



UNIVERSITÉ DE STRASBOURG

ÉCOLE DOCTORALE DES SCIENCES DE LA VIE ET DE LA SANTÉ

*Thèse présentée par*  
**Emilie CLER**

*pour obtenir le titre de*

**Docteur de l'Université de Strasbourg**  
*Sciences du Vivant-Aspects Moléculaires et Cellulaires de la Biologie*

*Caractérisation d'un sous-complexe de TFIID et de  
la fonction anti-oncogène de sa sous-unité TAF4  
dans les fibroblastes embryonnaires murins*

*Soutenue publiquement le 18 septembre 2009*

Membres du jury :

Directeur de Thèse : **Dr Irwin DAVIDSON**

Rapporteur interne : **Dr Cécile ROCHETTE - EGLY**

Rapporteur externe : **Dr Matthieu GERARD**

Rapporteur externe : **Dr Laurence VANDEL**

Membre invité : **Pr Philippe CARBON**



*A mes parents et à ma sœur,*

*Merci pour votre soutien si précieux !*

*A mon cousin Nico,*

*Qui m'a donné le courage d'affronter  
tous les instants...*



*Il faut toujours viser la lune car même en cas d'échec  
on atterrit dans les étoiles (Oscar Wilde).*



## Remerciements

*Mes premiers remerciements vont aux membres de mon jury qui ont accepté d'évaluer mon travail de thèse, notamment à Cécile Rochette-Egby, à Philippe Carbon, à Laurence Vandiel et à Matthieu Gérard, venus parfois de loin pour écouter et juger ma soutenance orale.*

*Je tiens également à remercier tout particulièrement le Dr Irwin Davidson qui m'a accueillie dans son laboratoire pour mon travail de thèse. Il a su faire preuve d'une grande disponibilité, d'une incroyable sérénité et d'un calme Olympien ! Avec un œil expert et une mémoire impressionnante, il m'a conseillée, et suivie dans mes investigations. Il m'a guidée dans l'univers des complexes de transcription tout en m'encourageant à poursuivre une partie de mon sujet en cancérologie. Merci pour m'avoir fait confiance et pour tous vos conseils scientifiques au long de ces 4 années.*

*Merci aussi à tous les membres du laboratoire qui m'ont soutenue au quotidien, avec qui j'ai pu partager fous rires et verres cassés (oups !), odeurs chimiques et petites histoires, grognements et bonne humeur et qui m'ont donné le courage de me battre jusqu'au bout. Un grand merci à Isabelle M. pour le réconfort qu'elle nous apporte quotidiennement, pour son écoute et sa compassion, à Manu pour sa disponibilité à toutes épreuves et son humour grinçant (et vive lab FM 4039, la meilleure des radios du monde !), à Dominique qui m'a plus que soutenue ces derniers temps dans la réalisation de mon petit projet secret aujourd'hui plus trop secret, à Ataaillah pour ses sages conseils et son soutien précieux dans mes évadées scientifiques, à Daniil pour les chouettes moments partagés au parapente (à nous, la Dune du Pyla !), à Thomas pour son éternel optimisme, à Amin et à Igor pour les petits apéros surprises, et à Gabriëlle pour la relecture du manuscrit et pour ses gâteaux délicieux !*

*Bonne chance pour la suite aux petits nouveaux, à Laure pour sa présence et sa gentillesse, à Ghislain (le fameux chevalier aux chaînes de feux), à Anne-Claire (on se retrouve en Guadeloupe ?) et à Sylvia la petite dernière !*

*Je tiens aussi à saluer les anciens, ex-IGBMciens : Laurence pour son enthousiasme débordant et son incroyable dynamisme, Isa B. pour les soirées Mojito et nos tribulations à travers toute l'Alsace mais aussi pour sa personnalité hors du commun et son amitié inestimable, Aurore pour son humour fin et décalé, Anas pour ses petites discussions autour d'un café, Ullas pour ses bonnes idées scientifiques, Raffaëlla pour nos petites virées dans Strasbourg et Philippe pour son soutien, ses nems et sa jolie petite fille aux yeux bridés et aux cheveux bouclés qui me donne toujours une folle envie de voyager très loin.*

*Mais tout ça n'aurait pas été possible sans une initiation au préalable au monde de la Science !*

*Un grand merci à Angela Giangrande et à toute son équipe qui m'ont accueillie dans leur laboratoire lorsque j'étais novice, qui m'ont guidée dans mes premiers pas scientifiques et qui m'ont fait découvrir l'univers surprenant des Drosophiles, à Bohdan Wasyluk qui m'a permis de préparer le DEA dans son laboratoire et surtout à mes ex-collègues formidables, amis d'aujourd'hui : Christian, Gilles, Atefeh, Alberto et Christina qui ont rendu cette année-là fort sympathique et qui m'ont donnée envie de poursuivre en thèse. Merci aussi à ceux qui leur ont succédé et qui sont devenus de précieux amis au fil de ces 4 années de thèse: notamment Jan le belge qui fait les meilleures moules-frites illkirchoises du monde entier, et Héléne S. et Céline C., mes garde-bonne-humeur, toujours présentes pour échanger quelques bonnes paroles réconfortantes ou humoristiques autour d'un café ou d'un thé à la menthe !*

*Ayant pu apprécier la Science au sein de 3 des 4 étages de cet institut, je tiens donc à remercier chaleureusement les nombreuses personnes que j'ai côtoyées de près ou de loin lors de ma thèse. Avec un merci spécial à ceux de l'étage où la vue est la plus belle (!) et notamment aux membres des labos Egly, Tora et Losson qui m'ont accueillie de temps à autre dans leur laboratoire, le temps d'une manip, d'un entremède ou d'une discussion ! Merci à mes voisins proches et mitoyens: Sascha, Rénier (le meilleur des confidents et le roi des Long Island !), à Nico, David, Bruno, Valentyn, Colm, à mes voisins d'en face : Florence (à quand cette raclette alors?!), Johan (pour tous les concerts, les festivals de rock indépendant et toutes les soirées inoubliables au Trou !), Elie, Emese, Guillaume et Jacques (ça SMART bien?), Arnaud, Didier et Lázlo, et à ceux du bout du couloir : Christine F. (pour les petites discussions autour de tout et de rien et pour nos progrès extraordinaires en snowboard !), Nathalie (pour son calme et son éternel sourire) et Dorothée pour son enthousiasme débordant !*

*Merci aux personnes des services communs et notamment à celles du service de microscopie pour l'ambiance formidable qui y règne. Merci à toi Marc qui a tout de suite compris ma détresse face aux microscopes et autres bestioles à objectifs ! à toi J-D qui m'a fait découvrir l'univers des petits poissons-zèbres cyclopes (à quand le barbecue dans ta nouvelle maison ?), au cousin Yves si serein et si méticuleux, à Didier pour ses récits de voyages lointains et à Jean-Luc toujours très attentif à tous nos projets scientifiques.*

*Bon courage à tous les rédacteurs assidus de la bibliothèque : Séb, Lama, Wassim, Josselin, Caroline, Axelle et les autres avec qui j'ai partagé l'angoisse de la page vide, du mémoire qui ne veut pas s'écrire tout seul, le stress des jours qui s'égrainent à vive allure, mais aussi plein d'agréables moments*



*autour d'une bonne tarte aux framboises, de jeux désopilants, et de délires farfelus et savoureux des derniers instants d'une thèse...*

*Et parce qu'il y a aussi une vie en dehors du laboratoire (!) je tiens à remercier tous mes amis aujourd'hui répartis aux quatre coins du monde : Claire P., ma formidable binôme de concert, qui m'a accompagnée dans les catacombes du rock indépendant, Claire C. et les après-midis passés à Strasbourg-Plage à raconter nos vies, Séverine et Antje avec qui j'ai pu aller découvrir les confins du bout du monde en Amérique latine et en Asie, Aurélie et Jimmy toujours à mes côtés, bien qu'ayant opté pour le Soleil du Sud, Marie P., Dorothée H. et Céline B. pour qui la vie est toujours bonne à croquer, Greg l'éternel boute-en-train, Claire D. pour nos innombrables virées dans la "forêt lointaine", Audrey M. pour nos interminables soirées autour du Trivial Pursuit! (la prochaine fois, je gagne!), Kévin pour son originalité à toutes épreuves et Gwen l'australienne, formidable photographe du monde entier (à quand le grand reportage ensemble?). N'oublions pas non plus tou(te)s mes ami(e)s du handball qui m'ont permis de passer des moments formidables et de me changer les idées, le temps d'un match, d'un footing ou d'une soirée bien arrosée !...*

*Une grosse pensée à toute ma famille, mes petites bulles d'air vosgiennes et haut-saônoises qui me font respirer et me permettent de me ressourcer de temps en temps !*

*Enfin, je remercie intensément mes parents qui ont largement contribué à l'aboutissement de ce travail, grâce à l'amour et à l'écoute qu'ils m'ont portés, et ma sœur Julie que j'apprécie tout particulièrement et qui a toujours trouvé les mots justes et sincères pour me reconforter quand le moment s'en est fait sentir.*

*Après 5 années passées au sein de l'IGBMC, le moment est venu de clore cette parenthèse enrichissante en tous points pour m'envoler vers d'autres horizons.*

*Merci à tous ceux que je n'ai pas su nommer et avec qui j'ai échangé une parole au détour d'un couloir! Et mon dernier grand merci va à Maïté, à Héléne, à Evelyne et à toutes les dames de l'accueil qui agrémentent tous les jours notre arrivée à l'institut d'un merveilleux sourire !*

*Ce travail a été soutenu par le ministère de la Recherche et de l'Enseignement Supérieur et par la Fondation pour la Recherche Médicale qui m'a soutenue financièrement pour cette quatrième année de thèse.*



# *ABREVIATIONS*



## Abréviations et Acronymes

### A

**ADN** : Acide désoxyribonucléique

**AF-1** : Activation Function 1

**AF-2** : Activation Function 2

**ApoE** : apolipoprotéine E

**AR** : Acide rétinoïque

**ARN** : Acide ribonucléique

**AR tout-trans** : Isomère tout-trans de l'acide rétinoïque

**AR 9-cis** : Isomère 9-cis de l'acide rétinoïque

**AR 13-cis** : Isomère 13-cis de l'acide rétinoïque

**ARAT** : acyl-CoA-rétinol acyltransférase

### B

**BRE** : TFIIB Recognition Element

### C

**CAGE** : Cap Analysis of Gene Expression

**CBP** : CREB Binding Protein

**CRABP** : cellular retinoic acid binding protein

**CRALBP** : cellular retinaldehyde binding protéin

**CRBP** : Cellular retinol-binding protein

**CTD** : domaine C-terminal

### D

**DBD** : DNA Binding Domain

**DCE** : Downstream Core Element

**DNMT** : DNA methyltransferase

**DPE** : Downstream Promoter Element

**DR** : Direct Repeat

### E

**EGF**: Epidermal Growth Factor

### G

**GCN5**:General Control Nonderepressible5

**GFP**: Green Fluorescent Protein

**GNATs** : Gcn5-related acetyltransferases)

**GTF** : General Transcription Factors

### H

**HAT** : Histone Acetyltransferase

**HD** : Huntington Disease

**HDAC**: Histone Desacetylase

**HFD** : Histone Fold Domain

**Hi** : Huntingtine

**HMT** : Histone Methyl Transférase

**14-HRR** : 14-hydroxy-4,14-rétro-rétinol

**HRE**: Hormone Response Element

**HSP** : heat shock protein

## **I**

**INR** : Initiator Element

**IR** : Inverted Repeat

## **L**

**LBD** : Ligand Binding Domain

**LCR** : Locus Control Region

**LPL** : Lipoprotéine lipase

**LRAT** : lécithine rétinol-acyl transférase

## **M**

**MEFs** : Mouse Embryonic Fibroblasts

**MMP-2** : Matrix Metalloproteinase 2

**MTE** : Motif Ten Element

**MPT** : Modifications post-traductionnelles

## **N**

**NCoR** : Nuclear hormone receptor co-repressor

**NRIP1**: Nuclear receptor interacting protein 1

## **P**

**PA** : pré-albumine

**PIC** : Pre-initiation Complex

**PKA** : Protéine kinase A

**PKC** : Protéine kinase C

**PRAME** : preferentially expressed antigen in melanoma

**PRC1** : Polycomb repressive complex 1

**PML** : ProMyelocytic Leukemia

**Pu** : Purine

**Py** : Pyrimidine

## **R**

**Ral** : Rétinaldéhyde

**RALDH** : Retinaldehyde déhydrogenase

**RAR** : Retinoic Acid Receptor

**RARE** : Retinoic Acid Response Element

**RBP** : Retinol Binding Protein

**RBP2** : RNA polymerase II gene

**RDH** : rétinol déhydrogénase

**RE** : esters de glycérol

**RIP140** : Receptor interacting protein of 140 kDa

**RN** : Récepteurs nucléaires

**RNAP II** : RNA Polymerase II

**Rol** : Rétinol

**RXR** : Retinoid X Receptor

## **S**

**SAGA** : Spt-Ada-GCN5 Acetyltransferase

**SALSA** : SAGA altered Spt8 absent

**SLIK** : SAGA-like complex

**SMRT** : silencing mediator for retinoic acid

**Sp/XKLF** : specificity protein/Krüppel-like factor

**Sp1** : Specify protein 1

**STAGA** : SPT3-TAF9-GCN5L Acetylase

**SWI/SNF**: faulty mating types witching sucrose non fermenting

## **T**

**TAFs** : TBP Associated Factors

**TAF4** : TBP Associated Factor 4

**TAFH**: TAF Homology domain

**TBP** : TATA Box Binding Protein

**TFIIA** : Transcription Factor IIA

**TFIIB** : Transcription Factor IIB

**TFIID** : Transcription Factor IID

**TFIIE** : Transcription Factor IIE

**TFIIF** : Transcription Factor IIF

**TFIIH** : Transcription Factor IIF

**TFTC** : TBP-freeTAF<sub>II</sub>s

containing complex

**TLF** : TBP-like Factor

**TLP** : TBP-like protein

**TR** : Thyroid hormone receptor

**TRF2** : TBP-related factor 2

**TS** : Thermosensible

**TSS** : Transcription Start Site

**TTR** : transthyréine

## **U**

**UAS** : Upstream Activating Sequence

## **V**

**VAD** : Vitamin A-deficiency

**VDR** : Vitamin D receptor

**VP16** : herpes simplex virus protein 16





*TABLE DES  
MATIÈRES*



# Table des Matières

<i>REMERCIEMENTS</i>	7
<i>ABREVIATIONS</i>	11
<i>TABLE DES MATIERES</i>	17
<i>INDEX DES FIGURES, CD</i>	25
<i>MOTS AUX LECTEURS</i>	33
<i>AVANT- PROPOS</i>	35
<i>CHAPITRE 1 : La transcription chez les Eucaryotes supérieurs</i>	
I. Un socle nécessaire à l'initiation de la transcription.....	45
A. Le promoteur minimal .....	45
1. Une plate-forme d'assemblage du PIC .....	45
2. Une redéfinition du promoteur minimal .....	46
B. Les éléments constitutifs du promoteur minimal .....	48
1. La boîte TATA.....	48
2. L'INR, élément initiateur .....	49
3. Le DPE, site de régulation en aval du TSS .....	50
4. Le BRE, élément de reconnaissance à TFIIB .....	50
5. Le MTE, motif à 10 éléments .....	51
6. Le DCE, élément de promoteur en aval du TSS .....	51
C. Les îlots CpG .....	52
II. Les acteurs impliqués dans l'initiation de la transcription.....	53
A. Un acteur essentiel aux gènes eucaryotes : l'ARN polymérase II.....	54
1. Les ARN polymérases chez les eucaryotes.....	54
2. L'ARN polymérase de type II.....	54
B. Les facteurs généraux de la transcription .....	58

C.	Les modèles d'assemblage des différents facteurs .....	58
1.	Le modèle séquentiel .....	60
2.	Modèle de l'holoenzyme.....	65
3.	Intermédiaire entre les deux modèles d'assemblage du PIC.....	65
D.	Le complexe Médiateur .....	67
III.	Les séquences régulatrices .....	68
A.	Les séquences régulatrices proximales.....	69
1.	Le motif CCAAT .....	69
2.	Le motif riche en GC .....	70
3.	Les éléments de réponse HRE .....	71
B.	Les séquences régulatrices distales.....	72
IV.	Les activateurs, les répresseurs et les cofacteurs de la transcription.....	74
A.	Le complexe Médiateur .....	75
B.	Les cofacteurs dérivés de l'UAS .....	76
C.	Les récepteurs nucléaires.....	76
1.	Les différentes catégories de récepteurs nucléaires .....	77
2.	Structure des récepteurs nucléaires .....	79
3.	Les mécanismes de la régulation transcriptionnelle .....	82
D.	Les modifications de la chromatine.....	88
1.	Les complexes de remodelage .....	89
2.	Les histones acétyltransférases (HAT) .....	90
3.	Phosphorylation des histones et activation de la transcription .....	97
4.	Méthylation des histones.....	98
5.	Ubiquitination des histones et activation de la transcription .....	99
6.	Les complexes.....	99

*CHAPITRE 2 : Zoom sur un cofacteur multipartite particulier TFIID*

I.	Vue d'ensemble du facteur TFIID .....	107
A.	TFIID et sa structure modulaire.....	107
1.	Observation du facteur TFIID en microscopie électronique.....	107
2.	La plasticité de la structure de TFIID .....	109
3.	Bilan des changements conformationnels de TFIID.....	110
B.	La composition intrinsèque du facteur TFIID .....	112
1.	TBP et les TAFs, des sous-unités de TFIID .....	112
2.	Les TAFs et le domaine de repliement de type histone .....	113
C.	L'assemblage de TFIID .....	117
1.	La masse moléculaire du complexe natif.....	117
2.	La stœchiométrie des sous-unités de TFIID .....	118
3.	Dynamique d'assemblage des TAFs au sein du TFIID .....	121
D.	Variabilité du complexe TFIID .....	124
1.	Les variants des TAFs.....	125
2.	Les facteurs apparentés à TBP .....	127
3.	Les complexes transcriptionnels sans TBP.....	128
II.	Les TAFs et leur véritable fonction dans la régulation des gènes.....	143
A.	Les TAFs assurent la stabilité de TFIID sur le promoteur .....	143
1.	L'occupation des TAFs sur les promoteurs .....	144
2.	Un recrutement spécifique des TAFs.....	145
B.	TFIID, une cible idéale pour les activateurs.....	146
1.	Les liens entre activateurs et TAFs .....	146
2.	Les TAFs, médiateurs de l'activation de la transcription .....	147
3.	Mécanismes d'action des TAFs dans l'activation de la transcription.....	148
4.	Changement de conformation de TFIID .....	149

C.	Les TAFs et la répression de la transcription .....	150
D.	Modifications de la structure de la chromatine par les TAFs.....	151
E.	Les TAFs, des protéines essentielles .....	152
F.	TAF4 et la voie TGF $\beta$ .....	153
III.	TAF4, à la croisée des voies de transcription et de la cancérologie.....	157
A.	Découverte de TAF4 .....	157
B.	Structure du facteur de transcription TAF4 ( CRI, CRII ).....	157
1.	Le domaine N-terminal .....	157
2.	La région conservée CRI.....	158
3.	La région conservée CRII .....	160
4.	Les domaines centraux riches en glutamine .....	163
C.	TAF4 et les maladies neurodégénératives .....	163
D.	TAF4 renforce l'activation transcriptionnelle .....	164
E.	TAF4 intervient dans la répression de la transcription.....	167
F.	TAF4 et la cancérogenèse .....	169

### *CHAPITRE 3 : Les rétinoïdes, des acteurs anti-cancérigènes*

I.	Aux origines de l'acide rétinoïque .....	175
A.	Historique .....	175
1.	Les travaux de Hopkins .....	175
2.	Le principe actif .....	176
B.	Du rétinol à l'acide rétinoïque .....	176
1.	La transformation du rétinol dans l'organisme .....	176
2.	Les rétinoïdes bioactifs .....	177
II.	Métabolisme des rétinoïdes .....	179
A.	Provenance de la vitamine A .....	179
B.	Devenir du rétinol dans l'organisme .....	180

1.	L'absorption du rétinol .....	181
2.	Le transport des rétinoïdes aux cellules cibles.....	181
3.	Les rétinoïdes dans le compartiment cellulaire.....	183
<b>III.</b>	<b>Les récepteurs des rétinoïdes.....</b>	<b>185</b>
A.	Composition de la famille des récepteurs des rétinoïdes.....	185
B.	Spécificité du ligand .....	186
C.	Éléments de réponse .....	186
D.	Possibilités de dimérisation .....	188
1.	Dimères impliquant RXR .....	188
2.	Dimères impliquant RAR .....	189
3.	Le rôle des RARs et RXRs et les phénotypes de leurs mutants.....	190
<b>IV.</b>	<b>Le rôle des rétinoïdes en cancérologie .....</b>	<b>192</b>
A.	La vitamine A, une molécule miracle ?.....	192
B.	Effets anti-néoplasiques des rétinoïdes.....	192
C.	Comment les rétinoïdes peuvent-ils inhiber la carcinogenèse ?.....	193
1.	Les rétinoïdes et leur action anti-tumorale.....	194
2.	L'AR et son effet anti-prolifératif.....	194
3.	L'AR et la différenciation .....	195
D.	Les pathologies associées aux récepteurs RAR.....	195
1.	La leucémie aigüe promyélocytaire .....	195
	-Acide rétinoïque et arsenic.....	199
2.	Le récepteur RAR $\beta$ et les cancers .....	200
E.	Les pathologies associées aux récepteurs RXR.....	203
F.	Les autres pathologies traitées par les rétinoïdes .....	204
1.	AR et médulloblastome.....	204
2.	AR et sarcomes .....	204

<i>Résultats</i> .....	209
<i>Revue (211), Manuscrit 1 (227), Manuscrit 2 (259), Publication DEA (Annexe) (301)</i>	
<i>Discussion et Perspectives</i> .....	313
I. TAF4 et TAF4b, deux facteurs essentiels ? .....	315
II. Un complexe isolé à partir de TAF4 et dépourvu de TBP .....	318
A. Un sous-complexe de TFIID ?.....	318
1. Combien y-a-t-il de TAFs ? .....	318
2. Que représente ce complexe à 5 TAFs ?.....	318
3. Une étude stoechiométrique des composants de ce complexe.....	320
B. Des questions en suspens.....	321
III. Le complexe à 5 TAFs, un co-activateur transcriptionnel ?.....	321
A. L'initiation de la transcription .....	321
B. Quels sont les promoteurs cibles pour ce complexe ?.....	323
1. Variation d'expression de gènes entre les cellules C3 taf4 <sup>-/-</sup> et C1.....	323
2. Comment déterminer des gènes-cibles directs ?.....	323
C. TAF4 et les modifications post-traductionnelles des histones .....	324
IV. Taf4, un nouveau gène suppresseur de tumeurs.....	325
A. L'inactivation de TAF4 et la transformation phénotypique .....	325
B. Induction de sarcomes .....	325
C. Un nouveau mode d'action de l'acide rétinoïque ?.....	325
D. Une voie de signalisation spécifique entraînant la tumorigenèse.....	326
E. Une piste en cours ?.....	327
V. Elargissement de l'étude sur la tumorigenèse induite par les fibroblastes TAF4 <sup>-/-</sup> .....	328
VI. Vers la recherche clinique ? .....	329
<i>Bibliographie</i> .....	331



*INDEX des*  
*FIGURES*



## Index des Figures

Figure 1: De l'ADN aux chromosomes mitotiques.....	37
Figure 2 : Décondensation de la chromatine .....	38
Figure 3 : Les six points essentiels de régulation de l'expression génique .....	39
Figure 4 : La balance dans la sélection du brin est influencée par le niveau de chaque transcrit .....	46
Figure 5 : Les promoteurs classés selon la distribution de leur TSS chez les Mammifères .....	47
Figure 6 : Représentation schématique des éléments de séquence du promoteur .....	48
Figure 7: Modèle illustrant la liaison d'une partie de TFIID au promoteur minimal .....	50
Figure 8 : Méthylation de l'ADN au niveau des îlots CpG .....	52
Figure 9: Vue d'ensemble du complexe de préinitiation de la transcription. ....	53
Figure 10: Structure aux rayons X de la RNA Pol II .....	54
Figure 11: Structure de l'ARN polymérase II de la Levure à 2,8 Å de résolution. ....	55
Figure 12: Les 3 états du complexe d'élongation de la transcription de la Pol II. ....	56
Figure 13: Importance du CTD dans la régulation de la transcription . ....	57
Figure 14 : GTFs et Pol II assemblés sur le promoteur du gène à transcrire;.....	58
Figure 15: Structure du complexe TBP lié à l'élément TATA. ....	60
Figure 16: Le complexe TFIIA/TBP/ADN .....	62
Figure 17: Structure du complexe TFIIIB-TFIIA-TFIID-ADN .....	63
Figure 18: Structure de l'hétérodimère RAP30/RAP74 de TFIIIF .....	64
Figure 19: Structure du complexe TFIIH .....	65
Figure 20: Les modèles d'assemblage du complexe d'initiation de la transcription. ....	66
Figure 21: Médiateur humain.....	67
Figure 22: Représentation schématique du promoteur et des séquences régulatrices .....	68
Figure 23 : Liaison du facteur Sp1 de la famille Sp/XKLF .....	70
Figure 24 : Eléments de réponse .....	71
Figure 25 : Fixation de récepteurs nucléaires .....	72
Figure 26: Modèle d'interaction entre le promoteur .....	73
Figure 27 : Représentation schématique d'un activateur proximal. ....	74
Figure 28 : Rôle des activateurs de la transcription .....	75
Figure 29 : L'activation de la transcription via le complexe du Médiateur .....	76
Figure 30 : Modèle hypothétique concernant l'évolution de la superfamille des RN .....	77
Figure 31 : Les différentes régions (A à F) et les domaines caractéristiques .....	79
Figure 32 : Diagramme en rubans .....	81
Figure 33 : Les récepteurs stéroïdiens et leur rôle de cofacteurs transcriptionnels. ....	83
Figure 34 : Mécanisme de répression de la transcription .....	85
Figure 35 : Mécanisme d'activation de la transcription grâce au recrutement de co-activateurs. ....	85
Figure 36: Modèle de répression par Prame de la voie de signalisation à l'acide rétinoïque .....	86

Figure 37 : Représentation schématique des domaines fonctionnels .....	87
Figure 38 : Le bourgeonnement des nucléosomes .....	89
Figure 39 : Rôle des HDACs et des HATs dans l'acétylation des histones .....	90
Figure 40 : Alignement des protéines de la superfamille GNAT .....	92
Figure 41 : Organisation des protéines CBP/p300 .....	93
Figure 42 : Rôle de CBP/p300 dans l'activation transcriptionnelle .....	94
Figure 43: Alignement des protéines de la famille MYST .....	95
Figure 44: Les différents types d'activités coactivatrices .....	100
Figure 45 : Aperçu de la transcription divergente .....	103
Figure 46: Le facteur TFIID associé à TFIIA, TBP et TFIIB .....	108
Figure 47: Modèles tridimensionnels du complexe hTFIID à 35 Å de résolution .....	108
Figure 48: Conformations alternatives du facteur TFIID endogène humain.....	109
Figure 49: Flexibilité de la structure de TFIID .....	110
Figure 50 : Reconstruction 3-D des interactions c-Jun-IID ; Sp1-IID et p53-IID.....	111
Figure 51: Dimère H2A-H2B.....	113
Figure 52 : Hétérodimère hTAF4-hTAF12.....	114
Figure 53 : Comparaison de hTAF4 et hTAF12 aux histones H2A and H2B.....	114
Figure 54 : Comparaison de tétramère de hTAF4-hTAF12 avec le nucléosome .....	115
Figure 55: Les domaines de repliement des histones et leurs homologues dans les TAFs . .....	116
Figure 56 : Détermination de la masse moléculaire de yTFIID .....	117
Figure 57: Organisation structurale de yTFIID à 23 Å de résolution .....	121
Figure 58: Modèle schématique d'assemblage du TFIID .....	122
Figure 59 : Analyse de la stabilité de TFIID in vivo.....	123
Figure 60: La formation de l'holo TFIID .....	124
Figure 61 : Représentation schématique de différents complexes de TAFs .....	126
Figure 62 : Facteurs appartenant à la famille TBP codés par différents génomes.....	128
Figure 63 : Comparaison des complexes sans TBP à TFIID.....	129
Figure 64 : Structure du complexe SAGA résolue par microscopie électronique.....	131
Figure 65 : SAGA, co-activateur transcriptionnel.....	133
Figure 66 : Représentation schématique des complexes TFIID et TFTC .....	139
Figure 67 : TFTC et TFIID, structures tridimensionnelles des complexes .....	140
Figure 68 : Contacts entre activateurs et composants de la machinerie basale. ....	143
Figure 69 : Occupation du promoteur par TBP et les TAFs.....	145
Figure 70 : Les activateurs de la transcription contactent la machinerie basale .....	147
Figure 71 : Différents stades de l'expression des gènes,.....	150
Figure 72 : Les différentes activités enzymatiques de la sous-unité TAF1.....	151
Figure 73 : Génération d'allèles TAF4 porteurs de sites LoxP .....	154
Figure 74 : Croissance des fibroblastes taf4 <sup>-/-</sup> .....	155
Figure 75 : Induction de la voie de signalisation de TGFβ lors de l'inactivation de TAF4 .....	155

Figure 76 : Structure de la protéine hTAF4 .....	157
Figure 77 : Structure globale du domaine hTAF4-TAFH .....	158
Figure 78 : alignement de TAF4 .....	161
Figure 79 : Perturbation de la transcription par un mutant de la protéine huntingtine .....	164
Figure 80 : Répression de la transcription via la séquestration de TAF4.....	168
Figure 81 : Galerie des vues les plus représentatives de yTFIID. ....	171
Figure 82 : Devenir du rétinol et du bêta-carotène dans l'organisme.....	176
Figure 83 : Les rétinoïdes bioactifs les plus étudiés : l'AR-T, l'Ar-9cis et l'AR-13cis .....	177
Figure 84 : Les rétinoïdes bioactifs chez les Vertébrés supérieurs.....	178
Figure 85 : Métabolisme des rétinoïdes .....	180
Figure 86 : Représentation en ruban de la structure tridimensionnelle de l'holo-RBP bovine .....	182
Figure 87: Les voies de signalisation impliquées dans l'anabolisme et le catabolisme de l'AR.....	183
Figure 88: Métabolisme des rétinoïdes dans la cellule .....	184
Figure 89 : Les récepteurs RXR et RAR .....	185
Figure 90 : Exemple de liaison des récepteurs des rétinoïdes sur des demi-sites en répétition.....	187
Figure 91 : Structure générale de l'hétérodimère LXR $\alpha$ -RXR $\beta$ .....	189
Figure 92 : Augmentation de la liaison à l'ADN du complexe PML/RAR $\alpha$ .....	196
Figure 93 : Recrutement des HDAC par PML-RAR $\alpha$ .....	197
Figure 94 : Interaction entre les domaines de RAR $\alpha$ et NcOR .....	197
Figure 95 : Recrutement des protéines HDAC par le complexe PML/RAR $\alpha$ en l'absence de ligand. ....	198
Figure 96 : Expérience d'ARN interférence contre TAF4b dans les cellules C3 et C1. ....	317
Figure 97 : Double IP d'un extrait nucléaire de cellules de Métazoaires. ....	319
Figure 98 : Exemple de test de transcription in vitro .....	322
Figure 99 : Vaisseaux sanguins approvisionnant la tumeur d'une souris nue .....	329



## Contenu du CD

Un **disque compact** a été joint à ce manuscrit. Il se trouve en **3<sup>ème</sup> page de couverture**.

Dans ce CD se trouvent des données supplémentaires concernant le travail de thèse ayant mené au manuscrit n°2.

**1) Un tableau Excel** regroupe toutes les données issues de l'analyse des **puces Affymétrie**. Lors de ces expériences, 3 types d'ADNc ont été hybridés (en triplicatas soit 9 échantillons). Ils proviennent de fibroblastes embryonnaires murins taf4<sup>-/-</sup> traités pendant 0h, 12h et 72h avec de l'acide rétinoïque tout-trans à la concentration de 10<sup>-6</sup>M finale.

Ainsi, les onglets situés en bas à gauche du tableau Excel permettent de retrouver les différences d'expression des gènes de la puce pour les comparaisons suivantes : RA0h-RA12h ; RA0h-RA72h ; RA12h-RA72h ; All Data

**2) Les expériences de blessure-cicatrisation** (Wound-Healing) suivies en temps réel sur 30 heures ont permis la réalisation de films en accéléré de 20sec chacun.

4 types cellulaires ont été suivis : les cellules TAF4<sup>lox/-</sup> (C1) et les cellules TAF4<sup>-/-</sup> (C3) traitées ou non pendant 3 jours avec RA (10<sup>-6</sup>M).

Les films 1 à 3 correspondent à la migration des C13, les films 4 à 6 à celle des C13+RA, les films 7 à 9 à celle des C11 et les films 10 à 12 à celle des C13+RA.

**3) Les expériences de suivi individuel de la migration des noyaux cellulaires** (Tracking) ont également été réalisées. Sur ce CD, deux films (n°13 et 14) montrent les chemins parcourus par ces noyaux, donnant une idée de la manière dont leur vitesse moyenne de migration a été calculée.





## Mots aux lecteurs

Quelques mots tout d'abord pour présenter la manière dont a été organisée ce manuscrit de thèse afin de vous en faciliter la lecture...

Le laboratoire du Dr Davidson au sein duquel j'ai travaillé pendant ma thèse s'intéresse à la transcription des gènes de classe II et plus particulièrement à l'étape d'initiation, dans la mesure où cette étape joue un rôle majeur dans la régulation de l'expression des gènes.

L'initiation de la transcription nécessite la formation d'un complexe de pré-initiation autour du site de démarrage de la transcription. Parmi les facteurs présents dans ce complexe, se trouve le facteur TFIID, un facteur multiprotéique composé de TBP (TATA binding protein) et de 14 TAFs (TBP-associated factors). Nos études se concentrent sur l'organisation et la structure de TFIID et le rôle des TAFs *in vivo* et *in vitro*.

Au sein du laboratoire, nous avons travaillé sur TAF4, une protéine de 135kDa qui présente un intérêt particulier du fait de son rôle présumé dans la signalisation par l'AMPc (Asahara et al, 2001), les récepteurs nucléaires (notamment RAR) (Mengus et al, 1997), et les stress kinases.

Afin de mieux comprendre la fonction de TAF4, le gène correspondant a été inactivé dans des modèles cellulaires *in vitro* et dans des processus physiologiques *in vivo* chez la souris. La perte de TAF4 dans les fibroblastes embryonnaires murins (MEFs) induit une activation de la voie de signalisation TGF-beta (Mengus et al., 2005). Elle n'entraîne pas d'arrêt du cycle cellulaire, ni même l'apoptose des cellules, contrairement à ce qui est observé lors de l'inactivation d'autres TAFs dans différentes lignées cellulaires. Au contraire, les cellules *Taf4*<sup>-/-</sup> deviennent capables de croissance autocrine, grâce à cette activation constitutive de la voie TGF-beta.

Lors de mon projet de thèse, l'étude du facteur de transcription TAF4 a d'abord été envisagée, au préalable, d'un point de vue transcriptionnel. Il s'agissait de déterminer les rôles respectifs de TAF4 et de TAF4b au sein de ces MEFs en élaborant des versions étiquetées de ces protéines afin de détecter leurs partenaires respectifs et leurs gènes cibles par des techniques d'immunoprécipitation, de spectrométrie de masse et de CHIP.

Très vite, l'étude de TAF4 a révélé la présence d'un nouveau complexe différent de TFIID sur lequel je me suis concentrée. Mon projet de thèse s'est enrichi par la suite d'une composante en cancérologie, dès lors que l'étude approfondie des MEFs *Taf4*<sup>-/-</sup> a révélé l'importance de la protéine TAF4 en tant qu'anti-oncogène. Aussi, mon projet de thèse comporte deux parties distinctes n'ayant en commun que l'importance du facteur de transcription TAF4.

Dans la seconde partie de mon travail, j'ai découvert que le traitement à l'acide rétinoïque (AR) de ces MEFs *Taf4*<sup>-/-</sup> conduit à une réversion de la tumorigénèse. En d'autres mots, l'acide rétinoïque supprime la formation de tumeurs induites par l'injection des MEFs *Taf4*<sup>-/-</sup> dans des souris immuno-déficientes mais via un mécanisme original encore non élucidé.

Afin de pouvoir présenter l'ensemble de mes résultats dans leur contexte, j'ai opté pour l'élaboration d'une introduction en 3 parties.

Le premier chapitre est un rappel des connaissances fondamentales en transcription qui nous permettront, à la fin du manuscrit, de discuter des implications de la présence de ce nouveau complexe au sein de la transcription.

Le second chapitre fait un état des lieux non exhaustif des connaissances actuelles concernant le facteur TFIID et sa sous-unité TAF4, en vue de mieux comprendre la double dimension de TAF4, cette protéine située à la croisée des voies de la transcription et de la cancérologie.

Le troisième chapitre aborde la voie de signalisation des rétinoïdes et leur implication au sein du domaine de la cancérologie. Il a été conçu pour donner des pistes de réflexions éventuelles quant au mécanisme d'action de l'AR dans la réversion tumorigénique.

Je vous souhaite une bonne lecture de ce manuscrit.

Emilie CLER

# *AVANT-PROPOS*



## La transcription chez les eucaryotes supérieurs

L'**ADN** est le support de l'information génétique. Il constitue le génome des êtres vivants. La cellule en est l'unité de base. Au cours de l'évolution sont apparus deux grands types d'organismes cellulaires décrits par Schwann et Schleiden en 1838. D'un côté, les cellules procaryotes, des êtres unicellulaires, dépourvus de noyau et bordés d'une membrane plasmique. De l'autre, les **cellules eucaryotes**, au sein desquelles l'ADN est compartimenté dans un noyau, délimité par une membrane nucléaire.

La **transcription** est un processus biologique cellulaire **ubiquitaire** qui consiste en la copie des régions dites codantes de l'ADN en molécules d'**ARN**. Car si la molécule d'ADN est le **support universel de l'information génétique**, ce sont les molécules d'ARN qui sont reconnues par la machinerie de traduction et qui seront traduites en **séquences protéiques**.

Chez les organismes eucaryotes, l'ADN atteint une telle taille (3 milliards de paires de bases pour le génome humain, équivalent à une longueur de 2 mètres) qu'il nécessite d'être compacté. C'est ainsi que ces cellules ont développé le **système chromatinien**, qui confère à l'ADN un très haut degré de compaction (Fig.1).

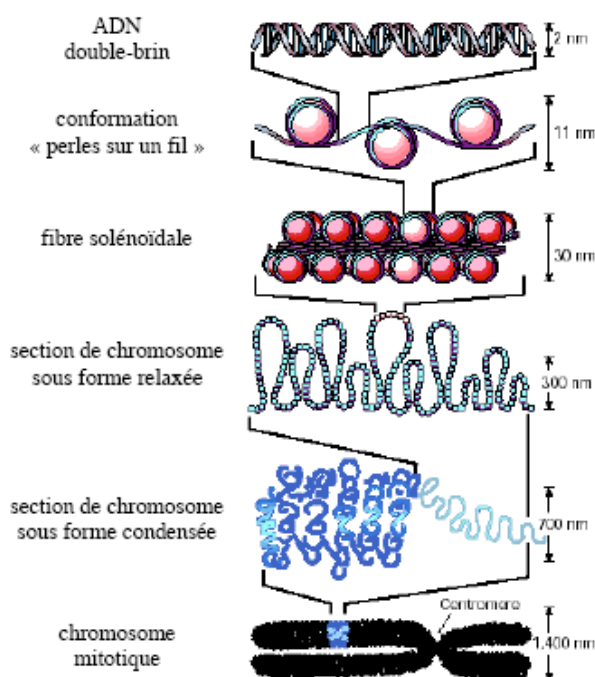


Figure 1: De l'ADN aux chromosomes mitotiques : la compaction de la chromatine (Felsenfeld and Groudine, 2003)

Pour ce faire, les deux longs filaments de nucléotides constituant la double hélice d'ADN vont se condenser autour de protéines spécifiques : les histones, eux-mêmes associés en nucléosomes. Plusieurs niveaux de compaction de la structure chromatinienne vont alors se mettre en place, allant de la structure en collier de perles à la section de **chromosome** en passant par la fibre solénoïdale (Fig.1).

Pendant longtemps, la **chromatine** n'a été considérée que comme une nécessité structurale imposée par un accroissement de la complexité et de la longueur du génome. Or, de nombreuses études menées ces dernières années tendent à démontrer que la chromatine n'est pas une structure inerte, un simple système d'empaquetage de l'ADN. Grâce aux nombreuses modifications post-traductionnelles des **histones**, ses composants protéiques majeurs, elle joue un rôle de premier plan dans la régulation de nombreux processus cellulaires impliquant des changements de compaction de l'ADN.

La **compaction de l'ADN** eucaryote influence beaucoup l'accessibilité de l'ADN en tant que matrice de la transcription, l'expression des gènes dans les cellules eucaryotes étant dépendante de la structure de la chromatine. Lorsque la chromatine est décondensée, elle devient alors accessible aux facteurs de transcription (Fig.2).

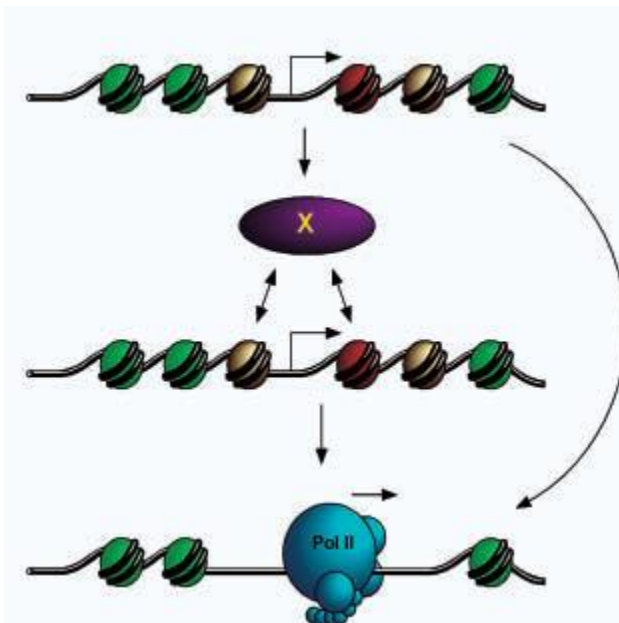


Figure 2 : Décondensation de la chromatine (Guillemette and Gaudreau, 2006)

Un facteur X modifie l'architecture du promoteur et prépare le gène à l'activation de la transcription. Lors de l'activation du gène, les nucléosomes sont rapidement enlevés pour laisser la place à l'ARN polymérase II (Pol II) et à ses facteurs associés.

Les séquences codantes des gènes pourront alors être soumises à la transcription en vue d'obtenir des ARN pré-messagers. Ceux-ci seront ensuite épissés en **ARN messagers**, puis traduits en **protéines** au sein du cytoplasme, via le ribosome. Les protéines sont des molécules indispensables au bon fonctionnement de notre organisme.

Elles assurent des fonctions architecturales, enzymatiques, motrices, de communication ou encore de transport au sein de la cellule. Pour devenir active, une protéine doit adopter une structure tridimensionnelle, souvent à l'aide de protéines chaperonnes.

Mais cette synthèse protéique doit être soumise à une **régulation**, car la plupart des protéines n'assurent une fonction nécessaire qu'à un instant donné et dans un ensemble de cellules déterminé, même si l'information dont elles sont issues est présente dans toutes les cellules. Par ailleurs, l'expérience a montré que seule une faible partie du génome est en cours de transcription à un instant donné dans une cellule. Les étapes intermédiaires entre le gène et la protéine constituent donc une opportunité de choix pour la régulation de l'expression des différents gènes (Fig.3).

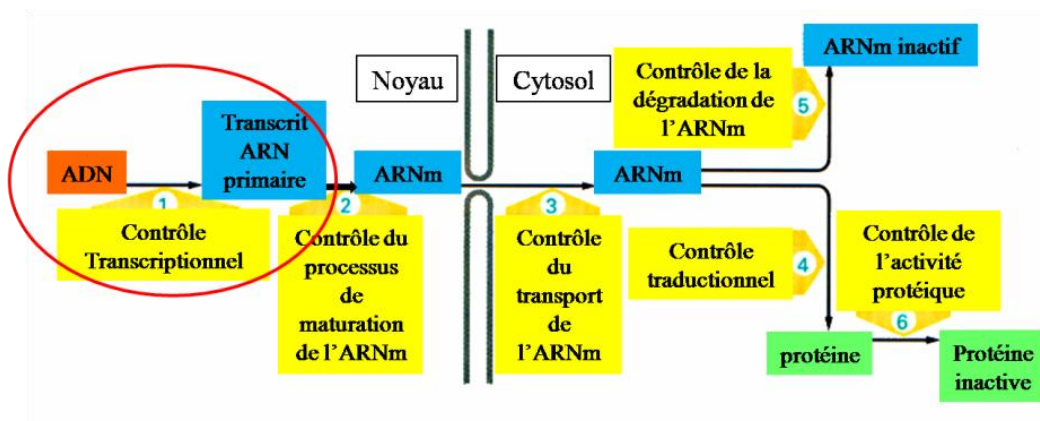


Figure 3 : Les six points essentiels de régulation de l'expression génique (d'après M. Denis, Université de Rennes)

Parmi les 6 points de contrôle de régulation de l'expression des gènes, l'étape de la **transcription** ou synthèse de l'ARN à partir d'une matrice d'ADN, en constitue le premier point. Elle se fait grâce à des enzymes : les ARN polymérase ADN dépendantes et est divisée en trois grandes étapes : l'**initiation** (synthèse des premiers nucléotides), l'**élongation** (allongement de la chaîne d'ARN) et la **terminaison** (libération du pré-ARN messager et départ de la machinerie de transcription de l'ADN).

Avec l'étude de l'opéron lactose, François Jacob, Jacques Monod et André Lwoff ont été les premiers scientifiques à décrire un système de régulation de la transcription des gènes chez les Procaryotes. Ils proposent l'existence de deux classes de gènes qu'ils différencient par leur fonction : les gènes structuraux et les gènes régulateurs. C'est à partir de ces travaux qu'est né le concept de la régulation génique. (Prix Nobel de physiologie et médecine en 1965).

Chez les Eucaryotes, il existe trois classes de gènes :

- les **gènes de classe I** codent pour les ARN ribosomiaux qui constituent les ARN les plus représentés dans la cellule. Ces ARNs 18S et 28S sont produits à partir d'un transcrit unique.
- les **gènes de classe II** codent pour les ARN messagers (ARNm) qui peuvent être traduits en protéines, les micro ARN et les ARN à petits noyaux (snRNA).
- les **gènes de classe III** codent pour l'ARN ribosomal 5S et les ARN de transfert (ARNt).

Le laboratoire au sein duquel j'ai travaillé pendant ma thèse s'intéresse à la transcription des **gènes de classe II** et plus particulièrement à **l'étape d'initiation**, dans la mesure où cette étape joue un rôle majeur dans le choix des gènes à transcrire au cours des différents processus cellulaires.

L'initiation de la transcription débute par l'assemblage de complexes multiprotéiques sur le promoteur. Le complexe **TFIID** semble reconnaître précocement le promoteur. Il est constitué de la protéine **TBP** (« *TATA-Binding Protein* ») capable de reconnaître une séquence d'ADN appelée « boîte TATA » d'une part, et d'un assemblage de protéines TAFs (« *class II TBP Associated Factors* »). C'est ce facteur multi-protéique **TFIID** et sa sous-unité **TAF4** qui feront l'objet de ce mémoire.



# *INTRODUCTION GENERALE*



# *CHAPITRE 1*

*L'INITIATION de la TRANSCRIPTION*

*des GENES de CLASSE II*

*chez les EUKARYOTES SUPERIEURS*



# CHAPITRE 1:

## L'INITIATION de la TRANSCRIPTION des GENES de CLASSE II chez les eucaryotes supérieurs

L'initiation de la synthèse de l'ARNm dans les cellules eucaryotes est un processus complexe et hautement régulé. Il nécessite la présence d'une séquence particulière d'ADN double brin appelée **promoteur** sur laquelle vient s'assembler un **complexe de pré-initiation** de la transcription (PIC), composé des facteurs généraux de transcription (GTFs) et d'une enzyme : l'ARN polymérase de type II (Pol II). L'assemblage de ce complexe d'initiation de la Pol II est sujet à de nombreux niveaux de régulation à travers un ensemble varié de facteurs protéiques et de cofacteurs. Des **séquences activatrices proximales et distales** de part et d'autre du site d'initiation de la transcription vont permettre la liaison de **cofacteurs** capables de moduler l'activité transcriptionnelle et d'intervenir dans le remodelage de la chromatine, rendant ainsi l'ADN génomique plus accessible aux GTFs. Au cours de ce chapitre, nous aborderons donc ces éléments clés de l'initiation de la transcription.

### I. Un socle nécessaire à l'initiation de la transcription

#### A. Le promoteur minimal

Le promoteur est la **région d'ADN minimale suffisante** pour assurer la transcription basale *in vitro* grâce à la Pol II, à des niveaux même très faibles, indépendamment de tout activateur. Les promoteurs des gènes de classe II codant pour les protéines s'étendent approximativement sur 40 paires de bases de part et d'autre du TSS

#### 1. Une plate-forme d'assemblage du PIC

Chez les Eucaryotes, le promoteur sert de **plate-forme** pour l'assemblage du PIC, c'est-à-dire pour le recrutement des GTFs et de la Pol II. Le rôle du promoteur est double. Il permet, d'une part, de positionner tous les éléments protéiques nucléaires nécessaires à l'initiation de la transcription, et d'autre part, d'orienter le sens de la transcription du fait de sa structure asymétrique. D'une manière générale, la transcription est fortement contrôlée au

niveau de l'étape d'initiation, c'est-à-dire de la fréquence de démarrage de la Pol II. Les chaînes naissantes d'ARN seront ensuite d'autant plus nombreuses que l'initiation et l'élongation auront été facilitées.

## 2. Une redéfinition du promoteur minimal

### a) Brin sens et brin anti-sens

A l'origine, le mot «promoteur» décrivait une séquence commune à tous les gènes, en amont de la séquence codante. Au vu des connaissances actuelles, il s'agit plutôt à présent d'une région de séquence variable, entourant le **site d'initiation de la transcription** ou TSS (noté +1). Par convention une numérotation négative est utilisée en amont du TSS et positive en aval (Fig.4). D'autre part, le **brin sens** ou brin non-matrice spécifie le brin d'ADN dont la séquence est identique à celle de l'ARN pré-messager. (Inversement, le brin anti-sens ou **brin matrice** correspondrait à un ARN pré-messager anti-sens).

Récemment pourtant, la transcription anti-sens, c'est-à-dire à partir du brin opposé à celui codant pour une protéine, a été découverte au sein du transcriptome des Mammifères (Kiyosawa *et al.*, 2003; Yelin *et al.*, 2003). Ces transcrits anti-sens seraient impliqués dans la dégradation de leurs transcrits correspondants de type sens (Fig.4).

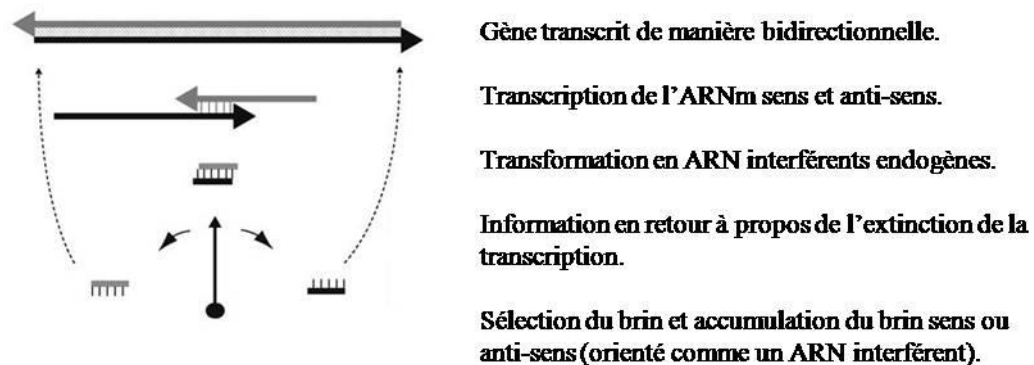


Figure 4 : La balance dans la sélection du brin est influencée par le niveau de chaque transcrit (Werner and Sayer, 2009).

Une analyse globale du transcriptome a montré que la proportion du génome pouvant produire des transcrits dans les deux sens, avait été largement sous-estimée auparavant (Carninci *et al.*, 2005; Chen *et al.*, 2004) et pouvait aller jusqu'à 72% du génome. Les

transcrits anti-sens serviraient donc de liens entre des gènes voisins afin d'en faire des unités transcriptionnelles à part entière (Katayama *et al.*, 2005).

### b) Plusieurs promoteurs pour un même gène

L'équipe de Carninci a montré que la plupart des gènes codant pour des protéines humaines ou murines sont associés à plus d'un promoteur (Carninci *et al.*, 2006). Ces **promoteurs alternatifs** génèrent des exons 5' alternatifs qui peuvent ou non contenir des codons d'initiation. L'usage de promoteurs et de sites de polyadénylation alternatifs contribuent à la complexité du protéome des Mammifères (Carninci *et al.*, 2006).

### c) Promoteur compact ou espacé

Une approche informatique de type CAGE (Cap Analysis of Gene Expression) a révélé que la plupart des promoteurs humains et murins n'avaient pas un TSS spécifique à un endroit précis du génome (Sandelin *et al.*, 2007) mais un ensemble de TSS situés à proximité les uns des autres s'étendant sur 50 à 100 paires de bases (Carninci *et al.*, 2005; Carninci *et al.*, 2006) (Fig.5).

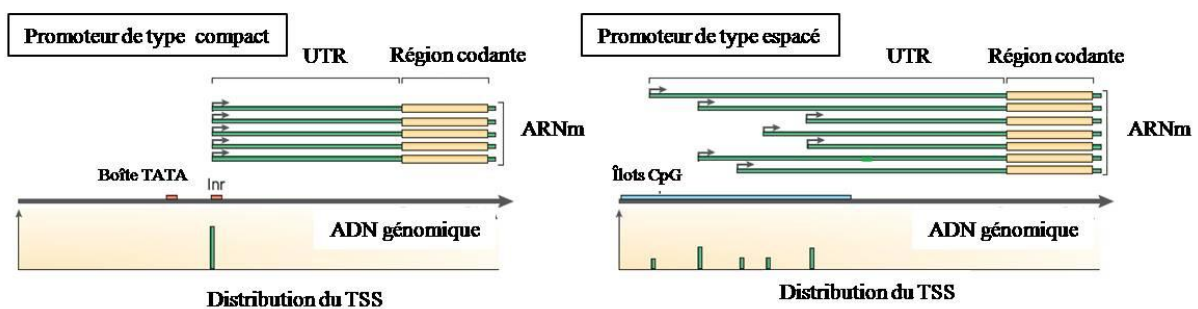


Figure 5 : Les promoteurs classés selon la distribution de leur TSS chez les Mammifères (Sandelin *et al.*, 2007).

Ces observations sont à la base d'un nouveau système de classification des TSS distinguant les promoteurs de type compact (sharp promoters) qui ont un TSS unique bien défini et les promoteurs de type espacé (broad promoters) dont les TSS sont distribués sur une large région (Ahsan *et al.*, 2009). Alors que la première catégorie de TSS corrèle avec les promoteurs ayant une boîte TATA, la seconde catégorie prévaut surtout pour les promoteurs à îlots CpG (Sandelin *et al.*, 2007).

Les promoteurs de type compact semblent être souvent requis pour l'expression de gènes spécifiques à un tissu, alors que les promoteurs de type espacé seraient associés avec des gènes exprimés de manière ubiquitaire.

### B. Les éléments constitutifs du promoteur minimal

Chez les Eucaryotes supérieurs, les promoteurs comprennent une combinaison des éléments suivants: **une boîte TATA** (site de liaison de TBP), un site de liaison à TFIIB (**BRE**), un élément initiateur (**INR**) et un élément de promoteur situé en aval du TSS (**DPE**). L'INR et le DPE semblent coopérer et se substituent parfois à la boîte TATA dans la liaison du complexe TFIID. Par ailleurs, le BRE chez l'Homme est divisé en deux séquences de part et d'autre de la boîte TATA (Fig.6). Chez les Archaeobactéries, il est essentiel à la bonne orientation de la transcription.

#### 1. La boîte TATA

Première structure commune identifiée au sein des promoteurs analysés, **la boîte TATA** contient une séquence consensus TATA(A/T)A(A/T) située de 28 à 34 paires de bases en amont du site d'initiation de la transcription chez les Eucaryotes supérieurs (Breathnach and Chambon, 1981). Cette séquence est reconnue par la protéine **TBP** (TATA Binding Protein) du facteur de transcription TFIID (Fig.6). Cette fixation permettrait la formation du **PIC** et elle pourrait également jouer un rôle dans la détermination du sens de la transcription.

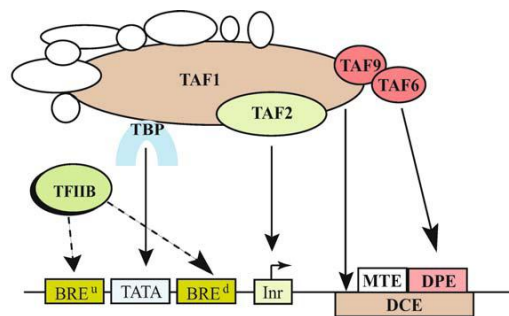


Figure 6 : Représentation schématique des éléments de séquence du promoteur des gènes de classe II et des facteurs de transcription qui s'y lient (Thomas and Chiang, 2006).

Core Promoter Element	Position	Consensus Sequence (5' to 3')	Bound Protein
BRE <sup>u</sup>	-38 to -32	(G/C)(G/C)(G/A)CGCC	TFIIB
TATA	-31 to -24	TATA(A/T)A(A/T)(A/G)	TBP
BRE <sup>d</sup>	-23 to -17	(G/A)T(T/G/A)(T/G)(G/T)(T/G)(T/G)	TFIIB
Inr	-2 to +5	PyPyAN(T/A)PyPy	TAF1/TAF2
MTE	+18 to +29	C(G/C)A(A/G)C(G/C)(G/C)AACG(G/C)	n.a.
DPE	+28 to +34	(A/G)G(A/T)CGTG	TAF6/TAF9
DCE	3 subelements +6 to +11 +16 to +21 +30 to +34	core sequence: S <sub>I</sub> CTTC S <sub>II</sub> CTGT S <sub>III</sub> AGC	TAF1



Cependant, des études informatiques sur le génome des Métazoaires ont suggéré que la prévalence de la boîte TATA avait été surestimée dans le passé et que seuls 10 à 20% des promoteurs de gènes humains examinés en étaient pourvus (Carninci *et al.*, 2005; Cooper *et al.*, 2006; Gershenson and Ioshikhes, 2005). D'ailleurs, dès 2001, d'autres éléments de séquences très distincts de la boîte TATA ont été identifiés au sein de promoteurs de Métazoaires suggérant ainsi l'existence d'une régulation plus fine de l'initiation de la transcription (Butler and Kadonaga, 2002; Smale, 2001).

## 2. L'INR, élément initiateur

Des études de mutagenèse ont montré que l'exactitude du site d'initiation de la transcription dépend de l'adénosine en position +1 mais également des nucléotides qui l'entourent. Une cytosine précède généralement le +1, et les nucléotides situés entre -2 et +5 constituent une séquence de reconnaissance appelée **élément initiateur** (INR) qui a un rôle similaire à celui de la boîte TATA (Smale and Baltimore, 1989). L'INR est en effet capable d'initier seul la transcription via la RNAP II *in vitro* et *in vivo*.

Découvert en 1989 au sein du promoteur du **gène murin TdT** (Terminal deoxynucleotidyltransferase) dénué de boîte TATA (Smale and Baltimore, 1989), l'INR entoure le TSS (Fig.6). Sa séquence consensus chez l'Homme est Py-Py(C)-A<sup>(+1)</sup>-N-T/A-Py-Py (Corden *et al.*, 1980).

Cet INR est suffisant pour initier la transcription et peut fonctionner de manière synergique avec la boîte TATA si la distance qui sépare ces deux éléments est comprise entre 25 et 30 pb (O'Shea-Greenfield and Smale, 1992). Cependant, de même que pour la boîte TATA, l'INR ne se trouve pas dans tous les promoteurs de gènes codant pour des protéines. Une analyse informatique des promoteurs de 1031 gènes humains a révélé la présence de la séquence consensus de l'INR dans 85% des promoteurs étudiés (Suzuki *et al.*, 2001). Cet élément initiateur est reconnu par le complexe TFIID (Bellorini *et al.*, 1996) et des interactions stables de l'INR avec un **complexe regroupant les facteurs TAF1 et TAF2** ont été démontrées (Chalkley and Verrijzer, 1999), ce résultat étant en accord avec le fait que la fixation de l'hétéro-trimère **TBP-TAF1-TAF2** sur l'INR puisse engendrer une activité transcriptionnelle dans des tests de transcription *in vitro* (Verrijzer *et al.*, 1995).

### 3. Le DPE, site de régulation en aval du TSS

Situé en aval du site d'initiation de la transcription, le DPE (Downstream Promoter Element) a été identifié au départ chez la Drosophile comme un motif nécessaire à la liaison sur des promoteurs dépourvus de boîte TATA du facteur purifié TFIID (Burke and Kadonaga, 1996). Il est conservé de la Drosophile à l'Homme, sa séquence consensus (A/G)G(A/T)CGTC étant localisée précisément à 28 nucléotides du site +1 de l'initiation de la transcription (Kutach and Kadonaga, 2000) (Fig.7).

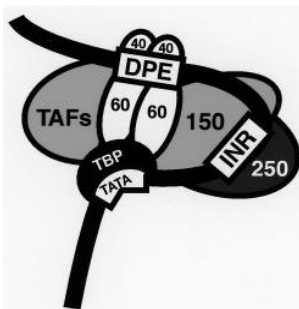


Figure 7: Modèle illustrant la liaison d'une partie de TFIID au promoteur minimal (d'après (Chalkley and Verrijzer, 1999). D'après la nouvelle nomenclature de 2002 : dTAF<sub>II</sub>250=TAF1 ; dTAF<sub>II</sub>150=TAF2 ; dTAF<sub>II</sub>60=TAF6 et dTAF<sub>II</sub>40=TAF9.

Cependant, le DPE existe aussi au sein de promoteurs contenant la boîte TATA. Dans une analyse de 205 promoteurs de Drosophile, il a été montré que 29% des promoteurs contenaient uniquement la boîte TATA, 26% uniquement le DPE, 14% les deux motifs et 31% aucun de ces deux motifs (Kutach and Kadonaga, 2000). Chez l'Homme, aucune étude de ce type n'a pour l'instant été réalisée. Le DPE est donc aussi commun que la boîte TATA chez la Drosophile. Cependant, il semble qu'une activité stimulant la transcription dépendante du DPE ait un **effet répressif** sur la boîte TATA (Willy *et al.*, 2000). En tous cas, la transcription basale dépendante du DPE implique une liaison extrêmement coordonnée du TFIID au DPE et à l'INR, le DPE interagissant avec les sous-unités **TAF6** et **TAF9** du TFIID (Fig.6 et 7).

### 4. Le BRE, élément de reconnaissance à TFIIB

En amont de la boîte TATA se trouve la séquence **BRE<sup>u</sup>** (TFIIB Recognition Element): (G/C)-(G/C)-(G/A)-C-G-C-C qui est un site de reconnaissance pour le facteur **TFIIB** (Lagrange *et al.*, 1998) (Fig.6). Une séquence BRE<sup>d</sup> en aval de la boîte TATA a également été identifiée ces dernières années (Deng and Roberts, 2005). L'analyse de 315 promoteurs contenant la boîte TATA a révélé que 12% des promoteurs contenaient au moins 5 des 7 nucléotides de sa séquence consensus. Alors que dans les Archaeobactéries, l'interaction de TFB, l'homologue archaeobactérien de TFIIB, avec la séquence BRE favorise

l'assemblage du complexe de pré-initiation de la transcription, chez l'Homme, il semblerait que cette interaction soit plus encline à **réprimer la transcription basale**, comme le montrent les essais de transcription *in vitro* réalisés sur des extraits nucléaires bruts humains (Evans *et al.*, 2001). Ces résultats suggèrent donc que la fonction de l'élément BRE pourrait s'être élargie au cours de l'évolution, présentant une activité transcriptionnelle **tantôt inhibitrice, tantôt activatrice** (Lagrange *et al.*, 1998). Initialement défini dans des promoteurs contenant la boîte TATA, il semble à présent que les séquences BRE<sup>u</sup> et BRE<sup>d</sup> puissent être présentes dans des promoteurs dépourvus de boîte TATA (Deng and Roberts, 2005). Cela suggère que la présence de TFIIB, dans les promoteurs à boîte TATA, renforcerait la liaison de TFIID à l'ADN en lui fournissant des points d'ancrage supplémentaires.

### 5. Le MTE, motif à 10 éléments

Le MTE (Motif Ten Element) a été découvert récemment suite à sa surreprésentation dans une analyse informatique portant sur 2000 promoteurs de *Drosophila* (Lim *et al.*, 2004; Ohler *et al.*, 2002). Chez la *Drosophila*, sa séquence consensus C(G/C)A(A/G)C(G/C)(G/C)AACG(G/C) a été découverte à la position +18, en aval du site d'initiation de la transcription (Fig.6). Le MTE a été conservé au cours de l'évolution, de la *Drosophila* à l'Homme. Il a certes besoin de l'INR pour activer la transcription par la RNAP II, mais peut, par contre, fonctionner de **manière indépendante** de la boîte TATA et du DPE, servant ainsi de substitut fonctionnel à ces motifs. En leur présence, le MTE peut également agir avec eux de manière synergique (Lim *et al.*, 2004). Pour l'instant, il reste encore à déterminer quel facteur de transcription pourrait se fixer spécifiquement sur le MTE.

### 6. Le DCE, élément de promoteur en aval du TSS

Originellement identifié dans le promoteur de la  $\beta$ -globine humaine (Lewis *et al.*, 2000), l'existence de cet élément a été révélée par la présence de 2 mutations à +22 et à +33 chez des individus atteints de beta-thalassémie. Le DCE (Downstream Core Element) est un élément tripartite qui s'étend approximativement de +6 à +34 en aval du site d'initiation de la transcription. La présence du DCE et du DPE semble être mutuellement exclusive (Fig.6), comme l'ont montré des analyses de séquences à partir de bases de données de promoteurs humains (Lee *et al.*, 2005). Le DCE fonctionne **de concert** avec la boîte TATA et l'INR, tout comme **en l'absence** de boîte TATA. La fixation de TAF1, composant de TFIID, sur le DCE,

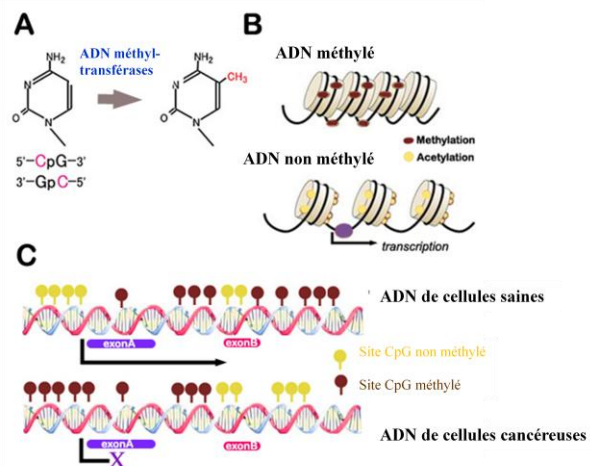
suggère que le TFIID joue un rôle non négligeable dans le fonctionnement de ce DCE (Lewis *et al.*, 2000).

### C. Les îlots CpG

Le génome des Mammifères est ponctué de séquences d'ADN contenant une fréquence anormalement élevée de sites riches en bases C et G, appelés **îlots CpG** (Bird, 1986; Gardiner-Garden and Frommer, 1987). Cette séquence remarquable s'étend sur 1kb en moyenne et chevauche les régions promotrices de 60 à 70% des gènes humains (Hellmann *et al.*, 2008) et de 93% d'un panel de gènes exprimés lors de **l'embryogenèse murine** (Ponger *et al.*, 2001).

Ces îlots CpG, dont le nombre est estimé à 27000 chez les Mammifères (Waterston *et al.*, 2002), sont généralement **hypométhylés**. Ces îlots CpG seraient donc vraisemblablement des **empreintes** laissées par la machinerie basale transcriptionnelle lors de l'embryogenèse, car ils empêcheraient l'accessibilité de la **DNMT** (DNA methyltransferase) à l'ADN, empêchant par là-même la méthylation *de novo* de l'ADN (Illingworth and Bird, 2009). De nombreux cas de méthylation de ces îlots CpG dans les cellules malignes ont été rapportés (Esteller, 2007), entraînant ainsi la **répression transcriptionnelle** de certains gènes de manière inappropriée (Fig.8).

**Figure 8 : A) Méthylation de l'ADN au niveau des îlots CpG par des ADN méthyltransférases. B) la déméthylation de l'ADN relâche la structure de la chromatine permettant l'acétylation des histones et la liaison de complexes transcriptionnels. C) Les cellules tumorales sont caractérisées par une hyperméthylation des îlots CpG alors que l'ADN sain est hypométhylé à cet endroit (adapté de Shirley M. Taylor, 2006).**



La méthylation de ces îlots intervient aussi dans des tissus normaux et participe à l'**inactivation** de nombreux gènes sur le chromosome X ainsi qu'à l'empreinte génomique

(Reik and Walter, 2001). Plus récemment, une méthylation différentielle entre les tissus a été détectée, les îlots CpG des cellules somatiques testées étant méthylés alors que ceux des cellules germinales ne le sont pas. Ceci suggère un rôle potentiel des îlots CpG dans la **régulation de la transcription** lors de la spécialisation de certaines cellules (Song *et al.*, 2005).

Le mode de fonctionnement de ces séquences manque toutefois de clarté. Dans la majorité des cas, les îlots CpG ne comportent ni boîte TATA, ni Inr, ni DRE ni aucun élément décrit comme impliqué dans la transcription. Il a néanmoins été observé de multiples sites de liaison à la **protéine Sp1**, entre 40 et 80 nucléotides en aval du site d'initiation de la transcription, associés aux îlots CpG (Blake *et al.*, 1990), ce qui suggère que Sp1 serait directement impliqué dans la formation du complexe de pré-initiation de tels promoteurs.

Sur ces éléments du promoteur minimal vont venir se fixer spécifiquement des protéines impliquées dans l'initiation de la transcription.

## II. Les acteurs impliqués dans l'initiation de la transcription

Des expériences de transcription *in vitro* réalisées à l'aide de fractions chromatographiques d'extraits nucléaires de cellules Hela ont révélé la formation d'un **complexe de pré-initiation** (PIC) composé des facteurs généraux de transcription TFIID, TFIIA, TFIIB, TFIIF, TFIIE, TFIIH facilitant la fixation de la Pol II (Fig. 9). Dans cette partie, nous présenterons brièvement chacun des acteurs nécessaires à l'initiation de la transcription des gènes de classe II puis nous aborderons la dynamique de leur liaison aux séquences de reconnaissance. Nous évoquerons ensuite les 2 modèles connus ainsi qu'un troisième plus hypothétique, mettant en jeu un nouveau complexe intermédiaire d'initiation de la transcription.

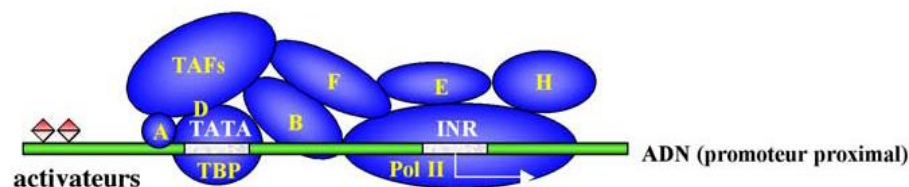


Figure 9: Vue d'ensemble du complexe de préinitiation de la transcription.

## A. Un acteur essentiel aux gènes eucaryotes : l'ARN polymérase II

### 1. Les ARN polymérases chez les eucaryotes

Chez les eucaryotes supérieurs, il existe 5 **ARN polymérases**, qui synthétisent de l'ARN à partir d'une matrice d'ADN. Les 3 premières ont été caractérisées très tôt grâce à des fractionnements chromatographiques et présentent des différences de sensibilité à l' $\alpha$ -amanitine. La 4<sup>ème</sup> (Pol IVa) et la 5<sup>ème</sup> (Pol IVb) découvertes plus tardivement dans les plantes, facilitent la production de siRNA impliqués dans la réduction au silence de la transcription et dans la formation de l'hétérochromatine (Herr *et al.*, 2005). Les ARN Pol I, II et III sont chacune spécifique d'une classe de gènes (Roeder and Rutter, 1969).

-L'ARN polymérase I synthétise les **ARN ribosomaux** qui constituent les ARN les plus représentés dans la cellule. Ces ARNs 18S et 28S sont produits à partir d'un transcrit unique.

-L'ARN polymérase de type III synthétise l'**ARN ribosomal 5S** et les **ARN de transfert**.

-L'ARN polymérase de type II est responsable de la synthèse des **ARNm codant pour des protéines** et c'est cette dernière qui va focaliser notre attention dans le cadre de l'initiation de la transcription des gènes de classe II.

### 2. L'ARN polymérase de type II

La Pol II représente la sous-unité **catalytique** clé du PIC (Fig.9). Son architecture et sa fonction sont hautement conservées de la Levure à l'Homme (Fig.10).

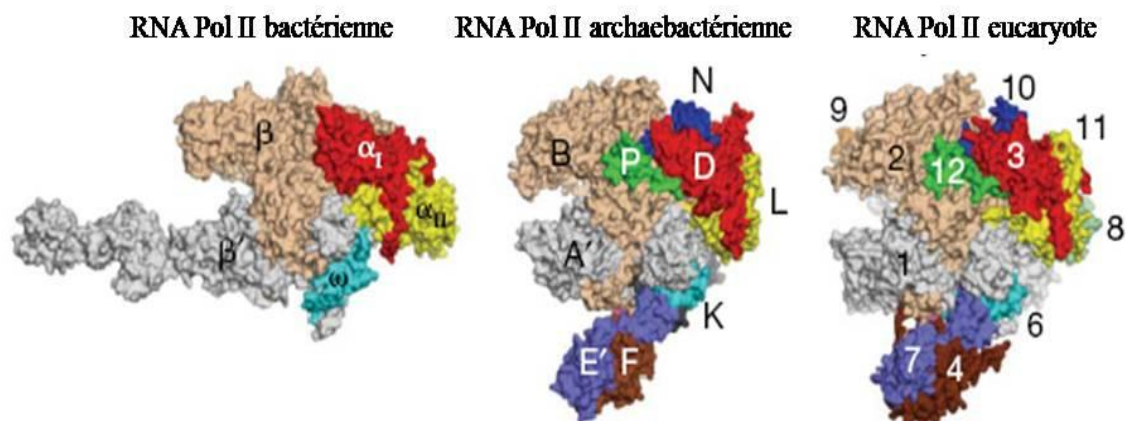


Figure 10: Structure aux rayons X de la RNA Pol II au sein de 3 espèces différentes (Hirata *et al.*, 2008).

Il s'agit d'un **complexe multiprotéique** de 0,5 MDa composé de 12 sous-unités **RPB** (RNA Polymerase B) nommées de 1 à 12 par ordre de décroissance de leurs masses moléculaires (Young, 1991). Parmi ces 12 sous-unités, 5 éléments (RPB5, RPB6, RPB8, RPB10, et RPB12) sont partagés par les 3 premières polymérases (Woychik *et al.*, 1990). Seules RPB 4, 7 et 9 sont uniques à la Pol II ainsi que l'extrémité carboxyterminale (domaine CTD) de RPB1.

a) Les sous-unités de l'ARN polymérase II

Grâce à la **cristallographie aux rayons X**, les structures de la Pol II, seule à 2,8 Å (Cramer *et al.*, 2000) ou en cours d'élongation à 3 Å (Armache *et al.*, 2003; Gnatt *et al.*, 2001) ont pu être résolues. L'arrangement des différentes sous-unités au sein de **Pol II** est bipartite : d'un côté, un **noyau** de 10 sous-unités (Fig.11) et de l'autre, un **hétérodimère** constitué des sous-unités RPB4 et RPB7 (Armache *et al.*, 2003; Bushnell and Kornberg, 2003). 8 ions  $Zn^{2+}$ , situés à la surface de Pol II, stabilisent l'enzyme et 3 ions  $Mg^{2+}$  se situent au sein du **site actif** de l'enzyme.

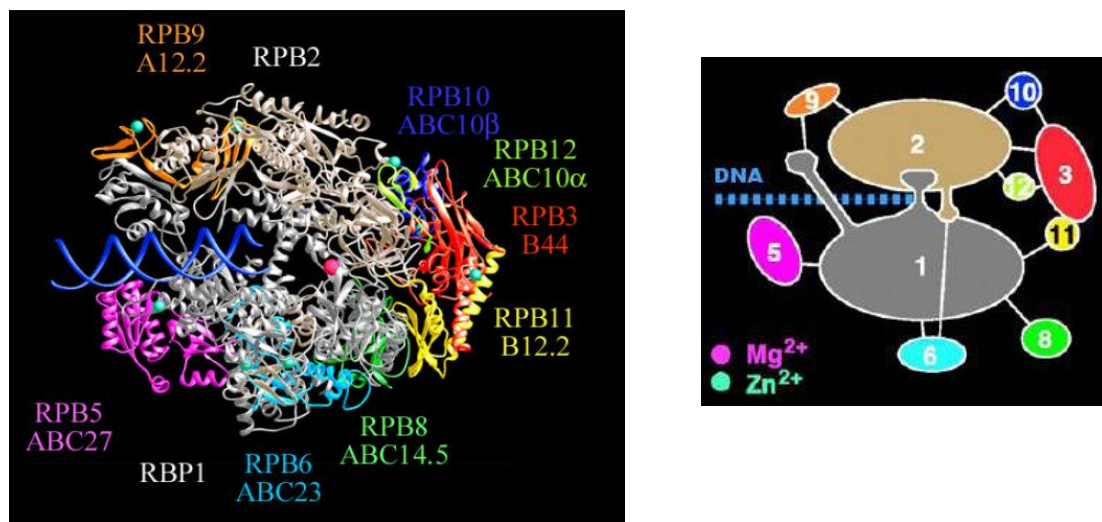


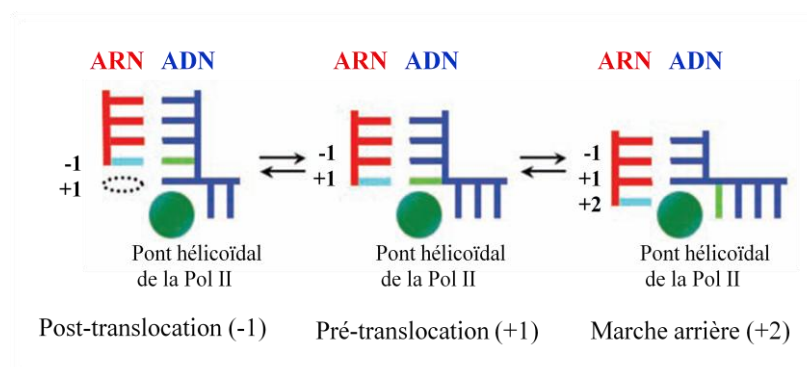
Figure 11: a) Structure de l'ARN polymérase II de la Levure à 2,8 Å de résolution.

b) Représentation schématique de l'ossature de l'ARN polymérase II

La structure globale de la Pol II ressemble à une sorte de **mâchoire** formée par les sous-unités RPB 2 et 9 en haut et par RPB1 en bas (Fig.11b), avec à l'intérieur une fissure centrale dans laquelle pénétrerait l'ADN. Le **site actif** de la Pol II est localisé au fond de la cavité, fermée à une extrémité par le «mur», un domaine de la sous-unité RPB2. Un module flexible appelé «**pince**» (extrémité N-terminale de RPB1 et C-terminale de RPB2) pivote autour du site actif. Une hélice  $\alpha$  de Rpb1 fait saillie dans le sillon en aval du site catalytique pour former un **pont hélicoïdal** entre les mâchoires et la pince. Ces dernières sont responsables de la fixation de la Pol.II à l'ADN.

L'hétérodimère RPB4/RPB7, dissociable du reste de la Pol II de Levure en conditions dénaturantes (Edwards *et al.*, 1991), se place à la base de cette pince dans une poche formée par RPB1, 2 et 6. Un de ses rôles consisterait à réduire fortement la mobilité de cette pince (Bushnell and Kornberg, 2003). Bien que la protéine RPB7 interagisse avec le noyau de Pol II (Armache *et al.*, 2003), la quasi-totalité de l'hétérodimère RPB4/RBP7 est exposée et **accessible pour des interactions** avec des protéines ou des acides nucléiques.

Ainsi l'ARN naissant pourrait se fixer sur une surface d'accrochage potentielle des acides nucléiques appartenant à RBP7 qui fait face à la «selle à cheval» de la Pol II à partir de laquelle l'ARN émerge (Todone *et al.*, 2001). L'ARN polymérase II catalyse la synthèse de l'ARN à des vitesses très élevées, allant de **20 à 70 nucléotides par seconde** le long d'une matrice d'ADN. Son activité n'est cependant pas linéaire, elle oscille entre des mouvements en avant et en arrière à chaque étape du processus de synthèse (Toulme *et al.*, 1999).



**Figure 12:** Les 3 états du complexe d'élongation de la transcription de la Pol II. L'ARN transcrit est rouge, l'ADN matrice est bleu. La base nucléotidique ajoutée à l'extrémité 3' de l'ARN est cyan et sa base d'ADN complémentaire est vert clair. L'ovale en pointillés délimite le site vide d'insertion des nouveaux nucléotides à l'état de post-translocation. Le rond vert représente le pont hélicoïdal de la Pol II.



3 étapes ont pu être caractérisées dans la formation du transcrit primaire (Fig.12): la première lorsqu'un **nucléotide était ajouté** dans le site d'addition en vue d'intégrer la chaîne d'ARN, la seconde lors du **déplacement de l'enzyme** vers l'avant afin d'insérer le nucléotide, et la dernière lors du **mouvement de recul** de la Pol II pour terminer la maturation de l'extrémité 3' de l'ARN (Wang *et al.*, 2009).

### 6) Le domaine CTD et ses phosphorylations

Au sein de la sous-unité RBP1 de la Pol II se trouve un domaine CTD (domaine C-terminal) contenant **25 à 52 répétitions en tandem** de la séquence heptamérique YSPTSPS avec des sites de phosphorylation sur la sérine 2 et sur la sérine 5. Ce domaine CTD est en effet la cible de **phosphorylations** et de déphosphorylations participant à la régulation de la transcription (Bentley, 2002).

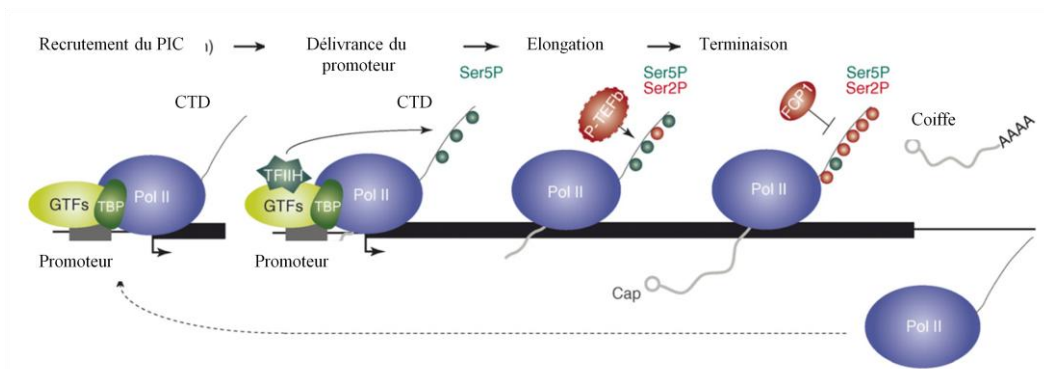


Figure 13: Importance du CTD dans la régulation de la transcription (Koch *et al.*, 2008).

La Pol II contenant un **domaine CTD** hypophosphorylé est tout d'abord recrutée au niveau des promoteurs lors de la formation du PIC puis elle est **phosphorylée** au niveau de la **sérine 5** lors de l'initiation de la transcription (Lee and Young, 2000). Deux kinases dépendant des cyclines Cdk7 et Cdk8 et appartenant au PIC, ciblent le CTD pour la phosphorylation (Liu *et al.*, 2004). Cette phosphorylation par des **kinases** déstabilise le PIC et permet le départ de la Pol II du promoteur. Après l'initiation, d'autres kinases telles que Cdk9/Ctk1 phosphoryleront **la sérine 2** (Fig.13), en vue de la maturation des ARN et de leur **polyadénylation** via le recrutement de facteurs sur la Pol II (Ahn *et al.*, 2004). Le CTD agit donc comme une plate-forme d'assemblage des différents facteurs régulant l'initiation,

l'élongation et la terminaison de la transcription ainsi que la maturation des ARNm (Hahn, 2004).

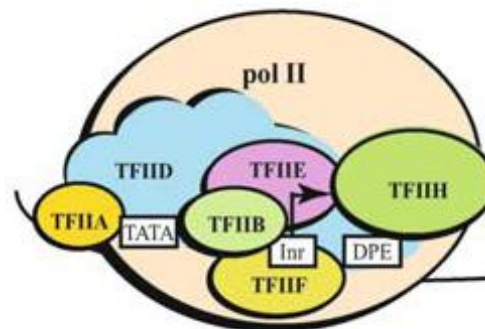
### B. Les facteurs généraux de la transcription

Chez les eucaryotes, la Pol II est incapable d'initier la transcription par elle-même, car sa fixation sur les promoteurs est très **instable** (Roeder *et al.*, 1976; Weil *et al.*, 1979).

Un ensemble de facteurs regroupés sous l'appellation de **facteurs généraux de la transcription** (GTF) de classe II assistent l'ARN polymérase II pendant les différentes étapes d'initiation et de démarrage de la transcription ainsi que l'élongation du transcrit.

Chacun de ces facteurs a été nommé par une lettre en fonction de la fraction chromatographique dans laquelle résidait son activité, de TFIIA à TFIIIF. (Sawadogo and Roeder, 1985a). La mise en place du complexe d'initiation de la transcription est une étape essentielle de l'activation des gènes. Une brève description de ces GTF figure dans le tableau 1. L'ensemble des facteurs est représenté schématiquement dans la figure ci-dessous (Fig.14).

Figure 14 : GTFs et Pol II assemblés sur le promoteur du gène à transcrire;



### C. Les modèles d'assemblage des différents facteurs

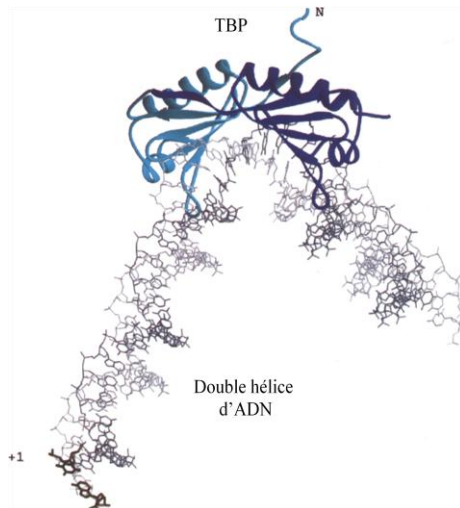
Des études biochimiques approfondies ont révélé la manière dont s'assemble le complexe de pré-amorçage de la transcription contenant la Pol II. Dans ces études, les techniques **d'empreinte à la DNase I** et de mesure des **changements de mobilité électrophorétiques** ont été utilisées, afin de déterminer l'ordre dans lequel la Pol II et les GTFs se fixent aux promoteurs des gènes à transcrire.

Facteur de transcription	Composition protéique	Fonctions assurées par ces facteurs
TFIIA	p35( $\alpha$ ), p19( $\beta$ ), p12( $\gamma$ )	<ul style="list-style-type: none"> <li>• Antirépresseur, Coactivateur</li> <li>• stabilise le complexe TBP-TATA</li> </ul>
TFIIB	p33	<ul style="list-style-type: none"> <li>• Sélection du site d'initiation</li> <li>• stabilise le complexe TBP-TATA, lie le BRE</li> <li>• permet le recrutement de Pol II/TFIIF</li> </ul>
TFIID	TBP+TAFs (TAF1 à TAF14)	<ul style="list-style-type: none"> <li>• Facteur de liaison au promoteur minimal</li> <li>• coactivateur</li> <li>• protéine kinase</li> <li>• activité d'ubiquitination et de conjugaison</li> <li>• histone acétyltransférase</li> </ul>
TFIIE	p56, p34	<ul style="list-style-type: none"> <li>• recrute TFIIH</li> <li>• Facilite la formation d'un complexe efficace d'initiation contenant Pol II</li> <li>• Impliqué dans le dégagement des promoteurs</li> </ul>
TFIIF	RAP30, RAP74	<ul style="list-style-type: none"> <li>• Se lie à Pol II et facilite son recrutement au niveau du promoteur</li> <li>• Recrute TFIIE et TFIIH</li> <li>• Fonctionne avec TFIIB et Pol II dans la sélection du site d'initiation</li> <li>• Facilite l'échappement de la Pol II</li> <li>• Augmente l'efficacité de l'élongation de la Pol II</li> </ul>
TFIIH	P89/XPB, p80/XPD, p62, p52, p44, p40/CDK7, p38/CyclinH, p34, p32/MAT1, p8/TFB5	<ul style="list-style-type: none"> <li>• Activité ATPase pour l'initiation de la transcription et pour la délivrance des promoteurs</li> <li>• Activité hélicase pour l'accessibilité au promoteur</li> <li>• Activité de réparation par excision de nucléotides couplée à la transcription</li> <li>• Activité de E3 ubiquitine ligase</li> </ul>
Pol II	RPB1-RPB12	<ul style="list-style-type: none"> <li>• Initiation, Elongation, Terminaison de la transcription</li> <li>• Recrutement d'enzymes de coiffe des ARNm</li> <li>• Recrutement couplé à la transcription de facteurs intervenant dans l'épissage et dans le traitement des extrémités 3'.</li> <li>• Phosphorylation du CTD, ubiquitination et glycosylation</li> </ul>

Tableau 1: Aperçu global des facteurs de l'initiation de la transcription chez les Eucaryotes supérieurs.

## 1. Le modèle séquentiel

Comme le facteur TFIID est relativement difficile à purifier avec ses nombreuses sous-unités, les études portant sur la chronologie de fixation des facteurs de transcription lors de l'initiation de la transcription ont été réalisées à l'origine avec TBP (TATA Binding Protein), la sous-unité de TFIID fixant la boîte TATA (Fig.15). *In vitro*, Pol II est capable de débiter la transcription en l'absence des autres sous-unités de TFIID.



**Figure 15: Structure du complexe TBP lié à l'élément TATA. Les domaines amino- et carboxy-terminaux de la molécule sont respectivement représentés en bleu clair et en bleu foncé (Orphanides *et al.*, 1996).**

TBP est la première protéine découverte ayant la capacité de se fixer à la boîte TATA via son domaine C-terminal. Celui-ci est très conservé chez les Eucaryotes et présente près de 80% de similarité entre la Levure et l'Homme. Au contraire, le domaine N-terminal est plus variable d'un Eucaryote à l'autre, en longueur comme en séquence. L'extrémité C-terminale adopte une structure en forme de selle de cheval qui fixe la séquence TATA et provoque la courbure de la double hélice d'ADN. Les deux parties du domaine C-terminal de TBP donnent à la molécule une symétrie globale en dyade, mais ne sont pourtant pas identiques (Nikolov *et al.*, 1992). TBP peut intervenir efficacement sur l'initiation de la transcription *in vitro* s'il est combiné aux GTF et à la Pol II (Hoey *et al.*, 1990; Horikoshi *et al.*, 1990).

### a) Reconnaissance du promoteur minimal par TFIID

L'étape initiale de la formation du PIC est la reconnaissance du promoteur.

*-les promoteurs contenant une boîte TATA*

Si le promoteur est pourvu d'une boîte TATA, il sera reconnu par le complexe TFIID et plus particulièrement par la protéine TBP. Celle-ci interagit avec le petit sillon de l'ADN et courbe ainsi considérablement la double hélice d'ADN (Fig.15) (Burley, 1996; Pugh, 2000). Cela permet, d'une part, le rapprochement des séquences situées en amont de la boîte TATA avec celles situées en aval et d'autre part, l'ouverture des deux brins d'ADN.

*-les promoteurs dépourvus de boîte TATA*

Dans les promoteurs possédant une boîte TATA fonctionnelle, la mutation de la boîte TATA entraîne une diminution importante de la transcription (Wobbe and Struhl, 1990). Ainsi, la mutation de la boîte TATA en une séquence riche en CG permet certes le recrutement au promoteur de toute la machinerie transcriptionnelle mais abolit entièrement la transcription (Ranish *et al.*, 1999). Pourtant seulement 10% des promoteurs humains possèdent cette boîte TATA (Bajic *et al.*, 2006). Il est donc probable que TBP puisse lier d'autres séquences avec une affinité plus faible, peut-être avec l'aide des facteurs de transcription.

Des études récentes de structure cristallographique à haute résolution de l'interface TBP-ADN ont montré que même de très faibles énergies d'interaction entre ces deux partenaires pouvaient conduire à des déformations très importantes de la double hélice. Cela suggère la mise en place de stratégies visant à améliorer l'affinité et la sélectivité de TBP pour l'ADN. Parmi ces stratégies, la suppression de chocs stériques induit une modification de l'interface TBP-ADN (Millan-Pacheco *et al.*, 2009). Les TAFs, ces protéines associées à TBP peuvent également faciliter la liaison de TFIID au promoteur grâce à des contacts directs avec l'ADN au niveau des éléments INR et DPE (Chalkley and Verrijzer, 1999).

**6) Reconnaissance du complexe TFIID/promoteur par TFIIA**

Le complexe TFIIA vient stabiliser la liaison ADN/TFIID en interagissant directement avec l'ADN en amont de la boîte TATA (Buratowski *et al.*, 1989; Cortes *et al.*, 1992).

Le rôle précis de TFIIA dans l'initiation de la transcription a été longtemps sujet à controverses. Généralement, les réactions de transcription *in vitro* réalisées avec du TBP recombinant ne nécessitent pas l'intervention de TFIIA (Cortes *et al.*, 1992), alors qu'au contraire, celles réalisées avec le facteur TFIID purifié (Fig.16) sont stimulées 2 à 10 fois par

TFIIA (Ozer *et al.*, 1994). Ainsi, plus le système de transcription est pur, moins TFIIA semble nécessaire, suggérant que ce facteur neutralise des répresseurs présents dans la réaction (Auble *et al.*, 1994; Inostroza *et al.*, 1992; Merino *et al.*, 1993).

TFIIA semble également être la cible de nombreux coactivateurs (Jacobson and Tjian, 1996; Ozer *et al.*, 1998), notamment de p300 qui interviendrait en acétylant le TFIIA (Mitsiou and Stunnenberg, 2003).

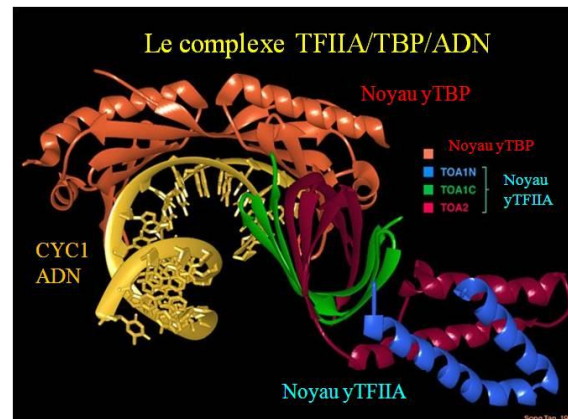


Figure 16: Le complexe TFIIA/TBP/ADN (Tan *et al.*, 1996).

### c) Reconnaissance de TFIID/TFIIA/promoteur par TFIIIB

TFIIIB est le GTF suivant à entrer dans la voie d'assemblage du PIC (Fig.17).

Il reconnaît le complexe TFIID/TFIIA/promoteur, interagissant aussi bien avec TBP (Nikolov *et al.*, 1995) qu'avec les séquences d'ADN situées de part et d'autre de la boîte TATA (Reese, 2003). La majeure partie de TFIIIB se trouve en aval de la boîte TATA créant alors une asymétrie qui pourrait être responsable de l'ancrage, de la stabilisation du PIC et de l'orientation de la transcription en direction du site d'initiation (Tsai and Sigler, 2000). Des études sur la levure appuient ce modèle, puisque des mutations dans le gène TFIIIB entraînent un déplacement du site d'initiation de la transcription (Wu *et al.*, 1999b).

TFIIIB subit des modifications conformationnelles importantes lors de ses interactions avec l'ADN ou avec des coactivateurs où son domaine « B-finger » joue un rôle primordial (Elsby and Roberts, 2004; Fairley *et al.*, 2002). Il favoriserait en effet le recrutement d'activateurs (Hawkes *et al.*, 2000) tout en interagissant avec la sous-unité RBP2 de la Pol II

(Chen and Hampsey, 2004). Ainsi TFIIB participerait au positionnement de l'ADN des gènes à transcrire dans la fissure de la Pol II (Chen and Hahn, 2004).

A côté de ces rôles de stabilisation et de spécification du site d'initiation de la transcription, TFIIB joue également un rôle important dans le recrutement du complexe TFIIF/Pol II (Fig.18).

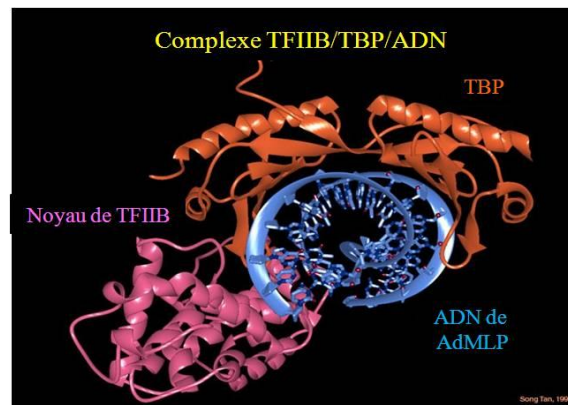


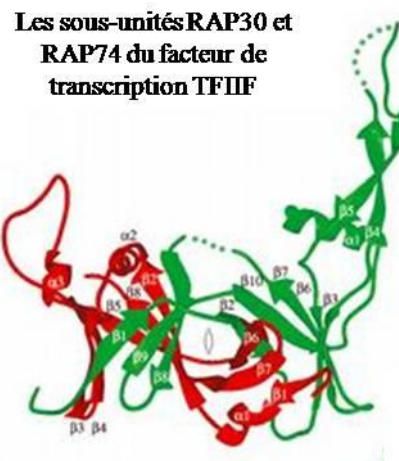
Figure 17: Structure du complexe TFIIB-TFIIA-TFIID-ADN

#### d) Recrutement du complexe TFIIF/ARN pol II

Lorsque TFIID et TFIIB sont assemblés sur le promoteur, l'ARN polymérase II peut alors s'associer au complexe en formation (Fig.20). Ceci se fait sous l'escorte de TFIIF (Young, 1991; Zawel *et al.*, 1995).

TFIIF est un facteur hétérotétramérique (Fig.18), constitué de 2 sous-unités de RAP74 et de 2 sous-unités de RAP30 (Flores *et al.*, 1990). RAP30 partage deux régions d'homologie avec le facteur bactérien  $\sigma$ , l'une capable d'interagir étroitement avec la Pol II (Sopta *et al.*, 1985) et l'autre avec l'ADN (Tan *et al.*, 1994) entre la boîte TATA et le site d'initiation INR (Coulombe *et al.*, 1994). TFIIF apporte ainsi des surfaces de contacts supplémentaires qui augmentent l'affinité de la Pol II pour le complexe TFIIB-TBP-promoteur (Robert *et al.*, 1998). Ce faisant, TFIIF induit des modifications topologiques de l'ADN qui favorisent l'enroulement de l'ADN autour de la Pol II. TFIIF serait donc responsable du positionnement correct de la polymérase en réduisant les interactions non spécifiques de celle-ci avec l'ADN (Conaway and Conaway, 1993) et en favorisant la formation d'un complexe transcriptionnel stable, résistant à l'inhibition par des répresseurs de la transcription (Hou *et al.*, 2000).

Par ailleurs, TFIIF joue également un rôle dans la sélection du site d'initiation, de par l'association de sa sous-unité RAP74 avec TFIIB (Fairley *et al.*, 2002; Ha *et al.*, 1993) et sa proximité du centre actif de Pol II (Bushnell *et al.*, 2004). Il peut également se lier à de nombreux partenaires protéiques, dont TBP, TAF1 et TAF5 (Dubrovskaya *et al.*, 1996). Enfin, TFIIF est aussi impliqué dans l'étape de transition suivant l'initiation où il empêche l'arrêt prématuré de la polymérase (Yan and Gralla, 1997) en travaillant de manière synergique avec le facteur d'élongation TFIIIS.



**Figure 18: Structure de l'hétérodimère RAP30/RAP74 de TFIIF (Gaiser *et al.*, 2000).**

(en vert : RAP 74, en rouge : RAP30)

### e) Liaison de TFIIE et TFIIH au PIC alors complet

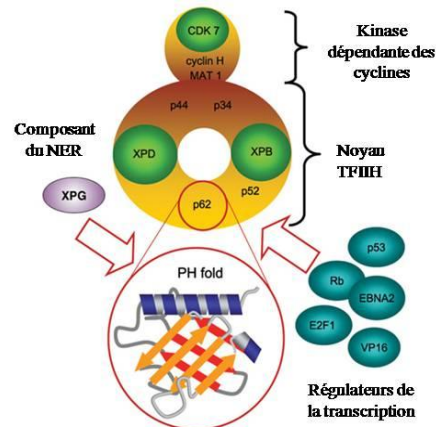
TFIIE rejoint le complexe en formation par des interactions avec la polymérase, avec TFIIF et avec le complexe TBP/TFIID (Orphanides *et al.*, 1996). TFIIE se lierait à une position très proche du site d'initiation via son domaine « forkhead » (Tanaka *et al.*, 2009), recrutant ensuite le complexe TFIIH (Flores *et al.*, 1990). La formation du PIC ainsi achevée (Fig.20) permettrait à TFIIH de commencer à ouvrir la double hélice d'ADN au niveau du promoteur en vue de la transcription.

Le complexe TFIIH occupe une place importante parmi les GTFs. En effet, plusieurs activités enzymatiques ont été associées à ce complexe. Celui-ci est composé de deux sous-complexes (Fig.19), le cœur et le complexe kinase-cycline-dépendant (CAK). Le cœur est composé d'une ADN hélicase dépendante de l'ATP, Xeroderma Pigmentosum B (XPB), de p34, p52, p62, p8 et p44 qui interagit avec XPD, et d'une hélicase ADN/ARN (Xeroderma Pigmentosum D). Le complexe CAK inclut une kinase dépendante de la cycline (Cdk7), accompagnée de la cycline H et de la protéine MAT1. Les activités hélicases jouent un rôle



majeur dans la formation du complexe ouvert lors de l'initiation de la transcription. La kinase Cdk7, quant à elle, est capable de phosphoryler le domaine CTD de l'ARN polymérase II, ce qui stimule l'élongation de la transcription. Outre son rôle dans la transcription en tant que facteur de base, le complexe TFIIH est aussi impliqué dans la régulation de la transcription (Zurita and Merino, 2003), dans la réparation de l'ADN par excision de nucléotides (NER), et dans la régulation du cycle cellulaire.

Figure 19: Structure du complexe TFIIH à l'interface des machineries de transcription et de réparation de l'ADN (Kwek *et al.*, 2004).



## 2. Modèle de l'holoenzyme

Une voie alternative de formation du PIC a été découverte lorsque plusieurs laboratoires ont réussi à purifier un complexe holoenzymatique contenant la Pol II, les GTF ainsi que d'autres protéines impliquées dans le remodelage chromatinien, ou la réparation de l'ADN (Chao *et al.*, 1996; Ossipow *et al.*, 1995; Parvin and Young, 1998). Dans ce modèle dit de « l'holoenzyme » (Fig.20), tous les facteurs seraient recrutés en une seule étape sur les promoteurs des gènes (Orphanides *et al.*, 1996). Cependant, parmi les différents complexes holoenzymatiques isolés, jamais le TFIID n'apparaît (Wu *et al.*, 1999a) ni le TFIIA. Cela suggère que le TFIID, en tant que facteur cœur de la liaison à l'ADN, pourrait faciliter l'entrée de l'holoenzyme Pol II dans la région promotrice.

## 3. Intermédiaire entre les deux modèles d'assemblage du PIC

Il est probable que les deux modèles (Fig.20), le modèle séquentiel et le modèle de l'holoenzyme coexistent *in vivo*, selon les facteurs impliqués et le contexte dans lequel se trouve le promoteur (Hampsey, 1998; Lee and Young, 2000; Lemon and Tjian, 2000). De plus, l'ordre de fixation des facteurs au niveau du promoteur semble également être variable selon le contexte. D'une manière logique, il a été proposé que les cofacteurs responsables du

remodelage de la structure de la chromatine se lie au promoteur avant les facteurs de base de la transcription, de manière à faciliter l'accès à l'ADN. Cependant, des expériences d'immunoprécipitation de la chromatine ont conduit à remettre en cause ce point de vue. Les composants de la machinerie transcriptionnelle de base peuvent promouvoir le remodelage de la chromatine et réciproquement (Featherstone, 2002; Zawel *et al.*, 1995). La limite entre facteurs de base de la transcription et cofacteurs semble se restreindre petit à petit. Le bon déroulement de l'initiation de la transcription est le résultat d'une dynamique d'assemblage très subtile d'une panoplie de facteurs qui dépend sans doute du contexte du promoteur.

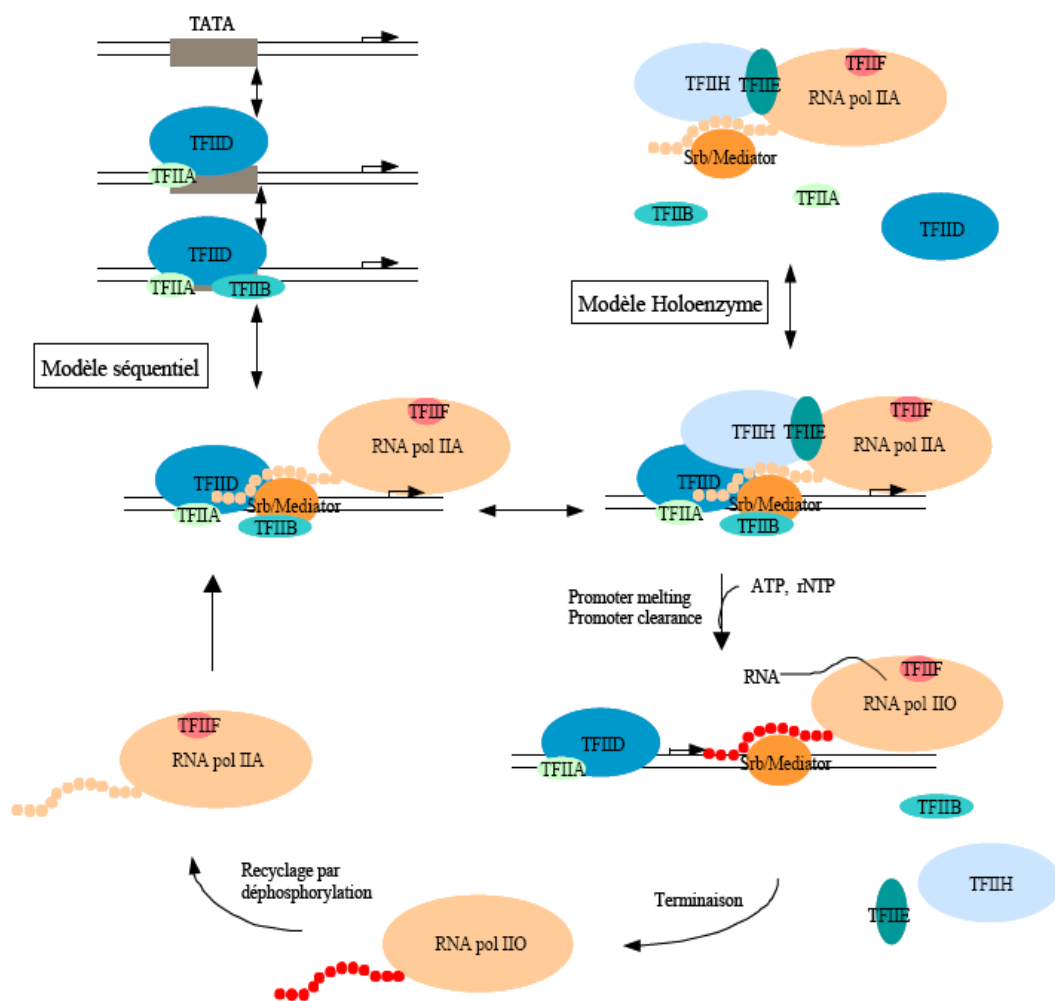
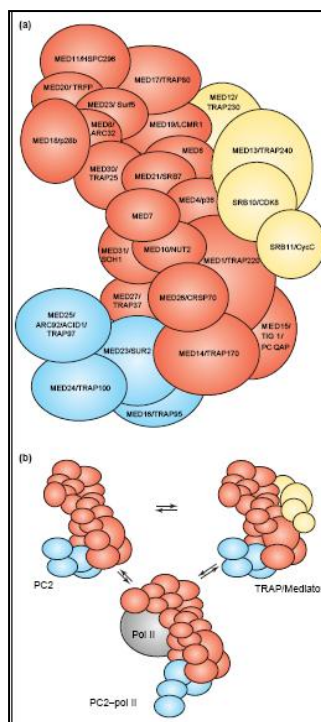


Figure 20: Les modèles d'assemblage du complexe d'initiation de la transcription.

#### D. Le complexe Médiateur

Dans les années 1990, des analyses biochimiques et génétiques chez la Levure *Saccharomyces cerevisiae* ont permis la découverte du complexe Médiateur (Thompson *et al.*, 1993), dont le nom évoque son rôle d'adaptateur entre les facteurs de transcription liés aux éléments de régulation, la Pol II et les GTF. Son existence dans la machinerie transcriptionnelle des Métazoaires a été démontrée plus tardivement (Malik and Roeder, 2000). Plusieurs types de médiateur ont été trouvés chez l'Homme et notamment PC2 (Malik *et al.*, 2000) et TRAP (Fondell *et al.*, 1996).



