73.
- Higashino, F., Yoshida, K., Fujinaga, Y., Kamio K. and Fujinaga, K. (1993).  
Isolation of a cDNA encoding the adenovirus E1A enhancer binding protein: a new human member of the *ets* oncogene family. *Nucleic Acids Res.* **21**, 547-553.

- Hill, C.S., Marais, R., John, S., Wynne, J., Dalton, S. and Treisman, R. (1993). Functional analysis of a growth factor-responsive transcription factor complex. *Cell* **73**, 395-406.
- Hipskind, R.A., Rao, V.N., Mueller, C.G.F., Reddy, E.S.P. and Nordheim, A. (1991). Ets-related protein *elk-1* is homologous to the *c-fos* regulatory factor p62TCF. *Nature* **354**, 531-534.
- Ho, I.C., Bhat, N.K., Gottschalk, L.R., Lindsten, T., Thompson, C.B., Papas, T.S. and Leiden, J.M. (1990). Sequence-specific binding of human *ets-1* to the T cell receptor  $\alpha$  gene enhancer. *Science* **250**, 814-818.
- Introna, M., Golay, L., Frampton, J., Nakano, T., Ness, S.A. and Graf, T. (1990). Mutations in *v-myb* alter the differentiation of myelomonocytic cells transformed by the oncogene. *Cell* **63**, 1287-1297.
- Ivanov, X., Mladenov, Z., Nedyalkov, S. and Todorov, T.G. (1962). Experimental investigations into avian leukosis I. Transmission experiments of certain diseases of the avian leukosis complex, found in Bulgaria. *Bull. Inst. Pathol. Comp. Animaux Domestiques* **9**, 5-36.
- Janknecht, R. and Nordheim, A. (1992). Elk-1 protein domains required for direct and SRF-assisted DNA-binding. *Nucleic Acids. Res.* **20**, 3317-3324.
- Janknecht, R., Ernst, W.H., Pingoud, V. and Nordheim, A. (1993). Activation of ternary complex factor Elk-1 by MAP kinases. *EMBO J.* **12**, 5097-5104.
- Jorcyk, C.L., Watson, D.K., Mavrothalassitis, G.J. and Papas, T.S. (1991). The human ETS1 gene : genomic structure, promoter characterization and alternative splicing. *Oncogene* **6**, 523-532.
- Jurdic, P., Benchaibi M., Grandillon, O. and Samarut, J. (1987). Transforming and mitogenic effects of avian leukemia virus E26 on chicken hematopoietic cells and fibroblasts, respectively, correlate with level of expression of the provirus. *J. Virol.* **61**, 3058-3065.
- Kabat, D. (1990). Molecular biology of Friend viral erythroleukemia. *Curr. Top. Microbiol. Immunol.* **148**, 1-42.

- Karim, F.D., Urness, L.D., Thummel, C.S., Klemsz, M.J., McKercher, S.R., Celada, A., Van Beveren, C., Maki, R.A., Gunther, C.V., Nye, J.A. and Graves, B.J. (1990).  
The ETS-domain : a new DNA-binding motif that recognizes a purine-rich core DNA sequence. *Genes and Dev.* **4**, 1451-1453
- Karim, F.D. and Thummel, C.S. (1991).  
Ecdysone coordinates the timing and amounts of E74A and E74B transcription in *Drosophila*. *Genes Dev.* **5**, 1067-1079.
- Klämbt C. (1993).  
The *Drosophila* gene *pointed* encodes two ETS-like proteins which are involved in the development of the midline glial cells. *Development* **117**, 163-176.
- Klempnauer, K.-H. and Sippel, A.E. (1987).  
The highly conserved amino-terminal region of the protein encoded by the *v-myb* oncogene functions as a DNA-binding domain. *EMBO J.* **6**, 2719-2725.
- Klempnauer, K.-H., Symonds, G., Evan, G.I. and Bishop, J.M. (1984).  
Subcellular localization of proteins encoded by oncogenes of avian myeloblastosis virus and avian leukemia virus E26 and by the chicken *c-myb* gene. *Cell* **37**, 537-547.
- Klemsz, M.J., Mc Kercher, S.R., Celada, A., Van Beveren, C. and Maki, R.A. (1990).  
The macrophage and B cell-specific transcription factor PU.1 is related to the *ets* oncogene. *Cell* **61**, 113-124.
- Koelle, M.R., Talbot, W.S., Segraves, W.A., Bender, M.T., Cherbas, P. and Hogness, D.S. (1991).  
The *Drosophila* EcR gene encodes an ecdysone receptor, a new member of the steroid receptor superfamily. *Cell* **67**, 59-77.
- Koizumi, S., Fisher, R.J., Fujiwara, S., Jorcyk, C., Bhat, N.K., Seth, A. and Papas, T.S. (1990).  
Isoforms of the human *ets-1* protein : generation by alternative splicing and differential phosphorylation. *Oncogene* **5**, 675-681.
- Kola, I., Brookes, D., Green, A.R., Garber, R., Tymms, M., Papas, T.S. and Seth, A. (1993).  
The Ets 1 transcription factor is widely expressed during murine embryo development and is associated with mesodermal cells involved in morphogenetic processes such as organ formation. *Proc. Natl. Acad. Sci. USA* **90**, 7588-7592.

