

50376
1993
346

50376
1993
346

**UNIVERSITE DES SCIENCES ET TECHNOLOGIES
DE LILLE**

Numéro d'ordre 1226

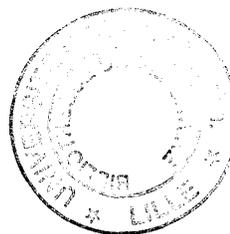
THESE

présentée pour l'obtention du grade de

DOCTEUR DE L'UNIVERSITE DES SCIENCES ET TECHNIQUES DE LILLE

par

Pascale CREPIEUX



STRUCTURE ET REGULATION DES PROMOTEURS DU LOCUS

***C-ETS-1* AVIAIRE**

Soutenu le 7 Décembre 1993 devant la commission d'examen

Président : Pr Jean-Pierre BOHIN

Rapporteurs : Pr Hans CLEVERS

Dr Françoise MOREAU-GACHELIN

Examineurs : Pr Jean KREMBEL

Dr Dominique LEPRINCE

Pr Dominique STEHELIN

Grand merci à Dominique Stéhelin, à qui je suis reconnaissante de quatre années (et plus encore) de vie intense parmi les fous. Et merci aussi pour la confiance qu'il a osé m'accorder.

Merci à Dominique Leprince (notre grand DL à tous, un garçon très bien...), parce que j'ai appris à son contact la "confiance dubitative".

Je remercie vivement Jean Coll de mettre au service de tous son regard alternatif, toujours avec bienveillance et souvent avec sagesse.

Merci à Anne-Claire Flourens dont l'infatigable énergie a permis de démêler mes clonages les plus aventureux.

Merci à Christian Lagrou qui, inlassablement et sans broncher, a accepté d'endosser pendant des mois le rôle ingrat de dealer de FEP.

Merci à Stéphane Ansieau d'éclairer chaque jour de son sourire le labo, même par gros temps.

Je remercie Vincent Laudet qui persuaderait un pipetman qu'il a ses chances au concours du CNRS, pour sa précieuse amitié, et pour sa passion de la vie.

Je remercie Véronique Fafeur, présidente d'honneur de notre micro-club littéraire autour du calumet de la pose, et autres "grands plaisirs".

Merci à Martine Duterque-Coquillaud, ma consoeur de la transcription et de la traduction, pour toutes ces années de complicité.

Merci à Christine Dozier qui a contribué à dégager la piste embourbée des promoteurs et m'a fait part de son expérience avec beaucoup de sollicitude.

Merci à Olivier Albagli pour tous les bons moments de la période bleue.

Merci à Elisabeth Ferreira pour l'aide attentive qu'elle m'a apportée pendant la période noire, et pour les luxueux goûters avec lesquels elle a choyé le groupe.

Merci à Yvan II, pour avoir guidé ma participation à l'établissement de la canadien connection.

Merci aussi à tous les jeunes qui ont su créer une ambiance-chaleureuse-d'entraide-et-d'amitié : Philippe Dhordain (la vie est belle !), Hélène Pelczar (je vous la recommande), Anne Hautefeuille (même quand les clonages la gonflent), Nathalie Soudant (la petite-mais-quand-même-insolente), Jean-Marc Vanacker (be the force with thou), Catherine Hänni et son orchestre jurassique, Frédérique Dewitte et son grand rire, Manos Perros et toute la bande. et bonne chance à eux tous.

Merci à Nicole Devassine qui a souvent su garder son calme devant ma stupeur indémodable devant le fax, la photocopieuse, le mac, etc, etc. Merci aussi à Marie-Kiki, euh...un peu pour les mêmes raisons.

Merci à Patrick Martin et à Thierry Dupressoir qui ont répondu à mes multiples signaux de détresse.

En résumé, grand merci à tous ceux qui m'ont aidée un jour ou l'autre, par conséquent à l'ensemble des membres du laboratoire.

Grand merci aussi à Mme Françoise Moreau-Gachelin, et à Mrs Hans Clevers, Jean-Pierre Bohin et Jean Krembel qui ont accepté avec beaucoup de diligence d'examiner le travail qui est présenté ici.

SOMMAIRE

ABRÉVIATIONS	6
RÉSUMÉ	8
PRÉSENTATION DU MÉMOIRE	9
<u>INTRODUCTION</u>	11
<u>1ère partie : L'énigme de la transactivation</u>	12
1. Présentation des protagonistes du complexe d'initiation.....	12
2. Présentation des modulateurs transcriptionnels.....	13
3. Relations avec le complexe d'initiation.....	15
4. Contra-répression nucléosomale.....	17
4.1.Exclusion permanente.....	18
4.2. Induction du déplacement du nucléosome.....	18
4.3. Répression transcriptionnelle.....	19
5. Conclusions.....	19
<u>2ème partie : Présentation du rétrovirus leucémogène</u>	
E26 et propriétés de l'oncogène <i>v-ets</i>	21
1. Découverte de l'oncogène viral <i>v-ets</i>	21
2. Structure et expression du génome viral.....	22
2.1. Fusion des séquences <i>v-ets</i> aux gènes <i>gag</i> et <i>v-myb</i>	22

2.2. Genèse du virus E26.....	22
2.2.1. Homologies entre les séquences <i>v-myb</i> ^{E26} et <i>c-myb</i>	22
2.2.2. Homologies entre les gènes <i>v-ets</i> et <i>c-ets-1</i>	22
2.2.3. Point de recombinaison entre les gènes <i>v-myb</i> et <i>v-ets</i>	24
2.2.4. Modes de recombinaison des séquences <i>v-myb</i> et <i>v-ets</i>	24
3. Transformation des cellules hématopoïétiques.....	25
3.1. Les cellules-cibles du virus E26.....	25
3.2. Contribution respective des gènes <i>v-ets</i> et <i>v-myb</i> au pouvoir transformant du virus E26.....	26
3.2.1. Rôle de l'oncogène <i>v-myb</i>	26
3.2.2. Contribution de l'oncogène <i>v-ets</i>	27
3.2.3. Effets des mutants thermosensibles <i>in vivo</i>	28
3.3. Importance de la fusion des séquences <i>v-myb</i> et <i>v-ets</i> sur le pouvoir transformant.....	28
3.4. Domaines fonctionnels importants pour la transformation hématopoïétique.....	29
3.5. Les gènes cibles hématopoïétiques de la P135 ^{Agag-myb-ets}	30
4. Influence de l'infection par le virus E26 sur la croissance des fibroblastes.....	31
5. Effet mitogène du virus E26 sur les cellules neurorétiniennes.....	32

3ème partie : La famille Ets, un ensemble multigénique de facteurs de transcription. Aspects moléculaires.....

1. Présentation de la famille Ets.....	34
1.1. Le domaine ETS.....	34
1.2. Localisation chromosomique.....	34
1.3. Histoire évolutive de la famille Ets.....	35
2. Les protéines Ets constituent une superfamille de facteurs de transcription.....	35
2.1. Reconnaissance de l'ADN.....	36
2.1.1. Reconnaissance macromoléculaire.....	36
2.1.2. Définition de la cible nucléotidique.....	36
2.2. Interactions avec l'ADN vues du côté protéique.....	37

2.2.1. Le domaine de liaison à l'ADN.....	37
2.2.2. Inhibition intramoléculaire de la liaison à l'ADN.....	38
2.3. Domaines transactivateurs.....	40
2.4. Interactions intermoléculaires.....	41
2.4.1. La fixation du partenaire transcriptionnel dépend de celle de la protéine Ets.....	42
2.4.2. La fixation de la protéine Ets dépend de celle de son partenaire transcriptionnel.....	44
2.4.3. La fixation des deux partenaires est indépendante.....	44
3. Régulation de l'activité des protéines Ets.....	45
3.1. Régulation du transport nucléaire.....	45
3.2. Modifications post-traductionnelles.....	46
3.2.1. Phosphorylations.....	46
3.2.2. Modification des groupements thiols.....	47
3.3. Régulation par épissage alternatif.....	48
3.3.1. Epissage alternatif modulant la fixation à l'ADN.....	48
3.3.2. Epissage alternatif modulant l'interactivité avec un cofacteur.....	48
3.3.3. Epissages dont les conséquences sont inconnues.....	49
3.4. Stabilité des protéines Ets.....	50
4. Conclusions.....	51

4ème partie : Rôles physiologiques des protéines de la famille Ets.....

1. Activité des protéines Ets dans les cellules immunocompétentes.....	52
1.1. Ontogenèse des lymphocytes T.....	52
1.2. Activation lymphocytaire T.....	55
1.2.1. Balance Ets-1/Ets-2.....	56
1.2.2. Gènes cibles potentiels.....	57
- régulation de gènes de réponse immune.....	57
- régulation de LTR de virus lymphotropes.....	58
1.3. Différenciation d'autres cellules immunocompétentes.....	58
1.3.1. Lymphocytes B.....	58
1.3.2. Cellules de la lignée monocyttaire.....	58
1.3.3. Mégacaryocytes.....	60

2. Implications de la famille Ets dans le développement embryonnaire.....	60
2.1. Développement de la drosophile.....	60
2.2. Développement du xénope.....	63
2.3. Développement du poulet.....	63
3. Propriétés oncogéniques.....	64
3.1. Transformation des cellules NIH3T3	64
3.2. Activation par insertion rétrovirale.....	65
3.3. Activation par translocation chromosomique.....	67
3.4. Implications plus indirectes dans d'autres tumeurs humaines.....	68
3.5. Hypothèses sur l'oncogénicité de v-Ets.....	68

5ème partie : Régulation transcriptionnelle de gènes de la famille Ets.....

1. Promoteur du gène <i>ets-2</i> humain.....	71
2. Promoteur du gène <i>ets-1</i> humain.....	72
3. Recherche de zones promotrices dans le locus <i>c-ets-1</i> aviaire, membre fondateur de la famille Ets.....	73

RESULTATS..... 75

1. Caractérisation des sites d'initiation de la transcription des ARNm <i>c-ets-1</i>	76
2. Isolement de deux promoteurs du gène <i>c-ets-1</i> de poulet.....	77
3. Autorégulation positive du promoteur p68 ^{c-ets-1}	78

ARTICLE..... 79

<u>DISCUSSION</u>	90
1. Structure des ARNm <i>c-ets-1</i>	92
2. Comparaison avec les autres promoteurs des gènes de la famille Ets.....	93
3. Autorégulation du locus <i>c-ets-1</i>	94
4. L'expression des protéines c-Ets-1 aviaires est-elle mutuellement exclusive ?	95
5. La régulation des protéines c-Ets-1 est-elle en partie transcriptionnelle ?	98
<u>REFERENCES BIBLIOGRAPHIQUES</u>	99

ABREVIATIONS

ADNc	ADN complémentaire de l'ARNm
AEV	avian erythroblastosis virus
AMPc	adénosine monophosphate cyclique
AMV	avian myeloblastosis virus
AP-1	activator protein 1
AP-2	activator protein 2
ARNm	ARN messenger
BFU-E	burst forming unit erythroid
CD	cluster de différenciation
CFU-E	colony forming unit erythroid
CFU-S	colony forming unit in spleen
cMGF	chicken myelomonocytic growth factor
CRE	cAMP responsive element
CSF-1	colony-stimulating factor
DN	double négatif (lymphocyte)
DP	double positif
DSE	dyad symetry element
FEP	fibroblastes embryonnaires de poulet
F-MuLV	Friend murine leukemia virus
HIV-1	human immunodeficiency virus 1
HLH	helix-loop-helix
HMG	high-mobility group
HSV	herpes simplex virus
HTLV-1	human T lymphotropic virus
IL2	interleukine 2
IP3	inositol tri-phosphate
kDa	kilodalton
kpb	kilopaire de bases
LTR	long terminal repeat
MAP kinase	mitogen activating protein kinase
MEP	myb-ets progenitor
MMTV	mouse mammary tumour virus
NK (cellules)	cellule "natural killer"
pb	paire de bases
PCR	polymerase chain reaction
PDGF	platelet-derived growth factor

PEA3	polyoma enhancer 3
PEST (région)	région riche en proline, glutamine, sérine et thréonine
PKA	protéine kinase A
PKC	protéine kinase C
RAV	Rous associated virus
RMN	résonance magnétique nucléaire
RSV	Rous sarcoma virus
SAAB	selected and amplified binding
SFFV	spleen focus-forming virus
SRE	serum responsive element
SRF	serum responsive factor
SV40	simian virus 40
TAF	TBP-associated factor
TBP	TATA-binding protein
TCF	ternary complex factor
TCF-1	T cell factor 1
TCR	T cell receptor
TdT	terminal désoxynucléotidyl transférase
TLCK	p-tosyl-L-lysine chloromethyl ketone
TPA	tumor promoting agent (12-O-tétradécanoyl-phorbol-13-acétate)
uPA	urokinase plasminogen activator

La famille des gènes *Ets* regroupe un vaste ensemble de gènes apparentés dont la conservation phylogénique s'étend des Arthropodes aux Vertébrés supérieurs. Ces gènes semblent tous coder des facteurs de transcription qui possèdent en commun un domaine de liaison à l'ADN d'un type inédit, baptisé pour cette raison le domaine ETS. Bon nombre de ces protéines exercent leur activité transactivatrice en reconnaissant une séquence riche en bases puriques au sein de promoteurs viraux et cellulaires aujourd'hui identifiés en grand nombre. Toutefois, la reconnaissance de ce motif ne semble pas suffire à leur spécificité d'action, puisqu'il apparaît de plus en plus clairement que les protéines *Ets* n'accomplissent pleinement leur fonction régulatrice des gènes qu'en association avec d'autres partenaires transcriptionnels appartenant à différentes grandes familles de facteurs de transcription.

Le locus *c-ets-1* aviaire a été caractérisé le premier, à la suite du clonage de l'oncogène *v-ets* transduit par le virus E26. A partir de ce gène sont synthétisés deux facteurs de transcription, p54^{c-ets-1} et p68^{c-ets-1} qui ne diffèrent que par leur extrémité amino-terminale : celle de la protéine p54^{c-ets-1} n'est codée que par un seul exon absent de l'oncogène *v-ets*, dénommé I⁵⁴, tandis que celle de la protéine p68^{c-ets-1} est codée par deux exons, α et β , colinéaires à l'extrémité 5' de *v-ets*. Cette unique différence entre les deux protéines suffit à attribuer à la protéine p68^{c-ets-1} un domaine transactivateur supplémentaire par rapport à la protéine p54^{c-ets-1} qui n'en comporte qu'un seul.

D'autre part, la distribution tissulaire des deux ARNm correspondants présente plusieurs aspects intéressants : chez le poulet, le niveau d'expression de l'ARNm p54^{c-ets-1} est relativement modéré dans la plupart des tissus, sauf dans les organes lymphoïdes dans lesquels il est très élevé. Par contre, l'expression de l'ARNm p68^{c-ets-1} est restreinte aux vaisseaux sanguins de la rate, ainsi qu'à d'autres tissus d'origine mésodermique, mais elle est rigoureusement réprimée dans les organes lymphoïdes, quelque soit le stade embryonnaire observé ou l'âge de l'animal.

Une meilleure compréhension des effecteurs positifs et négatifs qui permettent une expression appropriée des gènes nécessite d'examiner les mécanismes régulant l'expression même de ces régulateurs. Dans cette optique, la description des mécanismes moléculaires qui gouvernent l'expression du locus *c-ets-1* aviaire pouvait présenter un grand attrait. Les résultats rapportés ici indiquent que les deux transcrits issus de ce gène sont synthétisés à partir de deux promoteurs différents : la transcription s'initie en de nombreux sites d'initiation ce qui reflète l'absence de motif TATA dans les deux cas. Le promoteur initiant la synthèse de l'ARNm p54^{c-ets-1} est riche en bases G et C, et son organisation est globalement conservée chez l'homme. Il est également montré ici que l'activité du promoteur du transcrit p68^{c-ets-1} est stimulée par son propre produit. De plus, sa séquence indique la présence de nombreux sites potentiels de fixation pour des régulateurs d'expression strictement lymphocytaire qui pourraient contribuer à réprimer l'expression de l'ARNm p68^{c-ets-1} dans les organes lymphoïdes.

PRESENTATION DU MEMOIRE

Depuis plusieurs décennies, l'étude des rétrovirus oncogènes éclaire d'un jour nouveau la biologie moléculaire des cancers. Ces virus sont généralement dépourvus de capacités répliquatives autonomes. En effet, en lieu et place des gènes qui sont essentiels à leur réplication, ils ont recombinaison au cours de leurs cycles d'infection, des séquences géniques issues d'une cellule-hôte (oncogène viral ou *v-onc*). C'est en partie à cause des remaniements occasionnés par cette capture, et en partie aussi parce que les virus se répliquent très activement, grâce à une enzyme aux propriétés singulières, que la séquence d'ADN transduite présente des mutations plus ou moins étendues par rapport au gène cellulaire dont elle est issue. Et c'est pour cette raison que ces rétrovirus ont acquis des potentialités transformantes qui ont facilité leur isolement en laboratoire. A travers les oncogènes viraux peut donc être entrevue l'importance de leurs homologues cellulaires normaux (oncogènes cellulaires ou *c-onc*) dans la croissance et la différenciation cellulaires. A ce jour, un grand nombre d'oncogènes viraux ont été isolés à partir de rétrovirus transformants et leurs homologues cellulaires ont pu être décrits. Ces oncogènes cellulaires codent des protéines qui interviennent à différents stades du signal de prolifération/différenciation propagé de la membrane plasmique jusqu'aux confins du génome. Ces processus de croissance et de différenciation sont strictement contrôlés, et leur régulation s'étend sur plusieurs niveaux, dont le contrôle de l'initiation de la synthèse des ARNm et donc de l'expression génique n'est pas le moindre.

L'étude qui est présentée ici porte sur le proto-oncogène de poulet *c-ets-1*, découvert à la suite de la description moléculaire du rétrovirus leucémogène E26. Ce proto-oncogène code deux facteurs de transcription. A ce titre, il est impliqué dans la régulation transcriptionnelle des gènes.

Le contrôle de l'initiation de la transcription dépend de la régulation de l'activité des facteurs de transcription eux-mêmes. Un certain nombre de régulateurs géniques préexistent sous une forme inactive dans la cellule et sont activés en réponse à un stimulus externe. L'étape d'activation qui induit leur activité transcriptionnelle met en jeu des mécanismes de différente nature qui peuvent néanmoins coexister. Ainsi, elle peut être contrôlée par des interactions protéines-protéines. Elle peut également résulter de modifications post-traductionnelles, dont la phosphorylation est la plus couramment employée dans la cellule ou en tout cas la plus étudiée, et peut résulter de la cascade des kinases mise en oeuvre au cours de la transmission d'un signal physiologique.

Un autre moyen évident de contrôler l'activité d'un facteur de transcription est de réguler avec précision son propre promoteur. Ainsi, l'expression de nombreux facteurs de transcription est strictement restreinte à quelques tissus. L'expression de gènes spécifiques peut donc être dans ce cas considérée comme une manifestation de la distribution et de la disponibilité de facteurs de transcription tissu-spécifiques. Mais le problème peut rester entier puisque ce moyen de contrôle n'explique pas pourquoi l'ARNm du facteur considéré est lui-même tissu-spécifique, à moins que sa régulation n'implique un ensemble particulier de régulateurs d'expression moins restreinte dont la combinaison serait unique. On voit donc que de la disponibilité de cette combinatoire de facteurs de transcription à un moment donné du cycle cellulaire, va dépendre la nature des gènes exprimés. C'est cet aspect qui a été développé ici, à propos du proto-oncogène *c-ets-1*, afin de mieux préciser ses modalités d'expression.

Avant de présenter les résultats qui ont pu être obtenus à ce sujet, j'ai souhaité relater brièvement quelques données récentes qui illustrent notre compréhension actuelle de la mécanistique grâce à laquelle les gènes s'expriment de façon appropriée. Ensuite, la famille de facteurs de transcription à laquelle le gène *c-ets-1* appartient sera décrite, en commençant par l'oncogène viral qui a permis la découverte de cet ensemble multigénique.

INTRODUCTION

1ère partie : L'énigme de la transactivation

1. Présentation des protagonistes du complexe d'initiation :

Le contrôle de l'initiation de la transcription des gènes est assuré par des éléments promoteurs qui constituent un site accessible aux ARN polymérases ADN-dépendantes, et les orientent. Les ARN polymérases requièrent la fixation additionnelle de plusieurs polypeptides au voisinage du promoteur afin de démarrer la transcription, et l'ensemble constitue le complexe d'initiation ou encore la machinerie générale de transcription (Figure 1). Ainsi, pour déclencher la synthèse des ARNm d'Eucaryotes, et préalablement au recrutement de l'ARN polymérase II, l'édification progressive de ce complexe débute par la fixation d'une protéine du complexe TFIID, la protéine TBP (TATA-binding protein) (Van Dyke *et al.*, 1988 ; Buratowski *et al.*, 1989) sur une boîte TATA ou sur un autre type de séquence initiatrice de la région promotrice (Pugh et Tjian, 1991 ; Cormack et Struhl, 1992).

Les données qui émergent des études menées sur l'initiation de la transcription et sa régulation indiquent que ces mécanismes sont hautement conservés chez les Eucaryotes puisque certains composants sont interchangeable entre espèces éloignées sur le plan de l'évolution. Un exemple frappant à cet égard est illustré par l'observation que la molécule TBP de levure peut fonctionnellement se substituer au complexe humain *in vitro*, et réciproquement (voir revue de Guarente et Bermingham-McDonogh, 1992). Une avancée récente a vu le jour grâce à la détermination de la structure tridimensionnelle du facteur TBP par cristallographie aux rayons X. Une caractéristique frappante de cette molécule est son assymétrie par rapport à l'axe de la double hélice, qui suggère que TBP pourrait orienter la transcription dans une direction et non dans une autre, à partir de la boîte TATA (Nikolov *et al.*, 1992 ; Chasman *et al.*, 1993 ; Y. Kim *et al.*, 1993 ; J.L. Kim *et al.*, 1993). Sa structure en selle constitue une surface de contact avec les nombreuses autres protéines de la transcription. Pourtant, compte tenu du fait que sa taille est relativement modeste (38 kDa chez l'homme), il faut probablement envisager les interactions contractées avec TBP en terme d'équilibre dynamique, ou sous la forme de mécanismes de type "hit and run".

L'importance du rôle de TBP dans la transcription est soulignée par la découverte que ce facteur est indifféremment recruté au cours de la transcription dépendante des trois types d'ARN polymérases d'Eucaryotes (Sharp, 1992). L'ARN polymérase I promeut la transcription des ARN ribosomiaux, tandis que l'ARN polymérase III initie la synthèse des ARN de transfert, de l'ARN ribosomal 5S et de différents petits ARN cellulaires (voir revue de White et Jackson, 1992). Dans chaque cas, la protéine TBP s'associe à un ensemble de protéines TAFs (TBP-associated factors) propre à chaque type d'ARN polymérase (Comai *et al.*, 1992).

Cette première étape de formation du complexe d'initiation peut être facilitée par le facteur TFIIA (Buratowski et Zhou, 1992a ; Lee *et al.*, 1992) qui serait susceptible de déplacer un ou des inhibiteur(s) associé(s) à TBP (Cortes *et al.*, 1992). L'issue de la compétition pouvant s'établir entre TFIIA et ces inhibiteurs peut constituer une étape limitante de la formation du complexe d'initiation.

Dans un second temps, le facteur TFIIB s'associe au complexe TFIID-promoteur : d'une part, il permet le recrutement de l'ARN polymérase associée au complexe TFIIF ; d'autre part, il spécifie précisément le site de démarrage de la transcription (Buratowski et Zhou, 1992b ; Pinto *et al.*, 1992). De nombreux autres facteurs parmi lesquels TFIIE et H concourent successivement à la mise en place d'un complexe d'initiation de la transcription fonctionnel, bien que le rôle de ces deux derniers facteurs dans l'élongation de l'ARNm soit plus clair qu'au stade de l'initiation. Seuls quatre facteurs du complexe d'initiation ont été clonés : la composante TBP du complexe TFIID (Cavallini *et al.*, 1988 ; Hahn *et al.*, 1989 ; Horikoshi *et al.*, 1989 ; Schmidt *et al.*, 1989 ; Kao *et al.*, 1990), TFIIB (Ha *et al.*, 1991), la petite sous-unité de TFIIF et TFIIE (pour revue, voir Zawel et Reinberg, 1992). Il importe de préciser que le terme TFIIX désigne dans la plupart des cas non pas une protéine mais plutôt un complexe multimérique potentialisant l'activité de l'ARN polymérase II à un moment donné de la formation du complexe d'initiation : par exemple, si l'activité TFIIB implique une seule protéine, l'appellation TFIIH recouvre par contre un ensemble de facteurs d'activités complémentaires. Initialement, ces différents complexes ont été isolés par fractionnement chimique, et l'activité globale de chacun d'entre eux a été déterminée grâce à des expériences de reconstitution du complexe d'initiation *in vitro*.

A ce jour, la compréhension du mode de fonctionnement du complexe d'initiation n'est pas satisfaisante. Pourtant, elle constitue une étape importante pour décrypter la nature des interactions entre les protéines de ce complexe et les facteurs spécifiques de transcription.

2. Présentation des modulateurs transcriptionnels :

La machinerie générale de transcription ne suffit pas à assurer une spécificité de régulation des différents gènes, elle ne permet qu'un taux de transcription basal. Un signal spécifique est délivré aux gènes par l'intermédiaire de séquences régulatrices additionnelles, localisées tant en 5' qu'en 3' des promoteurs, dont la nature et la combinaison déterminent l'expression spatio-temporelle des gènes. En effet, ces éléments sont reconnus par des facteurs de transcription qui stimulent ou répriment l'initiation de la transcription. La diversité des réponses possibles est accrue par le fait qu'un même site nucléotidique de reconnaissance peut lier plusieurs facteurs appartenant à la même famille. Cette diversité repose également sur la capacité que

Type de domaine de liaison à l'ADN	Facteur de transcription	Site reconnu	Références
basique à "Leucine zipper" (bZIP)	GCN4	TGAGTCA	O'Shea <i>et al.</i> , 1991
homéodomaine	<i>engrailed</i>	ATTA	Kissinger <i>et al.</i> , 1990
hélice-boucle-hélice basique (bHLH)	Max	CACGTG	Ferré d'Amaré <i>et al.</i> , 1993
doigt de zinc en C4	récepteur des glucocorticoïdes	AGAACANNNTGTTCT	Luisi <i>et al.</i> , 1991
en C6	Gal4	CGGN11CCG	Kraulis <i>et al.</i> , 1990
autre	GATA-1	T/AGATAA/G	Omichinski <i>et al.</i> , 1993

Tableau 1 : Structure tri-dimensionnelle des domaines de liaison à l'ADN de quelques classes de facteurs de transcription eucaryotes.

possèdent un grand nombre de facteurs de transcription de s'organiser en homo-oligomères et surtout en hétéro-oligomères avec des protéines apparentées ou même complètement distinctes. Il semble que l'activation génique qui en résultera, selon que tel ou tel membre d'une famille de régulateurs est sélectionné, puisse dépendre en partie de la composition globale du promoteur et donc des interactions avec les autres facteurs de transcription recrutés.

Des avancées majeures de ces dix dernières années ont consisté à caractériser et bien souvent à cloner les gènes qui codent les régulateurs transcriptionnels. Elles ont permis l'identification des régions cruciales pour leur fonction que sont les domaines impliqués dans la fixation à l'ADN et les régions d'activation de la transcription.

Le domaine de liaison à l'ADN des facteurs de transcription est généralement une région bien individualisée dont la fixation sur une séquence-cible provoque d'importantes modifications spatiales de la double hélice. La compréhension détaillée de la façon dont les protéines contactent l'ADN requiert entre autres des études physico-chimiques comme celles que permet la cristallographie aux rayons X de complexes ADN/protéine ou la RMN. Ces technologies ont été appliquées avec succès à plusieurs facteurs de transcription dont quelques-uns sont énumérés dans le tableau 1. En dépit de leurs différences structurales, ces domaines exposent tous une hélice α dans le grand sillon de l'ADN. Toutefois, chaque classe de protéine présente cette hélice de reconnaissance d'une façon qui lui est propre, au sein d'une structure particulière bien définie (voir revue de Angrand, 1993).

S'inspirant des résultats obtenus avec des protéines chimériques constituées du domaine de liaison à l'ADN du répresseur bactérien LexA et de l'activateur de levure Gal4 dépourvu de sa région de fixation, de nombreuses études ont montré que les domaines transactivateurs des facteurs de transcription sont souvent séparables et indépendants du domaine de fixation à l'ADN. En effet, chez la levure, les protéines hybrides Gal4-LexA activent efficacement la transcription initiée à partir d'un promoteur contenant un élément de réponse à LexA. De telles expériences ont été étendues à plusieurs régulateurs géniques et indiquent que bon nombre des domaines transactivateurs peuvent être combinés à différents domaines de liaison à l'ADN, suggérant une structure modulaire de ces facteurs. Toutefois, cette observation, souvent vérifiée dans le cas des transactivateurs bactériens ou fongiques, ne s'applique pas systématiquement à ceux des organismes supérieurs : à titre d'exemple, le deuxième domaine transactivateur du récepteur des oestrogènes ne semble pas confiné à une région isolée, mais serait plutôt formé de plusieurs structures discontinues que l'architecture tri-dimensionnelle de la molécule rapprocherait (Lees *et al.*, 1989 ; Webster *et al.*, 1989). Et la définition des domaines d'activation transcriptionnelle des facteurs de transcription de Mammifères n'est pas aussi développée que celle des régions de reconnaissance de l'ADN, en partie parce que cette activité ne peut être

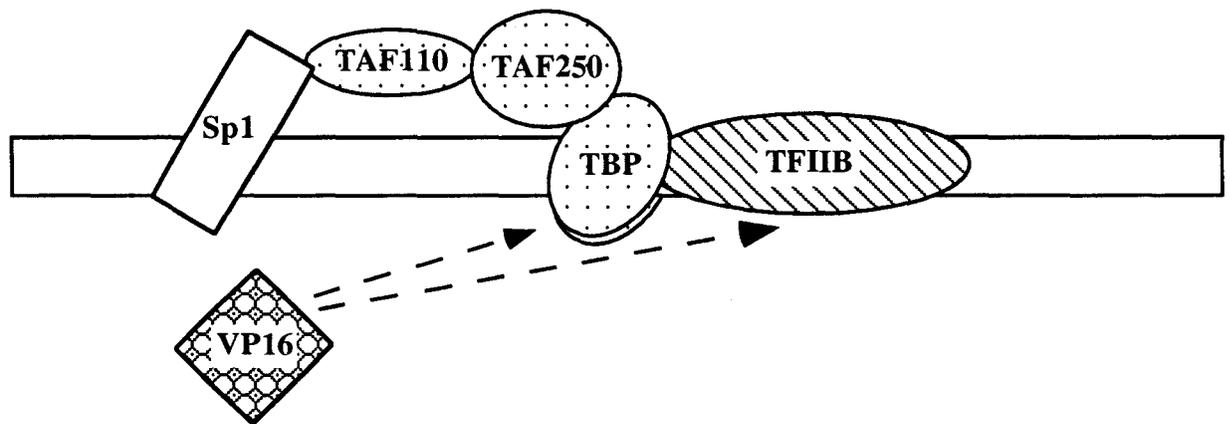


Figure 2 : Exemples d'interaction de deux facteurs de transcription, Sp1 et VP16, avec le complexe d'initiation.

testée qu'*in cellula*, en partie aussi parce que ces protéines peuvent comporter plusieurs sites transactivateurs généralement redondants.

Les analyses les plus détaillées de la nature des domaines transactivateurs ont porté sur des facteurs de levure comme Gal4 et GCN4 ou sur des facteurs viraux comme VP16, et ont mis à jour quelques aspects intéressants qui pu être appliqués à d'autres facteurs de transcription. Certains facteurs comme Sp1, comportent une région activatrice riche en résidus Glutamine. D'autres facteurs possèdent une région très riche en Proline qui constitue un domaine transactivateur. La richesse en acides aminés de type acide, que l'on considérait comme une caractéristique des domaines transactivateurs de facteurs comme Gal4 ou VP16 semble peu à peu céder le pas devant une autre propriété : la disposition de régions riches en acides aminés hydrophobes semble aujourd'hui une caractéristique au moins aussi importante que l'acidité. De fait, les résidus hydrophobes sont disposés de façon analogue au sein des domaines transactivateurs de VP16 et de Sp1 qui pourtant ne sont pas répertoriés dans la même catégorie (Cress et Triezenberg, 1991 ; Régier *et al.*, 1993).

En guise de conclusion, il semble que cette région soit moins structurée que le domaine de fixation à l'ADN et il est possible qu'elle ne dépende d'aucune structure primaire ni d'aucune organisation secondaire particulières, et l'on en vient à postuler que la seule contrainte est une structure suffisamment flexible pour pouvoir s'adapter à la surface des multiples protéines avec lesquelles le facteur peut être amené à interagir (Leuther *et al.*, 1993 ; Van Hoy *et al.*, 1993), ce qui est certes intéressant, mais néanmoins très imprécis.

Si l'on imagine relativement bien comment s'effectue la fixation d'un facteur de transcription à l'ADN, la transactivation n'est par contre appréhendée que par ses conséquences mesurables sur l'expression des gènes, et sa mécanique n'est encore que très hypothétique. Dans ce qui suit, je vais tenter de présenter quelques aspects particuliers que peut revêtir cette énigmatique propriété des facteurs de transcription.

3. Relations avec le complexe d'initiation

L'importance des interactions entre les modulateurs transcriptionnels et les facteurs généraux apparaît si l'on considère que la transactivation consiste à faciliter la liaison de TBP sur la boîte TATA ou la région initiatrice, ou à stabiliser cette liaison pour augmenter le nombre (la fréquence) de complexes productifs. Cette étape accélérerait à son tour le recrutement des autres facteurs généraux dont résulterait finalement de multiples cycles de reconnaissance et d'activation de l'ARN polymérase II. De fait, les transactivateurs viraux VP16 (Stringer *et al.*, 1990) et E1A (Horikoshi *et al.*, 1991) s'associent directement à TBP (Figure 2). Dans ces conditions, TBP pourrait

constituer le facteur limitant de la transcription dont la régulation suffirait à elle seule à dicter le niveau d'expression d'un gène.

Cependant, les protéines TBP humaine et de drosophile ne dirigent que la transcription basale et donc, contrairement au complexe TFIID complet, elles ne répondent pas aux activateurs transcriptionnels (Hoey *et al.*, 1990 ; Hoffman *et al.*, 1990 ; Peterson *et al.*, 1990 ; Pugh et Tjian, 1990 ; Smale *et al.*, 1990). D'une part, cette observation implique que TBP suffit aux interactions avec les protéines du complexe d'initiation. D'autre part, elle indique que les TAFs sont douées de propriétés co-activatrices (ou médiatrices ou adaptatrices), les désignant ainsi comme relais entre les modulateurs transcriptionnels et les facteurs généraux (Dynlacht *et al.*, 1991 ; Tanese *et al.*, 1991 ; Zhou *et al.*, 1992 et 1993 ; Hoey *et al.*, 1993). Parmi les trois TAFs qui ont été clonées chez la drosophile (dTAF80, dTAF110 et dTAF250) et chez l'homme (huTAF250), il apparaît que seule la TAF250 contacte directement TBP (Takada *et al.*, 1992 ; Hisatake *et al.*, 1993 ; Ruppert *et al.*, 1993 ; Weinzierl *et al.*, 1993 ; Zhou *et al.*, 1993). La protéine TAF110 quant à elle effectue un pontage entre le facteur Sp1 et la TAF250 (Hoey *et al.*, 1993 ; Weinzierl *et al.*, 1993) (Figure 2). Cette observation est de toute première importance car la TAF250 joue ici un rôle primordial dans l'assemblage des éléments du complexe TFIID et leur connection avec les facteurs de transcription. Elle l'est tout autant lorsque l'on sait que la TAF250 est codée par le gène CCG1 humain qui n'est autre qu'un facteur de progression dans le cycle cellulaire (Sekigushi *et al.*, 1988 et 1991). La TAF250 pourrait donc concourir à la restriction d'expression de certains gènes au cours de la phase G1 du cycle cellulaire.

Selon toute vraisemblance, les facteurs de transcription contactent le complexe d'initiation à différents stades de son édification. Ainsi, VP16 contacte TFIIB, avec plus d'affinité qu'il n'interagit avec TBP (Lin et Green, 1991 ; Lin *et al.*, 1991 ; Roberts *et al.*, 1993) (Figure 2). TFIIB constitue un autre facteur limitant de l'assemblage du complexe d'initiation (Lin et Green, 1991), d'autant plus que sa participation provoque le recrutement immédiat de l'ARN polymérase II et de TFIIF. Sa participation peut être sensiblement facilitée par la présence d'un transactivateur de type acide comme VP16. Des interactions entre TFIIB et le récepteur des hormones stéroïdes ont également été rapportées récemment (Tsai *et al.*, 1993). Dans les deux cas énoncés ici, ce n'est peut-être plus tant le recrutement de TBP sur la séquence initiatrice qui est stimulé, mais plutôt la stabilisation d'un complexe productif TFIIB-TBP-ADN.

La multiplicité présumée des interactions possibles (avec les TAFs en particulier) reflète sans doute le polymorphisme des domaines transactivateurs qui ont été identifiés jusqu'à présent, et pour lesquels aucune caractéristique structurale définie n'apparaît de façon évidente. On peut certes formuler l'hypothèse selon laquelle la sélectivité des gènes transcrits dépend d'un assemblage particulier de TAFs qui sera recruté par une combinaison appropriée de transactivateurs. Mais, en terme de

transactivation, cette hypothèse n'explique toujours pas avec précision le mécanisme qui déclenche la formation d'un complexe d'initiation fonctionnel stimulant l'activité de l'ARN polymérase II. Elle n'explique pas non plus comment les multiples régulateurs d'un seul promoteur et les TAFs supposées également nombreuses peuvent se sélectionner mutuellement et se rencontrer. L'existence de multiples complexes TFIID ou d'un complexe unique est toujours controversée, même si plusieurs types de complexe TFIID sont purifiables à partir d'une même cellule et peuvent être la cible des différentes classes de domaines transactivateurs (Tasset *et al.*, 1990 ; Davidson *et al.*, 1993). De plus, il convient de ne pas mésestimer des associations transitoires entre TBP et les TAFs qui ne peuvent être visualisées par copurification.

Une autre appellation des protéines du complexe d'initiation les désignent aussi sous le terme de "machinerie générale". En fait, il est permis d'imaginer que ne sont véritablement "généraux" que des composantes comme l'ARN polymérase II ou TBP, et peut-être TFIIA et B. Dans ces conditions, on peut considérer les TAFs non plus comme des facteurs généraux mais plutôt comme des régulateurs géniques à part entière liant éventuellement l'ADN. Sous cet angle, et comme le suggère d'ailleurs l'identité de la TAF250 et de la protéine CCG1, il est fort à parier que les TAFs seront en fait codées par des facteurs de transcription "non généraux". On pourra certes objecter que ces régulateurs transcriptionnels présentent souvent une certaine restriction tissulaire qui est elle-même une manifestation de la distribution et de la disponibilité d'une combinatoire de régulateurs à un moment donné de la physiologie cellulaire. Mais en fait, avec l'avènement de techniques de plus en plus sensibles comme la PCR par exemple, il n'est pas rare de détecter l'expression d'un facteur que l'on croyait absent d'un type cellulaire donné à l'aide de techniques plus archaïques. Par conséquent, cette observation pourrait permettre de postuler l'existence d'une hiérarchie d'expression des régulateurs géniques qui, exprimés constitutivement à taux très faible, ferait office de TAFs. Une fois leur propre transcription activée, ils rempliraient la fonction qui leur est communément attribuée. D'ailleurs, on a vu que le complexe d'initiation peut assurer constitutivement la transcription basale. Mais il semble que cette transcription basale soit souvent (mais pas toujours !) réprimée par les nucléosomes. On peut donc envisager un autre mode de fonctionnement des domaines transactivateurs, qui semble d'ailleurs se surimprimer au premier.

4. Contra-répression nucléosomale :

Depuis longtemps déjà, nul généticien n'ignore l'existence d'une corrélation entre l'activité transcriptionnelle des gènes et des modifications de l'aspect de la chromatine qui se traduisent par une décondensation chromatiniennne résultant

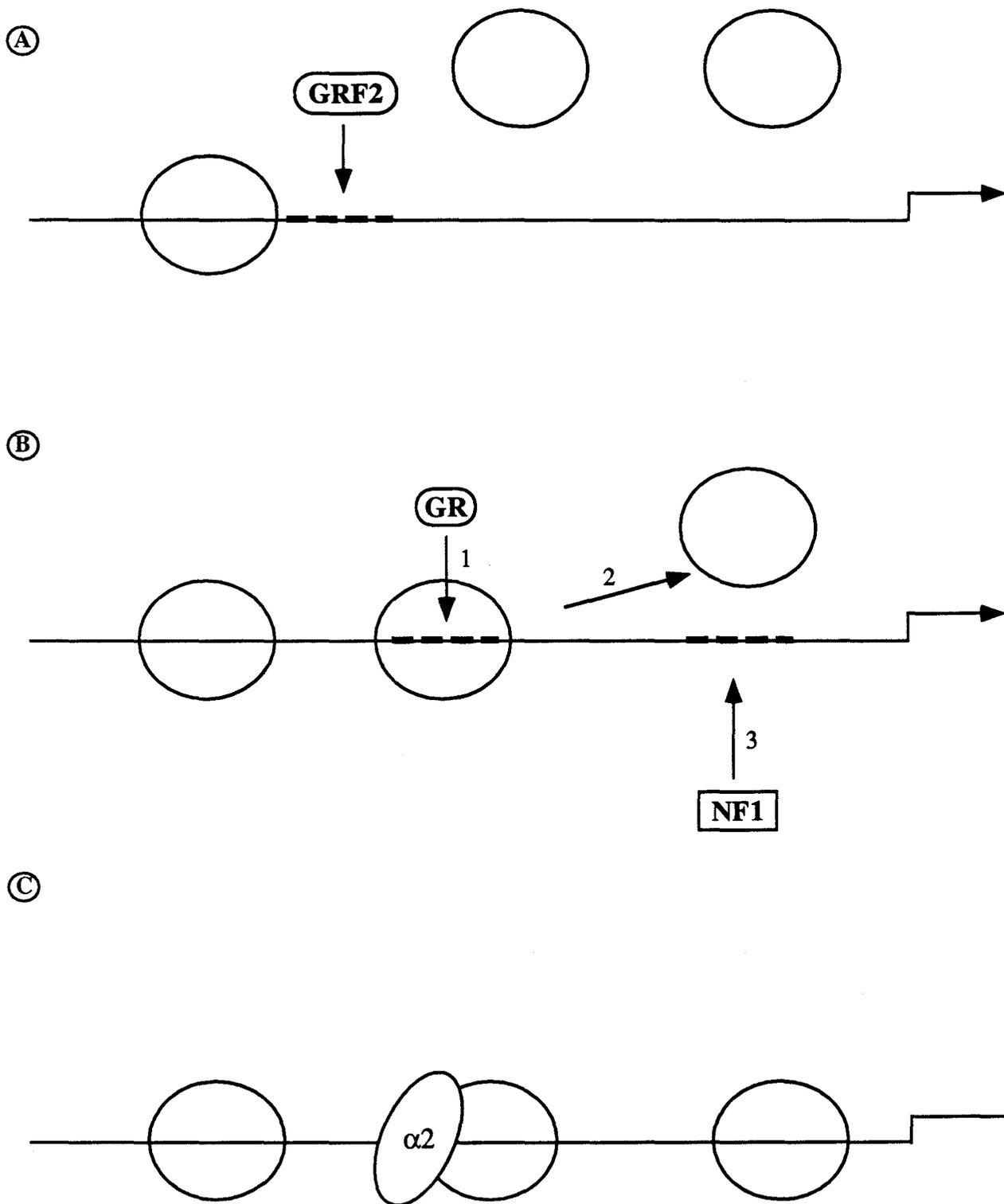


Figure 3 : Deux exemples de levée d'inhibition nucléosomale (A et B) et un exemple de son maintien (C). Les numéros indiquent l'ordre d'enchaînement des évènements. Les régions activatrices sont représentées en pointillés épais, les nucléosomes par des cercles.

d'une déstabilisation des structures nucléosomales. Ces modifications trahissent en fait un rôle dynamique des nucléosomes au cours de la transcription. Les zones d'activité sont d'ailleurs hypersensibles à l'action de la DNase, parce que localement, elles ne sont plus protégées par les histones. Dans ces conditions, il est tentant de définir l'activité transactivatrice, au moins en partie, comme la capacité à contre-balancer l'effet répressur des nucléosomes, soit avant que ceux-ci ne soient assemblés à l'ADN par compétition, soit après leur assemblage par déplacement, ou encore par compétition et déplacement successivement (pour revue, Svaren et Hörz, 1993 ; Workman et Buchman, 1993).

4.1. Exclusion permanente :

Les nucléosomes se fixent à l'ADN d'une manière périodique, indépendamment de la séquence nucléotidique qu'ils obstruent, contrairement aux facteurs transcriptionnels dont la fixation s'accompagne généralement d'une reconnaissance préalable de la séquence cible. Avant l'assemblage des nucléosomes, certains facteurs de transcription peuvent entrer en compétition avec les histones pour l'occupation des séquences régulatrices. Cette situation permet une expression basale constitutive. Elle peut être illustrée par l'exemple du facteur GRF2 de *S. cerevisiae*, qui maintient localement l'ADN sous forme nue : en fixant une séquence activatrice amont, il fait obstacle à l'encombrement nucléosomal des sites adjacents et de la région promotrice (Chasman *et al.*, 1990) (Figure 3A). GRF2 possède une faible activité transactivatrice isolément, mais entre très efficacement en synergie avec les facteurs de transcription dont il facilite la fixation aux séquences régulatrices. Le facteur GAGA de drosophile se fixe quant à lui à la séquence initiatrice du promoteur des gènes *hsp26* ou *hsp70*, et maintient les séquences régulatrices d'amont accessibles aux facteurs de transcription (Gilmour *et al.*, 1989). Ces gènes constitutivement activés à un taux basal conservent néanmoins leur inductibilité par les protéines de choc thermique, sur des zones promotrices dénuées de nucléosomes.

4.2. Induction du déplacement du nucléosome :

Lorsque les nucléosomes sont structurés sur l'ADN, la réactivation de l'expression d'un gène implique que certains facteurs de transcription trouvent néanmoins quelque accès au promoteur. Il existe au moins un facteur de transcription dont la capacité à reconnaître sa séquence-cible malgré la présence d'histones a été démontrée : il s'agit du récepteur aux glucocorticoïdes dont la fixation sur le promoteur du virus MMTV déplace les nucléosomes adjacents et promeut la libération d'un site de fixation pour le facteur NF1 (Pina *et al.*, 1990 ; Archer *et al.*, 1991) (Figure 3B).

4.3. Répression transcriptionnelle :

Les relations que les facteurs de transcription entretiennent avec les histones suggèrent que ces deux types de composants peuvent également rendre compte de la répression transcriptionnelle. En effet, le maintien de la structure nucléosomale ou son organisation par un facteur de transcription peut réprimer la transcription basale. Tel est le mode d'action de la protéine α_2 de levure qui éteint l'expression des gènes de détermination sexuelle en positionnant un ensemble de nucléosomes sur les zones de régulation positive (Shimizu *et al.*, 1991 ; Roth *et al.*, 1992) (Figure 3C).

Les histones constituent donc une barrière non négligeable à la transcription. Néanmoins, leur assemblage fait du nucléosome un élément dynamique de la régulation génique. Même si cette remarque peut sembler traduire des connaissances acquises depuis longtemps (au moins intuitivement), il faut noter que ces dernières peuvent aujourd'hui être considérées avec un regard nouveau. En effet, l'émergence de techniques de reconstitution nucléosomale et de transcription menées complètement *in vitro* permet de commencer à décrypter précisément un mécanisme qui semble essentiel à l'activité des régulateurs géniques. Dans cette optique, l'étude de la transactivation par une protéine chimère Gal4-VP16 est riche d'enseignement : d'une part, elle indique que le domaine transactivateur, porté par VP16, est indispensable non seulement au recrutement de TBP et de TFIIB, mais aussi pour contrer la répression causée par l'histone H1. D'autre part, il apparaît que la dérégulation s'opère en conjonction avec un co-anti-répresseur ribonucléoprotéique qui jouerait le rôle d'accepteur d'histone (Croston *et al.*, 1992).

5. Conclusions :

Le lecteur quelque peu naïf qui s'attendait à ce que soit dévoilée la face cachée de la transactivation doit assurément se trouver très désappointé. Il est vrai que l'on commence à décrire avec une précision croissante certains des protagonistes grâce auxquels cette propriété des facteurs de transcription peut s'exprimer. Les relations existant entre les régulateurs transcriptionnels et l'encombrement de l'ADN dans les régions promotrices peuvent également être appréhendées avec un regard nouveau. Mais, par exemple, nulle trace dans la littérature des réactions biochimiques mises en jeu. Plus encore, il a été montré plus haut que plusieurs étapes de la formation du complexe d'initiation peuvent être des étapes limitantes. Comment sont-elles contrôlées ? Si l'on admet que les domaines transactivateurs interagissent avec différentes protéines du complexe d'initiation, que se passe-t-il ensuite ? Pourquoi cette interaction devrait-elle stimuler la transcription ? On parle d'accroissement de la vitesse

d'échange ou de l'affinité de TBP sur la boîte TATA, mais est-ce bien le cas ou faut-il invoquer un autre processus ? Dans tous les cas, l'on pourrait envisager des modifications structurales de TBP par le transactivateur, qui n'ont jusqu'à présent jamais été rapportées.

Néanmoins, et pour conclure comme il se doit sur une note optimiste, il est clair que ces multiples questions, et bien d'autres encore, ne sont formulables que parce que les interactions qu'entretiennent certains effecteurs de la transcription commencent à être entrevues.

2ème partie : Présentation du rétrovirus leucémogène E26 et propriétés de l'oncogène *v-ets* :

1. Découverte de l'oncogène viral *v-ets* :

En 1983, l'étude du rétrovirus E26, responsable de leucémies aviaires, a permis la caractérisation simultanée par deux laboratoires de l'oncogène viral *v-ets* (Leprince *et al.*, 1983 ; Nunn *et al.*, 1983). Découvert lors d'un criblage de virus à haut potentiel oncogénique chez le poulet (Ivanov, 1962), le virus E26 a révélé plusieurs propriétés tout-à-fait remarquables : outre l'oncogène *v-ets*, son génome contient un deuxième oncogène, *v-myb*, préalablement identifié au sein du rétrovirus AMV (Roussel *et al.*, 1979). Le virus E26 n'est pas un cas unique à cet égard, puisque deux autres rétrovirus aviaires ayant transduit deux oncogènes distincts, les virus MH2 (Coll *et al.*, 1983 ; Jansen *et al.*, 1983 ; Kahn *et al.*, 1983) et AEV (Roussel *et al.*, 1979 ; Saule *et al.*, 1981), sont décrits à ce jour. Toutefois, il est le premier rétrovirus connu pour avoir capté les versions oncogéniques de deux facteurs de transcription modifiés dans son génome.

Contrairement aux autres rétrovirus à deux oncogènes, le virus E26 exprime les séquences qu'il a transduites sous forme d'une protéine de fusion tripartite de 135 kDa de masse moléculaire, la P135 Δ gag-myb-ets. Il utilise donc un nouveau type de subterfuge pour modifier les propriétés normales de chacun des oncogènes qu'il a détournés, illustré par la fusion des séquences *myb* et *ets* (Bister *et al.*, 1982).

Enfin, ses propriétés oncogéniques sont uniques, puisqu'il affecte le développement normal d'au moins deux lignées hématopoiétiques aviaires distinctes (pour revue, voir Pierce, 1989) en prenant pour cible un progéniteur hématopoiétique bipotent très immature, dénommé MEP (Myb-Ets Progenitor) : il provoque principalement une érythroleucémie aigue, et transforme également tant *in vivo* qu'*in vitro*, des cellules de la lignée myéloïde (Moscovici *et al.*, 1981 ; Radke *et al.*, 1982 ; Graf *et al.*, 1992).

Avec l'étude du rôle de *v-ets* dans la transformation des cellules hématopoiétiques par le virus E26, puis la découverte du facteur de transcription codé par le gène *c-ets-1*, débutait une fabuleuse histoire qui devait indiquer des mécanismes originaux par lesquels l'altération d'oncogènes cellulaires et leur coopération contrainte aboutit à la cancérogénèse. Cette étude a également contribué aux prémices de la compréhension de l'hématopoïèse chez le poulet, grâce à l'émergence d'élégants systèmes d'étude *in vitro*.

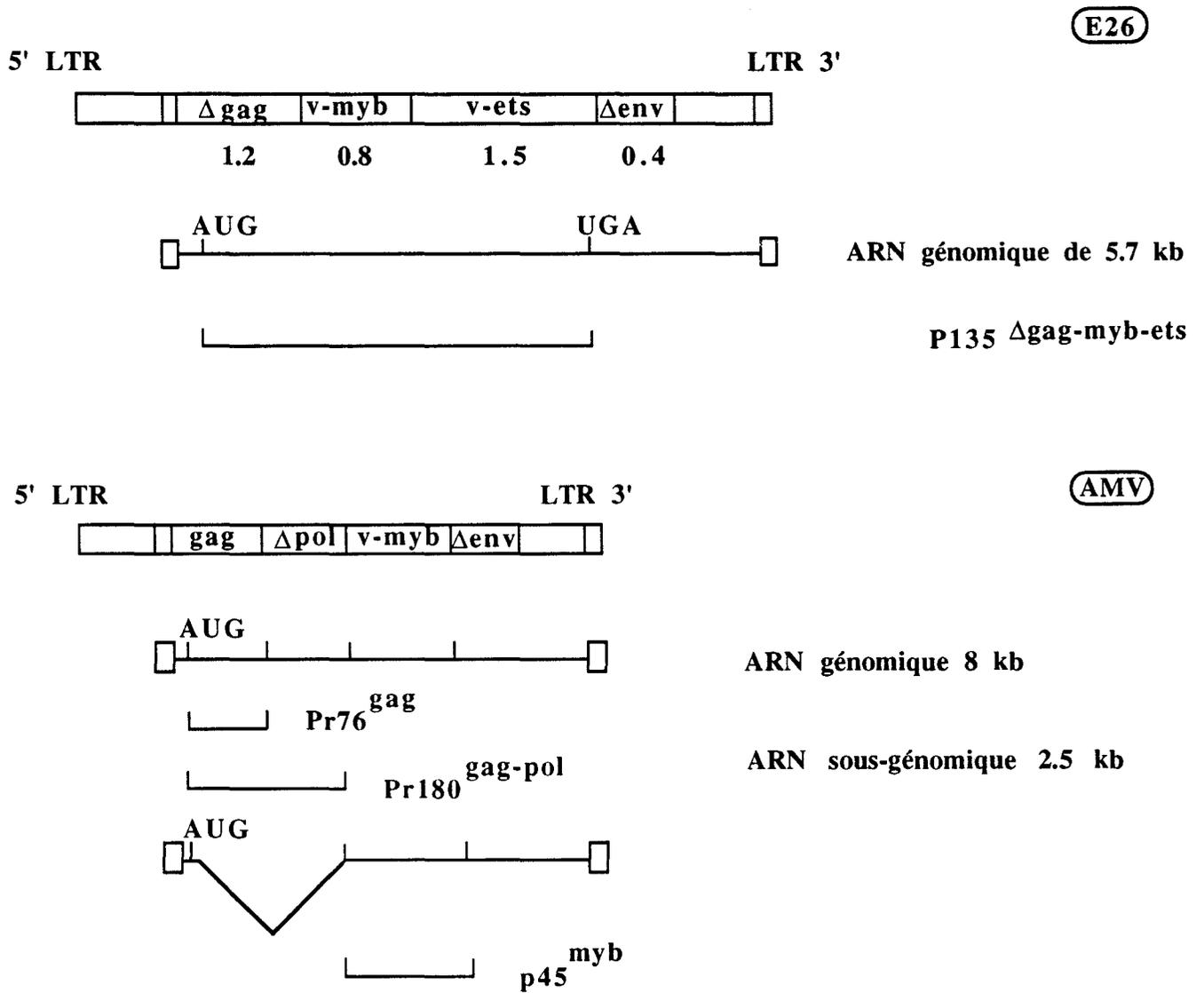


Figure 4 : Comparaison de la structure et de l'expression des rétrovirus E26 et AMV

2. Structure et expression du génome viral :

2.1. Fusion des séquences *v-ets* aux gènes *gag* et *v-myb*

Le virus E26 exprime ses séquences oncogènes sous forme d'un ARN génomique de 5.7 kb couvrant l'ensemble des séquences virales dans l'ordre 5' LTR- Δ gag-myb-ets- Δ env-LTR 3' (Figure 4). La P135 Δ gag-myb-ets est de localisation nucléaire (Klempnauer *et al.*, 1984). Sa traduction s'initie sur un ATG situé au niveau du gène *gag* et s'achève au codon stop situé trente-deux nucléotides en amont de la jonction *ets- Δ env* (Nunn *et al.*, 1983). Le virus E26 est jusqu'à présent le seul virus dont le génome inclue des déterminants *ets*. Par contre, le gène *v-myb* est également transduit par le virus AMV, rendant ce dernier responsable d'une leucémie de type myéloïde chez le poulet, en bloquant des cellules de la voie macrophagique à un stade intermédiaire entre le monoblaste et le macrophage (Durban et Boettiger, 1981). Le contexte structural des séquences *v-myb*^{E26} est donc différent de celui dans lequel se trouve le gène *v-myb*^{AMV}, ce qui rend compte à l'évidence des nouvelles propriétés transformantes acquises par le virus E26 par rapport au virus AMV. En effet, non content d'être fusionné à la séquence *v-ets*, *v-myb*^{E26} se trouve également en phase en 5' avec des déterminants du gène *gag*. D'autre part, sa structure propre présente quelques modifications, puisqu'il est tronqué de 106 nucléotides en 5' et de 240 nucléotides en 3' par rapport à l'oncogène *v-myb*^{AMV}. Au regard des séquences transduites, il est plus proche que *v-myb*^{AMV} de leur progéniteur cellulaire commun *c-myb*, dont il ne diffère que par une substitution nucléotidique (Nunn *et al.*, 1984).

2.2. Genèse du virus E26 :

2.2.1. Homologies entre les gènes *v-myb*^{E26} et *c-myb* :

L'oncogène *v-myb*^{E26} est homologue aux six premiers exons codants du gène cellulaire *c-myb*. Dans le génome du virus E26, le premier exon est amputé de son extrémité 5', et la jonction *gag-myb* se fait de telle sorte que le cadre de lecture correspondant à celui du gène *c-myb* soit préservé (Nunn *et al.*, 1984). En 3', l'homologie avec le gène *c-myb* s'achève dans l'exon 6, et le cadre de lecture se poursuit dans le gène *v-ets* (Leprince *et al.*, 1983 ; Nunn *et al.*, 1983).

