
THESE

Présentée à la faculté des
Sciences de la Vie
De l'Université Louis Pasteur
STRASBOURG-I

Pour l'obtention de titre de :

**DOCTEUR DE L'UNIVERSITE
LOUIS PASTEUR DE STRASBOURG**

Specialité : Aspects moléculaires et cellulaires de la biologie

Par

Anass JAWHARI

**Etude structure-fonction du facteur
de transcription / réparation TFIIH**

Soutenue le 4 Décembre 2003 devant le Jury composé de :

Pr Frédéric Dardel	Rapporteur externe
Dr Bertrand Séraphin	Rapporteur externe
Pr Eric Westhof	Rapporteur interne
Dr Dino Moras	Directeur de Thèse
Dr Jean Marc Egly	Examineur
Dr Arnaud Poterszman	Examineur

Liste des abréviations	1
Avant-propos	3
Chapitre I- Régulation de la transcription	6
1- La transcription de base	6
1.1- Les promoteurs de l'ARN polymérase II	7
1.2- L'assemblage séquentiel des facteurs de transcription	8
1.2.1- Le facteur TFIID	11
1.2.2- Le facteur TFIIB	12
1.2.3- Le facteur TFIIF et l'ARN polymérase II	12
1.2.4- Le facteur TFIIE et TFIIH	14
1.3- Modèle de l'holoenzyme	15
1.4- L'élongation	16
2- Facteurs de transcription spécifiques	17
2.1- Activateurs transcriptionnels	18
2.2- Coactivateurs transcriptionnels	19
2.2.1- TFIIA	19
2.2.2- Les TAFs	20
2.2.3- Le médiateur	20
3- Chromatine et Transcription	22
3.1- Structure de la chromatine	22
3.2- Remodelage de la chromatine par modifications covalentes	25
3.2.1- Acétylation des histones	26
3.2.2- Autres modifications des histones	27
3.3- Remodelage de la chromatine par modifications non covalente	28
Chapitre II- Transcription à la lumière de la structure	30
1- La structure du nucléosome	30
2- Le promoteur avant l'ARN polymérase II	32
2.1- TBP	32
2.2- Complexe TBP/ boîte TATA	32
2.3- Complexe TBP/ADN/TFIIB	34
2.4- Complexe TBP/ADN/TFIIA	36
2.5- Complexe TBP/TAF230	38
2.6- Complexe TBP/ADN/NC2	38
3- Structure de l'ARN polymérase	41
3.1- Structure du core ARN polymérase II	41
3.2- Conservation structurale	43
3.3- Structure de l'ARN polymérase II entière	45
3.4- Elongation de l'ARN polymérase II	47
4- Le médiateur	48
5- Structures moléculaires de facteurs de transcription multiprotéiques	51
5.1- Le facteur TFIID	51
5.2- Le facteur TFIIH	54
5.3- Le facteur TFIIE	56
Chapitre III- TFIIH, facteur multiprotéique	57
1- Les sous unités du core TFIIH	57
1.1- L'hélicase p89/XPB	57
1.2- La sous unité p62	59
1.3- La sous unité p52	62
1.4- La sous unité p44	62
1.5- La sous unité p34	64
2- Les sous unités du CAK	64
2.1- La sous-unité Cdk7	64
2.2- La sous-unité Cyclin H	66

2.3-La sous-unité Mat1	67
3- La sous-unité p80/XPD	68
Chapitre IV- TFIIH à l'interface de processus cellulaires fondamentaux	71
1- Rôle de TFIIH dans la transcription de base	71
1.1- Rôle de TFIIH dans l'ouverture du promoteur	71
1.2- Rôle de TFIIH dans le départ du promoteur	72
2- Rôle de TFIIH dans la transcription activée	72
3- Rôle de TFIIH dans la réparation par Excision-Resynthèse de Nucléotides	73
3.1- Généralités	73
3.2- Le mécanisme du NER	75
3.2.1- La reconnaissance de la lésion	75
3.2.2- L'ouverture autour de la lésion	78
3.2.3- La double incision	78
3.2.4- La resynthèse et la ligation d'ADN	78
3.3- Le rôle de TFIIH dans le mécanisme NER	79
3.4- TFIIH et la réparation couplée à la transcription	80
4- Rôle de TFIIH dans le cycle cellulaire	83
5- Rôle de TFIIH dans l'apoptose	84
6- TFIIH et d'autres phénomènes	84
7- Maladies génétiques rares associées à TFIIH	85
7.1- Le Xeroderma pigmentosum	87
7.2- Le syndrome de Cockayne	90
7.3- La Trichothiodystrophie	91
Chapitre V- résultats	92
Article 1	92
Article 2	94
Article 3	96
Article 4	98
Article 5	100
Chapitre VI- Etude cristallographique du complexe p34/p44 humain	102
I- Expression et Purification du complexe p34/p44	102
1- Définition du complexe stable	102
2- Expression	102
3- Purification	103
3.1- Chromatographie d'affinité	104
3.2- Chromatographie sur colonne d'interactions hydrophobes	105
3.3- Chromatographie de filtration sur gel	106
3.4- Rendement de purification	106
4- Caractérisation du complexe p34N/p44C	107
II- Cristallisation du complexe p34/p44	109
1- Généralités	109
2- Introduction à la cristallisation	110
2.1- Les étapes de la cristallisation	110
2.1.1- La sursaturation	110
2.1.2- La nucléation	110
2.1.3- La croissance cristalline	110
2.2- Les paramètres influençant la cristallisation	111
2.2.1- Les paramètres physico-chimiques	111
2.2.2- Les paramètres liés à la macromolécule	112
2.3- Les techniques de cristallisation	112
2.3.1- La cristallisation en batch	112
2.3.2- La cristallisation par diffusion de vapeur	113
2.3.3- La cristallisation par dialyse	114
2.3.4- L'ensemencement	114
2.4- Criblage de conditions de cristallisation	115

2.5-La cryoprotection.....	115
3-Cristallisation de complexe p34/p44.....	116
3.1-Cristallogénèse.....	116
3.2-Caractérisation des cristaux.....	118
3.3- Congélation.....	118
3.4- Collecte de données.....	119
3.5- phasage.....	122
4- Autres solutions.....	123
4.1- Mutagenèse.....	123
4.2- Purification.....	123
4.3- Vers un complexe plus grand.....	126
4.4- Changement d'espèces.....	129
Chapitre VII- Autres résultats.....	130
1- Architecture du core-TFIIH.....	130
2- Fonction du domaine PH/PTB de p62.....	131
3- Autre fonction du PH/PTB de p62.....	132
4- Etude de TFIIE.....	133
4.1-Protéolyse ménagée de TFIIE.....	133
4.2-Interaction entre TFIIE et TFIIH.....	134
Chapitre VIII- Discussion et perspectives.....	137
1-Expression et purification de TFIIH.....	137
2- Etude architecturale de TFIIH.....	140
3- p52 régule la fonction de XPB au sein de TFIIH.....	143
4-Etude structure-fonction de p62.....	147
5-Etude du facteur TFIIE.....	150
Annexes.....	151
Références bibliographiques.....	158

Remerciements

Avant tout j'aimerais remercier les Professeurs Frédéric Dardel, Bertrand Séraphin et Eric Westhof pour avoir accepté de juger ce travail de thèse.

Je suis particulièrement reconnaissant au Pr Dino Moras pour m'avoir donné la possibilité d'effectuer ma thèse dans un environnement scientifique exceptionnel. Je le remercie pour son soutien et son dynamisme.

Arnaud Poterszman a supervisé la progression de ma recherche depuis le DEA. Je le remercie vivement pour m'avoir fait confiance durant ces années et pour m'avoir fait profité de ses connaissances.

Je remercie Jean Marc Egly pour son intérêt pour ce travail, et pour les discussions et les précieux conseils dont il m'a fait bénéficier.

Merci à Jean Claude Thierry pour son intérêt constant pour TFIIH.

Une pensée très amicale pour tous les membres du laboratoire de génomique et biologie structurale, et plus particulièrement à mes compagnons de paillasse.

Une pensée à Alex et son riz collant, à Isabelle et son intérêt pour les randonnées de 4h, à Fabrice et les parties de squash très très.. ATP dépendantes, à Thomas et son âme d'artiste, à Gilbert et les petites virées à la capitale, à Natacha et le prince charmant, à Denis et sa sympathie, à Holger et Ralf et leur éternel optimisme, à Sylvia et son sourire attachant, à Sabrina et son gâteau au chocolat, à Sarah et sa bonne humeur, à Florence et sa sympathie, aux « stevenins » et leur potins, à Pascal et ses pensées, à Natacha et sa gentillesse, à Saraswathi et sa philosophie Zen, à Marat et sa modestie et à André et son high throughput.

Merci à la première génération de thésards « TFIIH », Valérie, Sébastien et Didier pour m'avoir transmis le virus baculo.

Sans transition, un grand Merci à Jean Luc Weikert et Isabelle Kolb-Cheynel pour les litres de baculovirus et les milliards de cellules d'insectes.

Le présent travail est le fruit d'une étroite collaboration avec Noëlle Potier et Stéphanie Boussert, pour la spectrométrie de masse, avec l'équipe de Jean Marc Egly pour les fonctions de TFIIH, avec Patrick Schultz et Muriel Uhring pour la microscopie et avec Bruno Kieffer et Virginie Gervais pour la RMN.

Merci à Anne Catherine Dock-Bregeon pour toutes nos discussions enrichissantes. Merci à Marc Ruff pour son aide et en particulier à propos « du *soaking* » des dérivés lourds.

Merci à Catherine Birck et Jean Pierre Samama pour m'avoir introduit dans le cercle fermé de l'Ultracentrifugation analytique.

Merci à Olivier Cuniculi et Odile Gardia pour m'avoir sensibilisé à la Bio-info.

Merci à Hatim et Didier pour leur lecture critique de ce manuscrit.

Merci également à Odile Lemble, Anne Ney et Christianne Rether pour leur aide et leur disponibilité.

Merci à mes collègues du 4ème étage, pour nos discussions, et plus particulièrement Sandy Dubaele, Jean Phillippe Lainé et Frédéric Coin.

Une pensée très amicale à la bande du DEA, Romain, Sandy, Bertrand, Omar, Guillaume et Alexi, au souvenir de soirées bien sympathiques.

Merci à Serge Vicair pour le séquençage d'ADN, Adrien Staub pour le séquençage N-terminal de protéine, Pascal Ebeling pour la synthèse de peptides et Franck Ruffenach et Edouard Troesch pour la synthèse d'oligonucléotides.

Merci à tous ceux qui rendent ce lieu très attachant malgré les difficultés de la recherche.

A tous ceux que j'ai rencontré à l'Institut et qui ont fait en sorte que cette période de thèse a été pour moi un moment très riche.

Ce travail a été en partie financé par le Centre National de la Recherche Scientifique et par l'Association de recherche sur le Cancer.

Et le meilleur pour la fin, je voudrais remercier mes parents pour m'avoir donné envie d'apprendre, pour avoir accepté mes orientations et pour m'avoir accompagné tout le long de ces années d'études. Merci à Yassir et Hatim pour leur aide ainsi qu' à Iphigénie pour son soutien et son encouragement sans faille.

Liste des abréviations

Å	Angström (1 Å = 10 ⁻¹⁰ m)
ADN	Acide désoxyribonucléique
ARN	Acide ribonucléique
ARNm	ARN messenger
ATP	Adénosine Triphosphate
CAK	Kinase activant les Cdk (<i>Cdk-Activating Kinase</i>)
Cdk	Kinase dépendante des cyclines (<i>Cyclin-Dependant Kinase</i>)
CHAPS	3-((3-cholamidopropyl) dimethylamonio)-1-propane sulfonate
CS	Syndrome de Cockayne
CTD	Domaine carboxy-terminal de l'ARN polymérase II
DTT	Dithiothreitol
ESRF	European Synchrotron Radiation Facility (Grenoble, France)
GGR	Réparation globale du génome (<i>Genome Global Repair</i>)
IPTG	Isopropyl-β-dithiogalactopyranoside
Inr	Elément initiateur
kDa	kiloDalton
MAT1	Ménage à trois 1
NER	Réparation par excision resynthèse de nucléotides (<i>Nucleotide Excision Repair</i>)
PEG	polyéthylène-glycol
RMN	Résonance magnétique nucléaire
RX	Rayons X

SDS	Sodium dodécyl sulfate
SDS-PAGE	Gel d'électrophorèse en conditions dénaturantes
TAF _{II}	Facteur de classe II associé à TBP (<i>TBP Associated Factor</i>)
TBP	Protéine de liaison à la boîte TATA (<i>TATA Binding Protein</i>)
TCR	Réparation couplée à la transcription (<i>Transcription Coupled Repair</i>)
TFIIH	facteur de transcription de classe II H
TTD	Trichotiodystrophie
UV	Rayonnement ultraviolet
XP	Xeroderma pigmentosum

Avant-propos

Le facteur TFIIH, formé de deux sous complexes (le CAK et le core-TFIIH), a fait l'objet d'études approfondies depuis une dizaine d'années. Ces études ont montré son implication dans différents processus cellulaires fondamentaux, notamment la transcription des gènes de classe I et de classe II, la réparation de l'ADN par excision re-synthèse de nucléotides et la régulation du cycle cellulaire. Une collaboration entre l'équipe du Dr Dino Moras et celle du Dr Jean Marc Egly (IGBMC) a été engagée dans le but de pousser la compréhension du rôle biologique de ce facteur au niveau atomique.

Dans ce sens, trois approches ont été envisagées. La cristallographie aux rayons X, la résonance magnétique nucléaire (RMN) et la microscopie électronique. Ces trois approches ont démarré en collaboration interne avec Arnaud Poterszman (Cristallographie), Bruno Kieffer (RMN) et Patrick Schultz (Microscopie électronique) et ont permis chacune d'apporter leur contribution. En effet, la cristallographie a permis de résoudre la structure de la cycline H (une sous-unité du CAK) (Andersen et al., 1997; Andersen et al., 1996). La RMN a permis de résoudre la structure du domaine *ring-finger* de MAT1 (une sous unité du CAK) et la structure du domaine carboxy-terminal de p44 (Fribourg et al., 2000; Gervais et al., 2001). La microscopie électronique, quant à elle, a permis de résoudre la structure moléculaire de TFIIH entier (Schultz et al., 2000).

Mon travail de thèse s'inscrit dans cette approche multi-disciplinaire.

Au cours de cette thèse, je me suis focalisé sur l'étude structure-fonction du sous complexe core-TFIIH :

La première étape de cette étude a été la mise en place d'un protocole de purification basé sur l'affinité au peptide Flag et l'expression de sous complexes de TFIIH dans les cellules d'insectes infectées par des baculovirus recombinants. Elle a permis d'initier l'étude de l'architecture du core-TFIIH. Ce système a été utilisé pour l'étude de la sous-unité p52 et m'a été utile pour montrer que cette dernière interagit directement avec XPB, conditionnant son ancrage au core-TFIIH, et régulant ainsi son

rôle dans la transcription et dans la réparation d'ADN. Je me suis ensuite intéressé à la sous-unité p62. J'ai pu montrer qu'elle est organisée en trois domaines dont deux ont été caractérisés en protéolyse ménagée, spectrométrie de masse, fluorescence et résonance magnétique nucléaire. La structure du domaine N-terminal de p62 a été résolue par RMN (Dr Gervais Virginie et Dr Valérie Lamour) et a révélé un repliement du type PTB/PH (*Phosphotyrosine Binding Protein/Pleckstrin Homology*). Mon travail à ce niveau a été d'établir une relation structure-fonction de ce domaine. J'ai pu lui assigner un rôle dans l'activité de réparation par excision resynthèse de nucléotide et dans une interaction spécifique avec l'endonucléase XPG. Une étude par protéolyse ménagée m'a permis de délimiter un domaine stable de 8 kDa dans la partie médiane de p62, correspondant à une région très conservée, contenant un motif FW (Phe Trp) et capable d'interagir avec l'hélicase XPD au sein de TFIIF. La structure du domaine médian est en cours de résolution par RMN et devrait nous permettre de mieux comprendre le mode de reconnaissance entre p62 et XPD.

Une étude architecturale du complexe p34/p44 a servi à établir que le domaine amino-terminal de p34 contacte le domaine carboxy-terminal de p44 (Dr Sébastien Fribourg). Ma contribution a été d'exprimer, purifier et cribler les conditions de cristallisation de ce complexe. J'ai réussi à produire de manière reproductible des cristaux qui diffractent à 3.3 Å dont j'ai pu enregistrer un certain nombre de jeux de données. La mauvaise qualité de diffraction et la difficulté d'obtention de phases, m'ont empêché d'aller plus loin dans la résolution de la structure.

Par ailleurs, une étude de pontage chimique et d'ultracentrifugation analytique du facteur de transcription TFIIF m'a permis de montrer qu'il se comporte comme dimère $\alpha\beta$ en solution, contrairement à ce qui est établi depuis des années. Sa structure, déterminée par microscopie électronique à 17 Å, montre une organisation en domaines distincts.

Avant de développer les résultats obtenus et le travail entrepris au cours de cette thèse, je me propose d'essayer de replacer ce travail dans son contexte en décrivant d'abord les mécanismes complexes de la régulation de la transcription, et en exposant l'apport de la biologie structurale pour la compréhension de ces mécanismes. Je présenterai ensuite les sous-unités du facteur TFIID ainsi que leur fonctions établies et j'essayerai enfin de montrer à quel point ce facteur est versatile et peut être à l'interface de plusieurs processus fondamentaux.

Chapitre I - Régulation de la transcription

1- La transcription de base

Les réponses physiologiques d'un organisme à son environnement sont associées à des mécanismes complexes de régulation génétique. La régulation de l'expression des gènes peut s'opérer en trois étapes essentielles : la réplication, la transcription et la traduction de l'information génétique contenue dans l'ADN. Parmi l'ensemble de ces régulations la transcription occupe une place primordiale, et sa régulation est une des plus finement élaborée.

Chez les organismes eucaryotes, il existe trois ARN polymérases en charge de transcrire l'ensemble des gènes de la cellule. En effet, l'ARN polymérase I transcrit les gènes codant pour les ARN ribosomiaux (18S, 5.8S, 28S). Les ARN de transfert ainsi que l'ARN ribosomal 5S et l'ARN U6 sont transcrits par l'ARN polymérase III. Et enfin, les gènes codant pour les protéines sont exclusivement transcrits par l'ARN polymérase II (ARN Pol II) en ARN messager (ARNm).

La régulation de l'expression des gènes au niveau transcriptionnel s'effectue plus particulièrement à l'étape de démarrage ou d'initiation. Les ARN polymérases à elles seules ne sont pas capables d'initier la transcription. Elles ont besoin d'un ensemble de protéines appelées facteurs généraux de transcription, qui se lient au promoteur avant le démarrage de la transcription. L'assemblage de ces facteurs au niveau du promoteur permet différentes possibilités de régulation de l'initiation de la transcription, en réponse aux signaux, et un grand nombre de protéines régulatrices agissent à ce niveau. D'autre part, il existe des protéines régulatrices pouvant reconnaître spécifiquement certaines séquences d'ADN et se fixer au niveau de séquences spécifiques, soit à quelques paires de bases du site d'initiation de la transcription (séquences proximales), soit à des dizaines de kilobases du promoteur (séquences distales). Chaque polymérase reconnaît un type de promoteur donné. Au cours de cet exposé, nous aborderons essentiellement l'initiation de la transcription

des gènes codant pour les protéines (gènes de classe II) et traiterons de la régulation par des activateurs ou répresseurs de ce même type de gènes.

La transcription s'effectue en trois étapes distinctes :

L'initiation

Elle implique la reconnaissance du promoteur et la fixation de l'ARN polymérase ainsi que des facteurs généraux de transcription. La double hélice d'ADN est ouverte entre la position -9 et +2 de part et d'autre du site d'initiation de la transcription (+1) pour former la bulle transcriptionnelle. Ceci permet la synthèse de la première liaison phospho-diester, l'appariement des ribonucléotides au brin matrice et l'initiation de la transcription.

L'élongation (ou allongement)

L'ARN polymérase se déplace le long de l'ADN en séparant les deux brins et en incorporant davantage de ribonucléotides.

La terminaison

Elle implique la reconnaissance par l'ARN polymérase d'une séquence de terminaison (*terminator*). Lorsque le dernier ribonucléotide a été incorporé, la polymérisation s'arrête et le complexe de transcription est dissocié libérant ainsi l'ARN néo-synthétisé et la machinerie de transcription.

1.1- Les promoteurs de l'ARN polymérase II

Les promoteurs des gènes transcrits par l'ARN polymérase II comportent un promoteur minimal correspondant aux séquences reconnues par les facteurs généraux de transcription et différents éléments de régulation proximaux (situés près du site d'initiation) ou distaux (plus éloignés du site d'initiation), spécifiques à chaque gène. Les éléments du promoteur minimal sont définis comme les séquences d'ADN nécessaires et suffisantes pour permettre un démarrage de la transcription par l'ARN Pol II et les facteurs généraux de la transcription *in vitro*.

