

50376
1996
326

Université des Sciences et Technologies de Lille

THESE

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

par Pierre-Antoine DEFOSSEZ

**MECANISMES D'ACTIVATION TRANSCRIPTIONNELLE
PAR ERM, UN NOUVEAU MEMBRE DE LA FAMILLE ETS**



Soutenue le 16 décembre 1996 devant la commission d'examen:

Président : Pr Xavier Desbiens
Rapporteurs : Dr Olivier Delattre
Dr Laszlo Tora
Examineurs : Dr Yvan de Launoit
Dr Dominique Leprince
Pr Dominique Stéhelin

Je voudrais d'abord remercier les membres du Jury, Président, Rapporteurs et Examineurs, qui ont accepté de juger ce travail.

Ma part de responsabilité dans leur calvaire est sans doute importante, mais rien n'aurait été possible sans les multiples complicités dont j'ai bénéficié. Nombreux sont ces "partners in crime" qui m'ont conseillé et/ou soutenu au cours des différentes étapes de ce travail.

En premier lieu, Dominique Stéhelin a accepté de m'accueillir dans son laboratoire. La conception et la réalisation des expériences doivent beaucoup aux conseils de Martin et d'Olivier Albagli. Ce dernier a été un juge parfois rude mais toujours avisé, et j'admire la richesse et la rigueur de ses vues scientifiques. Yvan et Jean-Luc m'ont toujours fait profiter de leur expérience. Le troisième expert du groupe, c'est bien sûr Anne Chotte, mon embryologiste préférée. Marie-Pierre a énormément contribué au lancement du sujet "transcription".

Le "club des thésards qui ont des hauts et des bas" (voire des mi-bas) a rendu la rédaction et l'édition de ce mémoire un peu moins difficile. Le stress est soluble dans le Bailey's, semble-t-il...

Enfin, une mention spéciale pour Guillaume qui m'a fait découvrir, pêle-mêle, l'illustration et les beautés de l'électroporation; le Cottage Cheese et l'organisation ergonomique des paillasses. Il est évident que sans lui cette thèse aurait été tout autre.

Bon courage Gromit, c'est bientôt ton tour...

A ma famille et mes amis.

Scientific progress goes "boink"
Hobbes

Résumé

La famille ETS est un ensemble de gènes codant des facteurs de transcription, dont on trouve des représentants chez tous les Eucaryotes métazoaires. Ses membres jouent des rôles fondamentaux dans la multiplication et la différenciation cellulaire. Certains d'entre eux sont des proto-oncogènes dont l'altération peut conduire au cancer.

Nous avons cloné un nouveau membre de la famille ETS, *erm*. Avec *pea3* et *er81*, deux gènes fort proches identifiés auparavant, il appartient au groupe PEA3 au sein des gènes ETS.

Nous avons étudié les mécanismes par lesquels Erm assure sa fonction de régulateur transcriptionnel. Son domaine ETS de 85 acides aminés suffit à assurer la liaison spécifique à l'ADN. Cette activité est inhibée de façon intramoléculaire par deux régions d'Erm.

Nous avons identifié deux domaines transactivateurs au sein d'Erm. AD1, situé côté N-terminal, est riche en résidus acides. Des tests d'interférence transcriptionnelle montrent qu'il est fonctionnellement apparenté aux domaines de transactivation acides comme celui de la protéine virale VP16. Le deuxième domaine transactivateur, AD2, est situé dans la queue C-terminale et se superpose à une région inhibitrice de la liaison à l'ADN. AD2 n'appartient pas à la classe des activateurs acides, et exerce un effet coopératif avec AD1.

Nous montrons que le transactivateur acide AD1 contient une hélice alpha d'une quinzaine d'acides aminés nécessaire à sa fonction. Il apparaît qu'une partie des domaines de transactivation acides contiennent une hélice alpha potentielle. Nous proposons que l'hétérogénéité fonctionnelle des transactivateurs acides reflète leur hétérogénéité structurale.

Enfin, l'activité de AD1 et AD2 est soumise à une inhibition intramoléculaire; nous suggérons qu'Erm peut présenter une conformation "fermée" dans laquelle il est inactif, et une conformation ouverte où les domaines fonctionnels sont accessibles. La transition entre ces états pourrait être modulée par modification post-traductionnelle ou interaction avec un cofacteur.

Table des matières

INTRODUCTION	10
Présentation du mémoire	11
1. Les gènes ETS sont des régulateurs de la différenciation et la multiplication cellulaire	13
1.1 Les protéines ETS dans le développement des photorécepteurs de drosophile	13
1.1.1 L'oeil de drosophile : un système génétique modèle	13
1.1.2 <i>pnt</i> et <i>yan</i> , deux membres antagonistes de la famille ETS	17
1.1.3 L'activité de Pnt et Yan est régulée par la kinase Rolled/Erk-A	17
1.1.4 <i>pnt</i> coopère avec <i>D-jun</i> pour induire la différenciation de R7	18
1.1.5 Quelques thèmes ETSiens	21
1.2 Protéines ETS et hématopoïèse	21
1.2.1 <i>spi-1/pu.1</i> est un gène essentiel au développement de différentes lignées hématopoïétiques	21
1.2.2 Ets-1, Ets-2 et l'activation des lymphocytes	26
1.3 Protéines ETS et remodelage tissulaire	28
1.3.1 Gènes ETS et migration cellulaire	28
1.3.2 Angiogenèse et vasculogenèse	29
1.3.3 Morphogenèse épithéliale	30
1.3.4 Les protéines ETS régulent la transcription de gènes impliqués dans la dégradation de la matrice extracellulaire et l'adhérence	30
1.3.5 L'induction de l'activité ETS peut être transcriptionnelle ou post-traductionnelle	32
2. Certains gènes ETS sont des proto-oncogènes (ou : quand la machine se détraque)	35
2.1 Spi-1, Fli-1, et les leucémies de Friend	35
2.1.1 <i>spi-1</i> est surexprimé dans les cellules infectées par les virus de Friend	35
2.1.2 L'expression de <i>spi-1</i> transforme les proérythroblastes	37
2.1.3 <i>fli-1</i> est surexprimé dans les leucémies induites par FMuLV	38
2.2 <i>v-ets</i> : un oncogène viral dérivé de <i>ets-1</i>	39
2.2.1 <i>v-ets</i> coopère avec <i>v-myb</i> pour induire des leucémies	39

2.2.2 v-Ets pourrait rentrer en compétition avec d'autres protéines ETS	41
2.2.3 Transformation par interférence transcriptionnelle	43
2.3 Quatre gènes ETS recombinaés dans les tumeurs d'Ewing	45
2.3.1 Les fusions EWS-ETS sont transformantes	45
2.3.2 Le pouvoir transformant des chimères EWS-Fli-1 est lié à la dérégulation de cibles de Fli-1	47
2.3.3 <i>erg</i> , <i>fli-1</i> , <i>er81</i> et <i>pea3</i> ont-ils des rôles équivalents dans les différents cas de sarcome d'Ewing?	50
3. Les mécanismes de régulation transcriptionnelle	53
3.1 Le remodelage de la chromatine	53
3.1.1 Les nucléosomes inhibent la transcription	53
3.1.2 Les complexes de remodelage de la chromatine	54
3.1.3 Acétylation et déacétylation des histones	58
3.1.4 Comment la machinerie de remodelage est-elle recrutée aux gènes induits ?	59
3.1.5 L'oeuf ou la poule ?	61
3.2 Le recrutement du complexe de préinitiation de transcription	61
3.2.1 Les multiples composants du complexe de préinitiation	61
3.2.2 Un paradigme : l'interaction TFIID-ADN	64
3.2.3 De multiples mécanismes permettent d'activer la formation du PIC	69
3.2.4 Effets de l'architecture du promoteur sur la préinitiation	71
3.3 Le contrôle de l'élongation transcriptionnelle	72
3.3.1 L'élongation peut être une étape limitante de la transcription des gènes	72
3.3.2 Des facteurs de transcription peuvent stimuler l'élongation	73
3.3.3 Comment la processivité vient à la polymérase	74
3.4 Que retenir de tout ça?	76
3.5 Post-scriptum : il y a une vie après l'élongation.	76
3.5.1 La régulation de l'épissage par des facteurs de transcription	77
3.5.2 Contrôle de la localisation des ARNs	77
3.5.3 Contrôle traductionnel par des facteurs de transcription	77
4. La fin du début	79

4.1 La spécificité peut dériver de la différence des patrons d'expression	79
4.2 La sélectivité de liaison à l'ADN distingue-t-elle les gènes ETS?	81
4.3 Les interactions protéine-protéine peuvent distinguer les protéines ETS	85
4.4 Le début de la fin	87
RESULTATS	88
Publication 1 : <u>clonage et caractérisation d'ERM, un nouveau membre de la famille ETS</u>	89
Discussion	100
1. Généalogie du groupe PEA3	100
2. Quelques éléments remarquables dans la séquence d'Erm	103
2.1 Certaines régions des protéines du groupe PEA3 ont peu divergé	103
2.2 Séquence PEST	103
2.3 Sites de modifications post-traductionnelles	104
2.3.1 Sites de phosphorylation par les kinases MAP	105
2.3.2 Sites de phosphorylation par la protéine kinase dépendante de l'AMPc	106
2.3.3 Sites de phosphorylation par la Caséine Kinase II	107
3. Carte d'expression comparée d'<i>erm</i>, <i>er81</i> et <i>pea3</i>	108
Publication 2 : <u>étude fonctionnelle des domaines transactivateurs d'ERM</u>	111
Discussion	125
1. L'inhibition intramoléculaire de l'activité d'Erm	125
1.1 Inhibition de la liaison à l'ADN	125
1.2 Inhibition du pouvoir transactivateur	126
1.3 L'activité d'Erm pourrait être régulée par modification de sa structure	128
2. Quel mécanisme pour la synergie AD1-AD2?	129
3. Conséquences "physiologiques" des différences AD1-AD2	130

Publication 3 : <u>analyse structurale du domaine transactivateur acide d'Erm</u>	133
Discussion	164
1. Il existe des différences structurales entre les domaines activateurs acides	164
1.1 Préambule historique	164
1.2 Quelle structure pour les domaines acides de GAL4 et GCN4?	164
1.3 Où en est-on?	165
2. Les différences structurales établissent-elles des classes fonctionnelles?	165
2.1 Certains activateurs interfèrent entre eux, mais ce n'est pas un renseignement précis	166
2.2 Il existe des différences fonctionnelles plus précises entre les domaines acides	168
2.3 Il est encore trop tôt pour établir un parallèle structure-fonction	170
3. Des domaines à structure non définie pourraient-ils activer la transcription?	171
CONCLUSION	172
REFERENCES	175

Liste des illustrations

1 Les gènes de la famille ETS, classés par degré de similitude	12
2 Des mutations des gènes ETS altèrent le développement de l'oeil de drosophile	14
3 Les acteurs de la différenciation du photorécepteur R7	16
4 L'activité de Pnt, Yan et D-Jun est modulée par Ras	20
5 Sur des promoteurs différents, l'activité transcriptionnelle de Pu.1 est différente	23
6 Ets-1 et Ets-2 pourraient jouer des rôles opposés dans l'activation des lymphocytes	27
7 Les protéines ETS régulent avec AP1 la transcription de gènes impliqués dans le remodelage tissulaire	34
8 La séquence des événements conduisant à la leucémie de Friend après infection par SFFV ou FMuLV	36
9 E26 transforme des cellules souches multipotentes, et requiert l'activité du domaine ETS	40
10 v-Ets diffère de Ets-1 par sa fusion à v-Myb, deux mutations ponctuelles, et la séquence de son extrémité C-terminale	42
11 p135 perturbe l'antagonisme AP1/RAR-T3R	44
12 Quatre gènes de la famille ETS sont fusionnés au gène EWS dans des sarcomes d'Ewing	46
13 Une hypothèse pour expliquer le pouvoir transformant d'EWS/Fli-1	48
14 Les étapes successives de la transcription peuvent être stimulées par les transactivateurs	52
15 Les complexes remodelant la chromatine chez la levure, l'homme et la drosophile	55
16 Sur le promoteur du gène hsp70, le facteur GAGA empêche la répression nucléosomale et la protéine HSF lève l'inhibition d'élongation par l'ARN polymérase II	57
17 Deux modèles pour le recrutement de la machinerie de remodelage de la chromatine et d'acétylation des histones	60
18 Deux modèles pour l'assemblage du complexe de préinitiation	62
19 Le complexe TFIID est la cible des activateurs	65
20 Un modèle pour le recrutement de TBP au promoteur via les TAFs	66
21 Un modèle triangulaire pour l'activation de l'initiation	70
22 Lef-1 courbe l'ADN de l'enhancer du TCR α	72
23 Les transactivateurs pourraient stimuler l'élongation via TFIIH	75
24 Quelques-uns des paramètres déterminant la fonction des gènes ETS	80
25 Les séquences optimales sélectionnées in vitro par divers membres de la famille ETS diffèrent peu	82
26 Les éléments de réponse aux protéines ETS observés dans les promoteurs présentent des séquences non-optimales	83
27 Trois régions des TCFs sont bien conservées	86
28 La différence de capacité à interagir avec SRF différencie Elk et Δ Elk	86
29 Les gènes du groupe PEA3 sont apparus suite à deux duplications d'un "ancêtre"	100

30 Quelques motifs dans la séquence d'Erm	102
31 L'activité d'Erm est modulée par des kinases de la voie Ras et par la PKA	105
32 Plusieurs possibilités pour l'action de BXB sur Erm	106
33 Une hypothèse pour la régulation de l'activité d'Erm par phosphorylation	127
34 Les propriétés des domaines activateurs déterminent en partie les gènes-cibles d'Erm	132
35 Le squelching	166
36 Qui interfère avec qui?	167
37 Deux cas où B squelche A sans avoir les mêmes cibles	168
38 Il existe des différences fonctionnelles entre les transactivateurs acides	169

INTRODUCTION

Présentation du mémoire

La famille ETS est un ensemble de gènes codant des facteurs de transcription dont on trouve des représentants chez tous les Eucaryotes métazoaires. Si sa taille ne suffisait pas à décourager les thésards férus d'exhaustivité (voir fig.1), son perpétuel accroissement rendrait bien vite caduc le fruit de leurs efforts. C'est donc le coeur léger que nous abandonnons toute prétention à un exposé *in extenso* sur les gènes de la famille ETS.

Nous les découvrirons d'abord à travers quelques exemples, qui nous permettront de dégager leur rôle dans la croissance et la différenciation cellulaire. Nous dévoilerons ensuite leur côté sombre, en présentant les cas où les gènes ETS deviennent des oncogènes. Puis, pour mieux comprendre les caractéristiques fonctionnelles des protéines ETS, nous nous livrerons à quelques rappels concernant la modulation transcriptionnelle.

Enfin, nous exposerons nos résultats assortis de quelques commentaires *a posteriori*.

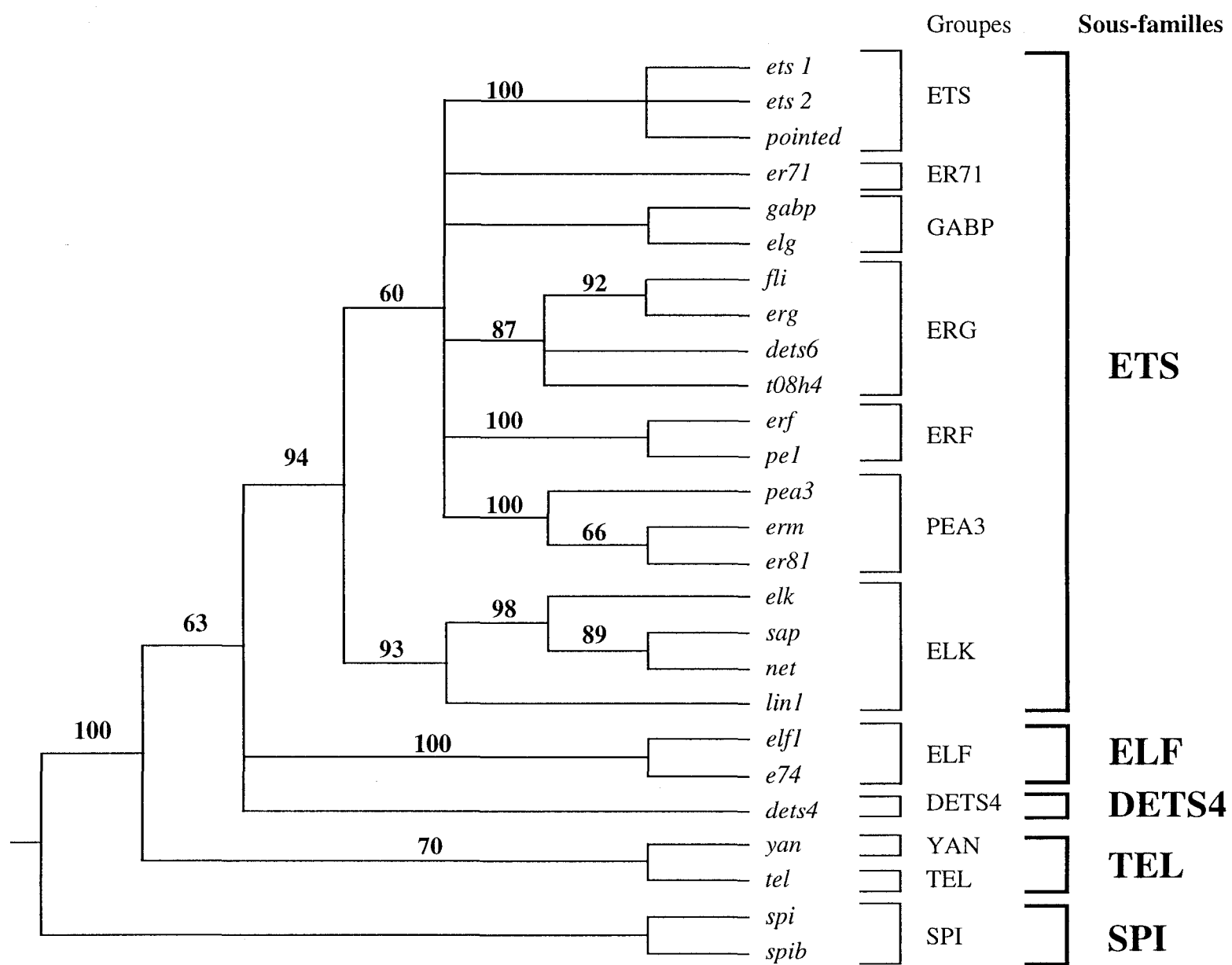


Fig.1 Les gènes de la famille ETS, classés par degré de similitude

Les chiffres indiquent le taux de confiance sur la parenté phylogénétique des gènes d'une branche.

(D'après Laudet et al., soumis)

1. Les gènes ETS sont des régulateurs de la différenciation et la multiplication cellulaire

Pour être brutal, la question est : pourquoi travailler sur les gènes de la famille ETS? Heureusement, les justifications semblent assez faciles à trouver.

L'objectif de cette partie sera de montrer que les gènes de la famille ETS sont des acteurs importants dans de nombreux cas de différenciation cellulaire ou d'interactions entre cellules. On connaît, dans certains cas, les mécanismes par lesquels ils intègrent les signaux extracellulaires, et les gènes-cibles qu'ils régulent.

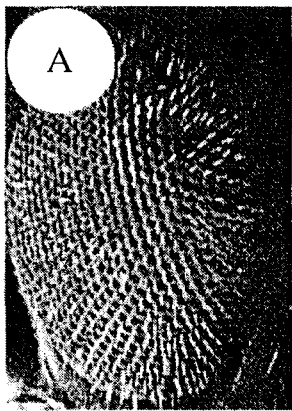
1.1 Les protéines ETS dans le développement des photorécepteurs de drosophile

Laissons-nous aller à un court voyage au pays des diptères. Cette excursion pourrait sembler nous éloigner de notre centre d'intérêt, les gènes ETS de mammifères, mais, on le verra, elle sera riche d'informations utiles.

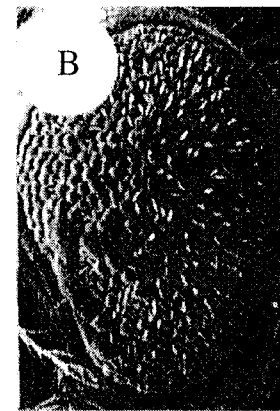
1.1.1 L'oeil de drosophile : un système génétique modèle

La drosophile présente un oeil composé comprenant environ 800 ommatidies arrangées régulièrement. Ces unités contiennent chacune 8 cellules nerveuses photoréceptrices, notées R1 à R8, que l'on peut distinguer par leur morphologie, leur position, ainsi que par différents marqueurs biochimiques (voir fig. 2, partie gauche). Elles contiennent également 4 cellules cônes et 8 cellules accessoires.

L'oeil est donc un organe relativement simple qui est facilement observable. Sa perte de fonction n'étant pas létale, on peut de plus y étudier des mutations incapacitantes. Il représente donc un système intéressant où appliquer la génétique à l'analyse du développement. De fait, son étude a été extrêmement fructueuse.



Sauvage



Irrégulier

yan⁻

yan⁺

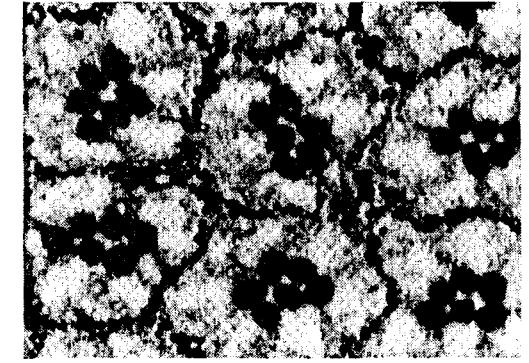
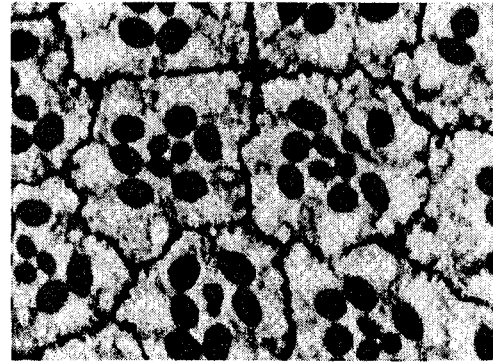
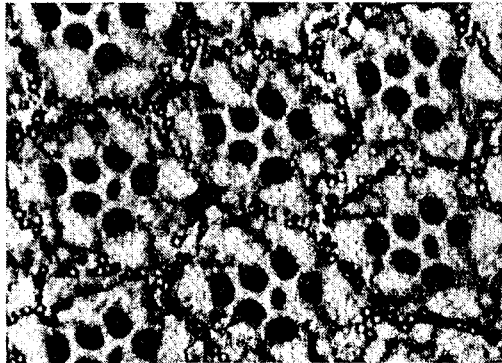


Fig. 2 Des mutations des gènes ETS altèrent le développement de l'oeil de drosophile

A) un oeil normal en microscopie électronique à balayage (en haut), et une coupe tangentielle de cet oeil (en bas). On distingue les masses sombres des rhabdomères de sept photorécepteurs, les R8 étant dans un autre plan de coupe. R7 est le plus petit récepteur, au centre de chaque ommatidie. Les cellules cônes n'apparaissent pas sans coloration.

B) un oeil de phénotype irrégulier, causé par la perte de fonction de *yan*. En bas, à gauche, une coupe sur laquelle on observe des R7 surnuméraires. A droite, une surexpression de *yan* provoque au contraire un manque de R7.

Avant d'évoquer certains de ces résultats, livrons-nous à un rappel succinct sur la différenciation de l'oeil de drosophile (présentée dans Zipursky et Rubin, 1994). La morphogénèse des ommatidies débute dans le disque imaginal oculaire, au passage du sillon morphogénétique. Le disque imaginal contient des cellules totipotentes dont la différenciation ultérieure résulte d'une série d'interactions avec les cellules voisines. Les photorécepteurs se différencient séquentiellement et R7 est le dernier photorécepteur recruté, suite à une interaction avec R8. Les précurseurs des cellules cônes peuvent également, dans certaines conditions, évoluer en R7. Ils forment donc avec le précurseur de R7 un "groupe d'équivalence R7".

La drosophile a été soumise à une mutagenèse intensive qui a conduit à la découverte de nombreux gènes dont l'altération perturbe l'apparition du R7. Ils ont été organisés en une cascade où les gènes aval peuvent compenser une perte de fonction en amont (fig. 3, revue dans Dickson, 1995).

Le premier fut *sevenless* (*sev*), dont l'inactivation entraîne l'absence de R7 dans toutes les ommatidies. Le produit de ce gène est un récepteur transmembranaire à activité tyrosine-kinase. Son ligand, identifié peu après, est le produit du gène *bride of sevenless* (*boss*); il est exprimé spécifiquement dans R8. C'est donc la proximité physique de R8 qui va permettre à une cellule du groupe d'équivalence R7 d'évoluer en photorécepteur. Les quatre autres cellules du groupe, qui ne contactent pas R8, deviendront par défaut des cellules cônes.

De nombreux autres gènes impliqués dans la différenciation du R7 ont été clonés, et on a maintenant une idée assez claire de la cascade biochimique qui résulte de l'activation de Sev par Boss, menant à la différenciation. Elle transite par Ras, Raf et la kinase Erk-A. Parmi les effecteurs nucléaires du signal Boss se trouvent deux membres de la famille ETS, Pointed (Pnt) et Yan, aussi appelé Pok.

Amont

boss

sevenless

drk

sos

E(Sev^{LS})3A (Dras)

polehole (Draf)

Dsor

rolled (Derk-A)

pointed

yan

prospero?

phyllopod?

Aval

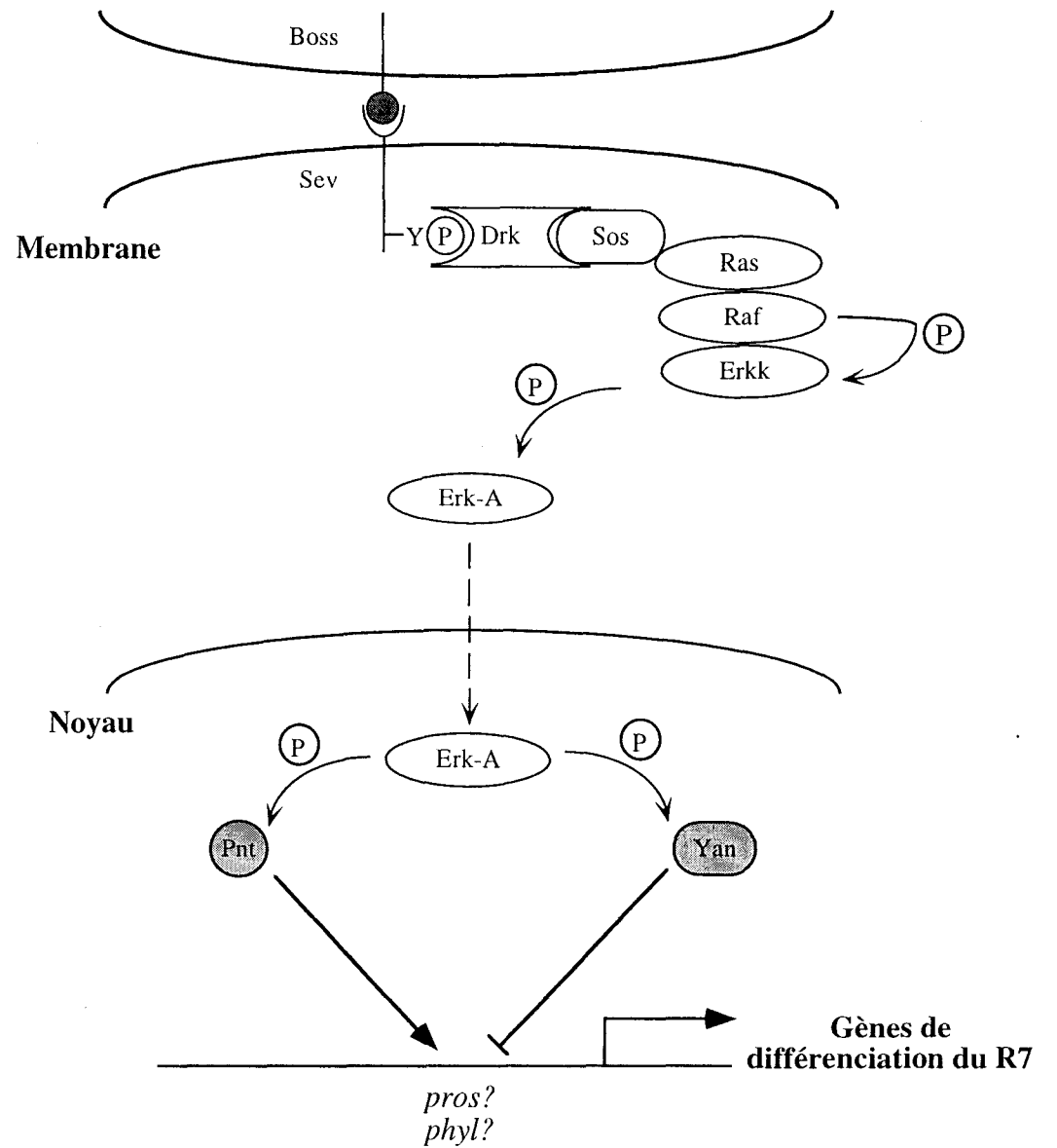


Fig. 3 Les acteurs de la différenciation du photorécepteur R7

A gauche, la cascade génétique, à droite son interprétation biochimique. Les flèches accompagnées d'un P représentent les phosphorylations.

Les "ramifications", impliquant par exemple Gap ou Ksr, ont été omises pour plus de clarté.

1.1.2 *pnt* et *yan*, deux membres antagonistes de la famille ETS

Des mutations dans *pnt* et *yan* affectent, entre autres, la formation de l'oeil (voir fig. 2). Ces gènes s'inscrivent dans la cascade de la figure 3, avec des rôles antagonistes : un allèle hypomorphe (faible) de *pnt* compense la suractivation de *ras* et réduit l'irrégularité de l'oeil. Un allèle faible de *yan* aggrave au contraire ce phénotype, mais peut être compensé par un allèle hypomorphe de *ras*. ***pnt* apparaît donc comme un effecteur positif du signal Boss, tandis que *yan* en est un inhibiteur.**

pnt et *yan* sont exprimés dans les précurseurs de R7 au moment où ceux-ci sont compétents pour le signal Boss, et agissent sur leur différenciation. Un allèle faible de *pnt* conduit à la disparition du R7 dans certaines ommatidies, car son précurseur évolue en cellule cône. Un allèle hypomorphe de *yan* entraîne l'apparition de R7 surnuméraires, qui dérivent de précurseurs de cellules cônes anormalement recrutés (fig. 2).

1.1.3 L'activité de Pnt et Yan est régulée par la kinase Rolled/Erk-A

Le gène immédiatement en amont de *pnt* et *yan* est *rolled* (*rll*), qui code la kinase Erk-A de drosophile. L'interaction génétique entre *rll*, *pnt* et *yan* (*rll* active *pnt* mais réprime *yan*), reflète des événements de phosphorylation des protéines ETS par la kinase. En effet, Yan et Pnt peuvent être phosphorylés *in vitro* par une protéine Rll activée. Dans les deux gènes ETS les sites de phosphorylation *in vitro* ont été identifiés et mutés. Un gène *pnt* dont le site de phosphorylation est éliminé ne répond plus à *rll*, et se comporte comme un inhibiteur du *pnt* endogène (Brunner *et al.*, 1994; O'Neill *et al.*, 1994).

Yan possède de nombreux sites de phosphorylation. En les mutant, on perd là aussi le potentiel de répondre à Rll, et on fait de *yan* un inhibiteur puissant de la différenciation du R7 (Rebay et Rubin, 1995).

En fait, ces variations d'activité induites par Rll correspondent à des variations du **pouvoir transactivateur de Pnt et Yan**. De fait, Pnt peut activer la transcription d'un gène rapporteur dont le promoteur contient des sites ETS multimérisés, et cette activité est accrue par Rll. A

l'inverse, Yan est un inhibiteur de la transcription induite par Pnt. Cet effet est aboli en présence de Rll (Treier *et al.*, 1995).

***pnt* et *yan* influent donc probablement sur le développement de R7 en régulant de façon opposée la transcription d'un ou de gène(s) cible(s) intervenant dans la différenciation.** Un gène est placé en aval de *pnt* et *yan* dans la cascade de la fig. 3, il s'agit de *prospero* (*pros*). En l'absence de *pnt*, ou en présence d'un *yan* hyperactif, *pros* n'est pas induit dans R7. Bien que les preuves directes manquent, ces arguments font de *pros* une cible possible de la régulation par *pnt* et *yan* (Kaufmann *et al.*, 1996). *phyllopod* (*phyl*) est une autre cible possible (Treier *et al.*, 1995).

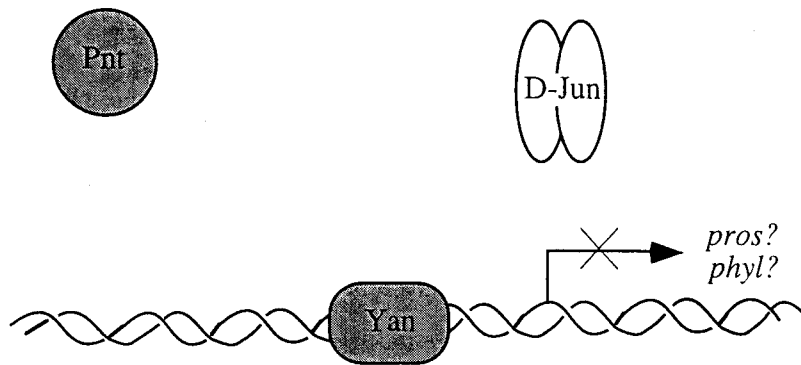
1.1.4 *pnt* coopère avec *D-jun* pour induire la différenciation de R7

La drosophile ne possède qu'un seul homologue de *jun*, *D-jun*, ainsi qu'un seul *fos*. *D-jun* est exprimé de façon transitoire dans les R7 lors de l'induction par Boss. Pour y mettre en évidence sa fonction, Bohmann *et al.* (1994) ont utilisé une version tronquée de D-Jun. Celle-ci contient le leucine-zipper qui permet la dimérisation mais pas le domaine de liaison à l'ADN. Cette protéine forme avec les protéines endogènes des hétérodimères inactifs, et se comporte donc comme un mutant dominant-négatif. Une mutation de *jun* étant létale, cette construction a été exprimée spécifiquement dans l'oeil, grâce au promoteur de *sev*. Elle y **bloque la différenciation de R7 induite par Ras**.

Chez les mammifères, on sait que l'activation de Jun se traduit par la phosphorylation de sites reconnus par les kinases Erk ou Jnk. En remplaçant les sites de phosphorylation par un acide aminé chargé, l'acide aspartique, on peut dans certains cas mimer la phosphorylation. Dans le cas de Jun, la mutation des sites en Asp conduit effectivement à l'activation en absence de phosphorylation (Papavassiliou *et al.*, 1995). Une telle protéine Jun activée a également été placée sous le contrôle du promoteur de *sev* dans l'oeil de drosophile. Elle y **induit la formation de R7 supplémentaires, à la condition que Pnt soit présent** (Treier *et al.*, 1995). *In vitro*, D-Jun coopère fortement avec Pnt pour activer la transcription d'un gène rapporteur

dont le promoteur contient des sites ETS et AP1, et cette activation est contrecarrée par Yan.
Treier *et al.*, (1995) proposent le modèle de la fig. 4.

En l'absence du signal Boss



Lors du signal Boss

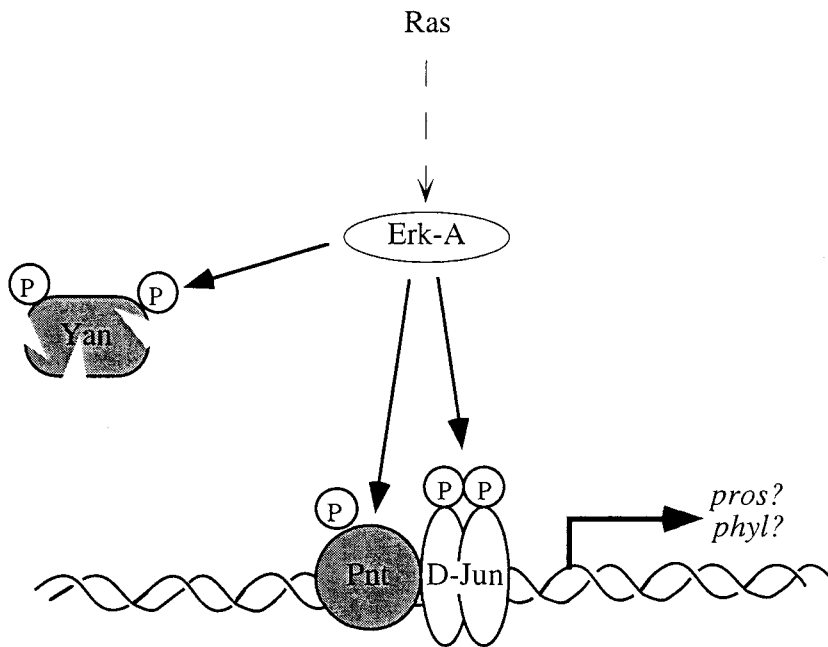


Fig. 4 L'activité de Pnt, Yan et D-Jun est modulée par Ras
En haut, Yan réprime la transcription des gènes spécifiques du R7.
En bas, après phosphorylation par Erk-A, Yan est dégradé, Pnt et D-Jun coopèrent pour induire la transcription de leurs gènes-cibles.

1.1.5 Quelques thèmes ETSiens

Si nous avons choisi de commencer par ces quelques données issues de la génétique de drosophile, c'est qu'elles illustrent plusieurs "leitmotivs fonctionnels" de la famille ETS.

- elles démontrent le **rôle crucial de gènes ETS dans des processus de différenciation cellulaire**.

- elles illustrent bien leur fonction **d'effecteurs de signaux extracellulaires**. La régulation post-traductionnelle de l'activité des protéines ETS par des kinases impliquées dans la transduction de signaux est un thème que nous évoquerons encore.

- enfin, la **capacité des protéines ETS à interagir avec d'autres modulateurs de transcription** (ici, D-Jun), semble être couramment mise en jeu pour accomplir les régulations transcriptionnelles subtiles des gènes impliqués dans la différenciation. Nous rencontrerons d'autres cas d'association entre protéines ETS et facteurs de transcription d'autres familles.

1.2 Protéines ETS et hématopoïèse

Armés de ces quelques préceptes, nous voilà prêts à aborder plus facilement un aspect du développement des vertébrés où le rôle des gènes ETS est particulièrement clair. Les données étant fort nombreuses sur ce sujet, nous nous limiterons à en exposer certaines qui ont valeur d'exemple. Pour plus d'informations, se rapporter à Wasylyk *et al.*, 1993; Romano-Spica *et al.*, 1994.

1.2.1 *spi-1/pu.1* est un gène essentiel au développement de différentes lignées hématopoïétiques

spi-1 a été cloné en qualité d'oncogène potentiel surexprimé dans les cellules leucémiques de souris infectées par le virus SFFV (Moreau *et al.*, 1988; voir paragraphe 2.1). *pu.1* était quant à lui identifié comme un facteur de transcription exprimé dans les cellules B et les macrophages,

et régulant la transcription des gènes du virus SV40 (Klemsz *et al.*, 1990). En fait, *spi-1* et *pu.1* ne font qu'un...

spi-1/pu.1 (que nous noterons plus brièvement *pu.1* dans cette partie) est un des membres de la famille ETS les mieux connus. Il est exprimé dans toutes les cellules sanguines, à l'exception des cellules T, mais ne semble pas l'être dans les autres tissus. De nombreux résultats soulignent son rôle crucial dans l'expression de gènes propres aux cellules hématopoïétiques (Pour une revue sur *pu.1*, voir Moreau-Gachelin, 1994).

a) Pu.1 régule l'expression de gènes spécifiques des lymphocytes B

L'expression tissu-spécifique des chaînes constituant les immunoglobulines est soumise à un strict contrôle (revue dans l'introduction de Eisenbeis *et al.*, 1995).

"L'enhancer" (en français, amplificateur) en 3' du gène de la chaîne légère κ des Immunoglobulines (noté E κ 3') lui confère son expression spécifique des cellules B. Il contient une séquence ETS qui lie Pu.1 (Fig. 5), adjacente à un site liant un facteur spécifique des cellules B, NF-EM5 (Pongubala *et al.*, 1992). Si l'on mute le site Pu.1, NF-EM5 ne peut se fixer à E κ 3'. Réciproquement, si l'on mute le site NF-EM5, Pu.1 se fixe à l'ADN mais n'active pas l'enhancer. C'est donc Pu.1 qui recrute NF-EM5, mais l'interaction des deux protéines est nécessaire à l'activation (Fig. 5).

L'interaction NF-EM5-Pu.1 implique chez ce dernier une région de 43 Acides Aminés contenant une séquence PEST et des sites de phosphorylation par la Caséine Kinase II (CKII). Pongubala *et al.* (1993) ont montré que la phosphorylation de ces sites, par une kinase qui est vraisemblablement la CKII, est essentielle pour l'interaction entre les deux protéines. Il se pourrait que l'interaction, en masquant la séquence PEST, protège Pu.1 de la protéolyse (Pongubala *et al.*, 1992).

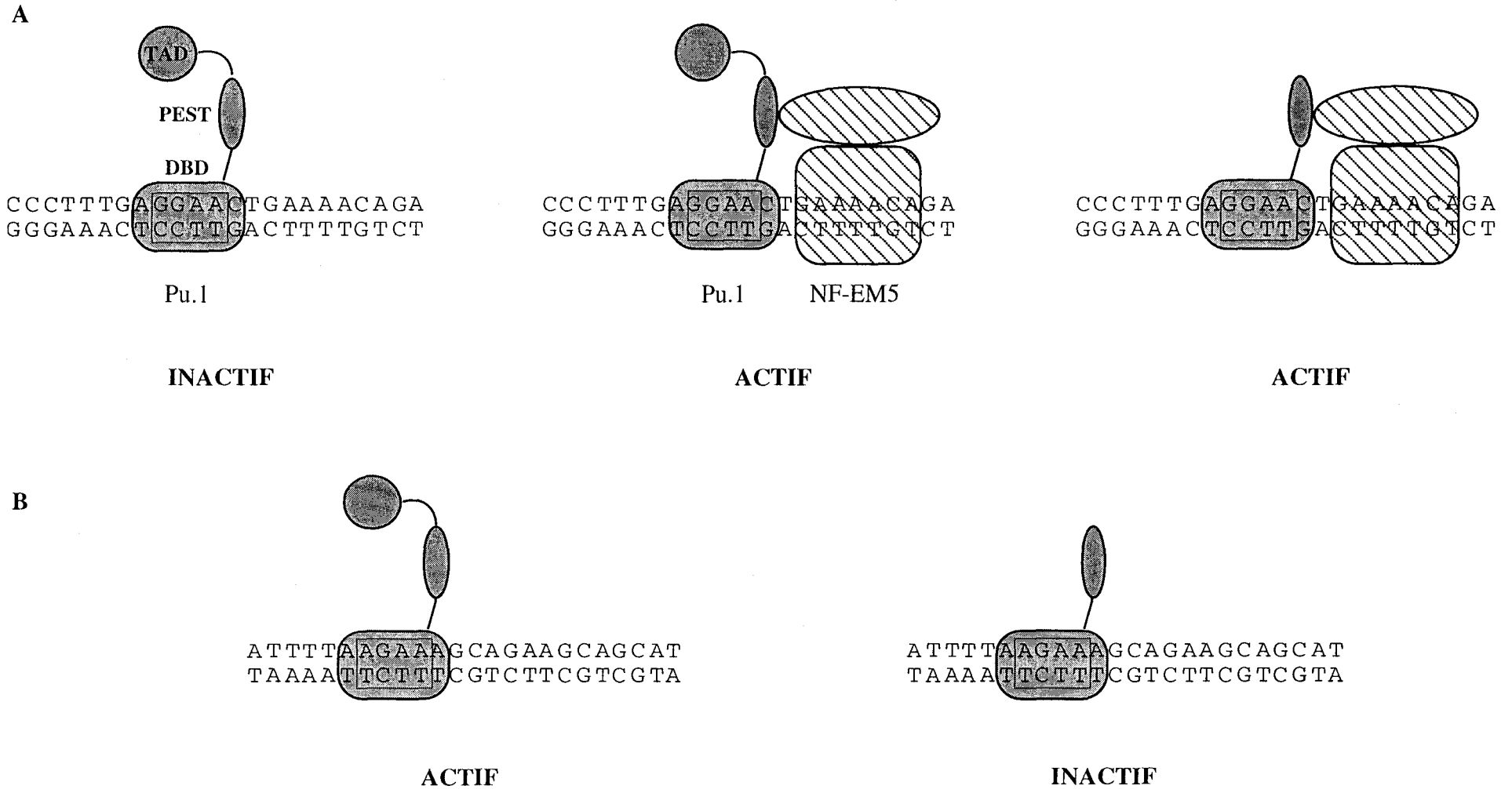


Fig. 5 Sur des promoteurs différents, l'activité transcriptionnelle de Pu.1 est différente

A) Sur "l'enhancer" 3' de la chaîne κ des Immunoglobulines, Pu.1 recrute NF-EM5 mais ne transactive pas de façon autonome
B) Sur le promoteur de la chaîne J des Immunoglobulines, la transactivation par Pu.1 est autonome et requiert le domaine de transactivation N-terminal

Une protéine ayant une activité NF-EM5 a été clonée par la suite et dénommée Pip1, pour Pu.1-interacting-protein. Elle est exprimée surtout dans les cellules B, mais aussi un peu dans les cellules T (Eisenbeis *et al.*, 1995).

La régulation de l'activité de E κ 3' est donc particulièrement représentative de la combinatoire transcriptionnelle impliquant les protéines ETS :

- la combinaison de deux facteurs dont les patrons d'expression ne se recoupent que dans les cellules B permet d'obtenir une expression tissulaire limitée à ces cellules.
- un niveau supplémentaire de régulation est apporté par la nécessité de la phosphorylation de Pu.1.

Cet exemple, adjoint à celui de l'interaction Pnt-D-Jun dans les photorécepteurs de drosophile, pourrait suggérer que les protéines ETS sont des facteurs "grégaire", qui sont peu efficaces pour moduler la transcription en l'absence d'un partenaire. Il est bon de discuter quelque peu cette idée. Dans de nombreux cas de régulation transcriptionnelle mettant en jeu les protéines ETS, on peut effectivement observer ou suspecter leur coopération avec d'autres régulateurs. Cependant, il est probable que cette situation est **favorisée** car on atteint avec des complexes protéiques des régulations plus subtiles qu'avec des protéines isolées, mais qu'elle **n'est pas une nécessité** fonctionnelle absolue.

De fait l'interaction de Pu.1 avec un partenaire n'est pas toujours requise pour sa fonction. Ainsi, toujours dans les cellules B, Pu.1 active la transcription du gène de la chaîne J des immunoglobulines (IgJ) en réponse à une stimulation par l'Interleukine 2 (Shin et Koshland, 1993). Cet effet dépend non d'un enhancer, mais de la séquence JB présente dans le promoteur de IgJ. Pu.1 s'y fixe, mais n'y recrute pas NF-EM5, et peut activer, seul, la transcription (Fig. 5). L'activation de transcription à partir de E κ 3' ne requiert pas les 100 premiers acides aminés de Pu.1. Par contre, cette région est nécessaire à l'activité sur JB. Elle contient un domaine

transactivateur qui contient des modules riches en glutamine et en résidus acides (Shin et Koshland, 1993; Klemsz et Maki, 1996).

Nous souhaitons dégager deux idées clés :

- **les protéines ETS sont des facteurs de transcription "polyvalents"**. Sur certains sites ils agissent seuls, sur d'autres ils interagissent avec d'autres régulateurs.

- **l'effet des domaines transactivateurs dépend du promoteur régulé**. Ainsi, sur E κ 3', les domaines activateurs de Pu.1 ne sont pas nécessaires. On ne sait pas encore s'ils sont inactifs dans ce contexte, ou s'ils sont redondants avec ceux de NF-EM5.

b) Pu.1 contrôle l'expression de nombreux gènes exprimés dans les lignées myéloïdes

Le rôle de Pu.1 dans les cellules hématopoïétiques ne se limite pas aux lymphocytes B. Pu.1 régule en effet un grand nombre de gènes spécifiques de lignées myéloïdes (revue dans Moreau-Gachelin, 1994). Citons parmi ceux-là deux gènes impliqués dans la maturation des monocytes et des macrophages : le Scavenger Receptor (SR), et le Granulocyte/Macrophage Colony Stimulating Factor (GM-CSF).

c) Un KO qui fait du bien

pu.1 est vraiment un cas d'école. Nous avons mentionné son rôle dans l'expression spécifique de nombreux gènes spécifiques des cellules hématopoïétiques. Nous avons vu qu'il pouvait fonctionner seul ou avec un partenaire, et que son activité était dans ce dernier cas régulée par phosphorylation. Et bien ce n'est pas tout : l'inactivation de *pu.1* par recombinaison homologe a livré des résultats gratifiants. Les souris homozygotes pour la mutation meurent peu avant la naissance. Elles sont dépourvues de lymphocytes, de macrophages, et de granulocytes. On l'a dit, *pu.1* n'est pas exprimé dans les cellules T, et pourtant elles sont absentes des souris

mutantes. Il est donc probable qu'un précurseur de ces cellules soit dépendant de Pu.1 (Scott *et al.*, 1994; McKercher *et al.*, 1996).

1.2.2 Ets-1, Ets-2 et l'activation des lymphocytes

ets-1 et *ets-2* présentent un patron d'expression dans les cellules sanguines qui est clairement distinct de celui de *pu.1*. *ets-1* est exprimé principalement dans les lymphocytes B et T, tandis que le champ d'expression de *ets-2* est nettement plus vaste, et inclut également des cellules érythroïdes et myéloïdes (Romano-Spica *et al.*, 1994).

Il semble y avoir une intéressante relation Ets-1/Ets-2 dans les lymphocytes B et T. Dans les cellules quiescentes, Ets-1 est détectable, mais Ets-2 ne l'est quasiment pas (Bhat *et al.*, 1990). L'activation mitogénique des lymphocytes inverse ce rapport, en diminuant la transcription de *ets-1* et en augmentant celle de *ets-2*. Elle a également des effets post-traductionnels, car elle provoque la phosphorylation de Ets-1 par une kinase dépendante du Calcium. Cette modification réduit l'affinité de la protéine pour l'ADN (Rabault et Ghysdael, 1994).

L'activation se traduit donc par une induction de *ets-2* et une répression, tant transcriptionnelle que post-traductionnelle, de *ets-1*. Le retour à la quiescence inverse ce processus.

Cet apparent antagonisme entre deux gènes ETS pour réguler la différenciation n'est pas sans rappeler les frères ennemis *yan* et *pnt*. On peut spéculer sur un rôle de Ets-1 dans le maintien de l'état quiescent, opposé au rôle de Ets-2 dans l'activation.

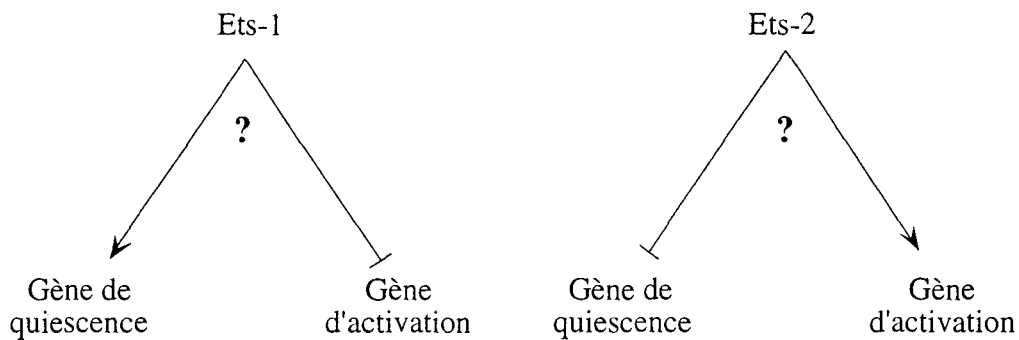


Fig.6 Ets-1 et Ets-2 pourraient jouer des rôles opposés dans l'activation des lymphocytes

Plusieurs résultats supportent ce modèle :

- d'abord, dans les cellules T, Ets-1 active la transcription de certains gènes, mais en réprime d'autres. Dans la première catégorie, on trouve le gène de la chaîne α du TCR (Prosser *et al.*, 1992); dans la deuxième "l'enhancer" de la chaîne β du même récepteur (Prosser *et al.*, 1992). Dans les deux cas une fixation coopérative avec CBF (alias PEBP2 α , voir fig. 22) est impliquée (Wotton *et al.*, 1994; Giese *et al.*, 1995). Un effet antagoniste de Ets-1 et Ets-2 sur certaines cibles est donc envisageable. On ne sait pas encore, cependant, si Ets-2 à l'inverse de Ets-1 active "l'enhancer" de la chaîne β .

Notons qu'il n'est pas nécessaire que Ets-1 soit un répresseur actif : il pourrait simplement rentrer en compétition avec des activateurs pour des sites tels que "l'enhancer" de la chaîne β .

- le gène *ets-1* a été inactivé par recombinaison homologe dans des cellules souches embryonnaires ES. Ces cellules recombinées ont été introduites dans des blastocystes de souris mutées dans le gène *rag2*, qui est indispensable au développement des lymphocytes. Les cellules B et T trouvées dans les souris chimères dérivent donc nécessairement des cellules ES injectées et sont dépourvues de Ets-1 sauvage. On observe dans les animaux chimères des défauts majeurs, dont un déficit en lymphocytes T (Bories *et al.*, 1995; Muthusamy *et al.*, 1995). Ces lymphocytes T mis en culture sont particulièrement enclins à l'apoptose. La situation des lymphocytes B dépourvus de Ets-1 est un peu différente : les cellules sont présentes en quantité normale mais se différencient de façon aberrante en cellules sécrétrices

d'IgM. Ces deux observations sont compatibles avec l'hypothèse que Ets-1 contribue à maintenir ces cellules en phase G0, et que la perte de ce contrôle les conduise par défaut vers l'apoptose ou une différenciation anormale.

Le rôle de Ets-2 dans le processus d'activation est moins clair. Des souris chez lesquelles l'expression relativement ubiquiste de Ets-2 est induite par transgénèse présentent des anomalies des os et du cartilage, mais les auteurs ne citent pas de défaut des cellules sanguines, sans préciser toutefois s'ils ont examiné ce caractère (Sumarsono *et al.*, 1996).

1.3 Protéines ETS et remodelage tissulaire

Passons du liquide au solide. La mise en place de tissus organisés et leur évolution supposent des interactions bien définies entre cellules du tissu, pour assurer leur cohésion, et des interactions entre les cellules et leur environnement, qui permettront les migrations. Les gènes ETS interviennent dans ces processus complexes.

1.3.1 Gènes ETS et migration cellulaire

Le mouvement c'est la vie. Ou, plus académiquement, la migration cellulaire est un événement nécessaire à la morphogenèse. Elle intervient par exemple au cours de la gastrulation ou, plus tard, dans la formation du sclérotome par des cellules essaimant depuis les somites.

Plusieurs gènes de la famille ETS semblent impliqués dans les migrations cellulaires :

- chez l'oiseau, la formation des plumes a lieu dans les dômes plumaires. Ces foyers se forment par aggrégation de cellules dermiques, qui migrent et se regroupent. *ets-1* est fortement exprimé

dans le derme d'oiseau quand ces mouvements ont lieu, et son expression cesse une fois les dômes plumaires constitués (Desbiens *et al.*, 1991; Quéva *et al.*, 1993).

La corrélation entre l'expression de *ets-1* et la migration cellulaire n'est pas limitée au derme : ce gène est également transcrit lorsque les cellules ventrales des somites se dissocient et essaient pour former le sclérotome. Une fois encore, l'expression du gène cesse avec la fin de la migration (Fafeur *et al.*, soumis). L'expression de *ets-1* est également associée à la migration des cellules de crête neurale (Fafeur *et al.*, soumis).

- la migration requiert bien sûr la motilité des cellules, mais aussi la présence d'un "chemin de migration" aux caractéristiques définies. *fli-1* pourrait intervenir dans la genèse de ces "pistes cellulaires". Rappelons que la formation des reins commence avec l'établissement d'un canal pronéphrétique. Celui-ci dérive en partie de cellules précurseurs migrant le long de l'axe antéropostérieur. Chez le Xénope, Meyer *et al.* (1995) ont décrit l'expression de *fli-1* non dans les cellules migrant, mais le long de leur voie de migration. Des embryons de xénope ayant subi l'injection d'ARNm codant Fli présentent de multiples défauts cohérents avec l'hypothèse de l'intervention de Fli dans certains événements migratoires (Rémy *et al.*, 1996).

1.3.2 Angiogenèse et vasculogenèse

Il existe une corrélation très claire entre l'expression de *ets-1* et la formation des vaisseaux sanguins, par angiogenèse (ramification de vaisseaux pré-existants) ou vasculogenèse (néo-formation).

Au cours du développement de l'embryon de poulet, *c-ets-1* est exprimé dans les cellules endothéliales lors de la formation des capillaires mais aussi des vaisseaux de plus fort diamètre, comme l'aorte; l'expression cesse une fois les vaisseaux formés (Vandenbunder *et al.*, 1989; Quéva *et al.*, 1993). Un patron similaire est observé lors de l'embryogenèse murine (Grévin *et al.*, 1993).

De nouveaux vaisseaux peuvent se former chez l'homme adulte, après une blessure ou lors de la vascularisation de tumeurs. Dans ces deux cas, *ets-1* est ré-exprimé dans les cellules endothéliales impliquées (Wernert *et al.*, 1992).

1.3.3 Morphogenèse épithéliale

Cachez ce sein, que je ne saurais voir... Il est intéressant, pourtant, de relever l'expression de *ets-1* dans le développement de la glande mammaire. La croissance des canalicules mammaires et leur ramification résulte d'interactions entre l'épithélium des tubules et le mésenchyme environnant. L'expression de *ets-1* est détectable à l'extrémité des canalicules, mais aussi dans le mésenchyme (Delannoy-Courdent *et al.*, soumis).

On sait que ce dernier délivre aux cellules épithéliales un signal de croissance et de ramification, véhiculé entre autres par le Scatter Factor (SF, aussi connu comme Hepatocyte Growth Factor ou HGF). Des cellules épithéliales mammaires en culture n'expriment pas *ets-1*. Si on les traite avec un milieu conditionné contenant du SF, *ets-1* est induit (Delannoy-Courdent *et al.*, soumis).

L'inductibilité de *ets-1* par des molécules impliquées dans la morphogenèse épithéliale est également retrouvée dans un autre système modèle. Les cellules canines MDCK, qui dérivent d'épithélium rénal, forment des agrégats lors de la culture *in vitro*. Ceux-ci se dissocient rapidement sous l'effet du SF. Fafeur *et al.* (soumis) ont montré que le SF induisait la transcription de *ets-1* et l'accumulation de la protéine correspondante.

1.3.4 Les protéines ETS régulent la transcription de gènes impliqués dans la dégradation de la matrice extracellulaire et l'adhérence

Les événements morphogénétiques que nous avons cités : migration, angiogenèse, et morphogenèse des épithéliums, semblent bien disparates. Ils présentent néanmoins un

"dénominateur commun" : **tous requièrent la réorganisation de la matrice extracellulaire (MEC) et/ou la modification des caractéristiques d'adhérence des cellules.**

Plusieurs gènes impliqués dans ces deux processus sont régulables par les protéines ETS.

a) Gènes impliqués dans le remodelage de la MEC

Les gènes de l'activateur de l'*urokinase-plasminogen activator* (u-PA), de la *stromélysine*, de la *collagénase* de type I et de type IV répondent à des protéines ETS (Nerlov *et al.*, 1992; Wasyluk *et al.*, 1991; Gutman et Wasyluk, 1990; Higashino *et al.*, 1995). L'inhibition de certaines protéases est un autre moyen de contrôler la composition de la MEC. Un inhibiteur de protéases, *timp-1* (Tissue Inhibitor of MetalloProteinases-1), est également régulé par des protéines ETS (Logan *et al.*, 1996).

Il existe un point commun remarquable à tous ces promoteurs : ils sont régulables par les protéines de la famille AP1, et les sites de fixation de ces facteurs de transcription sont situés à proximité des sites ETS (certains sont représentés en fig. 7). Il est difficile de croire que cette récurrence soit fortuite et, de fait, **il existe une coopération entre protéines ETS et AP1 pour la transactivation.**

Cette collaboration dépend de l'espacement sur le promoteur des sites fixant les partenaires. Sur le promoteur du gène de la *collagénase I*, ils sont distants d'une dizaine de paires de bases et l'effet de Ets-1 et AP1 pour la stimulation transcriptionnelle est additif. Si par mutagenèse on rend les sites adjacents, on augmente la coopérativité Ets-1-AP1 (Gutman et Wasyluk, 1990). Ces observations suggèrent des interactions entre les protéines, qui ont effectivement été mises en évidence (Buttice *et al.*, 1996). En fait, l'affinité ETS-AP1 qui, rappelons-le, existe déjà chez la drosophile, peut être mise en évidence sur différents promoteurs, et entre différentes protéines de ces deux familles (Bassuk et Leiden, 1995; Nakae *et al.*, 1995; Sieweke *et al.*, 1996).

b) Gènes codant des molécules d'adhérence

Ces molécules permettent aux cellules d'interagir entre elles, ainsi qu'avec la MEC. Elles agissent donc sur la migration cellulaire mais aussi sur la cohésion des tissus. Plusieurs gènes codant des intégrines sont régulables via des sites ETS dans leur promoteur (Bottinger *et al.*, 1994; Rosen *et al.*, 1994). Notre groupe a montré qu'Erm active la transcription du gène *icam-1* (Intercellular Adhesion Molecule-1, de Launoit *et al.*, Soumis).

1.3.5 L'induction de l'activité ETS peut être transcriptionnelle ou post-traductionnelle

Il existe donc des cibles potentielles des gènes ETS qui pourraient rendre compte de leur rôle dans les événements de morphogénèse. Il nous manque toutefois un lien entre le "signal" morphogénétique, par exemple le SF, et l'expression des protéines qui affecteront les caractéristiques de la cellule ou de la MEC. Ce lien, c'est l'activation des protéines ETS par les signaux extracellulaires morphogénétiques.

Montrons que cette modulation se porte sans doute à deux niveaux :

- de façon rapide, par phosphorylation de la protéine ETS,
- avec une latence plus longue, par une régulation de sa transcription.

a) Phosphorylation de Ets-1 et Ets-2 en réponse aux stimuli extracellulaires

Le signal véhiculé par le SF, mais aussi de nombreux autres facteurs de croissance ou de morphogénèse, est transduit jusqu'au noyau par une cascade de kinases impliquant Ras et Raf, similaire à celle présentée dans la fig. 3.