2.2.2. Homologies entre les séquences *v-ets* et *c-ets-I* :

Les séquences *v-ets* transduites par le virus E26 sont originaires d'un gène cellulaire de poulet, le gène *c-ets-I*. Nous ne détaillerons pas ici la structure de ce gène, puisqu'il représente notre principal sujet d'étude et qu'à ce titre, ses propriétés seront présentées de façon approfondie plus loin. Néanmoins, certains éléments doivent d'ores

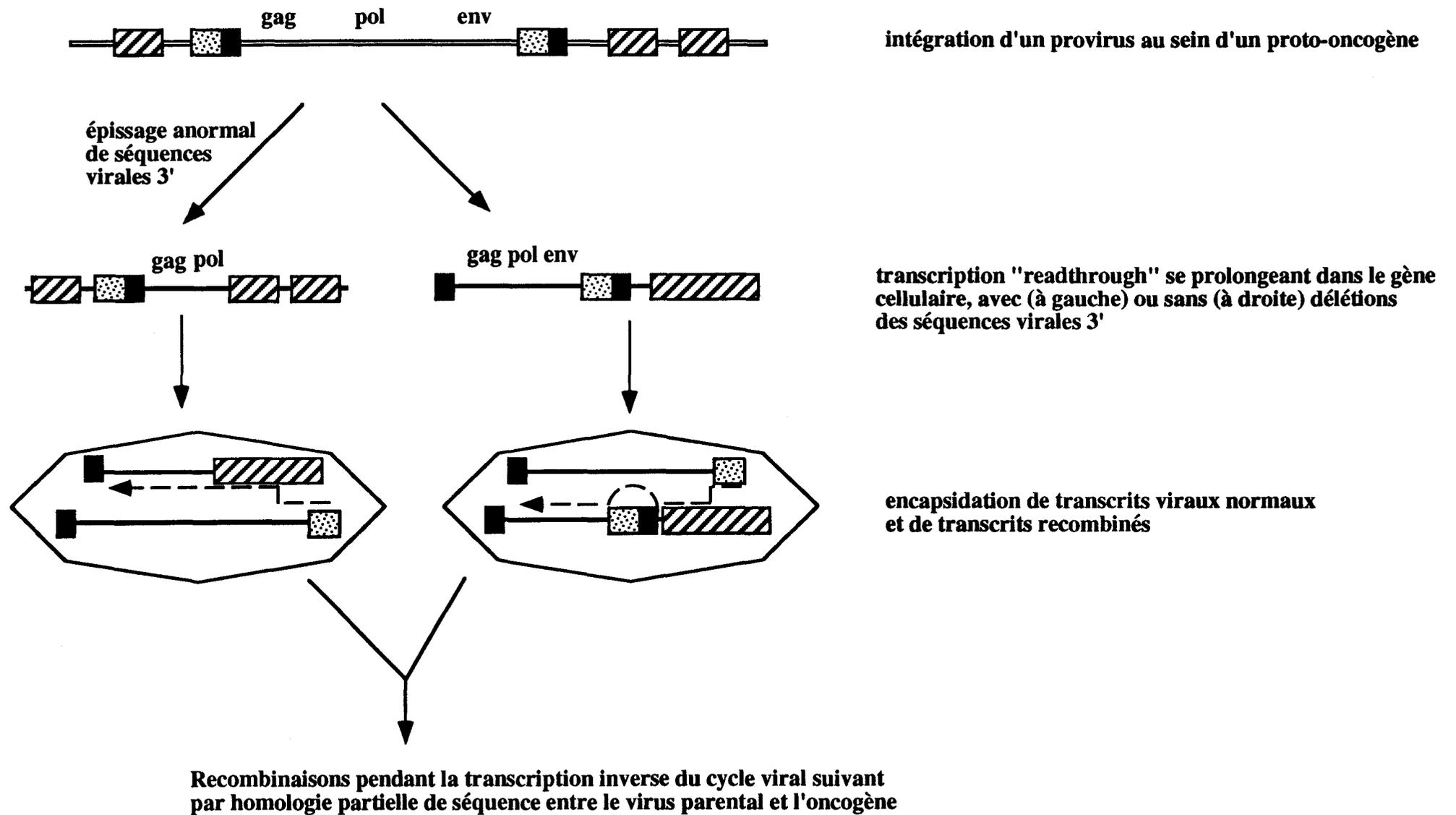


Figure 5 : Mécanismes hypothétiques de formation des rétrovirus oncogènes

et déjà être dévoilés, afin d'approcher les mécanismes qui ont présidé à la création du virus E26.

Deux équipes ont récemment testé les différentes hypothèses émises sur les mécanismes de transduction des oncogènes (Swain et Coffin, 1992 ; Zhang et Temin, 1993a ; voir aussi revue de Temin, 1993) (Figure 5). Après recombinaison d'un provirus au voisinage d'un proto-oncogène, se produit une transcription se prolongeant à travers le gène cellulaire au mépris du signal de polyadénylation viral (transcrits "readthrough"), accompagnée ou non d'un épissage anormal entre les gènes viraux et l'oncogène. Les ARN néoformés et les transcrits rétroviraux normaux sont encapsidés. Au cours du cycle viral suivant, des recombinaisons dues à des changements de matrices d'un brin d'ARN sur l'autre créent de nouveaux virus. Dans une revue récente, Zhang et Temin (1993b) les classent en deux catégories, selon que les deux séquences se sont couplées par recombinaison homologue (groupe 1) ou bien que la jonction 3' entre les séquences virales et cellulaires soit séparée par une insertion (groupe 2).

Les séquences *v-ets* couvrent l'ensemble des régions codantes du gène *c-ets-1* et ne présentent que deux mutations ponctuelles par rapport à leur homologue cellulaire. Cependant, les amino-acides de l'extrémité amino-terminale de v-Ets sont codés par les onze derniers nucléotides 5' non-codants de *c-ets-1*. D'autre part, les treize derniers acides aminés carboxy-terminaux codés par le gène *c-ets-1* sont remplacés par seize acides aminés complètement différents dans v-Ets (Figure 6A). Récemment, deux groupes ont découvert que ces acides aminés sont en fait codés par le brin opposé à celui qui code les treize derniers acides aminés de la protéine cellulaire Ets-1 (Lautenberger et Papas, 1993 ; Leprince *et al.*, 1993). Cette homologie s'achève au premier codon stop possible, situé 16 codons en aval du point de divergence *c-ets-1/v-ets* (Figure 6B). Il semblerait que *v-ets* soit le seul oncogène viral connu à ce jour ayant inversé l'extrémité carboxy-terminale de son progéniteur cellulaire. Il illustre un mécanisme nouveau et original qui, comme nous allons le voir par la suite, n'est pas sans incidence sur ses propriétés biologiques. S'il n'est pas encore possible d'expliquer comment cette inversion s'est opérée, il est néanmoins probable qu'un mécanisme de changement de matrice l'a permise.

Trente-deux nucléotides en 3' du codon stop se trouve ce qui avait été défini comme la frontière *ets/env* puisque commence alors une zone homologue au gène *env* du virus RSV PrC (Prague C) (Nunn *et al.*, 1983). Cette insertion a amené Zhang et Temin (1993b) à classer E26 parmi les virus du groupe 2. En fait, la comparaison des trente-deux nucléotides intermédiaires avec les séquences issues de banques informatiques de données (Lautenberger et Papas, 1993 ; Leprince *et al.*, 1993) a révélé une forte homologie avec le gène *env* d'un autre rétrovirus, RAV-2 (Bova *et al.*, 1988). Cette homologie se poursuit avec les derniers codons de *v-ets* et par voie de conséquence, elle est complémentaire du brin codant de *c-ets*, sur environ six bases

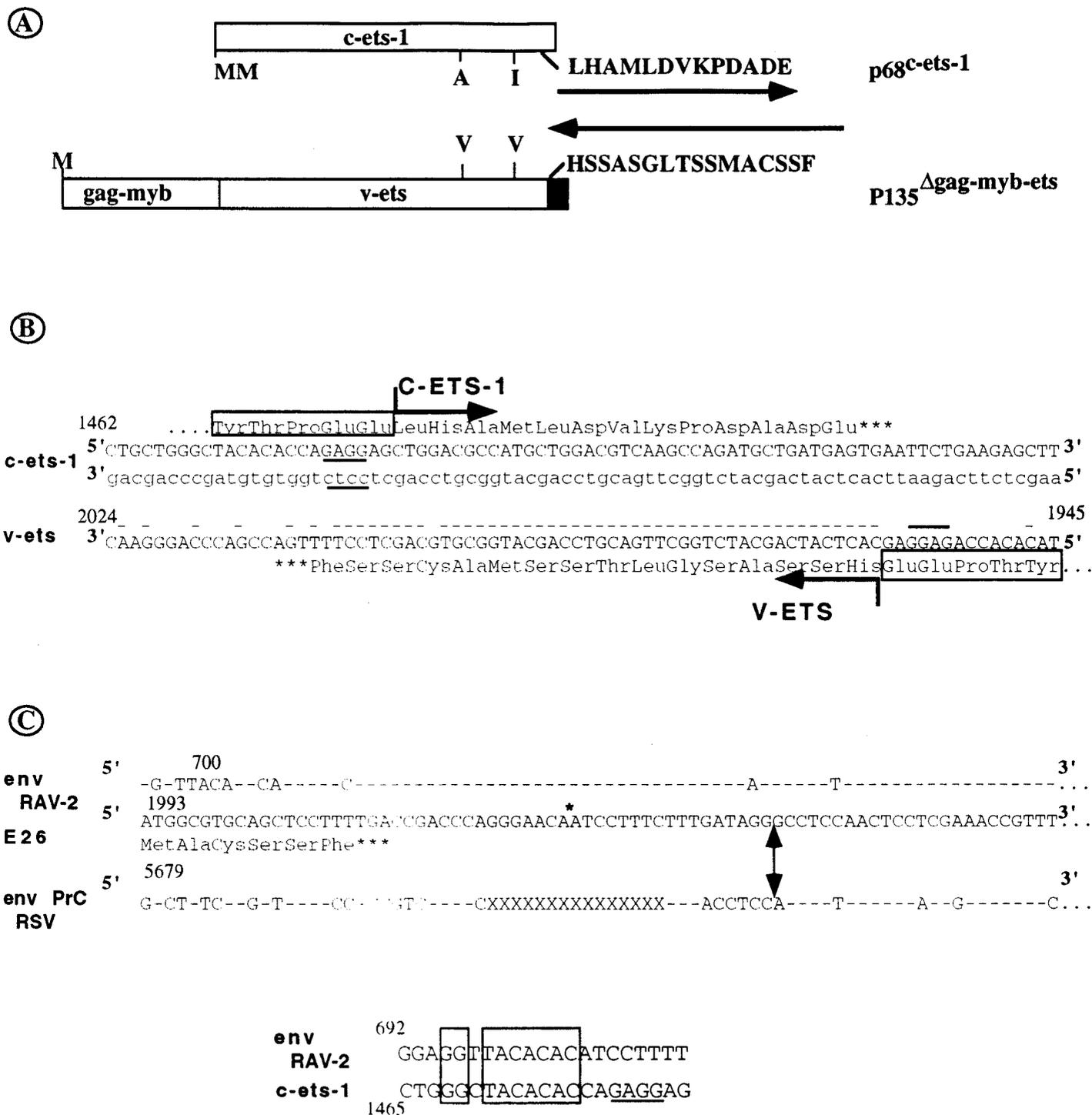


Figure 6 : Comparaison des extrémités carboxyterminales des protéines c-Ets-1 et P135 Δgag-myb-ets

situées juste en amont du point de divergence *v-ets/c-ets* (Figure 6C). Elle suggère que c'est un rétrovirus de type RAV-2 qui a présidé à la naissance du virus E26.

2.2.3. Point de recombinaison entre les gènes *v-myb* et *v-ets*:

La jonction *myb/ets* se fait au milieu de l'exon 6 du gène *c-myb*. L'examen de la séquence du point de recombinaison révèle une homologie avec une séquence consensus donneur d'épissage (CAAAGGgtagagt) (Mount, 1982 ; Leprince *et al.*, 1988). L'extrémité 5' de *v-ets* est colinéaire au premier exon codant de *c-ets-1*. C'est donc probablement un épissage aberrant ayant joint un site cryptique donneur dans l'exon 6 de *c-myb* et un site accepteur normal du premier exon codant de *c-ets-1* qui a permis la fusion en phase des séquences *myb* et *ets*.

Il est clair que si la description moléculaire de ce mécanisme permet d'appréhender la façon dont se sont fusionnées les séquences *v-ets* et *v-myb* au sein du génome du virus E26, de nombreuses questions restent en suspens. En particulier, quelle a été la première séquence transduite, *myb* ou *ets* ? Il importe de préciser que les rétrovirus AMV et E26 ont visiblement acquis leurs séquences *v-myb* indépendamment, comme l'indique par exemple le fait que leurs LTR respectifs ne présentent qu'une médiocre homologie. On ne peut donc pas considérer le virus E26 comme un virus de seconde génération, produit par suite de la transduction de séquences *v-ets* dans le virus AMV.

2.2.4. Modes de recombinaison des séquences *myb* et *ets* :

Il est probable que l'intégration d'un virus contenant déjà l'un des deux oncogènes au voisinage du second oncogène cellulaire explique la création du virus E26. En effet, il peut se produire dans n'importe quel type cellulaire et ne requiert pas la co-transcription dans une même cellule des gènes *myb* et *ets*. Toutefois, des résultats récents obtenus par hybridation *in situ* suggèrent la coexpression de ces deux types de transcrits au cours de l'embryogénèse du poulet (Quéva *et al.*, 1993), ce qui soulève la possibilité d'un épissage illégitime des ARNm *c-myb* et *c-ets-1*.

Le fait que l'extrémité carboxy-terminale de *v-Ets* est codée par le brin complémentaire de l'extrémité carboxy-terminale de *c-Ets* vient d'être évoqué. Faut-il pour autant imaginer que le virus E26 s'est formé graduellement et invoquer la possibilité de l'existence d'un virus intermédiaire possédant des séquences fusionnées *gag-myb-c-ets-1* ? Bien que les étapes successives de la genèse de ce rétrovirus tel que nous le connaissons actuellement demeureront spéculatives, puisque les premiers isolats, réalisés en 1962 en Bulgarie, sont définitivement perdus, il convient de noter que l'expérience de son étude montre qu'il est devenu de plus en plus transformants au fil des passages successifs. Par conséquent, l'existence de virus intermédiaires ayant

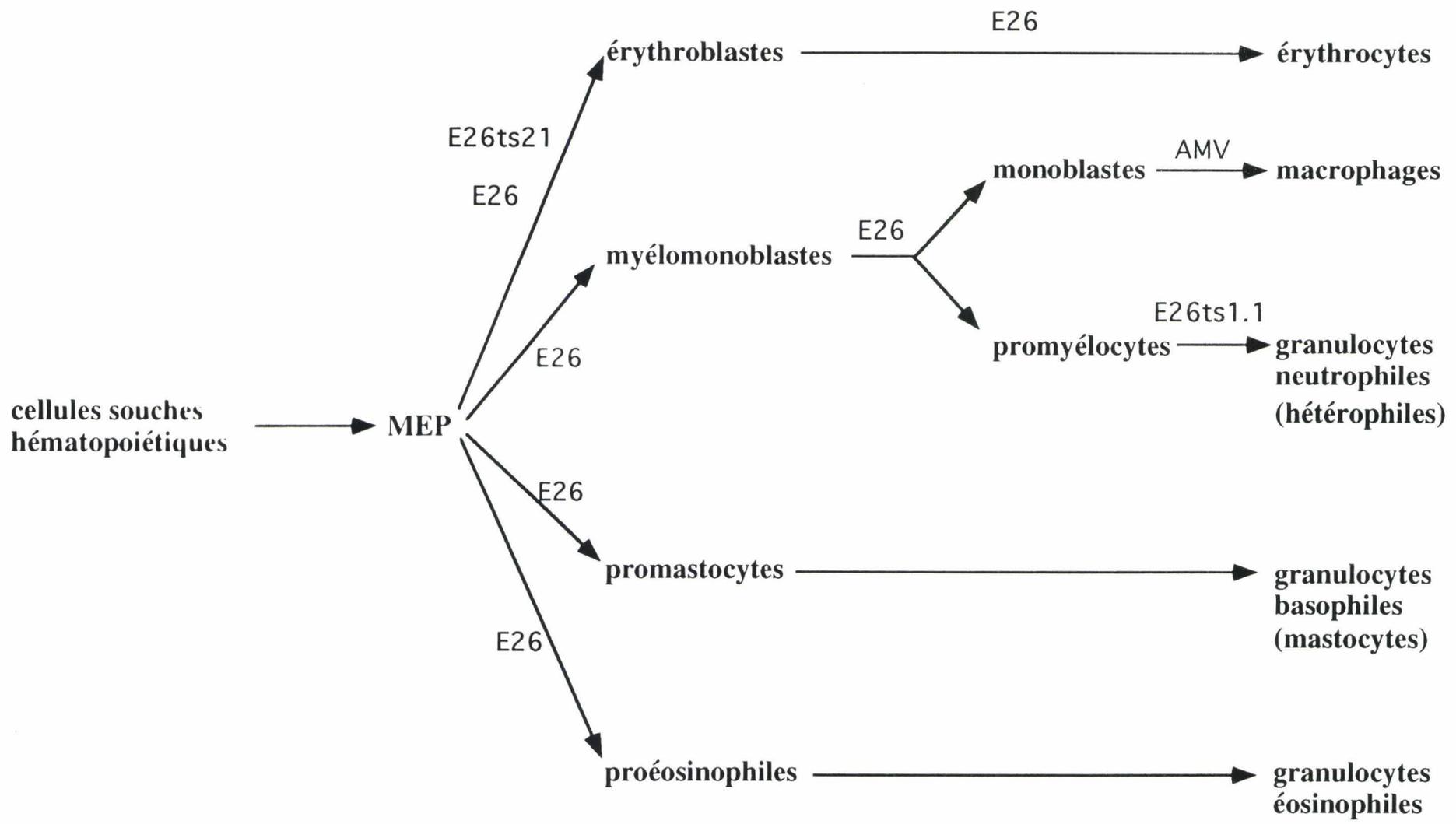


Figure 7 : Les étapes de la différenciation hématopoïétique du poulet bloquées par les virus aviaires E26 et AMV, ainsi que par les mutants thermosensibles E26ts 1.1 et ts 21.

accumulé des mutations dont celle de l'extrémité carboxy-terminale, est hautement probable.

3. Transformation des cellules hématopoiétiques :

3.1. Les cellules-cibles du virus E26 :

C'est dans un délai très bref, de l'ordre de 3 à 4 semaines, que s'installe une leucémie chez le poulet infecté par le virus E26. Le sang des animaux leucémiques est envahi par une proportion impressionnante de précurseurs de la lignée érythrocytaire, dont la différenciation est bloquée de façon hétérogène. Dans une moindre mesure, des myélomonocytes très immatures, précurseurs communs aux macrophages et aux granulocytes neutrophiles (hétérophiles, chez le poulet), sont observés en quantité anormale, aggravant les troubles de l'érythropoïèse d'une discrète myéloblastose (la figure 7 résume l'hématopoïèse du poulet). Ces effets peuvent être reconstitués *in vitro*, après infection de cellules de moelle osseuse ou dérivées de blastoderme en présence de cMGF dont dépendent la survie et la prolifération des précurseurs myéloïdes, ou de sérum de poulet anémié nécessaire au maintien en culture des cellules érythroïdes.

Ces observations amenèrent rapidement à s'interroger sur la nature des cellules-cibles du virus E26 : le virus est-il capable d'infecter et de transformer indépendamment des cellules-souches érythroïdes et des précurseurs myéloïdes ? Ou bien s'attaque-t-il plutôt à un progéniteur bipotent, pas encore engagé dans l'une de ces deux voies ?

Sur la base de critères morphologiques et immunocytochimiques, trois types de colonies ont pu être identifiés à partir de cultures de blastoderms infectés par le virus E26 : des colonies de type érythroïde, des colonies myéloïdes et des colonies mixtes composées de cellules exprimant soit des marqueurs érythroïdes soit des marqueurs myéloïdes (Moscovici *et al.*, 1983 ; Graf *et al.*, 1992). Les colonies mixtes ne résultent pas de la contamination d'un type cellulaire par l'autre, ainsi que l'a démontré l'analyse des sites d'intégration rétrovirale (Graf *et al.*, 1992). En effet, le fait que toutes les cellules d'une même colonie, qu'elles soient de type myéloïde ou érythroïde, présentent un site d'intégration unique et identique établit sans conteste leur clonalité. Cette population, dénommée MEP, constitue donc un ensemble de précurseurs hématopoiétiques très immatures et multipotents (Graf *et al.*, 1992).

En absence de tout traitement particulier, ces cellules expriment des marqueurs de type érythroïde, tels l'histone H5, le facteur de transcription GATA-1 ou l'antigène Ebl (erythroblast antigen) (McNagny *et al.*, 1992). Dans le sang des animaux leucémiques, des précurseurs érythroïdes bloqués à tous les stades de leur différenciation peuvent être observés. La MEP constitue-t-elle une cellule-cible unique du virus E26 dont la maturation serait médiocrement bloquée par le virus ? Ou bien le

virus E26 peut-il infecter de surcroît des cellules plus différenciées ? Le fait que les MEP ne présentent qu'un faible taux de différenciation spontanée *in vitro* constitue un argument favorable à la deuxième hypothèse.

Sous l'action d'inducteurs tels le TPA ou après surinfection par un oncogène de type kinase (*v-sea*, *v-erbB*), les MEP s'engagent dans la voie myélomonocytaire ou éosinophilique. Cette dernière observation n'est qu'expérimentale puisque, pour le moment, des précurseurs d'éosinophiles ne sont pas détectés dans le cours naturel de la leucémie. Elles permettent néanmoins de s'interroger sur la chronologie normale d'apparition des éosinophiles et des myéloblastes, et soulèvent la possibilité d'une conversion phénotypique des deux lignées.

Enfin, des études récentes ont permis de reconnaître, au sein de cultures de moelle osseuse infectée par le virus E26, une population cellulaire présentant certains critères morphologiques et histologiques propres aux granulocytes basophiles (mastocytes) (Domenget *et al.*, 1992). Par conséquent, à ce jour, il semble que toutes les voies de la différenciation hématopoiétique soient affectées par la transformation par le virus E26, à l'exception notable des voies lymphoïdes. Plus précisément, des marqueurs précoces de ces des précurseurs lymphoïdes ne sont pas décrits chez le poulet, aussi serait-il difficile de reconnaître des précurseurs transformés. Une autre possibilité est aussi que ces cellules potentiellement transformées soient peu viables *in vitro*.

3.2. Contribution respective des gènes *v-ets* et *v-myb* au pouvoir transformant du virus E26 :

3.2.1. Rôle de l'oncogène *v-myb* :

La présence de deux oncogènes dans un rétrovirus capable de transformer deux lignées cellulaires distinctes a très tôt suggéré que chaque oncogène était responsable de la transformation d'une des deux lignées. Trois arguments majeurs contribuaient à associer le gène *v-myb* à la transformation myélomonocytaire :

- *v-myb* est la seule séquence exogène transduite par le virus AMV. Or, ce virus bloque uniquement la différenciation des cellules hématopoiétiques à un stade intermédiaire entre les monoblastes et les macrophages (Moscovici, 1975).

- un mutant du virus E26, Δ 4E26, délété de la majeure partie des séquences *v-ets*, demeure capable de transformer des myéloblastes mais n'induit plus de leucémies érythroïdes (Nunn et Hunter, 1989).

- enfin, a pu être isolé un mutant du virus E26 thermosensible pour la transformation myéloïde, dénommé ts 21 (Beug *et al.*, 1984). A la température restrictive de 42°C, des myéloblastes infectés par le virus E26 ts21 se différencient *in vitro* en cellules de type macrophages résistants au cMGF, peut-être par absence de

récepteurs, mais capables de sécréter ce facteur de croissance. La transformation des érythroblastes n'est pas affectée. L'analyse de la séquence du provirus ts 21 dévoile trois mutations localisées dans le gène *v-myb*, dont deux sont silencieuses, tandis que la troisième est située dans son domaine de liaison à l'ADN (Frykberg *et al.*, 1988). Réintroduite dans un virus sauvage, cette seule mutation suffit à conférer un phénotype thermosensible à des myéloblastes infectés (Li *et al.*, 1989).

3.2.2. Contribution de l'oncogène *v-ets* :

Les arguments énoncés précédemment attribuent donc la transformation des cellules myéloïdes au gène *v-myb*, imputant ainsi par élimination la transformation des cellules érythroïdes au gène *v-ets*. La création d'un autre mutant thermosensible pour la transformation érythroïde devait définitivement ratifier cette proposition (Golay *et al.*, 1988). *In vitro*, les érythroblastes infectés par le mutant ts1.1 se différencient en érythrocytes, à la température restrictive de 42°C. L'examen de la séquence nucléotidique du provirus ts 1.1 révèle qu'une seule mutation ponctuelle, dans le domaine de liaison à l'ADN du gène *v-ets*, est responsable du phénotype thermosensible des cellules transformées. De plus, la construction d'un virus chimérique gag-*v-erbA/v-ets* a récemment montré que l'oncogène *v-ets* libère les cellules érythroïdes transformées de leur dépendance vis-à-vis des facteurs de croissance (Metz et Graf, 1992). A cet égard, *v-Ets* se substituerait fonctionnellement à une tyrosine kinase du type de celle qui est codée par *v-erbB*, qui est le partenaire normal de *v-erbA* au sein du rétrovirus leucémogène AEV.

Cependant et de façon tout-à-fait inattendue, les cellules myéloïdes transformées par le mutant ts 1.1 présentent un phénotype altéré par rapport à leurs équivalents transformés par le virus E26 sauvage. En effet, à 37°C aussi bien qu'à 42°C, ces cellules ressemblent à des promyélocytes, c'est-à-dire à des précurseurs d'hétérophiles. Donc, par le fait d'une simple mutation ponctuelle dans le gène *v-ets*, le virus E26ts 1.1 transforme des cellules d'une autre lignée que celle qui est affectée par le virus sauvage. De plus, des macrophages infectés par ce mutant subissent une transdifférenciation en promyélocytes, ce qui suggère que les cellules des deux voies granulocytaire et monocyttaire sont très étroitement apparentées. Une autre implication de ces observations est qu'un progéniteur bipotent (ici, le myéloblaste) engagé dans l'une des deux voies ne l'est donc pas de façon irréversible.

Si les résultats obtenus grâce à la création du virus mutant ts 1.1 montrent à l'évidence que l'oncogène *v-ets* est responsable de la transformation des cellules érythroïdes *in vitro*, ils indiquent de surcroît que son rôle dans la transformation des précurseurs myéloïdes n'est pas anodin. En fait, il apparaît que ce gène, sans être nécessaire à la transformation myéloïde, module pourtant le phénotype des cellules

myéloïdes transformées. L'influence du gène *v-ets* sur la transformation myéloïde était déjà suggérée par la comparaison des myéloblastes transformés par le virus AMV *versus* le virus E26. Les premiers peuvent se différencier spontanément en macrophages (Moscovici et Gazzolo, 1982), tandis que la maturation des seconds est strictement bloquée. L'oncogène *v-ets* pourrait donc être, au moins en partie, responsable de cet effet drastique sur la différenciation des myéloblastes transformés par le virus E26.

3.2.3. Effets des mutants thermosensibles *in vivo* :

Les poulets ayant une température interne de 41, 5°C, ils peuvent être utilisés pour tester la perte d'oncogénicité des virus thermosensibles à la température restrictive. Le virus ts 21, muté dans le gène *v-myb*, est parfaitement leucémogène, mais l'érythroblastose se déclenche avec un temps de latence légèrement plus important qu'avec le virus sauvage (une quarantaine de jours contre une trentaine), ce qui pourrait suggérer une influence inattendue de l'oncogène *v-myb* sur la transformation érythropoïétique. Les poulets infectés développent également une leucémie myéloïde parce que dans certaines zones, leur température corporelle n'excède pas 37°C, mais les cellules transformées conservent le phénotype thermosensible lorsqu'elles sont replacées *in vitro* (Beug *et al.*, 1984).

Les altérations du pouvoir transformant du mutant ts 1.1, muté dans *v-ets*, observées *in vitro* sont reflétées *in vivo* par le phénotype des cellules transformées, érythroïde et promyélocytaire. L'érythroblastose apparaît deux fois moins rapidement que chez les animaux infectés par le virus E26 sauvage et, comme on peut s'y attendre, les érythroblastes récoltés du sang des animaux leucémiques semblent beaucoup plus avancés dans leur différenciation.

Par conséquent, si l'on peut conclure que le gène *v-myb* est suffisant pour transformer les myéloblastes, tandis que l'intégrité structurale du gène *v-ets* semble nécessaire au blocage des érythroblastes à un stade immature, les effets du virus E26 sur la transformation des deux lignées ne peuvent pas être séparés de façon aussi radicale.

3.3. Importance de la fusion des séquences *v-myb* et *v-ets* sur le pouvoir transformant :

La prolifération incontrôlée de précurseurs hématopoïétiques très immatures, les MEP, est la cause prédominante de décès des poulets infectés par le virus E26 et, même si le gène *v-ets* en est directement responsable, le gène *myb* est capable d'en moduler la latence. Ces deux observations soulèvent la possibilité d'une coopération entre les deux oncogènes. Mais selon quelles modalités ? Leurs effets respectifs sont-ils

Construction	Taux de croissance des érythroblastes	Durée de vie
v-myb	+	+
v-ets	+	ND
myb/ets	++	++
myb-ets	+++	+++

Tableau 2 : Importance de la fusion protéique sur les propriétés transformantes du virus E26 *in vitro*

simplement additifs ou bien leur fusion permet-elle l'émergence de nouvelles propriétés transformantes encore supérieures à la somme du potentiel oncogénique de chacun des deux gènes viraux ? Pour apprécier les conséquences de la fusion des deux oncogènes, des mutants du virus E26 exprimant les protéines Δ gag-myb et ets séparément (mutant Myb/Ets) ont été construits (Metz et Graf, 1991a) (Tableau 2). Leur pouvoir transformant a été comparé à celui de virus n'exprimant que la protéine Δ gag-myb seule ou la protéine v-Ets seule, ainsi qu'à celui du virus E26 sauvage (désigné Myb-Ets). Les résultats obtenus *in vitro* indiquent que tous les virus exprimant des séquences v-myb sont capables de transformer des cellules myéloïdes, mais aussi, dans une moindre mesure, des cellules érythroïdes issues de l'infection de blastodermes. Les mutants myb/ets transforment les précurseurs érythroïdes. Cependant, si le virus E26 sauvage ne transforme que des cellules érythroïdes immatures, les mutants myb/ets permettent d'obtenir des cellules dont le potentiel de différenciation spontanée est élevé. Enfin, le taux de croissance et la durée de vie de ces précurseurs sont supérieurs s'ils sont infectés par le mutant Myb/Ets plutôt que par les virus n'exprimant que l'un des deux oncogènes, ce qui démontre que leurs effets sont cumulatifs. Croissance et durée de vie s'accroissent encore si les cellules sont infectées par le virus E26 sauvage. Par conséquent, les deux oncogènes coopèrent également en *cis*, au sein de la protéine de fusion. *In vivo*, seul le virus E26 sauvage est leucémogène (Metz et Graf, 1991b). Cependant, quelques animaux infectés par le mutant Myb/Ets développent des leucémies après un temps de latence important. L'analyse de la séquence du génome de ces mutants indique qu'en fait, un événement de recombinaison a naturellement refusionné les protéines Δ gag-myb et Ets, en déléant des portions variables de l'extrémité carboxy-terminale de v-Myb et de l'extrémité amino-terminale de v-Ets. Cette observation inattendue suggère qu'une importante pression de sélection pèse sur les multiples cycles réplicatifs des mutants pour favoriser l'émergence et la dissémination d'un virus refusionné de type E26 sauvage.

3.4. Domaines fonctionnels importants pour la transformation hématopoiétique :

Par analogie avec les protéines codées par les gènes cellulaires *c-ets-1* et *c-myb*, une fonctionnalité particulière a pu être attribuée à différents domaines de la P135 Δ gag-myb-ets de manière spéculative mais néanmoins hautement probable. Le domaine de liaison à l'ADN (DB) de la partie v-Myb est en phase avec la partie Δ gag à l'extrémité amino-terminale, tandis que celui de la composante v-Ets est relégué à l'extrémité carboxy-terminale de la protéine. Les domaines transactivateurs (TA), à raison d'un pour v-Myb, deux pour v-Ets, se retrouvent vers la partie médiane de la protéine (Figure 8).

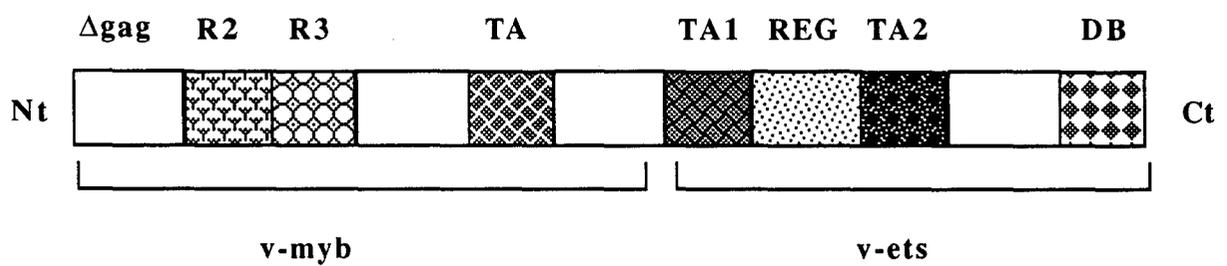


Figure 8 : Domaines fonctionnels présumés de la P135 Δ gag-myb-ets. R2 et R3 = 2 motifs répétés du domaine de liaison à l'ADN de v-Myb. TA = domaine transactivateur de v-Myb. TA1 et TA2 = domaines transactivateurs de v-Ets. DB = domaine de liaison à l'ADN de v-Ets. REG = domaine régulateur de l'activité transactivatrice.

L'analyse des mutants de Metz et Graf exprimant des protéines Myb et Ets séparément, mais recombinés *in vivo*, indique que les virus néoformés ont subi des délétions de taille variable, situées à la jonction Myb-Ets. En terme de domaines fonctionnels, il s'avère qu'une destruction des deux domaines TA codés par le gène *v-ets* n'abolit en rien la leucémogénicité. Le domaine TA de la composante v-Myb est alors capable à lui seul d'assurer les propriétés transactivatrices de la molécule entière (Metz et Graf, 1991b) (Figure 8). Réciproquement, la délétion du domaine TA codée par le gène *v-myb* est compensée par les domaines homologues codés par le gène *v-ets*, et la molécule résultante est toujours douée de propriétés transformantes, à la fois vis-à-vis des cellules érythroïdes et myéloïdes (Domenget *et al.*, 1992). Par contre, comme l'évoquaient déjà les effets des mutants thermosensibles, les domaines de liaison à l'ADN des parties v-Myb et de v-Ets sont absolument indispensables à la transformation de chaque lignée (Tableau 3).

3.5. Les gènes cibles hématopoiétiques de la P135 Δ gag-myb-ets

La protéine oncogène du rétrovirus E26 est capable d'altérer le déroulement normal de la différenciation hématopoiétique. Cette protéine est une chimère associant deux facteurs de transcription altérés. Ses propriétés transformantes se manifestent vraisemblablement par la régulation aberrante de certains gènes importants pour la prolifération et la différenciation hématopoiétique. Pour mieux comprendre les mécanismes d'action de la P135 Δ gag-myb-ets et par là même pour décrypter le cours normal de l'hématopoïèse, il est donc essentiel de caractériser ses gènes-cibles. Ce sont les études menées sur le virus mutant E26 ts 21 (Beug *et al.*, 1984) qui ont permis de décrire un gène-cible de la P135 Δ gag-myb-ets pour la première fois (Ness *et al.*, 1989). Ce gène a pu être isolé après hybridation différentielle de banques d'ADNc obtenues à partir de myéloblastes infectés par le mutant E26 ts 21, et cultivés soit à 42°C, soit à 37°C. Dénommé *mim-1* (Myb-induced myeloid protein 1), ce gène est exprimé massivement, sous forme d'un ARNm de 1 kb environ, dans les myéloblastes (à 37°C), mais cesse de l'être quand ceux-ci sont engagés dans la voie monocytaire (à 42°C). Le fait que *mim-1* ait été isolé de banques soustraites d'ADNc n'est pas fortuit, mais reflète bien au contraire l'expression de la P135 Δ gag-myb-ets dans les cellules myéloïdes. En effet, au cours d'expériences de cotransfection entreprises dans des macrophages HD11, l'expression de la P135 Δ gag-myb-ets s'accompagne d'une stimulation de l'activité du promoteur exogène du gène *mim-1*. Des résultats analogues sont obtenus avec les protéines c-Myb et v-Myb^{AMV}, suggérant que la régulation du promoteur du gène *mim-1* est imputable à la partie v-Myb de la P135 Δ gag-myb-ets (Ness *et al.*, 1989). Cependant, des monoblastes transformés par AMV n'expriment pas *mim-1*. Et, *in vitro*,

Protéine oncogène	Domaine déléte	Transformation <i>in vitro</i>		Références
		myéloïde	érythroïde	
v-Myb	DB	-	+	(Domenget <i>et al.</i> , 1992)
	TA	+	+	(Metz et Graf, 1991a)
v-Ets	DB	+	-	(Nunn et Hunter, 1989)
	TA	+	+	(Metz et Graf, 1991a)

Tableau 3 : Importance des différents domaines fonctionnels de la P135 Δ gag-myb-ets dans la transformation hématopoïétique. TA = domaine transactivateur. DB = domaine de liaison à l'ADN.

la protéine Ets-2 cellulaire concourt avec la protéine c-Myb à une induction maximale du promoteur du gène *mim-1* (Dudek *et al.*, 1992).

Le différentiel d'expression de *mim-1* constaté entre des monoblastes infectés par AMV, ou des myéloblastes infectés par le virus E26, pourrait suggérer que ce gène n'est pas un marqueur spécifique de la transformation cellulaire. Le gène *mim-1* représenterait plutôt un marqueur de la différenciation des myéloblastes vers la voie granulocytaire.

Récemment, un second gène cible de la P135 Δ gag-myb-ets a été découvert. Yuan *et al.* (1989) ont astucieusement construit un équivalent murin du rétrovirus E26, en insérant les déterminants *myb* et *ets* dans un vecteur dérivé du virus leucémogène d'Abelson (Abelson MuLV). Ce virus provoque des érythroleucémies chez le souriceau nouveau-né. Les cellules transformées sont des précurseurs très immatures et ne peuvent être établies en lignée qu'en présence d'érythropoïétine. Elles expriment prématurément le récepteur de l'érythropoïétine. Le facteur GATA-1, l'un des effecteurs de la régulation transcriptionnelle du gène de ce récepteur, constitue une cible de la P135 Δ gag-myb-ets (Aurigemma *et al.*, 1992 ; Ruscetti *et al.*, 1992), bien qu'une interaction directe entre la protéine virale et le promoteur du gène GATA-1 reste à démontrer. C'est toutefois probable puisqu'au moins la protéine Ets-1 cellulaire transactive ce promoteur (Seth *et al.*, 1993). Le promoteur du gène du récepteur de l'érythropoïétine n'est pas directement régulé par la P135 Δ gag-myb-ets, mais il semble que cette dernière accroisse l'activité transcriptionnelle de ce gène par interaction coopérative avec le facteur GATA-1. Ces résultats présentent l'intérêt de poser les bases d'un modèle de croissance autocrine des cellules érythroïdes transformées par le virus E26. Toutefois, la transposition de ce modèle au système aviaire n'est qu'hypothétique dans la mesure où les équivalents fonctionnels de l'érythropoïétine et de son récepteur ne sont pas identifiés chez le poulet.

4. Influence de l'infection par le virus E26 sur la croissance des fibroblastes :

Sans être totalement transformés, des fibroblastes infectés par le virus E26 voient certaines de leurs caractéristiques morphologiques et physiologiques modifiées (Jurdic *et al.*, 1987). En particulier, leur taux de croissance est plus élevé et leur durée de vie plus longue que ceux de fibroblastes normaux cultivés en milieu classique. Ils peuvent également croître en milieu pauvre en sérum et former de petites colonies sur milieu semi-solide. Des délétions des différents domaines de la P135 Δ gag-myb-ets montrent que les deux oncogènes coopèrent à l'acquisition des nouvelles performances prolifératives des fibroblastes (Tableau 4) (Ravel-Chapuis *et al.*, 1991). Les domaines TA et DB codés par le gène *v-myb* contribuent différemment à la croissance en milieu appauvri en sérum, et à la formation de colonies en milieu semi-solide. Un mutant dans la partie v-Ets auquel manque le domaine TAI provoque des effets mitogéniques

Protéine oncogène	Domaine délété	Croissance 0,1 % sérum	Petites colonies
myb-ets	aucun	+	+
Myb	TA	+/-	-
	DB	+/-	+
Ets	TAII	++	++
	REG	+	+
	DB	+	+
	TAII+REG	-	-
	DB+TAII	++	++
	DB+TAII+REG	-	-

Tableau 4 : Influence des domaines présumés de la P135 Δ gag-myb-ets sur la croissance des fibroblastes

accrus par rapport au virus E26 sauvage. Le domaine TAI_{II} a donc un effet négatif sur la croissance des fibroblastes. De plus, il semble qu'une coopération des différents domaines des déterminants v-Ets soit nécessaire (voir mutants DB-TA-REG⁻). Enfin, le domaine DB de la partie v-Ets ne semble pas impliqué, ce qui pourrait indiquer que c'est plutôt par des interactions protéine-protéine (soit v-Myb/v-Ets, soit v-Ets/autre facteur de transcription) et non par une reconnaissance spécifique de l'ADN que la composante v-Ets exerce ses effets sur les fibroblastes.

L'infection de fibroblastes par le virus E26 engage donc ces cellules vers les premières étapes de la transformation, sans toutefois la parachever complètement. Il est possible que pour transformer les cellules hématopoïétiques, la P135 Δ gag-myb-ets interagisse avec des cofacteurs absents des fibroblastes. Les voies menant à la transformation sont probablement différentes dans deux types cellulaires distincts, interrogation qui met sans doute en péril la notion de gènes clés de la transformation. Compte tenu du très faible niveau d'expression de l'ARNm de la P135 Δ gag-myb-ets (vingt fois moins que dans les myéloblastes infectés par E26), il est possible qu'une surexpression soit nécessaire pour observer des effets analogues à ceux que le virus E26 inflige aux cellules hématopoïétiques (Jurdic *et al.*, 1987). Il est intéressant de noter que la P135 Δ gag-myb-ets est indétectable par immunoprécipitation de lysat de fibroblastes infectés. Peut-être est-ce parce qu'elle est produite en trop faible quantité voire pas du tout. Dans ce cas, on peut se risquer à penser que l'activité transformante du virus E26 pourrait aussi s'exercer au niveau de son ARNm, peut-être par hybridation antisens de l'extrémité 3' inversée de la partie v-ets avec des ARNm voisins ou analogues à *ets-1* dont les produits assureraient la quiescence des fibroblastes.

5. Effet mitogène du virus E26 sur les cellules neurorétiniennes :

Parallèlement aux études menées sur le système hématopoïétique, le développement d'autres systèmes cellulaires sensibles à l'infection par le virus E26 peut permettre de déchiffrer le fonctionnement de v-Myb et de v-Ets en réponse à différents facteurs de croissance et dans un contexte plus maniable. Ainsi, E26 accroît également le potentiel prolifératif des cellules de neurorétine aviaires comme il le fait sur les fibroblastes (Amouyel *et al.* 1989). Dans ces cellules, le transcrit viral de 5.7 kb est fortement exprimé ce qui infirme l'hypothèse selon laquelle l'intensité du pouvoir transformant du virus serait directement corrélée à son taux d'expression (Jurdic *et al.*, 1987).

Ce système apporte des éléments sur le rôle de v-Myb essentiellement puisque l'intégrité de v-Ets n'est pas requise pour la croissance des cellules neurorétiniennes en milieu minimum, bien que le virus E26 ait un effet plus prononcé sur ces cellules que le virus AMV. De plus, c'est v-Myb qui transmet au niveau nucléaire une réponse

proliférative accrue provoquée par l'ajout de bFGF, et son domaine transactivateur semble indispensable à cet effet (Garrido *et al.*, 1992).

Outre sa création même, qui permet de formuler un certain nombre d'hypothèses sur le plan évolutif, le rétrovirus E26 est un modèle qui a permis des avancées dans différents domaines de haut intérêt : en figeant des cellules hématopoiétiques à des stades précis de leur développement, ce virus permet la description de précurseurs cellulaires dont l'observation serait naturellement inaccessible. Grâce à ce type de virus, nombre de laboratoires espèrent décrypter les voies majeures qui aboutissent à la transformation cellulaire. Les mutants de délétion et les mutants ponctuels de la P135 Δ gag-myb-ets devraient être informatifs à cet égard, car il pourrait être possible en effet de décrire leur mode d'interaction avec l'ADN et, en particulier, de préciser leur(s) séquence(s) cible(s). Les relations entre les composantes protéiques myb et ets restent encore mystérieuses, et bien que leurs différents domaines semblent montrer une certaine redondance, on ignore pourtant si elles exercent une stimulation ou une répression synergique sur leurs gènes cibles, même si elles coopèrent à la transformation des cellules. L'isolement d'un gène cible de la P135 Δ gag-myb-ets répondra enfin à toutes les attentes, car sa régulation devrait permettre de concilier le mode d'action de la protéine au niveau moléculaire et les propriétés biologiques du virus E26. Elle procurera également des comparaisons intéressantes avec le fonctionnement normal des oncogènes que le virus a détournés.

		Hélice α		Charnière			Domaine basique				
		10	20	30	40	50	60	70	80	90	100
HETS1	GSG---PI	QLWQFLLELLTDKS	--CQSFISWTG-DGWEFKLS--DPDEVA	RRWGKPK-NKPKMNYEKL	SRGLRYYYDKNI	IHKTA-GKRYVYRFVCDI	QSLLG	93			
CETS1	GSG---PI	QLWQFLLELLTDKS	--CQSFISWTG-DGWEFKLS--DPDEVA	RRWGKPK-NKPKMNYEKL	SRGLRYYYDKNI	IHKTA-GKRYVYRFVCDI	QSLLG	93			
VETS	GSG---PI	QLWQFLLELLTDKS	--CQSFISWTG-DGWEFKLS--DPDEVA	RRWGKPK-NKPKMNYEKL	SRGLRYYYDKNI	IHKTA-GKRYVYRFVCDI	QSLLG	93			
XETS1A	GSG---PI	QLWQFLLELLTDKS	--CQSFISWTG-DGWEFKLS--DPDEVA	RRWGKPK-NKPKMNYEKL	SRGLRYYYDKNI	IHKTA-GKRYVYRFVCDI	QSLLG	93			
XETS1B	GSG---PI	QLWQFLLELLTDKS	--CQSFISWTG-DGWEFKLS--DPDEVA	RRWGKPK-NKPKMNYEKL	SRGLRYYYDKNI	IHKTA-GKRYVYRFVCDI	QSLLG	93			
HETS2	GSG---PI	QLWQFLLELLSDKS	--CQSFISWTG-DGWEFKLA--DPDEVA	RRWGKPK-NKPKMNYEKL	SRGLRYYYDKNI	IHKTS-GKRYVYRFVCDI	QNLLG	93			
CETS2	GSG---PI	QLWQFLLELLTDKS	--CQSFISWTG-DGWEFKLA--DPDEVA	RRWGRRK-NKPKMNYEKL	SRGLRYYYDKNI	IHKTS-GKRYVYRFVCDI	QNLLG	93			
XETS2A	GSG---PI	QLWQFLLELLTDKS	--CQSFISWTG-DGWEFKLT--DPDEVA	RRWGKPK-NKPKMNYEKL	SRGLRYYYDKNI	IHKTS-GKRYVYRFVCDI	HNLLG	93			
XETS2B	GSG---PI	QLWQFLLELLTDKS	--CQSFISWTG-DGWEFKLA--DPDEVA	RRWGKPK-NKPKMNYEKL	SRGLRYYYDKNI	IHKTS-GKRYVYRFVCDI	HNLLG	93			
SUETS2	GSG---PI	QLWQFLLELLTDKT	--CQHIISWTG-DGWEFKLS--DPDEVA	RRWGKPK-NKPKMNYEKL	SRGLRYYYDKNI	IHKTA-GKRYVYRFVCDI	QSLLG	93			
DETS2	GSG---PI	QLWQFLLELLTDKT	--CQSFISWTG-DGWEFKLT--DPDEVA	RRWGIRK-NKPKMNYEKL	SRGLRYYYDKNI	IHKTA-GKRYVYRFVCDI	QNLVG	93			
ERG	GSG---QI	QLWQFLLELLSDSS	--NSSCITWEG-TNGEFKMT--DPDEVA	RRWGERK-SKPNMNYDKL	SRALRYYYDKNIMTKVII	GKRYAYKF--DF	HGIAQ	91			
FLI	GSG---QI	QLWQFLLELLSDSA	--NASCITWEG-TNGEFKMT--DPDEVA	RRWGERK-SKPNMNYDKL	SRALRYYYDKNIMTKVII	GKRYAYKF--DF	HGIAQ	91			
SUERG	?SG---QI	QLWQFLLELLSDSS	--NANCITWEG-TNGEFKMT--DPDEVA	RRWGERK-SKPNMNYDKL	SRALRYYYDKNIMTKVII	GKRYAYKF--DF	AGLAQ	90			
DETS3	GSG---QI	QLWQFLLELLSDSN	--NASCITWEG-TNGEFKLT--DPDEVA	RRWGERK-SKPNMNYDKL	SRALR????????????????????	?????	64				
DETS4	GG---SHI	HLWQFLKELLASPQ	VNGTA-IRWIDRSKGIFKIE--DSVRVA	KLWGRRK-NRPAMNYDKL	SRSIRQYYKKGIMKKTTERS	QRLVYQF-CHF	---YS	92			
DETS6	GSG---GQI	QLWQFLLELLADSS	--NANAISWEG-QSGEFRLI--DPDEVA	RRWGERK-AKPNMNYDKL	SRALRYYYDKNIMTKVII	GKRYAYKF--DF	HGLMA	92			
GABP	G--NNGQI	QLWQFLLELLTDKD	--ARDCISWVG-DEGEFKLN--QPELVA	QKWGQRK-NKPTMNYEKL	SRALRYYDGDMI	CKVQ-GKRFVYKFVCDI	KTLIG	94			
DELG	GSGNNGQV	QLWQFLLEYLTDCE	--HTDVIEWVG-TEGEFKLT--DPDRVA	RLWGEKK-NKPMNYEKL	SRALRYYDGDMI	SKVS-GKRFAYKFDCDI	KLLIG	96			
POK	NTN--GRL	-LWDFLQQLLNDRN	QKYSDLIAWKCRDTGVFKIV--DPAGLA	KLWGIQK-NHLSMNYDKM	SRALRYYYRVNI	LRKVQ-GERHCYQFLRNI	TELKN	96			
SAP1	---MDSA1	TLWQFLQLLQKQPQ	--NKHMICWTSMDG-QFKLL--QAEEVA	RLWGIK-NKPNMNYDKL	SRALRYYYVKNII	KKVN-GQKFVYKFVS-Y	PEILN	92			
EIK	---MDPSV	TLWQFLQLLREQG	--NGHIIISWTSRDGGFKLV--DAEEVA	RLWGLRK-NKTNMNYDKL	SRALRYYYDKNII	IRKVS-GQKFVYKFVS-Y	PEVAG	93			
PEA3	PYQRRGAL	QLWQFLVALLDDPT	--NAHFIAWTG-RGMEFKLI--EPEEVA	RLWGIQK-NRPAMNYDKL	SRSRLRYYYEKGIMQKVA	GERYVYKIVCEP	EALFS	93			
ELF1	K-G--NTI	YLWFLIALLLQDKA	T-CPKYIKWTNREKGVFKLV--DSKAVS	RLWGLRK-NKPNMNYE	TMGRALRYYYQ	RGI LAKVE-GGRLVYQF	KEMF	95			
E74	REG--STT	YLWFLKLLQDRE	Y-CPRFIKWTNREKGVFKLV--DSKAVS	RLWGLRK-NKPNMNYE	TMGRALRYYYQ	RGI LAKVD-GQRLVYQF	VDFV	96			
SP11	GSKKK--I	RLYQFLDLRLSGD	MK--DSIWWVDKDKGTFQFSSKIIKEALA	IRWGIQKGNRKKMTYQK	MARALRNYGKT	GEVKKVKK--KLTYQF	SGEV	95			
SP12	GTRKK--L	RLYQFLGLLTRGD	MR--ECVWVVEPGAGVQFSSKIIKELIA	RRWQOKGNRKRMTYQK	LARALRNYAKT	GEIRKVKR--KLTYQF	DSAI	95			
CONS		L* *FL *L	* W * F*	* W * K * M*Y * R *R*Y*	* K * Y F						

Figure 9 : Alignement du domaine ETS de la plupart des membres de la famille Ets. Les amino-acides invariants sont indiqués à la ligne CONS. Les astérisques désignent les résidus conservés dans trois protéines sur quatre.

3ème partie : La famille Ets, un ensemble multigénique de facteurs de transcription. Aspects moléculaires.

1. Présentation de la famille Ets :

1.1. Le domaine ETS :

Obtenues à partir du rétrovirus E26, des sondes *v-ets* ont permis l'isolement et la caractérisation chez le poulet d'un gène cellulaire homologue, baptisé *c-ets-1* (Leprince *et al.*, 1983). Ce gène code deux protéines et leurs isoformes dont les propriétés seront détaillées plus loin. Il est apparu comme le membre fondateur d'une famille pléthorique de gènes apparentés, isolés dans des espèces aussi distinctes sur le plan évolutif que l'homme ou la drosophile. L'appartenance à cette famille est dictée par la présence d'un domaine protéique d'environ 85 acides aminés, le domaine ETS, plus ou moins divergé par rapport à celui des protéines *c-Ets-1* (Karim *et al.*, 1990). L'alignement des séquences de 27 de ces domaines (Figure 9) indique que 18 % des acides aminés sont parfaitement conservés tandis que 14 autres % sont retrouvés dans la plupart des cas (Lautenberger *et al.*, 1992 ; Laudet *et al.*, 1993). Ces zones conservées se répartissent aux extrémités amino- et carboxy-terminales du domaine, de part et d'autre d'une région charnière de séquence variable, dans laquelle deux acides aminés seulement sont préservés entre toutes les protéines Ets.

1.2. Localisation chromosomique :

La localisation chromosomique de la plupart des gènes de la famille Ets est connue, avec précision pour certains, comme l'indique le Tableau 5. Chez l'homme, les gènes *ets-1* et *ets-2* sont localisés sur le même chromosome que les gènes *fli-1* et *erg-1* respectivement : *ets-1* et *fli-1* sont disposés en tandem en position 11q23, tandis qu'*ets-2* et *erg-1* sont organisés de même en 21q22 . De plus, il apparaît que les gènes *ets-1*, *ets-2* et *fli-1* sont localisés sur des chromosomes synténiques d'autres espèces comme le chat ou la souris (la localisation chromosomique du gène *erg-1* n'est déterminée que chez l'homme). Le gène *spi-1/PU.1* est lui aussi localisé sur le chromosome 11 chez l'homme, mais sur le chromosome 2 de souris. Bien que ces chromosomes ne soient pas synténiques, l'analyse de liaison montre que le gène *spi-1* se situe au sein d'un ensemble multigénique qui s'étend dans une région équivalente sur ces deux chromosomes, et qui comprend notamment le gène de la phosphatase acide lysosomiale.

Gène	Espèce	Chromosome	Références
<i>ets-1</i>	homme	11q23-24	De Taisne <i>et al.</i> , 1984
<i>ets-1</i>	souris	9	Watson <i>et al.</i> , 1986
<i>ets-1</i>	chat	D1	Watson <i>et al.</i> , 1986
<i>ets-2</i>	homme	21q22.1-22.3	Watson <i>et al.</i> , 1985 et 1986
<i>ets-2</i>	souris	16	Watson <i>et al.</i> , 1986
<i>ets-2</i>	chat	C2	Watson <i>et al.</i> , 1986
<i>fli-1</i>	homme	11q23-24	Baud <i>et al.</i> , 1991
<i>fli-1</i>	souris	9	Ben David <i>et al.</i> , 1990
<i>spi-B</i>	homme	19q13.3-19q13-4	Ray <i>et al.</i> , 1992
<i>sp-1</i>	homme	11p11.22	Nguyen <i>et al.</i> , 1990
<i>spi-1</i>	souris	2E3	Moreau-Gachelin <i>et al.</i> , 1989 Paul <i>et al.</i> , 1991
<i>elk-1</i>	homme	Xp11.2	Rao <i>et al.</i> , 1989
<i>elk-2</i>	homme	14q32.3	Rao <i>et al.</i> , 1989
<i>erg</i>	homme	21q22.3	Rao <i>et al.</i> , 1987 et 1988

Tableau 5 : Localisation chromosomique de plusieurs gènes de la famille Ets

1.3. Histoire évolutive de la famille Ets :

L'analyse des taux de divergence entre les différents domaines ETS (Lautenberger *et al.*, 1992 ; Laudet *et al.*, 1993) indique que la famille Ets constitue très vraisemblablement un groupe très ancien de gènes. Les gènes *E74*, *elg*, *D-ets-2* et *D-ets-3* sont des gènes de drosophile dont les équivalents sont supposés connus chez les Vertébrés : il s'agit respectivement des gènes *elf-1*, *GABP α* , *c-ets-1/c-ets-2*, *erg-1/fli-1* (Figure 10). Il est donc probable que la diversité observée actuellement au sein des gènes *ets* était déjà apparente avant la divergence Arthropodes/Vertébrés, c'est-à-dire il y a plus de 500 millions d'années. Puisque la séquence des gènes *D-ets-4*, *D-ets-6* et *pokkuri* n'est proche d'aucun groupe de gènes, il est permis de supposer que leurs équivalents restent à découvrir chez les Vertébrés.

Cependant, chez la drosophile, un certain nombre d'évènements de duplication ne se sont pas produits : les gènes *c-ets-1* et *c-ets-2* d'une part, *erg* et *fli-1* d'autre part, présentent des séquences nucléotidiques extrêmement proches particulièrement régionalisées en 5' et en 3' des parties codantes (se reporter à la figure 13). L'analyse des homologies de séquences indique qu'il est possible que les gènes *D-ets-2* et *D-ets-3* en soient respectivement les ancêtres. Si tel est le cas, on n'attend donc qu'un seul gène de type *ets* et un seul gène de type *erg* chez les Arthropodes, ce qui doit encore être vérifié.

La disposition en tandem sur des chromosomes différents des gènes *ets-1/fli-1* et *ets-2/erg* évoque la probabilité d'une duplication d'un gène ancestral *ets/erg* sur le même chromosome. Cet évènement se serait produit avant la divergence Arthropodes/Vertébrés, puisque les gènes de type *ets* et *erg* sont séparés chez la drosophile. La duplication du tandem sur un chromosome différent s'est produite, quant à elle, avant l'apparition des Amphibiens (il y a 400 millions d'années), puisque le xénope possède déjà des exemplaires de gènes *c-ets-1* et *c-ets-2* séparés (Figure 11).