- Kuppuswamy, M., Subramanian, T., Srinivasan, A. and Chinnadural, G. (1989). Multiple functional domains of Tat, the trans-activator of HIV-1, defined by mutational analysis. *Nucleic Acids Res.* **17**, 3551-3561.
- Laget, M.-P., Callebaut, I., de Launoit, Y., Stéhelin D. and Mornon, J.-P. (1993). Predicted common structural features of DNA-binding domains from Ets, Myb and HMG transcription factors. *Nucleic Acids Res.* **21**, 5987-5996.
- Lai, Z.C. and Rubin, G.M. (1992). Negative control of photoreceptor development in *Drosophila* by the product of the *yan* gene, an ETS domain protein. *Cell* **70**, 609-620.
- La Marco, K., Thompson, C.C., Byers, B.P., Walton, E.M. and McKnight, S.L. (1991). Identification of Ets- and Notch-related subunits in GA binding protein. *Science* **253**, 789-792.
- Lameere, A. (1929-1942). Précis de Zoologie. Desoer Ed., Liège (7 vol).
- Landschultz, W., Johnson, P.F. and McKnight, S.L. (1988). The leucine zipper : a hypothetical structure common to a new class of DNA binding proteins. *Science* **240**, 1759-1764.
- Laudet, V., Hänni, C., Coll, J., Catzeflis, F. and Stéhelin, D. (1992). Evolution of the nuclear receptor gene superfamily. *EMBO J.* **11**, 1003-1013.
- Laudet, V., Niel, C., Duterque-Coquillaud, M., Leprince, D. and Stéhelin, D. (1993). Evolution of *ets* gene family. *Biochem. Biophys. Res. Com.* **190**, 8-14.
- Lautenberger, J.A., Burdett, L.A., Gunnell, M.A., Qi, S., Watson, D.K., O'Brien, S.J. and Papas, T.S. (1992). Genomic dispersal of the *ets* gene family during metazoan evolution. *Oncogene* **7**, 1713-1719.
- Lautenberger, J.A. and Papas, T.S. (1993). Inversion of a chicken *ets-1* proto-oncogene segment in avian leukemia virus E26. *J. Virol.* **67**, 610-612.
- Lemarchandel, V., Ghysdael, J., Mignotte, V., Rahuel, C. and Romeo, P.H. (1993). GATA and Ets cis-acting sequences mediate megakaryocyte-specific expression. *Mol. Cell. Biol.* **13**, 668-676.

- Leprince, D., Gégonne A., Coll, J., de Taisne, C., Schneeberger, A., Lagrou, C. and Stéhelin, D. (1983).  
A putative second cell-derived oncogene of the avian leukemia virus E26. *Nature* **306**, 395-397.
- Leprince D., Duterque-Coquillaud, M., Li, R.P., Henry, C., Flourens A., Debuire, B. and Stéhelin, D. (1988).  
Alternative splicing within the chicken *c-ets-1* locus : implications for transduction within the E26 retrovirus of the *c-ets* proto-oncogene. *J. Virol.* **62**, 3233-3241.
- Leprince D., Gesquiere, J.C. and Stéhelin, D. (1990).  
The chicken cellular progenitor of the *v-ets* oncogene, *p68<sup>c-ets-1</sup>*, is a nuclear DNA-binding protein not expressed in lymphoid cells of the spleen. *Oncogene Res.* **5**, 255-265.
- Leprince, D., Crépieux, P. and Stéhelin, D. (1992).  
*c-ets-1* DNA binding to the PEA3 motif is differentially inhibited by all the mutations found in *v-ets*. *Oncogene* **7**, 9-17.
- Leprince D., Crépieux, P., Laudet, V., Flourens, A. and Stéhelin, D. (1993).  
A new mechanism of oncogenic activation : E26 retroviral *v-ets* oncogene has inverted the C-terminal end of the transcription factor *c-ets-1*. *J. Virol.* **194**, 855-857.
- Leung, S., McCracken, S., Ghysdael, J. and Miyamoto, N.G. (1993).  
Requirement of an ETS-binding element for transcription of the human *lck* type I promoter. *Oncogene* **8**, 989-997.
- Levine, M. and Hoey, T. (1988).  
Homeobox proteins as sequence-specific transcription factors. *Cell* **55**, 537-540.
- Lewin, B. (1991).  
Oncogenic conversion by regulatory changes in transcription factors. *Cell* **64**, 303-312.
- Lim, F., Kraut, N., Frampton, J. and Graf, T. (1992).  
DNA binding by *c-ets-1*, but not *v-ets*, is repressed by an intramolecular mechanism. *The EMBO J.* **11**, 643-652.
- Mac Leod, K., Leprince, D. and Stéhelin, D. (1992).  
The *ets* gene family. *Trends Biol. Sci.* **17**, 251-256.
- Majérus M.-A., Bibollet-Ruche, F., Telliez, J.-B., Wasyluk, B. and Bailleul, B. (1992).  
Serum, AP-1 and Ets-1 stimulate the human *ets* promoter. *Nucleic Acids Res* **20**, 2699-2703.