L'élément principal est la boîte TATA (séquence consensus TATAa/tAa/t). La position du site de démarrage de la transcription se situe le plus souvent de 25 à 30 bp en aval de la boîte TATA, qui est le site de fixation de la protéine TBP (TATA Binding Protein), sous unité du facteur général de transcription TFIID.

Le deuxième élément du promoteur minimal est l'initiateur. Il s'agit de séquences riches en pyrimidines (séquences consensus YYANt/aYY) localisées au niveau du site de démarrage de la transcription (Javahery et al., 1994).

Il existe néanmoins des promoteurs à boîte TATA dépourvus d'initiateur. En l'absence de boîte TATA, l'initiateur joue un rôle particulièrement important. En effet, l'ARN polymérase II ainsi que les facteurs TFIID et TFIIB sont capables de reconnaître les éléments initiateurs, forment le complexe de préinitiation et permettent le démarrage de la transcription (Hampsey, 1998).

Un troisième élément du promoteur minimal de séquence faiblement conservée, le DPE (*downstream promoter element*) est localisé environ 30 bp en aval du site de démarrage de la transcription du promoteur dépourvu de boîte TATA. Le DPE fonctionne de manière synergique avec l'initiateur et est reconnu par le facteur TFIID (Burke and Kadonaga, 1997). La séquence BRE (*TFIIB Responsive Element*) est une séquence riche en guanine, située en amont de la boîte TATA, et reconnu spécifiquement par le facteur général de transcription TFIIB (Lagrange et al., 1998). En résumé, il existe des promoteurs de boîte TATA avec ou sans initiateur, des promoteurs avec initiateur mais sans boîte TATA, accompagnés ou non d'un élément DPE.

1.2- L'assemblage séquentiel des facteurs de transcription

La première étape de la transcription correspond à la formation au niveau du promoteur minimal d'un complexe de très grande taille (4 MDa) appelé complexe de préinitiation (PIC). Les étapes de l'assemblage du PIC ont été reconstituées *in vitro* avec des facteurs purifiés, ce qui a permis de proposer un modèle d'assemblage séquentiel, selon lequel les différents facteurs de transcription de base s'assemblent dans un ordre précis autour du promoteur (Figure 1A et table 1). Ces expériences ont fourni des informations importantes sur les interactions entre les différents facteurs au sein du PIC (Hampsey, 1998).

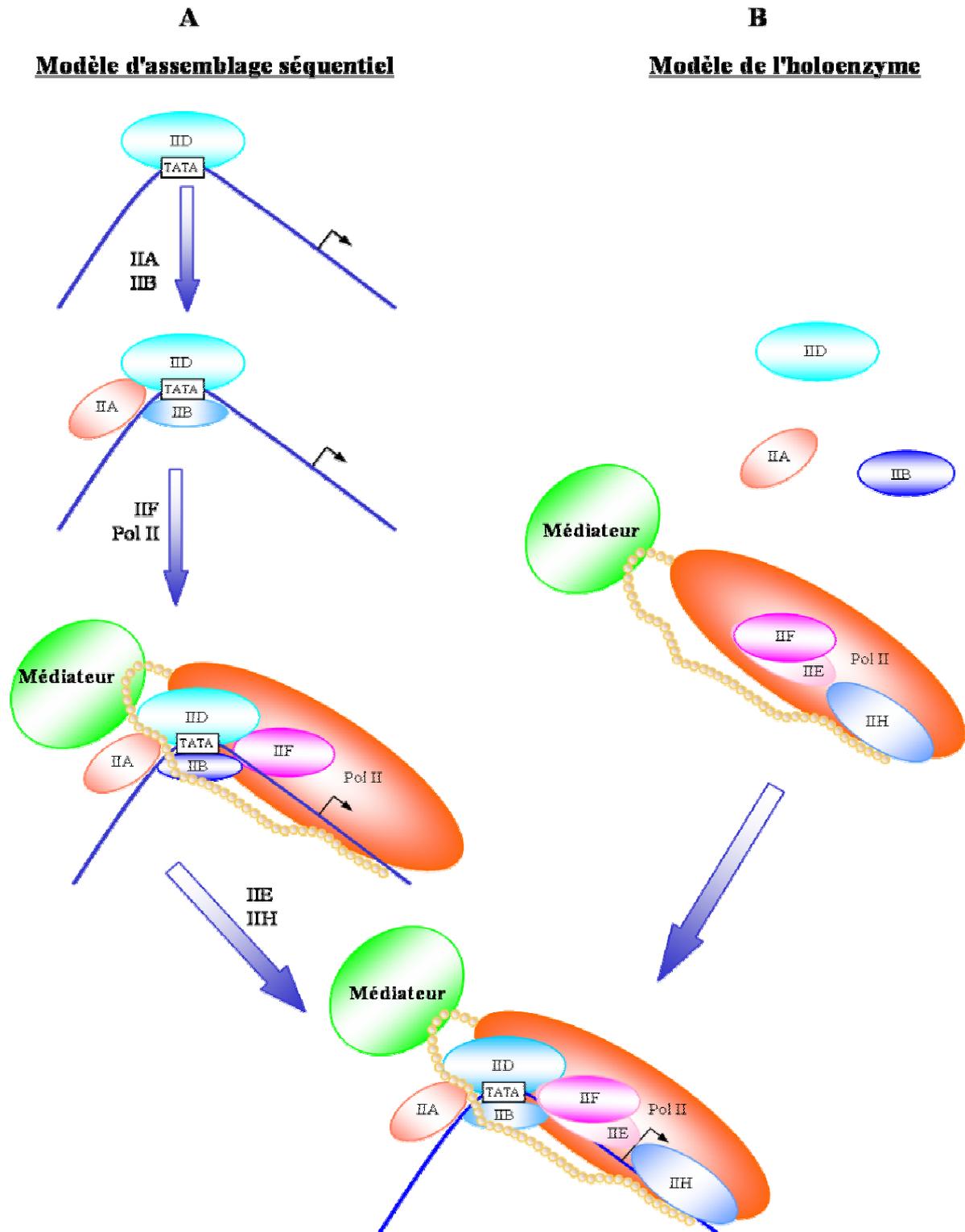


Figure 1 : Modèles d'assemblage du complexe de pré-initiation de la transcription par l'ARN polymérase II.

A- Modèle d'assemblage séquentiel.

B- Modèle de l'holoenzyme.

(Adapté d'après (Orphanides et al., 1996)).

Facteur général de transcription	Sous-unités	Propriétés
TFIIA	37kDa (α)	
	19kDa (β)	
	13kDa (γ)	
TFIIB	35kDa	Fixation asymétrique au promoteur
TFIID (1000kDa)	TBP 38kDa	Lie la boîte TATA
	TAF _{II} 250 250kDa	Activité Kinase
	TAF _{II} 150 150kDa	Lie le promoteur en amont
	TAF _{II} 135 135kDa	
	TAF _{II} 100 100kDa	Répétitions WD
	TAF _{II} 80 80kDa	Similarité à l'histone H4
	TAF _{II} 65 65kDa	
	TAF _{II} 55 55kDa	
	TAF _{II} 31 31kDa	Similarité à l'histone H3
	TAF _{II} 30 30kDa	
	TAF _{II} 28 28kDa	Histone-fold
	TAF _{II} 20 20kDa	Similarité à l'histone H2B
	TAF _{II} 18 18kDa	Histone-fold
TFIIE	56kDa (α)	Recrutement de TFIIH avec TFIIE β
	34kDa (β)	Lie l'ADN
TFIIF	58kDa (RAP74)	Stimule l'élongation
	30kDa (RAP30)	σ homologie
TFIIH (470kDa)	Voir Chapitre III	Voir Chapitre III
RNA polymérase II (500kDa)	hRPB1 217kDa	
	hRPB2 130kDa	
	hRPB3 30kDa	
	hRPB4 14.5kDa	
	hRPB5 22kDa	
	hRPB6 14kDa	
	hRPB7 17.8kDa	
	hRPB8 15.4kDa	
	hRPB9 13.8kDa	
	hRPB10 α 10kDa	
	hRPB10 β 9kDa	
hRPB11 13.3kDa		

Table 1 : Les facteurs de transcription généraux, leurs sous-unités et leurs propriétés.

1.2.1- Le facteur TFIID

Le facteur général de transcription TFIID est un complexe multiprotéique constitué de la TBP et d'au moins huit facteurs associés, les TAF (*TBP-associated factors*) : TAF250, TAF150, TAF130, TAF100, TAF80, TAF55, TAF30, TAF18, pour les principaux chez l'homme, nommés selon leur poids moléculaire (Burley and Roeder, 1996). La TBP est un facteur de transcription universel, requis pour le démarrage de la transcription par les ARN polymérases I, II et III. Elle reconnaît la boîte TATA et se fixe dans le petit sillon de l'ADN, formant un complexe stable, provoquant une courbure de l'ADN (voir chapitre II) (Kim et al., 1993). TBP est aussi capable d'interagir avec d'autres facteurs de transcription au sein du complexe de préinitiation. TBP est formé d'un domaine carboxy-terminal (*TBPcore*) bien conservé, servant à fixer l'ADN et interagissant avec certains facteurs de transcription au sein du PIC. Le domaine amino-terminal de TBP est beaucoup moins conservé, mais néanmoins nécessaire pour la viabilité chez la levure. Ce domaine déstabilise le complexe *TBPcore*/ADN, en jouant un rôle inhibiteur dans la fixation de *TBPcore* à l'ADN (Lee and Struhl, 2001). Des expériences d'anisotropie de fluorescence ont montré que ces deux domaines de TBP étaient capables d'interagir en solution (Khrapunov et al., 2002; Perez-Howard et al., 1995). La fixation de TBP au promoteur constitue la première étape dans la formation du PIC, et est une étape limitante du démarrage de la transcription.

Certains TAFs sont capables de reconnaître des séquences situées en aval de la boîte TATA, comme les initiateurs. Les TAFs sont indispensables à la transcription de base des gènes dont le promoteur ne possède pas la boîte TATA. Les TAF250 et TAF150 sont nécessaires à la fixation aux initiateurs et aux DPE (Verrijzer and Tjian, 1996). Quelques TAFs possèdent une activité enzymatique associée qui régule la transcription. C'est le cas de TAF_{II} 250 qui possède une activité kinase capable de phosphoryler la sous-unité RAP74 du facteur TFIIF (O'Brien and Tjian, 1998).

D'autres TAFs, possèdent un motif « *histone fold* ». C'est le cas de TAF 17, 48, 60 et 61 de levure (Gangloff et al., 2001; Reese et al., 2000). Ce motif permet de former des

paires de TAFs. La structure cristallographique de certaines de ces paires a été résolue et a révélé une grande homologie structurale avec l'octamère d'histone (Birck et al., 1998; Werten et al., 2002; Xie et al., 1996). Une étude de reconstitution de TFIID de levure a montré que TAF17, TAF60, TAF61c et TAF48 formaient un « *octamer TAF* » comportant deux copies des quatre TAFs. Cet octamère est similaire en terme de stoechiométrie et de spécificité de partenaires à l'octamère d'histone trouvé dans la chromatine (Luger et al., 1997; Selleck et al., 2001). Ainsi, TFIID interagirait avec l'ADN non seulement via TBP, mais aussi grâce aux TAFs.

1.2.2- Le facteur TFIIB

Le facteur TFIIB est le deuxième facteur général après TFIID à participer à la formation du PIC. Il est constitué d'un seul polypeptide de 35 kDa (Lin et al., 1991) qui interagit avec TBP et avec l'ADN. La courbure de l'ADN provoquée par la fixation de TBP augmente l'affinité de TFIIB pour l'ADN aussi bien en amont qu'en aval de la boîte TATA. La partie centrale de TFIIB contient un domaine de liaison à l'ADN du type hélice-coude-hélice (*helix-turn-helix*). Le site consensus (g/cg/cg/aCGCC) situé immédiatement en amont de la boîte TATA et appelé BRE (*TFIIB Recognition Element*) correspond au site de fixation de TFIIB (Lagrange et al., 1998). Ces mêmes auteurs, suggèrent que l'interaction entre BRE et TFIIB joue un rôle dans la détermination directionnelle (5'-3') de l'assemblage du PIC. En effet, du fait de la très forte symétrie du complexe TBP/ADN, TFIIB une fois fixé, orienterait l'assemblage du reste du PIC.

1.2.3- Le facteur TFIIF et l'ARN polymérase II

L'ARN polymérase seule n'est pas capable de se fixer de manière stable au PIC. Elle s'y associe en s'accompagnant du facteur TFIIF, qui est un hétérotétramère composé de sous-unités RAP30 et RAP75, de masse moléculaire 30 et 58 kDa, respectivement (Flores et al., 1988). Les deux sous-unités de TFIIF interagissent avec TFIIB mais seule RAP30 interagit avec l'ARN polymérase II. Des interactions entre TFIIF et des séquences d'ADN situées entre la boîte TATA et le site de démarrage de la transcription, ont été mises en évidence et pourraient participer à la stabilisation du

PIC. En plus de son rôle dans l'initiation de la transcription, l'association stable de TFIIF avec l'ARN Polymérase II joue un rôle dans le processus d'élongation (Orphanides et al., 1996). L'enzyme clé de la transcription des gènes codant pour les protéines est l'ARN polymérase II. Elle est composée de 10 à 14 sous-unités, selon les espèces (Young, 1991). On peut distinguer trois types de sous-unités : (1) les sous-unités Rpb1, 2, 3 et 11 dont on trouve des homologues dans les ARN polymérase I et III. Les trois premières sont aussi homologues à l'ARN polymérase bactérienne ; (2) cinq sous-unités (Rpb5, 6, 8, 10 et 12) sont communes aux ARN polymérases I, II et III ; (3) les sous-unités Rpb4, Rpb7 et Rpb9 sont spécifiques à l'ARN polymérase II (Hampsey, 1998). La plus grande sous unité, Rpb1, contient dans sa partie carboxy-terminale la répétition en tandem (26 chez *Saccharomyces cerevisiae* et 52 chez l'homme) d'un heptapeptide Tyr-Ser-Pro-Thr-Ser-Pro-Ser (Allison et al., 1988; Corden et al., 1985), qui est très conservé chez les organismes eucaryotes. Il est essentiel pour l'activité de l'ARN Polymérase II *in vivo* (Allison et al., 1988; Bartolomei et al., 1988; Nonet and Young, 1989). Ce domaine carboxy-terminal (CTD) est le substrat de plusieurs kinases, notamment le facteur TFIIH et le facteur d'élongation p-TEFb. Une délétion de ce domaine est létale chez la levure, la drosophile et la souris (Allison et al., 1988; Bartolomei et al., 1988; Zehring et al., 1988). Cependant, Le CTD n'est pas toujours indispensable pour la transcription *in vitro* (Kim and Dahmus, 1989; Zehring et al., 1988). L'état de phosphorylation du CTD, définit deux formes de l'ARN Polymérase II : IIA et IIO correspondant à la forme phosphorylée et à la forme non phosphorylée, respectivement (Dahmus, 1981). Ces deux formes jouent un rôle dans les étapes de transcription. L'ARN Pol IIA est présente lors de la formation du complexe de pré-initiation de la transcription (Chesnut et al., 1992; Laybourn and Dahmus, 1990; Lu et al., 1991), alors que l'ARN Pol IIO intervient lors de la transition entre l'initiation et l'élongation (Cadena and Dahmus, 1987; Laybourn and Dahmus, 1990; Zhang and Corden, 1991a; Zhang and Corden, 1991b).

La déphosphorylation du CTD, joue également un rôle dans la régulation de la transcription. En effet, le passage de la forme hypo-phosphorylée à la forme hyper-phosphorylée permet à l'ARN Pol II de recommencer un cycle de transcription. Une

phosphatase du CTD, la protéine phosphatase FCP1 (pour *TFIIF interacting CTD Phosphatase*) a été identifiée chez *Saccharomyces cerevisiae* (Chambers and Kane, 1996; Kobor et al., 1999) et chez l'homme (Chambers and Dahmus, 1994; Cho et al., 1999).

1.2.4- Le facteur TFIIE et TFIIH

L'ARN Polymérase II se trouvant au sein du complexe intermédiaire TFIID-TFIIB-ARN Pol II-TFIIF est incapable de démarrer la synthèse d'ARN et requiert pour ce faire la présence de deux autres facteurs généraux de transcription TFIIE et TFIIH. Le facteur TFIIE est composé de deux sous-unités α et β de poids moléculaire respectif de 56 et 34 kDa (Ohkuma et al., 1990). Une étude basée sur une purification du complexe par filtration sur gel a montré qu'il s'agirait d'un hétéro-tétramère $\alpha_2\beta_2$ (Ohkuma et al., 1990). Des études de pontage chimique et de centrifugation analytique ont montrés que TFIIE en solution se comporte comme un hétéro-dimère $\alpha\beta$. Sa structure moléculaire a été déterminée par microscopie électronique (article en préparation). TFIIE interagit avec la forme hyper-phosphorylée de l'ARN polymérase II (IIA) et avec les deux sous unités de TFIIF. Il interagit directement avec TFIIH et stimule la phosphorylation TFIIH-dépendante du CTD (Hampsey, 1998). TFIIH est un des facteurs généraux de transcription les plus complexes. En effet, formé de 9 sous-unités, il contient deux activités hélicases et une activité kinase (voir Chap III et IV) (Flores et al., 1992; Gerard et al., 1991). L'activité hélicase est impliquée dans l'ouverture de l'ADN autour du site d'initiation. TFIIH contrôle la transition entre l'initiation et l'élongation par son activité kinase qui permet de phosphoryler le CTD de l'ARN Pol II et le passage de l'ARN Pol II de la forme hypo-phosphorylée IIO à la forme hyper-phosphorylée IIA (Hampsey, 1998). TFIIH est impliqué dans la transcription des gènes de classe I (Iben et al., 2002), mais son rôle a été mieux caractérisé dans la transcription des gènes de classe II.

En fait, l'arrivée des facteurs TFIIE et TFIIH au sein du PIC constitue la dernière étape de l'initiation de la transcription. A cette étape, l'hélicase XPB de TFIIH est indispensable pour l'ouverture du promoteur (Coin et al., 1999) en se fixant sur l'ADN en aval du site d'initiation de +1 à +38 (Douziech et al., 2000; Kim et al., 2000). La bulle de transcription est ainsi formée et au fur et à mesure que la transcription

progressive, la bulle s'étend en aval d'une paire de base de la position du dernier ribonucléotide (rNTP) ajouté (Holstege et al., 1997). Lors de l'ajout du quatrième rNTP, le complexe d'initiation est stabilisé et TFIIH n'est plus nécessaire au maintien de la bulle de transcription. A la position +10, l'ARN polymérase peut encore entrer dans un cycle d'initiation avorté (*abortive initiation*). A la position +11, la polymérase entre dans un mode d'initiation productive. Entre l'initiation et l'élongation, les facteurs généraux de transcription, à l'exception de TFIIF, sont progressivement relâchés et la polymérase peut alors s'associer aux facteurs d'élongation. Ces derniers préviennent son arrêt durant la synthèse d'ARNm (Conaway et al., 2000).

1.3- Modèle de l'holoenzyme

Le modèle de l'holoenzyme a été proposé suite à la purification d'un complexe stable de levure appelé holoenzyme, comportant l'ARN polymérase II et d'autres protéines régulatrices de la transcription (co-activateurs et complexe *médiateur*). Deux complexes holoenzymes différents ont été isolés : le premier contient l'ARN Pol II, TFIIF, TFIIB, TFIIH et certains facteurs du complexe SRB (pour *Suppressor of RNA Polymerase II*) qui a une fonction de coactivateur (Koleske and Young, 1995) ; le deuxième complexe correspond à un sous-complexe du premier auquel il manque TFIIB et TFIIH (Figure 1B). En outre, l'holoenzyme peut contenir des co-activateurs comme SUG1 ou encore certains composants du médiateur comme SIN4 (Kim and Burley, 1994). Le modèle de l'holoenzyme propose qu'une fois ces complexes formés, ils soient ensuite recrutés en une seule étape sur le promoteur où TFIID est déjà fixé. La présence de protéines du complexe SRB suggère que l'activité de l'ARN Pol II est régulée spécifiquement. Chez l'homme, un complexe similaire a été immuno-purifié et caractérisé à partir de cellules HeLa. Il contient tout les homologues des facteurs trouvés chez la levure, ainsi que d'autres facteurs notamment de réparation de l'ADN (Chao et al., 1996; Maldonado et al., 1996; Neish et al., 1998). L'holoenzyme peut parfois contenir le complexe SWI/SNF impliqué dans le remodelage de la chromatine (Wilson et al., 1996). En somme, il n'existe pas un complexe holoenzyme universel, mais plutôt différents holoenzymes dont la composition est variable en

fonction du mode de purification, de l'organisme, du tissu et probablement aussi de la phase du cycle cellulaire. Ce modèle est très séduisant et peut constituer une économie très importante pour la cellule qui, en réponse à un stimulus, recrute en une seule étape le complexe de transcription. Cependant, ce modèle ne tient pas compte du caractère flexible de la réponse aux différents régulateurs de manière spécifique.