2. Les protéines Ets constituent une superfamille de facteurs de transcription :

C'est aux protéines c-Ets-1 aviaires, prototypes de la famille Ets, puis murines et humaines qu'a initialement été assigné un rôle de facteurs de transcription parce qu'elles régulent l'expression de certains gènes après reconnaissance et fixation d'un motif nucléotidique particulier (Bosselut *et al.*, 1990 ; Gunther *et al.*, 1990 ; Ho *et al.*, 1990 ; Wasylyk *et al.*, 1990). La plupart des autres protéines de la famille Ets ont été découvertes sur la base d'une homologie importante avec les protéines c-Ets-1, qui recouvre le domaine ETS dans tous les cas. S'il s'avère, comme on peut le supposer, que toutes les protéines Ets connues à ce jour sont douées d'une activité transcriptionnelle se posent immédiatement un certain nombre de questions ayant trait à leur spécificité de

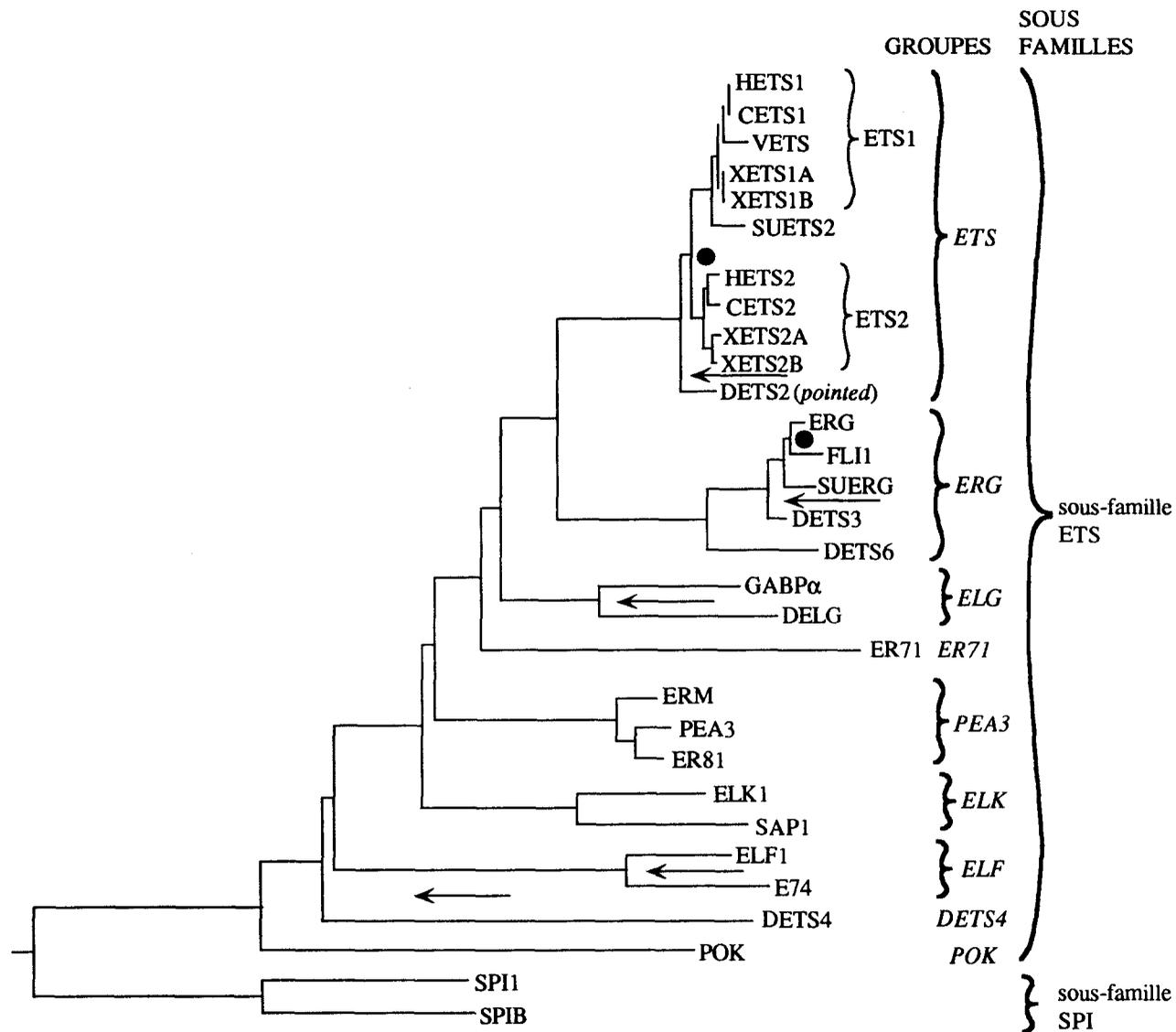


Figure 10 : Arbre phylogénétique de la famille de gènes Ets. Les flèches indiquent les équivalences possibles entre gènes de drosophile et gènes de Vertébrés. ● = duplication des gènes ancestraux de type *ets* et *erg* (tiré de Laudet *et al.*, 1993).

reconnaissance vis-à-vis des gènes qu'elles sont individuellement censées réguler. La définition de cette spécificité implique la connaissance de certains de leurs gènes cibles, et leur modes de régulation tant transcriptionnelle que post-traductionnelle. Et tout d'abord, elle nécessite un premier aperçu des interactions que ces protéines contractent avec l'ADN. Bien que cette gigantesque entreprise n'en soit qu'à son prélude, certains résultats laissent entrevoir les bases de la spécificité d'action des protéines Ets.

2.1. Reconnaissance de l'ADN :

2.1.1. Reconnaissance macromoléculaire :

Des expériences d'interférence de méthylation ont démontré que la protéine p63^{c-ets-1} murine contacte à la fois le grand sillon et les petits sillons adjacents de la molécule d'ADN sur une vingtaine de bases environ (Nye *et al.*, 1992). Comparé aux interactions avec l'ADN d'autres familles de facteurs de transcription connus, le mode de fixation de la protéine p63^{c-ets-1} semble unique. Si, comme cela est hautement probable, il est possible d'extrapoler ces données aux autres protéines Ets, elles confirment l'émergence d'une famille de facteurs de transcription totalement différente de celles qui étaient connues auparavant.

2.1.2. Définition de la cible nucléotidique :

Il est peu probable qu'existe une corrélation systématique qui permette d'attribuer à chaque membre de la famille Ets un site de reconnaissance privilégié. Les premiers résultats définissant une cible nucléotidique des protéines Ets-1 aviaires ou murine indiquaient la reconnaissance d'une séquence centrée autour d'un motif GGAA/T retrouvé au sein de certains enhanceurs viraux (Bosselut *et al.*, 1990 ; Gunther *et al.*, 1990 ; Wasylyk *et al.*, 1990). Certes, il est vrai que des études comparant la spécificité de liaison à l'ADN de protéines de la famille Ets indiquent que, dans certains cas, leurs séquences cibles ne sont pas interchangeables : la comparaison des propriétés de liaison à des séquences cibles naturelles des protéines Ets-1 et Elf-1 (Tableau 6) illustre l'importance que peut revêtir la nature des nucléotides entourant le motif GGA (Wang *et al.*, 1992), et indique que, comme son ancêtre présumé la protéine E74, Elf-1 ne reconnaît que le noyau GGAA. La protéine Elk-1 reconnaît préférentiellement un motif de type E74 elle aussi ; donc, chez les Vertébrés, les protéines Elf-1 et Elk-1 pourraient constituer un sous-groupe de reconnaissance. Une autre étude a comparé les séquences de fixation de la protéine Ets-1 et du produit du gène *Spi-1*, la protéine PU.1 (Wasylyk *et al.*, 1992). Elle suggère que la liaison à l'ADN de la protéine Ets-1 est moins exigeante que celle de PU.1. Enfin, Ascione *et al.* (1992) ont décrit des

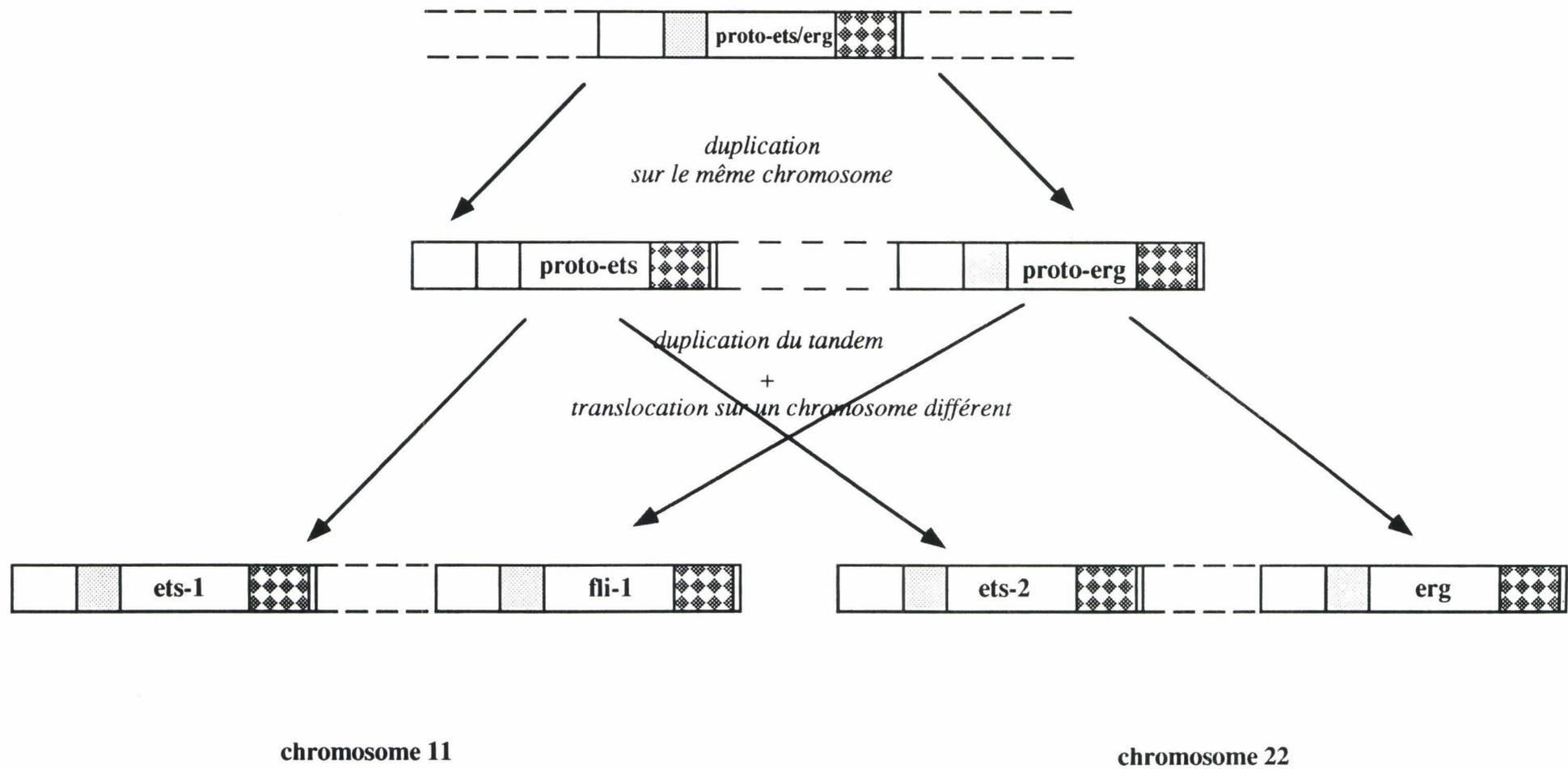


Figure 11 : Evolution des sous-groupes de gènes *ets* et *erg*

différences d'affinité des protéines Ets-1 et Ets-2 selon les promoteurs reconnus, en fonction des nucléotides proximaux.

Sur la base de la reconnaissance du motif GGAA/T, et pour définir un consensus idéal de fixation de chacun des membres de la famille de protéines Ets, un deuxième type d'approche a été développé. Il consiste à sélectionner les sites consensus d'une protéine donnée, à partir d'un ensemble d'oligonucléotides de séquence aléatoire et marqués (technique appelée SAAB pour Selected And Amplified Binding). Les complexes protéine-ADN sont séparés par retard sur gel et les oligonucléotides reconnus sont amplifiés après plusieurs cycles de PCR puis sont séquencés. Ce type de technique a permis de définir *in vitro* des consensus pour la protéine Ets-1 aviaire (Woods *et al.*, 1992), murine (Nye *et al.*, 1992) et humaine (Fisher *et al.*, 1991), pour la protéine E74 de drosophile (Urness et Thummel, 1990), pour les protéines murines GABP α , ER71, ER81 (Brown et McKnight, 1992) et Erg humaine (Murakami *et al.*, 1993). L'examen de leur séquence respective (Tableau 7) indique que dans tous les cas la séquence comporte le motif GGA, à l'exception du site de fixation de la protéine E74 qui, en plus, présente obligatoirement une adénine en position +3. De fait, l'ensemble des séquences sélectionnées est très homogène : le facteur GABP α devrait reconnaître sans peine le motif nucléotidique sélectionné par la protéine Ets-1 par exemple. On peut donc supposer que la fixation des différentes protéines Ets à un motif de type GGAA/T sera nécessaire à leur action sans toutefois constituer une caractéristique discriminante. Il convient donc de chercher ailleurs les bases de la spécificité d'action de chaque protéine Ets sur ses gènes cibles : ainsi, une distribution tissulaire propre à chaque membre ou à certaines combinaisons d'entre eux. L'importance de leurs partenaires transcriptionnels sera développée plus loin. Enfin, la possibilité d'une redondance de la fonctionnalité de certaines protéines Ets, mais pas de toutes, est aussi à envisager.

2.2. Interactions avec l'ADN vues du côté protéique :

2.2.1. Le domaine de liaison à l'ADN :

Ce sont des travaux entrepris avec de l'ADN-cellulose qui ont initialement permis de localiser le domaine de fixation à l'ADN à l'extrémité carboxy-terminale de la protéine c-Ets-1 aviaire (Boulukos *et al.*, 1989) ou humaine (Reddy et Rao, 1990). Ultérieurement, la construction de mutants de délétions a permis de définir le motif protéique minimal capable de reconnaître une cible nucléotidique spécifique de type GGA. Dans tous les cas étudiés, qu'il soit situé en position carboxy-terminale, amino-terminale (Elk-1, SAP-1) ou médiane (Elf-1), il recouvre entièrement le domaine ETS (Tableau 8).

L'examen de la structure primaire du domaine ETS ne permet pas de prévoir l'existence d'une conformation analogue à celle d'autres domaines de liaison à l'ADN

		Ets-1	Elf-1
T α 2	AGAGGATGTG	+	-
	 A	+	+
MSV-LTR	AGCGGAAGCG	+	+
	 T	+	-
	 AA	-	+

Tableau 6 : Comparaison de la liaison à l'ADN des protéines Ets-1 et Elf-1 sur des éléments de réponse issus de l'enhancer du gène de la chaîne α du TCR et du LTR du virus MSV.

portés par des facteurs de transcription différents, tels les domaines à hélice-boucle-hélice (Murre *et al.*, 1989), basique à "Leucine zipper" (Landschultz *et al.*, 1988), à doigts de Zinc (Evans et Hollenberg, 1988), ou à homéodomaine (Levine et Hoey, 1988), pour ne citer que ceux-là. Cependant, l'analyse de la partie conservée amino-terminale du domaine ETS (se reporter à la Figure 9) évoque la possibilité de former une hélice α (Wang *et al.*, 1992). La richesse en acides aminés de nature basique (environ 30 %) de la partie conservée carboxy-terminale pourrait constituer une caractéristique que partagent les protéines Ets avec d'autres facteurs de transcription comme c-Myc ou c-Fos. L'analyse informatique voit néanmoins une structure de type hélice-tour-hélice se former entre les acides aminés 371-417 de la protéine c-Ets-1 humaine (Reddy et Rao, 1990).

La comparaison des séquences de toutes les protéines Ets montre également la présence quasi-unanime de trois Tryptophanes espacés de 17 ou 18 acides aminés, répartis dans la région amino-terminale, dans la région charnière et dans la région basique du domaine. Cette particularité représente aux yeux de certains auteurs une analogie avec le domaine de liaison à l'ADN de la protéine c-Myb (Anton et Frampton, 1988 ; Karim *et al.*, 1990). Toutefois, celui-ci comporte 3 triplets de Tryptophanes dans chacun de ses trois motifs répétés R1, R2, R3, ce qui constitue une situation tant soit peu différente. De plus, la protéine PU.1, bien que possédant une Tyrosine à la place du Tryptophane le plus amino-terminal, présente des propriétés de liaison à l'ADN analogues à celles des autres membres de la famille. Enfin, le remplacement du Tryptophane médian par un acide aminé hydrophobe ne perturbe en rien la capacité de fixation à l'ADN de la protéine c-Ets-1 (Wang *et al.*, 1992) malgré les contraintes structurales qu'il est censé engendrer.

A nouveau, il apparaît donc clairement que le domaine ETS est un domaine de reconnaissance et de fixation à l'ADN d'un type inédit. C'est avec une grande impatience qu'en est attendue la structure tridimensionnelle qui permettra de visualiser les acides aminés impliqués dans cette liaison.

2.2.2. Inhibition intramoléculaire de la liaison à l'ADN :

Les propriétés de liaison à l'ADN des protéines Ets semblent en fait plus complexes qu'il n'apparaît au premier abord, du moins dans le cas de certaines d'entre elles. En effet, les protéines Ets-1 et Ets-2 de pleine longueur semblent reconnaître l'ADN beaucoup moins efficacement que leur version respective amputée d'une large région amino-terminale (Reddy et Rao, 1990 ; Hagman *et al.*, 1992 ; Lim *et al.*, 1992 ; Wasylyk *et al.*, 1992) (Figure 12A et B). Cette inhibition est intramoléculaire puisqu'elle peut être levée par délétion d'une région présentant 55 % d'homologie entre les protéines Ets-1 (acides aminés 207 à 280) et Ets-2 (acides aminés 282 à 368) aviaires, située immédiatement en amino-terminal du domaine ETS. La protéine Elk-1

Espèce		Séquence de fixation optimisée							Références				
		-3	-2	-1	0	+1	+2	+3	+4	+5	+6		
Ets-1	poulet	A	C	A/T	G	G	A	A/T	A/G	T	T	Woods <i>et al.</i> , 1992	
	souris	G/A	C	C/A	G	G	A	A/T	A/G	C/T	N	Nye <i>et al.</i> , 1992	
	homme	G/A	G/C	C/A	G	G	A	A	G	T	C/T	Fisher <i>et al.</i> , 1991 (GGAA fixé)	
E74	drosophile	C/T	C	A	G	G	A	A	G	T		Urness et Thummel, 1990	
GABP α	rat	G/A	C/G	C/A	G	G	A	A/T	G/A	T/C			
ER71	rat	G/A	C/G	C/A	G	G	A	A/T	G/A	G/A	T/C	C	Brown et McKnight, 1992
ER81	rat	G/C	C/G	C/A	G	G	A	A/T	G/A	T/C	N		
Erg	homme	C/G	C/A	G	G	A	A	G/A	T			Murakami <i>et al.</i> , 1993 (GGAA fixé)	

Tableau 7 : Séquences de fixation optimisées de quelques protéines de la famille Ets déterminées par la technique de sélection d'oligonucléotides aléatoires. Dans le cas des protéines Ets-1 et Erg humaines, un motif GGAA est imposé au sein de bases adjacentes variables. Les protéines dont le site de fixation a été déterminé par d'autres méthodes ne sont pas mentionnées sur cette figure.

présente elle aussi un domaine qui l'empêche de se fixer de façon autonome au motif SRE du promoteur du gène *c-fos*, bien qu'elle soit par contre capable de lier seule la séquence de reconnaissance de la protéine E74 de drosophile (Rao et Reddy, 1992). De la même façon, la troncature d'une région carboxy-terminale par rapport au domaine ETS permet à la protéine SAP-1 de se lier au motif SRE en absence de SRF (Dalton et Treisman, 1992). La protéine PU.1, dont le domaine ETS diverge beaucoup par rapport à celui des autres protéines de la famille, ne présente pas de domaine inhibiteur de la liaison à l'ADN (Wasylyk *et al.*, 1992).

Ces observations présentent un aspect tout-à-fait intéressant puisqu'elles proposent un modèle d'activation oncogénique des protéines c-Ets-1. En effet, il semble que la composante v-Ets de la P135 Δ gag-myb-ets du virus E26 ne soit pas sujette à cette inhibition intramoléculaire (Lim *et al.*, 1992) (Figure 12C). Or, cette protéine est colinéaire à la p68^{c-ets-1}, à deux mutations ponctuelles près et surtout à l'exception de ses 16 acides aminés carboxy-terminaux (se reporter à la 1ère partie de ce mémoire). Il semble donc que l'extrémité carboxy-terminale spécifique des protéines Ets-1 ou Ets-2 (76 % d'homologie) ait aussi un rôle à jouer dans l'inhibition de fixation de l'ADN. Cette inhibition pourrait être levée de façon régulée, par l'interaction de la protéine Ets avec des cofacteurs qui "débobineraient" la molécule (Figure 12B). La protéine v-Ets se comporterait donc comme un facteur de transcription constitutivement activé puisqu'incapable d'être régulé par répression intramoléculaire.

Il est urgent néanmoins de modérer ce séduisant modèle par d'autres observations rapportées de différents laboratoires. Ainsi, la recherche de protéines fixant le LTR du virus MSV (murine sarcoma virus) par criblage d'une banque d'ADNc de souris a désigné la protéine c-Ets-1 comme candidat. Cette protéine est certes tronquée du côté amino-terminal mais elle couvre néanmoins largement la région inhibitrice (Gunther *et al.*, 1990). De même, les premières expériences visant à établir le rôle de facteur de transcription des protéines c-Ets-1 et c-Ets-2 aviaires montraient de façon parfaitement claire une liaison de ces protéines au site riche en bases puriques de l'enhancer précoce du virus du polyome (Wasylyk *et al.*, 1990). Ces résultats ont également été reproduits au laboratoire, à de multiples reprises (Leprince *et al.*, 1992 ; Duterque-Coquillaud *et al.*, 1993). Ces auteurs constatent que l'extrémité carboxy-terminale des protéines Ets-1, voire Ets-2 (Wasylyk *et al.*, 1992), est nécessaire à la fixation, tandis que la région analogue de la protéine v-Ets est inhibitrice. Leur hypothèse pour expliquer l'oncogénicité de v-Ets reposerait plutôt sur un défaut de reconnaissance des gènes cibles et donc une perte de fonction.

A l'évidence, il faut imputer ces divergences de données à une variabilité expérimentale de chaque laboratoire. Si la clé de ce dilemme est d'ordre technique, comme l'on peut le supposer, elle n'en est pas moins intéressante. En effet, si des modifications chimiques des milieux réactionnels (salinité, potentiel d'oxydo-

	Domaine minimal de liaison à l'ADN	Production	Sonde	Références
Ets-1 (murin)	94 AA Ct (336-440)	bactéries	motif Ets-1 optimisé par SAAB	Nye <i>et al.</i> , 1992
GABPα (murin)	96 AA Ct (313-407)	bactéries	promoteur du gène ICP4 du virus HSV1	Thompson <i>et al.</i> , 1991
Elk-1 (humain)	89 AA Nt (1-89)	bactéries	motif E74	Janknecht et Nordheim, 1992 Rao et Reddy, 1992
Ets-2 (aviaire)	115 AA Ct (364-479)	lysat réticulocytes	motif PEA3 de l'enhancer du virus Polyome	Wasylyk <i>et al.</i> , 1992
PU.1 (murin)	106 AA Ct (160-266)	lysat réticulocytes	motif PU.1	Wasylyk <i>et al.</i> , 1992

Tableau 8 : Définition du domaine minimal de liaison à l'ADN de cinq protéines de la famille Ets

réduction...) (Wasylyk et Wasylyk, 1993) sont responsables des divergences de résultats concernant la fixation à l'ADN de la protéine Ets-1, il est probable qu'elles reflètent en réalité des modifications équivalentes opérant *in vivo* pour réguler cette activité.

2.3. Domaines transactivateurs :

Certains domaines transactivateurs ne présentent, on l'a vu dans la première partie, aucune caractéristique structurale particulière (Tasset *et al.*, 1990 ; revue de Hahn, 1993). C'est à cette catégorie qu'appartiennent semble-t-il les domaines transactivateurs des protéines Ets. A l'exception des protéines ERM (Monté *et al.*, 1993), ER81 et ER71 (Brown et McKnight, 1992) pour lesquelles les expériences sont sans doute en cours, toutes les protéines Ets de Mammifères semblent douées de propriétés transactivatrices, que ce soit vis-à-vis d'un élément de réponse artificiel ou d'un promoteur apparemment naturel. Les domaines transactivateurs ont été circonscrits avec précision dans la plupart des cas, grâce à la comparaison des activités activatrices de différents mutants de délétion.

La protéine p68^{c-ets-1} est dotée de deux domaines d'activation de la transcription notés ici par commodité TA1 et TA2. La protéine p58^{c-ets-2} présente une organisation analogue (Schneikert *et al.*, 1992 ; Chumakov *et al.*, 1993). Par contre, la protéine p54^{c-ets-1}, qui diffère de la p68^{c-ets-1} à l'extrémité amino-terminale, ne possède que le domaine TA2 (Gégonne *et al.*, 1992 ; Schneikert *et al.*, 1992) (Figure 13).

Entre les domaines TA1 et TA2, la région conservée entre les protéines Ets-1 et Ets-2 d'une part, et Erg et Fli-1 d'autre part, constituerait une zone de modulation de l'activité transcriptionnelle des deux domaines précédemment décrits (Schneikert *et al.*, 1992). Cette région régulatrice favoriserait les propriétés transactivatrices du domaine TA1 et inhiberait celles du domaine TA2. Elle adopterait une conformation d'hélice-boucle-hélice. Les facteurs de transcription MyoD et E12 s'associent grâce à ce type de domaine. Par analogie, le domaine hélice-boucle-hélice des protéines Ets-1 et Ets-2 pourrait être le siège d'interactions avec des cofacteurs transcriptionnels.

La composition en amino-acides des domaines TA1 et TA2 n'est analogue à celle d'aucune des catégories de domaines transactivateurs connus, si ce n'est qu'une succession de huit cystéines dispersées au sein de TA1 évoque le domaine transactivateur en doigt de zinc de la protéine Tat du virus HIV-1 (Kuppuswamy *et al.*, 1989 ; Ruben *et al.*, 1989 ; Schneikert *et al.*, 1992). Toutefois, cette hypothèse semble infirmée par le fait que seules trois de ces cystéines sont conservées au cours de l'évolution (Albagli *et al.*, 1992).

Les protéines Erg et Fli-1 possèdent également deux domaines transactivateurs, mais l'un des deux correspond curieusement à la zone régulatrice des protéines Ets-1 et

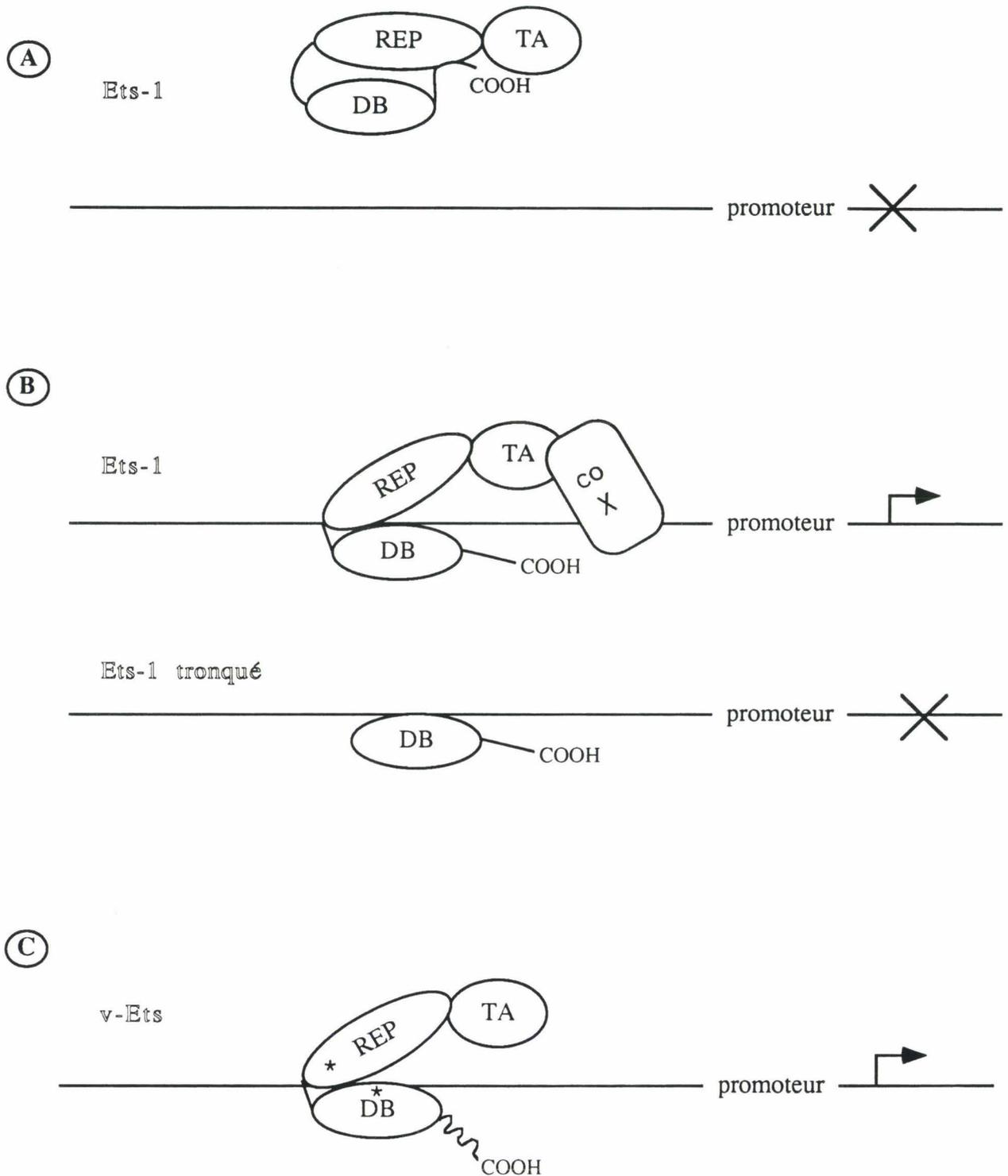


Figure 12 : Modèle d'inhibition intra-moléculaire de la fixation à l'ADN de la protéine p68^{c-ets-1} (tiré de Lim *et al.*, 1992). DB = domaine de liaison à l'ADN ; TA = domaine transactivateur aminoterminal ; REP = domaine répresseur. CO X = cofacteur X. Les aminoacides mutés dans v-Ets par rapport à c-Ets sont indiqués par une astérisque.

Ets-2, tandis que la seconde est située à l'extrémité carboxy-terminale de chaque molécule (Rao *et al.*, 1993 ; Siddique *et al.*, 1993 ; Zhang *et al.*, 1993) (Figure 13).

Le domaine transactivateur de la protéine PU.1 se restreint essentiellement aux 75 amino-acides amino-terminaux (Hagemeier *et al.*, 1993). Celui de Elf-1 fait exception à la règle (à l'absence de règle ?) des protéines Ets puisque son domaine transactivateur est clairement de type acide (C.Y. Wang et J.M. Leiden, observations non publiées).

Il apparaît donc que les protéines de la famille Ets possèdent des domaines transactivateurs localisés à divers endroits de chaque molécule, et dont la structure n'obéit à aucun critère vraiment particulier, à quelques exceptions près qui sont peut-être plutôt le fait de l'acharnement passionné des chercheurs que de la réalité.

En fait, si l'on compare les niveaux d'activation obtenus aux cours des tests *in vitro*, il apparaît clairement que les protéines Ets constituent un groupe de transactivateurs extrêmement faibles : on est bien loin des activations de 50 à 100 fois du complexe AP-1 ! Par contre, plusieurs de ces protéines agissent synergiquement avec d'autres facteurs de transcription comme nous allons le voir. Et au moins l'une d'elle, la protéine PU.1, est capable d'interagir directement avec le complexe TFIID par son domaine transactivateur (Hagemeier *et al.*, 1993). L'ensemble de ces observations étaye l'intérêt de lancer les recherches à venir sur la piste des cofacteurs des protéines Ets.

2.4. Interactions intermoléculaires :

Un grand nombre de facteurs de transcription présentent la particularité de s'associer à l'ADN sous forme d'homodimères (voir revue de Jones, 1990). Ils peuvent également constituer un complexe hétérodimérique avec des facteurs de transcription de la même famille comme le font entre autres les protéines à motif hélice-boucle-hélice basique de type c-Myc (Blackwood et Eisenman, 1991; Blackwood *et al.*, 1992) ou les protéines à motif "Leucine zipper". La spécificité de reconnaissance de l'ADN peut ainsi être largement modulée par la combinatoire de différents partenaires de dimérisation disponibles dans un type cellulaire donné. A titre d'exemple, le complexe MyoD/E12 a un effet positif sur la régulation des gènes de la différenciation musculaire tandis que le dimère MyoD/Id exerce un effet opposé (Benezra *et al.*, 1990). A l'intérieur d'une famille multigénique, ces interactions permettent d'accroître les possibilités de régulations des gènes en réponse à un signal transmis de la membrane cellulaire vers la noyau, par la mise en oeuvre d'un grand nombre de combinaisons protéiques.

Les protéines c-Ets-1 de différentes espèces semblent se fixer à l'ADN sous forme monomérique (Fisher *et al.*, 1991 et 1992 ; Hagman *et al.*, 1992 ; Lim *et al.*, 1992 ; Nye *et al.*, 1992). Par contre, ces protéines ainsi que d'autres protéines Ets apparentées présentent des interactions avec des protéines issues de familles

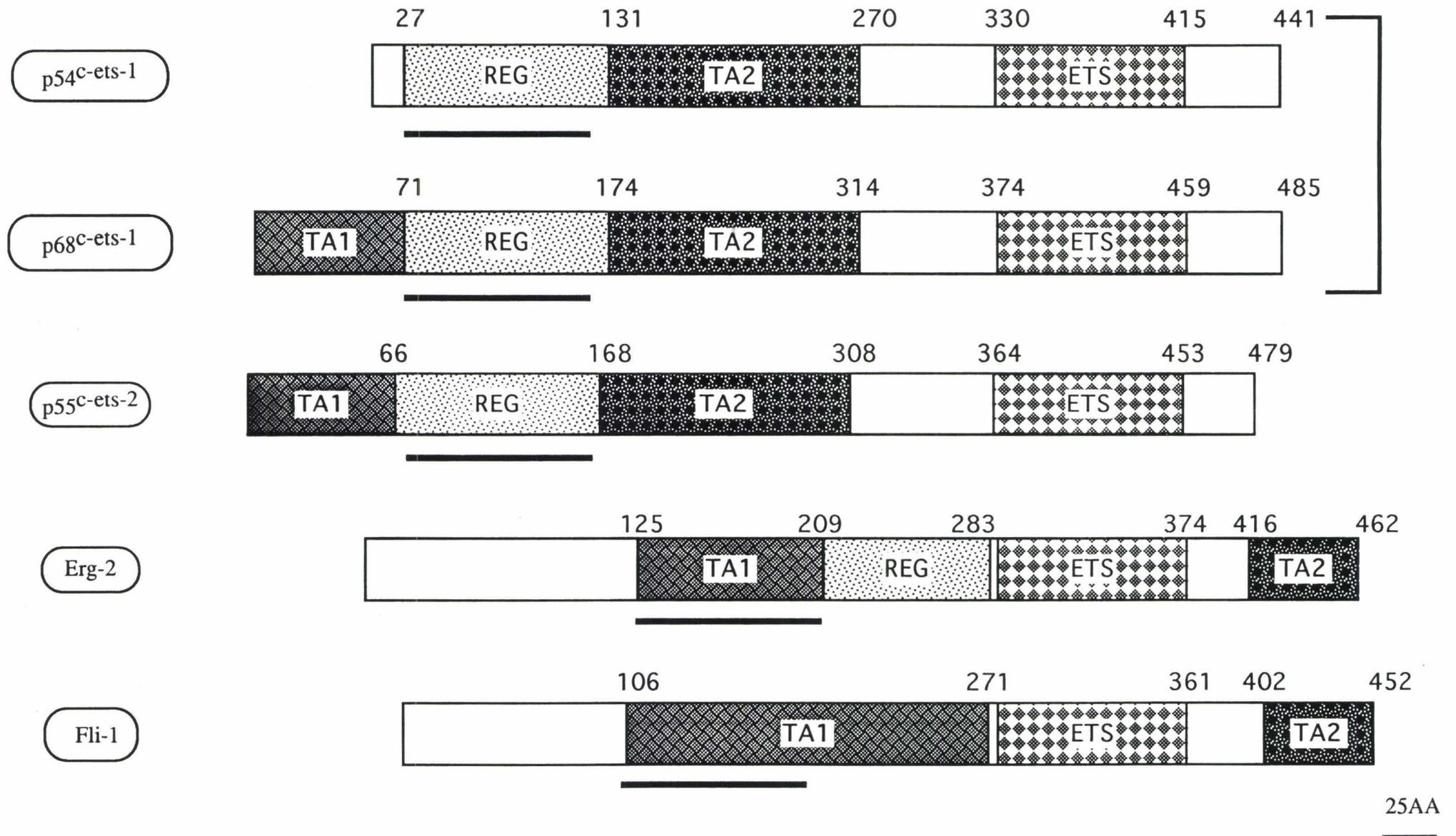


Figure 13 : Organisation des domaines fonctionnels de 4 types de protéines Ets étroitement apparentées. Le crochet souligne les 2 isoformes issues du gène *c-ets-1* aviaire, la barre horizontale indique la région aminoterminale homologue entre les protéines Ets-1, Ets-2 aviaires et Erg et Fli-1 humaines. TA1 = domaine transactivateur aminoterminal ; TA2 = domaine transactivateur carboxyterminal ; REG = domaine modulateur de l'activité transactivatrice ; ETS = domaine ETS. Les aminoacides sont numérotés.

différentes, impliquées dans des mécanismes physiologiques variés, ce qui accroît considérablement leur pouvoir transactivateur. D'après les données obtenues jusqu'à présent concernant quelques membres de la famille Ets, on peut arbitrairement décrire trois types de complexes auxquels elles sont associées, sur la base de leur mode de fixation à la séquence-cible.

2.4.1. La fixation du partenaire transcriptionnel dépend de celle de la protéine Ets :

L'étude de l'enhancer du gène codant la protéine ICP4 (Infected Cells Protein 4) du virus HSV a permis la description moléculaire d'un complexe transcriptionnel dont l'un des partenaires est une protéine de la famille Ets (Thompson et McKnight, 1992). Cet enhancer est composé de deux types d'éléments régulateurs : le premier est une séquence TAATGAG/AAT qui fixe un complexe Oct1/VP16/HCF (Host Cell Factor) (Goding et O'Hare, 1989 ; Katan *et al.*, 1990). La seconde séquence régulatrice est un dimère riche en bases puriques (CGGAAG/A)₂ (Figure 14). La purification de protéines capables de se lier à ce motif a révélé une composition multimérique (Triezenberg *et al.*, 1988 ; LaMarco et McKnight, 1989). Le clonage moléculaire de ce complexe, baptisé GABP pour GA binding protein, a permis l'identification de deux types de sous-unités, GABP α et GABP β (LaMarco *et al.*, 1991 ; Thompson *et al.*, 1991). Sur le plan fonctionnel, la sous-unité GABP α fixe le motif CGGAAG/A avec une affinité modérée, tandis que la sous-unité β ne peut se lier à ce motif isolément. Toutefois, le dimère α/β constitue un complexe de plus haute affinité pour l'ADN que la forme α seule (Thompson *et al.*, 1991).

L'examen de la manière dont se forme le complexe GABP indique trois interfaces moléculaires hautement spécifiques. Le domaine de GABP α nécessaire à l'interaction avec GABP β inclue un domaine ETS ainsi qu'un ensemble contigu vers l'extrémité carboxy-terminale de 35 amino-acides. Du côté de GABP β , ce sont quatre répétitions de trente-trois amino-acides localisés dans sa moitié amino-terminale qui suffisent à l'interaction avec GABP α . Cette interaction est hautement spécifique puisque GABP β n'interagit avec aucune autre protéine Ets (Brown et McKnight, 1992). De façon assez surprenante, ces quatre répétitions présentent des homologies de séquence significatives avec les motifs répétés de la protéine de levure SWI6 (Breedon et Nasmyth, 1987). De tels motifs répétés sont également retrouvés au sein de protéines transmembranaires comme les protéines Notch de drosophile (Wharton *et al.*, 1985), Glp-1 et Lin-12 de Nématodes (Yochem *et al.*, 1988 ; Yochem et Greenwald, 1989), de plusieurs polypeptides qui régulent la séquestration cytoplasmique du facteur de transcription NF- κ B (Ohno *et al.*, 1990 ; Haskill *et al.*, 1991), et de l'ankyrine, un constituant multifonctionnel du cytosquelette des hématies (Lux *et al.*, 1990). Compte tenu de la fonction des protéines dans lesquelles est retrouvé ce motif répété, il devient

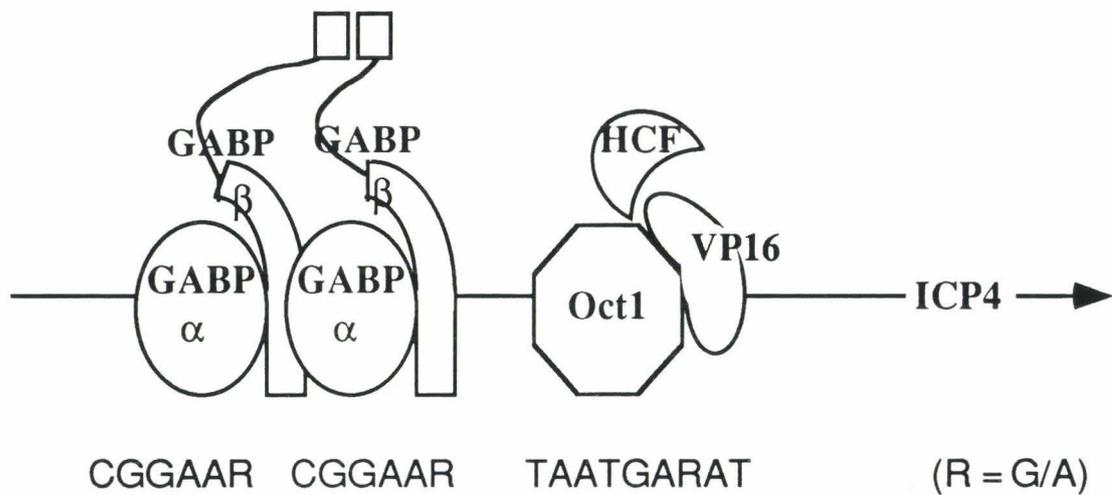


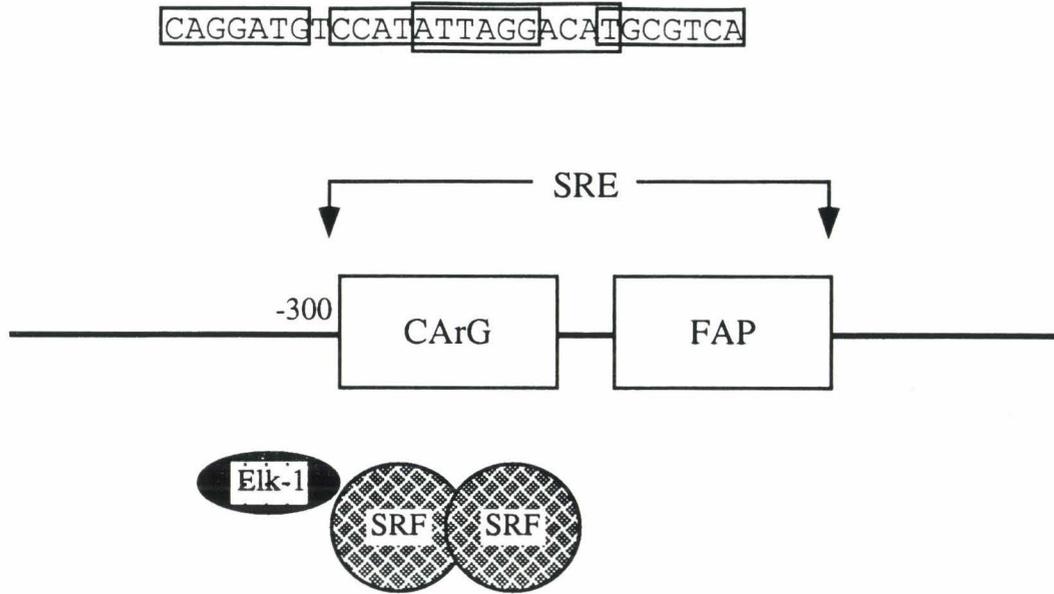
Figure 14 : Représentation schématique des complexes protéiques liés à l'enhancer du gène ICP4 du virus HSV-1

possible d'envisager un certain nombre de régulations de la formation du complexe GABP. Ainsi, le complexe pourrait être séquestré momentanément à l'extérieur du noyau, grâce à la reconnaissance du motif répété par un ligand cytoplasmique.

La seconde interface d'interaction est représentée par la liaison du complexe avec le motif CGGAAG/A. C'est bien sûr le domaine ETS de GABP α , et ce sont encore les motifs répétés de GABP β qui contactent l'ADN. Enfin, la dernière zone de contact impliquée dans l'assemblage d'un complexe ADN-GABP fonctionnel est un domaine de dimérisation localisé à l'extrémité carboxy-terminale de GABP β (Figure 14). GABP β forme des homodimères en solution, (Thompson *et al.*, 1991) et en présence de monomères de GABP α , chaque dimère capture deux molécules de GABP α pour former un complexe hétérotétramérique adapté à la reconnaissance du motif (CGGAAG/A)₂ (Bzik et Preston, 1986 ; Triezenberg *et al.*, 1988 ; Spector *et al.*, 1990). L'exemple du complexe GABP illustre parfaitement la façon dont peut s'opérer la discrimination entre différentes protéines Ets vis-à-vis d'un gène-cible. C'est grâce à la capacité que déploient GABP α et GABP β de se reconnaître spécifiquement qu'est assurée la sélectivité. De plus, GABP β est le premier cofacteur d'une protéine de la famille Ets qui semble pourvu d'une fonctionnalité hybride : d'une part, son motif répété pourrait lui conférer un rôle de relais cytoplasmique, d'autre part, il présenterait une caractéristique des facteurs de transcription puisque, d'après des résultats non publiés (Thompson et McKnight), c'est lui qui porterait l'activité transactivatrice. Compte tenu du fait que l'expression de GABP α et celle de GABP β est simultanément détectable dans de nombreux tissus différents (LaMarco *et al.*, 1991), il est très probable que leur interaction réciproque soit modulée au niveau post-traductionnel. Récemment, un rapport faisait état de la phosphorylation du complexe GABP *in vitro*, et de la sous-unité GABP β *in vivo* par la kinase Raf-1 activée, par l'intermédiaire d'une MAP kinase (Bruder *et al.*, 1993).

A l'inverse, la protéine PU.1 interagit avec un facteur qui, parce qu'il est strictement tissu-spécifique, pourrait assurer la sélectivité d'action du complexe : il s'agit de la protéine NF-EM5 qui stimule l'activité de PU.1 sur l'enhancer 3' du gène de la chaîne κ des immunoglobulines (Pongubala *et al.*, 1992). Il semble que les modalités d'interactions entre PU.1 et TFIID d'une part, et PU.1 et NF-EM5 d'autre part, sont bien différentes : NF-EM5 ne se lie pas à l'ADN sans PU.1. De plus, c'est une séquence PEST (riche en Proline, Glutamine, Sérine et Thréonine) située entre les amino-acides 118 et 160 de PU.1 qui est le siège de l'interaction avec NF-EM5. La présence de ce type de séquence semble corrélée à une dégradation enzymatique accrue des protéines intracellulaires (Rogers *et al.*, 1986). Il est donc possible d'envisager que le facteur NF-EM5 stabilise la protéine PU.1 en la protégeant de la protéolyse, favorisant ainsi son activité.

Ⓐ



Ⓑ

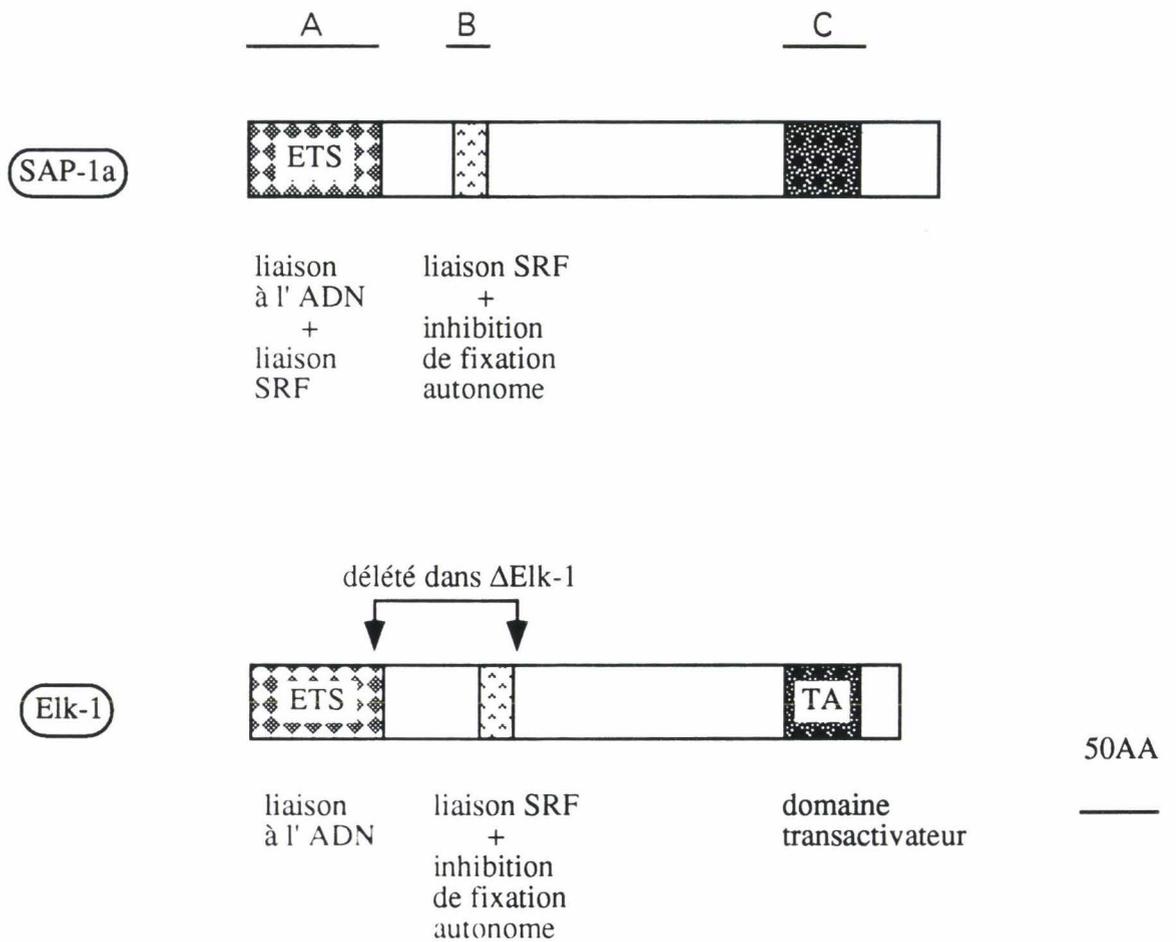


Figure 15 : A. Représentation schématique du motif SRE du promoteur du gène *c-fos* humain. B. Description des trois régions d'homologie des protéines SAP-1 et Elk-1.

2.4.2. La fixation de la protéine Ets dépend de celle de son partenaire transcriptionnel :

Le motif SRE (Serum Responsive Element) est une pièce maîtresse du promoteur du gène *c-fos*. En son absence, point d'expression. Il est composé d'une séquence appelée boîte CARG ou DSE (Dyad Symetry Element), adjacente à un élément de réponse aux protéines CREB, ATF ou AP-1, dénommé FAP (pour Fos/AP-1) (Figure 15A) (Treisman, 1990). L'intégrité de chacune des deux parties du SRE est indispensable à une inductibilité maximale du gène. La boîte CARG est le site de reconnaissance d'un dimère de la protéine SRF (Serum Responsive Factor). Mitoyen de l'extrémité 5' du site de liaison du SRF se trouve une séquence qui est reconnue par la protéine p62^{TCF} (Ternary Complex Factor). Il est démontré depuis peu que deux facteurs de transcription de la famille Ets sont doués de propriétés de liaison au SRE analogues à celles de la protéine p62^{TCF} : il s'agit de Elk-1 (Hipskind *et al.*, 1991) et de SAP-1 (Dalton et Treisman, 1992), et la liaison de chacune d'entre elles est strictement dépendante de la fixation préalable ou simultanée du SRF sur le site SRE. De plus, bien que la protéine Elk-1 soit vraisemblablement le véritable homologue de p62^{TCF}, les deux protéines Elk-1 et SAP-1 coopèrent individuellement avec le facteur SRF pour activer le promoteur du gène *c-fos*, probablement en stabilisant la liaison du SRF au motif SRE (Dalton et Treisman, 1992 ; Hill *et al.*, 1993). Les protéines Elk-1 et SAP-1 ont en commun trois régions d'homologie notées A, B et C (Figure 15B) : les deux régions amino-terminales A et B sont indispensables respectivement à la fixation à l'ADN et à la formation du complexe ternaire (Dalton et Treisman, 1992 ; Janknecht et Nordheim, 1992 ; Rao et Reddy, 1992). Au sein de la protéine Elk-1, le domaine C est un domaine d'activation transcriptionnelle qui est phosphorylé sur de multiples sites par une MAP kinase, tant *in vitro* qu'*in vivo* (Gille *et al.*, 1992 ; Marais *et al.*, 1993). Cette phosphorylation ne module pas la fixation au SRE, qui semble plutôt constitutive (Herrera *et al.*, 1989), mais par contre elle induit l'activité transcriptionnelle de Elk-1. Par conséquent, la protéine Elk-1 est le premier facteur de la famille dont l'activité est directement corrélée à la transmission d'un signal membranaire qui passe par l'activation des protéines Ras puis Raf, et enclenche la cascade des kinases cytoplasmiques dont les MAP kinases en fin de parcours.

2.4.3. La fixation des deux partenaires est indépendante :

Cette situation met en présence des protéines Ets et des facteurs de transcription *bona fide*, mais leurs modes d'interactions sont dans l'ensemble moins détaillés. Ainsi, la coexpression des protéines *c-Ets-1* et du complexe AP-1 aboutit à une stimulation synergique d'un certain nombre de gènes-cibles, dont l'enhancer des gènes précoces du virus du polyome (Wasylyk *et al.*, 1990) ou le promoteur de l'urokinase (Nerlov *et al.*, 1992) entre autres. Des protéines en position intermédiaire entre un site Ets/AP-1 distal

et un site AP-1 proximal jouent un rôle crucial pour assurer une coopérativité à longue distance de ces deux éléments de réponse lorsque le promoteur de l'uPA est stimulé par le TPA (Nerlov *et al.*, 1992).

Récemment, Gégonne *et al.* (1993) ont décrit précisément les mécanismes de coopérativité entre les protéines p54^{c-ets-1} et Sp1, tant au niveau de la transactivation que de la liaison à la région ERR1 (Ets responsive region 1) du LTR du virus HTLV1 (human T lymphotropic virus). Cette coopération met en jeu les 110 amino-acides carboxy-terminaux de la protéine c-Ets-1, c'est-à-dire une partie du domaine TA2 et le domaine ETS. La protéine p54^{c-ets-1} coopère également avec le facteur hématopoïétique GATA-1 pour assurer l'expression tissu-spécifique de la glycoprotéine mégacaryocytaire GPIIB (Lemarchandel *et al.*, 1993). Cette fois, les effets des deux facteurs de transcription mis en jeu sont additifs et non plus synergiques. Il est intéressant de noter qu'au voisinage de la région promotrice régulée par les facteurs c-Ets-1 et GATA-1 se trouve un site probablement modulé par Sp1 qui a un effet inhibiteur. Bien que les relations entre les sites Ets et Sp1 n'étaient pas à l'ordre du jour des investigations de ces auteurs, elles pourraient illustrer le jeu d'une même combinaison de facteurs de transcription ayant un effet positif ou négatif en fonction du gène ciblé.

3. Régulation de l'activité des protéines Ets :

3.1. Régulation du transport nucléaire :

C'est le transport nucléaire des protéines c-Ets-1 aviaire et murine qui a été le mieux examiné jusqu'à présent. Après transfection de cellules COS-1 de singe par un vecteur exprimant les protéines p54^{c-ets-1} ou p68^{c-ets-1}, la localisation nucléaire de ces protéines a pu être visualisée par immunofluorescence (Boulukos *et al.*, 1989 ; Pognonec *et al.*, 1989 ; Leprince *et al.*, 1990). Le domaine responsable de cette localisation est un signal peptidique de séquence GKRKNKP débutant à l'acide aminé 376 de la protéine p54^{c-ets-1} aviaire. La délétion de toute la région amino-terminale de la protéine entière incluant ce peptide la restitue en majeure partie au cytoplasme. Cependant, une fluorescence résiduelle localisée dans le noyau indique que des amino-acides situés en position carboxy-terminale par rapport au signal de translocation nucléaire seraient aussi impliqués dans cette activité.

Ce type de motif protéique est similaire au domaine minimum de translocation nucléaire de l'antigène T du virus SV40 et d'autres protéines nucléaires comme la protéine c-Myc (Richardson *et al.*, 1986 ; Dang et Lee, 1988). L'analyse informatique de sa conformation prévoit qu'il forme une boucle émergeant à la surface de la protéine p54^{c-ets-1} (Pognonec *et al.*, 1989). Par analogie avec le modèle proposé pour l'antigène

T du virus SV40 (Goldfarb *et al.*, 1986), la protéine pourrait être prise en charge par le biais de cette boucle externe par un transporteur actif. Une hypothèse alternative, mais non antagoniste de la précédente, consiste à imaginer un facteur cytoplasmique qui séquestrerait la protéine en fonction des stimulations extra-cellulaires, comme cela est déjà observé pour les facteurs de transcription de la famille NF- κ B (Bauerle et Baltimore, 1988). Cette hypothèse permettrait d'expliquer pourquoi, sous certaines conditions expérimentales, la protéine p54^{c-ets-1} a pu être observée dans le cytoplasme (Ghysdael *et al.*, 1986b ; Fujiwara *et al.*, 1988).