D'autre part, *ets-1* et *ets-2* sont les gènes ETS les plus proches de *pnt* de drosophile. Comme Pnt est la cible de Erk-A en réponse à la stimulation de Ras, il était légitime de rechercher un événement "homologue" de phosphorylation de Ets-1 et Ets-2 suite à l'activation de Ras.

On peut réaliser cette activation par différents moyens : en stimulant les cellules par du Sérum ou des facteurs de croissance comme le SF, ou en transfectant des mutants constitutivement actifs de Ras. Dans tous ces cas, on induit la phosphorylation de Ets-1 et Ets-2, et cet

événement augmente leur pouvoir transactivateur (Rabault *et al.*, 1996; Yang *et al.*, 1996; V. Fafeur, communication personnelle). Cependant, la kinase phosphorylant les protéines ETS ne semble pas être Erk-1 ou Erk-2. Il est clair que des pans de la cascade Ras/Raf.../Erk restent à découvrir (voir par exemple Zhou *et al.*, 1995)

L'idée importante, en tout cas, est que, comme Pnt chez la drosophile, **Ets-1 et Ets-2 sont des effecteurs du signal Ras**. On peut maintenant l'avouer : tous les sites présentés dans la figure 7 sont connus depuis longtemps comme des éléments de réponse à Ras (Ras Response Elements, RRE). Toutefois, ce n'est que récemment qu'on a montré que l'induction était une conséquence de la phosphorylation des protéines ETS. Bien sûr, les protéines AP1 ne sont pas des spectateurs passifs de l'induction : leur activité est également régulée par phosphorylation.

b) Régulation de la transcription de *ets-1*

On l'a dit, l'induction de l'activité de Ets-1 par le SF est également transcriptionnelle. Il est possible que cette induction passe par les séquences du promoteur de *ets-1*, qui répondent au sérum (Majérus *et al.*, 1992; Seth et Papas, 1990). D'autre part, comme Ets-1 régule positivement la transcription de son propre gène (Majérus *et al.*, 1992), il est possible que le SF, en induisant post-traductionnellement l'activité de Ets-1, établisse une boucle d'autorégulation positive.

2. Certains gènes ETS sont des proto-oncogènes (ou : quand la machine se détraque)

Certains gènes assurant des fonctions importantes dans les processus de différenciation ou de multiplication cellulaire sont de paisibles Dr Jekyll qui peuvent cependant se transformer en de dangereux Mr Hyde. Pour citer un cas célèbre, la protéine p53 est un "gardien du génome" qui veille à l'intégrité de l'ADN mais la même protéine, une fois mutée, suffit à transformer des cellules. Les gènes ETS n'échappent pas à cette dualité, et une altération de leur fonction peut les rendre oncogéniques.

Dans le premier cas que nous évoquerons, la leucémie de Friend, la transformation dérive de la surexpression d'une protéine ETS, qui toutefois n'est pas mutée. Les deux suivants, leucémie induite par le virus aviaire E26 et sarcomes d'Ewing, impliquent des membres de la famille ETS dont les propriétés sont modifiées par la fusion à d'autres protéines ou par des mutations.

2.1 Spi-1, Fli-1, et les leucémies de Friend

Les leucémies de Friend sont des pathologies causées par des rétrovirus chez la souris. Leur premier symptôme est la prolifération de proérythroblastes non tumorigènes. Il est suivi de l'apparition de clones proérythroblastiques immortels et tumorigènes, causant une leucémie fatale (revue dans Moreau-Gachelin, 1994). Deux protéines ETS sont intimement liées à ces leucémies érythroïdes.

2.1.1 *spi-1* est surexprimé dans les cellules infectées par les virus de Friend

La leucémie de Friend intervient après l'infection de souriceaux ou de souris adultes par un complexe viral composé de FMuLV (Friend Murine Leukemia Virus, un virus compétent pour la réplication), et de SFFV (Spleen Focus Forming Virus, un virus déficient). On connaît assez

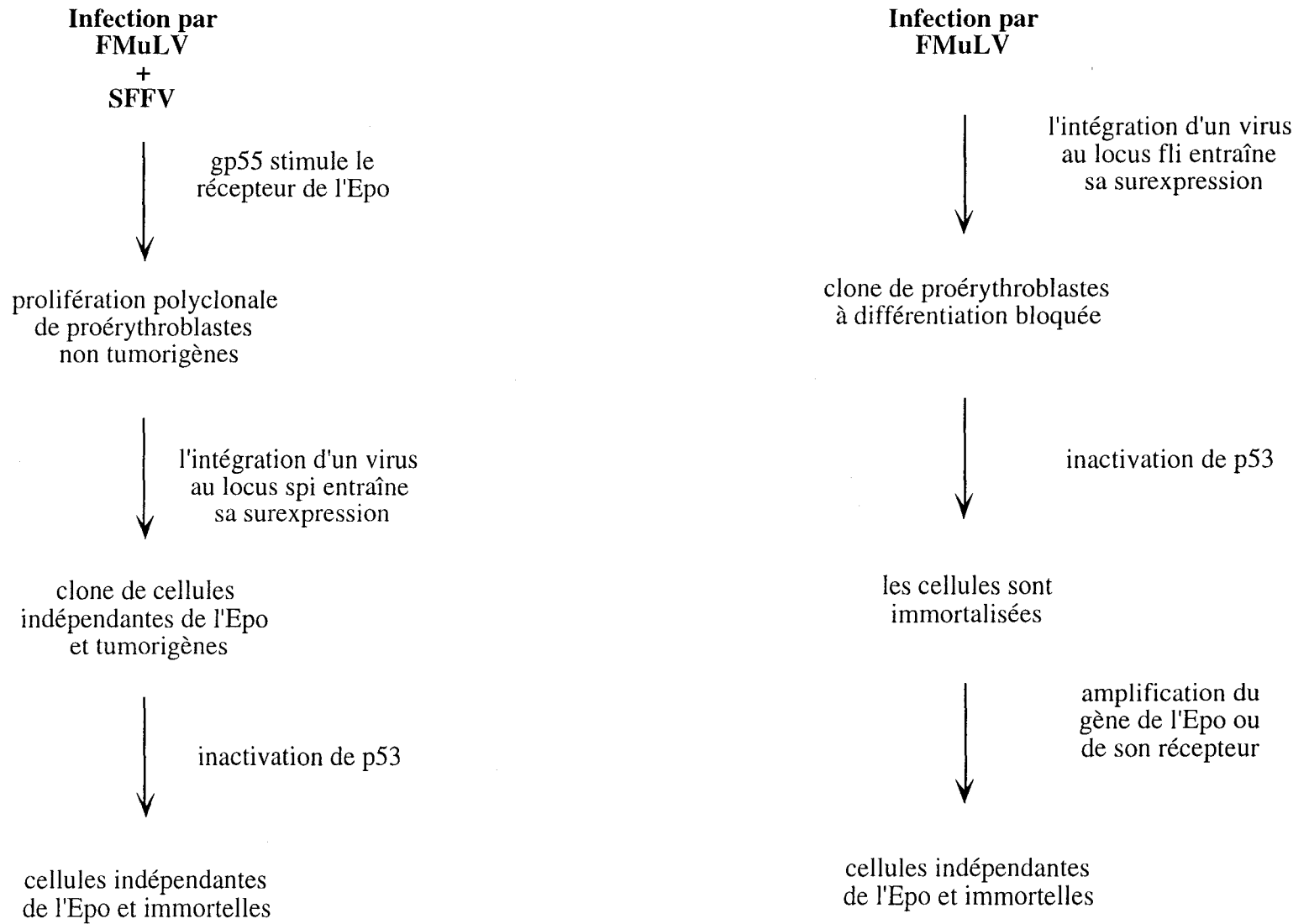


Fig. 8 La séquence des événements conduisant à la leucémie de Friend après infection par SFFV ou FMuLV
 Epo : Erythropoïétine

bien les mécanismes par lesquels les virus induisent ces leucémies (voir fig. 8, et Ben-David et Bernstein, 1991, pour revue).

L'étape proliférative résulte de la stimulation du récepteur de l'érythropoïétine (Epo-R), par la protéine d'enveloppe virale gp55. Les cellules malignes apparaissant ensuite sont des clones, ce qui suggère qu'un événement secondaire a conféré un avantage de croissance aux cellules et a été sélectionné. L'analyse de ces clones a montré que, dans la quasi-totalité des cas, la transformation est corrélée à l'intégration du provirus SFFV à un site spécifique du génome. Ce locus a été nommé *spi-1* (pour SFFV Provirus Integration-1; Moreau *et al.*, 1988). **L'insertion s'accompagne de la surexpression d'un gène situé à proximité du site d'intégration, appelé par extension *spi-1*. Il s'agit d'un membre de la famille ETS (Moreau *et al.*, 1988). *spi-1* est identique à *pu.1* découvert indépendamment par (Klemsz *et al.*, 1990), voir paragraphe 1.2.1.**

2.1.2 L'expression de *spi-1* transforme les proérythroblastes

Pour avoir une preuve du rôle causal de *spi-1* dans les leucémies de Friend, il a fallu le surexprimer dans des cellules. Et, de fait, la transformation de cellules de moelle de souris par un rétrovirus portant *spi-1* conduit à une prolifération de précurseurs érythroblastiques (Schuetze *et al.*, 1993). Des souris transgéniques exprimant *spi-1* sous le contrôle du LTR de SFFV développent également une leucémie à érythroblastes rappelant tout à fait la leucémie de Friend (Moreau-Gachelin *et al.*, 1996).

On peut avancer une hypothèse quant à l'origine du pouvoir oncogénique de *spi-1* dans les proérythroblastes. Les lignées MEL issues de leucémie de Friend expriment fortement *spi-1*. On peut pousser ces lignées à la différenciation érythroblastique et à la perte de l'immortalité en les traitant avec du DMSO, ou du HMBA. La différenciation s'accompagne d'une forte réduction de la quantité de Spi-1 (Schuetze *et al.*, 1992; Galson *et al.*, 1993). Si, réciproquement, on diminue la quantité de Spi-1 dans les cellules en utilisant des oligonucléotides antisens, on diminue leur potentiel réplicatif, sans induire leur différenciation

(Delgado *et al.*, 1994). **On pense donc que Spi-1 maintient les précurseurs érythroblastiques dans un état de prolifération qui leur interdit la différenciation.**

Il est vraisemblable qu'au cours de l'hématopoïèse cette activité mitogène de Spi-1 est fermement contrôlée, soit par modification de la protéine, soit par extinction de son expression. Lors de l'intégration par SFFV, *spi-1* est largement surexprimé et doit échapper à ces mécanismes de régulation.

2.1.3 *fli-1* est surexprimé dans les leucémies induites par FMuLV

Le virus FMuLV peut provoquer des leucémies de façon indépendante de SFFV. Il cause différents néoplasmes hématopoïétiques, dont des érythroleucémies aux caractéristiques très proches de celles causées par les complexes FLV.

Une fois encore, la transformation des cellules est associée à l'intégration du virus à un site spécifique du génome : dans 75% des clones leucémiques étudiés, FMuLV est inséré au même site, appelé *fli-1* (Friend Leukemia Integration-1), qui est distinct de *spi-1* (Ben-David *et al.*, 1990). Les LTR viraux induisent la **surexpression d'un gène proche du site d'insertion, appelé par extension *fli-1*. Il s'agit également d'un membre de la famille ETS** (Ben-David *et al.*, 1991).

spi-1 et *fli-1* sont très dissemblables (voir fig.1), et il est peu vraisemblable qu'ils transforment les cellules érythroïdes par le même mécanisme. Le rôle potentiel de *fli-1* dans la différenciation érythrocytaire est mal connu et, à l'heure actuelle, il semble difficile d'expliquer son pouvoir leucémogène. Les phénotypes des animaux mutants sont peu informatifs sur ce sujet : des souris transgéniques surexprimant *fli-1* présentent des anomalies lymphocytaires conduisant à une maladie auto-immune, mais pas de leucémies (Zhang *et al.*, 1995; on ne sait pas, cependant, si le transgène est exprimé dans les précurseurs érythroïdes). Le gène *fli-1* a été muté par recombinaison homologue chez la souris (Melet *et al.*, 1996). FMuLV s'intègre encore

au locus *fli-1* chez les animaux mutants, et induit des leucémies un peu plus tardives. Il semblerait donc que l'expression de *fli-1* ne soit pas indispensable à la transformation par FMuIV.

Ces interprétations sont toutefois sujettes à caution car les souris recombinaées produisent, en faible quantité, une version tronquée de *Fli-1* qui contient encore le domaine ETS (Melet *et al.*, 1996). Difficile de dire, dans ces conditions, si le gène est effectivement inactivé.

2.2 *v-ets* : un oncogène viral dérivé de *ets-1*

C'est l'étude des virus oncogènes qui a mené à la découverte du premier membre de la famille ETS en 1983. C'est à cette date qu'a été obtenue la séquence du rétrovirus aviaire E26, isolé pour sa capacité à provoquer des leucémies chez le poulet. Ce virus contient, fusionnés au gène viral *gag*, deux gènes exogènes : *v-myb* et *v-ets* (Leprince *et al.*, 1983; Nunn *et al.*, 1983). Ces gènes sont homologues à *c-myb* et *ets-1*.

2.2.1 *v-ets* coopère avec *v-myb* pour induire des leucémies

L'infection de cellules de moëlle par E26 *in vitro* génère des cellules transformées. Celles-ci peuvent, selon les conditions, se différencier en membres de trois lignées sanguines différentes : érythrocytes, myéloblastes et éosinophiles (fig. 9). **Elles sont donc bloquées par E26 à un stade extrêmement précoce de la différenciation.** On appelle MEP (Myb-Ets Progenitor) ces cellules multipotentes.

Golay *et al.* (1988) ont montré que la fonction du domaine ETS est nécessaire au blocage de la différenciation des MEPs. Ces auteurs ont obtenu un mutant thermosensible de E26, l'allèle *ts1.1*, qui porte une mutation dans le domaine ETS. Des MEPs infectées par E26 *ts1.1* peuvent se différencier, dès lors qu'on les cultive à température restrictive pour *v-ets* (fig. 9)

***v-ets* est donc, de façon un peu similaire à *yan*, un inhibiteur de différenciation.** Notons cependant que sa fonction dépend fortement de celle de *v-myb*. En effet, un virus portant *v-ets*

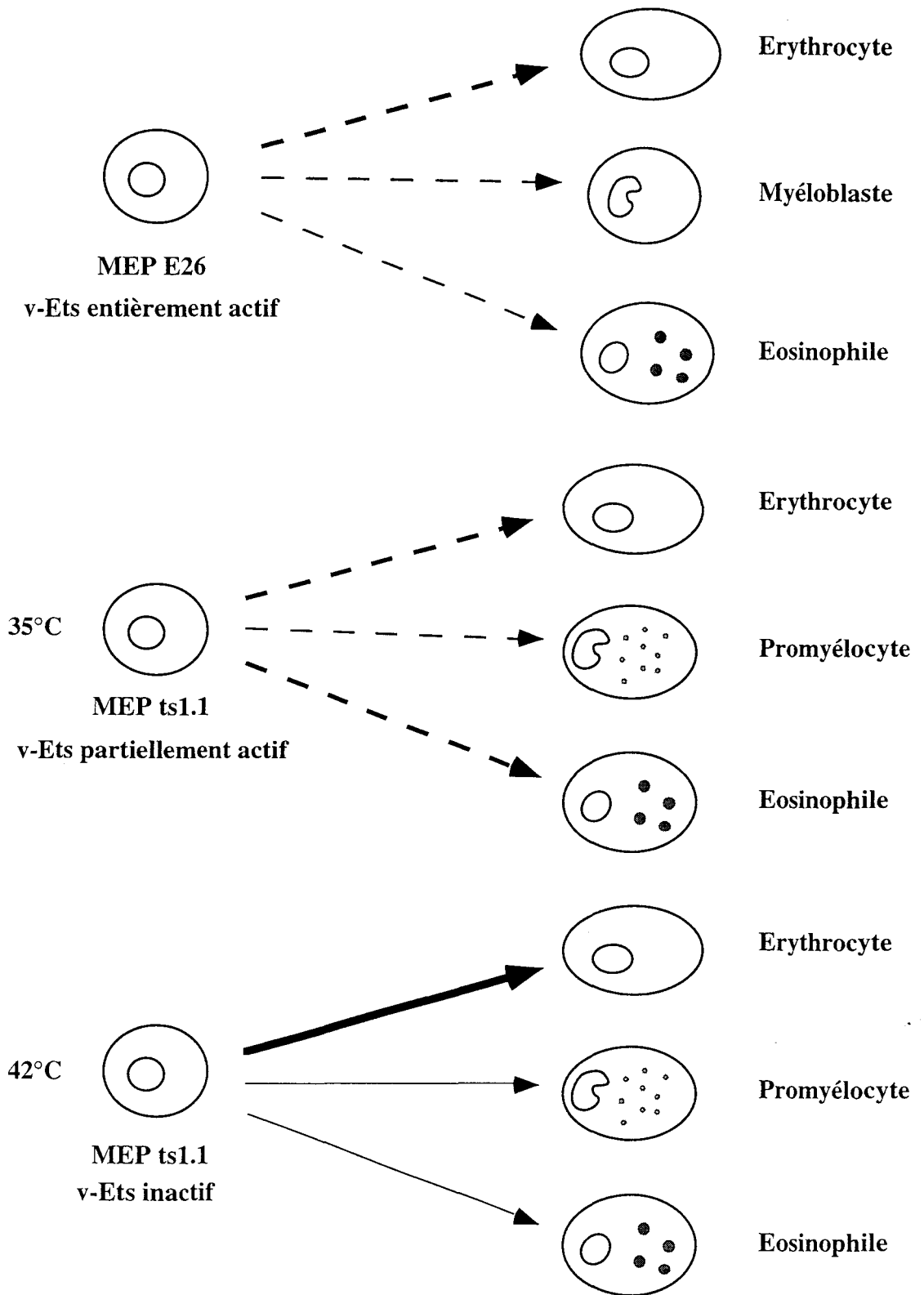


Fig. 9 E26 transforme des cellules souches multipotentes, et requiert l'activité du domaine ETS
 En haut : une faible fraction des cellules transformées par E26 se différencie, et peut donner différents types cellulaires.
 Au milieu et en bas : E26 ts1.1 porte une mutation thermosensible dans le domaine ETS de v-Ets. A température restrictive pour le mutant, les cellules se différencient spontanément.
 MEP : progéniteur hématopoïétique multipotent. D'après Kraut *et al.* (1994).

seul n'est que faiblement transformant (Metz et Graf, 1991b). La **fusion** des deux oncogènes est également importante : un virus portant *v-myb* et *v-ets* séparément est peu transformant (Metz et Graf, 1991a).

2.2.2 v-Ets pourrait rentrer en compétition avec d'autres protéines ETS

Comment expliquer l'effet transformant de *v-ets* ?

Outre le fait d'être fusionné à v-Myb, v-Ets diffère de Ets-1 par deux points (fig. 10):

- il présente deux mutations ponctuelles, dont une dans le domaine ETS, en position 445.
- les derniers acides aminés de la protéine, adjacents au domaine ETS, sont complètement différents de ceux de Ets-1. En fait, la séquence codant ce domaine dans *ets-1* a été inversée dans *v-ets* (Leprince *et al.*, 1993).

Ces différences semblent agir de concert pour modifier la spécificité de liaison de la protéine à l'ADN. *In vitro*, la queue C-terminale de Ets-1 inhibe fortement sa fixation à l'ADN (Hagman et Grosschedl, 1992; Lim *et al.*, 1992; Donaldson *et al.*, 1996). Si on la remplace par la région équivalente de v-Ets, l'inhibition est levée. La mutation du résidu 445 quant à elle diminue la sélectivité de liaison à l'ADN (toujours *in vitro*) (Soudant *et al.*, 1994). Il y a donc simultanément augmentation de l'affinité et diminution de la spécificité. Tout porte à croire que v-Ets reconnaisse des sites plus nombreux que Ets-1, et qu'il s'y lie plus efficacement (Hahn et Wasylyk, 1994).

On peut donc imaginer deux raisons pour expliquer la transformation par v-Ets : 1) la compétition avec d'autres membres de la famille ETS pour la régulation de gènes de différenciation 2) un effet sur des gènes qui ne sont pas normalement régulés par des protéines ETS.

Il est pour l'instant difficile d'argumenter sur la validité de ces hypothèses. Notons toutefois qu'aucune ne rend compte de la nécessité de la fusion v-Myb-v-Ets pour le pouvoir leucémogène *in vivo*. Ces propositions restent valables, mais n'expliquent donc pas tout.

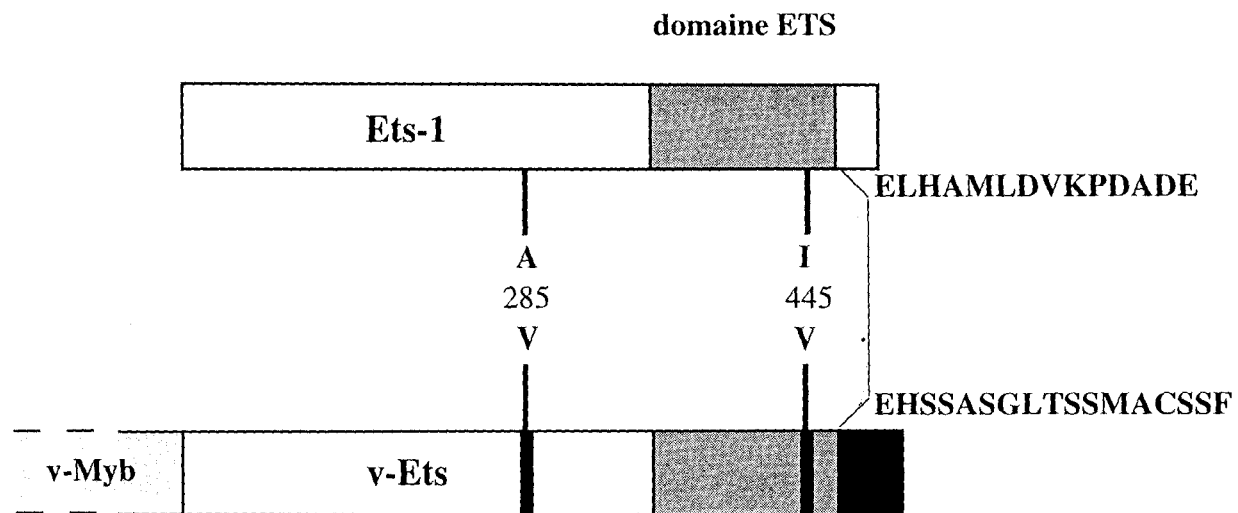


Figure 10 : v-Ets diffère de Ets-1 par sa fusion à v-Myb, deux mutations ponctuelles, et la séquence de son extrémité C-terminale

2.2.3 Transformation par interférence transcriptionnelle

Un autre mécanisme semble intervenir dans la transformation par E26. Il est moins proche de l'intuition que l'idée de compétition ou d'élargissement du spectre de fixation à l'ADN de v-Ets, et requiert quelques rappels préliminaires.

Il existe un antagonisme fonctionnel entre certains récepteurs nucléaires, comme le récepteur de l'hormone thyroïdienne T3 (T3R) ou celui de l'acide rétinoïque (RAR), et les protéines de la famille AP1 (Jun, Fos et leurs homologues) (voir l'introduction de Kamei *et al.*, 1996). L'expression des premiers diminue le pouvoir transactivateur des seconds, et réciproquement, sans qu'il semble y avoir d'interaction physique entre les protéines antagonistes (références citées dans Rasclé *et al.*, 1996). Cet effet pourrait résulter d'une compétition entre les facteurs pour un coactivateur transcriptionnel présent en quantité limitante. La protéine CBP est un bon candidat pour ce rôle (Kamei *et al.*, 1996).

v-ErbA est une protéine oncogénique dérivant de T3R. Elle inhibe la transcription des gènes induits par T3R ou RAR, et, consécutivement, augmente l'effet activateur des protéines AP1 (Desbois *et al.*, 1991). Les cellules de poulet transformées par v-ErbA ont un phénotype comparable à celles qui expriment la p135gag-myb-ets (notée p135). Rasclé *et al.* (1996) ont donc tenté de déterminer si la p135 ne reproduisait pas l'effet de v-ErbA.

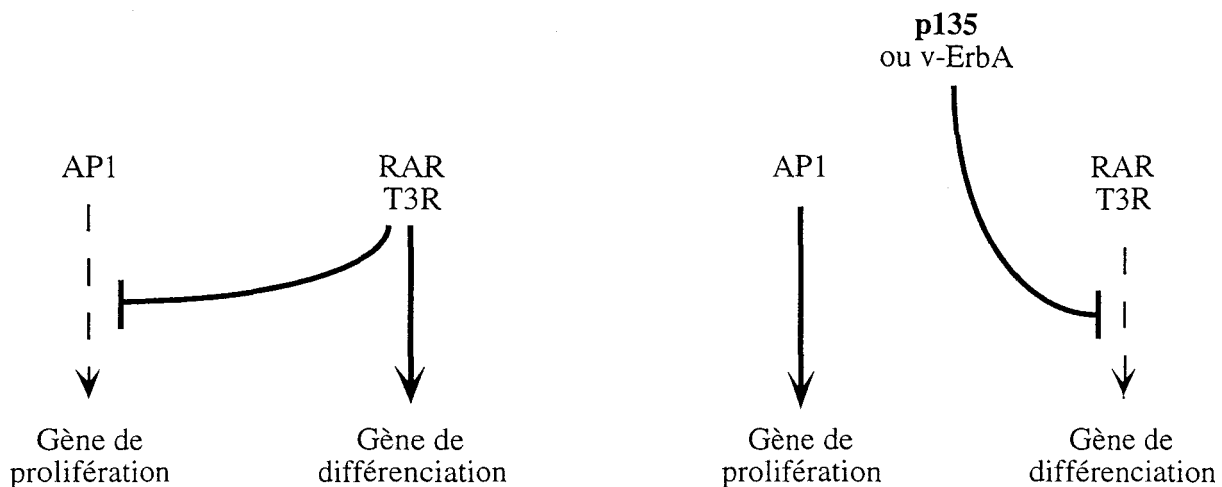


Fig. 11 p135 perturbe l'antagonisme AP1/RAR-T3R

Et, de fait, il apparaît que la **protéine p135 inhibe le pouvoir transactivateur de T3R, mais augmente celui de AP1**, sans interagir physiquement avec ces protéines (fig. 11, Rascle *et al.*, 1996). Cet effet requiert une région de v-Ets qui n'inclut pas le domaine de liaison à l'ADN. Sachant que ce domaine est indispensable à la leucémogénèse *in vivo* (Golay *et al.*, 1988; Kraut *et al.*, 1994), on comprend que l'interférence transcriptionnelle n'est pas non plus suffisante à expliquer la transformation.

De façon remarquable, la présence simultanée de v-Myb et v-Ets ne diminue pas l'activité de T3R : leur fusion est requise (Rascle *et al.*, 1996). On ne peut que spéculer sur les raisons de cet état de fait. En tout cas, il est représentatif de la complexité des relations entre partenaires de fusion dans les protéines chimériques. Soyons donc modestes dans nos prétentions à expliquer le pouvoir transformant des protéines de fusion, qu'il s'agisse de la p135 ou des protéines contenant EWS...

2.3 Quatre gènes ETS recombinaés dans les tumeurs d'Ewing

Les tumeurs d'Ewing sont un groupe de sarcomes et de tumeurs péricphériques du neurectoderme primitif qui touchent surtout les enfants et les jeunes adultes, et dont les caractéristiques cliniques sont assez variables (Delattre *et al.*, 1994). Leur point commun est la prsence dans les cellules tumorales de translocations affectant le chromosome 22. Ces translocations s'accompagnent du rarrangement d'un gène situé sur ce chromosome, baptisé EWS pour Ewing Sarcoma (Delattre *et al.*, 1992). Suite au remaniement chromosomique, la partie N-terminale d'EWS se trouve fusionnée à différents gènes situés sur d'autres loci. **Ces gènes : *fli-1*, *erg*, *er81* (alias *etv1*) et *pea3* (alias E1A-F), sont tous des membres de la famille ETS** (voir fig. 12) (Delattre *et al.*, 1992; Sorensen *et al.*, 1994; Jeon *et al.*, 1995; Urano *et al.*, 1996).

2.3.1 Les fusions EWS-ETS sont transformantes

La translocation peut *a priori* conduire à la production de deux espces de protines chimères : EWS-ETS, ou ETS-EWS. Si nous n'avons pas reprsenté ces dernires dans la figure 12, c'est qu'elles ne sont pas dtectées dans les tumeurs, et n'interviennent sans doute pas dans l'apparition des sarcomes. En revanche, il y a des preuves directes du pouvoir oncogénique des chimères EWS-ETS : leur expression suffit à transformer des cellules *in vitro* (May *et al.*, 1993; Braun *et al.*, 1995; Lessnick *et al.*, 1995).

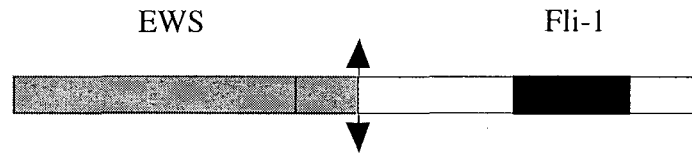
Le site de translocation dans le gène EWS est gnralement localis entre ses exons 7 et 8. Côté ETS la variabilit est bien plus grande : la partie transloquée de *fli-1* peut dbuter aux exons 4, 5, 6, 7 ou 8 (Zucman *et al.*, 1993). Ces diffrences ne semblent pas corrles simplement au phnotype du sarcome. Il est vraisemblable que ce soient les parties communes à toutes les chimères, c'est-à-dire la rgion "CTDII" d'EWS, et le domaine de liaison à l'ADN et la queue C-terminale des protines ETS, qui jouent le rôle principal dans la gnese du sarcome. Abordons notre discussion par un examen des proprits de la chimère EWS-Fli-1, qui est la mieux connue. Nous étendrons ensuite nos observations aux autres protines de fusion.

Translocation

Protéine chimère

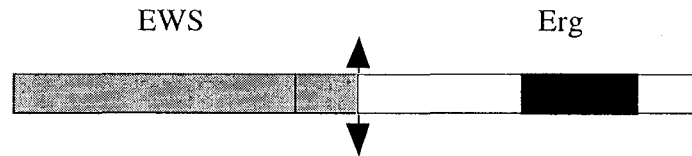
**Fréquence
(% des cas)**

t(11;22)



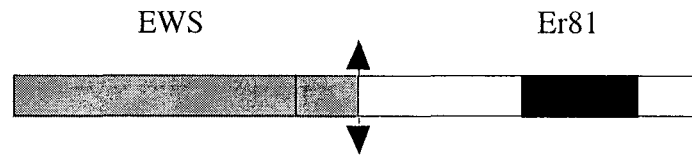
85%

t(21;22)



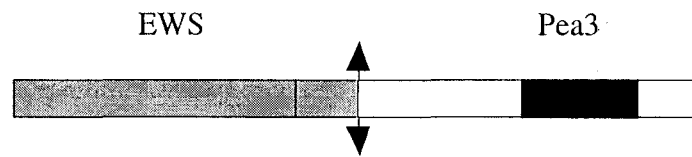
14%

t(7;22)



1%

t(17;22)



<1%



Fig. 12 Quatre gènes de la famille ETS sont fusionnés au gène EWS dans des sarcomes d'Ewing. Le domaine de liaison à l'ADN des protéines ETS est représenté en noir. En bas, la protéine EWS sauvage. La fréquence des cas est une communication personnelle d'Olivier Delattre.

2.3.2 Le pouvoir transformant des chimères EWS-Fli-1 est lié à la dérégulation de cibles de Fli-1

Le caractère transformant des chimères impliquant EWS peut a priori résulter :

- d'une modification de la fonction d'EWS. La chimère pourrait se comporter comme un mutant dominant-négatif vis-à-vis d'EWS, par exemple.
- d'un changement de la localisation et/ou du taux d'expression de *fli-1* qui, après translocation, est contrôlée par le promoteur d'EWS.
- d'une modification de la fonction de *fli-1*.

Les deux premières hypothèses semblent pouvoir être écartées. D'abord, une chimère VP16-Fli-1 est transformante. Donc il n'est pas nécessaire d'interférer avec la fonction d'EWS pour transformer les cellules. Deuxièmement, la protéine Fli-1 sauvage, à la différence d'EWS-Fli-1, n'est pas transformante (May *et al.*, 1993; Braun *et al.*, 1995). Dans le même ordre d'idées, des souris transgéniques pour Fli-1 ne développent pas de tumeurs (Zhang *et al.*, 1993). Il ne suffit donc pas de surexprimer Fli-1 pour transformer des cellules.

On peut donc considérer que ce sont les propriétés de la protéine EWS-Fli-1 qui la rendent intrinsèquement oncogénique. Les deux possibilités les plus simples à envisager sont que la chimère ait 1) des caractéristiques de liaison à l'ADN différentes de celle de Fli-1 2) une capacité de réguler la transcription différente de celle de la protéine sauvage.

On le verra, ces deux modifications semblent en fait se combiner pour générer le pouvoir transformant.

a) Rôle activateur d'EWS

La protéine EWS-Fli-1 peut se fixer aux cibles de Fli-1 (May *et al.*, 1993; Bailly *et al.*, 1994). De plus, elle active la transcription d'un gène dont le promoteur comporte des sites ETS, et ceci avec une "force" supérieure à celle de la protéine Fli-1 sauvage (May *et al.*, 1993; Bailly *et al.*,

1994). On pourrait donc considérer EWS-Fli-1 comme un "super-transactivateur" des gènes-cibles de Fli-1 :

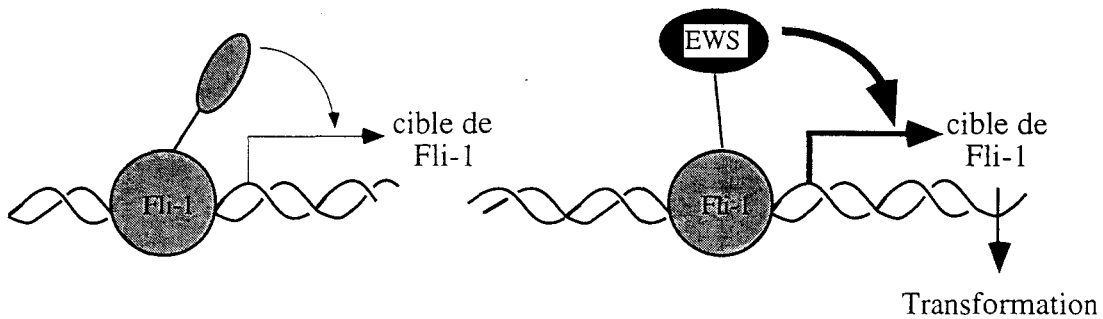


Fig.13 Une hypothèse pour expliquer le pouvoir transformant d'EWS/Fli-1
La chimère stimulerait plus efficacement l'expression des cibles de Fli-1

Au crédit de cette hypothèse, si l'on fusionne la partie de Fli-1 impliquée dans les protéines transformantes à un domaine transactivateur, on obtient dans certains cas une protéine transformante (Lessnick *et al.*, 1995). Cependant, la situation est plus complexe, car on ne peut relier simplement la "force" des domaines transactivateurs au pouvoir transformant des protéines chimères obtenues. Certains activateurs forts génèrent des chimères peu transformantes, et réciproquement certains activateurs modestes forment des protéines très transformantes (Lessnick *et al.*, 1995).

En fait l'explication ne réside peut-être pas dans une variation quantitative de l'activation, mais plutôt dans un effet qualitatif plus subtil. On pourrait proposer que la chimère EWS-Fli-1 soit transformante par sa capacité à stimuler l'élongation transcriptionnelle.

Plusieurs résultats nous poussent à cette audace spéculative :

- on a récemment révélé une identité de séquence surprenante entre EWS et TAF_{II}68. TAF_{II}68 a la particularité de contacter non seulement le complexe TFIID (décrit dans le paragraphe 3.2.2), mais aussi l'ARN polymérase II (Bertolotti *et al.*, 1996). Cette dernière association est particulièrement stable, et il est probable que TAF_{II}68 reste associé à la polymérase au cours de

l'élongation. On peut émettre l'hypothèse que TAF_{II}68, et par extension EWS, puissent jouer un rôle dans l'élongation.

- SF-1 est une protéine associée au facteur Tat du virus HIV1. Celui-ci peut activer l'élongation transcriptionnelle, et cet effet est augmenté par SF-1 (Zhou et Sharp, 1996). La région N-terminale de SF-1 est apparentée à EWS.

- enfin, il existe un autre cas où un facteur stimulant l'élongation, ELL, est fusionné à un domaine de liaison à l'ADN dans une protéine oncogénique (Thirman *et al.*, 1994; Shilatifard *et al.*, 1996). Rappelons aussi que le syndrome de von Hippel-Lindau, une prédisposition héréditaire à différents types de cancer, est dû à une mutation de la protéine VHL, laquelle module l'activité du facteur d'élongation ElonginSIII (voir Li et Green, 1996).

Le lien que nous suggérons entre transformation par EWS-Fli-1 et stimulation de l'élongation est certes ténu, mais il semble pouvoir être mis à l'épreuve assez facilement. On pourrait, par exemple, comparer le pouvoir transformant de chimères activant l'initiation transcriptionnelle (comme TBP-Fli-1) à celui de chimères activant l'élongation transcriptionnelle (TAF_{II}68-Fli-1 ou ELL-Fli-1).

b)EWS-Fli-1 et Fli-1 reconnaissent-ils les mêmes cibles?

In vitro, EWS-Fli-1 semble reconnaître les mêmes séquences que Fli-1 (May *et al.*, 1993; Bailly *et al.*, 1994; Mao *et al.*, 1994). **Cependant, il est clair que par le biais d'interactions avec des partenaires, les protéines ETS peuvent être recrutées à des sites "non-optimaux"** (interaction Ets-1-CBF, paragraphe 1.2.2 et fig. 22, voir aussi paragraphes 4.2 et 4.3). Les tests *in vitro* se déroulent sans ces partenaires potentiels, et négligent donc cette possibilité.

Il se pourrait qu'*in vivo* EWS-Fli-1 interagisse avec des partenaires différents de ceux de Fli-1, et se fixe ainsi à d'autres promoteurs. De fait, en utilisant une technique proche du criblage différentiel, Braun *et al.* (1995) ont mis en évidence des gènes régulés par EWS-Fli-1 qui ne semblent pas répondre à Fli-1 sauvage. L'un de ces gènes est celui de la *stromélysine-1*. On ne

sait pas, toutefois, si Fli-1 se fixe sur ce promoteur sans l'activer, ou s'il ne le reconnaît pas du tout.

2.3.3 *erg*, *fli-1*, *er81* et *pea3* ont-ils des rôles équivalents dans les différents cas de sarcome d'Ewing?

Dans les sarcomes d'Ewing recensés, on trouve la recombinaison EWS/*fli-1* environ 6 fois plus souvent que EWS/*erg*, et presque 100 fois plus souvent que EWS/*er81* ou EWS/*pea3* (Communication personnelle d'O. Delattre, voir fig. 12). On peut interpréter cette distribution de deux façons :

- soit toutes les chimères sont également transformantes, mais le locus *fli-1*, pour des raisons de structure génomique, est plus enclin à la recombinaison.
- soit les loci ont la même capacité à recombiner, mais la chimère EWS-Fli-1, étant plus transformante, est "sélectionnée" plus fréquemment.

Les pouvoirs transformants des 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). Ceci semble supporter notre deuxième hypothèse. Bien sûr, ces tests *in vitro* ne reproduisent que partiellement les conditions de la transformation. Il serait peut-être rassurant, pour confirmer ces données, de comparer la pénétrance des sarcomes dans des souris transgéniques pour les différentes chimères.

On peut donc supposer que les gènes-cibles de Fli-1, Erg, Pea3 et Er81 fusionnés à EWS sont distincts ou modulés avec des intensités différentes par les différentes protéines ETS. Cette idée simple semble effectivement valable : *in vitro*, les chimères les moins transformantes sont celles qui activent le moins la transcription du gène de la *stromélysine-1* (Braun *et al.*, 1995). Encore une fois on ne sait pas si elles se fixent moins bien au promoteur, ou si elles s'y fixent aussi efficacement, mais sans activer la transcription.

Il faut rappeler que dans la partie C-terminale de Erg, Fli, Er81 et Pea3 existe un domaine activateur (Rao *et al.*, 1993; Siddique *et al.*, 1993; Laget *et al.*, 1996; J. H. Chen, communication personnelle). Ses propriétés, et son éventuelle coopération avec le domaine activateur amené par EWS, pourraient être déterminantes pour l'induction des gènes liés à la transformation. Les propriétés fonctionnelles des domaines activateurs sont un des thèmes que nous développerons plus avant (voir aussi discussion de l'article 2).

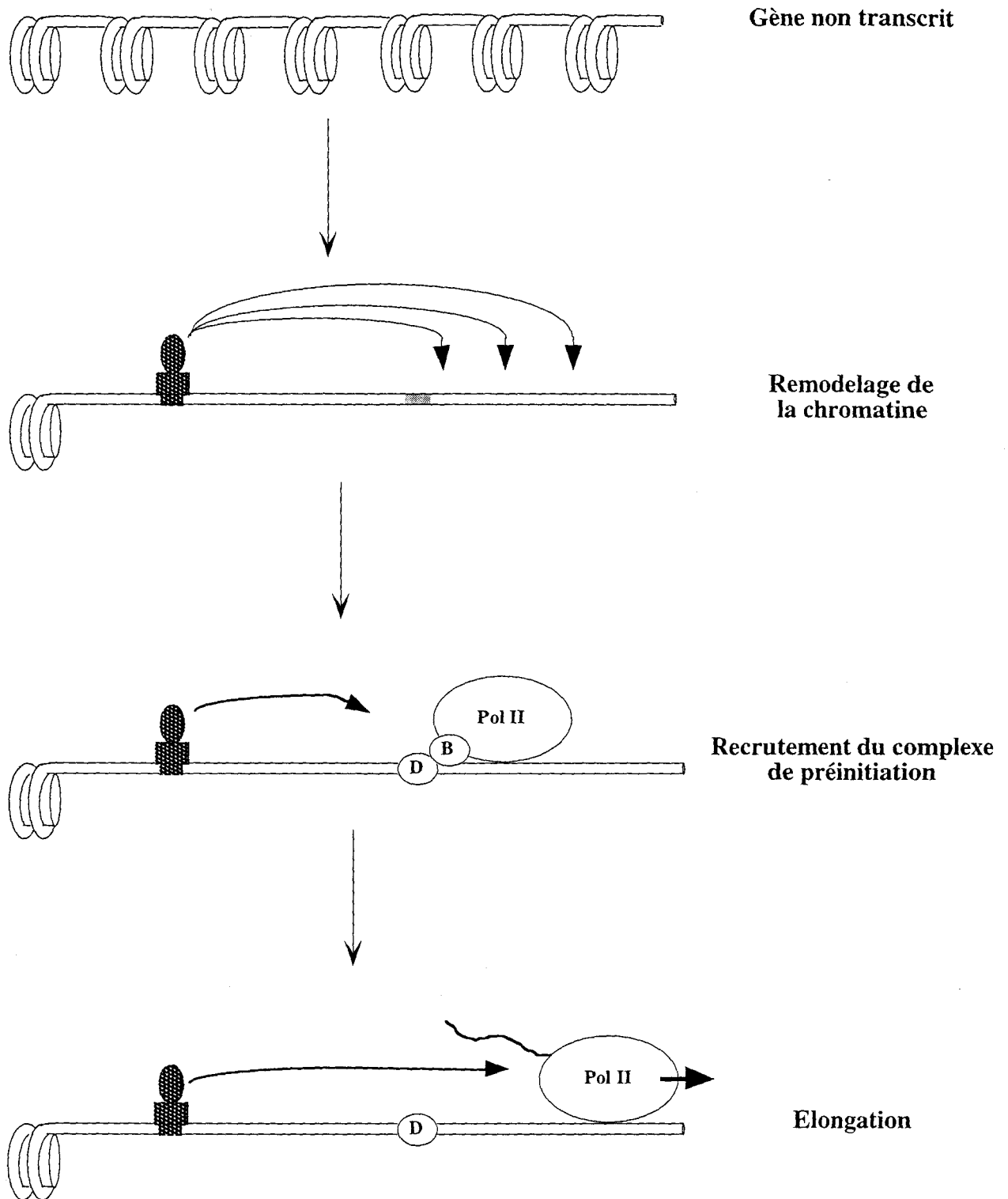


Fig. 14 Les étapes successives de la transcription peuvent être stimulées par les transactivateurs
 En noir, un activateur de transcription. En grisé, l'élément basal du promoteur (TATA Box et/ou Initiateur). TFIID (D), TFIIB (B) sont indiqués. Par simplicité, on a représenté l'ARN polymérase II et ses nombreux cofacteurs comme une seule molécule. Pour une image plus détaillée, voir fig. 18.

3. Les mécanismes de régulation transcriptionnelle

"La transcription par l'ARN polymérase II touche presque tout dans la biologie des Eucaryotes. Elle sous-tend le développement et la différenciation. Elle est la cible des cascades de transduction de signaux. Elle remodèle constamment la cellule en fonction des besoins métaboliques et des informations extérieures." Il est difficile, pour ceux qui étudient les facteurs de transcription, de ne pas partager l'enthousiasme égocentrique de Roger Kornberg...

Cependant, la justification de ces quelques rappels sur la régulation transcriptionnelle n'est pas qu'affective... Notre but, ici, sera de répondre à trois questions qui ressortent des faits précédemment cités :

- **comment des domaines activateurs peuvent-ils coopérer?**
- **pourquoi un domaine activateur peut-il fonctionner dans un type cellulaire et pas dans un autre?**
- **pourquoi, dans une même cellule, peut-il stimuler la transcription d'un gène mais pas d'un autre?**

Pour y répondre, il nous faudra considérer les différentes étapes de la transcription, schématisées dans la fig. 14.

3.1 Le remodelage de la chromatine

3.1.1 Les nucléosomes inhibent la transcription

L'ADN des eucaryotes est enroulé autour de nucléosomes, pour former la chromatine. Les nucléosomes inhibent tant l'initiation de la transcription que l'élongation. **La première étape de l'activation transcriptionnelle est donc la désorganisation des nucléosomes ou leur déplacement.**

Les nucléosomes sont des complexes de huit molécules d'histones (deux de chaque H2A, H2B, H3 et H4). Ces histones peuvent connaître de nombreuses modifications post-traductionnelles, parmi lesquelles la phosphorylation, l'ADP-ribosylation, la méthylation ou l'acétylation (revue dans Bradbury, 1992). De nombreuses études ont montré que l'**acétylation** réduisait l'effet inhibiteur des histones (Lee *et al.*, 1993; par exemple. Voir Brownell et Allis, 1996, pour une revue). Le mécanisme par lequel l'acétylation change le comportement des nucléosomes est encore mal connu; on propose que les groupements acétyls, qui sont greffés sur des lysines et donc neutralisent une charge positive, diminuent l'interaction nucléosome-ADN par un effet électrostatique. Par ailleurs, chez la levure il semble que l'acétylation réduise l'interaction des histones avec les protéines SIR3 et SIR4, qui sont sans doute responsables de la compaction de l'hétéro-chromatine (Hecht *et al.*, 1995). Cette compaction inhibe fortement la transcription (voir Wollfe et Pruss, 1996; Kornberg et Lorch, 1995, pour deux revues sur la compaction de la chromatine, domaine vaste que nous ne ferons qu'évoquer).

L'étude des mécanismes de remodelage de la chromatine est un domaine très actif, et l'on a identifié différents complexes protéiques capables de désorganiser ou de déplacer les nucléosomes. Plus récemment, ce sont les enzymes acétylant ou déacétylant les histones qui ont été impliquées dans les processus transcriptionnels.

3.1.2 Les complexes de remodelage de la chromatine

Chez la levure, les gènes du **complexe SWI/SNF** ont été identifiés en tant qu'acteurs nécessaires à l'induction de nombreux gènes (voir Carlson et Laurent, 1994). Chez des mutants des gènes SWI ou SNF, la chromatine n'est plus remodelée en réponse aux stimuli de

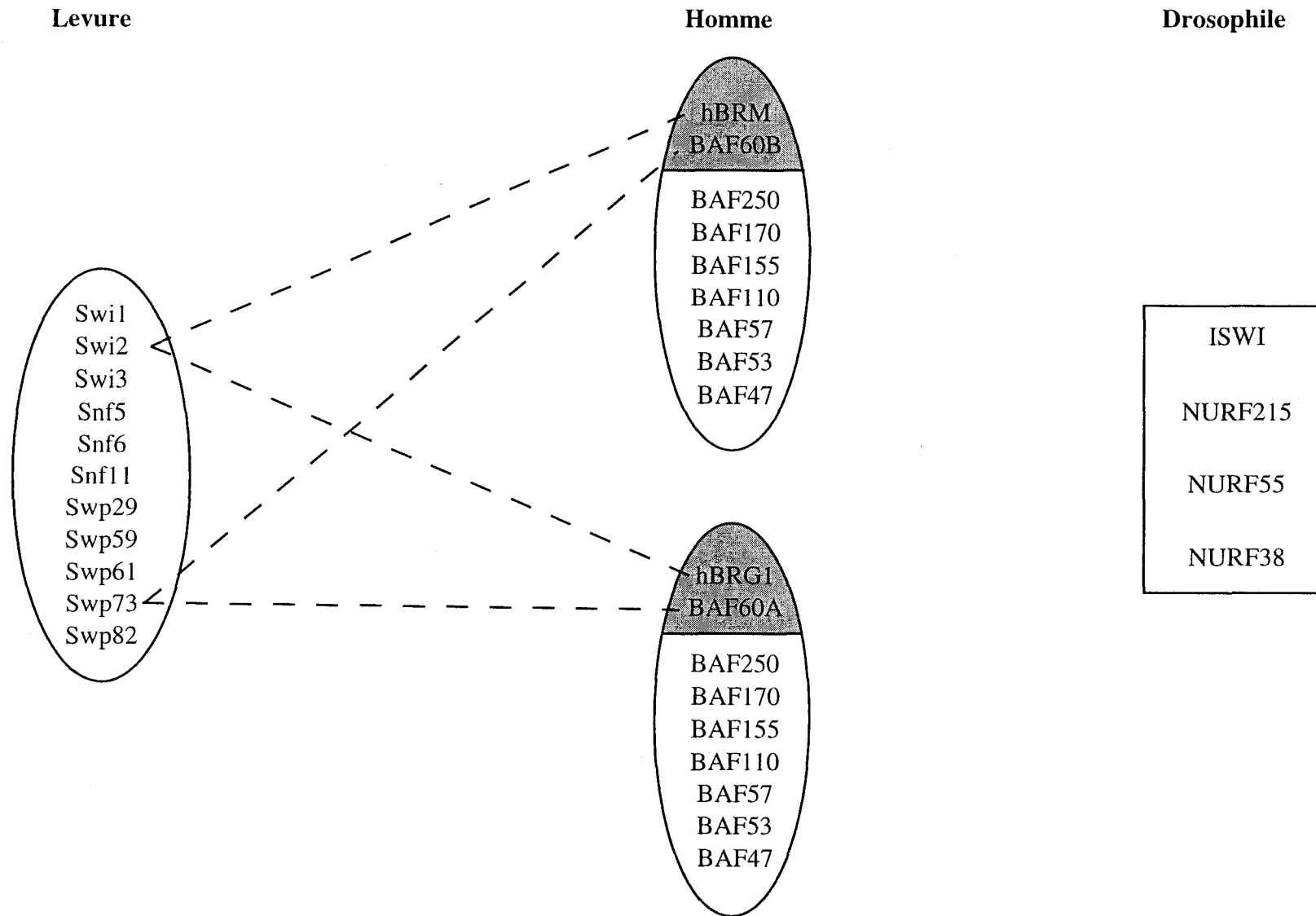


Fig. 15 Les complexes remodelant la chromatine chez la levure, l'homme et la drosophile

Il existe chez la levure un complexe unique dont les sous-unités sont connues. Chez l'homme on trouve au moins deux complexes, dont certaines sous-unités sont communes (en clair), et d'autres distinctes (en gris) Par clarté, on n'a indiqué les homologies homme/levure que pour les sous-unités variables. Le complexe NURF, qui remodèle les nucléosomes chez la drosophile contient le produit d'un gène apparenté à SWI2, mais a des caractéristiques différentes des complexes SWI/SNF.

transcription (Hirschorn *et al.*, 1992). On a pu montrer que les produits des gènes SWI/SNF forment un complexe qui, en présence d'ATP, peut déplacer les nucléosomes ou les désorganiser, et ainsi faciliter le contact entre ADN et protéines impliquées dans la transcription (Coté *et al.*, 1994; voir Peterson et Tamkun, 1995, pour revue). Le complexe SWI/SNF se fixe à l'ADN quadruple brin tel qu'il pourrait se présenter aux points d'entrée et de sortie de nucléosomes (Quinn *et al.*, 1996), mais on ne connaît pas encore le mécanisme par lequel il les déstabilise.

On a identifié dans les cellules humaines un complexe SWI/SNF présentant les mêmes propriétés (Kwon *et al.*, 1994; Imbalzano *et al.*, 1994). Ce complexe a été purifié par chromatographie et certaines de ses sous-unités clonées (voir Wang *et al.*, 1996a; Wang *et al.*, 1996b). Il contient des produits de gènes homologues aux gènes SWI/SNF de levure (voir fig. 15). La situation est toutefois plus complexe chez l'homme que chez la levure, au sens où dans une même cellule humaine il existe plusieurs complexes SWI/SNF, contenant des sous-unités distinctes. De plus, la composition des complexes varie avec les types cellulaires.

Cette situation n'est pas sans rappeler celle des TAFs, les protéines associées à TBP. Alors que chez la levure, il existe un unique complexe TFIID, chez l'homme les TAFs sont plus variées et constituent des complexes hétérogènes (voir paragraphe 3.2.2).

Curieusement, la machinerie de remodelage de la chromatine identifiée chez la drosophile n'est pas un homologue du complexe SWI/SNF de levure.

Le groupe de Carl Wu étudie depuis longtemps le promoteur du gène *hsp70*, dont la transcription est très rapidement induite en réponse au choc thermique. La vitesse de cette réponse tient au fait que, même avant le choc, le promoteur de *hsp70* est occupé par un complexe TFIID et une ARN polymérase qui a déjà synthétisé un court transcrit (Schéma 16, et revue dans Lis et Wu, 1993). Ce promoteur contient des sites pour le facteur de transcription GAGA, que l'on peut donc suspecter de maintenir *in vivo* cet état dégagé des nucléosomes. *In vitro*, en présence d'un extrait d'embryons, on peut former des nucléosomes sur un plasmide portant le promoteur de *hsp70*. L'ajout du facteur GAGA à ce plasmide entraîne un

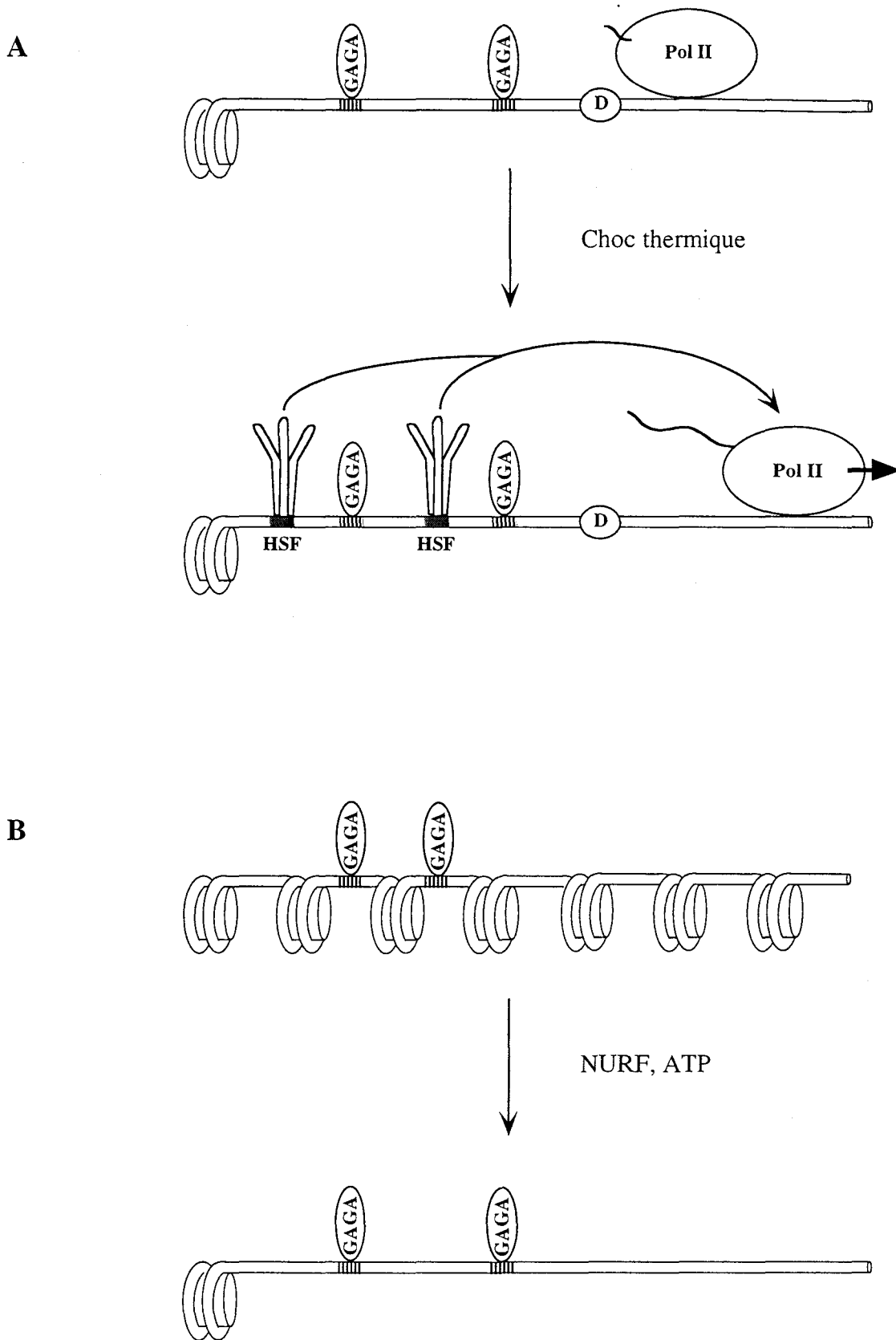


Fig. 16 Sur le promoteur du gène *hsp70*, le facteur GAGA empêche la répression nucléosomale, et la protéine HSF lève l'inhibition d'élongation imposée à l'ARN polymérase II

A) In vivo, à l'état non-activé, le promoteur du gène *hsp70* est dégagé des nucléosomes. Il contient des protéines GAGA et une ARN polymérase engagée transcriptionnellement. Lors du choc thermique, la protéine HSF se fixe au promoteur et libère la polymérase.

B) In vitro, la fixation du facteur GAGA suffit à déplacer des nucléosomes préassemblés sur le promoteur *hsp70*. Ce déplacement requiert de l'ATP et l'activité NURF (NUcleosome Remodeling Factor).

déplacement des nucléosomes (Tsukiyama *et al.*, 1994; voir schéma 16). Ce déplacement requiert de l'ATP et une activité baptisée **NURF** (NUcleosome Remodeling Factor). NURF a été purifié biochimiquement et certains de ses composants clonés (Tsukiyama et Wu, 1995). Bien qu'il contienne ISWI, une protéine apparentée à SWI2/SNF2, le complexe NURF a une taille et des propriétés clairement différentes de celles du complexe SWI/SNF (Tsukiyama *et al.*, 1995; voir schéma 15).

Il existe cependant chez la drosophile un complexe distinct de NURF qui contient l'homologue de SNF2, et que l'on pense être un complexe SWI/SNF (Tsukiyama *et al.*, 1995). Réciproquement, il est probable qu'un équivalent de NURF existe chez l'homme, et peut-être aussi chez la levure.

3.1.3 Acétylation et déacétylation des histones

GCN5 est chez la levure un cofacteur nécessaire à la stimulation transcriptionnelle par des domaines de transactivation acides (Marcus *et al.*, 1994). Cette protéine est engagée dans un complexe avec ADA2 et ADA3, deux autres coactivateurs. Le clonage d'une acétyltransférase d'histones de *Tetrahymena* a révélé que cette molécule était homologue à GCN5, et que GCN5 lui-même possédait une activité acétylase (Brownell *et al.*, 1996). La région de GCN5 qui est nécessaire à l'activation porte l'activité acétyltransférase, ce qui montre la relation entre modification des histones et activation transcriptionnelle.

CBP, et la molécule apparentée p300, sont des coactivateurs de nombreux facteurs de transcription (voir Janknecht et Hunter, 1996). Il est récemment apparu que ces protéines interagissaient avec **P/CAF**, une histone acétyltransférase (Yang *et al.*, 1996). La protéine E1A de l'adénovirus est dotée d'un pouvoir transformant et perturbe la transcription de gènes dont l'activation dépend de CBP (Battacharya *et al.*, 1996). Elle agit sans doute en rentrant en compétition avec P/CAF pour la fixation à CBP (Yang *et al.*, 1996).

Plus récemment encore, on a décrit deux protéines de levure, **SAS2** et **SAS3**, qui possèdent également une zone d'homologie avec les acétyltransférases. Ces protéines ressemblent aussi au produit du gène TIP60, qui interagit avec le transactivateur Tat du virus HIV-1 (Reifsnyder *et al.*, 1996). Notons que SAS2 et SAS3 ne sont pas impliquées dans l'activation, mais la répression transcriptionnelle. Il n'y a donc pas toujours équivalence entre l'acétylation des histones et l'activation transcriptionnelle. Il semble de fait que la répression exercée par les nucléosomes nécessite l'acétylation de certains résidus des histones (Braunstein *et al.*, 1996).

3.1.4 Comment la machinerie de remodelage est-elle recrutée aux gènes induits ?

La machinerie de remodelage et d'acétylation des histones n'a pas, intrinsèquement, de capacité à se fixer à l'ADN de façon spécifique. Il faut donc qu'un recrutement s'opère aux régions qui vont être induites. Deux hypothèses sont envisagées pour ce recrutement (voir schéma 17) :

- il pourrait résulter d'une **interaction directe** entre machinerie de remodelage et transactivateurs. On sait ainsi que le récepteur des glucocorticoïdes (GR) contacte le complexe SWI/SNF (Yoshinaga *et al.*, 1992), et que les domaines activateurs acides interagissent avec le complexe ADA2/ADA3/GCN5 (Silverman *et al.*, 1994). On classera dans la même catégorie le recrutement réalisé via CBP, comme dans le cas de c-Jun (Yang *et al.*, 1996).

- il pourrait également être une **conséquence du recrutement de l'holoenzyme**. Chez la levure, les complexes ADA2/ADA3/GCN5 et SWI/SNF semblent être associés à l'ARN polymérase II dans un "super-complexe" appelé holoenzyme (Wilson *et al.*, 1996). C'est donc peut-être en contactant la machinerie basale que les activateurs recrutent, indirectement, les complexes de remodelage. Le déroulement de ces événements, symbolisés dans le schéma 17, n'est pas encore clair. Notons qu'on a pu identifier chez l'homme un holoenzyme (Chao *et al.*, 1996), mais on ne sait pas encore s'il contient le complexe SWI/SNF ou les acétyltransférases.