1.4- L'élongation

L'étape de l'initiation de la transcription est suivie des étapes d'élongation et de terminaison. Au cours de l'élongation, l'ARN polymérase se déplace le long de l'ADN en séparant les deux brins et en incorporant davantage de ribonucléotides.

Le rôle des facteurs d'élongation est d'empêcher que l'ARN polymérase s'arrête en cours de l'élongation. Les facteurs d'élongation rentrent en action une fois les facteurs d'initiation re-largués. Ces derniers sont classés en différentes catégories :

- 1) Facteurs qui augmentent le taux global d'élongation de la chaîne d'ARN : TFIIF, élongine et ELL.
- 2) Facteur qui permet à l'ARN Pol II de repartir lorsqu'elle est arrêtée : TFIIS.
- 3) Facteur qui stimule l'élongation par phosphorylation du CTD : p-TEFb
- 4) Facteurs qui favorisent l'arrêt de la polymérase : DSIF, NELF, facteur 2.
- 5) Facteur qui régule le taux d'élongation à travers les nucléosomes : FACT (*FAcilitates Chromatin Transcription*). Ce facteur permet à l'ARN Pol II arrêtée au niveau du nucléosome de reprendre la transcription indépendamment de l'hydrolyse de l'ATP et des facteurs TFIIF et TFIIS (Orphanides et al., 1998).

Le facteur TFIIF est capable de s'associer à l'ARN Pol II, induisant un changement conformationnel, empêchant son arrêt au niveau des sites transitoires (Aso et al., 1995; Lei et al., 1999). Le facteur SIII ou élongine, constitué de trois polypeptides A, B et C de poids moléculaire 110, 18 et 15 kDa respectivement, augmente la vitesse de polymérisation de l'ARN et supprime les pauses de l'ARN Pol II (Bradsher et al., 1993; Reines et al., 1996). ELL est un facteur d'élongation de 80 kDa, qui augmente aussi la vitesse de synthèse de l'ARN (Reines et al., 1996).

TFIIS est un facteur de 38 kDa, capable d'interagir directement avec la grande sous-unité de l'ARN pol II et ainsi lui permet de repartir lorsqu'elle est arrêté (Nudler et al., 1994; Rudd et al., 1994).

Le facteur p-TEFb (*Positive Transcription Elongation Factor b*) formé de cdk9 et de la cycline T, est capable de phosphoryler le CTD de l'ARN Pol II et ainsi permettre l'allongement de l'ARN synthétisé (Marshall et al., 1996). DSIF (*DRB Sensitivity Inducing Factor*) et NELF (*Negative Elongation Factor*) sont capables de s'associer à la polymérase et d'empêcher sa progression (Yamaguchi et al., 1999). La phosphorylation du CTD par p-TEFb favorise la dissociation de ces deux facteurs.

Le facteur 2 est une ATPase capable d'induire la séparation du transcrit par la polymérase, déstabilisant ainsi le complexe et aboutissant à la terminaison prématurée de la transcription (Xie and Price, 1997).

2- Facteurs de transcription spécifiques

Les facteurs de transcription spécifiques peuvent moduler le démarrage de la transcription de plusieurs manières. Ils peuvent notamment être impliqués dans le déplacement des répresseurs fixés sur le promoteur, dans le recrutement de facteurs généraux de l'ARN Pol II sur le promoteur, dans la modification de conformation du PIC par modifications post-traductionnelles de certaines protéines de ce complexe, ou encore dans le dégagement du promoteur.

Les facteurs de transcription spécifiques possèdent d'une part, un domaine de liaison à l'ADN qui reconnaît une séquence spécifique et d'autre part, un domaine activateur qui interagit avec des cofacteurs (activateurs ou répresseurs) et/ou certains facteurs généraux de transcription (Figure 2).

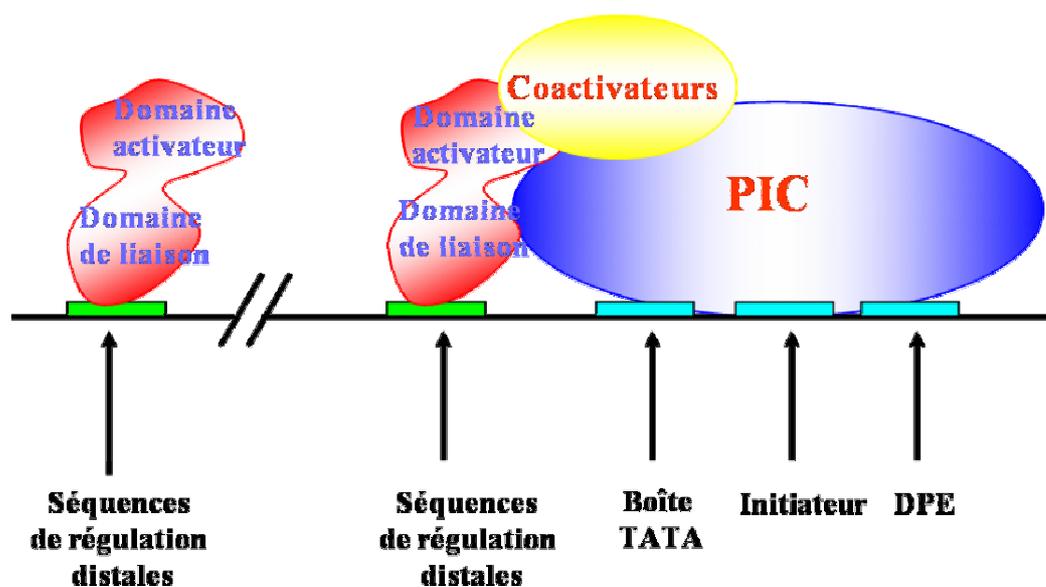


Figure 2 : Activation de la transcription.

Représentation schématique de l'activation de la transcription par la fixation de facteurs de transcription spécifiques sur les séquences de régulation proximales et distales du promoteur. L'activation de la transcription par les domaines activateurs se fait soit directement par l'interaction avec certains facteurs généraux de transcription, soit par l'intermédiaire de coactivateurs. (DPE pour *downstream promoter element* et PIC pour *Pre-initiation complex*).

2.1- Activateurs transcriptionnels

Le mode d'action des activateurs sur la machinerie de transcription de base est encore mal connu. Les activateurs agissent sur la transcription grâce à des contacts établis avec les facteurs généraux de transcription. En effet, un certain nombre d'interactions directes ont été reportées. Ainsi TBP est capable d'interagir *in vitro* avec divers activateurs tels que VP16, E1A, c-Rel, Tax1, p53, Sp1, Oct1 et Oct2 (Ranish and Hahn, 1996; Zawel and Reinberg, 1995) même si certaines mutations de TBP affectent l'interaction avec certains de ces activateurs et n'empêchent pas ces derniers de remplir leur rôle d'activateurs (Tansey and Herr, 1995). TFIIB est aussi la cible directe d'activateurs de la transcription tels que VP16, Sp1 et Krüpel. Une mutation dans TFIIB, inhibant l'association avec VP16, empêche l'activation de la transcription (Roberts et al., 1993). VP16 induit un changement de conformation de TFIIB au sein du PIC, ce qui entraîne un contact avec TFIIF et l'ARN Pol II (Roberts and Green, 1994). D'autres interactions ont été mises en évidence entre TFIIH et VP16 d'une part et p53 d'autre part (Xiao et al., 1994). Cdk7, la kinase du facteur TFIIH

phosphoryle le domaine AF1 du récepteur RAR- α (Bastien et al., 2000; Keriél et al., 2002; Rochette-Egly et al., 1997). Il a aussi été mis en évidence que la dimérisation des facteurs de transcription c-Jun et c-Fos permet l'interaction avec TFIIE et TFIIIF (Martin et al., 1996).

2.2- Coactivateurs transcriptionnels

La grande majorité des activateurs agissent le plus souvent sur le complexe de transcription par l'intermédiaire de coactivateurs (ou corépresseurs). A la différence des facteurs généraux de transcription, ces coactivateurs transcriptionnels, appelés aussi médiateurs ou adaptateurs, ne sont pas nécessaires pour la transcription de base, et contrairement aux facteurs de transcription spécifiques, la plupart ne se fixent pas directement à l'ADN. On distingue plusieurs types de coactivateurs : TFIIA, certains TAF ainsi que le complexe médiateur (encore appelé SRB/médiateur).

2.2.1- TFIIA

Le facteur général de transcription TFIIA est constitué de 3 sous unités (α , β et γ) dont les deux plus grandes (α et β) sont codés par le même gène et sont produites par clivage post-traductionnel (DeJong and Roeder, 1993; Ozer et al., 1994). Le facteur TFIIA interagit avec des séquences d'ADN situées en amont de la boîte TATA et avec le domaine N-terminal de la TBP. TFIIA se fixe sur le complexe TBP-ADN du côté opposé à celui impliqué dans l'interaction avec TFIIB (Tan et al., 1996). Sa fixation permet de stabiliser le complexe TBP-ADN. TFIIA semble aussi contribuer à supprimer l'effet répresseur de certains facteurs qui interagissent directement avec TBP et inhibent la fixation de TFIID sur l'ADN (Orphanides et al., 1996). Au départ, TFIIA a été décrit comme facteur général de transcription indispensable pour un grand nombre de gènes (Matsui et al., 1980). Une étude plus approfondie de ce facteur a permis de montrer que TFIIA n'est pas essentiel pour la transcription basale mais que son interaction avec TBP est nécessaire pour l'activation de la transcription *in vivo* par certains activateurs riches en résidus acides (activateurs acides), définissant ainsi TFIIA comme coactivateur (Bryant et al., 1996; Stargell and Struhl, 1995).

2.2.2- Les TAFs

La TBP seule est suffisante pour la reconnaissance du promoteur et donc pour l'assemblage du PIC. Cependant, les TAFs sont indispensables pour l'activation de la transcription par les facteurs spécifiques (Hoffman et al., 1990; Pugh and Tjian, 1990). Chez la drosophile, il a été montré que les activateurs Sp1 et Bicoid (domain activateur riche en glutamine) interagissent avec TAF_{II}110 et que les activateurs acides VP16 et p53 interagissent avec TAF_{II}40 et TAF_{II}60. Le récepteur aux oestrogènes interagit avec TAF_{II}30 et le facteur CTF (domaine d'activation riche en proline) interagit avec TAF_{II}55 (Verrijzer and Tjian, 1996). L'ensemble de ces résultats suggère que les TAFs permettent à TFIID d'intégrer des signaux provenant de plusieurs activateurs. Cependant, des expériences d'inactivation de gènes chez la levure, montrent que l'inactivation de yTAF_{II}145 et yTAF_{II}90 n'affectait l'expression que de 17 et de 8% des gènes, respectivement (Holstege et al., 1998). De plus, des expériences effectuées sur des cellules de hamster portant une mutation thermo-sensible sur le gène de TAF_{II}250 montrent une inactivation du gène codant pour la cycline A alors que le gène c-fos est normalement régulée à température non permissive (Ruppert et al., 1993; Wang and Tjian, 1994), ce qui suggère un rôle plus spécifique des TAFs dans l'activation de la transcription.

2.2.3- Le médiateur

Le complexe SRB/médiateur est un complexe multi-protéique qui a été mis en évidence pour la première fois chez la levure (Flanagan et al., 1991; Kelleher et al., 1990). Ce complexe fait partie de l'holoenzyme. Il est associé avec le CTD de l'ARN Pol II (Nonet and Young, 1989). Ayant un poids moléculaire d'un méga-dalton, le médiateur est composé de 20 polypeptides différents dont certains sont indispensables pour la viabilité de la cellule. Ils appartiennent à trois groupes de protéines : Med, Sin4/Rgr1 et Srb (Table 2).

	Factor	Mass (kDa)	Gene(s)	Essential	Metazoan homologs
Groupe Srb	Srb2	23	<i>SRB2</i>	No	
	Srb4	78	<i>SRB4</i>	Yes	
	Srb5	34	<i>SRB5</i>	No	
	Srb6	14	<i>SRB6</i>	Yes	
	Srb7	16	<i>SRB7</i>	Yes	Yes
	Srb8	167	<i>SRB8, SSN5, ARE2</i>	No	
	Srb9	160	<i>SRB9, SSN2, UME2</i>	No	
	Srb10	63	<i>SRB10, SSN3, UME5, ARE1</i>	No	
	Srb11	36	<i>SRB11, SSN8, UME3</i>	No	
	Gal11	38	<i>GAL11, SPT13, SDS4, RAR3</i>	No	
	Groupe Sin4/Rgr1	Sin4	111	<i>SIN4, SSN4, TSF3</i>	No
Rgr1		123	<i>RGR1</i>	Yes	Yes
Rox3 ^b		25	<i>ROX3, SSN7</i>	Yes	
Med1		64	<i>MED1</i>	No	
Groupe Med	Med2	48	<i>MED2</i>	No	
	Med3, Pgd1, Hrs1	47	<i>PGD1, HRS1</i>	No	
	Med4	32	<i>MED4</i>	Yes	
	Med6	33	<i>MED6, MTR32</i>	Yes	Yes
	Med7	32	<i>MED7</i>		Yes
	Med8	25	<i>MED8</i>	Yes	

Table 2 : Composition du médiateur.

Le médiateur est formé de 20 polypeptides organisés en trois groupes de protéines représentés par trois couleurs différentes.

Adapté d'après (Hampsey, 1998).

Le groupe Med est formé de Med 1, 2, 3, 4, 5, 6, 7 et 8. Le groupe Sin4/Rgr1 est formé de Gal 11, Sin 4, Rgr 1 et Rox 3. Finalement le groupe Srb est composé de Srb 2, 4, 5, 6, 7, 8, 9, 10 et 11. Contrairement aux TAFs, le complexe SRB/médiateur joue un rôle général dans l'activation de la transcription. Des études récentes chez les eucaryotes supérieurs ont permis de mettre en évidence des complexes homologues au médiateur (TRAP/SMCC, PC2, DRIP, ARC, CRSP, NAT). Ces complexes, isolés par des auteurs différents à partir de systèmes cellulaires différents, sont clairement apparentés puisqu'on y trouve un certain nombre de sous-unités communes. Le complexe TRAP/SMCC (*Thyroid hormone Receptor Associated Protein / SRB and MED-Containing Cofactor complex*) est un des mieux caractérisés ; il a été isolé sur la base de son interaction avec le récepteur de l'hormone thyroïdienne lié à son ligand (Ito et al., 1999). Ce complexe, d'un poids moléculaire de 1.5 MDa, est constitué de 25 sous-unités incluant des protéines homologues aux sous-unités Srb 7, Srb 10, Srb 11, Med 6 et Med 7 de levure. Ce complexe sert de coactivateur à plusieurs activateurs tels que le récepteur aux hormones thyroïdiennes, VP16 et p53 (Gu et al., 1999; Ito et al., 1999). Du point de vue moléculaire, la structure en microscopie électronique du

médiateur en présence et en absence d'ARN polymérase a montré un passage d'une forme compacte à une forme allongée, provoqué par l'interaction avec la polymérase (Dotson et al., 2000).

3- Chromatine et Transcription

La présentation ci-dessus des facteurs impliqués dans la transcription et leurs modes d'action est basée sur un ensemble de résultats obtenus essentiellement dans des systèmes reconstitués *in vitro* avec de l'ADN « nu ». Une telle situation rappelle en fait plutôt celle que l'on trouve chez les procaryotes. Chez ces organismes, les activateurs peuvent interagir directement avec des composants de la machinerie de transcription. Dans les cellules eucaryotes, l'ADN est associé à un grand nombre de protéines (dont les histones) et se présente sous forme de structure compacte, la chromatine. En effet, cette structure particulière sert à compacter la fibre d'ADN, ce qui ajoute un niveau de complexité supplémentaire à la régulation de la transcription. En fait, la compaction de l'ADN en chromatine la rend non accessible à la machinerie de transcription et provoque une répression constitutive de la transcription. Les cellules eucaryotes ont donc développées des mécanismes pour modifier la structure de la chromatine, faciliter ainsi l'accessibilité à l'ADN et permettre la transcription des gènes.

3.1- Structure de la chromatine

La fibre de chromatine est un filament de 10 nm de diamètre qui apparaît comme un collier de perles en microscopie électronique. Ces perles sont des éléments de base de la chromatine et sont appelés nucléosomes (Figure 3A). Ils sont séparés par une région d'ADN dont la taille varie de 40 à 50 paires de bases (pb) selon les espèces et le type cellulaire. La fibre de chromatine (de 10 nm) s'enroule sur elle-même en une structure hélicoïdale pour former une fibre de 30 nm. Cette compaction supplémentaire implique l'histone H1 dont la stoechiométrie est d'une molécule par nucléosome. Ces molécules d'histones H1 entrent en contact avec l'ADN de liaison situé entre les nucléosomes et compactent la chromatine en s'associant les unes aux autres.

Le nucléosome contient 146 pb d'ADN enroulées 1.65 fois autour d'un octamère d'histones formé de 2 hétérodimères d'histones H2A-H2B et un tétramère d'histones H3-H4). Les histones sont des petites protéines basiques de 11 à 15 kDa. Sa structure a été déterminée par cristallographie aux rayons X et a révélée une organisation tripartite dans laquelle un tétramère central (H3-H4)₂ est entouré de deux dimères H2A-H2B (Arents et al., 1991). Les différentes histones présentent un motif structural conservé appelé *histone fold* (Figure 3B). L'assemblage nucléoprotéique du nucléosome est verrouillé par les histones H1 et H5, qui interagissent avec l'ADN. Au total, 12 points de contact avec l'ADN sont répartis de façon périodique à la surface du nucléosome.

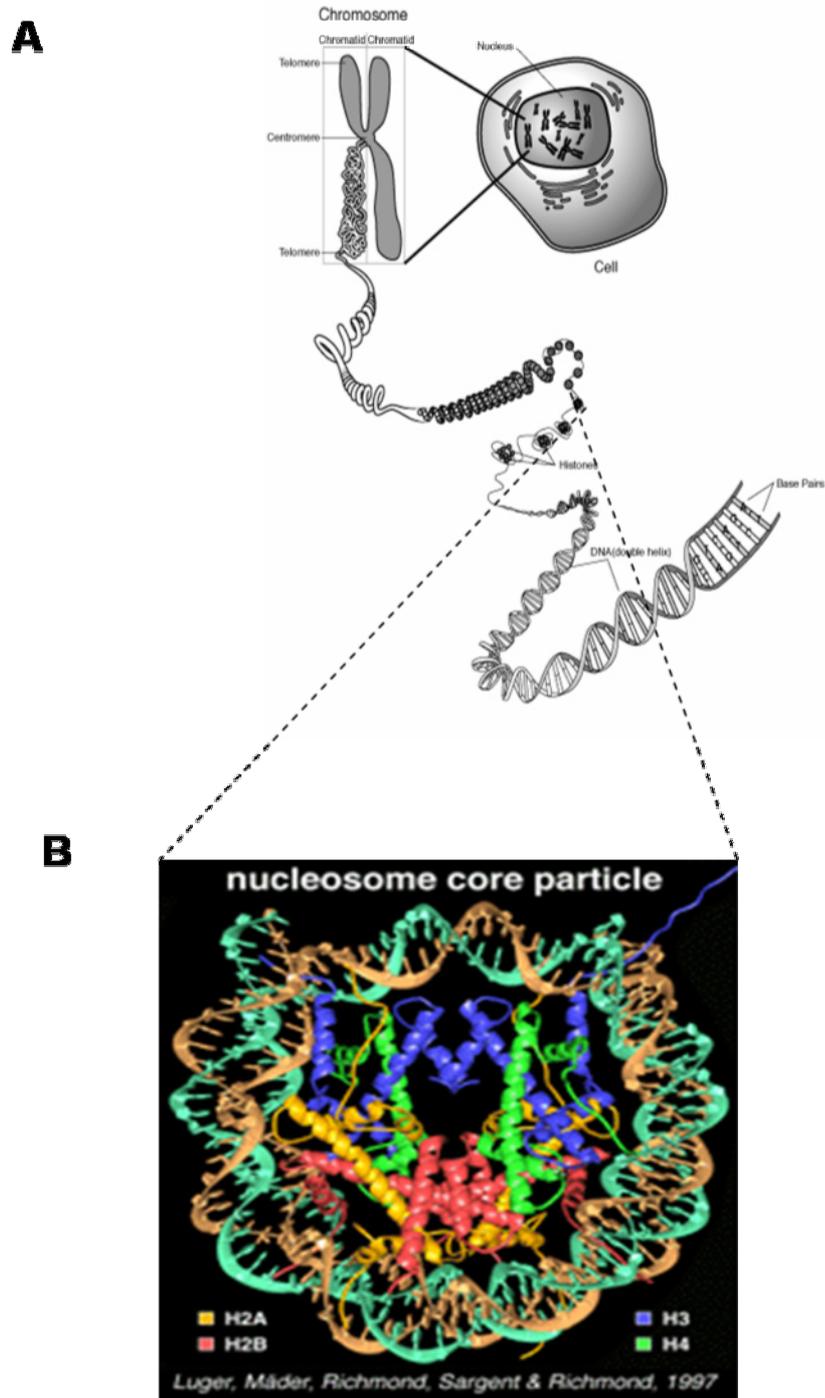


Figure 3 : Compaction de l'ADN.

