

50376  
1998  
169

Université des Sciences et Technologies de Lille I



**THESE**  
présentée pour l'obtention du grade de  
Docteur de l'Université en Sciences de la Vie et de la Santé

par

**Séverine CARRERE**

**Spécificité et redondance fonctionnelles  
des facteurs de transcription Erg,  
membres de la famille Ets**

Présentée le 6 juillet 1998 devant la commission d'examen :

**Président : Pr Xavier DESBIENS**

**Rapporteurs : Dr Olivier DELATTRE  
Dr Robert HIPSKIND**

**Examineurs Dr Martine DUTERQUE-COQUILLAUD  
Pr Dominique STEHELIN**

# *Je remercie*

Tout particulièrement, *les membres du jury* qui ont accepté d'évaluer mon travail.

*Dominique Stéhelin* de m'avoir acceptée au sein de son laboratoire. Je n'oublierai pas ses mots de bienvenue par une très belle journée de juillet: "Aujourd'hui, le temps est exceptionnel mais Lille est une ville super."

*Martine Duterque*, "Mégachef", qui a su autant me dévoiler les mystères de la science que celui des saisons dans le Nord? A ce propos, c'est bientôt le 14 juillet!

*Anne Flourens*, notre maman à tous, mais surtout une vraie ch'ti qui sait faire aimer son pays. Merci pour tous tes coups de main et ton soutien.

*Alexis Forger*, le seul homme du groupe c'est-à-dire une lourde responsabilité! Il est quand-même là pour une chansonnette, une discussion philosophique, ou pour me fournir du chocolat en cas de grosse crise de stress... Je ne peux que te dire : "Show Time!"

*Sandrine Phopin*, ma copine des bancs de facs, des sièges de paille et des salles d'écriture. Je crois que j'ai bien résumé la situation...

*Sophana Ung* pour son inlassable goût pour la plaisanterie et pour m'avoir donné une seconde famille loin de chez moi.

*Phantal Mercamer* qui pourrait être ma 3ème maman, sauf quand elle se noie dans la Lesse.

*Agnès Bègue* pour ses précieux conseils et toutes ses petites histoires.

*Delphine Watiez* et son sacré coup de main. Tu mériterais de faire la une!

*Vincent Laudet* et ses remarques toujours à propos.

*Virginie Flaeminck* pas mauvaise pour fournir un certificat de dispense de piscine.

*Patrick Dumont* pour avoir guidé mes premiers pas en animalerie.

*Albin Rourtier* et ses délicieuses petites phrases qui illuminent une journée.

Il a aussi toutes mes copines qui rendent la vie d'un laboratoire beaucoup plus agréable, *Anne-Marie Ayral*, *Isabelle Damour*, *Cateline Delmarre*, *Sophie Deltour*, *Frédérique Dewitte*, *nos secrétaires de choc* et surtout *les dames de la laverie*.

Je n'oublie pas *mes parents* et ma "*grande sœur*" pour leur soutien (financier, si si, et moral, quand-même), et bien sûr *Laurent*.

## Table des matières

Le stage de thèse.....	3
Publications.....	3
Présentation du mémoire.....	4
Abréviations et conventions.....	5
 <b>INTRODUCTION</b>	
<b>1ère partie : La famille des gènes <i>ets</i></b> .....	<b>6</b>
I-L'oncogène viral <i>v-ets</i> .....	6
II-La famille des gènes <i>ets</i> .....	6
1-Phylogénie des membres de la famille <i>ets</i> .....	8
2-Les domaines d'homologie autres que le domaine ETS.....	8
3-Une origine pour la famille <i>ets</i> ?.....	9
III-Le domaine ETS.....	11
1-Les séquences cibles.....	11
2-La structure du domaine ETS.....	14
a-Structure du domaine ETS de la protéine Fli-1 humaine.....	16
b-Structure du domaine ETS de la protéine Ets-1 murine.....	16
c-Structure du domaine ETS de la protéine Ets-1 humaine.....	16
d-Structure du domaine ETS de la protéine Pu-1 humaine.....	18
e-Structure du domaine ETS de la protéine Gabp $\alpha$ murine.....	18
IV-Le domaine PNT.....	20
1-Les tracas de sa dénomination.....	20
2-Les fonctions potentielles du domaine PNT.....	21
<b>2ème partie : Les protéines Fos et Jun, partenaires privilégiés des protéines Ets</b> .....	<b>23</b>
I-II était une fois.....	23
II-Les protéines Fos et Jun courbent l'ADN.....	25
III-Les protéines Fos et Jun dans des complexes protéiques.....	26
IV-Les études de fonction.....	27
1-Les gènes <i>c-fos</i> et <i>fos-B</i> .....	27
2-Le gène <i>c-jun</i> .....	27
<b>3ème partie : La fonction des protéines Ets</b> .....	<b>30</b>
I-Les facteurs de transcription Ets.....	30
1-Les domaines de régulation de la transcription.....	30
a-Les domaines transactivateurs.....	30
b-Les domaines inhibiteurs.....	32
2-Les protéines Ets coopèrent avec divers partenaires.....	34
a-La protéine Ets-1 sur le promoteur du gène <i>GM-CSF</i> .....	34
b-Des protéines Ets interfèrent avec des récepteurs nucléaires.....	35
3-Des interactions protéine-protéine.....	36
a-Les protéines Elk-1 et Sap-1 sur l'élément de réponse au sérum, SRE.....	36
b-Etude de l'"enhancer" minimal du gène <i>TcR<math>\alpha</math></i> .....	38
c-Coopération des familles Ets et AP-1.....	38
d-Les protéines MafB et Ets-1 dans la différenciation érythrocytaire.....	39
e-Ets-1 et USF-1 sur l'"enhancer" de HIV-1 .....	41
II-L'étude des fonctions des gènes <i>ets in vivo</i> .....	41
1-L'hématopoïèse.....	43
a-Le gène <i>ets-1</i> .....	43
b-Le gène <i>fli-1</i> .....	44
c-Le gène <i>pu-1</i> .....	45
d-Le gène <i>spi-B</i> .....	45
2-L'angiogenèse.....	46
3-La formation du squelette.....	46
<b>4ème partie : Le gène <i>erg</i></b> .....	<b>49</b>
I-Sa découverte.....	49
II-Ses propriétés oncogéniques.....	49
1-L'oncogène <i>erg</i> .....	49

2-Des translocations chromosomiques.....	49
a-La translocation t(16;21).....	51
b-La translocation t(21;22).....	51
3-Le syndrome de Down.....	52
III-Son expression.....	54
1-Les études par "Northern-blot".....	54
2-Les études par hybridation <i>in situ</i> .....	54
a-Coexpression du gène <i>erg</i> avec d'autres membres de la famille.....	54
b-Le gène <i>erg</i> impliqué dans la mise en place du cartilage.....	56
3-"Des petites billes".....	58
IV-Les produits du gène <i>erg</i> .....	60
1-Les ADNc issus du gène <i>erg</i> .....	60
2-Les ADNc <i>erg</i> et <i>fli-1</i> .....	61
V-Les propriétés transactivatrices des protéines Erg.....	61
1-Les protéines Erg se lient à l'ADN.....	61
2-Les protéines Erg transactivent.....	62
3-Les protéines Erg et leurs cofacteurs.....	63

## RESULTATS

I-Publication 1.....	65
1-Les protéines Ets-2 et Ergp55 diffèrent.....	66
2-Les protéines Ets et Fos/Jun coopèrent.....	67
3-Les MMP sont des gènes cibles potentiels des protéines Ets.....	68
II-Publication 2.....	70
1-Les protéines Erg se fixent en monomères.....	71
a-Le promoteur de la stromélysine.....	72
b-Un site, une protéine, deux sites, une protéine.....	72
2-Les domaines impliqués dans la dimérisation des protéines Erg.....	73
a-Le domaine ETS.....	73
b-Le domaine PNT.....	74
c-Les domaines ETS et PNT requis dans d'autres types d'interaction.....	74
3-L'implication fonctionnelle de ces interactions.....	74
4-Interaction des protéines Erg avec le dimère Fos/Jun.....	75
5-Différences entre les protéines Erg.....	76
6-Les isoformes.....	77
a-Un exemple dans la famille Ets.....	77
b-Les isoformes SRC1 et le récepteur des oestrogènes.....	77
c-Deux isoformes au cours de la myogenèse.....	78
III-Publication 3.....	80
1-Introduction.....	80
2-Matériels et méthodes.....	81
3-Résultats.....	83
a-L'expression des isoformes Erg varie au cours du temps.....	83
b-Les gènes <i>erg</i> et <i>fli-1</i> sont coexprimés de manière précoce.....	85
4-Discussion.....	85
a-L'ADNc <i>ergp55</i> semble majoritaire.....	86
b-Le gène <i>fli-1</i> est à l'origine d'un seul transcrit.....	86
c-L'expression des gènes <i>erg</i> et <i>fli-1</i> .....	87

## DISCUSSION GENERALE

## REFERENCES BIBLIOGRAPHIQUES

## Le stage de thèse

Quand je suis arrivée en septembre 1994 dans le laboratoire d'Oncologie Moléculaire, sous la direction de Dominique STEHELIN, j'ai intégré le groupe dirigé par Martine DUTERQUE-COQUILLAUD. Après avoir obtenu le DEA sciences de la vie et de la santé en juin 1995, j'ai bénéficié d'une bourse cofinancée par l'Institut Pasteur de Lille et le conseil régional Nord-Pas de Calais pour effectuer ma thèse de doctorat.

Durant ces quatre années, j'ai effectué mes recherches avec principalement l'aide technique de Anne FLOURENS. Si le thème de recherche de notre équipe consistait à étudier la fonction du gène *erg*, à la fois *in vivo* et *in vitro*, ma part du travail était de déterminer le rôle spécifique des isoformes issues du gène *erg* et leur mode d'action *in vitro*.

## Publications

Butticé G., Duterque-Coquillaud M., Basuyaux J.P., Carrère S., Kurkinen M. and Stéhelin D.

"Erg differentially regulates human collagenase1 and stromelysin1 genes by interacting with the Fos/Jun complex." *Oncogene* (1996) **13**, 2297-2306

Carrère S., Verger A., Flourens A., Stéhelin D. and Duterque-Coquillaud M.

"Erg proteins, transcription factors of the Ets family, form homo, heterodimers and ternary complex *via* two distinct domains." *Oncogene* (1998) **16**, 3259-3266

Carrère S., Vlaeminck-Guillem V., Flourens A., De Witte F., Dumont P., Desbiens X., Stéhelin D. and Duterque-Coquillaud M.

"Mouse *erg* gene, an *ets* family member, has at least two splice variants expressed very early in mouse development." *in preparation*

## Présentation du mémoire

Pourquoi étudier des facteurs de transcription dans un laboratoire intitulé "Oncologie moléculaire"?

Nous savons désormais que le cancer ne représente qu'une "déviation" des mécanismes fondamentaux, tels la division, la croissance et la différenciation cellulaires. Un modèle de "genèse du cancer" existe aujourd'hui sur lequel tout le monde s'accorde. Les cellules deviennent cancéreuses par suite d'altérations génétiques qui résultent de mutations spontanées ou induites par des cancérigènes chimiques ou des radiations. Des gènes normaux deviennent alors des "gènes du cancer" ou oncogènes. Depuis les travaux initiaux effectués en 1976, de nombreux oncogènes ont été découverts dans des espèces animales très variées. Ils partagent les propriétés suivantes : i/ ce sont des gènes normaux de notre ADN, exprimés dans certains tissus, qui remplissent des fonctions fondamentales dans notre physiologie comme la croissance, la différenciation ; ii/ ils ont été conservés au cours de l'évolution depuis plus de 500 millions d'années. De ces deux propriétés découlent deux conséquences importantes : i/ l'identification d'un oncogène chez l'animal permet d'identifier rapidement son homologue chez l'homme ; ii/ ces gènes sont impliqués dans des cancers qui possèdent des mécanismes semblables chez l'homme et chez l'animal. Par ailleurs, l'étude biochimique des produits des oncogènes identifiés a démontré que plusieurs d'entre eux codaient des protéines se liant à l'ADN et régulant la transcription de gènes impliqués dans la croissance et la différenciation cellulaires.

Ces oncogènes forment de véritables familles nombreuses qui ne cessent de croître. Une de ces familles de gènes est étudiée depuis longtemps dans notre laboratoire, la famille des gènes *ets*. La découverte des protéines Ets en tant que facteurs de transcription fournit une base pour la compréhension du potentiel oncogénique de ces gènes. De plus, ces protéines deviennent un modèle pour étudier les mécanismes moléculaires de la transcription.

Dans notre équipe, nous nous intéressons en particulier au gène *erg* appartenant à la famille *ets*. Différentes approches *in vivo* et *in vitro* sont utilisées pour cerner son rôle. Pour ma part, je me suis intéressée aux modes d'action des différentes protéines isoformes issues de ce gène. En effet, nous avons cherché à comprendre la fonction propre de chacune d'entre elles. Ainsi, les aspects de leur spécificité pourrait être lié i/ à des interactions avec des cofacteurs différents, ii/ à leur affinité vis-à-vis de différents promoteurs cellulaires, iii/ à leur localisation et leur taux d'expression. C'est ce qui fera l'objet de ce mémoire après vous avoir présenté la famille *ets* puis plus particulièrement le gène *erg*.

## Abréviations et Conventions

<b>ADN</b>	Acide désoxyribonucléique
<b>ADNc</b>	Acide désoxyribonucléique complémentaire
<b>AP-1</b>	"Activator Protein-1" ou protéine activatrice-1 qui désigne le complexe Fos/Jun ou le site de liaison à l'ADN de ce même complexe
<b>ARN</b>	Acide ribonucléique
<b>ARNm</b>	Acide ribonucléique messenger
<b>EBS</b>	"Ets binding site", site de liaison des protéines Ets
<b>ES cell</b>	"Embryonic Stem cell" ou cellule embryonnaire souche
<b>FMuLV</b>	"Friend murine Leukemia Virus" ou virus leucémogène de Friend
<b>GST</b>	Glutathione-S-transférase
<b>HLH</b>	"Helix-Loop-Helix" ou hélice-boucle-hélice
<b>HTH</b>	"Helix-Turn-Helix" ou hélice-tour-hélice
<b>kb</b>	kilobase
<b>kDa</b>	kilodalton
<b>KO</b>	"Knock-out" ou mutation ciblée
<b>M</b>	mole/litre
<b>PCR</b>	"polymerase chain reaction" ou amplification en chaîne par la polymérase
<b>pdb</b>	paires de bases
<b>PNT</b>	domaine POINTED
<b>RMN</b>	résonance magnétique nucléaire
<b>wHTH</b>	"winged Helix-Turn-Helix" ou hélice-tour-hélice de type ailé

Par convention, le nom des gènes est noté en italique et celui des protéines commence par une majuscule.

Les termes anglais d'usage courant, préférés aux termes français rarement utilisés, figurent entre guillemets.

# **Introduction**



## Première partie : La famille des gènes *ets*

### I- L'oncogène viral *v-ets*

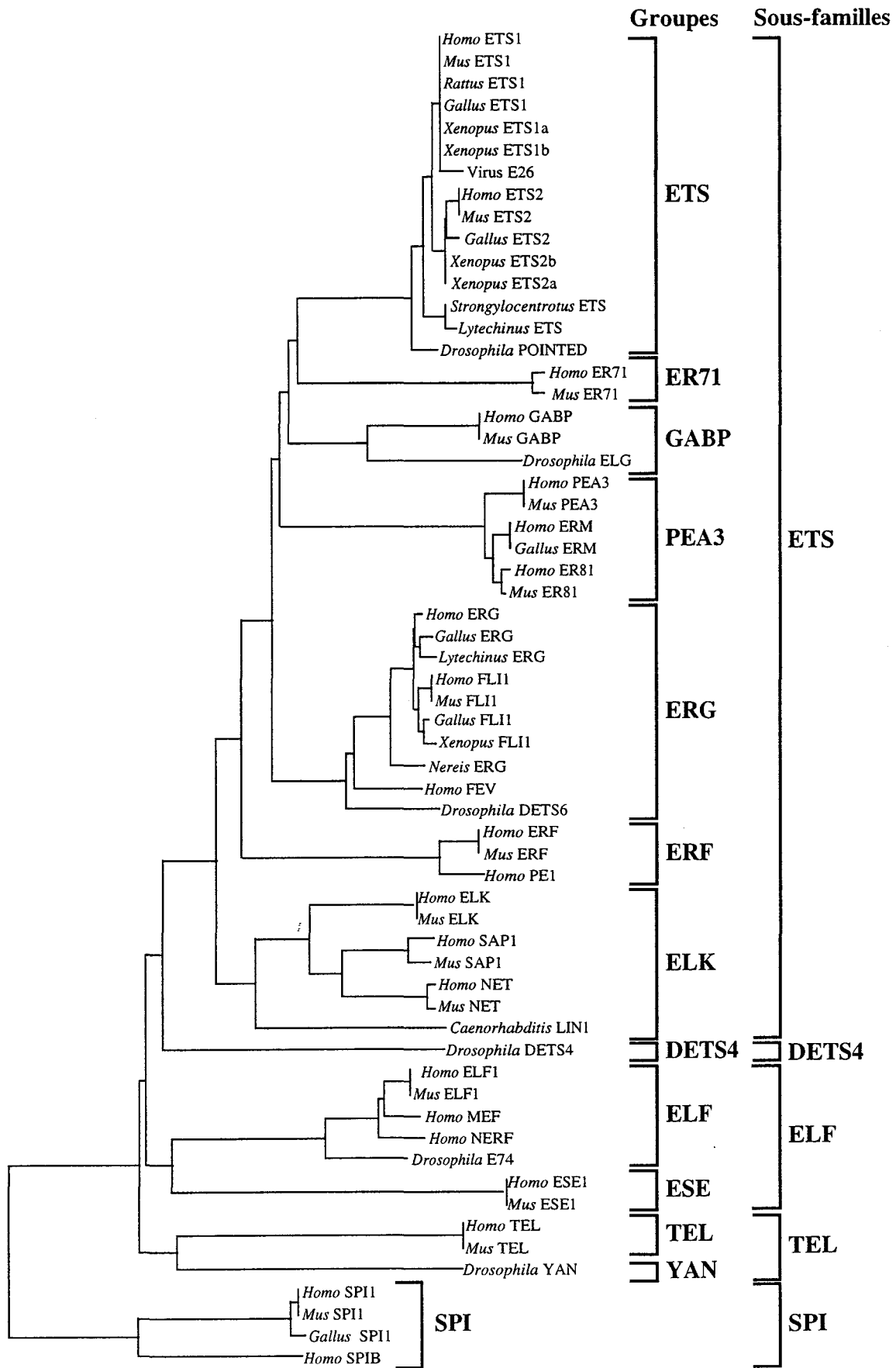
Le modèle des rétrovirus oncogènes aviaires a été le premier moyen d'étudier les mécanismes moléculaires qui président à la transformation cellulaire. En effet, le génome de ces rétrovirus comporte des séquences appelées oncogènes viraux qui ont pour origine des gènes, ou proto-oncogènes, appartenant à l'hôte infecté. Ces oncogènes diffèrent par mutation ou délétion (ces deux phénomènes n'étant pas mutuellement exclusifs) de leurs homologues cellulaires et s'expriment sous le contrôle de séquences virales. Ainsi, ils donnent la capacité aux rétrovirus qui les contiennent de transformer les cellules infectées, alors qu'un rétrovirus ne présentant pas d'oncogène n'est pas, ou très faiblement, tumorigène.

E26 est un de ces rétrovirus aviaires oncogènes. Il provoque chez le poulet des leucémies mixtes érythroïdes-myéloïdes et il transforme, *in vitro*, des érythroblastes et des myéloblastes. Son génome contient deux oncogènes, *v-myb* et *v-ets* ("E-twenty-six" d'après le nom du rétrovirus E26), dont les séquences sont fusionnées entre elles et à celle du gène viral *gag* (Leprince & al., 1983 ; Nunn & al., 1983). Elles s'expriment sous la forme d'une protéine virale de fusion, *p135gag-myb-ets*. Les oncogènes *v-myb* et *v-ets* coopèrent pour conférer au rétrovirus E26 son pouvoir transformant (Golay & al., 1988).

L'étude du rétrovirus E26 a ainsi permis la caractérisation de *v-ets*, le membre fondateur de la famille des gènes *ets*. Des sondes obtenues à partir de cet oncogène viral ont permis l'isolement et la caractérisation d'un gène cellulaire homologue, *c-ets-1*, chez le poulet puis chez l'homme (Gégonne & al., 1987).

### II- La famille des gènes *ets*

L'identification de cet oncogène viral *v-ets* et l'isolement de son homologue cellulaire *c-ets-1* ont été le point de départ de la caractérisation d'une nouvelle famille toujours croissante, la famille des gènes *ets*. Ils codent des facteurs de transcription dont la signature est un domaine protéique très conservé de 85 acides aminés, responsable de la fixation à l'ADN, le domaine ETS (Karim et al., 1990). Un des moyens d'étudier cette famille est de considérer les relations qui existent entre ses membres. Pour cela, nous nous intéressons à leur phylogénie.



**Figure 1** : Arbre phylogénétique de la famille des gènes *ets*. D'après Laudet et al., sous presse.

## 1- Phylogénie des membres de la famille *ets*

A ce jour, la famille des gènes *ets* comprend au moins vingt-deux membres différents décrits chez les mammifères (*ets-1*, *ets-2*, *er71*, *gabpa*, *pea3*, *erm*, *er81*, *erg*, *fli-1*, *fev*, *erf*, *pe1*, *elk1*, *sap1*, *net*, *elf1*, *mef*, *nerf*, *ese-1*, *tel*, *pu-1* et *spiB*), sept chez la drosophile (*pointed*, *elg*, *dets3*, *dets6*, *E74*, *dets4* et *yan*), un chez *Caenorhabditis elegans* (*lin1*) et six autres, pas encore caractérisés, mis en évidence à partir du projet de séquençage du génome de nématode (Laudet et al., sous presse).

L'analyse des séquences complètes et de celles, hautement conservées, du domaine ETS ont permis la mise en place d'une classification de la famille des protéines Ets. Par exemple, la protéine Pu-1 est considérée comme le membre le plus éloigné de la protéine Ets-1, puisque son domaine ETS ne présente que 37% d'identité de séquence en acides aminés avec celui de la protéine Ets-1. En revanche, le domaine ETS de la protéine Ets-2, qui présente 96% d'identité avec celui de la protéine Ets-1, est considéré comme le membre le plus proche. Nous pouvons aussi remarquer que les domaines ETS des protéines Erg et Fli-1 ont 70% d'identité avec celui de la protéine Ets-1 mais présentent 98% d'identité entre eux.

Ainsi, l'analyse du taux de divergence entre les différents membres montre que la famille peut être divisée en plusieurs groupes (Fig 1). Une première équipe avait observé six groupes : ETS, ERG, GABP, ELK, E74 et SPI (Lautenberger et al., 1992). Des groupes supplémentaires ont été recensés par la suite : PEA3, DETS4 et YAN (Laudet et al., 1993), ER71, ERF, ESE et TEL (Laudet et al., sous presse). Ces derniers correspondent chacun à un seul gène nouvellement identifié, *er71* (Brown and McKnight, 1992), *erf* (Sgouras et al., 1995), *ese* (Oettgen et al., 1997) et *tel* (Golub et al., 1994). Cela suggère donc que la description de la famille Ets est loin d'être complète et qu'il reste d'autres gènes et groupes à décrire. Ainsi, le nouveau membre de la famille caractérisé, *ehf* ("ets homologous factor"), appartiendrait au groupe ESE. Ces treize groupes ont été associés en cinq sous-familles : ETS, DETS4, ELF, TEL et SPI.

Les membres de la famille *ets* sont des gènes qui, par leurs similitudes, dérivent d'un seul ancêtre commun. Ils sont donc définis comme orthologues ou paralogues. Les orthologues sont des versions d'un même gène dans des espèces différentes tandis que les paralogues coexistent dans un même génome et sont dérivés de duplications de gènes.

## 2- Les domaines d'homologie autres que le domaine ETS

A l'intérieur des groupes paralogues, les différents membres se ressemblent beaucoup, même en dehors du domaine ETS. C'est le cas du groupe PEA3. Les produits des gènes *pea3*, *erm* et *er81* possèdent un domaine acide et une région riche en résidus

glutamine (de Launoit et al., 1996). Enfin, les gènes *erg* et *fli-1* qui appartiennent au même groupe ERG sont parmi les membres les plus proches à l'intérieur de cette grande famille. En effet, les produits de ces gènes présentent un pourcentage d'identité de séquence en acides aminés de 67% sur l'ensemble de la protéine qui s'élève à 98% dans le domaine ETS.

De plus, entre les groupes paralogues, nous constatons des identités de séquence. Ainsi, les gènes *ets-1*, *ets-2*, *gabpa*, *erg*, *fli-1*, *ese-1*, *tel*, *pnt*, *elg* et *yan*, appartenant à différents groupes et sous-familles, présentent, dans la partie amino-terminale de leurs produits, un second domaine conservé, le domaine Pointed (PNT). Avant l'analyse phylogénétique, il aurait été tentant d'associer dans une même sous-famille tous ces gènes. Cependant, la divergence de séquence de leur domaine ETS ne permet pas de classer les gènes *tel* et *yan* dans cette sous-famille.

L'hypothèse la plus simple pour expliquer la présence, parmi les paralogues, de ces séquences similaires en dehors du domaine ETS serait une duplication suivie de rétention de séquences.

### 3- Une origine pour la famille *ets*?

L'origine de la famille Ets est très ancienne et pourrait remonter aux tous premiers métazoaires. En effet, ces gènes sont présents dans tous les phyla testés : chez les métazoaires inférieurs c'est-à-dire les éponges, les cnidaires, les anémones, les plathelminthes et les nématodes, chez les invertébrés et les vertébrés. Par contre, aucun gène n'a été identifié chez les protozoaires, les plantes, les champignons (Degnan et al., 1993) ou même dans le génome maintenant entièrement connu de *Saccharomyces cerevisiae*. Les gènes *ets* sont donc spécifiques des animaux multicellulaires.

De plus, il est étonnant d'observer une grande conservation aussi bien au niveau de la séquence des protéines issues de ces gènes, principalement dans le domaine ETS, que du nombre de groupes dans la famille Ets, depuis le commencement de la vie métazoaire. Cela suggère clairement que les gènes *ets*, dont le rôle de certains commence à être connu, ont des fonctions de base dans la signalisation cellulaire et dans le développement embryonnaire. Par ailleurs, la grande diversité de ces gènes, très conservés dans la région qui code le domaine ETS, permet d'envisager des fonctions différentes mais aussi de possibles redondances (Laudet et al., sous presse).

En ce qui concerne la racine de l'arbre phylogénétique de cette famille Ets (Fig 1), des études récentes de la structure du domaine ETS ont relié les gènes *ets* avec des membres de la famille des facteurs de transcription hélice-tour-hélice de type ailé (wHTH), famille très ancienne déjà présente chez les procaryotes (Donaldson et al., 1994 ; Liang et al., 1994a).

GROUPE	PROTEINE	CONSENSUS	REFERENCE
ETS	m-Ets-1	ACCGGAAGT G A TAC	(Nye <i>et al.</i> , 1992)
	h-Ets-1	GCGGAAGT CA C	(Fisher <i>et al.</i> , 1991)
ER71	m-Er71	GCCGGAAGTGC CGA TAC	(Brown and McKnight, 1992)
GABP	m-Gabp $\alpha$	GCCGGAAGT AGA TAC	(Brown and McKnight, 1992)
PEA3	m-Er81	GGCGGAAGT ACA TAC	(Brown and McKnight, 1992)
ERG	h-Erg	CCGGAAGT GA A	(Murakami <i>et al.</i> , 1993)
	m-Fli-1	TNGACCGGAAGTAG G A ACGA	(Mao <i>et al.</i> , 1994)
ELK	m-Sap-1	ACCGGAAGT TAC	(Shore and Sharrocks, 1995)
ELF	m-Elk-1	AACCGGAAGTG A	(Shore and Sharrocks, 1995)
	d-E74	AACCCGGAAGT T A	(Urness and Thummel, 1990)
SP1	m-Elf-1	AACCCGGAAGTA TA C G	(John <i>et al.</i> , 1996)
	m-Pu.1	AAAAAGAGGAAGTAG TCC C GC G	(Ray-Gallet <i>et al.</i> , 1995)
	m-Spi-B	AAAAAGAGGAAGTA TCC CT	(Ray-Gallet <i>et al.</i> , 1995)

**Figure 2** : Séquences consensus sélectionnées par différentes protéines Ets.  
Le "coeur" très conservé GGAA/T est encadré.

Les séquences de ces gènes ont été intégrées dans un autre arbre phylogénétique beaucoup plus large. Cette étude permet alors de situer la racine de l'arbre de la famille Ets entre la sous-famille SPI et toutes les autres. Ces résultats sont en accord avec la faible homologie entre les gènes de la sous-famille SPI et ceux des autres sous-familles. Cependant, la question de l'origine de la famille Ets reste toujours ouverte (Laudet et al., sous presse).

Nous venons d'évoquer deux domaines protéiques conservés, le domaine ETS dans tous les membres de la famille des protéines Ets, et le domaine PNT présent chez quelques membres. Nous allons maintenant détailler les connaissances acquises à ce jour sur leur structure et leur fonction respectives. Ces connaissances seront des éléments de réponse à la grande question qui est : "Comment cette grande famille de facteurs de transcription homologues peut réguler différemment différents gènes dans différents types cellulaires et en réponse à différents signaux extracellulaires?".

### **III- Le domaine ETS**

Une des spécificités d'action des facteurs de transcription correspond à leur capacité de reconnaître des séquences d'ADN spécifiques. Le domaine ETS, caractéristique commune à toutes les protéines Ets, est nécessaire et suffisant pour la liaison à l'ADN ce qui en fait un domaine à la fois structural et fonctionnel (Karim et al., 1990 ; Graves et Petersen, 1998). Afin de comprendre cette spécificité de liaison à l'ADN à un niveau moléculaire, des études des séquences nucléotidiques cibles et des informations structurales sont requises.

#### **1- Les séquences cibles**

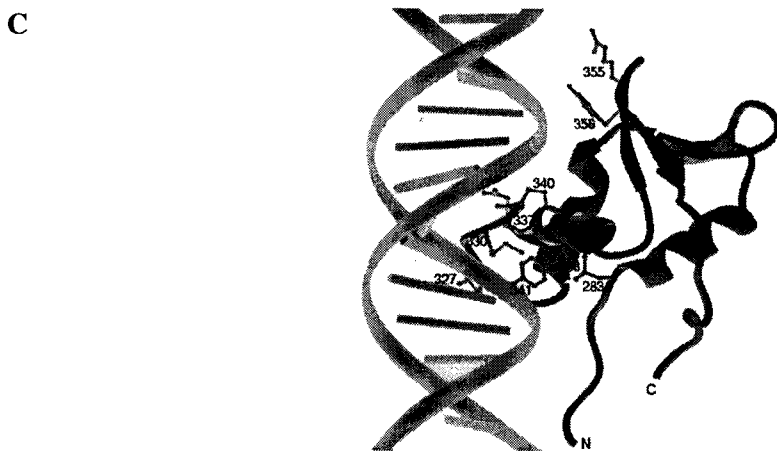
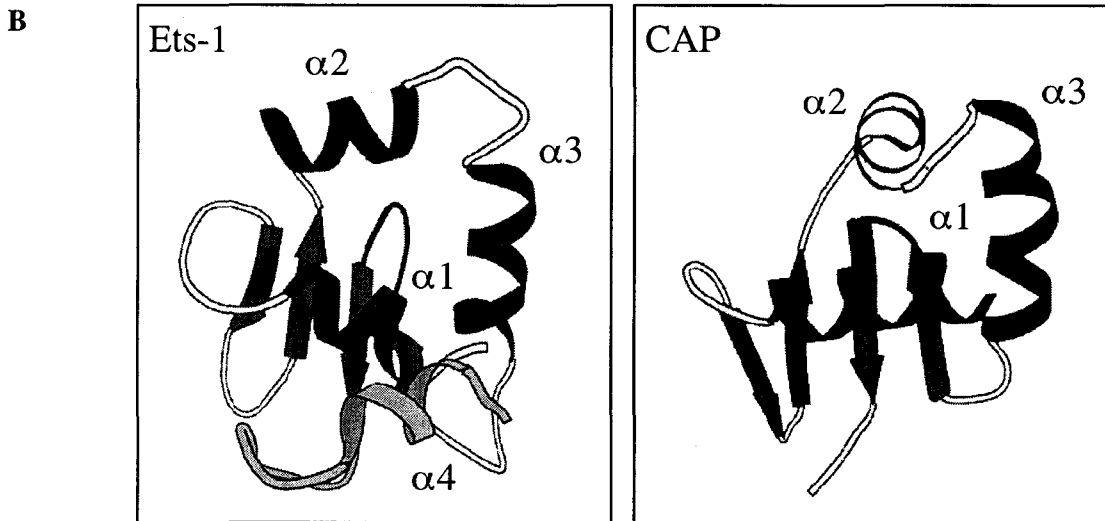
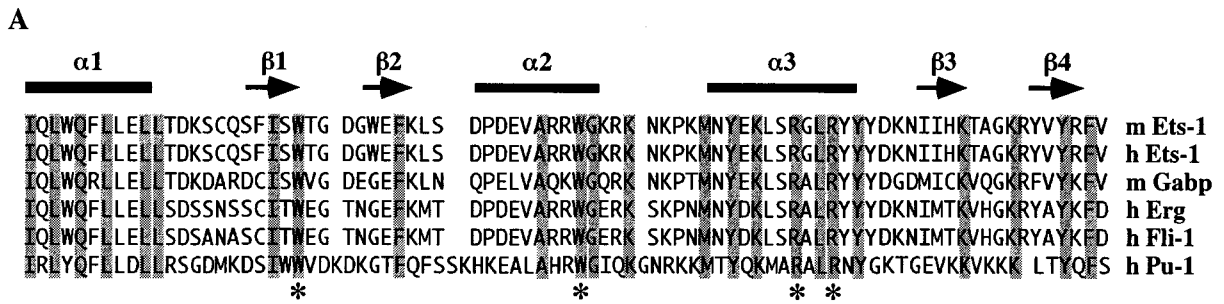
Du fait de la grande conservation des séquences protéiques du domaine ETS, les protéines Ets reconnaissent des séquences d'ADN très similaires. Toutefois, une sélection *in vitro* d'oligonucléotides aléatoires ainsi qu'une comparaison des sites de liaison détectés dans les promoteurs ou "enhancers" ont permis à plusieurs équipes de déterminer la séquence consensus préférentielle pour des protéines de la famille Ets (Fig 2). Le motif central de reconnaissance, 5'- GGAA/T -3', est présent dans la vaste majorité de tous les sites biologiques (Graves and Petersen, 1998). L'interaction de ces protéines avec seulement 4 paires de bases ne peut permettre une haute affinité de liaison à l'ADN. De plus, si les membres de la famille régulent différents gènes cibles par des propriétés de liaison distinctes, des contacts supplémentaires avec l'ADN sont impliqués. Des études ont démontré que le domaine ETS interagit avec une seule face de l'ADN ce qui implique un large domaine, principalement le grand sillon de la double hélice mais aussi les petits sillons

adjacents. C'est pourquoi, Nye et al (1992) suggèrent un nouveau mode de fixation à l'ADN. Ainsi, la sélection du site de liaison correspond à une approche différente. En effet, les nucléotides qui encadrent la séquence consensus GGAA/T sont nécessaires non seulement pour la fixation à l'ADN mais aussi pour la spécificité de reconnaissance des protéines Ets.

Concernant les nucléotides variables de part et d'autre de la séquence consensus GGAA/T, les résultats suivants permettent de donner un premier élément de réponse afin de déterminer la spécificité des protéines Ets. Les protéines Ets-1 et Elf-1, coexprimées dans les cellules T activées ou non, reconnaissent de manière préférentielle des cibles différentes (Wang et al., 1992). En effet, les nucléotides situés en 3' de la séquence consensus leur permettent une fine spécificité de liaison à l'ADN. La protéine Ets-1 se lie indifféremment à la cible GGAA/T alors que la protéine Elf-1 ne peut se lier qu'à la cible GGAA. De même, les sites consensus suivis des nucléotides NCG ou NTG fixent Ets-1 alors que ceux suivis par NAA ou NAG ne peuvent pas. Ce mécanisme pourrait tenir compte, au moins en partie, de l'expression simultanée de ces deux protéines dans les cellules T. Ainsi, les "enhancers" des gènes *TCR $\alpha$*  et *TCR $\beta$*  lient la protéine Ets-1 mais pas la protéine Elf-1 (Wang et al., 1992), tandis que l'"enhancer" du gène *IL-2* lie la protéine Elf-1 et pas la protéine Ets-1. (John et al., 1996).

En fait, l'ensemble des séquences sélectionnées est relativement homogène autour du "coeur" très conservé GGAA/T (Fig 2). Il semble cependant que la discrimination *in vivo* des éléments de différents promoteurs est gouvernée non seulement par ces subtiles différences dans la liaison à l'ADN mais aussi par d'autres mécanismes. En réalité, nous ne connaissons pas la relation entre l'affinité pour l'ADN d'un facteur Ets mesurée *in vitro* et son efficacité en transactivation *in vivo*. Cette affinité pour l'ADN n'est sans doute pas seule en cause puisque des interactions protéine-protéine pourrait participer à la sélection nécessaire pour discriminer différents promoteurs par divers membres de la famille Ets (Brown and McKnight, 1992).

Par ailleurs, les protéines Ets ont toujours été décrites comme se fixant en monomères (Crépieux et al., 1994 ; Graves and Petersen, 1998). Seule la protéine Gabp $\alpha$  se fixe en hétérodimères puisqu'elle se lie, avec une plus forte affinité, à des sites de type GGAA/T en présence de la protéine Gabp $\beta$  (Lamarco et al., 1991 ; Thompson et al., 1991 ; Batchelor et al., 1998). Une explication de la spécificité des protéines Ets pourrait venir de la disposition



**Figure 3** : Le domaine ETS "Hélice-Tour-Hélice de type ailé". **A**- Alignement des séquences du domaine ETS des membres de la famille Ets dont la structure a été résolue. La structure secondaire est indiquée par des traits pour les hélices  $\alpha$  et des flèches pour les feuillets  $\beta$ . Les résidus tryptophane et arginine très conservés dans la famille sont indiqués par une astérisque. **B**- Diagrammes comparant le domaine de liaison à l'ADN de la protéine Ets-1 murine et de la protéine bactérienne CAP. D'après Donaldson et al., 1994. **C**- Représentation schématique de la structure tridimensionnelle du domaine ETS de la protéine Fli-1 humaine en présence d'ADN. D'après Liang et al., 1994b.

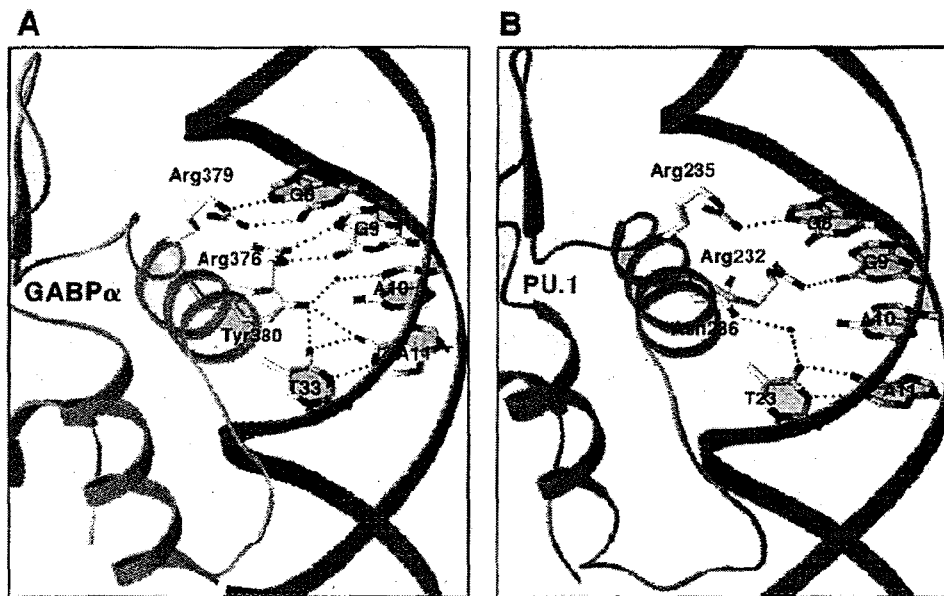


dans l'espace de ce domaine ETS lié à sa cible, c'est pourquoi différents auteurs se sont penchés sur l'étude de sa structure.

## 2- La structure du domaine ETS

Jusqu'alors, les comparaisons de séquences primaires des domaines ETS n'avaient jamais révélé une homologie convaincante avec les autres motifs de liaison à l'ADN et bien caractérisés comme le "doigt de zinc", le domaine "homéo", le motif "hélice-tour-hélice" ou encore la "glissière à leucine" (Karim & al., 1990). Des observations avaient conduit différents auteurs à conclure que la famille des protéines Ets comportait un nouveau motif structural de liaison à l'ADN (Gunther et al., 1990 ; Nye et al., 1992). Cependant, de nouvelles études de structure secondaire du domaine ETS des protéines Ets-1 murine et Fli-1 humaine ont démontré qu'il est composé de trois hélices  $\alpha$  et de quatre feuillets  $\beta$  (Donaldson et al., 1994 ; Liang et al., 1994a). Les trois hélices ( $\alpha 1$ -  $\alpha 3$ ) et les quatre feuillets ( $\beta 1$ -  $\beta 4$ ) sont disposés dans l'ordre  $\alpha 1$ -  $\beta 1$ -  $\beta 2$ -  $\alpha 2$ -  $\alpha 3$ -  $\beta 3$ -  $\beta 4$  (Fig 3A). Ces résultats ont montré que le domaine ETS est structurellement similaire à un domaine de liaison à l'ADN d'abord découvert chez la protéine activatrice du gène bactérien *catabolite*, CAP (Fig 3B). En fait, Donaldson et al. (1994) et Liang et al. (1994a) suggèrent que les protéines Ets appartiennent à une vaste famille, celle des protéines de type "hélice-tour-hélice de type ailé" ou "wHTH". Dans ce contexte, la région des feuillets  $\beta$  associée à des boucles est considérée comme les ailes de ce domaine HTH. Cette comparaison a aussi permis aux auteurs de suggérer que l'hélice  $\alpha 3$  interagit avec le grand sillon de la double hélice de l'ADN et que la région ailée pourrait intervenir dans des contacts supplémentaires avec le petit sillon de l'ADN (Donaldson et al., 1994 ; Liang et al., 1994a).

La structure tertiaire du domaine ETS a été décrite par résonance magnétique pour les protéines Fli-1 humaine (Liang et al., 1994b), Ets-1 murine (Donaldson et al., 1996) et Ets-1 humaine (Werner et al., 1995 et 1997), et par cristallographie pour les protéines Pu-1 humaine (Kodandapini et al., 1996) et Gabp $\alpha$  murine liée à Gabp $\beta$  (Batchelor et al., 1998). Dans tous les cas, les trois hélices  $\alpha$  forment un motif HTH, c'est-à-dire un noyau hydrophobe qui s'appuie sur l'"échafaudage" formé par quatre brins d'un petit feuillet  $\beta$  antiparallèle (Fig 3B). Deux résidus tryptophane, très conservés dans la famille (Fig 3A), sont "enterrés" dans le coeur hydrophobe de la structure où ils apparaissent jouer un rôle structural. C'est l'hélice  $\alpha 3$  qui reconnaît la séquence consensus GGAA/T, dans le grand sillon de la double hélice de l'ADN. Deux acides aminés arginine invariants chez tous les membres de la famille, situés dans cette hélice de reconnaissance (Fig 3A), interagissent avec les guanines de la séquence cible (Liang et al., 1994b ; Donaldson et al., 1996 ;



**Figure 4** : Les liaisons hydrogène entre l'hélice  $\alpha$ 3 du domaine ETS et la séquence cible consensus GGAA/T. **A-** Interaction de l'hélice  $\alpha$ 3 du domaine ETS de Gabp $\alpha$  avec le grand sillon de la double hélice de l'ADN. **B-** Interaction de l'hélice  $\alpha$ 3 du domaine ETS de Pu-1 avec le grand sillon. D'après Batchelor et al., 1998.

Kodandapani et al., 1996 ; Batchelor et al., 1998). Lors de la description des différents domaines ETS, des particularités ont cependant été observées.

#### a- Structure du domaine ETS de la protéine Fli-1 humaine

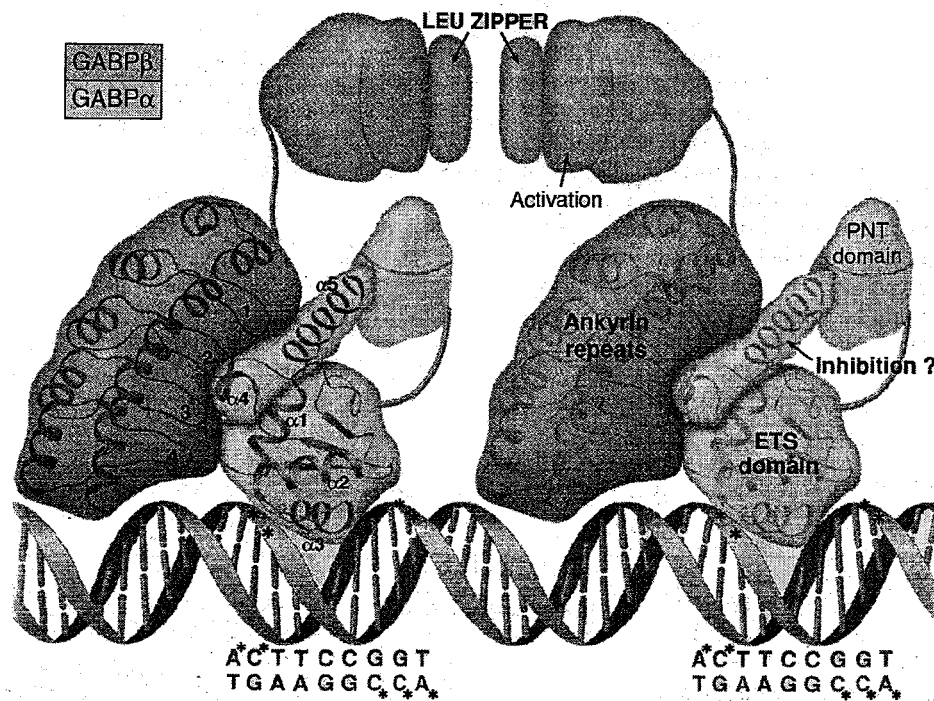
Pour la protéine Fli-1 humaine, en plus des deux acides aminés arginine décrits précédemment, un résidu tyrosine localisé dans la même hélice  $\alpha 3$ , interagit avec l'ADN (Fig 3C) (Liang et al., 1994b). Par ailleurs, des analyses d'interférence et de protection chimique ont montré que le domaine ETS interagit avec le grand sillon et les petits sillons adjacents de la double hélice de l'ADN (Nye et al., 1992). Les auteurs suggèrent ici que la boucle entre les feuillets  $\beta 3$  et  $\beta 4$  interagit avec un petit sillon adjacent d'un côté et celle entre les hélices  $\alpha 2$  et  $\alpha 3$  avec la partie amino-terminale de l'hélice  $\alpha 1$  dans le petit sillon opposé (Liang et al., 1994b).

#### b- Structure du domaine ETS de la protéine Ets-1 murine

Pour la protéine Ets-1 murine, l'étude du domaine ETS a été faite sans ADN et montre que les boucles supposées interagir avec les petits sillons de l'ADN sont plus grandes que pour la protéine Fli-1 (Donaldson et al., 1996). En effet, la taille et la séquence des tours entre les hélices  $\alpha 2$  et  $\alpha 3$  et entre les feuillets  $\beta 3$  et  $\beta 4$  diffèrent parmi les protéines Ets. Ce sont donc ces régions qui participeraient la spécificité d'interaction de chaque protéine Ets vis-à-vis d'une séquence cible. Cependant, dans cette étude, les auteurs n'ont pas utilisés de fragment d'ADN ce qui ne permet pas de déterminer l'orientation de la molécule par rapport à la double hélice de l'ADN. De plus, la particularité de la protéine Ets-1 est que l'extrémité carboxy-terminale forme une hélice  $\alpha 4$  qui se rabat sur l'hélice  $\alpha 1$  du domaine ETS (Fig 3B). Elle correspond en fait à une séquence inhibitrice de la liaison à l'ADN, très conservée chez les protéines Ets-1, Ets-2 et Gabp $\alpha$  (Hagman and Grosschedl, 1992). L'inhibition pourrait en effet s'expliquer par cette structure particulière qui empêcherait le domaine ETS de se fixer correctement.

#### c- Structure du domaine ETS de la protéine Ets-1 humaine

En ce qui concerne la protéine Ets-1 humaine, Werner et al. (1995 et 1997) ont précisé que la boucle située entre les feuillets  $\beta 3$  et  $\beta 4$  interagit avec le petit sillon situé en 5' de la séquence consensus tandis que celle située entre les hélices  $\alpha 2$  et  $\alpha 3$  est en contact avec le petit sillon en 3'. Par contre, ils n'ont pas pu mettre en évidence les interactions des deux résidus arginine invariants avec la séquence GGAA/T.



**Figure 5 :** La spécificité de liaison à l'ADN de la protéine Gabpα est déterminée par la formation d'un hétérotétramère (Gabpα-Gabpβ)<sub>2</sub>. D'après Graves, 1998.