3.2. Modifications post-traductionnelles :

3.2.1. Phosphorylations :

La stimulation mitogénique de lymphocytes T (Pognonec *et al.*, 1988) et B (Fisher *et al.*, 1991) a pour conséquence une phosphorylation immédiate et intense des protéines p54^{c-ets-1} aviaire et p63^{c-ets-1} murine. Cette hyper-phosphorylation résulte d'une augmentation du taux de calcium intra-cellulaire et touche essentiellement des résidus Sérines (Pognonec *et al.*, 1988). La phosphorylation des protéines c-Ets-1 n'est pas altérée après inactivation de la PKC par un inhibiteur comme la staurosporine (Fisher *et al.*, 1991) et elle n'est pas induite par le TPA dans ces cellules (Pognonec *et al.*, 1988). La stimulation des protéine kinases dépendantes de l'AMPc n'influence pas non plus leur taux de phosphorylation. L'examen de la séquence protéique révèle la présence de deux sites consensus de phosphorylation par la calcium/calmoduline kinase II (RXXS/T) et un site potentiel de phosphorylation par la caséine-kinase II. Cependant l'intervention de la caséine kinase II est peu probable puisqu'elle n'est activée au niveau du noyau qu'assez tard au cours de l'activation lymphocytaire (Gehalen et Harrison, 1984), ce qui est incompatible avec une rapide phosphorylation des protéines c-Ets-1. Quant à la calcium/calmoduline kinase II, son association avec le noyau n'a pour l'instant été démontrée que dans une lignée tumorale (Delcayre *et al.*, 1989). Par conséquent, le mystère reste entier, d'autant plus que, dans d'autres cellules comme les cellules endothéliales, le signal de phosphorylation de la protéine Ets-1 pourrait passer par la PKC (V. Fafeur, communication personnelle).

La phosphorylation a pour conséquence majeure de diminuer de façon notable l'affinité des protéines Ets-1 pour l'ADN (Pognonec *et al.*, 1989). Ses résultats ont été obtenus sur de l'ADN-cellulose, et l'on ignore encore si la fixation de ces protéines sur un motif nucléotidique de type GGAA est également modulée par phosphorylation. Toutefois, une équipe a récemment visualisé une forme hyperphosphorylée de la protéine Ets-1 pendant la mitose de lignées lymphocytaire T. Or, il ne semble pas que la fixation de cette isoforme à son site de reconnaissance soit altérée (Fleischman *et al.*, 1993). On peut donc imaginer que la régulation de ses gènes-cibles par la protéine Ets-1

peut être modulée par son taux de phosphorylation en fonction du cycle cellulaire. Cette modification de la protéine Ets-1 s'appliquant au domaine transactivateur carboxyterminal (TA2, Figure 13), elle pourrait également en altérer la conformation et, par suite, l'activité.

La phosphorylation influence également la liaison à l'ADN de la protéine apparentée Ets-2 (Boulukos *et al.*, 1990). Toutefois, cette protéine est phosphorylée en réponse aux activateurs de la PKC (Fujiwara *et al.*, 1988), tout comme l'est la protéine Erg humaine (Murakami *et al.*, 1993).

Les conséquences de la phosphorylation sur la protéine PU.1 sont différentes (Pongubala *et al.*, 1993). En effet, il semble que dans ce cas, la phosphorylation de PU.1 sur la Sérine 148 soit indispensable à l'interaction avec le co-facteur NF-EM5 pour former un complexe de transactivation efficace de l'enhancer 3' du gène de la chaîne κ des immunoglobulines. Ce résidu serait phosphorylé tant *in vitro* qu'*in vivo* par la caséine kinase II. Par conséquent, il est permis de penser que les protéines Ets-1 et PU.1 pourraient jouer des rôles antagonistes pendant l'activation lymphocytaire B (la protéine PU.1 n'est pas exprimée dans les lymphocytes T), dont l'équilibre serait perpétré par le jeu des protéines kinase.

Enfin, rappelons que c'est par phosphorylation par une MAP kinase que s'exprime l'activité transcriptionnelle de la protéine Elk-1 vis-à-vis du promoteur du gène *c-fos*.

3.2.2. Modification des groupements thiols :

Il a été montré récemment que inhibiteurs de protéases modifient les groupements thiols des protéines Ets-1 humaines, de façon covalente (Fisher *et al.*, 1992). Les candidats à ce type de modifications sont les Cystéines 99, 112 et 169, toutes situées en dehors du domaine de liaison à l'ADN. De fait, modifiées par le TLCK (p-tosyl-L-lysine chloromethyl ketone), ces protéines sont toujours capables de reconnaître le motif PEA3 par exemple. Il semble qu'*in vivo*, les protéines Ets-1 passeraient alternativement d'une forme oxydée à une forme réduite qui seule pourrait reconnaître un motif de reconnaissance (Wasylyk et Wasylyk, 1993). Toutefois, les effecteurs de cette régulation restent obscurs.

Un autre intérêt de l'étude de Fisher *et al.* (1992) est qu'elle suggère de surcroît que les protéines Ets-1 humaines pourraient dimériser en solution, sans doute covalamment, mais la reconnaissance de l'ADN ne peut se faire sous cette forme. Si une telle dimérisation se vérifie effectivement *in vivo*, elle constituerait un mode de régulation tout-à-fait original pour un facteur de transcripton de sa liaison à l'ADN.

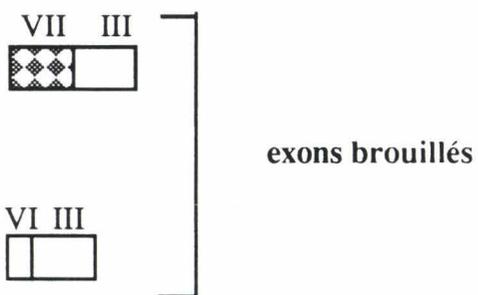
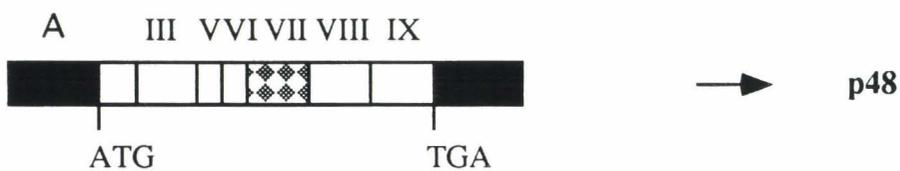
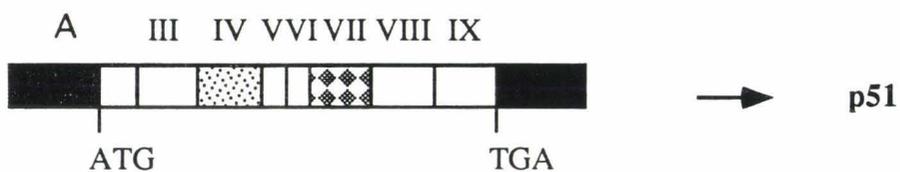
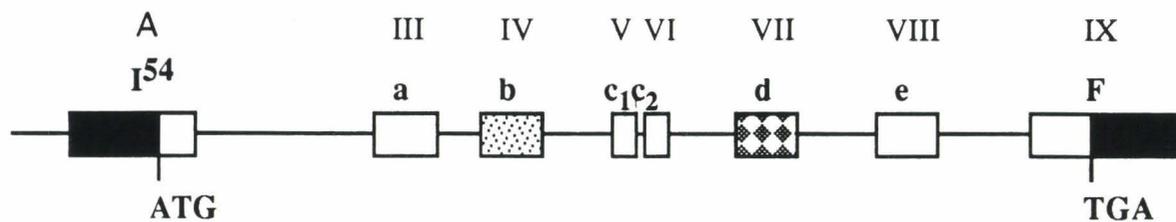


Figure 16 : Les différents épissages du locus *ets-1* humain. En lettres minuscules, nomenclature équivalente chez le poulet. Les régions non codantes sont indiquées en noir.

3.3. Régulation par épissage alternatif :

L'épissage alternatif des ARNm est une stratégie employée par le génome pour engendrer, à partir d'un seul gène, de multiples protéines douées de propriétés biologiques qui peuvent être antagonistes ou du moins distinctes. A cet égard, un exemple significatif de modulation de la fonction d'un facteur de transcription est illustré par l'épissage différentiel de l'ARNm *FosB* qui peut, en plus de la protéine FosB de pleine longueur, engendrer une protéine tronquée, Δ FosB, dépourvue de domaine transactivateur. Comme la protéine Δ FosB possède encore le domaine de dimérisation et de liaison à l'ADN, elle exerce un effet dominant négatif sur la protéine FosB et, par suite, inhibe l'activité du complexe AP-1 (Mumberg *et al.*, 1991 ; Nakabeppu et Nathans, 1991).

Le mécanisme d'épissage alternatif peut également créer une diversité des extrémités non codantes et de ce fait moduler la stabilité des ARNm ou leur traductibilité.

Les protéines de la famille Ets n'échappent pas à ce type de régulation, mais les conséquences de l'épissage différentiel des gènes qui les codent ne sont pas toujours très claires.

3.3.1. Epissage alternatif modulant la fixation à l'ADN :

Le mécanisme d'épissage alternatif peut aussi aboutir à la création de facteurs de transcription dont les propriétés de liaison à l'ADN sont modifiées. C'est le cas des protéines Ets-1 humaines (Reddy et Rao, 1988 ; Koizumi *et al.*, 1990 ; Jorcyk *et al.*, 1991) qui peuvent être amputées des amino-acides codés par l'exon VII (Figure 16), par l'exon IV ou par les deux exons conjointement. La perte de l'exon IV aboutit à la production d'une protéine p48 partiellement dépourvue du domaine équivalent au domaine de régulation du domaine transactivateur de la protéine aviaire. L'épissage de l'exon VII supprime les signaux majeurs de phosphorylation, et si l'on se souvient de l'influence des phosphorylations sur la perte d'affinité pour l'ADN, on comprend alors que la protéine p42 dépourvue des amino-acides codés par cet exon pourrait fixer constitutivement sa séquence-cible. De fait, Pognonec *et al.* (1990) ont isolé un variant murin équivalant à la protéine humaine sans exon VII, dont les propriétés de liaison à l'ADN ne sont pas affectées par phosphorylation. Dans le même ordre d'idée, l'épissage de l'exon VII supprime la répression intra-moléculaire de la fixation à l'ADN observée sur la protéine de pleine longueur (Wasylyk *et al.*, 1992).

3.3.2. Epissage alternatif modulant l'interactivité avec un cofacteur :

L'exemple le plus spectaculaire de la création de deux protéines Ets à rôles antagonistes est celui de l'épissage différentiel de l'ARNm de la protéine Elk-1. Nous

avons vu que cette protéine s'associe en compagnie du facteur SRF au motif SRE pour activer le promoteur du gène *c-fos*. Or, il existe une autre forme de la protéine Elk-1, baptisée Δ Elk-1 (Rao et Reddy, 1993), amputée de la totalité du domaine qui contribue à la fois à l'interaction avec le SRF et à l'inhibition de la fixation à l'ADN (Figure 15). La perte de onze amino-acides carboxy-terminaux de son domaine ETS n'empêche pas la protéine Δ Elk-1 de reconnaître le motif SRE de façon autonome. Elle pourrait donc exercer un effet répresseur du complexe Elk-1/SRF par compétition pour la fixation au motif SRE, et donc empêcher par la même l'activation du gène *c-fos*. De plus, la perte conjointe de ses sites de phosphorylation pourrait entraver sa régulation.

La protéine SAP-1 existe sous deux isoformes distinctes, SAP-1a et SAP-1b (Dalton et Treisman, 1992). L'ARNm codant la protéine SAP-1b présente une région 3' non codante différente de celui qui code la forme SAP-1a, ce qui pourrait être révélateur d'une différence de stabilité. Dans la partie codante, la troisième région d'homologie avec Elk-1 (du côté carboxy-terminal) est plus courte que celle de SAP-1a. En termes de fonctionnalité, l'interaction de SAP-1b avec le SRF serait plus efficace que celle de SAP-1a. Cette observation suggère donc un rôle régulateur vis-à-vis de l'interaction avec le SRF de l'extrémité carboxy-terminale. Toutefois, les deux formes semblent activer le promoteur du gène *c-fos* de façon équivalente.

Chez le poulet, l'épissage des exons correspondant aux exons IV et VII (exons b et d, Figure 16) n'est pas décrit pour l'instant. Par contre, l'ARNm *c-ets-1* aviaire présente un épissage tout-à-fait intéressant des exons 5'. De cet épissage résulte la traduction d'une $p68^{c-ets-1}$ et d'une $p54^{c-ets-1}$ (Leprince *et al.*, 1988) dont l'organisation sera décrite sur la figure 26. Grâce à ce mécanisme, la $p68^{c-ets-1}$ possède un domaine transactivateur supplémentaire dont la $p54^{c-ets-1}$ est dépourvue, ce qui pourrait indiquer que l'activité des deux domaines TA1 et TA2 est additive, et expliquer que la $p68^{c-ets-1}$ transactive mieux le promoteur de la stromélysine par exemple que ne fait la $p54^{c-ets-1}$ (Wasylyk *et al.*, 1991). Il est possible de formuler l'hypothèse selon laquelle ce domaine supplémentaire constituerait une cible pour des co-facteurs différents de ceux de la $p54^{c-ets-1}$.

3.3.3. Epissages dont les conséquences sont inconnues :

Un autre membre de la famille Ets pour lequel des protéines alternatives ont été décrites est le gène *erg* humain. L'ARNm *erg2* est plus long de 218 pb codantes en 5' que la forme *erg1*, et son site de polyadénylation est plus précoce (Rao *et al.*, 1987). Des événements d'épissages différentiels d'exons plus internes ont également été décrits (Duterque-Coquillaud *et al.*, 1993). Toutefois, aucune modulation de la fonctionnalité des différents produits n'a pu être observée, tant en ce qui concerne leur spécificité de fixation à l'ADN que leurs propriétés transactivatrices. Cependant, il faut noter que pour l'instant, aucun véritable gène-cible des protéines Erg n'est connu. Par conséquent, il

n'est pas possible, comme dans le cas idéal du promoteur du gène *c-fos*, de tester leur comportement dans le cadre d'un complexe transcriptionnel.

Enfin, dans des lignées cellulaires, sont observés des épissages très curieux du gène *ets-1* humain, puisqu'ils modifient l'ordre des exons, tout en préservant le cadre de lecture : au niveau de l'ARNm, en amont de l'exon III peut se trouver soit l'exon VI, soit l'exon VII (Cocquerelle *et al.*, 1992) (Figure 16). Ces ARN (messagers ?) présentent également l'originalité d'être circulaires, et ils ne sont pas polyadénylés (Cocquerelle *et al.*, 1993). Résultent-ils d'une erreur de lecture ou d'épissage, ou bien ont-ils une réelle fonctionnalité ? Un autre gène dont la maturation de l'ARNm aboutit à une inversion d'exons est le gène DCC (Deleted in colorectal carcinomas) (Nigro *et al.*, 1991), mais deux des trois produits formés perdent le cadre de lecture, ce qui soulève un doute profond sur la fonctionnalité biologique de tels produits.

3.4. Stabilité des protéines Ets :

De même que bon nombre de facteurs de transcription ont une durée de vie courte, celle de certaines protéines de la famille Ets comme la protéine Ets-1 n'excède pas 45 à 90 minutes (Pognonec *et al.*, 1988 ; Koizumi *et al.*, 1990). Dans certains cas, la stabilité pourrait être modulée par phosphorylation comme tel est le cas de la protéine Ets-2 dont la durée de vie passe de 20 minutes à plus de deux heures sous l'action des activateurs de la PKC (Fujiwara *et al.*, 1988).

Cette rapide dégradabilité est sans doute due au fait que la plupart des protéines de la famille Ets présente une séquence PEST. Au sein des protéines Ets-1, une région particulièrement riche en amino-acides de type PEST se situe entre le domaine TA2 et le domaine ETS. S'il s'avère que la protéolyse est effective à ce niveau, ce qu'il reste à démontrer, les deux fragments qui en résulteraient portant chacun un domaine fonctionnel, pourraient exercer un effet transdominant négatif de l'activité des protéines de pleine longueur. De fait, la plupart des expériences de retard sur gel qui ont été menées jusqu'à présent révèle l'existence d'un complexe de haute mobilité électrophorétique qui pourrait correspondre à un produit de protéolyse, toujours doué de reconnaissance d'un site Ets mais détaché de son extrémité amino-terminale.

Cependant, le cas de la protéine Erg contredit ce mécanisme. En effet, la protéine entière est relativement riche en résidus P, E, S et T, surtout au niveau d'une région qui lui est spécifique située entre le domaine TA et le domaine ETS (amino-acides 230 à 296 environ, figure 13). Malgré cela, elle semble beaucoup plus stable que les protéines Ets-1 et Ets-2, puisque sa durée de vie est supérieure à 21 heures (Murakami *et al.*, 1993). On ignore si la phosphorylation module sa stabilité, en la protégeant de l'action protéolytique.

4. Conclusions :

Des analyses fonctionnelles des différents domaines des protéines de la famille Ets ont permis, dans l'ensemble des cas, de leur attribuer au moins un domaine transactivateur et un domaine de liaison à l'ADN. Ces caractéristiques incitent à les classer dans la catégorie des facteurs de transcription. Mais ils tirent leur grande originalité du fait que leur éventuelle spécificité d'action respective repose peu ou pas du tout sur la reconnaissance d'une cible nucléotidique individuelle. L'interaction avec des partenaires transcriptionnels semble beaucoup plus à même de discriminer la protéine Ets qui sera recrutée sur un gène-cible donné, un peu à la manière des hormones stéroïdes sur les récepteurs nucléaires. Toutefois, leur activité semble polymorphe puisque certaines de ces protéines facilitent la liaison à l'ADN du co-facteur (cas de GABP α), d'autres voient leur fixation facilitée par celle du co-facteur, comme c'est le cas de la protéine Elk-1. Ce dernier facteur remplit également le rôle de relais nucléaire des informations transduites de la membrane vers le noyau, pour activer un gène de réponse précoce, le gène *c-fos*.

Il est également possible d'envisager d'autres rôles comme la reconnaissance d'éléments initiateurs (Inr) requis pour initier la transcription à partir de promoteurs sans boîte TATA (Smale et Baltimore, 1989), éléments généralement constitués de l'enchaînement de bases puriques (pyrimidiques sur le brin complémentaire) rappelant le motif GGA. D'ailleurs, une récente communication a fait état du fonctionnement d'un motif de reconnaissance pour les protéines Ets comme élément initiateur, en l'absence de séquences promotrices/activatrices additionnelles (Yu *et al.*, 1993).

Certaines protéines Ets pourraient également, comme le fait la protéine GAGA de drosophile (Gilmour *et al.*, 1989) maintenir des régions d'ADN dans une conformation invulnérable à la répression nucléosomale.

Pour finir, le fait que les protéines Ets compensent un pouvoir transactivateur généralement faible par une activité synergique de plusieurs autres facteurs de transcription rappelle le comportement du facteur de levure GRF2 dont la fixation à l'ADN provoque la création d'une zone dépourvue de nucléosomes et donc accessible aux autres facteurs de transcription avec lesquels GRF2 coopère (Chasman *et al.*, 1990).

	thymus	rate	BLc	TLC	cellules érythroïdes	macrophages	poumon	foie	cerveau	rein	coeur	testicule	glande mammaire
Ets-1	+	+	+	+	-	-	+	+		+	+	+	
Ets-2	+	+	+	+		+	+	+		+	+	+	
Elf-1	+	+	+	+									
Fli-1	+	+		+			+		-	-	+	-	
PU.1	-	+	+	-	+	+	-	-	-	-	-		
Spi-B	-	+	+	-	+	+							
GABP α	+	+		+				+	+	+		+	
PEA3	-		-	-		-			+			+	+
ER81	+	+					+	-	+	+	+	+	
ER71	-						-	-	-	-	-	+	
Elk-1							-		+			+	
ERMp			-	-			+	-	+	+	+	+	

Tableau 9 : Expression de certains membres de la famille Ets dans les tissus de différentes espèces animales.

4ème partie : Rôles physiologiques des protéines de la famille Ets

Pour être clair, ce titre est quelque peu présomptueux. En effet, à l'heure actuelle, les implications des protéines Ets dans la physiologie cellulaire ne sont qu'entr'aperçues. Leur approche se fait selon deux axes principaux : en premier lieu, la description précise de la distribution tissulaire de chaque membre désigne des types cellulaires dans lesquels ces protéines ou certaines combinaisons d'entre elles ont des chances d'être actives. Le tableau 9 résume cet aspect, bien qu'il ne soit donné qu'à titre indicatif, sans tenir compte de la sensibilité de la méthode qui en a fourni le contenu (Northern blot, hybridation *in situ*, transcription inverse/PCR), ni du fait que des cellules hématopoiétiques peuvent éventuellement contaminer ces tissus (c'est probablement le cas de Fli-1 par exemple). D'après ce tableau, il apparaît qu'aucun tissu n'est exempt de sa protéine Ets. Par conséquent, un certain nombre de systèmes cellulaires transposables *in vitro* ont permis d'identifier de gènes potentiellement cibles. Hélas, dans la plupart des cas, et compte tenu du fait que différents membres sont exprimés dans une même cellule, ces études ne proposent guère que des candidats sans désigner pour autant de véritables effecteurs de la régulation.

C'est pourquoi, en attendant l'inactivation séquentielle de différents gènes de la famille Ets par recombinaison homologue dans un avenir que l'on espère proche, les seules ébauches de modèles disponibles pour l'instant sont essentiellement basées sur des corrélations d'expression et d'activation d'un gène cible *in vitro*.

1. Activité des protéines Ets dans les cellules immunocompétentes :

Le système immunitaire constitue un outil de choix pour étudier la régulation génique par les protéines Ets, en partie parce que certaines d'entre elles sont surtout exprimées dans les cellules immunitaires (Ets-1, Spi-B, PU.1, Elf-1, Fli-1) (Tableau 9). D'autre part, les différentes étapes qui concourent à la mise en place séquentielle d'un certain nombre de marqueurs phénotypiques au cours du développement embryonnaire et au cours des processus d'activation sont relativement connues, du moins chez la souris.

1.1. Ontogenèse des lymphocytes T :

Le thymus est le site majeur de maturation fonctionnelle des lymphocytes T. En effet, c'est au sein de cet organe que les thymocytes apprennent à discriminer le soi du non-soi. C'est aussi au contact de l'épithélium thymique qu'ils expriment les glycoprotéines membranaires CD4 et CD8 qui signent leur différenciation en sous-populations fonctionnellement distinctes. Le thymus est initialement peuplé de

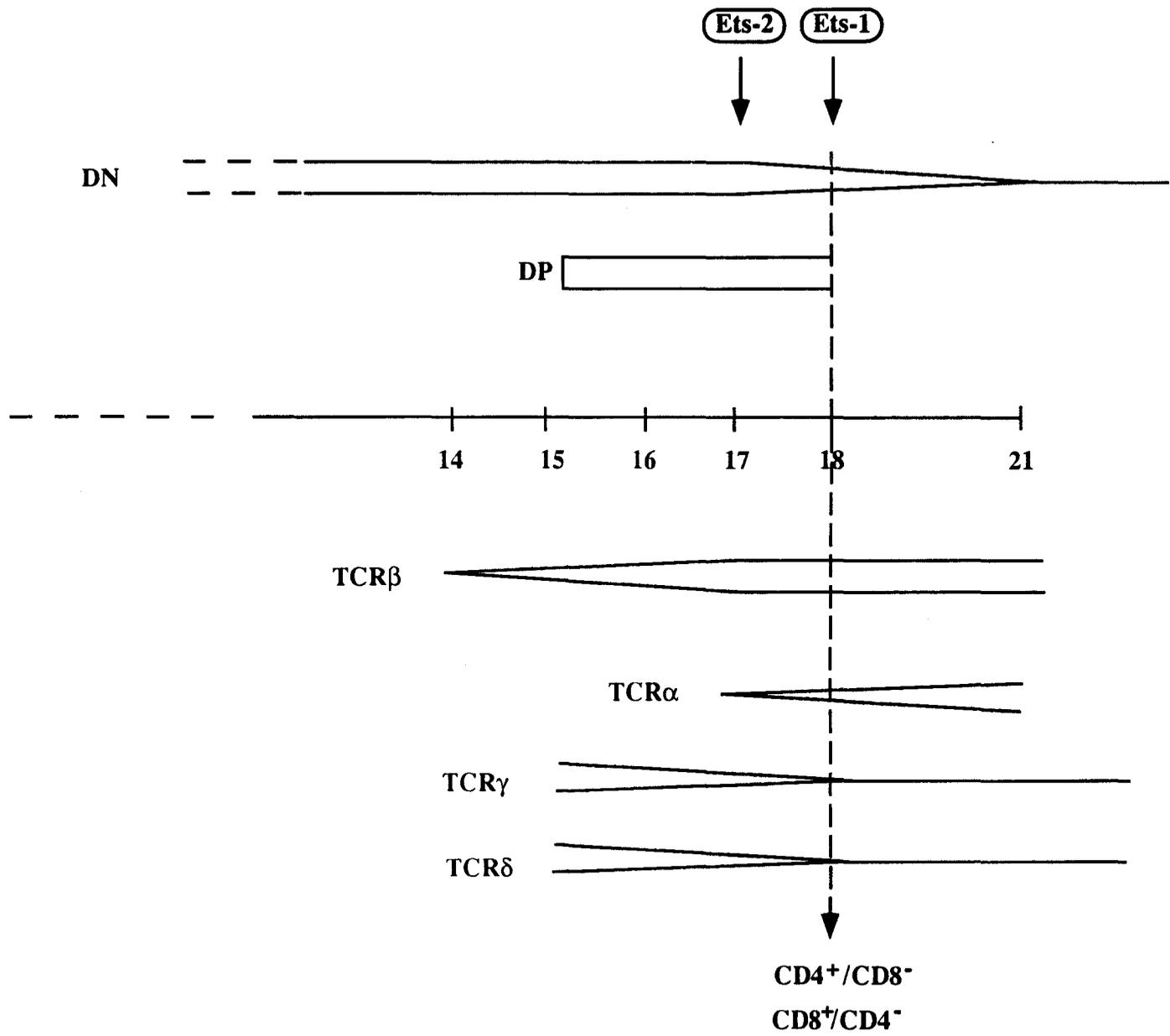


Figure 17 : Schéma de l'ontogénie des différentes chaînes constitutives des TCR chez la souris

thymocytes doubles négatifs CD4⁻/CD8⁻ (DN) dont certains deviennent doubles positifs CD4⁺/CD8⁺ (DP) vers le 15^{ème} jour de gestation (Figure 17). C'est à ce stade que s'ébauche le répertoire lymphocytaire T par sélection des clones auto-réactifs ou dénués de réactivité (Kisielow et von Boehmer, 1990).

Vers le 18^{ème} jour de gestation, des thymocytes se différencient en cellules présentant des caractéristiques fonctionnelles matures CD4⁺/CD8⁻ et CD8⁺/CD4⁻ (von Boehmer *et al.*, 1990). Ces deux sous-populations migrent du thymus vers les organes lymphoïdes périphériques et y exercent une fonction régulatrice de la réponse immunitaire (lymphocytes T auxiliaires CD4⁺) ou un rôle à la fois cytotoxique vis-à-vis de cellules infectées ou tumorales et suppresseur de la réponse immunitaire (lymphocytes T cytotoxiques/suppresseurs CD8⁺) (Sprent *et al.*, 1986). La fonction des marqueurs phénotypiques CD4 et CD8 consiste à restreindre la reconnaissance des lymphocytes T à un interlocuteur cellulaire respectivement porteur de molécules du complexe majeur d'histocompatibilité de classe II (cellules immunocompétentes) ou de classe I (toutes les autres cellules).

Parce que de nombreux auteurs ont observé une expression prédominante du gène *c-ets-1* aviaire (Chen, 1985 ; Ghysdael *et al.*, 1986b ; Quéva *et al.*, 1993), murin (Bhat *et al.*, 1987) ou humain (Chen, 1985 ; Bhat *et al.*, 1990) dans les tissus lymphoïdes, une corrélation entre cette expression et la chronologie d'apparition des marqueurs de différenciation lymphocytaire a été recherchée. D'une façon générale, les ARNm *c-ets-1* et *c-ets-2* sont détectés en plus grande abondance dans les thymocytes (Bhat *et al.*, 1987 et 1989), mais l'on ignore si la réduction d'expression de ces gènes dans les lymphocytes adultes résulte d'un renouvellement plus rapide des ARNm ou d'une activation plus faible de leur promoteur.

Chez l'embryon de souris, l'apparition de l'ARNm *c-ets-1* coïncide avec l'apparition des thymocytes matures de phénotype auxiliaire et cytotoxique/suppresseur. L'expression du gène *c-ets-2* de souris est déjà constitutive à ce stade puisqu'elle apparaît plus précocément. Une expression préférentielle de ces deux gènes est observée dans la sous-population CD4⁺, et se prolonge au niveau des lymphocytes circulants (Bhat *et al.*, 1989). D'ailleurs, les deux protéines sont capables d'activer le promoteur du gène *CD4 in vitro* (Salmon *et al.*, 1993).

C'est aussi au cours du développement des thymocytes que se réarrangent les gènes des chaînes constitutives du récepteur T (TCR pour T cell receptor). Les réarrangements aléatoires des multiples segments V, D et J qui sont productifs aboutissent à la création de TCR fonctionnels de type α/β ou γ/δ sur lesquels reposent la spécificité de reconnaissance de l'antigène et la destruction des clones potentiellement auto-réactifs. La régulation transcriptionnelle des gènes correspondants commence à être bien connue et elle suit des modalités similaires dans chacun des cas : les quatre gènes sont exprimés à partir de promoteurs relativement faibles, qui

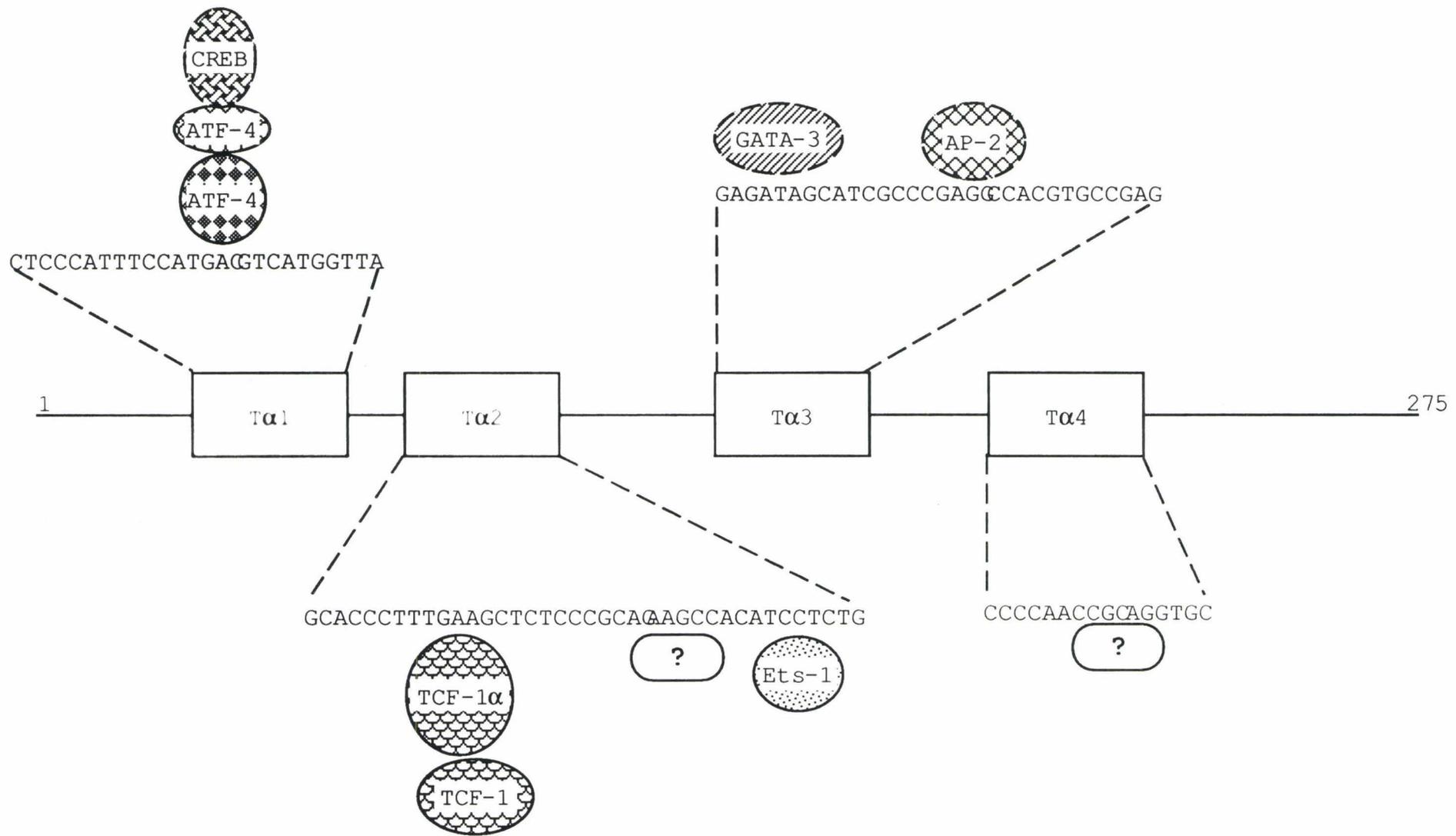


Figure 18 : Organisation des zones régulatrices de l'enhancer 3' du gène de la chaîne α du TCR

n'assurent pas la tissu-spécificité : l'expression lymphocytaire est due à une combinaison d'éléments régulateurs tels que des enhanceurs puissants stimulés par des facteurs de transcription dont certains sont spécifiques des cellules lymphoïdes. Des éléments silenciers pourraient restreindre l'expression des chaînes α/β d'une part, et γ/δ d'autre part à chaque sous-population. Enfin, la régulation des recombinaisons des différents segments V, D, J, C pourrait être indépendante de l'expression des chaînes de récepteur et donc constituer ainsi une première étape de spécialisation (pour revue, voir Leiden, 1993).

La chronologie de réarrangements des chaînes de TCR (Figure 17) indique que l'expression de la protéine Ets-2 coïncide avec l'apparition de la chaîne α réarrangée. De fait, bien que cette protéine n'ait pas été testée directement, l'élément T α 2 de l'enhancer 3' du gène codant la chaîne α du TCR (TCR α) nécessite la liaison de protéines de type Ets en conjonction avec d'autres facteurs de transcription pour être complètement activé (Ho *et al.*, 1990 ; Leiden, 1992). Cet élément enhancer est d'ailleurs reconnu *in vitro* par la protéine Ets-1. Un grand nombre de facteurs de transcription appartenant à des familles de gènes différentes est vraisemblablement impliqué dans la régulation complexe de cet enhancer (Figure 18). Leur activité semble redondante puisque des mutations affectant l'activité des motifs T α 1 et T α 2 peuvent être compensées si l'intégrité des éléments T α 3 et T α 4 est respectée, bien que ces derniers éléments ne soient pas doués isolément d'une activité enhancer importante (Ho et Leiden, 1990a). En fait, il apparaît que le motif T α 2 est le pivot de l'expression du gène TCR α , d'abord parce qu'il semble être le seul élément qui soit reconnu de façon nette par un facteur de transcription strictement lymphocytaire (TCF). Ensuite parce que ce motif testé isolément en test de transactivation pratiqué dans la lignée cellulaire Jurkat est répresseur (Ho et Leiden, 1990b). L'hypothèse a donc été formulée selon laquelle, dans le contexte de l'enhancer natif, l'expression du gène est réprimée par le seul complexe fixé par l'élément T α 2. *In vivo*, l'activation transcriptionnelle du gène serait liée à la fixation conditionnelle sur les motifs adjacents des autres complexes qui, en interagissant avec les protéines liées à T α 2, abrogent la répression constitutive. Par conséquent, la protéine de type Ets fixée sur le motif T α 2 pourrait interagir avec l'un ou l'autre des facteurs fixés sur les éléments T α 1, T α 3 ou T α 4 de l'enhancer du gène TCR α .

Le gène de la chaîne β réarrangée du TCR possède également un enhancer dont l'organisation est analogue à celle du gène TCR α (4 motifs T β 1 à T β 4) et qui répond aussi aux gènes de la famille Ets. La protéine c-Ets-1 aviaire cotransfectée avec l'élément T β 2 réprime à la fois l'activité basale conférée par le motif enhancer et l'activité induite par le TPA (Prosser *et al.*, 1992). *In vivo*, il apparaît que l'expression de Ets-1 dans l'embryon de souris est concomitante de l'époque à laquelle la transcription du gène TCR β devient à peu près constante (Figure 17). Si cette protéine

est véritablement impliquée dans la régulation transcriptionnelle du gène TCR β , cette observation pourrait indiquer que, dans le contexte de l'enhancer natif, le rôle répresseur de Ets-1 est contre-balançé par les facteurs liant les autres motifs enhancers du gène, et elle expliquerait qu'*in vivo*, l'expression de la protéine Ets-1 ne semble pas être incompatible avec celle du TCR β .

Un autre gène potentiellement cible de la famille Ets est le gène codant l'enzyme terminal désoxynucléotidyl transférase (TdT) (Ernst *et al.*, 1993). Cette enzyme est exprimée transitoirement à l'aube du développement lymphocytaire et cesse de l'être au fil de la maturation de ces cellules (Chang, 1971 ; Hermans *et al.*, 1989). Elle intervient lors des processus de recombinaison des gènes de TCR en insérant des nucléotides aléatoirement, aux jonctions des segments recombinés (Kallenbach *et al.*, 1992).

Un élément de régulation du promoteur du gène de la TdT est reconnu à affinité égale par les protéines Ets-1 et Fli-1 *in vitro*. Il est peu vraisemblable que la protéine Ets-1 soit le véritable effecteur de la régulation de ce gène *in vivo*, puisqu'elle est exprimée beaucoup plus tardivement (Bhat *et al.*, 1989). Quant à la protéine Fli-1, les données disponibles concernant son expression ne sont pour l'instant que trop parcellaires pour qu'aucune hypothèse puisse être émise à son sujet.

Il semble donc que les protéines Ets sont impliquées dans la régulation de gènes clés de la maturation lymphocytaire. A nouveau, les exemples rapportés ici illustrent l'importance des interactions qu'entretiennent cette famille protéique avec les autres facteurs de transcription.

1.2. Activation lymphocytaire T:

L'interaction d'un déterminant antigénique présenté dans le contexte d'une molécule d'histocompatibilité de classe appropriée par une cellule présentatrice d'antigène avec le TCR induit, en quelques minutes de contact, une série d'évènements ordonnés dont résultent des modifications morphologiques et prolifératives qui permettent la réponse immunitaire de lymphocytes circulants, quiescents jusqu'alors. Cette interaction peut être mimée par des ionophores calciques, des lectines, des anticorps dirigés contre le TCR, ou par le TPA. Douze heures après le début de l'activation, la cellule T revêt une morphologie blastique, puis elle se divise en 24 à 48 heures pour finalement se différencier au fur et à mesure que sont séquentiellement activés les gènes de réponse immune (voir Crabtree, 1989 pour revue).

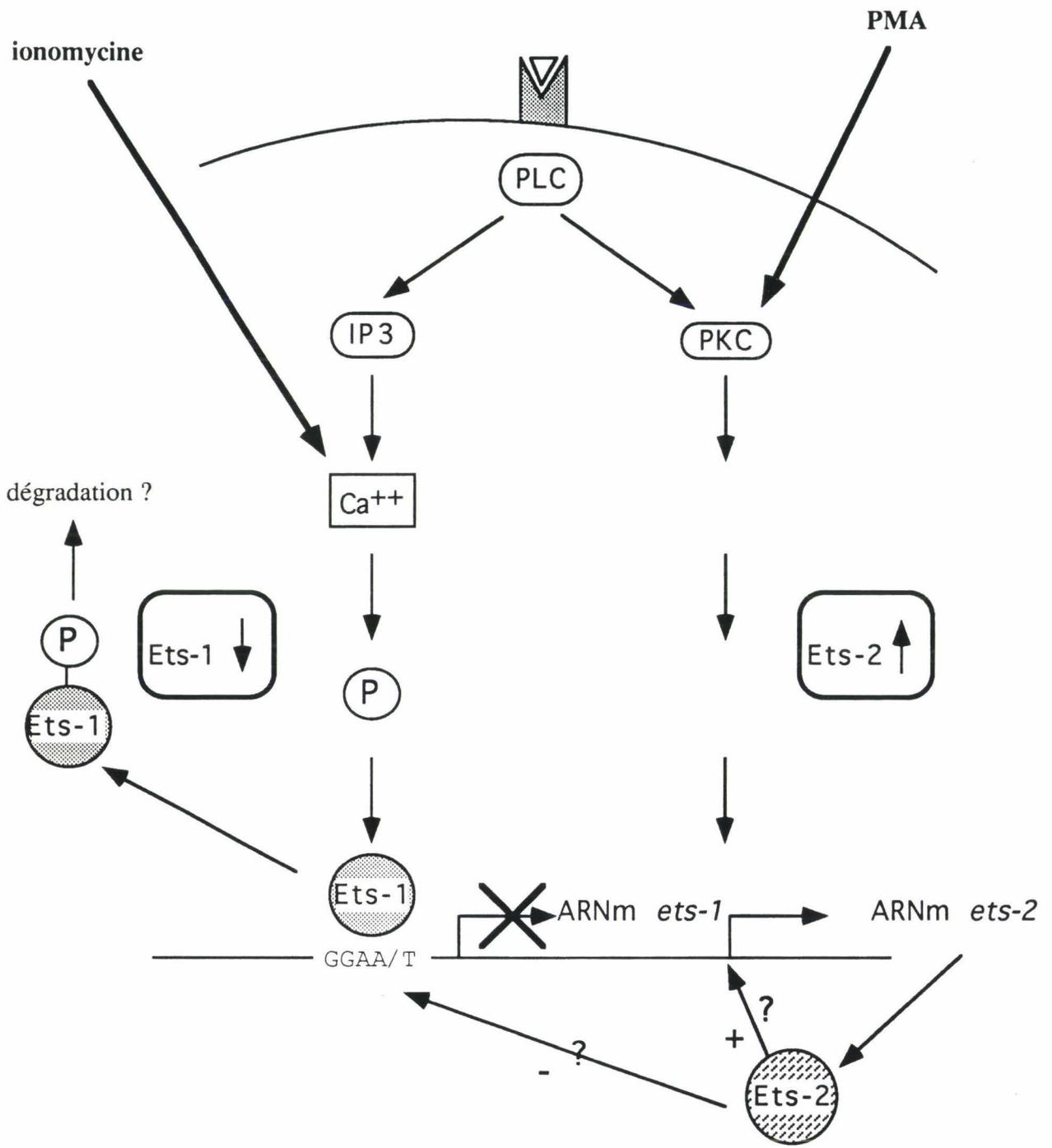


Figure 19 : Modèle de régulation de l'expression des protéines Ets-1 et Ets-2 humaines au cours de l'activation lymphocytaire

1.2.1. Balance Ets-1/Ets-2 :

En culture, des lymphocytes humains quiescents expriment une quantité relativement élevée d'ARNm et de protéine Ets-1, mais très peu de Ets-2. Réciproquement, l'activation des mêmes cellules par un anticorps CD3 inverse les rapports d'expression des gènes *ets-1* et *ets-2*, sauf si les synthèses protéiques sont bloquées par un traitement à la cycloheximide (Bhat *et al.*, 1990). D'après ces données, *ets-2* peut être considéré comme un gène de réponse précoce dont l'activité semble être corrélée à la prolifération cellulaire. C'est ce que suggéraient déjà des résultats préalables montrant que l'expression de cette protéine est induite très précocément dans le foie en régénération (Seth *et al.*, 1989). A l'inverse, l'expression de la protéine Ets-1 est plutôt corrélée au maintien de l'état quiescent des lymphocytes. Peut-être active-t-elle la transcription de gènes inhibiteurs de l'entrée dans le cycle cellulaire, ou réprime-t-elle l'expression de gènes impliqués dans la prolifération. On se souviendra à ce stade que la phosphorylation de la protéine Ets-1 résultant de l'élévation du taux de calcium intracellulaire constitue un événement immédiat de l'activation lymphocytaire et a pour conséquence d'abolir sa fixation à l'ADN. L'ensemble de ces résultats permet d'élaborer une ébauche de modèle, résumé sur la figure 19 : la reconnaissance d'un déterminant antigénique par le TCR active d'une part la PKC et donc enclenche la cascade des kinases cytoplasmiques, d'autre part *via* la production d'IP3 (Inositol tri-phosphate), elle provoque une augmentation du calcium intracellulaire. Cette seconde voie suffit à altérer la protéine Ets-1 au niveau post-traductionnel et la détache de l'ADN auquel elle est fixée pendant la quiescence cellulaire. Un second niveau de contrôle s'exercerait au niveau transcriptionnel puisqu'un facteur néo-synthétisé réprime l'expression du messenger *ets-1*, comme l'indique les expériences employant la cycloheximide. Le même facteur, ou un facteur différent mais inductible, pourrait être simultanément responsable de l'activation conjointe du gène *ets-2*.

Ce modèle est intéressant mais appelle quelques commentaires. En effet, la protéine Ets-2 phosphorylée indirectement par la PKC, perd son affinité pour l'ADN (Boulukos *et al.*, 1990). La transcription accrue de l'ARNm *ets-2* codant une protéine qui perdrait à cette occasion son activité biologique peut sembler paradoxale. Il est néanmoins permis d'envisager que la protéine est douée d'une activité non élucidée qui soit indépendante de sa liaison à l'ADN : par exemple, elle pourrait piéger un co-facteur qui contribuerait à bloquer la prolifération cellulaire plus ou moins directement. Ou bien, les résultats obtenus par Northern blot, qui détectent tous les types d'ARNm *ets-2*, ne rendraient pas compte des formes différemment épissées qui pourraient perdre leurs sites de phosphorylation. Cependant, pour l'instant, les différents ARNm *ets-2* connus ne diffèrent que par leurs sites de polyadénylation (Watson *et al.*, 1990).

Il convient finalement de garder à l'esprit que la perte d'affinité des protéines Ets-1 pour l'ADN, en réponse à des activateurs lymphocytaires, n'a pas encore été établie vis-à-vis de la cible nucléotidique de type GGAA/T de ces protéines.

1.2.2. Gènes cibles potentiels

- régulation de gènes de réponse immune :

Conjointement à la description de la répartition spatio-temporelle des différentes protéines Ets, il est nécessaire de pouvoir leur attribuer des gènes cibles si l'on veut comprendre la portée de leur intervention aux différents stades de la maturation lymphocytaire et de la réponse immunitaire. Le gène *lck* code une tyrosine kinase lymphocytaire dont l'association avec les molécules CD4 ou CD8, et avec la chaîne β du récepteur de l'IL-2 suggère qu'elle intervient au cours de l'activation et de la régulation de la prolifération des cellules T (Rudd *et al.*, 1988 ; Hatakeyama *et al.*, 1991). Un site de reconnaissance pour les protéines Ets est essentiel à l'activité du promoteur de type I de ce gène dans les cellules Jurkat (Leung *et al.*, 1993). Bien que la protéine Ets-1 reconnaisse ce motif au cours d'expériences de retard sur gel et qu'elle transactive le promoteur entier *in vitro*, à nouveau, elle n'est pas nécessairement la protéine impliquée dans l'activation de ce promoteur *in vivo*. En effet, de nombreuses protéines de la famille Ets sont exprimées dans les cellules Jurkat, et d'autre part, ce site est stimulé dans des cellules n'exprimant pas la protéine Ets-1.

Le gène de l'interleukine 2 (IL-2) constitue un modèle d'étude des mécanismes moléculaires dont dépend l'inductibilité de certains gènes eucaryotes. Or, ce gène est clairement inductible puisqu'il n'est pas transcrit dans les lymphocytes T quiescents mais le devient par contre une heure après délivrance du signal d'activation. Sa transcription est modulée par un enhancer complexe spécifique des cellules T. Il est en partie régulé par deux motifs nucléotidiques riches en purines, NFAT-1 et NFIL-2B, qui incluent des sites potentiels de reconnaissance pour des protéines de la famille Ets. De fait, si les protéines Ets-1 et Ets-2 ne peuvent fixer ces motifs, la protéine Elf-1 en est capable (Thompson *et al.*, 1992). Toutefois, cette protéine n'est pas responsable à elle seule de l'inductibilité du gène de l'IL-2 puisque son expression est constitutive dans les lymphocytes T tant quiescents qu'activés. En fait, le motif NFAT-1 est reconnu par un complexe multi-protéique dont deux autres membres identifiés sont les protéines JunB et Fra-1 (Boise *et al.*, 1993). La cinétique d'induction du gène *junB* est identique à celle de l'IL-2, tandis que le gène *fra-1* est plutôt d'expression constitutive dans ces cellules. C'est donc apparemment de l'expression du gène *junB* que dépendrait l'inductibilité du gène de l'IL-2. Parallèlement, il est conjointement possible qu'un complexe NFAT partiel se forme dans le cytoplasme et que le signal d'activation induise sa translocation vers le noyau où l'attend la protéine JunB néo-formée (Jain *et al.*, 1992), ce qui représenterait un niveau supplémentaire de régulation. De fait, la

protéine Elf-1 est liée à la protéine Rb non phosphorylée dans les lymphocytes quiescents, et se voit relarguée en conséquence de la phosphorylation de Rb dès l'entrée des cellules dans le cycle cellulaire (Wang *et al.*, 1993).

-régulation de LTR de virus lymphotropes :

L'activation transcriptionnelle du génome d'un provirus intégré nécessite la participation de facteurs de transcription de la cellule-hôte qui se lient au LTR viral. Si le facteur NF- κ B semble suffire à stimuler l'expression des gènes du rétrovirus HIV-1 dans les lymphocytes T, il n'en est pas de même en ce qui concerne son cousin HIV-2 dont les manifestations symptomatiques sont du reste différentes. De fait, l'activation du LTR de HIV-2 nécessite également l'action d'une protéine voisine de Elf-1, qui pourrait coopérer avec NF- κ B (Leiden *et al.*, 1992). Mais l'incubation de complexes formés de protéines nucléaires et du LTR de HIV-2 avec un anticorps dirigé contre la protéine Elf-1 n'en retarde pas la migration en gel d'acrylamide. Donc, c'est probablement une protéine apparentée à elf-1 qui participe au complexe.

Ce résultat pourrait expliquer la sensibilité de ce virus à la cyclosporine. En effet, cette drogue agirait au niveau transcriptionnel sur un facteur qui modifierait la protéine "Elf-1-like" de telle sorte qu'elle ne puisse plus être transloquée vers le noyau, ou qu'elle soit incapable de former un complexe fonctionnel avec Fra-1 et JunB (Randak *et al.*, 1990 ; Jain *et al.*, 1992). La protéine JunB étant la seule composante inductible du complexe NF-AT, elle constitue une cible acceptable de l'action de la cyclosporine.

Le rétrovirus HTLV-1 est un autre virus qui utilise entre autres effecteurs des protéines Ets à seule fin d'activer la transcription de son génome (Bosselut *et al.*, 1990 ; Gitlin *et al.*, 1991). La protéine Ets-1 en est en partie responsable, en synergie avec le facteur Sp1 (Gégonne *et al.*, 1993). Si le virus met à profit l'activation lymphocytaire pour s'exprimer, on voit que ces résultats contredisent ceux de Bhat *et al.* (1990) qui montraient une chute de l'expression de Ets-1 au tout début du processus. Faut-il plutôt incriminer la protéine Ets-2 dans l'activation du LTR du virus HTLV-1 *in vivo*, d'autant plus que cette protéine est également capable de le stimuler *in vitro* (Bosselut *et al.*, 1990) ? En fait, des résultats récents indiquent que c'est plutôt la protéine Elf-1 qui est responsable de l'activation de ce LTR (Clark *et al.*, 1993).

Par conséquent, il semble que la protéine Elf-1 ou une protéine apparentée soit un effecteur acceptable des trois promoteurs pris ici en exemple. Il a été dit, dans la troisième partie de ce mémoire que la spécificité de cible nucléotidique peut constituer un facteur de sélection des protéines Ets-1 et Elf-1. Compte tenu de l'expression majoritairement embryonnaire de la protéine Ets-1 chez la souris, on peut imaginer qu'elle sera impliquée dans l'établissement d'un phénotype différencié tandis que l'activité de la protéine Elf-1 s'appliquera essentiellement aux gènes exprimés dans les lymphocytes matures.

1.3. Différenciation d' autres cellules immuno-compétentes :

1.3.1. Lymphocytes B

Le facteur PU.1, d'expression strictement hématopoiétique à l'exception des cellules T (Klemsz *et al.*, 1990 ; Galson *et al.*, 1993) est présent en quantité notable dans les lymphocytes B. Il a été dit plus haut que son activité dans ces cellules est dictée par la présence du facteur strictement tissu-spécifique, NF-EM5. Le complexe PU.1/NF-EM5 régule l'enhancer 3' du gène de la chaîne légère κ des immunoglobulines (Grant *et al.*, 1992 ; Pongubala *et al.*, 1992 et 1993). Au fil de la différenciation des lymphocytes B, se produit un accroissement de l'activité de cet enhancer (Pongubala et Atchinson, 1991) qui rend compte de la sécrétion d'anticorps par les plasmocytes ; une augmentation parallèle de l'expression de NF-EM5 pourrait être responsable de cette activation. C'est aussi au cours de l'ontogénie des lymphocytes B que PU.1 coopère synergiquement avec Ets-1 pour activer l'enhancer de la chaîne μ non réarrangée des immunoglobulines (Nelsen *et al.*, 1993). Le gène correspondant constitue le premier exemple d'un site de coopération de deux protéines de la famille Ets dont la distribution tissulaire est souvent distincte, mais parfois chevauchante comme c'est ici le cas.

1.3.2. Cellules de la lignée monocyttaire

Rares sont pour l'instant les données qui permettraient de comprendre les mécanismes moléculaires régulant l'expression de gènes spécifiques de la différenciation de la lignée myéloïde. C'est pourquoi une équipe s'est intéressée récemment aux modalités d'expression de la molécule CD11b. L'hétérodimère CD11b/CD18 est une intégrine exprimée exclusivement à la surface des monocytes/macrophages, granulocytes et cellules tueuses NK (natural killer). Si la chaîne CD18 de l'intégrine est synthétisée par tous les leucocytes (Law *et al.*, 1987), c'est le composant CD11b qui en restreint la distribution au trois lignées citées plus haut (Todd *et al.*, 1981). La synthèse de l'ARNm et de la protéine CD11b augmente au fil de la différenciation, et donc cette molécule peut être considérée comme un marqueur de la différenciation monocyttaire, régulé en partie au niveau transcriptionnel (Rosmarin *et al.*, 1989 ; Pahl *et al.*, 1992). L'étude du promoteur de ce gène met en lumière l'importance d'une région fixatrice des facteurs PU.1 et Sp1 dans son expression tissu-spécifique (Chen *et al.*, 1993 ; Pahl *et al.*, 1993). Le fait que CD11b ne soit pas synthétisé par les lymphocytes B qui sont pourtant riches en protéine PU.1 suggère que dans les cellules monocytaires, Sp1 remplace fonctionnellement NF-EM5. Sp1 est généralement considéré comme un facteur ubiquiste, bien que des données récentes indiquent que son expression (Saffer *et al.*, 1991), son affinité pour l'ADN (Borellini *et al.*, 1991) et les modifications post-traductionnelles qu'il subit (Jackson et Tjian, 1988 ; Schaufele *et al.*, 1990) peuvent moduler son activité. Par conséquent, le statut de Sp1

pourrait restreindre l'activité du facteur PU.1 aux cellules myéloïdes. Cette hypothèse peut être complétée par l'observation selon laquelle seules des expériences de footprint *in vivo* rendent compte de l'occupation tissu-spécifique des sites de fixation des facteurs Sp1 et PU.1 de ce promoteur dans les cellules myéloïdes (Chen *et al.*, 1993). Il est possible que la protéine PU.1 modifie la chromatine de façon à faciliter l'accessibilité de son site de liaison au facteur Sp1. L'exemple du promoteur du gène de l'intégrine CD11b pourrait donc illustrer de façon plus générale les mécanismes par lesquels un facteur de transcription d'expression ubiquiste peut réguler la transcription de gènes tissu-spécifiques.

La protéine Ets-2 est elle aussi majoritairement exprimée dans les macrophages différenciés et quiescents (Ghysdael *et al.*, 1986a ; Boulukos *et al.*, 1988) et son expression est activée indirectement par un facteur essentiel à la survie des macrophages aviaires en culture, le cMGF (Boulukos *et al.*, 1990). Le statut de Ets-2 dans les macrophages indique donc que la fonction de cette protéine n'est pas nécessairement liée à la prolifération cellulaire, comme dans le cas des cellules hépatiques ou des cellules T, mais peut aussi être associée au maintien de l'état quiescent et à la survie des cellules.

1.3.3. Mégacaryocytes

Rappelons simplement pour mémoire l'importance d'une région promotrice du gène codant une glycoprotéine membranaire spécifique des mégacaryocytes, la GPIIb (Lemarchandel *et al.*, 1993). Cette région est responsable de l'expression tissu-spécifique de la GPIIb grâce à la fixation et à la stimulation additive des protéines GATA-1 et Ets-1 à laquelle Fli-1 peut se substituer mais pas Ets-2.

2. Implications de la famille Ets dans le développement embryonnaire :

2.1. Développement de la drosophile :

La métamorphose qui signe le passage de l'état larvaire au stade nymphal chez la drosophile est initiée par un pic de production d'ecdysone, une hormone stéroïde, qui se produit à la fin du troisième stade larvaire. Ce pic hormonal déclenche des changements coordonnés du développement des tissus imaginaux qui se différencieront progressivement en tissus adultes, et des tissus larvaires qui pour la plupart disparaîtront. Les investigations portant sur la régulation génique de ces modifications dépendantes de l'ecdysone mettent à profit l'accessibilité d'observation des chromosomes polyténiques géants des glandes salivaires de la drosophile. En effet, ces chromosomes de glandes salivaires explantées et exposées à l'ecdysone, au troisième stade larvaire, présentent aussitôt des vagues de renflements successifs qui constituent

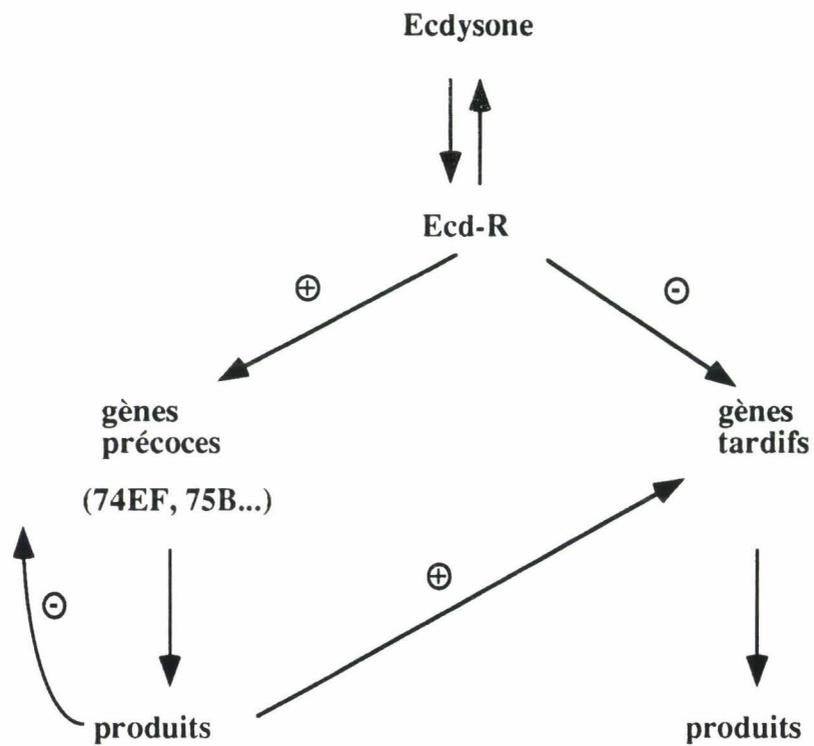


Figure 20 : Modèle de régulation génique par l'ecdysone au troisième stade larvaire du développement de la drosophile, d'après Ashburner.
Ecd-R = récepteur de l'ecdysone

autant de manifestations d'une activité transcriptionnelle intense, accessible à l'observation directe. Une étude détaillée des conditions requises à l'orchestration de cette activité génique a permis l'élaboration d'un modèle selon lequel la liaison de l'ecdysone à son récepteur induirait l'activité transitoire des gènes précoces en inhibant celle des gènes tardifs. Dans un second temps, les produits des gènes précoces activeraient une centaine de gènes tardifs et réprimeraient leur propre expression (Figure 20) (Ashburner *et al.*, 1974). Ce modèle postule que les produits précoces sont des régulateurs géniques d'expression certes ubiquiste, mais dont la combinatoire variée à différents stades du développement nymphal permet l'expression coordonnée des gènes d'expression tissu-spécifique (pour revue, voir Andres et Thummel, 1992).