A- Schéma représentatif de la compaction de l'ADN dans le noyau de la cellule (Adapté d'après (Nelson et al., 1986)).

B- Structure cristallographique du nucléosome à 2.8 Å. La double hélice d'ADN de 146 Pb (représenté en marron et vert) est enroulée autour d'un octamère histones (H3 : bleu ; H4 : vert ; H2A : jaune ; H2B : rouge). (Luger et al., 1997).

Le domaine amino-terminal des histones formant le nucléosome encore appelé « queue N-terminale » est très flexible et est dirigé vers l'extérieur du nucléosome. Il comporte plusieurs résidus chargés positivement, qui peuvent être modifiés par acétylation ou méthylation, ainsi que des sérines qui peuvent être phosphorylées.

En plus des histones, d'autres protéines peuvent être extraites de la chromatine. Il s'agit notamment des protéines HMG (High Mobility Group). Ces derniers possèdent un faible poids moléculaire (< 30 kDa) et sont riches en acides aminés chargés (Bustin and Reeves, 1996). Leur rôle est de faciliter l'accès de facteurs de transcription à l'ADN. Ainsi, HMG-1 facilite la liaison du récepteur à la progestérone à son site de fixation en modifiant la courbure de l'ADN cible. De même, HMG-I(Y) favorise les interactions protéine-protéine au sein du PIC (Bustin and Reeves, 1996).

Des expériences réalisées *in vitro* et *in vivo* ont montré que la présence des nucléosomes sur l'ADN réprime la transcription (Kornberg & Lorch., 1999; Grunstein & al., 1990). De plus, il a été démontré que les nucléosomes inhibent la formation du PIC sur le promoteur. L'addition d'activateurs à la réaction de transcription permet de lever cette répression (Owen-Hughes and Workman, 1994). Des expériences d'immunoprécipitation de la chromatine, montrent que TBP ne peut pas s'associer à l'ADN empaqueté (Kuras and Struhl, 1999; Li et al., 1999), indiquant que l'initiation de la transcription ne peut se faire sur une matrice chromatinienne. D'où la nécessité de réguler l'accessibilité de l'ADN dans ce contexte chromatinien.

3.2- Remodelage de la chromatine par modifications covalentes

Le terme « remodelage de la chromatine » recouvre l'ensemble des changements structuraux qui affectent le nucléosome et leur distribution sur l'ADN. Ces changements de conformation sont catalysés par des protéines responsables d'addition ou de suppression d'un groupement chimique. Il s'agit d'enzymes capables de modifier covalamment les histones par acétylation, phosphorylation ou méthylation essentiellement.

3.2.1- Acétylation des histones

Une corrélation a été établie il y a presque quarante ans entre le niveau d'acétylation des histones et l'activité transcriptionnelle dans les cellules eucaryotes (Allfrey et al., 1968). Ultérieurement, les histones acétylées ont été montrées être préférentiellement associées à la chromatine transcriptionnellement active (Struhl, 1998). L'état d'acétylation des histones résulte d'un équilibre entre deux types d'activités antagonistes : l'activité histone acétyl transférase (HAT), qui consiste à ajouter un groupement acétyl sur certaines lysines des queues des histones, et une activité histone déacétylase (HDAC) qui consiste à enlever ces groupements acétyl. Quatre lysines sont acétylables dans le cas des histones H2B, H3 et H4 alors qu'une seule peut l'être dans l'histone H2A (Mizzen and Allis, 1998). L'acétylation a pour effet de neutraliser la charge positive des lysines et ainsi de diminuer leur affinité pour l'ADN, ce qui augmente l'accessibilité d'autres protéines (Lee and Garrard, 1992; Oliva et al., 1990; van Holde et al., 1992; Vettese-Dadey et al., 1996).

Il existe deux groupes de protéines à activité HAT. Elles sont soit nucléaires (groupe A) soit cytoplasmiques (groupe B). Dans le groupe A, trois familles ont été identifiées : GNAT, MYST et p300/CBP. Toutes ces protéines possèdent un motif très conservé de fixation à l'acétyl coA. Ces HATs fonctionnent *in vivo* au sein de complexes multi-protéiques de fonctions biologiques distinctes. C'est le cas de yADA, ySAGA, hPCAF, yNuA3 et yNuA4.

Les HAT de type B acétylent seulement les histones H3 et H4 néo-synthétisées avant leur incorporation à la chromatine (Sobel et al., 1995).

Certains co-activateurs transcriptionnels possèdent une activité histone acétyl transférase ; c'est le cas de p300, PCAF et CBP (Ogryzko et al., 1996). Certains facteurs généraux de transcription possèdent aussi une activité HAT, c'est le cas de TAF_{II}250 du facteur TFIID (Mizzen et al., 1996). D'autre part, l'implication de facteurs de transcription de l'ARN polymérase III tels que TFIIC220, TAFIIC110 et TAFIIC90 dans l'acétylation des histones, montre que ce mode de régulation n'est pas limité uniquement à la transcription des gènes codant pour les protéines (Sternier and Berger, 2000).

3.2.2- Autres modifications des histones

Outre l'acétylation au niveau de la queue N-terminale, les histones peuvent subir d'autres modifications post-traductionnelles telles que la phosphorylation et la méthylation. La découverte de leur importance dans le contrôle de la transcription est plus récente et leur rôle précis reste encore à déterminer. La phosphorylation des histones H1 et H3 est connue depuis longtemps et était initialement impliqué dans la condensation des chromosomes. Néanmoins, des études récentes ont remis en cause cette hypothèse. En effet, il a été montré que cette phosphorylation se situait au niveau de régions ouvertes et transcriptionnellement actives de la chromatine. La protéine Rsk-2 a été identifiée comme étant responsable de la phosphorylation de la serine 10 de l'histone H3 (Mizzen and Allis, 1998). D'autre part, le rôle de la méthylation des histones est encore mal connu. Chez l'Homme, la protéine responsable de la méthylation de la lysine 9 de l'histone H3 a été identifiée. Il s'agit de SUV39H1 (Rea et al., 2000). Cette méthylation de l'histone H3 serait un signal pour une déacétylation de l'histone H4 (Strahl and Allis, 2000). Strahl et collaborateurs ont suggéré que des combinaisons de différentes modifications post-traductionnelles des queues N-terminales des histones pourraient constituer un code de reconnaissance, appelé code histone, pour des protéines interagissant avec les histones. En effet, il a été démontré par exemple que l'acétylation de la lysine 14 (caractérisant un état transcriptionnellement inactif) précède la méthylation de la lysine 9, qui est elle-même précédée de la phosphorylation de la serine 10. Ce code histone tient donc compte de la nature des modifications et aussi de leur chronologie (Berger, 2002). Il fait partie intégrante de la voie de régulation de la transcription par la régulation du remodelage de la chromatine.

3.3- Remodelage de la chromatine par modifications non covalente

Le remodelage de la chromatine par modifications non covalentes est catalysé par des complexes multi-protéiques appelés facteurs de remodelage. Ces facteurs ont tous en commun une activité ATPase et présentent une homologie de séquence avec les ADN-hélicases. Cette activité ATPase permet au complexe de modifier la structure chromatinienne grâce à l'énergie libérée par l'hydrolyse de l'ATP. Les complexes de remodelage sont classés selon l'identité de leur sous-unités. SWI2/SNF2 (pour *SWItching mating Type/Sucrose Non-Fermenting*) définit la première famille du complexe, appelée SWI/SNF (Cote et al., 1994). L'ATPase ISWI (pour *Initiation of SWItch*) définit la deuxième famille du complexe. La troisième famille connue à ce jour est appelée Mi2/NuRD (pour *Nucleosome Remodeling Histone Deacetylase Complex* par référence à la sous-unité catalytique Mi2) (Vignali et al., 2000). D'autres complexes ont ensuite été isolés, tels que le complexe RSC (pour *Remodeling the Structure of Chromatin*), le complexe NURF (pour *NUcleosome Remodeling Factor*), le complexe CHRAC (pour *Chromatin Accessibility Complex*) ou encore le complexe ACF (pour *ATP-dependent Chromatin Factor*) chez la levure. Deux modèles pour le mécanisme ATP-dépendant du remodelage de la chromatine ont été proposés : (i) un glissement de l'ADN (*Sliding*) et (ii) un changement de conformation de la structure chromatinienne (Lorch et al., 1999; Studitsky et al., 1994). Les complexes de remodelage peuvent parfois interagir avec des activateurs de transcription, c'est le cas du récepteur aux glucocorticoïdes et du récepteur à l'oestrogène qui interagit avec le complexe SWI/SNF (Deroo and Archer, 2001; Hassan et al., 2001; Vignali et al., 2000). Des mutations des résidus acides de l'activateur VP16 empêchent l'interaction avec SWI/SNF (Hassan et al., 2001; Peterson and Workman, 2000), mais pas avec l'ARN Pol II holoenzyme (Wilson et al., 1996), montrant ainsi le lien entre les complexes de remodelage de la chromatine et la machinerie de transcription.

Les complexes de remodelage non covalents ont été trouvés associés avec des complexes à activité HAT ou HDAC, suggérant ainsi que ces deux types de complexes puissent fonctionner ensemble sur les promoteurs.

De manière similaire, il n'est pas impossible que la fixation des facteurs de transcription facilite le remodelage de la chromatine et *vice versa*.

En somme, l'ensemble des travaux décrits dans ce chapitre, met en évidence un contrôle étroit et une fine régulation de la transcription. Ainsi dans l'objectif d'apporter une explication moléculaire à cette régulation, un certain nombre de structures ont été résolus, et contribuent à une meilleure compréhension de la transcription.

Chapitre II – La transcription à la lumière de la structure

Trente ans après la découverte de l'ARN polymérase II, l'enzyme clé de la transcription des gènes codant pour les protéines, la biologie structurale a joué un rôle considérable dans la compréhension du mécanisme moléculaire de la transcription. En effet, la résolution de la structure cristallographique de l'ARN polymérase II eucaryote et procaryote, du nucléosome et des complexes dimériques, trimériques ou tétramériques contenant TBP, représente une avancée très importante pour une meilleure compréhension de l'architecture et de la dynamique du promoteur au cours des phases d'initiation de la transcription. La microscopie électronique a contribué essentiellement à avoir une image globale des complexes multi-protéiques impliqués dans la transcription en particulier TFIID, TFIIH et le médiateur. Tout le long de ce chapitre, nous nous proposons de montrer en quoi la contribution de la biologie structurale a été primordiale pour une meilleure compréhension de la régulation de la transcription et plus particulièrement l'étape de l'initiation.

1-La structure du nucléosome

Comme nous l'avons vu dans le chapitre I, la chromatine possède une organisation structurale hiérarchisée. La plus petite unité structurale dans cette hiérarchie est le nucléosome. Sa structure cristallographique a été résolue par l'équipe de Richmond à 2.8 Å à partir d'ADN palindromique dérivé d'ADN α -satellite humain, et de protéines histones recombinantes (Figure 4A). Cette structure révèle l'assemblage de l'octamère d'histone, entouré de 146 paires de bases d'ADN sur-enroulé en « super-hélice » (Luger et al., 1997). Les interactions histones/histones et histones/ADN dépendent des motifs *histone fold*, ainsi que des régions d'extension structurées de ces motifs. L'octamère d'histone est divisé en deux dimères d'*histone fold* définis par les paires d'histones H3-H4 et H2A-H2B (Figures 4B et C). La région centrale de l'*histone fold* des quatre protéines histones est formé de même motif structural de trois hélices α reliées par deux boucles ($\alpha 1$ -L1-c2-L2- $\alpha 3$). L'hétérodimérisation des protéines d'histone est à l'origine de l'unité de fixation à l'ADN (Luger et al., 1997).

Très récemment, la même équipe a résolu la structure de l'ADN au sein du core-nucléosome, préparé à partir de chromatine endogène. Cette structure à haute résolution (1.9\AA) de 147 paires de bases a été comparée à la structure de 146 paires de bases décrite précédemment et a montré une altération du « *twist* » de l'ADN, qui pourrait expliquer la possibilité de formation de fibre chromatinienne ou encore le remodelage de la chromatine *in vivo* (Richmond and Davey, 2003).

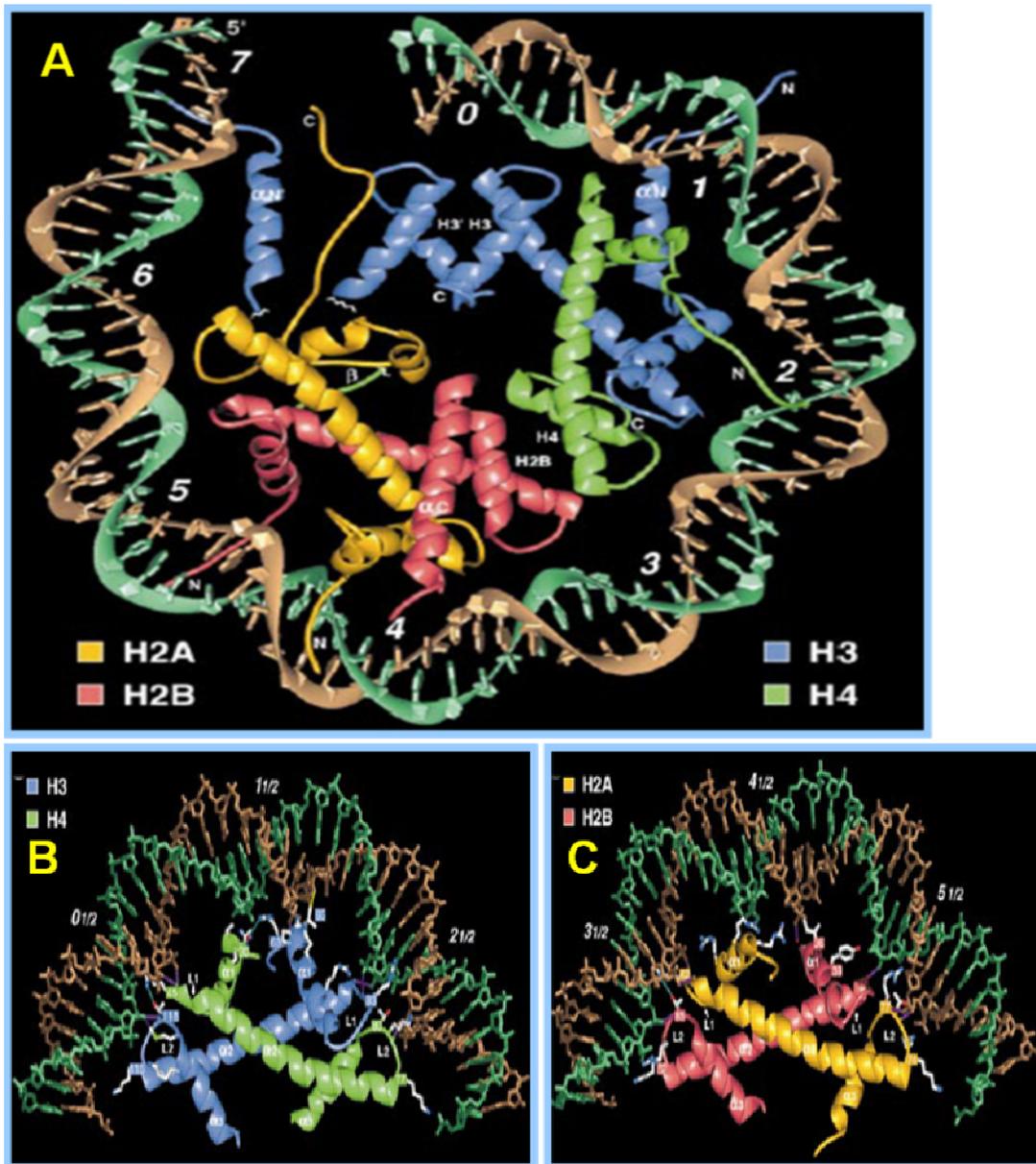


Figure 4 : Structure du nucléosome.

A- Structure cristallographique du « *nucleosome core particle* » à 2.8\AA . Les *histone folds* de H2A, H2B, H3 et H4 sont indiqués en jaune, rouge, bleu et vert, respectivement. La super-hélice d'ADN est indiquée en vert et en marron.

B- Structure de la paire d'*histone fold* H3-H4 montrant le motif caractéristique $\alpha 1-L1-\alpha 2-L2-\alpha 3$.

C- Structure de la paire d'*histone fold* H2A-H2B (Luger et al., 1997)..

2-Le promoteur avant l'ARN polymérase II

2.1- TBP

La structure de TBP d'*Arabidopsis thaliana* a été résolue par cristallographie aux rayons X à 2.6 Å (Figure 5A). Cette structure a révélée un nouveau repliement de fixation à l'ADN ressemblant à une structure « en fer à cheval ». La protéine est formée de deux modules identiques composés de 5 feuillets- β anti-parallèles et 2 hélices α . Ces 2 modules sont reliés par une symétrie d'ordre 2, ce qui avait déjà été suggéré par la conservation de séquence. Une boucle basique permet de les connecter. La surface de liaison à l'ADN est courbée et formée de feuillets- β anti-parallèles. La structure de TBP de levure et humaine a été résolue par la suite et a montré une très grande homologie structurale suggérant une très bonne conservation au cours de l'évolution (Chasman et al., 1993; Nikolov et al., 1996).

2.2- Complexe TBP/boîte TATA

La structure cristallographique du complexe TBP/ADN a été résolue à 2.25Å (Figure 5B). Cette structure a montré que TBP d'*Arabidopsis thaliana* induit un changement dans la conformation de la double hélice d'ADN du promoteur tardif de l'adénovirus formant un angle de 100°. Les contacts entre ADN et TBP concernent exclusivement les phospho-riboses et s'effectue au niveau du sillon mineur de l'ADN. Aucune interaction avec le sillon majeur n'a été observée, ce qui a été suggéré auparavant par des expériences de pontage chimiques et de *footprinting*. La transition de la forme B d'ADN à une forme courbée est induite par l'interaction des Phe 148, Phe 165 d'une part et Phe 57 ainsi que Phe 74 d'autre part, avec A (-30), A (-31), A (-25) et G (-24), respectivement. Les dix feuillets- β , forment une surface hydrophobe d'interaction avec l'ADN. Les deux hélices α situées à l'extérieur de la structure « en fer à cheval » pourraient être impliquées dans une interaction avec d'autres facteurs de transcription. Cette structure TBP/ADN représente le premier exemple d'un mode de reconnaissance protéine/ADN séquence spécifique, sans avoir recours à une interaction via le sillon majeur.

La structure cristallographique de TBP de levure et d'*Arabidopsis* complexée à l'élément TATA a permis de montrer une très grande homologie structurale des deux complexes et par conséquent un mécanisme commun de reconnaissance protéine/ADN (Kim and Burley, 1994; Nikolov et al., 1996). En effet, ces trois complexes mettent en évidence la même déformation de l'ADN, même si le premier domaine de TBP semble être relativement moins conservé structurellement que le deuxième entre l'homme, la levure et la plante.