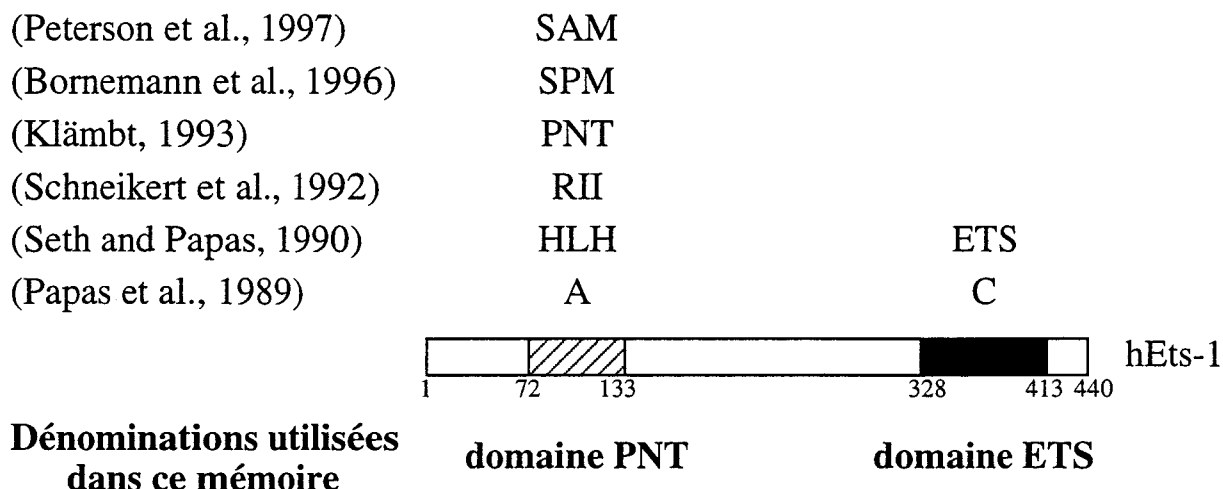
#### d- Structure du domaine ETS de la protéine Pu-1 humaine

Le domaine ETS de la protéine Pu-1 humaine a été cristallographié avec un fragment d'ADN comportant la séquence consensus. Bien que la protéine Pu-1 soit le membre le plus éloigné phylogénétiquement des autres déjà étudiés, son domaine de liaison à l'ADN garde les principales caractéristiques déjà décrites. Des spécificités lui sont quand-même octroyées. Ainsi, sur les deux résidus arginine déjà décrits qui établissent des liaisons hydrogène avec les deux guanines de la séquence cible consensus, le deuxième se lie plus faiblement (Fig 4). Ceci permet à la protéine de reconnaître, en plus du motif GGAA/T, le motif AGAA/T (Batchelor et al., 1998 ; Ely et al., 1998). Trois autres acides aminés situés dans l'hélice  $\alpha 3$  permettent des interactions avec l'ADN *via* des molécules d'eau : une thréonine, une glutamine et une asparagine. Ces résidus ne sont pas conservés dans la famille ce qui suggère que ces interactions sont spécifiques à la protéine Pu-1. Par ailleurs ce sont les mêmes régions "ailées" qui entrent en contact avec les petits sillons, les auteurs ont même pu déterminer un résidu lysine dans chaque "aile" qui interagit avec le squelette de l'ADN. Ici, la boucle située entre les hélices  $\alpha 2$  et  $\alpha 3$  est beaucoup plus grande que chez les autres membres déjà étudiés (Kodandapani et al., 1996).

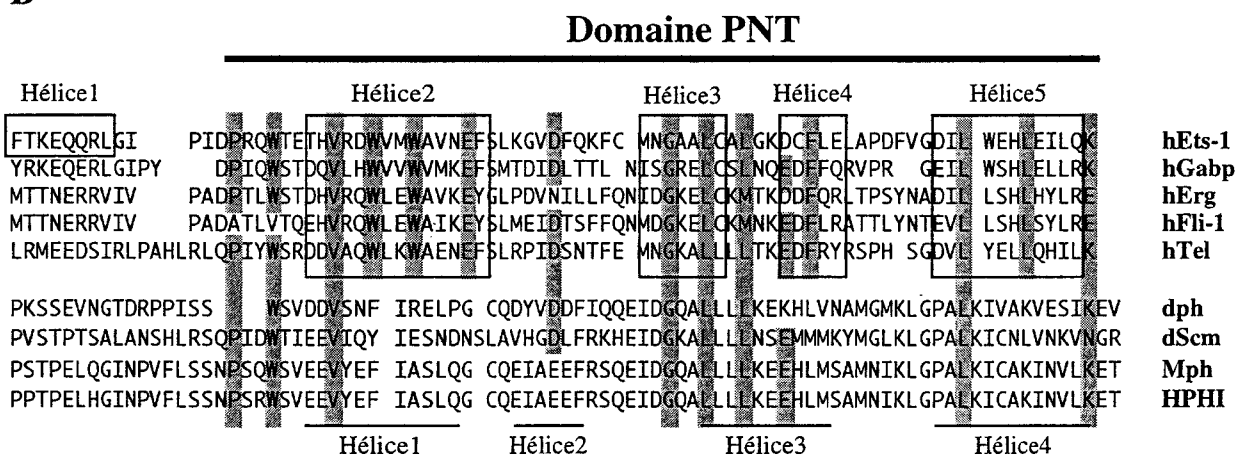
#### e- Structure du domaine ETS de la protéine Gabp $\alpha$ murine

La dernière structure déterminée par cristallographie concerne le domaine ETS de la protéine murine Gabp $\alpha$  ainsi que la trentaine d'acides aminés suivants (Batchelor et al., 1998 ; Ely et al., 1998). Ces derniers forment une extrémité carboxy-terminale qui permet à la protéine, une fois liée à l'ADN, d'interagir avec la protéine murine Gabp $\beta$ . Cette dernière s'associe au domaine ETS de la protéine Gabp $\alpha$  en insérant ses quatre domaines ankyrine et demi (deux hélices  $\alpha$  suivies d'une grande boucle perpendiculaire avec un feuillet  $\beta$  à son extrémité) dans une dépression formée par l'hélice  $\alpha 1$  et les deux hélices  $\alpha$  de l'extrémité carboxy-terminale. Bien que la protéine Gabp $\beta$  soit proche de l'ADN, elle ne rentre pas en contact direct avec lui. La protéine Gabp $\alpha$ , elle, se fixe à l'ADN grâce à l'hélice  $\alpha 3$ , comme les autres domaines ETS, et une de ses "ailes" correspondant aux hélices  $\alpha 1$  et  $\alpha 2$  contacte le petit sillon en 3' de la séquence consensus GGAA/T. Tout comme pour la protéine Pu-1, une légère courbure de l'ADN permet d'améliorer le contact. De plus, une liaison optimale à l'ADN de la protéine Gabp $\alpha$  requiert davantage qu'une interaction directe entre le domaine ETS et les répétitions ankyrine (Graves, 1998). Un motif de glissière à leucine de la protéine Gabp $\alpha$  absent de la structure cristallographiée permet la formation d'un hétérotétramère (Fig 5). Cette configuration reconnaît deux sites consensus et permet une meilleure affinité de liaison à l'ADN. Enfin, nous avons vu que l'extrémité carboxy-terminale de la protéine Gabp $\alpha$ , très conservée chez les protéines Ets-1 et Ets-2, correspond

**A**



**B**



**Figure 6** : Le domaine PNT. **A-** Récapitulatif des différentes dénominations utilisées dans la bibliographie. La protéine Ets-1 humaine est prise comme référence, sa taille en acides aminés est indiquée dessous. **B-** Alignement des séquences du domaine PNT des protéines Ets et d'un domaine apparenté des protéines appartenant au groupe *Polycomb* ; chacun de ces domaines est représenté avec leur partie amino-terminale. Les hélices  $\alpha$  déterminées par RMN sont encadrées (Graves and Petersen, 1998), celles prédites sont soulignées (Alkema et al., 1996).

à une séquence inhibitrice de liaison à l'ADN. Dans le cas de *Gabp $\alpha$* , l'interaction avec *Gabp $\beta$*  oriente de manière très différente cette extrémité et l'éloigne de l'ADN. Ainsi, à la différence de ce qui est observé pour la protéine *Ets-1*, elle n'induirait pas une inhibition de liaison à l'ADN. Dans le cas de *Ets-1*, un cofacteur potentiel pourrait jouer le même rôle que la protéine *Gabp $\beta$*  et repousser l'hélice carboxy-terminale loin de l'ADN.

#### IV- Le domaine PNT

La présence d'un second domaine conservé est constatée chez plusieurs membres de la famille *Ets*. Ce domaine de 65 acides aminés est présent dans les protéines issues des gènes *ets-1*, *ets-2*, *gabp $\alpha$* , *erg*, *fli-1*, *ese* et *tel* chez les vertébrés et *pnt*, *elg* et *yan* chez la drosophile. Son nom et sa fonction restent encore controversés.

##### 1- Les tracas de sa dénomination

A l'origine, la comparaison des séquences protéiques des premières protéines *Ets* caractérisées a mis en évidence deux domaines A et C (Fig 6A), le domaine C se révélant être le domaine ETS (Papas et al., 1989). Plus tard, cette région A fut appelée RII pour région II quand les domaines fonctionnels de la protéine *Ets-1* furent cartographiés (Schneikert et al., 1992). Enfin, elle porta le nom de domaine POINTED lorsque la description du nouveau membre de la famille *pointed* (*pnt*) chez la drosophile fournit une homologie plus étendue pour ce domaine (Klämbt, 1993). Par la suite, les prédictions de structure secondaire proposèrent une ressemblance structurale et fonctionnelle avec le domaine de dimérisation hélice-boucle-hélice des facteurs de transcription bHLH, d'où la dénomination de domaine HLH (Seth et Papas, 1990). En fait, des études plus récentes de structure secondaire par RMN du domaine PNT de la protéine *Ets-1* suggèrent la présence de quatre hélices  $\alpha$  ( $\alpha 2$ -  $\alpha 5$ ), sans identité de structure avec une domaine HLH. Une cinquième hélice  $\alpha$  ( $\alpha 1$ ) est prédite dans la partie amino-terminale du domaine PNT de la protéine *Ets-1* mais pas forcément pour les autres protéines *Ets* vu la faible homologie dans cette région (Graves and Petersen, 1998 ; Barbara Graves, communication personnelle).

La figure 6B montre l'alignement des domaines PNT des différentes protéines *Ets*. D'une part, nous constatons que le pourcentage global d'identité est très faible, de l'ordre de 22%. D'autre part les résidus communs sont situés aussi bien dans les hélices  $\alpha$  que dans les régions intermédiaires ce qui laisse supposer une fonction conservée aussi bien pour les hélices que pour les boucles.

Un alignement récent (Alkema et al., 1996) montre une homologie de ce domaine PNT avec des séquences conservées de protéines n'appartenant pas à la famille *Ets* (Fig

6B). En effet, nous pouvons noter la présence de domaines semblables dans des protéines du groupe des gènes *Polycomb* (PcG) tels que ph ("polyhomeotic") et Scm ("Sex comb on midleg") de la drosophile, Mph1 ("Murine polyhomeotic") chez la souris, et HPH1 et HPH2 ("Human polyomeotic") chez l'homme, mais aussi dans quatre protéines cytoplasmiques impliquées dans la différenciation sexuelle chez la levure (Bornemann et al., 1996 ; Alkema et al., 1996 ; Gunster et al., 1997 ; Peterson et al., 1997). Dans chacun de ces cas, un nom a été donné à ce domaine, SPM pour le module commun aux protéines Scm et ph, SEP pour "yeast Sterility-Ets related-PcG proteins" ou encore SAM pour "Sterile Alpha Module". Quand nous considérons l'alignement du domaine PNT des protéines Ets et du groupe *Polycomb* (Fig 6B), nous constatons que l'homologie ne situe pas seulement au niveau des acides aminés mais surtout au niveau des structures secondaires prédites (Alkema et al., 1996).

Compte tenu de ces nouvelles observations, nous aurions pu appeler cette région, domaine SEP ("yeast Sterility-Ets related-PcG proteins"), car ce nom révèle bien l'homologie entre les protéines Ets, les protéines PcG et celles impliquées dans la différenciation sexuelle chez la levure. Cependant, vu la faible homologie de séquence entre toutes ces protéines et en concertation avec Barbara Graves qui a étudié la structure de ce domaine, nous avons adopté dans ce mémoire la terminologie de domaine POINTED et l'abréviation PNT.

## 2- Les fonctions potentielles du domaine PNT

Des analyses fonctionnelles indiquent que le domaine PNT des protéines Ets-1, Ets-2 et Fli-1 modulent leurs propriétés transcriptionnelles (Gégonne et al., 1992 ; Schneikert et al., 1992 ; Rao et al., 1993). Plus récemment, de nouvelles études ont démontré que ce domaine pourrait permettre des interactions protéine-protéine (Jousset et al., 1997). En effet, la protéine Tel est capable de s'oligomériser *via* ce domaine. Cependant, cette propriété semble être spécifique de la protéine Tel puisque le domaine PNT des protéines Ets-1, Fli-1 et Gabp $\alpha$ , réintroduit dans la protéine Tel, ne permet plus son oligomérisation (Jousset et al., 1997). Par ailleurs, les membres Ets-1 et Fli-1 interagissent avec la protéine huUBC9 ("Ubiquitin Conjugating enzyme") *via* leur domaines ETS et PNT (Hahn et al., 1997). L'association Ets-1/huUBC9 stimule *in vitro* l'activité transcriptionnelle de la protéine Ets-1. Ce domaine pourrait alors intervenir dans les interactions des protéines Ets avec d'autres protéines tels que des corépresseurs, des coactivateurs ou encore des composants du complexe d'initiation de la transcription.



Ce domaine, chez les protéines *ph* et *Scm* de *Drosophila*, intervient dans la fonction des protéines du groupe des *Polycomb* (PcG). Les différents membres de ce groupe sont des régulateurs de l'expression des gènes homéotiques responsables de l'axe antéro-postérieur chez l'embryon. En effet, Borneman et al. (1996) ont démontré que les protéines PcG se fixent en complexes multimériques au niveau des régions régulatrices des gènes à réprimer. Diverses expériences, aussi bien chez la *Drosophila* que chez les mammifères, ont démontré que ces protéines PcG sont colocalisées et font partie d'un même complexe multimérique. Des études plus poussées de ces gros complexes sont en cours afin de déterminer si les mécanismes moléculaires de la répression des gènes homéotiques sont similaires dans des organismes aussi divers que la *Drosophila*, la souris ou l'homme (Borneman et al., 1996). De plus, les interactions protéine-protéine à l'intérieur de ces complexes multimériques sont dues, entre autres, à ce domaine homologue au domaine PNT des protéines Ets (Peterson et al., 1997). Pourrions-nous alors supposer, à partir de ces résultats, que les protéines Ets, possédant ce domaine PNT, forment des complexes multimériques *via* ce même domaine?

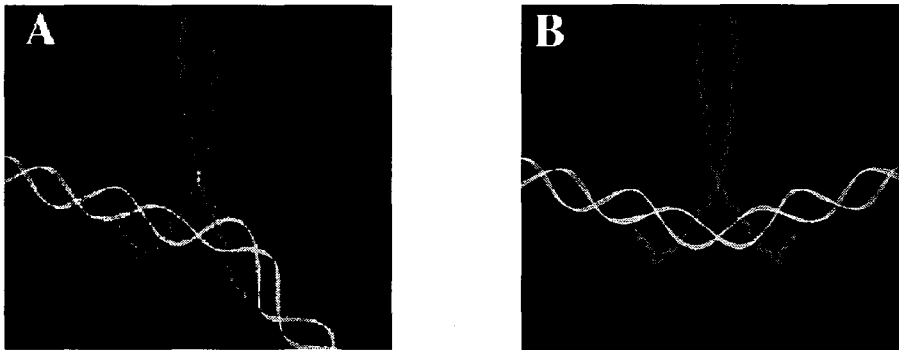
Cette possibilité d'interaction *via* le domaine PNT, associée à la spécificité de fixation à l'ADN, pourrait alors déterminer des mécanismes moléculaires d'activation de la transcription des protéines Ets en coopération avec des cofacteurs. Nous allons voir que différents membres de la famille Ets coopèrent et/ou interagissent avec de nombreuses protéines dont les facteurs de transcription Fos et Jun. De plus, dans la discussion, nous détaillerons l'interaction des protéines Erg avec le complexe AP-1. C'est pourquoi, nous présentons rapidement dans le chapitre suivant les protéines Fos et Jun, avant de détailler les coopérations et les interactions observées avec différents membres de la famille Ets.

## **Deuxième partie : Les protéines Fos et Jun, partenaires privilégiés des protéines Ets**

La famille des facteurs de transcription qui constituent le complexe AP-1 ("Activating Protein") est constituée de protéines bZIP ("basic region Zipper Interaction Protein"), c'est-à-dire des protéines qui possèdent une région basique pour la liaison à l'ADN et une "glissière à leucine" pour la dimérisation. La famille de gènes correspondants est formée d'au moins quatre sous-familles, Fos, Jun, ATF et Maf. Chez les mammifères, la sous-famille Jun est constituée de trois membres homologues au produit de l'oncogène *v-jun* (abréviation de "ju-nana", dix-sept en japonais car *v-jun* a été isolé à partir du rétrovirus E17), *c-jun*, *junB* et *junD*, quant à la sous-famille Fos, elle comprend quatre membres homologues au produit de l'oncogène *v-fos* ("Finkel Biskis Jinkins - mouse sarcoma virus Oncogene Specific"), *c-fos*, *fosB*, *fra1* ("Fos Related Antigen") et *fra2*. Il existe aussi la sous-famille ATF ("Activating Transcription Factor") avec les membres *ATF2*, *ATF3/LRF1* et *B-ATF* ainsi que la sous-famille Maf ("Musculo Aponeurotic Fibrosarcoma") composée de *v-Maf*, *c-Maf*, *MafB*, *MafF*, *MafG* et *MafK* (Karin et al., 1997).

### **I- Il était une fois...**

Une fraction nucléaire issue de cellules HeLa, capable d'activer la transcription du promoteur humain de la métallothionéine, a été dénommée AP-1 ("Activator Protein-1"). Cette fraction protéique se fixe sur une séquence cible, 5'- TGACTCA -3', appelée site AP-1, commune au promoteur de la métallothionéine et à la séquence activatrice du virus simien, SV40. Son analyse a révélé qu'elle était composée de protéines au poids moléculaire apparent compris entre 41 et 50 kDa (Lee et al., 1987a et 1987b). Des anticorps spécifiques de la protéine virale v-Jun reconnaissent des protéines AP-1 or la protéine c-Jun est capable de se lier à un site AP-1. Bohman et al. (1987) avaient donc conclu que le gène *c-jun* codait le facteur AP-1. Cependant, la protéine Fos est elle aussi capable de se lier à des sites AP-1 ce qui supposait son interaction avec des protéines connues pour lui être associées telle p39 ou avec une protéine AP-1 (Franza et al., 1988). En fait, comme nous le savons aujourd'hui, les protéines c-Fos et c-Jun forment un complexe qui se lie à des sites



**Figure 7** : La fixation des protéines Fos et Jun induisent la courbure de l'ADN. Les hétérodimères Fos/Jun (A) et les homodimères Jun/Jun (B) courbent la double hélice d'ADN dans des orientations inverse. D'après Kerppola et al., 1991b.

AP-1, et donc p39 isolée à l'origine avec la protéine c-Fos n'est autre que le produit du proto-oncogène *c-jun* (Rauscher et al., 1988).

Par la suite, les auteurs ont constaté que le complexe AP-1 rassemble une collection de dimères composés des protéines Jun, Fos, ATF et Maf qui interagissent entre elles *via* leur glissière à leucine (Busch and Sassone-Corsi, 1990 ; Karin et al., 1997). Les protéines Jun se montrent les plus versatiles. Elles sont capables de former des dimères Jun/Jun, Jun/Fos et Jun/ATF. Les protéines Fos ne s'homodimérisent pas et de façon remarquable, l'hétérodimère Fos/Jun est beaucoup plus stable que l'homodimère Jun. Les protéines v-Maf et c-Maf forment des hétérodimères avec Jun et Fos tandis que MafB, MafF, MafG et MafK ne peuvent s'associer qu'avec Fos (Busch and Sassone-Corsi, 1990 ; Karin et al., 1997). Enfin, des facteurs inhibiteurs peuvent faire partie d'un complexe AP-1 comme les produits des gènes *JDP1 et 2* ("Jun Dimerization Protein") qui forment des hétérodimères avec la protéine c-Jun (Aronheim et al., 1997).

Tous ces dimères se lient à l'ADN *via* leur domaine basique. Concentrons-nous sur les protéines Fos et Jun dont les fonctions sont multiples.

## **II- Les protéines Fos et Jun courbent l'ADN**

La régulation de la transcription implique des contacts entre différents facteurs de transcription dont des composés de la machinerie transcriptionnelle basale. Des interactions entre protéines liées à l'ADN sur des sites séparés requièrent une courbure de l'hélice d'ADN. Bien que des longs fragments puissent se comporter comme des chaînes flexibles, des petits fragments ont la caractéristique d'un bâton rigide. Dans ce dernier cas, les interactions des protéines liées à des sites séparés mais assez proches sont contraintes par des énergies défavorables de courbure de l'ADN. Les facteurs qui induisent de telles courbures peuvent réduire la barrière thermodynamique et faciliter ainsi les interactions protéine-protéine. Les protéines Fos et Jun font partie de ces facteurs (Kerppola and Curran, 1991a et 1991b). En effet, les dimères Jun/Jun et Fos/Jun, grâce à leurs domaines basiques, se lient à l'ADN *via* une séquence cible consensus AP-1. Ils courbent alors l'ADN mais dans des orientations opposées (Fig 7).

En fait, la courbure résulte des domaines transactivateurs et non de liaison à l'ADN. Sur les protéines Fos et Jun, ces domaines transactivateurs sont situés à l'opposé par rapport aux domaines de liaison à l'ADN, ce qui fait que les homodimères et hétérodimères courbent l'ADN dans des orientations inverses (Kerppola and Curran, 1991a et 1991b). Cependant, cette théorie est très controversée. Tout d'abord, d'autres auteurs n'ont pu mettre en évidence cette courbure et affirment que les complexes AP-1 ne courbent pas l'ADN de manière significative (Sitlani et al., 1996 et 1998). Enfin, une équipe a

constaté cette courbure mais dans la même orientation quelque soit le complexe protéique étudié (Glover et al., 1995). En fait, dans cette dernière étude faite par cristallographie, les homo- et hétérodimères sont figés ce qui pourrait gêner l'analyse des courbures.

### III- Les protéines Fos et Jun dans des complexes protéiques

Les protéines Jun sont le point focal de multiples interactions protéiques. La plupart de ces interactions font intervenir la "glissière à leucine" et impliquent l'hétérodimérisation avec d'autres protéines bZIP. Cependant, de nombreuses protéines très différentes s'associent avec le complexe AP-1 *via* Jun.

Dans le cytomégalovirus humain, la protéine IE2 86, issue du gène viral *IE* ("Immediate Early gene"), va activer la transcription d'un gène viral. Cette activation nécessite une liaison directe avec le promoteur *via* une séquence cible mais aussi des interactions protéiques (Scully et al., 1995). En effet, la protéine IE2 86 est capable de s'associer avec les protéines Jun mais pas avec les protéines Fos à moins qu'elles ne soient en hétérodimères. De la même manière, la protéine E7 produite par le papillomavirus humain de type 16 interagit *in vivo* avec la protéine c-Jun et augmente l'activité transactivatrice du complexe AP-1 (Antinore et al., 1996).

D'autre part, une équipe a isolé un extrait nucléaire de cellules T qui stimule fortement la liaison à l'ADN du complexe AP-1 (Powers et al., 1996). La purification à partir de cet extrait nucléaire de la protéine AF-1 ("AP-1 associated Factor") démontre qu'elle forme un complexe ternaire avec les dimères AP-1. Ce complexe nécessite l'interaction physique entre AF-1 et Jun ce qui explique une association différentielle entre AF-1 et les homo- ou hétérodimères.

Hormis les facteurs de transcription, des protéines coactivatrices sont capables d'interagir avec le complexe AP-1. Ainsi, le coactivateur  $\alpha$ -NAC ("Nascent polypeptide-Associated Complex"), spécifiquement exprimé dans les ostéoblastes, potentialise l'activité transcriptionnelle des homodimères Jun tandis qu'il n'affecte en aucun cas celle des hétérodimères Jun/Fos (Moreau et al., 1998). En effet,  $\alpha$ -NAC stabilise la fixation de l'homodimère sur l'ADN et avec le complexe d'initiation de la transcription en interagissant directement avec Jun et le complexe basal.

Cette liste des protéines qui interagissent avec le complexe AP-1 *via* le facteur de transcription Jun est loin d'être exhaustive. De nouvelles études mettent en évidence des associations de facteurs avec le dimère Fos/Jun *via* Fos. C'est le cas d'une des isoformes Erg comme nous le discuterons dans les résultats.

#### IV- Les études de fonction

Les gènes qui codent les protéines du complexe AP-1 sont des gènes de réponse précoce dont la caractéristique principale est leur induction rapide après stimulation de la prolifération cellulaire. Les différentes mutations ciblées des gènes *fos* et *jun*, réalisées chez la souris, apportent des informations supplémentaires sur leurs fonctions.

##### 1- Les gènes *c-fos* et *fosB*

Les souris mutées pour le gène *c-fos* sont viables mais présentent des retards de croissance (Johnson et al., 1992). Elles présentent des problèmes de comportement mais surtout une ostéopétrose associée à des déficiences dans le remodelage de l'os. Elles présentent aussi des anomalies dans l'éruption des dents. La protéine c-Fos ne semble pas requise pour la prolifération de la plupart des cellules *in vivo*, contrairement à ce que laissaient supposer les résultats obtenus *in vitro* (Johnson et al., 1992). Les auteurs suggèrent donc que les autres protéines Fos compensent la plupart des fonctions du gène *c-fos*. De plus, par injection d'anticorps spécifiques, il a été démontré que les protéines Fos ne sont pas essentielles individuellement mais que la neutralisation des quatre, c'est-à-dire c-Fos, FosB, Fra1 et Fra2, bloque la prolifération cellulaire (Karin et al., 1997).

En ce qui concerne les souris recombinées *fosB*<sup>-/-</sup>, elles ne présentent qu'une absence de comportement maternel (Brown et al., 1996). Comme dans l'étude de KO précédente, les auteurs proposent une redondance des protéines Fos.

En revanche, le gène *c-fos* surexprimé dans les tissus osseux provoque des ostéosarcomes (Wang et al., 1995). La surexpression simultanée de ce même gène avec le gène *c-jun* augmente le pourcentage d'apparition d'ostéosarcomes (Van Dam et al., 1993). Cette étude démontre donc une coopération *in vivo* entre les protéines c-Jun et c-Fos dans le développement de ces tumeurs.

##### 2- Le gène *c-jun*

A la différence du gène *c-fos*, le gène *c-jun* est nécessaire pour le développement normal de la souris. En effet, les souris mutées *c-jun*<sup>-/-</sup> meurent après 12 à 14 jours de développement embryonnaire c'est-à-dire au milieu de la gestation. Les embryons mutants montrent une érythropoïèse du foie altérée, une mauvaise hépatogenèse et des oedèmes généralisés (Hilberg et al., 1993). De plus, les cellules embryonnaires souches *c-jun*<sup>-/-</sup> restent capables de se différencier en cellules somatiques ou germinales fonctionnelles à l'exception des hépatocytes (Hilberg et al., 1993).

Par ailleurs, les fibroblastes primaires des embryons *c-jun*<sup>-/-</sup> ne peuvent pas proliférer (Johnson et al., 1993). En fait, la protéine c-Jun est le premier facteur de

transcription décrit comme absolument nécessaire pour la prolifération des fibroblastes primaires.

Dans ce chapitre sur les protéines Fos et Jun, nous avons pu voir que ce sont des facteurs de transcription impliqués dans de nombreux types de régulation. Ils interagissent fonctionnellement avec les protéines Ets et pourraient être un de leurs modulateurs. De plus, Fos et Jun courbent l'ADN, ce qui pourrait être nécessaire à l'activité transactivatrice de nombreuses protéines dont les protéines Ets. Ainsi, nous allons constater que le complexe AP-1 et les protéines Ets ne sont pas indifférents l'un vis-à-vis des autres.

## Troisième partie : La fonction des protéines Ets

### I- Les facteurs de transcription Ets

La fonction biologique des protéines Ets est restée inconnue jusqu'en 1990. Les protéines Ets sont en fait des facteurs de transcription puisque les protéines Ets-1 et Ets-2 sont capables d'activer la transcription une fois fixées, *via* leur domaine ETS, sur leur séquence cible GGAA/T (Bosselut et al., 1990 ; Gunther et al., 1990 ; Wasylyk et al., 1990). Par la suite, on a constaté que les protéines Ets agissent comme des activateurs ou des répresseurs de la transcription. Cependant, à la différence de beaucoup de facteurs de transcription, elles sont caractérisées par de faibles capacités transactivatrices (Crépieux et al., 1994 ; Graves and Petersen, 1998).

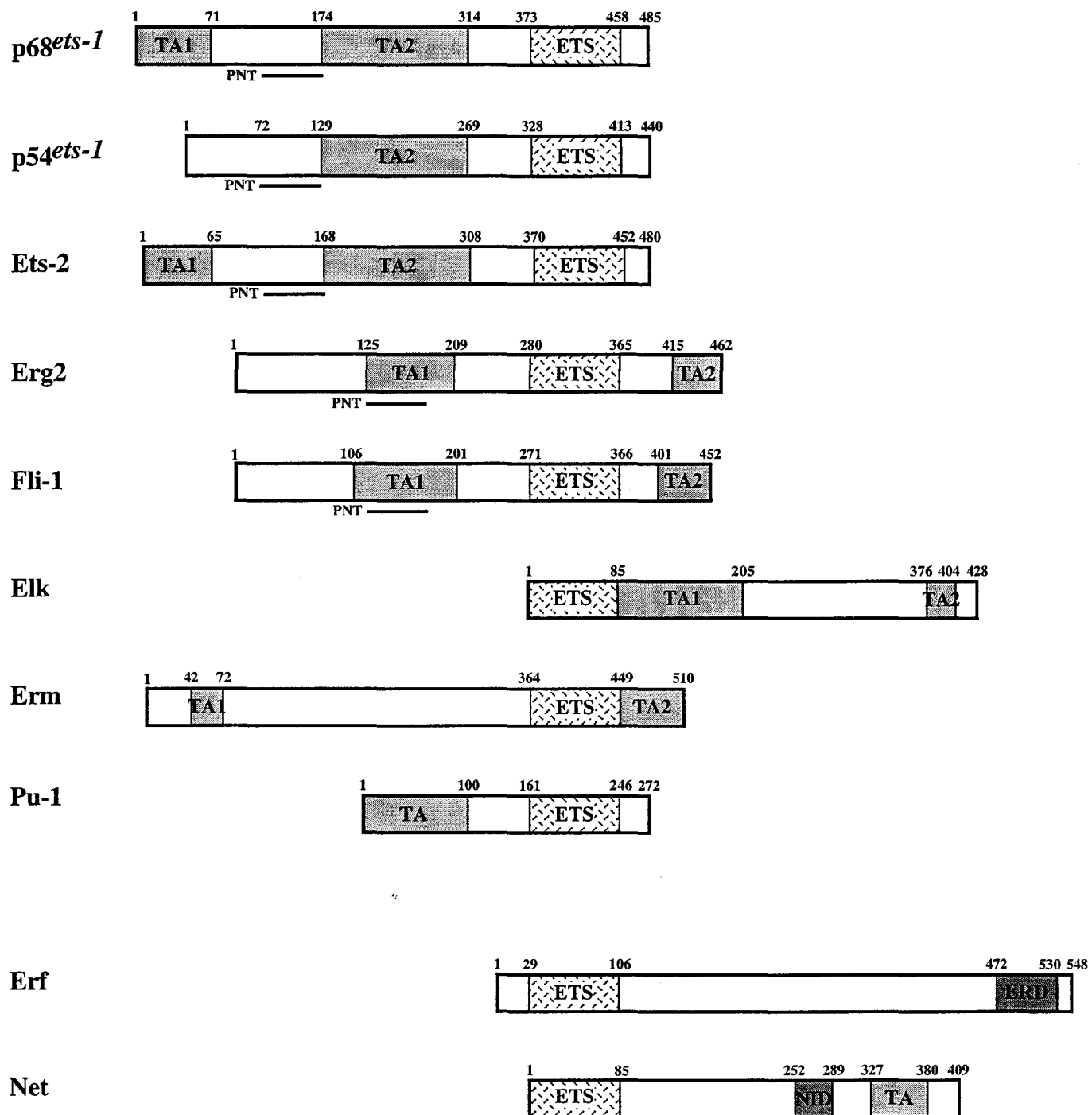
#### 1- Les domaines de régulation de la transcription

Des domaines transactivateurs et inhibiteurs ont été cartographiés chez différentes protéines Ets (Fig 8). Ils ne sont pas conservés, ni dans leur séquence, ni dans leur localisation, à l'intérieur de la famille, ce qui suggère des fonctions distinctes pour les membres. De plus, ils n'ont pas toujours une caractéristique définie. En effet, un domaine transactivateur est souvent caractérisé par l'abondance d'acide aminé particulier, il y a les domaines riches en proline, en proline et sérine, en glutamine ou encore en résidus acides (Graves and Petersen, 1998). De la même manière, un domaine inhibiteur peut être défini par la présence de nombreux résidus alanine, glutamine et/ou proline (Cowell, 1994).

#### a- Les domaines transactivateurs

Une des protéines Ets-1 aviaire, p68, et la protéine Ets-2, bien que différentes entre elles, possèdent deux domaines transactivateurs TA1 et TA2, peu conservés entre les deux membres (Fig 8). Ils sont apparemment indépendants, puisque Schneikert et al. (1992) n'ont pu déterminer une synergie entre eux. Nous pouvons remarquer qu'une des deux isoformes issue du gène *ets-1*, p54 seule présente chez les mammifères, ne possède pas le domaine amino-terminal TA1 (Schneikert et al., 1992 ; Crépieux et al., 1994). De plus, à l'intérieur de la sous-famille ETS, nous constatons la présence d'un domaine transactivateur dans la partie amino-terminale des protéines, TA2 pour Ets-1 et TA1 pour Erg et Fli-1, de nature acide (Reddy et al., 1987), qui ne correspond pas strictement au domaine PNT conservé. Par contre, les protéines Erg et Fli-1 présentent la même région activatrice, riche





**Figure 8** : Organisation générale des domaines fonctionnels chez certaines protéines Ets. ETS = domaine de liaison à l'ADN, TA = domaine de transactivation, ERD = "Erf Repression Domain" et NID = "Nerf Inhibitory Domain". En dehors des protéines Erg et Fli-1, les domaines TA ne sont pas homologues entre deux protéines distinctes. L'emplacement du domaine PNT est indiqué par une barre pour les protéines Ets-1, Erg et Fli-1. D'après Crépieux et al., 1994.

en résidus proline et sérine, à l'extrémité carboxy-terminale (Fig 8) (Siddique et al., 1993 ; Rao et al., 1993).

Dans le groupe PEA3, les membres possèdent en commun un domaine transactivateur acide, TA1. La protéine Erm présente, en plus, dans sa partie carboxy-terminale un domaine à la fois activateur de la transcription et inhibiteur de la liaison à l'ADN, TA2. A l'intérieur de cette protéine, les deux domaines transactivateurs montrent une synergie fonctionnelle ce qui fait penser à une activation de la transcription *via* deux mécanismes différents (Laget et al., 1996).

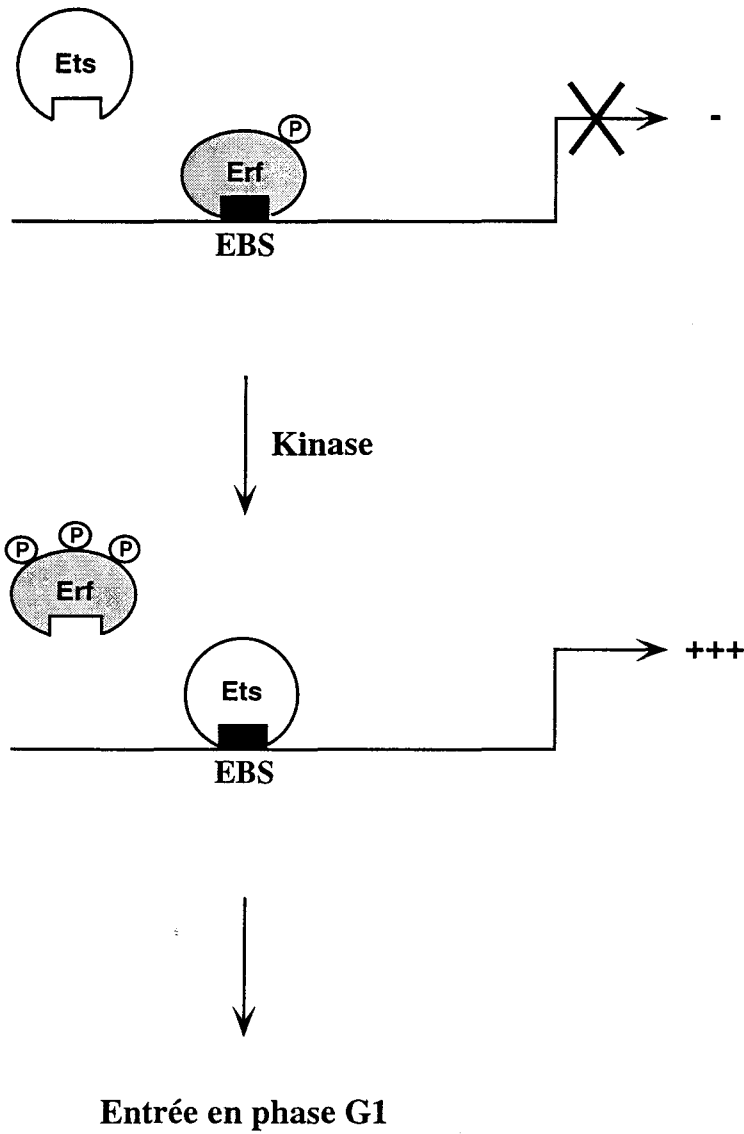
Les protéines Elk et Sap-1 ont aussi deux domaines transactivateurs autonomes situés dans les parties amino- et carboxy-terminales. Ces deux domaines sont tous les deux riches en résidus proline (Bhattacharaya et al., 1993). La partie transactivatrice juste en aval du domaine ETS, TA1, contient la boîte B permettant l'interaction avec le facteur de réponse au sérum, SRF (Dalton and Treisman, 1992).

Enfin, dans la protéine Pu-1, un seul domaine transactivateur a été mis en évidence dans la première moitié amino-terminale de la protéine. Sa particularité est qu'il se divise en deux parties, et la délétion d'une de ses deux régions diminue considérablement la capacité transactivatrice de la protéine. Par ailleurs, ces deux parties ont des caractéristiques connues de domaines transactivateurs. La première et la plus grande est de nature acide tandis que la seconde est riche en résidus glutamine (Klemsz et al., 1996).

#### b- Les domaines inhibiteurs

Trois protéines de la famille Ets sont des répresseurs de la transactivation : il s'agit des protéines Yan chez la drosophile (Rebay and Rubin., 1995), Erf chez l'homme (Sgouras et al., 1995) et Net chez la souris (Maira et al., 1996). Pour la protéine Yan, aucun domaine inhibiteur n'a été mis en évidence. Il est à noter que la protéine humaine Fev ("Fifth Ewing Variant") contient dans sa partie carboxy-terminale un domaine dont les caractéristiques ont été observées chez divers inhibiteurs de la transcription (Peter et al., 1997). Cependant, des études sont en cours pour démontrer ces potentielles propriétés inhibitrices.

La protéine Erf ("Ets-2 Repressor Factor") a été mise en évidence grâce à sa capacité à lier le promoteur du gène *ets-2*. Elle possède dans sa partie carboxy-terminale un domaine de 70 acides aminés qui, associé au domaine ETS, est capable d'inhiber la transcription de gènes possédant dans leur promoteur des sites EBS (Fig 8). Ce domaine, appelé ERD ("Erf



**Figure 9 :** Modèle simplifié de la fonction de la protéine Erf.  
D'après Sgouras et al., 1995.

Repressor Domain"), ne présente aucune des caractéristiques particulières d'un domaine inhibiteur (Sgouras et al., 1995). Cette protéine, dont l'expression est constante au cours du temps, voit sa fonction dépendre de sa phosphorylation. En effet, la protéine hyperphosphorylée n'inhiberait plus l'expression de ses gènes cibles et permettrait à la cellule de rentrer en phase G1 (Fig 9).

La protéine Net, homologue de Elk-1 et Sap-1, est elle aussi capable de former un complexe ternaire avec le dimère SRF ("Serum Responsive Factor") sur l'élément de réponse au sérum, SRE. Nous détaillerons cet aspect un peu plus loin. En fait, la protéine Net, dans des expériences de transactivation *in vitro* et selon le type cellulaire utilisé, réprime la transcription d'un promoteur artificiel composé de plusieurs sites EBS tandis que Sap-1 n'a aucun effet et que Elk-1 active. Le domaine inhibiteur, NID ("Net Inhibitory Domain"), de 50 acides aminés, est situé en amont du domaine transactivateur commun aux membres du groupe ELK (Maira et al., 1996). Après comparaison de séquences, Maira et al. (1996) suggèrent que ce domaine NID adopte une structure HLH.

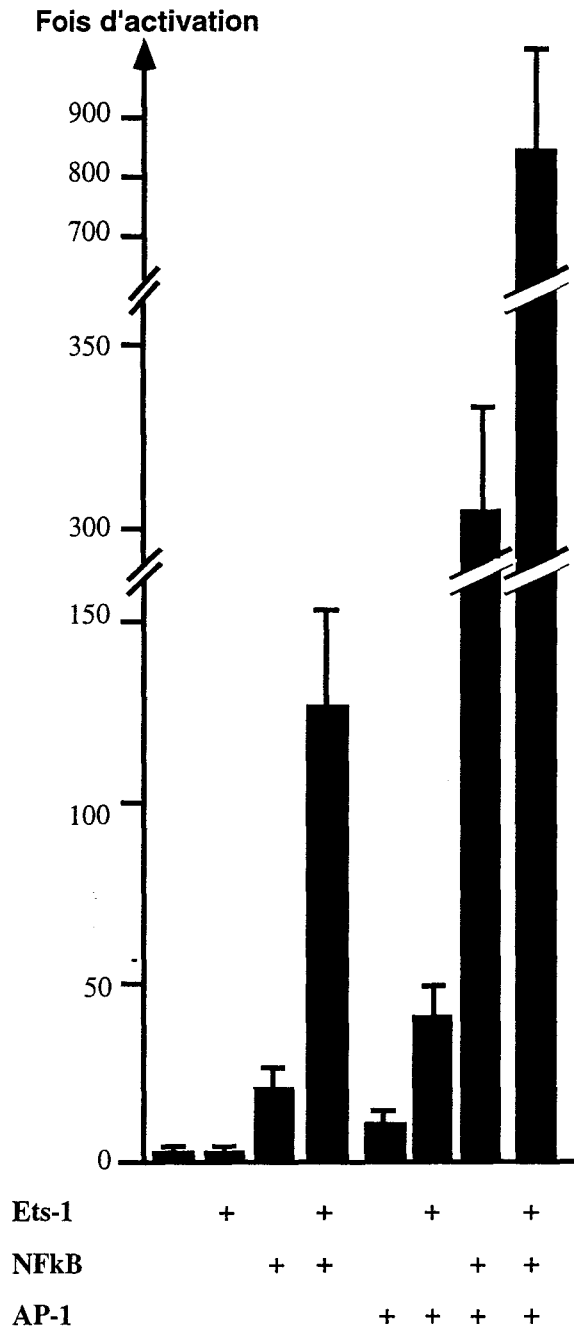
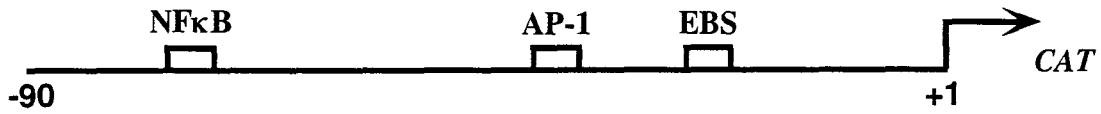
## 2- Les protéines Ets coopèrent avec divers partenaires

Les membres de la famille Ets sont des facteurs de transcription avec de faibles capacités transactivatrices. Ces capacités peuvent augmenter lorsqu'ils agissent en coopération avec d'autres facteurs de transcription afin de réguler de nombreux promoteurs et "enhancers" cellulaires et viraux. Cependant, les mécanismes impliqués ne sont toujours pas identifiés.

### a- La protéine Ets-1 sur le promoteur du gène *GM-CSF*

Lorsque les cellules T effectrices sont stimulées, elles initient une réponse immunitaire. Leur activation est à l'origine d'une cascade d'événements intracellulaires qui aboutit à l'induction de la transcription et à la sécrétion d'un certain nombre de cytokines. La protéine GM-CSF ("Granulocyte-Macrophage Colony-Stimulating-Factor") est une de ces cytokines impliquées dans la différenciation, la prolifération ou l'activation du système hématopoïétique. L'expression du gène *GM-CSF* dans les cellules T activées est régulée par au moins deux mécanismes. La transcription de ce gène est contrôlée par des facteurs constitutifs ou inductibles qui interagissent soit avec le promoteur proximal soit avec un "enhancer" commun aux gènes *GM-CSF* et *IL3* (Thomas et al., 1995).

Dans le promoteur du gène *GM-CSF*, des éléments de liaison aux protéines Ets, AP-1 et NFκB ont été identifiés et démontrés comme importants pour l'activation de la transcription. La protéine Ets-1, exprimée dans les cellules T, est capable d'activer la



**Figure 10 :** Etude de transactivation des protéines Ets-1, NFκB et AP-1 sur une partie du promoteur du gène *GM-CSF*. D'après Thomas et al., 1997.

transcription du promoteur du gène *GM-CSF* dans des cellules Jurkat stimulées par le phorbol-myristate-acétate et l'ionomycine (Thomas et al., 1995 et 1997). Cependant, dans des cellules Jurkat non stimulées, elle ne transactive fortement qu'en présence des protéines AP-1 et NF $\kappa$ B exogènes. Il existe alors une synergie entre ces trois facteurs (Fig 10). De plus, si les sites de liaison pour les protéines AP-1 et NF $\kappa$ B sont mutés, la transactivation par la protéine Ets-1 dans les cellules Jurkat stimulées est fortement réduite. Cela suggère que les protéines AP-1 et NF $\kappa$ B sont les facteurs spécifiquement induits qui agissent en synergie avec la protéine Ets-1. Malgré l'interaction avérée entre la protéine Ets-1 et les protéines AP-1 (Bassuk et al., 1995) et NF $\kappa$ B (Bassuk et al., 1997), les auteurs n'ont pas montré ici une interaction physique entre ces différents facteurs (Thomas et al., 1995 et 1997).

#### b- Des protéines Ets interfèrent avec des récepteurs nucléaires

Deux équipes différentes ont démontré que des membres de la famille Ets, Pu-1 (Gauthier et al., 1993) et Fli-1 (Darby et al., 1997), inhibent l'activité transactivatrice de récepteurs nucléaires, et que ces mêmes récepteurs peuvent interférer avec l'effet transactivateur des protéines Ets.

La protéine Pu-1 est capable de réprimer, dans divers types cellulaires, l'activation transcriptionnelle induite par le récepteur aux glucocorticoïdes, GR, sur un promoteur présentant un élément de réponse aux glucocorticoïdes, GRE. Son action est la même vis-à-vis des récepteurs à l'acide rétinoïque, RAR $\alpha$ , et à l'hormone T3, T3R (Gauthier et al., 1993). Inversement, GR, en présence d'hormone, réprime l'activation transcriptionnelle induite par la protéine Pu-1 sur des éléments de réponse aux protéines Ets. Le mécanisme de ces interférences fait intervenir le domaine de liaison à l'ADN des récepteurs (Gauthier et al., 1993). En effet, un récepteur nucléaire constitutif, c'est-à-dire sans son domaine de liaison du ligand, peut toujours réprimer l'activité de Pu-1. Par contre, une mutation qui abolit la liaison à l'ADN ne permet plus au récepteur d'interférer. Comme la protéine Pu-1 est exprimée dans diverses lignées cellulaires hématopoïétiques sauf dans les cellules T, les auteurs ont suggéré qu'elle pouvait inhiber l'action des hormones dans la différenciation érythrocytaire (Gauthier et al., 1993).

De la même façon, il a été décrit que la protéine Fli-1, mais pas les protéines Ets-1, Gabp $\alpha$  et Elf-1, réprime la transactivation initiée par l'acide rétinoïque, l'hormone T3 et la dexaméthasone en présence de leur récepteur respectif. Inversement, RAR $\alpha$ , T3R et GR, répriment, en présence de leur ligand, l'activité transcriptionnelle de Fli-1 (Darby et al.,

1997). Ces interférences ne sont pas dépendantes de la liaison à l'ADN car la protéine Fli-1 et les récepteurs nucléaires se fixent sur leur propre site et n'entrent pas en compétition, de plus elles ne feraient pas intervenir des interactions directes entre ces deux facteurs. Par ailleurs, l'utilisation de mutants de délétion de la protéine Fli-1 et du récepteur à l'acide rétinoïque a démontré que l'effet observé est lié d'une part à la partie amino-terminale de la protéine Fli-1 possédant le domaine PNT, et d'autre part à trois régions du récepteur : les domaines de liaison à l'ADN et au ligand et la partie amino-terminale. Nous pouvons constater une différence avec la première étude où seul le domaine de liaison à l'ADN des récepteurs est requis. Une seconde différence vient du fait que ces interférences n'interviennent que dans des cellules hématopoïétiques, ce que les auteurs expliquent par un facteur de liaison entre les deux protéines, présents uniquement dans ce type de cellules (Darby et al., 1997).

