

50376  
1992  
89

64401

50376  
1992  
89

UNIVERSITE DES SCIENCES ET TECHNIQUES DE LILLE  
FLANDRES ARTOIS

Numéro d'ordre 902

**THESE**

présentée pour l'obtention du grade de

**DOCTEUR DE L'UNIVERSITE  
DES SCIENCES ET TECHNIQUES DE LILLE**

par

**Vincent LAUDET**



**STRUCTURE ET EVOLUTION DES  
RECEPTEURS NUCLEAIRES**



Soutenue le 15 Avril 1992 devant la commission d'examen

**Président : Pr. André VERBERT**

**Rapporteurs : Pr. Walter WAHLI  
Pr. Jean KREMBEL**

**Examineurs : Pr. Pierre FORMSTECHE  
Pr. Dominique STEHELIN  
Dr. François CATZEFLIS**



*A Catherine*

*A mes parents...*

*"Nous sommes en ce monde pour nous efforcer d'apprendre toujours, pour nous éclairer les uns les autres au moyen des conversations, et pour nous appliquer à faire de plus en plus progresser les sciences et les arts."*

*Wolfgang Amadeus Mozart*

*Lettre au Padre Martini,*

*Salzbourg, le 4 septembre 1776.*

*"Je m'élève contre le mythe selon lequel la science est en soi une entreprise objective qui n'est menée à bien que lorsque les savants peuvent se débarrasser des contraintes de leur culture et regarder le monde tel qu'il est réellement. (...) Je pense plutôt que l'on doit appréhender la science comme un phénomène social, comme une entreprise humaine dynamique, et non comme le travail de robots programmés pour recueillir de pures informations. (...) La science, puisque ce sont des individus qui la font, est une activité qui plonge ses racines dans la société. Elle progresse par pressentiment, vision et intuition."*

*Stephen Jay Gould*

*La mal-mesure de l'homme.*

*"Faire ce que l'on peut là où on est, je ne connais pas d'autre morale. Et ne pas s'arrêter de chercher à comprendre..."*

*François Mitterrand*

*Ici et Maintenant*

## REMERCIEMENTS

Cette thèse est une étape importante de cinq années passées dans le laboratoire de Dominique Stéhelin. Je voudrais le remercier de m'avoir accepté dans son laboratoire. Et surtout, surtout, je voudrais le remercier de laisser la porte de son bureau ouverte lorsqu'il part en voyage. Cette porte ouverte symbolise, pour moi, sa confiance et le climat de liberté et de bonne humeur qu'il sait créer et entretenir dans son équipe. J'espère que cette porte restera encore longtemps ouverte, Dominique...

Je remercie les membres du Jury d'avoir bien voulu accorder un peu de leur temps à la lecture de ce travail. Je suis très sensible à l'honneur qu'il me font en acceptant de le juger.

Que Simon Saule trouve ici ma gratitude. Bien des choses nous séparent à présent, mais il m'a lancé sur erbA et les récepteurs nucléaires et, au fond, c'est ça le plus important ...

Je voudrais remercier, pour des raisons différentes, trois piliers du laboratoire. Jean Coll parce que, fidèle ami, il est chaque jour disponible pour discuter, critiquer et surtout enrichir toutes les idées qui me passent par la tête, Agnès Bègue pour son inlassable dévouement, sans lequel jamais notre travail n'aurait pu aboutir et enfin Christian Lagrou pour avoir accepté, avec bonne humeur de me voir jour après jour massacrer des millions de cellules !

Tout au long du travail mené sur erbA et les récepteurs nucléaires, nous avons bénéficié de l'aide d'équipes extérieures au laboratoire. Je voudrais donc remercier Françoise Dieterlen et son groupe à Nogent sur Marne et François Catzeflis qui a accepté de guider un pauvre biologiste moléculaire dans les labyrinthes de la phylogénie.

Je voudrais dire un petit mot particulier à Catherine Hänni grâce à qui l'idée d'évolution fait son chemin dans notre laboratoire. Je voudrais aussi la remercier pour tout le reste ...

Les séduisantes secrétaires de ce laboratoire ont beaucoup souffert avec mon écriture, mes fautes d'orthographe et mes commandes urgentes. Merci Nini, Merci Marie-Christine. Merci aussi à Mômman de prendre soin de Popov. Je refuse de remercier Thierry Dupressoir avant la soutenance, il m'a promis de faire beaucoup de diapos-surprises et j'ai peur... J'ai travaillé (et rigolé) avec beaucoup de monde au laboratoire. Dans le désordre, et en espérant n'oublier personne, Merci à Dominique Leprince (de m'engueuler chaque jour avec tant de gentillesse), Pascale Crépieux (ski, ets et waterzoi ...), Véronique Fafeur (eh Chiquita !), Marie-Berthe et Marie-Anne (qui savent déposer, avec efficacité, une potion dégonflante sur mes chevilles), Brigitte Quatannens (BQ, BQ !!), Jean Marc Vanacker (tu peux parler plus fort ?), Olivier Albagli (la référence que tu cherches c'est Leprince et al., ...), Martine Duterque (et le placard ...), Anne Claire Flourens (qui a dit, une seule fois, que j'étais beau !). Avant d'arrêter, je voudrais saluer les ex p'tits Myckey (PM et PA), mes compagnons thésards (Christophe, Stéphane etc ...), les petits nouveaux du labo (welcome Guillaume !) et tout l'équipage de notre beau navire ...

## SOMMAIRE DETAILLE

<b>RESUME</b> .....	7
<b>PUBLICATIONS</b> .....	8
<b>ABREVIATIONS</b> .....	10
<b>PRESENTATION DU MEMOIRE</b> .....	11
<b>I - <u>INTRODUCTION</u></b> .....	14
<b>1<sup>ère</sup> PARTIE : PRELIMINAIRES</b> .....	15
<b>A) LES HORMONES</b> .....	15
1) Petite histoire des hormones.....	15
2) Le système endocrinien ou "le paradigme du rétrocontrôle".....	17
3) Les réponses aux hormones.....	18
4) Stéroïdes, rétinoïdes et hormones thyroïdiennes.....	19
<b>B) LES RECEPTEURS NUCLEAIRES DU COTE BIOCHIMIQUE</b> ..	20
1) Les complexes hormone-récepteur.....	20
2) La localisation subcellulaire des récepteurs.....	20
3) La "transformation" en deux étapes des récepteurs.....	21
4) Les premiers récepteurs clonés.....	22
<b>2<sup>ème</sup> PARTIE : BIOLOGIE MOLECULAIRE DES RECEPTEURS NUCLEAIRES</b> .....	23
<b>A) UNITE ET DIVERSITE DES RECEPTEURS NUCLEAIRES</b> .....	23
1) Une superfamille de gènes.....	23
2) Les récepteurs orphelins.....	24
3) Rechercher la nature du ligand.....	26

4) D'autres récepteurs ?.....	27
<b>B) STRUCTURE DES RECEPTEURS NUCLEAIRES.....</b>	<b>28</b>
<b>C) LA FIXATION A L'ADN.....</b>	<b>29</b>
1) Anatomie moléculaire des doigts de zinc.....	29
a) Doigts de zinc et doigts de zinc .....	29
b) Comment reconnaître l'ADN ?.....	30
c) Le domaine C est multifonctionnel.....	31
d) Doigts de zinc et spécificité ?.....	31
2) Les éléments de réponse sur l'ADN (HRE) et la spécificité.....	33
a) Les éléments de réponse aux hormones ont une structure palindromique.....	33
b) HRE et spécificité.....	35
3) RMN et cristallographie : un modèle d'interaction protéine-ADN.....	37
4) La fixation à l'ADN <i>in vivo</i> .....	38
<b>D) LE DOMAINE DE FIXATION DU LIGAND.....</b>	<b>39</b>
1) Anatomie du domaine E.....	39
2) La fixation du ligand.....	40
3) La dimérisation.....	41
4) Les rôles de la région Ti.....	42
a) Interactions avec les protéines hsp 90.....	42
b) Domaine Ti et transactivation.....	43
<b>E) LES SIGNAUX DE TRANSLOCATION NUCLEAIRES.....</b>	<b>44</b>
<b>F) LA REGULATION DE LA TRANSCRIPTION.....</b>	<b>45</b>
1) Localisation des domaines transactivateurs.....	45
2) Transactivation et hormones.....	46
3) La synergie transcriptionnelle entre récepteurs.....	47
a) La fixation à l'ADN.....	47
b) Synergie transcriptionnelle.....	48

<b>G) LES INTERACTIONS AVEC LES FACTEURS DE TRANSCRIPTION.</b>	49
1) Fixation des facteurs sur des sites différents : le modèle du LTR du MMTV.....	49
2) La fixation sur les mêmes sites.....	51
3) Les interactions protéines-protéines.....	52
4) La titration de facteurs communs.....	53

### **3<sup>ème</sup> PARTIE : PHYSIOLOGIE CELLULAIRE ET RECEPTEURS**

<b>NUCLEAIRES</b> .....	55
<b>A) LA SPECIFICITE DES EFFETS PHYSIOLOGIQUES</b> .....	55
1) Les récepteurs des glucocorticoïdes et des minéralocorticoïdes.....	55
2) Les effets des minéralocorticoïdes dans les reins.....	56
3) Les effets des glucocorticoïdes dans le cerveau.....	56
<b>B) L'ACIDE RETINOIQUE ET LA MORPHOGENESE DES MEMBRES</b> .....	57
1) Le rôle morphogène des rétinoïdes : découverte et discussions.....	57
a) La morphogénèse des membres.....	57
b) ZPA et gradient.....	58
c) L'acide rétinoïque est un morphogène.....	58
d) L'acide rétinoïque n'est pas un morphogène.....	59
2) Les récepteurs de l'acide rétinoïque.....	60
a) Caractérisation.....	60
b) De multiples isoformes des RARs.....	60
c) Expression dans le bourgeon de membre.....	61
d) Les gènes RXR.....	61
3) Les CRABP et CRBP.....	62

4) Les gènes régulés par l'acide rétinoïque dans le bourgeon de membre....	62
a) Des cibles intrigantes : les gènes homéotiques.....	62
b) Les gènes Hox dans le bourgeon de membre.....	63
c) Gènes Hox et acide rétinoïque : "je t'aime, moi non plus" !.	63
5) L'acide rétinoïque ... au milieu du gué.....	64
<b>C) LES MALADIES LIEES AUX RECEPTEURS NUCLEAIRES....</b>	<b>65</b>
1) Les syndromes de résistance aux hormones.....	65
2) Récepteurs nucléaires et cancer : le cas des leucémies aiguës à promyélocytes.....	66
a) La maladie.....	66
b) La translocation t(15:17) et le récepteur $\alpha$ de l'acide rétinoïque.....	66
c) Un traitement à base d'acide rétinoïque ?.....	68

## **4<sup>ème</sup> PARTIE : LES HORMONES THYROIDIENNES ET LEURS RECEPTEURS.....**

<b>A) LES HORMONES THYROIDIENNES ET LEURS EFFETS.....</b>	<b>69</b>
<b>B) LE RETROVIRUS AEV ET LA DECOUVERTE DES RECEPTEURS DES HORMONES THYROIDIENNES.....</b>	<b>70</b>
1) AEV : le virus de l'érythroblastose aviaire.....	70
2) Les récepteurs des hormones thyroïdiennes.....	71
a) Découverte des deux gènes <u>c-erbA</u> .....	71
b) Plusieurs isoformes de <u>c-erbA-1</u> .....	72
c) Plusieurs isoformes de <u>c-erbA-2</u> .....	72
3) Expression des récepteurs des hormones thyroïdiennes.....	73
a) Expression des gènes <u>c-erbA</u> .....	73
b) Régulation par la T3 de l'expression des gènes <u>c-erbA</u> .....	74
4) Les domaines fonctionnels des récepteurs des hormones thyroïdiennes...	74

5) Les complexes avec les récepteurs de l'acide rétinoïque.....	75
<b>C) LES MECANISMES D'ACTIONS DE <u>v-erbA</u>.....</b>	<b>76</b>
1) Le modèle classique des cellules érythroïdes.....	76
2) Le rôle du domaine C.....	78
3) <u>v-erbA</u> , les levures et l'acide rétinoïque.....	79
<b>D) LA REPONSE AUX HORMONES THYROIDIENNES.....</b>	<b>80</b>
1) Les gènes cibles.....	80
2) La régulation du gène de l'hormone de croissance.....	81
3) Un processus physiologique : la métamorphose des amphibiens.....	82
<b>E) CONCLUSION.....</b>	<b>85</b>
<b>II - <u>RESULTATS</u>.....</b>	<b>86</b>
<b>A) ARTICLE 1 : ORGANISATION GENOMIQUE DU GENE HUMAIN CODANT LE RECEPTEUR DE TYPE <math>\alpha</math> DES HORMONES THYROIDIENNES (<u>c-erbA-1</u>).....</b>	<b>87</b>
1) Article.....	89
2) Résumé des résultats.....	97
<b>B) ARTICLE 2 : EVOLUTION DE LA SUPERFAMILLE DES RECEPTEURS NUCLEAIRES.....</b>	<b>103</b>
1) Article.....	105
2) Résumé des résultats.....	116
<b>III - <u>DISCUSSION</u>.....</b>	<b>122</b>
<b>A) STRUCTURE ET EVOLUTION DU LOCUS <u>c-erbA-1/ear-1</u>.....</b>	<b>123</b>
1) Les rôles possibles de l'imbrication des gènes <u>c-erbA-1</u> et <u>ear-1</u> .....	123
a) Des contraintes structurales.....	123

b) Des modèles pour le chevauchement des deux gènes.....	124
c) Le taux d'expression de <u>ear-1</u> semble influencer sur l'épissage de <u>c-erbA-1</u> ..	124
d) D'autres mécanismes possibles.....	125
e) Quelle est la fonction de <u>ear-1</u> ?.....	126
2) Evolution des récepteurs des hormones thyroïdiennes et du locus imbriqué....	127
a) Evolution de la fonction thyroïdienne.....	127
b) Apparition du locus imbriqué.....	128
c) Epissage alternatif humain et formes N-terminales tronquées du poulet...129	
<b>B) EVOLUTION DES RECEPTEURS NUCLEAIRES.....</b>	<b>130</b>
1) Apparition des récepteurs nucléaires.....	130
a) La formation du premier récepteur nucléaire.....	130
b) Ancienneté fonctionnelle des récepteurs nucléaires.....	131
c) Des récepteurs nucléaires chez d'autres organismes ?.....	132
d) Deux modèles d'évolution précoce.....	133
e) Les événements plus tardifs.....	134
2) Les trois sous-familles de récepteurs.....	135
a) Relation structure/évolution.....	135
b) Les "récepteurs orphelins" sont-ils orphelins ?.....	137
c) Rôle des récepteurs orphelins.....	138
3) Les récepteurs chimères.....	139
a) Les exemples de récepteurs chimères.....	139
b) Comment confirmer l'existence de récepteurs chimères ?.....	140
c) Les mécanismes possibles ayant pu conduire aux chimères.....	142
4) Histoire et récepteurs nucléaires.....	143
<b>IV - REFERENCES.....</b>	<b>145</b>

## RESUME

De nombreux facteurs régulant la transcription des gènes eucaryotes ont été caractérisés ces dernières années et, parmi ceux ci, des facteurs agissant sous le contrôle d'une hormone : les récepteurs nucléaires d'hormones. Ces molécules ont la même structure de base constituée de domaines fonctionnels relativement indépendants. Les deux régions les mieux caractérisées étant le domaine C, responsable de la fixation spécifique à l'ADN via deux structures rappelant les "doigts de zinc" et le domaine E situé en C-terminal, responsable de la fixation du ligand et, souvent, de la régulation de la transcription. Cette superfamille de gènes contient notamment les récepteurs des hormones stéroïdes, les récepteurs de l'acide rétinoïque, ceux des hormones thyroïdiennes mais aussi un certain nombre de récepteurs putatifs dont les ligands éventuels ne sont pas connus et qui sont appelés "récepteurs orphelins".

Les récepteurs des hormones thyroïdiennes codés par les gènes c-erbA-1 et c-erbA-2 nous ont particulièrement intéressés parce que l'oncogène v-erbA qui fait partie du génome du rétrovirus oncogène aviaire AEV est une version altérée du récepteur des hormones thyroïdiennes de type  $\alpha$  (c-erbA-1). Nous avons caractérisé l'organisation génomique du gène c-erbA-1 qui est constitué de 10 exons répartis sur 27 kb d'ADN du chromosome 17 chez l'homme. Au cours de cette étude nous avons montré que ce gène peut coder deux formes différentes de récepteurs suite à un épissage alternatif qui implique les deux derniers exons du gène. De plus, nous avons observé que le gène c-erbA-1 chevauchait un autre gène appartenant également à la famille des récepteurs nucléaires et appelé ear-1.

L'étude de l'organisation génomique de c-erbA-1 et la mise en évidence de cette structure chevauchante très inhabituelle nous ont poussé à étudier de plus près l'évolution de la superfamille des récepteurs nucléaires. En construisant des arbres phylogéniques basés sur la séquence des deux domaines conservés des récepteurs nucléaires (les domaines C et E) nous avons pu construire un modèle d'évolution de cette famille de gènes. Ce modèle implique que les récepteurs nucléaires sont apparus très tôt au cours de l'évolution et se sont probablement diversifiés rapidement au moment des stades d'évolution explosive des tout premiers métazoaires. Nous avons pu grouper les récepteurs nucléaires en trois sous-familles : celle des récepteurs des hormones thyroïdiennes, de l'acide rétinoïque et des gènes apparentés comme ear-1, celle des récepteurs orphelins et enfin celle des récepteurs des stéroïdes. Le regroupement de la majorité des récepteurs orphelins a été une surprise qui permet peut être d'imaginer un rôle commun pour ces molécules. Enfin, en comparant l'évolution des deux domaines structuraux nous avons pu montrer que les récepteurs nucléaires ont en général évolué par simple duplication de gènes. Cependant il existe des exceptions importantes à ce modèle comme les gènes de drosophile du groupe *knirps* ou comme les gènes des récepteurs de la vitamine D et de l'ecdysone qui semblent avoir une histoire évolutive plus complexe.

## PUBLICATIONS

P.AMOUYEL, V.LAUDET, P.MARTIN, R.LI, B.QUATANNENS, D.STEHELIN and S.SAULE.

Two nuclear oncogenic proteins, P135*gag-myb-ets* and p61/63*myc* cooperate to induce transformation of chicken neuroretina cells.

Journal of Virology **1989**, 63 : 3382-3388.

C.HÄNNI, V.LAUDET, M.SAKKA, A.BEGUE et D.STEHELIN.

Amplification de fragments d'ADN mitochondrial à partir de dents et d'os humains anciens.

C.R.Acad. Sci. III **1990**, 310 : 365-375.

\* V.LAUDET, A.BEGUE, C.HENRY-DUTHOIT, A.JOUBEL, P.MARTIN, D.STEHELIN et S. SAULE.

Genomic organization of the human thyroid hormone receptor  $\alpha$  (c-erbA-1) gene.

Nucleic Acids Research **1991**, 19 : 1105-1112

H.CAZLARIS, N.LE MARER, V.LAUDET, C.LAGROU, Q.ZHU, J.MONTREUIL et P.DELANNOY.

Modifications de la sialylation des cellules BHK 21/C13 après transfection *in vitro* par l'oncogène humain c-Ha-ras.

C.R.Acad. Sci. III **1991**, 312 : 293-300.

N.BACHNOU, V.LAUDET, T.JAFFREDO, B.QUATANNENS, S.SAULE et F.DIETERLIN-LIEVRE.

Chondrogenic transdifférenciation mediated by the cooperation of v-myc and v-erbA oncogenes in chick embryo heart rhabdomyosarcomas.

Oncogene **1991**, 6 : 1041-1048.

N.LE MARER, V.LAUDET, E.SVENSSON, H.CAZLARIS, B.VAN HILLE, C.LAGROU, D.STEHELIN, J.MONTREUIL, A.VERBERT et P.DELANNOY.

The c-Ha-ras oncogene induces an increase of the expression of the  $\beta$ -galactoside  $\alpha$ 2,6 sialyltransferase in rat fibroblast (FR3T3) cells.

Glycobiology, **1992**, 2 : 49-56.

\* V.LAUDET, C.HÄNNI, J.COLL, F.CATZEFLIS et D.STEHELIN.

Evolution of the nuclear receptor gene superfamily.

The EMBO Journal, **1992**, 11 : 1003-1013.

V.VANDAMME, H CAZLARIS, N.LE MARER, V.LAUDET, C.LAGROU, A.VERBERT et P.DELANNOY.

Comparison of sialyl- and  $\alpha$ -1,3galactosyltransferase activity in NIH3T3 cells transformed with ras oncogene : increased  $\beta$ -galactoside  $\alpha$ -2,6-sialyltransferase.

Biochimie, **1992**, in press.

\* Articles présentés dans ce mémoire.

### ARTICLES DE REVUE

C.HÄNNI, V.LAUDET et D.STEHELIN.

Paléontologie et biologie moléculaire : la rencontre de deux mondes.

Biofutur **1991**, 28 55-58.

C.HÄNNI, V.LAUDET et D.STEHELIN.

Les gènes des os : la récupération de l'ADN des os anciens.

Pour la Science **1991**, 161 12-13.

## ABREVIATIONS

<b>AEV</b>	Avian Erythroblastosis Virus
<b>APL</b>	Leucémie aiguë à promyélocytes
<b>cDNA</b>	Copie ADN double brin d'un ARN
<b>CLUSTAL</b>	Programme d'alignement de séquences
<b>CRABP</b>	Protéine cellulaire fixant l'acide rétinoïque
<b>CRBP</b>	Protéine cellulaire fixant le rétinol
<b>DM</b>	Domaine de dimérisation situé à côté du domaine Ti, dans le domaine E
<b>ECR</b>	Récepteur de l'ecdysone
<b>ECRE</b>	Élément de réponse à l'ecdysone
<b>ER</b>	Récepteur des oestrogènes
<b>ERE</b>	Élément de réponse aux oestrogènes
<b>FITCH</b>	Méthode de Fitch et Margoliash de construction d'arbres phylogéniques (connue aussi sous le nom de "méthode des moindres carrés")
<b>GH</b>	Hormone de croissance
<b>GR</b>	Récepteur des glucocorticoïdes
<b>GRE</b>	Élément de réponse aux glucocorticoïdes
<b>Hox</b>	Gènes homéotiques de vertébrés contenant la typique "boîte homéo"
<b>HRE</b>	Élément de réponse aux hormones
<b>hsp</b>	Protéine de choc thermique
<b>kpb</b>	Milliers de paires de bases de l'ADN
<b>MMTV</b>	Mouse Mammary Tumor Virus
<b>NJ</b>	Méthode Neighbor-Joining de construction d'arbres phylogéniques
<b>NL</b>	Signal de translocation nucléaire
<b>pb</b>	Paire de base de l'ADN
<b>RAR</b>	Récepteur de l'acide rétinoïque
<b>RARE</b>	Élément de réponse à l'acide rétinoïque
<b>TAF</b>	Domaine transactivateur
<b>Ti</b>	Domaine responsable de l'inhibition de la transactivation en absence d'hormone, situé au milieu du domaine E
<b>TRAP</b>	Thyroid Hormone Receptor Auxiliary Protein
<b>TRE ou T3RE</b>	Élément de réponse aux hormones thyroïdiennes
<b>THR ou TR</b>	Récepteur des hormones thyroïdiennes
<b>T3</b>	3,5,3'-L-triiodothyronine
<b>T4</b>	L-thyronine
<b>VDR</b>	Récepteur de la vitamine D

## PRESENTATION DU MEMOIRE

Les récepteurs nucléaires d'hormones sont des facteurs de transcription qui dépendent d'un ligand pour leur action et qui régulent directement l'expression de certains gènes cibles en se fixant sur des séquences d'ADN précises de leurs régions régulatrices. La conjonction d'un intérêt fondamental évident (comment les cellules eucaryotes régulent-elles l'expression de leur génome ?) et les nombreuses possibilités d'applications médicales (comment relier des maladies à des altérations des récepteurs nucléaires ?) ont conduit fondamentalistes et cliniciens à se pencher sur ces molécules. On sait à présent que les récepteurs nucléaires sont codés par une vaste famille de gènes. Cependant, la fonction de la grande majorité des membres de cette superfamille reste tout à fait mystérieuse. C'est en particulier le cas des "récepteurs orphelins" pour lesquels aucun ligand n'est encore décrit - si toutefois ces molécules ont des ligands...

De nombreux travaux ont étudié les récepteurs nucléaires, la plupart du temps en disséquant soigneusement l'un d'entre eux. C'est ce que nous avons fait en étudiant l'organisation génomique du gène du récepteur de type  $\alpha$  des hormones thyroïdiennes. Mais nous avons choisi au cours de ce travail de privilégier également une autre approche en tenant compte du fait que les gènes des récepteurs nucléaires appartiennent à une même famille. La notion de famille est basée sur des rapprochements structuraux entre les différents gènes mais elle débouche naturellement sur l'idée d'une origine évolutive commune des membres de cette famille. Le choix fondamental de notre approche est donc d'observer les récepteurs nucléaires comme une famille de gènes, de comprendre le mieux possible leur histoire évolutive et, à partir des connaissances accumulées sur les différents membres de cette famille, d'essayer de dégager ce qui les réunit : un modèle général de leurs modes d'action et de leurs rôles dans la cellule.

Avant d'étudier les récepteurs nucléaires sous cet angle, il nous fallait donc présenter assez précisément ce que l'on sait de leurs fonctions : c'est ce que nous avons essayé de faire dans l'introduction de ce mémoire. Cependant comme l'étude des récepteurs nucléaires est en pleine explosion, pour parvenir à présenter toute la superfamille des récepteurs nucléaires nous avons donc dû faire un certain nombre de choix qu'il nous faut expliquer quelque peu.

Dans une partie préliminaire de l'introduction nous avons présenté l'essentiel des idées depuis celle d'hormone (comment parler des récepteurs nucléaires en escamotant leurs ligands ?) jusqu'à celle de domaines fonctionnels, qui ont conduit les biologistes moléculaires à tenter de cloner les gènes des récepteurs nucléaires. Pour l'étude détaillée du mode d'action des récepteurs nucléaires nous avons pris comme "point zéro" le clonage des premiers gènes de récepteurs

nucléaires. Ceci implique que, d'une manière parfaitement injuste, nous n'avons que fort peu mentionné les nombreux travaux de biochimie réalisés avant cette date. Pourtant, les travaux de biologie moléculaire reposent toujours sur une base de biochimie et de biologie cellulaire peut être moins spectaculaire mais complètement indispensable et trop souvent ignorée.

Un deuxième choix important a été de présenter dans le détail les résultats obtenus sur le mode d'action de ces molécules. Dans la deuxième partie de l'introduction ("Biologie moléculaire des récepteurs nucléaires"), nous avons donc examiné les fonctions des récepteurs nucléaires. Cette approche a peut-être l'inconvénient d'être perçue comme un long catalogue, mais elle permet de mettre en place, peu à peu, un modèle général concernant le mode d'action des récepteurs nucléaires. Au cours de cette présentation, nous avons volontairement fait ressortir les points d'ombres ou ceux faisant l'objet de discussions encore vives.

Le modèle "classique" d'action des récepteurs nucléaires est tout jeune (il a 5 ans seulement !) mais déjà il est en pleine évolution. En effet la découverte des interactions entre récepteurs nucléaires et facteurs de transcription permet d'imaginer toute une série de fonctions nouvelles de ces molécules sans doute aussi importantes que celles de régulateurs directs de la transcription. On ne soupçonne sans doute encore qu'une toute petite portion de cet aspect. Nous avons essayé de faire passer dans l'introduction ce climat particulièrement stimulant en présentant les travaux récents, leurs conclusions encore provisoires et aussi leurs limites.

L'implication des récepteurs nucléaires dans la physiologie cellulaire est encore mal comprise et on arrive mal à faire un lien solide au niveau moléculaire entre un effet physiologique donné et les récepteurs nucléaires. Nous avons insisté sur le modèle de la morphogenèse des membres parce que son étude a permis de révéler peu à peu l'extraordinaire complexité des systèmes de réponse à l'acide rétinoïque. Nous avons aussi présenté comment l'étude de certains processus pathologiques permet de mieux comprendre le rôle des récepteurs nucléaires dans la cellule. Il semble évident que, dans ce domaine plus que dans tous les autres, les années à venir vont nous permettre d'obtenir de nouvelles et passionnantes informations.

Enfin, dans la dernière partie de l'introduction nous avons décortiqué un récepteur nucléaire qui nous est particulièrement cher, entre autre parce qu'il est le progéniteur cellulaire d'un oncogène viral : le récepteur des hormones thyroïdiennes. Nous l'avons fait toujours avec la même idée, dégager des lignes de forces sans négliger les exceptions et les interrogations.

Nous avons choisi de présenter nos résultats sous forme d'articles directement insérés dans la thèse. Cependant, à chaque fois nous avons expliqué quelles étaient les questions posées et comment nous avons essayé d'y répondre. Chaque article est suivi d'une présentation des résultats et des conclusions auxquelles nous sommes parvenus. Nous présentons d'abord nos résultats sur la structure du gène c-erbA-1. L'étude de l'organisation de ce gène a permis de révéler une imbrication entre deux gènes : c-erbA-1 et un récepteur orphelin, ear-1. Dans la deuxième partie nous avons présenté notre analyse de l'évolution de la superfamille des récepteurs nucléaires.

Au cours de la discussion, nous avons présenté des hypothèses sur le rôle possible de la très inhabituelle structure imbriquée du locus c-erbA-1 humain. Pour la partie concernant l'évolution des récepteurs nucléaires, nous avons présenté un scénario chronologique de l'émergence de cette superfamille de gènes. Nous nous sommes appuyés sur le maximum de données différentes en tenant par exemple compte des connaissances acquises sur le rôle des domaines fonctionnels des récepteurs. Ainsi doit émerger une image globale, quoiqu'encore imprécise, de l'histoire de la superfamille des récepteurs nucléaires ...

## **INTRODUCTION**

## PRELIMINAIRES

### A) LES HORMONES.

#### 1) Petite histoire des hormones.

L'histoire commence en 1628 lorsqu'un chirurgien anglais, William Harvey, découvre la circulation sanguine. Après Copernic et Gallilée qui viennent, à un siècle d'intervalle, de démontrer que la terre tourne autour du soleil, voici que le chirurgien personnel des rois d'Angleterre se met à imaginer que le corps humain est le siège d'un courant de liquide sanguin ! Cette découverte constitue le premier élément de notre système : les hormones sont véhiculées par le sang et pour cela, le sang doit circuler ! Peu après, Antonie Van Leeuwenhoek met au point le microscope et observe pour la première fois des cellules : spermatozoïdes, globules sanguins, micro-organismes ...

Très vite, alors que le corps humain dans son ensemble et surtout la manière dont il fonctionne restent largement mystérieux, les notions d'interdépendance, de relations intimes, entre les différentes parties du corps surgissent. Ainsi, Thomas Willis, propose que des "ferments des glandes génitales" sont véhiculés par le sang et tiennent sous leur dépendance toute une série de phénomènes par exemple la mue de la voix ou les menstruations. En 1775, Théophile de Bordeu, dans son livre "*Analyse médicale du sang*" explique que chaque organe "*sert de laboratoire à une humeur particulière qu'il renvoie dans le sang après lui avoir donné un caractère radical*". Il insiste sur le rôle des testicules dont "*la semence, repompée et envoyée dans la masse des humeurs, donne un ton mâle et ferme à toutes les parties*". Malgré la justesse de ce genre de remarque, ceci reste général et, pendant bien longtemps, la situation n'évoluera pas. Ainsi, en 1843, Henle, en étudiant les glandes vasculaires sanguines dépourvues de canaux excréteurs, suppose encore qu'elles déversent dans le sang des substances élaborées dans leur parenchyme.

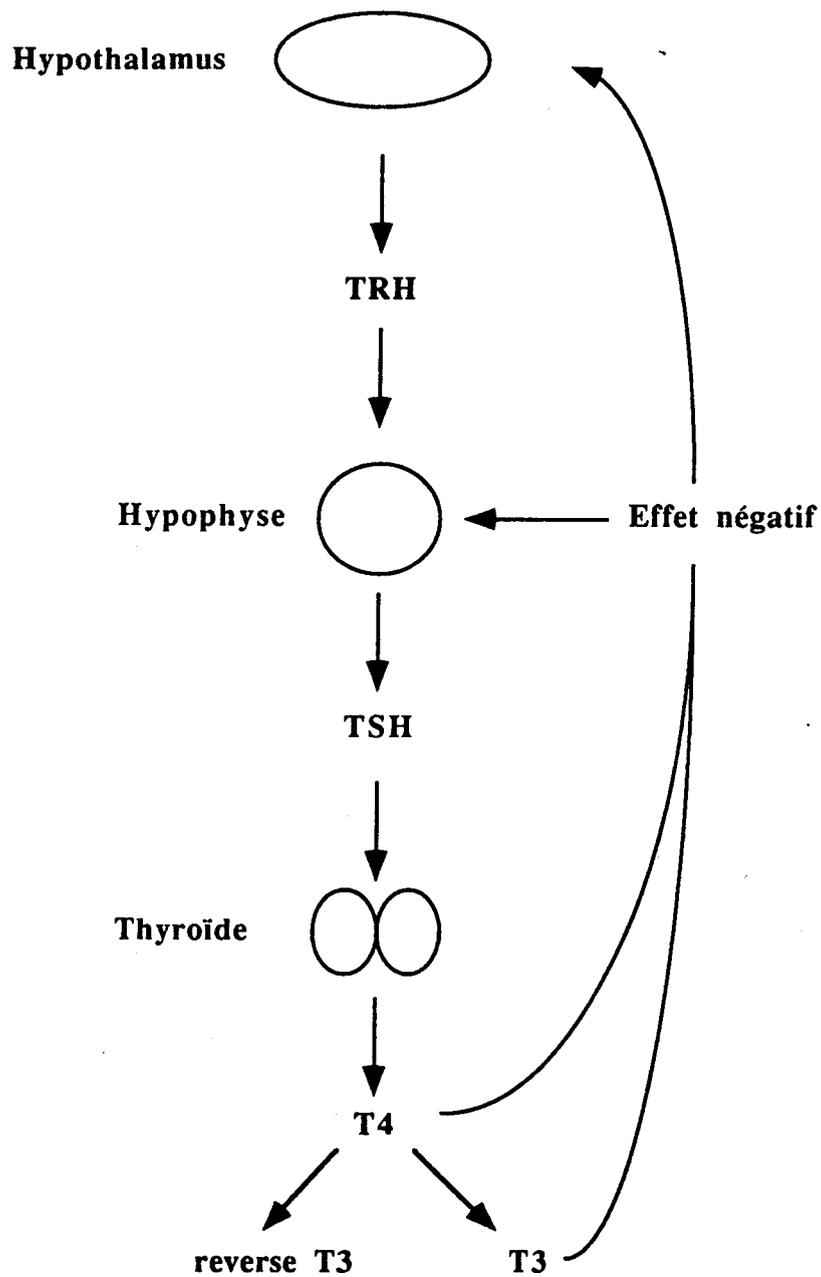
Il faut attendre le développement de la méthode dite anatomo-clinique, développée essentiellement par René Laennec au début du 19ème siècle, pour que des informations capitales sur ces fameuses glandes dépourvues de canaux excréteurs, qui s'appelleront les glandes endocrines, soient obtenues. Cette méthode consistera à confronter une maladie et l'observation des altérations de certains tissus ou organes. A l'aide de cette approche, (qui finalement est restée tout à fait d'actualité dans son essence même ; c'est à dire étudier les anomalies d'un système pour le comprendre) toute une série de maladies liées aux glandes endocrines va être décrite. Maladie de Basedow, maladie bronzée d'Addison, idiotie myxoédémateuse, acromégalie, etc ... sont associées au cours du 19ème siècle à l'altération de certaines glandes. Pourtant, le rôle même de ces glandes est encore très mal compris ; les médecins pensent encore qu'une substance nocive est élaborée par la glande ou n'est pas neutralisée par elle. C'est le cas par exemple de Pierre Marie qui découvre

l'acromégalie, un développement pathologique des os de la face et des extrémités des membres, lié à un hyper-fonctionnement de l'hypophyse. En 1925, à propos des glandes endocrines, Gley parle encore de "déviation trophique".

C'est dans la deuxième partie du 19<sup>ème</sup> siècle qu'un certain nombre d'observations va réellement prouver que les glandes endocrines sécrètent bien des substances libérées dans le sang. Il faut dire que tout un cadre conceptuel est alors en place. Ceci essentiellement grâce aux travaux de Claude Bernard qui insiste sur la notion de milieu intérieur et qui, en 1855, parle de sécrétion interne pour décrire la production et le rejet dans le sang du glucose par le foie. Dans cet ordre d'idées Berthold, démontre expérimentalement l'action des testicules sur la composition du sang en castrant des coqs et en transplantant la glande. En 1856, Vulpian observe le passage dans les veines surrénales d'une substance colorée en vert par le perchlorure de fer. Conscient de l'importance de cette observation, il écrit : *"ainsi serait prouvée pour la première fois et d'une façon décisive l'hypothèse qui regarde les capsules surrénales comme des glandes dites sanguines, c'est à dire versant directement dans le sang leur produit de sécrétion"*.

Observations décisives en effet, qui seront suivies de toute une série de preuves expérimentales de ce mode d'action des glandes dites sanguines. Reverdin et Kocher en 1883, puis Schiff en 1884, respectivement chez l'Homme et chez l'animal, montrent l'identité des effets de la thyroïdectomie et des symptômes du myxoedème. Cette idée ira à son aboutissement naturel grâce à Vassale et Gley qui, dans les années 1890, injectent des extraits thyroïdiens à un chien thyroïdectomisé et observent l'amélioration de son état. Bien sûr, ces travaux déboucheront sur les premiers traitements de myxoedème humain : le rôle fondamental des hormones dans l'harmonisation fonctionnelle des différentes parties de l'organisme est de mieux en mieux compris. Ainsi, c'est à cette époque que la théorie hormonale de la différenciation sexuelle est mise en place à la suite des travaux de Bouin et Ancel, et de Frank Lillie (Jost, 1991).

La première démonstration, presque officielle, de l'existence des hormones et de leur action, provient des expériences de Bayliss et Starling au tout début de notre siècle. Ils montrent en effet que le contenu gastrique acide surgissant du duodénum, à l'issue de la digestion, va entraîner la muqueuse duodénale (la glande) à libérer une substance, la gastrine (l'hormone), qui véhiculée par le sang, va agir au niveau du pancréas (le tissu cible), en stimulant la sécrétion des sucs pancréatiques. Cette fois-ci, la démonstration est complète. C'est de ce travail de Bayliss et Starling que provient le nom même d'hormone (du grec ορμαω : "je mets en mouvement, j'excite") et sa définition : *"toute substance produite normalement dans les cellules de n'importe quelle partie du corps et transportée par le courant sanguin dans des régions éloignées, sur lesquelles elle agit pour le bien de l'organisme entier"*. Il ne restait plus, si l'on peut dire, qu'à déterminer la composition chimique des hormones. Les hormones à structure simple furent en premier lieu purifiées et leur structure chimique élucidée : l'adrénaline en 1894 par Oliver et Schafer, puis la thyroxine par Kendall en 1914. La composition des hormones plus complexes, comme les hormones stéroïdes mettra plus de temps à être connue, mais ce sera chose faite dans les années 1950. Si la T4 a été



**Figure 1 : Un exemple classique de régulation en retour ou "feedback" : le contrôle de la sécrétion des hormones thyroïdiennes par l'hypophyse et l'hypothalamus (d'après Wilson et Foster, 1985).**

découverte assez vite, il a fallu attendre quarante ans pour que, au début des années cinquante, Rosalind Pitt-Rivers et Jack Gross à Londres, à la suite des travaux des équipes de Frédéric Joliot et Jean Roche, à Paris, identifient un dérivé déiodiné de la T4, la triiodothyronine ou T3, comme étant la forme active des hormones thyroïdiennes (Tata, 1990).

## 2) Le système endocrinien ou "le paradigme du rétro-contrôle".

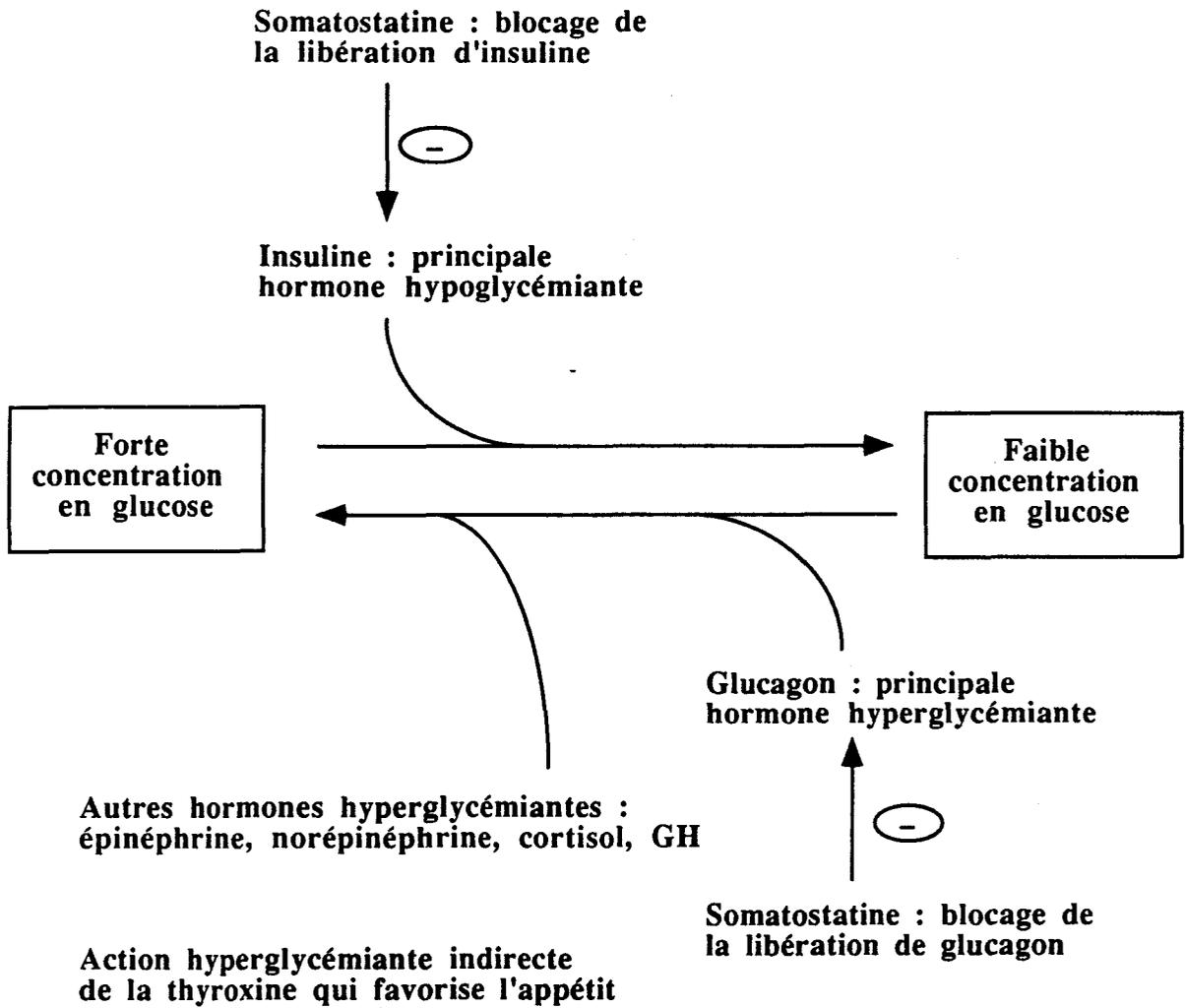
La définition précise d'une hormone et sa distinction avec d'autres molécules signal n'est pas toujours facile (Wilson et Foster, 1985). Par exemple il n'est pas simple d'établir une différence claire entre hormones et neurotransmetteurs : certains des médiateurs chimiques utilisés par le système nerveux répondent dans certains cas fort bien à la définition d'hormone. Le même problème se pose entre les notions d'hormone et de facteur de croissance. Classiquement, la distinction entre les deux types de molécule repose sur leur rayon d'action respectif. Mais les contre-exemples sont nombreux. Par exemple, la testostérone peut agir sur les cellules adjacentes à son lieu de synthèse dans les testicules, comme un facteur de croissance. La règle reste cependant que les hormones, véhiculées par le sang, agissent à longue distance, loin de leur lieu de synthèse.

Les hormones peuvent se diviser en deux groupes en fonction de leur nature chimique : les hormones de nature protéique, peptidique, ou dérivées d'un seul acide aminé (comme la sérotonine), classe à laquelle appartiennent les hormones thyroïdiennes (dérivées d'un dipeptide) et les hormones stéroïdes provenant du cholestérol. Toutes ont cependant trois caractéristiques communes :

1) Elles sont présentes dans le sang à très faible concentration. Pour les hormones stéroïdes, la gamme de concentration va de la picomole à la micromole par litre, tandis que pour les hormones peptidiques elle varie de 1 à 100 femtomoles par litre.

2) Elles sont dirigées vers leurs sites d'actions par des mécanismes spécifiques très efficaces qui permettent d'obtenir une quantité d'hormone dans le tissu cible bien supérieure à celle dans le plasma. Ce ciblage spécifique est obtenu grâce à la présence dans les tissus cibles de récepteurs à haute affinité. Cependant, l'idée de réponse "tout ou rien" du tissu cible est de moins en moins acceptable parce que la notion même de tissu cible est remise en question. En effet, souvent, les récepteurs spécifiques sont présents dans la plupart des organes : c'est le cas par exemple des récepteurs aux hormones thyroïdiennes (bien évidemment il y a de nombreux contre-exemples : le récepteur de l'aldostérone a une distribution tissulaire limitée). Il semble en fait que le ciblage de l'action de l'hormone est d'ordre quantitatif et non pas qualitatif. L'hormone va influencer l'ordre de grandeur de la réponse plutôt que déterminer s'il y a ou non une réponse. Un tissu va donc être un endroit de réponse majeur ou mineur à une hormone donnée.

Les récepteurs hormonaux ne sont pas les seuls à participer au ciblage de l'action des hormones (Wilson et Foster, 1985). D'autres types de mécanismes existent, qui vont



**Figure 2 : Le contrôle hormonal de la concentration de glucose (d'après Cryer, 1985).**

nous pousser à affiner la notion de cible. Ainsi, la proximité géographique entre un lieu de synthèse et un lieu "d'utilisation" de l'hormone a souvent une importance considérable. Dans ce cas, le tissu cible est sensible, non pas parce qu'il possède une grande quantité de récepteurs, mais parce qu'il est exposé à des doses plus fortes d'hormones. C'est le cas de l'hypophyse qui, via le système porte-hypophysaire est exposé à de fortes concentrations de facteurs peptidiques sécrétés par l'hypothalamus. Un autre mécanisme possible est la formation locale d'une molécule d'hormone active à partir d'un précurseur inactif circulant, le plus souvent plus stable. Le système des hormones thyroïdiennes T4 (précurseur) et T3 (molécule active) dont la synthèse est contrôlée par les déiodinases cellulaires nous en fournit un bon exemple.

3) L'une des caractéristiques distinctives du système endocrinien est la régulation en retour ou "feedback" de la production d'hormone (figure 1). Ces systèmes de rétro-contrôle atteignent un degré de minutie et d'efficacité impressionnant. Un exemple classique d'un tel contrôle est celui des hormones thyroïdiennes (Ingbar, 1985). Une fois sécrétées, ces hormones (T3 et T4) vont agir sur l'hypothalamus en diminuant la synthèse d'une hormone peptidique, la TRH (TSH Releasing Hormone). Cette molécule est responsable de l'induction de la synthèse de TSH (Thyroid Stimulating Hormone) par l'hypophyse. La TSH libérée dans le sang, va stimuler la glande thyroïde à sécréter les hormones thyroïdiennes. A chaque instant, la concentration en hormones thyroïdiennes sera donc le reflet des concentrations relatives en TRH et TSH. Les hormones thyroïdiennes vont également avoir un effet direct, négatif, sur la concentration en TSH. Virtuellement, toutes les hormones voient leur synthèse régulée par ce type de mécanisme. Ce concept de rétro-contrôle est si important en endocrinologie qu'on peut même le considérer comme le signe distinctif du système plus que les hormones elles-mêmes.

### 3) Les réponses aux hormones.

Dans tous ces cas de réponse physiologique à une hormone, l'équilibre obtenu n'est jamais le simple reflet de l'action univoque d'un signal donné. Il est évident qu'une hormone aura de multiples actions mais qu'en retour une fonction donnée sera régulée par plusieurs hormones (Wilson et Foster, 1985). Tous ces systèmes sont multiparamétriques.

Un exemple poussé à l'extrême de processus physiologique régulé de façon multiparamétrique par de nombreuses hormones est celui de la maintenance de la concentration plasmatique en glucose (Cryer, 1985). Cette concentration est soumise à deux exigences contradictoires : elle doit être assez élevée pour permettre un bon fonctionnement du système nerveux central et assez faible pour éviter les effets néfastes de l'hyperglycémie. Deux hormones majeures, l'insuline et le glucagon, vont donc moduler, de façon opposée, la concentration plasmatique en glucose (figure 2). L'insuline va diminuer la production de glucose, par exemple en favorisant son stockage dans le foie sous forme de glycogène et le glucagon va protéger le système nerveux central du manque de glucose, en stimulant sa production par le foie. Jusqu'ici le système

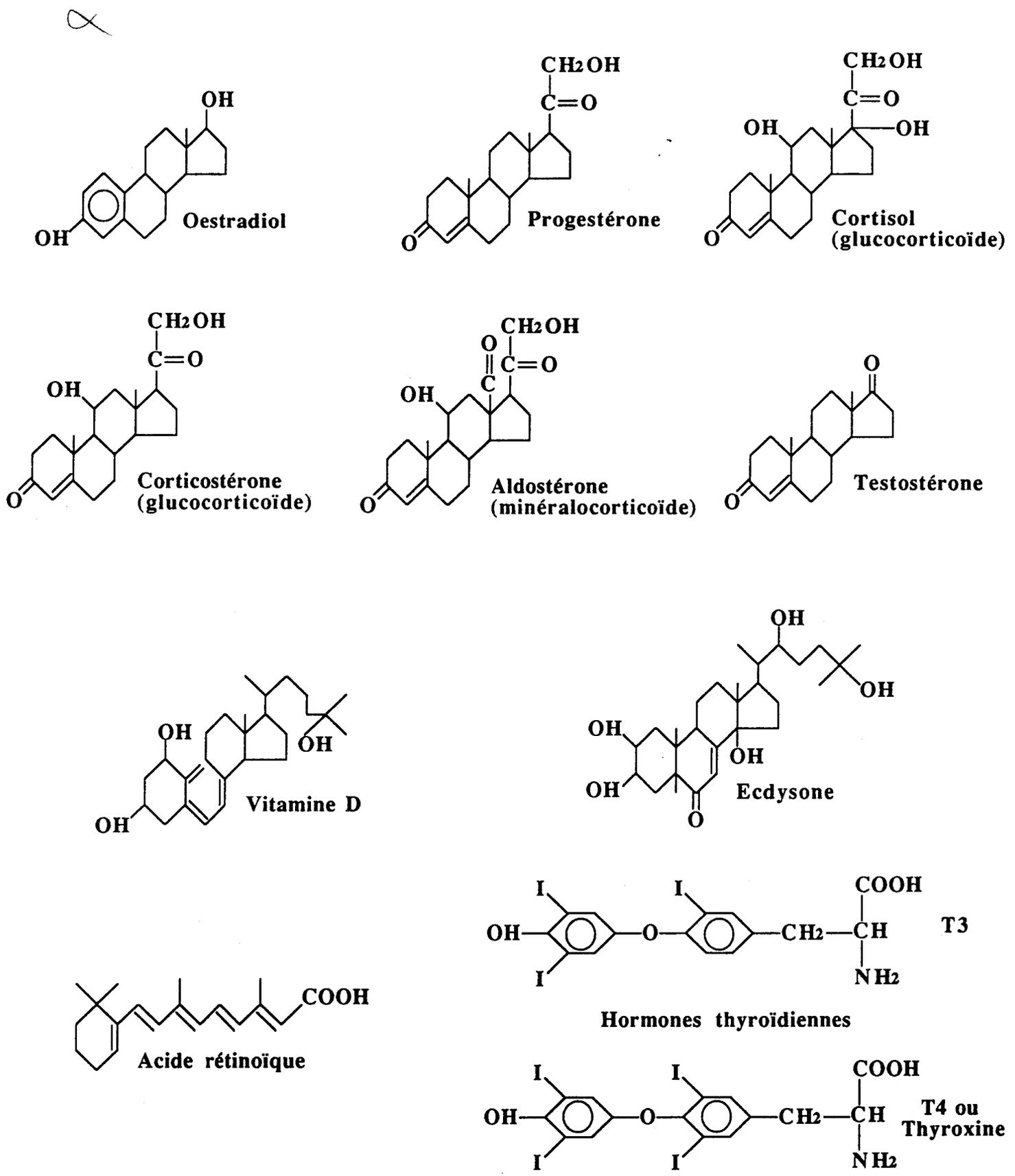


Figure 3 : Structure chimique des ligands des récepteurs nucléaires.

est assez simple. Mais, étant donné que l'hypoglycémie est un risque beaucoup plus grave que l'hyperglycémie, toute une série de molécules hyperglycémiantes vont être libérées dans le plasma si, par malheur, le glucose vient à manquer. Ce sont l'épinéphrine, la norépinéphrine, le cortisol et l'hormone de croissance (GH). Voici donc, déjà 6 molécules impliquées dans le maintien de la concentration en glucose ! De plus, d'autres hormones vont agir indirectement : la thyroxine va influencer l'appétit, la somatostatine va bloquer la libération de glucagon et d'insuline, et diminuer l'absorption des nutriments etc... Ces exemples illustrent bien la notion de redondance des régulations hormonales. Dans bien des cas le système est prêt à pallier la déficience anormale en l'un de ses composants...

#### 4) Stéroïdes, rétinoides et hormones thyroïdiennes.

Il y a cinq grandes classes d'hormones stéroïdes (figure 3) : les oestrogènes dont la forme la plus active est le  $17\beta$ -oestradiol, la progestérone, les glucocorticoïdes (cortisol chez l'Homme, corticostérone chez les rongeurs), les androgènes (surtout la dihydrotestostérone) et les minéralocorticoïdes dont l'aldostérone représente la forme majoritaire. Ces molécules ont des voies de synthèse imbriquées et reliées au cholestérol (Clark et al., 1985). A cette voie métabolique, on peut ajouter la vitamine D (connue aussi sous le nom de calciférol) et ses dérivés (la 1,25-dihydroxyvitamine D<sub>3</sub> ; Aurbach et al., 1985). Les hormones thyroïdiennes sont, pour leur part, sécrétées essentiellement sous deux formes d'inégales activités biologiques : la T<sub>4</sub> (L-thyroxine) et la T<sub>3</sub> (3,5,3'-L-triiodothyronine). Toutes ces molécules ont en commun avec l'acide rétinol (un dérivé de la vitamine A) le fait de diffuser "librement" à travers la membrane plasmique et d'avoir des récepteurs appartenant à la superfamille des récepteurs nucléaires.

Dans le modèle classique, les hormones stéroïdes sont véhiculées par le sang, soit sous forme libre (c'est la forme que l'on considère disponible pour se fixer aux récepteurs, puisqu'elle peut diffuser dans la cellule cible), soit sous forme liée à des composants du sang avec des affinités très variables ( $K_d$  de 1 mM jusqu'à 0,1 nM). Certaines de ces protéines fixatrices d'hormones stéroïdes, comme l' $\alpha$ -foetoprotéine qui fixe les oestrogènes, sont présentes à forte concentration et peuvent donc limiter considérablement la quantité d'hormone active dans le sang. Il est facile d'imaginer que la présence de telles protéines va permettre un niveau de régulation supplémentaire de l'action de l'hormone considérée (voir revue de Clark et al., 1985).

Après leur entrée dans la cellule, ces hormones vont se fixer à leurs récepteurs. C'est sur l'action des hormones dans la cellule et plus particulièrement sur leurs récepteurs que nous souhaitons maintenant concentrer notre attention.

## B) LES RECEPTEURS NUCLEAIRES, COTE BIOCHIMIQUE.

### 1) Les complexes hormone-récepteur.

Les récepteurs nucléaires se caractérisent par leur faible quantité dans la cellule : ils constituent en effet de 0,001% (récepteur des androgènes) à 0,1% (récepteur de la progestérone) des protéines solubles totales de la cellule. Leur purification a donc été difficile à réaliser (voir revue de Jensen, 1991).

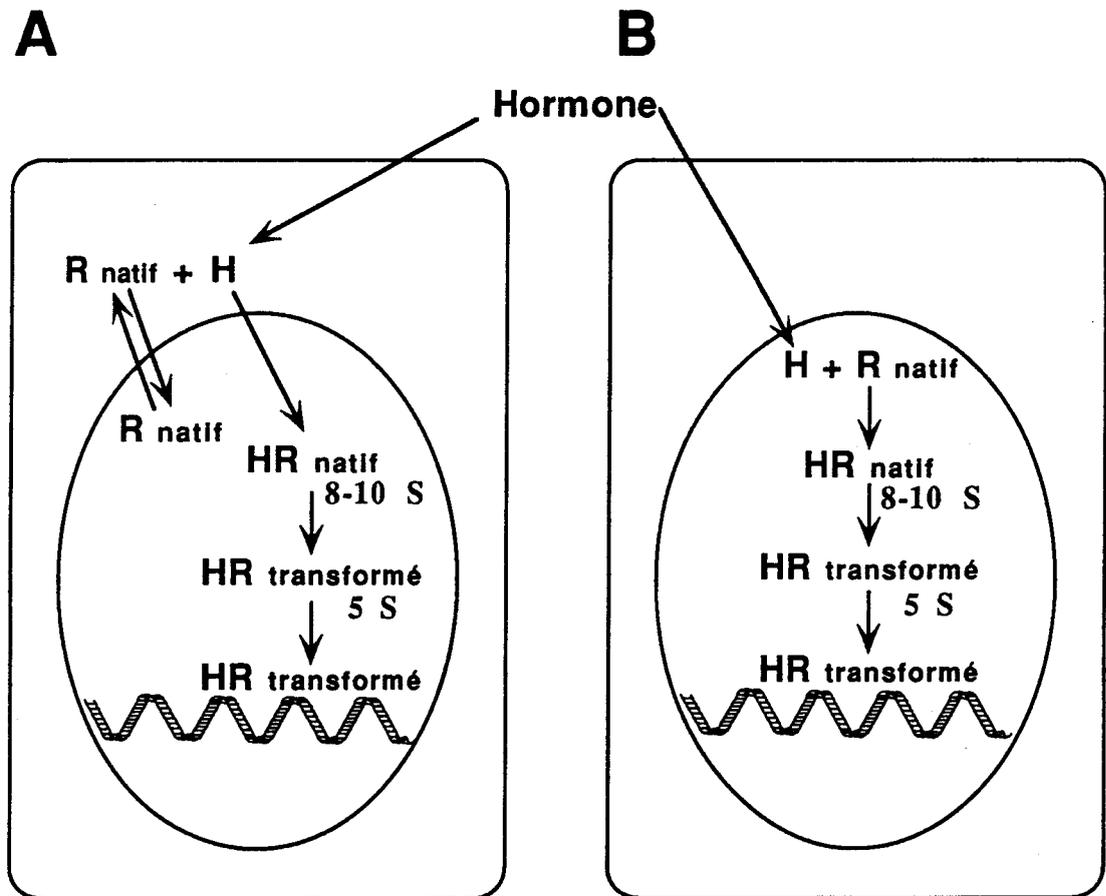
De nombreuses études biochimiques ont montré que l'affinité des récepteurs des stéroïdes pour le noyau est augmentée en présence de leur ligand. Certains récepteurs purifiés, le récepteur des oestrogènes par exemple, sont incapables de fixer l'ADN en absence de leur ligand. D'autres, comme le récepteur de la progestérone, fixent l'ADN en absence d'hormone. Ceux-ci voient cependant leur affinité pour l'ADN augmenter en présence de l'hormone (Jensen, 1991 ; Clark et al., 1985).

Par ultracentrifugation on a pu démontrer que les récepteurs des stéroïdes se trouvent dans différentes formes de complexes selon la force ionique du milieu. Ainsi, en haute force ionique (0,3M KCl) le récepteur des glucocorticoïdes est présent dans un complexe qui sédimente à 4 S, tandis qu'en faible force ionique le même récepteur est sous forme d'un complexe de 8S (Vedeckis, 1981 ; revue dans Jensen, 1991). Les deux formes de complexe hormone-récepteur observées selon la force ionique du milieu correspondent tout à fait aux complexes que l'on retrouve *in vivo* après exposition d'un tissu à l'hormone marquée. On retrouve une fraction cytosolique soluble à faible force ionique sous forme d'un complexe de 8 à 10 S et une fraction plus fortement associée au noyau, qui sédimente entre 3,5 et 4 S. Ce complexe correspond à la forme active du récepteur lié à l'hormone puisqu'il est capable de fixer l'ADN. Dans le cas des récepteurs des hormones thyroïdiennes ou de l'acide rétinoïque, on ne retrouve que la forme nucléaire très affine pour l'ADN (Jensen, 1991).

### 2) La localisation subcellulaire des récepteurs.

Puisque les récepteurs libres apparaissent le plus souvent dans la fraction cytosolique des homogénats cellulaires, il est logique de penser que le récepteur libre se retrouve effectivement dans le cytoplasme. Donc, la fixation de l'hormone à son récepteur doit induire une translocation nucléaire du complexe.

Les études d'autoradiographie à l'aide du ligand marqué ou les études d'immunohistochimie à l'aide d'anticorps monoclonaux ont montré que les récepteurs étaient, de manière très prédominante, à localisation nucléaire. Ces résultats ont été confirmés par des expériences d'énucléation des cellules cibles, où l'on a retrouvé l'essentiel des récepteurs testés dans les nucléoplastes plutôt que dans le cytoplasme. Cependant, ce type de méthode donne



**Figure 4 : Représentation schématique du modèle "à deux étapes". Dans un premier temps, il y a reconnaissance du récepteur natif par l'hormone, soit dans le cytoplasme (A), soit dans le noyau (B), selon les cas. On obtient alors un complexe hormone-récepteur qui, dans une deuxième étape, va être transformé en un complexe capable de se fixer à l'ADN.**

souvent des résultats contradictoires : ainsi, Barsony et al., 1990, en utilisant une méthode de fixation des cellules à l'aide de micro-ondes, retrouvent le récepteur de la vitamine D dans le cytoplasme alors que Clemens et al., 1988, ont montré qu'il était largement nucléaire (voir aussi les exemples de Rundle et al., 1989 et Martins et al., 1991).

Une constante de tous ces types d'étude reste cependant que, en présence du ligand, les récepteurs sont convertis en une forme majoritairement nucléaire ; ce qui confirme parfaitement les données biochimiques obtenues. Ce point est important car il peut expliquer certains résultats contradictoires. En effet, selon les conditions expérimentales, on pourra avoir une translocation du récepteur dans le noyau et donc retrouver un marquage nucléaire artificiel.

Malgré les incertitudes il semble raisonnable de séparer trois types de récepteurs, en fonction de leur localisation subcellulaire :

- Les récepteurs des glucocorticoïdes, des minéralocorticoïdes et, peut-être de la vitamine D (c'est beaucoup moins clair dans ce cas), sont de manière prédominante mais pas exclusive, présents dans le cytoplasme de la cellule cible non traitée et subissent une translocation nucléaire en présence d'hormone.

- Les récepteurs des hormones sexuelles (oestrogènes, progestérone, androgènes) sont présents dans les deux compartiments, mais plus fortement ancrés au noyau. Ainsi des études récentes sur les signaux de translocation nucléaire suggèrent que le récepteur des oestrogènes a des signaux nucléaires plus forts que celui des glucocorticoïdes (Guiochon-Mantel et al., 1989 ; Picard et al., 1990).

- Les récepteurs de l'acide rétinoïque et des hormones thyroïdiennes sont toujours nucléaires et seraient fixés à l'ADN indépendamment de la présence de l'hormone.

### 3) La "transformation" en deux étapes des récepteurs.

Les relations entre les formes cytoplasmiques à haut coefficient de sédimentation et les formes nucléaires à plus faible coefficient de sédimentation ont été très étudiées. Il a été démontré, voici de nombreuses années, que le complexe oestradiol-récepteur fixé au noyau de la cellule cible n'était pas produit directement, mais résulte d'un mécanisme plus complexe, à deux étapes, que l'on a appelé "transformation" ou "activation" du récepteur (Gorski et al., 1968 ; Jensen et al., 1968 ; Jensen, 1990). Lorsque l'hormone rencontre son récepteur à l'état natif, elle forme un complexe avec celui-ci. Ce complexe est alors activé et va se fixer à l'ADN (figure 4).

La transformation du récepteur semble impliquer qu'un certain nombre de composants du complexe "cytoplasmique" 8-10 S quitte ce complexe (Milgrom, 1981 ; Pratt et al., 1989). De fait, la transformation du récepteur des oestrogènes est obtenue simplement par filtration sur gel ou par dialyse. On sait maintenant que le récepteur natif est en association avec un certain nombre de protéines cellulaires, par exemple les protéines de choc thermique hsp 90 qui quittent le complexe au cours de la transformation du récepteur (Catelli et al., 1985). Cette association avec les hsp90

requiert la présence d'ions métalliques, ce qui explique qu'une simple dialyse entraîne la transformation. De ce modèle il découle que, d'une manière encore mal comprise, l'association avec les hsp90 masque la région de fixation à l'ADN du récepteur (Pratt et al., 1989).

Énoncé par Jensen voici plus de vingt ans, le modèle d'activation des récepteurs nucléaires reste tout à fait d'actualité, même si les discussions restent vives quant à la localisation subcellulaire précise de ces différents récepteurs (revue dans Jensen, 1991). De même, son application à tous les membres de la famille, y compris les récepteurs des hormones thyroïdiennes ou de l'acide rétinoïque, est encore controversée.

#### 4) Les premiers récepteurs clonés.

Le premier récepteur nucléaire cloné fut celui des glucocorticoïdes (Hollenberg et al., 1985). En effet, ce récepteur avait été purifié au début des années 1980 chez le rat (Simons et al., 1981) et chez l'homme (Gehring et Hotz, 1983). On savait qu'il s'agissait d'une protéine de 94 kDa qui avait des portions fixatrices de l'ADN et de l'hormone distinctes. À l'aide des récepteurs purifiés, de nombreux anticorps avaient été fabriqués (Harmon et al., 1984 ; Gametchu et Harrison, 1984). On connaissait même un certain nombre de gènes (LTR du virus MMTV, métallothionéine, etc ...) directement régulés par les glucocorticoïdes et on avait pu déterminer une séquence "consensus" sur laquelle se fixait le récepteur en amont de ces gènes.

Grâce aux anticorps, le groupe de Ronald Evans à San Diego a isolé des clones cDNA d'une banque de cellules IM-9, une lignée lymphoïde humaine qui répond aux glucocorticoïdes (Weinberger et al., 1985 ; Hollenberg et al., 1985). Les clones obtenus représentaient bien le récepteur aux glucocorticoïdes, car la transcription, puis la traduction *in vitro* des clones, permettait d'obtenir une protéine d'un poids moléculaire apparent de 94 Kd qui fixait les glucocorticoïdes de manière spécifique.

L'analyse de la séquence des clones montra que le récepteur aux glucocorticoïdes avait des homologies avec l'oncogène v-erbA (Hollenberg et al., 1985) issu du rétrovirus aviaire AEV (Roussel et al., 1979 ; Debuire et al., 1984). Cette observation, qui ouvrait un champ nouveau pour l'étude de v-erbA, permettait de suggérer que les deux gènes avaient un ancêtre commun. L'oncogène v-erbA est donc en fait le premier membre connu de la superfamille des récepteurs nucléaires !

TABLEAU 1

## Liste des récepteurs nucléaires

Abréviation	Nom	Espèce	Références
AR	Androgen Receptor	Human	Tilley et al., 1989
ARP-1	(Orphan receptor)	Human	Ladias and Karathanasis, 1991
COUP	= ear 3 (Orphan receptor)	Human	Wang et al., 1989, Miyajima et al., 1988
E75	(Orphan receptor)	Drosophila	Segraves and Hogness, 1990
EAR1	= Rev-ErbA (Orphan receptor)	Human	Miyajima et al., 1989 ; Lazar et al., 1989
EAR2	(Orphan receptor)	Human	Miyajima et al., 1988
ECR	Ecdysone receptor	Drosophila	Koelle et al., 1991
EGON	(Orphan receptor)	Drosophila	Rothe et al., 1989
ER	Estrogen Receptor	Human	Green et al., 1986
ERR1	(Orphan receptor)	Human	Giguère et al., 1988
ERR2	(Orphan receptor)	Human	Giguère et al., 1988
FTZ-F1	(Orphan receptor)	Drosophila	Lavorgna et al., 1991
GR	Glucocorticoid Receptor	Human	Hollenberg et al., 1985
H2RIIBP	(Orphan receptor)	Mouse	Hamada et al., 1989
HNF4	(Orphan receptor)	Rat	Sladek et al., 1990
KNI	Knirps (Orphan receptor)	Drosophila	Nauber et al., 1988
KNRL	Knirps-related (Orphan receptor)	Drosophila	Oro et al., 1988
MR	Mineralocorticoid Receptor	Human	Arriza et al., 1987
NGF1B	= nur 77 = N10 (Orphan receptor)	Rat	Milbrandt, 1988 ; Ryseck et al., 1989 ; Hazel et al., 1988
PPAR	(Orphan receptor)	Mouse	Issemann and Green, 1990
PR	Progesterone Receptor	Human	Misrahi et al., 1987
RARA	Retinoic acid receptor $\alpha$	Human	Petkovich et al., 1987 ; Giguere et al., 1987
RARAX	Retinoic acid receptor $\alpha$	Xenopus	Ragsdale et al., 1989
RARB	Retinoic acid receptor $\beta$	Human	De Thé et al., 1987
RARG	Retinoic acid receptor $\gamma$	Mouse	Zelent et al., 1989
RARGX	Retinoic acid receptor $\gamma$	Xenopus	Ragsdale et al., 1989
RXR	(Orphan receptor)	Human	Mangelsdorf et al., 1990
SVP	Sevenup (Orphan receptor)	Drosophila	Mlodzik et al., 1990
TLL	Tailless (Orphan receptor)	Drosophila	Pignoni et al., 1990
TR2	(Orphan receptor)	Human	Chang and Kokontis, 1988 ; Chang et al., 1989
THRA	Thyroid hormone Receptor $\alpha$ (= <u>c-erbA-1</u> )	Human	Thompson et al., 1987 ; Laudet et al., 1991
THRAC	Thyroid hormone Receptor $\alpha$ (= <u>c-erbA-1</u> )	Chicken	Sap et al., 1986
THRAXA	Thyroid hormone Receptor $\alpha$ (= <u>c-erbA-1</u> )	Xenopus	Yaoita et al., 1990
THRAXB	Thyroid hormone Receptor $\alpha$ (= <u>c-erbA-1</u> )	Xenopus	Yaoita et al., 1990
THR $\beta$	Thyroid hormone Receptor $\beta$ (= <u>c-erbA-2</u> )	Human	Weinberger et al., 1986
THRBC	Thyroid hormone Receptor $\beta$ (= <u>c-erbA-2</u> )	Chicken	Forrest et al., 1990
THR $\beta$ XA	Thyroid hormone Receptor $\beta$ (= <u>c-erbA-2</u> )	Xenopus	Yaoita et al., 1990
THR $\beta$ XB	Thyroid hormone Receptor $\beta$ (= <u>c-erbA-2</u> )	Xenopus	Yaoita et al., 1990
USP	Ultraspiracle = 2C1-3 = CF1 (Orphan receptor)	Drosophila	Oro et al., 1990 ; Henrich et al., 1990 ; Shea et al., 1990
VERBA	<u>v-erbA</u> gene of Avian Erythroblastosis Virus	Chicken	Debuire et al., 1984 ; Damm et al., 1987
VDR	Vitamine D Receptor	Human	Baker et al., 1988

(Pour les Mammifères, seule une espèce a été mentionnée)

# BIOLOGIE MOLECULAIRE DES RECEPTEURS NUCLEAIRES

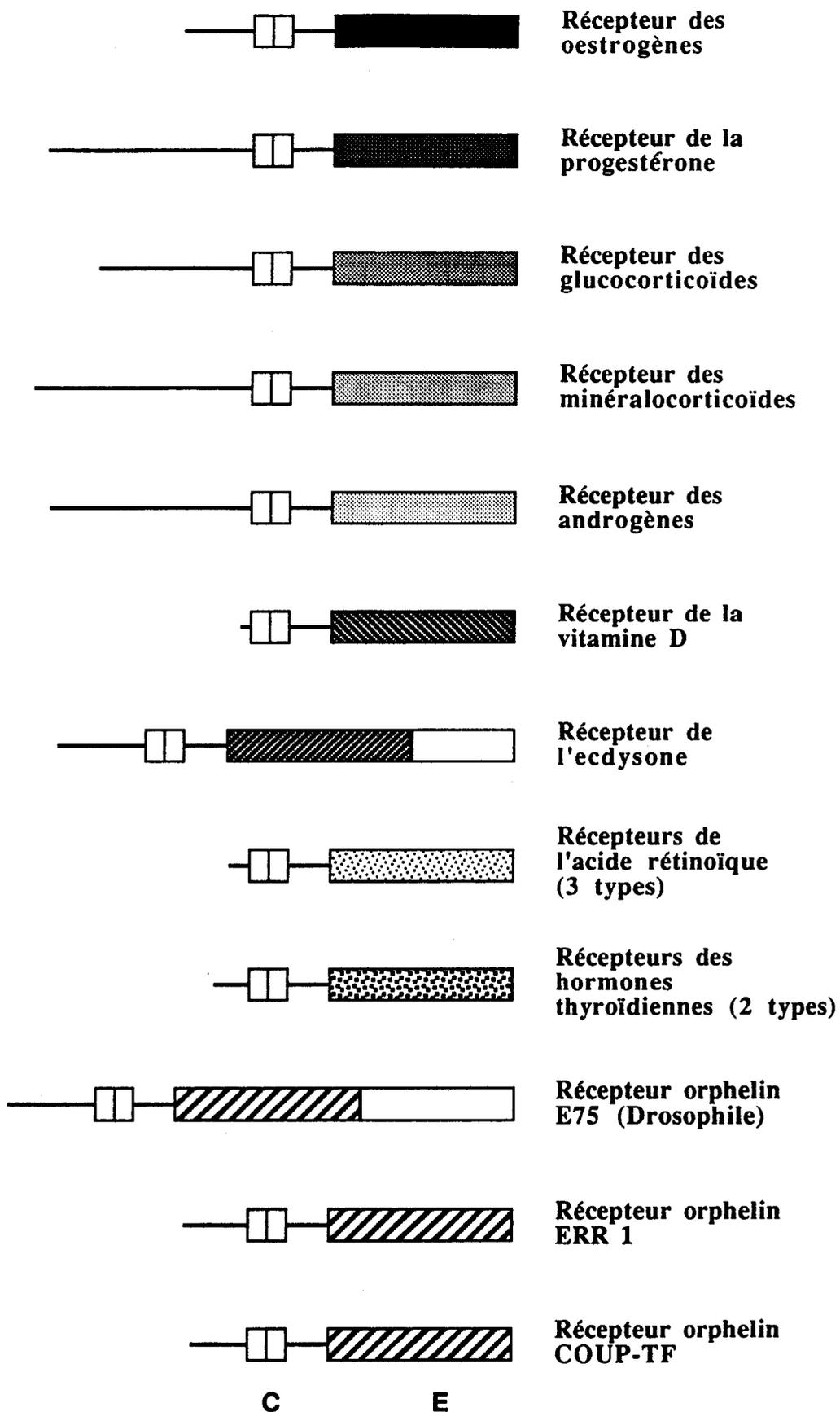
## A) UNITE ET DIVERSITE DES RECEPTEURS NUCLEAIRES.

### 1) Une superfamille de gènes.

Les récepteurs nucléaires forment une famille homogène de molécules capables de réguler la transcription des gènes de manière ligand-dépendante bien que, pour un certain nombre des membres de cette famille, la nature précise du ligand (voire même l'existence d'un ligand) reste encore mystérieuse. Dans un premier temps, à la suite du clonage des premiers récepteurs, on a cru avoir à faire à une famille de gènes de taille relativement modeste constituée des récepteurs des hormones stéroïdes et des récepteurs des hormones thyroïdiennes. Mais plus tard, il est apparu que cette famille comportait aussi de nombreux autres gènes, d'où l'appellation actuelle de superfamille.

Actuellement, la superfamille des récepteurs nucléaires comporte 32 gènes clonés qui présentent tous la même organisation en domaines fonctionnels caractéristiques notés A/B, C, D et E (figure 5 et tableau 1). On y trouve les récepteurs des hormones stéroïdes (glucocorticoïdes, oestrogènes, progestérone, androgènes et minéralocorticoïdes), le récepteur de la vitamine D3, les deux récepteurs des hormones thyroïdiennes et les trois récepteurs de l'acide rétinoïque (Evans, 1988). Depuis peu, le groupe de Hogness a cloné chez la drosophile le récepteur de l'ecdysone qui fait également partie de cette superfamille (Koelle et al., 1991). L'ecdysone a un rôle important pour la mue chez de nombreux insectes et peut être également chez d'autres invertébrés (Nirde et al., 1983; Hoffman et Joosse, 1991). En plus de ces récepteurs, dont la fonction est bien caractérisée, on trouve de nombreux récepteurs orphelins qui ne possèdent pas de ligand décrit et pour lesquels l'appellation de "récepteur" est, à ce jour, un abus de langage.

Jusqu'à tout récemment (voir Yu et al., 1992) aucun ligand réellement convaincant n'avait été décrit pour un récepteur orphelin sauf dans le cas du récepteur  $\beta$  de l'acide rétinoïque. Bien avant le clonage du premier récepteur de l'acide rétinoïque (Petkovich et al., 1987 ; Giguere et al., 1987), l'équipe d'Anne Dejean à l'Institut Pasteur avait découvert et caractérisé un site d'intégration du virus de l'hépatite B (HBV) à l'origine d'un hépatocarcinome (Dejean et al., 1986). Ce site a été séquencé et a montré une homologie significative avec l'oncogène v-erbA. Ceci en faisait donc un membre présumé de la famille des récepteurs nucléaires pour lequel aucun ligand n'était connu (De Thé et al., 1987). C'est dans un second temps que, en collaboration avec le laboratoire de Pierre Chambon, Anne Dejean et son équipe ont montré que ce gène code pour un deuxième type de récepteur de l'acide rétinoïque (Brand et al., 1988 ; Benbrook et al., 1988). On



**Figure 5 : Conservation de la structure en domaines de quelques récepteurs nucléaires. La coloration différente des domaines E indique des ligands différents ou inconnus (pour les récepteurs orphelins).**

sait maintenant qu'il en existe un troisième type cloné d'abord chez la souris (Zelent et al., 1989) puis chez l'homme (Krust et al., 1989) et chez le Xénope (Ragsdale et al., 1989 ; Giguere et al., 1989).

## 2) Les récepteurs orphelins.

A part le cas que nous venons de voir, ce sont les gènes ERR1 et ERR2 (pour "Estrogen Related Receptors") qui sont parmi les premiers récepteurs orphelins clonés (Giguere et al., 1988 ; voir le tableau 1). Ces deux gènes, comme leur nom l'indique, sont proches du récepteur des oestrogènes mais leurs produits ne fixent pas ces molécules. Ils ont été clonés en utilisant une sonde du récepteur des oestrogènes et en criblant une banque de testicules humains. Cet exemple permet d'illustrer ce que fut la stratégie la plus prolifique de clonage de membres de la famille. Par cette méthode, et en utilisant soit des sondes définies dans la région responsable de la fixation à l'ADN d'un récepteur donné, soit en utilisant des oligonucléotides dégénérés (définis également dans la région fixatrice de l'ADN) un très grand nombre de récepteurs orphelins a été cloné (Tableau 1). On peut par exemple noter les gènes de la série ear (pour erbA related) découverts par le groupe de Kumao Toyoshima : ear-1, ear-2 et ear-3 (Miyajima et al., 1988 ; Miyajima et al., 1989). Curieusement, si le gène ear-1 est effectivement assez proche des gènes c-erbA-1 et c-erbA-2 ce n'est pas le cas des deux autres. On sait maintenant que le gène ear-3 code le facteur de transcription COUP-TF ("chicken ovalbumine upstream promotor-transcription factor") (Wang et al., 1989) et que le gène ear-2 est proche de ear-3/COUP-TF.

L'exemple du gène ear-3/COUP-TF illustre bien comment le même récepteur nucléaire a pu être découvert en même temps par plusieurs équipes avec des approches différentes. Ainsi ear-3 a été découvert grâce à son hybridation croisée avec v-erbA (Miyajima et al., 1989), mais également en tant que facteur de transcription impliqué dans la régulation du gène de l'ovalbumine de poulet par le groupe de Bert O'Malley (Wang et al., 1989). En effet, depuis longtemps, cette équipe cherchait à caractériser les facteurs de transcription impliqués dans la régulation de l'expression des gènes eucaryotes. Le modèle choisi par ce groupe est le gène de l'ovalbumine de poulet parce que c'est un bon exemple de régulation tissu-spécifique (l'ovalbumine n'est exprimée que dans l'oviducte des poules). Un long travail sur le promoteur de ce gène a permis de montrer que sa région distale (appelée COUP et qui contient une répétition en tandem des bases GTCAAA), est essentielle pour la transcription du gène. Une protéine se fixe spécifiquement à cette séquence (Sagami et al., 1986). La purification de ce facteur a permis de montrer que non pas une, mais deux protéines différentes étaient impliquées dans cette régulation (Sagami et al., 1986 ; Baghi et al., 1987). Ces deux protéines sont : COUP-TF, une protéine de 45 à 50 kDa capable de se fixer au tandem GTCAAA, sans doute sous forme de dimère (Tsai et al., 1987), et le facteur S300-II, d'environ 90 à 100 Kd qui n'a pas d'activité de fixation à l'ADN. Il semble que S300-II, qui est absolument requis *in vitro* pour l'activation de la transcription, augmente la durée de vie du

complexe COUP-TF/ADN (Sagami et al., 1986 ; Tsai et al., 1987). La combinaison de trois approches (criblage d'une banque d'expression grâce à des anticorps, sélection des clones positifs par les propriétés de fixation à la séquence COUP et confirmation par comparaison avec une microséquence protéique) a permis d'isoler un clone correspondant à COUP-TF. C'est alors que, de manière très surprenante, on s'est aperçu que COUP-TF faisait partie des récepteurs nucléaires et était en fait identique au gène *ear-3* récemment décrit par l'équipe de Toyoshima !

On sait maintenant que, en association avec S300-II ou non, le facteur COUP régule de nombreux gènes, par exemple celui de l'insuline II de rat (Hwung et al., 1988) ou le LTR du rétrovirus humain HTLV-1 (Cooney et al., 1991). De plus, il existe au moins deux gènes très proches de COUP : le gène humain ARP-1 (Ladias et Karathanasis, 1991) cloné comme un régulateur des gènes d'apolipoprotéines, et le gène de drosophile *sevenup* (Mlodzik et al., 1991).

Nous venons donc d'illustrer deux grandes voies d'approche utilisées pour isoler des membres de la famille des récepteurs nucléaires : les homologies de séquence dans le domaine de fixation à l'ADN et l'activité de facteur de transcription. Dès que la diversité de cette famille a été fortement suspectée, un grand nombre de chercheurs a tenté de cloner de nouveaux récepteurs, en utilisant des sondes correspondant à des séquences conservées, ce qui a permis la découverte de nombreux récepteurs orphelins.

L'étude du groupe des récepteurs orphelins de type COUP (COUP, ARP-1 et *sevenup*) fournit un bon exemple d'une autre manière de trouver un récepteur orphelin : le hasard ! En effet, certains récepteurs n'ont pas été découverts en cherchant un récepteur nucléaire, mais en étudiant un problème biologique complètement différent. Ainsi, le gène *sevenup* découvert par les groupes de Corey Goodman et Gerald Rubin a été isolé à l'aide de la méthode de capture d'éléments activateurs ("enhancer trap") en cherchant des gènes impliqués dans la détermination des cellules de l'oeil de la drosophile (Mlodzik et al., 1990). Au cours de ce procédé on introduit dans des mouches un vecteur contenant un élément transposable et un gène *lacZ* (qui code une  $\beta$ -galactosidase d'origine bactérienne) fusionné à un promoteur minimal (O'Kane et Gehring, 1987). Si cette construction est intégrée près d'un élément activateur de la transcription (ou "enhancer"), on observera une expression du gène *lacZ*. Cette expression est facile à visualiser, puisque la  $\beta$ -galactosidase va catalyser la formation d'un produit de couleur bleue. Par cette méthode, un locus situé sur le troisième chromosome de la drosophile, et qui s'exprime dans des cellules spécialisées de l'oeil a été mis en évidence. En clonant le gène situé dans ce locus ces chercheurs ont montré que *sevenup* code un récepteur nucléaire très proche de COUP-TF. Le fait que ce gène, qui est considéré comme l'homologue de drosophile de COUP-TF, soit impliqué dans la différenciation de certaines cellules nerveuses est particulièrement intéressant. En effet, on sait depuis peu que COUP-TF est capable de répondre en activant la transcription d'un gène cible suite à l'adjonction de dopamine, un neurotransmetteur (Power et al., 1991a).

### 3) Rechercher la nature du ligand.

Pour tenter de trouver quels sont les ligands des récepteurs orphelins, un certain nombre de méthodes a été utilisé. La plus connue est la méthode dite de "l'échange des doigts" ou "finger swap". Elle met à profit l'indépendance relative qui existe entre les différents domaines fonctionnels des récepteurs nucléaires. Si l'on cherche à déterminer le ligand qui se fixe sur un récepteur donné "X", il faut trouver une manière de mesurer la réponse de ce récepteur à un ligand. Le plus simple est de mesurer l'activité transcriptionnelle du récepteur en présence ou en absence de ligand. Mais pour ce faire, on doit connaître la séquence spécifique de fixation à l'ADN du récepteur X. On peut donc réaliser un récepteur chimère entre X et un récepteur dont on connaît le motif de reconnaissance sur l'ADN, par exemple le récepteur des oestrogènes. Si on remplace le domaine responsable de la fixation à l'ADN de X par celui du récepteur des oestrogènes, on obtiendra une molécule chimérique capable de fixer le ligand de X, mais également capable de reconnaître sur l'ADN les éléments de réponse aux oestrogènes. Ainsi, en présence du ligand de X, le récepteur chimère va activer la transcription d'un construit indicateur dont l'activité est facilement mesurable. Cette technique a été utilisée dans de nombreux cas par exemple pour le récepteur  $\beta$  de l'acide rétinoïque que nous avons évoqué plus haut (Brand et al., 1988).

Cependant cette méthode doit être interprétée avec précaution parce qu'un ligand capable d'augmenter l'activité transactivatrice d'un récepteur n'est pas forcément le vrai ligand de ce récepteur. Par exemple, si la dopamine est capable d'activer COUP-TF, elle n'en est pas le ligand puisqu'aucune fixation de la dopamine radioactive sur COUP-TF n'a pu être observée (Power et al., 1991a). Dans ce cas, l'activation doit passer par des intermédiaires, comme, pour COUP-TF, la phosphorylation de certaines protéines cellulaires. Cet effet de la dopamine a d'ailleurs été récemment démontré pour d'autres récepteurs nucléaires comme celui de la progestérone qui est phosphorylé en C-terminal suite à l'action de la dopamine (Power et al., 1991b). Le même type de débat est actuellement en cours pour le récepteur aux rétinoïdes RXR (Mangelsdorf et al., 1990 ; Yu et al., 1991 ; Levin et al., 1992).

Ces dernières années, la méthode du "finger swap" est devenue moins utile parce que, suite aux travaux sur la manière dont les récepteurs reconnaissent spécifiquement leur cible sur l'ADN, on peut désormais prévoir, en analysant la séquence peptidique du domaine de fixation à l'ADN, la séquence de fixation d'un récepteur donné (tableau 2). Il suffit donc d'utiliser la séquence cible que l'on pense être reconnue par le récepteur (Mangelsdorf et al., 1990).

### 4) D'autres récepteurs ?

La famille des récepteurs nucléaires est-elle actuellement complète ? Certainement pas. Si de nombreux membres de cette famille sont encore orphelins, il existe également un certain nombre de ligands dont on peut suggérer qu'ils sont reconnus par un récepteur nucléaire. Dans un certain

nombre de cas, l'appartenance des récepteurs de ces ligands à la superfamille des récepteurs nucléaires, ne fait guère de doute (hormone juvénile) mais dans d'autres cas la situation a l'air plus complexe (dioxine). Il est possible qu'il existe aussi des récepteurs nucléaires chez des organismes plus "primitifs" .

L'hormone juvénile, dérivée des terpénoïdes, est une hormone très importante dans la physiologie des arthropodes. Elle est impliquée dans la métamorphose et dans la maturation de leur système reproducteur. On ignore encore si le récepteur de l'hormone juvénile est un récepteur nucléaire, mais certains auteurs ont suggéré, sans preuve directe, que le récepteur orphelin de drosophile E75 (isolé par le groupe de Hogness parce qu'il est régulé par l'ecdysone ; Segraves et Hogness, 1990), pourrait être le récepteur de cette hormone (Ashburner, 1990). Récemment, un groupe a annoncé avoir cloné chez *Manduca* (un papillon) un récepteur potentiel de l'hormone juvénile très proche des récepteurs de l'acide rétinoïque. Si l'on montre qu'effectivement cette molécule est le récepteur de l'hormone juvénile, cela confirmerait les données récentes qui suggèrent que certains rétinoïdes de *Manduca* ont une faible activité d'hormone juvénile (cité dans Hoffman et Joosse, 1991).

Dans le cas de la dioxine, qui bien sûr n'est pas une molécule "naturelle", la situation est plus complexe. Cette molécule a tout pour être le ligand d'un récepteur nucléaire. Son récepteur se fixe directement à l'ADN, via des éléments qui font penser aux éléments de réponse aux hormones, et peut même se complexer avec les protéines de choc thermique de type hsp90, comme le récepteur des glucocorticoïdes (Wilhelmsson et al., 1990 ; Cuthill et al., 1991). De plus, le récepteur de la dioxine semble activer directement certains gènes codant des enzymes à cytochrome P450. Certains auteurs ont proposé que le récepteur orphelin PPAR (Issemann et Green, 1990), qui est activé par des molécules capables d'induire la prolifération des peroxyosomes, pourrait être un récepteur de la dioxine (Johnson, 1991). En effet, ces molécules, (la nafénopine par exemple), sont également capables d'activer la transcription de gènes codant des enzymes à cytochrome P450 . Mais on ne sait pas si le produit du gène PPAR lui même est capable d'induire la transcription de ces gènes (Johnson 1990 ; voir aussi les résultats récents de Tugwood et al., 1992).

De plus, on a isolé récemment un des composants de ce récepteur, appelé Arnt, qui n'est pas un récepteur nucléaire mais une protéine contenant une structure typique "région basique-hélice-boucle-hélice". Cette protéine Arnt présente des homologies avec le produit de l'oncogène *myc*. Elle est donc sans doute capable de fixer l'ADN et de dimériser avec des protéines dont, peut être, l'énigmatique récepteur de la dioxine (Hoffman et al., 1991). Il semble donc qu'elle soit un composant d'un complexe récepteur formé de plusieurs sous-unités. Les relations entre cette protéine Arnt, le récepteur de la dioxine et le récepteur orphelin PPAR restent à étudier.

## B) STRUCTURE DES RECEPTEURS NUCLEAIRES.

C'est la comparaison des tout premiers récepteurs clonés qui a permis de mieux comprendre l'organisation structurale de ces molécules. L'étude des homologies entre les récepteurs des oestrogènes humains et aviaires a permis de mettre en place une terminologie qui est utilisée pour tous les membres de la superfamille (Krust et al., 1986). En effet, si l'on compare les récepteurs des oestrogènes humain et aviaire, qui ont globalement 80% de similitude en acides aminés, on peut noter différentes régions présentant des degrés d'homologie différents. De l'extrémité N-terminale à l'extrémité C-terminale, on note l'existence de 6 domaines, notés de A à F, présentant des degrés d'homologie divers : les domaines A, C et E sont les plus conservés (87%, 100% et 94% d'homologie entre eux respectivement) et les domaines B, D et F sont les moins conservés (56%, 38% et 41% d'homologie respectivement).

Si on compare les récepteurs des oestrogènes (ER), des glucocorticoïdes (GR) et v-erbA on note que, à peu de chose près, les mêmes régions définissent des zones plus ou moins conservées. En effet, on retrouve le domaine C conservé (61,5% d'homologie entre ER et GR et 53% entre ER et v-erbA), le domaine E moins conservé (30% entre ER et GR, 17% entre ER et v-erbA) et les domaines B, D et F qui ne sont pas conservés d'un récepteur à l'autre. La seule différence provient du domaine A qui, s'il est très conservé entre le récepteur des oestrogènes humain et aviaires, ne présente plus d'homologie entre ER, GR et v-erbA (Krust et al., 1986).

Cette étude des homologies de séquence entre domaines a permis de confirmer qu'il existe plusieurs domaines importants au sein des récepteurs nucléaires. Par clivage protéolytique, on a par exemple isolé un peptide de 39 kDa capable de fixer à la fois l'ADN et l'hormone (Wrange et al., 1984). Un peptide contenant à la fois les domaines C, D et E a un poids moléculaire de 39 kDa, ce qui confirme très bien ces données biochimiques. La caractérisation fonctionnelle de ces domaines a parfaitement confirmé ces idées (Giguere et al., 1986 ; Kumar et al., 1986 ; 1987). Par des expériences de délétions, ces auteurs ont pu montrer que le domaine C correspondait bien à la région interagissant avec l'ADN, tandis que le domaine E fixe l'hormone via une hypothétique "poche hydrophobe". Le domaine D est une charnière flexible entre les domaines C et E. On peut y introduire des mutations sans dommage pour la fonction du récepteur (Kumar et al., 1987)

Ainsi le cadre est en place. Les récepteurs nucléaires forment une vaste superfamille de protéines complexes toutes construites sur le même modèle, avec des domaines fonctionnels indépendants. Ces récepteurs fonctionnent en général comme des facteurs de transcription dépendant de leur ligand. Ils interagissent directement avec les régions régulatrices de leurs gènes cibles en se fixant spécifiquement sous forme de dimères sur de courtes séquences appelées HRE pour "Hormone Responsive Elements".

Il convient maintenant de mieux comprendre comment fonctionne dans le détail un récepteur nucléaire pour transduire le message hormonal.

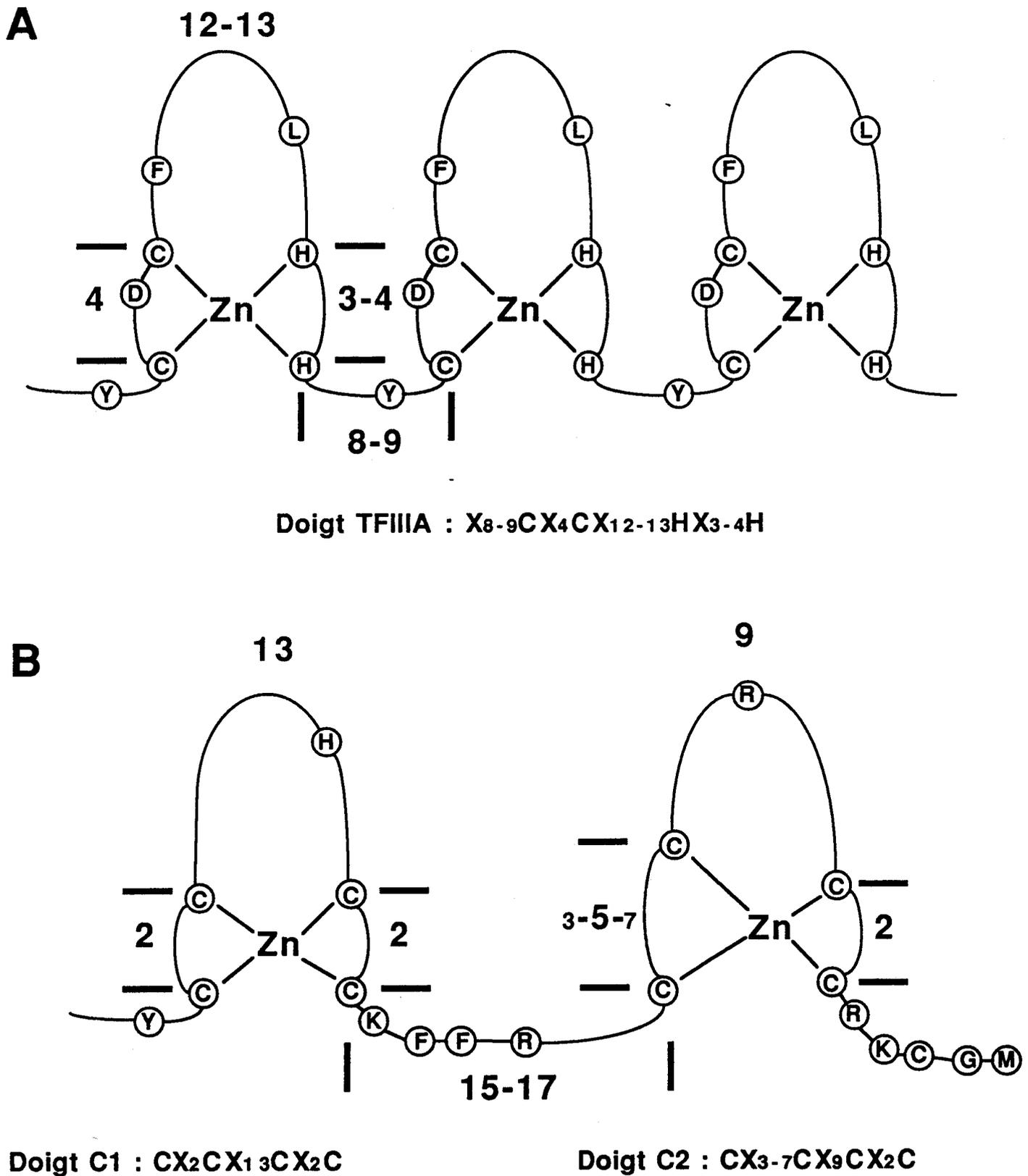


Figure 6 : (A) Structure des doigts de zinc "classiques" (type TFIIIA) et (B) structure des doigts de zinc des récepteurs nucléaires. Pour chaque cas, on a résumé la structure du doigt par une formule cabalistique simple. Pour les récepteurs nucléaires, seuls les acides aminés strictement conservés ont été indiqués.