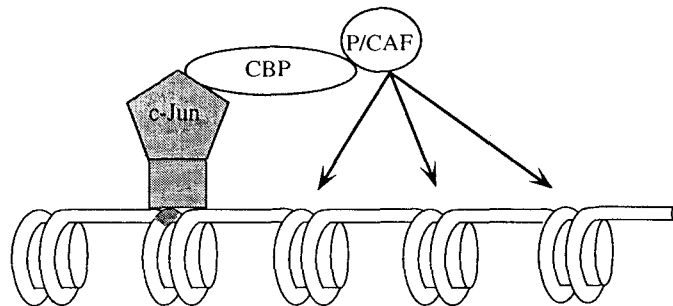
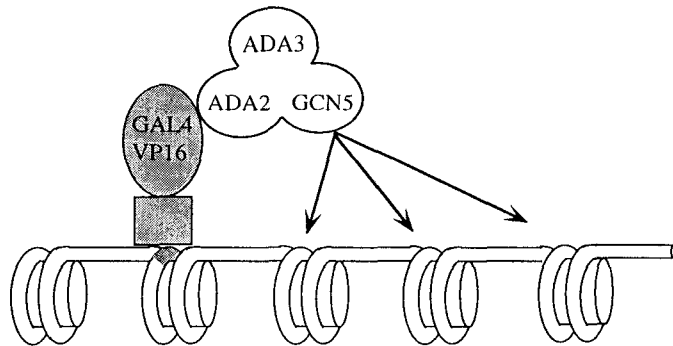
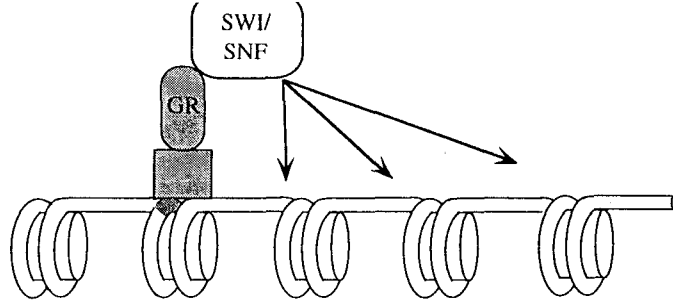
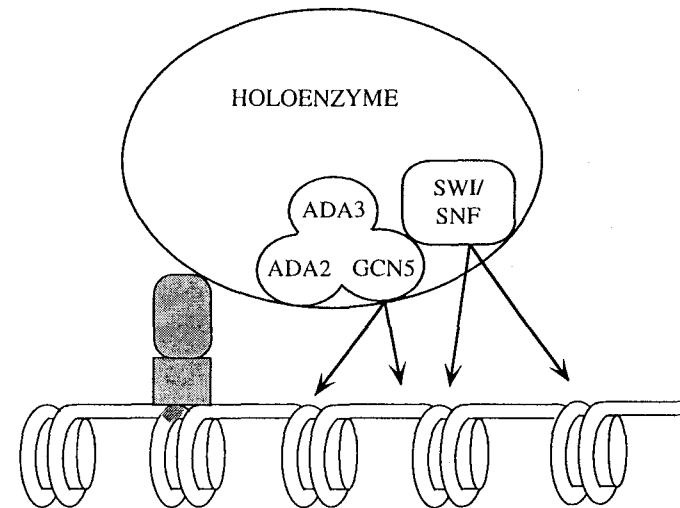
A**B**

Fig. 17 Deux modèles pour le recrutement de la machinerie de remodelage de la chromatine et d'acétylation des histones
 A) recrutement direct par interaction transactivateur-machinerie de remodelage. GR : récepteur aux glucocorticoïdes.
 B) recrutement via l'holoenzyme. On sait que des transactivateurs peuvent recruter l'holoenzyme par l'intermédiaire de certains de ses composants (voir fig. 21).

3.1.5 L'oeuf ou la poule ?

Ces résultats indiquent que la machinerie de remodelage est recrutée aux promoteurs par les activateurs de transcription. Pour que cette situation soit possible, il nous faut admettre que des facteurs de transcription peuvent se fixer à l'ADN même en présence de nucléosomes, puis recruter la machinerie de remodelage.

Il semble en effet, tant chez la levure que le xénope ou l'homme, que **les facteurs de transcription puissent se fixer à l'ADN nucléosomal**. Toutefois, ils ne peuvent stimuler la transcription tant que la chromatine n'a pas été remodelée (voir, par exemple Prioleau *et al.*, 1995). On imagine donc que la séquence des événements aux premières étapes soit:

- 1) Fixation des transactivateurs, qui sont cependant encore incapables de stimuler la transcription.
- 2) Recrutement (direct ou via l'holoenzyme) de la machinerie de remodelage, qui permettra l'accès ultérieur de la machinerie basale de transcription au promoteur.
- 3) Activation de l'initiation de la transcription.

Il est possible que les deux derniers événements soient chronologiquement très proches, étant donné l'association de la machinerie de remodelage avec l'holoenzyme.

3.2 Le recrutement du complexe de préinitiation de transcription

3.2.1 Les multiples composants du complexe de préinitiation

On sait depuis 1979 reconstituer *in vitro* l'initiation de la transcription avec des composants protéiques purifiés (pour toute cette partie, voir revue dans Roeder, 1996). Sept fractions nécessaires à la transcription basale de la plupart des promoteurs ont été purifiées biochimiquement, puis leurs composants identifiés et clonés.

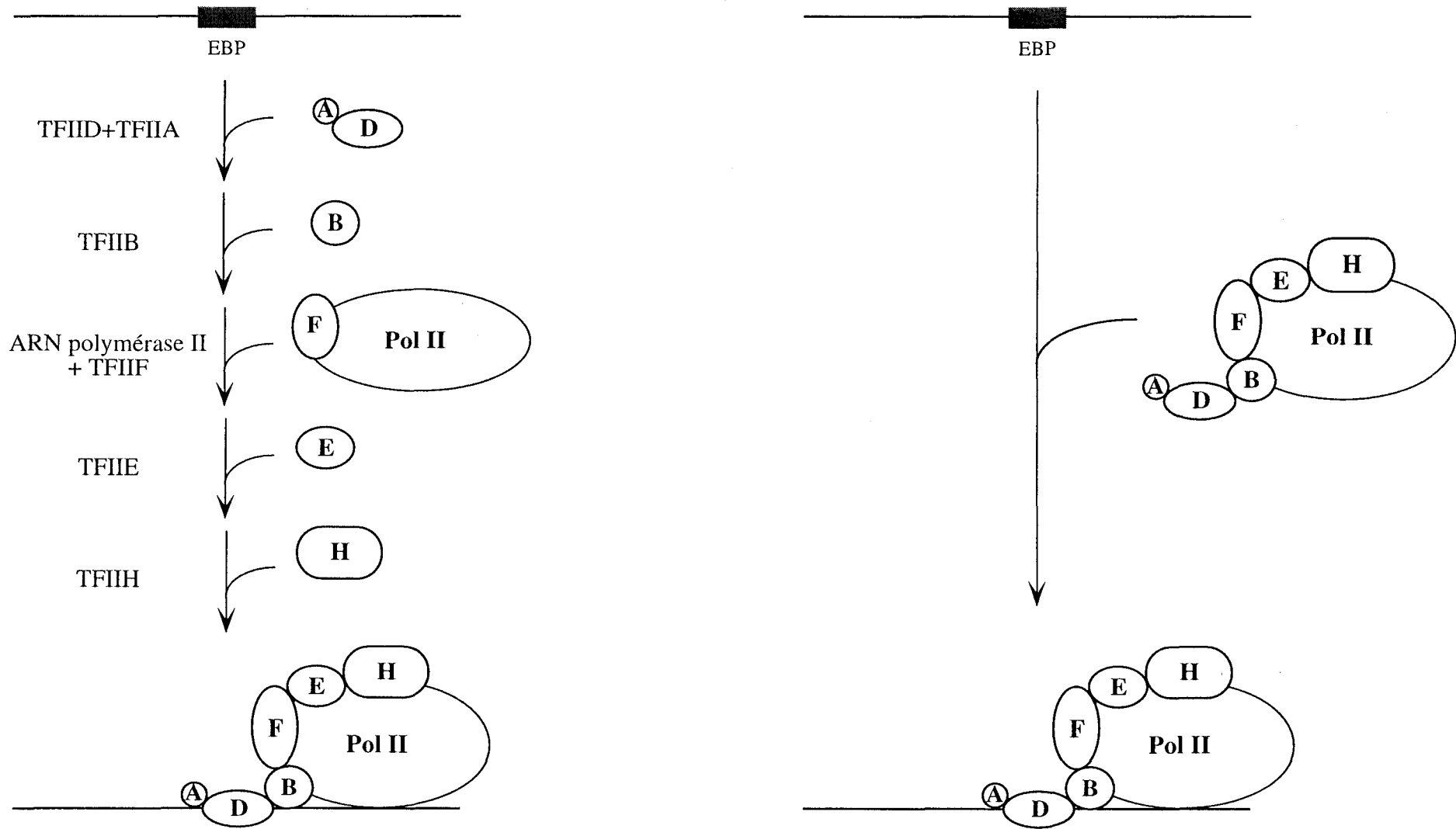


Fig. 18 Deux modèles pour l'assemblage du complexe de préinitiation

In vitro, l'assemblage se fait étape par étape (à gauche). *In vivo*, on observe des complexes préassemblés (à droite). Ces deux modèles ne sont pas nécessairement exclusifs, car la totalité des molécules de polymérase n'est peut-être pas engagée dans des complexes préformés. EBP = Elément Basal du Promoteur.

Les reconstitutions *in vitro* ont permis de définir l'ordre d'action des composants de la machinerie basale (voir aussi schéma 18) :

- le complexe TFIID contient la TATA-Box Binding Protein (TBP), et des protéines associées, les TBP-Associated Factors (TAFs). La première étape de l'initiation est la liaison de TFIID à l'élément basal de transcription, TATA-Box ou Initiateur (Inr). Cette étape est stimulée par TFIIA.

- une interaction avec TFIID permet ensuite à TFIIB de se fixer sur le promoteur.

-TFIIB recrute alors un complexe ARN polymérase II-TFIIF.

-TFIIE interagit avec TFIIF et rejoint alors le complexe.

-TFIIH complète cet édifice en reconnaissant TFIIE.

Ce **complexe de préinitiation** (Pre-Initiation Complex, ou PIC), en présence d'ATP et de nucléotides, peut provoquer la séparation des deux brins d'ADN au site de départ de la transcription, et commencer la polymérisation de ribonucléotides.

Notons que ce schéma de préinitiation connaît d'importantes variations qualitatives en fonction des séquences promotrices. Ainsi, la transcription *in vitro* à partir du promoteur tardif de l'adénovirus, AdMLP, nécessite TFIID et TFIIA, TFIIB, l'ARN polymérase II, TFIIF, TFIIE et TFIIH. Dans les mêmes conditions, la transcription à partir du promoteur de la chaîne lourde des immunoglobulines, IgH, est indépendante de TFIIE (Parvin *et al.*, 1992). Si IgH est transcrit à partir d'un plasmide superenroulé, seuls TFIID, TFIIB et l'ARN polymérase sont requis (Parvin et Sharp, 1993). On imagine que l'énergie libre fournie par le superenroulement se substitue alors à l'activité de TFIIF, TFIIE et TFIIH.

L'holoenzyme déjà mentionné contient, outre l'ARN polymérase II, certains des facteurs généraux de transcription. Il existe donc des **complexes de préinitiation préformés** qui ne suivent pas le modèle séquentiel (voir schéma 18, partie droite). De tels complexes ont été observés chez la levure, le rat et l'homme (revue dans Koleske et Young, 1995; Halle et Meisterernst, 1996), mais on ne s'accorde pas encore sur leur composition. Le schéma séquentiel n'est peut-être pas invalidé, car on ne connaît pas avec certitude la proportion d'ARN polymérase II engagée dans des complexes préformés.

3.2.2 Un paradigme : l'interaction TFIID-ADN

Nous avons décrit la formation du PIC comme un processus séquentiel. Pour l'activer, il suffit donc de rendre plus rapide une de ses étapes (à condition qu'elle soit cinétiquement limitante). La fixation de TFIID à l'ADN est l'une de ces étapes régulables.

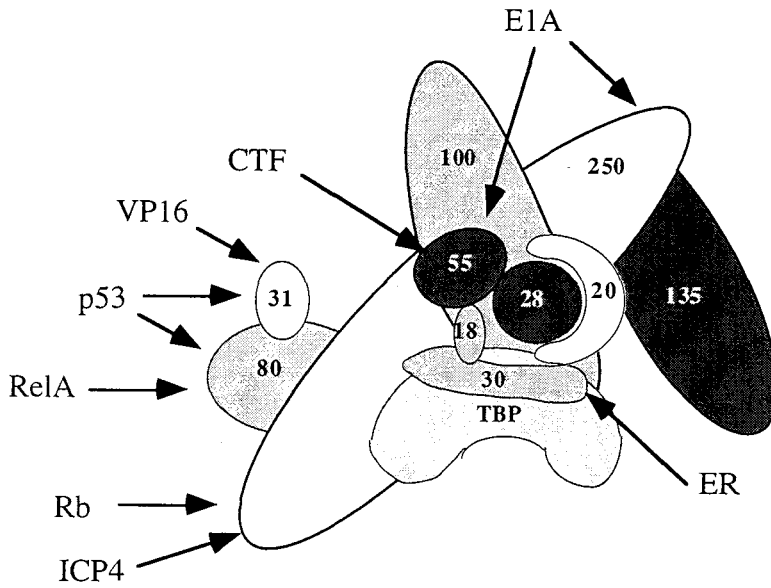
a) Cette étape est limitante

Comme nous l'avons précisé, seuls des points **limitants** de la formation du PIC peuvent *a priori* être activés. Or cette étape est effectivement limitante *in vitro*, et peut être activée par des facteurs de transcription (Chi et Carey, 1993; par exemple). *In vivo* on peut également montrer que le recrutement de TFIID au promoteur en l'absence d'un activateur est très lent (Klein et Struhl, 1994), et qu'il suffit d'amener TBP au promoteur, par exemple en la fusionnant à un domaine de liaison à l'ADN, pour activer la transcription (Chatterjee et Struhl, 1995; Klages et Strubin, 1995; Xiao *et al.*, 1995).

b) Le rôle des TAFs

De nombreux activateurs peuvent interagir physiquement avec TBP. Cependant, la transcription reproduite *in vitro* avec une TBP recombinante n'est pas stimulable par des activateurs à moins d'être additionnée de la fraction TFIID. Cette fraction contient, outre TBP, des protéines qui lui sont associées, les TAFs (TBP-Associated Factors, voir fig. 19; revue courte sur ce sujet dans Verrijzer et Tjian, 1996; plus détaillée dans Burley et Roeder, 1996). On les note TAF_{II}X, où X

A



B

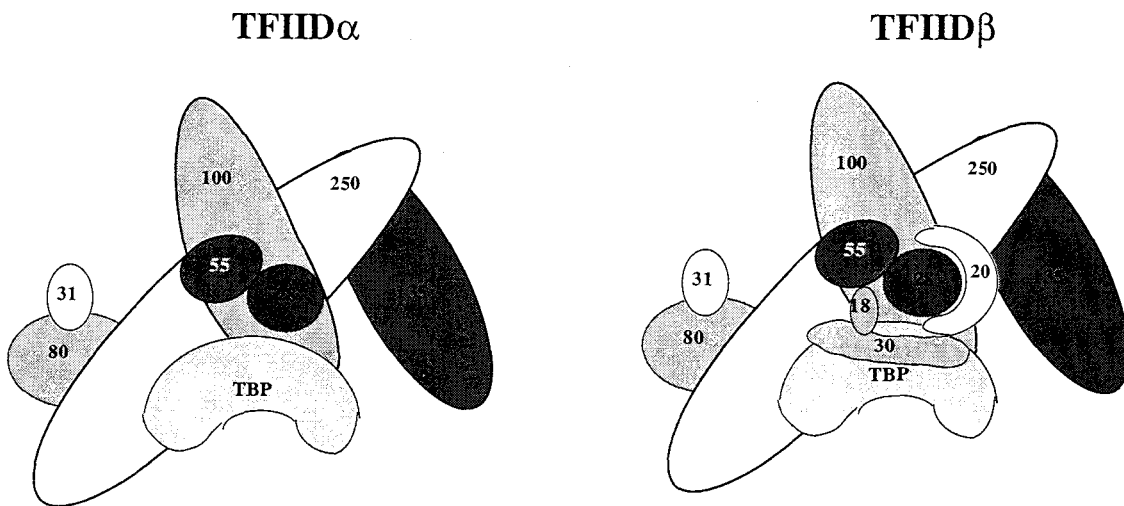
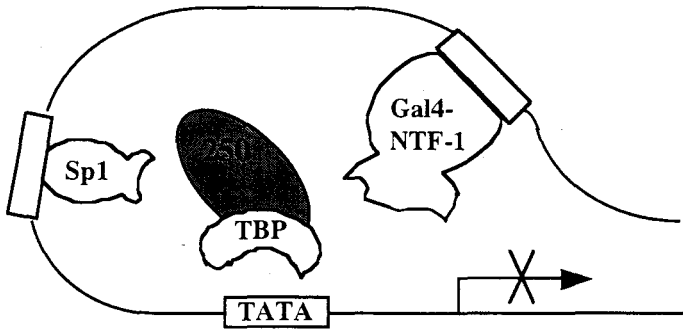


Fig. 19 Le complexe TFIIID est la cible des activateurs

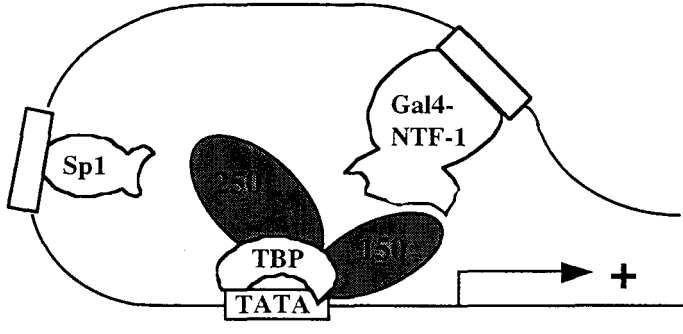
A) Un complexe TFIIID humain (tous les TAFs ne sont pas représentés). On a indiqué les masses des TAF et le nom des activateurs qui les contactent.

B) Dans une même cellule, il existe différents complexes TFIIID. En particulier, certains contiennent TAF₁₈, 20 et 30, et d'autres non.

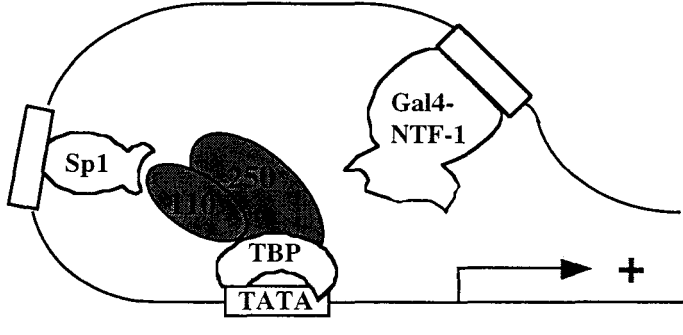
D'après Dubrovskaya *et al.*, (1996).



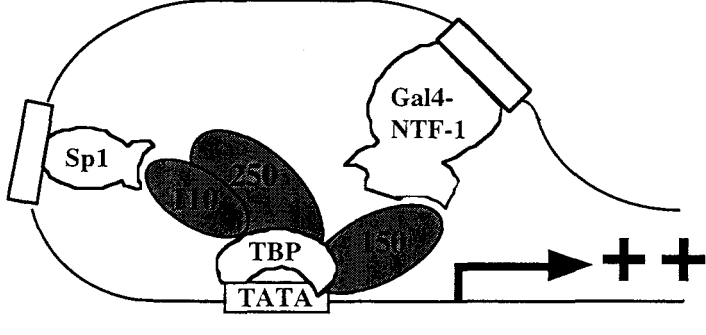
La transcription in vitro réalisée avec TBP et TAF_{II}250 n'est activable ni par Sp1 ni par GAL4-NTF1



GAL4-NTF1, mais pas Sp1, peut activer la transcription en présence de TAF_{II}150



Sp1 transactive en présence de TAF_{II}110



Quand TAF_{II}110 et TAF_{II}150 sont présents, la stimulation est plus importante

Fig. 20 : un modèle pour le recrutement de TBP au promoteur via les TAFs

Sp1 interagit avec TAF_{II}110, et GAL4-NTF1 avec TAF_{II}150. Quand les deux TAFs sont présents, le recrutement est plus efficace et l'on observe une synergie. TAF_{II}250 sert de lien entre ces TAFs et TBP. Les TAFs cités sont ceux de la drosophile, leur masse diffère des TAFs humains.

est leur masse apparente, et II rappelle qu'ils interviennent dans la transcription par l'ARN polymérase II. De nombreux facteurs de transcription interagissent avec les TAFs (fig. 19). **On explique leur effet activateur par un recrutement de TFIID via les TAFs** (voir le modèle proposé par Robert Tjian en fig. 20).

L'interaction de différents activateurs avec des TAFs distincts du complexe TFIID permettrait un recrutement plus efficace et expliquerait certains effets synergiques (Chen *et al.*, 1994; Sauer *et al.*, 1995a; Sauer *et al.*, 1995b).

Ce modèle a plusieurs conséquences fondamentales :

- le grand nombre des TAFs (au moins 10 chez l'homme) permet d'offrir de **multiples cibles potentielles** aux modulateurs de transcription. De fait de nombreuses interactions activateurs - TAFs ont été décrites (Fig. 19).

- chez les mammifères il existe dans une même cellule des complexes TFIID hétérogènes, contenant des TAFs distincts (Jacq *et al.*, 1994; Mengus *et al.*, 1995). Il semble que la séquence du promoteur détermine en partie la nature du complexe TFIID fixé (Verrijzer *et al.*, 1996). La composition du complexe, elle, peut déterminer quel activateur sera actif sur ce gène : si l'on reprend l'exemple de la figure 20, en recrutant un complexe TFIID dépourvu de TAF II110, on empêchera l'action de Sp1. **On peut donc comprendre pourquoi, dans une même cellule, un activateur peut activer la transcription d'un gène mais pas celle d'un autre.**

- un niveau supérieur de complexité dérive du patron d'expression des TAFs qui montre d'importantes variations selon les tissus (Dikstein *et al.*, 1996b). **Il se pourrait donc que, dans un tissu donné, un activateur soit inefficace par suite de l'absence du TAF approprié**. Des résultats expérimentaux appuient cette idée importante. Ainsi, le domaine activateur AF2 du récepteur X de l'acide rétinoïque (RXR) n'active pas la transcription dans les cellules COS. Si on transfecte le gène codant TAFII28 dans ces cellules, on rend AF2 actif (May *et al.*, 1996).

c) Quelques bémols à apporter

La réalité est un peu plus complexe que ne le laisse supposer ce séduisant modèle. Quelques faits méritent d'être mentionnés :

1) D'abord, les interactions transactivateurs-TAFs n'ont pas pour unique effet de **recruter** TFIID. Elles semblent également conduire à un **changement conformationnel** du complexe TFIID, qui pourrait faciliter les étapes ultérieures de la préinitiation (Horikoshi *et al.* , 1988a; Horikoshi *et al.*, 1988b).

2) Ensuite, **il est sans doute simpliste de limiter le rôle des TAFs au recrutement de TFIID** . De fait, les TAFs contactent des facteurs généraux de la transcription autres que TBP. Citons à titre d'exemple le cas de TAF_{II} 68, qui est associé à l'ARN polymérase II, et pourrait ainsi être impliqué dans le contrôle de l'élongation (Bertolotti *et al.* , 1996). TAF_{II} 100 interagit quant à lui avec TFIIF (Dubrovskaya *et al.* , 1996). D'autre part, TAF_{II} 250 possède une activité kinase qui peut s'exercer sur TFIIF (Dikstein *et al.* , 1996a). Ces diverses interactions pourraient expliquer le fait que les TAFs apparaissent stimuler des étapes de la transcription postérieures à la liaison de TFIID (Choy et Green, 1993).

3) Enfin, le plus surprenant. **Chez la levure, la disparition des TAFs n'empêche pas l'activation transcriptionnelle de la majorité des gènes** (Apone *et al.*, 1996; Moqtaderi *et al.*, 1996; Walker *et al.* , 1996). Sachant que les TAFs sont absolument nécessaires à la vie de la cellule (la délétion de la plupart d'entre eux est létale), on peut proposer deux interprétations de ces observations :

- soit les TAFs sont nécessaires à l'activation de la transcription d'un **nombre restreint de gènes** (pas encore identifiés), mais sont dispensables pour l'induction transcriptionnelle de la plupart des autres.

- soit, chez la levure, les TAFs sont en fait **complètement redondants pour l'activation transcriptionnelle**, les activateurs pouvant interagir avec de nombreuses autres cibles de la machinerie de transcription. Leur fonction serait alors tout autre. On sait par exemple que la mutation de certains TAFs entraîne un blocage du cycle cellulaire (Wang et Tjian, 1994; Apone *et al.*, 1996). Ce rôle dans le cycle pourrait être indépendant de la transcription et expliquer le caractère essentiel des TAFs.

Quittons notre gros plan sur cette première étape de l'initiation en précisant que si elle apparaît **souvent** limitante, elle ne l'est peut-être pas toujours. Ainsi, *in vitro*, si la fixation de TFIID à la TATA Box est effectivement limitante dans le promoteur MinL d'EBV (Lieberman et Berk, 1991), elle ne semble pas l'être sur le promoteur E4 de l'adénovirus (Choy et Green, 1993). Quant au gène *hsp70*, *in vivo*, ce n'est clairement pas l'initiation qui limite son expression (fig. 16).

3.2.3 De multiples mécanismes permettent d'activer la formation du PIC

Comment réconcilier les deux modèles d'assemblage du PIC présentés dans la fig. 18? Chez la levure, un modèle semble se dégager (présenté dans Struhl, 1996). La majorité des molécules de polymérase paraît être engagée dans des complexes d'holo-enzyme, mais TFIID n'y est pas présent. Il faut donc, pour réaliser la transcription, recruter indépendamment TFIID et l'holoenzyme. Les différents activateurs peuvent faciliter l'un ou l'autre de ces recrutements, en interagissant physiquement avec une protéine présente dans les complexes (fig. 21). Dans les deux cas, de multiples cibles leur sont offertes :

- **les nombreux TAFs** associés à TBP dans TFIID,
- **les protéines associées à la polymérase dans l'holoenzyme**. Parmi celles-ci se trouvent les facteurs généraux de transcription, notamment TFIIB que de nombreux activateurs contactent. De plus, un ensemble de protéines nécessaires à l'activation forme un complexe appelé "médiateur", qui lie la queue Carboxy-terminale de la polymérase. Plusieurs protéines du

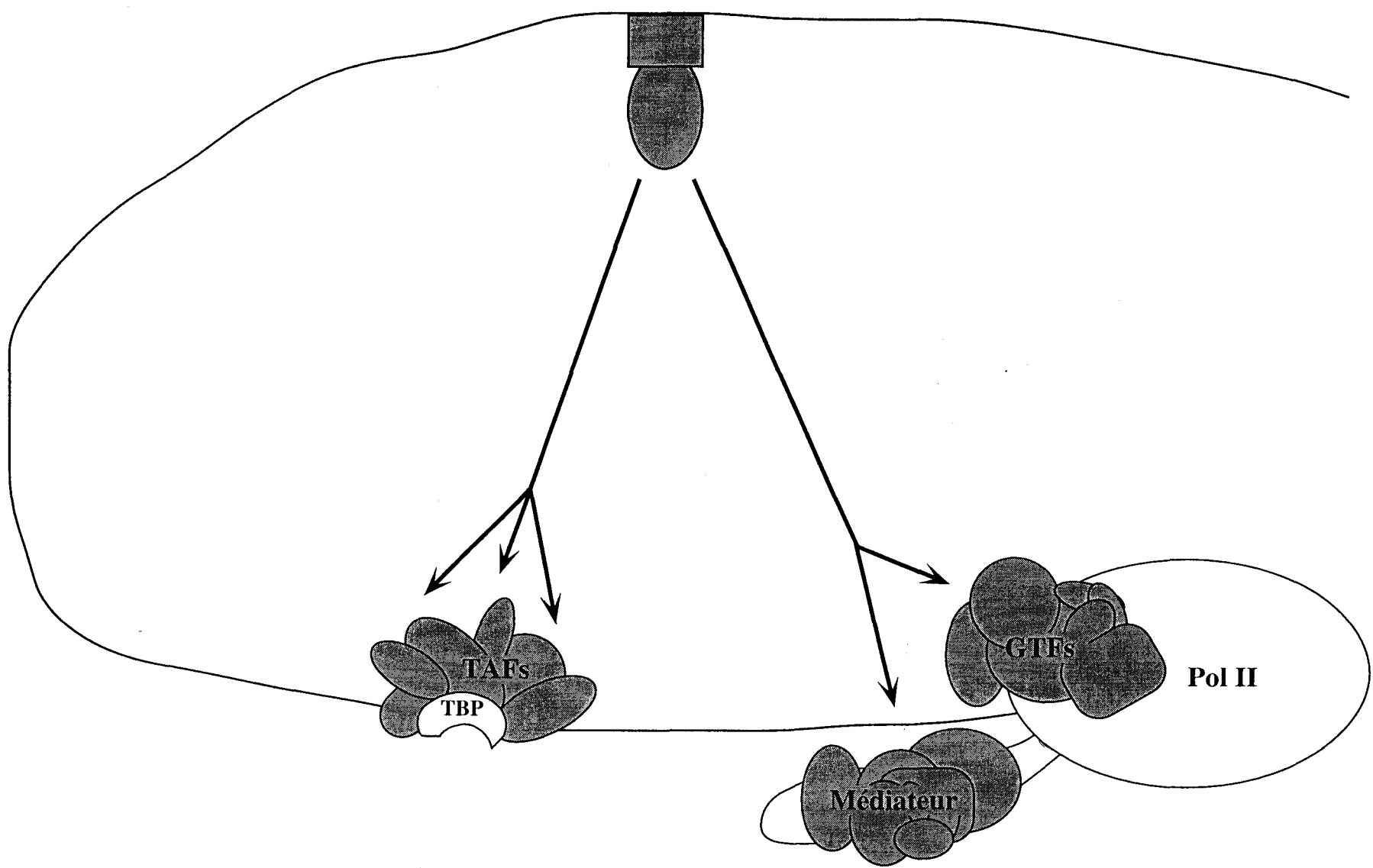


Fig. 21 : un modèle triangulaire pour l'activation de l'initiation

L'initiation requiert la présence de TFIID (à gauche) et de l'holoenzyme (à droite). Les activateurs peuvent fonctionner en recrutant l'un ou l'autre de ces complexes. Dans les deux cas, de multiples possibilités d'interaction activateur-cible existent. GTFs : facteurs généraux de la transcription. D'après Struhl (1996).

médiateur interagissent directement avec des activateurs (Barberis *et al.*, 1995; Farrell *et al.*, 1996; revue dans Bjorklund et Kim, 1996).

3.2.4 Effets de l'architecture du promoteur sur la préinitiation

Nous avons jusqu'ici évoqué des transactivateurs qui interagissent avec les protéines des complexes de transcription. Certains modulateurs de transcription agissent **non en contactant la machinerie de transcription, mais en courbant l'ADN** pour faciliter la formation de complexes nucléoprotéiques. On les appelle facteurs "architecturaux".

Nous avons rappelé plus haut que, selon le degré de superenroulement de la matrice à transcrire, l'exigence vis-à-vis des facteurs généraux de transcription pouvait varier. D'une manière générale, **la conformation de l'ADN joue un rôle fondamental dans l'initiation de la transcription.**

Rappelons que c'est le coude induit par la fixation de TBP qui permet la fixation subséquente de TFIIB, qui ne pourrait reconnaître l'ADN "droit" (Nikolov *et al.*, 1995). De même, il semble que les TAFs puissent jouer un rôle structural : hTAF_{II}80, hTAF_{II}31 et hTAF_{II}20/15, qui possèdent des motifs "histone" pourraient "mimer" partiellement un nucléosome autour duquel l'ADN pourrait s'enrouler (Hoffmann *et al.*, 1996; Xie *et al.*, 1996; Oelgeschlager *et al.*, 1996).

La protéine HMG2 est un transactivateur architectural. Elle facilite le recrutement de TFIID au promoteur (Shykind *et al.*, 1995), sans avoir de domaine de transactivation identifiable. C'est son domaine de liaison à l'ADN qui courbe la double hélice, et l'on sait que, dans certains cas, un promoteur pré-courbé s'associe plus facilement à TFIID (Parvin *et al.*, 1995).

D'autres facteurs architecturaux n'influencent pas la machinerie basale, mais permettent aux autres modulateurs de transcription d'interagir entre eux, ou avec la machinerie de transcription. Une illustration claire de cette idée est fournie par l'enhancer de la chaîne α du récepteur des cellules T (TCR α), qui contient des sites de fixation pour les protéines ATF/CREB, LEF-1, PEBP2 α et Ets-1. LEF-1 est un autre membre de la famille de protéines HMG qui n'active pas

intrinsèquement la transcription, mais est nécessaire à l'effet des autres transactivateurs. Giese *et al.* (1995) ont montré que cet effet de LEF-1 s'expliquait par la courbure qu'il provoque dans l'ADN, et qui facilite les interactions entre les activateurs :

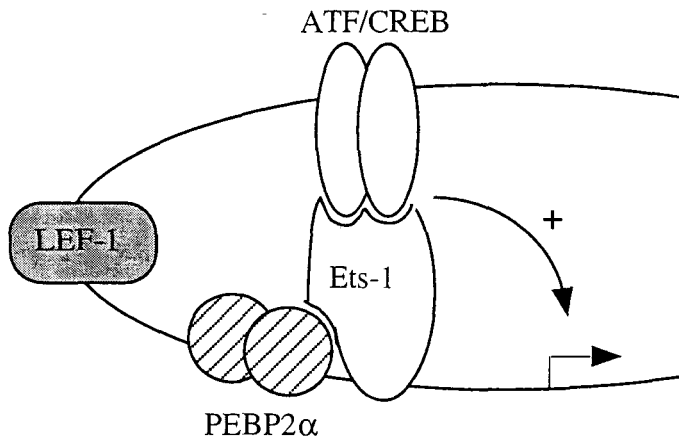


Fig. 22 LEF-1 courbe l'ADN de l'enhancer du TCR α

La conformation de l'ADN induite par LEF-1 rend possibles les interactions entre transactivateurs, et permet la stimulation de l'expression du gène.

Il est intéressant de relever le rôle de Ets-1 dans la régulation de cet enhancer où les effets architecturaux sont visiblement cruciaux. Werner *et al.*, (1995) proposent pour le domaine de liaison à l'ADN de Ets-1 une structure dans laquelle un Tryptophane s'enfonce profondément dans le petit sillon, engendrant un coude important dans l'ADN. Ils suggèrent donc que Ets1 pourrait jouer un rôle architectural (Werner *et al.*, 1996). Notons toutefois que la structure qu'ils ont publiée diffère nettement de celle de Fli-1 ou Pu.1, et qu'elle est controversée (voir Kodandapani *et al.*, 1996).

3.3 Le contrôle de l'élongation transcriptionnelle

3.3.1 L'élongation peut être une étape limitante de la transcription des gènes

Il est apparu assez tôt qu'il **ne suffit pas toujours d'accomplir l'initiation pour induire l'expression des gènes**. Nous avons déjà mentionné le promoteur du gène de réponse au choc

thermique *hsp70* où, avant le choc thermique, un complexe de transcription est déjà assemblé. L'ARN polymérase II y a déjà synthétisé un court messenger, mais est arrêtée à une position définie. Elle est libérée de ce blocage lors du choc thermique, qui induit la fixation de HSF au promoteur (voir fig. 16 et revue dans Lis et Wu, 1993).

Le gène *c-myc* est un autre exemple clair de régulation post-initiation. La technique de "run-on" permet d'estimer la localisation des molécules d'ARN polymérase sur un gène à un moment précis. Appliquée à *c-myc*, elle montre que l'état réprimé est associé à une faible densité de polymérases dans la partie 3' du gène. Lors de l'activation, la quantité de polymérases engagées à proximité du promoteur, qui reflète l'intensité de l'initiation, est peu modifiée. Par opposition, le nombre de polymérases présentes dans la partie 3' du gène, reflétant la processivité, est fortement augmenté (Bentley et Groudine, 1986).

C'est donc la **processivité** des polymérases, et non leur **recrutement**, qui limite l'expression des gènes dans ces deux cas. Notons que ce mécanisme permet une stimulation rapide, puisque l'étape d'initiation est réalisée préalablement à l'induction.

La présence d'ARN polymérases non processives n'est pas une "curiosité" liée à la séquence codante de *c-myc* ou *hsp70*, mais dépend de leurs éléments promoteurs. Ainsi, le promoteur de *hsp70* peut induire la présence de polymérases bloquées sur un gène hétérologue placé en aval (Lee *et al.*, 1992). Il est donc vraisemblable que de nombreux autres gènes connaissent une régulation au niveau de l'élongation. A l'heure actuelle, on sait que tel est le cas pour *hsp26*, *gapdh* 1 et 2, et la *tubuline* B1 chez la drosophile; et pour la *transthyréline* et les gènes de HIV chez les mammifères (revue dans Bentley, 1995).

3.3.2 Des facteurs de transcription peuvent stimuler l'élongation

La transcription initiée au promoteur de HIV-1 génère des produits de pleine longueur ainsi que des produits résultant d'une terminaison précoce de transcription. La protéine E1A active la

transcription de HIV, mais ne modifie pas la nature des transcrits : une partie d'entre eux sont encore tronqués. Par contre, la protéine virale Tat active la transcription en augmentant la quantité de transcrits, mais également en augmentant le pourcentage de transcrits de pleine longueur (Laspia *et al.*, 1989). **Tat est donc un facteur stimulant l'élongation.**

La protéine Tat est particulière en ce sens qu'elle se fixe à l'ARN, sur la séquence TAR, pour activer la transcription. Southgate et Green (1991) ont montré que Tat gardait son activité s'il était recruté sur le promoteur par le domaine de liaison à l'ADN de GAL4. Son pouvoir stimulateur n'est plus alors limité aux séquences transcrites à partir du promoteur de HIV-1. Tat semble donc se comporter comme un facteur de transcription "classique" (c'est à dire, qui se lie à l'ADN).

En fait, d'autres transactivateurs "classiques" activent également l'élongation (Yankulov *et al.*, 1994). On peut même proposer une classification fonctionnelle des facteurs de transcription en trois groupes, selon qu'au cours de la transcription ils activent l'initiation, l'élongation, ou les deux (Blau *et al.*, 1996) :

Classe	Facteurs de transcription	Stimulation de l'initiation	Stimulation de l'élongation
I	Sp1, CTF	+	-
II	Tat	-	+
III	VP16, p53, E2F1	+	+

3.3.3 Comment la processivité vient à la polymérase

L'ARN polymérase II possède un domaine C-terminal (CTD) constitué d'heptapeptides répétés, et qui peut être phosphorylé sur de nombreux résidus. Il existe une corrélation entre la **phosphorylation du CTD** et la capacité processive de l'ARN polymérase (O'Brien *et al.*, 1994). TFIIF possède une activité protéine kinase qui peut s'exercer sur le CTD (Lu *et al.*, 1992). Sachant de plus que TFIIF est nécessaire à l'élongation (Yankulov *et al.*, 1996) et qu'il

interagit avec certains domaines transactivateurs stimulant l'élongation (Xiao *et al.*, 1994; Blau *et al.*, 1996), on peut proposer le modèle de la fig. 23. Il s'accorde bien avec l'observation que Tat ne peut activer l'élongation menée par une polymérase dépourvue de CTD (Okamoto *et al.*, 1996).

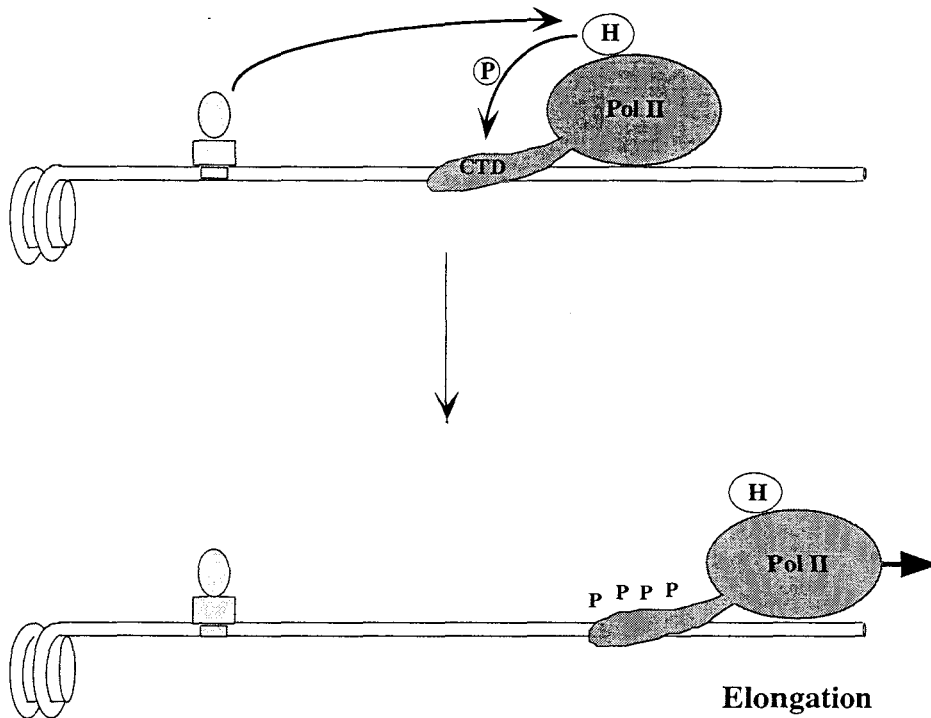


Fig. 23 Les transactivateurs pourraient stimuler l'élongation via TFIIH

Les stimulateurs d'élongation peuvent interagir avec TFIIH. Celui-ci pourrait, en réponse, phosphoryler le domaine C-terminal de l'ARN polymérase II (CTD). Cette phosphorylation (symbolisée par une flèche marquée d'un P) semble faciliter l'élongation.

Il est vraisemblable que TFIIH ne soit pas la seule kinase qui, en phosphorylant le CTD, active l'élongation (voir Dahmus, 1996). Herrmann et Rice, (1995) proposent une activité similaire pour TAK, une kinase associée à Tat.

TFIIF est le seul facteur qui reste associé à la polymérase pendant toute l'élongation, et il stimule la polymérisation *in vitro* (Gu et Reines, 1995). Il interagit également avec certains transactivateurs (Joliot *et al.*, 1995), mais le mécanisme par lequel cette interaction pourrait stimuler l'élongation est encore inconnu.

Quelques autres polypeptides sont connus pour stimuler l'élongation transcriptionnelle (revue dans Reines *et al.*, 1996). Cependant, les mécanismes par lesquels ces facteurs pourraient être recrutés et/ou activés lors de la transcription de certains gènes par des facteurs de transcription sont encore inconnus.

3.4 Que retenir de tout ça?

Nous voilà, semble-t-il, en mesure d'apporter des réponses à nos questions :

- la transcription est une **succession d'étapes dont chacune est régulable**. Des activateurs peuvent donc coopérer en stimulant des étapes différentes, voire la même étape par des mécanismes différents.
- **les mêmes étapes ne sont pas forcément limitantes lors de la transcription de gènes différents**, ou d'un même gène dans des situations différentes. Selon l'étape qu'il régule, un activateur peut donc être actif ou non selon le contexte considéré.
- **l'activité d'un domaine transactivateur dépend de cofacteurs**. L'absence de ces cofacteurs peut le rendre inopérant.

Que ces conclusions péremptoires ne nous abusent pas. De nombreux aspects de la fonction des domaines activateurs sont encore mal connus (voir discussion de la publication n°3). Terminons sur cette note d'indécision, en rappelant que les facteurs de transcription ne sont pas que des facteurs de transcription...

3.5 Post-scriptum : il y a une vie après l'élongation.

La transcription est, clairement, un point de régulation très important, mais il y a loin de l'ARN prémessager à la protéine. Les frontières entre les étapes successives de l'expression génique -transcription, maturation des ARN, transport des ARN, traduction...- sont au fil du temps de

plus en plus fragiles. Les facteurs de transcription n'échappent pas à cette interpénétration des domaines, et semblent impliqués dans de multiples aspects du métabolisme des ARN. Pour nous en convaincre, rappelons ces quelques exemples d'activité "extra-transcriptionnelle" de facteurs de transcription :

3.5.1 La régulation de l'épissage par des facteurs de transcription

L'épissage est un phénomène vraisemblablement concomitant à la transcription. Ainsi, le CTD de l'ARN polymérase II interagit avec des protéines impliquées dans l'épissage (Yuryev *et al.*, 1996). Il est donc compréhensible que des facteurs modulant les premières étapes de la transcription influent également sur le métabolisme des ARN pré-messagers.

Ainsi, Spi/Pu, membre de la famille ETS qui peut activer la transcription, lie l'ARN et la protéine d'épissage P54nrb. La conséquence de ces interactions semble être une diminution de l'épissage (Hallier *et al.*, 1996).

Notons également que la protéine à doigts de Zinc WT1, qui peut activer ou réprimer la transcription, possède un domaine de liaison à l'ARN rappelant celui de la protéine d'épissage U1A, mais dont le rôle n'est pas encore établi (Kennedy *et al.*, 1996).

3.5.2 Contrôle de la localisation des ARN

La protéine TFIIIA est impliquée dans la transcription des gènes de classe III, dont les petits ARN ribosomiaux 5S. On a pu montrer que TFIIIA interagissait avec l'ARN 5S, et que cette interaction était nécessaire à l'exportation hors du noyau des particules ribonucléoprotéiques contenant l'ARN5S RNP (Guddat *et al.*, 1990).

3.5.3 Contrôle traductionnel par des facteurs de transcription

Remarquons enfin que les facteurs de transcription peuvent, parfois, être des "facteurs de traduction"...

Le cas le plus spectaculaire est peut-être fourni par le gène *bicoid* de drosophile. La protéine Bcd possède un domaine de liaison à l'ADN de type *homéo*, et active la transcription de plusieurs gènes. Lors du développement de la mouche, la protéine Cad, produit du gène *caudal*, est seulement exprimée dans les compartiments dont Bcd est absent. Dubnau et Struhl (1996) ont montré que cette exclusion résulte de la liaison de l'homéodomaine de Bcd à l'ARNm de *cad*, qui inhibe la traduction de ce dernier.

4. La fin du début

Nous avons livré jusqu'ici une somme d'informations sans doute difficile à synthétiser. On peut tenter d'en effectuer une intégration, en se posant une question assez simple. Nous avons évoqué quelques-uns des nombreux gènes ETS, et mentionné certaines de leurs fonctions. Quels sont les mécanismes qui permettent une telle diversité de fonctions? Ou, en d'autres termes, qu'est ce qui différencie deux gènes ETS?

Nous avons représenté dans la fig. 24 certains paramètres déterminant "l'identité" de chacun des gènes ETS. Nous souhaitons, à titre de conclusion, discuter quelque peu des trois premiers et montrer, notamment, que les interactions protéines-protéines déterminent en partie la spécificité de liaison à l'ADN.

4.1 La spécificité peut dériver de la différence des patrons d'expression

On l'a dit plus haut (page 9), des lymphocytes B ou T dépourvus de Ets-1 ont des défauts majeurs (Bories *et al.*, 1995; Muthusamy *et al.*, 1995). **Donc, l'absence de *ets-1* dans les cellules recombinées n'a pu être compensée par un autre gène de la famille ETS.** On peut donc penser :

-soit qu'aucune protéine ETS ne peut se substituer à Ets-1.

-soit qu'il existe des protéines ayant des similarités fonctionnelles avec Ets-1, mais qu'elles ne sont pas exprimées où et quand il le faut pour pouvoir s'y substituer.

Certains arguments portent à croire que la seconde hypothèse est vraisemblable. Ets-2, qui est le plus proche parent de Ets-1, est à l'évidence un candidat intéressant pour une éventuelle similarité fonctionnelle. Il se trouve que **Ets-1 et Ets-2 ont effectivement des fonctions partiellement similaires.** La preuve en est fournie par d'élégantes expériences chez la Drosophile où les défauts induits par une altération de *pointed* peuvent être partiellement

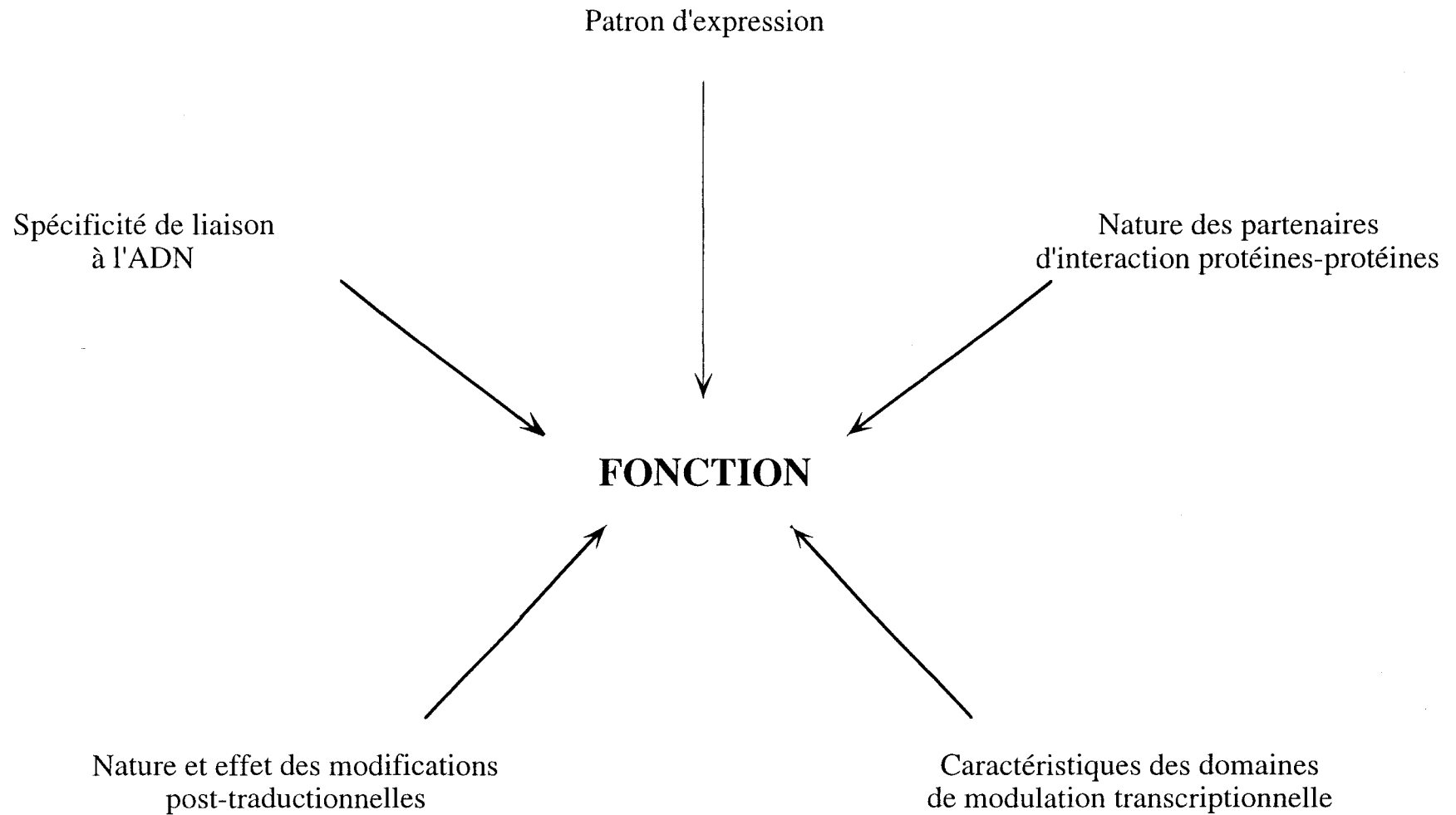


Fig. 24 Quelques-uns des paramètres déterminant la fonction des gènes ETS

corrigés par l'expression d'un transgène codant les protéines Ets-1 ou Ets-2 de poulet (Albagli *et al.*, 1996). Ets-1, Ets-2 et Pnt ont donc certaines capacités fonctionnelles en commun.

Comme Ets-2 est absent des lymphocytes quiescents, on peut donc penser que c'est l'exclusion d'expression qui l'empêche de se substituer à Ets-1. En d'autres termes, **c'est en partie leur patron d'expression qui confère à *ets-1* et *ets-2* leur rôle spécifique.**

Cependant ces deux gènes portent de larges régions non-conservées, et il est probable que d'autres caractéristiques contribuent à les différencier. Une manière plus directe de mesurer l'étendue de la conservation fonctionnelle *ets-1-ets-2* serait de remplacer par recombinaison homologue chez la souris le gène *ets-1* par *ets-2* (ou plutôt, son ADN complémentaire). Si ces souris ne présentent pas les anomalies des animaux *ets-1* *-/-*, la preuve sera faite de la redondance. De telles expériences ont montré qu'*engrailed-2* pouvait totalement se substituer à *engrailed-1* (Hanks *et al.*, 1995). Le pourcentage d'identité *en-1/en-2* n'est pourtant que de 55% sur la longueur de la protéine.

4.2 La sélectivité de liaison à l'ADN distingue-t-elle les gènes ETS?

On a représenté dans la fig. 25 les séquences "optimales" sélectionnées *in vitro* par différentes protéines ETS. S'il n'est pas surprenant d'observer des consensus identiques pour Ets-1 et Ets-2, dont les domaines ETS ne diffèrent que de trois acides aminés sur 85, il est plus étonnant de constater que Ets-1 et Er81, qui ne présentent "que" 70% d'identité dans le domaine de liaison à l'ADN, sélectionnent des séquences fort semblables.

Seule la protéine la plus éloignée de Ets-1, c'est-à-dire Spi-1, semble avoir des préférences nettement distinctes des précédentes. La sélection qu'elle opère en amont du "coeur" GGAA est plus orientée sur des bases puriques, notamment A.

Protéine	Séquences sélectionnées										Référence		
Ets-1			A/G	C	C/A	G	G	A	A	G/A	T/C	N	Fisher et al. (1991)
Ets-2			C	C	A	G	G	A	A	G	T	G	Woods et al. (1992)
Fli-1			A	C	C	G	G	A	A	G	T/C		Zhang et al. (1993)
Erg2				C/G	C/A	G	G	A	A	G/A	T		Murakami et al. (1993)
Er71			G/C	C/G	C/A	G	G	A	T/A	G/A	T/C		Brown et McKnight (1992)
Er81			G/A	G/C	C/A	G	G	A	A/T	G/A	T/C		Brown et McKnight (1992)
Spi-1	A	A	A	G	A/C/G	G	G	A	A	G/c	T	A/G	Ray-Gallet et al. (1995)
Spi-B	A	A	A	G	A	G	G	A	A	G	T	A/T	Ray-Gallet et al. (1995)

Fig. 25 Les séquences optimales sélectionnées *in vitro* par divers membres de la famille ETS diffèrent peu

Protéine

Séquences sélectionnées

Ets-1	In vitro	A/G	C/G	C/A	G	G	A	A	G	T/C		
	LTR HTLV1	G	G	A	G	G	A	A	A	T		
	Enhancer virus polyome	G	C	A	G	G	A	A	G	T		
	TCR alpha	A	G	A	G	G	A	T	G	T		
	GM CSF	A	G	A	G	G	A	A	A	T		
	Ig mu		G	C	G	G	A	A	G	C		
	strom 1	G	C	A	G	G	A	A	G	C		
	ets1	A	A	A	G	G	A	A	G	G		
Consensus	A/G	C/G	A/c	G	G	A	A	G/A	T/C			
Ets-2	In vitro	C	C	A	G	G	A	A	G	T		
	LTR HTLV1	G	G	A	G	G	A	A	A	T		
	PEA3	G	C	A	G	G	A	A	G	T		
	GM CSF	A	G	A	G	G	A	A	A	T		
	strom 1	G	C	A	G	G	A	A	G	C		
	ets2	G	G	A	G	G	A	A	G	T		
	Consensus	G	C/G	A	G	G	A	A	G/A	T/c		
Spi-1	In vitro	A	A	G	A/C/G	G	G	A	A	G/c	T	A/G
	enhancer SV40	A	A	G	A	G	G	A	A	G	C	T
	Ig kappa	T	T	G	A	G	G	A	A	C	T	G
	Ig mu	T	T	T	G	G	G	A	A	A		
	scavenger receptor	A	A	G	G	G	G	A	A	G	A	A
	Ig lambda	A	A	A	A	G	G	A	A	G	T	G
	Il 1	A	A	G	C	A	G	A	A	G	T	A
	Consensus	A/T	A/T	G/t	G/A/c	G/a	G	A	A	G/c	T/a	A/G

In vitro : séquences optimales obtenues par sélection, LTR : Long Terminal Repeat, TCR : T Cell Receptor
 GM CSF : Granulocyte/Macrophage CytoStatic Factor, Ig : immunoglobuline, strom : stromélysine
 Il : Interleukine. D'après Romano-Spica et al. (1994) et Kominato et al. (1995)

Fig. 26 Les éléments de réponse aux protéines ETS observés dans les promoteurs présentent des séquences "non optimales"

Il vient immédiatement à l'esprit que :

- les séquences "optimales" sélectionnées *in vitro* ne sont peut-être pas les mêmes qui sont liées *in vivo*.

- d'autre part, il peut y avoir une nette différence entre les sites liant la protéine *in vitro*, et ceux qui permettent à celle-ci d'exercer son effet de modulation de la transcription. Huang *et al.* (1996) ont particulièrement bien illustré ce principe avec MyoD, qui ne transactive que fort peu quand il est fixé à la séquence "optimale" obtenue *in vitro*.

Examinons donc quelques sites "naturels", effectivement régulés par des protéines ETS (fig. 26, compilation trouvée en partie dans Romano-Spica *et al.*, 1994; Kominato *et al.*, 1995). Il apparaît, particulièrement pour Pu.1, que les sites reconnus s'approchent du consensus *in vitro*, mais qu'une même protéine peut reconnaître des sites fort différents. Il semble donc y avoir une "flexibilité" de reconnaissance de l'ADN notable.

Permettons-nous ici un parallèle. Il existe d'autres familles de facteurs de transcription où les différents membres jouent manifestement des rôles distincts, alors que leur spécificité de liaison à l'ADN est presque indistingable *in vitro*. C'est le cas, notamment, des protéines à homéodomaine (Hayashi et Scott, 1990). La structure de leur domaine de liaison à l'ADN, organisé en Hélice-Boucle-Hélice, se rapproche d'ailleurs un peu du domaine ETS.

Il est clair que les protéines à homéodomaine reconnaissent chacune des sites spécifiques *in vivo* (Walter et Biggin, 1996). L'apparition de ces différences est liée à l'interaction avec des cofacteurs qui favorisent la fixation sur certaines séquences. Le mieux connu de ces cofacteurs est d'ailleurs une protéine à homéodomaine, Extradenticle (van Dijk et Murre, 1994; revue dans Mann et Chan, 1996).

De même, pour les protéines ETS, il est vraisemblable que **les différences intrinsèques de spécificité de liaison à l'ADN ne suffisent pas à rendre compte des différences fonctionnelles**. On pourrait dire que les protéines ETS ont une sélectivité "innée", qui sans

doute est distincte entre Ets-1 et Pu.1, mais pas entre Ets-1 et Ets-2, et une sélectivité "acquise" par recrutement à l'ADN via des interactions protéines-protéines différentes.

La liaison de Elk à la séquence SRE, que nous traiterons maintenant, en est un exemple convaincant.

4.3 Les interactions protéines-protéines peuvent distinguer les protéines ETS

En réponse au sérum ou aux facteurs mitogènes, de nombreux gènes sont induits de façon très rapide. Dans les promoteurs de la plupart de ces gènes, on trouve un élément commun qui médie l'induction : c'est la séquence SRE (Serum Response Element). Le prototype de ces gènes "de réponse immédiate" est *c-fos*, qui présente un SRE à proximité du site d'initiation de la transcription (revue dans Treisman, 1995).

Les SRE sont constitués d'une séquence de liaison pour ETS juxtaposée au site de fixation d'un homodimère de la protéine SRF (Serum Response Factor).

Trois protéines ETS peuvent former avec SRF un complexe ternaire sur le SRE, on les appelle collectivement TCFs (Ternary Complex Factors). Il s'agit de Elk, Sap1 (qui a deux isoformes, Sap1a et Sap1b) et Sap2, qui est la protéine humaine homologue à Erp et Net (Hipskind *et al.*, 1991; Dalton et Treisman, 1992; Lopez *et al.*, 1994; Giovane *et al.*, 1994). Ces trois protéines présentent trois régions de forte identité de séquence, les boîtes A, B et C (fig. 27). La boîte A contient le domaine ETS.

On retrouve chez les TCFs le caractère "polyvalent" décrit pour Pu.1 (fig. 5). Ces protéines peuvent se fixer de façon autonome à certains sites ETS, comme la séquence E74, mais sur la séquence SRE elles doivent nécessairement être recrutées par SRF. L'interaction SRF-TCF met en jeu le domaine de liaison à l'ADN du premier, et la boîte B du second.