- Marais, R., Wynne, J. and Treisman, R. (1993).  
The SRF accessory protein elk-1 contains a growth factor-regulated transcriptional activation domain. *Cell* **73**, 381-393.
- Martinez, E., Givel, F. and Wahli, W. (1991).  
A common ancestor DNA motif for invertebrate and vertebrate hormone response elements. *EMBO J.* **10**, 263-268.
- Mavrothalassitis, G.J., Watson, D.K. and Papas, T.S. (1990).  
Molecular and fonctionnal characterisation of the promoter of *ETS-2*, the human *c-ets-2* gene. *Proc. Natl. Acad. Sci. USA* **87**, 1047-1051.
- Mavrothalassitis, G.J. and Papas, T.S. (1991).  
Positive and negative factors regulate the transcription of the *ets-2* gene via an oncogene-responsive-like unit within the *ets-2* promoter region. *Cell Growth and Diff.* **2**, 215-224.
- May, W.A., Gishizky, M.L., Lessnick, S.L., Lunsford, L.B., Lewis, B.C., Delattre, O., Zucman, J., Thomas, G. and Denny, C.T. (1993a).  
Ewing sarcoma 11 ; 22 translocation produces a chimeric transcription factor that requires the DNA-binding domain encoded by *FLI-1* for transformation. *Proc. Natl. Acad. Sci. USA* **90**, 5752-5756.
- May, W.A., Lessnick, S.L., Braun B.S., Klemsz M., Lewis, B.C., Lunsford L.B., Hromas R. and Denny C.T. (1993b).  
The Ewing's sarcoma *EWS/FLI-1* fusion gene encodes a more potent transcriptional activator and is a more powerful transforming gene than *FLI-1*. *Molec. Cell. Biol.* **13**, 7393-7398.
- Metz, T. and Graf, T. (1991a).  
*v-myb* and *c-ets* transform chicken erythroid cells and cooperate both in *trans* and in *cis* to induce distinct differentiation phenotypes. *Genes Dev.* **5**, 369-380.
- Metz, T. and Graf, T. (1991b).  
Fusion of the nuclear oncoproteins *v-Myb* and *v-Ets* is required for the leukemogenicity of E26 virus. *Cell* **66**, 95-105.
- Moreau-Gachelin, F., Tavitian, A. and Tambourin, P. (1988).  
*Spi-1* is a putative oncogene in virally induced murine erythroleukemia. *Nature* **331**, 277-280.
- Moreau-Gachelin, F., Ray, D., Mattei, M.G., Tambourin, P. and Tavitian, A. (1989).  
The putative oncogene *Spi-1* : murine chromosomal localization and transcriptional activation in murine acute erythroleukemia. *Oncogene* **4**, 1449-1456.

- Moscovici C. (1975).  
Leukemic transformation with avian myeloblastosis virus ; present status.  
*Curr. Top. Microbiol. Immunol* **71**, 79-101.
- Moscovici C., Samarut, J., Gazzolo, L. and Moscovici, M.G. (1981).  
Myeloid and erythroid neoplastic responses to avian defective leukemia viruses in chicken and in quail. *Virology* **113**, 765-768.
- Moscovici M.G., Jurdic, P., Samarut, J., Gazzolo, L., Mura, C.V. and Moscovici, C. (1983).  
Characteristics of the hemopoietic target cells for the avian leukemia virus E26. *Virology* **129**, 65-78.
- Mount, S.M. (1982).  
A catalogue of splice junction sequences. *Nucleic Acids Res.* **10**, 459-472.
- Murakami, K., Mavrothalassitis, G., Bhat, N.K., Fisher, R.J. and Papas, T.S. (1993).  
Human ERG-2 is a phosphorylated DNA-binding protein, a distinct member of the *ets* family. *Oncogene* **8**, 1559-1566.
- Murre, C., Schonleber McCaw, P. and Baltimore, D. (1989).  
A new DNA binding and dimerization motif in immunoglobulin enhancer binding, *daughterless*, myoD and myc proteins. *Cell* **56**, 777-783.
- Murtha, M.T., Leckman, J.F. and Ruddle, F.H. (1991).  
Detection of homeobox genes in development and evolution. *Proc. Natl. Acad. Sci. USA* **88**, 10711-10715.
- Nelsen, B., Tian, G., Erman, B., Gregoire, J., Maki, R., Graves, B. and Sen, R. (1993).  
Regulation of lymphoid-specific immunoglobulin  $\mu$  heavy chain gene enhancer by ETS-domain proteins. *Science* **261**, 82-86.
- Nerlov, C., De Cesare, D., Pergola, F., Caracciolo, A., Blasi, F., Johnsen, M. and Verde, P. (1992).  
A regulatory element that mediates cooperation between a PEA3-AP-1 element and an AP-1 site is required for phorbol ester induction of urokinase enhancer activity in HepG2 hepatoma cells. *EMBO J.* **11**, 4573-4582.
- Ness, S.A., Marknell, A. and Graf, T. (1989).  
The v-myb oncogene product binds to and activates the promyelocyte-specific mim-1 gene. *Cell* **59**, 1115-1125.
- Nunn, M.F., Seeburg, P.H., Moscovici C. and Duesberg, P.H. (1983).  
Tripartite structure of the avian erythroblastosis virus E26 transforming gene. *Nature* **306**, 391-395.