## C) LA FIXATION A L'ADN.

Pour commencer notre dissection des récepteurs nucléaires nous souhaitons nous pencher d'abord sur le domaine qui est à la fois le mieux conservé et le mieux connu : le domaine C qui permet la fixation à l'ADN. Dans un second temps, nous aborderons le domaine E et nous discuterons de fonctions plus versatiles des récepteurs nucléaires comme la transactivation.

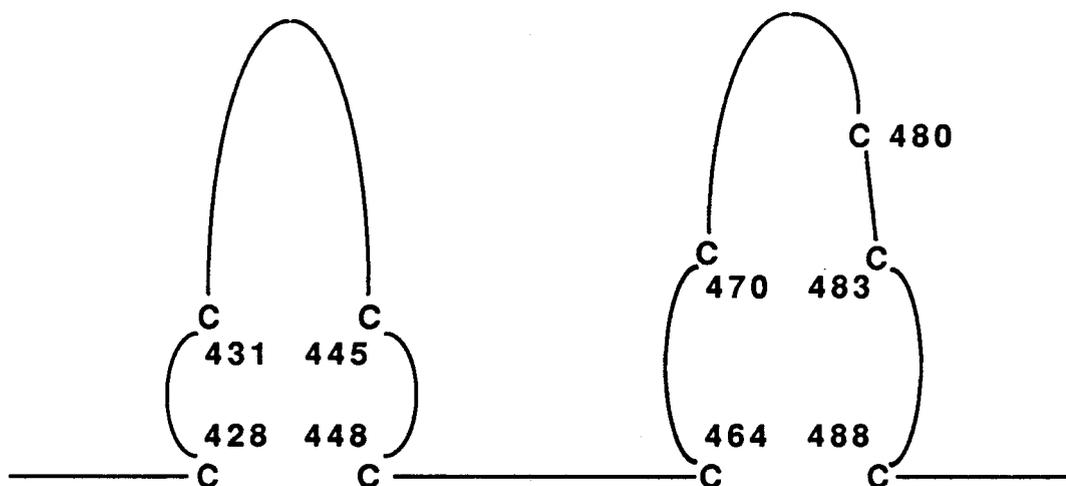
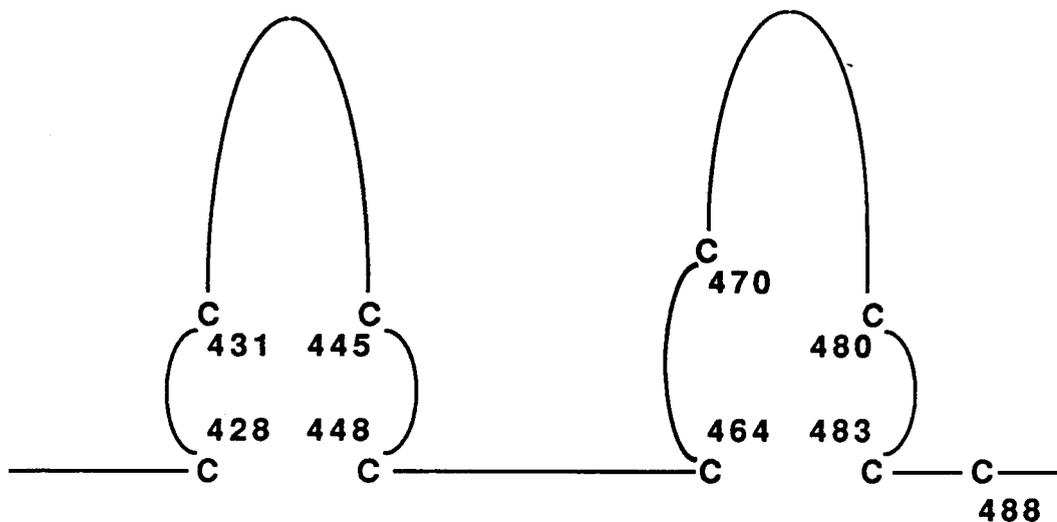
### 1) Anatomie moléculaire des doigts de zinc.

#### a) Doigts de zinc et doigts de zinc...

En scrutant la séquence du domaine C, les groupes d'Evans et Chambon ont noté l'existence de cystéines conservées et répétées d'une façon relativement régulière (Weinberger et al., 1985 ; Krust et al., 1986). L'existence de ces cystéines permettait de faire une analogie entre les récepteurs nucléaires et des facteurs de transcription bien connus qui possèdent des structures dites en "doigt de zinc" avec des cystéines et des histidines répétées (Klug et Rhodes, 1987). L'exemple type de cette famille est le facteur de transcription TFIIIA, qui possède 9 de ces doigts (Ginsberg et al., 1984 ; Miller et al., 1985). Lors de la découverte de ce facteur impliqué, chez le xénope, dans la transcription des ARN 5S, le groupe d'Aaron Klug a noté que la protéine TFIIIA fixe des atomes de zinc et qu'elle possède des répétitions de cystéines et d'histidines - des acides aminés connus pour fixer des atomes de Zinc. Chaque structure en doigt contient 2 cystéines et 2 histidines, l'atome tétravalent de zinc se fixant à ces quatre résidus. Depuis, d'autres protéines à doigt de zinc ont été découvertes et certaines contiennent uniquement des cystéines, comme la protéine de levure GAL 4 ou comme les récepteurs nucléaires (voir la revue de Klug et Rhodes, 1987). Les doigts de zinc "canoniques", comme ceux de TFIIIA, se présentent en général sous la forme d'une séquence de 30 acides aminés environ, ayant la structure présentée sur la figure 6. Un certain nombre d'acides aminés sont conservés en plus des cystéines et des histidines impliquées dans la fixation de l'atome de zinc. Chaque doigt fixe un atome de zinc et, pour TFIIIA par exemple, le nombre de doigts (9) correspond bien au nombre d'atomes de zinc fixé par protéine (entre 7 et 11 ; Miller et al., 1985). Souvent, les doigts de zinc sont présents en assez grand nombre dans la protéine, l'exemple le plus spectaculaire venant de la protéine Xfin du xénope (Ruiz i Atalba et al., 1987) qui contient 37 doigts de zinc !

Il faut déjà insister sur l'originalité des doigts de zinc des récepteurs nucléaires par rapport aux nombreux doigts presque identiques d'une protéine comme TFIIIA. Il y a en effet seulement deux doigts de zinc chez les récepteurs nucléaires et ces deux doigts n'ont pas la même structure. Comme le montre la figure 6 les acides aminés conservés dans chaque doigt diffèrent entre C1, C2 et TFIIIA. Brièvement, C1 est plus riche en acides aminés hydrophobes, tandis que C2 (comme TFIIIA) est plus riche en acides aminés basiques (Green et Chambon, 1988).

Pour affirmer que la structure du domaine C des récepteurs nucléaires correspondait



**Figure 7 : Les deux schémas possibles de coordination de l'atome de zinc du doigt C2. Pour les récepteurs des glucocorticoïdes et des oestrogènes, c'est le schéma du haut qui est correct. La numérotation se réfère au récepteur des glucocorticoïdes de Souris (d'après Danielsen, 1991).**

bien à des doigts de zinc, deux points ont été démontrés : (i) que ce domaine était responsable de la fixation spécifique à l'ADN ce qui a été très vite fait et (ii) que deux atomes de zinc se fixaient sur les récepteurs nucléaires. En produisant de grandes quantités du domaine C du récepteur des glucocorticoïdes, il a été possible de réaliser des expériences de spectroscopie par EXAFS ("Extended X-ray Absorption Fine Structure") qui ont permis de montrer que ce domaine isolé fixe bien deux atomes de zinc, chacun étant coordonné par quatre cystéines (Freedman et al., 1988). Ces deux atomes de zinc sont d'ailleurs requis pour la conformation correcte de la molécule dont dépend la fixation spécifique à l'ADN.

#### b) Comment reconnaître l'ADN ?

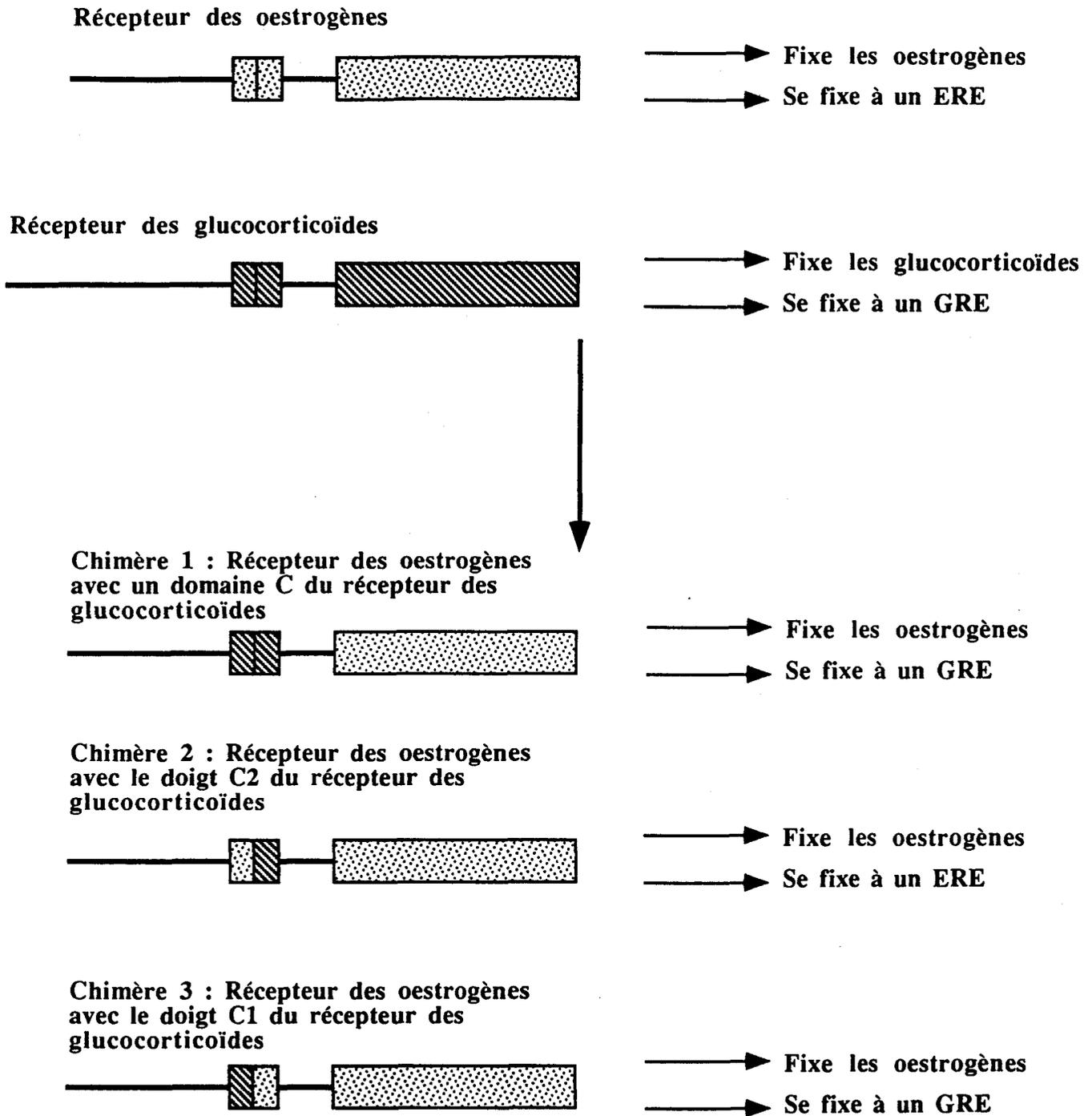
Quelles sont les zones des doigts de zinc C1 et C2 qui sont importantes pour la fixation à l'ADN ? A l'évidence, on peut les classer en trois groupes : (i) les résidus cystéines qui sont les signaux distinctifs du doigt de zinc (il est important de noter qu'il y a 9 cystéines [4 dans C1 et 5 dans C2] strictement conservées entre tous les récepteurs nucléaires et non pas 8 comme on aurait pu s'y attendre) (ii) les résidus très conservés (plus de 90%) entre les différents récepteurs et enfin (iii) les autres résidus.

Pour simplifier, nous prendrons un exemple type, celui du récepteur des glucocorticoïdes pour lequel un résumé très complet des différentes mutations effectuées est disponible (voir Danielsen, 1991 et aussi Green et Chambon, 1991 pour le même genre d'informations pour le récepteur des oestrogènes).

(i) Les cystéines. Bien sûr elles sont toutes les 9 requises pour la fixation correcte à l'ADN. Certaines cystéines semblent être remplaçables par des histidines pour former des doigts de zinc mixtes à His/Cys, mais pas toutes (Severne et al., 1988 ; Green et Chambon, 1987).

Deuxième interrogation concernant ces cystéines : dans le doigt C2, quelles sont celles effectivement impliquées dans la fixation de l'atome de zinc ? Puisqu'il y a 5 cystéines dans C2, deux schémas de coordination semblent possibles a priori (figure 7), et ont été proposés (Freedman et al., 1988 ; Severne et al., 1988). Lequel des deux est-il correct ? Par des expériences de mutagenèse qui ont été confirmées depuis par la résolution de la structure tridimensionnelle, on a pu montrer que c'est le modèle excluant la cys 488 qui est le bon (Severne et al., 1988 ; Härd et al., 1990 ; Luisi et al., 1991 ; voir Danielsen, 1991 pour la numérotation).

(ii) Les résidus conservés. Bien sûr, dans de nombreux cas la mutation d'un acide aminé conservé inactive la molécule. Cependant, des expériences de mutagenèse dans le domaine C ont permis d'observer que la mutation de certains acides aminés bien conservés n'inactivait pas forcément les récepteurs (Danielsen, 1991). C'est le cas de l'histidine 439 absolument conservée chez tous les récepteurs nucléaires clonés à ce jour. Sa mutation en tyrosine, glycine ou glutamine chez le récepteur des glucocorticoïdes n'affecte pas l'activité transcriptionnelle du récepteur (Severne et al., 1988). La fixation directe à l'ADN de ce mutant n'a pas été mesurée,



**Figure 8 : Principe de l'échange de doigt ou "finger swap". En réalisant ces expériences avec les doigts C1 ou C2 seuls, Green et al., 1988 ont démontré que C1 reconnaissait spécifiquement la séquence des HRE.**

mais si le récepteur est fonctionnel en terme d'activité transcriptionnelle, il est raisonnable de penser qu'il se fixe à peu près correctement à l'ADN.

(iii) Les autres résidus. La plupart de ces résidus n'ont pas été testés. Il n'est pas logique de commencer une expérience de mutagenèse par les résidus qui sont apparemment les moins importants. A part les acides aminés impliqués dans la spécificité de fixation à l'ADN sur lesquels nous reviendrons et qui sont évidemment variables d'un récepteur à l'autre, il a été montré que certains résidus peu ou pas conservés semblaient néanmoins importants pour le maintien de l'activité du récepteur. C'est par exemple le cas de la sérine 432 qui est une asparagine chez le récepteur des oestrogènes et une glycine chez la majorité des autres récepteurs. (Schena et al., 1989).

#### c) Le domaine C est multifonctionnel.

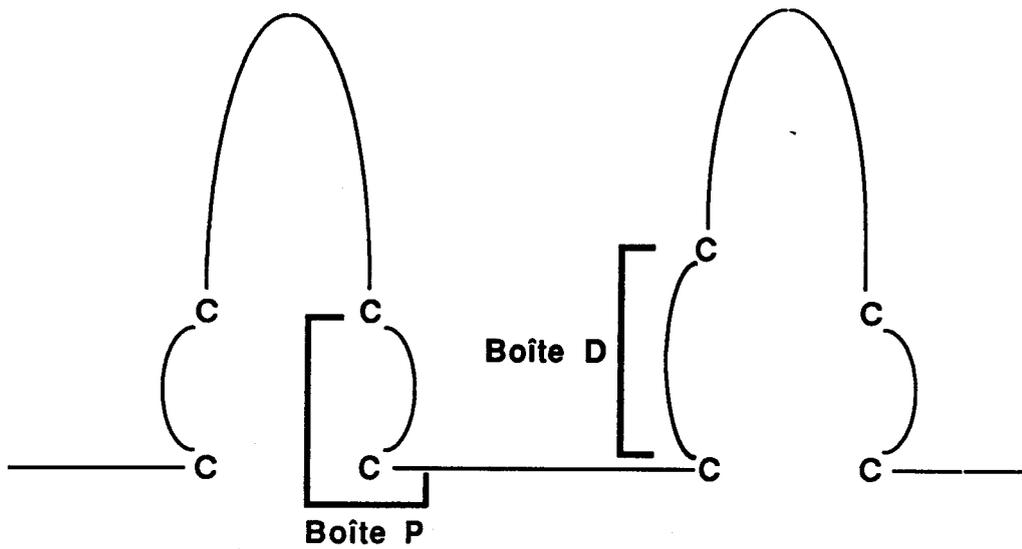
De nombreuses expériences suggèrent qu'outre la fixation à l'ADN, le domaine C est important pour d'autres fonctions puisqu'on obtient souvent des molécules capables de se fixer à l'ADN mais dénuées d'activité transcriptionnelle. On peut mentionner par exemple la transactivation, l'interaction avec d'autres protéines et la dimérisation (les trois choses peuvent être liées).

On sait que le domaine N-terminale (noté A/B) du récepteur des glucocorticoïdes est requis pour l'activité transcriptionnelle du récepteur (Giguere et al., 1986). Cependant, certains auteurs ont démontré qu'une partie de la capacité transactivatrice du récepteur des glucocorticoïdes se trouve dans le domaine C (Godowski et al., 1988). De plus, par une élégante approche de mutagenèse au hasard suivi d'un criblage phénotypique chez la levure, Schena et al., 1989 ont identifié certaines mutations ponctuelles situées au sommet du doigt C2 qui pourraient affecter des liaisons protéines-protéines importantes pour l'activité du récepteur. Des mutations en ces endroits (R476, R477, N479 mais aussi P481 et A482 identifiées par Godowski et al., 1989) produisent des récepteurs possédant toujours une activité de fixation à l'ADN, mais qui sont cependant beaucoup moins actifs sur le plan transcriptionnel. De manière intéressante, cette région ainsi définie (Schena et al., 1989) correspond à la région dotée d'une activité transactivatrice observée précédemment par d'autres auteurs et appelée enh1 (Godowski et al., 1988).

Après avoir démontré que les récepteurs nucléaires sont capables de dimériser (Kumar et Chambon, 1988 ; Tsai et al., 1988) on a observé que le domaine C jouait un rôle important dans ce processus en association avec le domaine de fixation du ligand (domaine E). En effet, des récepteurs délétés de leur domaine E sont toujours capables de dimériser, comme le sont également des domaines C seuls.

#### d) Doigts de zinc et spécificité ?

Les résultats que nous venons de présenter sont tout à fait en accord avec les expériences qui suggèrent que les doigts de zinc C1 et C2 ont des rôles fonctionnels différents et



	SEQUENCE	FIXATION	
	Boîte P	GRE	ERE
GR	CGSCKV	+	-
GR Mutant (Umesono et Evans, 1989)	CESCKV	+	+
GR Mutant (Danielsen et al., 1989)	CEGCKV	-	+
ER	CEGCKA	-	+
ER Mutant (Mader et al., 1989)	CGSCKV	+	-

Figure 9 : Localisation des boîtes P et D sur les doigts de zinc des récepteurs nucléaires et résumé simplifié des expériences prouvant le rôle de la boîte P dans la spécificité de fixation aux HRE. La boîte D joue également un rôle dans ce processus (voir texte).

complémentaires et peuvent être considérés comme des "sous-domaines" (Green et al., 1988). En effet, tout comme on peut réaliser par "finger swap" des récepteurs chimères possédant des domaines C et E venant de récepteurs différents, il est possible de faire des chimères plus fines où l'on met uniquement le doigt C1 ou le doigt C2 du récepteur des glucocorticoïdes à la place du doigt correspondant du récepteur des oestrogènes (figure 8). On peut ainsi disséquer les rôles de C1 et de C2. A l'aide de ce type d'analyse, les groupes de Pierre Chambon et de Walter Wahli ont pu montrer que le doigt C1 était directement impliqué dans la fixation spécifique à l'ADN, tandis que C2 avait plutôt un rôle de "soutien" en étant impliqué dans des liaisons non spécifiques à l'ADN (Green et al., 1988). Ces expériences suggèrent également que C2 peut avoir un rôle à jouer dans la dimérisation des récepteurs fixés sur l'ADN.

Nous avons vu comment les récepteurs reconnaissent l'ADN, et comment les deux doigts semblaient "se répartir les rôles". Il nous faut donc maintenant étudier comment se fait la reconnaissance spécifique des éléments de réponse (HRE ou "Hormone Responsive Element") par les récepteurs. Il faut noter que la connaissance des mécanismes de reconnaissance des HRE par les récepteurs a permis, plus tard, de mieux cerner la spécificité des HRE eux-mêmes. Il y a donc sans arrêt dialogue, confrontation, entre ces deux approches : l'étude du récepteur et l'étude de sa cible. Cela permet, à l'aide des structures tridimensionnelles récemment proposées, d'établir un modèle qui explique en partie la spécificité d'action des récepteurs nucléaires.

Donc, comment les récepteurs reconnaissent-ils les trois grands types de HRE décrits (GRE, ERE et TRE) ? Quels sont les acides aminés impliqués dans ce phénomène de reconnaissance ? D'après les expériences de Green et al., 1988, il semblait bien que ce soient des acides aminés situés dans C1, puisque ce doigt est responsable de la fixation spécifique à l'ADN. Trois groupes sont parvenus en même temps à montrer comment se faisait la différence entre le GRE et le ERE par les deux récepteurs considérés : celui des oestrogènes et celui des glucocorticoïdes (Mader et al., 1989 ; Danielsen et al., 1989 ; Umesono et Evans, 1989 ; voir les revues de Berg, 1989 et Danielsen, 1991). Bien que ERE et GRE aient des séquences très proches, la différence fonctionnelle entre ces deux séquences se fait clairement, tant *in vitro* qu'*in vivo*. Ces trois groupes ont montré que l'on pouvait changer la spécificité d'un récepteur en changeant seulement deux acides aminés. L'élément de surprise a été la localisation de ces acides aminés. En effet, ils sont situés dans l'articulation C-terminale de C1 ("knuckle" en anglais), entre les deux cystéines (figure 9), alors que la plupart des chercheurs pensaient que ces acides aminés étaient dans le doigt proprement dit. Ainsi, si on mute le récepteur des glucocorticoïdes à cet endroit en remplaçant les résidus glycine et sérine en résidus glutamate et glycine (présents dans le récepteur aux oestrogènes), on obtient un récepteur capable de transactiver un vecteur reporter via un ERE et non via un GRE (Danielsen et al., 1989). L'opération inverse marche aussi et les chercheurs du laboratoire de Pierre Chambon ont ainsi modifié la spécificité du récepteur des oestrogènes en celle du récepteur des glucocorticoïdes en modifiant trois acides aminés : le glutamate et la glycine déjà cités et une alanine située deux résidus sous la dernière cystéine de C1 (changement de la séquence

**TABLEAU 2**

**Classification des récepteurs nucléaires selon leur spécificité de fixation à l'ADN**

<b>Sequence de la boîte P</b>	<b>Classes (Forman and Samuels, 1990)</b>	<b>Gènes</b>
<b>GSCKV</b>	<b>I</b>	<b>AR GR MR PR</b>
<b>EGCKG</b>	<b>II</b>	<b>E75 EAR1 ECR RXR H2RIIBP USP TR2 VDR NGF1B PPAR RARA RARB RARG THRA THRB</b>
<b>EGCKS</b>	<b>III</b>	<b>ARP-1 COUP EAR2 EGON KNI KNRL SVP VERBA</b>
<b>EGCKA</b>	<b>IV</b>	<b>ER ERR1 ERR2</b>
<b>ESCKG</b>	<b>NC</b>	<b>FTZ-F1</b>
<b>DGCKG</b>	<b>NC</b>	<b>HNF4</b>
<b>DGCAG</b>	<b>NC</b>	<b>TLL</b>

Classification des récepteurs nucléaires selon leur spécificité de fixation à l'ADN (séquence de la boîte P ; d'après Forman et Samuels, 1990. Les récepteurs qui n'ont pas été classés par ces auteurs sont indiqués par NC.

cEGckA en cGSckV ; Mader et al., 1989). De plus, le stade intermédiaire où un seul acide aminé (le premier) seulement de l'articulation de C1 est modifié (cESckV) permet d'obtenir un récepteur fixant à la fois un ERE et un GRE (Umesono et Evans, 1989). Cette boîte qui code la spécificité des récepteurs a été appelée boîte P (pour proximale) par le groupe de Ronald Evans.

La situation n'est tout de même pas aussi simple et une autre région, appelée boîte D (pour distale), est également importante (figure 9). Cette région est située dans la première articulation de C2 (Umesono et Evans, 1989 ; Danielsen et al., 1989). Ainsi, pour fabriquer un récepteur des glucocorticoïdes capable de se fixer au TRE, il faut non seulement changer la boîte P (changement cGSckV en cEGckA), mais aussi modifier la boîte D constituée de 5 acides aminés (changement cAGRNDc en cKYEGKc). Cette région est impliquée dans la dimérisation (voir le tableau 2).

On peut donc schématiser l'association ADN-récepteur de la manière suivante : les récepteurs se fixent sous forme de dimère sur leur site HRE palindromique, chaque récepteur est fixé à un demi-palindrome (Green et al., 1988 ; Green et Chambon, 1991 ; Danielsen 1991). Le doigt C1 reconnaît spécifiquement les bases d'ADN spécifiques du HRE. Les deux récepteurs se font face sur l'ADN, et chacun des C2, d'une part renforce la fixation à l'ADN par des liaisons non spécifiques et d'autre part, via la boîte D, est en contact avec le C2 de son jumeau par des interactions protéines-protéines. La spécificité de reconnaissance d'un récepteur en ce qui concerne la dimérisation (de manière à éviter la formation d'hétéro-dimères intempestifs), est donc codée par le deuxième doigt. L'étude de la structure tridimensionnelle a pour l'essentiel confirmé ce modèle (Schwabe et al., 1991 ; Kaptein, 1991 ; figure 13).

## 2) Les éléments de réponse sur l'ADN (HRE) et la spécificité.

### a) Les éléments de réponse aux hormones ont une structure palindromique.

L'idée selon laquelle les récepteurs nucléaires se fixent à l'ADN de manière spécifique est assez vieille, puisqu'elle a été proposée par Keith Yamamoto et Bruce Alberts au milieu des années 1970 (Yamamoto et Alberts, 1976). Par exemple, en utilisant des virus dont l'expression est régulée par les stéroïdes, comme le MMTV (Mouse Mammary Tumor Virus), le groupe de Miguel Beato a mis en évidence la fixation du récepteur des glucocorticoïdes sur des sites précis de la région promotrice (LTR) de ce virus (Scheidereit et al., 1983). Des séquences consensus de fixation de récepteurs ont pu être proposées, en comparant les promoteurs de gènes régulés ou en réalisant des expériences de délétions (Mulvihill et al., 1982; Scheideret et al., 1983).

Le premier site de réponse aux hormones (appelé HRE pour Hormone Response Element) réellement bien caractérisé, a été le GRE (Glucocorticoid Response Element) qui est constitué de 15 paires de bases présentant une structure de palindrome imparfait avec la séquence consensus AGA/GACANNNTGTACC (voir la figure 10 ; Martinez et Wahli, 1991). D'une manière surprenante, il a été démontré que la séquence du GRE est la même que celle nécessaire

**HRE : AGGN<sub>a</sub>CAN<sub>x</sub>TGN<sub>b</sub>CCT**



Si : Na = A

Nb = T

Nx = 3

**GRE : AGGACANNNTGTCCT**



Glucocorticoid Responsive Element  
Correspond également à un PRE, un MRE ou un ARE

Si : Na = T

Nb = A

Nx = 3

**ERE : AGGTCANNNTGACCT**



Estrogen Responsive Element

Si : Na = T

Nb = A

Nx = 1

**EcRE : AGGTCANTGACCT**



Ecdysone Responsive Element

Si : Na = T

Nb = A

Nx = 0

**TRE : AGGTCATGACCT**



Thyroid hormone Responsive Element

Figure 10 : Une vision simplifiée des HRE, éléments de réponse aux hormones (d'après Martinez et Wahli, 1991).

pour l'activation transcriptionnelle de gènes cibles par la progestérone, les minéralocorticoïdes ou les androgènes. Le GRE est donc également un PRE, un MRE et un ARE ! La spécificité d'action de ces différentes hormones, n'est donc pas déterminée par le HRE.

L'identification de l'élément de réponse aux oestrogènes (ERE ; Klein-Hitpass et al., 1986 ; Martinez et al., 1987) puis celui des autres hormones (hormones thyroïdiennes, TRE ; Glass et al., 1987 ; Koenig et al., 1987 ; Vitamine D, VDRE ; Morrison et al., 1989 ; Acide rétinoïque, RARE ; Umesono et al., 1988) a confirmé la parenté structurale existant entre ces divers éléments. Comme pour le GRE, plusieurs voies hormonales aboutissent à l'activation d'éléments structuraux identiques sur l'ADN ; ainsi, il a été montré que le RARE était identique au TRE. Cependant, il existe aussi des RARE spécifiques, différents des TRE : c'est le cas du RARE présent dans le promoteur du gène du récepteur de l'acide rétinoïque  $\beta$  (De Thé et al., 1990b ; voir aussi les exemples de Vasios et al., 1989 et Lucas et al., 1991). Il est tout à fait clair que l'on peut facilement transformer un HRE d'un type en un autre type, par un petit nombre de mutations ponctuelles. C'est ce qui a été réalisé par de nombreuses équipes qui ont par exemple transformé un ERE en GRE en y introduisant seulement deux mutations ponctuelles (Klock et al., 1987 ; Martinez et al., 1987).

L'équipe de Walter Wahli à Lausanne s'est également penchée sur l'élément de réponse à l'ecdysone et a montré que ce EcdRE avait, lui aussi, la structure classique des HRE (Martinez et al., 1991). Cet élément ressemble en fait beaucoup à un ERE et ces auteurs ont pu, en introduisant deux nucléotides entre chaque demi-palindrome, transformer ce EcdRE en ERE : c'est donc l'espacement entre les deux demi-palindromes qui semble être un facteur crucial de distinction entre un EcdRE et un ERE. Ce travail montre de plus que, puisque les HRE sont semblables chez les vertébrés et chez les insectes, on peut en conclure que ce sont des éléments anciens, apparus très tôt dans l'évolution des êtres multicellulaires (Martinez et Wahli, 1991).

Si les HRE présentent le plus souvent une séquence palindromique ou apparentée, "dans la nature", les HRE ne sont pas strictement palindromiques. C'est ainsi que l'on peut transformer le TRE imparfaitement palindromique du gène de l'hormone de croissance de rat en un TRE parfaitement palindromique. Ce TRE palindromique permet une meilleure activation de la transcription qu'un TRE "naturel" (Glass et al., 1988). Le même genre de remarque est valable pour les ERE (Klein-Hitpass et al., 1986 ; Kumar et Chambon, 1988 ; Martinez et Wahli, 1989 ; Berry et al., 1989 ; revue dans Green et Chambon, 1991).

Puisque maintenant de très nombreux HRE sont connus (voir Martinez et Wahli, 1991, pour une liste exhaustive des HRE connus en 1990) Martinez et Wahli ont essayé de les classer et de dégager une structure de base pour chaque type (voir figure 10). Cependant, il ne faut pas considérer uniquement la fixation à l'ADN. Ainsi, le récepteur des hormones thyroïdiennes est capable de se fixer à un ERE classique, mais n'est pas capable d'induire une transactivation à partir de cet élément (Glass et al., 1988). De plus, il existe dans le génome de nombreux demi-HRE (un seul "bras" du palindrome est présent) et des HRE qui induisent une répression de l'activité des

Motif répété : AGGTCA =  $\longrightarrow$

1  $\longrightarrow$  N  $\longrightarrow$  RXRE

3  $\longrightarrow$  NNN  $\longrightarrow$  VDRE

4  $\longrightarrow$  NNNN  $\longrightarrow$  TRE

5  $\longrightarrow$  NNNNN  $\longrightarrow$  RARE

$\longrightarrow$  NNN  $\longleftarrow$  ERE

$\longrightarrow$  N  $\longleftarrow$  EcRE

$\longrightarrow$   $\longleftarrow$  TRE, RARE, VDRE

$\longrightarrow$  NNN  $\longleftarrow$  GRE

mais avec le motif répété AGGACA

Figure 11 : La "règle du 3-4-5" de Evans, permettant de déterminer la spécificité d'un HRE à structure de "répétition directe". Cette règle a été récemment complétée pour le récepteur RXR qui reconnaît une répétition directe séparée par 1pb (Umesono et al., 1991; Mangelsdorf et al., 1991). Ces HRE sont comparés aux HRE "classiques" (voir Fig. 9).

gènes qui ont, semble-t-il, une séquence différente des HRE classiques (Levine et Manley, 1989 ; Reinkawitz, 1990). Le moins que l'on puisse dire est qu'il règne une certaine confusion en ce qui concerne la caractérisation et la définition des HRE. Un bon HRE est un élément qui répond à l'hormone considérée ...!

b) HRE et spécificité.

Jusqu'à présent, la spécificité d'action des différentes hormones n'est pas bien comprise au niveau moléculaire. En effet, le TRE palindromique est capable d'induire la réponse des récepteurs des hormones thyroïdiennes et de l'acide rétinoïque. Comment, dans une telle situation, peut-on rendre compte de la spécificité d'action physiologique de ces deux molécules ? Modifier l'espacement entre les bras du palindrome est-il un moyen simple et efficace d'obtenir un HRE très spécifique. Un changement d'espacement d'une base modifie la distance entre les deux parties du site, chacune occupée par un monomère, de 3,4 Angström, et surtout induit une rotation des deux demi sites de 36° autour de la double hélice ; Umesono et al., 1991.

Cette idée a été menée à son aboutissement par les groupes de Ronald Evans et de Michael Rosenfeld en Californie. En même temps, ces chercheurs ont considérablement modifié et enrichi notre connaissance des séquences de réponse aux hormones. L'équipe de Ronald Evans a commencé par analyser le TRE présent dans le promoteur du gène de la chaîne lourde de myosine cardiaque de rat (MHC). Ce TRE a été décrit d'abord comme un élément imparfaitement palindromique qui avait la particularité intéressante de ne pas être inductible par l'acide rétinoïque (Izumo et Mahdavi, 1988). En fait, il contient une répétition directe de la séquence AGGTCA qui est l'élément qui répond effectivement à l'hormone. En changeant l'espacement entre chaque répétition, on modifie complètement le comportement de cet élément (figure 11). Ainsi une répétition directe de ce motif avec un espacement de 4 nucléotides (DR4) comme dans le gène de la MHC confère l'inductibilité à la T3 (c'est donc un TRE). Si les mêmes motifs sont séparés de 5 pb (DR5), on obtient un RARE comme dans le promoteur du récepteur de l'acide rétinoïque  $\beta$  (De Thé et al., 1990b). Si on les sépare de 3 pb (DR3), on obtient un élément de réponse à la vitamine D (VDRE). Ronald Evans propose ainsi la "règle du 3-4-5" qui permet de prédire l'activité d'une répétition directe de l'élément AGGTCA (figure 11). Il faut préciser que ces observations sont bien corrélées à l'activité de fixation des récepteurs sur ces différents éléments (Umesono et al., 1991). Enfin, il semble que la séquence de l'élément AGGTCA joue tout de même un rôle : toutes les substitutions possibles ne sont pas sans effet. Donc l'espacement des répétitions et la séquence de ces répétitions sont importantes et complémentaires.

Le seul petit problème est que le groupe de Michael Rosenfeld a, en même temps, publié une étude aussi détaillée abordant le même problème, d'une façon légèrement différente (Näär et al., 1991). Ces deux études mises côte à côte se complètent assez bien mais se compliquent aussi beaucoup mutuellement !! Ainsi, ces auteurs, ont voulu systématiquement tester l'effet des orientations différentes des séquences de fixation des récepteurs : palindrome, répétition

# Motif répété : AGGTCA

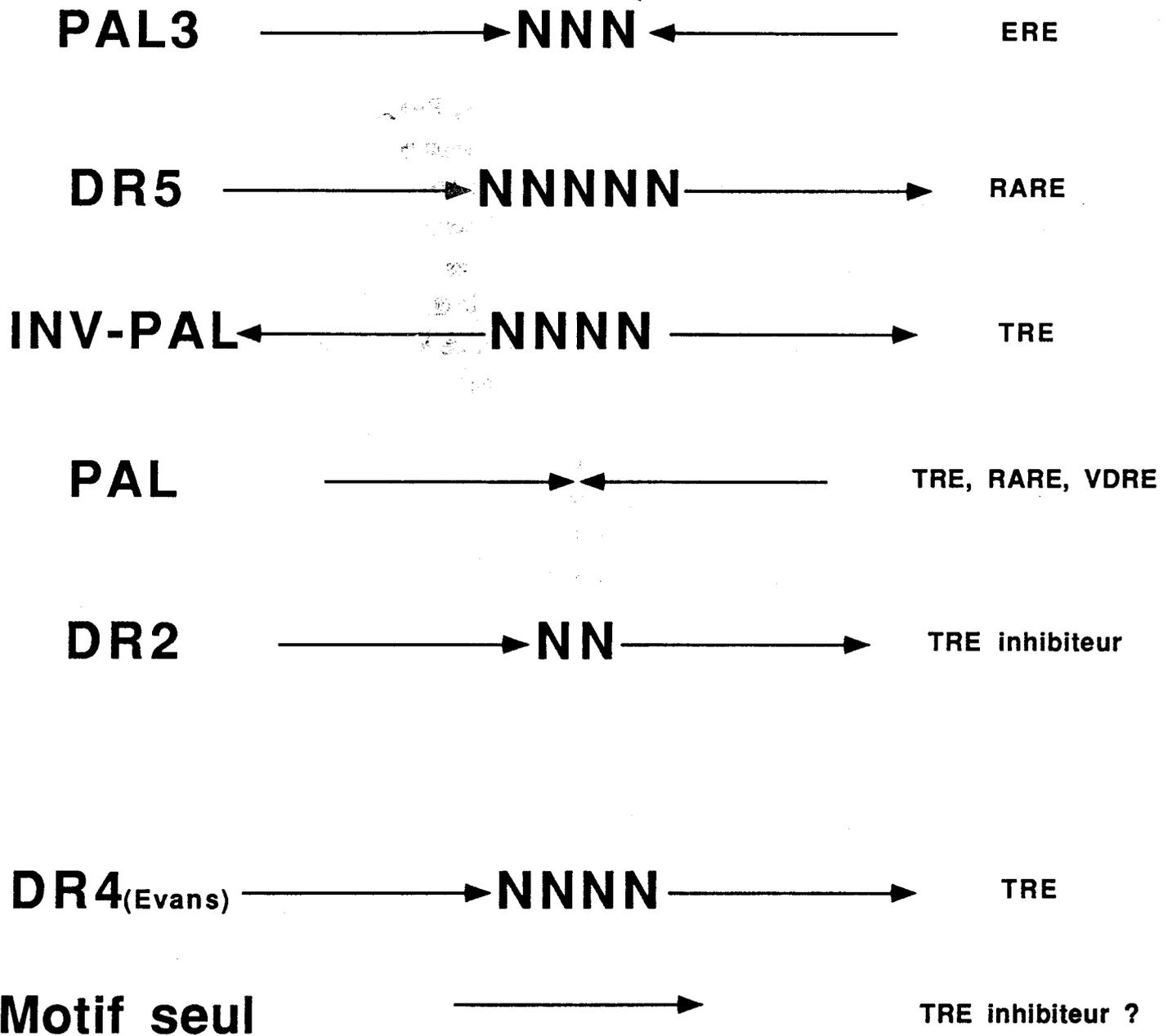
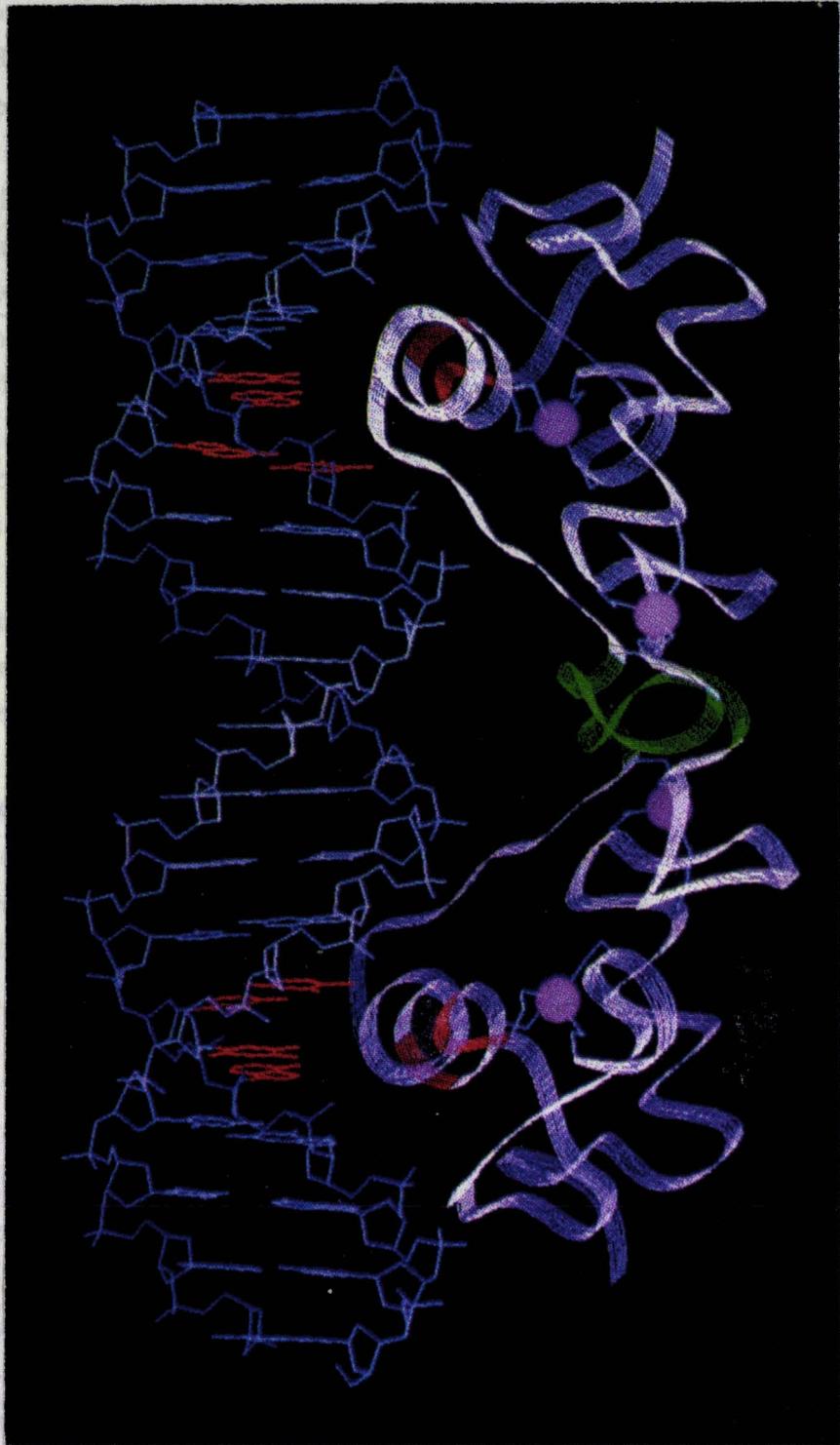


Figure 12 : Une autre vision des HRE : celle du groupe de Michael Rosenfeld (Näär et al., 1991). En bas de la figure deux autres TRE ont été figurés ce qui permet d'illustrer la grande variété de ces éléments.

directe, palindrome inversé, tous espacés de deux manières possibles (sans espace ou avec 3 bases; figure 12). Ils retrouvent grosso modo les mêmes résultats que le groupe d'Evans quant à l'importance de l'écartement des répétitions. Le récepteur des hormones thyroïdiennes se fixe de manière très différente aux constructions testées (il se fixe très fort aux INV-PAL, INV-PAL4 et DR4, il se fixe correctement au PAL et DR et très mal au PAL4 ; voir la figure 12). Si on délète l'extrémité C-terminale du récepteur, on obtient une fixation plus faible partout mais, surtout, on gomme les différences d'affinité entre les 6 éléments testés. C'est donc en partie, la région C-terminale de la molécule (que l'on sait être importante pour la dimérisation) qui joue un rôle dans ce processus de reconnaissance différentiel des séquences. Ceci recoupe assez bien des résultats récents de l'équipe de Magnus Pfahl, qui suggèrent que le domaine E influence la reconnaissance spécifique à l'ADN (Zhang et al., 1991a). Les travaux de Michael Rosenfeld et de son groupe montrent aussi que l'orientation des séquences testées "code" aussi la spécificité d'action des récepteurs : ainsi le PAL4 reconnaît surtout le récepteur des oestrogènes, le DR4 celui de l'acide rétinoïque et le INV-PAL4 celui des hormones thyroïdiennes. Une répétition directe sans espacement (DR) est reconnue par le récepteur aux hormones thyroïdiennes comme un élément régulateur négatif. Un récepteur de la T3 fixé à cet élément transactive en absence d'hormone et inhibe l'activité en présence des hormones thyroïdiennes.

On en arrive donc à une situation d'une grande complexité qui démontre clairement (*sic...*) la plasticité de ce système. D'une part, on a des TRE palindromiques qui obéissent aux règles "classiques" et qui répondent à la fois aux hormones thyroïdiennes, à l'acide rétinoïque et à la vitamine D. Ces éléments vont donc réguler des gènes communs à ces différentes réponses hormonales. D'autre part, on a les répétitions directes (DR3, DR4 ou DR5) qui vont chacune permettre la réponse spécifique à une hormone donnée (vitamine D, hormones thyroïdiennes et acide rétinoïque, respectivement). Il faut insister sur l'extraordinaire variété d'éléments qui sont capables d'être reconnus et activés par le récepteur des hormones thyroïdiennes qui semble capable de se livrer à des contorsions incroyables sur l'ADN (figure 12). Ce récepteur reconnaît en effet le TRE palindromique classique (PAL) qui est également reconnu par les récepteurs de l'acide rétinoïque, le INV-PAL (Näär et al., 1991 ; Baniahmad et al., 1990), le DR4 décrit par le groupe d'Evans (Umesono et al., 1991), l'élément négatif DR sans espace (Näär et al., 1991), et également le demi palindrome souvent décrit comme élément négatif. Clairement, cette extraordinaire complexité cache sûrement notre ignorance de lois importantes gouvernant cet aspect de la physiologie de la réponse hormonale. La règle du 3-4-5 de Evans a le mérite d'une relative simplicité (un élément palindromique est reconnu par plusieurs récepteurs, l'espacement des répétitions directes permet une réponse spécifique) mais hélas, de trop nombreuses exceptions viennent en limiter l'utilisation (Umesono et al., 1991). Des aspects importants, tels le grand nombre de récepteurs et d'isoformes de récepteurs possibles, la formation d'homo- et d'hétéro-dimères, l'interaction avec d'autres protéines etc... permettent à la réponse hormonale d'aboutir à l'extraordinaire variété que l'on observe *in vivo*.



**Figure 13 : Modèle structural du récepteur des glucocorticoïdes, fixé à son site GRE sous forme de dimère (d'après les expériences de RMN de Hård et al., 1990). Sur les récepteurs les boîtes P sont en rouges et les boîtes D en vert. Sur l'ADN, les bases importantes pour la spécificité du HRE sont en rouge**

### 3) RMN et cristallographie : Un modèle d'interaction protéine-ADN.

Le fait que la capacité de se fixer spécifiquement sur l'ADN soit codé par un domaine de petite taille (68 à 70 acides aminés) indépendant du reste de la molécule (Freedman et al., 1988), permet d'étudier la structure tridimensionnelle de ce domaine en utilisant les techniques classiques de biophysique structurale : spectroscopie aux rayons X et résonance magnétique nucléaire. Dans un premier temps, la structure du domaine C du récepteur des glucocorticoïdes en solution a été déterminée par RMN (Härd et al., 1990). Puis, la structure du même domaine du récepteur des oestrogènes, également déterminée par RMN, a permis de comparer les deux récepteurs (Schwabe et al., 1990). Enfin, récemment, l'analyse cristallographique du complexe ADN-récepteur des glucocorticoïdes a permis de valider et préciser les analyses par RMN (Luisi et al., 1991 ; revues dans Schwabe et Rhodes, 1991 ; Kaptein, 1991).

La structure déterminée a les caractéristiques suivantes (Kaptein, 1991 ; figure 13) : le domaine peut être séparé en deux modules qui correspondent aux deux doigts de zinc C1 et C2. Chaque module est constitué d'une zone de nucléation du zinc et d'une hélice alpha amphipatique immédiatement adjacente. Les deux modules sont imbriqués l'un dans l'autre pour former un seul domaine globulaire ; les hélices alpha de chaque module se croisent avec un angle d'environ 90° et se stabilisent l'une l'autre par des interactions entre résidus hydrophobes. La région de la boîte D et du deuxième doigt est mal résolue, sans doute parce qu'elle est très flexible, mais la fixation du récepteur à l'ADN la stabilise probablement lors d'interactions entre monomères. Il est clair que l'hélice alpha du module C1 impliquée dans la reconnaissance spécifique des HRE est plaquée dans le grand sillon de la molécule d'ADN et qu'ainsi, la boîte P, impliquée dans la reconnaissance spécifique du HRE est très proche des bases discriminantes du site (Kaptein, 1991).

Une des conclusions à tirer de cette structure est que les doigts de zinc des récepteurs nucléaires sont organisés différemment des doigts de zinc plus classiques (TFIIIA, par exemple) contenant cystéines et histidines. La structure tridimensionnelle des doigts de TFIIIA est en effet connue et on peut faire d'utiles comparaisons entre les deux structures (Lee et al., 1989 ; Pavletich et al., 1991 ; Schwabe et Rhodes, 1991). La ressemblance entre les différents types de doigts de zinc n'est valable que pour la structure primaire et non pour la structure dans l'espace. En effet, les doigts de zinc adoptent une structure contenant une hélice alpha qui s'étend au milieu même du doigt et non juste après comme c'est le cas pour les récepteurs nucléaires. De plus, cette structure comporte des zones en feuillet  $\beta$ . Enfin, les différents doigts des protéines de type TFIIIA n'interagissent pas entre eux. La structure des protéines à doigts de zinc fixant le DNA est donc variable et certains auteurs parlent de doigts de zinc pour TFIIIA et de "zinc twist" pour les récepteurs nucléaires (Vallee et al., 1991 ; Harrison, 1991).

Il semble bien, à la lumière de ces structures, que la reconnaissance spécifique d'une séquence d'ADN via une hélice alpha soit un phénomène assez fréquent. Pour les récepteurs nucléaires nous avons une hélice plaquée dans le grand sillon et une deuxième hélice en arrière

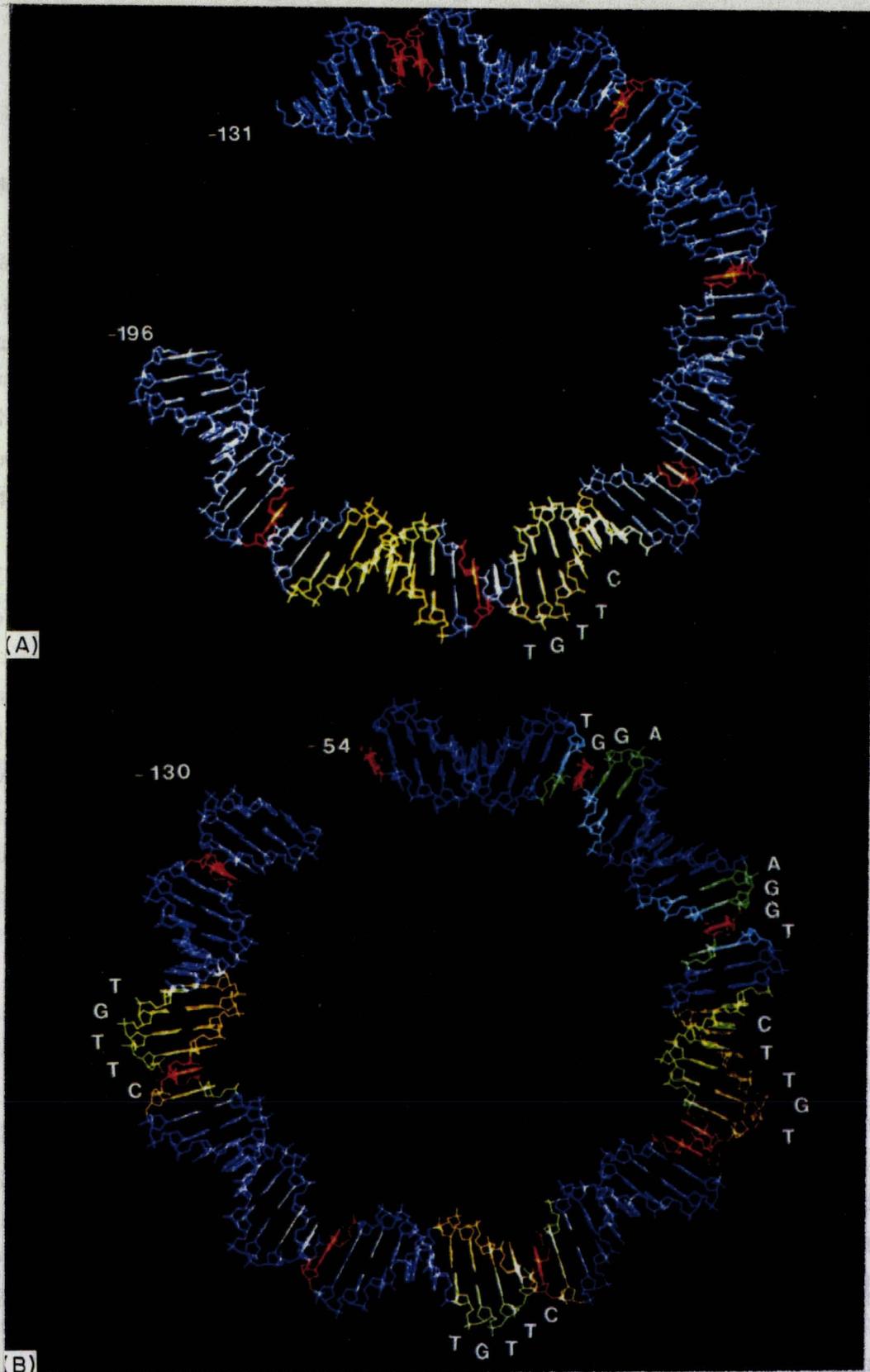


Figure 14 : Représentation de la super hélice formée par le LTR du virus MMTV autour des octamères d'histones. Les HRE sont en jaune, le site NF1 en vert. Les bases hypersensibles à la DNase sont en rouge (d'après Beato et al., 1991).

située perpendiculairement qui stabilise la structure via des interactions hydrophobes et permet, en outre, la reconnaissance de protéines diverses sur son autre face. En fait, cette structure ressemble plus à celle des protéines à "hélice-coude-hélice" (comme les protéines à boîtes homéo) qu'à celle des doigts de zinc "classiques" (Schawbe et Rhodes, 1991). Il y a bien sûr plus d'une manière de présenter une hélice alpha contre de l'ADN mais dans tous les cas, l'hélice située le plus près de l'ADN est plaquée contre une structure stable constituée d'autres hélices ou de feuilletts  $\beta$  rigides. On a ainsi un modèle général (qui bien sûr a ses exceptions !) permettant de rendre compte de la reconnaissance spécifique ADN/protéine.

#### 4) La fixation à l'ADN *in vivo*.

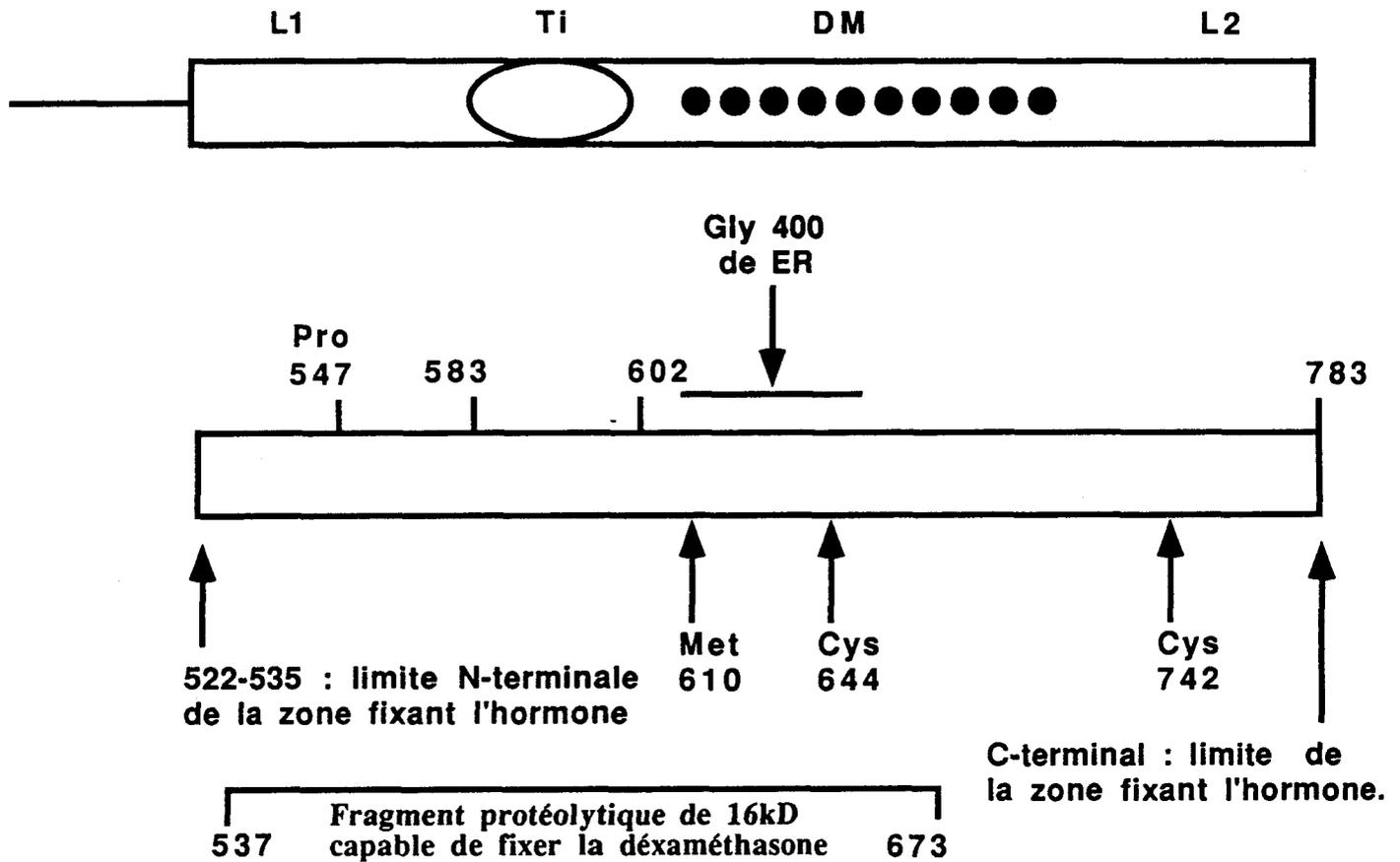
La cinétique de fixation à l'ADN des récepteurs des glucocorticoïdes et de la progestérone a été étudiée par le groupe de Miguel Beato qui a mesuré les vitesses d'association et de dissociation de ces molécules. Ainsi, la vitesse d'association du récepteur à l'ADN est augmentée de 5 fois lorsque l'hormone est présente. De même, la vitesse de dissociation est augmentée 20 à 30 fois en présence de ligand (Schauer et al., 1989 ; Beato et al., 1991). On peut donc supposer que l'hormone va ainsi permettre un balayage ("scanning") plus rapide de l'ADN cible pour trouver des HRE.

Qu'en est-il du rôle éventuel de l'ADN organisé en nucléosome dans la chromatine ? On sait par exemple que les séquences du LTR du provirus MMTV sont organisées en nucléosomes selon un mode très précis (Richard-Foy et Hager, 1987 ; Perlmann et Wrangle, 1988). La structure de cette zone, organisée en 6 nucléosomes, change après un traitement par l'hormone (Cordingley et al., 1987). Si l'on reconstruit *in vitro* un tel arrangement, en utilisant une région de 142pb du LTR du MMTV contenant les 4 HRE, on peut étudier quels sont les sites accessibles au récepteur (Perlmann et Wrangle, 1988 ; Pina et al., 1990a; 1990b). Ainsi, il a été montré que sur les quatre sites existants, les deux sites extrêmes seulement sont accessibles car le grand sillon de la molécule d'ADN est exposé à l'extérieur du nucléosome (Pina et al., 1990b ; Beato et al., 1991 ; figure 14). Ces observations sont également très intéressantes en ce qui concerne l'interaction éventuelle des récepteurs nucléaires avec d'autres facteurs de transcription.

Enfin, *in vivo* il est clair que les récepteurs nucléaires sont impliqués dans des complexes avec d'autres protéines, qui peuvent les aider à se fixer à l'ADN. Cela vient par exemple d'être montré pour le récepteur des hormones thyroïdiennes. On sait qu'une protéine appelée TRAP (Thyroid Hormone Receptor Auxiliary Protein) est capable d'augmenter la fixation de ce récepteur sur son site TRE (Burnside et al., 1990 ; Beebe et al., 1991 ; O'Donnel et al., 1991 ; Darling et al., 1991 ; Lazar et al., 1991). Le groupe de Michael Rosenfeld a décrit également des interactions entre le récepteur de l'acide rétinoïque de type alpha et des protéines nucléaires parmi lesquelles le nouveau récepteur des rétinoïdes appelé RXR $\beta$  (Yu et al., 1991). Ces interactions permettent de réguler la fixation spécifique du récepteur sur l'ADN (Glass et al., 1990 ; Näär et al., 1991).

# Domaine D

# Domaine E



**Pro 547 :** Point à partir duquel une délétion N-terminale du récepteur des glucocorticoïdes diminue la fixation du ligand

**Met 610 :** reconnue par un ligand photoactivable

**Cys 644 :** reconnue par un ligand photoactivable

**Cys 742 :** reconnue par un ligand photoactivable et muté dans un récepteur déficient.

**Gly 400 de ER :** mutation "naturelle" affectant la fixation de l'hormone de façon thermodépendante.

**Figure 15 :** Anatomie du domaine E et régions importantes pour la fixation du ligand. La numérotation est tirée du récepteur des glucocorticoïdes de Souris (d'après Danielsen, 1991).

## D) LE DOMAINE DE FIXATION DU LIGAND.

### 1) Anatomie du domaine E.

Le domaine E, C-terminal, est le plus grand domaine conservé des récepteurs nucléaires. Il contient environ 250 acides aminés qui sont, en règle générale, plutôt hydrophobes, ce qui semble cohérent avec leur rôle de fixation du ligand. A l'origine, on pensait que ce domaine était uniquement impliqué dans la fixation du ligand. On sait maintenant qu'il a de nombreuses autres fonctions importantes. On ne devrait pas parler de "sous-domaines" au sein du domaine E, mais plutôt de différentes "régions" ou "motifs" fortement impliqués dans un phénomène donné (Green et Chambon, 1991 ; Danielsen, 1991 ; Chin, 1991 ; Parker et Bakker, 1991).

Grossièrement et en oubliant volontairement la remarque faite plus haut sur les inter-relations entre les différentes parties du domaine E, celui-ci peut être découpé en 4 régions de l'extrémité N-terminale vers l'extrémité C-terminale (Forman et Samuels, 1990a ; 1990b ; figure 15) :

- La région L1 directement impliquée dans la fixation du ligand et qui diffère très fort d'un type de récepteur à l'autre.
- La région Ti responsable de la régulation de la transcription ligand-dépendante et de l'interaction avec certaines protéines, comme les protéines de choc thermique.
- La région de dimérisation que nous appellerons DM. Cette région est avec la région Ti la seule zone conservée entre presque tous les récepteurs nucléaires et, de ce point de vue, on peut souder ces deux zones en une seule région conservée du domaine E, appelé région Ti-DM.
- La région L2, C-terminale qui correspond plus ou moins au domaine F décrit chez les récepteurs des stéroïdes (Krust et al., 1986). Cette région est, avec L1, responsable de la fixation du ligand. Ces deux zones doivent se trouver associées par des repliements de la structure primaire des récepteurs pour former une "poche hydrophobe" putative, chargée de fixer le ligand. Ces deux régions L1 et L2 sont les plus conservées chez les récepteurs ayant la même spécificité de fixation d'un ligand donné (comme les deux types de récepteurs des hormones thyroïdiennes ; Chin, 1991)

Cette classification a été dressée pour les récepteurs des hormones thyroïdiennes, de l'acide rétinoïque et de la vitamine D. Il en est sans doute de même pour les autres types de récepteurs (Forman et Samuels, 1990a ; 1990b). En effet, pour les récepteurs des stéroïdes et pour les récepteurs orphelins, on retrouve la région Ti-DM conservée. Evidemment, il est difficile de parler de L1 et L2 pour les récepteurs orphelins... Il est toutefois important de noter que 3 gènes étroitement apparentés, (les gènes du groupe *knirps* impliqués dans la mise en place de la segmentation du corps chez la drosophile) ne possèdent pas de séquences correspondant au domaine Ti-DM et pas d'homologie avec les autres récepteurs dans le domaine E (Pankratz et

Jäckle, 1990 ; Segraves, 1991; voir les résultats et la discussion).

## 2) La fixation du ligand.

Le domaine E ayant été à l'origine décrit comme étant responsable de la fixation du ligand, il est apparu nécessaire de préciser ses limites. On s'est très vite aperçu que la région responsable de la fixation du ligand s'étendait sur la totalité de ce domaine, que ce soit pour le récepteur des oestrogènes (Green et Chambon, 1991) ou pour celui des glucocorticoïdes (Danielsen, 1991).

Il y a trois méthodes pour déterminer les zones importantes d'un tel domaine dans la fixation du ligand (figure 15 ; Danielsen 1991) :

- Les mutants de délétions : Ces mutants permettent de préciser en amont et en aval les limites du domaine. Des expériences de digestion par des protéases permettent également de tester la validité de ces expériences de mutagenèse (Simons et al., 1989 ; Katzenellenbogen et al., 1987). Si on délète le récepteur des glucocorticoïdes de rat de ses 547 acides aminés N-terminaux, on diminue énormément l'affinité du récepteur pour l'hormone (Rusconi et Yamamoto, 1987). L'étude d'autres mutants de ce type a permis de localiser la frontière N-terminale de la zone fixant l'hormone au niveau des acides aminés 522-535 du récepteur des glucocorticoïdes de rat (Danielsen, 1991). La frontière C-terminale est plus facile à déterminer : c'est l'extrémité C-terminale ! En effet, que ce soit pour le récepteur des glucocorticoïdes ou pour celui des oestrogènes, la délétion même d'un petit nombre (5) d'acides aminés (Rusconi et Yamamoto, 1987) détruit la capacité de fixer l'hormone (Hollenberg et al., 1985 ; Kumar et al., 1986 ; Danielsen, 1991 ; Green et Chambon, 1991).

- Les ligands photoactivables : Ces ligands sont capables de se fixer de façon covalente au récepteur. Ils permettent ainsi d'identifier les acides aminés directement impliqués dans la fixation du ligand ou situés très près de lui (environ 1,5 nm ; Danielsen, 1991; Simons et al., 1987 ; Carlstedt-Duke et al., 1988 ; Smith et al., 1988; Harlow et al., 1989; revue dans Samuels et al., 1988). De nombreux acides aminés ont ainsi été identifiés, comme la méthionine 610 et la cystéine 644 du récepteur des glucocorticoïdes. Ces acides aminés ne sont pas conservés d'un récepteur à l'autre, même au sein des récepteurs des stéroïdes.

- Les récepteurs déficients : Ponglikitmongkol et al., 1988, se sont aperçus, en étudiant l'organisation génomique du récepteur des oestrogènes, que dans les cellules MCF-7 (isolées d'une tumeur mammaire humaine) le récepteur des oestrogènes était muté sur la position 400 (une glycine est changée en valine). Cette mutation "naturelle" produit un récepteur dont l'affinité pour le ligand est thermosensible (Tora et al., 1989a). Ceci suggère que la glycine 400 est importante, bien qu'elle ne soit pas indispensable pour la fixation du ligand (Green et Chambon, 1991 voir aussi Danielsen et al., 1986 ; Byravan et al., 1990; Lubahn et al., 1989; Sultan et Lobaccaro, 1991; Sakurai et al., 1989 ; Usala et al., 1990).

Figure 16 : Les sous domaines des récepteurs des hormones thyroïdiennes, de l'acide rétinolique et de la vitamine D. Dans la zone de dimérisation, les répétitions en heptades ont été indiquées (d'après Forman et Samuels, 1990a).



De toutes ces études, il ressort clairement que les régions importantes pour la fixation du ligand sont disséminées tout au long du domaine E, les deux extrémités de ce domaine étant particulièrement importantes dans ce processus (voir Horowitz et al., 1989 et Lin et al., 1991 pour cette démonstration chez les récepteurs des hormones thyroïdiennes). Cependant, on ne dispose pas encore d'un modèle structural correct expliquant l'interaction hormone-récepteur. Bien que nous n'ayons pas abordé ce problème, un tel modèle permettrait d'expliquer et peut être de prévoir l'action d'analogues, agonistes ou antagonistes, des ligands des récepteurs nucléaires (voir les revues de Baulieu, 1989 ; Moudgil, 1988 ; Milgrom, 1991).

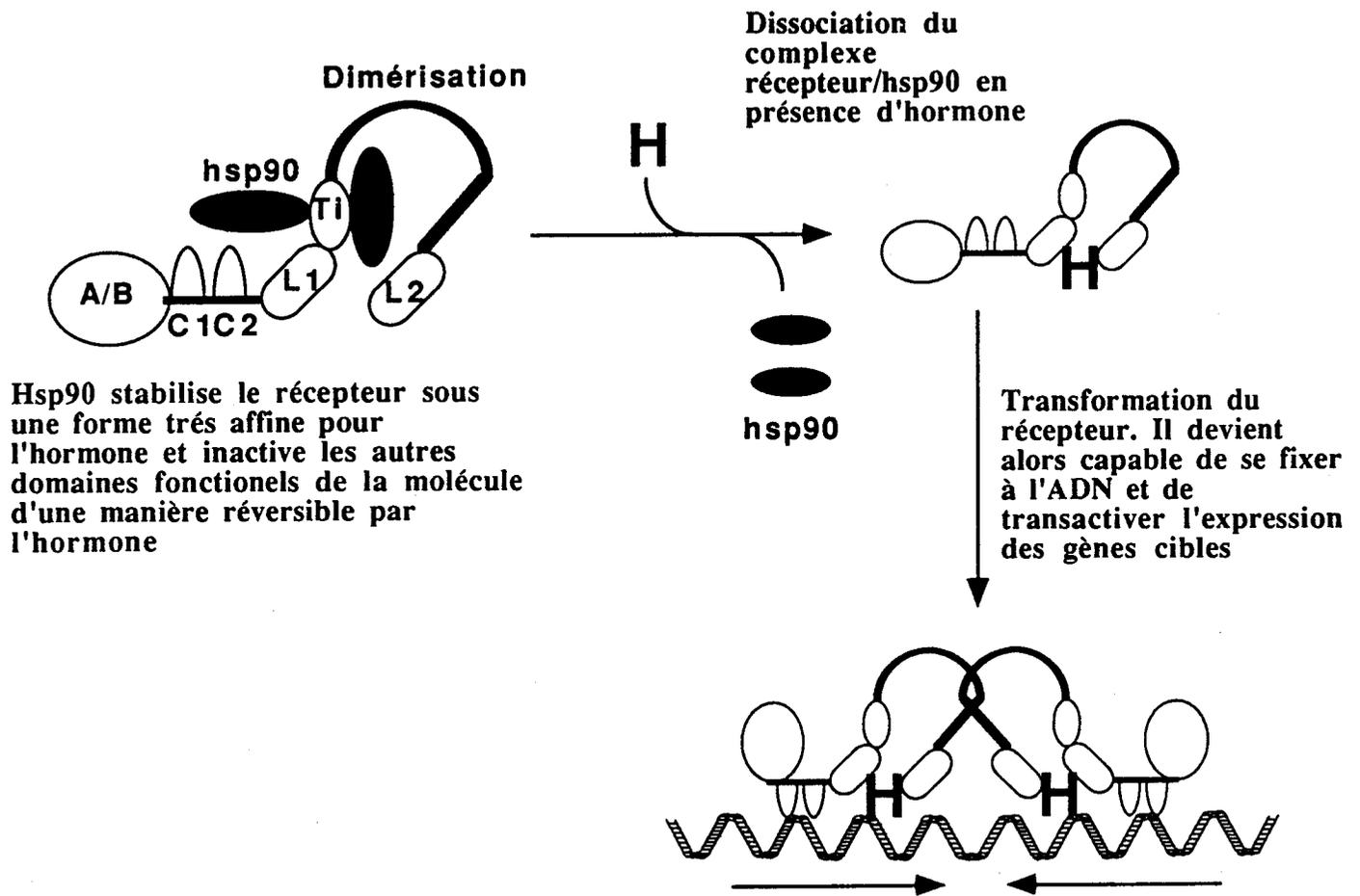
### 3) La dimérisation.

De nombreuses expériences indiquent que les récepteurs nucléaires se fixent à l'ADN sous forme de dimères (Kumar et Chambon, 1988; Tsai et al., 1988). Deux régions de dimérisation importantes ont été identifiées : le domaine C et le domaine E, le plus efficace et qui a été décrit comme étant ligand dépendant (Green et Chambon, 1988). Depuis ce type d'analyse a été étendu à d'autres récepteurs comme ceux des hormones thyroïdiennes (Forman et al., 1988 ; 1989).

Le groupe de Herbert Samuels à New York a remarqué que la zone centrale du domaine E était riche en acides aminés hydrophobes (leucine, isoleucine, valine, méthionine, phénylalanine) disposés régulièrement en 9 "heptades" (Forman et Samuels, 1990a ; 1990b). Repliées au sein d'une hélice alpha classique, de telles "heptades" vont former une surface hydrophobe d'un seul côté de l'hélice (figure 16). On sait que de telles surfaces sont très favorables à la formation de dimères entre molécules, comme c'est le cas pour les protéines à "leucine zipper" ou à motifs "hélice-boucle-hélice" (Busch et Sassone-Corsi, 1990). Ces motifs semblent exister également pour les récepteurs des stéroïdes (Fawell et al., 1990 ; Guiochon-Mantel et al., 1989; 1991). On peut donc imaginer que ces répétitions en "heptade" dans le domaine E des récepteurs nucléaires jouent un rôle dans le processus de dimérisation, ce qui est confirmé par des expériences fonctionnelles récentes menées sur le récepteur des hormones thyroïdiennes (Moore et Brent, 1991).

Il est assez bien accepté que les récepteurs des stéroïdes, comme le récepteur des glucocorticoïdes forment des dimères en solution, en absence d'ADN (Fawell et al., 1990 ; Guiochon-Mantel et al., 1989). Par contre, ce n'est pas le cas des récepteurs des hormones thyroïdiennes qui doivent être fixés à l'ADN pour interagir entre eux (Forman et Samuels, 1990a ; 1990b). Ces récepteurs se caractérisent également par une très grande plasticité puisqu'ils sont capables d'hétérodimériser avec de nombreux autres récepteurs (Forman et Samuels, 1990a ; 1990b).

**A** : Récepteurs des glucocorticoïdes, protéines hsp90 et modèle "à deux étapes"



**B** : Récepteur des hormones thyroïdiennes et protéine TRAP

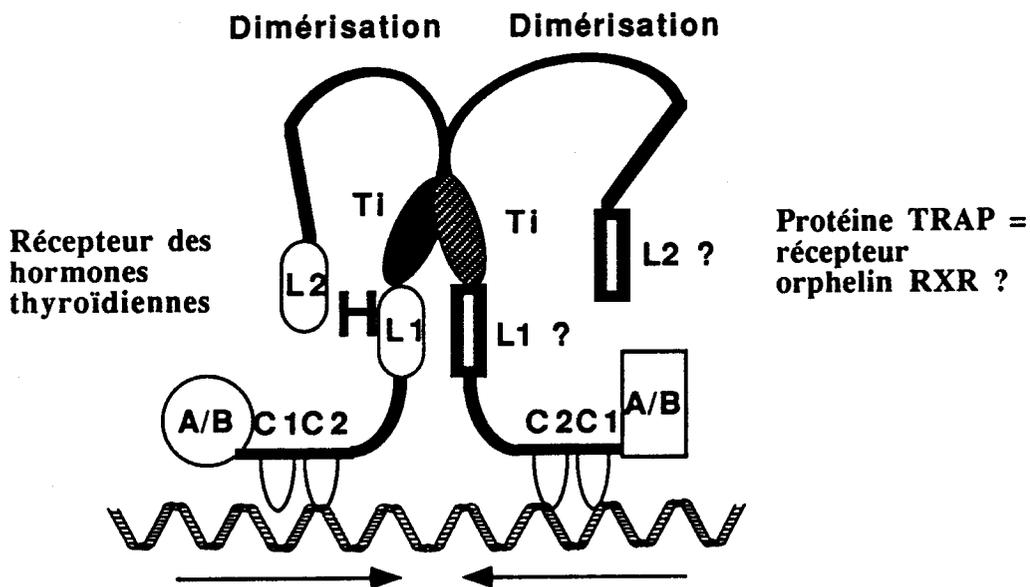


Figure 17 : Les rôles de la région Ti, illustrés pour (A), le récepteur des glucocorticoïdes et les protéines hsp90 ; et (B), le récepteur des hormones thyroïdiennes et la protéine TRAP/RXR.

#### 4) Les rôles de la région Ti.

##### a) Interactions avec les protéines hsp90.

On sait depuis longtemps que le récepteur des glucocorticoïdes est associé avec une protéine de 90 kDa qui ne fixe pas le ligand (Housley et al., 1985 ; Mendel et al., 1986). Cette protéine a été identifiée comme étant une protéine de choc thermique, appelée hsp90, qui est également associée aux autres récepteurs des stéroïdes lorsque ceux-ci ne sont pas liés à leur ligand (Sanchez et al., 1985 ; 1987 ; Catelli et al., 1985 ; Joab et al., 1984 ; Redeuilh et al., 1987). Le complexe protéine hsp90-récepteur est formé d'une molécule de récepteur de 94 kDa et de deux molécules de hsp90 (Denis et al., 1988; Rexin et al., 1988a ; 1988b). Le complexe récepteur-hsp90 n'est pas capable de se fixer à l'ADN ni de transactiver l'expression d'un gène cible. La fixation du ligand induit une transformation du récepteur en une forme 4S qui n'est plus liée aux hsp90 et qui devient capable de se fixer à l'ADN: c'est le modèle "à deux étapes" de Jensen et al., 1968. Il y a une relation entre la fixation de hsp90 sur le récepteur et la conversion de celui-ci sous une forme transactivatrice (figure 17A ; Hollenberg et al., 1987). Un argument en faveur de ce lien a été obtenu lorsque l'on s'est aperçu que la protéine hsp90 se fixe sur une séquence située dans le domaine de liaison à l'hormone et qui est conservée chez tous les récepteurs des stéroïdes (Pratt et al., 1988 ; Dalman et al., 1989 ; Howard et al., 1990). Certains auteurs ont supposé que cette séquence agissait comme un régulateur négatif de l'activité transactivatrice des récepteurs et ont donc proposé le nom de domaine Ti, pour "Transcriptional inhibition" (Danielsen et al., 1986 ; Godowski et al., 1987 ; Pratt et al., 1988 ; Forman et Samuels, 1990a ; 1990b). On pense qu'en se fixant sur le domaine Ti, les protéines hsp90 bloquent le récepteur sous une forme inactive. De plus, l'interaction avec les protéines hsp90 est importante pour contraindre le récepteur dans une forme très affine pour le ligand (figure 17A ; Nemoto et al., 1990 ; revue dans Danielsen, 1991). Notons que d'autres protéines de choc thermique comme la série des hsp70 sont également capables de s'associer aux récepteurs (Danielsen, 1991; Green et Chambon, 1991).

Un test direct du rôle *in vivo* de l'interaction récepteur-hsp90 a été développé par le groupe de Keith Yamamoto, chez la levure (Picard et al., 1990). Ces auteurs ont utilisé des souches de levure possédant un taux régulable de protéines hsp90. Ceci leur a permis de montrer que la capacité transactivatrice du récepteur en absence de hsp90 est grandement diminuée. Ces données suggèrent que l'interaction avec hsp90 a deux types d'effet sur la réponse hormonale : (i) elle stabilise le récepteur sous une forme très affine pour l'hormone et (ii) elle inactive d'une manière réversible par l'hormone les autres domaines fonctionnels de la molécule (Picard et al., 1990).

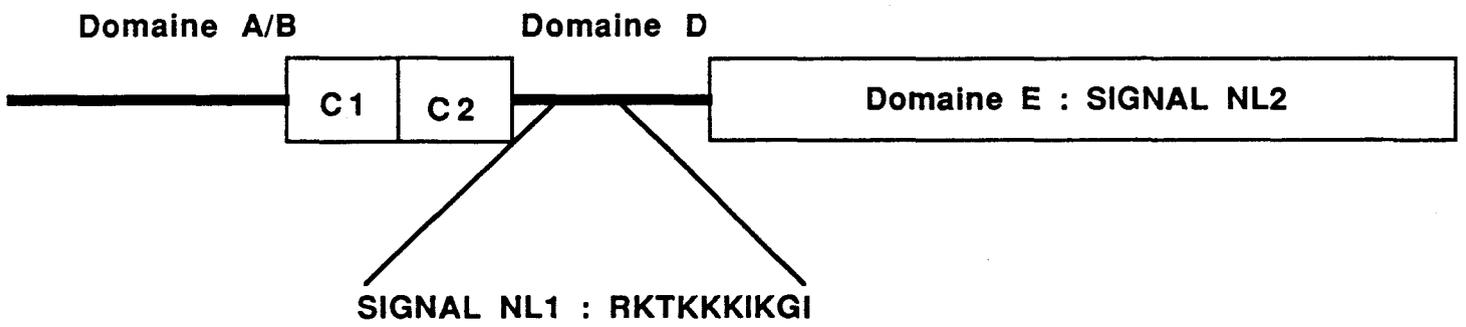
Jusqu'à tout récemment, l'interaction des récepteurs avec les protéines hsp90 semblait uniquement restreinte aux récepteurs des stéroïdes. Certaines expériences ont même fourni des évidences expérimentales allant dans ce sens puisqu'il a été montré que contrairement au récepteur des glucocorticoïdes, le récepteur des hormones thyroïdiennes est directement traduit sous une

forme non liée aux hsp90, capable de se fixer à l'ADN, (Dalman et al., 1990). Cependant le domaine Ti existe bien chez les récepteurs des hormones thyroïdiennes comme chez tous les récepteurs nucléaires, sauf chez les produits des gènes *knirps* de drosophile (Danielsen, 1991 ; Segraves, 1991). Quel est donc le rôle de ce domaine Ti ? Les protéines hsp90 se fixent-elles tout de même au récepteur des hormones thyroïdiennes ? Un travail récent de Martin Privalsky montre que les protéines hsp90 se fixent au produit du gène v-erbA, l'homologue viral du récepteur des hormones thyroïdiennes de type  $\alpha$  chez le poulet (Privalsky, 1991). Ce travail montre que, *in vivo*, la protéine codée par v-erbA forme majoritairement un complexe de haut poids moléculaire (150 à 400 kDa). Dans ce complexe on peut mettre en évidence, à l'aide d'agents "pontants", une association entre le produit de v-erbA et une hsp90 de poulet. Il semble que seules les protéines cytoplasmiques codées par v-erbA (soit 30 à 40% du stock cellulaire de ces protéines) soient liées aux hsp90. Comment peut-on réconcilier ces expériences avec celles de Dalman et al., 1990 ? Il faut rappeler que jusqu'à présent ce phénomène n'a été observé que pour v-erbA. Il existe dans le domaine Ti de v-erbA une mutation par rapport aux récepteurs des hormones thyroïdiennes codés par les gènes cellulaires. On peut imaginer que cette mutation permette à v-erbA de fixer hsp90 alors que c-erbA en serait incapable, ou qu'elle stabilise un complexe autrement trop fugace pour être observable. Alternativement on peut aussi suggérer que les produits des gènes c-erbA, presque exclusivement nucléaires, ne peuvent pas interagir avec hsp90 comme la fraction cytoplasmique de v-erbA. Cependant certaines données suggèrent que les hsp90 se rencontrent aussi dans le noyau (Gasc et al., 1990). La question de l'interaction des hsp90 avec les récepteurs des hormones thyroïdiennes, voire ceux de l'acide rétinoïque, reste donc ouverte...

#### b) Domaine Ti et transactivation.

Le domaine Ti, conservé chez tous les récepteurs, devrait logiquement avoir un rôle commun chez toutes ces molécules. Ce rôle a été étudié par une autre approche chez les récepteurs des hormones thyroïdiennes. Un certain nombre de mutations ponctuelles de cette région du récepteur  $\beta$  des hormones thyroïdiennes ont été réalisées. Dans la plupart des cas, ces mutants sont transcriptionnellement inactifs en présence d'hormone. De plus ces mutants sont capables d'inhiber l'action d'un récepteur normal. Il semble donc bien que ces mutations dans cette région définissent un domaine fonctionnel propre, impliqué dans la régulation de la transcription (O'Donnell and Koenig, 1990). Il faut noter que certaines mutations ont un effet sur la capacité de fixer l'hormone des récepteurs mutants (O'Donnell and Koenig, 1990). Ceci montre bien que le domaine de fixation de l'hormone n'est pas uniquement restreint aux extrémités C-terminales du domaine E.

Les auteurs de ce travail supposent qu'un facteur nucléaire serait capable de se fixer sur cette région et de réguler la transcription des gènes cibles. Une telle protéine, appelée TRAP, ("Thyroid hormone Receptor Auxiliary Protein") a été caractérisée (Burnside et al., 1990 ; Lazar et Berrodin, 1990 ; Beebe et al., 1991 ; O'Donnell et al., 1991; Darling et al., 1991). Il s'agit d'une



**Figure 18 : Les signaux de translocation nucléaires du récepteur des glucocorticoïdes.**

protéine capable d'augmenter la capacité de fixation des récepteurs des hormones thyroïdiennes à l'ADN, de fixer elle-même l'ADN, via un élément structurellement proche d'un TRE et d'interagir avec les récepteurs des hormones thyroïdiennes (figure 17B). Deux régions du récepteur sont importantes pour l'interaction avec TRAP : le deuxième doigt de zinc C2 et le domaine Ti (O'Donnell et al., 1991). Ces données suggèrent que la protéine TRAP a un rôle dans la régulation de la transcription et pas seulement dans la modulation de la fixation à l'ADN des récepteurs des hormones thyroïdiennes. Tout récemment, pour les récepteurs de l'acide rétinoïque, une protéine dont l'activité ressemble très fortement à TRAP a été identifiée : il s'agit du récepteur des rétinoïdes RXR $\beta$  (figure 17B ; Yu et al., 1991 ; Leid et al., 1992 ; Zhang et al., 1992 ; Kliewer et al., 1992).

Ainsi, le domaine Ti semble bien avoir un double rôle (au moins) : il constitue la zone de reconnaissance des récepteurs aux stéroïdes libres (cytoplasmiques ?) et des protéines hsp90, et il médie l'interaction entre certains récepteurs fixés à l'ADN et des protéines accessoires qui permettent une modulation de la transactivation.

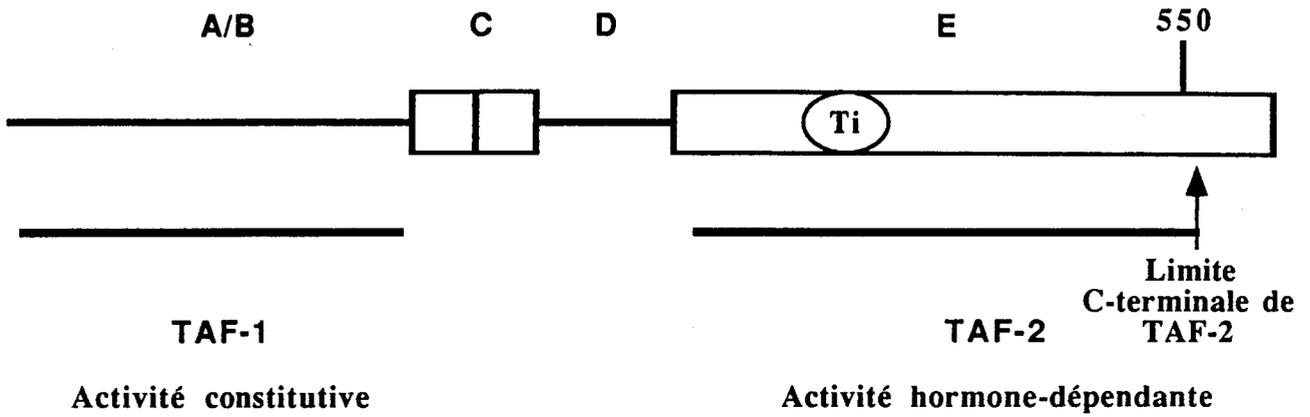
## E) LES SIGNAUX DE TRANSLOCATION NUCLEAIRES.

Nous avons déjà discuté de la localisation subcellulaire des récepteurs qui est assez variable d'un récepteur à l'autre. Il est tout de même clair qu'en présence de ligand les récepteurs nucléaires sont bien situés dans le noyau. Etant donné leur taille, il doit y avoir un mécanisme de transport actif qui les fait passer dans le noyau. Il doit donc exister des signaux de reconnaissance permettant à la protéine d'être prise en charge par ces systèmes de transport.

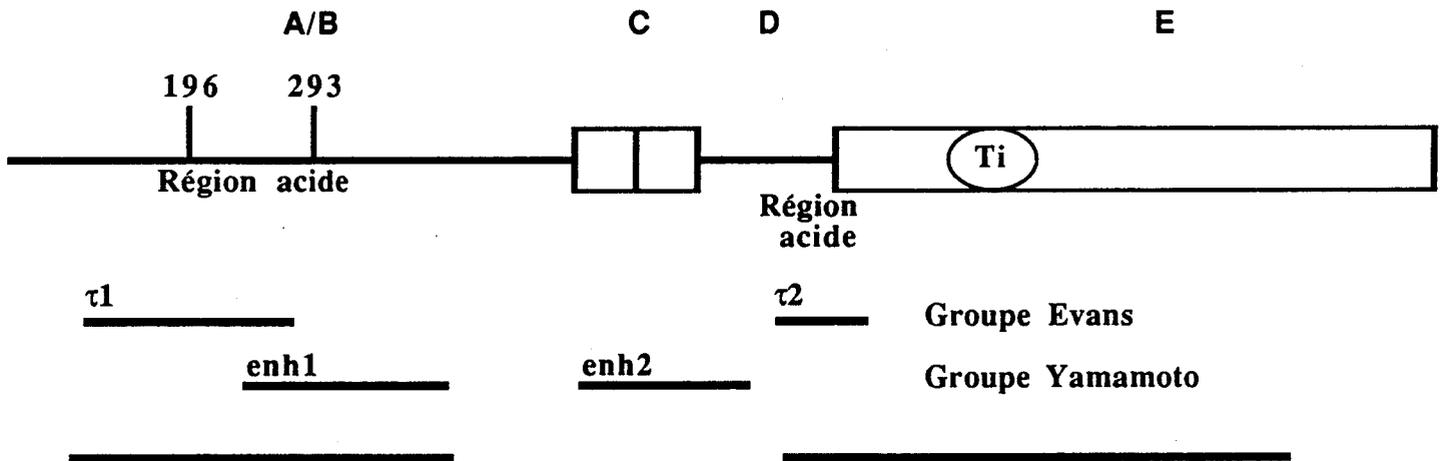
De tels signaux ont été identifiés par Picard et Yamamoto, 1987, dans le récepteur des glucocorticoïdes (figure 18). En construisant des mutants de délétion du récepteur et des protéines de fusion entre le récepteur (ou des parties de ce récepteur) et la  $\beta$ -galactosidase, ces auteurs ont pu mettre en évidence deux signaux de translocation nucléaire NL1 et NL2. Le signal NL1 est situé à proximité du domaine de fixation à l'ADN mais n'est pas confondu avec lui. Il est plutôt en C-terminal du deuxième doigt de zinc probablement lié à une courte séquence très riche en acides aminés basiques (RKT $\overline{\text{K}}$ KIKGI ; Picard et Yamamoto, 1987) qui ressemble assez au signal de translocation nucléaire de l'antigène T de SV40 (PPKKR $\overline{\text{K}}$ V ; Kalderon et al., 1984). NL2 lui, est situé dans le domaine E de fixation du ligand, mais on n'a pas pu préciser davantage sa localisation (figure 18). Il est possible que tout ce domaine soit important pour la translocation nucléaire du récepteur. Le signal NL2 semble être le plus efficace car, en présence d'hormone, il permet seul une translocation du récepteur dans le noyau aussi rapide qu'un produit contenant NL1 et NL2 (Ucker et Yamamoto, 1984).

Pour d'autres récepteurs des stéroïdes, il semble que l'on retrouve un signal situé au même endroit que le signal NL1 du récepteur des glucocorticoïdes mais d'autres domaines sont importants, par exemple pour le récepteur de la progestérone (Guiochon-Mantel et al., 1989). Il est possible que certains des domaines de localisation nucléaire soient en fait des zones d'interactions

# Récepteur des oestrogènes



# Récepteur des glucocorticoïdes



TAF-1 en deux parties : une acide, type VP16, l'autre non ?

TAF-2 en deux parties également ?

Figure 19 : Les domaines transactivateurs des récepteurs des oestrogènes et des glucocorticoïdes (d'après Green et Chambon, 1991 ; Danielsen, 1991, et Tasset et al., 1990).

avec des protéines nucléaires qui captureraient dans le noyau le récepteur entré par simple diffusion. Les deux mécanismes possibles d'entrée dans le noyau joueraient donc : transport spécifique via un signal de translocation nucléaire et diffusion avec rétention dans le noyau, après l'entrée (Guiochon-Mantel et al., 1989).

## **F) LA REGULATION DE LA TRANSCRIPTION.**

Poursuivons notre analyse des grandes fonctions des récepteurs nucléaires. Puisque nous venons d'étudier la fixation à l'ADN et celle du ligand, ainsi que les fonctions qui y sont associées, il nous faut à présent nous pencher sur la manière dont les récepteurs nucléaires régulent la transcription de leurs gènes cibles. Dans cet aspect, plus encore que dans tous les autres, règne une grande diversité selon les types de récepteurs et il n'est pas question de se livrer ici à une étude détaillée des zones transactivatrices cartographiées sur les récepteurs.

### **1) Localisation des domaines transactivateurs.**

Pour étudier de plus près cet aspect nous prendrons un exemple précis : le récepteur des oestrogènes (Green et Chambon, 1991). Deux domaines d'activation transcriptionnelle, TAF-1 et TAF-2, ont été détectés par le groupe de Pierre Chambon, chez le récepteur des oestrogènes (figure 19 ; Kumar et al., 1987). Le domaine TAF-1 est situé dans le domaine A/B et a une activité constitutive, tandis que le domaine TAF-2 est situé dans le domaine de fixation du ligand E et agit d'une manière hormono-dépendante (Lees et al., 1989). C'est le domaine TAF-2 qui supporte l'activité transcriptionnelle majeure du récepteur et qui est responsable de l'activation ligand-dépendante des gènes cibles (Tora et al., 1989b). Ces sites n'ont pas été précisément cartographiés sur le récepteur et il n'y a pas dans la séquence de celui-ci de régions réminiscentes de domaines transactivateurs acides ou riches en proline, comme ceux du facteur VP16 du virus de l'herpès (Mitchell et Tjian, 1989; Sadowski et al., 1988). On sait que le domaine est codé par les différents exons correspondants au domaine E et il est probable qu'il ne soit pas confiné à une région discrète mais soit formé par la mise en contact, dans la structure tridimensionnelle, de plusieurs régions différentes de E (Webster et al., 1989; Lees et al., 1989).