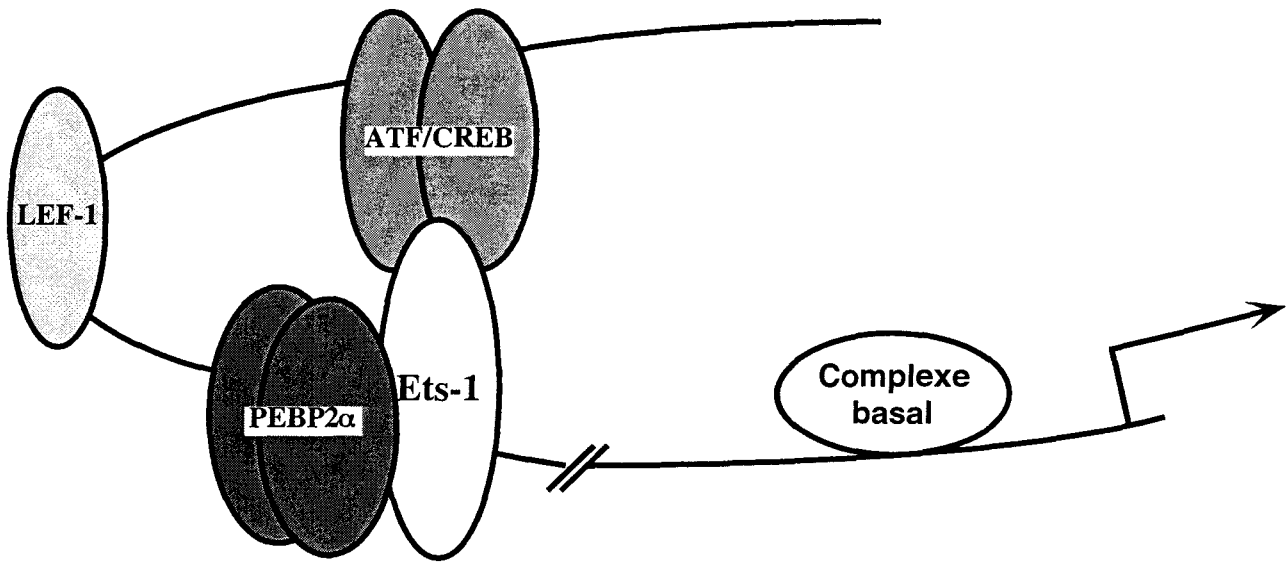
Nous venons de voir des mécanismes de coopération où les protéines impliquées n'interagissent pas directement entre elles. Nous allons voir que ce n'est pas toujours le cas.

### 3- Des interactions protéine-protéine

A travers la description des domaines ETS et PNT, nous avons déjà abordé les interactions protéine-protéine dans la famille des protéines Ets. La protéine Gabp $\alpha$  se fixe en hétérodimère avec le facteur Gabp $\beta$  sur une cible consensus des protéines Ets, et la protéine Tel est capable de s'oligomériser *via* son domaine PNT. Dans ce paragraphe, nous allons voir que différentes protéines Ets sont capables d'interagir avec divers facteurs. Nous ne ferons pas une liste exhaustive mais plutôt la présentation des cas les plus parlants.

#### a- Les protéines Elk-1 et Sap-1 sur l'élément de réponse au sérum, SRE

Un événement clé dans la réponse cellulaire à un signal de prolifération est une induction rapide et transitoire du gène *c-fos*. Cette activation se fait *via* un élément de réponse au sérum, appelé SRE, présent dans le promoteur du gène (Hipskind et al., 1991). Une protéine TCF ("Ternary Complex Factor") est capable de se lier *in vitro* au SRE en complexe avec un dimère du facteur de réponse au sérum, SRF. Cette interaction requiert d'abord la fixation de SRF sur sa séquence cible qui recrute TCF qui se lie alors à un consensus EBS GGAA/T, l'ensemble de ces deux sites définissant le SRE. L'étude des caractéristiques de liaison à l'ADN a permis d'identifier la protéine TCF au membre de la famille Ets, Elk-1. En effet, cette dernière interagit avec SRF sur SRE mais ne peut se fixer seule (Hipskind et al., 1991).



**Figure 11** : Modèle du complexe multiprotéique formé sur l'"enhancer" minimal du gène *TcR $\alpha$* . D'après Giese et al., 1995.



Par la suite, une autre équipe a utilisé, par la technique de double hybride, la protéine SRF comme appât dans le criblage d'une banque d'expression (Dalton and Treisman, 1992). Ils ont isolés un ADNc codant deux isoformes de la protéine Sap-1 ("SRF Accessory Protein-1"), Sap-1a et Sap-1b qui diffèrent dans leur partie carboxy-terminale. Ces protéines de la famille Ets sont très proches de la protéine Elk-1. Les auteurs ont déterminé un domaine protéique, une boîte B de 30 acides aminés située en aval du domaine ETS, qui permet l'interaction directe entre les membres de la famille Ets et le dimère SRF (Dalton and Treisman, 1992). Dalton and Treisman (1992) suggèrent que cette boîte B se présente sous la forme d'une hélice  $\alpha$  avec une face hydrophobe qui pourrait interagir avec SRF.

#### b- Etude de l'"enhancer" minimal du gène *TcR $\alpha$*

Dans cet exemple, Giese et al. (1995) ont montré que des interactions protéine-protéine et protéine-ADN contribuent à un assemblage coordonné et à une fonction synergique de multiples facteurs sur l'"enhancer" minimal du gène *TcR $\alpha$*  codant le récepteur des cellules T. Tout d'abord, le dimère du facteur PEBP2 $\alpha$  et la protéine Ets-1 reconnaissent des sites adjacents et se lient de manière coopérative à l'ADN. D'une part, les auteurs suggèrent que cette association instable implique une interaction protéique. D'autre part, ce complexe ternaire est stabilisé par l'intervention d'autres facteurs (Fig 11). Ainsi, la protéine Ets-1 interagit physiquement avec les facteurs ATF-2 et CREB qui se fixent en dimère sur leur site, en amont du site EBS. Cette interaction stabilise la fixation à l'ADN des protéines Ets-1 et PEBP2 $\alpha$ . Elle nécessite une courbure de l'ADN induite par la protéine LEF-1. *In vitro*, toutes ces protéines sont capables d'activer, en synergie, *via* l'"enhancer" minimal du gène *TcR $\alpha$*  (Giese et al., 1995).

La reconstitution de ce complexe a été faite dans des cellules non-lymphoïdes. Cependant, dans les cellules lymphoïdes, nous savons que plusieurs protéines Ets sont exprimées (Wang et al., 1992). Les auteurs ont montré que la protéine Ets-1, mais pas la protéine Fli-1, est capable d'entrer en synergie avec le dimère ATF/CREB sur cet "enhancer". De plus, la protéine Fli-1 n'interagit pas avec ce dimère. Tous ces résultats amènent les auteurs à penser que de multiples contacts protéine-protéine sont nécessaires pour le recrutement de la protéine Ets effectrice dans ce complexe fonctionnel (Giese et al., 1995).

#### c- Coopération des familles Ets et AP-1

Des interactions combinatoires entre des classes distinctes de facteurs de transcription jouent un rôle important dans la régulation de l'expression des gènes. De telles interactions peuvent être dictées par l'arrangement de sites de liaison sur l'ADN dans les promoteurs et

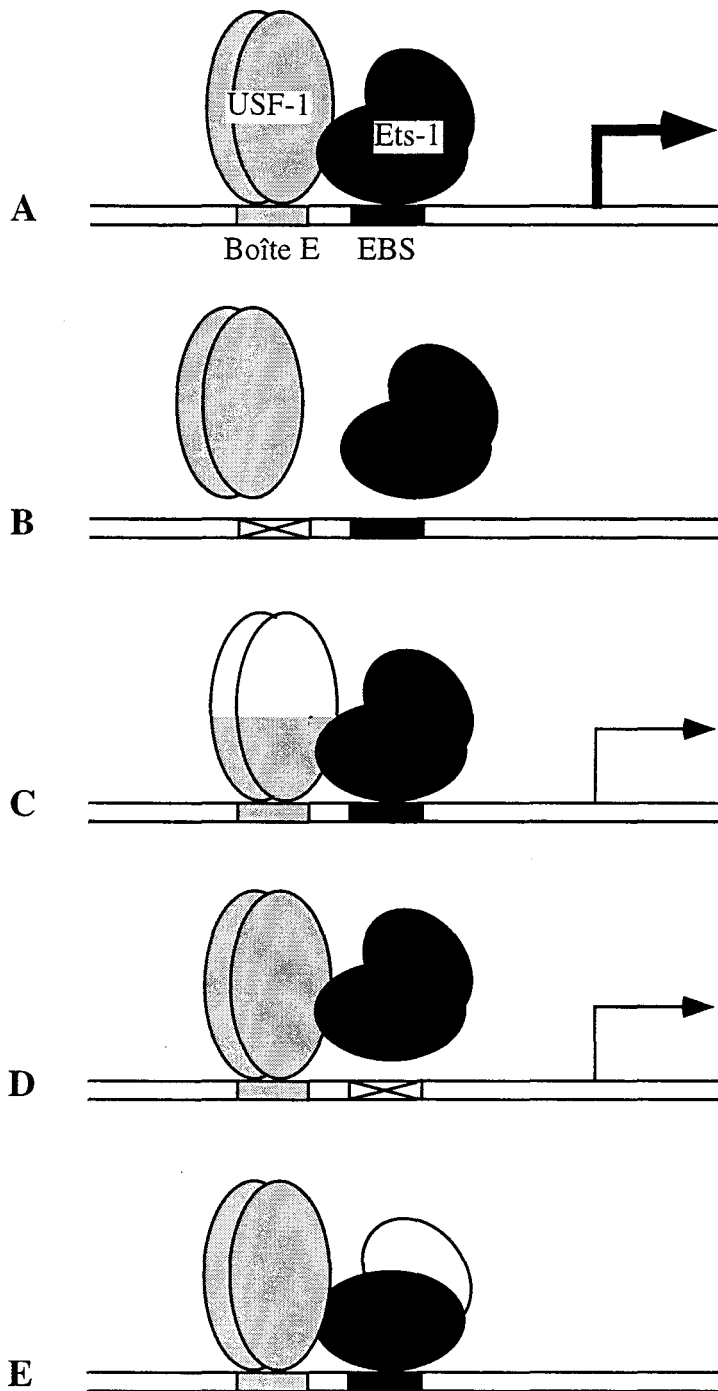
les "enhancers" eucaryotes ou par des motifs protéiques qui permettent des interactions physiques ou par les deux. Dans l'exemple suivant, nous pouvons distinguer les deux cas. Sur l'"enhancer" du virus du polyome, il existe des sites EBS et AP-1. Il a déjà été montré que la protéine Ets-1 active la transcription *via* ces éléments en coopération avec les protéines Fos/Jun (Wasylyk et al., 1990). Depuis, quatre membres de la famille Ets, Ets-1, Fli-1, Elf-1 et Pu-1, interagissent physiquement avec les protéines de la famille Jun. Ces associations existent en particulier dans les cellules T activées (Bassuk et al., 1995). Elles font intervenir le domaine ETS et le domaine basique des protéines Jun. Bien que les protéines Ets n'interagissent pas directement avec les membres de la famille Fos, les protéines Jun sont associées avec les protéines Fos grâce à la "glissière à leucine" qui permet la formation des complexes AP-1. Cette démonstration d'interaction physique directe en absence d'ADN suggère qu'elle peut servir à stabiliser la liaison du complexe ternaire sur l'ADN. Bassuk et al. (1995) confirment cette hypothèse en démontrant que des mutants de protéines Fos/Jun, liées à l'ADN mais incapables de se lier à une protéine Ets, ne peuvent transactiver. Par ailleurs, les sites AP-1 proches de sites EBS ne sont pas reconnus quand les complexes Fos/Jun ne sont pas associés à un membre de la famille Ets.

Une partie des résultats obtenus au cours de ma thèse concerne l'interaction des produits du gène *erg* et du complexe AP-1. Ces résultats seront développés dans la partie "Résultats".

#### d- Les protéines MafB et Ets-1 dans la différenciation érythrocytaire

Le phénotype d'une cellule différenciée est en général déterminé par l'ensemble de ses gènes activés. Puisque l'exécution d'un tel programme n'est jamais déclenché par un seul "gène maître", différents auteurs ont proposé qu'une action combinée de plusieurs facteurs de transcription était requise. L'effet additif des facteurs présents ainsi que des interactions protéiques pourraient modifier les activités individuelles de chacun des partenaires.

La protéine Ets-1 joue un rôle dans la différenciation érythroïde précoce (Quéva et al., 1993 ; Wernert et al., 1994 ; Calmels et al., 1995 ) mais nous ne savons pas comment son activité est régulée dans les cellules hématopoïétiques. Dans l'exemple choisi, Sieweke et al. (1996) ont mis en évidence une protéine bZIP appelée MafB, qui interagit, grâce à sa région basique avec le domaine ETS de la protéine Ets-1. Bien que la protéine MafB soit exprimée uniquement dans les cellules myélomonocytaires, sa surexpression dans des érythroblastes inhibe l'activité transactivatrice de Ets-1 sur le promoteur du gène de la transferrine et réprime la différenciation érythrocytaire. Les auteurs suggèrent donc que la protéine MafB réprime l'action de Ets-1 dans la régulation de l'expression de gènes spécifiques de la lignée myéloïde. Ils proposent que la protéine MafB puisse masquer la



**Figure 12:** Modèle du complexe Ets-1/USF-1 sur l'"enhancer" du LTR de HIV-1. A- Les protéines coopèrent pour la liaison à l'ADN et la transactivation. B- Pas de complexe si la boîte E est mutée. C- USF-1 sans son domaine transactivateur recrute mais modère la transactivation. D- Ets-1 ne se lie pas à l'ADN et modère la transactivation. E- Ets-1 sans son domaine transactivateur ne permet pas la transactivation. D'après Sieweke et al., 1998.

partie activatrice de la protéine Ets-1 et/ou l'empêcher de se fixer à l'ADN (Sieweke et al., 1996).

Nous pouvons voir que plusieurs membres appartenant à la grande famille des protéines bZIP, Fos/Jun et MafB, peuvent avoir des effets opposés sur une même protéine Ets pour réguler des gènes différents.

e- Ets-1 et USF-1 sur l'"enhancer" de HIV-1

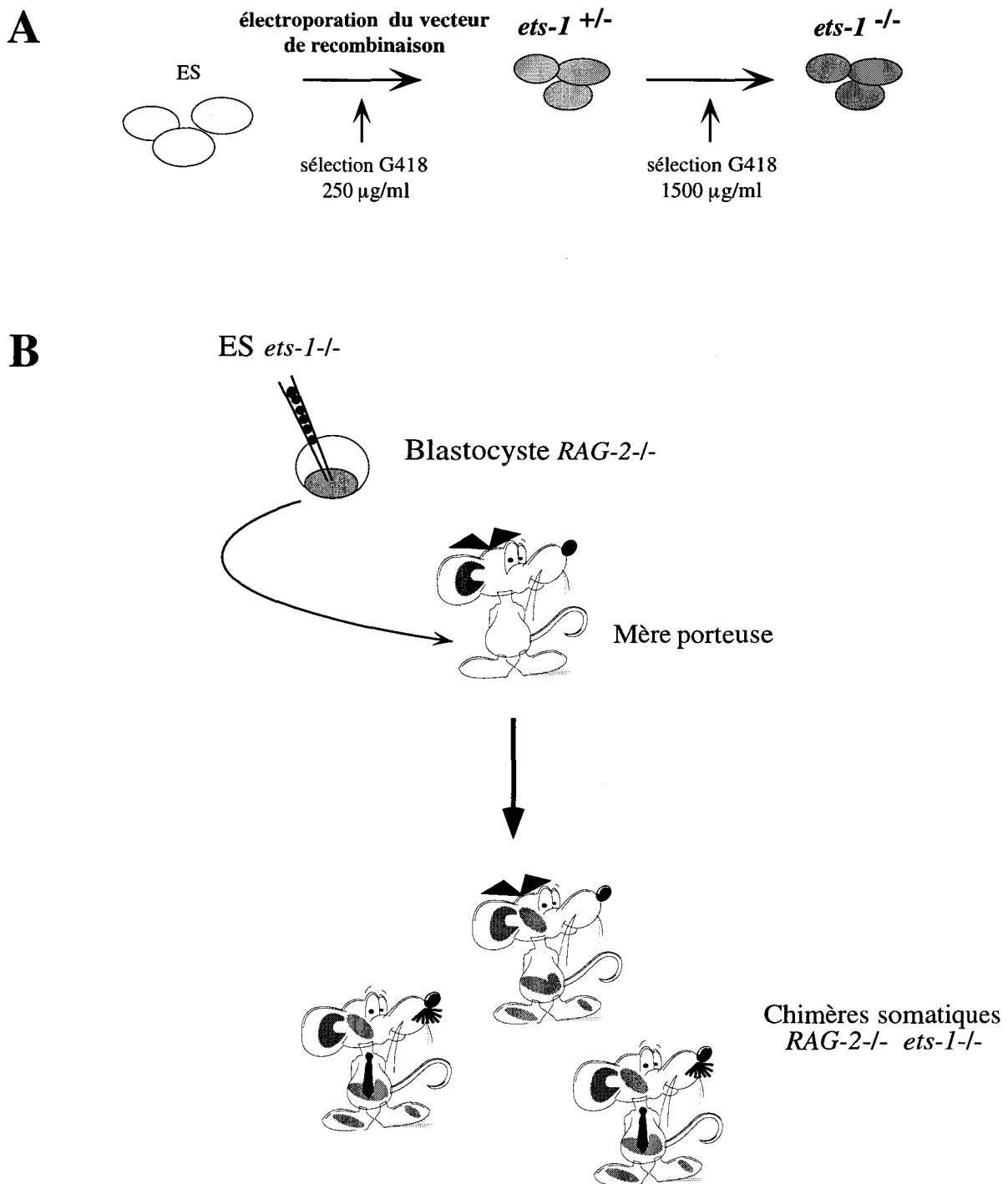
L'"enhancer" situé dans la région distale du LTR 5' ("Long Terminal Repeat") du rétrovirus HIV-1 ("Human Immunodeficiency Virus) est nécessaire pour l'activité transcriptionnelle et par conséquent pour la réplication virale. Il contient différents sites de liaison à l'ADN dont ceux des protéines Ets, EBS, et des protéines HLH, boîte E (Fig 12). Tout comme dans l'exemple précédent, Sieweke et al. (1998) ont utilisé le domaine ETS de la protéine Ets-1, lié à cet "enhancer" pour cribler une banque. Ils ont alors mis en évidence une interaction protéine-protéine fonctionnelle entre le membre de la famille Ets, Ets-1, et une protéine HLH, USF-1, *via* leur domaine de liaison à l'ADN respectif (Sieweke et al., 1998).

Les protéines Ets-1 et USF-1 ont des fonctions complémentaires dans la transactivation synergique de l'"enhancer" situé dans la région distale du LTR 5' de HIV-1. En effet, le domaine de liaison à l'ADN de USF-1 est requis pour un recrutement correct du complexe (Ets-1/USF-1) sur l'"enhancer" tandis que la protéine Ets-1 totale serait nécessaire pour établir des contacts optimaux avec la machinerie transcriptionnelle basale (Fig 12). En fait, les auteurs suggèrent que les protéines Ets-1/USF-1 interagissent avec d'autres facteurs de telle manière qu'elles interviendraient dans un complexe multiprotéique (Sieweke et al., 1998).

Les études *in vitro* que nous venons de décrire ne sont pas les seuls moyens de déterminer la fonction d'un gène. Son rôle peut être étudié *in vivo*, en particulier par les techniques puissantes de "knock-out" et de transgénèse. Ainsi, l'association des études *in vitro* et *in vivo* permettra une meilleure vue d'ensemble du rôle de chaque protéine Ets.

## **II- Les études de gènes *ets* *in vivo***

Dans ce chapitre, nous illustrons les spécificités biologiques de certaines protéines Ets en passant en revue les études récentes réalisées *in vivo* chez la souris. Dans ce but, les différentes équipes ont utilisées les techniques puissantes de KO ("knock-out") et de transgénèse. Nous constatons alors que les différents gènes étudiés pour l'instant sont impliqués dans l'hématopoïèse, l'angiogénèse et la formation du squelette.



**Figure 13** : Schéma résumant le système de complémentation des blastocystes *RAG-2*<sup>-/-</sup>.  
**A-** Les cellules ES *ets-1*<sup>-/-</sup> sont obtenues par deux étapes de recombinaison au cours desquelles la dose de G418 est augmentée lors de la sélection. **B-** Le phénotype des cellules B et T est étudié directement sur les animaux chimère.

## 1- L'hématopoïèse

Plusieurs gènes *ets* sont largement impliqués dans l'hématopoïèse, à différents stades et à différents niveaux. C'est le cas des gènes *ets-1*, *fli-1*, *pu-1* et *spi-B* dont nous décrivons l'étude ci-dessous.

### a- Le gène *ets-1*

L'expression du gène *ets-1* au cours de la maturation des cellules B et T, ainsi que l'identification de plusieurs gènes cibles potentiels parmi les gènes spécifiques de la lignée lymphocytaire (Crépieux et al., 1993) laissent penser que les protéines Ets-1 pourraient être des facteurs cruciaux pour le développement et la fonction des cellules B et T. Plusieurs équipes différentes ont entrepris de réaliser le KO du gène *ets-1* mais aucun résultat n'est encore publié. Deux équipes indépendantes ont par ailleurs utilisé des cellules ES *ets-1*<sup>-/-</sup> dans le système de complémentarité des blastocystes *RAG-2*<sup>-/-</sup> (Fig 13). Elles ont ainsi pu déterminer l'implication fonctionnelle des protéines Ets-1 dans le système immunitaire (Bories et al., 1995 ; Muthusamy et al., 1995).

La différenciation des lymphocytes B et T nécessite une série d'événements ordonnés qui inclut l'assemblage des produits de gènes codant les régions variables du récepteur des cellules T ou des immunoglobulines. Deux protéines sont responsables de cet assemblage, RAG-1 et RAG-2 ("Recombinaison Activating Gene"). Dans les souris homozygotes mutées pour l'un ou l'autre des gènes correspondants, la différenciation des cellules T est bloquée au stade double négatif, CD4<sup>-</sup> CD8<sup>-</sup>, dans le thymus, et celles des cellules B au stade pro-B dans la moëlle osseuse. Ces souris présentent alors une immunodéficiência sévère due à l'absence totale des lymphocytes B et T matures, sans autre défaut apparent. L'injection de cellules ES normales dans des blastocystes issus de souris dépourvues du gène *RAG-2* conduit à l'obtention de chimères somatiques qui contiennent des populations de lymphocytes normaux tous dérivés des cellules ES. Ceci démontre la capacité des cellules ES injectées à restaurer le phénotype *RAG-2*<sup>-/-</sup> par un procédé que les auteurs ont appelé la complémentarité des blastocystes *RAG-2*<sup>-/-</sup>. Un tel procédé peut être utilisé pour déterminer le rôle de n'importe quel gène dans la production et/ou la différenciation des cellules B ou T, en aval du stade de différenciation où intervient le gène *RAG-2*.

La première équipe a utilisé un plasmide avec le gène *ets-1* amputé de la région codant le domaine ETS. Les cellules ES *ets-1*<sup>-/-</sup> obtenues avec le plasmide correspondant sont introduites dans un blastocyste *RAG-2*<sup>-/-</sup> (Fig 13). Les souris chimères *RAG-2*<sup>-/-</sup> *ets-1*<sup>-/-</sup> possèdent des cellules B et T (Bories et al., 1995). Cependant, l'absence de protéines Ets-1 a un effet marqué sur le développement et la fonction de ces deux types de cellules. En effet, le nombre de cellules T est très réduit et ces cellules présentent une susceptibilité

accrue à la mort par apoptose *in vitro*. Quant aux cellules B, leur nombre n'est pas affecté mais la proportion des cellules plasmiques sécrétrices des IgM est anormalement élevée. Bories et al. (1995) en ont conclu que les protéines Ets-1 sont des facteurs de transcription essentiels pour le maintien des populations de cellules B et T au repos.

La seconde équipe a utilisé le gène *ets-1* amputé de l'exon III soit une délétion plus grande du gène. Elle a aussi obtenu des souris chimères *RAG-2<sup>-/-</sup> ets-1<sup>-/-</sup>*. Les résultats sont similaires à ceux de Bories et al. (1995), du moins pour les cellules T car les auteurs ne présentent pas les résultats sur les cellules B (Muthusamy et al., 1995). Il semble donc que les défauts observés soient bien liés à une mutation dans le gène *ets-1*.

#### b- Le gène *fli-1*

Le gène *fli-1* ("Friend Leukemia Integration-1") a été mis en évidence chez la souris suite à l'étude de leucémies érythroïdes induites par le rétrovirus murin FMuLV ("Friend Murine Leukemia Virus"). L'intégration de ce rétrovirus en amont du gène est à l'origine de sa surexpression "naturelle" (Ben-David et al., 1990 et 1991).

Dans le but d'étudier les fonctions biologiques de la protéine Fli-1 *in vivo*, deux équipes ont utilisé les techniques de transgénèse et de KO.

La première équipe a surexprimé le gène *fli-1* par la technique de transgénèse (Zhang et al., 1995). Elle a utilisé un promoteur murin constitutif et ubiquiste. Les souris transgéniques développent des anomalies lymphocytaires qui conduisent à une maladie auto-immune puis à la mort. Paradoxalement, ces souris ne présentent pas de leucémies alors que la transgénèse était sensée mimer l'effet du rétrovirus murin FMuLV à l'origine de leucémies érythroïdes. En fait, ces résultats laissent penser que la surexpression du gène *fli-1* ou son expression ectopique perturbe la fonction et l'apoptose des cellules lymphoïdes normales.

Quant aux souris recombinées *fli-1<sup>-/-</sup>*, elles ne présentent pas de phénotype altéré (Mélet et al., 1996). Les auteurs constatent cependant une réduction de taille du thymus, bien que toutes les populations lymphocytaires soient présentes. De plus, le rétrovirus FMuLV s'intègre, comme chez les souris normales, au niveau du locus *fli-1* et induit des leucémies cependant le délai d'apparition est plus long. Il semblerait donc que l'expression du gène *fli-1* ne soit pas indispensable au développement des animaux, ni à la transformation des cellules par FMuLV. Ces résultats doivent cependant être confirmés car les souris recombinées produisent, en faible quantité, une version tronquée de la protéine Fli-1 qui contient encore le domaine ETS. Le KO ne serait donc pas total.

### c- Le gène *pu-1*

Le gène *pu-1* est exprimé exclusivement dans les lignées hématopoïétiques. Une équipe a donc étudié sa surexpression sous le contrôle d'un "enhancer" spécifique des cellules hématopoïétiques. Les souris transgéniques développent des érythroleucémies (Moreau-Gachelin et al., 1996). Au début de la maladie, elles deviennent sérieusement anémiques car les proérythroblastes sont bloqués à ce stade de la différenciation. Une rémission complète mais pas durable de la maladie est observée après transfusion d'érythrocytes. Ceci suggère que l'expression constitutive du gène *pu-1* est en relation avec le blocage de la différenciation des proérythroblastes. Nous pouvons remarquer que la surexpression du gène *pu-1* n'altère pas les autres lignées hématopoïétiques.

Deux équipes ont travaillé sur le KO du gène *pu-1*, chacune en supprimant l'exon qui code le domaine ETS. Dans le premier cas, les souris recombinaées *pu-1*<sup>-/-</sup> meurent au stade embryonnaire entre 14,5 et 17 jours de développement (Scott et al., 1994 et 1997). En effet, la mutation affecte le développement des lymphocytes B et T, des monocytes et des granulocytes, la protéine Pu-1 intervenant dans le développement des lignées myéloïdes au niveau des progéniteurs multipotents au cours de l'hématopoïèse foetale. Dans la deuxième étude, les souris recombinaées naissent à terme mais meurent de septicémie dans les 24 à 48 heures qui suivent la naissance (McKercher et al., 1996 ; Tondravi et al., 1997). Des souris peuvent être maintenues en vie par un traitement antibiotique pour combler l'absence de macrophages mais elles présentent une ostéopétrose précoce due à l'absence d'ostéoclastes. Cette seconde expérience met en évidence un rôle de la protéine Pu-1 dans la différenciation en macrophages et ostéoclastes à partir d'une cellule progénitrice hématopoïétique. Dans les deux cas, les souris mutées présentent des défauts des lignées myéloïde et lymphocytaire mais les cellules absentes ne sont pas les mêmes. Ces différences peuvent venir du site d'insertion du gène de résistance légèrement différent, des cellules ES utilisées, d'une protéine résiduelle présente dans l'une ou l'autre étude mais l'explication n'est toujours pas connue.

### d- Le gène *spi-B*

Malgré la grande similarité entre les protéines Spi-B et Pu-1, leur rôle *in vivo* n'est pas redondant. En effet, si la protéine Spi-B compensait l'absence de la protéine Pu-1, les souris *pu-1*<sup>-/-</sup> n'auraient pas de phénotype. Pour déterminer le rôle propre au gène *spi-B*, une équipe a généré des souris mutées *spi-B*<sup>-/-</sup>. A la différence de celles mutées *pu-1*<sup>-/-</sup>, elles sont viables, fertiles et possèdent des lymphocytes B et T matures. Cependant, elles présentent de sévères anomalies. D'abord, bien que les cellules B répondent normalement à une stimulation *in vitro*, elles prolifèrent peu et meurent en réponse à une interaction entre



un antigène et son épitope. Ensuite, les animaux présentent des réponses antigéniques dépendantes des cellules T anormales et produisent en faible quantité certains antigènes spécifiques. En fait, les souris recombinantes *spi-B*<sup>-/-</sup> ont un grave défaut dans la maintenance et la formation des centres germinaux (Su et al., 1997).

## 2- L'angiogenèse

Des souris mutées *tel*<sup>-/-</sup> ont été générées afin de déterminer l'implication de ce gène dans le développement (Wang et al., 1997). Les embryons meurent entre 10,5 et 11,5 jours de développement. Ils présentent des défauts dans l'angiogenèse du sac vitellin et dans l'apoptose intra-embryonnaire des cellules neurales et mésenchymateuses. En revanche, l'hématopoïèse dans le sac vitellin ne semble pas affectée par l'absence du gène *tel*. Toute l'étude établit donc que la protéine Tel est requise dans la construction du réseau vasculaire et dans la survie de nombreux types cellulaires au cours du développement précoce de l'embryon. L'hypothèse proposée est que la protéine Tel régulerait l'apoptose.

L'étude de son expression ne faisait pas penser à cette fonction. En effet, le gène *tel* est largement exprimé, chez la souris, d'abord dans l'embryon et le sac vitellin et ensuite dans de nombreux tissus adultes (Wang et al., 1997). Par contre, les gènes *ets-1*, *erg* et *fli-1* sont connus pour être exprimés lors de l'angiogenèse qu'elle soit embryonnaire ou induite (Quéva et al., 1989 ; Dhordain et al., 1995 ; Mélet et al., 1996). Les résultats des différents KO nous permettront de mesurer son implication dans ce phénomène.

## 3- La formation du squelette

Sumarsono et al. (1996) ont généré une lignée de souris transgéniques afin de rechercher les conséquences de la surexpression du gène *ets-2*. Pour cela, ils ont utilisé le promoteur murin de la méthallothionéine, inductible par le sulfate de zinc. Les souris développent des anomalies dans les os du squelette dérivés de l'ossification intra-cartilagineuse, dans les sutures entre les os plats et dans les régions des os plats formés par l'ossification intra-membranaire. Ces défauts présentent des similitudes avec les anomalies squelettales dues à la trisomie 16 chez la souris, et à la trisomie 21 ou syndrome de Down chez l'homme. En effet, chez la souris, le gène *ets-2* est situé sur le chromosome 16 et chez l'homme sur le chromosome 21, au niveau de la région impliquée dans le syndrome de Down. Il est donc surexprimé dans chacune de ces trisomies (Sumarsono et al., 1996).

Ces résultats ont deux implications. D'abord, le gène *ets-2* semble être important dans le développement de l'os et du cartilage, ce qui concorde avec de précédentes données démontrant une expression de ce gène dans ces tissus en cours de développement (Maroulakou et al., 1994). Ensuite, la surexpression du gène *ets-2* entraîne la genèse

Membre de la famille <i>ets</i>	Altération génétique	Phénotype	Référence
<i>ets-1</i>	KO dans la lignée lymphoïde	Viable. Développement anormal des cellules B et T.	Bories et al., 1995 Muthusamy et al., 1995
<i>ets-2</i>	transgénèse	Viable. Anormalités du squelette.	Sumarsono et al., 1996
<i>fli-1</i>	transgénèse	Léthal. Maladie auto-immune..	Zhang et al., 1995
	KO	Viable. Nombre de lymphocytes diminué. Protéine résiduelle.	Mélet et al., 1996
<i>tel</i>	KO	Léthal. Défaut d'angiogenèse dans le sac vitellin. Apoptose anormale.	Wang et al., 1997
<i>pu-1</i>	transgénèse	Viable. Erythroleucémie.	Moreau-Gachelin et al., 1996
	KO	Léthal. Pas de différenciation lymphoïde et myéloïde.	Scott et al., 1994 McKercher et al., 1996
<i>spi-B</i>	KO	Viable. Anormalies au niveau des centres germinatifs lymphoïdes.	Su et al., 1997

**Figure 14 :** Tableau récapitulatif des mutations ciblées et des transgénèses effectuées à ce jour dans la famille *ets*.

d'anomalies squelettiques présentes dans le syndrome de Down. Ceci pourrait permettre d'établir une relation entre un gène situé sur le chromosome 21 humain et une pathologie spécifique du phénotype du syndrome de Down (Sumarsono et al., 1996).

Le gène *erg*, localisé aussi sur le chromosome 21 humain, est particulièrement exprimé lors de la mise en place du cartilage (Dhordain et al., 1995). Le KO de ce gène, en cours dans notre équipe, pourrait révéler des malformations du squelette. D'autre part, comme nous venons de le voir, l'absence du gène *pu-1* impliqué dans l'hématopoïèse génère aussi de l'ostéopétrose car les ostéoclastes sont dérivés de la lignée lymphocytaire.

Tout au long de ces trois premières parties, nous avons tenté de répondre, à partir des connaissances acquises, à la question : "Comment cette grande famille de facteurs de transcription homologues peut réguler différemment différents gènes dans différents types cellulaires et en réponse à différents signaux extracellulaires?".

Nous avons vu que la liaison de ces protéines à l'ADN sur des séquences cibles distinctes est un des premiers éléments de réponse. Des domaines particuliers tel le domaine PNT, spécifiques d'un membre ou d'un ensemble de membres, peuvent permettre l'interaction avec divers partenaires et une activation ou une répression de la transcription. Enfin, des études *in vivo* par KO ou transgénèse (Fig 14) suggèrent que chaque protéine Ets dirige individuellement un procédé biologique distinct mais de possibles redondances fonctionnelles peuvent exister.

## Quatrième partie : Le gène *erg*

### I- Sa découverte

Le gène *erg* ("Ets Related Gene") humain est un membre de la famille des gènes *ets*. Initialement, il a été mis en évidence par criblage d'une banque d'ADNc d'une lignée humaine d'adénocarcinome de colon, COLO320, avec une sonde correspondant à l'ADNc *ets-2* humain (Reddy et al., 1987).

Par analyse génétique de cellules somatiques, les auteurs ont situé le gène *erg* sur le chromosome 21 (Reddy et al, 1987). Parmi les gènes de la sous-famille ETS, le gène *erg* ainsi que les gènes *ets-1*, *ets-2* et *fli-1* présentent la particularité d'être localisés deux à deux sur les chromosomes humains : 11q23-24 pour *ets-1* et *fli-1* et 21q22.3 pour *ets-2* et *erg* (Watson et al., 1985 ; Baud et al., 1991 ; Rao et al., 1987).

Chez la souris, son organisation génomique n'est que partiellement connue (Fig 19) et présente au moins onze exons dont certains sont alternatifs ainsi qu'au moins deux promoteurs alternatifs (Adam Hart, Ismail Kola et Martine Duterque-Coquillaud, communications personnelles).

### II- Ses propriétés oncogéniques

Le gène *erg* est un proto-oncogène qui peut être impliqué dans plusieurs pathologies aussi diverses que des leucémies, des sarcomes et le syndrome de Down.

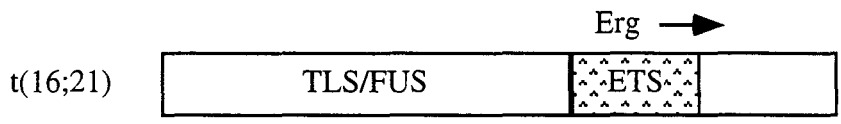
#### 1- L'oncogène *erg*

Le gène *erg* humain est un proto-oncogène (Hart et al., 1995). En effet, un des produits issus de ce gène possède une activité mitogénique et transformante puisqu'il est capable de transformer des cellules NIH3T3.

#### 2- Des translocations chromosomiques

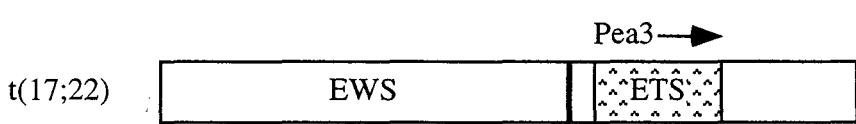
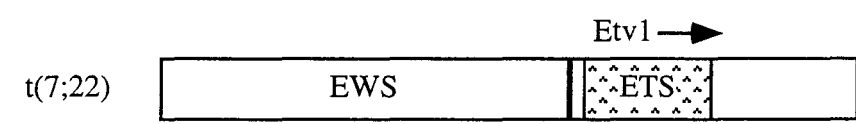
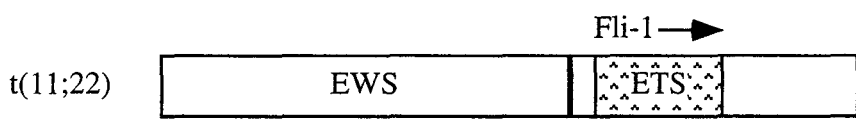
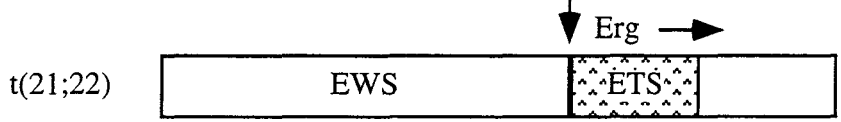
Le gène *erg* est impliqué dans deux types de translocations chromosomiques, t(16;21) et t(21;22) qui sont à l'origine de protéines chimères possédant des capacités oncogéniques.

**Leucémie myéloïde**



↑  
point de cassure  
↓

**Sarcomes d'Ewing**



**Figure 15** : Protéines chimères rencontrées dans diverses translocations chromosomiques, cause de leucémie ou de sarcome. D'après Ichikawa et al., 1994.

#### a- La translocation t(16;21)

Dans plusieurs types de leucémies myéloïdes, Shimizu et al. (1993) ont observé la translocation (16;21) qui rompt le gène *erg* toujours au même endroit. Cette translocation (16;21)(p11;q22) associe la partie 5' du gène *TLS-FUS* avec la partie 3' du gène *erg* codant le domaine ETS (Fig 15). L'association inverse qui donnerait le transcrite chimère *erg-TLS/FUS* n'a jamais été détectée (Ichikawa et al., 1994). Le gène *TLS/FUS* code une protéine qui présente une grande similarité avec la protéine EWS impliquée aussi dans des translocations (Croizat et al., 1993). En fait, les protéines TLS/FUS et EWS appartiennent à une famille de protéines qui se lient à l'ARN et l'ADN simple brin, la famille TET ("*TLS/FUS*, *EWS*, *TAF<sub>n</sub>68* ") nouvellement définie. Tout comme *TAF<sub>n</sub>68* ("*TATA-Binding-Protein Associated Factor*"), les facteurs TLS/FUS et EWS pourraient jouer un rôle spécifique dans l'initiation de la transcription car ils interagissent avec le complexe TFIID ("*basal Transcription Factor*") et la RNA polymérase II (Bertolotti et al., 1996 et 1997).

#### b- La translocation t(21;22)

De la même manière, dans le sarcome d'Ewing, tumeur de l'os long chez l'enfant, et plus largement dans des tumeurs d'origine neuroectodermique, on retrouve la translocation t(21;22)(q22;q12) qui associe la région amino-terminale de la protéine EWS avec le domaine ETS de la protéine Erg (Zucman et al., 1993 ; Sorensen et al., 1994).

Le transcrite chimère *EWS-erg* qui résulte de cette translocation est présent dans environ 15% seulement des sarcomes d'Ewing étudiés. En effet, différents domaines ETS sont fusionnés avec la partie amino-terminale de la protéine EWS, résultat de différentes translocations (Fig 15). Dans 85% des cas, la translocation t(11;22)(q24;q12) fait intervenir le gène *fli-1*, dont la séquence et l'organisation génomique sont très similaires à celles du gène *erg* (Delattre et al., 1992). Dans de plus rares cas, trois autres types de translocation ont été identifiés, t(7;22)(p22;q12), t(17;22)(q12;q12) et t(2;21;22)(q33;q22;q12). Ils font respectivement intervenir trois autres membres de la famille *ets*, le gène *er81* ("*Ets Related*" alias *etv1* "*Ets Translocation Variant*"), le gène *pea3* ("*Polyomavirus Enhancer Activator*" alias *e1af* "*adenovirus E1A Factor*") et le gène *fev* ("*Fifth Ewing Variant*") pour aboutir aux mêmes types de transcrits chimères (Jeon et al., 1995 ; Kaneko et al., 1996 ; Peter et al., 1997). Le fait que cinq gènes de la famille *ets* soient impliqués dans la même pathologie suggère une redondance de leur fonction dans le phénomène de cancérisation conduisant à ce type de tumeur.

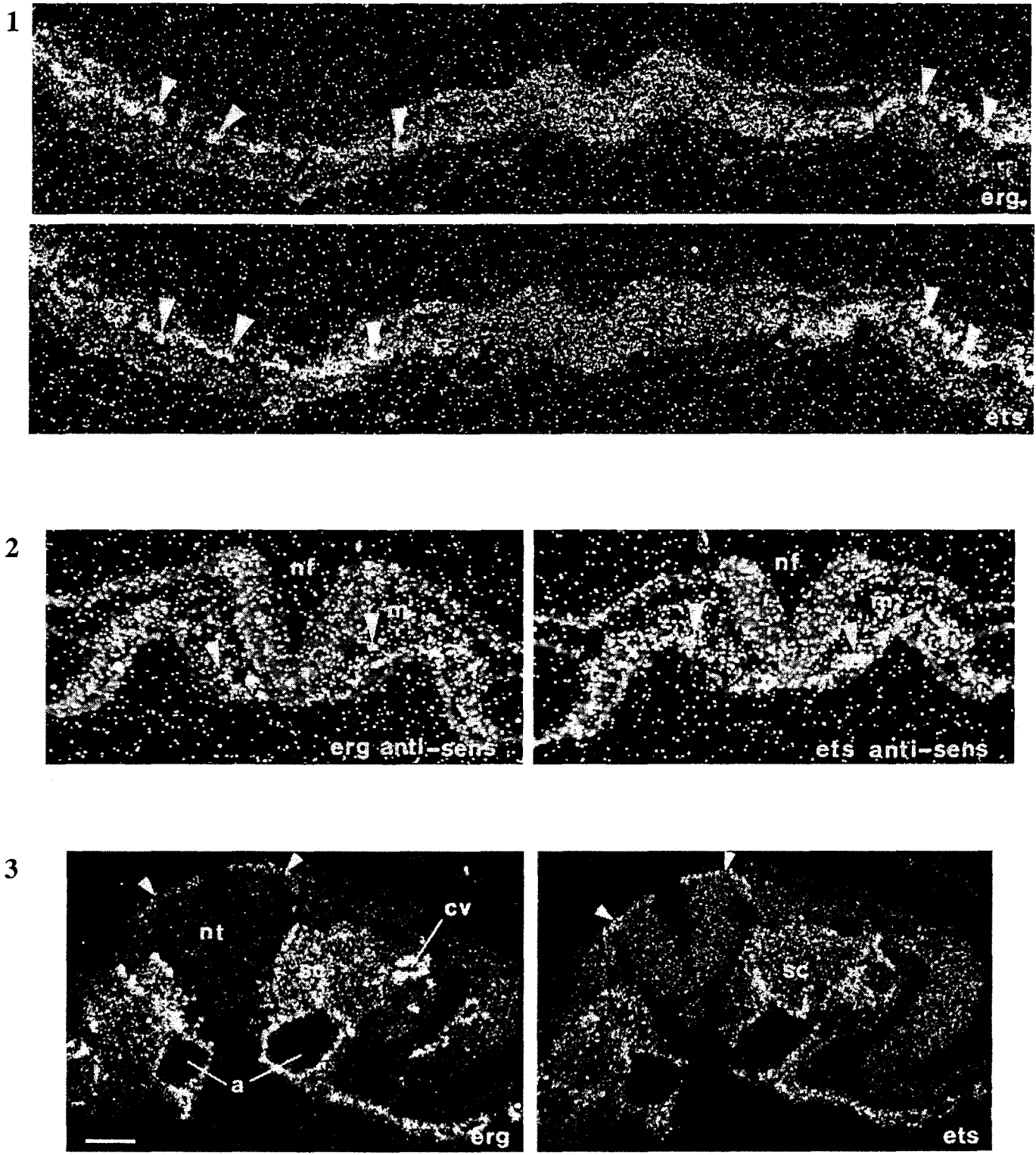
Différentes équipes ont démontré que la protéine chimère EWS-Fli-1 est nucléaire (May et al., 1993 ; Onho et al., 1994 ; Bailly et al., 1994 ; Magnaghi-Jaulin et al., 1996). Elle se

comporte comme un meilleur transactivateur que la protéine Fli-1 sauvage sur des promoteurs possédant des sites EBS. De plus, sa capacité de liaison à l'ADN n'est pas modifiée. Ces résultats suggèrent que la protéine chimère, facteur au pouvoir transactivateur accru, pourraient être entraînée vers des gènes cibles ou des cofacteurs inhabituels. De plus, à la différence de la protéine Fli-1, cette protéine chimère a une activité transformante quand elle est surexprimée dans des cellules NIH3T3. En effet, elle possède des propriétés intrinsèques qui la rendent oncogénique (May et al., 1993). Les pouvoirs transformants d'autres protéines chimères ont été testés sur des cellules en culture. Dans ces conditions, il semble que EWS-Erg soit moins transformant que EWS-Fli-1, et que EWS-Er81 ne le soit pas du tout (Braun et al., 1995). Bien sûr, ces études *in vitro* ne reproduisent que partiellement les conditions de transformation. Il est donc raisonnable de penser que les protéines chimères EWS-Erg, EWS-Er81, EWS-Pea3 et EWS-Fev sont, elles aussi, plus ou moins transformantes.

### 3- Le syndrome de Down

Chez l'homme, le syndrome de Down est le défaut de naissance le plus fréquent, environ un nouveau-né sur 700, et la cause de retard mental la plus importante. Dans la plupart des cas, il résulte d'une copie supplémentaire du chromosome 21 dans toutes les cellules de l'individu, et plus rarement, d'une trisomie 21 partielle. Les caractéristiques de ce syndrome sont un retard mental constant, un phénotype spécifique avec des anomalies morphologiques de la tête et des membres ainsi qu'une petite stature, une hypotonie, une hyperlaxité des ligaments, une malformation viscérale en particulier du cœur, un risque accru de leucémie et des développements précoces de neuropathologies comme la maladie d'Alzheimer.

Les gènes *erg* et *ets-2* sont localisés sur le chromosome 21, dans la région q22-3, celle-là même qui est impliquée dans le syndrome de Down et les pathologies qui lui sont associées (Crété et al., 1993 ; Dahmane et al., 1998). La cartographie précise de la région responsable de ce syndrome a permis de montrer que les gènes *erg* et *ets-2* ne sont pas parmi les gènes indispensables à l'expression de ce syndrome, mais ils pourraient jouer un rôle dans les pathologies qui lui sont associées (Nicole Créau, communication personnelle). En effet, des expériences de surexpression du gène *ets-2* dans des souris transgéniques ont montré que le phénotype de ces souris présente des similarités avec les malformations crâniennes observées chez des personnes atteintes du syndrome de Down (Sumarsono et al., 1996). Des observations similaires ont été constatées en surexprimant le gène *erg* dans des souris transgéniques (Adam Hart et Ismail Kola, communication personnelle).



**Figure 16 :** Hybridation *in situ* sur des embryons de poulet avec les sondes *erg* et *ets-1*. Au stade E1, le gène *erg* est exprimé dans les îlots sanguins dans l'ère extra-embryonnaire (1, flèches), et dans le mésenchyme (2). Au stade E2, il est exprimé en particulier au niveau des crêtes neurales (flèches, 3). D'après Dhordain et al., 1995.



### III- Son expression

L'étude du profil d'expression d'un gène est une étape clé dans la compréhension de son rôle biologique et tout particulièrement lorsqu'il appartient à une famille de gènes dont les rôles respectifs sont encore mal connus. En ce qui concerne le gène *erg*, plusieurs approches ont été utilisées pour définir ses domaines d'expression.