Il existe une forme alternative de Elk, Δ Elk, qui est dépourvue de boîte B (Rao et Reddy, 1993). Cette protéine lie toujours E74 mais, même en présence de SRF, elle ne reconnaît pas SRE. En

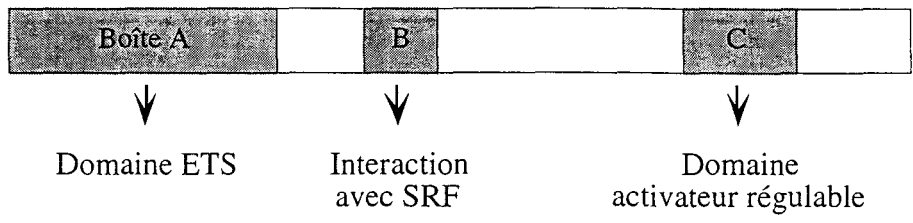


Fig. 27 Trois régions des TCFs sont bien conservées

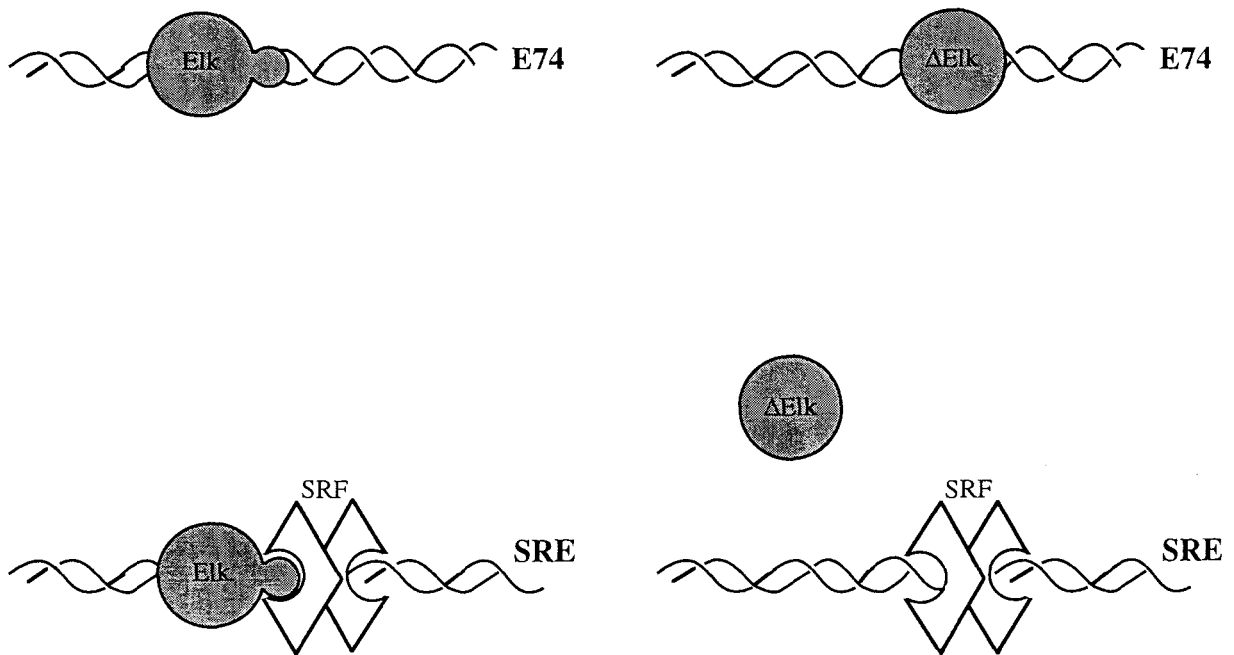


Fig. 28 La différence de capacité à interagir avec SRF différencie Elk et Δ Elk

conséquence elle est incapable d'activer la transcription du promoteur de *c-fos* (Rao Reddy, 1993). La différence fonctionnelle Elk/ Δ Elk tient donc à leur capacité différente à interagir avec SRF (fig. 28).

4.4 Le début de la fin

La manne des illustrations convaincante de notre théorie de la fig. 24 est déjà tarie. Il n'existe pas d'exemples très parlants où la différence entre deux protéines ETS tiennent clairement à la nature des modifications post-traductionnelles. Citons quand même Elk et Sap1 qui sont coexprimés dans de nombreux tissus, semblent reconnaître les mêmes séquences-cibles, forment des complexes ternaires avec la même efficacité, mais seraient les cibles de kinases différentes, (Hipskind *et al.*, 1994; Strahl *et al.*, 1996).

Quant au dernier point, il y a des cas où deux protéines ETS lient le même site *in vitro*, mais une seule l'active *in vivo* (Oettgen *et al.*, 1996; Miyazaki *et al.*, 1996). Ceci suggère une différence de caractéristiques des domaines activateurs, mais en fait il n'est pas prouvé que les deux protéines lient effectivement le promoteur *in vivo*. Hors de la famille ETS existent par contre des exemples très clairs de domaines activateurs actifs sur certains promoteurs mais pas sur d'autres (Nagpal *et al.*, 1992; Tanaka *et al.*, 1992; Das *et al.*, 1995).

Nous espérons avoir réussi à dégager dans cette introduction les principes qui nous guideront dans l'exposition des résultats. Gardons en tête que les gènes ETS s'inscrivent dans des cascades de transduction de signaux qui leur permettent de jouer des rôles clés dans le développement et l'oncogénèse. Leur rôle physiologique est déterminé par différentes caractéristiques fonctionnelles que l'on peut étudier séparément, mais qui s'influencent entre elles.

RESULTATS

Publication n°1

**Molecular cloning et characterization of human ERM, a new member of the Ets family
closely related to mouse PEA3 et ER81 transcription factors.**

Monté D., Baert J.L., Defossez P.A., de Launoit Y. et Stéhelin D.

Oncogene (1994), **9**, 1397-1406

Résumé:

Cet article décrit l'identification d'un nouveau membre de la famille ETS, *erm*.

Sa séquence l'apparente plus particulièrement à deux gènes ETS découverts auparavant, *pea3* et *er81*. Nous montrons que la protéine Erm peut se fixer à une séquence canonique ETS, via un domaine de 83 acides aminés.

L'expression d'*erm* est détectable en Northern Blot dans de nombreux tissus humains, dont le cerveau, le rein, les muscles squelettiques et le placenta. Ce patron d'expression ne recoupe que partiellement celui de *pea3* et d'*er81*.

Donc, même si les trois gènes présentent une grande similitude de séquence, il est probable que leurs patrons d'expression leur confèrent des rôles physiologiques distincts.

Molecular cloning and characterization of human ERM, a new member of the Ets family closely related to mouse PEA3 and ER81 transcription factors

Didier Monté, Jean-Luc Baert, Pierre-Antoine Defossez, Yvan de Launoit & Dominique Stéhelin

Unité d'Oncologie Moléculaire, CNRS URA 1160 - Institut Pasteur de Lille, 59019 Lille Cedex - France

The *ets*-related transcription factors PEA3 and ER81 have recently been isolated and characterized in the mouse. They share 95% identity in a 85 amino acid (AA) domain termed the ETS domain which is responsible for DNA binding, and therefore they form an Ets family group. By screening a human testis cDNA library with a probe containing the mouse PEA3 ETS domain, we isolated a 2.2 kb clone containing a 510 AA open reading frame. Since the ETS domain, which is localized in the carboxy terminal region of the encoded protein, is 95% and 96% identical to that of PEA3 and ER81, respectively, we named this new member 'Ets Related Molecule PEA3-like' (ERM). Although the first 120 AA in the amino-terminal region of ERM share 47% identity with PEA3 and 66% with ER81, ERM contains a central region of ≈ 35 AA not found in the two mouse proteins. Gel shift analysis indicates that the full-length ERM protein is able to bind specifically to an oligonucleotide containing the consensus nucleotide core sequence GGAA recognized by the Ets proteins. Moreover, *in vitro* translation of 83 AA of the ERM ETS domain led to the production of a truncated protein which also binds to DNA. Though differential expression is observed in primary tumors and normal lymphocytes do not express ERM, this gene is almost ubiquitously expressed in human normal tissues. ERM mRNA is highly expressed in brain as well as in placenta and, to a lesser degree, in lung, pancreas, and heart. Moreover, almost all human cell lines tested express it at varying levels. In mouse tissues, we showed that PEA3 and ER81 mRNAs display restricted expression, whereas ERM is almost ubiquitously expressed as observed for human tissues. Altogether these results indicate that ERM is clearly a new *ets* family member and not the human equivalent of PEA3 or ER81.

Introduction

The *ets*-related genes encode a family of transcription factors involved in the regulation of gene expression (MacLeod *et al.*, 1992; Seth *et al.*, 1992). *c-ets-1* (Leprince *et al.*, 1983; Nunn *et al.*, 1983) was the first gene

found in this increasingly growing family, that includes in Vertebrates *c-ets-2* (Watson *et al.*, 1988), *erg* (Reddy *et al.*, 1987), *elk-1* (Rao *et al.*, 1989), SPI-1 and SPI-B (Moreau-Gachelin *et al.*, 1988), SAP-1a (Dalton & Treisman, 1992), *fli-1* (Ben David *et al.*, 1991), GABP α (LaMarco *et al.*, 1991), *elf-1* (Thompson *et al.*, 1992), ER71 (Brown & McKnight, 1992), PEA3 (Xin *et al.*, 1992) and ER81 (Brown & McKnight, 1992). All the encoded proteins contain a DNA-binding domain of ≈ 85 AA, named the ETS domain (Karim *et al.*, 1990), which may be structurally related to the Myb DNA-binding domain (Laget *et al.*, 1993). The polyomavirus enhancer (PyE) is composed of multiple DNA sequence elements serving as binding sites for transcription factors. One of these proteins that binds to the PyE was defined as the *PyE* Activator 3 (PEA3), which is capable of protecting the core sequence 5'-AGGAAG-3' (the PEA3 motif) (Martin *et al.*, 1988). The recent cloning of PEA3 was performed using a PEA3 motif nucleotide probe to screen a mouse expression library (Xin *et al.*, 1992), thus confirming that the nucleotide sequence specificity of the PEA3 motif is high for this family of transcription factors. Most of the Ets family members containing an ETS domain bind to the PEA3 motif (MacLeod *et al.*, 1992).

The ETS domain of all Ets proteins is however variable from member to member, allowing classification of Ets members in different groups with respect to their relationships (Lautenberger *et al.*, 1992; Laudet *et al.*, 1993). For example, the ETS domain of Ets-1 and Ets-2, and Erg-1 and Fli-1 are 96% and 98% identical, respectively (MacLeod *et al.*, 1992), thus leading to the establishment of Ets and Erg groups, respectively. In contrast, the PEA3 ETS domain ranges from 66% AA identity when compared to that of Ets-2, to 36% of SPI-1 (Laudet *et al.*, 1993), thus establishing that this gene is most probably not closely phylogenetically related to the other *ets*-family members. The recent finding of ER81 which is 95% AA identical within the ETS domain and almost 50% identical with the full-length protein when compared to PEA3 (Brown & McKnight, 1992) permits the establishment of a new Ets group referred to as the PEA3 group.

Here, we report the isolation of a human cDNA clone that codes for a new PEA3 group member, named *Ets-Related Molecule PEA3-like* (ERM). After characterizing its DNA-binding capacity, we show that this gene is expressed in almost all mouse and human tissues examined to date.

Results

Isolation of ERM cDNA

Using the ³²P-labeled mouse PEA3 PCR fragment corresponding to nucleotides 1289 to 1866 and containing the ETS domain, five cDNAs were isolated from 10⁶ human fetal liver λgt11 recombinants. They were purified and size-characterized. The three clones longer than 1.5 kb were subcloned into pCRII vector, amplified and their 5'- and 3'-regions were sequenced using primers in the vector. Only one clone contained about 250 nucleotides of the PEA3 ETS domain and a long 3'-noncoding region. Using part of this clone as ³²P-labeled human homologous probe, we screened the testis library and isolated three positive clones containing inserts larger than 1.5 kb. One clone of about 2.2 kbp was entirely sequenced, and proved to be different from PEA3 (Figure 1a). The sequence of this cDNA presented an open reading frame of 1530 nucleotides (Figure 1b). The first in-frame ATG codon was at position 196 and was preceded by an in-frame terminator codon TGA 135 nucleotides upstream (Figure 1b). Although the 3'-untranslated region of this cDNA is about 400 bp long, the data obtained on the expression of this cDNA suggest that this clone is truncated in its 3'-untranslated region and that the corresponding full-length cDNA is about 4 kb. This may explain why no polyadenylation consensus AATAAA site (Wickens, 1990) was detected downstream from the stop codon. This new cDNA was named *Ets-Related Molecule PEA3-like* (ERM).

Deduced amino acid sequences and homology with PEA3-related members

As shown in Figure 1b, the sequence of the first in-frame initiating codon, namely, CACCATGG, corresponds to the consensus sequence required for optimal initiation by eukaryotic ribosomes (Kozak, 1991). Fifteen in-frame ATG codons are found in the sequence. The second in-frame ATG 11 AA downstream from the first one, although possessing a G at position +4, lacks a purine at position -3 which is essential for efficient translation (Kozak, 1991). The open reading frame starting at the first ATG could encode 510 contiguous AA. The calculated molecular mass of the product of this ORF is 57 800 Da. Translation of the ORF predicts a protein rich in proline (13.9%), glutamine (7.5%) and glutamic acid (6.9%) residues. These AA do not appear to be clustered in any particular region of the primary structure of the protein. Computer analysis of the deduced AA sequence shows that human ERM protein contains an ETS domain between AA 364 and 449. Sequence comparison reveals that human ERM ETS domain is 63% identical to that of Ets-1 (data not shown) and 95% identical to that of human (Higashino *et al.*, 1993) and mouse PEA3 (Xin *et al.*, 1992), and 96% to that of mouse ER81 (Brown & McKnight, 1992). This latter finding led to the establishment of a new Ets group, named the 'PEA3-related proteins' currently comprising ERM, ER81 and PEA3. Sequence alignment of the complete ERM, ER81 and PEA3 open reading frames indicates that PEA3 contains 79 additional putative AA upstream from the methionine common to the three members.

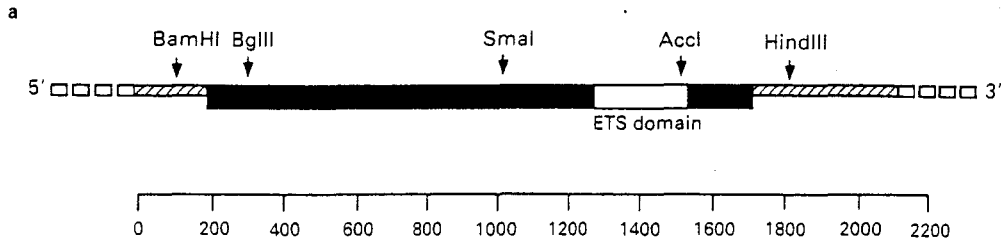
From this methionine to the stop codon, ERM, ER81 and PEA3 contain 510, 477 and 476 AA, respectively. Throughout this latter region, ERM is 52% identical to PEA3 and 58% to ER81. In the amino terminal region, the three members possess a well-conserved acidic region of 32 AA (Figure 2), which could correspond to a typical activation domain of transcription factors (Johnson & McKnight, 1989). In contrast, ERM possesses a specific region between AA 146 and 179, the ERM domain.

DNA-binding of ERM ETS domain

Full-length ERM was introduced into the pSG5 plasmid and was transcribed and translated *in vitro* with the rabbit reticulocyte lysate system using radiolabeled methionine. The size of the produced protein is detected at ≈ 70 kDa (Figure 3b, lane 1) instead of at the calculated molecular mass of ≈ 58 kDa, indicating that ERM could be submitted to post-translational modifications. However, Xin *et al.* have reported the same observation for the migration of mouse PEA3 product, explaining this phenomenon by the presence of many stretches of prolines (Xin *et al.*, 1992). In order to check the DNA-binding ability of ERM ETS domain *sensu stricto*, we constructed by PCR a fragment encoding the 83 AA peptide from AA 366 to 447 of ERM that contained an ATG and a stop codon at its 5'- and 3'-end, respectively (Figure 3a). After inserting this fragment into pSG5 vector, *in vitro* transcription and translation of the fragment with the reticulocyte lysate system using radiolabeled methionine revealed a product of ≈ 11 kDa (Figure 3b, lane 2). Neither of the two labeled proteins could be detected when pSG5 plasmid alone was submitted to the same procedure (data not shown). The different labeling intensity obtained for the two products is due not only to the presence of 15 methionines in the full-length ERM protein and only three in its ETS domain, but also to a higher transcription-translation efficiency of pSG5-ERM plasmid. As shown in Figure 3c (lane 2), non-radiolabeled *in vitro* translated ERM protein bound to a ³²P-labeled double-stranded oligonucleotide recognized by Ets family members. This binding is specific since displaced by a 100-fold nonradiolabeled oligonucleotide excess (Figure 3c, lane 3), whereas 100-fold excess of a non-specific oligonucleotide containing no Ets binding site had no effect (Figure 3c, lane 4). As control of ERM binding specificity, lane 1 of Figure 3c shows that reticulocyte lysate alone did not bind to the probe. As demonstrated in Figure 3c (lane 5) a truncated protein of 83 AA which encompassed the ERM ETS domain efficiently bound to the probe. A 100-fold excess of non-radiolabeled oligonucleotide competed the binding of the probe (Figure 3c, lane 6), whereas the non-specific oligonucleotide had no effect (Figure 3c, lane 7).

Northern blot analysis of ERM in human tissues and cell lines

As shown in Figure 4, analysis of human poly(A⁺) RNA samples (10 μg/lane) using the 650 bp probe containing the ERM domain revealed the presence of 4 kb mRNA transcripts. Human ERM mRNA is very



b

```

ggggcgagaagagacgcttggcggacgtccgcagttgggaggagggggcggggaagcagtttgaggggaatgt 72
ctgagagggcgtgccccagagcaccgggtagggggtaaatgacacaggaggatcccttttccccagaaa 144
ttactcaatgctgaaacctctcaaagtggattagagacgctgaaagcacc ATG GAC GGG TTT TAT 210
                                     M D G F Y
GAT CAG CAA GTC CCT TTT ATG GTC CCA GGG AAA TCT CGA TCT GAG GAA TGC AGA 264
D Q Q V P F M V P G K S R S E E C R
GGG CGG CCT GTG ATT GAC AGA AAG AGG AAG TTT TTG GAC ACA GAT CTG GCT CAC 318
G R P V I D R K R K F L D T D L A H
GAT TCT GAA GAG CTA TTT CAG GAT CTC AGT CAA CTT CAA GAG GCT TGG TTA GCT 372
D S E E L F Q D L S Q L Q E A W L A
GAA GCA CAA GTT CCT GAT GAT GAA CAG TTT GTC CCA GAT TTT CAG TCT GAT AAC 426
E A Q V P D D E Q F V P D F Q S D N
CTG GTG CTT CAT GCC CCA CCT CCA ACC AAG ATC AAA CGG GAG CTG CAC AGC CCC 480
L V L H A P P P T K I K R E L H S P
TCC TCT GAG CTG TCG TCT TGT AGC CAT GAG CAG GCT CTT GGT GCT AAC TAT GGA 534
S S E L S S C S H E Q A L G A N Y G
GAA AAG TGC CTC TAC AAC TAT TGT GCC TAT GAT AGG AAG CCT CCC TCT GGG TTC 588
E K C L Y N Y C A Y D R K P P S F
AAG CCA TTA ACC CCT CCA ACC CCC CTC TCA CCC ACC CAT CAG AAT CCC CTA 642
K P L T P P T T P L S P T H Q N P L
TTT CCC CCA CCT CAG GCA ACT CTG CCC ACC TCA GGG CAT GCC CCT GCA GCT GGC 696
F P P P Q A T L P T S G H A P A A G
CCA GTT CAA GGT GTG GGC CCC GCC CCC GCC CAT TCG CTT CCA GAG CCT GGA 750
P V Q G V G P A P A P H S L P E P G
CCA CAG CAG CAA ACA TTT GCG GTC CCC CGA CCA CCA CAT CAG CCC CTG CAG ATG 804
P Q Q Q T F A V P R P P H Q P L Q M
CCA AAG ATG ATG CCT GAA AAC CAG TAT CCA TCA GAA CAG AGA TTT CAG AGA CAA 858
P K M M P E N Q Y P S E Q R F Q R Q
CTG TCT GAA CCC TGC CAC CCC TTC CCT CAG CCA GGA GTT CCT GGA GAT AAT 912
L S E P C H P F P P Q P G V P G D N
CGC CCC AGT TAC CAT CGG CAA ATG TCA GAA CCT ATT GTC CCT GCA GCT CCC CCG 966
R P S Y H R Q M S E P I V P A A P P
CCC CCT CAG GGA TTC AAA CAA GAA TAC CAT GAC CCA CTC TAT GAA CAT GGC TTC 1020
P P Q G F K Q E Y H D P L Y E H G V
CCG GGC ATG CCA GGG CCC CCA GCA CAC GGG TTC CAG TCA CCA ATG GGA ATC AAG 1074
P G M P G P P A H G F Q S P M G I K
CAG GAG CCT CGG GAT TAC TGC GTC GAT TCA GAA GTG CCT AAC TGC CAG TCA TCC 1128
Q E P R D Y C V D S E V P N C Q S S
TAC ATG AGA GGG GGT TAT TTC TCC AGC AGC CAT GAA GGT TTT TCA TAT GAA AAA 1182
Y M R G G Y F S S S H E G F S Y E K
GAT CCC CGA TTA TAC TTT GAC GAC ACT TGT GTT GTG CCT GAG AGA CTG GAA GGC 1236
D P R L Y F D D T C V V P E R L E G
AAA GTC AAA CAG .GAG CCT ACC ATG TAT CGA GAG GGG CCC CCT TAC CAG AGG CGA 1290
K V K O E P T M Y R E G P P Y Q R R
GGT TCC CTT CAG CTG TGG CAG TTC CTG ACC CTT CTT GAT GAC CCA GCC AAT 1344
G S L Q L W Q F L V T L L D D P A N
GCC CAC TTC ATT GCC TGG ACA GGT CGA GGC ATG GAG TTC AAG CTG ATA GAA CCP 1398
A H F I A W T G R G M E F K L I E P
GAA GAG GTT GCT CGG CGC TGG GGC ATC CAG AAG AAC CGG CCA GCC ATG AAC TAT 1452
E E V A R R W G I Q K N R P A M N Y
GAC AAG CTG AGC CGC TCT CTC GCG TAT TAC TAT GAA AAG GGC ATC ATG CAG AAG 1506
D K L S R S L R Y Y Y E K G I M O K
GTG GCT GGA GAG CGA TAC GTC TAC AAA TTT GTC TGT GAC CCA GAT GCC CTC TTC 1560
V A G E R Y V Y K F V C D P D A L F
TCC ATG GCT TTC CCG GAT AAC CAG CGT CCG TTC CTG AAG GCA GAG TCC GAG TGC 1614
S M A F P D N Q R P F L K A E S E C
CAC CTC AGC GAG GAG ACC CTG CCG CTG ACC CAC TTT GAA GAC AGC CCC GCT 1668
H L S E E D T L P L T H F E D S P A
TAC CTC CTG GAC ATG GAC CGC TGC AGC AGC CTC CCC TAT GCC GAA GGC TTT GCT 1722
Y L L D M D R C S S L P Y A E G F A
TAC TAA gtttctgagtggcggatggcacaacctagagctagcagttccattcaggcaaacaggggc 1791
Y *
agtggttttgtttgtgtttttgtgttcttaagcttgccctttgagtattatctggagaacccaagctgt 511
ctctggattggcacccttaagacagatacattggctggggagtggaacaggggaggggcagaaaaccacca 1863
aaaggccagtgccctcaactcttgatctctgatgaggttctctggaagagatcaaatggagtctccttaccat 2007
ggacaatacatgcaagcaatctctgttcaggttagtaccgcgcaaacgggacatgatgtgacaatctgca 2079
tcgatcatggactactaatgcctttacatagaaaaaaaaaaaaaaaaaaaaa 2130
    
```

Figure 1 (a) Restriction map of a human ERM cDNA clone. The protein-coding region is represented by a **■** which contains the ETS domain, while the flanking 5'- and 3'-noncoding regions are shown as **▨**; the cloning vector pCRII being represented by **□**. A scale is shown below, where 1 corresponds to the first 5'-nucleotide of the ERM cDNA presented. (b) Nucleotide and predicted AA sequences of human ERM cDNA. Nucleotide and AA positions are indicated to the right of the sequence. The boxed region corresponds to the ETS domain

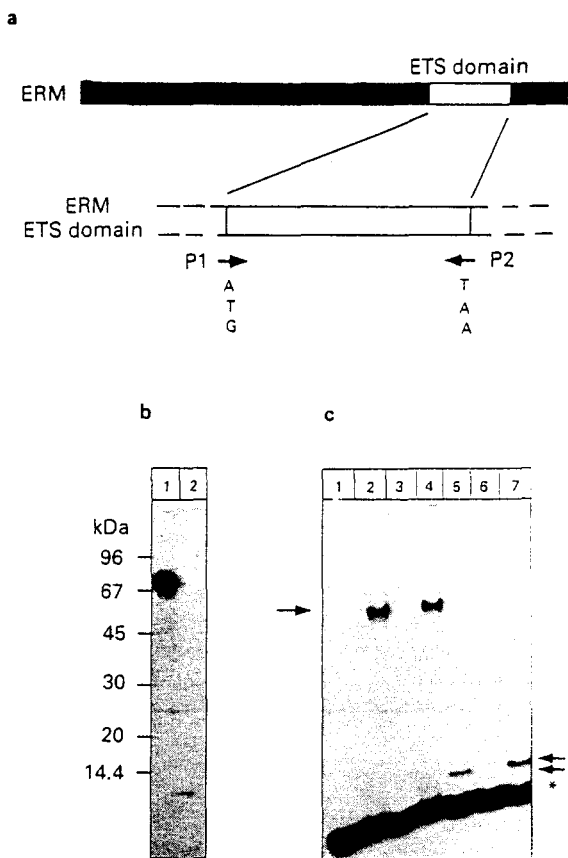


Figure 3 DNA-binding activity of overexpressed ERM and ERM ETS domain. (a) Map of the sequence amplified by PCR for ERM ETS domain overexpression. Using two primers (P1 and P2) described in Materials and methods section and containing an ATG and a stop codon, respectively, the PCR fragment was inserted downstream from the T7 promoter of pSG5 vector (pSG5-ED), in order to translate *in vitro* in rabbit reticulocyte lysates the ETS domain of 83 AA. (b) SDS-PAGE of ERM and ERM ETS domain. Synthesis of the ERM (lane 1) and ERM ETS domain of 83 AA (lane 2) was performed by submitting pSG5-ERM and pSG5.ED plasmids to *in vitro* transcription and translation using rabbit reticulocyte lysates containing radiolabeled methionine. The products were separated on a SDS gradient 10–30% polyacrylamide gel. The sizes (kDa) and positions of molecular weight standards are indicated on the left. (c) Gel retardation assay of ERM and ERM ETS domain synthesized in rabbit reticulocyte lysates. 32 P-labeled double stranded oligonucleotide containing a PEA3 motif described in Material and methods was used as specific probe, and the ERM protein-DNA (one arrow) or ERM ETS domain-DNA (two arrows) complexes were separated from the probe (*) on a non-denaturing 5–15% gradient polyacrylamide gel. The overexpressed ERM protein (lanes 2–4) and ERM ETS domain protein (lanes 5–7) were incubated by competing with a 100-fold molar excess of unlabeled specific (lanes 3 and 6) and non specific (lanes 4, 7) oligonucleotide. Lane 1 represents gel shift assay using rabbit reticulocyte lysate alone incubated with the probe.

contrast, the estrogen receptor negative cell line BT-20 moderately expressed ERM. The prostate PC3 and DU-145 cancer cell lines were moderately positive, also confirming ERM expression in two of the three prostate tumors tested. Table 1b shows the relative ERM expression in hematopoietic cell lines. All the infected myelocytic and leukemia or lymphoma tumor cell lines were weakly or moderately positive for ERM expression except the lymphoblastic leukemia MOLT-4 cell line which expressed high amount of ERM. In

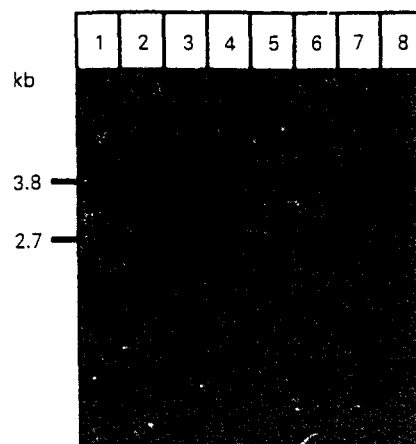


Figure 4 Northern blot analysis of ERM in human non-cancerous tissues. mRNA blot contained 10 μ g poly(A⁺) mRNA from human heart (lane 1), brain (lane 2), placenta (lane 3), lung (lane 4), liver (lane 5), skeletal muscle (lane 6), kidney (lane 7) and pancreas (lane 8). The blot was hybridized with a 650 bp probe from nucleotide 226 to 875 of ERM cDNA. RNA blot analysis was also performed as previously described (Zhao *et al.*, 1991). The sizes (kb) and positions of 3.8 and 2.7 kb of human ets-2 mRNAs (standards used) are indicated on the left

contrast, normal and peripheral blood leukocytes or activated T4 lymphocytes did not express ERM mRNA.

RNase protection analysis of ERM in human cell lines

As illustrated in Figure 6, the highly specific RNase protection technique confirmed the data obtained by Northern blot on human cell lines. Using a 456 nucleotide probe containing the acidic domain, a single band of 348 nucleotides corresponding to the ERM cDNA portion homologous to the probe is detected at a high level in SK-MEL, DU-145 and MOLT-4 cells, at moderate level in BT-20, ACHN and Dami cells, and at low level in HeLa and HCT-15 cells (Figure 6). Moreover, no band is detected in T-47D, MCF-7 and JAR cells (Figure 6). We also confirmed that unactivated and phytohemagglutinin-activated peripheral blood leukocytes do not express ERM (Figure 6). Using a 746 nucleotide probe containing the acidic domain as well as the ERM domain, a single band of 650 nucleotides corresponding to the ERM cDNA homologous to the probe is detected in DU-145 and MOLT-4 cells (Figure 7), thus confirming that the ERM domain from nucleotide 634 to nucleotide 736 is transcribed.

Northern blot analysis of ERM in normal mouse tissues

As shown in Figure 8, we also characterized the expression pattern of PEA3, ER81 and ERM in mouse tissues. Due to substantial sequence identities between these three cDNAs, and thus the risk of cross-hybridization, we used PCR to produce specific probes outside the ETS domain in the 5' region of mER81 and mPEA3 open reading frames. We also used the 650 nucleotide ERM probe containing the ERM domain described here above. Using the human ERM probe with stringent hybridization conditions, a 3.8 to 4.2 kb transcript was expressed in almost all mouse

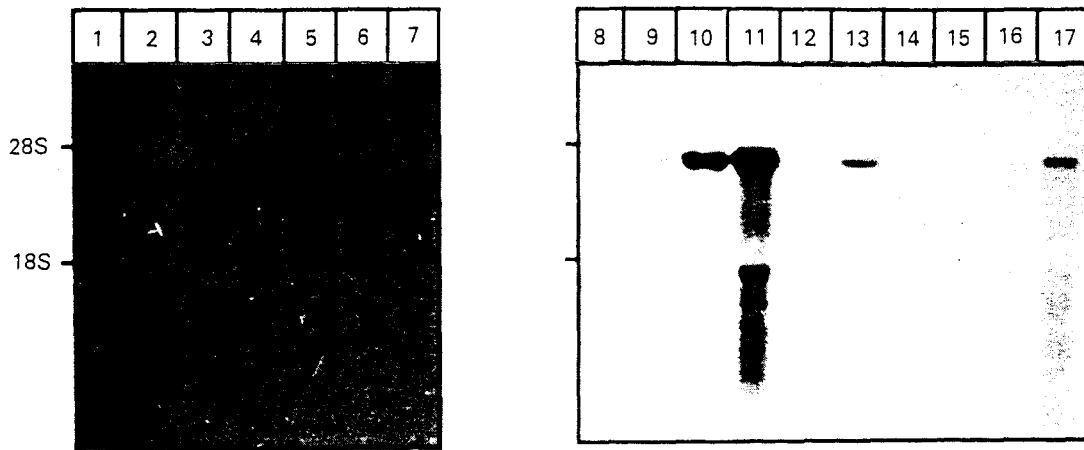


Figure 5 Northern blot analysis of ERM in human tumor tissues. Ten μ g total RNA from normal tissues such as placenta (lanes 1, 17), ovary (lane 8) and breast fibroadenoma (lane 13), and from tumoral tissues such as prostate adenocarcinoma (lanes 2, 16), meningioma (lanes 3–7), ovarian carcinoma (lanes 9, 10), melanoma metastasis (lane 11), vulva carcinoma (lane 12), breast adenocarcinoma (lane 14) and penis carcinoma (lane 15) were extracted as previously described (Monté *et al.*, 1992) and applied onto 1% denaturing agarose gel. After RNA transfer on nylon membranes, blots were hybridized, washed and autoradiographed as described in legend to Figure 4. The sizes and positions of 28S and 18S RNAs are indicated on the left

tissues tested (Figure 8a). In brain, ERM is highly expressed (Figure 8a, lane 6), and in spleen (Figure 8a, lane 2), testis (Figure 8a, lane 4) and kidney (Figure 8a, lane 5), it is moderately expressed. In contrast, PEA3 is only detected in brain with a weakly expressed transcript at 2.4 kb (Figure 8b, lane 6). Using the ER81 probe, we showed that two mRNA species of 6–7 and 3–3.4 kb are detected at high levels in kidney (Figure 8c, lane 5) and brain (Figure 8c, lane 6), whereas a third species was exclusively detected at \approx 2.4 kb in testis (Figure 8c, lane 4).

Discussion

The present report characterizes a human cDNA encoding a DNA-binding protein of the Ets family, the Ets-related molecule PEA3-like, ERM. Using sequence alignment methodology, we were able to define a group composed of three closely related members which were recently cloned, PEA3, ER81 and ERM. They share not only AA sequence identity in the ETS domain but also in the acidic domain observed in PEA3 (Xin *et al.*, 1992), as well as in the regions surrounding the ETS domain. Within this PEA3 group, the three proteins share about 50% identity all over their AA sequence. Nevertheless, the fact that the amino- and carboxyl-terminal regions of PEA3 members are highly similar indicate that the selection pressure was not only focused on the ETS domain located in the carboxyl terminus of the proteins. It is probably the case for other specific domains putatively responsible for transactivation, such as for Ets-1 and Ets-2 (Watson *et al.*, 1988). For example, a specific PEA3 region of 79 AA is present in its amino terminal region, and an ERM specific domain of about 35 AA is present in ERM. We found that this latter domain is not a cloning artifact, since using RNase protection technology we demonstrated its presence in the mRNA of human cell lines (Figure 7), and we also showed that it is present in the genomic DNA (data not shown).

Signature of the Ets family is a conserved DNA-binding domain of \approx 85 AA (Karim *et al.*, 1990). We

Table 1 Northern blot analysis of ERM in human cell lines

Tissue origin	Cell line	Expression
a. Solid Tumors		
Melanoma	SK-MEL	+++++
Renal adenocarcinoma	ACHN	+++
Colon adenocarcinoma	HCT-15	++
	LOVO	++
Mammary adenocarcinoma	ZR-75-1	–
	T-47D	–
	MCF-7	+/-
	BT-20	+++
Placenta choriocarcinoma	JAR	–
	JEG	–
	BEWO	+
Prostate adenocarcinoma	LNCaP	–
	DU-145	+++
	PC3	+++
Cervix carcinoma	HeLa	+
Bladder carcinoma	J82	++
	Scaber	++
	HT-1376	++
Hepatocarcinoma	HEPG2	++
Lung carcinoma	A427	+++
b. Hematopoietic Cells		
Peripheral blood leukocytes		–
Activated T4 lymphocytes		–
Infected lymphocytes		
B infected by EBV	BEBV	++
T infected by HTLV-1	MT4	+++
Tumor cells		
Lymphoblastic leukemia	CEM	+++
	MOLT-4	+++++
	Jurkat	++
Erythroleukemia	HEL	++
Promonocytic leukemia	U937	+++
Megakaryocytic leukemia	Dami	+++

Cells were cultured and collected at subconfluent stage. ERM expression was quantified by Northern blot analysis. The highest ERM expression was observed in SK-MEL melanoma cells (+++++). ERM expression in each cell line was appreciated relatively to that in SK-MEL. Arbitrary quantification of expression was as follows: none: –, very weak: +, weak: ++, moderate: +++, high: +++++ and very high: ++++++

have observed that the full-length ERM protein transcribed and translated *in vitro* is able to bind to specific DNA sequences, clearly confirming that this new Ets member is also a DNA-binding protein as previously

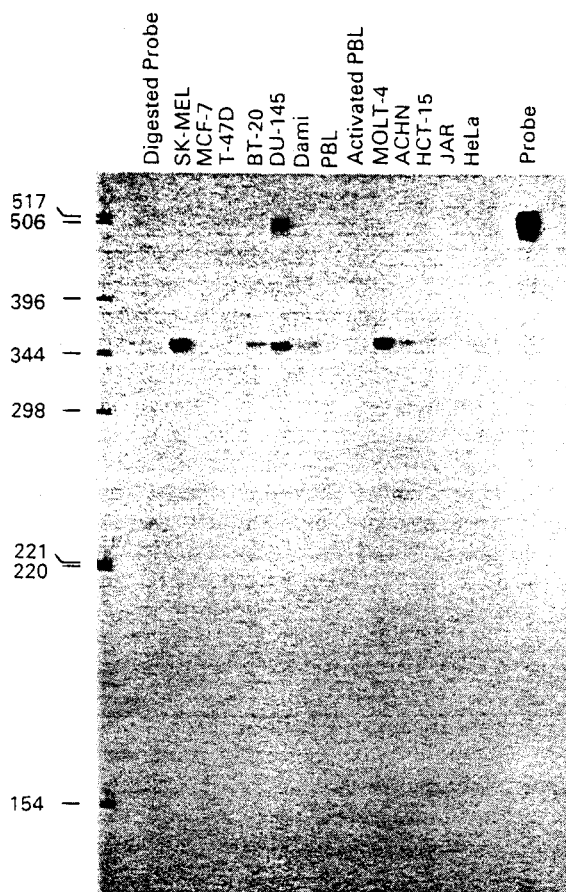


Figure 6 RNase protection analysis of ERM in human cell lines. Using a 456 nucleotide cRNA probe containing 108 nucleotides from the PCR II vector and 348 nucleotides corresponding to nucleotide 328 to nucleotide 675 of ERM cDNA, samples of 20 μ g RNA from the indicated cell lines were hybridized for 18 h at 60°C and digested with ribonuclease. The protected fragments were resolved on a 6% denaturing polyacrylamide-urea sequencing gel and then exposed to X-ray films for 4 days with intensifying screens

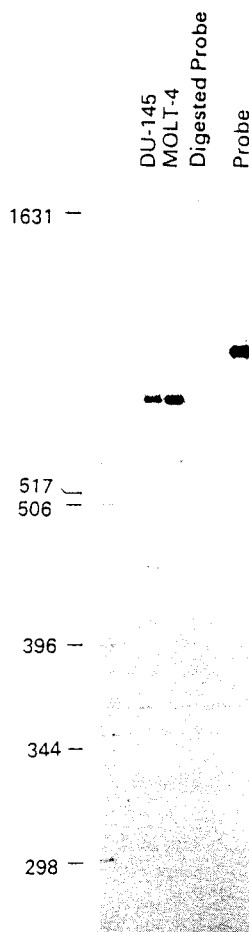


Figure 7 RNase protection for the presence of the ERM domain. Using a specific cRNA probe of 746 nucleotides, including 96 nucleotides from the PCR II vector and 650 nucleotides corresponding to nucleotide 226 to nucleotide 875 of ERM cDNA, samples of 10 μ g RNA from DU-145 and MOLT-4 cell lines were hybridized for 18 h at 70°C and treated as described in legend to Figure 6

demonstrated for PEA3 (Xin *et al.*, 1992) and ER81 (Brown & McKnight, 1992). The minimal ETS domain sufficient for DNA-binding has not been clearly established until now. Recent works have shown that a truncated protein containing the ETS domain supplemented by the carboxyl terminus region of Ets-1 (Gégonne *et al.*, 1992; Hagman & Grosschedl, 1992; Nye *et al.*, 1992) binds to DNA. The importance of Ets-1 carboxyl terminus domain has been established in studies comparing v-Ets and c-Ets-1 which are almost homologous in the full-length protein except in this region (Leprince *et al.*, 1992). However, to our knowledge, no studies have reported that the ETS domain of ≈ 85 AA alone is able to bind to DNA. We reported here that a 83 AA portion of the ERM ETS domain translated *in vitro* using rabbit reticulocyte lysates binds to the DNA in a specific manner. Recently, Langer *et al.* (1992) have shown that a 132 AA portion of Ets-2 containing the ETS domain inhibits the transactivation of intrinsic Ets proteins, thus suggesting that, also *in vivo*, the ETS domain alone binds to DNA, even when devoid of all the transactivator domains.

Concerning expression of this class of transcription factors, some display tissue-specificity and others are widely distributed. For example, *ets-2* (Bhat *et al.*, 1989; 1990) and GABP α (LaMarco *et al.*, 1991) seem to be ubiquitous, whereas ER71 (Brown & McKnight, 1992) and *elf-1* (Thompson *et al.*, 1992) seem to be expressed only in epididymis and T lymphocytes, respectively. Since the biopsies, or the surgical specimens used in the present study have been removed from cancerous patients, even if they were anatomopathologically diagnosed as normal tissues, tissue distribution of ERM should be taken in account with great care. A larger study on an homogenous series of human tumors is presently being conducted in order to potentially correlate ERM expression and tumor aggressiveness. Nevertheless, we show here that this gene is highly expressed in normal and tumoral brain tissue as it is the case for *elk-1* (Rao *et al.*, 1989), PEA3 (Xin *et al.*, 1992), ER81 (Brown & McKnight, 1992) and GABP α (LaMarco *et al.*, 1991). Our data also indicate that ERM could be expressed in normal or benign mammary tissues and not in the corresponding cancerous ones. We also show here that only a few

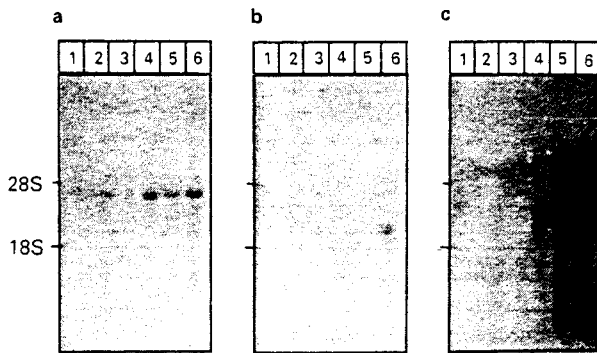


Figure 8 Northern blot analysis of ERM in adult male mouse tissues. Ten μ g total RNA from liver (lane 1), spleen (lane 2), intestine (lane 3), testis (lane 4), kidney (lane 5) and brain (lane 6) were separated and transferred as described. Three different membranes containing the same mouse tissue mRNA were hybridized with the ERM probe described in legend to Figure 4 (a), as well as with a 614 bp PEA3 probe from nucleotide 136–749 of mouse PEA3 cDNA (b) (Xin *et al.*, 1992), and a 318 bp ER81 probe from nucleotide 496–813 of mouse ER81 cDNA (c) (Brown & McKnight, 1992). The membranes were washed as described in legend to Figure 4 except for the blot hybridized with the human probe where the high stringency wash was $0.5 \times \text{SSC} - 0.1\%$ SDS at 55°C for 30 min

tissues tested were negative by Northern blot analysis on total RNA. These data thus tend to classify this gene as a ubiquitously expressed one.

Expression of ERM has also been detected in the mouse, clearly indicating the existence of this gene in this class of mammals. Its tissue distribution confirms our observations on human tissues, since ERM seems to be widely distributed in the mouse as well. It is thus of interest to note that the brain expresses ER81, ERM and PEA3. In contrast to the PEA3 expression observed by Xin *et al.* (1992), we were not able to detect PEA3 in mouse epididymis (data not shown). Let us emphasize that PEA3 expression previously observed (Xin *et al.*, 1992) was obtained using the very sensitive RNase protection technique on poly(A⁺) RNA, whereas the present data were obtained using Northern analysis. However, we did detect specific expression in certain seminiferous tubules of the mouse adult testis using *in situ* hybridization with a mouse PEA3 specific probe (Defossez, unpublished data). The discrepancy between our data and those of Xin *et al.* (1992) could be explained either by the mouse strain used, or more probably by the fact that ERM is associated with a specific stage in spermatogenesis. We could thus conclude that the three genes are also expressed in the testis. Expression in the same tissues of three genes from the same Ets group is of great interest for studying their putative expression relay. As a matter of fact, ER81 is highly expressed in the testis, ERM moderately and PEA3 weakly.

Due to the fact that the antibodies we have tested do not recognize PEA3-related proteins, we are not presently able to study the relationships between ERM, ER81 and PEA3 mRNA expression levels and their respective translation levels. However, when we carefully analyse nucleotide sequences encompassing the first ATG of PEA3 as well as the second in-frame one 79 AA downstream, neither of the two appears to be highly efficient for optimal translation by Kozak

rules (Kozak, 1991). In contrast, ER81 and ERM first ATGs are encompassed with nucleotides for optimal translation (Kozak, 1991).

It is difficult to imagine that these genes display sequence specificity on targeted promoters, since recent data using the 'Target Detection Assay' technique reveal that no clear specific sequences outside the GGAA/T core sequence are required for binding of most of the Ets family members (Nye *et al.*, 1992; Thompson *et al.*, 1992; Woods *et al.*, 1992). Since ER81 and Ets-1 which share only 63% identity in the ETS domain recognize the same sequence, it is thus more than likely that ERM, ER81 and PEA3 which share 95% identity in the ETS domain and $\approx 50\%$ within the full-length protein also bind to the same DNA sequence. We thus suggest that it is not the nucleotide target which is responsible for optimal binding of Ets proteins, but the specificity of the transcriptional cofactors which interact with Ets proteins. As demonstrated for GABP α with GABP β (LaMarco *et al.*, 1991), Elk-1 with SRF (Hipskind *et al.*, 1991) or Ets-1 with AP-1 (Wasyluk *et al.*, 1990), requirement of cofactor(s) with Ets proteins for optimal transcription could also be observed for the PEA3-related proteins. The putative domains of ERM for specific protein-protein interactions are currently under investigation. However, the first 79 AA of PEA3 not present in the two other members, could be of interest, as well as the ≈ 35 AA of ERM domain. These data and the finding of three genes almost identical in their ETS domain, but also possessing specific regions potentially involved in protein-protein interactions, open a new field of investigation on the fine tuning of transcription regulation by Ets proteins.

Materials and methods

Isolation, subcloning and sequencing of cDNA clones

In order to clone a human PEA3 cDNA, we generated a 578 bp cDNA fragment in the 3'-end of the open reading frame of mouse PEA3 cDNA containing the ETS domain (Xin *et al.*, 1992). This fragment was produced by reverse transcription PCR on mouse brain total RNA using the following primers 5'-TGAAGGAGACATCAAGCAGGAAG-3' and 5'-CTTCCAACCACAACACTGTCTGCTGCG-3' (Eurogentec, Belgium). Approximately 10^6 recombinant phages from an oligo(dT)-primed human fetal liver λ gt11 cDNA expression library were screened with a 1×10^6 c.p.m. per ml ^{32}P -labeled probe. Then, we screened a human adult testis cDNA library (Clontech, Palo Alto, CA) using a cDNA fragment cloned from the fetal liver library that contained high sequence identity in 250 bp of mouse PEA3 ETS domain. The filters were washed at $0.5 \text{SSC} - 0.1\%$ SDS at 60°C for 30 min and then autoradiographed 12 h at -80°C . After two purification steps, phages were isolated as described (Zhao *et al.*, 1991) and the inserts were excised by EcoRI digestion. The purified cDNA inserts were subcloned into the PCR II (Invitrogen, San Diego, CA) polylinker site.

Synthetic oligonucleotides, as well as T7 or SP6 vector primers, and modified T7 DNA polymerase were used to sequence both strands of double-stranded plasmid DNA with the dideoxy chain termination method with [^{35}S]ATP (USB, Cleveland, OH). Sequences were confirmed using an Applied Biosystem 370A automatic system with fluorescent dye-labeled cDNA sequence-specific primers and a Taq dye-primer sequencing kit (Applied Biosystem, Foster City, CA).

In vitro transcription and translation of ERM and ERM ETS domain in rabbit reticulocyte lysates

After digestion with EcoRI endonuclease, the full-length ERM fragment was subcloned into the unique EcoRI restriction site of pSG5 expression vector (Pharmacia, Piscataway, NJ). In order to obtain optimal translation, this vector was digested with BamHI endonuclease using the BamHI site of ERM cDNA at position 124 bp in the 5'-noncoding region of ERM, and the unique BamHI site of pSG5. This fragment was thereafter subcloned into the BamHI site of pSG5 downstream from the T7 promoter, to produce the pSG5-ERM recombinant plasmid. Using PCR, we also generated a fragment encoding 83 AA of ERM ETS domain using the following primers, 5'-GGGAGATCTATGGGTTCCCTT-CAGCTGTGGCAG-3' and 5'-AGACCCGGGTTTAAAA-TTTGTAGACGTATCGCTC-3'. The PCR product was subcloned into the PCRII vector, and, after EcoRI digestion, was inserted into the unique EcoRI site of pSG5 vector, to produce the pSG5-ED plasmid. One μg of each plasmid was then used for transcription and translation of ERM ETS domain with 10 units of T7 RNA polymerase, 25 μl rabbit reticulocyte lysates and 40 μCi [^{35}S]methionine in the appropriate buffer as specified by the manufacturer (Promega, Madison, WI) for 1.5 h at 30°C. In order to avoid interference of ^{35}S and ^{32}P radioactivity, non-labeled methionine was used for translation of truncated proteins used in gel shift assays. Translation of the full-length ERM protein and of the 83 AA truncated protein was verified by SDS-PAGE in the presence of 2 β -mercaptoethanol.

Electrophoretic mobility-shift assays

Binding reactions were performed in a final volume of 20 μl containing 2 ng of radiolabeled double-stranded DNA (4×10^4 c.p.m.), 1 μg poly(d(I-C)) (Pharmacia, Piscataway, NJ), $\approx 5 \mu\text{g}$ total protein from rabbit reticulocyte lysates in 20 mM HEPES pH 7.5, 4% Ficoll, 5 mM MgCl_2 , 40 mM NaCl, 0.2 mM EDTA, and 1 mM DTT. The reactions were carried out for 20 min at 4°C. After prerunning a 5–15% gradient polyacrylamide gel (30:1 ratio acrylamide/bis-acrylamide) in 0.25 \times TBE for 90 min at 10 mA, the samples were separated at 4°C for 3 h. A primer containing sequences of PyE 5'-GATCTTCGAAAGGAAGTTCGAG-3' was rendered double-stranded by elongation with Klenow and labeled with [γ - ^{32}P]ATP and polynucleotide kinase. When employed, non-specific double-stranded primer (5'-TTTACCGTGTGGTC-AGTTCTAAAAAT-3' was added at 100-fold excess.

RNA isolation and analysis

A commercial poly(A⁺) RNA blot was obtained from Clontech (Palo Alto, CA). Human tissues were also provided by Dr Michel Coibion (Dept. of Surgery, Institut J. Bordet, Brussels, Belgium) and Dr Pierre-Marie Martin (Lab. of Experimental Oncology, Faculty of Medicine, Marseille, France) from patients operated for tumor resection. Tissues were immediately frozen in liquid nitrogen and stored at -80°C. RNA extraction was carried out by homogenizing frozen tissues in 2–4 ml RNazol (Cinna-Biotex, Friendwood, TX) with an Ultraturax polytron. After addition of equal volumes chloroform, extracted RNA was precipitated with one volume isopropanol. RNA from cell cultures was obtained using the same procedure, except that prior to addition of RNazol, cells were washed twice with phosphate-buffered saline (pH 7.5) and no homogenization with the polytron procedure was necessary (Monté *et al.*, 1992). Ten μg of total RNA were fractionated by electrophoresis through a 1% denaturing agarose gel (Monté *et al.*, 1992), transferred to a Hybond N membrane (Amersham, Little Chalfont, UK), and UV-crosslinked as previously described (Monté *et al.*, 1992). For each blot, equal mRNA loading was checked

using ethidium bromide staining followed by UV visualization (not shown). The membranes were prehybridized in a solution containing 250 $\mu\text{g ml}^{-1}$ of denatured salmon sperm DNA, 50% formamide, 5 \times Denhardt's, 0.1% SDS, 5 \times SSC, 50 mM Na_2HPO_4 , for 2 h at 42°C, and hybridization was carried out in the same solution containing 5% dextran sulfate and one of the ^{32}P -labeled probes for 16 h at 42°C. The first probe of 2.0 kb comprises the ERM ETS domain and the 3'-noncoding region of the cDNA, whereas, using the primers 5'-TTTATGGTCCCAGGGAAT-3' and 5'-TGGCAGGGTTCAGACAGTTGTCTC-3', a 650 bp probe from nucleotide 226 to 875 of ERM cDNA was amplified by PCR in the 5' portion of the ERM open reading frame outside the ETS domain. A 614 bp PEA3 probe from nucleotide 136 to 749 of mouse PEA3 cDNA (Xin *et al.*, 1992) and a 318 bp ER81 probe from nucleotide 496 to 813 of mouse ER81 cDNA (Brown & McKnight, 1992) were also generated by PCR. The membranes were washed sequentially in 2 \times SSC containing 0.1% SDS at 55°C for 1 h, in 0.5 \times SSC - 0.1% SDS at 60°C for 30 min and in 0.1 \times SSC - 0.1% SDS at 65°C for 30 min. The membrane containing mouse RNA hybridized with the human probe was washed at 0.5 \times SSC - 0.1% SDS at 55°C for 30 min. Thereafter, the membranes were exposed to X-ray films with an intensifying screen at -80°C. When necessary, the probes were stripped off the membranes by incubating the blots in 20 mM NaOH and 1% SDS for 30 min at room temperature, and different probes were subsequently rehybridized.

Ribonuclease protection assay

A fragment from nucleotide 328 to nucleotide 675 of ERM cDNA was amplified by PCR using the following primers: 5'-GAGCTATTTCCAGGATCTCAGT-3' and 5'-TGAGGTGGGCAGAGTTGCCTG-3'. This fragment was subcloned into the PCRII vector as above described. The recombinant plasmid was linearized with BamHI and labeled with [α - ^{32}P]UTP using T7 RNA polymerase to generate a specific cRNA probe of 456 nucleotides, including 108 nucleotides from the vector and 348 nucleotides from ERM. Twenty μg of total RNA from human cell lines were hybridized with 2 $\times 10^5$ c.p.m. probe for 18 h at 60°C. The PCRII recombinant plasmid containing the 650 bp PCR fragment from nucleotide 226 to 875 of ERM cDNA above described was linearized with EcoRV and labeled with [α - ^{32}P]UTP using T7 RNA polymerase to generate a specific cRNA probe of 746 nucleotides, including 96 nucleotides from the vector and 650 nucleotides from ERM. Ten μg of total RNA from DU-145 and MOLT-4 cell lines were hybridized with 2 $\times 10^5$ c.p.m. probe for 18 h at 60°C. Ribonuclease digestion and resolution of RNase-resistant RNA-RNA hybrids on a 6% denaturing polyacrylamide-urea sequencing gel were performed as described by the supplier (Clontech, Palo Alto, CA). Autoradiography was performed at -80°C with two intensifying screens.

Acknowledgements

We thank Isabelle Angeli, Brigitte Quatannens and Sophana Ung for skillfull technical assistance. We are grateful to Jean Coll, Vincent Laudet, Dominique Leprince and Bernard Vandebunder for critical reading of the manuscript. This work was done according to grants awarded in part by the 'Centre National de la Recherche Scientifique' (France), and by the 'Association pour la Recherche contre le Cancer' (France). Y.de L. is Senior Research Associate with the 'Fonds National de la Recherche Scientifique' of Belgium.

The sequence data described in this paper have been submitted to the EMBL/GenBank data libraries under accession number X76184.

References

- Ben David, Y., Giddens, E.B., Letwin, K. & Bernstein, A. (1991). *Genes Dev.*, **5**, 908-918.
- Bhat, N.K., Komschlies, K.L., Fujiwara, S., Fisher, R., Mathieson, B.J., Gregorio, T.A., Young, H.A., Kasik, J.W., Ozato, K. & Papas, T.S. (1989). *J. Immunol.*, **142**, 672-678.
- Bhat, N.K., Thompson, C.B., Lindsten, T., June, C.H., Fujiwara, S., Koizumi, S., Fisher, R.J. & Papas, T.S. (1990). *Proc. Natl. Acad. Sci. USA*, **87**, 3723-3727.
- Brown, T.A. & McKnight, S.L. (1992). *Genes Dev.*, **6**, 2502-2512.
- Dalton, S. & Treisman, R. (1992). *Cell*, **68**, 597-612.
- Gégonne, A., Punyanmmalee, B., Rabault, B., Bosselut, R., Seneca, S., Crabeel, M. & Ghysdael, J. (1992). *New Biol.*, **4**, 512-519.
- Hagman, J. & Grosschedl, R. (1992). *Proc. Natl. Acad. Sci. USA*, **89**, 8889-8893.
- Higashino, F., Yoshida, K., Kamio, K. & Fujinaga, K. (1993). *Nucleic Acids Res.*, **21**, 547-553.
- Higgins, D.G. & Sharp, R.M. (1988). *Gene*, **73**, 237-244.
- Hipskind, R.A., Rao, V.N., Mueller, C.G.F., Reddy, E.S.P. & Nordheim, A. (1991). *Nature*, **354**, 531-534.
- Johnson, P.F. & McKnight, S.L. (1989). *Annu. Rev. Biochem.*, **58**, 799-839.
- Karim, F.D., Urness, L.D., Thummel, C.S., Klemsz, M.J., McKercher, S.R., Celada, A., Van Beveren, C., Maki, R.A., Gunther, C.V., Nye, J.A. & Graves, B.J. (1990). *Genes Dev.*, **4**, 1451-1453.
- Kozak, M. (1991). *J. Biol. Chem.*, **266**, 19867-19870.
- Laget, M.P., Callebaut, I., de Launoit, Y., Stéhelin, D. & Mornon, J.P. (1993). *Nucleic Acid Res.*, **21**, 5987-5996.
- LaMarco, K., Thompson, C.C., Byers, B.P., Walton, E.M. & McKnight, S.L. (1991). *Science*, **253**, 789-792.
- Langer, S.J., Bortner, D.M., Roussel, M.F., Sherr, C.J. & Ostrowski, M.C. (1992). *Mol. Cell. Biol.*, **12**, 5355-5362.
- Laudet, V., Niel, C., Duterque-Coquillaud, M., Leprince, D. & Stéhelin, D. (1993). *Biochem. Biophys. Res. Com.*, **190**, 8-14.
- Lautenberger, J.A., Burdett, L.A., Gunnell, M.A., Qi, S., Watson, D.K., O'Brien, S.J. & Papas, T.S. (1992). *Oncogene*, **7**, 1713-1720.
- Leprince, D., Crépieux, P. & Stéhelin, D. (1992). *Oncogene*, **7**, 9-17.
- Leprince, D., Gégonne, A., Coll, J., de Taisne, C., Schneeberger, A., Lagrou, C. & Stéhelin, D. (1983). *Nature*, **306**, 395-397.
- MacLeod, K., Leprince, D. & Stéhelin, D. (1992). *Trends Biochem. Sci.*, **17**, 251-256.
- Martin, M.E., Piette, J., Yaniv, M., Tang, W.J. & Folk, W.R. (1988). *Proc. Natl. Acad. Sci. USA*, **85**, 5839-5843.
- Monté, D., Groux, H., Raharinivo, B., Plouvier, B., Dewulf, J., Clavel, T., Grangette, C., Torpier, G., Auriault, C., Capron, A. & Ameisen, J.C. (1992). *Blood*, **79**, 2670-2678.
- Moreau-Gachelin, F., Tavitian, A. & Tambourin, P. (1988). *Nature*, **331**, 277-280.
- Nunn, M.F., Seeburg, P.H., Moscovici, C. & Duesberg, P.H. (1983). *Nature*, **306**, 391-395.
- Nye, J.A., Petersen, J.M., Gunther, C.V., Jonsen, M.D. & Graves, B.J. (1992). *Genes Dev.*, **6**, 975-990.
- Rao, V.N., Huebner, K., Isobe, M., Ar-Rushdi, A., Croce, C.M. & Reddy, E.S.P. (1989). *Science*, **244**, 66-70.
- Reddy, E.S.P., Rao, V.N. & Papas, T.S. (1987). *Proc. Natl. Acad. Sci. USA*, **84**, 6131-6135.
- Seth, A., Ascione, R., Fisher, R.J., Mavrothalassitis, G.J., Bhat, N.K. & Papas, T.S. (1992). *Cell Growth Diff.*, **3**, 327-334.
- Thompson, C., Wang, C.Y., Ho, I.C., Bohjanen, P.R., Petryniak, B., June, C.H., Miesfeldt, S., Zhang, L., Nabel, G.J., Karpinski, B. & Leiden, J.M. (1992). *Mol. Cell. Biol.*, **12**, 1043-1053.
- Wasylyk, B., Wasylyk, C., Flores, P., Bègue, A., Leprince, D. & Stéhelin, D. (1990). *Nature*, **346**, 191-193.
- Watson, D.K., McWilliams, M.J., Lapis, P., Lautenberger, J.A., Schweinfest, C.W. & Papas, T.S. (1988). *Proc. Natl. Acad. Sci. USA*, **85**, 7862-7866.
- Wickens, M. (1990). *Trends Biochem. Sci.*, **15**, 277-281.
- Woods, D.B., Ghysdael, J. & Owen, M.J. (1992). *Nucleic Acids Res.*, **20**, 699-704.
- Xin, J.H., Cowie, A., Lachance, P. & Hassell, J.A. (1992). *Genes Dev.*, **6**, 481-496.
- Zhao, H.F., Labrie, C., Simard, J., de Launoit, Y., Trudel, C., Martel, C., Rhéaume, E., Dupont, E., Luu-The, V., Pelletier, G. & Labrie, F. (1991). *J. Biol. Chem.*, **266**, 583-593.



Discussion

1. Généalogie du groupe PEA3

Il est bon, avant d'entamer l'étude fonctionnelle d'Erm, de planter le décor du groupe PEA3, et notamment de préciser les relations évolutives entre les gènes qui le composent.

L'analyse du taux de divergence entre les séquences d'*erm*, *er81* et *pea3* a été réalisée sur des régions extérieures au domaine de liaison à l'ADN, trop conservé pour être informatif. Elle montre que les deux premiers gènes sont plus étroitement apparentés l'un à l'autre qu'ils ne le sont à *pea3*. Il semble donc qu'un "ancêtre" des trois gènes ait subi deux duplications successives :

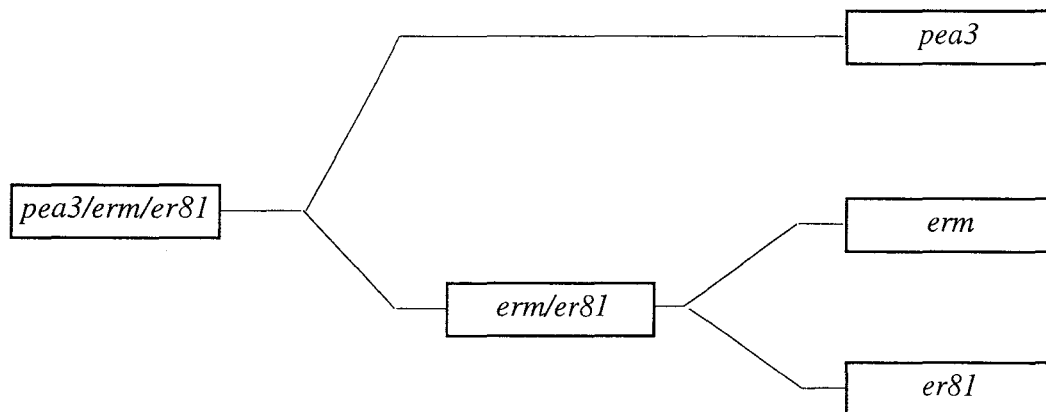


Fig. 29 Les gènes du groupe PEA3 sont apparus suite à deux duplications d'un "ancêtre"

Une vague de duplication des gènes de la famille ETS a eu lieu concomitamment à l'émergence des vertébrés (Laudet *et al.*, 1993; Laudet *et al.*, soumis). Si tel fut le cas pour le précurseur du groupe PEA3, on s'attend à trouver au moins deux gènes du groupe chez tous les vertébrés. De fait, nous avons isolé un équivalent d'*erm*, *er81* et *pea3* chez l'homme, la souris et le poulet (H. Pelczar, P. A. Defossez et Y. de Launoit, pour les gènes non encore publiés). Chez le poisson-zèbre, il existe au moins deux membres du groupe (A. Sharrocks, comm. pers.). Nous ne savons pas encore si la duplication *erm-er81* s'est produite avant la divergence poissons-reste des vertébrés.