Il est clair que les activités des deux domaines TAF-1 et TAF-2 dépendent du contexte du promoteur cible et du type cellulaire. Par exemple, il a été démontré que le fait de retirer TAF-1 d'un récepteur n'a pas d'effet lorsque le vecteur indicateur contient un ERE palindromique synthétique. Par contre, cela entraîne la perte de 80% de l'activité basale lorsque le vecteur reporter contient le promoteur du gène pS2 où se situe un ERE naturel, non palindromique (Berry et al., 1989). Il est donc clair que l'activité des deux domaines TAF-1 et TAF-2 doit être médiée par un certain nombre de facteurs capables d'interagir avec ces domaines et qui varient selon les cellules et selon le contexte du ERE étudié (Bocquel et al., 1989; Tora et al., 1989b; revue dans Green et

**TABLEAU 3**

Interférence transcriptionnelle des domaines transactivateurs des récepteurs des oestrogènes (ER) et des glucocorticoïdes (GR) entre eux et avec le domaine transactivateur de type acide de la protéine VP16

COMPETITEUR		ACTIVATEUR				
		ER		GR		VP16
		TAF-1	TAF-2	TAF-1	TAF-2	
ER	TAF-1	++	+	+++	+++	+++
	TAF-2	++	+	++	+++	++
GR	TAF-1	+	-	++	-	-
	TAF-2	+	+	+	+	-
VP16		-	-	+++	+/-	+++

Principe : Dans une cellule (Hela par ex), on mesure le taux de transactivation d'un construit indicateur induit par un activateur, comme le TAF-1 du récepteur des oestrogènes seul. Dans une expérience menée en parallèle, on introduit l'indicateur, TAF-1 de ER et un compétiteur, par exemple VP16. On compare les deux résultats et on en conclut que TAF-1 de ER et VP16 n'entrent pas en compétition. Ceci signifie que, probablement les cibles cellulaires de ces deux transactivateurs sont différentes. Dans la mesure où TAF-1 de GR entre en compétition avec TAF-1 de ER et avec VP16, on peut admettre que TAF-1 de GR contient deux domaines transactivateurs (peut être chevauchants) : l'un de type VP16 et l'autre de type TAF-1 de ER. Donc, TAF-1 de GR recrute à la fois les cibles de VP16 et celles de TAF-1 de ER. Evidemment, ces résultats vont varier d'un type cellulaire à l'autre ! (D'après Tasset et al., 1990).

Chambon, 1991).

Les domaines TAF-1 et TAF-2 sont capables d'agir en synergie, cette synergie pouvant aussi avoir lieu avec des domaines transactivateurs plus "classiques", comme le domaine acide de la protéine VP16. Au cours d'expériences d'interférences transcriptionnelles (ou "squenching" ; voir tableau 3; Ptashne, 1988) montrent que VP16 et le récepteur des oestrogènes interagissent avec des facteurs différents et spécifiques. Les TAFs sont donc bien différents structuralement et fonctionnellement des domaines acides (Meyer et al., 1989; Tasset et al., 1990). Par contre, si l'on considère les deux domaines TAF-1 et TAF-2 séparément, on peut montrer qu'ils entrent en compétition l'un avec l'autre pour la fixation d'un facteur indéterminé présent en quantité limitée dans la cellule.

La situation semble différente en ce qui concerne le récepteur des glucocorticoïdes, parce que dans ce cas, les domaines transactivateurs observés contiennent des régions riches en acides aminés acides (figure 19 ; Godowski et al., 1988 ; Hollenberg et Evans, 1988; Picard et al., 1988; Webster et al., 1988 ; Bocquel et al., 1989; revue dans Danielsen, 1991). De fait, les domaines transactivateurs de ce récepteur ont été beaucoup plus difficiles à cerner et les résultats varient beaucoup d'un laboratoire à l'autre...ce qui est un bon indicateur d'un certain flottement général ! Il semble à présent (et en simplifiant sans doute exagérément) que les TAFs du récepteur des glucocorticoïdes sont chacun constitués de deux régions (peut-être chevauchantes bien sûr...) ; l'une acide entrerait en compétition avec VP16 et l'autre non (Tasset et al., 1990). L'identification des différentes protéines impliquées dans les différentes voies de transactivation permettra peut-être d'y voir plus clair...

## 2) Transactivation et hormones.

Après bien des débats, il semble clair que les récepteurs nucléaires dépendent de la liaison de leurs ligands respectifs pour activer l'expression des gènes cibles (Becker et al., 1986 ; Willman et Beato, 1986 ; Bailly et al., 1986 ; revue dans Green et Chambon, 1986). Un problème demeure cependant à propos de l'activité basale du récepteur sans hormone. Le groupe de Pierre Chambon a montré qu'en absence d'hormone, le récepteur des oestrogènes ne se fixe pas spécifiquement à l'ADN et n'a pas d'activité transcriptionnelle basale. L'hormone induirait donc, d'une manière qui reste à préciser, deux phénomènes : la capacité de se fixer à l'ADN et l'activation de la transcription (Webster et al., 1988). Curieusement, en travaillant sur le récepteur des oestrogènes de souris, le groupe de Malcolm Parker ne confirme pas ces résultats. Ils observent en effet que le récepteur sans hormone est capable de se fixer spécifiquement à l'ADN et d'induire une activation de la transcription, faible mais détectable (Lees et al., 1989). Des données plus récentes de l'équipe de Magnus Pfahl (Tzukerman et al., 1990) suggèrent que ces différences ne sont pas dues au fait que les deux équipes utilisent des récepteurs d'espèces différentes, mais au fait que le récepteur humain

utilisé par le groupe de Pierre Chambon, isolé des cellules MCF-7, contient une mutation dans le domaine E (la glycine 400 est changée en valine ; Ponglikitmongkol et al., 1988 ; Tora et al., 1989a). Il semble que cette mutation induit une chute de l'activité basale du récepteur et ne permet plus la fixation spécifique à l'ADN en absence d'hormone, sans doute en empêchant la formation de dimères (Tzukerman et al., 1990). Le groupe de Pierre Chambon suggère que ce sont plutôt les procédures d'élimination des oestrogènes du sérum utilisé pour cultiver les cellules qui expliquent ces différences (Berry et al., 1990b). Si les oestrogènes ne sont pas complètement éliminés du milieu, on obtient une légère activité transactivatrice du récepteur que l'on peut confondre avec une activité basale. Le problème reste posé.

Le modèle général à retenir est que l'hormone a un double rôle : elle permet la transformation du récepteur en une forme nucléaire capable de se fixer spécifiquement à l'ADN et elle active les domaines transactivateurs, ce qui permet l'activation transcriptionnelle. Pour étudier ce type de problème, on dispose également d'analogues des hormones stéroïdes qui jouent, selon les cas, des rôles d'agonistes ou d'antagonistes (Baulieu, 1989; Green et Chambon, 1991 ; Parker et Bakker, 1991). L'étude de l'effet de ces antagonistes, que nous ne décrivons pas en détail, conforte le modèle énoncé plus haut (Fawell et al., 1990; Berry et al., 1990b).

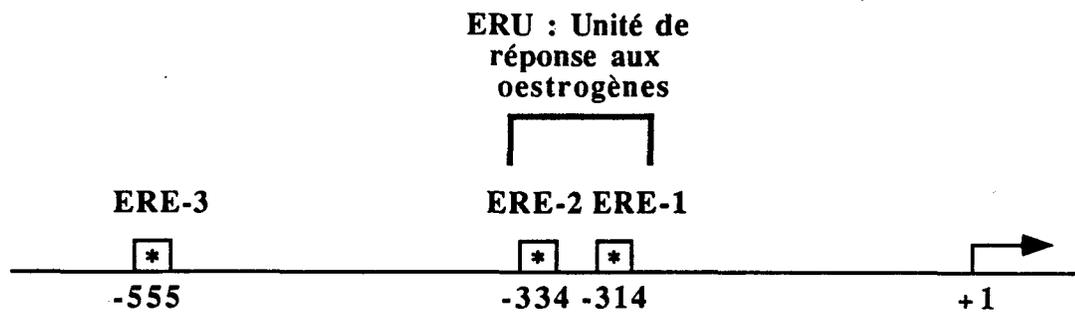
### 3) La synergie transcriptionnelle entre récepteurs.

Dans ce domaine, plus que dans tout autre, règne la loi du cas par cas. Il semble ne pas y avoir de règle générale ou alors cette règle n'est pas encore comprise (Beato, 1989; Muller et al., 1991; Martinez et Wahli, 1991). La question se pose à deux niveaux : est-ce qu'il y a fixation coopérative des récepteurs sur plusieurs sites adjacents, comme c'est le cas pour de nombreux facteurs de transcription, par exemple chez les bactéries (Ptashne, 1988) et deuxièmement, y a-t-il synergie transcriptionnelle entre ces différents sites ?

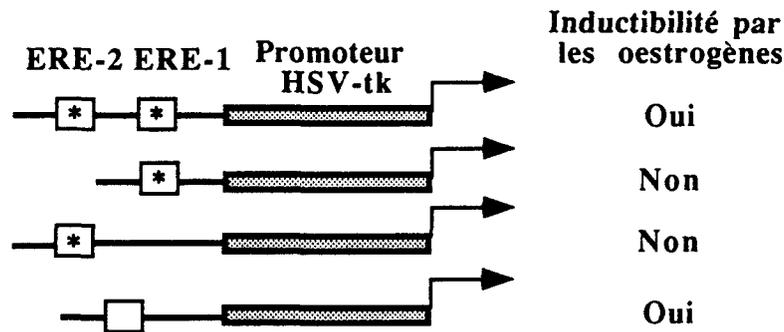
#### a) La fixation à l'ADN.

Il existe des exemples de gènes régulés de manière synergique où il y a fixation coopérative des deux récepteurs : la fixation de l'un facilite la fixation de l'autre (figure 20). C'est le cas par exemple du gène de la vitellogénine B1 qui contient deux ERE et où on a pu démontrer une fixation coopérative des deux récepteurs des oestrogènes sur chacun des ERE (Martinez et Wahli, 1989). Il faut cependant préciser que d'autres expériences de ce type, menées sur des constructions artificielles contenant plusieurs ERE parfaits ou non, espacés différemment, ne confirment pas ces résultats : il n'y a pas de fixation coopérative dans ce cas (Ponglikitmongkol et al., 1990). Il est probable que ce sont les systèmes expérimentaux différents et notamment les promoteurs utilisés qui expliquent ces divergences (Green et Chambon, 1991 ; Martinez et Wahli, 1991). Le message à retenir est donc : coopération au niveau de la fixation à l'ADN dans certains cas, pas dans d'autres.

Promoteur du gène de la vitellogénine B1 de Xenope (d'après Martinez et al., 1987).



- ERE imparfait
- ERE parfaitement palindromique



Etude des synergies entre ERE (d'après Ponglikitmongkol et al., 1990).

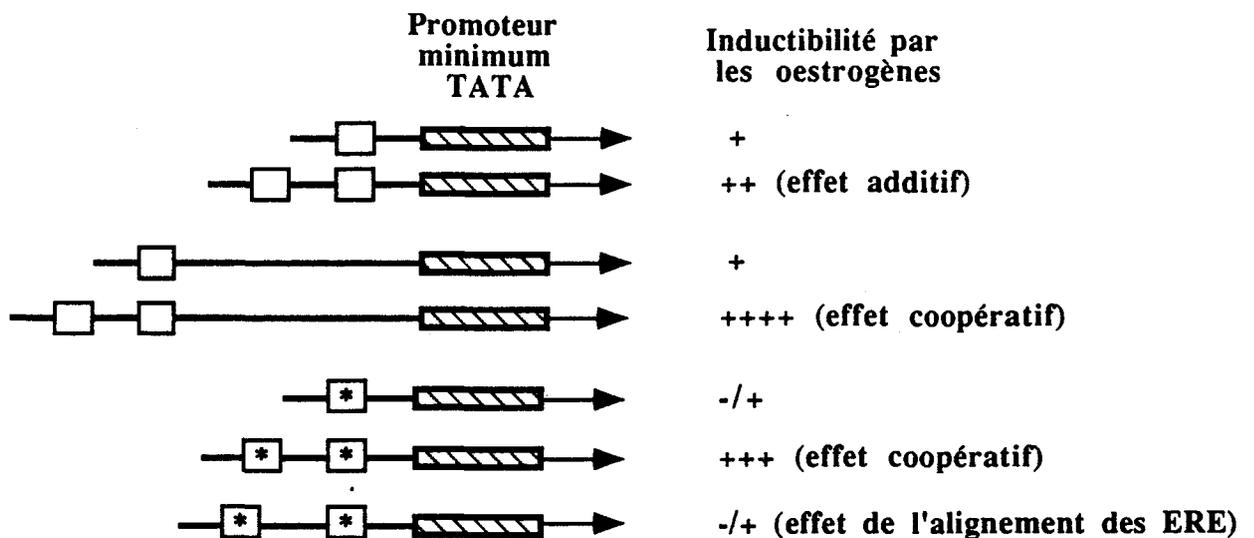


Figure 20 : La synergie transcriptionnelle entre HRE : exemple du récepteur des oestrogènes et des ERE (d'après Martinez et Wahli, 1991 ; Green et Chambon, 1991).

b) Synergie transcriptionnelle.

Au niveau de l'activité transcriptionnelle, la règle semble être en faveur d'une coopération entre les HRE (figure 20). C'est le cas pour le gène de la vitellogénine B1 de Xénope qui contient deux palindromes imparfaits représentant des ERE. Ces deux sites ERE-1 et ERE-2 sont très proches l'un de l'autre et forment ensemble une unité de réponse aux oestrogènes (ERU) aussi efficace qu'un ERE palindromique parfait seul. Le groupe de Walter Wahli a clairement montré qu'aucun de ces deux ERE ne fonctionne seul (Martinez et al., 1987). Par contre, les deux ensemble confèrent une bonne inductibilité aux oestrogènes, ceci étant sans doute lié à la fixation coopérative des récepteur à ces sites, ce qui a été démontrée par la suite (Martinez et Wahli, 1989). La coopération entre ces deux sites a été étudiée plus en détail et on a pu démontrer que les deux sites doivent être très proches l'un de l'autre pour que l'induction normale ait lieu.

Un test systématique des différentes possibilités d'interaction entre ERE a été réalisé par l'équipe de Pierre Chambon en utilisant différents ERE synthétiques, liés à un promoteur minimal contenant la boîte TATA du promoteur tardif de l'adénovirus 2 (figure 20 ; Ponglikitmongkol et al., 1990). Le comportement de ERE parfaitement palindromiques a été comparé au comportement de ERE imparfaits. Des ERE parfaits ne coopèrent que lorsqu'ils sont éloignés l'un et l'autre de la boîte TATA. S'ils sont proches, l'effet est seulement additif. Cette coopération de deux ERE parfaits éloignés de 200pb environ de la boîte TATA dépend de l'alignement de ces deux éléments : en effet, si les deux ERE sont séparés par des demi-tours d'hélice d'ADN (c'est-à-dire si chaque ERE est sur un côté de l'hélice) il n'y a pas de coopération. Pour les ERE imparfaits, il y a coopération même lorsque ces éléments sont proches de la boîte TATA, mais là encore l'effet est dépendant de l'alignement des ERE.

On peut expliquer ces résultats en admettant que pour qu'il y ait un effet transcriptionnel fort, il faut que le récepteur reste fixé à son site assez longtemps pour se "connecter" à la machinerie de base de la transcription et activer celle-ci. Si un récepteur reste fixé très longtemps à son site, le fait d'amener un deuxième site n'apportera rien. Par contre si le récepteur est fixé faiblement à son site, comme dans le cas des ERE imparfaits, on peut imaginer que l'apport d'un deuxième site améliore la cinétique d'interaction avec la machinerie transcriptionnelle (Ponglikitmongkol et al., 1990).

Pour généraliser quelque peu ce phénomène de coopération entre HREs il faut noter que des HRE correspondant à des hormones différentes peuvent entrer en synergie : c'est le cas du gène de la vitellogénine II de poulet qui contient un GRE et un ERE. Il a été montré que la transcription de ce gène est régulée de manière coopérative par la combinaison des deux hormones, les oestrogènes et les glucocorticoïdes (Ankenbauer et al., 1988) mais aussi les oestrogènes et la progestérone (Cato et al., 1988).

## G) LES INTERACTIONS AVEC LES FACTEURS DE TRANSCRIPTION.

Ce domaine est, par excellence, celui où l'on doit s'attendre à des avancées fulgurantes dans les années (voire les mois !) à venir. L'idée que les récepteurs nucléaires interagissent avec des facteurs de transcription semble pourtant évidente en soi. Encore faut-il démontrer expérimentalement cette évidence et mettre en place des modèles tant expérimentaux que conceptuels pour rendre compte de ces interactions. Grosso modo, on peut dire que les récepteurs nucléaires vont interférer avec les facteurs nucléaires soit au niveau de la fixation à l'ADN (compétition pour le même site ou interactions via la fixation à des sites différents), soit au niveau d'interactions protéines-protéines (fixation directe ou "titration de facteurs communs"). Examinons successivement des exemples de ces quatre mécanismes possibles (revues dans Muller et al, 1991; Schüle et Evans, 1991; Miner et Yamamoto, 1991).

### 1) Fixation des facteurs sur des sites d'ADN différents : le modèle du LTR du MMTV.

A ce niveau les exemples abondent ! Mais les différents types d'interactions détectés ne sont pas toujours finement caractérisés (voir les exemples de Danesch et al., 1987 ; Schüle et al., 1988; Akerblom et Mellon, 1991). L'exemple classique et incontournable de ce type d'interaction est le promoteur du virus MMTV ("Mouse Mammary Tumor Virus" ; voir la revue de Beato et al., 1991). Avant de discuter plus en détail les interactions moléculaires qui gouvernent l'expression de ce promoteur, il convient de replacer ce problème dans son contexte physiologique (figure 21). Ce rétrovirus de souris a un tropisme très particulier : il est la cause de carcinomes uniquement chez les souris femelles. Il n'est pas oncogène par lui-même. Mais en s'insérant dans le génome à des endroits précis, il dérégule l'expression de certains oncogènes (revue dans Ponta et al., 1985). Une fois intégré dans le génome de l'hôte, le virus s'exprime à son niveau maximum dans la glande mammaire pendant la gestation et la lactation, deux moments où les taux d'hormones stéroïdes, comme les glucocorticoïdes ou la progestérone, sont très élevés. Il a été clairement démontré que la transcription du génome viral est sous la dépendance de ces hormones via les quatre sites HRE situés dans son promoteur. D'autres facteurs induisent aussi l'expression de ce virus, par exemple la prolactine, mais seulement en présence d'hormones stéroïdes. Donc, en résumé, la régulation de la transcription de ce virus va se situer à trois niveaux, étroitement reliés : (i) la spécificité tissulaire (le virus est exprimé presque exclusivement dans la glande mammaire), (ii) la présence d'hormones stéroïdes et (iii) la présence d'autres facteurs en association avec des hormones. Le niveau de transcription observé aura intégré nécessairement ces trois niveaux.

L'interaction récepteur-promoteur est bien comprise au niveau moléculaire y compris *in vivo*, en tenant compte des repliements de l'ADN autour des octamères d'histones (Cordingley et al., 1987 ; Richard-Foy et Hager, 1987; revues de Beato et al., 1991 et Hager et Archer, 1991 ).

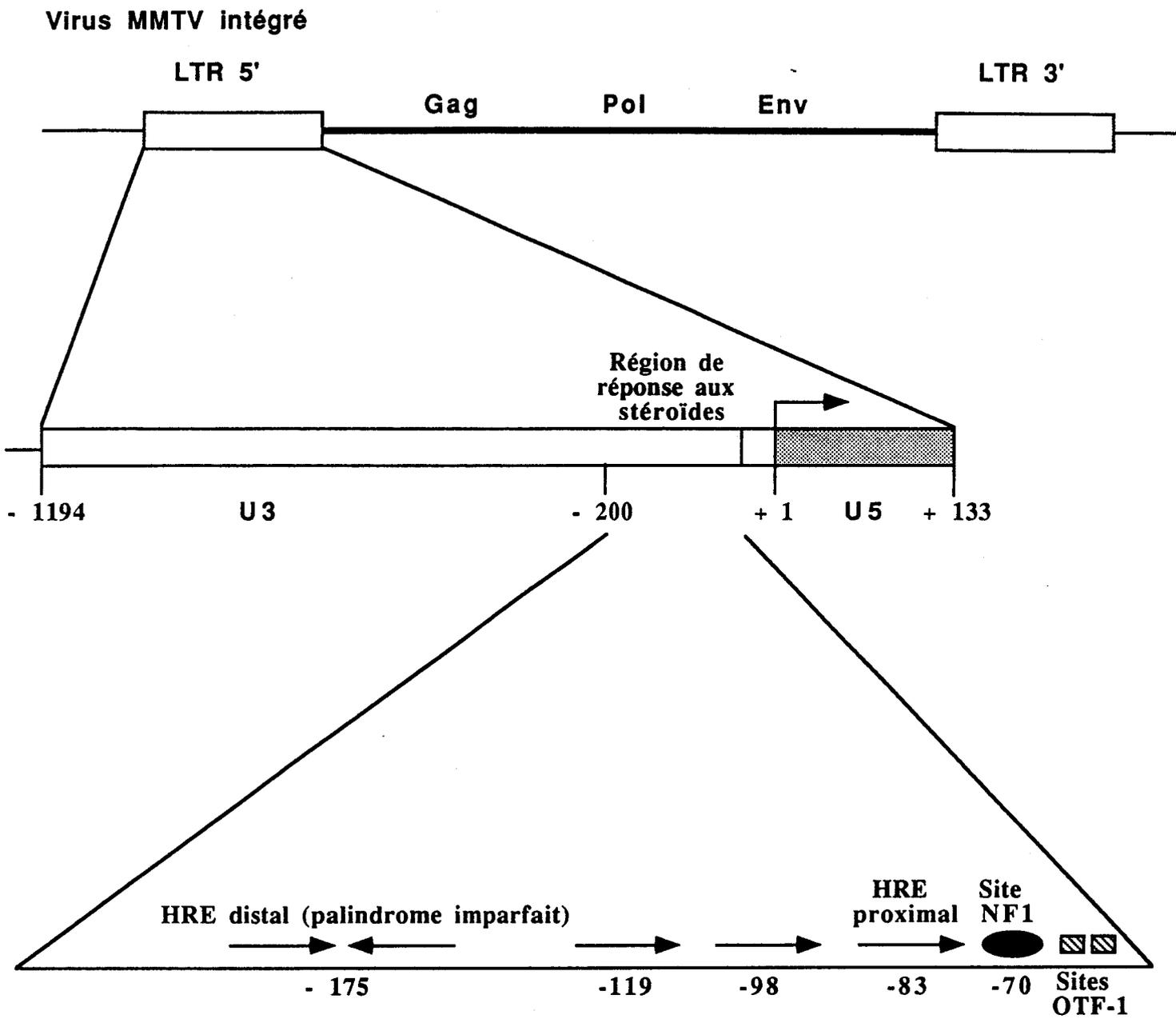


Figure 21 : Schéma du LTR du virus MMTV indiquant la position des sites HRE, NF-1 et OTF-1 (ou Oct-1) (d'après Beato et al., 1991 ; et Pina et al., 1990).

D'autre part, on sait de longue date que le facteur de transcription NF-1 joue aussi un rôle important dans la régulation de ce promoteur et notamment qu'il contrôle l'inductibilité par les stéroïdes (figure 21 ; Buetti et Kühnel, 1986 ; Miksicek et al., 1987 ; Srähle et al., 1988). Si, au cours d'expériences réalisées *in vitro* on mute le site NF-1, on inhibe très sévèrement la transcription induite par les glucocorticoïdes (Miksicek et al., 1987 ; Buetti et al., 1989). De plus, on sait que, *in vivo*, l'induction de la transcription par les glucocorticoïdes s'accompagne de la fixation du facteur NF-1 sur son site spécifique (Cordingley et al., 1987). Par des expériences de complémentation, en utilisant des cellules dépourvues de NF-1, le groupe de Miguel Beato a pu montrer que ce facteur était effectivement requis pour l'induction (Brüggemeier et al., 1990). Ce type d'effet a de même été mis en évidence dans des systèmes de transcription *in vitro* (Kalff et al., 1990). Il a également été démontré que le traitement par les glucocorticoïdes ne modifiait pas la quantité de facteur NF-1 présent dans des extraits nucléaires de cellules (Cordingley et Hager, 1988). Ceci suggère qu'avant traitement hormonal le site NF-1 est inaccessible au facteur et que l'adjonction d'hormone modifie la structure de la chromatine pour rendre ce site accessible. On peut donc penser que la fixation des récepteurs des glucocorticoïdes facilite l'accessibilité du site NF-1 à son facteur. Des interactions protéines-protéines rendent sans doute cette interaction plus facile. Alternativement, la fixation des récepteurs peut modifier la structure de la chromatine pour rendre le site NF-1 disponible.

Pourtant, il n'y a pas coopération des récepteurs des glucocorticoïdes et de NF-1 au niveau de la fixation à l'ADN mais plutôt compétition. En effet, lors des expériences réalisées *in vitro* sur de l'ADN linéaire ou super-enroulé, ces facteurs sont des antagonistes vis-à-vis de la fixation à l'ADN (Brüggemeier et al., 1990). Lorsque les récepteurs sont fixés, NF-1 ne se fixe pas et vice versa. Cette observation est en contradiction avec le modèle énoncé ci-dessus. D'un point de vue structural ce n'est pas si étonnant parce que le site NF-1 et le dernier HRE (appelé proximal, par rapport au site promoteur) sont seulement séparés de deux paires de bases. On peut donc tout à fait imaginer un encombrement stérique pour la fixation simultanée des deux types de molécules.

Que se passe-t-il donc ? Comment réconcilier ces différentes observations ? Pour cela, il faut se souvenir des expériences de fixation de ces facteurs sur le LTR du MMTV organisé en mini-chromosome. On sait en effet comme nous l'avons vu, que le promoteur du MMTV peut être organisé *in vitro* en nucléosomes qui sont arrangés selon un mode extrêmement précis (Perlmann et Wrangle, 1988 ; Pina et al., 1990a). Dans ces conditions, les récepteurs purifiés peuvent se fixer à leur site avec une bonne affinité tandis que NF-1 en est incapable (Pina et al., 1990a). L'adjonction d'hormone provoque *in vivo* la fixation de NF-1 sur son site, dans le LTR arrangé en nucléosomes (Cordingley et al., 1987). Dans ce cas, cependant, on ne visualise pas la fixation des récepteurs sur leur site, ce qui colle bien avec les expériences de compétitions réalisées *in vitro*. Il est donc possible d'imaginer un modèle où, *in vivo*, la fixation des récepteurs à leur site modifierait la structure de la chromatine et permettrait alors la fixation de NF-1. Ceci libérerait le promoteur de

l'action répressive des nucléosomes observés sur le taux de transcription basale du récepteur (Kalff et al., 1990). Alors NF-1 chasserait les récepteurs et jouerait son rôle de facteur de transcription. En poussant un peu loin ce modèle, on pourrait même imaginer qu'ici, les récepteurs jouent seulement un rôle structural, détruisant l'organisation précise des nucléosomes et permettent à un facteur de transcription actif de se fixer. L'activation transcriptionnelle observée serait le reflet de l'action de NF-1.

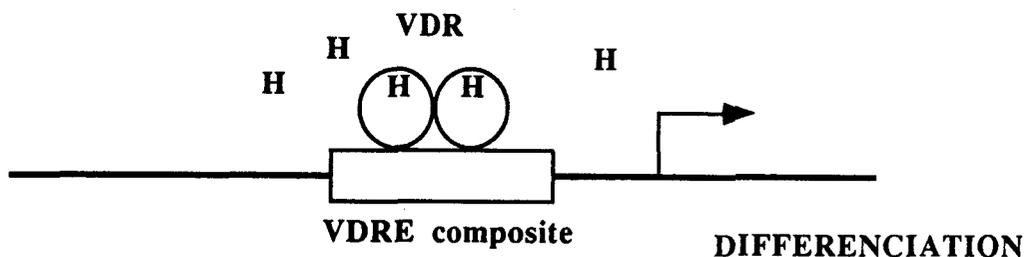
Ce modèle déjà complexe montre bien comment l'étude sous plusieurs angles *in vivo* et *in vitro* d'un promoteur donné, permet d'approcher la compréhension intime d'un tel système. Pourtant la situation est encore plus compliquée car ni le facteur NF-1, ni les récepteurs nucléaires ne sont suffisants pour aboutir à une induction hormonale correcte (figure 21). Il a été récemment démontré que juste en aval du site NF-1 se trouvaient deux sites capables de fixer efficacement le facteur de transcription Oct-1 (Brüggemeier et al., 1991). Contrairement à ce qui se passe avec NF-1, il y a fixation coopérative de Oct-1 et des récepteurs nucléaires, en présence d'hormone. On a donc des interactions complexes entre au moins trois partenaires : les récepteurs, NF-1 et Oct-1. D'autres régions du promoteur, plus en amont, sont également importantes pour l'inductibilité et pour la spécificité tissulaire (Gouilleux et al., 1991 ; Mink et al., 1991 ; Lefebvre et al., 1991).

## 2) La fixation sur les mêmes sites d'ADN.

Le cas le mieux décrit de compétition pour la fixation sur un même site d'ADN est celui du promoteur de l'ostéocalcine qui est régulé positivement à la fois par la vitamine D et l'acide rétinoïque, et qui est par contre régulé négativement par le facteur de transcription AP-1 (Schüle et al., 1990a). Il existe d'autres exemples de ce type, notamment le gène de la proliférine (Diamond et al., 1990).

L'exemple de l'ostéocalcine mérite d'être précisé quelque peu, car il montre bien l'antagonisme qui peut exister entre la différenciation cellulaire (souvent induite par les récepteurs nucléaires) et la prolifération (mise en route par les proto-oncogènes de type *fos* et *jun* ; Schüle et Evans, 1991). Cela dit, il faut se méfier très fort de ce type de simplification (différenciation/récepteurs contre prolifération/oncogènes) souvent abusive! Le gène de l'ostéocalcine est régulé positivement par la vitamine D3 et par l'acide rétinoïque via l'interaction de leurs récepteurs avec un site appelé au choix, VDRE ou RARE (figure 22). On sait depuis longtemps que cette induction est contrecarrée par diverses molécules, dont l'interleukine 1 et la calcitonine. La surprise a été de constater que le complexe *fos/jun* (autrement dit le facteur AP-1) qui peut médier l'action d'une cytokine comme l'interleukine 1 se fixe sur le même site que le récepteur de la vitamine D. Il s'agit en effet d'un VDRE un peu particulier qui combine à la fois un site AP-1 et un palindrome classique de type HRE (figure 22). Si l'une des protéine est fixée à son site elle empêche la fixation de l'autre. Fixé sur ce site, le facteur AP-1 n'active pas la transcription.

En présence de vitamine D : induction de l'ostéocalcine



Ajout de facteurs de croissance (IL-1, calcitonine etc...) → Induction du facteur AP-1

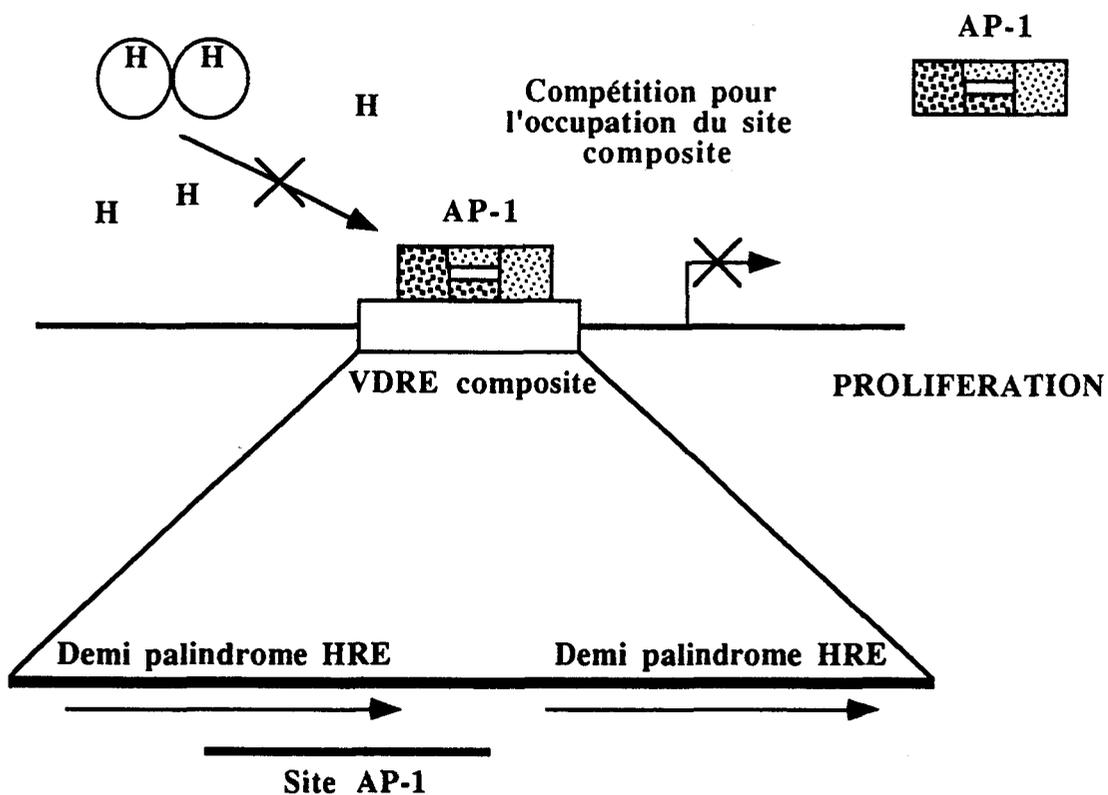


Figure 22 : Un exemple d'interaction entre récepteurs nucléaires et facteurs de transcription : compétition entre le récepteur de la vitamine D et le complexe AP-1 pour l'occupation d'un site HRE composite unique (d'après Schüle et al., 1990a).

Il y a donc à chaque instant la possibilité d'un effet de balance entre les quantités de complexe hormone-récepteur et de facteur AP-1. (Schüle et al., 1990a). Il est bien évident que ce type d'interaction entre récepteur et facteur de transcription repose sur l'existence de tels sites très particuliers.

### 3) Les interactions protéines-protéines.

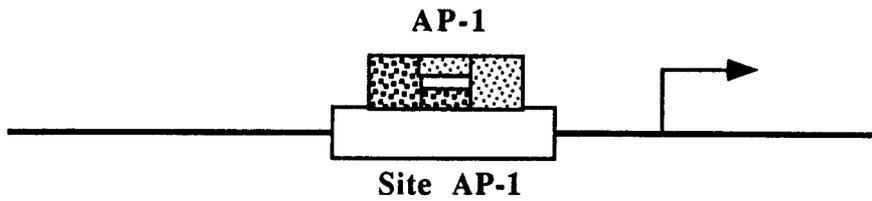
Tout récemment, un autre type possible d'interaction entre récepteurs nucléaires et facteurs de transcription a été mis en évidence par plusieurs groupes (Lucibello et al., 1990 ; Jonat et al., 1990 ; Yang-Yen et al., 1990 ; Schüle et al., 1990b ; revue dans Schüle et Evans, 1991). Ici, il n'y a pas d'intervention de la fixation à l'ADN mais directement des interactions entre protéines. La première mise en évidence de ce type de mécanisme a d'abord concerné le récepteur des glucocorticoïdes et le facteur AP-1. Depuis, cela a aussi été démontré pour d'autres membres de la superfamille des récepteurs nucléaires (récepteurs des oestrogènes, de l'acide rétinoïque ; Schüle et al., 1991 ; Nicholson et al., 1990 ; Desbois et al., 1991b ; Gaub et al., 1990 ; Doucas et al., 1991 ; Chatterjee et al., 1991 ) et d'autres facteurs de transcription comme NF-1 ou CREB (Martinez et al., 1991 ; Chatterjee et al., 1991 ).

C'est essentiellement en étudiant la régulation de la collagénase que ces nouvelles interactions ont été mises en évidence. Comme le gène de l'ostéocalcine, celui de la collagénase est régulé à la fois par le facteur AP-1 et par les glucocorticoïdes (figure 23). Cette fois-ci, il s'agit d'une régulation positive par AP-1 et d'une répression par les glucocorticoïdes. L'effet positif de AP-1 est médié par un site classique AP-1 situé dans le promoteur du gène de la collagénase. Bizarrement, la régulation négative par les glucocorticoïdes passe également par ce site, alors que le récepteur des glucocorticoïdes ne peut pas s'y fixer. La surprise a été de s'apercevoir que si on transfère le site AP-1 sur un promoteur hétérologue, on reproduit ce même effet de répression par les glucocorticoïdes (Lucibello et al., 1990 ; Schüle et al., 1990b).

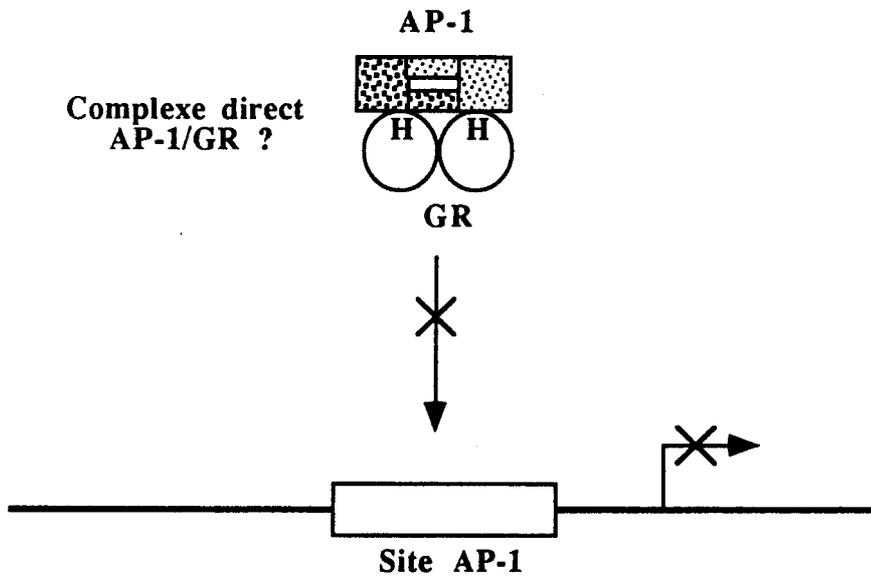
Les régions nécessaires à cet effet pour chacune des molécules impliquées ont été précisées par des expériences de mutagenèse. Pour le récepteur des glucocorticoïdes, c'est la région de fixation à l'ADN qui est requise bien que le récepteur n'ait pas besoin d'être fixé à l'ADN pour agir dans ce système. Ceci n'est pas très étonnant, car on sait depuis les expériences de mutagenèse fine de ce domaine, que certaines régions, comme la partie supérieure du doigt C2, sont impliquées dans des interactions protéines-protéines (Schéna et al., 1990). Par contre, pour le récepteur des oestrogènes également impliqué dans une interaction avec AP-1 le domaine de fixation à l'ADN n'est pas requis (Gaub et al., 1990). Il faut attendre d'autres expériences pour clarifier ce point.

Pour AP-1 la détermination des zones importantes est moins claire (si cela est possible !) et n'est pas complètement réalisée. Rappelons que AP-1 est constitué de deux protéines majeures codées par les gènes *c-fos* et *c-jun*. Pour *jun*, la région importante semble être le "leucine zipper", c'est à dire la région nécessaire pour la dimérisation avec d'autres protéines (Schüle et al., 1990b).

**Promoteur de la collagénase : régulation positive par AP-1**



**Régulation négative par les glucocorticoïdes, mais aucun GRE dans le promoteur de la collagénase**



**Figure 23 : Un autre exemple d'interaction entre récepteurs nucléaires et facteurs de transcription : formation d'un complexe, direct ou indirect, entre AP-1 et le récepteur des glucocorticoïdes (GR).**

Pour fos, il semble que la région importante pour cette interaction soit la partie N-terminale, dont la fonction n'est pas connue et qui n'est pas conservée pour les membres de la famille fos (comme fosB ou fra-1 ; Lucibello et al., 1990). En accord avec cette expérience, le produit du gène fosB n'est pas capable d'inhiber l'action du récepteur des glucocorticoïdes.

Est-ce que ces différentes molécules se complexent les unes avec les autres ou est-ce qu'elles entrent en compétition pour la fixation à un facteur important de la machinerie de base de la transcription ? Il semble bien qu'il y ait un complexe direct entre AP-1 et le récepteur des glucocorticoïdes. C'est en tout cas ce qui a été observé par certains auteurs (Jonat et al., 1990 ; Yang-Yen et al., 1990) mais pas par d'autres (Schüle et al., 1990b ; Lucibello et al., 1990). Ce n'est pas parce qu'on n'arrive pas à voir un complexe protéique que ce complexe n'existe pas : il peut être très fugace ou alors fragile et ne pas résister aux conditions expérimentales utilisées. Voilà encore un débat ouvert...

Cette découverte est vraiment importante, car elle conduit à réexaminer le rôle possible des hormones à action intra-cellulaire, comme les stéroïdes ou les rétinoïdes. Elle suggère deux modes d'action pour ces molécules (i) le mode d'action "classique" via l'interaction de leur récepteur avec l'ADN et la modulation de la transcription des gènes cibles ; (ii) l'interaction directe du récepteur avec des facteurs de transcription comme AP-1. En présence de l'hormone, il y aura donc activation d'un réseau de gène donné par fixation sur les promoteurs de ces gènes et répression de l'expression d'autres gènes contrôlés par AP-1 ou d'autres facteurs de transcription (Schüle et Evans, 1991). Il va falloir maintenant déterminer précisément quel est le rôle de l'hormone dans cette deuxième voie. Ce type de mécanisme n'est pas valable uniquement pour le récepteur des glucocorticoïdes, mais semble bien être une caractéristique générale des récepteurs nucléaires...

#### 4) La titration de facteurs communs.

Ce paragraphe est encore virtuel, parce qu'à l'heure actuelle, on n'a pas une compréhension fine des relations qui existent entre facteurs de transcription et récepteurs nucléaires. Mais tout le monde est persuadé que cela va bientôt venir ! Dans le cas que nous venons de voir, on soupçonne fortement un complexe direct ; dans d'autres cas, comme celui du gène de l'ovalbumine, on peut penser à un effet d'interférence transcriptionnelle ("squelching") via la titration de facteurs inhibiteurs ou activateurs des différentes molécules impliquées (Schüle et Evans, 1991).

Chez le poulet, le gène de l'ovalbumine est régulé positivement, à la fois par fos, jun et le récepteur des oestrogènes, d'une manière synergique. L'équipe de Pierre Chambon et de Paolo Sassone-Corsi viennent de démontrer que cette co-activation se produisait via l'interaction de fos et jun avec un site AP-1 situé dans le promoteur de ce gène (Tora et al., 1988 ; Gaub et al., 1990). C'est le facteur AP-1 qui se fixe sur le promoteur de l'ovalbumine et transactive de manière

oestrogène-dépendante. Le récepteur aux oestrogènes ne se fixe pas sur ce site et son domaine de fixation à l'ADN n'est pas requis pour la synergie avec AP-1. Bien que l'on ne puisse pas exclure l'existence d'un complexe direct entre AP-1 et le récepteur des oestrogènes, il est tentant d'imaginer d'autres scénarios. En effet, tout récemment, l'équipe de Paolo Sassone-Corsi a mis en évidence un facteur appelé IP-1 qui est un inhibiteur de fos et jun. IP-1, dont l'activité semble être modulée par phosphorylation, va empêcher fos et jun d'agir sur leurs gènes cibles. On peut imaginer que le récepteur des oestrogènes complexe IP-1 d'une manière hormono-dépendante et, ce faisant, permet une plus forte action du complexe AP-1 sur le promoteur du gène de l'ovalbumine de poulet.

## **PHYSIOLOGIE CELLULAIRE ET RECEPTEURS NUCLEAIRES**

Pour exercer leur action physiologique, les hormones stéroïdes, les hormones thyroïdiennes ou encore l'acide rétinoïque, vont modifier la transcription des gènes de la cellule cible. On estime à environ 1000, le nombre de gènes mis en route dans une cellule lorsqu'on applique les glucocorticoïdes. Chaque hormone ayant une action physiologique bien précise, il nous faut examiner quels sont les gènes mis en route. De plus, comme on vient de le voir, la spécificité d'action des hormones n'est pas strictement codée par leur récepteur, puisque ceux-ci semblent être capables de reconnaître des séquences similaires voire identiques sur l'ADN. Comment peut-on relier les connaissances que l'on a sur le mode de fonctionnement des récepteurs nucléaires et des effets physiologiques donnés, liés à ces récepteurs ?

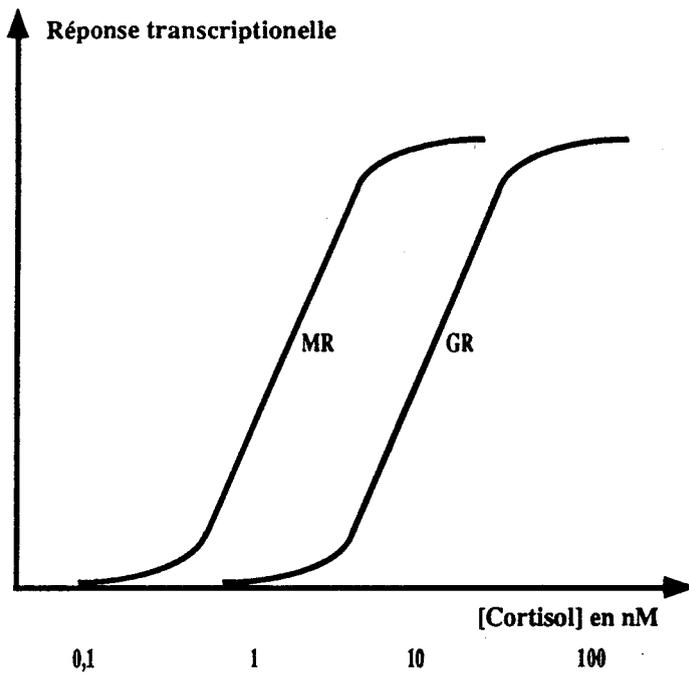
Il serait illusoire de présenter, même dans les grandes lignes, l'action physiologique de tous les récepteurs nucléaires. Pour traiter le problème des réseaux de gènes régulés, nous observerons de plus près l'action des glucocorticoïdes et des minéralocorticoïdes dans la cellule. Dans un deuxième temps nous essaierons de décortiquer un effet physiologique connu d'un ligand de récepteur nucléaire : l'action morphogène de l'acide rétinoïque au cours du développement du bourgeon de membre. Nous avons particulièrement développé ce modèle car il permet de comprendre comment l'étude fine d'un phénomène fait, peu à peu, émerger son immense complexité. Nous terminerons la discussion des effets physiologiques des récepteurs nucléaires par une évocation rapide de leurs implications dans deux types de maladies chez l'homme : les syndromes de résistance aux hormones et le cancer.

### **A) LA SPECIFICITE DES EFFETS PHYSIOLOGIQUES.**

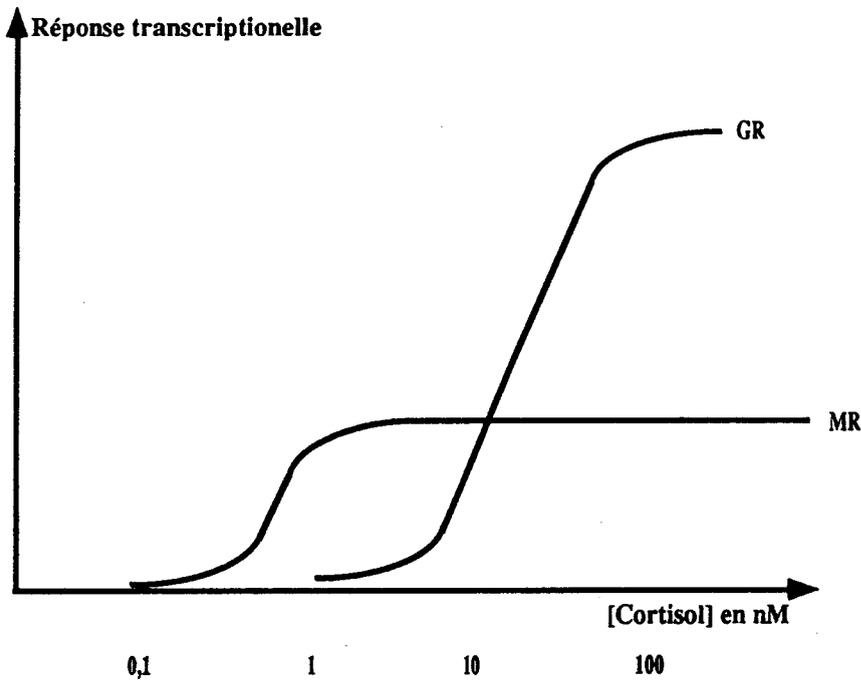
#### **1) Les récepteurs des glucocorticoïdes et des minéralocorticoïdes.**

Les minéralocorticoïdes et les glucocorticoïdes sont sécrétés par les glandes surrénales. Leurs effets physiologiques sont très différents. L'aldostérone (le minéralocorticoïde majeur chez l'homme) régule le métabolisme du calcium en favorisant sa réabsorption par les reins. Son action est très largement sous la dépendance du système rénine/angiotensine et est sensible à la tension artérielle. Les glucocorticoïdes ont une action beaucoup plus étendue qui s'étend du métabolisme des sucres jusqu'aux effets sur le développement (revue de Funder, 1987).

Avec le clonage des récepteurs nucléaires des glucocorticoïdes (Weinberger et al., 1985) et des minéralocorticoïdes (Arriza et al., 1987, 1988), on s'est aperçu que chacun des deux récepteurs répond aux deux types d'hormones, ce qui confirmait pleinement des données



**Courbe dose/réponse de MR et de GR en fonction de la quantité de cortisol (un glucocorticoïde). MR est sensible à des doses plus faibles en cortisol que GR.**



**Comparaison des réponses transcriptionnelles de MR et de GR à des doses croissantes de cortisol. MR répond faiblement à des doses faibles et GR répond fortement à des doses fortes en cortisol.**

**Figure 24 : Illustration des rôles complémentaires des récepteurs des minéralocorticoïdes (MR) et des glucocorticoïdes (GR) dans les organes où ils sont tous les deux présents (par exemple, le cerveau).**

biochimiques (revues dans Funder, 1987 ; Evans et Arriza, 1989 ; Bonvalet et al., 1991). Dans ces conditions, et sachant que les deux types de récepteurs reconnaissent tous les deux le même HRE dans le génome, comment peut-on observer *in vivo* une réponse physiologique différente pour chacune des hormones ? De plus ces données suggèrent qu'il y a deux types de récepteurs de haute affinité pour les glucocorticoïdes. Quelle est leur utilité ? Nous considérerons deux tissus cibles différents pour avoir un début de réponse à ces questions.

## 2) Les effets des minéralocorticoïdes dans les reins.

Il faut rappeler que dans le plasma il y a une concentration en glucocorticoïdes beaucoup plus élevée que celle en aldostérone. Paradoxalement, si on mesure la réponse transcriptionnelle des récepteurs à un ligand donné on observe que le récepteur des minéralocorticoïdes est sensible à des concentrations de glucocorticoïdes beaucoup plus faibles que le récepteur des glucocorticoïdes lui-même (figure 24). Dans ces conditions comment les reins, qui expriment les deux types de récepteurs peuvent-ils réagir spécifiquement à l'aldostérone qui est, rappelons-le, un minéralocorticoïde ?

On sait maintenant que les reins, le lieu majeur d'action des minéralocorticoïdes expriment une enzyme chargée de détruire les glucocorticoïdes présents (Stewart et al., 1987 ; Funder et al., 1988). Cette enzyme, la 11- $\beta$ -hydroxystéroïde déhydrogénase est exprimée spécifiquement dans certains types cellulaires et réduit le niveau de glucocorticoïdes présents en les transformant en métabolites non actifs. Par contre, l'aldostérone n'est pas dégradée par cette enzyme. Ainsi, le récepteur des minéralocorticoïdes sera occupé uniquement par la molécule idoïne, l'aldostérone, et une réponse cellulaire spécifique sera observée (Evans et Arriza, 1989). Ici, la spécificité d'action de l'hormone n'est pas codée par le récepteur, mais par une enzyme modifiant le taux des hormones présentes (Funder et al., 1988).

## 3) Les effets des glucocorticoïdes dans le cerveau.

Il est démontré que dans le cerveau certaines zones, comme l'hippocampe expriment les deux types de récepteurs (Bonvalet et al., 1991). Pourtant, les minéralocorticoïdes n'ont aucun effet au niveau du cerveau. La concentration en glucocorticoïdes circulant est très variable, non seulement en fonction d'un rythme circadien, mais aussi en fonction de l'environnement. Ainsi, elle peut passer très vite de 0,5 nM à 100 nM, sous l'effet d'un stress. La réponse physiologique doit intégrer ces variations de concentration et doit donc être adaptée à chaque circonstance. Jeffrey Arriza et Ronald Evans ont proposé que dans les neurones, les deux types de récepteurs agissent de façon coordonnée pour entraîner une réponse aux glucocorticoïdes (figure 24).

Les récepteurs des minéralocorticoïdes et des glucocorticoïdes reconnaissent tous deux les glucocorticoïdes, mais le récepteur des minéralocorticoïdes est sensible à des concentrations

plus basses en glucocorticoïdes (vers 1 nM) que le récepteur canonique des glucocorticoïdes (vers 10 nM ; Arriza et al., 1988). Par contre, la réponse du récepteur des minéralocorticoïdes est, à concentration égale, toujours plus faible (5 à 15 %) que celle du récepteur des glucocorticoïdes (Evans et Arriza, 1989). D'ailleurs, il semble que le récepteur des minéralocorticoïdes soit dépourvu d'un domaine transactivateur situé dans le domaine A/B du récepteur des glucocorticoïdes.

On peut donc imaginer que le récepteur des minéralocorticoïdes soit chargé de répondre faiblement aux faibles concentrations en cortisol, tandis que le récepteur des glucocorticoïdes répondrait aux fortes concentrations (figure 24). Les deux récepteurs, ensemble, permettraient ainsi d'avoir un spectre de réponse physiologique à cette hormone très étendu et très flexible. Il est tout à fait envisageable que, dans certaines conditions, chaque récepteur active des gènes différents, bien que, jusqu'à présent, il n'y ait pas de preuves expérimentales allant dans ce sens.

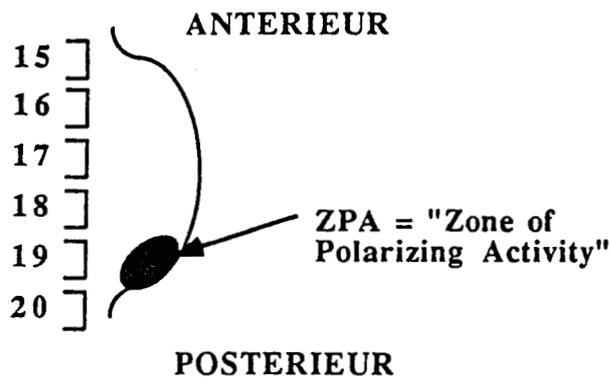
## **B) L'ACIDE RETINOIQUE ET LA MORPHOGENESE DES MEMBRES.**

### **1) Le rôle morphogène des rétinoïdes : découvertes et discussions.**

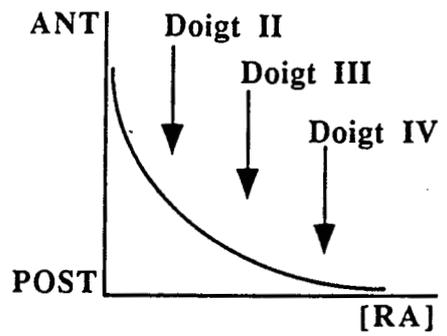
#### **a) La morphogénèse des membres.**

La morphogénèse, étape cruciale du développement des organismes supérieurs, associe trois mécanismes de base : la prolifération cellulaire, la différenciation et les communications cellulaires. Ensemble, dans un processus complexe en trois dimensions décrit par Clifford Tabin comme une chorégraphie, ces trois mécanismes permettent l'émergence de l'organisation spatiale du futur organisme (Tabin, 1991). Par sa facilité d'accès et sa taille, le bourgeon de membre (de poulet, de souris ou même d'amphibien) est un modèle de morphogénèse étudié de longue date par les embryologistes (revues dans Eichele, 1989 ; Brockes, 1989 ; Summerland et Maden, 1990 ; Tickle, 1991 ; Tabin, 1991 ; Ragsdale et Brockes, 1991). La plupart des études sur le développement des membres ont été réalisées sur l'embryon de poulet, parce que celui-ci est aisément accessible à l'expérimentateur. Chez cet embryon, le membre apparaît d'abord vers 3 jours de développement comme un bourgeon situé sur les flancs. Ce bourgeon est constitué de cellules mésenchymateuses situées au centre, entourées d'une couche de cellules épithéliales. A la suite de toute une série de "dialogues", dont les bases moléculaires ne sont pas comprises, le mésenchyme va induire l'épithélium à s'épaissir au niveau du bourgeon. Puis le bourgeon s'allonge rapidement, et après quelque jours, le mésenchyme se différencie d'une façon étroitement coordonnée, pour donner les différentes structures : cartilages, os, muscles etc...qui constituent le membre embryonnaire. Chez le poulet, dont le développement total dure 21 jours, la différenciation du mésenchyme est déjà largement engagée à 6 jours et, à 10 jours, la structure spatiale du membre est définitivement mise en place (Eichele, 1989 ; Tabin, 1991).

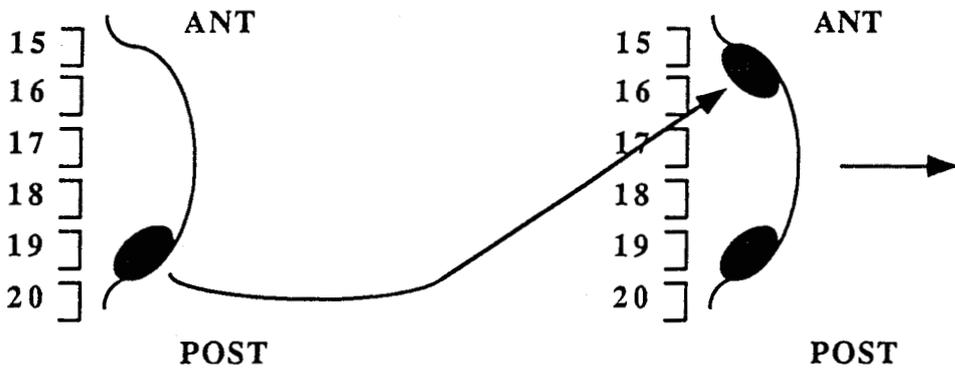
Comment une structure aussi complexe est-elle mise en place ? Plus précisément,



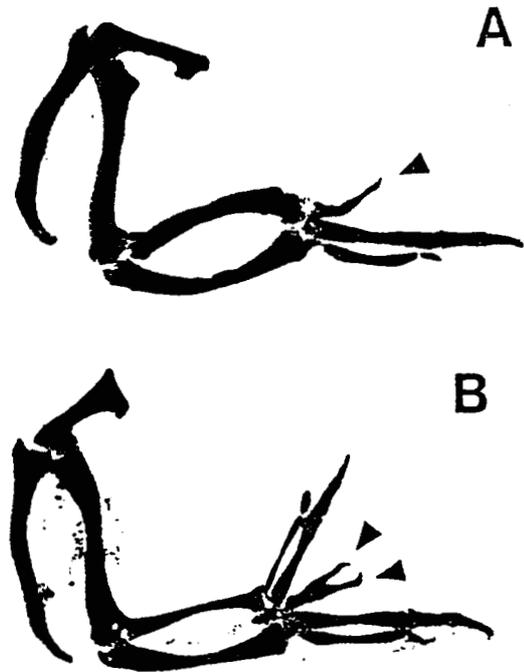
POULET A 3 JOURS



Seuils de concentration pour la formation des doigts



Grefte de ZPA



On observe le même phénomène si on greffe une microbille relarguant de l'acide rétinoïque

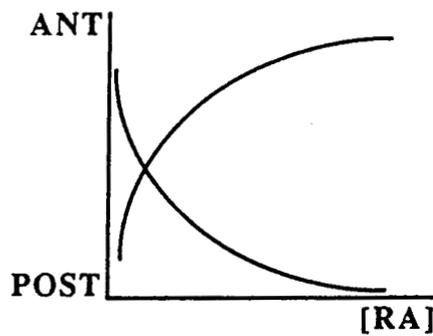


Figure 25 : La ZPA spécifie l'axe antéro-postérieur du bourgeon de membre de Poulet, à l'aide d'un gradient d'acide rétinoïque.

on peut se demander comment le bourgeon de membre "sait" où doivent se situer les doigts antérieurs et les doigts postérieurs dans le membre. Donc, comment l'axe antéro-postérieur est-il spécifié ? Pour se repérer plus aisément, il faut se souvenir que si l'on considère la main de l'homme le côté antérieur est le pouce tandis que le côté postérieur est le petit doigt.

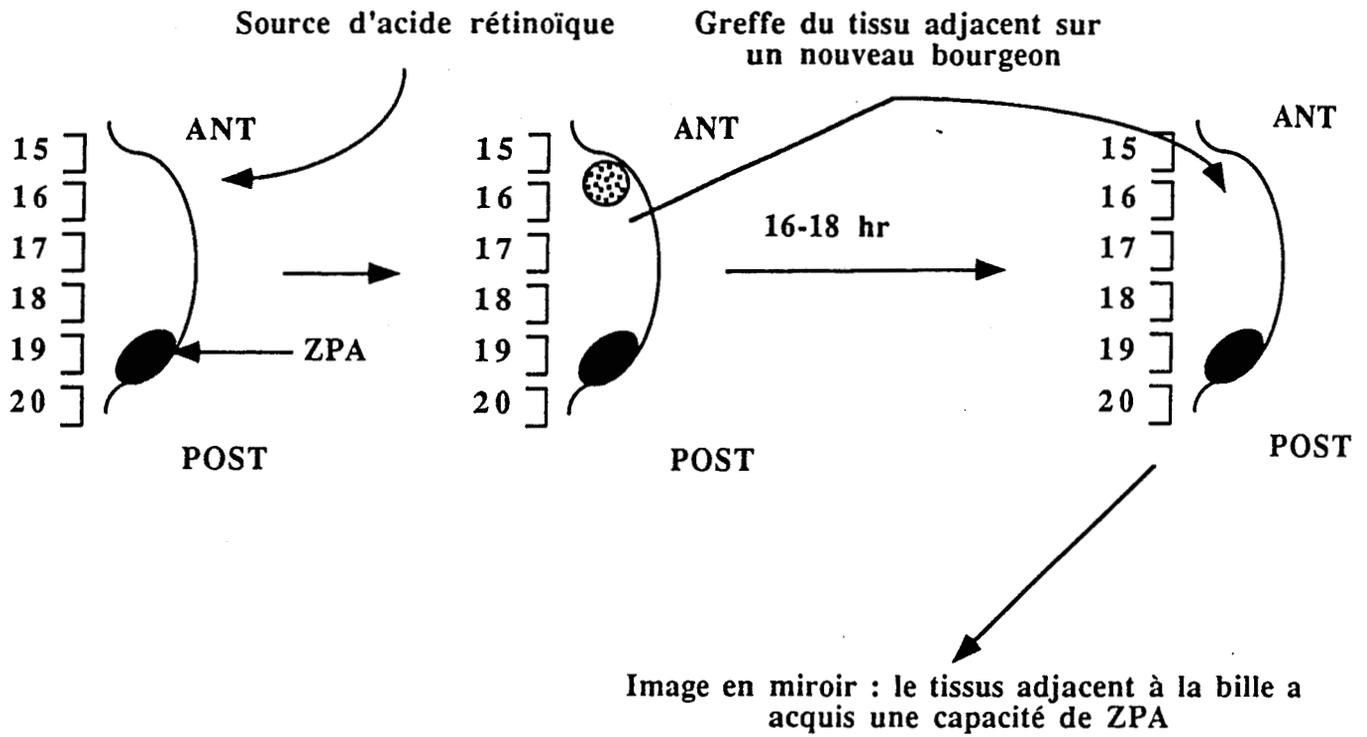
b) Zone d'activité polarisante et gradient.

Une fascinante expérience d'embryologie menée à la fin des années 1960 a donné les premiers indices pour expliquer ce phénomène en termes moléculaires. En effet, si on transplante une pièce du mésenchyme postérieur du bourgeon à trois jours dans la position antérieure d'un autre bourgeon du même âge on obtient, après développement, un membre anormal présentant une duplication en miroir du nombre des doigts (Saunders et Gasseling, 1968). Si on considère le nombre normal de doigts d'un membre de poulet, selon la série numérotée 234, on obtient, après transplantation, la série 432234 (figure 25). La région postérieure du bourgeon capable d'induire ces structures surnuméraires a été appelée zone d'activité polarisante ou ZPA ("Zone of Polarizing Activity"). Menée d'abord chez le poulet, cette expérience a été par la suite étendue à de nombreux autres vertébrés (Eichele, 1989). L'interprétation la plus classique de ce type d'expérience consiste à dire que la ZPA produit une substance diffusible qui va former un gradient le long de l'axe antéro-postérieur. Ce gradient d'une substance que l'on va appeler un morphogène va être perturbé lors des expériences de transplantation. Au cours de ces manipulations, on va créer en fait deux gradients de sens opposés dont résulte l'image dupliquée en miroir des doigts.

c) L'acide rétinoïque est un morphogène.

L'identité du morphogène est longtemps restée mystérieuse. On a pourtant assez vite soupçonné que les rétinoïdes avaient un effet sur ce système (Maden, 1982). C'est en 1987 que Christina Thaller et Gregor Eichele ont clairement montré que l'acide rétinoïque pouvait être le morphogène recherché (Thaller et Eichele, 1987). En effet, l'application d'acide rétinoïque sur le côté antérieur du bourgeon de membre à 3 jours de développement reproduit les mêmes effets que la greffe de la ZPA. De plus, la concentration d'acide rétinoïque dans le bourgeon de membre obéit à un gradient antéro-postérieur croissant d'environ 2,5 fois, ce qui correspond parfaitement à la théorie du gradient (figure 25 ; Thaller et Eichele, 1987 ; Slack, 1987). L'acide rétinoïque est présent dans d'autres endroits de l'embryon, par exemple le système nerveux central en cours de développement également sous forme de gradient de concentration. (Durstun et al., 1989 ; Wagner et al., 1990 ; Ruiz i Atalba et Jessell, 1991).

Pourtant, la situation est plus complexe. Par exemple, l'acide rétinoïque n'est pas le seul rétinoïde présent dans le bourgeon de membre. Il représente en effet l'extrémité d'une chaîne métabolique complexe qui part de la vitamine A (Blomhoff et al., 1990). Ses deux précurseurs immédiats sont le rétinol et le rétinol qui sont présents dans le bourgeon (Eichele, 1989). Le rétinol



Ce n'est pas l'acide rétinolique retenu dans le tissu adjacent qui provoque l'effet observé car cet effet ne se voit qu'après 16 à 18 heures, bien après que la concentration d'acide rétinolique autour de la bille ait atteint une concentration maximale.



Les doigts supplémentaires viennent bien de l'hôte et non du greffon (preuve à l'aide de greffes de Caille dans du Poulet).

Figure 26 : Une expérience récente suggérant que l'acide rétinolique n'est pas un morphogène, mais permet l'induction d'un morphogène qui reste à découvrir (d'après Tabin, 1991 et Wanek et al., 1991).

peut également donner un autre rétinoïde l'acide 3,4-didehydro-rétinoïque, qui a la même activité que l'acide rétinoïque et qui est également présent sous forme d'un gradient, mais à des concentrations 6 fois plus élevées que celles de l'acide rétinoïque (Thaller et Eichele, 1990 ; Hoffman, 1990).

Le fait que ce gradient s'étende sur une échelle de concentration assez limitée constitue une objection importante à ce modèle. Entre la zone postérieure et la zone antérieure, la différence de concentration en rétinoïdes est seulement de 2,5 fois (Eichele, 1989 ; Tabin, 1991). Le système de lecture d'un tel gradient doit donc impérativement amplifier d'une quelconque manière cette différence, pour donner à chaque cellule une information positionnelle correcte. Deux mécanismes possibles ont été proposés. Ils reposent tous deux sur l'idée que des protéines lient l'acide rétinoïque. Le premier de ces mécanismes utilise les protéines cytoplasmiques qui lient l'acide rétinoïque. On a en effet montré que ces protéines, sur lesquelles nous reviendrons plus en détail, sont présentes sous forme d'un gradient dans le bourgeon de membre. Ce gradient a une orientation inverse de celui d'acide rétinoïque (fortes concentrations du côté antérieur et faibles concentrations du côté postérieur). Ceci permettrait peut être d'accentuer très fort le gradient d'acide rétinoïque libre disponible pour jouer un rôle dans le noyau. Le deuxième mécanisme d'amplification repose sur les trois types de récepteurs nucléaires qui fixent l'acide rétinoïque. L'expression de certains de ces récepteurs est régulée par l'acide rétinoïque et on peut imaginer qu'un mécanisme de rétro-contrôle contribue à augmenter la sensibilité de la réponse (Tabin, 1991).

#### d) L'acide rétinoïque n'est pas un morphogène.

L'enthousiasme du début est pourtant assez vite retombé car un certain nombre d'expériences récentes remettent en cause le fait que l'acide rétinoïque soit le morphogène du bourgeon de membre (Brockes, 1991). Il aurait un rôle dans la mise en place d'un hypothétique gradient de morphogène, mais ne serait pas le morphogène lui-même.

Ainsi, si on greffe une source d'acide rétinoïque dans la région antérieure d'un bourgeon de membre à 3 jours et qu'après un temps variable, on transplante sur un nouveau receveur le tissu qui a été en contact avec la source d'acide rétinoïque, on peut montrer qu'il induit une activité ZPA dans cette région (figure 26 ; Wanek et al., 1991). Il faut cependant attendre 16 à 18 heures entre les deux transplantations pour observer cet effet. Les auteurs de ce travail ont pu exclure que les cellules de la deuxième greffe produisent (ou contiennent) de l'acide rétinoïque. Leur interprétation est que l'acide rétinoïque induit les cellules situées autour de la première greffe à acquérir une activité ZPA, peut être en sécrétant le vrai morphogène (Tabin, 1991 ; Brockes, 1991).

De plus, il y a une grosse différence entre la concentration d'acide rétinoïque exogène nécessaire pour obtenir un effet biologique observable, et les concentrations d'acide rétinoïque endogène dans le bourgeon de membre. Par exemple, en utilisant un gène indicateur

contenant un RARE efficace (celui du récepteur de l'acide rétinoïque de type  $\beta$  ; De Thé et al., 1990b) on peut montrer que ce RARE est effectivement induit par un implant d'acide rétinoïque installé dans le bourgeon de membre, mais pas par une transplantation de ZPA. Les concentrations d'acide rétinoïque ou d'acide 3,4-didehydrorétinoïque endogènes putativement sécrétées par la ZPA ne semblent donc pas suffisantes pour réguler effectivement un RARE *in vivo* (Noji et al., 1991).

Pour l'instant, il semble donc que l'acide rétinoïque ait "simplement" un rôle causal dans l'établissement d'un gradient d'un morphogène qui reste à identifier. Mais la situation peut encore évoluer et il ne faut pas prématurément exclure le rôle morphogène de l'acide rétinoïque ou de certains de ces métabolites. Nous discuterons ce point un peu plus loin.

## 2) Les récepteurs de l'acide rétinoïque.

### a) Caractérisation.

Trois récepteurs de l'acide rétinoïque, appelés  $RAR\alpha$ ,  $RAR\beta$  et  $RAR\gamma$  ont été successivement découverts tant chez les mammifères que chez d'autres organismes, comme par exemple le poulet ou le xénope (Petkovitch et al., 1987 ; Giguere et al., 1987 ; De Thé et al., 1987 ; Brand et al., 1988 ; Benbrook et al., 1988 ; Zelent et al., 1989 ; Krust et al., 1989 ; Ragsdale et al., 1989 ; Giguere et al., 1989 ; Noji et al., 1991 ; voir la revue de Ragsdale et Brockes, 1991). Tous ces récepteurs partagent la structure classique en domaines, des récepteurs nucléaires. Leurs plus proches homologues sont les récepteurs des hormones thyroïdiennes et le gène ear-1. On sait maintenant que les récepteurs de l'acide rétinoïque reconnaissent un HRE qui leur est propre et qui est appelé RARE. Les deux premiers RARE ont été caractérisés en amont du gène de la laminine B1 (Vassios et al., 1989) et du gène du  $RAR\beta$  (De Thé et al., 1990b ; Sucov et al., 1990 ; Hoffman et al., 1990).

### b) De multiples isoformes des RARs.

Non seulement il existe trois types de récepteurs de l'acide rétinoïque, dont les rôles respectifs ne sont pas encore compris mais, en plus, chaque gène peut être épissé de très nombreuses manières dans sa partie 5' (Ragsdale et Brockes, 1991). Ainsi il existe 7 isoformes 5' du  $RAR\gamma$  dont deux varient dans le domaine A/B. Les cinq autres formes ont des régions 5' non traduites différentes sans que l'on comprenne la signification de ces épissages alternatifs (Kastner et al., 1990). La même chose est vraie pour le  $RAR\alpha$  (7 isoformes 5' dont deux sont dans la région A/B codante ; Leroy et al., 1991a) et pour  $RAR\beta$  (3 isoformes différentes dans la région A/B ; Zelent et al., 1991). En tout, il y a donc 7 isoformes de récepteurs différentes dans la partie N-terminale. Il semble que souvent ce soit le sous domaine A qui change alors que la partie distale du domaine A/B reste constante. On ignore encore à quoi servent ces différentes isoformes mais il est probable qu'elles permettent d'obtenir des récepteurs ayant des domaines transactivateurs

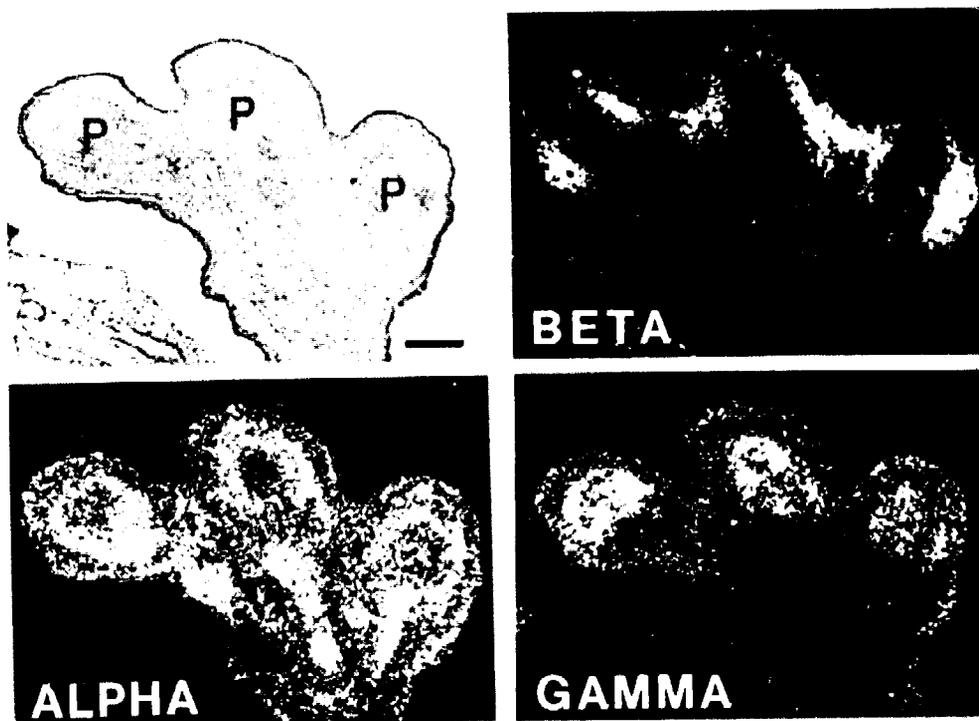


Figure 27 : Expression différentielle des récepteurs de l'acide rétinoïque dans le bourgeon de membre d'un embryon de souris à 14,5 jours de développement. P : formation osseuses embryonnaires. La barre représente 160  $\mu\text{m}$ .

différents et des traductibilités différentes. Enfin, il faut préciser que pour les RAR $\alpha$  et RAR $\beta$  ce sont des promoteurs différents qui régulent l'expression des différentes isoformes. Dans les deux cas, l'un des promoteurs est inductible par l'acide rétinoïque et l'autre pas (De Thé et al., 1990b ; Brand et al., 1990 ; Leroy et al., 1991a ; Zelent et al., 1991). Ceci permet d'imaginer une régulation très fine de l'expression de ces gènes.

#### c) Expression dans le bourgeon de membre.

L'expression des différents RAR (mais pas encore de leurs isoformes, sauf récemment pour le RAR $\gamma$  de Xénope ; Ragsdale et Brockes, 1991) a été étudiée par hybridation *in situ* d'une manière très complète (Giguere et al., 1989 ; Dollé et al., 1989 ; 1990 ; Ruberte et al., 1990 ; 1991 ; Smith et Eichele, 1991 ; Rossant et al., 1991). Concentrons-nous sur le bourgeon de membre de souris (figure 27). RAR $\alpha$  est exprimé de manière ubiquiste dans le bourgeon (quoique un peu plus faiblement dans les cartilages) comme, en première approximation, dans le reste de l'embryon. Par contre, l'expression des deux autres gènes semble être beaucoup plus finement régulée et, curieusement, RAR $\beta$  et RAR $\gamma$  semblent avoir des territoires d'expression complémentaires. Ainsi, RAR $\beta$  est exprimé dans un premier temps dans les parties proximales du bourgeon puis, dans un second temps, dans le mésenchyme interdigital, là où de nombreuses cellules sont engagées dans un processus d'apoptose (ou mort cellulaire programmée). A 10 jours de développement chez la souris, RAR $\gamma$  est distribué de manière homogène ; mais à 12,5 jours il devient spécifique du mésenchyme proximal et dans les condensations pré-cartilagineuses (Ragsdale et Brockes, 1991). Une information importante à retenir est qu'aucun des RARs ne semble être exprimé sous forme de gradient antéro-postérieur au cours de la morphogenèse du bourgeon de membre, contrairement aux protéines cytoplasmiques fixant l'acide rétinoïque.

#### d) Les gènes RXRs.

Comme si la situation n'était déjà pas assez compliquée avec les nombreuses isoformes possibles des trois types de récepteur de l'acide rétinoïque, l'équipe de Ronald Evans a récemment annoncé le clonage d'un nouveau récepteur orphelin, appelé RXR, qui peut être activé par l'acide rétinoïque (Mangelsdorf et al., 1990 ; Rowe et al., 1991). Les auteurs pensent que ce gène qui n'est pas proche des RARs du point de vue des homologies de séquence (seulement 61% et 28% avec les domaines C et E de RAR $\alpha$ ), serait un récepteur nucléaire soit à l'acide rétinoïque lui-même, soit plus probablement à un autre rétinoïde. D'après des études toutes nouvelles (Levin et al., 1992 ; Heyman et al., 1992), il semble que le ligand de ce récepteur soit l'isomère *cis* de l'acide rétinoïque (les RARs fixent la forme *trans*).

Il est en tout cas important de noter que le produit de RXR se fixe sur un HRE différent des RARE et TRE déjà décrits. Ce RXRE obéit en effet à la loi du 3-4-5 énoncée par Ron Evans et son équipe, mais l'espacement entre les deux répétitions de la séquence AGGTCA est dans ce cas de 1 nucléotide seulement (Mangelsdorf et al., 1991). Par cette voie, l'acide

**TABEAU 4**

**Les protéines fixant l'acide rétinoïque (d'après Blomhoff et al., 1990)**

<b>Protéines</b>	<b>PM</b>	<b>Ligand</b>	<b>Fonctions possibles</b>
RBP	21	Rétinol	Transport du rétinol dans le plasma
IRBP	140	Retinol et rétinol	Transport intercellulaire dans le cycle visuel
4 protéines sécrétées de l'utérus de porc	22	Retinol	Transport vers le foetus
2 protéines de l'épididyme de rat	20	Acide rétinoïque	Transport intercellulaire
CRBP1	16	Retinol	Transport intracellulaire
CRBP2	16	Retinol	Transport intracellulaire, concentration du retinol dans les organes utilisateurs
CRBP3	15	Retinol	Trouvée dans les yeux de Poisson Rôle inconnu
CRABP1	16	Acide rétinoïque	Transport intracellulaire, régulation de la concentration d'acide rétinoïque libre
CRABP2	15	Acide rétinoïque	Transport intracellulaire, régulation de la concentration d'acide rétinoïque libre
CRALBP	36	Rétinal	Réactions enzymatiques au cours du cycle visuel
RAR $\alpha$	50	Acide rétinoïque	Facteur de transcription ligand-dépendant
RAR $\beta$	50	Acide rétinoïque	Facteur de transcription ligand-dépendant
RAR $\gamma$	50	Acide rétinoïque	Facteur de transcription ligand-dépendant

*cis*-rétinoïque pourra donc activer une série de gènes différents des gènes régulés par les RARs. C'est le cas du gène de la protéine cellulaire de type II fixant le rétinol régulé par RXR mais non par les RARs (CRBP<sub>II</sub> ; Mangelsdorf et al., 1991). Le gène RXR $\alpha$  n'est pas exprimé dans le bourgeon de membre mais plutôt dans les cellules dérivées de la crête neurale, chez le poulet (Rowe et al., 1991). Mais il existe aussi deux autres gènes apparentés, RXR $\beta$  et RXR $\gamma$ , qui, logiquement, devrait avoir le même ligand (Hamada et al., 1989 ; Leid et al., 1992). Peut-être seront-ils exprimés dans le bourgeon de membre...

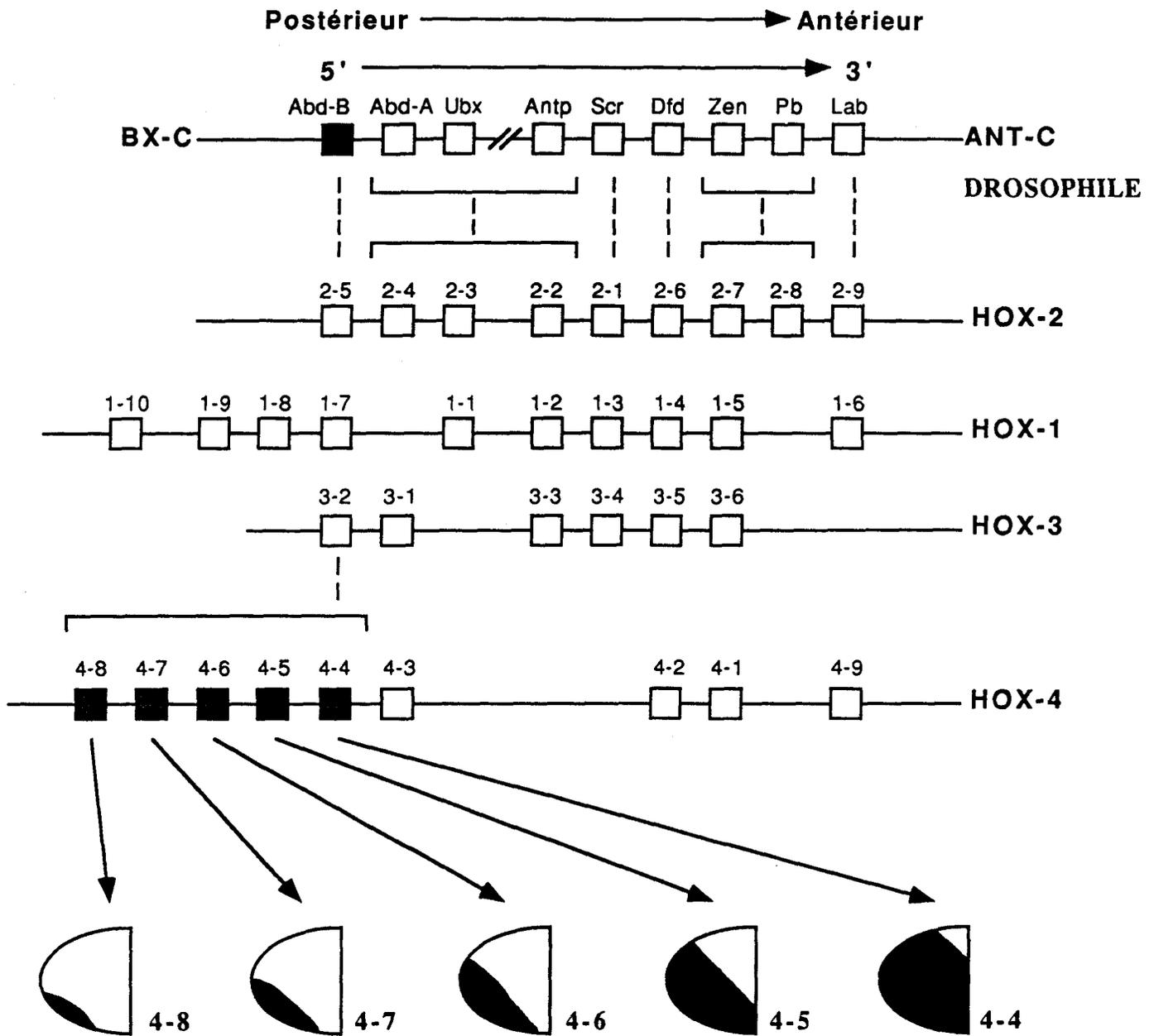
### 3) Les CRABP et CRBP.

De très nombreuses protéines, autres que les trois récepteurs nucléaires, fixent l'acide rétinoïque ou ses précurseurs (tableau 4; Blomhoff et al., 1990 ; Balling, 1991). Parmi ces protéines, ce sont les protéines cytoplasmiques fixant le rétinol (CRBP) et celles fixant l'acide rétinoïque (CRABP) qui nous intéressent ici plus particulièrement. Ces protéines sont de petite taille (environ 137 acides aminés) et sont structurellement non reliées aux récepteurs nucléaires. Il y a deux CRBPs (CRBP<sub>I</sub> et CRBP<sub>II</sub>) qui fixent le rétinol et deux CRABPs (CRABP<sub>I</sub> et CRABP<sub>II</sub>) qui fixent l'acide rétinoïque avec une bonne affinité. On a depuis longtemps pensé que ces protéines pouvaient réguler l'action de l'acide rétinoïque dans la cellule. Les CRBPs fixent le rétinol, précurseur de l'acide rétinoïque, et formeraient ainsi une réserve de rétinol intracellulaire, inactif mais rapidement mobilisable. Les CRABPs, elles, pourraient avoir un rôle de "tampon" vis-à-vis de la concentration en acide rétinoïque. Ces protéines pourraient également favoriser le transfert vers le noyau des molécules d'acide rétinoïque de manière à leur permettre de rencontrer leurs récepteurs. Il faut cependant noter que ces protéines ne sont pas indispensables pour le fonctionnement normal des RARs, puisque ceux-ci fonctionnent dans des cellules qui en sont dépourvues (revue dans Tabin, 1991). A notre connaissance, aucune activité enzymatique n'a été associée à ces molécules.

### 4) Les gènes régulés par l'acide rétinoïque dans le bourgeon de membre.

#### a) Des cibles intrigantes : les gènes homéotiques.

Parmi les gènes cibles de l'acide rétinoïque, figure une classe de gènes très importants : les gènes à boîte homéo de la série Hox (Gehring et al., 1990 ; Kessel et Gruss, 1990; Boncinelli et al., 1991). Ces gènes ont été découverts chez les vertébrés grâce à leurs homologies de séquence avec les gènes de drosophile appartenant aux complexes Bithorax et Antennapedia. Comme chez la drosophile, ces gènes sont présents sous forme de complexes chez les vertébrés (figure 28). Il y a 4 complexes de gènes Hox chez les vertébrés numérotés de Hox 1 à Hox 4 et situés sur les chromosomes 7, 17, 12 et 2 respectivement. Il a été clairement démontré que les gènes Hox fournissent à l'embryon des informations positionnelles le long de l'axe



Expression de chaque gène Hox4 dans le bourgeon de membre précoce : on arrive à un gradient d'expression antéro-postérieur du complexe des gènes Hox4-8 à Hox 4-4

Figure 28 : Les complexes des gènes Hox de Souris et l'expression des gènes de la série Hox4-8 à Hox4-4 dans les stades précoces du bourgeon de membre (d'après Tabin, 1991).

antéro-postérieur, que ce soit chez les insectes ou chez les vertébrés. Une caractéristique spectaculaire de ces gènes est que leurs domaines d'expression dans l'embryon reflètent leur position au sein du complexe où ils sont situés. Autrement dit, un gène situé en 5' sur le complexe Hox1A par exemple, sera exprimé plus postérieurement qu'un gène situé en 3' du complexe comme Hox1F (Kessel et Gruss, 1990 ; Boncinelli et al., 1991).

**b) Les gènes Hox dans le bourgeon de membre.**

De nombreux gènes Hox sont exprimés dans le bourgeon de membre. Parmi eux, les gènes Hox4.4, Hox4.5, Hox 4.6 et Hox4.7 sont particulièrement intéressants (Dollé et al., 1989b ; Izpisua-Belomonte et al., 1991). Au sein du bourgeon ils sont exprimés chacun dans une zone bien précise et forment ensemble un gradient d'expression antéro-postérieur (figure 28). Ainsi, le gène Hox4.4 situé en 3' dans cette série est exprimé dans la quasi totalité du bourgeon, le gène Hox4,5 est exprimé dans une région restreinte postérieurement et ainsi de suite (Tabin, 1991). Par contre, plus tard, les barrières d'expression entre les différents gènes Hox4 se distordent et deviennent presque parallèles à un axe proximal-distal.

D'autres gènes à boîte homéo sont exprimés dans les bourgeons de membres. C'est le cas du gène XIHbox1 (également appelé Hox3.3) dont l'expression suit un gradient antéro-postérieur qui pourrait tout à fait être créé par le gradient de morphogène (Oliver et al., 1988; 1989 ; Cho et De Robertis, 1990 ; De Robertis et al., 1991; voir aussi les exemples de La Rosa et Gudas, 1988a ; 1988b ; Eichele, 1989).

**c) Gènes Hox et acide rétinoïque: "je t'aime, moi non plus" !**

L'origine de l'expression étant située près de la ZPA, il était logique de penser que l'acide rétinoïque sécrété par la ZPA détermine le statut d'expression des gènes Hox4 qui, eux-mêmes, induiraient une série d'événements morphogénétiques spécifiques. Il est vrai que l'application d'acide rétinoïque exogène peut stimuler une expression coordonnée et correctement agencée dans l'espace de ces 4 gènes Hox4. Si on dispose l'acide rétinoïque du côté antérieur, on obtient une inversion de l'expression de ces gènes Hox4 après un long temps de latence (18 heures ; Nohno et al., 1991; Izpisua-Belmonte et al., 1991). Mais par contre l'expression de XIHbox1 n'est pas fondamentalement modifiée par l'inversement du gradient d'acide rétinoïque (Oliver et al., 1990).

Un certain nombre d'expériences réalisées *in vitro* montrent que les gènes Hox sont sensibles à l'acide rétinoïque et surtout, que cette sensibilité varie selon la position du gène considéré dans le complexe (De Robertis et al., 1991 ; Boncinelli et al., 1991 ; Tabin, 1991). Par exemple les gènes situés en 3' du complexe Hox2 sont stimulés par une concentration de 10nM en acide rétinoïque, alors que les gènes situés en 5' requièrent 10  $\mu$ M pour être induits. Ces expériences permettent de faire un lien entre l'acide rétinoïque, l'expression des gènes Hox et la morphogénèse. Ainsi, l'équipe de Peter Gruss a pu récemment montrer que le traitement par des

doses tératogènes d'acide rétinoïque induit des transformations homéotiques des vertèbres de souris. Les zones d'expression de certains gènes Hox sont également modifiées dans ces souris (Kessel et Gruss, 1991). Tout ceci va dans le sens du gradient d'acide rétinoïque morphogène.

Hélas, comme nous l'avons déjà fait remarquer plus haut, les choses ne sont pas aussi simples. C'était un trop beau modèle ! Il semble en effet que, les gènes de la série Hox4.4-Hox4.8, exprimés sous forme d'un gradient antéro-postérieur ne sont pas induits par l'acide rétinoïque *in vitro*. Si on traite des cellules de carcinome embryonnaire totipotentes par de l'acide rétinoïque, on observe une diminution de l'expression de ces gènes (Siméone et al., 1991). Cette observation, obtenue dans un système *in vitro*, différent du bourgeon de membre, ne suffit pas, à elle seule, pour exclure un contrôle de l'expression de ces gènes par l'acide rétinoïque. Mais d'autres résultats confirment que l'acide rétinoïque sous forme *trans* ne contrôle qu'indirectement la transcription des gènes Hox.

### 5) L'acide rétinoïque... au milieu du gué.

Cette histoire à épisodes montre bien la complexité des systèmes dynamiques comme celui la morphogénèse. L'acide rétinoïque n'est pas directement impliqué dans la genèse de l'activité polarisante et il ne régule pas directement les gènes Hox. Après avoir été hissé au pinacle de l'embryologie ("We have a morphogen !" ; Slack, 1987) il semble bien que l'acide rétinoïque doive en descendre ! ("We may not have a morphogen !" ; Brockes, 1991). Cependant, gardons-nous de conclure trop vite dans un sens ou dans un autre.

D'abord parce que nous ne connaissons pas tous les intervenants de ce système et de loin ! Par exemple, la découverte récente d'un ligand pour l'isomère *cis* de l'acide rétinoïque et l'existence d'enzymes capables de convertir la forme *trans* en forme *cis* (Levin et al., 1992) peut permettre d'imaginer (tout à fait gratuitement !) un rôle de la forme *cis* dans la morphogénèse. Peut-être le morphogène est-il le *cis* ?

On ne peut pas non plus glisser tranquillement sur la coexistence du gradient, même faible, d'acide rétinoïque et de celui de l'expression de certains gènes Hox. Ainsi, on peut imaginer qu'un faible gradient d'acide rétinoïque induise un fort gradient d'une autre molécule à identifier. Dans ce cas, qui est le "vrai" morphogène ?

## C) LES MALADIES LIEES AUX RECEPTEURS NUCLEAIRES.

### 1) Les syndromes de résistance aux hormones.

Les syndromes de résistance aux hormones stéroïdes ont été décrits dans la littérature médicale depuis très longtemps, mais la découverte de la cause de ces syndromes est récente et a été tout d'abord obtenue pour le récepteur de la vitamine D, dans le rachitisme hypocalcémique (Hughes et al., 1988 ; revues dans Sultan et Lobaccaro, 1991 et dans Hughes et O'Malley, 1991). Dans toutes ces affections héréditaires (qui se caractérisent le plus souvent par un haut taux plasmatique de l'hormone considérée et par une absence de réponse des tissus cibles à cette hormone) on a suspecté un défaut génétique lié au récepteur de l'hormone. Nous décrirons plus en détail les résistances liées à des anomalies des récepteurs des hormones thyroïdiennes.

La résistance aux hormones thyroïdiennes est un syndrome héréditaire autosomal dominant extrêmement rare (moins de 100 cas connus ; Refetoff et al., 1967 ; Refetoff, 1982). La plupart des sujets atteints se caractérise par la présence d'un goître et des niveaux élevés de T3 et de T4 plasmatique avec un métabolisme de ces hormones et de la TSH normal. Souvent des problèmes de métabolisme osseux et de retard mental sont associés à ce syndrome. L'extrême rareté de ce syndrome peut être expliquée soit par ses effets tragiques sur les individus affectés qui meurent très tôt, soit par le fait qu'il n'est pas toujours facile de distinguer ce syndrome des pathologies thyroïdiennes plus classiques (Refetoff, 1982). De nombreux travaux biochimiques ont permis de suspecter que la résistance aux hormones thyroïdiennes était sans doute liée à un défaut du récepteur et on a très vite lié ce syndrome au gène c-erbA-2 qui code pour le récepteur des hormones thyroïdiennes de type  $\beta$  présent sur le chromosome 3 en 3p21-25 (Usala et al., 1988a).

Pour l'instant, seules quelques études ont permis de mettre en évidence des défauts génétiques liés aux récepteurs. La première anomalie découverte a été une mutation ponctuelle dans le domaine de fixation de l'hormone du gène c-erbA-2 (Usala et al., 1990). Cette mutation change une proline en histidine tout à la fin du domaine E, une position conservée chez les gènes c-erbA-2 d'autres espèces. Dans la famille étudiée, seul un allèle est muté chez les patients atteints : la mutation est donc bien dominante. Ce n'est pas étonnant puisque des mutants régulateurs négatifs dominants dont l'extrémité C-terminale est modifiée ont été décrits pour ces gènes (Forman et Samuels, 1990a ; 1990b). L'effet régulateur négatif dominant de certains récepteurs mutés liés à des syndromes de résistance généralisée aux hormones thyroïdiennes a d'ailleurs été récemment décrit par Samuel Refetoff, Leslie De Groot et leurs équipes (Sakurai et al., 1990).

Un individu atteint à l'état homozygote et présentant un phénotype beaucoup plus anormal que les hétérozygotes a également été étudié (Usala et al., 1991). Il existe aussi un cas inhabituel où le syndrome est récessif et où il semble être lié à une délétion de la majeure partie du gène c-erbA-2 chez le patient atteint (Takeda et al., 1991). Enfin, tout récemment, chez des individus affectés de résistance aux hormones thyroïdiennes : des mutations dans c-erbA-1 ont été

mises en évidence (communication personnelle de Samuel Refetoff).

## 2) Récepteurs nucléaires et cancer : le cas des leucémies aiguës à promyélocytes.

### a) La maladie.

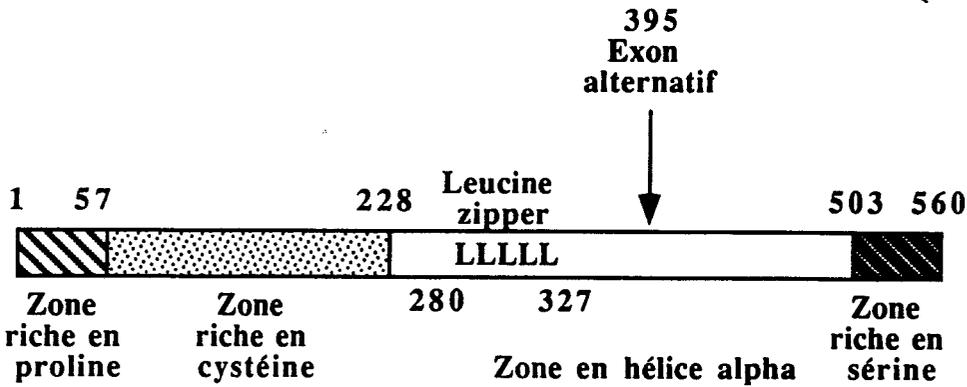
Que les récepteurs nucléaires, qui sont des molécules essentielles pour les communications entre cellules, soient impliqués dans des cancers (qui sont justement liés à des dérèglements de ces communications) n'a rien d'étonnant. Etudier le rôle des récepteurs nucléaires dans la transformation cellulaire va permettre de mieux comprendre comment ils fonctionnent. De plus, il semble que, au moins dans le cas que nous développerons (celui des leucémies aiguës à promyélocytes) la découverte d'un lien causal entre un récepteur et la maladie permettent d'envisager des voies thérapeutiques nouvelles à brève échéance. Voici un bel exemple d'inter-connection entre recherche fondamentale et applications cliniques.