#### 1- Les études par "northern-blot"

Les premiers résultats concernant l'expression du gène *erg* ont révélé la présence d'ARNm dans plusieurs types cellulaires : dans la lignée humaine d'adénocarcinome du colon COLO320, dans la lignée humaine lymphoblastique MOLT4, et dans plusieurs lignées de cellules pré-B et B. Ils se présentent sous la forme de deux transcrits majeurs de 3,2kb et 3,6kb et un mineur de 5kb (Rao et al., 1987 ; Rivera et al., 1993). Dans l'organisme adulte, parmi tous les tissus testés, seul le thymus présente une très faible expression. Cependant, seuls les transcrits de 3,2k et 3,6kb y ont été détectés (Watson et al., 1992).

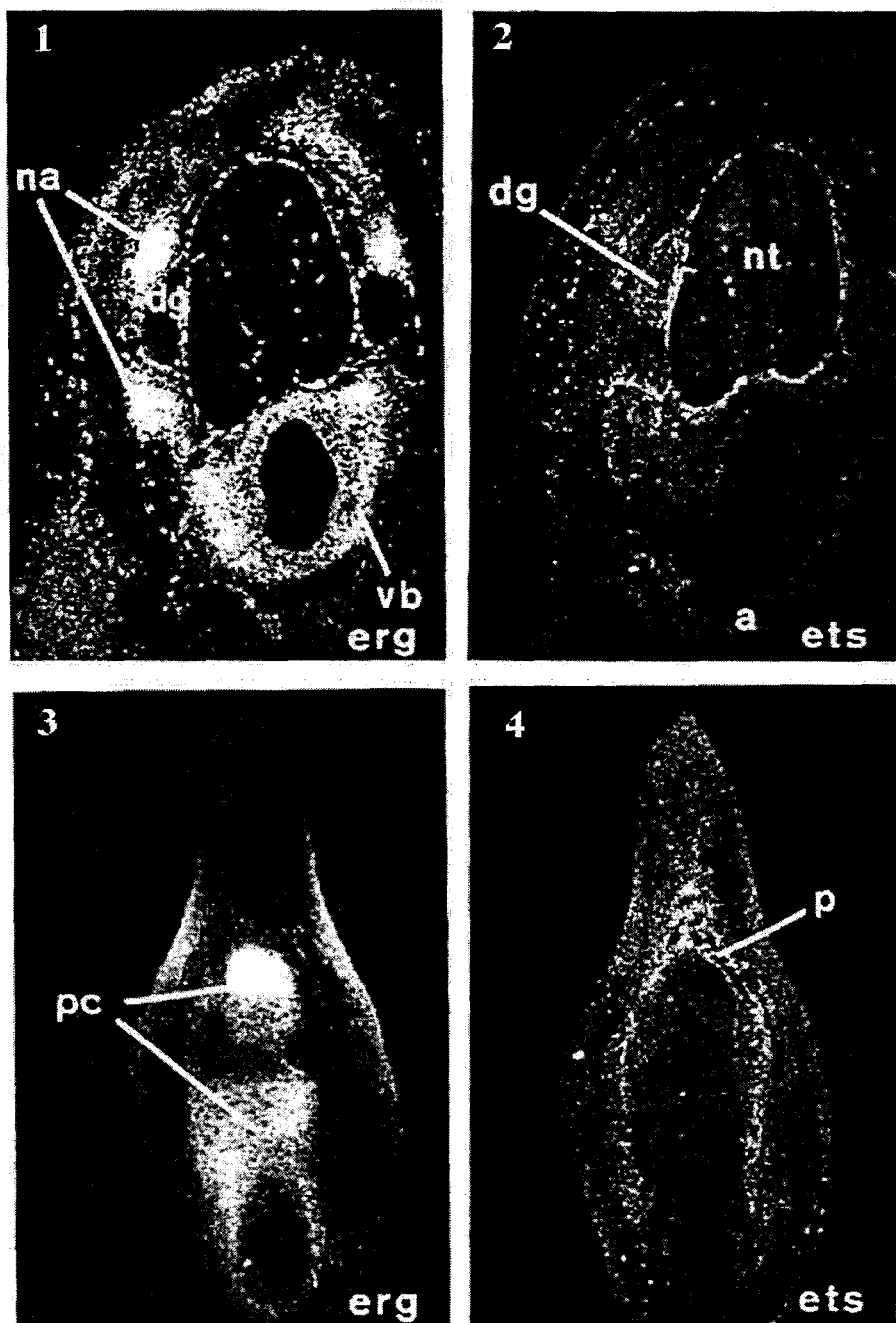
Dans le but de déterminer l'expression du gène *erg* au cours du développement, des analyses ont été effectuées sur des extraits d'embryons totaux ou de tissus de poulet (Dhordain et al., 1995). De manière surprenante, les transcrits de 3,2kb et 3,6kb sont présents à de fort taux dans les fibroblastes embryonnaires de poulet, aux stades de développement E4 à E7 et dans de nombreux tissus de nouveaux-nés. Par contre, ils sont en faible quantité dans le thymus et absents de la lignée lymphoïde MSB-1T. Le gène *erg* de poulet semble donc plus largement exprimé que son homologue humain.

#### 2- Les études par hybridation *in situ*

L'expression du gène *erg* au cours du développement embryonnaire restait flou. Elle a donc été étudiée par hybridation *in situ* chez le poulet. Le gène *erg* s'exprime d'une part de manière très précoce dans l'embryon où il partage des territoires d'expression avec le gène *ets-1*, et d'autre part, son expression est spécifiquement associée à la mise en place du tissu cartilagineux, à partir du troisième jour de développement (Dhordain et al., 1995).

##### a- Coexpression du gène *erg* avec d'autres membres de la famille

Dès les premières heures de développement du poulet, une forte expression du gène *erg* a été observée au niveau des îlots sanguins en formation dans l'aire extra-embryonnaire (Fig 16-1), puis dans le mésenchyme et les cellules endothéliales des vaisseaux en formation (Fig 16-2). A ce stade, cette expression correspond parfaitement au spectre d'expression du gène *ets-1* (Fig 16) et laisse supposer un rôle conjoint ou antagoniste de ces deux gènes



**Figure 17** : Hybridation *in situ* chez le poulet avec les sondes *ets-1* au niveau de la vertèbre (1 et 2) et du bourgeon de membre (3 et 4). L'expression du gène *erg* est restreinte aux zones de condensation du cartilage telles que la vertèbre elle-même (vb), l'arc neural (na) et les chondroblastes en différenciation (pc). D'après Dhordain et al., 1995.

dans la formation du système hématopoïétique, initié dans les îlots sanguins et dans la formation des vaisseaux (Quéva et al., 1993 ; Dhordain et al., 1995). Un deuxième site de coexpression de ces deux gènes, au deuxième jour de développement, concerne les cellules des crêtes neurales à l'origine de nombreux tissus dans l'embryon (Fig 16-3) (Dhordain et al., 1995). Ces cellules ont la particularité d'être en migration, phénomène dans lequel les gènes *erg* et *ets-1* pourraient être impliqués. En effet, la migration cellulaire implique l'intervention de nombreuses protéases dont l'expression pourrait être régulée par les protéines de la famille Ets (Wasylyk et al., 1991 ; Grévin et al., 1993 ; Buttice et al., 1996). D'autre part, aucun de ces deux gènes ne s'exprime dans le tissu épithélial ni dans le tissu nerveux quelque soit le stade de développement observé (Quéva et al., 1993 ; Dhordain et al., 1995).

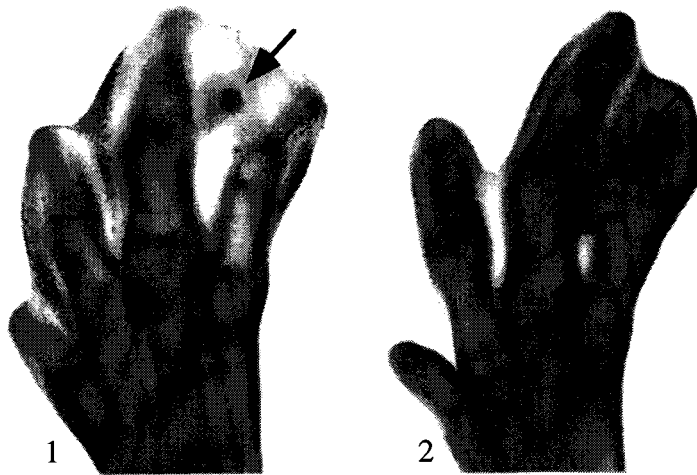
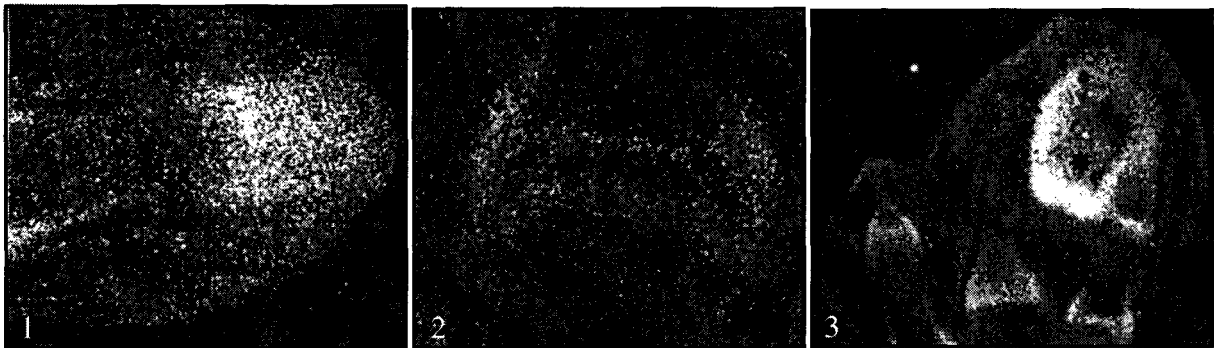
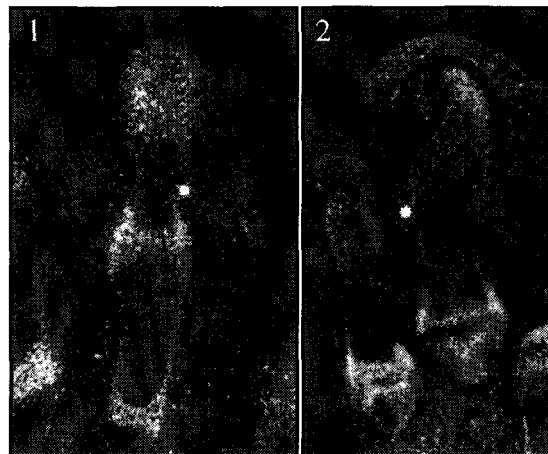
Des études identiques ont été faites chez la souris pour le gène *fli-1*. En fait, son profil d'expression est semblable à celui du gène *ets-1*. En effet, dans les stades d'embryogenèse très précoces, le gène *fli-1* est exprimé dans les lignées mésodermiques, tandis qu'après la morphogenèse, son expression est fortement diminuée dans le mésoderme mature (Mélet et al., 1996).

Malgré les modèles animaux différents, nous pourrions conclure que l'expression des gènes *erg*, *fli-1* et *ets-1* semble présenter des similitudes dans les premiers jours du développement de l'embryon, leurs sites d'expression se différenciant par la suite.

#### b- Le gène *erg* impliqué dans la mise en place du cartilage

L'une des observations les plus inattendues au cours de l'étude du gène *erg* chez le poulet par hybridation *in situ* fut la détection d'une forte expression du gène *erg* au niveau du cartilage en différenciation avant la formation du squelette. A partir du troisième jour de développement, le gène *erg* montre une aire d'expression spécifique mais qui ne se chevauche pas avec celle du gène *ets-1*.

Rappelons d'abord comment se forme le cartilage. Dès le troisième jour du développement, la formation du squelette débute par une étape transitoire de condensation des cellules du mésenchyme pour constituer le tissu précartilagineux. Cette étape est suivie de la différenciation cartilagineuse proprement dite, c'est-à-dire de la différenciation des préchondroblastes en chondroblastes puis en chondrocytes. Cette différenciation cellulaire s'accompagne du dépôt de la matrice extracellulaire spécifique du cartilage. De plus, un profond changement au niveau de l'expression des gènes est observé. En particulier, lors de la condensation du mésenchyme, une augmentation des ARNm du collagène de type II puis des ARNm du collagène de type IX est mise en évidence. Les deux protéines

**A****B****C**

**Figure 18** : Implantation de billes imprégnées de facteurs dans la patte d'embryon de poulet. Les billes sont indiquées par des flèches ou des astérisques. A- Billes dans l'espace interdigital avec (1) ou sans (2) facteurs chondrogéniques. B- Hybridation *in situ* avec une sonde *erg* dans les conditions normales dans un bout de pattes (1) ou au niveau d'une articulation (2), et 8h après implantation d'un bille imprégnée de TGF $\beta$  (3). C- Hybridation *in situ* avec une sonde *erg* dans un doigt , dans des conditions normales (1) ou après implantation d'une bille imprégnée d'un facteur pro-apoptotique.

correspondantes sont d'ailleurs des marqueurs de cette étape de condensation du cartilage (Hall et Miyake, 1995).

Une expression du gène *erg*, particulièrement forte, a été observée au cours de la condensation précartilagineuse. Cette observation est illustrée au cours de la formation de la vertèbre et du bourgeon de membre au sixième jour du développement (Fig 17). Cette expression, comme l'étape de condensation cellulaire, est transitoire, environ 24 h, et se restreint ensuite aux régions des articulations, seules zones de condensation et de croissance du cartilage qui subsistent (Dhordain et al., 1995).

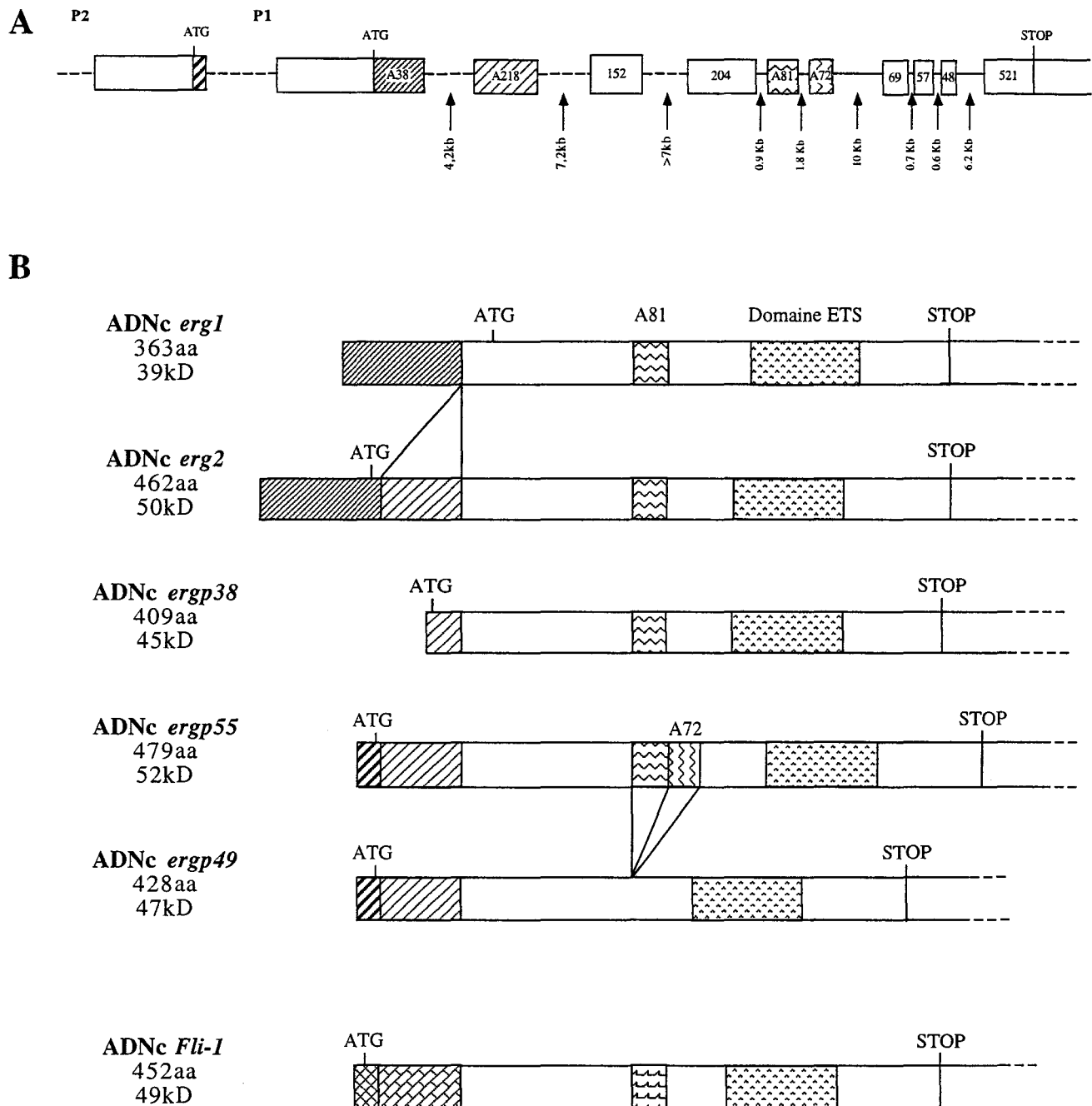
Parmi les gènes de la famille *ets*, seul *ets-2* est aussi exprimé lors de la mise en place du cartilage et des os (Maroulakou et al., 1994). Par ailleurs, il est largement exprimé au cours de l'embryogenèse et à un plus faible taux dans divers tissus adulte.

L'hybridation *in situ* sur coupe d'embryon ne permet pas à elle seule de définir l'implication du gène *erg* dans la mise en place du cartilage. C'est pourquoi une approche différente a été utilisée pour mettre en évidence l'expression de ce gène dans les étapes de la différenciation cartilagineuse. Elle a été réalisée en collaboration avec un chercheur italien, Patrizio Castagnola, capable de cultiver *in vitro* des cellules cartilagineuses de poulet et d'identifier chaque étape de la différenciation. En fait, ce sont les toutes premières étapes qui s'accompagnent d'une forte expression du gène *erg*, et cette expression décroît ensuite au cours de la différenciation (Martine Duterque-Coquillaud, communication personnelle).

### 3- "Des petites billes"

Enfin, une étude a permis de suivre l'expression du gène *erg* aviaire au cours de l'induction ectopique de cartilage. En effet, connaissant l'effet chondrogénique des facteurs de croissance TGF $\beta$ 1 et TGF $\beta$ 2, il est possible d'induire localement la formation de cartilage (Fig 18A). Cette induction s'effectue grâce à des billes imprégnées des facteurs de croissance et capables, une fois greffées, de diffuser ces molécules. Les billes sont greffées dans l'espace interdigital d'un membre d'un embryon de poulet de six jours, soit 20 à 30 heures avant la mise en place du cartilage (Gañan et al. 1996).

L'implantation de bille imprégnée de TGF $\beta$ 1 ou TGF $\beta$ 2, dans un espace interdigital, fournit un support expérimental intéressant pour mettre en évidence les facteurs impliqués dans la formation du cartilage ectopique résultant. L'expression du gène *erg* précède cette formation, ce qui le définit comme un des facteurs impliqués (Gañan et al. 1996). Les auteurs ont pu constater que l'expression normale est restreinte aux articulations et au bout des doigts, mais que la mise en place de la bille imprégnée de TGF $\beta$  est à l'origine d'une



**Figure 19 :** Le gène *erg*. **A-** Organisation génomique du gène *erg* humain. Les exons sont nommés d'après leur taille en paire de bases, les exons alternatifs sont représentés en hachuré et avec la lettre A. La taille des introns est donnée à titre indicatif. Les deux promoteurs potentiels sont indiqués par P1 et P2. D'après Martine Duterque-Coquillaud, Adam Hart et Ismail Kola, communications personnelles. **B-** Les ADNc isoformes issus des gènes *erg* humain et l'ADNc issu du gène *fli-1* humain. Leur taille en acides aminés et leur masse moléculaire correspondantes sont indiquées. D'après Duterque-Coquillaud et al., 1993.

forte expression du gène *erg* (Fig 18B). Dans ce dernier cas, l'expression est détectable dès la huitième heure après l'implantation de la bille et pendant 20 à 30 heures.

Dans une autre étude, les billes sont imprégnées des protéines BMP-2 ("Bone Morphogenetic Protein") ou OP-1 ("Osteogenic Protein") qui induisent l'apoptose et empêchent la formation du cartilage. L'hybridation *in situ* démontre que l'expression du gène *erg* est alors inhibée autour de la bille et qu'elle est aussi perturbée dans son aire habituelle (Fig 18C) (Macias et al., 1997).

Tous ces résultats sont en faveur d'un rôle important du gène *erg* dans l'initiation de la chondrogenèse. Les produits de ce gène sont donc parmi les facteurs de transcription les plus précoces connus dans ce phénomène. Par ailleurs, nous pouvons noter la localisation particulière de l'expression du gène *erg*, au sixième jour de développement, de part et d'autre du bourgeon de membre (Fig 17-3). Cette zone ne correspond pas à un tissu précartilagineux mais à la mise en place des tendons chez le poulet. L'implication du gène *erg* dans ce mécanisme est actuellement à l'étude.

#### **IV- Les produits du gène *erg***

Lors de la mise en évidence du gène *erg* à partir d'une banque d'ADNc de cellules COLO320, seuls deux ADNc correspondants à deux transcrits alternatifs *erg1* et *erg2* avaient été décrits (Reddy et al., 1987 ; Rao et al., 1987). Par la suite, à partir d'une banque d'ADNc de foie foetal humain et avec une sonde spécifique obtenue par RT-PCR à partir d'ARN de cellules COLO320, trois nouveaux ADNc ont été isolés, *ergp55*, *ergp49* et *ergp38* (Duterque-Coquillaud et al., 1993). Enfin, un ADNc appelé *erg3* a été mis en évidence et il correspondait à *ergp55* (Prasad et al., 1994).

##### 1- Les ADNc issus du gène *erg*

Au moins cinq ADNc différents sont issus du gène *erg*, *erg1*, *erg2*, *ergp55*, *ergp49* et *ergp38* (Fig 19). Ils diffèrent dans leurs parties 5' et centrale.

Tout d'abord, *erg1* et *erg2* diffèrent dans leur partie 5' par l'insertion de 218 nucléotides dans *erg2*. Cette insertion due à un événement de promotion alternative provoque un décalage dans le cadre de lecture ouverte de la partie 5', c'est-à-dire l'utilisation d'un codon d'initiation alternatif dans *erg2*. Rao et al. (1987) suggèrent aussi des sites de polyadénylation différents.

Par ailleurs, la partie 5' des ADNc *ergp55* et *ergp49* présente une séquence très différente de celle de *erg1* et *erg2* (Duterque-Coquillaud et al., 1993). En effet, nous pouvons constater la présence d'un nouvel exon d'initiation et par conséquent un nouveau

codon d'initiation. Cette grande différence entre d'un côté les ADNc *erg1* et *erg2* et de l'autre *ergp55* et *ergp49* est vraisemblablement due à l'existence de deux promoteurs distincts (Fig 19) (Adam Hart et Ismail Kola, communication personnelle). Ce phénomène a déjà été décrit dans la famille *ets* pour le gène *ets-1* aviaire (Crépieux et al., 1993).

Enfin, dans leur partie centrale, les ADNc présentent tous une insertion de 81 nucléotides excepté *ergp49*. Par contre, seul *ergp55* possède une insertion de 72 nucléotides en aval des 81 nucléotides. L'étude de l'ADN génomique a permis de montrer que ces deux régions de 81 nucléotides et de 72 nucléotides correspondent en fait à des exons alternatifs (Duterque-Coquillaud et al., 1993).

Par "northern-blot", seuls deux ARNm majeurs de 3,2kb et 3,6kb et un mineur de 5kb ont été détectés. Actuellement, aucun résultat expérimental ne permet d'associer ces ARNm aux différents ADNc mis en évidence.

## 2- Les ADNc *erg* et *fli-1*

A ce jour, un seul ADNc issu du gène *fli-1* a été décrit (Ben-David et al., 1990). Une forte similarité entre les ADNc *fli-1* et *erg2* a été observée. Sur l'ensemble de leur séquence codante, elle s'élève à 70% d'homologie (Ben-David et al., 1991). De plus, l'ADNc *fli-1* partage la même extrémité 5' que les ADNc *ergp55* et *ergp49* et ne possède qu'un des deux exons alternatifs, celui de 81 nucléotides (Fig 19) (Duterque-Coquillaud et al., 1993).

L'existence de cinq ADNc différents issus du gène *erg* plaide en faveur de l'existence d'autres ADNc *fli-1* générés par des promotions et des épissages alternatifs semblables à ceux rencontrés pour le gène *erg*. Nous en discuterons dans les résultats.

## **V- Les propriétés transactivatrices des protéines Erg**

*In vivo*, à l'aide d'un anticorps spécifique, seule une protéine phosphorylée de 52kDa a été mise en évidence dans la fraction nucléaire de cellules lymphoblastiques humaines. Elle correspond à l'ADNc *erg2* (Murakami et al., 1993). *In vitro*, tous les ADNc issus du gène *erg* sont à l'origine de protéines que nous appellerons Erg1, Erg2, Ergp55, Ergp49 et Ergp38. Elles présentent toutes des caractéristiques de facteurs de transcription mais sont de faibles transactivateurs (Rao et al., 1987 ; Duterque-Coquillaud, 1993).

### 1- Les protéines Erg se lient à l'ADN

La protéine Erg2 est nucléaire ainsi que les isoformes Ergp55, Ergp49 et Ergp38 (Murakami et al., 1993 ; Martine Duterque-Coquillaud, communication personnelle). De plus, les protéines Erg1 et Erg2 sont capables de se fixer à l'ADN *via* un site consensus



pour les protéines Ets (Reddy and Rao, 1991) par leur domaine ETS (Murakami et al., 1993) Il a été démontré que les trois isoformes Ergp55, Ergp49 et Ergp38 interagissent de manière comparable avec un site même EBS (Duterque-Coquillaud et al., 1993).

La protéine Erg2 se lie à la séquence cible consensus GGAA/T. La séquence cible optimale a été déterminée en utilisant une série d'oligonucléotides aléatoires : C/G C/a **GGAA** G/a T (Murakami et al., 1993). Bien que le domaine ETS des protéines Erg et Fli-1 soit homologue à 98%, leur séquence cible optimale n'est pas tout à fait la même. En effet, pour la protéine Fli-1, nous observons : T/g N G/A A C C **GGAA** G/a T/c A/G G/A (Mao et al., 1994). Nous pouvons envisager que la région en amont du domaine ETS, région la moins homologue entre les protéines, puisse influencer la liaison à l'ADN, ce qui expliquerait cette différence d'affinité.

Nous savons que les exons alternatifs sont localisés dans deux régions distinctes des protéines Erg. D'une part à l'extrémité amino-terminale et d'autre part au centre des protéines. Bien que le mécanisme d'épissage alternatif ne concerne pas le domaine ETS, il n'est pas exclu que la présence ou l'absence d'un ou deux exons affecte la fonction de fixation à l'ADN. En effet, ce phénomène a déjà été observé pour une des protéines isoformes issue du gène *ets-1* qui ne possède pas l'exon 7 situé en amont du domaine ETS (Wasylyk et al., 1992).

## 2- Les protéines Erg transactivent

Des expériences de co-transfections transitoires ont permis de démontrer que les protéines Erg1 et Erg2 étaient capables d'activer la transcription d'un gène rapporteur après fixation spécifique sur un site EBS artificiel (Reddy and Rao, 1991). Des études similaires montrent que les protéines Ergp55, Ergp49 et Ergp38 sont aussi des transactivateurs (Duterque-Coquillaud et al., 1993).

Plusieurs gènes cibles potentiels ont été étudiés par la suite. C'est le cas des gènes humains de la collagénase1 et de la stromélysine1 dont nous détaillerons l'étude dans les résultats.

Par ailleurs, une étude a été faite sur le promoteur du gène *vWF* ("von Willebrand Factor"), une des cibles potentielles des protéines Erg. La protéine *vWF* permet l'adhérence et l'agrégation des plaquettes et sa mutation provoque la maladie de von Willebrand, un des plus fréquents désordres de l'hémostase. La distribution cellulaire de ce facteur est restreinte aux cellules endothéliales, aux plaquettes et aux mégacaryocytes. Deux raisons

principales sont à l'origine de l'étude de son promoteur avec la protéine Ergp55 : le promoteur contient deux sites EBS potentiels et les gènes *ets-1* et *erg* sont aussi exprimés dans les cellules endothéliales (Schwachtgen et al., 1997). En fait, les protéines c-Ets-1 et Ergp55 sont capables de d'induire la transcription basale de ce promoteur en co-transfection.

Cependant, toutes ces protéines isoformes sont de faibles transactivateurs puisque leur capacité transactivatrice varie entre six et dix fois. Aucun effet additif, synergique ou antagoniste n'a été observé lorsque les protéines isoformes Erg sont co-transfectées. De plus, ces résultats sont obtenus dans un système artificiel de transfections transitoires vis-à-vis d'un consensus optimisé. Ils n'excluent donc pas l'existence à l'intérieur des promoteurs cellulaires, régulés *in vivo* par les protéines Erg, de séquences d'ADN spécifiquement reconnues par les différents produits du gène *erg*. Il est alors envisageable que les domaines alternatifs des isoformes permettent de moduler la reconnaissance et la fixation sur la séquence d'un promoteur particulier et ainsi de réguler la transcription de manière différente.

### 3- Les protéines Erg et leurs cofacteurs

L'activité transactivatrice et la spécificité des protéines Erg peuvent être modulées dans certaines conditions. En effet, elles sont susceptibles de s'associer physiquement avec d'autres facteurs de transcription. Préalablement, nous avons constaté que plusieurs protéines de la famille Ets interagissent avec divers facteurs. Les protéines Erg sont susceptibles d'interagir avec ces mêmes cofacteurs mais cela n'a pas été systématiquement étudié. De plus, les protéines Erg sont capables d'interagir physiquement avec les protéines Fos/Jun comme nous pourrions le voir dans les résultats.

Parmi les cofacteurs potentiels des protéines Erg, Rivera et al. (1993) ont proposé une interaction protéique entre les protéines Erg et un facteur de type HLH, la protéine E12.

En effet, la régulation de l'expression des immunoglobulines est dépendante de divers motifs présents dans les "enhancers" des gènes des chaînes lourdes et légères dont une classe de ces motifs sont les boîtes E. Ainsi, cinq boîtes E sont présentes sur l'"enhancer" du gène des chaînes lourdes, elles sont désignées par  $\mu E1$ ,  $\mu E2$ ,  $\mu E3$ ,  $\mu E4$  et  $\mu E5$ . Une famille de protéines de type HLH est capable de se lier à ces boîtes E. En plus des boîtes E très conservées, d'autres régions sont cruciales pour la régulation des gènes des immunoglobulines. C'est le cas d'un site appelé  $\pi$ , adjacent au site  $\mu E2$ , capable de fixer avec une grande affinité deux protéines de la famille Ets, Ergp55 et Fli-1. Des

expériences de transfections transitoires ont démontré que les protéines Ergp55 et Fli-1 sont capables d'activer, en synergie avec la protéine HLH E12, la transcription d'un gène rapporteur situé en amont d'un multimère des sites  $\mu$ E2 et  $\pi$ . La même étude avec l'"enhancer" complet du gène des chaînes lourdes a été infructueuse. Cependant, les auteurs n'ont pu déterminer une interaction physique entre les membres de la famille Ets et la protéine E12 de type HLH (Rivera et al., 1993).

L'étude du gène *erg* est loin d'être achevée et beaucoup de questions à propos de son rôle restent sans réponse. Quand je suis arrivée au laboratoire, le but de mon travail était d'étudier le mode d'action des protéines Erg et de déterminer leur redondance et leur spécificité. Pour cela, nous avons voulu savoir si elles avaient des activités transactivatrices différentes sur des promoteurs cellulaires de gènes cibles potentiels dues à la présence ou l'absence des régions codées par les exons alternatifs, si elles étaient capables d'interagir avec d'autres facteurs de transcription et enfin, si leur profil d'expression était le même. A toutes ces questions, nous avons apporté des réponses que nous allons développer dans le chapitre suivant.

# Résultats

## **I- Publication 1**

Dans cet article, nous montrons que les protéines humaines Ets-2 et Ergp55, seules ou en combinaison, régulent différemment les promoteurs des gènes humains de la collagénase1 et de la stromélysine1. La protéine Ets-2 induit les deux promoteurs tandis que la protéine Ergp55 active celui de la collagénase1 mais pas celui de la stromélysine1. De plus, Ergp55 inhibe totalement l'activité du promoteur de la stromélysine1 induite par Ets-2. Par ailleurs, la protéine Ergp55 seule est capable de se lier au promoteur de la stromélysine1 mais pas à celui de la collagénase1. La protéine Ergp55 est donc la première dans la famille Ets à être capable de distinguer les promoteurs de la collagénase1 et de la stromélysine1. C'est aussi la première fois qu'un phénomène d'antagonisme est mis en évidence entre deux protéines Ets.

Par mutagenèse dirigée, nous avons identifié un site EBS (-89) nécessaire à l'induction du promoteur de la collagénase1. De manière inattendue, la mutation d'un site AP-1 (-79) supprime également l'induction par Ergp55, ce qui suggère que Ergp55 coopère étroitement avec le complexe AP-1. En accord avec cette hypothèse, Ergp55 interagit physiquement avec le dimère Fos/Jun. En fait, Ergp55 induit la transcription du promoteur de la collagénase1 quand il est recruté par les protéines Fos/Jun, mais pas celui de la stromélysine1. La disposition différente des sites EBS et AP-1 dans ces promoteurs peut vraisemblablement être à l'origine de leur régulation différentielle.

# Erg, an Ets-family member, differentially regulates human collagenase1 (MMP1) and stromelysin1 (MMP3) gene expression by physically interacting with the Fos/Jun complex

Giovanna Buttice<sup>1</sup>, Martine Duterque-Coquillaud<sup>1</sup>, Jean Philippe Basuyaux<sup>1</sup>, Séverine Carrère<sup>1</sup>, Markku Kurkinen<sup>2</sup> and Dominique Stéhelin

<sup>1</sup>Laboratoire d'Oncologie Moléculaire, CNRS URA 1160, Institut Pasteur de Lille, 1 rue Calmette, BP. 245, 59019 Lille Cedex, France and <sup>2</sup>Center for Molecular Medicine and Genetics and Department of Pathology Wayne State University, School of Medicine, 5125 Biological Sciences Bldg, 5047 Gullen Mall, Detroit, Michigan 48202, USA

Collagenase1 (MMP1) and stromelysin1 (MMP3) are extracellular proteolytic enzymes that degrade connective tissue macromolecules and basement membranes. Both genes are regulated by the Ets and Fos/Jun families of transcription factors/oncoproteins. Here, we show that two members of the Ets-family, Ets2 and Erg and their combinations differentially regulate collagenase1 and stromelysin1 promoter activity. In transiently transfected cells, Ets2 activates both promoters whereas Erg induces collagenase1 but not stromelysin1 promoter activity. Moreover, Erg completely inhibits stromelysin1 promoter activation by Ets2. In gel shift assays however, the Erg protein bound little or not to the collagenase1 promoter, whereas it bound to the stromelysin1 promoter. By site-specific mutagenesis, we identified one major site at –88 that abolished collagenase1 promoter activation by Erg. Surprisingly, mutation of the collagenase1 AP1 site at –73 also abolished the activation by Erg suggesting that Erg cooperates with Fos/Jun in collagenase1 promoter regulation. Indeed, gel shift and *in vitro* protein interaction studies showed that Erg binds to the Fos/Jun complex. Thus, Erg represents the first example of a transcription factor that can distinguish between the collagenase1 and stromelysin1 promoters in that when Erg is recruited by Fos/Jun at the promoter, it transcriptionally activates collagenase1 gene but not stromelysin1 expression.

**Keywords:** extracellular matrix; metalloproteinases; oncogenes; gene regulation

## Introduction

The Ets family members comprise a large number of genes that encode transcription factors/oncoproteins (Crépieux *et al.*, 1994; Janknecht and Nordheim, 1993; Tymms and Kola, 1994; Wasylyk *et al.*, 1993) which are implicated in the generation of several types of cancers (Rabbitts, 1994). They are expressed widely during development and organogenesis. Ets1 is expressed in several cell types and is particularly abundant in endothelial cells during blood vessel formation (Wernert *et al.*, 1992). By contrast, Ets2

appears to be expressed in all proliferating tissues studied (Kola *et al.*, 1993) and can induce tumors in nude mice (Seth *et al.*, 1989). During chicken embryogenesis, the *ets* related gene *erg* is expressed in cells of mesodermal origin and may play a role in limb and skeleton formation (Dhordain *et al.*, 1995). It has also been shown that the *erg* gene can be rearranged in myeloid leukemia and Ewing's sarcoma, a long bone tumor of childhood (Ichikawa *et al.*, 1994; Sorensen *et al.*, 1994), suggesting that *erg* may participate in the regulation of particular genes implicated in the metastatic progression of cancer.

Different members of the Ets-family of transcription factors contain an homologous DNA binding domain, the Ets domain, that binds to the purine rich GGA(A/T) core sequence, but have divergent transcriptional activation domains (Karim *et al.*, 1990; Laudet *et al.*, 1993).

Ets transcription factors regulate several viral and cellular genes (Nerlov *et al.*, 1991; Wang *et al.*, 1994; Wasylyk *et al.*, 1990). It has been shown that the flanking regions of the GGA(A/T) core are responsible for the binding specificity of different members of the Ets-family (Ho *et al.*, 1990; Karim *et al.*, 1990; Nye *et al.*, 1992; Wasylyk and Wasylyk, 1992). However, as demonstrated in some cases, the binding of accessory proteins can confer such specificity to the different Ets proteins. Thus, SAP-1 and Elk are recruited by the serum response factor (SRF) at the *c-fos* serum response element (SRE) for promoter activation (Dalton and Treisman, 1992; Hipskind *et al.*, 1991; Shaw *et al.*, 1989) and NF-EM5 by Spi-1/PU1 for Igk enhancer activity (Pongubala *et al.*, 1992). In several reports (Bradford *et al.*, 1995; Dudek *et al.*, 1992; Gégonne *et al.*, 1993; Grant *et al.*, 1995; John *et al.*, 1995; Nelsen *et al.*, 1993; Wang *et al.*, 1994; Wasylyk *et al.*, 1990; Wu *et al.*, 1994) it is suggested that cooperation in the activation of several genes by the Ets-family members is due to an interaction between the Ets-members with other transcription factors. Despite some functional evidence, it still remains to be shown in several cases whether such interactions physically take place (Bradford *et al.*, 1995; Dudek *et al.*, 1992; Gégonne *et al.*, 1993; Grant *et al.*, 1995; John *et al.*, 1995; Nelsen *et al.*, 1993; Wasylyk *et al.*, 1990; Wu *et al.*, 1994).

As yet, little is known about the downstream cellular genes regulated by the Ets family of transcription factors. However, we (Buttice and Kurkinen, 1993,

1994) and others (Wasylyk *et al.*, 1991; Wasylyk and Wasylyk, 1992) have shown that c-Ets1 and c-Ets2 activate collagenase1 and stromelysin1 genes in transient transfection assays. These two genes encode enzymes that belong to a family of matrix-metalloproteinases (MMPs). MMPs are extracellular enzymes that degrade connective tissue and basement membranes and are expressed widely during growth and development (Birkedal-Hansen, 1995; Gack *et al.*, 1995). In mature adults, however, MMPs are rarely expressed except in places of disease or active tissue renewal/remodelling. For example, many MMPs are expressed in human tumors and are thought to play a critical role in several stages of tumor progression: tumor growth, tissue invasion and metastasis (MacDougall and Matrisian, 1995; Stetler-Stevenson *et al.*, 1993; Werb and Alexander, 1994).

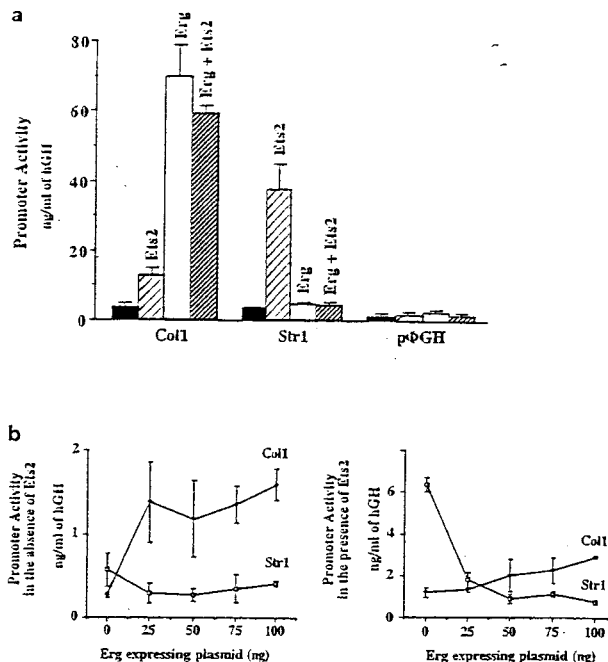
We have previously identified, by site-specific mutagenesis one functional AP1 (activator protein-1) site at -70 in the human stromelysin1 (MMP3) promoter that is necessary for basal activity and binds the Fos and Jun transcription factors (Buttice *et al.*, 1991). Stromelysin1 promoter (Quinones *et al.*, 1989) contains 13 Ets binding sites (EBS), however only two of them, that are linked tail-to-tail at -217 and -200 are functional for Ets2 activation and TPA induction (Buttice and Kurkinen, 1993). Similarly, the human collagenase1 (MMP1) promoter has been shown to contain an AP1 site necessary for TPA induction (Angel *et al.*, 1987) and Fos/Jun activation when studied in a minimal (-89 to -64) collagenase promoter context (Gutman and Wasylyk, 1990). The collagenase promoter (-610 to +61) contains also 11 GGA(A/T) core sequences that are potential binding sites for the Ets-family of transcription factors.

In this study, we show that the collagenase1 and stromelysin1 gene promoters are differentially regulated by Erg, one of the Ets members. We identify in the collagenase1 gene promoter an EBS (Ets-binding site) sequence motif at -88 for Erg transactivation. Also, we show that the Erg transactivation can operate through collagenase1 AP1 site at -73 and we present evidence for the binding of the Erg protein to the Fos/Jun complex.

## Results

### *Ets-related gene product, Erg, differentially activates the collagenase1 and stromelysin1 promoters*

The effect of the Ets-family oncoproteins on the human collagenase1 and stromelysin1 promoter activity was first studied in transient gene transfection assays. The collagenase1 (-610 to +61) or stromelysin1 (-437 to +4) promoter was linked to the human growth hormone reporter gene and co-transfected into HepG2 cells with an equal amount of plasmids expressing Ets2 or Erg. We used the human HepG2 cell line because collagenase1 and stromelysin1 genes are transcriptionally regulated in these cells (Buttice and Kurkinen, 1993 and unpublished data). As shown in Figure 1a, Erg activated collagenase1 but not stromelysin1 promoter while Ets2 activated both promoters. In control experiments, Erg and Ets2 had no effect on the promoterless reporter gene (Figure 1a). In summary, data from eleven independent experiments



**Figure 1** Erg activates collagenase1 but not stromelysin1 promoter. Erg inhibits stromelysin1 promoter activation by Ets2. (a) The human collagenase1 (-610 to +61) and stromelysin1 (-437 to +4) gene promoters were linked to the growth hormone reporter gene and studied in transiently transfected HepG2 cells ( $10^4$ ). Equal amounts of empty plasmid pSG5 (black bar) and plasmids expressing Erg, or Ets2, oncoproteins were cotransfected with the collagenase1 or stromelysin1 promoter constructs. After transfection, the cells were incubated for 24 h and the growth hormone secreted into the culture medium was quantitated by radioimmunoassay. The total amount of DNA was kept constant by adding the empty vector pSG5. In control experiments, the promoterless reporter gene (pΦGH) was similarly studied. The graph shows data from a representative experiment. The values are the average of triplicate transfections expressed as ng/ml of secreted growth hormone and error bar represents the variation between transfections. (b) Cells ( $10^4$ ) were co-transfected with collagenase1 (-◆-) or stromelysin1 promoter (-□-) constructs in the presence of increasing amounts of Erg. The vertical axis represents the promoter activity induced by Erg alone, or the promoter activity induced by Ets2 in the presence of increasing amounts of Erg. Note that the stromelysin1 promoter activity induced by 100 ng of Ets2 is repressed by low doses of plasmid expressing Erg. The bars represent the standard error from two independent experiments

reproducibly indicated that Ets2 activated at least twice better stromelysin1 than collagenase1 promoter (10-fold and 3.5-fold respectively in Figure 1a) whereas Erg was always a strong activator of collagenase1 promoter (20-fold). We also studied the effect of two other Ets members and found that Ets1 was a weak activator of collagenase1 and stromelysin1 promoters while Flt1 did not have any effect on either promoters (data not shown).

Interestingly, when Erg and Ets2 were transfected together the activation pattern of the collagenase1 and stromelysin1 promoter was very different. Ets2 did not augment nor prevent the collagenase1 promoter activation mediated by Erg (Figure 1a and b). In contrast, the activation of stromelysin1 gene by Ets2 was completely abolished with the co-expression of Erg (Figure 1a). As shown in Figure 1b, Erg repressed Ets2 induced transcription from the stromelysin1 promoter in a dose-dependent manner. Low levels of transfected Erg expression vector (25 ng) were sufficient to repress the stromelysin1 promoter activity induced by Ets2 and activate collagenase1 promoter. These results argued

against a squelching model for transcriptional repression (Ptashne, 1988) in which a small amount of protein would activate transcription but a large amount of protein would repress transcription by titrating limiting components of the transcriptional machinery. These findings suggest that Erg can act either as an activator or repressor of gene transcription depending on the promoter type.

*Erg target sites in the collagenase1 promoter*

From the transfection data it appears that Erg alone has no effect on stromelysin1 promoter activity

whereas it strongly activates the collagenase1 promoter. Therefore, we next determined which sequence motifs are involved in the Erg activation of the collagenase1 promoter. Inspection of the human collagenase1 promoter sequence (-610 to +61) revealed eleven GGA(A/T) core motifs as potential targets for the Ets-family of oncoproteins (Table 1). To characterize the GGA(A/T) target sequence(s) involved in the activation of the collagenase1 gene by Erg, we first prepared 5' end deletion mutants and studied their activity in transiently transfected HepG2 cells. Results from a representative transfection experiment are shown in Figure 2. The collagenase1 promoter (-610 to +61) and the mutant promoter (-297 to +61) were similarly activated, suggesting that the four GGA(A/T) motifs upstream of -297 are not involved in the promoter activation by Erg. By contrast, the promoter activation was decreased from 8.7-fold in the -297Coll promoter deletion to 2.9-fold in the -76Coll construct. Finally, the -37Coll promoter, still containing three GGA(A/T) core sequences and the TATA box, was not any more activated by Erg. Therefore, these data suggest that the GGA(A/T) target(s) for Erg activation of the collagenase1 promoter lies within the -297 to -76 sequence in the promoter.