- Nunn, M.F. and Hunter, T. (1989).  
The *ets* sequence is required for induction of erythroblastosis in chicken by avian retrovirus E26. *J. Virol.* **63**, 398-402.
- Nye, J.A., Petersen, J.M., Gunther, C.V., Jonsen, M.D. and Graves, B.J. (1992).  
Interaction of murine Ets-1 with GGA-binding sites establishes the ETS domain as a new DNA-binding motif. *Genes and Dev* **6**, 975-990.
- Oka, T., Rairkar, A. and Chen, J.H. (1991).  
Structural and functional analysis of the regulatory sequences of the *ets-1* gene. *Oncogene* **6**, 2077-2093.
- Ottillie, S., Raulf, F., Barnekow, A., Hannig, G. and Scharl, M. (1992).  
Multiple *src*-related kinase gene, *srk-1-4*, in the fresh water sponge *Spongilla lacustris*. *Oncogene* **7**, 1625-1630.
- Oro, A.E., Mc Keown, M. and Evans, E.M. (1990).  
Relationship between the *Drosophila ultraspiracle* locus and the vertebrate retinoid X receptor. *Nature* **347**, 298-301.
- Paz-Ares, J., Ghosal, D., Wienand, U., Peterson, P.A. and Saedler, H. (1987).  
The regulatory *cl* locus of *Zea mays* encodes a protein with homology to *myb* proto-oncogene products and with structural similarities to transcriptional activators. *EMBO J.* **6**, 3553-3558.
- Pendleton, J.W., Nagai, B.K., Murtha, M.T. and Ruddle, F.H. (1993).  
Expansion of the *Hox* gene family and the evolution of chordates. *Proc. Natl. Acad. Sci. USA* **90**, 6300-6304.
- Pognonec P., Boulukos K.E., Gesquière, J.C., Stéhelin, D. and Ghysdael, J. (1988).  
Mitogenic stimulation of thymocytes results in the calcium-dependent phosphorylation of *c-ets-1* proteins. *EMBO J.* **7**, 977-983
- Pognonec P., Boulukos K.E. and Ghysdael, J. (1989).  
The *c-ets-1* protein is chromatin associated and binds to DNA *in vitro*. *Oncogene* **4**, 691-697.
- Pognonec P., Boulukos K.E., Bosselut, R., Boyer, C., Schmitt-Verhulst, A.M. and Ghysdael, J. (1990).  
Identification of a Ets1 variant protein unaffected in its chromatin and *in vitro* DNA binding capacities by T cell antigen receptor triggering and intracellular calcium rises. *Oncogene* **5**, 603-610.
- Pongubala, J.M.R., Nagulaplali, S., Klemsz, M.J., Mc Kercher, S.R., Maki, R.A. and Atchison, M.L. (1992).  
PU-1 recruits a second nuclear factor to a site important for immunoglobulin  $\kappa$  3' enhancer activity. *Mol. Cell. Biol.* **12**, 368-378.

- Pongubala, J.M.R., Van Beveren, C., Nagulapalli, S., Klemsz, M.J., McKercher, S.R., Maki, R.A. and Atchison, M.K. (1993).  
Effect of PU-1 phosphorylation on interaction with NF-EM5 and transcriptional activation. *Science* **259**, 1622-1625.
- Porchet-Henneré, E., M'Béri, M., Dhainaut, A. and Porchet, M. (1987).  
Ultrastructural study of the encapsulation-response of the Polychaete Annelid *Nereis diversicolor*. *Cell Tissue Res.* **248**, 463-471.
- Prasad, D.D.K., Rao, V.N. and Reddy, E.S.P. (1992).  
Structure and expression of human *Fli-1* gene. *Cancer Res.* **52**, 5833-5837.
- Pribyl, L.J., Watson, D.K., McWilliams, M.J., Ascione, R. and Papas, T.S. (1987).  
Characterization of the *ets-2* gene in *Drosophila melanogaster*. In *Advances in Gene Technology : The Molecular Biology of Development* (Voellmy, R.W., Ahmad, F., Black, S., Burgess, D.R., Rotundo, R., Scott, W.A. and Whelan, W.J. Eds). Vol. 7, p.93. Cambridge Univ. Press, Cambridge.
- Pribyl, L.J., Watson, D.K., McWilliams, M.J., Ascione, R. and Papas, T.S. (1988).  
The *Drosophila ets-2* gene : molecular structure, chromosomal localization and developmental expression. *Dev. Biol.* **127**, 45-53.
- Pribyl, L.J., Watson, D.K., Schulz, R.A. and Papas, T.S. (1991).  
*D-eltg*, a member of the *Drosophila ets* gene family : sequence, expression and evolutionary comparison. *Oncogene* **6**, 1175-1183.
- Prosser, H.M., Wotton, D., Gégonne, A., Ghysdael, J., Wang, S., Speck, N.A. and Owen, M.J. (1992).  
A phorbol ester response element within the human T-cell receptor  $\beta$ -chain enhancer. *Proc. Natl. Acad. Sci. USA* **89**, 9934-9938.
- Quéva, C., Leprince, D. and Stéhelin, D. (1993).  
*p54<sup>c-ets-1</sup>* and *p68<sup>c-ets-2</sup>*, the two transcription factors encoded by the *c-ets-1* locus, are differentially expressed during the development of the chick embryo. *Oncogene* **8**, 2511-2520.
- Radke, K., Beug, H., Kornfeld, S. and Graf, T. (1982).  
Transformation of both erythroid and myeloid cells by E26, an avian leukemia virus that contains the *myb* gene. *Cell* **31**, 643-653.
- Rao, V.N. and Reddy, E.S.P. (1992a).  
A divergent *ets*-related protein, *elk-1*, recognizes similar *c-ets-1* proto-oncogene target sequences and acts as a transcriptional activator. *Oncogene* **7**, 65-70.