La leucémie aiguë à promyélocytes (ou APL "Acute Promyelocytic Leukemia") est une forme particulière de leucémie myéloïde aiguë qui se caractérise au niveau moléculaire par la présence d'une translocation chromosomique spécifique entre les bras longs des chromosomes 15 et 17 ( $t(15;17)(q22; q11.2-q21.1)$ ) ; Rowley et al., 1977 ; voir les revues de Clarkson, 1991 ; Degos et al., 1991). Si la délimitation de la zone transloquée du chromosome 15 a été assez facile à obtenir, il en a été tout autrement pour le chromosome 17 où différents points de translocation, allant de q11.2 à q21.1 ont été proposés. Dans la région de cassure du chromosome 17, se trouvent plusieurs gènes qui peuvent être impliqués dans ce processus malin, comme le gène c-erbB-2, le gène de la myéloperoxydase, le gène c-erbA-1, celui du G-CSF, et celui du récepteur de l'acide rétinoïque de type alpha ( $RAR\alpha$ ). A un moment ou à un autre, la plupart de ces gènes, et surtout c-erbA-1, ont été supposés jouer un rôle dans la genèse de l'APL (Van Tuinen et al., 1987 ; Weil et al., 1988 ; 1989 ; Miller et al., 1989 ; Donti et al., 1989 ; Spurr et al., 1984 ; Dayton et al., 1984 ; Sheer et al., 1985 ; Longo et al., 1990a).

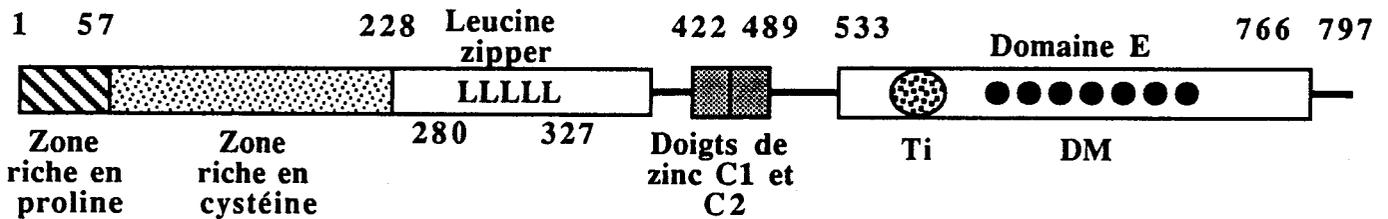
### b) La translocation t(15;17) et le récepteur alpha de l'acide rétinoïque.

La résolution de cette énigme est intervenue tout récemment grâce, entre autres, au groupe d'Anne Dejean à l'Institut Pasteur de Paris. En effet, c'est le gène du récepteur de l'acide rétinoïque de type alpha ( $RAR\alpha$ ) qui est, sur le chromosome 17, impliqué dans la translocation  $t(15;17)$  (Borrow et al., 1990 ; De Thé et al., 1990a ; Longo et al., 1990b ; Alcalay et al., 1991 ; Cleary, 1991). Plus précisément, le point de translocation est situé dans des îlots de séquence riches en dinucléotides CpG, dans l'intron qui se trouve entre les deux premiers exons codants du gène  $RAR\alpha$ . Dans les cellules leucémiques, il y a formation d'un ARN messager hybride qui code une protéine de fusion entre un locus situé sur le chromosome 15 et appelé *myl* ou, depuis peu *PML* et le gène codant le  $RAR\alpha$  (figure 29 ; De Thé et al., 1990a ; Pandolfi et al., 1991).

## PROTEINE PML NORMALE



## PROTEINE DE FUSION PML-RAR ALPHA.



## RAR ALPHA NORMAL

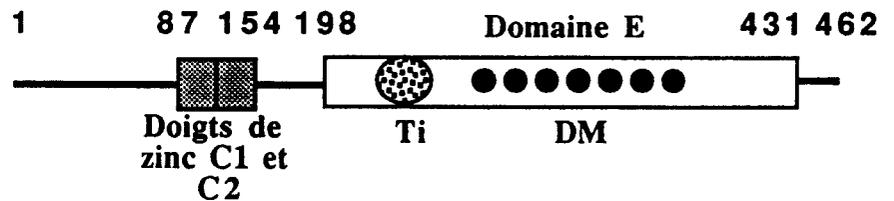


Figure 29 : La translocation t(15;17) présente dans les leucémies aigües à promyélocyte entraîne la fusion de deux facteurs de transcription, PML et RAR alpha (d'après de Thé et al., 1991 et Kakizuka et al., 1991).

L'implication d'un RAR dans un processus leucémogène n'est pas très étonnante en soi, vu l'action de l'acide rétinoïque qui contrôle la différenciation et la prolifération de nombreuses cellules, dont par exemple la lignée myéloïde HL60 (Breitman et al., 1980). On connaissait déjà un RAR, le RAR $\beta$ , impliqué dans la genèse d'hépatocarcinomes expérimentaux chez la marmotte, suite à l'intégration du virus de l'hépatite B en amont du gène, ce qui résulte en la synthèse d'une protéine de fusion (Dejean et al., 1986 ; De Thé et al., 1987).

La protéine de fusion *PML-RAR $\alpha$*  a un poids moléculaire de 103 kDa et contient les 3/4 N-terminaux de la protéine *PML* normale, puis les domaines B, C, D et E de RAR $\alpha$  (figure 29). Donc, dans la partie RAR $\alpha$ , seul manque le domaine A situé tout à fait en N-terminal. Par analogie avec les récepteurs des stéroïdes, on a émis l'hypothèse que ce domaine A contient un domaine transactivateur en plus du domaine E (Pandolfi et al., 1991 ; Beato, 1989). Evidemment, la protéine de fusion est sous la dépendance du promoteur de *PML* pour son expression. La première caractérisation fonctionnelle de cette protéine a montré qu'elle est capable de transactiver un promoteur hétérologue contenant un TRE classique, ce qui suggère qu'elle est encore capable de se fixer spécifiquement à l'ADN. De même, elle fixe toujours l'acide rétinoïque avec une bonne affinité (Pandolfi et al., 1991). Il faut noter que chez la souris la région A des RAR $\alpha$  est variable d'une isoforme à l'autre puisque le groupe de Pierre Chambon a montré que de nombreux épissages alternatifs peuvent modifier les exons situés en 5' (Leroy et al., 1991). On ignore si de telles isoformes existent chez l'homme, mais si tel est le cas on peut attendre avec une certaine curiosité de savoir comment elles se comportent lors de la translocation t(15;17).

Quels sont les mécanismes moléculaires qui conduisent une protéine de fusion *PML-RAR $\alpha$*  à être leucémogène ? La transactivation par *PML-RAR $\alpha$*  d'un promoteur contenant un RARE est beaucoup plus faible que celle médiée par le RAR $\alpha$  normal (elle est identique à celle d'un RAR $\alpha$  mutant ne contenant pas le domaine A, ce qui confirme bien l'importance de l'absence de ce domaine dans la protéine de fusion). Si on mélange le RAR $\alpha$  normal et la protéine de fusion *PML-RAR $\alpha$* , on s'aperçoit que cette dernière agit comme un régulateur négatif dominant de RAR $\alpha$  (Kakizuka et al., 1991 ; De Thé et al., 1991). On peut imaginer que les deux protéines se fixent sur les mêmes sites sur l'ADN et que la présence de *PML-RAR $\alpha$*  diminue l'activité normale du RAR $\alpha$  (Cleary, 1991).

Il existe une autre hypothèse qui d'ailleurs n'exclut pas la précédente. L'analyse attentive de la séquence du gène *PML* normal indique en effet que ce gène présente des homologies avec des facteurs de transcription (Kakizuka et al., 1991 ; De Thé et al., 1991 ; Kastner et al., 1992). Si *PML* est un facteur de transcription exprimé dans les promyélocytes, on peut imaginer qu'il a un rôle à jouer dans leur différenciation normale. Si la protéine *PML-RAR $\alpha$*  se fixe également au site de fixation de *PML* normal, on peut imaginer qu'elle bloque un programme de différenciation des cellules myéloïdes induit par *PML*.

c) Un traitement à base d'acide rétinoïque ?

La découverte de l'implication du gène  $RAR\alpha$  dans la translocation t(15;17) est particulièrement frappante parce qu'un groupe chinois a récemment montré que le traitement des APL par l'acide rétinoïque entraîne la différenciation des cellules promyélocytiques et la rémission pendant environ 8 mois des patients atteints d'APL (Huang et al., 1988 ; Wang et al., 1990 ; revues dans Clarkson, 1991 et Degos et al., 1991). Cette observation a été reproduite aux Etats Unis et en France par le groupe de Christine Chomienne et Laurent Degos qui s'attachent à coupler le traitement par l'acide rétinoïque et la chimiothérapie "classique" pour guérir les patients atteints d'APL (Castaing et al., 1990 ; Chomienne et al., 1990 ; Warrell et al., 1991). Evidemment, on se demande s'il existe une relation entre cette thérapeutique et la translocation affectant les gènes  $RAR\alpha$  et *PML*...

**TABEAU 5**

**Liste non exhaustive de quelques protéines fixant les hormones thyroïdiennes**

<b>Protéine</b>	<b>PM</b>	<b>Kd</b>	<b>Fonctions possibles Références</b>
TBG	36	T4 : 0,05nM T3 : 0,5nM	Transport dans le sang (Lissitzky, 1991)
TBPA	57	T4 : 1nM T3 : 10nM	Transport dans le sang (Lissitzky, 1991)
Albumine	60	T4 : 10 $\mu$ M T3 : 100 $\mu$ M	Transport dans le sang (Lissitzky, 1991)
Déiodinase de type 1	14	T4 : 2 $\mu$ M	Conversion de la T4 en T3 (Berry et al., 1990 ; 1991)
Déiodinase de type 2	?	T4 : 2nM	Conversion de la T4 en T3 (Berry et al., 1990)
TR $\alpha$	45	T4 : 0,5nM T3 : 0,03nM	Facteur de transcription ligand-dépendant (Sap et al., 1986 ; Thompson et al., 1987)
TR $\beta$	45	T4 : 0,5nM T3 : 0,05nM	Facteur de transcription ligand-dépendant (Weinberger et al., 1986)
Pyruvate kinase	58	T4 : 100nM T3 : 30nM	Protéine multifonctionnelle cytosolique fixant la T3 et la T4 (Kato et al., 1989)
Disulfite isomérase	57	?	Protéine multifonctionnelle cytosolique fixant la T3 et la T4 (Geetha-Habib et al., 1988)

## LES HORMONES THYROIDIENNES ET LEURS RECEPTEURS

Pour terminer notre panorama des récepteurs nucléaires nous souhaitons nous pencher particulièrement sur l'un d'eux. Bien sûr, nous avons choisi le, ou plutôt les récepteurs des hormones thyroïdiennes puisque nous avons étudié les gènes qui codent ces récepteurs.

### A) LES HORMONES THYROIDIENNES ET LEURS EFFETS.

Les hormones thyroïdiennes sont sécrétées par la glande thyroïde dès le troisième mois de la vie foetale. La synthèse de ces hormones est un processus complexe régulé de façon multifactorielle, où intervient la quantité d'iode ingérée par l'individu (Lissitzky, 1990). Dans la thyroïde, les cellules sécrétrices organisées en follicules sécrètent dans la lumière de ce follicule une protéine de haut poids moléculaire, la thyroglobuline. Cette protéine va être iodée sur les résidus tyrosine qui seront couplés pour former des dipeptides. La thyroglobuline iodée est alors recaptée par les cellules folliculaires et, au sein des cellules, elle est soumise à une digestion protéolytique précise qui va libérer les résidus hormonaux. L'efficacité de ce procédé est mauvaise puisque pour une molécule de thyroglobuline, on a une synthèse d'une molécule de T3 et de trois molécules de T4, qui seront libérées dans la circulation sanguine (Lissitzky, 1990).

Les hormones thyroïdiennes n'existent pratiquement pas sous forme libre dans le sang, mais sont associées à des protéines transporteuses de 3 types qui ont toutes une meilleure affinité pour la T4 que pour la T3 : la TBG (pour "Thyroxine Binding Protein") qui fixe 75 % de la T4 sérique, représente la forme majeure de transport des hormones thyroïdiennes et la TBPA ("Thyroid Binding PréAlbumine") qui fixe 15 % de la T4 circulante (tableau 5). Pour la petite histoire notons que la TBPA se complexe aussi à une protéine de 21 kD, la RBP, qui fixe ... le rétinol, précurseur de l'acide rétinoïque. On estime que seulement 0,025 % de T4 et 0,5 % de T3 sont libres dans le plasma.

Pendant longtemps on pensait que les hormones thyroïdiennes rentraient dans les cellules cibles par un mécanisme passif de passage à travers la bi-couche lipidique hydrophobe de la membrane plasmique. En fait, on a récemment suggéré qu'il existe un mécanisme de transport qui pourrait être le système T, qui transporte le tryptophane. Dans la cellule, les hormones thyroïdiennes, comme l'acide rétinoïque, peuvent se fixer à des protéines cytoplasmiques (voir tableau 6).

La seule forme active d'hormone thyroïdienne est la T3. Or, nous avons vu que la thyroïde sécrète surtout de la T4. Au niveau des tissus cibles, il existe donc des enzymes spécifiques capables de transformer la T4 en T3, c'est à dire de convertir une forme inactive en une forme active qui peut éventuellement être renvoyée dans le sang. On estime que 80 % de la T3

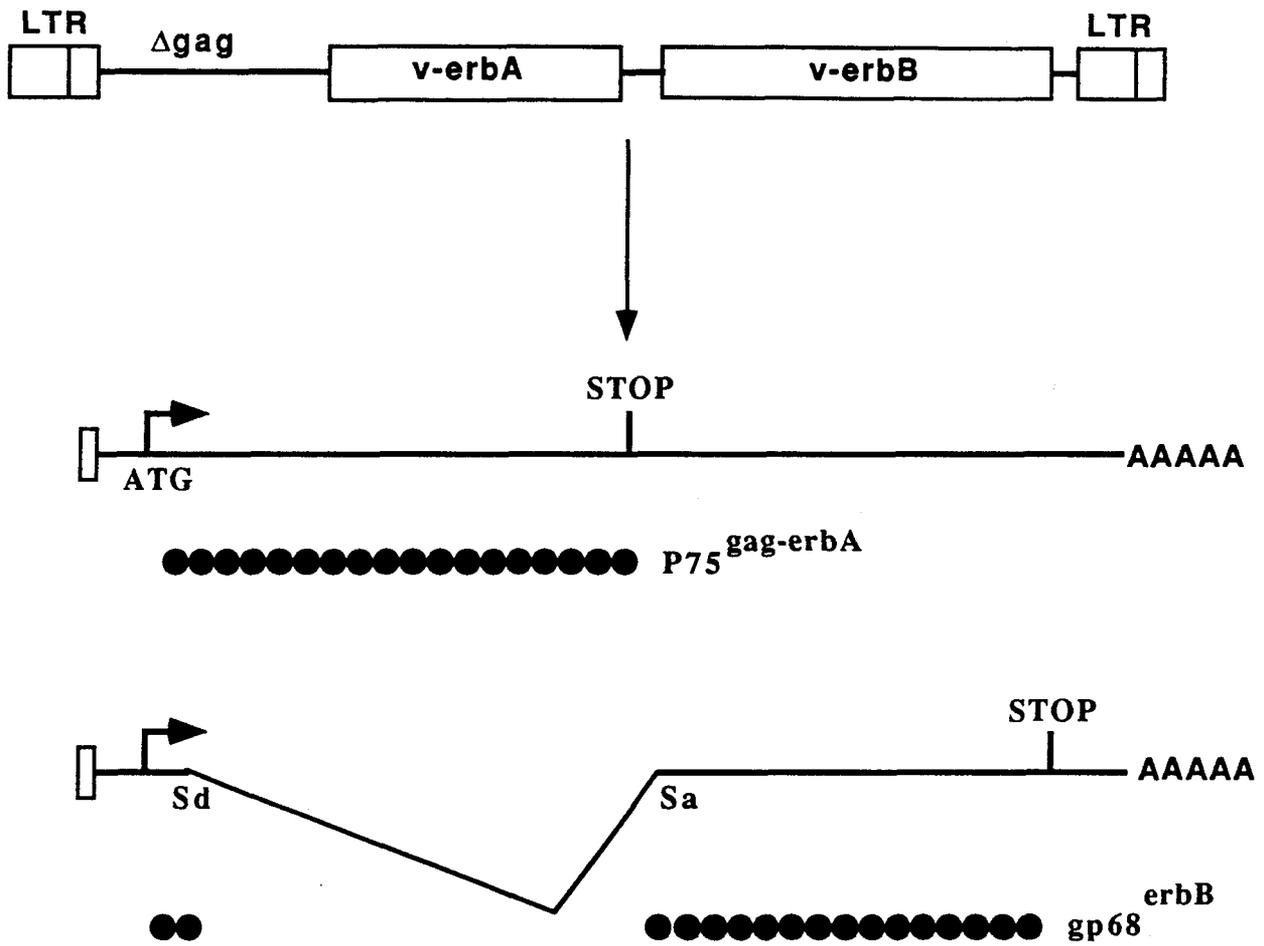


Figure 30 : Structure du rétrovirus aviaire AEV. Deux ARN différents sont produits à partir du provirus intégré dans le génome de la cellule hôte. Sd et Sa : sites accepteurs et donneurs d'épissage.

circulante ne provient pas de la thyroïde, mais de la conversion de T4 par des déiodinases cellulaires (Lissitzky, 1990 ; Berry et al., 1991; Silva et Larsen, 1983 ; Houstek et al., 1990 ; Reiter et al., 1990).

## B) LE RETROVIRUS AEV ET LA DECOUVERTE DES RECEPTEURS DES HORMONES THYROIDIENNES.

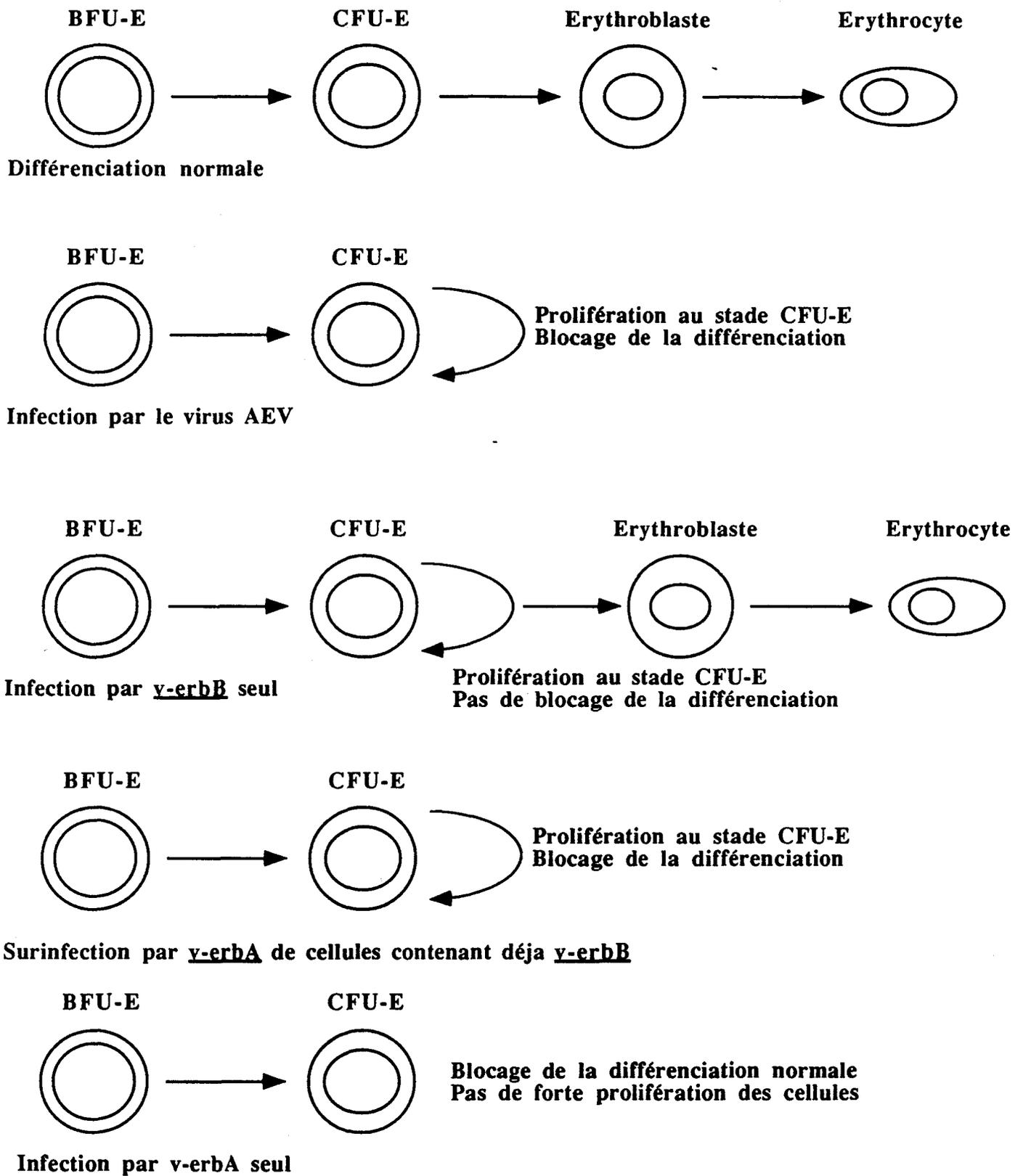
Dès 1972, c'est à dire 20 ans après la découverte de la T3, la preuve de l'existence d'un récepteur spécifique à localisation nucléaire des hormones thyroïdiennes était faite (voir revue de Oppenheimer, 1979). Cependant, tous les efforts pour purifier ce récepteur sont restés vains (Ichikawa et Degroot, 1987) et le récepteur a été identifié presque "par hasard" en étudiant l'un des oncogènes porté par un rétrovirus leucémogène aviaire, le virus AEV ("Avian Erythroblastosis Virus"). Un petit détour s'impose donc pour présenter ce nouvel intervenant.

### 1) AEV : le virus de l'érythroblastose aviaire.

Le rétrovirus AEV contient deux oncogènes v-erbA et v-erbB (Roussel et al., 1979; Sheiness et al., 1981 ; Saule et al., 1981 revue dans Beug et Vennström, 1991). Ces deux oncogènes coopèrent pour l'induction de l'effet pathogène du virus AEV, qui est responsable d'érythroblastoses et de sarcomes chez le poulet. Le gène v-erbA s'exprime sous forme d'une protéine de fusion avec les déterminants du gène de structure *gag* du rétrovirus pour former une protéine de fusion de 75 kDa, à localisation nucléaire (figure 30). Cependant, environ 40 % des protéines P75<sub>gag-erbA</sub> sont dans le cytoplasme et une notable proportion d'entre elles sont glycosylées (Privalsky, 1990). v-erbB est exprimé sous forme d'une protéine membranaire glycosylée.

C'est v-erbB qui représente l'oncogène transformant le plus fort de AEV. En effet v-erbB seul est capable de transformer des fibroblastes et des précurseurs hématopoïétiques *in vivo* et *in vitro* (Frykberg et al., 1983 ; Yamamoto et al., 1983). On a découvert, en 1984, que v-erbB code une version tronquée et activée du récepteur de l'EGF (Beug et al., 1982 ; Khazaie et al., 1988 ; revue dans Beug et Vennström, 1991 ).

L'oncogène v-erbA a un effet beaucoup plus modeste. Seul, il est incapable de causer des tumeurs chez des poulets infectés (Frykberg et al., 1983 ; Gandrillon et al., 1989). Cependant, le groupe de Jacques Samarut à Lyon a montré que la prolifération des fibroblastes embryonnaires de poulet contenant uniquement v-erbA est stimulée *in vitro*. Ces cellules sont capables de former des tumeurs lorsqu'on les dépose sur des membranes chorioallantoïdes d'oeufs (Gandrillon et al., 1987). L'effet de v-erbA seul, sur des cellules érythroïdes, est assez discuté, puisque certains auteurs observent un effet de v-erbA sur la croissance et la différenciation de ces cellules, tandis



**Figure 31 : La coopération entre v-erbA et v-erbB pour la transformation des érythroblastes aviaires**

que d'autres ne l'observent pas (Gandrillon et al., 1989 ; Schröder et al., 1990 ; Frykberg et al., 1983). Par contre, il est clair qu'en association avec v-erbB ou d'autres oncogènes transformants comme v-src ou v-Ha-ras, v-erbA est capable de bloquer la différenciation des précurseurs érythroïdes. En effet, l'équipe de Thomas Graf a montré que des cellules érythroïdes transformées par v-erbB seul sont encore capables de se différencier en érythrocytes. De plus, les érythroblastes transformés par v-erbB poussent mal en culture et seulement dans des conditions de pH et de force ionique bien précises (Kahn et al., 1986 ; Damm et al., 1987). Si on introduit v-erbA dans de telles cellules, on observe deux phénomènes (figures 31 et 32) : (i) un blocage de la différenciation de ces cellules à un stade précoce et (ii) une croissance plus rapide de ces cellules qui ont des exigences de pH et de force ionique beaucoup moins précises que les cellules contenant v-erbB seul. En résumé, on observe bien une claire coopération entre ces deux oncogènes viraux pour obtenir le phénotype observé *in vivo*. Si on schématise la différenciation érythroïde en cinq étapes (figures 31 et 32) on peut dire que AEV infecte les cellules au stade BFU-E et qu'il bloque ces cellules dans leur différenciation au stade CFU-E (Gazzolo et al., 1980 ; Beug et al., 1979 ; 1982 ; Graf et al., 1981 ; Hayman et al., 1982).

D'autres effets de v-erbA ont été observés. Par exemple, en association avec l'oncogène myc, v-erbA semble capable d'induire une transdifférenciation en nodules cartilagineux, à partir de cellules musculaires cardiaques transformées par l'effet de myc (Saule et al., 1987 ; Bachnou et al., 1991). Il semble aussi que v-erbA ait chez la caille des cellules cibles différentes de celles du poulet (Moscovici et al., 1989). Enfin, certains auteurs ont montré que v-erbA est capable de coopérer avec v-erbB, mais aussi avec c-erbB, qui semble avoir un rôle important dans la différenciation des cellules érythroïdes précoces (Khazaie et al., 1988 ; 1991 ; Pain et al., 1991).

## 2) Les récepteurs des hormones thyroïdiennes.

### a) Découverte des deux gènes c-erbA.

Après la découverte de v-erbA et à l'aide de la séquence complète de ce gène déterminée dès 1984, de nombreuses équipes se sont penchées sur son homologue cellulaire : c-erbA (Debuire et al., 1984 ; Mathieu-Mahul et al., 1985 ; Zarahoui et Cuny, 1987 ; Brauch et al., 1990 ; Ganly et Rabbitts, 1991 ; Sakurai et al., 1990). Très vite, il est apparu qu'il y avait dans le génome des vertébrés plusieurs gènes capables d'hybrider avec une sonde v-erbA (Jansson et al., 1983). C'est en étudiant le gène cellulaire le plus homologue à v-erbA, appelé c-erbA-1, que les groupes de Bjorn Vennström et de Ron Evans ont montré de façon indépendante que v-erbA représentait une version modifiée du récepteur des hormones thyroïdiennes (Sap et al., 1986 ; Weinberger et al., 1986). On sait maintenant que les deux groupes ont isolé deux gènes différents ; l'un (c-erbA-1) est situé, chez l'homme, sur le chromosome 17 en q11.2 et l'autre (c-erbA-2) sur le chromosome 3 en p24 (Spurr et al., 1984 ; Dayton et al., 1984 ; Sheer et al., 1985 ; Weinberger et al., 1986 ; Drabkin et al., 1988 ; Douglas et al., 1991). A la lumière des données biochimiques

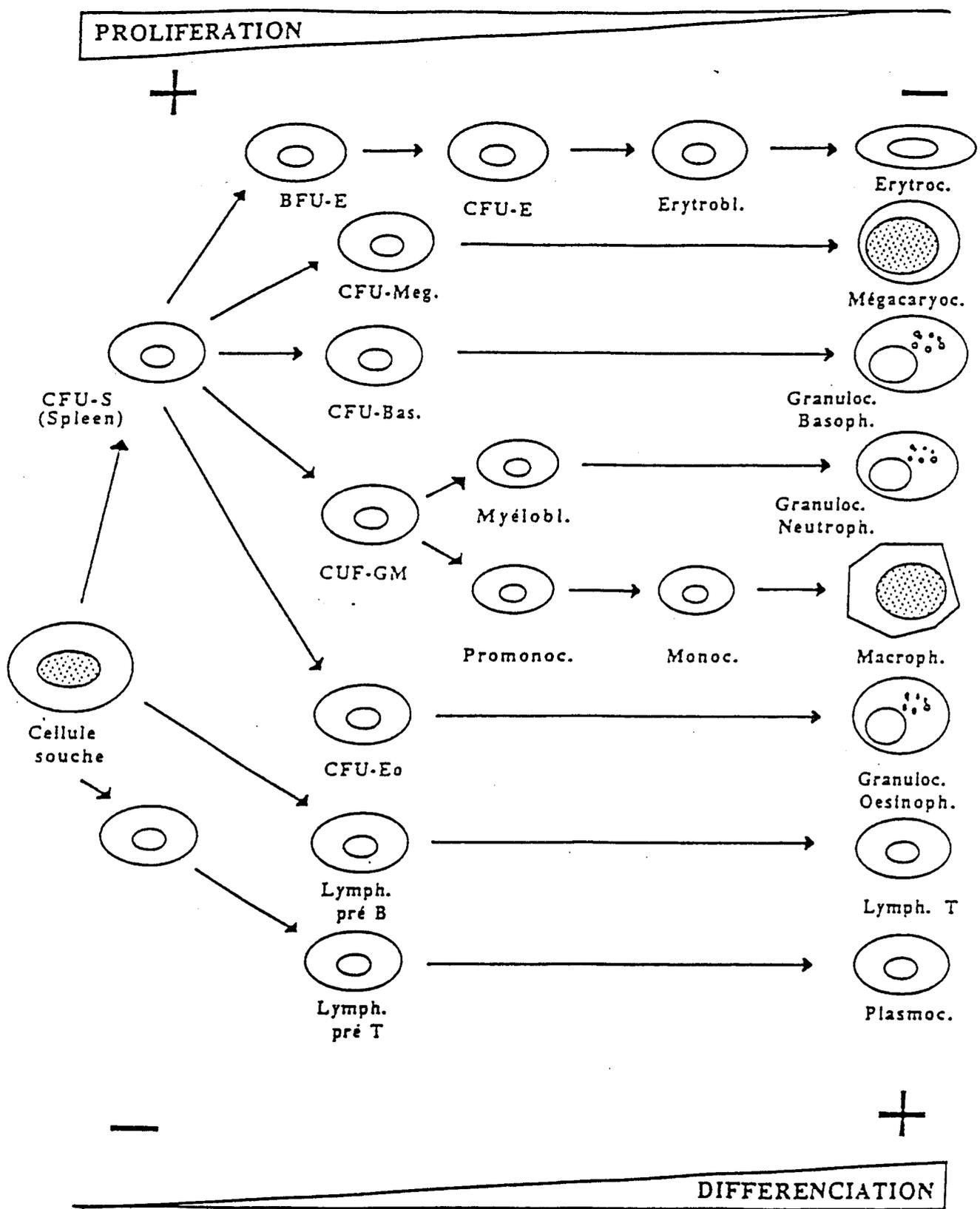


Figure 32 : Schéma des voies de différenciation des cellules hématopoïétiques.

anciennes, l'existence de plusieurs formes de récepteurs n'était pas étonnante car on avait déjà noté qu'il devait y avoir plusieurs types de molécules douées de caractéristiques différentes capables de fixer les hormones thyroïdiennes (Latham et al., 1976 ; Casanova et al., 1984). Les années passant, en ce qui concerne les multiples formes de récepteur des hormones thyroïdiennes, personne n'a été déçu...!

#### b) Plusieurs isoformes de c-erbA-1.

Dès l'isolement des premiers ADNc de c-erbA-1, plusieurs équipes ont cherché à savoir s'il existait d'autres gènes ou d'autres types de cDNAs codant des récepteurs des hormones thyroïdiennes (figure 33). C'est ainsi qu'entre 1987 et 1988 de nouveaux types de récepteurs codés par le gène c-erbA-1 ont été décrits (Thompson et al., 1987 ; Benbrook et Pfahl, 1987 ; Pfahl et Benbrook, 1987 ; Mitsushashi et al., 1988a ; 1988b ; Nakai et al., 1988a ; 1988b ; Prost et al., 1988 ; Murray et al., 1988 ; Koenig et al., 1988 ; 1989 ; Lazar et al., 1988 ; 1989b ; Izumo et Mahdavi, 1988 ; Schueler et al., 1990 ; Masuda et al., 1990). Certains de ces récepteurs étaient capables de fixer les hormones thyroïdiennes dans certains laboratoires (Benbrook et Pfahl, 1987) mais pas dans d'autres (Lazar et al., 1988 ; 1989b Koenig et al., 1989). Une période d'intense confusion s'en est suivie pour finalement déboucher sur une situation assez simple !

Il existe deux formes majeures alternativement épissées, codées par le gène c-erbA-1. Une de ces formes code, via un ARN de 5 kb, une protéine de 410 acides aminés capable de fixer les hormones thyroïdiennes et d'activer la transcription d'un gène cible (figure 33). Ce produit, un récepteur des hormones thyroïdiennes classique est appelé TR $\alpha$ 1. L'autre forme est différente dans sa partie 3' et est codée par un ARN de 2,7 kb, chez l'homme. Il s'agit d'une protéine de 490 acides aminés, TR $\alpha$ 2, incapable de fixer les hormones thyroïdiennes et incapable d'activer la transcription. De plus, dans des expériences de compétition, il a été démontré que cette forme est capable d'inhiber la transactivation dépendante de l'hormone médiée par la forme TR $\alpha$ 1 (Izumo et Mahdavi, 1988 ; Koenig et al., 1989). Cette situation est conservée chez tous les mammifères étudiés. Par contre, il semble, d'après des arguments indirects obtenus par Bjorn Vennström et son équipe, que la forme TR $\alpha$ 2 n'existe pas chez le poulet (Forrest et al., 1990b). D'autres formes alternativement épissées ont été décrites, notamment dans la partie C-terminale (Mitsushashi et al., 1988 ; Mitsushashi et Nikodem, 1989). Dans une étude systématique par protection à la RNase, de nombreuses formes résultant peut-être d'épissages alternatifs en 5' et en 3' ont été décrites, mais les clones cDNA correspondants n'ont pas été isolés (Izumo et al., 1990).

#### c) Plusieurs isoformes de c-erbA-2.

Dans le cas du gène c-erbA-2, il existe également différentes isoformes possibles (figure 33). La plus intéressante est spécifiquement exprimée dans l'hypophyse, lieu de sécrétion de la TSH qui régule la synthèse des hormones thyroïdiennes. Elle est différente du récepteur de type  $\beta$ 1 classique par son domaine A/B responsable de l'activation de la transcription (Hodin et al.,

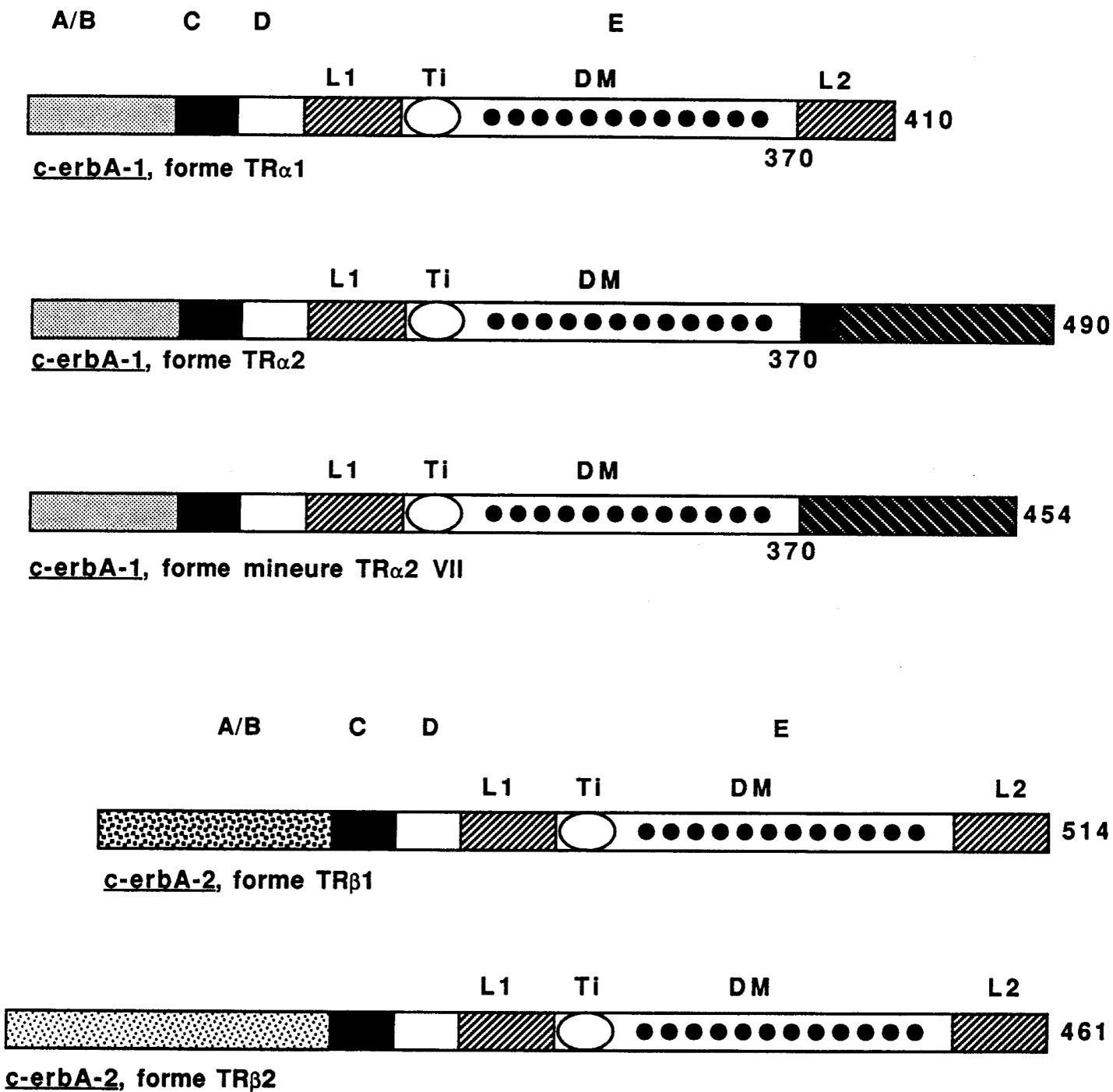


Figure 33 : Les différentes isoformes des récepteurs des hormones thyroïdiennes de mammifères.

1989). Récemment, cette forme TRB2 a été retrouvée chez le poulet et une troisième espèce a été découverte exprimée dans la rétine de cet espèce (Forrest et al., 1991; Sjöberg et al., 1992).

Il existe d'autres formes d'ARN, différentes en 5', codées par le gène c-erbA-2. Ainsi, chez le xénope, qui contient deux gènes c-erbA-1 et deux gènes c-erbA-2, il y a 9 formes différentes en 5' et chaque gène c-erbA-2 code au moins deux formes différentes en N-terminal. Cette situation ressemble étrangement à celle constatée par le groupe de Chambon pour les trois gènes de récepteurs de l'acide rétinoïque chez la souris (Kastner et al., 1990 ; Leroy et al., 1991 ; Zelent et al., 1991). On ignore si cette observation faite chez le xénope peut être généralisée chez les mammifères.

### 3) Expression des récepteurs des hormones thyroïdiennes.

#### a) Expression des gènes c-erbA.

Les premières études menées sur l'expression de c-erbA, réalisées avant le clonage des deux gènes c-erbA-1 et c-erbA-2, ont insisté sur le fait que c-erbA était exprimé dans de nombreux types cellulaires ou organes à des taux différents, ce qui laissait présager une régulation assez fine de cette expression (Coll et al., 1983 ; Hentzen et al., 1987). Dans les cellules érythroïdes du poulet, cible de l'oncogène v-erbA, le taux de transcrits c-erbA varie selon les stades de différenciation : il y a une faible expression dans les érythroblastes précoces (BFU-E et CFU-E) et une expression beaucoup plus forte dans les érythroblastes plus tardifs et les érythrocytes. Ceci est bien corrélé avec l'action de v-erbA, qui bloque la différenciation de ces cellules au moment où l'action de la T3 semble requise (Hentzen et al., 1987).

Les deux gènes c-erbA sont exprimés de manières différentes. En effet, en simplifiant, on peut dire que c-erbA-1 est exprimé faiblement dans la plupart des organes avec une prédominance dans le coeur, le cerveau et les cellules érythroïdes, alors que c-erbA-2 est exprimé un peu plus fortement et d'une façon plus variable d'un organe à l'autre (Mitsushashi et Nikodem, 1989 ; Sakurai et al., 1989). Cette remarque est valable chez les mammifères, et chez le poulet (Forrest et al., 1990) où, au cours du développement, le taux d'expression de c-erbA-1 reste relativement constant. L'expression de c-erbA-2 est par contre beaucoup plus finement régulée selon les organes et au cours du temps. Par exemple, l'expression de c-erbA-2 augmente de façon générale de 30 fois au moment de l'éclosion. Si on s'intéresse aux différentes isoformes de ces récepteurs, on note que les formes  $\alpha 1$  et  $\alpha 2$  de c-erbA-1 sont exprimées de manière à peu près identique :  $\alpha 2$  est généralement plus exprimé que  $\alpha 1$ . Le rapport entre les deux formes ( $\alpha 1/\alpha 2$ ) varie légèrement d'un organe à l'autre : il est de 1/1 dans la rate et de 1/6 dans le cerveau (Mitsushashi et Nikodem, 1989).

Il n'y a pas concordance parfaite entre ces expériences d'expression des ARN et des études antérieures réalisées avec de la T3 marquée pour mesurer la quantité de récepteurs dans les différents organes (Freake et al., 1988 ; Santos et al., 1988 ; Sakurai et al., 1989). Ainsi chez

le rat adulte, dans le cerveau, l'un des rares organes apparemment insensible aux hormones thyroïdiennes on note la présence d'un haut taux de transcrits c-erbA. On peut sans doute réconcilier ces deux observations si on se souvient que la forme  $\alpha 2$  de c-erbA-1, qui est très fortement exprimée dans le cerveau adulte, ne fixe pas les hormones thyroïdiennes et inhibe l'action des récepteurs normaux (Mitsubishi et Nikodem, 1989 ; Lazar et Chin, 1988).

Le cerveau foetal a été très étudié, surtout par hybridation *in situ*, mais également par immunofluorescence avec des anticorps spécifiques des formes  $\alpha 2$  et  $\beta 1$  (Bradley et al., 1989 ; Forrest et al., 1991 ; Strait et al., 1991). La règle générale (c-erbA-1 plutôt ubiquiste, c-erbA-2 plus spécifique et variable) est relativement bien confirmée au cours du développement de cet organe. Chez les oiseaux, il semble que le gène c-erbA-1 soit exprimé très tôt, à des moments où les hormones thyroïdiennes ne sont pas encore produites. Par contre, l'expression de c-erbA-2 démarre au début de la période hormono-sensible (Forrest et al., 1991 ; Strait et al., 1991).

#### b) Régulation de l'expression des gènes c-erbA par la T3.

Les hormones thyroïdiennes sont-elles capables de contrôler l'expression des gènes de leurs récepteurs ? Les données biochimiques étaient très contradictoires à ce sujet (Samuels et al., 1977 ; Hodin et al., 1990). En règle générale, dans des expériences menées sur des rats hypothyroïdiens ou hyperthyroïdiens, il semble que les hormones thyroïdiennes diminuent l'expression des ARN c-erbA (Mitsubishi et Nikodem, 1989 ; Lazar et Chin, 1988 ; Hodin et al., 1990). Cela dit, il existe des exceptions notables comme l'ont remarqué les chercheurs de l'équipe de William Chin (Hodin et al., 1990). Ainsi, dans l'hypophyse, les formes  $\beta 1$  et  $\beta 2$  de c-erbA-2 sont régulées de façon différentielle par la T3 : l'expression de la forme  $\beta 1$  augmente de 3,5 fois, tandis que  $\beta 2$ , qui est une forme spécifique de l'hypophyse (Hodin et al., 1990), voit son expression diminuée de 50 %. Les transcrits c-erbA-1 obéissent à la règle de régulation négative de l'expression par la T3 sauf dans le cerveau où l'expression demeure constante, qu'il y ait ou non de la T3.

#### 4) **Les domaines fonctionnels des récepteurs des hormones thyroïdiennes.**

Comme les récepteurs des hormones thyroïdiennes ont la même structure en domaines fonctionnels que les récepteurs des stéroïdes, l'identification d'un domaine particulier a souvent été faite par comparaison avec ceux-ci.

La comparaison entre c-erbA et v-erbA offrait un excellent moyen d'étudier les zones du domaine E importantes pour la fixation du ligand, puisque le produit de c-erbA-1 chez le poulet fixe la T3, alors que celui de v-erbA ne la fixe pas (Sap et al., 1986). Une telle étude a permis de montrer que l'extrémité C-terminale du récepteur était impliquée dans la fixation de l'hormone. Si l'on remplace la partie C-terminale de v-erbA par celle de c-erbA, on obtient une molécule capable de fixer l'hormone (Munoz et al., 1988). On sait maintenant que les zones de fixation du ligand

sont situées à chaque extrémité du domaine E, bien que des régions situées plus vers le centre de ce domaine, comme la région Ti, semblent aussi importantes (O'Donnel et Koenig, 1990 ; Forman et Samuels, 1990a ; 1990b ; Horowitz et al., 1989).

L'équipe de Ron Evans s'est attachée à étudier l'activation transcriptionnelle induite par les différents produits des gènes c-erbA (Thompson et Evans, 1989). Il semblerait que le produit  $\alpha 1$  de c-erbA-1 ait une activité transcriptionnelle trois fois supérieure à celle du produit  $\beta 1$  de c-erbA-2. De plus, l'activation maximale de c-erbA-2 est obtenue pour une concentration en T3 deux fois plus faible que celle nécessaire à une activation maximale de c-erbA-1. On peut donc imaginer que l'un des récepteurs (codé par c-erbA-2) réponde faiblement à des petites doses d'hormones tandis que l'autre (c-erbA-1) réponde plus fortement, mais à des doses plus fortes. Ce modèle ressemble à celui que nous avons vu pour les récepteurs des glucocorticoïdes et des minéralocorticoïdes, dans le cerveau. Le problème est bien sûr de savoir si un tel résultat est valable *in vivo*.

Où sont situées les régions transactivatrices dans les récepteurs des hormones thyroïdiennes ? Par analogie avec les récepteurs des stéroïdes, on a souvent baptisé le domaine A/B "domaine transactivateur", mais cela est inexact. La délétion de ce domaine pour c-erbA-1 ou c-erbA-2 n'a presque aucun impact sur l'activation transcriptionnelle (Thompson et Evans, 1989). Cette fonction se trouve en fait plutôt dans la partie C-terminale de ces récepteurs.

### 5) Les complexes avec les récepteurs de l'acide rétinoïque.

Nous avons déjà noté que les récepteurs des hormones thyroïdiennes et ceux de l'acide rétinoïque se fixent à des sites proches, sinon identiques sur l'ADN. On peut résumer rapidement la situation en disant qu'un TRE peut être activé par les deux types de récepteurs, alors qu'un RARE ne sera activé que par un RAR (Glass et al., 1989 ; Umesono et al., 1988 ; 1991 ; De Thé et al., 1990b ; Näär et al., 1991).

Les récepteurs de l'acide rétinoïque et ceux des hormones thyroïdiennes peuvent faire des hétérodimères. Ainsi, un récepteur des hormones thyroïdiennes mutant à activité négative dominante, qui inhibe la transcription d'un récepteur des hormones thyroïdiennes sauvage, peut avoir le même type d'effet sur un récepteur de l'acide rétinoïque et, même sur un récepteur de la vitamine D (Forman et al., 1989 ; Forman et Samuels 1990a ; 1990b). Les hétérodimères des deux types de récepteurs (RAR et TR) ont été visualisés par liaison avec des agents pontants par Christopher Glass et Michael Rosenfeld. Il a été suggéré que dans certains cas les hétérodimères pouvaient avoir des effets transcriptionnels différents (voir la revue de Jacobs et Moxham, 1991 sur le concept général des récepteurs hybrides). Si on exprime à faible taux un RAR et un THR dans une même cellule, on observe une activation de la transcription en utilisant un promoteur lié à un TRE palindromique. Par contre, si on utilise un promoteur lié à un TRE imparfait (provenant par exemple du gène de la chaîne lourde de myosine cardiaque), on obtient une répression de la

transactivation (Glass et al., 1989). Ce type d'effet varie d'un promoteur à l'autre (Hudson et al., 1990 ; Moore et Brent, 1991).

Barry Marc Forman et Herbert Samuels, à New York, ont proposé un modèle structural pour expliquer la formation de tels dimères entre RAR et TR, mais aussi entre RARs ou entre TRs. Dans la partie centrale de ces récepteurs, il existe en effet des répétitions d'acides aminés hydrophobes qui font penser aux "leucine zipper" des facteurs de transcription comme fos et jun (Landschulz et al., 1988 ; Forman et Samuels, 1990a ; 1990b) et qui pourraient médier la dimérisation observée (figure 16). De nombreuses expériences de mutations confirment bien ce modèle (Glass et al., 1989 ; Forman et al., 1989 ; Hudson et al., 1990). Ceci a poussé certains auteurs à proposer que le domaine E, en participant à la spécification des dimérisations possibles, modulait l'activité de la fixation à l'ADN (Zhang et al., 1991a).

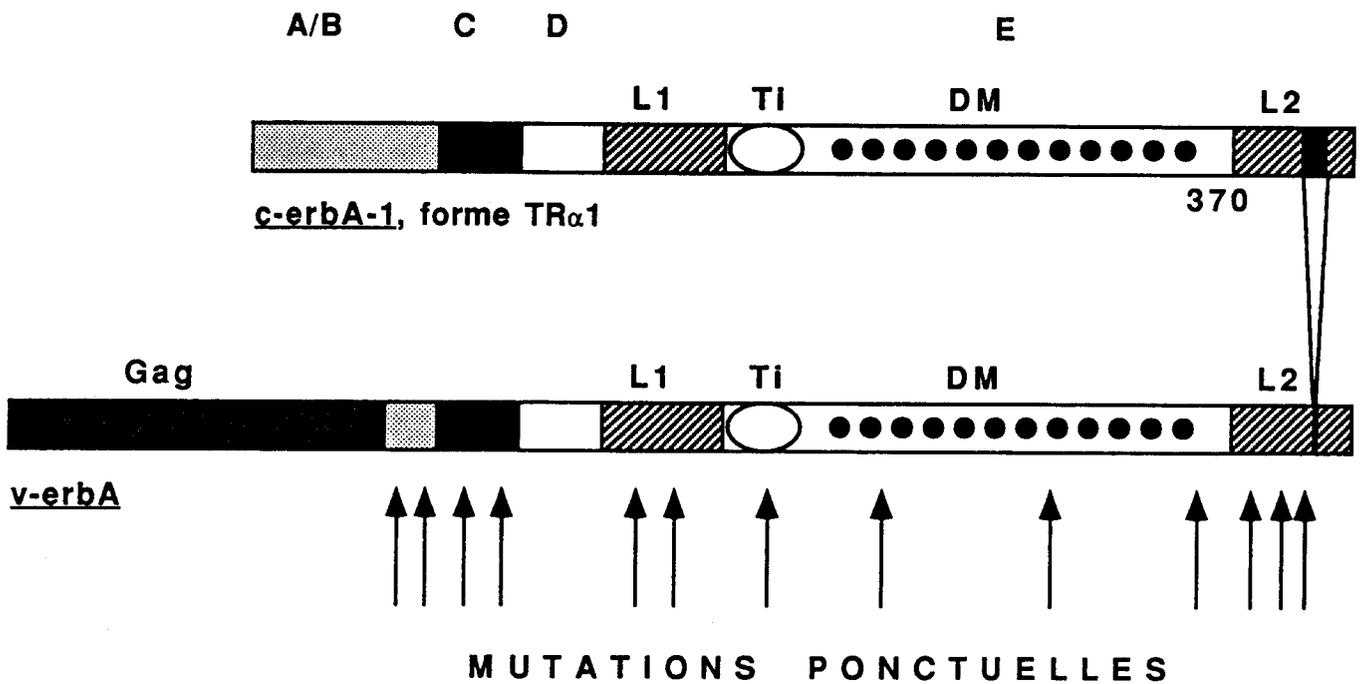
Incidemment, l'hétérodimérisation et ses antagonismes fonctionnels éventuellement mis en évidence ne se rencontrent pas uniquement entre récepteurs ayant des spécificités de ligand différentes. On a par exemple mis en évidence des effets antagonistes entre le RAR $\gamma$ 1 et d'autres RAR ( $\beta$ ,  $\gamma$ 2) sur la transcription du gène RAR $\beta$  qui contient un RARE. Tous les RARs testés activent ce promoteur, sauf RAR $\gamma$ 1 (Husmann et al., 1991). RAR $\gamma$ 1 et RAR $\gamma$ 2 ne diffèrent que par leur domaine N-terminal A qui a donc des effets transcriptionnels très différents (Kastner et al., 1990).

Dans une cellule, à chaque instant, la population des différents types de récepteurs présents (TR $\alpha$ 1, TR $\alpha$ 2, TR $\beta$ 1, TR $\beta$ 2, RAR $\alpha$ , RAR $\beta$ , RAR $\gamma$ , sans compter les nombreuses isoformes décrites des RAR) va déterminer, selon un système de combinatoire moléculaire, les types de complexes qui vont se fixer sur un promoteur donné. De plus, la découverte récente de l'hétérodimérisation de tous ces récepteurs avec les produits des gènes RXR ajoute de nouvelles possibilités d'interactions complexes (Yu et al., 1991). Alors, selon les promoteurs, les sites qu'ils contiennent, les facteurs qui y sont fixés etc.. la réponse transcriptionnelle pourra varier du tout au tout. La flexibilité du système est immense et rend bien compte de la diversité des réponses physiologiques observées...

### C) LES MECANISMES D'ACTION DE v-erbA.

#### 1) Le modèle classique des cellules érythroïdes.

Sachant comment agit c-erbA, c'est à dire comme un activateur transcriptionnel ligand-dépendant, que peut-on dire du mode d'action du produit de v-erbA, la P75gag-erbA? Le fait que cette protéine ne fixe pas la T3 a une conséquence très importante : en se fixant sur les mêmes sites d'ADN que les produits de c-erbA, elle agit comme un répresseur constitutif de l'expression des gènes régulés par les hormones thyroïdiennes (Sap et al., 1986 ; 1989; Munoz et al., 1988 ; Boucher et al., 1988 ; Privalsky et al., 1988 ; Damm et al., 1989). On sait en effet que les



Différences entre *c-erbA-1* et *v-erbA* : 1) fusion avec le gène viral gag

2) 13 mutations ponctuelles

3) délétion C-terminale de 9 acides aminés

(Mutations *c-erbA-1* de poulet/*v-erbA* : 24 Arg/His ; 44 Tyr/Cys ; 73 Gly/Ser ; 90 Lys/Thr ; 186 Lys/Arg ; 191 Pro/Leu ; 203Pro/Leu ; 233 Lys/Asn ; 342 Thr/Ser ; 363 Pro/Ser ; 370 Thr/Ala ; 378 Cys/Tyr ; 395 Phe/Ser).

Figure 34 : Les différences entre *c-erbA-1* et *v-erbA*.

récepteurs des hormones thyroïdiennes codés par c-erbA sont des répresseurs en absence de leur ligand (Damm et al., 1989 ; Graupner et al., 1989 ; Brent et al., 1989). Le modèle simple qui se dégage de ces observations est donc que le produit de v-erbA va entrer en compétition avec ceux des gènes c-erbA, pour l'occupation des TRE. Comme la P75<sup>gag-erbA</sup> est incapable de fixer la T3, elle va réprimer en permanence l'expression de ces gènes. On peut expliquer l'action de v-erbA sur le blocage de la différenciation érythroïde en imaginant que la T3 est importante pour la mise en route de certains gènes à des étapes clefs de ce processus de différenciation. Si v-erbA bloque ces gènes, il y a blocage de la différenciation.

Ce modèle est bien joli mais encore faut-il trouver des gènes dont l'expression est bloquée par v-erbA et qui sont *in vivo* régulés par les hormones thyroïdiennes. On a effectivement trouvé de tels gènes réprimés par v-erbA (Zenke et al., 1988). Ces gènes sont ceux qui codent "BAND 3", un transporteur d'anion des érythrocytes, l'anhydrase carbonique de type II (ACOII) qui est importante pour le métabolisme du CO<sub>2</sub> et l'acide δ-aminolévulinique synthétase (ALA-S), une enzyme impliquée dans la biosynthèse de l'hème (Zenke et al., 1988). Depuis il a été confirmé que les hormones thyroïdiennes activaient l'expression d'un de ces gènes (celui de l'ACOII) *in vivo*, que c-erbA et v-erbA étaient capables de se fixer au promoteur de ce gène et de réguler sa transcription (Pain et al., 1990 ; Disela et al., 1991). Le produit de c-erbA chez le poulet active la transcription de ce gène, tandis que celui de v-erbA la bloque. Si on introduit en même temps une quantité identique de v-erbA et de c-erbA on peut encore induire ce gène par la T3 ce qui montre qu'à quantité égale, c-erbA domine sur v-erbA (Disela et al., 1991). Mais, au sein d'un rétrovirus aviaire comme AEV, v-erbA s'exprime beaucoup plus fortement que le gène c-erbA-1 endogène exprimé normalement dans les érythroblastes (Hentzen et al., 1987).

Entre temps, nous l'avons vu, de nombreux auteurs ont montré que les récepteurs des hormones thyroïdiennes étaient capables de dimériser entre eux, mais aussi avec les récepteurs de l'acide rétinoïque (Glass et al., 1989 ; Forman et al., 1989). On pouvait donc imaginer que les protéines codées par v-erbA et c-erbA seraient capables de faire des dimères inactifs qui donc empêcheraient la mise en route des gènes de différenciation dans les cellules érythroïdes. Martin Zenke, Hartmund Beug et leur équipe, à Vienne, ont montré qu'effectivement, v-erbA pouvait bloquer l'action des produits de c-erbA sur des gènes de différenciation (Zenke et al., 1990).

Ces auteurs ont clairement mis en cause l'extrémité C-terminale de v-erbA dans son activité biologique. Si on compare v-erbA à son homologue cellulaire c-erbA-1, chez le poulet, on peut en effet noter un certain nombre de différences (figure 34; Debuire et al., 1984 ; Sap et al., 1986 ; Damm et al., 1987). Le travail de l'équipe de Zenke et Beug a permis de démontrer que c'est la délétion des 9 acides aminés C-terminaux qui est importante pour l'activité de répression de la transactivation de v-erbA. Cette délétion est située juste dans la deuxième moitié du domaine de fixation du ligand et on pouvait imaginer que l'absence de fixation de la T3 par la P75<sup>gag-erbA</sup> suffise à expliquer la répression de l'expression observée du gène de band 3 ou de l'ACOII. En fait, cette même équipe a réussi à séparer la fixation de l'hormone et la répression. On pense donc

que cette délétion enlève sans doute aussi un domaine transactivateur présent dans c-erbA (Zenke et al., 1990 ; Schroeder et al., 1990 ; Forrest et al., 1990a).

Un des éléments qui confirme très bien ce modèle est l'existence, au moins chez les mammifères, de l'isoforme TR $\alpha$ 2, chez qui l'extrémité C-terminale du récepteur normal est remplacée par un nouvel exon. Cette forme est précisément incapable de fixer la T3 et d'activer la transcription (Izumo et Mahdavi, 1988 ; Koenig et al., 1989). En toute première approximation, (pourtant très largement inexacte, nous allons le voir !) on pourrait dire que cette forme se comporte comme v-erbA. Evidemment, et heureusement, cette forme n'est pas oncogène, ce qui montre bien qu'on ne peut pas inconsidérément forcer l'analogie avec v-erbA.

On comprend mieux ce qui se passe dans les cellules érythroïdes depuis que l'on sait que la P75gag-erbA n'est pas capable de se complexer aux récepteurs des hormones thyroïdiennes. Ceci est dû à une mutation ponctuelle de v-erbA, située dans le domaine E, à la fin des répétitions en heptades découvertes par le groupe de Samuels et importantes pour la dimérisation des récepteurs (Forman et Samuels, 1990a ; 1990b ; Selmi et Samuels, 1991). Cette découverte signifie que dans le modèle de blocage de la différenciation des cellules érythroïdes, la P75gag-erbA ne formerait pas des hétérodimères inactifs avec les récepteurs normaux, mais entrerait en compétition avec eux pour l'occupation des TRE. Ceci expliquerait pourquoi v-erbA doit être surexprimé pour bloquer l'action des récepteurs des hormones thyroïdiennes (Disela et al., 1991).

## 2) Le rôle du domaine C.

Bien que le modèle présenté ci-dessus soit très cohérent et rende bien compte de la situation observée dans les érythroblastes transformés par v-erbA, un certain nombre d'auteurs a suspecté qu'en plus de la région C-terminale, d'autres régions de v-erbA étaient importantes pour son activité. Ainsi, Hubert de Verneuil et Daniel Metzger, à Strasbourg, en construisant des chimères entre v-erbA et le récepteur des oestrogènes, ont montré que la répression de la transcription induite par v-erbA était, au moins en partie, due à une mutation ponctuelle située dans le domaine C (De Verneuil et Metzger, 1990). De même, l'équipe de Martin Privalsky a suggéré que la spécificité de fixation à l'ADN de la P75gag-erbA était sans doute légèrement différente de celle du produit de c-erbA. Ceci implique que la séquence reconnue par les deux protéines n'est pas tout à fait la même, contrairement à ce que l'on pensait jusqu'à présent (Boucher et Privalsky, 1990; Bonde et al., 1990 ; 1991).

Il faut dire que les deux mutations présentes dans v-erbA, au niveau du domaine C, ont de quoi exciter l'imagination (figure 34). L'une d'entre elles est en effet le changement d'une glycine en sérine dans la boîte P qui dicte la reconnaissance spécifique des HRE des récepteurs nucléaires. L'autre mutation est un changement d'une lysine en thréonine, dans la boîte D située entre les deux premières cystéines du doigt C2 et qui est la deuxième région importante pour la reconnaissance spécifique des HRE. Il est tout à fait possible que ces changements modifient la

spécificité de fixation à l'ADN. Il faut remarquer que la mutation de la boîte P modifie v-erbA de façon à ce que cette boîte P devienne identique à celle de récepteurs orphelins, comme COUP ou RXR. Pour l'instant, bien qu'il soit clairement démontré que ce changement modifie légèrement la spécificité de fixation à l'ADN, on ignore encore ses implications physiologiques (Bonde et al., 1991). Mais on peut s'attendre à des surprises...

### 3) v-erbA, les levures et l'acide rétinoïque.

Tout récemment, des découvertes nous ont permis de rafraîchir agréablement nos idées sur le mode d'action de v-erbA qui, comme toujours, semble être plus complexe que ce que l'on pensait au départ.

Ainsi, Martin Privalsky, a récemment démontré qu'après introduction dans des levures, v-erbA se comporte comme un ... activateur de la transcription, dépendant de son ligand (Privalsky et al., 1990). Il semble bien que la P75gag-erbA ait conservé une faible capacité de fixer les hormones thyroïdiennes. Chez la levure, cette fixation suffit à entraîner une activation de la transcription. Ceci permet d'imaginer que, chez les vertébrés, la répression de la transactivation induite par v-erbA dépend d'un certain nombre de protéines qui n'existent pas chez la levure. Ce processus ne serait donc pas une propriété intrinsèque de v-erbA. De même Privalsky a observé que le produit de v-erbA était complexé à des protéines de choc thermique hsp90 dans les cellules de vertébrés (Privalsky, 1991). Ces protéines sont-elles responsables de l'extinction de l'activité de v-erbA dans ces cellules ? On sait que des analogues des hsp90 existent chez la levure. Peut-être sont-elles incapables de réprimer v-erbA ? Mystère.

En plus de la délétion C-terminale, du domaine de fixation à l'ADN, et du domaine de dimérisation, d'autres régions de v-erbA sont importantes pour son activité. C'est le cas du domaine N-terminal, où l'on trouve des sites de phosphorylation dont l'intégrité est requise pour l'action oncogène de v-erbA (Goldberg et al., 1988 ; Glineur et al., 1989 ; 1990).

Notre connaissance des mécanismes d'action de v-erbA s'est brusquement enrichie, à la suite de travaux récents qui montrent que v-erbA pourrait bloquer l'action des récepteurs de l'acide rétinoïque (figure 35). Ainsi, v-erbA interfère avec la transcription médiée par de nombreux autres récepteurs nucléaires, comme par exemple ceux de l'acide rétinoïque, ceux des oestrogènes mais pas ceux des glucocorticoïdes (Sharif et Privalsky, 1991). Pour cet effet, c'est le domaine de fixation à l'ADN qui serait le plus important, la partie C-terminale étant moins cruciale. Un mutant de v-erbA où la mutation de la boîte P a été "corrigée" n'est plus capable d'interférer avec la transactivation induite par les récepteurs de l'acide rétinoïque. Le groupe de Jacques Samarut, à Lyon, a récemment démontré qu'*in vivo* v-erbA était capable de supprimer l'effet inhibiteur de l'acide rétinoïque sur la croissance de fibroblastes de poulet (Desbois et al., 1991a). De plus, le même groupe a montré que v-erbA était capable d'empêcher l'inactivation du complexe AP-1 induite par les récepteurs de l'acide rétinoïque (Desbois et al., 1991b). Nous avons en effet noté

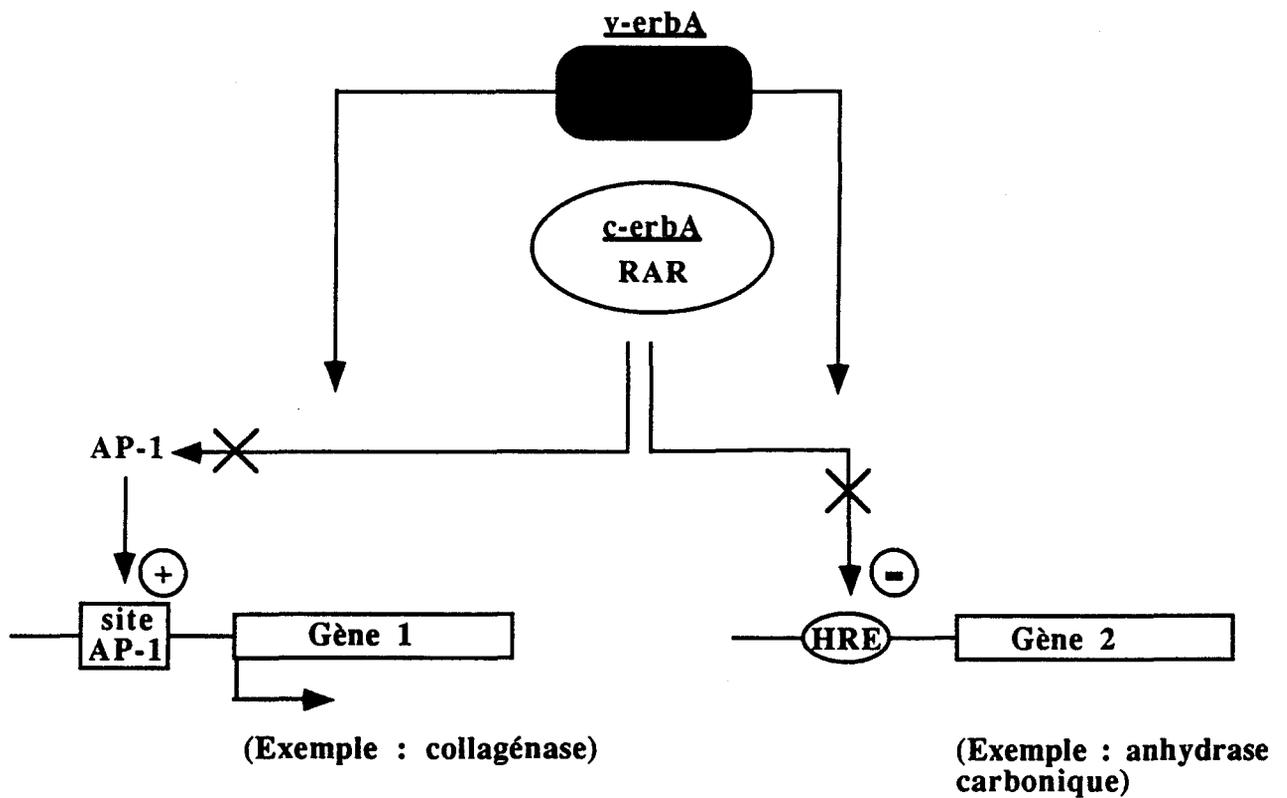
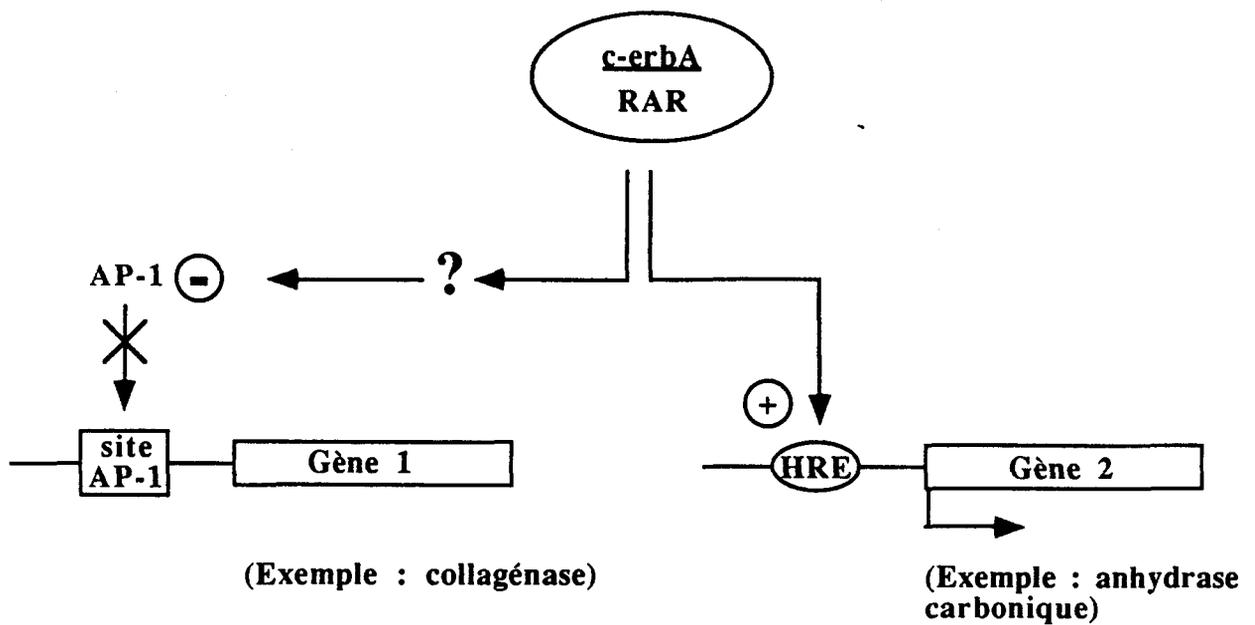


Figure 35 : Les mécanismes d'action possible de v-erbA. A gauche, le blocage par v-erbA de l'action inhibitrice des récepteurs des hormones thyroïdiennes et des récepteurs de l'acide rétinoïque sur AP-1 entraîne la mise en route des gènes régulés par AP-1. A droite, v-erbA empêche l'action régulatrice négative des récepteurs des hormones thyroïdiennes et de l'acide rétinoïque (d'après Desbois et al., 1991).

plus haut que certains récepteurs nucléaires comme ceux de l'acide rétinoïque sont capables d'empêcher l'action activatrice de certains facteurs de transcription comme AP-1. Par exemple, AP-1 active le promoteur de la collagénase et l'acide rétinoïque, via ses récepteurs, est capable de supprimer cet effet (figure 35). Ceci probablement en formant un complexe inactif avec AP-1 (Nicholson et al., 1990). Ce que montre le groupe de Jacques Samarut, c'est que v-erbA est capable de supprimer l'effet négatif des récepteurs de l'acide rétinoïque. Qui supprime un effet négatif obtient *in fine* un effet positif et, ainsi, v-erbA met-il en route inconsidérément des gènes impliqués dans la prolifération cellulaire.

Actuellement il y a donc deux niveaux d'action relativement indépendants de v-erbA (Desbois et al., 1991b). D'abord, dans les cellules érythroïdes, v-erbA entre en compétition avec les récepteurs des hormones thyroïdiennes et bloque ainsi la mise en route d'un programme de différenciation. Ensuite, et c'est probablement le cas dans les fibroblastes, v-erbA, bloque l'inactivation induite par les récepteurs de l'acide rétinoïque, de gènes de prolifération, contrôlés par les facteurs de transcription de la famille AP-1. Il y aura sûrement dans un avenir proche des raffinements supplémentaires apportés à ce modèle...

## D) LA REPONSE AUX HORMONES THYROIDIENNES.

### 1) Les gènes cibles.

A travers le décorticage de l'action de v-erbA que nous venons de réaliser, nous avons rencontré un certain nombre de gènes cibles possibles des récepteurs des hormones thyroïdiennes. C'est le cas par exemple du gène de l'anhydrase carbonique de type II, ou du gène qui code le transporteur d'anion BAND 3 (Zenke et al., 1988 ; 1990 ; Pain et al., 1990 ; Disela et al., 1991). Il est clair cependant qu'il y a gène cible et gène cible. Il faut discriminer un effet direct du récepteur qui va réguler un gène en se fixant sur son promoteur, d'un effet indirect où il va agir sur un gène donné par l'intermédiaire d'un autre. Par exemple, si l'action des récepteurs des hormones thyroïdiennes sur la transcription du gène de l'ACOII est directe (Pain et al., 1990 ; Disela et al., 1991), on ignore toujours si c'est vrai pour le gène de band 3.

Le tableau 6 présente une série de gènes qui sont des cibles soit directes, soit putatives (on n'a pas encore trouvé de TRE dans ces gènes) des récepteurs des hormones thyroïdiennes. Il est certain que la liste n'est ni exhaustive, ni close. De nombreux gènes (celui de la TRH, par exemple) sont suspectés depuis longtemps d'être régulés par les hormones thyroïdiennes (Samuels et al., 1988). C'est le cas du gène de l'hormone de croissance qui est induit par les hormones thyroïdiennes dans l'hypophyse antérieure de rat (Samuels et Shapiro, 1976 ; voir la revue de Samuels et al., 1988), mais aussi de l'enzyme malique stimulée dans le foie (Usala et al., 1988b ; Petty et al., 1989 ; Desvergne et al., 1991), des gènes "SPOT 11" et "SPOT 14" également dans le foie (revue dans Oppenheimer et al., 1987), et de la chaîne lourde de la myosine  $\alpha$  dans le

**TABLEAU 6**

**Une sélection de quelques gènes cibles des récepteurs des hormones thyroïdiennes**

Gène	Remarque	Références
TSH- $\alpha$	Régulation négative	Burnside et al., 1989 ; Chatterjee et al., 1989
TSH- $\beta$	Régulation négative directe ou indirecte	Krane et al., 1991 ; Wondisford et al., 1989 ; Wood et al., 1989 ; Carr et al., 1989
Lysozyme	Régulation négative directe	Baniahmad et al., 1990
Enzyme malique	Régulation positive directe	Desvergne et al., 1991 ; Petty et al., 1989 ; Usala et al., 1988
Band 3	Régulation positive Effet de v-erbA	Zenke et al., 1988
Anhydrase carbonique II	Régulation positive Effet de v-erbA	Pain et al., 1990 ; Disela et al., 1991
Myosine Cardiaque $\alpha$ (chaîne lourde)	Régulation positive directe	Tsika et al., 1990 ; Gustafson et al., 1987 ; Everett et al., 1983 ; Izumo et al., 1986
Deiodinase I	Régulation positive Directe ou indirecte ?	Berry et al., 1990
Spot 14	Régulation positive Directe ou indirecte ?	Hamblin et al., 1987 ; Wong et al., 1989
Adenine Nucleotide Translocase	Régulation positive Directe ou indirecte ?	Höppner et al., 1988
GRH	Régulation positive Directe ou indirecte ?	Downs et al., 1990
Fibronectine	Régulation positive Directe ou indirecte ?	Murata et al., 1990
Apolipoprotéines	Dépend des apolipoprotéines et de l'apolipopptose	Davidson et al., 1988

myocarde (Everett et al., 1983 ; Izumo et al., 1986). Certains gènes, qui sont réprimés par les hormones thyroïdiennes, sont également connus depuis assez longtemps : c'est le cas de la TSH (Kourides et al., 1984 ; Shupnik et al., 1985 ; Carr et al., 1989) et de la chaîne lourde de la myosine de type  $\beta$  (Sinha et al., 1982).

Le gène de l'hormone de croissance est la première cible précisément décrite du récepteur des hormones thyroïdiennes et nous allons nous attarder sur la régulation de ce gène pour montrer comment la réponse aux hormones thyroïdiennes s'inscrit dans un cadre plus général.

## 2) La régulation du gène de l'hormone de croissance.

La partie antérieure de l'hypophyse a pour particularité de sécréter, sous le contrôle de l'hypothalamus, de très nombreuses hormones peptidiques qui vont jouer des rôles très divers dans l'organisme. Parmi ces hormones, on peut citer la TSH, l'ACTH, la LH, mais aussi la GH et la prolactine. Au sein de la glande, une cellule donnée sera responsable de la synthèse d'un type d'hormone donné. On aura donc des cellules somatotrophes (productrices de GH), des cellules thyrotrophes (productrices de TSH) etc... Il faut donc trouver des facteurs tissus spécifiques se fixant dans les parties régulatrices du gène de la GH par exemple (Karin et al., 1990). De plus, chez le rat, on sait que la production de GH, à la fois *in vivo* et *in vitro* est sous la dépendance des hormones thyroïdiennes et, dans une moindre mesure des glucocorticoïdes (On a d'ailleurs localisé un GRE dans le premier intron du gène de la GH humaine Slater et al., 1985 ; Moore et al., 1985 ; Brent et al., 1988 ).

La régulation de la transcription du gène de la GH est variable selon l'espèce considérée. Ainsi, le gène de la GH de rat est dépendant des hormones thyroïdiennes, tandis que celui de l'homme ne l'est pas (Brent et al., 1988). Le gène de la GH est aussi influencé par différentes hormones, comme la GHRH (l'hormone hypothalamique qui contrôle par "feed-back" la sécrétion de GH) qui agit par la voie du second messenger l'AMP cyclique. L'acide rétinolique contrôle aussi de manière directe la transcription du gène de la GH (Bedo et al., 1989).

En amont du gène de la GH de rat, on a localisé des éléments conférant la spécificité tissulaire, d'autres qui sont contrôlés par l'AMP cyclique et bien sûr d'autres encore nécessaires pour la réponse aux hormones thyroïdiennes (figure 36 ; Wight et al., 1987 ; 1988 ; Ye et al., 1987 ; 1988 ; Glass et al., 1987 ; Brent et al., 1988 ; 1989 ; 1991 ; Forman et al., 1988 ; Copp et Samuels, 1989). Certains auteurs ont suggéré que les hormones thyroïdiennes via leur récepteur, permettaient d'augmenter l'effet activateur du facteur tissu-spécifique en l'aidant à former un complexe plus "actif" avec la machinerie de base de la transcription (Ye et al., 1988).

Le TRE du gène de la GH de rat a été localisé, après beaucoup de discussions, à environ 175 paires de bases en amont du site d'initiation de la transcription (Larsen et al., 1986 ; Koenig et al., 1987 ; Flug et al., 1987 ; Glass et al., 1987). Par la suite, l'équipe de Bjorn Vennström a montré qu'il existait dans le troisième intron de ce gène un TRE qui est semble-t-il, beaucoup plus

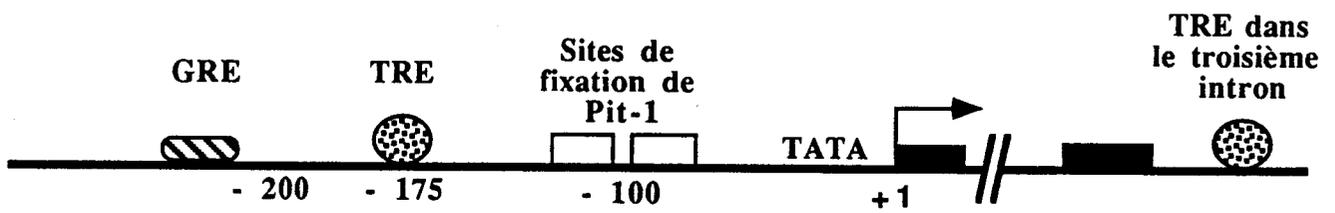


Figure 36 : Schéma du promoteur du gène de l'hormone de croissance de Rat.

efficace que celui présent dans le promoteur (Sap et al., 1990).

Le facteur tissu-spécifique qui contrôle la régulation du gène de la GH est le facteur Pit-1, lequel contient un domaine homéo responsable de la fixation à l'ADN et une boîte typique POU (Bodner et al., 1988 ; Ingraham et al., 1988 ; 1990 ; Fox et al., 1990; Karin et al., 1990). Ce facteur se fixe sur le promoteur de la GH sur deux sites situés en aval du TRE vers -80 et -120 (figure X). Un argument qui souligne l'importance de Pit-1 est que certaines souris affectées de la mutation "Dwarf" et à qui il manque trois types cellulaires dans l'hypophyse antérieure (dont les cellules somatotrophes) ont des délétions au sein des deux allèles de Pit-1 (Li et al., 1990 ; Castrillo et al., 1991). Il existe de nombreuses interactions entre Pit-1 et d'autres facteurs de transcription (Simmons et al., 1990). Par exemple, on sait qu'entre les deux sites de fixation de Pit-1, présent dans le promoteur de gène de l'hormone de croissance de rat, on trouve le site de fixation du facteur de transcription mis en route par la voie de l'AMP cyclique (figure 36 ; Dana et Karin, 1989; Karin et al., 1990; Courtois et al., 1990). En ce qui concerne le récepteur des hormones thyroïdiennes et ses interactions avec le facteur Pit-1, on reste encore sur notre faim puisque le modèle de potentialisation par les hormones thyroïdiennes de l'activation de la transcription médiée par Pit-1 n'a, à notre connaissance, pas encore été vérifié expérimentalement (Yé et al., 1988 ; Karin et al., 1990).

### 3) Un processus physiologique : la métamorphose des amphibiens.

Le modèle le plus classique et le plus spectaculaire d'effet des hormones thyroïdiennes est celui de la métamorphose des amphibiens. Au cours de cette série d'événements particulièrement complexes, on assiste à un remodelage complet du têtard avec résorption de la queue, développement des membres, modification du tractus digestif, etc... Ces événements sont largement contrôlés par les hormones thyroïdiennes, puisqu'un têtard thyroïdectomisé ne se métamorphose pas (Hourdry et Beaumont, 1985). Le clonage des récepteurs des hormones thyroïdiennes a permis d'approcher cette question en terme moléculaire.

On peut diviser le développement des amphibiens en trois périodes distinctes: (i) la pré-métamorphose, qui commence avec le développement de l'oeuf, et se poursuit avec l'éclosion et la phase de croissance des têtards. Chez les grenouilles et crapauds de nos régions, cette phase dure environ trois mois; (ii) la prométamorphose, qui constitue le début de la métamorphose proprement dite et qui se signale par un ralentissement de la croissance des têtards. C'est une période de transition, de préparation à la métamorphose qui dure entre deux et huit semaines.; (iii) le climax ou paroxysme de la métamorphose où l'on observe des transformations morphologiques et biochimiques rapides (moins de deux semaines). Au cours de cette étape, il y a disparition de certains organes (la queue, chez les Anoures), apparition d'autres organes (les membres) et remodelage de nombreux autres (l'hépatopancréas, l'oeil etc...) (Hourdry et Beaumont, 1985).

Grâce à la découverte des homologues d'amphibiens des gènes c-erbA (Brooks et al.,

1989 ; Yaoita et al., 1990 ; Schneider et Galton et al., 1991), on a pu étudier leur expression au cours de ces étapes. Le premier enseignement de ces études est que ces gènes sont exprimés au cours du développement beaucoup plus tôt que ce que l'on imaginait. En effet, les hormones thyroïdiennes ne sont décelables qu'au début de la prométamorphose. Pourtant, les récepteurs de type  $\alpha$ , codés par le gène c-erbA-1, sont présents bien plus tôt au cours du développement (Baker et Tata, 1990 ; Yaoita et Brown, 1990 ; Kawahara et al., 1991 ; Banker et al., 1991). Des travaux récents suggèrent même que des ARN c-erbA-1 pourraient être exprimés dans les oocytes, et donc se comporter comme des ARN à effet maternel (Kawahara et al., 1991 ; Banker et al., 1991). Il est intéressant de rappeler qu'en absence d'hormone les récepteurs se comportent, au moins chez les mammifères, comme des régulateurs négatifs de la transcription. On peut donc imaginer qu'au cours des stades précoces les protéines erbA produites répriment en permanence certains programmes génétiques. Il faut noter que cette expression précoce des gènes erbA, avant la détection d'une quelconque fonction thyroïdienne, est un phénomène qui semble assez général puisqu'il a été détecté également chez le poulet (Forrest et al., 1991).

Pour résumer l'expression des gènes c-erbA dans les étapes plus tardives, on peut dire que c-erbA-1 est beaucoup plus exprimé que c-erbA-2. Cette différence est très forte au début du développement (30 fois, selon certains auteurs ; Kawahara et al., 1991) pour ensuite devenir plus faible. Au moment du climax l'expression du gène c-erbA-2 augmente brusquement, sans jamais dépasser celle de c-erbA-1. Après la métamorphose le niveau des deux gènes retombe à un très faible niveau (Baker et Tata, 1990 ; Yaoita et Brown, 1990 ; Kawahara et al., 1991). D'un point de vue spatial, il semble que ces gènes soient exprimés plus fortement du côté de la tête mais les expériences d'hybridation *in situ* sont encore rares. Dans les organes impliqués à un stade donné dans des événements morphogénétiques particulièrement vigoureux, on observe une nette augmentation des signaux correspondants aux ARNs erbA (Kawahara et al., 1991). Comme chez les mammifères et les oiseaux, le gène c-erbA-2 semble être plus finement régulé que le gène c-erbA-1, à la fois dans le temps et dans l'espace.

Evidemment, comme la métamorphose est sous le contrôle des hormones thyroïdiennes, tout le monde a été convaincu de l'intérêt d'observer la régulation de la transcription des gènes erbA par ces hormones. Et bien ça marche ! Tous les auteurs sont d'accord pour dire que la transcription des gènes est augmentée par les hormones thyroïdiennes. Le cas le plus spectaculaire est celui du gène c-erbA-2 dont la transcription peut être augmentée de 20 fois, mais le gène c-erbA-1 voit également son expression doubler en présence de T3 (Baker et Tata, 1990 ; Yaoita et Brown, 1990 ; Schneider et Galton, 1991 ; Kawahara et al., 1991). Un certain nombre de ces travaux laissent penser que cette régulation est directe mais ce n'est pas encore définitivement prouvé. Il est remarquable que, contrairement aux systèmes étudiés chez les mammifères, la T3 induit une régulation positive de l'expression des gènes c-erbA.

Cette cartographie assez détaillée de l'expression des gènes erbA, n'est pas encore terminée et déjà se profile à l'horizon l'étude moléculaire des gènes cibles impliqués dans la

réponse à la T3 au cours de la métamorphose. Jamshed Tata et son équipe disposent d'élégants systèmes *in vitro* pour pêcher de tels gènes cibles. Ils ont en effet mis au point des cultures organotypiques de queues et de membres de Xénope en boîtes de Pétri (Tata et al., 1991). Ces organes répondent correctement à l'adjonction de T3 dans le milieu : les queues régressent et les membres poussent ! Avec ce genre de système, on peut, par exemple, rechercher les gènes qui sont mis en route par la T3 et qui déclenchent la mort cellulaire programmée des cellules de queue. La réponse observée à la T3 n'est sans doute pas simple et probablement pas directe dans la plupart des cas (Mathisen et Miller, 1989 ; Shi et Brown, 1990). Il est donc clair que si la métamorphose du xénope semble un merveilleux système d'étude des effets physiologiques de la T3, il s'agit sans aucun doute d'un système extrêmement complexe. Soyons patients...

## CONCLUSION

Au cours de ce petit voyage au sein du monde des récepteurs nucléaires nous avons fortement insisté sur les relations, les synergies, les antagonismes, les coopérations, les interactions, les liens enfin, qui unissent ces molécules entre elles et avec les autres molécules chargées de modifier, en fonction de l'environnement, le programme d'expression génétique de l'organisme. Ce programme va donc à chaque instant réagir d'une façon parfois très douce, parfois violente à cet environnement et doit donc intégrer, c'est le maître mot, tous les messages qui lui parviennent.

Etudier l'apparition d'un tel système, puis son évolution, au cours du temps et selon les espèces va avoir deux types d'intérêts. D'abord, amener des renseignements sur le système considéré en lui-même : ici la superfamille des récepteurs nucléaires. Comment une famille aussi complexe est-elle apparue ? Comment s'est-elle diversifiée ? Peut-on relier une réponse physiologique précise (la réponse aux hormones thyroïdiennes par exemple) à un moment précis de l'évolution de cette famille ? Y a-t-il des éléments redondants dans cette famille ? Quelle est la manière dont les différents domaines fonctionnels ont évolué ? Sont-ils indépendants ?

Ensuite, on peut également apporter des informations sur l'évolution des gènes en général et peut-être, c'est une perspective ambitieuse, réussir à relier un événement important de l'évolution, l'apparition des animaux multicellulaires, à un moment de l'évolution d'une superfamille de gènes - sans y voir nécessairement un lien causal direct.

Pour mener une telle démarche évolutive il est utile de bien connaître la structure des gènes que l'on étudie. Lorsque nous avons commencé ce travail dans le laboratoire de Dominique Stéhelin, seule la structure des gènes de certains récepteurs des stéroïdes (oestrogènes, progestérone; Ponglikitmongkol et al., 1988 ; Huckaby et al., 1987) était connue. Il nous a donc paru important de connaître également la structure d'un gène de récepteur nucléaire d'un autre type, celui du récepteur des hormones thyroïdiennes de type  $\alpha$ , le gène c-erbA-1 humain. Pour un laboratoire d'oncologie moléculaire ce gène présente également l'intérêt d'être étroitement apparenté à l'oncogène viral v-erbA.

Pour la rédaction de ce mémoire nous avons respecté l'ordre chronologique. Nous présenterons d'abord les résultats obtenus sur la structure du gène c-erbA-1 humain puis, les résultats obtenus sur l'évolution de la superfamille des récepteurs nucléaires. Il est clair que ce sont les résultats obtenus pendant l'étude du locus c-erbA-1 qui, peu à peu nous ont poussés à nous interroger sur l'évolution des récepteurs nucléaires.

## RESULTATS

**ARTICLE 1 :**

**ORGANISATION GENOMIQUE DU GENE HUMAIN**

**CODANT LE RECEPTEUR DE TYPE  $\alpha$  DES**

**HORMONES THYROIDIENNES (c-erbA-1).**

Les études fonctionnelles menées sur les récepteurs nucléaires montrent bien que ces récepteurs ont des domaines fonctionnels indépendants. Le fait que l'on puisse réaliser des récepteurs chimères comprenant à la fois un domaine du récepteur des hormones thyroïdiennes et les autres domaines du récepteurs des oestrogènes est une preuve de cette indépendance (Green et Chambon, 1988 ; Evans, 1988). Cette indépendance est vraie pour les différents domaines étudiés : fixation à l'ADN, fixation du ligand, ou transactivation (Thompson et Evans, 1989). Certains auteurs ont même réalisé des chimères entre des récepteurs nucléaires et des oncogènes viraux (E1A ou la protéine rev du virus HIV-1 ; Becker et al., 1989 ; Hope et al., 1990) ou cellulaires (myc ou fos ; Eilers et al., 1989 ; Superti-Furgas et al., 1991). On a même obtenu des oncogènes dont l'action transformante devenait hormono-dépendante.

Cette indépendance des domaines fonctionnels se retrouve-t-elle au niveau du génome ? En d'autres termes, chaque domaine est-il codé par un seul exon ou par des exons séparés ? Les premières analyses menées sur les gènes des récepteurs des stéroïdes permettaient de répondre partiellement à cette question. La structure du gène humain codant le récepteur des oestrogènes a permis d'observer, qu'en effet, les domaines de ce gène sont, en règle générale, codés par des exons séparés (Ponglikitmongkol et al., 1988). Dans ce gène, qui mesure au total plus de 140 kb, on trouve un grand exon codant le domaine A/B, puis un exon pour chaque doigt de zinc du domaine C et enfin, le domaine de fixation du ligand qui est découpé en cinq exons. Les seuls domaines qui ne sont pas codés par des exons séparés sont le domaine D et le domaine F situé en C-terminal, codé par le huitième et dernier exon. Il est intéressant de constater que la position de chaque exon est très bien conservée entre le gène du récepteur des oestrogènes et celui du récepteur de la progestérone (Huckaby et al., 1987 ; Ponglikitmongkol et al., 1988). Ceci montre que ces gènes sont étroitement apparentés et proviennent bien d'un ancêtre commun, par duplication. La comparaison de la structure de ces gènes avec la structure partielle du gène du récepteur des hormones thyroïdiennes de poulet (Zahraoui et Cuny, 1987) a permis de montrer que leur structure semblait avoir des points communs (chaque doigt de zinc est codé par des exons séparés) et des différences importantes (position et nombre des exons, taille, etc...).

Outre la comparaison de la structure génomique des différents gènes de la superfamille des récepteurs nucléaires, des raisons supplémentaires nous ont également poussés à réaliser cette étude. En effet, nous avons vu que les récepteurs des hormones thyroïdiennes pouvaient coder de nombreuses isoformes. En particulier, le gène c-erbA-1 peut donner, par épissage alternatif, deux isoformes majeures qui diffèrent dans leur partie 3' (Koenig et al., 1989 ; Izumo et Mahdavi, 1988; revue dans Chin, 1991). De plus, d'autres espèces minoritaires ont été décrites (Mitsubishi et al., 1988 ; Mitsubishi et Nikodem, 1989 ; Izumo et al., 1990). Il était donc intéressant d'étudier l'organisation de ce gène en exons et en introns pour situer les différents épissages. De plus, la connaissance de l'organisation d'un tel locus permet de prévoir l'existence de nouveaux épissages alternatifs, qui peuvent être mis en évidence par des expériences de protection à la RNase ou d'amplification de gènes (PCR). Par exemple, comme l'on sait qu'il existe de nombreux épissages alternatifs possibles en 5' des gènes c-erbA-2 du xénope (Yaoita et al., 1990), et des gènes des récepteurs de l'acide rétinoïque chez les rongeurs (Kastner et al., 1990 ; Leroy et al., 1991 ; Zelent et al., 1991). La connaissance de l'organisation génomique, peut permettre ainsi de tester si de tels épissages existent aussi pour le gène c-erbA-1 humain comme il est tentant de l'imaginer...

La détermination d'une telle structure permet de disposer d'une carte du gène dans son organisation normale. Nous avons vu, lors de l'étude des leucémies aiguës à promyélocytes dans lesquelles le gène  $RAR\alpha$  est impliqué, combien il pouvait être utile de disposer d'une carte fine d'un gène (Cleary, 1991). On peut ainsi avoir les outils nécessaires à la compréhension de la structure modifiée au cours d'un processus pathologique. Enfin, les études concernant la régulation de l'expression de ce gène sont également rendues plus faciles par l'étude de l'organisation génomique, surtout dans la partie 5'.

Pour ces raisons, il nous a semblé utile de déterminer l'organisation génomique du gène c-erbA-1 humain. Ainsi, nous avons donc d'abord caractérisé différents cDNA correspondant à ce locus. Puis, dans un second temps, nous avons isolé des phages d'ADN génomique humain en utilisant comme sonde, soit des fragments de ces cDNAs, soit des sondes v-erbA. Après avoir localisé au sein de ces phages les régions qui hybrident avec les cDNAs nous avons sous-cloné ces fragments d'exons et nous les avons séquencés. Nous disposons ainsi de la séquence complète des parties exprimées (*i.e.* les exons) du gène c-erbA-1.

## Genomic organization of the human thyroid hormone receptor $\alpha$ (*c-erbA-1*) gene

Vincent Laudet\*, Agnes Begue, Catherine Henry-Duthoit<sup>1</sup>, Anita Joubel, Patrick Martin, Dominique Stehelin and Simon Saule

INSERM U 186/ CNRS UA 04 1160, Institut Pasteur de Lille, 59019 Lille Cedex and <sup>1</sup>INSERM U 124 IRCL, place de Verdun, 59045 Lille cedex, France

Received November 16, 1990; Revised and Accepted February 5, 1991

EMBL accession nos X55004, X55005, and X55066 - X55074 (incl.)