Toutes nos tentatives d'isolement d'un quatrième gène du groupe PEA3, par criblage de banques, par Reverse Transcription suivie de PCR, ou par recherche dans les banques de données de courtes séquences exprimées (Expressed Sequence Tags), ont été infructueuses. Nous pensons donc qu'il n'existe que trois gènes dans le groupe PEA3.

2. Quelques éléments remarquables dans la séquence d'Erm

Nous nous livrerons ici à l'exercice périlleux des "expériences sur ordinateur". En d'autres termes, nous chercherons à exploiter quelques informations perceptibles dans la séquence d'Erm et, autant que faire se peut, à les corrélérer à des résultats obtenus par ailleurs.

2.1 Certaines régions des protéines du groupe PEA3 ont peu divergé

L'observation la plus immédiate est que certaines régions ont divergé moins vite que d'autres lors de l'évolution des protéines du groupe PEA3 (voir les acides aminés conservés dans la fig. 30). Deux régions sont particulièrement proches dans les trois membres : l'une (en noir sur la figure), constitue le domaine ETS qui confère le pouvoir de liaison à l'ADN (publication 1); l'autre (en gris) est riche en résidus à caractère acide et s'est avérée contenir un domaine activateur de transcription (voir publications 2 et 3). La région C-terminale adjacente au domaine ETS est partiellement conservée et contient un deuxième domaine transactivateur (publication 2).

2.2 Séquence PEST

Ces séquences induisent, dans certaines conditions, la dégradation rapide de la protéine qui les porte (revue dans Rechsteiner et Rogers, 1996). Une séquence PEST potentielle est présente dans la queue C-terminale d'Erm. Sa présence s'accorde à la courte demi-vie de la protéine. Celle-ci, mesurée par marquage métabolique de cellules puis chasse, n'excède pas une heure (J. L. Baert, non publié). A titre de comparaison, la protéine C-Jun, jugée instable, a une demi-vie de 90 minutes (Treier *et al.*, 1994).

Des séquences PEST ont été identifiées dans d'autres membres de la famille ETS, dont Ets-1 et Spi. "L'intérêt" d'imposer une demi-vie courte à ces protéines régulatrices est sans doute de permettre des fluctuations rapides de leur abondance dans la cellule. Notons, dans le même

ordre d'idées, que l'ARNm d'*erm* possède également des éléments déstabilisateurs (Monté *et al.*, 1996).

Remarquons que, lors de transfections transitoires, des mutants de délétion d'Erm privés de la région PEST ne semblent pas plus stables que des mutants la contenant (Publication n°3 fig. 5, P. A. Defossez et Y. de Launoit, non publié). Ce "paradoxe" semble simplement signifier que la séquence PEST détectée n'a pas de relevance fonctionnelle. En fait, il est peut-être hâtif de tirer cette conclusion : C-Jun contient une séquence PEST sur laquelle il est ubiquitinilé, ce qui entraîne sa dégradation . Cependant, lors de transfections, des mutants de délétion pourvus ou non de la région PEST s'accumulent à des niveaux similaires (Treier *et al.*, 1994).

Ne nions pas que ces observations sur la possible séquence PEST sont très préliminaires, et ne visent qu'à fournir des suggestions d'expériences.

2.3 Sites de modifications post-traductionnelles

On sait que des glycosylations sont appliquées à certains facteurs de transcription (Jackson et Tjian, 1988, par exemple), mais leurs conséquences fonctionnelles sont assez obscures. En revanche, comme évoqué dans l'introduction, la fonction des gènes ETS est fréquemment régulée par phosphorylation. C'est donc principalement cette modification que nous souhaitons mentionner.

Les travaux de notre groupe ont montré que l'activité transactivatrice d'Erm est modulable par la voie Ras/Raf/Erk, et la voie de la protéine kinase A (Janknecht *et al.*, (1996) et fig. 31). C'est également le cas d'Er81 (Janknecht, 1996); quant à Pea3, le groupe de John Hassell a montré que son activité était influencée par les kinases de la cascade Ras/Raf/Erk, mais aussi Rho/Rac/Pak/Jnk (O'Hagan *et al.*, 1996).

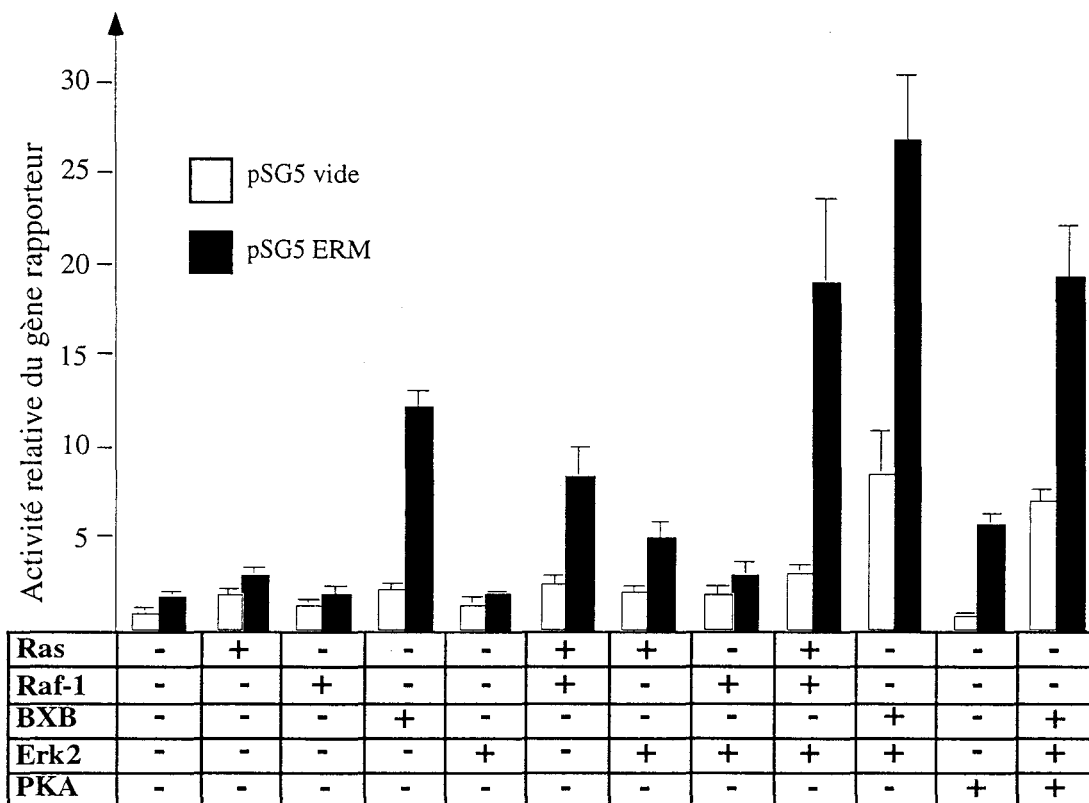


Fig. 31 L'activité d'Erm est modulée par des kinases de la voie Ras et par la Protéine Kinase A

On a représenté l'activité d'un gène rapporteur dont le promoteur porte trois sites ETS, après cotransfection transitoire des constructions indiquées, plus un vecteur vide (barres blanches), ou le même vecteur portant Erm (barres noires).

BXB est une forme activée de Raf. D'après Janknecht *et al.*, (1996)

2.3.1 Sites de phosphorylation par les kinases MAP

On désignera comme MAP Kinases (Mitogen-Activated Protein Kinases) des protéines de plusieurs groupes : les kinases du groupe Erk, activées par les mitogènes; celles du groupe Jnk, stimulées par les stress (UV, anisomycine, par exemple); et enfin la kinase p38 (voir Cahill *et al.*, 1995, pour revue). Ces enzymes peuvent phosphoryler des Sérines ou des Thréonines, quand celles-ci sont suivies d'une Proline (site S/T-P). Il existe six sites potentiels dans Erm (Fig. 30). Les kinases Erk connues à l'heure actuelle ont une exigence plus précise : leur activité est maximale aux sites P-X-S/T-P, X étant un acide aminé quelconque (Songyang *et al.*, 1996).

Notre équipe a porté son attention sur les **trois sites potentiels de phosphorylation par les kinases Erk** juxtaposés aux acides aminés 133-143 d'ERM, qui sont conservés dans les gènes de souris et de poulet. Ces sites sont les seuls phosphorylés par les kinases Erk *in vitro*, mais un mutant d'Erm dont ces sites sont absents est encore activable par la combinaison BXB/Erk2 *in vivo* (J. L. Baert et Y. de Launoit, non publié). Il y a plusieurs possibilités pour expliquer ces observations (voir ci-dessous). Jean-Luc Baert tente actuellement de découvrir si la phosphorylation induite par BXB est médiée par Erk ou une autre kinase, et si elle s'applique directement à Erm ou à un possible cofacteur (voir Johnson *et al.*, 1996, pour un tel cas où la phosphorylation augmente l'activité de la protéine MRF4 *via* un de ses partenaires).

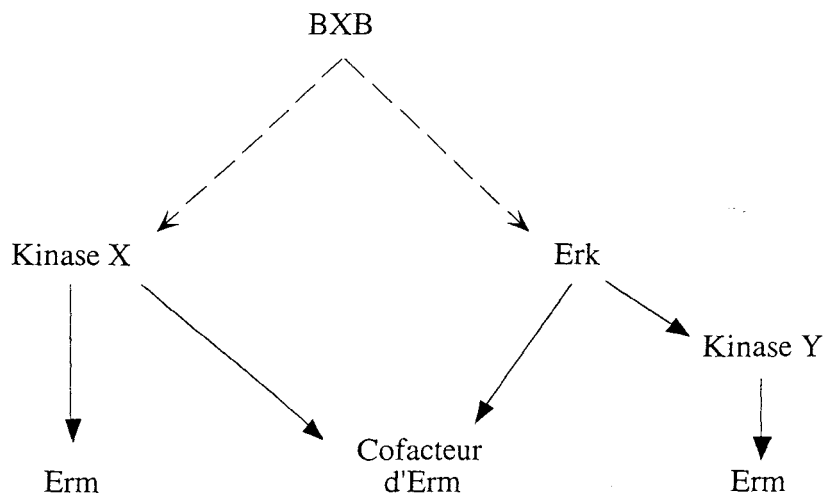


Fig. 32 Plusieurs possibilités pour l'action de BXB sur Erm

2.3.2 Sites de phosphorylation par la protéine kinase dépendante de l'AMPc

Un site de phosphorylation par la Protéine Kinase A (PKA) est présent au début du domaine ETS. Il se pourrait donc que l'effet activateur de la PKA résulte de la phosphorylation de ce site présent dans le domaine ETS, ce qui modulerait la fixation à l'ADN.

Rappelons que l'activité de liaison à l'ADN de Ets-1 est régulée par la phosphorylation par la Protéine Kinase C (PKC) de sites adjacents au domaine ETS (Rabault Ghysdael, 1994). **Deux sites PKC** sont présents dans le domaine ETS d'Erm et pourraient également constituer un point de régulation.

Des tests sont en cours pour étudier ces deux possibilités (Jean-Luc Baert, non publié).

2.3.3 Sites de phosphorylation par la Caséine Kinase II

On détecte de nombreux **sites putatifs de phosphorylation par la caséine kinase II (CKII)**. L'un d'eux mérite peut-être plus d'attention : situé dans la queue C-terminale, il ressemble à des sites effectivement régulés par la CKII *in vivo* (Glineur *et al.*, 1989, par exemple).

De plus, on sait que des sites CKII situés dans des régions PEST peuvent moduler le taux de dégradation des protéines. IκB en est un exemple convaincant (voir par exemple Schwarz *et al.*, 1996).

Notre équipe a montré que le domaine ETS d'Erm, exprimé indépendamment du reste de la protéine, est importé dans le noyau (Janknecht *et al.*, 1996). Il doit donc contenir un site de localisation nucléaire, et de fait on y trouve une séquence s'apparentant au consensus un peu flou déterminé pour ces signaux (Dingwall et Laskey, 1991). On a rapporté que la phosphorylation de sites CKII situés à proximité des signaux de localisation nucléaire pouvait moduler le transit Cytoplasme-Noyau de certaines protéines (Rihs *et al.*, 1991; revue dans Jans, 1995).

Evidemment, il ne s'agit là que de "pistes", sur lesquelles nous ne nous sommes pas encore aventurés...

3. Carte d'expression comparée d'*erm*, *er81* et *pea3*

Notre description a été jusqu'ici bien statique... L'expression des gènes ETS, qui régulent tant des aspects du développement, est évidemment un processus dynamique soumis à des régulations complexes.

Nous avons cherché à dresser une carte spatiotemporelle précise de l'expression d'*erm*, *er81* et *pea3*. Pour ce faire, nous avons précisé les données préliminaires obtenues par Northern Blot en menant des hybridations *in situ* sur embryons de souris à différents stades du développement (A. Chotteau-Lelièvre, X. Desbiens, H. Pelczar, P. A. Defossez et Y. de Launoit, soumis).

Les trois gènes sont coexprimés dans de nombreux organes (par exemple cerveau, rein ou muscle), mais pas toujours dans les mêmes sous-types cellulaires de ces organes. Il existe par ailleurs des sites d'expression spécifiques pour chacun des gènes.

pea3, *erm* et *er81* sont exprimés de façon importante au cours des événements précoces de prolifération et de migration cellulaire. Le motif le plus saillant est sans nul doute l'association récurrente de leur expression avec les **événements morphogénétiques impliquant des interactions épithélium-mésenchyme**. **Nous pensons que les membres du groupe PEA3 tiennent un rôle actif dans ces processus**, en régulant la transcription de gènes impliqués dans le remodelage de la matrice extracellulaire et les propriétés d'adhérence des cellules.

Plusieurs résultats soutiennent cette proposition :

- un événement morphogénétique est strictement corrélé à un patron d'expression spécifique des trois gènes. Par exemple, le rein, le poumon, la glande salivaire et la glande mammaire contiennent chacun des canaux issus de la ramification d'un épithélium sous l'influence du mésenchyme environnant. Dans chacun de ces organes, on observe qu'*erm* et *pea3* sont exprimés à l'extrémité des tubules en croissance, tandis qu'*er81* n'est détectable que dans le mésenchyme.

- la corrélation spatiale s'accompagne d'une corrélation temporelle qui n'est pas moins stricte : l'expression des gènes chute avec la différenciation épithéliale.

- on sait que la protéine Pea3 active la transcription de protéases dégradant la matrice extracellulaire (Higashino *et al.*, 1995). Le remodelage de la matrice est un événement crucial lors de l'organogenèse, mais aussi de l'invasion tumorale. C'est sans doute pourquoi l'expression de Pea3 dans des cellules mammaires tumorales suffit à leur conférer un pouvoir métastatique (Kaya *et al.*, 1996).

- enfin, Erm peut moduler la transcription d'un gène codant une molécule d'adhérence cellulaire, ICAM-1 (InterCellular Adhesion Molecule-1, de Launoit *et al.*, Soumis). Les propriétés d'adhérence des cellules influent sur leur mobilité lors des événements morphogénétiques, mais permettent aussi la cohésion des épithéliums.

Le développement de certains organes, comme le rein, le poumon ou la dent, peut partiellement être reproduit *in vitro* à partir d'explants. L'utilisation de ces systèmes modèles permettra d'apporter des preuves plus directes à l'appui de notre hypothèse. En perturbant l'expression de certains des gènes du groupe PEA3, par exemple à l'aide d'oligonucléotides antisens, on verra si l'on perturbe les phénomènes de morphogenèse.

Nous ne savons pas si les trois gènes *pea3*, *erm* et *er81*, sont fonctionnellement équivalents, ou en d'autres termes, si leur différence fonctionnelle ne tient qu'à leur patron d'expression spécifique. Nous savons, par contre, que les protéines Erm et Pea3 se fixent à des séquences d'ADN très similaires lors de tests de sélection *in vitro* (J. L. Baert, Y. de Launoit et A. Sharrocks, non publié). Les résultats exposés dans les publications 2 et 3 montrent que leurs domaines transactivateurs ont sans doute des caractéristiques proches.

Nous espérons obtenir des éléments de réponse à la question de la conservation de fonction des gènes lors de l'obtention de souris dont les gènes *erm* et/ou *er81* auront été inactivés par recombinaison homologue (Travaux en cours dans notre groupe).

Publication n°2

Two functionally distinct domains cooperate for transactivation by the Ets family member ERM.

Laget M.P.*, Defossez P.A.*, Albagli O., Baert J.L., Dewitte F.,
Stéhelin D. et de Launoit Y.

(* : premiers auteurs associés)

Oncogene (1996), **12**, 1325-1336

Résumé:

Dans cet article nous étudions les mécanismes de liaison à l'ADN et de transactivation par Erm.

Nous montrons :

- que la fixation à l'ADN d'Erm est inhibée de façon intramoléculaire. Deux régions sont impliquées dans ce phénomène : la première réside au centre de la protéine, et la deuxième est adjacente au domaine ETS, et constitue la queue C-terminale d'Erm.

- qu'Erm active la transcription d'un gène dont le promoteur contient des sites ETS multimérisés.

- que cet effet peut être attribué à deux domaines transactivateurs, l'un situé dans les 72 premiers acides aminés d'Erm, et l'autre dans la queue C-terminale. Ces deux domaines présentent des différences fonctionnelles claires, et exercent un effet synergique.

- enfin, que le potentiel transactivateur d'Erm est également régulé de façon intramoléculaire.

Quand la protéine n'est pas liée à l'ADN, l'action des deux domaines transactivateurs est inhibée par le domaine ETS.

Nous proposons un modèle dans lequel Erm transite entre une conformation "fermée" et inactive, et une conformation "ouverte".



Two functionally distinct domains responsible for transactivation by the Ets family member ERM

Marie-Pierre Laget, Pierre-Antoine Defossez, Olivier Albagli*, Jean-Luc Baert, Frédérique Dewitte†, Dominique Stehelin and Yvan de Launoit

Unité d'Oncologie Moléculaire, CNRS URA 1160, Institut Pasteur, 1 rue Calmette, 59019 Lille Cedex, France

The recently cloned human Ets transcription factor ERM is closely related to the ER81 and PEA3 genes. Here, we report the functional analysis of the DNA-binding and transactivation properties of ERM. Specific DNA-binding by ERM requires the ETS domain, conserved in all members of the Ets family and is inhibited by an 84 residue long central region and the carboxy-terminal tail. Two fragments of ERM are transferrable activation domains: α , which sits in the 72 first residues and encompasses the acidic domain conserved between ERM, ER81 and PEA3, and the carboxy-terminal tail which also bears a DNA-binding inhibition function. Deletion of α strongly reduces transactivation by ERM. Moreover, α and the carboxy-terminal tail exhibit functional synergism, suggesting that they activate transcription through different mechanisms. In support of this idea, we demonstrate that VP16 quelches transactivation by α but not by the carboxy-terminal tail. This result also indicates that α and VP16 may share common limiting cofactors. α and the carboxy-terminal tail do not seem to be conserved within the whole Ets family, indicating that the specificity of ERM may rely on interactions with distinct cofactors.

Keywords: Ets gene; ERM; DNA-binding; transactivation

Introduction

Initiation of messenger RNA synthesis is a primary control point for the spatial and temporal regulation of specific gene expression and, in higher eukaryotes, transcriptional modifications can have profound consequences on development and differentiation. As a group of transcription factors, the Ets family is involved in normal development of various species: *pointed* and *yan* have opposite effects on *Drosophila* eye photoreceptor determination (O'Neill *et al.*, 1994), *ets-2* expression is required for *Xenopus* oocyte maturation (Chen *et al.*, 1992) and Spi-1/PU-1 is necessary for the development of mouse hematopoietic lineages (Romano-Spica *et al.*, 1994; Scott *et al.*, 1994; de Crépieux *et al.*, 1994 for a review). They also participate in the regulation of viral genes: c-Ets-1

activates the HTLV1 LTR in T cells (Bosselut *et al.*, 1990). GABP α is found as part of a polypeptide complex bound to the enhancer of a herpes simplex virus immediate early (HSV IE) gene (LaMarco *et al.*, 1991; Thompson and McKnight, 1992). Some of these genes can turn oncogenic when mutated (*v-ets* being the transforming counterpart of *c-ets-1*; Metz and Graf, 1991), when overexpressed following retroviral insertion [*fli-1*, (Ben David *et al.*, 1991); *spi-1*, (Moreau-Gachelin *et al.*, 1988)], or when included in a chimeric protein as a consequence of translocation: EWS-Fli-1 (Delattre *et al.*, 1992), EWS-Erg (Zucman *et al.*, 1993), t(12;21) ERG (Shimizu *et al.*, 1993), EWS-ETV1 (Jeon *et al.*, 1995) and Tel-PDGFR β (Golub *et al.*, 1994).

Ets transcription factors are classified according to their sequence identity in the DNA-binding domain which is widely conserved during evolution (Lautenberger *et al.*, 1992; Degnan *et al.*, 1993; Laudet *et al.*, 1993). The structure of this 85 residue domain, named the ETS domain, that enables the transcriptional regulatory proteins to be anchored to a promoter or enhancer in a sequence specific manner, has been recently determined through NMR analysis (Donaldson *et al.*, 1994; Liang *et al.*, 1994a,b; Shore *et al.*, 1995). It places the Ets transcription factors into the helix-turn-helix DNA-binding protein superfamily, along with a wide array of prokaryotic factors and homeodomain proteins involved in eukaryotic development (Qian *et al.*, 1989; Kissinger *et al.*, 1990; Wolberger *et al.*, 1991; see Pabo and Sauer, 1992; Gehring *et al.*, 1994 for reviews).

Based on the extent of amino acid sequence identity in the ETS domain, Ets family members are divided into nine different groups (Laudet *et al.*, 1993). The PEA3 group of the Ets family so far comprises three members, ER81, PEA3 and ERM (Ets Related Molecule). Mouse ER81 is widely expressed (Brown and McKnight, 1992) and is present in all tissues tested with high mRNA level, in brain, testis, lung and heart (Jeon *et al.*, 1995; Monté *et al.*, 1995). Its human counterpart, ETV1, has been recently shown to be implicated in Ewing's sarcoma as a fusion protein with EWS (Jeon *et al.*, 1995).

In contrast, in adult mice, the expression pattern of PEA3 is highly restricted: PEA3 mRNA is mainly detected in the epididymis and the brain, and, to a lesser extent, in the mammary gland (Xin *et al.*, 1992). PEA3 can activate the promoters of three different matrix metalloproteinase genes while an elevated expression of these enzymes is known to affect normal mechanisms of embryogenesis and wound healing (Higashino *et al.*, 1995). Moreover, PEA3 is overexpressed in metastatic mammary adenocarcinomas of

Correspondence: Y de Launoit
Present addresses: *U124 INSERM, IRCL, Place de Verdun 59045 Lille Cedex, France. †Laboratoire d'Endocrinologie Moléculaire, H.U.L. Research Center, 2705 boulevard Laurier, G1V4G2 Sainte-Justine, Quebec, Canada
M-P Laget and P-A Defossez contributed equally to this paper
Received 2 October 1995; revised 13 December 1995; accepted 13 December 1995

neu oncogene – or Polyomavirus middle T antigen – carrying transgenic mice (Trimble *et al.*, 1993). Taken together, these results suggest that the PEA3 group members could be required for normal developmental processes and that their overexpression may participate in transformation and metastasis.

ERM has a wide expression pattern and encodes a 510 amino acid protein that is 95% and 96% identical in the ETS domain to PEA3 and ER81, respectively. The ERM ETS domain (residue 363 to residue 449) is necessary and sufficient for binding to an ETS specific probe (Monté *et al.*, 1994). Outside the ETS domain,

ERM, PEA3 and ER81 also share some sequence identity in an acidic region located at the amino-terminus and in the carboxy-terminal tail.

In this report, we have functionally dissected the ERM protein, and DNA-binding inhibition and transactivation could be attributed to various parts of ERM. We have generated deletion constructs with modified DNA-binding abilities, and consequently delineated the central part and the carboxy-tail of the protein as inhibitors of the DNA-binding ability. We have identified two domains bearing autonomous transactivation function as the first 72 amino acids

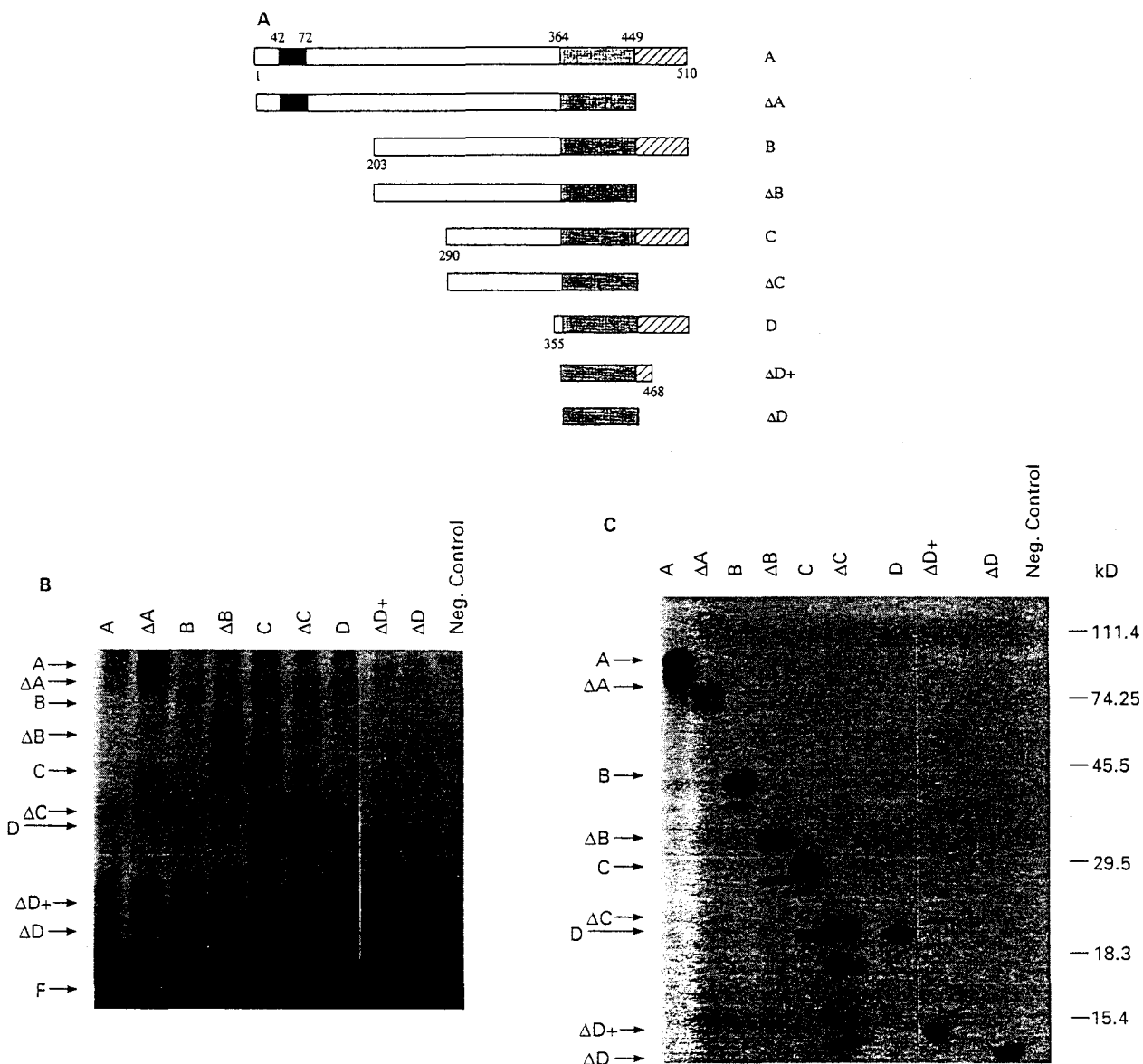


Figure 1 Sequences outside the ERM ETS domain inhibit its DNA-binding affinity. (A) Schematic representation of the ERM full-length open reading frame and of the deletion proteins. The ETS domain is represented by the grey box, while the acidic domain and the carboxy-terminal tail are indicated as the black box and the hatched box, respectively. (B) The deletion proteins were synthesized in rabbit reticulocyte lysate, and their DNA-binding capacity assessed performing electrophoretic mobility-shift assays with an equal amount of each protein. We used as a specific probe a ³²P-labeled double-stranded oligonucleotide containing the PEA3 motif described in Materials and methods. The protein-DNA complexes (arrows) were separated from the free probe (F) on a non-denaturing 5–20% gradient polyacrylamide (29/1) gel and visualized on an X-ray film after 18 h exposure at –70°C. (C) A representative experiment showing the *in vitro* ³⁵S-Methionine labeled translation products of the deletion proteins, separated on a 10–30% gradient SDS-PAGE (see Materials and methods). Proteins corresponding to the translated product starting at the initiation methionine are indicated by arrows. The relative amount of the proteins was estimated by quantifying the radioactivity of the bands with a PhosphorImager (Molecular Dynamics), and making corrections to compensate for the different numbers of methionines in the constructs

encompassing the acidic region, and the less potent carboxy-terminal tail. Using GAL4 DNA-binding domain chimeras, we showed that deletion of α strongly affects transactivation by ERM. These two domains act synergistically. Since we found that the 72 amino acid domain but not the carboxy-terminal tail could be squelched by the Herpes virus VP16 transactivation domain, we propose that the basis of this synergism is the interaction of those two domains with distinct targets of the transcriptional machinery.

Results

ERM contains two inhibitory domains of DNA-binding adjacent to the ETS domain

We had previously observed that the ERM ETS domain more strongly bound to its sequence specific target than the full size ERM protein, leading to the conclusion that it may contain DNA-binding modulatory domains. In order to detect those regions, we synthesized various carboxy- and amino-terminal deletion mutants of ERM (Figure 1A). Their DNA-binding capacities were compared using equal amounts of proteins (Figure 1C).

Full length ERM (A, Figure 1B) bound to DNA to a lesser extent than the isolated ETS domain (Δ D).

Indeed, the more the protein was deleted in the amino-terminal part, the stronger the binding to the ETS probe was observed (compare A, B or C with D in Figure 1B and, for the constructs lacking the carboxy-terminal tail Δ A, Δ B or Δ C with Δ D).

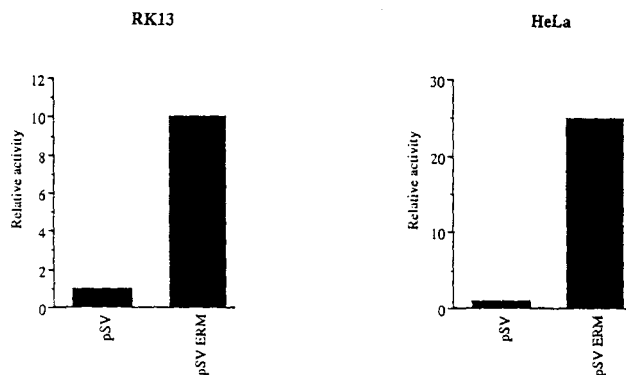


Figure 2 A representative experiment showing that human ERM activates the transcription of a reporter plasmid containing ETS binding sites. Rabbit kidney carcinoma (RK13) and HeLa cervix carcinoma (HeLa) cells were transfected as described in Materials and methods with 0.8 μ g pSV or pSV ERM expression vector and 0.3 μ g 3 \times TORU-tk-CAT reporter. Transfection assays by the lipofectamin technique were performed at least three times, in duplicate on at least two different plasmid preparations. The data are calculated relative to the control plasmid (pSV)

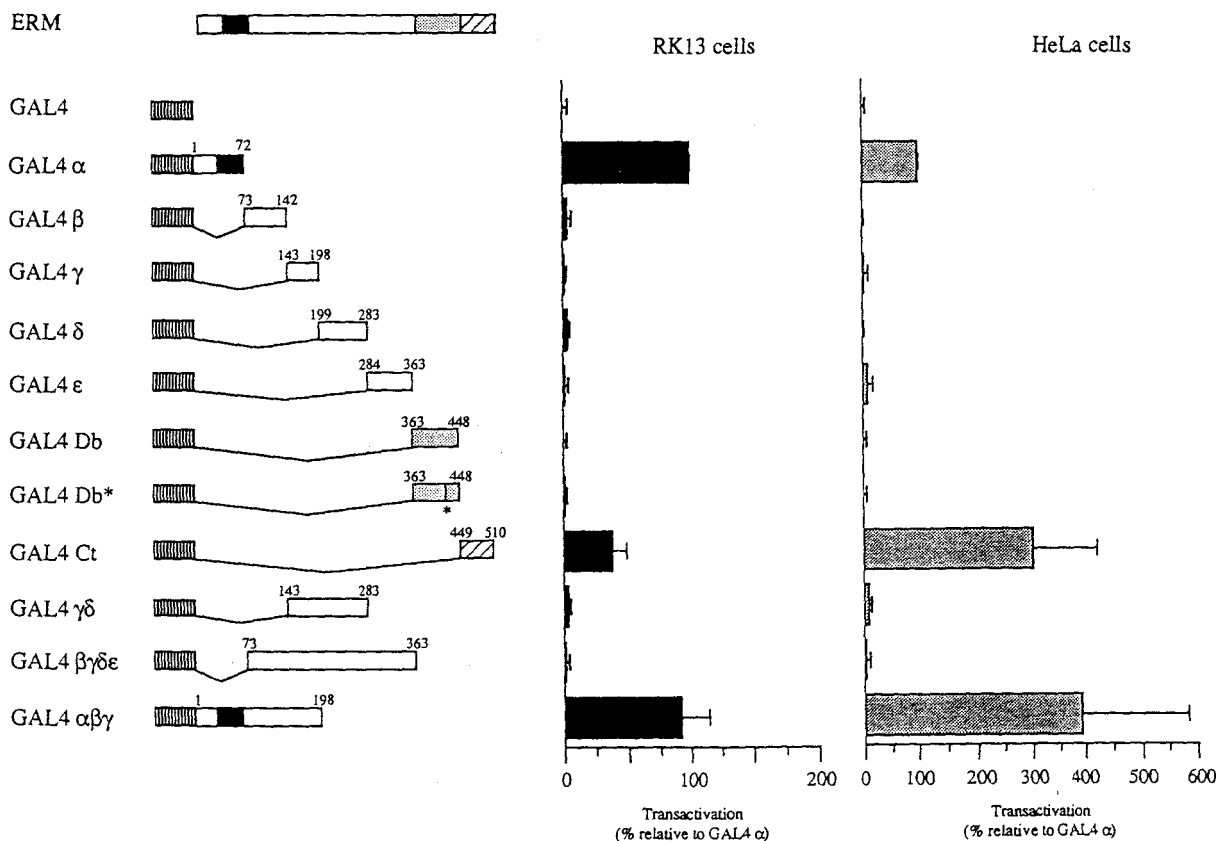


Figure 3 Identification of ERM fragments harboring intrinsic transactivating capacity. The various fragments we fused to the GAL4 DNA-binding domain are represented, the numbers indicated being the corresponding ERM amino acids. The transfection experiments were performed as described in the legend to Figure 2 but with a 5 \times GAL4-tk-CAT reporter plasmid. The background activity of this reporter is very low in the cell lines used, making it difficult to calculate fold increases relative to the control (pSG424 vector). The transactivation activities are therefore calculated relative to that of α (100%) which is the fragment exhibiting the strongest activity. The data are presented as mean \pm standard error of the mean

Moreover, since A and B behaved similarly while C bound the ETS probe stronger (Figure 1B, compare lanes A or B with lane C), we suggest that a region spanning residues 203 to 290 inhibits DNA-binding. This hypothesis is strengthened by the fact that the corresponding constructs lacking the carboxy-terminal tail behave in the same manner (compare ΔB and ΔC).

We also observed that each of the A, B, C and D deletion mutants bound to DNA much less avidly than their respective Δ counterparts (Figure 1B). These results implied that the carboxy-terminal tail of ERM exerts a negative influence on its DNA-binding ability, an effect which does not depend on the sequences upstream from the DNA-binding domain.

Amino acid sequence alignment of the three PEA3 group members revealed that a region of 19 amino acids downstream from the ETS domain is highly conserved. We thus generated the $\Delta D+$ truncated protein which contains the ETS domain flanked by this conserved region. As shown in Figure 1B, $\Delta D+$ had the same DNA-binding capacity as ΔD , suggesting that the conserved region on its own is not sufficient to inhibit DNA-binding but that the carboxy-terminal inhibition we observed was mediated by the less conserved region.

We thus showed that DNA-binding ability of ERM is inhibited by the central part of the protein as well as by its carboxy-terminal tail.

The α region and the carboxy-terminal tail show intrinsic transactivating ability

As shown in Figure 2, the wild-type ERM protein activates transcription of a reporter gene placed downstream of an artificial promoter containing three ETS binding sites ($3 \times \text{TORU-tk-CAT}$). This transactivation effect relative to the control is 10- and 25-fold in RK13 and HeLa cells, respectively.

In order to delineate the regions of ERM responsible for this transactivating ability, we generated a series of fragments spanning the whole ERM open reading frame and fused to the DNA-binding domain of the GAL4 protein, and tested these constructs for their capacity to modulate the transcription of a CAT reporter plasmid containing five GAL4 binding sites ($5 \times \text{GAL4-tk-CAT}$).

These constructions are depicted in Figure 3 and their activity is indicated relatively to that of α . As previously shown, the transfection assays were carried out in the RK13 and HeLa cell lines. It appeared that the first 72 amino-terminal residues of ERM, termed α , and the carboxy-terminal tail (Ct) of ERM enhanced transcription of the GAL4-reporter plasmid when fused to the GAL4 DNA-binding domain, whereas all the other fragments (β , γ , δ , ϵ and Db) failed to do so (Figure 3). Moreover, the transactivation level of GAL4 Ct is 37.5% and 300% that of GAL4 α in RK13 and HeLa cells, respectively, thus indicating that these two transactivation domains behave differently, depending on the cell type.

To test whether the presence of an active ETS domain could interfere with the activating properties of GAL4 Db, we looked for a point mutation in the ETS domain that totally abrogated its DNA-binding ability. By electrophoretic mobility-shift assay on an ETS probe derived from the Polyomavirus enhancer (Figure

4A), we showed that any mutation of Y419 into alanine, phenylalanine, proline or tryptophan completely abolished the DNA-binding ability of the ETS domain (ΔD). Mutation of Y443 had various effects depending on the amino acid chosen; compared to ΔD ,

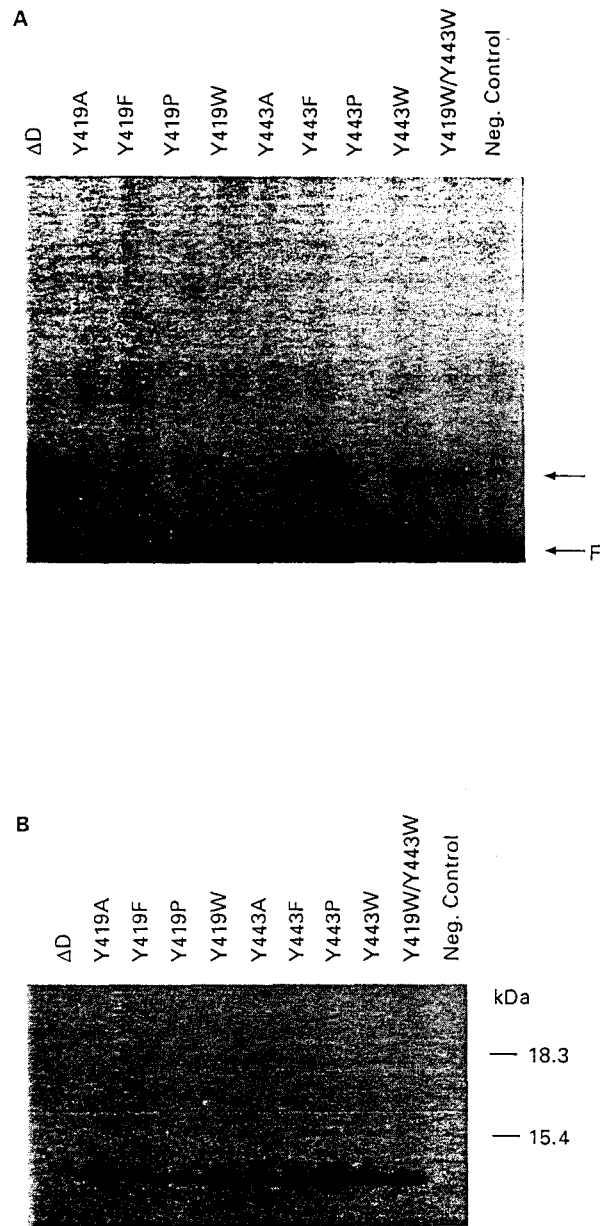


Figure 4 Mutations of two conserved tyrosines in the ERM ETS domain (ΔD) have different consequences on its DNA-binding ability. As indicated above each lane, Y419 and Y443 in the ERM ETS domain (ΔD) were mutated into alanine (Y419A and Y443A), phenylalanine (Y419F and Y443F), proline (Y419P and Y443P) and tryptophan (Y419W and Y443W). Y419W/Y443W correspond to the double mutation of Y419 and Y443 into tryptophan (see Materials and methods). (A) The ETS domain point mutants were synthesized in rabbit reticulocyte lysate, and analysed by an electrophoretic mobility-shift assay with an ETS probe. As in Figure 1, rabbit reticulocyte lysate of pSG5 vector was used as negative control (Neg. Control). F is the free probe and the DNA-protein complexes are indicated by arrow. (B) A representative electrophoresis of the ^{35}S -Methionine labeled wild-type and mutated ΔD proteins overexpressed in rabbit reticulocyte lysate. Proteins were separated on a tricine-SDS-PAGE (50) and visualized on X-ray film after 3 days exposure at -70°C . Their relative amount was estimated as described in Figure 1

Y443A and Y443W have dramatically decreased DNA-binding ability, while Y443P was unable to bind DNA. Finally, mutation of Y443 into phenylalanine had no effect on the DNA-binding ability. Interestingly, double mutation of Y419 and Y443 into tryptophan resulted in a binding comparable to the binding of Y443W, while Y419W did not bind DNA. These results were not due to a difference of protein expression level (Figure 4B) since electrophoretic mobility-shift assays were completed with equal amounts of protein.

Mutation of tyrosine 419 into proline was consequently introduced into the GAL4 Db construct, yielding to the GAL4 DB* fusion protein. This chimera still could not activate transcription (Figure 6), indicating that the ETS domain itself is not a transactivating region or that the GAL4 Db* is not in a conformation fit for transactivation.

The possibility remained that a region located between α and Db has transactivating properties, but that the fragments generated represented only parts of this region, and thus were devoid of activity. To investigate this possibility, we assayed the activity of two chimeras: GAL4 $\gamma\delta$, GAL4 $\beta\gamma\delta\epsilon$. Moreover GAL4 $\beta\gamma\delta\epsilon$ Db, referred to below as GAL4 ERM- α , is also devoid of any transactivating ability (Figure 7). Neither of these constructs activated transcription (Figure 3). We also observed that the GAL4 $\alpha\beta\gamma$ chimera displayed a transactivation potency comparable to

that of GAL4 α in RK13 cells, while it was much higher in HeLa cells, thus indicating the differential behavior of the GAL4 chimeras depending on the cell type.

It thus appears that α and the 61 amino acid long Ct domains are the only two ERM regions bearing transactivation properties.

Definition of ERM minimal regions sufficient for activation

In an attempt to map the ERM transactivating regions with more accuracy, we assayed the activity of subregions of the activating domains hereabove defined. In particular, the last 31 amino acids of α constitute the 'acidic domain' (aD), which is highly conserved in PEA3 and ER81 and could be hypothesized to be an intrinsic transactivating region due to its richness in acidic amino acids. We thus fused amino acid 41 to amino acid 72 of ERM to the GAL4 DNA-binding domain (yielding GAL4 aD). This chimera showed a transcriptional activity reduced to about 10% and 30% of that of α in RK13 and HeLa cells, respectively (Figure 5). However, this activity is significantly higher than the GAL4 chimeras showing no transcriptional activity (Figures 3 and 5). This weak intrinsic transactivation establishes that the aD region by itself is active but is not fully responsible for the transactivation achieved by α .

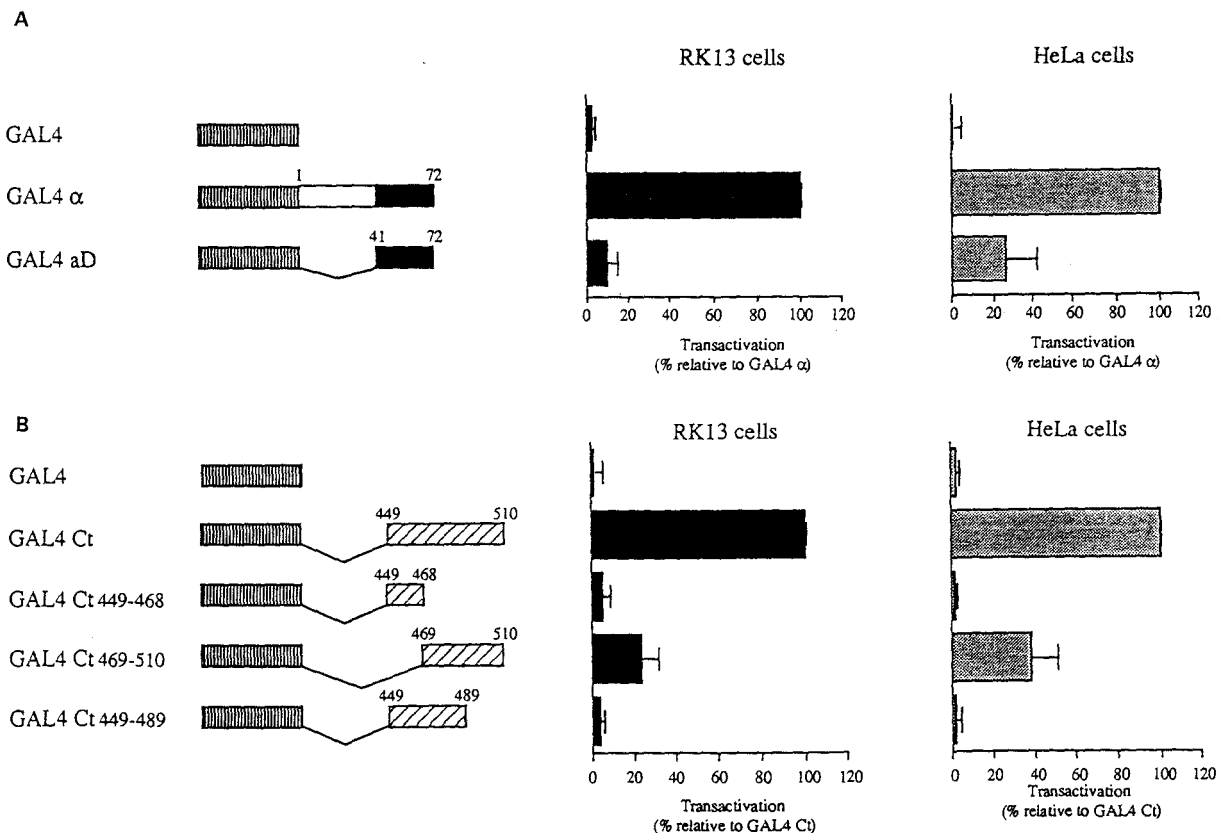


Figure 5 Delineation of the minimal domains capable of transactivation. Subdomains of α and Ct we fused to the GAL4-DNA-binding domain are presented, the numbers indicating the corresponding ERM amino acids. The transfection experiments were performed as described in the legend to Figure 2. (A) The activity of aD (acidic domain *stricto sensu*) is calculated relatively to that of α (100%). (B) The activities of Ct deletion constructs (449-468, 469-510 and 449-489) are calculated relatively to that of full-length Ct (100%)

We then sought to identify subregions of Ct that retained activity. We produced a deletion construct spanning amino acids 449 to 468 of ERM, GAL4 Ct 449–468, which encompasses the region of Ct which is contiguous to the ETS domain and highly similar among the PEA3 group. As illustrated in Figure 5, we observed that neither GAL4 Ct 449–468 nor GAL4 Ct 449–489 could activate transcription. GAL4 Ct 469–510 still transactivated about 20% and 40% relatively to that of GAL4 Ct in RK13 and HeLa cells, respectively. We thus conclude that the subregion of Ct which is not conserved among the PEA3 group is partly responsible for transactivation, but that it requires residues located in the conserved region to achieve full level activation.

Synergistic effect of α and Ct domains

As shown in Figure 6, we simultaneously expressed the GAL4 chimeras containing α or Ct, both described as ERM transactivation domains. We found the coexpression of GAL4 Ct to potentiate GAL4 α -mediated transactivation. Indeed the level of activity achieved by cotransfection was approximately four times higher than the sum of the activities measured separately for the GAL4 α and the GAL4 Ct chimeras (Figure 6), demonstrating a functional synergy between α and Ct. Note that the transfection assays were performed with only 100 ng/well reporter plasmid. We suppose that reducing the ratio of reporter to expression vector increases the probability that both GAL4 α and GAL4 Ct fusion proteins are found on the same promoter element, thereby being able to cooperate more efficiently. In fact, with 300 ng/well reporter plasmid, synergistic effect of α and Ct was significant but less important (data not shown). We also tested as a control all the other GAL4 chimeras assayed for intrinsic transactivation potential. When fused to GAL4, β , γ , δ , ϵ and Db were unable to increase transactivation by GAL4 α (data not shown). In order to further substantiate our results, we fused α to the heterologous DNA-binding domain of Reverb-A. This construct was cotransfected with GAL4 Ct and a reporter plasmid bearing the corresponding two target sequences. The effect was also five times higher than the activity measured for the α chimera alone, confirming a significant synergy between α and Ct (data not shown).

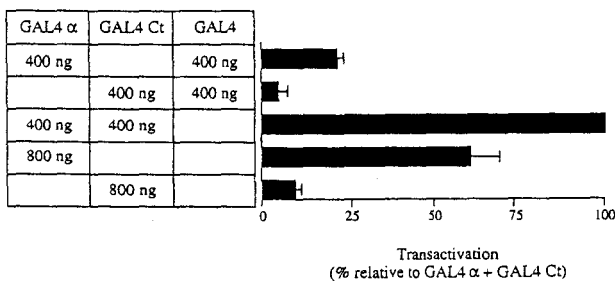


Figure 6 A synergistic effect between α and Ct. Except for the use of 100 ng reporter plasmid instead of 300 ng, the transactivation experiments were performed in RK13 cells as described in the legend to Figure 2, with 800 ng and 400 ng GAL4 α or GAL4 Ct. Cotransfection was assayed with 400 ng GAL4 α and 400 ng GAL4 Ct. Transactivation activities are calculated relatively to that of 400 ng GAL4 α + 400 ng GAL4 Ct (100%)

Negative influence of the ETS domain on transactivation

We sought to investigate the effect of deletion of α and Ct, both bearing some intrinsic transactivation potential, on the ERM transactivation ability. However, our previous results indicated that Ct bears a DNA-binding inhibitory function. Binding through the heterologous yeast GAL4 domain allowed us to overcome the problem of variable DNA-binding ability by the deletion mutants. The complete ERM open reading frame was inserted in-frame into the pSG424 vector, yielding the GAL4 ERM construct (Figure 7A). Deletion of regions exhibiting intrinsic transactivation potential, namely α which encompasses the acidic domain (amino acids 42 to 72), or Ct, were engineered either independently (GAL4 ERM- α , GAL4 ERM-Ct) or in combination (GAL4 ERM- α Ct). Figure 7B shows that in RK13 and HeLa cells, GAL4 ERM, GAL4 ERM- α , GAL4 ERM-Ct and GAL4 ERM- α Ct chimeras were all unable to significantly transactivate

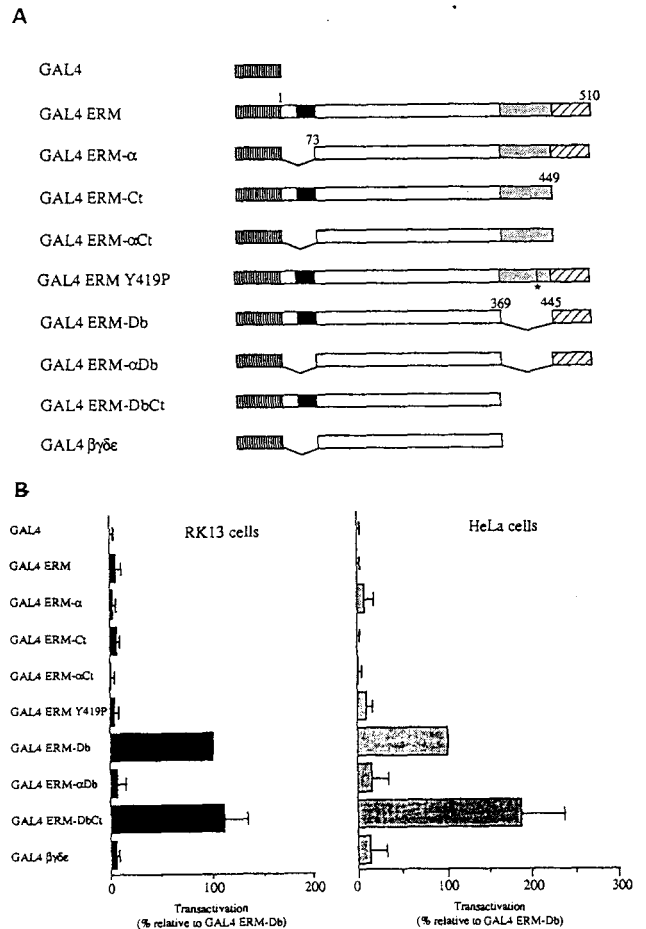


Figure 7 An amino-terminal domain is essential for transactivation by ERM. (A) Representation of the chimeric protein GAL4 ERM and the deletion constructs GAL4 ERM- α (lacking the first 72 amino acids), GAL4 ERM-Ct (lacking the carboxy-terminal tail) and GAL4 ERM- α Ct (lacking both). Their counterparts lacking the entire ETS domain are described below. GAL4 ERM Y419P is a point mutant exhibiting no more DNA-binding ability. (B) The GAL4 chimeras were tested in RK13 and HeLa cells for their ability to increase transcription from a 5 \times GAL4-tk-CAT reporter. The transfection experiments were performed as described in the legend to Figure 2. We therefore calculated the transactivation activities relatively to that of GAL4 ERM-Db (100%)

to a level superior to that achieved by the GAL4 DNA-binding domain. The inability of the GAL4 ERM protein to activate transcription from GAL4 sites could be due to the presence in the chimera of two domains showing specific DNA-binding activities. The presence of the functional ERM ETS domain could directly tether chimeras to endogenous Ets binding sites, thus decreasing the amount of fusion proteins bound to the reporter GAL4 sites and, consequently, the measured transactivation. Mutation of Y419 into proline was consequently introduced into the GAL4 ERM chimera, yielding the GAL4 ERM Y419P chimera where the mutation of Y419 into proline totally abolished the DNA-binding ability of ERM (data not shown). GAL4 ERM-Db protein where the whole DNA-binding domain has been deleted and GAL4 ERM Y419P were both used to determine the importance of the ETS domain in transactivation by the GAL4 ERM chimeras (Figure 7A). GAL4 ERM Y419P still failed to activate transcription from the 5 × GAL4-tk-CAT reporter (Figure 7B). Thus, the sole presence of the ETS binding domain, functional or not, is sufficient to inhibit transactivation. In the absence of its recognition sequence, the DNA-binding domain could adopt a conformation which prevents ERM from transactivating. Consistent with this hypothesis, GAL4 ERM-Db, lacking the entire ETS domain was able to

transactivate from the 5 × GAL4-tk-CAT reporter. We conclude that the ETS domain of ERM, functional or not, prevents transcriptional activation by the entire protein. Perhaps an intramolecular interaction involving the 203–290 region could play a role in this inhibition phenomenon. Alternatively, it could be hypothesized that the interaction between the ETS domain and a cofactor inhibits transactivation by ERM.

GAL4 ERM-DbCt, lacking the ETS domain and the carboxy-terminal tail was still able to transactivate to a level comparable to that of GAL4 ERM-Db. However, GAL4 ERM-αDb failed to transactivate. These data suggest that the first 72 amino acids and not the carboxy-terminal tail are necessary for transactivation by ERM on the GAL4 reporter. Alternatively, the action of Ct is somehow impeded in the GAL4 ERM-Db chimera.

VP16 squelches factors necessary for the action of α, but not for that of Ct

Sequence comparison, shown in Figure 8A, has revealed some sequence identity between part of α (which encompasses the aD acidic domain), and one of the transcription activation regions of the Herpes virus VP16 protein (Cress and Triezenberg, 1991). To

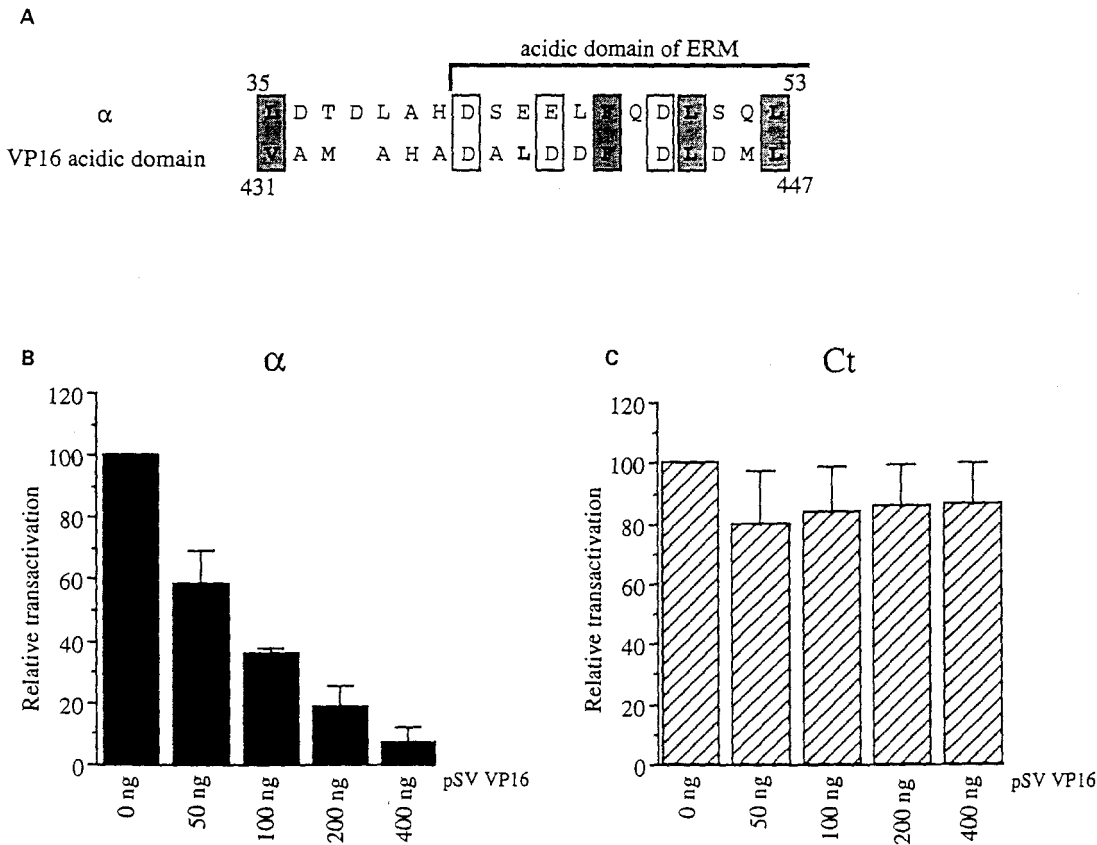


Figure 8 Expression of the VP16 transactivation domain squelches transactivation by α, but not by the carboxy-terminal tail of ERM. (A) Sequence alignment between an α subregion (ERM amino acids 35 to 53) and the transactivation domain of VP16 (amino acids 431 to 447). The boxed amino acids correspond to the conservative alignment. Grey boxes correspond to bulky hydrophobic residues important for VP16 function (11). (B) 200 ng of GAL4 α were cotransfected with increasing amounts of pSV VP16 vector (50 ng to 400 ng) which contains the 76 carboxy-terminal amino acids of the VP16 transactivation domain (kindly provided by G Adelmant). The transfection experiments were performed in RK13 cells as described in the legend to Figure 2. The transactivation activities are calculated relatively to that of GAL4 α alone (100%). Cotransfections of GAL4 α with pSV vector were performed as a negative control (data not shown). (C) Cotransfections of GAL4 Ct with pSV VP16 vector were performed and activities are calculated as described in (B) except that the transactivation activities were calculated relatively to that of GAL4 Ct alone (100%)

address the possibility that this sequence resemblance reflected similar activating mechanisms, we tested whether the coexpression of VP16 transactivation domain, by decreasing the availability of common cofactors, could impede the activating effect of α (a phenomenon called squelching). As shown in Figure 8B, expression of the VP16 transactivation domain did significantly decrease the effect of GAL4 α on transcription. In fact, when 400 ng of VP16 plasmid are cotransfected with 200 ng of GAL4 α , the transactivation induced by GAL4 α alone is reduced by 93% (Figure 8). Moreover, this effect is specific of α mediated transactivation since no significant inhibition by VP16 was observed on Ct mediated transactivation. We thus demonstrate that, as could be proposed based on sequence similarity, α and the viral protein VP16 interact with common limiting cofactors for transactivation.