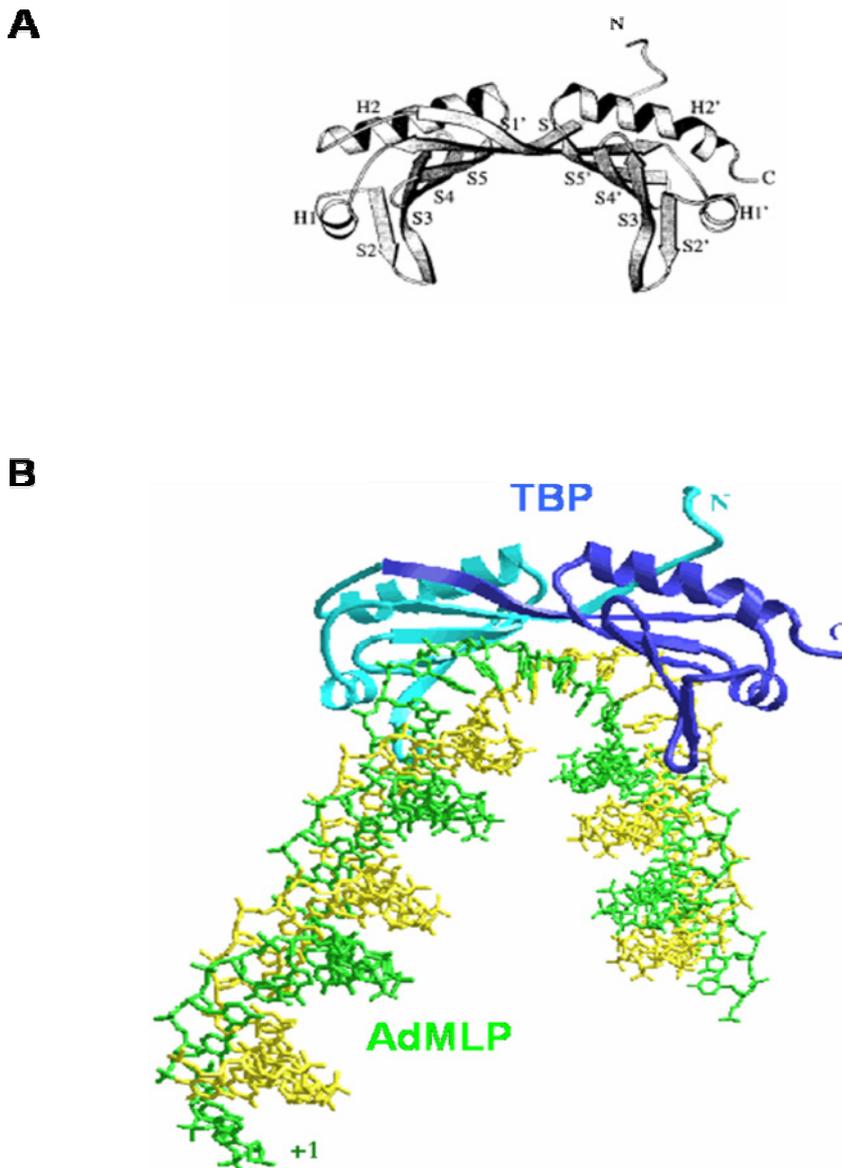


Figure 5 : Structures de TBP seul et lié à l'ADN.

A- Structure cristallographique de TBP d'*Arabidopsis thaliana* à 2.6Å (Nikolov et al., 1992).

B- Structure cristallographique du complexe TBP d'*Arabidopsis thaliana* en complexe avec l'ADN du promoteur tardif de l'adénovirus à 2.25 Å (Kim et al., 1993).

2.3- Complexe TBP/ADN/TFIIB

La structure du complexe ternaire composé du facteur de transcription humain TFIIB, TBP d'*Arabidopsis thaliana* et l'ADN du promoteur tardif de l'adénovirus a été résolue à 2.7 Å (Figure 6B). Cette structure révèle que le core-TFIIB (cTFIIB) est composé de deux domaines hélicoïdaux qui adoptent un repliement ressemblant à celui de la cycline A et interagit avec la partie carboxy-terminale de TBP. Le cTFIIB reconnaît le complexe TBP/ADN grâce à des interactions protéine-protéine et protéine-ADN. En effet, le core-TFIIB fait aussi des contacts avec les phosphoriboses du core promoteur en amont et en aval du centre de l'élément TATA. Cette structure constitue un point de départ pour l'étude d'autres facteurs de régulation de la transcription tels que les TAFs, les activateurs, co-activateurs ou encore d'autres facteurs d'initiation de la transcription.

Une étude structurale d'un homologue archaebactérien de TFIIB, appelé TFB en complexe avec TBP et le promoteur du virus SSV1 (pour *Solfolobus shibatae virus 1*) a permis d'étudier le stéréo-spécificité d'interaction entre l'élément de réponse à TFIIB et le motif hélice-boucle-hélice de TFB et a montré que la fixation de TFB est à l'origine de la polarité du complexe de préinitiation chez les archaebactéries (Littlefield et al., 1999).

Tsai et Sigler ont résolu la structure de TFIIB/TBP en complexe avec un ADN allongé du promoteur tardif de l'adénovirus. Cette structure a révélé que TFIIB se fixait non seulement au grand sillon en amont de la boîte TATA, mais aussi au petit sillon en aval de cette boîte. Cette fixation asymétrique de TFIIB pourrait être à l'origine de l'assemblage unidirectionnel de la machinerie d'initiation de la transcription (Tsai and Sigler, 2000).

Un complexe ternaire de la transcription des gènes de classe III, composé de Brf1 (*TFIIB-related factor*), de TBP et de l'ADN du promoteur U6 a été étudié structuralement (Juo et al., 2003). La résolution de sa structure à 2.95 Å montre que Brf1 interagit spécifiquement avec la surface convexe et latérale de TBP, confirmant ainsi le rôle universel de TBP dans la transcription des gènes de classe II et III (Figure 8A).

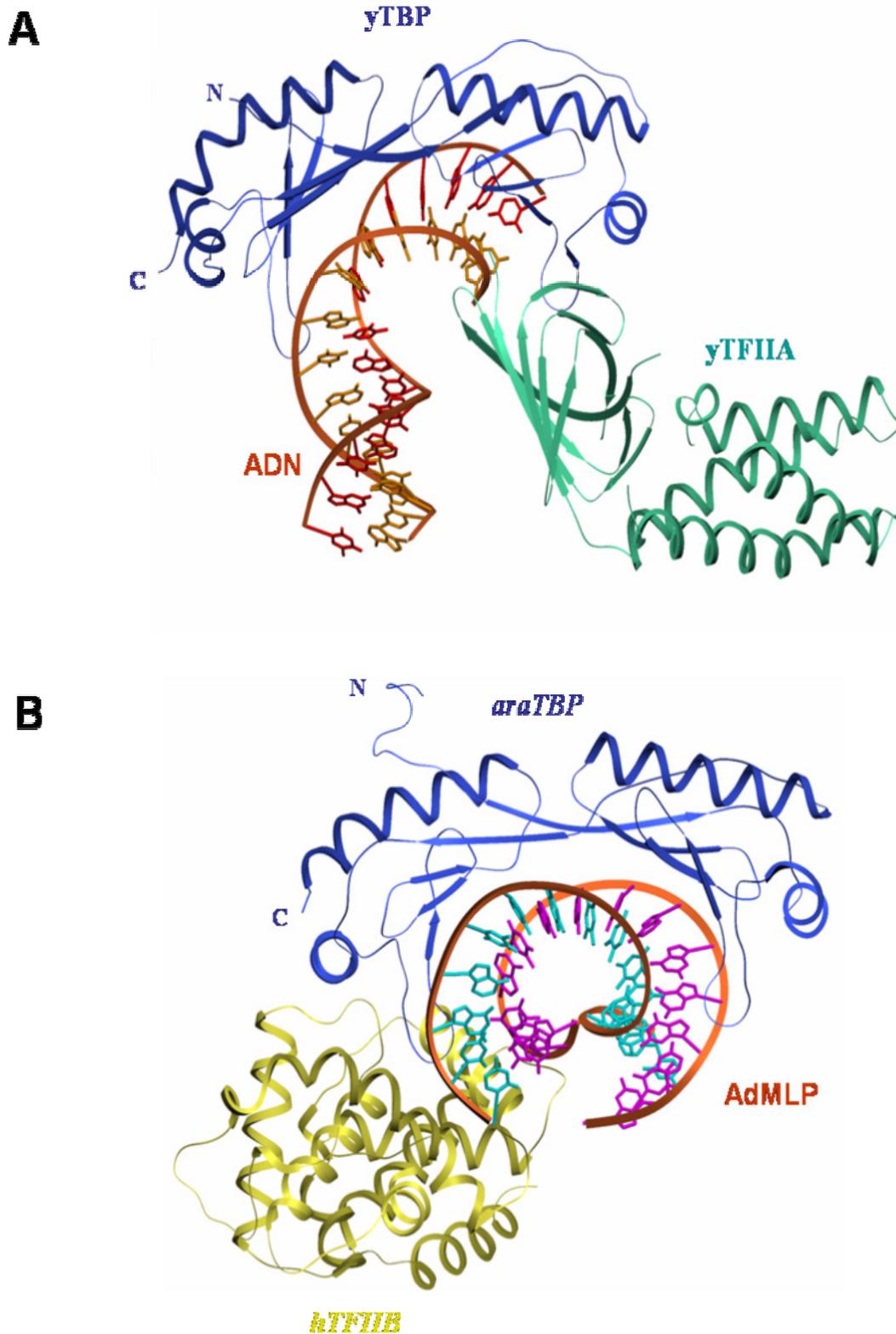


Figure 6 : Structures cristallographiques de complexes de transcription de base.

A- Structure du complexe TFIIA/TBP/ADN à 2.5 Å (Tan et al., 1996).

B- Structure du complexe TFIIB/TBP/ADN à 2.7 Å (Nikolov et al., 1995).

2.4- Complexe TBP/ADN/TFIIA

La structure cristallographique du core du facteur de transcription de levure TFIIA en complexe avec l'élément TATA du promoteur CYC1 et TBP de levure a été résolue à 2.5Å (Figure 6A). Cette structure a révélée que TFIIA se fixe du côté opposé de celui où se fixe TFIIB et qu'il adopte un repliement en « *boot-like* » composé de deux domaines : un « *β-barrel* » de 12 feuillets-β et d'un motif « *bundle* » de 4 hélices. Le « *β-barrel* » interagit avec le domaine amino-terminal de TBP via deux feuillets parallèles. TFIIA fixe l'ADN exclusivement par des contacts avec les phosphates au niveau du grand sillon de l'ADN. La fixation de TFIIA n'altère pas l'interaction TBP/ADN. Cependant, l'orientation de TBP vis-à-vis de l'ADN a été décalée de 2 paires de bases, en comparaison au complexe TBP/ADN de levure. Ce complexe TFIIA/TBP/ADN libère une surface importante pour permettre la fixation d'autres facteurs de transcription. Par exemple la jonction entre TBP et TFIIA, pourrait être la cible d'activateurs de la transcription. La structure du même complexe ternaire a été résolue par un autre groupe et abouti au même résultat (Geiger et al., 1996).

L'ensemble de ces résultats cristallographiques a permis aux auteurs d'extrapoler un modèle pour le complexe de pré-initiation de la transcription (Figure 7). Dans ces modèles, TFIIA et TFIIB sont de part et d'autres de TBP qui se retrouve au centre d'un réseau d'interaction avec TFIIA, TFIIB, l'ADN et Brf1. Ce qui confirme le rôle central joué par TBP au sein du complexe de pré-initiation, tel qu'il a été démontré au préalable par la génétique et la biologie moléculaire.

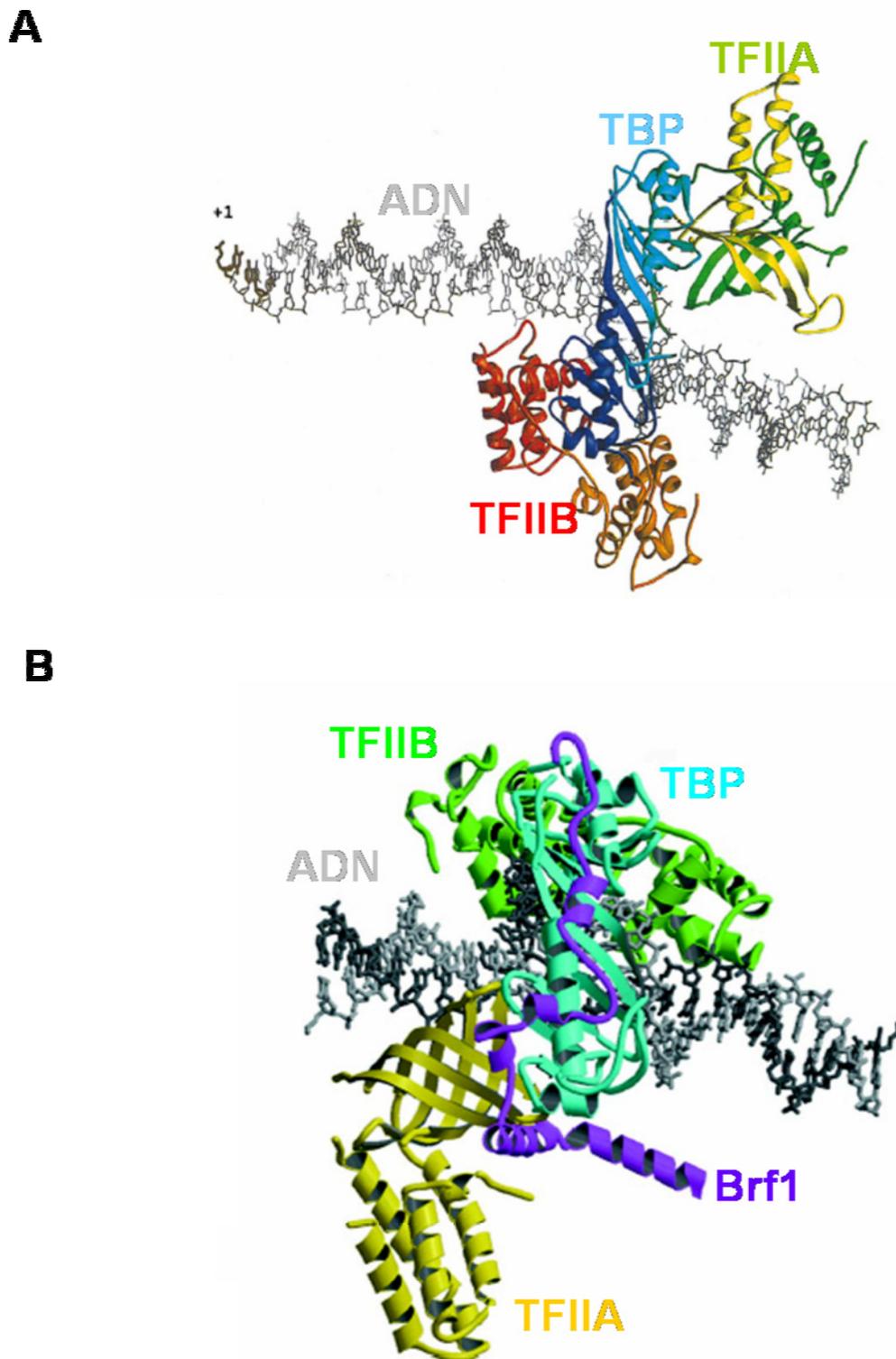


Figure 7 : Modèle du complexe d'initiation de la transcription.

A- TFIIB/TFIIA/TBP/ADN (Geiger et al., 1996).

B- TFIIB/TFIIA/TBP/Brf1/ADN (Juo et al., 2003).

2.5- Complexe TBP/TAF230

Nous avons vu dans le chapitre I, que TFIID est formé de TBP et des TAFs. La structure du complexe TBP de levure avec le bromodomaine de TAF230 (de 11 à 77) de drosophile a été résolue par résonance magnétique nucléaire et a montré que TBP adopte le même repliement que ce qui a été observé précédemment et que l'interaction avec TAF230 se fait au niveau de la région hydrophobe d'interaction avec l'ADN. La structure montre que TAF230 (de 11 à 77) est formé de 3 hélices α et un β -*hairpin*, formant un core, qui occupe la surface de fixation à l'ADN de TBP (Figure 8B).

2.6-Complexe TBP/ADN/NC2

NC2 est un cofacteur négatif qui a été isolé dans les extraits de cellules tumorales humaines qui a la capacité d'inhiber la transcription en interagissant avec le complexe binaire TBP/ADN (Inostroza et al., 1992; Meisterernst and Roeder, 1991). La structure cristallographique du complexe TBP/ADN/NC2 a été résolue à 2.6 Å et a révélée que le domaine N-terminal de NC2 α et β présentent une homologie structurale avec l'histone H2A et H2B respectivement, et forment un hétérodimère qui fixe l'ADN double hélice du côté opposé à celui où se fixe TBP et ce via des interactions électrostatiques (Figure 9). La boucle L1 de NC2 α interagit avec la partie N-terminale de TBP en faisant un certain nombre de contacts (voir le site 1, figure 9B). NC2 β contribue à l'inhibition de la transcription à travers une interaction de son hélice C-terminale avec la surface supérieure hydrophobe et conservée de TBP (voir le site 2, figure 9B), empêchant ainsi la reconnaissance du complexe TBP/ADN par le facteur de transcription TFIIB. L'homologie structurale de NC2 avec l'hétérodimère H2A/H2B pourrait permettre à NC2 de former des paires d'histones avec certains TAFs via leur motif *histone fold* et réguler ainsi la transcription.

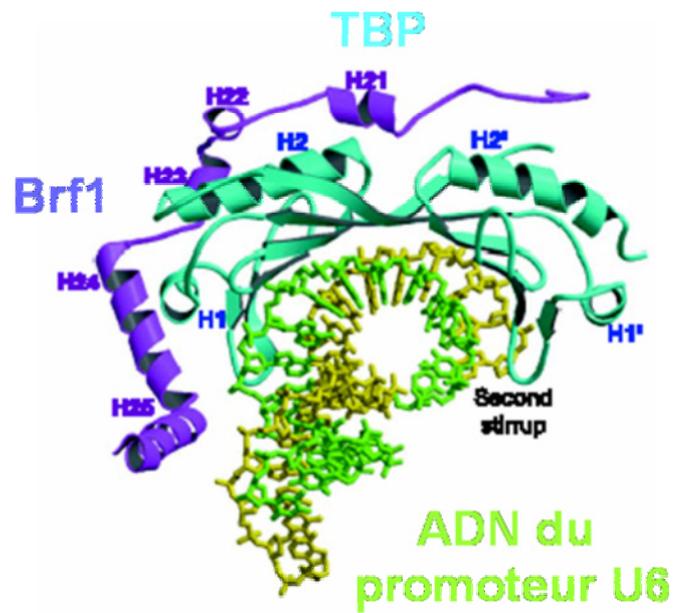
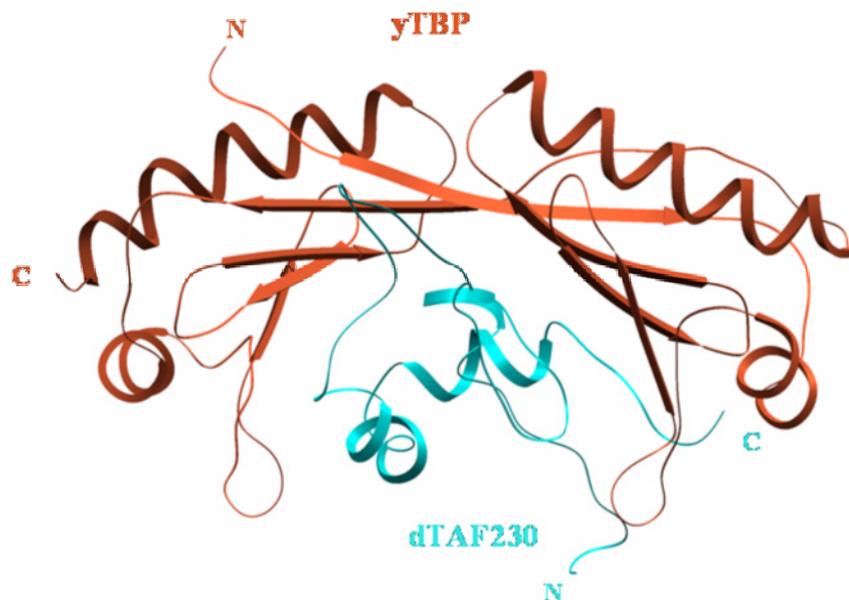
A**B**

Figure 8 : Structures cristallographique de TBP au sein de deux complexes.
A- Structure cristallographique du complexe ternaire Brf1/TBP/ADN (Juo et al., 2003).
B- Structure en solution du complexe TBP/TAF230 (de 11 à 77) par RMN (Liu et al., 1998).

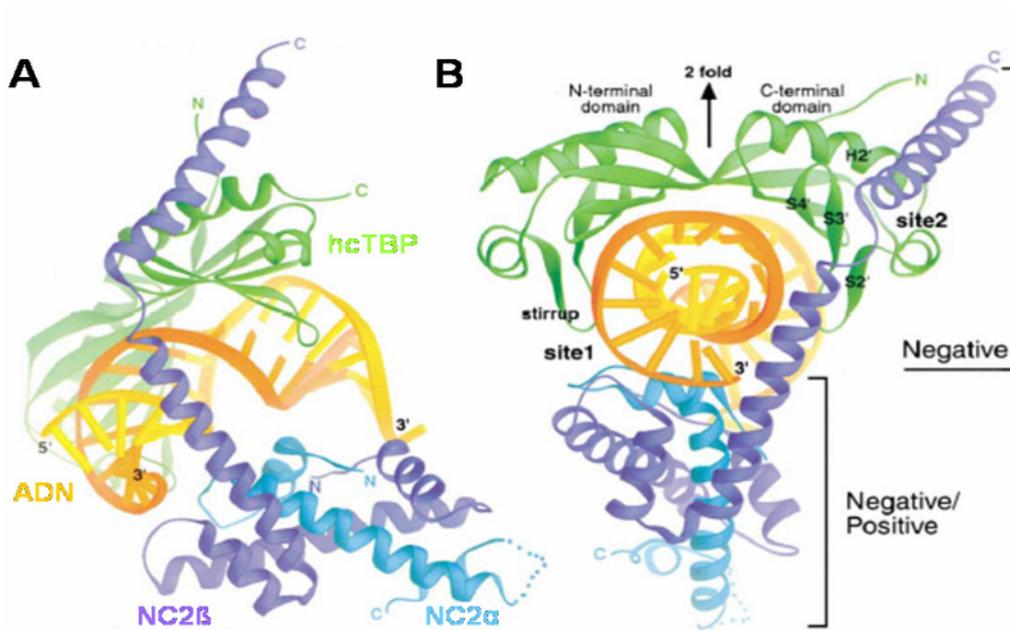


Figure 9 : Structure cristallographique du complexe ternaire TBP/ADN/NC2.