Le gène *E74*, transcrit au niveau du locus *74EF* satisfait aux critères qui en font l'un des six gènes de réponse précoce de la métamorphose nymphale. A partir de trois promoteurs individuels sont synthétisés trois types d'ARNm qui codent deux protéines d'extrémités amino-terminales dissemblables, *E74A* et *E74B*. Ces protéines possèdent une partie commune qui inclue un domaine ETS à l'extrémité carboxy-terminale, et un domaine amino-terminal spécifique de chaque protéine, riche en acides aminés de type acide (Burtis *et al.*, 1990 ; Urness et Thummel, 1990). L'expression de *E74A* est induite transitoirement en réponse à chacun des six pics d'ecdysone qui se produisent tout au long du développement de la drosophile. Elle est d'ailleurs directement régulée par l'ecdysone (Thummel *et al.*, 1990), probablement au niveau d'un élément de réponse situé dans le premier intron (Burtis *et al.*, 1990 ; Karim et Thummel, 1991). La transcription du locus *74EF* est directement corrélée à l'apparition consécutive des transcrits issus des gènes tardifs, bien qu'aucun gène cible n'en soit clairement défini. D'ailleurs, le traitement d'embryon par la cycloheximide n'altère en rien l'activité transcriptionnelle du locus *74EF* mais entrave par contre l'activation des gènes tardifs, ce qui prouve bien la corrélation.

Le gène *E74A* est très rapidement réprimé, sauf après un traitement par la cycloheximide et bien que la protéine soit capable de reconnaître *in vitro* trois sites riches en bases GA dans un intron situé vers le milieu du gène (Urness et Thummel, 1990), il ne semble pas qu'elle soit l'effecteur de cette répression. C'est apparemment le facteur *E75*, une autre protéine de réponse précoce appartenant à la superfamille des récepteurs nucléaires (Segrave et Hogness, 1990) qui en est responsable. Cette répression de leur propre expression permet aux produits des gènes précoces de déterminer la durée de leur activité et donc, leur quantité accumulée en réponse à un pic hormonal.

Il semble que la transcription de *E74B* ne soit pas régie par les mêmes mécanismes que celle de *E74A*, puisque son expression peut ou non être concomitante de celle de *E74A*, tant en ce qui concerne leur répartition spatiale que leur expression temporelle (Figure 21). Surtout, l'apparition des transcrits *E74B* ne

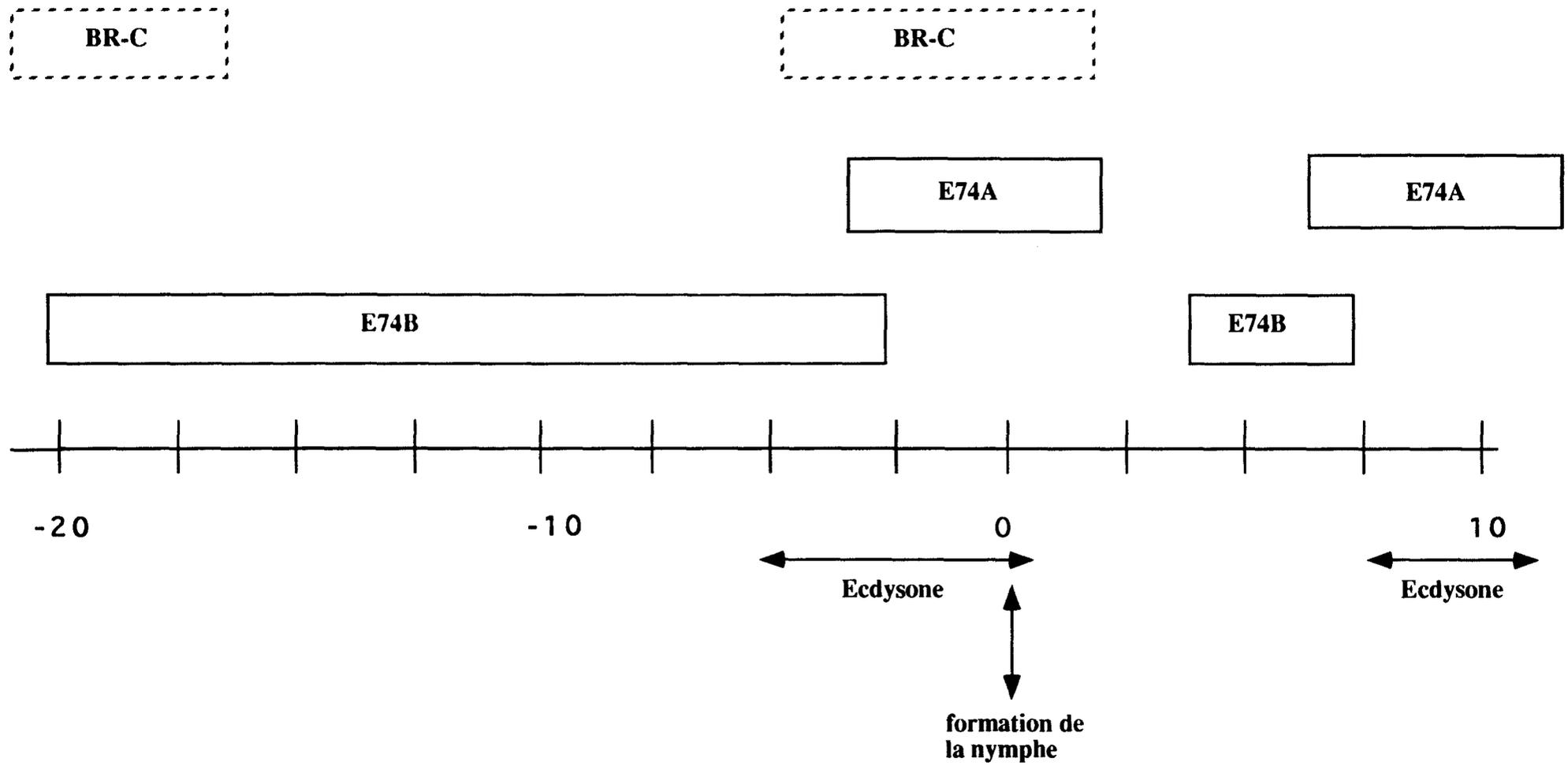


Figure 21 : Coordination des activités géniques en fonction des pics d'ecdysone.

coïncide pas systématiquement avec un pic d'ecdysone mais peut plutôt être activée en réponse à des doses infinitésimales d'hormone (Karim et Thummel, 1991). De fait, E74B, comme les protéines codées par le locus *BR-C* et comme le récepteur de l'ecdysone, est activé à une concentration de 2.10^{-9} M d'hormone. Au fil de l'augmentation du titre hormonal (2.10^{-8} M) sont exprimées les protéines E74A, E75A, E75B et à nouveau BR-C qui réprimerait l'expression des gènes du développement précoce (Karim et Thummel, 1992). Il est clair que le récepteur de l'ecdysone doit être rapidement et massivement produit pour permettre l'épanouissement de la réponse hormonale. Toutefois, on comprend difficilement que la réponse précoce ne soit pas inhibée par la cycloheximide si la synthèse du récepteur n'est pas accélérée. Il faut alors envisager le relais possible d'un autre récepteur nucléaire moins sensible à la quantité d'ecdysone, comme l'une des protéines issues du locus *BR-C* par exemple.

L'expression des autres gènes de la famille Ets identifiés jusqu'alors chez la drosophile ne présente pas les mêmes caractéristiques que celle de E74. Les gènes *D-ets-2* (Pribyl *et al.*, 1988), *D-elg* (Pribyl *et al.*, 1991), *D-ets-3*, *D-ets-4* et *D-ets-6* (Chen *et al.*, 1992) ne répondent pas à l'ecdysone. Les transcrits *D-ets-3* et *D-ets-6* sont trouvés à niveau égal pendant tout le développement. Les gènes *D-ets-2*, *D-ets-4* et *D-elg* sont également exprimés pendant la majeure partie du développement, bien qu'un peu plus discrètement pendant les stades larvaires. Pendant l'embryogénèse, c'est la répartition des ARNm *D-ets-2* qui semble la plus complexe, et surtout la plus évolutive. Une restriction de son expression particulièrement intéressante est dévoilée par l'analyse des mutants *pointed* qui présentent une déficience du système nerveux central par altération des interactions entre cellules gliales en migration et cellules nerveuses de la ligne médiane (Klämmt, 1993).

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

Les gènes pour lesquels une tissu-spécificité aura été montrée pendant le développement de la drosophile présentent exclusivement semble-t-il, un tropisme nerveux. C'est en effet également le cas d'un autre gène de la famille Ets, nommé *pokkuri* ou *yan* selon les auteurs (Lai *et al.*, 1992 ; Tei *et al.*, 1992) dont la déficience provoque une différenciation erronée de certaines cellules nerveuses de l'oeil en photorécepteur de type R7 surnuméraires.

Des expériences de Southern blot menées en conditions douces de haute salinité pourraient indiquer qu'il reste peu de gènes de la famille Ets à cloner chez la drosophile (Chen *et al.*, 1992). Cependant, cette affirmation n'engage que ses auteurs puisqu'il eut été préférable, pour dénombrer les protéines Ets de cet organisme, de pratiquer une immunoprécipitation à l'aide d'un anticorps dirigé contre le domaine de liaison à l'ADN.

2.2. Développement du xénope :

Seuls quatre gènes de la famille *Ets* sont identifiés chez le xénope à l'heure actuelle : il s'agit des gènes *X-ets-1* (Stiegler *et al.*, 1990) et *X-ets-2* (Wolff *et al.*, 1990 ; Burdett *et al.*, 1992), qui existent sous deux versions, notées A et B, du fait de la tétraploïdie du génome de certaines espèces d'Amphibiens (Kobel et Du Pasquier, 1986). La production des ARNm *X-ets-1* est plutôt corrélée à l'oogénèse et à l'embryogénèse, puisqu'elle est indétectable dans les tissus somatiques adultes, exceptions faites de l'ovaire et, curieusement, de la rate (Stiegler *et al.*, 1990 et 1993). Les ARNm *X-ets-1* sont intensément transcrits dans l'oocyte jusqu'au stade III et, pendant l'embryogénèse, au moment de la neurulation.

Les transcrits issus du gène *X-ets-2* présentent une distribution tissulaire beaucoup plus vaste (Wolff *et al.*, 1990). Toutefois leur rôle dans l'oogénèse a pu être clairement démontré après microinjection intra-oocytaire d'oligonucléotides anti-sens de l'ARNm *X-ets-2*, dont l'effet principal est d'abolir la dissolution de l'enveloppe nucléaire qui constitue normalement un marqueur de maturation méiotique (Chen *et al.*, 1990). La méiose peut cependant reprendre son cours normal après microinjection d'un ADNc *ets-2* humain, ce qui montre en plus de sa conservation fonctionnelle dans deux espèces éloignées, que la protéine *Ets-2* est l'un des facteurs clés de la maturation méiotique.

2.3. Développement du poulet :

Des expériences d'hybridation *in situ* chez l'embryon de poulet ont permis la description de la répartition spatio-temporelle des ARNm de la famille *Ets*, en particulier des deux transcrits codant les protéines p54 et p68, issues du gène *c-ets-1*. Globalement, l'expression de ces deux transcrits peut être observée dans bon nombre de dérivés mésodermiques (Vandenbunder *et al.*, 1989 ; Quéva *et al.*, 1993) parmi lesquels les vaisseaux sanguins en cours de formation, que celle-ci se fasse par angiogénèse (bourgeonnement de capillaires préexistants) ou par vasculogénèse (émergence *in situ* à partir d'une cellule souche), les cellules migratrices de la crête neurale, et les cellules du bourgeon plumaire en cours de condensation (Desbiens *et al.*, 1991). Si l'on veut tenter d'établir une règle claire portant sur l'expression du locus *c-ets-1* de poulet, on pourra simplement ajouter que ces transcrits sont complètement exclus des épithéliums, quelque soit leur origine, mésodermique, endodermique ou ectodermique (Vandenbunder *et al.*, 1989). L'examen de leur expression dans le bourgeon plumaire montre que celle-ci n'est pas liée à la prolifération cellulaire, mais plutôt à la migration et à la différenciation, situation probablement analogue dans les cellules de crête neurale (Desbiens *et al.*, 1991).

Les données présentées ici relatives à l'expression de gènes de la famille *Ets* chez la drosophile, le xénope et le poulet pourraient indiquer que les protéines qu'ils codent jouent un rôle durant les processus morphogénétiques du développement de ces organismes. Mais si leur implication en est clairement attestée pour certains gènes de drosophile dont il existe des mutants phénotypiques, ou chez le xénope grâce aux expériences pratiquées sur des oocytes à l'aide d'oligonucléotides antisens spécifiques, elle l'est beaucoup moins chez le poulet, espèce chez laquelle seule la répartition de l'ARNm a été décrite. Or, l'expression d'un ARNm ne présage en rien de sa traductibilité. Et surtout, il est primordial de garder à l'esprit la multiplicité des protéines qui peuvent être issues d'un seul locus, soit par épissage alternatif, soit par modifications post-traductionnelles. Ces protéines pourraient exercer des effets antagonistes sur la régulation génique (voire pas d'effet du tout), ce qui ne déprécie pas nécessairement leur importance, mais empêche néanmoins de préjuger de la finalité de leur influence. A titre d'exemple, si l'on étudie l'expression globale des transcrits codant les protéines p54^{c-ets-1} et p68^{c-ets-1}, on oblitère totalement les tissus au niveau desquels seul l'un des deux transcrits est présent et donc où l'une des deux protéines n'a aucune probabilité d'être produite.

3. Propriétés oncogéniques :

La présence de l'oncogène *v-ets* au sein du rétrovirus transformant E26 démontre l'implication probable des gènes apparentés dans l'oncogénèse. L'activité de certaines des protéines codées par les gènes *ets* est en effet corrélée à l'état prolifératif des cellules : c'est le cas, on l'a vu, de la protéine Ets-2 dans le foie en régénération ou dans les lymphocytes T activés. Des expériences rudimentaires pratiquées *in vitro*, ont été mises en oeuvre pour démontrer l'oncogénicité de certaines de ces protéines. De plus, *in vivo*, des gènes de la famille *Ets* sont activés dans certaines formes de leucémies murines, ou transloqués dans au moins une tumeur humaine.

3.1. Transformation des cellules NIH 3T3 :

Les cellules de souris NIH 3T3 transfectées par un vecteur plasmidique exprimant les protéines aviaires p54^{c-ets-1} (Seth et Papas, 1990) ou p55^{c-ets-2} (Seth *et al.*, 1989) voient leurs caractéristiques prolifératives modifiées : elles acquièrent la capacité de croître en milieu appauvri en sérum et, parce qu'elles ne sont plus ancrées à leur substrat, forment des colonies en milieu semi-solide. Il apparaît donc que ces deux protéines confèrent aux cellules 3T3 une croissance autonome indépendante de tout contrôle, en interférant probablement avec l'expression physiologique des facteurs de

croissance ou de leur récepteurs, ainsi qu'avec l'expression normale de molécules membranaires dont dépend l'adhésion cellulaire.

Une démonstration explicite du rôle de la protéine Ets-2 humaine dans le contrôle de la prolifération cellulaire est née de l'utilisation d'un mutant protéique ne comportant que le domaine de fixation à l'ADN. Cette étude a permis d'entrevoir un schéma de transduction du signal prolifératif délivré par le CSF-1. En effet, l'action de ce facteur de croissance stimule indirectement les protéines Ras/GAP (Bortner *et al.*, 1991) qui modulent à leur tour l'activité des protéines Raf-1 et MAP (Thomas *et al.*, 1992 ; Wood *et al.*, 1992). Or, il est établi que certaines cibles nucléaires de la voie ras/raf sont constituées par des éléments de réponse au complexe AP-1 et aux protéines Ets (Wasylyk *et al.*, 1989 ; Gutman et Wasylyk, 1991 ; Bruder *et al.*, 1992). La transfection de cellules NIH 3T3 transformées par les oncogènes v-Ras ou c-Fms (récepteur du CSF-1) par une protéine Ets-2 tronquée de ses domaines transactivateurs résulte en la réversion de la réponse proliférative vis-à-vis du CSF-1 (Langer *et al.*, 1992). Ces expériences situent Ets-2 sur la même voie de transduction que c-Myc, dont l'expression est affectée par l'action du mutant dominant négatif ; par contre, cette voie est différente de celle du complexe AP-1 dont l'activité n'est pas modifiée.

3.2. Activation par insertion rétrovirale :

L'érythroleucémie murine de Friend représente un excellent exemple d'activation oncogénique de gènes de la famille Ets par insertion rétrovirale. Le complexe de Friend se compose d'un virus déficient pour la réplication, le SFFV (Spleen Focus-Forming Virus), et d'un virus auxiliaire compétent, le F-MuLV (Friend Murine Leukemia Virus) (pour revue, Ben David et Bernstein, 1991). Ce complexe présente la particularité d'infecter des animaux adultes, tandis que le F-MuLV seul infecte l'animal plus précocement, dès la naissance.

Deux types d'évènements majeurs président à l'établissement de la malignité. Chez l'adulte, il y a d'abord prolifération préleucémique polyclonale de cellules érythroïdes dont l'état de différenciation se situe entre le stade BFU-E tardif et le stade CFU-E. Cette croissance est indépendante de la présence d'érythropoïétine. En fait, il est probable que la glycoprotéine d'enveloppe du SFFV active constitutivement le récepteur de l'érythropoïétine (Yoshimura *et al.*, 1990 ; Hino *et al.*, 1991 ; Zon *et al.*, 1992 ; Wang *et al.*, 1993). Un avantage prolifératif serait ainsi conféré aux cellules infectées et permettrait parmi celles-ci l'émergence d'un clone malin, à condition que l'animal joue de malchance. En effet, une deuxième série d'évènements plus tardifs peut être mise en branle si, par hasard, en s'intégrant dans le génome cellulaire, le virus provoque d'importantes altérations géniques. Parmi celles-ci doivent être citées l'inactivation de la protéine anti-oncogène p53 et l'intégration du SFFV en amont du

gène *spi-1*, dans une orientation inversée (Moreau-Gachelin *et al.*, 1988 ; Moreau-Gachelin *et al.*, 1989 ; Paul *et al.*, 1991). Dans ces cellules, la structure de la protéine est préservée, seule sa surexpression suffit à bloquer la différenciation des progéniteurs érythroïdes (Paul *et al.*, 1991), ainsi que le reconstruit l'infection par un rétrovirus exprimant la protéine PU.1 (Schuetze *et al.*, 1993). De fait, si les mêmes cellules sont contraintes à se différencier sous l'effet d'un agent chimique tel le DMSO, le taux d'expression de la protéine PU.1 décline aussitôt (Schuetze *et al.*, 1992).

Chez l'animal infecté par le complexe de Friend, il n'y a pas d'intégration du SFFV au voisinage du gène *spi-1*. Par contre, le virus auxiliaire F-MuLV s'intègre préférentiellement en amont du gène *fli-1*, en orientation inversée (Ben-David *et al.*, 1990 et 1991). Rappelons que le gène *fli-1* est situé sur le même chromosome que le gène *c-ets-1*, à 240 kbp environ. L'effet enhanceur du à l'activité du LTR du provirus intégré ne semble pas s'appliquer à *c-ets-1* et, en première approximation, seule la surexpression du gène *fli-1* est une étape importante de l'établissement de la malignité. Le maintien en culture d'une lignée variante de cellules érythroleucémiques murines (la lignée ELM) nécessite la présence de cellules nourricières issues de la moelle osseuse. La protéine Erg participe à l'établissement d'une croissance autonome de ces cellules (Nibbs *et al.*, 1993).

Dans aucune des leucémies de Friend développées chez l'animal adulte n'est observée d'intégration de la composante SFFV au voisinage du gène *fli-1*. Les deux virus SFFV et F-MuLV présentent donc probablement une spécificité d'intégration, respectivement en amont du gène *spi-1* et du gène *fli-1*. Il est vrai que parler d'intégration préférentielle peut paraître abusif dans la mesure où il est envisageable que seules soient sélectionnées les cellules présentant ce type d'intégration, les autres cellules infectées étant le siège de recombinaisons dont les conséquences phénotypiques ne sont pas apparentes. Des cellules d'érythroleucémie murine dont la différenciation peut être bloquée par surexpression de plusieurs autres oncogènes comme *c-myb* par exemple (Sheng-Ong, 1990) devraient également être détectées, ce qui n'est pourtant pas le cas. C'est pourquoi une intégration préférentielle des composantes du complexe de Friend au voisinage des gènes *spi-1* et *fli-1* constitue une hypothèse probable, même si, jusqu'à présent, aucune homologie de séquence entre les séquences recombinées n'a été rapportée.

Ces différences d'intégration signifient-elles que les virus se propagent à partir de cellules-cibles distinctes qu'ils bloquent à différents stades de l'hématopoïèse ? On peut supposer que bien que les précurseurs érythroïdes soient infectés par les deux virus simultanément, la conformation de la chromatine variera en fonction du stade de développement (nouveau-né *versus* adulte) et donc que l'accessibilité des gènes *Spi-1* et *fli-1* sera plus ou moins favorisée.

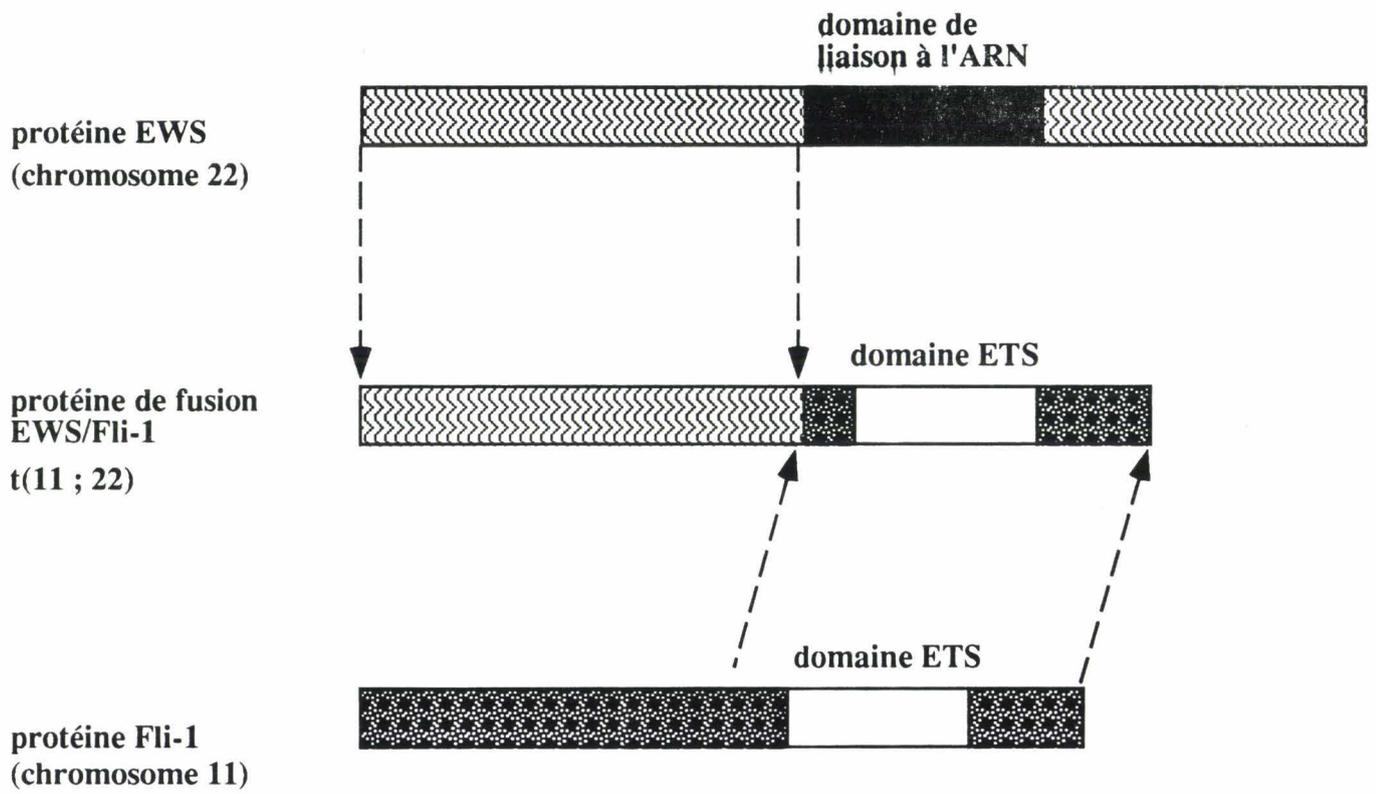


Figure 22 : Représentation schématique des protéines codées par les gènes *EWS* et *fli-1* normaux, et par le gène de fusion *EWS/fli-1*

Une description plus détaillée du phénotype des cellules érythroïdes transformées soit par le SFV soit par le F-MuLV devrait permettre de savoir si les voies de transformation sont analogues. En d'autres termes, les gènes *fli-1* et *spi-1* codant des facteurs de transcription dont les domaines de liaison à l'ADN sont apparentés, ont-ils une activité redondante vis-à-vis des gènes dont ils régulent l'expression, ou bien ont-ils tous deux une activité indépendante qui leur permettra de réguler un ensemble de gènes différents dont l'expression abusive mène à la transformation ? Compte tenu de leur distribution tissulaire respective (voir Tableau 9) et également parce que la spécificité de reconnaissance de l'ADN ne semble pas être un facteur discriminant la fonctionnalité des protéines de la famille Ets, la première hypothèse est la plus probable. En fait, la comparaison de différents promoteurs potentiellement reconnus par PU.1 ou par Fli-1 suggère qu'il pourrait y avoir simultanément régulation individuelle ou chevauchante (Zhang *et al.*, 1993).

Des souris mutantes aux locus *W* et *Sl* sont résistantes au développement d'une érythroleucémie de Friend. Les deux gènes en question codent respectivement le récepteur transmembranaire à activité tyrosine kinase c-Kit et son ligand Steel. Il pourrait être intéressant de savoir si ces deux protéines sont modulées au niveau transcriptionnel par les facteurs PU.1 et Fli-1 au cours de l'hématopoïèse normale et, le cas échéant, si elles le sont au même degré au cours de l'érythroleucémie de Friend.

3.3. Activation par translocation chromosomique :

L'implication de gènes de la famille Ets dans des translocations chromosomiques humaines est le fait d'une seule étude qui porte sur le sarcome d'Ewing, tumeur d'origine neuroectodermique. Les cellules malignes subissent une translocation réciproque des chromosomes 11 et 22 dont résulte la synthèse, à partir du chromosome 22, d'une protéine de fusion qui raboute un domaine aminoterminal homologue à 40 % avec l'extrémité carboxyterminale de l'ARN polymérase II, au domaine ETS de la protéine Fli-1 (Delattre *et al.*, 1992) (Figure 22). Il n'apparaît pas clairement dans la littérature que soit synthétisée, à partir du chromosome 11, une protéine "réciproque" portant le domaine transactivateur de Fli-1 associé à un domaine de liaison à l'ARN. Les deux domaines de la protéines EWS sont indispensables à l'activité transformante de la protéine de fusion (May *et al.*, 1993), probablement parce qu'ils s'associent pour former un facteur de transcription aux propriétés transactivatrices altérées.

Ce type de translocation n'est observé que dans le cas du sarcome d'Ewing. Son analyse précise constitue la première démonstration du rôle oncogénique de gènes de la

famille Ets dans des processus tumoraux qui ne résultent pas d'une intervention rétrovirale et ne touchent pas les cellules hématopoïétiques.

Bien que les régions chromosomiques 11q23 et 21p22 semblent transloquées au cours de différentes leucémies humaines (Diaz *et al.*, 1986 ; Rovigatti *et al.*, 1986 ; Sacchi *et al.*, 1986), la démonstration d'une dérégulation des gènes *ets-1* et *ets-2* respectivement exige encore des analyses plus fines pour leur attribuer clairement un rôle dans le déclenchement de ces leucémies.

Pour conclure, il convient de se demander si les translocations de gènes *ets* que l'on observe (sarcome d'Ewing) ou que l'on croit observer (leucémies) sont véritablement l'élément causal de l'établissement de la maladie, ou si elles ne constituent qu'une manifestation contingente des profonds bouleversements chromosomiques qui touchent toute cellule cancéreuse.

3.4. Implications plus indirectes dans d'autres tumeurs humaines :

Dans des tumeurs humaines malignes ou bénignes, les transcrits *ets1* ne sont pas détectés dans les cellules cancéreuses proprement dites mais plutôt dans les vaisseaux sanguins qui irriguent la tumeur, et dans les fibroblastes stromaux (Wernert *et al.*, 1992). Le gène *c-ets-1* est exprimé dans les vaisseaux en voie de formation de l'embryon de poulet (Vandenbunder *et al.*, 1989). Les vaisseaux sanguins qui irriguent les tissus tumoraux sont eux aussi néoformés puisqu'ils sont recrutés par la tumeur *via* des facteurs angiogéniques, afin de la nourrir et de promouvoir sa croissance.

D'autre part, la détection d'ARNm *c-ets-1* dans les fibroblastes du stroma est une observation intéressante, parce qu'*in vitro*, en association avec le complexe AP-1, la protéine Ets-1 stimule l'expression de protéases dont le rôle consiste à dégrader la matrice extra-cellulaire, comme la collagénase, la stromélysine, l'uPA (Gutman et Wasyluk, 1990 ; Rorth *et al.*, 1990 ; Wasyluk *et al.*, 1991).

Bien qu'il faille démontrer que la protéine c-Ets-1 régule effectivement ces protéases *in vivo*, il est tentant d'imaginer que le blocage de son expression pourrait d'une part permettre la régression des tumeurs par défaut de croissance des vaisseaux sanguins, et d'autre part, empêcher l'invasion tumorale par défaut de dégradation de la matrice extra-cellulaire.

3.5. Hypothèses sur l'oncogénicité de v-Ets :

Si les propriétés oncogéniques du gène *v-ets*, partie intégrante de la protéine P135 Δ gag-myb-ets du virus E26 sont clairement illustrées par la transformation de

précurseurs hématopoiétiques *in vitro* et *in vivo*, les mécanismes moléculaires qui traduisent les altérations qu'il a subies par rapport à son homologue cellulaire sont encore sujets à de multiples spéculations. Il serait peu judicieux d'attribuer aux déterminants *myb* et *ets* une activité autonome quand l'on connaît l'importance de la fusion des deux séquences pour la leucémogénicité (Metz et Graf, 1991b). L'architecture moléculaire de la P135 Δ gag-*myb-ets* est donc celle d'une nouvelle protéine dont on peut entrevoir qu'elle présente simultanément un sous-ensemble des propriétés des parties qui la constituent, mais aussi qu'elle a acquis des propriétés inédites. Cette remarque rappelle à l'évidence les propriétés de la protéine EWS du sarcome d'Ewing.

Deux aspects fonctionnels de la P135 Δ gag-*myb-ets* peuvent être testés *in vitro* : ses propriétés transactivatrices et sa reconnaissance spécifique de l'ADN. Malheureusement, bien qu'il ait été montré *in vivo* que les domaines transactivateurs de Myb et de Ets sont redondants (Metz et Graf, 1991b), les propriétés transactivatrices de la protéine de fusion n'ont pas été explorées plus avant. Les seules études disponibles concernent la protéine v-Ets synthétisée *in vitro* et indiquent que celle-ci transactive les promoteurs des gènes de la collagénase ou de la stromélysine avec une efficacité voisine de celle de la protéine c-Ets-1 (Wasylyk et Wasylyk, 1992). Cet effet est indépendant de la liaison de v-Ets à un site Ets, mais est plutôt relayé par le complexe AP-1, par un mécanisme indirect. Il est probable que l'interface d'interaction mette en jeu le domaine par lequel le complexe AP-1 interagit avec le récepteur au glucocorticoïdes (Jonat *et al.*, 1990 ; Lucibello *et al.*, 1990 ; Schüle *et al.*, 1990 ; Yang-Ye *et al.*, 1990) puisque ce récepteur est antagoniste de l'activation du site AP-1 par v-Ets. Ces résultats suggèrent que la transformation par v-Ets pourrait passer par l'activation inadéquate du complexe AP-1, selon un comportement différent de celui de la protéine normale qui, pour l'instant, ne semble coopérer avec ce complexe que fixée à un site Ets adjacent à un site AP-1. Ils indiquent également que la protéine v-Ets peut transactiver sans reconnaître une séquence spécifique : il est possible que la P135 Δ gag-*myb-ets* possède cette caractéristique et qu'elle puisse de surcroît agir directement *via* une séquence qui lui est propre.

Dans tous les cas, qu'elle agisse directement ou non, l'éventualité selon laquelle, par son haut niveau de production, elle puisse épuiser le stock d'un ou plusieurs facteurs de transcription présents dans la cellule transformée par "squenching" mérite également considération. Cette hypothèse expliquerait qu'un virus exprimant une protéine artificielle gag-Myb-c-Ets-1 soit toujours douée d'une activité transformante importante vis-à-vis des cellules érythroïdes (T. Graf, communication personnelle). Les gènes cibles de la P135 Δ gag-*myb-ets* seraient alors peut-être moins cruciaux pour la transformation que ceux qui ne peuvent être régulés faute de régulateurs adéquats disponibles.

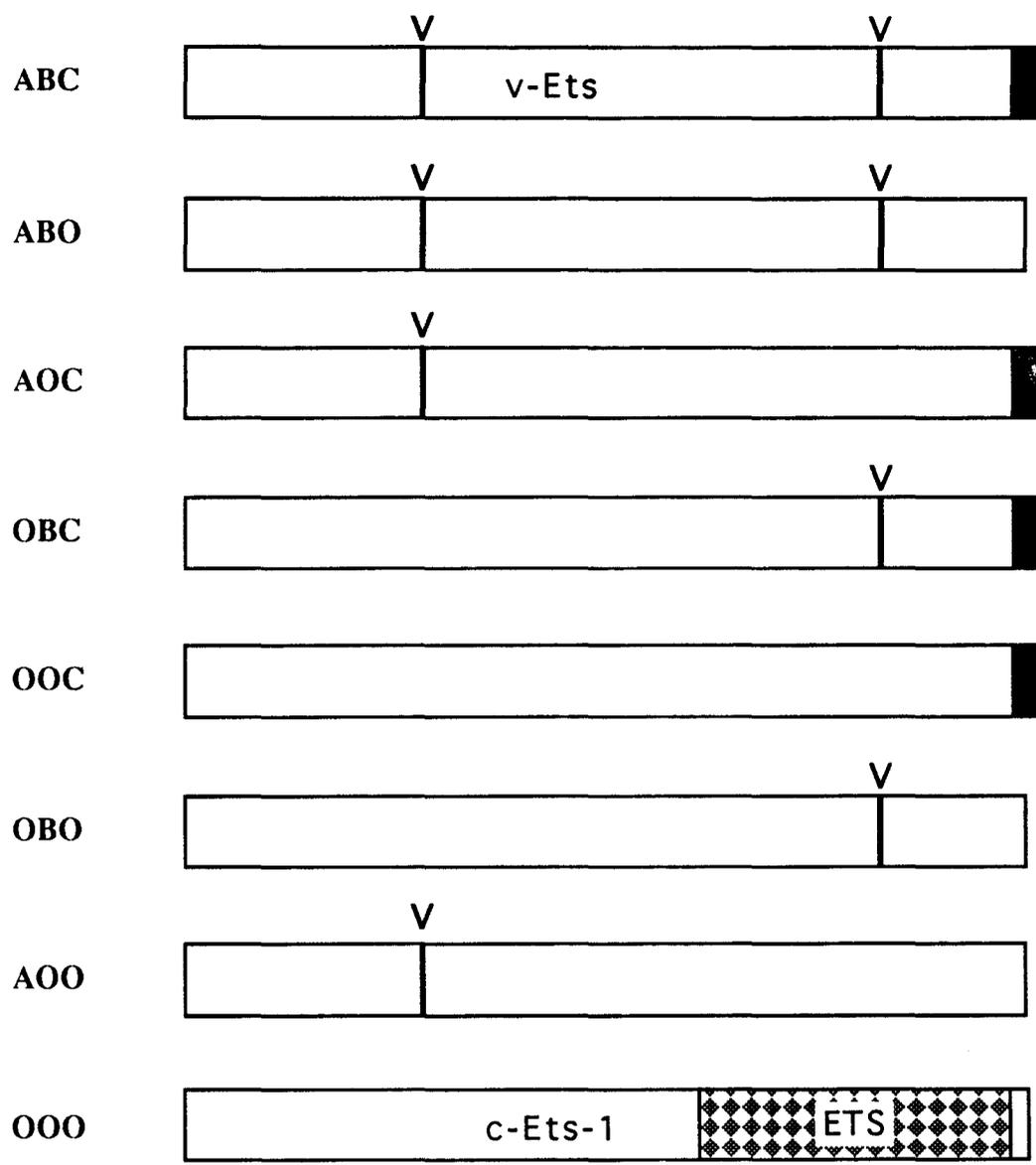


Figure 23 : Représentation schématique des différents mutants de substitution combinant les trois mutations principales observées entre l'oncogène v-Ets et sa contrepartie cellulaire. A = Valine amino-terminale de v-Ets. B = Valine de v-Ets située dans le domaine de fixation à l'ADN. C = extrémité carboxy-terminale spécifique de v-Ets. O = mutation correspondante dans c-Ets.

Les quelques rares données concernant la liaison à l'ADN de la P135 Δ gag-myb-ets permettent de formuler plusieurs hypothèses, qui ne concilient pas les propriétés de la protéine v-Ets et celles de la protéine de fusion. En effet, une protéine v-Ets de 35 kDa comportant toute la partie carboxy-terminale de c-Ets-1 semble incapable de reconnaître le motif PEA3 *in vitro* (Leprince *et al.*, 1992). La réversion successive des deux mutations ponctuelles (une dans le domaine ETS, à proximité de la mutation du virus ts 1.1, l'autre en dehors) et de la queue carboxy-terminale de v-Ets, par les amino-acides correspondants dans la protéine normale (Figure 23) rétablit peu à peu la reconnaissance de cette séquence cible. A partir de ces résultats, il est donc tentant d'imaginer que v-Ets est transformant parce qu'il reconnaît des gènes que c-Ets-1 n'est pas censé réguler dans le cours normal de l'hématopoïèse. Mais un autre mécanisme peut également être proposé, sur la base de l'homologie de séquence entre l'extrémité carboxyterminale de v-Ets avec la région correspondante sur le brin non codant de *c-ets-1* développée dans la deuxième partie de ce texte. C'est en premier lieu cette mutation qui est responsable de la perte de reconnaissance du motif PEA3. *In vivo*, il est possible que non seulement v-Ets reconnaisse une autre séquence cible mais que de surcroît, l'extrémité 3' de l'ARNm qui la code s'hybride avec l'ARNm codant la protéine c-Ets-1 endogène. La formation de ces hybrides, dégradés par les RNases intracytoplasmiques, réduirait considérablement la quantité de protéine Ets-1 dans la cellule.

A défaut de pouvoir tirer profit de mutations naturelles (mutants *pokkuri* et *pointed* de drosophile, chimère EWS du sarcome d'Ewing ou "mariage forcé" entre Myb et Ets-1 dans la P135 Δ gag-myb-ets), il est probable que les perspectives les plus réjouissantes émergeront de l'inactivation de chaque gène de la famille Ets par recombinaison homologue. D'une part, cette approche devrait indiquer le degré de redondance de leur activité. D'autre part, elle pourrait désigner des terrains d'activité physiologique privilégiés de chaque membre ou de plusieurs membres conjointement. La plupart des études entreprises jusqu'à présent, essentiellement basées sur la topographie d'expression des protéines Ets, a néanmoins eu le mérite de fournir des modèles cellulaires ouvrant la voie aux mécanismes moléculaires grâce auxquels l'activité de ces facteurs de transcription peut s'exercer.

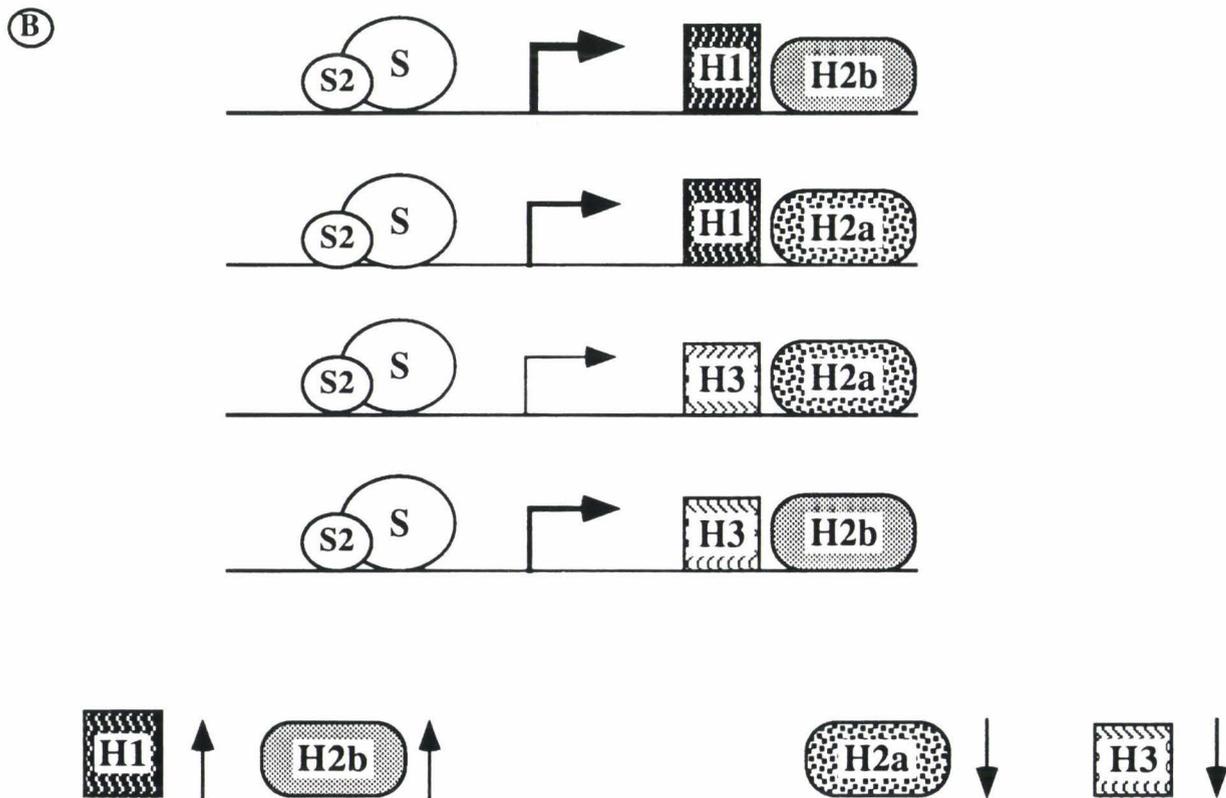
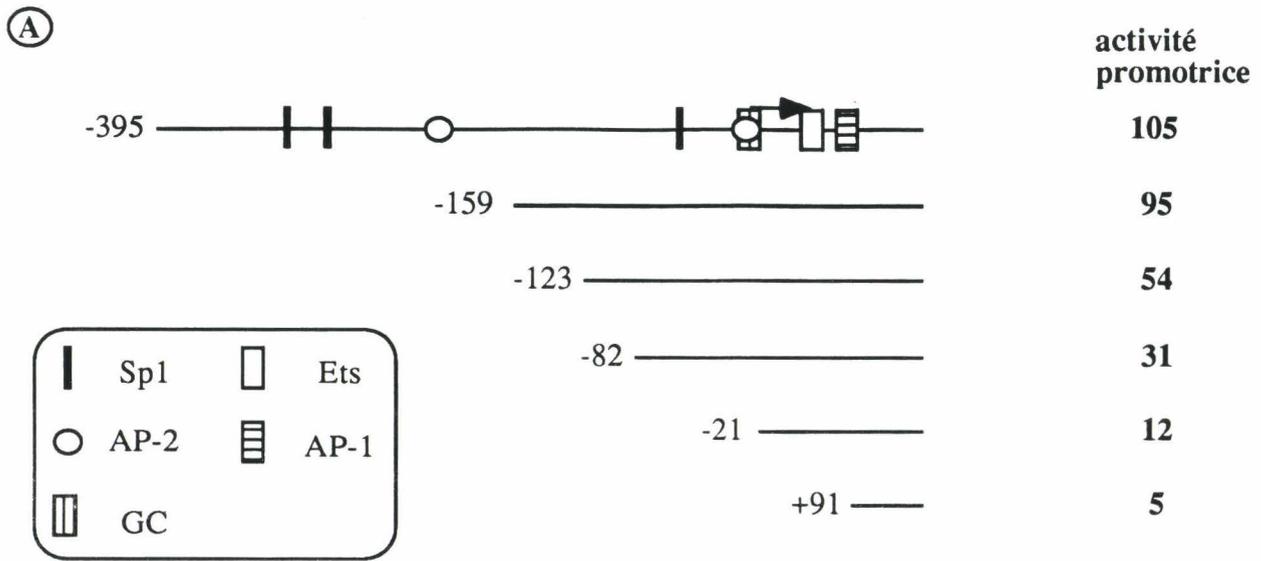


Figure 24 : Régulation transcriptionnelle du gène *ets-2* humain. A : Activité promotrice de différents mutants de délétion. B : Modèle reliant l'intensité de l'expression de l'ARNm *ets-2* à la combinatoire des facteurs disponibles dans les différentes cellules étudiées. (tiré de Mavrothalassitis *et al.*, 1991).

5ème partie : Régulation transcriptionnelle de gènes de la famille Ets :

Compte tenu des implications plus que probables de protéines de la famille Ets dans des processus cellulaires fondamentaux, de multiples aspects de leur régulation sont actuellement à l'étude. Les promoteurs de quelques gènes de la famille ont été clonés, mais dans l'ensemble, les données concernant leur régulation transcriptionnelle sont encore fragmentaires. Ainsi, les promoteurs des gènes *fli-1* et *PU.1* sont isolés, et leur séquence est connue (Moreau-Gachelin *et al.*, 1989 ; Barbeau *et al.*, 1992). Tous deux dirigent la transcription de leur ARNm respectif à partir de multiples sites d'initiation résultant probablement de l'absence de boîte TATA et CAAT en position usuelle chez les Eucaryotes (-30 et -60 environ). La seule région de régulation avérée est constituée d'un enchaînement de bases G+A au sein du promoteur du gène *fli-1* qui semble reconnu par des protéines nucléaires non identifiées extraites de la lignée érythroïde K562 (Barbeau *et al.*, 1993).

1. Promoteur du gène *ets-2* humain :

Il a été indiqué plus haut que le gène *ets-2* s'exprime de façon relativement ubiquiste. Toutefois, selon les tissus, des variations du taux d'ARNm *ets-2* sont observées. Le clonage de son promoteur a permis de circonscrire des régions d'ADN génomiques pouvant éventuellement rendre compte de ces fluctuations quantitatives. Ce promoteur initie la transcription en absence d'élément TATA ou CAAT à partir de plusieurs sites concentrés sur une dizaine de paires de bases (Mavrothalassitis *et al.*, 1990a). L'analyse de constructions plasmidiques incluant des délétions séquentielles en 5' du promoteur transfectées dans des cellules Hela, met en lumière une région d'importance pour l'activité promotrice puisque sa suppression diminue celle-ci considérablement (Figure 24A). Cette région est délimitée par les nucléotides -159 et -21 par rapport au site d'initiation majeur, et elle comprend un site potentiel de fixation pour le facteur Sp1, un autre pour le facteur AP-2, ainsi qu'un site Ets et un site AP-1 situés en aval des sites d'initiation (Mavrothalassitis *et al.*, 1990b). Le TPA est un inducteur connu de l'activité du facteur AP-2 (Imagawa *et al.*, 1987). Néanmoins, il n'induit pas le gène *ets-2* au niveau transcriptionnel dans les cellules Hela (Mavrothalassitis *et al.*, 1990b), bien que ce gène soit induit par les activateurs de la PKC dans les macrophages (Boulukos *et al.*, 1990). Des positions -293 à -21 s'étend une succession de 270 nucléotides G+A qui inclue donc les sites qui viennent d'être décrits, et peut de plus stimuler l'activité d'un promoteur hétérologue.

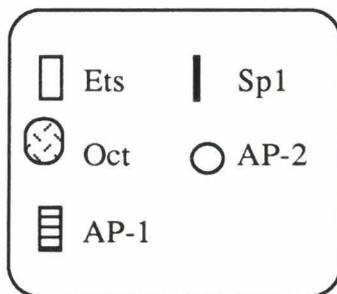
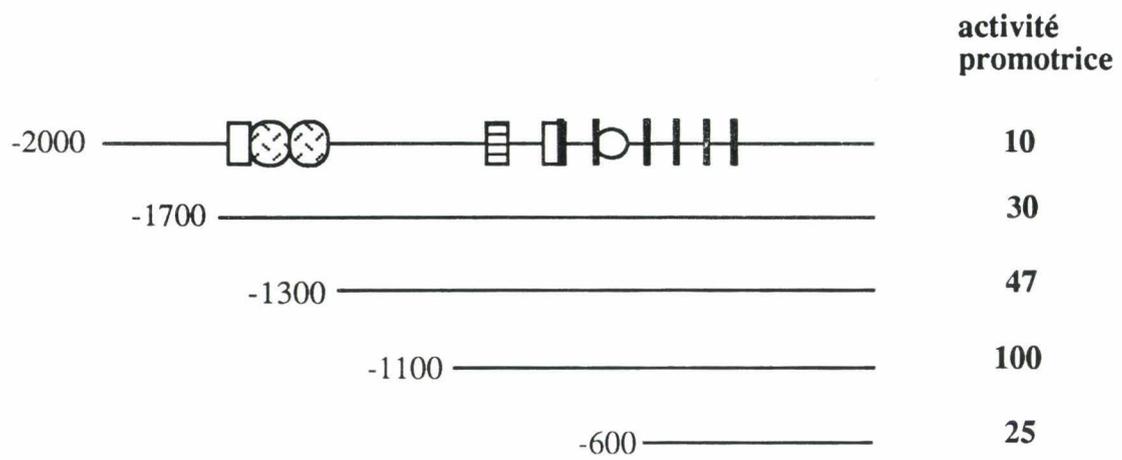


Figure 25 : Promoteur du gène *ets-1* humain : activité promotrice de différents mutants de délétion dans des cellules Daudi.

Des expériences de retard sur gel menées sur des sondes représentatives de la zone de démarrage de la transcription (nucléotides -67 à +55), incluant respectivement un site potentiel de fixation riche en bases G+C, ainsi que les sites AP-1 et Ets révèlent la formation de six complexes différents (Mavrothalassitis *et al.*, 1991). Les résultats obtenus avec des protéines nucléaires extraites de plusieurs types de lignées cellulaires mettent en lumière une corrélation entre le taux d'expression de l'ARNm *ets-2* et la formation préférentielle des différents complexes possibles (Figure 24B). Les complexes notés S, S2, H1 et H2a se forment dans toutes les lignées étudiées, tandis que le complexe H3 est spécifique des cellules hématopoiétiques à l'exception des cellules Jurkat, et que le complexe H2b est de distribution tissulaire encore plus restreinte. Les complexes S et S2 correspondent à la fixation de protéines sur la zone d'initiation riche en bases G+C qui assureraient un taux de transcription basal. Le site AP-1 pourrait interagir avec un facteur inhibiteur (complexe H2a) ou activateur (H2b). De la même façon, le site Ets serait reconnu par une protéine inhibitrice (complexe H3) ou stimulatrice (H1) qui n'est probablement ni Ets-1 ni Ets-2. C'est la combinaison des différentes protéines effectrices sur le promoteur selon leur disponibilité dans la cellule considérée qui conditionnerait le taux de synthèse de l'ARNm *ets-2*, ce qui ne constitue pas une profonde révélation. Quoiqu'il en soit, aucun de ces facteurs n'est identifié précisément jusqu'à présent.

2. Promoteur du gène *ets-1* humain :

Ce promoteur présente quelques analogies avec le promoteur du gène *ets-2*, telles que la présence de multiples sites d'initiation et l'absence de boîte TATA et CAAT (Jorcyk *et al.*, 1991 ; Oka *et al.*, 1991 ; Majérus *et al.*, 1992). Il est actif dans des lignées cellulaires de types différents, même dans la lignée K562 dans laquelle l'ARNm *ets-1* endogène n'est pourtant pas détectable, ce qui suggère d'autres niveaux de contrôle telle qu'une atténuation transcriptionnelle en particulier.

Des délétions séquentielles de l'extrémité 5' du promoteur et des délétions internes révèlent l'alternance de zones régulatrices inhibitrices et activatrices (Chen *et al.*, 1993) (Figure 25). Ainsi, les constructions qui préservent les sites AP-1, AP-2 et Ets localisés respectivement en position -830, -540 et -749 conservent une activité promotrice importante. De fait, des expériences de cotransfection du promoteur *ets-1* avec des vecteurs exprimant soit AP-1, soit AP-2 soit Ets-1 indiquent une stimulation des zones régulatrices par ces facteurs (Oka *et al.*, 1991 ; Majérus *et al.*, 1992). Une partie de ces résultats est en accord avec l'autorégulation positive du gène *ets-1* humain rapportée préalablement (Seth et Papas, 1990). La région qui inclue ces trois sites fonctionnels est inductible par le sérum (Majérus *et al.*, 1992).

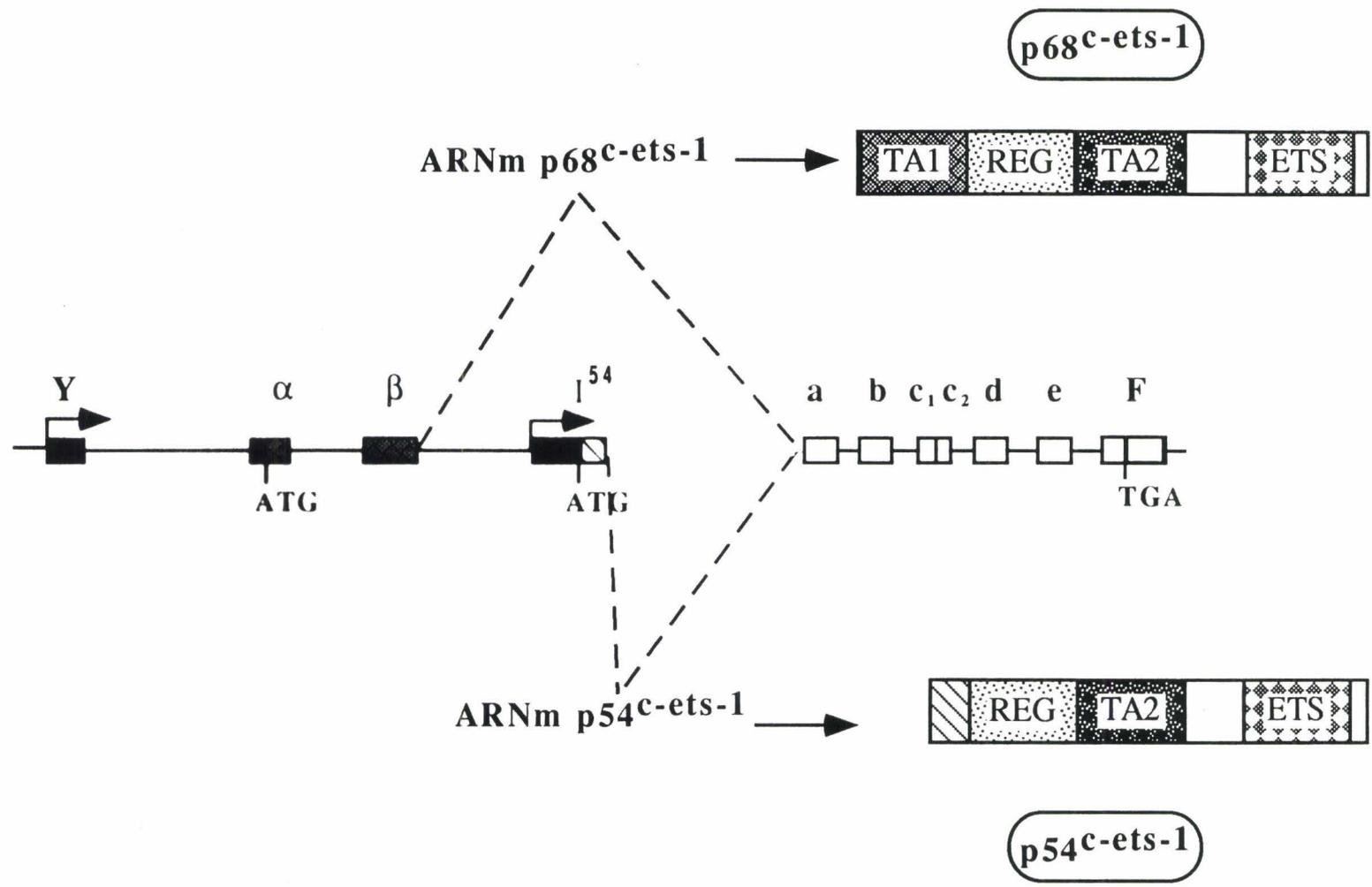


Figure 26 : Organisation génomique et expression du locus *c-ets-1* aviaire. Les séquences non-codantes sont en noir, les exons communs en blanc et les exons spécifiques de chacune des deux protéines sont signalés par des figurés particuliers. La nomenclature des domaines protéiques est la même que sur la figure 13.