### ABSTRACT

The thyroid hormone receptor  $\alpha$  (THRA or *c-erbA-1*) gene belongs to a family of genes which encode nuclear receptors for various hydrophobic ligands such as steroids, vitamin D, retinoic acid and thyroid hormones. These receptors are composed of several domains important for hormone-binding, DNA-binding, dimerization and activation of transcription. We show here that the human THRA gene is organized in 10 exons distributed along 27 kbp of genomic DNA on chromosome 17. The position of the introns in human THRA is highly conserved when compared to the chicken gene despite their differing lengths. The N-terminal A/B domain as well as the 5' untranslated region is encoded by two exons. Interestingly, each of the putative zinc fingers of the receptor DNA-binding domain is encoded by one exon and the hormone-binding domain is assembled from three exons. The two last exons of the gene are alternatively spliced to generate two different messenger RNAs. In addition, we confirm that another gene, belonging to the nuclear receptor superfamily, *ear-1*, overlaps with the 3' region of THRA in an opposite transcriptional orientation.

### INTRODUCTION

The thyroid hormone receptor  $\alpha$  (THRA) belongs to a family of ligand-activated transcription factors which regulate the expression of target genes by binding to specific cis-acting sequences (1-3). This family includes nuclear receptors for several hydrophobic ligands, such as glucocorticoids, estrogens, progesterone, vitamin D or retinoic acid and also for dioxin and ecdysone. In addition to these receptors for known ligands, several authors have described 'orphan receptors' which are putative receptors to ligands still to be identified (1). The total number of genes described in this growing family is near 30. The family of nuclear receptors has been classically divided in two subfamilies on the basis of structural homologies. One subfamily includes the genes encoding steroid hormone receptors and the other comprises the receptors for thyroid hormone and

retinoic acid as well as several orphan receptors. To date, this subfamily contains three retinoic acid receptors,  $\alpha$ ,  $\beta$  and  $\gamma$  (4-7), two thyroid hormone receptors,  $\alpha$  (THRA on chromosome 17q21) and  $\beta$  (THRB on chromosome 3p21) (8,9) and three *c-erbA* related genes named *ear-1*, *ear-2* and *ear-3* which are located on human chromosomes 17, 19 and 5 respectively (10,11). The THRA and THRB are similar in overall structure, being most related in the cysteine-rich DNA-binding and C-terminal hormone-binding regions and more divergent in the A/B domain which has been suggested to represent at least part of the transactivation domain. Nucleotide sequence comparison has revealed that the chicken THRA gene (12,13) is the cellular progenitor of the *v-erbA* oncogene found in the Avian Erythroblastosis Virus (AEV Refs : 12, 14-17). The *v-erbA* oncogene does not by itself cause tumors in animals, but enhances the transforming potential of oncogenes that induce sarcomas (18-19).

In addition to this diversity of thyroid hormone receptor (TR) genes, isolation of rat and human THR cDNA clones has revealed that a variety of TRs may be produced (8,12,20-23). Alternative splicing of the two THR genes leads to the generation of functionally different receptors (11, 24-28). In the case of the THRA gene, two receptor forms are generated which differ in their C-terminal region. One receptor (hTHRA1), is composed of 410 amino-acids, binds T3, and activates the transcription of target genes such as the growth hormone gene. The other receptor (hTHRA2), composed of 490 amino-acids, does not bind T3 and acts as a dominant negative regulator of TRs. Dominant negative regulation by hTHRA2 is exerted by suppression of the transactivation mediated by products of either the hTHRA1 or THRB genes (26,27). Recent studies have shown that the *v-erbA* product acts as a thyroid hormone independent repressor of target gene expression much like hTHRA2. The *v-erbA* product is able to block the thyroid hormone-mediated activation of a target gene when coexpressed with the normal receptor (18).

To understand the evolution of this superfamily of ligand-dependent transactivation factors and to facilitate the analysis of their transcriptional regulation, the chromosomal genes must be cloned. To date, knowledge of the complete genomic organization

\* To whom correspondence should be addressed

of members of this family is restricted to members of the steroid hormone receptor subfamily (chicken progesterone [29], human estrogen [30] and human androgen receptor genes). For the thyroid hormone/retinoic acid receptor subfamily, only partial gene structures have been described for THRA and RAR  $\beta$  and  $\gamma$  genes (11, 33–36). In this paper, we show that the human THRA gene is composed of a transcriptional unit spanning 27 kbp split into 10 exons. In addition, we show that another gene, already described and termed *ear-1* or *REVerb* (11, 34) is detected in the same locus and overlaps in a tail-to-tail orientation with the 3' region of THRA.

## MATERIALS AND METHODS

### Molecular Probes

Different fragments were derived from either *v-erbA* and THR cDNAs or genomic DNA and used as probes for cloning different parts of the locus:

The *v-erbA* probes used were (i) a 1.4 kbp *HincII* fragment representing the major part of the *v-erbA* gene and a part of the *v-erbB* gene and (ii) a 0.5 kbp *Sall*-*SacI* 3' fragment of *v-erbA*. These two probes were isolated from the AEV provirus cloned in pKH47 (37).

The cDNA and genomic probes were (i) the entire 14C THRA cDNA (ii) an *EcoRI*-*BstEII* fragment of 14C containing the first 550 nucleotides of this cDNA, (iii) a 0.9 kbp *EcoRI*-*KpnI* fragment representing a large portion of exon 9 from  $\lambda$ E2A2, (iv) a 1 kbp *XhoI*-*KpnI* fragment of the genomic clone  $\lambda$ E2A3 containing parts of exons 8 and 9.

### Library screening

The 14C and 14J cDNAs were isolated from a K562 cell cDNA library in  $\lambda$ gt10 (provided by E. Canaani) using the 3' *v-erbA* *Sall*-*SacI* probe and the 0.9 kbp *EcoRI*-*KpnI* genomic probe respectively. The 14E cDNA was isolated from a fetal liver cDNA library cloned in Charon 21, provided by J.J. Toole and screened with the 550 bp *EcoRI*-*BstEII* probe representing the 5' part of 14C.

We have isolated four overlapping clones named  $\lambda$ E2A3,  $\lambda$ E2A2,  $\lambda$ E2A1 and  $\lambda$ E2A0 from three human genomic libraries. The  $\lambda$ E2A3 clone was isolated from a human placental DNA library provided by Tom Maniatis (38). This library was screened with the 1.4 kbp *v-erbA* probe. Because we were unable to isolate a phage containing the 5' part of the THRA gene in this library, we decided to construct a new genomic DNA library using ML3 cells DNA partially digested with *SauIII*A and inserted into the EMBL4 vector. From this library the  $\lambda$ E2A2 and  $\lambda$ E2A1 clones were isolated. The  $\lambda$ E2A2 was obtained after screening with the entire 14C probe. The  $\lambda$ E2A1 clone was isolated using the 5' *EcoRI*-*BstEII* probe derived from the 14C cDNA. Finally, the  $\lambda$ E2A0 clone was derived from DNA of ovarian carcinoma IGROV-2 cells, partially digested with *Sau3A* and inserted into the EMBL4 vector. The clone was isolated by screening with the 5' *EcoRI*-*BstEII* probe. Cloning, phage and plasmid growth or isolation were performed using standard procedures (39).

### Sequencing experiments

The different cDNA or genomic fragments were subcloned into the M13mp18 or M13mp19 sequencing vectors using standard techniques and an Applied Biosystem 370A automatic sequencer. Exon 8 and the 5' half of 14C were sequenced using the base-specific chemical method described by Maxam and Gilbert (40).

### PCR amplification on messenger RNA

Messenger RNA from MRC5 cells was extracted using guanidinium thiocyanate according to standard methods (39). RNA was retrotranscribed and amplified by PCR using the Perkin Elmer Cetus GeneAmp RNA PCR kit and the classical method of Wang et al. (41). Briefly, 500 ng of polyA+ RNAs were incubated for 15 min at 42°C in the buffer furnished by the supplier with 5mM  $MgCl_2$ , 1mM of each dNTP, 50 pmol of oligonucleotide A, 1 unit of RNasin and 2.5 units of the reverse transcriptase from Perkin Elmer Cetus in a volume of 20  $\mu$ l. The mixture was then denatured at 99°C for 5 min and ice-cooled for 5 min. The PCR amplification was then conducted in the same buffer at a final concentration of 2 mM  $MgCl_2$  with 50pmol of primer, 2.5 units of TAQ polymerase. The PCR reaction was conducted in a Perkin Elmer Cetus DNA Thermal Cycler with the following 35 cycles: 1 min at 95°C, 90 sec at 60°C. During the last cycle, a synthesis step at 72°C was performed for 7 min. The samples were analyzed on agarose gels stained with ethidium bromide. Nested PCR was conducted using the same protocol and reagents on 2  $\mu$ l of the primary amplification mixture as a target.

The oligonucleotides used were synthesised by Institut Pasteur de Lille Oligonucleotide Facility and were as follows:

(A) GTG GCC AGG ACC CTG TTG TG which corresponds to the beginning of 14J cDNA, (B) AAA AAC TGC CCA TGT TCT CCG which corresponds to the end of exon 7, (C) CCA GCT GAA TTC ACC CTC TGG CCG CCT GAG GCT which corresponds to the end of the coding part of 14C in exon 9 and (D) GCT CTG CTG CAG GCT GTG CTG CTA ATG T which was used for nested PCR and which corresponds to the end of exon 8. The sequences of the oligonucleotides are underlined in Fig.3

## RESULTS

### Cloning of THRA cDNAs

A K562 cell cDNA library was screened using a  $^{32}P$ -labeled 0.5 kbp *Sall*-*SacI* probe which corresponds to the 3' portion of *v-erbA*. Two positive clones were obtained from  $4 \times 10^6$  plaques and designated 14C and 14E. The 14C clone was further characterized by restriction endonuclease mapping and sequence analysis. The restriction map is shown in Fig.1A. The 2288 bp

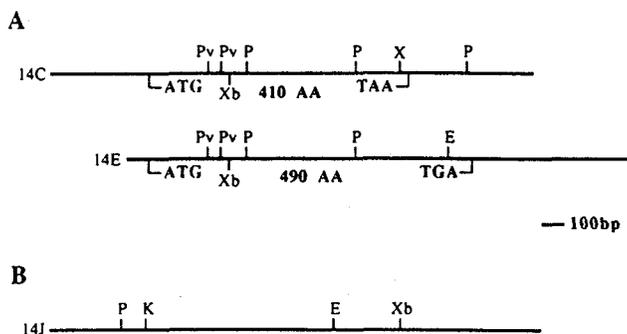


Fig. 1. Map of the three cDNA clones used in this study. (A) 14C and 14E clones have two thirds of their 5' regions in common, as revealed by endonuclease mapping. (B) Map of the 14J cDNA. ATG: Initiation codon. TGA and TAA: Stop codon. Enzymes used: Pv, Pvu2; P, Pst1; X, Xho1; Xb, Xba1; E, EcoR1; K, Kpn1.



EXON 10 \* End of ear-1 exon F  
 acacaccacagaagCCAGCTCAGCTGTGACTATTGGATTGGAGACAGGACAGAACCAATCAGAGGGCCAGGGGAGGGTGTGGGGGAGACAGAGTGGTTAATAGGGGAGGGGGG 120  
 AGTTCCGGTATGGGGAGGGAGGCGAGGTATTACAGAGAGGCTCAGGGGGCCAGAGGGCTCATCTTGGAAATATTTATACRAATATAAATAGATTCTGGTTTGGCTTTCCCTTTTCGTC 240  
 GTAARCGAGAGAGAGTCCAGAGTTCGATCTGTACAGGGGGCAGCGGCGAGAGGGCCGGCCGGGGCTCACTGGCGCTCCACCCGGAGGACACAGCTTCTCGGAATGCAITGTTGT 360  
 E R E U Q S S I L Y K G A A A E G R P G G S L G U H P E G Q Q L L G N H U U  
 CAGGGTCCGAGGTCGGCAGCTTGAGCAGCAGCTTGGTGAGCGGGAGTCTCCAGGGCCGGTCTCAGCACCAGAGCCCGAGAGGCCCCGAGCAGCGTCTCTGGAGCTGCTCCAC 480  
 Q G P Q U R Q L E Q Q L G E A G S L Q G P U L Q H Q S P K S P Q Q R L L E L L H  
 Start of ear-1 exon F \*  
 CGAAGCGAATTCCTCCATGCCCGAGCGGTCTGTGGGAGAGCAGCAGCGTGAAGCGGACTCCCCAGCTCTCTGAGGAGGACCGGAGGTCTGCGAGGACCTGGCAGGCAATGCAGCC 600  
 R S G I L H A A A U C G E D D S S E A D S P S S S E E E P E U C E D L A G N A A  
 TCTCCCTGAGCCCCCAGAGGCCGATGGGAGGAGAGGAGTGCATACCTTCTCCAGGCCCTCTGCCCAAGAGCGAGGAGTGCCTGAAAGCTGGGAGCGCTGGCTCAGCAGGGCT 720  
 S P \*  
 GGTCACTCCCATCCGTAGAGGACCTCCCTTCTCAGCAGGCCAARCATGGCCAGACTCCCTTGTCTTTTGTGTGTAGTTCCTCTGCTGGGATGCCCTTCCCTTCTCTGCC 840  
 TGGCAACATCTAGTTGTCCFTTGAGGCCCAACTCAGTGTCACTCCTTCCCGAGCTCCCCAGGCAGAAATAGTTGTCTGTGCTTCTTGGTTTCATCGCTCTACTGTGACACTATC 960  
 TGACTGTTTATAATTAGTCGGGCATGAGTCTGTTTCCCAAGCTAGACTGTCTGATCATGTCTGTATCCCAAGTCCCGGTGACAGGCCCTGGCATAGAGTAGGTACTCCATAAAGG 1080  
 \* End of ear-1 exon F-1  
 TGTGTTGaaTgaactgcgtctgcttccccgggtcagggagagagcctgacctacCTGCAGAGACAGCACCAGCGGTGAGAGGCCAGCTCTCTCGGTAAGCGCCAGGGAGT 1100  
 TGAGCTTCTCGCTGAAGTCGACATGGCACCTGAGCAGGTCTCCATGCCCATGCCACCAAGCTCTCGAGGCTGATAGGTGGTGGCGCTTAGGAACATCACTGTCTGGTCTTCAGTTGA 1320  
 \* Start of ear-1 exon F-1  
 ACAACGAGCAACGCCACcctcagtcacctaggagggaaggggaacagtc 1370

Fig. 3. Sequence of the 10 exons of THRA gene. Exons are typed in upper case letters except for the 147 bp junction between 14C and 14J cDNAs in exon 9 (844-961) which is in underlined, lower case letters. Each exon is numbered individually. Intronic sequences are in lower case letters. Oligonucleotides A (exon 9), B (exon 7), C (exon 9) and D (exon 8) are underlined. Putative polyadenylation signals are shown by arrowheads located under the first base of the site. For the genomic portion containing exon 10, the *ear-1* exons 'F' and 'F-1' are shown in italics. The translation product of these two exons is not shown since it is identical to that in reference 11. The beginning of the THRA exon 10 is underlined. The alternative splice donor site located 128 bp downstream from the beginning of exon 10 is indicated.

kb mRNA species in MRC5 cells (data not shown). To exclude the possibility that the two probes recognize two different RNAs of the same size, we performed PCR experiments which showed that the 14C and 14J sequences both contributed to the same 5.0 kb transcript. It was difficult to directly amplify the missing 147 nucleotide junction between 14C and 14J because this fragment was entirely encoded by exon 9 and we could not exclude the possibility that genomic DNA contaminating the RNA preparation would give the expected signal. We therefore chose an indirect strategy depicted in Fig.4A wherein RNAs isolated from MRC5 cells expressing the 5.0 kb mRNA were retrotranscribed using oligonucleotide primer A. This oligonucleotide corresponds to the very 5' part of the 14J cDNA and was chosen to be specific for these sequences (an extensive computer search has revealed no significant homology with 14C). The single strand DNA molecules obtained were used as templates for PCR experiments using oligonucleotides B and C corresponding to sequences of exons 7 and 9. The expected 556 bp band was obtained (data not shown) and was reamplified by nested PCR using primer C already described and primer D corresponding to the end of exon 8 and yielded the expected 307 bp product. This band co-migrates with the product of the same PCR reaction conducted with MRC5 RNAs retrotranscribed with oligonucleotide C (Fig. 4, panel B). To exclude the possibility that the observed bands were due to carryover contamination (49), we performed control amplifications where no target DNAs were added. As expected, this control did not reveal any band after amplification with either oligonucleotides B and C (not shown) or C and D (Fig. 4, panel B). This experiment shows that after priming in the 14J sequences we have been able to amplify parts of 14C. The use of oligonucleotides which correspond to different exons (namely 7, 8 and 9) for primary as well as nested PCR prevents the amplification of genomic DNA contaminating the RNA preparation. We thus conclude that 14C and 14J sequences both contribute to the 5.0 kb mRNA species.

As these experiments do not exclude the possibility that a small intron could lie in the 147 bp missing fragment between 14C and 14J, we performed amplification experiments of that genomic DNA fragment and of MRC5 mRNA retrotranscribed with

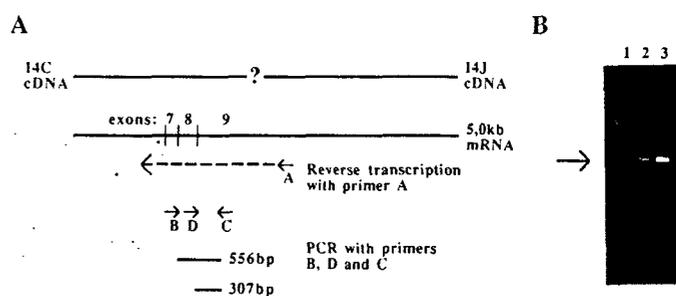


Fig. 4. (A) Strategy for the PCR experiments. Arrows show the oligonucleotides used: A and C correspond to sequences in exon 9, and B and D are from exons 7 and 8, respectively. This strategy prevents the amplification of genomic DNA. (B) visualization on 2% agarose gel stained with ethidium bromide of the nested PCR using primers D and C. The arrow shows the 307 bp band. Lane 1: negative control without target; lane 2: amplification of the 307 bp band using as target the MRC5 RNAs retrotranscribed with primer A; lane 3: amplification conducted with an EcoRI-digested 14C cDNA as target.

oligonucleotide A. In the two samples we show that the bands obtained after amplification using primers flanking the 147 bp fragment displayed the same size, suggesting that there is no intron lying within this fragment (data not shown).

#### Alternative splicing in the THRA locus

The two products that can be generated from the THRA locus from the two mRNAs species differ in their C-terminal regions. The two products may be generated through an alternative splicing event in exon 9 (Fig.5). The 410 amino acid product may be generated by a messenger RNA which begins at exon 1 and terminates at the 3' end of exon 9. The stop codon used for this product lies in the 3' region of exon 9 (251 bp) (see Fig.3). The 490 amino acid product may be generated from the RNA produced by an alternative splicing event that uses a splice donor site present at 128 bp after the beginning of exon 9 and the splice acceptor site of exon 10. The stop codon used to

terminate this product is located 364 bp downstream from the beginning of exon 10. This event would replace the 40 final amino acids of the hTHRA1 with 120 amino acids specific for hTHRA2.

#### A nuclear receptor gene in an inverted orientation overlaps with THRA at its 3' end

A region of homology with the human *ear-1* cDNA is positioned in an inverted orientation relative to THRA (11, 34). This region corresponds to the most 3' exon of the human *ear-1* gene (named F here) which overlaps with exon 10 of THRA over 1028 bp (263 nucleotides of coding region for both genes; see fig.2 and 3). The overlapping region does not extend in the 3' untranslated part of exon 9 (data not shown).

We have also located two others *ear-1* exons. The first (denoted 'F-1' in Fig.2) is immediately adjacent to the human exon F and

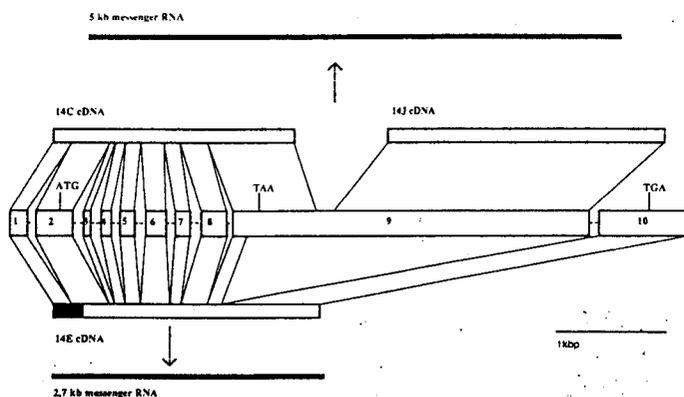


Fig. 5. Alternative splicing of the THRA gene. In exon 9 an alternative splice site is present and used to generate the 2.7 kb messenger RNA containing exon 10. The black boxes at the beginning of 14E show sequences absent from our clone but identified in the literature (20). The 5.0 kb messenger RNA contains sequences from exon 1 to 9. The 2.7 kb mRNA contains sequences from exon 1 to 8, the 128 first nucleotides of exon 9 and exon 10. The 5.0 kb RNA encodes the 410 amino acid product and the 2.7 kb RNA species encodes the 490 amino acid product. The black box at the beginning of the 14E cDNA represents missing sequences in our clone.

		* REV-ERB SPLICE ACCEPTOR											
REV-erb	200	TTTGAAGA	AAGGTGGT	TATCACCTAC	ATTGGCTCCA	GCGGATCCTC	CCCCAGCCGG	ACCAGCCAG	AGTCCCTCTA	TAGTGACAGC	TCCAATGGCA	GCTTCAGTC	
HR2	1	ccccctttcc	ttttgttctc	tttccaatgg	CATCACCTCC	ATTGGCTCCA	GTGGCTCCTC	CCCAAGCCGC	ACCAGCCCTG	AATCCCTCTA	TAGTGACAAC	TCCAATGGCA	GCTTCAGTC
EAR-1	647	ACAACAACA	CAGGTGGCGT	CATCACCTAC	ATTGGCTCCA	GTGGCTCCTC	CCCAAGCCGC	ACCAGCCCTG	AATCCCTCTA	TAGTGACAAC	TCCAATGGCA	GCTTCAGTC	
				I T Y	I G S	S G S	P S R	T S P	E S L Y	S D N	S N G	S F Q S	
				----- EAR-1 SPLICE ACCEPTOR									
REV-erb	309	CCTGACTCAA	GGTGTGCCA	CATACTTCCC	ACCATCACCT	ACTGGCTCCC	TCACCCAAGA	CCCTGCCCGC	TCATTTGGCA	CTGTGCCACC	CAGCCTCAGT	GATGATAGCT	CCCTTCTTC
HR2	121	CCTGACCCAA	GGCTGTGCCA	CCTACTTCCC	ACCATCCCCC	ACTGGCTCCC	TCACCCAAGA	CCCGTTGCC	TCCTTTGGGA	GCATTCACCC	CAGCCTGAGT	GATGAGGCT	CCCTTCTTC
EAR-1	756	CCTGACCCAA	GGCTGTGCCA	CCTACTTCCC	ACCATCCCCC	ACTGGCTCCC	TCACCCAAGA	CCCGCTGCC	TCCTTTGGGA	GCATTCACCC	CAGCCTGAGT	GATGAGGCT	CCCTTCTTC
		L T Q	G C P	T Y F P	P S P	T G S	L T Q D	P A R	S F G	S I P P	S L S	D D G	S P S S
REV-erb	429	TGCTTCATCA	TCATCATCATC	TTCTCTCCC	TCCTTCTATA	ATGGGAGGCC	CCCCGGAGT	CTACAAGTGG	CCATGGAAGA	CAGCAGCCGA	GTGTCCOCCA	GCAAGGGCAC	CAGCAACAT
HR2	241	--CT-CGTCT	TCCTCGTCGC	ATCTCTCCC	TCCTTCTGTG	ATGGGAGGCC	CCCTGGAGT	CTACAAGCGG	CCATGGAGGA	CAGCAOCCGA	GTGTCCOCCA	GCAAGAGCAC	CAGCAACAT
EAR-1	876	--CT-CATCT	TCCTCGTCGC	ATCTCTCCC	TCCTTCTATA	ATGGGAGGCC	CCCTGGAGT	CTACAAGTGG	CCATGGAGGA	CAGCAGCCGA	GTGTCCOCCA	GCAAGAGCAC	CAGCAACAT
		S S	S S S S	S S S	S F Y	N G S P	P G S	L Q V	A M E D	S S R	V S P	S K S T	S N I
REV-erb	549	T K L	N G M	V L L									
		CACTACCAAG	CTFAACGGCA	TGGTACT	575								
HR2	358	CACCAgtgag	tacttccagc	tgatattt	384								
EAR-1	996	CACCAAGCTG	AATGGCATGG	TGTACT	1022								
		T K L	N G M	V L L									
		* SPLICE DONOR											

Fig. 6. Sequence of the HR2 exon of *ear-1* compared with the human *ear-1* cDNA (11) and the rat *REV-erb* cDNA. The initiation codon used to generate the *REV-erb* gene product is underlined as is the splice donor site. Intronic sequences are in lower case letters (see also Ref 47).

corresponds to exon 7 of *REVerbA*, the rat homologue of *ear-1* (34). This exon is separated by 50 bp from the end of the THRA exon 10. The second (denoted HR2 for 'Homologous to Rat exon 2' in Fig.2) corresponds to the final exon of the A/B domain of *ear-1*. This region is homologous to exon 2 of the *REVerbA* described in rats (34) (see Fig.6). The precise location of the splice acceptor site of this exon has not been determined since there is no sequence reminiscent of a canonical splice acceptor consensus in this region (Fig.6)

## DISCUSSION

### Genomic organization of THRA

In these studies, we determined the exon/intron organization of the human THRA gene. The THRA gene is split into 10 exons distributed along 27 kbp of chromosome 17 and overlaps in its 3' region with the structurally related *ear-1* gene. In addition, we provide evidence that the 410 amino acid product of THRA is encoded by a large messenger RNA which contains a lengthy 3' untranslated region.

The four phage inserts aligned to determine this structure were isolated from three different genomic libraries (see Material and Methods). In these libraries (derived from human placental, ML3 and IGROV-2 cells DNA) no rearrangement of the THRA gene has been described. By extensive endonuclease mapping and sequencing, we determined that the phage sequences overlap and have no broad differences in their common regions. Moreover, the structure we determined in this paper is confirmed by previous results obtained in the study of an EcoRI RFLP site present in the human THRA gene (44).

We have cloned two cDNAs, 14C and 14J, both of which hybridize to a 5.0 kb mRNA species. We demonstrate by PCR experiments that the two cDNAs may be attributed to the same messenger RNA. Indeed, after specific reverse transcription of 14J 5' related sequences we were able to amplify 14C 3'-related sequences. Thus, we conclude that the 410 amino acid product of THRA is translated from an RNA of 5.0 kb which contains a long 3' untranslated region of at least 3106 nucleotides. This feature is common among nuclear receptor genes. For example, the human glucocorticoid receptor contains a 3' untranslated

region of 2.3 kb (45) and the rat THRB gene possesses a similar 2.7 kb region. As numerous reports have suggested that sequences present in the 3' non coding parts of various mRNAs may play a regulatory role in the turnover of these molecules (see 44 and references therein), it is tempting to speculate that secondary structures in the long, 3'-untranslated region of the THRA 5.0 kb mRNA may positively or negatively modulate its degradation.

The putative polyadenylation sites of 14J do not correspond to any known consensus sequence for polyadenylation (47). We nonetheless feel we have identified the authentic terminus of the RNA because the corresponding site in the genomic DNA, contains no A-rich sequence that may be used for internal priming during library construction. The role of these alternative polyadenylation sites on the stability or on the translation rates of the various THRA mRNAs remains to be addressed (46, 47).

### The rat and human *ear-1* genes

Nuclear receptor genes are generally highly conserved during evolution especially among mammals. However, whereas the initiation codon of the rat *REV-erbA* gene is located within exon 2 which corresponds to our human HR2 exon (34), the ATG initiation codon proposed for the human sequence lies in the preceding exon (11, 49). Indeed, the human product described (11) is 106 amino acid longer at its 5' end than the rat product (34) and the 5' portions of the two published cDNAs have no homology upstream of the HR2 exon. This observation is interesting since, to our knowledge, this represents the major variation between rat and human nuclear receptor genes in a domain (A/B) which is generally implicated in trans-activation.

### Comparison with other nuclear receptor genes

The genomic organization of the human THRA gene shares similar structural features with its avian homologue (33). The structure of the two genes is similar within their coding region, with the initiation codon at the end of an exon encoding 17 amino acids in both cases (the entire length of this initiation exon is 202 nucleotides for chicken exon 3, and 350 nucleotides for human exon 2). The following 5 exons are of similar size for the two species and their boundaries are located at similar positions. No sequencing data of the 3' region of the chicken gene are available in the literature, and RNase mapping experiments suggest that the equivalent of human exon 10 does not exist or is not transcribed in the chicken (13). The major differences between the two genes are as follows: (i) The 5' untranslated regions share no homology and are split into three exons in the chicken and two in the human genes. Nevertheless, in the chicken as in humans, the most 5' exon is far from the next one (4.5 kbp in chickens, 11 kbp in humans) (ii) The size of the 5' two-thirds of the gene is 21 kbp for humans and 12 kbp for chickens, the difference being a consequence of the relative sizes of introns. For example, the intron lying between the two zinc finger-coding exons is 80 bp in the chicken and 5 kbp in the human gene. Since the complete chicken gene has yet to be cloned, we cannot compare the two entire genes. A comparison is also possible with the partially determined organization of the 3' region of the rat THRA gene. In this region, human and rat genes possess a nearly identical organization. In particular, the 263 nucleotide overlap found between the coding regions of *ear-1* and THRA genes is perfectly conserved between rats and humans (11, 34). This argues in favor of an important role for the unusual overlapping structure.

When THRA is compared with other nuclear receptor genes of the steroid hormone receptor sub-family, several observations

can be made. First, the steroid hormone receptor genes, such as the chicken progesterone receptor (29), the human estrogen receptor genes (30) or the human androgen receptor gene (31, 33) appear longer than human THRA. For instance, the human estrogen receptor gene is more than 140 kbp in length (30). Second, like all known nuclear receptor genes except *nur77* (a growth factor inducible 'orphan receptor' gene; see 43), the two zinc finger domains of THRA are encoded by separate exons. In both THRA and retinoic acid receptors genes an intron is located one amino acid after the last cysteine residue of the first zinc finger C1 (2, 5, 35). In contrast, the intron is located in a different position in the middle of the region lying between the two zinc fingers in the steroid hormones receptor genes and *nur77* (2, 29, 30, 43). Third, the A/B domain of THRA is split in two exons (exons 2 and 3), whereas in steroid hormone receptors gene (human estrogen receptor and chicken progesterone receptor genes) and *nur77*, this domain is encoded by a single exon (29, 30, 43). Fourth, the organization of the 5' region of the thyroid hormone/retinoic acid receptor genes seems variable: the RAR $\alpha$  and  $\gamma$  5' region are split into several exons (36, 50) like THRA, whereas the same portions (*i.e.* 5' untranslated region and A/B domain) of RAR $\beta$  are composed of only one exon (V. Giguere, personal communication). Finally, as pointed out by Ponglikitmongkol et al. (30), there is an intron between domain D and E in THRA which is not present in estrogen or progesterone receptor genes. In these latter genes, the exon encoding domain D exhibits a mosaic structure encoding the end of C2, D and the first 63 amino acids of E. This difference may be explained by the probable loss in the steroid hormone receptor sub-family, of an intron which is still present in THRA genes. A comparison of the genomic structures of various nuclear receptor genes reveals a common mode of organization which appears to have evolved through gene duplication from a common ancestral progenitor to give rise to the three sub-families presently known: steroid hormone receptor genes, thyroid hormone/retinoic acid receptor genes and the *nur77* family. Additional work on other nuclear receptor genes, including determination of their genomic organization, is needed to better understand the evolutionary history of this superfamily of genes.

### ACKNOWLEDGEMENTS

We thank Eli Canaani, John Toole and Tom Maniatis for providing human cDNA or genomic libraries, Jean Coll and Jean Marc Vanacker for helpful discussions, Brigitte Quatannens for the excellent technical assistance, Marie-Christine Bouchez and Nicole Devassine for piloting the word processor apparatus and Thierry Dupressoir for help in figure preparation. We are also grateful to Ruth Globus and Philip Boucher for critical reading of the manuscript.

V.L. holds a fellowship from the French Ministère de la Recherche et de la Technologie. We thank the Association de la Recherche contre le Cancer and Ligue Nationale contre le Cancer for financial support.

### REFERENCES

1. Evans, R.M. (1988) *Science*, **240**, 889-895.
2. Green, S. and Chambon, P. (1988) *TIG*, **4**, 309-314.
3. Beato, M. (1989) *Cell*, **56**, 335-344.
4. Petkovich, M., Brand, N.J., Krust, A. and Chambon, P. (1987) *Nature*, **330**, 444-450.
5. Brand, N., Petkovich, M., Krust, A., Chambon, P., de Thé, H., Marchio, A., Tiollais, P. and Dejean, A. (1988) *Nature*, **332**, 850-853.

1112 *Nucleic Acids Research*, Vol. 19, No. 5

6. Zelent, A., Krust, A., Petkovich, M., Kastner, P. and Chambon, P. (1989) *Nature*, **339**, 714–717.
7. Krust, A., Kastner, P., Petkovich, M., Zelent, A. and Chambon, P. (1989) *Proc. Natl. Acad. Sci. USA*, **86**, 5310–5314.
8. Weinberger, C., Thompson, C.C., Ong, E.S., Lebo, R., Gruol, D.J. and Evans, R.M. (1986) *Nature*, **324**, 641–646.
9. Spurr, N.K., Solomon, E., Jansson, M., Sheer, D., Goodfellow, P.N., Bodmer, W.F. and Vennström, B. (1984) *EMBO J.*, **3**, 159–163.
10. Miyajima, N., Kadowaki, Y., Fukushige, S., Shimizu, S., Semba, K., Yamanashi, Y., Matsubara, K., Toyoshima, K. and Yamamoto, T. (1988) *Nucl. Acids Res.*, **16**, 11057–11074.
11. Miyajima, N., Horiuchi, R., Shibuya, Y., Fukushige, S., Matsubara, K., Toyoshima, K. and Yamamoto, T. (1989) *Cell*, **57**, 31–39.
12. Sap, J., Munoz, A., Damm, K., Goldberg, Y., Ghysdael, J., Leutz, A., Beug, H. and Vennström, B. (1986) *Nature*, **324**, 635–640.
13. Forrest, D., Sjöberg, M. and Vennström, B. (1990) *EMBO J.*, **9**, 1519–1528.
14. Saule, S., Roussel, M., Lagrou, C. and Stéhelin, D. (1981) *J. Virol.*, **38**, 409–419.
15. Debuire, B., Henry, C., Benaissa, M., Biserte, G., Claverie, J.M., Saule, S., Martin, P. and Stéhelin, D. (1984) *Science*, **224**, 1456–1459.
16. Kahn, P., Frykberg, L., Brady, C., Stanley, I., Beug, H., Vennström, B. and Graf, T. (1986) *Cell*, **45**, 349–356.
17. Gandrillon, O., Jurdic, P., Benchaibi, M., Xiao, J.H., Ghysdael, J. and Samarut, J. (1987) *Cell*, **49**, 687–697.
18. Gandrillon, O., Jurdic, P., Pain, B., Desbois, C., Madjar, J.J., Moscovici, M.G., Moscovici, C. and Samarut, J. (1989) *Cell*, **58**, 115–121.
19. Damm, K., Thompson, C.C. and Evans, R.M. (1989) *Nature*, **339**, 593–597.
20. Benbrook, D. and Pfahl, M. (1987) *Science*, **238**, 788–791.
21. Thompson, C.C., Weinberger, C., Lebo, R. and Evans, R.M. (1987) *Science*, **237**, 1610–1614.
22. Koenig, R.J., Warne, R.L., Brent, G.A., Harney, J.W., Reed Larsen, P. and Moore, D.D. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 5031–5035.
23. Murray, M.B., Zilz, N.D., McCreary, N.L., MacDonald, M.J. and Towle, H.C. (1988) *J. Biol. Chem.* **263**, 12770–12777.
24. Mitsuhashi, T., Tennyson, G.E. and Nikodem, V.M. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 5804–5808.
25. Nakai, A., Sakurai, A., Bell, G.I. and DeGroot, L.J. (1988) *Molec. Endocrinol.*, **2**, 1087–1092.
26. Izumo, S. and Madhavi, V. (1988) *Nature*, **334**, 539–543.
27. Koenig, R.J., Lazar, M.A., Hodin, R.A., Brent, G.A., Reed Larsen, P., Chin, W.W. and Moore, D.D. (1989) *Nature*, **337**, 659–661.
28. Hodin, R.A., Lazar, M.A., Wintman, B.L., Darling, D.S., Koenig, R.J., Reed Larsen, P., Moore, D.D. and Chin, W.W. (1989) *Science*, **244**, 76–79.
29. Huckaby, C.S., Conneely, O.M., Beattie, W.G., Dobson, A.D.W., Tsai, M.J. and O'Malley, B.W. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 8380–8384.
30. Ponglikitmongkol, M., Green, S. and Chambon, P. (1988) *The EMBO Journal*, **7**, 3385–3388.
31. Kuiper, G.G.J.M., Faber, P.W., van Rooij, H.C.J., van der Korput, J.A.G.M., Ris-Stalpers, C., Klaassen, P., Trapman, J. and Brinkmann, A.O. (1989). *J. Mol. Endocr.*, **2**, R1–R4.
32. Tilley, W.D., Marcelli, M. and McPhaul, M.J. (1990). *J. Biol. Chem.* **265**, 13776–13781.
33. Zahraoui, A. and Cuny, G. (1987) *Eur. J. Biochem.*, **166**, 63–69.
34. Lazar, M.A., Hodin, R.A., Darling, D.S. and Chin, W.W. (1989) *Mol. Cell. Biol.* **9**, 1128–1136.
35. Dejean, A., Bougueleret, L., Grzeschik, K.H. and Tiollais, P. (1986), *Nature*, **322**, 70–72.
36. Kastner, P., Krust, A., Mendelsohn, C., Garnier, J.M., Zelent, A., Leroy, P., Staub, A. and Chambon, P. (1990) *Proc. Natl. Acad. Sci. USA*, **87**, 2700–2704.
37. Vennstrom, B., Fanshier, L., Moscovici, C. and Bishop, J.M. (1980) *J. Virol.*, **336**, 575–585.
38. Lawn, R.M., Fritsch, E.F., Parker, R.C., Blake, G. and Maniatis, T. (1978) *Cell*, **15**, 1157–1174.
39. Maniatis, T., Fritsch, E.F. and Sambrook, J. (1989) *Molecular Cloning: A laboratory Manual*, 2nd edition, Cold Spring Harbor Laboratory, New York
40. Maxam, A.M. and Gilbert, W. (1980) *Methods Enzymol.* **65**, 499–560.
41. Wang, A.M., Doyle, M.V. and Mark, D.F. (1989) *Proc. Natl. Acad. Sci. USA*, **86**, 9717–9721.
42. Breathnach, R. and Chambon, P. (1981) *Annu. Rev. Biochem.*, **50**, 349–383.
43. Ryseck, R.P., Macdonald-Bravo, H., Mattei, M.G., Ruppert, S. and Bravo, R. (1989) *EMBO J.*, **8**, 3327–3335.
44. Mathieu-Mahul, D., Xu, D.Q., Saule, S., Lidereau, R., Galibert, F., Berger, R., Mauchauffé, M. and Larsen, C.J. (1985) *Hum. Genet.*, **71**, 41–44.
45. Hollenberg, S.M., Weinberger, C., Ong, E.S., Cerelli, G., Oro, A., Lebo, R., Thompson, B., Rosenfeld, M.G. and Evans, R.M. (1985) *Nature*, **318**, 635–641.
46. Jackson, R.J. and Standart, N. (1990) *Cell*, **62**, 15–24.
47. Wickens, M. (1990) *Trends Biol. Sci.*, **15**, 277–281.
48. Higuchi, R. and Kwok, S. (1989) *Nature*, **339**, 237–238.
49. Lazar, M.A., Jones, K.E. and Chin, N. (1990) *DNA*, **9**, 77–83.
50. Brand, N.J., Petkovich, M. and Chambon, P. (1990) *Nucl. Acids Res.*, **18**, 6799–6806

## RESUME DES RESULTATS DE L'ARTICLE 1

Notre travail a permis d'obtenir deux types d'informations :

1) Le gène c-erbA-1 humain est épissé en 10 exons repartis sur environ 27 kb d'ADN génomique (figure 37). Le premier exon, de petite taille (160 pb) est entièrement non codant et il est très riche en bases G et C. Le deuxième exon contient le codon d'initiation situé à la bordure 3' de cet exon (17 codons avant le site donneur d'épissage). Le troisième exon est le plus petit du gène et, avec l'exon 2, il code le domaine A/B. Les exons 4 et 5 codent pour les doigts de zinc du récepteur alors que l'exon 6, qui contient les parties codant le domaine D, code également pour l'un des signaux de translocation nucléaire localisé chez v-erbA (séquence humaine RERRRK; Boucher et al., 1988; Boucher et Privalsky, 1990). Cet exon code aussi le début du domaine E. Les trois exons 7, 8 et 9, codent le domaine E. Le domaine L1 est codé par les exons 6 et 7, tandis que le domaine Ti est codé par les exons 7 et 8. Le grand domaine de dimérisation est codé par les exons 8 et 9. Le domaine L2, co-responsable avec L1 de la fixation du ligand, est codé par la fin de la partie codante de l'exon 9. Il faut noter que l'exon 9 est, de très loin, le plus grand du gène, puisqu'il comprend une partie 3' non codante de 3106 nucléotides. Enfin, le dixième et dernier exon est spécifique d'une isoforme particulière du récepteur et sa fonction est largement inconnue.

Le gène c-erbA-1 peut coder deux isoformes majeures à partir de deux ARN de tailles très différentes. Un ARN de 5 kb, correspondant aux exons 1 à 9, va coder un produit de 410 acides aminés (figure 38). Ce produit est un récepteur normal des hormones thyroïdiennes qui fixe la T3, se lie spécifiquement à l'ADN et transactive l'expression d'un gène cible (Koenig et al., 1989 ; Izumo et Mahdavi, 1988). Un autre ARN, de 2,6 kb, correspond aux exons 1 à 8, puis aux 127 premiers nucléotides de l'exon 9. A ce niveau, un site donneur d'épissage est reconnu, et il y a jonction à l'exon 10. Le produit de 490 acides aminés, codé par cet ARN, n'est plus capable de fixer l'hormone parce que le domaine L2 qui, avec L1, est responsable de cette fixation, est absent et il est remplacé par des acides aminés codés par l'exon 10. Ce produit se comporte donc en permanence comme le récepteur normal en absence d'hormone : il réprime l'expression des gènes cibles. De plus, comme cette molécule est capable de dimériser avec les récepteurs normaux codés par c-erbA-1 ou c-erbA-2 (et même avec les récepteurs de l'acide rétinoïque) il va se comporter comme un régulateur négatif dominant de ces molécules. Cet épissage alternatif est donc particulier à deux titres. Premièrement, il donne deux produits, qui ont une action biologique opposée : un récepteur des hormones thyroïdiennes, et un inhibiteur du récepteur des hormones thyroïdiennes. Deuxièmement, il est étonnant d'un point de vue mécanistique parce qu'il n'y a pas ajout ni suppression d'un exon, mais reconnaissance d'un site "cryptique" au milieu de l'exon 9, et ensuite terminaison différentielle de la transcription. En effet, chaque ARN messager se termine par un

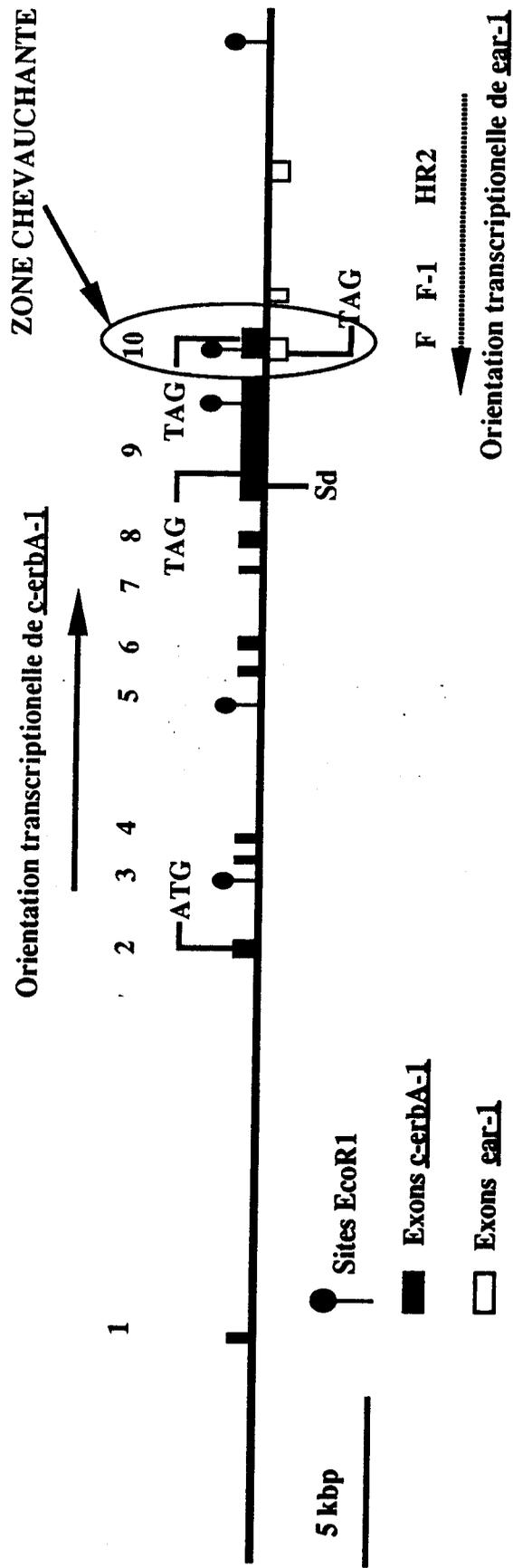


Figure 37 : Organisation génomique de *c-erbA-1*. La zone de chevauchement avec l'exon F de *ear-1* a été entourée.

signal de polyadénylation différent. Le messenger de 2,6 kb a un signal relativement classique (CATAAAA alors que le consensus est AATAAAA) alors que le messenger de 5 kb n'a pas de signal consensus évident. Les seuls qui s'en rapprochent sont ATTTTAAA ou ATATTTTAA. Ces sites différents sont peut être des points de régulation supplémentaires de l'épissage. De même, les longues séquences non traduites 3' ont peut être un rôle dans la régulation de la stabilité des ARN messagers. Il faut noter que la plupart des récepteurs nucléaires sont codés par des ARN messagers possédant de très longues séquences 3' non codantes.

Il est probable que le locus c-erbA-1 puisse donner lieu à d'autres épissages alternatifs (Mitsubishi et al., 1988 ; Mitsubishi et Nikodem, 1989 ; Izumo et al., 1990) comme par exemple celui qui donne la forme mineure rTR $\alpha$ VII décrite par l'équipe de Vera Nikodem. Ce produit ressemble beaucoup à la forme  $\alpha$ 2 codée par l'ARN de 2,6 kb, mais dans ce cas le site accepteur d'épissage reconnu dans l'exon 10 est situé 117 nucléotides en aval. Ceci donne un produit final de 451 acides aminés et non pas de 490, comme le produit classique de l'exon 10. Ce produit ne fixe pas les hormones thyroïdiennes. Ce n'est pas étonnant, puisque dans ce cas aussi, le domaine L2 a été remplacé par des séquences différentes provenant de l'exon 10 (figures 37 et 38).

Par des expériences de protection à la RNase, d'autres épissages alternatifs ont été détectés chez le rat, notamment dans la partie 5' du gène (Izumo et al., 1990). Ces résultats vont dans le même sens que des résultats préliminaires obtenus au laboratoire par amplification génique sur le gène humain. Toutes ces expériences suggèrent que la partie 5' de ce gène est soumise à certains épissages alternatifs. Il est intéressant de rappeler que les exons 1 et 2 sont séparés, chez l'homme, d'environ 11kb et que dans cette région nous avons observé des zones d'homologie faible avec l'exon 1 (65% sur 160 nucléotides très riches en bases G et C). En bref, il existe peut-être des exons alternatifs dans cette région, exons qui pourraient donner des transcrits ayant des parties 5' non codantes différentes, voire même des domaines A/B différents. Une analyse détaillée de la transcription de la partie 5' du gène c-erbA-1 est donc tout à fait nécessaire.

2) Le gène c-erbA-1 chevauche un autre gène, appelé ear-1, placé dans une orientation transcriptionnelle inverse, et appartenant, lui aussi, à la famille des récepteurs nucléaires (figure 37). Cette structure très inhabituelle, où les derniers exons connus de chaque gène se chevauchent sur 263 nucléotides, est unique chez les eucaryotes (Lazar et al., 1989a ; Miyajima et al., 1989). En effet si l'on connaît d'autres gènes imbriqués les uns dans les autres, ou même des gènes entièrement codés dans les introns d'autres gènes plus grands, c'est à notre connaissance le seul cas décrit où les deux gènes impliqués dans un tel complexe sont étroitement apparentés (Spencer et al., 1986 ; Williams et Fried, 1986 ; Morel et al., 1989). Cette imbrication des deux gènes est conservée chez les deux espèces de mammifères étudiées (le rat et l'homme ; Lazar et al., 1989a ; Miyajima et al., 1989).

En plus de l'exon impliqué dans le chevauchement avec l'exon 10 de c-erbA-1 que nous avons appelé "F" (car c'est l'exon final de ear-1), nous avons localisé deux autres exons correspondants à ce gène. Signalons d'ailleurs que nous n'avons pas numéroté les exons de ear-1,

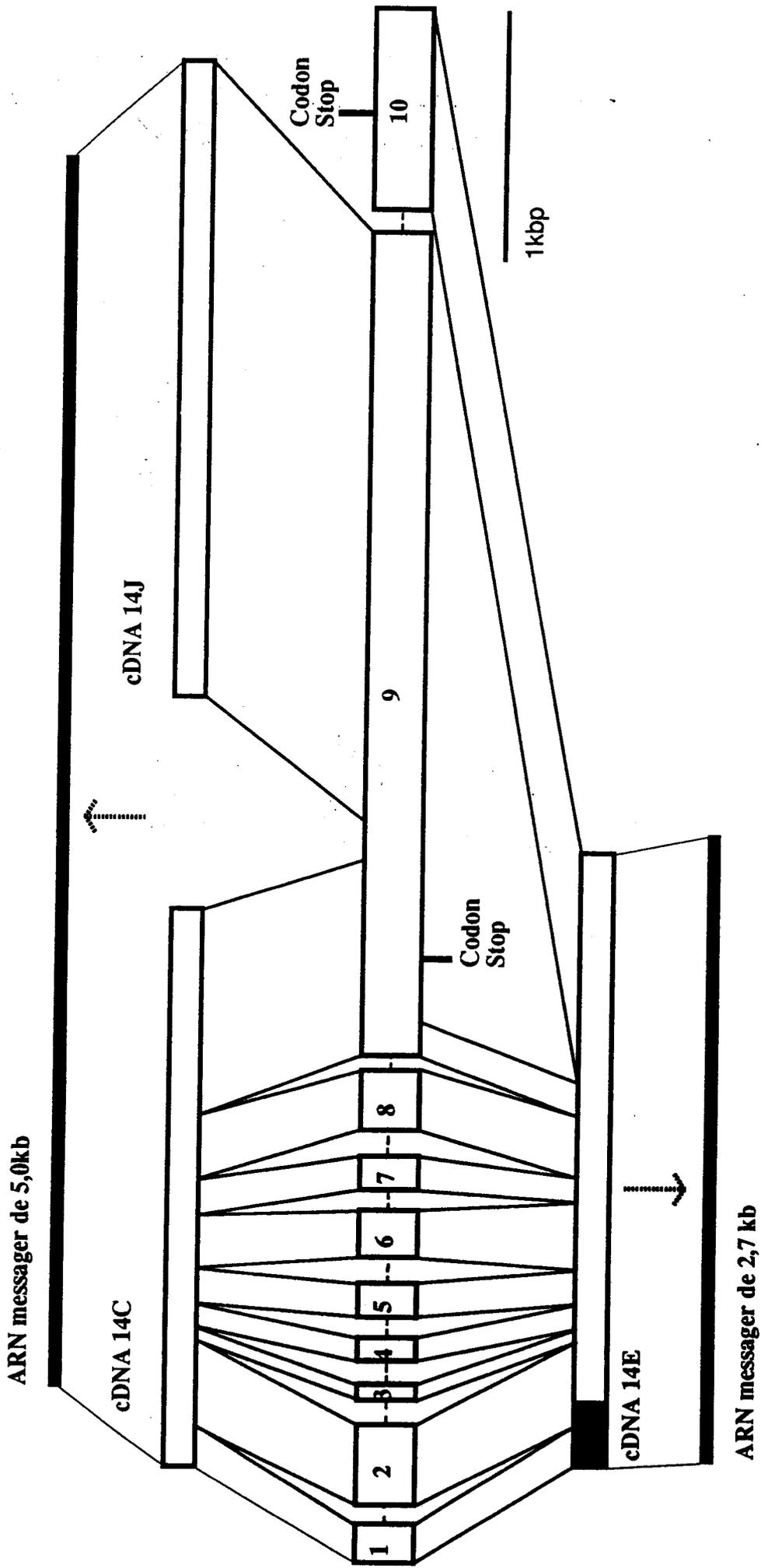


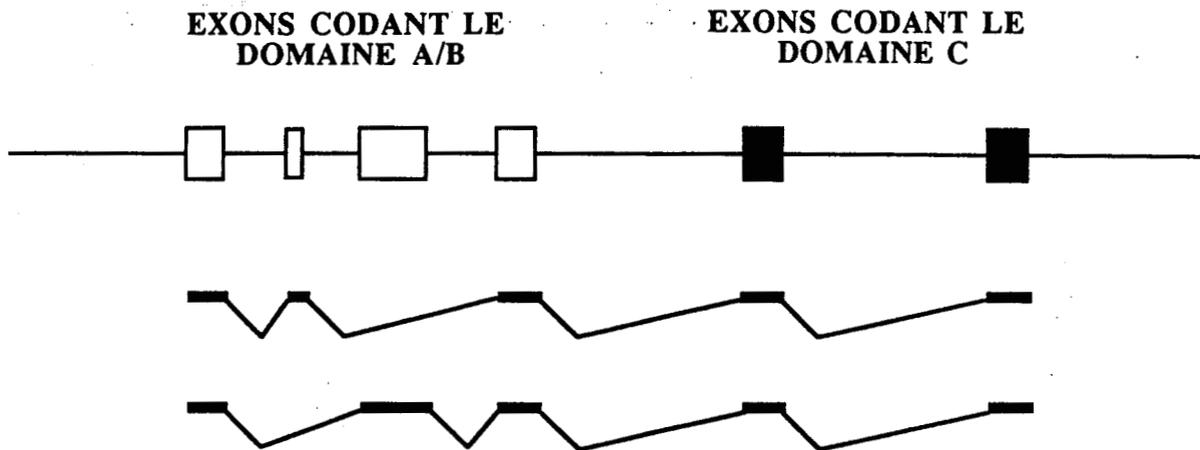
Figure 38: Les deux isoformes majeures de c-erbA-1 proviennent d'un événement d'épissage alternatif situé dans l'exon 9.



car l'organisation génomique complète de ce gène n'est pas déterminée chez l'homme. L'un de ces exon est appelé "F-1" car il est adjacent à l'exon F (figure 37). Il n'est séparé de l'extrémité de l'exon 10 de c-erbA-1 que par 50 paires de bases. L'autre exon de ear-1 est situé beaucoup plus loin en 5' et semble être l'homologue de l'exon 2 du gène ear-1 de rat (Lazar et al., 1989). Nous l'avons donc appelé HR2 pour "homologous to rat exon 2". Curieusement, les gènes ear-1 de rat et d'homme sont différents à ce niveau (Lazar et al., 1990a). En effet, le codon d'initiation du gène ear-1 humain est situé dans un exon localisé en amont de HR2 et dont nous ignorons la position exacte (Miyajima et al., 1989). Par contre, le codon d'initiation du gène ear-1 de rat se trouve au niveau de l'exon 2 (Lazar et al., 1989a ; 1990a). Si on compare les séquences des exons 2 de rat et HR2 de l'homme, on observe une forte homologie, y compris en amont de l'ATG de rat (qui est conservé chez l'homme). Cependant, il existe dans la séquence du gène de rat des mutations ponctuelles qui décalent le cadre de lecture. Il n'est pas très facile, à l'heure actuelle, d'expliquer ces résultats. Une erreur de séquence faite par un des laboratoires ayant séquencé cette portion semble être exclue, puisque les mêmes séquences ont été obtenues et confirmées par les trois laboratoires impliqués (Toyoshima, au Japon, Chin à Boston et notre équipe à Lille). On peut donc simplement constater une forte divergence dans le domaine A/B de l'homme et de rat qui constituerait l'une des divergences les plus fortes observées dans un récepteur nucléaire chez deux espèces de mammifères. Une deuxième hypothèse est que, là encore, comme pour les récepteurs de l'acide rétinoïque, il existe différentes isoformes 5' de ce gène, et que les différentes équipes ont isolé des isoformes différentes (Lazar et al., 1990a). Un argument en défaveur de cette hypothèse est le fait qu'il existe une forte homologie entre l'exon 2 de rat et l'exon HR2 de l'homme, ce qui suggère que ces deux exons sont bien homologues.

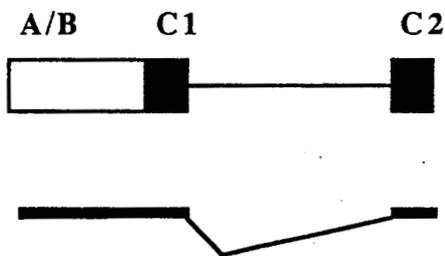
La connaissance de la structure génomique du gène c-erbA-1 humain permet de comparer cette structure avec celle des autres récepteurs nucléaires connus. Outre l'organisation génomique des récepteurs des oestrogènes et de la progestérone (Huckaby et al., 1987; Ponglikitmongkol et al., 1988), on connaît actuellement celles des récepteurs des androgènes (Kuiper et al., 1989), de la vitamine D (Hughes et al., 1988), des hormones thyroïdiennes de type  $\beta$  (c-erbA-2 ; Sakurai et al., 1990), de l'acide rétinoïque (RAR $\gamma$  ; Lehman et al., 1991) et du récepteur orphelin NGF1B (Ryseck et al., 1989). En outre, un certain nombre d'organisations partielles sont déterminées par exemple pour le gène ear-1, pour les gènes des récepteurs de l'acide rétinoïque ou pour le gène c-erbA-1 de poulet (Lazar et al., 1989a ; Miyajima et al., 1989 ; Kastner et al., 1990 ; Leroy et al., 1991b ; Zelent et al., 1991 ; Zahraoui et Cuny, 1987). Nous avons montré (voir article II) que, outre le domaine C, la partie centrale du domaine E, constituée des sous-domaines Ti et DM, était conservée chez tous les récepteurs, sauf pour les gènes de drosophile du groupe *knirps* (Pankratz et Jäckle, 1990). De plus, pour les gènes des récepteurs de l'acide rétinoïque et des hormones thyroïdiennes, la forte conservation existant entre les différents récepteurs permet une comparaison plus détaillée.

## PREMIER GROUPE : THR, RAR, EAR



Domaine A/B codé par plusieurs exons variables ; nombreuses isoformes 5'

## DEUXIEME GROUPE : NGF1B



Domaine A/B codé par un seul exon

Exon codant C1 fusionné avec l'exon codant le domaine A/B

## TROISIEME GROUPE : RECEPTEURS DES STEROIDES



Domaine A/B codé par un seul exon

Exon codant le domaine A/B séparé de l'exon codant C1

Figure 39 : Classification des récepteurs nucléaires en fonction de l'organisation génomique de leur partie 5' qui code les domaines A/B.

### 1) Le domaine A/B et les parties 5' non traduites.

Les récepteurs nucléaires se divisent en trois groupes, lorsque l'on examine l'organisation génomique de leurs parties 5' non traduites (figure 39).

- Dans le **premier groupe**, on trouve les récepteurs des hormones thyroïdiennes et de l'acide rétinoïque, mais également le récepteur orphelin ear-1 (Zahraoui et Cuny, 1987 ; Kastner et al., 1990 ; Leroy et al., 1991 ; Zelent et al., 1991 ; Lehman et al., 1991 ; Lazar et al., 1989a ; Miyajima et al., 1989). Chez ces gènes, il existe plusieurs exons en amont de celui codant le doigt de zinc C1. Le domaine A/B est donc codé par différents exons. De plus, il y a dans de nombreux cas différentes isoformes alternatives en 5'.

Il est intéressant d'observer de plus près l'organisation des exons 5' de cette sous-famille. En effet, de nombreux auteurs ont remarqué que chez les récepteurs de l'acide rétinoïque, les différentes isoformes 5' se raboutent sur la zone commune, au même endroit. Le même phénomène est vrai pour le gène c-erbA-2 chez le xénope et chez l'homme (Hodin et al., 1989 ; Yaoita et al., 1990). Ces positions ne sont cependant pas strictement conservées entre les récepteurs de l'acide rétinoïque et ceux des hormones thyroïdiennes. En effet, pour les récepteurs de l'acide rétinoïque, c'est au milieu du domaine A/B (à 26 acides aminés de la première cystéine du doigt C1) que tous les transcrits deviennent communs tandis que, chez c-erbA-2, c'est plus près du domaine C (à 13 acides aminés de la première cystéine de C1). Mais dans ces deux cas la jonction se fait avec l'exon codant le premier doigt de zinc. Toute la région 5' est donc très variable chez ces récepteurs et, dans les deux groupes (THR et RAR), la frontière entre cette zone variable et le domaine C n'est pas tout à fait identique en terme de séquence, mais est conservée au niveau des exons. L'exon codant le doigt C1 est plus étendu en 5' chez les RAR que chez les THR. Il reste sûrement des isoformes 5' à découvrir (en particulier pour les récepteurs des hormones thyroïdiennes) et elles devraient permettre de réaliser une meilleure comparaison de ces régions variables.

- Dans le **deuxième groupe**, on trouve le récepteur orphelin NGF1B (Ryseck et al., 1989) chez qui le domaine A/B est codé par un seul exon et, où cet exon code aussi le premier doigt de zinc. Mentionnons une autre différence: il existe un exon 5' non codant situé juste en aval du promoteur. Pour l'instant, à notre connaissance, le gène NGF1B fait seul partie de ce groupe. C'est donc peut être un cas particulier ou, éventuellement, le prototype d'une sous-famille de récepteurs encore à décrire.

- Dans le **troisième groupe**, on trouve les récepteurs des stéroïdes: oestrogènes, androgènes, progestérone et, sans doute, vitamine D (Ponglikitmongkol et al., 1988 ; Kuiper et al., 1989 ; Huckaby et al., 1987 ; Hughes et al., 1988). Chez tous ces gènes, le domaine A/B, qui est la plupart du temps très long (sauf dans le cas du récepteur de la vitamine D), est codé par un seul exon, le premier du gène. Cet exon est séparé des exons codant le domaine C. Il semble ne pas y avoir d'exon 5' non codant.

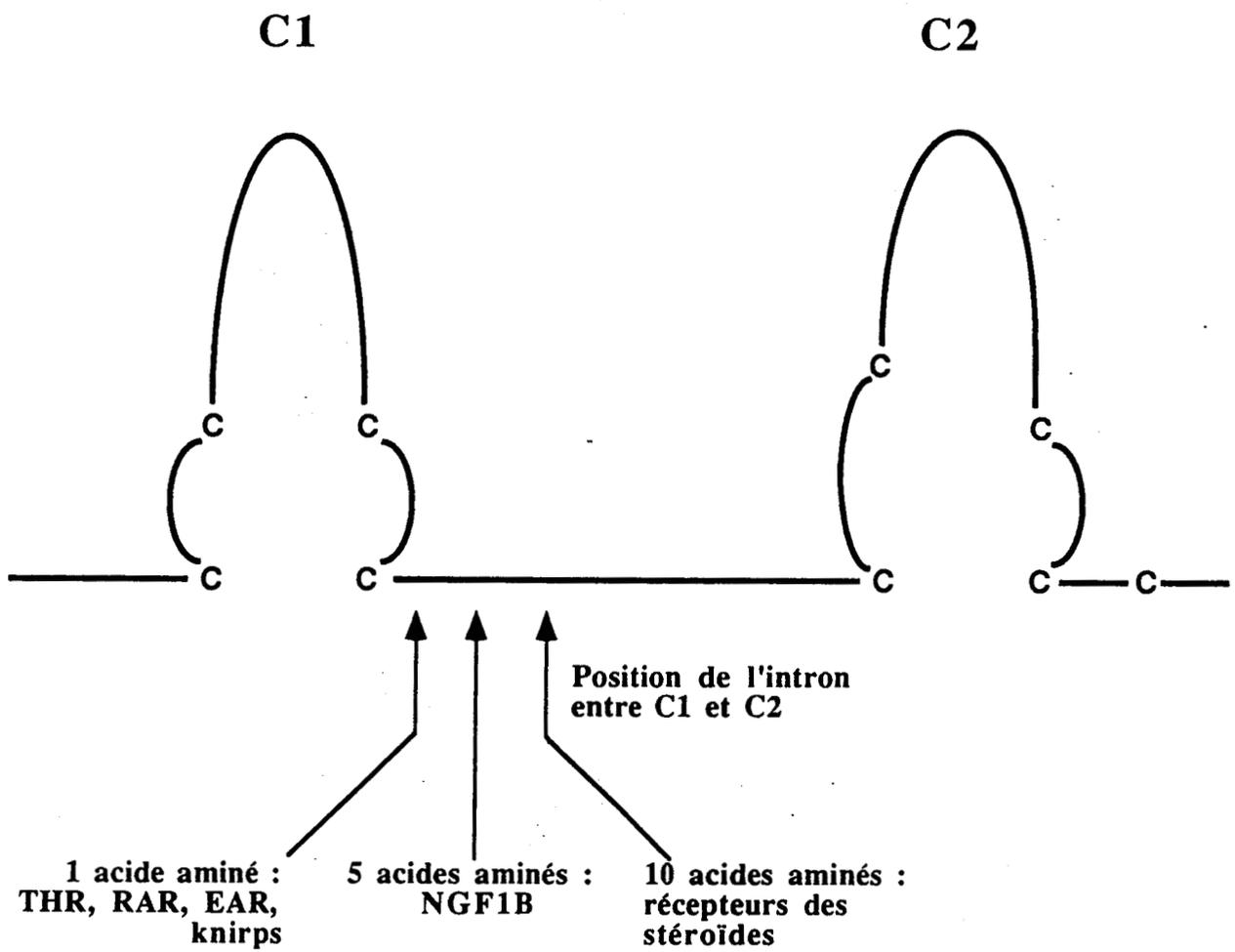


Figure 40 : Classification des récepteurs nucléaires en fonction de la position de l'intron entre les doigts de zinc C1 et C2.

## 2) Le domaine C.

Chez tous les récepteurs, les doigts C1 et C2 sont codés par des exons différents. La position 3' du domaine est remarquablement conservée puisque l'exon contenant C2 s'arrête 3 acides aminés après la séquence glycine-méthionine qui marque la fin du domaine C. On peut cependant séparer les récepteurs en trois groupes en fonction de la position de l'intron qui se trouve entre l'exon codant C1 et celui codant C2 (figure 40).

- Pour les récepteurs THR, RAR, ear-1 et également chez les gènes de drosophile du groupe *knirps* (*knirps*, *knirps-related* et *embryonic gonad*; Nauber et al., 1988 ; Oro et al., 1988 ; Rothe et al., 1989), cet intron se situe juste après la dernière cystéine de C1.

- Pour le gène NGF1B, cet intron se situe 5 acides aminés après la dernière cystéine de C1.

- Pour les récepteurs des stéroïdes (y compris le récepteur de la vitamine D) cet intron se situe 10 acides aminés après la dernière cystéine de C1.

Avant la découverte de la position intermédiaire de l'intron situé entre C1 et C2 dans le gène NGF1B, certains auteurs avaient proposé que le domaine C ancestral ait été codé par 3 exons : C1, un exon tout petit (10 AA) et C2 (Ponglikitmongkol et al., 1988). Selon les cas, c'est l'un ou l'autre des introns qui aurait été perdu au cours de l'évolution. Si on veut conserver ce modèle avec NGF1B, il faut imaginer qu'il y ait eu 4 exons primitifs dont deux de quinze nucléotides environ. Ce modèle, qui considère implicitement que les exons sont peu à peu perdus au cours de l'évolution, sera discuté plus loin, à la lumière des données évolutives obtenues (voir article II et discussion).

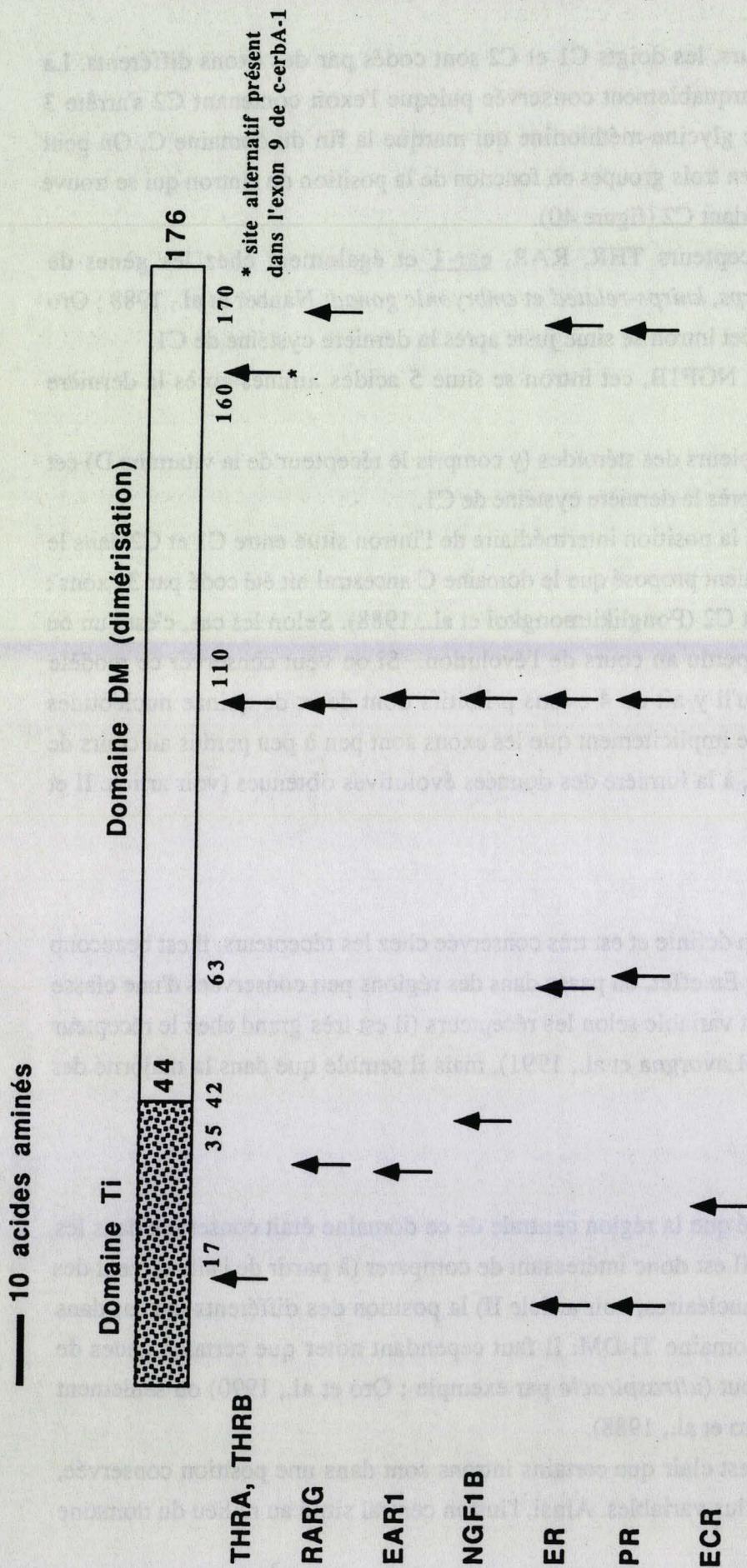
## 3) Le domaine D.

Si sa frontière 5' est bien définie et est très conservée chez les récepteurs, il est beaucoup plus difficile de situer sa limite 3'. En effet, on passe dans des régions peu conservées d'une classe de récepteurs à l'autre. Sa taille est variable selon les récepteurs (il est très grand chez le récepteur orphelin de Drosophile FTZ-F1 ; Lavorgna et al., 1991), mais il semble que dans la majorité des cas il soit codé par un seul exon.

## 4) Le domaine E.

Nous avons déjà signalé que la région centrale de ce domaine était conservée dans les différentes classes de récepteurs. Il est donc intéressant de comparer (à partir de l'alignement des séquences de tous les récepteurs nucléaires; voir article II) la position des différents introns dans cette région que nous appelons domaine Ti-DM. Il faut cependant noter que certains gènes de drosophile n'ont pas d'intron du tout (*ultraspiracle* par exemple ; Oro et al., 1990) ou seulement un, dans le domaine C (*knirps* ; Oro et al., 1988).

D'après la figure 41, il est clair que certains introns sont dans une position conservée, alors que d'autres sont beaucoup plus variables. Ainsi, l'intron central situé au milieu du domaine



**Figure 41 : Position des introns dans la zone conservée du domaine E (appelée domaine Ti-DM) pour les récepteurs nucléaires dont on connaît l'organisation génomique. Les positions respectives des introns ont été déterminées après alignement des séquences en acides aminés du domaine Ti-DM.**  
 THRA : récepteur des hormones thyroïdiennes de type alpha ; THRB : récepteur des hormones thyroïdiennes de type beta ; RARG : récepteur de l'acide rétinolique de type gamma ; EAR1 : récepteur orphelin ear-1 ; NGF1B : récepteur orphelin NGF1B ; ER : récepteur des oestrogènes ; PR : récepteur de la progestérone ; ECR : récepteur de l'ecdysone.

de dimérisation DM semble tout à fait conservé (à un acide aminé près, ce qui est probablement dû à des imperfections de l'alignement utilisé, ou à des insertions ou délétions d'acides aminés dans les exons) dans les gènes des récepteurs nucléaires que nous avons étudiés. De même, l'intron 3' est assez bien conservé, par exemple entre le récepteur de l'acide rétinoïque  $\gamma$  et le récepteur des oestrogènes. Ceci n'est pas toujours vrai et cet intron n'existe ni chez les récepteurs des hormones thyroïdiennes ni dans le gène ear-1 (Ponglikitmongkol et al., 1988; Sakurai et al., 1990 ; Lazar et al., 1990a ; Lehman et al., 1991). Par contre, la position des introns est très variable du côté N-terminal du domaine Ti-DM, et tout spécialement dans le domaine Ti lui même (pourtant la zone la mieux conservée du domaine E). Elle diffère par exemple entre les récepteurs des hormones thyroïdiennes et ceux de l'acide rétinoïque. Par contre, elle est identique entre ces derniers et le récepteur orphelin ear-1, comme entre le récepteur des oestrogènes et celui de la progestérone. Les introns situés dans la zone du gène codant le domaine Ti sont encore à des positions différentes pour NGF1B et pour le récepteur de l'ecdysone. Dans tout ces cas, une imperfection éventuelle de l'alignement ne peut pas expliquer les différences observées car le degré d'homologie est élevé. Du point de vue de la conservation des positions des introns, il y a donc un clair déséquilibre entre la partie Ti et la partie DM. Si, pour expliquer la situation actuelle, on voulait appliquer un modèle de stricte duplication de gène et de perte des introns au cours de l'évolution, il faudrait faire intervenir au moins 8 introns dans le domaine Ti-DM. Nous reviendrons plus tard sur les conséquences de ces observations sur l'évolution des récepteurs nucléaires (voir la discussion).

## ARTICLE 2 :

# EVOLUTION DE LA SUPERFAMILLE DES RECEPTEURS NUCLEAIRES.

L'étude de l'organisation génomique d'un membre d'une famille de gène comme celle des récepteurs nucléaires amène à se poser des questions sur l'évolution de cette famille de gènes. Quelle est l'origine des premiers récepteurs nucléaires ? Comment se sont-ils diversifiés ? Existents-ils dans d'autres organismes que les insectes et les vertébrés ? Peut-on retracer leur histoire évolutive ? Telles sont les questions auxquelles nous avons tenté d'apporter une réponse.

Dans la littérature, les récepteurs nucléaires sont communément répartis en deux sous-familles : d'abord celle des récepteurs des hormones thyroïdiennes, de l'acide rétinoïque, de la vitamine D et des récepteurs orphelins de type ear-1 (ear-1, PPAR et le gène de *Drosophile* E75 ; Miyajima et al., 1989 ; Lazar et al., 1989a ; Segraves et Hogness, 1990 ; Isseman et Green, 1990) et ensuite celle des récepteurs des hormones stéroïdes et des récepteurs orphelins ERR1 et ERR2 proches du récepteur des oestrogènes (Forman et Samuels, 1990a; 1990b; Parker et Bakker, 1991; Giguere et al., 1988). Cette répartition est fondée sur des arguments structuraux (place de l'intron entre les exons codant les doigts de zinc, etc...) et fonctionnels (existence d'hétérodimères, etc...). Mais, la plupart des récepteurs orphelins ne rentrent pas clairement dans cette classification et, dans la plupart des cas, on ignore à quelle sous-famille de récepteurs ils se rattachent. La découverte du gène NGF1B (également appelé N10 ou nur77 ; Milbrandt, 1988 ; Hazel et al., 1988; Ryseck et al., 1989) dont l'organisation génomique est unique (surtout dans la région des exons codant les doigts de zinc) a incité certains auteurs à suggérer qu'il définissait une troisième sous-famille de récepteurs. Clairement, il semblait intéressant d'analyser toutes ces informations pour essayer de bâtir une classification cohérente des membres de la superfamille.

Nous avons décrit la constitution des récepteurs nucléaires en quatre grands domaines fonctionnels indépendants, agissant de concert pour moduler harmonieusement la transcription des gènes cibles. De ces quatre domaines (A/B, C, D et E), deux (A/B et D), ne sont pas du tout conservés au sein de la famille, et même, d'un récepteur à l'autre. Par contre, le domaine C et, dans une moindre mesure la partie centrale du domaine E, se retrouvent dans la très grande majorité des récepteurs. Le domaine E est le plus grand de tous les domaines fonctionnels des récepteurs. Il semble que, au moins chez les récepteurs des hormones thyroïdiennes et de l'acide rétinoïque, les deux extrémités de ce domaine fixent le ligand, tandis que la partie centrale est responsable de la régulation de la transcription dépendante de l'hormone et de la dimérisation (modèle du "regulatory zipper" de Forman et Samuels, 1990a ; 1990b). L'observation des homologies existant entre les

récepteurs va dans le sens de ce modèle, puisque les parties L1 et L2 responsables de la fixation du ligand sont très conservées pour les récepteurs qui ont la même spécificité, mais n'ont pas de similitude apparente dans des types de récepteurs différents. Par contre, la zone centrale qui comprend le domaine Ti et la région de dimérisation, présente des homologies significatives chez la quasi totalité des récepteurs. En effet, les pourcentages d'homologie entre les récepteurs dans la région Ti vont de 20 à 45% environ (Wang et al., 1989 ; O'Donnel et Koenig, 1990). Nous appellerons cette zone centrale conservée du domaine E, le domaine Ti-DM.

L'existence de deux domaines conservés différents et indépendants chez tous les récepteurs nucléaires (ou presque) permettrait d'aborder la question de l'évolution des différents domaines fonctionnels. Y a-t-il eu évolution par simple duplication de gènes à partir d'un progéniteur commun des récepteurs nucléaires ou, au contraire, les domaines ont-ils évolué de manière indépendante pour donner lieu à des récepteurs chimères ? La comparaison de l'histoire évolutive des deux domaines devrait permettre de répondre à cette question. En effet, si les différents domaines ont évolué ensemble, les arbres évolutifs construits à partir des séquences de chaque domaine doivent avoir une topologie superposable (ce qui ne préjuge pas d'une vitesse d'évolution identique des domaines ; cette vitesse étant par exemple liée à des contraintes structurales différentes). Par contre, si les deux domaines ont évolué différemment, et si on a des molécules chimères, on doit observer des différences notables entre les deux types d'arbres évolutifs.

Pour ces raisons nous avons construit des arbres phylogéniques qui retracent, tout au moins dans ses grandes lignes, l'histoire évolutive de la superfamille des récepteurs nucléaires.

## Evolution of the nuclear receptor gene superfamily

Vincent Laudet<sup>2</sup>, Catherine Hänni, Jean Coll, François Catzeflis<sup>1</sup> and Dominique Stehelin

INSERM U186/CNRS URA 1160, Institut Pasteur, 1 Rue Calmette, 59019 Lille Cedex and <sup>1</sup>Laboratoire de Paléontologie, Institut des Sciences de l'Evolution (UA 327 CNRS), USTL, Place Eugène Bataillon, 34095 Montpellier Cedex, France

<sup>2</sup>Corresponding author

Communicated by D.Stehelin

Nuclear receptor genes represent a large family of genes encoding receptors for various hydrophobic ligands such as steroids, vitamin D, retinoic acid and thyroid hormones. This family also contains genes encoding putative receptors for unknown ligands. Nuclear receptor gene products are composed of several domains important for transcriptional activation, DNA binding (C domain), hormone binding and dimerization (E domain). It is not known whether these genes have evolved through gene duplication from a common ancestor or if their different domains came from different independent sources. To test these possibilities we have constructed and compared the phylogenetic trees derived from two different domains of 30 nuclear receptor genes. The tree built from the DNA binding C domain clearly shows a common progeny of all nuclear receptors, which can be grouped into three subfamilies: (i) thyroid hormone and retinoic acid receptors, (ii) orphan receptors and (iii) steroid hormone receptors. The tree constructed from the central part of the E domain which is implicated in transcriptional regulation and dimerization shows the same distribution in three subfamilies but two groups of receptors are in a different position from that in the C domain tree: (i) the *Drosophila knirps* family genes have acquired very different E domains during evolution, and (ii) the vitamin D and ecdysone receptors, as well as the FTZ-F1 and the NGF1B genes, seem to have DNA binding and hormone binding domains belonging to different classes. These data suggest a complex evolutionary history for nuclear receptor genes in which gene duplication events and swapping between domains of different origins took place.

**Key words:** gene evolution/nuclear receptors/phylogeny/transcription factors/zinc fingers

### Introduction

Nuclear receptors are ligand-activated transcription factors which regulate the expression of target genes by binding to specific *cis*-acting sequences (Evans, 1988; Green and Chambon, 1988; Beato, 1989). This family includes nuclear receptors for several hydrophobic ligands such as steroids, vitamin D, thyroid hormones, retinoic acid and also for dioxin and ecdysone. In addition to these receptors for known

ligands, numerous authors have described 'orphan' receptors which are putative receptors for ligands still to be identified (Evans, 1988; Moore, 1990). The family of nuclear receptor genes which to date contains 32 genes has been classically divided into two subfamilies on the basis of structural homologies (Green and Chambon, 1988). One subfamily includes the genes encoding steroid hormone receptors (receptors for glucocorticoids, androgens, oestrogens, mineralocorticoids and progesterone) and the other comprises receptors for thyroid hormone, retinoic acid and vitamin D (Forman and Samuels, 1990) as well as the *ear1* group (the *ear1/Rev-erb* gene and the *Drosophila E75* gene) (Miyajima *et al.*, 1989; Lazar *et al.*, 1989; Segraves and Hogness, 1990). The various orphan receptors so far described have not been classified among these subfamilies except for the recently discovered PPAR gene which is a close relative of *ear1* (Issemann and Green, 1990). In addition to these two subfamilies, Ryseck *et al.* (1989) have described a growth factor inducible orphan receptor gene named *nur77* (but also known as the NGF1B gene) which could define a third subfamily.

To perform their extremely diverse functions in homeostasis, reproduction, development and differentiation, nuclear hormone receptors link extracellular signals directly to transcriptional responses. All nuclear receptors are composed of at least four domains (Figure 1A) which are differentially conserved between the various products. The A/B domain implicated in transactivation and the hinge (D) domain are poorly conserved between the various subfamilies. The DNA binding (C) domain is composed of two zinc finger structures named CI and CII which play complementary roles in the DNA binding process (Green *et al.*, 1988). The C domain is the most conserved between nuclear receptors. The carboxy terminal E domain is the largest one and has a complex structure (Figure 1A). It is required for hormone binding, dimerization and regulation of transcription. Several authors have suggested that, at least for the thyroid hormone/retinoic acid receptor subfamily, the two extremities of this domain bind the ligand (Forman and Samuels, 1990; Glass *et al.*, 1989). In this model, the central part of the E domain is devoted to dimerization and to the so-called Ti domain implicated in the ligand-dependent regulation of transcription. This 'regulatory zipper' model (Forman and Samuels, 1990) seems suitable for an examination of the extent of the homologies between the various receptor sub-classes in this domain. In fact, the ligand binding part of E domain is nearly identical among receptors of the same binding specificity but does not have any apparent similarity among different receptor types. In contrast, the Ti domain exhibits 20-45% conservation among all receptors (Wang *et al.*, 1989; O'Donnell and Koenig, 1990).

The modular organization of nuclear receptors, the various degrees of conservation between their respective domains and the fact that the domains are encoded by different exons

(see Ponglikitmongkol *et al.*, 1988; Laudet *et al.*, 1991 and references therein) were prerequisites for investigating whether these genes had evolved by successive duplications from a common progenitor or if they had a chimeric structure originating from independent duplications of DNA binding and ligand binding domains from different origins.

In this study, we have taken advantage of the relative conservation of the DNA binding C domain and the Ti and dimerization domains (Ti-DM domain) among nuclear receptors to examine their evolutionary history. The phylogenetic trees derived from our study lead to the conclusion that the three subfamilies of nuclear receptor genes that we define have evolved through a simple duplication model except for some examples such as the *knirps* group or the receptors for ecdysone or vitamin D which may represent two types of 'evolutionary chimeras'.

Our work also shows that the three subfamilies of nuclear receptor genes emerged at an early stage during evolution.

## Results

### Alignment of nuclear receptors sequences in C and Ti-DM domains

The nuclear receptor genes included in this study are listed in Table I. To our knowledge this list is complete and all nuclear receptor genes thus far identified to date are studied in this paper.

The C domain sequences were manually aligned (Figure 1B). The C domain is composed of a variable number of amino acids (aa): from 67 for PPAR to 70 for THR $\alpha$ s, *v-erbA* and *tailless*. The majority of nuclear receptor genes possess 68 aa in this domain and it is likely that the

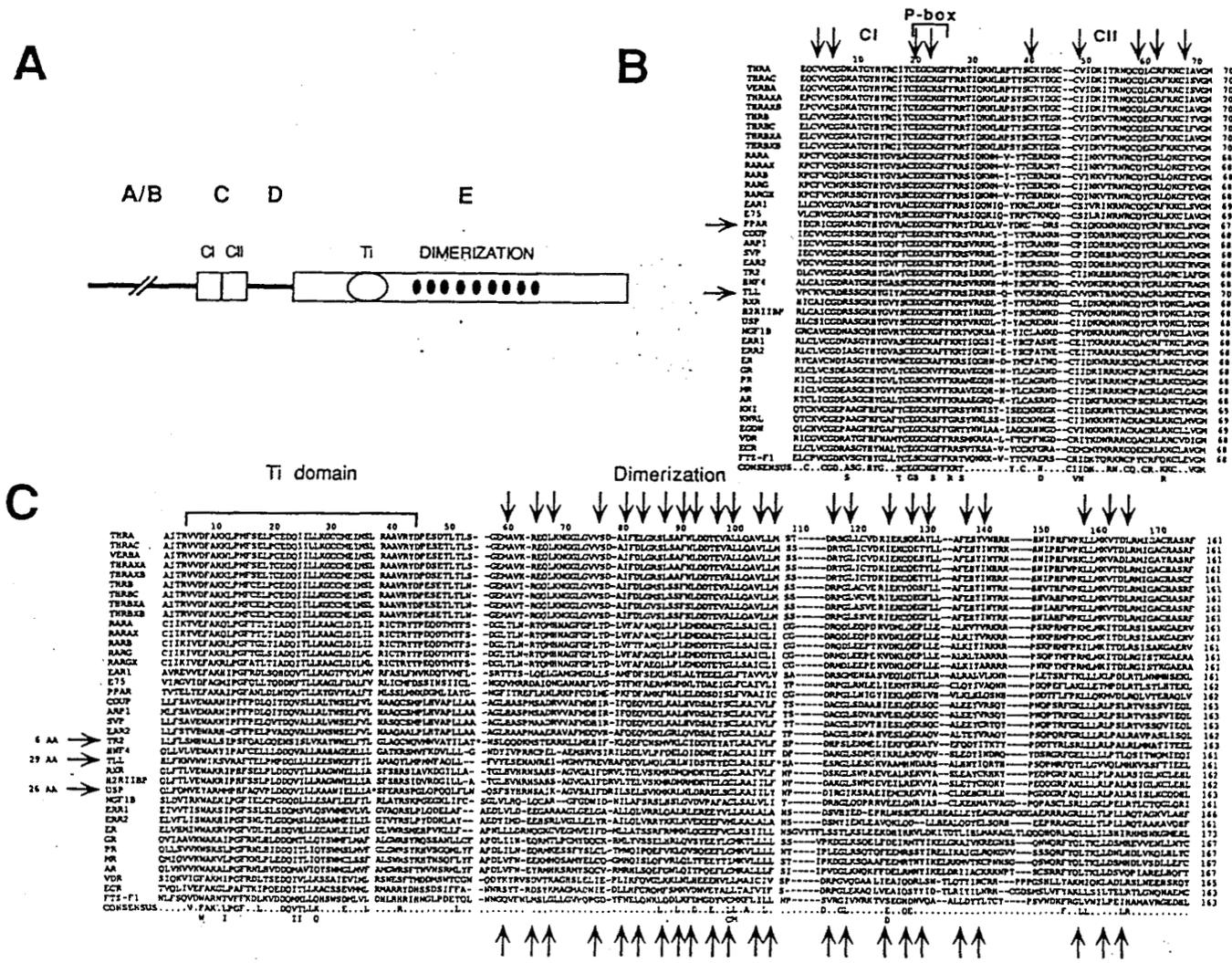
Table I. Sequences used in this study

Abbreviation	Name	Species	References
AR	Androgen receptor	Human	Tilley <i>et al.</i> (1989)
ARP-1	(Orphan receptor)	Human	Ladias and Karathanasis (1991)
COUP	= <i>ear3</i> (orphan receptor)	Human	Wang <i>et al.</i> (1989); Miyajima <i>et al.</i> (1988)
E75	(Orphan receptor)	<i>Drosophila</i>	Segraves and Hogness, 1990
EAR1	= <i>Rev-ErbA</i> (orphan receptor)	Human	Miyajima <i>et al.</i> (1989); Lazar <i>et al.</i> (1989)
EAR2	(Orphan receptor)	Human	Miyajima <i>et al.</i> (1988)
ECR	Ecdysone receptor	<i>Drosophila</i>	Koelle <i>et al.</i> (1991)
EGON	(Orphan receptor)	<i>Drosophila</i>	Rothe <i>et al.</i> (1989)
ER	Oestrogen receptor	Human	Green <i>et al.</i> (1986)
ERR1	(Orphan receptor)	Human	Giguère <i>et al.</i> (1988)
ERR2	(Orphan receptor)	Human	Giguère <i>et al.</i> (1988)
FTZ-F1	(Orphan receptor)	<i>Drosophila</i>	Lavorgna <i>et al.</i> (1991)
GR	Glucocorticoid receptor	Human	Hollenberg <i>et al.</i> (1985)
H2RIIBP	(Orphan receptor)	Mouse	Hamada <i>et al.</i> (1989)
HNF4	(Orphan receptor)	Rat	Sladek <i>et al.</i> (1990)
KNI	<i>knirps</i> (orphan receptor)	<i>Drosophila</i>	Nauber <i>et al.</i> (1988)
KNRL	<i>knirps-related</i> (orphan receptor)	<i>Drosophila</i>	Oro <i>et al.</i> , 1988
MR	Mineralocorticoid receptor	Human	Arriza <i>et al.</i> (1987)
NGF1B	= <i>nur77</i> = <i>N10</i> (orphan receptor)	Rat	Milbrandt (1988); Ryseck <i>et al.</i> (1989); Hazel <i>et al.</i> (1988)
PPAR	(Orphan receptor)	Mouse	Issemann and Green (1990)
PR	Progesterone receptor	Human	Misrahi <i>et al.</i> (1987)
RARA	Retinoic acid receptor $\alpha$	Human	Petkovich <i>et al.</i> (1987); Giguère <i>et al.</i> (1987)
RARAX	Retinoic acid receptor $\alpha$	<i>Xenopus</i>	Ragsdale <i>et al.</i> (1989)
RARB	Retinoic acid receptor $\beta$	Human	De Thé <i>et al.</i> (1987)
RARG	Retinoic acid receptor $\gamma$	Mouse	Zelent <i>et al.</i> , 1989
RARGX	Retinoic acid receptor $\gamma$	<i>Xenopus</i>	Ragsdale <i>et al.</i> (1989)
RXR	(Orphan receptor)	Human	Mangelsdorf <i>et al.</i> (1990)
SVP	<i>sevenup</i> (orphan receptor)	<i>Drosophila</i>	Mlodzik <i>et al.</i> (1990)
TLL	<i>tailless</i> (orphan receptor)	<i>Drosophila</i>	Pignoni <i>et al.</i> (1990)
TR2	(Orphan receptor)	Human	Chang and Kokontis (1988); Chang <i>et al.</i> (1989)
THRA	Thyroid hormone receptor $\alpha$ (= <i>c-erbA-1</i> )	Human	Thompson <i>et al.</i> (1987); Laudet <i>et al.</i> (1991)
THRAC	Thyroid hormone receptor $\alpha$ (= <i>c-erbA-1</i> )	Chicken	Sap <i>et al.</i> (1986)
THRAXA	Thyroid hormone receptor $\alpha$ (= <i>c-erbA-1</i> )	<i>Xenopus</i>	Yaoita <i>et al.</i> (1990)
THRAXB	Thyroid hormone receptor $\alpha$ (= <i>c-erbA-1</i> )	<i>Xenopus</i>	Yaoita <i>et al.</i> (1990)
THRB	Thyroid hormone receptor $\beta$ (= <i>c-erbA-2</i> )	Human	Weinberger <i>et al.</i> (1986)
THRBC	Thyroid hormone receptor $\beta$ (= <i>c-erbA-2</i> )	Chicken	Forest <i>et al.</i> (1990)
THRBXA	Thyroid hormone receptor $\beta$ (= <i>c-erbA-2</i> )	<i>Xenopus</i>	Yaoita <i>et al.</i> (1990)
THRBXB	Thyroid hormone receptor $\beta$ (= <i>c-erbA-2</i> )	<i>Xenopus</i>	Yaoita <i>et al.</i> (1990)
USP	Ultraspiracle = 2C1-3 = CF1 (orphan receptor)	<i>Drosophila</i>	Oro <i>et al.</i> (1990); Henrich <i>et al.</i> (1990); Shea <i>et al.</i> (1990)
VERBA	<i>v-erbA</i> gene of avian erythroblastosis virus	Chicken	Debuire <i>et al.</i> (1984); Damm <i>et al.</i> (1987)
VDR	Vitamin D receptor	Human	Baker <i>et al.</i> (1988)

Sequences used in this study. The names (and published synonyms when available) of the gene sequences are indicated, as well as their abbreviation used in the text and in the figures.

ancestral nuclear receptor C domain had the same number of amino acids. As most phylogenetic tree reconstruction procedures necessitate that all compared sequences be of the same size, it was necessary to introduce some gaps in the sequences. On recent structural observation it appears that the gaps introduced in numerous receptor sequences between C1 and C2 are localized in a loop structure (Schwabe and Rhodes, 1991). The cysteine residues of the zinc fingers (vertical arrows in Figure 1B) are absolutely conserved at an invariable position in the whole family, except for the first cysteine of the second finger CII in PPAR (shift of 2 aa downstream) and in the *Drosophila* gene *tailless* (shift of 2 aa upstream). Several other residues are conserved in the C domain, as indicated in the 'consensus' line in Figure 1B. Recently, Forman and Samuels (1990) have classified nuclear receptors into four classes (I-IV in Table II) on the basis of their DNA binding specificity. It is well known that DNA binding is mediated, at least in part, by

the three variable residues of a region called the P box which is situated at the end of the first finger. The alignment of Figure 1B gives an opportunity to complete this classification with three additional classes (NC, Table II). Interestingly, the orphan receptors all exhibit closely related P boxes (EGCKS or EGCKG) with the exception of the HNF4/*tailless* tandem and of the FTZ-F1 orphan receptor from *Drosophila*. The THR and RAR also have such a P box (EGCKG). Indeed, orphan receptors, THR, RAR, ECR and VDR all bind distantly related sequences, such as AGTCATGACCT for THR, RAR and VDR class II genes of Forman and Samuels (1990) or GTGTCAAAGTCA for orphan receptors COUP/*knirps/ear2* class III genes (Green and Chambon, 1988). This is confirmed by several recent reports: for example, the ARP-1 orphan receptor is able to bind to a thyroid hormone responsive element (Ladiaz and Karathanasis, 1991). All these sequences are formed by partially conserved inverted



repeats of 5 bp half sites separated by a variable number of residues. Direct repeats and half sites alone are also possible recognition sites.