A- Le domaine C-terminal de TBP humain (en vert) en complexe avec l'ADN du promoteur tardif de l'adénovirus (en orange) et le co-facteur négatif humain (NC2).

B- Vue perpendiculaire à l'axe de symétrie d'ordre 2 de TBP montrant l'hélice H5 de NC2β avec le domaine C-terminal de TBP (site 2). La boucle L1 de NC2α interagit avec la partie N-terminale de TBP (site 1). La possibilité de régulation positive ou positive/négative est indiquée (Kamada et al., 2001).

3- Structure de l'ARN polymérase

L'ARN polymérase II est sans conteste l'enzyme clé de la transcription, requise pour la synthèse de l'ARNm à partir de l'ADN. Durant les 10 dernières années, des structures cristallographiques de l'ARN polymérase de *Thermus aquaticus* et de *Saccharomyces cerevisiae* ont été résolues. Ces structures constituent une base pour la compréhension du mécanisme de transcription.

3.1- Structure du core ARN polymérase II

La structure de l'ARN polymérase de levure a été résolue à 2.8Å (Cramer et al., 2001) Elle a permis l'identification de domaines ou modules structuraux, souvent composés de parties de différentes sous-unités (Figure 10). C'est le cas d'une paire de « mâchoires » supérieure (*upper jaw* partie Rpb1 et Rpb9) et inférieure (*lower jaw*, Rpb5) qui borde l'entrée du sillon (*cleft*, Rpb1) au fond duquel se trouvent deux ions Mg²⁺, indiquant l'emplacement du site actif. Les acides nucléiques prendraient place dans ce sillon fortement chargé positivement. A proximité du site actif, une hélice de Rpb1, l'hélice de pontage (*bridge helix*), traverse le sillon et rejoint Rpb2. Cette hélice est importante pour la translocation de l'enzyme le long de l'ADN lors de l'élongation. Proche du site actif, se trouve un canal qui s'étend au fond du sillon et qui servirait à l'entrée des rNTPs à destination du site actif.

Les structures de la polymérase libre (Cramer et al., 2001) et en association avec le duplex ADN-ARN (Gnatt et al., 2001) diffèrent principalement au niveau d'une structure mobile appelée « pince » ou *clamp*, formée du domaine d'interaction entre Rpb1 et Rpb2. Cinq boucles (*switches*) de Rpb1 sont à la base de cette « pince », et joueraient un rôle dans le changement de conformation. En effet, si les boucles 1 et 2 de Rpb1 sont peu ordonnées et la boucle 3 désordonnées dans la structure de la polymérase libre, toutes trois s'ordonnent dans la structure en élongation. Cette structure montre le trajet d'un fragment d'ADN de 13 dNTPs (de -9 à +4) appariés à 9 rNTPs. Ainsi, on peut voir que l'ADN double-brin situé en aval (+4 à +20) se retrouve dans le sillon entre la « pince » et Rpb2. En fait, l'ADN fait d'abord contact avec « la mâchoire inférieure » et passe ensuite entre le lobe de Rpb2 et une partie de la « pince » formée par Rpb1.

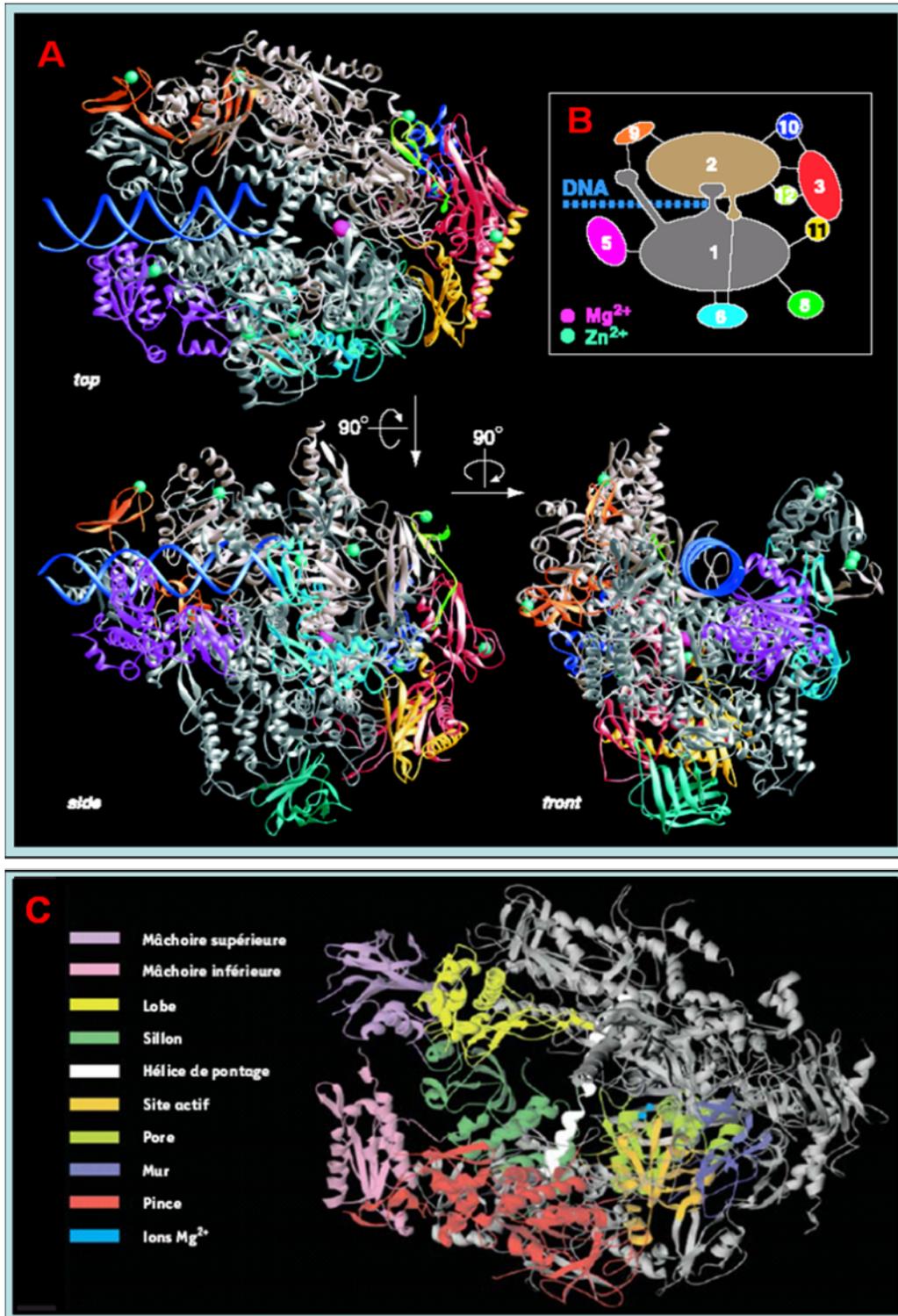


Figure 10: Structure de l'ARN polymérase II de levure à 10 sous-unités.

A-Structure cristallographique de l'ARN polymérase II de levure à 2.4 Å (Cramer et al., 2001)

B-Diagramme d'interaction des 10 sous-unités de l'ARN polymérase II de levure.

C- Description de la structure cristallographique du core ARN polymérase II.

L'enzyme contient un sillon (Rpb1) bordé à son entrée par une paire de « mâchoires », supérieure (partie de Rpb1 et Rpb9) et inférieure (Rpb5), et au fond duquel se trouvent deux ions Mg^{2+} qui indiquent l'emplacement du site actif. Près du site actif, l'hélice de pontage (Rpb1) traverse le sillon pour rejoindre Rpb2. Le canal localisé à l'arrière de l'enzyme permettrait l'entrée des rNTPs et la sortie de l'extrémité 3' libre de l'enzyme. Ces modules structuraux sont indiqués par un code couleur (Langelier., 2003).

En tout, une quinzaine de régions protéiques appartenant à Rpb1 et Rpb2 est impliquée dans des interactions avec l'hybride ARN/ADN. Cette zone contribue à la stabilité du complexe de transcription et donc à l'efficacité de la polymérase durant l'élongation. Cependant, bien que les interactions doivent permettre à l'enzyme de rester accrochée, celle-ci doit tout de même pouvoir se transloquer le long de la matrice d'ADN. Différentes observations permettent d'expliquer cela. Tout d'abord, les contacts protéiques sont faits avec le squelette sucre-phosphate et non avec les bases. Ensuite, la plupart des chaînes latérales interagissent avec deux groupements phosphates simultanément, ce qui abaisse la barrière énergétique de la translocation. Finalement, des résidus chargés positivement forment une « coquille » autour de l'hybride à une distance de 4 à 8 Å, qui le maintient dans le sillon sans bloquer le mouvement de l'enzyme.

L'hélice de pontage pourrait avoir en effet deux conformations : soit droite, soit courbée. L'hélice courbée entrerait en interaction avec le nucléotide en position +1, ce qui permettrait le déplacement des acides nucléiques lors de l'élongation de l'ARN. La résolution de la structure de l'ARN pol II en élongation a aussi permis de mieux comprendre comment l'incorporation du 11e rNTP de l'ARNm provoque la transition des cycles d'initiation avortés vers une élongation productive.

3.2- Conservation structurale

L'ARN polymérase II est une enzyme fortement conservée au cours de l'évolution. En effet, la polymérase humaine présente 53% d'identité de séquence avec celle de levure. Les résidus conservés sont distribués de manière homogène sur la structure (Figure 11A). Le pourcentage d'identité entre la polymérase de levure et de bactérie (*E.coli*) est beaucoup plus bas qu'entre la pol II de levure et celle de l'homme (Figure 11B). L'homologie de séquence de la polymérase de levure et celle de la bactérie, se situe essentiellement au niveau de la région du site actif, c'est-à-dire à l'interface d'interaction entre les deux grandes sous-unités β et β' (Rpb1 et Rpb2 chez la levure), formant un grand canal et un site catalytique (Figure 11C).

Une très grande homologie structurale est observée entre la polymérase de levure et de bactérie (Figure 11D). Elle est à la base d'une conservation du mécanisme catalytique de base entre la levure et la bactérie. De plus, la différence structurale observée à la périphérie de l'enzyme peut être corrélée avec la spécificité de recrutement des facteurs de transcription ou facteurs de régulation qui varie d'une espèce à une autre.

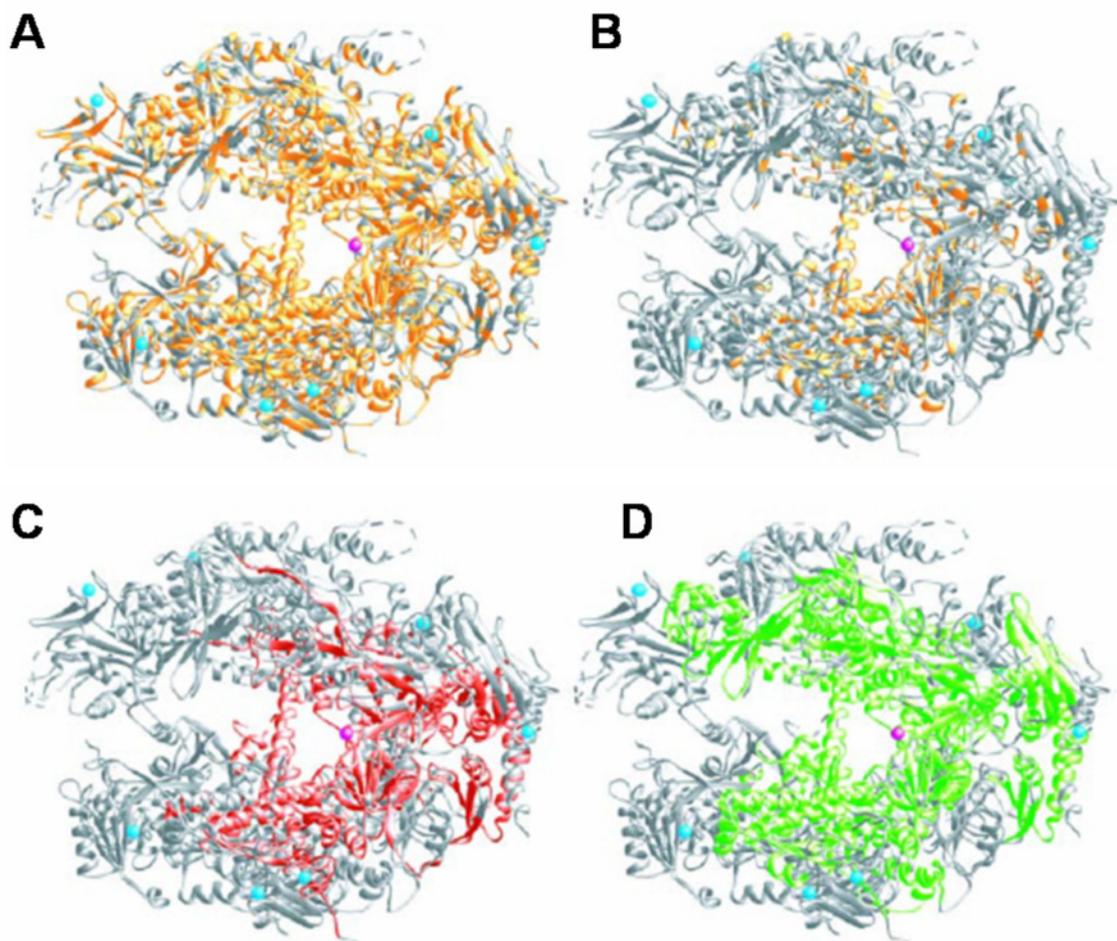


Figure 11: Conservation de séquence et de structure des ARN polymérases.

A-Représentation des résidus conservés entre l'ARN Pol II humaine et celle de levure.

Les résidus identiques sont représentés en orange.

B-Représentation de résidus identiques entre l'ARN Pol II de levure et celle de *E.coli*.

C-Homologie de séquence entre les deux plus grandes sous-unités bactérienne et de levure (en rouge).

D-Homologie structurale entre l'ARN polymérase bactérienne et celle de levure (en vert)

(Cramer et al., 2001).

3.3- Structure de l'ARN polymérase II entière

L'ARN polymérase II de levure peut être isolée sous deux formes : le core ARN polymérase formé de 10 sous-unités, qui est la forme majoritaire et la plus stable, et le complexe entier formé en plus du core d'un hétérodimère Rpb4/Rpb7. Cet hétérodimère est requis pour l'initiation de la transcription. L'étude structurale du complexe entier a été limitée par la difficulté d'obtenir un échantillon contenant le complexe Rpb4/Rpb7 à stoechiométrie équivalente aux autres sous-unités du core. L'équipe de R. Kornberg a résolu ce problème en utilisant une souche de levure modifiée permettant l'expression de l'ARN polymérase II avec une fusion *TAP-tag* (*Tandem Affinity Protein-Tag*) en C-terminus de Rpb4. Ainsi, après purification par immuno-affinité, des cristaux de la polymérase entière diffractant à 4.1Å, ont été obtenus et la structure a ensuite été résolue (Bushnell and Kornberg, 2003). Le groupe de P. Cramer a résolu la structure du complexe entier également, après purification du core ARN polymérase endogène et co-cristallisation avec l'hétérodimère Rpb4/Rpb7 recombinant surexprimé dans *E.coli*. La structure a été résolue en superposant les structures déjà déterminées de l'hétérodimère et du core polymérase à la carte de densité électronique de 4.2Å (Armache et al., 2003).

Ces deux structures montrent que l'hétérodimère Rpb4/Rpb7 maintient la polymérase dans une conformation fermée observée précédemment dans la structure du complexe core polymérase en élongation (Gnatt et al., 2001), en position d'interagir avec les facteurs d'initiation de la transcription et le médiateur. Le changement conformationnel du *clamp* est la seule différence structurale significative révélée par la comparaison du core polymérase et du complexe entier, induite par l'interaction avec Rpb4/Rpb7. Des contacts sont observés entre Rpb7 et Rpb6 et entre Rpb4 et Rpb1. La localisation de Rpb4/Rpb7 est en accord avec le modèle de cryo-microscopie obtenu sur des particules isolées et refroidies d'ARN polymérase entière (Craighead et al., 2002). Cependant, les données de microscopie et de cristallographie ne sont pas en accord sur l'orientation de l'hétérodimère. En effet, la structure en microscopie montre que Rpb4 se fixe au core-polymérase, alors que la structure cristallographique montre que c'est plutôt Rpb7 qui s'y lie (Figure 12).

3.4- Elongation de l'ARN polymérase II

Les facteurs d'élongation procaryotes Gre A et Gre B ainsi que le facteur eucaryote TFIIS appartiennent à la même famille. Ces facteurs jouent un rôle important dans la stimulation du taux d'élongation de l'ARN polymérase en la convertissant en ribonucléase 3'OH, aboutissant sur une nouvelle re-synthèse d'ARNm (Awrey et al., 1997; Rudd et al., 1994). Pour mieux comprendre, le mécanisme moléculaire qui est à la base de cette conversion et stimulation, une étude structurale des complexes ARN polymérase/GreB bactérien d'une part et ARN polymérase II/TFIIS de levure d'autre part, a été menée très récemment par les groupes de Darst et de Cramer, respectivement. La structure du complexe bactérien a été résolue par cryo-microscopie et traitement d'images de cristaux hélicoïdaux à 15 Å. La structure de haute résolution de la polymérase et celle de Gre B déterminée au préalable, ont été superposées dans la figure 13. Le domaine *coiled-coil* amino-terminal de Gre B est long de 45 Å et est situé directement dans le canal au niveau du site actif de la polymérase. Des résidus acides conservés, situés au niveau d'une boucle du *coiled-coil* de GreB interagissent avec des résidus du site actif de la polymérase ainsi qu'avec un ion magnésium catalytique, et semblent jouer un rôle clé dans l'activité de Gre B. Le rôle de ces résidus a été confirmé par des expériences de mutagenèse (Opalka et al., 2003). La structure du complexe de levure a été résolue à 3.8 Å, par diffusion du facteur TFIIS dans les cristaux de l'ARN polymérase II. TFIIS se situe à 100 Å la surface de la polymérase et interagit via un canal jusqu'au site actif. Deux résidus acides essentiels situés dans une boucle de TFIIS sont en contact avec le site actif de la polymérase et pourraient repositionner l'ion métallique et les molécules d'eau pour le clivage de l'ARN. La fixation de TFIIS provoque un changement de conformation de la polymérase, permettant un repositionnement des acides nucléiques dans le site actif. L'ensemble de ces deux études montrent qu'en dépit de la différence dans la séquence primaire et la structure de TFIIS avec Gre B, le mécanisme moléculaire de conversion de la polymérase en ribonucléase reste similaire et semble être conservé entre la bactérie et la levure (Kettenberger et al., 2003; Opalka et al., 2003). De plus, ces deux structures apportent des arguments en

faveur de l'existence d'un seul site actif pour la polymérisation et le clivage des ARNm pour l'ARN polymérase II, alors que l'ADN polymérase en possède deux.

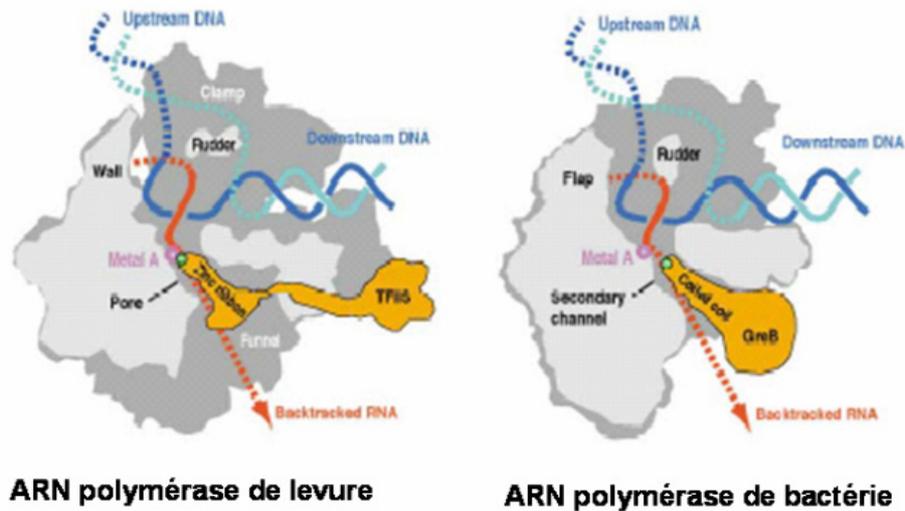


Figure 13 : Représentation schématique de l'ARN polymérase bactérienne en complexe avec Gre B et celle de levure en complexe avec TFIIS.