Discussion

In this report, we have attributed DNA-binding inhibitory and transactivating functions to various parts of ERM, the first member of the PEA3 group to be functionally dissected. The DNA-binding activity of the ETS domain is decreased by a flanking region localized between residues 203 and 290. In addition, we

demonstrated that the carboxy-terminal tail exerts an inhibitory effect on DNA-binding and that this inhibition can be relieved by partial deletion of this domain i.e. the last 42 residues of ERM. It is noteworthy that this deletion corresponds to the removal of the less conserved part of the carboxy-terminal tail. This inhibition function also exists in PEA3 (data not shown), and since it seems to rely on the less conserved part of the carboxy-terminal tail, a differential inhibition mechanism between ERM, PEA3 and ER81 could be possible. This intramolecular inhibition of the DNA-binding ability is reminiscent of that observed in other Ets family members such as Net (Giovane *et al.*, 1994; Lopez *et al.*, 1994), as well as c-Ets-1 and c-Ets-2 (Hagman and Grosschedl, 1992; Lim *et al.*, 1992; Wasyluk *et al.*, 1992) where both the carboxy-terminal tail and a central region of the protein negatively influence the DNA-binding activity of the ETS domain. However, no sequence identity could be detected between these regions and their analogues in ERM. Our results also show that the concomitant presence of both domains is not required for DNA-binding inhibition and that each domain acts separately, suggesting that these inhibitory domains, if they act through an intramolecular mechanism, form complexes with other regions of the protein (probably the ETS domain) rather than between them. ERM seems to be an example of the widespread mechanism

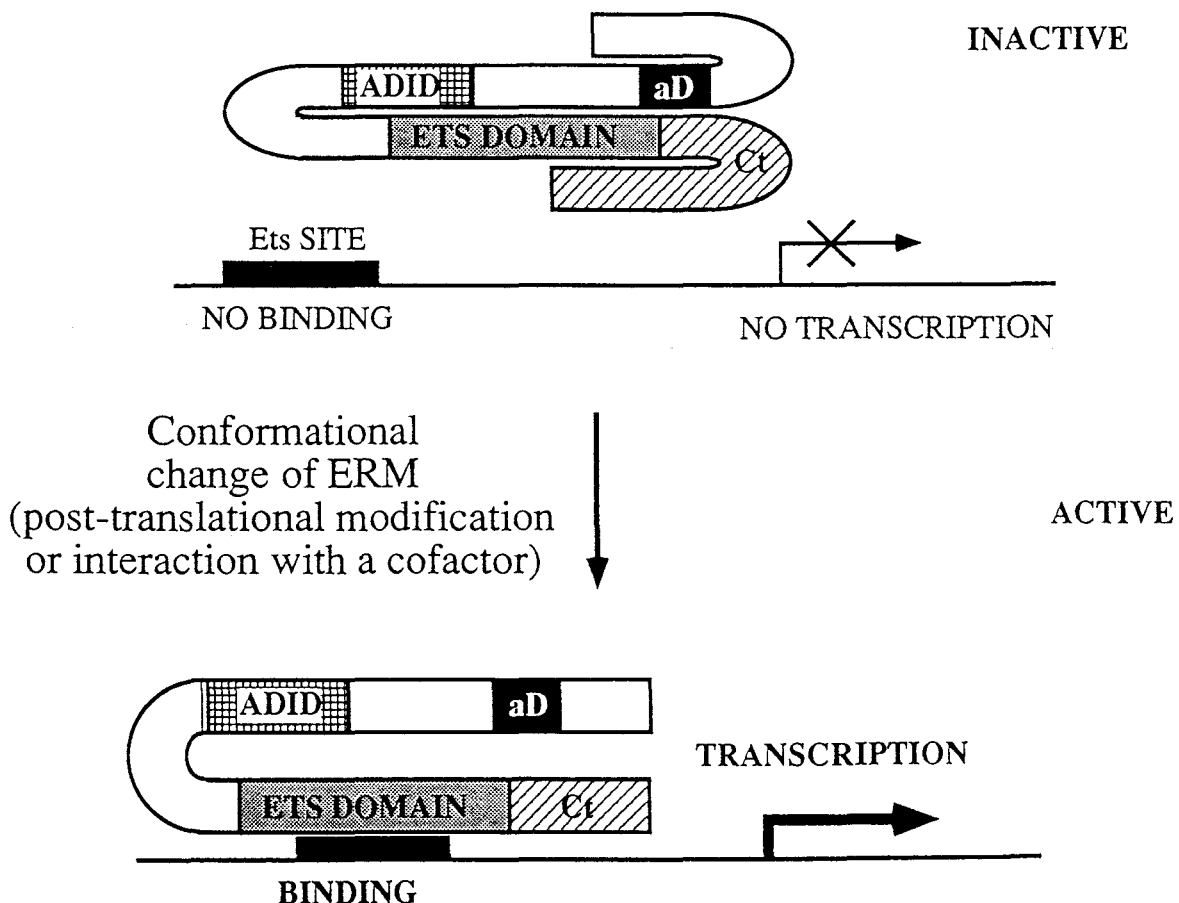


Figure 9 A schematic model for ERM intramolecular regulation of the DNA-binding activity and synergistic transactivation by the amino-terminal (α) and the carboxy-terminal (Ct) domains. Black, grey and hatched boxes represent the acidic domain, the ETS domain and Ct, respectively. Checkered box corresponds to the Amino-terminal DNA-binding Inhibitory Domain encompassing amino acids 203 to 290 (ADID)

transactivation regulation through masking the DNA-binding domain and impeding DNA-protein interactions (Figure 9) (Kowentz-Leutz *et al.*, 1994). This effect could be relieved by post-translational modification of the transcription factor or by interaction with cofactors required for the transactivation of a specific gene.

We delineated two transactivation domains in ERM. One lies in the first 72 amino acids (α) containing the so-called acidic domain and could be classified as an 'acidic' transactivation domain. The second one, Ct, correspond to the carboxy-terminal tail and does not seem to share any feature of classical transactivation domains. Ct contains 21% acidic amino acids, 13% leucine and 10% proline, alanine and serine, thus corresponding to a 'hydrophobic-acidic' domain. Possession of two transactivation domains, as in ERM, is a common property among Ets factors outside the PEA3 group; such is the case for c-Ets-1 and c-Ets-2 (Schneikert *et al.*, 1992; Chumakov *et al.*, 1993; Albagli *et al.*, 1994), Erg (Siddique *et al.*, 1993), Eli-1 (Rao *et al.*, 1993) and Elk-1 (Bhattacharya *et al.*, 1993). For the latter two Ets proteins, both transactivation domains are also devoid of any apparent sequence identity with known transactivating domains.

We tried to precise the boundaries of the two ERM transactivation domains but further deletion led to an important decrease of transactivation, indicating that we had delineated the minimal regions sufficient for maximal transactivation. It could also be hypothesized that the acidic domain *stricto sensu* (α D), or any of the Ct deletion constructs tested here, when fused to the GAL4 DNA-binding domain, are not in an appropriate conformation to interact with the transcriptional machinery, and that a linker between the GAL4 domain and the ERM fragment would increase its efficiency.

Except for the α and Ct domains, no other ERM fragment was able to transactivate. Neither β (amino acids 73 to 142) nor γ (amino acids 143 to 198) did modulate transcription by itself, nevertheless we noticed that, in HeLa cells, $\alpha\beta\gamma$ (amino acids 1 to 8) was a much stronger activator than α , indicating that the $\beta\gamma$ fragment could contain a cryptic transactivation domain active only in HeLa cells or positively modulate α induced transactivation by recruiting a specific cofactor found in HeLa but not RK13 cells. $\beta\gamma$ could possibly play the same role as region II of c-Ets-1, which is not able to transactivate but increases transactivation by region I and negatively affects transactivation by region I (Schneikert *et al.*, 1992).

As demonstrated here, α and Ct are intrinsic transactivation domains and not two parts of a sole transactivation domain. They are able to synergize, indicating that they probably act through different mechanisms involving specific cofactors or specific domains of the same cofactor.

Deletion constructs fused to that heterologous domain failed to transactivate efficiently but deletion of the DNA-binding domain resulted in a major transactivation increase (Figure 7). This phenomenon has been already described for other Ets proteins (Chumakov *et al.*, 1993; Janknecht *et al.*, 1994) and for transcription factors (Weintraub *et al.*, 1991;

Williams and Tjian, 1991; Hill *et al.*, 1994). Neither GAL4 ERM nor GAL4 ERM Y419P, the DNA-binding capacity of the latter being fully abrogated, did transactivate efficiently. Thus, if the structural change induced by mutation of Y419 into proline has a drastic effect on the DNA-binding ability of ERM, it has no influence on transactivation; Y419, which does not belong to a transactivation domain, is obviously not directly involved in the transactivation mechanism. When the ETS domain is not bound to its sequence target or when a mutation totally abrogates its DNA-binding capacity, it seems to act as a negative regulatory element of transactivation. Indeed, for other transcription factors like SRF, the model established for DNA-binding requirement in transcriptional activation (Johansen and Prywes, 1993) involves sequence specific DNA-binding as an allosteric regulation of the activity of the transactivating domain. DNA-binding mutations with altered specificity or absence of the target sequence (as for the GAL4 chimeras), modify the interactions between the three partners, and thus impede transactivation.

While deletion of α seems to exert a drastic effect on transactivation, the construct lacking the carboxy-terminal tail still transactivates quite efficiently (Figure 7). This situation is reminiscent from that of the amino-terminal BEC containing activation domain which, in p68^{c-ets-1}, is supposed to work in a strictly context dependent manner (Albagli *et al.*, 1994). Other transcription factors contain two autonomous transactivation domains, only one being indispensable for transactivation by the wild-type protein on certain responsive elements (Nagpal *et al.*, 1993). Alternatively, Ct could be inhibited in the GAL4 ERM-Db protein.

α and Ct exhibit functional differences. Indeed, they have different cell type dependent behaviours: the Ct/ α activity ratio varies with the cell line tested (37.5% and 300.0% in RK13 and HeLa cells, respectively). The strongest support for a functional difference is that a common limiting cofactor for both VP16 and the α transactivation domain was evidenced by squelching experiments. These data, together with the fact that α and the VP16 transactivation domain show some sequence identity, indicate that these two transactivating domains could be functionally related, suggesting a different transactivation mechanism for α and Ct. It could also be hypothesized that VP16, α and Ct need a common cofactor. Different from that shared only by VP16 and α , this cofactor could be in non limiting concentration and consequently not detectable by the competition assay. Factors of the general transcription machinery such as TFIIB (Lin and Green, 1991), two subunits of TFIID, TBP (Stringer *et al.*, 1990), TAF_{II}31 (Klemm *et al.*, 1995; Lu and Levine, 1995) and TAF_{II}40 (Goodrich *et al.*, 1993), TFIIF (Xiao *et al.*, 1994) as well as an adaptator named ADA2 (Berger *et al.*, 1992) have been shown to directly interact with VP16. These cofactors and coactivators could be candidates for an interaction with α , an hypothesis we are currently testing.

In this report, we performed experiments demonstrating that the 72 amino-terminal residues play a role only in transactivation, whereas the carboxy-terminal tail and specially the last 45 residues of ERM, seem to exert both a negative effect of DNA-binding ability and a positive effect on transactivation. This situation is

reminiscent of another Ets family member, *c-ets-1* for which the VII/d exon (residues 243 to 330) bears both DNA-binding inhibition (residues 207 to 280) and transactivation functions (Wasylyk et al., 1992). This suggests that, even if they do not share any sequence identity outside the ETS domain, some Ets transcription factors could act through similar mechanisms. Based on the above results, we are able to propose a model for the binding and transactivation activities of ERM (Figure 9). Two regions, the carboxy-tail and the 203–290 fragment (termed ADID for Amino-terminal DNA-binding Inhibitory Domain) are responsible for the decrease of the DNA-binding ability. Post-translational modification of ERM or interference of a cofactor could induce conformational changes of the protein, thus releasing DNA-binding inhibition and allowing efficient transactivation by ERM through a synergy between α and Ct. Transcription by ERM may rely on a mechanism similar to that of C/EBP β for which transcription and inactivation of inhibitory functions are contemporaneous events (Kowentz-Leutz et al., 1994; Williams et al., 1995). Some Ets family members have been demonstrated to require outer factors for their DNA binding or transactivation abilities. Indeed, c-Ets-1 and CBF form higher affinity complexes with their targets on the human T-cell receptor gene enhancer than either protein alone (Wotton et al., 1994). Similarly, cooperation of c-Ets-1 and Sp1 for transactivation, strictly depending on the integrity of their respective binding sites, has been demonstrated (Gégonne et al., 1993). Elk-1 (Hipskind et al., 1991; Janknecht and Nordheim, 1992; Shore and Sharrocks, 1994), SAP-1 (Dalton and Treisman, 1992) and SAP-2 directly interact with SRF (Giovane et al., 1994), GABP α with GABP β (LaMarco et al., 1991; Thompson et al., 1991) and PU-1 with NF-EM5 (Pongubala et al., 1992).

The localization of the functional domains of ERM will provide us with new clues to its transactivation mechanism. Further studies of the ERM intramolecular interactions and interactions between the different domains with potential cofactors should help determine the physiological role of ERM.

Materials and methods

Construction of ERM truncated protein expression vectors

N-terminal truncated proteins were generated by PCR with upstream oligonucleotides overlapping original or artificially designed internal methionines of ERM and the following downstream oligonucleotide 5'-ACCAAAC-CACTGCCCTTGTGGC-3'. Indeed, M203 (oligonucleotide 5'-CAGCCCCTGGCCATGCCAAAGATGA-3'), M290 (oligonucleotide 5'-TCACCAATGGGAATCAAG-CAGGAGC-3') and M355 (oligonucleotide 5'-TACCATG-TATCGAGAGGGGCCCT-3') gave rise to the so called B, C and D truncated ERM proteins (Figure 1A). The translation product of the full length ERM sequence gave rise to A. Using the same upstream oligonucleotides and a downstream oligonucleotide situated at the end of the ETS domain (5'-ATCTGGTTATCAGCAGACAAATTTGTA-G-3'), we generated ΔA , ΔB , ΔC and ΔD constructs, ΔD corresponding to the ETS domain. $\Delta D+$, corresponding to the ETS domain followed by the first half of the carboxy-terminal domain was also amplified by PCR using the following downstream oligonucleotide 5'-TAGGATCCT-

TACTTCAGGAACGGACGCTGG-3'. All PCR products were cloned into the pCRII vector, entirely sequenced and then cloned into pSG5 using the *EcoRI* restriction sites of the pCRII vector.

The translation product of the full length ERM sequence was also introduced into the pSG5 derived vector pSV (kindly provided by Guillaume Adelmant) using the *EcoRI* and *XbaI* sites.

GAL4 fusion chimeras

All the following GAL4 chimeras α , β , γ , δ , ϵ , Db, Db*, Ct, aD, Ct 449–468, Ct 449–489, Ct 469–510, $\alpha\beta\gamma$, $\gamma\delta$ and $\beta\gamma\delta\epsilon$ were constructed by cloning various PCR products (oligonucleotides are listed in Table 1), downstream from and in frame with the GAL4 DNA-binding domain (1–147) by using the pSG424 vector (Sadowski and Ptashne, 1989) kindly provided by Dr Achim Leutz. All the PCR products were first subcloned into the pCRII vector, entirely sequenced and cloned into *Asp718*–*XbaI* digested pSG424 vector. For the GAL4 ERM chimera, we first amplified a fragment corresponding to nucleotides 196 to 1370, subcloned it in the *Asp718*–*SmaI* digested pCRII vector containing the full length ERM coding region. The obtained construct was then cloned into the pSG424 vector. GAL4 ERM- α construct, where the first 72 residues are removed, was obtained using the same strategy. *AccI*–*XbaI* digestion followed by blunting of the GAL4 ERM and GAL4 ERM- α chimeras gave rise to GAL4 ERM-Ct (where the last 61 residues are removed) and GAL4 ERM- α Ct (lacking the first 72 and the last 61 residues), respectively. *NdeI*–*AccI* digestion and blunting of GAL4 ERM, GAL4 ERM- α and GAL4 ERM-Ct gave rise to GAL4 ERM-Db, GAL4 ERM- α Db and GAL4 ERM-DbCt, respectively, where the ETS domain is removed. A *NdeI*–*AccI* mutated fragment (Y419 into proline) was generated by double PCR as described above (see Table 1 for oligonucleotides) using the internal oligonucleotide 5'-GCCATGAACCCCTGACAAGCTG-3' and its antisense counterpart, and cloned into the GAL4 ERM chimera in order to obtain GAL4 ERM Y419P.

Construction of mutated ERM ETS domain expression vectors

We generated by PCR a fragment encoding 82 amino acids of ERM ETS domain using the upstream primer 5'-ATGAATTCAATGGGTTCCCTTCAGCTGTGG-3' and the downstream primer 5'-TAGGATCCTTATTAATAA-TTTGTAGACGTATCG-3'. The PCR product was subcloned into the pCRII vector (Invitrogen) and used as matrix for synthesis of ERM ETS domain mutants. Y419 mutants were generated by double PCR method: oligonucleotide 5'-GCCATGAACNNNGACAAGCTG-3' for which NNN correspond to GCT, CCT, TTC and TGG and its antisense counterpart, generated a specific mutation of Y419 into alanine, proline, phenylalanine and tryptophan, respectively. Y443 mutants were synthesized using an upstream primer 5'-TAGGATCCTTATTAATAAATTTGTA-GACNNNTCG-3' where NNN correspond to AGC, GGG, GAA and CCA, generating a specific mutation of Y443 into alanine, proline, phenylalanine and tryptophan, respectively. The double mutant Y419W/Y443W was constructed by double PCR method using the mutant Y419W as matrix. Each construct was then entirely sequenced and subcloned into the pSG5 vector (Green et al., 1988) using the *EcoRI* and *BamHI* restriction sites of the primers.

In vitro translation and protein analysis

One μ g of each pSG5 plasmid was used for transcription and translation of ERM constructs **120** 10 units of T7

Table 1 Oligonucleotides primers used for PCR amplification

Construction	Upstream primer (5' → 3')	Downstream primer (5' → 3')
AL4 α	TGGTACCCCATGGACGGGTTTTATGAT	TTCTAGATTAATCTGGGACAAACTGTTCA
AL4 β	TGGTACCATTTCAGTCTGATAACCTGGT	TTCTAGATTATGAGAGGGGGTTGTAGGA
AL4 γ	TGGTACCCACCCACCCATCAGAATCCCT	TTCTAGATTAATGTGGTGGTGGGGGACC
LA4 δ	TGGTACCATCAGCCCCGACAGTGC AAA	TTCTAGATTATGCTGGGGCCCTGGCATG
AL4 ε	TGGTACCCACACGGGTTCCAGTCACCAAT	TTCTAGATTACTGGTAAGGGGGCCCTCT
AL4 Db	TGGTACCAGAGGCGAGGTTCCCTTCAGCT	TTCTAGATCAGACAAATTTGTAGACGTAT
AL4 Db*	TGGTACCAGAGGCGAGGTTCCCTTCAGCT	TTCTAGATCAGACAAATTTGTAGACGTAT
AL4 Ct	TGGTACCTCTGTGACCCAGATGCCCTCT	TTCTAGATTAGTAAGCAAAGCCTTCGGCA
AL4 aD	TGGTACCACGATTTCTGAAGAGCTATTTTCAG	TTCTAGATTAATCTGGGACAAACTGTTCA
AL4 Ct 449-468	TGGTACCTCTGTGACCCAGATGCCCTCT	TTCTAGATTACTTCAGGAACGGACGCTGG
AL4 Ct 469-510	TGGTACCACGAGAGTCCGAGTGCACCT	TTCTAGATTAGTAAGCAAAGCCTTCGGCA
AL4 Ct 449-489	TGGTACCTCTGTGACCCAGATGCCCTCT	TTCTAGATTAGTGTCTTCAAAAGTGGT
AL4 αβγ	TGGTACCCCATGGACGGGTTTTATGAT	TTCTAGATTAATGTGGTGGTGGGGGACC
AL4 γδ	TGGTACCCACCCACCCATCAGAATCCCT	TTCTAGATTATGCTGGGGCCCTGGCATG
AL4 βγδε	TGGTACCATTTCAGTCTGATAACCTGGT	TTCTAGATTACTGGTAAGGGGGCCCTCT
AL4 ERM	TGGTACCCCATGGACGGGTTTTATGAT	ATTCCCATTTGGTGACTGGAAC
AL4 ERM-α	TGGTACCATTTCAGTCTGATAACCTGGT	ATTCCCATTTGGTGACTGGAAC
del - AccI fragment	ATCATATGAAAAAGATCCCCCGA	TAGGATCCTTATTAATAATTTGTAGACGTATCG

NA polymerase, 25 μl Rabbit reticulocyte lysate and 0.5 μCi [³⁵S]methionine in the appropriate buffer as specified by the manufacturer (Promega) for 1.5 h at 30°C. Translation of the truncated proteins was verified on a 10-30% gradient SDS-PAGE and the use of a Phosphorimager and ImageQuant software analysis. (Molecular Dynamics) allowed equal protein loading for electrophoretic mobility-shift assays. Wild-type and mutated ETS domains were visualized on a tricine-SDS-PAGE as described (Schägger and von Jagow, 1987).

Electrophoretic mobility-shift assays

Proteins were incubated for 1 h on ice in a final volume of 20 μl with 20 mM HEPES pH 7.9, 20% glycerol, 0.1 mM DTT, 1 mM DDT, 1 μg/20 μl polyIdC, 50 mM NaCl and excess probe rendered double-stranded and containing part of the sequence of the Polyomavirus enhancer 5'-ATCTTCGAAAGGAAGTTCGAG-3'. The samples were then loaded on a 5-20% gradient polyacrylamide (29/1) gel in 0.2 × TBE and run overnight at 150 V (4°C).

Transfection assays

Human cervix carcinoma HeLa and rabbit kidney carcinoma RK13 cell cultures were routinely maintained in Dulbecco's medium with 10% fetal calf serum, at 37°C in H₂O saturated, 5% CO₂ atmosphere. Three to 5 × 10⁵ cells were transfected in 35 mm diameter dishes using 10 μl of polyfectamin (Gibco-BRL) with 0.8 μg expression vector, 1 μg pSV2luc vector, 0.3 μg of the reporter plasmid pGAL4-tk-CAT or 0.1 μg of the 3 × TORU-tk-CAT reporter (both kindly provided by Dr K. MacLeod).

References

Albagli O, Soudant N, Ferreira E, Dhordain P, Dewitte F, Begue A, Flourens A, Stehelin D and Leprince D. (1994). *Oncogene*, **9**, 3259-3271.
 An D, David Y, Giddens EB, Letwin K and Bernstein A. (1991). *Genes Dev.*, **5**, 908-918.
 Berger SL, Pina B, Silverman N, Marcus GA, Agapite J, Regier JL, Triezenberg SJ and Guarente L. (1992). *Cell*, **70**, 251-265.
 Chattacharya G, Lee L, Reddy ES and Rao VN. (1993). *Oncogene*, **8**, 3459-3464.
 Cress WD, Duvall JF, Gégonne A, Bailly M, Hémard A, Brady J and Ghysdael J. (1990). *EMBO J.* **9**, 3137-3144.

TORU (for TPA Oncogene Responsive Unit) is a Polyomavirus enhancer derived sequence 5'-GATCTC-GAGGAAGTGACTAACTTGAGCA-3' (Wasylyk et al., 1990). In order to take into account transfection efficiency variations, all assays were normalized to the luciferase activity of the pSV2luc vector. Cells were harvested 24 h after transfection in a reporter lysis buffer (Promega). Luciferase and CAT activities were determined as previously described (Crépieux et al., 1993) and the ratio of acetylated/non acetylated [¹⁴C]chloramphenicol quantified using a Phosphorimager and Image Quant software analysis (Molecular Dynamics).

Protein expression of each GAL4 chimera construct was verified by immuno-precipitation as described (Albagli et al., 1993) and Western blotting using a GAL4 antibody (Santa Cruz Biotechnology).

Acknowledgements

The authors insist on the fact that PAD and MPL equally contributed to this work. We are indebted to Guillaume Adelmant for considerable involvement into this work and for kindly providing pSV and pSV VP16 constructs. We thank Ralf Janknecht for RK13 cells, Anne Lelièvre-Chotteau for critical reading of this manuscript. PAD, MPL, OA and FD are supported by the Ministère de la Recherche et de l'Espace, the Association de Recherche sur le Cancer (ARC), the 'CNRS-région' and the Institut Pasteur de Lille, respectively. This work was done according to grants awarded in part by the CNRS, the ARC and the 'Ligue contre le cancer' France.

Brown TA and McKnight SL. (1992). *Genes and Dev.*, **6**, 2502-2512.
 Chen ZQ, Burdett LA, Seth AK, Lautenberger JA and Papas TS. (1992). *Science*, **250**, 1416-1418.
 Chumakov AM, Chen DL, Chumakova EA and Koeffler HP. (1993). *J. Virol.*, **67**, 2421-2425.
 Crépieux P, Coll J and Stehelin D. (1994). *Crit. Rev. Oncology*, **5**, 615-638.
 Crépieux P, Leprince D, Flourens A, Albagli O, Ferreira E and Stehelin D. (1993). *Gene Expression*, **3**, 215-225.
 Cress WD and Triezenberg SJ. (1991). *Science*, **251**, 87-90.
 Dalton S and Treisman R. (1992). *Cell*, **68**, 597-612.

- Degnan BM, Degnan SM, Nagagnuma T and Morse DE. (1993). *Nucleic Acids Res.*, **21**, 3479–3484.
- Delattre O, Zucman J, Plougastel B, Desmaze C, Melot T, Peter M, Kovar H, Joubert I, de Jong P, Rouleau G, Aurias A and Thomas G. (1992). *Nature*, **359**, 162–165.
- Donaldson LW, Petersen JM, Graves BJ and McIntosh LP. (1994). *Biochemistry*, **33**, 13509–13516.
- Gegonne A, Bosselut R, Bailly RA and Ghysdael J. (1993). *EMBO J.*, **12**, 1169–1178.
- Gehring WJ, Qian YQ, Billeter M, Furukubo-Tokunaga K, Schier AF, Resendez-Perez D, Affolter M, Otting G and Wüthrich K. (1994). *Cell*, **78**, 211–223.
- Giovane A, Pintzas A, Maira SM, Sobieszczuk P and Wasylyk B. (1994). *Genes Dev.*, **8**, 1502–1513.
- Golub T, Barker GF, Lovett M and Gilliland DG. (1994). *Cell*, **77**, 307–316.
- Goodrich JA, Hoey T, Thut CJ, Admon A and Tjian R. (1993). *Cell*, **75**, 519–530.
- Green S, Issemann I and Sheer E. (1988). *Nucleic Acids Res.*, **16**, 369.
- Hagman J and Grosschedl R. (1992). *Proc. Natl. Acad. Sci. USA*, **89**, 8889–8893.
- Higashino F, Yoshida K, Noumi T, Seiki M and Fujinaga K. (1995). *Oncogene*, **10**, 1461–1463.
- Hill CS, Wynne J and Treisman R. (1994). *EMBO J.*, **13**, 5421–5432.
- Hipskind RA, Rao VN, Mueller CGF, Reddy ESP and Nordheim A. (1991). *Nature*, **354**, 531–534.
- Janknecht R and Nordheim A. (1992). *Nucleic Acids Res.*, **20**, 3317–3324.
- Janknecht R, Zinck R, Ernst WH and Nordheim A. (1994). *Oncogene*, **9**, 1273–1278.
- Jeon I-S, Davis JN, Braun BS, Sublett JE, Roussel MF, Denny CT and Shapiro DN. (1995). *Oncogene*, **10**, 1229–1234.
- Kissinger CR, Liu B, Martin-Blanco E, Kornberg TB and Pabo CO. (1990). *Cell*, **63**, 579–590.
- Kowentz-Leutz E, Twamley G, Ansieau S and Leutz A. (1994). *Genes Dev.*, **8**, 2781.
- LaMarco K, Thompson CC, Byers BP, Walton EM and McKnight SL. (1991). *Science*, **253**, 789–791.
- Laudet V, Niel C, Duterque M, Leprince D and Stéhelin D. (1993). *Biochem. Biophys. Res. Com.*, **190**, 8–14.
- Lautenberger JA, Burdett LA, Gunnell MA, Shengmei Q, Watson DK, O'Brien SJ and Papas TS. (1992). *Oncogene*, **7**, 1713–1719.
- Liang H, Mao X, Olejniczak ET, Nettesheim DG, Yu L, Meadows RP, Thompson CB and Fesik SW. (1994a). *Nature Struct. Biol.*, **1**, 871–876.
- Liang H, Olejniczak ET, Mao X, Nettesheim DG, Yu L, Thompson CB and Fesik SW. (1994b). *Proc. Natl. Acad. Sci. USA*, **91**, 11655–11659.
- Lim F, Kraut N, Frampton J and Graf T. (1992). *EMBO J.*, **11**, 643–652.
- Lin YS and Green MR. (1991). *Cell*, **64**, 971–981.
- Metz T and Graf T. (1991). *Genes Dev.*, **5**, 369–380.
- Monté D, Baert JL, Defossez PA, de Launoit Y and Stéhelin D. (1994). *Oncogene*, **9**, 1397–1406.
- Monté D, Coutte L, Baert JL, Angeli I, Stéhelin D and de Launoit Y. (1995). *Oncogene*, **11**, 771–780.
- Moreau GF, Tavitian A and Tambourin P. (1988). *Nature*, **331**, 277–280.
- Naggal S, Friant S, Nakshatri H and Chambon P. (1993). *EMBO J.*, **12**, 2349–2360.
- O'Neill EM, Rebay I, Tjian R and Rubin GM. (1994). *Cell*, **78**, 137–147.
- Pabo CO and Sauer RT. (1992). *Annu. Rev. Biochem.*, **61**, 1053–1095.
- Pongubala JMR, Nagulapalli S, Klemsz MJ, McKercher SR, Maki RA and Atchinson ML. (1992). *Mol. Cell. Biol.*, **12**, 368–378.
- Qian YQ, Billeter M, Otting G, Müller M, Gehring WJ and Wüthrich K. (1989). *Cell*, **59**, 573–580.
- Rao VN, Ohno T, Prasad DDK, Batthacharya G and Reddy ES. (1993). *Oncogene*, **8**, 2167–2173.
- Romano-Spica V, Suzuki H, Georgiou P, Chen S-L, Ascione R, Papas TS and Bhat NK. (1994). *Int. J. Oncol.*, **4**, 521–531.
- Sadowski I and Ptashne M. (1989). *Nucleic Acids Res.*, **17**, 7539.
- Schägger H and Von Jagow G. (1987). *Anal. Biochem.*, **166**, 368–379.
- Schneikert J, Lutz Y and Wasylyk B. (1992). *Oncogene*, **7**, 249–256.
- Scott EW, Simon MC, Anastasi J and Singh H. (1994). *Science*, **265**, 1573–1577.
- Shimizu K, Ichikawa H, Tojo A, Kaneko Y, Maseki N, Hayashi Y, Ohira M, Asano S and Ohki M. (1993). *Proc. Natl. Acad. Sci. USA*, **90**, 10280–10284.
- Shore P, Bisset L, Lakey J, Waltho JP, Virden R and Sharrocks AD. (1995). *J. Biol. Chem.*, **270**, 5805–5811.
- Shore P and Sharrocks AD. (1994). *Mol. Cell. Biol.*, **14**, 3283–3291.
- Siddique HR, Rao VN, Lee L and Reddy ESP. (1993). *Oncogene*, **8**, 1751–1755.
- Soudant N, Albagli O, Dhordain P, Flourens A, Stéhelin D and Leprince D. (1994). *Nucleic Acids Res.*, **22**, 3871–3879.
- Stringer KF, Ingles CJ and Greenblatt J. (1990). *Nature*, **345**, 783–786.
- Thompson CC, Brown TA and McKnight SL. (1991). *Science*, **253**, 762–767.
- Thompson CC and McKnight SL. (1992). *Trends Genet.*, **8**, 232–236.
- Trimble MS, Xin JH, Guy CT, Muller WJ and Hassell JA. (1993). *Oncogene*, **8**, 3037–3042.
- Wasylyk B, Wasylyk C, Flores P, Bègue A, Leprince D and Stéhelin D. (1990). *Nature*, **346**, 191–193.
- Wasylyk C, Kerckaert JP and Wasylyk B. (1992). *Genes Dev.*, **6**, 965–974.
- Wasylyk C and Wasylyk B. (1993). *Nucleic Acids Res.*, **21**, 523–529.
- Weintraub H, Dwarki VJ, Verma I, Davis R, Hollenberg S, Snider L, Lassar A and Tapscott SJ. (1991). *Genes Dev.*, **5**, 1377–1386.
- Williams SC, Baer M, Dillner AJ and Johnson PF. (1995). *EMBO J.*, **14**, 3170–3183.
- Williams T and Tjian R. (1991). *Genes Dev.*, **5**, 670–682.
- Wolberger C, Vershon AK, Liu B, Johnson AD and Pabo CO. (1991). *Cell*, **67**, 517–528.
- Wotton D, Ghysdael J, Wang S, Speck NA and Owen MJ. (1994). *Mol. Cell. Biol.*, **14**, 840–850.
- Xiao H, Pearson A, Coulombe B, Truant R, Zhang S, Regier JL, Triezenberg SJ, Reinberg D, Flores O, Ingles CJ and Greenblatt J. (1994). *Mol. Cell. Biol.*, **14**, 7013–7024.
- Xin JH, Cowie A, Lachance P and Hassell JA. (1992). *Genes and Dev.*, **6**, 481–496.
- Zucman J, Melot T, Desmaze C, Ghysdael J, Plougastel B, Peter M, Zucker JM, Triche TJ, Sheer D, Turc-Carel C, Ambros P, Combaret V, Lenoir G, Aurias A, Thomas G and Delattre O. (1993). *EMBO J.*, **12**, 4481–4487.

Discussion

Un point sémantique : lors de la réalisation de ces expériences, les localisations des domaines activateurs d'Erm n'était pas connue. Nous avons désigné les portions N-terminales d'Erm par des lettres grecques, de α à ϵ , le domaine de liaison à l'ADN par DBD et la queue C-terminale par Ct. Nous avons montré que α et Ct contenaient chacun un domaine transactivateur, notés respectivement AD1 et AD2, par simplicité. Il y a donc équivalence entre les termes α et **AD1**; et **Ct** et **AD2**.

1. L'inhibition intramoléculaire de l'activité d'Erm

L'objectif de ce travail était d'identifier les régions d'Erm impliquées dans ses fonctions de liaison à l'ADN et de modulation transcriptionnelle. Nous avons donc été amenés à pratiquer une dissection de la protéine pour considérer indépendamment ses domaines. Tentons maintenant d'être un peu moins réductionnistes, et d'offrir une vue de la situation des domaines fonctionnels d'Erm au sein de la molécule native.

1.1 Inhibition de la liaison à l'ADN

Nos résultats montrent que deux régions d'Erm inhibent sa fixation à l'ADN. La première de ces régions est située entre les AA 203 et 290 et correspond à peu près à la région δ , l'autre réside dans la queue C-terminale. Notons que nos tests ne sont pas quantitatifs et ne nous permettent pas d'établir, par exemple, si l'inhibition est plus intense quand les deux régions sont présentes plutôt qu'une seule.

Nos expériences ayant été réalisées *in vitro*, l'inhibition ne requiert aucun cofacteur. L'hypothèse la plus simple pour rendre compte de nos constatations est donc d'imaginer des contacts intramoléculaires entre les régions inhibitrices (δ et Ct), et le domaine ETS. Cette idée s'accorde avec l'observation que l'effet inhibiteur de Ct ne peut être transposé au domaine de liaison à l'ADN du récepteur nucléaire Rev-erb α (M. P. Laget et P. A. Defossez, non publié).

Nous avons étudié l'interaction des différents domaines d'Erm par la technique de double-hybride chez la levure. Nos premiers résultats montrent que δ et Ct n'interagissent pas entre eux (P. A. Defossez, non publié). Reste à savoir s'ils interagissent avec le domaine de liaison à l'ADN.

1.2 Inhibition du pouvoir transactivateur

La chimère GAL4- α augmente fortement la transcription d'un gène rapporteur dont le promoteur contient des sites GAL4. La chimère GAL4- $\alpha\beta\gamma\delta\epsilon$ est bien moins active, tandis que la chimère GAL4-Erm ne l'est pas du tout. Il existe donc des régions d'Erm qui inhibent la fonction du domaine activateur α .

On peut imaginer trois mécanismes pour cet effet :

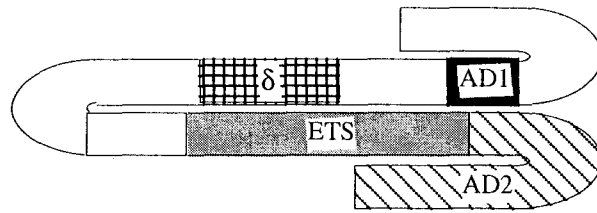
- ces régions pourraient exercer une inhibition intramoléculaire,
- elles pourraient recruter des "corépresseurs" de α ,
- enfin, elles pourraient inhiber de façon autonome la transcription.

La troisième proposition est peu probable. Nous avons coexprimé avec GAL4- α chacune des autres chimères (M. P. Laget et P. A. Defossez, non publié), et aucune d'entre elles n'a diminué l'étendue de la transactivation.

Si la deuxième hypothèse était vraie, en exprimant les domaines inhibiteurs, on pourrait capter les corépresseurs, et augmenter le pouvoir de transactivation de GAL4- $\alpha\beta\gamma\delta\epsilon$ (voir Baichwal *et al.*, 1991, pour un mécanisme similaire chez C-Jun). Nous ne l'avons pas non plus observé (M. P. Laget et P. A. Defossez, non publié).

Nous pensons donc que, comme pour le domaine de fixation à l'ADN, la fonction des domaines transactivateurs est soumise à une inhibition intramoléculaire.

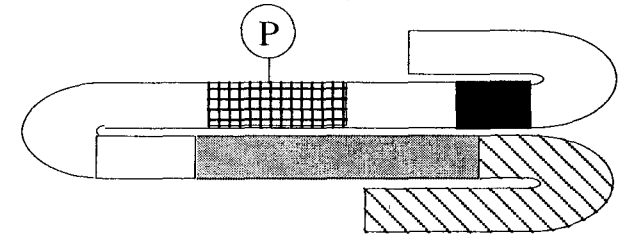
Forme "fermée"



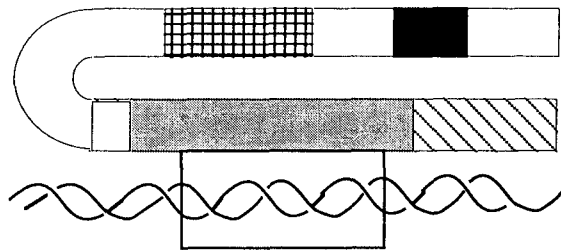
Signaux extracellulaires



Phosphorylation de δ



Forme "ouverte"



SITE ETS

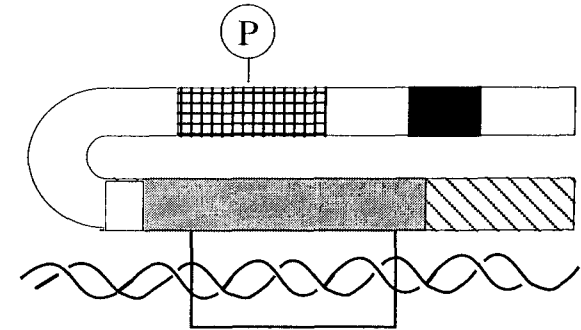


Fig. 33 Une hypothèse pour la régulation de l'activité d'ERM par phosphorylation

ERM évolue entre une conformation "fermée" et une conformation "ouverte" (à gauche). Le domaine δ est en partie responsable de l'inhibition intramoléculaire, et il est la cible de kinases. Celles-ci pourraient activer Erm en facilitant la transition fermé-ouvert.

1.3 L'activité d'Erm pourrait être régulée par modification de sa structure

Les résultats discutés ci-dessus suggèrent qu'Erm peut **transiter d'une conformation "fermée", dans laquelle il n'active pas la transcription, à une conformation "ouverte" où les domaines activateurs sont démasqués**. C'est peut-être la transition fermé-ouvert qui permet la fixation à l'ADN, mais il est aussi possible que la liaison à l'ADN "déplie" la protéine par un effet allostérique (voir fig. 33). Cette situation n'est pas sans rappeler celle observée pour Ets-1. Dans ce cas une hélice se replie sur le domaine ETS et inhibe son pouvoir de liaison à l'ADN. La reconnaissance de la séquence cible s'accompagne d'un déploiement de l'hélice (Petersen *et al.*, 1995; Jonsen *et al.*, 1996). Cette région inhibitrice de Ets-1 ne présente pas de similarité de séquence avec les domaines δ ou Ct.

Un moyen d'augmenter l'activité d'Erm est de déplacer l'équilibre (fermé)-(lié à l'ADN-ouvert) en déstabilisant la forme fermée, ou en stabilisant la forme ouverte.

La région inhibitrice δ est un bon substrat pour la kinase p38 (J. L. Baert, non publié). Il est possible que des stimuli extracellulaires entraînent la phosphorylation de δ , déstabilisent la forme fermée et augmentent l'activité d'Erm. C/EBP β est un cas connu où l'inhibition intramoléculaire est levée par phosphorylation des domaines inhibiteurs (Kowentz-Leutz *et al.*, 1994).

Il est également possible que l'activation passe par un cofacteur d'Erm stabilisant la forme ouverte ou fermée.

2. Quel mécanisme pour la synergie AD1-AD2?

La présence simultanée de AD1 et AD2 sur un promoteur augmente leur capacité à activer la transcription (fig. 6 de la publication. Nous avons répété cette expérience en fusionnant AD1 et AD2 à deux domaines de liaison à l'ADN différents et en utilisant un rapporteur portant les sites de fixation de chacun des domaines. Les résultats obtenus étaient comparables à ceux décrits dans la publication). La collaboration ne résulte pas d'une interaction physique entre les domaines (P. A. Defossez, non publié), mais doit être issue de la combinaison de leurs effets sur la transcription.

Quelles hypothèses peut-on poser pour expliquer la synergie observée?

Notons préalablement que **deux activateurs ayant le même mode d'activation peuvent avoir des effets synergiques**. Ainsi, en augmentant le nombre de modules activateurs de VP16 fusionnés au domaine de liaison à l'ADN de GAL4, on augmente fortement l'effet activateur de la chimère (Emami et Carey, 1992). On peut trouver deux interprétations non exclusives à cette observation :

- soit chaque domaine peut interagir avec plusieurs cibles. Deux domaines pourraient simultanément contacter les deux cibles et améliorer les interactions avec la machinerie de transcription.
- soit les interactions activateurs-cibles sont transitoires. En augmentant le nombre de domaines, on augmenterait statistiquement la fréquence de ces interactions (voir Herschlag et Johnson, 1993, pour une discussion plus approfondie de ces considérations cinétiques).

On ne peut donc pas conclure de la synergie AD1-AD2 que ces domaines ont des modes d'actions différents. Par contre, le fait qu'ils aient 1) un pouvoir transactivateur qui varie de façon différente dans différents types cellulaires et 2) une sensibilité différente à la surexpression du domaine activateur de VP16 laisse peu de doute sur la **dissimilarité de leur mécanisme d'activation**. Notons également qu'AD2, à la différence d'AD1, n'active pas la transcription dans des levures (P. A. Defossez, non publié).

Nous avons évoqué en introduction les multiples points activables de la transcription. **AD1 et AD2 pourraient réguler différents points, ou le même point via des mécanismes différents.** La séquence d'AD1 évoque des régulateurs de la préinitiation et de l'élongation (voir publication n°3, et sections 3.2 et 3.3 de l'introduction), mais celle d'AD2 n'est guère informative.

Des travaux non publiés de notre groupe (J. L. Baert, Y. de Launoit et R. Janknecht), montrent que la protéine CBP (CREB-Binding Protein) peut moduler le pouvoir transactivateur d'Erm. Elle peut se fixer directement à la partie N-terminale d'Erm, qui contient AD1. Par contre, elle ne reconnaît pas la partie C-terminale où se trouve AD2. On peut donc penser qu'AD1 active la transcription de façon dépendante de CBP, tandis qu'AD2 en est indépendant. Des expériences sont en cours pour tenter de découvrir ces mécanismes.

Notons que l'activité de CBP est régulable par plusieurs kinases, dont les kinases Erk et la PKA (revue dans Janknecht et Hunter, 1996). On comprend donc que l'effet des kinases Erk et PKA sur le pouvoir transactivateur d'Erm pourrait être indirect, et s'exercer en fait sur CBP.

3. Conséquences "physiologiques" des différences AD1-AD2

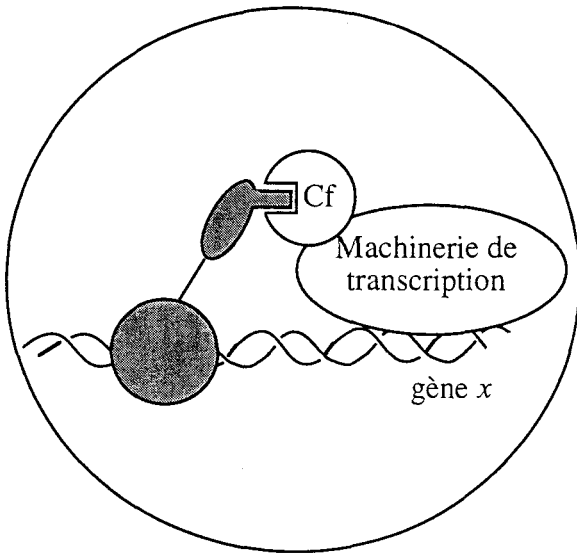
Nous souhaitons, pour conclure, montrer que les différences démontrées entre AD1 et AD2 sont plus qu'une curiosité transcriptionnelle sans conséquence.

Le pouvoir transactivateur d'AD1 et AD2 varie de façon non-coordonnée selon les types cellulaires. Nous avons également mis en évidence des effets promoteurs-spécifiques de ces deux domaines. Ainsi, dans un même type cellulaire, la chimère GAL4-AD1 est plus efficace que GAL4-AD2 pour activer la transcription d'un gène rapporteur dont le promoteur contient une TATA Box. Si on remplace cette séquence par un Initiateur, le pouvoir activateur de GAL4-AD2 devient le plus important (P. A. Defossez, non publié). (Des observations

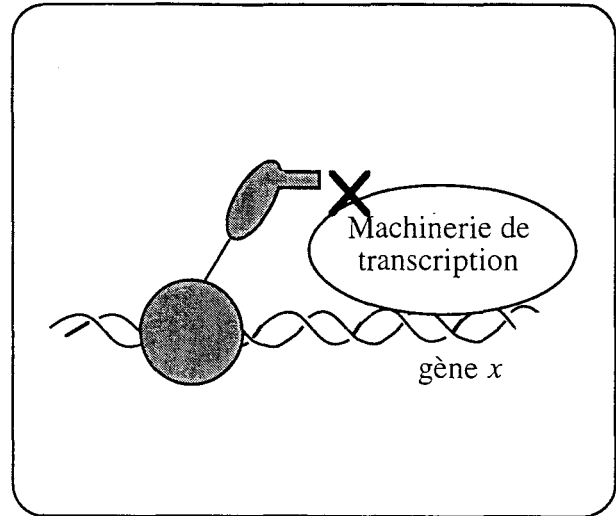
similaires ont été rapportées pour d'autres domaines activateurs, voir par exemple Emami *et al.*, 1995).

Une conséquence de cet état de fait est que, selon la nature de la cellule considérée et le contexte du promoteur où Erm est recruté, l'un, l'autre ou les deux domaines seront actifs. En d'autres termes, **les caractéristiques des domaines de transactivation déterminent en partie quels pourront être les gènes-cibles d'Erm, et dans quelles cellules ils seront régulés** (fig. 34).

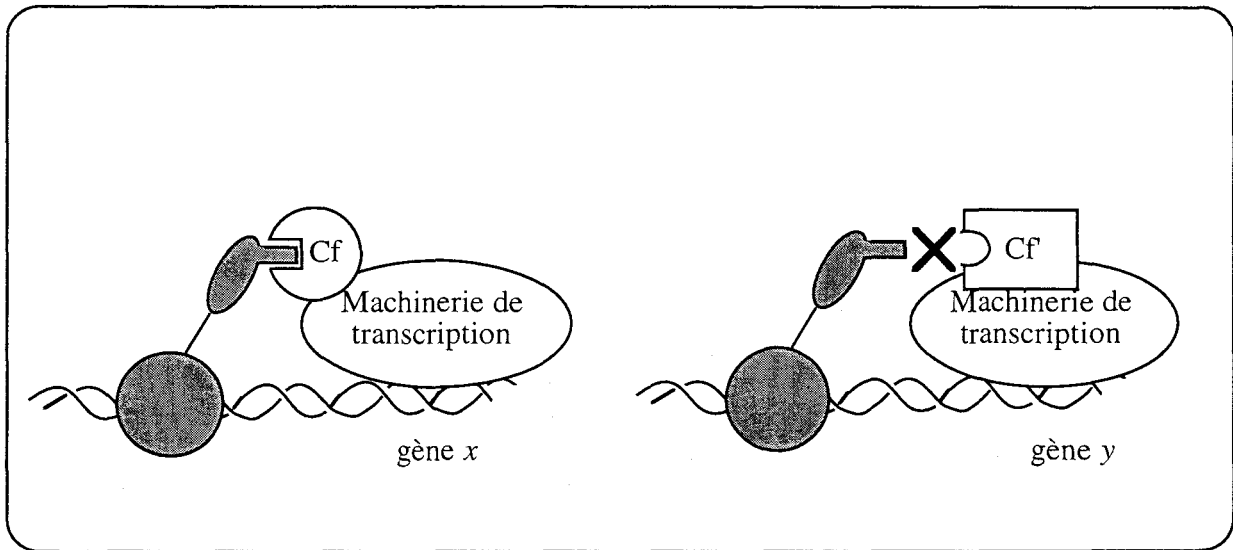
Bien sûr nous ne sommes pas à l'abri d'effets artefactuels liés à l'utilisation de protéines chimères. Il faudra donc répéter ces expériences en utilisant la protéine Erm de pleine longueur, fixée sur un site ETS. Rappelons toutefois que la région AD2/Ct est ambivalente : elle peut activer la transcription mais aussi moduler l'affinité d'Erm pour l'ADN. Il est donc difficile d'interpréter les variations de pouvoir transactivateur observées lorsque l'on ampute Erm d'AD2 (voir par exemple publication n°3, fig. 5).



Cellule A:
Erm active la transcription
du gène *x*, via un cofacteur (Cf)



Cellule B :
le cofacteur est absent,
pas d'activation de transcription



La transcription du gène *y* ne requiert pas le cofacteur
contacté par Erm, la transcription n'est pas activée.

Fig. 34 Les propriétés des domaines activateurs déterminent en partie les gènes-cibles d'Erm
Pour plus de clarté, un seul domaine activateur est représenté. Selon ses caractéristiques, Erm peut:
- induire un gène dans un type cellulaire donné, mais pas dans un autre
- dans la même cellule, induire ou ne pas induire l'expression des gènes auxquels il est recruté

Publication n°3

The ETS family member ERM contains an alpha helical acidic transactivation domain.

Defossez P.A., Monnot M. et de Launoit Y.

Soumis.

Résumé:

Cette publication propose une étude structurale du domaine transactivateur AD1 d'ERM.

Nous établissons que la fonction de ce domaine dépend d'une région d'une vingtaine de résidus à caractère acide, qui a le potentiel de former une hélice α . Un peptide de 32 acides aminés recouvrant cette région peut être induit à former une hélice α .

En testant le pouvoir transactivateur de mutants ponctuels de ce domaine acide, nous montrons que l'intégrité structurale de l'hélice est nécessaire à l'activité d'AD1. Sa richesse en résidus acides est moins cruciale.

En examinant la séquence d'autres domaines transactivateurs acides, nous observons que plusieurs d'entre eux contiennent une hélice α potentielle. Nous proposons d'expliquer l'hétérogénéité fonctionnelle des transactivateurs acides par une hétérogénéité structurale.

The ETS family member ERM contains an alpha helical
acidic transactivation domain.

Pierre-Antoine Defossez, Monique Monnot[°] and Yvan de Launoit*.

Unité d'Oncologie Moléculaire, CNRS URA 1160, Institut Pasteur, 1 rue Calmette, 59019 Lille
Cedex, France.

[°] CNRS URA 147, Institut Gustave Roussy, 39 rue Camille Desmoulins, 94805 Villejuif
Cedex, France

* to whom correspondence should be addressed : phone : (33) 20 87 78 78
fax : (33) 20 87 79 08
e.mail : Sprite@Citi2.fr

Running title : Alpha helical acidic transactivation domains

Abstract

Our investigations focus on ERM, a eukaryotic transcription factor of the ETS family. ERM harbors two transactivating domains (TADs) : AD1 lies in its first 72 AA, and AD2 in the last 62. Here we show that AD1 contains a 20 AA stretch predicted to form an α helix that is found conserved in the related PEA3 and ER81 proteins. As the structural constraints put on transactivating regions, if any, are mostly unknown, we investigated the structure of AD1 further. Circular Dichroism analysis revealed that an AD1-derived 32 AA peptide is unstructured in water but folds into an helix when the hydrophobic solvent TriFluoroEthanol is added. The isolated helix activated transcription when tethered to an heterologous DNA-binding domain. Helix-breaking mutations dramatically affected AD1-driven transactivation, whereas mutations decreasing its acidity had more gentle effects. A phenylalanine residue within the helix was particularly sensitive to mutations. We also show that AD1 is a *bona fide* acidic TAD, for it activated transcription in yeast cells. We have recognized a similar helical motif with important aromatic residues in a set of acidic TADs and believe that, along with AD1 of ERM, they define a structural subclass of the acidic TADs.

Introduction

The ETS family is a set of eukaryotic genes encoding transcription factors with a conserved DNA-binding region named the ETS domain (1). ETS genes are present throughout the metazoans (2) and play key developmental roles. For instance, *pointed* and *yan* control eye development in *Drosophila* (3), while in the mouse the B and T lymphocytes fail to differentiate properly in the absence of a functional *ets1* (4).

Beside their physiological roles, genes of the ETS family can also be involved in oncogenesis. Indeed *v-ets* is one of the oncogenes responsible for the transforming properties of the E26 chicken leukemia virus (5). Some other ETS genes, including *fli*, *pul* or *tel*, also are proto-oncogenes for their misregulation or their structural alteration can lead to cell transformation (6, see 7 for a review).

The regulation of the activity of ETS transcription factors has in the recent past become clearer, and some of them are now inserted in well-defined signal transduction pathways. A series of genetic or biochemical investigations have for instance exposed the Pointed (3, 8), Ets1 and Ets2 (9) proteins as downstream effectors of the Ras/Raf/ERK pathway. Another ETS gene product, Elk, associates with SRF to manage the serum induction of *c-fos*. Elk can be activated by Ras through the ERK kinases, but is also targeted by the JNK kinases (see 10 for a review). The Ras pathways potentiate the transactivating functions of these ETS proteins by promoting their phosphorylation at well-defined sites. However, as little is known about the structure of the ETS proteins transactivating domains and their targets, the mechanisms underlying their function and regulation by phosphorylation remain elusive.

We have focused our interest on the PEA3 subgroup of the ETS family. To date it contains three genes with high sequence similarity : PEA3 (11), ER81 (12) and ERM (13).

The members of the PEA3 group appear to be involved in different aspects of cancer disorder: ER81 is found fused to the EWS protein as a result of the (7;22)(p22;q12) chromosomal translocation in some cases of Ewing's sarcomas (14) and thus is a proto-oncogene. Several lines of evidence also point to PEA3 as a participant in the metastasis process : both in *neu*-induced mammary tumors (15) and in squamous cell carcinoma cell lines (16), its expression correlates with metastatic potential. Furthermore it can stimulate the transcription of several genes encoding matrix metalloproteinases (17), and its expression is sufficient to turn a non-invasive breast cancer cell line into an invasive one (18).

We had previously engaged in a study of the transactivation properties of ERM, and found it harbors two transactivating domains (TADs) : AD1 (formerly α) lies in the first 72 AA of ERM and AD2 in the last 62 (19). The two TADs are functionally dissimilar and exhibit a synergistic effect on transcription.

AD1 (also described in 20) seems to have been conserved within the PEA3 group. Indeed the homologous region of ER81 and PEA3 also can activate transcription (21, J.H. Chen, pers. comm.). This functional conservation prompted us to further investigate its characteristics.

Here we show that AD1 has functional features of the "acidic" class of TADs, which for example contains TADs from the Herpesvirus VP16 protein, or the yeast regulators GAL4 and GCN4. However we demonstrate that, unlike the VP16 TAD, which seems to be unstructured (22, 23), or that of GAL4 and GCN4 that adopt β -sheet conformation (24, 25), AD1 depends on a 15 AA α helix for function. This helix, fully conserved in ERM, ER81 and PEA3, requires a proper protein environment for folding and function, and constitutes a single exon, raising the possibility that it has been inserted into an ERM/ER81/PEA3 precursor gene by exon shuffling. Acidic transactivation domains were previously recognized on primary structure criteria and were not generally believed to adopt specific structures, save for a few identified cases (26, 27, 28, for instance). We show here that several other acidic TADs are predicted to contain an α helix bearing functionally crucial aromatic residues. This finding argues for the assignment of these TADs, together with AD1 of ERM, to a specific subgroup of "helical-aromatic" acidic TADs.

Materials and Methods

Plasmids

To ease the construction of GAL4(1-147) chimeras, we engineered a new vector we call pGAP. It has a pSG5 backbone and promoter, an IRES sequence to increase mRNA translatability, the GAL4(1-147) coding sequence and an extended cloning linker.

pSG424AD1_{F47P} was obtained by PCR of pSG424AD1 (formerly pSG424 α , see 19) with mutant primers changing a TTT codon to CCT. AD1, AD1_{F47P}, and AD2 were recloned from pSG424 into pGAP424 as EcoRI/XbaI restriction fragments. The helix within AD1 of ERM was cloned by PCR. Mutagenesis of pGAPAD1 was performed using the Chameleon double-strand mutagenesis kit (Stratagene). The unrelated linker peptide QSRAATAVELGTRSY was introduced into GAL4helix by a site-directed mutagenesis that removed the first stop codon.

ERM and its mutant derivatives were cloned into pSV (a gift of Guillaume Adelmant, see 29), an expression vector containing an HA tag at the amino-terminus. ERM was excised from pSG424ERM (19) and inserted into pSV as an EcoRI/XbaI partial restriction fragment. The F47P mutation was introduced into ERM by site-directed mutagenesis. PCR with internal primers was performed on ERM and ERM_{F47P} to remove the carboxy-terminal activation domain AD2, yielding pSVERM Δ AD2 and pSVERM_{F47P} Δ AD2.

For expression in yeast cells AD1, AD1_{F47P}, and AD2 were fused to GAL4(1-147) in the pGBT9 vector, as EcoRI-PstI (for the first two constructs) or EcoRI-BamHI fragments (for the last).

The 5xGAL4 Luc plasmid containing five GAL4 binding sites upstream of the luciferase gene was a kind gift of Dr Jay Gralla (pG5TATA Luc in reference 30). The 3xTORU Luc and ICAM(-176/-44) Luc plasmids have been described (19, 31, respectively).

All PCR and mutagenesis products were sequenced on an automated sequencer (Applied Biosystems 373A).

Cell culture, transfection, and labelling.

RK13, Hela and HC11 cells were maintained in DMEM supplemented with 10% Fetal Bovine Serum at 37°, in water-saturated atmosphere with 5% CO₂.

Subconfluent cells were transfected in 6-well plates with lipofectamin (Boehringer Mannheim) as previously described (19). They were harvested 24 hours after transfection, and the luciferase activity measured using a Berthold LB 9501 chimiluminometer. An RSVLacZ plasmid was cotransfected, and the LacZ activity also assessed (Tropix Galacto-Light kit) to normalize data for transfection efficiency.

Metabolic labeling of transiently transfected cells was performed 24 hours after transfection, for three hours in medium containing 100 μ Ci/ml ³⁵S-Met/Cys. Immunoprecipitation was as described in (32).

Yeast manipulation.

The strain Y187, which has an integrated LacZ gene downstream of GAL4 binding sites, was made competent and transformed by the PEG/LiAc method, and then cultured in YPD lacking the appropriate amino-acids. Transformed cells were disrupted by bead-beating and the LacZ activity in the extract measured using the Tropix GalactoLight kit. It was normalized to the protein concentration in the extract. Protein extracts used to perform immunoblots were made as in (33).

Structure predictions.

Structure predictions were performed with the PHD program (34), available as predictprotein@EMBL-Heidelberg.de.

Circular Dichroism.

Circular dichroic spectra were recorded on a Jobin-Yvon CD6 Dichrograph in a 1 mm path length quartz cell, with a thermostat cell holder. The Acidic Domain peptide had the sequence GYDSEELFQDLSQLQETWLAEAQVPDNDEQFVPD. It was dissolved in aqueous solution with NaOH at pH 10.6, at a concentration of $1.4 \cdot 10^{-5}$ M.

The effect of TFE titration was studied between 0 and 90% TFE (TFE Volume/Total Volume) in aqueous alkaline solution at 20°C.

CD was expressed as mean residue ellipticity $[\theta]$ in $\text{deg.cm}^2.\text{dmol}^{-1}$. $[\theta]$ is related to the dichroic increment $\Delta\epsilon$ as follows : $[\theta] = (3300 \times \Delta\epsilon) / N$, where N is the number of residues.

Results

A single exon-encoded α helix is shared by the N-terminal TADs of ERM, ER81 and PEA3.

We and others have detected a TAD in the aminotermminus of ERM, AD1 (formerly α) (19, 20). The homologous region of ER81 and PEA3 can also activate transcription (21, J.H.Chen, pers. comm.). This functional conservation prompted us to investigate the features of AD1. A 32 AA region rich in acidic residues (30% are either D or E), the "Acidic Domain" (13), is particularly conserved between ERM, ER81 and PEA3 (see fig. 1A). A structure prediction using the PHD program (34) was performed on the ERM sequence and predicted with 90% confidence the presence of an α helical segment at the N-terminal side of this Acidic Domain (fig. 1B). The putative helix precisely fits within an exon, exon 4 of ERM (35), and homologous exons are also found in PEA3 and ER81 (our unpublished results).

We sought more direct evidence for the presence of an α helix in AD1. We had in parallel studies generated a peptide encompassing the whole Acidic Domain of ER81, which contains the putative helix, and is nearly identical to that of ERM (see fig. 1A). The results of the Circular Dichroism analysis that was performed on this peptide are shown in fig.2. In aqueous alkaline solution, a strong minimum appears at 200 nm, a characteristic of a largely random structure. The small negative value at 222 nm could reflect a faint α helix proportion (about 6%). Additional recordings were made in the presence of the hydrophobic solvent TriFluoroEthanol (TFE), which is known to ease the formation of potential α helices. As the TFE/water ratio increased, a characteristic α helical signature appeared as a double minimum at 207 and 222 nm and a positive maximum at 190 nm (fig. 2). An isodichroic point occurs at 204 nm, indicating that TFE induces a two-state, mainly coil to helix, transition. With the assumption of a balance between the Tyr and Trp far-UV CD contributions (36), we could calculate a 39% α helix content at 90% TFE from the $[\theta]_{222}$ value (37). This could mean either that at any given time 39% of the molecules are fully helical, or that all molecules are helical over 39% of their length, that is, about 15 AA. Both the Chou-Fasman (38) and Robson-Garnier (39) algorithms agreed with PHD on predicting a 15 to 20 AA α helix component in the N-Terminal region of the 32 AA-long Acidic Domain, but failed to detect an organized secondary structure in the remaining fragment. We thus favour the second interpretation and conclude that an α helix at least 15 AA long exists between AA 43 and 62 of ERM AD1.

A 20 AA domain containing the α helix is sufficient for transactivation but requires neighbouring sequence elements for maximum activity.