- Rao, V.N. and Reddy, E.S.P. (1992b).  
Elk-1 domains responsible for autonomous DNA binding, SRE : SRF interaction and negative regulation of DNA binding. *Oncogene* **7**, 2335-2340.
- Rao, V.N. and Reddy, E.S.P. (1993).  
 $\Delta$ elk-1, a variant of elk-1, fails to interact with the serum response factor and binds to DNA with modulated specificity. *Canc. Res.* **53**, 215-220.
- Rao, V.N., Papas, T.S. and Reddy, E.S.P. (1987).  
erg, a human ets-related gene on chromosome 21 : alternative splicing, polyadenylation, and translation. *Science* **237**, 635-639.
- Rao, V.N., Huebner, K., Isobe, M., Ar-Rushdi A., Croce, C.M. and Reddy, E.S.P. (1989).  
elk, tissue-specific ets-related genes on chromosomes X and 14 near translocation breakpoints. *Science* **244**, 55-70.
- Rao, V.N., Ohno, T., Prasad, D.D.K., Bhattacharya, G. and Reddy, E.S.P. (1993).  
Analysis of the DNA-binding and transcriptional activation functions of human *Fli-1* protein. *Oncogene* **8**, 2167-2173.
- Ravel-Chapuis, P., Leprince, D., Pain, B., Li, R., Domenget, C., Stéhelin, D., Samarut, J. and Jurdic, P. (1991).  
*v-myb* and *v-ets* cooperate for the mitogenic stimulation of primary fibroblasts by avian E26 retrovirus. *J. Virol.* **65**, 3928-3931.
- Reddy, E.S.P. and Rao, V.N. (1988).  
Structure, expression and alternative splicing of the human *c-ets-1* proto-oncogene. *Onc. Res.* **3**, 239-246.
- Reddy, E.S.P. and Rao, V.N. (1990).  
Localization and modulation of the DNA-binding activity of the human *c-ets-1* proto-oncogene. *Canc. Res.* **50**, 5013-5016.
- Reddy, E.S.P., Rao, V.N. and Papas, T.S. (1987).  
The *erg* gene : A human gene related to the *ets* oncogene. *Proc. Natl. Acad. Sci. USA* **84**, 6131-6135.
- Richardson, W.D., Roberts, B.L. and Smith, A.E. (1986).  
Nuclear localization signals in polyoma virus large-T. *Cell* **44**, 77-85.
- Riddihough, G. and Pelham, H.R.B. (1987).  
An ecdysone response element in the *Drosophila hsp27* promoter. *EMBO J.* **6**, 3729-3734.