The alignment procedure for the Ti-DM domain was much more difficult due to the low similarity scores between the genes (see Materials and methods). We have chosen to favour the conserved hydrophobic residues (vertical arrows in Figure 1C) implicated in the heptad repeats described by Forman and Samuels (1990) and Glass *et al.* (1989). Moreover, we have attributed a high penalty score for gap insertions in the CLUSTAL alignment algorithm (Higgins and Sharps, 1988). This procedure is valid when the level of amino acid identity is low, especially when the sequences compared are from different sub-classes of nuclear receptor genes. For alignment of reverse transcriptase sequences, Xiong and Eickbush (1990) have also used a high gap penalty in a procedure they called 'conserved residues' alignment. As with the Ti-DM domain in nuclear receptor genes, reverse transcriptase sequences have a low similarity score.

The Ti-DM domain alignment presented in Figure 1C shows some well conserved positions, although only two amino acids are strictly conserved between all nuclear receptor genes: an F and a D in positions 15 and 22 of the Ti sub-domain. But even in the DM domain, ~20 residues are conserved in >80% of the genes (see consensus line in Figure 1C). Very often such conserved residues are hydrophobic and likely to be implicated in the dimerization process (Glass *et al.*, 1989).

Several papers have emphasized that in addition to the C domain, several regions of the ligand binding domain of the nuclear receptor genes are conserved. A detailed analysis of such conserved regions appears in Segraves and Hogness (1990) for the E75 gene, in Pignoni *et al.* (1990) for the *tailless* gene and recently in Koelle *et al.* (1991) for the ecdysone receptor gene. In all these cases, the conserved regions lie in the Ti-DM domain as defined here.

Numerous reports implicate this domain in dimerization for the THR/RAR/VDR proteins as well as the *earl* gene product, such a detailed analysis was recently also done for the steroid hormone receptors (Fawell *et al.*, 1990). Our alignment shows that for these genes, only a fraction of the conserved residues of THR/RAR/VDR are conserved in the steroid hormone receptors. Nevertheless, it is possible to recognize the Ti-DM domain in the steroid receptor gene subfamily. A significant degree of homology exists in that region in all nuclear receptor genes (Pignoni *et al.*, 1990; Segraves and Hogness, 1990).

We were not able to align the *knirps* group of genes in the Ti-DM domains. Neither the Ti nor the DM leucine-rich sub-domains are visible in these molecules and their E domains have no homologies with E domains of other genes.

#### Generation of a phylogenetic tree for the C domain

Based on the alignment of C domains presented in Figure 1B, the Fitch least square method (Fitch, 1981; Swofford and Olsen, 1990 for review) was used to generate a phylogenetic tree of the 32 nuclear receptor genes, which is shown in Figure 2. This tree shows that the nuclear receptor gene family can be subdivided into three subfamilies: (i) the THR/RAR/*earl* subfamily; (ii) the orphan receptor subfamily, which contains the COUP, RXR and HNF4/TLL type orphan receptors; and (iii) the steroid

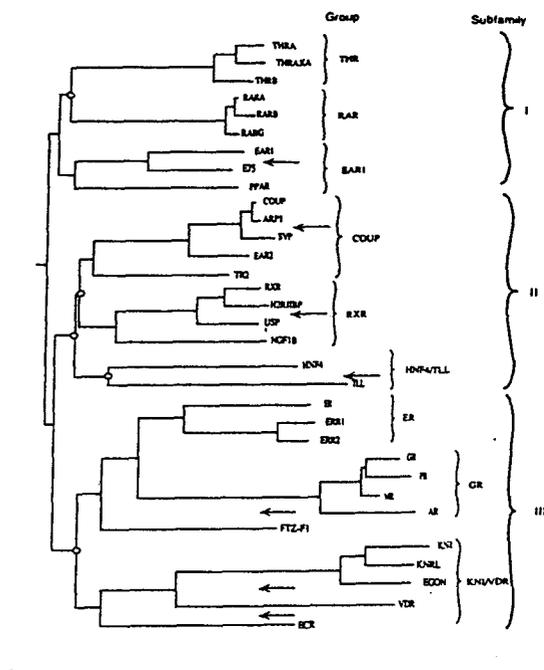


Fig. 2. Rooted phylogenetic Fitch tree for nuclear receptor genes based on C domain sequences. The bar represents a branch length of 10 units. The deepest branch between the internal nodes connecting all the nuclear receptor genes was divided at mid-length in order to root the tree. Arrows point to the mammalian and *Drosophila* genes which cluster together. Open circles point out the differences between Fitch and NJ trees. Groups and subfamilies are indicated by brackets.

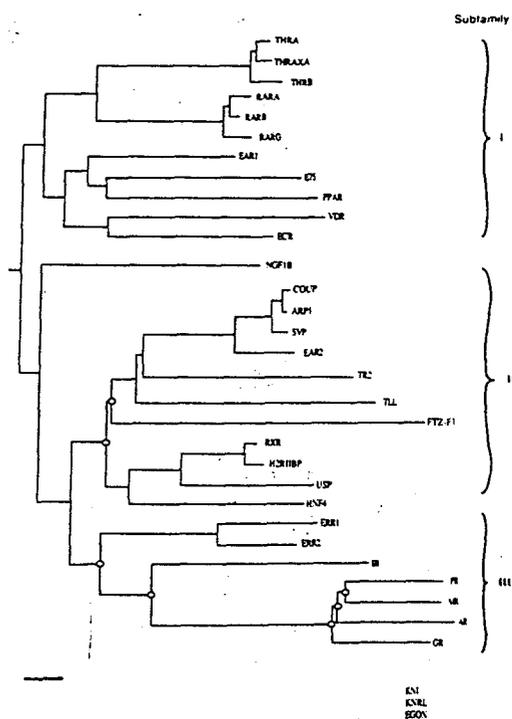


Fig. 3. Rooted phylogenetic Fitch tree for nuclear receptor genes based on Ti-DM domain sequences. The bar represents a branch length of 10 units. The deepest branch between the internal nodes connecting all the nuclear receptor genes was divided at mid-length in order to root the tree. Open circles point out the differences between Fitch and NJ trees. The three main subfamilies are bracketed. The KNI, KNRL and EGON genes which do not have recognizable Ti-DM domain are indicated at the bottom of the figure.

hormone receptor subfamily. All nuclear receptors published to date can be classified into one of these three subfamilies; this observation is in favour of the tree of Figure 2.

This Fitch tree was confirmed by a neighbour-joining (NJ) analysis of the 32 DNA binding domains (Saitou and Nei, 1987). The NJ tree we obtained has the same topology as the Fitch tree except for the position of the VDR/*knirps* group, which clusters with the THR group in the NJ tree (Figure 2 and data not shown). In fact, ECR has been described as being close to THRB and to VDR (Koelle *et al.*, 1991). The position of the ECR gene near the THR group in the NJ tree, with very long branches, is poorly confirmed by bootstrap analysis: this association was found in only 18% of bootstrap trials (data not shown). In addition, in our previous tree analysis, done before the publication of ECR, the VDR/*knirps* group was always stably associated with the subfamily III genes as in our present Fitch tree. Thus, the rooting of this group with the THR group may be due to a convergence of ECR toward the THRs (see Discussion). Another difference between the Fitch and NJ trees is the precise position of HNF4 and *tailless*. These two genes form a clade in subfamily II in the Fitch tree (Figure 2) but root together with the COUP group in the NJ tree (data not shown). More information will be required in order to locate these two genes more precisely.

We have arbitrarily rooted this Fitch tree at mid-length of the deepest branch between subfamily I and subfamilies II and III. This order of first gene duplication is tentative, as an initial dichotomy between subfamilies I and II versus III cannot be ruled out (until an adequate outgroup can be included). Until more sequence data on these and additional genes are made available, the precise order of the first and second gene duplications which gave rise to the three ancestors of the nuclear receptor gene subfamilies will remain undetermined.

The finding that the majority of the orphan receptors (exceptions are *ear1*, E75, PPAR, ERR1, ERR2, FTZ-F1 and the *knirps* group) cluster together is surprising because there is no previous report of any particular similarity between these genes. This observation may have important repercussions for the study of the physiological roles of these molecules (see Discussion). Moreover, the Fitch tree in Figure 2 shows that grouping together the nuclear receptors based on similarity within their DNA binding domain sequences correlates with their ligand binding specificity, where this is known: receptors of the THR/RAR subfamily (class I) bind thyroid hormone and retinoic acid respectively and members of the steroid receptor subfamily (class III) bind diverse derivatives of the steroid hormones. This supports the hypothesis of a common evolution of DNA binding and ligand binding domains.

Each of the three subfamilies recognized in Figure 2 may, in turn, be divided into groups: THR, RAR and *ear1* groups for class I receptor genes, COUP, RXR and HNF4/*tailless* groups for orphan class II receptor genes, and ER, GR and VDR/*knirps* groups for class III genes.

In terms of nuclear receptor gene evolution, it is interesting to note that, in each of the three subfamilies described here, mammalian and *Drosophila* genes are grouped together: for example, this is the case for E75 and *ear1* class I genes. Moreover, each of the three groups of genes among the subfamily II orphan receptors has a cluster of human and *Drosophila* genes: *sevenup* for COUP, *ultraspiracle* for RXR

and *tailless* for HNF4. This suggests that the three subfamilies of the nuclear receptor genes already existed before the divergence of the arthropod and vertebrate lineages.

A careful examination of Figure 2 leads to several additional observations: the most surprising is that the ecdysone and vitamin D receptors and the *knirps* group (*knirps*, *knirps-related* and *egon* (Oro *et al.*, 1988; Nauber *et al.*, 1988) cluster together. This observation was unexpected since there is no report of a close similarity between these genes. The fact that they are grouped together, after a rather long common ancestral branch length (see Figure 2), may be the result of a real ancient, common origin or, less probably, of a convergence phenomenon. In the NJ tree, the VDR and ECR genes also cluster with the *knirps* family (data not shown), confirming our observation based on the Fitch tree analysis, although the position of this whole group in the two trees is different, as mentioned previously.

#### Generation of a phylogenetic tree for the Ti-DM domain

To compare the evolution of DNA binding and Ti-DM domains we have constructed, using the same methods (i.e. Fitch least squares analysis confirmed by an NJ tree) a tree based on the sequence alignment of the Ti-DM domains. Figure 3 shows the topology of this tree which exhibits roughly the same distribution into three subfamilies as found in the C domain analysis, i.e. THRs/RARs (subfamily I), orphan receptors (II) and steroid hormone receptors (III). Thus, the three subfamilies we have defined with the C domain tree are also valid in the Ti-DM domain derived tree. It is interesting to note that the same topology is obtained when we construct a Fitch or an NJ tree based on the C and Ti-DM sequences treated together (data not shown).

Nevertheless, it is obvious from Figure 3 that some important differences exist between the two molecular phylogenies. The first point is that three genes cannot be included in the phylogenetic analysis of the other nuclear receptors: these are the three genes of the *Drosophila knirps* group (*knirps*, *knirps-related* and *egon*: reviewed in Pankratz and Jäckle, 1990). These genes have acquired completely new E domains that are unrelated to the Ti-DM domain of the other nuclear receptors. This observation fits well with the data of several authors who have described these genes (Oro *et al.*, 1988; Nauber *et al.*, 1988; Rothe *et al.*, 1989).

The other major difference concerns the VDR, ECR, FTZ-F1, NGF1B, *tailless* and HNF4 genes. In the C domain Fitch tree (Figure 2), VDR and ECR are associated with the *knirps* family in subfamily III, but in the Ti-DM domain tree (Figure 3) they appear to belong to subfamily I. Numerous data suggest that the VDR gene product has a physiological behaviour resembling that of the THR/RAR gene products (reviewed in Forman and Samuels, 1990). The fact that the VDR and the ECR genes belong to different subfamilies according to C and Ti-DM domain trees was confirmed by an NJ tree based on Ti-DM sequences (data not shown). This point will be developed further in the Discussion.

The FTZ-F1 gene belongs to subfamily III for the C domain, whereas in the Fitch tree constructed from the Ti-DM domain it belongs to the COUP group within the subfamily II with a very long branch (Figure 3). This position is not confirmed by the NJ tree of Ti-DM domain

where this gene clusters with all the subfamily II genes. Thus, although clearly differing in C and Ti-DM domain trees, the precise location of FTZ-F1 needs further work. The NGF1B gene also shows a different partition between C and Ti-DM domain trees. In the C domain tree, NGF1B belongs to the subfamily II orphan receptor genes in close association with the group of the retinoid responsive RXR genes. Surprisingly, in the Ti-DM domain-derived tree, NGF1B cannot be assigned to a particular subfamily since it branches early before the divergence of subfamilies II and III (see Figure 3). The *tailless* and HNF4 genes show a slightly different situation: these two genes cluster differently according to the tree construction procedure (NJ or Fitch) for C domain. For Ti-DM domains they are always separated: the *tailless* gene groups with the COUP group and the HNF4 gene clusters with the RXR group.

Finally, the comparison of the two NJ trees shows a few additional, more subtle differences. In the C domain of subfamily III, the GR and PR genes form a rather young sister group and are then joined by MR and AR respectively, whereas for the Ti-DM domain, the PR and MR genes first cluster together and are then associated with AR and GR. A note of caution should be made since the internodal, ancestral, segments are very short (Figures 3), and the relationships between PR, MR and GR cannot be resolved unambiguously (trichotomy). The same type of discrepancies are seen within the *ear1* and ER groups of genes. Again, trichotomy events cannot be ruled out. These differences may possibly be attributed to different rates of evolution among and between the two domains involved in this comparison: for example, the C domain, which is smaller than the Ti-DM domain, is also much more conserved between the nuclear receptor genes.

#### Evolution of the THR/RAR gene complex

Figure 4 presents a Fitch tree obtained for the Ti-DM domains of all representatives of THR/RAR subgroups of genes. This is of interest since it is known that the genes encoding THR and RARs have similar chromosomal locations (namely chromosome 17q21 for THRA and RARA and chromosome 3p24 for THRB and RARB), which suggests a common evolution for these genes. Furthermore, several authors have emphasized that the human and rat THRA genes partially overlap with the orphan receptor *ear1* gene (Lazar *et al.*, 1989; Miyajima *et al.*, 1989; Laudet *et al.*, 1991). Finally, the evolution of this gene subfamily is worth studying because the *v-erbA* oncogene represents a virally transduced and modified version of THRA which has evolved independently from its cellular progenitor (Debuire *et al.*, 1984).

The tree in Figure 4 confirms the pattern observed in the preceding trees (Figure 3), i.e. in the class I subfamily the individual genes of the *ear1* subgroup, namely ECR, VDR, *ear1*, PPAR and E75, have diverged at a very early stage, approximately at the same time as the THR and RAR subgroups. For the THR group, several additional observations can be made: first, the *v-erbA* oncogene, after its presumed divergence from the chicken THRA, has evolved more quickly than the other genes of its group. This was expected since this gene, transduced by the avian erythroblastosis virus (AEV), has been subjected to a high rate of replication by reverse transcriptase, which is known to have a very low fidelity. Curiously, the chicken THRA

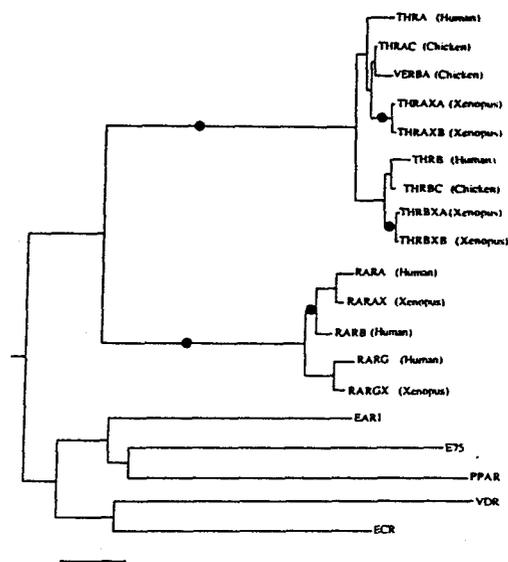


Fig. 4. Fitch tree for all sequences of subfamily I for Ti-DM domain, which shows the emergence of the various THR and RAR genes. The bar represents a branch length of 10 units. For THR and RAR groups, when duplications appear unambiguous, they are marked by dots along ancestral branches.

gene does not root with the human THRA gene, but clusters with the *Xenopus* genes. Nevertheless, internodal segments defining the different clusters in this subgroup are very short, and a polytomy such as [(THRA, THRAC, *v-erbA*), (THRAXA, THRAXB)] cannot be ruled out. Secondly, it is worth emphasizing the relationships of the four THR genes described in *Xenopus*. These genes have, according to Figure 4, been duplicated rather recently in the *Xenopus* lineage; this recency represents a specific difference between *Xenopus* and other vertebrates. Similar gene duplication phenomena in *Xenopus* have been observed for many other genes (tetraploidization, see Kobel and Du Pasquier, 1986 and Bisbee *et al.*, 1977).

## Discussion

### The three nuclear receptor subfamilies

Our study allows to classify the nuclear receptor genes into three subfamilies: subfamily I joins together the THR, RAR and *ear1* subgroups of genes, subfamily II groups most of the orphan receptor genes and subfamily III contains the steroid hormone receptor genes and the *knirps* group. This observation, based on an evolutionary tree of C domain sequences, can be substantiated by structural data available on nuclear receptor genes (Baker *et al.*, 1988; Hughes *et al.*, 1988; Laudet *et al.*, 1991; Lehman *et al.*, 1991; Miyajima *et al.*, 1989; Oro *et al.*, 1990; Ponglikitmongkol *et al.*, 1988; Rothe *et al.*, 1989; Ryseck *et al.*, 1989; Yaoita *et al.*, 1990; see also Green and Chambon, 1988 for review). It is well known that for all vertebrate nuclear receptor genes, except NGF1B, the A/B domain and each of the zinc fingers CI and CII are encoded by separate exons. Furthermore, the position of the intron lying between these two exons is varies among the sub-classes; for the THR/RAR/*ear1* subfamily, this intron (number 4 for THRA; Laudet *et al.*, 1991) is located one amino acid after the last cysteine of CI, whereas for the steroid hormone receptor genes, this intron (number 2 for ER; Ponglikitmongkol *et al.*, 1988)

is situated 10 aa after this last cysteine of CI. The NGF1B orphan receptor gene is the only known exception to this observation since its CI finger is encoded by an exon which also encodes the A/B domain. Moreover, the intron separating CI and CII lies between the two positions known for THR/RAR and steroid receptors i.e. 5 aa after the last cysteine of CI (Ryseck *et al.*, 1989). Consequently, NGF1B—based on structural features—seems to define on its own a third subfamily of nuclear receptor genes. How do our data fit with these structural observations? The C domain tree does show clearly that orphan receptors genes, including NGF1B, cluster together (except for the *knirps*, *ear1* and ERR groups). This interesting observation strongly supports the proposal that the nuclear receptor genes can be organized into three subfamilies as suggested by Ryseck *et al.* (1989) based on the study of only one gene (NGF1B). The members of subfamilies I and III (as defined by us, i.e. containing the subgroups THR, RAR, *ear1* and steroid receptors including VDR respectively) have the same exon/intron structures within their own subfamily. Unfortunately, the *Drosophila* ECR gene has no intron between its two zincfingers (Koelle *et al.*, 1991). For the vertebrates subfamily II genes, only the genomic structure of NGF1B is known and it is tempting to speculate that other orphan receptor genes may have the same structure in their C domain. To our knowledge, the only other gene of subfamily II whose genomic organization has been determined is the *ultraspiracle* gene, which has no introns. But it is well known that intron position and number are often more variable in *Drosophila* genes than in their vertebrate homologues. It is possible that the *ultraspiracle* gene has lost all its introns during the evolution of *Drosophila*. The only structural information which does not fit with our C domain tree also comes from *Drosophila* genes: the *knirps* group have the same exon–intron position between the two fingers as the THR/RAR/*ear1* genes, although they cluster with class III steroid hormone receptors. Nevertheless, the *knirps* genes have only one intron and, again, this observation must be tempered by the fact that the *knirps* genes are *Drosophila* nuclear receptors with a very unusual evolutionary history (see below). Of course, there are other possible explanations for the actual intron positions in the various genes of the family, and especially in *Drosophila*. For example, some organisms such as *Drosophila* could have gained introns during evolution and these introns could have been inserted at putative 'proto-splice' sites as in the case of the tubulin and actin genes studied by Dibb and Newman (1989). The description of other genes from 'primitive' organisms in the future should enable us to test such a hypothesis.

The fact that the orphan receptor genes (except for the ERR, *ear1* and *knirps* groups of genes) are all grouped together to form subfamily II strongly suggests that these genes have a common ancestor. The main problem to be solved regarding orphan receptors function is whether or not they bind a ligand and, if so, what is the identity of these ligands. Moore (1990) has hypothesized that terpenoids may be ligands of orphan receptors. *Drosophila* juvenile hormone may be such ligand and the plant hormones gibberellic and abscissic acids are other examples of ligands of receptors yet to be found. Another possibility is that orphan receptors do not have ligands at all, but act as hormone-independent transcriptional regulators. These molecules, which all have

Table II. Classification of nuclear receptors for their DNA binding specificity

Sequence of the P box	Class	Genes
GSCKV	I	AR, GR, MR, PR
EGCKG	II	E75, EAR1, ECR, RXR, H2RIIBP, USP, TR2, VDR, NGF1B, PPAR, RARA, RARB, RARG, THRA, THRB
EGCKS	III	ARP-1, COUP, EAR2, EGON, KNI, KNRL, SVP, VERBA
EGCKA	IV	ER, ERR1, ERR2
ESCKG	NC	FTZ-F1
DGCKG	NC	HNF4
DGCAG	NC	TLL

Classification of nuclear receptors according to their DNA binding specificity (sequence of the P box). This classification is based on that of Forman and Samuels (1990). Receptors which were not classified by these authors (namely HNF4, TLL and FTZ-F1) are indicated as NC (new class).

very similar DNA binding specificities (see Table II), may bind to the same DNA sequences as the canonical nuclear receptors. As an illustration of this hypothesis, the THR/RAR proteins have the same DNA binding sequence (P box) as the majority of subfamily II orphan receptors. The fact that the ARP-1 orphan receptor, a very close relative of COUP, can bind the thyroid hormone responsive elements (T3REs) suggests that the majority of the orphan receptors, which possess the same P box, could also bind T3REs. As all subfamily II orphan receptors seem to have diverged from a common ancestor (Figures 2 and 3), it is tempting to speculate that these orphan receptors have diverged from THR/RAR-like ancestral genes to become hormone independent regulators of THR/RAR action by competing, dimerizing or synergizing with them. Indeed, numerous proteins are known to regulate the action of THRs and RARs (O'Donnell and Koenig, 1990). For example, the recently identified thyroid hormone receptor auxiliary protein (TRAP) (Beebe *et al.*, 1991; O'Donnell *et al.*, 1991; Darling *et al.*, 1991), which is able to bind a T3RE and to dimerize with THRs via the Ti-DM domain and the CII zinc finger, may indeed be an orphan receptor. The answer to this question awaits the cloning of the gene encoding TRAP. We believe that the common origin of members of subfamily II orphan receptors argues for this model and permits us to hypothesize a complex regulation of the action of T3 and retinoids.

#### Comparison of C domain and Ti-DM domain trees

The comparison we have made between the C and Ti-DM domain trees (Figures 2 and 3) confirms the notion that the major mode of evolution of nuclear receptor genes has been by gene duplication. Nevertheless, we have identified some nuclear receptor genes which may have had a more complex history.

The first example is the *knirps* family (*Drosophila knirps*, *knirps-related* and *egon* genes). These genes lack the Ti-DM domain which is replaced by an unrelated domain. It is postulated that these genes arose from an incomplete duplication of an ancestor gene. This may result in nuclear receptor-like proteins which have no dimerization and

possibly no ligand binding motifs and which may act as classical transcription factors as suggested by others (Nauber *et al.*, 1988; Oro *et al.*, 1988; Rothe *et al.*, 1989). Alternatively, it is possible that these genes encode nuclear receptors for ligands that are chemically unrelated to steroids, retinoids and thyroid hormones. This hypothesis might explain the observation that their E domains are clearly different from other nuclear receptor genes.

The second example of an independent evolution between C and Ti-DM domains comes from genes such as VDR, ECR, FTZ-F1, NGF1B, *tailless* and HNF4, which belong to one subfamily when the C domain is considered and to another based on the Ti-DM domain analysis. The most dramatic examples of this are the VDR and ECR genes which were classified into subfamily III based on their C domains and into subfamily based on their Ti-DM domains. This latter fact is not surprising and was already mentioned by others (see Forman and Samuels, 1990 and references therein). Furthermore, bootstrap analyses on the monophyletic character of the various groups which contain VDR and ECR give reliable results except, as previously mentioned, for the rooting of the VDR/*knirps* group with THR genes in the C domain NJ tree (data not shown). The placement of VDR in subfamily III from its C domain is confirmed by gene structure analysis since this gene has an exon-intron boundary between CI and CII that is diagnostic of the steroid hormone receptor from subfamily III. It is not surprising that the DNA binding specificity of VDR and ECR is like that of THRs and RARs (i.e. subfamily I members) because the specificity of DNA binding is encoded only by a very small number of amino acids (P box, see Table II). Thus, one can hypothesize an alteration (or 'correction') of binding specificity gained from a small number of mutations within the C domain of type III which has consequently been switched towards a subfamily I type physiological behavior. This gives a large evolutionary flexibility to the nuclear receptor superfamily of genes. It has to be noted, however, that our C domain tree is based on only 70 aa residues which have kept a high degree of homology. Phylogenetic trees based on this short domain are very sensitive to local phenomena of evolutionary convergence. This may explain the NJ tree pattern observed for the VDR/*knirps* group which may be driven to the THR group by the homology of ECR with the THR genes. It is important to note that our conclusion on the position of these genes in the C domain trees should be considered, for the moment, as tentative. Nevertheless, the fact that structural and phylogenetical observations fit together suggests that the ECR and VDR genes indeed had a complex evolutionary history. One can hypothesize that an incomplete gene duplication event, or a homologous recombination-like event, may explain the appearance of such nuclear receptor chimeras with C and Ti-DM domains arising from different ancestors.

The lack of structural and functional characterization of FTZ-F1, NGF1B and *tailless*/HNF4 limits the discussion of our observations. It should be noted, however, that the NGF1B gene was previously described as an exception since its C domain exon-intron structure is unique (Ryseck *et al.*, 1989). It is possible that the NGF1B gene represents an exceptional chimera among nuclear receptor genes; alternatively, for an unknown reason, its C and Ti-DM domains may have been subjected to very different rates of substitution. For the *tailless* and HNF4 genes, we strongly

favour the hypothesis that they belong to different groups in the subfamily II of orphan receptors (COUP group for *tailless* and RXR group for HNF4). The fact that they cluster together in subfamily II in the C domain tree (Figure 2) might be the result of an evolutionary convergence between the C domains of these genes. Indeed for this domain, the NJ and Fitch trees do not group these genes in the same manner, suggesting that too much homoplasy prevents us from placing these genes correctly. Another possibility which cannot be excluded is that, again, these genes represent two types of chimeras. More information, such as physiological data, will be necessary in order to infer the precise phylogenetic relationship between these genes. It is important to make clear that the C domain tree is based on rather short and highly conserved sequences and so is much more unstable than Ti-DM domain tree. The precise position of some deeply rooted sequences such as HNF4, *tailless* and FTZ-F1 may be poorly resolved by this tree.

In conclusion, our analysis suggests that in addition to canonical gene duplication, other events (recombination, translocation or exon shuffling) may have arisen during the evolution of the ancestral nuclear receptor genes in order to yield the complex gene family presently known. Some genes such as *knirps*, ECR, VDR, FTZ-F1 and NGF1B may represent different sorts of evolutionary chimeras, increasing the sensitivity of the various hormone response pathways.

#### *A hypothesis for the evolution of the nuclear receptor gene superfamily*

It is obvious from our work that the different subfamilies of nuclear receptors appeared before the divergence of the arthropod and vertebrate lineages which occurred at least 500 million years ago (Hartland *et al.*, 1982). Indeed, numerous vertebrate nuclear receptor genes have a *Drosophila* homologue belonging to their subgroup clade, suggesting that all the preceding divergences took place before the separation of the arthropods and vertebrates. This is the case for the *ear1* group and for all the subfamily II orphan receptor groups (RXR, COUP and HNF4). From this observation alone, it is unclear whether THRs/RARs and steroid receptors co-existed, as they presently do, before the arthropod/vertebrate divergence. However, the recent cloning of the *Drosophila* ecdysone receptor showed that canonical nuclear receptors are functional in this fly (Koelle *et al.*, 1991). In *Drosophila*, other molecules which are potential ligands for nuclear receptors have been described (Segraves, 1991). Recently, several authors have hypothesized that the *Drosophila* E75 gene may be the juvenile hormone receptor (Ashburner, 1990; Segraves, 1991). It is clear that some of the *Drosophila* orphan receptors are probably receptors for molecules known to mediate numerous physiological responses in the fly. So we can conclude that the function of nuclear receptors (i.e. transduction of cell-cell messages via small hydrophobic molecules) already existed before the divergence of the arthropod and vertebrate lineages.

Figure 5 presents an evolutionary hypothesis for the possible evolution of nuclear receptor genes, in which putative switches and incomplete duplication events are indicated. It is obvious from this figure that the nuclear receptor superfamily diverged (mainly by duplication) early on, and has since been well conserved with the exception of some minor events. This raises the intriguing question

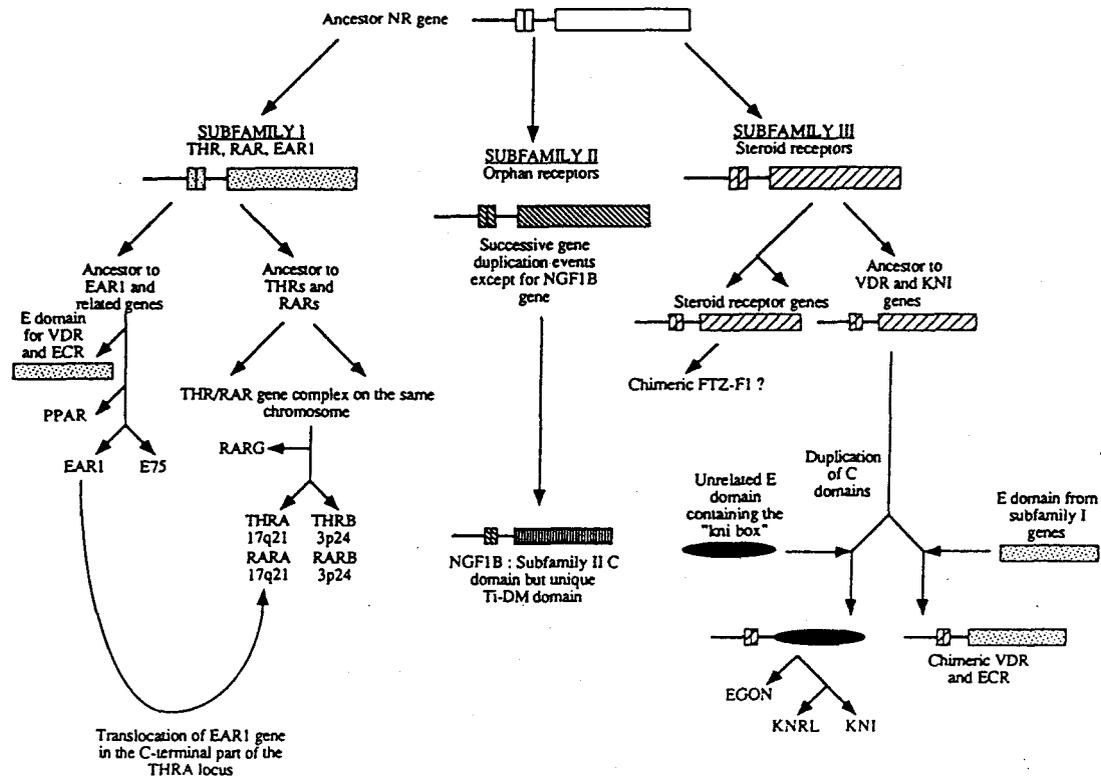


Fig. 5. An evolutionary scenario for the nuclear receptor genes, in which only important events and features are indicated. Each subfamily of nuclear receptor genes is represented by a different shading. The VDR and ECR genes may be formed with a subfamily III C domain and a subfamily I Ti-DM domain. The NGF1B gene has a subfamily II C domain and a Ti-DM domain which cannot be classified between subfamilies II and III. The FTZ-F1 gene is perhaps also an example of chimera. The *knirps* group of genes have an E domain unrelated to other nuclear genes but all the *knirps* genes have the same 'kni box' situated just after the second zinc finger CII. The order of duplication of THRs, RARs and EAR1 genes with VDR and ECR is taken from Ti-DM domain tree. The translocation event which has placed the EAR1 gene in the THRA locus is indicated: this event seems to have arisen just before the apparition of mammalian lineage (see text for details).

of the origin of this gene superfamily. O'Malley (1989) and Moore (1990) have speculated that nuclear receptor genes originated very early on by fusion of DNA binding sequences and genes implicated in steroid binding in the cytosol recently described in bacteria (see Moore, 1990 for references). Nuclear receptor genes are not yet known to exist in organisms other than vertebrates and insects. However, it is obvious that, at least for worms, such genes do exist, as the identification of nuclear receptor genes has been announced recently in the nematode *Caenorhabditis elegans* (cited in Moore, 1990) and as ecdysteroids are known to trigger physiological responses in the helminth parasite *Schistosoma* (Nirdé *et al.*, 1983). Potential nuclear receptors found in primitive organisms are the receptor for the DIF factor, a hydrophobic morphogen, of the slime mould *Dictyostelium* which was recently described as a nuclear protein (Insall and Kay, 1990) and the receptor for the steroid hormone antheridiol from the filamentous fungus *Achlya ambisexualis* which is able to bind heat shock proteins as vertebrates steroid receptors (Brunt *et al.*, 1990). Our data indicate that most of the evolution of this family was nearly complete 500 million years ago, i.e. at the arthropod/vertebrate split. We can imagine that there was a rapid evolution of these genes and then, when the different classes of receptors were fixed in their present function (i.e. transduction of a signal from other cells), only minor rearrangements, occurred leading to an overall strong gene conservation. Numerous reports have recently emphasized the speed of early metazoan evolution in the early Cambrian

(for review see Valentine, 1977). From our data it is tempting to speculate that, as nuclear receptor genes are implicated in cell-cell communication, a crucial need for metazoans, the rapid divergence of early nuclear receptor genes coincides with this 'burst' of metazoan evolution [see Erwin (1991) for a review].

The evolution of the THRs/RARs/*ear1* subfamily presented in Figure 5 needs further comments: THRA and RARA are situated on the same chromosome (17q21), and are THRB and RARB are on the chromosomal segment, 3p24. Furthermore, the *ear1* gene partially overlaps with the mammalian THRA gene (Miyajima *et al.*, 1989; Lazar *et al.*, 1989; Laudet *et al.*, 1991). It was then of interest to study the history of that group of genes carefully. From Figure 4, it is clear that the *ear1* group diverged early from the THR/RAR group, before the arthropod/vertebrate divergence, since the *ear1* gene has a *Drosophila* homologue, E75, clustering with it. Then, the undifferentiated THR/RAR ancestor diverged to give THR and RAR progenitors likely to be on the same chromosome (probably an ancestor of chromosome 3 or 17). Except for the RARG gene, whose chromosomal location is still unknown, it is clear that this THR/RAR gene complex was finally duplicated to give the four present genes, which were already present in early vertebrates since they exist in *Xenopus* (Yaoita *et al.*, 1990; Ragsdale *et al.*, 1989). The overlap which exists between THRA and *ear1* genes seems to be a recent event which may be the result of a translocation which has placed *ear1* in the same locus as THRA in a tail-

- Green, S., Kumar, V., Theulaz, I., Wahli, W. and Chambon, P. (1988) *EMBO J.*, **7**, 3037-3044.
- Hamada, K., Gleason, S.L., Levi, B.Z., Hirschfeld, S., Appella, E. and Ozato, K. (1989) *Proc. Natl. Acad. Sci. USA*, **86**, 8289-8293.
- Hartland, W.B., Cox, A.V., Llewellyn, P.G., Pickton, C.A.G., Smith, A.G. and Walters, R. (1982) *A Geologic Time Scale*. Cambridge University Press, Cambridge.
- Hazel, T.G., Nathan, D. and Lau, L.F. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 8444-8448.
- Henrich, V.C., Sliter, T.J., Lubahn, D.B., Macintyre, A. and Gilbert, L.I. (1990) *Nucleic Acids. Res.*, **18**, 4143-4148.
- Higgins, D.G. and Sharp, P.M. (1988) *Gene*, **73**, 237-244.
- Hollenberg, S.M., Weinberger, C., Ong, E.S., Cerelli, G., Oro, A., Lebo, R., Thompson, E.B., Rosenfeld, M.G. and Evans, R.M. (1985) *Nature* **318**, 635-641.
- Hughes, M.R., Malloy, P.J., Kieback, D.G., Kesterson, R.A., Pike, J.W., Feldman, D. and O'Malley, B.W. (1988) *Science*, **242**, 1702-1705.
- Insall, R. and Kay, R.R. (1990) *EMBO J.*, **9**, 3323-3328.
- Issemann, I. and Green, S. (1990) *Nature*, **347**, 645-650.
- Kobel, H.R. and Du Pasquier, L. (1986) *Trends Genet.*, **2**, 310-315.
- Koelle, M.R., Talbot, W.S., Segraves, W.A., Bender, M.T., Cherbak, P. and Hogness, D.S. (1991) *Cell*, **67**, 59-77.
- Ladias, J.A.A. and Karathanasis, S.K. (1991) *Science*, **251**, 561-565.
- Laudet, V., Bègue, A., Henry-Dutoit, C., Joubel, A., Martin, P., Stèhelin, D. and Saule, S. (1991) *Nucleic Acids. Res.*, **5**: 1105-1112.
- Lavorgna, G., Ueda, H., Clos, J. and Wu, C. (1991) *Science*, **252**, 848-851.
- Lazar, M.A., Hodin, R.A., Darling, D.S. and Chin, W.W. (1989) *Mol. Cell. Biol.*, **9**, 1128-1136.
- Lehmann, J.M., Hoffmann, B. and Pfahl, M. (1991) *Nucleic Acids. Res.*, **19**, 573-578.
- Mangelsdorf, D.J., Ong, E.S., Dyck, J.A. and Evans, R.M. (1990) *Nature*, **345**, 224-229.
- Milbrandt, J. (1988) *Neuron*, **1**, 183-188.
- Misrahi, M., Atger, M., D'auriol, L., Loosfelt, H., Meriel, C., Fridlansky, F., Guiochon-Mantel, A., Galibert, F. and Milgrom, E. (1987) *Biochem. Biophys. Res. Commun.*, **143**, 740-748.
- Miyajima, N., Kadowaki, Y., Fukusjoge, S.I., Shimizu, S.I., Semba, K., Yamashita, Y., Matsubara, K.I., Toyoshima, K. and Yamamoto, T. (1988) *Nucleic Acids. Res.*, **16**, 11057-11066.
- Miyajima, N., Horiuchi, R., Shibuya, Y., Fukushige, S.I., Matsubara, K.I., Toyoshima, K. and Yamamoto, T. (1989) *Cell*, **57**, 31-39.
- Mlodzik, M., Hiromi, Y., Weber, U., Goodman, C.S. and Rubin, G.M. (1990) *Cell*, **60**, 211-224.
- Moore, D.D. (1990) *New Biol.*, **2**, 100-105.
- Nauber, U., Pankratz, M.J., Kienlin, A., Seifert, E., Klemm, U. and Jäckle, H. (1988) *Nature* **336**, 489-492.
- Nei, M. (1987) *Molecular Evolutionary Genetics*. Columbia University Press, New York.
- Nirdé, P., Torpier, G., De Reggi, M.L. and Capron, A. (1983) *FEBS Lett.*, **151**, 223-227.
- O'Donnell, A.L. and Koenig, R.J. (1990) *Mol. Endocrinol.*, **4**, 715-720.
- O'Donnell, A.L., Rosen, E.D., Darling, D.S. and Koenig, R.J. (1991) *Mol. Endocrinol.*, **5**, 94-99.
- O'Malley, B.W. (1989) *Endocrinology*, **125**, 1119-1120.
- Oro, A.E., Ong, E.S., Margolis, J.S., Posakony, J.W., McKeown, M. and Evans, R.M. (1988) *Nature*, **336**, 493-496.
- Oro, A.E., McKeown, M. and Evans, R.M. (1990) *Nature*, **347**, 298-299.
- Pankratz, M.J. and Jäckle, H. (1990) *Trends Genet.*, **6**, 287-292.
- Petkovitch, M., Brand, N.J., Krust, A. and Chambon, P. (1987) *Nature*, **330**, 444-450.
- Pignoni, F., Baldarelli, R.M., Steingrimsson, E., Diaz, R.J., Patapoutian, A., Merriam, J.R. and Lengyel, J.A. (1990) *Cell*, **62**, 151-163.
- Ponglikitmongkol, M., Green, S. and Chambon, P. (1988) *EMBO J.*, **7**, 3385-3388.
- Ragsdale, C.W., Petkovitch, M., Gates, P.B., Chambon, P. and Brockes, J.P. (1989) *Nature*, **341**, 654-657.
- Rothe, M., Nauber, U. and Jäckle, H. (1989) *EMBO J.*, **8**, 3087-3094.
- Ryseck, R.P., Macdonald-Bravo, H., Mattéi, M.G., Ruppert, S. and Bravo, R. (1989) *EMBO J.*, **8**, 3327-3335.
- Saitou, N. and Nei, M. (1987) *Mol. Biol. Evol.*, **4**, 406-425.
- Sap, J., Munoz, A., Damm, K., Goldberg, Y., Ghysdael, J., Leutz, A., Beug, H. and Vennström, B. (1986) *Nature*, **324**, 635-640.
- Schwabe, J.W.R. and Rhodes, D. (1991) *Trends Biochem. Sci.*, **16**, 291-296.
- Segraves, W.A. (1991) *Cell*, **67**, 225-228.
- Segraves, W.A. and Hogness, D.S. (1990) *Genes Dev.*, **4**, 204-219.
- Shea, M.J., King, D.L., Conboy, M.J., Mariani, B.D. and Kafatos, F.C. (1990) *Genes Dev.*, **4**, 1128-1140.
- Sladek, F.M., Zhong, W., Lai, E. and Darnell Jr, J.E. (1990) *Genes Dev.*, **4**, 2353-2365.
- Sneath, P.H.A. and Sokal. (1973) *Numerical Taxonomy*. W.H. Freeman, San Francisco.
- Swofford, D.L. and Olsen G.J. (1990) In Hillis, D.M. and Moritz, C. (eds), *Molecular Systematics*. Sinauer Associates Inc., Sunderland, MA, pp.411-501.
- Thompson, C.C., Weinberger, C., Lebo, R. and Evans R.M. (1987) *Science*, **237**, 1610-1614.
- Tilley, W.D., Marcelli, M., Wilson, J.D. and McPhaul, M.J. (1989) *Proc. Natl. Acad. Sci. USA*, **86**, 327-331.
- Valentine, J.W. (1977) In Hallam, A. (ed.), *Pattern of Evolution*. Elsevier, Amsterdam, pp. 27-58.
- Wang, L.H., Tsai, S.Y., Cook, R.G., Beattie, W.G., Tsai, M.J. and O'Malley, B.W. (1989) *Nature*, **340**, 163-166.
- Weinberger, C., Thompson, C.C., Ong, E.S., Lebo, R., Gruol, D.J. and Evans, R.M. (1986) *Nature*, **324**, 641-646.
- Xiong, Y. and Eickbush, T.H. (1988) *Mol. Biol. Evol.*, **5**, 675-690.
- Xiong, Y. and Eickbush, T.H. (1990) *EMBO J.*, **9**, 3353-3362.
- Yaoyita, Y., Shi, Y.B. and Brown, D.D. (1990) *Proc. Natl. Acad. Sci. USA*, **87**, 7090-7094.
- Zelent, A., Krust, A., Petkovitch, M., Kastner, P. and Chambon, P. (1989) *Nature*, **339**, 714-717.

Received on July 30, 1991; revised on November 22, 1991

## RESUME DES RESULTATS DE L'ARTICLE 2

Notre analyse a donc consisté dans un premier temps, à réaliser un alignement soigneux des séquences des domaines C et E des récepteurs nucléaires. En ce qui concerne le domaine C, l'alignement est assez facile à réaliser, soit manuellement, soit à l'aide d'un programme d'alignement multiple de séquences, comme le programme CLUSTAL (ou sa version plus élaborée CLUSTAL V ; Higgins et Sharp, 1988). Les résultats obtenus par ces deux méthodes sont très semblables, parce que l'homologie existant entre les différents récepteurs dans cette région est forte. La taille du domaine C varie de 67 acides aminés (chez le récepteur orphelin PPAR ; Isseman et Green, 1990) à 70 acides aminés chez les récepteurs de l'acide rétinoïque et des hormones thyroïdiennes (figure 42). La grande majorité des récepteurs possède 68 acides aminés dans ce domaine. Il a donc été nécessaire de placer des coupures ("gaps" c'est à dire "trous" en anglais) dans les différentes séquences entre les deux doigts de zinc dans une région où le degré de conservation est bas. La position des coupures a été déterminée en tenant compte du fait que les récepteurs possédant un plus grand nombre d'acides aminés dans le domaine C ont dû acquérir ces acides aminés supplémentaires de façon indépendante les uns des autres. Il n'y a donc pas de raison de penser que ces insertions aient eu lieu aux mêmes endroits pour les récepteurs des hormones thyroïdiennes et pour les récepteurs orphelins de type *ear-1* par exemple. Les résidus cystéine présents dans ce domaine sont bien sûr strictement conservés chez tous les récepteurs nucléaires. Cependant, il existe, chez deux récepteurs (PPAR et le gène de *Drosophile tailless*; Isseman et Green, 1990 ; Pignoni et al., 1990) une petite variation sur la position exacte de la première cystéine du deuxième doigt de zinc C2 (figure 42) que l'on peut expliquer par une délétion (PPAR) ou une insertion (*tailless*) de deux acides aminés entre les deux premières cystéines de C2.

En ce qui concerne le domaine Ti-DM, beaucoup plus grand et divergent entre les récepteurs, l'alignement a été réalisé en utilisant le programme CLUSTAL, avec des paramètres qui défavorisent l'apparition de coupures trop nombreuses ou trop longues entre les séquences. L'alignement obtenu a parfois été retouché à la main, de façon à favoriser l'alignement des résidus impliqués dans les répétitions en "heptades", importantes pour la dimérisation des récepteurs (figure 42 ; Glass et al., 1989 ; Forman et al., 1989 ; Forman et Samuels, 1990a ; 1990b ; Fawell et al., 1990). Ce type d'alignement est utilisé lorsque les séquences étudiées sont très homologues à l'intérieur d'un groupe (par exemple les récepteurs des hormones thyroïdiennes) mais très divergentes entre groupes (les récepteurs des hormones thyroïdiennes et les récepteurs des stéroïdes). La même analyse a été utilisée par d'autres auteurs pour étudier l'évolution des transcriptases inverses, pour lesquelles un phénomène équivalent existe (Xiong et Eickbush, 1988 ; 1990).

Le pourcentage d'homologie entre les différentes séquences est beaucoup plus élevé dans la région Ti. Ainsi, dans ce domaine Ti, il existe deux résidus conservés chez tous les récepteurs



étudiés. On note tout de même que, pour tout le domaine Ti-DM, il existe environ un vingtaine de résidus conservés dans plus de 80% des gènes.

Cet alignement montre qu'il existe clairement deux domaines conservés chez les récepteurs nucléaires : le domaine C et le domaine Ti-DM. Cependant, cette loi générale a ses exceptions. En effet, le domaine Ti-DM n'existe pas chez trois récepteurs nucléaires de drosophile, tous proches les uns des autres: ce sont les gènes *knirps* (KNI; Nauber et al., 1988), *knirps-related* (KNRL; Oro et al., 1988) et *embryonic gonad* (EGON ; Rothe et al., 1989 ; revue dans Pankratz et Jäckle, 1990). Il a été impossible de trouver la moindre similitude entre ces gènes et les autres récepteurs dans le domaine E. Ceci a également été observé par les différents chercheurs qui les ont décrit. A la place du domaine Ti-DM, ces trois récepteurs orphelins ont un domaine E différent qui, de plus, ne présente que très peu d'homologie entre les trois gènes (10 à 13% inégalement répartis). On note cependant l'existence d'une courte séquence commune en C terminal de leur domaine C : la boîte KNI (Rothe et al., 1989). Il semble cependant qu'il existe une petite convergence entre le domaine E de ces gènes et le domaine DM des récepteurs classiques. En effet, le domaine E de KNI, KNRL et EGON est assez riche en résidus leucine, comme le domaine DM. La signification de cette richesse en leucine reste à explorer. Ces trois gènes ont une caractéristique commune supplémentaire : ils sont tous les trois dépourvus de domaine A/B. La première cystéine de C1 est située à 4 acides aminés du codon d'initiation pour KNI et EGON et à 14 acides aminés pour KNRL. Cette observation renforce l'hypothèse d'une origine commune de ces trois gènes malgré la faible homologie constatée dans le domaine E.

A partir de ces alignements, qui représentent une étape capitale dans toute analyse phylogénique de séquences, nous avons construit des arbres en utilisant deux méthodes différentes: d'une part celle "des moindres carrés", également appelée méthode de Fitch et Margoliash (Fitch et Margoliash, 1967 ; Fitch, 1981) et d'autre part la méthode "des plus proches voisins" (ou "Neighbor Joining" ; Saitou et Nei, 1987). Ces méthodes utilisent toutes deux des matrices de distances entre les séquences prises deux à deux, en partant des pourcentages d'homologie entre ces séquences. En utilisant ces deux méthodes qui doivent se compléter et se valider mutuellement, nous avons étudié indépendamment l'évolution des deux domaines C et Ti-DM. De cette analyse, nous pouvons tirer quatre séries de conclusions:

**1) Les récepteurs nucléaires peuvent être classés en trois sous familles (figure 43).** La sous-famille I regroupe les récepteurs des hormones thyroïdiennes (THR) et de l'acide rétinoïque (RAR) et les récepteurs orphelins de type *ear-1* (*ear-1*, E75 et PPAR). La sous-famille II comprend uniquement des récepteurs orphelins : ceux du groupe COUP, ceux de type RXR et les récepteurs orphelins HNF4 et *tailless*. Enfin, dans la sous-famille III, on trouve les récepteurs des stéroïdes, ceux de la vitamine D et de l'ecdysone, et les gènes du groupe *knirps*. Cette classification est valable pour les deux domaines C et Ti-DM bien que, nous allons le voir, un

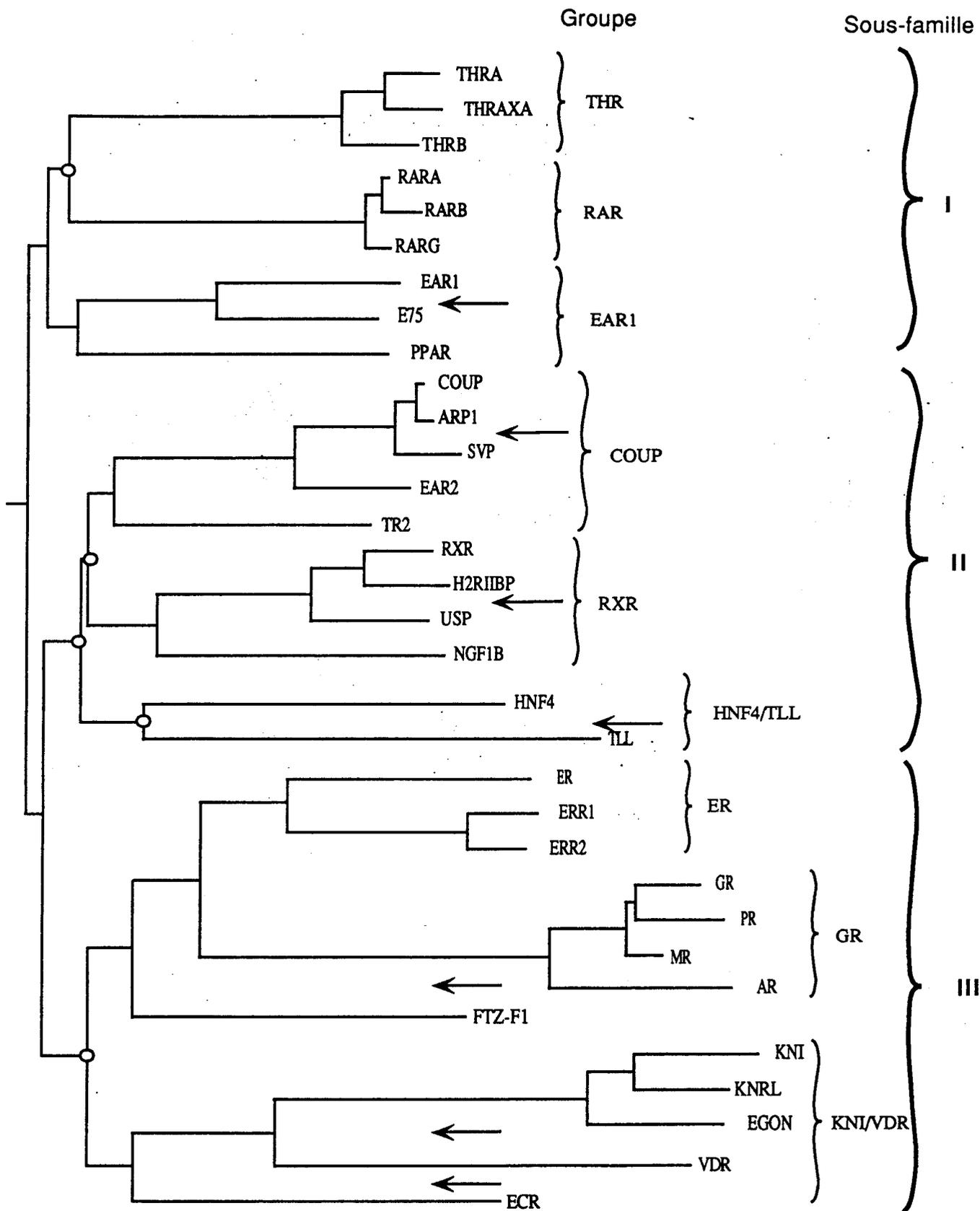


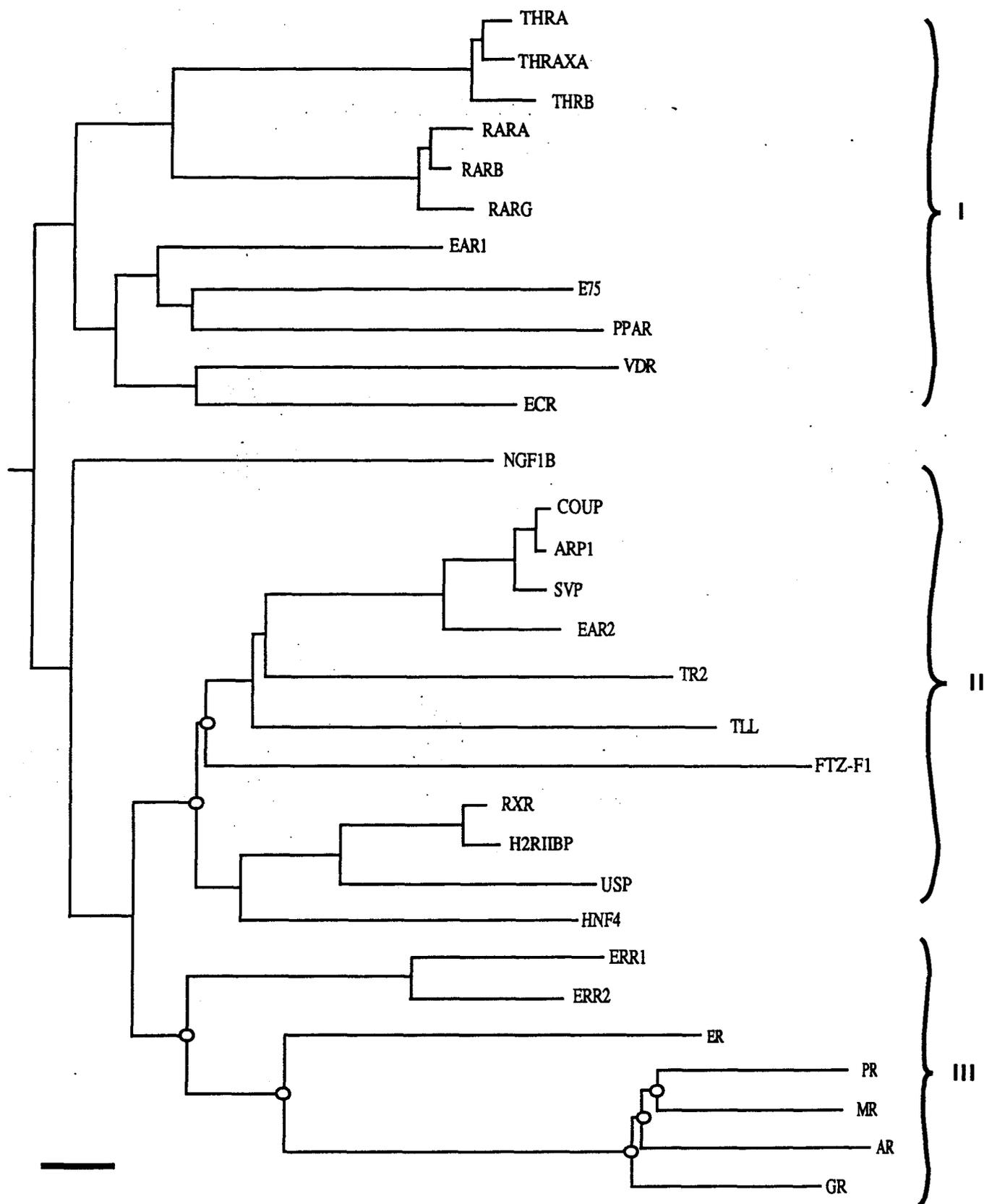
Figure 43 : Arbre phylogénique construit par la méthode de Fitch (moindres carrés) avec les séquences du domaine C. Les petits cercles indiquent les points de divergence entre les arbres construits par la méthode de Fitch et ceux construits par la méthode NJ. Les flèches indiquent la dichotomie arthropodes/vertébrés

certain nombre de différences importantes existent entre les arbres bâtis en utilisant les séquences de chacun des deux domaines. On peut même séparer chaque sous-famille en trois groupes, à partir de l'analyse réalisée sur le domaine C. Ainsi pour la sous-famille I, on a les groupes THR, RAR et EAR-1. Pour la sous-famille II, on trouve les groupes COUP, RXR et HNF4/TLL. Enfin, pour la sous-famille III, on note l'existence des groupes ER, GR et VDR/KNI. Dans les arbres construits avec la méthode Neighbor-Joining ce dernier groupe VDR/KNI se branche entièrement sur les récepteurs des hormones thyroïdiennes dans la sous-famille I. Cette divergence importante entre les deux méthodes sera discutée en détail plus loin.

La sous-famille II est particulièrement intéressante, parce qu'elle regroupe des gènes qui n'étaient préalablement classés dans aucune sous-famille et entre lesquels on n'a jamais décrit de relations structurales ou fonctionnelles particulières. Il est tout à fait singulier que le gène NGF1B, qui possède une organisation génomique unique (Ryseck et al., 1989), fasse partie de cette sous-famille. Rappelons que ces auteurs ont imaginé que ce récepteur pouvait définir à lui seul une nouvelle sous-famille au sein des récepteurs nucléaires. Est-ce à dire que tous les autres récepteurs appartenant à cette sous-famille ont les mêmes particularités structurales ? Nous l'ignorons pour l'instant. Tous ces récepteurs orphelins ont pourtant un ancêtre commun facilement visible sur les arbres construits à partir du domaine C ou du domaine Ti-DM. Si l'on fait une analogie avec les deux autres sous familles, on peut donc suggérer que, en plus de cette filiation évolutive, ces différents gènes doivent partager "quelque chose" au niveau fonctionnel. Ligand, dimérisation, sites de fixation sur l'ADN, il est difficile de raisonner pour l'instant sur ce quelque chose...

Il est important de noter, à ce stade, que les deux analyses utilisées donnent des arbres sans racine. Le positionnement d'une racine dans de tels arbres ne peut être fait correctement qu'en introduisant un groupe externe (c'est-à dire des séquences assez proches pour montrer une homologie, mais néanmoins clairement différentes). Un tel groupe n'est pas disponible pour les récepteurs nucléaires et donc nous avons placé la racine de manière arbitraire mais cohérente avec ce que l'on connaît des différents groupes de récepteurs. La racine a ainsi été placée entre la sous-famille I et les deux autres. Ce positionnement précis des racines est seulement une tentative pour donner une image aussi cohérente que possible de la réalité, mais peut tout à fait être discutée. Ainsi on aurait pu placer la racine entre la sous-famille II et les autres, ou encore entre la sous-famille III et les deux autres. Il est clair que l'ordre des toutes premières duplications de "l'ancêtre commun" de tous les récepteurs nucléaires ne peut pas être déterminée pour l'instant.

**2) Les récepteurs nucléaires sont apparus très tôt au cours de l'évolution (figure 43).** Si l'on compare l'apparition des récepteurs nucléaires de vertébrés avec ceux d'arthropodes (connus uniquement chez la drosophile) on note que leur divergence au sein des arbres est plutôt tardive. On note, de plus, que toutes les sous familles contiennent des récepteurs de vertébrés et d'arthropodes. Dans certains cas, comme la sous-famille II, chaque groupe contient des gènes de vertébrés et d'arthropodes. On sait que la divergence entre les lignées qui ont donné



**Figure 44 : Arbre phylogénique construit par la méthode de Fitch (moindres carrés) avec les séquences du domaine Ti-DM. Les petits cercles indiquent les points de divergence entre les arbres construits par la méthode de Fitch et ceux construits par la méthode NJ.**

KN1  
KNRL  
EGON

naissance aux arthropodes et aux vertébrés a eu lieu il y a environ 500 millions d'années (Hartland et al., 1982). Ceci suggère que l'évolution de la superfamille des récepteurs nucléaires, en terme de sous familles et même de groupes de gènes, était déjà pratiquement faite à cette date. Nous tirerons les importantes conséquences de cette observation dans la discussion.

**3) Les récepteurs nucléaires ont évolué par simple duplication de gènes, mais il existe d'intéressantes exceptions (figure 44).** La comparaison des arbres bâtis à partir des séquences du domaine C ou du domaine Ti-DM montre clairement que, pour leur grande majorité, les gènes de récepteurs nucléaires se sont dupliqués au cours de l'évolution pour donner lieu à la diversification impressionnante que l'on observe à présent. En effet, les deux types d'arbres sont pour l'essentiel superposables d'un point de vue topologique : l'ordre des différentes duplications est le même. Il existe pourtant des exceptions importantes qui peuvent être regroupées en deux types majeurs :

- Les gènes du groupe KNI ne possèdent pas de domaine Ti-DM. Il ne sont donc pas classables au sein des récepteurs nucléaires pour ce domaine. Pourtant, ils possèdent un domaine C tout à fait typique qui les place dans la sous-famille III avec les récepteurs des stéroïdes. Il semble clair qu'au cours de l'évolution ces trois gènes ont perdu leur domaine Ti-DM et l'ont remplacé par un domaine différent. Cet événement ne s'est sans doute produit qu'une seule fois au cours de l'évolution, lors des étapes qui ont précédé l'apparition des drosophiles actuelles. En effet, les trois gènes du groupe KNI sont relativement proches les uns des autres et ont tous en commun une séquence particulière appelée boîte KNI. Une fois ce remplacement du domaine Ti-DM par un domaine différent, il est probable que les séquences qui ont remplacé le domaine Ti-DM ont, pour une raison inconnue, divergées très vite pour donner les gènes actuels qui possèdent 3 domaines E différents. Un tel événement (le remplacement d'un domaine par un autre complètement différent) est assez facile à expliquer par une simple translocation chromosomique. Le fait que l'on n'a, pour l'instant, pas trouvé de gène de vertébré homologue aux gènes du groupe KNI est un bon argument pour placer cet événement spécifiquement dans la lignée des drosophiles. Nous avons signalé que les membres du groupe *knirps* n'ont pas de domaine A/B. Il est amusant de constater que leur plus proche homologue, le gène codant pour le récepteur de la vitamine D, a un domaine A/B court (une vingtaine d'acides aminés). Il y a donc probablement eu perte totale ou partielle de ce domaine chez un ancêtre commun à tous ces gènes. Chez les autres membres de la sous-famille III, le domaine A/B est en effet de grande taille.

- Les autres récepteurs n'obéissant pas à la règle de la simple duplication sont plus nombreux, et leur apparition est plus difficile à expliquer. Les deux exemples les plus spectaculaires sont les récepteurs de la vitamine D (VDR) et de l'ecdysone (ECR) qui appartiennent à la sous-famille III des récepteurs des stéroïdes (si l'on considère le domaine C) et à la sous-famille I des THR et RAR (si l'on observe le domaine Ti-DM). Ces deux gènes peuvent donc être considérés comme des "chimères évolutives". D'autres exemples moins spectaculaires existent.

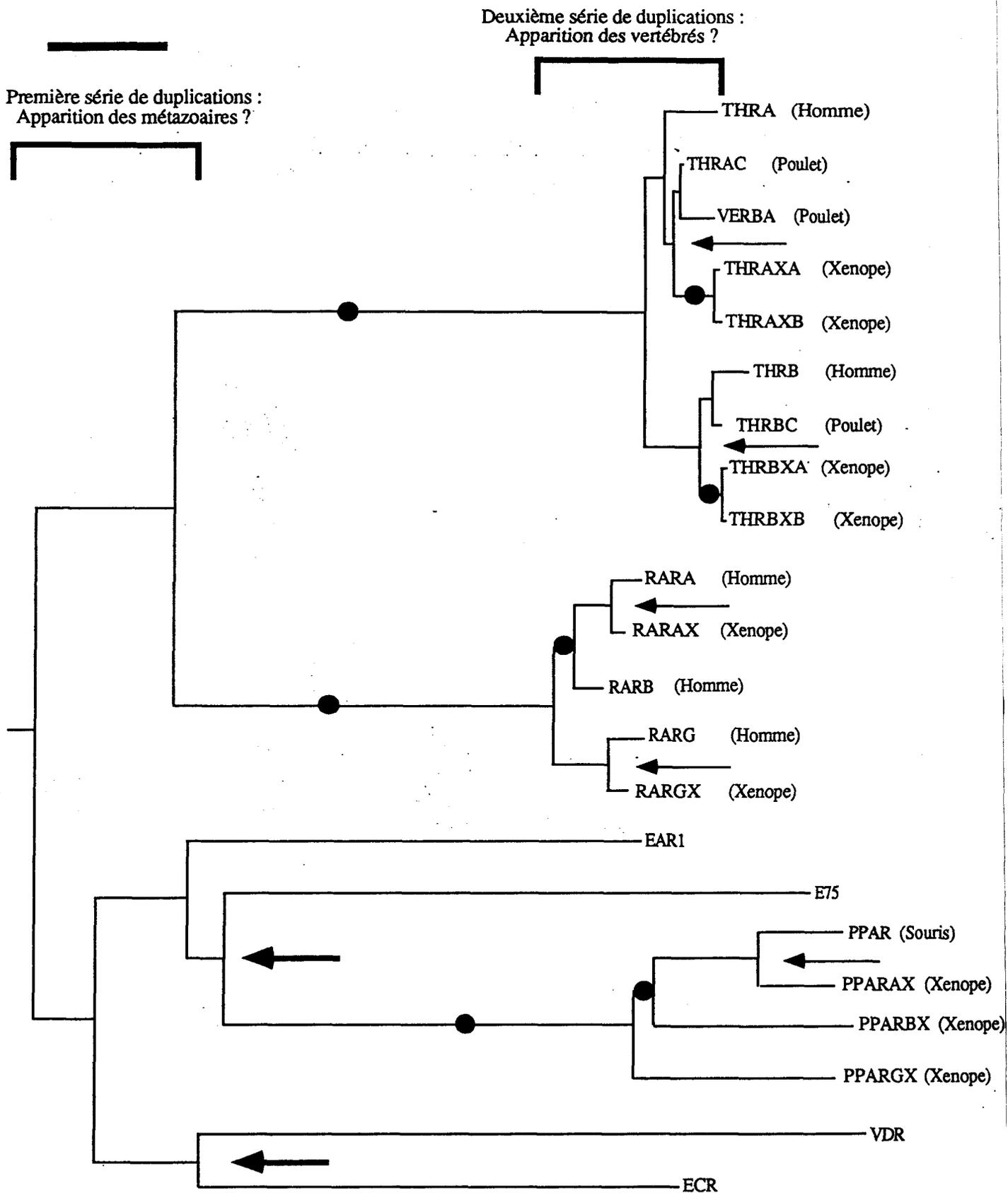


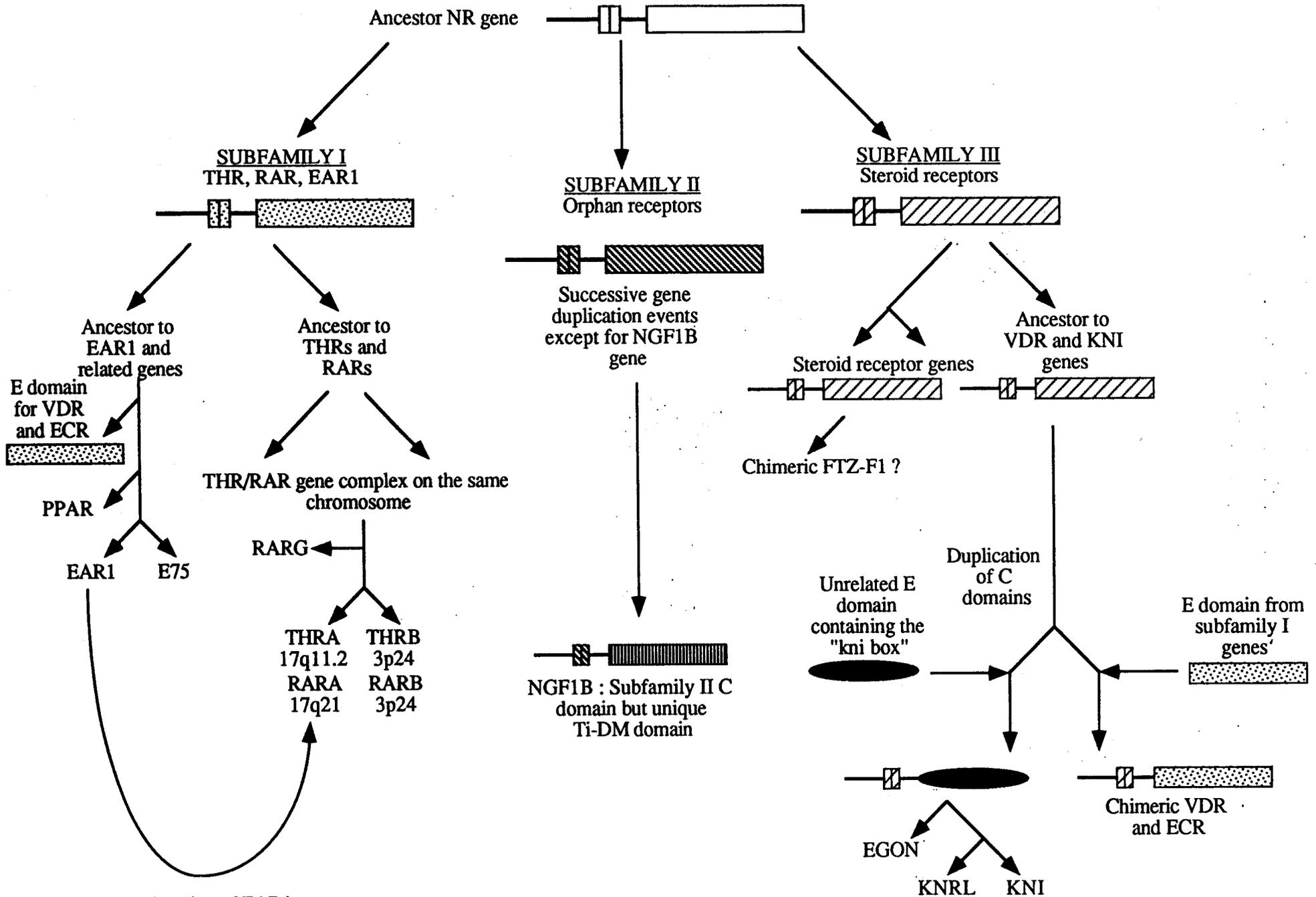
Figure 45 : Arbre phylogénique construit par la méthode de Fitch (moindres carrés) avec les séquences du domaine TI-DM des membres de la sous famille I, y compris les trois nouveaux récepteurs PPAR de Xénope (XPPARA, XPPARB et XPPARG) (W.Wahli, com. pers.). Les cercles noirs indiquent, pour les groupes THR, RAR et PPAR, les événements de duplication de gènes. Les petites flèches indiquent la dichotomie amphibiens/mammifères et les grosses flèches, la dichotomie arthropodes/vertébrés).

C'est semble-t-il le cas des récepteurs orphelins NGF1B et FTZ-F1 et, dans une moindre mesure des gènes HNF4 et *tailless*. Dans tous ces cas, et surtout ceux qui, comme HNF4 et *tailless* semblent être des chimères "à petite échelle", il faut faire soigneusement la différence entre ce qui est le résultat d'un évènement évolutif spectaculaire conduisant à un récepteur chimère, et une convergence évolutive ou une vitesse d'évolution des deux domaines très différente permettant d'expliquer plus simplement la disposition observée. Dans chaque cas, il faut faire la part des choses et tenter de confirmer l'hypothèse de la chimère par d'autres arguments, si cela est possible. Il faut noter ici que la majorité des récepteurs montrant une évolution différente des domaines C et Ti-DM appartiennent à la sous-famille III pour le domaine C (ECR, VDR, FTZ-F1, gènes du groupe KNI). Cette observation a peut-être une signification, mais laquelle ?

4) Les THR et RAR ont évolué en parallèle (figure 45). Il était intéressant, au cours de cette étude d'observer plus précisément l'évolution de la sous-famille I contenant les THR, les RAR et les membres du groupe EAR-1. En effet, la localisation chromosomique de certains membres de cette famille laissait penser que ces gènes avaient pu se dupliquer en même temps. Ainsi, sur le bras long du chromosome 17, on trouve le gène codant THRA (c'est le gène c-erbA-1, qui est en 17q11.2) et celui codant RARA (en 17q21). De même sur le bras court du chromosome 3 en p24) se trouvent les gènes THRB (c-erbA-2) et RARB. Ceci suggère que ces gènes ont subi ensemble des événements de duplication successifs. L'analyse de l'arbre construit avec les séquences des différents membres de la sous-famille I (figure 45) dans le domaine Ti-DM à confirmé cette hypothèse. On remarque en effet une première duplication de gène qui a séparé les récepteurs des hormones thyroïdiennes (THR) de ceux de l'acide rétinoïque (RAR). Probablement, les deux gènes issus de cette duplication sont restés sur le même chromosome à proximité l'un de l'autre (figure 46). Alors, le gène RARG (le plus divergent des RAR) s'est séparé de son ancêtre et est venu se placer sur le chromosome 12 actuel. Plus tard, ce locus entier s'est dupliqué pour donner THRA et RARA sur le chromosome 17 et THRB et RARB sur le chromosome 3. Il est probable qu'ensuite, au moins dans le chromosome 17, d'autres gènes soient venus s'insérer entre THRA et RARA, ce qui explique la distance assez grande qui les sépare.

Une raison supplémentaire qui nous a poussés à observer de plus près l'évolution de cette sous-famille est l'existence du curieux locus c-erbA-1/ear-1 imbriqué, présent chez les mammifères. Il est clair, d'après l'arbre que nous avons construit, que ces deux gènes se sont séparés depuis fort longtemps. En effet, leur ancêtre commun est l'ancêtre de toute la sous-famille I. L'évènement d'imbrication est donc un évènement récent, ce qui est bien confirmé par son absence chez le poulet.

A titre d'anecdote, on peut noter que le gène v-erbA a, comparativement à son progéniteur le gène c-erbA-1 de poulet, évolué d'une manière particulièrement rapide. Ceci n'est pas étonnant, car au sein du génome du rétrovirus AEV, le gène v-erbA a dû subir un nombre très élevé de réplifications virales successives. Or, on sait que l'enzyme responsable de ce processus, la



Translocation of EAR1 gene in the C-terminal part of the THRA locus

Figure 46 : Un scénario possible d'évolution de la superfamille des récepteurs nucléaires.

transcriptase inverse, commet beaucoup d'erreurs.

Enfin, l'arbre construit avec toutes les séquences des gènes THRA (c-erbA-1) et THRB (c-erbA-2) de vertébrés (pour les mammifères, seules les séquences humaines ont été incluses parce que la conservation est très forte) permet de visualiser l'évènement de duplication du génome observé chez le xénope (figure 45). En effet, chez cet animal, il existe deux gènes c-erbA-1 et deux gènes c-erbA-2 (Yaoita et Brown, 1990a). Cet évènement de tétraploïdisation est bien connu chez le xénope (Kobel et Du Pasquier, 1986 ; Bisbee et al., 1977). Si ce que nous avons dit sur l'évolution concomitante des gènes THRA et RARA d'une part et THRB et RARB d'autre part, est vrai, alors on doit s'attendre à ce qu'il existe deux gènes RARA et deux gènes RARB chez le xénope sachant que la tétraploïdisation a affecté la grande majorité du génome de cet animal. On ignore encore ce qu'il en est...

## DISCUSSION

## A) STRUCTURE ET EVOLUTION DU LOCUS c-erbA-1/ear-1.

Nous avons donc observé deux gènes, sur le chromosome 17, appartenant tous deux à la superfamille des récepteurs nucléaires, qui se chevauchent partiellement. Le fait qu'une telle organisation, très inhabituelle, soit conservée au moins chez les mammifères suggère qu'elle a une importance pour la physiologie de ces organismes.

### 1) Les rôles possibles de l'imbrication des gènes c-erbA-1 et ear-1.

#### a) Des contraintes structurales.

Il faut rappeler que dans ce locus imbriqué, le chevauchement a lieu entre un exon alternatif de c-erbA-1 (l'exon 10) et le dernier exon de ear-1 (appelé F). De plus, cet épissage est particulier, parce qu'il implique deux événements distincts : la reconnaissance alternative d'un site donneur d'épissage au début de l'exon 9 et, en 3', une polyadénylation alternative concernant le site situé à la fin de l'exon 9, ou le site situé à la fin de l'exon 10. Si, comme on l'affirme communément (Wickens, 1990 ; Maniatis, 1991), la polyadénylation a bien lieu avant l'épissage des ARN, c'est cet événement de polyadénylation alternatif qui va décider de l'ARN messager (5kb ou 2,6kb) finalement produit. En effet, si c'est le site de polyadénylation situé à la fin de l'exon 9 qui est utilisé, on ne pourra plus fabriquer l'ARN de 2,6 kb. Outre l'épissage, les polyadénylations alternatives entraînent le "choix" de séquences 3' non codantes de tailles très différentes (3000 bases contre 360 environ) qui peuvent influencer la traduction des ARN (pour une revue voir Jackson et Standart, 1990).

Ces événements de polyadénylation et d'épissage alternatif sont importants à prendre en compte pour l'étude de la signification fonctionnelle de l'imbrication des gènes c-erbA-1 et ear-1. En effet, seul l'exon 10 est dans la zone imbriquée. L'exon 9 de c-erbA-1 est en effet situé environ 250 paires de bases en amont de l'exon F de ear-1. Comme la terminaison de la transcription, un mécanisme mal connu chez les eucaryotes, semble pouvoir se produire très loin des sites de polyadénylation il est possible que, d'un point de vue transcriptionnel, d'autres exons des deux gènes soient impliqués dans le chevauchement. On peut concevoir que l'un des mécanismes contrôlant l'épissage alternatif de c-erbA-1 soit la transcription du locus ear-1 (Proudfoot, 1989). Si ce locus est transcrit à très haut taux, à un moment précis ou dans un tissu donné, on pourra éventuellement favoriser la formation de l'ARN de 5 kb qui n'est pas impliqué dans le chevauchement.

b) Des modèles pour le chevauchement des deux gènes.

Un des problèmes posés par le chevauchement des parties 3' des gènes c-erbA-1 et ear-1 est de savoir comment ces deux gènes peuvent s'exprimer. En effet, tout au long de la partie chevauchante, la RNA polymérase II ne peut pas parvenir à transcrire en même temps les deux brins d'ADN, tout simplement pour des problèmes d'encombrement stérique. A un niveau ou à un autre, il faut donc un mécanisme régulateur de l'expression de ces deux gènes. L'étude de l'expression de ces deux gènes devrait permettre de savoir à quels niveaux se fait cette régulation. Sans préjuger des mécanismes impliqués, on peut imaginer trois niveaux possibles de régulation :

- La régulation peut se faire au niveau des tissus. Dans ce modèle, chaque gène serait exprimé au sein de tissus différents d'une manière mutuellement exclusive. C'est le cas le plus favorable, parce qu'il est facilement visualisable à partir d'une étude assez simple de "Northern blot". On peut dire tout de suite que ce modèle est excessivement simpliste car on connaît de nombreux tissus où les deux gènes (ear-1 et la version  $\alpha 2$  de c-erbA-1) s'expriment.

- La régulation peut se faire au niveau des cellules. On peut imaginer qu'au sein d'un tissu donné, certaines cellules expriment ear-1 et d'autres c-erbA-1. Une étude par "Northern blot" ne permettra pas de répondre facilement à cette question et seule la technique d'hybridation *in situ* peut le faire. Même avec cette technique, ce ne sera pas chose aisée, parce qu'il sera techniquement difficile de démontrer qu'une cellule donnée exprime un seul des deux gènes.

- La régulation peut se faire au sein de chaque cellule. Dans ce cas, le taux d'expression observé sera la résultante d'un phénomène de "balance" entre les taux d'expression de ces deux gènes. Selon les conditions expérimentales, on doit donc pouvoir modifier cette balance et observer comment réagissent les deux gènes. Si les deux gènes sont exprimés ensemble dans la même cellule, les deux ARN messagers transcrits vont pouvoir s'hybrider et ainsi être dégradés par des RNases cellulaires spécifiques des ARN doubles brins.

c) Le taux d'expression de ear-1 semble influencer sur l'épissage de c-erbA-1.

Les données obtenues à ce jour sur l'expression des gènes de ce locus montrent clairement qu'il n'y a pas d'expression tissulaire spécifique de c-erbA-1 ou de ear-1. Au contraire, les deux gènes sont exprimés dans les mêmes tissus (par exemple le coeur ou le cerveau) avec des intensités tout à fait comparables (Lazar et al., 1989a; 1990a; Miyajima et al., 1989). Aucune donnée d'hybridation *in situ* permettant de savoir si une même cellule peut exprimer les deux gènes en même temps n'a été publiée. Cependant, il est infiniment probable que ce soit le cas, puisque, nous allons le voir, la transcription d'un des deux gènes peut influencer la transcription de l'autre dans des cellules en culture.

Le gène ear-1 est très fortement exprimé dans le tissu musculaire strié squelettique ainsi que dans le tissu adipeux brun du rat (Lazar et al., 1989a). Ces deux tissus sont les deux sites majeurs de la thermogenèse chez les mammifères. Ce phénomène est sous la dépendance des hormones thyroïdiennes (Silva, 1986; Klingenberg, 1990). Curieusement, dans ces deux tissus

l'expression de l'ARN c-erbA-1 $\alpha$ 2 est très faible. Ceci pouvait laisser penser qu'un taux d'expression très élevé de ear-1 favorisait l'épissage conduisant à l'expression de la forme  $\alpha$ 1 de 5 kb. Une preuve directe de cette influence a été obtenue par Mitchel Lazar en traitant des cellules avec la cycloheximide (Lazar et al., 1990b). Cette drogue augmente en effet l'expression de ear-1, tant au niveau transcriptionnel qu'au niveau de la stabilité des ARNs. Ceci suggère que l'expression de ear-1 est sous la dépendance d'un répresseur labile qui diminue fortement son expression. Ce qui est intéressant, c'est que le rapport  $\alpha$ 1/  $\alpha$ 2 est augmenté de 10 fois lorsqu'on augmente le taux d'expression de ear-1 avec la cycloheximide. Pourtant, ni le taux de transcription, ni la stabilité des ARN qui sont codés par c-erbA-1, ne sont modifiés par la cycloheximide. Cette étude suggère donc que c'est bien l'augmentation de transcription de ear-1 qui, en défavorisant l'épissage de l'ARN de 2,6 kb, augmente par contrecoup la formation de l'ARN de 5 kb et donc le rapport  $\alpha$ 1/  $\alpha$ 2. Le fait qu'il n'y ait pas d'effet de la cycloheximide sur le taux d'expression (mesuré par élongation *in vitro* de chaînes d'ARN préformées ou "run-on") de la forme c-erbA-1 $\alpha$ 2 laisse penser que la diminution de la quantité de cet ARN est directement due à un effet sur l'épissage ou sur la polyadénylation.

#### d) D'autres mécanismes possibles.

Cette première étude fonctionnelle du locus imbriqué c-erbA-1/ear-1 permet comme nous venons de le voir de bâtir un modèle raisonnable de régulation post-transcriptionnelle de l'expression de ces gènes. Cependant la question reste encore largement ouverte et d'autres mécanismes possibles peuvent être testés.

L'effet post-transcriptionnel nécessite encore une preuve directe. Il est en effet important de montrer que l'augmentation de la transcription d'un des deux gènes influe sur celle de l'autre. Par exemple, dans le cerveau, où la forme  $\alpha$ 2 de c-erbA-1 est très fortement exprimée, on devrait s'attendre à une faible expression de ear-1. C'est semble-t-il ce qui se passe puisque l'expression de ear-1 et c-erbA-1 (la forme  $\alpha$ 1 ou la forme  $\alpha$ 2) visualisée par hybridation *in situ* montrent que ces gènes sont exprimés à des endroits différents du cerveau. Dans le télencéphale, par exemple, la forme  $\alpha$ 2 de c-erbA-1 et ear-1 ont des sites d'expression mutuellement exclusifs (Bradley et al., 1989).

Il est envisageable que les produits des gènes c-erbA-1 et ear-1 régulent réciproquement leurs expressions au niveau de l'initiation de la transcription. Dans ce modèle, chaque promoteur contiendrait un site de fixation pour le produit de l'autre gène. L'étude du promoteur humain de c-erbA-1, actuellement en cours au laboratoire, devrait permettre de répondre au moins partiellement à cette question. De plus, si on dispose de ce promoteur, on pourra étudier les événements qui régulent la transcription de c-erbA-1 et peut être, dans un proche avenir, les comparer avec ceux qui régulent la transcription de ear-1.

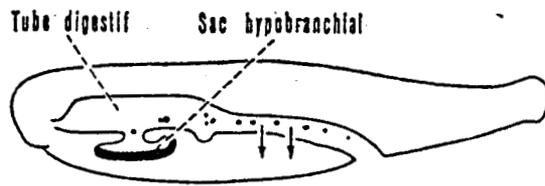
e) Quelle est la fonction de ear-1 ?

Evidemment, la signification de ce locus ne pourra être comprise que si l'on connaît la fonction de ear-1 dans l'organisme. Visiblement, ear-1 doit avoir un rôle important, puisque son expression est régulée au moins à deux niveaux: initiation de la transcription et stabilité des ARNs. Deux gènes sont assez proches de ear-1 : le gène de drosophile E75 inductible par l'ecdysone (Segraves et Hogness, 1990), et le gène de souris PPAR.

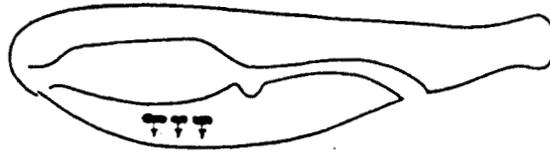
Le gène E75 semble être un homologue de drosophile de ear-1 ce qui suggère que ear-1 existe depuis très longtemps dans le génome. Comme ear-1, l'expression de E75 est très fortement augmentée par la cycloheximide. Curieusement, la partie 5' du gène E75 est l'objet d'un épissage alternatif unique chez les récepteurs nucléaires, qui supprime le premier doigt de zinc C1 pour donner une forme différente appelée E75B. Certains auteurs ont imaginé que E75 pourrait être le récepteur de l'hormone juvénile (Ashburner, 1990), mais pour l'instant aucune preuve n'est venue étayer cette hypothèse. La seule chose sûre, c'est que E75 fait partie des gènes des "puffs" de drosophile, exprimés très précocément après traitement par l'ecdysone.

Le gène PPAR de souris semble plus éloigné des deux autres, mais il fait tout de même partie du même groupe de gènes (voir article II). Le groupe de Walter Wahli, à Lausanne vient d'annoncer le clonage de trois PPAR de xénope dont deux sont des nouveaux gènes (W. Wahli commun. pers.; voir Dreyer et al., 1992). On sait que les produits des PPAR sont activés par des molécules carcinogènes (par exemple la nafénopine) qui provoquent la prolifération des peroxysomes et qui ont un effet hypolipidémiant (Issemann et Green, 1990; Dreyer et al., 1992). On ne sait pas encore si les produits de PPAR fixent ou non ces molécules, mais il existe une relation claire entre le lieu d'expression de PPAR alpha et les organes touchés par ces drogues (le foie et les reins). De même, on ignore les relations exactes qui unissent cette molécule et le nébuleux récepteur nucléaire de la dioxine (Johnson, 1990; Hoffman et Joosse, 1991).

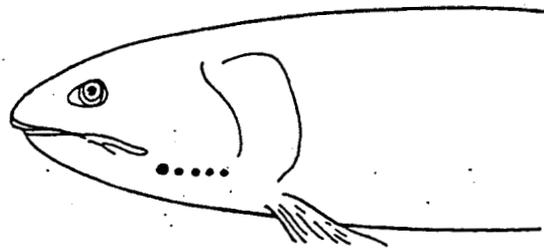
Il est raisonnable d'imaginer que le gène ear-1 ait un rôle relativement semblable à celui de ses deux congénères, mais ce n'est pas évident, car ces trois gènes sont assez distants les uns des autres. Par exemple, à première vue, on peut penser que les trois types de récepteurs orphelins vont se fixer sur des sites semblables sur l'ADN, puisqu'ils ont la même boîte P qui détermine la spécificité de fixation à l'ADN. Cependant, il faut rappeler que la boîte P du récepteur orphelin PPAR est très différente de celle des autres récepteurs, car elle a 2 acides aminés en moins. Cette région étant impliquée dans la dimérisation des récepteurs fixés à l'ADN, on peut donc imaginer que le produit de PPAR se fixe à l'ADN sous forme de dimères (Issemann et Green, 1990). Dans le même ordre d'idée, les domaines E des trois molécules n'ont qu'environ 40% d'homologie contre environ 80% entre les domaines E des différents récepteurs des hormones thyroïdiennes et il n'est pas du tout évident que ces molécules partagent le même type de ligand. En plus du faible pourcentage d'homologie existant dans le domaine E, il faut noter l'existence d'un grand domaine D chez ear-1 (246 acides aminés) et pas chez PPAR (115 acides aminés), et d'un très long domaine supplémentaire F de 635 acides aminés complètement divergent chez E75. Les



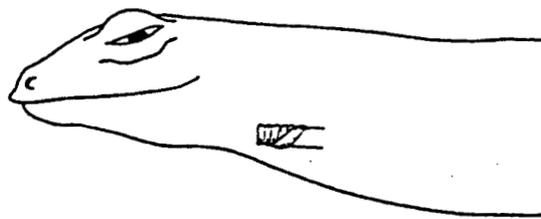
Ammocète : sac hypobranchial larve de Lamproie, Cyclostome.



Lamproie après métamorphose.



Poisson Téléostéen



Type général de Vertébré

Passage d'une glande exocrine (sécrétion dans le tube digestif) à une glande endocrine (sécrétion dans le milieu intérieur).

Figure 47 : Anatomie comparée de la glande thyroïde chez les Procordés et les premiers Vertébrés.

relations fonctionnelles qui existent éventuellement entre ces gènes restent donc largement à définir.

## 2) Evolution des récepteurs des hormones thyroïdiennes et du locus imbriqué.

Il semble clair qu'ear-1 a un homologue chez la drosophile, qui est le gène E75. Donc ear-1 semble ancien et doit être apparu très tôt. La question qui se pose est de savoir à quand remonte l'imbrication des locus c-erbA-1 et ear-1. Cette question pose, en corollaire, celle de l'apparition des récepteurs des hormones thyroïdiennes qui, à notre connaissance, n'existent pas chez la Drosophile.

### a) Evolution de la fonction thyroïdienne.

Chez les invertébrés, on connaît des tissus qui élaborent des molécules iodées qui peuvent faire penser aux hormones thyroïdiennes, mais ces substances ne semblent pas avoir de rôle hormonal (figure 47 ; Barrington, 1964; Fontaine, 1969). C'est chez les prochordés, avec par exemple les tuniciers (les "violets" de Méditerranée en font partie), que les hormones thyroïdiennes semblent acquérir ce rôle hormonal. En effet, chez ces organismes, on les trouve circulant dans le milieu intérieur. Cependant, elles ne sont pas fabriquées par une glande spécialisée, mais par divers tissus comme le manteau (ou endostyle) des ascidies. Chez les prochordés, les hormones thyroïdiennes sont sécrétées de manière exocrine et non pas endocrine. L'étude de la fonction thyroïdienne chez les tous premiers vertébrés comme les agnates (les "poissons" sans machoire que sont les lamproies et les myxines) nous offre une image très suggestive du passage d'une glande exocrine à une glande endocrine. En effet, chez les larves de lamproie c'est le sac hypobranchial (l'équivalent de l'endostyle des ascidies, qui provient du plancher du pharynx) qui élabore les hormones thyroïdiennes. Ce sac communique avec le pharynx et il y déverse une sécrétion muqueuse contenant des hormones thyroïdiennes. Ainsi celle-ci sont-elles libérées dans le tube digestif et non pas dans le sang : il s'agit bien d'une sécrétion exocrine. Lors de la métamorphose de la larve ammocète en lamproie adulte, les vésicules thyroïdiennes qui seront à sécrétion uniquement interne, se forment à partir du sac hypobranchial. Chez les lamproies comme chez les poissons, la glande thyroïde n'est pas un organe pleinement individualisé, mais elle se présente sous forme de follicules plus ou moins dispersés dans du tissu conjonctif, près de l'aorte ventrale.

On peut donc découper cette histoire en trois étapes : (i) existence de molécules d'hormones thyroïdiennes sans fonction thyroïdienne connue, chez les invertébrés; (ii) apparition d'une glande exocrine synthétisant les hormones thyroïdiennes chez les prochordés et les larves d'agnates, et enfin (iii) apparition d'une fonction thyroïdienne endocrine présente chez tous les vertébrés (figure 47). Il est bien sûr difficile de relier les arbres évolutifs que nous avons construits avec ces différentes étapes. Il est pourtant intéressant de préciser que (si la longueur des branches

de l'arbre est proportionnelle au temps, ce qui n'est pas sûr du tout !) l'ancêtre des récepteurs des hormones thyroïdiennes semble avoir évolué longtemps avant de se dupliquer en c-erbA-1 et c-erbA-2 (voir figure 4 de l'article II). Il est difficile de savoir si ce long passage sous forme de gène unique s'est produit avant l'apparition de la fonction endocrine de la glande thyroïdienne. Il faut néanmoins faire remarquer que chez les amphibiens il existe déjà deux gènes c-erbA bien individualisés qui sont très conservés si on les compare à leurs homologues respectifs chez les mammifères. Il faudrait savoir si ces gènes existent déjà chez les agnates, et surtout chez les prochordés, pour répondre à cette question.

#### b) Apparition du locus imbriqué.

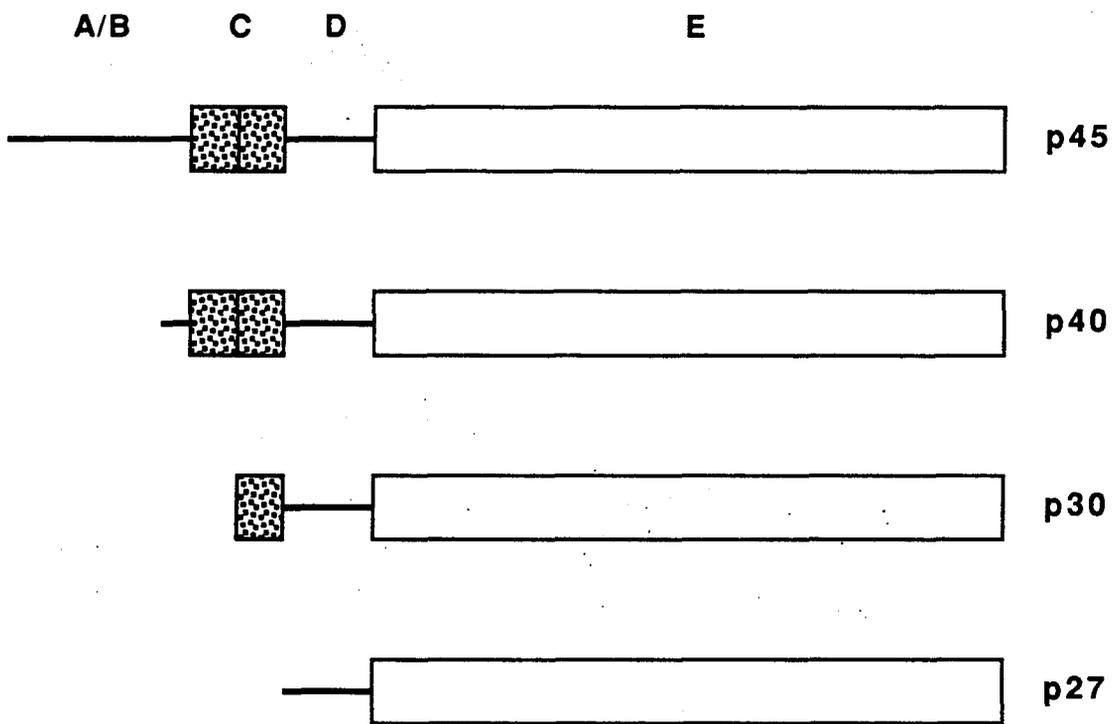
Puisqu'il ne semble pas y avoir de gène imbriqué avec c-erbA-2 on peut dire avec quelque assurance que l'imbrication de c-erbA-1 et de ear-1 est postérieure à la duplication de l'ancêtre des deux gènes c-erbA. D'après l'arbre phylogénique construit avec les séquences des différents membres de la sous-famille I des récepteurs nucléaires, cette duplication s'est effectuée bien après l'individualisation des ancêtres des gènes erbA de ceux des récepteurs de l'acide rétinoïque.