The helix we detected appeared as an autonomous structural unit. We tried to determine whether it also was functionally autonomous and constituted a TAD on its own. To that end its 20 AA were fused to the GAL4(1-147) DNA-binding-domain, which itself has little transcriptional modulation properties (40). Transient cotransfections were then performed in

mammalian cells to monitor the effect of this chimera on the expression of a reporter gene bearing GAL4 operators in its promoter. For comparison purposes, we performed the same experiments with chimeras including the full-size 72 AA AD1, or the 32 AA Acidic Domain (see fig. 3A).

As shown in fig. 3B, the GAL4helix fusion was found to activate transcription, albeit with a threefold lower efficacy than the complete Acidic Domain, itself being about fourfold less efficient than the entire AD1 (a result shown in 19, and reproduced here for comparison). This was not the consequence of a lower expression, as we verified all chimeras to be expressed at comparable levels (fig. 3C).

We thought that the helix could merely require additional protein length on its C-terminal side for appropriate folding, and thus added to our construct a linker sequence to reconstitute a peptide with a length similar to that of the Acidic Domain. This did not, however, increase the transactivating potency of the helix (compare the activity of GAL4helix to that of GAL4helix-linker). We thus conclude that the helix is an autonomous activation domain that cooperates with neighbouring sequences for maximum activity. As will be shown below, these cooperating sequences are not themselves capable of transactivation.

Mutations affecting the α helix of AD1 abrogate its transactivation capacity.

Though the helix was sufficient for activity, there was no proof that helical structure in the Acidic Domain was necessary for transactivation. We thus engineered in the GAL4AD1 construct a set of mutations chosen to disturb or not the helix (fig. 4A), and assessed the activity of the mutants as described above.

As shown in fig. 4B, we observed that both mutations predicted to impair the formation of the α helix (F47P and L50P) had very severe effects on GAL4AD1 transactivation levels, rendering them merely superior to that of the control GAL4(1-147). An equivalent F to P mutation introduced at residue 69, which lies outside the predicted helix, and a mutation replacing acidic residues but compatible with helical structure, E44S/D49A, had in contrast quite moderate consequences : these mutant TADs indeed are only one-third less active than their wild-type counterpart.

The F47L mutation, though predicted to spare the helix, had an effect as dramatic as helix-breaking mutations. F47 could nevertheless be replaced by the W residue. It thus appears that an aromatic residue within the helix is crucial, a situation reminiscent of that observed in RelA (28, 41, see Discussion for further details).

The transcriptional defects in the mutant proteins were observed at all plasmid doses tested (not shown), and are not, for instance, due to intense self-squelching effects. We also verified the wild-type and mutant proteins to be expressed at comparable levels by immunoprecipitation (fig. 4C).

We thus conclude that the presence of the helix is necessary for activation by GAL4AD1. As a point mutation within the helix totally deprives AD1 of transactivation potential, we also conclude that regions of AD1 outside the helix do not on their own activate transcription.

Mutations in the α helix of AD1 impair transactivation by full-size ERM.

AD1 of ERM does not lie in its native environment in the GAL4 chimera, and the fusion could render it artificially sensitive to the introduced mutations. We thus tested whether one of the point mutations characterized above, F47P, had a comparable effect when inserted into the native ERM molecule (fig. 5A).

We have recently found that transcription from the ICAM-1 (Inter Cellular Adhesion Molecule 1) gene promoter can be stimulated by ERM (de Launoit et al., submitted). The (-176/-44) fragment of this promoter bears only one high-affinity ETS site, and was used to drive the expression of the Luciferase reporter gene. Transient cotransfections were performed in the HC11 mouse mammary cell line. Under these conditions we found ERM to drive a 35-fold increase in reporter gene activity over the control vector, while the potency of ERM_{F47P} was about fourfold lower than that of ERM (fig. 5B). We have previously shown that ERM contains a second TAD, AD2, in its carboxy-terminal tail (19). To know whether the residual activity of ERM_{F47P} could be attributed to AD2, we removed AD2 from ERM_{F47P} and indeed found the corresponding construct ERM_{F47P} Δ AD2 to be a very poor transactivator, with about 10% the activity of ERM. Note that deletion of AD2 from ERM is not in itself sufficient to deprive it of transactivating ability (see lane ERM Δ AD2). These effects were not dose-dependent (not shown).

Mutations in transactivating modules can sometimes be partly compensated by intermolecular cooperation when several mutated molecules are simultaneously recruited to the promoter (42). To determine whether that applied to mutations in AD1 in the context of the full-size ERM, we assayed the activity of ERM and its derivatives on the 3xTORU Luc reporter plasmid, which harbors an artificial promoter with three ETS binding sites, in rabbit RK13 cells as previously described (19). Under these conditions, transfection of ERM invoked a 35-fold increase in reporter gene activity, while the effect of ERM_{F47P} was about six fold lower (fig. 5B). The activity of ERM_{F47P} Δ AD2 was hardly detectable.

We believe the F47P mutation selectively impairs transcriptional activation and not another functional feature of ERM, for we verified the mutant proteins i) to be expressed at a level similar to that of the wild-type (fig. 5C) ii) to be correctly addressed to the nucleus (immunofluorescence not shown) and iii) to have the same *in vitro* DNA binding capacity as their wild-type counterpart (EMSA shown in fig. 5D).

We thus conclude that a point mutation in the α helix in AD1 of ERM is indeed sufficient to cripple its transactivation potential. This effect is not promoter nor cell-type specific and cannot be overcome by intermolecular cooperation.

AD1 of ERM activates transcription in yeast cells, and this effect is abrogated by the F47P mutation.

A hallmark of TADs of the acidic class, as opposed to the "glutamine-rich" ones for instance, seems to be their potential to activate transcription in yeast cells. Sequence observation could lead us to surmise that AD1 was a TAD of the acidic type. To learn if this was functionally true, we tested whether a GAL4AD1 chimera could increase the expression of a lacZ gene containing GAL4 operators in its promoter, in the yeast strain Y187 which lacks endogenous GAL4 (fig. 6). We found that it was indeed the case. As observed in mammalian cells, introducing the F47P mutation into GAL4AD1 deprived it of most of its transactivating ability. This was not due to a poor expression for the chimeric proteins were verified to be similarly expressed by immunoblotting (not shown).

This supports the widely evidenced hypothesis that the mechanisms of transcriptional activation have been well conserved from yeast to mammalian cells (see 43, for instance).

A subset of acidic TADs contain potential α helices with functionally crucial aromatic residues.

Three genes have previously been shown to contain acidic TADs whose activity depends on an α helix : the Glucocorticoid Receptor (26), p65 RelA (28) and the helix-loop-helix myogenic factor E2A (27). We have investigated whether that structural feature could be found in other acidic TADs, and structure predictions performed with the PHD program indeed revealed potential helices in a number of these (fig. 7). Notably, we failed to detect similar structures in activators that nevertheless bear primary structure similarity to ERM, such as GAL4, GCN4 or VP16. An aromatic residue is found at a conserved position and, in the cases where its importance was investigated, has proved crucial. We therefore believe that AD1 of ERM, along with the depicted TADs and undoubtedly others, exemplifies a newly defined subclass of the acidic TADs, characterized by an helical motif with a conserved aromatic residue.

Discussion

An α helix as the functional core of AD1

A region within AA 43-62 of ERM has the potential to structure into an α helix (fig. 1, 2), and the results of our mutagenesis studies clearly show that helical potential in this region is required for AD1 activity, negative charge being of a lesser importance (fig. 4, 5). Yet a peptide containing the potential helix is unstructured in water. Two interpretations could reconcile this finding with our model of an helical activation domain.

The first one surmises that protein sequences present in native ERM, but absent from the peptide used in CD spectroscopy, are required for the helix to fold into its structure. Nevertheless this is hardly compatible with the observation that the same peptide, when tethered to the GAL4 DNA-binding domain, activates transcription, unless one assumes that the GAL4 protein itself induces folding in the helix.

Alternatively, it could be that the helix acquires structure by an "induced-fit" mechanism. The transcriptional activator VP16 has been shown to invoke conformational changes in TFIIB upon contact (44); it is thus conceivable that, conversely, the activating helix folds when reaching its target. It has for instance recently appeared that the VP16 TAD, which contains highly mobile segments, becomes structurally constrained upon contact with TBP (45). Another example supporting this hypothesis comes from studies on RelA. RelA indeed contains an helical acidic TAD, and a 32 AA peptide spanning this region is unstructured in water but folds when TFE is added (28). NMR proved this activating region also to be unstructured in water in the context of a 123 AA RelA fragment (28). Besides, a RelA C-terminal activation domain-derived peptide in water, though unstructured, can interact with its targets (41). Although more detailed studies would be necessary to ascertain this point, we believe a comparable situation occurs in AD1 of ERM : the helix may be unstructured at rest but fold upon meeting its interaction partner.

The helix contained within exon 4 of ERM likely constitutes the active core of AD1. Indeed a point mutation at AA47 totally abrogates transactivation by AD1, implying that regions other than the helix are not on their own sufficient for transactivation (fig. 4 and 5). Nevertheless, it appears in the context of the GAL4 chimera that removing either the 1-42 or the 63-72 regions notably reduces the helix-driven transactivation (fig. 3). How, then, could it be that these segments act as potentiators while being themselves inactive? Possibly the helix, when directly tethered to GAL4, is constrained in a non-optimal situation, and the neighbouring regions merely act as inert linkers providing flexibility. We judge this unlikely, though, for a peptide as short as 15 AA has already been shown to activate transcription when tethered to GAL4(1-147) (46). Also, more conclusively, the sequences C-terminal to the helix could not be replaced by a peptide chain of similar length without a clear activity decrease (fig. 3).

A recent report has brought a striking illustration that secondary structure does not only depend on AA composition, but also on neighboring sequences (47). We thus favour the hypothesis that the regions flanking the helix provide a propitious environment for helix formation. This may in particular explain the high degree of sequence conservation between ERM, ER81 and PEA3 in the 63-72 region of the Acidic Domain, and in the 30-41 region between ERM and ER81.

The acidic class of TADs.

Most TADs have up to now been assigned to one of the Proline-rich, Glutamine-rich or Acidic classes of TADs (48, however see 49, 50, among others, for exceptions to that rule). As their names imply, these categories were primarily defined after sequence features. They nevertheless have a functional relevance, for members of the same group share a number of traits.

Hallmarks of the acidic TADs include the ability to activate transcription in yeast cells, a property common to Proline-rich TADs, but not Glutamine-rich ones (48). Also, the observation that acidic TADs squelch each other's activity strongly argues for the existence of shared cofactors (51).

A set of results nevertheless challenges the homogeneity of the acidic class of TADs.

First, not all of them interact with the same targets within the transcriptional machinery (see for instance (52) for a comparison of the two VP16 activation domains; and (53) for a study of EBNA-2 and VP16). Even more compellingly, experiments of a genetic nature have pointed out that not all acidic TADs have the same cofactor requirements. Mutations in the yeast adaptor ADA2 for instance severely cripple the transactivation potential of VP16, but leave that of HAP1 mostly unabated (54).

It is striking in the perspective of acidic TAD disparity that they assume different structures. Indeed, while the VP16 activation domain is unstructured (22, 23), those of GAL4 and GCN4 form β sheets under proper conditions (24, 25); more recently, acidic TADs from the RelA (28), Glucocorticoid Receptor (26), E2A (27) and ERM (this study) proteins have been shown to adopt an α helical conformation. We have performed structure predictions on a variety of acidic TADs and found several others were likely to contain an α helix bearing resemblance to that already described (fig. 7). However we do not at present have a proof that the three structural classes of acidic TADs (unstructured, forming a β sheet or an α helix) mirror functional categories. The fact that the VP16 TAD, which is unstructured, squelches helical acidic TADs such as those of RelA (28) or ERM (19) seems to argue against this possibility. It should not be overlooked, though, that the VP16 TAD might contact the same target(s) as ERM or RelA, but through different interfaces, and nevertheless cause squelching.

It might also be that TADs of the three structural classes have evolved from a common ancestor, which would explain their sequence similarity, and have maintained some common features while evolving new ones.

TAD diversity in the ETS family.

ETS family genes share a well-conserved DNA-binding domain and, at least *in vitro*, some display overlapping DNA-binding specificities (55). Their functional specificity *in vivo* may stem from different features : difference of expression pattern, of cofactor recruitment, or of transactivation characteristics. Indeed the ETS proteins have evolved a variety of TADs (see (56, 57, 58, 59, 60 for illustrations), and it is known that different TADs may have distinct cell-type and promoter element specificities (61, 62, 63). Acquisition of unrelated TADs by different ETS genes thus may permit regulation of target genes with dissimilar promoter contexts, in diverse cell types.

Also, one can notice that many transcription factors, including most of the ETS family members, possess several TADs. The acquisition of additional TADs within the same molecule may be beneficial for these domains may cooperate to yield higher inductions of target gene expression (64).

We have identified the functional core of ERM AD1 as an α helix that is encoded by a single exon, ERM exon 4. The remaining part of the Acidic Domain, also highly conserved in ER81 and PEA3, positively regulates the activity of the helix, and is also encoded by a single exon, exon 5 of ERM (35). The intron between exons 4 and 5 is only 128 bp long. As exons 3 and 6 of ERM have compatible reading frames and are separated by a particularly large intron, we view it as possible that the (exon 4-intron-exon 5) block became inserted into the PEA3-ERM-ER81 progenitor by an exon-shuffling mechanism. Notably the (exon 4-intron-exon 5) module is flanked by phase-1 introns, as are most of the modules building up mosaic proteins (65).

We think the acquisition of AD1 by the PEA3 group members exemplifies the way ETS proteins have become increasingly diverse. This diversity, along with the evolution of various promoter and enhancer elements, now allows them to regulate developmental events as distant as hematopoiesis (66) and skeletal development (67).

References

1. Karim, F. D., Urness, L. D., Thummel, C. S., Klemsz, M. J., McKercher, S. R., Celada, A., Van Beveren, C., Maki, R. A., Gunther, K., Nye, J. A., and Graves, B. J., (1990) *Genes & Dev.* **4**, 1451-1453
2. Degan, B. M., Degan, S. M., Nagagnuma, T., and Morse, D. E., (1993) *Nucleic. Acids. Res.* **21**, 3479-3484
3. O'Neill, E. M., Rebay, I., Tjian, R., and Rubin, G. M., (1994) *Cell* **78**, 137-147
4. Bories, J.-C., Willerford, D. M., Grévin, D., Davidson, L., Camus, A., Martin, P., Stéhelin, D., and Alt, F. W., (1995) *Nature* **377**, 635-638
5. Leprince, D., Gegonne, A., Coll, J., de Taisne, C., Schneeberger, A., Lagrou, C., and Stehelin, D., (1983) *Nature* **306**, 395-397
6. Golub, T., Barker, G. F., Lovett, M., and Gilliland, D. G., (1994) *Cell* **77**, 307-316
7. Wasyluk, B., Hahn, S. L., and Giovane, A., (1993) *Eur. J. Biochem.* **211**, 7-18
8. Brunner, D., Dücker, K., Oellers, N., E., H., Scholz, H., and Klämbt, C., (1994) *Nature* **370**, 386-389
9. Yang, B.-S., Hauser, C. A., Henkel, G., Colman, M. S., Van Beveren, C., Stacey, K. J., Hume, D. A., Maki, R. A., and Ostrowski, M. C., (1996) *Mol. Cell. Biol.* **16**, 538-547
10. Treisman, R., (1995) *EMBO J.* **14**, 4905-4913
11. Xin, J. H., Cowie, A., Lachance, P., and Hassell, J. A., (1992) *Genes and Dev* **6**, 481-496
12. Brown, T. A., and McKnight, S. L., (1992) *Genes and Dev* **6**, 2502-2512
13. Monté, D., Baert, J. L., Defossez, P. A., de Launoit, Y., and Stéhelin, D., (1994) *Oncogene* **9**, 1397-1406
14. Jeon, I.-S., Davis, J. N., Braun, B. S., Sublett, J. E., Roussel, M. F., Denny, C. T., and Shapiro, D. N., (1995) *Oncogene* **10**, 1229-1234
15. Trimble, M. S., Xin, J. H., Guy, C. T., Muller, W. J., and Hassell, J. A., (1993) *Oncogene* **8**, 3037-3042
16. Shindo, M., Higashino, F., Kaya, M., Yasuda, M., Funaoka, K., Hanzawa, M., Hida, K., Kohgo, T., Amemiya, A., Yoshida, K., and Fujinaga, K., (1996) *Am. J. Pathol.* **148**, 693-700
17. Higashino, F., Yoshida, K., Noumi, T., Seiki, M., and Fujinaga, K., (1995) *Oncogene* **10**, 1461-1463
18. Kaya, M., Yoshida, K., Higashino, F., Mitaka, T., Ishii, S., and Fujinaga, K., (1996) *Oncogene* **12**, 221-227
19. Laget, M. P., Defossez, P.-A., Albagli, O., Baert, J. L., Dewitte, F., Stehelin, D., and De Launoit, Y., (1996) *Oncogene* **12**, 1325-1336
20. Nakae, K., Nakajima, K., Inazawa, J., Kitaoka, T., and Hirano, T., (1995) *J. Biol. Chem.* **270**, 23795-23800
21. Janknecht, R., (1996) *Mol. Cell. Biol.* **16**, 1551-1556
22. Cress, W. D., and Triezenberg, S. J., (1991) *Science* **251**, 87-90

23. O'Hare, P., and Williams, G., (1992) *Biochemistry* **31**, 4150-4156
24. Leuther, K. K., Salmeron, J. M., and Johnston, S. A., (1993) *Cell* **72**, 575-585
25. Van Hoy, M., Leuther, K. K., Kodadek, T., and Johnston, S. A., (1993) *Cell* **72**, 587-594
26. Dahlman-Wright, K., Bauman, H., McEwan, I. J., Almlöf, T., Wright, A. P. H., Gustafsson, J.-A., and Hard, T., (1995) *Proc. Natl. Acad. Sci. USA* **92**, 1699-1703
27. Massari, M. E., Jennings, P. A., and Murre, C., (1996) *Mol. Cell. Biol.* **16**, 121-129
28. Schmitz, M. L., dos Santos Silva, M. A., Altmann, H., Czisch, M., Holak, T. A., and Bauerle, P. A., (1994) *J. Biol. Chem.* **269**, 25613-25620
29. Adelmant, G., Bègue, A., Stéhelin, D., and Laudet, V., (1996) *Proc. Natl. Acad. Sci. USA* **93**, 3553-3558
30. Chang, C., and Gralla, J. D., (1993) *Mol. Cell. Biol.* **13**, 7469-7475
31. Voraberger, G., Schäfer, R., and Stratowa, C., (1991) *J. Immunol.* **147**, 2777-2786
32. Albagli, O., Soudant, N., Ferreira, E., Dhordain, P., Dewitte, F., Begue, A., Flourens, A., Stéhelin, D., and Leprince, D., (1994) *Oncogene* **9**, 3259-3271
33. Silverman, N., Agapite, J., and Guarente, L., (1994) *Proc. Natl. Acad. Sci. USA* **91**, 11665-11668
34. Rost, B., (1996) *Meth. in Enzym.* **266**, 525-539
35. Monté, D., Coutte, L., Dewitte, F., Defossez, P. A., Le Coniat, M., Stéhelin, D., Berger, R., and de Launoit, Y., (1996) *Genomics* **35**, 236-240
36. Chakrabarty, A., Kortemme, T., Padmanhaban, S., and Baldwin, R. L., (1993) *Biochemistry* **32**, 5560-5565
37. Padmanhaban, S., Marqusee, S., Ridgeway, T., Laue, T. M., and Baldwin, R. L., (1990) *Nature* **344**, 268-270
38. Chou, P. Y., and Fasman, G. D., (1978) *Annu. Rev. Biochem.* **47**, 251-276
39. Garnier, J., Osguthorpe, D. J., and Robson, B., (1978) *J. Mol. Biol.* **120**, 97-120
40. Sadowski, I., and Ptashne, M., (1989) *Nucleic Acids Res.* **17**, 7539
41. Blair, W. S., Bogerd, H. P., Madore, S. J., and Cullen, B. R., (1994) *Mol. Cell. Biol.* **14**, 7226-7234
42. Walker, S., Greaves, R., and O'Hare, P., (1993) *Mol. Cell. Biol.* **13**, 5233-5244
43. Reese, J. C., Apone, L., Walker, S. S., Griffin, L. A., and Green, M. R., (1994) *Nature* **371**, 523-527
44. Roberts, S. G. E., and Green, M. R., (1994) *Nature* **371**, 717-720
45. Shen, F., Triezenberg, S. J., Hensley, P., Porter, D., and Knutson, J. R., (1996) *J. Biol. Chem.* **271**, 4827-4837
46. Giniger, E., and Ptashne, M., (1987) *Nature* **330**, 670-672
47. Minor, D. L., and Kim, P. S., (1996) *Nature* **380**, 730-734
48. Triezenberg, S. J., (1995) *Curr. Op. Gen. Dev.* **5**, 190-196
49. Attardi, L. D., and Tjian, R., (1993) *Genes Dev.* **7**, 1341-1353

50. Duncan, R., Collins, I., Tomonaga, T., Zhang, T., and Levens, D., (1996) *Mol. Cell. Biol.* **16**, 2274-2282
51. Oehler, T., and Angel, P., (1992) *Mol. Cell. Biol.* **12**, 5508-5515
52. Goodrich, J. A., Hoey, T., Thut, C. J., Admon, A., and Tjian, R., (1993) *Cell* **75**, 519-530
53. Tong, X., Wang, F., Thut, C. J., and Kieff, E., (1995) *J. Virol.* **69**, 585-588
54. Berger, S. L., Pina, B., Silverman, N., Marcus, G. A., Agapite, J., Regier, J. L., Triezenberg, S. J., and Guarente, L., (1992) *Cell* **70**, 251-265
55. Wang, C. Y., Petryniak, B., Ho, I. C., Thompson, C. B., and Leiden, J. M., (1992) *J. Exp. Med.* **175**, 1391-1399
56. Chumakov, A. M., Chen, D. L., Chumakova, E. A., and Koeffler, H. P., (1993) *J. Virol.* **67**, 2421-2425
57. Gegonne, A., Punyanmalee, B., Rabault, B., Bosselut, R., Seneca, S., Crabeel, M., and Ghysdael, J., (1992) *New Biol.* **4**, 512-519
58. Hagemeyer, C., Bannister, A. J., Cook, A., and Kouzarides, T., (1993) *Proc. Natl. Acad. Sci. USA* **90**, 1580-1584
59. Klemsz, M. J., and Maki, R. A., (1996) *Mol. Cell. Biol.* **16**, 390-397
60. Schneikert, J., Lutz, Y., and Wasylyk, B., (1992) *Oncogene* **7**, 249-256
61. Argenton, F., Arava, Y., Aronheim, A., and Walker, M. D., (1996) *Mol. Cell. Biol.* **16**, 1714-1721
62. Das, G., Hinkley, C. S., and Herr, W., (1995) *Nature* **374**, 657-660
63. Emami, K. H., Navarre, W. W., and Smale, S. T., (1995) *Mol. Cell. Biol.* **15**, 5906-5916
64. Hahn, S., (1993) *Nature* **363**, 672-673
65. Patthy, L., (1991) *Curr. Op. Struct. Biol.* **1**, 351-361
66. Scott, E. W., Simon, M. C., Anastasi, J., and Singh, H., (1994) *Science* **265**, 1573-1577
67. Sumarsono, S. H., Wilson, T. J., Tymms, M. J., Venter, D. J., Corrick, C. M., Kola, R., Lahoud, M. H., Papas, T. S., Seth, A., and Kola, I., (1996) *Nature* **379**, 534-537
68. Xiao, H., Friesen, J. D., and Lis, J. T., (1994) *Mol. Cell. Biol.* **14**, 7507-7516
69. Hwang, J. J., Chambon, P., and Davidson, I., (1993) *EMBO J.* **12**, 2337-2348
70. Nieto-Sotelo, J., Wiedderecht, G., Okuda, A., and Parker, C. S., (1990) *Cell* **62**, 807-817
71. Sutherland, J. A., Cook, A., Bannister, A. J., and Kouzarides, T., (1992) *Genes Dev.* **6**, 1810-1819
72. Brokaw, J. L., Blanco, M., and McBride, A., (1996) *J. Virol.* **70**, 23-29

Acknowledgements

Dr Jay Gralla provided the 5xGAL4 Luc plasmid used in this work. We thank Guy Lippens, Patrick Martin, Jean-Luc Baert, H el ene Pelczar, Anne Chotteau and other members of our group for discussion. Claire Montpellier, Isabelle Brun, Elisabeth Ferreira and Brigitte Quatannens offered friendly advice, and Laurent Pouilly efficient assistance. J.H. Chen accepted to share unpublished results. J. P. Levillain kindly initiated our collaboration with the Institut Gustave Roussy. Olivier Albagli, Simon Saule and Andrew Sharrocks made insightful comments on the manuscript. We feel particularly indebted to Guillaume Adelmant for his unvaluable support.

Funded in part by grants from the Centre National de la Recherche Scientifique, the Association pour la Recherche sur le Cancer, the Ligue contre le Cancer, the FEGEFLUC and the Institut Pasteur de Lille.

Legends to figures

Fig. 1 A potential α helix is conserved in the acidic transactivation domains of ERM, ER81 and PEA3.

A) Sequence comparison of the AD1 activation domain of human ERM and the homologous regions in mouse ER81 and mouse PEA3.

Amino acids on the CONS line are shared by all three genes. The boxed region is the "Acidic Domain" (13). AA numbers are indicated.

B) A potential α helix present in the Acidic Domain is encoded by a single exon of ERM.

The sequence of the N-terminal AD1 activation domain of ERM (central line, see 13) is compared with the exon boundaries (top line and vertical bars, 35), and with a secondary structure prediction using the PHD algorithm (34) (bottom line). E = Extended sheet, H = α Helix, no letter = random coil. The confidence of the helix prediction in this region is about 90%.

Fig.2 A peptide encompassing the Acidic Domain has α helix forming potential.

Circular Dichroism spectrum of a 34 amino-acid peptide encompassing the whole Acidic Domain of human ER81. The spectra were recorded at the indicated TFE/H₂O ratios, at a 1.4 10⁻⁵M peptide concentration, in a 1 mm path length quartz cell. The temperature was held at 20°. The peptide is mainly unstructured at 0% TFE as appears from the strong minimum at 200 nm. The signature of an helical structure appears as the TFE percentage increases (double minimum at 207 and 222 nm, positive maximum at 190 nm).

Fig. 3 The AD1 helix is sufficient for transactivation but requires flanking regions for maximal activity.

A) AD1, the Acidic Domain, and the AD1 helix of ERM followed or not by a linker (sequence in italics, see text) were fused to the GAL4 DNA-binding domain.

B) The chimeras were transiently cotransfected with a Luciferase reporter plasmid bearing five GAL4 binding sites (5xGAL4 Luc) in rabbit RK13 cells. The effector/reporter ratio was optimized at 500/300ng. 200ng of RSVLacZ were also cotransfected to allow normalization for transfection efficiency. The results are displayed as mean \pm SEM of the normalized luciferase activity relative to the control GAL4 plasmid, set as 1. The transfections were repeated three times with each of two independent DNA preparations.

C) RK13 cells transfected with the indicated plasmids were metabolically labeled, and an immunoprecipitation was performed with the cell lysates using an anti GAL4 polyclonal antibody (Upstate Biotechnology). The recovered proteins were separated by SDS PAGE, and all appear to be expressed at similar levels.

Fig. 4 Mutations altering the helix abrogate the transactivating capacity of AD1.

A) A series of point mutations predicted to alter (L50P, F47P) or not (E44S/D49A, F47L, F47W, F69P) the helix were introduced into GAL4AD1.

B) The activity of the wild type construct is compared to that of the indicated point mutants. The transfections were performed as described in the legend to fig.3.

C) Transfection, labeling and immunoprecipitation were performed as in fig. 3C. The point mutants of GAL4AD1 are expressed at a level comparable to that of the wild type.

Fig. 5 A mutation in the AD1 α helix impairs transactivation by full-size ERM.

A) Schematic representation of ERM, the point mutant ERM_{F47P} and their derivatives lacking the carboxy-terminal AD2 activation domain (ERM Δ AD2 and ERM_{F47P} Δ AD2), the activity of which was assessed by transient cotransfection with different reporter plasmids.

B) Left, the results of cotransfecting the ERM derivatives with the ICAM(-176/-44) Luc reporter plasmid (31) in mouse HC11 cells.

Right, similar experiments performed with the 3xTORU Luc plasmid bearing three copies of the ETS binding site TORU (19) in rabbit RK13 cells.

All experiments were repeated at least three times with each of two independent DNA preparations. Immunofluorescence showed all proteins were nuclearly localized (not shown).

C) HC11 cells were transfected with the indicated plasmids, labeled and used for an immunoprecipitation using a polyclonal anti HA antibody (BAbCo). ERM and its mutant derivatives are expressed at similar levels.

D) Rabbit reticulocyte lysates were programmed with the indicated plasmids. The expressed proteins were quantified with a PhosphorImager (Molecular Dynamics), and equimolar amounts were used in an Electrophoretic Mobility Shift Assay, using an ETS specific oligonucleotide. We had previously observed the deletion of AD2 to increase the DNA binding capacity of ERM (19). Nevertheless the F47P mutation does not alter the DNA binding ability of ERM.

Fig.6 AD1 activates transcription in yeast cells and the transactivation is impaired by the F47P mutation.

Yeast cells of the strain Y187, which bears an integrated LacZ reporter downstream of GAL4 binding sites, were transformed with the indicated plasmids. They were disrupted and the LacZ activity in the lysate measured and normalized to protein content. All chimeras were expressed at similar levels, as judged by an immunoblot (not shown).

Fig.7 Several TADs within the acidic class are helical and display conserved sequence features, including an aromatic residue.

The helicity of the underlined peptides has been proved by CD or NMR. All other peptides are predicted to be helical by the PHD program. In the cases where mutagenesis has been performed, the essential residues are circled.

The boxes indicate positions where conservative residue changes are observed. The dark gray box points out to the aromatic core, light grey boxes represent apolar residues and clear boxes acidic ones.

References, from top to bottom : (19, 28, 26, 56, 68, 69, 70, 27, 71, 72).

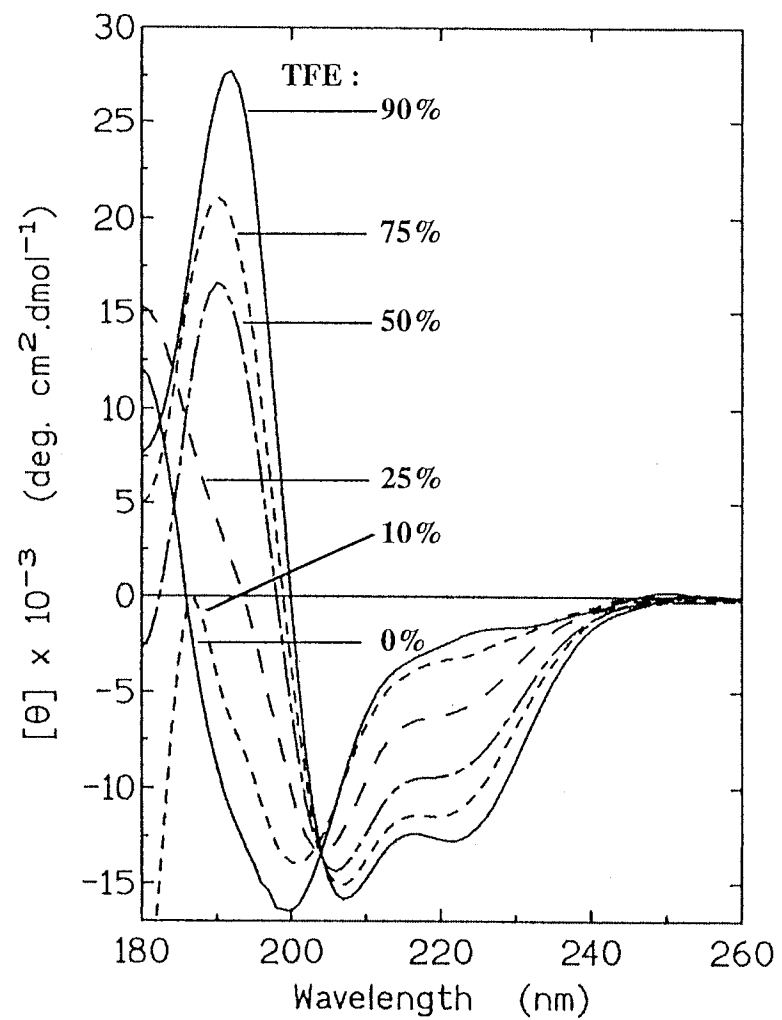


Figure 2

A

GAL4 AD1	MDGFYDQQVPFMVPVGKSRSEECRGRPVIDRKRKFLDSDLAH	DSEELFQDLSQLQEAWLAEAQVPDDEQFVPD
GAL4 Acidic Domain		DSEELFQDLSQLQEAWLAEAQVPDDEQFVPD
GAL4 helix		DSEELFQDLSQLQEAWLAEA
GAL4 helix-linker		DSEELFQDLSQLQEAWLAEAQSRAATAVELGTRSY

B

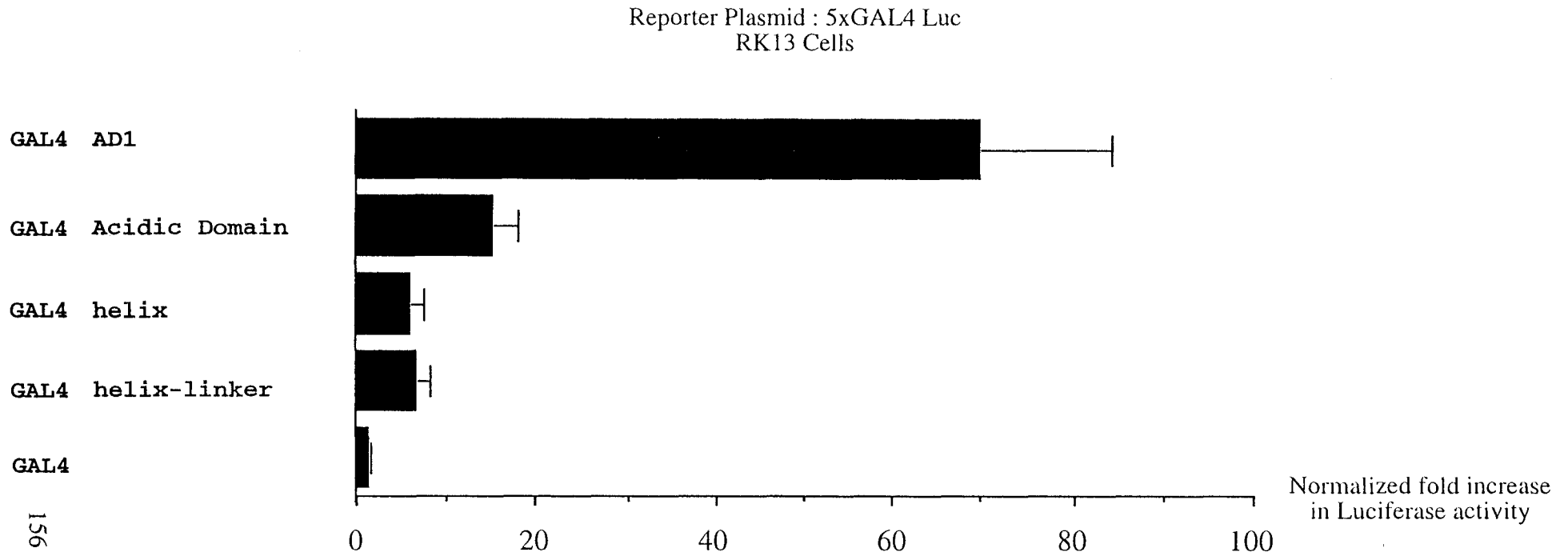
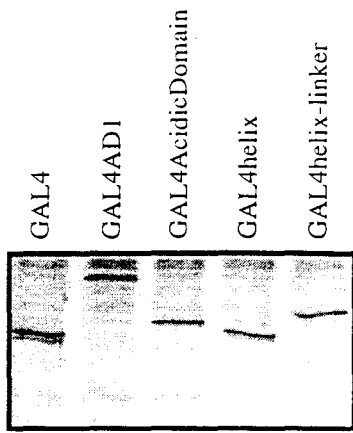


Figure 3

C



A

GAL4 AD1 MDGFYDQQVPPFMVPGKSRSEECRGRPVIDRKRKFLDSDLAHDSSEELFQDLSQLQEAWLAEAQVPDDEQFVPD
 GAL4 L50P -----P-----
 GAL4 E44S/D49A -----S-----A-----
 GAL4 F47L -----L-----
 GAL4 F47P -----P-----
 GAL4 F47W -----W-----
 GAL4 F69P -----P-----

B

Reporter Plasmid : 5xGAL4 Luc
RK13 Cells

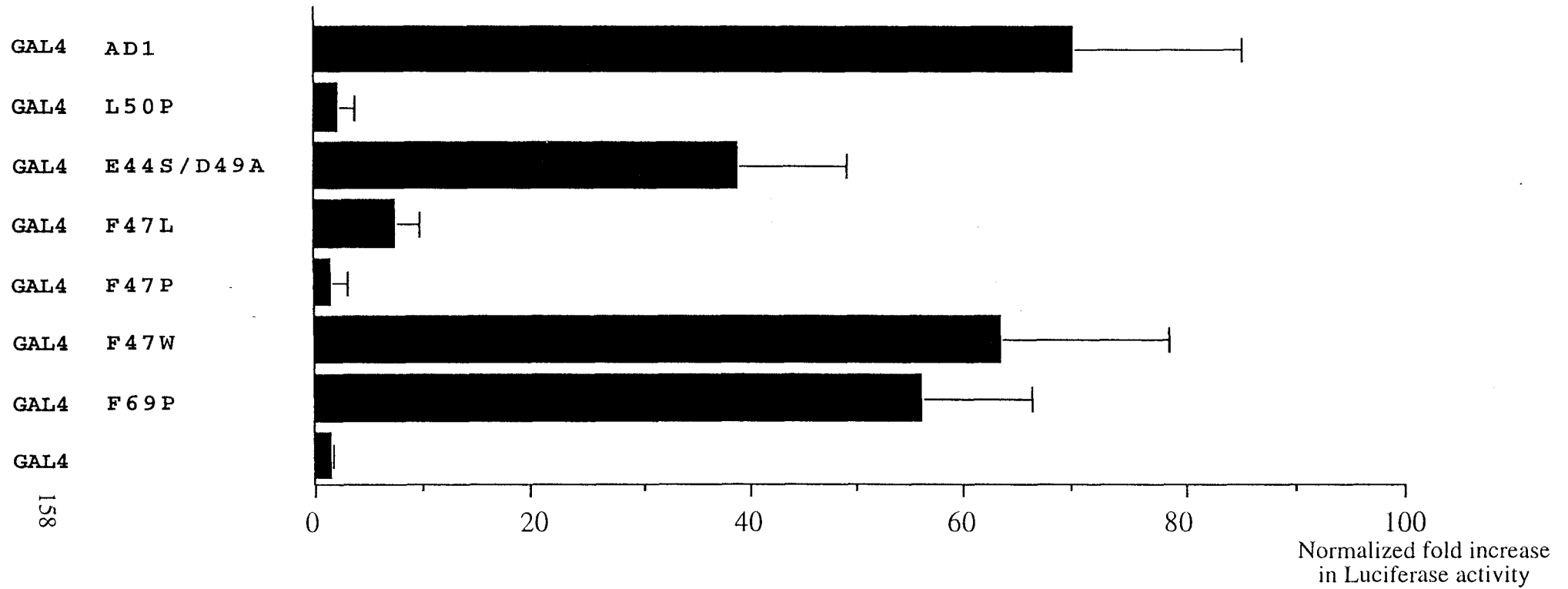
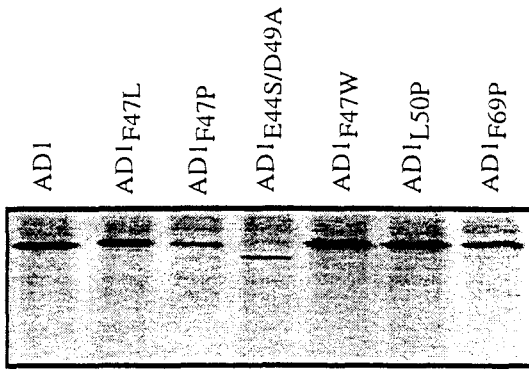
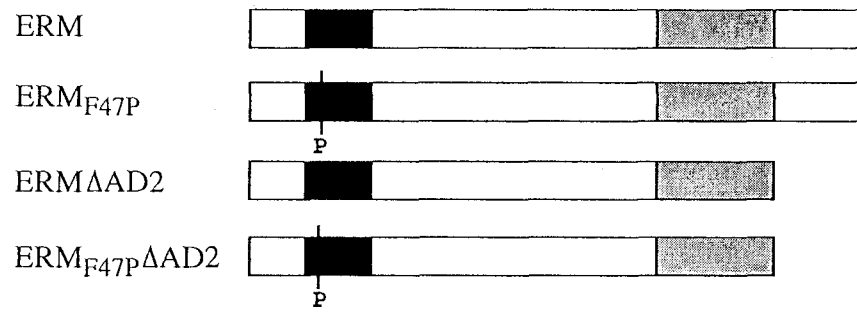


Figure 4

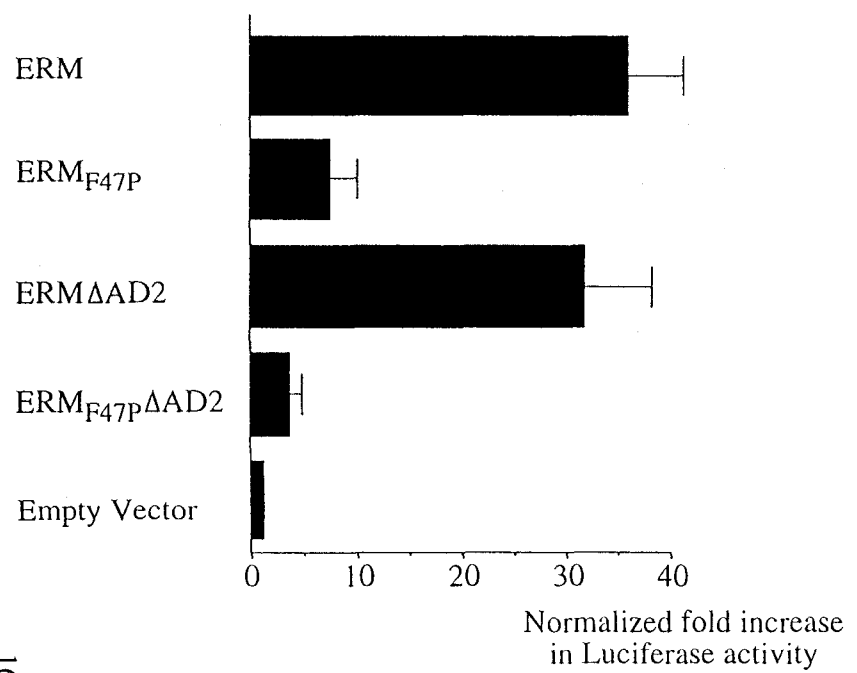
C



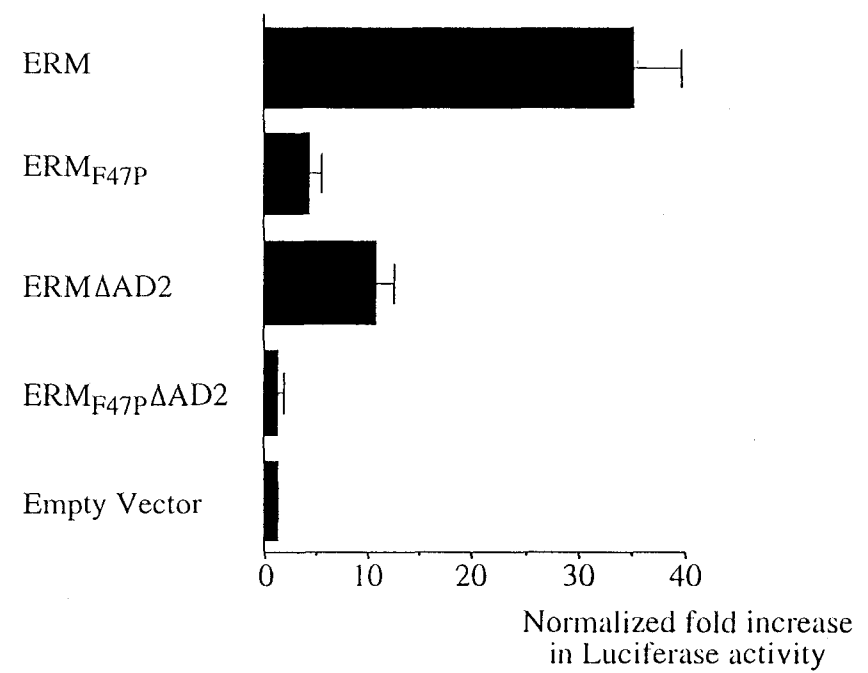


B

Reporter plasmid : ICAM(-176/-44) Luc
HC11 Cells



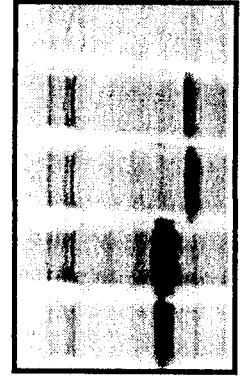
Reporter Plasmid : 3xTORU Luc
RK13 Cells



160

Figure 5

C



Empty Vector

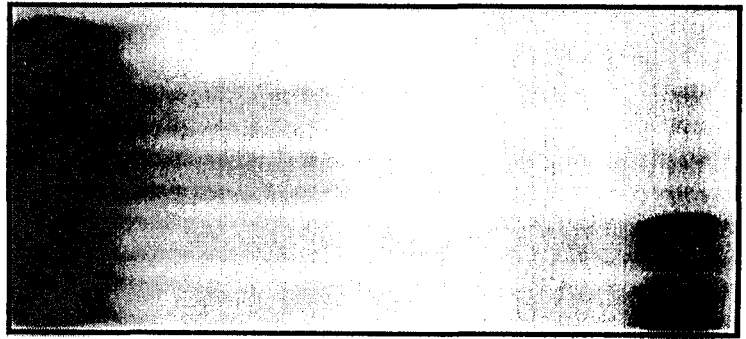
ERM

ERM_{F47P}

ERM Δ AD2

ERM_{F47P} Δ AD2

D



Empty Vector

ERM

ERM_{F47P}

ERM Δ AD2

ERM_{F47P} Δ AD2

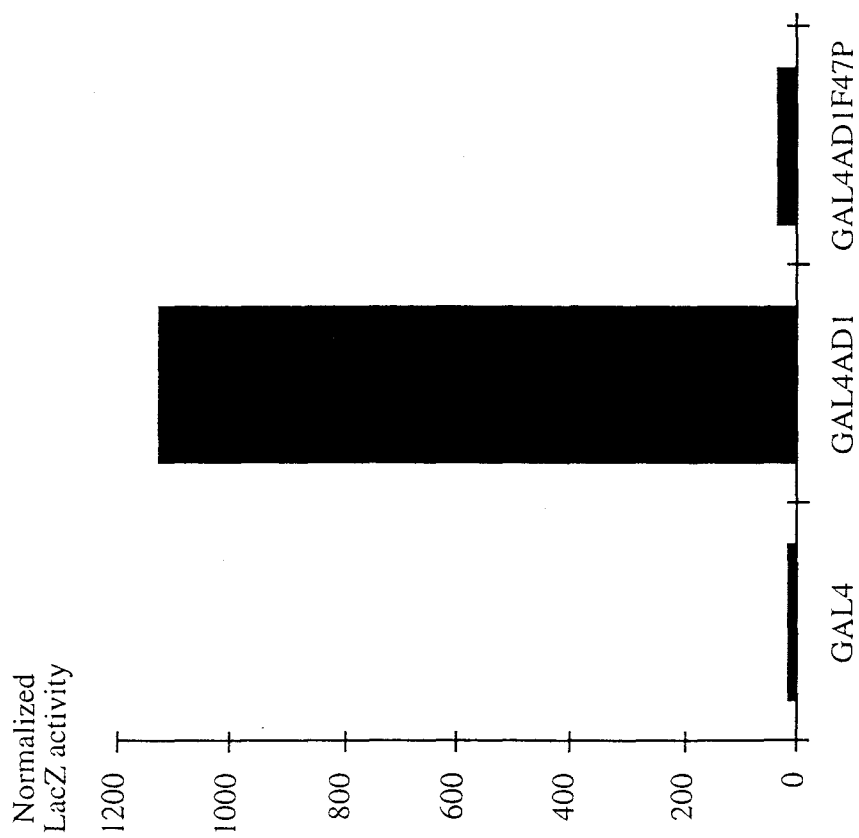


Figure 6

ERM
RelA
hGR
Ets2
RNAPolII
TEF1
HSF1
E2A
Jun
BPV E2

E	E	F	Q	D	L	S	Q	L	
D	E	F	S	S	I	A	D	M	
D	L	F	S	S	G	S	P	G	
D	I	L	W	E	H	L	E	Q	M
E	I	L	F	E	A	G	A	S	A
E	Q	S	F	Q	E	A	L	A	I
E	L	L	W	Q	E	N	M	M	A
L	L	D	F	S	M	M	P	L	P
A	E	G	F	V	R	A	E	L	E
I	D	Y	W	K	H	M	R	L	E

Figure 7

Discussion

1. Il existe des différences structurales entre les domaines activateurs acides

1.1 Préambule historique

Le contraste entre les données structurales obtenues pour les domaines de liaison à l'ADN et les domaines transactivateurs est saisissant. On sait que la conformation des premiers est clairement définie, et qu'elle répond à l'impératif d'interactions spécifiques protéines-ADN. La structure des seconds est bien plus floue, et le sobriquet de "blobs acides ou nouilles négatives" que l'on a utilisé pour définir les domaines transactivateurs acides a marqué les esprits. Cette absence apparente de structure bien définie semble difficile à concilier avec la fonction des domaines transactivateurs, qui contactent différentes protéines de la machinerie de transcription (Voir Hahn, 1993; Triezenberg, 1995, pour une discussion sur la structure des domaines transactivateurs).

Giniger et Ptashne (1987) ont montré qu'une hélice amphipatique avec un côté acide pouvait activer la transcription, suggérant que ceux des activateurs acides qui étaient structurés étaient peut-être hélicoïdaux. La description des domaines acides de GAL4 et GCN4 comme des feuilletts β était dans ce contexte particulièrement provocante (Leuther *et al.*, 1993; Van Hoy *et al.*, 1993). Elle est également controversée, et l'intérêt que nous portons à la structure des domaines activateurs mérite que nous exposions les termes du débat.

1.2 Quelle structure pour les domaines acides de GAL4 et GCN4?

Les observations du groupe de Johnston (Leuther *et al.*, 1993; Van Hoy *et al.*, 1993) vont à l'encontre d'une publication préalable (Hope et Struhl, 1986). Dans celle-ci, on montrait que la délétion progressive du domaine activateur de GCN4 menait à une décroissance régulière de son pouvoir transactivateur. Si ce domaine contenait un module structural important, comme le

feuillet β proposé par Johnston et ses collègues, on s'attendrait à une chute brutale de l'activité dès que ce motif serait touché.

Wu *et al.* (1996) proposent une explication pour cette discordance. Leuther *et al.* (1993), craignant de possibles artefacts liés à l'utilisation de chimères ont muté le domaine activateur dans le contexte de la protéine GAL4 sauvage, dont l'expression dépendait du promoteur naturel. Wu *et al.* (1996) arguent que dans ce cas, la protéine GAL4 ne sature pas ses sites de fixation. Ceci provoque des "effets de seuil" invalidant les conclusions quant au pouvoir transactivateur des mutants du domaine acide. De fait, en exprimant les protéines mutantes à un niveau suffisant, ils observent que la délétion progressive du domaine activateur de GAL4 entraîne une perte graduelle de son activité, comme pour GCN4. Il semble donc que, si un feuillet β est présent dans le domaine activateur de GAL4, son intégrité n'est pas nécessaire à la fonction transactivatrice.

1.3 Où en est-on?

L'état des lieux de nos connaissances structurales est donc assez vite dressé :

- certains activateurs contiennent des hélices α nécessaires à leur fonction. C'est le cas d'Erm, E2A (Massari *et al.*, 1996), GR (Dahlman-Wright *et al.*, 1995), RelA (Schmitz *et al.*, 1994). L'activateur artificiel AH forme également une hélice (Van Hoy *et al.*, 1993).
- GAL4 et GCN4 contiennent potentiellement des feuillets β , mais ceux-ci ne sont pas indispensables à la transactivation (voir ci-dessus).
- enfin, il est clair que la fonction de VP16 ne requiert pas la présence d'hélices α ou de feuillet β dans le domaine transactivateur (Cress et Triezenberg, 1991).

2. Les différences structurales établissent-elles des classes fonctionnelles?

L'hypothèse que nous suggérons dans cette publication est que la **diversité structurale** existant parmi les domaines de transactivation acide est en partie responsable de leur **diversité**

fonctionnelle. Pour en débattre, il est bon de rappeler certaines des données qui nous permettent d'affirmer l'hétérogénéité fonctionnelle des domaines acides, en précisant clairement les limites qui se posent à l'interprétation de ces expériences.

2.1 Certains activateurs interfèrent entre eux, mais ce n'est pas un renseignement précis

Il est clair que le critère de la séquence primaire, ici la richesse en résidus à caractère acide, est peu satisfaisant pour établir une classe de domaines de transactivation. Il l'est d'autant moins que dans plusieurs cas que les résidus acides ne sont pas indispensables à la fonction des transactivateurs acides (Hope et Struhl, 1986; Cress et Triezenberg, 1991; Drysdale *et al.*, 1995, par exemple).

Un critère plus fonctionnel est heureusement disponible : celui du "sqelching" (ou, en français, "interférence transcriptionnelle". Nous avons choisi, exceptionnellement, de garder le terme anglais, plus concis, et avec lequel les lecteurs sont sans doute familiers).

On dénomme ainsi les cas où un activateur non lié à l'ADN, en se liant à un composant de la machinerie de transcription, empêche la fonction d'un autre activateur qui contacte la même cible:

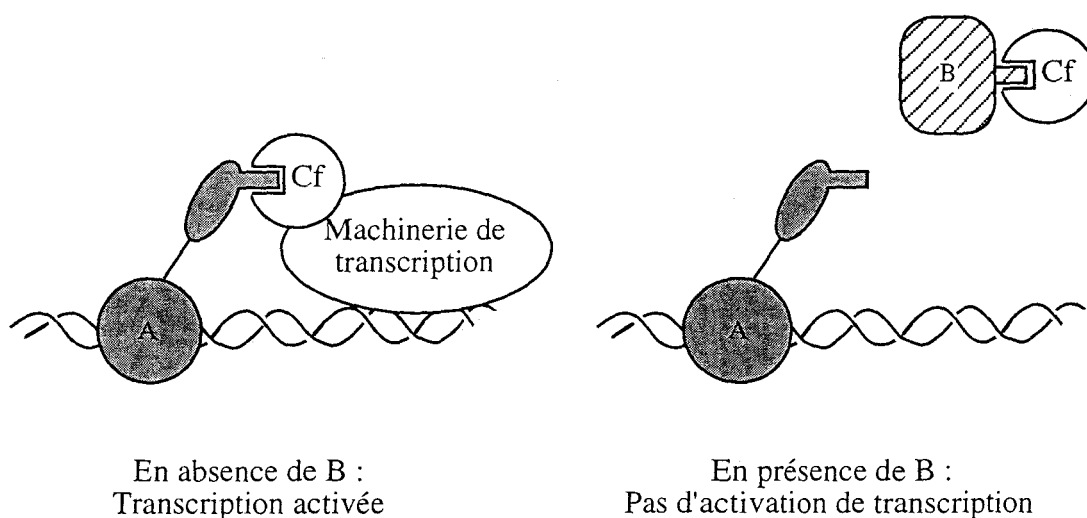


Fig. 35 Le sqelching : le transactivateur B gène l'action du transactivateur A en établissant une compétition pour un cofacteur limitant (Cf)

Il n'y a pas, à notre connaissance, d'étude extensive sur les capacités des domaines acides à interférer l'un avec l'autre. Nous pouvons juste proposer un rappel de quelques-uns des résultats obtenus :

Interférence	Pas d'interférence
Domaine N-terminal de C-Jun et VP16, Gal4 et JunB Oehler et Angel (1992)	Rta et VP16 Hardwick <i>et al.</i> (1992)
AD1 d'Erm et VP16 Laget <i>et al.</i> (1996)	
AD1 de RelA et VP16 Schmitz <i>et al.</i> (1994), Blair <i>et al.</i> (1994)	
AD1 de RelA et AD2 de RelA Schmitz <i>et al.</i> (1995)	

Fig. 36 Qui interfère avec qui?

Pour ne pas rendre la lecture trop fastidieuse, nous nous sommes permis un abus de langage: "VP16" signifie en fait "région activatrice acide de VP16".
AD : Activation Domain, Rta : transactivateur R du virus d'Epstein-Barr
RelA est aussi appelé p65 NFkB

Il y a donc un premier critère pour affirmer que deux activateurs acides n'ont pas le même fonctionnement, c'est qu'ils n'interfèrent pas l'un avec l'autre. Cependant, nous souhaitons montrer que la conclusion réciproque n'est pas valide, et que **ce n'est pas parce que deux domaines acides interfèrent l'un avec l'autre qu'ils fonctionnent de façon identique**. En effet, pour des raisons stériques, un activateur peut interférer avec un autre, sans avoir exactement le même mode d'action :

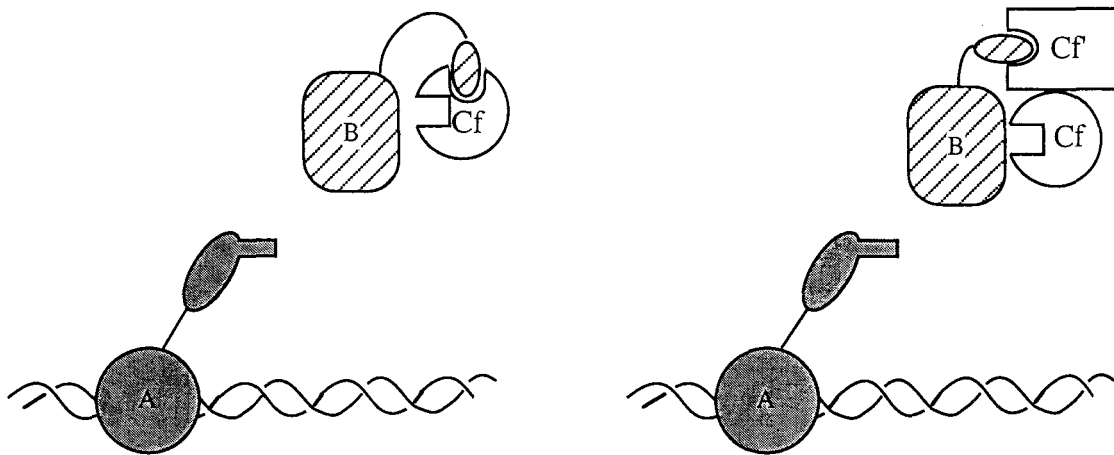


Fig. 37 Deux cas où B quelche A sans avoir les mêmes cibles

Par conséquent nos conclusions sont :

- des activateurs acides qui n'interfèrent pas entre eux (par exemple Rta et VP16), ont sans doute un mode d'action distinct,
- par contre des activateurs qui interfèrent entre eux (par exemple, AD1 d'ERM et VP16) n'ont pas nécessairement le même mode d'action.

Aussi, nous jugeons que le quelching n'est pas une donnée fonctionnelle très précise, et que d'autres tests fonctionnels sont nécessaires pour classifier les domaines acides.

2.2 Il existe des différences fonctionnelles plus précises entre les domaines acides

La première est que **tous les activateurs ne sont pas également affectés par la disparition de certains cofacteurs**. Ainsi, si l'on supprime chez la levure le gène du coactivateur ADA2, l'activité transactivatrice du domaine acide de VP16 est fortement réduite, tandis que celle de HAP4 n'est que peu modifiée (Fig. 38).

Expérience

Observations

Délétion du gène du cofacteur ADA2 Berger et al (1992)	L'activation par les domaines acides de VP16, GAL4 et GCN4 est fortement réduite	L'activation par le domaine acide de HAP4 n'est pas réduite
Test de stimulation de l'élongation transcriptionnelle Ghosh et al (1996)	Le sous-domaine C de la région acide de VP16 est inactif	Le sous-domaine N est actif
Surexpression du cofacteur Sub1 Knaus et al (1996)	L'activation par GAL4 n'est pas modifiée	L'activation par GCN4 et le complexe HAP est augmentée
Ajout du cofacteur PC2 lors de la transcription <i>in vitro</i> Kretzschmar et al (1994)	L'activation par le domaine acide de VP16 n'est pas modifiée	L'activation par l'activateur acide AH est augmentée

Fig. 38 Il existe des différences fonctionnelles entre les transactivateurs "acides"
La liste n'est nullement exhaustive.

Encore une fois, nous souhaitons préciser que **même si la perte d'un coactivateur affecte simultanément l'activité de deux domaines transactivateurs, on ne peut en conclure qu'ils ont le même mode d'action**. En effet, ils pourraient :

- interagir avec des régions différentes du coactivateur (comme schématisé dans la partie gauche de la figure ci-dessus),
- interagir avec deux protéines différentes d'un même complexe (partie droite du schéma). La délétion d'une de ces protéines pourrait déstabiliser le complexe, et donc rendre inactifs les deux activateurs.

En résumé, si l'on peut conclure que les domaines acides de VP16 et de HAP4 ne fonctionnent pas de manière similaire car ils n'ont pas la même sensibilité à la perte d'ADA2 (Berger *et al.*, 1992), il est risqué d'affirmer que VP16 et GCN4, tous deux affectés, utilisent le même mécanisme pour activer la transcription. Pour établir que deux activateurs utilisent de façon identique un coactivateur, il est plus concluant d'obtenir des **mutations ponctuelles** du coactivateur qui affectent simultanément les deux activateurs. Il faut bien sûr qu'il ne s'agisse pas de mutations "triviales" qui inactivent totalement ce coactivateur.

D'autres différences fonctionnelles entre domaines acides ont été rapportées (fig. 38).

2.3 Il est encore trop tôt pour établir un parallèle structure-fonction

C'est cette conclusion mesurée, et sans doute un peu décevante pour le lecteur, qu'il semble falloir adopter. En effet, nous ne connaissons pas d'étude qui ait visé à définir précisément les classes fonctionnelles d'activateurs acides par certains des moyens évoqués ci-dessus. L'établissement des classes structurales est également au stade préliminaire.

La démonstration, ou l'infirmité de notre théorie ne pourra avoir lieu qu'avec l'accroissement des connaissances dans ces deux domaines. Soulignons, toutefois, qu'on ne peut écarter la possibilité que les relations structure-fonction dans les domaines activateurs soient plus

complexes que notre modèle ne le laisse supposer. Exposons cette possibilité, à titre de conclusion.