Les domaines de l'ARN polymérase sont indiqués. TFIIS et Gre B sont représentés en orange, l'ion Mg est indiqué en rose. La position des résidus acides conservés est indiquée en vert. Le transcrit est indiqué en rouge, et la matrice d'ADN en bleu (Conaway et al., 2003).

4- Le médiateur

Le médiateur est un complexe multi-protéique d'un méga-dalton. Chez la levure, il est composé de 20 sous-unités (Kim et al., 1994). Sa structure moléculaire a été résolue à basse résolution par microscopie électronique et a révélé une structure compacte en forme de triangle (Figure 14). La formation de l'holoenzyme aboutit sur une structure plus allongée du médiateur, formée de trois domaines : *head*, *middle* et *tail* (Figure 14). Cette structure montre que le médiateur interagit avec l'ARN polymérase II via les régions *head* et *middle*, le *tail* s'orientant vers l'extérieur de l'holoenzyme. Asturias & collaborateurs ont établi un modèle pour le changement conformationnel du médiateur, mettant en évidence une interaction entre les domaines *head* et *middle*, induite par l'arrivée de l'ARN polymérase II (Figure 15)

La structure cristallographique de l'ARN polymérase II (Cramer et al., 2001) a été superposée à la carte de densité électronique de l'holoenzyme (Davis et al., 2002) et a montré que seulement 4 domaines du médiateur contactent l'ARN polymérase : *tail/middle*, *middle*, *head 1* et *head 2*. Le *tail/middle* entre en contact avec Rpb1, Rpb3,

Rpb11 et Rpb6; le *middle* avec Rpb3 et Rpb11; le *head 1* avec Rpb2 et le *head 2* avec Rpb12. Cette étude montre que le changement conformationnel du médiateur augmente la surface d'interaction avec d'autres facteurs de transcription.

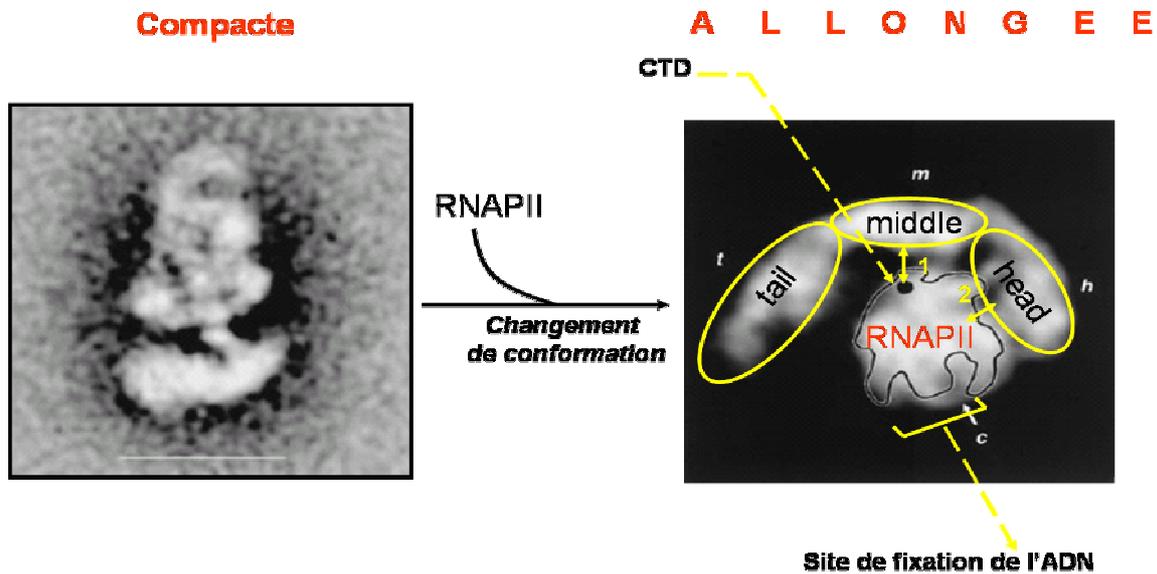


Figure 14 : Structure moléculaire du médiateur avec et sans ARN Polymérase II. Cette structure montre un changement de conformation due à l'interaction avec la polymérase. La forme allongée du médiateur contient 3 domaines *head*, *middle* et *tail*. L'ARN polymérase II, le CTD et le site de fixation de l'ADN ont été indiqués (Dotson et al., 2000).

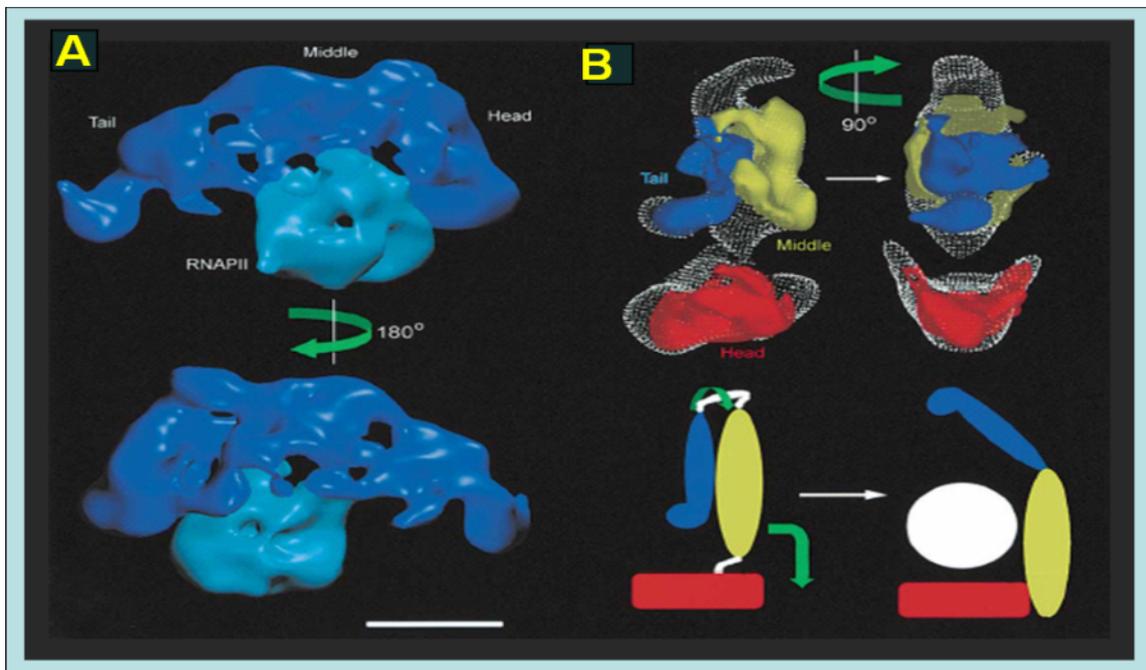


Figure 15 : Modèle du changement de conformation du médiateur induit par l'arrivée de l'ARN polymérase II.

A-Structure de l'holoenzyme.

B-Structure des domaines *head*, *middle* et *tail* du médiateur et modèle de changement de conformation. La sphère blanche représente l'ARN Polymérase II. En rouge, jaune et bleu sont représentés le *head*, le *middle* et le *tail*, respectivement (Davis et al., 2002).

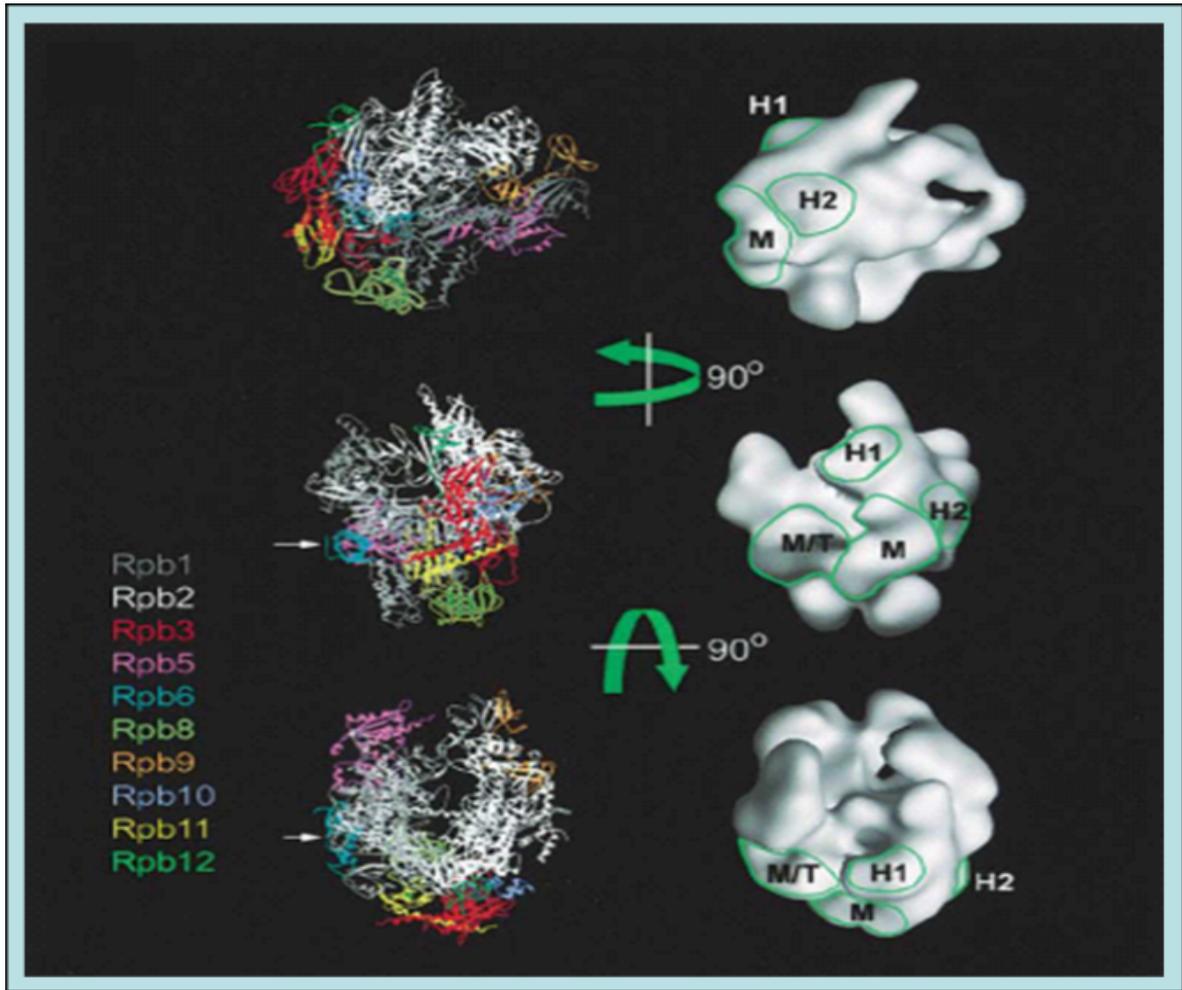


Figure 16 : Etude de l'interface d'interaction entre le médiateur et l'ARN polymérase II.

Superposition de la structure cristallographique de l'ARN polymérase II (Cramer & al., 2001) et de la structure moléculaire de l'holoenzyme (Davis & al., 2002). Les domaines du médiateur qui se fixent à la polymérase sont indiqués sur sa surface en vert. M, H1, H2 et T correspondent aux domaines *middle*, *head1*, *head 2* et *tail* (Davis et al., 2002).

5- Structures moléculaires de facteurs de transcription multiprotéiques.

L'étude cristallographique des complexes multi-protéiques humains impliqués dans la transcription est une tâche relativement difficile. En effet, cela est en partie lié à la difficulté de purifier ces complexes en quantité suffisante.

La microscopie, nécessitant des protéines à plus faible concentration a permis d'obtenir les premières images à basse résolution des facteurs TFIID, TFIIDH et du médiateur (Andel et al., 1999; Brand et al., 1999; Chang and Kornberg, 2000; Dotson et al., 2000; Schultz et al., 2000).

5.1- Le facteur TFIID

La structure de TBP décrite précédemment a été la première structure d'une sous-unité de TFIID. La première structure de paires de TAFs a été celle de l'hétérodimère TAF62/TAF42 de drosophile (Xie et al., 1996). Cette structure a montré un repliement de type histone (*histone fold*) trouvé initialement dans le core hétérotétramérique de l'octamer d'histone. Ce repliement caractéristique, trouvé dans d'autres paires de TAFs tels que TAF18/TAF28 et plus récemment TAF4/TAF12, serait impliqué dans une interaction avec l'ADN (Birck et al., 1998; Wertzen et al., 2002) (Figure 17).

Les premières images du facteur TFIID et de TFIC (*TBP-Free TAF-containing Complex*) ont été obtenues par microscopie électronique et reconstruction d'images (Figure 18). La structure de TFIID à 35 Å révèle une organisation en trois lobes reliés par des régions moins denses aux électrons, suggérant une certaine flexibilité et donc une liberté de mouvement (Brand et al., 1999). L'immuno-localisation de TBP a montré son positionnement au niveau de la cavité du lobe central (Andel et al., 1999) (Figure 18B). La structure moléculaire de TFIID humain a montré une architecture « en pince moléculaire » fortement comparable à celle de TFIID de levure (Figure 18C). Une analyse par immunomarquage montre que la majorité des TAFs se situent sur deux lobes de TFIID révélant une nouvelle organisation moléculaire de TFIID (Leurent et al., 2002). Ce qui est en opposition avec l'hypothèse d'une structure en

« octamer-like » (Selleck et al., 2001). L'étude du complexe dimérique TFIID/TFIIB ou trimérique TFIID/TFIIB/TFIIA montre que TFIIA et TFIIB se fixent sur deux lobes opposés de TFIID à proximité des régions faisant le pont. TFIIA occupe une partie de la cavité centrale et est connu pour interagir avec TBP (Figure 18B).

La structure du facteur TFIC montre une grande homologie structurale avec celle de TFIID de levure, à l'exception près que TFIC possède un domaine externe. Ce dernier de masse moléculaire approximative de 400 kDa, donne une architecture plus allongée à TFIC (Figure 18D) (Brand et al., 1999).

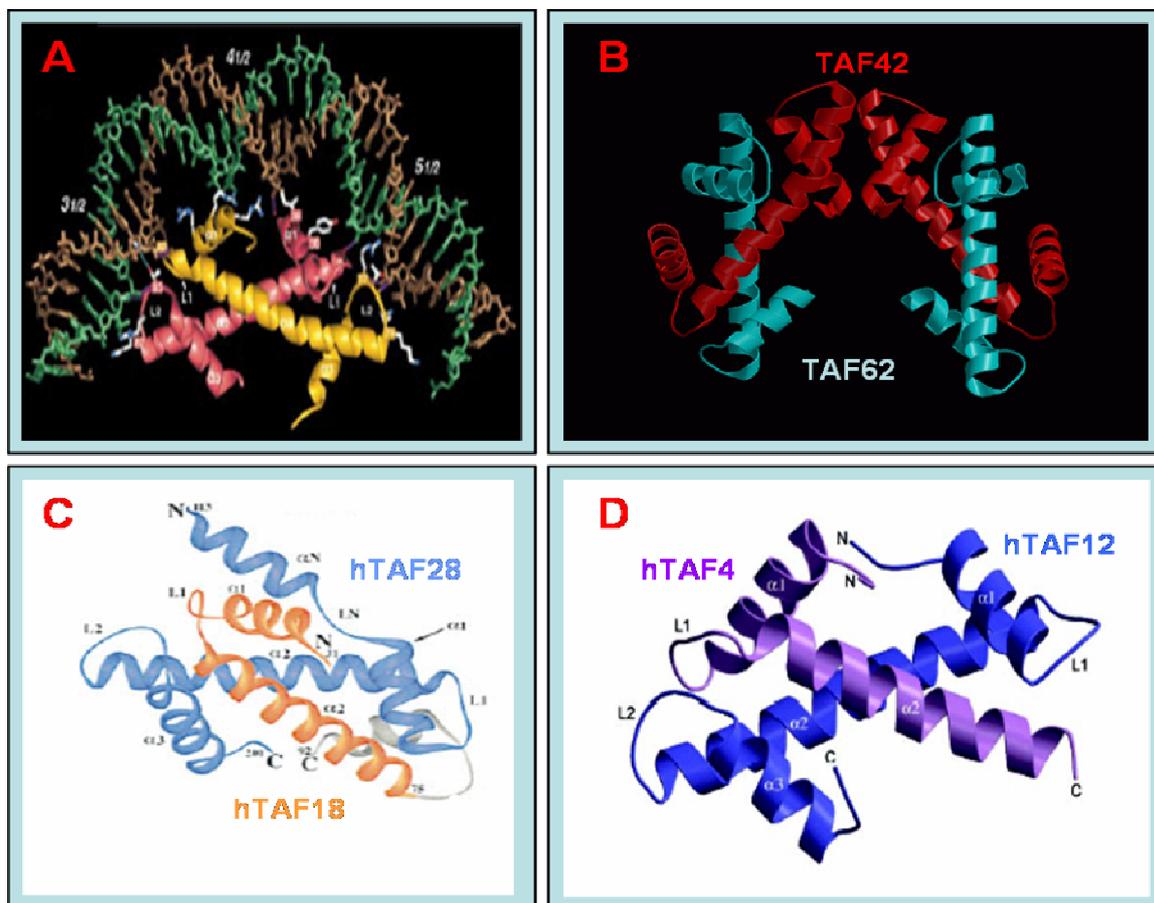


Figure 17 : Structure des *histone folds* des TAFs.

A- Structure cristallographique de paire d'*histone-fold* H2A/H2B en complexe avec l'ADN au sein de la structure du nucléosome (Luger et al., 1997).

B- Structure cristallographique de TAF62/TAF42 de drosophile (Xie et al., 1996).

C- Structure cristallographique de TAF18/TAF28 humain (Birck et al., 1998).

D- Structure cristallographique de TAF4/TAF12 humain (Werten et al., 2002).

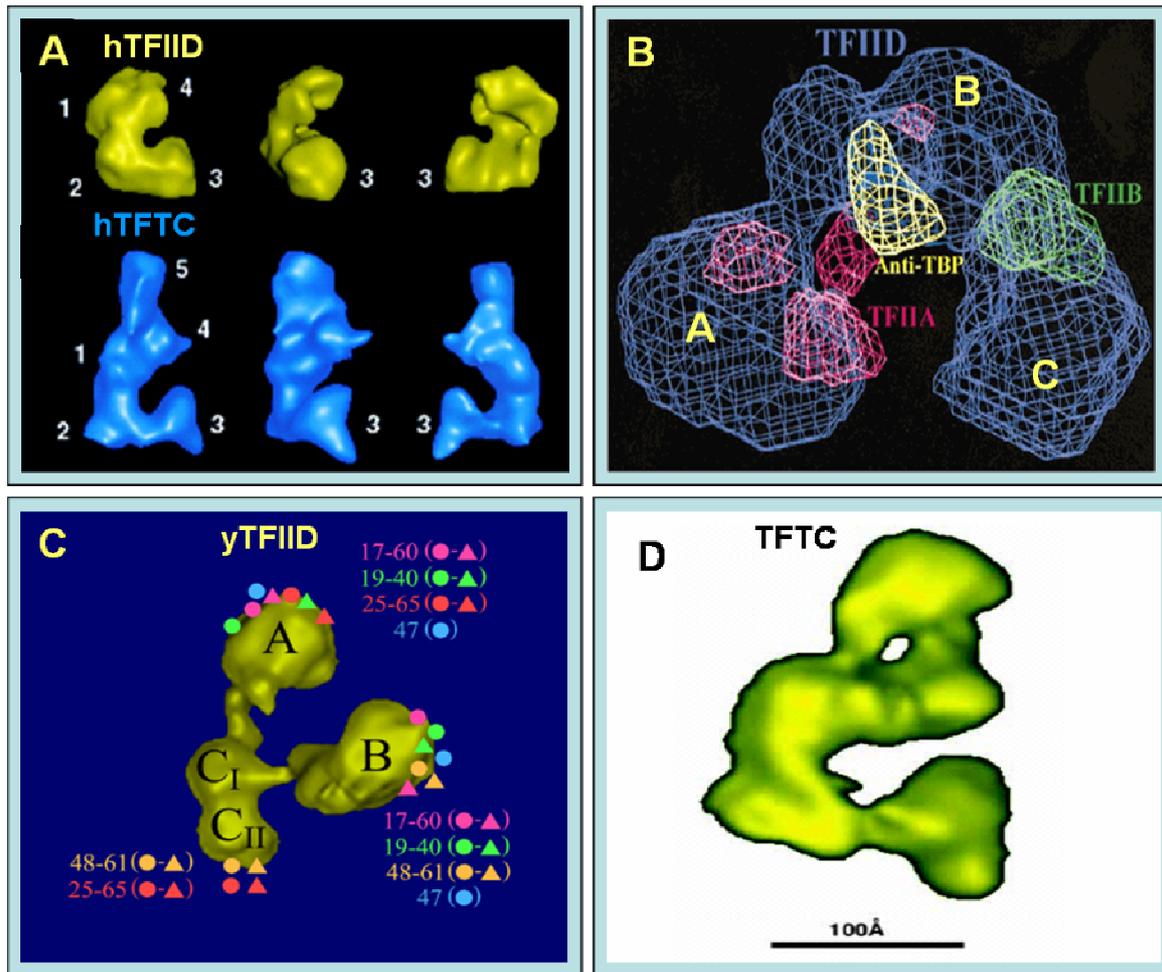


Figure 18 : Structure moléculaire des facteurs TFIID et TFIIIC.