**Figure 21: Médiateur humain**

Ces complexes sont constitués de 28 à 30 sous-unités (Sato *et al.*, 2004), regroupées en 3 modules :

- le noyau central, en rouge sur la Fig.21 correspond aux 2 premiers modules du Médiateur (tête et corps médian) (Dotson *et al.*, 2000). Cette partie peut interagir avec le complexe TFIIF-Pol II (Lariviere *et al.*, 2006) et subit des changements conformationnels drastiques suite à cette liaison.
- la partie bleutée (Fig.21) correspond au troisième module aussi appelé « queue » qui peut interagir avec des activateurs proximaux. L'ensemble des trois premiers modules forme le complexe PC2.
- la partie jaune (Fig.21) s'ajoute aux trois premiers modules pour former le complexe TRAP.

A l'origine considéré comme un cofacteur de l'holoenzyme Pol II (Myers and Kornberg, 2000), il est reconnu à présent que son rôle prédominant consiste à promouvoir l'assemblage du PIC (Fig.20); soit en augmentant l'efficacité de l'assemblage (Cantin *et al.*, 2003), soit en agissant directement sur le recrutement de la Pol II ou des GTF (Wu *et al.*, 2003). D'ailleurs, des interactions ont été montrées avec le CTD de la Pol II (Naar *et al.*, 2002) ainsi qu'avec le facteur TFIIB (Baek *et al.*, 2006).

### III. Les séquences régulatrices

*In vitro*, les éléments du promoteur minimal jouent le rôle de plate-forme d'assemblage du PIC et permettent à l'initiation de la transcription de se dérouler correctement.

*In vivo*, l'ADN est compacté dans la chromatine et la présence des éléments du PIC sur le promoteur minimal ne suffit plus à assurer l'initiation de la transcription, car leur accès à l'ADN s'en trouve perturbé. Il y a donc répression de la transcription basale.

La présence de protéines activatrices se fixant sur des séquences spécifiques est indispensable pour déréguler la transcription. Aussi, le promoteur minimal semble plus complexe, car d'autres motifs peuvent être nécessaires à l'initiation et à l'activation de la transcription. Certains éléments sont situés à proximité du TSS et portent le nom de séquences régulatrices proximales, et d'autres sont plus éloignés, elles sont qualifiées alors de séquences régulatrices distales (Fig.22).

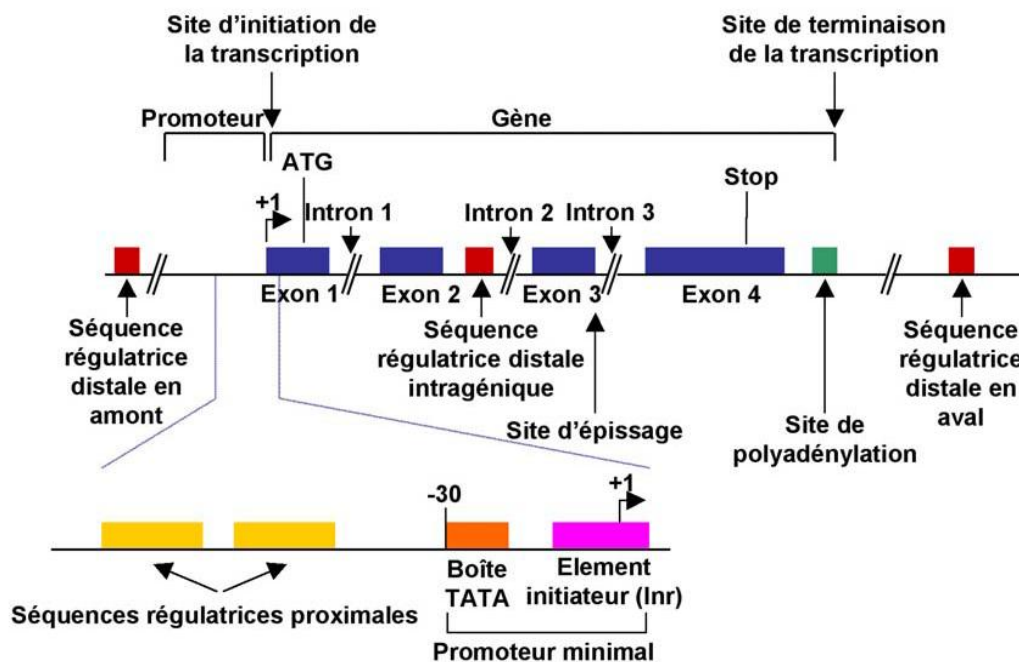


Figure 22: Représentation schématique des différentes séquences d'ADN impliquées dans la régulation de la transcription (promoteur, séquences régulatrices proximales et distales) ou dans la maturation des ARN (sites d'épissage et de polyadénylation des ARN). Les trinucléotides « ATG » et « stop » codent pour l'initiation et l'arrêt de la traduction de la protéine, respectivement.

## A. Les séquences régulatrices proximales

Habituellement situées entre 40 et 110 paires de bases en amont du site d'initiation, les **séquences régulatrices proximales** sont reconnues par des facteurs de transcription spécifiques. Ces facteurs peuvent agir indirectement sur la machinerie transcriptionnelle en recrutant d'autres facteurs responsables de la modification de la chromatine, ou directement en interagissant avec les éléments de la machinerie transcriptionnelle. Cette interaction protéine-ADN permet donc une régulation plus fine de l'initiation de la transcription.

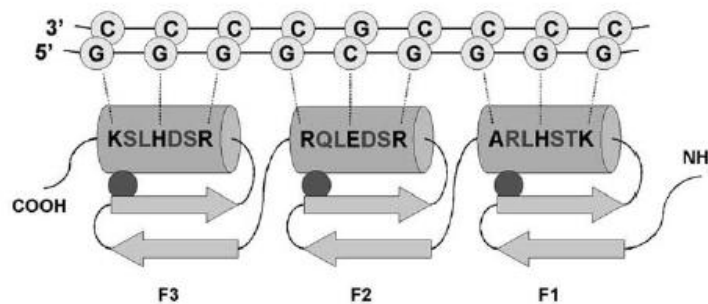
Les séquences régulatrices peuvent avoir un effet **activateur** (UAS soit Upstream Activating Sequence) ou **répresseur** (URS ou Upstream Repressing Sequence) selon la protéine qui les reconnaît. Plusieurs séquences sont maintenant bien caractérisées, comme les **boîtes CAAT** et les **motifs riches en GC**, les éléments de réponse aux cytokines comme la boîte ISRE (IFN Stimulatory Response Element), les éléments de réponse à des facteurs de croissance comme la boîte SRE (Serum Responsive Element) ou encore les éléments de réponse à des signaux hormonaux comme la **boîte RARE** (Retinoic Acid Responsive Element).

### 1. Le motif CCAAT

La boîte **CCAAT** est l'une des séquences régulatrices proximales les plus communes des promoteurs eucaryotes ; elle a été découverte très tôt, dès les années 1980 (Benoist *et al.*, 1980; Efstratiadis *et al.*, 1980). Située de 80 à 100 nucléotides, ou de 40 à 80 nucléotides en amont du site d'initiation de la transcription selon que le promoteur contient ou non une boîte TATA, elle peut être **orientée dans un sens ou dans l'autre**. Sous sa forme inverse ATTGG, sa fréquence s'élève à 60% dans les promoteurs dépourvus de boîte TATA (Mantovani, 1998). Parmi les différents facteurs protéiques qui interagissent avec cette séquence (CBF, HA2/3/4/5), seul NF-Y requiert pour sa liaison les 5 nucléotides de cette boîte. L'interaction de ce facteur avec la boîte CCAAT influencerait de manière sélective le sens de la transcription dans des promoteurs bidirectionnels (Zanotto *et al.*, 2009). Elle faciliterait également la mise en place d'une **conformation ouverte de la chromatine** via l'acétylation de l'histone H3, permettant ainsi la fixation d'autres facteurs de transcription sur l'ADN tels que FGFR2 (Sun *et al.*, 2009).

## 2. Le motif riche en GC

La **boîte GC** est une séquence régulatrice proximale qui s'étend de -50 à -100 pdb et qui contient une série de répétitions de nucléotides GC (Fig.23). Les boîtes GC sont significativement plus grandes en aval du TSS qu'en amont (58% contre 53%) et il y a une forte corrélation négative entre le contenu en GC et le nombre de promoteurs alternatifs putatifs (Ma *et al.*, 2009). Cela suggère donc que les **boîtes GC situées à proximité du TSS** jouent un rôle important dans la **régulation** des promoteurs alternatifs.



**Figure 23 : Liaison du facteur Sp1 de la famille Sp/XKLF grâce à ses structures en doigts de zinc (Bouwman and Philipsen, 2002).**

Ces boîtes GC sont des sites de liaison spécifiques pour la famille de protéines **Sp/XKLF** (specificity protein/Krüppel-like factor). Celle-ci est unie dans une configuration particulière : chacun de ses membres possède une combinaison de 3 doigts de zinc Cys2His2 (Fig.23) particulièrement conservés (Philipsen and Suske, 1999). Appelés Krüppel chez la drosophile (Schuh *et al.*, 1986), le premier facteur de ce type trouvé chez l'homme est Sp1 (Kadonaga *et al.*, 1987). La partie C-terminale de chaque doigt de zinc forme des hélices  $\alpha$  qui lient l'ADN, et la partie N-terminale forme des feuillets  $\beta$ . La partie non conservée en C-terminal de chaque doigt est responsable de la **reconnaissance spécifique** de séquences d'ADN au niveau des gènes cibles. Sp1 se fixe à l'ADN en reconnaissant des séquences de 10 pdb riches en GC. Les amino-acides de Sp1 qui contactent l'ADN sont KHA, RER et RHK, situés respectivement dans chacun des 3 doigts de zinc (Fig.23). Sp1 est un **transactivateur** du promoteur de SV40. Celui-ci porte 6 exemplaires d'un élément riche en GC voisin du promoteur précoce. Tous les membres Sp/XKLF reconnaissent les mêmes séquences d'ADN, qui se décomposent sous les variations suivantes : **GC-(GGGGCGGGG)** et **GT-(GGTGTGGGG)**. Les membres de la famille Sp/XKLF ont toutefois des affinités différentes

pour le motif GC à cause des substitutions qui peuvent exister parmi les acides aminés des doigts de zinc.

Les boîtes GC et GT jouent un rôle important pour l'expression de très nombreux gènes ubiquitaires aussi bien que pour des gènes spécifiques d'un tissu ou viraux (Philipsen and Suske, 1999). Ces motifs GC sont impliqués dans la maintenance des états de méthylation libres des îlots CpG et sont protégés des méthylations de novo par la famille Sp (Brandeis *et al.*, 1994). Récemment, il a été montré que la Thalidomide induit des changements texturaux importants de la chromatine ; les cellules sensibles à cette drogue ayant des boîtes CG non méthylées, alors que les cellules qui lui sont résistantes, possèdent des boîtes CG hyperméthylées (Trussardi-Regnier *et al.*, 2009).

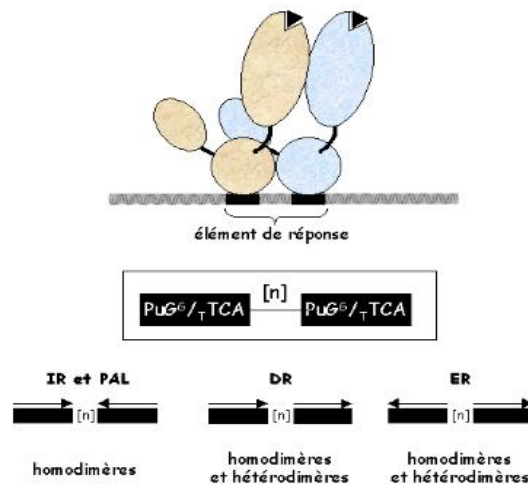
### 3. Les éléments de réponse HRE

L'analyse des régions promotrices des gènes cibles a permis de décrire les différents HRE (Leid *et al.*, 1992). Une nomenclature est utilisée pour représenter ce nombre de nucléotides. Par exemple, les éléments en répétition inverse espacés par zéro et un nucléotide sont appelés IR0 et IR1 respectivement. L'**efficacité** et la **sélectivité** de l'élément de réponse dépend : (Mader *et al.*, 1993) de l'orientation (répétition inverse, directe, éverse ou palindrome), de l'espacement, c'est-à-dire du nombre de nucléotides séparant les demi-sites, de la séquence des nucléotides contenue entre les répétitions des demi-sites, de la séquence des motifs consensus et de la séquence des nucléotides entourant l'élément de réponse (Fig.24).

Orientation	Elément de réponse	Exemples de récepteurs
Répétition directe (DR)		Homodimère : RXR/RXR Hétérodimères : RAR/RXR, RAR/TR, RAR/VDR, RXR/VDR, TR/VDR
Répétition inverse (IR) Palindrome		Homodimères : AR/AR, ER/ER, GR/GR, PR/PR Hétérodimères : RAR/RXR, TR/RXR, RAR/TR
Répétition éverse (ER)		Homodimère : TR/TR Hétérodimères : RAR/RXR, TR/RXR

**Figure 24 : Éléments de réponse avec des exemples de récepteurs pour chaque configuration (Gronemeyer and Laudet, 1995).** AR : récepteur des androgènes ; ER : récepteur des oestrogènes ; GR : récepteur des glucocorticoïdes ; PR : récepteur de la progestérone ; RAR : Récepteur de l'acide rétinoïque ; RXR : récepteur des rétinoides ; TR : Récepteur des hormones thyroïdiennes ; VDR : récepteur de la vitamine D3.

Pour réguler la transcription du gène cible, deux récepteurs nucléaires (RN) doivent se fixer sur l'ADN, un récepteur sur chaque séquence hexa-nucléotidique. Ils le font sur des séquences spécifiques d'ADN appelées **éléments de réponse aux hormones (HRE)**. Les RN peuvent agir sous la forme d'**homodimères** (les deux récepteurs nucléaires sont identiques) ou d'**hétérodimères** (les deux récepteurs nucléaires sont différents) (Fig.25). A titre d'exemple, les récepteurs des hormones stéroïdiennes agissent exclusivement en homodimères alors que les récepteurs RARs, VDR et TRs agissent le plus souvent en hétérodimères avec RXR comme partenaire.



**Figure 25 : Fixation de récepteurs nucléaires sous forme d'hétérodimères sur la boîte HRE.**

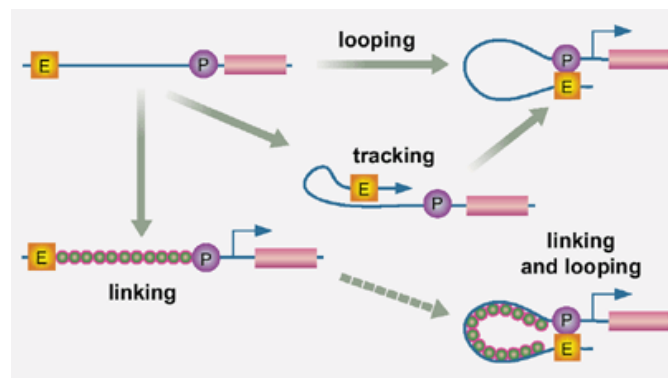
### B. Les séquences régulatrices distales

Certaines séquences d'ADN sont parfois situées jusqu'à plusieurs milliers de paires de bases du promoteur pour contrôler l'expression d'un gène. Ces séquences sont appelées «**enhancer**» lorsqu'elles sont activatrices et «**silencer**» lorsqu'elles sont inhibitrices.

Un exemple classique du rôle de séquences régulatrices distales provient de l'étude de l'expression du gène de la  $\beta$ -globine. Son expression est contrôlée, au moins en partie, par une région située 20 à 30 kb en amont et nommée **LCR**, pour région de contrôle du locus (Fraser and Grosveld, 1998; Vakoc et al., 2005). Trois modèles coexistent pour expliquer la régulation de la transcription par ces séquences (Bulger and Groudine, 2002; Tang et al., 2002).



Un premier modèle, appelé «**looping**», suggère que l'ADN se courbe et forme une boucle pour permettre aux activateurs fixés sur les régions régulatrices de contacter les facteurs de base de la transcription (Fig.26). Des analyses récentes d'immunoprécipitation de la chromatine sur les promoteurs des gènes actifs Met16 et Ino1 de la Levure, supportent ce modèle mais tendent à démontrer que l'interaction entre l'activateur CF1 et le facteur TFIIB serait à la base de cette configuration enroulée (El Kaderi *et al.*, 2009). Cette boucle faciliterait ainsi la réinitiation de la transcription en permettant le transfert direct de la Pol II du site de terminaison au site d'initiation contenant les facteurs de transcription déjà en place.



**Figure 26: Modèle d'interaction entre le promoteur (P) d'un gène et un activateur de la transcription ou enhancer (E)°(Bulger and Groudine, 2002)**

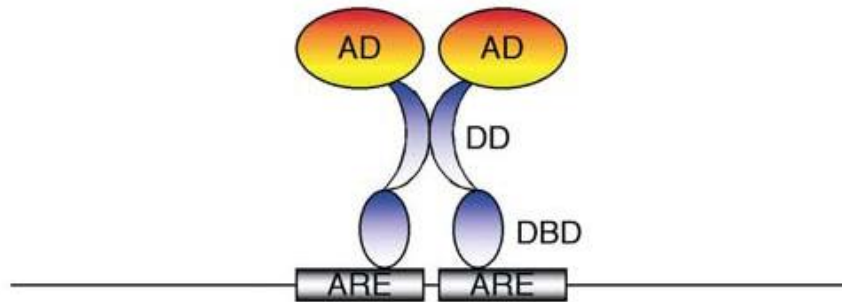
Un deuxième modèle, appelé «**linking**», suggère que les régions de contrôle permettent d'établir une structure chromatinienne (activatrice ou inhibitrice) qui se propage jusqu'au gène à réguler (Bulger and Groudine, 1999).

Le troisième modèle, appelé «**sequential**», suggère que les domaines de régulation permettent la **relocalisation du gène** dans une région du noyau riche en facteurs déstabilisants pour la compaction de la chromatine, et l'activation de la transcription du gène (Francastel *et al.*, 2001; Tang *et al.*, 2002).

Le génome compte ainsi de nombreuses séquences d'ADN qui participent à la régulation de l'expression des gènes. Pour être fonctionnelles, ces séquences d'ADN doivent être reconnues par leurs cibles protéiques. Voyons à présent quels acteurs participent à la reconnaissance de ces séquences et à l'induction de la transcription.

#### IV. Les activateurs, les répresseurs et les cofacteurs de la transcription

Les activateurs transcriptionnels possèdent une structure modulaire constituée au minimum d'un **domaine de liaison à l'ADN (DBD)** et d'un **domaine d'activation (AD)**. La plupart des protéines activatrices possèdent en outre un domaine de dimérisation (DD) qui leur permet de s'homo- ou de s'hétérodimériser (Fig.27). Chacun de ces domaines est échangeable avec le domaine correspondant d'un autre activateur (Sadowski *et al.*, 1988).



**Figure 27 : Représentation schématique d'un activateur proximal.**

AD : domaine d'activation ; DD : domaine de dimérisation ; DBD : domaine de liaison à l'ADN qui reconnaît spécifiquement un site de liaison appelé ARE pour « élément de réponse spécifique à l'activateur ».

Un des activateurs les plus employés dans les systèmes de transcription *in vitro* est la protéine de levure GAL4. Les domaines DBD et AD de cet activateur ont largement été utilisés sous forme de fusion avec différents domaines afin de comparer directement l'efficacité d'activation de multiples activateurs dans un système commun.

Les Eucaryotes supérieurs possèdent une très grande variété d'activateurs ubiquitaires ou spécifiques de certaines cellules. Les récepteurs nucléaires qui se lient à l'ADN par un domaine en doigt à zinc possèdent deux domaines d'activation : l'un dépend de la liaison par un ligand, l'autre est indépendant (Steinmetz *et al.*, 2001). Cette famille d'activateurs comprend les récepteurs aux hormones thyroïdes, stéroïdes, à la vitamine D et les récepteurs à l'acide rétinoïque (Acevedo and Kraus, 2004; Umesono *et al.*, 1988) (voir chapitre 3).

Le rôle de ces activateurs est varié : Ils peuvent faciliter le dégagement du promoteur, déstabiliser l'organisation des nucléosomes, favoriser le recrutement de GTFs, induire des changements conformationnels des GTFs ou des modifications covalentes de la chromatine au niveau des queues des histones (Fig.28).

L'activation de la transcription passe par la liaison de protéines activatrices aux séquences activatrices proximales, qui régulent alors l'activité de complexes de remodelage de la chromatine et celle de l'appareil transcriptionnel. Alors que les activateurs peuvent recruter les GTFs sur le promoteur, en accélérant le processus d'assemblage, les répresseurs peuvent inhiber la transcription en bloquant celui-ci.

L'assemblage des GTFs est lui aussi sujet à régulation par la fixation de cofacteurs positifs ou négatifs. Les cofacteurs incluent les TAFs du TFIID (sur lesquels nous reviendront dans le chapitre II), le complexe du Médiateur et les activateurs proximaux.

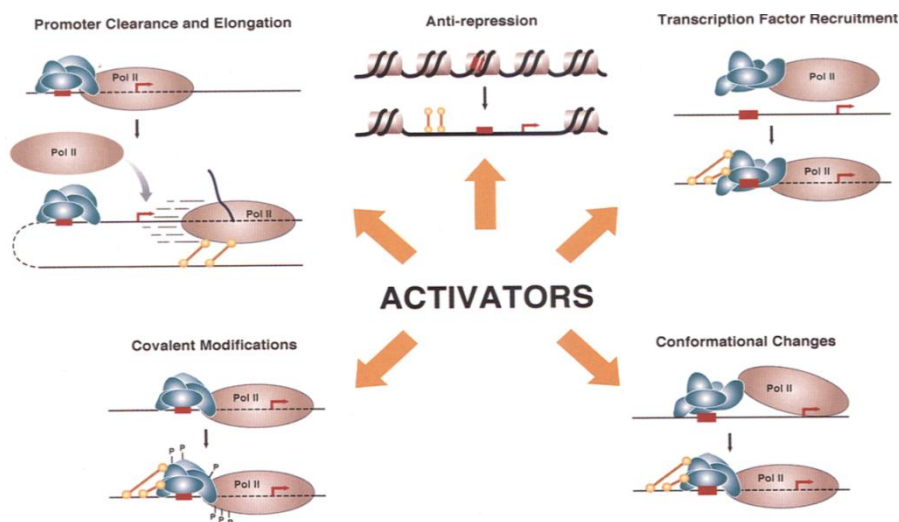


Figure 28 : Rôle des activateurs de la transcription (Orphanides *et al.*, 1996).

### A. Le complexe Médiateur

Le Médiateur est un authentique coactivateur, capable de stimuler à la fois la transcription basale (Mittler *et al.*, 2001) et la transcription dépendante des activateurs (Kim *et al.*, 1994). Cette dernière faculté repose essentiellement sur le fait que le médiateur sert de molécule de liaison dans la transduction des signaux depuis l'activateur grâce à son module de la queue, jusqu'à la Pol II via son module de la tête (Fig.29).

Les sous-unités du Médiateur interagissent pour une grande part avec des activateurs : la sous-unité MED220, par exemple, est la cible des récepteurs nucléaires TR, VDR, RAR $\alpha$ , RXR $\alpha$ , PPAR $\gamma$ , ER et AR. L'importance physiologique de ces interactions a été confirmée par des expériences KO chez *M. musculus*. Les sous-unités MED100 et MED220 sont essentielles au

développement embryonnaire, et le KO de la seconde se caractérise par la dérégulation de la transcription dépendante de certains récepteurs nucléaires (Ito *et al.*, 2000). Le complexe du Médiateur a donc indéniablement un rôle de co-activation de la transcription.

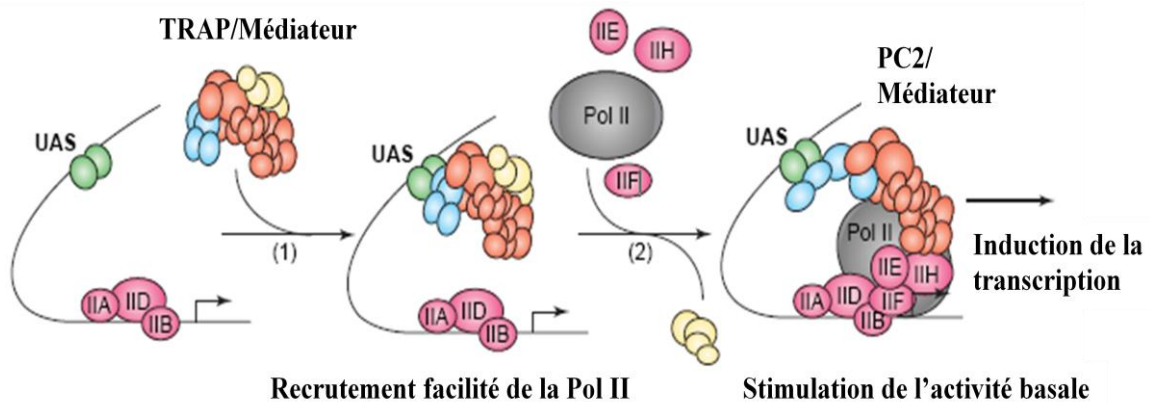


Figure 29 : L'activation de la transcription via le complexe du Médiateur (Malik and Roeder, 2005).

### B. Les cofacteurs dérivés de l'UAS

Le fractionnement biochimique des extraits nucléaires de cellules Hela (Meisterernst *et al.*, 1991) a laissé entrevoir des fractions capables de stimuler la transcription dépendante des activateurs nommées UAS (Upstream Activating Sequence) (Fig.44). Plusieurs cofacteurs positifs ont été identifiés (PC1, PC2, PC3) ainsi qu'un cofacteur négatif (NC1) (Kaiser and Meisterernst, 1996). Ces acteurs augmentent le niveau de transcription lorsque les activateurs sont présents mais le répriment en cas d'absence (Thomas and Chiang, 2006).

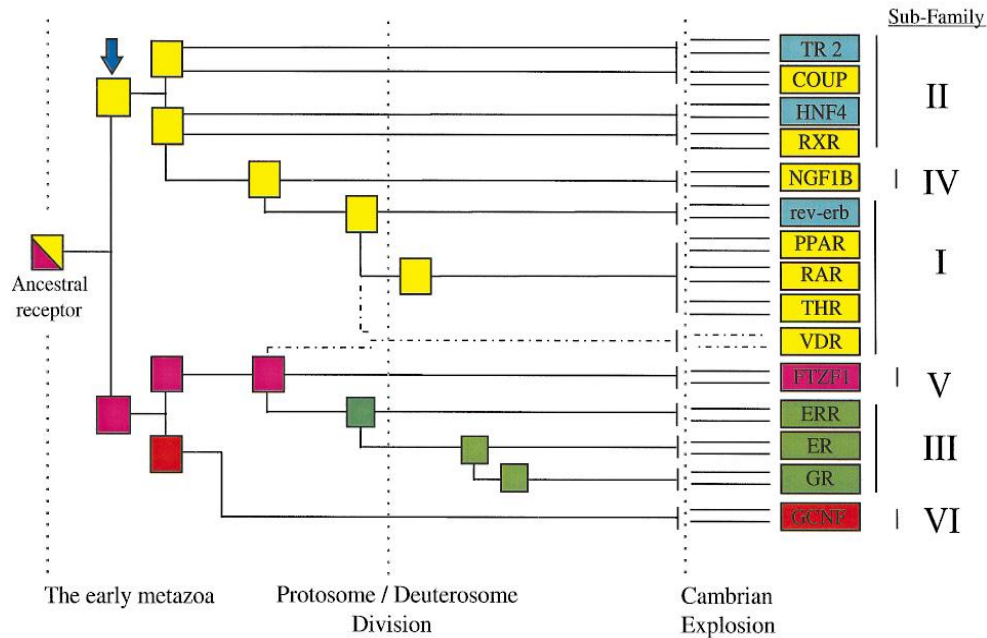
### C. Les récepteurs nucléaires

Les récepteurs nucléaires (RN) possèdent un domaine de fixation à l'ADN et des domaines d'activation de la transcription. Ils forment donc une superfamille de facteurs de transcription activés par des ligands qui régulent de nombreuses fonctions physiologiques, du développement à l'homéostasie chez les Métazoaires. Ce sont donc des activateurs de la transcription.

Des études phylogénétiques ont montré que l'apparition des RN s'était produite très tôt au cours du développement des Métazoaires et avait suivi une évolution très rapide (Escriva *et al.*, 2000).

## 1. Les différentes catégories de récepteurs nucléaires

Des alignements de séquences et la construction d'un arbre phylogénétique ont permis de définir 6 grandes sous-familles de RN (Laudet, 1997), basées sur la conservation des domaines de liaison au ligand (LBD) et à l'ADN (DBD) au cours de l'évolution (Fig.30).



**Figure 30 : Modèle hypothétique concernant l'évolution de la superfamille des RN (Owen and Zelent, 2000).**

Ces 6 grandes sous-familles sont les suivantes :

- Sous-famille 1: des hormones thyroïdiennes/vitamines
- Sous famille 2: Retinoid X Receptor-like
- Sous Famille 3: Récepteur des hormones stéroïdes
- Sous Famille 4: Nerve Growth Factor IB-like
- Sous Famille 5: Steroidogenic Factor-like
- Sous Famille 6: Germ Cell Nuclear Factor-like

Cette superfamille des RN contient non seulement des RN pour des ligands connus mais aussi un nombre important de récepteurs orphelins qui n'ont pas de ligands ou pour

lesquels aucun ligand n'a, à ce jour, été identifié. 48 RN ont été identifiés dans le génome humain (Chambon, 2004).

Chez les Métazoaires, il existe trois familles de RN (Tableau 2) :

- les RN des **hormones stéroïdiennes** (RN aux œstrogènes ou aux androgènes par exemple).
- les RN des **hormones non-stéroïdiennes** (RN des hormones thyroïdiennes ou de l'AR).
- les RN dits "**orphelins**" car aucun ligand susceptible de moduler leur activité transcriptionnelle n'a été identifié. La capacité de ces RN à lier un ligand a peut-être été perdue au cours de l'évolution (perte de fonction secondaire) ou au contraire été induite ce qui permettrait une liaison permanente de ces récepteurs à leurs ligands. Parmi ces RN, se trouvent les récepteurs orphelins associés aux récepteurs des rétinoïdes (ROR) ou ceux associés aux récepteurs des œstrogènes (ERR) (Aranda and Pascual, 2001; Gronemeyer *et al.*, 2004).

Ligand	Récepteur
<u>Hormones stéroïdiennes</u>	
Œstrogène	ER
Androgène	AR
Progestérone	PR
Glucocorticoïde	GR
<u>Hormones non-stéroïdiennes</u>	
Hormones thyroïdiennes	TR
Acide rétinoïque	RAR
Rétinoïdes	RXR
Vitamine D	VDR
Acides gras / eicosanoïdes	PPAR
Orphelin Pas de ligand connu	COUP-TF, NGFI-B, ERR ROR, SF1

**Tableau 2 : Quelques membres de la famille des récepteurs nucléaires et leur ligand.**

AR : récepteur d'androgène ; COUP-TF : facteur de transcription du promoteur de l'ovalbumine de poulet ; ER : récepteur des œstrogènes; ERR : récepteur orphelin associé aux récepteurs des œstrogènes ; GR : récepteur des glucocorticoïdes ; NGFI-B : facteur de croissance neuronale 1B ; PPAR: Récepteur des peroxyosomes en prolifération ; PR : récepteur de progestérone ; RAR : récepteur de l'acide rétinoïque ; ROR : récepteur orphelin associé aux récepteurs des rétinoïdes ; RXR : récepteur des rétinoïdes ; SF1 : récepteur orphelin, facteur stéroïdogène 1 ; VDR : récepteur de la vitamine D3.

## 2. Structure des récepteurs nucléaires

Les récepteurs nucléaires ont une **structure modulaire commune** formée de 5 à 6 régions (Fig.31), chaque région possédant un rôle bien précis (Chambon, 1996; Gronemeyer and Laudet, 1995).

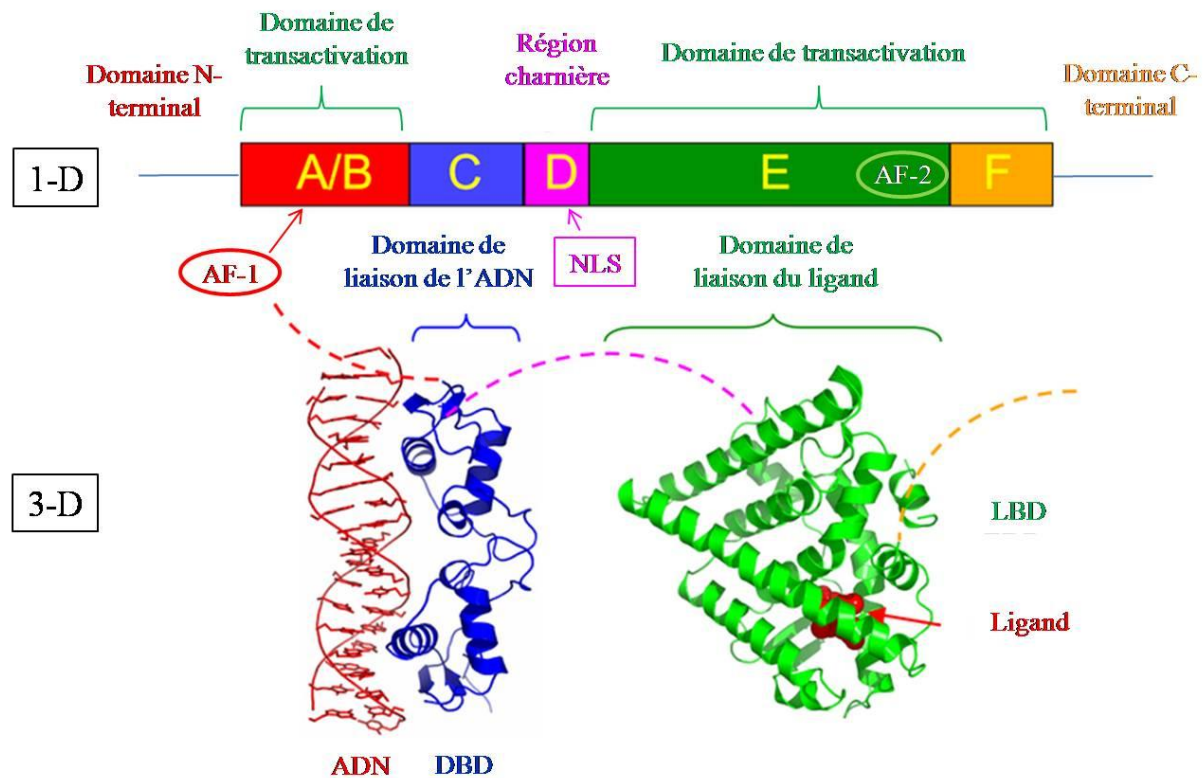


Figure 31 : Les différentes régions (A à F) et les domaines caractéristiques (DBD, LBD, AF1 et AF2) d'un récepteur nucléaire.

### a) La région A/B

La **région A/B** est située du côté amino-terminal (NH<sub>2</sub>) (Fig.31). Cette région est extrêmement variable d'un récepteur nucléaire à l'autre en termes de taille, de séquence d'acides aminés et de présence ou non du **domaine de transactivation AF-1** (Activation Function 1) (Krust *et al.*, 1986). Les capacités transactivatrices de AF-1 varient selon le récepteur considéré, l'élément de réponse, le type cellulaire et enfin selon le contexte du promoteur (possibilité d'interaction avec le facteur général de transcription TFIIB) (Bocquel *et al.*, 1989; Dowhan and Muscat, 1996; Hadzic *et al.*, 1995). Ce domaine de transactivation est indépendant de la liaison du ligand (Nagpal *et al.*, 1993). Il peut activer la transcription des gènes cibles même si aucun ligand ne s'est fixé au récepteur et peut également être reconnu

par des co-activateurs, et/ou d'autres facteurs de transcription comme la protéine de liaison à la boîte TATA (TBP) et la protéine de liaison à CREB (CBP) (Almlof *et al.*, 1998; Warnmark *et al.*, 2003).

### b) La région C

La **région C** est présente chez tous les récepteurs nucléaires (Green *et al.*, 1988) (Fig.31). Elle est structurée par deux **doigts de zinc** contenant le DBD (Freedman *et al.*, 1988; Schena *et al.*, 1989). Ce domaine permet au récepteur de reconnaître des séquences spécifiques de l'ADN, de s'y lier et d'activer les gènes cibles.

### c) La région D

La **région D** est une région charnière entre les régions C et E (Fig.31). Cette région varie d'un récepteur nucléaire à l'autre. Elle peut contenir un **signal de localisation nucléaire** (NLS) (Ylikomi *et al.*, 1992). Dans le cas du RN des hormones thyroïdiennes, la mutation de certains résidus de la région D inhibe l'interaction avec des corépresseurs (Damm *et al.*, 1993). Dans le cas du récepteur RXR, cette région est aussi impliquée dans la liaison à l'ADN (Lee *et al.*, 1993).

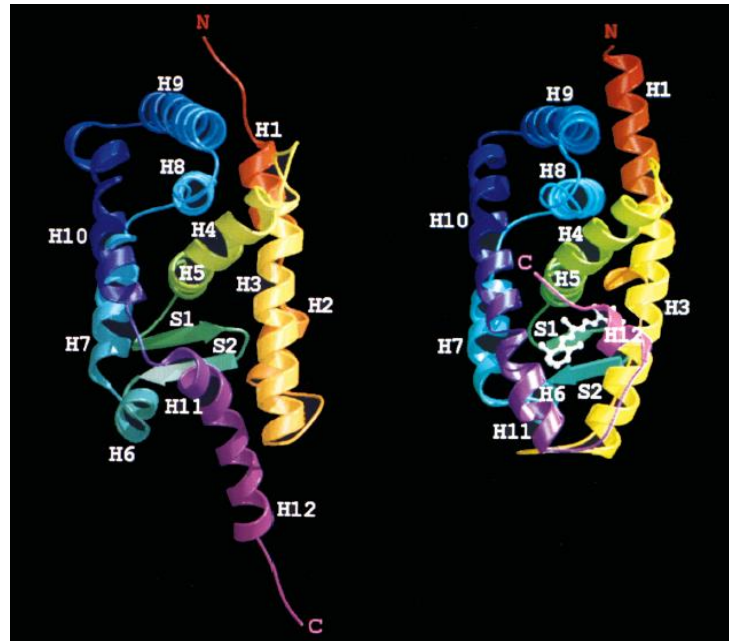
### d) La région E

La **région E** présente une très forte conservation inter- et intra-espèces (Laudet *et al.*, 1992) (Fig.31). Cette région contient le LBD et un **domaine de transactivation nommé AF-2** (Activation Function 2) (Chambon, 1996). L'activité transactivatrice de AF-2 dépend de la liaison du ligand (Nagpal *et al.*, 1993). Ceci signifie que ce domaine est actif à condition que le ligand soit lié au récepteur. Le domaine AF-2 est la cible de nombreux co-activateurs et corépresseurs (Warnmark *et al.*, 2003).

La région E contient également une **interface de dimérisation des récepteurs nucléaires** (Chambon, 1996; Leid *et al.*, 1992; Leng *et al.*, 1995). Dans certains cas, cette région contient un **signal de localisation nucléaire** (Picard and Yamamoto, 1987; Zhou *et al.*, 1994). Dans le cas des récepteurs des hormones stéroïdiennes, la région E contient aussi un domaine d'interaction avec les protéines HSP (heat shock protein) (Chambraud *et al.*, 1990; Housley *et al.*, 1990). Enfin, il a été démontré pour RXR et RAR que ce domaine contient des sites de **phosphorylation** (Bastien *et al.*, 2002; Taneja *et al.*, 1997).



Chez la plupart des récepteurs, la région E adopte un **repliement conservé** comprenant douze hélices  $\alpha$  (H1 à H12) et un feuillet- $\beta$  composé de deux brins antiparallèles localisé entre les hélices H5 et H6. L'hélice H12 correspond au domaine de transactivation AF-2. La figure 32 représente l'enchaînement des hélices  $\alpha$  et des brins du feuillet  $\beta$  pour le cas du récepteur nucléaire RXR $\alpha$  sans ligand (apo-LBD) et RAR $\gamma$  avec son ligand (holo-LBD) (Renaud *et al.*, 1995).



**Figure 32 : Diagramme en rubans des formes humaines de RXR $\alpha$  sans ligand (apo-LBD) à gauche et de RAR $\gamma$  avec son ligand (holo-LBD) à droite (Chambon, 2004).**

Le changement majeur observé entre ces deux RN est un **repositionnement de l'hélice H12** située à l'extrémité carboxy-terminale de la région E (Fig.31). Dans la forme de RXR $\alpha$  en l'absence de ligand, l'hélice H12 adopte une conformation étendue qui l'éloigne du reste de la protéine, alors que dans la forme de RAR $\gamma$  liée au ligand, cette même hélice vient se plaquer contre la structure et piège le ligand à l'intérieur de la poche de fixation (Fig.31). De plus, l'hélice H2 présente dans la forme "apo" est déstructurée dans la forme "holo" du récepteur.

L'équipe d'Egea a montré en comparant les formes apo et holo du récepteur RXR $\alpha$  que la fixation du ligand entraîne également d'**autres changements de conformation** de la région E (Egea *et al.*, 2000). Suite à la fixation du ligand et au repositionnement de l'hélice H12 qui en découle, des remaniements structuraux affectent les hélices H3 et H11 qui se

repositionnent (Fig.32). L'hélice H3 se déplace vers l'intérieur de la protéine et permet ainsi à sa partie amino-terminale d'interagir directement avec le ligand. Au contraire, l'hélice H11 est repoussée vers l'extérieur. Dans une moindre mesure, l'hélice H6 ainsi que les brins du feuillet  $\beta$  sont également concernés. Ces changements structuraux stabilisent l'ensemble de la région E.

#### e) La région F

La **région F** des récepteurs nucléaires est située du côté carboxy-terminal (COOH) (Fig.31). Elle est absente chez certains récepteurs et son taux de conservation est faible. Dans le cas de RARalpha, la présence de cette région semble indispensable à l'activité du domaine AF-2 (Tate et al., 1996). Elle pourrait modifier les propriétés des RAR grâce à ses nombreuses phosphorylations (Bastien *et al.*, 2000), stabiliser l'hélice 12 (AF-2) du sous-type de récepteur RAR $\alpha$  dans une conformation ouverte, permettant ainsi la liaison de corépresseurs (Farboud and Privalsky, 2004). Pour les RXRs, l'absence de la région F les transforme en inhibiteurs de la transcription (Damm *et al.*, 1993).

### 3. Les mécanismes de la régulation transcriptionnelle

Les éléments de réponse ont tout d'abord été localisés dans la région régulatrice en 5' des gènes cibles associés (Green and Chambon, 1988). Mais il est maintenant bien connu que le HRE peut être localisé dans **toutes les régions régulatrices**. Les récepteurs nucléaires, une fois fixés sur leurs éléments de réponse ont la capacité d'**activer** ou de réprimer l'expression de gènes cibles. Ils régulent ainsi leur transcription.

#### a) Régulation transcriptionnelle par les cofacteurs

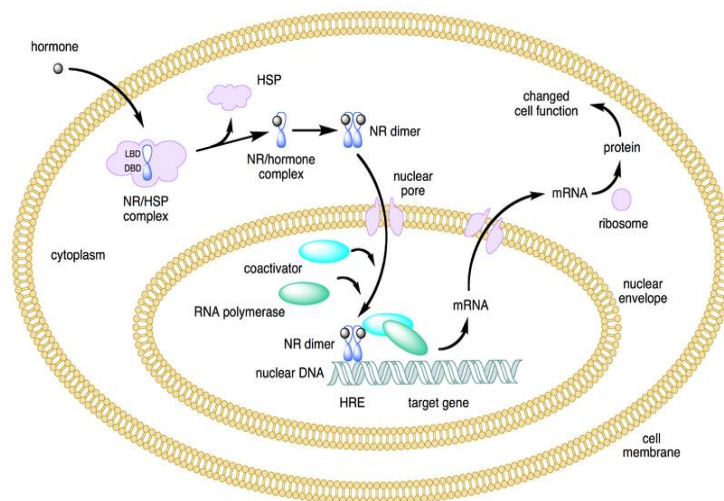
Les modèles classiques séparant le mode d'action des deux types de récepteurs tendent à s'estomper avec l'apparition de la **notion de recrutement cyclique** des récepteurs stéroïdiens au niveau de leur HRE.

#### -Les récepteurs stéroïdiens

Dans le modèle classique faisant intervenir des récepteurs stéroïdiens, les récepteurs sans ligand sont liés à des **protéines de choc thermique** (HSP) dans le cytoplasme (Fig.33). La liaison à l'hormone stéroïde enclenche la dissociation du RN et de la HSP, la dimérisation

du RN et sa translocation à l'intérieur du noyau où il pourra aller se fixer à son HRE (Pratt *et al.*, 2004).

Le complexe RN-ADN formé recrute alors des **cofacteurs transcriptionnels** tels que la famille p160 (Xu and Li, 2003) ou CBP/p300 (Wang *et al.*, 2008). La famille p160 des co-activateurs comprend les protéines SRC1 (NCoA-1), SRC2 (TIF-2, GRIP-1) et SRC3 (pCIP, ACTR, AIB1, TRAM1, RAC3) (Glass and Rosenfeld, 2000; Perissi *et al.*, 1999). Ces membres vont servir d'adaptateurs pour le recrutement de complexes protéiques ayant des activités enzymatiques (Lefebvre *et al.*, 2005). Ces derniers seront détaillés un peu plus loin.



**Figure 33 : Les récepteurs stéroïdiens et leur rôle de cofacteurs transcriptionnels.**

Cependant, ce modèle tend à être balayé par des travaux récents. G. Hager a étudié, lors d'une expérience de photo-bleaching avec le récepteur aux glucocorticoïdes (GR) lié à la GFP (Green Fluorescent Protein), la liaison de ce RN et de ses co-activateurs à son HRE sur le promoteur murin du virus tumoral MMTV (mouse mammary tumour virus).

Ses expériences *in vivo* montrent que le GR occupe son HRE pendant quelques secondes seulement (Archer *et al.*, 1994). Le modèle proposé de "**retour à la matrice**" suggère que les séquences régulatrices sont occupées de manière **cyclique** par GR, qui recrute le complexe SWI-SNF (Auwerx *et al.*, 2003). Cela est suivi d'un remodelage chromatinien et d'une dissolution du complexe pour permettre l'initiation d'un autre cycle de fixation de GR. Il s'agit donc là d'un **contraste majeur avec la vision classique de l'activation** de la transcription (John *et al.*, 2009). Cependant, des expériences d'immunoprécipitations de la chromatine sur d'autres RN stéroïdiens dans des systèmes reconstitués ou sur des cellules

vivantes suggèrent plutôt une périodicité de l'ordre de 10 minutes. Si la relation entre des échanges rapides de l'ordre de la seconde et des longs cycles d'occupation du promoteur demeure étrange, une même cinétique de **recrutement cyclique** a été trouvée pour le récepteur ER sur le promoteur pS2 (Denger *et al.*, 2008). Ce recrutement pourrait se faire en présence ou en l'absence de ligand. La cyclicité dépendrait seulement de l'activité du complexe du protéasome, également requis pour la transcription.

Ces expériences indiquent que la question pour un RN n'est plus de savoir s'il est, ou non, lié à l'ADN, puisque les RN stéroïdiens s'associent de manière cyclique avec leur HRE même en l'absence de ligand. Il s'agit donc plutôt d'examiner si des **cofacteurs** dépendant du ligand sont recrutés sur le complexe activé RN-ligand (Hong *et al.*, 2009) et si tel est le cas, il serait intéressant de connaître le temps pendant lequel ce complexe est associé au promoteur. L'image **dynamique** qui émerge à présent, suggère que la régulation transcriptionnelle est un processus continuellement ré-évalué, dans lequel les protéines et l'ADN s'associent et se dissocient régulièrement afin d'**ajuster les taux de transcription** aux conditions immédiates. Dans le cas de la transcription dépendant du RN, la concentration en ligand semble être déterminante. Contrairement aux modèles statiques classiques, cette vision dynamique de l'action des NR souligne qu'un autre paramètre est important dans ce domaine : **le temps**.

### *-Les récepteurs non stéroïdiens*

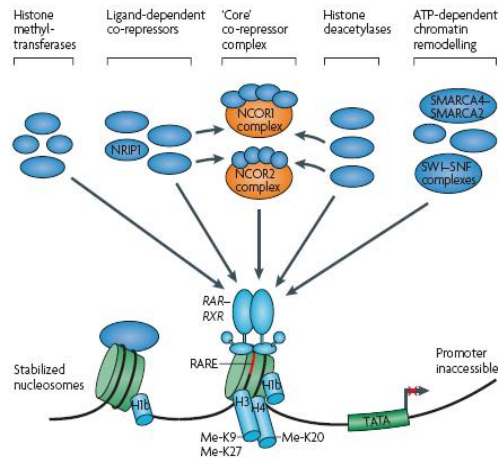
Ils sont situés dans le noyau de la cellule. A l'exception des homodimères RXR/RXR, les RN non stéroïdiens sont capables de se lier à leur HRE que leur ligand soit présent ou non (Pfahl, 1993).

### *--Absence de ligand*

En l'**absence de ligand**, l'activation des gènes est réprimée par des co-répresseurs. Les premiers corépresseurs qui ont été identifiés ont été appelés **N-CoR** (corépresseur des récepteurs nucléaires) (Horlein *et al.*, 1995; Kurokawa *et al.*, 1995) et **SMRT** (médiateur silencieux des hormones rétinoïdiennes et thyroïdiennes) (Chen & Evans, 1995).

Ces corépresseurs (Fig.34) interagissent avec le récepteur nucléaire par l'intermédiaire de son LBD (Leo and Chen, 2000). Ils se lient efficacement aux RAR et aux TR. Ils permettent une répression de la transcription grâce au recrutement de complexes ayant des activités histones **désacétylases et méthyltransférases** qui stabilisent la structure du

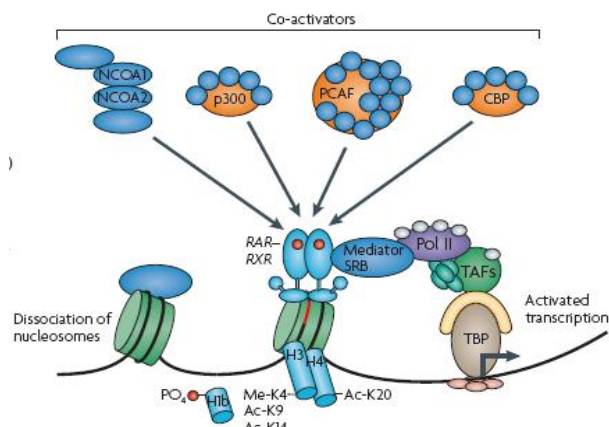
nucléosome, empêchant ainsi toute accessibilité à l'ADN. Il n'y a donc pas de transcription possible (Niederreither and Dolle, 2008).



**Figure 34 : Mécanisme de répression de la transcription par la liaison de corépresseurs à l'hétérodimère RAR/RXR en absence de ligand (Niederreither and Dolle, 2008).**

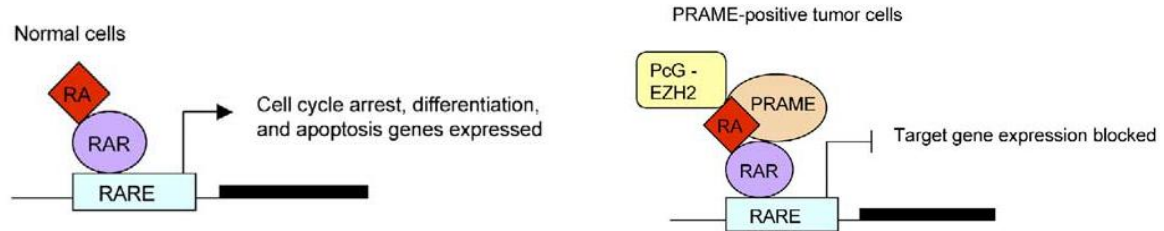
### --Présence du ligand

La fixation du ligand induit un **changement conformationnel** dans la structure du LBD permettant le départ des corépresseurs et le recrutement de complexes co-activateurs (Chambon, 1996; Germain *et al.*, 2004; Lefebvre *et al.*, 2005) sur le complexe de transcription (McKenna *et al.*, 1999; Xu *et al.*, 1999). Alors que certains co-activateurs **interagissent avec les facteurs** de la transcription basale, d'autres assurent le **remodelage de la chromatine** (Fig. 35). La fonction de ces différents co-activateurs sera expliquée de manière plus détaillée dans la partie suivante intitulée « les modifications covalentes de la chromatine ».



**Figure 35 : Mécanisme d'activation de la transcription grâce au recrutement de co-activateurs.**

Il arrive dans certains cas que les ligands recrutent des protéines telles que RIP140, TIF1 $\alpha$  ou PRAME (Fig. 36) qui interagissent alors avec les récepteurs afin de réprimer la transcription (Cavailles *et al.*, 1995; Epping *et al.*, 2005; Le Douarin *et al.*, 1996).



**Figure 36: Modèle de répression par Prame de la voie de signalisation à l'acide rétinoïque (Epping *et al.*, 2005).**

Dans ce cas, la fixation du ligand induit une répression de la transcription des gènes cibles. Cette situation se rencontre surtout lors de la tumorigenèse cellulaire (Fig.35).

### 6) Régulation transcriptionnelle via la phosphorylation

La **phosphorylation** constitue un niveau supplémentaire de contrôle de l'activité de la transcription des récepteurs nucléaires. Il s'agit d'une **modification post-traductionnelle** commune aux protéines. La plupart des membres de la famille des récepteurs nucléaires sont phosphorylables.

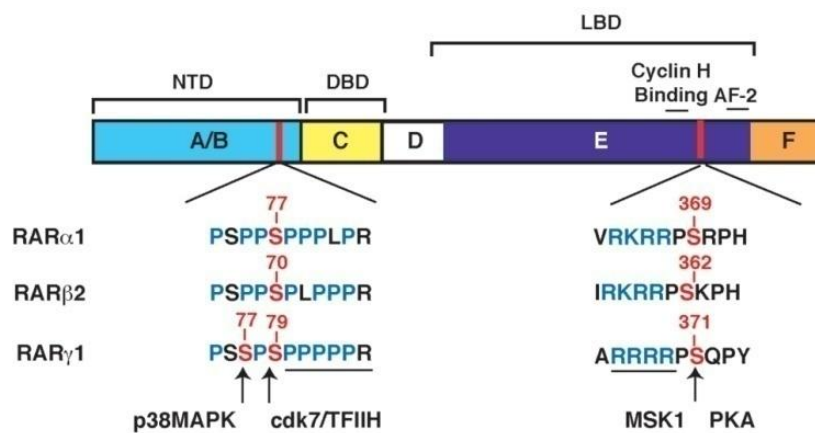
#### -la phosphorylation des récepteurs nucléaires

*In vivo*, les récepteurs des hormones stéroïdiennes (androgènes, oestrogènes, glucocorticoïdes, progestérone), le récepteur de la vitamine D3, les récepteurs d'hormones thyroïdiennes et les récepteurs de l'acide rétinoïque sont phosphorylables (Auricchio, 1989; Fahrner *et al.*, 1990; Gaub *et al.*, 1992; Glineur *et al.*, 1990; Rochette-Egly *et al.*, 1991; van Laar *et al.*, 1991; Weigel, 1996).

Plusieurs kinases peuvent **phosphoryler les récepteurs nucléaires** sur différents résidus. Parmi ces kinases, on trouve la **caséine kinase II** (Bodwell *et al.*, 1991), la **MAP kinase** (Katoch and Moreland, 1995), la **protéine kinase A** (Chen *et al.*, 1999; Rochette-Egly *et al.*, 1995), la protéine kinase C (Tahayato *et al.*, 1993) ou la **kinase dépendante des cyclines cdk7**, une sous-unité du facteur général de transcription et de réparation de l'ADN, TFIIF (Bastien *et al.*, 2000; Rochette-Egly *et al.*, 1997).

-le cas des RARs

Le **DBD** de RAR $\alpha$  par exemple est phosphorylable au niveau de la sérine 96 par la kinase Akt (Srinivas *et al.*, 2006). Cette sérine est très conservée parmi les isoformes de RAR. Sa phosphorylation ne nuit pas à la fixation des hétérodimères RAR $\alpha$ /RXR $\alpha$  au niveau de leurs éléments de réponse mais entraîne une **diminution de la transactivation** des gènes rapporteurs en présence d'AR. Le DBD de RAR $\alpha$  peut également être phosphorylé au niveau de la sérine 157 par la protéine kinase C (PKC) (Delmotte *et al.*, 1999). Cette région est importante pour la **dimérisation des récepteurs** (Quack *et al.*, 1998) ; ce résidu de sérine est bien conservé et toute mutation à ce niveau peut engendrer des troubles de dimérisation des récepteurs ou de fixation à l'ADN.



**Figure 37 : Représentation schématique des domaines fonctionnels et des sites de phosphorylation majeurs des récepteurs nucléaires à l'acide rétinoïque (Bastien and Rochette-Egly, 2004).**

Les **domaines AF-1 et AF-2** des RARs n'échappent pas non plus aux phosphorylations. La kinase cdk7 phosphoryle en effet les sérines 77 de RAR $\alpha$  et 68 de RAR $\gamma$ 2 situées au niveau de la région B (Fig.37). Tout défaut de cette phosphorylation à cause d'une mutation de TFIIH aboutit à une diminution considérable de l'expression des gènes cibles des rétinoïdes (Keriel *et al.*, 2002). De plus, les sérines phosphorylables situées dans AF-1 se trouvent au cœur d'un motif **riche en proline** qui est un motif d'interaction avec les protéines possédant un **domaine SH3 ou WW**. Toute modification de la phosphorylation aura donc des effets directs sur la liaison de ces protéines (Macias *et al.*, 2002). De même, la **protéine kinase A (PKA)** phosphoryle les sérines 369 de RAR $\alpha$  et 360 de RAR $\gamma$ 2, pouvant entraîner ainsi le départ de corépresseurs liés à RAR (Fig.37).

### -le rôle de cette phosphorylation ?

L'état de phosphorylation des récepteurs nucléaires peut considérablement moduler leurs fonctions (Jackson, 1992).

Tout d'abord, il peut modifier la capacité du récepteur à **se lier sur son élément** de réponse. Par exemple, Lefebvre et al. ont montré que la déphosphorylation de RAR $\alpha$  ou de RXR $\beta$  altère la liaison du récepteur sur l'ADN (Lefebvre *et al.*, 1995). De même, la phosphorylation du résidu sérine 369 de RAR $\alpha$  augmente la capacité de l'hétérodimère RAR/RXR à se fixer sur l'ADN (Rochette-Egly *et al.*, 1995).

Ensuite, l'état de phosphorylation peut réguler **le recrutement de coactivateurs** ou de **corépresseurs** (Dilworth *et al.*, 2000). Par exemple, la phosphorylation de ER alpha (Hammer *et al.*, 1999) et du récepteur orphelin SF-1 (Tremblay *et al.*, 1999) permet l'interaction avec les coactivateurs SRC-1 et GRIP-1 respectivement.

Enfin, l'état de phosphorylation contrôle la **stabilité des récepteurs nucléaires** et régule ainsi leur quantité. Les phosphorylations de RAR, PPAR et ER jouent un rôle essentiel dans la dégradation de ces récepteurs par l'ubiquitine-protéasome (Floyd and Stephens, 2002).

#### D. Les modifications de la chromatine

Les histones sont les pièces maîtresses de la compaction de l'ADN en chromatine et jouent un rôle majeur dans la régulation des fonctions du génome. Elles sont les cibles de multiples modifications post-traductionnelles (MPT) qui apportent une information épigénétique. Les MPT les plus courantes sont les suivantes : l'acétylation et la méthylation des lysines et des arginines, la phosphorylation des sérines et des thréonines, l'ubiquitination, la sumoylation, la ribosylation des lysines ou encore la citrullination des arginines (Cuthbert et al., 2004). L'ensemble de ces modifications constituerait un "code histone", permettant d'associer à chaque combinaison de modifications un état particulier de la chromatine.

Ces modifications régulent la structure et la fonction de la chromatine en recrutant des protéines additionnelles (Ajiro and Allis, 2002) et en facilitant la liaison des activateurs et des GTF à l'ADN chromosomique. Des **co-activateurs recrutés** par les activateurs et capables d'agir sur les histones, contribuent en plus à faciliter l'activation de la transcription.



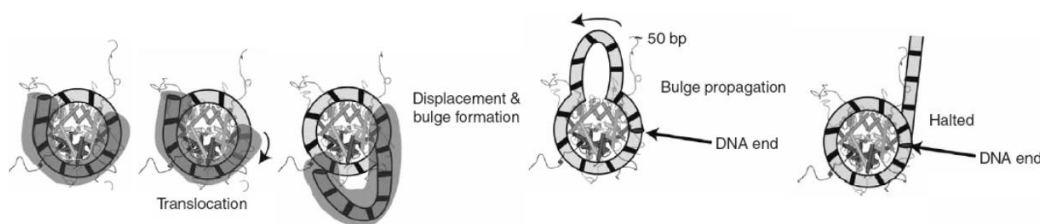
Parmi les familles de co-activateurs qui modulent la structure chromatinienne, on distingue les HAT, les arginines méthyltransférases (HMTs) comme CARM1 et PRMT1, les ubiquitinasés/dés-ubiquitinasés et les complexes de remodelage des nucléosomes comme SWI/SNF.

### 1. Les complexes de remodelage

Les complexes eucaryotes protéiques utilisent de l'énergie dérivée de l'**hydrolyse de l'ATP** pour réorganiser la structure de la chromatine et déplacer les contacts entre histones et ADN (Kornberg and Lorch, 1999; Workman and Kingston, 1998). Cette activité est considérée comme **co-activatrice indirecte** puisqu'elle favorise la liaison d'activateurs et promeut la transcription en altérant (Bouazoune *et al.*, 2009) ou en déplaçant les structures répressives chromatinienes (Fig.38).

Le complexe SWI/SNF en est le prototype majeur. Il possède une activité ATPase intrinsèque (Laurent *et al.*, 1991) perturbant les nucléosomes *in vitro* et facilitant la fixation de facteurs de transcription (Kwon *et al.*, 1994) ou réprimant la fixation de répresseurs transcriptionnels (Tyler and Kadonaga, 1999).

Les complexes **SWI/SNF** peuvent déplacer les nucléosomes, soit d'un brin d'ADN à un autre, soit sur le même brin d'ADN, par glissement (Whitehouse *et al.*, 1999). SWI/SNF peut encore rendre l'ADN accessible, soit en conservant les histones en place, soit en les remplaçant par des variants (Mohrmann and Verrijzer, 2005).



**Figure 38 : Le « bourgeonnement des nucléosomes (Zofall *et al.*, 2006).**

Outre le fait que SWI/SNF pourrait altérer les contacts entre ADN et histones, une nouvelle hypothèse, depuis 2006, fait état de la formation de "**bourgeons**" ou de "**boucle puis recapture**" (Fig.38) pour expliquer l'avancée des nucléosomes et la **décompaction** de la chromatine. L'ADN se dissocierait à un coin du nucléosome, formant une bulle d'ADN à la surface de l'octamère. La boucle d'ADN se propagerait ensuite à travers la surface de

l'octamère comme une onde. Il en résulterait un repositionnement de l'ADN sans modification dans le nombre total de contacts entre histone et ADN.

L'expression de certains gènes requiert une **action concertée** des **complexes remodelant la chromatine** SWI/SNF et de protéines impliquées dans **l'acétylation des histones** telles que Gcn5.

## 2. Les histones acétyltransférases (HAT)

L'acétylation des histones se caractérise par le transfert d'un groupement acétyl provenant d'un donneur tel que l'acétylcoenzyme A sur les lysines des histones. De nombreuses HAT (Histone Acetyl Transferase) et HDAC (Histone Deacetylase) ont été découvertes ces dernières années. Des cofacteurs transcriptionnels, tels que p300, CBP, hGCN5 ou PCAF en possèdent. Le recrutement d'activateurs ayant une activité HAT permet l'acétylation des histones (Berndsen and Denu, 2008; Selvi and Kundu, 2009) et contribue à la dérégulation locale de la chromatine (Fig.39).

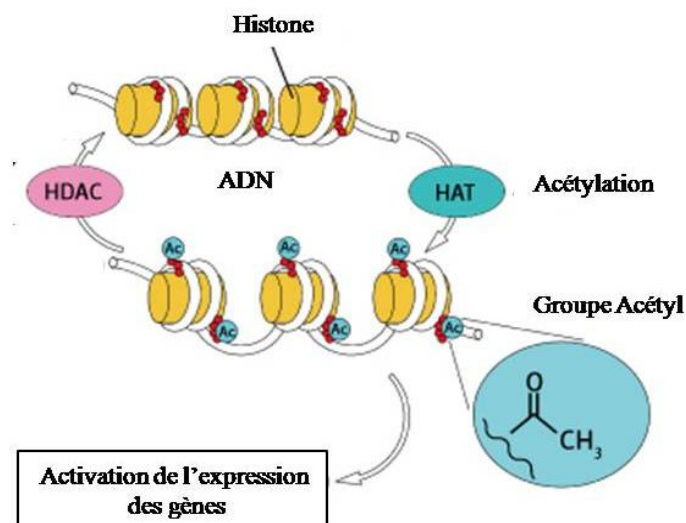


Figure 39 : Rôle des HDACs et des HATs dans l'acétylation des histones et implication dans la transcription des gènes (Wade, 2001).

### a) L'implication de HAT dans la transcription

L'acétylation constituerait une sorte de signal pour recruter les facteurs impliqués dans l'initiation de la transcription. Il a été proposé que l'acétylation des histones chromatiniens au

niveau des promoteurs conduit à une structure de la chromatine moins compacte, permettant **l'accessibilité de l'ADN aux facteurs de transcription**. Comment les modifications en N-terminal des histones peuvent-elles modifier la structure de la chromatine ?

Certains modèles proposent une implication directe de l'acétylation des histones dans les changements de la structure de la chromatine :

- l'acétylation des histones permettrait de **restructurer la région promotrice**, nécessaire au recrutement et à la stabilisation de toute la machinerie de transcription (Anderson et al., 2001; Sewack et al., 2001) ;
- l'acétylation des histones pourrait **induire des changements de charges** (neutralisation) entraînant une diminution du contact ADN/histones, ou une altération des interactions histone/histone entre nucléosomes voisins.

Des études physiques ont montré qu'in vivo, l'hyperacétylation des histones était sans conséquence sur les propriétés hydrodynamiques et sur l'enroulement de l'ADN autour du nucléosome (Peterson and Laniel, 2004). L'acétylation des histones n'est peut-être pas directement responsable d'une modification de la structure chromatinienne. Les modifications spécifiques des histones exercent un certain contrôle sur la liaison de diverses protéines à la fibre de la chromatine. C'est l'hypothèse du **code des histones**, qui est maintenant bien admise.

Que le rôle des HATs soit direct ou indirect, leur importance pour établir **un contexte chromatinien optimal** pour initier la transcription n'est plus à mettre en doute. Ainsi ces dix dernières années, on a pu déceler des activités HATs pour de nombreux coactivateurs, comme GCN5, PCAF (Yang et al., 1996), p300/CBP (Bannister and Kouzarides, 1996 ; Ogryzko et al., 1996), ACTR/Src1(Chen et al., 1997 ; Spencer et al., 1997), TIP60 (Yamamoto and Horikoshi, 1997) et TAF 1(Mizzen et al., 1996). Ces différentes enzymes possèdent de nombreuses homologues de séquences, ce qui a permis de regrouper ces protéines en trois superfamilles : GNAT, MYST et p300/CBP.

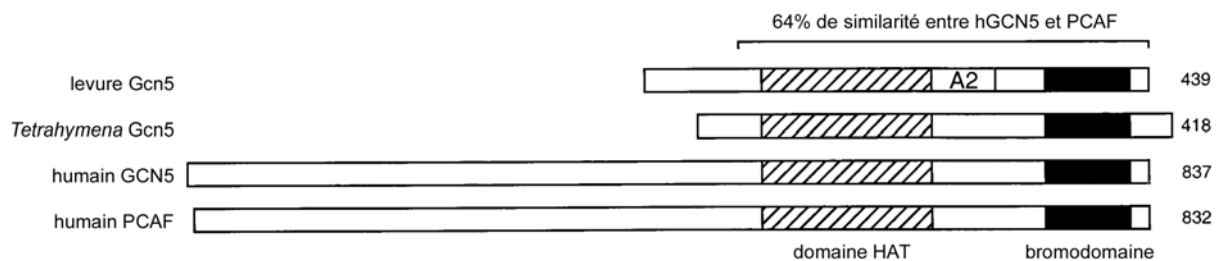
#### *6) La superfamille GNATs (Gcn5-related acetyltransferases)*

Ces protéines ont été regroupées sur la base de quelques régions homologues, notamment du **motif lié à l'acétylation**. Ce groupe comprend les protéines de levure Gcn5,

Hat1, Elp3 et Hpa2, et une variété d'autres HATs eucaryotes (GCN5, PCAF) et procaryotes à substrats différents. Ces protéines sont donc très **conservées** au cours de l'évolution.

### -les motifs conservés

Cette superfamille est définie par la conservation de 4 motifs, C, D, A et B (de N-terminal vers C-terminal) dont la fonction n'est pas complètement comprise (Fig.40). Le motif C est retrouvé dans la majorité des membres de la famille GNAT, mais pas dans la plupart des HATs connues. Le motif A constitue la **région la plus conservée** et elle est retrouvée dans d'autres familles de HATs, comme la famille MYST. Ce motif contient un segment Arg/Gln-X-X-Gly-X-Gly/Ala qui est spécifiquement impliqué dans la reconnaissance et la liaison d'un substrat : l'acétyl-CoA (Dutnall et al., 1998).



**Figure 40 : Alignement des protéines de la superfamille GNAT, le sous-groupe Gcn5 d'après (Sterner and Berger, 2000)**

Ces HATs acétylent principalement l'histone H3. Les protéines de cette famille sont importantes pour la régulation de la **croissance cellulaire et du développement**. Il a été montré que ces protéines GNAT font partie de complexes possédant des rôles de coactivateurs transcriptionnels et de réparation de l'ADN.

### -GCN5 et PCAF

Chez les mammifères, deux protéines homologues de Gcn5 chez la levure ont été découvertes, GCN5 et PCAF (p300/CREB-binding Associated Factor), dont les séquences sont hautement conservées (70 % d'identité et 80% de similarité). Ces protéines humaines semblent capables d'exercer un rôle d'adaptateur et d'acétyler les histones, comme la protéine de levure Gcn5. Si le **domaine HAT** est bien sûr indispensable à la fonction d'acétylation de GCN5 (Fig.40), deux autres domaines semblent capables d'influencer l'activité HAT ainsi que le choix du substrat. Le bromodomaine N-terminal de cette protéine est important dans la reconnaissance du substrat (Xu et al., 1998b). Lorsque ce **bromodomaine** se lie à une autre

protéine comme la protéine kinase DNA/PK, l'activité HAT de GCN5 se trouve alors inhibée (Barlev et al., 1998). Chez les Vertébrés, deux formes distinctes de GCN5 sont produites par épissage alternatif : une forme courte, qui n'a jamais été bien caractérisée et une forme longue (hGCN5-L). La forme longue **hGCN5-L** et la **protéine PCAF** se ressemblent beaucoup : leurs tailles sont identiques, leurs séquences homologues, elles interagissent toutes les deux avec p300/CBP. Ces deux protéines recombinantes sont capables d'acétyler *in vitro* des histones H3 (Lys14) et d'une manière moins importante H4 (Lys8), qu'elles soient sous forme libres ou bien agencées en nucléosomes (Schiltz et al., 1999 ; Yang et al., 1996).

c) *p300/CBP*

**CBP et p300** sont des protéines co-activatrices quasi-homologues qui lient un certain nombre de séquences spécifiques d'activateurs transcriptionnels. Ces deux protéines ont tout d'abord été isolées indépendamment, p300 pour son interaction avec **E1A**, et CBP pour son interaction avec une forme phosphorylée du facteur de transcription CREB. Ces protéines se sont avérées interchangeables dans la plupart de leurs fonctions (Fig.41).

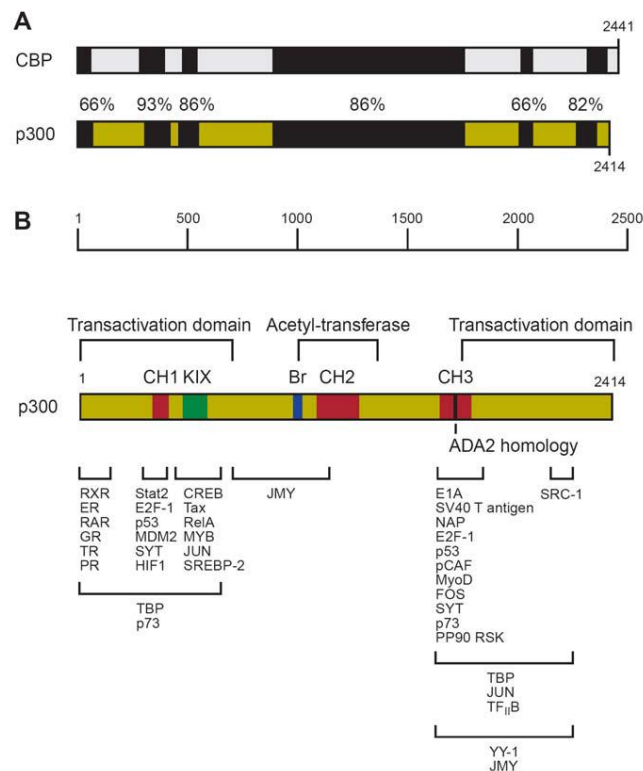


Figure 41 : Organisation des protéines CBP/p300, d'après (Chan and La Thangue, 2001).

Dans diverses études, *in vitro* et *in vivo*, la présence de p300 s'est avérée cruciale pour l'activation de la transcription de nombreux gènes. L'activation de la transcription par les récepteurs à l'acide rétinoïque RAR/RXR, les récepteurs thyroïdiens (Lee et al., 2003b), les récepteurs aux oestrogènes, à la progestérone ou aux glucocorticoïdes (Acevedo and Kraus, 2003) est augmentée par la présence de la protéine p300. p300 favorise donc l'activité de récepteurs nucléaires (Fig.42).

De même, p300 est capable d'interagir avec toute une panoplie de facteurs impliqués dans l'activation de la transcription, comme c-jun, c-myb, c-fos, myoD (Polesskaya et al., 2001), E2F-1 (Dyson, 1998), SRC-1 (Yao et al., 1996), CREB (Yuan and Gambia, 2001), p53 (An et al., 2004; Barlev et al., 2001 ; Espinosa et al., 2003 ) et leur acétylation par p300 est indispensable à la fonction de transactivation de ces protéines, puisque des mutations dans le domaine HAT de p300 entraînent la perte de transactivation.

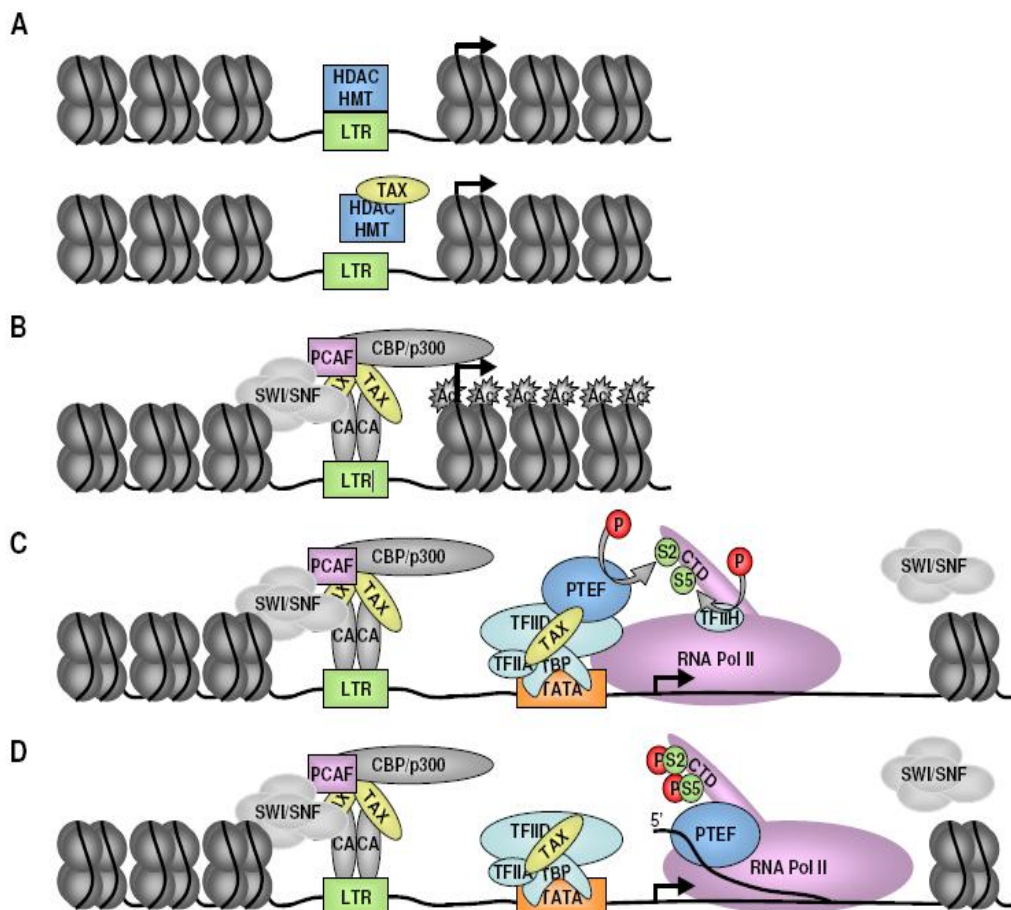


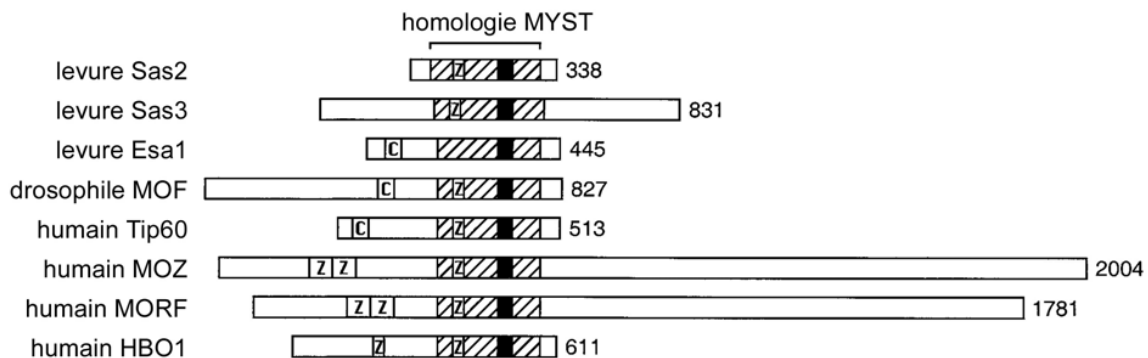
Figure 42 : Rôle de CBP/p300 dans l'activation transcriptionnelle

La protéine p300 participe donc à l'activation de la transcription de nombreux gènes, en adéquation avec ses fonctions physiologiques. Le rôle de p300 dans le contrôle du cycle cellulaire peut s'expliquer en partie par son implication dans l'expression de gènes induite par E2F-1. Egalement, le rôle de p300 dans la différenciation des cellules en cellules musculaires pourrait s'expliquer par son implication dans l'activation de la transcription aidée par MyoD, qui déclenche le programme de différenciation des cellules en cellules musculaires.

Le transfert de groupements acétyles peut également avoir lieu sur d'autres substrats et favoriser l'activité transcriptionnelle de gènes spécifiques (Gu and Roeder, 1997). Ainsi, des coactivateurs existant dans des complexes à plusieurs sous-unités tels que STAGA (SPT3-TAF9-GCN5), TFTC (TBP-free TAF complex) ou PCAF (Nagy and Tora, 2007), homologues fonctionnels du complexe SAGA (Spt-ADA-GCN5) (Baker and Grant, 2007) chez la Levure, possèdent une activité acétyltransférase. Ils agissent ainsi comme des cofacteurs positifs pour l'activation de facteurs nucléaires *in vivo*.

#### d) La famille MYST

Cette famille de protéines est nommée MYST, sigle composé des initiales des quatre protéines MOZ, Ybf2/Sas3, Sas2, Tip60, découvertes en premier (Fig.43).



**Figure 43: Alignement des protéines de la famille MYST, d'après (Sterner and Berger,2000).**

Par la suite, la famille MYST s'est agrandie aux protéines Esa1 chez la levure, MOF chez la drosophile, HBO1 et MORF chez les mammifères. Les membres de cette famille présentent une grande homologie dans leur domaine HAT, qui comprend un site de liaison à l'acétyl-CoA (apparenté au motif A de la famille GNAT), ainsi qu'un doigt à Zn atypique et

un motif E-R (Esa1-Rpd3), tous deux requis pour l'activité enzymatique et la reconnaissance du substrat (Akhtar and Becker, 2001). Bien que les protéines de cette famille possèdent des régions très similaires en séquences, elles sont impliquées dans une grande variété de fonctions régulatrices dans divers organismes.

La région d'homologie comprend un site de liaison à l'acétyl CoA représenté par une boîte noire, Z, un motif en doigt à zinc, atypique. TIP60 impliquée dans l'activation de la transcription par p53. **La protéine TIP60** et le suppresseur de tumeur p53 partagent certaines fonctions. Ces deux protéines sont impliquées dans la réponse cellulaire déclenchée par des **dommages d'ADN**. Elles sont dégradées par le protéasome après **ubiquitination** aidée par Mdm2, et s'accumulent après une irradiation UV. Chez les mammifères, p53 participe au bon déroulement de la réponse des cellules après induction de dommages à l'ADN. p53 est un facteur de transcription qui, après induction de dommages à l'ADN, régule la transcription de nombreux gènes impliqués dans le contrôle de la prolifération des cellules et dans l'apoptose ou la réparation de l'ADN. En l'absence de la protéine p53, la réponse cellulaire aux cassures d'ADN est anormale et conduit à l'accumulation de mutations dans le génome (Michael and Oren, 2002).

En absence d'ADN endommagé, l'activité de p53 est contrôlée par le proto-oncogène Mdm2. Mdm2 se lie directement à p53 et réprime son activité en accélérant sa dégradation via le protéasome, et en **inhibant l'activation de la transcription par p53**. Lorsque l'ADN est endommagé, p53 est phosphorylée par la voie de signalisation ATM/ATR chk1/chk2, ce qui entraîne sa dissociation de Mdm2. p53 est alors libre d'activer la transcription de certains gènes. TIP60 est impliqué dans la voie de signalisation p53 à plusieurs niveaux. Tout d'abord, elle contribue à maintenir un niveau basal de p53 dans les cellules en contrecarrant la dégradation de p53. Puis, après induction de cassures d'ADN, TIP60 fonctionne comme coactivateur de p53 (Legube et al., 2004). Chez la levure, le complexe NuA4, qui contient l'HAT Esa1, est également requis pour l'activation de la transcription par p53 (Eisen et al., 2001).

TIP60 participe à l'activation de la transcription. Pourtant, elle est également douée d'une activité co-répressive de la transcription, notamment lors de l'activation de la transcription par STAT3, quand TIP60 se lie à une enzyme déacétylant les histones, HDAC7 (Xiao et al., 2003). TIP60 joue aussi un rôle dans la transcription des gènes ribosomiques



(Halkidou et al., 2004), ce qui implique TIP60 dans la régulation de la transcription d'un large éventail de gènes.

Outre son activité HAT, TIP60 est capable d'acétyler d'autres substrats que les histones par exemple le récepteur aux androgènes AR. L'acétylation de AR va conduire à l'activation de la transcription de gènes cibles (Gaughan et al., 2002). Ainsi TIP60 active la transcription en acétylant soit directement les histones, soit divers activateurs ou coactivateurs. La multiplicité de fonctions qu'exerce TIP60 suggère que l'environnement (protéines associées) de cette protéine puisse modifier son activité.

### *-Équilibre entre les activités HAT et HDAC*

Si bien des progrès ont été réalisés dans la compréhension de la structure et de la fonction des HATs et des HDAC individuellement, il reste à expliquer comment l'activité opposée de ces deux classes d'enzymes est régulée. Les HATs et les HDACs peuvent se trouver assez proches dans des cellules *in vivo*. Des HATs comme PCAF ou GCN5 ont même été retrouvées en complexe avec HDAC1. D'après la même étude, le complexe formé par les HATs et les HDACs serait détruit lors de la différenciation cellulaire, libérant les HATs qui peuvent alors activer la transcription de certains gènes (Yamagoe *et al.*, 2003). La position globale et le recrutement des HATs et des HDACs dans le génome de levure ont été étudiés : Gcn5 et Esa1 sont globalement recrutés au niveau des promoteurs de gènes activement transcrits, tandis que les déacétylases Hst1 et Rpd3 sont recrutées plus spécifiquement au niveau des promoteurs de gènes impliqués dans des fonctions particulières (Robert et al., 2004). Puisque les HATs et les HDACs peuvent être recrutées ensemble au niveau d'un promoteur, il faudrait qu'un signal particulier active une des deux classes d'enzymes à un moment précis, soit pour activer, soit pour inhiber la transcription des gènes. L'équilibre entre activité HAT et activité HDAC doit être très sensible et donc contribuer à réguler finement l'expression de gènes particuliers.

### *3. Phosphorylation des histones et activation de la transcription*

Dès les années 1990, on a suspecté une implication de la phosphorylation des histones dans l'induction de la transcription des gènes dits précoces, comme le gène *c-Fos* chez l'humain (Mahadevan et al., 1991). Ces études ont été suivies par la mise en évidence d'une famille de protéines, Rsk/Msk, qui régule directement la phosphorylation des histones

(Clayton and Mahadevan, 2003). Cette **modification post-traductionnelle** est associée à la condensation des chromosomes et à leur ségrégation lors de la mitose et de la méiose (Gurley et al., 1978). Elle prend son origine au niveau de l'hétérochromatine péricentromérique et se disperse au niveau de l'ensemble du génome lors de la transition de phase G2/M (Hendzel et al., 1997). En outre, la phosphorylation de S10-H3 est nécessaire à l'initiation de la **condensation des chromosomes** chez les mammifères. L'induction de gène heat-shock chez la drosophile s'accompagne d'une forte augmentation de la phosphorylation de la Ser3 de l'histone H3 (Nowak and Corces, 2004). Ainsi la phosphorylation de la Ser10 de l'histone H3 est associée à une transcription active. A l'inverse, la phosphatase PP2A (Protein Phosphatase 2A) qui retire le phosphate de la Ser10 de l'histone H3 affecte la transcription en l'inhibant (Nowak and Corces, 2004). L'implication de ces **protéines kinases** dans la transcription suggère qu'elles soient spécifiquement recrutées au niveau de promoteurs, en tant que coactivateurs de la transcription. Une étude chez la levure montre que la phosphorylation de l'histone H3 facilite le recrutement de la protéine TBP au niveau du promoteur (Lo et al., 2005).

#### 4. Méthylation des histones

Les enzymes qui méthylent les histones peuvent être classées en deux catégories, selon leur spécificité de substrat : les lysines ou les arginines de l'extrémité N-terminale des histones. Il existe deux types d'HMT (Histone Methyl Transférase) capables de modifier les histones sur des résidus arginine : CARM1 (Co-Activator-associated arginine Méthyltransferase 1) et PRMT1 (Protein Arginine Méthyltransferase 1). CARM1 est un coactivateur des récepteurs nucléaires qui peut méthyler *in vitro* les histones H3 et H2B sur plusieurs résidus arginine, ce qui constitue un événement important pour l'activation de la transcription de gènes cibles (Chen et al., 1999a). En effet, **CARM1** peut, par exemple, méthyler CBP sur son arginine 742, permettant ainsi à CBP d'exercer sa fonction de co-activateur pour l'activation de gènes induits par les hormones stéroïdes (Chevillard-Briet *et al.*, 2002). Les **PRMTs**, pour leur part, sont des HATs spécifiques du résidu arginine en position 3 de l'histone H4. La méthylation de l'Arg3 facilite l'acétylation de l'histone H4 par p300/CBP, ce qui va conduire à l'activation de la transcription (Wang et al., 2001a).

Des méthylations peuvent également survenir sur les **résidus lysines** des histones et un résidu peut accepter un, deux, ou trois groupes méthyls. Plusieurs études décrivent la

méthylation de la Lys9 de l'histone H3 comme une marque de la **répression de la transcription**. L'enzyme Suv39H1, spécifique de cette méthylation, est impliquée dans le "silencing" péri-centromérique, mais également dans le silencing des gènes localisés au niveau de l'euchromatine (Vandel et al., 2001). Or, certaines enzymes, contenant un **chromodomaine**, sont capables de reconnaître spécifiquement la Lys9 de l'histone H3 méthylée. C'est le cas de la protéine HP1 qui recrute une déacétylase probablement capable de déacétyler les nucléosomes voisins. Ceci pourrait expliquer comment la méthylation de la Lys9 de l'histone H3 conduit à l'inhibition de la transcription. Les régions inactives en transcription sont souvent associées à des méthylations de l'ADN qui s'effectuent particulièrement sur les cytosines des dinucléotides CpG. Chez les vertébrés, environ 60 à 90 % des dinucléotides sont méthylés, et beaucoup de dinucléotides non méthylés sont situés dans des régions appelées îlots CpG qui regroupent un grand nombre de promoteurs fonctionnels (Razin and Riggs, 1980).

#### 5. Ubiquitination des histones et activation de la transcription

L'ubiquitination des histones peut également jouer un rôle dans la transcription. TAF1 (la sous-unité HAT de TFIID) est impliquée dans l'ubiquitination de l'histone H1. Or, certaines mutations de TAF1 pour l'activité d'ubiquitination chez la drosophile entraînent une perturbation de l'activation de la transcription de certains gènes (Pham and Sauer, 2000). Les formes de H2A et H2B ubiquitinées ont également été retrouvées dans les loci actifs en transcription (Jason et al., 2002). Deux processus contraires **d'ubiquitination et de dé-ubiquitination** seraient impliqués dans l'activation de la transcription, et la sous-unité Ubp8 du complexe SAGA serait douée d'une activité de dé-ubiquitination de l'histone H2B (Wyce et al., 2004). Le fait que l'ubiquitination et la dé-ubiquitination soient tous deux impliqués dans l'activation de la transcription paraît assez surprenant au vu des autres modifications post-traductionnelles dont les effets s'opposent la plupart du temps (acétylation/désacétylation). Quoiqu'il en soit, l'ubiquitination des histones participe à la régulation de la transcription.

#### 6. Les complexes

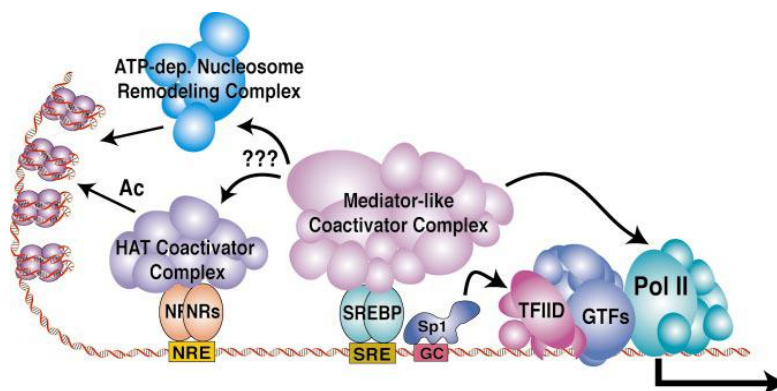
Les complexes GNAT ou SAGA-like : SAGA, SLIK, PCAF, STAGA, TFTCGcn5, Pcaf, p300, TAF250 ont été très largement étudiés pour leur capacité à acétyler autant les histones que d'autres substrats. Ils ne fonctionnent correctement qu'entourés d'autres facteurs,

c'est-à-dire en **complexes multi-protéiques**. Ces partenaires primordiaux confèrent une plus grande stabilité, une meilleure spécificité à l'enzyme, ou la dirigent vers certains sites du génome. Ainsi, scGcn5 est inactive sur la chromatine si elle est isolée (Kuo et al., 1996). Néanmoins, la présence d'ADA2 et ADA3 lui permet d'acétyler des histones organisées en nucléosomes. Ces trois facteurs participent aux complexes de levure SAGA et SLIK, mais forment aussi une entité indépendante qui peut être purifiée (Balasubramanian et al., 2002).

Parmi les complexes capables de modifier les queues amino-terminales des histones, et de réguler ainsi l'état plus ou moins actif de la chromatine, ce groupe de complexes se distingue par le fait qu'ils contiennent des TAFs. Ceux-ci forment, avec quelques autres facteurs, un noyau commun auquel s'ajoute dans chacun des complexes des sous-unités spécifiques. Leur enzyme HAT est Gcn5 ou Pcaf (Tableau 3).

### Résumé du chapitre 1 :

L'initiation de la transcription chez les eucaryotes nécessite l'assemblage d'un complexe de pré-initiation sur les promoteurs des gènes à transcrire (Fig.44). Celui-ci est soumis à régulation via un vaste ensemble d'activateurs et de répresseurs transcriptionnels spécifiques mais aussi de cofacteurs facilitant la transmission de signaux, la décompaction de la chromatine et le recrutement de la machinerie transcriptionnelle au niveau de promoteur de gènes spécifiques. Les interactions entre ces différents acteurs de la transcription sont résumées sur le schéma suivant :



**Figure 44:** Les différents types d'activités coactivatrices régulant la transcription des gènes de classe II (Naar *et al.*, 2001).

TFIID	sc SAGA	sc SLIK	dm SAGA	hs TFTC	hs PCAF	hs STAGA	Commentaires
TBP	/	/	/	/	/	/	
TAF1	/	/	/	/	/	/	
TAF2	/	?	dTAF2 ?	hTAF2	?	?	
TAF3	?	?	?	hTAF3	?	?	
TAF4	/	?	dTAF4 ?	hTAF4	?	/	HFD (TAF12)
TAF5	yTAF5	yTAF5	dTAF5 ?	hTAF5	?	/	
TAF6	yTAF6	yTAF6	dTAF6 ?	hTAF6	?	?	HFD (TAF9)
TAF7	?	?	?	hTAF7	?	/	
TAF8	?	?	?	hTAF8L	?	?	HFD (TAF10 ou Spt7)
TAF9	yTAF9	yTAF9	dTAF9 ?	hTAF9	hTaf9	hTaf9	HFD (TAF6 ou TAF6L)
TAF10	yTAF10	yTAF10	dTAF10b dTAF10a	hTAF10	hTaf10	hTaf10	HFD (TAF ou SPT7)
TAF11	/	/	/	/	/	/	
TAF12	yTAF12	yTAF12	dTAF12	hTAF12 hTAF6L hTAF5L	hTaf12 hTaf6L hTAF5L	hTaf12 hTaf6L hTAF5L	HFD (TAF4 ou ADA1)
TAF13	/	/	/	/	/	/	
TAF14*	/	/	/	/	/	/	
TAF15**	/	/	/	/	/	/	
	Tra1	Tra1	dtra1	hTRRAP	hTRRAP	hTRRAP	cible des coactivateurs
	yGCN5	yGCN5	dGCN5	hGCN5L	hPCAF	hGCN5	HAT
	yADA1	yADA1	?	hADA1	?	hADA1	HFD(TAF12)
	yADA2	yADA2	(dADA2a) ?	?	hADA2a	hADA2a	doigts de zinc et SANT : régulation de l'HAT
	/	/	dADA2b	(hAdA2b) /	(hADA2b)/	(hADA2b)/	doigts de zinc et SANT : régulation de l'HAT
	yADA3	yADA3	dADA3	hADA3	hADA3	hADA3	interaction avec TBP
	ySPT3	ySPT3	?	hSPT3	hSpt3	hSpt3	HFD
	?	?	?	hSAP130		hSAP130	
	ySPT7	ySPT7 tronqué	?	hSPT7L (STAF65g)		hSPT7L (STAF65g)	HFD (TAF10), bromodomain. Interaction génétique avec TBP
	ySPT8	/	?	?	?	?	
	ySPT20	ySPT20	?	?	?	?	interaction avec TBP
	Sgf73	?	?	Ataxin7	?	?	extention polyQ dans des ataxies spinocerebeuses
	Sgf29	Sgf29	?	?	?	?	
	Ubp8	?	?	?	?	?	Ubiquitine ligase
	Sgf11	?	?	?	?	?	liaison de Ubp8
	?	Rtg2	?	?	?	?	
	Chd1	Chd1	?	?	?	?	
	Sus1		e(y)2 ?	e(y)2 ?	e(y)2 ?	e(y)2 ? STAF36 STAF46	export des ARNm

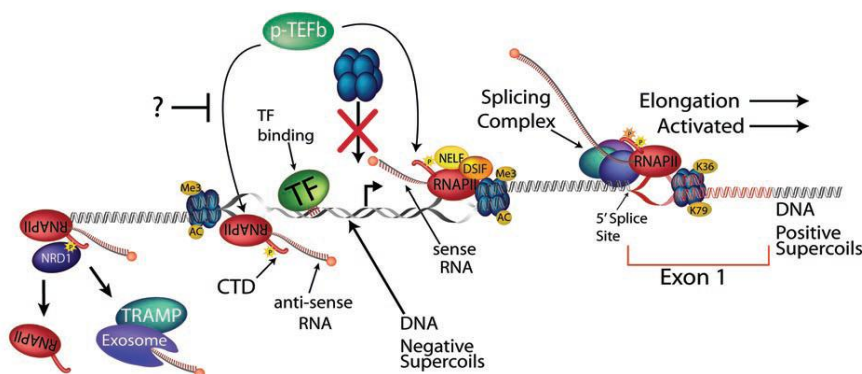
Tableau 3 : Composition et caractéristiques des complexes contenant des TAFs.



## Conclusion du chapitre 1 :

L'initiation de la transcription des gènes eucaryotes est un processus complexe nécessitant un nombre d'étapes surprenant et mettant en jeu des acteurs très variés : des GTFs à la Pol II, en passant par les co-activateurs, les co-répresseurs et les modifications covalentes des histones et des facteurs de transcription, le tout afin d'assurer une modulation fine et précise de cette étape si cruciale pour la vie des organismes.

Des découvertes très récentes y ajoutent pourtant un niveau de complexité supplémentaire. La transcription sur un même promoteur pourrait s'effectuer dans deux directions opposées, grâce à deux polymérases, après un arrêt de celles-ci sur les sites +1 et -1 des nucléosomes.



**Figure 45 : Aperçu de la transcription divergente (d'après Seila AC et al., 2009)**

Cette double transcription simultanée ou transcription divergente favoriserait la formation de supertours négatifs de l'hélice d'ADN, améliorant ainsi l'accès des GTFs aux régions promotrices et induisant très rapidement l'étape d'élongation des transcrits sens et anti-sens.

Ce modèle remet en question la vision classique de l'unité transcriptionnelle de base. Pourtant, s'il dégage un nouveau coin d'ombre pour l'origine des ARNm anti-sens, il ne résout toujours pas les questions qui entouraient le modèle le précédant. De nombreuses interrogations mécanistiques demeurent encore, notamment sur le rôle des différentes sous-unités protéiques de tous les complexes dans la régulation de la transcription ainsi que sur les différences de modulation de la transcription au sein des différents types cellulaires.

Dans le second chapitre, nous aborderons de manière plus détaillée le complexe TFIID et nous nous focaliserons sur le rôle de sa sous-unité TAF4.





# *CHAPITRE 2*

*ZOOM SUR UN COFACTEUR*

*MULTIPARTITE PARTICULIER:*

*TFIID*



## CHAPITRE 2

### Zoom sur un cofacteur multipartite particulier : TFIID

Requis lors de l'initiation de la transcription de l'ADN eucaryote par l'ARN Pol II, le facteur général de transcription TFIID est un complexe multiprotéique. Il est composé de TBP et de 14 TAFs (TBP-associated factors) qui sont conservées de la levure à l'homme. Chez la levure *Saccharomyces cerevisiae*, des mutations thermosensibles ont révélé que 13 des 14 TAFs étaient essentiels. Chez les Eucaryotes supérieurs, l'inactivation d'allèles pour les composants de TFIID mène le plus souvent à la létalité dès l'embryogenèse précoce (Mohan *et al.*, 2003). Quelle est la structure de TFIID ? Le facteur TFIID est-il soumis à une dynamique d'assemblage et de fonctionnement ? Comment les TAFs, ses composants, interviennent-ils dans l'activation de la transcription ? Ce facteur, si important pour l'initiation de la transcription, laisse derrière lui quelques zones d'ombres en ce qui concerne l'activation de la transcription. Que sait-on aujourd'hui de TFIID ?

#### I. Vue d'ensemble du facteur TFIID

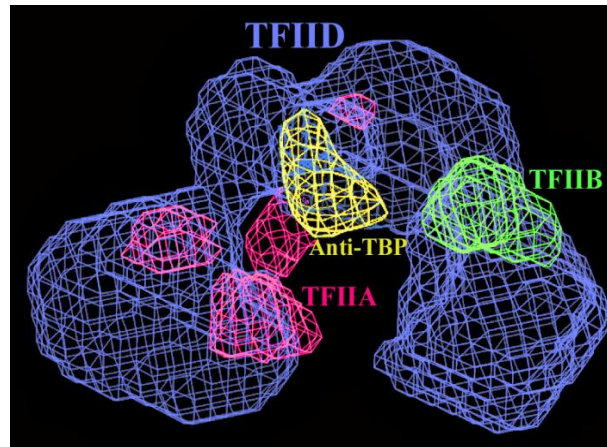
Malgré son rôle central dans l'initiation de la transcription et dans sa régulation, les connaissances structurales de TFIID ont été limitées jusqu'à présent aux études de microscopie électronique et aux colorations négatives des échantillons. De par sa taille importante et la difficulté de produire de grandes quantités de complexe actif, l'organisation générale du facteur TFIID demeure en effet évasive. Voyons les premières images de TFIID à faible résolution que ces techniques ont tout de même permis d'acquérir, ainsi que les solutions proposées pour améliorer le décryptage de ce facteur TFIID.

#### A. TFIID et sa structure modulaire

##### 1. Observation du facteur TFIID en microscopie électronique

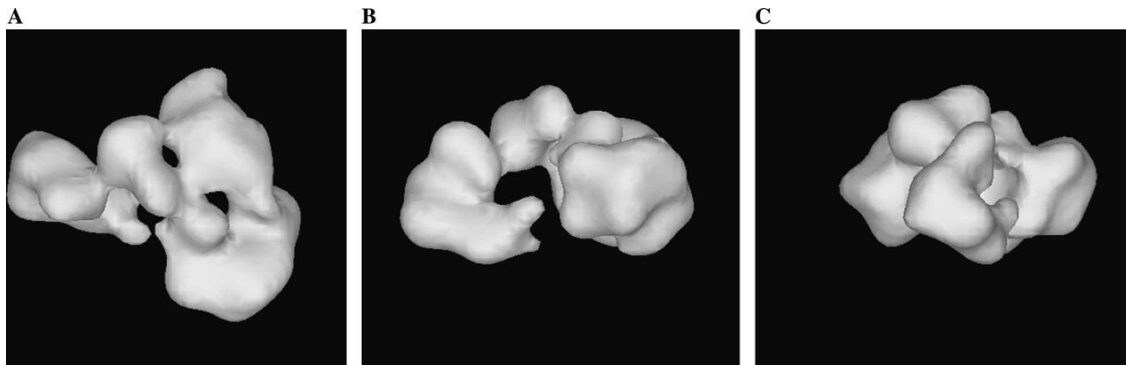
Deux études indépendantes d'observation de TFIID en microscopie électronique (Andel *et al.*, 1999; Brand *et al.*, 1999a) ont permis de constater que ce complexe multiprotéique était composé de 4 lobes (A, B, C1 et C2) séparés par des ponts très étroits,

donnant une impression globale de structure en fer à cheval autour d'un canal central (Brand *et al.*, 1999a) (Fig.46). L'analyse de la structure de TFIID associée à TFIIA ou à TFIIB (Andel *et al.*, 1999), ainsi que l'utilisation d'un anticorps anti-TBP ont permis de localiser TBP et de suggérer que ce canal central était l'interface de liaison avec l'ADN (Fig.44).



**Figure 46:** Le facteur TFIID associé à TFIIA, TBP et TFIIB (Andel *et al.*, 1999).

Initialement observée à faible résolution, l'image de TFIID basée sur la combinaison de la carte de densité 3D et des paramètres d'orientation des projections 2D a été améliorée grâce à un algorithme Quasi-Newton (Yang *et al.*, 2005) (Fig.47).



**Figure 47:** Modèles tridimensionnels du complexe hTFIID à 35 Å de résolution (Yang *et al.*, 2005).

Ces études constituent un pas important dans la compréhension de l'architecture globale de TFIID. Pourtant la qualité basale des résultats est souvent limitée par les effets de la coloration négative et de la déshydratation du complexe. Pour parer à ces biais, une autre technique permettant la caractérisation structurale du complexe TFIID humain (Fig.48) a été

réalisée (Grob *et al.*, 2006). Lors de cette étude, TFIID a été immunopurifié à partir de cellules Hela, hydraté, congelé puis traité par microscopie cryoélectronique.

Cette étude suggère que 2 conformations majeures existent pour le facteur TFIID. La première est dite « ouverte » ou « vide »; la seconde, plus stable, est qualifiée de « fermée » ou de « pleine » (Fig. 48). La seconde conformation pourrait avoir des contacts additionnels avec d'autres protéines la stabilisant, favorisant ainsi cette forme par rapport à la conformation ouverte.

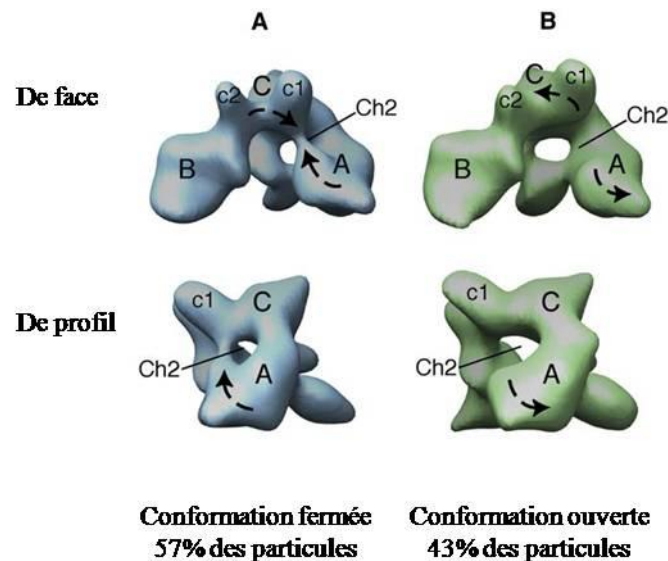


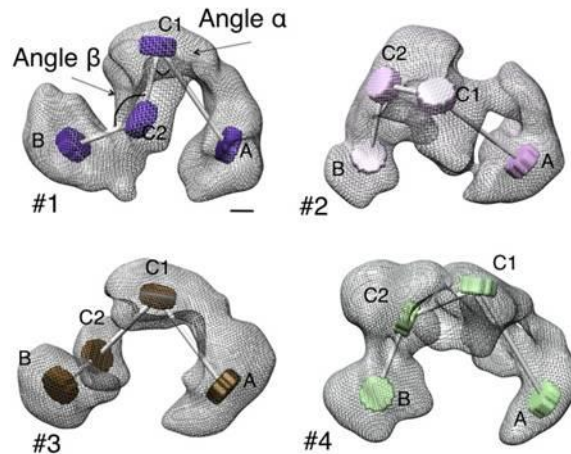
Figure 48: Conformations alternatives du facteur TFIID endogène humain (Grob *et al.*, 2006).

Des études complémentaires de covariance 3D (Penczek *et al.*, 2006) sur ces modèles ont permis d'identifier des mouvements coordonnés au sein du complexe TFIID qui font penser à des mouvements respiratoires. Ceux-ci reflèteraient le potentiel du complexe TFIID à se réorganiser constamment, afin de pouvoir lier différentes protéines de manière coordonnée (Grob *et al.*, 2006).

## 2. La plasticité de la structure de TFIID

Afin de mieux percevoir la structure du facteur TFIID malgré sa flexibilité apparente, une étude de tomographie électronique a été réalisée sur des modèles dépourvus ou non de la protéine TAF2. Des essais de purification en tandem de TFIID via l'immunoprécipitation des sous-unités TAF1 et TAF2 ont en effet mis en évidence le fait que la sous-unité TAF2 était souvent détachée du complexe TFIID (Papai *et al.*, 2009). Après une observation en coloration négative sous cryomicroscope électronique et une fixation des molécules selon la méthode GraFix (Kastner *et al.*, 2008), la comparaison des différents modèles obtenus pour

yTFIID avec ou sans TAF2 fait apparaître des variations dans les angles séparant les lobes A, C1 et C2 ainsi que dans la rotation des lobes autour de leurs connections (Papai *et al.*, 2009). Ceci suggère des changements de conformation du lobe B par rapport au lobe C, alors que le lobe A resterait en position fixe (Fig.49).



**Figure 49: Flexibilité de la structure de TFIID, représentation des 4 conformations (#1,#2,#3 et #4) majeures obtenues (Papai *et al.*, 2009).**

Les 4 états identifiés peuvent représenter des conformations stables du facteur TFIID ou des clichés instantanés d'une transition structurale continue. Si cela s'avère exact, un des rôles potentiels de TAF2 pourrait donc être de stabiliser TFIID dans l'une ou l'autre de ces conformations.

### 3. Bilan des changements conformationnels de TFIID

La présence de plusieurs états conformationnels de TFIID constitue une limitation majeure dans l'acquisition d'information structurale à haute résolution concernant ce facteur. La signification fonctionnelle de ce phénomène reste encore à élucider. Plusieurs hypothèses tendent tout de même à expliquer la présence de ces différentes formes et notamment :

- la structure de TFIID serait adaptable pour permettre la reconnaissance d'une large variété de promoteurs (Grob *et al.*, 2006; Papai *et al.*, 2009).

- des modifications post-traductionnelles des sous-unités de TFIID pourraient participer à la dynamique de régulation de la fixation de TFIID sur son promoteur (Boyer-Guittaut *et al.*, 2005; Kouskouti *et al.*, 2004).