To identify more precisely the Erg target motif(s) in the collagenase1 promoter, we separately mutated each of the GGA(A/T) core sequences and inactivated the putative Ets-binding sites. The two guanosines were substituted by two adenosines using site-directed mutagenesis. The activation of these mutant promo-

Table 1 Potential Ets-binding sites of the collagenase1 promoter

Potential Ets-binding sites in the collagenase1 promoter	Position relative to the transcription start site*
attctgGGATtatagg	-595
M1 attaagGGAAgccatg	-548
M2 cagaagGGAActcag	-494
tagtaaTTCCaccctc	-386
M3 caaataGGAaagaat	-256
M4 aagtgaTTCCaaataa	-151
M5 tcaagaGGATgttata	-88
M6 ctttctGGAAgggcaa	-48
M7 gggagcTTCCtagctg	-11
M8 tagctgGGATattgga	-4
M9 aggctgGGAagccatc	+20

The table shows the potential EBS and their flanking sequences. The EBS were mutated in the 5'-GGA(A/T)-3' core and the mutants named from M1 to M9. The indicated position refers to the first G in the GGA(A/T) motifs, or to the first C in the inverted motifs 5'-(T/A)TCC-3'

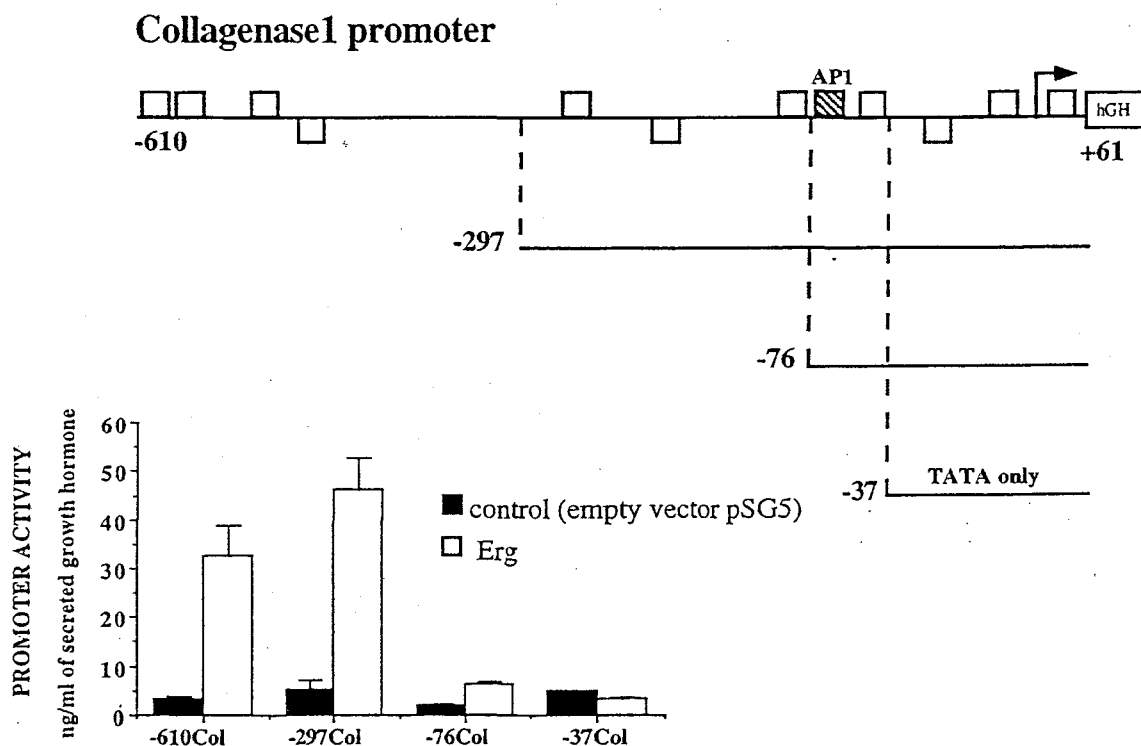
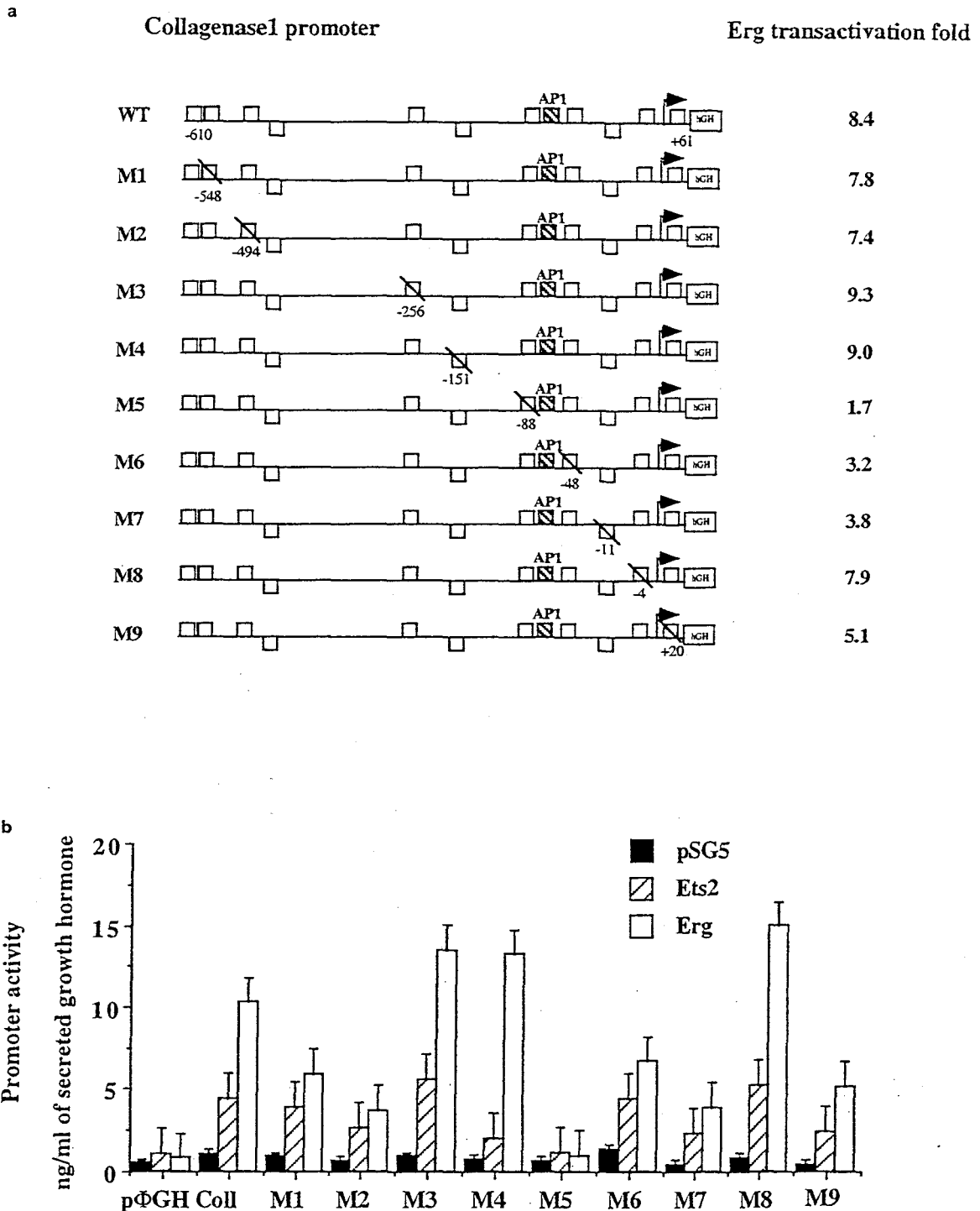


Figure 2 Deletion analysis and transactivation of the human collagenase1 gene by Erg. The collagenase1 promoter deletion mutants (-610, -297, -76, -37/+61) are indicated and a representation of the promoter shows the API site and the potential targets, GGA(A/T) core sequences, for the Ets-family members. Each promoter construct was cotransfected with a control plasmid (pSG5) or a plasmid expressing Erg. The histogram represents a typical result of five independent experiments and the bars indicate the error of quadruplicate transfections. The graph shows that Erg activates the collagenase1 promoter through a promoter region extending from -297 to -76. This region contains four potential targets for the members of the Ets family. The collagenase1 promoter mutants (-610, -297, -76, -37/+61) were transactivated by Erg oncoprotein, 9.6-, 8.7-, 2.9- and 0.7-fold respectively



ters by Erg was then compared in transient transfection assays (Table 1 and Figure 3). Mutation of the GGA(A/T) motif at position -88 (M5) almost completely abolished the promoter activation by Erg.

The average transactivation fold from several independent experiments shows that the four proximal EBS mutants M5 (-88), M6 (-48), M7 (-11) and M9 (+20) were activated by Erg respectively 1.7-, 3.2-, 3.8-



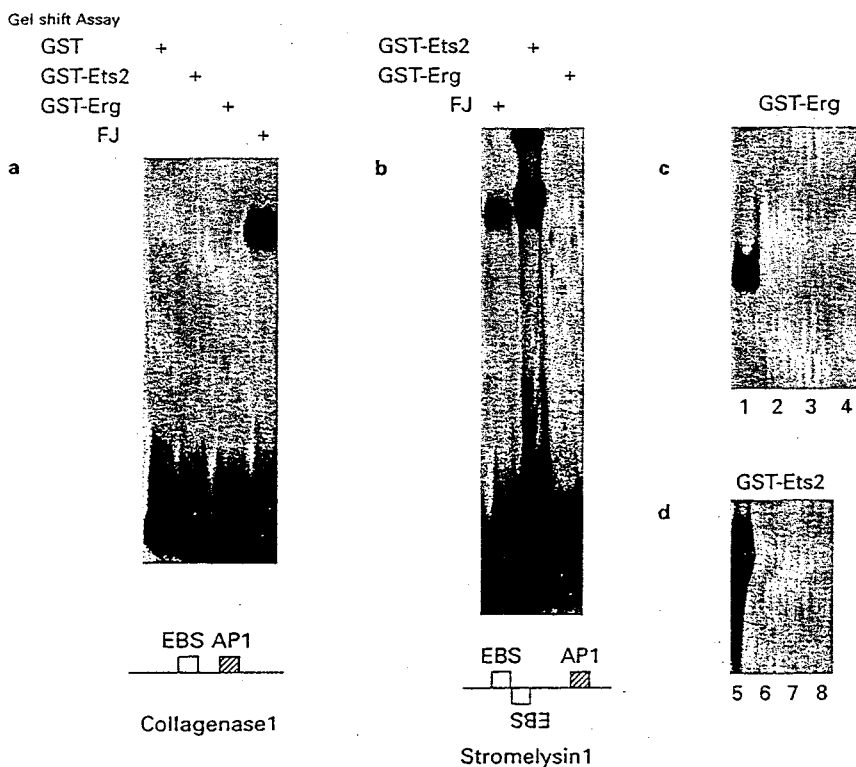
**Figure 3** Collagenase1 target motif for Erg transcriptional activation. Erg and Ets2 transactivate collagenase1 gene via the same core element. (a) The GGA(A/T) core sequences in the collagenase1 promoter were mutated to AAAA. The wild type and the mutant promoters were transfected with control or Erg expressing plasmids. Results from five independent experiments, expressed as transactivation fold, show that mutation of the distal GGA(A/T) cores (M1 to M4) does not significantly affect Erg activation of the collagenase1 promoter. In contrast, mutation in the three proximal GGA(A/T) cores M5, M6, and M7 reduces Erg transactivation. The M5 and M7 mutation reduces both the basal and Erg mediated activation, whereas the M6 mutation increases the basal activity twofold compared to the wild type expression. (b) The collagenase1 (-610 to +61) and mutant promoter constructs (M1 to M9 as in (a)) were cotransfected with a control plasmid (pSG5) or plasmids expressing Erg or Ets2. The graph shows that among the eleven potential Ets binding sites, the one at position -88 (M5) is a major motif in the Erg transactivation of the collagenase1 gene promoter. The same site is also a target for Ets2 transactivation. Results from one representative experiment are the average of two wells. Note that M3, M4 and M8 were better trans-activated by Erg compared to the wild type promoter (see also a)

and 5.1-fold compared to 8.4-fold for the wild type promoter. The M5, M7 and M9 mutation reduced both the basal and Erg mediated activation, while the M6 mutation increased the basal activity by twofold, compared to the wild type (see Figure 3b). Therefore, the activation fold by Erg of this mutant was apparently decreasing. We also noted that Ets2, although a weaker activator of collagenase1 promoter than Erg, displayed a pattern of activation of the mutants similar to that of Erg (Figure 3b). We therefore conclude that the same GGAT containing sequence at position -88 (M5) is the major target site for collagenase1 promoter activation by Erg and Ets2.

*Erg binds to the stromelysin1 promoter but not to the collagenase1 promoter*

To investigate whether the differential activation of the collagenase1 and stromelysin1 promoters by Erg was simply due to a different binding of Erg, we performed *in vitro* gel shift assays. We note that the collagenase1 and stromelysin1 promoters contain similar consensus sequence motifs such as EBS and AP1 sites. However, the number and position of the motifs differ between the two promoters. The stromelysin1 promoter has two copies (-217 and -200) of EBS, spaced by 130 nucleotides from the AP1 site (-70) (Buttice and

Kurkinen, 1993). The collagenase1 promoter contains, as demonstrated here, one functional EBS (-88) where the GGAT sequence is spaced by ten nucleotides from the AP1 motif (-73). For gel shift assays, we used promoter fragments of about 300 bp. Restriction fragments for collagenase1 (-296 to +61) and stromelysin1 (-274 to -11) were end-labelled with <sup>32</sup>P to a similar specific activity and incubated with bacteria-produced GST-fusion proteins for Ets2, Erg and/or with *in vitro* synthesized Fos and Jun. Surprisingly, the GST-Erg fusion protein bound to the stromelysin1 (Figure 4b) but not to the collagenase1 promoter (Figure 4a). The complex of the stromelysin1 promoter with Erg was very consistent but not as strong as with Ets2 (Figure 4b). Interestingly, Erg and Ets2 binding to the stromelysin1 promoter required both copies of the Ets-binding sites (Figure 4c and d). In contrast, the collagenase1 promoter did not show any detectable binding to the GST-Erg or to the GST-Ets2 fusion protein, at these conditions (Figure 4a) or even at low ionic salt conditions (data not shown). The lack of binding of the Ets-fusion proteins to the collagenase1 promoter cannot be attributed to the experimental conditions since the *in vitro* synthesized Fos and Jun bind strongly to the collagenase1 and stromelysin1 promoters (Figure 4a and b). Similarly, it seems unlikely that the lack of

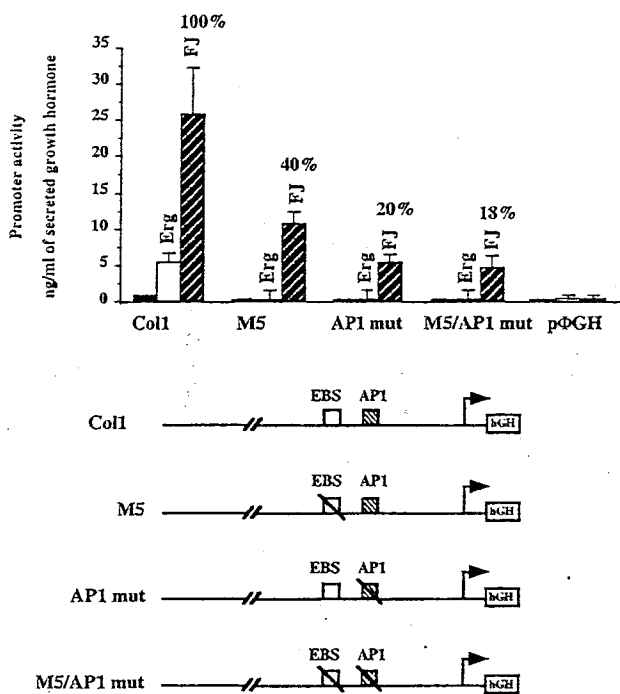


**Figure 4** Binding of Ets2, and Erg to the collagenase1 and stromelysin1 promoter fragments. (a and b) The promoter fragments of the human collagenase1 (-296 to +61) and stromelysin1 (-274 to -11) genes were end-labelled with <sup>32</sup>P and incubated with 100 ng of purified recombinant GST-Ets2, GST-Erg, or 1-2 μl of *in vitro* translated Fos and Jun proteins (FJ). The complexes were separated from the free probe by 6% polyacrylamide gel electrophoresis in Tris-glycine buffer. Note that the upper band present in the lane of the stromelysin1 promoter shifted by GST-Erg is an artefact. It is a contaminant product copurified with the GST-Erg protein. (c and d) The stromelysin1 promoter fragments (-274 to -101) corresponding to the 5' end of the (-274 to -11) fragment and containing only the two wild types EBS (lane 1 and 5), the mutated distal (lane 2 and 6) or proximal site (lane 3 and 7) and both (lane 4 and 8), as previously described (Buttice and Kurkinen, 1993), were incubated with GST-Erg (lane 1 to 4) or with GST-Ets2 (lane 5 to 8). The complexes were resolved by 6% gel in Tris-glycine buffer (c) or in 0.25 × TBE (d). Note that two intact EBS are required for the binding of Ets2 and Erg proteins to the stromelysin1 promoter. A representation of the collagenase1 and stromelysin1 promoter shows the functional EBS (empty box) and the AP1 sites (shaded box)

binding observed in the collagenase promoter could be due to the conformation of GST-fusion proteins, since they (GST-Ets2 and GST-Erg) bind to the stromelysin promoter (Figure 4b, c and d). These results suggest that the transcription factor binding to the promoter does not necessarily correlate with the transcriptional gene activation.

*The AP1 site of the collagenase promoter is required for Erg activation*

The AP1 site (-73) of the collagenase promoter is spaced by ten nucleotides from the Erg target motif. Previously, it has been suggested that the AP1 site is necessary for efficient activation by v-Ets and c-Ets1 (p68) (Wasylyk and Wasylyk, 1992). From our results, it appears that Erg activates the collagenase promoter through one specific motif at -88. However, Erg does not bind *in vitro* to this site or other sites in the collagenase promoter. To explore the role of the collagenase AP1 site in the promoter activation by Erg, we first performed transfection experiments with the wild-type collagenase promoter and the AP1

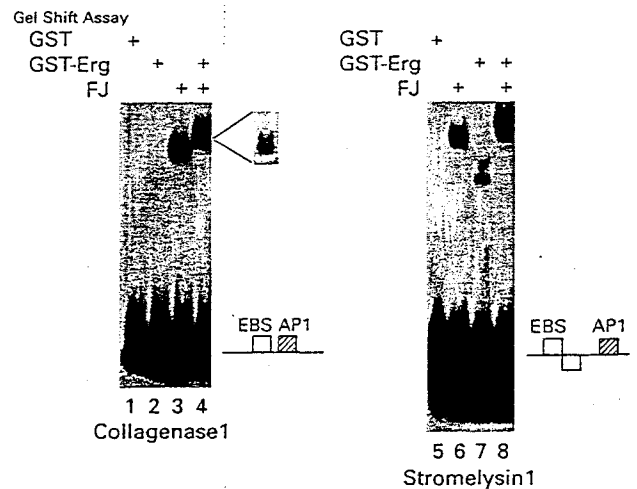


**Figure 5** The AP1 site is necessary for the collagenase promoter activation by Erg. HepG2 cells were transiently transfected with the collagenase promoter, the M5 mutant (EBS mut), or the AP1 mutant or the double mutant (EBS/AP1 mut) together with equal amount of Erg or Fos and Jun expression plasmids. Note that the EBS is only required for Erg but not for Fos/Jun activation (FJ), while mutation of the AP1 site abolishes Erg activation (see AP1 mut). The graph shows the absolute values of promoter activity with error bars from four transfections. The percentage of the Fos/Jun induced activity as compared to the wild type is indicated on top of the bars. We note that, in comparison with the Col1 promoter, the AP1 mutant showed an absolute reduction (20%) of Fos/Jun induced activity. Remark: the actual ratios of Fos/Jun-induced promoter activities divided by the basal promoter activities appear difficult to interpret since the basal levels themselves are also affected by the mutation. A schematic representation of the promoter constructs, with the functional target sites for the oncoproteins studied, is also indicated

promoter mutant. In this mutant, the first thymidine of the AP1 consensus site (TGAGTCA) was substituted by adenosine. The activation of these promoters by Erg and Fos/Jun in transfected cells was then compared. As shown in Figure 5, collagenase promoter was transactivated by both Erg and Fos/Jun. As before, mutation at -88 (M5) prevented the Erg activation, however it did not have any significant effect on the activation by Fos/Jun. Surprisingly, mutation of the AP1 site also abolished the Erg activation, indicating that the AP1 site is involved in the Erg regulation of collagenase promoter. It is of interest that the promoter AP1 mutant and the double mutant EBS/AP1 are still activated by Fos/Jun, suggesting that other sites in the collagenase promoter (-610 to +61) can function as targets for Fos/Jun. We note that, along the entire collagenase promoter studied here, there are three other putative AP1-like sites that can account for the residual Fos/Jun activation. However, the functional role of these sites need still to be explored.

*Erg interacts with the Fos/Jun complex*

The involvement of the AP1 site in the collagenase promoter activation by Erg prompted us to investigate a possible mechanism of interaction between Erg, Fos and Jun proteins. To study the role of the promoter DNA in such interactions we used gel shift assays.



**Figure 6** Erg supershifts the Fos/Jun complex formed with collagenase1 and stromelysin promoter fragments. The promoter fragments of the human collagenase1 (-296 to +61) and stromelysin1 (-274 to -11) were end-labelled with <sup>32</sup>P and incubated with 100 ng of purified recombinant GST-Erg (lanes 2 and 7) or GST alone as control (lanes 1 and 5). For the binding with Fos and Jun, 1-2 μl of *in vitro* translated products were used (lanes 3 and 6). Co-incubation of GST-Erg (100 ng) and Fos/Jun (1-2 μl) were simultaneously performed with collagenase1 or stromelysin1 probe (lanes 4 and 8). The complexes were resolved in 6% polyacrylamide gel electrophoresis in Tris-glycine buffer. The gels were dried and exposed 8 h. In control experiments Fos/Jun did not bind the collagenase AP1 mutant while it bound to the collagenase EBS mutant (data not shown). Note the shift in molecular weight of Fos/Jun complex when coincubated with GST-Erg (lanes 4 and 8), and the two bands more visible in a shorter exposure (2 h) of the same gel (small boxed picture). The upper band in lane 7 is an artefact due to a high molecular weight contaminant product copurified with the GST-Erg protein

Labelled collagenase1 probe was incubated with GST-Erg fusion protein alone, or in combination with *in vitro* synthesized Fos and Jun (Figure 6). In contrast to the stromelysin1 promoter (Figures 4 and 6), collagenase1 promoter did not bind to the GST-Erg protein alone whereas it showed a strong binding to the Fos/Jun heterodimer. In control experiments Fos/Jun did not bind the collagenase AP1 mutant while it bound to the collagenase EBS mutant (data not shown). Interestingly, co-incubation of the GST-Erg with Fos and Jun yielded two distinct complexes at the collagenase1 promoter, that both migrated with a slower mobility than the complex formed with Fos/Jun alone. Although we cannot distinguish the nature of these two complexes, one possibility is that the slower migrating complex represents a complex of Fos, Jun and Erg. As seen also before (Figure 4), no detectable complex was formed with Erg alone (Figure 6), suggesting that the Fos/Jun complex at the AP1 site may help Erg to bind to the collagenase1 promoter. Of interest is that the Erg complex at the stromelysin1 was supershifted when co-incubated with Fos/Jun, indicating that Erg may associate with Fos and Jun.

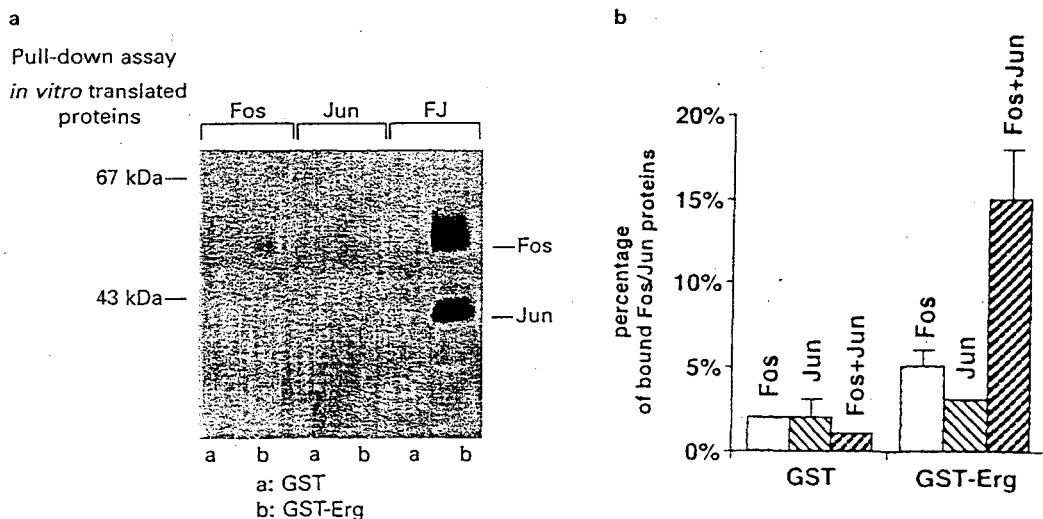
To further investigate the molecular mechanism of the AP1 site involvement in Erg activation, we hypothesized that the two motifs, AP1 at -73 and EBS at -88, could facilitate direct protein interaction of Erg and Fos/Jun. We used the *in vitro* protein interaction assay (pull-down) with bacteria-produced GST-Erg fusion protein and  $^{35}\text{S}$ -labelled *in vitro* synthesized Fos and Jun from reticulocyte lysate. The GST-Erg protein was immobilized to Sepharose beads and the bound Fos, Jun and Fos/Jun products were eluted and analysed in SDS-PAGE gel (Figure 7). The Fos or Jun protein alone gave only a faint signal (~5% and ~2.5% of the total input respectively), suggesting that they bind poorly to the Erg protein. In contrast, when Fos and Jun were incubated together, they were able to strongly bind the Erg protein

(Figure 7, last lane) by reaching 15% of the total input. As shown in Figure 7, no signal was detectable with the GST protein attached to the Sepharose beads used as a negative control. In addition, an unrelated  $^{35}\text{S}$ -labelled protein such as luciferase did not bind to the GST-Erg (data not shown), suggesting that the binding of Fos/Jun to the Erg protein is specific. These data are consistent with our gel shift studies (Figure 6) and suggest that Erg can directly interact with the Fos/Jun complex. Although the results from gel shift and pull-down experiments do not necessarily reflect the *in vivo* interactions that may occur in transient gene transfection assays, they suggest a possible regulatory mechanism, implying both protein-DNA and protein interactions.

## Discussion

Our previous studies demonstrated that the human collagenase1 and stromelysin1 promoters are similarly regulated in transfected cells by several transcription factors/oncoproteins that are genetic markers for different types of cancers (Buttice and Kurkinen, 1994; Curran and Vogt, 1992). The data described in this study represent the first example of a transcription factor/oncoprotein, Erg (an Ets-family member) that activates collagenase1 promoter but not stromelysin1 promoter. Rather, Erg completely prevents stromelysin1 promoter activation by other Ets and Fos/Jun oncoproteins (data not shown). We also provide evidence that the AP1 site is necessary for collagenase1 promoter activation by Erg, and that Erg interacts with the Fos/Jun heterodimer.

Collagenase1 and stromelysin1 genes share similar or identical promoter elements that are targets for regulation by oncogene products. The AP1 site, that binds the Fos and Jun transcription factors, is a well conserved sequence-motif in collagenase1 and stromely-



**Figure 7** Erg interacts with the Fos/Jun complex. (a) The figure shows an *in vitro* protein interaction assay (pull-down) between GST-Erg fusion proteins and  $^{35}\text{S}$ -labelled Fos and Jun proteins. GST-Erg (residues 1-462) fusion protein (125 pmol) or GST alone was immobilized on Glutathione Sepharose beads and incubated with *in vitro* translated Fos, Jun, and Fos/Jun. After extensive washing, the bound proteins were analysed by SDS-PAGE, visualized by autoradiography (3 h exposure) and the amount of bound Fos and Jun was quantitated by PhosphorImager analysis. Note the lack of binding of Fos/Jun to GST alone. The [ $^{35}\text{S}$ ]luciferase used as an unrelated protein did not bind to the GST-Erg (not shown). Molecular size markers are shown in kilodaltons. (b) The graph shows the amount of bound Fos, Jun or Fos/Jun to the GST or the GST-Erg proteins expressed in percentage of the total input (not shown). The bars indicate the standard error from two independent experiments

sin1 genes while the Ets-binding sites differ by sequence and position between the two genes. Previously, based only on sequence homology with the PEA3 element of the polyoma virus enhancer, a GGAT motif was identified in the collagenase1 promoter (-89 to -82). This motif was further studied as a multimerized sequence linked to the  $\beta$ -globin promoter and thus defined as a target for Ets1 (p68) transactivation (Gutman and Wasylyk, 1990; Wasylyk and Wasylyk, 1992). However, additional GGA(A/T) motifs with different flanking sequences are present in the human collagenase1 promoter and thus the function and role of the other GGA(A/T) core sequences in the collagenase1 gene regulation were unknown.

We demonstrate here that an Ets related gene product, Erg, activates the collagenase1 promoter *via* a GGAT target motif at position -88. Our results also suggest that, among the other eleven GGA(A/T) motifs with different flanking sequences (Table 1), only the motifs proximal to the transcription start site at -48 and -11 also contribute to the Erg regulation. Therefore, it is possible that the close proximity of these target motifs to the transcription start site reflects transcription factor interactions with the transcription apparatus. Surprisingly, the collagenase1 AP1 site at position -73 is also required for Erg activation, suggesting that the transcription factors (Fos/Jun) that bind to it might serve as a mediator of interaction between the Ets factors and the basal transcription complex. This hypothesis is consistent with our novel finding that the Erg protein physically interacts with the Fos/Jun complex *in vitro* and is in agreement with recent studies showing that Fos interacts with the TATA-binding protein (TBP) of the transcription apparatus (Metz *et al.*, 1994).

Although Erg activates collagenase1 promoter *via* the GGAT motif at -88, we could not detect in gel shift assay any binding of the Erg protein to the promoter. However, the binding may take place because of protein interaction with Fos/Jun bound to the nearby AP1 site. This notion is consistent with our gel shift data showing that in the presence of Fos/Jun, Erg supershifts the Fos/Jun complex of the promoter. The effects of Erg on the collagenase1 promoter strongly resemble that of Elk, another Ets family member, at the Fos promoter. Elk does not bind to the EBS near the serum responsive element (SRE), unless it is recruited by the serum responsive factor (SRF) at the adjacent site SRE (Shaw *et al.*, 1989). So far, this represents an interesting aspect of two distantly related Ets family member, Elk and Erg that display similar features of binding and gene activation following an interaction with a second partner, SRF for Elk and Fos/Jun for Erg.

In contrast to the collagenase1 promoter, Erg binds in gel shift assay to the stromelysin1 promoter and yet it does not activate this gene. A simple interpretation of our data would be that Erg at stromelysin1 promoter, masks sites that are targets for activators. Secondly, the sequence and the EBS arrangement in stromelysin1 promoter are significantly different from the collagenase1 promoter, suggesting that the particular tail-to-tail arrangement of the stromelysin1 pair of EBS brings together two monomers of Erg which would have inhibiting effect on promoter activity. On the other hand, we cannot exclude the possibility that Erg forms heterodimers with Ets2 at the stromelysin1 EBS pair and, as a result, the dimeric

complex would be transcriptionally inactive. This would also explain how Erg can inhibit stromelysin1 activation by Ets2, but it would imply that Erg and Ets2 must interact to form the heterodimer. Indeed, we have recently detected by gel shift and pull-down assays that in the presence of Ets2, Erg supershifts the stromelysin1 promoter fragment and that Erg physically associates with distinct domains of Ets2 (JP Basuyaux *et al.*, in preparation). Thus, although protein interactions between different Ets-family members may be required for differential gene regulation, they cannot be sufficient as regulation also depends on the promoter sequence.

Finally, the spacing between the EBS and the AP1 motifs differs between the stromelysin1 and collagenase1 genes (respectively 130 and 10 nucleotides). The 130 nucleotide stretch contains several motifs that may bind transcription factors including NF-IL6 (nuclear factor-IL6), NIP (nuclear inhibitor protein), or high mobility group (HMG) proteins that function as architectural components. Of interest is that NF-IL6 has been shown to interact with Fos and Jun *via* their leucine zippers (Hsu *et al.*, 1994) and that HMG protein can induce a sharp bend in DNA (for review, see Landsman and Bustin, 1993). In line with this, Giese *et al.* have recently provided evidence that a multiprotein complex of Ets1, PEBP2 $\alpha$  (polyomavirus enhancer binding protein), ATF/CREB (activating transcription factor/CRE-binding protein) and LEF-1 (lymphoid enhancer factor-1) formed at the minimal TCR $\alpha$  (T cell receptor  $\alpha$ ) enhancer contributes to the synergistic function of these factors (Giese *et al.*, 1995). The authors demonstrate that although Ets1 does not bind alone to its binding site, it can physically associate with PEBP2 $\alpha$  at the adjacent site. Ets1 can also contact the ATF/CREB factor at the distant site. These multiple interactions are facilitated by the presence of the regulatory lymphocyte enhancer-binding factor1 (LEF-1) that bends DNA. Therefore, it is possible that interactions of different transcription factors that are interspersed between the EBS and AP1 sites of the stromelysin1 promoter could explain why stromelysin1 gene is differently regulated from collagenase1 gene. Combinatorial protein interactions play an important part in gene regulation, since only minute changes in transcription factors can lead to synergistic activation or repression of gene expression and therefore provide precise control for transcription.

It is significant that we have now found that Erg provides a major switch between the expression of stromelysin1 and collagenase1 genes in transfected cells. Matrix metalloproteinases share homologous gene sequences and similar promoter elements. Despite their similarity, they show differential expression in response to several stimuli, probably by recruiting different sets of transcription factors. In addition, collagenase1 and stromelysin1 are enzymes with distinct substrate specificities. For instance, stromelysin1 has the widest range of substrates and can degrade almost any extracellular matrix component including type IV collagen, the main component of basement membranes. Stromelysin1 is also a potent activator of other MMPs such as collagenase1 (Murphy *et al.*, 1987), and the 92 kDa type IV procollagenase (Shapiro *et al.*, 1995) and therefore can act as a major control point and upstream regulator of the extracellular matrix degradation and

tissue remodelling. On the other hand, collagenase1 digests only fibril-forming collagens type I, II and III (Imai *et al.*, 1995), which are the most abundant proteins in the body, mainly present in bone and cartilage. In summary, our results show that collagenase is an Erg-regulated gene. Of interest is that Erg is expressed during embryogenesis in skeletal tissues. Therefore, it will be important to study and differentiate the collagenase1 and stromelysin1 gene regulation by Erg in cells *in vivo*, in the context of normal ontogenesis as well as in tumor progression.

## Materials and methods

### Constructs

The human collagenase1 promoter (-610 to +61) was cut with HaeIII, ligated to HindIII linkers and cloned into the HindIII site of the human growth hormone reporter gene, pΦGH (S Quinones and MK, unpublished data). To generate the 5' deletion mutants, this construct was linearized with ScaI, FokI or HinfI and the ends generated by FokI and HinfI were then filled-in with T7 polymerase (Pharmacia). The fragments with 5' blunt ends were released by cutting at the 3' end with XbaI and cloned into pΦGH vector which was blunted at the SallI site and sequentially digested at the XbaI site. The resulting constructs named -297Col, -76Col and -37Col contain promoter fragments extending respectively from -297, -76, -37 to +61 in respect to the +1 transcription start site.

The collagenase1 promoter mutants of the Ets binding sites were constructed by the polymerase chain reaction (PCR) (Ho *et al.*, 1989). The collagenase1 gene promoter (-610 to +61) linked to the human growth hormone reporter gene was used as template. The following primers were used: the 5' flanking primer 1 (+) (5'-AAGCTTCCTCCTGAAATCTGGGATTATAG-3') at -610 of the collagenase1 promoter sequence contained a tailed HindIII site; the 3' flanking primer 2 (-) (5'-GAGCTGTCCACAGGACCCT-3') at position 50 of the growth hormone reporter gene (pΦGH). Forward primers and the complementary reverse primers with two central AA or TT were used to mutate the GG of the GGA(A/T) core sequences. The following forward primer (5'-GTTATAAAGCAGGAGTCAGACAGCCT-3') and the complementary reverse primer (5'-AGAGGCTGTCTGACTCCTGCTTTA-TAAC-3') were used for the APl mutant. The PCR products were purified by agarose gel, sub-cloned in PCRII vector (Invitrogen) and the sequence confirmed by the dideoxy termination method. The mutant collagenase1 promoter fragments were then cloned into the HindIII site of the pΦGH reporter gene.

The stromelysin1 promoter constructs were prepared as described (Buttice and Kurkinen, 1990, 1993).

### Cell culture

Human HepG2 hepatocarcinoma cells (HB8065) from American Type Culture Collection were cultured in Dulbecco's Modified Eagle's Medium (DMEM) containing 10% fetal calf serum (FCS), streptomycin (100 µg/ml), penicillin (100 units/ml).

### Transient transfection and reporter gene assay

The human growth hormone was used as a reporter to study the regulation of the human collagenase1 and stromelysin1 genes in transiently transfected cells in culture. For the experiments, HepG2 cells were grown on 96-well microtiter plates (Buttice and Kurkinen, 1990).

100 ng of total DNA per well was transfected overnight using calcium phosphate procedure with the cell transfect kit (Pharmacia). Cultures were then rinsed twice and incubated for 24 h in DMEM containing 10% FCS plus antibiotics after which the culture medium was removed and assayed for secreted growth hormone using a solid phase radioimmunoassay kit (Nichols Institute). Experiments were carried out in triplicate or quadruplicate wells and repeated several times using two different DNA preparations. In cotransfections, the DNA consisted of equal amounts of a target gene (100 ng/well) and (100 ng/well) plasmids expressing c-Ets2, c-Erg, or c-Fos/c-Jun. To construct the c-ets2 expression plasmid (Buttice and Kurkinen, 1993), an EcoRI insert from the human c-ets2 cDNA, a kind gift of Dennis Watson (Watson *et al.*, 1988), was isolated from low melting point agarose gel and ligated to SV40 early promoter in pSVK3 vector (Pharmacia). The cloning of c-erg expression vectors (pSG5) containing respectively the human c-erg (isoform p55<sup>erb</sup>) has been previously described (Duterque-Coquillaud *et al.*, 1993). The pSV-c-jun and pBK28 expressing mouse c-Jun (Lamph *et al.*, 1988) and human c-Fos (Sassone-Corsi *et al.*, 1988), respectively, were kind gifts of Inder Verma. All DNA preparations were purified by centrifugation twice through a CsCl gradient.

### In vitro synthesis of oncoproteins

Rat cDNAs for c-fos and c-jun (Rauscher *et al.*, 1988) were kindly provided by Tom Curran. The transcription and translation of the cDNAs were simultaneously performed using the TNT coupled reticulocyte lysate system and [<sup>35</sup>S]methionine, according to the instructions of the manufacturer (Promega). The translation products were visualized in SDS-minigel and quantified by Phosphor Imager (Molecular Dynamics). The GST-Ets2, and GST-Erg fusion proteins were expressed in DH5α *E coli*, purified on Glutathione-Sepharose beads (Pharmacia) as described (JP Basuyaux *et al.*, in preparation).

### Gel shift

Restriction fragments from the collagenase1 and stromelysin1 promoters were isolated by agarose gel electrophoresis and end-labelled to similar specific activity using T7 DNA polymerase and [<sup>32</sup>P]dNTP. The binding reaction mixture (20 µl) contained 0.4 ng of probe (20 000 c.p.m.), 1-2 µl of the *in vitro* translation products Fos and Jun, or 100 ng of GST-Ets2, GST-Erg, fusion proteins and 17 µl of binding reaction buffer (12 mM HEPES pH 7.9; 4 mM Tris-HCl pH 7.9; 50 mM NaCl, 10 mM KCl; 1 mM EDTA, 1 mM DDT, 200 µg/ml bovine serum albumine, 12% glycerol, 100 µg/ml poly(dI-dC; dI-dC)). Protein-DNA complexes were resolved by 6% polyacrylamide (acrylamide/bis-acrylamide 37.5:1, Biorad) gel in 1× Tris-glycine gel shift buffer (50 mM Tris-HCl, 380 mM glycine, 2 mM EDTA) at constant current of 30 mA. Autoradiography was performed on dried gels using an extra film to quench radioactivity arising from the <sup>35</sup>S-labelled proteins.

### Acknowledgements

We thank Frédérique Dewitte for technical assistance, Drs Simon Saule and Jean Coll for critical reading of the manuscript and Nicole Devassine for patient typing. This work was supported by a grant no 1093 from the Association pour la Recherche sur le Cancer.

## References

- Angel P, Imagawa M, Chiu R, Stein B, Imbra RJ, Rahmsdorf HJ, Jonat C, Herrlich P and Karin M. (1987). *Cell*, **49**, 729–739.
- Birkedal-Hansen H. (1995). *Curr. Opin. Cell Biol.*, **7**, 728–735.
- Bradford AP, Conrad KE, Wasylyk C, Wasylyk B and Gutierrez HA. (1995). *Mol. Cell. Biol.*, **15**, 2849–2857.
- Butticè G and Kurkinen M. (1990). *Nucleic Acids Res.*, **18**, 3079.
- Butticè G and Kurkinen M. (1993). *J. Biol. Chem.*, **268**, 7196–7204.
- Butticè G and Kurkinen M. (1994). *Contrib. Nephrol.*, **107**, 101–107.
- Butticè G, Quinones S and Kurkinen M. (1991). *Nucleic Acids Res.*, **19**, 3723–3731.
- Crépieux P, Coll J and Stéhelin D. (1994). *Critic. Rev. Oncogenesis*, **5**, 615–638.
- Curran T and Vogt PK. (1992). Cold Spring Harbor Laboratory Press, NY.
- Dalton S and Treisman R. (1992). *Cell*, **68**, 597–612.
- Dhordain P, Dewitte F, Desbiens X, Stéhelin D and Duterque-Coquillaud M. (1995). *Mec. Dev.*, **50**, 17–28.
- Dudek H, Tantravahi RV, Rao VN, Reddy ESP and Reddy EP. (1992). *Proc. Natl. Acad. Sci. USA*, **89**, 1291–1295.
- Duterque-Coquillaud M, Niel C, Plaza S and Stéhelin D. (1993). *Oncogene*, **8**, 1865–1873.
- Gack S, Vallon R, Schmidt J, Grigoriadis A, Tuckermann J, Schenkel J, Weiher H, Wagner EF and Angel P. (1995). *Cell Growth Differ.*, **6**, 759–767.
- Gégonne A, Bosselut R, Bailly RA and Ghysdael J. (1993). *EMBO J.*, **12**, 1169–1178.
- Giese K, Kingsley C, Kirschner JR and Grosschedl R. (1995). *Genes Dev.*, **9**, 995–1008.
- Grant PA, Thompson CB and Pettersson S. (1995). *EMBO J.*, **14**, 4501–4513.
- Gutman A and Wasylyk B. (1990). *EMBO J.*, **9**, 2241–2246.
- Hipskind RA, Rao VN, Mueller CGF, Reddy ESP and Nordheim A. (1991). *Nature*, **354**, 531–534.
- Ho I-C, Bhat NK, Gottschalk LR, Lindsten T, Thompson CB, Papas TS and Leiden JM. (1990). *Science*, **250**, 814–818.
- Ho SN, Hunt HD, Horton RM, Pullen JK and Pease LR. (1989). *Gene*, **77**, 51–59.
- Hsu W, Kerppola TK, Chen PL, Curran T and Chen KS. (1994). *Mol. Cell. Biol.*, **14**, 268–276.
- Ichikawa H, Shimizu K, Hayashi Y and Ohki M. (1994). *Cancer Res.*, **54**, 2865–2868.
- Imai K, Yokohama Y, Nakanishi I, Ohuchi E, Fujii Y, Nakai N and Okada Y. (1995). *J. Biol. Chem.*, **270**, 6691–6697.
- Janknecht R and Nordheim A. (1993). *Biochem. Biophys. Acta*, **1155**, 346–356.
- John S, Reeves RB, Lin JX, Child R, Leiden JM, Thompson CB and Leonard WJ. (1995). *Mol. Cell. Biol.*, **15**, 1786–1796.
- Karim FD, Urness LD, Thummel CS, Klemsz MJ, McKeicher SR, Celada A, Beveren CV, Maki RA, Gunther CV, Nye JA and Graves BJ. (1990). *Genes Dev.*, **4**, 1451–1453.
- Kola I, Brookes S, Green AR, Garber R, Tymms M, Papas TS and Seth A. (1993). *Proc. Natl. Acad. Sci. USA*, **90**, 7588–7592.
- Lamph WW, Wamsley P, Sassone-Corsi P and Verma IM. (1988). *Nature*, **334**, 629–631.
- Landsman D and Bustin M. (1993). *Bioessays*, **15**, 539–546.
- Laudet V, Niel C, Duterque-Coquillaud M, Legrain D and Stéhelin D. (1993). *Biochem. Biophys. Res. Commun.*, **190**, 8–14.
- MacDougall JR and Matrisian LM. (1995). *Cancer Metastasis Rev.*, **14**, 351–362.
- Metz R, Bannister AJ, Sutherland JA, Engelmeier C, O'Rourke EC, Cook A, Bravo R and Karamidas T. (1994). *Mol. Cell. Biol.*, **14**, 6021–6029.
- Murphy G, Cockett MI, Stephens PE, Smith BJ and Docherty A. (1987). *Biochem. J.*, **248**, 265–268.
- Nelsen B, Tian G, Erman B, Gregoire J, Maki R, Graves B and Sen R. (1993). *Science*, **261**, 32–36.
- Nerlov C, Rorth P, Blasi F and Johnson M. (1991). *Oncogene*, **6**, 1583–1592.
- Nye JA, Petersen JM, Gunther CV, Jonson MD and Graves BJ. (1992). *Genes Dev.*, **6**, 975–990.
- Pongubala JMR, Nagulapalli S, Klemsz MJ, McKeicher SR, Maki RA and Atchinson ML. (1992). *Mol. Cell. Biol.*, **12**, 368–373.
- Ptashne M. (1988). *Nature*, **683**–689.
- Quinones S, Saus J, Otani Y, Harris JED and Kurkinen M. (1989). *J. Biol. Chem.*, **264**, 8339–8344.
- Rabbitts TH. (1994). *Nature*, **372**, 143–143.
- Rauscher FJ, Voulalas PJ, Franza BRJ and Curran T. (1988). *Genes Dev.*, **2**, 1687–1699.
- Sassone-Corsi P, Lamph WW, Kamps M and Verma IM. (1988). *Cell*, **54**, 553–560.
- Seth A, Watson DK, Blair DG and Papas TS. (1989). *Proc. Natl. Acad. Sci. USA*, **86**, 7833–7837.
- Shapiro S, Fliszar JC, Broekelmann TJ, Meehan RP, Senior RM and Welgus WG. (1995). *J. Biol. Chem.*, **270**, 6351–6356.
- Shaw PE, Schröter H and Nordheim A. (1989). *Cell*, **56**, 563–572.
- Sorensen PH, Lessnick SL, Lopez TD, Lin KF, Treiche TJ and Denny CT. (1994). *Nat. Genet.*, **6**, 145–151.
- Stetler-Stevenson WG, Aznavoorian S and Liotta LA. (1993). *Ann. Rev. Cell Biol.*, **9**, 541–573.
- Tymms MJ and Kola I. (1994). *Mol. Reprod. Dev.*, **39**, 208–214.
- Wang CY, Bassuk AG, Boise LH, Thompson CB, Bravo R and Leiden JM. (1994). *Mol. Cell. Biol.*, **14**, 1153–1159.
- Wasylyk B, Hahn SL and Giovane A. (1995). *Eur. J. Biochem.*, **211**, 7–18.
- Wasylyk B, Wasylyk C, Flores P, Begue A, Legrain D and Stéhelin D. (1990). *Nature*, **346**, 191–193.
- Wasylyk C, Gutman A, Nicholson R and Wasylyk B. (1991). *EMBO J.*, **10**, 1127–1134.
- Wasylyk C and Wasylyk B. (1992). *Cell Growth Differ.*, **3**, 617–625.
- Watson DK, McWilliams MJ, Lapis P, Lamminger JA, Schweinfest CW and Papas TS. (1983). *Proc. Natl. Acad. Sci. USA*, **85**, 7862–7866.
- Werb Z and Alexander CM. (1994). *Proteases and matrix degradation*. WB Saunders Co.
- Wernert N, Racs MB, Lassalle P, Dehotock MP, Gosselin B, Vandenbunder B and Stéhelin D. (1992). *Am. J. Pathol.*, **140**, 119–127.
- Wu H, Moulton K, Horval A, Parik S and Glass C. (1994). *Mol. Cell. Biol.*, **14**, 2129–2139.

## Discussion de la publication 1

La collagénase1 et la stromélysine1 appartiennent à la famille des MMP ("Matrix-MetalloProteinases"). Les MMP sont des enzymes extracellulaires capables de dégrader les composants des membranes basales et de la matrice extracellulaire des tissus conjonctifs. Elles sont impliquées dans de nombreux processus physiologiques normaux ou pathologiques, ce qui motive l'étude de la régulation de leur expression.

En tant que modèle d'étude, nous avons voulu nous intéresser en particulier à la régulation de la collagénase1 et de la stromélysine1 par les protéines Ets et le dimère Fos/Jun. En effet, une étude a démontré que les protéines Ets-1 et Ets-2 sont capables de coopérer *in vitro* avec le complexe AP-1 via un domaine de réponse aux oncogènes formé par un site EBS et un site AP-1 (Wasylyk et al., 1990). Ce motif a depuis été retrouvé et étudié dans de nombreux promoteurs dont ceux de la collagénase1 (Gutman and Wasylyk, 1990 ; Schneikert et al., 1996) et de la stromélysine1 (Wasylyk et al., 1991 ; Buttice and Kurkinen, 1993).

Dans l'article que nous vous présentons, nous proposons un modèle d'interaction protéine-protéine et une différence dans le mode d'action de deux protéines Ets.

### 1- Les protéines Ets-2 et Ergp55 diffèrent

Nous avons démontré que les protéines humaines Ergp55 et Ets-2 transactivent de manière similaire le promoteur de la collagénase1 humaine. De même, les protéines Ets-1, Ets-2 et Fli-1 activent de manière similaire l'"enhancer" de HIV-1 (Hodge et al., 1996) ou encore les protéines Ets-1 et Ergp55 ont une action comparable sur le promoteur du gène *von Willebrand* (Schwachtgen et al., 1997). Jusqu'alors, les études *in vitro* ne démontraient pas de différence significative entre les divers membres de la sous-famille ETS.

Pour la première fois, nous avons mis en évidence une différence d'activité *in vitro* sur un même promoteur entre deux membres de la famille Ets, les protéines humaines Ergp55 et Ets-2. En effet, dans cet article nous constatons que la protéine Ergp55 n'a aucune activité sur le promoteur de la stromélysine1 humaine tandis que la protéine Ets-2 en a une. De plus, Ergp55 est capable d'inhiber l'activité transactivatrice de Ets-2 sur ce même promoteur. Les mécanismes qui expliquent cet antagonisme de fonction ne sont pas encore clairs et plusieurs hypothèses peuvent être émises. Ainsi, des interactions protéiques, entre



les protéines Ets-2 et Erg elles-mêmes, avec des cofacteurs spécifiques de chacune, peuvent expliquer cet antagonisme. Une régulation par des phosphorylations est aussi envisageable.

Cet antagonisme de fonction est un premier élément qui permet d'envisager un des aspects de la spécificité de chacune des protéines Ets, tout particulièrement lorsqu'elles sont coexprimées.

## 2- Les protéines Ets et Fos/Jun coopèrent

Le promoteur de la collagénase1 humaine, contenant un site EBS fonctionnel situé 89 paires de bases en amont du site d'initiation de la transcription, a été utilisé dans des expériences de transfections transitoires avec la protéine humaine Ergp55. Cette dernière active le promoteur de la collagénase1 vingt fois. En revanche, par la technique de retard sur gel, elle est incapable de se fixer seule à ce site EBS fonctionnel. Ces résultats peuvent paraître contradictoires puisque la protéine Ergp55 est capable de transactiver le promoteur de la collagénase1 sans se fixer seule sur son site. Cependant, sa liaison l'ADN peut être facilitée par d'autres facteurs.