3. Des domaines à structure non définie pourraient-ils activer la transcription?

Nous suggérons que les classes fonctionnelles reflètent des classes structurales. Est-il possible que cette proposition soit fausse, et que les domaines activateurs puissent contacter la machinerie de transcription sans avoir de structure secondaire définie? La réponse est "oui", nous ne souhaitons pas le nier.

On peut en effet citer au moins deux processus biologiques où des interactions spécifiques reposent sur des domaines qui visiblement n'ont pas de structure secondaire organisée :

- les signaux de localisation nucléaire (Nuclear Localization Signals, ou NLS) sont des déterminants d'une dizaine d'acides aminés dans la séquence des protéines qui leur permettent d'être importées du cytoplasme vers le noyau. Ce mouvement dépend de l'interaction du NLS avec une protéine présente dans les pores nucléaires. Le consensus pour la séquence des NLS est extrêmement flou, sa caractéristique la plus saillante étant la richesse en Lysine et Arginine. Il n'y a pour l'instant aucune preuve que la région du NLS interagissant avec les protéines du pore nucléaire ait une structure secondaire définie (pour une définition du consensus voir Jans, 1995; Makkerh *et al.*, 1996).

- enfin, considérons un exemple bien plus distant, au moins géographiquement, car il ne se tient pas dans le noyau. Les interactions entre un antigène protéique et un anticorps sont spécifiques. Cependant elles mettent en jeu des portions courtes de l'antigène (10 à 15 acides aminés), qui n'ont nullement besoin d'avoir une structure secondaire définie.

Nous ne négligeons donc pas la possibilité que les interactions activateurs-machinerie de transcription puissent avoir lieu sans que les premiers aient une structure secondaire définie.

CONCLUSION

Il est difficile d'être précis sans être trop long, et de rendre compte de la complexité sans être obscur. Nous espérons dans ces deux domaines être arrivés à un compromis acceptable.

L'étude des mécanismes d'activation transcriptionnelle par Erm n'est pas une fin en soi, et notre "petit pan de mur jaune" n'a pas grand sens s'il n'est placé dans un paysage plus vaste, et étayé par les données acquises dans d'autres domaines.

Nous pensons que l'analyse "structurale" d'Erm que nous avons menée s'inscrit dans une logique dictée par plusieurs interrogations :

- quels sont les partenaires de régulation d'Erm?

- comment son activité est-elle modulée?

- quels sont ses gènes-cibles?

Voilà certes un vaste programme, mais nous devrions, petit à petit, obtenir des débuts de réponse à ces questions. Le phénotype des souris dont les gènes *erm* et/ou *er81* et/ou *pea3* seront inactivés par recombinaison homologue sera une source d'informations fondamentales sur les rôles physiologiques des gènes du groupe PEA3 et leur éventuelle redondance. Elles seront interprétables grâce aux données acquises sur les patrons d'expression spatio-temporelle de ces trois gènes. Il faudra en parallèle établir des systèmes modèles *in vitro*, où les régulations transcriptionnelles et post-traductionnelles des gènes seront étudiées en réponse à des stimuli bien caractérisés.

Erm est, bien sûr, un morceau de choix pour la recherche de partenaires d'interaction par la technique de double-hybride chez la levure. Les mêmes constructions utilisées astucieusement pourront également nous éclairer sur les régions impliquées dans le maintien de la conformation fermée d'Erm, ou sur ses coactivateurs transcriptionnels.

Passé l'inévitable débroussaillage préliminaire, les années qui viennent seront, n'en doutons pas, riches en enseignements sur la fonction des gènes du groupe PEA3.

Je les suivrai de loin, mais avec intérêt...

REFERENCES

- Albagli, O., Klaes, A., Ferreira, E., Leprince, D. et Klämbt, C.** (1996). Function of ets genes is conserved between vertebrates and *Drosophila*. **Mechanisms of Development** *In press*,
- Apone, L. M., Virbasius, C. M. A., Reese, J. C. et Green, M. R.** (1996). Yeast TAF_{II}90 is required for cell-cycle progression through G2/M but not for general transcription activation. **Genes Dev.** 10, 2368-2380.
- Baichwal, V. R., Park, A. et Tjian, R.** (1991). v-Src and EJ Ras alleviate repression of c-Jun by a cell-specific inhibitor. **Nature** 352, 165-168.
- Bailly, R. A., Bosselut, R., Zucman, J., Cormier, F., Delattre, O., Roussel, M., Thomas, G. et Ghysdael, J.** (1994). DNA-binding and transcriptional activation properties of the EWS-FLI-1 fusion protein resulting from the t(11;22) translocation in Ewing Sarcoma. **Mol. Cell. Biol.** 14, 3230-3241.
- Barberis, A., Pearlberg, G., Simkovich, N., Farrell, S., Reinagel, P., Bamdad, C., Sigal, G. et Ptashne, M.** (1995). Contact with a component of the polymerase II holoenzyme suffices for gene activation. **Cell** 81, 359-368.
- Bassuk, A. G. et Leiden, J. M.** (1995). A direct physical association between ETS and AP-1 transcription factors in normal human T cells. **Immunity** 3, 223-37.
- Battacharya, S., Eckner, R., Grossman, S., Oldread, E., Arany, Z., D'Andrea, A. et Livingston, D. M.** (1996). Cooperation of Stat2 and p300/CBP in signalling induced by interferon- α . **Nature** 383, 344-347.
- Ben-David, Y. et Bernstein, A.** (1991). Friend virus-induced erythroleukemia and the multistage nature of cancer. **Cell** 66, 831-834.
- Ben-David, Y., Giddens, E. B. et Bernstein, A.** (1990). Identification and mapping of a common proviral integration site Fli-1 in erythroleukemia cells by Friend murine leukemia virus. **Proc. Natl. Acad. Sci. USA** 87, 1332-1336.
- Ben-David, Y., Giddens, E. B., Letwin, K. et Bernstein, A.** (1991). Erythroleukemia induction by friend murine leukemia virus : insertional activation of a new member of the ets gene family, fli-1, closely linked to c-ets1. **Genes Dev.** 5, 908-918.
- Bentley, D. L.** (1995). Regulation of transcriptional elongation by RNA polymerase II. **Curr. Op. Gen. Devt** 5, 210-216.
- Bentley, D. L. et Groudine, M.** (1986). A block to elongation is largely responsible for decreased transcription of c-myc in differentiated HL60 cells. **Nature** 321, 702-706.

Berger, S. L., Pina, B., Silverman, N., Marcus, G. A., Agapite, J., Regier, J. L., Triezenberg, S. J. et Guarente, L. (1992). Genetic isolation of ADA2 : a potential transcriptional adaptor required for function of certain acidic activation domains. *Cell* 70, 251-265.

Bertolotti, A., Lutz, Y., Heard, D. J., Chambon, P. et Tora, L. (1996). hTAF_{II}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.

Bhat, N. K., Thompson, C. B., Lindsten, T., June, C. H., Fujiwara, S., Koizumi, S., Fisher, R. J. et Papas, T. S. (1990). Reciprocal expression of human ets1 and ets2 genes during T-cell activation : regulatory role for the proto-oncogene ets1. *Proc. Natl. Acad. Sci. USA* 87, 3723-3727.

Bjorklund, S. et Kim, Y. J. (1996). Mediator of transcriptional regulation. *Trends Biochem. Sci.* 21, 335-337.

Blair, W. S., Bogerd, H. P., Madore, S. J. et Cullen, B. R. (1994). Mutational analysis of the transcription activation domain of RelA : identification of a highly synergistic minimal acidic activation module. *Mol. Cell. Biol.* 14, 7226-7234.

Blau, J., Xiao, H., McCracken, S., O'Hare, P., Greenblatt, J. et Bentley, D. (1996). Three functional classes of transcription activation domains. *Mol. Cell. Biol.* 16, 2044-2055.

Bohmann, D., Ellis, M. C., Staszewski, L. M. et Mlodzik, M. (1994). Drosophila jun mediates ras-dependent photoreceptor determination. *Cell* 78, 973-986.

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

Bottinger, E. P., Shelley, C. S., Farokhzad, O. C. et Arnaout, M. A. (1994). The human beta 2 integrin CD18 promoter consists of two inverted Ets cis elements. *Mol. Cell. Biol.* 14, 2604-2615.

Bradbury, E. M. (1992). Reversible histone modifications and the chromosome cell cycle. *BioEssays* 14, 9-16.

Braun, B. S., Frieden, R., Lessnick, S. L., May, W. A. et Denny, C. T. (1995). Identification of target genes for the Ewing's sarcoma EWS/FLI fusion protein by representational difference analysis. *Mol. Cell. Biol.* 15, 4623-30.

- Braunstein, M., Sobel, R. E., Allis, C. D., Turner, B. M. et Broach, J. R.** (1996). Efficient transcriptional silencing in *Saccharomyces cerevisiae* requires a heterochromatin histone acetylation pattern. *Mol. Cell. Biol.* 16, 4349-4356.
- Brown, T. A. et McKnight, S. L.** (1992). Specificities of protein-protein and protein-DNA interaction of GABP α and two newly defined *ets*-related proteins. *Genes Dev.* 6, 2502-2512.
- Brownell, J. E. et Allis, C. D.** (1996). Special HATs for special occasions : linking histone acetylation to chromatin assembly and gene activation. *Curr. Op. Gen. Dev.* 6, 176-184.
- Brownell, J. E., Zhou, J., Ranalli, T., Kobayashi, R., Edmondson, D. G., Roth, S. Y. et Allis, C. D.** (1996). Tetrahymena histone acetyltransferase A : a homolog to yeast GCN5p linking histone acetylation to gene activation. *Cell* 84, 843-851.
- Brunner, D., Dücker, K., Oellers, N., E., H., Scholz, H. et Klämbt, C.** (1994). The ETS domain protein pointed-P2 is a target of MAP kinase in the sevenless signal transduction pathway. *Nature* 370, 386-389.
- Burley, S. K. et Roeder, R. G.** (1996). Biochemistry and structural biology of transcription factor IID (TFIID). *Annu. Rev. Biochem.* 65, 769-799.
- Buttice, G., Duterque-Coquillaud, M., Basuyaux, J. P., Carrère, S., Kurkinen, M. et Stéhelin, D.** (1996). Erg, an Ets-family member, differentially regulates human Collagenase1 (MMP1) and Stromelysin1 gene expression by physically interacting with the Fos/Jun complex. *Oncogene* *In press*.
- Cahill, M. A., Janknecht, R. et Nordheim, A.** (1996). Jack of all cascades. *Curr. Biol.* 6, 16-19.
- Carlson, M. et Laurent, B. C.** (1994). The SWI/SNF family of global transcriptional activators. *Curr. Op. Cell. Biol.* 6, 296-302.
- Chao, D. M., Gadbois, E. L., Murray, P. J., Anderson, S. F., Sonu, M. S., Parvin, J. D. et Young, R. A.** (1996). A mammalian SRB protein associated with an RNA polymerase II holoenzyme. *Nature* 380, 82-85.
- Chatterjee, S. et Struhl, K.** (1995). Connecting a promoter-bound protein to TBP bypasses the need for a transcriptional activation domain. *Nature* 374, 820-822.
- Chen, J.-L., Attardi, L. D., Verrijzer, C. P., Yokomori, K. et Tjian, R.** (1994). Assembly of recombinant TFIID reveals different coactivator requirements for distinct transcriptional activators. *Cell* 79, 93-105.

- Chi, T. et Carey, M.** (1993). The ZEBRA activation domain : modular organization and mechanism of action. *Mol. Cell. Biol.* 13, 7045-7055.
- Choy, B. et Green, M. R.** (1993). Eukaryotic activators function during multiple steps of preinitiation complex assembly. *Nature* 366, 531-536.
- Coté, J., Quinn, J., Workman, J. L. et Peterson, C. L.** (1994). Stimulation of a GAL4 derivative binding to nucleosomal DNA by the yeast SWI/SNF complex. *Science* 265, 53-60.
- Cress, W. D. et Triezenberg, S. J.** (1991). Critical structural elements of the VP16 transcriptional activation domain. *Science* 251, 87-90.
- Dahlman-Wright, K., Bauman, H., McEwan, I. J., Almlöf, T., Wright, A. P. H., Gustafsson, J.-A. et Hard, T.** (1995). Structural characterization of a minimal functional transactivation domain from the human glucocorticoid receptor. *Proc. Natl. Acad. Sci. USA* 92, 1699-1703.
- Dahmus, M. E.** (1996). Reversible phosphorylation of the C-terminal domain of RNA polymerase II. *J. Biol. Chem.* 271, 19009-19012.
- Dalton, S. et Treisman, R.** (1992). Characterization of SAP-1, a protein recruited by Serum Response Factor to the c-fos Serum Response Element. *Cell* 68, 597-612.
- Das, G., Hinkley, C. S. et Herr, W.** (1995). Basal promoter elements as a selective determinant of transcriptional activator function. *Nature* 374, 657-660.
- de Launoit, Y., Audette, M., Plaza, S. et Baert, J. L.** (Soumis). The ETS transcription factor ERM activates InterCellular Adhesion Molecule-1 transcription.
- Delannoy-Courdent, A., Fauquette, W., Dong-Le Bourhis, X. F., Vandebunder, B., Boilly, B. et Desbiens, X.** (soumis). Reciprocal interactions trigger c-ets-1 and uPA expressions in both epithelial and mesenchymal cells during normal or pathological development of the mammary gland.
- Delattre, O., Zucman, J., Melot, T., Garau, X. S., Zucker, J. M., Lenoir, G. M., Ambros, P. F., Sheer, D., Turc, C. C., Triche, T. J. et al.** (1994). The Ewing family of tumors-a subgroup of small-round-cell tumors defined by specific chimeric transcripts. *N Engl J Med* 331, 294-9.
- Delattre, O., Zucman, J., Plougastel, B., Desmaze, C., Melot, T., Peter, M., Kovar, H., Joubert, I., de Jong, P., Rouleau, G., Aurias, A. et Thomas, G.** (1992). Gene fusion with an

ETS DNA-binding domain caused by chromosome translocation in human tumors. **Nature** 359, 162-165.

Delgado, M. D., Hallier, M., Menecoeur, T., Tavitian, A. et Moreau-Gachelin, F. (1994). Inhibition of Friend cells proliferation by spi-1 antisense nucleotides. **Oncogene** 9, 1723-1727.

Desbiens, X., Queva, C., Jaffredo, T., Stéhelin, D. et Vandebunder, B. (1991). The relationship between cell proliferation and the transcription of the nuclear oncogenes c-myc, c-myb and c-ets-1 during feather morphogenesis in the chick embryo. **Development** 111, 699-713.

Desbois, C., Aubert, D., Legrand, C., Pain, B. et Samarut, J. (1991). A novel mechanism of action for v-erbA : abrogation of the inactivation of transcription factor AP1 by retinoic acid and thyroid hormone receptors. **Cell** 67, 731-740.

Dickson, B. (1995). Nuclear factors in sevenless signalling. **Trends Genet.** 11, 106-11.

Dikstein, R., Ruppert, S. et Tjian, R. (1996a). TAF_{II}250 is a bipartite protein kinase that phosphorylates the basal transcription factor RAP74. **Cell** 84, 781-90.

Dikstein, R., Zhou, S. et Tjian, R. (1996b). Human TAF_{II}105 is a cell-type specific TFIID subunit related to hTAF_{II}130. **Cell** 87, 137-146.

Dingwall, C. et Laskey, R. A. (1991). Nuclear targeting sequences-a consensus? **Trends Biochem. Sci.** 16, 478-481.

Donaldson, L. W., Petersen, J. M., Graves, B. J. et McIntosh, L. P. (1996). Solution structure of the ETS domain from murine ets-1: a winged helix-turn-helix DNA binding motif. **EMBO J.** 15, 125-134.

Drysdale, C. M., Duenas, E., Jackson, B. M., Reusser, U., Braus, G. R. et Hinnebusch, A. G. (1995). The transcriptional activator GCN4 contains multiple activatin domains that are critically dependent on hydrophobic amino acids. **Mol. Cell. Biol.** 15, 1220-1233.

Dubnau, J. et Struhl, G. (1996). RNA recognition and translational regulation by a homeodomain protein. **Nature** 379, 694-9.

Dubrovskaya, V., Lavigne, A. C., Davidson, I., Acker, J., Staub, A. et Tora, L. (1996). Distinct domains of hTAF_{II}100 are required for functional interacton with transcription factor TFIIF β (RAP30) and incorporation into the TFIID complex. **EMBO J.** 14, 3702-3712.

Eisenbeis, C. F., Singh, H. et Storb, U. (1995). Pip, a novel IRF family member, is a lymphoid specific, Pu.1-dependent transcriptional activator. **Genes Dev.** 9, 1377-1387.

- Emami, K. et Carey, M.** (1992). A synergistic increase in potency of a multimerized VP16 activation domain. *EMBO J.* 11, 5005-5012.
- Emami, K. H., Navarre, W. W. et Smale, S. T.** (1995). Core promoter specificities of the Sp1 and VP16 activation domains. *Mol. Cell. Biol.* 15, 5906-5916.
- Fafeur, V., Tulasne, D., Quéva, C., Vercamer, C., Dimster, V., Mattot, V., Stéhelin, D., Desbiens, X. et Vandebunder, B.** (soumis). The ETS1 transcription factor is expressed during epithelial-mesenchymal transitions in the chick embryo and is activated in Scatter-Factor-stimulated MDCK epithelial cells.
- Farrell, S., Simkovich, N., Wu, Y., Barberis, A. et Ptashne, M.** (1996). Gene activation by recruitment of RNA polymerase II holoenzyme. *Genes Dev.* 10, 2359-2367.
- Fisher, R. J., Mavrothalassitis, G., Kondoh, A. et Papas, T. S.** (1991). High-affinity protein-DNA interactions of the cellular ETS1 protein : determination of the ETS binding motif. *Oncogene* 6, 2249-2254.
- Galson, D. L., Hensold, J. O., Bishop, T. R., Schalling, M., D'Andrea, A. D., Jones, C., Auron, P. E. et Housman, D. E.** (1993). Mouse β -globin DNA-binding protein B1 is identical to a proto-oncogene, the transcription factor Spi-1/Pu.1, and is restricted in expression to hematopoietic cells and the testis. *Mol. Cell. Biol.* 13, 2929-2941.
- Ghosh, S., Toth, C., Peterlin, B. M. et Seto, E.** (1996). Synergistic activation of transcription by the mutant and wild-type minimal transcriptional activation domain of VP16. *J. Biol. Chem.* 271, 9911-9918.
- Giese, K., Kingsley, C., Kirshner, J. R. et Grosschedl, R.** (1995). Assembly and function of a TCR α enhancer complex is dependent on LEF-1-induced DNA bending and multiple protein-protein interactions. *Genes Dev.* 9, 985-1008.
- Giniger, E. et Ptashne, M.** (1987). Transcription in yeast activated by a putative amphipatic alpha helix linked to a DNA binding unit. *Nature* 330, 670-672.
- Giovane, A., Pintzas, A., Maira, S. M., Sobieszczuk, P. et Wasylyk, B.** (1994). Net, a new *ets* transcription factor that is activated by Ras. *Genes Dev.* 8, 1502-1513.
- Glineur, C., Bailly, M. et Ghysdael, J.** (1989). The c-erb-a encoded thyroid hormone receptor is phosphorylated in its amino-terminal domain by Casein Kinase II. *Oncogene* 4, 1247-1254.
- Golay, J., Introna, M. et Graf, T.** (1988). A single point mutation in the v-ets oncogene affects both erythroid and myelomonocytic cell differentiation. *Cell* 55, 1147-58.

- Grévin, D., Chen, J. H., Raes, M. B., Stéhelin, D., Vandebunder, B. et Desbiens, X.** (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.
- Gu, W. et Reines, D.** (1995). Identification of a decay in transcription potential that results in elongation factor dependence of RNA polymerase II. *J. Biol. Chem.* 270, 11238-44.
- Guddat, U., Bakken, A. H. et Pieler, T.** (1990). Protein-mediated nuclear export of RNA : 5S rRNA containing small RNPs in *Xenopus* oocytes. *Cell* 60, 619-628.
- Gutman, A. et Wasylyk, B.** (1990). The collagenase gene promoter contains a TPA and oncogene-responsive unit encompassing the PEA3 and AP-1 binding sites. *EMBOJ.* 9, 2241-6.
- Hagman, J. et Grosschedl, R.** (1992). An inhibitory carboxyl-terminal domain in Ets-1 and Ets-2 mediates differential binding of ETS family factors to promoter sequences of the *mb-1* gene. *Proc. Natl. Acad. Sci. USA* 89, 8889-8893.
- Hahn, S.** (1993). Structure (?) and function of acidic transcriptional activators. *Cell* 72, 481-483.
- Hahn, S. L. et Wasylyk, B.** (1994). The oncoprotein v-ETS is less selective in DNA binding than c-ETS-1 due to the C-terminal sequence change. *Oncogene* 9, 2499-2512.
- Halle, J. P. et Meisterernst, M.** (1996). Increasing evidence for a transcriptosome. *Trends Genet.* 12, 161-163.
- Hallier, M., Tavitian, A. et Moreau, G. F.** (1996). The transcription factor Spi-1/PU.1 binds RNA and interferes with the RNA-binding protein p54nrb. *J. Biol. Chem.* 271, 11177-81.
- Hanks, M., Wurst, W., Anson, C. L., Auerbach, A. B. et Joyner, A. L.** (1995). Rescue of the *En-1* mutant phenotype by replacement of *En-1* with *En-2*. *Science* 269, 679-82.
- Hardwick, J. M., Tse, L., Applegren, N., Nicholas, J. et Veluona, M.** (1992). The Epstein-Barr virus R transactivator (Rta) contains a complex, potent activation domain with properties different from those of VP16. *J. Virol.* 66, 5500-5508.
- Hayashi, S. et Scott, M. P.** (1990). What determines the specificity of action of Drosophila homeodomain proteins? *Cell* 63, 883-894.
- Hecht, A., Laroche, T., Strahl-Bolsinger, S., Gasser, S. M. et Grunstein, M.** (1995). Histone H3 and H4 N-termini interact with SIR3 and SIR4 proteins : a molecular model for the formation of heterochromatin in yeast. *Cell* 80, 583-592.

Herrmann, C. H. et Rice, A. P. (1995). Lentivirus Tat proteins specifically associate with a cellular protein kinase, TAK, that hyperphosphorylates the Carboxy-terminal domain of the large subunit of RNA polymerase II : candidate for a Tat cofactor. **J. Virol.** 69, 1612-1620.

Herschlag, D. et Johnson, F. B. (1993). Synergism in transcriptional activation : a kinetic view. **Genes Dev.** 7, 173-179.

Higashino, F., Yoshida, K., Noumi, T., Seiki, M. et Fujinaga, K. (1995). Ets-related protein E1A-F can activate three different matrix metalloproteinase gene promoters. **Oncogene** 10, 1461-1463.

Hipskind, R. A., Buscher, D., Nordheim, A. et Baccarini, M. (1994). Ras/MAP kinase-dependent and independent signaling pathways target distinct ternary complex factors. **Genes Dev.** 8, 1803-1816.

Hipskind, R. A., Rao, V. N., Mueller, C. G. F., Reddy, E. S. P. et Nordheim, A. (1991). Ets-related protein Elk-1 is homologous to the c-fos regulatory factor p62. **Nature** 354, 531-534.

Hirschorn, J. N., Brown, S. A., Clark, C. D. et Winston, F. (1992). Evidence that SNF2/SWI2 and SNF5 activate transcription by altering chromatin structure. **Genes Dev.** 6, 2288-2298.

Hoffmann, A., Chiang, C. M., Oelgeschlager, T., Xie, X., Burley, S. K., Nakatani, Y. et Roeder, R. G. (1996). A histone octamer-like structure within TFIID. **Nature** 380, 356-9.

Hope, I. A. et Struhl, K. (1986). Functional dissection of a eukaryotic transcriptional activator protein, GCN4 of yeast. **Cell** 46, 885-894.

Horikoshi, M., Carey, M. F., Kakidani, H. et Roeder, R. G. (1988a). Mechanism of action of a yeast activator : direct effect of GAL4 derivatives on mammalian TFIID-promoter interactions. **Cell** 54, 665-669.

Horikoshi, M., Hai, T., Lin, Y. S., Green, M. R. et Roeder, R. G. (1988b). Transcription factor ATF interacts with the TATA factor to facilitate establishment of a preinitiation complex. **Cell** 54, 1033-1042.

Huang, J., Blackwell, T. K., Kedes, L. et Weintraub, H. (1996). Differences between MyoD DNA-binding and activation site requirements revealed by functional random site selection. **Mol. Cell. Biol.** 16, 3893-3900.

Imbalzano, A. N., Kwon, H., Green, M. R. et Kingston, R. E. (1994). Facilitated binding of the TATA-binding protein to nucleosomal DNA. **Nature** 370, 481-485.

- Jackson, S. P. et Tjian, R.** (1988). O-glycosylation of eukaryotic transcription factors: implications for mechanisms of transcriptional regulation. *Cell* 55, 125-33.
- Jacq, X., Brou, C., Lutz, Y., Davidson, I., Chambon, P. et Tora, L.** (1994). Human TAF_{II}30 is present in a distinct TFIID complex and is required for transcriptional activation by the estrogen receptor. *Cell* 79, 107-117.
- Janknecht, R.** (1996). Analysis of the ERK-stimulated transcription factor ER81. *Mol. Cell. Biol.* 16, 1551-1556.
- Janknecht, R. et Hunter, T.** (1996). Versatile molecular glue. *Curr. Biol.* 6, 951-954.
- Janknecht, R., Monté, D., Baert, J.-L. et De Launoit, Y.** (1996). The ETS-related transcription factor ERM is a nuclear target of signaling cascades involving MAPK and PKA. *Oncogene* 13, 1745-1754.
- Jans, D. A.** (1995). The regulation of protein transport to the nucleus by phosphorylation. *Biochem. J.*
- Jeon, I.-S., Davis, J. N., Braun, B. S., Sublett, J. E., Roussel, M. F., Denny, C. T. et Shapiro, D. N.** (1995). A variant Ewing's sarcoma translocation (7;22) fuses the EWS gene to the ETS gene ETV1. *Oncogene* 10, 1229-1234.
- Johnson, S. E., Wang, X., Hardy, S., Taparowski, E. J. et Konieczny, S. F.** (1996). Casein kinase II increases the transcriptional activities of MRF4 and MyoD independently of their direct phosphorylation. *Mol. Cell. Biol.* 16, 1604-1613.
- Joliot, V., Demma, M. et Prywes, R.** (1995). Interaction with RAP74 subunit of TFIIF is required for transcriptional activation by serum response factor. *Nature* 373, 632-5.
- Jonsen, M. D., Petersen, J. M., Xu, Q. P. et Graves, B. J.** (1996). Characterization of the cooperative function of inhibitory sequences in Ets-1. *Mol. Cell. Biol.* 16, 2065-73.
- Kamei, Y., Xu, L., Heinzl, T., Torchia, J., Kurokawa, R., Gloss, B., Lin, S. C., Heyman, R. A., Rose, D. W., Glass, C. K. et Rosenfeld, M. G.** (1996). A CBP integrator complex mediates transcriptional activation and AP-1 inhibition by nuclear receptors. *Cell* 85, 403-14.
- Kaufmann, R. C., Li, S., Gallagher, P. A., Zhang, J. et Carthew, R. W.** (1996). Ras1 signaling and transcriptional competence in the R7 cell of drosophila. *Genes Dev.* 10, 2167-2178.

Kaya, M., Yoshida, K., Higashino, F., Mitaka, T., Ishii, S. et Fujinaga, K. (1996). A single ets-related transcription factor, E1AF, confers invasive phenotype on human cancer cells. *Oncogene* 12, 221-227.

Kennedy, D., Ramsdale, T., Mattick, J. et Little, M. (1996). An RNA recognition motif in Wilms' tumour protein (WT1) revealed by structural modelling. *Nat. Genet.* 12, 329-31.

Klages, N. et Strubin, M. (1995). Stimulation of RNA polymerase II transcription initiation by recruitment of TBP in vivo. *Nature* 374, 822-823.

Klein, C. et Struhl, K. (1994). Increased recruitment of TATA-binding protein to the promoter by transcriptional activation domains in vivo. *Science* 266, 280-2.

Klemsz, M. J. et Maki, R. A. (1996). Activation of transcription by the PU.1 requires both acidic and glutamine-rich domains. *Mol. Cell. Biol.* 16, 390-397.

Klemsz, M. J., McKercher, S. R., Celada, A., Van, B. C. et Maki, R. A. (1990). The macrophage and B-cell specific transcription factor Pu.1 is related to the ETS oncogene. *Cell* 61, 113-124.

Knaus, R., Pollock, R. et Guarente, L. (1996). Yeast SUB1 is a suppressor of TFIIB mutations and has homology to the human co-activator PC4. *EMBO J.* 15, 1933-40.

Kodandapani, R., Pio, F., Ni, C. Z., Piccialli, G., Klemsz, M., Mckercher, S., Maki, R. A. et Ely, K. R. (1996). A new pattern for helix-turn-helix recognition revealed by the PU.1 ETS-domain-DNA complex. *Nature* 380, 456-460.

Koleske, A. J. et Young, R. A. (1995). The RNA polymerase II holoenzyme and its implications for gene regulation. *Trends Biochem. Sci* 20, 113-6.

Kominato, Y., Galson, D., Waterman, W. R., Webb, A. C. et Auron, P. E. (1995). Monocyte expression of the human prointerleukin 1 beta gene (IL1B) is dependent on promoter sequences which bind the hematopoietic transcription factor Spi-1/PU.1. *Mol. Cell. Biol.* 15, 59-68.

Kornberg, R. D. et Lorch, Y. (1995). Interplay between chromatin structure and transcription. *Curr. Op. Cell Biol.* 7, 371-375.

Kowentz-Leutz, E., Twamley, G., Ansieau, S. et Leutz, A. (1994). Novel mechanism of C/EBP β (NF-M) transcriptional control : activation through derepression. *Genes Dev.* 8, 2781.

Kraut, N., Frampton, J., McMagny, K. M. et Graf, T. (1994). A functional ETS DNA-binding domain is required to maintain multipotency of hematopoietic progenitors transformed by Myb-Ets. *Genes Dev.* **8**, 33-44.

Kretzchmar, M., Stelzer, G., Roeder, R. G. et Meisterernst, M. (1994). RNA polymerase II cofactor PC2 facilitates activation of transcription by GAL4-AH in vitro. *Mol. Cell. Biol.* **14**, 3927-3937.

Kwon, H., Imbalzano, A. N., Khavari, P. E., Kingston, R. E. et Green, M. R. (1994). Nucleosome disruption and enhancement of activator binding by a human SWI/SNF complex. *Nature* **370**, 477-481.

Laget, M. P., Defossez, P.-A., Albagli, O., Baert, J. L., Dewitte, F., Stehelin, D. et De Launoit, Y. (1996). Two functionally distinct domains cooperate for transactivation by the ETS family member ERM. *Oncogene* **12**, 1325-1336.

Laspia, M. F., Rice, A. P. et Matthews, M. B. (1989). HIV-1 Tat protein increases transcriptional initiation and stimulates elongation. *Cell* **59**, 283-292.

Laudet, V., Niel, C., Duterque, M., Leprince, D. et Stehelin, D. (1993). Evolution of the *ets* gene family. *Biochem. Biophys. Res. Comm.* **190**, 8-14.

Lee, D. Y., Hayes, J. J., Pruss, D. et Wolffe, A. P. (1993). A positive role for histone acetylation in transcription factor access to nucleosomal DNA. *Cell* **72**, 73-84.

Lee, H. S., Kraus, K. W., Wolfner, M. F. et Lis, J. T. (1992). DNA sequence requirements for generating paused polymerase at the start of *hsp70*. *Genes Dev.* **6**, 284-295.

Leprince, D., Crepieux, P., Laudet, V., Flourens, A. et Stehelin, D. (1993). A new mechanism of oncogenic activation: E26 retroviral v-ets oncogene has inverted the C-terminal end of the transcription factor c-ets-1. *Virology* **194**, 855-7.

Leprince, D., Gegonne, A., Coll, J., de Taisne, C., Schneeberger, A., Lagrou, C. et Stehelin, D. (1983). A putative second cell-derived oncogene of the avian leukemia virus E26. *Nature* **306**, 395-397.

Lessnick, S. L., Braun, B. S., Denny, C. T. et May, W. A. (1995). Multiple domains mediate transformation by the Ewing's sarcoma EWS/FLI-1 fusion gene. *Oncogene* **10**, 423-431.

Leuther, K. K., Salmeron, J. M. et Johnston, S. A. (1993). Genetic evidence that an activation domain of GAL4 does not require acidity and may form a β -sheet. *Cell* **72**, 575-585.

- Li, X. Y. et Green, M. R.** (1996). Transcriptional elongation and cancer. **Curr. Biol.** 6, 943-944.
- Lieberman, P. M. et Berk, A. J.** (1991). The Zta trans-activator protein stabilizes TFIID association with promoter DNA by direct protein-protein interaction. **Genes Dev.** 5, 2441-2454.
- Lim, F., Kraut, N., Frampton, J. et Graf, T.** (1992). DNA binding by c-ETS-1, but not v-ETS, is repressed by an intramolecular mechanism. **EMBO J.** 11, 643-652.
- Lis, J. et Wu, C.** (1993). Protein traffic on the Heat Shock Promoter : parking, stalling and trucking around. **Cell** 74, 1-4.
- Logan, S. K., Garabedian, M. J., Campbell, C. E. et Werb, Z.** (1996). Synergistic transcriptional activation of the tissue inhibitor of metalloproteinase-1 promoter via functional interaction of AP1 and ETS transcription factors. **J. Biol. Chem.** 271, 774-782.
- Lopez, M., Oettgen, P., Akbarali, Y., Dendorfer, U. et Libermann, T. A.** (1994). ERP, a new member of the ets transcription factor/oncoprotein family : characterization and differential expression during B-lymphocyte development. **Mol. Cell. Biol.** 14, 3292-3309.
- Lu, H., Zawel, L., Fisher, L., Egly, J.-M. et Reinberg, D.** (1992). Human general transcription factor TFIIF phosphorylates the C-terminal domain of RNA polymerase II. **Nature** 358, 641-645.
- Majerus, M.-A., Bibollet-Ruche, F., Telliez, J.-B., Wasylyk, B. et Bailleul, B.** (1992). Serum, AP1 and Ets1 stimulate the human ets-1 promoter. **Nucl. Acids Res.** 20, 2699-2703.
- Makkerh, J. P. S., Dingwall, C. et Laskey, R. A.** (1996). Comparative mutagenesis of nuclear localization signals reveals the importance of neutral and acidic amino acids. **Curr. Biol.** 6, 1025-1027.
- Mann, R. S. et Chan, S. K.** (1996). Extra specificity from extradenticle : the partnership between HOX and PBX/EXD proteins. **Trends Genet.** 12, 258-262.
- Mao, X., Miesfeldt, S., Yang, H., Leiden, J. M. et Thompson, C. B.** (1994). The FLI-1 and chimeric EWS-FLI-1 oncoproteins display similar DNA binding specificities. **J. Biol. Chem.** 269, 18216-22.
- Marcus, G. A., Silverman, N., Berger, S. L., Horiuchi, J. et Guarente, L.** (1994). Functional similarity and physical association between GCN5 and ADA2 : putative transcriptional adaptators. **EMBO J.** 13, 4807-4815.

- Massari, M. E., Jennings, P. A. et Murre, C. (1996). The AD1 transactivation domain of E2A contains a highly conserved helix which is required for its activity in both *S. cerevisiae* and mammalian cells. **Mol. Cell. Biol.** 16, 121-129.
- May, M., Mengus, G., Lavigne, A. C., Chambon, P. et Davidson, I. (1996). Human TAF_{II}28 promotes transcriptional stimulation by Activation Function 2 of the Retinoid X Receptor. **EMBO J.** 15, 3093-3104.
- May, W. A., Lessnick, S. L., Braun, B. S., Klemsz, M., Lewis, B. C., Lunsford, L. B., Hromas, R. et Denny, C. T. (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., Torbett, B. E., Anderson, K. L., Henkel, G. W., Vestal, D. J., Baribault, H., Klemsz, M., Feeney, A. J., Wu, G. E., Paige, C. J. et Maki, R. A. (1996). Targeted disruption of the Pu.1 gene results in multiple hematopoietic abnormalities. **EMBO J.** 15, 5647-5658.
- Melet, F., Motro, B., Rossi, D. J., Zhang, L. et Bernstein, A. (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, 2708-18.
- Mengus, G., May, M., Jacq, X., Staub, A., Tora, L., Chambon, P. et Davidson, I. (1995). Cloning and characterization of hTAF_{II}18, hTAF_{II}20 and hTAF_{II}28 : three subunits of the human transcription factor TFIID. **EMBO J.** 14, 1520-1531.
- Metz, T. et Graf, T. (1991a). Fusion of the nuclear oncoproteins v-myb and v-ets is required for the leukemogenicity of E26 virus. **Cell** 66, 95-105.
- Metz, T. et Graf, T. (1991b). v-myb and v-ets transform chicken erythroid cells and cooperate both in *trans* and *cis* to induce distinct differentiation phenotypes. **Genes Dev.** 5, 369-380.
- Meyer, D., Stiegler, P., Hindelang, C., Mager, A. M. et Rémy, P. (1995). Whole-mount in situ hybridization reveals the expression of the XI-Fli gene in several lineages of migrating cells in *Xenopus* embryos. **Int. J. Dev. Biol.** 39, 909-919.
- Miyazaki, Y., Sun, X., Uchida, H., Zhang, J. et Nimer, S. (1996). MEF, a novel transcription factor with an Elf-1 like DNA-binding domain but distinct transcriptional activating properties. **Oncogene** 13, 1721-1729.

- Monté, D., Coutte, L., Dewitte, F., Defossez, P. A., Le Coniat, M., Stéhelin, D., Berger, R. et de Launoit, Y. (1996). Genomic organization of the human ERM gene, a member of the PEA3 group of ETS transcription factors. *Genomics* 35, 236-240.
- Moqtaderi, Z., Bai, Y., Poon, D., Weil, P. A. et Struhl, K. (1996). TBP-associated factors are not generally required for transcriptional activation in yeast. *Nature* 383, 188-191.
- Moreau, G. F., Tavitian, A. et Tambourin, P. (1988). Spi-1 is a putative oncogene in virally induced murine erythroleukemias. *Nature* 331, 277-280.
- Moreau-Gachelin, F. (1994). Spi-1/PU.1 : an oncogene of the Ets family. *Biochim. Biophys. Acta* 1198, 149-163.
- Moreau-Gachelin, F., Wendling, F., Molina, T., Denis, N., Titeux, M., Grimber, G., Briand, P., Vainchenker, W. et Tavitian, A. (1996). Spi-1/Pu.1 transgenic mice develop multistep erythroleukemia. *Mol. Cell. Biol.* 16, 2453-2463.
- Muthusamy, R., Barton, K. et Leiden, J. (1995). Defective activation and survival of T cells lacking the Ets-1 transcription factor. *Nature* 377, 639-642.
- Nagpal, S., Saunders, M., Kastner, P., Durand, B., Nakshatri, H. et Chambon, P. (1992). Promoter-context and response-element dependent specificity of the transcriptional activation and modulation functions of retinoic acid receptors. *Cell* 70, 1007-1019.
- Nakae, K., Nakajima, K., Inazawa, J., Kitaoka, T. et Hirano, T. (1995). ERM, a member of the PEA3 subfamily of Ets transcription factors, can cooperate with c-jun. *J. Biol. Chem.* 270, 23795-23800.
- Nerlov, C., De, C. D., Pergola, F., Caracciolo, A., Blasi, F., Johnsen, M. et Verde, P. (1992). A regulatory element that mediates co-operation between a PEA3-AP-1 element and an AP-1 site is required for phorbol ester induction of urokinase enhancer activity in HepG2 hepatoma cells. *EMBO J.* 11, 4573-82.
- Nikolov, D. B., Chen, H., Halay, E. D., Usheva, A. A., Hisatake, K., Lee, D. K., Roeder, R. G. et Burley, S. K. (1995). Crystal structure of a TFIIB-TBP-TATA element ternary complex. *Nature* 377, 119-28.
- Nunn, M. F., Seeburg, P. H., Moscovici, C. et Duesberg, P. H. (1983). Tripartite structure of the avian erythroblastosis virus E26 transforming gene. *Nature* 306, 391-395.
- O'Brien, T. S., Hardin, A., Greenleaf, A. et Lis, J. T. (1994). Phosphorylation of RNA polymerase II C-terminal domain and transcriptional elongation. *Nature* 370, 75-77.

- O'Hagan, R. C., Tozer, R. G., Symons, M., McCormick, F. et Hassell, J. A.** (1996). The activity of the transcription factor PEA3 is regulated by two distinct MAPK cascades. *Oncogene* 13, 1323-1333.
- O'Neill, E. M., Rebay, I., Tjian, R. et Rubin, G. M.** (1994). The activities of two Ets-related transcription factors required for Drosophila eye development are modulated by the Ras/MAPK pathway. *Cell* 78, 137-147.
- Oehler, T. et Angel, P.** (1992). A common intermediary factor (p52/54) recognizing "acidic blob"-type domains is required for transcriptional activation by the Jun proteins. *Mol. Cell. Biol.* 12, 5508-5515.
- Oelgeschlager, T., Chiang, C. M. et Roeder, R. G.** (1996). Topology and reorganization of a human TFIID-promoter complex. *Nature* 382, 735-738.
- Oettgen, P., Akbarali, Y., Boltax, J., Best, J., Kunsch, C. et Libermann, T.** (1996). Characterization of NERF, a novel transcription factor related to the ETS factor Elf-1. *Mol. Cell. Biol.* 16, 5091-5106.
- Okamoto, H., Sheline, C. T., Corden, J. L., Jones, K. A. et Peterlin, B. M.** (1996). Transactivation by human immunodeficiency virus Tat protein requires the C-terminal domain of RNA polymerase II. *Proc. Natl. Acad. Sci. USA* 93, 11575-11579.
- Papavassiliou, A., Treier, M. et Bohmann, D.** (1995). Intramolecular signal transduction in c-Jun. *EMBO J.* 14, 2014-2019.
- Parvin, J. D., McCormick, R. J., Sharp, P. A. et Fisher, D. E.** (1995). Pre-bending of a promoter sequence enhances affinity for the TATA-binding factor. *Nature* 373, 724-727.
- Parvin, J. D. et Sharp, P. A.** (1993). DNA topology and a minimal set of basal factors for transcription by RNA polymerase II. *Cell* 73, 533-540.
- Parvin, J. D., Timmers, H. T. M. et Sharp, P. A.** (1992). Promoter specificity of basal transcription factors. *Cell* 68, 1135-1144.
- Petersen, J. M., Skalicky, J. J., Donaldson, L. W., McIntosh, L. P., Alber, T. et Graves, B. J.** (1995). Modulation of transcription factor Ets-1 DNA binding: DNA-induced unfolding of an alpha helix. *Science* 269, 1866-9.
- Peterson, C. L. et Tamkun, J. W.** (1995). The SWI-SNF complex: a chromatin remodeling machine? *Trends Biochem. Sci* 20, 143-6.

- Pongubala, J. M. R., Nagulapalli, S., Klemsz, M. J., McKercher, S. R., Maki, R. A. et Atchinson, M. L. (1992). PU.1 recruits a second nuclear factor to a site important for immunoglobulin κ 3' enhancer activity. *Mol. Cell. Biol.* 12, 368-378.
- Pongubala, J. M. R., Van Beveren, C., Nagulapalli, S., Klemsz, M. J., McKercher, S. R., Maki, R. A. et Atchison, M. L. (1993). Effect of PU1 phosphorylation on interaction with NF-EM5 and transcriptional activation. *Science* 259.
- Prioleau, M. N., Buckle, R. S. et Méchali, M. (1995). Programming of a repressed but committed chromatin structure during early development. *EMBO J.* 14, 5073-5084.
- Prosser, H. M., Wotton, D., Gégonne, A., Ghysdael, J., Wang, S., Speck, N. A. et Owen, M. J. (1992). A phorbol ester response element within the human T- Cell receptor Y-chain enhancer. *Proc. Natl. Acad. Sci. USA* 89, 9934-9938.
- Quéva, C., Leprince, D., Stéhelin, D. et Vandebunder, B. (1993). p54^{c-ets-1} and p68^{c-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.
- Quinn, J., Fyrberg, A. M., Ganster, R. W., Schmidt, M. C. et Peterson, C. L. (1996). DNA-binding properties of the yeast SWI/SNF complex. *Nature* 379, 844-7.
- Rabault, B. et Ghysdael, J. (1994). Calcium-induced phosphorylation of ETS1 inhibits its specific DNA binding activity. *J. Biol. Chem.* 269, 28143-28151.
- Rabault, B., Roussel, M. F., Tran Quang, C. et Ghysdael, J. (1996). Phosphorylation of Ets1 regulates the complementation of a CSF-1 receptor impaired in mitogenesis. *Oncogene* 13, 877-881.
- Rao, V. N., Ohno, T., Prasad, D. D. K., Batthacharya, G. et Reddy, E. S. (1993). Analysis of the DNA-binding and transcriptional activation functions of human *Fli-1* protein. *Oncogene* 8, 2167-2173.
- Rao, V. N. et Reddy, E. S. (1993). Delta elk-1, a variant of elk-1, fails to interact with the serum response factor and binds to DNA with modulated specificity. *Cancer Res* 53, 215-20.
- Rasclé, A., Ferrand, N., Gandrillon, O. et Samarut, J. (1996). Myb-Ets fusion oncoprotein inhibits thyroid-hormone Receptor/c-Erb-A and Retinoic Acid Receptor functions : a novel mechanism of action for leukemogenic transformation by E26 avian retrovirus. *Mol. Cell. Biol.* 16, 6338-6351.

- Ray-Gallet, D., Mao, C., Tavitian, A. et Moreau, G. F.** (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, 303-13.
- Rebay, I. et Rubin, G. M.** (1995). Yan functions as a general inhibitor of differentiation and is negatively regulated by activation of the Ras1/MAPK pathway. *Cell* 81, 857-866.
- Rechsteiner, M. et Rogers, S. W.** (1996). PEST sequences and regulation by proteolysis. *Trends Biochem. Sci.* 21, 267-271.
- Reifsnyder, C., Lowell, J., Clarke, A. et Pillus, P.** (1996). Yeast SAS silencing genes and human genes associated with AML and HIV-1 Tat interactions are homologous with acetyltransferases. *Nature Genetics* 14, 42-49.
- Reines, D., Conaway, J. W. et Conaway, R. C.** (1996). The RNA polymerase II general elongation factors. *Trends Biochem. Sci.* 21, 351-355.
- Rémy, P., Sénan, F., Meyer, D., Mager, A. M. et Hindelang, C.** (1996). Overexpression of the *Xenopus* XL-fli gene during early embryogenesis leads to anomalies in head and heart development and erythroid differentiation. *Int. J. Dev. Biol.* 40, 577-589.
- Rihs, H. P., Jans, D. A., Fan, H. et Peters, R.** (1991). The rate of nuclear cytoplasmic protein transport is determined by the casein kinase II site flanking the nuclear localization sequence of the SV40 T-antigen. *EMBO J.* 10, 633-639.
- Roeder, R. G.** (1996). The role of general initiation factors in transcription by RNA polymerase II. *Trends Biochem. Sci.* 21, 327-335.
- Romano-Spica, V., Suzuki, H., Georgiou, P., Chen, S.-L., Ascione, R., Papas, T. S. et Bhat, N. K.** (1994). Expression of *ets* family genes in hematopoietic cells. *Int. J. Oncol.* 4, 521-531.
- Rosen, G. D., Barks, J. L., Iademarco, M. F., Fisher, R. J. et Dean, D. C.** (1994). An intricate arrangement of binding sites for the *c-ets-1* family transcription factors regulates the activity of the alpha 4 integrin gene promoter. *J. Biol. Chem.* 269, 15652-15660.
- Sauer, F., Fondell, J. D., Ohkuma, Y., Roeder, R. G. et Jäckle, H.** (1995a). Control of transcription by Krüppel through interaction with TFIIB and TFIIE β . *Nature* 375, 162-164.
- Sauer, F., Hansen, S. K. et Tjian, R.** (1995b). DNA template and activator-coactivator requirements for transcriptional synergism by *Drosophila* bicoid. *Science* 270, 1825-8.

Schmitz, M. L., dos Santos Silva, M. A., Altmann, H., Czisch, M., Holak, T. A. et Bauerle, P. A. (1994). Structural and functional analysis of the NF- κ B p65 C terminus. *J. Biol. Chem.* 269, 25613-25620.

Schmitz, M. L., Stelzer, G., Altmann, H., Meisterernst, M. et Bauerle, P. A. (1995). Interaction of the COOH-terminal transactivation domain of NF- κ B with TATA-binding protein, TFIIB and coactivators. *J. Biol. Chem.* 270, 7219-7226.

Schuetze, S., Paul, R., Gliniak, B. C. et Kabat, D. (1992). Role of the Pu.1 transcription factor in controlling differentiation of Friend erythroleukemia cell lines. *Mol. Cell. Biol.* 12, 2967-2975.

Schuetze, S., Stenberg, P. E. et Kabat, D. (1993). The Ets-related transcription factor Pu.1 immortalizes erythroblasts. *Mol. Cell. Biol.* 13, 5670-5678.

Schwarz, E. M., Van, A. D. et Verma, I. M. (1996). Constitutive phosphorylation of IkappaB α by casein kinase II occurs preferentially at serine 293: requirement for degradation of free IkappaB α . *Mol. Cell. Biol.* 16, 3554-9.

Scott, E. W., Simon, M. C., Anastasi, J. et Singh, H. (1994). Requirement of transcription factor PU-1 in the development of multiple hematopoietic lineages. *Science* 265, 1573-1577.

Seth, A. et Papas, T. S. (1990). The c-ets-1 proto-oncogene has oncogenic activity and is positively autoregulated. *Oncogene* 5, 1761-7.

Shilatifard, A., Lane, W. S., Jackson, K. W., Conaway, R. C. et Conaway, J. W. (1996). An RNA polymerase II elongation factor encoded by the human ELL gene. *Science* 271, 1873-1876.

Shin, M. K. et Koshland, M. E. (1993). Ets-related protein PU-1 regulates expression of the immunoglobulin J-chain gene through a novel Ets-binding element. *Genes Dev.* 7, 2006-2015.

Shykind, B. M., Kim, J. et Sharp, P. A. (1995). Activation of the TFIID-TFIIA complex with HMG-2. *Genes Dev.* 9, 1354-1365.

Siddique, H. R., Rao, V. N., Lee, L. et Reddy, E. S. P. (1993). Characterization of the DNA binding domain and transcriptional activation domains of the erg protein. *Oncogene* 8, 1751-1755.

Sieweke, M. H., Tekotte, H., Frampton, J. et Graf, T. (1996). MafB is an interaction partner and repressor of Ets-1 that inhibits erythroid differentiation. *Cell* 85, 49-60.

Silverman, N., Agapite, J. et Guarente, L. (1994). Yeast ADA2 protein binds to the VP16 protein activation domain and activates transcription. *Proc. Natl. Acad. Sci. USA* 91, 11665-11668.

Songyang, Z., Lu, K. P., Kwon, Y. T., Tsai, L. H., Filhol, O., Cochet, C., Brickey, D. A., Soderling, T. R., Bartleson, C., Graves, D. J., DeMaggio, A. J., Hoekstra, M. F., Blenis, J., Hunter, T. et Cantley, L. C. (1996). A structural basis for substrate specificities of protein Ser/Thr kinases : primary sequence preference of Casein Kinases I and II, NIMA, Phosphorylase Kinase, CDK II, CDK5, and Erk1. *Mol. Cell. Biol.* 16, 6486-6493.

Sorensen, P. H., Lessnick, S. L., Lopez, T. D., Liu, X. F., Triche, T. J. et Denny, C. T. (1994). A second Ewing's sarcoma translocation, t(21;22), fuses the EWS gene to another ETS-family transcription factor, ERG. *Nat. Genet.* 6, 146-51.

Soudant, N., Albagli, O., Dhordain, P., Flourens, A., Stéhelin, D. et Leprince, D. (1994). A residue in the ETS domain mutated in the v-ets oncogene is essential for the DNA-binding and transactivating properties of the Ets-1 and Ets-2 proteins. *Nucleic Acids Res.* 22, 3871-3879.

Southgate, C. D. et Green, M. R. (1991). The HIV-1 Tat protein activates transcription from an upstream DNA-binding site : implications for Tat function. *Genes Dev.* 5, 2496-2507.

Strahl, T., Gille, H. et Shaw, P. E. (1996). Selective response of ternary complex factor Sap1a to different MAP kinase subgroups. *Proc. Natl. Acad. Sci. USA* 93, 11563-11568.

Struhl, K. (1996). Chromatin structure and RNA polymerase II connection : implications for transcription. *Cell* 84, 179-182.

Sumarsono, S. H., Wilson, T. J., Tymms, M. J., Venter, D. J., Corrick, C. M., Kola, R., Lahoud, M. H., Papas, T. S., Seth, A. et Kola, I. (1996). Down's syndrome like skeletal abnormalities in Ets2 transgenic mice. *Nature* 379, 534-537.

Tanaka, M., Lai, J.-S. et Herr, W. (1992). Promoter-selective activation domains in Oct-1 and Oct-2 direct differential activation of an snRNA and mRNA promoter. *Cell* 68, 755-767.

Thirman, M. J., Levitan, D. A., Kobayashi, H., Simon, M. C. et Rowley, J. D. (1994). Cloning of ELL, a gene that fuses to MLL in a t(11;19)(q23;p13.1) in acute myeloid leukemia. *Proc. natl. Acad. Sci. USA* 91, 12110-12114.

Treier, M., Bohmann, D. et Mlodzik, M. (1995). JUN cooperates with the ETS domain protein pointed to induce photoreceptor R7 fate in the Drosophila eye. *Cell* 83, 753-60.

- Treier, M., Staszewski, L. M. et Bohmann, D.** (1994). Ubiquitin-dependent c-Jun degradation is mediated by the δ domain. *Cell* 78, 787-798.
- Treisman, R.** (1995). Journey to the surface of the cell : Fos regulation and the SRE. *EMBO J.* 14, 4905-4913.
- Triezenberg, S. J.** (1995). Structure and function of transcriptional activation domains. *Curr. Op. Gen. Dev.* 5, 190-196.
- Tsukiyama, T., Becker, P. B. et Wu, C.** (1994). ATP-dependent nucleosome disruption at a heat-shock promoter mediated by binding of GAGA transcription factor. *Nature* 367, 525-532.
- Tsukiyama, T., Daniel, C., Tamkun, J. et Wu, C.** (1995). ISWI, a member of the SWI2/SNF2 ATPase family, encodes the 140 kDa subunit of the nucleosome remodeling factor. *Cell* 83, 1021-6.
- Tsukiyama, T. et Wu, C.** (1995). Purification and properties of an ATP-dependent nucleosome remodeling factor. *Cell* 83, 1011-20.
- Urano, F., Umezawa, A., Hong, W., Kikuchi, H. et Hata, J. H.** (1996). A novel chimera gene between EWS and E1AF, encoding the adenovirus E1A enhancer binding protein, in extraosseous Ewing's sarcoma. *Biochem. Biophys. Res. Comm.* 219, 608-612.
- van Dijk, M. A. et Murre, C.** (1994). extradenticle raises the DNA-binding specificity of homeotic selector gene products. *Cell* 78, 617-624.
- Van Hoy, M., Leuther, K. K., Kodadek, T. et Johnston, S. A.** (1993). The acidic activation domains of the GCN4 and GAL4 proteins are not α helical but form β sheets. *Cell* 72, 587-594.
- Vandenbunder, B., Pardanaud, L., Jaffredo, T., Mirabel, M. A. et Stehelin, D.** (1989). Complementary patterns of expression of *c-ets1*, *c-myb* and *c-myc* in the blood-forming system of the chick embryo. *Development* 107, 265-274.
- Verrijzer, C. P., Chen, J. L., Yokomori, K. et Tjian, R.** (1996). Binding of TAFs to core promoter elements mediate promoter selectivity by RNA polymerase II. *Cell* 81, 1115-1125.
- Verrijzer, C. P. et Tjian, R.** (1996). TAFs mediate transcriptional activation and promoter selectivity. *Trends Biochem. Sci.* 21, 338-342.
- Walker, S. S., Reese, J. C., Apone, L. M. et Green, M. R.** (1996). Transcription activation in cells lacking TAFs. *Nature* 383, 185-188.

Walter, J. et Biggin, M. D. (1996). DNA-binding specificity of two homeodomain proteins in vitro and in vivo. *Proc. Natl. Acad. Sci. USA* 93, 2680-2685.

Wang, E. et Tjian, R. (1994). Promoter-selective transcriptional defect in cell cycle mutant ts13 rescued by hTAF_{II}250. *Science* 263, 811-814.

Wang, W., Cote, J., Xue, Y., Zhou, S., Khavara, P. A., Biggar, S. R., Muchardt, C., Kalpana, G. V., Goff, S. P., Yaniv, M., Workman, J. L. et Crabtree, G. R. (1996a). Purification and biochemical heterogeneity of the mammalian SWI/SNF complex. *EMBO J.* 15, 5370-5382.

Wang, W., Xue, Y., Zhou, S., Kuo, A., Cairns, B. R. et Crabtree, G. R. (1996b). Diversity and specialization of mammalian SWI/SNF complexes. *Genes Dev.* 10, 2117-2130.

Wasylyk, B., Hahn, S. L. et Giovane, A. (1993). The Ets family of transcription factors. *Eur. J. Biochem.* 211, 7-18.

Wasylyk, C., Gutman, A., Nicholson, R. et Wasylyk, B. (1991). The c-Ets oncoprotein activates the stromelysin promoter through the same elements as several non-nuclear oncoproteins. *EMBO J.* 10, 1127-34.

Werner, M. H., Clore, G. M., Fisher, C. L., Fisher, R. J., Trinh, L., Shiloach, J. et Gronenborn, A. M. (1995). The solution structure of the human ETS1-DNA complex reveals a novel mode of binding and true side-chain intercalation. *Cell* 83, 761-771.

Werner, M. H., Gronenborn, A. M. et Clore, G. M. (1996). Intercalation, DNA kinking, and the control of transcription. *Science* 271, 778-84.

Wernert, N., Raes, M. B., Lassale, P., Dehouck, M. P., Gosselin, B., Vandebunder, B. et Stéhelin, D. (1992). c-ets-1 protooncogene is a transcription factor expressed in endothelial cells during tumor vascularization and other forms of angiogenesis in humans. *Am. J. Pathol.* 140, 119-127.

Wilson, C. J., Chao, D. M., Imbalzano, A. N., Schnitzler, G. R., Kingston, R. E. et Young, R. A. (1996). RNA polymerase II holoenzyme contains SWI/SNF regulators involved in chromatin remodeling. *Cell* 84, 235-44.

Wollfe, A. P. et Pruss, D. (1996). Chromatin : hanging on to histones. *Curr. Biol.* 6, 234-237.

Woods, D. B., Ghysdael, J. et Owen, M. J. (1992). Identification of nucleotide preferences in DNA sequences recognized specifically by c-Ets-1 protein. *Nucleic Acid Res* 20, 699-704.

Wotton, D., Ghysdael, J., Wang, S., Speck, N. A. et Owen, M. J. (1994). Cooperative binding of Ets-1 and Core Binding Factor to DNA. *Mol. Cell. Biol.* 14, 840-850.

Wu, Y., Reece, R. J. et Ptashne, M. (1996). Quantitation of putative activator-target affinities predicts transcriptional activating potential. *EMBO J.* 15, 3951-3963.

Xiao, H., Friesen, J. D. et Lis, J. T. (1995). Recruiting TATA-binding protein to a promoter: transcriptional activation without an upstream activator. *Mol. Cell. Biol.* 15, 5757-61.

Xiao, H., Pearson, A., Coulombe, B., Truant, R., Zhang, S., Regier, J. L., Triezenberg, S. J., Reinberg, D., Flores, O., Ingles, C. J. et Greenblatt, J. (1994). Binding of basal transcription factor TFIID to the acidic activation domains of VP16 and p53. *Mol. Cell. Biol.* 14, 7013-7024.

Xie, X., Kokubo, T., Cohen, S. L., Mirza, U. A., Hoffmann, A., Chait, B. T., Roeder, R. G., Nakatani, Y. et Burley, S. K. (1996). Structural similarity between TAFs and the heterotetrameric core of the histone octamer [see comments]. *Nature* 380, 316-22.

Yang, B.-S., Hauser, C. A., Henkel, G., Colman, M. S., Van Beveren, C., Stacey, K. J., Hume, D. A., Maki, R. A. et Ostrowski, M. C. (1996a). Ras-mediated phosphorylation of a conserved threonine residue enhances the transactivation activities of c-ets-1 and c-ets-2. *Mol. Cell. Biol.* 16, 538-547.

Yang, X.-J., Ogryzko, V. V., Nishikawa, J.-I., Howard, B. H. et Nakatani, Y. (1996b). A p300/CBP-associated factor that competes with the adenoviral oncoprotein E1A. *Nature* 382, 319-324.

Yankulov, K., Blau, J., Purton, T., Roberts, S. et Bentley, D. L. (1994). Transcription elongation by RNA polymerase II is stimulated by transactivators. *Cell* 77, 749-759.

Yankulov, K. Y., Pandes, M., McCracken, S., Bouchard, D. et Bentley, D. L. (1996). TFIID functions in regulating transcriptional elongation by RNA Polymerase II in *Xenopus* oocytes. *Mol. Cell. Biol.* 16, 3291-3299.

Yoshinaga, S. K., Peterson, C. L., Herskowitz, I. et Yamamoto, K. R. (1992). Roles of SWI1, SWI2, and SWI3 proteins for transcriptional enhancement by steroid receptors. *Science* 258, 1598-604.

Yuryev, A., Patturajan, M., Litington, Y., Joshi, R. V., Gentile, C., Gebara, M. et Corden, J. L. (1996). The C-terminal domain of the largest subunit of RNA polymerase II interacts with a novel set of serine/arginine-rich proteins. *Proc. Natl. Acad. Sci. U S A* 93, 6975-80.

Zhang, L., Eddy, A., Teng, Y. T., Fritzler, M., Kluppel, M., Melet, F. et Bernstein, A. (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.* 15, 6961-6970.

Zhang, L., Lemarchandel, V., Romeo, P. H., Ben-David, Y. et Greer, P. (1993). The Fli-1 proto-oncogene, involved in erythroleukemia and Ewing's sarcoma, encodes a transcriptional activator with DNA-binding specificities distinct from other Ets family members. *Oncogene* 8, 1621-1630.

Zhou, G., Bao, Z. Q. et Dixon, J. E. (1995). Components of a new human protein kinase signal transduction pathway. *J. Biol. Chem.* 270, 12665-12669.

Zhou, Q. et Sharp, P. A. (1996). Tat-SF1 : cofactor for stimulation of transcriptional elongation by HIV-1 Tat. *Science* 274, 605-610.

Zipursky, S. L. et Rubin, G. M. (1994). Determination of neuronal cell fate: lessons from the R7 neuron of *Drosophila*. *Annu Rev Neurosci* 17, 373-97.

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