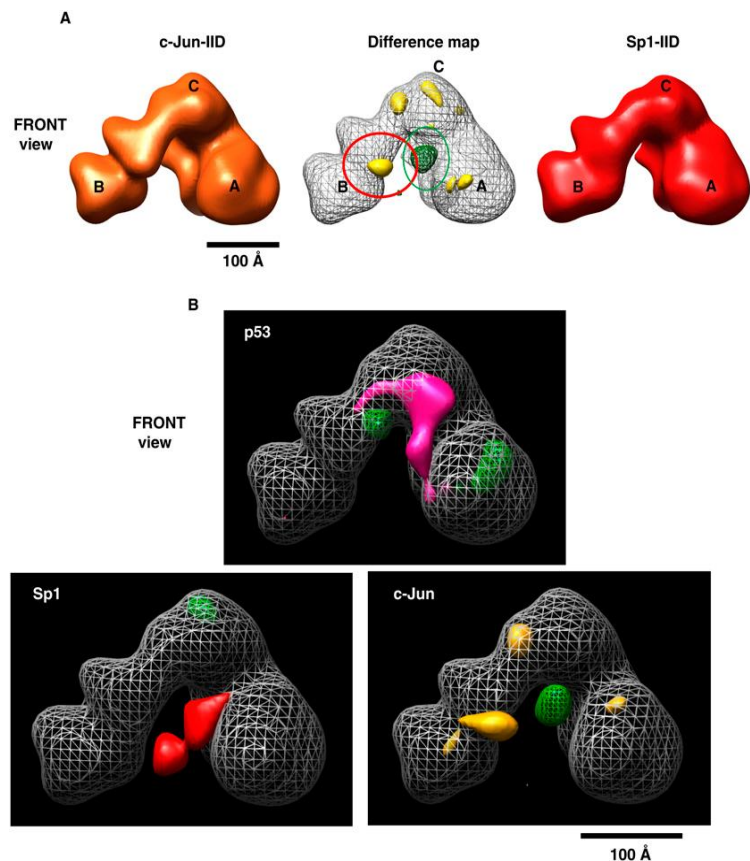
-la fixation d'activateurs ou d'inhibiteurs de la transcription aurait pu être à l'origine d'une réorganisation de ce complexe (Pugh and Tjian, 1990). Mais les études récentes de microscopie électronique caractérisant la structure de TFIID en présence de 3 activateurs montrent qu'il n'en est rien. Contrairement au complexe du médiateur qui subit des changements structuraux drastiques et globaux lors de la fixation d'un activateur, la structure globale de TFIID ne subit que quelques petites variations locales lorsque Sp1, p53 ou c-Jun s'y fixent (Liu *et al.*, 2009) (Fig.50).

A partir des structures 3-D obtenues pour c-Jun-IID et Sp1-IID, une image en superposition a été élaborée appelée carte en densité différentielle. Sur cette dernière, des différences positives (en jaune) et négatives (en vert) ont été mises en évidence entre les deux structures à partir du squelette de l'un des deux composés (en gris).

**Figure 50 : Reconstruction 3-D des interactions c-Jun-IID ; Sp1-IID et p53-IID (Liu *et al.*, 2009).**

(Les zones vertes sont les différences négatives entre les deux structures).

Chaque activateur semble contacter des TAFs spécifiques au sein de TFIID. TAF1 et TAF4 servent de co-activateurs pour Sp1, TAF1, TAF6 et TAF9 interviennent dans la transactivation dépendante de p53 et TAF1 et TAF7 seraient des co-activateurs pour c-Jun.



Quelque soit l'activateur fixé, il semble donc que l'architecture de TFIID ne varie guère. La position unique de chacun de ces trois activateurs sur le TFIID reflète la variété des points de contacts potentiels des activateurs avec le TFIID.

## B. La composition intrinsèque du facteur TFIID

### 1. TBP et les TAFs, des sous-unités de TFIID

La structure cristallographique de la protéine TBP associée à l'ADN a été déterminée (Kim and Burley, 1994) et présentée au cours du premier chapitre (Fig.15). Cette protéine possède une structure pseudosymétrique en forme de selle à cheval, avec une face concave qui interagit avec l'ADN, et une face convexe indispensable à l'interaction avec les TAFs. Nous allons donc nous intéresser plus particulièrement aux TAFs.

Outre TBP, TFIID contient 14 TAFs conservés d'une espèce à l'autre, de poids moléculaires allant de 18 à 250kDa.

Ces différents TAFs sont regroupés au sein du tableau 4. Leur appellation a évolué avec l'apparition en 2002 d'une nomenclature unifiée, permettant de mieux se repérer entre les différentes espèces (Tora, 2002).

New name	<i>H. sapiens</i> (hs)	<i>D. melanogaster</i> (dm)	<i>C. elegans</i> (ce)		<i>S. cerevisiae</i> (sc)	<i>S. pombe</i> (sp)
			previous name	new name		
TAF1	TAF <sub>II</sub> 250	TAF <sub>II</sub> 230	<i>taf-1</i> (W04A8.7)	<i>taf-1</i>	Taf145/130	TAF <sub>II</sub> 111
TAF2	TAF <sub>II</sub> 150	TAF <sub>II</sub> 150	<i>taf-2</i> (Y37F11B.4)	<i>taf-2</i>	Taf150 or TSM1	(T38673)
TAF3	TAF <sub>II</sub> 140	TAF <sub>II</sub> 155 or BIP2	(C11G6.1)	<i>taf-3</i>	Taf47	
TAF4	TAF <sub>II</sub> 130/135	TAF <sub>II</sub> 110	<i>taf-5</i> (R119.6)	<i>taf-4</i>	Taf48 or MPT1	(T50183)
TAF4b	TAF <sub>II</sub> 105					
TAF5	TAF <sub>II</sub> 100	TAF <sub>II</sub> 80	<i>taf-4</i> (F30F8.8)	<i>taf-5</i>	Taf90	TAF <sub>II</sub> 72
TAF5b						TAF <sub>II</sub> 73
TAF5L	PAF65β	Cannonball				
TAF6	TAF <sub>II</sub> 80	TAF <sub>II</sub> 60	<i>taf-3.1</i> (W09B6.2)	<i>taf-6.1</i>	Taf60	(CAA20756)
TAF6L	PAF65α	(AAF52013)	<i>taf-3.2</i> (Y37E11AL.8)	<i>taf-6.2</i>		
TAF7	TAF <sub>II</sub> 55	(AAF54162)	<i>taf-8.1</i> (F54F7.1)	<i>taf-7.1</i>	Taf67	TAF <sub>II</sub> 62/PTR6
TAF7L	TAF2Q		<i>taf-8.2</i> (Y111B2A.16)	<i>taf-7.2</i>		
TAF8	(BAB71460)	Prodos	(ZK1320.12)	<i>taf-8</i>	Taf65	(T40895)
TAF9	TAF <sub>II</sub> 32/31	TAF <sub>II</sub> 40	<i>taf-10</i> (T12D8.7)	<i>taf-9</i>	Taf17	(S62536)
TAF9L	TAF <sub>II</sub> 31L (AAG09711)					
TAF10	TAF <sub>II</sub> 30	TAF <sub>II</sub> 24	<i>taf-11</i> (K03B4.3)	<i>taf-10</i>	Taf25	(T39928)
TAF10b		TAF <sub>II</sub> 16				
TAF11	TAF <sub>II</sub> 28	TAF <sub>II</sub> 30β	<i>taf-7.1</i> (F48D6.1)	<i>taf-11.1</i>	Taf40	(CAA93543)
TAF11L			<i>taf-7.2</i> (K10D3.3)	<i>taf-11.2</i>		
TAF12	TAF <sub>II</sub> 20/15	TAF <sub>II</sub> 30α	<i>taf-9</i> (Y56A4.3)	<i>taf-12</i>	Taf61/68	(T37702)
TAF13	TAF <sub>II</sub> 18	(AAF53875)	<i>taf-6</i> (C14A4.10)	<i>taf-13</i>	Taf19 or FUN81	(CAA19300)
TAF14					Taf30	
TAF15	TAF <sub>II</sub> 68					
<i>B-TFIID</i>						
BTAF1	TAF <sub>II</sub> 170/TAF-172	Hel89B	(F15D4.1)	<i>btaf-1</i>	Mot1	(T40642)

**Tableau 4: Les TAFs composant le complexe multiprotéique TFIID.**



## 2. Les TAFs et le domaine de repliement de type histone

Sur les 14 TAFs, 9 possèdent un domaine de repliement de type histone appelé HFD (Histone Fold Domain). Il s'agit de TAF3, TAF4, TAF6, TAF8, TAF9, TAF10, TAF11, TAF12 et TAF13. Ce HFD est un domaine d'interaction protéine-protéine très bien caractérisé comprenant 3 hélices  $\alpha$  reliées par 2 boucles. Ce motif minimal aurait surgi du duplicata du motif hélice-boucle-hélice qui aurait fusionné pour donner l'hélice centrale  $\alpha 2$ . Il est hautement conservé. Le motif HFD n'est pas seulement trouvé dans les histones mais également au sein de nombreux facteurs de transcription (Burley *et al.*, 1997; Sullivan *et al.*, 2000).

### a) La formation de dimères de TAFs

Des analyses de cristallographie par rayons X et des expériences biochimiques ont mis en évidence le rôle important de ce motif de repliement de type histone, qui, lié à d'autres HFD ou à d'autres domaines protéiques, permet une hétérodimérisation des protéines.

La présence de HFD au sein des sous-unités de TAFs permet à celles-ci de s'associer tête-bêche deux par deux (Fig.51). Chronologiquement, les couples de TAFs découverts qui interagissent ensemble sont les suivants :

-TAF6 et TAF9, dont des homologies respectives avec les histones H4 et H3 ont été constatées très tôt (Xie *et al.*, 1996).

-TAF11 et TAF13 (Birck *et al.*, 1998), qui ne sont pas apparentés aux histones du nucléosome.

-TAF4 et TAF12 (Gangloff *et al.*, 2000) qui sont respectivement homologues à H2A et à H2B.

-TAF8-TAF10 et TAF3-TAF10 (Gangloff *et al.*, 2001b).



**Figure 51: Dimère H2A-H2B.**

La structure du dimère TAF4-TAF12 a été isolée par l'équipe de D. Moras (Werten *et al.*, 2002). hTAF12 forme bien un hétérodimère avec hTAF4 comme le préconisaient les études de double-hybride (Gangloff *et al.*, 2000) et les analyses de co-expression bactérienne (Fribourg *et al.*, 2001)(Fig.52).

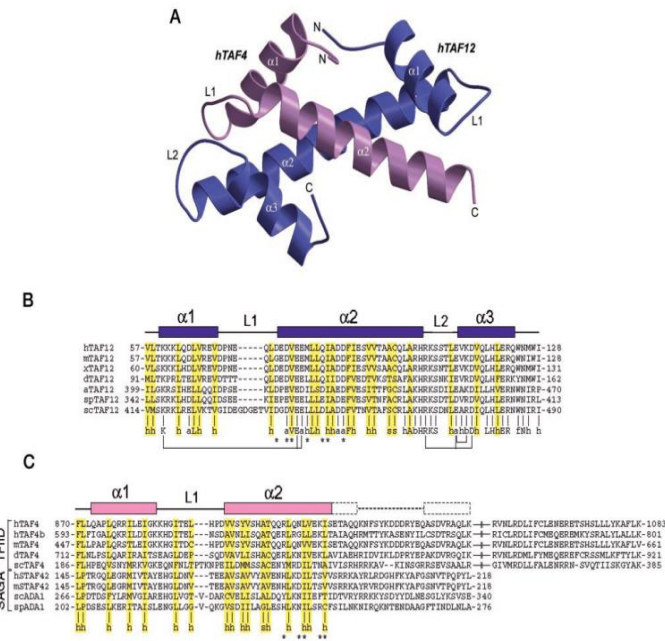


Figure 52 : Hétérodimère hTAF4-hTAF12 (Werten *et al.*, 2002).

La superposition des structures hélicoïdales de hTAF4 et hTAF12 avec celles des histones H2A et H2B révèle que l'hélice  $\alpha 3$  est manquante chez hTAF4 (Fig.51).

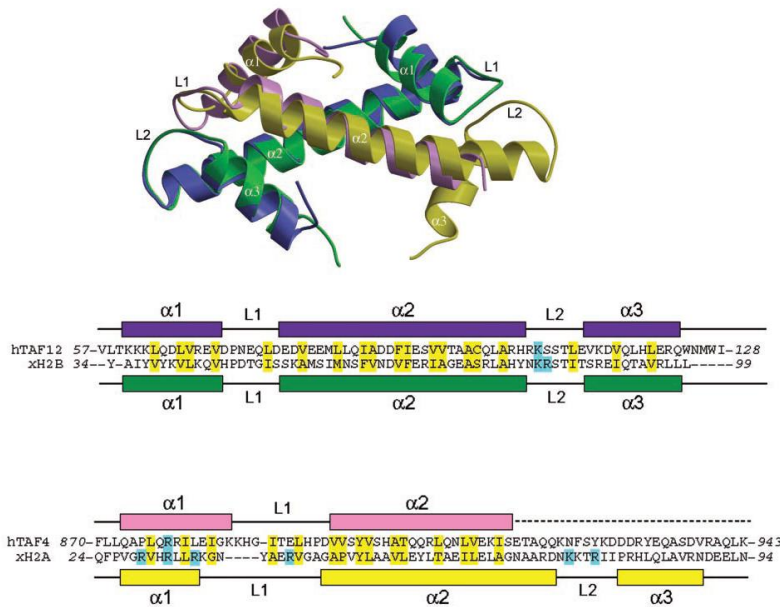
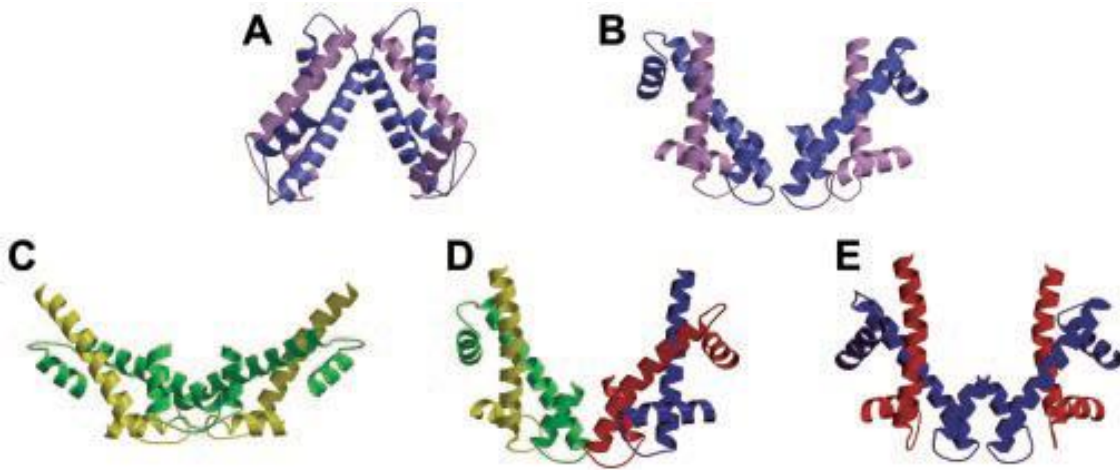


Figure 53 : Comparaison de hTAF4 et hTAF12 aux histones H2A and H2B.

Ces hétérodimères peuvent s'associer entre eux et former des dimères de dimères, ou tétramères. Le fait que le même hétérotétramère soit trouvé dans deux cristaux différents indique une interaction serrée qui pourrait en principe correspondre à une sous-structure appropriée de TFIID (Fig.53).

Cependant, les hétérotétramères formés par hTAF4 et hTAF12 diffèrent sensiblement des hétérotétramères trouvés au sein du nucléosome (Fig.54). Il y aurait une asymétrie de la structure de l'hétérotétramère formé par hTAF4 et hTAF12. De plus, l'interface entre les deux hétérodimères hTAF4 et hTAF12 consiste en des résidus hydrophobes, qui, dans une structure canonique d'hétérodimère d'histones, ne serait pas exposée en surface mais serait recouverte par la troisième hélice de hTAF4.



**Figure 54 :** Comparaison de la formation en tétramère de hTAF4-hTAF12 avec les contacts entre les hétérodimères observés dans le nucléosome (Werten *et al.*, 2002). Couleurs: rose : hTAF4; bleu : hTAF12; jaune: histone H2A ; vert: histone H2B; bleu foncé : histone H3; rouge : histone H4.

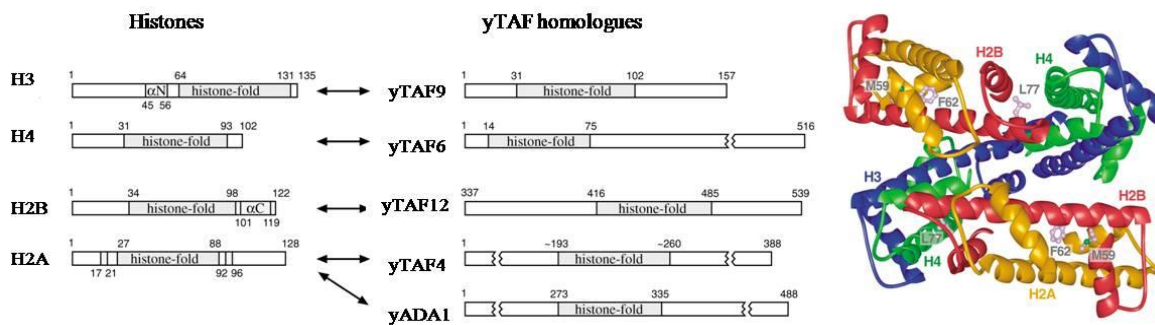
A. Tétramère formé par hTAF4 et hTAF12 au sein des cristaux. B. Autre tétramère formé entre hTAF4 et hTAF12. C. Contacts entre les hétérodimères H2A-H2B dans le nucléosome (L'hélice 3 de H2A a été enlevée du modèle). D : Contacts entre les hétérodimères H2A-H2B et H3-H4 dans le nucléosome. (L'hélice 3 de H2A a été enlevée du modèle). E. Contacts entre les hétérodimères H3-H4 dans le nucléosome (L'hélice 3 de H4 a été enlevée du modèle).

### 6) La controverse autour du pseudo-nucléosome

Les protéines TAFs possèdent des motifs de repliement de type histone qui adoptent une structure semblable à celle des histones. Sur la base biochimique de ces interactions entre TAFs, l'équipe de M.Hoffmann a proposé une sous-structure de TFIID de type nucléosomale, qui consisterait en un tétramère de TAF6 et TAF9 plus deux homodimères de TAF12 (Hoffmann and Roeder, 1996). Pour eux, lorsque ces protéines s'associent, elles miment la structure du nucléosome et le remplacent sans perturber la périodicité de l'agencement

histone-ADN (Hoffmann and Roeder, 1996). La structure formée serait alors qualifiée de pseudo-nucléosome.

Ce modèle a donné lieu à de nombreuses investigations biochimiques et génétiques concernant la fonction de TFIID. A présent, il est vraisemblable que si une telle structure nucléosomale devait exister, elle contiendrait plutôt deux hétérodimères de TAF4-TAF12 (Werten *et al.*, 2002). Conformément à cette idée, un complexe octamérique *in vitro* a été reconstitué par l'équipe de M. Selleck. Il est composé de yTAF6, yTAF9, yTAF4 et yTAF12 (Selleck *et al.*, 2001) (Fig.55).



**Figure 55:** Les domaines de repliement des histones et leurs homologues dans les TAFs (Selleck et al., 2001).

D'après M. Selleck, des mutations ponctuelles à l'intérieur du HFD détruisent la structure de l'octamère *in vitro* et ces mutations sensibles à la chaleur peuvent être spécifiquement supprimées en surexprimant d'autres composants d'octamères *in vivo*. Ces résultats indiquent que l'octamère de TAFs est similaire en stœchiométrie et en interactions de HFD à l'octamère d'histones composant la chromatine.

Ces résultats sont tout de même à prendre avec précaution, car les hétérodimères TAF4-TAF12 et TAF6-TAF9 existent aussi de manière indépendante ou groupés avec d'autres hétérodimères. Ainsi la structure octamérique, si elle existe, cohabite avec d'autres structures plus simplifiées.

### c) *HFD, un élément crucial dans l'organisation de TFIID*

Les 5 hétérodimères TAF4-TAF12 ; TAF6-TAF9 ; TAF8-TAF10 ; TAF3-TAF10 et TAF11-TAF13 représentent un motif structural important puisqu'on retrouve ces paires de TAFs au sein du TFIID natif. Ces motifs HFD pourraient contribuer à la liaison de TFIID au

promoteur. En effet, l'interaction protéine-protéine facilitée par les HFD peut augmenter et stabiliser l'interaction des TAFs avec l'ADN. C'est le cas de TAF9 et TAF6, qui par leurs domaines de liaison similaires à celui des histones H3 et H4 respectivement (Xie *et al.*, 1996), augmentent nettement leur activité de liaison à l'ADN. En effet, il a été montré qu'une structure en pseudo-nucléosome composée de TAF6, TAF12, TAF9, TAF4, correspondant aux histones respectifs H4, H2B, H3 et H2A (Selleck *et al.*, 2001), a plus d'affinité avec l'ADN que chacune des TAFs isolées (Shao *et al.*, 2005).

### C. L'assemblage de TFIID

#### 1. La masse moléculaire du complexe natif

La détermination de la masse moléculaire du complexe TFIID purifié a été réalisée par filtration sur gel et par gradient de sucrose linéaire (10 à 30%) suivi d'une analyse des zones de sédimentation (Sanders *et al.*, 2002). Conformément à la masse moléculaire calculée pour le monomère TFIID en fonction des données stœchiométriques, l'ensemble des TAFs migrent de concert en un complexe de 1,2 kDa lors de la filtration sur gel (Fig.56).

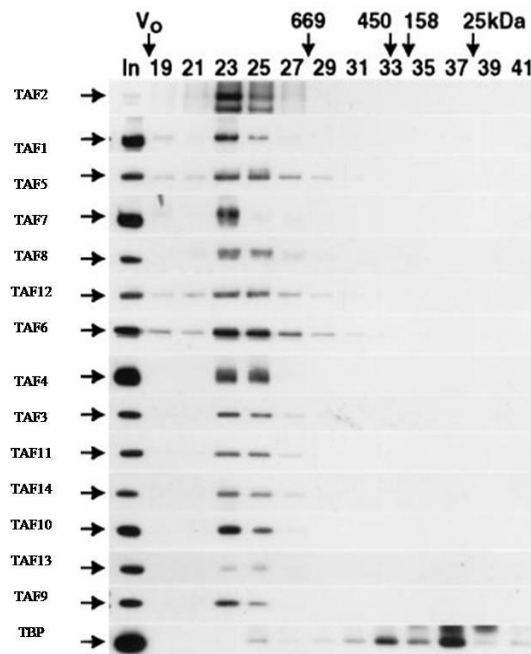


Figure 56 : Détermination de la masse moléculaire de  $\gamma$ TFIID (Sanders *et al.*, 2002).

Seul TBP affiche un profil d'éluion différent avec un pic à 30kDa (Fig.56). Ce résultat signifie que soit TBP n'est associé que de manière lâche aux autres sous-unités de TFIID, soit il a été dissocié lors de la chromatographie. Mais la première hypothèse paraît peu probable car TBP s'associe à TAF1 avec une forte affinité (Bai *et al.*, 1997). La deuxième hypothèse paraît plus séduisante : elle suggère que TBP peut s'associer ou se dissocier de TFIID et qu'il aurait été séparé physiquement via la chromatographie (intégrant l'intérieur des billes) sans qu'il puisse par la suite se réassocier au reste du complexe TFIID. Au contraire, le passage de TFIID dans des colonnes échangeuses d'ions préserverait TBP de cette séparation physique. Le bénéfice majeur retiré de cette association ou dissociation rapide de TBP vis-à-vis du complexe TFIID est le fait que TBP puisse se répartir sur différents complexes transcriptionnels, passant rapidement de TFIID à TFIIB par exemple en fonction des conditions environnementales et de croissance cellulaire.

Les expériences de gradient de sucrose ont apporté des résultats contradictoires, attribuant à TFIID la masse de 700kDa (Sanders *et al.*, 2002). Bien que l'explication donnée repose sur la forme ellipsoïde de  $\gamma$ TFIID favorisant une sédimentation plus lente que les molécules globulaires, il se peut que la masse moléculaire obtenue ne reflète pas celle de l'holo-TFIID mais d'un complexe primitif constitué de quelques TAFs essentiels qui s'associeraient par la suite à d'autres TAFs pour former l'holo-TFIID.

Une étude plus approfondie de la dynamique d'assemblage du facteur TFIID apparaît donc fondamentale et essentielle à la compréhension de la régulation de la transcription des différents gènes.

## 2. La stœchiométrie des sous-unités de TFIID

Afin de positionner correctement les couples de TAFs au sein du TFIID et de déterminer si les sous-unités de TFIID sont présentes en une seule ou plusieurs copies au sein de ce facteur, des immunoprécipitations de  $\gamma$ TAF1, couplé à l'étiquette HA, chez la Levure *Saccharomyces cerevisiae* ont été réalisées par l'équipe d'Anthony Weil. Les éluions obtenues ont été résolues sur des gels SDS PAGE colorés au bleu de Coomassie ou au SYPRO Ruby. Ces gels ainsi colorés ont ensuite été soumis à un scanner densitomètre et la quantité relative de chacune des sous-unités de TFIID a été mesurée par rapport à celle de  $\gamma$ TAF1 (Sanders *et al.*, 2002) (Tableau 5).

Les études de génétique et de biochimie ont permis de définir 5 paires de TAFs à l'intérieur de TFIID (Gangloff *et al.*, 2001a). Conformément à ces études, les partenaires présents dans ces paires apparaissent à des quantités approximativement équimolaires. Sauf pour TAF4 qui migrerait en triplet de bandes à cause de multiples phosphorylations auquel il est soumis *in vivo* (Sanders and Weil, 2000) et pour TAF10 qui a plusieurs partenaires d'hétérodimérisation (Gangloff *et al.*, 2001b).

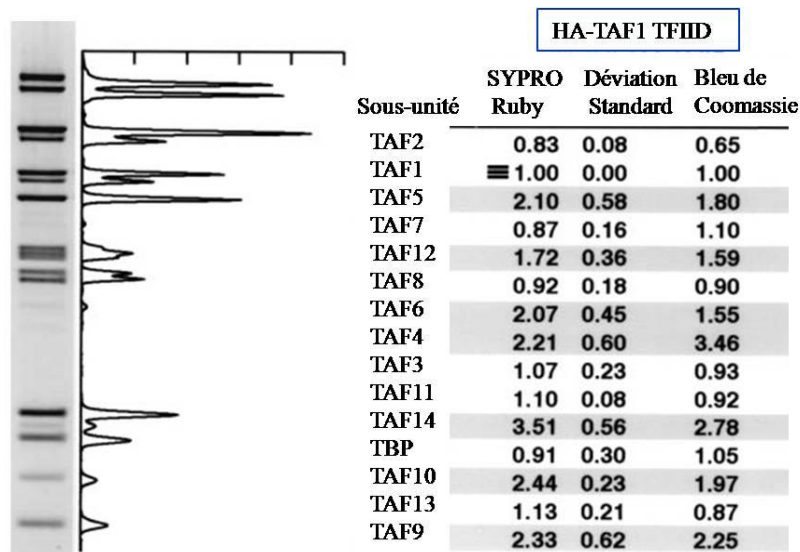


Tableau 5: Stœchiométrie des sous-unités de TFIID (Sanders *et al.*, 2002).

Les résultats obtenus les plus surprenants portent sur le différentiel de valeur stœchiométrique apparente entre les sous-unités spécifiques à  $\gamma$ TFIID (TAF2, TAF1, TAF7, TAF8, TAF3, TAF11 et TBP) dont la quantité relative est de 1 mole par mole de TFIID, et les sous-unités communes à TFIID et SAGA (TAF9, TAF10, TAF6, TAF12, TAF5) dont la quantité relative est égale à 2 moles par mole de TFIID (Tableau 5). TAF4 fait figure d'exception avec sa valeur relative égale à un peu plus de 2 moles par mole de TFIID, alors qu'il n'apparaît pas dans le complexe SAGA.

Afin de déterminer plus précisément le nombre de copies de chaque TAF au sein du TFIID, des expériences de co-immunoprécipitations ont été réalisées sur une série de TAFs contenus dans des lignées diploïdes de Levure (Sanders *et al.*, 2002) et portant une étiquette HA. Alors que les protéines, étiquetées avec HA, dont la valeur stœchiométrique est estimée à 2 ou 3 moles par mole de TFIID co-immunoprécipitent avec leur forme apparentée non

étiquetée aussi bien qu'avec d'autres substrats, les autres TAFs, dont la valeur stœchiométrique avoisine les 1 mole par mole de TFIID, ne réalisent aucune association homodimérique (Tableau 6).

**Tableau 6 : Les sous-unités de TFIID et leur association homodimérique.**

Sous-unité	Association homodimérique
TAF2	Non
TAF1	Non
TAF5	Oui
TAF7	Non
TAF12	Oui
TAF8	Non
TAF6	Oui
TAF4	Oui
TAF3	Non
TAF11	Non
TAF14	Oui
TBP	Non
TAF10	Oui
TAF13	Non
TAF9	Oui

Au final, toutes ces expériences génétiques, biochimiques et de microscopie électronique couplées à de l'immunohistologie ont permis d'identifier clairement deux sous-populations de TAFs au sein du TFIID, sur la base de leur stœchiométrie.

Alors que TAF1, TAF2, TAF7 et TBP sont présents en une copie unique au sein du TFIID, la plupart des autres TAFs sont recensés en doubles exemplaires. Ces résultats ont été confirmés par immuno-électromicroscopie, mettant en évidence la présence de deux copies au minimum pour TAF5, TAF10, TAF6/TAF9, TAF4/TAF12 (Leurent *et al.*, 2002). L'immunolocalisation a permis de définir plus précisément la composition de chaque lobe.

L'organisation finale de TFIID repose essentiellement sur TAF5 qui joue le rôle d'un échafaudage joignant les différents lobes et TAF1 qui scelle l'holo-TFIID final (Leurent *et al.*, 2002). Les TAFs présents en une simple copie se situent surtout au niveau des lobes A et C (Fig.57). On trouve par exemple la partie C-terminale de TAF1 et TAF7 au sommet du lobe A, alors que la partie N-terminale de TAF1 et la protéine TBP se trouvent plutôt à l'interface entre les lobes A et C, de même que TAF2.

L'organisation générale du TFIID ne semble donc pas se limiter pas à un simple octamère d'histones dont la fonction serait de récapituler simplement celle du nucléosome lorsqu'il est fixé sur l'ADN (Hoffmann *et al.*, 1997).



La distribution de tous ces hétérodimères au sein de TFIID est en fait bien plus complexe qu'il n'y paraît, car la structure cristalline des TAFs n'est connue que pour quelques domaines (Romier *et al.*, 2007; Werten *et al.*, 2002).

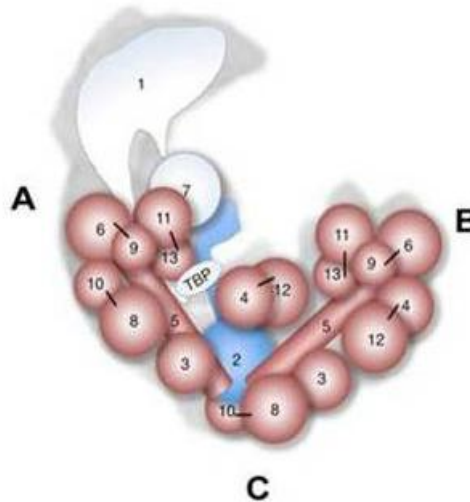


Figure 57: Organisation structurale de yTFIID à 23 Å de résolution (Cler *et al.*, 2009).

### 3. Dynamique d'assemblage des TAFs au sein du TFIID

#### a) Un assemblage séquentiel

Au vu des connaissances actuelles, l'assemblage du facteur TFIID se fait en plusieurs étapes, détaillées dans la revue n°1, et résumées ici.

Des études biochimiques ont révélé des associations entre certaines TAFs. TAF5 forme un sous-complexe stable avec TAF6 et TAF9 (Tao *et al.*, 1997) et se lie aussi fortement à TAF1, TAF7 et TAF11 (Dubrovskaya *et al.*, 1996). Comme TAF5 se lie à de nombreuses TAFs, il a été proposé que cette sous-unité soit à l'origine de la formation d'un complexe de TAFs. En outre, TAF3 ainsi que TAF8 et SPT7 hétérodimérisent avec TAF10 (Gangloff *et al.*, 2001). Ainsi, il est très probable que les TAFs s'agencent en sous-complexes avant d'être incorporées aux complexes TFIID ou TFTC.

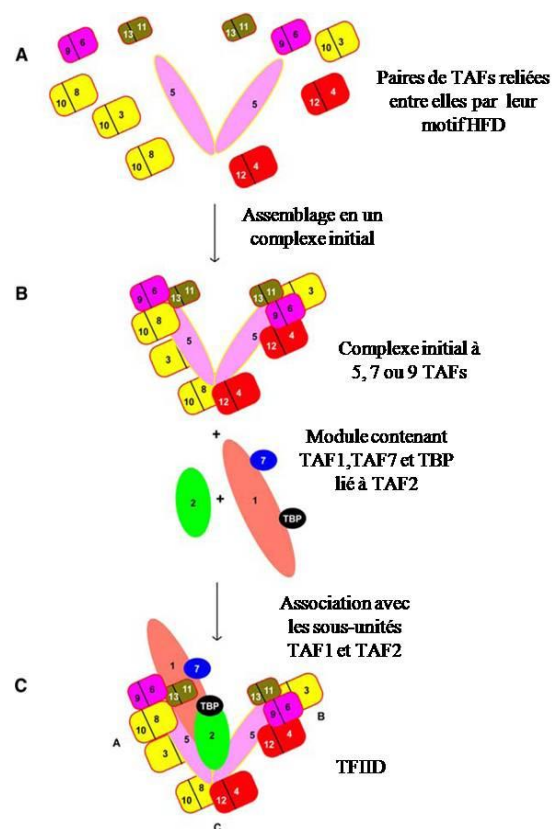
Les TAFs contenant des HFD forment des hétérodimères (Fig.57 et 58). Ils s'associeraient donc avec TAF5 pour former un complexe primordial (voir revue n°1). La

capacité de TAF5 à s'homodimériser fait l'objet de plusieurs recherches structurales encourageantes (Bhattacharya *et al.*, 2007; Romier *et al.*, 2007).

TAF5 et les paires d'hétérodimères finiraient par s'assembler, formant ainsi un petit complexe de transcription (Fig.58).

Ce petit complexe initial est celui que l'on retrouve dans les cellules S2 de *Drosophile* (Wright *et al.*, 2006) mais aussi dans les cellules fibroblastiques de souris (voir manuscrit n°1). D'autres complexes à 7 TAFs (Leurent *et al.*, 2004) et à 9 TAFs ont été reconstitués respectivement dans des baculovirus *in vitro* et dans la Levure *in vivo*.

Ce complexe initial interagit alors à son tour avec le module TBP-TAF1-TAF7 pour former le facteur TFIID.

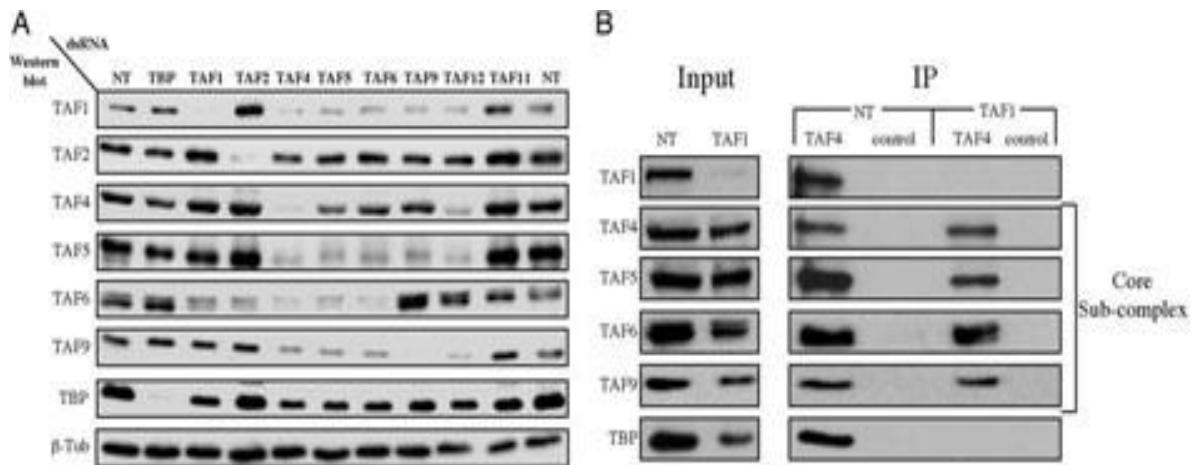


**Figure 58: Modèle schématique d'assemblage du TFIID**

TAF2 s'associerait seulement ensuite à TFIID, de manière indépendante, à travers une interaction avec TAF1, afin de stabiliser la conformation de TFIID (Papai *et al.*, 2009). Mais le rôle de TAF2 ne semble pas s'arrêter là. Son interaction avec l'élément initiateur (Kaufmann *et al.*, 1996; Verrijzer *et al.*, 1995) et sa liaison avec l'élément INR du promoteur tardif majeur de l'adénovirus dans le contexte de hTFIID plus particulièrement (Oelgeschlager *et al.*, 1996) laissent penser que TAF2 occupe une large part dans le choix du site d'initiation de certains promoteurs.

### 6) *Un sous-complexe stable de 5 TAFs ?*

L'utilisation d'ARN interférents pour inactiver TAF4 dans les cellules S2 (Schneider line) de *Drosophile* montre que la perte de TAF4 entraîne une diminution très forte de l'expression de toutes les autres TAFs testées, celle de TAF2 exceptée (Fig.59).



**Figure 59 : Analyse de la stabilité de TFIIID in vivo.**

A. Les cellules S2 ont été laissées sans traitement (NT) ou traitées avec des ARN interférents ciblant la sous-unité de TFIIID indiquée en haut du tableau de gauche pendant 3 jours. Des lysats de cellules entières ont été soumis à une analyse par Western Blot avec des anticorps dirigés contre les sous-unités situées à gauche du tableau. La tubuline sert de témoin de dépôt.

B. Un anticorps monoclonal contre TAF4 a été utilisé pour immunoprécipiter les complexes contenant TAF4 à partir des extraits nucléaires préparés à partir de cellules S2 soit non traitées (NT) soit traitées avec un ARN interférent contre TAF1. Les protéines ont été éluées et soumises à une analyse par Western Blot avec des anticorps dirigés contre les sous-unités situées à gauche du tableau. A gauche : le dépôt. A droite : l'éluat. Des protéines G couplées à des billes sépharose ont été utilisées comme contrôle non spécifique.

Au contraire, l'inactivation de TAF1 ne déstabilise pas les autres TAFs.

TAF4 apparaît donc comme une sous-unité clé, assurant la stabilité du complexe TFIIID. Ce sous-complexe serait décoré par la suite avec d'autres sous-unités périphériques et notamment avec le sous-complexe TAF1-TAF2-TBP (Fig.56) ou TAF7-TAF1-TBP (Singh et al., 2004). La déplétion TAF5, TAF6 et TAF9 déstabilisent aussi grandement le complexe TFIIID. Ainsi ces TAFs font vraisemblablement partie d'un sous-complexe stable de TFIIID. Après la formation de ce sous-complexe, l'holo-TFIIID serait généré d'après la suggestion des auteurs, par addition de TBP, TAF1, TAF2 et TAF11 comme sous-unités périphériques (Fig.60).

L'utilisation d'ARN interférents pour tester la stabilité du complexe TFIIID dans des cultures cellulaires de *Drosophila* révèle l'existence de composés critiques au sein de TFIIID et d'un sous-complexe à 5 TAFs, composé de TAF4/TAF12, TAF6/TAF9 et TAF5 (Wright et al., 2006). Mais cette étude reste limitée car il n'est pas exclu que ce complexe ne possède pas quelques sous-unités en plus. Dans ce cas, on pourrait se retrouver avec un complexe à 7

TAFs plutôt que 5, comme le préconisent les études de coinfections de baculovirus de Leurent (Leurent *et al.*, 2004).

On peut regretter dans cette expérience que de nombreuses sous-unités n'aient pas fait l'objet de tests avec ces ARN interférents. Qu'en est-il alors des variations d'expression de TAF3, TAF7, TAF8, TAF10, TAF11, TAF12 et TAF13 lorsque TAF4 est inactivée?

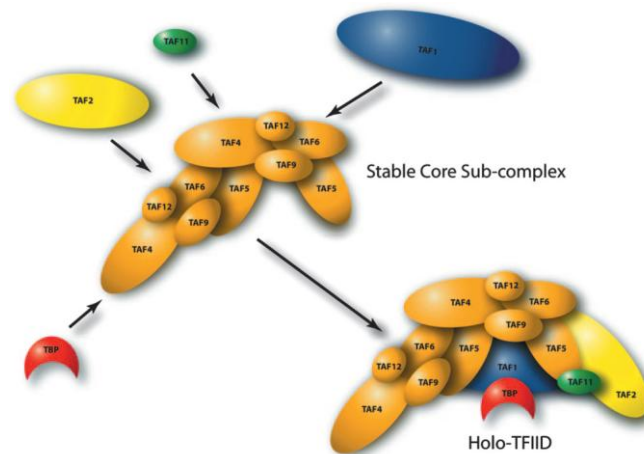


Figure 60: La formation de l'holo TFIID (Wright *et al.*, 2006).

#### D. Variabilité du complexe TFIID

Progressivement, de nouveaux TAFs ont été découverts et de nouveaux complexes transcriptionnels contenant ces TAFs ont été isolés. De ce fait, la transcription par la Pol II a pu être initiée au niveau de certains promoteurs par des PIC ne contenant qu'un ensemble partiel de TAFs, voir même de GTFs (Parvin *et al.*, 1994). Une certaine variabilité existe donc au sein des complexes de transcription.

Certains composants du TFIID ont des gènes paralogues. Après duplication d'un gène, une des deux formes a évolué. Ces gènes, bien que différents, appartiennent à la même famille et codent pour des protéines apparentées qui seront exprimées de manière spécifique dans certains tissus ou au cours du développement. C'est le cas de certains TAFs mais aussi de TBP. Voyons quels sont ces différents variants et quelle est leur implication dans la transcription dépendant de Pol II.

## 1. Les variants des TAFs

La diversité des complexes TFIID peut être à l'origine de la régulation très fine de l'expression des gènes. En effet, la composition des sous-unités du complexe TFIID peut varier en fonction du stade de différenciation de la cellule, ainsi que du tissu dans lequel elle se trouve. Certains variants présentent une grande similarité avec les TAFs et sont spécifiques de types cellulaires particuliers. Ils sont parfois capables de remplacer leur paralogue sous certaines conditions. En voici quelques exemples :

### a) TAF7L

TAF7L, de localisation cytoplasmique dans les spermatogonies et les stades précoces des spermatocytes, devient nucléaire dans les spermatocytes de stade pachytène et dans les spermatides haploïdes ronds et cet import dans le noyau coïncide avec la baisse de l'expression de TAF7 et l'augmentation de celle de TBP nucléaire (Pointud *et al.*, 2003). Il semble donc que TAF7L remplace TAF7 au sein de TFIID au cours de la spermatogenèse (Fig.61). De même, TAF1L agirait comme un substitut fonctionnel de TAF1 au cours de la méiose des cellules germinales mâles humaines (Wang and Page, 2002).

### b) TAF4b

TAF4b, un paralogue de TAF4. L'inactivation de ce gène chez la souris entraîne deux phénotypes : une stérilité chez les femelles où il est nécessaire pour les cellules de la granulosa de l'ovaire et une stérilité chez les mâles où il est nécessaire pour les cellules germinales mâles puisqu'il assure l'auto-renouvellement des spermatogonies. (Freiman *et al.*, 2002). On le trouve exprimé en quantités importantes dans le testicule et dans les lymphocytes B humains (Fig.61). C'est d'ailleurs à partir de ces cellules qu'a été isolé un complexe TFIID contenant TAF4b (Dikstein *et al.*, 1996). La surexpression de TAF4b dans ces cellules mène à des changements transcriptionnels sur certains promoteurs. Cela suggère que des TAFs spécifiques de certaines cellules sont impliqués dans la régulation spécifique de gènes intervenant lors du développement ou dans des tissus différenciés (Bell and Tora, 1999).

### c) TAF6 $\delta$

TAF6 possède plusieurs variants :  $\alpha$ ,  $\beta$ ,  $\gamma$ ,  $\delta$ . L'expression de l'isoforme TAF6 $\delta$  augmente dans les cellules soumises à des stimuli apoptotiques. Le complexe TFIID

contenant cette sous-unité TAF6 $\delta$  est donc impliqué dans l'apoptose (Bell *et al.*, 2001). En réponse à un stress génotoxique, la stabilisation et la modification post-traductionnelle du facteur de transcription p53 mènent à l'induction de l'expression de plusieurs gènes pro-apoptotiques tels que Bax (Ryan *et al.*, 2001). L'activation de la transcription du facteur p53 nécessite l'intervention d'autres facteurs comme les facteurs modifiant l'état de la chromatine, le médiateur, ou TFIID. Dans le hTFIID classique, TAF6 intervient comme co-activateur de la transcription pour la protéine suppresseur de tumeur p53. Le remplacement de TAF6 par son isoforme TAF6 $\delta$  supprime le recrutement de TAF9 au sein du TFIID (Fig.61) et déclenche l'apoptose dans les cellules tumorales, même si elles sont exemptes de p53. TAF6 $\delta$  contrôle donc l'expression de gènes menant à l'apoptose en l'absence de p53 (Wilhelm *et al.*, 2008) prouvant ainsi que des signaux toxiques extracellulaires peuvent entraîner des changements dans la composition en sous-unités de TFIID et contribuer à l'induction de programmes particuliers d'expression des gènes.

d) Les différents complexes hTFIID

Les différents complexes de TAFs impliqués dans l'initiation de la transcription

Avec TBP

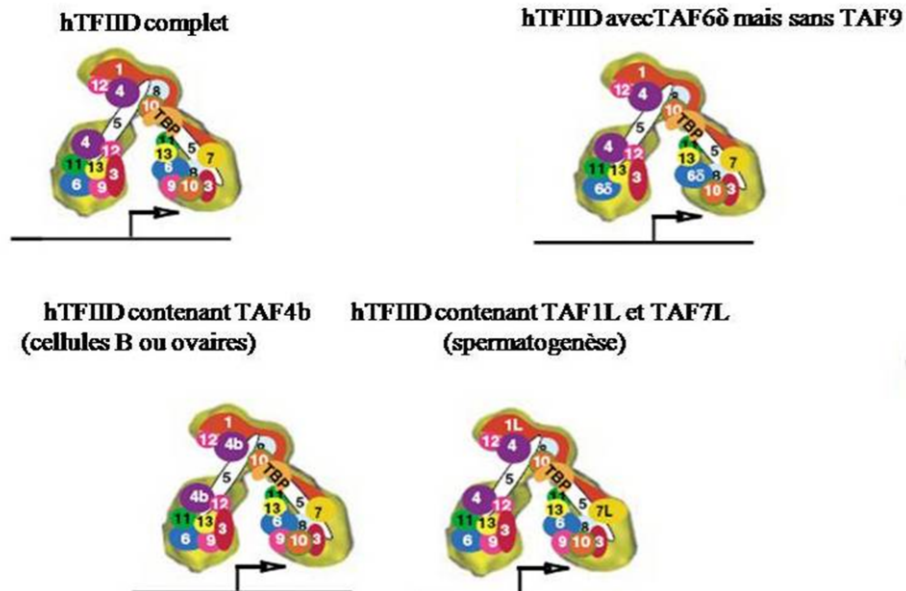


Figure 61 : Représentation schématique de différents complexes de TAFs isolés de cellules humaines (Muller and Tora, 2004).

L'existence de paralogues des facteurs TAFs contribue donc à augmenter la complexité et les possibilités combinatoires des TAFs au sein du facteur TFIID (Fig.61). Ainsi le spectre des divers TAFs permet aux cellules eucaryotes de former des complexes capables de reconnaître des promoteurs complexes, différant en composition et en fonction. La composition du TFIID au sein des cellules vivantes est donc dynamique et sensible aux signaux cellulaires (Muller and Tora, 2004).

## 2. Les facteurs apparentés à TBP

La famille de TBP chez les Métazoaires comprend plusieurs sous-groupes : les plus connus sont TBP, TLF et TBP2.

### a) TLF, TRF2, TLP

TLF (TBP-Like Factor) également appelé TRF2 (TBP-related factor 2) (Rabenstein *et al.*, 1999) et TLP (TBP-like Protein) (Ohbayashi *et al.*, 1999) présente une grande similarité à TBP mais ne provient pas du même arbre phylogénétique (Dantonel *et al.*, 1999). TLFs et TBPs ont 60% d'homologie de séquences. Par contre, leurs domaines spécifiques ont été conservés différemment au cours de l'évolution : à 80% pour TBP et seulement à 40% pour TLF (Dantonel *et al.*, 1999). Cela suggère que les facteurs TBP et TLF ont été soumis à des pressions évolutives différentes, TBP évoluant très peu de la Levure à l'Humain alors que les TLFs semblent assurer une fonction tolérant plus de diversité (Fig.62).

Cependant, la plupart des résidus d'acides-aminés importants pour la liaison de TRF2 avec TFIIA et TFIIB sont restés inchangés et permettent l'assemblage d'un PIC sur les promoteurs dépourvus de boîte TATA (Ohbayashi *et al.*, 2003)

Cette propriété de TRF2 lui permet aussi de fonctionner comme un répresseur, en empêchant la formation du PIC initiée par TBP ou TFIID sur des promoteurs contenant des boîtes TATA, certainement grâce à la compétition envers des quantités limitées de TFIIA et TFIIB (Chong *et al.*, 2005).

Contrairement aux Xénopes, les souris TRF2<sup>-/-</sup> sont viables et ne présentent que des troubles de la spermiogenèse (Martianov *et al.*, 2001). Le rôle biologique de TRF2 semble donc être espèce-spécifique. D'ailleurs, un autre membre de ce sous-ensemble ressemblant à TBP a été identifié chez la Drosophile : il s'agit de TRF1 (Fig.62). *In vivo*, il pourrait même

s'associer avec des facteurs spécifiques des neurones, les nTAFs pour former un complexe multiprotéique distinct de TFIID actif transcriptionnellement (Holmes and Tjian, 2000).

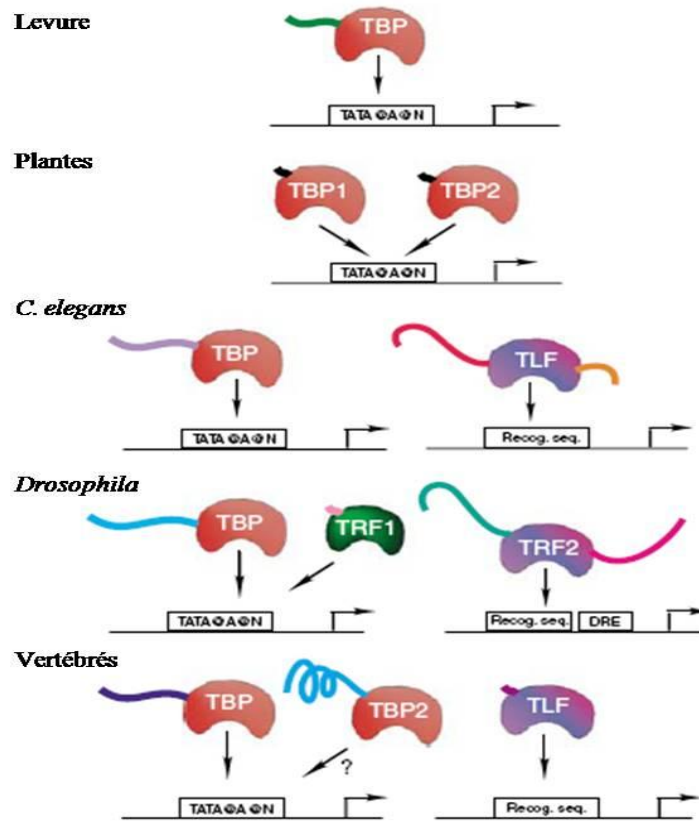


Figure 62 : Facteurs appartenant à la famille TBP codés par différents génomes (Muller and Tora, 2004).

### 6) TBP2

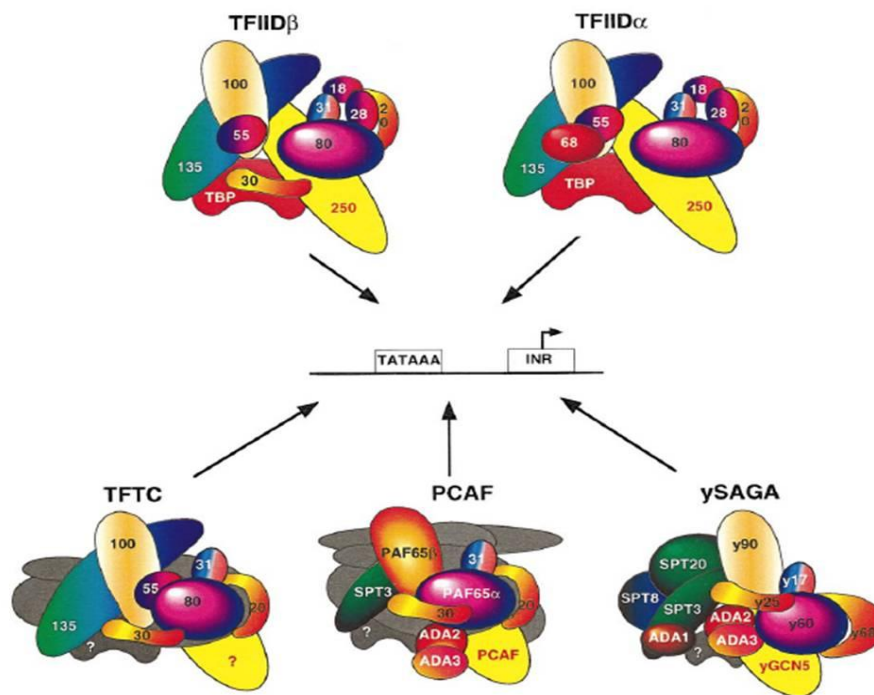
Le deuxième sous-ensemble de TBP a été plus récemment identifié chez les Vertébrés (Veenstra and Wolffe, 2001) et ses protéines ont été nommées TBP2s (Fig.62). Elles partagent une très grande similarité avec TBP avec plus de 92% d'identité. Il est probable que les organismes aient développé différentes isoformes de TBP afin de faciliter une expression génique différentielle assurant la régulation de voies de développement distinctes.

### 3. Les complexes transcriptionnels sans TBP

Pourtant, des complexes dépourvus de TBP ont également pu être isolés. L'initiation de la transcription *in vitro* est donc possible sans le facteur TBP en présence de ces



complexes. Ceux-ci remplacent alors le facteur TFIID sur des promoteurs sans boîte TATA (Wieczorek *et al.*, 1998).



**Figure 63 :** Comparaison des complexes sans TBP à TFIID. (Pour la nomenclature des TAFs, voir Tableau n°4).

Selon leur définition, les TAFs étaient censés n'exister qu'en présence de TBP. Pourtant, la validité de cette définition des TAFs a été sérieusement remise en question depuis que plusieurs complexes contenant des TAFs et dépourvus de TBP ont été isolés chez la Levure, la Drosophile ou dans des cellules humaines (Fig.63).

Parmi les plus étudiés, on trouve TFTC (TBP-free TAF<sub>II</sub>s-containing complex), le complexe TFTC-related PCAF/GCN5, SAGA (Spt-Ada-Gcn5 Acetyl transferase) et son homologue STAGA, SLIK (SAGA-like complex), SALSA (SAGA altered Spt8 absent) et PRC1 (Polycomb repressive complex 1) (Fig.63).

TFTC, PCAF et SAGA sont des complexes multiprotéiques tous trois dépourvus de TAF1 et de TBP. Ils contiennent des sous-ensembles de TAFs et parmi eux, 3 possèdent un HFD, contrairement à TFIID, ainsi qu'une sous-unité à activité histone acétyltransférase (HAT) (Grant *et al.*, 1998). Cette HAT est contenue par la protéine GCN5 de *Saccharomyces cerevisiae* (Brownell *et al.*, 1996) dans 2 de ces complexes et a été découverte par homologie

avec la protéine p55 du Cilié *Tetrahymena thermophila* (Brownell and Allis, 1995). La découverte de cette fonction HAT chez GCN5 a permis de faire le lien entre l'acétylation des nucléosomes et l'activation de l'expression de certains gènes.

a) *Le complexe de levure SAGA*

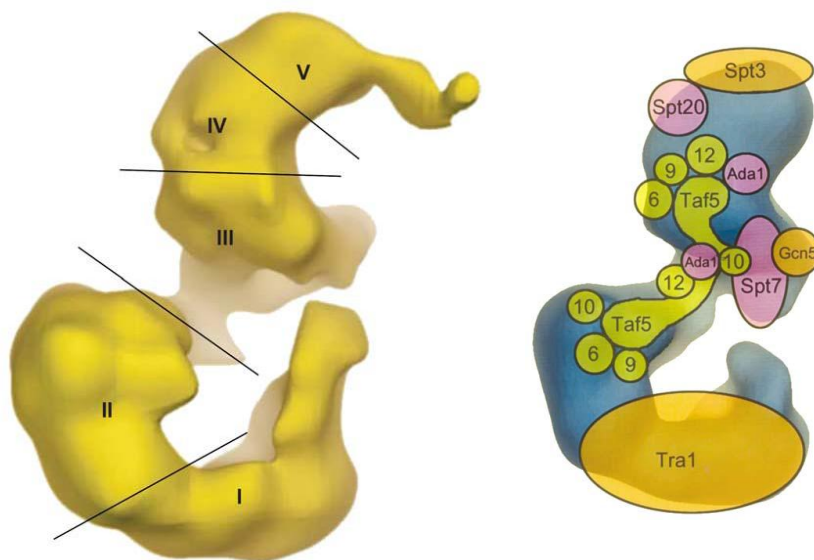
*-Composition et structure du complexe SAGA*

Le complexe SAGA, isolé chez la levure (Fig.63), est le mieux caractérisé des complexes SAGA/TFTC/STAGA/PCAF. C'est un large complexe de 1,8 MD composé d'une quinzaine de sous-unités qui réunit divers groupes de protéines : les adaptateurs de transcription (Ada1, Ada2, Ada3, Ada5, et Gcn5), des protéines Spt (Spt3, Spt7, Spt8 et Spt20), une partie des TAFs (TAF5, TAF6, TAF9, TAF10 et TAF12), la protéine Tra1 homologue de la protéine hTRRAP humaine. En plus de ces sous-unités, de nouvelles sous-unités ont été identifiées : Ubp8, une protéase spécifique de la dé-ubiquitination de H2B, Sgf11 requis pour l'incorporation de Ubp8 dans les complexes (Ingvarsdottir et al., 2005), Sgf29 et Sgf73, dont les fonctions sont inconnues, et la protéine Sus1 qui fait aussi partie de la machinerie d'export nucléaire de l'ARNm.

La structure du complexe SAGA a été résolue à 30 Å par microscopie électronique et l'emplacement des sous-unités, déterminé par immunolocalisation, a permis d'attribuer des fonctions spécifiques à chaque domaine du complexe SAGA (Wu et al., 2004) (Fig.64). Un modèle a été proposé dans lequel 5 domaines sont décrits (Fig.64). Le domaine 1, qui contient la protéine Tra1, l'homologue de hTRRAP, peut servir d'interface avec les activateurs. Les domaines II, III et IV, au centre de la structure, contiennent principalement des TAFs (le domaine II contient TAF5 et TAF6, le domaine III TAF10, TAF5 Gcn5, Spt7 et Ada1, et le domaine IV Ada1 et TAF6). On remarque que le domaine III contient les sous-unités Gcn5 et Spt7, il n'est donc pas uniquement un domaine structural important, mais aussi fonctionnel puisqu'il est doué d'activité HAT. Enfin, le dernier domaine, dont la structure semble plus flexible se compose des sous-unités Spt3, Spt20 et probablement Spt8, délimite un module capable d'interagir avec la protéine TBP.

L'analyse de mutants de sous-unités du complexe SAGA a permis de classer les protéines Ada et Spt en plusieurs catégories selon leur rôle structural et fonctionnel. Des mutations à l'intérieur des gènes codant pour **Ada1, Spt7 ou Spt20/Ada5** ont conduit à la

destruction des complexes SAGA et à l'inhibition de la croissance cellulaire, ce qui accorde un rôle crucial à ces sous-unités dans la structure et la fonction du complexe SAGA. Deux de ces sous-unités, Spt7 et Ada1, ont été immunolocalisées en microscopie électronique au niveau du domaine III, soit au centre de la structure (Fig.64). Il paraît donc logique que leur délétion déstabilise complètement la formation du complexe SAGA. Lorsque d'autres sous-unités comme Ada2/Ada3/Gcn5 ou Spt3/Spt8 sont mutées, le complexe garde son intégrité et les phénotypes observés sont très discrets (Stern et al., 1999), ce qui suggère que ces sous-unités jouent des rôles annexes.



**Figure 64 : Structure du complexe SAGA résolue par microscopie électronique, d'après (Timmers and Tora, 2005; Wu et al., 2004 ).**

A – Représentation schématique de la localisation des sous-unités du complexe SAGA sur la structure 3D. Les Taf5 sont représentées en jaune. L'extension de Taf5, ainsi que la localisation de Taf9 et de Taf12 ont été déduites de précédentes immunolocalisations au niveau de TFIID (Leurent et al., 2004 ; Leurent et al., 2002). Les sous-unités importantes pour l'architecture du complexe, ainsi que celles comportant un domaine de repliement de type « histone fold » sont représentées en violet, tandis que les sous-unités impliquées à différents niveaux de la transcription sont représentées en orange. B – Structure tridimensionnelle du complexe SAGA avec un module flexible. La taille de ce complexe est de 18 x 28 nm et comprend cinq domaines. Les cinq domaines contiennent différentes sous-unités de SAGA et sont séparés artificiellement par des barres.

- SAGA, coactivateur transcriptionnel

*In vivo* et *in vitro*, les composants du complexe SAGA sont essentiels au bon déroulement de la transcription de certains gènes.

*In vitro*, le complexe SAGA purifié est capable d'activer la transcription dans un contexte chromatinien, grâce d'une part à son activité HAT et d'autre part à sa capacité à se lier aux activateurs acides tels que VP16 ou Gcn4 (Bhaumik et al., 2004; Qiu et al., 2005 ; Utley et al., 1998 ).

*In vivo*, l'examen de mutants des sous-unités du complexe SAGA a permis d'identifier des gènes dont la transcription est activée par le complexe SAGA, comme *Gall*, *Trp3*, *His3* (Larschan and Winston, 2001). D'autres approches *in vivo* ont été utilisées pour étudier la liaison des sous-unités du complexe SAGA au niveau de certains promoteurs. Des études d'immunoprécipitation de la chromatine (ChIP) ont montré que le domaine d'activation Gal4 recrute SAGA au niveau d'une séquence activatrice en amont (UAS) du promoteur du gène *Gall*.

Les effets de l'activateur Gal4, comme le recrutement de SAGA ou la stimulation de l'assemblage du PIC, sont dépendants d'une partie des sous-unités du complexe SAGA. Ces études ont permis d'élaborer un modèle d'activation de la transcription par Gal4 (Fig. 65) dans lequel Gal4 se fixe d'abord à la séquence UAS, puis recrute le complexe SAGA, qui à son tour recrute TBP au niveau de la boîte TATA, ce qui permet alors l'arrivée des facteurs généraux de la transcription (Bhaumik and Green, 2001).

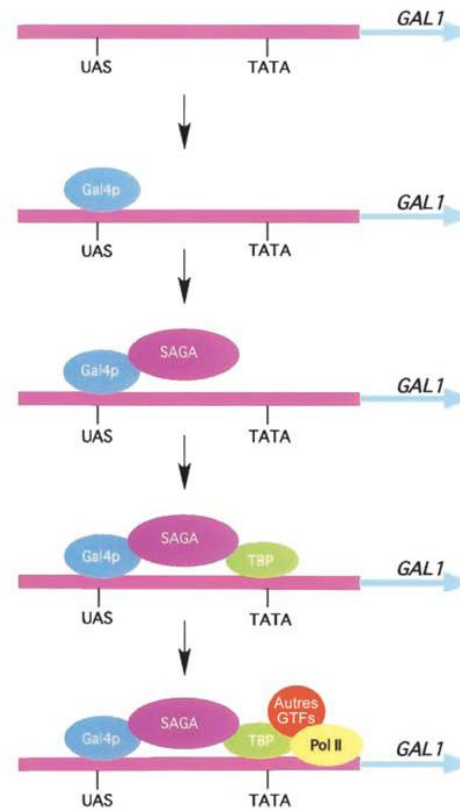
D'autres études *in vivo* par **FRET** (*Fluorescence Resonance Energy Transfert*) ont montré une interaction directe entre l'activateur acide Gal4 et Tra1, la plus grande sous-unité du complexe SAGA (Bhaumik et al., 2004).

Pour le génome entier de levure, SAGA est impliqué dans la transcription de **10 % des gènes** qui correspondent à **des gènes inductibles après stress** (Huisinga and Pugh, 2004). Au niveau de ces promoteurs, SAGA est nécessaire au recrutement de la machinerie transcriptionnelle de base, ainsi qu'au recrutement probable de la protéine TBP, ces promoteurs étant pourvus d'une boîte TATA) (Bhaumik and Green, 2002 ; Huisinga and Pugh, 2004; Larschan and Winston, 2001 ). SAGA, contrairement à TFIID ne se fixe pas au cœur du promoteur, mais via des séquences activatrices en amont (UAS) (Bhaumik and

Green, 2001; Larschan and Winston, 2001), en accord avec son implication dans la transcription de gènes inductibles. Quoiqu'il en soit, SAGA semble participer à l'assemblage du PIC, comme TFIID, et s'impose alors comme facteur essentiel pour initier la transcription de certains gènes.

**Figure 65 : SAGA, co-activateur transcriptionnel.**

Résumé de l'assemblage des complexes multiprotéiques au niveau du promoteur *GAL1*, d'après (Bhaumik and Green, 2001). Les études cinétiques n'ont pas permis de déterminer l'ordre de recrutement entre SAGA et TBP. Toutefois, le résultat d'études de mutations indiquent que SAGA est capable de se fixer au niveau de ce promoteur en l'absence de TBP. Dans ce modèle, le recrutement des GTFs est représenté après celui de TBP, mais il est possible que certaines GTFs se fixent conjointement à TBP.



SAGA et TFIID interviennent-ils ensemble ou séparément pour initier la transcription ?

Tandis que la transcription de 70% des gènes est dépendante des sous-unités partagées par TFIID et SAGA, la transcription de seulement 12% des gènes est affectée par l'absence des sous-unités spécifiques de SAGA, et 30% par l'absence de sous-unités spécifiques de TFIID (Lee et al., 2000).

### -Le complexe SAGA, doué d'activité HAT

SAGA participe également à l'activation de la transcription en facilitant l'accessibilité à l'ADN de la machinerie transcriptionnelle de base grâce à sa protéine Gcn5 capable d'acétyler les histones. L'intégration de Gcn5 dans le complexe SAGA lui confère une spécificité de substrat toute particulière. Gcn5 dans SAGA peut reconnaître les histones

agencées en nucléosomes, s’y fixer et acétyler préférentiellement l’histone H3 mais aussi l’histone H2B. La reconnaissance du substrat nucléosomal implique le bromodomaine de la protéine Gcn5. Ce domaine s’associe avec les protéines Ada, ce qui souligne l’importance des autres sous-unités du complexe SAGA dans la reconnaissance des nucléosomes. De plus, le profil d’acétylation de la protéine Gcn5 intégrée à SAGA est plus étendu : en plus de la lysine 14, d’autres résidus lysines sont acétylés à l’extrémité N-terminale de l’histone H3. La fonction de la protéine Gcn5 est ainsi modifiée en fonction des protéines qui lui sont liées.

#### -Le complexe ADA

Un autre complexe de 800 Kda, ADA, partage quelques protéines avec le complexe SAGA, comme Gcn5, Ada2 et Ada3, et possède la même capacité d’acétyler les histones agencées en nucléosomes. Toutefois, la spécificité d’acétylation est quelque peu différente, puisque le complexe ADA acétyle uniquement les résidus 14 et 18 sur l’histone H3, tandis que SAGA acétyle les résidus 9, 14, 18 et 23 (Eberharter et al., 1999 ; Grant et al., 1998a). Ce complexe n’est pas simplement un module de SAGA, car il possède une sous-unité spécifique nommée Ahc1 (pour ADA HAT complex component 1) impliquée dans l’intégrité structurale du complexe. Néanmoins, les fonctions spécifiques de ce complexe sont encore mal connues (Eberharter et al., 1999).

#### -Les TAFs au sein du complexe SAGA

Le complexe SAGA est composé d’un ensemble de **TAFs** important pour la fonction d’activation de la transcription et d’acétylation *in vitro*. Le rôle exact des TAFs au sein du complexe SAGA n’est pas encore connu. Il est probable que ces protéines jouent un rôle architectural majeur au sein du complexe, puisqu’elles sont localisées à des positions clés, formant le squelette du complexe. On peut ainsi imaginer que l’importance fonctionnelle des TAFs découle de leur rôle structural, que ces protéines permettent d’intégrer les autres sous-unités pour rendre le complexe fonctionnel. Les TAFs peuvent également avoir des fonctions plus spécifiques, comme participer à l’interaction avec TBP ou être impliquées dans le déplacement des histones, mais ces hypothèses restent à démontrer. La protéine **Tra1** joue un rôle structural de premier ordre dans l’édifice du complexe SAGA et interagit avec des activateurs acides tels que VP16, ce qui lui confère un rôle de coactivateur transcriptionnel (Bhaumik et al., 2004 ; Qiu et al., 2005).

-un complexe dynamique

Des travaux plus récents ont montré que la composition de ce complexe est dynamique et ses fonctions variées, et toutes deux dépendent des différentes conditions dans lesquelles la cellule se trouve. Les complexes SAGA « classiques » sont formés lorsque les levures sont dans un milieu riche, qui réprime la transcription de *His3*, tandis qu'en les plaçant dans des conditions où *His3* est transcrit, un complexe SAGA altéré est formé (dépourvue de protéine Spt8) (Belotserkovskaya et al., 2000). Ces résultats vont dans le sens d'autres études qui ont montré qu'un variant du complexe SAGA, appelé **SALSA** ou **SLIK** (sans Spt8 et avec une forme tronquée de la protéine Spt7) constitue la forme fonctionnelle pour activer la transcription chez la levure (Sterner et al., 2002a). Ces résultats ont permis de construire un modèle d'activation de la transcription des complexes multiprotéiques de type SAGA, dans lequel une forme spécifique du complexe serait prédominante pour activer la transcription au niveau de certains promoteurs, alors que l'autre forme n'est visualisée que lorsque les gènes ne sont pas transcrits. Il reste à comprendre comment les complexes modifient leur composition pour passer d'une forme inactive à une forme active en transcription.

Le complexe SAGA, sous diverses formes, semble réguler la transcription d'ensemble de gènes différents (Pray-Grant et al., 2002). Ainsi les complexes SAGA et **SLIK** (SALSA) ont en commun certaines activités, alors que d'autres leur sont spécifiques et restent à clarifier (Wu and Winston, 2002). La sous-unité Spt8, présente dans le complexe SAGA et non dans SLIK, joue un rôle important dans l'interaction entre SAGA et TBP. *In vitro*, les mutants *spt8Δ* sont incapables d'interagir avec TBP (Saleh et al., 1997). La protéine Spt3, impliquée elle aussi dans l'interaction avec TBP, est requise pour la formation du complexe de préinitiation au niveau du promoteur *GALI* (Bhaumik and Green, 2001 ; Dudley et al., 1999). Les interactions Spt3-TBP ou Spt8-TBP inhibent la transcription au niveau de certains promoteurs (Warfield et al., 2004). L'interaction entre Spt8 et TBP est soit nécessaire à l'initiation de la transcription au niveau de certains promoteurs, soit elle inhibe la transcription au niveau d'autres promoteurs. Cette « double » fonction de Spt8 confère une spécificité fonctionnelle au complexe SAGA, mais pas au complexe SLIK, qui en est dépourvu.

En résumé, d'après le modèle construit à partir de la structure tridimensionnelle de SAGA (microscopie électronique), certaines sous-unités du complexe SAGA, plus mobiles,

peuvent modifier la composition de ce complexe. Ainsi, grâce à ce vaste réseau d'interactions entre protéines, toujours dynamique, il est probable que ce complexe soit impliqué dans des activités assez variées.

### -Action concertée de SAGA, du Médiateur et de SAGA/Swi/Snf

Le complexe SAGA s'agence à d'autres complexes de protéines dans le but d'activer la transcription, tels que le complexe Médiateur et les complexes Swi-Snf. Les sous-unités du complexe Médiateur de levure MED12-CycC jouent un rôle dans l'activation de la transcription du gène *Gall* comme SAGA. Des études d'immunoprécipitation de la chromatine suggèrent que MED12-CycC ne s'associe au promoteur du gène *Gall* qu'en présence du complexe SAGA. De même, les liaisons de TBP et MED12-CycC au niveau du promoteur *gall* sont interdépendantes (Larschan and Winston, 2005). D'autres études plus globales vont dans ce sens, car elles ont montré que la majorité des promoteurs régulés par SAGA requiert également la présence du complexe Médiateur (Basehoar et al., 2004 ; Huisinga and Pugh, 2004 ; Qiu et al., 2005).

SAGA et le complexe Swi/Snf (impliqué dans le remodelage de la chromatine) interagissent ensemble et participent à l'activation de la transcription de manière complémentaire ou partiellement redondante (Topalidou et al., 2004).

Ces expériences soulignent l'importance de recruter plusieurs complexes de protéines au niveau de certains promoteurs pour activer la transcription, notamment par l'altération de la structure chromatiniennne.

### 6) Le complexe PCAF

Le complexe PCAF comprend une protéine douée d'activité HAT nommée PCAF (Fig.63) (p300/CBP Associated Factor) et partage de nombreuses sous-unités protéiques avec les complexes GCN5, STAGA et TFTC.

La protéine p300/CBP, qui interagit avec la protéine PCAF, n'est pas présente en quantité stœchiométrique dans le complexe PCAF purifié, ce qui indiquerait que l'interaction ne peut s'établir qu'au niveau de certains promoteurs. Ce complexe a été isolé grâce à l'expression de PCAF portant un épitope Flag dans des cellules de type HeLa. Une vingtaine de sous-unités ont été identifiées dont 11 sont homologues au complexe de levure SAGA. Ce



complexe comprend par exemple les sous-unités hADA2, hADA3, SPT3, hTRRAP, TAF9, TAF10, TAF12, TAF6L et TAF5L (Vassilev et al., 1998), homologues de sous-unités de SAGA.

Par contre, certaines sous-unités du complexe SAGA sont absentes du complexe PCAF, comme la protéine TAF5, ce qui est particulièrement surprenant puisque TAF5 est localisée dans une région cruciale pour l'intégrité de l'architecture du complexe SAGA chez la levure (Wu et al., 2004). Ce complexe comprend cependant les protéines TAF5L et TAF6L, très similaires aux protéines TAF5 et TAF6 respectivement. Ces protéines pourraient remplacer les TAFs pour former une structure ressemblant à celle des nucléosomes (Schiltz and Nakatani, 2000).

De nombreuses similarités sont observées entre les complexes PCAF et les complexes qui contiennent la protéine GCN5, tels que STAGA et TFIIIC, mais leurs sous-unités ne sont pas toutes identiques. Les protéines PCAF et GCN5 s'expriment dans des tissus spécifiques chez la souris. Il est alors envisageable que ces complexes multiprotéiques diffèrent au niveau de leur composition ou de leur activité en fonction des tissus ou bien du stade de développement cellulaire.

c) *Le complexe STAGA (SPT3-TAF(II)31-GCN5L acetylase)*

*-Composition de STAGA*

C'est un homologue du complexe SAGA chez la levure. Il contient l'HAT GCN5, un certain nombre de TAFs (TAF10, TAF12, TAF9), SPT7, SPT3, ADA1, ADA3, SAP 130, TAF5L et TAF6L (Fig.63). A la différence du complexe TFIIIC isolé au laboratoire, ce complexe ne semble pas contenir de TAFs de haut poids moléculaire telles que TAF5 ou bien TAF6. Or, l'immunolocalisation de TAF5 dans SAGA attribue une place centrale à TAF5 dans l'architecture du complexe, ce qui laisse penser que sans cette protéine, le complexe est détruit (Wu et al., 2004). Le complexe STAGA contient les protéines TAF5L et TAF6L (PAF65 $\beta$  et PAF65 $\alpha$ , respectivement) qui, homologues des TAF5 et 6, peuvent les remplacer (Martinez et al., 2001). La composition des complexes TFIIIC et STAGA est donc très proche et il devient difficile de les distinguer. Plusieurs groupes ont nommé l'ensemble de ces complexes « TFIIIC-like ».

### -Fonctions du complexe STAGA

Il existe peu de données fonctionnelles pour ce complexe. Sans surprise, STAGA s'avère doué d'**activité HAT** ciblée principalement sur l'histone H3, tout comme GCN5. Des tests d'acétylation *in vitro*, sur des histones libres ou agencées en nucléosomes, ont en effet révélé la capacité de STAGA à acétyler l'histone H3 (Martinez et al., 2001). La capacité d'activer la transcription *in vitro* a été testée par la suite. Premièrement, **STAGA** se fixe sur l'activateur **VP16** *in vitro* (chromatographie d'affinité GST-VP16). Ensuite, des expériences d'activation de la transcription ont été réalisées avec VP16 et elles montrent l'incapacité de STAGA à activer la transcription sur une matrice d'ADN nue. Par contre, lorsque l'expérience est réalisée avec un substrat d'ADN chromatinien, STAGA facilite l'activation de la transcription par VP16. Ces résultats concordent avec la fonction d'acétylation des histones H3, dont on présume qu'elle facilite la transcription uniquement dans un contexte chromatinien (Martinez et al., 2001).

STAGA est par ailleurs impliqué dans **l'activation de la transcription par c-myc**. Des expériences d'interaction *in vitro* (immunoprécipitation et chromatographie d'affinité GST-Myc) révèlent une interaction stable entre c-myc et les sous-unités du complexe STAGA. Grâce à des expériences de transfection transitoires, les auteurs ont montré que TRRAP et GCN5 sont tous deux requis pour l'activation de la transcription par c-myc (Liu et al., 2003b).

Ainsi le complexe STAGA joue un rôle dans la transcription en tant que **coactivateur**. D'autre part, ce complexe est impliqué dans la **réparation de l'ADN par excision de nucléotides**. Il peut lier différentes protéines impliquées dans l'épissage alternatif et dans la réparation par excision de nucléotides (Martinez et al., 2001). Certaines des fonctions du complexe STAGA sont communes au complexe TFIC, comme l'activation de la transcription par VP16, ou bien l'implication dans la réparation des cassures d'ADN.

#### d) Le complexe TFIC

### -Caractérisation biochimique et structurale

Une fraction sans la protéine TBP a été identifiée au sein du laboratoire Tora (Fig.66). Elle contenait des TAFs associées à d'autres facteurs mais pas à TBP. Après caractérisations biochimiques et structurales, il s'est avéré que cette fraction contenait un nouveau complexe

multiprotéique stable appelé **TFTC** pour « *TBP Free TAF Containing Complex* ». La purification de TFTC s'articule en plusieurs parties. Tout d'abord, des extraits de cellules HeLa sont immunoprécipités à l'aide d'un anticorps contre TAF10 ; le matériel élué est ensuite immunoprécipité à l'aide d'un anticorps contre TBP ; ainsi l'éluat de deuxième immunoprécipitation contient le complexe TFIID. Le surnageant de cette deuxième IP est par la suite immunodéplété en TBP, et cette fraction contient le complexe TFTC (Wieczorek et al., 1998).

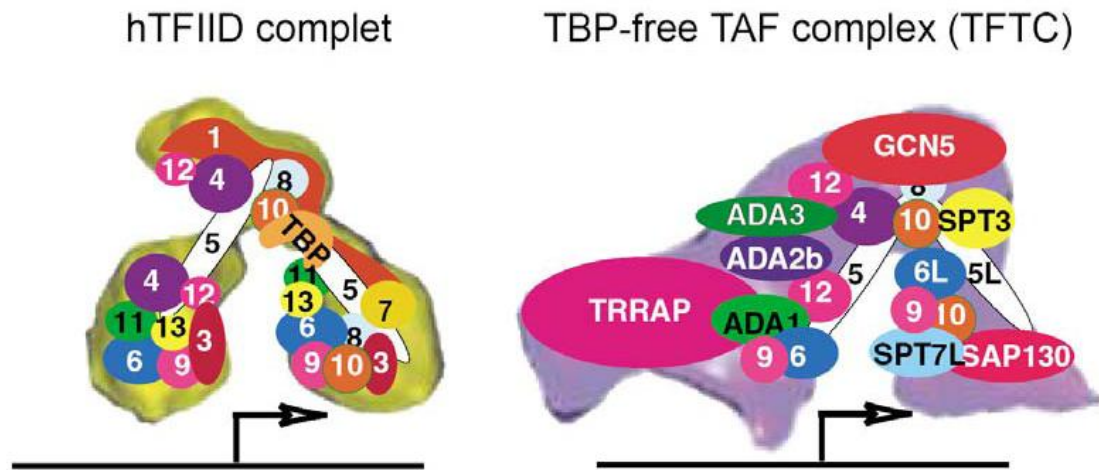


Figure 66 : Représentation schématique des complexes TFIID et TFTC, d'après (Muller and Tora, 2003).

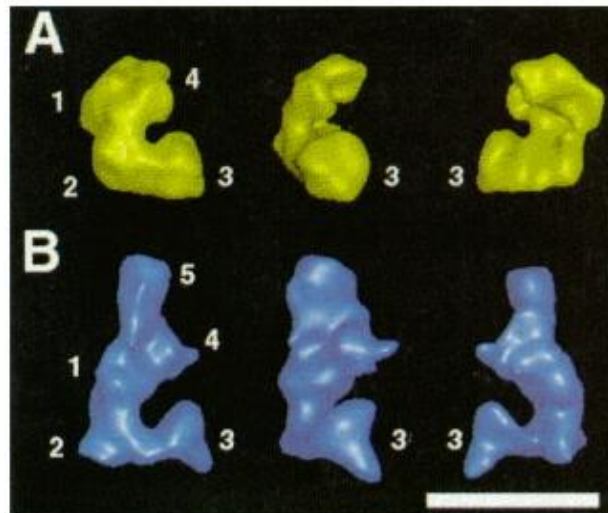
### -La composition en protéines

Ce complexe partage certaines sous-unités avec le complexe TFIID, comme TAF2, TAF4, TAF5, TAF6, TAF7, TAF9, TAF10 et TAF12. Par contre, certaines TAFs sont présentes spécifiquement dans TFIID, TAF1 et TAF11. D'une manière logique, les TAFs présentes uniquement dans TFIID interagissent avec TBP. TFTC comprend des sous-unités spécifiques, comme hTRRAP, l'HAT GCN5, ADA1, ADA3, SPT3, SPT7-like, SAP130, TAF5L et TAF6L qui sont retrouvées dans le complexe STAGA et constituent des homologues de sous-unités du complexe SAGA chez la levure.

### -La structure du complexe TFTC

La structure du complexe multiprotéique TFTC a été résolue par microscopie électronique au laboratoire, en collaboration avec le laboratoire du Dr Schultz (Brand et al., 1999a) (Fig.67). La structure tridimensionnelle de TFTC, résolue à 35Å en microscopie

électronique ressemble, comme celle de TFIID, à un serre-joint moléculaire mais avec 5 domaines globulaires au lieu de 3. Elle présente une cannelure dans laquelle peut s'insérer l'ADN double-brin (Brand *et al.*, 1999a). Cette configuration suggère que TFTC a adopté une conformation de liaison à l'ADN relativement similaire à celle de TFIID (Andel *et al.*, 1999).



**Figure 67 : TFTC et TFIID, structures tridimensionnelles des complexes TFIID en haut (A) et TFTC en bas (B), résolues par microscopie électronique, d'après (Brand *et al.*, 1999a).**

Les trois orientations correspondent à une rotation d'un angle de 80° autour d'un axe vertical. Les domaines sont numérotés de 1 à 4 pour TFIID et de un à cinq pour TFTC. A – 115 vues différentes ont été utilisées pour construire le modèle 3D de TFIID. B- Un total de 112 vues de TFTC ont été utilisées pour construire ce modèle tridimensionnel. La barre blanche représente 20 nm.

La structure du complexe SAGA, homologue de TFTC, a depuis été résolue et certaines sous-unités, immunolocalisées. Par analogie, il est possible d'imaginer maintenant la localisation de certaines protéines dans TFTC. La résolution de la structure du complexe TFTC confirme visuellement l'existence de ce complexe.

### -Etudes fonctionnelles

#### \*Test de transcription *in vitro*

Tout comme TFIID, TFTC est capable d'intervenir dans la transcription basale ainsi que dans celle induite par Gal4-VP16 *in vitro* sur des promoteurs contenant ou non une boîte TATA. Cela se ferait apparemment grâce à la reconnaissance du promoteur minimal par les

TAFs et par l'interaction avec d'autres composants de la machinerie transcriptionnelle (Wieczorek *et al.*, 1998).

Suite à la mise en évidence d'une implication de TFTC dans l'initiation de la transcription, il a été suggéré que le complexe TFTC puisse se fixer au promoteur comme TFIID. Des expériences de protection à la DNaseI ont alors montré que TFTC peut reconnaître des séquences de l'élément initiateur du promoteur AdMLP. Les TAFs peuvent jouer un rôle important dans la liaison de TFTC au niveau de l'élément initiateur, puisqu'elles peuvent également se fixer au niveau de cette région (Wieczorek *et al.*, 1998).

#### \*Test HAT

Le complexe TFTC, qui comprend la protéine GCN5 douée d'activité HAT, est capable d'acétyler *in vitro* les histones libres H3 et H1 (Brand *et al.*, 1999b). Chez la levure, la protéine recombinante Gcn5 acétyle très bien les histones libres, mais pas celles agencées en nucléosomes. L'activité HAT de Gcn5 semble modulée par les protéines qui lui sont liées. En effet, lorsque la protéine native Gcn5 est intégrée à des complexes du type SAGA, elle devient capable d'acétyler les nucléosomes *in vitro* (Grant *et al.*, 1997). De même, l'ensemble du complexe TFTC est également requis pour acétyler l'histone H3 agencé en nucléosomes (Brand *et al.*, 1999b). Ainsi la protéine GCN5 a besoin d'autres facteurs pour pouvoir acétyler les histones agencés en nucléosomes.

#### \*La fonction de co-activation de la transcription

*In vitro*, TFTC est capable d'activer la transcription par VP16 sur une matrice d'ADN nue (Wieczorek *et al.*, 1998). TFTC semble également impliqué dans l'activation de la transcription par les récepteurs nucléaires en présence de leurs ligands. *In vivo*, certaines sous-unités du complexe TFTC se lient au récepteur aux oestrogènes ER $\alpha$  uniquement en présence de ligand (E2). Des expériences de transfection ont ensuite souligné l'importance de la présence des sous-unités hTRRAP et GCN5 pour activer la transcription par l'intermédiaire des récepteurs nucléaires (Yanagisawa *et al.*, 2002). Ces études montrent une implication probable des complexes de type TFTC dans l'activation de la transcription par les oestrogènes.

#### \*la réparation par excision de nucléotides

Le complexe TFIC contient une sous-unité de 130 Kda, SAP130 (spliceosome associated protein 130) fortement homologue à la protéine DDB1, qui est recrutée au niveau de lésions d'ADN induites par les UV *in vivo* (Brand et al., 2001; Green and Almouzni, 2003). *In vitro*, le complexe TFIC se fixe préférentiellement au niveau de l'ADN soumis aux UV et n'est alors plus disponible pour activer la transcription. Ainsi, le complexe TFIC, en plus de son rôle dans l'initiation de la transcription, pourrait être impliqué dans la réparation des cassures d'ADN par excision de nucléotides. Les deux processus de transcription et de réparation sont en compétition, et, en présence de lésions au niveau de l'ADN, TFIC préfère se lier sur des lésions pour acétyler les histones et pour faciliter la réparation (Brand et al., 2001).

e) Bilan sur ces complexes sans TBP

Pris dans leur ensemble, tous ces résultats suggèrent que les différents complexes apparentés à TFIID ont des rôles spécifiques :

- dans la reconnaissance de différents promoteurs
- dans l'interaction avec des ensembles de facteurs de transcription différents
- dans la mise en place de réponses différentielles à des activateurs transcriptionnels distincts.

La diversité des promoteurs de base fait qu'ils peuvent être reconnus par différentes protéines ou différents complexes protéiques, chacun d'entre eux contribuant à la transcription d'un faible pourcentage de gènes. Cette hypothèse a émergé des propriétés respectives des facteurs de la famille de TBP (TRFs; Dantonel et al. 1999; Lemon and Tjian 2000) et de TFIC (Wieczorek et al. 1998).

Le complexe TFIC contient de nombreux TAFs en commun avec TFIID et il peut se substituer à lui pour des essais de transcription *in vitro* même s'il manque de TBP. Bien que cette absence de TBP suggère que TFIC s'occupe des promoteurs dépourvus de boîte TATA, TFIC et TFIID stimulent la transcription avec une égale efficacité que ce soit sur des promoteurs dépourvus de boîte TATA ou non (Wieczorek et al. 1998).

Ces résultats suggèrent que TFIC ne semble pas discriminer les promoteurs ayant des structures différentes ; il apporte alors probablement des contributions différentes à chacun de ces promoteurs en terme de régulation de la transcription.

L'existence de ces différents complexes suggère que les différents TAFs ne régulent qu'un ensemble limité de gènes spécifiques. La présence de ces complexes et leur rôle dans la transcription en l'absence de TBP indiquent clairement qu'il existe un degré supérieur de complexité dans la régulation de l'expression des gènes eucaryotes.

## II. Les TAFs et leur véritable fonction dans la régulation des gènes

L'intégration d'un grand nombre de signaux extra et intracellulaires et leur conversion en une réponse transcriptionnelle nécessitent la présence d'activateurs qui contacteraient directement les composants de la machinerie basale, afin d'influencer l'assemblage du PIC et de permettre la régulation des gènes (Fig.68).

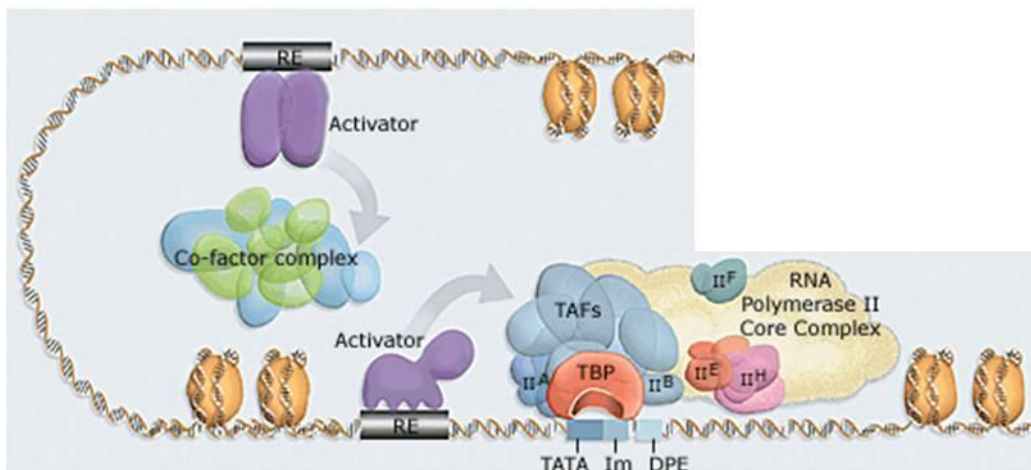


Figure 68 : Contacts entre activateurs et composants de la machinerie basale.

### A. Les TAFs assurent la stabilité de TFIID sur le promoteur

La fixation des TAFs sur le promoteur, en plus du facteur TBP, permet tout d'abord d'élargir la séquence spécifique d'ADN reconnue par le TFIID. En effet, comme nous l'avons vu au cours du premier chapitre, TFIID peut contacter directement la boîte TATA grâce à TBP, mais il peut également contacter l'élément initiateur via TAF1/TAF2, le DPE via TAF6/TAF9 et le DCE via TAF1 (Fig.6). Ce faisant, il génère ainsi une région de protection étendue, de part et d'autre de la boîte TATA, plus importante que celle définie par le seul facteur TBP, comme le montre les expériences d'empreintes à la DNase I sur le promoteur

tardif majeur de l'adénovirus (Chiang *et al.*, 1993; Zhou *et al.*, 1992). De plus, le contact extensif établi par TFIID pourrait fournir une stabilité supplémentaire à la fixation du PIC et au recrutement de la Pol II au niveau du promoteur (Naar *et al.*, 2001). Ces interactions TAFs-ADN peuvent intervenir sur la plupart des promoteurs, mais sont vraisemblablement particulièrement importantes pour les promoteurs dépourvus de séquence canonique TATA (Burke and Kadonaga, 1997). Grâce à ces multiples interactions, l'affinité de TFIID qui en résulte pour les séquences du promoteur minimal pourrait être augmentée et favoriser la mise en place d'une activité transcriptionnelle de haut niveau. Car si les TAFs interviennent avant tout dans la stabilisation de la liaison de TFIID sur le promoteur grâce à des contacts directs avec l'ADN, au niveau des éléments INR et DPE (Chalkley and Verrijzer, 1999), ils peuvent aussi servir d'adaptateurs entre la machinerie transcriptionnelle de base et certains activateurs.

### 1. L'occupation des TAFs sur les promoteurs

Des expériences d'immunoprécipitations de la chromatine ont permis d'entrevoir que l'activité transcriptionnelle chez la Levure était fortement corrélée avec l'occupation des promoteurs par la Pol II et par les GTFs. Si TFIID est recruté de manière constante au niveau du promoteur, présent à la fois sur les promoteurs dépendants et indépendants des TAFs, le ratio d'occupation des TAF/TBP doit être constant sur tous les promoteurs en conditions physiologiques. Des expériences d'immunoprécipitation de la chromatine démontrent au contraire que ce ratio TAF/TBP varie considérablement selon les promoteurs que ce soit pour TAF1, TAF2, TAF11 ou TAF9 (Kuras *et al.*, 2000). Il existerait donc deux types de promoteurs, ceux qui dépendent des TAFs pour leur activation et ceux qui en sont indépendants.

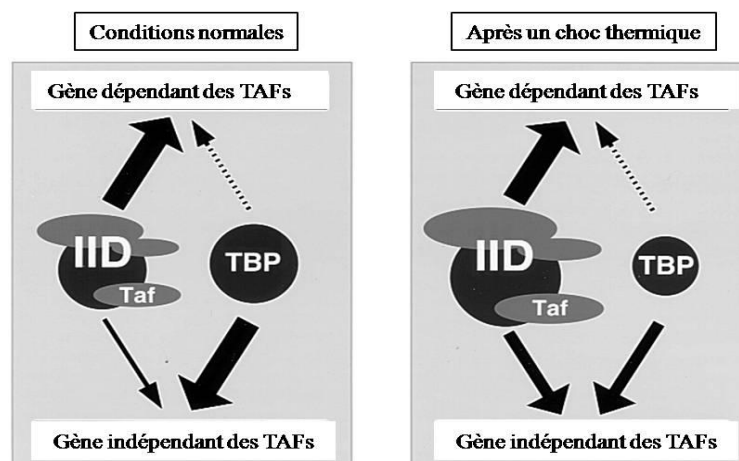
Les équipes de Struhl et de Green ont également montré que, dans le cas de promoteurs dépendants des TAFs, cette corrélation était aussi valable pour l'occupation des promoteurs par les TAFs (Kuras *et al.*, 2000; Li *et al.*, 2000). Par ailleurs, si l'unique présence de TBP suffit à activer la transcription basale en l'absence de TAFs lors de tests de transcription *in vitro*, dans le cas de promoteurs dépourvus de boîte TATA, les contacts entre TAFs et ADN se révèlent indispensables et semblent jouer un rôle particulièrement important pour la transcription *in vitro* de ces promoteurs. Dans le contexte de TFIID, l'inactivation d'une sous-unité individuelle (un TAF) chez la Levure n'affecte qu'un nombre restreint de



gènes (Moqtaderi *et al.*, 1996; Moqtaderi *et al.*, 1998), ce qui, à première vue, a laissé penser que les TAFs ne jouaient pas un rôle majeur dans l'activation de la transcription.

## 2. Un recrutement spécifique des TAFs

Afin de mieux entrevoir la fonction de TBP et celle des TAFs, des mutants sensibles à la température ont été élaborés. Pour tous les promoteurs testés, qu'ils possèdent ou non une boîte TATA, la transcription décroît rapidement et spécifiquement lorsque l'on inactive TBP (Cormack and Struhl, 1992). Cela suggère que TBP est directement impliqué dans le processus d'activation de la transcription. Cependant, l'inactivation de TBP n'abolit pas la liaison des TAFs aux promoteurs TAFs-dépendants. Le recrutement des TAFs au niveau de ces promoteurs ne repose donc pas sur TBP.



**Figure 69 : Occupation du promoteur par TBP et les TAFs en conditions normales et après un choc thermique (Kuras *et al.*, 2000).**

Afin de savoir si le recrutement de TFIID dépend des activateurs de la transcription, des tests ont été réalisés sur l'occupation des TAFs sur des promoteurs par dont la transcription est induite par des chocs thermiques. Cette occupation du promoteur par les TAFs augmente sensiblement sous l'effet du choc thermique pour les promoteurs TAFs-dépendants et indépendants. Cela signifie que des activateurs agissant lors du choc thermique augmentent le recrutement au promoteur de TFIID (Fig.69). De plus, l'association des TAFs et de TBP sur le promoteur est perdue si ces activateurs sont supprimés, indiquant clairement

que les activateurs sont nécessaires au recrutement des TAFs (Li *et al.*, 2000). Les TAFs seraient donc des cibles directes de certains activateurs, ce qui confirme les études biochimiques préalables (Verrijzer and Tjian, 1996).

Pour les promoteurs TAFs dépendants, les TAFs sont donc recrutés via des activateurs et leur niveau d'expression est comparable à celui de TBP, alors que pour les promoteurs dits TAFs indépendants, les TAFs sont présents à un niveau avoisinant le bruit de fond et seul TBP semble être recruté pour l'activation de la transcription (Li *et al.*, 2000).

### B. *TFIID, une cible idéale pour les activateurs*

Plusieurs caractéristiques de TFIID font de lui un candidat idéal pour être la cible de ces activateurs. Tout d'abord, il possède la capacité de se lier à l'ADN sur des séquences spécifiques de celui-ci. Premier composant du PIC qui s'assemble sur le promoteur, il joue le rôle de point d'ancrage pour le recrutement de la Pol II et des autres GTFs (Pugh, 1996). Ensuite, son interaction avec des activateurs a été démontrée dès les années 80's, notamment avec USF et Gal4 (Horikoshi *et al.*, 1988; Sawadogo and Roeder, 1985b). De plus, la fraction TFIID entière contient des activités additionnelles qui, bien que non essentielles pour la transcription basale, sont indispensables pour la transcription régulée. En effet, si le facteur TBP peut à lui seul remplacer la fraction TFIID lors de la transcription basale (Hahn *et al.*, 1989; Horikoshi *et al.*, 1989), sans les TAFs il échoue pour l'activation de la transcription (Pugh, 1996).

#### 1. *Les liens entre activateurs et TAFs*

Le fait de contacter différents TAFs au sein du TFIID permettrait d'améliorer l'intégration de multiples signaux provenant de distincts activateurs et répresseurs de la transcription (Verrijzer and Tjian, 1996) tout en renforçant la stabilité du PIC (Fig.70). La liaison activateur-TAFs entraîne une augmentation du recrutement du facteur TFIID sur le promoteur. Plusieurs études vont dans ce sens, soulignant que des activateurs peuvent favoriser le recrutement de TFIID, TFIID-TFIIA ou même de TFIIB (Choy and Green, 1993).

De nombreuses interactions entre activateurs et TAFs ont été décrites. Les différentes classes d'activateurs se lient à des TAFs spécifiques au sein du complexe TFIID. Par exemple, chez la Drosophile, les activateurs riches en Gln tels que Sp1 et Bicoid se lient à

TAF4 (Pugh, 1996); les activateurs à domaine acide comme VP16 et p53 se lient à TAF6 et TAF9 (Uesugi *et al.*, 1997; Uesugi and Verdine, 1999) et ceux riches en isoleucine comme NTF-1 se lient à TAF2 et TAF6 (Dymlacht *et al.*, 1991). La stimulation de la transcription par Sp1 ou NTF-1 est dépendante de la présence des TAFs qui lient cet activateur (Chen *et al.*, 1994).

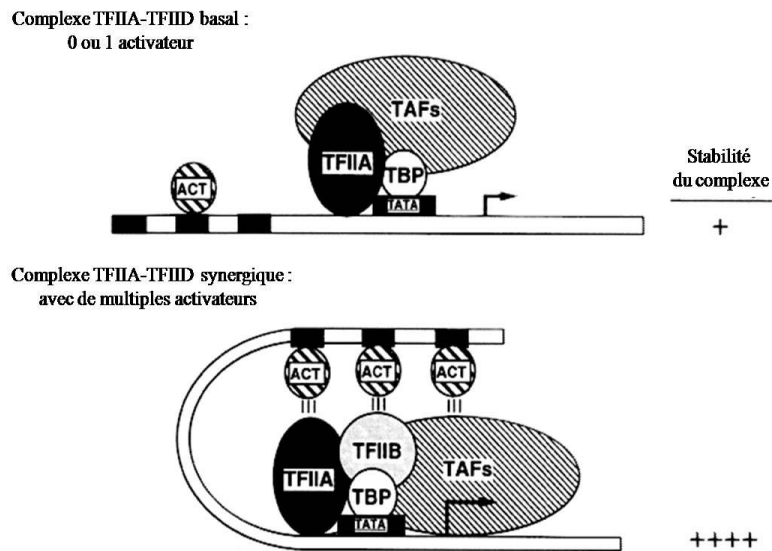


Figure 70 : Les activateurs de la transcription contactent la machinerie basale et stabilisent le PIC.

## 2. Les TAFs, médiateurs de l'activation de la transcription

Les TAFs sont capables d'interagir avec des activateurs de la transcription (Jager *et al.*, 2001a) ; ils contribuent à renforcer la liaison de TFIID au promoteur (Abmayr *et al.*, 1988 ; Workman *et al.*, 1988) et facilitent ainsi la formation du PIC (Albright and Tjian, 2000 ; Hoffmann *et al.*, 1997). Par exemple, TAF9 interagit *in vitro* avec l'activateur acide viral VP16, et TAF10 avec le récepteur aux oestrogènes.

Cependant, ces expériences ne permettent pas de savoir si les TAFs sont recrutées seules ou avec d'autres protéines pour former les complexes de type TFIID ou de type TFIIIC.

*In vivo*, la fonction des TAFs reste plus floue. Dans les cellules de mammifères, des expériences de transfection ont montré que TAF11 et TAF4 sont impliqués dans l'activation de la transcription par les récepteurs nucléaires (May *et al.*, 1996a ; Mengus *et al.*, 1997). Les TAFs semblent également impliqués dans le contrôle du cycle cellulaire. Des mutations dans TAF2 provoquent un arrêt du cycle cellulaire en G2/M (Martin *et al.*, 1999). Des mutations

dans TAF1 provoquent un arrêt du cycle en G1 (Sekiguchi et al., 1991). Les TAFs pourraient jouer des rôles différents tout au long du cycle cellulaire. Or la phosphorylation des TAFs et de TBP, dépendante du cycle cellulaire, pourrait inhiber sélectivement la capacité de TFIID à activer la transcription (Segil et al., 1996). Les TAFs sont très probablement impliqués dans la régulation de la transcription de gènes qui s'expriment en fonction du cycle cellulaire.

Chez la levure, la plupart des délétions des TAFs est létale (Poon et al., 1995; Reese et al., 1994) Ces résultats ont permis de concevoir l'existence de deux groupes de promoteurs. Un premier groupe de promoteurs, nommé dépendants des TAFs, requiert la présence de TAFs pour initier correctement la transcription (Li et al., 2000), tandis qu'un deuxième groupe, nommé indépendants des TAFs ne lie pas les TAFs (ou en quantité faible par rapport à TBP) *in vivo* (expérience d'immunoprécipitation de la chromatine) (Kuras et al., 2000 ; Li et al., 2000). Les promoteurs dépendants ou indépendants des TAFs se distinguent également par leurs séquences activatrices localisées en amont du promoteur, qui ne permettent pas de recruter les TAFs dans le cas des promoteurs indépendants des TAFs (Li et al., 2002).

### 3. Mécanismes d'action des TAFs dans l'activation de la transcription

Diverses théories ont été proposées pour tenter d'expliquer le rôle des TAFs dans l'activation de la transcription.

Il a tout d'abord été suggéré que les TAFs jouent un rôle important dans la transcription *in vitro* de gènes dont les promoteurs sont dépourvus de boîte TATA (Martinez et al., 1994). TBP est capable, seule, de se substituer à l'ensemble du complexe TFIID *in vitro* dans un système basal de reconstitution de la transcription, excepté lorsque les promoteurs sont dépourvus de boîte TATA. Ainsi si les TAFs ne peuvent pas se lier à la boîte TATA, elles seront capables de contacter d'autres séquences d'ADN, comme l'initiateur ou bien les éléments en aval du promoteur. Ainsi les TAFs permettraient la transcription de gènes qui ne contiennent pas de boîte TATA dans leur promoteur.

Ensuite, il a été suggéré que les TAFs soient impliqués dans l'activation de la transcription de gènes inductibles. Lorsque des levures sont soumises à un choc thermique, la proportion de complexe TFIID augmente par rapport à celle de TBP seule (Kuras et al., 2000). Lorsque les TAFs interviennent dans la transcription de gènes inductibles et/ou de gènes dont le promoteur est dépourvu de boîte TATA, il est possible que ces TAFs s'associent

en l'absence de TBP au sein d'un complexe multiprotéique différent de TFIID comme le complexe TFTC.

Enfin, il a été proposé que les TAFs soient impliqués dans la sélection de promoteurs à transcrire. Des expériences de protection à la Dnase I avec des promoteurs distincts suggèrent que les TAFs se lient à différents promoteurs avec des affinités différentes.

Une théorie suppose que les TAFs provoquent un changement de conformation du complexe ADN/TFIID qui lui permettrait de passer de l'état inactif à l'état actif. Par exemple, la protéine TAF1 inhibe les interactions entre l'ADN et TBP puisqu'elle se lie au domaine de liaison à l'ADN de TBP (Liu et al., 1998). Ce modèle propose que TAF1 change de conformation suite à la liaison d'activateurs, libérant alors le domaine de liaison à l'ADN de TBP. TBP pourrait alors se fixer à l'ADN et l'initiation de la transcription démarrer.

#### 4. *Changement de conformation de TFIID*

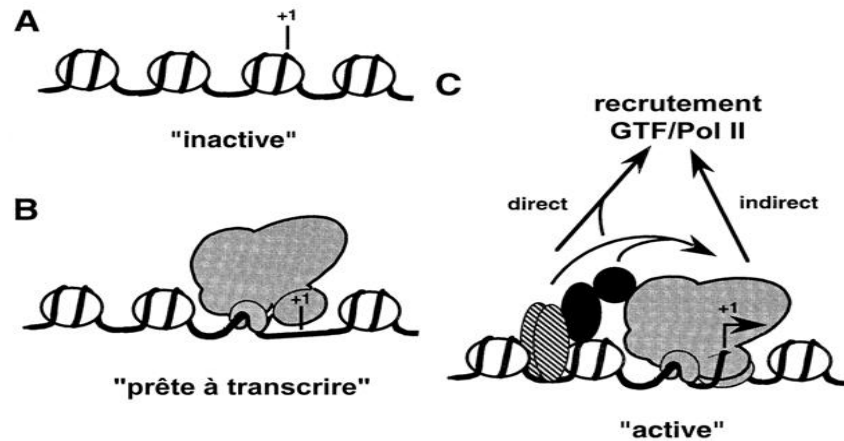
Ainsi le complexe TFIID/ADN pourrait changer de conformation pour passer d'un état inactif en transcription à un état actif (Fig.71). Plusieurs états d'activité du complexe hTFIID pour les gènes de classe II (Hoffmann et al., 1997) ont été décrits :

- 1 – promoteurs ne contenant pas hTFIID, qui sont transcriptionnellement inactifs et qui ne peuvent pas être activés sans remodelage de la chromatine.
- 2 – « poised genes » où hTFIID est présent au niveau du promoteur à un stade inactif. Des stimulus peuvent alors rapidement rendre la transcription de ces gènes active.
- 3 – gènes activement transcrits contenant hTFIID.

Selon ce modèle, la présence de TBP au niveau de promoteurs n'est pas synonyme d'une transcription active. Certaines études *in vivo* d'immunoprécipitation de la chromatine vont dans ce sens puisqu'elles ont pu montrer qu'au niveau de certains promoteurs, TBP est lié à la boîte TATA avant l'intervention d'autres facteurs tels que les TAFs ou la polymérase II (Soutoglou and Talianidis, 2002).

Des expériences d'immunoprécipitation de la chromatine et de fluorescence ont même montré que TBP peut s'associer à la chromatine très condensée en mitose (Christova and Oelgeschlager, 2002). TBP serait présente au niveau de promoteurs avant même que la

transcription ne débute, et servirait ainsi de repère pour la réinitiation de la transcription (Davidson, 2003). D'autres études *in vitro* ont aussi proposé que les TAFs soient impliqués dans la réinitiation de la transcription (Oelgeschlager et al.1998).



**Figure 71 : Différents stades de l'expression des gènes, d'après (Hoffmann et al., 1997)**

A - Les gènes inactifs sont retrouvés au niveau de la chromatine compactée et sont inaccessibles aux facteurs de transcription. Dans ce modèle, l'assemblage du PIC ainsi que l'initiation de la transcription est précédé par l'action de facteurs capables de remodeler la chromatine.

B - TFIID est fixé au niveau du promoteur de gènes « prêts à être transcrits ». L'expression de ces gènes n'est pas active, mais rapidement inductible. La conformation de ce complexe en l'absence de stimuli rend impossible le recrutement de l'ARN polymérase II ainsi que des facteurs généraux de transcription.

C - gènes activement transcrits par la liaison d'activateurs au niveau de séquences spécifiques, ce qui permet le recrutement, direct ou indirect, de la machinerie de transcription.

### C. Les TAFs et la répression de la transcription

Le récepteur alpha de l'hormone thyroïdienne humaine hTR $\alpha$  peut réprimer la transcription lorsqu'il est dépourvu de ligand, en inhibant la formation du PIC sur les promoteurs possédant des éléments de liaison à l'hormone thyroïdienne (TR). L'hormone thyroïdienne alpha peut contacter directement la protéine TBP ce qui a pour conséquence de diminuer *in vitro* la répression engendrée par le récepteur hTR alpha. Le récepteur va se fixer à TBP, soit au cours de la formation du complexe TBP/TFIIA, soit au cours de celle du complexe TBP/TFIIA/TFIIB. Son domaine de liaison interagit avec la partie C-terminale de TBP. Ces interactions directes entre le récepteur à l'hormone thyroïdienne et TBP peuvent être inhibées par la présence de l'hormone. Ces résultats suggèrent donc un modèle dans

lequel hTR $\alpha$  contacte TBP lié au promoteur et interfère avec les dernières étapes de la formation du complexe d'initiation de la transcription (Fondell et al., 1996a ; Hou et al., 2000).

#### D. Modifications de la structure de la chromatine par les TAFs

Outre leur rôle de pont moléculaire entre les activateurs et la machinerie transcriptionnelle de base, les TAFs participeraient aux modifications de la structure de la chromatine pour activer la transcription de certains gènes.