Des expériences de retard sur gel supplémentaires ont démontré que la protéine Ergp55 s'associe au complexe formé par les protéines Fos/Jun sur un fragment du promoteur. Ce dernier contient le site EBS étudié ainsi qu'un site AP-1 fonctionnel, dix paires de bases en aval du premier. Nous avons aussi constaté que la protéine Ergp55 interagit physiquement avec les protéines Fos/Jun. Nous pouvons en déduire que la liaison de la protéine Ergp55 sur son site nécessite la liaison préalable du dimère Fos/Jun sur l'ADN ainsi qu'une interaction directe avec le complexe AP-1. Par la suite, nous cartographierons les régions protéiques impliquées et nous déterminerons le rôle de l'ADN dans cette association Ergp55/Fos/Jun.

En fait, pour transactiver le promoteur de la collagénase1 humaine, la liaison de la protéine Ergp55 au site EBS fonctionnel est nécessaire, pour cela, elle est aidée par le complexe AP-1. Nous pouvons donc rapprocher le comportement de la protéine Ergp55 sur le promoteur de la collagénase1 à celui d'un autre membre de la famille Ets, la protéine Elk-1, sur l'élément SRE, un site EBS adjacent à un site de liaison pour le facteur SRF, du promoteur de *c-fos*. En effet, la protéine Elk-1 ne lie pas, seule, le site EBS mais est recrutée par le facteur SRF lié à son site (Hipskind et al., 1991). De manière inverse, la protéine Pu-1, membre de la famille Ets, permet le recrutement de la protéine Pip à son site de liaison sur l'"enhancer" 3' du gène des immunoglobulines  $\kappa$  (Pongubala et al., 1992 ; Eisenbeis et al., 1995 ; Perkel et al., 1998).

### 3- Les MMP sont des gènes cibles potentiels des protéines Ets

Par ces études *in vitro*, nous avons démontré que les gènes humains de la collagénase1 et de la stromélysine1 sont potentiellement des gènes cibles des protéines Ets. Cependant, pour qu'un gène soit la cible du produit d'un autre, la coexpression de ces deux gènes est nécessaire.

Les gènes de la collagénase1 et de la stromélysine1 ne sont pas parfaitement coexprimés dans le temps au cours de l'embryogenèse murine (Gack et al., 1995). En effet, le gène de la collagénase1 est exprimé à partir de E14 dans les aires d'ossification enchondrale et membranaire. Dans ces mêmes zones, aucune expression du gène de la stromélysine1 n'est détectable. Seul un facteur de transcription exprimé de manière ubiquiste et continue au cours du développement de l'embryon, s'il existe, pourrait être susceptible de réguler les promoteurs de ces deux gènes.

Le gène *ets-2* est largement exprimé dans différents tissus tout au long du développement de l'embryon (Maroulakou et al., 1994). Son produit est donc susceptible de réguler la transcription des gènes de la collagénase1 et de la stromélysine1.

En ce qui concerne le gène *erg*, son expression est très précoce (Dhordain et al., 1995) et ne correspond pas à celle des ces deux MMP (Gack et al., 1995). Cependant, le gène de la collagénase1, tout comme le gène *erg*, est exprimé dans le cartilage. L'expression du premier concerne le début de l'ossification tandis que celle du deuxième est impliquée dans la mise en place du cartilage (Gack et al., 1995 ; Dhordain et al., 1995). L'expression du gène *erg* précède donc celle du gène de la collagénase1. Une interférence entre une protéine Erg et le promoteur du gène de la collagénase1, dans ce cas, est donc peu probable.

Ainsi, *in vivo*, les protéines Ets-2 et Ergp55 n'agissent pas forcément sur la collagénase1 et la stromélysine1.

La collagénase1 et la stromélysine1 ne sont pas les seules MMP régulées par les membres de la famille Ets (Majka et al., 1997 ; Himelstein et al., 1998).

En ce qui concerne la protéine Ergp55, nos études pourraient alors se diriger vers d'autres MMP en tant que gènes cibles et non plus en tant que modèles d'étude de son mode d'action. En effet, la plupart des membres de la famille des MMP possèdent dans leur promoteur de nombreux sites EBS potentiellement fonctionnels. Parmi ces membres, la gélatinaseA, la MMP-14 ou encore la collagénase3 (Knauper et al., 1996) seraient intéressantes à étudier. Les deux premières activent par protéolyse la dernière (Knauper et al., 1996), elle-même exprimée dans le cartilage mais plus précocement que la collagénase1 (Reboul et al., 1996).

Cette étude, à défaut d'être le reflet de ce qui se passe *in vivo*, aura fourni deux modèles d'étude et nous aura permis de mettre en évidence des mécanismes moléculaires de régulation génique.

## II- Publication 2

Dans ce deuxième article, nous démontrons que les protéines isoformes Erg forment des homo- et hétérodimères *in vitro*. Ainsi, la protéine Ergp55 se dimérise avec elle-même et avec les deux autres isoformes Ergp49 et Ergp38. Grâce à la construction de mutants de délétion, nous avons défini deux domaines distincts impliqués de manière indépendante dans cette association. Le premier est localisé dans la partie amino-terminale de la protéine contenant le domaine PNT. Le second correspond au domaine ETS. Par contre, la partie centrale de la protéine se comporte comme un domaine inhibiteur de cette dimérisation.

De plus, la protéine Ergp55 est capable de s'hétérodimériser avec d'autres protéines de la famille Ets. Parmi ces dernières, les membres Fli-1, Ets-2, Er81 et Pu-1 sont capables d'interagir physiquement avec Ergp55.

Enfin, nous poursuivons l'étude du complexe ternaire Erg/Fos/Jun mise en évidence dans la première publication. Cette étude a révélé que le complexe Ergp55/Fos/Jun implique le domaine ETS et la protéine Jun tandis que celui Ergp49/Fos/Jun implique la protéine Fos.

Par ailleurs, des études de transfections transitoires avec les protéines Erg et les mutants de délétion ont permis de décrire les domaines protéiques impliqués dans l'activité transcriptionnelle d'un promoteur artificiel et de celui de la collagénase1 humaine. Nous avons défini la présence d'un domaine transactivateur dans la partie carboxy-terminale des protéines. De plus, nous avons caractérisé le domaine central comme inhibiteur de la transactivation.

Sur le promoteur de la collagénase1, les protéines Erg coopèrent avec le complexe AP-1. Nous avons démontré que le domaine ETS impliqué dans l'interaction physique entre ces protéines est suffisant pour permettre la coopération. Cependant, il existe d'autres domaines capables de moduler cette coopération, la délétion du domaine carboxy-terminal l'inhibe tandis que celle du domaine central l'augmente.

# Erg proteins, transcription factors of the Ets family, form homo, heterodimers and ternary complexes *via* two distinct domains

Séverine Carrère, Alexis Verger, Anne Flourens, Dominique Stéhelin and Martine Duterque-Coquillaud

CNRS UMR 319, Institut de Biologie de Lille, BP447, 1 rue Calmette, 59021 Lille Cedex, France

The *ets* genes family encodes a group of proteins which function as transcription factors under physiological conditions. We report here that the Erg proteins, members of the Ets family, form homo and heterodimeric complexes *in vitro*. We demonstrate that the Ergp55 protein isoform forms dimers with itself and with the two other isoforms, Ergp49 and Ergp38. Using a set of Erg protein deletion mutants, we define two distinct domains independently involved in dimerization. The first one is located in the amino-terminal part of the protein containing the *pointed* domain (PNT), conserved in a subset of Ets proteins. The second one resides within the ETS domain, the DNA-binding domain. We also show that the Erg protein central region behaves as an inhibitory domain of dimerization and its removal enhances the Ergp55 transactivation properties. Furthermore, Ergp55 forms heterodimers with some other Ets proteins. Among the latter, we show that Fli-1, Ets-2, Er81 and PU-1 physically interact with Erg. Finally, we show that the formation of the previously described ternary complex Ergp55/Fos/jun is mediated by ETS domain and Jun protein, while the ternary complex Ergp49/Fos/Jun is mediated by Fos protein.

**Keywords:** Ets family; protein interaction; transcription regulation

## Introduction

The *ets* family contains a large number of genes that encode related transcriptional regulators involved in a variety of developmental mechanisms (Crépieux *et al.*, 1994). The proto-oncogene *c-ets1* is the cellular progenitor of the *v-ets* oncogene of the avian leukemia virus E26 and the founder member of the burgeoning *ets* family (LePrince *et al.*, 1983; Nunn *et al.*, 1983). The common feature of all Ets proteins is their DNA-binding domain (85 amino acid residues), named ETS domain and structured as a winged helix–turn–helix motif (Donaldson *et al.*, 1994, 1996; Liang *et al.*, 1994; Kodandapani *et al.*, 1996). This domain can bind to the purine rich GGA(A/T) core sequence (Karim *et al.*, 1990; Wasylyk *et al.*, 1990). In addition, a second domain is conserved in a subset of Ets proteins, Ets-1, Ets-2, Erg, Fli-1, Gabpa and Tel in vertebrates and Pntp2, Elg and Yan in *Drosophila*. This region of about 65 amino acid

residues, localized in the amino-terminal part of the Ets proteins, has been referred as the *pointed* domain (PNT) (Klambt, 1993). A recent alignment reveals that it is related to a region present in *Polycomb* protein family and in cytoplasmic proteins involved in yeast sexual differentiation (Borneman *et al.*, 1996; Alkema *et al.*, 1996; Graves and Petersen, 1998), so it was named SEP (yeast Sterility, Ets-related, Polycomb proteins). It was previously described as a putative HLH domain because of a weak resemblance to the helix–loop–helix motif (Seth and Papas, 1990) but recent secondary structure data reveal that the structure does not match a HLH motif (C Slupsky, L Gentile, L Donaldson, C Mackereth, J Seitel, B Graves and L McIntosh; personal communication). In this report, we will name this conserved region PNT domain. However, in Ets family, no clear function have been associated with this domain except in Tel protein, where it is able to mediate specific oligomerization of Tel (Jousset *et al.*, 1997).

We have focused our attention on the *erg* gene, a member of the *ets* family, which encodes at least five proteins, Erg-1, Erg-2, Ergp55, Ergp49 and Ergp38 as a result of differential splicing, polyadenylation or initiation codon. These protein isoforms are weak transactivators and present similar properties of binding and transactivation (Rao *et al.*, 1987; Reddy *et al.*, 1987; Duterque-Coquillaud *et al.*, 1993). We have previously shown that Ergp55, one of the isoforms encoded by the human *erg* gene, could differentially regulate the two matrix-metalloproteinase promoters of collagenase1 and stromelysin1 (Butticè *et al.*, 1996). We found a functional antagonism between Ergp55 and Ets-2 proteins where Ergp55 does not activate the human stromelysin1 but can repress the activation by Ets-2. This antagonism can not be explained by differential DNA-binding affinity for the double EBS (Ets Binding Site) since both Ergp55 and Ets-2 bind these sites. We suggested a possible mechanism of interaction between the Ergp55 and Ets-2 proteins. By contrast to stromelysin1, Ergp55 activates human collagenase1 promoter and is able to cooperate with Fos/Jun complex *via* functional EBS and (API) sites. This cooperation involved protein interaction between Ergp55 and (API) complex.

In the present study, we show that Erg protein isoforms encoded by the human *erg* gene are able to form heterodimers with some, but not all, Ets proteins. We also demonstrate that they are able to form homodimers *via* two distinct domains. We found that the ETS domain is involved in the formation of ternary complex Erg/Fos/Jun and that this formation is directed by different (API) partners for the various Erg isoforms.

Results and discussion

Formation of dimeric complexes by Erg and Ets proteins in vitro

Our previous regulation studies on human stromelysin promoter have recently revealed that Erg is able to repress Ets-2 activation (Butticè *et al.*, 1996). One of the possible mechanisms for this antagonism might be a physical interaction between Erg and Ets-2 proteins (Basuyaux *et al.*, 1997). To test whether the Erg and other Ets proteins form homo and heterodimeric complexes *in vitro*, we performed pull-down assays between GST-Ergp55, one of the five isoforms encoded by the human *erg* gene fused to the GST protein, and various Erg derived proteins (Figure 1) or Ets family members. First, as shown in Figure 2a (lanes 6-3),

GST-Ergp55 interacted with Ergp55 itself and with the two Erg isoforms tested, Ergp49 and Ergp38, but not with luciferase (lane 10) or GST alone (lanes 11-15). In order to confirm these homodimerizations, we performed co-immunoprecipitation assays with GST-Ergp55 incubated with <sup>35</sup>S-labeled *in vitro* translated proteins. Complexes were precipitated by a GST-specific monoclonal antibody. Each protein isoform co-precipitated with GST-Ergp55 (data not shown). Second, we also showed that Fli-1 protein, the most closely Erg related member among the Ets proteins, was able to interact with GST-Ergp55 (Figure 2a, lane 9). In order to see if more distantly related proteins in the same family were also able to interact with Erg protein, we performed additional pull down assays with several other Ets proteins. Among the proteins tested, only Ets-2, Fli-1, Er81 and Pu-1 proteins

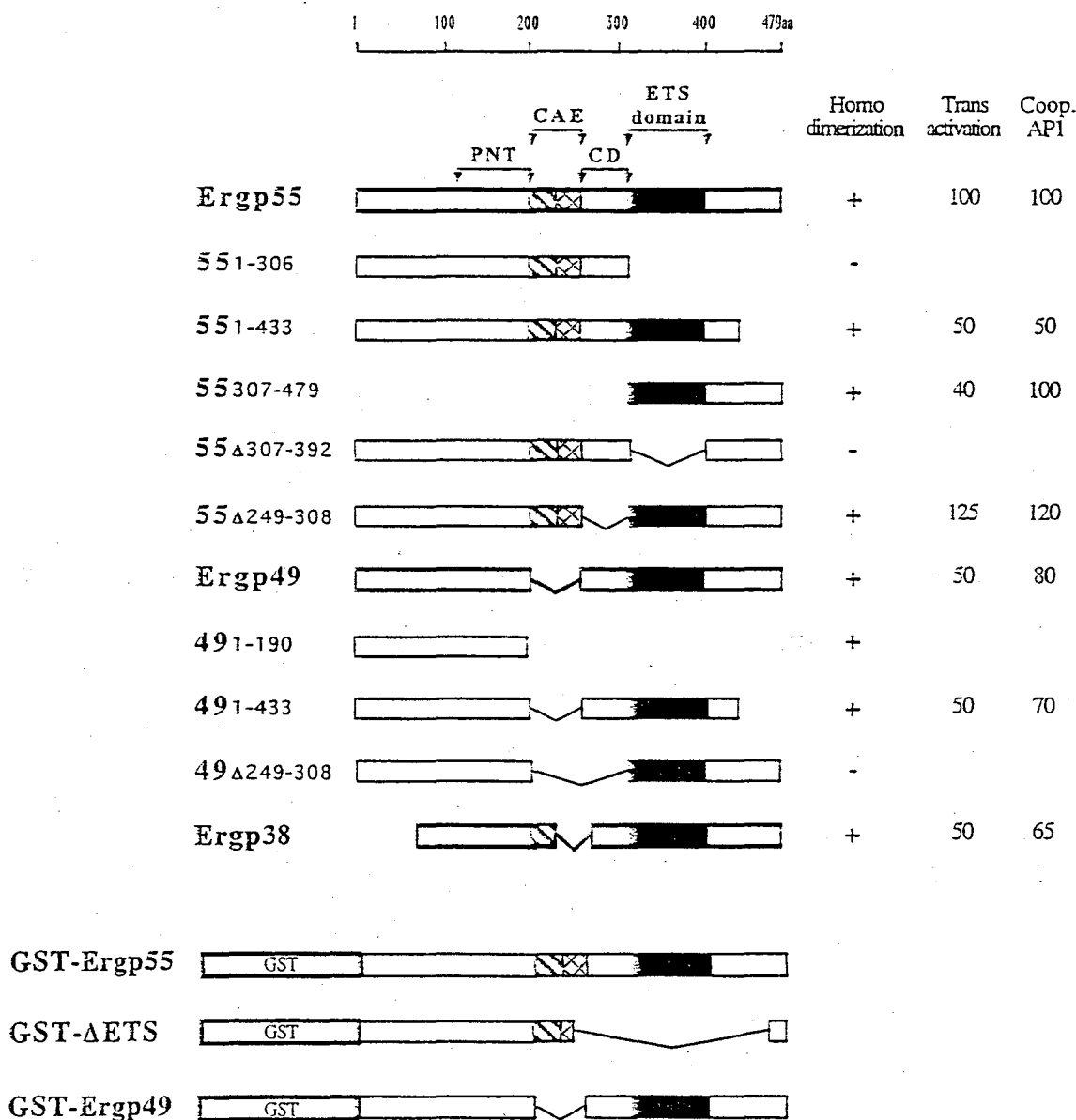


Figure 1 Schematic structure of proteins used in this study. Ergp55, Ergp49 and Ergp38 correspond to three human Erg isoforms. The ETS domain is shown as a black box and the alternative exons as hatched boxes. PNT is the amino-terminal domain conserved in Ets-1, Ets-2, Erg, Fli-1, Gabpa and Tel. CAE is the region coded by central alternative exons. CD stands for central domain. Transcriptional activation on Py enhancer elements and cooperation with AP1 complex on collagenase 1 promoter of Ergp55 is taken as 100%.

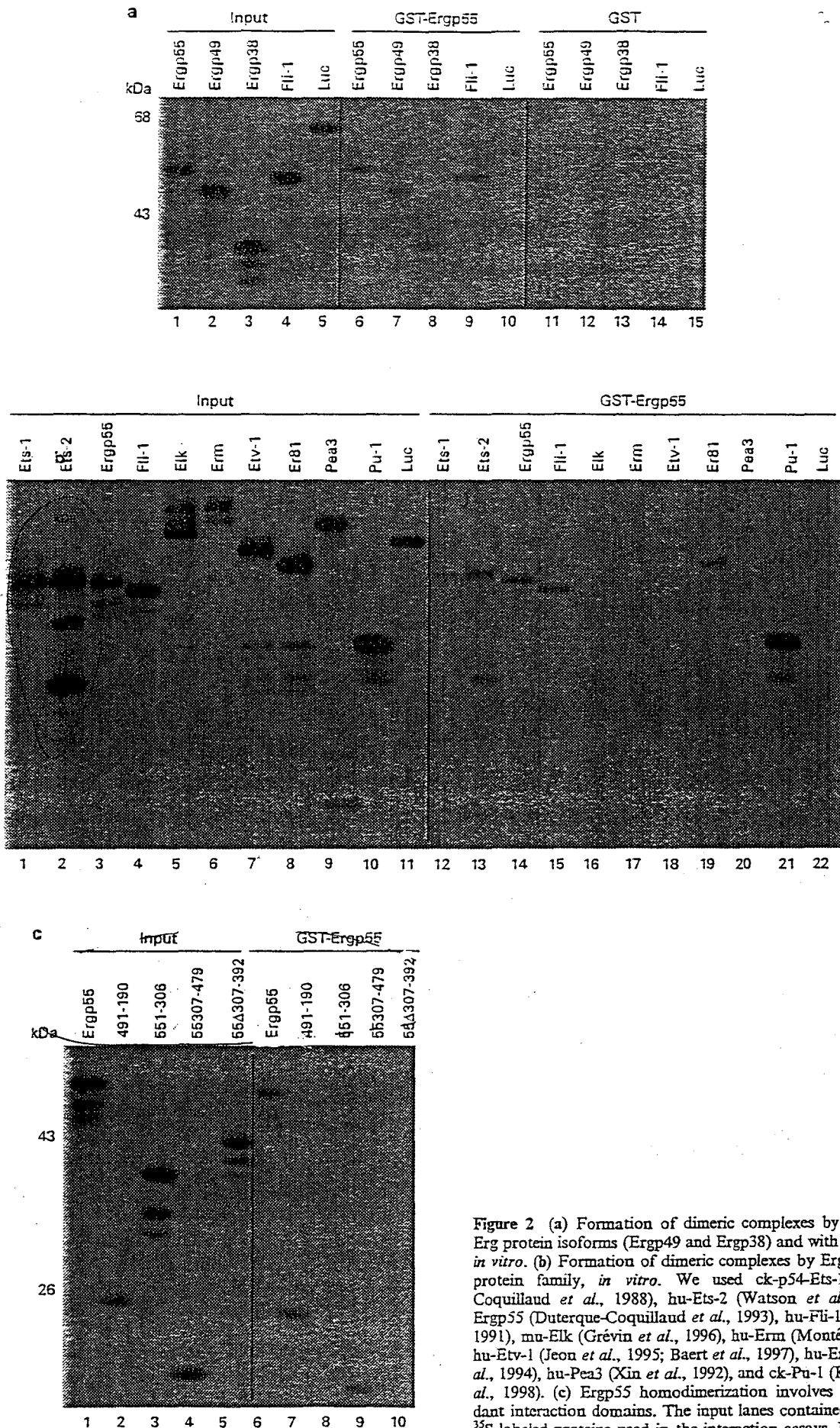


Figure 2 (a) Formation of dimeric complexes by Ergp55 with Erg protein isoforms (Ergp49 and Ergp38) and with Fli-1 protein, *in vitro*. (b) Formation of dimeric complexes by Ergp55 with Ets protein family, *in vitro*. We used ck-p54-Ets-1 (Duterque-Coquillard *et al.*, 1988), hu-Ets-2 (Watson *et al.*, 1988), hu-Ergp55 (Duterque-Coquillard *et al.*, 1993), hu-Fli-1 (Baud *et al.*, 1991), mu-Elk (Grévin *et al.*, 1996), hu-Erm (Monté *et al.*, 1994), hu-Etv-1 (Jeon *et al.*, 1995; Baert *et al.*, 1997), hu-Er81 (Monté *et al.*, 1994), hu-Pea3 (Xin *et al.*, 1992), and ck-Pu-1 (Kherrouche *et al.*, 1998). (c) Ergp55 homodimerization involves two independent interaction domains. The input lanes contained 10% of the <sup>35</sup>S-labeled proteins used in the interaction assays. Note the lack of Ets protein binding to GST alone and luciferase (Luc), as negative controls.

interacted significantly with Ergp55 (Figure 2b, lanes 13, 15, 19 and 21). The interaction with Er81 and Pu-1, which do not have the aminoterminal conserved PNT domain, demonstrated that this domain is not essential for heterodimerization. Poor or no interaction was detected with Ets-1, Elk-1, Erm, Ety-1 and Pea3 proteins (Figure 2b, lanes 12, 16-18 and 20) indicating that the ETS domain shared by all Ets proteins is not sufficient for interaction. These results suggest that unconserved domains, perhaps involved in conformational structures of these proteins, are essential for heterodimerization.

To summarize these interactions, three types of complexes were formed: (i) homodimeric complexes between individual Ergp55; (ii) homo-iso-dimeric complexes between Ergp55 and various Erg isoforms; (iii) heterodimeric complexes between Ergp55 and members of the Ets family. However, our experiments do not allow us to rule out either possible oligomerization or multimerization of Erg and Ets proteins as it has been suggested for other Ets proteins (Jousset *et al.*, 1997). Moreover, in mobility shift assays, no altered mobility of complexes were observed using various Ets Binding Sites incubated with a full length Erg protein and either other Ets proteins nor Erg protein mutants (data not shown). This observation confirms that Ets proteins recognize DNA as monomers (Nye *et al.*, 1992) and suggests that dimeric or multimeric forms are inactive.

#### Homodimerization of Erg proteins involves two distinct protein domains

In order to define the domain responsible for the homodimerization, we used a series of truncated Erg proteins for another set of experiments (Figure 1). In Figure 2c, we showed that 49<sub>1-190</sub> mutant (lane 7), which corresponds to the amino-terminal part of the Ergp55 or Ergp49 proteins, was able to bind GST-Ergp55 in a way similar to Ergp55. Moreover, the 55<sub>307-479</sub> mutant (lane 9), which corresponds to the carboxy-terminal part of Ergp55 with the ETS domain, was also able to bind GST-Ergp55. Using a Ergp55 protein mutant deleted in its ETS domain, named 55<sub>Δ307-392</sub> (lane 10), no interaction was detectable suggesting that the ETS domain is necessary for Erg proteins homodimerization. Conversely, the deletion of the carboxy-terminal part of the protein in Erg (55<sub>1-433</sub> and 49<sub>1-433</sub> mutants) did not affect the interaction with GST-Ergp55 (data not shown). We therefore concluded that the Ergp55 protein can homodimerize through two distinct and independent domains, located either in the amino-terminal part or in the carboxy-terminal part including the ETS domain. The amino-terminal part of the Ergp55 protein required for interaction is located in the 200 first residues. This region comprises the PNT domain which was recently described to be sufficient for oligomerization of Tel protein, a member of the Ets family (Jousset *et al.*, 1997). This property has been described to be specific of Tel protein. By contrast, Ets-2 protein, which contains the PNT domain, appears unable to homodimerize (Basuyaux *et al.*, 1997). Moreover, the SEP domain previously described in *Polycomb* protein family and in

cytoplasmic proteins involved in yeast sexual differentiation, related to PNT, has been shown to be involved in protein interactions (Borneman *et al.*, 1996; Gunster *et al.*, 1997).

To investigate the role of the alternative exons in dimerization, we used the 55<sub>1-306</sub> mutant, comprising the region coded by central alternative exons (CAE) and central domain (CD) of the Erg protein. The absence of interaction with GST-Ergp55 (Figure 2c, lane 8) suggested that CD (aa 249-306) could behave as an inhibitory domain for the interaction of Erg amino-terminal part with Ergp55. Moreover, CAE seemed to act in connection with this central region since 55<sub>Δ249-308</sub> interacted with GST-Ergp55 but 49<sub>Δ249-308</sub> did not (data not shown). These observations also suggest that the inhibitory effect of CD acts on dimerization through amino-terminal part of Erg protein but not when both amino and carboxy-terminal parts are present (55<sub>Δ249-308</sub>). These observations also suggest conformational differences, in the presence or absence of CAE and CD, potentially associated with distinct intramolecular interactions. The absence of both CAE and CD in 49<sub>Δ249-308</sub> mutant may drastically modify the protein conformation and prevent dimers formation.

#### Negative influence of the central domain (CD) of Erg proteins on transactivation

As an approach to the relevance of the detected interactions, we next investigated the functional consequences of Erg protein deletions on their respective transcriptional activities. The deletion mutants used in this study were analysed for their DNA-binding specificity and no significant difference was observed between them (data not shown). We excluded the 49<sub>Δ249-308</sub> mutant which failed to bind to a canonical Ets binding site. Transient transfections were performed with Cos-1 and HepG2 cell lines by using a reporter plasmid containing four copies of the artificial Py enhancer consensus element (Pye) cloned upstream of the TK promoter and a luciferase reporter gene (Figure 3a). Similar results were obtained in both cell lines. The Pye, containing EBS and API sites, has previously been shown to be regulated by Ets protein (Gutman and Wasyluk, 1990). As shown in Figure 3a, transfection of the Ergp55 expression plasmid resulted in significant transcriptional activation of the Pye reporter plasmid (20-fold) compared with transfection of the Ergp49 and Ergp38 expression plasmids (tenfold). The transactivation effect of these two isoforms was comparable with the effect induced by the two Erg protein mutants (55<sub>1-433</sub> and 49<sub>1-433</sub>) deleted in their carboxy-terminal part. Moreover, the deletion of the amino-terminal moiety of Erg proteins comprising CAE and CD (55<sub>307-479</sub>), also showed transactivation effect (about eightfold). These results confirm the presence of the carboxy-terminal transcriptional activation domain previously described (Siddique *et al.*, 1993). The CD deletion (55<sub>Δ249-308</sub>) conferred to the mutant protein a better transactivation capacity on the Pye element than Ergp55 wild type protein. These results suggest that CD which acts as an inhibitory domain on dimerization, plays an inhibitory role in transcriptional activation by Erg proteins.



*The Erg protein interacts with the AP1 complex via DNA-binding domain*

We have previously shown that Ergp55 interacted with an AP1 complex (Butticè *et al.*, 1996). In GST pull down assays, we co-incubated either Fos and Jun proteins alone or both proteins together with GST-Ergp55 protein (Figure 4). Thus, we confirmed that Ergp55 interacted principally with Fos/Jun complex (Figure 4, lane 7) and that a low interaction was detectable with Jun protein alone (lane 6) and no interaction with Fos protein alone (lane 5). Since several Ets proteins have been described to interact with AP1 complex (Bassuk and Leiden, 1995), we further investigated the involvement of the ETS domain in these interactions. As shown in Figure 4, a GST-ΔETS mutant which corresponds to GST-Ergp55 protein without a region including the ETS domain (Figure 1) was unable to interact either with Fos and Jun proteins alone (lanes 9 and 10) or with Fos/Jun complex (lane 11). In parallel, we also studied the ability of Ergp49 to interact with AP1 complex. Surprisingly, by contrast with GST-Ergp55, GST-Ergp49 was able to strongly interact not only with AP1 complex (lane 15) but also with Fos and Jun proteins alone

(lanes 13 and 14). The Ergp49 isoform which lacks the two central alternative exons (CAE) is the first Ets protein able to interact with Fos protein alone

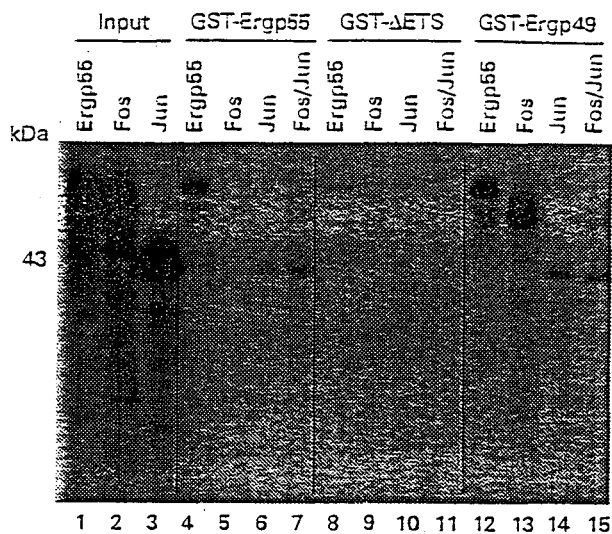


Figure 4 Formation of dimeric complexes by GST-Ergp55, GST-ΔETS and GST-Ergp49 with Fos and Jun proteins alone and with Fos/Jun complex, *in vitro*. The input lanes contained 10% of the <sup>35</sup>S-labeled proteins used in the interaction assays.

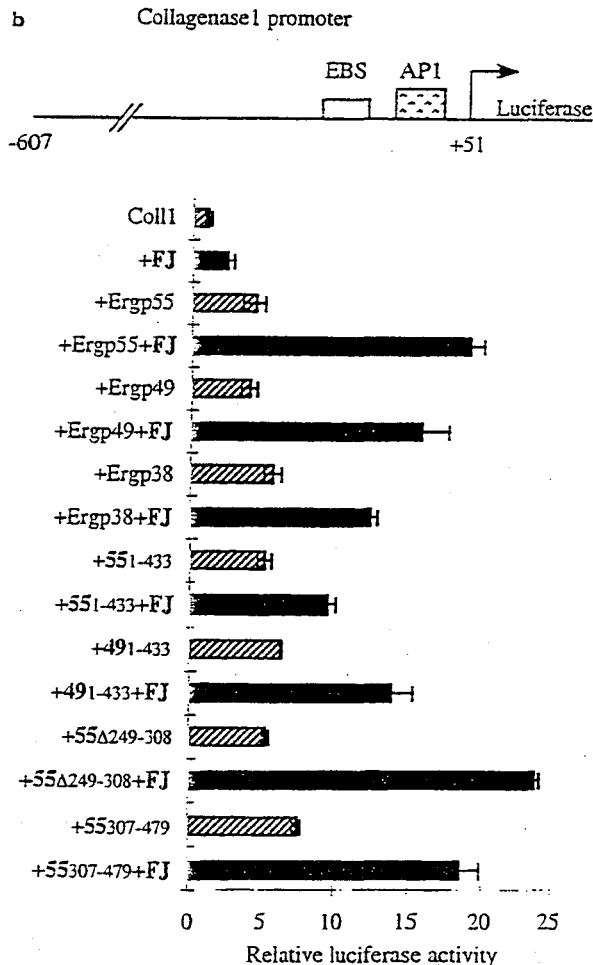
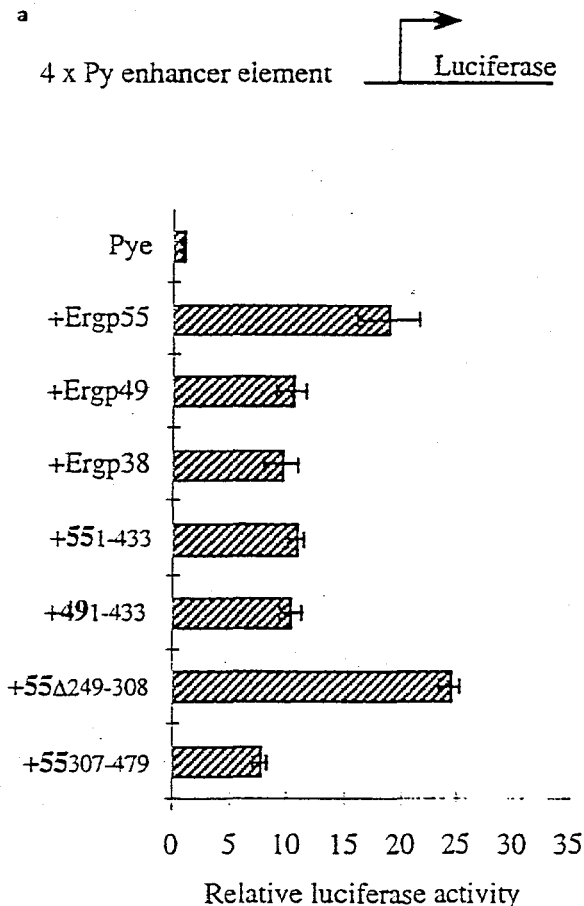


Figure 3 (a) Functional analysis of Erg isoforms and deletion mutants. (b) Erg proteins cooperate with Fos/Jun complex. Transient transfection assays were performed in Cos-1 cells by using the reporter and expression plasmids shown. Data are expressed as fold luciferase activity (as detailed in Materials and methods); representative assays are shown with bars indicating the error between triplicates.

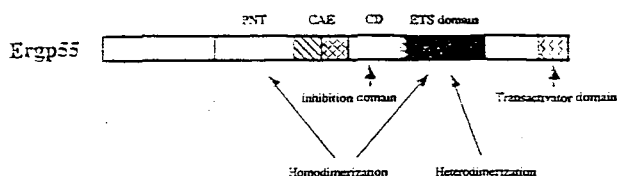


Figure 5 Schematic representation of the functional domains of the human Ergp55 protein and their involvement in dimerization and transactivation properties. PNT: amino-terminal conserved domain. CAE: domain coded by central alternative exons. CD: central domain.

since all the Ets proteins studied so far interact with Jun family proteins (Bassuk and Leiden, 1995). This result constituted the first functional difference observed between Erg protein isoforms which usually exhibit similar properties.

In conclusion, the ETS domain of Ergp55 protein appears involved in the formation of a ternary complex with Fos/Jun. Thus this observation taking together with the mapping of the Erg domains responsible for the homodimerization shows that the ETS domain serves a key dual function, DNA-binding and in protein-protein interactions. It is also the case for other transcriptional factors, for example, the Runt family proteins are able to physically interact with other factors *via* their RUNT domain, also responsible of DNA-binding properties (Ogawa *et al.*, 1993; Giese *et al.*, 1995).

#### The carboxy-terminal part of Erg proteins, comprising the ETS domain, is sufficient to mediate Erg/API cooperation

We and other have previously shown that Ets proteins and API complex can cooperatively activate transcription from a reporter construct containing a promoter or an enhancer element (Wasylyk *et al.*, 1990; Gutman and Wasylyk, 1990; Buttice *et al.*, 1996). Human collagenase1 promoter contains adjacent EBS and API sites, both of which are required for this cooperative transcriptional regulation (Buttice *et al.*, 1996). To assess directly the functional role of the physical interaction between Erg proteins and API complex, we performed a series of transient transfection assays in which we compared the ability of the wild-type Erg isoforms and their deletion mutants to activate collagenase1 promoter transcription cooperatively with Fos and Jun proteins. We transfected Cos-1 cells with a luciferase reporter construct containing collagenase1 promoter in conjunction with various combinations of Erg and Fos/Jun proteins (Figure 3b). Overexpression of the three Erg isoforms tested alone resulted in similar transactivation effect, respectively five-, four- and sixfold. Overexpression of Fos/Jun proteins alone resulted in threefold transactivation effect. Co-transfection of Ergp55, Ergp49, Ergp38 and Fos/Jun proteins resulted in 20-, 16- and 13-fold transactivation respectively which revealed a cooperative effect of these factors. Similar results were described with others Ets proteins family on an artificial promoter (Bassuk and Leiden, 1995). Co-transfection of Fos/Jun proteins with the 55<sub>307-499</sub> mutant, which corresponds to the carboxy-terminal

part of Erg proteins comprising the ETS domain, resulted in a 18-fold transactivation. This result demonstrated that this part of Erg proteins, which is involved in physical interaction between Erg and Fos/Jun proteins, is sufficient to mediate Erg/API cooperation. However, co-transfection of various Erg mutants with Fos/Jun indicated that cooperative transactivation by these factors was modulated by other domains without direct implication in protein interaction. First, the absence of the carboxyterminal part of Ergp55 and Ergp49 (55<sub>1-433</sub> and 49<sub>1-433</sub> mutants) resulted in a reduction of their cooperative effect with API complex (ten- and 14-fold). Second, CD acted as an inhibitory domain of cooperation since co-transfection with 55<sub>Δ249-308</sub> mutant resulted in a better cooperative effect than with the wild-type Ergp55 (24-fold), these results are in favor of conformation influence of the Erg isoforms where CD is possibly involved in specific protein interactions. In Figure 6, we summarize the functions of the Ergp55 domains defined in this paper.

#### Concluding remarks

Based on these observations, we can envisage several models concerning the functional consequences of Erg heterodimerization with Ets proteins. First, formation of heterodimeric complexes between members of the same family could be a mean to regulate their transcriptional activity. For example, Erg/Ets-2 dimer formation may prevent Ets-2 protein from acting as transcriptional activator, as it was previously shown on stromelysin1 promoter (Buttice *et al.*, 1996; Basuyaux *et al.*, 1997). In this case, monomeric proteins would be functionally active while heterodimeric complexes would be inactive by masking their transactivation domains or modulating their availability to accessory protein factors or other transcription factors like API. Second, dimerization may prevent Ets proteins from binding to genomic DNA target sites since Ets proteins have always been described to bind in monomer form (Crépieux *et al.*, 1994). Therefore, these two possibilities are not mutually exclusive but could be complementary. Moreover, various conditions would necessarily modulate these regulatory interactions. One of them is the competition for homodimerization *versus* heterodimerization which depends on the relative intracellular concentrations of Ets proteins. Another one concerns the accessibility and the affinity of DNA target sites for maintaining the appropriate balance of active monomers relative to inactive dimers.

In other respects, we and others have previously shown that *ets-1*, *erg* and *fli-1* are co-expressed during chicken or mouse embryo development, particularly in mesenchyme, blood vessel formation and neural crest cell migration (Vandenbunder *et al.*, 1989; Dhordain *et al.*, 1995; Mélet *et al.*, 1996). Particularly, *ets-2* and *erg* are co-expressed in cartilaginous tissue during chicken and mouse embryonic development (Maroulakou *et al.*, 1994; Dhordain *et al.*, 1995). Based both on their overlapping spatiotemporal expression patterns during embryogenesis and on our biochemical data, we propose that the Ets proteins ability to heterodimerize to each other provides a regulatory mechanism by which Ets proteins mutually regulate their transcriptional actions *in vivo*. Similar regulation models have

References

- Alkema MJ, Bronk M, Verhoeven E, Otte A, van't Veer LJ, Berns A and van Lohuizen M. (1996). *Genes & Dev.*, 11, 226-240.
- Baert JL, Monté D, Musgrove EA, Albagli O, Sutherland RL and de Launoit Y. (1997). *Int. J. Cancer*, 70, 590-597.
- Bassuk AG and Leiden JM. (1995). *Immunity*, 3, 223-237.
- Basuyaux JP, Ferreira E, Stéhelin D and Buttice G. (1997). *J. Biol. Chem.* in press.
- Baud V, Lipinski M, Rassart E, Poliquin L and Bergeron D. (1991). *Genomics*, 11, 223-224.
- Borneman D, Miller E and Simon J. (1996). *Development*, 122, 1621-1630.
- Butticé G, Duterque-Coquillaud M, Basuyaux JP, Carrère S, Kurkinen M and Stéhelin D. (1996). *Oncogene*, 13, 2297-2306.
- Crépieux P, Coll J and Stéhelin D. (1994). *Crit. Rev. Onc.*, 5, 615-638.
- Dhordain P, Dewitte F, Desbiens X, Stéhelin D and Duterque-Coquillaud M. (1995). *Mech. Dev.*, 50, 17-23.
- Donaldson LW, Petersen JM, Graves BJ and McIntosh LP. (1994). *Biochemistry*, 33, 13509-13516.
- Donaldson LW, Petersen JM, Graves BJ and McIntosh LP. (1996). *EMBO J.*, 15, 125-134.
- Duterque-Coquillaud M, Leprince D, Flourens A, Henry C, Ghysdael J, Debuire B and Stéhelin D. (1988). *Oncogene Res.*, 2, 335-344.
- Duterque-Coquillaud M, Nié C, Plaza S and Stéhelin D. (1993). *Oncogene*, 8, 1865-1873.
- Giese K, Kingsley C, Kirshner JR and Grosschedl R. (1995). *Genes & Dev.*, 9, 995-1008.
- Graves BJ and Petersen JM. (1998). *Adv. Cancer Res.*, 75, in press.
- Grévin D, Ung S, Denhez F, Dehem M, Quatennens B, Bègue A, Stéhelin D and Martin P. (1996). *Gene*, 174, 185-188.
- Gunster MJ, Satijn DPE, Hamer KM, den Blaauwen JL, de Bruijn D, Alkema MJ, van Lohuizen M, van Driel R and Otte AP. (1997). *Mol. Cell. Biol.*, 17, 17-23.
- Gutman A and Wasylyk B. (1990). *EMBO J.*, 9, 2241-2246.
- Jeon IS, Davis JN, Braun BS, Sublett JE, Rousset M17, Denny CT and Shapiro DN. (1995). *Oncogene*, 10, 1229-1234.
- Jousset C, Carron C, Boureux A, Tran Quang C, Oury C, Dusanter FI, Charon M, Levin J, Bernard O and Ghysdael J. (1997). *EMBO J.*, 16, 69-82.
- Karim FD, Urness LD, Thummel CS, Klemsz MJ, McKercher SR, Celada A, Van Beveren C, Maki C, Gunther RA, Nye JA and Graves B. (1990). *Genes & Dev.*, 4, 1451-1453.
- Kherrouche Z, Beuscart A, Huguet C, Flourens A, Moreau-Gachelin F, Stéhelin D and Coll J. (1998). *Oncogene*, 17, 163-176.
- Klambt C. (1993). *Development*, 117, 163-176.
- Kodandapani R, Pio F, Ni CZ, Picialli G, Klemsz M, McKercher S, Maki RA and Ely KR. (1996). *Nature*, 380, 456-460.
- Leprince D, Gégonne A, Coll J, De Taisne C, Schmeisberger A, Lagrou C and Stéhelin D. (1983). *Nature*, 306, 395-397.
- Liang H, Olejniczak ET, Mao X, Nettesheim DG, Yu L, Thompson CB and Fezik SW. (1994). *Proc. Natl. Acad. Sci. USA*, 91, 11655-11659.
- Maroulakou IG, Papas TS and Green JE. (1994). *Oncogene*, 9, 1551-1565.
- Mélet F, Motro B, Rossi DJ, Zhang L and Bernstein A. (1996). *Mol. Cell. Biol.*, 16, 2708-2718.
- Monté D, Baert J-L, Defossez P-A, De Launoit Y and Stéhelin D. (1994). *Oncogene*, 9, 1397-1406.
- Nunn MF, Seeburg PH, Moscovici C and Duesberg PH. (1983). *Nature*, 306, 391-395.
- Nye JA, Petersen JM, Gunther CV, Jonsen MD and J. GB. (1992). *Genes & Dev.*, 6, 975-990.
- Ogawa E, Maruyama M, Kagoshima H, Inuzuka M, Lu J, Satake M, Shigeseda K and Ito Y. (1993). *Proc. Natl. Acad. Sci. USA*, 90, 6859-6863.
- Rao VN, Papas TS and Reddy ES. (1987). *Science*, 237, 635-639.
- Reddy ES, Rao VN and Papas TS. (1987). *Proc. Natl. Acad. Sci. USA*, 84, 6131-6135.
- Seth A and Papas TS. (1990). *Oncogene*, 5, 1761-1767.
- Siddique HR, Rao VN, Lee L and Reddy ESP. (1993). *Oncogene*, 3, 1751-1755.
- Taggart AKP and Pugh BF. (1996). *Science*, 272, 1331-1333.
- Vandenbunder B, Pardanaud L, Jaffredo T, Mirabel MA and Stéhelin D. (1989). *Development*, 106, 265-274.
- Wasylyk B, Wasylyk C, Flores P, Bègue A, Leprince D and Stéhelin D. (1990). *Nature*, 346, 191-193.
- Watson DK, McWilliams MJ, Lapis P, Lautenberger JA, Schweinfest CW and Papas TS. (1988). *Proc. Natl. Acad. Sci. USA*, 85, 7862-7866.
- Xin JH, Cowie A, Lachance P and Hassell JA. (1992). *Genes & Dev.*, 6, 481-496.
- Zhang H, Hu G, Wang H, Scivolino P, Iler N, Shen MM and Abate-Shen C. (1997). *Mol. Cell. Biol.*, 17, 2920-2932.

been proposed in various transcription factor families. Indeed, for example, Msx and Dlx homeoproteins, which exhibit overlapping expression patterns, form heterodimers with respect to their functional antagonism (Zhang *et al.*, 1997). Moreover, TBP dimerization have been shown to inhibit DNA-binding and, consequently, to provide a mode of regulating TBP-DNA interaction (Taggart and Pugh, 1996). The similarity of these observations and our results raises the possibility that formation of inactive homo or heterodimers could be a general mode of regulating transcription factors activity *in vivo*.

## Materials and methods

### Plasmid constructs

In order to construct recombinant pGEX-2TK (Pharmacia), the cDNA encoding Ergp55 and Ergp49 were amplified by PCR, using pSG5-HuErgp55 and pSG5-HuErgp49 as template (Duterque-Coquillaud *et al.*, 1993), with the 5' amplicon 5'-GAGGATCCCGGGTATGGC-CAGCACTATAAG and the 3' amplicon 5'-TTGAGTCCCGGGCTTTAGTAGTAAGTGC. Each amplicon contained *Sma*I sites, such that the resulting PCR products first subcloned into pCRII (Invitrogen), and could be ligated, upon digestion with *Sma*I, in frame into the unique *Sma*I site present within the pGEX-2TK polylinker. The recombinant pGEX-2TK containing cDNA encoding Ergp55 was digested by *Nde*I restriction enzyme and religated to delete the ETS domain coding region.

Deletion mutants were obtained by PCR amplification and subcloned into pCRII. Mutants 55<sub>1-306</sub> and 49<sub>1-306</sub> were cloned via *Eco*RI sites into pSG5. Template was Ergp55 cDNA for all 55<sub>x,y</sub> mutants and Ergp49 cDNA for all 49<sub>x,y</sub> mutants. The 5' amplicon was 5'-AGACAGATCTCAATAAACTTGAT-TGC and the 3' amplicon was, 5'-ACTTCATGGATTG-CAAGGCGCTACTTG for 55<sub>1-306</sub> mutant and 5'-CCTGGGGGCTCATATGGTTAAT for 49<sub>1-306</sub> mutant. The 55<sub>1-433</sub>, 49<sub>1-433</sub>, 55<sub>Δ307-392</sub>, 55<sub>Δ249-308</sub> and 49<sub>Δ249-308</sub> mutants were cloned via *Bgl*III sites into pSG5. The 5' amplicon was still the same. For 55<sub>1-433</sub> and 49<sub>1-433</sub>, the 3' amplicon was 5'-AGATCTACAAAGTTCATCTTCTGT. For 55<sub>Δ307-392</sub>, 55<sub>Δ249-308</sub> and 49<sub>Δ249-308</sub>, the 3' amplicon was 5'-CTAGATCTCTCCCGGCTTCC; the internal amplicons were 5'-GCCGCTTGCAAATCCAGACTCCAGGGGATCGCCAG-GCC for 55<sub>Δ307-392</sub> 5'-CTGGATCTGGCCACTGCCTGGC-CTAGTTGTAATTCT for 55<sub>Δ249-308</sub> and 5'-CTGGATCTG-GCCA CTGCCCTCTGAGGTAGTGGAG for 49<sub>Δ249-308</sub>. Mutant 55<sub>307-479</sub> was cloned via *Eco*RI sites into pSG5. The 5' amplicon was 5'-AGATCTGCATTATGGCCGGCAGT-GGCCAGTCCAGC and the 3' amplicon was 5'-CTA-GATCTCTTCCCGGCTTCC.

### In vitro proteins synthesis

The *in vitro* T7 driver transcription and translation of the cDNAs were simultaneously performed using TNT coupled reticulocyte lysate system and [<sup>35</sup>S]methionine, according to the instructions of manufacturer (Promega). The translation products were visualized in SDS-PAGE and quantified by Phosphor Imager (Molecular Dynamics).