- Rogers, S., Wells, R. and Rechsteine, M. (1986).  
Amino acids sequences common to rapidly degraded proteins : the PEST hypothesis. *Science* **234**, 364-368.
- Rorth, P., Nerlov, C., Blasi, F. and Johnson, M. (1990).  
Transcription factor PEA3 participates in the induction of urokinase plasminogen activator transcription in murine keratinocytes stimulated with epidermal growth factor or phorbol-ester. *Nucleic Acids Res.* **18**, 5009-5017.
- Roussel, M., Saule, S., Lagrou, C., Rommens, C., Beug, H., Graf, T. and Stéhelin, D. (1979).  
Three new types of viral oncogenes of cellular origin specific for hematopoietic cell transformation. *Nature* **281**, 452-455.
- Rovigatti, U., Watson, D.K. and Yunis, J.J. (1986).  
Amplification and rearrangement of Hu-*ets-1* in leukemia and lymphoma with involvement of 11q23. *Science* **232** 398-400.
- Rowe, A. and Propst, F. (1992).  
*Ets-1* and *Ets-2* protooncogene expression in theca cells of the adult mouse ovary. *Exptl Cell Res.* **202**, 199-202.
- Ruscetti, S., Aurigemma, R.E., Yuan, C.C., Sawyer S. and Blair, D.G. (1992).  
Induction of erythropoietin responsiveness in murine hematopoietic cells by the gag-myb-ets-containing ME26 virus. *J. Virol.* **66**, 20-26.
- Sacchi, N., Watson, D.K., Guertis Van Kessel, Ad H.M., Hagemeyer A., Kersey, J., Drabkin, H.D., Patterson, D. and Papas, T.S. (1986).  
Hu-*ets-1* and Hu-*ets-2* genes are transposed in acute leukemias with (4 ; 11) and (8 ; 21) translocations. *Science* **231**, 379-382.
- Sacchi, N., Wendtner, C.M. and Thiele, C.J. (1991).  
Single-cell detection of *ets-1* transcripts in human neuroectodermal cells. *Oncogene*, **6**, 2149-2154.
- Saitou, N. and Nei, M. (1987).  
The neighbor-joining method : a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4**, 406-425.
- Salmon, P., Giovane, A., Wasyluk, B. and Klatzmann, D. (1993).  
Characterization of the human CD4 gene promoter : transcription from the CD4 gene core promoter is tissue-specific and is activated by Ets proteins. *Proc. Natl. Acad. Sci. USA* **90**, 7739-7743.
- Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989).  
Molecular cloning, A laboratory manual. Cold Spring Harbor Laboratory Press.

- Sanger, F., Nicklen, S. and Coulson, A.R. (1977).  
DNA sequencing with chain termination inhibitors. *Proc. Natl. Acad. Sci. USA* **74**, 5463-5464.
- Schneikert, J., Lutz, Y. and Wasyluk, B. (1992).  
Two independent activation domains in *c-ets-1* and *c-ets-2* located in non-conserved sequences of the *ets* gene family. *Oncogene* **7**, 249-256.
- Schulz, R.A., The, S.M., Hogue, D.A., Galewsky S. and Guo, Q. (1993a).  
*Ets* oncogene-related gene *Elg* functions in *Drosophila* oogenesis. *Proc. Natl. Acad. Sci. USA*, **90**, 10076-10080.
- Schulz, R.A., Hogue, D.A. and The, S.M. (1993b).  
Characterization of lethal alleles of D-elg, an *ets* proto-oncogene related gene with multiple functions in *Drosophila* development. *Oncogene* **8**, 3369-3374.
- Seneca, S., Punyammalee, B., Bailly, M., Ghysdael, J. and Crabeel, M. (1991).  
Ets1, when fused to the GAL4 DNA binding domain, efficiently enhances galactose promoter dependent gene expression in yeast. *Oncogene* **6**, 357-360.
- Seth, A. and Papas, T.S. (1990).  
The *c-ets-1* proto-oncogene has oncogenic activity and is positively autoregulated. *Oncogene* **5**, 1761-1767.
- Seth, A., Watson, D.K., Blair, D.G. and Papas, T.S. (1989).  
*c-ets-2* protooncogene has mitogenic and oncogenic activity. *Proc. Natl. Acad. Sci. USA* **86**, 7833-7837.
- Seth, A., Ascione, R., Fisher, R.J., Mavrothalassitis, G.J, Bhat, N.K. and Papas, T.S. (1992).  
The *ets* Gene Family. *Cell Growth Diff.* **3**, 327-334.
- Seth, A., Robinson, L., Thompson, D.M., Watson, D.K. and Papas, T.S. (1993).  
Transactivation of GATA-1 promoter with ETS1, ETS2 and ERGB/Hu-FLI-1 proteins : stabilization of the ETS1 protein binding on GATA-1 promoter sequences by monoclonal antibody. *Oncogene* **8**, 1783-1790.
- Shenk, M.A. and Steele, R. (1993).  
A molecular snapshot of the metazoan "Eve". *TIBS* , 459-463.
- Siddique, H.R., Rao, V.N., Lee, L. and Reddy, E.S.P. (1993).  
Characterization of the DNA binding and transcriptional activation domains of the *erg* protein. *Oncogene* **8**, 1751-1755.