D'emblée, si la structure des TAFs peut mimer celle des nucléosomes, il est facile d'imaginer que TFIID modifie la structure de la chromatine en s'incorporant à la structure en "collier de perles". Or, remplacer les histones par des TAFs peut induire des modifications chimiques locales conduisant à un contexte plus favorable à l'initiation de la transcription.

TAF1 est impliquée dans la régulation de la transcription grâce à ses activités enzymatiques d'acétylation (Mizzen et al., 1996). TAF1 est également douée d'une activité enzymatique de type kinase qui peut être reliée à l'activation de la transcription. En effet, des études chez la Drosophile ont montré que TAF1 peut phosphoryler la sérine 33 de l'histone H2B. Or cette modification se retrouve dans la chromatine active en transcription et dans la chromatine au stade de la mitose, ce qui laisse penser que l'activité kinase de TAF1 serait impliquée dans l'activation de la transcription (Maile et al., 2004).

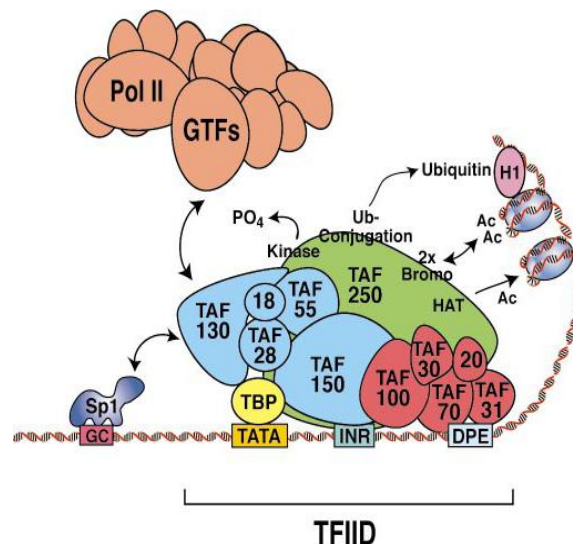


Figure 72 : Les différentes activités enzymatiques de la sous-unité TAF1.

Enfin, TAF1 est douée d'une troisième activité enzymatique, de mono-ubiquitination, ciblant *in vivo* et *in vitro*, l'histone H1 (Pham and Sauer, 2000) (Fig.72). Cette modification sur les histones H2A et H2B est enrichie au niveau des régions de la chromatine active. L'ubiquitination constitue une première marque pour la dégradation des protéines via le protéasome. Si les histones H1 sont dégradées, il est probable que cela permette la décondensation de la structure de la chromatine. Ainsi les activités enzymatiques associées à TAF1 peuvent contribuer à modifier la structure de la chromatine, la rendant plus accessible aux facteurs d'initiation de la transcription.

En résumé, l'action des TAFs semble passer par le recrutement de la machinerie de base via les activateurs, ou par la modification de la structure de la chromatine. La contribution des TAFs comme co-activateurs transcriptionnels semble varier d'un promoteur à l'autre, reflétant en cela la complémentarité ou la redondance des activités co-activatrices présentes dans un extrait nucléaire. Les TAFs et les sous-unités du Médiateur sont nécessaires pour activer la transcription (Guermah et al., 2001). Leurs rôles de co-activateurs a été beaucoup étudié, mais lors d'expériences de transcription *in vitro*, TBP seule s'avère parfois plus efficace en transcription qu'associée aux TAFs. Cette différence pourrait s'expliquer par un effet corépresseur des TAFs (Guermah et al., 2001). De plus, certaines études ont montré que la présence des TAFs peut être utile lors de la répression de la transcription de certains gènes (Fondell et al., 1996a ; (Hou *et al.*, 2000)

### E. Les TAFs, des protéines essentielles

#### a) Pour le cycle cellulaire

L'étude d'un ensemble de mutations sensibles à la température (TS) chez les TAFs indique que plus de 84% des gènes de la Levure ont besoin d'un ou de plusieurs TAFs pour leur expression (Shen *et al.*, 2003).

La mutation tsBN462 de TAF1 dans des cellules de hamster par exemple empêche les cellules de passer de la phase G<sub>1</sub> à S. Une décroissance brutale de la synthèse protéique a lieu, qui commence au milieu de la phase G1 et devient très prononcée à la fin de celle-ci. Placés en conditions restrictives, ces mutants voient l'expression de leurs cyclines A, D1, D3, E réduire, ainsi que l'activité de leur kinase cdk2 associée à l'histone H1 par action de p21. Il



s'en suit une phosphorylation partielle de Rb inhibant alors le cycle cellulaire (Sekiguchi *et al.*, 1996).

Mais ces mutations n'entraînent pas toujours l'arrêt au même stade du cycle cellulaire. Des mutants TS de  $\gamma$ TAF5 présentent en grande majorité un arrêt lors de la transition de la phase G<sub>2</sub> à M. Ces mutations provoquent des dégâts sévères, non seulement sur la viabilité des cellules, mais aussi sur l'intégrité des complexes TFIID et SAGA (lorsque la mutation touche la partie C-terminale de  $\gamma$ TAF5) ainsi que sur la transcription (Durso *et al.*, 2001). Quant aux mutants TS de TAF10, ils présentent des arrêts de synthèse protéique en phase G<sub>1</sub> ou G<sub>2</sub>/M (Kirschner *et al.*, 2002). Des mutations TS ciblant la même région de  $\gamma$ TAF10 peuvent entraîner des modifications distinctes du phénotype de la Levure et de l'expression des gènes.

### 6) Lors de l'embryogenèse

L'inactivation de TAF8 ou de TAF10 chez la Souris est létale pour l'embryon au stade blastocyste (Mohan *et al.*, 2003). Dès le stade E3.75, la condensation de la chromatine est visible dans les cellules de la masse interne de l'embryon taf8<sup>-/-</sup>. Cultivées *in vitro*, ces cellules arrêtent rapidement de proliférer et meurent peu de temps après (Voss *et al.*, 2000). Ces resultants montrent clairement que Taube Nuss (Taf8) est nécessaire à la survie des cellules de la masse interne et que son inactivation entraîne l'apoptose.

L'inactivation de TAF4 chez la souris n'est pas immédiatement létale. Les embryons se développent jusqu'au stade E9.5 avant de mourir. Si TAF4 est inactivé, qui prend le relais ?

### F. TAF4 et la voie TGF $\beta$

Afin de mieux comprendre la fonction de TAF4 dans un contexte *in vivo*, le gène correspondant a été inactivé dans des modèles cellulaires *in vitro* et dans des processus physiologiques *in vivo* chez la souris. Des allèles modifiés, "floxés", ont été générés par recombinaison homologe et des fibroblastes embryonnaires murins (MEFs) taf4<sup>-/-</sup> ont été isolés (Fig.73).

Au contraire des autres TAFs, les fibroblastes embryonnaires murins (MEFs) inactivés pour la protéine TAF4 survivent (Fig.74). TAF4 n'est donc pas essentiel pour la viabilité des MEFs puisqu'il n'entraîne pas un arrêt du cycle cellulaire et n'induit pas l'apoptose.

La viabilité de ces cellules est due à la présence du paralogue de TAF4 appelé TAF4b. TAF4b a tout d'abord été considéré comme spécifique des lymphocytes B (Wu *et al.*, 2005) mais a été observé par la suite dans de nombreuses autres lignées (Freiman *et al.*, 2001). Dans les MEFs sauvages, TAF4b est exprimée mais seule une faible partie est associée à TFIID. Au contraire, en absence de TAF4, TAF4b est incorporé dans le TFIID, assurant son intégrité et la viabilité cellulaire.

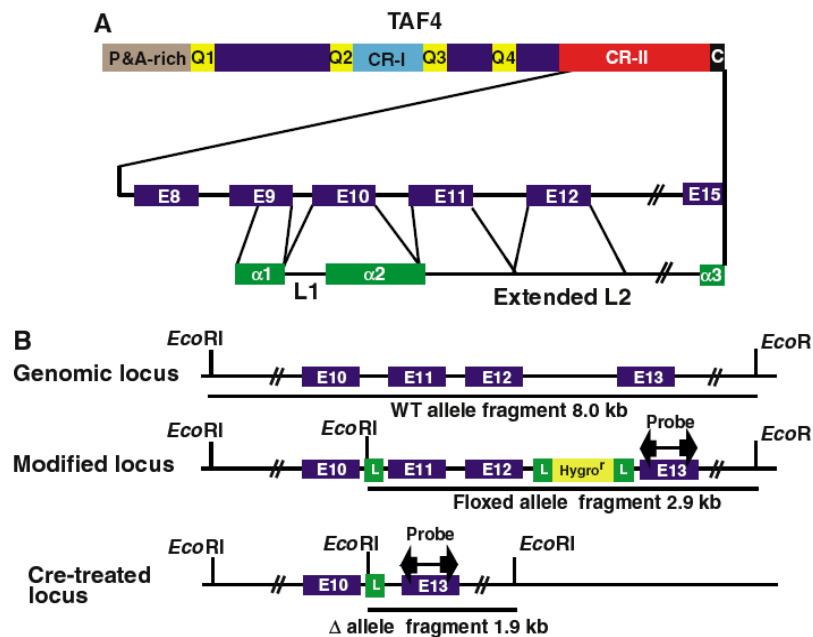


Figure 73 : Génération d'allèles TAF4 porteurs de sites LoxP (Mengus *et al.*, 2005)

L'inactivation de TAF4 entraîne un changement de morphologie par rapport aux MEFs sauvages (Fig.74). Les MEFs dépourvus de TAF4 ont une forme allongée voire étirée et ne présentent plus d'inhibition de contact. Ils sont doués d'une croissance autocrine sérum-indépendante, accompagnée d'une dérégulation de plus de 1000 gènes.

Parmi les gènes les plus affectés, se trouvent beaucoup de facteurs de croissance tels que TGFβ1, TGFβ3 et CTGF. Ce dernier est induit par TGFβ et il a des propriétés de croissance intrinsèque. Il agit comme un médiateur de la réponse proliférative des fibroblastes à TGFβ. Il interagit directement avec TGFβ1 et facilite sa liaison à son récepteur de surface.(Abreu *et al.*, 2002).

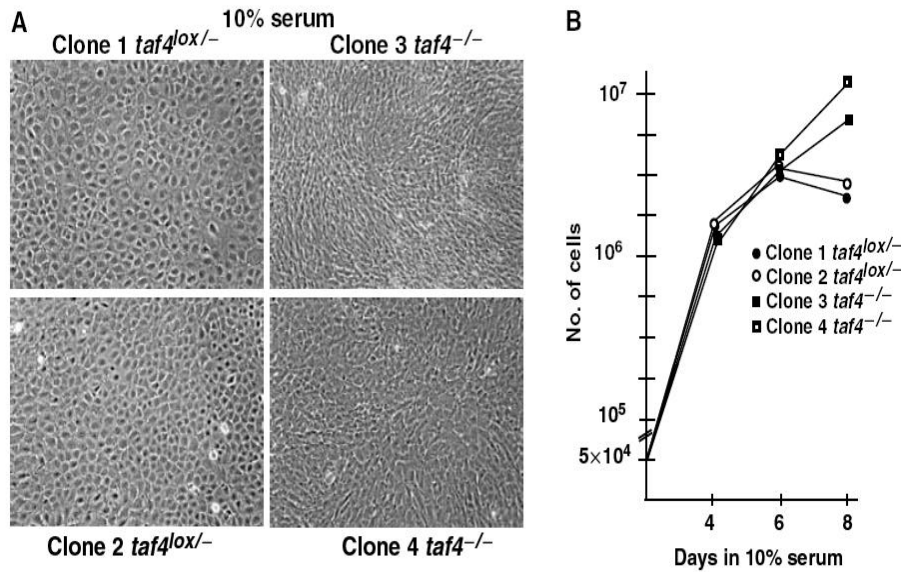


Figure 74 : Croissance des fibroblastes *taf4<sup>-/-</sup>* (Mengus *et al.*, 2005).

De plus, la perte de TAF4 induit les facteurs MMP2 et THBS1 qui sont connus pour activer la voie de signalisation de TGF $\beta$ . En conséquence, les cellules *taf4<sup>-/-</sup>* secrètent fortement du TGF $\beta$  qui active la voie SMAD et induit une boucle positive d'auto-régulation de la voie TGF $\beta$ . La perte de TAF4 dans les MEFs induit l'activation de la voie de signalisation TGF $\beta$ .

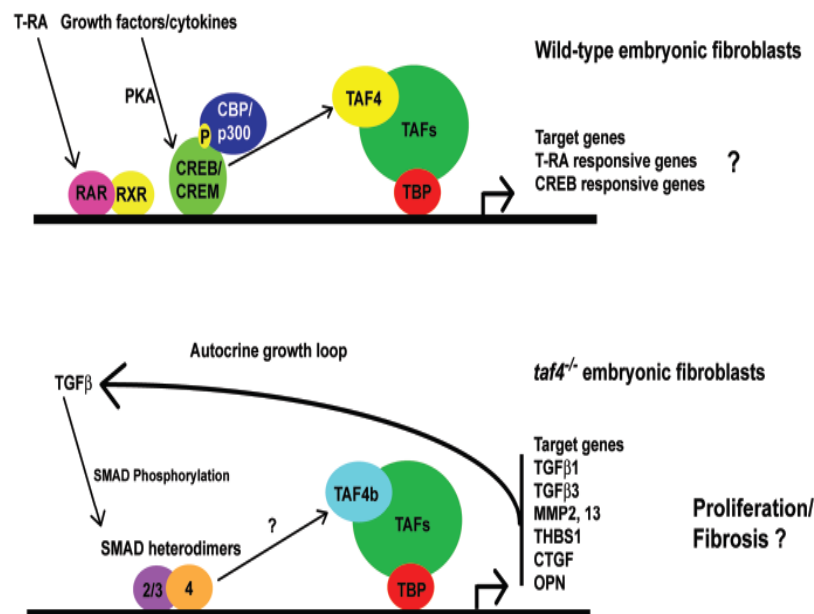


Figure 75 : Induction de la voie de signalisation de TGF $\beta$  lors de l'inactivation de TAF4 dans les MEFs (Davidson *et al.*, 2005).

Les fonctions moléculaires de TAF4b dans la réponse transcriptionnelle à la voie TGF $\beta$  restent encore à élucider. Des études d'interaction protéiques et d'immunoprécipitation de la chromatine seront donc nécessaires pour déterminer si TAF4b interagit avec les SMADs et s'il est recruté au niveau du promoteur de gènes cibles.

La surexpression de TAF4b induit l'expression de TGF $\beta$  et de CTGF même en présence de TAF4, ce qui suggère que TAF4 régule négativement leur expression en entrant en compétition avec TAF4b pour l'intégration du complexe TFIID. TAF4b est cofacteur d'un ou de plusieurs composants de la voie TGF $\beta$ . L'augmentation de la transcription des gènes nécessaires pour cette boucle de croissance autocrine serait induite par le facteur TFIID contenant TAF4b. Un équilibre compétitif entre TAF4 et TAF4b régule alors l'expression des gènes contrôlant la prolifération cellulaire (Fig.75).

### III. TAF4, à la croisée des voies de transcription et de la cancérologie

#### A. Découverte de TAF4

TAF4 est une sous-unité du complexe TFIID. Elle a été isolée initialement chez la *Drosophile* à partir de TFIID (sous son ancien nom : TAF<sub>II</sub>110) (Hoey *et al.*, 1993). Chez l'homme, le clonage de l'ADNc codant pour la sous-unité humaine TAF4 a nécessité le criblage d'une banque d'ADN génomique, du fait de sa teneur extrêmement importante en nucléotides GC (jusqu'à 95% sur 60 paires de bases) (Mengus *et al.*, 1997). La protéine TAF4 de levure a été la dernière orthologue à être identifiée (Reese *et al.*, 2000).

#### B. Structure du facteur de transcription TAF4 (CRI, CRII, domaine riche en glutamine)

hTAF4 est une protéine de 1083 acides-aminés.

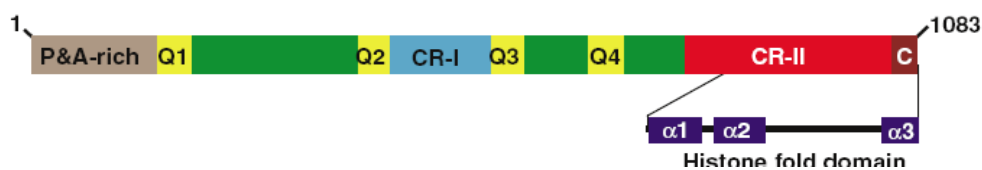


Figure 76 : Structure de la protéine hTAF4 (Mengus *et al.*, 1997).

Elle comprend 4 régions principales : (Fig.76).

- Une région N-terminale riche en proline et en alanine.
- Une région correspondant au premier domaine CRI restreint aux Métazoaires
- 4 régions riches en glutamine
- Une deuxième région conservée CRII

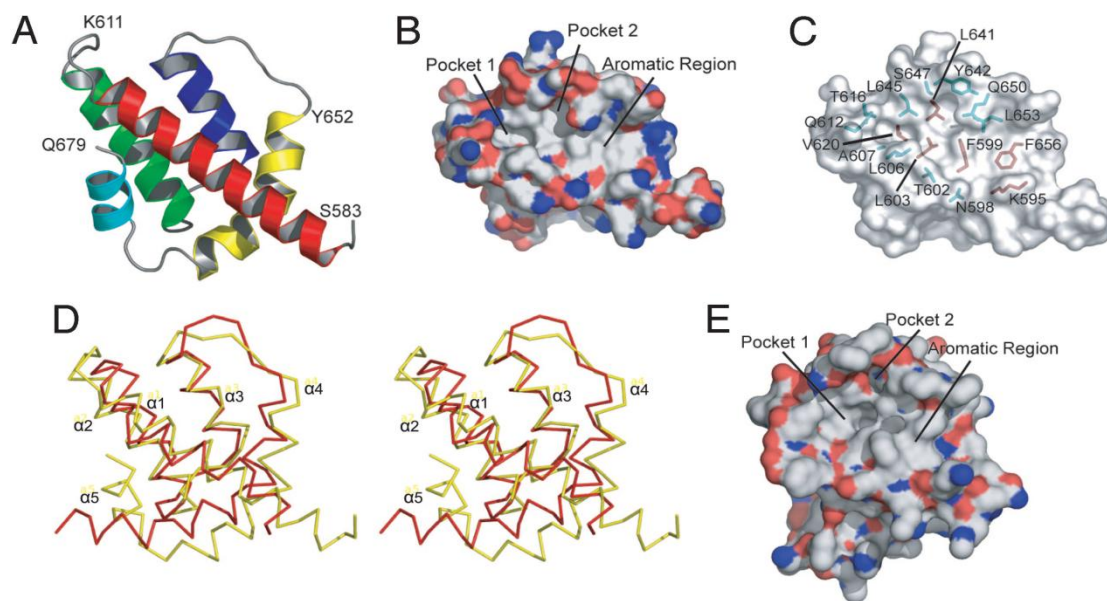
#### 1. Le domaine N-terminal

Le domaine N-terminal de TAF4 est peu conservé (voir les alignements de séquences de TAF4 entre différentes espèces). Des analyses fonctionnelles ont montré que cette région

n'était pas nécessaire à la viabilité des levures *Saccharomyces cerevisiae* (Thuault *et al.*, 2002).

## 2. La région conservée CRI

La fonction de cette région hautement conservée est peu claire. Sa structure 3-D a été élucidée récemment (Wang *et al.*, 2007). Le CRI de TAF4 est aussi connu comme étant le domaine TAFH (TAF Homology Domain). Ce domaine se retrouve chez tous les membres de la famille ETO (Davis *et al.*, 2003), la famille ETO étant composée de protéines impliquées dans la répression transcriptionnelle et exprimées dans la plupart des cellules.



**Figure 77 : Structure globale du domaine hTAF4-TAFH et comparaison avec le domaine ETO-TAFH.**

B) Diagramme en ruban du domaine hTAF4-TAFH (CRI). B) Représentation en surface du domaine hTAF4-TAFH, ayant subi une rotation de 90° le long de l'axe des X par rapport à la représentation en A). C) Représentation des résidus hydrophobes formant la poche du domaine hTAF4-TAFH. Les résidus conservés au sein de tous les domaines TAFH sont représentés en rouge alors que ceux conservés dans les domaines TAFH de TAF4 ou de TAF4B sont représentés en cyan. D) Vue stéréo des traces de C<sup>α</sup> de hTAF4-TAFH (jaune) et de ETO-TAFH (rouge). E) Représentation en surface du domaine ETO-TAFH (Wang *et al.*, 2007).

### a) Le motif hTAF4-TAFH

Etrangement, après alignement des domaines TAFH, la proline P669 qui permet la formation de l'hélice 5 n'a été détectée que dans les sous-unités TAF4 et TAF4b des

Vertébrés. Par homologie avec les RAR, elle pourrait être la cible de phosphorylations éventuelles, à la base d'une réponse transcriptionnelle fine, déterminante pour la structure des Métazoaires.

Le domaine hTAF4-TAFH est constitué de **5 hélices  $\alpha$**  (Fig.77) qui forment une pince contenant 3 éléments principaux : 2 poches et une région aromatique (Fig.77). Une large cannelure dépourvue de liaison hydrogène donneur ou accepteur est présente entre les hélices 1 et 4. Cette cannelure est formée de résidus hydrophobes des hélices 1, 2, 3, 4. Bien que les 5 hélices ne participent pas toutes à la formation de ce canal hydrophobe, elles jouent un rôle important dans l'empaquetage général de ce noyau hydrophobe.

La plupart des résidus formant la pince sont hautement conservés dans le domaine TAFH de TAF4, TAF4b et des membres de la famille ETO. Cette surface d'interaction permet la reconnaissance d'un **motif à 4 résidus hydrophobes de type  $\Phi\Phi\sigma\sigma\Phi\Phi$**  où  $\Phi$  représente un résidu hydrophobe et  $\sigma$  un autre résidu quelconque. Ce motif de reconnaissance est présent au sein de nombreux régulateurs transcriptionnels qui pourraient être des co-activateurs ou des corépresseurs capables de se lier à TAF4 et dont beaucoup auraient un rôle dans le développement des Métazoaires.

#### *6) Les protéines interagissant avec ce motif*

**NCoR** interagit *in vitro* avec ETO-TAFH (Amann *et al.*, 2001; Lausen *et al.*, 2006) et le réprime dans les essais de gènes rapporteurs à la luciférase. L'un des deux domaines d'activations de la boîte E des facteurs de transcription HEB et E2A interagit aussi avec ce CRI (Wei *et al.*, 2007; Zhang *et al.*, 2004). Les deux protéines **LZIP** et **ZF** contactent également la protéine TAF4 via leurs motifs **DLLxxLL**, ce qui suggère qu'un complexe co-activateur contenant TAF4 et HCF-1 pourrait se former lors de l'activation transcriptionnelle de ZF et LZIP ou de Sp1.

Parmi la liste des interactants potentiels se trouvent un grand nombre d'enzymes impliquées dans des MPT ou dans le remodelage de la chromatine comme par exemple **CRSP130**, mais plusieurs HDAC, méthylases et kinases impliquées dans le contrôle du cycle cellulaire. La région d'homologie des TAFs pourrait donc intervenir dans la régulation de la transcription de manière positive ou négative.

### c) Un motif spécifique des Métazoaires

Par contre, ce domaine  $\Phi\Phi\sigma\sigma\Phi\Phi$  ne semble pas être nécessaire à l'assemblage de TFIID contrairement à CRII (Furukawa and Tanese, 2000). Dans l'étude de Wright, il semble que la région CRII soit suffisante pour permettre l'élaboration d'un sous-complexe de TFIID ou de TFIID en entier. De plus, chez la levure, il existe un complexe TFIID mais pas de domaine TAFH mais ce qui irait donc à l'encontre d'un éventuel rôle de TAFH dans l'assemblage de TFIID.

#### 3. La région conservée CRII



La région C-terminale CRII est une région très conservée au cours de l'évolution, au sein de la famille de TAF4. Elle est essentielle pour l'incorporation de TAF4 dans le TFIID et permet son interaction avec d'autres protéines telles que CREM, TFIIA ou E1A.

La technique du double-hybride chez la levure a permis de montrer que cette région CRII contient un HFD qui permet l'hétérodimérisation de TAF4 avec TAF12 (Gangloff *et al.*, 2000). L'association réalisée par les HFD est déterminante pour la liaison à une séquence spécifique du promoteur. Les membres de la famille TAF4 possèdent un motif atypique de repliement des histones (voir comparaison de séquences de TAF4) constitué d'un espace long entre la deuxième et la troisième hélice et d'une troisième hélice  $\alpha 3$  située au sein d'un domaine C-terminal très conservé (Thuault *et al.*, 2002).

Enfin, la liaison de TAF4b à l'ADN requerrait ce motif structural et une partie du domaine d'espacement entre les deux dernières hélices. Et contrairement à l'histone H2A (Luger *et al.*, 1997), c'est la deuxième hélice et non pas la première qui s'attacherait à l'ADN. Cela expliquerait pourquoi les résidus d'interaction avec l'ADN de l'histone H2A n'ont pas été préservés au sein de la famille de TAF4 (Shao *et al.*, 2005).

Le CRII est important pour l'interaction avec des régions riches en polyglutamines comme celle du domaine Q2 de CREB. Il est également nécessaire pour la liaison de TAF4 à TFIIA (Guermah *et al.*, 2001) et à E1a, qui nécessite de surcroît le domaine CCTD



hsTAF4 1 MAAGSDLDEVFNSEVDEKVVSDLVCSLESOLAASAAHHHHLAPRTPEVRAAAAGATGNHVVSGSFAAAGAFAAFAECAFCAAFEPFPAGRA  
 mmTAF4 1 MAAGSDLDEVFNSEVDEKVVSDLVCSLESOLAASAAHHHHLAPRTPEVRAAAAGATGNHVVSGSFAAAGAFAAASAECAFPAAFEPFPAGRA  
 hsTAF4b 1 .....MVFOLKCSCEPFIYISGOHLVSV EEDARSEDEFEEDVKLLSISGKOSALVEVAFEFHKKKOLITAFIANDKVOEDKONLEFWETSIVYCBT  
 mmTAF4b .....  
 dmTAF4 .....  
 ceTAF4 .....  
 scTAF4 .....  
 ecTAF4 .....  
 dmNHT .....

hsTAF4 96 RFGGGGFORPFPSPRRRLVPA GPAPPAKLRPPPEGSAGSCAPV PAAAAVAA GPPEPAFAGPAKPA GPAALAARA GFGEGPFP GEGPFGPCKPA  
 mmTAF4 96 RFGGGGFRREGS SSSERRLLVPA GPAPPAKLRLEPEASAG... PCEFAAAAAATAAAVAGPD... PAAAAAPAGFG... ASKFGPFGP  
 hsTAF4b 90 DSFFTIHKADEITGVIEKKNKFLPSPVTSRRLRSLAN..... HAACSRGNEGRWRREGRVREVSAAWIKHCKWR..... GRAACPCOSR  
 mmTAF4b .....  
 dmTAF4 1 ..... NTSOTAAGNRITFTSOEIPNGTISIAG  
 ceTAF4 .....  
 scTAF4 .....  
 ecTAF4 .....  
 dmNHT .....

hsTAF4 191 GFGAAQILNGSAALLNSHHAAPAVSLVNNGPAALLPLPKPAAGCTVITQTPPEVCAAAPPAPAAPSPPAAPAPAAAPAAAA PPPPPAPATLARPPG  
 mmTAF4 177 GFGAAQILNGSAAPPATPHAAFAVSEVNNGPAP.....PPAAGSVIQ..... AAPPAPAPSPPAAPAPAAFP... PPPPAPPAPVAFARPPG  
 hsTAF4b 167 VGGCGHS GEGSREGREECEGRACPAVORA..... PVTDFAAAAPAPVPSDAVVEPEEHSRRIEPRAKPPT  
 mmTAF4b .....  
 dmTAF4 29 NFGAVISTAOLEN TTTIKTIQAGLGGHOGILOVHHVQOQOQSQOQOQOQOQITQSAQGFLLNSMLPAGVVVGRROCAFQOQOQKNVPEINELSRVV  
 ceTAF4 .....  
 scTAF4 .....  
 ecTAF4 .....  
 dmNHT .....

hsTAF4 286 HPAG...PPTAAFAVPPPAAANNGSAGAAPAPAPAAAGFPAGVSGQFGGAAAAAPAGVKAESKRVYCAAPPAAQTILAASGFASTAASMVIGPT  
 mmTAF4 254 HPAAFAFPVSPAAAAAANAANCAAPAPAAASGCAAGT...GQFG...AGACGAKAFAKKEGFGEGPCEFAAAAAAGVAGGCAAGSVLLGPA  
 hsTAF4b 229 TSAPGTAAPKLELSPFGMPAGLIEPAGAAPPAAVSASG.....TVMAPAGALPVVVESTPVALGAVIKAPVSVQVPEPT  
 mmTAF4b 1 .....MPAGLIEPAGAAPPAAVSASG.....AVTMAPVAALPVVVEGTEVAIGAVIKAPVSVQVPEPT  
 dmTAF4 124 INSHMACVREQSPSITLSTLNTICOTPALLVKTDNCFOLLRVCTTTCPPTVTITITNTSNNSNTTSTTNEPTTQIRLIQIVPAAAAMTNTTATSNI  
 ceTAF4 .....  
 scTAF4 .....  
 ecTAF4 .....  
 dmNHT .....

Q1

hsTAF4 379 NCGALPSEAAVPPPAEGTPTGTEKGAAGAVTOSLSRIFETATTSGLRATLITPTVLAERIEPOPPNPNTINONFOIPPGMVLVRSENGOLLMIPQOAL  
 mmTAF4 344 NCGALEG.....PAECTPAGLAKGAAATATSLPRTPAATTGATRATLITPTVLAERIEPOPPNPNTINONFOIPPGMVLVRSENGOLLMIPQOAL  
 hsTAF4b 303 ASOPLRS.....PVGLVTVKVPVPSAPPKVVSSGPRIPAPQIVAVKAFENTTTIIOFP.....ANLOLPPCTVLIKSNSGFLMLVSPQ...  
 mmTAF4b 58 AFOPLRA.....PVGLVTVKVPVPTALEKLGSPRLIPAPQIVAVKAFEGTTTIIOLE.....ANLOLPPCTVLIKSNSGOLMLVSPQ...  
 dmTAF4 219 LVNSVASSGYANSSQPPHITQNAQAEQLEQITQITPAAQQSQOQOQNNVSSAGGTATAVSSSTAATTTOGNTKEKQRKFLANLIEHSTREBK  
 ceTAF4 1 .....MGLATTESTVITLGSIKGQHFPLFSAHMSIERRFLVQG.....  
 scTAF4 .....  
 ecTAF4 .....  
 dmNHT .....

Q1

Q2

hsTAF4 474 AQMOAQAHAQPEOTMAERPATPTGAPPVOISTVQAPGTPPIIAR...OVITPTTIKQVSOAQTTVQPSA...TLORSECVQPOLVVLGG.....AAQ  
 mmTAF4 433 AQMOAQAHAQPEOSTMAERPATPTGAPPVOISTVQAPGTPPIIAR...OVITPTTIKQVSOAQTTVQPTT...TLORSECVQPOLVVLGG.....SAQ  
 hsTAF4b 378 ...OTVTRAEFTSNITSRPAVPANEQIVKICITVPNSSSOLIKKVAVTEPVKKLAQIGTTVVTTVEKPS...SSVQSVAVPTSVVTVTEKPELNTVITLK  
 mmTAF4b 132 ...OAVTGAKTTSNITSRPAVPANTQIVKICITVPNSSSOLMKKVTVAVPKNLIQIGTTVATTASST...SSGQPVVALPSSVITVTEPAKLVNTVSTLK  
 dmTAF4 314 FVEKNVRLIOQLVNAVVEFEFCDRIERLINASEPCLIGFUKKSIFLLROALYTKELVIEGFKPPPQHVILGLACISQQLLEKIQAOIRPFGESQ  
 ceTAF4 38 ...KALGERSTEGVSTREPAPEQIKQFVDYODAHOMAEFEVEAEQANHOYQPPROPIQOQVQHFQ.....  
 scTAF4 .....  
 ecTAF4 .....  
 dmNHT .....

CRI

hsTAF4 558 TASIQTATAVQIGTEPORTVEGATTTSSAAETMENVKKCKNFTSLIKLASSKOSTETAANVKELVONLLDCKIEAEFDTSRIYREINSSPOFY  
 mmTAF4 516 TASIQTATAVQIGTEPORTVEGASTTSSAAETMENVKKCKNFTSLIKLASSKOSTETAANVKDLVONLLDCKIEAEFDTSRIYREINSSPOFY  
 hsTAF4b 469 PSSLGASSTPSNEPILKAENSAAVQINLSPTIMLENVKKCKNFTLAMLIKLACSGSOSPEMGONVKKLVEOLLDAKIEAEFEFTRKLYVEIKSSPOPH  
 mmTAF4b 223 SSSIQLVSTESNDARLKAETSVAAGQATAPPVLENVKCKNFTSMLIKLACSGSOSPEMGONVKKRLVEOLLDAKIEAEFEFTRKLYTELKSAPOPH  
 dmTAF4 409 TTTIGTQVVRMITHENAICTERTHIGHTHISKQPPNIRLPT...APRLVNTGGIRTOIPSLQVEGOANIVQIRCEQHAQIQRTCSVQIRATRPPNS  
 ceTAF4 101 .....SESPMAEGPPECIPONSAAAAAASDDKNVTKQVRFKILINLSNNDDEEMPKAARVKELEIRGVIVYETTAEEFFTRNLQOVLKSOAOPH  
 scTAF4 .....  
 ecTAF4 .....  
 dmNHT .....

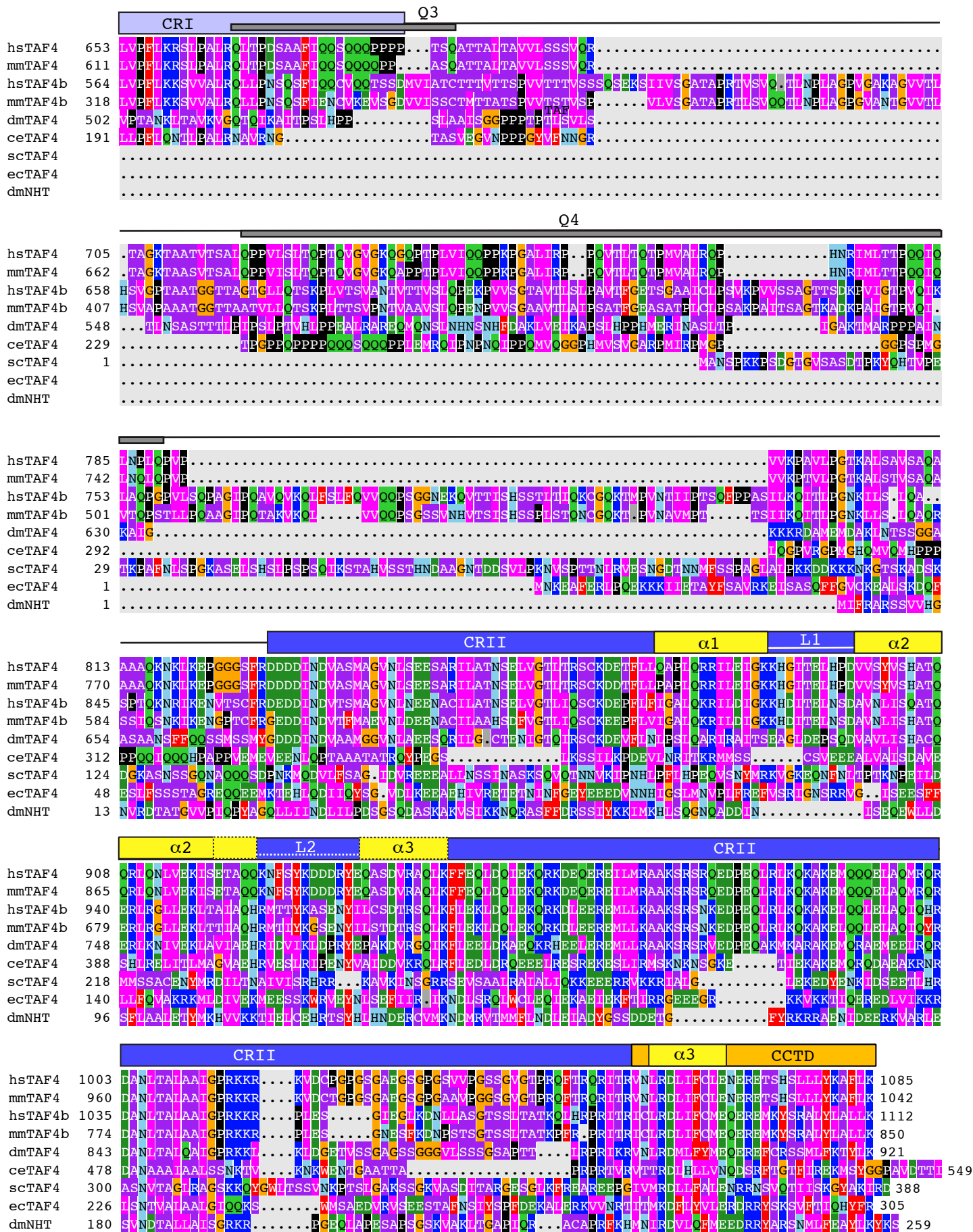


Figure 3 :

Alignement des séquences peptidiques des TAF4s humain (hs, *Homo sapiens*), de souris (mm, *Mus musculus*), de drosophile (dm, *Drosophila melanogaster*), de nématode (ce, *Caenorhabditis elegans*), de levure (sc, *Saccharomyces cerevisiae*), d'*Encephalitozoon cuniculi* (ec), des TAF4bs humain et de souris et de NoHitter (NHT) de drosophile.

Les 4 domaines riches en glutamines sont indiqués Q1 à Q4 en gris. Les régions conservées 1 et 2 (CRI et CRII) sont indiquées en bleu. La localisation des hélices  $\alpha$  du HFD est indiquée en jaune. Les boucles qui les relient sont notées L en blanc. L'hélice  $\alpha 3$  putative est entourée de pointillés. Le HFD de TAF4 est atypique : l'hélice  $\alpha 3$  se trouve en fait dans la région carboxy-terminale conservée (CCTD) représentée en orange, séparée de l'hélice  $\alpha 2$  par une très longue boucle.

Merci à Christophe Romier pour la réalisation de cette figure.

#### 4. Les domaines centraux riches en glutamine

TAF4 possède 4 domaines riches en glutamines identifiés de Q1 à Q4 (Fig.65). Ces domaines qui contiennent entre 20 et 30% de glutamine, sont impliqués dans les interactions avec les activateurs transcriptionnels Sp1 et CREB (Asahara *et al.*, 2001; Gangloff *et al.*, 2000).. Chez la drosophile, les régions polyQ interagissent avec l'homéodomaine de la protéine Bicoïd, impliquée dans le développement précoce de l'embryon de drosophile. Cependant, si cette interaction entre dTAF4 et Bcd aide l'activation transcriptionnelle observée *in vitro*, elle ne semble pas essentielle pour l'embryogenèse de la drosophile (Schaeffer *et al.*, 1999).

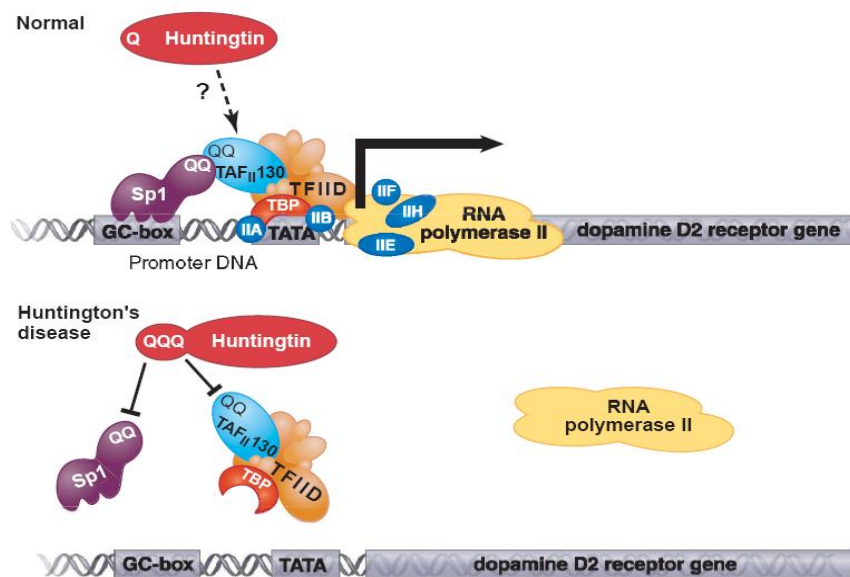
#### C. TAF4 et les maladies neurodégénératives

8 maladies neurodégénératives incluant la maladie de Huntington (HD) et plusieurs formes de l'ataxie spinocérébelleuse sont actuellement connues pour être causées par des répétitions de triplets CAG codant pour la polyglutamine (Ross, 1997).

HD est une maladie neurodégénérative héritée caractérisée par des troubles moteurs progressifs ainsi que par des déficiences cognitives qui s'aggravent au fil du temps. Ces troubles mènent à la mort de l'individu. Des années de recherche se sont passées avant que la forme mutante de la protéine Huntingtine (Hi) ne soit identifiée comme étant la cause de HD (par The Huntington's Disease Collaborative Research). L'expansion de répétitions du trinuécléotide CAG dans le gène HD résulte en une forme mutante et allongée de la protéine Huntingtine. L'âge auquel se déclare la maladie varie avec la longueur des expansions glutamine. Bien que l'augmentation des répétitions soit une caractéristique de nombreuses maladies, la fonction et le rôle exacte de la Hi normale dans la cellule son toujours méconnues. L'hypothèse émise par le groupe de Dunah est que les expansions polyglutamines de la Hi mutante tendraient à perturber la transcription normale dans les neurones du cerveau humain (Dunah *et al.*, 2002). En voici le mécanisme supputé :

Ordinairement, le facteur de transcription Sp1 se fixe sur des éléments de séquences riches en GC. Il s'associe à TAF4 par l'intermédiaire de ses domaines riches en polyglutamines. Cette sous-unité de TFIID est nécessaire au recrutement de la machinerie générale de transcription. Cette interface de glutamines sert de pont à Sp1 pour recruter la Pol

II via les GTFs. La transcription peut alors commencer dès lors que tous les acteurs sont réunis au niveau du promoteur du gène de la dopamine D2 (Fig.79).



**Figure 79 : Perturbation de la transcription par un mutant de la protéine huntingtine (Freiman and Tjian, 2002)**

Dans la HD, l'expansion glutamine de la Huntingtine perturbe l'activation transcriptionnelle ordinaire de la dopamine D2. Un fragment amino-terminal de la huntingtine contenant de nombreux résidus de glutamine s'accumule dans le noyau. La protéine mutante s'associe alors à Sp1 et à TAF4 et empêche le contact entre Sp1 et la boîte GC. De manière ultime, le mutant Hi dissocie même la liaison entre Sp1 et TAF4. Les GTFs ne sont alors plus recrutés sur le promoteur de la dopamine D2 et par conséquent, la Pol II ne peut pas s'y fixer. Il n'y a donc plus de transcription possible du gène de la dopamine D2. Cette dérégulation génique constitue l'un des premiers pas du processus neurodégénératif qui se met en place dans le cerveau des malades HD (Fig. 79).

TAF4 ainsi séquestrée par le mutant Hi ne peut alors plus participer à l'activation transcriptionnelle de la voie de CREB. Il en résulte un dérèglement de cette voie (Shimohata *et al.*, 2000).

#### D. TAF4 renforce l'activation transcriptionnelle

TAF4 est le premier des TAFs pour lequel une activité co-actriatrice a été découverte (Chen *et al.*, 1994).

a) De Sp1 ?

C'est en 1993 que Hoey a pressenti le premier les fonctions co-activatrices de TAF4. La présence de domaines riches en glutamine et en sérine/thréonine au sein de la structure de TAF4 rappelle en effet les domaines d'interaction protéine-protéine de Sp1 impliqués dans l'activation et dans la multimérisation. Chez la levure comme chez la drosophile, TAF4 interagit avec les domaines riches en glutamine de Sp1. De plus, Sp1 purifié lie le recombinant de TAF4 *in vitro*. L'ensemble de ces résultats suggérait donc que TAF4 puisse fonctionner comme un co-activateur en servant de site de contact entre le TFIID et des activateurs comme Sp1 (Hoey *et al.*, 1993).

La surexpression des domaines de TAF4 riches en polyglutamine lors d'expériences de transfections transitoires inhibent l'activation de Sp1 (Gill *et al.*, 1994). Au contraire, dans les fibroblastes embryonnaires murins (MEFs), l'activation de Sp1 est identique dans les cellules taf4<sup>-/-</sup> et dans les cellules sauvages. TAF4 ne semble donc pas essentiel à la fonction de Sp1 dans ces cellules.

b) De CREB et CREM

Des essais d'expression transitoire *in vitro* ont suggéré que TAF4 pourrait être un co-activateur pour les facteurs de transcription CREB et RAR (Asahara *et al.*, 2001; Mengus *et al.*, 1997).

CREB et CREM sont des facteurs qui activent la transcription en réponse à une élévation de la concentration en AMPc dans le milieu intracellulaire. La protéine kinase A (PKA) phosphoryle CREB sur la sérine S133 de son domaine KID. Cela permet le recrutement du co-activateur p300/CBP (Chrivia *et al.*, 1993). CREB possède un domaine d'activation constitutive riche en glutamine appelé Q2 qui interagit avec TAF4 via son domaine CRII contenant le HFD (Asahara *et al.*, 2001). CREM possède également un domaine Q2. Dans les MEFs, l'activité transcriptionnelle de CREB et CREM induite par la PKA est fortement diminuée dans les cellules taf4<sup>-/-</sup>. TAF4 est donc un co-activateur important coopérant avec CBP/p300 pour améliorer la fonction de CREB (Mengus *et al.*, 2005). En outre, la fonction de CREB peut être restaurée dans les cellules taf4<sup>-/-</sup> exprimant uniquement le domaine CRII, ce qui conforte ce domaine comme étant le site prioritaire d'interaction avec CREB.

c) Des récepteurs nucléaires

La surexpression de hTAF4 induit la transactivation dépendante des ligands pour les récepteurs nucléaires (RN) à l'acide rétinoïque, à l'hormone thyroïdienne et à la vitamine D (Gangloff *et al.*, 2000; Mengus *et al.*, 1997). Ainsi, TAF4 joue le rôle de co-activateur pour les RN de classe II.

Chez la drosophile par contre, TAF4 n'a pas d'effet spécifique sur l'activité de RAR, ce qui laisse suggérer que les TAFs chez les Vertébrés ont évolué afin d'acquérir des fonctions spécifiques de co-activation de la transcription. Bien que TAF4 agisse comme co-activateur avec le domaine AF-2 de RAR (Mengus *et al.*, 1997), aucune interaction directe n'a pu être détectée entre eux. Dans cette optique, il est tentant de spéculer que hTAF4 puisse interagir avec les TIFs (transcriptional intermediary factors) spécifiques de RAR, VDR et TR.

Quelques années plus tard, l'activation dépendante du ligand par RAR a pu être testée dans les cellules taf4<sup>-/-</sup>. Elle est abolie et peut être restaurée dans ces mêmes cellules exprimant le CRII.

Par contre pour VDR, l'activation dépendante du ligand n'est pas affectée par la perte de TAF4 (Mengus *et al.*, 2005). Comme pour Sp1, les expériences utilisant une construction artificielle de gène rapporteur ne peuvent pas exclure un effet cellule-spécifique ou promoteur-spécifique.

d) Des oncoprotéines

TAF4 interagit également avec des oncoprotéines tels que le papilloma virus E7, l'antigène T de SV40 et l'adénovirus E1A (Geisberg *et al.*, 1995; Mazzei *et al.*, 1995). Toutes ces oncoprotéines ont des séquences homologues et peuvent interférer sur la croissance cellulaire en interagissant avec des membres de la famille du suppresseur de tumeur Rb (Retinoblastoma) (DeCaprio *et al.*, 1988).

Le facteur nucléaire des cellules T activées (NFAT) a un effet pivot sur la régulation de la transcription des gènes codant pour des cytokines immuno-modulaires. Le mécanisme par lequel il induit l'activation de la transcription passera par une interaction directe avec hTAF4 (Rao *et al.*, 1997). Ces événements pourraient induire le recrutement ou la stabilisation de TFIID au niveau des promoteurs.

La protéine p45NF-E2 joue un rôle essentiel pour l'expression à haut niveau du gène de la globine dans les cellules érythroïdes (Lu *et al.*, 1994). Elle interagit directement avec hTAF4 suggérant un mécanisme possible d'activation transcriptionnelle à longue distance (Amrolia *et al.*, 1997).

### E. TAF4 intervient dans la répression de la transcription

TAF4 aurait également un rôle dans la répression de la transcription.

#### a) HP1

La famille HP1 est composée de 3 isoformes chez les Mammifères :  $\alpha$ ,  $\beta$ ,  $\gamma$  qui se distinguent par leur localisation différente au sein du noyau. HP1 $\alpha$  et  $\beta$  sont localisées préférentiellement au niveau de l'hétérochromatine alors que HP1 $\gamma$  est localisée à la fois dans l'hétérochromatine et dans l'euchromatine (Minc *et al.*, 2000). Les protéines HP1 jouent un rôle dans le « silencing » de l'hétérochromatine mais aussi dans la répression des gènes de l'euchromatine (Jones *et al.*, 2000). *In vitro*, des interactions ont été montrées entre hTAF4 et HP1 $\alpha$  et  $\gamma$  mais pas avec HP1 $\beta$  (Vassallo and Tanese, 2002).

Dans les cellules de Mammifère et chez la levure, le chromodomaine Swi6 de HP1 se lie à l'histone H3 méthylé par la méthyltransferase SUV39H1 au niveau de sa Lysine 9. De son côté, Rb cible la protéine SUV39H1 sur l'histone H3 et recrute HP1. Cela entraîne une répression transcriptionnelle du promoteur de la cycline E endogène. *In vitro*, des essais de transcription reconstitués avec de la chromatine et des extraits nucléaires ont montré que la répression de la transcription dépendant de HP1 se fait sur une matrice méthylée par SUV39H1. hTAF4 interagirait donc avec HP1 $\alpha$  et  $\gamma$  pour réguler la répression de gènes cibles. Le mécanisme complet reste encore à élucider.

#### b) Le cas d'ATF7

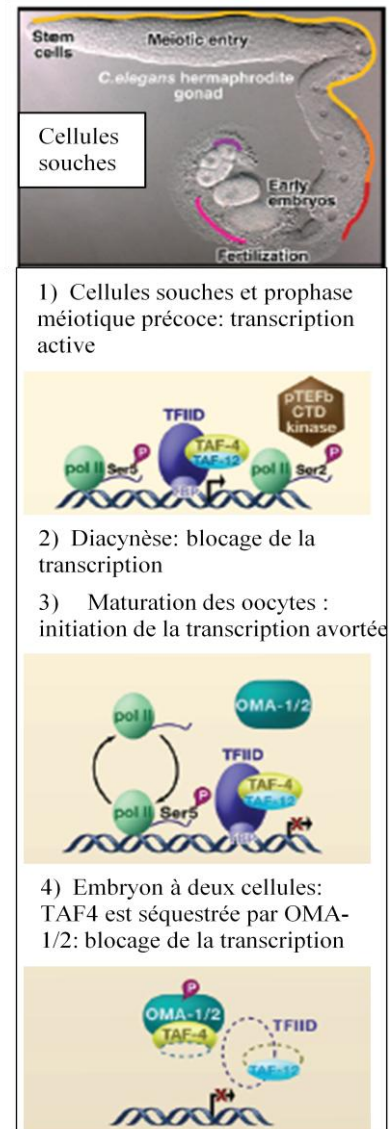
Le facteur de transcription ATF7 est un membre de la famille ATF7/CREB. Il forme des homodimères et se lie à la même séquence que CREB : au CRE. Il peut aussi former des hétérodimères avec les membres de la famille Jun et Fos et se lier à des TRE (Chatton *et al.*, 1994). ATF7 interagit avec les kinases de stress JNK2 (Bocco *et al.*, 1996), avec la tyrosine phosphatase PRL1 et elle est impliquée dans le contrôle de la croissance et de E1A. Son

activité transcriptionnelle est induite par l'expression de TAF12, mais pas par celle de son hétérodimère TAF4. ATF7 et TAF12 s'associent sur le promoteur du gène de l'E-sélectine.

TAF4 s'associe aussi à ATF7 mais sa surexpression entraîne l'inhibition de l'effet co-activateur de TAF12. TAF4 serait donc un régulateur négatif d'ATF7 dont l'effet pourrait être supprimé en surexprimant la protéine TAF12. Ainsi pour deux protéines CREB et ATF7 liant une même séquence, TAF4 agit comme un cofacteur positif dans le cas de CREB et comme un cofacteur négatif dans le cas d'ATF7.

### c) OMA-1

Un mystère a longtemps plané sur la manière dont la transcription est rendue silencieuse dans les embryons précoces. Lors de leur entrée en méiose, les oocytes de l'animal hermaphrodite *C. elegans* sont transcriptionnellement actifs et cette transcription est réalisée par l'intermédiaire du facteur TFIID. TAF4 est maintenu dans le noyau grâce à son interaction avec TAF12, un autre composant de TFIID (Fig.80(2)). Lors de la diacynèse, le dernier stade de la prophase méiotique, l'activité transcriptionnelle est bloquée. Les protéines OMA sont phosphorylées par la kinase MBK-2 qui ciblent de nombreuses protéines maternelles en vue de leur dégradation pendant la transition de la mère au zygote (Stitzel *et al.*, 2006). Les protéines à doigt de zinc OMA-1 et OMA-2 phosphorylées vont se lier à TAF4 et l'inhiber en le séquestrant dans le cytoplasme (Fig.80(4)) (Blackwell and Walker, 2008).



**Figure 80 : Répression de la transcription via la séquestration de TAF4 (Blackwell and Walker, 2008).**



Cette inhibition est rendue possible grâce à un mécanisme surprenant de mimétisme moléculaire de l'HFD de TAF12 par un domaine des protéines OMA. Ainsi la transcription est réduite à néant lors de ces stades précoces par un mécanisme de séquestrement de TAF4 qui n'est alors plus disponible pour l'initiation de la transcription.

TAF4 est un composant de TFIID/TFTC dont le rôle de co-activateur transcriptionnel pour RAR et CREB a été démontré ainsi que celui de répresseur de la voie TGF- $\beta$  et de la prolifération cellulaire dans les MEFs.

#### *F. TAF4 et la cancérogenèse*

L'inactivation de TAF4 dans les kératinocytes basaux de l'épiderme du fœtus se traduit par une diminution de l'expression des gènes nécessaires à la fonction de la peau en tant que barrière immunitaire. Elle se traduit par une mort néonatale. Au contraire, chez l'adulte, elle mène à la perte extensive des poils de la souris due à un cycle de renouvellement du poil aberrant caractérisé par une phase anagène défectueuse.

Alors que l'épiderme mutant contient très peu de poils en phase anagène, de nombreux gènes exprimés normalement au cours de cette phase sont surexprimés dans l'épiderme mutant de manière très forte, ce qui traduit une désynchronisation de l'expression des gènes. L'épiderme TAF4 mutant chez les adultes montre également une hyperplasie interfolliculaire associée à une surexpression des mitogènes de la famille d'EGF et à une activation de la voie des MAP kinases ; De plus, la perte de TAF4 induit la transformation maligne, caractérisée par des papillomes induits lors d'un traitement au DMBA et des mélanomes invasifs.

TAF4 est donc un important régulateur de la prolifération des kératinocytes. Il possède une activité suppresseur de tumeur. C'est un régulateur majeur de la prolifération puisqu'il contrôle l'expression de deux familles distinctes de facteurs de croissance dans des types cellulaires différents.



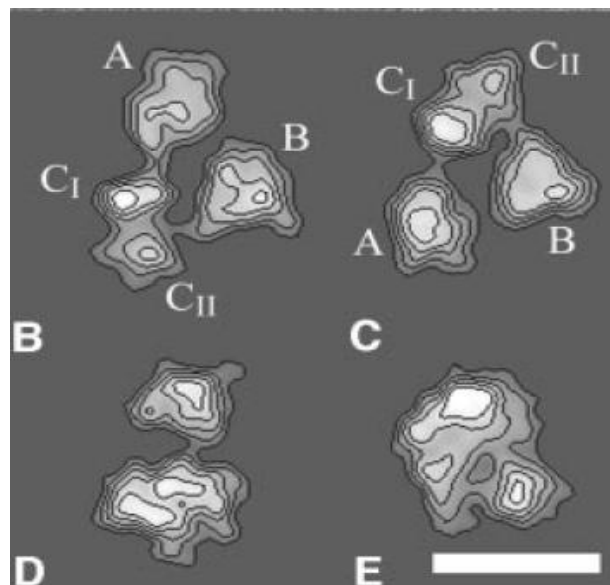
## Conclusion du Chapitre 2

L'existence d'éléments de promoteurs orphelins, interagissant avec des protéines inconnues suggère que d'autres facteurs de la transcription basale sont capables de se lier à ces éléments et de les activer.

Les facteurs composant TFIID apparaissent comme étant les candidats les mieux placés. A ce jour, la purification biochimique de TFIID par immunoprécipitation de TBP à partir de cellules en culture ou d'embryons a été la seule méthode d'identification *de novo* des sous-unités de TFIID (Dynlacht *et al.*, 1991; Takada *et al.*, 1992) (Fig.81).

**Figure 81 : Galerie des vues les plus représentatives de  $\gamma$ TFIID.**

A cause des limitations engendrées par cette approche, seules les sous-unités très abondantes et largement exprimées au sein du TFIID ont été identifiées. La **caractérisation fonctionnelle** de l'ensemble des TAFs fait encore largement défaut. Tout comme celle des sous-unités des autres complexes d'initiation de la



transcription. Or si certains TAFs peuvent être à la fois impliqués dans l'initiation et dans l'activation de la transcription, d'autres sous-unités ont peut-être aussi un **double rôle**. Des analyses comparatives de protéomes chez les Métazoaires peuvent être nécessaires afin d'identifier les domaines fonctionnels les plus courants qui participent à l'activation et à la répression de la transcription. Une collection de TAFs ou de protéines similaires va bientôt voir le jour. Il s'agira alors de discriminer les facteurs de transcription selon leur mode d'expression, de distinguer ceux qui sont exprimés de manière spécifique spatialement dans certaines cellules ou certains tissus, de ceux qui sont exprimés temporellement à des périodes précises du développement. L'étude du rôle spécifique des TAFs par des méthodes biochimiques classiques ou des analyses informatiques permettra peut-être à l'avenir de détecter de nouveaux activateurs de la transcription.



# *CHAPITRE 3*

## *LES RETINOÏDES, DES ACTEURS ANTI- CANCERIGÈNES*



## CHAPITRE 3

### Les rétinoïdes, des acteurs anti-cancérigènes

Les **rétinoïdes** forment une famille de composés chimiques, ligands d'une famille de récepteurs nucléaires tels que **RAR**, **RXR** et même **PPAR**, qui modulent l'activité transcriptionnelle de gènes cibles. Ce faisant, ils régulent les niveaux d'expression de certains gènes impliqués dans la morphogenèse, la croissance, la différenciation cellulaire et l'homéostasie. La connaissance des mécanismes dans lesquels ils sont impliqués a permis d'élaborer des stratégies thérapeutiques visant à contrecarrer l'avancée de maladies comme le cancer. Voici donc un état des lieux non exhaustif sur l'acide rétinoïque, son mécanisme d'action dans la cellule et la manière dont il est aujourd'hui utilisé en cancérologie.

#### I. Aux origines de l'acide rétinoïque

##### A. Historique

La **vitamine A** est une des premières vitamines à avoir été identifiée. Son histoire remonte bien au-delà de sa découverte, la cécité crépusculaire étant mentionnée dans la littérature ancienne aussi bien chinoise qu'indienne, grecque ou romaine.

##### 1. Les travaux de Hopkins

Au début du XX<sup>ème</sup> siècle, les chercheurs **Hopkins, Collum et Stepp** découvrent que des substances présentes dans l'alimentation sont indispensables à la croissance des rats (Wolf and Carpenter, 1997). Reprenant les travaux de Pekelharing, Gowland Hopkins (1861-1947) constate en 1906 que des rats soumis à des régimes dits « synthétiques », consistant en un mélange de protéines, de graisses, d'hydrates de carbone et de sels minéraux ne réussissent pas à grandir et à se développer normalement. L'ajout à leur régime d'une quantité infime de lait (2cc) favorise l'assimilation des protéines contenues dans ces mélanges artificiels et permet aux rats de survivre (Hopkins, 1912). Il recherche alors des facteurs de croissance dans le lait mais sans succès.

## 2. Le principe actif

La découverte du **principe actif** de la vitamine A date de 1909. L'Américain Elmer Mc Collum (1879-1967) découvre qu'un facteur actif présent dans le jaune d'œuf et dans la graisse du beurre peut être utilisé pour compléter les animaux carencés. Les rats supplémentés ont alors une durée de vie allongée et une croissance plus importante que les animaux carencés. Ce facteur soluble dans les graisses est appelé "**liposoluble A**" en 1912 (Day, 1997). Des carences en cette substance liposoluble A provoquent des altérations de la vision nocturne. Du fait de son implication dans la vision et de la nécessité de se la procurer dans son alimentation, elle est à présent appelée communément rétinol ou vitamine A. En 1931, le chercheur suisse Capper découvre son précurseur : le  $\beta$ -carotène et sa formule chimique (Capper *et al.*, 1931) ce qui permet, dès 1947, la synthèse industrielle du **rétinol**.

### B. Du rétinol à l'acide rétinoïque

#### 1. La transformation du rétinol dans l'organisme

La **vitamine A**, essentielle à la croissance et à la différenciation de plusieurs tissus, est stockée dans le foie sous forme estérifiée. Elle sera transformée et sécrétée en **rétinol** ; celui-ci est oxydé dans ses tissus cibles en **acide rétinoïque**, un de ses métabolites actifs majeurs, grâce à l'action successive de deux enzymes : la rétinol déshydrogénase puis la rétinol déshydrogénase (Fig.82).

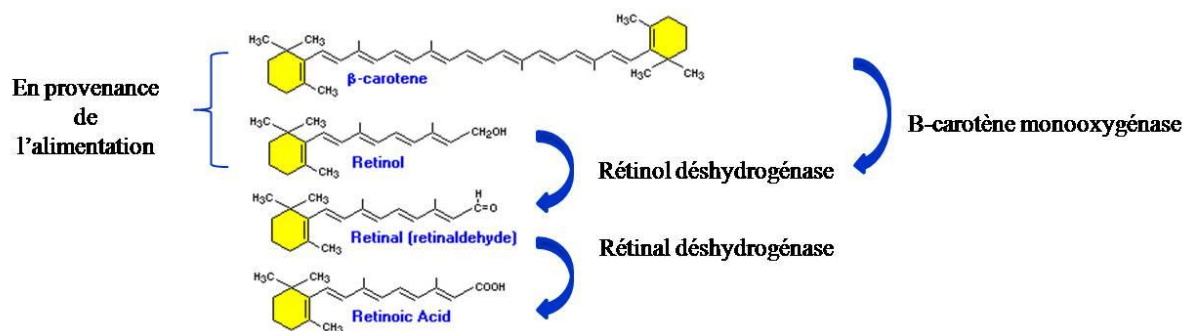


Figure 82 : Devenir du rétinol et du bêta-carotène dans l'organisme.

Ces transformations vont permettre au rétinol d'être absorbé, transporté, stocké et utilisé par l'organisme. L'ensemble des composés formés, qu'ils soient naturels ou synthétiques, constitue le groupe des **rétinoïdes**.

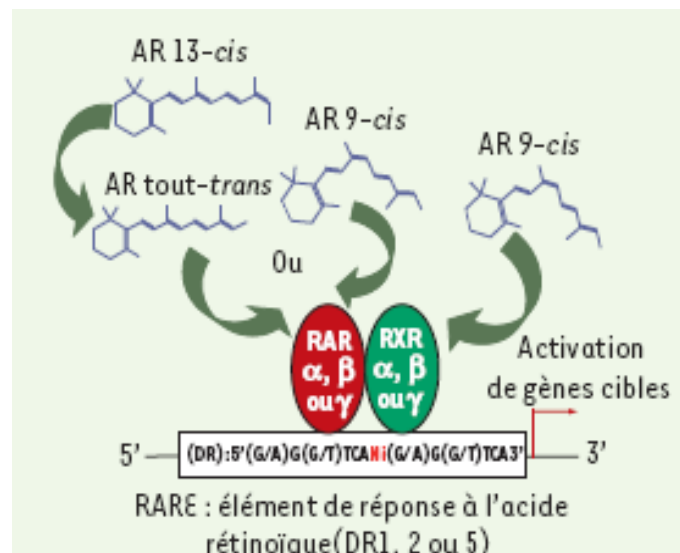


## 2. Les rétinoïdes bioactifs

A l'échelle moléculaire, les rétinoïdes naturels ont tous la même structure de base. Ils comportent un groupe cyclique non aromatique, une longue chaîne polyènique et un groupement polaire final pouvant prendre la forme d'un alcool (**rétinol**) (Roi), d'un ester (**ester de rétinol**) (RE), d'un aldéhyde (**rétinaldéhyde**) (Ral) ou d'un acide carboxylique (**acide rétinoïque**) (AR). Cette structure leur confère un caractère hydrophobe et liposoluble et leur permet de traverser facilement la bicouche lipidique membranaire.

### a) L'acide rétinoïque

L'**AR** est synthétisé dans les tissus par deux étapes successives d'oxydation à partir du rétinol (vitamine A). C'est la principale forme active de rétinoïde avec le **rétinal 11-cis**, qui participe au cycle visuel.



**Figure 83** : Les rétinoïdes bioactifs les plus étudiés : l'AR-T, l'Ar-9cis et l'AR-13cis (Parisotto et al., 2006).

Deux isomères de l'AR, les formes **tout-trans** et **9-cis**, sont des ligands de deux familles distinctes de récepteurs nucléaires : les **RAR (Retinoic Acid Receptor)** et les **RXR (Retinoid X Receptor)**. L'isomère tout-trans (AR-T) n'active que les récepteurs nucléaires RAR alors que l'isomère 9-cis (AR 9-cis) peut activer les deux familles connues de récepteurs nucléaires des acides rétinoïques, les RAR et les RXR (Fig.83). Il est à noter cependant que l'existence *in vivo* de l'AR 9-cis reste controversée. Un autre isomère naturel, l'**AR 13-cis**, ne lie pas ces récepteurs directement mais peut le faire après une isomérisation en AR tout-trans.

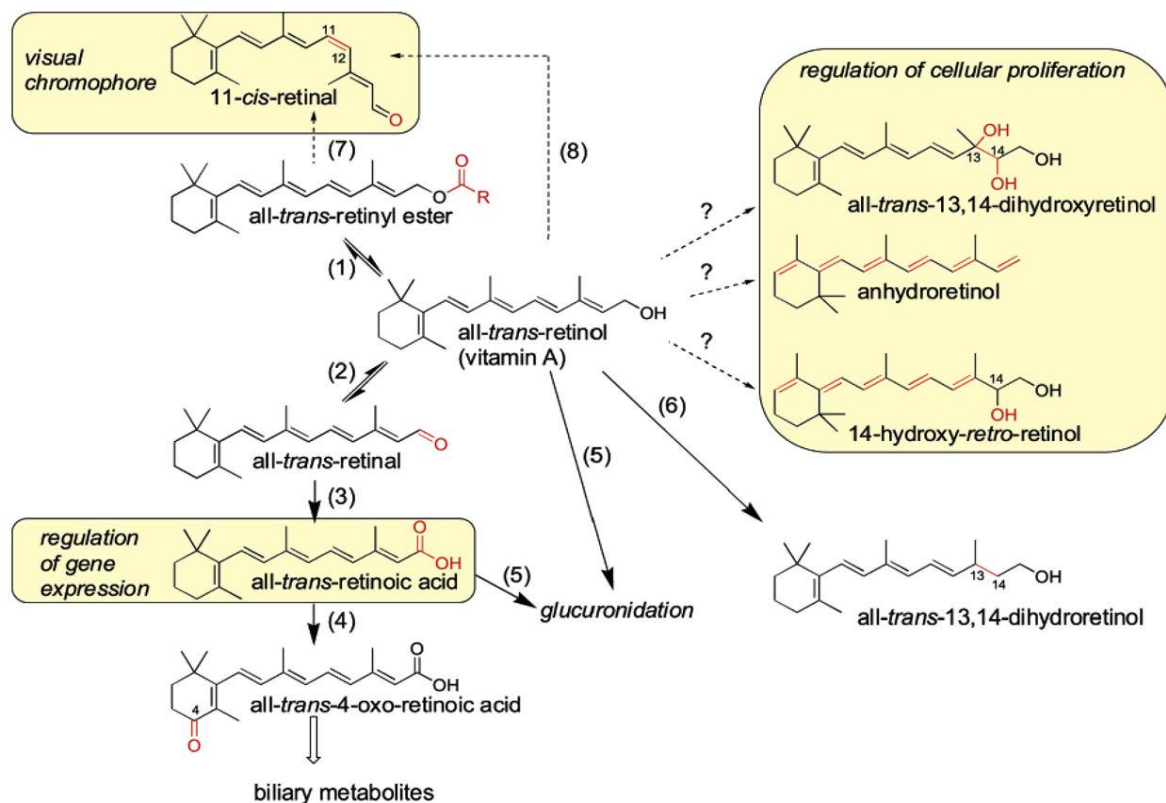
### 6) Les autres métabolites du rétinol

D'autres métabolites du RoI ont également montré des activités biologiques *in vitro* notamment (Fig.84):

-les **rétro-rétinoïdes** : l'anhydrorétinol et le 14-hydroxy-4,14-rétro-rétinol (14-HRR) qui sont impliqués dans la régulation de la prolifération des lymphocytes (Buck *et al.*, 1991; Buck *et al.*, 1993)

- le tout-trans-13,14-dihydroxy-rétinol dont l'activité de régulation de la prolifération des cellules promyélocytopiques HL-60 est comparable à celle du 14-HRR et qui est impliqué dans l'activation des cellules T (Derguini *et al.*, 1995).

L'ensemble **des métabolites bioréactifs** connus dérivant de l'AR-T est regroupé sur la figure ci-dessous (Fig.84) :



**Figure 84 :** Les rétinoides bioactifs chez les Vertébrés supérieurs. Les composés endogènes bioactifs connus sont encadrés et incluent l'AR-T qui régule l'expression des gènes via RAR, le 11-cis-rétinal qui est le chromatophore du photorécepteur de la molécule rhodopsine et le all-trans-13,14-dihydroxyrétinol, l'anhydrorétinol, le 14-hydroxy-rétro-rétinol qui contrôlent la prolifération cellulaire.

### c) Les besoins en vitamine A

Les besoins en vitamine A sont estimés à environ **5 000 UI par jour**, une unité internationale de vitamine A correspondant à 0,3µg de Rol.

Chez la souris et le rat de laboratoire, la **carence** en vitamine A conduit à un retard de croissance, à la dédifférenciation des épithéliums avec diminution de la résistance aux infections, à l' affection de certaines glandes endocrines, à une **dégénérescence** de la rétine (Dowling and Wald, 1958) et des motoneurones, (Corcoran *et al.*, 2002; Jokic *et al.*, 2007), ainsi qu'à une accumulation intracellulaire de neurofilaments et à l'apparition d'une astroglie (van Neerven *et al.*, 2008).

Au contraire, les apports excessifs en vitamine A peuvent avoir des effets toxiques. La surconsommation de vitamine A ou **hypervitaminose A** conduit en effet à des dysfonctionnements importants, à un retard de croissance et une fragilisation des os chez le rat (Johansson *et al.*, 2002). Chez l'homme, elle stimule la résorption des os et inhibe leur formation, favorisant alors la perte osseuse et pouvant contribuer au développement de l'ostéoporose (Binkley and Krueger, 2000).

De plus, la surconsommation de vitamine A sous forme de RE peut également avoir des **effets tératogènes**. Chez le rat, elle entraîne des malformations chez le fœtus comme la formation de fentes palatines ainsi qu'une réduction du nombre de doigts ou une ectromélie (absence de l'extrémité du membre) (Inomata *et al.*, 2005).

## II. Métabolisme des rétinoïdes

### A. Provenance de la vitamine A

L'homme est incapable de synthétiser le Rol. Il doit donc se le procurer dans son **alimentation**. Les aliments d'origine animale (viandes, produits laitiers et surtout foie) contiennent du Rol et **des RE** alors que les végétaux contiennent essentiellement **des carotènes** qui sont des précurseurs du rétinol. Une molécule de bêta-carotène, par hydrolyse de la liaison 15-15' sous l'influence d'une caroténoïde monooxygénase ( $\beta$ -carotène 15,15' monooxygénase), donne deux molécules de vitamine A. Par contre, les deux autres carotènes (alpha et gamma) ne donnent naissance qu'à une seule molécule de vitamine A.

Dans les pays en voie de développement, la carence en vitamine A est la première cause de cécité ; elle est également responsable d'une mortalité infantile élevée après infection principalement par la rougeole. Des applications thérapeutiques utilisent à présent des **agonistes et des antagonistes des rétinoïdes** naturels et synthétiques des différents RAR et RXR, afin de supplémer les déficiences en vitamine A dans les pays en voie de développement.

### B. Devenir du rétinol dans l'organisme

Les rétinoïdes délivrés oralement sont efficacement absorbés par l'intestin. Ils pénètrent dans les entérocytes et sont transformés en **RE** (Fig.85). Ces RE sont alors transportés vers le foie. Le RE peut être **stocké** dans le foie sous forme de RE ou être acheminé vers les autres tissus afin d'être **utilisé** par les cellules cibles.

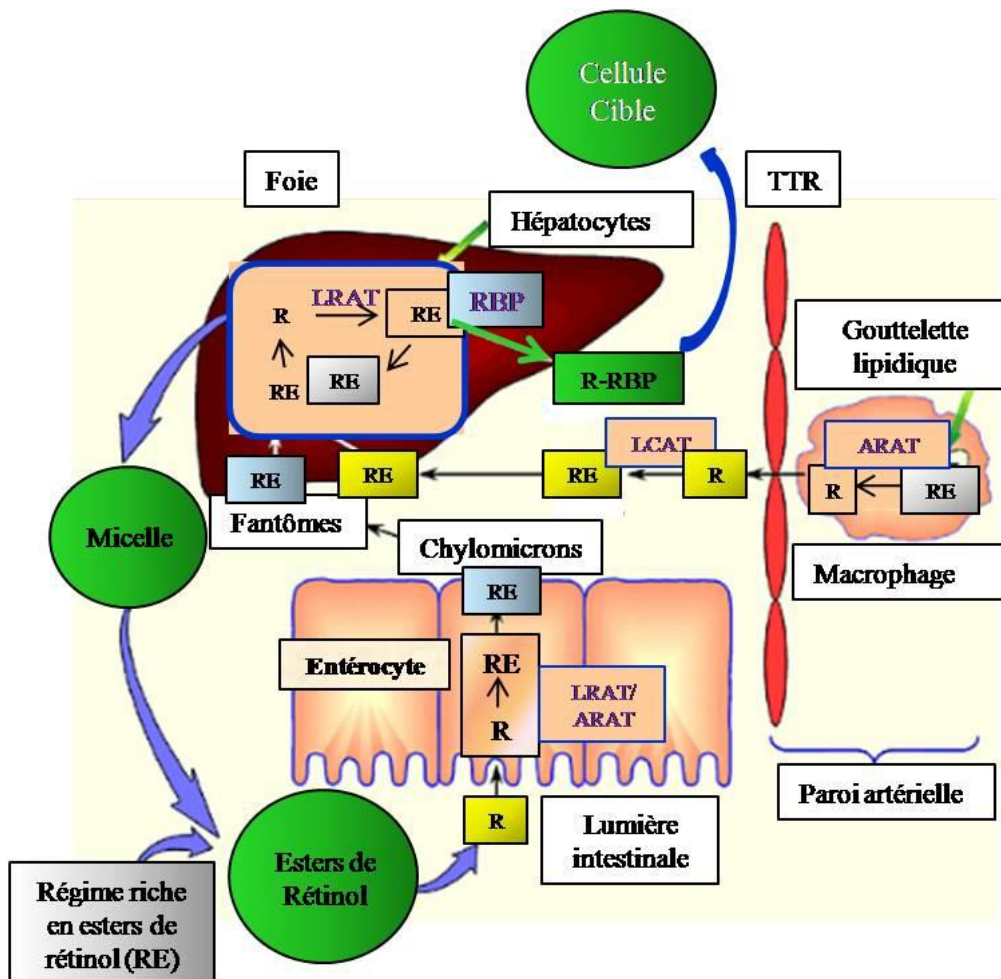


Figure 85 : Métabolisme des rétinoïdes. R : Rétinol ; RE : esters de rétinol ; LRAT : lécithine-rétinal-acyl-transférase ; ARAT : acyl-CoA-rétinol acyltransférase ; RBP : retinol binding protein.

## 1. L'absorption du rétinol

Les RE sont hydrolysés en Rol (Fig.85) dans la lumière intestinale par les actions de la **lipase triglycéride pancréatique** (van Bennekum *et al.*, 2000) ou par la **phospholipase B**, une enzyme fixée à la bordure en brosse de l'intestin (Rigtrup *et al.*, 1994).

Le Rol nouvellement généré, tout comme celui qui peut provenir de régimes diététiques ou thérapeutiques particuliers, sera incorporé dans l'entérocyte. La protéine **CRBP-II** (Cellular retinol-binding protein II) se lie au Rol avec une forte affinité (Ong, 1994) et facilite son apport comme substrat à la **lécithine rétinol–acyl transférase** (LRAT) (Fig.85). Cette enzyme ré-estérifie alors le Rol en catalysant la transacylation d'un acide gras de la lécithine au rétinol (MacDonald and Ong, 1988; Ruiz *et al.*, 1999).

Une autre enzyme, l'**acyl-CoA-rétinol acyltransférase** (ARAT) peut aussi prendre le relais dans les souris *Lrat*<sup>-/-</sup> ou pour la formation de RE dans le foie, la glande mammaire ou les tissus adipeux (O'Byrne *et al.*, 2005). Au final, les RE nouvellement synthétisés par la LRAT ou la ARAT seront incorporés avec d'autres lipides dans des **chylomicrons** (Fig.85) naissant qui seront sécrétés à l'extérieur de l'entérocyte dans le système lymphatique (Blomhoff *et al.*, 1990).

## 2. Le transport des rétinoïdes aux cellules cibles

### a) Vers le foie

Les chylomicrons du système lymphatique passent progressivement dans la circulation générale où ils sont décomposés pour former des **fantômes de chylomicrons** (Fig.85). Cette transformation fait appel à l'action de la **lipoprotéine lipase** (LPL). Cette enzyme va hydrolyser les triglycérides en acides gras libres, facilitant ainsi le recrutement ultérieur de l'**apolipoprotéine E** (apoE), un élément indispensable pour la reconnaissance future des fantômes de chylomicrons par les récepteurs hépatiques (Cooper, 1997; Redgrave, 2004).

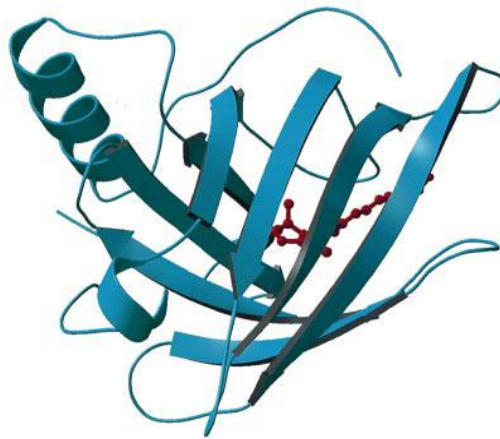
Environ 66 à 75% des RE contenus dans les chylomicrons sont pris en charge par le foie et sécrétés à nouveau dans la circulation associés à la protéine du sérum **RBP (retinol-binding protein)** ou stockés dans les cellules hépatiques stellaires en tant qu'esters (Fortuna *et al.*, 2003; Goodman *et al.*, 1965). Les 25–33% de RE restant sont répartis entre le tissu adipeux, le cœur, les muscles, les poumons, les organes reproductifs et la moelle osseuse. Les

chylomicrons se déchargent en rétinoïdes dans les tissus cibles essentiellement grâce à la LPL (van Bennekum *et al.*, 1999).

L'AR est moins hydrophobe que le Rol ou ses RE. Il peut passer dans le **système porte** plutôt que dans les chylomicrons (Noy *et al.*, 1992). Les RBP se combinent alors à la **pré-albumine** (PA) dans un rapport stœchiométrique. Le complexe PA-RBP se dissocie en deux composantes quand le rétinol est capté par les cellules cibles. Des études pharmacocinétiques indiquent que de nombreux analogues synthétiques de l'AR sont également absorbés par le système porte. Cependant, cela dépend surtout du caractère lipophile du composé et de ses propriétés chimiques (Johnson, 1999).

### 6) Dans le compartiment sanguin

Dans le compartiment sanguin, le Rol, hydrophobe, circule conjugué à des protéines de transport comme les **RBP (Retinol Binding Protein)** (Kanai *et al.*, 1968) (Fig.86). RBP est essentiellement synthétisée et sécrétée dans le foie, là où la majorité des rétinoïdes sont stockés (Blomhoff *et al.*, 1985). Cependant, elle peut être sécrétée en moindres quantités dans des autres tissus comme l'œil, les adipocytes ou les reins qui contiennent eux-aussi des stocks de RE.



**Figure 86 : Représentation en ruban de la structure tridimensionnelle de l'holo-RBP bovine (en bleu) liée au rétinol (en rouge) (Noy, 2000).**

Le complexe Rol-RBP s'associe ensuite à une autre protéine sanguine : la **transthyréline (TTR)** (Fig.85) qui empêche par la suite que le glomérule rénal ne se débarrasse du complexe Rol-RBP (Melhus *et al.*, 1991; Peterson, 1971).

### 3. Les rétinoïdes dans le compartiment cellulaire

#### a) Mécanisme d'entrée dans la cellule

Le mécanisme par lequel les tissus extra-hépatiques captent le RoI présent dans la circulation sanguine reste sujet à controverses. Composé hydrophobe en équilibre dynamique avec la protéine RBP, le rétinol pourrait se dissocier rapidement de RBP, **s'incorporer à la membrane lipidique** et diffuser dans le compartiment intracellulaire (Noy, 2000) ou se lier à un récepteur membranaire appelé **STRA6** (Maden, 2007).

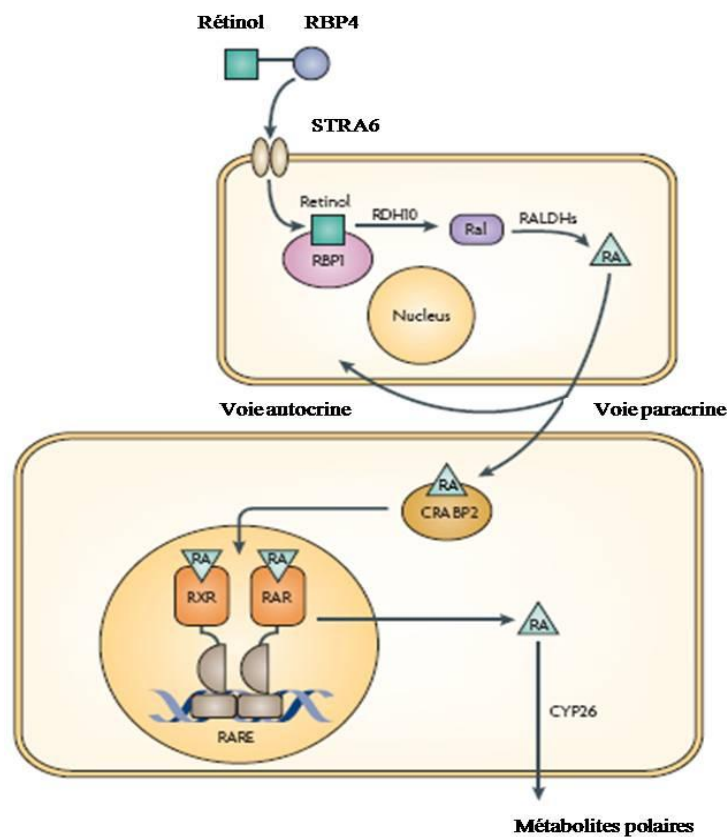


Figure 87: Les voies de signalisation impliquées dans l'anabolisme et le catabolisme de l'AR (Maden, 2007).

Sur ce schéma (Fig.87), le RoI est lié à la protéine plasmique RBP4 (retinol-binding protein 4). Celle-ci interagit avec le récepteur membranaire STRA6 et facilite l'entrée du RoI dans la cellule. Le RoI est métabolisé en Ral par la **rétinol déhydrogénase** (RDH10) puis en AR par les **rétinaldéhyde déhydrogénases** (RALDHs).

### b) Métabolisme des rétinoïdes dans la cellule

Une fois dans la cellule cible, le rétinol peut subir toute une série de transformations (Fig.88). Etant liposolubles, le rétinol et ses isomères se lient à des protéines de transport cellulaire qui les dirigent vers des enzymes bien définies. Ainsi le rétinol se lie à **CRBP** (**cellular retinol binding protein**) (Bashor *et al.*, 1973), le rétinol à **CRALBP** (**cellular retinaldehyde binding protéin**) (Stubbs *et al.*, 1979) et l'acide rétinoïque à **CRABP** (**cellular retinoic acid binding protéin**) (Ong and Chytil, 1978).

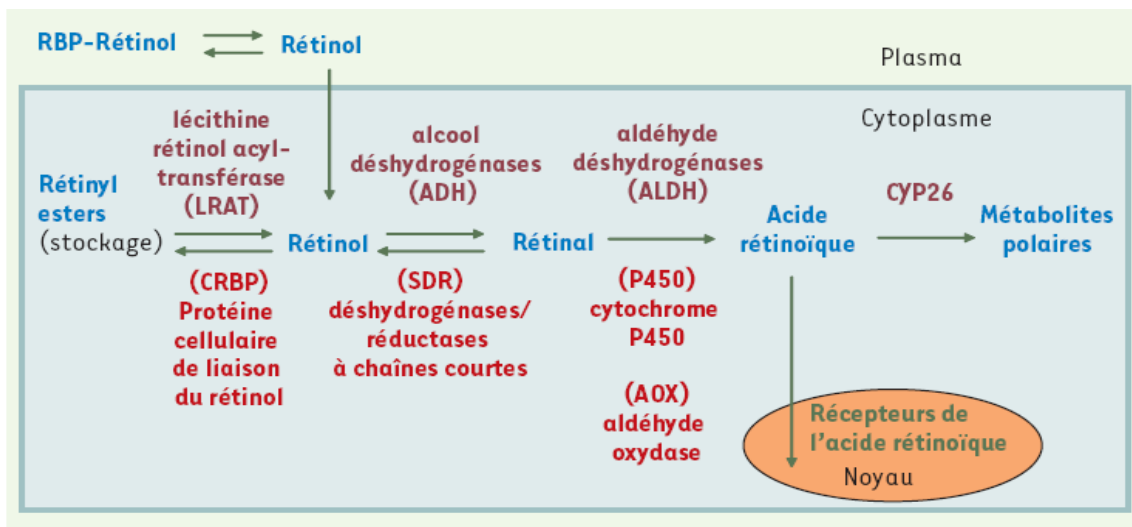


Figure 88: Métabolisme des rétinoïdes dans la cellule : de leur absorption jusqu'à l'activation de la transcription (Parisotto *et al.*, 2006).

### c) L'AR dans le noyau de la cellule

L'AR peut être transmis aux cellules voisines (**voie paracrine**) ou agir en retour sur la cellule dont il provient (**voie autocrine**) (Fig.87).

Il peut être dégradé en acide 4-oxo-rétinoïque par des cytochromes P450 hydroxylases comme CYP26 (Abu-Abed *et al.*, 1998). C'est la première étape du catabolisme de l'AR. Celle-ci va aboutir à la formation de métabolites polaires et à leur excrétion.

Il peut entrer dans le noyau de la cellule, assisté par sa protéine de transport **CRABP2**. Il agit ensuite comme un **ligand** et se lie aux récepteurs RAR ou RXR qui sont capables de s'hétérodimériser et de lier une séquence d'ADN, le RARE (retinoic acid-response element). Il en résulte une activation de la transcription de gènes cibles via la liaison de l'AR à son récepteur nucléaire.



### III. Les récepteurs des rétinoïdes

#### A. Composition de la famille des récepteurs des rétinoïdes

Les récepteurs des rétinoïdes RAR et RXR ont la structure caractéristique des récepteurs nucléaires décrite, à savoir l'organisation en différentes régions A à F pour les RARs et A à E pour les RXRs.

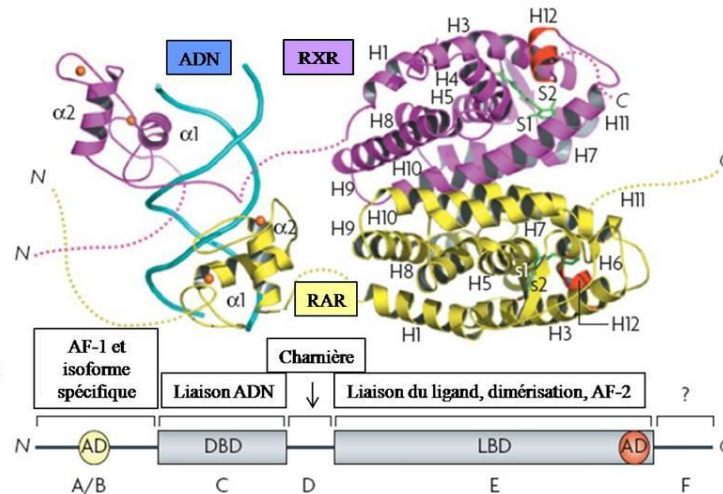


Figure 89 : Les récepteurs RXR et RAR (de Lera *et al.*, 2007).

Les RARs et les RXRs existent tout deux sous la forme de trois isotypes différents : alpha, beta et gamma. L'expression de ces six récepteurs (RAR  $\alpha$ ,  $\beta$ ,  $\gamma$  et RXR $\alpha$ ,  $\beta$ ,  $\gamma$ ) varie selon les cellules. Cependant, la plupart des cellules, sinon toutes, expriment au moins l'un de ces récepteurs.

Chaque RAR ( $\alpha$ ,  $\beta$ , et  $\gamma$ ) existe sous plusieurs isoformes qui sont répertoriées dans le Tableau 6. Il existe deux isoformes majeurs pour RAR $\alpha$  ( $\alpha 1$  et  $\alpha 2$ ), deux pour RAR $\gamma$  ( $\gamma 1$  et  $\gamma 2$ ) et quatre pour RAR $\beta$  ( $\beta 1$  à  $\beta 4$ ) (Kastner *et al.*, 1990; Leroy *et al.*, 1991; Zelent *et al.*, 1991) (Tableau 7). Ces isoformes diffèrent par leur région A du côté amino-terminal (Fig.89). Ils proviennent de l'utilisation de promoteurs différents et/ou sont générés par l'épissage alternatif de l'ARN pré-messager (Leid *et al.*, 1992).

De même que pour les RARs, les RXRs ( $\alpha$ ,  $\beta$ , et  $\gamma$ ) existent sous plusieurs isoformes. Il existe deux isoformes majeures pour RXR $\alpha$  ( $\alpha 1$  à  $\alpha 2$ ), deux pour RXR $\beta$  ( $\beta 1$  et  $\beta 2$ ) et deux pour RXR $\gamma$  ( $\gamma 1$  et  $\gamma 2$ ) (Brocard *et al.*, 1996; Liu & Linney, 1993; Nagata *et al.*, 1994). Ces

isoformes diffèrent par leur région A/B du côté amino-terminal et leur région D. Ils sont issus uniquement de l'utilisation de promoteurs différents (Kastner et al., 1994a).

### B. Spécificité du ligand

Les RARs et les RXRs ne lient pas tous les rétinoïdes. Comme le montre le Tableau 7, les RARs et les RXRs fixent différents isomères de l'acide rétinoïque (Chambon, 1996). Alors que les récepteurs RAR sont activés par l'**acide rétinoïque all-trans** et par son isomère **9-cis**, les récepteurs RXR ne répondent qu'au ligand 9-cis. Chaque famille est composée de 3 sous-types  $\alpha$ ,  $\beta$ ,  $\gamma$  codés par des gènes indépendants (Leid *et al.*, 1992).

Gènes	Isoformes majeures	Ligand
RAR $\alpha$	$\alpha 1, \alpha 2$	AR-T, AR-9cis
RAR $\beta$	$\beta 1, \beta 2, \beta 3, \beta 4$	AR-T, AR-9cis
RAR $\gamma$	$\gamma 1, \gamma 2$	AR-T, AR-9cis
RXR $\alpha$	$\alpha 1, \alpha 2$	AR-9cis
RXR $\beta$	$\beta 1, \beta 2$	AR-9cis
RXR $\gamma$	$\gamma 1, \gamma 2$	AR-9cis

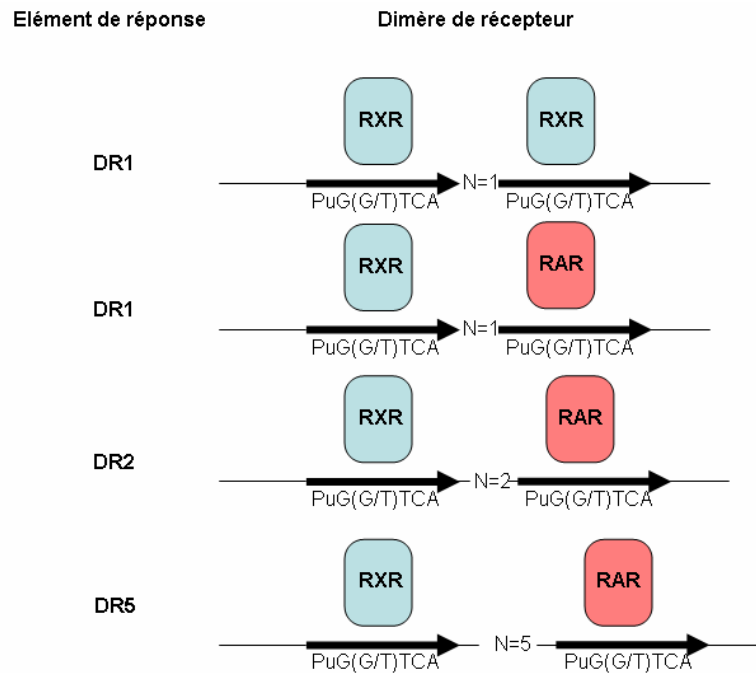
Tableau 7 : Isoformes majeures de chaque récepteur des rétinoïdes RAR ( $\alpha 1-2, \beta 1-4, \gamma 1-2$ ) et RXR ( $\alpha 1-2, \beta 1-2, \gamma 1-2$ ) ainsi que leurs ligands respectifs (acide rétinoïque tout-trans et/ou 9-cis).

Même si l'acide rétinoïque 9-cis a été décrit comme ligand naturel de RXR (Heyman et al., 1992), les analyses montrent qu'il est difficile de le détecter *in vivo* (Kurlandsky et al., 1995; Werner & DeLuca, 2001). Depuis le début de ce siècle, un intérêt grandissant porte sur la recherche d'autres ligands de RXR. Lengqvist et al. ont montré que certains acides gras insaturés sont des ligands naturels des RXRs (Lengqvist et al., 2004).

### C. Éléments de réponse

La **boîte RARE (Retinoic Acid Response Element)** est localisée dans les régions régulatrices des gènes cibles de l'acide rétinoïque. Elle est constituée de deux répétitions du motif hexamérique PuG(G/T)TCA organisés en répétition directe DR (Germain *et al.*, 2004;

Leid *et al.*, 1992; Mangelsdorf and Evans, 1995) et permet la fixation des RARs (Fig.90). Ces motifs répétés sont séparés par 1, 2, ou 5 paires de bases. Classiquement, ces répétitions sont séparées par 5 paires de bases et appelées **DR5**, mais il arrive également qu'une seule paire de bases (DR1) ou deux (DR2) les séparent. La liaison de l'AR aux récepteurs nucléaires qui agissent sous forme d'hétérodimères RXR/RAR (Kastner *et al.*, 1997), peut se faire sur des DR mutés ou plus espacés que la normale (Delacroix *et al.*, 2009 en soumission).



**Figure 90 : Exemple de liaison des récepteurs des rétinoïdes sur des demi-sites en répétition directe (DR) avec l'homodimère RXR/RXR et l'hétérodimère RXR/RAR.**

Les récepteurs de l'AR agissent comme des facteurs de transcription, s'associant sous forme d'homodimères de RXR ou d'hétérodimères RAR-RXR à des séquences spécifiques d'ADN.

*In vitro*, l'hétérodimère RAR/RXR a une efficacité de liaison croissante avec le nombre de nucléotides séparant les demi-sites (ce nombre est représenté par N=1, N=2 ou N=5 sur la Figure 90) (Chambon, 1996). Par exemple, l'efficacité de liaison de RAR/RXR est plus faible sur les DR1, plus forte sur les DR2 et encore plus importante sur les DR5. De plus, RXR se lie avec une affinité plus importante lorsqu'il est sous forme d'hétérodimère avec RAR que sous forme d'homodimère (Bugge *et al.*, 1992).

Un élément de réponse à l'acide rétinoïque de type DR5 est présent dans le promoteur du gène *Rarb2* qui est un gène inductible par l'acide rétinoïque (de The et al., 1990; Sucov et al., 1990). Cependant, il faut noter que la présence d'une telle séquence dans la région promotrice d'un gène peut générer *in vitro* une réponse à l'acide rétinoïque sans pour autant être la preuve qu'il s'agisse là d'un gène cible de l'acide rétinoïque *in vivo*.

#### D. Possibilités de dimérisation

En règle générale, les récepteurs nucléaires régulent l'expression de gènes cibles en se liant sur des séquences spécifiques de l'ADN sous forme d'homodimères ou d'hétérodimères et plus rarement sous forme de monomères (Glass, 1994). Plus particulièrement, les récepteurs des rétinoïdes se fixent sur l'ADN sous forme d'homodimères ou d'hétérodimères pour le cas des RXRs et semblent uniquement se fixer sous forme d'hétérodimères pour le cas des RARs. La complexité combinatoire des hétérodimères provenant de l'existence de plusieurs formes de RAR et de RXR (trois gènes  $\alpha$ ,  $\beta$  et  $\gamma$  et plusieurs isoformes pour chaque gène) assure une régulation transcriptionnelle complexe des gènes cibles par l'AR.

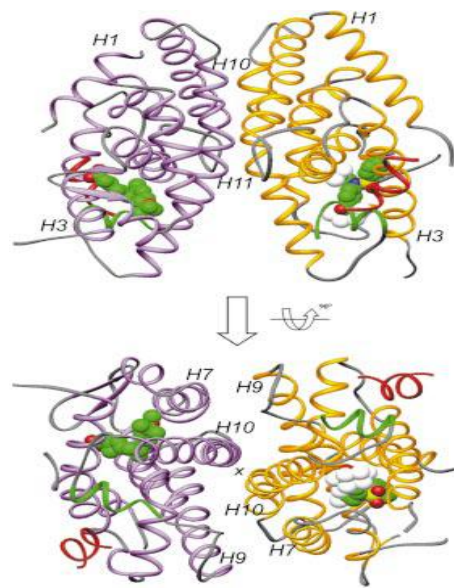
##### 1. Dimères impliquant RXR

Les RXRs sont impliqués dans de nombreuses voies de signalisation. Ils peuvent agir sous forme d'homodimères mais ils s'hétérodimérisent le plus souvent avec d'autres récepteurs nucléaires comme les COUP-TF (facteur de transcription du promoteur de l'ovalbumine de poulet), FXR (récepteur des farnésoïdes), LXR (récepteur X du foie), NGF1B (facteur de croissance neuronale 1B), PPAR (récepteur des péroxisomes en prolifération), RAR (récepteur de l'acide rétinoïque), TR (récepteur des hormones thyroïdiennes), ou VDR (récepteur de la vitamine D3) (Chambon, 1996; Laudet & Gronemeyer, 2002; Mangelsdorf & Evans, 1995). Certains hétérodimères impliquant RXR peuvent activer la transcription des gènes par la simple liaison du ligand sur RXR. C'est le cas des hétérodimères RXR/FXR, RXR/LXR (Fig.91) ou RXR/PPAR (Blumberg & Evans, 1998).

Cependant, dans d'autres cas d'hétérodimérisation impliquant les RXRs (RXR/RAR, RXR/TR, RXR/VDR), la transcription des gènes n'est pas activée par la liaison du ligand sur RXR (Forman et al., 1995). La régulation de la transcription est plus complexe et implique que chaque partenaire ait lié son ligand. Par exemple, l'hétérodimère RAR/RXR est

transcriptionnellement actif seulement si les deux partenaires du dimère ont fixé leur ligand respectif (Altucci & Gronemeyer, 2001). De plus, le complexe ligand/RXR ne peut recruter des coactivateurs que si RAR a déjà lié son ligand. Ce phénomène est appelé subordination du RXR. Germain et al. ont proposé un mécanisme moléculaire pouvant expliquer le fonctionnement de la subordination des RXRs. Dans ce mécanisme, RXR est subordonné par la fixation de corépresseurs sur RAR, qui forme un environnement stérique empêchant le recrutement de coactivateurs sur RXR (Germain et al., 2002). Cette subordination n'est pas observée dans le cas des hétérodimères PPAR/RXR et LXR/RXR (Forman et al., 1995; Mukherjee et al., 1997).

**Figure 91 : Structure générale de l'hétérodimère LXR $\alpha$ -RXR $\beta$  présenté sur deux vues séparées de 90° (Svensson et al., 2003).**



## 2. Dimères impliquant RAR

Même si les RARs peuvent se fixer *in vitro* sous forme d'homodimères à des séquences spécifiques sur l'ADN de gènes cibles (promoteur P2 de RAR $\beta$ 2, DR5), cette liaison n'est réellement efficace que par l'intervention de facteurs externes. Les premiers facteurs à avoir été mis en évidence sont les RXRs (Leid et al., 1992). Ainsi, l'**hétérodimère RAR/RXR** permet d'activer la transcription des gènes cibles. *In vitro*, la comparaison des homodimères RAR/RAR par rapport aux hétérodimères RAR/RXR a permis de montrer que l'hétérodimérisation permet d'augmenter l'efficacité mais aussi la sélectivité de la liaison aux différents éléments de réponse (Mader et al., 1993).

Deux autres récepteurs nucléaires peuvent s'hétérodimériser avec les RARs (Glass, 1994). Ces hétérodimères ont été mis en évidence uniquement *in vitro*. C'est le cas de

l'**hétérodimère RAR/VDR**. Cet hétérodimère peut tout d'abord se lier à l'élément de réponse 5'-GGGTGA-3' en répétition directe. L'acide rétinoïque tout-trans, ligand de RAR, suffit alors pour induire l'expression du gène en aval. L'hétérodimère RAR/VDR peut également se lier à l'élément de réponse de l'acide rétinoïque (RARE). Dans ce cas, la transcription est activée par une association d'acide rétinoïque tout-trans (ligand de RAR) et de vitamine D3 (ligand de VDR) (Schrader et al., 1993). Garcia-Villalba et al. ont proposé que la liaison de l'hétérodimère RAR/VDR sur l'élément de réponse de l'acide rétinoïque inhiberait la liaison de l'hétérodimère RAR/RXR qui est plus actif (Garcia-Villalba *et al.*, 1996). Ces données suggèrent une interaction directe entre les voies de signalisation de l'acide rétinoïque et celles de la vitamine D3. RAR peut aussi s'hétérodimériser avec le récepteur des hormones thyroïdiennes (TR) (Glass et al., 1989). L'hétérodimère RAR $\alpha$ /TR $\beta$  peut induire ou inhiber l'expression des gènes en fonction de l'élément de réponse auquel il se lie (Glass et al., 1989) et des co-facteurs qu'il recrute (Lee and Privalsky, 2005).

Toutes ces possibilités de dimérisation des RARs augmentent et complexifient les possibilités de régulation de l'expression génique par l'intermédiaire des récepteurs nucléaires. Elles fournissent un moyen de couplage de différentes voies de signalisation hormonales. Cependant, ces observations ont été menées *in vitro* et la fonctionnalité *in vivo* de ces hétérodimères reste à démontrer.

### 3. Le rôle des RARs et RXRs et les phénotypes de leurs mutants

Afin de mieux comprendre le rôle joué par RARs et RXRs dans la transduction du signal des rétinoïdes, l'**invalidation des gènes** de chacune des isoformes des RARs et des RXRs a été effectuée chez la souris (Kastner *et al.*, 1995). Le tableau 7 résume brièvement les phénotypes des souris invalidées pour un seul type de RAR ou de RXR.

#### a) Les mutants

Chez la souris, les mutants Rar $\alpha$ , Rar $\beta$  et Rar $\gamma$  sont viables (Mark *et al.*, 2009). Mais ils présentent quelques **malformations congénitales** ainsi que certains aspects du syndrome fœtal de déficience en vitamine A (VAD) (Tableau 8). Après inactivation d'un type de RAR, des malformations spécifiques apparaissent. Chaque isoforme de RAR exerce donc un rôle bien précis dans la voie de signalisation de l'AR.

Souris invalidée	Phénotype
RAR	
RAR $\alpha$ -/-	<ul style="list-style-type: none"> <li>• Diminution de viabilité, déficience de la croissance et stérilité des mâles (anomalies de la spermatogenèse)</li> <li>• Malformations congénitales variées (cervicales, doigts)</li> <li>• Anomalies des effets exogènes des rétinoïdes dans les cellules myéloïdes</li> </ul>
RAR $\beta$ -/-	<ul style="list-style-type: none"> <li>• Défauts de la locomotion (détérioration de la voie de signalisation de la dopamine)</li> <li>• Détérioration des fonctions cognitives</li> <li>• Anomalies des corps vitreux des yeux (RAR<math>\beta</math>2-/-)</li> </ul>
RAR $\gamma$ -/-	<ul style="list-style-type: none"> <li>• Phénotype proche de celui observé chez les souris RAR<math>\alpha</math>-/-</li> <li>• Diminution de viabilité, déficience de la croissance, et stérilité des mâles</li> <li>• Anomalies dans la prostate et des vésicules séminales</li> <li>• Anomalies congénitales</li> <li>• Résistance aux effets tératogènes de l'acide rétinoïque</li> </ul>
RXR	
RXR $\alpha$ -/-	<ul style="list-style-type: none"> <li>• Létalité durant l'embryogenèse (défauts cardiaques et placentaires)</li> <li>• Anomalies oculaires</li> <li>• Anomalies dans le développement de la peau (renouvellement du poil)</li> </ul>
RXR $\beta$ -/-	<ul style="list-style-type: none"> <li>• 50 % de létalité embryonnaire</li> <li>• Stérilité des mâles (défauts testiculaires et spermatogenèse anormale)</li> <li>• Défauts de la locomotion (détérioration de la voie de signalisation de la dopamine)</li> </ul>
RXR $\gamma$ -/-	Résistance aux effets de l'hormone thyroïdienne

**Tableau 8 : Caractéristiques phénotypiques des souris invalidées pour un type de RAR ou de RXR (Chomienne C. ; Despouy G., 2005).**

### *6) La redondance des récepteurs aux rétinoïdes*

L'ensemble de ces anomalies se trouve dans les tissus cibles où sont exprimés ces récepteurs normalement. Or, les malformations observées ne correspondent pas en totalité aux anomalies décrites dans le **syndrome VAD**. Certains récepteurs doivent donc être capables de prendre en charge l'absence d'autre isoformes. L'hypothèse d'une **redondance** entre les différents récepteurs est envisageable (Kastner *et al.*, 1995; Mascrez *et al.*, 1998).

Afin de tester cette hypothèse, deux isoformes de RAR de souris ont été mutés simultanément. Les double-mutants Rar $\alpha/\beta^-$ , Rar $\alpha/\gamma^-$  et Rar $\beta/\gamma^-$  ne sont pas viables et meurent *in utero* ou à la naissance à cause de défauts développementaux sévères. Les rétinoïdes sont donc **essentiels au développement embryonnaire (Dolle, 2009)**; ce sont eux qui les **RAR/RXRs qui sont responsables de la traduction du signal de l'AR**.

#### IV. Le rôle des rétinoïdes en cancérologie

##### A. La vitamine A, une molécule miracle ?

En 1922 avant que la structure chimique de la vitamine A ne soit connue, le docteur Mori observe qu'un **régime pauvre en vitamine A** chez des les enfants des pays en voie de développement s'accompagne de l'ébauches de métaplasies épithéliales du système respiratoire et des glandes salivaires (Wolf, 1998).

En 1925, Wolbach décrit la présence d'une prolifération excessive et d'un défaut de différenciation cellulaire au niveau du tractus digestif et du système génito-urinaire chez des animaux **carencés en vitamine A** (Wolbach and Howe, 1978). Une métaplasie épithéliale de type squameuse est observée, de type semblable aux changements morphologiques induits par certains oncogènes.

Au contraire, un épiderme de poulet soumis à un milieu de culture contenant un **excès en vitamine A** voit son l'épithélium kératinisé se transformer en un épithélium métastatique et muqueux (Fell and Mellanby, 1953).

Le rapprochement entre la vitamine A et le cancer devient encore plus évident lorsque Fujimake démontre en 1926 l'apparition de carcinomes de l'estomac chez des rats privés de vitamine A dans leur alimentation. 50 ans après, Lasnitski inhibe *in vitro* chez la souris des lésions pré-cancéreuses prostatiques par divers rétinoïdes (Lasnitzki, 1976). On commence alors à saisir l'importance d'approfondir le rôle des rétinoïdes sur la carcinogenèse, et de nombreuses études *in vivo* chez le rongeur vont rechercher les effets possibles anti-néoplasiques des rétinoïdes sur divers organes (Moon and Mehta, 1990).

##### B. Effets anti-néoplasiques des rétinoïdes

Progressivement les rétinoïdes ont été testés chez la souris, afin d'étudier leurs effets anti-tumorigéniques. Plusieurs modes d'action apparaissent :

-L'acide 13-cis rétinoïque incite les cellules du carcinome thyroïdien à se redifférencier (Kim *et al.*, 2009) alors que l'utilisation d'AR **inhiberait l'angiogenèse et la croissance des cellules tumorales** (Hoffmann *et al.*, 2007).



-Dans le système digestif, les rétinoïdes préviennent l'apparition de carcinomes gastriques induits par des hydrocarbures polycycliques et inhibent les tumeurs œsophagiennes dues aux nitrosamines. Ils induisent **l'arrêt du cycle cellulaire et l'apoptose**. Par ailleurs, l'incidence des néoplasies hépatiques et pancréatiques est réduite.

-Dans le système urinaire, les rétinoïdes s'avèrent posséder des **propriétés anti-carcinogènes** sur les tumeurs vésicales du rongeur.