GST fusion proteins were isolated from *E. coli* XL1blue bacteria containing the appropriate expression vectors. 20 ml of an overnight culture was inoculated into 1 l of 2XYT and incubated at 37°C until an absorption of 1 at 600 nm was reached. Isopropyl-D-galactopyranoside (Boehringer) was added to 1 mM final concentration and the culture was incubated for additional 5 h. After centrifugation, bacterial pellet was suspended in 10 ml PBS-Triton X-100 1%

containing a quarter of a protease inhibitors cocktail tablet (Boehringer) and lysed by passing through a French press twice. After centrifugation at 10 000 g for 10 min at 4°C, DNaseI (Boehringer) was added to 1 mM final concentration to the supernatant containing soluble proteins. GST fusion proteins were then purified on glutathione-Sepharose beads column (Pharmacia) in accordance with the manufacturer's instructions. The concentrations and purities of the GST fusion proteins were estimated by SDS-PAGE, followed by Coomassie blue staining.

### Pull-down assays

50 μl of Glutathione Sepharose 4B beads (Pharmacia) were incubated with 1 μg of either GST or GST-Ergp55 or GST-Ergp49 or GST-ΔETS fusion proteins in NETN (20 mM Tris HCl pH 8, 100 mM NaCl, 1 mM EDTA, 0.5% NP40), for 1 h at 4°C. Beads were washed three times with incubation buffer (12 mM HEPES pH 7.9, 4 mM Tris HCl pH 7.9, 50 mM NaCl, 10 mM KCl, 1 mM EDTA) and resuspended in 30 μl of a mixture containing <sup>35</sup>S-labeled protein expressed in reticulocyte lysate and incubation buffer. After 1 h at 4°C, beads were washed five times with NETN and mixed with SDS-sample buffer. The bound proteins were analysed by SDS-PAGE followed by autoradiography.

### Transactivation assays

Human HepG2 hepatocarcinoma cells (H8065) from American Type Culture Collection and Cos-1 cells were cultured for 24 h in DMEM (Life Technologies Inc.) containing 10% fetal calf serum (FCS), gentamycin (100 units/ml). They were rinsed and incubated in 1 ml of OptiMEM (Life Technologies, Inc.) with a mixture of DNA (maximal amount 1 μg) and Poly Ethylene Imine (4 μl, Euromedex). In each experiment, plates were incubated with the same total plasmid amount (1 μg) completed as necessary with the corresponding empty vector. Cells were incubated overnight, harvested and disrupted in lysis buffer (Promega). The supernatant was assayed for protein content (Bio-Rad protein assay) and processed for luciferase assay. Reaction mixtures, containing 20 μl of cell extract and 80 μl of luciferase assay buffer (25 mM Tris phosphate pH 7.8, 8 mM MgCl<sub>2</sub>, 1 mM DTT, Triton X-100 1% v/v) were measured in a Berthold Lumat LB9501 luminometer following injection of luciferine assay solution (0.25 mM luciferine and 2 mM ATP in luciferase assay buffer). Triplicate samples were performed in two or three experiments, and the values obtained were verified by normalization to the proteins content of each cell extract. In addition, we used pSV2-Luc plasmid as positive control of transfection. For each protein expression vector transiently transfected in the two cell lines used in this study, their expression level was shown similar by immunoprecipitating labeled protein extracts with Erg specific polyclonal antibodies (Dhordain *et al.*, 1995).

### Acknowledgements

The authors thanks J. Baert, G. Buttice, J. Coll, Y. de Launoit, D. Grévin, M. Lipinski and I. Verma for reagents; F. De Witte for technical advises; AM. Ayrat, J. Coll, P. Crépieux and D. Régnier for critical reading of the manuscript. We thank B. Graves and L. McIntosh for communicating results prior to publication. SC and AV are predoctoral fellows supported by the Institut Pasteur de Lille-Région Nord-Pas-de-Calais and by Ministère de l'Éducation Supérieure et de la Recherche. This work was supported by grants from CNRS and ARC (2058).

## Discussion de la publication 2

Dans l'article précédent, nous avons mis en évidence un antagonisme de fonction entre les protéines Ergp55 et Ets-2 sur le promoteur de la stromélysine1. Cet antagonisme de fonction ne peut-être du à une différence d'affinité des deux protéines vis-à-vis de l'ADN puisqu'elles se fixent de manière similaire au double site EBS. Une des explications de cet antagonisme pourrait donc être une interaction physique entre ces deux facteurs de transcription. En effet, nous avons mis en évidence pour la première fois des interactions protéine-protéine à l'intérieur même de la famille Ets. Ainsi, dans cet article, nous décrivons la particularité de la protéine Ergp55 d'interagir *in vitro* avec des membres de la famille Ets mais pas avec tous. Depuis, aucun autre exemple n'a été défini dans la littérature, mais nous pouvons imaginer que d'autres hétérodimères qui n'impliquent pas les protéines Erg existent dans la famille.

De plus, nous démontrons la capacité des protéines Erg à former des dimères : i/ des homodimères composés de protéines Ergp55 ; ii/ des homo-iso-dimères entre les protéines isoformes Ergp55, Ergp49 et Ergp38 ; iii/ des hétérodimères entre la protéine Ergp55 et divers membres de la famille Ets.

Pour toutes ces interactions protéiques étudiées, nous avons suggéré une dimérisation. En fait, nous ne pouvons pas exclure l'hypothèse d'une oligomérisation, c'est-à-dire l'association d'au moins deux protéines, ou même d'une multimérisation. En effet, par les techniques de co-rétention protéique ou de co-immuno-précipitation utilisées, nous ne pouvons déterminer le nombre de protéines impliquées dans l'association.

### 1- Les protéines Erg se fixent en monomères

Les protéines Erg sont capables de s'homo- et s'hétérodimériser. Afin de comprendre si cette association requiert la présence d'ADN, nous avons effectué diverses études. Nous avons effectué des co-rétentions protéiques, expériences au cours desquelles nous avons ajouté ou non une séquence d'ADN cible, en absence ou en présence de bromure d'éthidium. Les résultats n'ont pas mis en évidence une augmentation ou une diminution des dimères (résultats non publiés). Ces derniers ne nécessitent donc pas la présence d'ADN et il ne semble pas y avoir de compétition entre la formation de dimères et la liaison à l'ADN. De plus, des études de retards sur gel qui associent la protéine sauvage à divers mutants de délétion des protéines Erg n'ont jamais montré des retards intermédiaires témoins d'associations entre mutants de délétions.

La formation de dimères et la fixation à une séquence cible semblent donc deux événements distincts pour les protéines Erg. Comme la plupart des membres de la famille Ets décrits dans l'introduction de ce mémoire, elles se fixent donc en monomères sur leur séquence cible.

La fixation des protéines Ets sur l'ADN a été étudié sur de simples sites EBS mais il existe des doubles sites. Dans le premier article, nous avons étudié la fixation des protéines Ets-2 et Ergp55 sur le double site EBS situé sur le promoteur de la stromélysine1 humaine. Nous avons pu comparer nos résultats avec ceux d'études sur d'autres promoteurs.

#### a- Le promoteur du gène de la stromélysine1

Sur le promoteur de la stromélysine1, deux sites EBS sont présents, tête-bêche et séparés par quatre paires de bases. Nous aurions pu imaginer pour ces deux sites deux protéines qui se fixent en dimère. En fait, bien que les deux sites EBS soient nécessaires pour la fixation des protéines Ets-2 et Ergp55, nous n'avons pas constaté la formation de dimères. Malgré la présence de deux sites EBS, là encore, une seule protéine Ets se fixe.

#### b- Un site, une protéine, deux sites, une protéine

Des études antérieures ont démontré la présence de doubles sites EBS, régulés par des protéines Ets, à répétition directe sur l'"enhancer" de HIV-1 et tête-bêche sur les promoteurs des gènes *GATA-1* et *p53* (Hodge et al., 1996 ; Venanzoni et al., 1996). Chacun de ces doubles sites sont espacés de quatre paires de bases tout comme pour le promoteur du gène de la stromélysine1. De la même manière, les protéines Ets-1, Ets-2 et Fli-1 étudiées se fixent en monomères.

A l'intérieur de ces promoteurs, les doubles sites ne sont espacés que de quatre paires de bases. Un encombrement stérique pourrait donc empêcher la fixation de deux protéines sur ces deux sites. En fait, l'espacement entre les doubles sites n'est pas à l'origine d'un manque de place mais plutôt d'une variation d'affinité des protéines Ets. En effet, la liaison de la protéine Fli-1 n'est pas affectée par l'orientation des sites mais seulement par la proximité des double sites (Hodge et al., 1996). Plus ils sont éloignés, jusqu'à vingt paires de bases, meilleure est la liaison. Par contre, en ce qui concerne les protéines Ets-1 et Ets-2, elles se lient mieux lorsque l'espacement est faible dans des répétitions directes ou lorsque l'espacement est grand dans des sites tête-bêche (Hodge et al., 1996 ; Venanzoni et al., 1996).

Un site, une protéine, deux sites, une protéine. Cela semble étonnant mais les données actuelles ne permettent pas d'expliquer ce phénomène. Une chose est sûre, nous

confirmons que les protéines Ets étudiées ici se fixent en monomères sur l'ADN et que la formation de dimères est un événement bien distinct.

## 2- Les domaines impliqués dans la dimérisation des protéines Erg

L'homodimérisation et l'homo-iso-dimérisation des protéines Erg font intervenir de manière indépendante deux régions, une dans la partie amino-terminale de la protéine comprenant le domaine PNT et une dans la partie carboxy-terminale avec le domaine ETS.

Des résultats complémentaires nous ont montré que l'homodimérisation de la protéine Fli-1, protéine très proche du membre Erg à l'intérieur de la famille Ets, implique uniquement le domaine ETS. En revanche, l'hétérodimère formé par les protéines Ergp55 et Fli-1 nécessite l'intervention des domaines ETS des deux protéines et seulement le domaine PNT de la protéine Ergp55 (résultats non publiés). En fait, bien que ces deux membres soient très proches, leur mode d'interaction respectif divergerait.

Nous avons aussi constaté la formation d'hétérodimères entre la protéine Ergp55 et les membres Ets-2, Er81 et Pu-1. Le seul domaine commun à tous ces membres étant le domaine ETS, nous pouvons supposer que ce dernier est impliqué dans l'hétérodimérisation. Cependant, les régions adjacentes spécifiques à chaque protéine Ets pourraient influencer la dimérisation, ce qui expliquerait que tous les membres de la famille Ets n'interagissent pas avec Ergp55.

### a- Le domaine ETS

Cette région, définie comme nécessaire et suffisante pour la liaison à l'ADN, s'est révélée être aussi impliquée dans des interactions protéine-protéine. Comme nous pourrions le voir par la suite, le complexe ternaire formé par les protéines Erg, Fos et Jun implique l'ADN. En revanche, d'après les résultats précédents, la liaison d'une protéine Erg à l'ADN ou à une autre protéine Ets semblent être deux phénomènes mutuellement exclusifs. D'après de nouvelles études non publiées, il semble que ce soit la partie amino-terminale du domaine ETS qui permette l'interaction protéique tandis que nous savons déjà que l'hélice  $\alpha 3$  est impliquée dans la liaison à l'ADN (Fig 3B) (Liang et al., 1994a). Même si ce n'est pas la même région impliquée dans ces deux phénomènes, une orientation différente des hélices  $\alpha$  du domaine ETS, décrite par des études de structure, pourrait être à l'origine de leur séparation. En effet, la dimérisation pourrait modifier la structure tertiaire du domaine ETS, ce qui l'empêcherait alors de se fixer à l'ADN.

En ce qui concerne l'hétérodimérisation de la protéine Ergp55 avec différents membres de la famille Ets, le seul domaine en commun entre tous ces partenaires est le domaine ETS. Cependant, toutes les protéines Ets n'interagissent pas avec la protéine Ergp55. Nous

pouvons alors imaginer que les régions autour du domaine ETS pourraient influencer cette hétérodimérisation. En effet, elles sont déjà à l'origine d'une inhibition à l'ADN pour diverses protéines Ets dont Ets-1 (Wasylyk et al., 1992) et Erm (Laget et al., 1996). Dans le cas de Ets-1, une hélice  $\alpha$  située juste en aval du domaine ETS se rabat sur ce dernier (Hagman and Grosschedl, 1992) et gêne ainsi la reconnaissance de la séquence cible. Pourquoi ne pas envisager alors une inhibition de l'hétérodimérisation?

#### b- Le domaine PNT

Ce domaine PNT a récemment été démontré comme suffisant pour l'oligomérisation de Tel (Jousset et al., 1997). Cependant, le domaine PNT de la protéine Ets-1 mis dans le contexte de la protéine Tel ne permet plus cette oligomérisation. Par ailleurs, la protéine Ets-2 qui possède un domaine PNT est incapable de se dimériser (Basuyaux et al., 1997). Dans le cas de la protéine Erg, ce domaine ne semble pas suffisant pour une homodimérisation, et pas requis dans une hétérodimérisation (résultats non publiés).

Le domaine PNT a été défini par alignement de séquence entre divers membres de la famille Ets. L'identité reste faible en acides aminés et serait plus conservée en ce qui concerne la structure sans que cela ne soit confirmé. C'est sans doute pour toutes ces raisons que le domaine PNT ne présente pas les mêmes fonctions chez plusieurs membres Ets qui le possède. Nous pourrions donc penser que l'alignement de séquence à l'origine de la caractérisation de ce domaine n'explique pas tout. Ainsi, au cours de l'évolution, ce domaine aurait pu obtenir des fonctions distinctes.

#### c- Les domaines ETS et PNT requis dans d'autres types d'interactions

Hahn et al. (1997) ont démontré que les domaines ETS et PNT pouvaient être aussi impliqués dans des associations avec des facteurs n'appartenant pas à la famille Ets. En effet, la protéine huUBC9 ("UBiquitin Conjugating enzyme") interagit avec les membres de la famille Ets, Ets-1 et Fli-1, *via* le domaine ETS et le domaine PNT. Cependant, le domaine PNT ne semble pas essentiel pour cette interaction car les protéines Elf-1, Net et Pu-1, démunies de ce domaine, interagissent.

A l'intérieur de ces domaines, nous recherchons maintenant les motifs structuraux ainsi que les résidus impliqués dans ces dimérisations. La suite de ce travail est prise en charge par Alexis Verger, en thèse dans notre groupe.

### 3- L'implication fonctionnelle de ces interactions

L'implication fonctionnelle de ces dimérisations reste à étudier.



Tout d'abord, la formation d'homo- ou hétérodimères de protéines Ets peut dépendre de plusieurs conditions *in vivo*. Entre autres, la présence dans une même cellule d'un nombre suffisant de protéines qui permettrait la formation de dimères. Ces derniers seront fonction aussi des divers membres de la famille présents et accessibles. Ce dernier point fait intervenir l'étude comparative de l'expression des différents gènes *ets*.

Par ailleurs, la formation de dimères pourrait correspondre à un mécanisme de régulation de la fonction de ces protéines. Dans ce cas, les protéines monomériques pourraient être actives alors que les dimériques seraient inactives. En effet, comme nous l'avons déjà exposé, la dimérisation pourrait empêcher les protéines Erg et Fli-1 de se lier à l'ADN puisque les protéines Ets se lient à l'ADN sous forme de monomères. D'un autre côté, la formation de complexes pourrait masquer les domaines transactivateurs ou moduler les capacités d'interaction avec des cofacteurs. Ces deux possibilités ne sont pas mutuellement exclusives et pourraient même être complémentaires.

La dimérisation des protéines Ets et leur fixation à l'ADN seraient donc bien deux événements distincts et incompatibles qui feraient partie de la régulation de leur activité transcriptionnelle. Des exemples dans la littérature sont déjà en faveur de cette hypothèse. En effet, les homéoprotéines Msx et Dlx interagissent entre elles, indépendamment et de manière non coopérative, par leur domaine de fixation à l'ADN. Leur dimérisation est de plus bloquée en présence de leur site de liaison (Zhang et al., 1997). De même, le facteur de transcription général TFIID forme des homodimères qui doivent se dissocier pour permettre la liaison à l'ADN (Taggart et al., 1996).

#### 4- Interaction des protéines Erg avec le dimère Fos/Jun

Sur le promoteur de la collagénase1, nous avons mis en évidence une coopération entre la protéine Ergp55 et le complexe AP-1. Nous avons déjà vu dans le premier article que la protéine Ergp55 est capable d'interagir physiquement avec le complexe Fos/Jun. L'utilisation de mutants de délétion nous a montré que la formation du complexe Ergp55/Fos/Jun implique le domaine ETS de Ergp55 et la protéine Jun tandis que celui Ergp49/Fos/Jun implique la protéine Fos. Enfin, l'étude de ce complexe soulève deux points importants.

Le premier concerne l'implication du domaine ETS. Ce phénomène commence à être connu par plusieurs équipes. En effet, Nakae et al. (1995) ont suggéré une interaction du domaine ETS de la protéine Erm avec le facteur de transcription c-Jun. Par ailleurs, Bassuk et al. (1995) ont démontré la capacité des protéines Ets-1, Fli-1, Elf-1, et Pu-1 à s'associer au complexe Fos/Jun *via* la protéine Jun. Cette association implique les deux domaines de liaison à l'ADN : le domaine ETS et le domaine basique des protéines Jun. Enfin, Sieweke

et al. (1996) ont étudié la protéine MafB, un membre de la famille des facteurs de transcription bZIP. Elle interagit *via* son domaine basique avec le domaine ETS de la protéine Ets-1. L'implication du domaine de liaison à l'ADN dans des interactions protéiques ne semble donc pas particulier aux protéines Erg et pourrait être une des caractéristiques du domaine ETS.

Le deuxième point mis en évidence par l'étude du complexe Erg/Fos/Jun démontre pour la première fois l'implication de la protéine Fos. En effet, jusqu'à maintenant, il a été démontré que les diverses protéines qui s'associent au complexe AP-1 interagissent physiquement avec Jun. Dans cet article, nous démontrons *in vitro* une interaction directe de la protéine Ergp49 avec le complexe AP-1 *via* Fos. Nous savons qu'il existe différents types de complexes AP-1 formés par des interactions combinatoires entre les différents membres de la famille des protéines bZIP. Ces dimères distincts augmentent la diversité des gènes cibles et la régulation par des stimuli extracellulaires (Aronheim et al., 1997). Pourquoi n'existerait-il pas alors divers complexes ternaires Erg/AP-1? La formation de ces complexes pourrait être fonction de l'expression des isoformes Erg et/ou des gènes cibles à réguler. Ils permettraient alors une régulation plus fine des mécanismes dans lesquels ils sont impliqués.

#### 5- Différences entre les protéines Erg

Grâce aux études de transfections transitoires, nous envisageons pour la première fois une différence entre les protéines isoformes Ergp55, Ergp49 et Ergp38. En effet, nous avons constaté des différences de transactivation sur le promoteur formé par quatre éléments de l'"enhancer" du polyome ainsi que des différences de coopération sur le promoteur de la collagénase1 avec les protéines Fos/Jun.

Nous avons ainsi pu constater, dans les deux types cellulaires étudiés, que la protéine Ergp55 est plus efficace que les isoformes Ergp49 et Ergp38 en transactivation, la capacité des deux dernières étant comparable. Avec le complexe AP-1, il y a un ordre décroissant d'efficacité de coopération pour les protéines Ergp55, Ergp49 et Ergp38. Dans cette étude, les domaines codés par les exons alternatifs semblent nécessaires à la coopération. Ces derniers peuvent être impliqués dans l'interaction protéique soit de manière directe, soit indirectement par une modification de la structure tridimensionnelle ou par association avec d'autres facteurs.

La présence ou l'absence des domaines codés par les exons alternatifs peuvent être à l'origine des différences observées entre les isoformes issues du gène *erg*.

## 6- Les isoformes

Les spécificités des isoformes peuvent dépendre de cofacteurs, de phosphorylations ou même des séquences cibles. Peu d'exemples de rôle spécifique de protéines isoformes sont encore connus dans la littérature. Nous présentons ici quelques études qui pourraient être des modèles afin de définir la fonction des protéines Erg.

### a- Un exemple dans la famille Ets

Quand la protéine rhombotin-2, RBTN-2, est exprimée ectopiquement dans des cellules T, il en résulte une leucémie. Le mécanisme moléculaire n'est pas encore bien compris mais les fonctions normales de la protéine RBTN-2 nécessitent des interactions protéine-protéine. Afin de définir si l'implication de RBTN-2 dans des leucémies de cellules T fait aussi intervenir des interactions, Wilkinson et al. (1997) ont utilisé la protéine comme appât dans un criblage de banque de tumeur thymique murine. Cette étude a été l'occasion de mettre en évidence un nouveau membre de la famille Ets, la protéine Elf-2, appelée ainsi parce que son domaine ETS est quasiment identique à celui de la protéine Elf-1. Le gène *elf-2* donne naissance, par épissage alternatif, à deux transcrits. Les protéines correspondantes Elf-2a et Elf-2b diffèrent dans leur partie amino-terminale. La protéine Elf-2b ne possède pas un domaine acide, potentiellement transactivateur. Les auteurs suggèrent donc que les deux isoformes ont des effets opposés lors du développement des lymphocytes T. En effet, elles possèdent le même domaine ETS, le même domaine de liaison à la protéine RBTN-2, mais l'une a un domaine transactivateur absent chez l'autre (Wilkinson et al., 1997).

Les isoformes Erg possèdent deux, un ou aucun des exons alternatifs à l'intérieur de leur domaine transactivateur amino-terminal. Celui-ci pourrait donc être plus ou moins efficace ce qui pourrait se rapprocher de l'exemple des protéines Elf-2a et Elf-2b, avec ou sans domaine transactivateur.

### b- Les isoformes de SRC1 et le récepteur aux oestrogènes

SRC1 ("Steroid Receptor Co-activator") est un des coactivateurs capable de potentialiser l'activité des récepteurs nucléaires dont le récepteur aux oestrogènes, ER ("Estrogen Receptor"). Kalkhoven et al. (1998) ont mis en évidence deux isoformes humaines, SRC1a et SRC1e, qui divergent d'un point de vue structural, à leurs extrémités carboxy-terminales et d'un point de vue fonctionnel, dans leur rôle de médiateur de la transcription de ER.

Une conséquence de leur différence carboxy-terminale est la présence d'un second domaine de transactivation dans SRC1e, indépendant de CBP ("CREB Binding Protein"). En effet, bien que la protéine CBP soit reconnue comme essentielle dans la régulation par des récepteurs nucléaires, elle n'est pas suffisante à SRC1a pour potentialiser la transcription du

gène *ER*. Les auteurs suggèrent en fait que le domaine transactivateur indépendant de CBP permet à SRC1e d'interagir avec un autre facteur important (Kalkhoven et al., 1998).

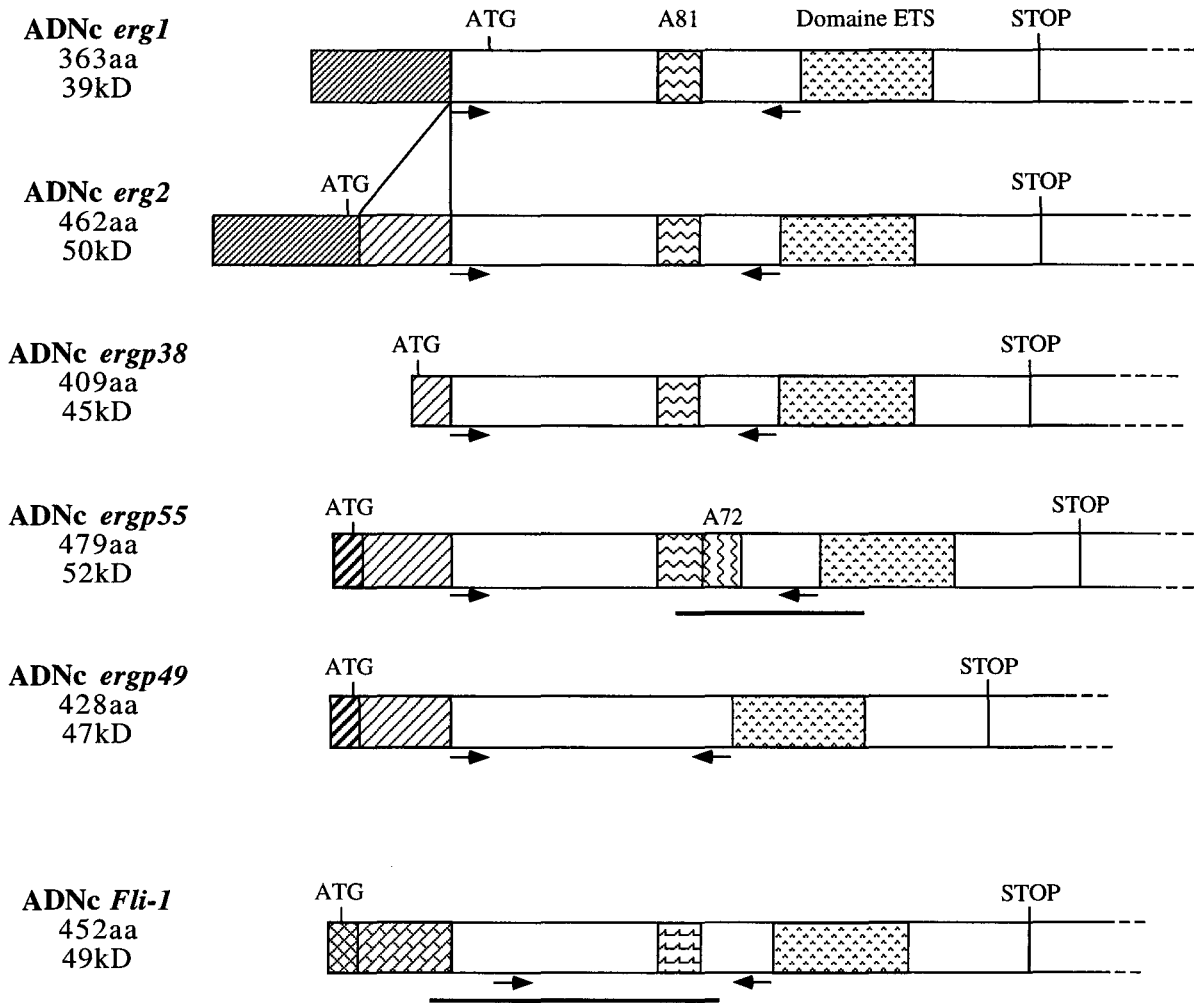
La présence d'un domaine supplémentaire dans une des isoformes permet l'interaction avec un nouveau cofacteur. Cette constatation est tout à fait envisageable avec les protéines Erg qui possèdent divers exons alternatifs. Cependant, les cofacteurs des protéines Erg restent à mettre en évidence.

#### c- Deux isoformes au cours de la myogenèse

La protéine MNF ("Myocyte Nuclear Factor") a été identifiée comme un facteur de transcription de type "hélice-tour-hélice de type ailé", tout comme les membres de la famille Ets. En fait, le gène correspondant est à l'origine de deux isoformes connues MNF $\alpha$  et MNF $\beta$  qui diffèrent dans leur partie carboxy-terminale. Les proportions relatives de chacune de ces isoformes divergent à différents stades de la myogenèse. La protéine MNF $\alpha$  est exprimée dans les myoblastes indifférenciés et en prolifération. Malgré la présence d'un domaine de liaison à l'ADN fonctionnel, elle se lie très faiblement à l'ADN et est incapable de réguler directement des promoteurs *in vivo*. Par contre, la protéine MNF $\beta$  est régulée positivement lors de la différenciation myogénique et de l'arrêt du cycle cellulaire. Elle se lie fortement à son site de reconnaissance sur l'ADN et fonctionne alors comme un répresseur (Yang et al., 1997).

Lors de la différenciation terminale au cours de la myogenèse, les auteurs ont observés une transition de l'expression des deux isoformes. Ils spéculent donc que MNF $\alpha$  et MNF $\beta$  pourraient exercer des fonctions opposées. Ainsi, par une régulation de l'épissage alternatif du gène *MNF*, un ensemble de gènes cibles est commuté dans la prolifération des myoblastes vers la différenciation des myotubes (Yang et al., 1997).

Nous ne connaissons pas l'expression des isoformes issues du gène *erg*. Cette étude pourrait donc mettre en évidence une nouvelle spécificité des protéines Erg. C'est ce que nous étudions dans le prochain article.



**Figure 20 :** Les ADNc isoformes issus des gènes *erg* humain et l'ADNc issu du gène *fli-1* humain. Leur taille en acides aminés et leur masse moléculaire correspondantes sont indiquées. D'après Duterque-Coquillaud et al., 1993. Les flèches indiquent l'emplacement des oligonucléotides utilisés en RT-PCR, et les lignes les fragments d'ADNc utilisés en hybridation *in situ*.

### III- Publication 3

Cet article est en préparation sous le titre de : "Le gène *erg* murin, membre de la famille *ets*, est à l'origine d'au moins deux transcrits alternatifs exprimés de manière très précoce dans le développement embryonnaire. "

#### 1- Introduction

Dans le but d'élucider la fonction d'un gène, les études d'expression au cours de l'embryogenèse sont une des premières étapes à réaliser. En ce qui concerne le gène *erg*, l'étude de son expression a été réalisée au cours du développement embryonnaire du poulet (Dhordain et al., 1995). Il est exprimé de manière très précoce, aux stades E1 et E3, dans le mésoderme et ses dérivés, les cellules endothéliales et les cellules des crêtes neurales en migration. A ces stades, ce profil est quasiment identique à celui du gène *ets-1*. Un peu plus tard, au stade E6, l'expression du gène *erg* se différencie de celle du gène *ets-1*. Elle se restreint aux zones précartilagineuses ce qui suppose l'implication du gène *erg* dans la mise en place du cartilage.

Le gène *erg* est à l'origine, par des mécanismes de promotions et d'épissages alternatifs, de différentes protéines Erg (Rao et al., 1987 ; Duterque-Coquillaud et al., 1993). Il en existe au moins cinq, Erg1, Erg2, Ergp55, Ergp49 et Ergp38, qui diffèrent toutes dans leurs parties amino-terminale et centrale (Fig 20). L'activité de régulation de ces protéines *in vivo* est fonction de divers mécanismes dont l'un d'entre eux peut être l'expression spécifique dans un tissu ou à un stade de développement. C'est pourquoi il est nécessaire de connaître leur profil d'expression pour mieux comprendre leur spécificité et redondance spécifiques.

En revanche, le gène *fli-1*, le plus proche du gène *erg* dans la famille *ets*, ne semble pas donner naissance à plusieurs ARNm puisqu'un seul ADNc a été mis en évidence (Ben-David et al., 1990). La protéine correspondante présente 70% d'identité avec la protéine Erg2 (Ben-David et al., 1991) mais sa partie amino-terminale se rapproche plus de celle des protéines Ergp55 et Ergp49 (Fig 20). Le profil d'expression du gène *fli-1* est semblable à celui du gène *ets-1* (Quéva et al., 1989). En effet, dans l'embryogenèse très précoce, le gène *fli-1* est exprimé dans les lignées mésodermiques, tandis qu'après la morphogenèse, son expression est fortement diminuée dans le mésoderme mature (Mélet et al., 1996). Malgré les modèles animaux différents utilisés dans chacune de ces études, la coexpression des gènes *erg* et *fli-1* semble exister dans les premiers jours du développement de l'embryon, leurs sites d'expression se différenciant par la suite.

Dans cette étude, nous recherchons, chez la souris, l'expression correspondant aux cinq ADNc *erg* et à l'ADNc *fli-1*, très proches les uns de autres. L'approche choisie est de

comparer l'expression des gènes *erg* et *fli-1* au cours de l'embryogenèse murine par deux techniques complémentaires, la RT-PCR et l'hybridation *in situ*.

## 2- Matériels et méthodes

### a- RT-PCR

Les ARN totaux de souris B6SJL ou (C57Bl/6 x SJL)F1 à 8, 10, 13, 15 et 17 jours de développement embryonnaire ont été extraits par la méthode de thiocyanate de guanidinium, suivie d'une centrifugation sur coussin de chlorure de césium en solution à 5,7M (Sambrook et al., 1989).

Pour la transcription reverse, 15µg d'ARN totaux qsp 25µl sont dénaturés à 90°C 2mn. La réaction est effectuée 1h à 37°C avec 500ng d'oligonucléotide polydT, le tampon approprié (50mM Tris HCl pH 8,3, 75mM KCl, 3mM MgCl<sub>2</sub> et 0,02M DTT) et 200 unités de transcriptase reverse (Superscript Gibco BRL). La réaction est arrêtée à 100°C 5mn.

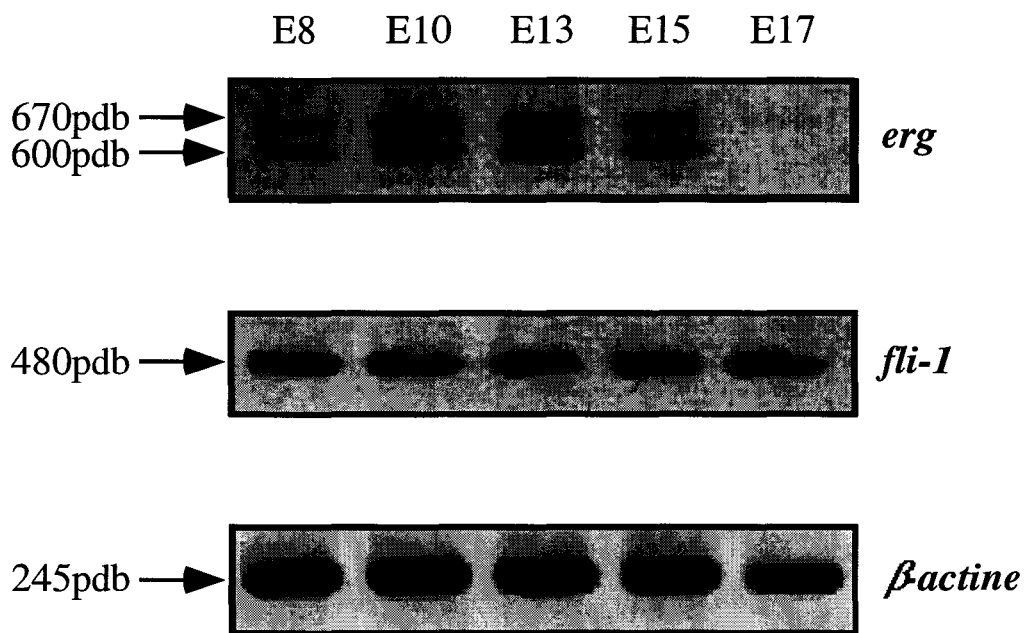
Pour la PCR, un dixième de la réaction de transcription reverse est amplifié avec 2,5 unités de Taq polymérase (Gold Perkin Elmer-Cetus) dans un volume final de 100µl contenant le tampon approprié (75mM Tris HCl pH 9, 20mM sulfate de diammonium, 0,01% Tween 20 et 15mM MgCl<sub>2</sub>), 10µM de chaque dNTP et 100ng de chaque oligonucléotide. La PCR s'effectue sur 30 cycles où la dénaturation préalable se fait à 94°C pendant 8mn, la dénaturation à 94°C pendant 2mn, l'hybridation des amorces à 60°C pendant 1mn, l'élongation à 72°C pendant 1mn et l'élongation finale à 72°C pendant 10mn. Un deuxième tour de PCR est refait avec un vingtième de la première PCR. 10µl de la seconde PCR sont déposés sur un gel agarose 2% coloré au bromure d'éthidium.

Les oligonucléotides pour obtenir les ADNc de *erg* sont 5'-CTCACCTGATGAGTGCAGTGTGAACAAAGG-3' et 5'-ATTAGCAAGGCGGCTACTGGTCCGTCAG-3', pour l'ADNc de *fli-1* 5'-AGCCAACCCCATGAACTATAATAGCTACAT-3' et 5'-CTGCTCAGTGTCTTGCCCATGGTCTGTGAT-3' et pour l'ADNc de *β-actine* 5'-CGTGGGCCGCTAGGCACCA-3' et 5'-TTGGCCTTAGGGTTCAGGGGGG-3'.

### b- Hybridation *in situ*

L'hybridation *in situ* est exécutée sur des sections de tissus d'embryons de souris à divers stades de développement comme décrit dans Quéva et al. (1989).

Les sondes ARN antisens et sens marquées au <sup>35</sup>S ont été synthétisées à partir d'un fragment de l'ADNc *ergp55* murin et de l'ADNc *fli-1* murin (Fig 20).



**Figure 21 :** Profil d'expression des gènes *erg*, *fli-1* et  $\beta$ -actine au cours de divers stades du développement embryonnaire. Les expériences de RT-PCR ont été réalisées à partir d'ARN totaux de souris à 8, 10, 13, 15 et 17 jours de développement.



### 3- Résultats

Dans cette étude, nous avons choisi d'étudier, chez la souris, les stades entre 8 jours et 17 jours de développement embryonnaire. Ce choix est fonction de premiers résultats publiés. En effet, Dhordain et al. (1995) ont décrit l'expression du gène *erg* chez le poulet et Mélet et al. (1996) celle du gène *fli-1* chez la souris.

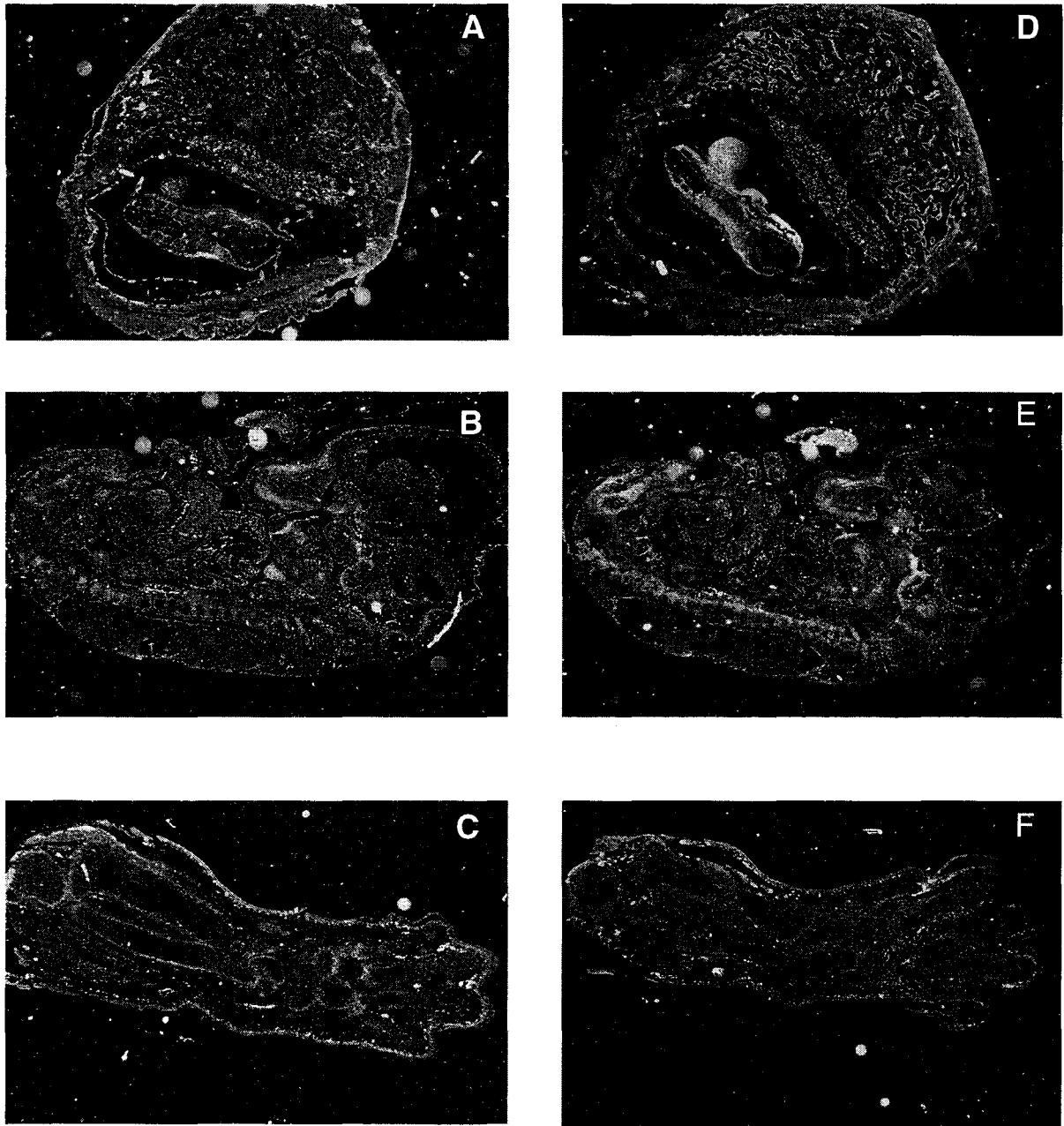
#### a- L'expression des isoformes *erg* varie au cours du temps

Le choix des oligonucléotides, dans la technique de RT-PCR, a été difficile puisque la séquence du gène *erg* murin n'est que partiellement connue. Cependant, l'étude de l'organisation génomique de ce gène et des expériences de RT-PCR déjà réalisées dans le groupe ont permis de choisir des oligonucléotides murins qui encadrent les exons alternatifs centraux et qui permettent d'amplifier des fragments d'ADN dont la taille révélera la présence des deux exons alternatifs centraux, d'un seul ou d'aucun (Fig 20). Par ailleurs, la séquence murine de l'ADNc *fli-1* étant publiée (Ben-David et al., 1990), nous avons choisi des oligonucléotides qui encadrent la même région (Fig 20) afin de mieux comparer les deux études. Nous avons choisi d'amplifier la  $\beta$ -actine comme témoin d'expression, les oligonucléotides ont déjà été décrits par Telford et al. (1990).

Nous avons ainsi mis en évidence au moins deux isoformes issues du gène *erg*. Les tailles des fragments obtenus démontrent la présence de deux exons alternatifs dans la bande de 670 pdb et d'un seul dans celle de 600 pdb (Fig 21). Ces fragments d'ADN ont été clonés puis séquencés ce qui a confirmé leur origine murine et la présence des exons alternatifs. Les fragments de 670 pdb pourraient correspondre à l'ADNc *ergp55*, et ceux de 600 pdb aux ADNc *erg1*, *erg2* et *ergp38*. Par contre, l'isoforme *ergp49*, qui ne possède aucun des exons alternatifs (bande attendue à 530 pdb), ne peut-être détectée à ces stades.

Nous constatons que l'expression des deux types de transcrits est constante de E8 à E15. Par contre, à E17, l'expression des isoformes ne présentant qu'un seul exon alternatif disparaît et celle de *ergp55* perdure mais plus faiblement. Même si nous n'avons pas fait ici de PCR quantitatives, ces variations d'expression peuvent être prises en compte car le pourcentage d'ADNc de  $\beta$ -actine amplifié est essentiellement le même pour chaque stade de développement étudié.

En ce qui concerne le gène *fli-1*, une seule isoforme est détectée (Fig 21). La taille du fragment d'ADN est en accord avec celle attendue et avec la présence de l'exon central. Le clonage et le séquençage de ce fragment le définissent comme murin. A la différence du gène *erg*, son expression est constante au cours des différents stades étudiés.



**Figure 22** : Localisation de l'expression des ARNm issus des gènes *erg* (A, B et C) et *fli-1* (D, E et F) par hybridation *in situ* au cours de l'embryogenèse murine. A 10,5 jours du développement dans l'embryon et le placenta (A et D), à 13 jours dans l'embryon (B et E) et à 15 jours dans le bourgeon de membre (C et F).

b- Les gènes *erg* et *fli-1* sont coexprimés de manière très précoce

Le profil d'expression du gène *erg* étudié ici chez la souris est semblable à celui observé antérieurement dans l'embryon de poulet (Dhordain et al., 1995). En revanche, le développement de la souris implique la mise en place préalable du placenta qui s'est révélé un site majeur d'expression de ce gène. Cette étude sera présentée en parallèle avec celle du gène *fli-1*.

Expression à E10,5 dans l'embryon et le placenta

La figure 22A et D montre une coupe longitudinale d'embryon en vue dorsale dans son placenta à 10,5 jours de développement, marquée avec les sondes antisens *erg* (Fig 22A) et *fli-1* (Fig 22D). Les sites d'expression des deux gènes ont des points communs. Cependant, le signal correspondant au marquage du gène *fli-1* est plus intense ce qui laisse supposer une plus forte expression de ce dernier. Ils sont exprimés dans les cellules endothéliales qui bordent les sinus veineux du placenta mais pas du tout dans le trophoblaste. Au niveau de l'embryon lui-même, nous constatons un marquage important au niveau du bourgeon de membre et dans le mésenchyme autour du tube neural.

Expression dans l'embryon total à E13

La figure 23B et E présente une coupe longitudinale d'un embryon à 13 jours de développement. Les sites d'expression des deux gènes sont toujours chevauchants avec un signal plus intense pour le gène *fli-1*. Nous constatons un marquage de la condensation précartilagineuse, dans les disques intervertébraux, le squelette céphalique, le cartilage de Meckel et la capsule optique. Ils sont aussi exprimés dans l'endocarde et le mésenchyme rénal.

Expression dans les pattes à E15

Sur des coupes tangentielles d'un membre antérieur à 15 jours de développement (Fig 22C et F), les sites d'expression des deux gènes se différencient. Le gène *erg* est exprimé dans le derme, le périchondrium et dans le cartilage. A la différence des stades précédents, l'expression du gène *fli-1* est plus faible, avec cependant un site important dans le derme.

#### 4- Discussion

Dans cette étude, nous avons utilisé deux techniques complémentaires, la RT-PCR et l'hybridation *in situ*, pour décrire l'expression des gènes murins *erg* et *fli-1*. Ils sont coexprimés de manière très précoce puis leur sites d'expression diffèrent. Le gène *erg* est à l'origine de plusieurs transcrits tandis que le gène *fli-1* ne donne naissance qu'à un ARN.

#### a- L'ADNc *ergp55* semble majoritaire

Par la technique de RT-PCR, nous avons mis en évidence au moins deux ARNm entre les stades E8 et E15. La bande de 670 pb représente des fragments d'ADN comprenant les deux exons alternatifs qui peuvent correspondre à l'ADNc *ergp55* (Fig 20). Cependant, l'étude de la partie 5' des transcrits n'étant pas faite, nous ne pouvons affirmer qu'il n'existe pas un ARNm possédant deux exons alternatifs centraux et une partie 5' semblable à celle des ARNm *erg1* et *erg2*. Cette étude est encore impossible car nous ne connaissons pas la séquence des parties codantes du gène *erg* murin pour choisir des oligonucléotides murins. La bande de 600 pb représente des fragments avec un seul exon alternatif. Ici, nous pouvons avoir trois possibilités, les ADNc *erg1*, *erg2* et *ergp38* (Fig 20). Pour trancher, nous devrions pouvoir amplifier les parties 5' des ARNm afin de définir la présence ou l'absence des exons alternatifs correspondants.

Plusieurs arguments sont en faveur d'une majorité du transcrit *ergp55* sur les autres isoformes. En effet, l'ADNc correspondant a été isolé chez l'humain ainsi que les autres membres (Duterque-Coquillaud et al., 1993) mais c'est le seul mis en évidence chez le poulet (Dhordain et al., 1995). Il a aussi été mis en évidence chez le chat (Martine Duterque-Coquillaud, communication personnelle).

A E17, seule une bande perdure faiblement. Elle correspond à l'ADNc *ergp55*. Ce résultat est en accord avec la publication de Dhordain et al. (1995). En effet, à ce stade là, les auteurs ont démontré que l'expression du gène *erg* est restreinte aux cellules endothéliales des vaisseaux en formation et aux zones cartilagineuses interdigitales, c'est-à-dire un faible taux d'expression sur l'embryon total.

Aux stades de développement embryonnaire étudiés ici, nous n'avons pas détecté l'expression de l'ARNm *ergp49* qui ne contient aucun exon alternatif central. Nous pouvons supposer que cette isoforme est exprimée à un autre stade voire à l'état adulte et que son expression correspond à une autre fonction. Ces résultats, tout comme les études de dimérisation avec les protéines Erg, suggèrent l'existence d'une spécificité de cette isoforme. En effet, l'isoforme Ergp49 interagit avec le complexe AP-1 via la protéine Fos tandis que Ergp55, comme la plupart des facteurs qui s'associent aux protéines Fos et Jun, interagit avec Jun.

#### b- Le gène *fli-1* est à l'origine d'un seul transcrit

Pour le gène *fli-1*, un seul fragment est amplifié. Un seul ADNc a été mis en évidence à ce jour et notre résultat le confirme. Les gènes *erg* et *fli-1* étant très proches, nous aurions pu penser qu'ils existaient plusieurs ARNm *fli-1* comparables à ceux issus du gène *erg*.

Cependant, nous pourrions imaginer qu'au cours de l'évolution, l'organisation génomique est restée sensiblement la même mais que les régions correspondant aux sites d'épissage de ces exons alternatifs n'aient pas été conservées. De plus, cette région de la protéine Fli-1 correspond à la partie codante la plus divergente de la protéine Erg.

#### c- L'expression des gènes *erg* et *fli-1*

Les gènes *erg* et *fli-1* présentent de larges sites de coexpression aux stades précoces de développement embryonnaire, E10 et E15. Ils sont tous deux exprimés dans les dérivés mésodermiques et dans les cellules endothéliales des vaisseaux en formation. De telles coexpressions suggèrent une fonction partiellement redondante dans les tissus correspondants à moins qu'ils aient des effets antagonistes sur leurs gènes cibles.

A E15, leur coexpression se restreint au derme dans le bourgeon de membre. En revanche, l'expression du gène *erg* s'associe aux zones précartilagineuses résiduelles ainsi qu'au périchondrium. Il a déjà été montré que ce gène est impliqué dans la mise en place du cartilage (Gañan et al., 1996).