D'après ce que nous savons, il n'existe pas d'épissage alternatif 3' du gène c-erbA-1 de poulet comparable à celui qui existe chez les mammifères (Forrest et al., 1990b). Il semble raisonnable de lier l'existence de cet épissage alternatif et celle de l'imbrication des gènes c-erbA-1 et ear-1. Dans ce cas, l'imbrication n'existerait pas chez le poulet. Hélas, l'organisation génomique du locus c-erbA-1 chez cette espèce n'est connue que dans la partie 5' (Zahraoui et Cuny, 1987) et il serait très intéressant de savoir ce qui se passe en 3' de ce gène. On peut imaginer trois situations, plus ou moins probables :

- Bien que l'épissage 3' du locus c-erbA-1 ne semble pas exister chez le poulet, la structure imbriquée du locus existe. Ceci conduirait à dissocier les deux événements, alors qu'il nous paraît beaucoup plus simple d'envisager que c'est la translocation de ear-1 dans c-erbA-1 qui a créé l'exon 10 alternatif de c-erbA-1. Dans ce modèle, l'exon 10 alternatif de c-erbA-1 aurait d'abord appartenu à ear-1.

- L'imbrication n'existe pas, mais les deux gènes c-erbA-1 et ear-1 sont liés. C'est un simple événement de retournement de ear-1 (ou de c-erbA-1) qui a créé le locus imbriqué tel qu'on le connaît. Cette hypothèse laisse penser qu'on peut éventuellement trouver un gène homologue à ear-1 proche de c-erbA-2. Ces deux gènes proviennent d'un même ancêtre assez récent.

- Le gène ear-1 existe chez le poulet, mais à un endroit complètement différent, par exemple sur un autre chromosome. Dans ce cas, qui nous semble être l'hypothèse la plus crédible, le gène ear-1 s'est transloqué dans c-erbA-1, en créant du même coup l'exon 10, au début de l'émergence des mammifères. Le fait que ear-1 semble, chez ces organismes, jouer un rôle dans la thermogénèse sans frisson (un processus physiologique



**Figure 48 : Les différentes formes N-terminales tronquées du récepteur des hormones thyroïdiennes alpha chez le Poulet (d'après Bigler et Eisenman, 1988).**

spécifique des mammifères), peut être un argument très indirect en faveur de cette hypothèse.

La question reste posée pour l'instant, mais l'hypothèse selon laquelle le gène ear-1 se serait transloqué dans c-erbA-1 lors de l'émergence des mammifères, a notre préférence. C'est en effet la plus économique et elle rend bien compte de ce que l'on observe. Dans cette hypothèse, un évènement unique et assez simple, la translocation de ear-1 dans c-erbA-1, suffit à expliquer l'apparition de l'épissage alternatif 3' de c-erbA-1 et l'imbrication des deux gènes c-erbA-1 et ear-1.

### c) Epissage alternatif humain et formes N-terminales tronquées du poulet.

Le fait que l'épissage alternatif 3' n'existe pas chez le poulet nous oblige à un commentaire. En effet, cet épissage produit un variant du récepteur des hormones thyroïdiennes incapable de fixer ces hormones, et qui agit comme un régulateur négatif dominant des récepteurs normaux. On suspecte cette molécule d'être un régulateur endogène de l'action du récepteur des hormones thyroïdiennes, par exemple dans le cerveau où elle est exprimée très abondamment. Comment alors expliquer son absence chez le poulet ? Une hypothèse hardie mais intéressante, serait que chez cet animal, d'autres variants du récepteur des hormones thyroïdiennes jouent le même rôle régulateur, tout en étant structurellement différents.

Chez le poulet, Jeanette Bigler et Robert Eisenman, à Seattle, ont en effet observé l'existence de multiples protéines codées par les ARNs c-erbA-1 qui codent la forme classique du récepteur de 410 acides aminés (figure 48). Ces formes, plus courtes du côté N-terminal, résultent d'initiations alternatives de la traduction sur différents codons d'initiation en phase présents dans cet ARN (Bigler et Eisenman, 1988). On obtient ainsi, en plus du récepteur de 45 kDa, 3 formes mineures, respectivement de 40, 30 et 27 kDa, qui sont particulièrement bien exprimées *in vivo* dans les érythrocytes d'embryons de poulets. Ces évènements d'initiations alternatifs existent aussi dans le gène du récepteur de la progestérone (Conneely et al., 1989; Kastner et al., 1990b) et également dans d'autres gènes (voir Kozak, 1986 pour une revue).

Les formes les plus courtes de 30 et 27 kDa ne possèdent plus le domaine de fixation à l'ADN et ne sont donc plus capables de se lier à un TRE et d'activer la transcription. Mais, ajoutées en quantité importante à la protéine de 45 kDa, elles empêchent celle-ci de se fixer à l'ADN et d'activer la transcription. Donc, bien qu'elles soient parfaitement capables de fixer la T3, elles se comportent comme des régulateurs négatifs dominants de l'action du récepteur normal...exactement comme le produit de l'ARN c-erbA-1 $\alpha$ 2 chez les mammifères. Chez le poulet, il existe donc des molécules qui semblent remplir le rôle physiologique du produit de c-erbA-1 $\alpha$ 2 qui, lui, n'existe que chez les mammifères.

Si l'on retient l'hypothèse du lien entre l'épissage alternatif 3' de c-erbA-1 chez les mammifères avec la translocation du locus ear-1 (cette translocation créant du même coup l'épissage alternatif), on peut expliquer pourquoi la mise en place inopinée de cet épissage n'a pas

eu d'effet dévastateur. En effet, la fonction (régulation négative des récepteurs des hormones thyroïdiennes) était déjà présente. L'adjonction du produit de 490 acides aminés de c-erbA-1 $\alpha$ 2 n'a fait que rajouter une possibilité de régulation supplémentaire. Il faut noter que toutes les méthionines impliquées dans la formation des protéines tronquées en N-terminal sont conservées chez les amphibiens, les oiseaux et les mammifères. Il semble donc que cette fonction soit apparue très tôt par rapport à la date d'apparition présumée des récepteurs des hormones thyroïdiennes.

## **B) EVOLUTION DES RECEPTEURS NUCLEAIRES.**

Le modèle d'évolution des récepteurs nucléaires que nous avons proposé apporte trois séries d'informations particulièrement importantes. Premièrement, les récepteurs nucléaires étaient déjà très diversifiés au moment de la séparation des lignées qui donneront les arthropodes et les vertébrés. Il semble donc que les premières étapes de l'évolution de cette famille aient été très rapides. Deuxièmement, trois sous-familles de récepteurs se sont mises en place et l'une de ces sous-familles ne comporte, semble-t-il, que des récepteurs orphelins. Il faut donc se poser la question de la fonction de ces molécules. Sont-ils des récepteurs ? Sinon, que peut-on imaginer comme fonction pour ces molécules ? Enfin, si les duplications simples de gènes semblent expliquer la majeure partie de la diversification des récepteurs, dans certains cas, la situation semble plus complexe. Par exemple, les gènes des récepteurs de la vitamine D et de l'ecdysone sont visiblement des chimères évolutives. Dans ce cas, ils sont donc le résultat d'un événement (accidentel ?) qui a mis côte à côte le domaine C et le domaine E de deux récepteurs ancestraux appartenant à des sous familles différentes. Il convient donc de savoir si d'autres arguments renforcent cette hypothèse. De plus, il nous faut imaginer quelle peut être la nature de cet événement et quelles en ont été les conséquences. Ce sont ces différents points que nous souhaitons discuter maintenant.

### **1) Apparition des récepteurs nucléaires.**

#### **a) La formation du premier récepteur nucléaire.**

Il est très difficile de mettre en place un scénario d'apparition des récepteurs nucléaires, parce que l'on manque de faits expérimentaux sur ce que pourraient être les premiers récepteurs nucléaires. Notre analyse permet tout de même de confirmer ce que l'on suspectait déjà depuis longtemps : les récepteurs nucléaires sont des molécules très anciennes. Il y a 500 millions d'années, les récepteurs nucléaires avaient déjà une longue histoire derrière eux, puisqu'à cette date les trois sous familles de récepteurs nucléaires existaient déjà. Par comparaison, il est intéressant de noter que les premiers organismes multicellulaires sont apparus au début du Cambrien, il y a environ 570 millions d'années (Valentine, 1977; Hartland et al., 1982).

Nous pouvons donc affirmer que le "premier" récepteur nucléaire est apparu chez un organisme simple, très probablement unicellulaire. Un tel organisme était, à coup sûr, dépourvu d'organes spécialisés, tels que des glandes. Avec un tel raisonnement, Bert O'Malley a imaginé que les récepteurs nucléaires primitifs ne devaient pas avoir la fonction qu'on leur connaît actuellement, mais devaient plutôt être des molécules chargées de fixer des ligands intra-cellulaires. Ces ligands, dont la nature chimique pouvait être différente des ligands actuels, étaient peut être des nutriments amenés de l'extérieur, ou des substrats métaboliques (O'Malley, 1989). Dans un tel modèle, les récepteurs nucléaires pourraient "avertir" la cellule d'une baisse de la concentration d'un métabolite essentiel (un acide aminé par exemple) et ainsi mettre en route un programme de dormance ou de sporulation.

David Moore a récemment insisté sur le fait que tous les ligands de récepteurs nucléaires actuellement connus dépendent de la présence d'oxygène au cours de leur voie biosynthétique (Moore, 1990). Il estime donc probable que les récepteurs nucléaires et leurs ligands ont évolué à partir de l'émergence de la photosynthèse. Il imagine donc un modèle selon lequel, dans un eucaryote aérobic primitif, il y a eu fusion entre les exons codant des doigts de zinc et d'autres exons codant une protéine fixant des stéroïdes. De telles protéines, fixant les oestrogènes, existent chez les levures (Feldman et al., 1984; Skowroski et Feldman, 1989; Burshell et al., 1984). Les domaines des récepteurs nucléaires sont très indépendants les uns des autres. On peut, par exemple greffer un domaine de fixation des glucocorticoïdes sur un domaine d'une protéine complètement différente (comme *c-myc* ou E1A) et obtenir une molécule chimère dont la capacité transactivatrice est dépendante des glucocorticoïdes (Eilers et al., 1989; Becker et al., 1989; Hope et al., 1990; Superti-Furga et al., 1991). On peut également faire ce genre de chimères entre un domaine E de récepteur nucléaire et le domaine catalytique d'une protéine kinase et obtenir une kinase dont l'activité dépend d'un ligand. Cette grande indépendance des domaines permet de suggérer qu'avant une éventuelle fusion avec un domaine de fixation à l'ADN, les domaines E des récepteurs avaient une autre fonction, plus "simple", par exemple la régulation d'une activité catalytique.

#### b) Ancienneté fonctionnelle des récepteurs nucléaires.

Ce n'est pas parce que les récepteurs nucléaires existent depuis longtemps que leur fonction actuelle (la réception et la transduction d'un signal hormonal *via* le contrôle direct de la transcription) est conservée depuis aussi longtemps. Les récepteurs nucléaires de drosophile auraient pu avoir un rôle différent. Ceci a été décrit par exemple pour deux gènes structurellement proches : le facteur de croissance *int-1* et son homologue de drosophile, le gène *Wingless* (Bender et Peifer, 1987). Mais la découverte récente du récepteur de l'ecdysone a réglé cette question (Koelle et al., 1991). Il existe bien des récepteurs nucléaires tout à fait classiques chez la drosophile. Leurs éléments de réponse sur l'ADN ont la même structure que leurs homologues de vertébrés (Martinez et al., 1991). Donc la fonction des récepteurs nucléaires a été conservée,

comme leur structure, depuis au moins 500 millions d'années.

c) Des récepteurs nucléaires chez d'autres organismes ?

On peut apporter des éléments de réponse à la question de l'ancienneté et de l'origine des récepteurs nucléaires, en recherchant l'existence de tels récepteurs chez des organismes "primitifs" ou ayant divergé très tôt des ancêtres des animaux multicellulaires. La connaissance de tels récepteurs "primitifs" pourrait apporter des éléments d'information intéressants pour comprendre l'émergence d'une telle famille de gènes. Nous allons donc, à partir des données de la littérature tenter de retrouver la trace de ces molécules chez d'autres organismes.

Chez les bactéries, la situation est simple : nous n'avons trouvés aucune trace de récepteur nucléaire ni de protéine à doigt de zinc se fixant à l'ADN. De même, il ne semble pas exister de protéines capables de fixer les stéroïdes et dans la littérature nous n'avons pas trouvé d'effet physiologique pour une telle molécule.

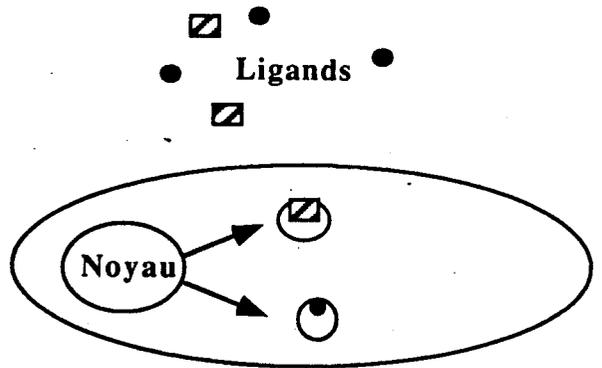
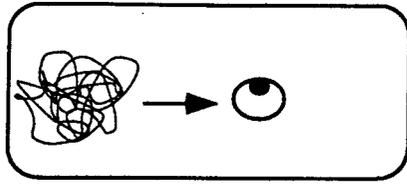
Chez les levures, la situation est un peu plus complexe. En effet, il existe des protéines cytoplasmiques capables de fixer les oestrogènes, par exemple chez *Saccharomyces cerevisiae* (Feldmann et al., 1984) ou chez *Candida albicans* (Skowroski et Feldman, 1989). Il est clair que les oestrogènes ont un effet physiologique chez les levures, par exemple en régulant le cycle cellulaire (Tanaka et al., 1989). Mais pour le moment, il n'y a pas de trace claire de récepteur nucléaire.

Par contre, chez la moisissure *Achlya ambisexualis*, il existe des stéroïdes, (l'anthéridiol, par exemple) qui sont capables de réguler le développement sexuel, (Mac Morris et Barksdale, 1967), et qui sont fortement soupçonnés d'avoir des systèmes de reconnaissance ressemblant à ceux des stéroïdes classiques (Brunt et al., 1990). On a pu mettre en évidence que l'un des composants du complexe récepteur de l'anthéridiol, chez ce micro-organisme aquatique, est une protéine de choc thermique (hsp85) proche des hsp83 de drosophile et des hsp90 des vertébrés. De plus, le complexe récepteur de l'anthéridiol, d'un poids moléculaire de 192 kDa, a un coefficient de sédimentation de 8S et peut être converti en complexe 4S à la suite d'un traitement par des sels (Riehl et al., 1984). Evidemment, il est tentant de faire un parallèle avec les hsp90 des vertébrés qui, elles aussi, se complexent aux récepteurs des stéroïdes et dont les coefficients de sédimentation changent également en fonction de la salinité ...

Chez les protistes comme l'amibe *Dictyostelium discoïdium*, il existe des petites molécules hydrophobes capables de diffuser "librement" à travers la membrane plasmique. C'est le cas, par exemple, du DIF (Differentiation Inducing Factor = 1-(3,5-dichloro-2,6-dihydroxy-4-methoxyphenyl)-1-hexanone), un facteur lipidique essentiel pour certains processus d'agrégation chez *Dictyostelium*. On le décrit comme un morphogène puisque son action sur l'expression de certains gènes semble être régulée en fonction de la position dans l'espace (gradient antérieur-postérieur) (Firtel et al., 1991 ; Berks et al., 1991). Insall et Kay, 1990 ont recherché le récepteur du DIF et ont montré qu'il s'agit d'une molécule surtout cytoplasmique, ayant un Kd de 1

## PREMIER SCENARIO

Apparition des récepteurs nucléaires chez une bactérie ou chez un protiste primitif.

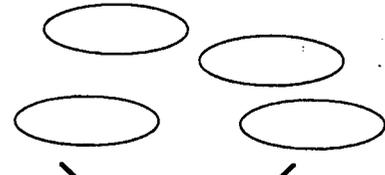


Diversification lente des récepteurs nucléaires

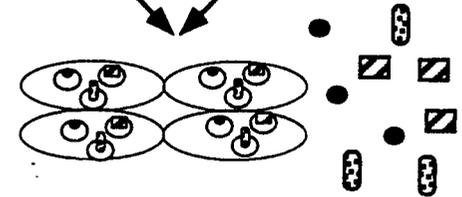
Rôle "d'avertisseur physiologique" : présence d'un nutriment ou d'une substance toxique



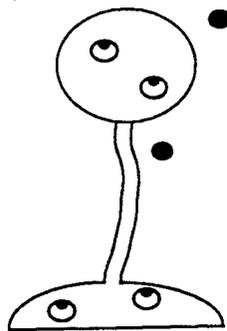
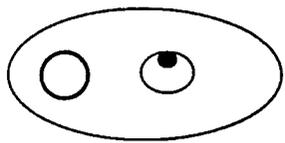
Formation des métazoaires



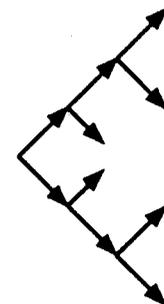
Diversification supplémentaire des récepteurs



## DEUXIEME SCENARIO



Formation des métazoaires



Diversification "explosive" des métazoaires...et des récepteurs nucléaires

Stade multicellulaire (ex : sporulation) : rôle des récepteurs dans la régulation de ces étapes

Figure 49 : Deux scénarios possibles d'apparition des récepteurs nucléaires. Ces deux scénarios ne diffèrent que dans la vitesse des phases précoces de l'évolution de ces gènes : lente dans le premier cas, rapide, voire même "explosive", dans le second cas.

à 2 nM : une valeur très comparable à celle des récepteurs des stéroïdes pour leurs ligands respectifs par exemple. Les caractéristiques biochimiques de cette molécule sont plus proches de celle d'un récepteur intracellulaire du type des récepteurs nucléaires, que d'une molécule cytoplasmique ayant une affinité pour un ligand donné (comme les protéines cellulaires fixant l'acide rétinoïque, les CRABP) (Insall et Kay, 1990).

Evidemment, chez les végétaux, il existe des systèmes de reconnaissance hormonale très élaborés. Les hormones végétales, telles que les auxines, les cytokinines, les gibbérellines ou l'acide abscissique, sont très souvent de petites molécules hydrophobes parfois dérivées des terpénoïdes (comme les stéroïdes) dont certains ont peut-être des récepteurs nucléaires (Moore, 1990). Dans le cas de l'auxine par exemple, on ne sait toujours pas très bien si le, ou les récepteurs, sont membranaires ou cytoplasmiques (Napier et Venis, 1991; Cross, 1991). On sait que les gibbérellines induisent des changements dans l'expression de certains gènes, et dans certains cas des plantes mutantes qui résistent à l'action de ces hormones ont été isolées (Moore 1990). Il n'est donc pas impossible d'imaginer que certaines de ces molécules ait des récepteurs appartenant à la superfamille des récepteurs nucléaires.

L'existence de récepteurs nucléaires chez tous les animaux multicellulaires à trois feuilletts embryonnaires (les métazoaires triploblastiques) peut être supposée sans grand risque d'erreur. Outre les arthropodes et les vertébrés on sait qu'il en existe chez le nématode *Caenorhabditis elegans* (cité dans Moore, 1990). L'ecdysone, hormone de la mue chez les insectes, a été découverte chez le parasite plathelminthe *Schistosoma mansonii* (Nirdé et al., 1983) sans que l'on sache encore si elle y a un rôle physiologique (voir Hoffman et Joose, 1991). Elle a pourtant sûrement une fonction qui reste à découvrir...

#### d) Deux modèles d'évolution précoce.

Selon que les récepteurs nucléaires existent ou non en dehors des métazoaires triploblastiques, on peut bâtir deux modèles décrivant l'évolution précoce de la superfamille. Ces deux modèles diffèrent surtout par les vitesses d'évolution qu'ils impliquent (figure 49).

Dans le premier modèle, les récepteurs apparaissent très tôt chez un protiste primitif ou même chez une bactérie. Dans ce cas, ils se diversifient assez lentement au cours du Précambrien en jouant le rôle "d'avertisseurs physiologiques" simples qui permettent à un organisme de savoir ce qui se passe dans son environnement et d'adapter son comportement en conséquence. L'avantage de posséder un récepteur nucléaire pour un unicellulaire est que ces molécules ne requièrent pas toute une voie de transduction élaborée pour agir. Dans ce modèle, les récepteurs nucléaires sont les systèmes d'écoute du pauvre !

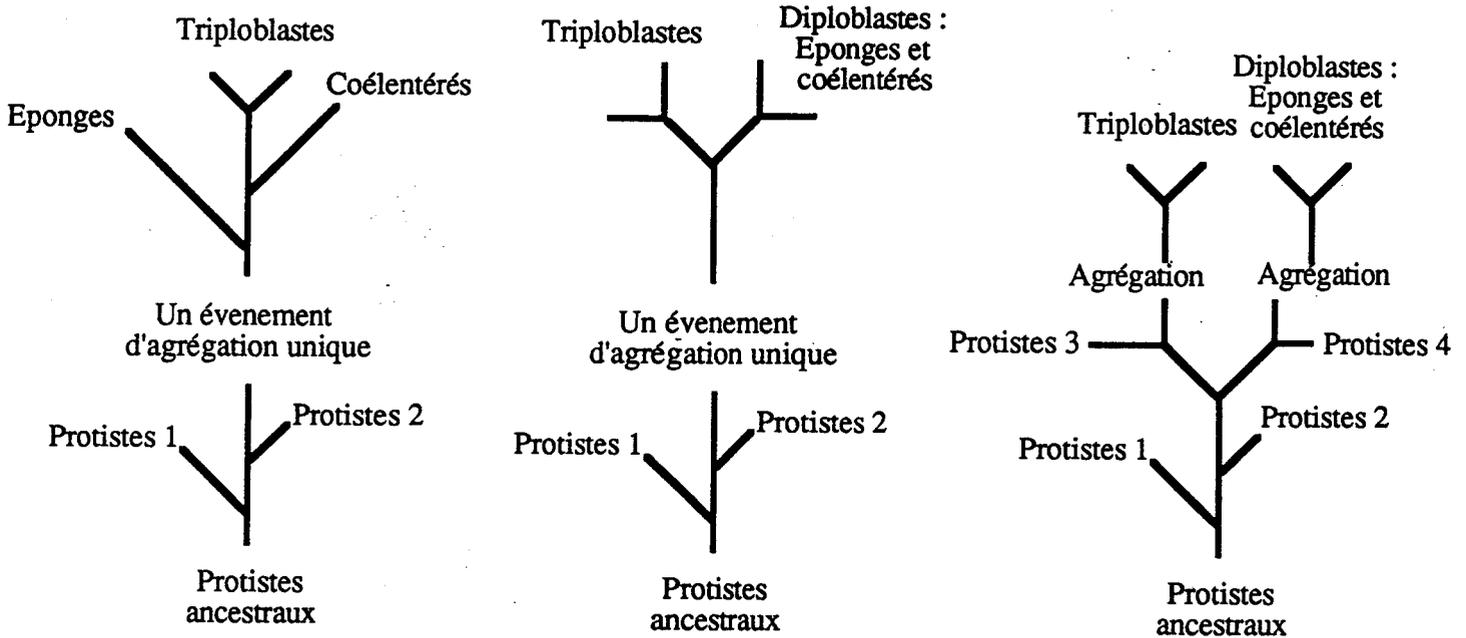
Le second modèle implique que les récepteurs nucléaires soient strictement restreints aux multicellulaires (figure 49). Dans ce cas, le récepteur apparaît chez un protiste et lui est bien utile, par exemple pour réguler les stades multicellulaires (sporulation etc...) de son cycle de vie. Ici les récepteurs nucléaires sont des systèmes de communication et d'écoute des

**Différentes hypothèses pour l'origine des métazoaires**

A : La théorie classique : une origine commune de tous les métazoaires

B : Une variante : origine commune des métazoaires et divergence précoce des deux grands groupes

C : Deux événements différents de multicellularisation



La phylogénie dressée à partir de l'étude du gène de l'ARN 28S : plusieurs événements d'agrégation distincts (hypothèse C confirmée)

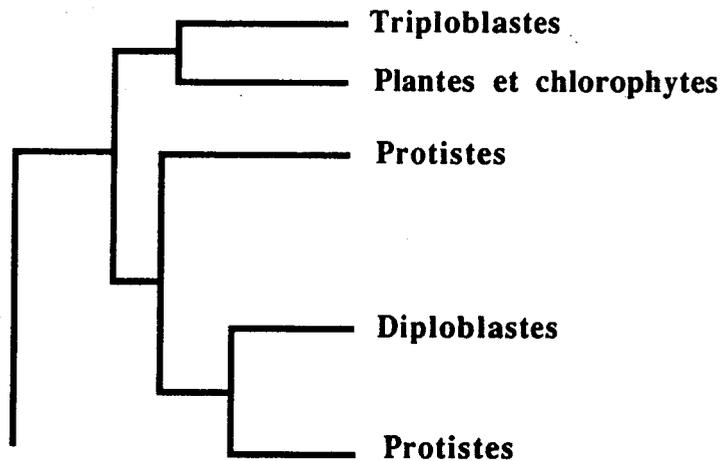


Figure 50 : L'origine des métazoaires est toujours l'objet de controverses. Parmi trois scénarios possibles, l'étude récente de Christen et al., 1991 privilégie le troisième. Mais la discussion n'est pas close...

congénères, là encore assez économiques. Ils favorisent de plus en plus le développement de la vie multicellulaire. Puis, lors des événements d'agréations qui vont donner les premiers multicellulaires, les récepteurs nucléaires se diversifient très vite, permettant ainsi aux relations intercellulaires d'acquérir de plus en plus de finesse. C'est alors la phase d'évolution très rapide et de diversification extrême que nous avons observée aux étapes précoces de nos arbres phylogéniques.

D'après les travaux de nombreux paléontologues, les phases d'évolution précoce des métazoaires se sont produites d'une manière explosive (Valentine, 1991; Erwin, 1991). En un laps de temps très court, la quasi totalité des plans d'organisation des métazoaires se sont mis en place (Gould, 1991). Il est tentant d'imaginer que, dans la mesure où les récepteurs nucléaires sont impliqués dans les communications entre cellules (un besoin crucial pour les métazoaires), la divergence rapide que nous observons coïncide avec ces étapes précoces de l'évolution des métazoaires.

Ces deux modèles ne sont pas mutuellement exclusifs. Les récepteurs nucléaires peuvent être apparus très tôt sans s'être multipliés, puis, au début de l'évolution des métazoaires, se diversifier très vite. Il faut noter que, d'après des travaux récents de phylogénie moléculaire réalisés par le groupe d'André Adoutte à Orsay, les animaux métazoaires triploblastiques et les plantes proviennent d'un ancêtre commun (figure 50 ; Christen et al., 1991; Barnes, 1991). On peut donc imaginer que, si l'on trouve des récepteurs nucléaires chez les plantes, le modèle de la diversification explosive puisse tout de même fonctionner. Dans ce cas, l'ancêtre commun des plantes et des triploblastes aurait eu des récepteurs nucléaires et, au moment des phénomènes d'agréation, ces récepteurs se seraient alors diversifiés. Il n'est pas impossible que dans ce cas l'évolution des récepteurs nucléaires chez les plantes et les triploblastes ait produit des récepteurs très différents dans ces deux règnes.

#### e) Les événements plus tardifs.

Nous avons tenté d'étudier ce qui s'est produit après cette première série d'événements, pour une sous-famille particulière, la sous-famille I, qui groupe les récepteurs des hormones thyroïdiennes, ceux de l'acide rétinoïque et les gènes du groupe EAR1 (EAR1, E75, PPAR). Cette sous-famille constitue un modèle d'étude intéressant parce que l'on connaît de nombreuses versions de chaque gène (mammifères, oiseaux et souvent amphibiens), ce qui permet de dater certains événements. De plus, certains gènes de cette sous-famille sont probablement des "chimères évolutives". Mieux comprendre les différentes étapes de cette histoire permettra sans doute d'éclaircir la formation de ces chimères.

Cette sous-famille se sépare en trois groupes de gènes, chacun contenant plusieurs gènes étroitement apparentés. Ceci est depuis longtemps bien connu pour les récepteurs des hormones thyroïdiennes et pour les récepteurs de l'acide rétinoïque, mais c'est tout à fait nouveau pour le groupe des gènes EAR1. Jusqu'à présent, ce groupe ne contenait que les gènes EAR1, E75

et PPAR. Mais tout récemment, trois nouveaux gènes PPAR viennent d'être clonés chez le xénope, par les groupes de Walter Wahli et Gerd Helftenbein (communication personnelle de Walter Wahli; Dreyer et al., 1992). L'un de ces gènes (XPPARA) est l'homologue de xénope du gène PPAR de souris, précédemment décrit par Stephen Green. Les deux autres (XPPARB et XPPARG) sont par contre des nouveaux gènes PPAR. L'arbre construit avec les membres de la sous-famille I par la méthode de Fitch et Margoliash, montre clairement que, comme le suggérait déjà l'observation scrupuleuse des séquences, ces deux nouveaux gènes ne résultent pas d'un événement spécifique du xénope, mais ont divergé plus tôt. Il doit donc logiquement exister des gènes PPAR  $\beta$  et  $\gamma$  chez les mammifères.

En étudiant l'arbre obtenu pour cette sous-famille (figure 45), on peut noter que les trois grands groupes de gènes de la sous-famille I semblent se diversifier en même temps. Cependant, la longueur des branches d'un arbre Fitch n'est pas proportionnelle au temps, mais à la distance (en terme de pourcentage d'homologie) entre les gènes. Ceci implique que si les gènes ont divergé à des vitesses différentes (ce qui est probablement le cas) le fait que l'on observe que ces gènes divergent "en même temps" sur l'arbre ne signifie pas une dichotomie contemporaine. Cependant, l'utilisation d'autres méthodes (celle de Kitch, par exemple), qui supposent une vitesse identique d'évolution des gènes étudiés, permettent de penser qu'effectivement, les trois groupes de la sous-famille I se sont diversifiés à peu près en même temps. Ceci n'est pas très étonnant pour les récepteurs des hormones thyroïdiennes et de l'acide rétinoïque, parce que l'on suspecte que ces gènes sont sur les mêmes zones d'ADN et ont évolué ensemble, suite à des duplications de grandes régions d'ADN. Si cette analyse, encore préliminaire, se confirme aussi pour les gènes PPAR, cela signifierait que nous pourrions dater cette période de diversification des groupes. Il est tentant d'imaginer que cette date correspond au moment de l'apparition des vertébrés mais il est encore trop tôt pour en être sûr.

## 2) Les trois sous familles de récepteurs.

### a) Relation structure/évolution.

Nous avons classé la famille des récepteurs nucléaires en trois sous familles (figure 43). Cette classification peut-elle être confirmée par l'analyse de l'organisation génomique des récepteurs ?

Pour les domaines A/B et C, nous avons regroupé les récepteurs nucléaires en trois types structuraux (voir résultats, article I). Le premier type contient des récepteurs à domaine A/B variable, codé par de nombreux exons et dont l'intron situé entre les exons codant les doigts de zinc C1 et C2 est situé tout près de l'exon codant C1. Dans ce groupe, on trouve les récepteurs des hormones thyroïdiennes, ceux de l'acide rétinoïque et les trois orphelins ear-1, E75 et PPAR. Au premier type structural correspond donc bien la première sous-famille de récepteurs.

Le deuxième type structural est celui du gène NGF1B, avec un seul exon codant

A/B et C1 et un intron situé entre les exons codant C1 et C2, en position médiane. Le gène NGF1B est un membre de la sous-famille II qui comprend les récepteurs orphelins. Il est difficile de savoir si le type structural NGF1B est valable pour le gène NGF1B seul, ou si on le trouve pour d'autres récepteurs orphelins. Les seules organisations génomiques de membres de la sous-famille II qui ont été déterminées, à notre connaissance, viennent de gènes de drosophile, comme *ultraspiracle*, qui est dépourvu d'intron (Oro et al., 1990).

Enfin, le troisième type structural est celui des gènes à domaine A/B codé par un seul exon, et où l'intron situé entre les exons codant C1 et C2 est du côté de C2. A ce type, correspondent les gènes de la sous-famille III (récepteurs des stéroïdes, de la vitamine D et de l'ecdysone, gènes du groupe *knirps*). Ici, cette classification semble un peu trop simple. Par exemple, le récepteur de l'ecdysone n'a pas d'intron entre C1 et C2. Ce n'est pas très étonnant, car cet intron a pu être perdu au cours de l'évolution de la drosophile. Par contre, il est plus ennuyeux de remarquer que le seul intron présent dans les gènes du groupe *knirps* soit situé entre C1 et C2, à une position diagnostique de...la sous-famille I! On peut expliquer cette observation de deux façons. La place des gènes du groupe *knirps* dans l'arbre évolutif du domaine C est incorrecte et résulte d'un phénomène de convergence, sans signification évolutive réelle. Il faut préciser à ce sujet, que les gènes *knirps* n'ont pas de domaine Ti-DM et que leur appartenance à la classe III repose uniquement sur l'arbre du domaine C, plus fragile que l'arbre du domaine Ti-DM. L'autre solution est que la position de l'intron entre C1 et C2 ne soit pas significative d'un point de vue évolutif. On peut par exemple imaginer que cet exon ait été acquis par l'ancêtre des gènes *knirps*, tout à fait indépendamment des gènes de la sous-famille I. On sait en effet que au cours de l'évolution des introns peuvent être acquis par les gènes. Très souvent cette acquisition semble se faire non pas au hasard, mais à des sites favorisés appelés sites de proto-épissage. On peut ainsi expliquer que deux introns soient placés indépendamment au même endroit, dans deux gènes appartenant à la même famille (Dibb et Newman, 1989; Cavalier-Smith, 1991). Il faut enfin ajouter que l'histoire évolutive des gènes du groupe *knirps* semble très inhabituelle, et que l'on doit donc rester très prudent dans des conclusions basées sur la structure de ces gènes.

En conclusion, sauf exception, il semble que notre classification des récepteurs nucléaires en trois sous familles soit relativement bien confirmée par l'analyse de la position des introns situés en 5' de ces gènes. Cependant, il faut attendre que l'on connaisse la structure d'autres gènes (comme ceux des récepteurs orphelins) pour être plus affirmatif. De plus, l'analyse, (actuellement en cours) des introns situés en 3', dans le domaine Ti-DM devrait également apporter des informations intéressantes. Peut-être pourra t-on étudier comment les introns des récepteurs nucléaires ont évolué comme cela est actuellement discuté pour de nombreuses autres familles de gènes (Dibb et Newman, 1989 ; Cavalier-Smith, 1991 ; Doris et Gilbert, 1991). Y a t-il eu des pertes ou des gains d'introns au cours de l'évolution de ces gènes ?

b) Les "récepteurs orphelins" sont-ils orphelins ?

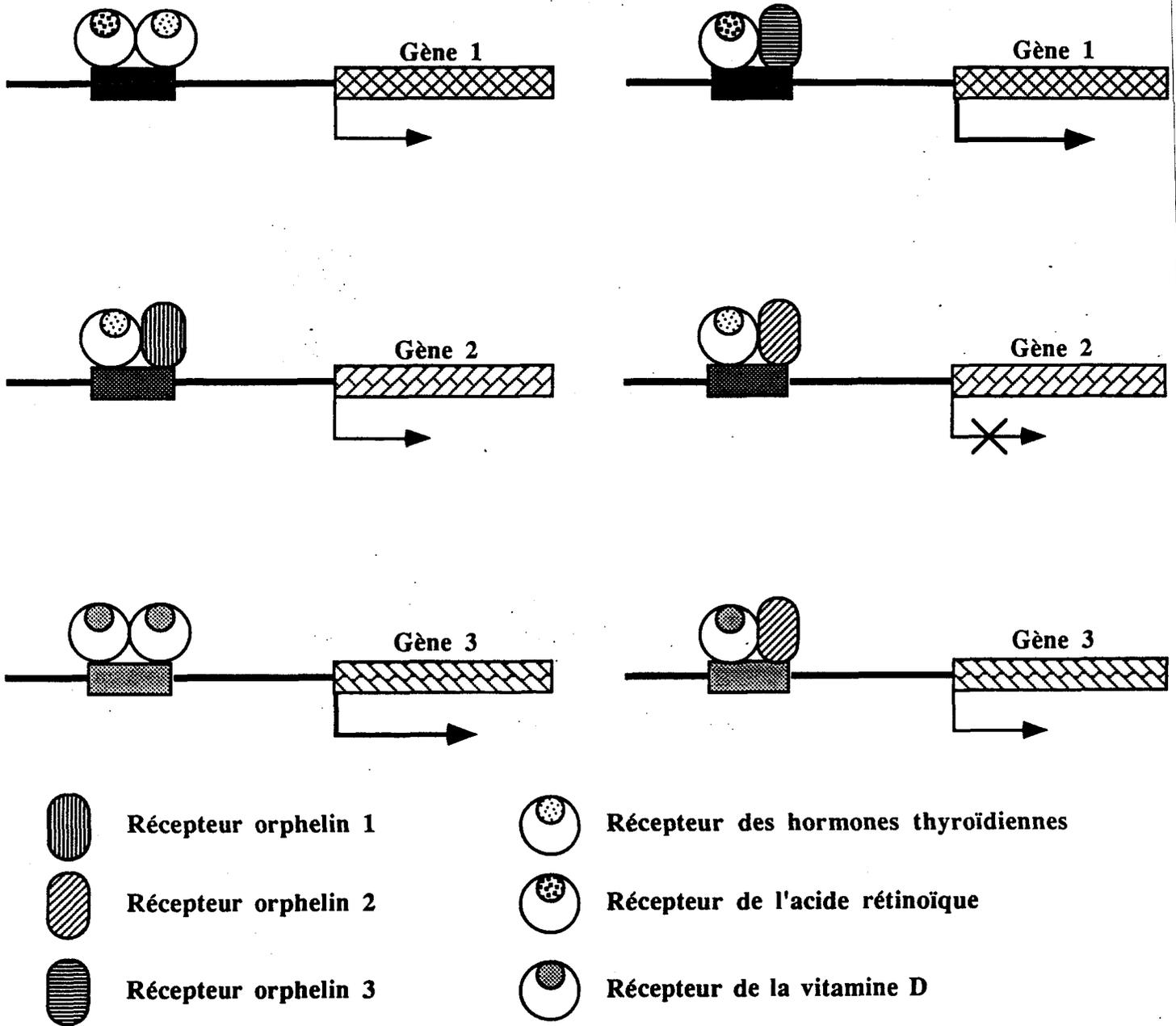
La classification de nombreux récepteurs orphelins dans la sous-famille II des récepteurs nucléaires demande quelques commentaires. En effet, pour les deux autres sous-familles, il semble exister une identité de fonction pour les gènes appartenant à un même groupe. Quelle peut donc être la fonction commune des récepteurs de la sous-famille II ? Deux fonctions possibles sont envisageables: ce sont des récepteurs à des ligands inconnus, ou ce ne sont pas des récepteurs.

La sous-famille II peut être découpée en trois groupes: le groupe des gènes COUP, le tandem HNF4/*tailless* et enfin le groupe des gènes RXR. Pour les deux premiers cas il n'y a aucune information quant à la nature d'un ligand quelconque reconnu par les récepteurs codés par ces gènes. Dans le cas de COUP, on a bien décrit une activation du produit de ce gène par la dopamine, mais c'est un effet indirect au cours duquel COUP est phosphorylé de manière dépendante de la dopamine (Power et al., 1991a; 1991b). Il est clair que la dopamine ne se fixe pas à COUP. Par contre, il semble bien que les membres du groupe RXR reconnaissent l'isoforme *cis* de l'acide rétinoïque (Levin et al., 1992). Voici des récepteurs orphelins qui ne le sont plus !

L'histoire des gènes RXR mérite d'être examinée plus en détail. Le premier membre de ce groupe, RXR $\alpha$ , a été décrit comme une molécule activable par l'acide rétinoïque de type *trans* à des doses très élevées (Mangelsdorf et al., 1990). Pourtant, RXR $\alpha$  n'est pas proche des RAR du point de vue des homologues de séquence (seulement 61% et 28% avec les domaines C et E de RAR $\alpha$ ). Très curieusement, les capacités transactivatrices de RXR $\alpha$  sont activées en présence d'acide rétinoïque *trans* de manière non saturable. De plus, aucune fixation de rétinoïdes sur RXR $\alpha$  n'avait pu être obtenue.

On a tout d'abord pensé qu'il était un récepteur nucléaire à un autre rétinoïde comme l'acide 3,4-didéhydrorétinoïque. Cependant, l'acide 3,4-didéhydrorétinoïque n'est pas capable d'activer RXR $\alpha$  (Eager et al., 1991). Si RXR est un récepteur d'un rétinoïde il faut croire que des formes différentes, mais proches, de rétinoïdes ont des effets complètement différents sur le récepteur, ce qui ne se retrouve pas pour les RAR classiques. De tels effets paraissent mieux s'expliquer par une action indirecte des rétinoïdes via leur récepteur, comme pour le modèle de COUP et de la dopamine. Il existe un homologue de RXR $\alpha$  chez la Drosophile, le gène *ultraspiracle* ou USP, qui n'est pas activé par les rétinoïdes (Oro et al., 1990). Ce gène, cloné en même temps par plusieurs équipes, a 49% d'homologie avec RXR dans le domaine E (Henrich et al., 1990; Shea et al., 1990). Ces données ne simplifiaient pas la situation car on ne voyait pas très bien comment la capacité de fixer les rétinoïdes serait apparue brusquement et spécifiquement pour le gène RXR $\alpha$ , et pas pour son homologue de Drosophile. Une activation indirecte, par phosphorylation par exemple, pouvait par exemple être envisagée.

La solution a été obtenue tout récemment, puisqu'un gène étroitement apparenté à RXR $\alpha$  et appelé RXR $\beta$  (ou H2RIIBP; Hamada et al., 1989; Rowe et al., 1991)) fixe la forme *cis*



**Figure 51 : Les récepteurs orphelins sont des régulateurs de l'action des récepteurs de classe 1 (THR, RAR, EAR, VDR etc...). Selon un jeu complexe de combinatoires, des gènes seront régulés différemment par des complexes de récepteurs différents.**

de l'acide rétinoïque (Levin et al., 1992 ; Heyman et al., 1992). En effet, il semble que l'acide rétinoïque sous forme *trans* puisse être transformé en forme *cis* dans la cellule, sans doute par des isomérases encore inconnues. Ceci permet d'expliquer que l'activation par l'acide rétinoïque *trans* de RXR $\alpha$  ne soit pas saturable. En effet, seule une faible proportion de cette molécule doit pouvoir être isomérisée et on ne doit pas obtenir assez de forme *cis* pour saturer le récepteur. Les produits des gènes RXR ont donc bien un ligand. L'avenir dira si les autres membres de ce groupe sont ou non orphelins...

### c) Rôle des récepteurs orphelins.

A quoi pourraient bien servir des récepteurs orphelins ? Une hypothèse raisonnable est d'imaginer qu'ils modulent l'action des autres récepteurs (et cela qu'ils aient, ou non, un ligand). En effet, ces molécules ont-elles des sites de fixation à l'ADN proches, sinon identiques, à ceux de certains récepteurs comme les membres de la sous-famille I (voir le tableau 2 basé sur l'analyse des boîtes P). Par exemple, le facteur de transcription ARP-1 est capable de se fixer à un TRE (Ladias et Karathanasis, 1991).

Le cas du gène NGF1B est particulièrement intéressant. Le produit de ce gène est en effet capable de se fixer à un TRE (Wilson et al., 1991). Or, le gène NGF1B est inductible par les facteurs de croissance, dans les cellules nerveuses (Milbrandt, 1988; Hazel et al., 1988; Ryseck et al., 1989; Watson et Milbrandt, 1990; Wisden et al., 1990). De plus, le produit de ce gène est très rapidement modifié après sa traduction, toujours en réponse à des stimuli par des facteurs de croissance ou par la dépolarisation de la membrane dans les cellules PC12 (Fahrner et al., 1990; Hazel et al., 1991). Enfin, cette protéine peut activer la transcription, même en absence de son domaine de fixation du ligand, ce qui peut faire penser qu'il n'a pas de ligand (il y a bien sûr d'autres interprétations possibles ; Wilson et al., 1991). Voici donc un gène dont l'action pourrait être de moduler très rapidement l'action des récepteurs de la sous-famille I, en réponse à des stimuli externe. Ceci permettrait de rendre compte de l'extrême plasticité de la réponse à certaines hormones, comme les hormones thyroïdiennes.

En se fixant sur les mêmes sites sur l'ADN, les récepteurs orphelins pourraient éventuellement, soit entrer en compétition avec eux, soit au contraire, les aider à se fixer (figure 51). Par exemple, le facteur TRAP (Beebe et al., 1991; O'Donnel et al., 1991; Darling et al., 1991) qui, présent dans des extraits nucléaires, augmente la fixation des récepteurs des hormones thyroïdiennes au TRE, pourrait être un récepteur orphelin comme cela a été récemment suggéré (Moore et Brent, 1991). En effet TRAP est capable de dimériser avec les récepteurs des hormones thyroïdiennes et les sites importants du récepteur pour cette dimérisation sont le domaine Ti-DM et le doigt C2. Or ces deux sites sont justement ceux impliqués dans la dimérisation de ces récepteurs entre eux et avec les RAR (Green et Chambon, 1988; Forman et Samuels, 1990a; 1990b). Il faut attendre le clonage de TRAP pour confirmer ce modèle, mais il permettrait d'imaginer un niveau de régulation supplémentaire de l'action des hormones thyroïdiennes et de l'acide rétinoïque.

### Groupe knirps



### Récepteur de la vitamine D, récepteur de l'ecdysone



### Récepteur orphelin NGF1B

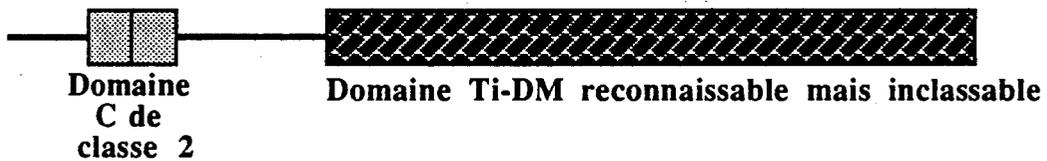


Figure 52 : Trois exemples de récepteurs nucléaires que l'on considère comme des "chimères évolutives".

L'équipe de Michael Rosenfeld a cherché des protéines ayant, pour les récepteurs de l'acide rétinoïque, une activité similaire à celle de TRAP pour les récepteurs des hormones thyroïdiennes. Ces chercheurs sont ainsi tombés sur...le récepteur orphelin RXR $\beta$  qui est donc capable d'augmenter l'affinité des RAR pour l'ADN (Yu et al., 1991). RXR $\beta$  est également capable de former des dimères avec les RAR, les THR et le récepteur de la vitamine D. Cette découverte importante a été confirmée par les trois autres équipes qui ont en même temps cloné un troisième type de RXR, RXR $\gamma$  (Leid et al., 1992 ; Zhang et al., 1992 ; Kliewer et al., 1992) . Tout ceci suggère que certains des récepteurs orphelins sont des composants de la (ou des) protéine(s) TRAP, ce qui confirme bien le modèle énoncé ci-dessus. On peut imaginer que l'acide rétinoïque sous forme *cis* régule ainsi, via son récepteur RXR $\beta$ , la réponse à l'acide rétinoïque sous forme *trans* (et vice-versa).

### 3) Les récepteurs chimères.

#### a) Les exemples de récepteurs chimères.

La comparaison des arbres construits d'après les séquences des domaines C et Ti-DM nous a conduit à envisager l'existence de gènes chimères, produits de différentes histoires évolutives. Nous pouvons envisager deux cas: celui des gènes *knirps*, qui ont un domaine E complètement différent et non homologue à ceux des autres récepteurs, et celui des gènes comme le récepteur de la vitamine D (VDR), celui de l'ecdysone (ECR) ou encore les récepteurs orphelins NGF1B, FTZ-F1, HNF4 et TLL (figure 52). Dans ce dernier cas, on a soit des gènes dont les deux domaines structuraux étudiés appartiennent à deux sous familles différentes (VDR, ECR, FTZ-F1), soit des gènes comme NGF1B, qui ont un domaine Ti-DM inclassable, soit enfin des gènes qui, au sein de la même sous-famille, changent de groupe (HNF4 et TLL).

Les trois gènes du groupe *knirps* représentent des exemples assez simples de gènes chimères qui peuvent être simplement le résultat d'un événement de translocation chromosomique qui aurait séparé le domaine C d'un récepteur ancestral, de son domaine Ti-DM. Le domaine C aurait ainsi été associé à un nouveau domaine structural pour donner une molécule fonctionnelle. La question est bien sûr de savoir quelle est la fonction de ces gènes (voir la revue de Pankratz et Jackle, 1990).

Avant de discuter plus précisément les trois cas les plus spectaculaires de gènes chimères (VDR, ECR et FTZ-F1 ; voir figure 52), il convient de dire un petit mot des trois autres exemples (NGF1B, HNF4 et TLL). Il faut en effet être prudent dans l'interprétation des différences que l'on obtient sur les deux types d'arbres. Ces différences (surtout à petite échelle), peuvent être dues à des vitesses d'évolution différentes des domaines, ou à des phénomènes de convergence évolutive qui ne permettent pas de tirer des conclusions sur l'origine du domaine considéré. Ainsi, les gènes HNF4 et TLL se placent différemment sur les arbres construits à partir du domaine C ou du domaine Ti-DM. Mais selon la méthode de construction d'arbre utilisée (la

méthode de Fitch ou celle appelée "Neighbor Joining") ces deux gènes ne se placent pas de la même façon dans les arbres construits à partir des séquences du domaine C. L'emplacement précis de HNF4 et TLL reste donc à définir.

Pour NGF1B, nous avons également des réticences à le considérer comme une vraie chimère évolutive. Si l'on considère le domaine C, ce gène fait partie du groupe de RXR, dans la sous-famille II des récepteurs orphelins. Par contre, pour le domaine Ti-DM, il est impossible de savoir s'il fait partie de la sous-famille II ou de la sous-famille III. Pour ce domaine, NGF1B semble avoir divergé très tôt de l'ancêtre commun des sous familles II et III. Mais ce résultat étonnant peut s'expliquer par une vitesse d'évolution très rapide de ce domaine. Tant que nous n'aurons pas d'informations physiologiques supplémentaires, il convient d'être prudent sur l'emplacement précis de NGF1B pour le domaine Ti-DM.

Il faut insister sur le fait que notre comparaison de l'évolution des deux domaines repose sur des arbres construits à partir de séquences très différentes. Pour le domaine Ti-DM, on a des séquences assez longues (environ 170 acides aminés) qui ont peu d'homologie. Par contre, pour le domaine C, les séquences sont plus courtes (70 acides aminés) et très homologues entre elles. Un tel domaine sera donc beaucoup plus difficile à traiter dans une analyse phylogénique. Le taux d'homoplasie (substitutions sans signification évolutive) risque d'être très élevé et les arbres construits sont donc plus instables que ceux construits à partir du domaine Ti-DM.

Pour la suite de cette discussion sur les chimères évolutives, nous nous intéresserons surtout aux trois gènes VDR, ECR et FTZ-F1, qui représentent les trois exemples les plus forts d'évolution indépendante des domaines des récepteurs nucléaires.

#### b) Comment confirmer l'existence de récepteurs chimères ?

Existe-t-il des éléments permettant de confirmer la nature chimérique de ces gènes? En fait, pour FTZ-F1 on en a assez peu. Ce récepteur orphelin est décrit depuis trop peu de temps pour que l'on dispose d'informations précises à son sujet. La seule chose que l'on sait, c'est qu'il régule le gène de segmentation de drosophile *fushi-tarazu* (Lavorgna et al., 1991). La séquence de sa "boîte P" (ESCKG) est différente de toutes celles connues pour les autres récepteurs. Ceci explique que son site de fixation sur l'ADN semble différent de ceux des autres récepteurs. Il est tout de même amusant de constater que cette boîte P est un hybride entre la boîte P des récepteurs des hormones thyroïdiennes ou de nombreux récepteurs orphelins comme RXR (EGCKG), et celle des récepteurs des stéroïdes, comme celui des glucocorticoïdes (GSCKV). Bien que cette boîte P soit à un endroit où s'insèrent certains des introns entre C1 et C2, la signification de cette observation n'est pas très claire...

Pour le récepteur de l'ecdysone (ECR), on se heurte au même problème que pour FTZ-F1, ce gène a été décrit très récemment, et on manque encore d'informations à son propos (Koelle et al., 1991). Il a été annoncé qu'il pouvait produire de nombreuses formes alternativement épissées, mais on aimerait savoir précisément quelles sont ces formes (Hoffman et Joose, 1991).

Il possède la même boîte P et a une spécificité de fixation proche de celle des membres de la sous-famille I (Koelle et al., 1991; Martinez et al., 1991). Du point de vue de l'organisation génomique, il convient d'être méfiant, parce que c'est un gène de drosophile et que les règles que nous avons fixées plus haut (voir la discussion de l'article I) ne semblent pas très respectées chez ces insectes. Ainsi, ECR, n'a pas d'intron entre C1 et C2 alors qu'il a plusieurs exons dans sa partie 5', notamment dans le domaine A/B. Mais est-ce le diagnostic d'une classe particulière, ou est-ce le reflet de l'histoire individuelle de ce gène ? Nous l'ignorons. A propos de ECR, il faut remarquer sa forte homologie de séquence avec les récepteurs des hormones thyroïdiennes. Cette homologie explique sans doute pourquoi l'une des méthodes de phylogénie utilisée, la méthode "Neighbor-Joining" place ECR à côté de ces récepteurs dans l'arbre du domaine C. Comme ECR est également très proche de VDR, l'arbre Neighbor-Joining (NJ) déplace aussi ce gène (et les gènes du groupe *knirps*) pour les mettre tout à côté des récepteurs des hormones thyroïdiennes. Par contre, l'arbre construit par la méthode de Fitch et Margoliash, qui est généralement considérée comme plus fiable que la méthode NJ (Swofford et Olsen, 1990), place le groupe {ECR, VDR et *knirps*} dans la sous-famille III, comme indiqué sur la figure 43. Nous pensons que cette localisation est celle qui est correcte et que la disposition calculée par la méthode NJ est le résultat de la convergence existant entre ECR et les THR. Deux arguments vont dans ce sens : d'abord, si on enlève ECR, VDR et le groupe *knirps* sont associés à la sous-famille III de manière stable. Ensuite, des calculs statistiques de confiance sur les arbres, menés par la méthode du "bootstrap" indiquent que le placement de ECR près des THR, dans l'arbre NJ, est très mal confirmé. Ceci suggère donc qu'ECR et ses acolytes VDR, KNI, KNRL et EGON, sont bien des membres de la sous-famille III, pour le domaine C. La position de ECR et de VDR dans la sous-famille I, au sein du groupe EAR-1 est, elle, clairement confirmée dans toutes les analyses que nous avons réalisées. Cet exemple illustre combien il faut interpréter prudemment les données de phylogénie moléculaire et essayer, lorsque cela est possible, de les confirmer par des arguments structuraux ou physiologiques. Pour l'instant, ce type d'argument manque encore pour ECR.

Le récepteur de la vitamine D offre un meilleur exemple de gène chimère, parce que d'autres arguments vont dans le même sens. Ainsi, la position de l'intron entre C1 et C2 et le nombre d'exons dans le domaine A/B sont clairement en faveur de l'appartenance de ces deux domaines à la sous-famille III (Hughes et al., 1988). Or, l'analyse de phylogénie moléculaire place le domaine C de VDR au sein de cette sous-famille. Pour le domaine Ti-DM, l'analyse de la position des introns est moins informative, parce que ces introns sont (nous l'avons vu dans les résultats) placés de manière assez variable d'un gène à l'autre. Une analyse plus précise est actuellement en cours pour voir si l'on peut tout de même amener des éléments d'information nouveaux. Par contre, d'un point de vue fonctionnel, il semble admis que VDR se rapproche des THR et RARs, par exemple pour la dimérisation. Ainsi, certains auteurs ont indiqué qu'il semblait y avoir des dimérisations entre VDR et les THR (Forman et al., 1989; Forman et Samuels, 1990a; 1990b). Le seul argument fonctionnel qui ne colle pas avec l'analyse phylogénique est la spécificité

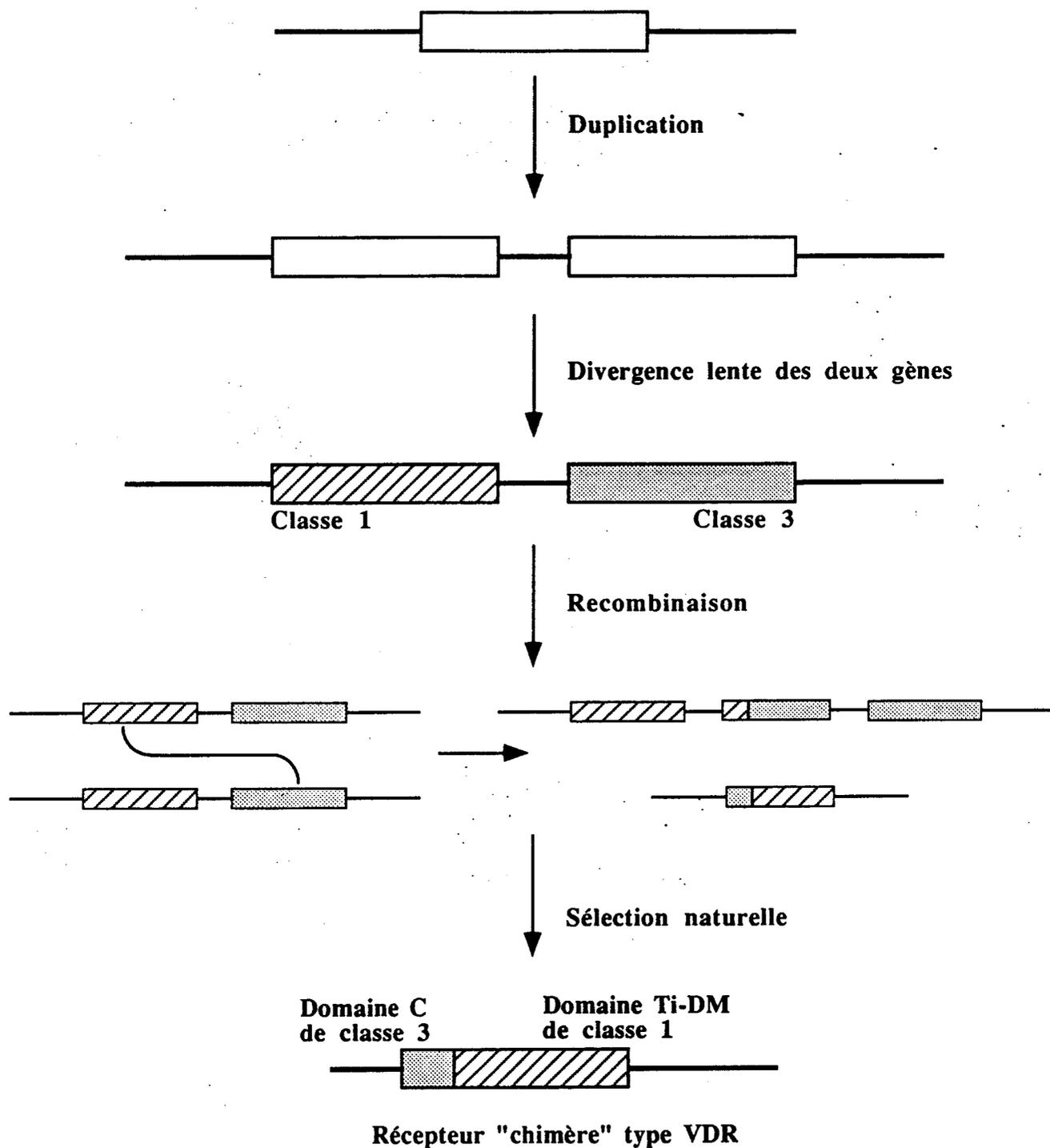


Figure 53 : Un mécanisme possible permettant d'expliquer la genèse de récepteurs chimères (VDR, par exemple).

de fixation à l'ADN. Nous avons vu que pour le domaine C l'analyse phylogénique place VDR avec la sous-famille III. Or, VDR a une boîte P identique à celle des THR<sub>s</sub> et RAR<sub>s</sub> de la sous-famille I. L'évènement qui a conduit à la formation d'un VDR chimère doit s'être produit il y a très longtemps. Pendant tout ce temps, il a été facile, pour le VDR muté, de "corriger" ou, de manière plus neutre, de "modifier" sa boîte P (constitué seulement de 3 acides aminés variables) pour que le récepteur chimère conserve une affinité pour les TRE et une capacité d'interaction avec les produits des THR<sub>s</sub>, des RAR<sub>s</sub> ou d'autres récepteurs de spécificité voisine. Le fait que la reconnaissance spécifique d'une séquence soit codée par un tout petit nombre d'acides aminés donne ainsi une grande flexibilité évolutive et fonctionnelle aux récepteurs nucléaires.

c) Les mécanismes possibles ayant pu conduire aux chimères.

Si effectivement des chimères évolutives des récepteurs nucléaires existent, il faut essayer d'imaginer par quel type de mécanismes elles ont été formées. Nous venons de voir que ces récepteurs peuvent sans doute réagir, avec une grande plasticité, à la mise côte à côte de deux domaines C et Ti-DM différents. Mais comment un tel mécanisme a-t'il pu avoir lieu ? Il est en effet beaucoup plus facile d'imaginer un mécanisme expliquant la structure actuelle des gènes du groupe *knirps* : une simple translocation chromosomique suivie d'une perte des introns au cours de l'évolution, et le tour est joué ! Par contre, pour des chimères contenant deux domaines différents mais provenant tous les deux de récepteurs nucléaires ancestraux, ceci est plus difficile à expliquer!

Etant donné le nombre de gènes présents dans le génome (au moins 50 000 ?), il faudrait une sacrée "chance" pour qu'une telle chimère existe par le biais du même mécanisme... Certes, les deux chimères les plus convaincantes (VDR et ECR) semblent dériver d'un ancêtre commun et on peut imaginer qu'un tel évènement ait eu lieu une fois et une seule, uniquement par hasard. Mais, si on confirme qu'il existe d'autres chimères (comme FTZ-F1 ou NGF1B), ce mécanisme ne fonctionne plus. Il n'est donc, au moins pour l'instant, pas très satisfaisant pour l'esprit. Essayons de trouver autre chose...

Les processus de recombinaison générale sont suspectés d'avoir un rôle dans l'évolution et dans les évènements de remodelage des génomes qui l'accompagnent (Clarke et al., 1986). Le mécanisme illustré sur la figure 53 peut entraîner la formation d'une chimère, comme le gène du VDR. Ce mécanisme nécessite deux évènements distincts. Dans un premier temps, il y a duplication d'un gène ancêtre. Les deux copies produites restent côte à côte. Si, sur de telles répétitions en tandem, des évènements de recombinaison ou de conversion génique inégale se produisent, on pourra obtenir une copie de gène chimère.

Cependant, ce mécanisme n'est pas entièrement satisfaisant, car on sait que, suite à ce genre d'évènement, les séquences d'ADN répétées en tandem ont tendance à rester identiques parce que des évènements de conversion génique brassent les copies ainsi formées (Smith, 1976; Kourilsky, 1986). Les mutations introduites dans une copie donnée auront tendance à se disperser

dans les autres copies. On aura donc une homogénéisation des séquences d'ADN apparentées. Or, pour expliquer la genèse d'un gène comme VDR, il faut supposer que les deux ancêtres qui ont recombinaison devaient être bien différents. On doit donc imaginer qu'un laps de temps assez long s'est écoulé entre les deux événements (la duplication et la recombinaison inégale). Ce laps de temps a ainsi permis aux deux gènes de diverger.

Comment ces deux gènes, différents mais appartenant à la même famille, ont alors pu recombinaison reste un mystère. Eventuellement, on peut émettre l'hypothèse d'un mécanisme dérivé de la recombinaison spécifique de site. C'est ce type de recombinaison qui permet l'intégration des éléments transposables du génome (comme les séquences Alu ou les séquences LINE) et des rétrovirus (Deininger et Daniels, 1986; Coen et Carpenter, 1986). Dans deux gènes proches mais différents, il doit exister de nombreuses zones très courtes de forte homologie entre séquences. S'il y a eu un mécanisme de recombinaison spécifique de site, on peut imaginer que deux gènes situés côte à côte recombinaison pour former une copie chimérique.

On pourrait ainsi continuer de spéculer longtemps sur le mécanisme qui a engendré les récepteurs chimères comme VDR ou ECR. Il est clair que, au moins pour ces deux gènes, il y a eu un événement unique impliquant un ancêtre des deux, avant la dichotomie arthropodes/vertébrés. Cet ancêtre chimère a également donné les gènes du groupe *knirps* qui n'ont pas de domaine Ti-DM. Peut-être ces deux événements sont-ils liés ? Il est en effet curieux de noter que dans cette branche de l'arbre du domaine C, tous les gènes semblent avoir une histoire complexe. Il a dû se passer un événement fortuit qui a conduit à la curieuse situation que l'on observe actuellement.

#### 4) Histoire et récepteurs nucléaires.

La situation actuelle observée chez un organisme, que ce soit la structure d'un gène de récepteur nucléaire ou la morphologie d'un organe est, bien sûr, le résultat de l'évolution. Or, l'évolution est un processus historique, une suite d'événements complexes qui se succèdent au cours du temps. On peut retrouver, parfois expliquer, ces événements, soit en observant les traces des organismes passés, soit en comparant les formes actuelles. Si on parle de processus historique, cela implique qu'un événement donné dépend de ce qui a précédé. Les organismes vivants sont les produits de la sélection naturelle qui leur permet de s'adapter à leur environnement mais ils sont aussi le produit de l'histoire des organismes qui les ont précédés. La sélection naturelle va agir dans le cadre de l'histoire des organismes (Gould, 1991).

Les récepteurs nucléaires sont, à mon sens, un bel exemple de cette idée de contingence historique comme composante essentielle de l'histoire évolutive des organismes vivants. Personne ne sait ce qui s'est passé à leur origine, mais peut-être qu'un événement fortuit a réuni des exons codant des doigts de zinc et des exons codant un domaine de fixation des stéroïdes pour donner un

## REFERENCES

- AKERBLOM, I.E. and P.L.MELLON. (1991). Repression of gene expression by steroid and thyroid hormone. In Nuclear Hormone Receptors. Ed. Parker, M.G. Academic Press : 175-191.
- ALCALAY, M., D.ZANGRILLI, P.P.PANDOLFI, L.LONGO, A.MENCARELLI, A.GIACOMUCCI, M.ROCCHI, A.BIONDI, A.RAMBALDI, F.LOCOCO, D.DIVERIO, E.DONTI, F.GRIGNANI and P.G.PELICCI. (1991). Translocation breakpoint of acute promyelocytic leukemia lies within the retinoic acid receptor  $\alpha$  locus. Proc. Natl. Acad. Sci. USA **88** : 1977-1981.
- ALTABA, A.R.I. and T.JESSEL. (1991). Retinoic acid modifies mesodermal patterning in early *Xenopus* embryos. Genes and Dev. **5** : 175-187.
- ANKENBAUER, W., U.STRÄHLE and G.SCHÜTZ. (1988). Synergistic action of glucocorticoid and estradiol responsive elements. Proc. Natl. Acad. Sci. USA **85** : 7526-7530.
- ARRIZA, J.L., C.WEINBERGER, G.CERELLI, T.M.GLASER, B.L.HANDELIN, D.E.HOUSMAN and R.M.EVANS. (1987). Cloning of human mineralocorticoid receptor complementary DNA : structural and functional kinship with the glucocorticoid receptor. Science **237** : 268-275.
- ARRIZA, J.L., R.B.SIMERLY, L.W.SWANSON and R.M.EVANS. (1988). Neuronal mineralocorticoid receptor as a mediator of glucocorticoid response. Neuron **1** : 887-900.
- ASHBURNER, M. (1990). Puffs, genes, and hormones revisited. Cell **61** : 1-3.
- AURBACH, G.D., S.J.MARX and A.M.SPIEGEL. (1985). Parathyroid hormone, calcitonin, and the calciferols. In Williams Textbook of Endocrinology. Seventh Edition. Eds. J.D.Wilson and D.W.Foster : 1137-1217.
- BONVALET, P., N.FARMAN, M. LOMBES et M.E. OBLIN. (1991). Compte-rendu du Symposium international sur l'aldostérone. Médecine/Sciences **7** : 748-749.
- BACHNOU, N., V.LAUDET, T.JAFFREDO, B.QUATANNENS, S.SAULE and F.DIETERLEN-LIEVRE. (1991). Cooperative effect of v-myc and v-erbA in the chick embryo. Oncogene **6** : 1041-1047.
- BAGCHI, M.K., S.Y.TSAI, M.J.TSAI and B.W.O'MALLEY. (1987). Purification and characterization of chicken ovalbumin gene upstream promoter transcription factor from homologous oviduct cells. Mol. Cell. Biol. **7** : 4151-4158.
- BAILLY, A., C.LEPAGE, M.RAUCH and E.MILGROM. (1986). Sequence-specific DNA binding of the progesterone receptor to the uteroglobin gene : effects of hormone, anti-hormone and receptor phosphorylation. The EMBO J. **5** : 3235-3241.
- BAKER, A.R., D.P.McDONNELL, M.HUGHES, T.M.CRISP, D.J.MANGELSDORF, M.R.HAUSSLER, J.WESLEY PIKE and J.SHINE. (1988). Cloning and expression of full-length cDNA encoding human vitamin D receptor. Proc. Natl. Acad. Sci. USA. **85** : 3294-3298.
- BAKER, B.S. and J.R.TATA. (1990). Accumulation of proto-oncogene c-erbA related transcripts during *Xenopus* development : association with early acquisition of response to thyroid hormone and estrogen. The EMBO J. **9** : 879-885.

- BALLING,R. (1991). CRABP and the teratogenic effects of retinoids.  
TIG 7 : 35-36.
- BANIAHMAD,A., C.STEINER, A.C.KOHNE and R.RENKAWITZ. (1990). Modular structure of a chicken lysozyme silencer : involvement of an unusual thyroid hormone receptor binding site.  
Cell 61 : 505-514.
- BANKER,D.E., J.BIGLER and R.N.EISENMAN. (1991). The thyroid hormone receptor gene (*c-erbA $\alpha$* ) is expressed in advance of thyroid gland maturation during the early embryonic development of *Xenopus laevis*.  
Mol. Cell. Biol. 11 : 5079-5089.
- BARNES,R.S.K. (1991). Two-layered awakening.  
Nature 350 : 561.
- BARRINGTON,E.J.W. (1964). Hormones and Evolution, The English Universities Press, Londres .
- BARSONY,J., J.W.PIKE, H.F.DELUCA and S.J.MARX. (1990). Immunocytology with microwave fixed fibroblasts shows  $1\alpha,25$ -dihydroxyvitamin D<sub>3</sub>, dependent rapid and estrogen dependent slow reorganization of vitamin D receptors.  
J. Cell Biol. 11 : 2385-2395.
- BAULIEU,E.E. (1989). Contragestion and other clinical applications of RU 486, an antiprogestone at the receptor.  
Science 245 : 1351-1357.
- BEATO,M. (1989). Gene regulation by steroid hormones.  
Cell 56 : 335-344.
- BEATO,M., D.BARETTINO, U.BRUGGEMEIR, G.CHALEPAKIS, R.J.G.HACHE, M.KALFF, B.PINA, M.SCHAUER, E.P.SLATER and M.TRUSS. (1991). Characterization of DNA receptor interactions. In Nuclear Hormone Receptors. Ed.Parker,M.G. Academic Press : 197-211.
- BECKER,P.B., B.GLOSS, W.SCHMID, U.STRAHLE and G.SCHUTZ. (1986). *In vivo* protein-DNA interactions in a glucocorticoid response element.  
Nature 324 : 686-688.
- BECKER,D.M., S.M.HOLLENGERG and R.P.RICCIARDI. (1989). Fusion of adenovirus E1A to the glucocorticoid receptor by high-resolution deletion cloning creates a hormonally inducible viral transactivator.  
Mol. Cell. Biol. 9 : 3878-3887.
- BEDO,G., P.SANTISTEBAN and A.ARANDA. (1989). Retinoic acid regulates growth hormone gene expression.  
Nature 339 : 231-234.
- BEEBE,J.S., D.S.DARLING and W.W.CHIN. (1991). 3,5,3'-triiodothyrosine receptor auxiliary protein (TRAP) enhances receptor binding by interactions within the thyroid hormone response element.  
Mol. Endo 5 : 85-93.
- BENBROOK,D. and M.PFAHL. (1987). A novel thyroid hormone receptor encoded by a cDNA clone from a human testis library.  
Science 238 : 788-791.

- BENBROOK, D., E. LERNHARDT and M. PFAHL. (1988). A new retinoic acid receptor identified from a hepatocellular carcinoma. *Nature* **333** : 669-672.
- BENDER, W. and M. PFEIFER. (1987). Oncogenes take wing. *Cell* **50** : 519-520.
- BERG, J.M. (1989). DNA binding specificity of steroid receptors. *Cell* **57** : 1065-1068.
- BERKS, M., D. TRAYNOR, I. CARRIN, R.H. INSALL and R.R. KAY. (1991). Diffusible signal molecules controlling cell differentiation and patterning in *Dictyostelium*. *Development* **1** : 131-139.
- BERRY, M., A.M. NUNEZ and P. CHAMBON. (1989). Estrogen-responsive element of the human pS2 gene is an imperfectly palindromic sequence. *Proc. Natl. Acad. Sci. USA* **86** : 1218-1222.
- BERRY, M.J., A.L. KATES and P. REED LARSEN. (1990). Thyroid hormone regulates type I deiodinase messenger RNA in rat liver. *Mol. Endo.* **4** : 743-748.
- BERRY, M.J., D. METZGER and P. CHAMBON. (1990). Role of the two activating domains of the estrogen receptor in the cell-type and promoter-context dependent agonistic activity of the anti-estrogen 4-hydroxytamoxifen. *The EMBO J.* **9** : 2811-2818.
- BEUG, H., A. von KIRCHBACH, G. DODERLEIN, J.F. CONSCIENCE and T. GRAF. (1979). Chicken hematopoietic cells transformed by seven strains of defective avian leukemia viruses display three distinct phenotypes of differentiation. *Cell* **18** : 375-390.
- BEUG, H., G. DOEDERLEIN, C. FREUDENSTEIN and T. GRAF. (1982). Erythroblast cell lines transformed by a temperature-sensitive mutant of avian erythroblastosis virus : a model system to study erythroid differentiation *in vitro*. *J. Cell. Phys.* **1** : 195-207.
- BEUG, H. and B. VENNSTROM. (1991). Avian erythroleukaemia : possible mechanisms involved in v-erbA oncogene function. In *Nuclear Hormone Receptors*. Ed. Parker, M.G. Academic Press : 355-372.
- BIGLER, J. and R.N. EISENMAN. (1988). c-erbA encodes multiple proteins in chicken erythroid cells. *Mol. Cell. Biol.* **8** : 4155-4161.
- BISBEE, C.A. M.A. BAKER, A.C. WILSON, I. HADJI-AZIMI and M. FISCHBERG (1977). Albumin phylogeny for clawed frogs (*Xenopus*). *Science* **195** : 785-787.
- BLOMHOFF, R., M.H. GREEN, T. BERG and K.R. NORUM. (1990). Transport and storage of vitamin A. *Science* **250** : 399-404.
- BOCQUEL, M.T., V. KUMAR, C. STRICKER, P. CHAMBON and H. GRONEMEYER. (1989). The contribution of the N- and C-terminal regions of steroid receptors to activation of transcription is both receptor and cell-specific. *Nucl. Acids. Res.* **17** : 2581-2595.

- COONEY,A.J., S.Y.TSAI, B.W.O'MALLEY and M.J.TSAI. (1991). Chicken ovalbumin upstream promoter transcription factor binds to a negative regulatory region in the human immunodeficiency virus type 1 long terminal repeat.  
J. Virol. 65 : 2853-2860.
- COPP,R.P. and H.H.SAMUELS. (1989). Identification of an adenosine 3', 5'-monophosphate (cAMP)-responsive region in the rat growth hormone gene : evidence for independent and synergistic effects of cAMP and thyroid hormone on gene expression.  
Mol. Endo. 3 : 790-796.
- CORDINGLEY,M.G., A.T.RIEGEL and G.L.HAGER. (1987). Steroid-dependent interaction of transcription factors with the inducible promoter of mouse mammary tumor virus *in vivo*.  
Cell 48 : 261-270.
- CORDINGLEY,M.G. and G.L.HAGER. (1988). Binding of multiple factors to the MMTV promoter in crude and fractionated nuclear extracts.  
Nucl. Acids Res. 16 : 609-630.
- COURTOIS,S.J., D.A.LAFONTAINE, F.P.LEMAIGRE, S.M.DURVIAUX and G.G.ROUSSEAU. (1990). Nuclear factor-1 and activator protein-2 bind in a mutually exclusive way to overlapping promoter sequences and *trans*-activate the human growth hormone gene.  
Nucl. Acids Res. 18 : 57-64.
- CROSS,J.W. (1991). Cycling of auxin-binding protein through the plant cell : pathways in auxin signal transduction.  
The New Biol. 3 : 813-819.
- CRYER,P.E. (1985). Glucose homeostasis and hypoglycemia. In Williams Textbook of Endocrinology. Seventh Edition. Eds. J.D.Wilson and D.W.Foster : 989-1017.
- CULLEN,K.J., D.YEE, S.PAIK, B.HAMPTON, J.F.PERDUE, M.E.LIPPMAN and N.ROSEN. (1988). Insulin-like growth factor II expression and activity in human breast cancer.  
Abstract 947, Proc. Annu. Meet. Am. Assoc. Cancer Res. 29 : 238.
- CUTHILL,S., A.WILHELMSSON and L.POELLINGER. (1991). Role of the ligand in intracellular receptor function : receptor affinity determines activation *in vitro* of the latent dioxin receptor to a DNA-binding form.  
Mol. Cell. Biol. 11 : 401-411.
- DALMAN,F.C., R.J.KOENIG, G.H.PERDEW, E.MASSA and W.B.PRATT. (1990). In contrast to the glucocorticoid receptor, the thyroid hormone receptor is translated in the DNA binding state and is not associated with hsp90.  
J. Biol. Chem. 265 : 3615-3618.
- DAMM,K., H.BEUG, T.GRAF and B.VENNSTROM. (1987). A single point mutation in erbA restores the erythroid transforming potential of a mutant avian erythroblastosis virus (AEV) defective in both erbA and erbB oncogenes.  
The EMBO J. 6 : 375-382.
- DAMM,K., C.C.THOMPSON and R.M.EVANS. (1989). Protein encoded by v-erbA functions as a thyroid-hormone receptor antagonist.  
Nature 339 : 593-597.
- DANA,S. and M.KARIN. (1989). Induction of human growth hormone promoter activity by the adenosine 3', 5'-monophosphate pathway involves a novel responsive element.  
Mol. Endo. 3 : 815-821.
- DANESH,U., B.GLOSS, W.SCHMID, G.SCHÜTZ and R.RENKAWITZ. (1987). Glucocorticoid induction of the rat tryptophan oxygenase gene is mediated by two widely separated glucocorticoid-responsive elements.  
The EMBO J. 6 : 625-630.

- DANIELSEN, M., L. HINCK and G. M. (1989). Two amino acids within the knuckle of the first zinc finger specify DNA response element activation by the glucocorticoid receptor. *Cell* **57** : 1131-1138.
- DANIELSEN, M. (1991). Structure and function of the glucocorticoid receptor. In *Nuclear Hormone Receptors*. Ed. Parker, M. G. Academic Press : 39-68.
- DARLING, D. S., J. S. BEEBE, J. BURNSIDE, E. R. WINSLOW and W. W. CHIN. (1991). 3,5,3'-triiodothyronine (T3) receptor-auxiliary protein (TRAP) binds DNA and forms heterodimers with the T3 receptor. *Mol. Endo* **5** : 73-84.
- DAVIDSON, N. O., CARLOS, R. C., M. J. DREWEK and T. G. PARMER. (1988). Apolipoprotein gene expression in the rat is regulated in a tissue-specific manner by thyroid hormone. *J. Lipid Res.* **29** : 1511-1522.
- DAYTON, A. I., J. R. SELDEN, G. LAWS, D. J. DORNEY, J. FINAN, P. TRIPPUTI, B. S. EMANUEL, G. ROVERA, P. C. NOWEL and C. M. CROCE. (1984). A Human c-erbA oncogene homologue is closely proximal to the chromosome 17 breakpoint in acute promyelocytic leukemia. *Proc. Natl. Acad. Sci. USA* **81** : 4495-4499.
- DEBUIRE, B., C. HENRY, M. BENAÏSSA, G. BISERTE, J. M. CLAVERIE, S. SAULE, P. MARTIN and D. STEHELIN. (1984). Sequencing the erbA gene of avian erythroblastosis virus reveals a new type of oncogene. *Science* **224** : 1456-1459.
- DEGOS, L., S. CASTAIGNE, P. FENAUX, M. T. DANIEL and C. CHOMIENNE. (1991). Le traitement des leucémies aiguës à promyélocytes par l'acide tout-trans rétinolique. *Médecine/Science* **5** : 460-464.
- DEININGER, P. L., G. R. DANIELS. (1986). The recent evolution of mammalian repetitive DNA elements. *TIG* **2** : 76-80.
- DEJEAN, A., L. BOUGUELERET, K. H. GRZESCHIK and P. TIOLLAIS. (1986). Hepatitis B virus DNA integration in a sequence homologous to v-erb-A and steroid receptor genes in a hepatocellular carcinoma. *Nature* **322** : 70-72.
- DENIS, M., J. A. GUSTAFSSON and A. C. WIKSTROM. (1988). Interaction of the Mr = 90,000 heat shock protein with the steroid-binding domain of the glucocorticoid receptor. *J. Biol. Chem.* **263** : 18520-18523.
- DE ROBERTIS, E. M., E. A. MORITA and K. W. Y. CHO. (1991). Gradient fields and homeobox genes. *Development* **112** : 669-678.
- DESBOIS, C., B. PAIN, C. GUILHOT, M. BENCHAIIBI, M. FFRENCH, J. GHYSDAEL, J.-J. MADJAR and J. SAMARUT. (1991a). v-erbA oncogene abrogates growth inhibition of chicken embryo fibroblasts induced by retinoic acid. *Oncogene* **6** : 2129-2135.
- DESBOIS, C., D. AUBERT, C. LEGRAND, B. PAIN and J. SAMARUT. (1991b). A novel mechanism of action for v-erbA : abrogation of the inactivation of transcription factor AP-1 by retinoic acid and thyroid hormone receptors. *Cell* **67** : 731-740.
- DESVERGNE, B., K. J. PETTY and V. M. NIKODEM. (1991). Functional characterization and receptor binding studies of the malic enzyme thyroid hormone response element. *J. Biol. Chem.* **266** : 1008-1013.

- DE THE,H., A.MARCHIO, P.THIOLLAIS and A.DEJEAN. (1987). A novel steroid thyroid hormone receptor-related gene inappropriately expressed in human hepatocellular carcinoma. *Nature* **330** : 667-670.
- DE THE,H., C.CHOMIENNE, M.LANOTTE, L.DEGOS and A.DEJEAN. (1990a). The T(15;17) translocation of acute promyelocytic leukaemia fuses the retinoic acid receptor  $\alpha$  gene to a novel transcribed locus. *Nature* **347** : 558-561.
- DE THE,H., M.M.VIVANCO-RUIZ, P.TIOLLAIS, H.STUNNENBERG and A.DEJEAN. (1990b). Identification of a retinoic acid responsive element in the retinoic acid receptor  $\beta$  gene. *Nature* **343** : 177-180.
- DE THE,H., C.LAVAU, A.MARCHIO, C.CHOMIENNE, L.DEGOS and A.DEJEAN. (1991). The PML-RAR $\alpha$  fusion mRNA generated by the t(15;17) translocation in acute promyelocytic leukemia encodes a functionally altered RAR. *Cell* **66** : 675-684.
- DE VERNEUIL,H. and D.METZGER. (1990). The lack of transcriptional activation of the y-erbA oncogene is in part due to a mutation present in the DNA binding domain of the protein. *Nuc. Acids. Res.* **18** : 4489-4497.
- DIAMOND,M.I., J.N.MINER, S.K.YOSHINAGA and K.R.YAMAMOTO. (1990). Transcription factor interactions : selectors of positive or negative regulation from a single DNA element. *Science* **249** : 1266-1272.
- DIBB,N.J. and A.J.NEWMAN. (1989). Evidence that introns arose at proto-splice sites. *The EMBO J.* **8** : 2015-2021.
- DISELA,C., C.GLINEUR, T.BUGGE, J.SAP, G.STENGL, J.DODGSON, H.STUNNENBERG, H.BEUG and M.ZENKE. (1991). overexpression is required to extinguish c-erbA function in erythroid cell differentiation and regulation of the erbA target gene CAII. *Genes and Dev.* **5** : 2033-2047.
- DOLLE,P., J.C.IZPISUA-BELMONTE, H.FALKENSTEIN, A.RENUCCI and D.DUBOULE. (1989). Coordinate expression of the murine Hox complex homeobox-containing genes during limb pattern formation. *Nature* **342** : 767-772.
- DOLLE,P., E.RUBERTE, P.KASTNER, M.PETKOVICH, C.M.STONER, L.J.GUDAS and P.CHAMBON. (1989). Differential expression of genes encoding  $\alpha$ ,  $\beta$  and  $\gamma$  retinoic acid receptors and CRABP in the developing limbs of the mouse. *Nature* **342** : 702-705.
- DOLLE,P., E.RUBERTE, P.LEROY, G.MORRISS-KAY and P.CHAMBON. (1990). Retinoic acid receptors and cellular retinoid binding proteins. *Development* **110** : 1133-1151.
- DORIT,R.L. and W.GILBERT. (1991). The limited universe of exons. *Cur. Opinion Genet. and Dev.* **1** : 464-469.
- DOUCAS,V., G.SPYROU and M.YANIV. (1991). Unregulated expression of c-jun or c-fos proteins but not jun D inhibits estrogen receptor activity in human breast cancer derived cells. *The EMBO J.* **10** : 2237-2245.
- DOUGLAS,J.B., M.C.DALY and P.H.RABBITS. (1991). A pulsed-field gel electrophoresis map locates the polymorphic probes for erbA2 and erbA $\beta$  within 120 kb of each other, confirming that THRB (formerly erbA2) maps to chromosome 3. *Genomics* **9** : 398-400.

- DOWNS, T.R., P. CHOMCZYNSKI and L.A. FROHMAN. (1990). Effects of thyroid hormone deficiency and replacement on rat hypothalamic growth hormone (GH)-releasing hormone gene expression *in vivo* are mediated by GH.  
Mol. Endo. **4** : 402-408.
- DRABKIN, H., F-T. KAO, J. HARTZ, I. HART, A. GAZDAR, C. WEINBERGER, R. EVANS and M. GERBER. (1988). Localization of human erbA-2 to the 3p22--->3p24.1 region of chromosome 3 and variable deletion in small cell lung cancer.  
Proc. Natl. Acad. Sci. **85** : 9258-9262.
- DREYER, C., G. KREY, H. KELLER, F. GIVEL, G. HELFTENBEIN and W. WAHLI. (1992) Control of the peroxisomal  $\beta$ -oxidation pathway by a novel family of nuclear hormone receptors.  
Cell (in press).
- DURSTON, A.J., J.P.M. TIMMERMANS, W.J. HAGE, H.F.J. HENDRIKS, N.J. de VRIES, M. HEIDEVELD and P.D. NIEUWKOOP. (1989). Retinoic acid causes an anteroposterior transformation in the developing central nervous system.  
Nature **340** : 140-144.
- EAGER, N.S.C., A. ROWE and P.M. BRICKELL. (1991). A member of the chicken RXR family of nuclear receptors activates transcription in response to retinoic acid.  
FEBS Lett. **292** : 103-106.
- EICHELE, G. (1989). Retinoids and vertebrate limb pattern formation.  
TIG **5** : 246-251.
- EILERS, M., D. PICARD, K.R. YAMAMOTO and J.M. BISHOP. (1989). Chimaeras of myc oncoprotein and steroid receptors cause hormone-dependent transformation of cells.  
Nature **340** : 66-68.
- ERWIN, D.H. (1991). Metazoan phylogeny and the cambrian radiation.  
TREE **6** : 131-134.
- EVANS, R. (1988). The steroid and thyroid hormone receptor superfamily.  
Science **240** : 889-895.
- EVANS, R.M. and J.L. ARRIZA. (1989). A molecular framework for the actions of glucocorticoid hormones in the nervous system.  
Neuron **2** : 1105-1112.
- EVERETT, A.W., W.A. CLARK, R.A. CHIZZONITE and R. ZAK. (1983). Change in the synthesis rates of  $\alpha$ - and  $\beta$ -myosin heavy chains in rabbit heart after treatment with thyroid hormone.  
J. Biol. Chem. **258** : 2421-2425.
- FAHRNER, T.J., S.L. CARROLL and J. MILBRANDT. (1990). The NGFI-B protein, an inducible member of the thyroid/steroid receptor family, is rapidly modified posttranslationally.  
Mol. Cell. Biol. **10** : 6454-6459.
- FAWELL, S.E., J.A. LEES, R. WHITE and M.G. PARKER. (1990). Characterization and colocalization of steroid binding and dimerization activities in the mouse estrogen receptor.  
Cell **60** : 953-962.
- FELDMAN, D., L.G. TOKES, P.A. STATHIS, S.C. MILLER, W. KURZ and D. HARVEY. (1984). Identification of 17 $\beta$ -estradiol as the estrogenic substance in *Saccharomyces cerevisiae*.  
Proc. Natl. Acad. Sci. USA **81** : 4722-4726.
- FIRTEL, R.A. (1991). Signal transduction pathways controlling multicellular development in *Dictyostelium*.  
TIG **7** : 381-388.

- FITCH, W.M. and E.MARGOLIASH. (1967). Construction of phylogenetic trees. *Science* **155** : 279-284.
- FITCH, W.M. (1981). A non-sequential method for constructing trees and hierarchical classifications. *J.Mol. Evol.* **18** : 30-37.
- FLUG, F., R.P.COPP, J.CASANOVA, Z.D.HOROWITZ, L.JANOCKO, M.PLOTNICK and H.H.SAMUELS (1987). Cis-acting elements of the rat growth hormone gene which mediate basal and regulated expression by thyroid hormone. *J.Biol.Chem.* **262** : 6373-6382.
- FONTAINE, M. (1969). Glandes endocrines et hormones. Encyclopédie La Pleiade, Physiologie. Ed.Fontaine M. Paris, 1491-1547.
- FORMAN, B.M., C-R.YANG, F.STANLEY, J.CASANOVA and H.H.SAMUELS. (1988). c-erbA proto-oncogenes mediate thyroid hormone-dependent and independent regulation of the rat growth hormone and prolactin genes. *Mol. Endo.* **2** : 902-911.
- FORMAN, B.M., C.R.YANG, M.AU, J.CASANOVA, J.GHYSDAEL and H.H.SAMUELS. (1989). A domain containing leucine-zipper-like motifs mediate novel *in vivo* interactions between the thyroid hormone and retinoic acid receptors. *Mol. Endo.* **3** : 1610-1626.
- FORMAN, B.M. and H.H.SAMUELS. (1990a). Dimerization among nuclear hormone receptors. *The New Biol.* **2** : 587-594.
- FORMAN, B.M. and H.H.SAMUELS. (1990b). Interactions among a subfamily of nuclear hormone receptors : the regulatory zipper model. *Mol. Endo.* **4** : 1293-1301.
- FORREST, D., A.MUNOZ, C.RAYNOSCHEK, B.VENNSTRÖM and H.BEUG. (1990a). Requirement for the C-terminal domain of the v-erbA oncogene protein for biological function and transcriptional repression. *Oncogene* **5** : 309-316.
- FORREST, D., M.SJOBERG and B.VENNSTRÖM. (1990b). Contrasting developmental and tissue-specific expression of  $\alpha$  and  $\beta$  thyroid hormone receptor genes. *The EMBO J.* **9** : 1519-1538.
- FORREST, D., F.HALLBOOK, H.PERSSON and B.VENNSTRÖM. (1991). Distinct functions for thyroid hormone receptors  $\alpha$  and  $\beta$  in brain development indicated by differential expression of receptor genes. *The EMBO J.* **10** : 269-275.
- FOX, S.R., M.T.C.JONG, J.CASANOVA, Z.S.YE, F.STANLEY and H.H.SAMUELS. (1990). The homeodomain protein, Pit-1/GHF-1, is capable of binding to and activating cell-specific elements of both the growth hormone and prolactin gene promoters. *Mol. Endo.* **4** : 1069-1075.
- FREAKE, H.C., A.SANTOS, Y.GOLDBERG, J.GHYSDAEL and J.H.OPPENHEIMER. (1988). Differences in antibody recognition of the triiodothyronine nuclear receptor and c-erbA products. *Mol. Endo.* **2** : 986-991.
- FREEDMAN, L.P., B.F.LUISI, Z.R.KORSZUN, R.BASAVAPPA, P.B.SIGLER and K.R.YAMAMOTO. (1988). The function and structure of the metal coordination sites within the glucocorticoid receptor DNA binding domain. *Nature* **334** : 543-546.

- FRYKBERG,L., S.PALMIERI, H.BEUG, T.GRAF, M.J.HAYMAN and B.VENNSTRÖM. (1983). Transforming capacities of avian erythroblastosis virus mutants deleted in the erbA or erbB oncogenes.  
Cell 32 : 227-238.
- FUNDER,J.W. (1987). Adrenal steroids : new answers, new question.  
Science 237 : 236-237.
- FUNDER,J.W., P.T.PEARCE, R.SMITH and I.SMITH. (1988). Mineralocorticoid action : target tissue specificity is enzyme, not receptor, mediated.  
Science 242 : 583-585.
- GAMETCHU,B. and R.W.HARRISON. (1984). Characterization of a monoclonal antibody to the rat liver glucocorticoid receptor.  
Endocrinology 114 : 274-279.
- GANDRILLON,O., P.JURDIC, M.BENCHAIBI, J-H.XIAO, J.GHYDAEL and J.SAMARUT. (1987). Expression of the v-erbA oncogene in chicken embryo fibroblasts stimulates their proliferation *in vitro* and enhances tumor growth *in vivo*.  
Cell 49 : 687-697.
- GANDRILLON,O., P.JURDIC, B.PAIN, C.DESBOIS, J.J.MADJAR, M.G.MOSCOVICI, C.MOSCOVICI and J.SAMARUT. (1989). Expression of the v-erbA product, an altered nuclear hormone receptor, is sufficient to transform erythrocytic cells *in vitro*.  
Cell 58 : 115-121.
- GANLY,P.S. and P.H.RABBITTS. (1991). Polymerase chain reaction (PCR) for detection of MspI and DraI polymorphism at the THRB gene.  
Nuc. Acids. Res. 19 : 13
- GAUB,M.P., M.BELLARD, I.SCHEUER, P.CHAMBON and P.SASSONE-CORSI. (1990). Activation of the ovalbumin gene by the estrogen receptor involves the fos-jun complex.  
Cell 63 : 1267-1276.
- GAZZOLO,L., J.SAMARUT, M.BOUABDELLI and J.P.BLANCHET. (1980). Early precursors in the erythroid lineage are the specific target cells of avian erythroblastosis virus *in vitro*.  
Cell 22 : 683-691.
- GEHRING,U. and A.HOTZ. (1983). Photoaffinity labeling and partial proteolysis of wild-type and variant glucocorticoid receptors.  
Biochemistry 22 : 4013-4018.
- GEHRING,W.J., M.MÜLLER, M.AFFOLTER, A.PERICIVAL-SMITH, M.BILLETER, Y.Q.QIAN, G.OLTING and K.WÜTBRICH. (1990). The structure of the homeodomain and its functional implications.  
TIG 6 : 309-341.
- GIGUERE,V., S.M.HOLLENBERG, M.G.ROSENFELD and R.M.EVANS. (1986). Functional domains of the human glucocorticoid receptor.  
Cell 46 : 645-652.
- GIGUERE,V., E.S.ONG, P.SEGUI and R.M.EVANS. (1987). Identification of a receptor for the morphogen retinoic acid.  
Nature 330 : 624-629.
- GIGUERE,V. N.YANG, P.SEGUI and R.M.EVANS. (1988). Identification of a new class of steroid hormone receptors.  
Nature 331 : 91-94.