-Dans la glande mammaire, certains rétinoïdes peuvent inhiber chez le rat le développement de cancers induits par le MNU (N-méthyl-N-nitrosurée) et le DMBA (7,12-diméthylbenz [a] anthracène).

Les rétinoïdes sont des candidats très prometteurs dans la prise en charge des cancers en s'inscrivant dans une thérapie s'associant à d'autres agents anti-cancéreux, comme par exemple avec l'interféron (Lippman *et al.*, 1993). Il existe donc une réelle nécessité de mettre au point un dérivé de la vitamine A possédant une spécificité d'organe et une toxicité acceptable.

### C. *Comment les rétinoïdes peuvent-ils inhiber la carcinogenèse ?*

D'une manière générale, les mécanismes d'action des rétinoïdes dans la prévention et la thérapie cancéreuse agissent sur 5 niveaux différents de la carcinogenèse : sur le carcinogène lui-même, dans la phase de promotion de la tumeur, sur la perte de différenciation cellulaire, sur la réponse immune et dans la phase de progression tumorale. Gollnick, dans sa publication intitulée « Theoretical aspects of the use of retinoids as anticancer agents » définit les propriétés suivantes comme étant caractéristiques des rétinoïdes :

- la capacité de stopper la transformation d'un pro-carcinogène en carcinogène.
- l'inhibition de l'ornithine décarboxylase (ODC), enzyme qui intervient dans le métabolisme des polyamines donc dans la réplication cellulaire.
- l'expression à la surface cellulaire de récepteurs à l'EGF (epidermal growth factor).
- la modulation des protéines de surfaces membranaires et des jonctions inter-cellulaires.

- la stimulation des cellules de Langerhans dont le nombre diminue lors de néoplasie cutanée ou après exposition aux UVB ; de plus, on observe une modulation des lymphocytes T, B et des natural killers (NK) ainsi qu'une stimulation de la sécrétion d'IL-2.
- la diminution de l'angiogenèse et des métalloprotéases participant à la progression de la tumeur et à la formation de métastases.

### 1. *Les rétinoïdes et leur action anti-tumorale*

Depuis que l'on connaît leur importance sur le contrôle de la différenciation cellulaire des épithéliums, les rétinoïdes suscitent un intérêt considérable dans leur application potentielle pour **la thérapie ou la prophylaxie anti-tumorale** puisque la plupart des néoplasies se forment à partir d'une structure épithéliale et s'associent à une perte de la différenciation cellulaire. C'est donc précisément grâce à la recherche oncologique que nous devons initialement le développement de rétinoïdes synthétiques. Ceux-ci sont d'ores et déjà utilisés pour inhiber la tumorigenèse induite par des carcinogènes au niveau de différents tissus (Miller, 1998; Sun and Lotan, 2002).

Carcinogenèse cutanée. Les souris soumises à une alimentation supplée en rétinol développent moins de papillomes cutanés induits de manière exogène par le DMBA que les témoins. Leurs lésions peuvent même régresser dans ces conditions (Davies, 1967). En 1971, Bollag démontre que l'administration systémique de palmitate de rétinyle possède un effet thérapeutique sur les tumeurs cutanées de la souris (Bollag, 1971) puis il confirme cette découverte avec d'autres rétinoïdes (Bollag, 1972); plus tard, il met en évidence avec Matter l'effet inhibiteur des rétinoïdes sur la carcinogenèse cutanée induite par les esters de phorbols et établit un index thérapeutique de divers rétinoïdes basé sur leur efficacité anti-tumorale et leur toxicité (Bollag and Matter, 1981).

### 2. *L'AR et son effet anti-prolifératif*

L'AR sous sa forme "tout-trans" est indispensable à certaines phases du développement ainsi qu'au maintien de l'homéostasie chez l'adulte. De nombreuses données suggèrent que l'AR peut également avoir un effet anti-prolifératif sur diverses formes de tumeurs exprimant son récepteur naturel.

### 3. L'AR et la différenciation

L'acide rétinoïque joue un rôle déterminant dans la différenciation des cellules épithéliales en culture. En absence d'acide rétinoïque, les cellules s'orientent vers d'une différenciation squameuse (Jetten, 1992). Par contre, en présence d'acide rétinoïque, les marqueurs de la différenciation squameuse sont supprimés et, il est observé entre autres, une diminution de la prolifération et augmentation de la sécrétion de glycoconjuguée (Chopra, 1989).

#### D. Les pathologies associées aux récepteurs RAR

Les anomalies des récepteurs RAR sont essentiellement décelées dans les cancers ou les états précancéreux. Ces anomalies aboutissent toutes à des pertes de fonction de ces récepteurs, qui mènent à la **différenciation** et la **survie des cellules**. La réinduction de l'expression du récepteur déficient peut aboutir à la reprise du processus de différenciation, à l'arrêt de prolifération et à l'induction de mort cellulaire.

#### 1. La leucémie aigüe promyélocytaire

L'implication la plus directe des RARs dans une pathologie humaine a été décelée avec la leucémie aigüe promyélocytaire également appelée LAM3.

RAR $\alpha$  est normalement exprimé dans la cellule souche myéloïde et permet la différenciation granulocytaire, c'est-à-dire le passage du stade de myéloblaste à celui de promyélocyte, en présence de concentrations physiologiques d'AR.

##### a) La translocation réciproque

La LAM3 se caractérise par une translocation réciproque acquise impliquant systématiquement le gène RAR $\alpha$  sur le chromosome 17 et dans la majorité des cas, le gène PML (ProMyelocytic Leukemia) localisé sur le chromosome 15 (de The *et al.*, 1991).

Dans des cas plus rares de LAM3, le gène RAR $\alpha$  peut être fusionné à quatre autres gènes : PLZF (Promyelocytic Leukemia Zinc Finger) situé en 11q23, NPM (NucleoPhosMin) situé en 5q31, NuMA (Nuclear Mitotic Apparatus) situé en 11q23 et STAT5b (Signal Transducer and Activator of Transcription 5b) situé en 17q11 (Pandolfi, 2001). Toutes ces

anomalies du gène codant pour RAR $\alpha$  aboutissent à un blocage au stade de promyélocyte, caractéristique phénotypique de ces leucémies.

Les translocations dans les LAM3 sont réciproques et génèrent des protéines de fusion X/RAR $\alpha$  et RAR $\alpha$ /X, X correspondant aux différents gènes partenaires de RAR $\alpha$ . Ces dernières sont impliquées dans la genèse des leucémies et déterminent le pronostic de la maladie (Pollock *et al.*, 1999). Les protéines de fusion X/RAR $\alpha$  exercent une inhibition dominante négative des voies de régulation transcriptionnelle contrôlées par RAR $\alpha$  et par son partenaire de fusion tel que PML ou PLZF (Pandolfi, 2001).

### 6) L'affinité pour l'AR et le recrutement de RXR

Toutes les protéines de fusion X/RAR $\alpha$  conservent des séquences identiques de RAR $\alpha$  comprenant les domaines LBD et DBD. Ces protéines de fusion X/RAR $\alpha$  ont une affinité pour l'AR comparable à celle de RAR $\alpha$  et peuvent fixer un élément de réponse de type RARE sous forme d'homodimère ou sous forme d'un complexe multimérique contenant les RXRs (Lin *et al.*, 2001) (Fig.92). La présence de RXR $\alpha$  au sein de ces complexes tendrait à renforcer leur liaison à l'ADN (Zhu *et al.*, 2007) (Fig.67). Le complexe PML/RXR $\alpha$  peut agir sur RXR $\alpha$  et induire sa sumoylation via l'enzyme Ubc9. Une fois sumoylé, il n'est plus silencieux et devient alors un répresseur transcriptionnel (Choi *et al.*, 2006). L'activation de la transcription pourrait, d'après l'auteur, avoir lieu grâce à des modifications post-traductionnelles du RXR $\alpha$  telles que la phosphorylation. Celle-ci pourrait entraîner son export nucléaire, et sa dégradation protéosome-dépendante.

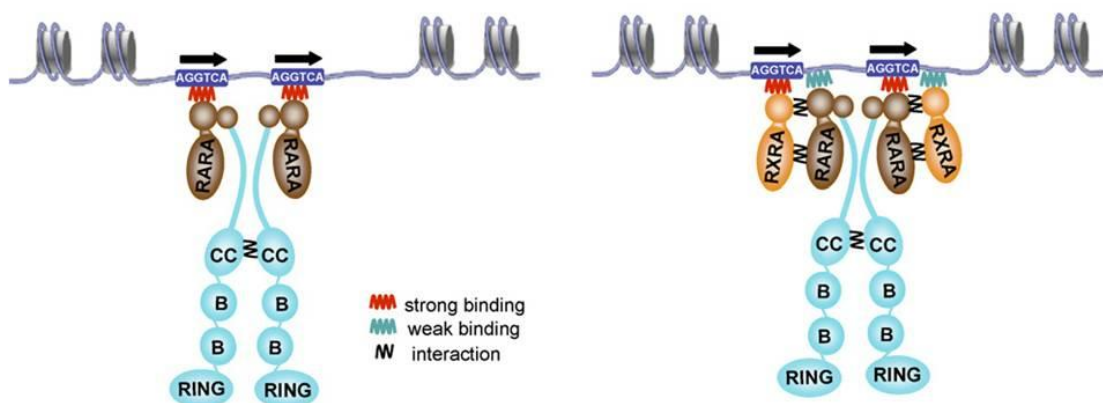


Figure 92 : Augmentation de la liaison à l'ADN du complexe PML/RAR $\alpha$  grâce à la présence de RXR $\alpha$  (Zhu *et al.*, 2007).

c) Activité transcriptionnelle

A des concentrations physiologiques d'AR ( $10^{-9}$  à  $10^{-8}$ M), l'expression de PML/RAR $\alpha$  **bloque l'activation transcriptionnelle** par l'AR des gènes cibles et empêche ainsi la différenciation de lignées cellulaires myéloïdes leucémiques (Fig.93). Par contre, des doses élevées d'AR ( $10^{-7}$  à  $10^{-6}$  M) activent la transcription des gènes et la différenciation cellulaire induite par PML/RAR $\alpha$  (Lin *et al.*, 2001).

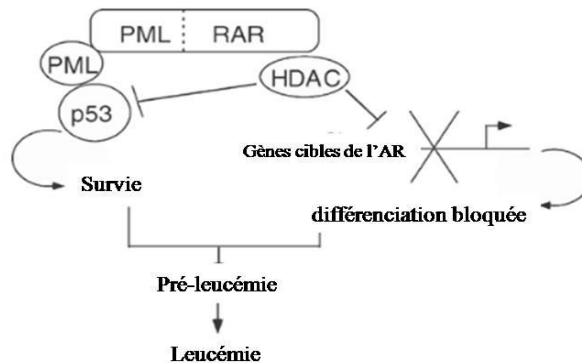


Figure 93 : Recrutement des HDAC par PML-RAR $\alpha$  (Insinga *et al.*, 2004).

Les deux protéines de fusion PML/RAR $\alpha$  et PLZF/RAR $\alpha$  interagissent fortement avec les corépresseurs N-CoR et SMRT et recrutent les protéines HDAC en l'absence de ligand et en présence de doses physiologiques d'AR (Fig.93 et 94).

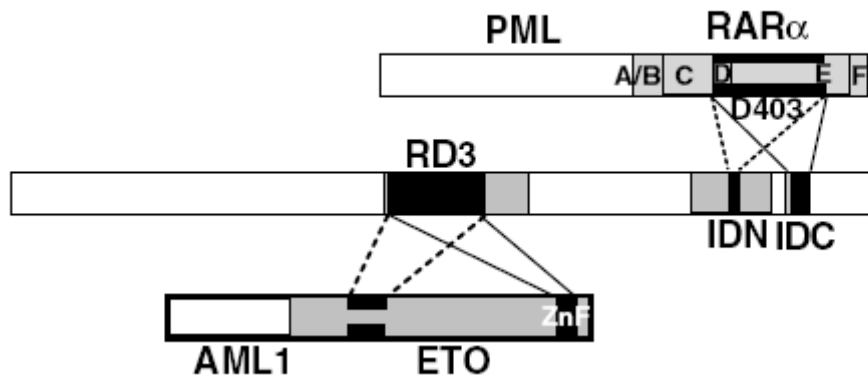


Figure 94 : Interaction entre les domaines de RAR $\alpha$  (en haut) et NCoR (au centre).

L'interaction entre la protéine de fusion et son corépresseur est très forte, car chacune des deux protéines fusionnées possède un site de liaison pour son corépresseur, ce qui n'est pas le

cas pour l'hétérodimère RAR/RXR. Comme pour le modèle classique des rétinoïdes, l'inhibition de la transcription va se faire via le recrutement spécifique d'enzymes

Le recrutement des complexes HDAC (Fig.95) semble être crucial dans la fonction oncogénique des X/RAR $\alpha$  puisque des inhibiteurs des HDACs en combinaison avec l'AR peuvent lever l'activité répressive des X/RAR $\alpha$  et déclencher la différenciation terminale des cellules de LAM3 résistantes à l'AR. La combinaison d'inhibiteurs des HDACS et d'AR a permis des rémissions cliniques chez des patients atteints d'une LAM3 exprimant PML/RAR $\alpha$  (Petti *et al.*, 2002).

PML/RAR $\alpha$  et PLZF/RAR $\alpha$  sont associées à de multiples complexes corépresseurs sous des concentrations physiologiques d'AR. L'augmentation locale de la concentration de corépresseurs au niveau des gènes cibles de l'AR entraînerait une répression transcriptionnelle et un blocage de la différenciation promyélocytaire.

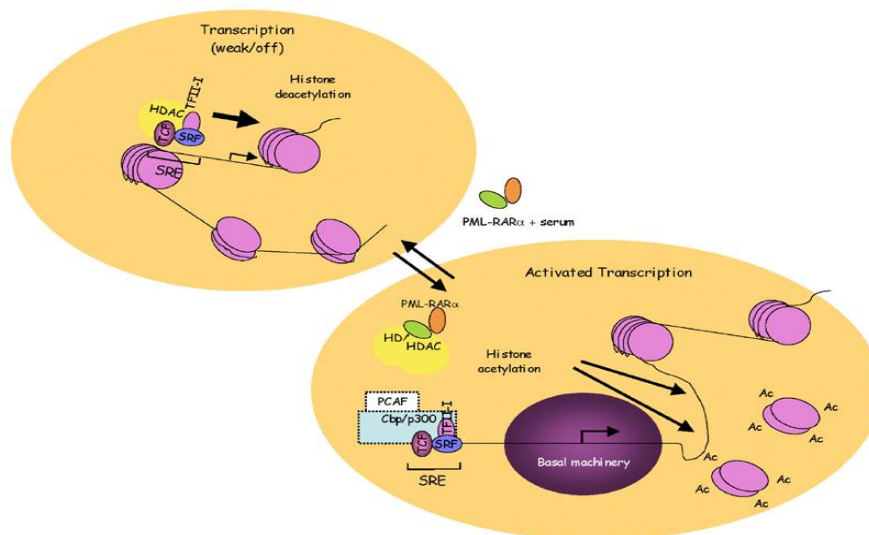


Figure 95 : Recrutement des protéines HDAC par le complexe PML/RAR $\alpha$  en l'absence de ligand.

#### d) La réparation de la cellule cancéreuse

##### -Action de l'acide rétinoïque

Les travaux récompensés de l'équipe française reposent sur le traitement par différenciation des maladies malignes humaines. L'originalité du concept tient dans une remise en question du dogme selon lequel le cancer est un état anormal et irréversible. A l'origine de cette nouvelle orientation: la démonstration de l'action thérapeutique de l'AR.

La collaboration établie entre Anne Dejean-Assemat (Institut Pasteur, Paris) et Laurent Degos (Hôpital St-Louis, Paris) a consisté à démontrer l'action du traitement par différenciation sur le désordre génétique de la leucémie aiguë promyélocytaire. Le récepteur de l'acide rétinoïque perd toutes ses facultés thérapeutiques au contact du PML et autorise ainsi le développement de la leucémie aiguë promyélocytaire.

En effet, apporté en doses importantes, l'acide rétinoïque provoque le déplacement des éléments PML. Une fois isolé, le récepteur de l'acide rétinoïque ré-active le programme de maturation des globules blancs : la cellule cancéreuse peut ainsi être transformée en une cellule normale. Par leurs travaux, Laurent Degos et Anne Dejean-Assemat confèrent à la réparation de la cellule cancéreuse une légitimité et de nouvelles perspectives dans la guérison d'autres leucémies et d'autres tumeurs malignes.

Des doses pharmacologiques d'AR déclencheraient la dissociation des complexes corépresseurs fixés sur PML/RAR $\alpha$ , conduisant à l'activation des gènes, dont le gène RAR $\alpha$  (Chomienne *et al.*, 1991) à la différenciation des cellules leucémiques et à la rémission clinique. Contrairement aux autres types de LAM, les patients atteints de LAM3 répondent à un traitement à l'AR à des concentrations pharmacologiques. Paradoxalement, l'acide rétinoïque lève ce blocage et induit la différenciation, aboutissant à la rémission morphologique complète de la leucémie, avec, en particulier, réorganisation des structures nucléaires et réapparition des corps nucléaires. Des études cliniques ont montré que la rémission était due à une **différenciation** des blastes leucémiques (Castaigne *et al.*, 1990). La sensibilité de la LAM3 à l'AR-T a constitué le premier exemple d'une thérapie basée sur la différenciation cellulaire. L'association de l'AR-T à des inhibiteurs des HDACs offre des nouvelles voies thérapeutiques efficaces, notamment dans les leucémies PLZF-RAR (Lowenberg *et al.*, 2003).

### *-Acide rétinoïque et arsenic*

L'arsenic, bien connu comme poison, peut cibler spécifiquement le produit d'une lésion génétique à l'origine d'un type de leucémie... C'est à ce résultat paradoxal qu'a aboutie la collaboration entre une équipe chinoise de l'Institut d'hématologie de Shanghai et une équipe française, dirigée par le professeur Hugues de Thé. La leucémie aiguë promyélocytaire est actuellement traitée par l'AR, associé à une chimiothérapie, ce qui constitue le premier

exemple d'une thérapeutique directement ciblée sur l'oncogène. Des équipes chinoises ont mis en évidence un effet thérapeutique de l'arsenic chez les malades atteints par ce type de leucémie (Chen *et al.*, 1996). Sur plus de 70 patients traités, cette médication a donné 65 % de réponses cliniques complètes (mais transitoires) et 25 % de guérisons à dix ans. Compte tenu des difficultés du traitement de cette maladie en Chine, une telle survie était exceptionnelle.

Le tri-oxyde d'arsenic,  $As_2O_3$ , est le principe actif de ces médicaments. Contrairement à l'AR qui déclenche la différenciation cellulaire, l'arsenic induit la mort programmée, ou apoptose. Cette collaboration franco-chinoise s'est penchée sur le mode d'action de l'arsenic et a montré que celui-ci, comme l'AR, provoque la dégradation de l'oncogène PML/RARa, soulignant une surprenante similitude des effets de ces deux agents. Si l'acide rétinoïque agit sur la partie RAR de la fusion, l'arsenic, lui, agit sur sa partie PML (Zhu *et al.*, 1997). Ces données permettent l'élaboration d'un modèle physio-pathologique de cette maladie, dans lequel RAR contrôlerait la différenciation et PML l'apoptose. Sur des sujets en état de résistance aux rétinoïdes, l'injection intraveineuse d' $As_2O_3$  induit des rémissions complètes et prolongées. Il n'y a donc pas de résistance croisée entre ces deux agents. Cela permet donc d'envisager des traitements combinés.

Les mécanismes moléculaires concernant le traitement à l'arsenic ont été récemment découverts avec notamment l'induction par l'arsenic de la sumoylation d'une lysine de la protéine PML par RNF4 (RING-domain-containing ubiquitin E3 ligase). Cette modification conduit à la dégradation de la protéine. L'inactivation de RNF4 ou l'inhibition du protéasome entraîne une accumulation de chaînes protéiques polyubiquitinylées. Dans ce cas, l'action de l'arsenic est réduite à néant et les protéines PML modifiées par SUMO (small ubiquitin-like modifiers) s'accumulent alors au sein du noyau (Tatham *et al.*, 2008).

## 2. *Le récepteur $RAR\beta$ et les cancers*

### a) *$RAR\beta$ , un suppresseur de tumeur*

Le gène codant pour le récepteur  $RAR\beta$  est considéré comme un suppresseur de tumeur (Soprano *et al.*, 2004). La répression de  $RAR\beta$  a été trouvée dans de nombreux types de cancers (Altucci and Gronemeyer, 2001) et dépend surtout de la méthylation de l'ADN de son promoteur (Kwong *et al.*, 2002). Dans le cancer de la prostate, des altérations



épigénétiques, comme l'hyperméthylation de gènes tels que GSTP1, APC, RAR $\beta$ 2, Cyclin D2, MDR1, and PTGS2 ont été détectées. Elles pourraient désormais être utilisées comme des biomarqueurs pour ce cancer (Costa *et al.*, 2007).

La perte de l'expression de ce récepteur a été montrée par Northern blot et par hybridation in situ dans de nombreuses tumeurs solides. L'équipe de Schmutzler a par exemple mis en évidence sa faible expression dans les lignées de cancer folliculaire de la thyroïde, alors que RAR $\beta$  est fortement exprimé dans le tissu thyroïdien normal (Schmutzler *et al.*, 1998). L'expression du RAR $\beta$  est supprimée chez de nombreuses patientes atteintes de cancer du sein (Xu *et al.*, 2001).

### b) Une diminution de l'ARNm de RAR $\beta$ dans les cancers

D'autres études ont montré que, dans plusieurs lignées dérivées de carcinome épithélial, le niveau d'expression de l'ARNm de RAR $\beta$ 2, l'une des quatre isoformes de ce récepteur, est diminuée ou indétectable (Nervi *et al.*, 1991). De la même façon, une diminution significative, par rapport au tissu normal, se rencontre dans les cancers de la tête et du cou (Xu *et al.*, 1995). Ces résultats évoquent que la diminution de l'expression de RAR $\beta$  est associée aux lésions dysplasiques et parfois au développement de carcinomes de la tête et du cou. Lotan *et al.* mettent en évidence que la diminution de l'expression du transcrit RAR $\beta$  se produit au stade précoce de la carcinogenèse du cancer de la prostate (Lotan *et al.*, 2000). Il en est de même pour les cancers du poumon où RAR $\beta$  est également impliqué. Les travaux de Houle (Houle *et al.*, 1993) mettent en évidence une tumorigénicité moindre lorsque des souris immunodéficientes SCID sont transfectées avec le RAR $\beta$ . D'autre part, l'apparition de cancer pulmonaire est également réduite chez des souris transgéniques exprimant un anti-sens RAR $\beta$ 2 (Berard *et al.*, 1996). La diminution de l'expression de ce récepteur est aussi repérée dans des cancers de l'œsophage (Qiu *et al.*, 1999) et de la prostate (Lotan *et al.*, 2000).

### c) Les causes de cette diminution

Des **réarrangements** et des **altérations** dans la région codante du gène ont été trouvés dans des lignées de cancer pulmonaire. La réinduction de l'expression de RAR $\beta$  par le 13cis-AR prouve cependant qu'un des 2 allèles n'est pas altéré. D'autres mécanismes sont également avancés, tels que la **suppression de l'expression** du gène, à la suite soit de la sur-expression d'un autre récepteur qui diminuerait celle de RAR $\beta$ , comme il l'a été montré dans

des cellules épithéliales, ou soit par l'expression de **dominant négatif RAR ou RXR** qui antagoniserait le RARE de RAR $\beta$  (Blanco *et al.*, 1996).

De plus, la transcription du gène étant dépendante de corépresseurs et de coactivateurs, des modifications à leurs niveaux ont également été évoquées, comme **la perte du EA1-like factor**, nécessaire à l'activation de la transcription du gène, par son interaction avec d'autres cofacteurs liant le site de fixation à l'AMPc, dans le promoteur de RAR $\beta$ . Les niveaux d'expression et les interactions entre nur77 et COUP-TF, deux récepteurs orphelins, sont des déterminants importants dans l'expression de RAR $\beta$  et dans son induction par l'AR-T dans des cellules de cancers pulmonaires. En effet, COUP-TF en réprimant l'activité basale de transactivation des RAREs, sensibilise les cellules à l'AR-T, alors que des cellules exprimant nur77 sont résistantes à ce rétinol.

La **méthylation de l'ADN**, événement épigénétique fréquent, fait partie des mécanismes rendant silencieux les gènes suppresseurs de tumeurs dans les cancers. Plusieurs études ont montré la méthylation du gène codant pour RAR $\beta$  dans des lignées des tumeurs primaires et de cancer du sein (Widschwendter *et al.*, 2000) et le traitement de ces lignées par un agent déméthylant ou en présence de l'AR-T restaure l'expression de RAR $\beta$ . De plus, la désacétylation de l'histone H3 sur le promoteur de RAR $\beta$  corrèle avec le caractère réfractaire à l'AR des lignées de cancer du poumon. En effet, dans leur étude, Suh et al. montrent que l'AR augmente l'acétylation des histones H3 et H4 au niveau du promoteur RAR $\beta$ 2 dans ces lignées sensibles à l'AR. En revanche, dans des lignées résistantes, l'AR induit uniquement l'acétylation de l'histone H4, suggérant ainsi que la perte de l'acétylation de l'histone H3 est impliquée dans le phénomène de résistance à l'AR. La combinaison de l'AR avec d'autres agents thérapeutiques, comme des inhibiteurs de la méthylation ou des histones désacétylases, permettrait ainsi de traiter les tumeurs résistantes à l'AR.

L'induction du RAR $\beta$  par l'AR est également corrélée à une inhibition de la prolifération et à un effet apoptotique dans des lignées de cancer du poumon et du sein. De plus, la transfection de RAR $\beta$  dans des lignées de cancer de la tête et du cou, résistantes aux rétinoïdes entraîne un arrêt de la croissance via une induction de l'apoptose. Dans la thyroïde, l'étude d'Haugen et al. montre qu'une inhibition de la prolifération par le 9cis-AR est observée dans les lignées de cancers papillaires, mais uniquement lorsque celles-ci expriment à l'état basal non seulement RAR $\beta$ , mais aussi un autre récepteur, le RXR $\gamma$ .

Enfin, l'expression de ce récepteur peut dépendre de la concentration cellulaire en rétinoïde puisque son promoteur contient un RARE de type DR5. L'expression de RAR $\beta$  diminue dans des organes déficients en vitamine A, et augmente avec des rétinoïdes.

Ces résultats soulignent la diminution du récepteur RAR $\beta$  à des stades précoces de la tumorigenèse dans différents cancers. L'expression de RAR $\beta$  pourrait ainsi être utilisée en tant que marqueur biologique de la carcinogenèse et marqueur prédictif d'une réponse thérapeutique, puisque son induction lors de traitements par des rétinoïdes est associée à une réponse clinique, notamment dans des cancers du rein (Thacher *et al.*, 2000).

### E. Les pathologies associées aux récepteurs RXR

Du fait de son hétérodimérisation avec de nombreux autres récepteurs nucléaires, le récepteur RXR pourrait être impliqué dans une grande variété de pathologies. Dans ce contexte, des ligands synthétiques de RXR ont été développés pour le traitement de plusieurs pathologies comme le psoriasis, les cancers et pour certaines pathologies métaboliques comme l'obésité et le diabète. De plus, le réxinoïde (ligand synthétique de RXR) réduirait la résistance à l'insuline chez les souris diabétiques et obèses (Mukherjee *et al.*, 1997). Par ailleurs, il aurait aussi un effet sur la différenciation terminale des cellules humaines de lipocarcinome qui est également induite par le ligand de PPAR (Tontonoz *et al.*, 1994).

L'observation des souris RXR $\beta^{-/-}$  montre un défaut dans la voie de signalisation de la dopamine suggérant un rôle potentiel de ces récepteurs dans la maladie de Parkinson et la schizophrénie (Krezel *et al.*, 1998). Actuellement, un nouvel agoniste semble être prometteur dans le traitement des tumeurs solides : le bexarotène (Targretin<sup>®</sup>), agoniste du/des RXRs. Une étude de phase I montre 2 réponses sur 9 cas de patients atteints de lymphome cutané et une stabilisation dans un cas de cancer bronchique et dans un carcinome épidermoïde de la tête et du cou. Le composé est également testé pour le psoriasis et il est actuellement en phase II des essais cliniques (Breneman *et al.*, 2007).

Les anomalies des récepteurs des rétinoïdes aboutissent le plus souvent à des pathologies malignes liées probablement au rôle capital des rétinoïdes et de ces récepteurs dans la différenciation, la prolifération et la mort cellulaire. L'absence complète ou l'altération de la voie du rétinoïde dans une cellule altère ce contrôle cellulaire et participe aux événements oncogéniques. Le décryptage des différentes voies de signalisation de l'AR au

sein des différents tissus, ainsi que la synthèse de rétinoïdes spécifiques de ces voies d'action devraient permettre de mieux cibler les cancers susceptibles de répondre aux rétinoïdes.

## F. Les autres pathologies traitées par les rétinoïdes

### 1. AR et médulloblastome

L'équipe du Dr Yan s'est aperçu que des copies supplémentaires multiples du gène, appelées OTX2, avaient été retrouvées au sein des cellules tumorales retirées chirurgicalement chez les patients ayant un médulloblastome. Aux Etats-Unis, le médulloblastome représente environ 30% de toutes les tumeurs pédiatriques de cerveau.

Le gène OTX2 contribue normalement au développement **embryonnaire** et à la croissance de certains secteurs du cerveau, tels que le cervelet, mais le gène est généralement inactif après la naissance. Le Dr Yan et ses collègues ont pourtant détecté un segment anormalement **amplifié** (jusqu'à 28 fois) parmi les cellules de tumeur sur le **chromosome 14**. Les scientifiques ont également déterminé que 60-70% des cellules de médulloblastome avait une activité importante au niveau de l'ARN du gène OTX2, alors que les cellules normales ne montrent généralement aucune activité pour ce gène.

L'AR-T, déjà utilisé dans le traitement de la **LAM3** inhibe la croissance et induit la mort cellulaire parmi les cellules de tumeur exprimant **OTX2**. Plus de la moitié des médulloblastomes en laboratoire a réagi au traitement à l'AR-T (Di *et al.*, 2005).

### 2. AR et sarcomes

Le sarcome synovial est un sarcome d'origine encore inconnue. Le profil d'expression des gènes dérégulés dans ce sarcome s'est avéré être très proche de celui des tumeurs des gaines neurales malignes périphériques (Ishibe *et al.*, 2008) et le signal FGF appartenait à une des voies d'induction les plus dérégulées. L'origine neurale des sarcomes synoviaux a été confirmée en utilisant des cellules tumorales en culture et des tumeurs primaires. Le traitement avec BMP4 ou FGF2 augmente l'expression des gènes neuraux et le traitement avec l'AR-T fait de même, engendrant l'expression de Tuj-1, une protéine spécifique des neurones. Un effet d'inhibition de croissance a été constaté sur les xénogreffes tumorales dans des souris sans thymus. Le traitement simultané avec des inhibiteurs de FGF augmente encore plus l'effet d'inhibition de croissance dû au traitement à l'AR. Ainsi la combinaison de

l'inhibition des voies de signalisation de la croissance ainsi que de différenciation pourrait représenter une solution pour traiter ces sarcomes synoviaux.

Tous ces exemples de traitement de cancers avec de l'acide rétinoïque reposent surtout sur la faculté de ce métabolite à entraîner la différenciation cellulaire ou à diriger les cellules cancéreuses vers l'apoptose. Mais déjà d'autres voies d'action commencent à être étudiées, notamment au niveau de la matrice extra-cellulaire (MEC) des cellules. Ainsi des protéines attachées à la MEC se révèlent être régulées par l'action de l'AR. RAMP (Retinoic acid-regulated nuclear matrix-associated protein) en fait partie (Li *et al.*, 2009) et est surexprimée dans 7 cancers gastriques. Son inhibition pourrait être une approche prometteuse dans le traitement des cancers gastriques. De même, l'étude du traitement à l'AR-T des cellules tumorales mammaires MCF7 a révélé une inhibition par l'AR-T de MMP-2 (Dutta *et al.*, 2009), une métalloprotéase directement en relation avec les intégrines et les cadhérines E. Il serait donc intéressant d'étudier l'effet de l'AR-T sur des molécules qui pourraient être impliquées dans la dégradation de la MEC et dans la progression tumorale maligne.



## *Conclusion du chapitre 3 :*

Les rétinoïdes, métabolites de la vitamine A, modulent les taux de transcription des gènes impliqués dans la prolifération cellulaire et la différenciation.

Leur activité chimio-thérapeutique et chimio-préventive est déjà utilisée dans un grand nombre de cancers humains. Déjà, plusieurs thérapies basées sur la prise de rétinoïdes existent et d'autres très prometteuses sont envisagées dans un futur proche, comme le traitement de la dégénérescence maculaire par le 11-cis rétinol.

Pourtant, la thérapie basée sur les rétinoïdes est souvent confrontée à la toxicité de ces composés, même utilisés à des doses pharmacologiques.

Des observations récentes suggèrent que les facteurs qui régulent le prélèvement et le stockage des rétinoïdes dans les cellules pourraient servir de cibles à de nouvelles thérapies stratégiques. C'est le cas notamment de la LRAT, l'enzyme qui catalyse l'estérification des rétinoïdes et des CRBPs, les protéines coopérant avec la LRAT dans la régulation de l'utilisation des rétinoïdes et de leur stockage.

Modifier les niveaux ou les activités de ces protéines pourrait permettre un meilleur contrôle du taux de rétinoïdes bioactifs dans les tissus, réduisant les quantités à des doses indispensables pour des thérapies efficaces et ciblant ainsi des tissus spécifiques.

Des voies de signalisation encore méconnues de l'AR restent également à explorer et l'action des protéines de la MEC sur les cellules tumorales pourrait bien être une approche particulièrement intéressante à étudier.





# *RESULTATS*



# REVUE



## Revue sur la structure et la fonction de TFIID

**Emilie CLER, Gabor PAPAI, Patrick SCHULTZ et Irwin DAVIDSON** (12 pages)

La régulation de l'initiation de la transcription par l'ARN Polymérase de type II requiert la formation d'un complexe de pré-initiation (PIC). Ce PIC s'assemble sur le promoteur des gènes à réguler au niveau du site de démarrage de la transcription. En plus de la Pol II, le PIC comprend tout un ensemble de facteurs généraux de la transcription : TFIIA, B, D, E, F et H. Le facteur TFIID est lui-même un complexe multiprotéique, composé de la protéine TBP et de 13 à 14 TAFs qui ont été conservés au cours de l'évolution.

Cette revue fait un tour d'horizon des avancées biochimiques, génétiques, et de microscopie électronique qui ont eu lieu au cours des dernières années, à propos du facteur de transcription général TFIID.

Elle décrit tout d'abord la composition du TFIID puis la manière dont se passe l'assemblage de toutes ses sous-unités. Elle présente également les informations structurales récentes provenant de l'étude en microscopie électronique du facteur TFIID chez la Levure à la résolution de 23 Å.

Elle met ensuite l'accent sur la formation possible de sous-complexes de TFIID à 5 ou à 7 TAFs, observés respectivement chez la Drosophile et chez le Ver *Caenorhabditis elegans*, concluant la partie structurale en soulignant que la stabilité de TFIID dépend d'une somme d'interactions dans laquelle plusieurs TAFs jouent un rôle critique.

Enfin l'aspect fonctionnel du facteur TFIID est envisagé, tout d'abord au sein de cellules germinales avec l'intervention de paralogues de TAFs, puis au niveau des promoteurs dépourvus de boîte TATA. Cette seconde partie récapitule donc les multiples interactions avec l'ADN et avec les activateurs renforçant le recrutement de TFIID et souligne l'importance des interactions de TFIID avec des histones méthylés et acétylés.

Plus globalement, l'ensemble des études génétiques suggère un modèle où TFIID est nécessaire à l'initiation de la transcription après la mitose ou au cours du développement, mais ne sert pas au maintien de la transcription d'un promoteur dans un état actif ni à la réinitiation de sa transcription.



## Recent advances in understanding the structure and function of general transcription factor TFIID

Emilie Cler · Gabor Papai · Patrick Schultz · Irwin Davidson

Received: 21 January 2009 / Revised: 19 February 2009 / Accepted: 23 February 2009 / Published online: 24 March 2009  
© Birkhäuser Verlag, Basel/Switzerland 2009

**Abstract** The general transcription factor TFIID is a macromolecular complex comprising the TATA-binding protein (TBP) and a set of 13–14 TBP associated factors (TAFs). This review discusses biochemical, genetic and electron microscopic data acquired over the past years that provide a model for the composition, organisation and assembly of TFIID. We also revisit ideas on how TFIID is recruited to the promoters of active and possibly repressed genes. Recent observations show that recognition of acetylated and methylated histone residues by structural domains in several TAFs plays an important role. Finally, we highlight several genetic studies suggesting that TFIID is required for initiation of transcription, but not for maintaining transcription once a promoter is in an active state.

**Keywords** *Drosophila* · Yeast · Electron microscopy · Chromatin · Acetylation · Methylation

### Introduction

Regulated initiation of transcription by RNA polymerase II requires the formation of a macromolecular preinitiation complex (PIC) that assembles over the transcription start site. In addition to RNA polymerase II (pol II), the PIC comprises a set of general transcription factors TFIIA, B, D, E, F and H. In yeast, the mediator (Med) and NC2 complexes are also found at the promoters of most active

genes [1, 2]. TFIID is itself a macromolecular complex composed of the TATA-binding protein (TBP) and a set of 13–14 TBP-associated factors (TAFs). TFIID has been the subject of intense study over the last 20 years, and many review articles have illustrated how our understanding of its function has progressed during this time (see, for example, [3–7]). A subset of TAFs is also present in the Spt-Ada-Gcn5-acetyltransferase (SAGA)-type transcriptional regulatory complexes, and several recent reviews describe the important roles of these complexes in transcription regulation [7–10].

In this review, we will discuss results concerning the organisation and assembly of the TFIID complex and how it is recruited to active promoters, not only through interactions with DNA or activators, but also as a reader of covalent histone modifications. Lastly, we will address the question of when the function of TFIID is actually required. Several studies provide genetic data that should modify our view of TFIID's role in transcription.

### TFIID comprises core and peripheral modules

TFIID was first characterised in *Drosophila* and in mammals where immunoprecipitation with antibodies against the TBP subunit revealed the presence of a set of tightly associated TAF proteins [11–13]. TFIID was subsequently biochemically characterised in the yeast *Saccharomyces cerevisiae* (hereafter yeast) [14–16]. The genes encoding the TAFs were isolated in each of these organisms, allowing the identification of structural and functional domains that are conserved from yeast to humans. While initial results suggested that there may be differences in the composition of TFIID in each species, an extensive series of biochemical and genetic studies rather showed that the

E. Cler · G. Papai · P. Schultz · I. Davidson (✉)  
Institut de Génétique et de Biologie Moléculaire et Cellulaire,  
CNRS/INSERM/ULP, 1 Rue Laurent Fries,  
67404 Illkirch Cedex, France  
e-mail: irwin@titus.u-strasbg.fr

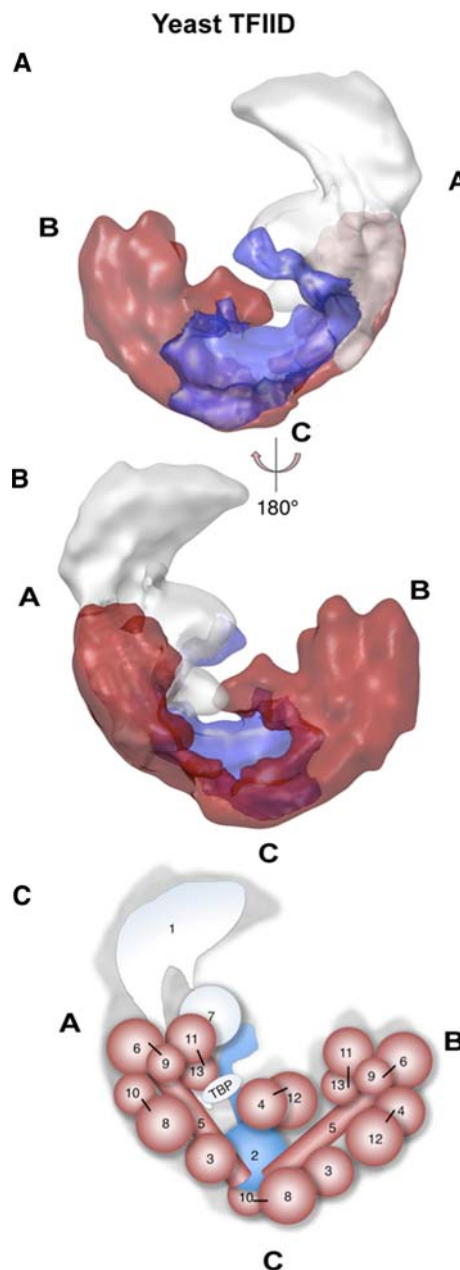
core composition is well conserved between yeast and mammals [5, 17]. This conservation has allowed the adoption of a unified nomenclature, TAF1-TAF14, that clearly designates TAF orthologues in different species [18]. Yeast, TFIID comprises 14 TAFs, and with the exception of TAF14 the orthologues of each of these proteins have been identified as *bone fide* TFIID subunits in *Drosophila* and mammals.

Biochemical, structural and genetic techniques have all been used to address the molecular organisation of TFIID. An initial analysis of the amino acid sequences of TAF6, 9 and 12 revealed a striking similarity to the core H4, H3 and H2B histones, suggesting the existence of a histone octamer structure within TFIID [19]. Further studies showed however that 9 of the 13 TAFs contained a histone fold domain (HFD) specifying the formation of five distinct TAF heterodimers within TFIID [5, 17]. These heterodimers were also found in the context of native TFIID, thus underlining that the HFD is a fundamental building block of TFIID [20].

Genetic, biochemical and electron microscopy (EM) experiments clearly identified two classes of subunits based on their stoichiometries [20–22]. TAF1, TAF2, TAF7 and TBP are present as a single copy, whereas almost all of the others are present in at least two copies. Therefore, the total number of heterodimers and their distribution within TFIID (see below) show that the overall organisation is more complex than a simple octamer-like core.

Despite the importance of TFIID in transcription initiation, there is a remarkable paucity of structural data on TFIID subunits, and little is known of the mechanisms directing its supramolecular assembly. Structural information at the atomic level is currently only available for a few TAF subdomains (see, for example, [23–25]). EM has however provided detailed information on the overall shape of native TFIID and the localisation of TAF subunits. Single molecule analysis and image reconstruction show that yeast (y)TFIID and human (h)TFIID have a similar overall organisation comprising three major lobes linked by connecting regions to form a ‘horseshoe’-shaped molecular clamp [20, 26, 27]. To date these studies are limited to low resolution, 32Å for hTFIID [28] and 23Å for yTFIID (see Fig. 1a; [29]), which are not sufficient to allow docking with the known crystal structures. Structural heterogeneity, in particular variable TAF2 content, and dynamic rearrangements of the complex are limitations to improved resolution.

EM coupled to immunolabelling of complete TFIID or of subpopulations have allowed the localisation of individual TAFs within this structure defining the composition of the lobes, each of which comprises a unique TAF combination [20, 22, 29] (see Fig. 1c). TAFs present as a single copy are found mainly in lobes A and C. The



**Fig. 1** Structural organisation of yeast TFIID; (a, b) two opposite views of the TFIID complex showing the location of TAF1, TAF7 and TBP (white), TAF2 (blue) and a quasi symmetric core module (red). c Schematic representation of the TAF localisation within TFIID. The approximate positions of the subunits are derived from antibody-labelling experiments. The size of the spheres is proportional to the molecular mass of the proteins. Black lines represent documented protein-protein interactions of histone fold domain containing TAFs

C-terminus of TAF1 and TAF7 localise on the top of lobe A, while the N-terminus of TAF1 and TBP localise at the interface of lobes A and C. TAF2 is also located in the interface between lobes A and C with its N-terminus close to the C-lobe and its C-terminus close to the A-lobe where

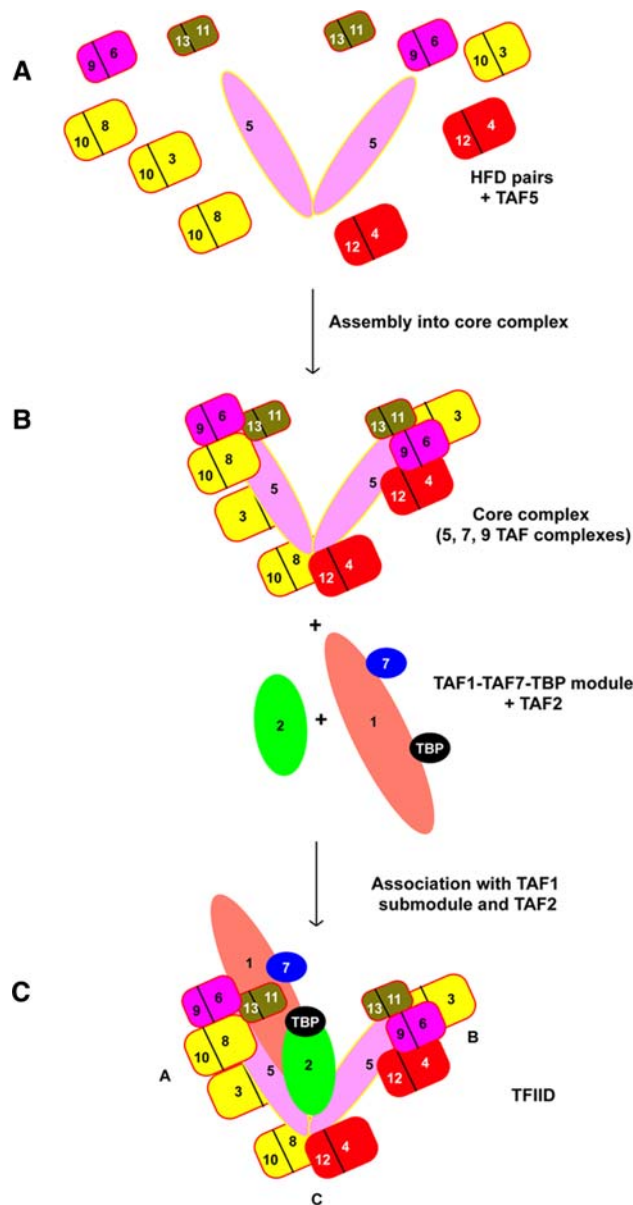


it could interact with TAF1. TAFs that are documented to be present in at least two copies are located in the bottom part of lobe A and in lobes B and C. The localisation of the HFD-containing TAFs in native  $\gamma$ TFIID is consistent with the previous genetic, biochemical and structural data on the formation of specific TAF heterodimers. An exception is the TAF3/TAF10 heterodimer. In EM immunolabelling, TAF3 can be observed in lobe B, but no associated TAF10 can be seen, while in lobe A it is not possible to discriminate between the TAF3/TAF10 and the TAF8/TAF10 heterodimers.

Interestingly, when the potential protein densities of TAF1, TAF2, TAF7 and TBP are removed from the 23Å  $\gamma$ TFIID model (blue and white in Fig. 1a, b), the shape of the remaining structure is reminiscent of that of a stable in vitro reconstituted complex composed of TAF5 and the three HFD-TAF heterodimers TAF4/12, TAF6/9 TAF8/10 [22] and presents an almost symmetric crescent-shaped structure (red in Fig. 1a). Residual asymmetry could result from the binding of the remaining HFD-containing TAFs. Altogether, the higher resolution structure and the labelling studies strongly suggest that the 3D architecture of TFIID is composed of two subcomplexes: (1) a core complex containing TAF5 and most of the HFD-containing TAF heterodimers (TAF6/9, TAF4/12, TAF8/10, TAF11/13) that adopts a crescent-shaped twofold symmetric structure; (2) a subcomplex containing TAF1, TAF7, TAF2 and TBP that is recruited to the core complex to form full TFIID (Fig. 2). The TAFs shared between TFIID and SAGA are all found in the core complex. In SAGA, the TAF8/TAF10 and TAF4/TAF12 heterodimers are replaced by the SPT7/TAF10 and ADA1/TAF12 heterodimers. This core domain is therefore pivotal in the assembly of both of these complexes. Perhaps competition between the SAGA and TFIID-specific heterodimerisation partners regulates the relative abundance of each of these complexes in the cell.

TBP is located in the linker region that lies within the major cavity of the clamp and is flanked by TAF1, TAF2, TFIIA and TFIIB, suggesting that this is the principal DNA binding site [27]. The structure of TFIID complexed with DNA has not yet been determined, but would be extremely informative and help to put in perspective the multiple TAF-DNA-chromatin interactions that have been described (see below).

TFIID structure may not be static, but rather appears to be dynamic as distinct 'open' and 'closed' conformations have been observed. Cryo-electron microscopy identifies two distinct conformations where the relative positions of the three lobes change in a coordinated and reproducible fashion resulting in an opening or closing of the central cavity [28]. The significance of these distinct states in terms of function or PIC formation has not been elucidated.



**Fig. 2** Schematic model for TFIID assembly. **a** The histone fold containing TAFs form heterodimers and associates with TAF5 to form the core complex. The ability of TAF5 to homodimerise is still open to question, and it is as yet unclear whether the histone-like heterodimers associate to form higher order 'lobe' structures in the absence of TAF5. **b** TAF5 and the histone-like heterodimers associate to form the core complex. Variants of this complex in the form of 5 TAF (*Drosophila* S2 cells), 7 TAF (baculovirus reconstitution in vitro) or 9 TAF (yeast in vivo) complexes have been described and are discussed in the text. **c** The core complex interacts with the TAF1-TAF7-TBP sub-module to form TFIID. TAF2 may associate independently with TFIID through interaction with TAF1, but may not be always present in TFIID

Also it is not yet clear whether these distinct states result from an inherent flexibility in the structure or whether they represent TFIID with different subunit compositions or TAFs bearing post-translational modifications.

Evidence that TAF-composition may influence TFIID topology comes from comparison of the EM structure of distinct TFIIDs containing either two molecules of TAF4 or TAF4 along with its paralogue TAF4b. TAF4 and TAF4b are encoded by related genes and share a central conserved region I, called the TAFH domain, as well as the larger conserved C-terminal region II containing an atypical HFD. TAF4 and TAF4b both heterodimerise with TAF12 and interact with TFIIA [24, 30–33]. A comparison of the EM structures of TFIID from HeLa cells containing predominantly TAF4 with that purified from Daudi cells containing TAF4 and TAF4b shows that the complexes containing TAF4b adopt a more open conformation [34]. This difference is achieved by changes in the relative positions of lobes A and B. TAF4b localises to lobes B and C, and TFIIA spans the lobe A–C channel. Consequently, it has been proposed that incorporation of TAF4b into TFIID leads to an ‘open’ conformation of lobe A, facilitating TAF4b interaction with TFIIA and various transcriptional activators. These observations thus provide evidence that differences in subunit composition affect TFIID conformation.

The status of TAF2 as a bona fide subunit of TFIID has been under debate since it was either missing or present at a reduced molar ratio in hTFIID preparations [35]. A more recent study [29] shows that the presence or absence of TAF2 also affects TFIID structure. Analysis of immunopurified yTFIID preparations that are heterogeneous in TAF2 content revealed complex conformational changes of the TFIID structure where at least four distinct states could be observed. The presence of TAF2 selectively stabilises one of these conformations, whereas in the absence of TAF2 significant domain reorganizations were observed, especially in the central lobe, as was noted for hTFIID. The TAF2-TFIID interaction is therefore labile, and while TAF2 can be dissociated without compromising TFIID integrity, its presence or absence can have a major influence on TFIID conformation. How these differences in subunit composition and in conformational state influence TFIID recognition and recruitment to promoters with different core sequences remains to be fully understood. The possibility that the structure of TFIID can be adapted to allow for the recognition of a large variety of promoters, each with distinct activator binding site distributions, is particularly attractive.

### Several TAFs are critical for assembly and stability of TFIID

Although we now have a better understanding of TFIID subunit composition and organisation, the mechanism of assembly is much less well understood. An initial series of

in vitro reconstitution studies put forward the idea that TAF1 formed a scaffold for TFIID assembly through its interactions with TBP and several other TAFs [36]. However, these studies were performed before the full subunit composition of TFIID was determined, and several TAF subunits were not included in these assays. More recent results reveal a more complex mechanism.

Genetic and biochemical studies in yeast have addressed the role of TAF1 in TFIID assembly. Biochemical analysis of a series of pseudodiploid yeast strains harbouring deletion mutants in a tagged allele of TAF1 shows that deletion of the N-terminal domain of TAF1 leads to diminished association with TBP, but not with the TAFs [37]. In contrast, deletion of the region between amino acids 200–303 leads to a loss of interaction with all of the TAFs except TAF7, but does not affect interaction with TBP. These data are consistent with the observation that the N-terminus of TAF1 (the TAND domain) interacts with TBP [38], while TAF7 interacts directly with several regions in the C-terminal portion of TAF1 [39, 40].

Two hybrid assays further show that the TAF1 200–303 region interacts with TAF4 and TAF6, and biochemical analysis shows that expression of this region alone is sufficient to nucleate the formation of a partial TFIID complex. Temperature-sensitive (TS) TAF1 mutants with amino acid substitutions in this region were also isolated and shown to affect interaction of the TAF1/TBP/TA7 module with the other TAFs. However, immunoprecipitation of TAF4 from the mutant strain shows that although interaction with TAF1 is lost, TAF4 is associated with other TAFs in a stable complex.

TFIID stability has also been investigated in *Drosophila* Schneider 2 (S2) cells where expression of TAF subunits was disrupted by siRNA. In this approach, the effect of siRNA knock-down of individual TAF subunits on TFIID integrity and an associated proteolytic degradation of the other TAFs was evaluated [41]. Knock-down of TAF1 did not destabilise the other tested TAFs, and in its absence a stable complex comprising TAF4, TAF5, TAF6, TAF9 and TAF12 was observed. In contrast, siRNA knock-down of TAF4 or its heterodimerisation partner TAF12 leads to degradation of TAF1 and most other TAFs, with the exception of TAF2. This observation shows that TAF4 is critical for TFIID assembly consistent with the fact that it is present in two of the globular lobes. Similar observations were made upon knock-down of TAF5 and TAF6 that also led to degradation of TAF1, suggesting that TFIID integrity is compromised. Moreover, expression of the conserved HFDs of TAF4 and TAF6 is sufficient to complement the loss of the corresponding native proteins for assembly and stability of the TFIID complex.

The idea that TAF4 is critical for TFIID assembly/stability is relevant to several functional observations. For

example, in *C. elegans*, siRNA of TAF4 had the most potent effect on transcription compared to other TAFs [42]. More recently it has been shown that TAF4 is a target of the *C. elegans* oocyte maturation (OMA)-1 and OMA-2, zinc finger proteins of the CCCH class that bind to RNA and are important for oocyte maturation and early embryonic development [43]. In early embryos, the OMA proteins bind to TAF4 by mimicking the HFD domain interaction with TAF12 and sequestering TAF4 in the cytoplasm, thus silencing transcription at these early stages. Similarly, it has been proposed that sequestration of TAF4 by variants of the Huntington protein carrying expanded polyglutamine domains interferes with transcription and contributes to neurodegenerative disease [44, 45]. Given the important role of TAF4 in TFIID assembly, its sequestration would therefore provide a mechanism to control the overall function of TFIID through targeting of a single subunit.

Further insight into TFIID assembly comes from reconstitution of TAF complexes by baculovirus coinfections. In these studies, a stable and homogenous seven-subunit complex comprising TAF5 and the TAF4/12 TAF6/9 and TAF8/10 HFD heterodimers could be isolated [22]. In the absence of TAF5, the stability of this complex is compromised. EM images of this sub-complex show a trilobed structure that is reminiscent of the TFIID core domain. These observations suggest that TAF5 may dimerise through its N-terminal domain, while its C-terminal domain containing the beta-transducin (WD40) repeats interacts with the HFD TAFs to form the globular lobes. The dimerisation of TAF5 is still under discussion as conflicting results have been obtained [25, 46].

Do the HFD-TAF heterodimers assemble by themselves into higher order 'lobe' structures or does TAF5 play an active role in this process? TS mutations in the WD40 repeats of  $\gamma$ TAF5 were found to cause broad transcription defects showing that these motifs are critical in maintaining the integrity of both the TFIID and SAGA complexes [47]. However, in S2 cells, TAF5 knock-down leads to degradation of TAF1, but not other TAFs. Thus, either some type of higher order complexes form in the absence of TAF5, or perhaps heterodimer formation ensures their stability even without their assembly in 'lobe' structures. The full set of interactions required for the formation of the core TAF subcomplex therefore remains to be determined.

To further complicate matters, genetic experiments in mammalian cells revealed that TAF10 is critical for TFIID integrity. Somatic inactivation of TAF10 by Cre-mediated deletion in F9 embryonic carcinoma cells, in early embryos and in adult liver have shown that loss of TAF10 leads to disassembly of TFIID [48–50]. In contrast, inactivation of TAF4 in cells that express TAF4b facilitates integration of TAF4b into TFIID without affecting its stability [51].

Taken altogether, the studies in yeast, *Drosophila* and in vitro all suggest a similar modular model of TFIID organisation and assembly. In each approach, a stable HFD-containing TAF subcomplex can be formed lacking TAF1, TAF2, TAF7 and TBP (see Fig. 2a, b). In yeast, this complex comprises most other TAFs and corresponds to the pseudo-symmetric core complex described by EM. In *Drosophila*, a subcomplex containing only five TAFs was described, but the presence of several others was not assayed and so could also be present in this complex. There is compelling evidence that TAF4, TAF5 and TAF10 all play primary roles in the assembly and stability of the core complex and hence of TFIID. The core subcomplex associates via TAF4 and TAF6 with a second submodule comprising TAF1-TBP-TAF7 to assemble the TFIID complex (see Fig. 2c).

These observations can be better understood in the context of the known EM structure. TAF4 and TAF10 are both present in at least two distinct lobes. Their loss presumably destabilises these lobes, leading to release of the TAF1-TBP-TAF7 module, which in S2 cells is unstable and degraded. TAF5 seems to link lobes A and B together and may be required to assemble the HFD-heterodimers into stable lobe structures. In contrast, TAF1, TAF7 and TBP are located on top of the pseudo-symmetric TFIID core, and this module may be unstable in the absence of the other TAFs forming the core structure.

TAF2 is not essential for TFIID assembly, but probably associates with TFIID via interactions with TAF1. Moreover, TAF1 and TAF2 have been shown to form a subcomplex with TBP that specifically binds to promoter DNA in vitro [52]. Although the existence of such a subcomplex has yet to be demonstrated in vivo, the *P. falciparum* genome encodes orthologues of TAF1, TAF2 and TBP, but not of other TAFs (with the possible exception of TAF10), further highlighting the idea of a functional TAF1-TAF2-TBP subcomplex [83].

The precise roles of TAF3, TAF7, TAF11 and TAF13 in the stability and assembly of TFIID remain to be investigated. As mentioned above, TAF7 is known to interact with TAF1 and to negatively regulate its HAT activity [39]. TAF7 also interacts with TFIID and Positive Transcription Elongation Factor b (pTEFb) to inhibit their ability to phosphorylate the carboxy-terminal domain of the largest subunit of RNA polymerase II, but is released upon entry of the polymerase into the preinitiation complex [53]. TAF7 may therefore function as a check-point regulator suppressing premature transcription initiation [54]. These observations are consistent with the above model of TFIID organisation where TAF7 can dissociate from the complex without loss of integrity.

In conclusion, the above results show that TFIID stability depends on a complex set of interactions in which not

one, but several TAFs that form the core structure play a critical role.

### Functional TAF subcomplexes in male germ cells

While TFIID has been purified and studied from cell extracts, the question arises whether the five or seven TAF subcomplexes of the core structure naturally exist in cells and whether they have any specific functions or serve simply as assembly intermediates. Evidence that such subcomplexes exist and may have specific functions comes from studies of transcription in male germ cells.

In addition to the core TAFs described above, the *Drosophila* genome encodes five additional testis (t)TAF paralogues that are specifically expressed in a coordinated manner in spermatocytes. *No hitter (nht)* (TAF4L), *canonball (can)* (TAF5L), *meiosis I arrest (mia)* (TAF6L), *spermatocyte arrest (sa)* (TAF8L) and *ryan express (rye)* (TAF12L) are paralogues of TAF4, TAF5, TAF6, TAF8 and TAF12, respectively [55, 56]. These paralogues share the same structural domains as the corresponding core TAFs and therefore may form stable 5 or 7-TAF complexes. Such an organisation would require additional partners (TAF9 and TAF10a or TAF10b) or as yet unidentified components as heterodimerisation partners for the HFDs of dTAF6L and dTAF8L.

A biochemical analysis of their function shows that most of the tTAF protein localises to the nucleolus where they are required for a nucleolar relocalisation of the polycomb (Pc), polyhomeotic (Ph) and dRING (really interesting new gene) subunits of the *Drosophila* PRC1 repressor complex that takes place in spermatocytes [57]. Sequestration of PRC1 facilitates activation of a series of genes required for germ cell differentiation.

The mechanism described above is not conserved in mammals. However, the mouse genome encodes two TAF paralogues involved in spermatogenesis. TAF7L is a protein with high sequence similarity to somatically expressed TAF7 that is expressed in spermatogonia and in early primary spermatocytes, where it is localised in the cytoplasm [40, 58]. During spermatocyte development, TAF7L is imported into the nucleus and accumulates strongly in post-meiotic round spermatids where it is associated with TBP.

The import of TAF7L into the nucleus is coordinated with both a loss of TAF7 expression and a potent upregulation of TBP. In addition, the expression of TAF4 and TAF10 are strongly downregulated in round spermatids. As a consequence, haploid round spermatids strongly express TAF7L and TBP, but little or no TAF4 and TAF10, suggesting that they do not contain high levels of intact TFIID. In agreement with this, TAF6 is not associated with TBP in

these cells. Therefore, in contrast to what has been discussed above, in haploid spermatids a stable and functional TAF7L-TAF1-TBP complex may exist.

Furthermore, a critical role for TAF1 in spermatogenesis is suggested by the observation that in old world monkeys, apes and humans, there is a retrotransposed copy of TAF1 encoding TAF1L that has been selected to evade meiotic sex chromosome inactivation of the ancestral TAF1 gene present on the X chromosome [59]. The study of male germ cells therefore reveals potential functions for both the HFD-containing TAFs and the TAF1-TAF7-TBP submodules.

### Multiple interactions with DNA, activators and covalent histone modifications are involved in promoter recruitment of TFIID

While TBP binding to canonical TATA elements has been well studied both biochemically and structurally [60], the vast majority of promoters do not contain a recognisable TATA element. A number of observations indicate that TFIID recognises promoters through additional interactions of TAFs with other DNA promoter elements and also through interaction of TAFs with acetylated and methylated histone lysine residues.

Genetic evidence showing that TAFs contribute to promoter recognition was first obtained studying the effect of TS TAF mutations in yeast. A TS mutation in TAF1 affects the expression of only a limited number of promoters, for example those of cyclin (CLN)2 or ribosomal protein (RP)S5 [61]. The dependence of these promoters on functional TAF1 was shown to reside in the core promoter sequence and not in the upstream activating sequence (UAS) element where the gene regulatory factors interact. The region conferring TAF1 dependence was mapped to a region around the TATA element, but a precise sequence element could not be determined.

Using biochemical approaches, multiple TAF-promoter DNA contacts due to 'wrapping' of DNA around TFIID [62] have been described along with several more specific TAF-promoter interactions. Cross-linking and electrophoretic mobility shift assay-mediated binding site selection has been used to show that the recombinant TAF1-TAF2 complex has specificity for interaction with DNA containing the Initiator (Inr) sequence, a loosely defined sequence with a 5'-YC/TANT/AYY-3' consensus [35, 52]. In vitro reconstitution and transcription studies also indicate a role for TAF2 in transcription from Inr containing promoters. On the other hand, cross-linking has been used to show that the TAF6-TAF9 heterodimer may contact the downstream promoter element (DPE), an element found downstream of the transcription start site that has been well

characterised in a subset of *Drosophila* promoters and is likely also present in mammalian promoters (for review see [63, 64]). The presence of a DPE is often associated with that of an Inr, and both elements functionally cooperate. Most of the TAFs involved in the interaction with promoter elements are located at the base of lobe A (Fig. 1c) and are therefore likely to define the major DNA binding interface within TFIID.

Promoters containing combinations of TATA, Inr or DPE or other well-defined and -characterised promoter elements generally belong to the 'sharp' class characterised by a single transcription start site [65–67]. The majority of promoters, however, fall into the 'broad' class characterised by multiple start sites, the lack of identifiable promoter elements and the presence of a CpG island. How does TFIID get recruited and stabilised at such promoters? Several lines of evidence suggest that protein-protein interactions may play a critical role.

A large body of evidence indicates that transcription activator proteins can interact with TAFs and recruit TFIID to the promoter. Well-characterised examples are the interactions between SP1 or CREB (cyclic AMP response element binding) and TAF4 [68–70]. More importantly it has been shown that activators can interact with TAFs in the context of native TFIID and recruit it to promoters [71, 72].

In addition to the above, interactions between TAFs and covalently modified histone lysines have been identified that probably also play an important role in TFIID promoter recognition. A large body of evidence has clearly established that transcriptionally active promoters are characterised by the presence of specific covalent modifications of histone residues (for recent reviews, see [73, 74]). Several marks are tightly associated with active promoters, such as the trimethylation of lysine 4 on histone H3 (H3K4me3) and acetylation of lysine 9 on histone H3 (H3K9ac). These and other covalent modifications are recognised by structural domains present in proteins that interact with chromatin (for an extensive review, see [73]). Amongst these domains, three are present in the TAFs, a double bromodomain in TAF1, a plant homeodomain (PHD) in TAF3 and a WD40 repeat domain in TAF5.

The first TAF domain to be characterised as interacting with modified histones was the double bromodomain in TAF1. Structural and biochemical studies showed that the TAF1 bromodomains did not bind the unacetylated H4 tail, but recognised with low affinity the H4 tail acetylated on K16, and with a much higher affinity H4 tails doubly acetylated at positions K5/K12 or K8/K16 [75]. These marks are characteristic of actively transcribed euchromatin, suggesting that TFIID may be recruited and/or stabilised at active promoters through interaction of TAF1 with nucleosomes bearing these modifications. As TAF1

has also been reported to be a histone acetyl transferase [76], it may be both a reader and writer of histone modifications.

Metazoan TAF3 comprises a C-terminal PHD domain. Stable isotope labelling with amino acids in cell culture (SILAC) proteomics experiments identified TAF3 and other TFIID subunits interacting specifically with histone tails carrying the H3K4me3 modification via this PHD domain [77]. Previously it had been reported that the PHD domains of subunits of a transcriptional repressor complex (inhibitor of growth, ING2) and the chromatin remodelling complex (bromodomain PHD finger transcription factor, BPTF) also interacted specifically with this mark (see [73] and references therein). However, the affinity of TAF3 for H3K4me3 is 10–20-fold higher than for ING2 and BPTF [77, 78]. In addition, the combination of H3K4me3 with K9 and K14 acetylation strongly increases the interaction with TFIID, presumably through recognition of the acetylated residues by the TAF1 double bromodomains. The affinity of the TAF3 PHD domain with the H3K4me3 mark is significantly higher than the bromodomains for the acetylated residues, suggesting that the PHD-mediated interaction is dominant, while the bromodomain interactions further stabilise the histone-TFIID complex. In contrast, binding of TAF3 is reduced when the H3K4me3 modification is coupled with asymmetric dimethylation of the adjacent arginine (R)2 residue, a mark that anti-correlates with transcriptional activity.

Together these studies suggest a model where interactions with transcriptional activators and with appropriately modified histone tails act to recruit and stabilise TFIID at active promoters. These interactions can complement those of TBP and TAFs with promoter DNA elements or substitute for these interactions at CpG island-type promoters. Interestingly, although these histone modifications exist in yeast, yTAF1 and yTAF3 lack the bromo and PHD domains, respectively.

As a further complication to the above model, the TAF4 subunit has been shown to interact with heterochromatin protein (HP)1 $\alpha$  and HP1 $\gamma$ , but not HP1 $\beta$  [79]. The HP1s interact with a variety of partner proteins through a conserved PXVXL motif (for recent reviews, see [80, 81]). Such a motif has been identified in the C-terminal domain of TAF4 and shown to be required for interaction with HP1. The HP1 proteins were first described as heterochromatin-associated proteins recruited through interactions with H3K9me3 and involved in gene silencing. More recent studies however have also shown that HP1s can be found in euchromatin and be involved in both activation and repression. It therefore remains to be determined whether the TAF4-HP1 interaction contributes to recruitment of TFIID to a subset of active promoters or is involved in TAF-dependent repression (see below).

Two further TAF domains may be involved in chromatin interactions. As mentioned above, the C-terminus of TAF5 comprises a WD40 repeat domain. WD40 domains form a beta-propellor type structure. The WD40 domain in the WDR5 subunit of the SET1 [*Drosophila* Su(var) 3-9, Enhancer of zeste (E(z)), and Trithorax (trx)] methylation complexes has been shown to interact with H3K4 in the di-, tri- and un-methylated states, with a preference for dimethylation and also with the H4 tail (for review, see [82]). It will be interesting to determine whether the TAF4 WD40 domain can also interact with modified histones and provide yet another surface for TFIID interaction with chromatin.

Lastly, the amino terminus of TAF2 comprises an enzymatically inactive amino-peptidase fold [83]. A similar fold exists in the suppressor of P-element transcription (SPT)16 subunit of the factor required for transcription elongation on chromatin templates (FACT) complex and has been shown to interact with the globular core and tails of H3 and H4 [84]. It remains to be determined whether TAF2 can also interact with histones.

### Is TFIID required to maintain promoters in an active state?

Once recruited to active promoters, what is the function of TFIID? In the classical view TFIID nucleates PIC formation and promotes transcription initiation. In vitro studies also suggested that TFIID remains at the promoter once transcription has been initiated as part of the reinitiation scaffold [85]. Several recent observations suggest a rather different model.

Inactivation of TBP in early mouse embryos leads to arrest of proliferation at the 32 cell stage after depletion of the maternal TBP and subsequent apoptosis [86]. Nevertheless, the levels of pol II transcription in the arrested cells of the TBP knockout embryo were comparable to that seen in normal embryos at the same stage. These observations led to the idea that TBP/TFIID may be differentially required in proliferating and post-mitotic cells [87]. During cell division, active transcription complexes are dissociated and must be reformed in the daughter cells. TBP plays a central role in this process as it remains associated with a subset of promoters during mitosis [88]. Comprehensive chromatin immunoprecipitation experiments show that TBP remains bound to many chromosomal sites during mitosis, thus 'bookmarking' promoters for re-expression in interphase [89]. TBP interacts with the protein phosphatase (PP)2A to locally inactivate condensin at these sites and inhibit their compaction. The results of the TBP knockout in post-mitotic cells suggest that once the active transcription complex has been established, TBP/TFIID may

be dispensable for the reinitiation step and persistence of the active state, but that cells can no longer undergo mitosis.

This model is also supported by observations of genetic knockout of TAF subunits. For example, inactivation of TAF10 in proliferating F9 embryonal carcinoma cells leads to TFIID disassembly, cell cycle arrest and apoptosis, whereas non-dividing differentiated F9 cells survive [49]. Similarly, TAF10 is required in the proliferating inner cell mass of the early embryo, but not in post-mitotic trophoblast cells [48]. Again, these results suggest an essential function for TFIID in proliferating, but not post-mitotic cells.

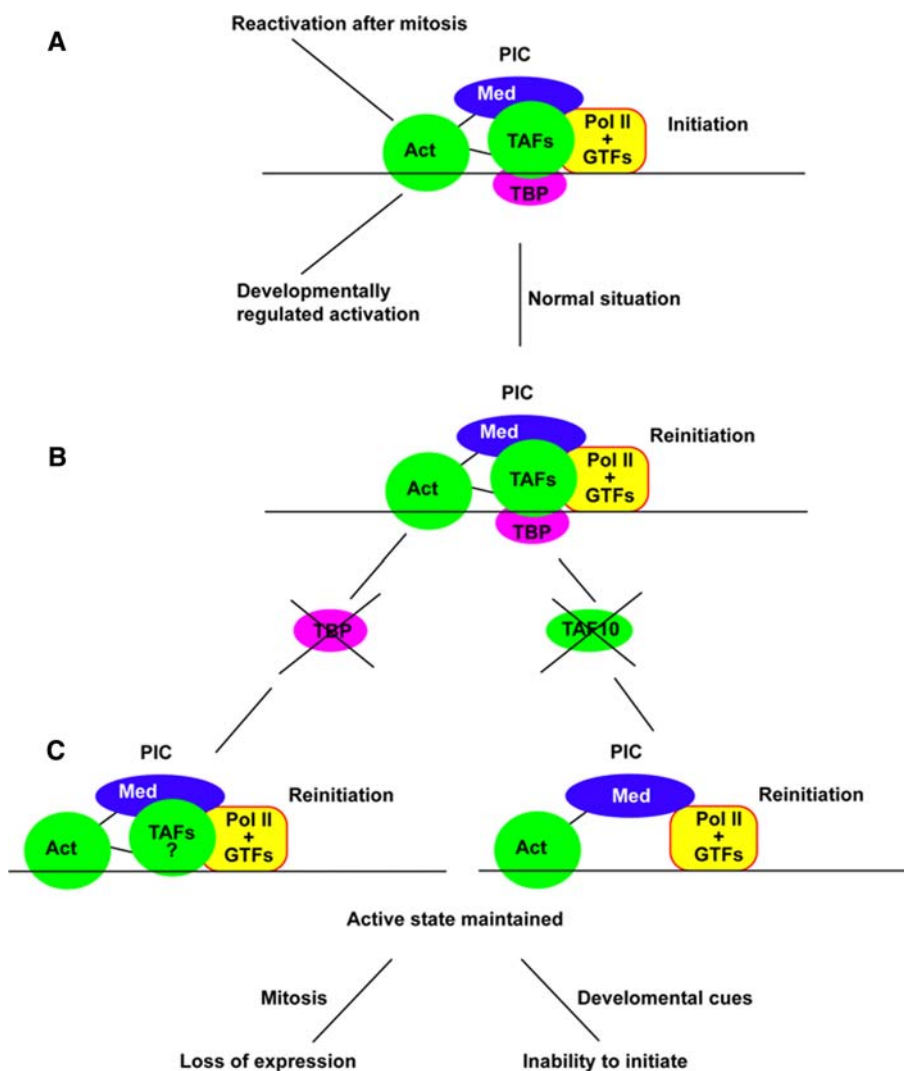
Further evidence for this idea has been provided by the inactivation of TAF10 in mouse hepatocytes. Knockout of TAF10 in proliferating embryonic hepatoblasts affects their proliferation and differentiation and leads to subsequent defective liver organogenesis [50]. In contrast in the adult liver, loss of TAF10 does not lead to an immediate apoptosis of the non-dividing hepatocytes. Transcriptome analysis showed that the expression of only a small number of genes was affected corresponding to those that are specifically activated in the adult hepatocytes, while genes that were already active in embryonic hepatocytes were not affected upon TAF10 inactivation. A biochemical analysis confirmed that TFIID is disassembled in the absence of TAF10 and that neither TBP nor TAFs are present at the promoters of active genes and are not recruited at the appropriate time to the developmentally regulated promoters.

In contrast, several genes expressed in embryonic hepatocytes and normally silenced in adult cells are re-expressed in the absence of TAF10, suggesting that TFIID plays an active repressive role. This idea is supported by the fact that while pol II is lost from these promoters during the normal developmental silencing, TAFs remain associated with these promoters. Perhaps it is in this type of situation that the TAF4-HP1 interaction would be required to recruit or stabilise TFIID.

Together the above results suggest a model whereby intact TFIID is required to mediate gene activation either after mitosis or in a developmentally regulated fashion (Fig. 3). However, once a promoter is active and even although in normal circumstances TFIID remains associated with the promoter, its function is no longer required to maintain the promoter in an active state and to promote transcription reinitiation.

One set of observations that do not obviously fit this model comes from the inactivation of TAF10 in the basal keratinocytes of the epidermis [90]. Loss of TAF10 in adult keratinocytes has no obvious effect on skin homeostasis, hair cycle or wound healing. Presumably loss of TAF10 leads to TFIID disassembly in keratinocytes as in other cell

**Fig. 3** Function of TFIID.  
**a** In wild-type cells TFIID is recruited along with pol II and the general transcription factors to facilitate PIC formation and promoter activation following mitosis or upon developmental cues. The transcriptional activators (*Act*), mediator complex (*Med*), *TBP* and *TAFs*, *RNA pol II* and the other general transcription factors (*Pol II + GTFs*) are all schematically depicted.  
**b** In normal cells, TFIID remains associated with the active promoter during the reinitiation step.  
**c** Following genetic inactivation of *TBP* or *TAF10*, the promoter remains in an active state until mitosis when the PIC is disassembled upon chromatin condensation and is unable to reform at the next interphase. Post-mitotic cells are unable to activate new sets of genes upon developmental cues or in response to signalling pathways. While inactivation of *TAF10* has been shown to lead to loss *TBP* and *TAFs* at active promoters, it is not yet known whether *TAFs* remain at promoters in the absence of *TBP*.



types, yet this has no detrimental effect on the above processes all of which require extensive cell proliferation.

Much of our understanding of TFIID's role in transcription regulation has come from studies in cultured animal cells or in yeast, both of which are proliferating systems. The observations cited above concerning inactivation of *TBP* or *TAF10* in vivo show that the lessons we have learned from cell culture or yeast cannot be simply extrapolated to more complex physiological situations. These considerations should modify our classical view of the role of TFIID in transcription regulation in post-mitotic cells.

**Acknowledgments** Work in the Davidson and Schultz laboratories is supported by grants from the CNRS, the INSERM, the Association pour la Recherche contre le Cancer and the Ligue Nationale et Départementale Région Alsace contre le Cancer. ID is an 'équipe labélisée' of the Ligue Nationale contre le Cancer, the Fondation pour la Recherche Médicale (FRM), the ANR the European SPINE programme QLG-CT-00988 and European Union grant

RTN-2001-00026 and integrated programme grant EuTRACC. E. Cler was supported by a fellowship from the FRM, and G. Papai from the ANR.

## References

- Andrau JC, van de Pasch L, Lijnzaad P, Bijma T, Koerkamp MG, van de Peppel J, Werner M, Holstege FC (2006) Genome-wide location of the coactivator mediator: binding without activation and transient Cdk8 interaction on DNA. *Mol Cell* 22:179–192
- van Werven FJ, van Bakel H, van Teeffelen HA, Altelaar AF, Koerkamp MG, Heck AJ, Holstege FC, Timmers HT (2008) Cooperative action of NC2 and Mot1p to regulate TATA-binding protein function across the genome. *Genes Dev* 22:2359–2369
- Burley SK, Roeder RG (1996) Biochemistry and structural biology of transcription factor IID (TFIID). *Annu Rev Biochem* 65:769–799
- Albright SR, Tjian R (2000) TAFs revisited: more data reveal new twists and confirm old ideas. *Gene* 242:1–13
- Gangloff Y, Romier C, Thuault S, Werten S, Davidson I (2001) The histone fold is a key structural motif of transcription factor TFIID. *Trends Biochem Sci* 26:250–257

6. Matangkasombut O, Auty R, Buratowski S (2004) Structure and function of the TFIID complex. *Adv Protein Chem* 67:67–92
7. Thomas MC, Chiang CM (2006) The general transcription machinery and general cofactors. *Crit Rev Biochem Mol Biol* 41:105–178
8. Timmers HT, Tora L (2005) SAGA unveiled trends. *Biochem Sci* 30:7–10
9. Pijnappel WW, Timmers HT (2008) Dubbing SAGA unveils new epigenetic crosstalk. *Mol Cell* 29:152–154
10. Baker SP, Grant PA (2007) The SAGA continues: expanding the cellular role of a transcriptional co-activator complex. *Oncogene* 26:5329–5340
11. Dynlacht BD, Hoey T, Tjian R (1991) Isolation of coactivators associated with the TATA-binding protein that mediate transcriptional activation. *Cell* 66:563–576
12. Kokubo T, Takada R, Yamashita S, Gong DW, Roeder RG, Horikoshi M, Nakatani Y (1993) Identification of TFIID components required for transcriptional activation by upstream stimulatory factor. *J Biol Chem* 268:17554–17558
13. Brou C, Chaudhary S, Davidson I, Lutz Y, Wu J, Egly JM, Tora L, Chambon P (1993) Distinct TFIID complexes mediate the effect of different transcriptional activators. *EMBO J* 12:489–499
14. Moqtaderi Z, Yale JD, Struhl K, Buratowski S (1996) Yeast homologues of higher eukaryotic TFIID subunits. *Proc Natl Acad Sci USA* 93:14654–14658
15. Poon D, Weil PA (1993) Immunopurification of yeast TATA-binding protein and associated factors. Presence of transcription factor IIIB transcriptional activity. *J Biol Chem* 268:15325–15328
16. Reese JC, Apone L, Walker SS, Griffin LA, Green MR (1994) Yeast TAFIIS in a multisubunit complex required for activated transcription. *Nature* 371:523–527
17. Gangloff YG, Pointud JC, Thuault S, Carre L, Romier C, Muratoglu S, Brand M, Tora L, Couderc JL, Davidson I (2001) The TFIID components human TAF(II)140 and *Drosophila* BIP2 (TAF(II)155) are novel metazoan homologues of yeast TAF(II)47 containing a histone fold and a PHD finger. *Mol Cell Biol* 21:5109–5121
18. Tora L (2002) A unified nomenclature for TATA box binding protein (TBP)-associated factors (TAFs) involved in RNA polymerase II transcription. *Genes Dev* 16:673–675
19. Xie X, Kokubo T, Cohen SL, Mirza UA, Hoffmann A, Chait BT, Roeder RG, Nakatani Y, Burley SK (1996) Structural similarity between TAFs and the heterotetrameric core of the histone octamer. *Nature* 380:316–322
20. Leurent C, Sanders S, Ruhlmann C, Mallouh V, Weil PA, Kirschner DB, Tora L, Schultz P (2002) Mapping histone fold TAFs within yeast TFIID. *EMBO J* 21:3424–3433
21. Sanders SL, Garbett KA, Weil PA (2002) Molecular characterization of *Saccharomyces cerevisiae* TFIID. *Mol Cell Biol* 22:6000–6013
22. Leurent C, Sanders SL, Demeny MA, Garbett KA, Ruhlmann C, Weil PA, Tora L, Schultz P (2004) Mapping key functional sites within yeast TFIID. *EMBO J* 23:719–727
23. Birck C, Poch O, Romier C, Ruff M, Mengus G, Lavigne AC, Davidson I, Moras D (1998) Human TAF(II)28 and TAF(II)18 interact through a histone fold encoded by atypical evolutionary conserved motifs also found in the SPT3 family. *Cell* 94:239–249
24. Werten S, Mitschler A, Romier C, Gangloff YG, Thuault S, Davidson I, Moras D (2002) Crystal structure of a subcomplex of human transcription factor TFIID formed by TATA binding protein-associated factors hTAF4 (hTAF(II)135) and hTAF12 (hTAF(II)20). *J Biol Chem* 277:45502–45509
25. Romier C, James N, Birck C, Cavarelli J, Vivares C, Collart MA, Moras D (2007) Crystal structure, biochemical and genetic characterization of yeast and *E. cuniculi* TAF(II)5 N-terminal domain: implications for TFIID assembly. *J Mol Biol* 368:1292–1306
26. Brand M, Leurent C, Mallouh V, Tora L, Schultz P (1999) Three-dimensional structures of the TAFII-containing complexes TFIID and TFIIIC. *Science* 286:2151–2153
27. Andel F 3rd, Ladurner AG, Inouye C, Tjian R, Nogales E (1999) Three-dimensional structure of the human TFIID-IIA-IIB complex. *Science* 286:2153–2156
28. Grob P, Cruse MJ, Inouye C, Peris M, Penczek PA, Tjian R, Nogales E (2006) Cryo-electron microscopy studies of human TFIID: conformational breathing in the integration of gene regulatory cues. *Structure* 14:511–520
29. Papai G, Tripathi MK, Ruhlmann C, Werten S, Crucifix C, Weil PA, Schultz P (2009) Mapping the initiator binding TAF2 subunit in the structure of hydrated yeast TFIID Structure (in press)
30. Dikstein R, Zhou S, Tjian R (1996) Human TAFII 105 is a cell type-specific TFIID subunit related to hTAFIII130. *Cell* 87:137–146
31. Mengus G, May M, Carre L, Chambon P, Davidson I (1997) Human TAF(II)135 potentiates transcriptional activation by the AF-2s of the retinoic acid, vitamin D3, and thyroid hormone receptors in mammalian cells. *Genes Dev* 11:1381–1395
32. Thuault S, Gangloff YG, Kirchner J, Sanders S, Werten S, Romier C, Weil PA, Davidson I (2002) Functional analysis of the TFIID-specific yeast TAF4 (yTAF(II)48) reveals an unexpected organization of its histone-fold domain. *J Biol Chem* 277:45510–45517
33. Guermah M, Tao Y, Roeder RG (2001) Positive and negative TAF(II) functions that suggest a dynamic TFIID structure and elicit synergy with traps in activator-induced transcription. *Mol Cell Biol* 21:6882–6894
34. Liu WL, Coleman RA, Grob P, King DS, Florens L, Washburn MP, Geles KG, Yang JL, Ramey V, Nogales E, Tjian R (2008) Structural changes in TAF4b-TFIID correlate with promoter selectivity. *Mol Cell* 29:81–91
35. Kaufmann J, Smale ST (1994) Direct recognition of initiator elements by a component of the transcription factor IID complex. *Genes Dev* 8:821–829
36. Chen JL, Attardi LD, Verrijzer CP, Yokomori K, Tjian R (1994) Assembly of recombinant TFIID reveals differential coactivator requirements for distinct transcriptional activators. *Cell* 79:93–105
37. Singh MV, Bland CE, Weil PA (2004) Molecular and genetic characterization of a Taf1p domain essential for yeast TFIID assembly. *Mol Cell Biol* 24:4929–4942
38. Liu D, Ishima R, Tong KI, Bagby S, Kokubo T, Muhandiram DR, Kay LE, Nakatani Y, Ikura M (1998) Solution structure of a TBP-TAF(II)230 complex: protein mimicry of the minor groove surface of the TATA box unwound by TBP. *Cell* 94:573–583
39. Gegonne A, Weissman JD, Singer DS (2001) TAFII55 binding to TAFII250 inhibits its acetyltransferase activity. *Proc Natl Acad Sci USA* 98:12432–12437
40. Pointud JC, Mengus G, Brancorsini S, Monaco L, Parvinen M, Sassone-Corsi P, Davidson I (2003) The intracellular localisation of TAF7L, a paralogue of transcription factor TFIID subunit TAF7, is developmentally regulated during male germ-cell differentiation. *J Cell Sci* 116:1847–1858
41. Wright KJ, Marr MT 2nd, Tjian R (2006) TAF4 nucleates a core subcomplex of TFIID and mediates activated transcription from a TATA-less promoter. *Proc Natl Acad Sci USA* 103:12347–12352
42. Walker AK, Rothman JH, Shi Y, Blackwell TK (2001) Distinct requirements for *C. elegans* TAF(II)s in early embryonic transcription. *EMBO J* 20:5269–5279
43. Guven-Ozkan T, Nishi Y, Robertson SM, Lin R (2008) Global transcriptional repression in *C. elegans* germline precursors by regulated sequestration of TAF-4. *Cell* 135:149–160



44. Shimohata T, Nakajima T, Yamada M, Uchida C, Onodera O, Naruse S, Kimura T, Koide R, Nozaki K, Sano Y, Ishiguro H, Sakoe K, Ooshima T, Sato A, Ikeuchi T, Oyake M, Sato T, Aoyagi Y, Hozumi I, Nagatsu T, Takiyama Y, Nishizawa M, Goto J, Kanazawa I, Davidson I, Tanese N (2000) Expanded polyglutamine stretches interact with TAFII30, interfering with CREB-dependent transcription. *Nat Genet* 26:29–36
45. Freiman RN, Tjian R (2002) Neurodegeneration. A glutamine-rich trail leads to transcription factors. *Science* 296:2149–2150
46. Bhattacharya S, Takada S, Jacobson RH (2007) Structural analysis and dimerization potential of the human TAF5 subunit of TFIID. *Proc Natl Acad Sci USA* 104:1189–1194
47. Durso RJ, Fisher AK, Albright-Frey TJ, Reese JC (2001) Analysis of TAF90 mutants displaying allele-specific and broad defects in transcription. *Mol Cell Biol* 21:7331–7344
48. Mohan IW Jr, Scheer E, Wendling O, Metzger D, Tora L (2003) TAF10 (TAF(II)30) is necessary for TFIID stability and early embryogenesis in mice. *Mol Cell Biol* 23:4307–4318
49. Metzger D, Scheer E, Soldatov A, Tora L (1999) Mammalian TAF(II)30 is required for cell cycle progression and specific cellular differentiation programmes. *EMBO J* 18:4823–4834
50. Tatarakis A, Margaritis T, Martinez-Jimenez CP, Kouskouti A, Mohan WS 2nd, Haroniti A, Kafetzopoulos D, Tora L, Talianidis I (2008) Dominant and redundant functions of TFIID involved in the regulation of hepatic genes. *Mol Cell* 31:531–543
51. Mengus G, Fadloun A, Kobi D, Thibault C, Perletti L, Michel I, Davidson I (2005) TAF4 inactivation in embryonic fibroblasts activates TGFbeta signalling and autocrine growth. *EMBO J* 24:2753–2767
52. Chalkley GE, Verrijzer CP (1999) DNA binding site selection by RNA polymerase II TAFs: a TAF(II)250-TAF(II)150 complex recognizes the initiator. *EMBO J* 18:4835–4845
53. Gegonne A, Weissman JD, Lu H, Zhou M, Dasgupta A, Ribble R, Brady JN, Singer DS (2008) TFIID component TAF7 functionally interacts with both TFIIF and P-TEFb. *Proc Natl Acad Sci USA* 105:5367–5372
54. Gegonne A, Weissman JD, Zhou M, Brady JN, Singer DS (2006) TAF7: a possible transcription initiation check-point regulator. *Proc Natl Acad Sci USA* 103:602–607
55. Hiller M, Lin T-Y, Wood C, Fuller MT (2001) Developmental regulation of transcription by a tissue-specific TAF homolog. *Genes Dev* 15:1021–1030
56. Hiller M, Chen X, Pringle MJ, Suchorolski M, Sancak Y, Viswanathan S, Bolival B, Lin TY, Marino S, Fuller MT (2004) Testis-specific TAF homologs collaborate to control a tissue-specific transcription program. *Development* 131:5297–5308
57. Chen X, Hiller M, Sancak Y, Fuller MT (2005) Tissue-specific TAFs counteract Polycomb to turn on terminal differentiation. *Science* 310:869–872
58. Wang PJ, McCarrey JR, Yang F, Page DC (2001) An abundance of X-linked genes expressed in spermatogonia. *Nat Genet* 27:422–426
59. Wang PJ, Page DC (2002) Functional substitution for TAF(II)250 by a retroposed homolog that is expressed in human spermatogenesis. *Hum Mol Genet* 11:2341–2346
60. Burley SK (1996) The TATA box binding protein. *Curr Opin Struct Biol* 6:69–75
61. Shen WC, Green MR (1997) Yeast TAF(II)145 functions as a core promoter selectivity factor, not a general coactivator. *Cell* 90(4):615–624
62. Oelgeschlaeger T, Chiang CM, Roeder RG (1996) Topology and reorganization of a human TFIID-promoter complex. *Nature* 382:735–738
63. Smale ST, Kadonaga JT (2003) The RNA polymerase II core promoter. *Annu Rev Biochem* 72:449–479
64. Kadonaga JT (2002) The DPE, a core promoter element for transcription by RNA polymerase II. *Exp Mol Med* 34:259–264
65. Juven-Gershon T, Hsu JY, Theisen JW, Kadonaga JT (2008) The RNA polymerase II core promoter - the gateway to transcription. *Curr Opin Cell Biol* 20:253–259
66. Carninci P, Sandelin A, Lenhard B, Katayama S, Shimokawa K, Ponjavic J, Semple CA, Taylor MS, Engstrom PG, Frith MC, Forrest AR, Alkema WB, Tan SL, Plessy C, Kodzius R, Ravasi T, Kasukawa T, Fukuda S, Kanamori-Katayama M, Kitazume Y, Kawaji H, Kai C, Nakamura M, Konno H, Nakano K, Mottagui-Tabar S, Arner P, Chesi A, Gustincich S, Persichetti F, Suzuki H, Grimmond SM, Wells CA, Orlando V, Wahlestedt C, Liu ET, Harbers M, Kawai J, Bajic VB, Hume DA, Hayashizaki Y (2006) Genome-wide analysis of mammalian promoter architecture and evolution. *Nat Genet* 38:626–635
67. Sandelin A, Carninci P, Lenhard B, Ponjavic J, Hayashizaki Y, Hume DA (2007) Mammalian RNA polymerase II core promoters: insights from genome-wide studies. *Nat Rev Genet* 8:424–436
68. Gill G, Pascal E, Tseng ZH, Tjian R (1994) A glutamine-rich hydrophobic patch in transcription factor Sp1 contacts the dTAFII110 component of the *Drosophila* TFIID complex and mediates transcriptional activation. *Proc Natl Acad Sci USA* 91:192–196
69. Rojo-Niersbach E, Furukawa T, Tanese N (1999) Genetic dissection of hTAF(II)130 defines a hydrophobic surface required for interaction with glutamine-rich activators. *J Biol Chem* 274:33778–33784
70. Asahara H, Santoso B, Guzman E, Du K, Cole PA, Davidson I, Montminy M (2001) Chromatin-dependent cooperativity between constitutive and inducible activation domains in CREB. *Mol Cell Biol* 21:7892–7900
71. Garbett KA, Tripathi MK, Cencki B, Layer JH, Weil PA (2007) Yeast TFIID serves as a coactivator for Rap1p by direct protein-protein interaction. *Mol Cell Biol* 27:297–311
72. Reeves WM, Hahn S (2005) Targets of the Gal4 transcription activator in functional transcription complexes. *Mol Cell Biol* 25:9092–9102
73. Taverna SD, Li H, Ruthenburg AJ, Allis CD, Patel DJ (2007) How chromatin-binding modules interpret histone modifications: lessons from professional pocket pickers. *Nat Struct Mol Biol* 14:1025–1040
74. Sugauma T, Workman JL (2008) Crosstalk among histone modifications. *Cell* 135:604–607
75. Jacobson RH, Ladurner AG, King DS, Tjian R (2000) Structure and function of a human TAFII250 double bromodomain module. *Science* 288:1422–1425
76. Mizzen CA, Yang XJ, Kokubo T, Brownell JE, Bannister AJ, Owen-Hughes T, Workman J, Wang L, Berger SL, Kouzarides T, Nakatani Y, Allis CD (1996) The TAF(II)250 subunit of TFIID has histone acetyltransferase activity. *Cell* 87:1261–1270
77. Vermeulen M, Mulder KW, Denisov S, Pijnappel WW, van Schaik FM, Varier RA, Baltissen MP, Stunnenberg HG, Mann M, Timmers HT (2007) Selective anchoring of TFIID to nucleosomes by trimethylation of histone H3 lysine 4. *Cell* 131:58–69
78. van Ingen H, van Schaik FM, Wienk H, Ballering J, Rehmann H, Dechesne AC, Kruijzer JA, Liskamp RM, Timmers HT, Boelens R (2008) Structural insight into the recognition of the H3K4me3 mark by the TFIID subunit TAF3. *Structure* 16:1245–1256
79. Vassallo MF, Tanese N (2002) Isoform-specific interaction of HP1 with human TAFII130. *Proc Natl Acad Sci USA* 99:5919–5924
80. Kwon SH, Workman JL (2008) The heterochromatin protein 1 (HP1) family: put away a bias toward HP1. *Mol Cells* 26:217–227
81. Lomberk G, Wallrath L, Urrutia R (2006) The Heterochromatin Protein 1 family. *Genome Biol* 7:228

82. Suganuma T, Pattenden SG, Workman JL (2008) Diverse functions of WD40 repeat proteins in histone recognition. *Genes Dev* 22:1265–1268
83. Callebaut I, Prat K, Meurice E, Mornon JP, Tomavo S (2005) Prediction of the general transcription factors associated with RNA polymerase II in *Plasmodium falciparum*: conserved features and differences relative to other eukaryotes *BMC. Genomics* 6:100
84. Stuwe T, Hothorn M, Lejeune E, Rybin V, Bortfeld M, Scheffzek K, Ladurner AG (2008) The FACT Spt16 “peptidase” domain is a histone H3–H4 binding module. *Proc Natl Acad Sci USA* 105:8884–8889
85. Yudkovsky N, Ranish JA, Hahn S (2000) A transcription reinitiation intermediate that is stabilized by activator. *Nature* 408:225–229
86. Martianov I, Viville S, Davidson I (2002) RNA polymerase II transcription in murine cells lacking the TATA binding protein. *Science* 298:1036–1039
87. Davidson I (2003) The genetics of TBP and TBP-related factors. *Trends Biochem Sci* 28:391–398
88. Christova R, Oelgeschlager T (2002) Association of human TFIID-promoter complexes with silenced mitotic chromatin in vivo. *Nat Cell Biol* 4:79–82
89. Xing H, Vanderford NL, Sarge KD (2008) The TBP-PP2A mitotic complex bookmarks genes by preventing condensin action. *Nat Cell Biol* 10:1318–1323
90. Indra AK, Mohan WS 2nd, Frontini M, Scheer E, Messaddeq N, Metzger D, Tora L (2005) TAF10 is required for the establishment of skin barrier function in foetal, but not in adult mouse epidermis. *Dev Biol* 285:28–37

# *MANUSCRIT N°1*



## Manuscrit n°1

**Emilie CLER, Isabelle MICHEL, Ataillah BENHADDOU, Dominique Kobi, Anne Gonzales, Jyoti Choudhary, Gabrielle MENGUS et Irwin DAVIDSON**

Le facteur général de transcription TFIID est un complexe macromoléculaire composé de la protéine TBP (TATA-binding protein) et d'un ensemble de 13 facteurs associés à TBP (TAFs). 9 des 13 TAFs contiennent un domaine de repliement des histones (HFD) et 5 forment des hétérodimères distincts à l'intérieur de TFIID.

Des approches biochimiques, structurales et génétiques suggèrent que l'architecture 3-D de TFIID soit faite de 2 sous-complexes, un complexe de base contenant TAF5 et les hétérodimères de TAFs contenant le HFD et un second module contenant TAF1, TAF7, TAF2 et TBP qui serait recruté sur le complexe de base afin de former le TFIID complet.

Dans cette étude, nous montrons que les fibroblastes embryonnaires murins contiennent une deuxième catégorie de complexe naturel de base contenant TAF4 ou TAF4b associé à TAF5, TAF6, TAF9 and TAF12. Ce complexe est bien plus abondant que le TFIID dans ces cellules qui expriment un niveau très faible de TBP. Au contraire, les mélanomes 501Mel expriment des niveaux bien plus élevés de TBP et dans ces cellules, la majorité de TAF4 est associée avec TFIID. Ces résultats démontrent l'existence d'un sous-complexe naturel de TFIID et indique que son abondance varie entre les types cellulaires de manière inverse à l'expression de TBP.



**Characterisation of naturally occurring TAF4 and TAF4b-containing TFIID subcomplexes in mouse embryonic fibroblasts.**

Emilie Cler, Isabelle Michel, Ataillah Benhaddou, Dominique Kobi, Anne Gonzales<sup>1</sup>, Jyoti Choudhary<sup>2</sup>, Gabrielle Mengus, and Irwin Davidson#.

Institut de Génétique et de Biologie Moléculaire et Cellulaire.

CNRS/INSERM/UDS.

1 Rue Laurent Fries,

67404 Illkirch Cédex.

France.

FAX: 33 3 88 65 32 01. TEL: 33 3 88 65 34 40 (45)

1. Proteomics and Mass Spectrometry of Biomolecules, 205 route de Narbonne

31077 Toulouse, Cedex04

2. Proteomic Mass Spectrometry, Wellcome Trust Genome Campus, Hinxton, Cambridge,

CB10 1SA, UK

# To whom correspondence should be addressed

E mail : [irwin@titus.u-strasbg.fr](mailto:irwin@titus.u-strasbg.fr)

Running Title : TFIID-subcomplex

Key words : TATA binding protein, melanoma cells, HeLa cells, gel filtration.

**Abstract.**

General transcription factor TFIID is a macromolecular complex composed of the TATA-binding protein (TBP) and a set of 13 evolutionary conserved TBP associated factors (TAFs). Nine of the 13 TAFs contain histone fold domains and form 5 distinct heterodimers within TFIID. Biochemical, structural and genetic approaches suggest that the 3-D architecture of TFIID is made up by two subcomplexes, a core complex containing TAF5 and the histone like TAF heterodimers and a second submodule containing TAF1, TAF7, TAF2 and TBP that is recruited to the core complex to form full TFIID. Here we show that mouse embryonic fibroblasts contain two species of a naturally occurring core-like complex containing either TAF4 or TAF4b along with TAF5, TAF6, TAF9 and TAF12. This complex is much more abundant than TFIID in these cells which also express low levels of TBP. In contrast, 501Mel cells express much higher levels of TBP and almost all of the TAF4 is associated with TFIID. These results demonstrate the existence of naturally occurring TFIID subcomplexes and indicate that their abundance varies between cell-types inversely with the expression of TBP.



## **Introduction.**

Regulated initiation of transcription by RNA polymerase II requires the formation of a macromolecular preinitiation complex (PIC) that assembles over the transcription start site. In addition to RNA polymerase II (pol II), the PIC comprises a set of general transcription factors TFIIA, B, D, E, F, and H. In yeast, the mediator (Med) and NC2 complexes are also found at the promoters of most active genes (Andrau et al., 2006; van Werven et al., 2008). TFIID is itself a macromolecular complex composed of the TATA-binding protein (TBP) and a set of 13-14 evolutionary conserved TBP associated factors (TAFs).

The molecular organisation of TFIID has been addressed in biochemical, structural and genetic approaches. These studies have shown that 9 of the 13 TAFs contain a histone fold domain (HFD) specifying the formation of 5 distinct TAF heterodimers within TFIID (Gangloff et al., 2001a; Gangloff et al., 2001b). Immuno-electron microscopy (EM) further demonstrated that these heterodimers are also found in the context of native TFIID thus underlining that the HFD is a fundamental building block of TFIID (Leurent et al., 2002). EM also revealed that yeast (y)TFIID and human (h)TFIID have a similar overall organisation comprising three major lobes linked by connecting regions to form of a 'horseshoe' shaped molecular clamp (Andel et al., 1999; Brand et al., 1999; Leurent et al., 2002). From these observations a model has been proposed suggesting that the 3-D architecture of TFIID is composed of two subcomplexes: (i) a core complex containing TAF5 and most of the HFD-containing TAF heterodimers (TAF6/9, TAF4/12, TAF8/10, TAF11/13) that adopts a crescent-shaped two-fold symmetric structure, (ii) a subcomplex containing TAF1, TAF7, TAF2 and TBP that is recruited to the core complex to form full TFIID. Further evidence for this comes from genetic studies in yeast, siRNA TAF knockdown in *Drosophila* S2 cells, and reconstitution of sub-complexes from recombinant subunits. In each approach, the existence

of a 7-TAF subcomplex corresponding to the core or a 5-TAF sub-core complex independently of the TAF1-submodule.

While the above results show the existence of TAF sub-complexes in cells where the expression of one or several TAFs has been altered, the presence of such naturally occurring subcomplexes in normal cells has not been described. In this study we have characterised a naturally occurring 5-TAF subcomplex in mouse embryonic fibroblasts (MEFs). We identify two sub-complexes that contain either TAF4 or its paralogue TAF4b along with their heterodimerisation partner TAF12, and TAF5, TAF6 and TAF9. In MEFs, these sub-complexes are much more abundant than full TFIID, are much less abundant than TFIID in HeLa or 501Mel melanoma cells, where almost all of the TAF4 is present in TFIID. This correlates with the observation that HeLa or 501Mel cells contain much higher levels of TBP than MEFs. These observations indicate that in cells with low levels of TBP, a 5-TAF intermediate subcomplex accumulates suggesting that TBP expression levels determine the ratio between full TFIID and the core-type complex.

## **Materials and Methods.**

### **Generation of cell lines.**

The C1*Taf4*<sup>lox/-</sup> and C3*Taf4*<sup>-/-</sup> TAF4-null MEFs are as previously described and were cultured in Dulbecco's minimal essential media (DMEM) supplemented with glutamax and 10% foetal calf serum. HeLa cells were cultured in DMEM with 5% foetal calf serum, and 501Mel cells were cultured in RPMI, supplemented with 10% foetal calf serum in 7% CO<sub>2</sub>.

The recombinant pBABE based retroviruses were constructed by classical cloning methodology using existing plasmids as sources of the TAF4 and TAF4b cDNAs and the final vectors were verified by DNA sequencing. Virus stocks were prepared in Phoenix helper cells by standard procedures. The MEF lines were then infected and cell populations stably expressing the tagged TAF4 and TAF4b were selected using blastocidine or puromycin (3µg/mL).

### **Preparation of cell extracts: Immunoprecipitation and Western blot.**

Nuclear extracts from the different cell lines were prepared essentially as previously described (Brou et al., 1993) and dialysed against buffer A (50 mM Tris-HCl, pH 7.9, 20% Glycerol, 1 mM EDTA, 5 mM MgCl<sub>2</sub>, 1 mM DTT, 0.5 mM PMSF and protinase inhibitor cocktail (Complete, EDTA-free, Roche Diagnostics) containing 0.1 M KCl. SDS-PAGE and Western blots were performed by standard procedure and proteins detection chemiluminescence and autoradiography. The antibodies used are monoclonals 32TA (TAF4, (Mengus et al., 1997)), 3G3 (TBP, (Lescure et al., 1994)), 1TA1C2 (TAF5), 25TA (TAF6), TAF4b was detected using the polyclonal antibody 2057 (Mengus et al., 2005), and TAF10 was detected with monoclonal antibody MAB3648 (Millipore). 2439 polyclonal TAF1, Anti-Flag antibody was from monoclonal antibody F1804 (Sigma).

Immunoprecipitations were performed with anti-Flag M2 resin (Sigma) or 5TF12C1 (TBP) and protein G sepharose for TBP. Cell extracts were first pre-cleared for 2 hours at 4°C with 50-100 ul of protein G-sepharose before incubation for 1 hour at 4°C with rotation with 50-100 ul of pre-washed M2 resin. Beads were collected by centrifugation and washed 6 times with 1 ml of buffer A containing 0.3 M KCl and finally 1 time with 1 ml of buffer A containing 0.1 M KCl. Proteins were eluted 2 times from the M2 resin using 0,2 mg/ml (final concentration) of a tandem Flag peptide for 30 min at 4°C and the last time at room temperature for 30min. The eluted fractions were either analysed directly by Western blot or reprecipitated with 5TF12C1 antibody. The 3G3 precipitates were washed as described above and the proteins eluted with the peptide PB242 corresponding to the 5TF12C1 epitope as previously described (Brou et al., 1993).

### **Mass spectrometry**

Protein bands were excised from the gel, de-stained completely and digested with trypsin (sequencing grade, Roche). Peptides were extracted using 0.5% formic acid-50% acetonitrile and dried in a Speed Vac (Thermo) and re-dissolved in 0.5% formic acid (FA) and analysed with on-line nanoLC-MS/MS on an LTQ FT (Thermo Fisher Scientific), a hybrid linear ion trap and a 7 Tesla Fourier transform ion cyclotron resonance mass spectrometer, coupled with an Ultimate 3000 Nano/Capillary LC System (Dionex). Samples were first loaded and desalted on a PepMap C18 trap column (0.3 mm id x 5 mm, LC Packings) at 25 µL/min with 0.1% formic acid for 5 min then separated on an analytical column, BEH C18 (75 µm id x 10 cm, Waters) at a flow rate of 300 nL/min over a 30, 45 or 60 min linear gradient of 4-32% CH<sub>3</sub>CN/0.1% FA based on the gel band stain intensity.

The LTQ FT mass spectrometer was operated in the standard data dependent acquisition mode controlled by Xcalibur 1.4. The survey scans (m/z 400-1500) were acquired on the FT-ICR at a resolution of 100,000 at m/z 400 and one microscan per spectrum. The

three most abundant doubly and triply charged ions (2+ and 3+) with a minimal intensity at 1000 counts were subject to MS/MS in the linear ion trap. The dynamic exclusion width was set at  $\pm 10$  ppm with one repeat within 45 sec and excluding for 60 sec.

The raw files were processed with BioWorks 3.2 to generate the dta files. The dta files from the raw files of all bands from each sample were then merged to a single mgf file and submitted to Mascot search with an in-house non-identical, non-redundant mouse IPI database. The search parameters were: trypsin with one mis-cleavage, 20 ppm peptide mass tolerance, 0.5 Da fragment mass tolerance, with variable modifications of carbamidomethyl (C), N-acetylation (protein), oxidation (M) and pyro-glu (N-term Q), deamidation (NQ), formyl (N-terminal) and methyl (E).

#### **Gel filtration and glycerol gradient centrifugation.**

For gel filtration 1mL of nuclear extract was injected into a S6 gel filtration column (10/600mm) at a flow rate of 0,3 ml/min (Amersham Biosciences), which was previously equilibrated in buffer A containing 0.1 M KCl. The column was eluted with 35 X1ml fractions and alternate fractions were analysed by Western blot.

Gel filtration was also performed with a SMART column PC 3,2/30 pre-packed with Superose<sup>TM</sup>6 designed for sensitive and highly resolving gel filtration separations of proteins from Amersham Biosciences.

For glycerol gradient centrifugation 500 ul of nuclear extracts were loaded onto a 5ml 19-37% glycerol (w/v) gradient. The gradients were formed from 9 layers containing 20mM TrisHCl pH7,9; 150mM NaCl, 1mM MgCl<sub>2</sub>, 0,1mM EDTA and 19 to 35% glycerol. The gradients were centrifuged in a Beckman SW60Ti rotor at 49.000 rpm for 16h, with slow acceleration and deceleration. Fractions of 500 $\mu$ l were collected via a hole punched in the bottom of the tubes. Molecular mass standards (NativeMark, Invitrogen) were loaded in one of the tubes in the rotor and analysed by 3-12% SDS-PAGE and silver nitrate staining.

## Results.

### Characterisation of a TAF4-containing complex that lacks TBP.

We have previously described MEF cell lines with a null and a floxed allele of the gene encoding TAF4 (*C1Taf4<sup>lox/-</sup>*) and TAF4-null MEFs (*C3Taf4<sup>-/-</sup>*, hereafter C1 and C3) that were derived from them after expression of the Cre recombinase (Mengus et al., 2005). Immunoprecipitation experiments showed that upon loss of TAF4, TAF4b is found associated with TBP and maintains TFIID integrity. However, most TAF4b was not associated with TBP and must therefore be present in another complex.