Il existe également deux zones inhibitrices de la transcription du gène *ets-1* dont l'une est située en amont du site Ets le plus amont (-1820/-1570), l'autre en aval des sites Oct (-1450/-1150) (Figure 25) qui peuvent également réprimer un promoteur hétérologue (Chen *et al.*, 1993). Par empreinte à la DNase, il a été démontré que ces sites sont effectivement protégés par des protéines PEA3, Oct-1 et Oct-2 produites par synthèse bactérienne (Chen, 1993). Par gel retard, il apparaît que ces deux régions inhibitrices peuvent se complexer à des protéines extraites de différentes lignées hématopoïétiques, sauf de la lignée promyélocytaire HL60 bien que le promoteur *ets-1* n'ait pas d'activité dans ces cellules non plus (Oka *et al.*, 1991).

3. Recherche de zones promotrices dans le locus *c-ets-1* aviaire, membre fondateur de la famille Ets :

La caractérisation moléculaire de l'oncogène *v-ets* a permis l'isolement du gène cellulaire *c-ets-1* dont il est issu. Ce gène s'étend sur 80 kpb chez le poulet (Gégonne *et al.*, 1987) et est morcelé en 10 exons codants (Figure 26). L'isolement d'ADNc colinéaires de ces exons montre qu'au moins deux types différents d'ARNm *c-ets-1* sont exprimés à partir de ce locus. Les extrémités 5' sont spécifiques de chacun d'entre eux, mais ils ont en commun les séquences codées par les exons α à F (I à IX selon la nomenclature américaine) (Gégonne *et al.*, 1987 ; Duterque-Coquillaud *et al.*, 1988 ; Leprince *et al.*, 1988). L'ARNm qui débute par les exons α et β code la protéine p68^{c-ets-1} (485 amino-acides) et il correspond aux séquences progénitrices de l'oncogène *v-ets* ; celui qui inclue l'exon I⁵⁴ code la protéine p54^{c-ets-1} (441 amino-acides). Ces ARNm ont tous deux une taille de 7.5 kb. Cependant, les séquences codant la p54^{c-ets-1} sont également retrouvées dans deux petits ARNm de 2.2 et 2.0 kb. L'existence d'un épissage alternatif des exons internes tel qu'il est observé chez l'homme (Reddy et Rao, 1988 ; Koizumi *et al.*, 1990 ; Jorcyk *et al.*, 1991) n'est pas connue chez le poulet.

Il a été évoqué dans la troisième partie de ce mémoire le fait que les deux protéines issues du locus *c-ets-1* diffèrent par la présence d'un domaine transactivateur supplémentaire codé par les exons α et β spécifiques de la p68^{c-ets-1} qui expliquerait les meilleures performances transactivatrices de cette protéine (Figure 26). Or, la p68^{c-ets-1} ne semble pas conservée chez les Mammifères (Watson *et al.*, 1988 ; Albagli *et al.*, 1992), mais par contre, les exons α et β sont détectés jusqu'au sein du génome des Reptiles (Albagli *et al.*, 1992). L'évolution n'a conservé que le transactivateur le plus faible, la p54^{c-ets-1}, puisque seul l'exon I⁵⁴ est hautement conservé, des Amphibiens aux Mammifères (Figure 27).

α

Poulet	MMSYYMDTTIGSTGPYPLARPGVM
Caille (30)	*****
Canard (80)	*****A***A*****
Tortue (250)	*****A***PA*****

β

Poulet	QGASSCCEDPWMPCLQSACPPRSCPPWDEAAIQEVPTGLEHYST
Caille	***A*****L**T*****S*****G*
Canard	***V*S*****T*GF***YQ**A*Y*L***S*****

Figure 27: Conservation phylogénétique des exons α et β spécifiques de la protéine p68^{c-ets-1} dans différentes espèces. Les distances temporelles par rapport au poulet sont indiquées entre parenthèses, en millions d'années.

Les travaux qui vont être décrits ici tirent leur origine des différences et des analogies qui caractérisent l'expression respective des deux types de transcrits issus du locus *c-ets-1* aviaire. La description des conditions dans lesquelles leur activité peut être redondante, voire additive, mais aussi antagoniste, passe, au moins en partie, par l'étude de leur régulation transcriptionnelle. A cette fin, il s'est avéré important d'entreprendre une étude moléculaire portant sur l'isolement et la caractérisation du ou des promoteur(s) impliqué(s).

RÉSULTATS

Lorsque cette étude fut initiée, deux types d'ADNc *c-ets-1* avaient été isolés de banques de poulet. Par la présence d'un grand cadre de lecture ouvert, leur séquence devait indiquer qu'ils couvraient toute la partie codante des ARNm *c-ets-1* (Duterque-Coquillaud *et al.*, 1988 ; Leprince *et al.*, 1988). Cependant, leur taille n'excédait pas 1.5 kbp, alors que celle de chaque ARNm *c-ets-1* est de 7.5 kb. C'est dire que de longues régions non codantes (environ 6 kb) restaient encore à découvrir et que l'organisation génomique de ces zones non traduites devaient encore être précisée.

L'ADNc codant la protéine p54^{c-ets-1} était colinéaire de l'ADN génomique cloné à l'époque, c'est-à-dire sur une centaine de paires de bases en amont de l'ATG initiateur. Par contre, celui qui code la protéine p68^{c-ets-1} présentait 69 nucléotides en 5' du codon initiateur de la traduction situé dans l'exon α , dont seulement 11 étaient colinéaires à l'ADN génomique en amont de cet exon. Ces 11 nucléotides représentaient l'extrémité 5' non codante de l'exon α , tandis que les 58 autres suggéraient l'existence d'au moins un exon 5' non-codant.

Ces données nous ont servi de base pour caractériser deux zones promotrices du locus *c-ets-1* dont les propriétés sont détaillées dans l'article ci-joint. Voici un résumé des principaux apports qui s'en dégagent.

1. Caractérisation des sites d'initiation de la transcription des ARNm *c-ets-1* :

La caractérisation des extrémités 5' des ARNm codant les deux produits majeurs issus du gène *c-ets-1* de poulet, par les techniques d'extension d'amorce et de protection à la RNase, a permis d'aboutir aux conclusions suivantes :

- la transcription des ARNm codant les protéines p54^{c-ets-1} et p68^{c-ets-1} s'initie dans les deux cas en de multiples sites. Le fait que les deux types d'extrémités 5' non codantes ne présentent aucune communauté structurale indique l'existence d'au moins deux promoteurs différents.

- l'ARNm codant la p54^{c-ets-1} s'initie dans une quarantaine de pb situées à 240 pb maximum en 5' de l'ATG initiateur localisé dans l'exon I⁵⁴. Cette partie non codante ne représente pas un exon supplémentaire mais prolonge plutôt l'exon I⁵⁴ en 5', puisqu'elle est colinéaire à l'ADN génomique.

- la transcription de l'ARNm codant la p68^{c-ets-1} débute sur plusieurs sites localisés 125 pb maximum en amont de l'ATG initiateur inclus dans l'exon α . Mais cette région comprend 11 nucléotides non codants qui prolongent l'exon α en 5', et 114 nucléotides qui signent l'existence d'un exon intégralement non codant. Cet exon,

baptisé "Y", a été retrouvé dans l'ADN génomique à une dizaine de kpb en amont de l'exon α .

- le fait que les deux types d'ARNm présentent des régions 5' non codantes relativement courtes indique que la quasi-totalité des 6 kpb d'information non codante est distribuée dans la partie 3' du gène *c-ets-1*.

2. Isolement de deux promoteurs du gène *c-ets-1* de poulet :

L'ADN génomique situé en amont des sites d'initiation des ARNm p54^{c-ets-1} et p68^{c-ets-1} a été isolé et ses propriétés promotrices étudiées. Cette analyse comporte plusieurs aspects intéressants :

- la séquence des deux promoteurs révèlent l'absence de boîte TATA ou CAAT.

- l'ARNm codant la p54^{c-ets-1} s'initie dans une région riche en bases G et C. Sur 0.7 kpb, son organisation est similaire à celle qui a été décrite par plusieurs équipes chez l'homme (voir 5ème partie de ce mémoire). En particulier, de nombreux sites AP-2 et Sp1 sont retrouvés, ainsi qu'une région exclusivement composée de nucléotides G et A, située en aval des sites d'initiation de la transcription.

- l'activité de ce promoteur est équivalente à 2 à 3 fois celle du promoteur minimal du gène de la thymidine kinase, lorsqu'il est placé en amont d'un gène indicateur (ADNc de la luciférase) et transfecté dans des fibroblastes embryonnaires de poulet (FEP).

- le promoteur initiant la synthèse de l'ARNm codant la p68^{c-ets-1} n'est pas particulièrement riche en bases G et C, bien qu'il ne comporte pas non plus de boîte TATA ou CAAT. Par contre, la transcription est initiée au sein d'une dizaine de bases C/T.

La séquence de 1.5 kpb en amont des sites d'initiation révèle également la présence de plusieurs sites potentiels de reconnaissance d'un certain nombre de facteurs de transcription spécifiques des cellules lymphoïdes, tels que Elf-1, PU.1, Ets-1, TCF-1 et GATA-3.

- l'activité promotrice d'un fragment de 1.1 kpb, étudiée dans des FEP, est faible. L'analyse de différents fragments du promoteur délétés de portions 5' variables indique que le promoteur minimal comprend 90 pb environ en 5' du site d'initiation le plus distal.

3. Autorégulation positive du promoteur p68^{c-ets-1} :

Par co-transfection d'un vecteur exprimant la protéine p68^{c-ets-1} et d'un fragment de 1.1 kpb de son promoteur dans des FEP, nous montrons que ce facteur peut exercer un effet régulateur positif sur sa propre expression, d'une manière dose-dépendante. Deux sites Ets situés en positions -666 et -225 pourraient permettre la médiation de cet effet.

The two functionally distinct amino termini of chicken *c-ets-1* products arise from alternative promoter usage

Pascale Crepieux, Dominique Leprince, Anne Flourens, Olivier Albagli, Elisabeth Ferreira, and Dominique Stéhelin

Unité d'Oncologie Moléculaire, CNRS URA 1160, Institut Pasteur, Lille, France

The chicken *c-ets-1* locus gives rise to two distinct transcription factors differing by structurally and functionally unrelated N-termini. p54^{c-ets-1} shows a striking phylogenetic conservation from *Xenopus* to humans, while p68^{c-ets-1}, the cellular counterpart of the E26-derived *v-ets* oncogene, is apparently restricted to avian and reptilian species. In the chick embryo, both mRNAs are expressed in a wide array of tissues of mesodermal origin; however, in the embryo and after hatching, p68^{c-ets-1} is excluded from lymphoid cells where p54^{c-ets-1} accumulates.

In this report, we define the basis of the differential expression of the chicken *c-ets-1* products to assess their different potentials as transcription factors. We demonstrate that the two distinct N-termini arise from alternative promoter usage within the chicken *c-ets-1* locus. Examination of both promoters reveals that transcription initiates from multiple sites, consistent with the absence of TATA and CAAT elements. Of these two regulatory regions, only the one that initiates the p54^{c-ets-1} mRNA synthesis is of the G+C-rich type, and its organization is conserved in humans. The avian-specific p68^{c-ets-1} promoter activity was enhanced by its own product. In addition, we identify numerous potential binding sites for lymphoid-specific transcription factors that might contribute to a tight repressor effect in lymphoid tissues.

The *ets* gene family groups a number of related transcription factors that are conserved from *Drosophila* to humans (Lautenberger et al., 1992; MacLeod et al., 1992; Laudet et al., 1993). This family is defined by the presence of a new type of DNA-binding domain, the ETS domain (Karim et al., 1990), most often located at the carboxy-terminus of the protein with some exceptions, including the elk-1 (Rao et al., 1989), elf-1 (Thompson et al., 1992), and SAP-1 (Dalton and Treisman, 1992) proteins. The *ets* family members exert their function as transcription factors by interacting in a sequence-

specific manner with purine-rich motifs they recognize with variable affinities (Wang et al., 1992) in the promoters and enhancers of several viral and cellular genes (Gunther et al., 1990; Ho et al., 1990; Wasylyk et al., 1990; Virbasius and Scarpulla, 1991; Wasylyk et al., 1991). Nevertheless, a dual function for *ets* family members has recently come to light from the observation that some of them can form ternary complexes, in association with unrelated transcriptional factors such as the SRF (Hipskind et al., 1991; Dalton and Treisman, 1992).

The roles of *ets*-related protein regions out-

Received May 7, 1993; revision accepted July 7, 1993.

Correspondence: Dominique Leprince, CNRS URA 1160, Institut Pasteur, 1 rue Calmette, 59019 Lille, France Tel 33-20-87-78-79 Fax 33-20-87-79-08

© 1993 by the University of Health Sciences/The Chicago Medical School. All rights reserved. 1052-2166/93/0302/215-11\$2.00

side of the DNA-binding domain still remain largely unclear, though they were partly elicited by recent studies on the avian proteins produced by the *c-ets-1* gene (Schneikert et al., 1992). The chicken *c-ets-1* locus, identified as the cellular counterpart of the E26 *v-ets* oncogene (Leprince et al., 1983; Nunn et al., 1983), was the first characterized member of the *ets* gene family. This locus gives rise to two different transcription factors, p54^{c-ets-1} and p68^{c-ets-1}, which differ only by unrelated N-termini, respectively encoded by a single exon absent from *v-ets*, I⁵⁴, or by the α and β exons, homologous to the 5' part of *v-ets* (Leprince et al., 1988). These alternative exons are fused to a common set of 3' exons named A to F. The common exons encode an N-terminal regulatory domain, a transactivating domain, and the C-terminal DNA-binding domain. While the highly hydrophobic α - and β -encoded amino acids consist of an additional transactivating domain, the function of the hydrophilic I⁵⁴-encoded amino acids remains undefined (Schneikert et al., 1992). We can therefore speculate that each N-terminus interacts with different transcription regulators to fulfill its function, possibly in a cell-specific manner. Consistent with this hypothesis is the fact that p54^{c-ets-1} and p68^{c-ets-1} display differences in their expression pattern: in chicken, p54^{c-ets-1} is widely expressed, with moderate levels in most tissues, but high levels in lymphoid cells (Ghysdael et al., 1986). In contrast, p68^{c-ets-1} expression is restricted to a blood vessel-containing fraction of the spleen (Leprince et al., 1990) and other mesodermal tissues such as the embryonic dermis at E6, but remains undetectable in lymphoid tissues (Quéva et al., 1993).

A better understanding of the positive and negative controls that underlie a cell-specific pattern of gene expression requires careful examination of the mechanism regulating the expression of the regulators themselves. In this report, we have initiated such studies by describing the molecular mechanism governing the expression of the chicken *c-ets-1* locus. To unravel the respective regulation of p54^{c-ets-1} and p68^{c-ets-1} expression, we asked if they arise from alternative splicing of a precursor mRNA initiated at a single promoter, or whether a differential promoter usage would account for the divergences in their expression patterns. The study reported here ascertains the second mechanism.

Materials and methods

Molecular cloning

Isolation of a new chicken *c-ets-1* genomic clone. A recombinant DNA library in the EMBL4 vector constructed with a partial *Sau* 3A digest of total chicken embryo DNA was used to isolate the promoter region of p68^{c-ets-1} mRNA. It was screened first with a 5.4 kbp *Hind* III genomic probe including the α exon (Gégonne et al., 1987). The positive clones were then counter-selected with a labeled oligonucleotide corresponding to the 5' end of the p68^{c-ets-1} cDNA (primer 1: 5' ACA AGT GTG GGG AGC CGT GGA GGA 3'). We obtained a 14.0 kbp long genomic insert, which is referred to as clone 68.1ETS. The previously described lambda *c-ets* B recombinant phage (Gégonne et al., 1987) containing sequences localized within the 40.0 kbp of genomic DNA including the I⁵⁴ exon was subjected to promoter analysis after restriction enzymes mapping.

Reporter and expression plasmids used in the promoters functional assays. The *c-ets-1* promoter regions and the p68^{c-ets-1} promoter deletion mutants were subcloned into the pLUCDSS luciferase reporter plasmid, generously provided by Dr. F. Gouilleux (Gouilleux et al., 1991), who derived it from pSPLUC2 (DeWet et al., 1987). The fragments were inserted either into the *Pvu* II-*Hind* III (p68^{c-ets-1}), or *Pvu* II (p54^{c-ets-1}) restriction sites. In order to allow in-frame synthesis of the luciferase protein during the functional assay, the p54^{c-ets-1} promoter fragment subcloned into the M13 mp19 vector was mutated in the translation initiation codon (ATG→AAG) according to standard methods (Sayers et al., 1988) using the Amersham oligonucleotide in vitro mutagenesis kit. The ptkFLUC vector (DeWet et al., 1987; Gouilleux et al., 1991) including the thymidine kinase minimal promoter was used as a positive control. The chicken p68^{c-ets-1} cDNA (Leprince et al., 1988) was cloned into the *Eco*R I restriction site of the pSG5 expression vector (Wasylyk et al., 1990).

Sequencing analysis. The *c-ets-1* regulatory regions were sequenced on both strands by the dideoxy termination method (Sanger et al., 1977), after cloning into the polylinker restriction sites of M13mp18 and mp19 derivatives.

Mapping of the *c-ets-1* mRNA 5' ends

RNase protection analysis was performed as previously described (Leprince et al., 1988) with the use of a Promega kit. A 400 bp $\alpha^{32}\text{P}$ -CTP labeled p68^{c-ets-1} RNA probe was transcribed from a corresponding BamH I–Hind III DNA fragment cloned into the pSP64 vector (Fig. 4). The probe was hybridized at 50°C to 20 μg of total RNA from chicken spleen, thymus, or embryonic dermis extracted using the CsCl/guanidinium isothiocyanate method (Maniatis et al., 1982). The RNA–RNA hybrids were digested by a mixture of T1 (Gibco BRL) and A (Boehringer Mannheim) RNases, and the resulting products were electrophoresed on a 6% sequencing acrylamide gel. Their size was determined by comparison with an M13 sequence ladder.

Primer extension experiments were completed as described (Maniatis et al., 1982). The following 5' end-labeled oligonucleotide, complementary to the p54^{c-ets-1} mRNA, was used as a primer:

primer 2 : 5' TCC TCT TCC TCC TCC TTC CTT CCT CGC TCT 3'

This primer was annealed at 60°C, and retrotranscription was completed using MMLV RTase (Gibco BRL). The elongation products were run on a 6% sequencing gel.

DNA transfection

Chicken embryonic fibroblasts (CEF) were plated at 10^6 cells per 60 mm plate the day before transfection. Transfections were achieved using DOTAP (Boehringer Mannheim). Forty-eight hours after transfection with 10 μg of DNA, the cells were washed three times and scraped into PBS. The cells were pelleted and resuspended in a lysis buffer composed of 25 mM Tris phosphate (pH 7.8), 8 mM MgCl_2 , 1 mM DTT, Triton X-100 (1% v/v), and glycerol (15% v/v). After 3 freeze-thawing cycles, the cellular debris was pelleted by centrifugation, and the supernatants were recovered for the luciferase assay.

Luciferase activity assay

In the standard assay, 25 μl of 1 mM luciferin was added to a cell lysate containing 20 μg of proteins in lysis buffer. Ten μl of 10 mM ATP were injected, thus triggering the light output of the tube over a 30 second interval. The Bio-Orbit 1251 Luminometer background was typically in the range of 20 light units. The values obtained were normalized to the level of hu-

man growth hormone (hGH) released in the media of transfected cells resulting from cotransfection of 1.5 μg of ptkGH expression plasmid (Selden et al., 1986).

Results

Conservation of the organization of the p54^{c-ets-1} promoter between humans and chicken

The avian *c-ets-1* gene gives rise to two different types of mRNAs (Duterque-Coquillaud et al., 1988; Leprince et al., 1988), as summarized in Figure 1. The p54^{c-ets-1} mRNA joins the I⁵⁴ exon to the common set of 3' exons, namely a to F, while the p68^{c-ets-1} mRNA results from the fusion of the Y (see below), α , and β exons to the a to F exons. Given that both types of mRNA display differential expression (Leprince et al., 1988; Quéva et al., 1993), we suspected that they might be subject to independent regulation, possibly via tissue-specific alternative splicing and/or differentially regulated specific promoters.

Our first goal was to compare the transcription regulation region of the p54^{c-ets-1} mRNA between humans and chicken. We carried out primer extension analysis on total chicken RNAs extracted from spleen, thymus, and the RP9 B-lymphoma cell line. Primer 2 (see Materials and Methods and Fig. 3) initiates retrotranscription 121 bp 5' of the ATG codon located in the I⁵⁴ exon. This oligonucleotide is complementary to region +121/+150 in our sequence (Fig. 3), close to the 5' end of the p54^{c-ets-1} cDNA (Duterque-Coquillaud et al., 1988). In each tissue tested, three major bands of 150, 116, and 110 bp were detected (Fig. 2), thus localizing the major start sites 201 to 241 bp 5' from the translation start codon found in the I⁵⁴ exon (Fig. 3). Minor bands were also observed, indicative of additional minor cap sites.

To determine whether these 5' ends were derived from I⁵⁴ or from one or more upstream exon(s), we next performed RNase protection analysis. A 14 kbp genomic DNA fragment containing the I⁵⁴ exon has been previously isolated (Gégonne et al., 1987; Duterque-Coquillaud et al., 1988). A 732 bp Hind III–Cla I fragment derived from this genomic DNA and including the 5' half of I⁵⁴ was transcribed into a cRNA probe suitable for hybridization with total RNAs from spleen and thymus. The sizes of the pro-

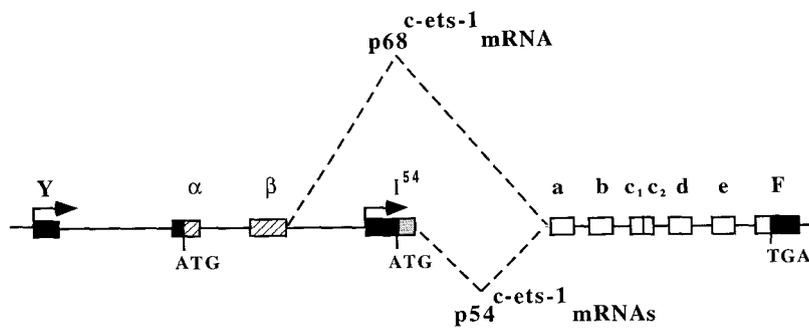


Figure 1. Genomic organization of the chicken *c-ets-1* locus. Open boxes mark the common exons, namely a to F; the alternative exons are indicated by striped (α and β) or dashed (I^{54}) boxes, and the non-coding sequences are boxed in black. The two types of mRNA, which differ in the 5' part of the gene, are shown. The Y exon and the localization of the promoters are described in this paper.

tected bands obtained confirmed the multiplicity of the start sites, along with their positions, and identified the I^{54} exon as the first exon of the $p54^{c-ets-1}$ mRNAs (data not shown). Clustered initiation start sites have also been described in the case of the human I^{54} exon, although there are some discrepancies about their

positions (Jorcyk et al., 1991; Oka et al., 1991; Majérus et al., 1992).

The promoter activity of the 732 bp Hind III–Cla I genomic sequence was assayed by transfection into CEF. For this purpose, the DNA fragment was subcloned after mutagenesis of the $p54^{c-ets-1}$ initiation codon (see Fig. 3) in both orientations into the pLUCDSS vector, in front of the coding sequence for the firefly luciferase cDNA (Table 1). LUCDSS and ptkFLUC plasmids were also transfected as controls. The RLU (relative light units) value obtained when the 732 bp Hind III–Cla I fragment is assayed in

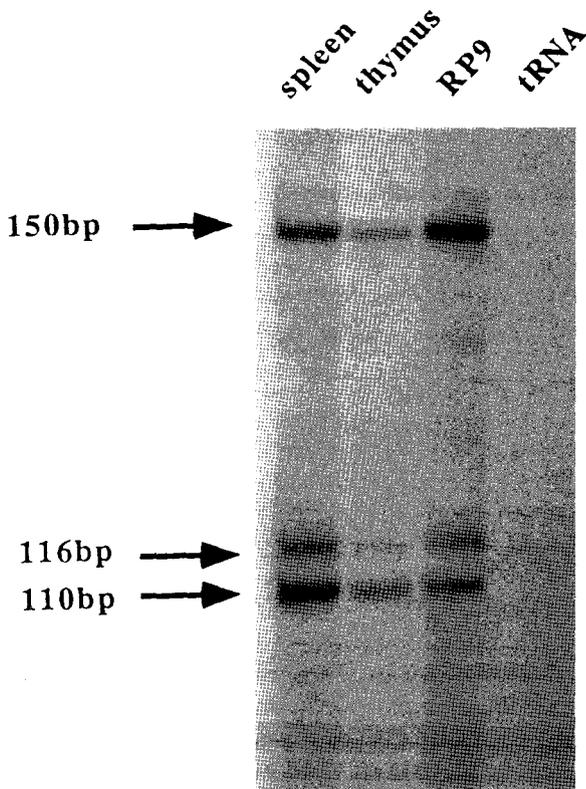


Figure 2. Mapping of the $p54^{c-ets-1}$ transcription start sites by primer extension analysis on total RNAs extracted from chicken spleen, thymus, or the RP9 B cell line. No signal is detected in the yeast tRNA control lane. The three major start sites map 241, 207, and 201 bp upstream from the I^{54} initiation codon.

Table 1. Functional activity of the two avian *c-ets-1* promoters in CEF (chicken embryo fibroblasts).

The DNA fragments using the nomenclature defined in Figures 3 and 5B were cloned in the sense orientation unless the AS letters are added. The pLUCDSS vector has no intrinsic promoter activity. The ptkFLUC vector is used as a positive control. Results are normalized for transfection efficiency and for protein amount. They represent mean values obtained from several independent experiments and are expressed as a percentage of the thymidine kinase promoter activity. Standard deviations are indicated.

Construct	Activity in CEF
p54 ^{c-ets-1}	
-732	258 ± 56
-732AS	6.2 ± 3
p68 ^{c-ets-1}	
-1123	60 ± 13
-533	62 ± 17
-322	83 ± 24
-90	47 ± 20
-28	6 ± 2
+12	12 ± 7
LUCDSS	10 ± 2.5
ptkFLUC	100

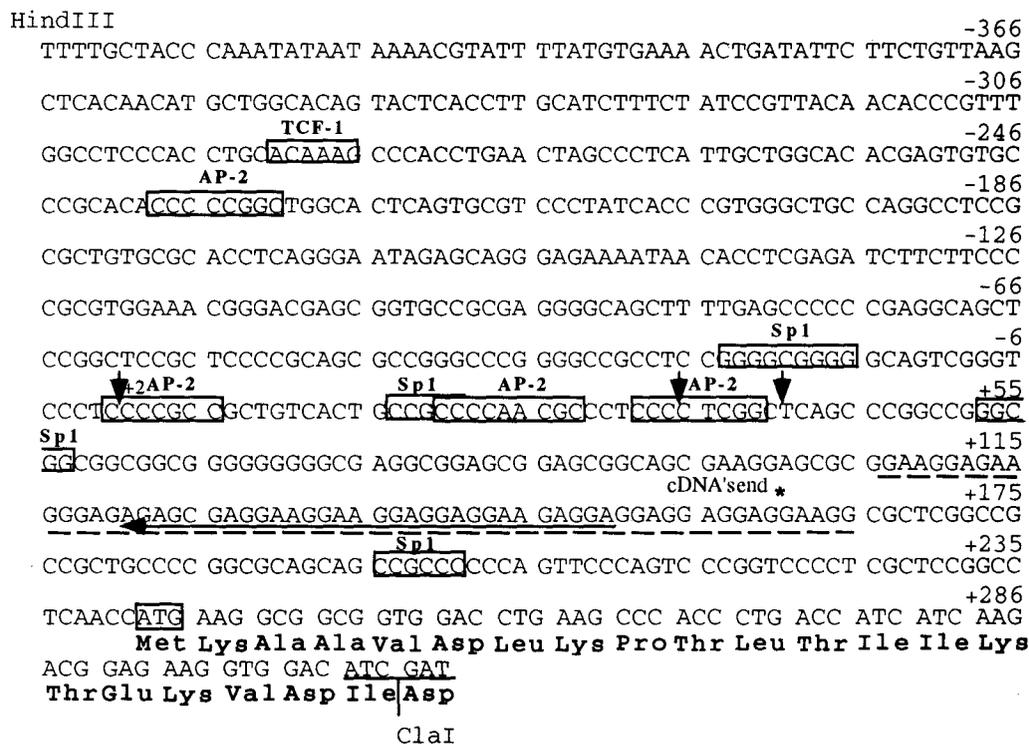


Figure 3. Nucleotide sequence of the 732 bp Hind III–Cla I DNA fragment. The mRNA major cap sites are indicated by vertical arrows, while the p54^{c-ets-1} cDNA 5' end is noted with an asterisk. The G+A stretch is underlined with dashes, and the translation initiation codon of p54^{c-ets-1} is boxed. The position of primer 2 is indicated by a horizontal arrow. Putative recognition sequences for transcription factors are boxed. The bases are numbered with respect to the most upstream start site.

the sense orientation is equivalent to more than twice the thymidine kinase minimal promoter activity, thus confirming that we have indeed cloned a promoter. Neither the insert in an anti-sense orientation nor the LUCDSS vector displays significant promoter activity in this assay.

This 732 bp Hind III–Cla I fragment was completely sequenced (Fig. 3). Similarly to its human homologue, the avian promoter lacks a TATA box or a CAAT box; rather, it has a very high G+C base pair content (more than 65% among the 430 bp located 5' from the ATG codon). As a consequence of the G+C abundance, four Sp1 (GGCG/TG/AG/A; Evans et al., 1988) and four AP-2 (CCNCNCG/CG/CG/C; Imagawa et al., 1987) cognate motifs are found in the vicinity of the cap sites, at positions –24, +17, +53, +196 and –238, –1, +20, +32, respectively. Interestingly, AP-2 has been shown to activate the human promoter in a transient transfection assay performed in the Daudi cell

line (Oka et al., 1991). Another potential binding site seems to be found exclusively in chicken: a recognition sequence at position –291 for TCF-1, an HMG-box protein specifically expressed in T-cells (Waterman et al., 1991). Within the +106/+165 region, the sequence is composed exclusively of G+A base pairs, similarly to the region 1604/1723 bp described in the human promoter (Jorcyk et al., 1991). Purine stretches have been proposed to form triple helix DNA structures (Hanvey et al., 1988).

The p54^{c-ets-1} mRNA is initiated about 10.0 kbp downstream of the p68^{c-ets-1}-specific α and β exons group (Fig. 1). Turning to the complete organization of the chicken *c-ets-1* locus, we could then conclude that the mRNA encoding p68^{c-ets-1} initiates its synthesis from a second promoter just upstream from the already identified α and β exons, or one or more 5' non-coding exon(s). Our study then focused on characterizing this putative second promoter region.

Mapping of initiation sites of the p68^{c-ets-1} transcript and molecular cloning of the 5'-flanking region encompassing the p68^{c-ets-1} first exon

Cloning of the ca. 1.5 kbp p68^{c-ets-1} cDNA, isolated from a chicken spleen cell library, has been previously reported (Leprince et al., 1988). It includes 59 nucleotides upstream from the 5' boundary of the α exon, probably corresponding to at least one unidentified noncoding exon. This putative exon was named Y. Primer extension analysis performed from an oligonucleotide overlapping this putative exon and the α exon 5' boundary suggested that the cDNA previously isolated was full-length in 5', since about 70 bp were extended by this method (data not shown). This result further suggested that sequences just upstream of the Y exon might include the putative p68^{c-ets-1} mRNA regulatory region. By differential screening of a total chicken embryo DNA library with a 5.4 kbp Hind III genomic probe encompassing the α exon, and with an oligonucleotide colinear to the 5' end of the cDNA, we isolated genomic DNA including the Y exon. Restriction map analysis showed its location about 10.0 kbp upstream from the α exon (Figs. 1 and 4). Its 3' boundary was sequenced and displayed typical

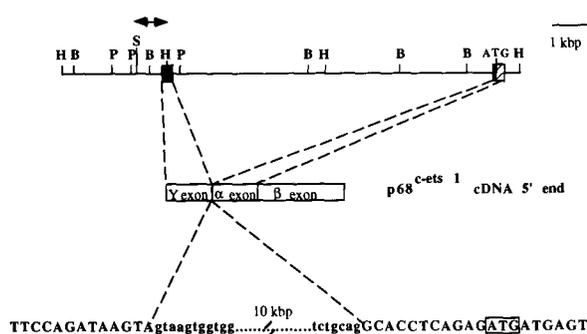


Figure 4. Genomic organization of the 5' part of the chicken *c-ets-1* locus. Relevant restriction sites are indicated as follows: B = BamH I; H = Hind III; P = Pst I; S = Ssp I. The Y and α exons are marked by boxes; the coding part of the α exon is striped. A double-headed arrow covers the 1204 bp DNA fragment, which was cloned into the pLUCDSS vector for testing promoter activity (fragment -1123). The three most 5' exons included in the p68^{c-ets-1} cDNA are shown below. At the bottom, the sequence of the Y and α exons splice junctions is shown. The intron 5' and 3' boundaries are noted in lowercase letters. The conserved gt and ag nucleotides are underlined, and the initiation codon is boxed.

features of a true splice donor site that should be fused to the splice acceptor site of the α exon to yield the p68^{c-ets-1} cDNA (Fig. 4).

To confirm the transcription start sites of the p68^{c-ets-1} mRNA and to ensure that the Y exon is the unique 5' noncoding exon, we performed RNase protection analysis on total RNA extracted from chicken spleen, thymus, and embryonic dermis. A ³²P-labeled RNA probe complementary to the coding strand of a genomic 400 bp BamH I–Hind III DNA fragment comprising the 5' part of the Y exon was used (see Figure 4). The result of a representative experiment is shown in Figure 5A. Protected fragments were detected only in the spleen and in embryonic dermis, where p68^{c-ets-1} is expressed, but not in the thymus, where p68^{c-ets-1} mRNA is not detected, consistent with tissue distribution previously described (Leprince et al., 1990; Quéva et al., 1993). We obtained multiple protected bands—one major band of 86 bp and 4 minor bands of 89, 85, 83, and 77 bp respectively—suggestive of multiple transcription start sites rather than multiple exons, given the small differences in the size of the bands observed. This result indicates that the Y exon is between 105 and 117 bp in size. It should be borne in mind that the size of cDNAs as well as that of primer extension or tailed PCR products can be determined by strong stops during the retrotranscription process. Despite extensive efforts, we could never correlate primer extension with RNase protection data, which probably signifies a premature retrotranscription arrest during the synthesis of the cDNA strand. This observation could also explain why primer extension results (data not shown) matched the size of the longest p68^{c-ets-1} cDNA isolated (see above).

Given the small size of their 5' noncoding regions, we demonstrate in addition the presence of a common unexpectedly long 3' noncoding region, ca. 6.0 kbp, for both the p54^{c-ets-1} and p68^{c-ets-1} mRNAs. These long 3' noncoding regions might allow additional regulation at the posttranscriptional level.

Functional activity of the p68^{c-ets-1} mRNA promoter region in fibroblasts

In order to demonstrate that we have isolated an efficient promoter region for the transcription of p68^{c-ets-1} mRNA, we inserted a 1204 bp Ssp I–Hind III restriction fragment into the

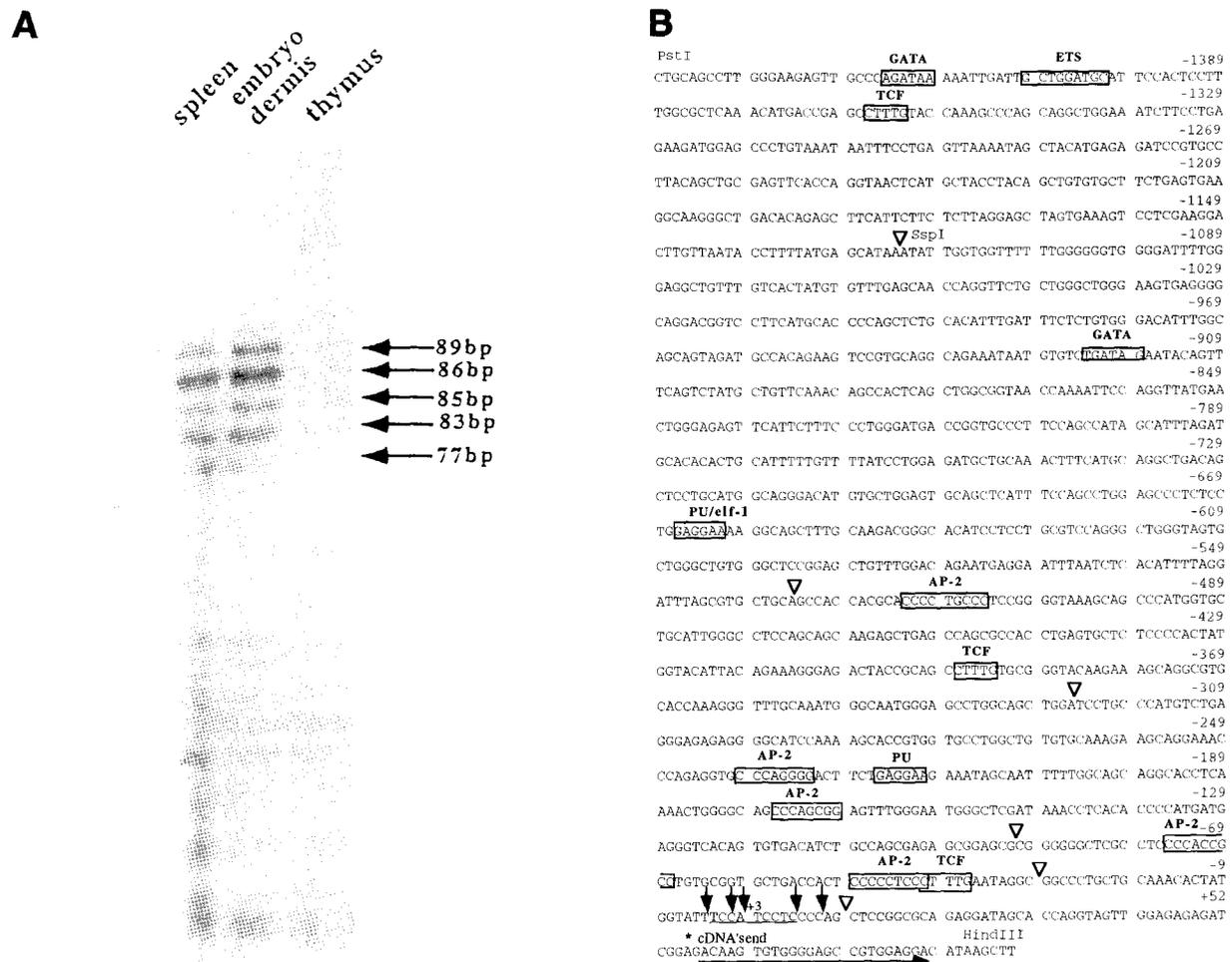


Figure 5. A. Mapping of the p68^{c-ets-1} transcripts start sites by RNase protection analysis on total RNA extracted from chicken spleen, E6 embryonic dermis, or thymus. Five start sites extend the Y exon beyond the Hind III restriction site of 77 to 89 bp, as indicated by arrows. **B.** Nucleotide sequence of the 1.5 kbp chicken genomic DNA spanning the first exon and the 5' upstream region of the p68^{c-ets-1} mRNA. It is numbered with respect to the major start site. The nomenclature is the same as in Figure 3, except that the horizontal arrow marks primer 1. The Inr region is underlined. The open triangles indicate the 5' endpoints of the deletions tested in the functional assay (see Table 1).

pLUCDSS vector. Its 3' border is the Hind III restriction site located in the Y exon (Fig. 4), and this fragment is referred to as fragment -1123 on the basis of the position of the most upstream cap site. This reporter plasmid was transfected into CEF, parallel to the pLUCDSS vector or the ptkFLUC plasmid as controls. We also assayed the promoter activity of different deletion constructs, in order to define crucial regions for transcriptional regulation.

Mean values of several experiments are indicated in Table 1. These results show that the -1123 DNA fragment display weak (60% of the activity of the minimal tk promoter on aver-

age) but significant promoter activity, when compared to that of the pLUCDSS vector (6-fold activation). The levels of the -1123 and -533 fragments' activity are identical, while activity is slightly increased upon deletion of another 211 bp. In contrast, promoter activity of the -90 construct is slightly reduced, and the -28 deletion mutant activity decreases dramatically (10-fold inhibition compared to the full-length promoter), indicative of a positively regulated region located between positions -90 and -28. The low values obtained with the +12 construct corroborate the predicted position of the cap sites.

Nucleotidic sequence of the p68^{c-ets-1} promoter

Sequencing of the 1.5 kbp genomic DNA fragment upstream from the Hind III restriction site located within the Y exon was performed. Examination of the sequence surrounding the predicted cap sites reveals no canonical TATA or CAAT boxes, as had been previously observed with the human and chicken p54^{c-ets-1} promoters. However, in striking contrast to these two promoters, a high G+C content was not seen in the p68^{c-ets-1} promoter (Fig. 5B). Rather, a pyrimidine-rich initiator-like sequence (Inr region), previously described as a potential recognition site for the transcription initiation complex (Smale and Baltimore, 1989), overlaps the position of the start sites. Preliminary DNase I footprinting experiments indicate a protection of this region (data not shown).

Other features of interest are three putative binding sites for ets-related factors at positions -1409, -666, and -225. These sites are centered around a GGAA/T core sequence (Woods et al., 1992). They are suggestive of autoregulation, or of cross-regulation by different members of the *ets* family. Another possible aspect of the regulation of the p68^{c-ets-1} expression is suggested by the presence of several consensus binding sites for lymphotropic transcription factors. We found three putative binding sites for the lymphoid-specific TCF-1 (-1366, -397, -40), a PU box (Klemsz et al., 1990) included in a binding site for the elf-1 protein (Thompson et al., 1992; Wang et al., 1992) at position -666, and two GATA sites (Evans and Felsenfeld, 1991; Ko et al., 1991) at positions -1424 and -923. These DNA motifs might account for the repression of p68^{c-ets-1} mRNA in lymphoid cells. Some binding sites for more ubiquitously expressed transcription factors were also found. Although the overall base composition of the p68^{c-ets-1} mRNA promoter is not of the G+C-rich type, its sequence reveals multiple AP-2 binding sites (Imagawa et al., 1987) in positions -522, -239, -176, -75, and -48. Therefore, some of them are in the vicinity of the predicted cap sites.

The p68^{c-ets-1} protein enhances its own expression

The presence of several binding sites for ets-related transcription factors prompted us to investigate the possibility of autoregulation of the p68^{c-ets-1} promoter, as previously demonstrated for a number of other transcription factors.

Cotransfections in CEF of the -1123 construct with a pSG5 expression vector producing the p68^{c-ets-1} protein resulted in a slight but reproducible (twofold \pm 0.3 SD with 1 μ g of expression plasmid) and dose-dependent activation (Fig. 6). Values obtained in lanes 4 and 5 indicate a lack of effect of up to 4 μ g of pSG5 p68^{c-ets-1} on the pLUCDSS vector. Such twofold activation has been reported as autoregulation in the case of other promoters, among which is the human *c-ets-1* gene (Oka et al., 1991). Such weak activation may be due to the low level of *c-ets-1* proteins exogenously expressed in fibroblasts (our unpublished data). We are currently attempting to establish whether this effect is direct and to identify the DNA sequence involved.

Discussion

Previous studies have demonstrated that the expression of the chicken *c-ets-1* locus leads to the

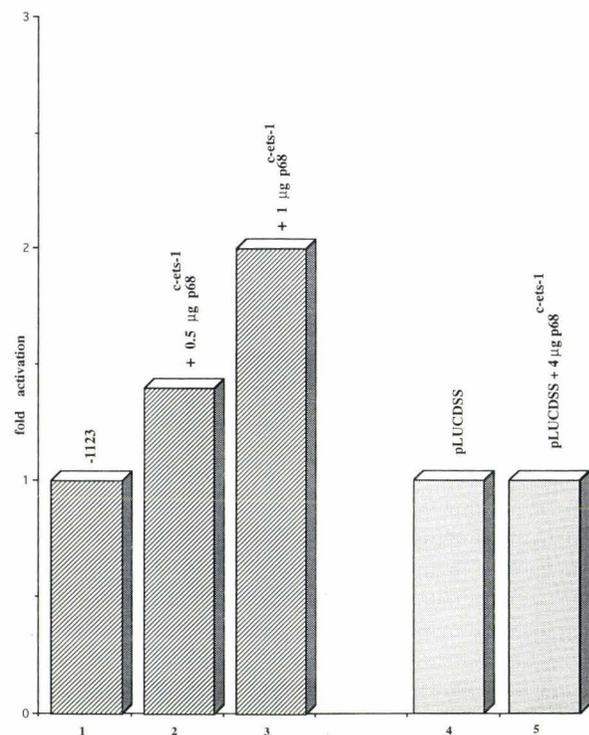


Figure 6. Effect of the p68^{c-ets-1} protein on its own promoter assayed in CEF. The -1123 deletion mutant subcloned into the pLUCDSS was transfected alone (lane 1), or with 0.5 or 1 μ g of pSG5p68^{c-ets-1} expression vector (lanes 2 and 3, respectively). Lanes 4 and 5 show the lack of effect of the expression vector on pLUCDSS activity.

synthesis of two proteins differing in their N-termini, p54^{c-ets-1} and p68^{c-ets-1} (Leprince et al., 1988). Initially, their respective expression, described by immunoprecipitation and RNase protection analysis, suggested marked differences in their level of expression and in their tissue distribution, with preferential localization of p54^{c-ets-1} in lymphoid cells (Ghysdael et al., 1986) and restriction of p68^{c-ets-1} to blood vessels of the spleen (Leprince et al., 1990). But the highly sensitive *in situ* hybridization method reveals a more subtle chicken *c-ets-1* mRNA expression pattern in the embryo (Quéva et al., 1993). In summary, during embryonic life, the tissue distribution of p54^{c-ets-1} and p68^{c-ets-1} largely overlaps. Strikingly, however, p68^{c-ets-1} is completely excluded from chick thymocytes as well as from lymphocytes, while p54^{c-ets-1} is highly expressed in these cells.

To investigate whether transcriptional regulation could account for this differential expression, we were interested in defining the promoter regions involved. In this study, we show that the two major transcripts produced from the chicken *c-ets-1* locus are driven by two distinct promoter regions. The promoter regions directing the respective transcription of p68^{c-ets-1} and of p54^{c-ets-1} mRNAs both exhibit multiple start sites, possibly due to the absence of canonical TATA and CAAT motifs. Nevertheless, they differ in several aspects, among which is the G+C-rich composition of the p54^{c-ets-1} promoter, in contrast to the C+T-rich p68^{c-ets-1} promoter region. In addition, the putative consensus DNA motifs for transcription factors differ somewhat. All of these divergences might account for the differential expression of both *c-ets-1* major transcripts in avian tissues.

We demonstrate that the p54^{c-ets-1} mRNA 5'-flanking region displays a G+C-rich sequence comparable to the human *c-ets-1* and *c-ets-2* promoters previously reported (Mavrothalassitis et al., 1990; Jorcyk et al., 1991; Oka et al., 1991). This type of promoter was initially thought to be typical of housekeeping genes, although it now appears that G+C-rich regions also characterize numerous tissue-specific promoters, such as the hematopoietic *pim-1* gene promoter (Meeker et al., 1990). p54^{c-ets-1} has been identified as a protein remarkably conserved throughout evolution (Watson et al., 1988). We show here that the conservation extends to the promoter region, which appears strikingly simi-

lar in sequence and organization from chicken to human (Jorcyk et al., 1991; Oka et al., 1991). Thus, a 360 bp region (from nucleotides -60 to the ATG codon) in the avian promoter is 65% homologous on average to a colinear region in the human promoter. As for some other proto-oncogene promoters, several of the transcription factor consensus motifs are conserved, but their position varies in the overall sequence. Nevertheless, these observations suggest similar transcriptional regulation between the two species. This regulation may also extend to autoregulation, supported by the observation that the murine p63^{c-ets-1} promoter is positively regulated by the avian p54^{c-ets-1} protein (Seth and Papas, 1990), which is reminiscent of autoregulation of the human promoter (Oka et al., 1991; Majérus et al., 1992).

In contrast to the well-conserved I⁵⁴ exon, the α and β exons, specific for p68^{c-ets-1}, have not been identified in any mammalian species until now (Albagli et al., 1992). The p68^{c-ets-1} mRNA 5'-flanking region reported here lacks characteristic structural features of usual promoters. Nevertheless, examples of transcription regulation sequences lacking a TATA box or a high G+C content have been well documented for a number of developmentally regulated genes, such as *Drosophila* homeotic genes (Biggin and Tjian, 1988) or lymphocyte differentiation genes (Anderson et al., 1988), whose transcripts are driven from a C+T-rich initiator sequence (the so-called Inr).

We identified several putative binding sites for known transcription factors in this promoter, the functionality of which remains speculative at this time. Because the most striking divergence in the patterns of expression of both *c-ets-1* mRNA is illustrated in lymphoid organs, experiments are now under way to test the functionality of lymphoid-specific transcription factors possibly involved in this differential regulation. Among the candidates, the T-cell factor TCF-1 (Waterman et al., 1991; van de Wetering et al., 1991; Allen et al., 1992), GATA-3 (Ko et al., 1991), elf-1 (Thompson et al., 1992), or p54^{c-ets-1} proteins might be effectors that inhibit p68^{c-ets-1} expression or activate p54^{c-ets-1} in lymphoid cells. The important role of the p54^{c-ets-1} protein in lymphoid organs is suggested by the discovery of target genes representative of the onset of lymphocyte-specific biological functions during T cell ontogeny (Prosser et al., 1992;

Ho et al., 1990; Leiden, 1992). The lymphoid-specific repression of p68^{c-ets-1} appears puzzling, since p54^{c-ets-1} and p68^{c-ets-1} are structurally related transcription factors that bind their DNA cognate motif with analogous specificity and affinity. They differ only in their N-termini. The α and β exons are translated into a hydrophobic peptide, which possesses autonomous transactivating properties when fused to a heterologous DNA-binding domain, while the amino acids encoded by I⁵⁴ are mostly hydrophilic and are devoid of such properties (Schneikert et al., 1992). Activation domains interact with the general transcription apparatus either directly or through bridging co-activators. The differences between the two *c-ets-1* proteins may reside in the mechanism by which they interact with other transcriptional activators or co-activators, and thus alter the spectrum of regulated genes. Thus, one might speculate that the strict inhibition of p68^{c-ets-1} expression in lymphoid cells avoids competition for common target genes that both proteins would regulate differently.

The p68^{c-ets-1} mRNA is initiated from a weak promoter. Other genes have been described that show reduced promoter activities highly strengthened by a distant control element. In fact, we did not expect either a potent promoter or enhancers, since previous experiments aimed at measuring the p68^{c-ets-1} mRNA level in tissues always indicated very low levels of expression. In this respect, this promoter is akin to the interleukin-1 receptor gene (Ye et al., 1993). Given the restricted pattern of expression of p68^{c-ets-1}, we could not test its mRNA promoter in a cell type in which the endogenous protein is expressed. We assayed CEF as an easily transfectable and commonly used experimental model, although p68^{c-ets-1} is not expressed in fibroblasts (our unpublished observations). Promoter activities of other tissue-specific genes, such as GATA-1 (Hannon et al., 1991) or *c-rel* (Hanninck et al., 1990), have been already observed in CEF, although no explanation for these results is available at the moment. In fact, when we assayed a DNA fragment extending 3.0 kbp upstream of the Y exon, the promoter activity was dramatically reduced (data not shown), possibly reflecting the lack of function of the endogenous promoter in these cells.

The various binding sites for *ets*-related transcription factors are possibly involved in the

mediation of the slight positive autoregulation of the p68^{c-ets-1} promoter we report. Experiments are now in progress to find a cell in which this autoregulation is enhanced, in case an accessory protein, absent in the CEF, helps p68^{c-ets-1} to autoregulate its expression positively. Because *in situ* hybridization experiments show localized and high-level expression of p68^{c-ets-1} in a small subset of cells in the embryonic dermis (Quéva et al., 1993), we are currently trying to derive an *in vitro* system from this tissue. This model should allow us to address several interesting questions, among which is the status of both promoters in the few cells that synthesize both messengers. Dermis cell culture should also provide a system that would enhance the p68^{c-ets-1} promoter activity, thus showing the regions crucial for transcriptional regulation.

Acknowledgments

We would like to thank members of the GEPHO for their constant interest and support of this work, and especially S. Ansieau, J. Coll, C. Dozier, and M. Duterque-Coquillaud. We are also grateful to R. Lafayette, V. Fafeur, V. Laudet, and C. Quéva for numerous helpful discussions and critical reading of the manuscript. C. Lagrou is warmly thanked for his efficient help in cell cultures.

The financial support of this work was provided by the Centre National de la Recherche Scientifique, by the Association pour la Recherche contre le Cancer, and by the Institut Pasteur de Lille.

The costs of publishing this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 USC Section 1734 solely to indicate this fact.

References

- O. Albagli, A. Flourens, P. Crépieux, A. Bègue, D. Stéhelin, and D. LePrince (1992), *Oncogene* 7, 1435-1439.
- J. M. Allen, K. A. Forbush, and R. M. Perlmutter (1992), *Mol Cell Biol* 12, 2758-2768.
- S. J. Anderson, H. S. Chou, and D. Y. Loh (1988), *Proc Natl Acad Sci USA* 85, 3551-3554.
- M. D. Biggin and R. Tjian (1988), *Cell* 53, 699-711.
- S. Dalton and R. Treisman (1992), *Cell* 68, 597-612.
- J. R. DeWet, K. V. Wood, M. Deluca, D. R. Helinski, and S. Subramant (1987), *Mol Cell Biol* 7, 725-737.
- M. Duterque-Coquillaud, D. LePrince, A. Flourens, C. Henry, J. Ghysdael, B. Debuire, and D. Stéhelin (1988), *Oncogene Res* 2, 335-344.

- T. Evans, T. DeChiara, and A. Efstratiadis (1988), *J Mol Biol* 199, 61–81.
- T. Evans and G. Felsenfeld (1991), *Mol Cell Biol* 11, 843–853.
- A. Gégonne, D. Leprince, M. Duterque-Coquillaud, B. Vandembunder, A. Flourens, J. Ghysdael, B. Debuire, and D. Stéhelin (1987), *Mol Cell Biol* 7, 806–812.
- J. Ghysdael, A. Gégonne, P. Pognonec, D. Dernis, D. Leprince, and D. Stéhelin (1986), *Proc Natl Acad Sci USA* 83, 1714–1718.
- F. Gouilleux, B. Sola, B. Couette, and H. Richard-Foy (1991), *Nucleic Acids Res* 19, 1563–1569.
- C. V. Gunther, J. A. Nye, R. S. Bryner, and B. J. Graves (1990), *Genes Dev* 4, 667–679.
- M. Hanninck and H. Temin (1990), *Oncogene* 5, 1843–1850.
- R. Hannon, T. Evans, G. Felsenfeld, and R. Gould (1991), *Proc Natl Acad Sci USA* 88, 3004–3008.
- J. C. Hanvey, M. Shimizu, and R. D. Wells (1988), *Proc Natl Acad Sci USA* 85, 6292–6296.
- R. A. Hipskind, V. N. Rao, C. G. F. Mueller, E. S. P. Reddy, and A. Nordheim (1991), *Nature* 354, 531–534.
- I. C. Ho, N. K. Bhat, L. R. Gottschall, T. Lindsten, C. B. Thompson, T. S. Papas, and J. M. Leiden (1990), *Science* 250, 814–818.
- M. Imagawa, R. Chiu, and M. Karin (1987), *Cell* 51, 251–260.
- C. L. Jorcyk, D. K. Watson, G. J. Mavrothalassitis, and T. S. Papas (1991), *Oncogene* 6, 523–532.
- F. D. Karim, L. D. Urness, C. S. Thummel, M. J. Klemsz, S. R. McKircher, A. Celada, C. Van Beveren, R. A. Maki, C. V. Gunther, J. A. Nye, and B. J. Graves (1990), *Genes Dev* 4, 1451–1453.
- M. J. Klemsz, S. R. McKircher, A. Celada, C. Van Beveren, and R. Maki (1990), *Cell* 61, 113–124.
- L. J. Ko, M. Yamamoto, M. W. Leonard, K. M. George, P. Ting, and J. D. Engel (1991), *Mol Cell Biol* 11, 2778–2784.
- V. Laudet, C. Niel, M. Duterque-Coquillaud, D. Leprince, and D. Stéhelin (1992), *Biochem Biophys Res Commun* 190, 8–14.
- J. A. Lautenberger, L. A. Burdett, M. A. Gunnell, S. Qi, D. K. Watson, S. J. O'Brien, and T. S. Papas (1992), *Oncogene* 7, 1713–1719.
- J. M. Leiden (1992), *Immunol Today* 13, 22–30.
- D. Leprince, A. Gégonne, J. Coll, C. de Taisne, A. Schneeberger, C. Lagrou, and D. Stéhelin (1983), *Nature* 306, 395–397.
- D. Leprince, M. Duterque-Coquillaud, R. P. Li, C. Henry, A. Flourens, B. Debuire, and D. Stéhelin (1988), *J Virol* 62, 3233–3241.
- D. Leprince, J. C. Gesquière, and D. Stéhelin (1990), *Oncogene Res* 5, 255–265.
- K. MacLeod, D. Leprince, and D. Stéhelin (1992), *Trends Biochem Sci* 17, 251–256.
- M. A. Majérus, F. Bibollet-Ruche, J. B. Telliez, B. Waslyk, and B. Bailleul (1992), *Nucleic Acids Res* 20, 2699–2703.
- T. Maniatis, E. F. Fritsch, and J. Sambrook (1982), *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- G. J. Mavrothalassitis, D. K. Watson, and T. S. Papas (1990), *Proc Natl Acad Sci USA* 87, 1047–1051.
- T. C. Meeker, J. Loeb, M. Ayres, and W. Sellers (1990), *Mol Cell Biol* 10, 1680–1688.
- M. Nunn, P. H. Seeburg, C. Moscovici, and P. H. Duesberg (1983), *Nature* 306, 391–395.
- T. Oka, A. Rairkar, and J. H. Chen (1991), *Oncogene* 6, 2077–2083.
- H. M. Prosser, D. Wotton, A. Gégonne, J. Ghysdael, S. Wang, N. A. Speck, and M. J. Owen (1992), *Proc Natl Acad Sci USA* 89, 9934–9938.
- C. Quéva, D. Leprince, D. Stéhelin, and B. Vandembunder (1993), *Oncogene*, in press.
- V. N. Rao, K. Huebner, M. Isobe, A. Ar-Rushdi, C. M. Croce, and E. S. P. Reddy (1989), *Science* 244, 66–70.
- F. S. Sanger, S. Nicklen, and A. R. Coulson (1977), *Proc Natl Acad Sci USA* 74, 5463–5467.
- J. R. Sayers, W. Schmidt, and F. Eckstein (1988), *Nucleic Acids Res* 16, 791–802.
- J. Schneikert, Y. Lutz, and B. Waslyk (1992), *Oncogene* 7, 249–256.
- R. F. Selden, K. B. Howie, M. E. Rowe, H. M. Goodman, and D. D. More (1986), *Mol Cell Biol* 6, 3173–3179.
- A. Seth and T. S. Papas (1990), *Oncogene* 5, 1761–1767.
- S. T. Smale and D. Baltimore (1989), *Cell* 57, 103–113.
- C. B. Thompson, C. Wang, I. Ho, P. R. Bohjanen, B. Petryniak, C. H. June, S. Miesfeldt, L. Zhang, G. J. Nabel, B. Karpinski, and J. M. Leiden (1992), *Mol Cell Biol* 12, 1043–1053.
- M. van de Wetering, M. Oosterwegel, D. Dooijes, and H. Clevers (1991), *EMBO J* 10, 123–132.
- J. V. Virbasius and R. C. Scarpulla (1991), *Mol Cell Biol* 11, 5631–5638.
- C. Y. Wang, B. Petryniak, I. C. Ho, C. B. Thompson, and J. M. Leiden (1992), *J Exp Med* 175, 1391–1399.
- B. Waslyk, C. Waslyk, P. Flores, A. Bègue, D. Leprince, and D. Stéhelin (1990), *Nature* 346, 191–193.
- C. Waslyk, A. Gutman, R. Nicholson, and B. Waslyk (1991), *EMBO J*, 5, 1127–1134.
- M. L. Waterman, W. H. Fisher, and K. A. Jones (1991), *Genes Dev* 5, 656–669.
- D. K. Watson, M. J. McWilliams, P. Lapis, J. A. Lautenberger, C. W. Schweinfest, and T. S. Papas (1988), *Proc Natl Acad Sci USA* 85, 7862–7866.
- D. B. Woods, J. Ghysdael, and M. Owen (1992), *Nucleic Acids Res* 20, 699–704.
- K. Ye, C. A. Dinarello, and B. D. Clark (1993), *Proc Natl Acad Sci USA* 90, 2295–2299.