- GIGUERE, V., E.S.ONG, R.M.EVANS and C.J.TABIN. (1989). Spatial and temporal expression of the retinoic acid receptor in the regenerating amphibian limb. *Nature* **337** : 566-569.
- GINSBERG, A.M., B.O.KING and R.G.ROEDER. (1984). *Xenopus* 5S gene transcription factor, TFIIIA : characterization of a cDNA clone and measurement of RNA levels throughout development. *Cell* **39** : 479-489.
- GLASS, C.K., R.FRANCO, C.WEINBERGER, V.R.ALBERT, R.M.EVANS and M.G.ROSENFELD. (1987). A c-erb-A binding site in rat growth hormone gene mediates trans-activation by thyroid hormone. *Nature* **329** : 738-741.
- GLASS, C.K., J.M.HOLLOWAY, O.V.DEVARY and M.G.ROSENFELD. (1988). The thyroid hormone receptor binds with opposite transcriptional effects to a common sequence motif in thyroid hormone and estrogen response elements. *Cell* **54** : 313-323.
- GLASS, C.K., S.M.LIPKIN, O.V.DEVARY and M.G.ROSENFELD. (1989). Positive and negative regulation of gene transcription by a retinoic acid-thyroid hormone receptor heterodimer. *Cell* **59** : 697-708.
- GLASS, C.K., O.V.DEVARY and M.G.ROSENFELD. (1990). Multiple cell type-specific proteins differentially regulate target sequence recognition by the  $\alpha$  retinoic acid receptor. *Cell* **63** : 729-738.
- GLINEUR, C., M.BAILLY and J.GHYSDAEL. (1989). The c-erbA  $\alpha$ -encoded thyroid hormone receptor is phosphorylated in its amino terminal domain by casein kinase II. *Oncogene* **4** : 1247-1254.
- GLINEUR, C., M.ZENKE, H.BEUG and J.GHYSDAEL. (1990). Phosphorylation of the v-erbA protein is required for its function as an oncogene. *Genes and Development* **4** : 1663-1676.
- GODOWSKI, P.J., S.RUSCONI, R.MIESFELD and K.R.YAMAMOTO. (1987). Glucocorticoid receptor mutants that are constitutive activators of transcriptional enhancement. *Nature* **325** : 365-369.
- GODOWSKI, P.J., D.PICARD and K.R.YAMAMOTO. (1988). Signal transduction and transcriptional regulation by glucocorticoid receptor-LexA fusion proteins. *Science* **214** : 812-816.
- GOLDBERG, Y., C.GLINEUR, J.C.GESQUIERE, A.RICOUART, J.SAP, B.VENNSTROM and J.GHYSDAEL. (1988). Activation of protein kinase C or cAMP-dependent protein kinase increases phosphorylation of the c-erbA-encoded thyroid hormone receptor and of the v-erbA encoded protein. *The EMBO J.* **7** : 2425-2433.
- GORSKI, J., D.TOFT, G.SHYAMALA, D.SMITH and A.NOTIDES. (1968). Hormone receptors : studies on the interaction of estrogen with the uterus. *Recent Prog. Horm. Res.* **24** : 45-80.
- GOUILLEUX, F., B.SOLA, B.COUETTE and H.RICHARD-FOY. (1991). Cooperation between structural elements in hormone-regulated transcription from the Mouse Mammary Tumor Virus promoter. *Nuc. Acids. Res.* **19** : 1563-1569.
- GOULD, S.J. (1991). *La vie est belle*. Coll. "Science ouverte" Editions du Seuil. Paris.

- GRAF, T., A. VON KIRCHBACH and H. BEUG. (1981). Characterization of the hematopoietic target cells of AEV, MC29 and AMV avian leukemia viruses. *Exp. Cell. Res.* **131** : 331-343.
- GRAUPNER, G., K. N. WILLS, M. TZUKERMAN, X. K. ZHANG and M. PFAHL. (1989). Dual regulatory role for thyroid-hormone receptors allows control of retinoic-acid receptor activity. *Nature* **340** : 653-656.
- GREEN, S. and P. CHAMBON. (1986). A superfamily of potentially oncogenic hormone receptors. *Nature* **324** : 615-617.
- GREEN, S., P. WALTER, V. KUMAR, A. KRUST, J. M. BORNERT, P. ARGOS and P. CHAMBON. (1986). Human estrogen receptor cDNA : sequence, expression and homology to *v-erb-A*. *Nature* **320** : 134-139.
- GREEN, S. and P. CHAMBON. (1987). Estradiol induction of a glucocorticoid-responsive gene by a chimaeric receptor. *Nature* **325** : 75-78.
- GREEN, S. and P. CHAMBON. (1988). Nuclear receptors enhance our understanding of transcription regulation. *TIG* **4** : 309-314.
- GREEN, S., V. KUMAR, I. THEULAZ, W. WAHLI and P. CHAMBON. (1988). The N-terminal DNA-binding 'zinc finger' of the estrogen and glucocorticoid receptors determines target gene specificity. *The EMBO J.* **7** : 3037-3044.
- GREEN, S. and P. CHAMBON. (1991). The estrogen receptor : from perception to mechanism. In *Nuclear Hormone Receptors*. Ed. Parker, M. G. Academic Press : 15-33.
- GRONEMEYER, H., TURCOTTE, B., C. QUIRIN-STRICKER, M. T. BOCQUEL, M. E. MEYER, Z. KROZOWSKI, J. M. JELTSCH, T. LEROUGE, J. M. GARNIER and P. CHAMBON. (1987). The chicken progesterone receptor : sequence, expression and functional analysis. *The EMBO J.* **6** : 3985-3994.
- GUIOCHON-MANTEL, A., H. LOOSFELT, P. LESCOP, S. SAR, M. ATGER, M. PERROT-APPLANAT and E. MILGROM. (1989). Mechanisms of nuclear localization of the progesterone receptor : evidence for interaction between monomers. *Cell* **57** : 1147-1154.
- GUIOCHON-MANTEL, A., P. LESOP, S. CHRISTIN-MAITRE, H. LOOSFELT, M. PERROT-APPLANAT and E. MILGROM. (1991). Nucleocytoplasmic shuttling of the progesterone receptor. *The EMBO J.* **10** : 3851-3859.
- GUSTAFSON, T. A., B. E. MARKHAM, J. J. BAHL and E. MORKIN. (1987). Thyroid hormone regulates expression of a transfected  $\alpha$ -myosin heavy-chain fusion gene in fetal heart cells. *Proc. Natl. Acad. Sci. USA* **84** : 3122-3126.
- HAGER, G. L. and T. K. ARCHER. (1991). The interaction of steroid receptors with chromatin. In *Nuclear Hormone Receptors*. Ed. Parker, M. G. Academic Press : 217-232.
- HAMADA, K., S. L. GLEASON, B. Z. LEVI, S. HIRSCHFELD, E. APPELLA and K. OZATO. (1989). H-2RIIBP, a member of the nuclear hormone receptor superfamily that binds to both the regulatory element of major histocompatibility class I genes and the estrogen response element. *Proc. Natl. Acad. Sci. USA* **86** : 8289-8293.

- HAMBLIN, P.S., A.SANTOS, N.C.W.WONG, H.L.SCHWARTZ and J.H.OPPENHEIMER. (1987). Triiodothyronine regulation of multiple rat hepatic genes : requirement for ongoing protein synthesis.  
Mol. Endo. 1 : 397-402.
- HÄRD, T., E.KELLENBACH, R.BOELENS, B.A.MALER, K.DAHLMAN, L.P.FREEDMAN, J.CARLSTEDT-DUKE, K.R.YAMAMOTO, J-A.GUSTAFSSON and R.KAPTEIN. (1990). Solution structure of the glucocorticoid receptor DNA-binding domain.  
Science 249 : 157-160.
- HARLOW, K.W., D.N.SMITH, J.A.KATZENELLENBOGEN, G.L.GREENE and B.S.KATZENELLENBOGEN. (1989). Identification of cysteine 530 as the covalent attachment site of an affinity-labeling estrogen (ketononestrol aziridine) and antiestrogen (tamoxifen aziridine) in the human estrogen receptor.  
J.Biol.Chem. 264 : 17476-17485.
- HARMON, J.M., H.J. EISEN, S.T. BROWER, S.S. SIMONS Jr, C.L.LANGLEY and E.B. THOMPSON. (1984). Identification of human leukemic glucocorticoid receptors using affinity labelling and anti-human glucocorticoid receptor antibodies.  
Cancer Res. 44 : 4540-4547.
- HARRISON, S.C. (1991). A structural taxonomy of DNA-binding domains.  
Nature 353 : 715-719.
- HARTLAND, W.B., A.V.COX, P.G.LLEWELLYN, C.A.G.PICKTON, A.G.SMITH and R.WALTERS. (1982).  
A Geologic Time Scale. Cambridge University Press, Cambridge
- HAYMAN, M.J., H.BEUG and K.W.SAVIN. (1982). Changes in the expression of membrane antigens during the differentiation of chicken erythroblasts.  
J. Cell. Biochem. 18 : 351-362.
- HAZEL, T.G., D.NATHANS and L.F.LAU. (1988). A gene inducible by serum growth factors encodes a member of the steroid and thyroid hormone receptor superfamily.  
Proc. Natl. Acad. Sci. USA 85 : 8444-8448.
- HAZEL, T.G., R.MISRA, I.J.DAVIS, M.E.GREENBERG and L.F.LAU. (1991). Nur77 is differentially modified in PC12 cells upon membrane depolarization and growth factor treatment.  
Mol. Cell. Biol. 11 : 3239-3246.
- HENRICH, V.C., T.J.SLITER, D.B.LUBAHN, A.MACINTYRE and L.I.GILBERT. (1990). A steroid/thyroid hormone receptor superfamily member in *Drosophila melanogaster* that shares extensive sequence similarity with a mammalian homologue.  
Nuc.Acids Res. 18 : 4143-4148.
- HENTZEN, D., A.RENUCCI, D.Le GUELLEC, M.BENCHAIBI, P.JURDIC, O.GANDRILLON and J.SAMARUT. (1987). The chicken c-erbA proto-oncogene is preferentially expressed in erythrocytic cells during late stages of differentiation.  
Mol. Cell. Biol. 7 : 2416-2424.
- HEYMAN, R.A., D.J.MANGELSDORF, J.A.DYCK, R.B.STEIN, G.EICHELE, R.M.EVANS and C.THALLER. (1992). 9-Cis Retinoic Acid is a high affinity ligand for the retinoid X receptor.  
Cell 68 : 397-406.
- HIGGINS, D.G. and P.M.SHARP. (1988). Clustal : a package for performing multiple sequence alignment on a microcomputer.  
Gene 73 : 237-244.

- HODIN,R.A., M.A.LAZAR, B.I.WINTMAN, D.S.DARLING, R.J.KOENIG, P.R.LARSEN, D.D.MOORE and W.W.CHIN. (1989). Identification of a thyroid hormone receptor that is pituitary-specific.  
Science **244** : 76-79.
- HODIN,R.A., M.A.LAZAR and W.W.CHIN. (1990). Differential and tissue-specific regulation of the multiple rat *c-erbA* messenger RNA species by thyroid hormone.  
J. Clin Invest. **85** : 101-105.
- HOFFMAN,B., J.M.LEHMANN, X-K.ZHANG, T.HERMANN, M.HUSMANN, G.GRAUPNER and M.PFAHL. (1990). A retinoic acid receptor-specific element controls the retinoic acid receptor- $\beta$  promoter.  
Mol. Endo. **4** : 1727-1736.
- HOFFMAN,M. (1990). The embryo takes its vitamins.  
Science **250** : 372-373.
- HOFFMANN,J.A. and J.JOOSSE. (1991). Chinless wonders.  
The New Biol. **3** : 336-340.
- HOLLENBERG,S.M., C.WEINBERGER, E.S.ONG, G.CERELLI, A.ORO, R.LEBO, E.B.THOMPSON, M.G.ROSENFELD and R.M.EVANS. (1985). Primary structure and expression of a functional human glucocorticoid receptor cDNA.  
Nature **318** : 635-641.
- HOLLENBERG,S.M., V.GIGUERE, P.SEQUI and R.M.EVANS.(1987). Colocalization of DNA binding and transcriptional activation functions in the human glucocorticoid receptor.  
Cell **49** : 39-46.
- HOLLENBERG,S.M. and R.M.EVANS. (1988). Multiple and cooperative trans-activation domains of the human glucocorticoid receptor.  
Cell **55** : 899-906.
- HOPE,T.J., X.HUANG, D.McDONALD and T.G.PARSLOW. (1990). Steroid-receptor fusion of the human immunodeficiency virus type 1 rev transactivator : mapping cryptic functions of the arginine-rich motif.  
Proc. Natl. Acad. Sci USA **87** : 7787-7791.
- HOPNER,W., U.B.RASMUSSEN, G.ABUERREISH, H.WOHLRAB and H.J.SEITZ. (1988). Thyroid hormone effect on gene expression of the adenine nucleotide translocase in different rat tissues.  
Mol. Endo. **2** : 1127-1131.
- HOROWITZ,Z.D., C.R.YANG, B.M.FORMAN, J.CASANOVA and H.H.SAMUELS. (1989). Characterization of the domain structure of chick *c-erbA* by deletion mutation : *in vitro* translation and cell transfection studies.  
Mol. Endo. **3** : 148-156.
- HOURDRY,J. et A.BEAUMONT. (1985). Les métamorphoses des amphibiens. Fondation Singer-Polignac : 1-273.
- HOUSELY,P.R., E.R.SANCHEZ, H.M.WESTPHAL, M.BEATO and W.B.PRATT. (1985). The molybdate-stabilized L-cell glucocorticoid receptor isolated by affinity chromatography or with a monoclonal antibody is associated with a 90-92 kDa non-steroid binding phosphoprotein.  
J.Biol.Chem. **260** : 13810-13817.
- HOUSTEK,J., S.PAVELKA, M.BAUDYSOVA and J.KOPECKY. (1990). Induction of type II iodothyronine 5'-deiodinase and mitochondrial uncoupling protein in brown adipocytes differentiated in cell culture.  
FEBS Lett. **274** : 185-188.

- HOWARD, K.J., S.J.HOLLEY, K.R.YAMAMOTO and C.W.DISSELHORST. (1990). Mapping the Hsp90 binding region of the glucocorticoid receptor. *J. Biol. Chem.* **265** : 11928-11935.
- HUANG, M., H.C.YE, S.R.CHEN, J.R.CHAI, J.X.LU, L.ZHOA, L.J.GU and Z.Y.WANG. (1988). Use of all-trans retinoic acid in the treatment of acute promyelocytic leukemia. *Blood* **72** : 567-572.
- HUCKABY, C.S., O.M.CONNELLY, W.G.BEATTIE, A.D.W.DOBSON, M.J.TSAI and B.W.O'MALLEY. (1987). Structure of the chromosomal chicken progesterone receptor gene. *Proc. Natl. Acad. Sci. USA* **84** : 8380-8384.
- HUDSON, L.G., J.B.SANTON, C.K.GLASS and G.N.GILL. (1990). Ligand-activated thyroid hormone and retinoic acid receptors inhibit growth factor receptor promoter expression. *Cell* **62** : 1165-1175.
- HUGHES, M.R., P.J.MALLOY, D.G.KIEBACK, R.A.KESTERSON, J.WESLEY PIKE, D.FELDMAN and B.W.O'MALLEY. (1988). Point mutations in the human vitamin D receptor gene associated with hypocalcemic rickets. *Science* **242** : 1702-1705.
- HUGHES, M.R. and B.W.O'MALLEY. (1991). Genetic defects of receptors involved in disease. In *Nuclear Hormone Receptors*. Ed. Parker, M.G. Academic Press : 321-348.
- HUSMANN, M., J.LEHMANN, B.HOFFMANN, T.HERMANN, M.TZUKERMAN and M.PFAHL. (1991). Antagonism between retinoic acid receptors. *Mol. Cell. Biol.* **11** : 4097-4103.
- HWUNG, Y.P., D.T.CROWE, L.H.WANG, S.Y.TSAI and M.J.TSAI. (1988). The COUP transcription factor binds to an upstream promoter element of the rat insulin II gene. *Mol. Cell. Biol.* **8** : 2070-2077.
- ICHIKAWA, K. and L.J. DeGROOT. (1987). Purification and characterization of rat liver nuclear thyroid hormone receptors. *Proc. Natl. Acad. Sci. USA* **84** : 3420-3424.
- INGBAR, S.H. (1985). The thyroid gland. In *Williams Textbook of Endocrinology*. Seventh Edition. Eds. J.D.Wilson and D.W.Foster : 682-817.
- INGRAHAM, H.A., R.CHEN, H.J.MANGALAM, H.P.ELSHOLTZ, S.E.FLYNN, C.R.LIN, D.M.SIMMONS, L.SWANSON and M.G.ROSENFELD. (1988). A tissue-specific transcription factor containing a homeodomain specifies a pituitary phenotype. *Cell* **55** : 519-529.
- INGRAHAM, H.A., S.E.FLYNN, J.W.VOSS, V.R.ALBERT, M.S.KAPILOFF, L.WILSON and M.G.ROSENFELD. (1990). The POU-specific domain of Pit-1 is essential for sequence-specific, high affinity DNA binding and DNA-dependent Pit-1/Pit-1 interactions. *Cell* **61** : 1021-1033.
- INSALL, R. and R.R.KAY. (1990). A specific DIF binding protein in *Dictyostelium*. *The EMBO J.* **9** : 3323-3328.
- ISSEMANN, I. and S.GREEN. (1990). Activation of a member of the steroid hormone receptor superfamily by peroxisome proliferators. *Nature* **347** : 645-650.
- IZPISUA-BELMONTE, J.C., C.TICKLE, P.DOLLE, L.WOLPERT and D.DUBOULE (1991). Expression of the homeobox Hox-4 genes and the specification of position in chick wing development. *Nature* **350** : 585-589.

- IZUMO,S., B.NADAL-GINARD and V.MAHDAVI. (1986). All members of the myosin heavy chain multigene family respond to thyroid hormone in a highly tissue specific manner. *Science* **231** : 597-600.
- IZUMO,S and V.MAHDAVI. (1988). Thyroid hormone receptor  $\alpha$  isoforms generated by alternative splicing differentially activate myosin HC gene transcription. *Nature* **334** : 539-543.
- IZUMO,S., B.NADAL-GINARD and V.MAHDAVI. (1990). The thyroid hormone receptor  $\alpha$  gene generates functionally different protein isoforms by alternative splicing. *Mol. Biol. of the Cardiovascular. Syst.* 111-123.
- JACKSON,R.J. and N.STANDART. (1990). Do the poly(A) tail and 3' untranslated region control mRNA translation ? *Cell* **62** : 15-24.
- JACOBS,S. and C.MOXHAM. (1991). Hybrid receptors. *The New Biol.* **3** : 110-115.
- JANSSON,M., L.PHILIPSON and B.VENNSTROM. (1983). Isolation and characterization of multiple human genes homologous to the oncogenes of avian erythroblastosis virus. *The EMBO J.* **2** : 561-565.
- JELTSCH,J.M., Z.KROZOWSKI, C.QUIRIN-STRICKER, H.GRONEMEYER, R.J.SIMPSON, J.M.GARNIER, A.KRUST, F.JACOB and P.CHAMBON. (1986). Cloning of the chicken progesterone receptor. *Proc. Natl. Acad. Sci. USA* **83** : 5424-5428.
- JENSEN,E.V., T.SUZUKI, T.KAWASHIMA, W.E.STUMPF, P.W.JUNGBLUT and E.R.DESOMBRE. (1968). A two-step mechanism for the interaction of estradiol with rat uterus. *Proc. Natl. Acad. Sci. USA* **59** : 632-638.
- JENSEN,E.V. (1991). Overview of the nuclear receptor family. In *Nuclear Hormone Receptors*. Ed.Parker,M.G. Academic Press : 1-10.
- JOHNSON,E.F. (1990). A partnership between the dioxin receptor and a basic helix-loop-helix protein. *Science* **252** : 924-925.
- 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.
- JOST,A. (1991). Les péripéties d'une recherche : l'étude de la différenciation sexuelle. *La Vie des Sciences, Comptes rendus*, **8** : 67-87.
- KAHN,P., L.FRYKBERG, C.BRADY, I.STANLEY, H.BEUG, B.VENNSTRÖM and T.GRAF. (1986). *v-erbA* cooperates with sarcoma oncogenes in leukemic cell transformation. *Cell* **45** : 349-356.
- KAKIZUKA,A., W.H.MILLE, Jr, K.UMESOMO, R.P.WARRELL,Jr., S.R.FRANKEL, V.V.V.S.MURTY, E.DMITROVSKY and R.M.EVANS. (1991). Chromosomal translocation t(15;17) in human acute promyelocytic leukemia fuses RAR $\alpha$  with a novel putative transcription factor, PML. *Cell* **66** : 663-674.
- KALDERON,D., B.L.ROBERTS, W.D.RICHARDSON and A.E.SMITH. (1984). A short amino acid sequence able to specify nuclear location. *Cell* **39** : 499-509.

- KALFF, M., B. GROSS and M. BEATO. (1990). Progesterone receptor stimulates transcription of Mouse Mammary Tumour Virus in a cell free system.  
Nature **344** : 360-362.
- KAPTEIN, R. (1991). Distinguishing features.  
Current Biol. **1** : 336-338.
- KARIN, M., J.-L. CASTRILLO and L.E. THEILL. (1990). Growth hormone gene regulation : a paradigm for cell-type-specific gene activation.  
TIG **6** : 92-96.
- KASTNER, P., A. KRUST, C. MENDELSON, J.M. GARNIER, A. ZELENT, P. LEROY, A. STAUB and P. CHAMBON. (1990). Murine isoforms of retinoic acid receptor  $\gamma$  with specific patterns of expression.  
Proc. Natl. Acad. Sci. USA **87** : 2700-2704.
- KASTNER, P., A. KRUST, B. TURCOTTE, U. STROPP, L. TORA, H. GRONEMEYER and P. CHAMBON. (1990). Two distinct estrogen-regulated promoters generate transcripts encoding the two functionally different human progesterone receptor forms A and B.  
The EMBO J. **9** : 1603-1614.
- KASTNER, P., A. PEREZ, Y. LUTZ, C. ROCHETTE-EGLY, M.P. GAUB, B. DURAND, M. LANOTTE, R. BERGER and P. CHAMBON. (1992). Structure, localization and transcriptional properties of two classes of retinoic acid receptor  $\alpha$  fusion proteins in acute promyelocytic leukemia (APL) : structural similarities with a new family of oncoproteins.  
The EMBO J. **11** : 629-642.
- KATZENELLENBOGEN, B.S., J.F. ELLISTON, F.J. MONSMA, P.A. SPRINGER, Y.S. ZIEGLER and G.L. GREENE. (1987). Structural analysis of covalently labeled estrogen receptors by limited proteolysis and monoclonal antibody reactivity.  
Biochemistry **26** : 2364-2373.
- KAWAHARA, A., B.S. BAKER and J.R. TATA. (1991). Developmental and regional expression of thyroid hormone receptor genes during *Xenopus* metamorphosis.  
Development **112** : 933-943.
- KESSEL, M. and P. GRUSS. (1990). Murine developmental control genes.  
Science **249** : 374-379.
- KHAZAIE, K., T.J. DULL, T. GRAF, J. SCHLESSINGER, A. ULLRICH, H. BEUG and B. VENNSTRÖM. (1988). Truncation of the human EGF receptor leads to differential transforming potentials in primary avian fibroblasts and erythroblasts.  
The EMBO J. **7** : 3061-3171.
- KHAZAIE, K., G. PANAYOTOU, A. AGUZZI, J. SAMARUT, L. GAZZOLO and P. JURDIC. (1991). EGF promotes *in vivo* tumorigenic growth of primary chicken embryo fibroblasts expressing *v-myc* and enhances *in vitro* transformation by the *v-erbA* oncogene.  
Oncogene **6** : 21-28.
- KLEIN-HITPASS, L., M. SCHORPP, U. WAGNER and G.U. RYFFEL. (1986). An estrogen-responsive element derived from the 5' flanking region of the *Xenopus* vitellogenin A2 gene functions in transfected human cells.  
Cell **46** : 1053-1061.
- KLIEWER, S.A., K. UMESONO, D.J. MANGELSDORF and R.M. EVANS. (1992). Retinoid X receptor interacts with nuclear receptors in retinoic acid, thyroid hormone and vitamin D3 signalling.  
Nature **355** : 446-449/

- KLINGENBERG, M. (1990). Mechanism and evolution of the uncoupling protein of brown adipose tissue.  
TIBS 15 : 108-112.
- KLOCK, G., U.STRÄHLE and G.SCHÜTZ. (1987). Estrogen and glucocorticoid responsive elements are closely related but distinct.  
Nature 339 : 734-736.
- KLUG, A. and D.RHODES. (1987). "Zinc fingers" : a novel protein motif for nucleic acid recognition.  
TIBS 12 : 464-469.
- KOELLE, M.R., W.S.TALBOT, W.A.SEGRAVES, M.T.BENDER, P.CHERBAS and D.S.HOGNESS. (1991). The *Drosophila* EcR gene encodes an ecdysone receptor, a new member of the steroid receptor superfamily.  
Cell 67 : 59-77.
- KOENIG, R.J., G.A.BRENT, R.L.WARNE, P.R.LARSEN and D.D.MOORE. (1987). Thyroid hormone receptor binds to a site in the rat growth hormone promoter required for induction by thyroid hormone.  
Proc.Natl.Acad.Sci..USA 84 : 5670-5674.
- KOENIG, R.J., R.L.WARNE, G.A.BRENT, J.W.HARNEY, P.R.LARSEN and D.D.MOORE. (1988). Isolation of a cDNA clone encoding a biologically active thyroid hormone receptor.  
Proc. Natl. Acad. Sci. USA 85 : 5031-5035.
- KOENIG, R.J., M.A.LAZAR, R.A.HODIN, G.A.BRENT, P.R.LARSEN, W.W.CHIN and D.D.MOORE. (1989). Inhibition of thyroid hormone action by a non-hormone binding c-erbA protein generated by alternative mRNA splicing.  
Nature 337 : 659-661.
- KOURIDES, I.A.J.A.GURR and O.WOLF. (1984). The regulation and organization of thyroid stimulating hormone genes.  
Recent Prog. Horm.Res. 40 : 79-120.
- KOURILSKY, P. (1986). Molecular mechanisms for gene conversion in higher cells.  
TIG 2 : 60-63.
- KOZAK, M. (1986). Bifunctional messenger RNAs in eukaryotes.  
Cell 47 : 481-483.
- KRANE, I.M., E.R.SPINDEL and W.W.CHIN. (1991). Thyroid hormone decreases the stability and the poly(A) tract length of rat thyrotropin  $\beta$ -subunit messenger RNA.  
Mol. Endo. 5 : 469-475.
- KRIEGER, D.T., J.ASCHOFF. (1979). Endocrine and other biological rhythms. In : DeGroot, L.J., G.F.Cahil Jr, L.Martini et al., eds. Endocrinology. Vol 3. New York : Grune and Stratton : 2079-2109.
- KRUST, A., S.GREEN, P.ARGOS, V.KUMAR, P.WALTER, J.M.BORNERT and P.CHAMBON. (1986). The chicken estrogen receptor sequence : homology with v-erbA and the human oestrogen and glucocorticoid receptors.  
The EMBO J. 5 : 891-897.
- KRUST, A., P.KASTNER, M.PETKOVICH, A.ZELEN and P.CHAMBON. (1989). A third human retinoic acid receptor, hRAR- $\gamma$ .  
Proc. Natl. Acad.Sci. USA 86 : 5310-5314.

- KUIPER, G.G.J.M., P.W.FABER, H.C.J.VAN ROOIJ, J.A.G.M. van der KORPUT, C.RIS-STALPERS, P.KLAASSEN, J.TRAPMAN and A.O.BRINKMANN. (1989). Structural organization of the human androgen receptor gene.  
J. Mol. Endo. 2 : R1-R4.
- KUMAR, V., S.GREEN, A.STAUB and P.CHAMBON. (1986). Localisation of the oestradiol-binding and putative DNA-binding domains of the human estrogen receptor.  
The EMBO J. 5 : 2231-2236.
- KUMAR, V., S.GREEN, G.STACK, M.BERRY, J.R.JIN and P.CHAMBON. (1987). Functional domains of the human estrogen receptor.  
Cell 51 : 941-951.
- KUMAR, V. and P.CHAMBON. (1988). The estrogen receptor binds tightly to its responsive element as a ligand-induced homodimer.  
Cell 55 : 145-156.
- LADIAS, J.A.A. and S.K.KARATHANASIS. (1991). Regulation of the apolipoprotein AI gene by ARP-1, a novel member of the steroid receptor superfamily.  
Science 251 : 561-565.
- LANDSCHULZ, W.H., 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.
- LAROSA, G.J. and L.J.GUDAS. (1988a). An early effect of retinoic acid : cloning of an mRNA (Era-1) exhibiting rapid and protein synthesis-independent induction during teratocarcinoma stem cell differentiation.  
Proc. Natl. Acad. Sci. USA 85 : 329-333.
- LAROSA, G.J. and L.J.GUDAS. (1988b). Early retinoic acid-induced F9 teratocarcinoma stem cell gene ERA-1 : alternative splicing creates transcripts for a homeobox-containing protein and one lacking the homeobox.  
Mol. Cell. Biol. 8 : 3906-3917.
- LARSEN, P.R., J.W.HARNEY and D.D.MOORE. (1986). Sequences required for cell-type specific thyroid hormone regulation of rat growth hormone promoter activity.  
J.Biol.Chem. 261 : 14373-14376.
- LATHAM, K.R., J.C.RING and J.D.BAXTER. (1976). Solubilized nuclear "receptors" for thyroid hormones.  
J. Biol. Chem. 251 : 7388-7397.
- LAVORGNA, G., H.UEDA, J.CLOS and C.WU. (1991). FTZ-F1, a steroid hormone receptor-like protein implicated in the activation of *fushi tarazu*.  
Science 252 : 848-851.
- LAZAR, M.A. and W.W.CHIN. (1988). Regulation of two c-erbA messenger ribonucleic acids in rat GH3 cells by thyroid hormone.  
Mol. Endo. 2 : 479-484.
- LAZAR, M.A., R.A.HODIN, D.S.DARLING and W.W.CHIN. (1988). Identification of a rat c-erbA $\alpha$ -related protein which binds deoxyribonucleic acid but does not bind thyroid hormone.  
Mol. Endo. 2 : 893-901.
- LAZAR, M.A., R.A.HODIN, D.S.DARLING and W.W.CHIN. (1989a). A novel member of the thyroid/steroid hormone receptor family is encoded by the opposite strand of the rat c-erbA $\alpha$  transcriptional unit.  
Mol. Cell. Biol. 9 : 1128-1136.

- LAZAR, M.A., R.A.HODIN and W.W.CHIN. (1989b). Human carboxyl-terminal variant of  $\alpha$ -type c-erbA inhibits trans-activation by thyroid hormone receptors without binding thyroid hormone.  
Proc. Natl. Acad. Sci. USA **86** : 7771-7774.
- LAZAR, M.A., K.E.JONES and W.W.CHIN. (1990a). Isolation of a cDNA encoding human Rev-erbA $\alpha$  : transcription from the noncoding DNA strand of a thyroid hormone receptor gene results in a related protein that does not bind thyroid hormone.  
DNA and Cell Biol. **9** : 77-83.
- LAZAR, M.A. and T.J.BERRODIN. (1990). Thyroid hormone receptors form distinct nuclear protein-dependent and independent complexes with a thyroid hormone response element.  
Mol. Endo **4** : 1627-1634.
- LAZAR, M.A., R.A.HODIN, G.CARDONA and W.W.CHIN. (1990b). Gene expression from the c-erbA $\alpha$ /rev-erbA $\alpha$  genomic locus.  
J. Biol. Chem. **265** : 12859-12863.
- LAZAR, M.A., T.J.BERRODIN and H.P.HARDING. (1991). Differential DNA binding by monomeric, homodimeric, and potentially heteromeric forms of the thyroid hormone receptor.  
Mol. Cell. Biol. **11** : 5005-5015.
- LEE, M.S., G.P.GIPPERT, K.V.SOMAIN, D.A.CASE and P.E.WRIGHT. (1989). Three-dimensional solution structure of a single zinc finger DNA-binding domain.  
Science **245** : 635-636.
- LEES, J.A., S.E.FAWELL and M.G.PARKER. (1989). Identification of two transactivation domains in the mouse estrogen receptor.  
Nuc. Acids. Res. **17** : 5477-5488.
- LEFEBVRE, P., D.S.BERARD, M.G.CORDINGLEY and G.L.HAGER. (1991). Two regions of the Mouse Mammary Tumor Virus long terminal repeat regulate the activity of its promoter in mammary cell lines.  
Mol. Cell. Biol. **11** : 2529-2537.
- LEHMANN, J.M., B.HOFFMANN and M.PFAHL. (1991). Genomic organization of the retinoic acid receptor  $\gamma$  gene.  
Nuc. Acids. Res. **19** : 573-578.
- LEID, M., P.KASTNER, R.LYONS, H.NAKSHATRI, M.SAUNDERS, T.ZACHAREWSKI, J.Y.CHEN, A.STAUB, J.M.GARNIER, S.MADER and P.CHAMBON. (1992). Purification, cloning, and RXR identity of the HeLa cell factor with which RAR or TR heterodimerizes to bind target sequences efficiently.  
Cell **68** : 377-395.
- LEROY, P., A.KRUST, A.ZESENT, C.MENDELSON, J-M.GARNIER, P.KASTNER, A.DIERICH and P.CHAMBON. (1991a). Multiple isoforms of the mouse retinoic acid receptor  $\alpha$  are generated by alternative splicing and differential induction by retinoic acid.  
The EMBO J. **10** : 59-69.
- LEROY, P., H.NAKSHATRI and P.CHAMBON. (1991b). Mouse retinoic acid receptor  $\alpha$ 2 isoform is transcribed from a promoter that contains a retinoic acid response element.  
Proc. Natl. Acad. Sci. USA **88** : 10138-10142.
- LEVIN, A.A., L.J.STURZENBECKER, S.KAZMER, T.BOSAKOWSKI, C.HUSELTON, G.ALLENBY, J.SPECK, C.KRATZEISEN, M.ROSENBERGER, A.LOVEY and J.F.GRIPPO. (1992). 9-Cis retinoic acid stereoisomer binds and activates the nuclear receptor RXR $\alpha$ .  
Nature **355** : 359-361.

- LEVINE, M. and J.L.MANLEY. (1989). Transcriptional repression of eukaryotic promoters. *Cell* **59** : 405-408.
- LI, S., E.B.CRENSHAW III, E.J.RAWSON, D.M.SIMMONS, L.W.SWANSON and M.G.ROSENFELD. (1990). Dwarf locus mutants lacking three pituitary cell types result from mutations in the POU-domain gene *pit-1*. *Nature* **347** : 528-533.
- LIN, K.-H., C.PARKISON, P.McPHIE and S.-Y.CHENG. (1991). An essential role of domain D in the hormone-binding activity of human  $\beta 1$  thyroid hormone nuclear receptor. *Mol. Endo.* **5** : 485-492.
- LISSITZKY, S. (1990). Thyroid hormones. In *Hormones from molecules to disease*. Eds. E.E.Baulieu and P.A.Kelly. Published by Hermann, Publishers in arts and science, Paris : 343-381.
- LONGO, L., E.DONTI, A.MENCARELLI, G.AVANZI, L.PEGORARO, G.ALIMENA, A.TABILIO, G.VENTI, F.GRIGNANI and P.G.PELICCI. (1990). Mapping of chromosome 17 breakpoints in acute myeloid leukemias. *Oncogene* **5** : 1557-1563.
- LUBAHN, D.B., T.R.BROWN, J.A.SIMENTAL, H.N.HIGGS, C.J.MIGEON, E.M.WILSON and F.S.FRENCH. (1989). Sequence of the intron/exon junctions of the coding regions of the human androgen receptor gene and identification of a point mutation in a family with complete androgen insensitivity. *Proc.Natl.Acad.Sci. USA* **86** : 9534-9538.
- LUCAS, P.C., B.M.FORMAN, H.H.SAMUELS and D.K.GRANNER. (1991). Specificity of a retinoic acid response element in the phosphoenolpyruvate carboxykinase gene promoter : consequences of both retinoic acid and thyroid hormone receptor binding. *Mol. Cell. Biol.* **11** : 5164-5170.
- 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*. *The 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.
- MACMORRIS, T.C. and A.W.BARKSDALE. (1967). Isolation of a sex hormone from the water mold *Achlya ambisexualis*. *Nature* **215** : 320-321.
- MADEN, M. (1982). Vitamin A and pattern formation in the regenerating limb. *Nature* **295** : 672-675.
- MADER, S., V.KUMAR, H. DE VERNEUIL and P.CHAMBON. (1989). Three amino acids of estrogen receptor are essential to its ability to distinguish an estrogen from a glucocorticoid-responsive element. *Nature* **338** : 271-274.
- MANGELSDORF, D.J., E.S.ONG, J.A.DYCK and R.M.EVANS. (1990). Nuclear receptor that identifies a novel retinoic acid response pathway. *Nature* **345** : 224-229.

- MANGELSDORF, D.J., K. UMESONO, S.A. KLIEWER, U. BORGMEYER, E.S. ONG and R.M. EVANS. (1991). A direct repeat in the cellular retinol-binding protein type II gene confers differential regulation by RXR and RAR.  
Cell 66 : 555-561.
- MANIATIS, T. (1991). Mechanisms of alternative Pre-mRNA splicing.  
Science 251 : 33-34.
- MARTINEZ, E., F. GIVEL and W. WAHLI. (1987). The estrogen-responsive element as an inducible enhancer : DNA sequence requirements and conversion to a glucocorticoid-responsive element.  
The EMBO J. 6 : 3719-3727.
- MARTINEZ, E. and W. WAHLI. (1989). Cooperative binding of estrogen receptor to imperfect estrogen-responsive DNA elements correlates with their synergistic hormone-dependent enhancer activity.  
The EMBO J. 8 : 3781-3791.
- MARTINEZ, E., Y. DUSSERE, W. WAHLI and N. MERMOD. (1991a). Synergistic transcriptional activation by CTF/NF-1 and the estrogen receptor involves stabilized interactions with a limiting target factor.  
Mol. Cell. Biol. 11 : 2937-2945.
- MARTINEZ, E., F. GIVEL and W. WAHLI. (1991b). A common ancestor DNA motif for invertebrate and vertebrate hormone response elements.  
The EMBO J. 10 : 263-268.
- MARTINEZ, E. and W. WAHLI. (1991). Characterization of hormone response elements. In Nuclear Hormone Receptors. Ed. Parker, M.G. Academic Press : 125-146.
- MARTINS, V.R., W.B. PRATT, L. TERRACIO, M.A. HIRST, G.M. RINGOLD and P.R. HOUSLEY. (1991). Demonstration by confocal microscopy that unliganded overexpressed glucocorticoid receptors are distributed in a nonrandom manner throughout all planes of the nucleus.  
Mol. Endo. 5 : 217-225.
- MASUDA, M., S. YASUHARA, M. YAMASHITA, M. SHIBUYA and T. ODAKA. (1990). Nucleotide sequence of the murine thyroid hormone receptor ( $\alpha$ 1) cDNA.  
Nuc. Acids Res. 18 : 3055.
- MATHIEU-MAHUL, D., D.Q. XU, S. SAULE, R. LIDEREAU, F. GALIBERT, R. BERGER, M. MAUCHAUFFE and C.J. LARSEN. (1985). An EcoRI restriction fragment length polymorphism (RFLP) in the human *c-erbA* locus.  
Hum. Gen. 71 : 41-44.
- MATHISEN, P.M. and L. MILLER. (1989). Thyroid hormone induces constitutive keratin gene expression during *Xenopus laevis* development.  
Mol. Cell. Biol. 9 : 1823-1831.
- MENDEL, D.B., J.E. BODWELL, B. GAMETCHU, R.W. HARRISON and A. MUNCK. (1986). Molybdate-stabilized nonactivated glucocorticoid-receptor complexes contain a 90-kDa non-steroid-binding phosphoprotein that is lost on activation.  
J. Biol. Chem. 261 : 3758-3763.
- METZGER, D., J.H. WHITE and P. CHAMBON. (1988). The human estrogen receptor functions in yeast.  
Nature 334 : 31-36.

- MIYAJIMAN, N., R. HORIUCHI, Y. SHIBUYA, S.-I. FUKUSHIGE, K.-I. MATSUBARA, K. TOYOSHIMA and T. YAMAMATO. (1989). Two erbA homologs encoding proteins with different T3 binding capacities are transcribed from opposite DNA strands of the same genetic locus.  
Cell 57 : 31-39.
- MLODZIK, M., Y. HIROMI, U. WEBER, C.S. GOODMAN and G.M. RUBIN. (1990). The drosophila seven-up gene, a member of the steroid receptor gene superfamily, controls photoreceptor cell fates.  
Cell 60 : 211-224.
- MOORE, D.D., A.R. MARKS, D.I. BUCKLEY, G. KAPLER, F. PAYVAR and H.M. GOODMAN. (1985). The first intron of the human growth hormone gene contains a binding site for glucocorticoid receptor.  
Proc. Natl. Acad. Sci. USA 82 : 699-702.
- MOORE, D.D. (1989). Promiscuous behaviour in the steroid hormone receptor superfamily.  
TINS 12 : 165-168.
- MOORE, D.D. (1990). Diversity and unity in the nuclear hormone receptors : a terpenoid receptor superfamily.  
The New Biol. 2 : 100-105.
- MOORE, D.D. and G.A. BRENT. (1991). Thyroid hormone : half-sites and insights.  
The New Biol. 3 : 835-844.
- MOREL, Y., J. BRISTOW, S.E. GITELMAN and W.L. MILLER. (1989). Transcript encoded on the opposite strand of the human steroid 21-hydroxylase/complement component C4 gene locus.  
Proc. Natl. Acad. Sci. USA 86 : 6582-6586.
- MORRISON, N.A., J. SHINE, J.C. FRAGONAS, V. VERKEST, M.L. McMENEMY and J.A. EISMAN. (1989). 1,25-Dihydroxyvitamin D3 responsive elements and glucocorticoid repression in the osteocalcin gene.  
Science 246 : 1158-1161.
- MOSCOVICI, M.G., M.L. SIEGEL and C. MOSCOVICI. (1989). Avian erythroblastosis virus transforms a novel mast cell-basophil precursor target in the japanese quail.  
J. Virol. 63 : 2335-2339.
- MOUDGIL, V.K. (1988). Steroid receptors in health and disease.  
FEBS Let. 226 : 213-216.
- MULLER, M., C. BANIAHMAD, C. KALTSCHMIDT, R. SCHULE and R. RENKAWITZ. (1991). Cooperative transactivation of steroid receptors. In Nuclear Hormone Receptors. Ed. Parker, M.G. Academic Press : 155-170.
- MULVIHILL, E.R., J.-P. LEPENNEC and P. CHAMBON. (1982). Chicken oviduct progesterone receptor : location of specific regions of high-affinity binding in cloned DNA fragments of hormone-responsive genes.  
Cell 24 : 621-632.
- MUNOZ, A., M. ZENKE, U. GEHRING, J. SAP, H. BEUG and B. VENNSTRÖM. (1988). Characterization of the hormone-binding domain of the chicken c-erbA/thyroid hormone receptor protein.  
The EMBO J. 7 : 155-159.
- MUNOZ, A., W. HOPFNER, J. SAP, G. BRADY, K. NORDSTROM, H.J. SEITZ and B. VENNSTRÖM. (1990). The chicken c-erbA  $\alpha$ -product induces expression of thyroid hormone-responsive genes in 3,5,3'-triiodothyronine receptor-deficient rat hepatoma cells.  
Mol. Endo. 4 : 312-320.

- MURATA, Y., H. SEO, K. SEKIGUCHI, T. IMAI, J. LEE and N. MATSUI. (1990). Specific induction of fibronectin gene in rat liver by thyroid hormone.  
Mol. Endo. 4 : 693-699.
- MURRAY, M.B., N.D. ZILZ, N.L. Mc CREARY, M.J. MAC DONALD and H.C. TOWLE. (1988). Isolation and characterization of rat cDNA clones for two distinct thyroid hormone receptors.  
J. Biol. Chem. 263 : 12770-12777.
- NÄÄR, A.M., J.M. BOUTIN, S.M. LIPKIN, V.C. YU, J.M. HOLLOWAY, C.K. GLASS and M.G. ROSENFELD. (1991). The orientation and spacing of core DNA-binding motifs dictate selective transcriptional responses to three nuclear receptors.  
Cell 65 : 1267-1279.
- NAKAI, A., Z. SAKURAI, G.I. BELL and L.J. DE GROOT. (1988). Characterization of a third human thyroid hormone receptor coexpressed with other thyroid hormone receptors in several tissues.  
Mol. Endo 2 : 1087-1092.
- NAKAI, A., S. SEINO, A. SAKURAI, I. SZILAK, G.I. BELL and L.J. DE GROOT. (1988). Characterization of a thyroid hormone receptor expressed in human kidney and other tissues.  
Proc. Natl. Acad. Sci. USA 85 : 2781-2785.
- NAPIER, R.M. and M.A. VENIS. (1991). From auxin-binding protein to plant hormone receptor  
TIBS 16 : 72-78
- NAUBER, U., M.J. PANKRATZ, A. KIENLIN, E. SEIFERT, U. KLEMM and H. JÄCKLE. (1988). Abdominal segmentation of the *Drosophila* embryo requires a hormone receptor-like protein encoded by the gap gene knirps.  
Nature 336 : 489-492.
- NEMOTO, T., Y. OHARA-NEMOTO, M. DENIS and J.A. GUSTAFSSON. (1990). The transformed glucocorticoid receptor has a lower steroid-binding affinity than the nontransformed receptor.  
Biochemistry 29 : 1880-1886.
- NICHOLSON, R.C., S. MADER, S. NAGPAL, M. LEID, C. ROCHETTE-EGLY and P. CHAMBON. (1990). Negative regulation of the rat stromelysin gene promoter by retinoic acid is mediated by an AP1 binding site.  
The EMBO J. 9 : 4443-4454.
- NIRDE, P., G. TORPIER, M.L. DE REGGI and A. CAPRON. (1983). Ecdysone and 20 hydroxyecdysone : new hormones for the human parasite *Schistosoma mansoni*.  
FEBS Lett. 151 : 223-227.
- NOJI, S., T. NOHNO, E. KOYAMA, K. MUTO, K. OHYAMA, Y. AOKI, K. TAMURA, K. OHSUGI, H. IDE, S. TANIGUCHI and T. SAITO. (1991). Retinoic acid induces polarizing activity but is unlikely to be a morphogen in the chick limb bud.  
Nature 350 : 83-86.
- O'DONNELL, A.L. and R.J. KOENIG. (1990). Mutational analysis identifies a new functional domain of the thyroid hormone receptor.  
Mol. Endo. 4 : 715-720.
- O'DONNELL, A.L., E.D. ROSEN, D.S. DARLING and R.J. KOENIG. (1991). Thyroid hormone receptor mutations that interfere with transcriptional activation also interfere with receptor interaction with a nuclear protein.  
Mol. Endo. 5 : 94-99.

- O'KANE,C. and W.J.GEHRING. (1987). Detection in situ of genomic regulatory elements in *Drosophila*.  
Proc. Natl. Acad. Sci. USA 84 : 9123-9127.
- OLIVER,G., C.V.E.WRIGHT, J.HARDWICKE and E.M.DE ROBERTIS. (1988). A gradient of homeodomain protein in developing forelimbs of *Xenopus* and mouse embryos.  
Cell 55 : 1017-1024.
- OLIVER,G., N.SIDELL, W.FISKE, C.HEINZMANN, T.MOHANDAS, R.S.SPARKES and E.M.DE ROBERTIS. (1989). Complementary homeo protein gradients in developing limb buds.  
Genes Dev. 3 : 641-650.
- OLIVIER,G., E.M.DE ROBERTIS, L.WOLPERT and C.TICKLE. (1990). Expression of a homeobox gene in the chick wing bud following application of retinoic acid and grafts of polarizing region tissue.  
The EMBO J. 9 : 3093-3099.
- O'MALLEY.W. (1989). Editorial : did eucaryotic steroid receptors evolve from intracrine gene regulators ?  
Endocrinology 125 : 1119-1120.
- OPPENHEIMER,J.H. (1979). Thyroid hormone action at the cellular level.  
Science 203 : 971-979.
- OPPENHEIMER,J.H., H.L.SCHWARTZ, C.N.MARIASH, W.B.KINLAW, N.C.W.WONG and H.C.FREAKE. (1987). Advances in our understanding of thyroid hormone action at the cellular level.  
Endocr. Rev. 8 : 288-308.
- ORO,A.E., E.S.ONG, J.S.MARGOLIS, J.W.POSAKONY, M.McKEOWN and R.M.EVANS. (1988). The *Drosophila* gene knirps-related is a member of the steroid-receptor gene superfamily.  
Nature 336 : 493-496.
- ORO,A.E., M.McKEOWN and R.M.EVANS. (1990). Relationship between the product of the *Drosophila* ultraspiracle locus and the vertebrate retinoid X receptor.  
Nature 347 : 298-301.
- PAIN,B., F.MELET, P.JURDIC and J.SAMARUT. (1990). The carbonic anhydrase II gene, a gene regulated by thyroid hormone and erythropoietin, is repressed by the v-erbA oncogene in erythrocytic cells.  
The New Biol. 2 : 1-11.
- PAIN,B., C.M.WOODS, J.SAEZ, T.FLICKINGER, M.RAINES, S.PEYROL, C.MOSCOVICI, M.G.MOSCOVICI, H-J.KUNG, P.JURDIC, E.LAZARIDES and J.SAMARUT. (1991). EGF-R as a hemopoietic growth factor receptor : the c-erbB product is present in chicken erythrocytic progenitors and controls their self-renewal.  
Cell 65 : 37-46.
- PANDOLFI,P.P., F.GRIGNANI, M.ALCALAY, A.MENCARELLI, A.BIONDI, F.LOCOCO, F.GRIGNANI and P.G.PELICCI. (1991). Structure and origin of the acute promyelocytic leukemia myl/RAR $\alpha$  cDNA and characterization of its retinoid-binding and transactivation properties.  
Oncogene 6 : 1285-1292.
- PANKRATZ,M.J. and H.JACKLE. (1990). Making stripes in the *Drosophila* embryo.  
TIG 6 : 287-292.
- PARKER,M.G. and O.BAKKER. (1991). Nuclear hormone receptors : concluding remarks. In Nuclear Hormone Receptors. Ed.Parker,M.G. Academic Press : 377-389.

- PAVLETICH,N.P., and C.O,PABO. (1991). Zinc finger-DNA recognition : crystal structure of a Zif268-DNA complex at 2.1 Å.  
Science **252** : 809-817.
- PERLMANN,T. and O.WRANGE. (1988). Specific glucocorticoid receptor binding to DNA reconstituted in a nucleosome.  
The EMBO J. **7** : 3073-3079.
- PETKOVICH,M., N.J.BRAND, A.KRUST and P.CHAMBON. (1987). A human retinoic acid receptor which belongs to the family of nuclear receptors.  
Nature **330** : 444-450.
- PETTY,K.J., H.MORIOKA, T.MITSUHASHI and V.M.NIKODEM. (1989). Thyroid hormone regulation of transcription factors involved in malic enzyme gene expression.  
J. Biol. Chem. **264** : 11483-11490.
- PFAHL,M. and D.BENBROOK. (1987). Nucleotide sequence of cDNA encoding a novel human thyroid hormone receptor.  
Nuc. Acids Res. **15** : 9613.
- PICARD,D. and K.R.YAMAMOTO. (1987). Two signal mediate hormone-dependent nuclear localization of the glucocorticoid receptor.  
The EMBO J. **6** : 3333-3340.
- PICARD,D., S.J.SALSER and K.R.YAMAMOTO. (1988). A movable and regulable inactivation function within the steroid binding domain of the glucocorticoid receptor.  
Cell **54** : 1073-1080.
- PICARD,D., B.KHURSHEED, M.J.GARABEDIAN, M.G.FORTIN, S.LINDQUIST and K.R.YAMAMOTO. (1990). Reduced levels of hsp90 compromise steroid receptor action *in vivo*.  
Nature **348** : 166-168.
- PIGNONI,F., R.M.BALDARELLI, E.STEINGRIMSSON, R.J.DIAZ, A.PATAPOUTIAN, J.R.MERRIAM and J.A.LENGYEL. (1990). The drosophila gene *tailless* is expressed at the embryonic termini and is a member of the steroid receptor superfamily.  
Cell **62** : 151-163.
- PINA.B., U.BRUGGEMEIER and M.BEATO. (1990a). Nucleosome positioning modulates accessibility of regulatory proteins to the Mouse Mammary Tumor Virus promoter.  
Cell **60** : 719-731.
- PINA,B., R.J.G.HACHE, J.ARNEMANN, G.CHALEPAKIS, E.P.SLATER and M.BEATO. (1990b). Hormonal induction of transfected genes depends on DNA topology.  
Mol. Cell. Biol. **10** : 625-633.
- PONGLIKITMONGKOL,M., S.GREEN and P.CHAMBON. (1988). Genomic organization of the human estrogen receptor gene.  
The EMBO J. **7** : 3385-3388.
- PONGLIKITMONGKOL,M., J.H.WHITE and P.CHAMBON. (1990). Synergistic activation of transcription by the human estrogen receptor bound to tandem responsive elements.  
The EMBO J. **9** : 2221-2231.
- PONTA,H., W.H.GÜNZBURG, B.SALMONS, B.GRONER and P.HERRLICH.(1985). Mouse Mammary Tumour Virus : a proviral gene contributes to the understanding of eukaryotic gene expression and mammary tumorigenesis.  
J.Gen.Virol. **66** : 931-943.
- POWER,R.F., J.P.LYDON, O.M.CONNEELY and B.W.O'MALLEY. (1991a). Dopamine activation of an orphan of the steroid receptor superfamily.  
Science **252** : 1546-1547.

- POWER, R.F., S.K.MANI, J.CODINA, O.M.CONNEELY and B.W.O'MALLEY. (1991b). Dopaminergic and ligand-independent activation of steroid hormone receptors. *Science* **254** : 1636-1679.
- PRATT, W.B., D.J.JOLLY, D.V.PRATT, S.M.HOLLENBERG, V.GIGUERE, F.M.CADEPOND, G.SCHWEIZER-GROYER, M.G.CATELLI, R.M.EVANS and E.E.BAULIEU. (1988). A region in the steroid binding domain determines formation of the non-DNA-binding, 9 S glucocorticoid receptor complex. *J. Biol. Chem.* **263** : 267-273.
- PRATT, W.B., E.R.SANCHEZ, E.H.BRESNICK, S.MESHINCHI, K.L.SCHERRER, F.C.DALMAN and M.J.WELSH. (1989). Interaction of the glucocorticoid receptor with the Mr 90.000 heat shock protein : an evolving model of ligand-mediated receptor transformation and translocation. *Cancer Res.* **49** : 2222-2229.
- PRIVALSKY, M.L., P.BOUCHER, A.KONING and C.JUDELSON. (1988). Genetic dissection of functional domains within the avian erythroblastosis virus v-erbA oncogene. *Mol. Cell. Biol.* **8** : 4510-4517.
- PRIVALSKY, M. (1990). A subpopulation of the avian erythroblastosis virus v-erbA protein, a member of the nuclear hormone receptor family, is glycosylated. *J. Virol.* **64** : 463-466.
- PRIVALSKY, M.L., M.SHARIF and K.R.YAMAMOTO. (1990). The viral erbA oncogene protein, a constitutive repressor in animal cells, is a hormone-regulated activator in yeast. *Cell* **63** : 1277-1286.
- PRIVALSKY, M.L. (1991). A subpopulation of the v-erbA oncogene protein, a derivative of thyroid hormone receptor, associates with heat shock protein 90. *J. Biol. Chem.* **266** : 1456-1462.
- PROST, E., R.J.KOENIG, D.D.MOORE, P.R.LARSEN and R.G.WHALEN. (1988). Multiple sequences encoding potential thyroid hormone receptors isolated from mouse skeletal muscle cDNA libraries. *Nuc. Acids Res.* **16** : 6248.
- PROUDFOOT, N.J. (1989). How RNA polymerase II terminates transcription in higher eukaryotes. *TIBS* **14** : 105-110.
- PTASHNE, M. (1988). How eukaryotic transcriptional activators work. *Nature* **335** : 683-689.
- RAGSDALE, C.W., Jr., M.PETKOVICH, P.B.GATES, P.CHAMBON and J.P.BROCKES. (1989). Identification of a novel retinoic acid receptor in regenerative tissues of the newt. *Nature* **341** : 654-657.
- RAGSDALE, C.W., Jr., and J.P.BROCKES. (1991). Retinoic acid receptors and vertebrate limb morphogenesis. In *Nuclear Hormone Receptors*. Ed. Parker, M.G. Academic Press : 269-291.
- REFETOFF, S., L.T.DEWIND and L.J.DE GROOT. (1967). Familial syndrome combining deaf-mutism, stippled epiphyses, goiter and abnormally high PBI : possible target organ refractoriness to thyroid hormone. *J. Clin. Endocrinol. Metab.* **27** : 279-294.
- REFETOFF, S. (1982). Syndromes of thyroid hormone resistance. *Endocrinol. Metab.* **6** : E88-E98.
- REINKAWITZ, R. (1990). Transcriptional repression in eukaryotes. *TIG* **6** : 192-196.

- REITER,R.J., S.KLAUS, C.EBBINGHAUS, G.HELDMAIER, U.REDLIN, D.RICQUIER, M.K.VAUGHAN and S.STEINLECHNER. (1990). Inhibition of 5'-deiodination of thyroxine suppresses the cold-induced increase in brown adipose tissue messenger ribonucleic acid for mitochondrial uncoupling protein without influencing lipoprotein lipase activity. *Endocrinol.* **126** : 2550-2554.
- RICHARD-FOY,H. and G.L.HAGER. (1987). Sequence-specific positioning of nucleosomes over the steroid-inducible MMTV promoter. *The EMBO J.* **6** : 2321-2328.
- RIEHL,R.M., D.O.TOFT, M.D.MEYER, G.L.CARLSON and T.C.McMORRIS. (1984). Detection of a pheromone-binding protein in the aquatic fungus *Achlya ambisexualis*. *Exp. Cell Res.* **153** : 544-549.
- ROSSANT,J., R.ZIRNGIBL, D.CADO, M.SHAGO and V.GIGUERE. (1991). Expression of a retinoic acid response element-hsplacZ transgene defines specific domains of transcriptional activity during mouse embryogenesis. *Genes and Dev.* **5** : 1333-1344.
- ROTHE,M., U.NAUBER and H.JÄCKLE. (1989). Three hormone receptor-like *Drosophila* genes encode an identical DNA-binding finger. *The EMBO J.* **8** : 3087-3094.
- ROUSSEL,M., S.SAULE, C.LAGROU, C.ROMMENS, H.BEUG, T.GRAF and D.STEHELIN. (1979). Three new types of viral oncogene of cellular origin specific for haematopoietic cell transformation. *Nature* **281** : 452-455.
- ROWE,A., N.S.C.EAGER and P.M.BRICKELL. (1991). A member of the RXR nuclear receptor family is expressed in neural-crest-derived cells of the developing chick peripheral nervous system. *Development* **111** : 771-778.
- ROWLEY,J.D., H.M.GOLUMB and C.DOUGHERTY. (1977). 15/17 translocation, a consistent chromosomal change in acute promyelocytic leukaemia. *Lancet* **i** : 549-550.
- RUBERTE,E., P.DOLLE, A.KRUST., A.ZELEN, G.MORRIS-KAY and P.CHAMBON. (1990). Specific spatial and temporal distribution of retinoic acid receptor gamma transcripts during mouse embryogenesis. *Development* **108** : 213-222.
- RUBERTE,E., P.DOLLE, P.CHAMBON and G.MORRIS-KAY. (1991). Retinoic acid receptors and cellular retinoid binding proteins. *Development* **111** : 45-60.
- RUNDLE,S.E., A.I.SMITH, D.STOCKMAN and J.W.FUNDER. (1989). Immunocyto-chemical demonstration of mineralocorticoid receptors in rat and human kidney. *J.Steroid Biochem.* **33** : 1235-1242.
- RUSCONI,S. and K.R.YAMAMOTO. (1987). Functional dissection of the hormone and DNA binding activities of the glucocorticoid receptor. *The EMBO J.* **6** : 1309-1315.
- RUIZ I ALTABA,A., H.PERRY-O'KEEFE and D.A.MELTON. (1987). Xfin : an embryonic gene encoding a multifingered protein in *Xenopus*. *The EMBO J.* **6** : 3065-3070.

- RYSECK, R.P., H. MACDONALD-BRAVO, M.G. MATTEI, S. RUPPERT and R. BRAVO. (1989). Structure, mapping and expression of a growth factor inducible gene encoding a putative nuclear hormonal binding receptor. *The EMBO J.* **8** : 3327-3335.
- SABBAH, M., F. GOUILLEUX, B. SOLA, G. REDEUILH and E.E. BAULIEU. (1990). Structural differences between the hormone and antihormone estrogen receptor complexes bound to the hormonal response element. *Proc. Natl. Acad. Sci. USA* **88** : 390-394.
- SADOWSKI, I., J. MA, S. TRIEZENBERG and M. PTASHNE. (1988). GAL4-VP16 is an unusual potent transcriptional activator. *Nature* **233** : 563-565.
- SAGAMI, I., S. Y. TSAI, H. WANG, M. J. TSAI and B. W. O'MALLEY. (1986). Identification of two factors required for transcription of the ovalbumin gene. *Mol. Cell. Biol.* **6** : 4259-4267.
- SAITO, T. and S. TANIGUCHI. (1991). Involvement of the chox-4 chicken homeobox genes in determination of anteroposterior axial polarity during limb development. *Cell* **64** : 1197-1205.
- SAITOU, N. and M. NEI. (1987). The neighbor-joining method : a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4** : 406-425.
- SAKURAI, A., T. MIYAMOTO, S. REFETTOFF and L. J. DE GROOT. (1990). Dominant negative transcriptional regulation by a mutant thyroid hormone receptor- $\beta$  in a family with generalized resistance to thyroid hormone. *Mol. Endo.* **4** : 1988-1994.
- SAKURAI, A., A. NAKAI and L. J. DE GROOT. (1989). Expression of three forms of thyroid hormone receptor in human tissues. *Mol. Endo.* **3** : 392-399.
- SAKURAI, A., A. NAKAI and L. J. DE GROOT. (1990). Structural analysis of human thyroid hormone receptor  $\beta$  gene. *Mol. Cell. Endocrinol* **71** : 83-91.
- SAMUELS, H.H. and L.E. SHAPIRO. (1976). Thyroid hormone stimulates *de novo* growth hormone synthesis in cultured GH1 cells : evidence for the accumulation of a rate limiting RNA species in the induction process. *Proc. Natl. Acad. Sci. USA* **73** : 3369-3373.
- SAMUELS, H.H., F. STANLEY and L.E. SHAPIRO. (1977). Modulation of thyroid hormone nuclear receptor levels by 3,5,3'-triiodo-L-thyronine in GH1 cells. *J. Biol. Chem.* **252** : 6052-6060.
- SAMUELS, H.H., B.M. FORMAN, Z.D. HOROWITZ and Z-S. YE. (1988). Regulation of gene expression by thyroid hormone. *J. Clin. Invest.* **81** : 957-967.
- SANCHEZ, E.R., D.O. TOFT, M.J. SCHLESINGER and W.B. PRATT. (1985). Evidence that the 90 kDa phosphoprotein associated with the untransformed L-cell glucocorticoid receptor is a murine heat shock protein. *J. Biol. Chem.* **260** : 12398-12401.

- SANCHEZ,E.R., S.MESHINCHI, M.J.SCHLESINGER and W.B.PRATT. (1987). Demonstration that the 90-kilodalton heat shock protein is bound to the glucocorticoid receptor in its 9S nondeoxynucleic acid binding form.  
Mol.Endo. 1 : 908-912.
- SANTOS,A., H.C.FREAKE, M.E.ROSENBERG, H.L.SCHWARTZ and J.H.OPPENHEIMER. (1988). Triiodothyronine nuclear binding capacity in rat tissues correlates with a 6.0 kilobase (kb) and not a 2.6 kb messenger ribonucleic acid hybridization signal generated by a human c-erbA probe.  
Mol. Endo 2 : 992-998.
- SAP,J., A.MUNOZ, K.DAMM, Y.GOLDBERG, J.GHYSDAEL, A.LEUTZ, H.BEUG and B.VENNSTRÖM. (1986). The c-erb-A protein is a high-affinity receptor for thyroid hormone.  
Nature 324 : 635-640.
- SAP,J., A.MUNOZ, J.SCHMITT, H.STUNNENBERG and B.VENNSTROM. (1989). Repression of transcription mediated at a thyroid hormone response element by the v-erbA oncogene product.  
Nature 340 : 242-244.
- SAP,J., L.DE MAGISTRIS, H.STUNNENBERG and B.VENNSTRÖM. (1990). A major thyroid hormone response element in the third intron of the rat growth hormone gene.  
The EMBO J. 9 : 887-896.
- 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.
- SCHAUER,M., G.CHALEPKAIS, T.WILLMANN and M.BEATO. (1989). Binding of hormone accelerates the kinetics of glucocorticoid and progesterone receptor binding to DNA.  
Proc. Natl. Acad. Sci. USA 86 : 1123-1127.
- SCHEIDEREIT,C., S.GEISSE, H.M.WESTPHAL and M.BEATO. (1983). The glucocorticoid receptor binds to defined nucleotide sequences near the promoter of Mouse Mammary Tumour Virus.  
Nature 304 : 749-752.
- SCHENA,M., L.P.FREEDMAN and K.R.YAMAMOTO. (1989). Mutations in the glucocorticoid receptor zinc finger region that distinguish interdigitated DNA binding and transcriptional enhancement activities.  
Genes and Dev. 3 : 1590-1601.
- SCHNEIDER,M.J. and V.A.GALTON. (1991). Regulation of c-erbA- $\alpha$  messenger RNA species in tadpole erythrocytes by thyroid hormone.  
Mol. Endo. 5 : 201-208.
- SCHROEDER,C., C.RAYNOSCHEK, U.FUHRMANN, K.DAMM, B.VENNSTROM and H.BEUG. (1990). The v-erbA oncogene causes repression of erythrocyte-specific genes and an immature, aberrant differentiation phenotype in normal erythroid progenitors.  
Oncogene 5 : 1445-1453.
- SCHUELER,P.A., H.L.SCHWARTZ, K.A.STRAIT, C.N.MARIASH and J.H.OPPENHEIMER. (1990). Binding of 3,5,3'-triiodothyronine (T3) and its analogs to the *in vitro* translational products of c-erbA protooncogenes : differences in the affinity of the  $\alpha$  and  $\beta$  forms for the acetic acid analog and failure of the human testis and kidney  $\alpha$ -2 products to bind T3.  
Mol. Endo. 4 : 227-234.
- SCHÜLE,R., M.MULLER, H.OTSUKA-MURAKAMI and R.RENKAWITZ. (1988). Cooperativity of the glucocorticoid receptor and the CACCC-box binding factor.  
Nature 332 : 87-90.

- SCHÜLE,R., K.UMESONO, D.J.MANGELSDORF, J.BOLADO, J.W.PIKE and R.M.EVANS. (1990a). Jun-Fos and receptors for vitamins A and D recognize a common response element in the human osteocalcin gene.  
Cell 61 : 497-504.
- SCHÜLE,R., P.RANGARAJAN, S.KLIEWER, K.J.RANSONE, J.BOLADO, N.YANG, I.M.VERMA and R.M.EVANS. (1990b). Functional antagonism between oncoprotein c-jun and the glucocorticoid receptor.  
Cell 62 : 1217-1226.
- SCHÜLE E,R. and R.M.EVANS. (1991). Cross-coupling of signal transduction pathways : zinc finger meets leucine zipper.  
TIG 7 : 377-381.
- SCHÜLE,R., P.RANGARAJAN, N.YANG, S.KLIEWER, L.J.RANSONE, J.BOLADO, I.M.VERMA and R.M.EVANS. (1991). Retinoic acid is a negative regulator of AP-1-responsive genes.  
Proc. Natl. Acad. Sci. USA 88 : 6092-6096.
- SCHWABE,J.W.R., D.NEUHAUS and D.RHODES. (1990). Solution structure of the DNA-binding domain of the estrogen receptor.  
Nature 348 : 458-461.
- SCHWABE,J.W.R. and D.RHODES. (1991) Beyond zinc fingers : steroid hormone receptors have a novel structural motif for DNA recognition.  
TIBS 16 : 291-296.
- 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 and Dev. 4 : 204-219.
- SEGRAVES,W.A. (1991). Something old, some things new : the steroid receptor superfamily in *Drosophila*.  
Cell 67 : 225-228.
- SELMI,S. and H.H.SAMUELS. (1991). Thyroid hormone receptor/and v-erbA.  
J. Biol. Chem. 266 : 11589-11593.
- SEVERNE,Y., S.WIELAND, W.SCHAFFNER and S.RUSCONI. (1988). Metal binding "finger" structures in the glucocorticoid receptor defined by site-directed mutagenesis.  
The EMBO J. 7 : 2503-2508.
- SHARIF,M. and M.L.PRIVALSKY. (1991). v-erbA oncogene function in neoplasia correlates with its ability to repress retinoic acid receptor action.  
Cell 66 : 885-893.
- SHEA,M.J., D.L.KING, M.J.CONBOY, B.D.MARIANI and F.C.KAFATOS. (1990). Proteins that bind to *Drosophila* chorion cis-regulatory elements : a new C2H2 zinc finger protein and a C2C2 steroid receptor-like component.  
Genes and Dev. 4 : 1128-1140.
- SHEER,D., D.M.SHEPPARD, M. Le BEAU, J.D.ROWLEY, C.S.ROMAN and E.SOLOMON. (1985). Localization of the oncogene c-erbA1 immediately proximal to the acute promyelocytic leukaemia breakpoint on chromosome 17.  
Ann. Hum. Genet. 49 : 167-171.
- SHEINESS,D., B.VENNSTRÖM and J.M.BISHOP. (1981). Virus-specific RNAs in cells infected by avian myelocytomatosis virus and avian erythroblastosis virus : modes of oncogene expression.  
Cell 23 : 291-300.

- SHI, Y.B. and D.D.BROWN. (1990). Developmental and thyroid hormone-dependent regulation of pancreatic genes in *Xenopus laevis*.  
Genes and Dev. 4 : 1107-1113.
- SHUPNIK, M.A., W.W.CHIN, J.F.HABENER and E.C.RIDGWAY. (1985). Transcriptional regulation of the thyrotropin subunit genes by thyroid hormone.  
J.Biol.Chem. 260 : 2900-2903.
- SILVA, J.E. and P.R.LARSEN. (1983). Adrenergic activation of triiodothyronine production in brown adipose tissue.  
Nature 305 : 712-713.
- SILVA, J.E. (1986). Brown adipose tissue : an extrathyroidal source of triiodothyronine.  
News Physiol. Sci. 1 : 119-122.
- SIMEONE, A., D.ACAMPORA, V.NIGRO, A.FAIELLA, M.D'ESPOSITO, A.STORNAIUOLO, F.MAVILIO and E.BONCINELLI. (1991). Differential regulation by retinoic acid of the homeobox genes of the four Hox loci in human embryonal carcinoma cells.  
Mechanisms of Dev. 33 : 215-228.
- SIMMONS, D.M., J.W.VOSS, H.A.INGRAHAM, J.M.HOLLOWAY, R.S.BROIDE, M.G.ROSENFELD and L.W.SWANSON. (1990). Pituitary cell phenotypes involve cell-specific Pit-1 mRNA translation and synergistic interactions with other classes of transcription factors.  
Genes and Dev. 4 : 695-711.
- SIMONS, S.S. and E.B.THOMPSON. (1981). Dexamethasone 21-mesylate : an affinity label of glucocorticoid receptors from rat hepatoma tissue culture cells.  
Proc. Natl. Acad. Sci. USA 78 : 3541-3545.
- SIMONS, S.S., J.G.PUMPHREY, S.RUDIKOFF and H.I.EISEN. (1987). Identification of cysteine 656 as the amino acid of hepatoma tissue culture cell glucocorticoid receptors that is covalently labeled by dexamethasone 21-mesylate.  
J.Biol.Chem. 262 : 9676-9680.
- SIMONS, S.S., F.D.SISTARE and P.K.CHAKRABORTI. (1989). Steroid binding activity is retained in a 16-kDa fragment of the steroid binding domain of rat glucocorticoid receptors.  
J. Biol. Chem. 264 : 14493-14497.
- SINHA, A.M., P.K.UMEDA, C.J.KAVINSKY, C.RAJAMANICKAM, J.J.HSU, S.JAKOVIC and M.M.RABINOWITZ. (1982). Molecular cloning of mRNA sequences for cardiac  $\alpha$ - and  $\beta$ -form myosin heavy chains : expression in ventricles of normal, hypothyroid, and thyrotoxic rabbits.  
Proc.Natl.Acad.Sci.USA 79 : 5847-5851.
- SJÖBERG, M., B.VENNSTRÖM and D.FORREST. (1992). Thyroid hormone receptors in chick retinal development : differential expression of mRNAs for  $\alpha$  and N-terminal variant  $\beta$  receptors.  
Development 114 : 39-47.
- SKOWRONSKI, R., and D.FELDMAN. (1989). Characterization of an estrogen-binding protein in the yeast *Candida albicans*.  
Endocrinology 124 : 1965-1972.
- SLACK, J.M.W. (1987). We have a morphogen !  
Nature 327 : 553-554.
- SLADEK, F.M., W.ZHONG, E.LAI and J.E.DARNELL. (1990). Liver-enriched transcription factor HNF-4 is a novel member of the steroid hormone receptor superfamily.  
Genes and Dev. 4 : 2353-2365.

- SLATER, E.P., O.RABENAU, M.KARIN, J.D.BAXTER and M.BEATO. (1985). Glucocorticoid receptor binding and activation of a heterologous promoter by dexamethasone in the first intron of the human growth hormone gene. *Mol. Cell. Biol.* **5** : 2984-2992.
- SMITH, G.P. (1976). Evolution of repeated DNA sequences by unequal crossovers. *Science* **191** : 528-535.
- SMITH, L.I., J.E.BODWELL, D.B.MENDEL, T.CIARDELLI, W.G.NORTH and A.MUNCK. (1988). Identification of cysteine-644 as the covalent site of attachment of dexamethasone 21-mesylate to murine glucocorticoid receptors in WEHI-7 cells. *Biochemistry* **27** : 3747-3753.
- SMITH, S.M. and G.EICHELE. (1991). Temporal and regional differences in the expression pattern of distinct retinoic acid receptor- $\beta$  transcripts in the chick embryo. *Development* **111** : 245-252.
- SPENCER, C.A., R.D.GIETZ and R.B.HODGETTS. (1986). Overlapping transcription units in the dopa decarboxylase region of *Drosophila*. *Nature* **322** : 279-281.
- SPURR, N.K., E.SOLOMON, M.JANSSON, D.SHEER, P.N.GOODFELLOW, W.F.BODMER and B.VENNSTRÖM. (1984). Chromosomal localisation of the human homologous to the oncogenes *erbA* and B. *The EMBO J.* **3** : 159-163.
- STEWART, P.M., A.M.WALLACE, R.VALENTINO, D.BURT, C.H.L.SHAKELTON and C.R.W.EDWARDS. (1987). Mineralocorticoid activity of liquorice : 11-beta-hydroxysteroid deficiency comes of age. *Lancet* **11** : 821-824.
- STRÄHLE, U., W.SCHMID and G.SCHUTZ. (1988). Synergistic action of the glucocorticoid receptor with transcription factors. *The EMBO J.* **7** : 3389-3395.
- STRAIT, K.A., H.L.SCHWARTZ, V.S.SEYBOLD, N.C.LING and J.H.OPPENHEIMER. (1991). Immunofluorescence localization of thyroid hormone receptor protein  $\beta$ 1 and variant  $\alpha$ 2 in selected tissues : cerebellar purkinje cells as a model for  $\beta$ 1 receptor-mediated developmental effects of thyroid hormone in brain. *Proc. Natl. Acad. Sci. USA* **88** : 3887-3891.
- SUCOV, H.M., K.K.MURAKAMI and R.M.EVANS. (1990). Characterization of an autoregulated response element in the mouse retinoic acid receptor type  $\beta$  gene. *Proc. Natl. Acad. Sci. USA* **87** : 5392-5396.
- SULTAN, C. and J.M.LOBACCARO. (1991). Génétique moléculaire des syndromes d'insensibilité aux androgènes. *Médecine/Sciences* **7** : 697-704.
- SUMMERBELL, D. and M.MADEN. (1990). Retinoic acid, a developmental signalling molecule. *TINS* **13** : 142-147.
- SUPERTI-FURGA, G., G.BERGERS, D.PICARD and M.BUSSLINGER. (1991). Hormone-dependent transcriptional regulation and cellular transformation by *fos*-steroid receptor fusion proteins. *Proc. Natl. Acad. Sci. USA* **88** : 5114-5118.

- TABIN, C.J. (1991). Retinoids, homeoboxes, and growth factors : toward molecular models for limb development.  
*Cell* **66** : 199-217.
- TAKEDA, K., S.BALZANO, A.SAKURAI, L.J. DE GROOT and S.REFETTOFF. (1991). Screening of nineteen unrelated families with generalized resistance to thyroid hormone for known point mutations in the thyroid hormone receptor  $\beta$  gene and the detection of a new mutation.  
*J. Clin. Invest.* **87** : 496-502.
- TANAKA, S., S.HASEGAWA, F.HISHINUMA and S.I.KURATA. (1989). Estrogen can regulate the cell cycle in the early G1 phase of yeast by increasing the amount of adenylate cyclase mRNA.  
*Cell* **57** : 675-681.
- 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.
- TATA, J.R. (1990). Rosalind Pitt-Rivers and the discovery of T3.  
*TIBS* **15** : 282-284.
- TATA, J.R., A.KAWAHARA and B.S.BAKER. (1991). Prolactin inhibits both thyroid hormone-induced morphogenesis and cell death in cultured amphibian larval tissues.  
*Dev. Biol.* **146** : 72-80.
- THALLER, C. and G.EICHELE. (1987). Identification and spatial distribution of retinoids in the developing chick limb bud.  
*Nature* **327** : 625-628.
- THALLER, C. and G.EICHELE. (1990). Isolation of 3,4-didehydroretinoic acid, a novel morphogenetic signal in the chick wing bud.  
*Nature* **345** : 815-819.
- THOMPSON, C.C., C.WEINBERGER, R.LEBO and R.M.EVANS. (1987). Identification of a novel thyroid hormone receptor expressed in the mammalian central nervous system.  
*Science* **237** : 1610-1614.
- THOMPSON, C.C. and R.M.EVANS. (1989). Trans-activation by thyroid hormone receptors : functional parallels with steroid hormone receptors.  
*Proc. Natl. Acad. Sci. USA* **86** : 3494-3498.
- TICKLE, C. (1991). Retinoic acid and chick limb bud development.  
*Development* **1** : 113-121.
- TILLEY, W.D., M.MARCELLI, J.D.WILSON and M.J.McPHAUL. (1989). Characterization and expression of a cDNA encoding the human androgen receptor.  
*Proc. Natl. Acad. Sci. USA* **86** : 327-331.
- TORA, L., M.P.GAUB, S.MADER, A.DIERICH, M.BELLARD and P.CHAMBON. (1988). Cell-specific activity of a GGTC A half-palindromic estrogen-responsive element in the chicken ovalbumin gene promoter.  
*The EMBO J.* **7** : 3771-3778.
- TORA, L., A.MULLICK, D.METZGER, M.PONGLIKITMONGKOL, I.PARK and P.CHAMBON. (1989a). The cloned human estrogen receptor contains a mutation which alters its hormone binding properties.  
*The EMBO J.* **8** : 1981-1986.

- TORA, L., J. WHITE, C. BROU, D. TASSET, N. WEBSTER, E. SCHEER and P. CHAMBON. (1989b). The human estrogen receptor has two independent nonacidic transcriptional activation functions.  
Cell 59 : 477-487.
- TSAI, S. Y., I. SAGAMI, H. WANG, M. J. TSAI and B. W. O'MALLEY. (1987). Interactions between a DNA-binding transcription factor (COUP) and a non-DNA binding factor (S300-II).  
Cell 50 : 701-709.
- TSAI, S. Y., J. A. CARLSTEDT-DUKE, N. L. WEIGEL, K. DAHLMAN, J. A. GUSTAFSSON, M. J. TSAI and B. W. O'MALLEY. (1988). Molecular interactions of steroid hormone receptor with its enhancer element : evidence for receptor dimer formation.  
Cell 55 : 361-369.
- TSIKA, R. W., J. J. BAHL, L. A. LEINWAND and E. MORKIN. (1990). Thyroid hormone regulates expression of a transfected human  $\alpha$ -myosin heavy-chain fusion gene in fetal rat heart cells.  
Proc. Natl. Acad. Sci. USA 87 : 379-383.
- TUGWOOD, J. D., I. LISSEMAN, R. G. ANDERSON, K. R. BUNDELL, W. L. MCPHEAT and S. GREEN. (1992). The mouse peroxisome proliferator activated receptor recognizes a response element in the 5' flanking sequence of the rat acyl CoA oxidase gene.  
The EMBO J. 11 : 433-439.
- TZUKERMAN, M., X. K. ZHANG, T. HERMANN, K. N. WILLS, G. GRAUPNER and M. PFAHL. (1990). The human estrogen receptor has transcriptional activator and repressor functions in the absence of ligand.  
The New Biol. 2 : 613-620.
- UCKER, D. S. and K. R. YAMAMOTO. (1984). Early events in the stimulation of Mammary Tumor Virus RNA synthesis by glucocorticoids.  
J. Biol. Chem. 259 : 7416-7420.
- UMESONO, K., V. GIGUERE, C. K. GLASS, M. G. ROSENFELD and R. M. EVANS. (1988). Retinoic acid and thyroid hormone induce gene expression through a common responsive element.  
Nature 336 : 262-265.
- UMESONO, K. and R. M. EVANS. (1989). Determinants of target gene specificity for steroid/thyroid hormone receptors.  
Cell 57 : 1139-1146.
- UMESONO, K., K. K. MURAKAMI, C. C. THOMPSON and R. M. EVANS. (1991). Direct repeats as selective response elements for the thyroid hormone, retinoic acid, and vitamin D3 receptors.  
Cell 65 : 1255-1266.
- USALA, S. J., A. E. BALE, N. GESUNDHEIT, C. WEINBERGER, R. W. LASH, F. E. WONDISFORD, O. W. Mc BRIDE and B. D. WEINTRAUB. (1988a). Tight linkage between the syndrome of generalized thyroid hormone resistance and the human c-erbA $\beta$  gene.  
Mol. Endo. 2 : 1217-1218.
- USALA, S. J., W. S. YOUNG III, H. MORIOKA and V. M. NIKODEM. (1988b). The effect of thyroid hormone on the chromatin structure and expression of the malic enzyme gene in hepatocytes.  
Mol. Endo. 2 : 619-626.
- USALA, S. J., G. E. TENNYSON, A. E. BALE, R. W. LASH, N. GESUNDHEIT, F. E. WONDISFORD, D. ACCILI, P. HAUSER and B. D. WEINTRAUB. (1990). A base mutation of the c-erbA $\beta$  thyroid hormone receptor in a kindred with generalized thyroid hormone resistance.  
J. Clin. Inv. 85 : 93-100.

- USALA, S.J., J.B.MENKE, T.L.WATSON, F.E.WONDISFORD, B.D.WEINTRAUB, J.BERARD, W.E.C.BRADLEY, S.ONO, O.T.MUELLER and B.B.BERCU. (1991). A homozygous deletion in the c-erbA $\beta$  thyroid hormone receptor gene in a patient with generalized thyroid hormone resistance : isolation and characterization of the mutant receptor. *Mol. Endo.* **5** : 327-335.
- VALENTINE, J.W. (1977). In Hallam, A. (ed.), *Pattern of Evolution*. Elsevier, Amsterdam, pp 27-58
- VALLEE, B.L., J.E.COLEMAN and D.S.AULD. (1991). Zinc fingers, zinc clusters, and zinc twists in DNA-binding protein domains. *Proc. Natl. Acad. Sci. USA* **88** : 999-1003.
- VAN TUINEN, P., K.R.JOHNSON, S.A.LEDBETTER, R.L.NUSSBAUM, G.ROVERA and D.H.LEDBETTER. (1987). Localization of myeloperoxidase to the long arm of human chromosome 17 : relationship to the 15 ; 17 translocation of acute promyelocytic leukemia. *Oncogene* **1** : 319-322.
- VASSIOS, G.W., J.D.GOLD, M.PETKOVICH, P.CHAMBON and L.J.GUDAS. (1989). A retinoic acid-responsive element is present in the 5' flanking region of the laminin B1 gene. *Proc. Natl. Acad. Sci. USA* **86** : 9099-9103.
- VEDECKIS, W.V. (1981). Activation and chromatographic properties of the AtT-20 mouse pituitary tumor cell line glucocorticoid receptor. *Biochemistry* **20** : 7237-7245.
- WAGNER, M., C.THALLER, T.JESSELL and G.EICHELE. (1990). Polarizing activity and retinoid synthesis in the floor plate of the neural tube. *Nature* **345** : 819-822.
- WALTER, P., S.GREEN, G.GREENE, A.KRUST, J-M.BORNERT, J.M.JELTSCH, A.STAUB, E.JENSEN, G.SCRACE, M.WATERFIELD and P.CHAMBON. (1985). Cloning of the human estrogen receptor cDNA. *Proc. Natl. Acad. Sci. USA* **82** : 7889-7893.
- WANEK, N., D.M.GARDINER, K.MUNEOKA and S.V.BRYANT. (1991). Conservation by retinoic acid of anterior cells into ZPA cells in the chick wing bud. *Nature* **350** : 81-83.
- WANG, L.H., S.Y.TSAI, R.G.COOK, W.G.BEATTIE, M.J.TSAI and B.W.O'MALLEY. (1989). COUP transcription factor is a member of the steroid receptor superfamily. *Nature* **340** : 163-166.
- WANG, Z.Y., G.L.SUN, J.X.LU, L.J.GU, M.E.HUANG and S.R.CHEN. (1990). Treatment of acute promyelocytic leukemia with all-trans retinoic acid in China. *J.Exp.Clin.Hematol.* **32** : 34-36.
- WARRELL, R.P., S.R.FRANKEL, W.H.MILLER, D.A.SCHEINBERG, L.M.ITRI, W.N.HITTELMAN, R.VYAS, M.ANDREEFF, A.TAFURI, A.JAKUBOWSKI, J.GABRILOVE, M.S.GORDON and E.DMITROVSKY. (1991). Differentiation therapy of acute promyelocytic leukemia with tretinoin (all-trans retinoic acid). *N. Engl. J. Med.* **324** : 1385-1393.
- WATSON, M.A. and J.MILBRANDT. (1990). Expression of the nerve growth factor-regulated NGFI-A and NGFI-B genes in the developing rat. *Development* **110** : 173-183.
- WEBSTER, N.J., S.GREEN, J.R.JIN and P.CHAMBON. (1988). The hormone-binding domains of the estrogen and glucocorticoid receptors contain an inducible transcription activation function. *Cell* **54** : 1299-207.

- 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 estrogen receptor is not encoded in a single exon.  
The EMBO J. 8 : 1441-1446.
- WEIL, S.C., G.L.ROSNER, M.S.REID, R.L.CHISHOLM, R.S.LEMONS, M.S.SWANSON, J.J.CARRINO, M.O.DIAZ and M.M. LEBEAU. (1988). Translocation and rearrangement of myeloperoxidase gene in acute promyelocytic leukemia.  
Science 240 : 790-792.
- WEIL, S.C., M.S.REID, L.A.NILLES, R.L.CHISHOLM, G.L.ROSNER, M.S.SWANSON, J.J.CARRIDO, M.O.DIAZ, M.M.LEBEAU. (1989). The myeloperoxidase gene in acute promyelocytic leukemia.  
Science 244 : 823-826.
- WEINBERGER, C., S.M.HOLLENBERG, M.G.ROSENFELD and R.M.EVANS. (1985). Domain structure of human glucocorticoid receptor and its relationship to the v-erb-A oncogene product.  
Nature 318 : 670-672.
- WEINBERGER, C., C.C.THOMPSON, E.S.ONG, R.LEBO, D.J.GRUOL and R.M.EVANS. (1986). The c-erb-A gene encodes a thyroid hormone receptor.  
Nature 324 : 641-646.
- WICKENS, M. (1990). How the messenger got its tail : addition of poly(A) in the nucleus.  
TIBS 15 : 277-281.
- WIGHT, P.A., M.D.CREW, and S.R.SPINDLER. (1987). Discrete positive and negative thyroid hormone-responsive transcription regulatory elements of the rat growth hormone gene.  
J. Biol. Chem. 262 : 5659-5663.
- WIGHT, P.A., M.D.CREW and S.R.SPINDLER. (1988). Sequences essential for activity of the thyroid hormone responsive transcription stimulatory element of the rat growth hormone gene.  
Mol. Endo. 2 : 536-545.
- WILHELMSSON, A., S.CUTHILL, M.DENIS, A-C.WIKSTRÖM, J-A GUSTAFSSON and L.POELLINGER. (1990). The specific DNA binding activity of the dioxin receptor is modulated by the 90 kd heat shock protein.  
The EMBO J. 9 : 69-76.
- WILLIAMS, T. and M.FRIED. (1986). A mouse locus at which transcription from both DNA strands produces mRNAs complementary at their 3' ends.  
Nature 322 : 275-279.
- WILLMANN, T. and M.BEATO. (1986). Steroid-free glucocorticoid receptors binds specifically to Mouse Mammary Tumour Virus DNA.  
Nature 324 : 688-691.
- WILSON, T.E., T.J.FAHRNER, M.JOHNSTON and J.MILBRANDT. (1991). Identification of the DNA binding site for NGFI-B by genetic selection in yeast.  
Science 252 : 1296-1300.
- WILSON, J.D. and D.W.FOSTER. (1985). Hormones and hormones action : Introduction. In Williams Textbook of Endocrinology. Seventh Edition. Eds. J.D.Wilson and D.W.Foster : 1-8.
- WISDEN, W., M.L.ERRINGTON, S.WILLIAMS, S.B.DUNNETT, C.WATERS, D.HITCHCOCK, G.EVAN, T.V.P.BLISS and S.P.HUNT. (1990). Differential expression of immediate early genes in the hippocampus and spinal cord.  
Neuron 4 : 603-614.

- WONDISFORD, F.E., E.A.FARR, S.RADOVICK, H.J.STEINFELDER, J.M.MOATES, J.H.McCLASKEY and B.D.WEINTRAUB. (1989). Thyroid hormone inhibition of human thyrotropin  $\beta$ -subunit gene expression is mediated by a cis-acting element located in the first exon. *J. Biol. Chem.* **264** : 14601-14604.
- WONG, N.C.W., H.L.SCHWARTZ, K.STRAIT and J.H.OPPENHEIMER. (1989). Thyroid hormone-, carbohydrate, and age-dependent regulation of a methylation site in the hepatic S14 gene. *Mol. Endo.* **3** : 645-650.
- WOOD, W.M., M.Y.KAO, D.F.GORDON and E.CHESTER RIDGWAY. (1989). Thyroid hormone regulates the mouse thyrotropin  $\beta$ -subunit gene promoter in transfected primary thyrotropes. *J. Biol. Chem.* **264** : 14840-14847.
- WRANGE, Ö, S.OKRET, M.RADOJCIC, J.CARLSTEDT-DUKE and J.A.GUSTAFSSON. (1984). Characterization of the purified activated glucocorticoid receptor from rat liver cytosol. *J. Biol. Chem.* **259** : 4534-4541.
- YAMAMOTO, K.R. and B.M.ALBERTS. (1976). Steroid receptors : elements for modulation of eukaryotic transcription. *Annu. Rev. Biochem.* **45** : 721-746.
- 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.
- YAOITA, Y. and D.D.BROWN. (1990). A correlation of thyroid hormone receptor gene expression with amphibian metamorphosis. *Genes and Dev.* **4** : 1917-1924.
- YAOITA, Y., Y-B.SHI and D.BROWN. (1990). *Xenopus laevis*  $\alpha$  and  $\beta$  thyroid hormone receptors. *Proc. Natl. Acad. Sci. USA* **87** : 7090-7094.
- YE, Z-S. and H.H.SAMUELS. (1987). Cell- and sequence-specific binding of nuclear proteins to 5'-flanking DNA of the rat growth hormone gene. *J. Biol. Chem.* **262** : 6313-6317.
- YE, Z-S., B.M.FORMAN, A.ARANDA, A.PASCUAL, H-Y.PARK, J.CASANOVA and H.H.SAMUELS. (1988). Rat growth hormone gene expression. *J. Biol. Chem.* **263** : 7821-7829.
- YU, V.C., C.DELSERT, B.ANDERSEN, J.M.HOLLOWAY, O.V.DEVARY, A.M.NÄÄR, S.YUN KIM, J.M.BOUTIN, C.K.GLASS and M.G.ROSENFELD. (1991). RXR $\beta$  : a coregulator that enhances binding of retinoic acid, thyroid hormone, and vitamin D receptors to their cognate response elements. *Cell* **67** : 1251-1266.
- ZAHRAOUI, A. and G.CUNY. (1987). Nucleotide sequence of the chicken proto-oncogene c-erbA corresponding to domain 1 of v-erbA. *Eur. J. Biochem.* **166** : 63-69.
- ZELENT, A., A.KRUST, M.PETKOVICH, P.KASTNER and P.CHAMBON. (1989). Cloning of murine  $\alpha$  and  $\beta$  retinoic acid receptors and a novel receptor  $\gamma$  predominantly expressed in skin. *Nature* **339** : 714-717.

- ZELENT, A., C. MENDELSON, P. KASTNER, A. KRUST, J.-M. GARNIER, F. RUFFENACH, P. LEROY and P. CHAMBON. (1991). Differentially expressed isoforms of the mouse retinoic acid receptor  $\beta$  are generated by usage of two promoters and alternative splicing. *The EMBO J.* **10** : 71-81.
- ZENKE, M., P. KAHN, C. DISELA, B. VENNSTRÖM, A. LEUTZ, K. KEEGAN, M. J. HAYMAN, H.-R. CHOI, N. YEW, J. D. ENGEL and H. BEUG. (1988). v-erbA specifically suppresses transcription of the avian erythrocyte anion transporter (Band 3) gene. *Cell* **52** : 107-119.
- ZENKE, M., A. MUNOZ, J. SAP, B. VENNSTROM and H. BEUG. (1990). v-erbA oncogene activation entails the loss of hormone-dependent regulator activity of c-erbA. *Cell* **61** : 1035-1049.
- ZHANG, X. K., K. N. WILLS, G. GRAUPNER, M. TZUKERMAN, T. HERMANN and M. PFAHL. (1991a). Ligand-binding domain of thyroid hormone receptors modulates DNA binding and determines their bifunctional roles. *The New Biol.* **3** : 169-181.
- ZHANG, X. K., K. N. WILLS, M. HUSMANN, T. HERMANN and M. PFAHL. (1991b). Novel pathway for thyroid hormone receptor action through interaction with jun and fos oncogene activities. *Mol. Cell. Biol.* **11** : 6016-6025.
- ZHANG, X. K., B. HOFFMANN, P. B. V. TRAN, G. GRAUPNER and M. PFAHL. (1992). Retinoid X receptor is an auxiliary protein for thyroid hormone and retinoic acid receptors. *Nature* **355** : 441-446.



## RESUME

De nombreux facteurs régulant la transcription des gènes eucaryotes ont été caractérisés ces dernières années et, parmi ceux ci, des facteurs agissant sous le contrôle d'une hormone : les récepteurs nucléaires d'hormones. Ces molécules ont la même structure de base constituée de domaines fonctionnels relativement indépendants. Les deux régions les mieux caractérisées étant le domaine C, responsable de la fixation spécifique à l'ADN via deux structures rappelant les "doigts de zinc" et le domaine E situé en C-terminal, responsable de la fixation du ligand et, souvent, de la régulation de la transcription. Cette superfamille de gènes contient notamment les récepteurs des hormones stéroïdes, les récepteurs de l'acide rétinoïque, ceux des hormones thyroïdiennes mais aussi un certain nombre de récepteurs putatifs dont les ligands éventuels ne sont pas connus et qui sont appelés "récepteurs orphelins".

Les récepteurs des hormones thyroïdiennes codés par les gènes c-erbA-1 et c-erbA-2 nous ont particulièrement intéressés parce que l'oncogène v-erbA qui fait partie du génome du rétrovirus oncogène aviaire AEV est une version altérée du récepteur des hormones thyroïdiennes de type  $\alpha$  (c-erbA-1). Nous avons caractérisé l'organisation génomique du gène c-erbA-1 qui est constitué de 10 exons répartis sur 27 kb d'ADN du chromosome 17 chez l'homme. Au cours de cette étude nous avons montré que ce gène peut coder deux formes différentes de récepteurs suite à un épissage alternatif qui implique les deux derniers exons du gène. De plus, nous avons observé que le gène c-erbA-1 chevauchait un autre gène appartenant également à la famille des récepteurs nucléaires et appelé ear-1.

L'étude de l'organisation génomique de c-erbA-1 et la mise en évidence de cette structure chevauchante très inhabituelle nous ont poussé à étudier de plus près l'évolution de la superfamille des récepteurs nucléaires. En construisant des arbres phylogéniques basés sur la séquence des deux domaines conservés des récepteurs nucléaires (les domaines C et E) nous avons pu construire un modèle d'évolution de cette famille de gènes. Ce modèle implique que les récepteurs nucléaires sont apparus très tôt au cours de l'évolution et se sont probablement diversifiés rapidement au moment des stades d'évolution explosive des tout premiers métazoaires. Nous avons pu grouper les récepteurs nucléaires en trois sous-familles : celle des récepteurs des hormones thyroïdiennes, de l'acide rétinoïque et des gènes apparentés comme ear-1, celle des récepteurs orphelins et enfin celle des récepteurs des stéroïdes. Le regroupement de la majorité des récepteurs orphelins a été une surprise qui permet peut être d'imaginer un rôle commun pour ces molécules. Enfin, en comparant l'évolution des deux domaines structuraux nous avons pu montrer que les récepteurs nucléaires ont en général évolué par simple duplication de gènes. Cependant il existe des exceptions importantes à ce modèle comme les gènes de drosophile du groupe *knirps* ou comme les gènes des récepteurs de la vitamine D et de l'ecdysone qui semblent avoir une histoire évolutive plus complexe.