To better characterise the TAF4 and TAF4b complexes in the MEF lines, each cell line was infected with a retrovirus expressing either TAF4 or TAF4b with a 3XFlag-SBP (streptavidine binding protein) Tag at their C-terminus (designated TAF4-3F and TAF4b-3F, Fig. 1A) and cell populations stably expressing the tagged proteins were derived by puromycin selection. Western blot analysis with antibody against the Flag-Tag or TAF4 shows that following selection, expression of exogenous TAF4-3F can be seen in both the C1 and C3 cells (designated C1-TAF4-3F and C3-TAF4-3F in Fig. 1B). The level of exogenous TAF4-3F in the C3 cells is comparable to that of the endogenous TAF4 protein seen in the C1 cells (compare lanes 1 and 3), while it is higher in the C1-TAF4-3F cells expressing both the endogenous and exogenous proteins (lane 4). The expression of TAF4b is also somewhat reduced in these cells expressing higher amounts of TAF4. We also observed expression of TAF4b-3F in the C3 cells, however, despite several attempts with different preparations of retrovirus, we were unable to generate C1 cells expressing exogenous TAF4b-3F (Fig. 1C).

We have previously shown that TAF4 inactivation leads to a pronounced morphological change as the cells lose their rounded and regular morphology to adopt an

irregular shape (Mengus et al., 2005) and see Fig. 1D). This can be rescued by expression of TAF4-3F showing that the presence of the C-terminal Tag does not affect TAF4 functionality.

To identify TAF4 associated proteins, we immunoprecipitated TAF4 from nuclear extracts of C3-TAF4-3F cells with anti-Flag antibody. TAF4 was precipitated using the anti-Flag antibody along with TBP and TAF6 (Fig. 1E, lanes 1 and 2). In contrast, no TAF4b was coprecipitated with TAF4 showing that these two related proteins are not present in the same TFIID complexes and are not common subunits of any other complexes (see below). We also made tandem purifications, by reprecipitating the eluates of the Flag precipitations with streptavidine agarose. However, the second step precipitation was extremely inefficient as little TAF4-3F was found associated with the streptavidine beads (data not shown). We therefore abandoned the tandem purification approach and optimised the Flag precipitations by using only short incubation times with the M2 resin and multiplying the number of washes.

The Flag immunoprecipitates were also subjected to SDS-PAGE and stained with silver nitrate. Compared to the Flag immunoprecipitation performed on control C3 cells infected with the empty retrovirus vector, several species with molecular masses corresponding to TAF4, TAF5 and TAF6 were clearly visible (Fig. 2A). Nevertheless, no species with the mass of TAF1, TAF2 or TBP, that are clearly visible in TFIID preparations (see (Mengus et al., 1995) ) were visible in the Flag-TAF4 immunoprecipitations.

The identity of the visible proteins was verified by mass-spectrometry analysis. The same fractions were analysed by SDS-PAGE and colloidal blue staining (Fig. 2B) and then each entire lane was divided into 24 slices and analysed by mass-spectrometry. This analysis confirmed the visible stained proteins as TAF4, TAF5 and TAF6 and identified the presence of TAF12 as well as TAF9 and the related TAF9b, the heterodimerisation partners of TAF4 and TAF6, respectively (see Supplemental Table 1). These proteins of much smaller mass were not readily visible by silver nitrate staining. This analysis also revealed the presence of

TBP and all of the other TFIID TAFs with the exception of TAF2 and TAF4b, specifically in the anti-Flag precipitation of the C3-TAF4-3F, but not the control extracts. These results confirm those of the western blot showing that TBP, but not TAF4b was coprecipitated with TAF4. We also reproducibly detected low levels of several RNA polymerase II subunits in the anti-TAF4-3F immunoprecipitations. In contrast, no TFTC/STAGA components were detected in the Flag immunoprecipitations showing that TAF4 and TAF4b are not subunits of these complexes. Thus while precipitation of TAF4 results in the purification of TFIID, the silver nitrate staining shows that TFIID is not the major constituent of the precipitate which rather seems to be a complex of TAF4, TAF5, TAF6 and their heterodimerisation partners TAF9(b) and TAF12.

To investigate the possibility that the majority of TAF4, is not associated with TFIID, the fraction obtained by anti-Flag immunoprecipitation was reprecipitated using anti-TBP antibody. TBP was depleted from the supernatant fraction and was present in the immunoprecipitate (Fig. 1E, lanes 3 and 4). In contrast only a fraction of TAF4 and TAF6 were coprecipitated along with TBP, while the majority of these proteins remained in the supernatant fraction. Sequential Flag and TBP immunoprecipitation therefore allows the purification of TBP-associated and TBP-independent forms of TAF4.

To further characterise the TAF4 containing complexes in the C3-TAF4-3F cells, the extracts were loaded on a gel filtration column. Western blot analysis of the eluted fractions showed that TAF4 eluted in two peaks, a high molecular mass fraction that elutes around fraction 5 along with TBP and a second more abundant fraction eluting around fractions 13-15 (Fig. 3A). TAF6 and TAF5 also showed elution profiles similar to that of TAF4, while in addition to the high molecular mass complex, TBP was also present in lower molecular mass complexes eluting in fractions 17-19, that may correspond to the SL1, TFIIC and B-TFIID



complexes. TAF10 elutes almost exclusively in the higher molecular mass range consistent with the fact that it is a component of the TFIID and the STAGA/TFTC complexes.

These results are consistent with the idea that TAF4, TAF5 and TAF6 are present in a high molecular mass TFIID complex that elutes in fractions 4-6 and a more abundant TBP-free complex in fractions 14-16. To verify this, the fractions corresponding to each peak were pooled separately and sequentially precipitated with the anti-Flag antibody and then the TBP antibody. Flag-precipitation of the high molecular mass fraction precipitated TAF4 and TBP, and TAF4 could also be clearly detected in the subsequent TBP precipitation (Fig. 3C). In contrast, in Flag-precipitation of the lower mass complex lesser amounts of TBP could be detected, and almost no TAF4 could be seen in the TBP immunoprecipitate (Fig. 3D). Together these results reveal the presence of at least two distinct TAF4-containing complexes in these cells. This idea was further supported by a second FPLC gel filtration experiment using on a SMART system. Again, TAF4 and TAF6 elute in two distinct peaks and TBP was detected only in the higher molecular mass complex (Fig. 3E). The majority of TAF1 also elutes in the higher molecular mass fractions with little being present in the fractions containing the TAF4-containing subcomplex. Comparison with the elution of the molecular mass standards show that the TAF4-containing subcomplex elutes with an apparent mass of around 600 kDa similar to that of thyroglobulin.

As an alternative approach, we separated the complexes in the C1 MEF extracts using a 15-35% glycerol gradient. Here we used C1 MEF cells containing only the endogenous TAF4. Using this technique we also separated TAF4 into two distinct peaks, a higher molecular mass complex containing TBP and a lower mass complex that contains the majority of TAF4 and TAF6 (Fig. 4A). TAF1 also elutes in the higher molecular mass range in a pattern distinct from that of TAF4. Comparison with molecular mass markers that were centrifuged on the same rotor (Fig. 4C), showed that the abundant TAF4-TAF6 subcomplex

had an apparent mass of around 500kDa, while the second complex had a mass of 1mDa or higher consistent with that described for TFIID. Together the results of these analyses show that MEFs contain an abundant TAF4-containing TFIID subcomplex in addition to the less abundant TFIID complex.

#### **Characterisation of a TAF4b-containing complex that lacks TBP.**

We next analysed the status of TAF4b in cell extracts and the fractions from the gel filtration and glycerol gradients. Anti-Flag immunoprecipitation of extracts from the C3-TAF4b-3F cells showed coprecipitation of TAF6, but only upon strong exposure could TBP be seen (Fig 1E, lanes 5-6). This is consistent with our previous results showing that only small amounts of endogenous TAF4b were coprecipitated with TBP from the C3 cells (Mengus et al., 2005). Similarly, in the gel filtration of the C3-TAF4-3F extracts, the majority of the endogenous TAF4b eluted similar to TAF4 in the lower molecular mass complex, while TAF4b was evident in the fractions containing the TFIID complex only upon strong exposure (Fig. 3B). In agreement with this, in the glycerol gradient the majority of TAF4b eluted in the lower molecular mass range similar to TAF4 (Fig. 4A).

These results suggest that like TAF4, the majority of TAF4b is present in a subcomplex along with TAF5, TAF6, TAF9 and TAF12. To verify this, the anti-Flag precipitates from the C3-TAF4b-3F cells were analysed by SDS-PAGE, silver nitrate staining and mass-spectrometry. The anti-Flag immunoprecipitates of TAF4b were much less pure than those of TAF4 and contained more contaminants (compare Figs 2A and C). Nevertheless, TAF4b, TAF5, TAF6 and the 20 kDa and 15 kDa forms of TAF12 were all visible by silver nitrate staining and their identity was confirmed by mass-spectrometry. Mass spectrometry further identified TAF4b and all the other TAFs, with the exception of TAF4, specifically in the Flag precipitation from the cell lines expressing TAF4b-3F (Supplemental

Table 1). Thus as observed with TAF4, precipitation of tagged TAF4b leads to purification of full TFIID and an additional more abundant lower molecular mass subcomplex.

### **Cell type-specific ratios of TFIID and TAF subcomplexes.**

The above results show that in MEFs the majority of TAF4 is not associated with TFIID, but is present in a 5-TAF subcomplex. Previously however, it had been shown that depletion of TBP from HeLa cell extracts resulted also in TAF4 depletion showing that in these cells most if not all TAF4 is associated with TBP (Oelgeschlager et al., 1998). Furthermore, in our experiments, we noticed that detection of TBP in the different C1 and C3 MEF-derived fractions in general required lengthy exposures. This suggested that C1 and C3 MEFs may have low levels of TBP and that the TAF4/TBP ratio may differ from one cell type to another and thus influence formation of full TFIID versus TAF subcomplexes.

To verify the above idea, we performed western blots of nuclear extracts from the MEFs and other different cell types. Comparison of the TBP levels in C1 and C3 MEFs compared to HeLa and 501Mel (a human melanoma cell line, (Goodall et al., 2008)) showed that HeLa and 501Mel cells contained much more TBP than the MEF lines (Fig. 5A). Moreover, TBP levels were higher in the C1 MEFs than in the C3 MEFs. This was more clearly seen in titration experiments where the levels of TAF4 in the C1 MEFs are comparable or even higher than in HeLa and 501Mel cell extracts (Fig. 5B). In contrast, the levels of TBP are substantially lower in C1 MEFs. We also compared TBP and TAF4 levels in another independently isolated MEF cell line (P1-SV40). While the levels of TBP in this line were higher than in the C1 MEFs, they still remained much lower than in HeLa and 501Mel cell extracts. The differences in the levels of TBP may reflect the fact that the P1-SV40 cells were immortalised by expression of the SV40 T-antigen, while the C1 and C3

cells were derived from a population of MEFs that were isolated after spontaneous immortalisation.

The cell line that appears to exhibit the highest level of TBP is 501Mel. We subjected the extracts from this line to glycerol gradient centrifugation. In contrast to the MEFs, in 501Mel extracts almost all of the TAF4 and TAF6 sedimented to the higher molecular mass TFIID complex along with TBP and TAF1 (Fig. 4B). In contrast, only little TAF4b co-sedimented with TBP, the majority remained in a lower mass complex that sedimented comparably to that seen in MEFs. These results show that in 501Mel cells, most if not all TAF4 is associated with TBP, whereas TAF4b remains in a lower mass complex. Thus, the TBP/TAF4 ratio in different cell types modulates the formation of full TFIID compared to TAF subcomplexes.

## **Discussion.**

In this study we characterise a naturally occurring TFIID subcomplex by immunopurification of TAF4 from MEF cell lines. Immunopurification, gel filtration, and glycerol gradient centrifugation revealed that the majority of TAF4 in these cells was not associated with TBP, but was present in a lower molecular mass subcomplex comprising TAF5, TAF6, TAF9 and TAF12. Our gel filtration experiments show that TAF10 elutes in the high molecular weight fraction where TFIID and STAGA are found, whereas only smaller amounts are found in the lower mass fraction with TAF4 indicating that the TAF4-containing subcomplex probably does not in its majority comprise the TAF8/TAF10 heterodimer. We were unable to assess the presence of TAF8, TAF11 and TAF13 in the different fractions as the available antibodies did not recognise the murine proteins efficiently. These results suggest that the TAF4 and TAF4b are present in 5-TAF TFIID subcomplexes. This subcomplex appears to have a molecular mass between 500-600 kDa. A complex containing

one copy of each TAF4, 5, 6, 9 and 12 would have a mass of around 365 kDa, while two copies would have an apparent mass of 730 kDa. We suggest that the complex described here contains two copies of each TAF by analogy with the 7-TAF complex which in addition contains TAF8 and TAF10, that has been reconstructed *in vitro* using recombinant baculoviruses (Leurent et al., 2004). The 7-TAF subcomplex was shown by electron microscopy to correspond to a symmetrical ‘core’ TFIID complex, comprising two copies of TAF5 and of each of the HFD heterodimers (Cler et al., 2009; Grob et al., 2006; Papai et al., 2009). The MEF-cell subcomplex is analogous to a stable 5-TAF subcomplex that has been described in *Drosophila* S2 cells where TAF1 has been depleted by siRNA knock down (Wright et al., 2006). The important observation however is that the TFIID subcomplex described here is found to exist naturally in MEF cells and that it is more abundant than the full TFIID complex.

We also show that in MEFs the majority of the cellular pool of TAF4b is present in a TFIID subcomplex analogous to that of TAF4. Moreover, immunopurification of TAF4 does not coprecipitate TAF4b showing that MEFs contain two distinct populations of subcomplexes containing either TAF4 or TAF4b. This finding explains our previous observation that little TAF4b is coprecipitated with TBP even in TAF4-null cells (Mengus et al., 2005).

Our results further show that the abundance of the TAF4-containing subcomplex varies between cell lines. In MEFs expressing low amounts of TBP this subcomplex is more abundant than the full TFIID, while in 501Mel cells that express high levels of TBP, little or no TAF4-containing subcomplex is observed. In contrast, the TAF4b in 501Mel cells is not associated with TFIID, but is present in a lower molecular mass complex with properties analogous to that found in MEFs. There are thus several distinct scenarios : i) in C1 MEFs the TAF4 and TAF4b subcomplexes are more abundant than TFIID which comprises almost

uniquely TAF4; ii) in C3 MEFs where TAF4 has been inactivated, the TFIID contains only TAF4b, but the majority of TAF4b remains in the form of a subcomplex; iii) in 501 Mel cells, TFIID is much more abundant and comprises almost uniquely TAF4, while TAF4b remains in a subcomplex.

An additional striking observation is the low TBP and hence TFIID levels found in MEFs. A comparison between the C1 and C3 MEFs show that TBP levels are lower in the C3 cells than in the C1 cells despite the fact that the C3 cells are capable of serum-independent autocrine growth and are oncogenically transformed forming tumours in nude mice ((Mengus et al., 2005) and our unpublished data). The TBP levels in the C3 cells are notably substantially lower than in 501Mel cells, that although derived from a melanoma tumour, are very inefficient in forming tumours in nude mice (Goodall et al., 2008). Thus in this comparison, the cell line with the higher proliferation rate *in vitro* and more tumorigenic *in vivo* is that which expresses the lower amounts of TBP. These observations are somewhat in contradiction to previous findings that increased TBP levels promote proliferation and oncogenic transformation (Johnson et al., 2003).

The finding that the C1 and the C3 cells contain low amounts of TBP/TFIID raise the question concerning the potential function of the TAF4 and TAF4b-containing subcomplexes. It is possible that they are simply assembly intermediates of TFIID and when TBP (and/or TAF1) levels are limiting, they cannot go on to assemble a full TFIID complex, but have no function *per se*. Alternatively, they may have a function distinct from that of TFIID. We have tried to assess the functional properties of the purified TAF4-containing subcomplexes in *in vitro* transcription reactions, but under the conditions used, we did not observe any positive or negative effects (our unpublished data). It will rather be necessary to perform chromatin immunoprecipitation experiments to determine whether we can identify promoters or loci that are occupied by TAF4, but not by TBP reflecting sites that are occupied by the TAF4-

subcomplex. In an initial study using well characterised promoters, we have so far been unable to see such differential occupation (our unpublished data). Further studies will be necessary to determine if and what function these complexes may play within the cell.

### **Acknowledgements**

We thank, J-M. Garnier for construction of the recombinant retrovirus vectors, L. Tora for antibodies, F. Klein and J Bonnet for help with gel filtration, and A Hamiche for help with glycerol gradients. This work was supported by grants from the CNRS, the INSERM, the Association pour la Recherche contre le Cancer, the Ligue Nationale et Départementale Région Alsace contre le Cancer. ID is an 'équipe labellisée' of the Ligue Nationale contre le Cancer. EC was supported by a fellowship from the Fondation pour la Recherche Médicale.

### **References.**

- Andel, F., 3rd, Ladurner, A. G., Inouye, C., Tjian, R., and Nogales, E. (1999). Three-dimensional structure of the human TFIID-IIA-IIB complex. *Science* 286, 2153-2156.
- Andrau, J. C., van de Pasch, L., Lijnzaad, P., Bijma, T., Koerkamp, M. G., van de Peppel, J., Werner, M., and Holstege, F. C. (2006). Genome-wide location of the coactivator mediator: Binding without activation and transient Cdk8 interaction on DNA. *Mol Cell* 22, 179-192.
- Brand, M., Leurent, C., Mallouh, V., Tora, L., and Schultz, P. (1999). Three-dimensional structures of the TAFII-containing complexes TFIID and TFIIIC. *Science* 286, 2151-2153.
- Brou, C., Chaudhary, S., Davidson, I., Lutz, Y., Wu, J., Egly, J. M., Tora, L., and Chambon, P. (1993). Distinct TFIID complexes mediate the effect of different transcriptional activators. *EMBO J* 12, 489-499.
- Cler, E., Papai, G., Schultz, P., and Davidson, I. (2009). Recent advances in understanding the structure and function of general transcription factor TFIID. *Cell Mol Life Sci* 66, 2123-2134.

Gangloff, Y., Romier, C., Thuault, S., Werten, S., and Davidson, I. (2001a). The histone fold is a key structural motif of transcription factor TFIID. *Trends Biochem Sci* *26*, 250-257.

Gangloff, Y. G., Pointud, J. C., Thuault, S., Carre, L., Romier, C., Muratoglu, S., Brand, M., Tora, L., Couderc, J. L., and Davidson, I. (2001b). The TFIID components human TAF(II)140 and *Drosophila* BIP2 (TAF(II)155) are novel metazoan homologues of yeast TAF(II)47 containing a histone fold and a PHD finger. *Mol Cell Biol* *21*, 5109-5121.

Goodall, J., Carreira, S., Denat, L., Kobi, D., Davidson, I., Nuciforo, P., Sturm, R. A., Larue, L., and Goding, C. R. (2008). Brn-2 represses microphthalmia-associated transcription factor expression and marks a distinct subpopulation of microphthalmia-associated transcription factor-negative melanoma cells. *Cancer Res* *68*, 7788-7794.

Grob, P., Cruse, M. J., Inouye, C., Peris, M., Penczek, P. A., Tjian, R., and Nogales, E. (2006). Cryo-electron microscopy studies of human TFIID: conformational breathing in the integration of gene regulatory cues. *Structure* *14*, 511-520.

Johnson, S. A., Dubeau, L., White, R. J., and Johnson, D. L. (2003). The TATA-binding protein as a regulator of cellular transformation. *Cell Cycle* *2*, 442-444.

Lescure, A., Lutz, Y., Eberhard, D., Jacq, X., Krol, A., Grummt, I., Davidson, I., Chambon, P., and Tora, L. (1994). The N-terminal domain of the human TATA-binding protein plays a role in transcription from TATA-containing RNA polymerase II and III promoters. *EMBO J* *13*, 1166-1175.

Leurent, C., Sanders, S., Ruhlmann, C., Mallouh, V., Weil, P. A., Kirschner, D. B., Tora, L., and Schultz, P. (2002). Mapping histone fold TAFs within yeast TFIID. *Embo J* *21*, 3424-3433.

Leurent, C., Sanders, S. L., Demeny, M. A., Garbett, K. A., Ruhlmann, C., Weil, P. A., Tora, L., and Schultz, P. (2004). Mapping key functional sites within yeast TFIID. *Embo J* *23*, 719-727.



Mengus, G., Fadloun, A., Kobi, D., Thibault, C., Perletti, L., Michel, I., and Davidson, I. (2005). TAF4 inactivation in embryonic fibroblasts activates TGFbeta signalling and autocrine growth. *Embo J* 24, 2753-2767.

Mengus, G., May, M., Carre, L., Chambon, P., and Davidson, I. (1997). Human TAF(II)135 potentiates transcriptional activation by the AF-2s of the retinoic acid, vitamin D3, and thyroid hormone receptors in mammalian cells. *Gene Develop Vol 11, Iss 11*, 1381-1395.

Mengus, G., May, M., Jacq, X., Staub, A., Tora, L., Chambon, P., and Davidson, I. (1995). Cloning and characterization of hTAFII18, hTAFII20 and hTAFII28: three subunits of the human transcription factor TFIID. *EMBO J* 14, 1520-1531.

Oelgeschlager, T., Tao, Y., Kang, Y. K., and Roeder, R. G. (1998). Transcription activation via enhanced preinitiation complex assembly in a human cell-free system lacking TAFIIs. *MolCell* 1, 925-931.

Papai, G., Tripathi, M. K., Ruhlmann, C., Werten, S., Crucifix, C., Weil, P. A., and Schultz, P. (2009). Mapping the initiator binding Taf2 subunit in the structure of hydrated yeast TFIID. *Structure* 17, 363-373.

van Werven, F. J., van Bakel, H., van Teeffelen, H. A., Altelaar, A. F., Koerkamp, M. G., Heck, A. J., Holstege, F. C., and Timmers, H. T. (2008). Cooperative action of NC2 and Mot1p to regulate TATA-binding protein function across the genome. *Genes Dev* 22, 2359-2369.

Wright, K. J., Marr, M. T., 2nd, and Tjian, R. (2006). TAF4 nucleates a core subcomplex of TFIID and mediates activated transcription from a TATA-less promoter. *Proc Natl Acad Sci U S A* 103, 12347-12352.

## Legends to figures.

**Figure 1.** Expression of recombinant Flag-Tagged TAF4 and TAF4b in MEFs. **A.** Schematic representation of the structure of TAF4 and TAF4b. In TAF4, the proline and alanine-rich N-terminal domain and the glutamine-rich domains (Q1-Q4) are indicated. The locations of the conserved regions I and II (CR-I, CR-II) that are shared between TAF4 and TAF4b are also shown. C indicates the conserved C-terminal domain that contains the  $\alpha$ 3-helix of the histone fold. **B-C** Western blot performed on extracts from the C1 and C3 MEFs and the cell populations selected after retrovirus infection stably expressing either TAF4 or TAF4b revealed either by the Flag, TAF4 or TAF4b antibodies as indicated in the left of the figure. **D.** Phase contrast microscopy of cells grown in 10% serum showing the changes in morphology upon loss of TAF4 and restoration by expression of exogenous Flag-Tagged TAF4. 20X magnification. **E.** Sequential Flag-TBP immunoprecipitation of C3-TAF4-3F extracts. Lane 1 shows the starting cell extract, lane 2 the Flag immunoprecipitation (IP), lane 3 the TBP IP and lane 4 the supernatant (SN) of the TBP IP. Lanes 5 and 6 show the extract and Flag IP of the C3-TAF4b-3F cells.

**Figure 2.** Identification of TAF4-associated proteins. **A.** Detection of proteins in the Flag IPs from the C3-TAF4-3F and control cells following SDS-PAGE and silver nitrate staining. The identities of the visible proteins were determined by mass-spectrometry. **B.** The Flag-IPs were subjected to SDS-PAGE and colloidal blue staining prior to excision of the gel slices and mass-spectrometry analysis. **C.** Detection of proteins in the Flag IPs from the C3-TAF4b-3F and control cells following SDS-PAGE and silver nitrate staining.

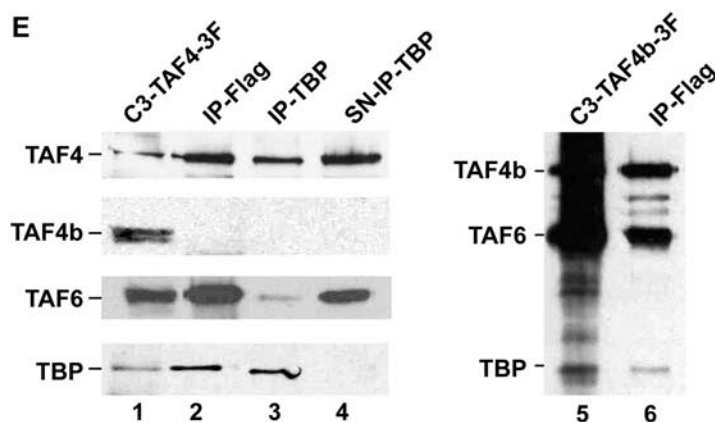
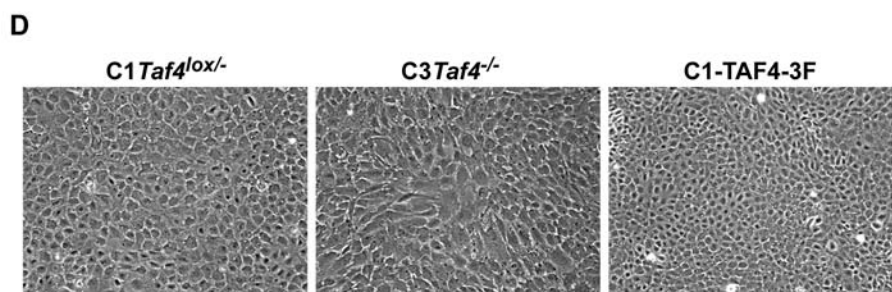
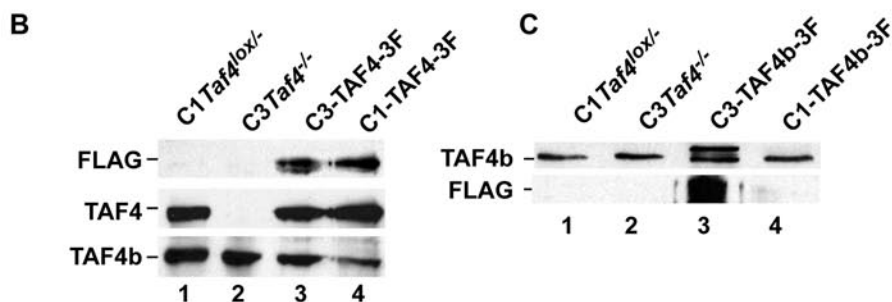
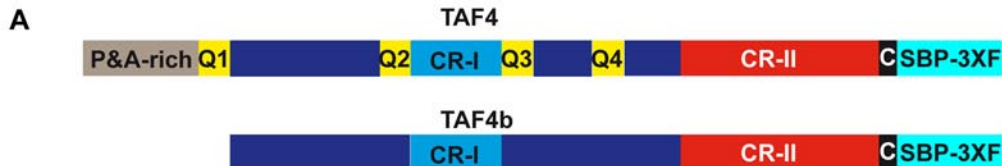
**Figure 3.** Gel filtration analysis of the C3-TAF4-3F cell extracts separates the TAF4-containing subcomplex from TFIID. **A.** Fractions from the gel filtration were analysed using antibodies against the proteins indicated to the left of the figure. For TAF5 a short and long

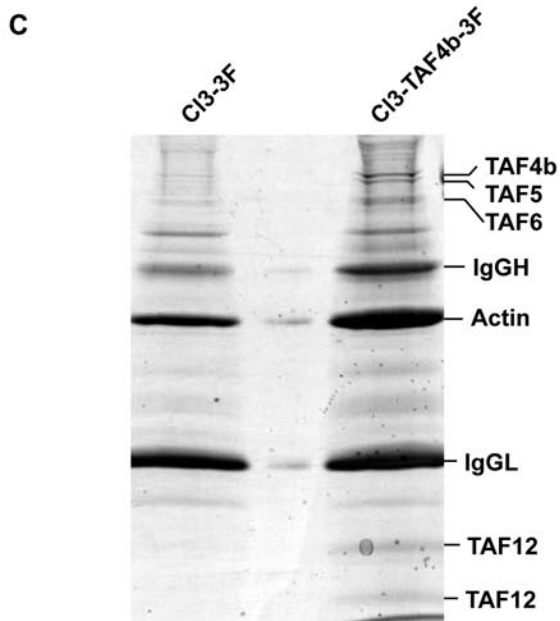
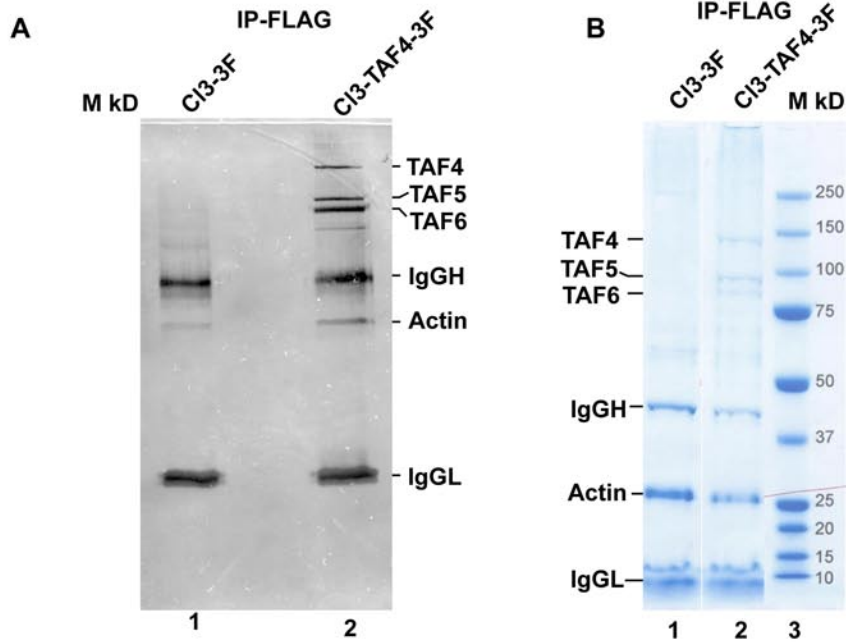
exposure are shown. **B.** The presence of TAF4b in the indicated fractions was analysed. **C-D.** The indicated fractions from the gel filtration were pooled and then subjected to sequential Flag and TBP immunoprecipitation. For TAF4, long and short exposures are shown, while for TBP only the long exposure is shown. **E.** HPLC gel filtration of the C3-TAF4-3F cell extracts using the SMART system. The fraction of elution of the mass marker is shown.

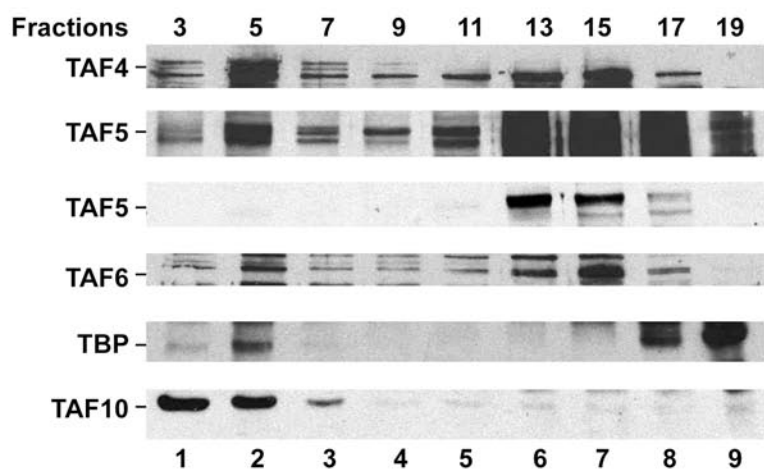
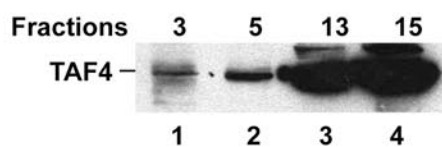
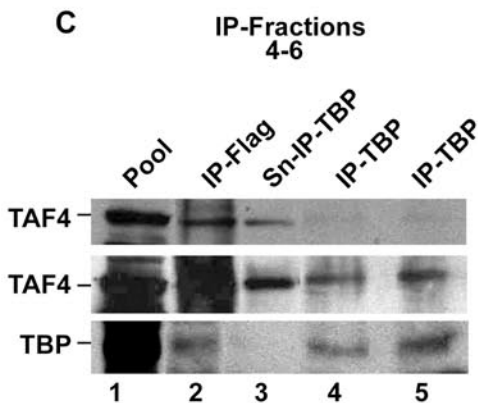
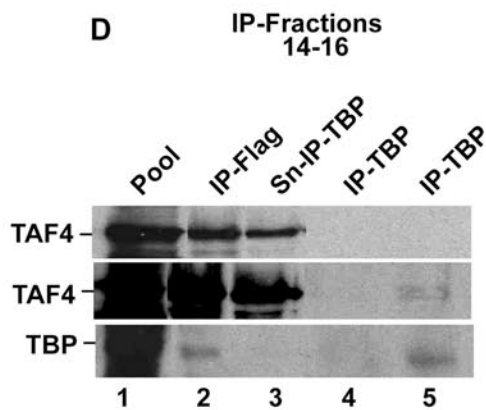
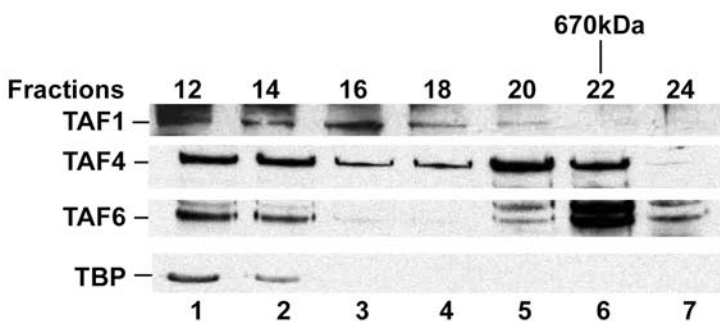
**Figure 4.** Glycerol gradient centrifugation of C1 MEF and 501Mel extracts reveals the differential association of TAF4 with TFIID. **A-B.** The fractions from the glycerol gradient centrifugation of the C1 MEF or 501Mel extracts were analysed for the presence of the proteins indicated to the left of the figure. **C.** The sedimentation of the molecular mass markers in the glycerol gradient was detected by silver nitrate staining following native gel electrophoresis.

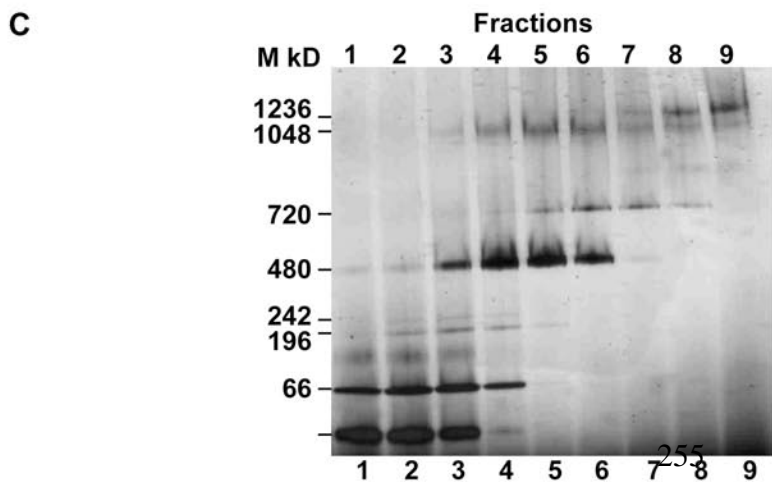
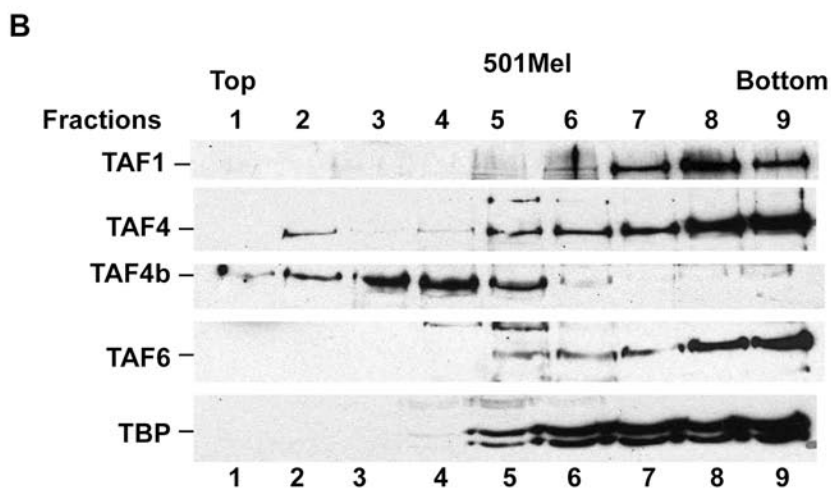
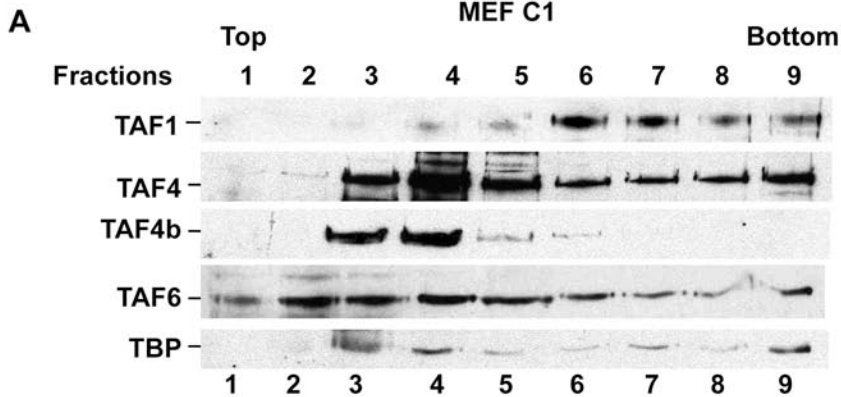
**Figure 5.** TBP/TAF4 ratios in different cell types. **A.** Western blot with the indicated antibodies using 20 ug of each cell extract. **B.** Western blot on increasing amounts (2, 5, 10 and 20 ug) of extracts from the indicated cell type.

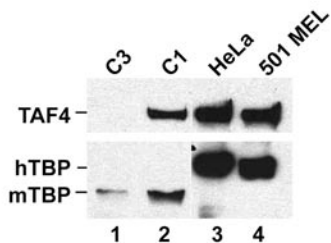
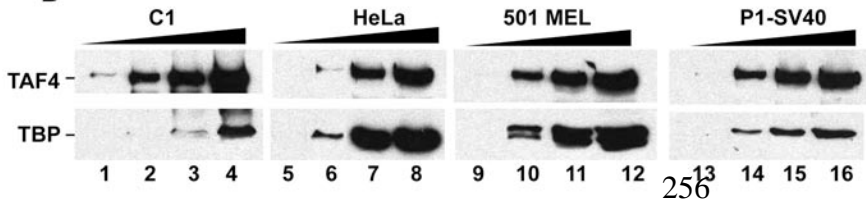
**Supplemental Table 1.** Excel table showing the results of the mass-spectrometry analysis of the TAF4 (page 1) and TAF4b (page 2) Flag IPs. The gene symbol description and number of identified peptides for each protein are indicated. In addition to the TFIID subunits, low amounts of several RNA polymerase II subunits were found in the TAF4 sample. To identify the proteins specifically present in the Flag TAF4(b)-IP. The protein identifications from the TAF4(b) and the respective control purifications were clustered based on 90% homology and proteins only detected in TAF4(b) but not control with FDR usually less than 1% are reported in the tables. Number of peptides over score threshold >1% FDR and used for protein identification.





**A****B****C****D****E**



**A****B**



MGI Gene symbol	Description	Number of peptides
<i>Taf9</i>	Transcription initiation factor TFIID subunit 9 - Mus musculus (Mouse)	24
<i>Taf6</i>	Transcription initiation factor TFIID subunit 6 - Mus musculus (Mouse)	53
<i>Taf5</i>	Transcription initiation factor TFIID subunit 5 - Mus musculus (Mouse)	51
<i>Taf7</i>	Transcription initiation factor TFIID subunit 7 - Mus musculus (Mouse)	26
<i>Taf11</i>	Transcription initiation factor TFIID subunit 11 - Mus musculus (Mouse)	14
<i>Taf9b</i>	Transcription initiation factor TFIID subunit 9b - Mus musculus (Mouse)	15
<i>Taf13</i>	Transcription initiation factor TFIID subunit 13 - Mus musculus (Mouse)	9
<i>Taf10</i>	Transcription initiation factor TFIID subunit 10 - Mus musculus (Mouse)	12
<i>Taf4a</i>	Transcription initiation factor TFIID subunit 4a - Mus musculus (Mouse)	40
<i>Taf1</i>	Transcription initiation factor TFIID subunit 1 - Mus musculus (Mouse)	98
<i>Taf8</i>	Transcription initiation factor TFIID subunit 8 - Mus musculus (Mouse)	11
<i>Taf12</i>	Transcription initiation factor TFIID subunit 12 - Mus musculus (Mouse)	7
<i>Taf3</i>	Transcription initiation factor TFIID subunit 3 - Mus musculus (Mouse)	36
<i>Tbp</i>	Transcription initiation factor TFIID subunit TATA binding protein - Mus musculus (Mouse)	10
<i>Polr2c</i>	polymerase (RNA) II (DNA directed) polypeptide C	2
<i>Polr2e</i>	polymerase (RNA) II (DNA directed) polypeptide E	2
<i>Polr2a</i>	polymerase (RNA) II (DNA directed) polypeptide A	7
<i>Polr2c</i>	polymerase (RNA) II (DNA directed) polypeptide C	2
<i>Polr2b</i>	polymerase (RNA) II (DNA directed) polypeptide B	3

<b>MGI Gene Symbol</b>	<b>Description</b>	<b>Number of peptides</b>
<i>Taf6</i>	Transcription initiation factor TFIID subunit 6 - Mus musculus (Mouse)	244
<i>Taf5</i>	Transcription initiation factor TFIID subunit 5 - Mus musculus (Mouse)	141
<i>Taf9</i>	Transcription initiation factor TFIID subunit 9 - Mus musculus (Mouse)	123
<i>Taf1</i>	Transcription initiation factor TFIID subunit 1 - Mus musculus (Mouse)	92
<i>Taf9b</i>	Transcription initiation factor TFIID subunit 9B - Mus musculus (Mouse)	32
<i>Taf2</i>	Transcription initiation factor TFIID subunit 2 - Mus musculus (Mouse)	50
<i>Taf7</i>	Transcription initiation factor TFIID subunit 7 - Mus musculus (Mouse)	44
<i>Taf4b</i>	Transcription initiation factor TFIID subunit 4B - Mus musculus (Mouse)	116
<i>Taf3</i>	Transcription initiation factor TFIID subunit 3 - Mus musculus (Mouse)	74
<i>Taf11</i>	Transcription initiation factor TFIID subunit 11 - Mus musculus (Mouse)	19
<i>Taf8</i>	Transcription initiation factor TFIID subunit 8 - Mus musculus (Mouse)	10
<i>Taf10</i>	Transcription initiation factor TFIID subunit 10 - Mus musculus (Mouse)	12
<i>Taf12</i>	Transcription initiation factor TFIID subunit 12 - Mus musculus (Mouse)	18
<i>Tbp</i>	Transcription initiation factor TFIID subunit TATA-box-binding protein - Mus m	11
<i>Taf13</i>	Transcription initiation factor TFIID subunit 13 - Mus musculus (Mouse)	4

# *MANUSCRIT N°2*



## Manuscrit n°2

**Emilie Cler, Isabelle Duluc, Marc Koch, Yves Lutz, Doulaye Dembélé, Jean-Noel Freund, and Irwin Davidson.**

L'inactivation de la sous-unité TAF4 du facteur général de transcription TFIID dans les fibroblastes embryonnaires murins (MEFs) conduit à l'activation de la voie de signalisation de TGF $\beta$  et à la croissance autocrine sérum-indépendante.

Dans cette étude, nous montrons que les MEFs *Taf4*<sup>-/-</sup> sont transformées de manière oncogénique et qu'elles forment des sarcomes sous-cutanés indifférenciés *in vivo* dans des souris immuno-déficientes. Le traitement de ces MEFs avec de l'acide rétinoïque tout-trans (AR), préalablement avant l'injection supprime la formation des tumeurs, mais n'affecte pas leur prolifération ou leur cycle cellulaire. Nos résultats suggèrent plutôt que l'AR induit un changement dans la morphologie, la motilité, l'alignement et l'adhérence des MEFs *in vitro*.

Une étude transcriptomique de l'expression des gènes affectés montre que plusieurs gènes sont impliqués dans le remodelage de la matrice extracellulaire (MEC) parmi les gènes dérégulés par l'AR. Nous suggérons donc que la suppression tumorale liée à l'ajout d'AR ne résulte pas d'un effet direct sur la prolifération et la différenciation mais plutôt d'un nouveau mécanisme engendrant des modifications du microenvironnement tumoral via la régulation de gènes impliqués dans le remodelage de la MEC.



**Retinoic acid suppresses tumour formation of TAF4-null fibroblasts in nude mice.**

Emilie Cler, Isabelle Duluc<sup>1</sup>, Marc Koch, Yves Lutz, Doulaye Dembélé, Jean-Noel Freund<sup>1</sup>,  
and Irwin Davidson#.

Institut de Génétique et de Biologie Moléculaire et Cellulaire.

CNRS/INSERM/UDS.

1 Rue Laurent Fries,

67404 Illkirch Cédex.

France.

FAX: 33 3 88 65 32 01. TEL: 33 3 88 65 34 40 (45)

1. Inserm U682 - Développement et Physiopathologie de l'Intestin et du Pancréas, 3 Avenue  
Molière, 67200, Strasbourg, France.

# To whom correspondence should be addressed

E mail : [irwin@titus.u-strasbg.fr](mailto:irwin@titus.u-strasbg.fr)

Running Title : TAF subcomplex

Key words : TFIID, TGM2, Col6A3, TGF $\beta$ , extracellular matrix.

**Abstract.**

Inactivation of the TAF4 subunit of general transcription factor TFIID in mouse embryonic fibroblasts (MEFs) leads to activation of the TGF $\beta$ -signalling pathway and serum-independent autocrine growth. In this study, we show that the *Taf4*<sup>-/-</sup> MEFs are oncogenically transformed and form undifferentiated sarcoma-like subcutaneous tumours *in vivo* in nude mice. Treatment of these MEFs with all-trans retinoic acid (RA) prior to injection suppresses tumour formation, but does not affect their proliferation or cell cycle. Our results rather show that, RA induces changes in morphology, motility, alignment and adhesion of the MEFs *in vitro*. A detailed transcriptome assay shows that amongst the genes most deregulated by RA, are several involved in remodelling of the extracellular matrix (ECM). We suggest that tumour suppression by RA does not result from a direct effect on MEF proliferation and differentiation, but rather through a novel mechanism involving modification of the tumour microenvironment via regulation of genes involved in ECM remodelling.



## Introduction.

General transcription factor TFIID is composed of the TATA-binding protein (TBP) and 13-14 evolutionary conserved TBP-associated factors (TAFs) (Tora, 2002). These factors assemble to form a large multi-protein complex comprising three large lobes around a central cavity (Grob et al., 2006 ; Leurent et al., 2004 ; Matangkasombut et al., 2004). TAF4 is a ubiquitously expressed TFIID subunit that is essential in yeast and plays a general role in transcription in *C.elegans* embryos (Sanders and Weil, 2000 ; Walker et al., 2001). TAF4 is a component of the TFIID core complex and is essential for TFIID integrity (Cler et al., 2009; Grob et al., 2006; Leurent et al., 2004; Papai et al., 2009; Wright et al., 2006). The TAF4 family also comprises TAF4b, that is required for testis and ovary function (Falender et al., 2005; Freiman et al., 2001).

Inactivation of TAF4 in the epidermis of the adult mice leads to enhanced EGF signalling and increased keratinocyte proliferation. TAF4 also has cell-autonomous and non-cell-autonomous tumour suppressor activity in the epidermis (Fadloun et al., 2007b). In addition, we have characterised *Taf4<sup>-/-</sup>* embryonic fibroblasts where TAF4 inactivation promotes altered cell morphology and serum-independent autocrine growth accompanied by a deregulation (both positive and negative) of more than 1000 genes (Mengus et al., 2005). Amongst the affected genes, are critical components of transforming growth factor (TGF) signalling pathway such as the TGF $\beta$ 1, and TGF $\beta$ 3 ligands (for review see (Shi and Massague, 2003) and connective tissue growth factor (CTGF) that has intrinsic growth factor properties and interacts directly with TGF $\beta$ 1 to facilitate binding to its surface receptor enhancing the proliferative response of fibroblasts to TGF $\beta$  (Abreu et al., 2002). In addition, loss of TAF4 induces matrix metalloproteases and thrombospondin that activate latent TGF $\beta$  (Annes et al., 2003 ; Yu and Stamenkovic, 2000). As a consequence, *Taf4<sup>-/-</sup>* cells

secrete high levels of active TGF $\beta$  inducing a positive autoregulatory feedback loop of TGF $\beta$  signalling.

We have also shown that TAF4 is a coactivator for the all-trans retinoic acid receptor (RAR) transcription factor (Gangloff et al., 2000; Mengus et al., 1997) as the activity of transfected RAR is strongly diminished in *Taf4<sup>-/-</sup>* cells (Mengus et al., 2005). Treatment of the *Taf4<sup>lox/-</sup>* and *Taf4<sup>-/-</sup>* cells with RA shows that TAF4 is required for regulation of most, but not all cellular genes (Fadloun et al., 2007a). Our results also revealed that, like *Taf4<sup>-/-</sup>* cells, RA treated *Taf4<sup>lox/-</sup>* cells exhibit TGF $\beta$ -dependent autocrine growth. Chromatin immunoprecipitation experiments indicated that RARs bind directly to elements in the TGF $\beta$ 3 and CTGF promoters and regulate their expression (Delacroix, 2009). We also showed that overexpression of CTGF is sufficient for induce autocrine growth (Fadloun et al., 2007a). The direct regulation of these two critical ligands by the RARs provides a mechanism of cross-talk between the RA and TGF $\beta$  pathways that leads to deregulated cell growth.

In this study, we show that the *Taf4<sup>-/-</sup>* MEFs not only show autocrine growth *in vitro*, but are oncogenically transformed and form tumours *in vivo* in nude mice. Although RA treatment of the *Taf4<sup>lox/-</sup>* MEFs mimics the loss of TAF4 by inducing their autocrine growth *in vitro*, they do not form tumours in nude mice. Moreover, RA suppresses the ability of the *Taf4<sup>-/-</sup>* MEFs to form tumours. RA does not however affect proliferation, cell cycle, and apoptosis of these cells *in vitro*. Our results rather shown that, RA induces changes in morphology, motility, alignment and adhesion of the cells *in vitro*. A detailed transcriptome assay identifies around 186 genes that are up or down-regulated after 3 days of exposure to RA, but only a subset of these genes show a strong regulation by RA. Amongst these, we identify several genes that may contribute to the tumour suppressor effect by affecting the interactions of the *Taf4<sup>-/-</sup>* MEFs with the extracellular matrix.

## **Material and Methods.**

### **Cell lines.**

The C1*Taf4*<sup>lox/-</sup> and C3*Taf4*<sup>-/-</sup> MEFs are as previously described (Mengus et al., 2005) and were cultured in Dulbecco's minimal essential media (DMEM) supplemented with glutamax and 10% foetal calf serum. 1205Lu cells were cultured in MCDB153/Leibovitz L15 3/1 with 4%FCS. Cells were treated with 10<sup>-6</sup> M RA dissolved in ethanol (DMSO) as indicated.

### **Nude mouse injection.**

C1*Taf4*<sup>lox/-</sup> and C3*Taf4*<sup>-/-</sup> MEFs were treated with RA or with solvent for 3 days prior to injection. The cells were harvested with trypsin, resuspended, and counted before injection. Nude mice (Charles River France) were subcutaneously injected with between 2.10<sup>6</sup> cells per injection point. Animals were sacrificed in accordance with the French national animal ethics regulations and the tumours were excised, measured and fixed in paraformaldehyde solution, embedded in paraffin and 5µm-thick sections were used for Hematoxylin-Eosin staining.

### **Proliferation assays.**

For cell counting, 5x10<sup>4</sup> cells were seeded in 10cm plates. At the indicated times the cells were trypsinised and counted. BrdU incorporation was performed using a BrdU Cell Proliferation Assay kit (QIA58) from Merck/CalBiochem as per the manufacturers instructions. Briefly, the cells were grown for the indicated times with or without RA and then BrdU incorporation for 24 hours was measured by an immuno-colometric assay.

### **Time-Lapse Microscopy (wound healing)**

Mouse embryonic fibroblasts cells were grown to confluence on 24-well polystyrene plates Flat Bottom with low evaporation lid (Becton Dickinson) 24 hours before imaging. A wound was created by scraping the monolayer with 200-µl disposable plastic pipette tips. The

plates were placed in a chamber fixed to a motorized platform of an inverted microscope (Leica DMIRE2) and maintained at 37°C with 5% CO<sub>2</sub> (Pecon GmbH). Images (10x objective, N Plan NA: 0.25; phase contrast, Leica Microsystems) were collected every 15 min for 30 h with a CoolSnap FX camera (Roper Scientific) using Metamorph software (Molecular Devices). Image analysis, editing and velocities measurements (kymograph function) were performed by the use of Metamorph software (Molecular Devices).

### **Time-Lapse Microscopy (tracking)**

Mouse embryonic fibroblasts cells were seeded on 24-well polystyrene plates Flat Bottom with low evaporation lid (Becton Dickinson) and incubated 5 hours with 0.02 µg/ml of Hoechst 33342 before imaging. The plates were placed in a chamber fixed to a motorized platform of an inverted microscope (Leica DMIRE2) and maintained at 37°C with 5% CO<sub>2</sub> (Pecon GmbH). In order to minimize phototoxicity and photodamage, fluorescence excitation was performed with the use of a metal halide lamp set at 25% of maximum intensity (EL6000, Leica Microsystems), with an exposure time of 30 msec and a fluorescence filterset adapted to Hoechst observation (Emission filter : LP 425 nm). Images (20x objective, N Plan NA: 0.40; phase contrast, Leica Microsystems) were collected every 15 min for 30 h with a CoolSnap HQ<sup>2</sup> camera (Roper Scientific) using Metamorph software (Molecular Devices). Images were processed with an average filter (radius 100) that was subtracted from the original images and then with a median filter using Timt software (IGBMC). Nuclei were tracked with the track objects application from Metamorph software (Molecular Devices).

### **Adhesion and invasion assays.**

Adhesion on the indicated substrates was performed using CYTOSELECT™ 48 well cell adhesion kits from Cell Biolabs (CBA-050, 052, 056 and 060) as per manufacturers instructions. Cells grown for 3 days in presence and absence of RA were trypsinised, counted and  $1.5 \cdot 10^4$  cells were seeded on each substrate. After 90 minutes incubation at 37°C, the cells

were fixed, stained with crystal violet and washed. The remaining stained cells were extracted and the amount of stain determined by OD at 560 nm. Invasion assays were performed as described (Javelaud et al., 2005). Briefly, cells were grown for 3 days in presence or absence of RA, trypsinised and seeded on the upper well of Tissue culture transwell inserts (8-um pore size, Falcon, Franklin Lakes, NJ, USA) coated with 25-100ug/ml growth factor-reduced Matrigel™ (Biocoat, BD Biosciences, San Jose, CA, USA). After 12 hours of incubation at 37°C, the cells on the matrigel surface were stained with Diff-Quik™ (Medion Diagnostics) and observed under the microscope. After 48 hours at 37°C, the upper surface of the membrane was removed using a cotton swab and the cells that had crossed the membrane to adhere to the lower surface were stained and visualised by microscopy.

#### **Transcriptome assay.**

Triplicate independent preparations of total RNA from the *C3taf4<sup>-/-</sup>* cells exponentially growing in 10% serum in the presence or absence of RA were analysed on the Affymetrix mouse Gene 1.0 ST array Chip comprising 770317 probe sets representing 28,853 mouse genes. RMA normalisation and statistical analysis were performed as previously described (Mengus et al., 2005). RT-qPCR was performed as previously described and the expression of the target genes normalised to that of the ribosomal gene *Rplp0* (Fadloun et al., 2007a). Primer sequences are available on request.

## Results.

### TAF4 is an anti-oncogene.

We have previously reported that TAF4 inactivation in *Taf4<sup>lox/-</sup>* MEFs leads to pronounced morphological changes and TGFβ-dependent autocrine growth (Mengus et al., 2005). Compared to the *taf4<sup>lox/-</sup>* MEFs (hereafter C1) that adopt a regular shape in cell culture, the *taf4<sup>-/-</sup>* MEFs (hereafter C3) are irregularly shaped (Fig. 1A and B). Moreover, we noted that the C3 cells seemed to have lost contact inhibition as they readily form three dimensional foci, but this was never observed with the C1 cells. These observations prompted us to test the ability of the C1 and C3 cells to grow in soft agar. While the C1 cells did not proliferate in soft agar, the C3 cells formed clearly visible colonies (Fig. 1C). TAF4 inactivation therefore induces profound changes in the growth properties of MEFs *in vitro*.

Autocrine growth, loss of contact inhibition, and growth in soft agar are characteristics of oncogenically transformed cells. To ask whether the C3 cells are oncogenically transformed, nude mice were injected at up to four distinct positions with either the C1 or C3 cells. The C3 cells consistently formed large subcutaneous tumours in nude mice, while no tumours were observed with the C1 cells (Fig. 2A and B). Oncogenic transformation is therefore directly associated with loss of TAF4 expression. As we had previously shown that TAF4b replaces TAF4 in the TFIID complex in the C3 cells, this suggests that it is TAF4b that has a pro-oncogenic function. Over-expression of TAF4b in the C3 cells did not affect their ability to form tumours, and even slightly retarded tumour growth (data not shown, summarised in Table 1).

Within 40 days most tumours had grown so large as to impose the euthanasia of the injected animals (Figs. 2C, D). Hematoxylin-Eosin stained sections from the tumours indicated that the cells formed large undifferentiated sarcoma-type tumours irrigated with

numerous blood vessels that were locally invasive in the underlying dermal muscle and adipocyte layers (Fig. 2E and F). However, we never observed spreading to adjacent lymph nodes nor distant metastases in the sacrificed animals. Also while tumours formed readily when the cells were injected subcutaneously, we did not observe tumours when the cells were injected in the liver or pancreas (data not shown).

These observations indicate that TAF4 has an anti-oncogene function as its inactivation in MEFs leads to their ability to form tumours *in vivo*.

### **RA suppresses tumour formation by *taf4*<sup>-/-</sup> MEFs.**

We have previously characterised the changes in gene expression that result from loss of TAF4 where the expression of several proto-oncogenes and growth factors are induced, amongst which are several critical components of the TGF $\beta$  signalling pathway (Mengus et al., 2005). We have further shown that RA treatment of the C1 MEFs mimics TAF4 inactivation by activating the TGF $\beta$ -signaling pathway leading to autocrine growth (Fadloun et al., 2007a). This observation suggests that RA-treatment may also confer on the C1 MEFs the ability to form tumours in nude mice. To test this possibility, the C1 and C3 cells were treated *in vitro* for 3 days with RA before injection. Similar to untreated C1 cells, RA treated C1 cells never formed tumours (Fig. 2A and B and Table 1). Strikingly, however RA treatment of the C3 cells completely suppressed their ability to form tumours. These results show that activation of TGF $\beta$  signalling and autocrine growth *in vitro* can be dissociated from the ability to form tumours *in vivo* (see Table 1).

To understand how RA suppresses the ability of the C3 cells to form tumours, we investigated the effect of RA treatment on their *in vitro* properties. RA treatment leads to a change in morphology of the C3 cells which adopt a more regular elongated shape, similar to

the C1 cells, and to a restoration of contact inhibition as the treated cells do not form three dimensional foci (Fig. 1A and B).

We investigated the effect of RA treatment on cell proliferation. Growth kinetics performed by cell counting did not show any significant difference between the untreated and RA treated C3 cells (Fig. 3A). Similarly, RA treatment did not lead to significant changes in BrdU incorporation over a 5 day period (Fig. 3B). Finally RA treatment did not modify the cell cycle as measured by FACS analysis (Fig. 3C). In particular, no significant G1/S arrest similar to what is seen with F9 cells is observed with the MEFs. These results show that RA does not suppress tumour formation by inducing cell cycle arrest or apoptosis of the C3 cells.

#### **RA induces changes in C3 cell motility and adhesion.**

We next verified whether RA treatment could affect other properties of the C3 cells known to affect tumorigenicity such as their motility, adhesion on different substrates, or invasive capacity. To do this we first performed wound healing assays. C3 cells were grown for 3 days as monolayers in the presence or absence of RA before the introduction of a wound in the monolayer using a pipette tip. The migration of cells into the cleared area was then followed by time lapse photography. With untreated cells, the two cell fronts required between 14-16 hours to merge (Fig. 4A and supplemental movies 1-3). In contrast, with the RA-treated cells, merging required only 10-12 hours (Fig. 4B and supplemental movies 4-6). In the above assays, we also noted that RA alters the alignment of the cells at the leading edge. The untreated cells tended to align themselves parallel to the cleared area, while the RA treated cells rather aligned perpendicular to the cleared area (see bars in Fig. 4A and B and supplemental movies). Thus, RA treatment of the C3 MEFs leads to an increase in motility and to changes in polarity of alignment with respect to the cell front.



Comparison of the untreated C1 and C3 cells showed that the C1 cells exhibited a faster migration than the C3 cells (compare Figs. 4A and 5A and supplemental movies 1-3 and 7-9). In contrast however to the above, we observed that RA had the opposite effect on the motility of the C1 MEFs. With untreated C1 cells, the two cell fronts merged between 10-12 hours, while more than 17 hours were required to achieve the same state with the RA treated cells (Fig. 5A and B and supplemental movies 7-9 compared to 10-12). RA treatment also affected alignment with respect to the cell front. As observed with the C3 cells, the untreated C1 cells, tended to align parallel to the leading edge, while the RA treated cells aligned perpendicular to the cell front. These results show that RA treatment has opposite effects on the motility of the C1 and C3 cells, but has a similar effect on cell alignment.

We further developed an assay to track the movement of individual cell nuclei by fluorescent time lapse microscopy (see Material and methods). In this assay the cells are plated at low density rather than as monolayers and individual cells are followed (see supplemental movies 13-14). This assay indicated that the C3 cells show a higher motility than the C1 cells (Table 2). However, no significant effect of RA on the C3 cells was observed, while RA appeared to increase the motility of the C1 cells, in contrast to what was observed in the wound healing assay.

We next investigated whether RA affected cell adhesion properties. The C3 cells were treated for 0-72 hours with RA and then plated on laminin, fibronectin, collagen I or collagen IV coated wells. After 2 hours the cells were fixed, stained and the number of adhered cells measured by absorbance at 550 nm. The untreated C3 cells showed similar adherence to all of the different substrates (Fig. 6A-D). RA treatment did not lead to significant changes in adherence on laminin and only a mild increase on fibronectin (Fig. 6A-B). In contrast, RA substantially increased adhesion on both the collagen substrates (Fig. 6C-D). The most

significant effects were seen after 48-72 hours of RA treatment, the time that is normally required to observe the differences in morphology.

The effect of RA on invasive capacity was also assessed using matrigel membranes. The C3 cells were treated for 0 or 5 days with RA and plated in transwell chambers with a Matrigel coated membrane. 12 hours after plating the cells adhered on the surface of the Matrigel were visualised and after 48 hours the upper surface of the membrane was removed and the number of cells that had crossed to the lower surface of the membrane was assessed. As a control we used 1205Lu cells, a melanoma derived cell line with known invasive and tumorigenic properties (Javelaud et al., 2007). After 12 hours we observed that RA significantly augmented the number of C3 cells that adhered on the Matrigel membrane (Fig. 6E). In contrast, neither the untreated nor RA treated C3 cells were capable of invading and crossing the Matrigel layer showing that they had no significant invasive capacity in this assay, whereas the 1205Lu cells could be clearly observed. We repeated this assay with different concentrations of matrigel (25-100 ug/ml), but no invasion of the Matrigel layers by the C3 cells was observed, while the 1205Lu cells consistently exhibited invasive capacity (data not shown).

The above results show that RA increases the adhesive capacity of the C3 cells on collagen and matrigel substrates, but the C3 cells do not show invasive capacity either in absence or presence of RA.

### **RA-induced changes in gene expression.**

The above results document the effects of RA on the motility and adherence properties of the C3 cells. To understand the molecular basis of these effects, we used Affymetrix gene arrays to investigate the changes in gene expression that are induced by RA. We have previously described the RA transcriptome of these cells after 24 hours of treatment.

However, the results described here indicate that the changes in morphology and other properties are evident only after at least 3 days in presence of RA. We therefore, re-examined the transcriptome after 12 hours of RA treatment to identify mostly the direct target genes, and then at 72 hours to characterise later events that take place when the effects of RA on the properties of the cells become evident.

After 12 hours of RA treatment, a greater than 2-fold increase in the expression of 75 genes was observed (Supplemental Table 1). Amongst these are several known direct target and responsive genes such as cytochrome P450, family 26, subfamily b, polypeptide 1 (*Cyp26b1*), the RAR beta receptor (*Rarb*), transglutaminase 2, C polypeptide (*Tgm2*) and basic helix-loop-helix family, member e40 (*Bhlhe40*, also known as *Bhlhb2* or *Stra13/14*). (Fig. 7A). After 72 hours, we observed deregulation of 186 genes (not counting the positive and negative array controls) of which 170 were up-regulated and 16 down-regulated (supplemental Table 1 and Fig. 7A). The changes in expression of several genes was verified by RT-qPCR. As observed on the array, *Tgm2* and dehydrogenase/reductase (SDR family) member 3 (*Dhrs3*) were up-regulated, while membrane metallo-endopeptidase (*Mme*), preproenkephalin (*Penk*), and collagen, type VI, alpha 3 (*Col6a3*) were down-regulated (Fig. 7B).

A closer inspection of the results shows that only a limited number of genes show strong changes in expression. Only 21 genes are up-regulated 4 fold or more and while only 2 are down-regulated 4 fold or more. These observations are in accordance with our previous results showing a reduced RA response of the C3 cells that do not express TAF4. However, comparison of this new data set with the previous one revealed a number of significant differences. Many genes identified as up-regulated or down-regulated on the previous set of arrays were not considered affected on the present arrays, and *vice-versa*. These discrepancies may be explained by the fact that the Affymetrix array design has been significantly changed

between the time of the two experiments and that the new arrays contain probes for transcripts that were not present on the older arrays. This does not however explain the fact that genes identified as regulated on the older arrays do not appear regulated on the new arrays. We therefore cannot exclude the possibility that we underestimate the number of regulated genes (see discussion).

## **Discussion.**

### **TAF4 is an anti-oncogene.**

We have previously shown that TAF4 inactivation in MEFs leads to profound changes in the growth properties of MEFs *in vitro*, however we were unable to show at that time that the C3 cells were capable of forming tumours *in vivo* (Mengus et al., 2005). In this study, we show that the C3 cells are capable of forming tumours *in vivo*. These cells reproducibly form large subcutaneous tumours that invade the local environment, but do not form distant metastases. The tumorigenic properties of these cells do not seem to be directly due only to the activation of the TGF $\beta$ -signalling pathway or to correlate with the autocrine growth phenotype seen *in vitro*, as RA-treated C1 cells do not form tumours *in vivo*, yet exhibit both of these phenotypes *in vitro*. TAF4 inactivation therefore affects an additional pathway(s) not affected by RA treatment that leads to oncogenic transformation.

A re-examination of our previous transcriptome results, focussing on the genes that are deregulated by TAF4 inactivation, but not affected by RA, identifies the gene encoding ubiquitin specific peptidase 18 as strongly induced (Fadloun et al., 2007a). USP18 overexpression has recently been shown to upregulate the synthesis of the epidermal growth factor receptor (EGFR) (Duex and Sorokin, 2009). It is also interesting to note that TAF4 inactivation upregulates heparin binding epidermal growth factor (HBEGF) expression. It is therefore possible that the ability of the C3 cells to form tumours results from increased EGF

signalling due to the coordinate upregulation of USP18 and HBEGF. This is reminiscent of the phenotype observed upon TAF4 inactivation in the epidermis where the EGF signalling pathway is activated (Fadloun et al., 2007b). Although our results show that activation of the TGF $\beta$ -signalling pathway is not sufficient to induce tumour formation, perhaps the TGF $\beta$  and EGF pathways act together to promote tumour formation by the C3 cells. These possibilities could be tested by using shRNAs to downregulate components of each pathway in C3 cells followed by assessment of the ability of the knockdown cells to form tumours in nude mice.

### **A novel mechanism of tumour suppression by RA.**

A major observation of this study is that RA can suppress the formation of tumours by C3 cells. Treatment of these cells with RA for 3 days prior to injection in nude mice leads to a complete loss of tumour formation. We have yet to determine whether shorter RA treatments can have the same effect or whether a prolonged exposure is required. Strikingly however, RA treatment does not significantly alter the growth properties of the C3 cells *in vitro* under standard growth conditions. We did not observe growth arrest and FACS analysis did not reveal changes in cell cycle or the appearance of a large apoptotic cell population. RA treatment of the C3 cells also did not abrogate their serum-free autocrine growth (Fadloun et al., 2007a). The anti-tumour effect of RA on the C3 MEFs is therefore fundamentally different from that seen in F9 embryonal carcinoma cells or in HL60 myeloid cells where RA treatment induces cell cycle arrest, differentiation and under some conditions apoptosis (Licht, 2006; Licht, 2009).

The above observations prompted us to ask whether RA affected other properties of the C3 cells to inhibit tumorigenesis. Our results show that RA increases the motility of the C3 cells. This is clearly seen in the wound healing assay where the RA treated cells rapidly migrate across the cleared area, but is not so clearly observed in the tracking assay.

Paradoxically, this is the opposite of what one would expect in terms of tumour suppression as oncogenically transformed cells often show increased motility. It is also interesting to note that RA has the opposite effect on the C1 cells where it slows their migration in the wound healing assay, but appears to increase motility in the tracking assay. Thus while RA significantly affects the migratory properties of the both the C1 and C3 cells this does not lead to any obvious insights into the tumour suppressor mechanism.

Comparison of the untreated C1 and C3 cells indicates that the C1 cells exhibit a more rapid migration than the C3 cells in the wound healing assay. However, in the tracking assay the opposite is observed as the C3 cells appear to have a higher motility than the C1 cells. The discrepancy with the wound healing assay may be explained by the fact the C3 cells tend to move in all directions, while the C1 cells move preferentially in the direction of the cleared area such that it is more rapidly closed. A higher motility of the C3 cells would be expected considering that they are tumourogenic, however the conflicting results observed in the two assays do not allow a clear interpretation of the results with respect to their differential abilities to form tumours.

Our results show that RA increases C3 cell adhesion to collagen matrices, but has little effect on adhesion to laminin and fibronectin substrates. The significance of the increased adherence of the C3 cells to collagen substrates for tumour formation remains to be determined, but it suggests that RA may act to inhibit tumourogenesis of the C3 cells by altering their interactions with the extracellular matrix (ECM) and their motile properties (see below). Thus while RA does not affect C3 cell proliferation under normal culture conditions *in vitro*, the altered adhesive properties may lead to significant effects on proliferation in the *in vivo* tumour environment. Indeed, it is well established that interactions between cells and the ECM, especially in a 3-dimensional matrix, has a major influence on cell growth and survival (Alavi and Stupack, 2007; Marastoni et al., 2008). In a preliminary experiment, we

observed that while the RA treated C3 cells showed increased adhesion to collagen substrates, they showed diminished long term survival. Further experiments will be required to determine whether this is indeed the case.

### **RA-regulated genes involved in tumour suppression.**

We performed a transcriptome analysis to identify genes whose deregulation may account for the effects of RA on motility and adhesion and tumour inhibition. This analysis identified a limited number of regulated genes of which only a subset show a strong changes in expression. Strikingly, we did not observe changes in expression of integrins that may account for the altered adhesive properties. RA does however induce the expression of collagen 8A1 that modifies adhesion and cell survival (Zhao et al., 2009) and Netrin 1 a secreted protein related to laminin that is known to influence adhesion and migration (Bernet and Fitamant, 2008; Mehlen and Furne, 2005).

Analysis of the gene expression changes also provide possible mechanisms for tumour suppression. RA strongly induces the expression of TGM2, a multi-functional protein that can remodel the ECM. Several studies have shown that TGM2 expression is down-regulated in transformed cells and can be reactivated by histone deacetylase inhibitors or the DNA demethylation agent 5-aza-2'-deoxycytidine leading to anti-tumour activity (Ai et al., 2008; Liu et al., 2007). This may be partly due to its ability to crosslink ECM proteins forming a barrier to tumour progression and inhibition of angiogenesis (Haroon et al., 1999; Jones et al., 2006; Mangala et al., 2005), but there is also evidence that TGM2 may also have intracellular functions notably acting to modulate apoptosis (Milakovic et al., 2004). It is therefore possible that RA induction of TGM2 may influence the adhesive and tumorigenic properties of the C3 cells.

A second gene that may play an important role in tumour suppression is collagen 6A3. This gene is strongly down-regulated after 72 hours of RA treatment. It has been shown that ovarian cancer cells resistant to cisplatin show a potent induction of the Col6A3 gene (Sherman-Baust et al., 2003). Moreover, overexpression of Col6A3 augments cisplatin resistance and high grade tumours *in vivo* express higher levels of Col6A3 than low grade tumours. These observations indicate that Col6A3 expression positively correlates with tumourogenicity. The potent repression of this gene may therefore contribute to the tumour suppressor effect of RA in C3 cells.

The two genes discussed above are amongst the most strongly regulated by RA and both influence ECM remodelling and tumour microenvironment. In addition, we also note that the most induced gene *Prss35* encodes a poorly characterised secreted serine protease normally expressed in the ovary and likely implicated in ECM remodelling. We therefore suggest that tumour suppression by RA in C3 cells does not result from a direct effect on cell proliferation and differentiation, but rather through a novel mechanism involving modification of the tumour microenvironment via regulation of genes involved in ECM remodelling.

This hypothesis could be tested by using shRNA to knockdown Col6A3 expression to determine whether this mimics the effect of RA with respect to tumour formation, or shRNA knockdown of TGM2 to ask if inhibition of its expression could abrogate the tumour suppressor effect of RA.

We noted in the transcriptome assay that prolonged exposure to RA resulted in upregulation of several members of the PRAME family of genes. PRAME is a protein expressed in melanoma cells that acts as a negative regulator of RARs (Epping et al., 2005). The upregulation of these genes perhaps acts to counteract the long term effects of RA in the C3 cells. We also noted that the results of the present transcriptome shows some significant



differences from that we previously reported. Some of these may result from differences in culture conditions, while others may reflect the different sensitivities of the Affymetrix array designs. We are at present re-evaluating the effects on gene expression by using high-throughput RNA sequencing rather than Affymetrix arrays. This is done by sequencing 3'-tags of the polyA-containing RNA in a SAGE type assay (digital gene expression). Comparison of these results with those of the Affymetrix arrays will allow us to compare the efficiencies of the two techniques and to achieve a more comprehensive description of the RA-induced changes in gene expression in C3 cells.

### **Acknowledgements**

We thank, J-L Vonesch, for help with imaging and C. Thibault for microarrays. This work was supported by grants from the CNRS, the INSERM, the Association pour la Recherche contre le Cancer, the Ligue Nationale et Départementale Région Alsace contre le Cancer. ID is an 'équipe labellisée' of the Ligue Nationale contre le Cancer. EC was supported by a fellowship from the Fondation pour la Recherche Médicale.

### **References.**

Abreu, J. G., Ketpura, N. I., Reversade, B., and De Robertis, E. M. (2002). Connective-tissue growth factor (CTGF) modulates cell signalling by BMP and TGF-beta. *Nat Cell Biol* 4, 599-604.

Ai, L., Kim, W. J., Demircan, B., Dyer, L. M., Bray, K. J., Skehan, R. R., Massoll, N. A., and Brown, K. D. (2008). The transglutaminase 2 gene (TGM2), a potential molecular marker for chemotherapeutic drug sensitivity, is epigenetically silenced in breast cancer. *Carcinogenesis* 29, 510-518.

Alavi, A., and Stupack, D. G. (2007). Cell survival in a three-dimensional matrix. *Methods Enzymol* 426, 85-101.

Annes, J. P., Munger, J. S., and Rifkin, D. B. (2003). Making sense of latent TGFbeta activation. *J Cell Sci* *116*, 217-224.

Bernet, A., and Fitamant, J. (2008). Netrin-1 and its receptors in tumour growth promotion. *Expert Opin Ther Targets* *12*, 995-1007.

Cler, E., Papai, G., Schultz, P., and Davidson, I. (2009). Recent advances in understanding the structure and function of general transcription factor TFIID. *Cell Mol Life Sci* *66*, 2123-2134.

Delacroix, L., Moutier, E., Altobelli, G., Legras, S., Poch, O., Choukrallah, M-A., Bertin, I., Jost, B., and Davidson, I. (2009). Cell-specific interaction of retinoic acid receptors with target genes in mouse embryonic fibroblasts and embryonic stem cells. *Molecular and Cellular Biology* *In Press*.

Duex, J. E., and Sorkin, A. (2009). RNA interference screen identifies Usp18 as a regulator of epidermal growth factor receptor synthesis. *Mol Biol Cell* *20*, 1833-1844.

Epping, M. T., Wang, L., Edel, M. J., Carlee, L., Hernandez, M., and Bernards, R. (2005). The human tumor antigen PRAME is a dominant repressor of retinoic acid receptor signaling. *Cell* *122*, 835-847.

Fadloun, A., Kobi, D., Delacroix, L., Dembele, D., Michel, I., Lardenois, A., Tisserand, J., Losson, R., Mengus, G., and Davidson, I. (2007a). Retinoic acid induces TGFbeta-dependent autocrine fibroblast growth. *Oncogene*.

Fadloun, A., Kobi, D., Pointud, J. C., Indra, A. K., Teletin, M., Bole-Feysot, C., Testoni, B., Mantovani, R., Metzger, D., Mengus, G., and Davidson, I. (2007b). The TFIID subunit TAF4 regulates keratinocyte proliferation and has cell-autonomous and non-cell-autonomous tumour suppressor activity in mouse epidermis. *Development* *134*, 2947-2958.

Falender, A. E., Freiman, R. N., Geles, K. G., Lo, K. C., Hwang, K., Lamb, D. J., Morris, P. L., Tjian, R., and Richards, J. S. (2005). Maintenance of spermatogenesis requires TAF4b, a gonad-specific subunit of TFIID. *Genes Dev* *19*, 794-803.

Freiman, R. N., Albright, S. R., Zheng, S., Sha, W. C., Hammer, R. E., and Tjian, R. (2001). Requirement of tissue-selective TBP-associated factor TAFII105 in ovarian development. *Science* 293, 2084-2087.

Gangloff, Y. G., Werten, S., Romier, C., Carre, L., Poch, O., Moras, D., and Davidson, I. (2000). The human TFIID components TAF(II)135 and TAF(II)20 and the yeast SAGA components ADA1 and TAF(II)68 heterodimerize to form histone-like pairs [In Process Citation]. *Mol Cell Biol* 2000Jan;20(1):340-51 20, 340-351.

Grob, P., Cruse, M. J., Inouye, C., Peris, M., Penczek, P. A., Tjian, R., and Nogales, E. (2006). Cryo-electron microscopy studies of human TFIID: conformational breathing in the integration of gene regulatory cues. *Structure* 14, 511-520.

Haroon, Z. A., Lai, T. S., Hettasch, J. M., Lindberg, R. A., Dewhirst, M. W., and Greenberg, C. S. (1999). Tissue transglutaminase is expressed as a host response to tumor invasion and inhibits tumor growth. *Lab Invest* 79, 1679-1686.

Javelaud, D., Delmas, V., Moller, M., Sextius, P., Andre, J., Menashi, S., Larue, L., and Mauviel, A. (2005). Stable overexpression of Smad7 in human melanoma cells inhibits their tumorigenicity in vitro and in vivo. *Oncogene* 24, 7624-7629.

Javelaud, D., Mohammad, K. S., McKenna, C. R., Fournier, P., Luciani, F., Niewolna, M., Andre, J., Delmas, V., Larue, L., Guise, T. A., and Mauviel, A. (2007). Stable overexpression of Smad7 in human melanoma cells impairs bone metastasis. *Cancer Res* 67, 2317-2324.

Jones, R. A., Kotsakis, P., Johnson, T. S., Chau, D. Y., Ali, S., Melino, G., and Griffin, M. (2006). Matrix changes induced by transglutaminase 2 lead to inhibition of angiogenesis and tumor growth. *Cell Death Differ* 13, 1442-1453.

Leurent, C., Sanders, S. L., Demeny, M. A., Garbett, K. A., Ruhlmann, C., Weil, P. A., Tora, L., and Schultz, P. (2004). Mapping key functional sites within yeast TFIID. *Embo J* 23, 719-727.

Licht, J. D. (2006). Reconstructing a disease: What essential features of the retinoic acid receptor fusion oncoproteins generate acute promyelocytic leukemia? *Cancer Cell* 9, 73-74.

Licht, J. D. (2009). Acute promyelocytic leukemia--weapons of mass differentiation. *N Engl J Med* 360, 928-930.

Liu, T., Tee, A. E., Porro, A., Smith, S. A., Dwarte, T., Liu, P. Y., Iraci, N., Sekyere, E., Haber, M., Norris, M. D., *et al.* (2007). Activation of tissue transglutaminase transcription by histone deacetylase inhibition as a therapeutic approach for Myc oncogenesis. *Proc Natl Acad Sci U S A* 104, 18682-18687.

Mangala, L. S., Arun, B., Sahin, A. A., and Mehta, K. (2005). Tissue transglutaminase-induced alterations in extracellular matrix inhibit tumor invasion. *Mol Cancer* 4, 33.

Marastoni, S., Ligresti, G., Lorenzon, E., Colombatti, A., and Mongiat, M. (2008). Extracellular matrix: a matter of life and death. *Connect Tissue Res* 49, 203-206.

Matangkasombut, O., Auty, R., and Buratowski, S. (2004). Structure and function of the TFIID complex. *Adv Protein Chem* 67, 67-92.

Mehlen, P., and Furne, C. (2005). Netrin-1: when a neuronal guidance cue turns out to be a regulator of tumorigenesis. *Cell Mol Life Sci* 62, 2599-2616.

Mengus, G., Fadloun, A., Kobi, D., Thibault, C., Perletti, L., Michel, I., and Davidson, I. (2005). TAF4 inactivation in embryonic fibroblasts activates TGFbeta signalling and autocrine growth. *Embo J* 24, 2753-2767.

Mengus, G., May, M., Carre, L., Chambon, P., and Davidson, I. (1997). Human TAF(II)135 potentiates transcriptional activation by the AF-2s of the retinoic acid, vitamin D3, and thyroid hormone receptors in mammalian cells. *Gene Develop* Vol 11, Iss 11, 1381-1395.

Milakovic, T., Tucholski, J., McCoy, E., and Johnson, G. V. (2004). Intracellular localization and activity state of tissue transglutaminase differentially impacts cell death. *J Biol Chem* 279, 8715-8722.

Papai, G., Tripathi, M. K., Ruhlmann, C., Werten, S., Crucifix, C., Weil, P. A., and Schultz, P. (2009). Mapping the initiator binding Taf2 subunit in the structure of hydrated yeast TFIID. *Structure* 17, 363-373.

Sanders, S. L., and Weil, P. A. (2000). Identification of two novel TAF subunits of the yeast *saccharomyces cerevisiae* TFIID complex [In Process Citation]. *J Biol Chem* 275, 13895-13900.

Sherman-Baust, C. A., Weeraratna, A. T., Rangel, L. B., Pizer, E. S., Cho, K. R., Schwartz, D. R., Shock, T., and Morin, P. J. (2003). Remodeling of the extracellular matrix through overexpression of collagen VI contributes to cisplatin resistance in ovarian cancer cells. *Cancer Cell* 3, 377-386.

Shi, Y., and Massague, J. (2003). Mechanisms of TGF-beta signaling from cell membrane to the nucleus. *Cell* 113, 685-700.

Tora, L. (2002). A unified nomenclature for TATA box binding protein (TBP)-associated factors (TAFs) involved in RNA polymerase II transcription. *Genes Dev* 16, 673-675.

Walker, A. K., Rothman, J. H., Shi, Y., and Blackwell, T. K. (2001). Distinct requirements for *C.elegans* TAF(II)s in early embryonic transcription. *Embo J* 20, 5269-5279.

Wright, K. J., Marr, M. T., 2nd, and Tjian, R. (2006). TAF4 nucleates a core subcomplex of TFIID and mediates activated transcription from a TATA-less promoter. *Proc Natl Acad Sci U S A* 103, 12347-12352.

Yu, Q., and Stamenkovic, I. (2000). Cell surface-localized matrix metalloproteinase-9 proteolytically activates TGF-beta and promotes tumor invasion and angiogenesis. *Genes Dev* 14, 163-176.

Zhao, Y., Jia, L., Mao, X., Xu, H., Wang, B., and Liu, Y. (2009). siRNA-targeted COL8A1 inhibits proliferation, reduces invasion and enhances sensitivity to D-limonene treatment in hepatocarcinoma cells. *IUBMB Life* 61, 74-79.



## **Legends to Figures.**

**Figure 1.** Effect of RA on cell growth and morphology. **A-B.** Phase contrast microscopy (20X magnification) of C1 and C3 cells grown for 3 days in presence or absence of RA as indicated. **C.** Growth of C1 and C3 cells after 10 days in soft agar.

**Figure 2.** C3 cells form tumours in nude mice. **A.** Photography of mice following injection of the indicated cell types to illustrate the presence or absence of tumours. **B.** Table summary of the number of mice, the number of individual injection points (several per animal), and the number of tumours for each cell line. **C.** Example of an excised tumour along with a scale rule. **D.** Graphic representation of the growth of 4 representative C3 cell tumours. **E-F** Hematoxylin-Eosin stained sections showing the histopathology of subcutaneous C3 tumours.

**Figure 3.** Effect of RA on C3 cell proliferation. **A.** Kinetics of cell growth in presence or absence of RA as evaluated by cell counting. **B.** Assessment of cell division by incorporation of BrdU on cells grown for the indicated periods in presence or absence of RA. **C.** Results of a representative FACS assay showing the % cells in each stage of cell cycle.

**Figure 4.** Effect of RA on C3 cell motility. **A-B.** Still images from timelapse films of wound healing assays on C3 cells in presence or absence of RA. The time of the still with respect to the wounding is indicated.

**Figure 5.** Effect of RA on C1 cell motility. **A-B.** Still images from timelapse films of wound healing assays on C1 cells in presence or absence of RA. The time of the still with respect to the wounding is indicated.

**Figure 6.** Effect of RA on C3 cell adhesion properties. **A-D.** C3 cells were treated for the indicated times before seeding on the different substrates. After 2 hours incubation fixing and staining the number of adhered cells is measured by OD at 550 nm. **E.** Phase contrast

microscopy (20 X magnification) of the surface of the Matrigel membrane after 12 hours of incubation (upper panel) or the lower surface of the Matrigel membrane after 48 hours incubation (lower panel). In the case of the C1 – and +RA (lower left and center panels), the dark black dots are the pores that allow contact with the medium in the lower chamber. These are seen as clear spots relative to the darkly stained 1205Lu cells in the bottom right panel.

**Figure 7.** RA induced changes in gene expression in C3 cells. **A.** Table resuming the changes of several strongly induced or repressed genes. Direct targets where binding of RARs to the corresponding promoters has been verified by chromatin immunoprecipitation are indicated. ? indicates that no RAR binding has yet been described. **B.** Results of RT-qPCR on the indicated genes in presence or absence of RA. Expression of the indicated genes was normalised against that of the ribosomal *Rplp0* gene that is not regulated by RA.

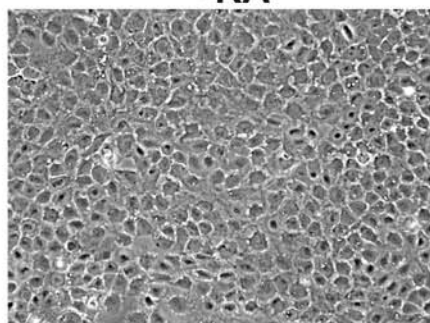
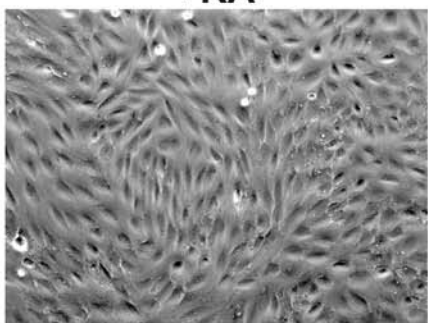
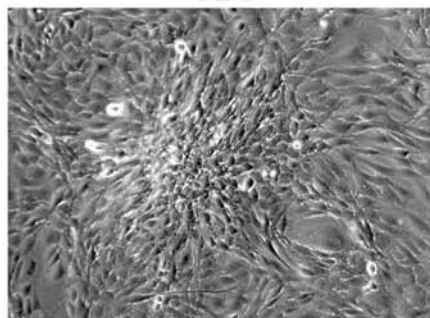
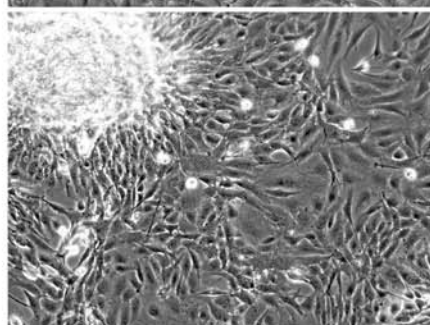
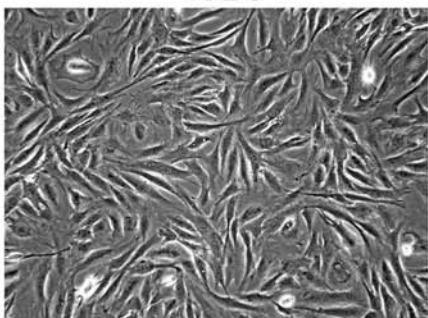
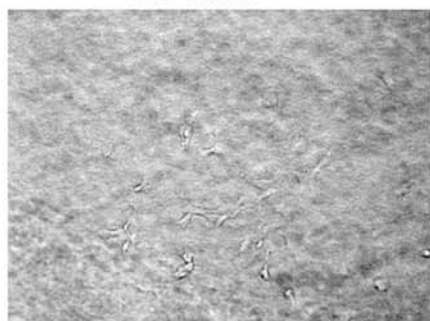
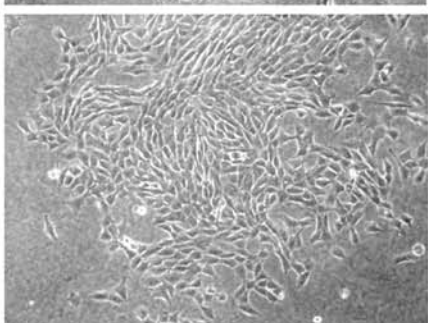
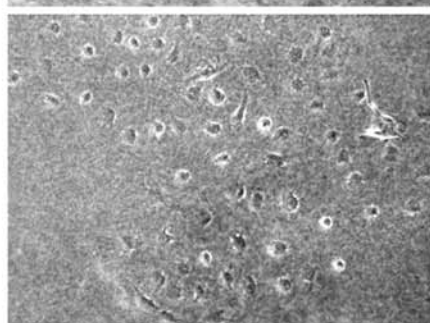
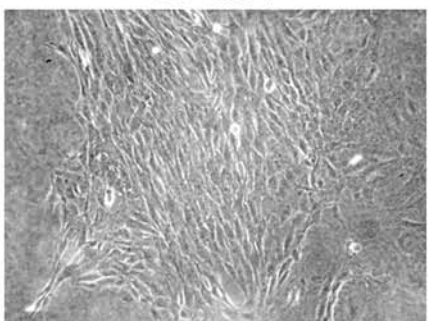
**Table 1.** Summary of the *in vitro* growth properties and tumorigenic capacity of the C1 and C3 cells as well as derived populations re-expressing wild-type TAF4 (*C3taf4<sup>-/-TAF4</sup>*) or expression exogenous TAF4b (*C3taf4<sup>-/-TAF4b</sup>*).

**Table 2.** Motility of C1 and C3 cells. Summary of the timelapse imaging of isolated C1 and C3 cell nuclei. The number of imaged nuclei for each condition is indicated along with the migration rate expressed as pixels/min.

**Supplemental Table 1.** Excel table of the Affymetrix array results. Page 1 shows genes de-regulated 2 fold or more after 12 hours of RA treatment, page 2 genes de-regulated after 72 hours and page 3, genes de-regulated between 12 and 72 hours. In each page, the Affymetrix probe set, MGI gene symbol, and mRNA accession number are shown as identifiers along with the fold change and the associated pvalue. In addition, gene ontology (GO) descriptions are provided for each gene. The final page shows the complete data set for all Affymetrix probes.



**Supplemental Movies. 1-3.** Time lapse imaging of untreated C3 cells following wounding of the monolayer. **4-6.** Time lapse imaging of RA-treated C3 cells. **7-9.** Time lapse imaging of untreated C1 cells. **10-12.** Time lapse imaging of RA-treated C1 cells. **13-14.** Fluorescent time lapse tracking of C3 cell nuclei.

**A****C1 *taf4*<sup>lox/-</sup>  
-RA****C1 *taf4*<sup>lox/-</sup>  
+RA****B****C3 *taf4*<sup>-/-</sup>  
-RA****C3 *taf4*<sup>-/-</sup>  
+RA****C****C1 *taf4*<sup>lox/-</sup> Soft Agar****C3 *taf4*<sup>-/-</sup>**

290

A



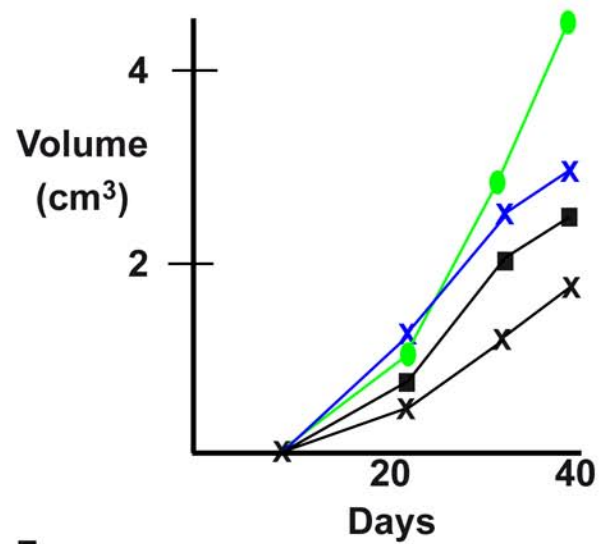
B

	C1 <i>Taf4</i> <sup>lox/-</sup> -RA	C3 <i>Taf4</i> <sup>-/-</sup> -RA	C3 <i>Taf4</i> <sup>-/-/taf4b</sup> -RA	C1 <i>Taf4</i> <sup>lox/-</sup> +RA	C3 <i>Taf4</i> <sup>-/-</sup> +RA
# Mice	7	9	5	4	7
# Injections	19	21	14	10	18
# Tumours	0	19	13	0	0

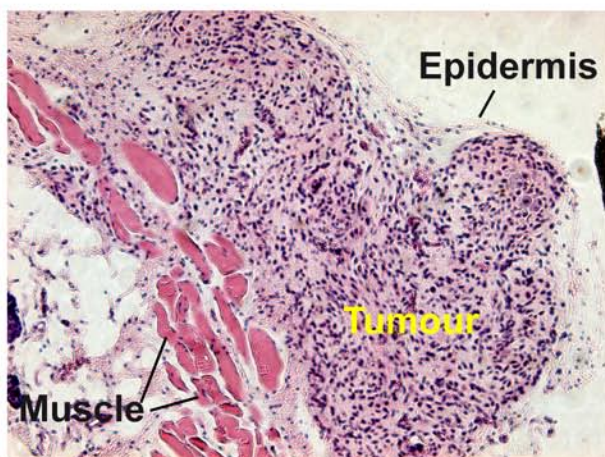
C



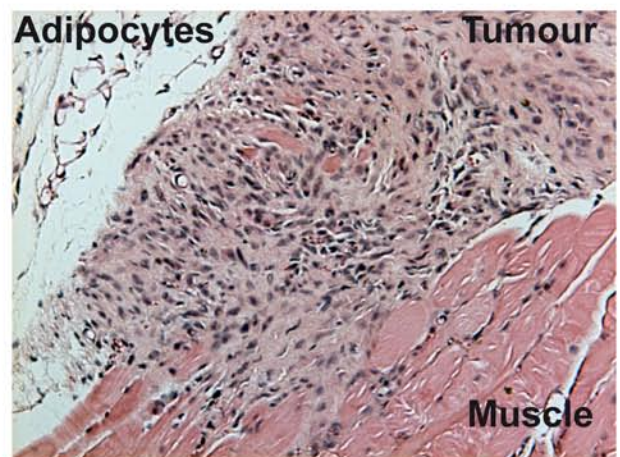
D

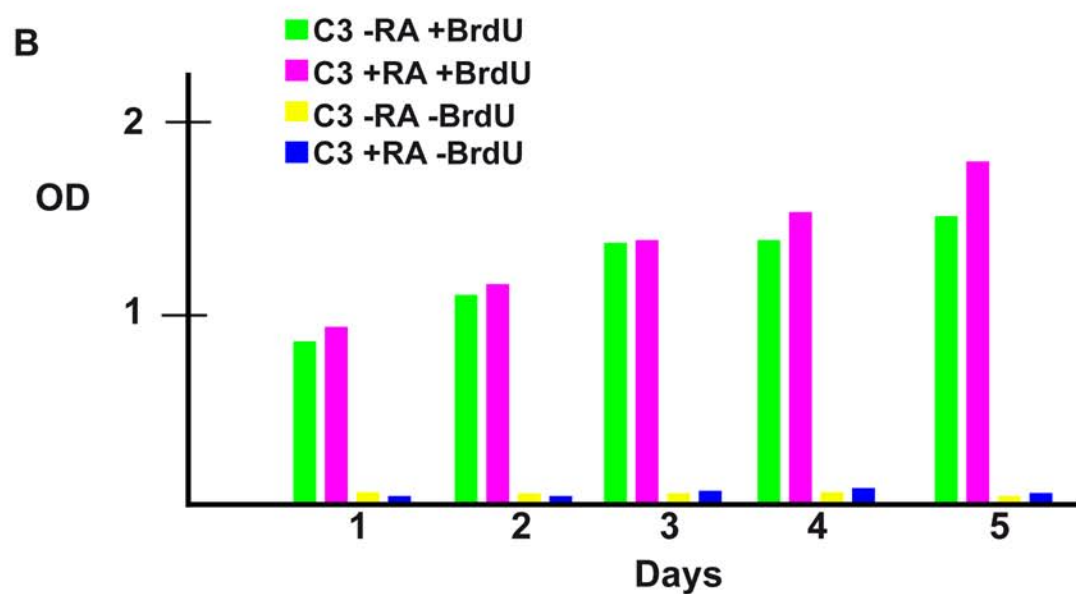
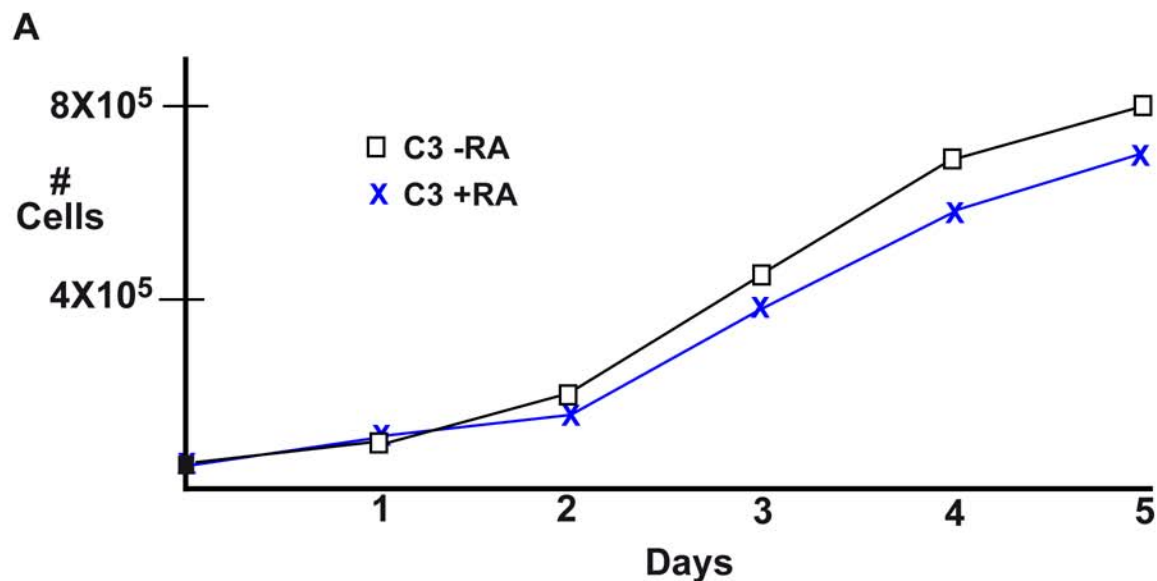


E



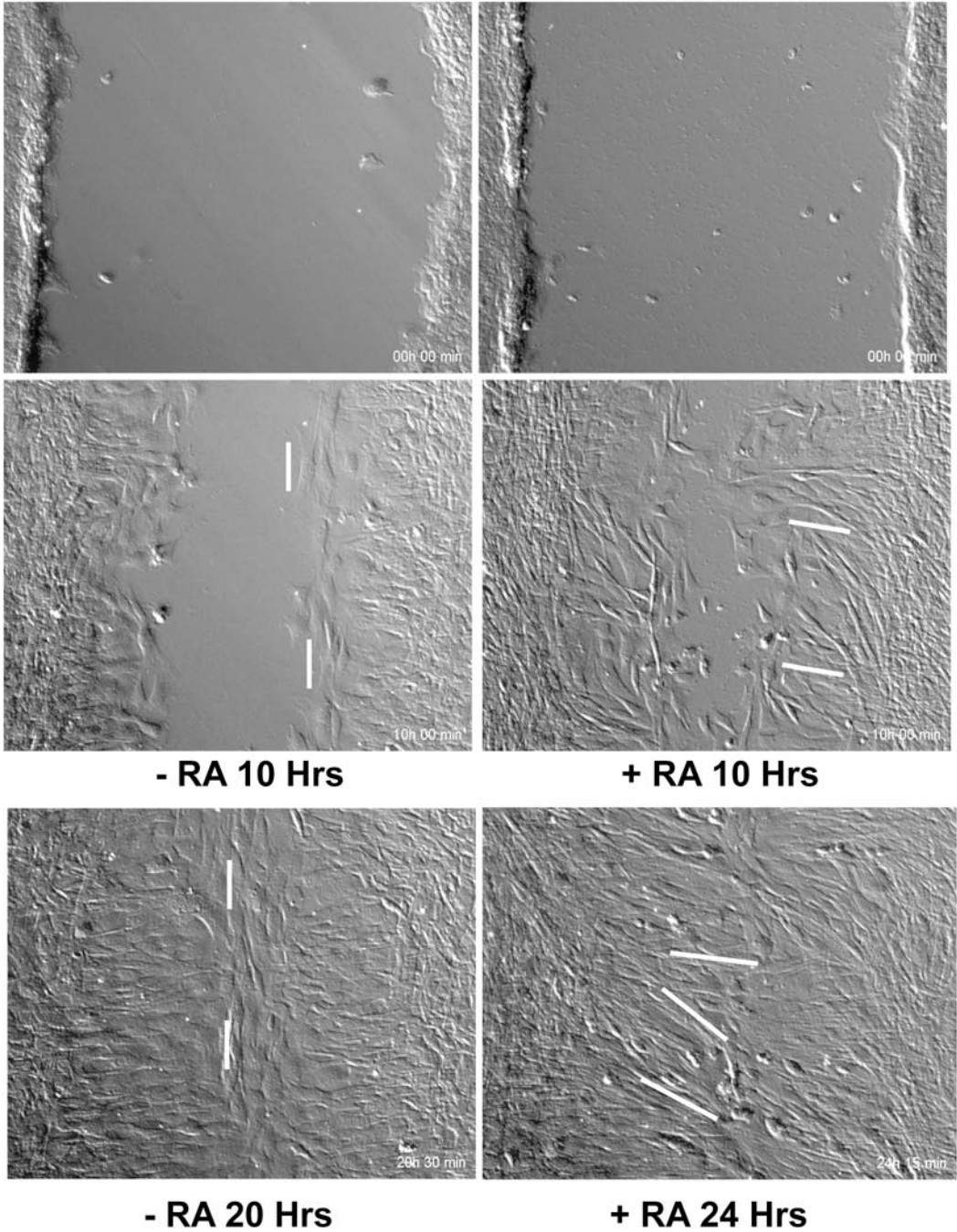
F

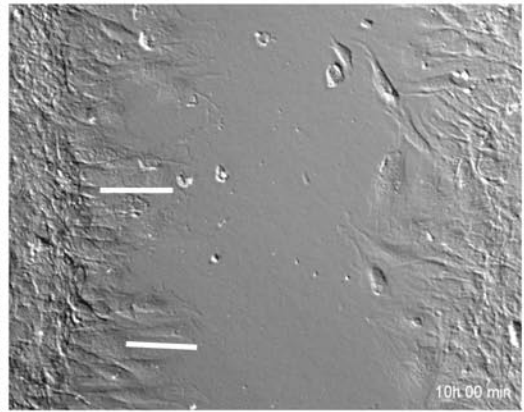
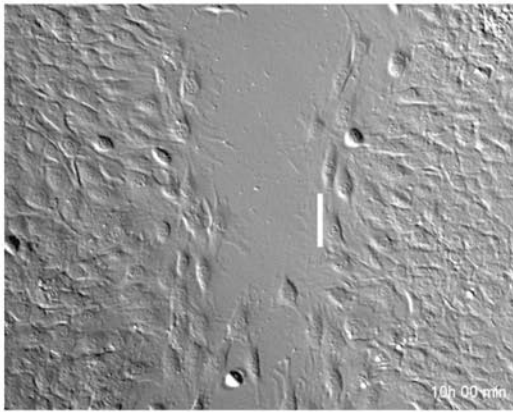
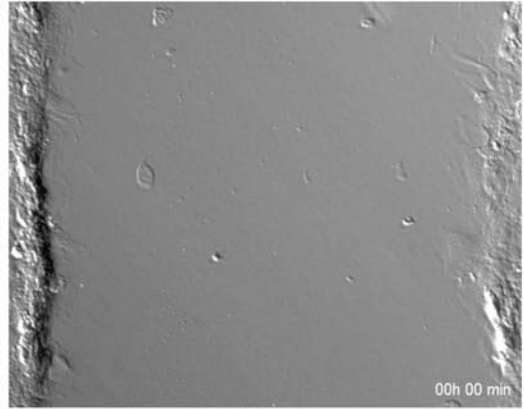
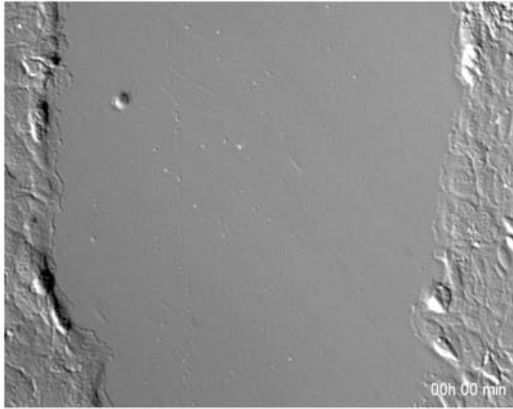
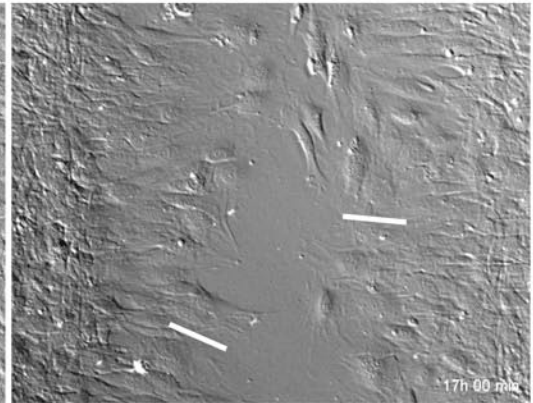
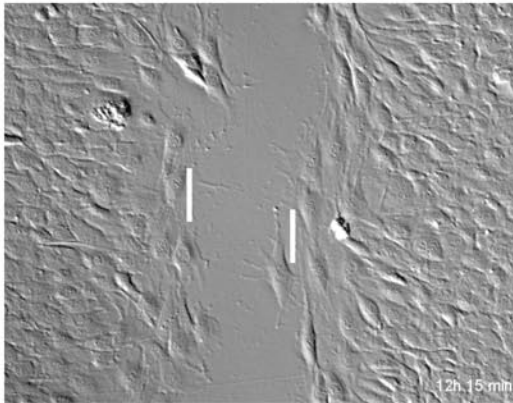


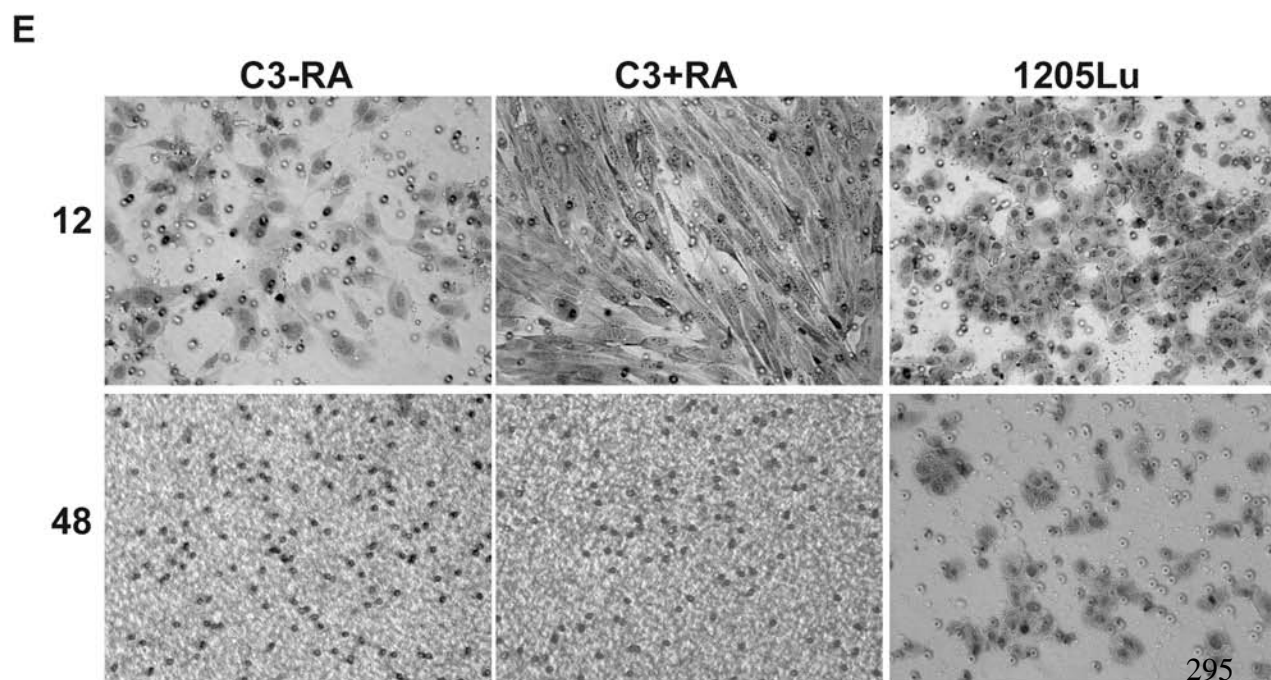
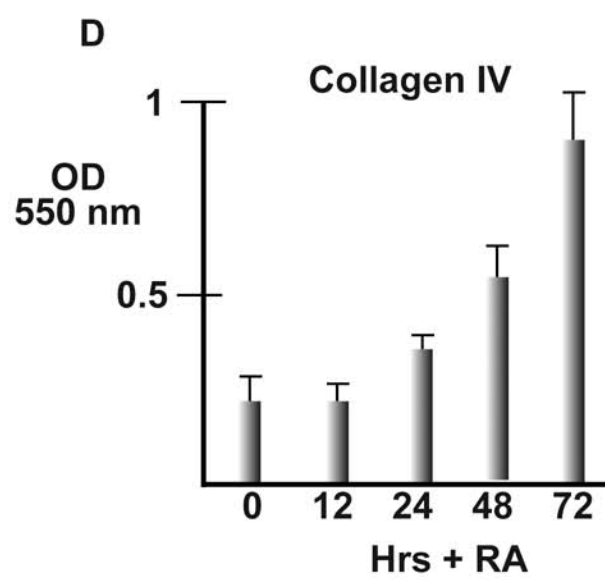
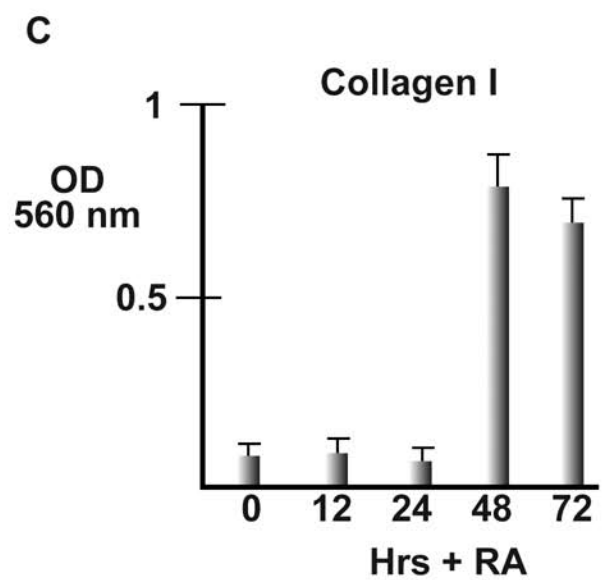
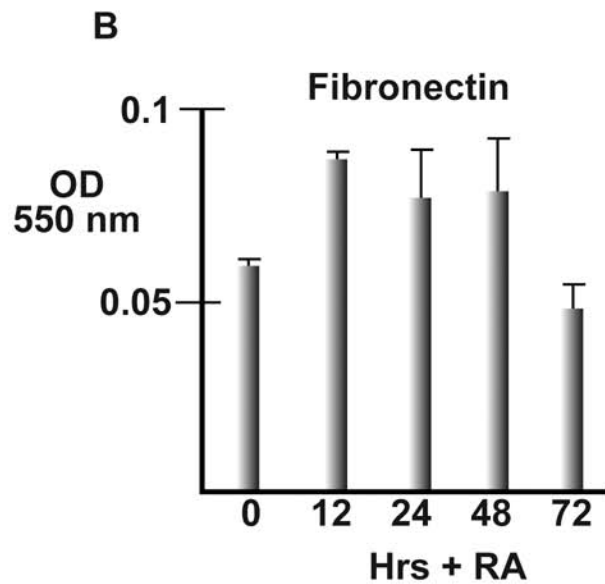
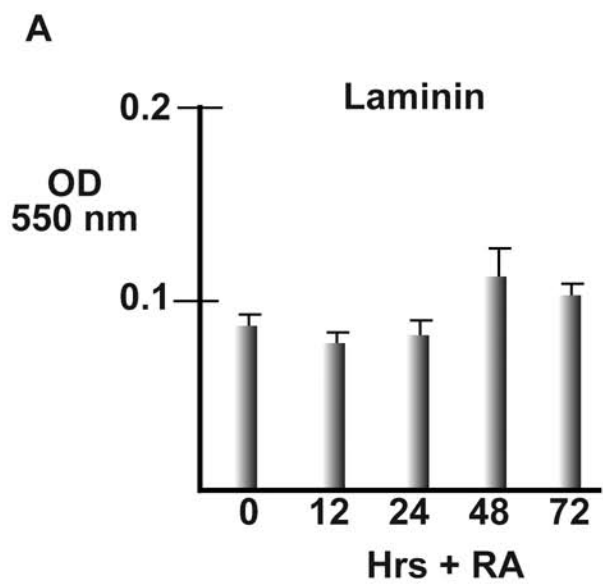


**C**

		G1	S	G2
-RA	24	46	26	26
	48	59	20	19
+RA	24	49	19	28
	48	56	23	19

**A****- RA 0 Hrs*****C3Taf4<sup>-/-</sup>*****+ RA 0 Hrs****B**

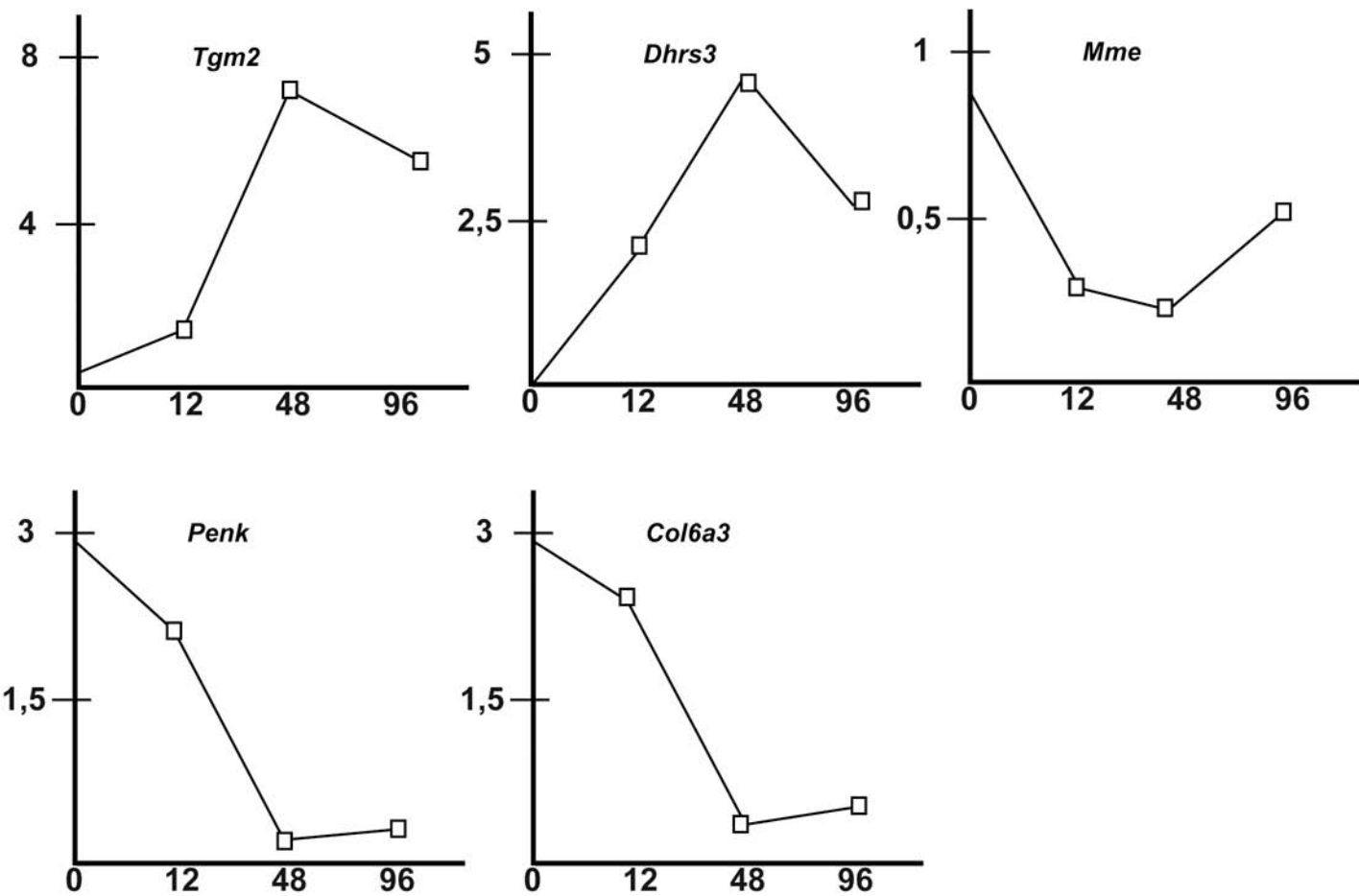
**A****- RA 0 Hrs*****C1Taf4<sup>lox/-</sup>* + RA 0 Hrs****B****- RA 10 Hrs****+ RA 10 Hrs****- RA 12 Hrs****+ RA 17 Hrs**



Cler et al., Fig. 6

Gene	Fold change 12 hours	Fold change 72 hours	Direct RAR target.
1. <i>Prss35</i>	12, 5	29,9	?
2. <i>Cyp26b1</i>	19,5	24,9	Yes
3. <i>Dhrs3</i>	8,3	16,0	Yes
4. <i>Tgm2</i>	7, 6	12,6	Yes
5. <i>Rarb</i>	7,8	11, 1	Yes
6. <i>Grpr</i>	4,6	10,5	?
9. <i>Gdpd5</i>	6,7	8,6	Yes
14. <i>Il1rn</i>	2,0	4,6	Yes
16. <i>Loxl2</i>	3,6	4,6	Yes
17. <i>F11r</i>	1,9	4,4	Yes
22. <i>Inmt</i>	5,7	4,1	Yes
48. <i>Bhlhb2</i>	2,0	3,0	Yes
12. <i>Prame 8</i>	0	4,8	?
13. <i>Prame 9</i>	0	4,7	?
205 <i>Anxa8</i>	-0,2	-3,7	?
206 <i>Cxcl7</i>	-0,5	-6,1	?
209 <i>Col6a3</i>	0	-6,3	?
210. <i>Penk</i>	-1,6	-11,7	?