DISCUSSION

	poussin		embryon		
	p54	p68	p54	p68	
thymus	+++	-	+++	-	E15
bourse	++	-	++	-	
rate	++	+	++	+	
rate			+++	+++	E6
rein			+++	+	
foie			+++	+	
derme en condensation			+	+++	
cellules endothéliales			++	++	E3
précurseurs érythroïdes primitifs			+	+	
bourgeon de membre			+	++	

Tableau 10 : Résumé de la répartition des ARNm p54^{c-ets-1} et p68^{c-ets-1} dans différents tissus de poussin et de d'embryon, par hybridation *in situ*.

Les premiers travaux visant à décrire la distribution tissulaire des ARNm *c-ets-1* chez le poulet indiquaient une expression de l'ARNm p54^{c-ets-1} relativement large à faible niveau, bien que prépondérante dans les organes lymphoïdes (Ghysdael *et al.*, 1986b ; Chen, 1985 et 1988 ; Duterque-Coquillaud *et al.*, 1988). A l'inverse, la présence de l'ARNm p68^{c-ets-1} apparaissait difficilement détectable et strictement restreinte aux vaisseaux sanguins de la rate (Leprince *et al.*, 1990). L'avènement de la technique d'hybridation *in situ* au laboratoire a permis de dresser une image plus précise de la répartition des deux types de transcrits chez l'embryon (Quéva *et al.*, 1993). Le tableau 10 résume l'ensemble des données obtenues par cette technique : les deux transcrits sont globalement observés dans les cellules du mésenchyme embryonnaire en interaction avec les épithéliums comme le derme au cours de la condensation des bourgeons plumaires (Desbiens *et al.*, 1991), et leur expression est le plus souvent liée à l'incidence de vaisseaux sanguins en formation (Vandenbunder *et al.*, 1989). Par contre, ils sont indétectables dans les ébauches des épithéliums, quelle que soit l'origine embryonnaire de ces derniers. Toutefois, malgré cette expression conjointe dans la plupart des tissus mésenchymateux, les marquages obtenus procurent des images différentes : au sein d'un même tissu, les messagers p54^{c-ets-1} sont répartis uniformément tandis que l'ARNm p68^{c-ets-1} est le plus souvent régionalisé à une sous-population cellulaire.

L'expression de l'ARNm p68^{c-ets-1} semble étendue à davantage de types cellulaires chez l'embryon que chez le poulet, bien qu'une technique aussi sensible que l'hybridation *in situ* appliquée aux tissus de poulet pourrait peut-être fournir le même type de résultats. Par contre, elle est complètement inexistante dans les cellules lymphoïdes, quel que soit le stade de développement.

L'ARNm p68^{c-ets-1} est également détecté dans les précurseurs érythroïdes primitifs des îlots sanguins extra-embryonnaires et dans les précurseurs de la lignée définitive du sac vitellin à E3 (Quéva *et al.*, 1993). Il s'éteint ensuite dans les cellules circulantes plus différenciées. Cette observation est intéressante en regard des mécanismes qui ont siégé à la genèse du rétrovirus E26. En effet, des transcrits du gène *c-myb* ont également été détectés dans ces cellules (Quéva *et al.*, 1992). Ces précurseurs constituent le premier système hématopoïétique identifié qui exprime simultanément les gènes *c-myb* et *c-ets-1* sous sa version p68^{c-ets-1}. Si des expériences futures de double marquage par anticorps confirment que les protéines correspondantes sont synthétisées, elles suggèreraient avec insistance que les deux facteurs coopèrent naturellement à la régulation de gènes spécifiques d'un stade précoce de la maturation érythroïde que le rétrovirus E26 aurait figé.

C'est de ces données que ce sont inspirés les travaux visant à isoler les promoteurs du gène *c-ets-1* de poulet, afin de définir si les divergences ainsi que les similitudes d'expression des deux transcrits sont basées sur des variations équivalentes

portant sur leur régulation transcriptionnelle. La caractérisation des promoteurs *c-ets-1* a permis non seulement de répondre à un certain nombre de questions d'ordre structural, mais aussi de fournir des pistes d'investigation concernant l'un des aspects de la régulation de deux facteurs de transcription aux propriétés vraisemblablement dissemblables.

1. Structure des ARNm *c-ets-1* :

A l'époque à laquelle ce travail fut initié, un ADNc de 1.5 kpb codant la protéine p54^{c-ets-1} avait été isolé d'une banque de rate de poulet. Un second ADNc de même taille comportant en 5' les exons α et β a été obtenu à partir d'une banque d'embryon de poulet. Les deux banques en question sont construites à base d'ARNm rétrotranscrits à partir d'une amorce nucléotidique complémentaire du site de polyadénylation. Or, l'analyse de la séquence des deux ADNc indiquait qu'ils ne comportaient pas de queue de polyA. Par conséquent, le premier problème qui se posait consistait à déterminer la répartition des 6 kb d'information non codante des deux types d'ARNm *c-ets-1*. Après quelques tentatives infructueuses de cartographie par la RNase H, la détermination précise des sites d'initiation de la transcription des ARNm *c-ets-1* devait y répondre : 6 kpb de séquences non codantes exoniques sont accumulées à l'extrémité 3' du gène *c-ets-1* de poulet.

D'une façon plus générale, les extrémités 3' non codantes des ARNm peuvent être impliquées dans le contrôle de la stabilité et de la traductibilité, par le biais d'une séquence de type UAUUU (pour revue, voir Jackson et Standart, 1990). Nous ignorons tout de cette extrémité du gène *c-ets-1*, hormis 0.3 kpb environ qui ne présente pas ce motif. Pourtant, la stabilité des ARNm *c-ets-1* est susceptible d'être modulée puisque, à titre d'exemple, dans les lymphocytes T activés et traités par la cycloheximide, la demi-vie de l'ARNm codant la protéine humaine est anormalement prolongée (Bhat *et al.*, 1990). La même observation est valable dans des cellules endothéliales bovines (V. Fafeur, communication personnelle). Ces données suggèrent donc l'existence de facteurs protéiques, des nucléases vraisemblablement, dont la mission consisterait à dégrader l'ARNm *c-ets-1* de façon régulée, et dont l'action pourrait être entravée artificiellement. Cependant, ces expériences seraient plus informatives si elles étaient également pratiquées en présence d'actinomycine D, un inhibiteur de la transcription, pour écarter la possibilité d'une superinduction transcriptionnelle.

Il pourrait donc être informatif d'isoler des ADNc complets, sans doute par criblage de banques construites selon d'autres modalités que précédemment, afin de préciser cet aspect. De plus, l'organisation génomique de cette information 3' non codante est, par la force des choses, totalement inconnue. Sa description pourrait encore nous réserver des surprises, au cas où, par exemple, de nouveaux épissages alternatifs étaient découverts.

2. Comparaison avec les autres promoteurs des gènes de la famille Ets :

Les deux promoteurs décrits dans ces pages présentent un certain nombre d'analogies avec des promoteurs des gènes de la famille Ets déjà connus. Comme les promoteurs des gènes humains *ets-1*, *ets-2*, *fli-1* et du gène *PU.1* de souris, les deux promoteurs du gène *c-ets-1* aviaire démarrent la transcription à partir de multiples sites d'initiation, corrélativement à l'absence de boîte TATA dans chacun de ces promoteurs. Selon les auteurs, les positions assignées aux sites d'initiation sont discordantes. Peut-être faut-il imputer ces divergences de résultats aux types cellulaires étudiés, lymphocytes B pour les uns (Oka *et al.*, 1991), lymphocytes T pour les autres (Jorcyk *et al.*, 1991), en fonction desquels le démarrage de la transcription pourrait varier. Les études concernant ces différents promoteurs ne sont certes pas encore suffisamment avancées pour qu'en soit tirées des conclusions vraiment nettes. Toutefois, ils semblent dans l'ensemble relativement peu modulables. Est-ce à dire que des éléments enhancer ou silencer distaux sont encore à découvrir ?

A l'inverse du promoteur p54^{c-ets-1}, le promoteur p68^{c-ets-1} n'est pas de type riche en bases G et C ; il initie la transcription à partir d'une petite zone d'une dizaine de nucléotides enchainant des bases C et T, analogue à celle à partir de laquelle les ARNm *lck* (Garvin *et al.*, 1988) ou codant la terminal désoxynucléotidyl transférase (Landau *et al.*, 1984) sont initiés. Les modalités d'assemblage du complexe d'initiation sur un promoteur TATA ou "TATA-less" semblent analogues (pour revue, Hernandez, 1993). Ainsi, sur un promoteur TATA-less, le complexe TFIID se fixera en position -30, quelque soit la séquence qui s'y trouve. Cependant, l'affinité du complexe risque de s'en trouver nettement amoindrie et le niveau basal d'expression diminué d'autant, plus la séquence de fixation s'éloigne d'un consensus TATA optimal. Si l'on place les séquences situées entre les positions -33 et -24 des promoteurs *c-ets-1* sur une échelle établie récemment qui corréle le niveau d'expression basale du promoteur du gène codant la TdT au type de séquence de fixation de TFIID (Zenzie-Grégory *et al.*, 1993), il apparaît qu'elles sont toutes deux très défavorables, ce qui expliquerait la faible activité basale de ces promoteurs.

Aujourd'hui, les caractéristiques qui viennent d'être énoncées ici ne sont plus, loin s'en faut, associées à l'expression de gènes de ménage, expression qui serait relativement peu modulable au niveau transcriptionnel : la littérature abonde d'exemples de gènes subtilement régulés à ce niveau, et dont les promoteurs sont dépourvus d'élément TATA.

Une autre caractéristique intéressante des gènes *ets-1*, *ets-2* et *fli-1* est qu'ils comportent tous une longue succession de nucléotides puriques localisés en 3' des sites d'initiation. L'extrémité non codante de l'ARNm p54^{c-ets-1} en présente une elle aussi. Ce type de structure peut adopter une conformation inhabituelle de triple hélice (ADN sous forme

H), qui laisse un brin non apparié (Hanvey *et al.*, 1988 ; Voloshin *et al.*, 1988). Fréquemment observée au sein des promoteurs de gènes eucaryotes, les triples hélices sont ultra-sensibles à l'action de la nucléase S1, ce qui procure un moyen de les détecter expérimentalement. Elles se forment préférentiellement à pH acide. Elles influenceraient, en modifiant sa structure, l'accessibilité de la chromatine aux protéines et pourraient donc intervenir au cours des processus de régulation génique.

3. Autorégulation du locus *c-ets-1* :

Semblablement au promoteur du gène *ets-1* humain, nos travaux montrent que l'un des promoteurs *c-ets-1* aviaire est stimulé par le produit du gène. Certes, l'action de la protéine p68^{c-ets-1} sur son propre promoteur est relativement modeste, bien que comparable aux niveaux de stimulation fréquemment rapportés dans la littérature et obtenus par le biais d'une méthodologie analogue. L'autorégulation positive d'un gène étant un moyen d'amplifier exponentiellement son produit quand la cellule a besoin d'en disposer en quantité importante, on s'explique difficilement la signification physiologique d'une amplification d'incrément 2. Toutefois, il convient d'apporter plusieurs commentaires à l'interprétation de cette expérience. En effet, les expériences de cotransfection qui ont donné lieu à ce résultat sont pratiquées avec un vecteur d'expression dans lequel est cloné l'ADNc de la protéine p68^{c-ets-1}. Or, nous avons vu que cet ADNc est amputé de la partie 3' de l'ARNm *c-ets-1*. Une différence de stabilité entre l'ARNm transcrit à partir de l'ADNc et l'ARNm de pleine longueur pourrait influencer sur la quantité de protéine p68^{c-ets-1} produite dans la cellule hôte. De fait, l'analyse des lysats cellulaires par la technique de Western blot n'a jamais permis de détecter la protéine, suggérant qu'elle est effectivement synthétisée à très faible taux. Il convient d'ajouter à ce stade que le vecteur d'expression utilisé, le pSG5, contient une origine de répllication du virus SV40 ainsi que des séquences promotrices de ce virus, en aval desquelles l'ADNc de la p68^{c-ets-1} est cloné. Ces attributs servent à amplifier l'expression de la protéine dans une cellule exprimant l'antigène T du virus SV40 et à permettre la répllication du plasmide. Or, les FEP sont des cellules primaires dans lesquelles il faut donc s'attendre à un très faible niveau d'expression de la protéine exogène p68^{c-ets-1}.

En supplément, il est également envisageable que le peu de protéine p68^{c-ets-1} exprimée dans ces conditions nécessite la co-expression d'un partenaire transcriptionnel qui, s'il était présent dans les FEP, lui permettrait d'accomplir pleinement sa mission en coopérant à l'activation complète du promoteur. En effet, il a été développé dans l'introduction de ce mémoire, l'importance des coopérations entre les protéines de la famille Ets et d'autres familles de facteurs de transcription.

Il importe de surcroît de définir précisément le site d'autorégulation sur le promoteur. Un fragment de 3.0 kpb en amont des sites d'initiation présente une activité promotrice très faible, qui reflète probablement l'absence d'expression de l'ARNm p68^{c-ets-1} dans les FEP. Cependant, ce fragment d'ADN est stimulé lui aussi par la protéine, selon un ordre de grandeur comparable au fragment de 1.1 kpb. Le site Ets le plus en amont que nous ayons identifié, en position -1409 n'est donc pas concerné par cette régulation. Il reste deux sites Ets situés en position -666 et -225.

Finalement, il est également possible que l'autorégulation du promoteur p68^{c-ets-1} n'ait pas de véritable réalité physiologique et qu'une autre protéine de la famille Ets soit mieux adaptée à la stimulation de ce promoteur. Avec toutes les restrictions qui s'imposent concernant la spécificité de reconnaissance de l'ADN des protéines Ets, des expériences préliminaires de retard sur gel les mettant en présence des deux sites Ets du promoteur p68^{c-ets-1} qui viennent d'être évoqués devraient fournir quelques pistes.

Nous avons également cherché à reproduire l'autorégulation du promoteur humain sur le promoteur p54^{c-ets-1} aviaire, sans résultat cependant. Le promoteur que nous avons analysé est un fragment de 0.7 kpb, qui ne contient apparemment pas de site Ets. Aussi est-il probable qu'il soit trop court en 5' ; c'est pourquoi nous tentons actuellement d'isoler la région génomique adjacente. A moins que la régulation du promoteur aviaire soit différente de celle du promoteur humain...

4. L'expression des protéines c-Ets-1 aviaires est-elle mutuellement exclusive ?

Cette question peut être abordée à la lumière de trois situations différentes, schématisées sur la figure 28 et qui constituent autant de modèles à explorer :

- dans les lymphocytes, seuls l'ARNm et la protéine p54^{c-ets-1} sont détectés, et l'expression de la p68^{c-ets-1} est clairement réprimée. La protéine p54^{c-ets-1} est-elle responsable de cette inhibition ou bien celle-ci est-elle le fait d'une autre protéine, dont l'expression serait restreinte aux organes lymphoïdes ? Des expériences préliminaires de cotransfection du promoteur p68^{c-ets-1} et d'un plasmide exprimant la p54^{c-ets-1} n'ont pas fourni de résultats très probants. De plus, chez la souris, l'équivalent de la protéine p54^{c-ets-1} n'apparaît que relativement tardivement au cours du développement. Au fil de l'activation lymphocytaire, l'expression de cette protéine se voit rétro-réglée négativement. Si l'on transpose audacieusement l'ontogenèse et l'activation des lymphocytes murins à celles des lymphocytes de poulet, l'hypothèse d'une protéine strictement lymphocytaire autre que la p54^{c-ets-1} semble la plus vraisemblable. En effet, quelque soit le stade ou l'état des lymphocytes, même en absence de p54^{c-ets-1}, la protéine ou du moins l'ARNm p68^{c-ets-1} n'est jamais détecté. Le facteur TCF-1, qui est

une protéine à domaine de liaison à l'ADN de type HMG mais qui reconnaît néanmoins une séquence spécifique, semblait être un bon candidat parce qu'il est strictement exprimé dans les lymphocytes T et très tôt au cours du développement des thymocytes (van de Wetering *et al.*, 1991). Mais, bien qu'il existe un site de haute affinité pour ce facteur dans le promoteur p68^{c-ets-1} en position -40 reconnu par une protéine voisine TCF-1 α en retard sur gel, nous avons obtenu des résultats préliminaires qui compromettent son implication.

Il est permis de penser que les protéines p54^{c-ets-1} et p68^{c-ets-1} ne sont pas interchangeables du point de vue de leur activité transcriptionnelle. En effet, bien que leur structure soit commune sur une grande partie de la molécule, elles diffèrent par leur extrémité amino-terminale, qui est globalement hydrophobe dans le cas de la p68^{c-ets-1} et plutôt hydrophile dans celui de la p54^{c-ets-1}. Il est donc tentant d'imaginer que les interactions qu'elles contractent avec d'autres facteurs de transcription sont spécifiques de chacune d'entre elles, ce qui empêcherait qu'elles n'entrent en compétition pour les mêmes gènes cibles, malgré le fait qu'elles se fixent apparemment à la même séquence de reconnaissance. Or, la possibilité que ces deux protéines régulent un ensemble de gènes différents semble important, puisqu'*in vitro*, elles présentent des pouvoirs transactivateurs différents, bien qu'agonistes : la protéine p54^{c-ets-1} fait figure de piètre transactivateur en comparaison de la p68^{c-ets-1}. Alternativement, elles pourraient se relayer pour réguler quantitativement un même ensemble de gènes cibles en fonction des besoins cellulaires. Peut-on imaginer que les cellules lymphoïdes soient les seules cellules de l'organisme qui contiendraient un facteur protéique capable d'interagir avec chacune des deux protéines c-ets-1 indistinctement, en leur conférant des activités antagonistes indésirables ? Pour l'instant, il est connu que le complexe AP-1 coopère aussi efficacement avec chacun des deux facteurs (Wasylyk *et al.*, 1990). L'existence d'une protéine pour laquelle les protéines p54^{c-ets-1} et p68^{c-ets-1} pourraient entrer en compétition n'est pas décrite, mais une réponse pourrait être apportée par le criblage, à l'aide de leur extrémité amino-terminale respective, de banques d'expression de lymphocytes ou de blots de protéines lymphocytaires.

Une autre alternative serait que l'exclusion de la p68^{c-ets-1} des cellules lymphoïdes ait plutôt pour effet de préserver l'activité transcriptionnelle de la protéine Ets-2 qui y est également abondamment exprimée. En effet, les protéines p68^{c-ets-1} et p55^{c-ets-2} aviaires entrent en compétition vis-à-vis d'un co-facteur commun (Schneikert *et al.*, 1992). Les deux protéines sont douées d'une organisation similaire et, à l'aide de tests de transactivation *in vitro*, aucune différence d'activité n'est décelée. Il a été dit plus haut que la production de protéine Ets-2 dépend du cycle cellulaire, mais rien n'est connu à ce sujet dans le cas de la protéine p68^{c-ets-1}. Si leur activité est purement redondante, il peut être nécessaire de limiter quantitativement cette activité dans les lymphocytes. On ignore si, dans les tissus exprimant la protéine p68^{c-ets-1}, l'expression

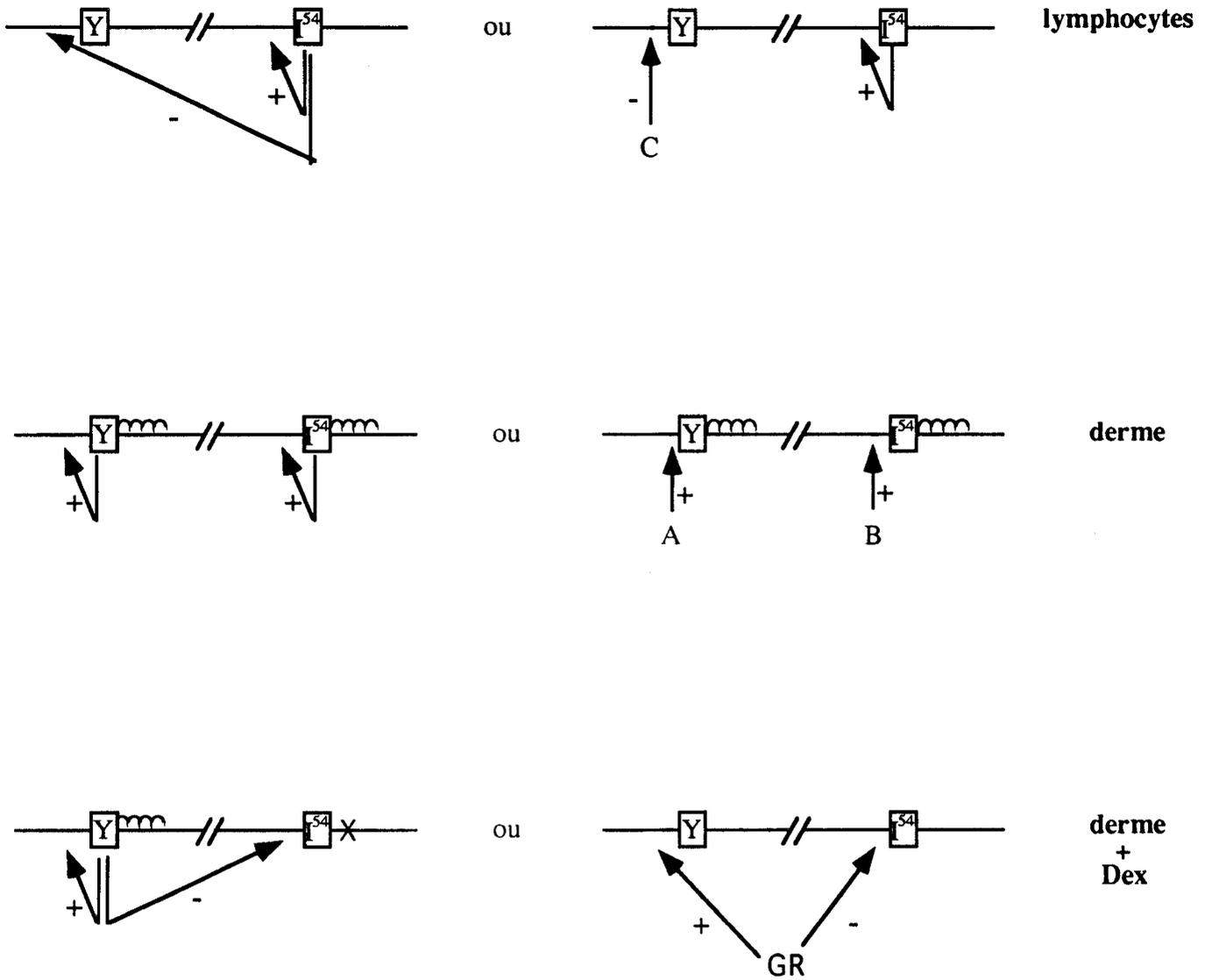


Figure 28 : Quelques-unes des nombreuses hypothèses de régulation transcriptionnelle des ARNm *c-ets-1* dans des situations significatives. Les protéines p54^{c-ets-1} et p68^{c-ets-1} sont figurées par leur exon 5' respectif.

de la protéine Ets-2 est à son tour réprimée, ce qui évoquerait l'existence de différences fonctionnelles non accessibles à l'expérimentation telle que nous la pratiquons aujourd'hui.

- par hybridation *in situ*, on détecte l'ARNm p54^{c-ets-1} et l'ARNm p68^{c-ets-1} dans le derme d'embryon de poulet au sixième jour de développement. Bien que l'expression de la p54^{c-ets-1} soit largement distribuée dans ce tissu et que celle de la p68^{c-ets-1} semble intense mais restreinte à une sous-population cellulaire, il est probable que les deux types d'ARNm cohabitent dans des cellules identiques. Par immunoprécipitation de lysats de derme d'embryon, peut être visualisée une quantité appréciable de protéine p54^{c-ets-1} et une quantité plus faible de protéine p68^{c-ets-1}. Les deux protéines peuvent vraisemblablement assurer leur propre expression par autorégulation positive de leur promoteur respectif. A moins que d'autres facteurs en soient responsables mais aucun élément ne permet de le supposer pour l'instant.

- le traitement d'embryon de poulet par la dexaméthasone désorganise la géométrie de répartition des ébauches de bourgeons plumaires, par altération du réseau de fibres de collagène le long desquels ces bourgeons se distribuent normalement. Conjointement est constatée une diminution du taux d'expression des ARNm *c-ets-1* et une augmentation de l'ARNm *c-jun* (N. Turque, 1991). L'étude des protéines effectuée par immunoprécipitation indique qu'en fait, le traitement d'embryon par la dexaméthasone provoque une vive répression de la protéine p54^{c-ets-1} et une augmentation de production de la protéine p68^{c-ets-1} (A. Hautefeuille et N. Turque, communication personnelle). Il est possible que seul le promoteur p68^{c-ets-1} réponde au récepteur, et que la protéine p68^{c-ets-1} soit seule responsable de sa propre stimulation et de la répression du promoteur p54^{c-ets-1}. Ou bien, les deux promoteurs pourraient être la cible directe du récepteur des glucocorticoïdes dont l'activité est stimulée par la dexaméthasone. Cependant, des données préliminaires indiquent que la dexaméthasone n'a pas d'effet sur l'activité des promoteurs *c-ets-1* transfectés dans des FEP. Il faut donc en conclure que soit la dexaméthasone n'exerce pas son effet au niveau transcriptionnel, soit elle l'exerce par l'intermédiaire d'un facteur de transcription différent des protéines c-Ets-1, du moins pour déclencher la réaction au niveau du promoteur p68^{c-ets-1}. La protéine c-Jun, dont la cinétique de réponse à la dexaméthasone est analogue à celle de la p68^{c-ets-1}, pourrait réguler son promoteur directement. Pourtant, les FEP expriment la protéine c-Jun. Faut-il y voir l'implication d'une autre protéine de la famille Jun, d'expression plus restreinte comme JunD par exemple ?

Il résulte des trois exemples cités ici que l'expression des protéines c-Ets-1 aviaires peut être mutuellement exclusive dans certains cas, illustrés par les lymphocytes et par le

derme embryonnaire traité par la dexaméthasone. Toutefois, les modalités de cet équilibre sont complètement spéculatives pour l'instant, et leur élucidation va nécessiter de préciser les zones promotrices qui pourraient être la cible des effets observés. Pour ce faire, la mise au point d'un système de culture de derme embryonnaire s'avère indispensable.

5. La régulation des ARNm. *c-ets-1* est-elle transcriptionnelle ?

Curieusement, aucun des auteurs ayant étudié les promoteurs de gènes de la famille Ets ne s'est véritablement attaché à résoudre cette question. A vrai dire, nos travaux sur les promoteurs aviaires ne permettent pas non plus d'y répondre. Il a été dit plus haut qu'existe vraisemblablement une régulation post-transcriptionnelle dans les lymphocytes T et dans les cellules endothéliales bovines, qui pourrait se manifester par une dégradation de l'ARNm *ets-1* par une RNase labile. Mais aucune expérience de "run-on", permettant d'écarter l'hypothèse d'une régulation par atténuation transcriptionnelle des ARNm *ets-1* n'a jamais été rapportée. Des cinétiques employant l'actinomycine D pour étudier leur durée de vie et son éventuelle stabilisation sous certaines conditions expérimentales n'apparaissent pas non plus dans la littérature. Par conséquent, l'hypothèse d'une stabilisation modulée d'un ARNm constitutivement synthétisé n'est pas à écarter.

Toutefois, il est permis de penser que la régulation des ARNm *c-ets-1* est en partie transcriptionnelle, au regard de la stimulation du promoteur *ets-1* humain par des facteurs comme Ets-1, AP-1, PEA3 ou Oct-1 qui semblent agir directement sur les motifs de reconnaissance qui sont les leurs dans ce promoteur. Par conséquent, force est de constater que la régulation des ARNm *ets-1* chez l'homme et probablement aussi chez le poulet, présente une composante transcriptionnelle et vraisemblablement une composante post-transcriptionnelle qui demeure un champ d'investigation ouvert.

RÉFÉRENCES BIBLIOGRAPHIQUES



- ALBAGLI O., A. FLOURENS, P. CREPIEUX, D. STEHELIN and D. LEPRINCE. (1992). Phylogeny of the p68^{c-ets-1} amino-terminal transactivating domain reveals some highly conserved structural features. *Oncogene*, **7**, 1435-1439.
- AMOUYEL P., V. LAUDET, P. MARTIN, R. LI, B. QUATANNENS, D. STEHELIN and S. SAULE. (1989). Two nuclear oncogenic proteins, P135^{gag-myb-ets} and p61/63^{myc}, cooperate to induce transformation of chicken neuroretina cells. *J. Virol.*, **63**, 3382-3388.
- ANDRES A.J. and C.S. THUMMEL. (1992). Hormones, puffs and flies : the molecular control of metamorphosis by ecdysone. *Trends Genet.*, **8**, 132-138.
- ANGRAND P.O. (1993). Les domaines de liaison de l'ADN des facteurs de transcription eucaryotes. *Med. Sci.*, **9**, 725-736.
- ANTON I.A. and J. FRAMPTON. (1988). Tryptophans in Myb proteins. *Nature*, **336**, 719.
- ARCHER T.K., M.G. CORDINGLEY, R.G. WOLFORD and G.L. HAGER. (1991). Transcription factor access is mediated by accurately positioned nucleosomes on the mouse mammary tumor virus promoter. *Mol. Cell. Biol.*, **11**, 688-698.
- ASCIONE R., D.M. THOMPSON, R. THOMAS, A. PANAYIOTAKIS, R. RAMSAY, M. TYMMS, I. KOLA and A. SETH. (1992). Influence of nucleotides flanking the -GAGA- core sequence on ETS1 and ETS2 DNA-binding activity and the mechanism of ETS1 autoregulation. *Int. J. Oncol.*, **1**, 631-637.
- ASHBURNER M., C. CHIHARA, P. MELTZER and G. RICHARDS. (1974). Temporal control of puffing activity in polytene chromosomes. *Cold Spring Harbor Symp. Quant. Biol.*, **38**, 655-662.
- AURIGEMMA R.E., D.G. BLAIR and S.K. RUSCETTI. (1992). Transactivation of erythroid transcription factor GATA-1 by a myb-ets-containing retrovirus. *J. Virol.*, **66**, 3056-3061.
- BARBEAU B., D. BERGERON, L. POLIQUIN and E. RASSART. (1992). Characterization of the FLI-1 integration site of the CAS-BR-E retrovirus and analysis of the FLI-1 promoter. 8th annual meeting on oncogenes, Hood College, Frederick, 429.
- BARBEAU B., D. BERGERON, Z. NADGEM, M. BEAULIEU and E. RASSART. (1993). Characterization of the mouse FLI-1 promoter. 9th annual meeting on oncogenes, Hood College, Frederick, 78.
- BAUD V., M. LIPINSKI, E. RASSART, L. POLIQUIN and D. BERGERON. (1991). The human homolog of the mouse common viral integration region FLI-1 maps to 11q23-q24. *Genomics*, **11**, 223-224.
- BAUERLE P.A. and D. BALTIMORE. (1988). I κ B : a specific inhibitor of the NF- κ B transcription factor. *Science*, **242**, 540-545.
- BEN-DAVID Y., E.B. GIDDENS and A. BERNSTEIN. (1990). Identification and mapping of a common proviral integration site Fli-1 in erythroleukemia cells induced by Friend murine leukemia virus. *Proc. Natl. Acad. Sci. USA*, **87**, 1332-1336.
- BEN-DAVID Y. and A. BERNSTEIN. (1991). Friend virus-induced erythroleukemia and the multistage nature of cancer. *Cell*, **66**, 831-834.

BEN-DAVID Y., E.B. GIDDENS, K. LETWIN and A. BERNSTEIN. (1991). Erythroleukemia induction by Friend murine leukemia virus : insertional activation of a new member of the ets gene family, Fli-1, closely linked to c-ets-1. *Genes Dev.*, **5**, 908-918.

BENEZRA R., R.L. DAVIS, D. LOCKSHON, D.L. TURNER and H. WEINTRAUB. (1990). The protein Id : a negative regulator of helix-loop-helix DNA binding proteins. *Cell*, **61**, 49-59.

BEUG H., A. LEUTZ, P. KAHN and T. GRAF. (1984). Ts mutants of E26 leukemia virus allow transformed myeloblasts, but not erythroblasts or fibroblasts, to differentiate at the nonpermissive temperature. *Cell*, **39**, 579-588.

BHAT N.K., R.J. FISHER, S. FUJIWARA, R. ASCIONE and T.S. PAPAS. (1987). Temporal and tissue-specific expression of mouse ets genes. *Proc. Natl. Acad. Sci. USA*, **84**, 3161-3165.

BHAT N.K., K.L. KOMSCHLIES, S. FUJIWARA, R.J. FISHER, B.J. MATHIESON, H.A. GREGORIO, J.W. YOUNG, K. KASIK, K. OZATO and T.S. PAPAS. (1989). Expression of ets genes in mouse thymocyte subsets and T cells. *J. Immunol.*, **142**, 672-678.

BHAT N.K., C.B. THOMPSON, T. LINDSTEN, C.H. JUNE, S. FUJIWARA, S. KOIZUMI, R.J. FISHER and T.S. PAPAS. (1990). Reciprocal expression of human ETS1 and ETS2 genes during T-cell activation : regulatory role for the protooncogene ETS1. *Proc. Natl. Acad. Sci. USA*, **87**, 3723-3727.

BISTER K., M. NUNN, C. MOSCOVICI, B. PERBAL, M.A. BALUDA and P.H. DUESBERG . (1982). Acute leukemia viruses E26 and avian myeloblastosis virus have related transformation-specific RNA sequences, but different genetic structures, gene products and oncogenic properties. *Proc. Natl. Acad. Sci. USA*, **79**, 3677-3681.

BLACKWOOD E.M. and R.N. EISENMAN. (1991). Max : a helix-loop-helix zipper protein that forms a sequence-specific DNA-binding complex with myc. *Science*, **251**, 1211-1217.

BLACKWOOD E.M., B. LUSCHER and R.N. EISENMAN. (1992). Myc and max associate *in vivo*. *Genes Dev.*, **6**, 71-80.

BOISE L.H., B. PETRYNIAK, X. MAO, C.H. JUNE, C.Y. WANG, T. LINDSTEN, R. BRAVO, K. KOVARY, J.M. LEIDEN and C.B. THOMPSON. (1993). The NFAT-1 DNA binding complex in activated T cells contains Fra-1 and junB. *Mol. Cell. Biol.*, **13**, 1911-1919.

BORELLINI F., Y.F. HE, A. AQUINO, G.YU, S.F. JOSEPHS and R.I. GLAZER. (1991). Increased DNA binding and transcriptional activity associated with transcription factor Sp1 in K562 cells transfected with the myeloid-specific c-fes tyrosine kinase gene. *J. Biol. Chem.*, **266**, 15850-15854.

BORTNER D.M., M. ULIVI, M.F. ROUSSEL and M.C. OSTROWSKI. (1991). The carboxy-terminal catalytic domain of GTPase-activating protein inhibits nuclear signal transduction and morphological transformation mediated by the CSF-1 receptor. *Genes Dev.*, **5**, 1777-1785.

BOSSELUT R., J.F. DUVALL, A. GEGONNE, M. BAILLY, A. HEMAR, J. BRADY and J. GHYSDAEL. (1990). The product of the c-ets-1 proto-oncogene and the related

Ets2 protein act as transcriptional activators of the long terminal repeat of human T cell leukemia virus HTLV-1. *EMBO*, **9**, 3137-3144.

BOULUKOS K.E., P. POGNONEC, A. BEGUE, F. GALIBERT, J.C. GESQUIERE, D. STEHELIN and J. GHYSDAEL. (1988). Identification in chickens of an evolutionary conserved cellular ets-2 gene (c-ets-2) encoding nuclear proteins related to the products of the c-ets proto-oncogene. *EMBO J.*, **7**, 697-705.

BOULUKOS K.E., P. POGNONEC, B. RABAULT, A. BEGUE and J. GHYSDAEL. (1989). Definition of an Ets1 protein domain required for nuclear localization in cells and DNA-binding activity *in vitro*. *Mol. Cell. Biol.*, **9**, 5718-5721.

BOULUKOS K.E., P. POGNONEC, E. SARIBAN, M. BAILLY, C. LAGROU and J. GHYSDAEL. (1990). Rapid and transient expression of Ets2 in mature macrophages following stimulation with cMGF, LPS, and PKC activators. *Genes Dev.*, **4**, 401-409.

BOVA C.A., J.C. OLSEN and R. SWANSTROM. (1988). The avian retrovirus env gene family : molecular analysis of host range and antigenic variants. *J. Virol.*, **62**, 75-83.

BREEDEN L. and K. NASMYTH. (1987). Similarity between cell-cycle genes of budding yeast and fission yeast and the Notch gene of *Drosophila*. *Nature*, **329**, 651-654.

BROWN T.A. and S.L. McKNIGHT. (1992). Specificities of protein-protein and protein-DNA interaction of GABP α and two newly defined ets-related proteins. *Genes Dev.*, **6**, 2502-2512.

BRUDER J.T., G. HEIDECKER and U.R. RAPP. (1992). Serum-, TPA-, and ras-induced expression from Ap-1/Ets-driven promoters requires Raf-1 kinase. *Genes Dev.*, **6**, 545-556.

BRUDER J.T., J. WESKE, T. BROWN, J.M. KYRIAKIS, J. AVRUCH and U.R. RAPP. (1993). GABP β interacts with P65 NF- κ B and activates HIV LTR expression in a raf-1 dependent fashion. Cold Spring Harbor meeting on cancer cells, 40.

BURATOWSKI S., S. HAHN, L. GUARENTE and P.A. SHARP. (1989). Five intermediate complexes in transcription initiation by RNA polymerase II. *Cell*, **56**, 549-561.

BURATOWSKI S. and H. ZHOU. (1992a). Transcription factor IID mutants defective for interaction with transcription factor IIA. *Science*, **255**, 1130-1132.

BURATOWSKI S. and H. ZHOU. (1992b). A suppressor of TBP mutations encodes an RNA polymerase III transcription factor with homology with TFIIB. *Cell*, **71**, 221-230.

BURDETT L.A., S. QI, Z.Q. CHEN, J.A. LAUTENBERGER and T.S. PAPAS. (1992). Characterization of the cDNA sequences of two *Xenopus* ets-2 proto-oncogenes. *Nucleic Acids Res.*, **20**, 371.

BURTIS K.C., C.S. THUMMEL, C. WELDON JONES, F.D. KARIM and D.S. HOGNESS. (1990). The *Drosophila* 74EF early puff contains E74, a complex ecdysone-inducible gene that encodes two ets-related proteins. *Cell*, **61**, 85-99.

BZIK D.J. and C.M. PRESTON. (1986). Analysis of DNA sequences which regulate the transcription of herpes simplex virus immediate early gene 3 : DNA sequences

required for enhancer-like activity and response to trans-activation by a virion polypeptide. *Nucleic Acids Res.*, **14**, 929-943.

CAVALLINI B., J. HUET, J.L. PLASSAT, A. SENTENAC, J.M. EGLY and P. CHAMBON. (1988). A yeast activity can substitute for the HeLa cell TATA box factor. *Nature*, **334**, 77-80.

CHANG L.M.S. (1971). Development of terminal nucleotidyl transferase in embryonic calf thymus gland. *Biochem. Biophys. Res. Commun.*, **44**, 124-131.

CHASMAN D.I., N.F. LUE, A.R. BUCHMAN, J.W. LAPOINTE, Y. LORCH and R.D. KORNBERG. (1990). A yeast protein that influences the chromatin structure of UAS_G and functions as a powerful auxiliary gene activator. *Genes Dev.*, **4**, 503-514.

CHASMAN D.I., K.M. FLAHERTY, P.A. SHARP and R.D. KORNBERG. (1993). Crystal structure of yeast TATA-binding protein and model for interaction with DNA. *Proc. Natl. Acad. Sci. USA*, **90**, 8174-8178.

CHEN H.M., H.L. PAHL, R.J. SCHEIBE, D.E. ZHANG and D.G. TENEN. (1993). The Sp1 transcription factor binds the CD11b promoter specifically in myeloid cells *in vivo* and is essential for myeloid-specific promoter activity. *J. Biol. Chem.*, **268**, 8230-8239.

CHEN J.H. (1985). The proto-oncogene *c-ets* is preferentially expressed in lymphoid cells. *Mol. Cell. Biol.*, **5**, 2993-3000.

CHEN J.H. (1988). Complementary DNA clones of chicken proto-oncogene *c-ets* : sequence divergence from the viral oncogene *v-ets*. *Onc. Res.*, **2**, 371-384.

CHEN J.H. (1993). PEA3, Oct 1 and Oct 2 are positive regulators for the ETS-1 promoter function. 9th annual meeting on oncogenes, Hood College, Frederick, 284.

CHEN J.H., S. JEHA and T. OKA. (1993). Negative regulatory elements in the human ETS1 gene promoter. *Oncogene*, **8**, 133-139.

CHEN T., M. BUNTING, F.D. KARIM and C.S. THUMMEL. (1992). Isolation and characterization of five *Drosophila* genes that encode an *ets*-related DNA binding domain. *Dev. Biol.*, **151**, 176-191.

CHEN Z.Q., L.A. BURDETT, A.K. SETH, J.A. LAUTENBERGER and T.S. PAPAS. (1990). Requirement of *ets-2* expression for *Xenopus* oocyte maturation. *Science*, **250**, 1416-1418.

CHUMAKOV A.M., D.L. CHEN, E.A. CHUMAKOVA and H.P. KOEFFLER. (1993). Localization of the *c-ets-2* transactivation domain. *J. Virol.*, **67**, 2421-2425.

CLARK N.M., M.J. SMITH, J.M. HILFINGER and D.M. MARKOVITZ. (1993). Activation of the human T-cell leukemia virus type I enhancer is mediated by binding sites for *elf-1* and the *pets* factor. *J. Virol.*, **67**, 5522-5528.

COCQUERELLE C., P. DAUBERSIES, M.A. MAJERUS, J.P. KERCKAERT and B. BAILLEUL. (1992). Splicing with inverted order of exons occurs proximal to large introns. *EMBO J.*, **11**, 1095-1098.

COCQUERELLE C., B. MASCREZ, D. HETUIN and B. BAILLEUL. (1993). Mis-splicing yields circular RNA molecules. *FASEB J.*, **7**, 155-160.

COLL J., M. RIGHI, C. DE TAISNE, C. DISSOUS, A. GEGONNE and D. STEHELIN. (1983). Molecular cloning of the avian acute transforming retrovirus MH2 reveals a novel cell-derived sequence (v-mil) in addition to the c-myc oncogene. *EMBO J.*, **2**, 2189-2194.

COMAI L., N. TANESE and R. TJIAN. (1992). The TATA-binding protein and associated factors are integral components of the RNA polymerase I transcription factor SL1. *Cell*, **68**, 965-976.

CORMACK B.P. and K. STRUHL. (1992). The TATA-binding protein is required for transcription by all three nuclear RNA polymerases in yeast cells. *Cell*, **69**, 685-696.

CORTES P., O. FLORES and D. REINBERG. (1992). Factors involved in specific transcription by mammalian RNA polymerase II : purification and analysis of transcription factor IIA and identification of transcription factor IIJ. *Mol. Cell. Biol.*, **12**, 413-421.

CRABTREE G.R. (1989). Contingent genetic regulatory events in T lymphocyte activation. *Science*, **243**, 355-361.

CRESS W.D. and S.J. TRIEZENBERG. (1991). Critical structural elements of the VP16 transcriptional activation domain. *Science*, **251**, 87-90.

CROSTON G.E., P.J. LAYBOURN, S.M. PARANJAPE and J.T. KADONAGA. (1992). Mechanism of transcriptional antirepression by GAL4-VP16. *Genes Dev.*, **6**, 2270-2281.

DALTON S. and R. TREISMAN. (1992). Characterization of SAP-1, a protein recruited by serum response factor to the c-fos serum response element. *Cell*, **68**, 597-612.

DANG C.V. and W.M.F. LEE. (1988). Identification of the human c-myc protein nuclear translocation signal. *Mol. Cell. Biol.*, **8**, 4048-4054.

DAVIDSON I., S. CHAUDHARY, C. BROU, J. WU, P. CHAMBON and L. TORA. (1993). Distinct TFIID complexes mediate the effect of different transcriptional activators. Cold Spring Harbor meeting on cancer cells, 61.

DELATTRE O., J. ZUCMAN, B. PLOUGASTEL, C. DESMAZE, T. MELOT, M. PETER, H. KOVAR, I. JOUBERT, P. de JONG, G. ROULEAU, A. AURIAS and G. THOMAS. (1992). Gene fusion with an ETS DNA-binding domain caused by chromosome translocation in human tumours. *Nature*, **359**, 162-165.

DELCAIRE A.X., A. FIANDINO, F. LYAMANI, M. BAREL, R. FRADE. (1989). Enhancement of Epstein-Barr virus/C3d receptor (EBV/C3dR or CR2) and nuclear p120 ribonuclear protein phosphorylation by specific EBV/C3dR ligands in subcellular fractions of the human B lymphoma cell line Raji. *Biochem. Biophys. Res. Com.*, **159**, 1213-1220.

DESBIENS X., C. QUEVA, T. JAFFREDO, D. STEHELIN and B. VANDENBUNDER. (1991). The relationship between cell proliferation and the transcription of the nuclear oncogenes c-myc, c-myb and c-ets-1 during feather morphogenesis in the chick embryo. *Development*, **111**, 699-713.

DE TAISNE C., A. GEGONNE, D. STEHELIN, A. BERNHEIM and R. BERGER. (1984). Chromosomal localization of the human proto-oncogene c-ets. *Nature*, **310**, 581-583.

DIAZ M.O., M.M. LE BEAU, P. PITHA and J.D. ROWLEY. (1986). Interferon and c-ets-1 genes in the translocation (9 ; 11) (p22 ; q23) in human acute monocytic leukemia. *Science*, **231**, 265-267.

DOMENGET C., D. LEPRINCE, B. PAIN, S. PEYROL, R.P. LI, D. STEHELIN, J. SAMARUT and P. JURDIC. (1992). The various domains of v-myb and v-ets oncogenes of E26 retrovirus contribute differently, but cooperatively, in transformation of hematopoietic lineages. *Oncogene*, **7**, 2231-2241.

DUDEK H., R.V. TANTRAVAHU, V.N. RAO, E.S.P. REDDY and E.P. REDDY. (1992). Myb and ets proteins cooperate in transcriptional activation of the mim-1 promoter. *Proc. Natl. Acad. Sci. USA*, **89**, 1291-1295.

DURBAN E.M. and D. BOETTIGER. (1981). Differential effects of transforming avian RNA tumor viruses on avian macrophages. *Proc. Natl. Acad. Sci. USA*, **78**, 3600-3604.

DUTERQUE-COQUILLAUD M., D. LEPRINCE, A. FLOURENS, C. HENRY, J. GHYSDAEL, B. DEBUIRE and D. STEHELIN. (1988). Cloning and expression of the chicken p54^{c-ets} cDNAs : the first p54^{c-ets} coding exon is located into the 40.0 kbp genomic domain unrelated to v-ets. *Onc. Res.*, **2**, 335-344.

DUTERQUE-COQUILLAUD M., C. NIEL, S. PLAZA and D. STEHELIN. (1993). New human erg isoforms generated by alternative splicing are transcriptional activators. *Oncogene*, **8**, 1865-1873.

DYNLACHT B.D., T. HOEY and R. TJIAN. (1991). Isolation of coactivators associated with the TATA-binding protein that mediate transcriptional activation. *Cell*, **66**, 563-576.

ERNST P., K. HAHM and S.T. SMALE. (1993). Both Lyf-1 and an Ets protein interact with a critical promoter element in the murine terminal transferase gene. *Mol. Cell. Biol.*, **13**, 2982-2992.

EVANS M.E. and S.M. HOLLENBERG. (1988). Zinc fingers : gilt by association. *Cell*, **52**, 1-3.

FERRE D'AMARE A.R., G.C. PRENDERGAST, E.B. ZIFF and S.K. BURLEY. (1993). Recognition by max of its cognate DNA through a dimeric b/HLH/Z domain. *Nature*, **363**, 38-45.

FISHER C.L., J. GHYSDAEL and J.C. CAMBIER. (1991). Ligation of membrane Ig leads to calcium-mediated phosphorylation of the proto-oncogene product, Ets-1. *J. Immunol.*, **146**, 1743-1749.

FISHER R.J., G. MAVROTHALASSITIS, A. KONDOH and T.S. PAPAS. (1991). High-affinity DNA-protein interactions of the cellular ETS1 protein: the determination of the ETS binding motif. *Oncogene*, **6**, 2249-2254.

FISHER R.J., S. KOIZUMI, A. KONDOH, J.M. MARIANO, G. MAVROTHALASSITIS, N.K. BHAT and T.S. PAPAS. (1992). Human ETS1 oncoprotein : purification, isoforms, -SH modification and DNA sequence-specific binding. *J. Biol. Chem.*, **267**, 17957-17965.

FLEISHMAN L.F., A.M. PILARO, K. MURAKAMI, A. KONDOH, R.J. FISHER and T.S. PAPAS. (1993). c-ets-1 protein is hyperphosphorylated during mitosis. *Oncogene*, **8**, 771-780.

FRYKBERG L., T. METZ, G. BRADY, M. INTRONA, H. BEUG, B. VENNSTROM and T. GRAF. (1988). A point mutation in the DNA-binding domain of the v-myb oncogene of E26 virus confers temperature sensitivity for transformation of myelomonocytic cells. *Onc. Res.*, **3**, 313-322.

FUJIWARA S., R.J. FISHER, N.K. BHAT, S. MORENO DIAZ de la ESPINA and T.S. PAPAS. (1988). A short-lived nuclear phosphoprotein encoded by the human ets-2 proto-oncogene is stabilized by activation of protein kinase C. *Mol. Cell. Biol.*, **8**, 4700-4706.

GALSON D.L., J.O. HENSOLD, T.R. BISHOP, M. SCHALLING, A.D. D'ANDREA, C. JONES, P.E. AURON and D.E. HOUSMAN. (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.

GARRIDO C., D. LEPRINCE, J.S. LIPSICK, D. STEHELIN, D. GOSPODAROWICZ and S. SAULE. (1992). Definition of functional domains in P135gag-myb-ets and p48^{v-myb} proteins required to maintain the response of neuroretina cells to basic fibroblast growth factor. *J. Virol.*, **66**, 160-166.

GARVIN A.M., S. PAWAR, J.D. MARTH and R.M. PERLMUTTER. (1988). Structure of the murine lck gene and its rearrangement in a murine lymphoma cell line. *Mol. Cell. Biol.*, **8**, 3058-3064.

GEGONNE A., D. LEPRINCE, M. DUTERQUE-COQUILLAUD, B. VANDENBUNDER, A. FLOURENS, J. GHYSDAEL, B. DEBUIRE and D. STEHELIN. (1987). Multiple domains for the chicken cellular sequences homologous to the v-ets oncogene of the E26 retrovirus. *Mol. Cell. Biol.*, **7**, 806-812.

GEGONNE A., B. PUNYAMMALEE, B. RABAULT, R. BOSSELUT, S. SENECA, M. CRABEEL and J. GHYSDAEL. (1992). Analysis of the DNA binding and transcriptional activation properties of the Ets1 oncoprotein. *New Biol.*, **4**, 512-519.

GEGONNE A., R. BOSSELUT, R.A. BAILLY and J. GHYSDAEL. (1993). Synergistic activation of the HTLV1 LTR Ets-responsive region by transcription factors Ets1 and Spl. *EMBO J.*, **12**, 1169-1178.

GEHALEN R.L. and M.L. HARRISON. (1984). Induction of a substrate for casein kinase II during lymphocyte mitogenesis. *Biochim. Biophys. Acta*, **804**, 169-

GHYSDAEL J., A. GEGONNE, P. POGNONEC, K. BOULUKOS, D. LEPRINCE, D. DERNIS, C. LAGROU and D. STEHELIN. (1986a). Identification in chicken macrophages of a set of proteins related to but distinct from the chicken cellular c-ets-encoded protein p54^{c-ets}. *EMBO J.*, **5**, 2251-2256.

GHYSDAEL J., A. GEGONNE, P. POGNONEC, D. DERNIS, D. LEPRINCE and D. STEHELIN. (1986b). Identification and preferential expression in thymic and bursal lymphocytes of a c-ets oncogene-encoded Mr 54,000 cytoplasmic protein. *Proc. Natl. Acad. Sci. USA*, **83**, 1714-1718.

GILLE H., A.D. SHARROCKS and P. SHAW. (1992). Phosphorylation of p62^{TCF} by MAP kinase stimulates ternary complex formation at c-fos promoter. *Nature*, **358**, 414-417.

GILMOUR D.S., G.H. THOMAS and S.C.R. ELGIN. (1989). *Drosophila* nuclear proteins bind to the regions of alternating C and T residues in gene promoters. *Science*, **245**, 1487-1490.

GITLIN S.D., R. BOSSELUT, A. GEGONNE, J. GHYSDAEL and J.N. BRADY. (1991). Sequence-specific interaction of the Ets1 protein with the long terminal repeat of the human T-lymphotropic virus type 1. *J. Virol.*, **65**, 5513-5523.

GODING C.R. and P. O'HARE. (1989). Herpes Simplex virus Vmw65-octamer binding protein interaction : a paradigm for combinatorial control of transcription. *Virology*, **173**, 363-367.

GOLAY J., M. INTRONA and T. GRAF. (1988). A single point mutation in the v-ets oncogene affects both erythroid and myelomonocytic cell differentiation. *Cell*, **55**, 1147-1158.

GOLDFARB D.S., J. GARIEPY, G. SCHOOLNICK and R.D. KORNBERG. (1986). Synthetic peptides as nuclear localization signals. *Nature*, **322**, 641-644.

GRAF T., K. McNAGNY, G. BRADY and J. FRAMPTON. (1992). Chicken "erythroid" cells transformed by the gag-myb-ets-encoding E26 leukemia virus are multipotent. *Cell*, **70**, 201-213.

GRANT P.A., V. ARULAMPALAM, L. AHRLUND-RICHTER and S. PETTERSSON. (1992). Identification of Ets-like lymphoid specific element within the immunoglobulin heavy chain 3' enhancer. *Nucleic Acids Res.*, **20**, 4401-4408.

GUARENTE L. and O. BERMINGHAM-McDONOGH. (1992). Conservation and evolution of transcriptional mechanisms in eukaryotes. *Trends Genet.*, **8**, 27-32.

GUNTHER C.V., J.A. NYE, R.S. BRYNER and B.J. GRAVES. (1990). Sequence-specific DNA binding of the proto-oncoprotein ets-1 defines a transcriptional activator sequence within the long terminal repeat of the Moloney murine sarcoma virus. *Genes Dev.*, **4**, 667-679.

GUTMAN A. and B. WASYLYK. (1990). The collagenase gene promoter contains a TPA and oncogene-responsive unit encompassing the PEA3 and AP-1 binding sites. *EMBO J.*, **9**, 2241-2246.

GUTMAN A. and B. WASYLYK. (1991). Nuclear targets for transcription regulation by oncogenes. *Trends Genet.*, **7**, 49-54.

HA I., W.S. LANE and D. REINBERG. (1991). Cloning of a human gene encoding the general transcription initiation factor IIB. *Nature*, **352**, 689-695.

HAGEMEIER C., A.J. BANNISTER, A. COOK and T. KOUZARIDES. (1993). The activation domain of transcription factor PU.1 binds the retinoblastoma (RB) protein and the transcription factor TFIID *in vitro* : RB shows sequence similarity to TFIID and TFIIB. *Proc. Natl. Acad. Sci. USA*, **90**, 1580-1584.

HAGMAN J. and R. GROSSCHELD. (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., S. BURATOWSKI, P.A. SHARP and L. GUARENTE. (1989). Isolation of the gene encoding the yeast TATA binding protein TFIID : a gene identical to the SPT15 suppressor of the Ty element insertions. *Cell*, **58**, 1173-1181.

HAHN S. (1993). Structure and function of acidic transcription activators. *Cell*, **72**, 481-483.

HANVEY J.C., M. SHIMIZU and R.D. WELLS. (1988). Intramolecular DNA triplexes in supercoiled plasmids. *Proc. Natl. Acad. Sci. USA*, **85**, 6292-6296.

HASKILL S., A.A. BEG, S.M. TOMPKINS, J.S. MORRIS, A.D. YUROCHKO, A. SAMPSON-JOHANNES, K. MONDAL, P. RALPH and A.S. BALDWIN Jr. (1991). Characterization of an immediate-early gene induced in adherent monocytes that encodes I κ B-like activity. *Cell*, **65**, 1281-1289.

HATAKEYAMA M., T. KONO, N. KOBAYASHI, A. KAWAHARA, S.D. LEVIN, R.M. PERLMUTTER and T. TANIGUSHI. (1991). Interaction of the IL-2 receptor with the src-family kinase p56^{lck} : identification of novel intermolecular association. *Science*, **252**, 1523-1528.

HERMANS M.H.A., H. HARTSUIKER and D. OPSTELTEN. (1989). An *in situ* study of B-lymphopoiesis in rat bone marrow : topographical arrangement of terminal nucleotidyl transferase-positive cells and pre-B cells. *J. Immunol.*, **142**, 67-73.

HERNANDEZ N. (1993). TBP, a universal eukaryotic transcription factor ? *Genes Dev.*, **7**, 1291-1308.

HERRERA R.E., P.E. SHAW and A. NORDHEIM. (1989). Occupation of the c-fos serum response element *in vivo* by a multiprotein complex is unaltered by growth factor induction. *Nature*, **340**, 68-70.

HILL C.S., R. MARAIS, S. JOHN, J. WYNNE, S. DALTON and R. TREISMAN. (1993). Functional analysis of a growth factor-responsive transcription factor complex. *Cell*, **73**, 395-406.