- Southern, E.M. (1975).  
Detection of specific sequence among DNA fragment separated by gel electrophoresis. *J. Mol. Biol.* **98**, 503-517.
- Spector, D., Purves, F. and Roizman, B. (1990).  
Mutational analysis of the promoter region of the  $\alpha 27$  gene of herpes simplex virus 1 within the context of the viral genome. *Proc. Natl. Acad. Sci. USA* **87**, 5268-5272.
- Stiegler, P., Wolff, C.M., Baltzinger, M., Hirtzlin, J., Senan, F., Meyer, D., Ghysdael, J., Stéhelin, D., Befort, N. and Remy, P. (1990).  
Characterization of *Xenopus laevis* cDNA clones of the *c-ets-1* proto-oncogene. *Nucleic Acids Res.* **18**, 5298.
- Stiegler, P., Wolff, C.M., Meyer, D., Senan, F., Durliat, M., Hourdry, J., Befort, N. and Remy, P. (1993).  
The *c-ets-1* proto-oncogenes in *Xenopus laevis* : expression during oogenesis and embryogenesis. *Mechanisms Develop.* **41**, 163-174.
- Swain, A. and Coffin, J.M. (1992).  
Mechanism of transduction by retroviruses. *Science* **255**, 841-844.
- Swofford, D.L. (1984).  
Phylogenetic Analysis Using Parsimony (PAUP). Illinois Natural History Survey ; Champaign.
- Tei, H., Nihonmatsu, I., Yokokura, T., Ueda, R., Sano, Y., Okuda, T., Sato, K., Hirata, K., Fujita, S.C. and Yamamoto, D. (1992).  
*Pokkuri*, a *Drosophila* gene encoding an E26-specific (Ets) domain protein, prevents overproduction of the R7 photoreceptor. *Proc. Natl. Acad. Sci. USA* **90**, 6856-6860.
- Thompson, C.C., Brown, T.A. and McKnight, S.L. (1991).  
Convergence of *ets*- and notch-related structural motifs in a heteromeric DNA binding complex. *Science* **253**, 762-768.
- Thompson, C.C., Wang, C.-Y., Ho, I.C., Bohjanen, P.R., Petryniak, B., June, C.H., Miesfeldt, S., Zhang, L., Nabel, G.J., Karpinski, B. and Leiden, J.M. (1992).  
cis-acting sequences required for inducible interleukin-2 enhancer function bind a novel Ets-related protein, Elf-1. *Mol. Cell. Biol.* **12**, 1043-1053.
- Thummel, C.S. (1990).  
Puffs and Gene Regulation - Molecular Insights into the *Drosophila* Ecdysone Regulatory Hierarchy. *BioEssays* **12**, 561-568.

- Thummel, C.S., Burtis, K.C. and Hogness, D.S. (1990).  
Spatial and temporal patterns of E74 transcription during *Drosophila* development. *Cell* **61**, 101-111.
- Triezenberg, S.J., Lamarco, K.L. and McKnight, S.L. (1988).  
Evidence of DNA : protein interaction that mediate HSV-1 immediate early gene expression *Genes Dev.* **3**, 730-742.
- Urness, L.D. and Thummel, C.S. (1990).  
Molecular interactions within the ecdysone regulatory hierarchy : DNA binding properties of the *Drosophila* ecdysone-inducible E74A protein. *Cell* **61**, 47-61.
- Vandenbunder, B., Pardanaud, L., Jaffredo, T., Mirabel, M.A. and Stéhelin, D. (1989).  
Complementary patterns of expression of *c-ets-1*, *c-myb* and *c-myc* in the blood forming system of the chick embryo. *Development* **106**, 265-274.
- Vison, C.R. and Garcia, K.C. (1992).  
Molecular model for DNA recognition by the family of basic-helix-loop-helix-zipper proteins. *New Biologist* **4**, 396-403.
- Wang, C.Y., Petryniak, B., Ho, I.C., Thompson, C.B. and Leiden, J.M. (1992).  
Evolutionary conserved ets family members display distinct DNA binding specificities. *J. Exp. Med.* **175**, 1391-1399.
- Wang, C.Y., Petryniak, B., Thompson, C.B. Kaelin, W.G. and Leiden, J.M. (1993).  
Regulation of the Ets-related transcription factor Elf-1 by binding to the retinoblastoma protein. *Science* **260**, 1330-1335.
- Wasylyk, C., Flores, P., Gutman, A. and Wasylyk B. (1989).  
PEA3 is a nuclear target for transcription activation by non-nuclear oncogenes. *EMBO J.* **8**, 3371-3378.
- Wasylyk B., Wasylyk, C., Flores, P., Begue, A., Leprince, D. and Stéhelin, D. (1990).  
The *c-ets* proto-oncogenes encode transcription factors that cooperate with c-Fos and c-Jun for transcriptional activation. *Nature* **346**, 191-193.
- Wasylyk C., Gutman, A., Nicholson, R. and Wasylyk, B. (1991).  
The c-Ets oncoprotein activates the stromelysin promoter through the same elements as several non-nuclear oncoproteins. *EMBO J.* **10**, 1127-1134.
- Wasylyk, C., Kerckaert, J.P. and Wasylyk, B. (1992).  
A novel modulator domain of Ets transcription factors. *Genes Dev.* **6**, 965-974