**B**





<b>Cell line</b>	<b>RA</b>	<b>Autocrine Growth</b>	<b>Tumour formation</b>
<b>C1<i>taf4</i><sup>lox/-</sup></b>	<b>-</b>	<b>No</b>	<b>No</b>
<b>C1<i>taf4</i><sup>lox/-</sup></b>	<b>+</b>	<b>Yes</b>	<b>No</b>
<b>C3<i>taf4</i><sup>-/-</sup></b>	<b>-</b>	<b>Yes</b>	<b>Yes</b>
<b>C3<i>taf4</i><sup>-/-</sup></b>	<b>+</b>	<b>Yes</b>	<b>No</b>
<b>C3<i>taf4</i><sup>-/-</sup>/TAF4</b>	<b>-</b>	<b>No</b>	<b>ND</b>
<b>C3<i>taf4</i><sup>-/-</sup>/TAF4</b>	<b>+</b>	<b>Yes</b>	<b>ND</b>
<b>C3<i>taf4</i><sup>-/-</sup>/TAF4<i>b</i></b>	<b>-</b>	<b>Yes</b>	<b>Yes</b>
<b>C3<i>taf4</i><sup>-/-</sup>/TAF4<i>b</i></b>	<b>+</b>	<b>Yes</b>	<b>ND</b>

<b>Cell line</b>	<b>RA</b>	<b># Nuclei</b>	<b>Rate Migration pix/min</b>
<b>C1taf4<sup>lox/-</sup></b>	<b>-</b>	<b>144</b>	<b>0,35</b>
<b>C1taf4<sup>lox/-</sup></b>	<b>+</b>	<b>122</b>	<b>0,52</b>
<b>C3taf4<sup>-/-</sup></b>	<b>-</b>	<b>170</b>	<b>0,53</b>
<b>C3taf4<sup>-/-</sup></b>	<b>+</b>	<b>126</b>	<b>0,64</b>

*ANNEXE*

*PUBLICATION DEA*



## Publication de DEA

**Christian Gross, Gilles Buchwalter, Hélène Dubois-Pot, Emilie Cler, Hong Zheng, and Bohdan Wasylyk**

L'hypoxie régule des processus tels que l'angiogenèse ou la migration cellulaire en induisant des gènes impliqués dans ces processus biologiques comme PAI-1, *c-fos*, *erg-1*, NOS-2, HO-1 et VEGF. Le facteur Net du complexe ternaire (TCF) fait de même, ce qui suggère que Net pourrait être régulé par la voie de signalisation de l'hypoxie.

Dans cette étude, nous montrons que l'hypoxie régule Net différemment des autres TCFs et que Net joue un rôle dans la réponse hypoxique *in vitro* des cellules endothéliales et *in vivo* chez la souris. L'hypoxie induit la suppression du facteur Net du promoteur de certains gènes, l'export nucléaire, l'ubiquitylation et la dégradation protéosomale.

Les prolyl-4-hydroxylases (PHDs), des facteurs clés de la réponse hypoxique, régulent le facteur Net. La diminution des PHDs dans des conditions d'oxygène atmosphériques, induit la dégradation de Net, alors qu'en hypoxie, la surexpression des PHDs retarde la dégradation de Net. L'inhibition de Net par la méthode de l'ARN interférent ou via des mutations conduit en hypoxie à une dérégulation des gènes cibles de Net et notamment de PAI-1, *c-fos* et *egr-1*.

Nous suggérons que l'hypoxie stimule la transcription de promoteurs cibles par le déplacement de leur répresseur Net. De manière intéressante, la réponse de l'hématocrite chez la souris au chlorure de cobalt, un agent chimique induisant des réponses similaires à l'hypoxie, est fortement altérée dans les souris mutante pour Net. Nos résultats montrent donc que Net fait partie intégrante de la réponse biologique à l'hypoxie chez la souris.



# The Ternary Complex Factor Net Is Downregulated by Hypoxia and Regulates Hypoxia-Responsive Genes<sup>∇†</sup>

Christian Gross, Gilles Buchwalter,<sup>‡</sup> H el ene Dubois-Pot,<sup>‡</sup> Emilie Cler,  
Hong Zheng, and Bohdan Wasyluk\*

*Institut de G en etique et de Biologie Mol eculaire et Cellulaire, CNRS/INSERM/ULP, 1 rue Laurent Fries,  
BP 10142, 67404 Illkirch cedex, France*

Received 3 October 2006/Returned for modification 11 December 2006/Accepted 23 March 2007

**Hypoxia and the Net ternary complex factor (TCF) regulate similar processes (angiogenesis, wound healing, and cellular migration) and genes (PAI-1, *c-fos*, *erg-1*, NOS-2, HO-1, and vascular endothelial growth factor genes), suggesting that they are involved in related pathways. We show here that hypoxia regulates Net differently from the other TCFs and that Net plays a role in the hypoxic response in vivo in mice and in cells. Hypoxia induces Net depletion from target promoters, nuclear export, ubiquitylation, and proteasomal degradation. Key mediators of the hypoxic response, the prolyl-4-hydroxylases containing domain proteins (PHDs), regulate Net. PHD downregulation in normoxia leads to Net degradation, and PHD overexpression delays Net downregulation by hypoxia. Net inhibition by RNA interference or mutation leads to altered regulation by hypoxia of the Net targets PAI-1, *c-fos*, and *egr-1*. We propose that hypoxia stimulates transcription of target promoters through removal of the repressor function of Net. Interestingly, the hematocrit response to a chemical inducer of hypoxia-like responses (cobalt chloride) is strongly altered in Net mutant mice. Our results show that the Net TCF is part of the biological response to hypoxia, adding a new component to an important pathological and physiological process.**

Hypoxia is a pleiotropic condition encountered during embryonic development and organogenesis and in many pathological situations, like wound healing, inflammation, and tumorigenesis. In cancer, proliferating tumor cells outstrip the vascular supply, resulting in hypoxia, which induces adaptive changes (7, 22), including activation of the transcription factor hypoxia-inducible factor 1 $\alpha$  (HIF-1 $\alpha$ ) (20, 31) and angiogenesis (35). Inhibition of angiogenesis is a novel aspect of cancer therapy, but more precise manipulation of tumor vasculature may improve treatment strategies (27). Elucidation of the molecular mechanisms by which cells respond and adapt to hypoxia is of great interest.

We have identified a new component of the hypoxic response, the Net (Elk-3) transcription factor. Net, together with Elk-1 and Sap-1, forms the subfamily of ternary complex transcription factors (TCFs) that are notably known for their participation in the early response of quiescent cells to growth factor stimulation. Activation of the growth factor–Ras–mitogen-activated protein kinase (MAPK) pathway leads to phosphorylation and activation of Net and the other TCFs. Under basal conditions, Net is a strong transcriptional repressor (8, 15, 19, 30). Net negatively regulates a number of proteins, including *c-fos* (44) and *egr-1* (2, 44), which are involved in the immediate early response; PAI-1, a serpin that regulates matrix remodeling (9); CTP:phosphocholine cytidyltransferase  $\alpha$  (41), an enzyme involved in the phosphatidylcholine biosyn-

thesis; and nitric oxide synthase 2 (10) and heme-oxygenase 1 (13), which are both implicated in the inflammatory response. Net is involved in various physiological processes. Net regulates cell migration through repression of PAI-1 (9). Mice that express mutant Net lacking its DNA binding domain exhibit respiratory distress, chylothorax (the accumulation of chyle in the thoracic cage) at autopsy (2), and delayed wound healing due to impaired angiogenesis (51).

The pathways that regulate the repressor function of Net remain to be identified. Several observations suggested that hypoxia could be one of these regulators. The physiological responses mediated by hypoxia and Net, including cell migration, inflammation, wound healing, and angiogenesis, are similar. Moreover, all the identified target genes of Net are known to be regulated by hypoxia (3, 11, 12, 21, 28, 34, 38, 39, 48). In this study, we show that the transcription repressor function of Net is specifically downregulated in hypoxia. Downregulation is mediated by the oxygen sensors, the prolyl-4-hydroxylases containing domain proteins (PHDs) (5), and involves Net degradation at the protein level. Net is required for the induction of gene expression by hypoxia and for the physiological response to a hypoxia mimic in mice. This study adds an additional component to the understanding of the hypoxic response and a potential new dimension to the control of tumor vasculature, providing hope for better treatments.

## MATERIALS AND METHODS

**Cell culture, transfections, and hypoxia treatment.** The culture conditions were as follows: modified Eagle's medium (MEM), 10% fetal calf serum (FCS), and 40  $\mu$ g/ml gentamicin for mouse skin endothelial (SEND) cells (a generous gift from Kari Alitalo); Dulbecco's MEM (DMEM)-Ham's F-12K medium (1:1), 10% FCS, and 40  $\mu$ g/ml gentamicin for PC-3 cells; MEM (Eagle), 10% FCS, 1 mM sodium pyruvate, 0.1 mM nonessential amino acids, and 40  $\mu$ g/ml gentamicin for DU 145 cells; and DMEM, 10% FCS, and 40  $\mu$ g/ml gentamicin for mouse embryonic fibroblasts, C11 cells, 444 cells, and COS-7 cells. COS-7 cells were

\* Corresponding author. Mailing address: Institut de G en etique et de Biologie Mol eculaire et Cellulaire, CNRS/INSERM/ULP, 1 rue Laurent Fries, BP 10142, Illkirch cedex 67404, France. Phone: 33 3 88 65 34 11. Fax: 33 3 88 65 32 01. E-mail: boh@igbmc.u-strasbg.fr.

<sup>†</sup> Supplemental material for this article may be found at <http://mcb.asm.org/>.

<sup>‡</sup> These authors contributed equally to this work.

<sup>∇</sup> Published ahead of print on 2 April 2007.

transfected with Jet-Pei (Polyplus Transfection) according to the manufacturer's instructions. 444 and SEND cells were transfected with Lipofectamine (Invitrogen) as previously described (9, 44). The hypoxia conditions were 1% O<sub>2</sub> in a ThermoForma model 3110 incubator.

**Immunoblotting and antibodies.** For immunoblotting, see reference 9. The antibody uses and dilutions were as follows: rabbit anti-mouse Net no. 1996 for immunoprecipitation (IP; 7 μl per IP) and Western blotting (WB; 1/2,000), mouse anti-mouse phospho(Ser364)-Net no. 2F3 for WB (1/2,000), rabbit anti-mouse Elk-1 no. 512 for WB (1/2,000), rabbit anti-mouse Sap-1a no. 644 for WB (1/2,000), mouse anti-mouse TATA binding protein (TBP; a gift from L. Tora) for WB (1/5,000), mouse anti-α-tubulin clone DM 1A for WB (1/10,000) (Sigma), mouse anti-mouse HIF-1α for WB (1/1,000) (Novus Biologicals), mouse anti-mouse hemagglutinin (HA; IGBMC core facilities) for WB (1/1,000), mouse anti-mouse ubiquitin (Ub) for WB (1/1,000) (Santa Cruz Biotechnologies), rabbit anti-mouse phospho-Erk1/Erk2 for WB (1/1,000) (P-ERK; Cell Signaling Technology), rabbit anti-p44/42 for WB (1/1,000) (Cell Signaling Technology), and rabbit anti-green fluorescent protein (anti-GFP) for WB (1/3,000) (Torrey Pines Biolabs, Inc.).

**Chemicals.** The chemicals were as follows: proteasome inhibitor MG132 (Calbiochem), cobalt chloride (CoCl<sub>2</sub>; Alfa Aesar), MEK inhibitor U0126 (Promega), dimethylxalylglycine (DMOG; Frontier Scientific), desferrioxamine (DFX) methanesulfonate (Sigma); ethyl 3,4-dihydroxybenzoate (Sigma), and cycloheximide (CHX; Calbiochem).

**Plasmids.** The plasmids were as follows: pTL2-Net, HA-tagged Ub (Ub-HA), pGFP-Net, PAI-1-luciferase reporters (9), pCMV-LacZ (IGBMC core facilities), and pCDNA3-FLAG-PHD1 (a generous gift from Frank S. Lee, University of Pennsylvania School of Medicine, Philadelphia, PA) (24).

**Immunofluorescence.** SEND cells were grown on coverslips, subjected to hypoxia for 8 h, fixed with 3.7% formaldehyde for 10 min at room temperature, permeabilized with 0.1% Triton X-100 for 5 min, and blocked with 3% bovine serum albumin for 40 min. The cells were incubated with a rabbit anti-mouse Net no. 1996 (1/500) antibody for 2 h at 37°C and incubated for 1 h at 37°C with Texas Red-conjugated anti-rabbit antibody (Beckman Coulter). Finally, the cells were stained for 1 min at room temperature with Hoechst (Sigma), covered with coverslips using mounting solution (5% propylgallate in glycerol), and observed with a fluorescence microscope (Leica; magnification, ×40).

**RNA interference (RNAi).** The small interfering RNAs (siRNAs) were as follows: mouse Net siRNA (Dharmacon, Inc.) (9); Human PHD1, PHD2, or PHD3 siRNA (siGENOME SMARTpool reagent; Dharmacon, Inc.); and GL2 luciferase control siRNA (9). SEND cells were transfected with siRNA duplexes (final concentration, 20 nM) by using Lipofectamine (Invitrogen, Carlsbad, CA) or Jet Pei (Polyplus Transfection) according to the manufacturer's guidelines.

**RT-PCR.** Total RNA was prepared using TRIzol reagent (Invitrogen). Reverse transcription (RT) was performed using a SuperScript II kit according to instructions from Invitrogen. The PCR steps were as follows: 5 min at 95°C and 30 cycles of 45 s at 95°C, 45 s at 60°C, 1 min at 72°C, and 5 min at 72°C. The PCR products were analyzed on a 2% agarose gel. The oligonucleotides were as follows: PHD1, 5'-CTGGGCAGCTATGTCATCAA and 5'-AAATGAGCAAC CCGTCAAAG; PHD2, 5'-GAAAGCCATGGTTGCTTGT and 5'-TTGGGT TCAATGTCAGCAA; PHD3, 5'-AGATCGTAGGAACCCACACG and 5'-C AGATTCAGAGCACGGTCA; 28S, see below.

**Q-RT-PCR.** Total RNA was prepared using TRIzol reagent (Invitrogen). Quantitative RT-PCR (Q-RT-PCR) was performed with the LightCycler system (Roche Diagnostics) and the SYBR green I (Roche Diagnostics) protocol. The reaction mixtures, containing 100 ng of RNA and 1× master mixture (0.5 μM primers, 4 mM MgCl<sub>2</sub>, 200 μM deoxynucleoside triphosphates, 1× PCR buffer [Sigma], 1 U/μl Superscript reverse transcriptase [Invitrogen], *Taq* polymerase [Promega], anti-*Taq* antibody diluted 1/200 (Taqstart antibody; Clontech), 4.3% glycerol, 0.15 mg/ml bovine serum albumin, and 0.25× SYBR green I), were reverse transcribed for 10 min at 55°C, denatured for 30 s at 95°C, and cycled 40 times, each time for 2 s at 95°C, 10 s at 61°C, and 10 s at 72°C. Amplification specificity was verified by melting-curve analysis, and the data were quantified with LightCycler software. The absence of genomic DNA contamination was verified by repeating the procedure with the same samples without reverse transcriptase. The oligonucleotides were as follows: Net, first couple, 5'-GGCC GAACACTTTTCCAG and 5'-GATTTCTGAGAGCTGGGGGA; Net, second couple, 5'-ACTAGCCCTGCTCTCTCAT and 5'-GTTTGTTCCTCCAC CACGG; PAI-1, 5'-CTCCGAGAATCCACACAG and 5'-ACTTTGAATC CCATAGCATC; *c-fos*, 5'-AAGGGAACGGAATAAGATGGC and 5'-CAAC GCAGACTTCTCATCTCAA; 28S, 5'-GGCGGCAAGCGTTTCATAGC and 5'-GCCAAGCACATACACCAAAT; *egr-1*, 5'-GCCGAGCGAACAACCTA and 5'-TCCACCATCGCCTTCTCATT; VEGF, 5'-GGGGGTACCTGGACCT

GGTGTCTGCGT and 5'-CCACCATTTTGACAAACAGCACAATCACAC CTTGCACGAAGAAG.

**ChIP.** Chromatin IP (ChIP) assays were performed as previously described (9) using a kit from Upstate Biotechnology. IPs were done with 10 μl of rabbit anti-mouse Net no. 1996 antibody or 10 μl of rabbit anti-clathrin heavy-chain H-300 (Santa Cruz Biotechnology). The coimmunoprecipitated DNA was quantified by semiquantitative PCR and quantitative PCR (denatured for 5 min at 95°C and cycled 35 times, each time for 5 s at 95°C, 30 s at 61°C, and 30 s at 72°C). Amplification specificity was verified by melting-curve analysis, and the data were quantified with LightCycler software.

**Luciferase assays.** SEND cells were transfected in triplicate with Jet Pei (Polyplus Transfection) according to the manufacturer's instructions in six-well clusters with 4 μg of DNA per well containing 1 μg of the PAI-1 reporters, 1 μg pCMV-LacZ, and 2 μg of pBSK. Forty-eight hours after transfection, the cells were subjected to hypoxia. Luciferase activities were measured with the Promega luciferase assay system (according to the manufacturer's protocol) and an EG&G Berthold luminometer and corrected for transfection efficiency by using β-galactosidase activity as an internal control.

**Co-IP.** SEND cells treated with MG132 and subjected to hypoxia or transfected COS-7 cells were washed with phosphate-buffered saline and lysed in radioimmunoprecipitation assay buffer containing 50 mM Tris-HCl (pH 7.5), 0.15 M NaCl, 1% Nonidet P-40, 0.1% deoxycholate, 0.1% sodium dodecyl sulfate, and a protease inhibitor cocktail (Roche Diagnostics). Immune complexes were formed by incubating the lysates overnight at 4°C with rabbit anti-mouse Net no. 1996, mixed for 1 h at 4°C with protein A/G Sepharose, washed three times with radioimmunoprecipitation buffer, resuspended in Laemmli buffer, boiled, separated by sodium dodecyl sulfate-10% polyacrylamide gel electrophoresis, and transferred electrophoretically to nitrocellulose membranes. The blots were blocked with 5% nonfat dry milk, incubated with horseradish peroxidase-conjugated secondary antibodies at room temperature for 1 h, visualized with the Pierce enhanced chemiluminescence detection system, and exposed to film.

**Animal studies.** Age- and sex-matched, 17- to 30-week-old, 129Sv strain wild-type and Net mutant mice were used. Fourteen mice (7 per genotype) were injected intraperitoneally three times per week with 30 mg/kg CoCl<sub>2</sub>, and blood samples were collected once a week during 4 weeks. Significance was evaluated by the Student *t* test.

**Hematology and plasma PAI-1.** Blood samples were obtained by retro-orbital sinus puncture. The hematological parameters were as follows. Thirty-microliter blood samples were collected in EDTA tripotassium salt microcuvettes (Sarstedt) and analyzed in a blood cell counter (AcT Diff Vet; Beckman Coulter). For PAI-1, 9 volumes of blood was collected in 1 volume of 0.1 M trisodium citrate and immediately centrifuged at 3,000 × *g* for 15 min. The plasma was transferred to a clean plastic tube, and total PAI-1 concentrations were determined by an enzyme-linked immunosorbent assay (MPAIKT-TOT; Innovative Research Inc.).

## RESULTS

### Hypoxia specifically downregulates Net at the protein level.

We studied Net regulation by hypoxia in endothelial cells, one of the cell types involved in the hypoxic response. Net protein levels decreased in SEND cells subjected to atmospheric hypoxia (1% O<sub>2</sub> in a hypoxic cell culture incubator), leading to complete loss of Net after 18 h (Fig. 1A). As expected, HIF-1α protein levels increased under the same conditions, peaking after 6 h of hypoxia. The levels of the two other TCFs, Elk-1 and Sap-1a, as well as the internal control TBP remained stable under the same conditions, showing that downregulation is specific for Net. Net degradation in response to hypoxia was not restricted to endothelial cells, since it was also observed in prostate cancer cells (PC3 and DU145), murine embryonic fibroblasts, and immortalized fibroblasts (C11) (Fig. 1B).

In order to determine whether hypoxia affects Net expression at the RNA level, Net mRNA from hypoxia-treated SEND cells was quantitated by real-time Q-RT-PCR. There were relatively small variations in Net mRNA levels (Fig. 2A) (similar results were obtained with a second primer pair [data



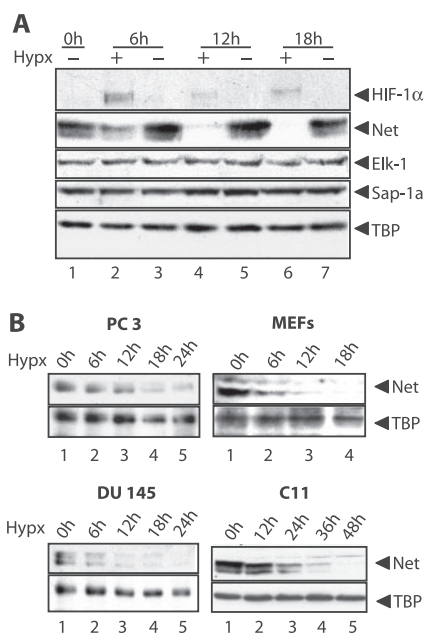


FIG. 1. Specific downregulation of Net protein levels under hypoxic conditions. SEND (A), PC3, murine embryonic fibroblast (MEF), DU145, and C11 (B) cells were placed under hypoxic (Hypx) atmospheric conditions (1% O<sub>2</sub>), and the levels of HIF-1α, Net, Elk-1, Sap-1a, and TBP were determined by WB analysis, as indicated.

not shown]), which could not account for the dramatic decrease at the protein level. Under the same conditions, the RNA levels for several other genes (*c-fos*, *PAI-1*, vascular endothelial growth factor [VEGF], and *egr-1* genes) increased (see below). To confirm that Net is regulated at the protein level, we studied the effect of hypoxia on the expression of Net from a heterologous promoter (simian virus 40 early region in pSG5) in the presence of an inhibitor of protein synthesis (CHX). The expression of the fusion protein GFP-Net in hypoxia was reduced compared to that in normoxia and compared to the expression of GFP alone (Fig. 2B and C). These results show that hypoxia downregulates Net at the protein level.

**PHDs are involved in Net degradation.** PHDs are cellular sensors of oxygen that require oxygen, iron, ascorbate, and 2-oxoglutarate to hydroxylate HIF-1α in normoxia (5). PHDs are inhibited by iron chelators, such as DFX; iron-displacing transition metals, such as cobalt; or synthetic 2-oxoglutarate antagonists, such as DMOG. In cobalt chloride (CoCl<sub>2</sub>)-treated cells, Net protein levels decreased in both a dose (Fig. 3A, panel a, 8 h of treatment)- and time (Fig. 3A, panel b, 250 μM of CoCl<sub>2</sub>)-dependent manner, whereas HIF-1α levels increased (Fig. 3A, panel b). Other PHD inhibitors (DMOG, DFX, and ethyldihydroxybenzoate) also decreased Net and increased HIF-1α protein levels (see Fig. S1 in the supplemental material). However, these chemicals are not completely specific. To confirm that the PHDs have a role in Net regulation, we used siRNAs against the three isoforms of the human PHDs. In 444 cells (described previously [44]), PHDs 1, 2, and 3 were efficiently and specifically downregulated, as shown by RT-PCR (Fig. 3B, panel a). Downregulation of PHD2 increased HIF-1α levels, whereas downregulation of the two other isoforms had only slight effects (Fig. 3B, panel b), in

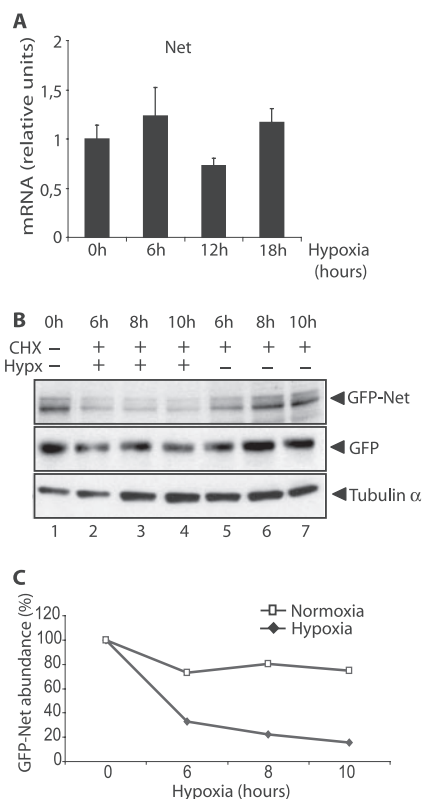


FIG. 2. Net protein stability is decreased in hypoxia (Hypx). (A) mRNA from cells incubated in 1% O<sub>2</sub> for 0 to 18 h was measured by Q-RT-PCR and normalized to 28S rRNA levels. (B, C) SEND cells were transfected with GFP or GFP-Net expression vectors, treated with CHX (50 μM) for 0 to 10 h as indicated, analyzed by WB for GFP and GFP-Net levels with a GFP antibody (B), and quantified with ChemiGenius XE (Syngene) (C).

agreement with another study (4). In contrast, downregulation of PHD2 as well as PHD3 had little effect on Net, whereas inhibition of PHD1 decreased Net expression (Fig. 3B, panel b). Overexpression of PHD1, which can counterbalance diminution of activity due to the lack of oxygen in hypoxia (32), delayed Net hypoxic downregulation (Fig. 3C, compare lanes 2, 3, and 4 with lanes 8, 9, and 10). These results implicate PHDs, and in particular PHD1, in Net regulation in hypoxia.

**Hypoxia induces Net nuclear export.** Many nuclear proteins that are degraded by the proteasome, such as p53, are concomitantly exported from the nucleus (50). Since Net has a nuclear export sequence (16), we studied the effects of hypoxia on its cellular localization. In SEND cells transfected with a GFP-Net expression vector, GFP-Net was nuclear in normoxia and nucleocytoplasmic after 8 h of hypoxia (Fig. 4A). Similarly, endogenous Net, detected by immunocytochemistry with a Net antibody in nontransfected cells, was nuclear in normoxia and nucleocytoplasmic after 8 h of hypoxia (Fig. 4B). We also studied whether nuclear export and the JNK pathway are involved in hypoxia-induced Net degradation. Inhibition of nuclear export with leptomycin B (LMB) delayed hypoxia-induced Net degradation (Fig. 4C, compare lane 5 with lane 4 and lane 7 with lane 6), suggesting that nuclear export is involved, at least in part, in Net degradation. In contrast, inhi-

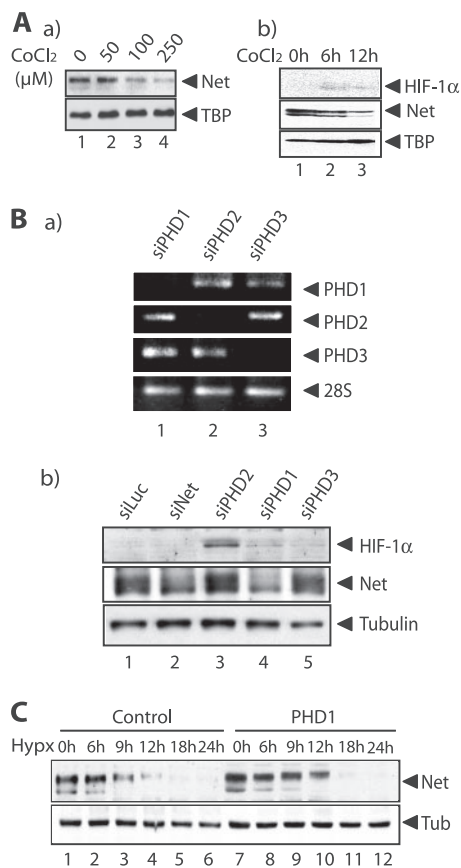


FIG. 3. PHDs are involved in Net regulation. (A) SEND cells were treated with cobalt chloride (0 to 250  $\mu$ M) for 8 h (a) and 250  $\mu$ M cobalt chloride (b) for 0 to 12 h, and extracts were analyzed by WB. (B) 444 cells were transfected with anti-PHD1, -PHD2, or -PHD3 siRNAs; PHD RNA levels were estimated by RT-PCR (a); and HIF-1 $\alpha$ , Net, and  $\alpha$ -tubulin protein levels were analyzed by WB (b). (C) SEND cells were transfected with an expression vector for PHD1 and subjected to hypoxia (Hypx) (1% O<sub>2</sub>) for different times. The protein levels of Net and  $\alpha$ -tubulin (Tub) were analyzed by WB.

bition of the JNK pathway with SP600125 did not affect Net levels (Fig. 4D), indicating that JNK signaling is not involved in hypoxia-induced Net degradation.

**Hypoxia induces ubiquitylation and proteasomal degradation of Net.** We used different protease inhibitors to elucidate the pathway of Net degradation. Whereas lysosomal and calpain inhibitors had no effect (data not shown), the proteasome inhibitor MG132 prevented Net protein degradation in hypoxia (Fig. 5A). The increase in Net levels (Fig. 5A, compare lanes 1 and 3) suggests that MG132 also inhibits Net turnover under normal conditions. Since proteasomes degrade ubiquitylated proteins (23), we studied whether Net can be ubiquitylated. COS-7 cells were transfected with Net and Ub-HA expression vectors and treated with MG132 as indicated, and extracts were analyzed by WB (Fig. 5B, panel a). Coexpression of Net with Ub-HA led to the appearance of a series of bands of increased molecular weights (Fig. 5B, panel a, lane 3) that was enhanced by the proteasome inhibitor (Fig. 5B, panel a, lane 4), consistent with them being ubiquitylated forms of Net. In order to confirm this hypothesis, Net immunoprecipitates

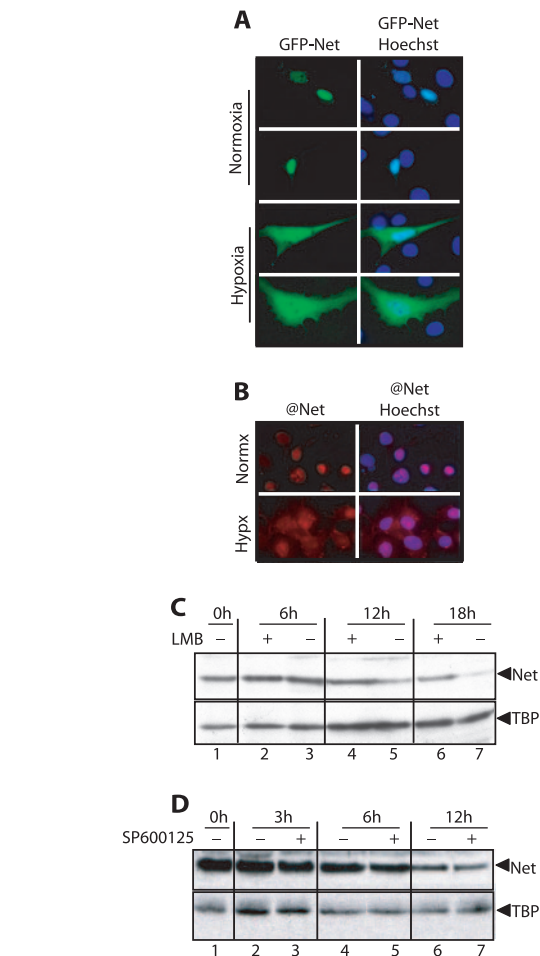


FIG. 4. Nuclear export contributes to hypoxia downregulation of Net in SEND cells. (A) Cells were transfected with a GFP-Net expression vector, incubated 24 h later in normoxia or hypoxia (1% O<sub>2</sub>) for 8 h, stained with Hoechst, and visualized by fluorescence microscopy (magnification,  $\times 40$ ). (B) Cells were incubated in normoxia (Normx) or hypoxia (Hypx) (1% O<sub>2</sub>) for 8 h, analyzed by immunocytochemistry with the Net antibody, stained with Hoechst, and visualized by fluorescence microscopy (magnification,  $\times 40$ ). (C, D) Cells were incubated in hypoxia (1% O<sub>2</sub>) for 0 to 18 h and treated with inhibitors of nuclear export (2 ng/ $\mu$ l LMB) or JNK1-3 (10  $\mu$ M SP600125), as indicated, and analyzed by WB with Net and TBP antibodies.

were analyzed with HA antibodies by WB. The higher-molecular-weight forms were detected only with COS-7 cells expressing both Net and Ub-HA (Fig. 5B, panel b, compare lane 3 with lanes 1 and 2), suggesting that Net can be ubiquitylated *in vivo*. To determine whether endogenous Net ubiquitylation is enhanced in hypoxia, we submitted SEND cells to 12 h of either hypoxia or normoxia and analyzed Net immunoprecipitates by WB with an Ub antibody (Fig. 5C). Ubiquitylated higher-molecular-weight forms of Net were detected in Net immunoprecipitates after 12 h of treatment with MG132 (Fig. 5C, panel b, lanes 2 and 3), and the levels of ubiquitylated Net were higher in hypoxia than in normoxia (Fig. 5C, panel b, lanes 2 and 3). This confirms that hypoxia stimulates the ubiquitylation of Net, which presumably accelerates proteasomal degradation. Since many proteins are phosphorylated prior to ubiquitylation (18), we tested whether Net phosphorylation by

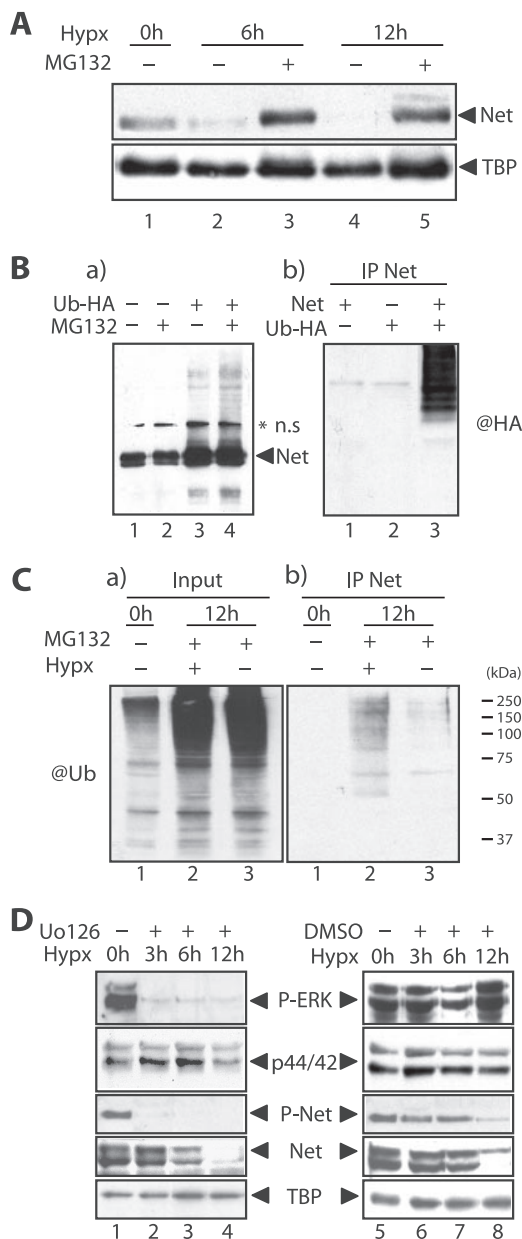


FIG. 5. Hypoxia-induced Net ubiquitylation enhances Net proteasomal degradation. (A) SEND cells were incubated under hypoxic (Hypx) conditions (1% O<sub>2</sub>) with or without 35 μM MG132 for 0 to 12 h. (B) COS-7 cells were transfected with Net and Ub-HA expression vectors and treated with 35 μM MG132 for 6 h, as indicated. The extracts were analyzed directly (a) or after IP with Net antibodies (b) by WB. (C) SEND cells were incubated for 12 h in hypoxia (1% O<sub>2</sub>) and treated with 35 μM MG132, and extracts were analyzed directly (a) or after IP with a Net antibody (b) by WB. (D) SEND cells were treated with hypoxia (1% O<sub>2</sub>) and U0126 for 0 to 12 h and analyzed by WB.

the extracellular signal-regulated kinase (ERK) pathway is required for degradation. SEND cells were incubated with the potent MEK inhibitor U0126 in hypoxia (Fig. 5D). U0126 decreased ERK phosphorylation but did not prevent Net degradation, showing that ERK phosphorylation, under these conditions, is not necessary for the decrease of Net levels in response to hypoxia.

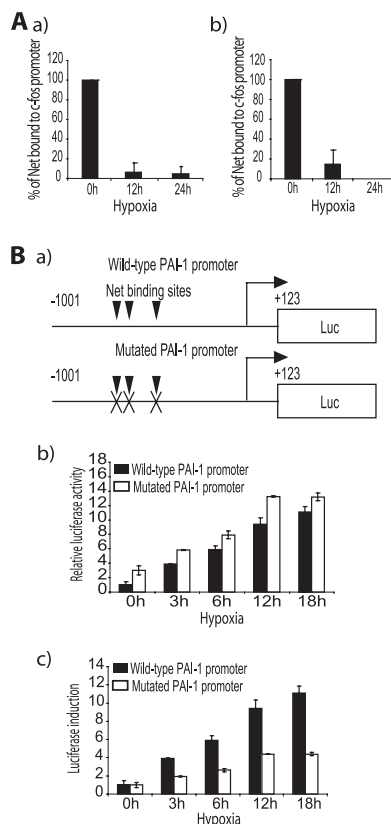


FIG. 6. Effect of hypoxia on binding of Net to the *c-fos* and PAI-1 promoters and regulation of PAI-1 promoter activity. (A) SEND cells incubated in hypoxia (1% O<sub>2</sub>) for the indicated times were analyzed by ChIP with Net and clathrin (control) antibodies. Input and coprecipitated (IP) DNA samples were analyzed by quantitative PCR, using primers for regions of the *c-fos* (–338 and –153) (a) and PAI-1 (–715 and –367) (b) promoters that contain Ets binding sites. The values were normalized to those for the inputs. (B) SEND cells were transfected with PAI-1 promoter luciferase reporters containing wild-type or mutated Net binding sites, and 48 h after transfection, the cells were incubated in hypoxia (1% O<sub>2</sub>) for 0 to 24 h. Luciferase activity was corrected for β-galactosidase activity expressed from the pCMV-LacZ internal control.

**Hypoxia induces loss of Net from target gene promoters.** In order to determine whether hypoxia decreases the amounts of Net on the *c-fos* (44) and PAI-1 (9) promoters, which are known to bind Net in normoxia, we performed ChIP assays with SEND cells by using antibodies against Net as well as against clathrin, which does not bind to DNA (as a negative control). The products were analyzed by both semiquantitative (data not shown) and quantitative PCR. As expected, in normoxia, Net was detected on the serum response element-containing region (–338 to –153) of the *c-fos* promoter (Fig. 6A, panel a) and the proximal region (–715 to –367) of the PAI-1 promoter (Fig. 6A, panel b). The negative control did not give specific products (data not shown). After exposure to hypoxia, binding of Net was lost from both promoters (Fig. 6A, panels a and b). The loss of Net as a transcriptional repressor could account for the hypoxic induction of these genes.

**Net is required for hypoxia induction of PAI-1 promoter activity.** To study Net regulation of PAI-1 promoter activity in

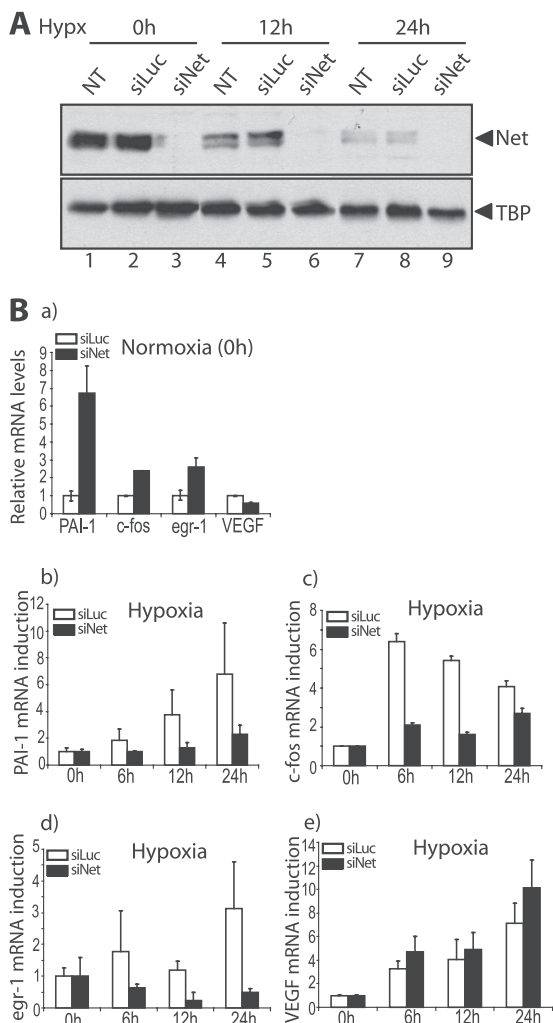


FIG. 7. Net is required for hypoxic induction of target genes. SEND cells were transfected with siRNAs against Net and luciferase (negative control), and 48 h later, the cells were incubated in hypoxia (1% O<sub>2</sub>) for 0 to 24 h. Protein extracts were analyzed by WB (A) and mRNA levels by Q-RT-PCR, and the levels were normalized to 28S rRNA levels (B). (a) The mRNA levels at 0 h in normoxia for cells transfected with siRNA Net were normalized to the control siLuc values. (b to e) The mRNA levels at various times in hypoxia were normalized to the levels observed at 0 h before induction. NT, not transfected.

response to hypoxia, we used a luciferase reporter with mutations that inhibit Net binding (9) (Fig. 6B, panel a). SEND cells were transfected with either wild-type or mutated reporters, and after 48 h, the cells were incubated in hypoxia for up to 24 h. In normoxia, the activity of the mutated PAI-1 promoter was three times higher than that of the wild type, as expected for loss of transcriptional repression by Net (Fig. 6B, panel b, 0 h). However, after 18 h of hypoxia, both promoters exhibited similar activities, as expected for hypoxia-induced relief of repression by Net, but only for the wild-type promoter. It appears that hypoxic induction of the wild-type promoter is due to relief of repression by Net, among other factors, whereas the mutated promoter is induced only by the other factors, explaining the similar activities of the two promoters after 18 h of

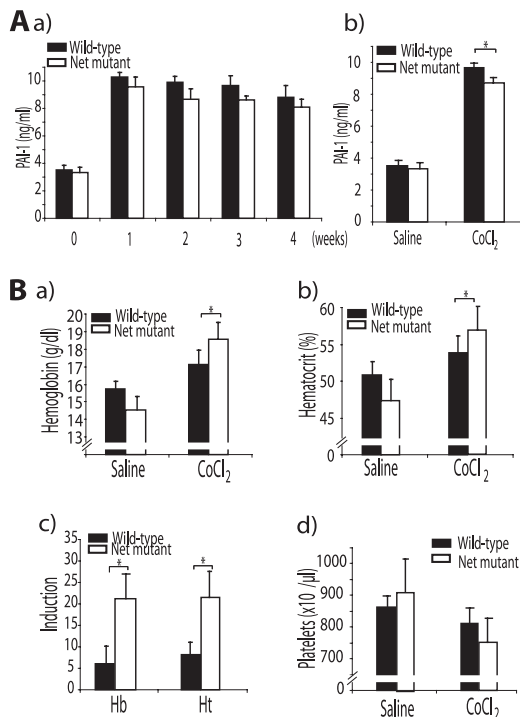
hypoxia. When luciferase activities were expressed relative to normoxic levels (0 h, induction), wild-type promoter activity increased about 11-fold after 18 h of hypoxia, whereas mutant promoter activity increased only about 4-fold (Fig. 6B, panel c). These results show that the Net binding sites are required in part for the induction of PAI-1 promoter activity by hypoxia and suggest that the loss of Net repression participates in this response.

**Net is required for the hypoxic induction of target genes.** To examine whether Net is required for the induction of PAI-1, *c-fos*, *egr-1*, and VEGF expression, we decreased Net levels with siRNA before incubation in hypoxia. Anti-Net siRNA abrogated Net protein expression (Fig. 7A, lane 3) in comparison with control siRNA (Fig. 7A, lane 2) or nontransfected (Fig. 7A, lane 1) cells. Hypoxia induced Net downregulation to similar extents in the nontransfected and control-transfected cells (Fig. 7A, lanes 4, 5, 7, and 8). mRNA levels were measured by Q-RT-PCR (Fig. 7B). In normoxia, at 0 h, Net is a transcriptional repressor of PAI-1, *c-fos*, and *egr-1* and an activator of VEGF (Fig. 7B, panel a), as already shown by previous studies (2, 9, 44, 51). Hypoxia induced a gradual increase in PAI-1 expression that was reduced by prior downregulation of Net, from about six- to twofold, after 24 h (Fig. 7B, panel b). Hypoxia induced a more rapid increase in *c-fos* expression, which was about sixfold at the peak (6 h) and then gradually decreased. The increase was less pronounced and only about threefold in cells lacking Net (Fig. 7B, panel c). The apparently more complex effect of hypoxia on *Egr-1* expression was strongly inhibited by downregulation of Net (Fig. 7B, panel d). Finally, the induction of VEGF was not significantly affected by prior downregulation of Net (Fig. 7B, panel e). These results show that Net contributes to hypoxia induction of PAI-1, *c-fos*, and *egr-1* expression and is not required for induction of VEGF.

**Net plays a role in the response to a hypoxia mimic in mice.** We studied the effect of Net mutation in mice on PAI-1 expression induced by cobalt chloride, which is often used as a hypoxia mimic in small rodents (37, 47). CoCl<sub>2</sub> (30 mg/kg) enhanced PAI-1 levels in plasma during 4 weeks (Fig. 8A, panel a). Overall, averaging all the data revealed a slight but significant decrease in PAI-1 induction in Net mutant mice (Fig. 8A, panel b). We also analyzed several hematological parameters that are known to be induced by hypoxia. As expected, hemoglobin and hematocrit were induced by CoCl<sub>2</sub> (Fig. 8B, panels a and b). Interestingly, there were marked differences ( $P < 0.05$ ) between wild-type and Net mutant mice in levels of both hemoglobin and hematocrit (Fig. 8B, panels a and b). Treatment with CoCl<sub>2</sub> induced both hemoglobin and hematocrit levels about 5 to 10% in wild-type mice and 20 to 25% in mutant mice (Fig. 8B, panel c). There were no significant differences in the numbers of platelets (Fig. 8B, panel d). These results show that Net is involved in regulation of hematological parameters in vivo.

## DISCUSSION

In this study, we have shown that hypoxia downregulates the TCF Net in several cell types without affecting the two other members of the subfamily (Elk-1 and Sap-1a). Hypoxia enhances Net ubiquitylation, nuclear export, and subsequent pro-



**FIG. 8.** Impaired response to a hypoxia inducer in Net mutant mice. (A) PAI-1 plasma levels. (a) Wild-type and Net mutant mice were injected intraperitoneally for 4 weeks with CoCl<sub>2</sub> (30 mg/kg), plasma was collected every week, and PAI-1 levels were determined by an enzyme-linked immunosorbent assay. (b) Average PAI-1 levels for all the treated mice shown in panel a. \*, *P* < 0.05 (Student *t* test). (B) Hematological parameters. Wild-type and Net mutant mice were injected intraperitoneally for 1 week with saline or CoCl<sub>2</sub> (45 mg/kg). Blood samples were collected before and after treatment. Hematological parameters, including those for hemoglobin (a), hematocrit (b), and platelets (d), were analyzed. (c) Inductions (%) of hemoglobin (Hb) and hematocrit (Ht) were calculated. \*, *P* < 0.05 (Student *t* test).

teasomal degradation. Net degradation in hypoxia is inhibited by overexpression of PHD1, and degradation in normoxia is induced by PHD1 downregulation, mimicking the loss of activity that occurs in hypoxia. These results indicate that the cellular oxygen sensor PHD1 is involved in Net regulation. Hypoxia induces loss of Net from the PAI-1 and *c-fos* gene promoters, and the Net binding sites of the PAI-1 promoter are necessary for hypoxic induction of promoter activity. Prior inhibition of Net by RNAi decreases hypoxic induction of PAI-1, *c-fos*, and *egr-1*. These results suggest that stimulation of transcription of these genes by hypoxia results in part from relief of transcription repression by Net. Our results also implicate Net in physiological responses to hypoxia in vivo, since Net mutant mice display altered responses to the hypoxia inducer cobalt chloride in terms of both plasma PAI-1 and hematocrit levels. In summary, this study implicates the TCF Net in the transcriptional and biological responses to hypoxia.

The transcriptional repressor Net can be switched to an activator upon phosphorylation. Previous studies have shown that posttranslational modifications, especially phosphorylation, play a key role in the cellular response to hypoxia (29). The transcription factor TAL1/SCL is ubiquitinated in hypoxia after phosphorylation by MAPK (42). The transcriptional co-

activator CREB is ubiquitinated after hypoxic repression of protein phosphatase 1 (43). HIF-1 $\alpha$  phosphorylation by p44/p42 MAPKs enhances its transcriptional activity (6, 36). Net is known to be phosphorylated through the ERK pathway (19). Interestingly, a specific inhibitor of the ERK pathway did not prevent Net degradation, suggesting that the ERK phosphorylation of Net is not required for hypoxia-induced degradation. Moreover, under our conditions, we did not detect hypoxia-induced changes in phosphorylation of Net on one of the important MAPK phosphorylation sites (serine 364). However, we cannot exclude that phosphorylation of Net could participate in the response to hypoxia, for example, at time points or sites that we have not tested.

The PHDs regulate HIF-1 $\alpha$  protein levels by hydroxylation of specific proline residues (25, 26). We show here that the PHDs also regulate Net protein levels. Inhibition of PHDs by different chemicals as well as by RNAi leads to downregulation of Net in normoxia. Furthermore, overexpression of PHD1 in hypoxia delays Net downregulation. These results suggest that decreased PHD activity in hypoxia, resulting from decreased cellular oxygen concentrations, is involved in Net downregulation. Under our experimental conditions, only PHD1 seems to regulate Net. Cellular localization of the PHDs may contribute to this apparent specificity. Net is lost from target promoters in the nucleus, suggesting that PHD regulation may occur in the nucleus. PHD1 is exclusively nuclear (32), PHD3 is distributed homogeneously in the cytoplasm, and PHD2 is mainly in the cytoplasm but can shuttle between the cytoplasm and the nucleus. HIF-1 $\alpha$  appears to be mainly regulated by PHD2 (4), although some studies suggest that all three isoforms can hydroxylate HIF-1 $\alpha$  with different efficiencies. Similarly, all three PHDs could regulate Net, but with different relative activities, depending on the cell type and the relative abundances of the PHDs (1). We are currently testing the possibility that Net is hydroxylated directly by the PHDs, using various approaches, including mass spectrometry.

Several observations raise the possibility that relief of Net repression by hypoxia may also involve its corepressor CtBP1, which interacts with the CID domain of Net (15). CtBP1 is upregulated in lung cancer cells in hypoxia (49), and high concentrations of NADH in hypoxia inhibit the interaction of hCtBP2 with Hdm2, resulting in relief of transcriptional inhibition of p53 (33). We did not detect changes in CtBP1 levels in hypoxia (data not shown). Furthermore, ChIP experiments indicate that Net is depleted from target promoters in response to hypoxia, which could be sufficient for relief of repression. However, we cannot exclude the possibility that the effects of hypoxia on CtBP1 could also modulate Net repression under certain conditions. Several studies have shown that hypoxia induces sumoylation and changes the balance between the sumoylation and ubiquitylation of several factors (14, 40). Net can be ubiquitinated (this study) and sumoylated, and sumoylation increases repression by Net (45). We did not detect any obvious changes in Net mobility in response to hypoxia that might correspond to sumoylation (data not shown). However, the possibility that hypoxia induces alterations in both the extents and the balance of Ub and SUMO modifications of Net needs to be investigated further. Interestingly, sumoylation can induce LMB-sensitive nuclear export of the Ets family transcription factor Tel (46). Nuclear export appears to be involved

in Net degradation, since an inhibitor of nuclear export decreases but does not prevent degradation. Nuclear export could enhance Net degradation by involving both nuclear and cytoplasmic proteasomes. Moreover, nuclear export could contribute, along with degradation, to the loss of Net promoter binding.

We report here that Net is implicated in the hypoxic induction of *c-fos* (44), PAI-1 (9), and *egr-1* (2). Our findings agree with a previous study that demonstrates that hypoxic induction of the *c-fos* promoter requires the TCF binding site and the serum response element (34). PAI-1 is known to be induced in hypoxia through HIF-1 $\alpha$  (17, 28). We show here that PAI-1 hypoxic induction is strongly impaired by mutation of the Net binding sites in the PAI-1 promoter and Net inhibition by RNAi. Net may similarly mediate induction of *egr-1* through the hypoxia-responsive Ets binding sites (48). Interestingly, another Net target gene, the nitric oxide synthase 2 gene (10), also responds to hypoxia (38), suggesting that Net also has a role in this process. HIF-1 $\alpha$  is classically considered to be the key transcription mediator of the hypoxic response, raising the question of how Net fits into this already well-characterized pathway. It will be interesting to compare the global contributions of HIF-1 $\alpha$  and Net to the transcriptional response to hypoxia, in order to help define their common and distinct physiological roles.

We provide evidence that Net plays a role in the hypoxic response in mice, suggesting that Net has perhaps many target genes involved in the physiological response to hypoxia. Net appears to be involved in hypoxic induction of PAI-1 in vivo, since Net mutant mice display lower plasma PAI-1 levels after CoCl<sub>2</sub> treatment. PAI-1 levels are similar in mutant and wild-type mice in normoxia, suggesting that PAI-1 is regulated by other mechanisms under these conditions. Hypoxic induction of PAI-1 in Net mutant mice is not dramatically impaired, suggesting that there are other compensatory mechanisms in mice in addition to the cellular models that we have studied. Interestingly, we found that Net contributes to the induction of several hematological parameters in mice, such as the hematocrit, raising the possibility that Net has a more global role in the hypoxic response. This observation opens new avenues for exploring the physiological and pathological mechanisms of the hypoxic response in vivo.

We have previously shown that the inhibition of Net in a tumor model decreases tumor mass and angiogenesis and increases hypoxia levels in tumors (51). These observations concord with our current study, which shows that inhibition of Net impairs hypoxic induction of genes involved in adaptive responses to hypoxia. During tumorigenesis, the hypoxic regulation of the TCF Net could be part of the pathways that are involved in tumor progression, such as the angiogenic response to hypoxia.

#### ACKNOWLEDGMENTS

We thank Frank S. Lee for providing the PHD expression vectors, William Magnant (IGBMC) for help in collecting blood samples, Marie-France Champy (MCI) for the blood analyses, Christine Wasylyk for providing mice, the Wasylyk laboratory members and Patricia Marchal for support and encouragement, Marion Poulard for her contribution in the beginning of this project, and the IGBMC core facilities.

C.G. received fellowships from the Ministère de la Recherche et Technologie, the Association pour la Recherche contre le Cancer and the CGE-DKF2 Regional Council of Alsace. G. Buchwalter received a fellowship from AICR (05-390). The Wasylyk laboratory is supported by the Ligue Nationale Française contre le Cancer (Equipe labélisée), the Ligue Régionale (Bas-Rhin) contre le Cancer and the Ligue Régionale (Haut-Rhin) contre le Cancer, the Association pour la Recherche contre le Cancer, the Centre National de la Recherche Scientifique, the Institut National de la Santé et de la Recherche Médicale, the EU (FP5 Procure project QLK6-2000-00159 and FP6 Prima project no 504587), INCA Cancéropole Grande Est (Axe IV), and AICR (05-390).

#### REFERENCES

- Appelhoff, R. J., Y. M. Tian, R. R. Raval, H. Turley, A. L. Harris, C. W. Pugh, P. J. Ratcliffe, and J. M. Gleadle. 2004. Differential function of the prolyl hydroxylases PHD1, PHD2, and PHD3 in the regulation of hypoxia-inducible factor. *J. Biol. Chem.* **279**:38458–38465.
- Ayadi, A., H. Zheng, P. Sobieszczuk, G. Buchwalter, P. Moerman, K. Alitalo, and B. Wasylyk. 2001. Net-targeted mutant mice develop a vascular phenotype and up-regulate *egr-1*. *EMBO J.* **20**:5139–5152.
- Basireddy, M., J. T. Lindsay, A. Agarwal, and D. F. Balkovetz. 2006. Epithelial cell polarity and hypoxia influence heme oxygenase-1 expression by heme in renal epithelial cells. *Am. J. Physiol. Renal Physiol.* **291**:F790–F795.
- Berra, E., E. Benizri, A. Ginouves, V. Volmat, D. Roux, and J. Pouyssegur. 2003. HIF prolyl-hydroxylase 2 is the key oxygen sensor setting low steady-state levels of HIF-1 $\alpha$  in normoxia. *EMBO J.* **22**:4082–4090.
- Berra, E., A. Ginouves, and J. Pouyssegur. 2006. The hypoxia-inducible-factor hydroxylases bring fresh air into hypoxia signalling. *EMBO Rep.* **7**:41–45.
- Berra, E., G. Pages, and J. Pouyssegur. 2000. MAP kinases and hypoxia in the control of VEGF expression. *Cancer Metastasis Rev.* **19**:139–145.
- Brahimi-Horn, C., E. Berra, and J. Pouyssegur. 2001. Hypoxia: the tumor's gateway to progression along the angiogenic pathway. *Trends Cell Biol.* **11**:S32–S36.
- Buchwalter, G., C. Gross, and B. Wasylyk. 2004. Ets ternary complex transcription factors. *Gene* **324**:1–14.
- Buchwalter, G., C. Gross, and B. Wasylyk. 2005. The ternary complex factor Net regulates cell migration through inhibition of PAI-1 expression. *Mol. Cell. Biol.* **25**:10853–10862.
- Chen, Y. H., M. D. Layne, S. W. Chung, K. Ejima, R. M. Baron, S. F. Yet, and M. A. Perrella. 2003. Elk-3 is a transcriptional repressor of nitric-oxide synthase 2. *J. Biol. Chem.* **278**:39572–39577.
- Choi, K. S., M. K. Bae, J. W. Jeong, H. E. Moon, and K. W. Kim. 2003. Hypoxia-induced angiogenesis during carcinogenesis. *J. Biochem. Mol. Biol.* **36**:120–127.
- Choy, P. C., M. Chan, G. Hatch, and R. Y. Man. 1992. Phosphatidylcholine metabolism in ischemic and hypoxic hearts. *Mol. Cell. Biochem.* **116**:53–58.
- Chung, S. W., Y. H. Chen, S. F. Yet, M. D. Layne, and M. A. Perrella. 2006. Endotoxin-induced down-regulation of Elk-3 facilitates heme oxygenase-1 induction in macrophages. *J. Immunol.* **176**:2414–2420.
- Comerford, K. M., M. O. Leonard, J. Karhausen, R. Carey, S. P. Colgan, and C. T. Taylor. 2003. Small ubiquitin-related modifier-1 modification mediates resolution of CREB-dependent responses to hypoxia. *Proc. Natl. Acad. Sci. USA* **100**:986–991.
- Criqui-Filipe, P., C. Ducret, S. M. Maira, and B. Wasylyk. 1999. Net, a negative Ras-switchable TCF, contains a second inhibition domain, the CID, that mediates repression through interactions with CtBP and de-acetylation. *EMBO J.* **18**:3392–3403.
- Ducret, C., S. M. Maira, A. Dierich, and B. Wasylyk. 1999. The Net repressor is regulated by nuclear export in response to anisomycin, UV, and heat shock. *Mol. Cell. Biol.* **19**:7076–7087.
- Fink, T., A. Kazlauskas, L. Poellinger, P. Ebbesen, and V. Zachar. 2002. Identification of a tightly regulated hypoxia-response element in the promoter of human plasminogen activator inhibitor-1. *Blood* **99**:2077–2083.
- Gao, M., and M. Karin. 2005. Regulating the regulators: control of protein ubiquitination and ubiquitin-like modifications by extracellular stimuli. *Mol. Cell* **19**:581–593.
- Giovane, A., A. Pintzas, S. M. Maira, P. Sobieszczuk, and B. Wasylyk. 1994. Net, a new ets transcription factor that is activated by Ras. *Genes Dev.* **8**:1502–1513.
- Greijer, A. E., P. van der Groep, D. Kemming, A. Shvarts, G. L. Semenza, G. A. Meijer, M. A. van de Wiel, J. A. Belien, P. J. van Diest, and E. van der Wall. 2005. Up-regulation of gene expression by hypoxia is mediated predominantly by hypoxia-inducible factor 1 (HIF-1). *J. Pathol.* **206**:291–304.
- Han, F., K. Takeda, S. Yokoyama, H. Ueda, Y. Shinozawa, K. Furuyama, and S. Shibahara. 2005. Dynamic changes in expression of heme oxygenases in mouse heart and liver during hypoxia. *Biochem. Biophys. Res. Commun.* **338**:653–659.

22. **Harris, A. L.** 2002. Hypoxia—a key regulatory factor in tumour growth. *Nat. Rev. Cancer* **2**:38–47.
23. **Hartmann-Petersen, R., M. Seeger, and C. Gordon.** 2003. Transferring substrates to the 26S proteasome. *Trends Biochem. Sci.* **28**:26–31.
24. **Huang, J., Q. Zhao, S. M. Mooney, and F. S. Lee.** 2002. Sequence determinants in hypoxia-inducible factor-1 $\alpha$  for hydroxylation by the prolyl hydroxylases PHD1, PHD2, and PHD3. *J. Biol. Chem.* **277**:39792–39800.
25. **Ivan, M., K. Kondo, H. Yang, W. Kim, J. Valiando, M. Ohh, A. Salic, J. M. Asara, W. S. Lane, and W. G. Kaelin, Jr.** 2001. HIF $\alpha$  targeted for VHL-mediated destruction by proline hydroxylation: implications for O<sub>2</sub> sensing. *Science* **292**:464–468.
26. **Jaakkola, P., D. R. Mole, Y. M. Tian, M. I. Wilson, J. Gielbert, S. J. Gaskell, A. Kriegsheim, H. F. Hebestreit, M. Mukherji, C. J. Schofield, P. H. Maxwell, C. W. Pugh, and P. J. Ratcliffe.** 2001. Targeting of HIF- $\alpha$  to the von Hippel-Lindau ubiquitylation complex by O<sub>2</sub>-regulated prolyl hydroxylation. *Science* **292**:468–472.
27. **Jain, R. K.** 2005. Normalization of tumor vasculature: an emerging concept in antiangiogenic therapy. *Science* **307**:58–62.
28. **Kietzmann, T., U. Roth, and K. Jungermann.** 1999. Induction of the plasminogen activator inhibitor-1 gene expression by mild hypoxia via a hypoxia response element binding the hypoxia-inducible factor-1 in rat hepatocytes. *Blood* **94**:4177–4185.
29. **Kumar, G. K., and J. B. Klein.** 2004. Analysis of expression and posttranslational modification of proteins during hypoxia. *J. Appl. Physiol.* **96**:1178–1186.
30. **Maira, S. M., J. M. Wurtz, and B. Wasylyk.** 1996. Net (ERP/SAP2) one of the Ras-inducible TCFs, has a novel inhibitory domain with resemblance to the helix-loop-helix motif. *EMBO J.* **15**:5849–5865.
31. **Manalo, D. J., A. Rowan, T. Lavoie, L. Natarajan, B. D. Kelly, S. Q. Ye, J. G. Garcia, and G. L. Semenza.** 2005. Transcriptional regulation of vascular endothelial cell responses to hypoxia by HIF-1. *Blood* **105**:659–669.
32. **Metzen, E., U. Berchner-Pfannschmidt, P. Stengel, J. H. Marxsen, I. Stolze, M. Klinger, W. Q. Huang, C. Wotzlaw, T. Hellwig-Burgel, W. Jelkmann, H. Acker, and J. Fandrey.** 2003. Intracellular localisation of human HIF-1  $\alpha$  hydroxylases: implications for oxygen sensing. *J. Cell Sci.* **116**:1319–1326.
33. **Mirnezami, A. H., S. J. Campbell, M. Darley, J. N. Primrose, P. W. Johnson, and J. P. Blaydes.** 2003. Hdm2 recruits a hypoxia-sensitive corepressor to negatively regulate p53-dependent transcription. *Curr. Biol.* **13**:1234–1239.
34. **Muller, J. M., B. Krauss, C. Kaltschmidt, P. A. Baeuerle, and R. A. Rupec.** 1997. Hypoxia induces c-fos transcription via a mitogen-activated protein kinase-dependent pathway. *J. Biol. Chem.* **272**:23435–23439.
35. **Pugh, C. W., and P. J. Ratcliffe.** 2003. Regulation of angiogenesis by hypoxia: role of the HIF system. *Nat. Med.* **9**:677–684.
36. **Richard, D. E., E. Berra, E. Gothie, D. Roux, and J. Pouyssegur.** 1999. p42/p44 mitogen-activated protein kinases phosphorylate hypoxia-inducible factor 1 $\alpha$  (HIF-1 $\alpha$ ) and enhance the transcriptional activity of HIF-1. *J. Biol. Chem.* **274**:32631–32637.
37. **Rosenberger, C., S. Mandriota, J. S. Jurgensen, M. S. Wiesener, J. H. Horstrup, U. Frei, P. J. Ratcliffe, P. H. Maxwell, S. Bachmann, and K. U. Eckardt.** 2002. Expression of hypoxia-inducible factor-1 $\alpha$  and -2 $\alpha$  in hypoxic and ischemic rat kidneys. *J. Am. Soc. Nephrol.* **13**:1721–1732.
38. **Rouet-Benzineb, P., S. Eddahibi, B. Raffestin, M. Laplace, S. Depond, S. Adnot, and B. Crozatier.** 1999. Induction of cardiac nitric oxide synthase 2 in rats exposed to chronic hypoxia. *J. Mol. Cell. Cardiol.* **31**:1697–1708.
39. **Sarri, E., D. Garcia-Dorado, A. Abellan, and J. Soler-Soler.** 2006. Effects of hypoxia, glucose deprivation and acidosis on phosphatidylcholine synthesis in HL-1 cardiomyocytes. CTP:phosphocholine cytidyltransferase activity correlates with sarcolemmal disruption. *Biochem. J.* **394**:325–334.
40. **Shao, R., F. P. Zhang, F. Tian, P. Anders Friberg, X. Wang, H. Sjolund, and H. Billig.** 2004. Increase of SUMO-1 expression in response to hypoxia: direct interaction with HIF-1 $\alpha$  in adult mouse brain and heart in vivo. *FEBS Lett.* **569**:293–300.
41. **Sugimoto, H., K. Okamura, S. Sugimoto, M. Satou, T. Hattori, D. E. Vance, and T. Izumi.** 2005. Sp1 is a co-activator with Ets-1, and Net is an important repressor of the transcription of CTP:phosphocholine cytidyltransferase  $\alpha$ . *J. Biol. Chem.* **280**:40857–40866.
42. **Tang, T., J. L. Arbiser, and S. J. Brandt.** 2002. Phosphorylation by mitogen-activated protein kinase mediates the hypoxia-induced turnover of the TAL1/SCL transcription factor in endothelial cells. *J. Biol. Chem.* **277**:18365–18372.
43. **Taylor, C. T., G. T. Furuta, K. Synnestvedt, and S. P. Colgan.** 2000. Phosphorylation-dependent targeting of cAMP response element binding protein to the ubiquitin/proteasome pathway in hypoxia. *Proc. Natl. Acad. Sci. USA* **97**:12091–12096.
44. **van Riggelen, J., G. Buchwalter, U. Soto, J. De-Castro Arce, H. Z. Hausen, B. Wasylyk, and F. Rosl.** 2005. Loss of net as repressor leads to constitutive increased c-fos transcription in cervical cancer cells. *J. Biol. Chem.* **280**:3286–3294.
45. **Wasylyk, C., P. Criqui-Filipe, and B. Wasylyk.** 2005. Sumoylation of the net inhibitory domain (NID) is stimulated by PIAS1 and has a negative effect on the transcriptional activity of Net. *Oncogene* **24**:820–828.
46. **Wood, L. D., B. J. Irvin, G. Nucifora, K. S. Luce, and S. W. Hiebert.** 2003. Small ubiquitin-like modifier conjugation regulates nuclear export of TEL, a putative tumor suppressor. *Proc. Natl. Acad. Sci. USA* **100**:3257–3262.
47. **Xi, L., M. Taher, C. Yin, F. Salloum, and R. C. Kukreja.** 2004. Cobalt chloride induces delayed cardiac preconditioning in mice through selective activation of HIF-1 $\alpha$  and AP-1 and iNOS signaling. *Am. J. Physiol. Heart Circ. Physiol.* **287**:H2369–H2375.
48. **Yan, S. F., J. Lu, Y. S. Zou, J. Soh-Won, D. M. Cohen, P. M. Buttrick, D. R. Cooper, S. F. Steinberg, N. Mackman, D. J. Pinsky, and D. M. Stern.** 1999. Hypoxia-associated induction of early growth response-1 gene expression. *J. Biol. Chem.* **274**:15030–15040.
49. **Zhang, Q., S. Y. Wang, A. C. Nottke, J. V. Rocheleau, D. W. Piston, and R. H. Goodman.** 2006. Redox sensor CtBP mediates hypoxia-induced tumor cell migration. *Proc. Natl. Acad. Sci. USA* **103**:9029–9033.
50. **Zhang, Y., and Y. Xiong.** 2001. Control of p53 ubiquitination and nuclear export by MDM2 and ARF. *Cell Growth Differ.* **12**:175–186.
51. **Zheng, H., C. Wasylyk, A. Ayadi, J. Abecassis, J. A. Schalken, H. Rogatsch, N. Wernert, S. M. Maira, M. C. Multon, and B. Wasylyk.** 2003. The transcription factor Net regulates the angiogenic switch. *Genes Dev.* **17**:2283–2297.





*DISCUSSION et  
PERSPECTIVES*



## Discussion et perspectives

Au cours de mon travail de thèse, je me suis penchée sur le facteur multiprotéique TFIID et plus précisément sur sa sous-unité TAF4. Au départ, curieuse de découvrir la fonction précise de TAF4 au sein du facteur de transcription TFIID, j'ai élargi par la suite mes investigations en cancérologie, après avoir découvert le rôle de TAF4 comme anti-oncogène.

### I. TAF4 et TAF4b, deux facteurs essentiels ?

Les TAFs jouent un rôle essentiel chez la levure *Saccharomyces cerevisiae* car les souches porteuses de **mutation thermosensibles** (TS) pour ces TAFs, subissent un arrêt du cycle cellulaire et meurent par apoptose. Pourtant la fonction des TAFs chez la levure a fait l'objet de nombreuses controverses, après que Moqtaderi ait montré en 1996 que les TAFs ne participent pas à l'activation de la transcription. Des études plus récentes et plus exhaustives des mutations TS ont contredit celle-ci, prouvant le caractère **essentiel** des TAFs dans la régulation du cycle cellulaire (Durso *et al.*, 2001; Kirschner *et al.*, 2002). Chez la drosophile, ces TAFs sont d'autant plus essentiels qu'ils sont également requis pour l'initiation et pour l'activation de la transcription (Wright *et al.*, 2006). Chez les Mammifères, l'inactivation des TAFs entraîne un **arrêt du cycle cellulaire** et **l'apoptose**. Les TAFs sont nécessaires pour le cycle cellulaire et pour la régulation de la transcription chez les Mammifères.

Chez la souris, l'inactivation de TAF4 *in vivo* conduit à la mort de l'embryon au stade E8.5 par suite d'un défaut de différenciation pas encore bien caractérisé. On peut préseumer qu'à partir de ce stade, une des fonctions exercées par TAF4 au cours du développement devient cruciale pour la suite du développement et ne peut plus être remplie par son **paralogue** TAF4b. Au contraire, l'inactivation de la protéine TAF4 dans les fibroblastes embryonnaires murins (MEFs) est viable.

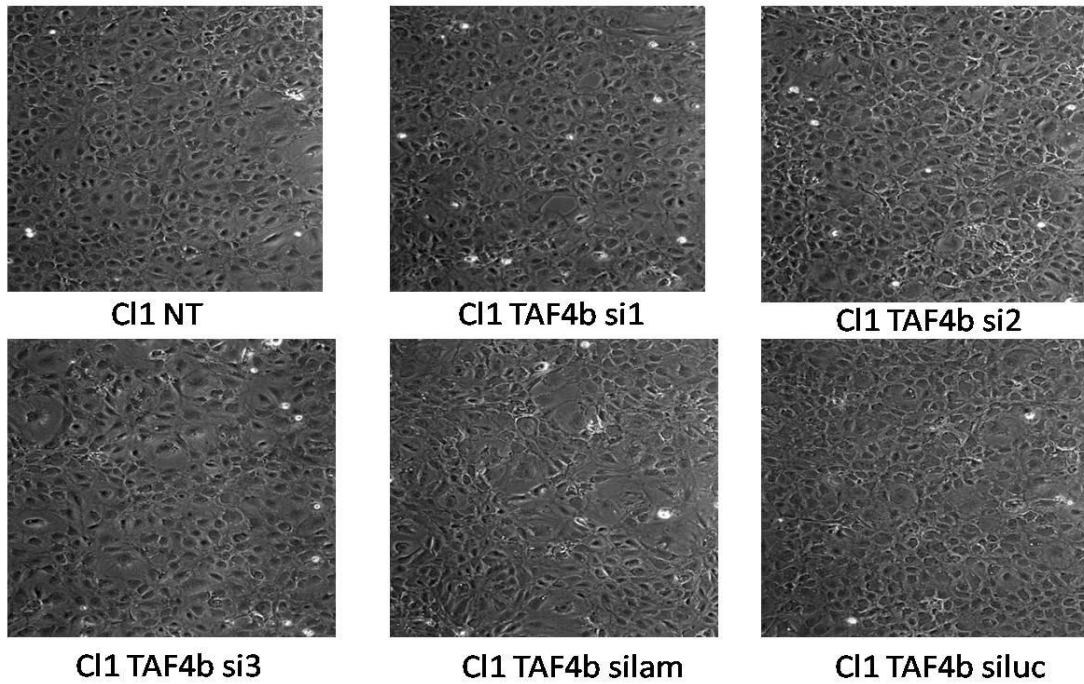
La perte de TAF4 dans les MEFs induit une surexpression des ligands TGF-beta, des récepteurs TGF-beta RI et RIII ainsi que du mitogène CTGF (connective tissue growth factor). Les cellules taf4<sup>-/-</sup> deviennent alors capables de croissance autocrine, via une activation constitutive de la voie TGF-beta. Cette perte de TAF4 dans les MEFs n'entraîne pas d'arrêt du cycle cellulaire, ni même l'apoptose des cellules, contrairement à la perte d'autres TAFs comme TAF9 dans les cellules DT40 de poulet (Chen and Manley, 2000) et

TAF8 et TAF10 chez la souris (Voss *et al.*, 2000). TAF4 a donc une action qui diffère des autres TAFs car elle ne semble pas être directement liée au cycle cellulaire. A moins que son paralogue TAF4b ne puisse prendre le relais. Ainsi, dans certains types cellulaires (les cellules germinales mâles et les cellules de la granulosa chez la femelle) et sous certaines conditions expérimentales à l'origine de la diminution de l'expression de la protéine TAF4, TAF4b semble compenser la perte de TAF4. Les MEFs Taf4<sup>-/-</sup> en sont d'ailleurs l'exemple le plus probant.

Une étude réalisée avec des **ARN interférents contre TAF4b** (Fig.96) montre clairement que l'inactivation de TAF4b dans les MEFs sauvages n'affecte pas leur viabilité, alors qu'elle induit la létalité pour les MEFs Taf4<sup>-/-</sup>. Ainsi les protéines TAF4 et TAF4b montrent peut-être une **redondance**, en tous cas elles semblent être toutes deux capables de maintenir la viabilité cellulaire et d'assurer la **complémentation** dans TFIID. Cependant, l'inactivation de TAF4 conduit à des changements de propriétés et de morphologie *in vitro* que ne semble pas entraîner la diminution de TAF4b (par interférence à l'ARN). De plus, ces deux protéines ont des fonctions différentes dans la transcription et régulent leur propre ensemble de gènes (Mengus *et al.*, 2005). Mais l'absence de ces 2 protéines au même moment conduit inévitablement à la mort cellulaire.

TAF4 et TAF4b ont donc des rôles redondants pour certaines fonctions, permettant peut-être la complémentation d'un complexe TFIID dépourvu de TAF4 avec TAF4b, comme le suggère les études réalisées au sein de l'équipe.

Pas de modification apparente de morphologie dans les clones 1 ( $Taf4^{lox/-}$ ) contenant TAF4



siRNA anti-TAF4b (1,2 et3) sur MEFs C13 ( $Taf4^{-/-}$ ) et C11 ( $Taf4^{lox/-}$ )  
Mortalité des C13 ( $Taf4^{-/-}$ ) lorsque TAF4b n'est pas exprimée.

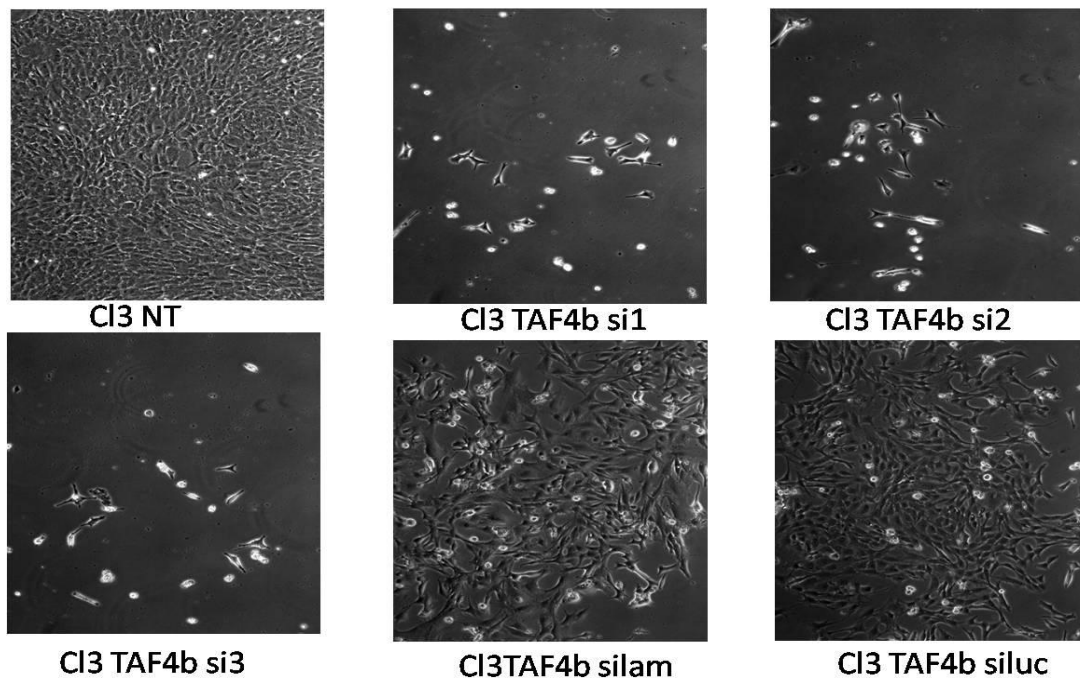


Figure 96 : Expérience d'ARN interférence contre TAF4b dans les cellules C3 et C1.

## II. Un complexe isolé à partir de TAF4 et dépourvu de TBP

### A. Un sous-complexe de TFIID ?

L'étude des TAFs dans les MEFs a mis en évidence la présence **d'un sous-complexe** de TFIID.

Si on fusionne TAF4 à une étiquette constituée de deux domaines d'affinité : SBP (Streptavidine Binding Protein) et 3 FLAG qu'on introduit dans des lignées de MEFs taf4<sup>-/-</sup>, par immunoprécipitation ou par filtration sur gel, on obtient deux complexes distincts. Le premier correspond vraisemblablement au complexe TFIID de 1,2MDa. Le second, de masse moléculaire plus faible (640kDa), est dépourvu de TBP.

#### 1. Combien y-a-t-il de TAFs ?

D'après les analyses réalisées, **au minimum 5 TAFs** sont présents dans ce sous-complexe : TAF4/TAF12, TAF6/TAF9 et TAF5 (voir manuscrit n°1).

Nous n'avons pas réussi à détecter TAF8 dans les MEFs après immunoprécipitations (IP) de TAF4. Mais nous ne disposions que d'un anticorps dirigé contre la protéine TAF8 humaine et il est probable que celui-ci ne reconnaisse pas correctement la protéine TAF8 murine. Son partenaire d'hétérodimérisation TAF10 est sorti plus tard lors de la filtration sur gel. Il semble donc peu probable que le dimère TAF8/TAF10 soit présent dans ce complexe.

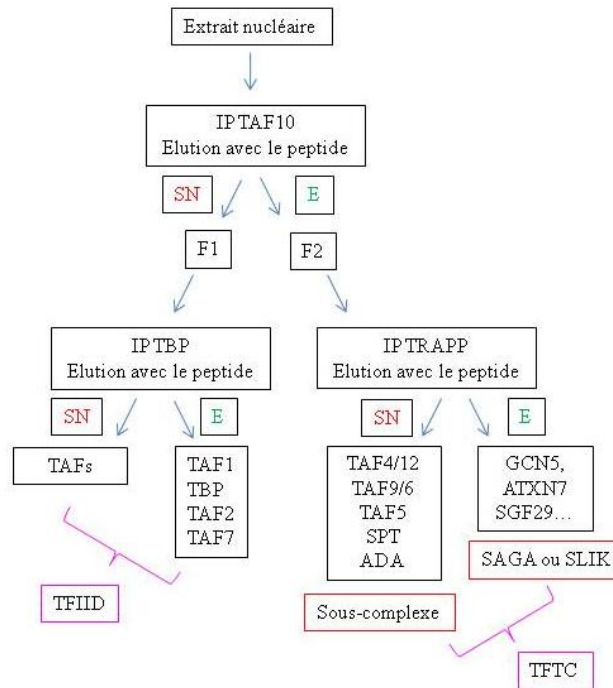
Le dimère TAF11/TAF13 n'a pas été suivi au cours de cette étude, mais ces protéines ne sont pas apparues de manière évidente lors des tests de coloration à l'argent après les séries d'IP sur TAF4. On ne peut donc pas savoir pour l'instant si TAF11 et TAF13 font partie de ce sous-complexe.

#### 2. Que représente ce complexe à 5 TAFs ?

La préparation de la fraction TFTC obtenue à partir des IP TAF10 est plutôt **hétérogène** (Timmers and Tora, 2005).

Des études plus récentes tendent à faire penser que TFTC est plutôt bipartite et pourrait être séparée en 2 fractions distinctes (Nagy and Tora, 2007) : l'une correspondant à une version homologue du complexe SLIK de 2MDa contenant GCN5, les protéines ATXN7,

SGF29 et SAP130, et l'autre correspondant au surnageant de la deuxième IP (IP TRAPP) (Fig.97) et contenant essentiellement les TAFs et les protéines SPT et ADA, assimilées aux TAFs.



**Figure 97 : Double IP d'un extrait nucléaire de cellules de Métazoaires. TFTC, un mélange de sous-complexes ?**

Après IP TRAPP de la fraction correspondant à TFTC, TAF5 et TAF6 ne figurent pas dans l'éluat. Après IP de la protéine ATXN7, TAF5 et TAF6 ne sont pas non plus décelées dans l'éluat, elles ne sont donc pas liées à l'Ataxine 7 (Devys D., unpublished results). La **deuxième fraction séparable à partir de TFTC** pourrait donc correspondre au sous-complexe que nous avons isolé dans les MEFs par IPs de la sous-unité TAF4 (manuscrit n°1).

Il peut s'agir d'un complexe à 7 TAFs mais TAF10 ne semble pas faire partie de ce complexe ou d'un complexe à 5 TAFs comme celui décrit par Wright en 2006 mais dans son étude comme dans la nôtre, toutes les sous-unités n'ont pas été examinées. Il est donc difficile pour l'instant d'exclure, de notre sous-complexe, les autres TAFs sans s'être au préalable assuré qu'elles ne sont pas présentes, même en quantités très restreintes.

### 3. Une étude stoechiométrique des composants de ce complexe

La somme des masses moléculaires de ces différents TAFs ne suffit pas à former la masse du complexe découvert. Il est donc peu probable que les TAFs de ce sous-complexe soient présents en une seule copie chacun. Si les 5 TAFs trouvés étaient tous présents en 2 copies, la masse de l'ensemble dépasserait celle du sous-complexe. A priori, si ce complexe est un sous-complexe de TFIID, pour former une **ossature pseudosymétrique**, 2 copies de TAF5 lui sont nécessaires. Que dire ensuite des autres TAFs ? La deuxième solution serait que tous les TAFs isolés possédant un HFD soient présents en doubles exemplaires, alors que TAF5 serait présent en un unique exemplaire. Cette **configuration stoechiométrique** est probable étant donné le motif structural WD40 spécifique à TAF5 qui lui permet d'interagir avec de nombreuses protéines à la fois.

Le développement de la **spectrométrie de masse quantitative** pourrait être d'une grande aide en matière de quantification du niveau d'expression des protéines de ce complexe. La détermination des interactions spécifiques entre protéines permet en effet de mieux comprendre l'assemblage de complexes protéiques. L'approche de ce complexe par la technique du **SILAC** (Mousson *et al.*, 2008) pourrait alors convenir ainsi que par toute autre technique quantitative. Dans la méthode SILAC, deux populations différant par la présence d'une protéine étiquetée, sont cultivées dans des conditions identiques de milieu basal. Seul un critère diffère : le premier milieu contient un isotope naturellement abondant d'un acide aminé sélectionné alors que le second contient une version stable et plus lourde de ce même isotope. Ces isotopes seront incorporés pendant la synthèse protéique des cellules. Les deux populations sont mélangées directement après récolte des cellules. **L'abondance relative** des protéines (ratio SILAC) est ensuite dérivée par comparaison de la masse intégrée par spectrométrie de l'aire des pics des formes marquées et non marquée d'un peptide. Nous possédons déjà des MEFs contenant TAF4 ou TAF4b étiquetées. Il est donc tout à fait possible d'envisager cette technique.

Une caractérisation structurale de ce complexe en **microscopie électronique** pourrait nous aider à mieux percevoir si ce complexe représente bien une sous-partie de TFIID. Cette étude a déjà débuté avec la collaboration de l'équipe du Dr Schultz (IGBMC).



### B. Des questions en suspens...

L'identification de la structure du sous-complexe nous permettra d'envisager les interactions protéiques possibles de complexe à 5 TAFs. Quelles sont les protéines qui se lient à ce sous-complexe ? Ce complexe est-il réellement destiné à former de l'holo-TFIID ou bien est-il **indépendant et fonctionnel** ?

La caractérisation de sa structure et de sa composition nous permettra aussi d'émettre des hypothèses quant à son rôle dans la cellule. Est-ce un complexe de **stockage** des différents TAFs en l'attente de formation du TFIID puis du PIC ? Est-ce un complexe fixé sur le promoteur des gènes dépourvus de boîte TATA, afin de faciliter la réinitiation de la transcription ? Ce complexe a-t-il besoin du soutien de la protéine TBP pour être actif ?

### III. Le complexe à 5 TAFs, un co-activateur transcriptionnel ?

#### A. L'initiation de la transcription

Pour se rendre compte si ce complexe à 5TAFs est capable d'initier la transcription, des études de transcription *in vitro*, sur une **matrice d'ADN nue**, au niveau du promoteur tardif de l'adénovirus majeur (AdML) ont été réalisées. Pour cela, nous avons isolé un extrait nucléaire (EN) de cellules Hela. Le TFIID homogène de cet EN a inactivé en chauffant l'extrait à 42°C. Pour s'assurer de l'efficacité du système de transcription *in vitro*, nous avons tenté de restaurer la transcription basale des **EN chauffés** en ajoutant la protéine TBP. Puis nous enrichi l'EN chauffé du complexe de TAFs purifié, afin de voir s'il restaure ou non la transcription basale.

Les premiers essais se sont révélés **négatifs** (Fig.98). Si les résultats sont décevants de prime abord, il faut tout de même relativiser les résultats obtenus en terme de vérité absolue. Plusieurs hypothèses peuvent expliquer ce résultat :

- ce complexe tout seul ne peut pas initier la transcription basale ; il lui faut donc recruter d'autres facteurs de transcription pour être actif. TBP est-il indispensable à l'activité transcriptionnelle de complexe dénué d'HAT ? Les expériences réalisées avec TFTC montrent que ce dernier est transcriptionnellement actif. Si ce sous-complexe ne l'est pas, qu'en est-il alors de l'autre fraction restante de TFTC (éluat de l'IP TRAPP) ?

- Le gène choisi pour cette expérience possède une boîte TATA. Ce complexe de TAFs n'est peut-être actif que sur des gènes dépourvus de boîte TATA.

-Il s'agit d'une expérience de transcription basale *in vitro* ; le complexe purifié répond peut-être à certains activateurs *in vivo* qui ne sont pas présents dans cette expérience.

-Peut-être est-ce un complexe répresseur plutôt qu'activateur ?

-Peut-être devrait-on reconstituer un système de transcription plus pur où l'on pourrait moduler la quantité des différents facteurs pour pouvoir initier la transcription de l'AdML ? Il faudrait donc recommencer les tests en ajoutant tous les facteurs purifiés du PIC à l'exception de TFIID. Une confirmation de ces tests semble nécessaire avec tous les éléments du PIC qui pourraient en outre contribuer à stabiliser ce complexe sur le promoteur. Si tel est le cas, l'implication du complexe à 5 TAFs dans l'activation de la transcription *in vitro* comme *in vivo*, sera à poursuivre. Ci-joint, une des expériences réalisées en transcription *in vitro* :

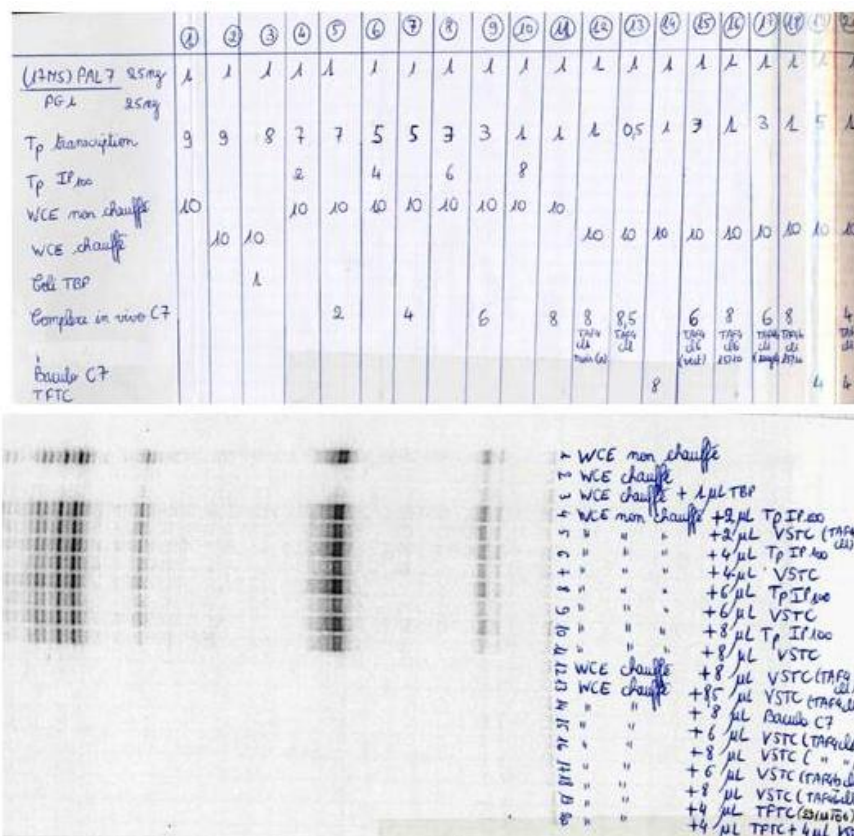


Figure 98 : Exemple de test de transcription *in vitro* réalisé avec le sous-complexe de TFIID purifié (VSTC very small TAFs complex) et des EN de cellules Hela.

## B. Quels sont les promoteurs cibles pour ce complexe ?

### 1. Variation d'expression de gènes entre les cellules C3 taf4<sup>-/-</sup> et C1

L'influence de TAF4 sur l'expression d'une grande quantité de gènes a été analysée à l'aide de puces à ADN. Sur 1000 transcrits, la perte de TAF4 résulte de la variation d'expression de 10% des gènes. La classification fonctionnelle de ces gènes via le logiciel Ingenuity indique que TAF4 influence divers processus cellulaires, tels que la progression du cycle, l'adhérence entre cellules, la biosynthèse des protéines, le cytosquelette, la transmission de signaux intracellulaires, ou encore le métabolisme. Ces données élargissent le champ d'action de TAF4 comme facteur de transcription.

### 2. Comment déterminer des gènes-cibles directs ?

La réalisation d'expériences **d'immunoprécipitations de la chromatine (ChIP)** devrait permettre de mieux comprendre quels gènes sont régulés par le facteur TAF4. A ce jour, des expériences ont été tentées avec les anticorps TAF4 (32TA) et (20TA) disponibles à l'institut ainsi qu'avec un anticorps commercial (Abcam) mais ces anticorps ne semblent pas fonctionner en ChIP. Des essais avec l'anticorps FLAG couplé à des billes sépharoses (SIGMA) ont également été réalisés sur des cellules TAF4<sup>-/-</sup> exprimant un TAF4 exogène muni d'une étiquette 3xFLAG (C3-TAF4-F). Si les conditions semblent réunies pour que le ChIP fonctionne (l'anticorps ayant déjà été testé avec succès pour le ChIP d'autres facteurs de transcription), les résultats n'ont toutefois pas été concluants jusqu'à présent. Les cellules C3-TAF4-F se révèlent également peu efficaces car au mieux, un enrichissement de 3 fois est observé sur le promoteur ribosomal étudié. Or, ce promoteur contient une boîte TATA et le gène est fortement exprimé de manière constitutive normalement...

L'équipe envisage donc de construire une nouvelle lignée de fibroblastes embryonnaires murins dans laquelle la protéine TAF4 sera fusionnée en C-terminal avec les **étiquettes HA et V5**. L'objectif est de déterminer l'ensemble des promoteurs sur lesquels se fixe TAF4 et de les comparer avec ceux sur lesquels se fixe TBP, afin d'en déduire des régions précises sur lesquelles TAF4 se fixe mais pas TBP.

Etant donné que TAF4 est partagé entre plusieurs complexes multiprotéiques (notamment TFIID et le sous-complexe de TAFs), il faudra désormais **discriminer la participation exacte de chacun de ces complexes** à la régulation génique. Le complexe à 5 TAFs se fixe-t-il spécifiquement au niveau des promoteurs dépourvus de boîte TATA ?

TAF4 et son paralogue TAF4b affichent des spécificités de gènes cibles et la compétition entre ces deux protéines régulerait l'expression des gènes contrôlant la prolifération cellulaire (Mengus *et al.*, 2005).

Lors de mes expériences d'IP et de gel filtration, j'ai pu mettre à jour un nouveau complexe autour de TAF4b. Des expériences de ChIP supplémentaires seront nécessaires pour déterminer les **gènes cibles régulés par TAF4b** et pouvoir les comparer à ceux de TAF4. De cette manière apparaîtront les gènes cibles communs aux deux facteurs paralogues dont les gènes régulateurs du cycle cellulaire font certainement partie, et les gènes cibles spécifiques.

Enfin, la réalisation d'expériences de double-immunoprécipitations de la chromatine serait une démarche élégante pour identifier les protéines liées à TAF4 durant le processus d'activation de la transcription, et apporterait une **vision plus globale et dynamique** des complexes présents au niveau de ces promoteurs.

### C. *TAF4 et les modifications post-traductionnelles des histones*

TAF4, en tant que **co-activateur**, participe à l'activation de la transcription de gènes cibles. En générant des complexes holo-TFIID altérés dans la Drosophile, l'équipe de Tjian a identifié le domaine ETO de TAF4 en tant que domaine de co-activation ciblé par Pygopus, une protéine requise pour la transcription du gène Wingless de la cuticule nue (Wright and Tjian, 2009). Ce résultat établit une fonction coactivatrice à TAF4 et fournit une stratégie permettant de disséquer les mécanismes de la fonction de TFIID et du petit complexe *in vivo*.

L'activation de la transcription de la plupart de ces gènes est couplée à l'acétylation des histones. Parmi la liste des **partenaires potentiels d'interaction** via le motif hTAF4-TAFH se trouvent des enzymes impliquées dans les modifications post-traductionnelles ou dans le remodelage de substrats chromatiniens comme par exemple la sous-unité CRSP 130

du Médiateur, plusieurs histones **déacetylases**, **déméthylases** et un nombre important de **kinases** impliquées dans le contrôle du cycle cellulaire (Wang *et al.*, 2007). Il se pourrait donc que TAF4 ait un rôle dans les modifications covalentes de la chromatine.

#### IV. Taf4, un nouveau gène suppresseur de tumeurs

##### A. L'inactivation de TAF4 et la transformation phénotypique

Mes travaux de recherche ont permis de montrer que TAF4 est un gène suppresseur de tumeur dans les fibroblastes embryonnaires murins (MEFs). En effet, son inactivation conduit à la formation de **sarcomes indifférenciés** dans des souris nues. Nous avons également mis en évidence l'effet anti-tumoral de l'acide rétinoïque (AR) sur des fibroblastes embryonnaires murins taf4<sup>-/-</sup>. L'ajout d'AR dans les MEFs a des effets contraires *in vitro* et *in vivo*. Alors qu'*in vitro*, l'AR **transforme le phénotype** des MEFs sauvages en un phénotype similaire à celui des MEFs taf4<sup>-/-</sup>, de manière surprenante, *in vivo*, il **supprime la formation de tumeur** dans les souris nues.

##### B. Induction de sarcomes

Dans les lignées fibroblastiques, l'inactivation de TAF4 entraîne la transformation des cellules et l'apparition de tumeurs dans les souris nues. Nous pourrions essayer d'extrapoler ce résultat avec des échantillons tumoraux issus de **cancers humains de type sarcomes**, afin de voir par Real-Time RT PCR d'une part, et par immunohistochimie d'autre part si le gène Taf4 est impliqué dans ces cancers, et si oui, s'il est dérégulé, réprimé voir muté.

Si tel est le cas, nous pourrions ensuite examiner les gènes dérégulés dans ces tumeurs humaines par la réalisation de puces Affymétrie puis comparer l'expression de ces gènes à celle que nous avons obtenue dans les sarcomes issus de nos lignées en culture. Nous pourrions alors mieux visualiser les gènes dont l'expression est dérégulée dans les deux types de sarcomes et essayer d'isoler la/les voie(s) de signalisation responsable(s) de l'apparition de ces tumeurs.

##### C. Un nouveau mode d'action de l'acide rétinoïque ?

Les rétinoïdes sont couramment employés dans le traitement des cancers, pour leur **action anti-proliférative**, ou leur capacité à **induire une différenciation**. Cependant dans

notre modèle, l'AR n'affecte pas la prolifération, l'apoptose ou le cycle cellulaire, mais semble agir au niveau de la migration, de l'adhésion et de l'invasion des cellules en jouant sur la **constitution de la matrice extracellulaire**.

Or, dans la littérature, l'AR est connu pour agir selon 3 modes distincts : en induisant l'apoptose, en arrêtant la croissance des cellules malignes ou en induisant une rétro-différenciation en cellules normales. Nous avons donc identifié les prémices d'un nouveau mécanisme anti-tumoral de l'AR, différent de ceux décrits par exemple dans les cellules de leucémie aigüe promyélocytaire ou de carcinome embryonnaire. Celui-ci semble basé plus spécifiquement sur le **remodelage de la matrice extracellulaire** et sur la modification des propriétés adhésives au sein du microenvironnement tumoral après ajout d'AR.

#### D. Une voie de signalisation spécifique entraînant la tumorigenèse

Des études supplémentaires via l'analyse des puces Affymetrix ont été réalisées afin de détecter la voie de signalisation responsable de cette réversion tumorigénique. Plusieurs gènes candidats nous paraissent intéressants, car ils pourraient intervenir dans les modifications cellulaires observées *in vitro* et dans la réversion de la tumorigenèse.