Les sites d'expression de ces gènes déterminés, nous pouvons alors envisager d'établir des lignées de souris transgéniques. Ainsi, avec des promoteurs connus pour spécifiques de certains tissus, nous pourrions surexprimer ces gènes dans les cellules endothéliales, le tissu précartilagineux. Nous pourrions alors avoir une idée précise sur le rôle de ces gènes dans ces tissus.

# **Discussion générale**

Jusqu'alors, peu de choses étaient connues sur le mode d'action des protéines Erg. Les études que nous avons entreprises ont apporté quelques réponses qui ont à leur tour soulevé d'autres questions.

#### Le mode d'action des protéines Erg

Le gène *erg* est à l'origine d'au moins cinq protéines isoformes. Elles présentent toutes, *in vitro* sur un promoteur artificiel, la même activité transactivatrice. Dans le but de déterminer leur mode d'action, nous avons mis en évidence la capacité de Ergp55 à transactiver différemment les promoteurs humains de la collagénase1 et de la stromélysine1. Sur ce dernier, elle présente de plus un antagonisme de fonction avec la protéine Ets-2. Le mode d'action précis de ces différents modes de fonctionnement n'est toujours pas clairement résolu. Cependant, nous avons démontré, pour la première fois, un antagonisme de fonction entre deux membres de la famille Ets. Ce type d'antagonisme n'est vraisemblablement pas un cas particulier dans la famille et pourrait être à l'origine d'une régulation des protéines Ets exprimées dans une même cellule.

D'autre part, nous avons caractérisé la capacité des protéines Erg à dimériser. Toutes les interactions protéine-protéine et leurs multiples combinaisons mises en évidence dans nos travaux pourraient correspondre à un vaste réseau de régulations. Les dimères étant incompatibles avec la fixation à l'ADN, l'activité de ces facteurs de transcription serait régulée par des associations protéiques. Les implications précises et les mécanismes moléculaires de ces potentielles régulations restent à déterminer *in vivo*.

#### Les isoformes Erg ne sont pas toutes coexprimées

Les fonctions potentiellement différentes de ces protéines Erg pourraient être modulées par leur expression. Ainsi, nous avons démontré que seule l'isoforme de type Ergp49 n'était pas exprimée entre les stades de développement embryonnaire E8 et E17. Son expression en d'autres sites et à d'autres stades reste à mettre en évidence et pourrait être à l'origine de sa spécificité. En revanche, les isoformes de type Erg1, Erg2, Ergp55 et Ergp38 semblent coexprimées. Leur fonction respective pourrait alors être régulée par la formation de dimères.

Ces premiers résultats définissent l'expression des transcrits possédant les exons alternatifs centraux. Ils doivent être complétés puisqu'il existe aussi des exons alternatifs en 5'. Cette deuxième expérience de RT-PCR, comparée à la première, permettra alors d'affirmer l'expression des différentes types d'isoformes. De plus, d'autres stades de développement embryonnaire ainsi que divers tissus adultes seront utilisés pour compléter

notre étude. Enfin, cette dernière ne sera complète que lorsque les deux promoteurs du gène *erg* seront isolés et caractérisés. Toutes ces études sont en cours dans notre groupe.

#### Le double rôle du domaine ETS

Nous avons démontré que le domaine ETS des protéines Erg est impliqué dans les interactions protéiques. Par ailleurs, Bassuk et al. (1995) et Hahn et al. (1997) ont mis en évidence cette propriété, nouvellement définie, chez les protéines Ets-1, Fli-1, Elk-1 et Pu-1. Jusqu'alors, ce domaine, signature de la famille Ets, possédait une fonction bien établie de liaison à l'ADN. Il aurait donc maintenant un double rôle : liaison à l'ADN et aux cofacteurs. Un argument en faveur de l'existence de ce double rôle dans toutes les protéines de la famille Ets pourrait expliquer la grande conservation de ce domaine ETS. Nous pouvons alors supposer que l'activité de tous ces facteurs de transcription est aussi régulée par des interactions protéine-protéine.

#### L'interface de dimérisation des protéines Erg

Afin de mieux comprendre le mode d'action des dimérisations des protéines Erg, la cartographie fine des régions impliquées est en cours et pourrait donner un premier élément de réponse. D'autre part, nous savons que quelques acides aminés peuvent suffire à une interaction protéique.

Ainsi, dans le cas des coactivateurs des récepteurs nucléaires, il a été démontré que le motif LXXLL, où L représente l'acide aminé leucine et X un acide aminé, est suffisant pour l'interaction entre les récepteurs nucléaires et leurs cofacteurs tels CBP, RIP-140, SRC-1, p/CIP ou NcoA-1 (Heery, 1997 ; Torchia et al., 1997). En fait, parmi ces acides aminés présents au sein d'une hélice  $\alpha$ , les leucines sont disposées sur une même face de l'hélice et forme une interface hydrophobe (Torchia et al., 1997).

Nous avons noté la présence d'un tel motif dans le domaine ETS qui pourrait être à l'origine de la dimérisation des protéines Erg. Bien sur, nous ne pouvons exclure la possibilité de l'existence d'un ou plusieurs autres motifs protéiques, non caractérisés, responsables des diverses interactions protéiques.

#### Le rôle des protéines Erg *in vivo*

Afin de déterminer la fonction des protéines Erg, des études *in vivo* sont requises. D'une part, l'inactivation du gène par recombinaison homologe est en cours dans notre groupe. D'autre part, par transgénèse, nous surexprimons la protéine sauvage et mutante de manière ciblée dans le cartilage et les cellules endothéliales. Ainsi, nous pourrions envisager la surexpression d'une protéine Erg incapable de dimériser mais se fixant toujours



à l'ADN, et évaluer les conséquences de l'absence de ces dimérisations. Cette étude serait à comparer à la surexpression d'une protéine incapable de fixer l'ADN mais toujours capable d'interaction protéiques. De telles protéines mutantes ne pourront être définies que lorsque des motifs protéiques suffisants pour les associations protéiques seront caractérisés.

# **Références bibliographiques**

- Alkema, M. J., M. Bronk, E. Verhoeven, A. Otte, L. J. van't Veer, A. Berns, and M. van Lohuizen.** 1996. Identification of Bmi1-interacting proteins as constituents of a multimeric mammalian Polycomb complex. *Genes & Dev.* **11**:226-240.
- Antinore, M. J., M. J. Birrer, D. Patel, L. Nader, and D. J. McCance.** 1996. The human papillomavirus type 16 E7 gene product interacts with and trans-activates the AP-1 family of transcription factors. *EMBO J.* **15**:1950-1960.
- Aronheim, A., E. Zandi, H. Hennemann, S. J. Elledge, and M. Karin.** 1997. Isolation of an AP-1 repressor by a novel method for detecting protein-protein interactions. *Mol. Cell. Biol.* **17**:3094-3102.
- Bailly, R. A., R. Bosselut, J. Zucman, F. Cormier, O. Delattre, M. Roussel, G. Thomas, and J. Ghysdael.** 1994. DNA-binding and transcriptional activation properties of EWS-FLI-1 fusion protein resulting from the t(11;22) translocation in Ewing sarcoma. *Mol. Cell. Biol.* :3230-3241.
- Bassuk, A. G., R. T. Anadappa, and J. M. Leiden.** 1997. Physical interactions between Ets and NF-KB/NFAT proteins play a role in their cooperative activation of the human immunodeficiency virus enhancer in Tcells. *J. Virol.* **71**:3563-3573.
- Bassuk, A. G., and J. M. Leiden.** 1995. A direct physical association between ETS and AP-1 transcription actors in normal human T cells. *Immunity* **3**:223-237.
- Basuyaux, J. P., E. Ferreira, D. Stéhelin, and G. Buttice.** 1997. The Ets transcription factors interact with each other and with the c-Fos/c-Jun complex *via* distinct protein domains in a DNA-dependent and -independent manner. *J. Biol. Chem.* **272**:26188-26195
- Batchelor, A. H., D. E. Piper, F. C. de la Brousse, S. L. McKnight, and C. Wolberger.** 1998. The structure of Gabp $\alpha/\beta$  : an ETS domain-ankyrin repeat heterodimer bound to DNA. *Science* **279**:1037-1041.
- Baud, V., M. Lipinski, E. Rassart, L. Poliquin, and D. Bergeron.** 1991. The human homolog of the mouse common viral integration region, *fli-1*, maps to 11q23-q24. *Genomics* **11**:223-224.
- Ben-David, Y., E. B. Giddens, and A. Bernstein.** 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., E. B. Giddens, K. Letwin, and A. Bernstein.** 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.
- Bertolotti, A., Y. Lutz, D. J. Heard, P. Chambon, and L. Tora.** 1996. hTAT<sub>II</sub>68, a novel RNA/ssDNA-binding protein with homology to the pro-oncoproteins TLS/FUS and EWS is associated with both TFIID and RNA polymerase II. *EMBO J.* **15**:5022-5031.
- Bertolotti, A., T. Melot, J. Acker, M. Vigneron, O. Delattre, and L. Tora.** 1998. EWS, but not EWS-Fli-1, is associated with both TFIID and RNA polymerase II :

interactions between two members of the TET family, EWS and hTAT<sub>II</sub>68, and subunits of TFIID and RNA polymerase II complexes. *Mol. Cell. Biol.* **18**:1489-1497.

**Bhattacharya, G., L. Lee, E. S. P. Reddy, and V. N. Rao.** 1993. Transcriptional activation domains of Elk-1, ΔElk-1 and Sap-1 proteins. *Oncogene* **8**:3459-3464.

**Bochert, M. A., L. A. Kleinbaum, L. Y. Sun and F. H. Burton.** 1998. Molecular cloning and expression of *ehf*, a new member of the *ets* transcription factor/oncoprotein gene family. *Biochem. Biophys. Res. Commun.* **246**:176-181.

**Bohmann, D., T. J. Bos, A. Admon, T. Nishimura, P. K. Vogt, and R. Tjian.** 1987. Human prot-oncogene c-jun encodes a DNA binding protein with structural and functional properties of transcription factor AP-1. *Science* **238**:1386-1392.

**Bories, J. C., D. M. Willerford, D. Grévin, L. Davidson, A. Camus, P. Martin, D. Stéhelin and F. W. Alt.** 1995. Increased T-cell apoptosis and terminal B-cell differentiation induced by inactivation of the *ets-1* proto-oncogene. *377*:635-638.

**Borneman, D., E. Miller, and J. Simon.** 1996. The *Drosophila Polycomb* group gene *Sex comb on midleg (Scm)* encodes a zinc finger protein similarity to polyhomeotic protein. *Development* **122**:1621-1630.

**Bosselut, R., J. F. Duvall, A. Gégonne, M. Bailly, A. Hemar, J. Brady, and J. Ghysdael.** 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**:3137-3144.

**Braun, B. S., R. Frieden, S. L. Lessnick, W. A. May, and C. T. Denny.** 1995. Identification of target genes for the Ewing's sarcoma EWS/FLI-1 fusion protein by representational difference analysis. *Mol. Cell. Biol.* **15**:4623-4630.

**Brown, J. R., H. Ye, R. T. Bronson, P. Dikkes, and M. E. Greenberg.** 1996. A defect in nurturing in mice lacking the immediate early gene *fosB*. *Cell* **86**:297-309.

**Brown, T. A., and S. L. McKnight.** 1992. Specificities of protein-protein and protein-DNA interaction of GABPα and two newly defined Ets-related proteins. *Genes & Dev.* **6**:2502-2512.

**Busch, S. J., and P. Sassone-Corsi.** 1990. Dimers, leucine zippers and DNA-binding domains. *TIG* **6**:36-40.

**Butticè, G., M. Duterque-Coquillaud, J. P. Basuyaux, S. Carrère, M. Kurkinen, and D. Stéhelin.** 1996. Erg, an Ets-family member, differentially regulates human collagenase1 (MMP1) and stromelysin1 (MMP3) gene by physically interacting with the Fos/Jun complex. *Oncogene* **13**:2297-2306.

**Butticè, G., and A. Kurkinen.** 1993. A polyomavirus enhancer A-binding protein-3 site and Ets-2 protein have a major role in the 12-O-tetradecanoylphorbol-13-acetate response of the human stromelysin gene. *J. Biol. Chem.* **268**:7196-7204.

**Calmels, T. P. G., V. Mattot, N. Wernert, B. Vanderbunder, and D. Stéhelin.** 1995. Invasive tumor induce c-ets-1 transcription factor expression in adjacent stroma. *Bio. Cell.* **84**:53-61.

**Cowell, I. G.** 1994. Repression versus activation in the control of gene transcription. *TIBS* **19**:38-42.

**Crépieux, P., J. Coll, and D. Stéhelin.** 1994. The ETS family of proteins : weak modulators of gene expression in quest for transcriptional partners. *Crit. Rev. of Onc.* **5**:615-638.

**Crépieux, P., D. Leprince, A. Flourens, O. Albagli, E. Ferreira, and D. Stéhelin.** 1993. The two functionally distinct amino-termini of chicken *c-ets-1* products arise from alternative promoter usage. *Gene Expression* **3**:215-225.

**Crété, N., P. H. Gosset, D. Teophile, M. Duterque-Coquillaud, J. L. Blouin, C. Vaysettes, P. M. Sinet, and N. Creau-Golderg.** 1993. Mapping the down syndrome chromosomal region (DCR) : establishment of a YAC contig spanning 1200 kb. *Eur. J. Hum. Genet.* **1**:51-63.

**Crozat, A., P. Aman, N. Mandahl, and D. Ron.** 1993. Fusion of CHOP to a novel RNA-binding protein in human myxoid liposarcoma. *Nature* **363**:640-644.

**Dahmane, N., G. A. Ghezala, P. Gosset, Z. Chamoun, M. C. Dufresne-Zacharia, C. Lopes, N. Rabatel, S. Gassanova-Maugenre, Z. Chettouh, V. Abramawski, E. Fayet, M. L. Yaspo, B. Korn, J. L. Blouin, H. Lehrach, N. Poutska, E. Antonarakis, P. M. Sinet, N. Créau, and J. M. Delabar.** 1998. Transcriptional map of the 2.5-MB CBR-ERG region of chromosome 21 involved in Down syndrome. *Genomics* **48**:12-23.

**Dalton, S., and R. Treisman.** 1992. Characterization of Sap-1, a protein recruited by serum response factor to the *c-fos* serum response element. *Cell* **68**:597-612.

**Darby, T. G., J. D. Meissner, A. Rühlmann, W. H. Mueller, and R. J. Scheibe.** 1997. Functional interference between retinoic acid or steroid hormone receptors and the oncoprotein Fli-1. *Oncogene* **15**:3067-3082.

**Degnan, B. M., S. M. Degnan, T. Nagamura, and D. E. Morse.** 1993. The *ets* multigene family is conserved throughout the Metazoa. *Nucl. Ac. Res.* **21**:3479-3484.

**de Launoit, Y., J. L. Baert, A. Chotteau, D. Monté, P. A. Defosse, L; Coutte, H. Pelzar, M. P. Laget and F. Leenders.** 1996. Characterization of the PEA3 group of Ets-related transcription factors : role in breast cancer. *Breast Cancer. Adv. Biol. Therap.* 115-122

**Delattre, O., J. Zucman, B. Plougastale, C. Desmaze, T. Melot, M. Peter, H. Kovar, I. Joubert, P. De Jonge, G. Rouleau, A. Aurias, and G. Thomas.** 1992. Gene fusion with an ETS DNA-binding domain caused by chromosome translocation in human tumours. *Nature* **359**:162-165.

**Dhordain, P., F. Dewitte, X. Desbiens, D. Stéhelin, and M. Duterque-Coquillaud.** 1995. Mesodermal expression of the chicken *erg* gene associated with precartilaginous condensation and cartilage differentiation. *Mech.of Dev.* **50**:17-23.

- Donaldson, L. W., J. M. Petresen, B. J. Graves, and L. P. McIntosh.** 1996. Solution structure of the ETS-domain from murine Ets-1 : a winged helix-turn-helix DNA binding motif. *Embo J.* **15**:125-134.
- Donaldson, L. W., J. M. Petersen, B. J. Graves, and L. P. McIntosh.** 1994. Secondary structure of the ETS domain places murine Ets-1 in the superfamily of winged helix-turn-helix DNA-binding proteins. *Biochemistry* **33**:13509-13516.
- Duterque-Coquillaud, M., C. Niel, S. Plaza, and D. Stéhelin.** 1993. New human Erg isoforms generated by alternative splicing are transcriptional activators. *Oncogene* **8**:1865-1873.
- Eisenbeis, C. F., H. Singh, and U. Storb.** 1995. Pip, a novel IRF family member, is a lymphoid-specific, PU.1-dependent transcriptional activator. *Genes & Dev.* **9**:1377-1387.
- Ely, K. R., and R. Kodandapani.** 1998. Ankyrin(g) ETS domains to DNA. *Nature Struct. Biol.* **5**:255-259.
- Fisher, R. J., G. Mavrothalassitis, A. Kondoh and T. S. Papas.** 1991. High-affinity DNA-protein interactions of the cellular Ets-1 protein : the determination of the ETS binding motif. *Oncogene* **6**:2249-2254.
- Franza, B. R., F. J. Rauscher III, S. F. Josephs, and T. Curran.** 1988. The Fos complex and Fos-related antigen recognize sequence elements that contain AP-1 binding sites. *Science* **239**:1150-1153.
- Gack, S., R. Vallon, J. Schmidt, A. Grigoriadis, J. Tuckermann, J. Schenkel, H. Weiher, E. F. Wagner, and P. Angel.** 1995. Expression of interstitial collagenase during skeletal development of the mouse is restricted to osteoblast-like cells and hypertrophic chondrocytes. *Cell Growth & Differ.* **6**:759-767.
- Gañan, Y., D. Macias, D.-C. M., M. A. Ros, and J. M. Hurle.** 1996. Role of TGFβs and BMPs as signals controlling the position of the digits and the areas of interdigital cell death in the developing chick limb autopod. *Development* **122**:2349-2357.
- Gauthier, J. M., B. Bourachot, V. Doucas, M. Yaniv, and F. Moreau-Gachelin.** 1993. Functional interference between the Spi-1/Pu.1 oncoprotein and steroid hormone or vitamin receptors. *EMBO J.* **12**:5089-5096.
- Gégonne, A., D. Leprince, M. Duterque-Coquillaud, B. Vandebunder, A. Flourens, J. Ghysdael, B. Debuire, and D. Stéhelin.** 1987. Multiple domains for chicken cellular sequences homologous to the *v-ets* oncogenes of the E26 retrovirus. *Mol. Cell. Biol.* **7**:806-812.
- Gégonne, A., B. Punyammalae, B. Rabault, R. Bosselut, S. Seneca, M. Crabeel, and J. Ghysdael.** 1992. Analysis of the DNA binding and transcriptional activation properties of the Ets-1 oncoprotein. *The New Biologist* **4**:512-519.
- Giese, K., C. Kingsley, J. R. Kirshner, and R. Grosschedl.** 1995. Assembly and function of a TCRα enhancer complex is dependent on LEF-1 induced DNA bending and multiple protein-protein interactions. *Genes & Dev.* **9**:995-1008.

- Glover, J. N. M., and S. C. Harrison.** 1995. Crystal structure of the heterodimeric bZIP transcription factor c-Fos/c-Jun bound to DNA. *Nature* **373**:257-260.
- Golay, J., M. Introna, and T. Graf.** 1988. A single point mutation in the *v-ets* oncogene affects both erythroid and myelomonocytic cell differentiation. *Cell* **55**:1147-1158.
- Golub, T. R., G. F. Barker, M. Lovett, and D. G. Gilliland.** 1994. Fusion of PDGF receptor beta to a novel *ets*-like gene, *tel*, in chronic myelomonocytic leukemia with t(5;12) chromosomal translocation. *Cell* **77**:307-316.
- Graves, B.** 1998. Inner workings of a transcription factor partnership. *Science* **279**:1000-1002.
- Graves, B. J., and J. M. Petersen.** 1998. Specificity within the Ets family of transcription factors. *Adv. Cancer Res.* **75**:in press.
- Grévin, D., J. H. Chen, M. B. Raes, D. Stéhelin, B. Vanderbunder, and X. Desbiens.** 1993. Involvement of the proto-oncogene *c-ets-1* and the urokinase plasminogen activator during mouse implantation and placentation. *Int. J. Dev. Biol.* **37**:519-529.
- Gunster, M. J., D. P. E. Satijn, K. M. Hamer, J. L. den Blaauwen, D. de Bruijn, M. J. Alkema, M. van Lohuizen, R. van Driel, and A. P. Otte.** 1997. Identification and characterization of interactions between the vertebrate Polycomb-group protein BMI1 and human homologs of polyhomeotic. *Mol. Cell. Biol.* **17**:2326-2335.
- Gunther, C. V., J. A. Nye, R. S. Bryner, and B. J. Graves.** 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 B. Wasylyk.** 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.
- Hagman, J., and R. Grosschedl.** 1992. An inhibitory carboxy-terminal domain in Ets-1 and Ets-2 mediates differential binding of Ets family factors to promoter sequences of the *mb-1* gene. *Proc. Natl. Acad. Sci. USA* **89**:8889-8893.
- Hahn, S. L., P. Criqui, and B. Wasylyk.** 1997. Modulation of Ets-1 transcriptional activity by huUBC9, a ubiquitin-conjugating enzyme. *Oncogene* **15**:1489-1495.
- Hall, B. K., and T. Miyake.** 1995. Divide, accumulate, differentiate : cell condensation in skeletal development revisited. *Int. J. Dev. Biol.* **39**:881-893.
- Hart, A. H., C. M. Corrick, M. J. Tymms, P. J. Hertzog, and I. Kola.** 1995. Human *erg* is a proto-oncogene with mitogenic and transforming activity. *Oncogene* **10**.

- Heery, D. M., E. Kalkhoven, S. Hoare, and M. G. Parker.** 1997. A signature motif in transcriptional co-activators mediates binding to nuclear receptors. *Nature* **387**:733-736.
- Hilberg, F., A. Aguzzi, N. Howells, and E. F. Wagner.** 1993. c-Jun is essential for normal development and hepatogenesis. *Nature* **365**:179-181.
- Himelstein, B. P., E. J. Lee, H. Sato, M. Seiki, and R. J. Muschel.** 1998. Tumor cell contact mediated transcriptional activation of the fibroblast matrix metalloproteinase-9 gene : involvement of multiple transcription factors including Ets and an alternative purine-pyrimidine repeat. *Clin. Exp. Metastasis* **16**:167-177.
- Hipskind, R. A., V. N. Rao, C. G. F. Mueller, E. S. P. Reddy, and A. Nordheim.** 1991. Ets-related protein Elk-1 is homologous to the c-fos regulatory factor p62<sup>TCF</sup>. *Nature* **354**:531-534.
- Hodge, D. R., L. Robinson, D. Watson, J. Lautenberger, X. K. Zhang, M. Venanzoni, and A. Seth.** 1996. Interaction of Ets-1 and ErgB/Fli-1 proteins with DNA is modulated by spacing between multiple binding sites as well as phosphorylation. *Oncogene* **12**:11-18.
- Ichikawa, H., K. Shimizu, Y. Hayashi, and M. Ohki.** 1994. An RNA-binding protein gene, *TLS/FUS*, is fused to *erg* in human myeloid leukemia with t(16 ; 21) chromosomal translocation. *Canc. Res.* **54**:1-4.
- Jeon, I. S., J. N. Davis, B. S. Braun, J. E. Sublett, M. F. Roussel, C. T. Denny, and D. N. Shapiro.** 1995. A variant Ewing's sarcoma translocation (7;22) fuses the *EWS* gene to the Ets gene *ETV1*. *Oncogene* **10**:1229-1234.
- John, S., R. Marais, R. Child, Y. Light, and W. J. Leonard.** 1996. Importance of the low affinity Elf-1 sites in the regulation of lymphoid-specific inducible gene expression. *J. Exp. Med.* **183**:743-750.
- Johnson, R. S., B. M. Spiegelman, and V. Papaioannou.** 1992. Pleiotropic effects of a null mutation in the *c-fos* proto-oncogene. *Cell* **71**:577-586.
- Johnson, R. S., B. van Lingen, V. Papaioannou, and B. M. Spiegelman.** 1993. A null mutation of the *c-jun* locus causes embryonic lethality and retarded cell growth in culture. *Genes & Dev.* **7**:1309-1317.
- Jousset, C., C. Carron, A. Boureux, C. Tran Quang, C. Oury, -. F. Dusanter, I., M. Charon, J. Levin, O. Bernard, and J. Ghysdael.** 1997. A domain of Tel conserved in a subset of Ets proteins defines a specific oligomerization interface essential to the mitogenic properties of the Tel-PDGFR $\beta$  oncoprotein. *EMBO J.* **16**:69-82.
- Kaelin Jr, W. G., W. Krek, W. R. Sellers, J. A. De Caprio, F. Ajchenbaum, C. S. Fuchs, T. Chittenden, Y. Li, P. J. Farnham, M. A. Blnar, D. M. Livingstone, and E. K. Flemington.** 1992. Expression cloning of a cDNA encoding a retinoblastoma-binding protein with E2F-like properties. *Cell* **70**:351-364.



- Kalkhoven, E., J. E. Valentine, D. M. Heery, and M. G. Parker.** 1998. Isoforms of steroid receptor co-activator 1 differ in their ability to potentiate transcription by oestrogen receptor. *EMBO J.* **17**:232-243.
- Kaneko, Y., K. Yoshida, M. Handa, Y. Toyoda, H. Nishihira, Y. Tanaka, Y. Sasaki, S. Ishida, F. Higashino, and K. Fujinaga.** 1996. Fusion of an *ets*-family gene, *elaf*, to *EWS* by t(17;22)(q12;q12) chromosome translocation in an undifferentiated sarcoma of infancy. *Genes Chr. Canc.* **15**:115-121.
- Karim, F. D., L. D. Urness, C. S. Thummel, M. J. Klemsz, S. R. McKercher, A. Celada, C. Van Beveren, C. Maki, R. A. Gunther, J. A. Nye, and B. Graves.** 1990. The ETS domain : a new DNA-binding motif that recognizes a purine-rich core DNA sequence. *Genes & Dev.* **4**:1451-1453.
- Karin, M., Z. Liu, and E. Zandi.** 1997. AP-1 function and regulation. *Curr. Opin. Cell. Biol.* **9**:240-246.
- Kerppola, T. K., and T. Curran.** 1991a. DNA bending by Fos and Jun : the flexible hinge model. *Science* **254**:1210-1214.
- Kerppola, T. K., and T. Curran.** 1991b. Fos-Jun heterodimers and Jun homodimers bend DNA in opposite orientations : implications for transcription factor cooperativity. *Cell* **66**:317-326.
- 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.
- Klemsz, M., and R. Maki.** 1996. Activation of transcription by PU-1 requires both acidic and glutamine domains. *Moll. Cell. Biol.* **16**:390-397.
- Knauper, V., H. Will, C. Lopez-Otin, B. Smith, S. J. Atkinson, H. Stanton, R. M. Hembry, and G. Murphy.** 1996. Cellular mechanisms for human procollagenase3 (MMP13) activation. Evidence that the MT1-MMP (MMP14) et gelatinaseA (MMP2) are able to generate active enzyme. *J. Biol. Chem.* **271**:17124-17131.
- Kodandapani, R., F. Pio, C. Z. Ni, G. Picialli, M. Klemsz, S. McKercher, R. A. Maki, and K. R. Ely.** 1996. A new pattern for helix-turn-helix recognition revealed by the PU-1 ETS-domain-DNA complex. *Nature* **380**:456-460.
- Laget, M. P., P. A. Defosse, O. Albagli, J. L. Baert, F. Dewitte, D. Stéhelin, and Y. de Launoit.** 1996. Two functionally distinct domains responsible for transactivation by the Ets family member ERM. *Oncogene* **12**:1325-1336.
- Lamarco, K., C. C. Thompson, B. P. Byers, E. M. Walton, and S. L. Mcknight.** 1991. Identification of Ets- and Notch-related subunits in GA binding protein. *Science* **253**:789-792.
- Laudet, V., C. Niel, M. Duterque-Coquillaud, D. Leprince and D. Stéhelin.** 1993. Evolution of the *ets* family. *Biochem. Biophys. Res. Commun.* **190**:8-14.
- Laudet, V., C. Hänni, D. Stéhelin, and M. Duterque-Coquillaud.** Molecular phylogeny of the *ets* gene family. *Oncogene in press.*

- Lautenberger, J. A., L. A. Burdett, M. A. Gunnell, S. Qi, D. K. Watson, S. J. O'Brien, and T. S. Papas.** 1992. Genomic dispersal of the *ets* gene family during metazoan evolution. *Oncogene* **7**:1713-1719.
- Lee, W., A. Haslinger, M. Karin, and R. Tjian.** 1987a. Activation of transcription by two factors that bind promoter and enhancer sequences of the human *metallothionein* gene and SV40. *Nature* **325**:368-372.
- Lee, W., P. Mitchell, and R. Tjian.** 1987b. Purified transcription factor AP-1 interacts with TPA-inducible enhancer elements. *Cell* **49**:741-752.
- Leprince, D., A. Gégonne, J. Coll, C. De Taisne, A. Schneeberger, C. Lagrou, and D. Stéhelin.** 1983. A putative second cell-derived oncogene of the avian leukaemia retrovirus E26. *Nature* **306**:395-397.
- Liang, H., X. Mao, E. T. Olejniczk, D. G. Nettesheim, L. Yu, R. P. Meadow, C. B. Thompson, and S. W. Fesik.** 1994b. Solution structure of the ETS-domain of FLI-1 when bound to DNA. *Nature Struct. Biol.* **1**:871-875.
- Liang, H., E. T. Olejniczak, X. Mao, D. G. Nettesheim, L. Yu, C. B. Thompson, and S. W. Fezik.** 1994a. The secondary structure of the ETS domain of human Fli-1 resembles that of the helix-turn-helix DNA-binding motif of the Escherichia coli catabolite gene activator protein. *Proc. Natl. Acad. Sci. USA* **91**:11655-11659.
- Macias, D., Y. Gañan, T. K. Sampath, M. E. Piedra, M. A. Ros, and J. M. Hurle.** 1997. Role of BMP-2 and OP-1 (BMP7) in programmed cell death and skeletogenesis during chick limb development. *Development* **124**:1109-1117.
- Magnaghi-Jaulin, L., H. Masutani, P. Robin, M. Lipinski, and A. Harel-Bellan.** 1996. SRE element are binding sites for the fusion protein EWS-FLI-1. *Nucl. Ac. Res.* **24**:1052-1058.
- Maira, S. M., J. M. Wurtz, and B. Wasylyk.** 1996. Net(ERP/SAP2), one of the Ras-inducible TCFs, has a novel inhibitory domain with resemblance to the helix-loop-helix motif. *EMBO J.* **15**:5849-5865.
- Majka, S. M., and P. G. McGuire.** 1997. Regulation of urokinase expression in the developing avian heart : a role for the Ets-2 transcription factor. *Mech. of Dev.* **68**:127-137.
- Mao, X., S. Miesfeld, H. Yang, J. M. Leiden, and C. B. Thompson.** 1994. The Fli-1 and chimeric EWS-Fli-1 oncoproteins display similar DNA-binding specificities. *J. Biol. Chem.* **269**:18216-18222.
- Maroulakou, I. G., T. S. Papas, and J. E. Green.** 1994. Differential expression of *ets-1* and *ets-2* proto-oncogenes during murine embryogenesis. *Oncogene* **9**:1551-1565.
- May, W. A., S. L. Lessnick, B. S. Braun, M. Klemsz, B. C. Lewis, L. B. Lunsford, R. Hromas, and C. T. Denny.** 1993. 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*. *Mol. Cell. Biol.* **13**:7393-7398.
- McKercher, S. R., B. E. Torbett, K. L. Anderson, G. W. Henkel, D. J. Vestal, H. Baribault, M. Klemsz, A. J. Feeney, G. E. Wu, C. J. Paige,**

and R. A. Maki. 1996. Targeted disruption of the Pu-1 gene results in multiple hematopoietic abnormalities. *Embo J.* **15**:5647-5658.

Mélet, F., B. Motro, D. J. Rossi, L. Zhang, and A. Bernstein. 1996. Generation of a novel Fli-1 protein by gene targeting leads to a defect in thymus development and a delay in Friend virus-induced erythroleukemia. *Mol. Cell. Biol.* **16**:2708-2718.

Moreau, A., W. V. Yotov, F. H. Glorieux, and R. St-Arnaud. 1998. Bone-specific expression of the alpha chain of the nascent polypeptide-associated complex, a coactivator potentiating c-Jun mediated transcription. *Mol. Cell. Biol.* **18**:1312-1321.

Moreau-Gachelin, F., F. Wendling, T. Molina, N. Denis, M. Titeux, G. Grimber, P. Briand, W. Vainchenker, and A. Tavitian. 1996. Spi-1/Pu-1 transgenic mice develop multistep erythroleukemias. *Mol. Cell. Biol.* **16**:2453-2463.

Murakami, K., G. Mavrothalassitis, N. K. Bhat, R. J. Fisher, and T. S. Papas. 1993. Human Erg-2 is a phosphorylated DNA-binding protein, a distinct member of the Ets family. *Oncogene* **8**:1559-1566.

Muthusamy, N., K. Barton, and J. M. Leiden. 1995. Defective activation and survival of T cells lacking the Ets-1 transcription factor. *Nature* **377**:639-642.

Nakae, K., K. Nakajima, J. Inazawa, T. Kitaoka, and T. Hirano. 1995. Erm, a PEA3 subfamily of Ets transcription factors, can cooperate with c-Jun. *J. Biol. Chem.* **270**:23795-23800.

Nunn, M. F., P. H. Seeburg, C. Moscovici, and P. H. Duesberg. 1983. Tripartite structure of the avian erythroblastosis virus E26 transforming gene. *Nature* **306**:391-395.

Nye, J. A., J. M. Petersen, C. V. Gunther, M. D. Jonsen, and G. B. J. 1992. Interaction of murine Ets-1 with GGA-binding sites establishes the ETS domain as a new DNA-binding motif. *Genes & Dev.* **6**:975-990.

Oettgen, P., R. M. Alani, A. Barcinski, L. Brown, Y. Akbarali, J. Boltax, C. Kunsch, K. Mungler, and T. A. Linermann. 1997. Isolation and characterization of a novel epithelium-specific transcription factor, ESE-1, a member of the Ets family. *Mol. Cell. Biol.* **17**:4419-4433.

Ohno, T., M. Ouchida, L. Lee, Z. Gatalica, V. N. Rao, and E. S. P. Reddy. 1994. The EWS gene, involved in Ewing family of tumors, malignant, melanoma of soft parts and desmoplastic small round cell tumors, codes for an RNA binding protein with novel regulatory domains. *Oncogene* **9**:3087-3097.

Papas, T. S., R. J. Fisher, N. Bhat, S. Fujiwara, D. K. Watson, J. Lautenberger, A. Seth, Z. Q. Chen, L. Burdett, L. Pribyl, C. W. Schweinfest, and R. Ascione. 1989. The *ets* family of genes : molecular biology and functional implications. *Curr. Topics Microbiol. and Immunol.* **149**:1159-1164.

Perkel, J. M., and M. L. Atchinson. 1998. A two-step mechanism for recruitment of Pip by Pu-1. *J. Immunol.* **160**:241-252.



- Peter, M., J. Couturier, H. Pacquement, J. Michon, G. Thomas, H. Magdelenat, and O. Delattre.** 1997. A new member of the Ets family fused to EWS in Ewing tumors. *Oncogene* **14**:1159-1164.
- Peterson, A. J., M. Kyba, D. Bornemann, K. Morgan, H. W. Brock, and J. Simon.** 1997. A domain shared by the *Polycomb* group proteins Scm and ph mediates heterotypic and homotypic interactions. *Mol. Cell. Biol.* **17**:6683-6692.
- Pongubala, J. M. R., S. Nagulapalli, M. J. Klemsz, S. R. Mckercher, R. A. Maki, and M. L. Atchison.** 1992. PU.1 recruits a second nuclear factor to a site important for immunoglobulin k 3' enhancer activity. *Mol.Cell.Biol.* **12**:368-378.
- Powers, C., H. Krutzch, and K. Gardner.** 1996. Modulation of JunD/AP-1 DNA binding activity by AP-1 associated factor 1 (AF-1). *J. Biol. Chem.* **271**:30089-30095.
- Prasad, D. D. K., V. N. Rao, L. Lee, and E. S. P. Reddy.** 1994. Differentially spliced *erg-3* product functions as a transcriptional activator. *Oncogene* **9**:669-673.
- Quéva, C., D. Leprince, D. Stéhelin, and B. Vandenbunder.** 1993. p54c-*ets-1* and p68c-*ets-1*, 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.
- Rao, V. N., T. Ohno, D. D. K. Prasad, G. Bhattacharya, and E. S. P. Reddy.** 1993. Analysis of the DNA-binding and transcriptional activation functions of human Fli-1 protein. *Oncogene* **8**:2167-2173.
- Rao, V. N., T. S. Papas, and E. S. Reddy.** 1987. *erg*, a human *ets*-related gene on chromosome 21 : alternative splicing, polyadenylation, and translation. *Science* **237**:635-639.
- Rauscher III, F. J., D. R. Cohen, T. Curran, T. J. Bos, P. K. Vogt, D. Bohmann, R. Tjian, and B. R. Franza.** 1988. Fos-associated protein p39 is the product of the *jun* prot-oncogene. *Science* **240**:1010-1016.
- Ray-Gallet, D., C. Mao, A. Tavitian, and F. Moreau-Gachelin.** 1995. DNA binding specificities of Spi-1/Pu-1 and Spi-B transcription factors and identification of a Spi-1/Spi-B binding site in the *c-fes/c-fps* promoter. *Oncogene* **11**:303-313.
- Rebay, I., and G. M. Rubin.** 1995. Yan functions as a general inhibitor of differentiation and is negatively regulated by activation of the Ras1/MAPK pathway. *Cell* **81**:857-866.
- Reboul, P., J. P. Pelletier, G. Tardif, J. M. Cloutier, and J. Martel-Pelletier.** 1996. The new collagenase, collagenase3, is expressed and synthesized by human chondrocytes but not synoviocytes. *J. Clin. Invest.* **97**:2011-2019.
- Reddy, E. S., V. N. Rao, and T. S. Papas.** 1987. The *erg* gene : a human gene related to the *ets* oncogene. *Proc. Natl. Acad. Sci. USA* **84**:6131-6135.
- Reddy, E. S. P., and V. N. Rao.** 1991. *erg*, an *ets*-related gene, codes for sequence-specific transcriptional activators. *Oncogene* **6**:2285-2289.

- Rivera, R. R., M. H. Stuver, R. Steenberger, and C. Murre.** 1993. Ets proteins : new factors that regulate immunoglobulin heavy-chain gene expression. *Mol. Cell. Biol.* **13**:7163-7169.
- Schneikert, J., H. Peterziel, P. A. Defosse, H. Klocker, Y. De Launoit, and A. C. B. Cato.** 1996. Androgen receptor-Ets protein interaction : a novel mechanism for steroid hormone mediated down modulation of matrix-metalloproteinase expression. *J. Biol. Chem.* **271**:23907-23913.
- Schneikert, J., Y. Lutz, and B. Wasylyk.** 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.
- Schwachtgen, J. L., N. Janel, L. Barek, M. Duterque-Coquillaud, J. Ghysdael, D. Meyer, and D. Kerbiriou-Nabias.** 1997. Ets transcription factors bind and transactivate the core promoter of the von Willerbrand factor gene. *Oncogene* **15**:3091-3102.
- Scott, E. W., R. C. Fischer, M. C. Olson, E. W. Kehrl, M. C. Simon, and H. Singh.** 1997. Pu-1 functions in a cell-autonomous manner to control the differentiation of multipotential lymphoid-myeloid progenitors. *Immunity* **6**:437-447.
- Scott, E. W., M. C. Simon, J. Anastasi, and H. Singh.** 1994. Requirement of transcription factor Pu-1 in the development of multiple hematopoietic. *Science* **265**:1573-1577.
- Scully, A. L., M. H. Sommer, R. Schwartz, and D. H. Spector.** 1995. The human cytomegalovirus IE2 86kD protein interacts with an early gene promoter *via* site-specific DNA binding and protein-protein associations. *J. Virol.* **69**:6533-6540.
- Seth, A., and T. S. Papas.** 1990. The *c-ets-1* proto-oncogene has oncogenic activity and is positively autoregulated. *Oncogene* **5**:1761-1767.
- Sgouras, D. N., M. A. Athanasiou, G. J. Beal Jr, R. J. Fischer, D. G. Blair, and G. J. Mavrothalassitis.** 1995. ERF : an ETS domain protein with strong transcriptional repressor activity, can suppress *ets*-associated tumorigenesis and is regulated by phosphorylation during cell cycle and mitogenic stimulation. *Embo J.* **14**:4781-4793.
- Shimizu, K., H. Ichikawa, A. Tojo, Y. Kaneno, N. Maseki, Y. Hayashi, M. Ohira, S. Asano, and M. Ohki.** 1993. An *ets*-related gene, *erg*, is rearranged in human myeloid leukemia with t(16 ; 21) chromosomal translocation. *Proc. Natl. Acad. Sci. USA* **90**:10280-10284.
- Shore, P., and A. D. Sharrocks.** 1995. The ETS-domain transcription factors Elk-1 and SAP-1 exhibit differential DNA binding specificities. *Nucl. Ac. Res.* **23**:4698-4706.
- Siddique, H. R., V. N. Rao, L. Lee, and E. S. P. Reddy.** 1993. Characterization of the DNA binding and transcriptional activation domains of the Erg protein. *Oncogene* **8**:1751-1755.
- Sieweke, M. H., H. Tekotte, J. Frampton, and T. Graf.** 1996. MafB is an interaction partner and repressor of Ets-1 that inhibits erythroid differentiation. *Cell* **85**:49-60.

- Sieweke, M. H., H. Tekotte, U. Jarosch, and T. Graf.** 1998. Cooperative interaction of Ets-1 with USF-1 for HIV-1 enhancer activity in T cells. *EMBO J.* **17**:1728-1739.
- Sitlani, A., and D. M. Crothers.** 1998. DNA-binding domains of Fos and Jun do not induce DNA curvature : an investigation with solution and gels methods. *Biochemistry* **95**:1404-1409.
- Sitlani, A., and D. M. Crothers.** 1996. Fos and Jun do not bend the AP-1 recognition site. *Proc. Natl. Acad. Sc. USA* **93**:3248-3252.
- Sorensen, P. H. B., S. L. Lessnick, D. Lopez-Terrada, X. F. Liu, T. J. Triche, and C. T. Denny.** 1994. A second Ewing's sarcoma translocation, t(21;22), fuses the *EWS* gene to another Ets family transcription factor, *erg*. *Nature Genetics* **6**:146-151.
- Su, G. H., H. M. Chen, N. Muthusamy, L. A. Garrett-Sinha, D. Baunoch, D. G. Tene, and M. C. Simon.** 1997. Defective B cell receptor-mediated responses in mice lacking the Ets protein, Spi-B. *Embo J.* **16**:7118-71298.
- Sumarsono, S. H., T. J. Wilson, M. J. Tymms, D. J. Venter, C. M. Corrick, R. Kola, M. H. Lahoud, T. S. Papas, A. Seth, and I. Kola.** 1996. Down's syndrome-like skeletal abnormalities in Ets-2 transgenic mice. *Nature* **379**:534-537.
- Taggart, A. K. P., and B. F. Pugh.** 1996. Dimerization of TFIIID when not bound to DNA. *Science* **272**:1331-1333.
- Telford, N. A., A. Hogan, C. R. Franz, and A. Schultz.** 1990. Expression of genes for insulin and insuline-like growth factors and receptors in early postimplantation mouse embryos and embryonal carcinoma cells. *Mol. Reprod. Dev.* **27**:81-92.
- Thomas, R. S., M. J. Tymms, L. H. McKinlay, M. F. Shannon, A. Seth, and I. Kola.** 1997. Ets-1, NFkB and AP1 synergistically transactivate the human GM-CSF promoter. *Oncogene* **14**:2845-2855.
- Thomas, R. S., M. J. Tymms, A. Seth, M. F. Shannon, and I. Kola.** 1995. Ets-1 transactivates the human GM-CSF promoter in Jurkat T cells stimulated with PMA and ionomycin. *Oncogene* **11**:2135-2143.
- Thompson, C. C., T. A. Brown, and S. L. Mcknight.** 1991. Convergence of Ets- and Notch-related structural motifs in a heteromeric DNA binding complex. *Science* **253**:762-768.
- Tondravi, M. M., S. R. McKercher, K. Anderson, J. M. Erdmann, M. Quirroz, R. Maki, and S. L. Teitelbaum.** 1997. Osteopetrosis in mice lacking haematopoietic transcription factor PU.1. *Nature* **386**:81-84.
- Torchia, J., D. W. Rose, J. Inostroza, Y. Kamei, S. Westin, C. K. Glass, and M. G. Rosenfeld.** 1997. The transcriptional co-activator p/CBP binds CBP and mediates nuclear-receptor function. *Nature* **387**:677-684.

- Urness, L. D., and C. S. Thummel. 1990. Molecular interactions within the ecdysone regulatory hierarchy: DNA binding properties of the *Drosophila* ecdysone-inducible E74A protein. *Cell* **63**:47-61.
- Van Dam, H., M. Duyndam, R. Rottier, A. Bosch, L. De Vries-Smits, P. Herrlich, A. Zantema, P. Angel, and A. Van der AB. 1993. Heterodimer formation of c-Jun and AYP-2 is responsible for induction of c-jun by the 243 amino-acid adenovirus E1A protein. *EMBO J.* **12**:479-487.
- Venanzoni, M. C., L. R. Robinson, D. R. Hodge, I. Kola, and A. Seth. 1996. Ets-1 and Ets-2 in p53 regulation : spatial separation of ETS binding sites (EBS) modulate protein : DNA interaction. *Oncogene* **12**:1199-1204.
- Wang, C. Y., B. Petryniak, I. C. Ho, C. B. Thompson, and J. M. Leiden. 1992. Evolutionary conserved Ets family members display distinct DNA binding specificities. *J. Exp. Med.* **175**:1391-1399.
- Wang, L. C., F. Kuo, Y. Fujiwara, D. G. Gilliland, T. R. Golub, and S. H. Orkin. 1997. Yolk sac angiogenic defect and intra-embryonic apoptosis in mice lacking the Ets-related factor Tel. *EMBO J.* **16**:4374-4383.
- Wang, Z. Q., K. Schellander, E. F. Wagner, and A. E. Grigoriadis. 1995. *c-fos* induced osteosarcoma formation in transgenic mice : cooperativity with *c-jun* and the role of endogenous *c-fos*. *Cancer Res.* **55**:6244-6251.
- Wasylyk, B., C. Wasylyk, P. Flores, A. Bègue, D. Leprince, and D. Stéhelin. 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., A. Gutman, R. Nicholson, and B. Wasylyk. 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., J. P. Kerckaert, and B. Wasylyk. 1992. A novel modulator domain of Ets transcription factors. *Genes & Dev.* **6**:965-974.
- Watson, D. K., M. J. McWilliams-Smith, M. F. Nunn, P. H. Duesberg, S. J. O'Brien, and T. S. Papas. 1985. The est 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. Sc. USA* **82**:7294-7298.
- Watson, D. K., F. E. Smyth, D. M. Thompson, J. Q. Cheng, J. R. Testa, T. S. Papas, and A. Seth. 1992. The *ergBfl1-1* gene : isolation and characterization of a new member of the family of human Ets transcription factors. *Cell Growth & Diff.* **3**:705-713.
- Werner, M. H., G. M. Clore, C. L. Fisher, R. J. Fischer, L. Trinh, J. Shiloach, and A. M. Gronenborn. 1997. Correction of the NMR structure of the Ets-1/DNA complex. *J. Bio. NMR* **10**:317-328.
- Werner, M. H., G. M. Clore, C. L. Fisher, R. J. Fischer, L. Trinh, J. Shiloach, and A. M. Gronenborn. 1995. The solution structure of the human Ets-1/DNA complex reveals a novel mode of binding and true side chain intercalation. *Cell* **83**:761-771 ; erratum, *Cell* (1996) 87.

Wernert, N., F. Gilles, V. Fafeur, F. Bouali, M. B. Raes, C. Pyke, T. Dupressoir, G. Seitz, B. Vandebunder, and D. Stéhelin. 1994. Stromal expression of c-ets-1 transcription factor correlates with tumor invasion. *Cancer Res.* **54**:5683-5688.

Wilkinson, D. A., G. A. M. Neale, S. Mao, C. W. Naeve, and R. M. Goorha. 1997. Elf-2, a rhombotin-2 binding Ets transcription factor : discovery and potential role in T cell leukemia. *Leukemia* **11**:86-96.

Yang, Q., R. Bassel-Duby, and R. S. Williams. 1997. Transient expression of a winged-helix protein, MNF- $\beta$ , during myogenesis. *Mol. Cell. Biol.* **17**:5236-5243.

Zhang, H., G. Hu, H. Wang, P. Sciavolino, N. Iler, M. M. Shen, and C. Abate-Shen. 1997. Heterodimerization of Msx and Dlx homeoproteins results in functional antagonism. *Mol. Cell. Biol.* **17**:2920-2932.

Zhang, L., A. Eddy, Y. T. Teng, M. Fritzler, M. Kluppel, F. Melet, and A. Bernstein. 1995. An immunological renal disease in transgenic mice that overexpress *Fli-1*, a member of the *ets* family of transcription factor genes. *Mol. Cell. Biol.* :6961-6970.

Zucman, J., T. Melot, C. Desmaze, J. Ghysdael, B. Plougastel, M. Peter, J. M. Zucker, T. J. Triche, D. Sheer, C. Turc-Carel, P. Ambros, V. Combaret, G. Lenoir, A. Aurias, G. Thomas, and O. Delattre. 1993. Combinatorial generation of variable fusion proteins in the Ewing family of tumours. *Embo J.* **12**:4481-4487.

175 références