- Wasylyk C. and Wasylyk, B. (1992).  
Oncogenic conversion alters the transcriptional properties of *ets*. *Cell Growth and Diff.* **3**, 617-625.
- Wasylyk C. and Wasylyk, B. (1993).  
Oncogenic conversion of *Ets* affects redox regulation *in-vivo* and *in-vitro*. *Nucleic Acids Res.* **21**, 523-529.
- Watson, D.K., Mc Williams, M.J., Nunn, M.F., Duesberg, P., O'Brien, S.J. and Papas, T.S. (1985).  
The *ets* sequence from the transforming gene of avian erythroblastosis virus, E26, has unique domains on human chromosomes 11 and 21 : both loci are transcriptionally active. *Proc. Natl. Acad. Sci. USA* **82**, 7294-7298.
- Watson, D.K., Mc Williams, M.J. and Papas, T.S. (1988a).  
Molecular organization of the chicken *ets* locus. *Virology* **164**, 99-105.
- Watson, D.K., Mc Williams, M.J., Lapis, P., Lautenberger, J.A., Schweinfest, C.W. and Papas, T.S. (1988b).  
Mammalian *ets-1* and *ets-2* genes encode highly conserved proteins. *Proc. Natl. Acad. Sci. USA* **85**, 7862-7866.
- Watson, D.K., Smyth, F.E., Thompson D.M., Cheng, J.Q., Testa, J.R., Papas, T.S. and Seth, A. (1992).  
The *ERGB/Fli-1* Gene : Isolation and Characterization of a New Member of the Family of Human *ETS* Transcription Factors. *Cell Growth Diff.* **3**, 705-713.
- Wattez-Combaz, A. (1985).  
Contrôle par la chaîne nerveuse de l'établissement de la polarité dorso-ventrale du régénérat caudal chez la Néréis (Annélide Polychète). *Thèse Université des Sciences et Techniques de Lille*, N° 654.
- Wernert, N., Raes, M.B., Lasalle, P., Dehouck, M.P., Gosselin, B., Vandebunder, B. and Stéhelin, D. (1992).  
*c-ets-1* proto-oncogene is a transcription factor expressed in endothelial cells during tumor vascularization and other forms of angiogenesis in humans. *Am. J. Pathol.* **140**, 119-127.
- Weston, K. and Bishop, J.M. (1989).  
Transcriptional activation by the *v-myb* oncogene and its cellular progenitor, *c-myb*. *Cell* **58**, 85-93.
- Wolff, C.M., Stiegler, P., Baltzinger, M., Meyer, D., Ghysdael, J. Stéhelin, D., Belfort, N. and Rémy, P. (1990).  
Isolation of two different *c-ets-2* proto-oncogenes in *Xenopus laevis*. *Nucleic Acids Res.* **18**, 4603-4604.

- Wolff, C.M., Stiegler, P., Baltzinger, M., Meyer, D., Ghysdael, J. Stéhelin, D., Belfort, N. and Rémy, P. (1991).  
Cloning, sequencing and expression of two *Xenopus laevis* *c-ets-2* proto-oncogenes. *Cell Growth Diff.* **2**, 447-456.
- Wolfner, M. (1980).  
Ecdysone-responsive genes of the salivary gland of *Drosophila melanogaster*. Ph. D. Thesis. Stanford University. Stanford, California. (Cité in Burtis, K.C. *et al.*, 1990).
- Woods, D.B., Ghysdael, J. and Owen, M.J. (1992).  
Identification of nucleotide preferences in DNA sequences recognised specifically by c-ETS-1 protein. *Nucl. Acids Res.* **20**, 699-704.
- Xin, J.H., Cowie, A., Lachance, P. and Hassell, J.A. (1992).  
Molecular cloning and characterization of PEA3, a new member of the *Ets* oncogene family that is differentially expressed in mouse embryonic cells. *Genes Dev.* **6**, 481-496.
- Yao, T.P., Segraves, W.A., Oro, A.E., Mc Keown, M. and Evans R.M. (1992).  
*Drosophila ultraspiracle* modulates ecdysone receptor function *via* heterodimer formation. *Cell* **71**, 63-72.
- Yuan, C.C., Kan, N., Dunn, K.J., Papas, T.S. and Blair, D.G. (1989).  
Properties of a murine retroviral recombinant of avian acute leukemia virus E26 : a murine fibroblast assay for v-*ets* function. *J. Virol.* **63**, 205-215.
- Zhang, J. and Temin, H. (1993a).  
Rate and mechanism of non homologous recombination during a single cycle of retroviral replication. *Science* **259**, 234-238.
- Zhang, J. and Temin, H. (1993b).  
3' junctions of oncogene-virus sequences and the mechanisms for formation of highly oncogenic retroviruses. *J. Virol.* **67** 1747-1751.
- Zhang, L., Lemarchandel, V., Romeo, P.H., Ben-David, Y., Greer, P. and Bernstein, A. (1993).  
The Fli-1 proto-oncogene, involved in erythroleukemia and Ewing's sarcoma, encodes a transcriptional activator with DNA-binding specificities distinct from other *Ets* family members. *Oncogene* **8**, 1621-1630.
- Zucman, J., Melot, T., Desmaze, C., Ghysdael, J., Plougastel, B., Peter, M., Zucker, J.M., Triche, T.J., Sheer, D., Turc-Carel, C., Ambros, P., Combaret, V., Lenoir, G., Aurias, A., Thomas, G. and Delattre, O. (1993).  
Combinatorial generation of variable fusion proteins in the Ewing family of tumours. *EMBO J.* **12**, 4481-4487.