HINO M., A. TOJO, Y. MISAWA, H. MORII, F. TAKAKU and M. SHIBUYA. (1991). Unregulated expression of the erythropoietin receptor gene caused by insertion of spleen focus-forming virus long terminal repeat in a murine erythroleukemia cell line. *Mol. Cell. Biol.*, **11**, 5527-5533.

HIPSKIND R.A., V.N. RAO, C.G.F. MUELLER , E.S.P. REDDY and A. NORDHEIM. (1991). Ets-related protein elk-1 is homologous to the c-fos regulatory factor p62^{TCF}. *Nature*, **354**, 531-534.

HISATAKE K., S. HASEGAWA, R. TAKADA, Y. NAKATANI, M. HORIKOSHI and R.G. ROEDER. (1993). The p250 subunit of native TATA box-binding factor TFIID is the cell-cycle regulatory protein CCG1. *Nature*, **362**, 179-181.

HO I.C., N.K. BHAT, L.R. GOTTSCHALK, T. LINDSTEN, C.B. THOMPSON, T.S. PAPAS and J.M. LEIDEN. (1990). Sequence-specific binding of human Ets-1 to the T cell receptor α gene enhancer. *Science*, **250**, 814-818.

HO I.C. et J.M. LEIDEN. (1990a). Regulation of the human T-cell receptor α gene enhancer : multiple ubiquitous and T-cell-specific nuclear proteins interact with four hypomethylated enhancer elements. *Mol. Cell. Biol.*, **10**, 4720-4727.

HO I.C. et J.M. LEIDEN. (1990b). The T α 2 nuclear protein binding site from the human T cell receptor α enhancer functions as both a T cell-specific transcriptional activator and repressor. *J. Exp. Med.*, **172**, 1443-1449.

HOEY T., B.D. DYNLACHT, M.G. PETERSON, B.F. PUGH and R. TJIAN. (1990). Isolation and characterization of the *Drosophila* gene encoding the TATA box binding protein TFIID. *Cell*, **61**, 1179-1186.

HOEY T., R.O.J. WEINZIERL, G. GILL, J.L. CHEN, B.D. DYNLACHT and R. TJIAN. (1993). Molecular cloning and functional analysis of *Drosophila* TAF110 reveal properties expected of coactivators. *Cell*, **72**, 247-260.

HOFFMAN A., M. HORIKOSHI, C.K. WANG, S. SCHROEDER, P.A. WEIL and R.G. ROEDER. (1990). Cloning of the *Schizosaccharomyces pombe* TFIID gene reveals a strong conservation of functional domains present in *Saccharomyces cerevisiae* TFIID. *Genes Dev.*, **4**, 1141-1148.

HORIKOSHI M., C.K. WANG, H. FUJII, J.A. CROMLISH, P.A. WEIL and R.G. ROEDER. (1989). Cloning and structure of a yeast gene encoding a general transcription initiation factor TFIID that binds to the TATA box. *Nature*, **341**, 299-303.

HORIKOSHI N., K. MAGUIRE, A. KRALLI, E. MALDONADO, D. REINBERG and R. WEINMANN. (1991). Direct interaction between adenovirus E1A protein and the TATA box binding transcription factor IID. *Proc. Natl. Acad. Sci. USA*, **88**, 5124-5128.

IMAGAWA M., R. CHIU and M. KARIN. (1987). Transcription factor AP-2 mediates induction by two different signal-transduction pathways : protein kinase C and cAMP. *Cell*, **51**, 251-260.

IVANOV X., Z. MLADENOV, S. NEDYALKOV and T.G. TODOROV. (1962). Experimental investigations into avian leukosis. I. Transmission experiments of certain diseases of the avian leukosis complex, found in Bulgaria. *Bulgarian Acad. Sci. Bull. Inst. Pathol. Comp. Animaux Domestiques*, **9**, 5-36.

JACKSON R.J. and N. STANDART. (1990). Do the poly(A) tail and 3' untranslated region control mRNA translation ? *Cell*, **62**, 15-24.

JACKSON S.P. and R. TJIAN. (1988). O-glycosylation of eukaryotic transcription factors : implications for the mechanisms of transcriptional regulation. *Cell*, **55**, 125-133.

JAIN J., P.G. McCAFFREY, V.E. VALGE-ARCHER and A. RAO. (1992). Nuclear factor of activated T cells contains Fos and Jun. *Nature*, **356**, 801-804.

JANKNECHT R. and A. NORDHEIM. (1992). Elk-1 protein domains required for direct and SRF-assisted DNA-binding. *Nucleic Acids Res.*, **20**, 3317-3324.

JANSEN H.W., B. RÜCKERT, R. LURZ and K. BISTER. (1983). Two unrelated cell-derived sequences in the genome of avian leukemia and carcinoma inducing retrovirus MH2. *EMBO J.*, **2**, 1969-1975.

JONAT C., H.J. RAHMSDORF, K.K. PARK, A.C.B. CATO, S. GEBEL, H. PONTA and P. HERRLICH. (1990). Antitumor promotion and antiinflammation : down-modulation of AP-1 (fos/jun) activity by glucocorticoid hormone. *Cell*, **62**, 1189-1204.

JONES N. (1990). Transcriptional regulation by dimerization : two sides of an incestuous relationship. *Cell*, **61**, 9-11.

JORCYK C.L., D.K. WATSON, G.J. MAVROTHALASSITIS and T.S. PAPAS. (1991). The human ETS1 gene : genomic structure, promoter characterization and alternative splicing. *Oncogene*, **6**, 523-532.

JURDIC P., M. BENCHAIIBI, O. GANDRILLON and J. SAMARUT. (1987). Transforming and mitogenic effects of avian leukemia virus E26 on chicken hematopoietic cells and fibroblasts, respectively, correlate with level of expression of the provirus. *J. Virol.*, **61**, 3058-3065.

KAHN N.C., C.S. FLORDELLIS, C.F. CARON, P.H. DUESBERG and T.S. PAPAS. (1983). Avian carcinoma virus MH2 contains a transformation-specific sequence, mht and shares the myc sequence with MC29, CMII and OK10 viruses. *Proc. Natl. Acad. Sci. USA*, **80**, 6566-6570.

KALLENBACH S., N. DOYEN, M. FANTON d'ANTON and F. ROUGEON. (1992). Three lymphoid-specific factors account for all junctional diversity characteristics of somatic assembly of T-cell receptor and immunoglobulin genes. *Proc. Natl. Acad. Sci. USA*, **89**, 2799-2803.

KAO C.G., P.M. LIEBERMAN, M.C. SCHMIDT, Q. ZHOU, R. PEI and A.J. BERK. (1990). Cloning of a transcriptionally active human TATA binding factor. *Science*, **248**, 1646-1650.

KARIM F.D., L.D. URNESS, C.S. THUMMEL, M.J. KLEMSZ, S.R. McKERCHER, A. CELADA, C. VAN BEVEREN, R.A. MAKI, C.V. GUNTHER, J.A. NYE and B.J. GRAVES. (1990). The ETS-domain : a new DNA-binding motif that recognizes a purine-rich core DNA sequence. *Genes Dev.*, **4**, 1451-1453.

KARIM F.D. and C.S. THUMMEL. (1991). Ecdysone coordinates the timing and amounts of E74A and E74B transcription in *Drosophila*. *Genes Dev.*, **5**, 1067-1079.

KARIM F.D. and C.S. THUMMEL. (1992). Temporal coordination of regulatory gene expression by the steroid hormone ecdysone. *EMBO J.*, **11**, 4083-4093.

KATAN M., A. HAIGH, C.P. VERRIJZER, P.C. van der VLIET and P. O'HARE. (1990). Characterization of a cellular factor which interacts functionally with Oct-1 in the assembly of a multicomponent transcription complex. *Nucleic Acids Res.*, **18**, 6871-6880.

KIM J.L., D.B. NIKOLOV and S.K. BURLEY. (1993). Co-crystal structure of TBP recognizing the minor groove of a TATA element. *Nature*, **365**, 520-527.

KIM Y., J.H. GEIGER, S. HAHN and P.B. SIGLER. (1993). Crystal structure of a yeast TBP/TATA-box complex. *Nature*, **365**, 512-520.

KISIELOW P. and H. von BOEHMER. (1990). Negative and positive selection of immature thymocytes : timing and the role of ligand for $\alpha\beta$ T cell receptor. *Sem. Immunol.*, **2**, 35-44.

KISSINGER C.R., B. LIU, E. MARTIN-BIANCO, T.B. KORNBERG and C.O. PABO. (1990). Crystal structure of an engrailed homeodomain-DNA complex at 2.8 Å resolution : a framework for understanding homeodomain-DNA interactions. *Cell*, **63**, 579-590.

KLAMBT C. (1993). The *Drosophila* gene *pointed* encodes two ETS-like proteins which are involved in the development of the midline glial cells. *Development*, **117**, 163-176.

- KLEMPNAUER K.M., G. SYMONDS, G.I. EVANS and J.M. BISHOP. (1984). Subcellular localization of proteins encoded by oncogenes of avian myeloblastosis virus and avian leukemia virus E26 and by the chicken *c-myb* gene. *Cell*, **37**, 537-547.
- KLEMSZ M.J., S.R. McKERCHER, A. CELADA, C. VAN BEVEREN and R.A. MAKI. (1990). The macrophage and B cell-specific transcription factor PU.1 is related to the *ets* oncogene. *Cell*, **61**, 113-124.
- KOBEL H.R. and L. DU PASQUIER. (1986). Genetics of polyploid *Xenopus*. *Trends Genet.*, **2**, 310-315.
- KOIZUMI S., R.J. FISHER, S. FUJIWARA, C. JORCYK, N.K. BHAT, A. SETH and T.S. PAPAS. (1990). Isoforms of the human *ets-1* protein : generation by alternative splicing and differential phosphorylation. *Oncogene*, **5**, 675-681.
- KRAULIS P.J., A.R.C. RAINE, P.L. GADHAVI and E.D. LAUE. (1990). Structure of the DNA-binding domain of the zinc GAL4. *Nature*, **356**, 448-50.
- KUPPUSWAMY M., T. SUBRAMANIAN, A. SRINIVASAN and G. CHINNADURAI. (1989). Multiple functional domains of Tat, the trans-activator of HIV-1, defined by mutational analysis. *Nucleic Acids Res.*, **17**, 3551-3561.
- LAI Z.C. and G.M. RUBIN. (1992). Negative control of photoreceptor development in *Drosophila* by the product of the *yan* gene, an ETS domain protein. *Cell*, **70**, 609-620.
- LAMARCO K. and S.L. McKNIGHT. (1989). Purification of a set of cellular polypeptides that bind to the purine-rich cis-regulatory element of herpes simplex virus immediate early genes. *Genes Dev.*, **3**, 1372-1383.
- LAMARCO K., C.C. THOMPSON, B.P. BYERS, E.M. WALTON and S.L. McKNIGHT. (1991). Identification of Ets- and Notch-related subunits in GA binding protein. *Science*, **253**, 789-792.
- LANDAU N.R., T.P. St JOHN, I.L. WEISSMAN, S.C. WOLF, A.E. SILVERSTONE and D. BALTIMORE. (1984). Cloning of terminal transferase cDNA by antibody screening. *Proc. Natl. Acad. Sci. USA*, **81**, 5836-5840.
- LANDSCHULTZ W., P.F. JOHNSON and S.L. McKNIGHT. (1988). The leucine zipper : a hypothetical structure common to a new class of DNA binding proteins. *Science*, **240**, 1759-1764.
- LANGER S.J., D.M. BORTNER, M.F. ROUSSEL, C.J. SHERR and M.C. OSTROWSKI. (1992). Mitogenic signaling by colony-stimulating factor 1 and *ras* is suppressed by the *ets-2* DNA-binding domain and restored by *myc* overexpression. *Mol. Cell. Biol.*, **12**, 5355-5362.
- LAUDET V., C. NIEL, M. DUTERQUE-COQUILLAUD, D. LEPRINCE and D. STEHELIN. (1993). Evolution of the *ets* gene family. *Biochem. Biophys. Res. Com.*, **190**, 8-14.
- LAUTENBERGER J.A., L.A. BURDETT, M.A. GUNNELL, S. QI, D.K. WATSON, S.J. O'BRIEN and T.S. PAPAS. (1992). Genomic dispersal of the *ets* gene family during metazoan evolution. *Oncogene*, **7**, 1713-1719.
- LAUTENBERGER J.A. and T.S. PAPAS. (1993). Inversion of a chicken *ets-1* proto-oncogene segment in avian leukemia virus E26. *J. Virol.*, **67**, 610-612.

- LAW S.K., J. GAGNON, J.E. HILDRETH, C.E. WELLS, A.C. WILLIS and A.J. WONG. (1987). The primary structure of the β -subunit of the cell surface adhesion glycoproteins LFA-1, CR3 and p150,95 and its relationship to the fibronectin receptor. *EMBO J.*, **6**, 915-919.
- LEE D.K., J. DEJONG, S. HASHIMOTO, M. HORIKOSHI and R. ROEDER. (1992). TFIIA induces conformational changes in TFIID via interactions with the basic repeat. *Mol. Cell. Biol.*, **12**, 5189-5196.
- LEES J.A., S.E. FAWELL and M.G. PARKER. (1989). Identification of two transactivation domains in the mouse oestrogen receptor. *Nucleic Acids Res.*, **17**, 5477-5488.
- LEIDEN J.M. (1992). Transcriptional regulation during T-cell development : the α TCR gene as a molecular model. *Immunol. Today*, **13**, 22-30.
- LEIDEN J.M., C.Y. WANG, B. PETRYNIAK, D.M. MARKOVITZ, G.J. NABEL and C.B. THOMPSON. (1992). A novel Ets-related transcription factor, elf-1, binds to human immunodeficiency virus type 2 regulatory elements that are required for inducible trans activation in T cells. *J. Virol.*, **66**, 5890-5897.
- LEIDEN J.M. (1993). Transcriptional regulation of T cell receptor genes. *Ann. Rev. Immunol.*, **11**, 539-570.
- LEMARCHANDEL V., J. GHYSDAEL, V. MIGNOTTE, C. RAHUEL and P.H. ROMEO. (1993). GATA and Ets cis-acting sequences mediate megakaryocyte-specific expression. *Mol. Cell. Biol.*, **13**, 668-676.
- LEPRINCE D., A. GEGONNE, J. COLL, C. DE TAISNE, A. SCHNEEBERGER, C. LAGROU and D. STEHELIN. (1983). A putative second cell-derived oncogene of the avian leukaemia retrovirus E26. *Nature*, **306**, 395-397.
- LEPRINCE D., M. DUTERQUE-COQUILLAUD, R.P. LI, C. HENRY, A. FLOURENS, B. DEBUIRE and D. STEHELIN. (1988). Alternative splicing within the chicken c-ets-1 locus : implications for transduction within the E26 retrovirus of the c-ets proto-oncogene. *J. Virol.*, **62**, 3233-3241.
- LEPRINCE D., J.C. GESQUIERE and D. STEHELIN. (1990). The chicken cellular progenitor of the v-ets oncogene, p68^{c-ets-1}, is a nuclear DNA-binding protein not expressed in lymphoid cells of the spleen. *Onc. Res.*, **5**, 255-265.
- LEPRINCE D., P. CREPIEUX and D. STEHELIN. (1992). c-ets-1 DNA binding to the PEA3 motif is differentially inhibited by all the mutations found in v-ets. *Oncogene*, **7**, 9-17.
- LEPRINCE D., P. CREPIEUX, V. LAUDET, A. FLOURENS and D. STEHELIN. (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**, 855-857.
- LEUNG S., S. McCracken, J. GHYSDAEL and N.G. MIYAMOTO. (1993). Requirement of an ETS-binding element for transcription of the human lck type I promoter. *Oncogene*, **8**, 989-997.
- LEUTHER K.K., J.M. SALMERON and S.A. JOHNSTON. (1993). Genetic evidence that an activation domain of GAL4 does not require acidity and may form a β sheet. *Cell*, **72**, 575-585.

- LEVINE M. and T. HOEY. (1988). Homeobox proteins as sequence-specific transcription factors. *Cell*, **55**, 537-540.
- LI R.P., M. DUTERQUE-COQUILLAUD, C. LAGROU, B. DEBUIRE, T. GRAF, D. STEHELIN and D. LEPRINCE. (1989). A single amino-acid substitution in the DNA-binding domain of the myb oncogene confers a thermolabile phenotype to E26-transformed myeloid cells. *Onc. Res.*, **5**, 137-141.
- LIM F., N. KRAUT, J. FRAMPTON and T. GRAF. (1992). DNA binding by c-ets-1, but not v-ets is repressed by an intramolecular mechanism. *EMBO J.*, **11**, 643-652.
- LIN Y.S. and M.R. GREEN. (1991). Mechanism of action of an acidic transcriptional activator *in vitro*. *Cell*, **64**, 971-981.
- LIN Y.S., I. HA, E. MALDONADO, D. REINBERG and M.R. GREEN. Binding of general transcription factor TFIIB to an acidic activating region. (1991). *Nature*, **353**, 569-571.
- LUCIBELLO F.C., E.P. SLATER, K.U. JOOSS, M. BEATO and R. MULLER. (1990). Mutual transrepression of fos and the glucocorticoid receptor : involvement of a functional domain in fos which is absent in fosB. *EMBO J.*, **9**, 2827-2834.
- LUISI B.F., W.X. XU, Z. OTWINOWSKI, L.P. FREEDMAN, K.R. YAMAMOTO and P.B. SIGLER. (1991). Crystallographic analysis of the interaction of the glucocorticoid receptor with DNA. *Nature*, **352**, 497-505.
- LUX S.E., M. JOHN and V. BENNETT. (1990). Analysis of cDNA for human erythrocyte ankyrin indicates a repeated structure with homology to tissue-differentiation and cell-cycle control proteins. *Nature*, **344**, 36-42.
- MAJERUS M.A., F. BIBOLLET-RUCHE, J.B. TELLIEZ, B. WASYLYK and B. BAILLEUL. (1992). Serum, AP-1 and Ets-1 stimulate the human ets-1 promoter. *Nucleic Acids Res.*, **20**, 2699-2703.
- MARAIS R., J. WYNNE and R. TREISMAN. (1993). The SRF accessory protein elk-1 contains a growth factor-regulated transcriptional activation domain. *Cell*, **73**, 381-393.
- MAVROTHALASSITIS G.J., D.K. WATSON and T.S. PAPAS. (1990a). Molecular and functional characterization of the promoter of ETS2, the human c-ets-2 gene. *Proc. Natl. Acad. Sci. USA*, **87**, 1047-1051.
- MAVROTHALASSITIS G.J., D.K. WATSON and T.S. PAPAS. (1990b). The human ETS-2 gene promoter : molecular dissection and nuclease hypersensitivity. *Oncogene*, **5**, 1337-1342.
- MAVROTHALASSITIS G.J. and T.S. PAPAS. (1991). Positive and negative factors regulate the transcription of the ETS2 gene *via* an oncogene-responsive-like unit within the ETS2 promoter region. *Cell Growth and Diff.*, **2**, 215-224.
- MAY W.A., M.L. GISHIZKY, S.L. LESSNICK, L.B. LUNSFORD, B.C. LEWIS, O. DELATTRE, J. ZUCMAN, G. THOMAS and C.T. DENNY. (1993). Ewing sarcoma 11 ; 22 translocation produces a chimeric transcription factor that requires the DNA-binding domain encoded by FL11 for transformation. *Proc. Natl. Acad. Sci. USA*, **90**, 5752-5756.

NELSEN B., G. TIAN, B. ERMAN, J. GREGOIRE, R. MAKI, B. GRAVES and R. SEN. (1993). Regulation of lymphoid-specific immunoglobulin μ heavy chain gene enhancer by ETS-domain proteins. *Science*, **261**, 82-86.

NERLOV C., D. DE CESARE, F. PERGOLA, A. CARACCILO, F. BLASI, M. JOHNSEN and P. VERDE. (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**, 4573-4582.

NESS S.A., A. MARKNELL and T. GRAF. (1989). The v-myb oncogene product binds to and activates the promyelocyte-specific mim-1 gene. *Cell*, **59**, 1115-1125.

NGUYEN V.C., D. RAY, M.S. GROSS, M.F. DE TAND, F. FREZAL and F. MOREAU-GACHELIN. (1990). Localization of the human oncogene Spi-1 on chromosome 11, region p11.22. *Hum. Genet.*, **84**, 542-546.

NIBBS R.J.B., K. ITOH, W. OSTERTAG and P.R. HARRISON. (1993). Differentiation arrest and stromal cell-independent growth of murine erythroleukemia cells are associated with elevated expression of ets-related genes but not with mutation of p53. *Mol. Cell. Biol.*, **13**, 5582-5592.

NIGRO J.M., K.R. CHO, E.R. FEARON, S.E. KERN, J.M. RUPPERT, J.D. OLINER, K.W. KINZLER and B. VOGELSTEIN. (1991). Scrambled exons. *Cell*, **64**, 607-613.

NIKOLOV D.B., S.H. HU, J. LIN, A. GASCH, A. HOFFMANN, M. HORIKOSHI, N.H. CHUA, R.G. ROEDER and S.K. BURLEY. (1992). Crystal structure of TFIID TATA-box binding protein. *Nature*, **360**, 40-46.

NUNN M.F., P.H. SEEBURG, C. MOSCOVICI and P.H. DUESBERG. (1983). Tripartite structure of the avian erythroblastosis virus E26 transforming gene. *Nature*, **306**, 391-395.

NUNN M., H. WEIHER, P. BULLOCK and P. DUESBERG. (1984). Avian erythroblastosis virus E26 : nucleotide sequence of the tripartite onc gene and of the LTR, and analysis of the cellular prototype of the viral ets sequence. *Virology*, **139**, 330-339.

NUNN M.F. and T. HUNTER. (1989). The ets sequence is required for induction of erythroblastosis in chickens by avian retrovirus E26. *J. Virol.*, **63**, 398-402.

NYE J.A., J.M. PETERSEN, C.V. GUNTHER, M.D. JONSEN and B.J. GRAVES. (1992). Interaction of murine Ets-1 with GGA-binding sites establishes the ETS domain as a new DNA-binding motif. *Genes Dev.*, **6**, 975-990.

OHNO H., G. TAKIMOTO and W. McKEITHAN. (1990). The candidate proto-oncogene bcl-3 is related to genes implicated in cell lineage determination and cell cycle control. *Cell*, **60**, 991-997.

OKA T., A. RAIRKAR and J.H. CHEN. (1991). Structural and functional analysis of the regulatory sequences of the ets-1 gene. *Oncogene*, **6**, 2077-2083.

OMICHIANSKI J.G., G.M. CLORE, O. SCHAAD, G. FELSENFELD, C. TRAINOR, E. APPELLA, S.J. STAHL and A.M. GRONENBORN. (1993). NMR structure of a specific DNA complex of Zn-containing DNA binding domain of GATA-1. *Science*, **261**, 438-446.

- O'SHEA E.K., J.D. KLEMM, P.S. KIM and T. ALBER. (1991). X-ray structure of the GCN4 leucine zipper, a two stranded, parallel coiled coil. *Science*, **254**, 539-544.
- PAHL H.L., A.G. ROSMARIN and D.G. TENEN. (1992). Characterization of the myeloid-specific CD11b promoter. *Blood*, **79**, 865-870.
- PAHL H.L., R.J. SCHEIBE, D.E. ZHANG, H.M. CHEN, D.L. GALSON, R.A. MAKI and D.G. TENEN. (1993). The proto-oncogene PU.1 regulates expression of the myeloid-specific CD11b promoter. *J. Biol. Chem.*, **268**, 5014-5020.
- PAUL R., S. SCHUETZE, S.L. KOZAK, C.A. KOZAK and D. KABAT. (1991). The Sfp1-1 proviral integration site of Friend erythroleukemia encodes the ets-related transcription factor PU.1. *J. Virol.*, **65**, 464-467.
- PETERSON M.G., N. TANESE, B.F. PUGH and R. TJIAN. (1990). Functional domains and upstream activation properties of cloned human TATA binding protein. *Science*, **248**, 1625-1630.
- PIERCE J.H. (1989). Oncogenes, growth factors and hematopoietic cell transformation. *Biochim. Biophys. Acta.* **989**, 179-208.
- PINA B., U. BRUGGEMEIER and M. BEATO. (1990). Nucleosome positioning modulates accessibility of regulatory proteins to the mouse mammary tumor virus promoter. *Cell*, **60**, 719-731.
- PINTO I., D.E. WARE and M. HAMPSEY. (1992). The yeast SUA7 gene encodes a homolog of human transcription factor TFIIB and is required for normal start site selection. *Cell*, **68**, 977-988.
- POGNONEC P., K.E. BOULUKOS, J.C. GESQUIERE, D. STEHELIN and J. GHYSDAEL. (1988). Mitogenic stimulation of thymocytes results in the calcium-dependant phosphorylation of c-ets-1 proteins. *EMBO J.*, **7**, 977-983.
- POGNONEC P., K.E. BOULUKOS and J. GHYSDAEL. (1989). The c-ets-1 protein is chromatin associated and binds to DNA *in vitro*. *Oncogene*, **4**, 691-697.
- POGNONEC P., K.E. BOULUKOS, R. BOSSELUT, C. BOYER, A.M. SCHMITT-VERHULST and J. GHYSDAEL. (1990). Identification of a Ets1 variant protein unaffected in its chromatin and *in vitro* DNA binding capacities by T cell antigen receptor triggering and intracellular calcium rises. *Oncogene*, **5**, 603-610.
- PONGUBALA J.M.R. and M.L. ATCHISON. (1991). Functional characterization of the developmentally controlled immunoglobulin kappa 3' enhancer : regulation by Id, a repressor of helix-loop-helix transcription factors. *Mol. Cell. Biol.*, **11**, 1040-1047.
- PONGUBALA J.M.R., S. NAGULAPALLI, M.J. KLEMSZ, S.R. McKERCHER, R.A. MAKI and M.L. ATCHISON (1992). PU.1 recruits a second nuclear factor to a site important for immunoglobulin κ 3' enhancer activity. *Mol. Cell. Biol.*, **12**, 368-378.
- PONGUBALA J.M.R., C. VAN BEVEREN, S. NAGULAPALLI, M.J. KLEMSZ, S.R. McKERCHER, R.A. MAKI and M.L. ATCHISON. (1993). Effect of PU-1 phosphorylation on interaction with NF-EM5 and transcriptional activation. *Science*, **259**, 1622-1625.
- PRIBYL L.J., D.K. WATSON, M.J. McWILLIAMS, R. ASCIONE and T.S. PAPAS. (1988). The *Drosophila* ets-2 gene : molecular structure, chromosomal localization and developmental expression. *Dev. Biol.*, **127**, 45-53.

PRIBYL L.J., D.K. WATSON, R.A. SCHULZ and T.S. PAPAS. (1991). D-*elg*, a member of the *Drosophila ets* gene family : sequence, expression and evolutionary comparison. *Oncogene*, **6**, 1175-1183.

PROSSER H.M., D. WOTTON, A. GEGONNE, J. GHYSDAEL, S. WANG, N.A. SPECK and M.J. OWEN. (1992). A phorbol ester response element within the human T-cell receptor β -chain enhancer. *Proc. Natl. Acad. Sci. USA*, **89**, 9934-9938.

PUGH B.F. and R. TJIAN. (1990). Mechanism of transcriptional activation by Sp1 : evidence for coactivators. *Cell*, **61**, 1187-1197.

PUGH B.F. and R. TJIAN. (1991). Transcription from a TATA-less promoter requires a multisubunit TFIID complex. *Genes Dev.*, **5**, 1935-1945.

QUEVA C., S.A. NESS, F.A. GRASSER, T. GRAF, B. VANDENBUNDER and D. STEHELIN. (1992). Expression patterns of *c-myb* and of *v-myb* induced myeloid-1 (*mim-1*) gene during the development of the chick embryo. *Development*, **114**, 125-133.

QUEVA C., D. LEPRINCE, D. STEHELIN and B. VANDENBUNDER. (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.

RADKE K., H. BEUG, S. KORNFELD and T. GRAF. (1982). Transformation of both erythroid and myeloid cells by E26, an avian leukemia virus that contains the *myb* gene. *Cell*, **31**, 643-653.

RANDAK C., T. BRABLETZ, M. HERGENROTHER, I. SOBOTTA and E. SERFLING. (1990). Cyclosporin A suppresses the expression of the interleukin 2 gene by inhibiting the binding of lymphocyte-specific factors to the IL-2 enhancer. *EMBO J.*, **9**, 2529-2536.

RAO V.N., T.S. PAPAS and E.S.P. REDDY. (1987). *erg*, a human *ets*-related gene on chromosome 21 : alternative splicing, polyadenylation, and translation. *Science*, **237**, 635-639.

RAO V.N., W.S. MODI, H.D. DRABKIN, D. PATTERSON, S.J. O'BRIEN, T.S. PAPAS and E.S.P. REDDY. (1988). The human *erg* gene maps to chromosome 21, band q22 : relationship to the 8 ; 21 translocation of acute myelogenous leukemia. *Oncogene*, **3**, 497-500.

RAO V.N., K. HUEBNER, M. ISOBE, A. AR-RUSHDI, C.M. CROCE and E.S.P. REDDY. (1989). *elk*, tissue-specific *ets*-related genes on chromosomes X and 14 near translocation breakpoints. *Science*, **244**, 66-70.

RAO V.N. and E.S.P. REDDY. (1992). Elk-1 domains responsible for autonomous DNA binding, SRE : SRF interaction and negative regulation of DNA binding. *Oncogene*, **7**, 2335-2340.

RAO V.N., T. OHNO, D.D.K. PRASAD, G. BHATTACHARYA and E.S.P. REDDY. (1993). Analysis of the DNA-binding and transcriptional activation functions of human Fli-1 protein. *Oncogene*, **8**, 2167-2173.

RAO V.N. and E.S.P. REDDY. (1993). Δ elk-1, a variant of elk-1, fails to interact with the serum response factor and binds to DNA with modulated specificity. *Canc. Res.*, **53**, 215-220.

RAVEL-CHAPUIS P., D. LEPRINCE, B. PAIN, R. LI, C. DOMENGET, D. STEHELIN, J. SAMARUT and P. JURDIC. (1991). v-myb and v-ets cooperate for the mitogenic stimulation of primary fibroblasts by avian E26 retrovirus. *J. Virol.*, **65**, 3928-3931.

RAY D., R. BOSSELUT, J. GHYSDAEL, M.G. MATTEI, A. TAVITIAN and F. MOREAU-GACHELIN. (1992). Characterization of Spi-B, a transcription factor related to the putative oncoprotein Spi-1/PU.1. *Mol. Cell. Biol.*, **12**, 4297-4304.

REDDY E.S.P. and V.N. RAO. (1988). Structure, expression and alternative splicing of the human *c-ets-1* proto-oncogene. *Onc. Res.*, **3**, 239-246.

REDDY E.S.P. and V.N. RAO. (1990). Localization and modulation of the DNA-binding activity of the human *c-ets-1* protooncogene. *Canc. Res.*, **50**, 5013-5016.

REGIER J.L., F. SHEN and S.J. TRIEZENBERG. (1993). Pattern of aromatic and hydrophobic amino acids critical for one of two subdomains of the VP16 transcriptional activator. *Proc. Natl. Acad. Sci. USA*, **90**, 883-887.

RICHARDSON W.D., B.L. ROBERTS and A.E. SMITH. (1986). Nuclear localization signals in polyoma virus large-T. *Cell*, **44**, 77-85.

ROBERTS S.G.E., I. HA, E. MALDONADO, D. REINBERG and M.R. GREEN. (1993). Interaction between an acidic activator and transcription factor TFIIB is required for transcriptional activation. *Nature*, **363**, 741-747.

ROGERS S., R. WELLS and M. RECHSTEINE. (1986). Amino acids sequences common to rapidly degraded proteins : the PEST hypothesis. *Science*, **234**, 364-368.

RORTH P., C. NERLOV, F. BLASI and M. JONHSEN. (1990). Transcription factor PEA3 participates in the induction of urokinase plasminogen activator transcription in murine keratinocytes stimulated with epidermal growth factor or phorbol-ester. *Nucleic Acids Res.*, **18**, 5009-5017.

RORTH S.Y., M. SHIMIZU, L. JOHNSON, M. GRUNSTEIN and R.T. SIMPSON. (1992). Stable nucleosome positioning and complete repression by the yeast alpha 2 repressor are disrupted by amino-terminal mutations in histone H4. *Genes Dev.*, **6**, 411-425.

ROSMARIN A.G., S.C. WEIL, G.L. ROSNER, J.D. GRIFFIN, M.A. ARNAOUT and D.G. TENEN. (1989). Differential expression of CD11b/CD18 (Mo1) and myeloperoxidase genes during myeloid differentiation. *Blood*, **73**, 131-136.

ROUSSEL M., S. SAULE, C. LAGROU, C. ROMMENS, H. BEUG, T. GRAF and D. STEHELIN. (1979). Three new types of viral oncogenes of cellular origin specific for hematopoietic cell transformation. *Nature*, **281**, 452-455.

ROVIGATTI U., D.K. WATSON and J. J. YUNIS. (1986). Amplification and rearrangement of Hu-ets-1 in leukemia and lymphoma with involvement of 11q23. *Science*, **232**, 398-400.

RUBEN S., A. PERKINS, R. PURCELL, K. JOUNG, R. SIA, R. BUGHOFF, W.A. HASELTINE and C.A. ROSEN. (1989). Structural and functional characterization of human immunodeficiency virus tat protein. *J. Virol.*, **63**, 1-8.

RUDD C.E., J.M. TREVILLYAN, J.D. DASGUPTA, L.L. WONG and S.F. SCHLOSSMAN. (1988). The CD4 receptor is complexed in detergent lysates to a protein-tyrosine kinase (pp58) from human T lymphocytes. *Proc. Natl. Acad. Sci. USA*, **85**, 5190-5194.

RUPPERT S., E.H. WANG and R. TJIAN. (1993). Cloning and expression of human TAF_{II}250 : a TBP-associated factor implicated in cell-cycle regulation. *Nature*, **362**, 175-179.

RUSCETTI S., R.E AURIGEMMA, C.C. YUAN, S. SAWYER and D.G. BLAIR. (1992). Induction of erythropoietin responsiveness in murine hematopoietic cells by the gag-myb-ets-containing ME26 virus. *J. Virol.*, **66**, 20-26.

SACCHI N., D.K. WATSON, AD H.M. GUERTIS VAN KESSEL, A. HAGEMEIJER, J. KERSEY, H.D. DRABKIN, D. PATTERSON and T.S. PAPAS. (1986). Hu-ets-1 and Hu-ets-2 genes are transposed in acute leukemias with (4 ; 11) and (8 ; 21) translocations. *Science*, **231**, 379-382.

SAFFER J.D., S.P. JACKSON and M.B. ANNARELLA. (1991). Developmental expression of Sp1 in the mouse. *Mol. Cell. Biol.*, **11**, 2189-2199.

SALMON P., A. GIOVANE, B. WASYLYK and D. KLATZMANN. (1993). Characterization of the human CD4 gene promoter : transcription from the CD4 gene core promoter is tissue-specific and is activated by Ets proteins. *Proc. Natl. Acad. Sci. USA*, **90**, 7739-7743.

SAULE S., M. ROUSSEL, C. LAGROU and D. STEHELIN. (1981). Characterization of the oncogene (erb) of avian erythroblastosis virus and its cellular progenitor. *J. Virol.*, **38**, 409-419.

SCHMIDT M.C., C.C. KAO, R. PEI and A.J. BERK. (1989). Yeast TATA box transcription factor gene. *Proc. Natl. Acad. Sci. USA*, **86**, 7785-7789.

SCHAUFLE F., B.L. WEST and T.L. REUDELHUBER. (1990). Overlapping Pit-1 and Sp1 binding sites are both essential to full rat growth hormone gene promoter activity despite mutually exclusive Pit-1 and Sp1 binding. *J. Biol. Chem.*, **265**, 17189-17196.

SCHNEIKERT J., Y. LUTZ and B. WASYLYK. (1992). Two independent activation domains in c-ets-1 and c-ets-2 located in non-conserved sequences of the ets gene family. *Oncogene*, **7**, 249-256.

SCHUETZE S., R. PAUL, B.C. GLINIAK and D. KABAT. (1992). Role of the PU.1 transcription factor in controlling differentiation of Friend erythroleukemia cells. *Mol. Cell. Biol.*, **12**, 2967-2975.

SCHUETZE S., P.E. STENBERG and D. KABAT. (1993). The Ets-related transcription factor PU.1 immortalizes erythroblasts. *Mol. Cell. Biol.*, **13**, 5670-5678.

SCHULE R., P. RANGARAJAN, S. KLIEWER, L.J. RANSONE, J. BOLADO, N. YANG, I.M. VERMA and R.M. EVANS. (1990). Functional antagonism between oncoprotein c-jun and the glucocorticoid receptor. *Cell*, **62**, 1217-1226.

SEGRAVES W.A. and D.S. HOGNESS. (1990). The E75 ecdysone-inducible gene responsible for the 75B early puff in *Drosophila* encodes two new members of the steroid receptor superfamily. *Genes Dev.*, **4**, 204-219.

SEKIGUCHI T., T. MIYATA and T. NISHIMOTO. (1988). Molecular cloning of the cDNA of human chromosomal gene (CCG1) which complements the temperature-sensitive G1 mutants, tsBN462 and ts13, of the BHK cell line. *EMBO J.*, **7**, 1683-1687.

SEKIGUCHI T., Y. NOHIRO, Y. NAKAMURA, N. HISAMOTO and T. NISHIMOTO. (1991). The human CCG1 gene essential for progression of the G1 phase encodes a 210-kilodalton nuclear DNA-binding protein. *Mol. Cell. Biol.*, **11**, 3317-3325.

SETH A., D.K. WATSON, D.G. BLAIR and T.S. PAPAS. (1989). *c-ets-2* protooncogene has mitogenic and oncogenic activity. *Proc. Natl. Acad. Sci. USA*, **86**, 7833-7837.

SETH A. and T.S. PAPAS. (1990). The *c-ets-1* proto-oncogene has oncogenic activity and is positively autoregulated. *Oncogene*, **5**, 1761-1767.

SETH A., L. ROBINSON, D.M. THOMPSON, D.K. WATSON and T.S. PAPAS. (1993). Transactivation of GATA-1 promoter with ETS1, ETS2 and ERGB/Hu-FLI-1 proteins : stabilization of the ETS1 protein binding on GATA-1 promoter sequences by monoclonal antibody. *Oncogene*, **8**, 1783-1790.

SHARP P.A. (1992). TATA-binding protein is a classless factor. *Cell*, **68**, 819-821.

SHEN-ONG G.L.C. (1990). The *myb* oncogene. *Biochim. Biophys. Acta*, **1032**, 39-52.

SHIMIZU M. S.Y. ROTH, C. SZENT GYORGYI and R.T. SIMPSON. (1991). Nucleosomes are positioned with base pair precision adjacent to the alpha 2 operator in *Saccharomyces cerevisiae*. *EMBO J.*, **10**, 3033-3041.

SIDDIQUE H.R., V.N. RAO, L. LEE and E.S.P. REDDY. (1993). Characterization of the DNA binding and transcriptional activation domains of the *erg* protein. *Oncogene*, **8**, 1751-1755.

SMALE S.T. and D. BALTIMORE. (1989). The initiator as a transcription control element. *Cell*, **57**, 103-113.

SMALE S.T., M.C. SCHMIDT, A.J. BERK and D. BALTIMORE. (1990). Transcriptional activation by Sp1 as directed through TATA or initiator : specific requirements for mammalian transcription factor IID. *Proc. Natl. Acad. Sci. USA*, **87**, 4509-4513.

SPECTOR D., F. PURVES and B. ROIZMAN. (1990). Mutational analysis of the promoter region of the $\alpha 27$ gene of herpes simplex virus 1 within the context of the viral genome. *Proc. Natl. Acad. Sci. USA*, **87**, 5268-5272.

SPRENT J., M. SCHAEFER, D. LO and R. KORNGOLD. (1986). Functions of purified L3T4⁺ and Lyt-2⁺ cells *in vitro* and *in vivo*. *Immunol. Rev.*, **91**, 195-218.

STIEGLER P., C.M. WOLFF, M. BALTZINGER, J. HIRTZLIN, F. SENAN, D. MEYER, J. GHYSDAEL, D. STEHELIN, N. BEFORT and P. REMY. (1990). Characterization of *Xenopus laevis* cDNA clones of the *c-ets-1* proto-oncogene. *Nucleic Acids Res.*, **18**, 5298.

STIEGLER P., C.M. WOLFF, D. MEYER, F. SENAN, M. DURLIAT, J. HOURDRY, N. BEFORT and P. REMY. (1993). The *c-ets-1* proto-oncogenes in *Xenopus laevis* : expression during oogenesis and embryogenesis. *Mech. Dev.*, 1-10.

STRINGER K.F., C.J. INGLES and J. GREENBLATT. (1990). Direct and selective binding of an acidic transcriptional activation domain to the TATA-box factor TFIID. *Nature*, **345**, 783-786.

SVAREN J., and W. HÖRZ. (1993). Histones, nucleosomes and transcription. *Curr. Op. Gen. Dev.*, **3**, 219-225.

SWAIN A. and J.M. COFFIN. (1992). Mechanism of transduction by retroviruses. *Science*, **255**, 841-844.

TAKADA R., Y. NAKATANI, A. HOFFMAN, N.T. KOKUBO, S. HASEGAWA, R.G. ROEDER and M. HORIKOSHI. (1992). Identification of human TFIID components and direct interaction between a 250-kD polypeptide and the TATA-box binding protein (TFIID τ). *Proc. Natl. Acad. Sci. USA*, **89**, 11809-11813.

TANESE N., B.F. PUGH and R. TJIAN. (1991). Coactivators for a proline-rich activator purified from the multisubunit human TFIID complex. *Genes Dev.*, **5**, 2212-2224.

TASSET D., L. TORA, C. FROMENTAL, E. SCHEER and P. CHAMBON. (1990). Distinct classes of transcriptional activating domains function by different mechanisms. *Cell*, **62**, 1177-1187.

TEI H., I. NIHONMATSU, T. YOKOKURA, R. UEDA, Y. SANO, T. OKUDA, K. SATO, K. HIRATA, S.C. FUJITA and D. YAMAMOTO. (1992). *Pokkuri*, a *Drosophila* gene encoding an E26-specific (Ets) domain protein, prevents overproduction of the R7 photoreceptor. *Proc. Natl. Acad. Sci. USA*, **89**, 6856-6860.

TEMIN H.M. (1993). Retrovirus variation and reverse transcription : abnormal strand transfers result in retrovirus genetic variation. *Proc. Natl. Acad. Sci. USA*, **90**, 6900-6903.

THOMAS S.M., M. DeMARCO, G. D'ARCANGELO, S. HALEGOUA and J.S. BRUGGE. (1992). *ras* is essential for nerve growth factor- and phorbol ester-induced tyrosine phosphorylation of MAP kinases. *Cell*, **68**, 1031-1040.

THOMPSON C.B., C.Y. WANG, I.C. HO, P.R. BOHJANEN, B. PETRYNIAK, C.H. JUNE, S. MIESFELDT, L. ZHANG, G.J. NABEL, B. KARPINSKI and J.M. LEIDEN. (1992). *cis*- acting sequences required for inducible interleukin-2 enhancer function bind a novel Ets-related protein, *elf-1*. *Mol. Cell. Biol.*, **12**, 1043-1053.

THOMPSON C.C., T.A. BROWN and S.L. McKNIGHT. (1991). Convergence of Ets- and Notch-related structural motifs in a heteromeric DNA binding complex. *Science*, **253**, 762-768.

THOMPSON C.C. and S.L. McKNIGHT. (1992). Anatomy of an enhancer. *Trends Genet.*, **8**, 232-236.

THUMMEL C.S., K.C. BURRIS and D.S. HOGNESS. (1990). Spatial and temporal patterns of E74 transcription during *Drosophila* development. (1990). *Cell*, **61**, 101-111.

TODD R.F., L.M. NADLER and S.F. SCHLOSSMAN. (1981). Antigens on human monocytes identified by monoclonal antibodies. *J. Immunol.*, **126**, 1435-1442.

TREISMAN R. (1990). The SRE : a growth factor responsive transcriptional regulator. *Sem. Canc. Biol.*, **1**, 47-58.

TRIEZENBERG S.J., K.L. LAMARCO and S.L. McKNIGHT. (1988). Evidence of DNA : protein interactions that mediate HSV-1 immediate early gene expression. *Genes Dev.*, **2**, 730-742.

TSAI S.Y., A. BANIAHMAD, I. HA, D. REINBERG, M.J. TSAI and B.W. O'MALLEY. (1993). Interaction of members of the steroid/thyroid hormone receptor super family with the basal transcription factor TFIIB. Cold Spring Harbor meeting on cancer cells, 235.

TURQUE N. (1991). Rôle du proto-oncogène c-ets-1 pendant le développement embryonnaire : recherches sur une éventuelle collaboration avec c-jun, c-fos pour le contrôle de la production de métalloprotéases au cours de la morphogénèse de la peau de l'embryon de poulet. Mémoire de DEA de Sciences de la Vie et de la Santé.

URNESS L.D. and C.S. THUMMEL. (1990). Molecular interactions within the ecdysone regulatory hierarchy : DNA binding properties of the *Drosophila* ecdysone-inducible E74A protein. *Cell*, **61**, 47-61.

VANDENBUNDER B., L. PARDANAUD, T. JAFFREDO, M.A. MIRABEL and D. STEHELIN. (1989). Complementary patterns of expression of c-ets-1, c-myb and c-myc in the blood-forming system of the chick embryo. *Development*, **106**, 265-274.

VAN DYKE M.W., R.G. ROEDER and M. SAWADOGO. (1988). Physical analysis of transcription preinitiation complex assembly on a class II gene promoter. *Science*, **241**, 1335-1338.

VAN HOY M., K.K. LEUTHER, T. KODADEK and S.A. JOHNSTON. (1993). The acidic activation domains of the GCN4 and GAL4 proteins are not α helical but form β sheets. *Cell*, **72**, 587-594.

VAN DE WETERING M., M. OOSTERWEGEL, D. DOOIJES and H. CLEVERS. (1991). Identification and cloning of TCF-1, a T lymphocyte-specific transcription factor containing a sequence-specific HMG box. *EMBO J.*, **10**, 123-132.

VOLOSHIN O.N., S.M. MIRKIN, V.I. LYAMICHEV, B.P. BELOTSEKOVSKII and M.D. FRANK-KAMENETSKII. (1988). Chemical probing of homopurine-homopyrimidine mirror repeats in supercoiled DNA. *Nature*, **333**, 475-476.

VON BOEHMER H. (1990). Developmental biology of T-cells in T-cell receptor transgenic mice. *Ann. Rev. Immunol.*, **8**, 531-556.

WANG C.Y., B. PETRYNIAK, I.C. HO, C.B. THOMPSON and J.M. LEIDEN. (1992). Evolutionary conserved ets family members display distinct DNA binding specificities. *J. Exp. Med.*, **175**, 1391-1399.

WANG C.Y., B. PETRYNIAK, C.B. THOMPSON, W.G. KAELIN and J.M. LEIDEN. (1993). Regulation of the Ets-related transcription factor Elf-1 by binding to the retinoblastoma protein. *Science*, **260**, 1330-1335.

WANG Y., S.C. KAYMAN, J.P. LI and A. PINTER. (1993). Erythropoietin receptor (EpoR)-dependent mitogenicity of spleen focus-forming virus correlates with viral pathogenicity and processing of env protein but not with formation of gp52-EpoR complexes in the endoplasmic reticulum. *J. Virol.*, **67**, 1322-1327.

WASYLYK B., C. WASYLYK, P. FLORES, A. BEGUE, D. LEPRINCE and D. STEHELIN. (1990). The c-ets proto-oncogenes encode transcription factors that cooperate with c-fos and c-jun for transcriptional activation. *Nature*, **346**, 191-193.

WASYLYK C., P. FLORES, A. GUTMAN and B. WASYLYK. (1989). PEA3 is a nuclear target for transcription activation by non-nuclear oncogenes. *EMBO J.*, **8**, 3371-3378.

WASYLYK C., A. GUTMAN, R. NICHOLSON and B. WASYLYK. (1991). The c-ets oncoprotein activates the stromelysin promoter through the same elements as several non-nuclear oncoproteins. *EMBO J.*, **10**, 1127-1134.

WASYLYK C., J.P. KERCKAERT and B. WASYLYK. (1992). A novel modulator domain of Ets transcription factors. *Genes Dev.*, **6**, 965-974.

WASYLYK C. and B. WASYLYK. (1992). Oncogenic conversion alters the transcriptional properties of ets. *Cell Growth and Diff.*, **3**, 617-625.

WASYLYK C. and B. WASYLYK. (1993). Oncogenic conversion of Ets affects redox regulation *in-vivo* and *in-vitro*. *Nucleic Acids Res.*, **21**, 523-529.

WATSON D.K., M.J. McWILLIAMS-SMITH, M.F. NUNN, P.H. DUESBERG, S.J. O'BRIEN and T.S. PAPAS. (1985). The ets sequence from the transforming gene of avian erythroblastosis virus E26 has unique domain on human chromosomes 11 and 21 : both loci are transcriptionally active. *Proc. Natl. Acad. Sci. USA*, **82**, 7294-7298.

WATSON D.K., M.J. McWILLIAMS-SMITH, C. KOZAK, R. REEVES, J. GEARHART, M.F. NUNN, W. NASH, J.R. FOWLE III, P. DUESBERG, T.S. PAPAS and S.J. O'BRIEN. (1986). Conserved chromosomal positions of dual domains of the ets protooncogene in cats, mice and humans. *Proc. Natl. Acad. Sci. USA*, **83**, 1792-1796.

WATSON D.K., M.J. McWILLIAMS, P. LAPIS, J.A. LAUTENBERGER, C.W. SCHWEINFEST and T.S. PAPAS. (1988). Mammalian ets-1 and ets-2 genes encode highly conserved proteins. *Proc. Natl. Acad. Sci. USA*, **85**, 7862-7866.

WATSON D. K., R. ASCIONE AND T. S. PAPAS. (1990). Molecular analysis of the ets genes and their products. *Crit. Rev. Oncog.*, **1**, 409-436.

WEBSTER N.J.G., S. GREEN, D. TASSET, M. PONGLIKITMONGKOL and P. CHAMBON. (1989). The transcriptional activation function located in the hormone-binding domain of the human oestrogen receptor is not encoded in a single exon. *EMBO J.*, **8**, 1441-1446.

WEINZIERL R.O.J., B.D. DYNLACHT and R. TJIAN. (1993). Largest subunit of *Drosophila* transcription factor IID directs assembly of a complex containing TBP and a coactivator. *Nature*, **362**, 511-517.

WERNERT N., M.B. RAES, P. LASALLE, M.P. DEHOUCQ, B. GOSSELIN, B. VANDENBUNDER and D. STEHELIN. (1992). c-ets1 proto-oncogene is a transcription factor expressed in endothelial cells during tumor vascularization and other forms of angiogenesis in humans. *Am. J. Pathol.*, **140**, 119-127.

WHARTON K.A., M. JOHANSEN, T. XU and S. ARTAVANIS-TSAKONAS. (1985). Nucleotide sequence from the neurogenic locus Notch implies a gene product that shares homology with proteins containing EGF-like repeats. *Cell*, **43**, 567-581.

- WHITE R.J. and S.P. JACKSON. (1992). The TATA-binding protein : a central role in transcription by RNA polymerases I, II and III. *Trends Genet.*, **8**, 284-288.
- WOLFF C.M., P. STIEGLER, M. BALTZINGER, D. MEYER, J. GHYSDAEL, D. STEHELIN, N. BEFORT and P. REMY. (1990). Isolation of two different c-ets-2 proto-oncogenes in *Xenopus laevis*. *Nucleic Acids Res.*, **18**, 4603-4604.
- WOOD K.W., C. SARNECKI, T.M. ROBERTS and J. BLENIS. (1992). ras mediates nerve growth factor receptor modulation of three signal-transducing protein kinases : MAP kinase, Raf-1 and RSK. *Cell*, **68**, 1041-1050.
- WOODS D.B., J. GHYSDAEL and M.J. OWEN. (1992). Identification of nucleotide preferences in DNA sequences recognised specifically by c-ets-1 protein. *Nucleic Acids Res.*, **20**, 699-704.
- WORKMAN J.L. and A.R. BUCHMAN. (1993). Multiple functions of nucleosomes and regulatory factors in transcription. *Trends Biol. Sci.*, **18**, 90-95.
- YANG-YEN H.F., J.C. CHAMBARD, Y.L. SUN, T. SMEAL, T.J. SCHMIDT, J. DROUIN and M. KARIN. (1990). Transcriptional interference between c-jun and the glucocorticoid receptor : mutual inhibition of DNA binding due to direct protein-protein interaction. *Cell*, **62**, 1205-1215.
- YOCHER J., K. WESTON and I. GREENWALD. (1988). The *Caenorhabditis elegans* lin-12 gene encodes a transmembrane protein with overall similarity to *Drosophila* Notch. *Nature*, **335**, 547-550.
- YOCHER J.K. and I. GREENWALD. (1989). glp-1 and lin-12, genes implicated in distinct cell-cell interactions in *C. elegans*, encode similar transmembrane proteins. *Cell*, **58**, 553-563.
- YOSHIMURA A., A.D. D'ANDREA and H.F. LODISH. (1990). Friend spleen focus-forming virus glycoprotein gp55 interacts with the erythropoietin receptor in the endoplasmic reticulum and affects receptor metabolism. *Proc. Natl. Acad. Sci. USA*, **8**, 4139-4143.
- YU M., X.Y. YANG and M.E. MARTIN. (1993). The PEA3/ETS motif functions as a bi-directional initiator element in the absence of additional promoter/activator elements. Cold Spring Harbor meeting on cancer cells, 253.
- YUAN C.C., N. KAN, K.J. DUNN, T.S. PAPAS and D.G. BLAIR. (1989). Properties of a murine retroviral recombinant of avian acute leukemia virus E26 : a murine fibroblast assay for v-ets function. *J. Virol.*, **63**, 205-215.
- ZAWEL L. and D. REINBERG. (1992). Advances in RNA polymerase II transcription. *Curr. Opin. Cell Biol.*, **4**, 488-495.
- ZENZIE-GREGORY B., A. KHACHI, I.P. GARRAWAY and S.T. SMALE. (1993). Mechanism of initiator-mediated transcription : evidence for a functional interaction between the TATA-binding protein and DNA in the absence of a specific recognition sequence. *Mol. Cell. Biol.*, **13**, 3841-3849.
- ZHANG J. and H. TEMIN. (1993a). Rate and mechanism of nonhomologous recombination during a single cycle of retroviral replication. *Science*, **259**, 234-238.
- ZHANG J. and H. TEMIN. (1993b). 3' junctions of oncogene-virus sequences and the mechanisms for formation of highly oncogenic retroviruses. *J. Virol.*, **67**, 1747-1751.

ZHANG L., V. LEMARCHANDEL, P.H. ROMEO, Y. BEN-DAVID, P. GREER and A. BERNSTEIN. (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 Q., P.M. LIEBERMAN, T.G. BOYER and A.J. BERK. (1992). Holo-TFIID supports transcriptional stimulation by diverse activators and from a TATA-less promoter. *Genes Dev.*, **6**, 1964-1974.

ZHOU Q., T.G. BOYER and A.J. BERK. (1993). Factors (TAFs) required for activated transcription interact with TATA-box binding protein conserved core domain. *Genes Dev.*, **7**, 180-187.

ZON L.I., J.F. MOREAU, J.W. KOO, B. MATHEY-PREVOT and A.D. D'ANDREA. (1992). The erythropoietin receptor transmembrane region is necessary for activation by the Friend spleen focus-forming virus gp55 glycoprotein. *Mol. Cell. Biol.*, **12**, 2949-2957.

RESUME

La famille des gènes *Ets* regroupe un vaste ensemble de gènes apparentés dont la conservation phylogénique s'étend des Arthropodes aux Vertébrés supérieurs. Ces gènes semblent tous coder des facteurs de transcription qui possèdent en commun un domaine de liaison à l'ADN d'un type inédit, baptisé pour cette raison le domaine ETS. Bon nombre de ces protéines exercent leur activité transactivatrice en reconnaissant une séquence riche en bases puriques au sein de promoteurs viraux et cellulaires aujourd'hui identifiés en grand nombre. Toutefois, la reconnaissance de ce motif ne semble pas suffire à leur spécificité d'action, puisqu'il apparaît de plus en plus clairement que les protéines *Ets* n'accomplissent pleinement leur fonction régulatrice des gènes qu'en association avec d'autres partenaires transcriptionnels appartenant à différentes grandes familles de facteurs de transcription.

Le locus *c-ets-1* aviaire a été caractérisé le premier, à la suite du clonage de l'oncogène *v-ets* transduit par le virus E26. A partir de ce gène sont synthétisés deux facteurs de transcription, $p54^{c-ets-1}$ et $p68^{c-ets-1}$ qui ne diffèrent que par leur extrémité amino-terminale : celle de la protéine $p54^{c-ets-1}$ n'est codée que par un seul exon absent de l'oncogène *v-ets*, dénommé I⁵⁴, tandis que celle de la protéine $p68^{c-ets-1}$ est codée par deux exons, α et β , colinéaires à l'extrémité 5' de *v-ets*. Cette unique différence entre les deux protéines suffit à attribuer à la protéine $p68^{c-ets-1}$ un domaine transactivateur supplémentaire par rapport à la protéine $p54^{c-ets-1}$ qui n'en comporte qu'un seul.

D'autre part, la distribution tissulaire des deux ARNm correspondants présente plusieurs aspects intéressants : chez le poulet, le niveau d'expression de l'ARNm $p54^{c-ets-1}$ est relativement modéré dans la plupart des tissus, sauf dans les organes lymphoïdes dans lesquels il est très élevé. Par contre, l'expression de l'ARNm $p68^{c-ets-1}$ est restreinte aux vaisseaux sanguins de la rate, ainsi qu'à d'autres tissus d'origine mésodermique, mais elle est rigoureusement réprimée dans les organes lymphoïdes, quelque soit le stade embryonnaire observé ou l'âge de l'animal.

Une meilleure compréhension des effecteurs positifs et négatifs qui permettent une expression appropriée des gènes nécessite d'examiner les mécanismes régulant l'expression même de ces régulateurs. Dans cette optique, la description des mécanismes moléculaires qui gouvernent l'expression du locus *c-ets-1* aviaire pouvait présenter un grand attrait. Les résultats rapportés ici indiquent que les deux transcrits issus de ce gène sont synthétisés à partir de deux promoteurs différents : la transcription s'initie en de nombreux sites d'initiation ce qui reflète l'absence de motif TATA dans les deux cas. Le promoteur initiant la synthèse de l'ARNm $p54^{c-ets-1}$ est riche en bases G et C, et son organisation est globalement conservée chez l'homme. Il est également montré ici que l'activité du promoteur du transcrit $p68^{c-ets-1}$ est stimulée par son propre produit. De plus, sa séquence indique la présence de nombreux sites potentiels de fixation pour des régulateurs d'expression strictement lymphocytaire qui pourraient contribuer à réprimer l'expression de l'ARNm $p68^{c-ets-1}$ dans les organes lymphoïdes.