Parmi ces gènes, nous trouvons **Col6a3**, un membre de la famille des collagènes. Il a déjà été démontré que l'expression de Col6a3 est fortement induite dans des lignées tumorales résistantes au cisplatine (Sherman-Baust *et al.*, 2003). L'induction de son expression corrèle avec une augmentation de l'agressivité des cellules. Dans notre modèle, son expression est fortement diminuée, ce qui pourrait modifier les propriétés d'adhésion et de motilité dans l'environnement sous-cutané. De plus, Col6a3 est un **gène cible direct de la voie TGF $\beta$**  via SMAD3 (Verrecchia *et al.*, 2001). Donc cette protéine pourrait jouer un rôle clef en étant située à l'intersection des voies de TGF $\beta$  et de l'AR.

Un autre gène nous paraît également intéressant : **la transglutaminase 2 (TGM2)**. C'est l'un des gènes les plus fortement induits dans nos cellules et c'est un gène cible direct des RARs (Ma *et al.*, 2003). TGM2 catalyse la transamidation et le pontage des kératines. Il a été démontré que son expression est fortement réprimée par l'oncogène myc, mais **induite par les inhibiteurs HDAC** lors de leur activité anti-tumorale (Liu *et al.*, 2007). TGM2 a également une activité pro-apoptotique et il inhibe la croissance des tumeurs du colon et

l'angiogenèse (Lentini, 2008) ; Jones, 2006. La surexpression de TGM2 pourrait donc limiter l'invasion des cellules et inhiber leur capacité à s'établir dans l'environnement sous-cutané.

D'autres gènes sont fortement induits ou réprimés. Par exemple, la **préproenképhaline** (PENK) est fortement inhibée. Le rôle de PENK est bien étudié dans le système nerveux mais son expression dans d'autres types cellulaires a été décrite. Son expression est induite par la voie TGF-beta lors de la prolifération des chondrocytes, mais réprimée par l'AR (Villiger and Lotz, 1992). Sa fonction reste inconnue.

Enfin, nous observons une forte induction de **PRSS35**, une protéase sécrétée. Les fonctions exactes de cette protéine ne sont pas encore connues, on sait simplement qu'elle est exprimée lors de la maturation des oocytes (Wahlberg, 2008).

### E. Une piste en cours ?

L'analyse des cellules *taf4<sup>-/-</sup>* *in vitro* a été réalisée afin d'évaluer l'effet de l'AR dans des tests d'adhérence sur différents substrats, tels que le **collagène**, la **fibronectine** et la **laminine**. L'AR n'affecte pas l'adhésion en conditions classiques, mais influence l'adhérence sur des substrats qui miment mieux l'environnement sous-cutané. Si l'on compare l'adhérence sur des matrices de collagène (I et IV) des fibroblastes *taf4<sup>-/-</sup>*, traités à l'AR (TAR) et non traités (NT), on s'aperçoit que les cellules NT résistent bien mieux que les TAR à ce substrat et prolifèrent davantage. L'effet contraire est observé sur le substrat fibronectine.

Afin de poursuivre cette approche, les gènes potentiellement impliqués dans l'effet de l'AR seront inactivés par siRNA et les effets sur la **morphologie, l'adhérence et la motilité** seront évalués. De plus, ces gènes seront surexprimés à partir d'un vecteur dont le promoteur est insensible à l'AR (afin que l'expression ne soit pas réprimée par l'AR) et l'effet sur les propriétés *in vitro* sera évalué. Nous générerons également des **lignées stables** par infection rétrovirale surexprimant ces protéines. Nous regarderons leur viabilité à plus long terme, comparerons leurs différences par rapport aux C3 *taf4<sup>-/-</sup>* pour former des colonies, proliférer, et survivre. Puis nous tenterons, si besoin est, des expériences d'injection de ces cellules en sous-cutané dans les souris nues comme celles réalisées pour les C3 *taf4<sup>-/-</sup>*.

L'effet de la surexpression de ces gènes sur la formation des tumeurs *in vivo* sera ensuite évalué. On pourrait aussi examiner les effets de l'AR sur la synthèse et la localisation subcellulaire de ces gènes dérégulés lors de la tumorigenèse. L'ensemble de ces expériences devrait permettre de mieux comprendre les mécanismes induisant l'effet anti-tumoral de l'AR.

*V. Elargissement de l'étude sur la tumorigenèse induite par les fibroblastes TAF4<sup>-/-</sup>*

Des **régions hypoxiques** sont présentes dans de nombreuses tumeurs solides et jouent un rôle majeur dans la croissance tumorale et la progression métastatique. Au sein des sarcomes observés, le centre des tumeurs est nécrosé, témoin d'une privation d'oxygène. La présence de régions hypoxiques au sein d'une tumeur est un désavantage en termes de pronostic. En effet, les tumeurs hypoxiques sont moins sensibles à la chimiothérapie et la radiothérapie et ont **un potentiel métastatique élevé**. Il serait donc intéressant de regarder si la perte de TAF4 participe à l'induction de l'un de ses gènes cibles clefs dans la réponse des cellules à l'hypoxie. Si le microenvironnement tumoral joue un rôle fondamental dans la réversion de la tumorigénicité, une étude plus approfondie de l'hypoxie et notamment de la variation de l'expression de **HIF1 $\alpha$**  dans les C3 et les C3+RA, ainsi que dans les régions d'implantation des tumeurs serait à envisager.

De plus, la formation de tumeurs est à l'origine du développement de nouveaux vaisseaux sanguins approvisionnant ces tumeurs, également appelé néo-angiogenèse. Sur la photo ci-dessous (Fig.99) prise après injection de fibroblastes TAF4<sup>-/-</sup> en sous-cutané dans une souris nue, on peut observer nettement l'important flux sanguin qui circule dans les vaisseaux approvisionnant la tumeur. D'ailleurs, les images d'histopathologie des tumeurs montrent clairement la présence de **nombreux vaisseaux sanguins** irriguant les tumeurs (Fig.2 du manuscrit 2).





**Figure 99 : Vaisseaux sanguins approvisionnant la tumeur d'une souris nue ayant été injectée avec  $2 \times 10^6$  cellules de type C3.**

Cette **néo-angiogenèse** est un événement précoce et important dans la progression tumorale. Elle permet l'apport de l'oxygène et des nutriments nécessaires à la croissance de la tumeur. Les vaisseaux qui irriguent les tumeurs présentent des caractéristiques structurales et fonctionnelles différentes de celles de vaisseaux normaux. La vascularisation au sein des tumeurs solides présente des ramifications tortueuses. Le réseau sanguin qui irrigue les tumeurs assure une perfusion inadéquate en raison de l'absence de structure hiérarchisée, de la présence de shunts artério-veineux et d'extrémités aveugles. Il serait donc particulièrement intéressant de regarder si parmi les gènes responsables de la néo-angiogenèse, certains sont régulés par la protéine TAF4.

#### *VI. Vers la recherche clinique ?*

La compréhension des mécanismes moléculaires permettant la transformation cellulaire est un enjeu capital dans la lutte contre le cancer. Nous avons montré l'implication de TAF4 dans la **transformation des MEFs, dans leur migration cellulaire et dans leur adhésion**. Ceci fait de TAF4 un candidat potentiellement très intéressant dans le processus de transformation cellulaire. La compréhension plus fine du rôle de TAF4 dans tous ces mécanismes pourrait à terme aboutir à la mise en place de drogues permettant d'induire

l'action de TAF4. L'AR est un bon candidat, il est d'ailleurs déjà utilisé dans plusieurs types de sarcomes :

Le **sarcome de Kaposi (SK)** est un exemple de sarcome traité d'ores et déjà avec une isoforme de l'AR. (L'action du RA sur ce sarcome n'est peut-être pas semblable à celle que nous retrouvons qui se produit lors des injections de MEFs C3+RA dans les souris immunodéficientes). Le SK est une maladie dont la présence amène un diagnostic de sida. Cette maladie cancéreuse peut se manifester sur la peau ou se développer à l'intérieur de l'organisme. Il peut également se déclarer au niveau des tissus muqueux, comme la muqueuse buccale, ainsi qu'au niveau des ganglions lymphatiques ou des organes internes comme les intestins, les poumons ou le foie. Le SK apparaît sous la forme de lésions ou de taches qui peuvent ressembler à des ecchymoses. Le type de traitement offert dépend de l'endroit, de la taille et de l'ampleur des lésions. Le **Panretin (acide 9-cis-rétinoïque), une forme gélifiée d'acide rétinoïque**, a récemment été approuvé aux États-Unis pour le traitement du SK. C'est un inhibiteur de l'angiogenèse.

Les thérapeutiques inhibitrices de la production des cytokines ont également fait l'objet de plusieurs publications récentes. **Un dérivé de l'acide rétinoïque, inhibiteur de la production de cytokines**, notamment de l'interleukine 6, a permis en administration locale d'obtenir 27 % de réponses objectives sur les lésions traitées dans une étude sur 115 patients. Le RA-9-cis administré par voie orale a permis d'observer également plus de 37 % de réponses positives chez 38 patients le plus souvent prétraités. Des réponses encourageantes ont également été observées avec un **rétinoïde sous forme liposomale administré par voie intraveineuse**. Ces différentes approches, basées sur une meilleure compréhension de la physiopathologie du sarcome de Kaposi, semblent donc prometteuses d'autant qu'elles sont associées à une toxicité limitée par rapport aux chimiothérapies conventionnelles administrées dans ces affections.

Avec cette découverte d'une nouvelle voie de régulation de l'AR, on peut envisager que l'utilisation de l'acide rétinoïque puisse être élargie dans le futur aux traitements d'autres maladies. Mais pour cela, il reste encore à déterminer tous les intervenants de cette nouvelle voie afin de comprendre plus précisément le rôle de l'acide rétinoïque sur la matrice extracellulaire.

# *BIBLIOGRAPHIE*



## Bibliographie

Abreu JG, Ketpura NI, Reversade B, De Robertis EM (2002). Connective-tissue growth factor (CTGF) modulates cell signalling by BMP and TGF-beta. *Nat Cell Biol* **4**: 599-604.

Abu-Abed SS, Beckett BR, Chiba H, Chithalen JV, Jones G, Metzger D *et al* (1998). Mouse P450RAI (CYP26) expression and retinoic acid-inducible retinoic acid metabolism in F9 cells are regulated by retinoic acid receptor gamma and retinoid X receptor alpha. *J Biol Chem* **273**: 2409-15.

Acevedo ML, Kraus WL (2004). Transcriptional activation by nuclear receptors. *Essays Biochem* **40**: 73-88.

Ahn SH, Kim M, Buratowski S (2004). Phosphorylation of serine 2 within the RNA polymerase II C-terminal domain couples transcription and 3' end processing. *Mol Cell* **13**: 67-76.

Ahsan B, Saito TL, Hashimoto S, Muramatsu K, Tsuda M, Sasaki A *et al* (2009). MachiBase: a *Drosophila melanogaster* 5'-end mRNA transcription database. *Nucleic Acids Res* **37**: D49-53.

Ajiro K, Allis CD (2002). [Histone code hypothesis]. *Tanpakushitsu Kakusan Koso* **47**: 753-60.

Almlöf T, Wallberg AE, Gustafsson JA, Wright AP (1998). Role of important hydrophobic amino acids in the interaction between the glucocorticoid receptor tau 1-core activation domain and target factors. *Biochemistry* **37**: 9586-94.

Altucci L, Gronemeyer H (2001). The promise of retinoids to fight against cancer. *Nat Rev Cancer* **1**: 181-93.

Amann JM, Nip J, Strom DK, Lutterbach B, Harada H, Lenny N *et al* (2001). ETO, a target of t(8;21) in acute leukemia, makes distinct contacts with multiple histone deacetylases and binds mSin3A through its oligomerization domain. *Mol Cell Biol* **21**: 6470-83.

Amrolia PJ, Ramamurthy L, Saluja D, Tanese N, Jane SM, Cunningham JM (1997). The activation domain of the enhancer binding protein p45NF-E2 interacts with TAFII130 and mediates long-range activation of the alpha- and beta-globin gene loci in an erythroid cell line. *Proc Natl Acad Sci U S A* **94**: 10051-6.

Andel F, 3rd, Ladurner AG, Inouye C, Tjian R, Nogales E (1999). Three-dimensional structure of the human TFIID-IIA-IIB complex. *Science* **286**: 2153-6.

Aranda A, Pascual A (2001). Nuclear hormone receptors and gene expression. *Physiol Rev* **81**: 1269-304.

- Archer TK, Lee HL, Cordingley MG, Mymryk JS, Fragoso G, Berard DS *et al* (1994). Differential steroid hormone induction of transcription from the mouse mammary tumor virus promoter. *Mol Endocrinol* **8**: 568-76.
- Armache KJ, Kettenberger H, Cramer P (2003). Architecture of initiation-competent 12-subunit RNA polymerase II. *Proc Natl Acad Sci U S A* **100**: 6964-8.
- Asahara H, Santoso B, Guzman E, Du K, Cole PA, Davidson I *et al* (2001). Chromatin-dependent cooperativity between constitutive and inducible activation domains in CREB. *Mol Cell Biol* **21**: 7892-900.
- Auble DT, Hansen KE, Mueller CG, Lane WS, Thorner J, Hahn S (1994). Mot1, a global repressor of RNA polymerase II transcription, inhibits TBP binding to DNA by an ATP-dependent mechanism. *Genes Dev* **8**: 1920-34.
- Auricchio F (1989). Phosphorylation of steroid receptors. *J Steroid Biochem* **32**: 613-22.
- Auwerx J, Drouin J, Laudet V (2003). Recepteurs a la Provencale. EMBO workshop on the biology of nuclear receptors. *EMBO Rep* **4**: 1122-6.
- Baek HJ, Kang YK, Roeder RG (2006). Human Mediator enhances basal transcription by facilitating recruitment of transcription factor IIB during preinitiation complex assembly. *J Biol Chem* **281**: 15172-81.
- Bai Y, Perez GM, Beechem JM, Weil PA (1997). Structure-function analysis of TAF130: identification and characterization of a high-affinity TATA-binding protein interaction domain in the N terminus of yeast TAF(II)130. *Mol Cell Biol* **17**: 3081-93.
- Bajic VB, Tan SL, Christoffels A, Schonbach C, Lipovich L, Yang L *et al* (2006). Mice and men: their promoter properties. *PLoS Genet* **2**: e54.
- Baker SP, Grant PA (2007). The SAGA continues: expanding the cellular role of a transcriptional co-activator complex. *Oncogene* **26**: 5329-40.
- Bashor MM, Toft DO, Chytil F (1973). In vitro binding of retinol to rat-tissue components. *Proc Natl Acad Sci U S A* **70**: 3483-7.
- Bastien J, Adam-Stitah S, Plassat JL, Chambon P, Rochette-Egly C (2002). The phosphorylation site located in the A region of retinoic X receptor alpha is required for the antiproliferative effect of retinoic acid (RA) and the activation of RA target genes in F9 cells. *J Biol Chem* **277**: 28683-9.
- Bastien J, Adam-Stitah S, Riedl T, Egly JM, Chambon P, Rochette-Egly C (2000). TFIIF interacts with the retinoic acid receptor gamma and phosphorylates its AF-1-activating domain through cdk7. *J Biol Chem* **275**: 21896-904.
- Bastien J, Rochette-Egly C (2004). Nuclear retinoid receptors and the transcription of retinoid-target genes. *Gene* **328**: 1-16.

- Bell B, Scheer E, Tora L (2001). Identification of hTAF(II)80 delta links apoptotic signaling pathways to transcription factor TFIID function. *Mol Cell* **8**: 591-600.
- Bell B, Tora L (1999). Regulation of gene expression by multiple forms of TFIID and other novel TAFII-containing complexes. *Exp Cell Res* **246**: 11-9.
- Bellorini M, Dantonel JC, Yoon JB, Roeder RG, Tora L, Mantovani R (1996). The major histocompatibility complex class II Ea promoter requires TFIID binding to an initiator sequence. *Mol Cell Biol* **16**: 503-12.
- Benoist C, O'Hare K, Breathnach R, Chambon P (1980). The ovalbumin gene-sequence of putative control regions. *Nucleic Acids Res* **8**: 127-42.
- Bentley D (2002). The mRNA assembly line: transcription and processing machines in the same factory. *Curr Opin Cell Biol* **14**: 336-42.
- Berard J, Laboune F, Mukuna M, Masse S, Kothary R, Bradley WE (1996). Lung tumors in mice expressing an antisense RARbeta2 transgene. *FASEB J* **10**: 1091-7.
- Berndsen CE, Denu JM (2008). Catalysis and substrate selection by histone/protein lysine acetyltransferases. *Curr Opin Struct Biol* **18**: 682-9.
- Bhattacharya S, Takada S, Jacobson RH (2007). Structural analysis and dimerization potential of the human TAF5 subunit of TFIID. *Proc Natl Acad Sci U S A* **104**: 1189-94.
- Binkley N, Krueger D (2000). Hypervitaminosis A and bone. *Nutr Rev* **58**: 138-44.
- Birck C, Poch O, Romier C, Ruff M, Mengus G, Lavigne AC *et al* (1998). Human TAF(II)28 and TAF(II)18 interact through a histone fold encoded by atypical evolutionary conserved motifs also found in the SPT3 family. *Cell* **94**: 239-49.
- Bird AP (1986). CpG-rich islands and the function of DNA methylation. *Nature* **321**: 209-13.
- Blackwell TK, Walker AK (2008). OMA-gosh, where's that TAF? *Cell* **135**: 18-20.
- Blake MC, Jambou RC, Swick AG, Kahn JW, Azizkhan JC (1990). Transcriptional initiation is controlled by upstream GC-box interactions in a TATAA-less promoter. *Mol Cell Biol* **10**: 6632-41.
- Blanco JC, Dey A, Leid M, Minucci S, Park BK, Jurutka PW *et al* (1996). Inhibition of ligand induced promoter occupancy in vivo by a dominant negative RXR. *Genes Cells* **1**: 209-21.
- Blomhoff R, Green MH, Berg T, Norum KR (1990). Transport and storage of vitamin A. *Science* **250**: 399-404.

- Blomhoff R, Rasmussen M, Nilsson A, Norum KR, Berg T, Blaner WS *et al* (1985). Hepatic retinol metabolism. Distribution of retinoids, enzymes, and binding proteins in isolated rat liver cells. *J Biol Chem* **260**: 13560-5.
- Bocco JL, Bahr A, Goetz J, Hauss C, Kallunki T, Kedinger C *et al* (1996). In vivo association of ATF $\alpha$  with JNK/SAP kinase activities. *Oncogene* **12**: 1971-80.
- Bocquel MT, Kumar V, Stricker C, Chambon P, Gronemeyer H (1989). The contribution of the N- and C-terminal regions of steroid receptors to activation of transcription is both receptor and cell-specific. *Nucleic Acids Res* **17**: 2581-95.
- Bodwell JE, Orti E, Coull JM, Pappin DJ, Smith LI, Swift F (1991). Identification of phosphorylated sites in the mouse glucocorticoid receptor. *J Biol Chem* **266**: 7549-55.
- Bollag W (1971). Therapy of chemically induced skin tumors of mice with vitamin A palmitate and vitamin A acid. *Experientia* **27**: 90-2.
- Bollag W (1972). Prophylaxis of chemically induced papillomas and carcinomas of mouse skin by vitamin A-acid. *Experientia* **28**: 1219-20.
- Bollag W, Matter A (1981). From vitamin A to retinoids in experimental and clinical oncology: achievements, failures, and outlook. *Ann N Y Acad Sci* **359**: 9-23.
- Bouazoune K, Miranda TB, Jones PA, Kingston RE (2009). Analysis of individual remodeled nucleosomes reveals decreased histone-DNA contacts created by hSWI/SNF. *Nucleic Acids Res*.
- Bouwman P, Philipsen S (2002). Regulation of the activity of Sp1-related transcription factors. *Mol Cell Endocrinol* **195**: 27-38.
- Boyer-Guittaut M, Birsoy K, Potel C, Elliott G, Jaffray E, Desterro JM *et al* (2005). SUMO-1 modification of human transcription factor (TF) IID complex subunits: inhibition of TFIID promoter-binding activity through SUMO-1 modification of hsTAF5. *J Biol Chem* **280**: 9937-45.
- Brand M, Leurent C, Mallouh V, Tora L, Schultz P (1999a). Three-dimensional structures of the TAFII-containing complexes TFIID and TFIIIC. *Science* **286**: 2151-3.
- Brand M, Yamamoto K, Staub A, Tora L (1999b). Identification of TATA-binding protein-free TAFII-containing complex subunits suggests a role in nucleosome acetylation and signal transduction. *J Biol Chem* **274**: 18285-9.
- Brandeis M, Frank D, Keshet I, Siegfried Z, Mendelsohn M, Nemes A *et al* (1994). Sp1 elements protect a CpG island from de novo methylation. *Nature* **371**: 435-8.
- Breathnach R, Chambon P (1981). Organization and expression of eucaryotic split genes coding for proteins. *Annu Rev Biochem* **50**: 349-83.



- Breneman D, Sheth P, Berger V, Naini V, Stevens V (2007). Phase II clinical trial of bexarotene gel 1% in psoriasis. *J Drugs Dermatol* **6**: 501-6.
- Brownell JE, Allis CD (1995). An activity gel assay detects a single, catalytically active histone acetyltransferase subunit in *Tetrahymena* macronuclei. *Proc Natl Acad Sci U S A* **92**: 6364-8.
- Brownell JE, Zhou J, Ranalli T, Kobayashi R, Edmondson DG, Roth SY *et al* (1996). *Tetrahymena* histone acetyltransferase A: a homolog to yeast Gcn5p linking histone acetylation to gene activation. *Cell* **84**: 843-51.
- Buck J, Derguini F, Levi E, Nakanishi K, Hammerling U (1991). Intracellular signaling by 14-hydroxy-4,14-retro-retinol. *Science* **254**: 1654-6.
- Buck J, Grun F, Derguini F, Chen Y, Kimura S, Noy N *et al* (1993). Anhydroretinol: a naturally occurring inhibitor of lymphocyte physiology. *J Exp Med* **178**: 675-80.
- Bulger M, Groudine M (1999). Looping versus linking: toward a model for long-distance gene activation. *Genes Dev* **13**: 2465-77.
- Bulger M, Groudine M (2002). TRAPping enhancer function. *Nat Genet* **32**: 555-6.
- Buratowski S, Hahn S, Guarente L, Sharp PA (1989). Five intermediate complexes in transcription initiation by RNA polymerase II. *Cell* **56**: 549-61.
- Burke TW, Kadonaga JT (1996). *Drosophila* TFIID binds to a conserved downstream basal promoter element that is present in many TATA-box-deficient promoters. *Genes Dev* **10**: 711-24.
- Burke TW, Kadonaga JT (1997). The downstream core promoter element, DPE, is conserved from *Drosophila* to humans and is recognized by TAFII60 of *Drosophila*. *Genes Dev* **11**: 3020-31.
- Burley SK (1996). The TATA box binding protein. *Curr Opin Struct Biol* **6**: 69-75.
- Burley SK, Xie X, Clark KL, Shu F (1997). Histone-like transcription factors in eukaryotes. *Curr Opin Struct Biol* **7**: 94-102.
- Bushnell DA, Kornberg RD (2003). Complete, 12-subunit RNA polymerase II at 4.1-A resolution: implications for the initiation of transcription. *Proc Natl Acad Sci U S A* **100**: 6969-73.
- Bushnell DA, Westover KD, Davis RE, Kornberg RD (2004). Structural basis of transcription: an RNA polymerase II-TFIIB cocystal at 4.5 Angstroms. *Science* **303**: 983-8.
- Butler JE, Kadonaga JT (2002). The RNA polymerase II core promoter: a key component in the regulation of gene expression. *Genes Dev* **16**: 2583-92.

- Cantin GT, Stevens JL, Berk AJ (2003). Activation domain-mediator interactions promote transcription preinitiation complex assembly on promoter DNA. *Proc Natl Acad Sci U S A* **100**: 12003-8.
- Capper NS, McKibbin IM, Prentice JH (1931). Carotene and vitamin A: The conversion of carotene into vitamin A by fowl. *Biochem J* **25**: 265-74.
- Carninci P, Kasukawa T, Katayama S, Gough J, Frith MC, Maeda N *et al* (2005). The transcriptional landscape of the mammalian genome. *Science* **309**: 1559-63.
- Carninci P, Sandelin A, Lenhard B, Katayama S, Shimokawa K, Ponjavic J *et al* (2006). Genome-wide analysis of mammalian promoter architecture and evolution. *Nat Genet* **38**: 626-35.
- Castaigne S, Chomienne C, Daniel MT, Ballerini P, Berger R, Fenaux P *et al* (1990). All-trans retinoic acid as a differentiation therapy for acute promyelocytic leukemia. I. Clinical results. *Blood* **76**: 1704-9.
- Cavaillès V, Dauvois S, L'Horset F, Lopez G, Hoare S, Kushner PJ *et al* (1995). Nuclear factor RIP140 modulates transcriptional activation by the estrogen receptor. *EMBO J* **14**: 3741-51.
- Chalkley GE, Verrijzer CP (1999). DNA binding site selection by RNA polymerase II TAFs: a TAF(II)250-TAF(II)150 complex recognizes the initiator. *EMBO J* **18**: 4835-45.
- Chambon P (1996). A decade of molecular biology of retinoic acid receptors. *FASEB J* **10**: 940-54.
- Chambon P (2004). How I became one of the fathers of a superfamily. *Nat Med* **10**: 1027-31.
- Chambraud B, Berry M, Redeuilh G, Chambon P, Baulieu EE (1990). Several regions of human estrogen receptor are involved in the formation of receptor-heat shock protein 90 complexes. *J Biol Chem* **265**: 20686-91.
- Chao DM, Gadbois EL, Murray PJ, Anderson SF, Sonu MS, Parvin JD *et al* (1996). A mammalian SRB protein associated with an RNA polymerase II holoenzyme. *Nature* **380**: 82-5.
- Chatton B, Bocco JL, Goetz J, Gaire M, Lutz Y, Kedinger C (1994). Jun and Fos heterodimerize with ATF $\alpha$ , a member of the ATF/CREB family and modulate its transcriptional activity. *Oncogene* **9**: 375-85.
- Chen BS, Hampsey M (2004). Functional interaction between TFIIB and the Rpb2 subunit of RNA polymerase II: implications for the mechanism of transcription initiation. *Mol Cell Biol* **24**: 3983-91.
- Chen D, Pace PE, Coombes RC, Ali S (1999). Phosphorylation of human estrogen receptor  $\alpha$  by protein kinase A regulates dimerization. *Mol Cell Biol* **19**: 1002-15.

- Chen GQ, Zhu J, Shi XG, Ni JH, Zhong HJ, Si GY *et al* (1996). In vitro studies on cellular and molecular mechanisms of arsenic trioxide (As<sub>2</sub>O<sub>3</sub>) in the treatment of acute promyelocytic leukemia: As<sub>2</sub>O<sub>3</sub> induces NB4 cell apoptosis with downregulation of Bcl-2 expression and modulation of PML-RAR alpha/PML proteins. *Blood* **88**: 1052-61.
- Chen HT, Hahn S (2004). Mapping the location of TFIIB within the RNA polymerase II transcription preinitiation complex: a model for the structure of the PIC. *Cell* **119**: 169-80.
- Chen J, Sun M, Kent WJ, Huang X, Xie H, Wang W *et al* (2004). Over 20% of human transcripts might form sense-antisense pairs. *Nucleic Acids Res* **32**: 4812-20.
- Chen JL, Attardi LD, Verrijzer CP, Yokomori K, Tjian R (1994). Assembly of recombinant TFIID reveals differential coactivator requirements for distinct transcriptional activators. *Cell* **79**: 93-105.
- Chen Z, Manley JL (2000). Robust mRNA transcription in chicken DT40 cells depleted of TAF(II)31 suggests both functional degeneracy and evolutionary divergence. *Mol Cell Biol* **20**: 5064-76.
- Chevillard-Briet M, Trouche D, Vandel L (2002). Control of CBP co-activating activity by arginine methylation. *EMBO J* **21**: 5457-66.
- Chiang CM, Ge H, Wang Z, Hoffmann A, Roeder RG (1993). Unique TATA-binding protein-containing complexes and cofactors involved in transcription by RNA polymerases II and III. *EMBO J* **12**: 2749-62.
- Choi SJ, Chung SS, Rho EJ, Lee HW, Lee MH, Choi HS *et al* (2006). Negative modulation of RXRalpha transcriptional activity by small ubiquitin-related modifier (SUMO) modification and its reversal by SUMO-specific protease SUSP1. *J Biol Chem* **281**: 30669-77.
- Chomienne C, Balitrand N, Ballerini P, Castaigne S, de The H, Degos L (1991). All-trans retinoic acid modulates the retinoic acid receptor-alpha in promyelocytic cells. *J Clin Invest* **88**: 2150-4.
- Chong JA, Moran MM, Teichmann M, Kaczmarek JS, Roeder R, Clapham DE (2005). TATA-binding protein (TBP)-like factor (TLF) is a functional regulator of transcription: reciprocal regulation of the neurofibromatosis type 1 and c-fos genes by TLF/TRF2 and TBP. *Mol Cell Biol* **25**: 2632-43.
- Choy B, Green MR (1993). Eukaryotic activators function during multiple steps of preinitiation complex assembly. *Nature* **366**: 531-6.
- Chrivia JC, Kwok RP, Lamb N, Hagiwara M, Montminy MR, Goodman RH (1993). Phosphorylated CREB binds specifically to the nuclear protein CBP. *Nature* **365**: 855-9.
- Cler E, Papai G, Schultz P, Davidson I (2009). Recent advances in understanding the structure and function of general transcription factor TFIID. *Cell Mol Life Sci* **66**: 2123-34.

- Conaway RC, Conaway JW (1993). General initiation factors for RNA polymerase II. *Annu Rev Biochem* **62**: 161-90.
- Cooper AD (1997). Hepatic uptake of chylomicron remnants. *J Lipid Res* **38**: 2173-92.
- Cooper SJ, Trinklein ND, Anton ED, Nguyen L, Myers RM (2006). Comprehensive analysis of transcriptional promoter structure and function in 1% of the human genome. *Genome Res* **16**: 1-10.
- Corcoran J, So PL, Maden M (2002). Absence of retinoids can induce motoneuron disease in the adult rat and a retinoid defect is present in motoneuron disease patients. *J Cell Sci* **115**: 4735-41.
- Corden J, Wasylyk B, Buchwalder A, Sassone-Corsi P, Kedinger C, Chambon P (1980). Promoter sequences of eukaryotic protein-coding genes. *Science* **209**: 1406-14.
- Cormack BP, Struhl K (1992). The TATA-binding protein is required for transcription by all three nuclear RNA polymerases in yeast cells. *Cell* **69**: 685-96.
- Cortes P, Flores O, Reinberg D (1992). Factors involved in specific transcription by mammalian RNA polymerase II: purification and analysis of transcription factor IIA and identification of transcription factor IIJ. *Mol Cell Biol* **12**: 413-21.
- Costa VL, Henrique R, Jeronimo C (2007). Epigenetic markers for molecular detection of prostate cancer. *Dis Markers* **23**: 31-41.
- Coulombe B, Li J, Greenblatt J (1994). Topological localization of the human transcription factors IIA, IIB, TATA box-binding protein, and RNA polymerase II-associated protein 30 on a class II promoter. *J Biol Chem* **269**: 19962-7.
- Cramer P, Bushnell DA, Fu J, Gnatt AL, Maier-Davis B, Thompson NE *et al* (2000). Architecture of RNA polymerase II and implications for the transcription mechanism. *Science* **288**: 640-9.
- Damm K, Heyman RA, Umesono K, Evans RM (1993). Functional inhibition of retinoic acid response by dominant negative retinoic acid receptor mutants. *Proc Natl Acad Sci U S A* **90**: 2989-93.
- Dantoni JC, Wurtz JM, Poch O, Moras D, Tora L (1999). The TBP-like factor: an alternative transcription factor in metazoa? *Trends Biochem Sci* **24**: 335-9.
- Davidson I, Kobi D, Fadloun A, Mengus G (2005). New insights into TAFs as regulators of cell cycle and signaling pathways. *Cell Cycle* **4**: 1486-90.
- Davies RE (1967). Effect of vitamin A on 7,12-Dimethylbenz(alpha)anthracene-induced papillomas in rhino mouse skin. *Cancer Res* **27**: 237-41.

- Davis JN, McGhee L, Meyers S (2003). The ETO (MTG8) gene family. *Gene* **303**: 1-10.
- Day HG (1997). Young rats need unknown growth factors (McCollum, 1913-1917). *J Nutr* **127**: 1029S-1031S.
- de Lera AR, Bourguet W, Altucci L, Gronemeyer H (2007). Design of selective nuclear receptor modulators: RAR and RXR as a case study. *Nat Rev Drug Discov* **6**: 811-20.
- de The H, Lavau C, Marchio A, Chomienne C, Degos L, Dejean A (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**: 675-84.
- DeCaprio JA, Ludlow JW, Figge J, Shew JY, Huang CM, Lee WH *et al* (1988). SV40 large tumor antigen forms a specific complex with the product of the retinoblastoma susceptibility gene. *Cell* **54**: 275-83.
- Delmotte MH, Tahayato A, Formstecher P, Lefebvre P (1999). Serine 157, a retinoic acid receptor alpha residue phosphorylated by protein kinase C in vitro, is involved in RXR.RARalpha heterodimerization and transcriptional activity. *J Biol Chem* **274**: 38225-31.
- Deng W, Roberts SG (2005). A core promoter element downstream of the TATA box that is recognized by TFIIB. *Genes Dev* **19**: 2418-23.
- Denger S, Reid G, Gannon F (2008). Expression of the estrogen receptor during differentiation of human osteoclasts. *Steroids* **73**: 765-74.
- Derguini F, Nakanishi K, Hammerling U, Chua R, Eppinger T, Levi E *et al* (1995). 13,14-Dihydroxy-retinol, a new bioactive retinol metabolite. *J Biol Chem* **270**: 18875-80.
- Di C, Liao S, Adamson DC, Parrett TJ, Broderick DK, Shi Q *et al* (2005). Identification of OTX2 as a medulloblastoma oncogene whose product can be targeted by all-trans retinoic acid. *Cancer Res* **65**: 919-24.
- Dikstein R, Zhou S, Tjian R (1996). Human TAFII 105 is a cell type-specific TFIID subunit related to hTAFII130. *Cell* **87**: 137-46.
- Dilworth FJ, Fromental-Ramain C, Yamamoto K, Chambon P (2000). ATP-driven chromatin remodeling activity and histone acetyltransferases act sequentially during transactivation by RAR/RXR In vitro. *Mol Cell* **6**: 1049-58.
- Dolle P (2009). Developmental expression of retinoic acid receptors (RARs). *Nucl Recept Signal* **7**: e006.
- Dotson MR, Yuan CX, Roeder RG, Myers LC, Gustafsson CM, Jiang YW *et al* (2000). Structural organization of yeast and mammalian mediator complexes. *Proc Natl Acad Sci U S A* **97**: 14307-10.

- Dowhan DH, Muscat GE (1996). Characterization of the AB (AF-1) region in the muscle-specific retinoid X receptor-gamma: evidence that the AF-1 region functions in a cell-specific manner. *Nucleic Acids Res* **24**: 264-71.
- Dowling JE, Wald G (1958). Vitamin a Deficiency and Night Blindness. *Proc Natl Acad Sci U S A* **44**: 648-61.
- Dubrovskaya V, Lavigne AC, Davidson I, Acker J, Staub A, Tora L (1996). Distinct domains of hTAFII100 are required for functional interaction with transcription factor TFIIF beta (RAP30) and incorporation into the TFIID complex. *EMBO J* **15**: 3702-12.
- Dunah AW, Jeong H, Griffin A, Kim YM, Standaert DG, Hersch SM *et al* (2002). Sp1 and TAFII130 transcriptional activity disrupted in early Huntington's disease. *Science* **296**: 2238-43.
- Durso RJ, Fisher AK, Albright-Frey TJ, Reese JC (2001). Analysis of TAF90 mutants displaying allele-specific and broad defects in transcription. *Mol Cell Biol* **21**: 7331-44.
- Dutta A, Sen T, Banerji A, Das S, Chatterjee A (2009). Studies on Multifunctional Effect of All-Trans Retinoic Acid (ATRA) on Matrix Metalloproteinase-2 (MMP-2) and Its Regulatory Molecules in Human Breast Cancer Cells (MCF-7). *J Oncol* **2009**: 627840.
- Dynlacht BD, Hoey T, Tjian R (1991). Isolation of coactivators associated with the TATA-binding protein that mediate transcriptional activation. *Cell* **66**: 563-76.
- Edwards AM, Kane CM, Young RA, Kornberg RD (1991). Two dissociable subunits of yeast RNA polymerase II stimulate the initiation of transcription at a promoter in vitro. *J Biol Chem* **266**: 71-5.
- Efstratiadis A, Posakony JW, Maniatis T, Lawn RM, O'Connell C, Spritz RA *et al* (1980). The structure and evolution of the human beta-globin gene family. *Cell* **21**: 653-68.
- Egea PF, Mitschler A, Rochel N, Ruff M, Chambon P, Moras D (2000). Crystal structure of the human RXRalpha ligand-binding domain bound to its natural ligand: 9-cis retinoic acid. *EMBO J* **19**: 2592-601.
- El Kaderi B, Medler S, Raghunayakula S, Ansari A (2009). Gene looping is conferred by activator-dependent interaction of transcription initiation and termination machineries. *J Biol Chem*.
- Elsby LM, Roberts SG (2004). The role of TFIIB conformation in transcriptional regulation. *Biochem Soc Trans* **32**: 1098-9.
- Epping MT, Wang L, Edel MJ, Carlee L, Hernandez M, Bernards R (2005). The human tumor antigen PRAME is a dominant repressor of retinoic acid receptor signaling. *Cell* **122**: 835-47.
- Escriva H, Delaunay F, Laudet V (2000). Ligand binding and nuclear receptor evolution. *Bioessays* **22**: 717-27.

- Esteller M (2007). Cancer epigenomics: DNA methylomes and histone-modification maps. *Nat Rev Genet* **8**: 286-98.
- Evans R, Fairley JA, Roberts SG (2001). Activator-mediated disruption of sequence-specific DNA contacts by the general transcription factor TFIIB. *Genes Dev* **15**: 2945-9.
- Fahrner TJ, Carroll SL, Milbrandt J (1990). The NGFI-B protein, an inducible member of the thyroid/steroid receptor family, is rapidly modified posttranslationally. *Mol Cell Biol* **10**: 6454-9.
- Fairley JA, Evans R, Hawkes NA, Roberts SG (2002). Core promoter-dependent TFIIB conformation and a role for TFIIB conformation in transcription start site selection. *Mol Cell Biol* **22**: 6697-705.
- Farboud B, Privalsky ML (2004). Retinoic acid receptor-alpha is stabilized in a repressive state by its C-terminal, isotype-specific F domain. *Mol Endocrinol* **18**: 2839-53.
- Featherstone M (2002). Coactivators in transcription initiation: here are your orders. *Curr Opin Genet Dev* **12**: 149-55.
- Fell HB, Mellanby E (1953). Metaplasia produced in cultures of chick ectoderm by high vitamin A. *J Physiol* **119**: 470-88.
- Felsenfeld G, Groudine M (2003). Controlling the double helix. *Nature* **421**: 448-53.
- Flores O, Ha I, Reinberg D (1990). Factors involved in specific transcription by mammalian RNA polymerase II. Purification and subunit composition of transcription factor IIF. *J Biol Chem* **265**: 5629-34.
- Floyd ZE, Stephens JM (2002). Interferon-gamma-mediated activation and ubiquitin-proteasome-dependent degradation of PPARgamma in adipocytes. *J Biol Chem* **277**: 4062-8.
- Fondell JD, Ge H, Roeder RG (1996). Ligand induction of a transcriptionally active thyroid hormone receptor coactivator complex. *Proc Natl Acad Sci U S A* **93**: 8329-33.
- Fortuna VA, Martucci RB, Trugo LC, Borojevic R (2003). Hepatic stellate cells uptake of retinol associated with retinol-binding protein or with bovine serum albumin. *J Cell Biochem* **90**: 792-805.
- Francastel C, Magis W, Groudine M (2001). Nuclear relocation of a transactivator subunit precedes target gene activation. *Proc Natl Acad Sci U S A* **98**: 12120-5.
- Fraser P, Grosveld F (1998). Locus control regions, chromatin activation and transcription. *Curr Opin Cell Biol* **10**: 361-5.

- Freedman LP, Luisi BF, Korszun ZR, Basavappa R, Sigler PB, Yamamoto KR (1988). The function and structure of the metal coordination sites within the glucocorticoid receptor DNA binding domain. *Nature* **334**: 543-6.
- Freiman RN, Albright SR, Chu LE, Zheng S, Liang HE, Sha WC *et al* (2002). Redundant role of tissue-selective TAF(II)105 in B lymphocytes. *Mol Cell Biol* **22**: 6564-72.
- Freiman RN, Albright SR, Zheng S, Sha WC, Hammer RE, Tjian R (2001). Requirement of tissue-selective TBP-associated factor TAFII105 in ovarian development. *Science* **293**: 2084-7.
- Freiman RN, Tjian R (2002). Neurodegeneration. A glutamine-rich trail leads to transcription factors. *Science* **296**: 2149-50.
- Fribourg S, Romier C, Werten S, Gangloff YG, Poterszman A, Moras D (2001). Dissecting the interaction network of multiprotein complexes by pairwise coexpression of subunits in *E. coli*. *J Mol Biol* **306**: 363-73.
- Furukawa T, Tanese N (2000). Assembly of partial TFIID complexes in mammalian cells reveals distinct activities associated with individual TATA box-binding protein-associated factors. *J Biol Chem* **275**: 29847-56.
- Gaiser F, Tan S, Richmond TJ (2000). Novel dimerization fold of RAP30/RAP74 in human TFIIF at 1.7 Å resolution. *J Mol Biol* **302**: 1119-27.
- Gangloff YG, Romier C, Thuault S, Werten S, Davidson I (2001a). The histone fold is a key structural motif of transcription factor TFIID. *Trends Biochem Sci* **26**: 250-7.
- Gangloff YG, Sanders SL, Romier C, Kirschner D, Weil PA, Tora L *et al* (2001b). Histone folds mediate selective heterodimerization of yeast TAF(II)25 with TFIID components yTAF(II)47 and yTAF(II)65 and with SAGA component ySPT7. *Mol Cell Biol* **21**: 1841-53.
- Gangloff YG, Werten S, Romier C, Carre L, Poch O, Moras D *et al* (2000). The human TFIID components TAF(II)135 and TAF(II)20 and the yeast SAGA components ADA1 and TAF(II)68 heterodimerize to form histone-like pairs. *Mol Cell Biol* **20**: 340-51.
- Garcia-Villalba P, Jimenez-Lara AM, Aranda A (1996). Vitamin D interferes with transactivation of the growth hormone gene by thyroid hormone and retinoic acid. *Mol Cell Biol* **16**: 318-27.
- Gardiner-Garden M, Frommer M (1987). CpG islands in vertebrate genomes. *J Mol Biol* **196**: 261-82.
- Gaub MP, Rochette-Egly C, Lutz Y, Ali S, Matthes H, Scheuer I *et al* (1992). Immunodetection of multiple species of retinoic acid receptor alpha: evidence for phosphorylation. *Exp Cell Res* **201**: 335-46.



- Geisberg JV, Chen JL, Ricciardi RP (1995). Subregions of the adenovirus E1A transactivation domain target multiple components of the TFIID complex. *Mol Cell Biol* **15**: 6283-90.
- Germain P, Kammerer S, Perez E, Peluso-Iltis C, Tortolani D, Zusi FC *et al* (2004). Rational design of RAR-selective ligands revealed by RARbeta crystal structure. *EMBO Rep* **5**: 877-82.
- Gershenson NI, Ioshikhes IP (2005). Synergy of human Pol II core promoter elements revealed by statistical sequence analysis. *Bioinformatics* **21**: 1295-300.
- Gill G, Pascal E, Tseng ZH, Tjian R (1994). A glutamine-rich hydrophobic patch in transcription factor Sp1 contacts the dTAFII110 component of the Drosophila TFIID complex and mediates transcriptional activation. *Proc Natl Acad Sci U S A* **91**: 192-6.
- Glass CK, Rosenfeld MG (2000). The coregulator exchange in transcriptional functions of nuclear receptors. *Genes Dev* **14**: 121-41.
- Glineur C, Zenke M, Beug H, Ghysdael J (1990). Phosphorylation of the v-erbA protein is required for its function as an oncogene. *Genes Dev* **4**: 1663-76.
- Gnatt AL, Cramer P, Fu J, Bushnell DA, Kornberg RD (2001). Structural basis of transcription: an RNA polymerase II elongation complex at 3.3 Å resolution. *Science* **292**: 1876-82.
- Goodman DW, Huang HS, Shiratori T (1965). Tissue Distribution and Metabolism of Newly Absorbed Vitamin a in the Rat. *J Lipid Res* **6**: 390-6.
- Grant PA, Duggan L, Cote J, Roberts SM, Brownell JE, Candau R *et al* (1997). Yeast Gcn5 functions in two multisubunit complexes to acetylate nucleosomal histones: characterization of an Ada complex and the SAGA (Spt/Ada) complex. *Genes Dev* **11**: 1640-50.
- Grant PA, Schieltz D, Pray-Grant MG, Steger DJ, Reese JC, Yates JR, 3rd *et al* (1998). A subset of TAF(II)s are integral components of the SAGA complex required for nucleosome acetylation and transcriptional stimulation. *Cell* **94**: 45-53.
- Green S, Chambon P (1988). Nuclear receptors enhance our understanding of transcription regulation. *Trends Genet* **4**: 309-14.
- Green S, Kumar V, Theulaz I, Wahli W, Chambon P (1988). The N-terminal DNA-binding 'zinc finger' of the oestrogen and glucocorticoid receptors determines target gene specificity. *EMBO J* **7**: 3037-44.
- Grob P, Cruse MJ, Inouye C, Peris M, Penczek PA, Tjian R *et al* (2006). Cryo-electron microscopy studies of human TFIID: conformational breathing in the integration of gene regulatory cues. *Structure* **14**: 511-20.
- Gronemeyer H, Gustafsson JA, Laudet V (2004). Principles for modulation of the nuclear receptor superfamily. *Nat Rev Drug Discov* **3**: 950-64.

- Gronemeyer H, Laudet V (1995). Transcription factors 3: nuclear receptors. *Protein Profile* **2**: 1173-308.
- Gu W, Roeder RG (1997). Activation of p53 sequence-specific DNA binding by acetylation of the p53 C-terminal domain. *Cell* **90**: 595-606.
- Guermah M, Tao Y, Roeder RG (2001). Positive and negative TAF(II) functions that suggest a dynamic TFIID structure and elicit synergy with traps in activator-induced transcription. *Mol Cell Biol* **21**: 6882-94.
- Guillemette B, Gaudreau L (2006). [H2A.Z: a histone variant that decorates gene promoters]. *Med Sci (Paris)* **22**: 941-6.
- Ha I, Roberts S, Maldonado E, Sun X, Kim LU, Green M *et al* (1993). Multiple functional domains of human transcription factor IIB: distinct interactions with two general transcription factors and RNA polymerase II. *Genes Dev* **7**: 1021-32.
- Hadzic E, Desai-Yajnik V, Helmer E, Guo S, Wu S, Koudinova N *et al* (1995). A 10-amino-acid sequence in the N-terminal A/B domain of thyroid hormone receptor alpha is essential for transcriptional activation and interaction with the general transcription factor TFIIB. *Mol Cell Biol* **15**: 4507-17.
- Hahn S (2004). Structure and mechanism of the RNA polymerase II transcription machinery. *Nat Struct Mol Biol* **11**: 394-403.
- Hahn S, Buratowski S, Sharp PA, Guarente L (1989). Isolation of the gene encoding the yeast TATA binding protein TFIID: a gene identical to the SPT15 suppressor of Ty element insertions. *Cell* **58**: 1173-81.
- Hammer GD, Krylova I, Zhang Y, Darimont BD, Simpson K, Weigel NL *et al* (1999). Phosphorylation of the nuclear receptor SF-1 modulates cofactor recruitment: integration of hormone signaling in reproduction and stress. *Mol Cell* **3**: 521-6.
- Hampsey M (1998). Molecular genetics of the RNA polymerase II general transcriptional machinery. *Microbiol Mol Biol Rev* **62**: 465-503.
- Hawkes NA, Evans R, Roberts SG (2000). The conformation of the transcription factor TFIIB modulates the response to transcriptional activators in vivo. *Curr Biol* **10**: 273-6.
- Hellmann N, Weber RE, Decker H (2008). Linked analysis of large cooperative, allosteric systems: the case of the giant HBL hemoglobins. *Methods Enzymol* **436**: 463-85.
- Herr AJ, Jensen MB, Dalmay T, Baulcombe DC (2005). RNA polymerase IV directs silencing of endogenous DNA. *Science* **308**: 118-20.
- Hirata A, Klein BJ, Murakami KS (2008). The X-ray crystal structure of RNA polymerase from Archaea. *Nature* **451**: 851-4.

- Hoey T, Dynlacht BD, Peterson MG, Pugh BF, Tjian R (1990). Isolation and characterization of the *Drosophila* gene encoding the TATA box binding protein, TFIID. *Cell* **61**: 1179-86.
- Hoey T, Weinzierl RO, Gill G, Chen JL, Dynlacht BD, Tjian R (1993). Molecular cloning and functional analysis of *Drosophila* TAF110 reveal properties expected of coactivators. *Cell* **72**: 247-60.
- Hoffmann A, Oelgeschlager T, Roeder RG (1997). Considerations of transcriptional control mechanisms: do TFIID-core promoter complexes recapitulate nucleosome-like functions? *Proc Natl Acad Sci U S A* **94**: 8928-35.
- Hoffmann A, Roeder RG (1996). Cloning and characterization of human TAF20/15. Multiple interactions suggest a central role in TFIID complex formation. *J Biol Chem* **271**: 18194-202.
- Hoffmann S, Rockenstein A, Ramaswamy A, Celik I, Wunderlich A, Lingelbach S *et al* (2007). Retinoic acid inhibits angiogenesis and tumor growth of thyroid cancer cells. *Mol Cell Endocrinol* **264**: 74-81.
- Holmes MC, Tjian R (2000). Promoter-selective properties of the TBP-related factor TRF1. *Science* **288**: 867-70.
- Hong W, Baniahmad A, Li J, Chang C, Gao W, Liu Y (2009). Bag-1M inhibits the transactivation of the glucocorticoid receptor via recruitment of corepressors. *FEBS Lett* **583**: 2451-6.
- Hopkins FG (1912). Feeding experiments illustrating the importance of accessory factors in normal dietaries. *J Physiol* **44**: 425-60.
- Horikoshi M, Carey MF, Kakidani H, Roeder RG (1988). Mechanism of action of a yeast activator: direct effect of GAL4 derivatives on mammalian TFIID-promoter interactions. *Cell* **54**: 665-9.
- Horikoshi M, Wang CK, Fujii H, Cromlish JA, Weil PA, Roeder RG (1989). Purification of a yeast TATA box-binding protein that exhibits human transcription factor IID activity. *Proc Natl Acad Sci U S A* **86**: 4843-7.
- Horikoshi M, Yamamoto T, Ohkuma Y, Weil PA, Roeder RG (1990). Analysis of structure-function relationships of yeast TATA box binding factor TFIID. *Cell* **61**: 1171-8.
- Hou SY, Wu SY, Zhou T, Thomas MC, Chiang CM (2000). Alleviation of human papillomavirus E2-mediated transcriptional repression via formation of a TATA binding protein (or TFIID)-TFIIB-RNA polymerase II-TFIIF preinitiation complex. *Mol Cell Biol* **20**: 113-25.
- Houle B, Rochette-Egly C, Bradley WE (1993). Tumor-suppressive effect of the retinoic acid receptor beta in human epidermoid lung cancer cells. *Proc Natl Acad Sci U S A* **90**: 985-9.

- Housley PR, Sanchez ER, Danielsen M, Ringold GM, Pratt WB (1990). Evidence that the conserved region in the steroid binding domain of the glucocorticoid receptor is required for both optimal binding of hsp90 and protection from proteolytic cleavage. A two-site model for hsp90 binding to the steroid binding domain. *J Biol Chem* **265**: 12778-81.
- Huisinga KL, Pugh BF (2004). A genome-wide housekeeping role for TFIID and a highly regulated stress-related role for SAGA in *Saccharomyces cerevisiae*. *Mol Cell* **13**: 573-85.
- Illingworth RS, Bird AP (2009). CpG islands--'a rough guide'. *FEBS Lett* **583**: 1713-20.
- Inomata T, Kiuchi A, Yoshida T, Hisamatsu S, Takizawa A, Kashiwazaki N *et al* (2005). Hypervitaminosis A resulting in DNA aberration in fetal transgenic mice (Muta Mouse). *Mutat Res* **586**: 58-67.
- Inostroza JA, Mermelstein FH, Ha I, Lane WS, Reinberg D (1992). Dr1, a TATA-binding protein-associated phosphoprotein and inhibitor of class II gene transcription. *Cell* **70**: 477-89.
- Insinga A, Monestiroli S, Ronzoni S, Carbone R, Pearson M, Pruneri G *et al* (2004). Impairment of p53 acetylation, stability and function by an oncogenic transcription factor. *EMBO J* **23**: 1144-54.
- Ishibe T, Nakayama T, Aoyama T, Nakamura T, Toguchida J (2008). Neuronal differentiation of synovial sarcoma and its therapeutic application. *Clin Orthop Relat Res* **466**: 2147-55.
- Ito M, Yuan CX, Okano HJ, Darnell RB, Roeder RG (2000). Involvement of the TRAP220 component of the TRAP/SMCC coactivator complex in embryonic development and thyroid hormone action. *Mol Cell* **5**: 683-93.
- Jackson SP (1992). Regulating transcription factor activity by phosphorylation. *Trends Cell Biol* **2**: 104-8.
- Jacobson RH, Tjian R (1996). Transcription factor IIA: a structure with multiple functions. *Science* **272**: 827-8.
- Johansson S, Lind PM, Hakansson H, Oxlund H, Orberg J, Melhus H (2002). Subclinical hypervitaminosis A causes fragile bones in rats. *Bone* **31**: 685-9.
- John S, Johnson TA, Sung MH, Biddie SC, Trump S, Koch-Paiz CA *et al* (2009). Kinetic complexity of the global response to glucocorticoid receptor action. *Endocrinology* **150**: 1766-74.
- Johnson IT (1999). IARC handbooks of cancer prevention volume 2: carotenoids and volume 3: vitamin A. *Eur J Clin Nutr* **53**: 830-4.
- Jokic N, Ling YY, Ward RE, Michael-Titus AT, Priestley JV, Malaspina A (2007). Retinoid receptors in chronic degeneration of the spinal cord: observations in a rat model of amyotrophic lateral sclerosis. *J Neurochem* **103**: 1821-33.

- Jones DO, Cowell IG, Singh PB (2000). Mammalian chromodomain proteins: their role in genome organisation and expression. *Bioessays* **22**: 124-37.
- Kadonaga JT, Carner KR, Masiarz FR, Tjian R (1987). Isolation of cDNA encoding transcription factor Sp1 and functional analysis of the DNA binding domain. *Cell* **51**: 1079-90.
- Kaiser K, Meisterernst M (1996). The human general co-factors. *Trends Biochem Sci* **21**: 342-5.
- Kanai M, Raz A, Goodman DS (1968). Retinol-binding protein: the transport protein for vitamin A in human plasma. *J Clin Invest* **47**: 2025-44.
- Kastner B, Fischer N, Golas MM, Sander B, Dube P, Boehringer D *et al* (2008). GraFix: sample preparation for single-particle electron cryomicroscopy. *Nat Methods* **5**: 53-5.
- Kastner P, Mark M, Chambon P (1995). Nonsteroid nuclear receptors: what are genetic studies telling us about their role in real life? *Cell* **83**: 859-69.
- Kastner P, Mark M, Ghyselinck N, Krezel W, Dupe V, Grondona JM *et al* (1997). Genetic evidence that the retinoid signal is transduced by heterodimeric RXR/RAR functional units during mouse development. *Development* **124**: 313-26.
- Katayama S, Tomaru Y, Kasukawa T, Waki K, Nakanishi M, Nakamura M *et al* (2005). Antisense transcription in the mammalian transcriptome. *Science* **309**: 1564-6.
- Katoch SS, Moreland RS (1995). Agonist and membrane depolarization induced activation of MAP kinase in the swine carotid artery. *Am J Physiol* **269**: H222-9.
- Kaufmann J, Verrijzer CP, Shao J, Smale ST (1996). CIF, an essential cofactor for TFIID-dependent initiator function. *Genes Dev* **10**: 873-86.
- Keriel A, Sary A, Sarasin A, Rochette-Egly C, Egly JM (2002). XPD mutations prevent TFIIH-dependent transactivation by nuclear receptors and phosphorylation of RARalpha. *Cell* **109**: 125-35.
- Kim WG, Kim EY, Kim TY, Ryu JS, Hong SJ, Kim WB *et al* (2009). Redifferentiation therapy with 13-cis retinoic acids in radioiodine-resistant thyroid cancer. *Endocr J* **56**: 105-12.
- Kim YJ, Bjorklund S, Li Y, Sayre MH, Kornberg RD (1994). A multiprotein mediator of transcriptional activation and its interaction with the C-terminal repeat domain of RNA polymerase II. *Cell* **77**: 599-608.
- Kirschner DB, vom Baur E, Thibault C, Sanders SL, Gangloff YG, Davidson I *et al* (2002). Distinct mutations in yeast TAF(II)25 differentially affect the composition of TFIID and SAGA complexes as well as global gene expression patterns. *Mol Cell Biol* **22**: 3178-93.

- Kiyosawa H, Yamanaka I, Osato N, Kondo S, Hayashizaki Y (2003). Antisense transcripts with FANTOM2 clone set and their implications for gene regulation. *Genome Res* **13**: 1324-34.
- Koch F, Jourquin F, Ferrier P, Andrau JC (2008). Genome-wide RNA polymerase II: not genes only! *Trends Biochem Sci* **33**: 265-73.
- Kornberg RD, Lorch Y (1999). Chromatin-modifying and -remodeling complexes. *Curr Opin Genet Dev* **9**: 148-51.
- Kouskouti A, Scheer E, Staub A, Tora L, Talianidis I (2004). Gene-specific modulation of TAF10 function by SET9-mediated methylation. *Mol Cell* **14**: 175-82.
- Krezel W, Ghyselinck N, Samad TA, Dupe V, Kastner P, Borrelli E *et al* (1998). Impaired locomotion and dopamine signaling in retinoid receptor mutant mice. *Science* **279**: 863-7.
- Krust A, Green S, Argos P, Kumar V, Walter P, Bornert JM *et al* (1986). The chicken oestrogen receptor sequence: homology with v-erbA and the human oestrogen and glucocorticoid receptors. *EMBO J* **5**: 891-7.
- Kuras L, Kosa P, Mencia M, Struhl K (2000). TAF-Containing and TAF-independent forms of transcriptionally active TBP in vivo. *Science* **288**: 1244-8.
- Kutach AK, Kadonaga JT (2000). The downstream promoter element DPE appears to be as widely used as the TATA box in Drosophila core promoters. *Mol Cell Biol* **20**: 4754-64.
- Kwek KY, O'Gorman W, Akoulitchev A (2004). Transcription meets DNA repair at a PH domain. *Nat Struct Mol Biol* **11**: 588-9.
- Kwon H, Imbalzano AN, Khavari PA, Kingston RE, Green MR (1994). Nucleosome disruption and enhancement of activator binding by a human SW1/SNF complex. *Nature* **370**: 477-81.
- Kwong J, Lo KW, To KF, Teo PM, Johnson PJ, Huang DP (2002). Promoter hypermethylation of multiple genes in nasopharyngeal carcinoma. *Clin Cancer Res* **8**: 131-7.
- Lagrange T, Kapanidis AN, Tang H, Reinberg D, Ebright RH (1998). New core promoter element in RNA polymerase II-dependent transcription: sequence-specific DNA binding by transcription factor IIB. *Genes Dev* **12**: 34-44.
- Lariviere L, Geiger S, Hoepfner S, Rother S, Strasser K, Cramer P (2006). Structure and TBP binding of the Mediator head subcomplex Med8-Med18-Med20. *Nat Struct Mol Biol* **13**: 895-901.
- Lasnitzki I (1976). Reversal of methylcholanthrene-induced changes in mouse prostates in vitro by retinoic acid and its analogues. *Br J Cancer* **34**: 239-48.

- Laudet V (1997). Evolution of the nuclear receptor superfamily: early diversification from an ancestral orphan receptor. *J Mol Endocrinol* **19**: 207-26.
- Laudet V, Hanni C, Coll J, Catzeflis F, Stehelin D (1992). Evolution of the nuclear receptor gene superfamily. *EMBO J* **11**: 1003-13.
- Laurent BC, Treitel MA, Carlson M (1991). Functional interdependence of the yeast SNF2, SNF5, and SNF6 proteins in transcriptional activation. *Proc Natl Acad Sci U S A* **88**: 2687-91.
- Lausen J, Liu S, Fliegau M, Lubbert M, Werner MH (2006). ELA2 is regulated by hematopoietic transcription factors, but not repressed by AML1-ETO. *Oncogene* **25**: 1349-57.
- Le Douarin B, Nielsen AL, Garnier JM, Ichinose H, Jeanmougin F, Losson R *et al* (1996). A possible involvement of TIF1 alpha and TIF1 beta in the epigenetic control of transcription by nuclear receptors. *EMBO J* **15**: 6701-15.
- Lee DH, Gershenzon N, Gupta M, Ioshikhes IP, Reinberg D, Lewis BA (2005). Functional characterization of core promoter elements: the downstream core element is recognized by TAF1. *Mol Cell Biol* **25**: 9674-86.
- Lee JS, Newman RA, Lippman SM, Huber MH, Minor T, Raber MN *et al* (1993). Phase I evaluation of all-trans-retinoic acid in adults with solid tumors. *J Clin Oncol* **11**: 959-66.
- Lee S, Privalsky ML (2005). Heterodimers of retinoic acid receptors and thyroid hormone receptors display unique combinatorial regulatory properties. *Mol Endocrinol* **19**: 863-78.
- Lee TI, Young RA (2000). Transcription of eukaryotic protein-coding genes. *Annu Rev Genet* **34**: 77-137.
- Lefebvre P, Gaub MP, Tahayato A, Rochette-Egly C, Formstecher P (1995). Protein phosphatases 1 and 2A regulate the transcriptional and DNA binding activities of retinoic acid receptors. *J Biol Chem* **270**: 10806-16.
- Lefebvre P, Martin PJ, Flajollet S, Dedieu S, Billaut X, Lefebvre B (2005). Transcriptional activities of retinoic acid receptors. *Vitam Horm* **70**: 199-264.
- Leid M, Kastner P, Chambon P (1992). Multiplicity generates diversity in the retinoic acid signalling pathways. *Trends Biochem Sci* **17**: 427-33.
- Lemon B, Tjian R (2000). Orchestrated response: a symphony of transcription factors for gene control. *Genes Dev* **14**: 2551-69.
- Leng N, Grasso P, Deziel MR, Reichert LE, Jr. (1995). A synthetic peptide corresponding to glycoprotein hormone alpha subunit residues 32-46 inhibits gonadotropin binding to receptor. *Pept Res* **8**: 272-7.
- Leo C, Chen JD (2000). The SRC family of nuclear receptor coactivators. *Gene* **245**: 1-11.

- Leurent C, Sanders S, Ruhlmann C, Mallouh V, Weil PA, Kirschner DB *et al* (2002). Mapping histone fold TAFs within yeast TFIID. *EMBO J* **21**: 3424-33.
- Leurent C, Sanders SL, Demeny MA, Garbett KA, Ruhlmann C, Weil PA *et al* (2004). Mapping key functional sites within yeast TFIID. *EMBO J* **23**: 719-27.
- Lewis BA, Kim TK, Orkin SH (2000). A downstream element in the human beta-globin promoter: evidence of extended sequence-specific transcription factor IID contacts. *Proc Natl Acad Sci U S A* **97**: 7172-7.
- Li J, Ng EK, Ng YP, Wong CY, Yu J, Jin H *et al* (2009). Identification of retinoic acid-regulated nuclear matrix-associated protein as a novel regulator of gastric cancer. *Br J Cancer* **101**: 691-8.
- Li XY, Bhaumik SR, Green MR (2000). Distinct classes of yeast promoters revealed by differential TAF recruitment. *Science* **288**: 1242-4.
- Lim CY, Santoso B, Boulay T, Dong E, Ohler U, Kadonaga JT (2004). The MTE, a new core promoter element for transcription by RNA polymerase II. *Genes Dev* **18**: 1606-17.
- Lin RJ, Sternsdorf T, Tini M, Evans RM (2001). Transcriptional regulation in acute promyelocytic leukemia. *Oncogene* **20**: 7204-15.
- Lippman SM, Glisson BS, Kavanagh JJ, Lotan R, Hong WK, Paredes-Espinoza M *et al* (1993). Retinoic acid and interferon combination studies in human cancer. *Eur J Cancer* **29A Suppl 5**: S9-13.
- Liu T, Tee AE, Porro A, Smith SA, Dwarto T, Liu PY *et al* (2007). Activation of tissue transglutaminase transcription by histone deacetylase inhibition as a therapeutic approach for Myc oncogenesis. *Proc Natl Acad Sci U S A* **104**: 18682-7.
- Liu WL, Coleman RA, Ma E, Grob P, Yang JL, Zhang Y *et al* (2009). Structures of three distinct activator-TFIID complexes. *Genes Dev* **23**: 1510-21.
- Liu Y, Kung C, Fishburn J, Ansari AZ, Shokat KM, Hahn S (2004). Two cyclin-dependent kinases promote RNA polymerase II transcription and formation of the scaffold complex. *Mol Cell Biol* **24**: 1721-35.
- Lotan Y, Xu XC, Shalev M, Lotan R, Williams R, Wheeler TM *et al* (2000). Differential expression of nuclear retinoid receptors in normal and malignant prostates. *J Clin Oncol* **18**: 116-21.
- Lowenberg B, Griffin JD, Tallman MS (2003). Acute myeloid leukemia and acute promyelocytic leukemia. *Hematology Am Soc Hematol Educ Program*: 82-101.
- Lu SJ, Rowan S, Bani MR, Ben-David Y (1994). Retroviral integration within the Fli-2 locus results in inactivation of the erythroid transcription factor NF-E2 in Friend erythroleukemias:



- evidence that NF-E2 is essential for globin expression. *Proc Natl Acad Sci U S A* **91**: 8398-402.
- Luger K, Mader AW, Richmond RK, Sargent DF, Richmond TJ (1997). Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* **389**: 251-60.
- Ma X, Li-Ling J, Huang Q, Chen X, Hou L, Ma F (2009). Systematic analysis of alternative promoters correlated with alternative splicing in human genes. *Genomics* **93**: 420-5.
- Ma Y, Koza-Taylor PH, DiMattia DA, Hames L, Fu H, Dragnev KH *et al* (2003). Microarray analysis uncovers retinoid targets in human bronchial epithelial cells. *Oncogene* **22**: 4924-32.
- MacDonald PN, Ong DE (1988). Evidence for a lecithin-retinol acyltransferase activity in the rat small intestine. *J Biol Chem* **263**: 12478-82.
- Macias MJ, Wiesner S, Sudol M (2002). WW and SH3 domains, two different scaffolds to recognize proline-rich ligands. *FEBS Lett* **513**: 30-7.
- Maden M (2007). Retinoic acid in the development, regeneration and maintenance of the nervous system. *Nat Rev Neurosci* **8**: 755-65.
- Mader S, Leroy P, Chen JY, Chambon P (1993). Multiple parameters control the selectivity of nuclear receptors for their response elements. Selectivity and promiscuity in response element recognition by retinoic acid receptors and retinoid X receptors. *J Biol Chem* **268**: 591-600.
- Malik S, Gu W, Wu W, Qin J, Roeder RG (2000). The USA-derived transcriptional coactivator PC2 is a submodule of TRAP/SMCC and acts synergistically with other PCs. *Mol Cell* **5**: 753-60.
- Malik S, Roeder RG (2000). Transcriptional regulation through Mediator-like coactivators in yeast and metazoan cells. *Trends Biochem Sci* **25**: 277-83.
- Malik S, Roeder RG (2005). Dynamic regulation of pol II transcription by the mammalian Mediator complex. *Trends Biochem Sci* **30**: 256-63.
- Mangelsdorf DJ, Evans RM (1995). The RXR heterodimers and orphan receptors. *Cell* **83**: 841-50.
- Mantovani R (1998). A survey of 178 NF-Y binding CCAAT boxes. *Nucleic Acids Res* **26**: 1135-43.
- Mark M, Ghyselinck NB, Chambon P (2009). Function of retinoic acid receptors during embryonic development. *Nucl Recept Signal* **7**: e002.
- Martianov I, Fimia GM, Dierich A, Parvinen M, Sassone-Corsi P, Davidson I (2001). Late arrest of spermiogenesis and germ cell apoptosis in mice lacking the TBP-like TLF/TRF2 gene. *Mol Cell* **7**: 509-15.

- Mascrez B, Mark M, Dierich A, Ghyselinck NB, Kastner P, Chambon P (1998). The RXRalpha ligand-dependent activation function 2 (AF-2) is important for mouse development. *Development* **125**: 4691-707.
- Mazzarelli JM, Atkins GB, Geisberg JV, Ricciardi RP (1995). The viral oncoproteins Ad5 E1A, HPV16 E7 and SV40 TAg bind a common region of the TBP-associated factor-110. *Oncogene* **11**: 1859-64.
- McKenna NJ, Xu J, Nawaz Z, Tsai SY, Tsai MJ, O'Malley BW (1999). Nuclear receptor coactivators: multiple enzymes, multiple complexes, multiple functions. *J Steroid Biochem Mol Biol* **69**: 3-12.
- Meisterernst M, Roy AL, Lieu HM, Roeder RG (1991). Activation of class II gene transcription by regulatory factors is potentiated by a novel activity. *Cell* **66**: 981-93.
- Melhus H, Nilsson T, Peterson PA, Rask L (1991). Retinol-binding protein and transthyretin expressed in HeLa cells form a complex in the endoplasmic reticulum in both the absence and the presence of retinol. *Exp Cell Res* **197**: 119-24.
- Mengus G, Fadloun A, Kobi D, Thibault C, Perletti L, Michel I *et al* (2005). TAF4 inactivation in embryonic fibroblasts activates TGF beta signalling and autocrine growth. *EMBO J* **24**: 2753-67.
- Mengus G, May M, Carre L, Chambon P, Davidson I (1997). Human TAF(II)135 potentiates transcriptional activation by the AF-2s of the retinoic acid, vitamin D3, and thyroid hormone receptors in mammalian cells. *Genes Dev* **11**: 1381-95.
- Merino A, Madden KR, Lane WS, Champoux JJ, Reinberg D (1993). DNA topoisomerase I is involved in both repression and activation of transcription. *Nature* **365**: 227-32.
- Millan-Pacheco C, Capistran VM, Pastor N (2009). On the consequences of placing amino groups at the TBP-DNA interface. Does TATA really matter? *J Mol Recognit*.
- Miller WH, Jr. (1998). The emerging role of retinoids and retinoic acid metabolism blocking agents in the treatment of cancer. *Cancer* **83**: 1471-82.
- Minc E, Courvalin JC, Buendia B (2000). HP1gamma associates with euchromatin and heterochromatin in mammalian nuclei and chromosomes. *Cytogenet Cell Genet* **90**: 279-84.
- Mitsiou DJ, Stunnenberg HG (2003). p300 is involved in formation of the TBP-TFIIA-containing basal transcription complex, TAC. *EMBO J* **22**: 4501-11.
- Mittler G, Kremmer E, Timmers HT, Meisterernst M (2001). Novel critical role of a human Mediator complex for basal RNA polymerase II transcription. *EMBO Rep* **2**: 808-13.
- Mohan WS, Jr., Scheer E, Wendling O, Metzger D, Tora L (2003). TAF10 (TAF(II)30) is necessary for TFIID stability and early embryogenesis in mice. *Mol Cell Biol* **23**: 4307-18.

- Mohrmann L, Verrijzer CP (2005). Composition and functional specificity of SWI2/SNF2 class chromatin remodeling complexes. *Biochim Biophys Acta* **1681**: 59-73.
- Moon RC, Mehta RG (1990). Cancer chemoprevention by retinoids: animal models. *Methods Enzymol* **190**: 395-406.
- Moqtaderi Z, Bai Y, Poon D, Weil PA, Struhl K (1996). TBP-associated factors are not generally required for transcriptional activation in yeast. *Nature* **383**: 188-91.
- Moqtaderi Z, Keaveney M, Struhl K (1998). The histone H3-like TAF is broadly required for transcription in yeast. *Mol Cell* **2**: 675-82.
- Mousson F, Kolkman A, Pijnappel WW, Timmers HT, Heck AJ (2008). Quantitative proteomics reveals regulation of dynamic components within TATA-binding protein (TBP) transcription complexes. *Mol Cell Proteomics* **7**: 845-52.
- Mukherjee R, Davies PJ, Crombie DL, Bischoff ED, Cesario RM, Jow L *et al* (1997). Sensitization of diabetic and obese mice to insulin by retinoid X receptor agonists. *Nature* **386**: 407-10.
- Muller F, Tora L (2004). The multicoloured world of promoter recognition complexes. *EMBO J* **23**: 2-8.
- Myers LC, Kornberg RD (2000). Mediator of transcriptional regulation. *Annu Rev Biochem* **69**: 729-49.
- Naar AM, Lemon BD, Tjian R (2001). Transcriptional coactivator complexes. *Annu Rev Biochem* **70**: 475-501.
- Naar AM, Taatjes DJ, Zhai W, Nogales E, Tjian R (2002). Human CRSP interacts with RNA polymerase II CTD and adopts a specific CTD-bound conformation. *Genes Dev* **16**: 1339-44.
- Nagpal S, Friant S, Nakshatri H, Chambon P (1993). RARs and RXRs: evidence for two autonomous transactivation functions (AF-1 and AF-2) and heterodimerization in vivo. *EMBO J* **12**: 2349-60.
- Nagy Z, Tora L (2007). Distinct GCN5/PCAF-containing complexes function as co-activators and are involved in transcription factor and global histone acetylation. *Oncogene* **26**: 5341-57.
- Nervi C, Vollberg TM, George MD, Zelent A, Chambon P, Jetten AM (1991). Expression of nuclear retinoic acid receptors in normal tracheobronchial cells and in lung carcinoma cells. *Exp Cell Res* **195**: 163-70.
- Niederreither K, Dolle P (2008). Retinoic acid in development: towards an integrated view. *Nat Rev Genet* **9**: 541-53.
- Nikolov DB, Chen H, Halay ED, Usheva AA, Hisatake K, Lee DK *et al* (1995). Crystal structure of a TFIIB-TBP-TATA-element ternary complex. *Nature* **377**: 119-28.

- Nikolov DB, Hu SH, Lin J, Gasch A, Hoffmann A, Horikoshi M *et al* (1992). Crystal structure of TFIID TATA-box binding protein. *Nature* **360**: 40-6.
- Noy N (2000). Retinoid-binding proteins: mediators of retinoid action. *Biochem J* **348 Pt 3**: 481-95.
- Noy N, Slosberg E, Scarlata S (1992). Interactions of retinol with binding proteins: studies with retinol-binding protein and with transthyretin. *Biochemistry* **31**: 11118-24.
- O'Byrne SM, Wongsiriroj N, Libien J, Vogel S, Goldberg IJ, Baehr W *et al* (2005). Retinoid absorption and storage is impaired in mice lacking lecithin:retinol acyltransferase (LRAT). *J Biol Chem* **280**: 35647-57.
- O'Shea-Greenfield A, Smale ST (1992). Roles of TATA and initiator elements in determining the start site location and direction of RNA polymerase II transcription. *J Biol Chem* **267**: 6450.
- Oelgeschlager T, Chiang CM, Roeder RG (1996). Topology and reorganization of a human TFIID-promoter complex. *Nature* **382**: 735-8.
- Ohbayashi T, Makino Y, Tamura TA (1999). Identification of a mouse TBP-like protein (TLP) distantly related to the drosophila TBP-related factor. *Nucleic Acids Res* **27**: 750-5.
- Ohbayashi T, Shimada M, Nakadai T, Wada T, Handa H, Tamura T (2003). Vertebrate TBP-like protein (TLP/TRF2/TLF) stimulates TATA-less terminal deoxynucleotidyl transferase promoters in a transient reporter assay, and TFIIA-binding capacity of TLP is required for this function. *Nucleic Acids Res* **31**: 2127-33.
- Ohler U, Liao GC, Niemann H, Rubin GM (2002). Computational analysis of core promoters in the Drosophila genome. *Genome Biol* **3**: RESEARCH0087.
- Ong DE (1994). Cellular transport and metabolism of vitamin A: roles of the cellular retinoid-binding proteins. *Nutr Rev* **52**: S24-31.
- Ong DE, Chytil F (1978). Cellular retinol-binding protein from rat liver. Purification and characterization. *J Biol Chem* **253**: 828-32.
- Orphanides G, Lagrange T, Reinberg D (1996). The general transcription factors of RNA polymerase II. *Genes Dev* **10**: 2657-83.
- Ossipow V, Tassan JP, Nigg EA, Schibler U (1995). A mammalian RNA polymerase II holoenzyme containing all components required for promoter-specific transcription initiation. *Cell* **83**: 137-46.
- Owen GI, Zelent A (2000). Origins and evolutionary diversification of the nuclear receptor superfamily. *Cell Mol Life Sci* **57**: 809-27.

- Ozer J, Lezina LE, Ewing J, Audi S, Lieberman PM (1998). Association of transcription factor IIA with TATA binding protein is required for transcriptional activation of a subset of promoters and cell cycle progression in *Saccharomyces cerevisiae*. *Mol Cell Biol* **18**: 2559-70.
- Ozer J, Moore PA, Bolden AH, Lee A, Rosen CA, Lieberman PM (1994). Molecular cloning of the small (gamma) subunit of human TFIIA reveals functions critical for activated transcription. *Genes Dev* **8**: 2324-35.
- Pandolfi PP (2001). In vivo analysis of the molecular genetics of acute promyelocytic leukemia. *Oncogene* **20**: 5726-35.
- Papai G, Tripathi MK, Ruhlmann C, Werten S, Crucifix C, Weil PA *et al* (2009). Mapping the initiator binding Taf2 subunit in the structure of hydrated yeast TFIID. *Structure* **17**: 363-73.
- Parisotto M, Brodeur H, Bhat PV, Mader S (2006). [Retinoid metabolism and cancer]. *Med Sci (Paris)* **22**: 1101-6.
- Parvin JD, Shykind BM, Meyers RE, Kim J, Sharp PA (1994). Multiple sets of basal factors initiate transcription by RNA polymerase II. *J Biol Chem* **269**: 18414-21.
- Parvin JD, Young RA (1998). Regulatory targets in the RNA polymerase II holoenzyme. *Curr Opin Genet Dev* **8**: 565-70.
- Paulson M, Press C, Smith E, Tanese N, Levy DE (2002). IFN-Stimulated transcription through a TBP-free acetyltransferase complex escapes viral shutoff. *Nat Cell Biol* **4**: 140-7.
- Penczek PA, Yang C, Frank J, Spahn CM (2006). Estimation of variance in single-particle reconstruction using the bootstrap technique. *J Struct Biol* **154**: 168-83.
- Perissi V, Dasen JS, Kurokawa R, Wang Z, Korzus E, Rose DW *et al* (1999). Factor-specific modulation of CREB-binding protein acetyltransferase activity. *Proc Natl Acad Sci U S A* **96**: 3652-7.
- Peterson PA (1971). Studies on the interaction between prealbumin, retinol-binding protein, and vitamin A. *J Biol Chem* **246**: 44-9.
- Petti MC, Fazi F, Gentile M, Diverio D, De Fabritiis P, De Propris MS *et al* (2002). Complete remission through blast cell differentiation in PLZF/RARalpha-positive acute promyelocytic leukemia: in vitro and in vivo studies. *Blood* **100**: 1065-7.
- Pfahl M (1993). Nuclear receptor/AP-1 interaction. *Endocr Rev* **14**: 651-8.
- Philipsen S, Suske G (1999). A tale of three fingers: the family of mammalian Sp/XKLF transcription factors. *Nucleic Acids Res* **27**: 2991-3000.

- Picard D, Yamamoto KR (1987). Two signals mediate hormone-dependent nuclear localization of the glucocorticoid receptor. *EMBO J* **6**: 3333-40.
- Pointud JC, Mengus G, Brancorsini S, Monaco L, Parvinen M, Sassone-Corsi P *et al* (2003). The intracellular localisation of TAF7L, a paralogue of transcription factor TFIID subunit TAF7, is developmentally regulated during male germ-cell differentiation. *J Cell Sci* **116**: 1847-58.
- Pollock JL, Westervelt P, Kurichety AK, Pelicci PG, Grisolano JL, Ley TJ (1999). A bcr-3 isoform of RARalpha-PML potentiates the development of PML-RARalpha-driven acute promyelocytic leukemia. *Proc Natl Acad Sci U S A* **96**: 15103-8.
- Ponger L, Duret L, Mouchiroud D (2001). Determinants of CpG islands: expression in early embryo and isochore structure. *Genome Res* **11**: 1854-60.
- Pratt WB, Galigniana MD, Morishima Y, Murphy PJ (2004). Role of molecular chaperones in steroid receptor action. *Essays Biochem* **40**: 41-58.
- Pugh BF (1996). Mechanisms of transcription complex assembly. *Curr Opin Cell Biol* **8**: 303-11.
- Pugh BF (2000). Control of gene expression through regulation of the TATA-binding protein. *Gene* **255**: 1-14.
- Pugh BF, Tjian R (1990). Mechanism of transcriptional activation by Sp1: evidence for coactivators. *Cell* **61**: 1187-97.
- Qiu H, Zhang W, El-Naggar AK, Lippman SM, Lin P, Lotan R *et al* (1999). Loss of retinoic acid receptor-beta expression is an early event during esophageal carcinogenesis. *Am J Pathol* **155**: 1519-23.
- Quack M, Szafranski K, Rouvinen J, Carlberg C (1998). The role of the T-box for the function of the vitamin D receptor on different types of response elements. *Nucleic Acids Res* **26**: 5372-8.
- Rabenstein MD, Zhou S, Lis JT, Tjian R (1999). TATA box-binding protein (TBP)-related factor 2 (TRF2), a third member of the TBP family. *Proc Natl Acad Sci U S A* **96**: 4791-6.
- Ranish JA, Yudkovsky N, Hahn S (1999). Intermediates in formation and activity of the RNA polymerase II preinitiation complex: holoenzyme recruitment and a postrecruitment role for the TATA box and TFIIB. *Genes Dev* **13**: 49-63.
- Rao A, Luo C, Hogan PG (1997). Transcription factors of the NFAT family: regulation and function. *Annu Rev Immunol* **15**: 707-47.
- Redgrave TG (2004). Chylomicron metabolism. *Biochem Soc Trans* **32**: 79-82.
- Reese JC (2003). Basal transcription factors. *Curr Opin Genet Dev* **13**: 114-8.

- Reese JC, Zhang Z, Kurpad H (2000). Identification of a yeast transcription factor IID subunit, TSG2/TAF48. *J Biol Chem* **275**: 17391-8.
- Reik W, Walter J (2001). Genomic imprinting: parental influence on the genome. *Nat Rev Genet* **2**: 21-32.
- Renaud JP, Rochel N, Ruff M, Vivat V, Chambon P, Gronemeyer H *et al* (1995). Crystal structure of the RAR-gamma ligand-binding domain bound to all-trans retinoic acid. *Nature* **378**: 681-9.
- Rigtrup KM, Kakkad B, Ong DE (1994). Purification and partial characterization of a retinyl ester hydrolase from the brush border of rat small intestine mucosa: probable identity with brush border phospholipase B. *Biochemistry* **33**: 2661-6.
- Robert F, Douziech M, Forget D, Egly JM, Greenblatt J, Burton ZF *et al* (1998). Wrapping of promoter DNA around the RNA polymerase II initiation complex induced by TFIIF. *Mol Cell* **2**: 341-51.
- Rochette-Egly C, Adam S, Rossignol M, Egly JM, Chambon P (1997). Stimulation of RAR alpha activation function AF-1 through binding to the general transcription factor TFIID and phosphorylation by CDK7. *Cell* **90**: 97-107.
- Rochette-Egly C, Lutz Y, Saunders M, Scheuer I, Gaub MP, Chambon P (1991). Retinoic acid receptor gamma: specific immunodetection and phosphorylation. *J Cell Biol* **115**: 535-45.
- Rochette-Egly C, Oulad-Abdelghani M, Staub A, Pfister V, Scheuer I, Chambon P *et al* (1995). Phosphorylation of the retinoic acid receptor-alpha by protein kinase A. *Mol Endocrinol* **9**: 860-71.
- Roeder RG, Rutter WJ (1969). Multiple forms of DNA-dependent RNA polymerase in eukaryotic organisms. *Nature* **224**: 234-7.
- Roeder RG, Schwartz LB, Sklar VE (1976). Function, structure, and regulation of eukaryotic nuclear RNA polymerases. *Symp Soc Dev Biol*: 29-52.
- Romier C, James N, Birck C, Cavarelli J, Vivares C, Collart MA *et al* (2007). Crystal structure, biochemical and genetic characterization of yeast and *E. cuniculi* TAF(II)5 N-terminal domain: implications for TFIID assembly. *J Mol Biol* **368**: 1292-306.
- Ross CA (1997). Intranuclear neuronal inclusions: a common pathogenic mechanism for glutamine-repeat neurodegenerative diseases? *Neuron* **19**: 1147-50.
- Ruiz A, Winston A, Lim YH, Gilbert BA, Rando RR, Bok D (1999). Molecular and biochemical characterization of lecithin retinol acyltransferase. *J Biol Chem* **274**: 3834-41.

- Ryan KM, Phillips AC, Vousden KH (2001). Regulation and function of the p53 tumor suppressor protein. *Curr Opin Cell Biol* **13**: 332-7.
- Sadowski I, Ma J, Triezenberg S, Ptashne M (1988). GAL4-VP16 is an unusually potent transcriptional activator. *Nature* **335**: 563-4.
- Sandelin A, Carninci P, Lenhard B, Ponjavic J, Hayashizaki Y, Hume DA (2007). Mammalian RNA polymerase II core promoters: insights from genome-wide studies. *Nat Rev Genet* **8**: 424-36.
- Sanders SL, Garbett KA, Weil PA (2002). Molecular characterization of *Saccharomyces cerevisiae* TFIID. *Mol Cell Biol* **22**: 6000-13.
- Sanders SL, Weil PA (2000). Identification of two novel TAF subunits of the yeast *Saccharomyces cerevisiae* TFIID complex. *J Biol Chem* **275**: 13895-900.
- Sato S, Tomomori-Sato C, Parmely TJ, Florens L, Zybaylov B, Swanson SK *et al* (2004). A set of consensus mammalian mediator subunits identified by multidimensional protein identification technology. *Mol Cell* **14**: 685-91.
- Sawadogo M, Roeder RG (1985a). Factors involved in specific transcription by human RNA polymerase II: analysis by a rapid and quantitative in vitro assay. *Proc Natl Acad Sci U S A* **82**: 4394-8.
- Sawadogo M, Roeder RG (1985b). Interaction of a gene-specific transcription factor with the adenovirus major late promoter upstream of the TATA box region. *Cell* **43**: 165-75.
- Schaeffer V, Janody F, Loss C, Desplan C, Wimmer EA (1999). Bicoid functions without its TATA-binding protein-associated factor interaction domains. *Proc Natl Acad Sci U S A* **96**: 4461-6.
- Schena M, Freedman LP, Yamamoto KR (1989). Mutations in the glucocorticoid receptor zinc finger region that distinguish interdigitated DNA binding and transcriptional enhancement activities. *Genes Dev* **3**: 1590-601.
- schmutzler C, Brtko J, Winzer R, Jakobs TC, Meissner-Weigl J, Simon D *et al* (1998). Functional retinoid and thyroid hormone receptors in human thyroid-carcinoma cell lines and tissues. *Int J Cancer* **76**: 368-76.
- Schuh R, Aicher W, Gaul U, Cote S, Preiss A, Maier D *et al* (1986). A conserved family of nuclear proteins containing structural elements of the finger protein encoded by Kruppel, a *Drosophila* segmentation gene. *Cell* **47**: 1025-32.
- Sekiguchi T, Noguchi E, Hayashida T, Nakashima T, Toyoshima H, Nishimoto T *et al* (1996). D-type cyclin expression is decreased and p21 and p27 CDK inhibitor expression is increased when tsBN462 CCG1/TAFII250 mutant cells arrest in G1 at the restrictive temperature. *Genes Cells* **1**: 687-705.



- Selleck W, Howley R, Fang Q, Podolny V, Fried MG, Buratowski S *et al* (2001). A histone fold TAF octamer within the yeast TFIID transcriptional coactivator. *Nat Struct Biol* **8**: 695-700.
- Selvi RB, Kundu TK (2009). Reversible acetylation of chromatin: implication in regulation of gene expression, disease and therapeutics. *Biotechnol J* **4**: 375-90.
- Shao H, Revach M, Moshonov S, Tzuman Y, Gazit K, Albeck S *et al* (2005). Core promoter binding by histone-like TAF complexes. *Mol Cell Biol* **25**: 206-19.
- Shen WC, Bhaumik SR, Causton HC, Simon I, Zhu X, Jennings EG *et al* (2003). Systematic analysis of essential yeast TAFs in genome-wide transcription and preinitiation complex assembly. *EMBO J* **22**: 3395-402.
- Sherman-Baust CA, Weeraratna AT, Rangel LB, Pizer ES, Cho KR, Schwartz DR *et al* (2003). Remodeling of the extracellular matrix through overexpression of collagen VI contributes to cisplatin resistance in ovarian cancer cells. *Cancer Cell* **3**: 377-86.
- Shimohata T, Nakajima T, Yamada M, Uchida C, Onodera O, Naruse S *et al* (2000). Expanded polyglutamine stretches interact with TAFII130, interfering with CREB-dependent transcription. *Nat Genet* **26**: 29-36.
- Singh MV, Bland CE, Weil PA (2004). Molecular and genetic characterization of a Taf1p domain essential for yeast TFIID assembly. *Mol Cell Biol* **24**: 4929-42.
- Smale ST (2001). Core promoters: active contributors to combinatorial gene regulation. *Genes Dev* **15**: 2503-8.
- Smale ST, Baltimore D (1989). The "initiator" as a transcription control element. *Cell* **57**: 103-13.
- Song F, Smith JF, Kimura MT, Morrow AD, Matsuyama T, Nagase H *et al* (2005). Association of tissue-specific differentially methylated regions (TDMs) with differential gene expression. *Proc Natl Acad Sci U S A* **102**: 3336-41.
- Soprano DR, Qin P, Soprano KJ (2004). Retinoic acid receptors and cancers. *Annu Rev Nutr* **24**: 201-21.
- Sopta M, Carthew RW, Greenblatt J (1985). Isolation of three proteins that bind to mammalian RNA polymerase II. *J Biol Chem* **260**: 10353-60.
- Srinivas H, Xia D, Moore NL, Uray IP, Kim H, Ma L *et al* (2006). Akt phosphorylates and suppresses the transactivation of retinoic acid receptor alpha. *Biochem J* **395**: 653-62.
- Steinmetz AC, Renaud JP, Moras D (2001). Binding of ligands and activation of transcription by nuclear receptors. *Annu Rev Biophys Biomol Struct* **30**: 329-59.

- Stitzel ML, Pellettieri J, Seydoux G (2006). The *C. elegans* DYRK Kinase MBK-2 Marks Oocyte Proteins for Degradation in Response to Meiotic Maturation. *Curr Biol* **16**: 56-62.
- Stubbs GW, Saari JC, Futterman S (1979). 11-cis-Retinal-binding protein from bovine retina. Isolation and partial characterization. *J Biol Chem* **254**: 8529-33.
- Sullivan SA, Aravind L, Makalowska I, Baxevanis AD, Landsman D (2000). The histone database: a comprehensive WWW resource for histones and histone fold-containing proteins. *Nucleic Acids Res* **28**: 320-2.
- Sun F, Xie Q, Ma J, Yang S, Chen Q, Hong A (2009). Nuclear factor Y is required for basal activation and chromatin accessibility of fibroblast growth factor receptor 2 promoter in osteoblast-like cells. *J Biol Chem* **284**: 3136-47.
- Sun SY, Lotan R (2002). Retinoids and their receptors in cancer development and chemoprevention. *Crit Rev Oncol Hematol* **41**: 41-55.
- Suzuki Y, Tsunoda T, Sese J, Taira H, Mizushima-Sugano J, Hata H *et al* (2001). Identification and characterization of the potential promoter regions of 1031 kinds of human genes. *Genome Res* **11**: 677-84.
- Svensson S, Ostberg T, Jacobsson M, Norstrom C, Stefansson K, Hallen D *et al* (2003). Crystal structure of the heterodimeric complex of LXRA $\alpha$  and RXR $\beta$  ligand-binding domains in a fully agonistic conformation. *EMBO J* **22**: 4625-33.
- Tahayato A, Lefebvre P, Formstecher P, Dautrevaux M (1993). A protein kinase C-dependent activity modulates retinoic acid-induced transcription. *Mol Endocrinol* **7**: 1642-53.
- Takada R, Nakatani Y, Hoffmann A, Kokubo T, Hasegawa S, Roeder RG *et al* (1992). Identification of human TFIID components and direct interaction between a 250-kDa polypeptide and the TATA box-binding protein (TFIID tau). *Proc Natl Acad Sci U S A* **89**: 11809-13.
- Tan S, Garrett KP, Conaway RC, Conaway JW (1994). Cryptic DNA-binding domain in the C terminus of RNA polymerase II general transcription factor RAP30. *Proc Natl Acad Sci U S A* **91**: 9808-12.
- Tan S, Hunziker Y, Sargent DF, Richmond TJ (1996). Crystal structure of a yeast TFIIA/TBP/DNA complex. *Nature* **381**: 127-51.
- Tanaka A, Watanabe T, Iida Y, Hanaoka F, Ohkuma Y (2009). Central forkhead domain of human TFIIE beta plays a primary role in binding double-stranded DNA at transcription initiation. *Genes Cells* **14**: 395-405.
- Taneja R, Rochette-Egly C, Plassat JL, Penna L, Gaub MP, Chambon P (1997). Phosphorylation of activation functions AF-1 and AF-2 of RAR alpha and RAR gamma is indispensable for differentiation of F9 cells upon retinoic acid and cAMP treatment. *EMBO J* **16**: 6452-65.

- Tang Y, Liu DP, Liang CC (2002). Further understanding of the beta-globin locus regulation at the molecular level: looping or linking models? *Genes Cells* **7**: 889-900.
- Tatham MH, Geoffroy MC, Shen L, Plechanovova A, Hattersley N, Jaffray EG *et al* (2008). RNF4 is a poly-SUMO-specific E3 ubiquitin ligase required for arsenic-induced PML degradation. *Nat Cell Biol* **10**: 538-46.
- Thacher SM, Vasudevan J, Chandraratna RA (2000). Therapeutic applications for ligands of retinoid receptors. *Curr Pharm Des* **6**: 25-58.
- Thomas MC, Chiang CM (2006). The general transcription machinery and general cofactors. *Crit Rev Biochem Mol Biol* **41**: 105-78.
- Thompson CM, Koleske AJ, Chao DM, Young RA (1993). A multisubunit complex associated with the RNA polymerase II CTD and TATA-binding protein in yeast. *Cell* **73**: 1361-75.
- Thuault S, Gangloff YG, Kirchner J, Sanders S, Werten S, Romier C *et al* (2002). Functional analysis of the TFIID-specific yeast TAF4 (yTAF(II)48) reveals an unexpected organization of its histone-fold domain. *J Biol Chem* **277**: 45510-7.
- Timmers HT, Tora L (2005). SAGA unveiled. *Trends Biochem Sci* **30**: 7-10.
- Todone F, Brick P, Werner F, Weinzierl RO, Onesti S (2001). Structure of an archaeal homolog of the eukaryotic RNA polymerase II RPB4/RPB7 complex. *Mol Cell* **8**: 1137-43.
- Tontonoz P, Graves RA, Budavari AI, Erdjument-Bromage H, Lui M, Hu E *et al* (1994). Adipocyte-specific transcription factor ARF6 is a heterodimeric complex of two nuclear hormone receptors, PPAR gamma and RXR alpha. *Nucleic Acids Res* **22**: 5628-34.
- Tora L (2002). A unified nomenclature for TATA box binding protein (TBP)-associated factors (TAFs) involved in RNA polymerase II transcription. *Genes Dev* **16**: 673-5.
- Toulme F, Guerin M, Robichon N, Leng M, Rahmouni AR (1999). In vivo evidence for back and forth oscillations of the transcription elongation complex. *EMBO J* **18**: 5052-60.
- Tremblay A, Tremblay GB, Labrie F, Giguere V (1999). Ligand-independent recruitment of SRC-1 to estrogen receptor beta through phosphorylation of activation function AF-1. *Mol Cell* **3**: 513-9.
- Trussardi-Regnier A, Lavenus S, Gorisse MC, Dufer J (2009). Thalidomide alters nuclear architecture without ABCB1 gene modulation in drug-resistant myeloma cells. *Int J Oncol* **35**: 641-7.
- Tsai FT, Sigler PB (2000). Structural basis of preinitiation complex assembly on human pol II promoters. *EMBO J* **19**: 25-36.

- Tyler JK, Kadonaga JT (1999). The "dark side" of chromatin remodeling: repressive effects on transcription. *Cell* **99**: 443-6.
- Uesugi M, Nyanguile O, Lu H, Levine AJ, Verdine GL (1997). Induced alpha helix in the VP16 activation domain upon binding to a human TAF. *Science* **277**: 1310-3.
- Uesugi M, Verdine GL (1999). The alpha-helical FXXPhiPhi motif in p53: TAF interaction and discrimination by MDM2. *Proc Natl Acad Sci U S A* **96**: 14801-6.
- Umesono K, Giguere V, Glass CK, Rosenfeld MG, Evans RM (1988). Retinoic acid and thyroid hormone induce gene expression through a common responsive element. *Nature* **336**: 262-5.
- Vakoc CR, Letting DL, Gheldof N, Sawado T, Bender MA, Groudine M *et al* (2005). Proximity among distant regulatory elements at the beta-globin locus requires GATA-1 and FOG-1. *Mol Cell* **17**: 453-62.
- van Bennekum AM, Fisher EA, Blaner WS, Harrison EH (2000). Hydrolysis of retinyl esters by pancreatic triglyceride lipase. *Biochemistry* **39**: 4900-6.
- van Bennekum AM, Kako Y, Weinstock PH, Harrison EH, Deckelbaum RJ, Goldberg IJ *et al* (1999). Lipoprotein lipase expression level influences tissue clearance of chylomicron retinyl ester. *J Lipid Res* **40**: 565-74.
- van Laar JH, Berrevoets CA, Trapman J, Zegers ND, Brinkmann AO (1991). Hormone-dependent androgen receptor phosphorylation is accompanied by receptor transformation in human lymph node carcinoma of the prostate cells. *J Biol Chem* **266**: 3734-8.
- van Neerven S, Kampmann E, Mey J (2008). RAR/RXR and PPAR/RXR signaling in neurological and psychiatric diseases. *Prog Neurobiol* **85**: 433-51.
- Vassallo MF, Tanese N (2002). Isoform-specific interaction of HP1 with human TAFII130. *Proc Natl Acad Sci U S A* **99**: 5919-24.
- Veenstra GJ, Wolffe AP (2001). Gene-selective developmental roles of general transcription factors. *Trends Biochem Sci* **26**: 665-71.
- Verrecchia F, Vindevoghel L, Lechleider RJ, Uitto J, Roberts AB, Mauviel A (2001). Smad3/AP-1 interactions control transcriptional responses to TGF-beta in a promoter-specific manner. *Oncogene* **20**: 3332-40.
- Verrijzer CP, Chen JL, Yokomori K, Tjian R (1995). Binding of TAFs to core elements directs promoter selectivity by RNA polymerase II. *Cell* **81**: 1115-25.
- Verrijzer CP, Tjian R (1996). TAFs mediate transcriptional activation and promoter selectivity. *Trends Biochem Sci* **21**: 338-42.

- Villiger PM, Lotz M (1992). Differential expression of TGF beta isoforms by human articular chondrocytes in response to growth factors. *J Cell Physiol* **151**: 318-25.
- Voss AK, Thomas T, Petrou P, Anastassiadis K, Scholer H, Gruss P (2000). Taube nuss is a novel gene essential for the survival of pluripotent cells of early mouse embryos. *Development* **127**: 5449-61.
- Wade PA (2001). Transcriptional control at regulatory checkpoints by histone deacetylases: molecular connections between cancer and chromatin. *Hum Mol Genet* **10**: 693-8.
- Wang D, Bushnell DA, Huang X, Westover KD, Levitt M, Kornberg RD (2009). Structural basis of transcription: backtracked RNA polymerase II at 3.4 angstrom resolution. *Science* **324**: 1203-6.
- Wang L, Tang Y, Cole PA, Marmorstein R (2008). Structure and chemistry of the p300/CBP and Rtt109 histone acetyltransferases: implications for histone acetyltransferase evolution and function. *Curr Opin Struct Biol* **18**: 741-7.
- Wang PJ, Page DC (2002). Functional substitution for TAF(II)250 by a retroposed homolog that is expressed in human spermatogenesis. *Hum Mol Genet* **11**: 2341-6.
- Wang X, Truckses DM, Takada S, Matsumura T, Tanese N, Jacobson RH (2007). Conserved region I of human coactivator TAF4 binds to a short hydrophobic motif present in transcriptional regulators. *Proc Natl Acad Sci U S A* **104**: 7839-44.
- Warnmark A, Treuter E, Wright AP, Gustafsson JA (2003). Activation functions 1 and 2 of nuclear receptors: molecular strategies for transcriptional activation. *Mol Endocrinol* **17**: 1901-9.
- Waterston RH, Lindblad-Toh K, Birney E, Rogers J, Abril JF, Agarwal P *et al* (2002). Initial sequencing and comparative analysis of the mouse genome. *Nature* **420**: 520-62.
- Wei Y, Liu S, Lausen J, Woodrell C, Cho S, Biris N *et al* (2007). A TAF4-homology domain from the corepressor ETO is a docking platform for positive and negative regulators of transcription. *Nat Struct Mol Biol* **14**: 653-61.
- Weigel NL (1996). Steroid hormone receptors and their regulation by phosphorylation. *Biochem J* **319** ( Pt 3): 657-67.
- Weil PA, Luse DS, Segall J, Roeder RG (1979). Selective and accurate initiation of transcription at the Ad2 major late promoter in a soluble system dependent on purified RNA polymerase II and DNA. *Cell* **18**: 469-84.
- Werner A, Sayer JA (2009). Naturally occurring antisense RNA: function and mechanisms of action. *Curr Opin Nephrol Hypertens* **18**: 343-9.
- Werten S, Mitschler A, Romier C, Gangloff YG, Thuault S, Davidson I *et al* (2002). Crystal structure of a subcomplex of human transcription factor TFIID formed by TATA binding

- protein-associated factors hTAF4 (hTAF(II)135) and hTAF12 (hTAF(II)20). *J Biol Chem* **277**: 45502-9.
- Whitehouse I, Flaus A, Cairns BR, White MF, Workman JL, Owen-Hughes T (1999). Nucleosome mobilization catalysed by the yeast SWI/SNF complex. *Nature* **400**: 784-7.
- Widschwendter M, Berger J, Hermann M, Muller HM, Amberger A, Zeschnick M *et al* (2000). Methylation and silencing of the retinoic acid receptor-beta2 gene in breast cancer. *J Natl Cancer Inst* **92**: 826-32.
- Wieczorek E, Brand M, Jacq X, Tora L (1998). Function of TAF(II)-containing complex without TBP in transcription by RNA polymerase II. *Nature* **393**: 187-91.
- Wilhelm E, Pellay FX, Benecke A, Bell B (2008). TAF6delta controls apoptosis and gene expression in the absence of p53. *PLoS One* **3**: e2721.
- Willy PJ, Kobayashi R, Kadonaga JT (2000). A basal transcription factor that activates or represses transcription. *Science* **290**: 982-5.
- Wobbe CR, Struhl K (1990). Yeast and human TATA-binding proteins have nearly identical DNA sequence requirements for transcription in vitro. *Mol Cell Biol* **10**: 3859-67.
- Wolbach SB, Howe PR (1978). Nutrition Classics. The Journal of Experimental Medicine 42: 753-77, 1925. Tissue changes following deprivation of fat-soluble A vitamin. S. Burt Wolbach and Percy R. Howe. *Nutr Rev* **36**: 16-9.
- Wolf G (1998). M. Mori's definitive recognition of vitamin A deficiency and its cure in children. *Nutrition* **14**: 481-4.
- Wolf G, Carpenter KJ (1997). Early research into the vitamins: the work of Wilhelm Stepp. *J Nutr* **127**: 1255-9.
- Workman JL, Kingston RE (1998). Alteration of nucleosome structure as a mechanism of transcriptional regulation. *Annu Rev Biochem* **67**: 545-79.
- Woychik NA, Liao SM, Kolodziej PA, Young RA (1990). Subunits shared by eukaryotic nuclear RNA polymerases. *Genes Dev* **4**: 313-23.
- Wright KJ, Marr MT, 2nd, Tjian R (2006). TAF4 nucleates a core subcomplex of TFIID and mediates activated transcription from a TATA-less promoter. *Proc Natl Acad Sci U S A* **103**: 12347-52.
- Wright KJ, Tjian R (2009). Wnt signaling targets ETO coactivation domain of TAF4/TFIID in vivo. *Proc Natl Acad Sci U S A* **106**: 55-60.
- Wu PY, Ruhlmann C, Winston F, Schultz P (2004). Molecular architecture of the *S. cerevisiae* SAGA complex. *Mol Cell* **15**: 199-208.

- Wu SY, Thomas MC, Hou SY, Likhite V, Chiang CM (1999a). Isolation of mouse TFIID and functional characterization of TBP and TFIID in mediating estrogen receptor and chromatin transcription. *J Biol Chem* **274**: 23480-90.
- Wu SY, Zhou T, Chiang CM (2003). Human mediator enhances activator-facilitated recruitment of RNA polymerase II and promoter recognition by TATA-binding protein (TBP) independently of TBP-associated factors. *Mol Cell Biol* **23**: 6229-42.
- Wu WH, Pinto I, Chen BS, Hampsey M (1999b). Mutational analysis of yeast TFIIB. A functional relationship between Ssu72 and Sub1/Tsp1 defined by allele-specific interactions with TFIIB. *Genetics* **153**: 643-52.
- Wu Y, Lu Y, Hu Y, Li R (2005). Cyclic AMP-dependent modification of gonad-selective TAF(II)105 in a human ovarian granulosa cell line. *J Cell Biochem* **96**: 751-9.
- Xie X, Kokubo T, Cohen SL, Mirza UA, Hoffmann A, Chait BT *et al* (1996). Structural similarity between TAFs and the heterotetrameric core of the histone octamer. *Nature* **380**: 316-22.
- Xu J, Li Q (2003). Review of the in vivo functions of the p160 steroid receptor coactivator family. *Mol Endocrinol* **17**: 1681-92.
- Xu L, Glass CK, Rosenfeld MG (1999). Coactivator and corepressor complexes in nuclear receptor function. *Curr Opin Genet Dev* **9**: 140-7.
- Xu W, Chen H, Du K, Asahara H, Tini M, Emerson BM *et al* (2001). A transcriptional switch mediated by cofactor methylation. *Science* **294**: 2507-11.
- Xu XC, Zile MH, Lippman SM, Lee JS, Lee JJ, Hong WK *et al* (1995). Anti-retinoic acid (RA) antibody binding to human premalignant oral lesions, which occurs less frequently than binding to normal tissue, increases after 13-cis-RA treatment in vivo and is related to RA receptor beta expression. *Cancer Res* **55**: 5507-11.
- Yamagoe S, Kanno T, Kanno Y, Sasaki S, Siegel RM, Lenardo MJ *et al* (2003). Interaction of histone acetylases and deacetylases in vivo. *Mol Cell Biol* **23**: 1025-33.
- Yan M, Gralla JD (1997). Multiple ATP-dependent steps in RNA polymerase II promoter melting and initiation. *EMBO J* **16**: 7457-67.
- Yang C, Ng EG, Penczek PA (2005). Unified 3-D structure and projection orientation refinement using quasi-Newton algorithm. *J Struct Biol* **149**: 53-64.
- Yang XJ, Ogryzko VV, Nishikawa J, Howard BH, Nakatani Y (1996). A p300/CBP-associated factor that competes with the adenoviral oncoprotein E1A. *Nature* **382**: 319-24.
- Yelin R, Dahary D, Sorek R, Levanon EY, Goldstein O, Shoshan A *et al* (2003). Widespread occurrence of antisense transcription in the human genome. *Nat Biotechnol* **21**: 379-86.

Ylikomi T, Bocquel MT, Berry M, Gronemeyer H, Chambon P (1992). Cooperation of proto-signals for nuclear accumulation of estrogen and progesterone receptors. *EMBO J* **11**: 3681-94.

Young RA (1991). RNA polymerase II. *Annu Rev Biochem* **60**: 689-715.

Zanotto E, Hakkinen A, Teku G, Shen B, Ribeiro AS, Jacobs HT (2009). NF-Y influences directionality of transcription from the bidirectional Mrps12/Sarsm promoter in both mouse and human cells. *Biochim Biophys Acta* **1789**: 432-42.

Zawel L, Kumar KP, Reinberg D (1995). Recycling of the general transcription factors during RNA polymerase II transcription. *Genes Dev* **9**: 1479-90.

Zhang J, Kalkum M, Yamamura S, Chait BT, Roeder RG (2004). E protein silencing by the leukemogenic AML1-ETO fusion protein. *Science* **305**: 1286-9.

Zhou Q, Lieberman PM, Boyer TG, Berk AJ (1992). Holo-TFIID supports transcriptional stimulation by diverse activators and from a TATA-less promoter. *Genes Dev* **6**: 1964-74.

Zhou ZX, Sar M, Simental JA, Lane MV, Wilson EM (1994). A ligand-dependent bipartite nuclear targeting signal in the human androgen receptor. Requirement for the DNA-binding domain and modulation by NH<sub>2</sub>-terminal and carboxyl-terminal sequences. *J Biol Chem* **269**: 13115-23.

Zhu J, Koken MH, Quignon F, Chelbi-Alix MK, Degos L, Wang ZY *et al* (1997). Arsenic-induced PML targeting onto nuclear bodies: implications for the treatment of acute promyelocytic leukemia. *Proc Natl Acad Sci U S A* **94**: 3978-83.

Zhu J, Nasr R, Peres L, Riaucoux-Lormiere F, Honore N, Berthier C *et al* (2007). RXR is an essential component of the oncogenic PML/RARA complex in vivo. *Cancer Cell* **12**: 23-35.

Zofall M, Persinger J, Kassabov SR, Bartholomew B (2006). Chromatin remodeling by ISW2 and SWI/SNF requires DNA translocation inside the nucleosome. *Nat Struct Mol Biol* **13**: 339-46.

Zurita M, Merino C (2003). The transcriptional complexity of the TFIID complex. *Trends Genet* **19**: 578-84.