A- Structure moléculaire de TFIID humain et de TFIIIC à 35 Å (Brand et al., 1999).

B- Structure moléculaire du complexe TFIID/IIA/IIB à 35 Å (Andel et al., 1999).

C- Structure moléculaire de TFIID de levure à 30Å (Leurent et al., 2002).

D- Structure moléculaire de TFIIIC humain à 35Å (Brand et al., 1999).

5.2- Le facteur TFIIH

TFIIH est un facteur multi-protéique composé de 9 sous-unités. Chez l'homme, il est organisé en deux sous-complexes : le core-TFIIH (p34, p44, p52, p62 et XPB) et le CAK (cdk7, cyclin H et Mat1). XPD est la sous-unité charnière entre les deux sous-complexes (Rossignol et al., 1997).

Les premières images du holo-TFIIH humain et du core-TFIIH de levure ont été obtenues par microscopie électronique en coloration négative et en cristallisation 2D, respectivement (Figure 19).

La structure moléculaire de TFIIH humain à 38 Å présente une taille 16 x12.5 x7.5 nm. Le core-TFIIH humain est organisé en anneau. Une protubérance externe observée dans l'holo-TFIIH, a été assignée au CAK. Cela a été confirmé par immunomarquage en utilisant des anticorps spécifiques contre cdk7, p44 et XPD.

La structure du core-TFIIH de levure a été résolue à 13 Å. Elle a révélé une structure en anneau organisée en trois régions distinctes : la région Tfb composée de Tfb1, Tfb2 et Tfb3 (chez l'homme p62, p52 et p34, respectivement), une région contenant Rad3 ainsi qu'une région faisant le pont entre les deux précédentes, correspondant à Ssl (Figure 19B).

La comparaison des deux structures de TFIIH humaine et de levure a montré une différence de forme et d'arrangement des sous-unités (Figure 19C). En effet, selon la structure humaine, XPD est localisé au niveau de l'anneau, ce qui n'est pas le cas de son homologue de levure. L'absence de Ssl2/XPB au sein du complexe de levure, pourrait être à l'origine de la différence structurale des deux complexes. Récemment, la composition de TFIIH de levure a été « révisé », puisque c'est Tfb4/p34 et non pas Tfb3/Mat1 qui a été montré comme appartenant au core-TFIIH de levure (Takagi et al., 2003). Ce résultat est capitale et permet de mettre en évidence une homologie structurale du core-TFIIH entre l'homme et la levure et de proposer un modèle de l'organisation moléculaire de TFIIH (voir chapitre Discussion).

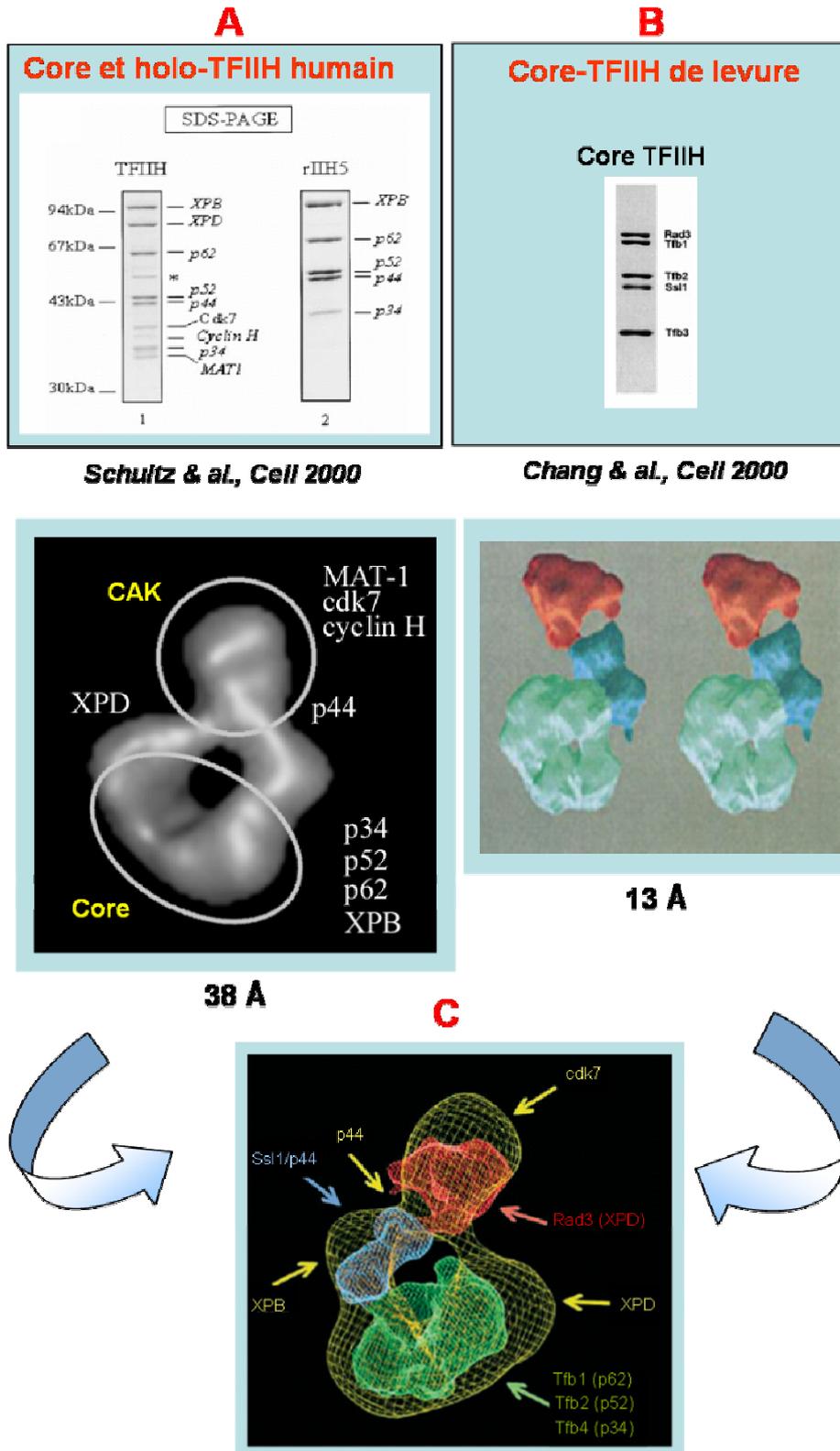


Figure 19 : Structures moléculaires du facteur TFIIH.

- A- Structure de TFIIH humain (Schultz et al., 2000).
- B- Structure du core-TFIIH de levure (Chang and Kornberg, 2000).
- C- Superposition des deux structures (Nogales, 2000).

5.3- Le facteur TFIIE

TFIIE est formé de deux sous-unités α et β de masse moléculaire respective de 34 et 56 kDa. La structure du domaine core centrale de TFIIE β , délimitée par protéolyse ménagée a été résolue par RMN (Okuda et al., 2000). Cette structure a révélée une organisation en trois hélices avec une épingle à cheveux β en extrémité carboxy-terminale (Figure 20B). Cette structure met en évidence une homologie structurale avec des protéines de type « *winged helix proteins* » et en particulier les domaines de liaison à l'ADN de DP2, HNF-3 γ et RAP30. La structure de ce domaine constitue la seule information structurale à l'échelle atomique disponible sur TFIIE. D'autre part, la structure moléculaire de TFIIE en interaction avec l'ARN polymérase II a été résolue à basse résolution (Figure 20A). Elle a permis de proposer que TFIIE interagit avec la polymérase à proximité du site actif au niveau de « la clé » induisant un changement de conformation de la polymérase (Leuther et al., 1996).

Très récemment une étude par pontage chimique et par ultracentrifugation analytique de TFIIE a été initié au laboratoire et a montré qu'en solution TFIIE se comporte comme un hétérodimère et non pas comme un hétérotétramère tel qu'il a été supposé pendant longtemps (voir chapitre V-résultats). La structure moléculaire de TFIIE à 17Å de résolution a révélé une structure allongée et une organisation en domaines distincts.

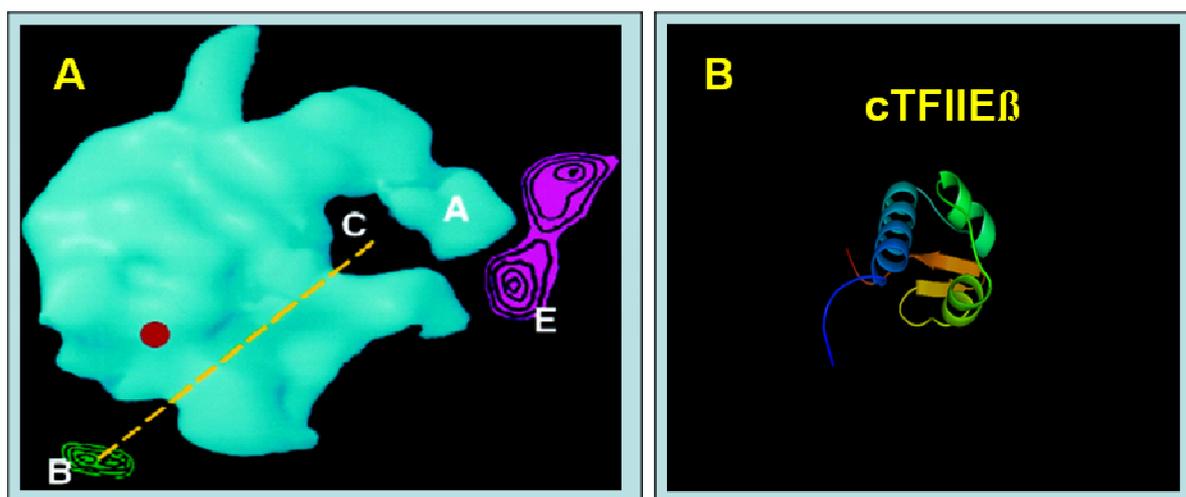


Figure 20 : Structures de TFIIE.

A- Structure de l'ARN polymérase de levure en complexe avec TFIIB et TFIIE (Leuther et al., 1996).
B- Structure du domaine core centrale de TFIIE β par RMN (Okuda et al., 2000).

Chapitre III – TFIIH, facteur multiprotéique

1- Les sous unités du *core* TFIIH

TFIIH est composé de 9 sous-unités dont le poids moléculaire varie de 34 à 89 kDa. Ce complexe multi-protéique est organisé en deux sous-complexes : le core-TFIIH et le CAK. La sous-unité XPD faisant le pont entre le core et le CAK (Table 3). TFIIH est un facteur multi-protéique impliqué dans plusieurs cellulaires importants : la transcription, la réparation de l'ADN par excision-resynthèse de nucléotides et le cycle cellulaire (Figure 21). Les sous-unités de TFIIH sont très conservées de l'homme à la levure (voir Annexes 1, 2 et 3).

<i>Composition</i>	<i>Poids moléculaire</i>	<i>Homologue de levure</i>	<i>Fonction/Caractéristique</i>
XPB	89kDa	Rad25/Ssl2	Hélicase 3'-5'
p62	62kDa	Tfb1	Réparation d'ADN
p52	52kDa	Tfb2	Ancre XPB
p44	44kDa	Ssl1	Lie l'ADN et le zinc, stimule l'activité de XPD
p34	34kDa	Tfb4	Motif de liaison au zinc
XPD	80kDa	Rad 3	Hélicase 5'-3'
Cdk7	38kDa	Kin28	Kinase
Cycline H	35kDa	Ccl1	Cycline associée à Cdk7
Mat1	33kDa	Rad5	Ancre le CAK au core

Table 3 : Composition du facteur TFIIH et rôle de ces différentes sous-unités.

Les poids moléculaires, les homologues de levure et les fonctions des différentes sous-unités sont indiqués. Les sous-unités du core-TFIIH sont représentés en bleu, les sous-unités du CAK en rouge et XPD la sous-unité charnière, en vert.

1.1- L'hélicase p89/XPB

La sous-unité XPB est un polypeptide de 762 acides aminés dont le poids moléculaire est de 89 kDa. Elle doit son nom à sa capacité de compléter les cellules de malades *xeroderma pigmentosum* de groupe B (Weeda et al., 1990a; Weeda et al., 1990b). XPB a été identifiée comme l'hélicase 5'-3' du facteur TFIIH (Schaeffer et al.,

1993). Elle possède sept domaines structuraux correspondant à des domaines hélicases, lui permettant d'avoir une activité hélicase de polarité 3'-5' couplée à une activité ATPasique ADN dépendante (Ma et al., 1994). XPB possède une séquence de localisation nucléaire (NLS), ainsi qu'un domaine de fixation à l'ADN de type hélice-tour-hélice dans sa partie amino-terminale. Des expériences de mutagenèse dirigée réalisées sur le site de fixation de l'ATP de XPB humain, montrent un effet dramatique sur la viabilité des cellules (Ma et al., 1994). De plus, certaines mutations sont à l'origine de maladies génétiques comme le *xeroderma pigmentosum*, le syndrome de Cockayne et la trichotiodystrophie (Coin et al., 1999; Hwang et al., 1996).

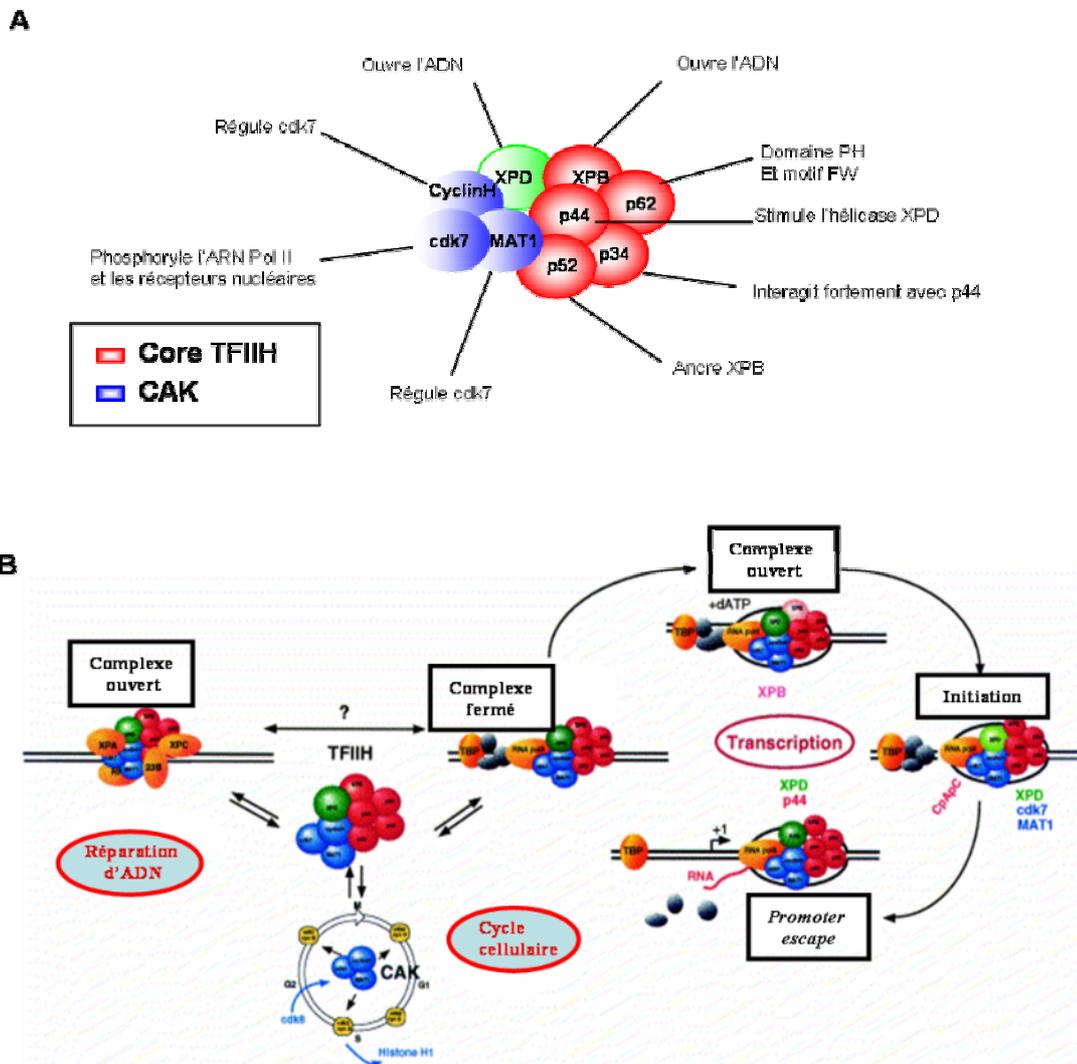


Figure 21: Schéma représentant les fonctions de TFIIH.

A- TFIIH est un facteur multi-protéique et multifonctionnel.

B- TFIIH est à l'interface de plusieurs processus cellulaires importants (Egly, 2001).

Le rôle de XPB est essentiel dans l'ouverture du promoteur lors de l'initiation de la transcription (*promoter melting*) (Holstege et al., 1996; Tirode et al., 1999), et aussi dans le déplacement de l'ARN polymérase II après la formation de la première liaison phosphodiester (*promoter escape*) (Fukuda et al., 2002; Moreland et al., 1999). Par ailleurs, XPB est impliqué au même titre que l'hélicase XPD dans l'ouverture de l'ADN autour de la lésion (Evans et al., 1997), jouant également un rôle dans l'étape d'incision en 5' (Sung et al., 1996).

1.2- La sous unité p62

p62 est un polypeptide de 548 acides aminés et d'un poids moléculaire de 62 kDa. P62 a été la première sous-unité de TFIIH à être cloné (Fischer et al., 1992). Son homologue de levure *tfb1* a été montré comme indispensable pour la viabilité de la cellule, important pour la transcription et pour la réparation de l'ADN (Gileadi et al., 1992). TFB1/p62 est impliquée dans une interaction avec Ssl1/p44 au sein de TFIIH et une défaillance dans cette interaction entraîne une hypersensibilité au rayonnement ultraviolet (Matsui et al., 1995). p62 semble être important pour le mécanisme de réparation par NER chez la levure et la drosophile (Castro et al., 2002; Wang et al., 1995).

p62 est la cible d'un ensemble de protéines régulatrices de la transcription, telle que VP16, p53, E2F-1, HBx, EBNA2 et FBP ou encore le récepteur à l'œstrogène ER α (Chen et al., 2000; Vandel and Kouzarides, 1999; Xiao et al., 1994).

Récemment, un domaine conservé appelé BSD (pour *BTF2-like transcription factors*, *Synapse-associated and DOS2-like protein*) a été localisé dans la partie médiane de p62 (Doerks et al., 2002). Ce domaine est constitué d'un motif conservé de répétition FW (phénylalanine et tryptophane).

P62 est organisé en trois domaines : amino-terminal, médian et carboxy-terminal (Figure 22). La structure du domaine amino-terminal de p62 a été résolue par résonance magnétique nucléaire et révèle un repliement similaire à celui des protéines de type PH et PTB (pour *Plekstrin Homology* et *Protein Tyrosine Binding proteins*) (voir chapitre-V résultats). Nous avons montré que ce domaine PH n'est pas

important pour l'architecture de TFIIH, puisqu'une délétion de ce domaine n'empêche pas les autres sous-unités de TFIIH de s'associer et former le complexe. Des tests effectués *in vitro*, montre que ce complexe muté est actif en transcription, mais pas en réparation d'ADN, ce qui suggère un rôle spécifique de ce domaine dans la réparation d'ADN. Nous avons ensuite montré que ce domaine interagit avec XPG, facteur de réparation impliqué dans l'incision en 3' (voir chapitre-V résultats). Le domaine carboxy-terminal est important pour l'architecture de TFIIH, puisqu'il permet une interaction avec le complexe p34/p44.

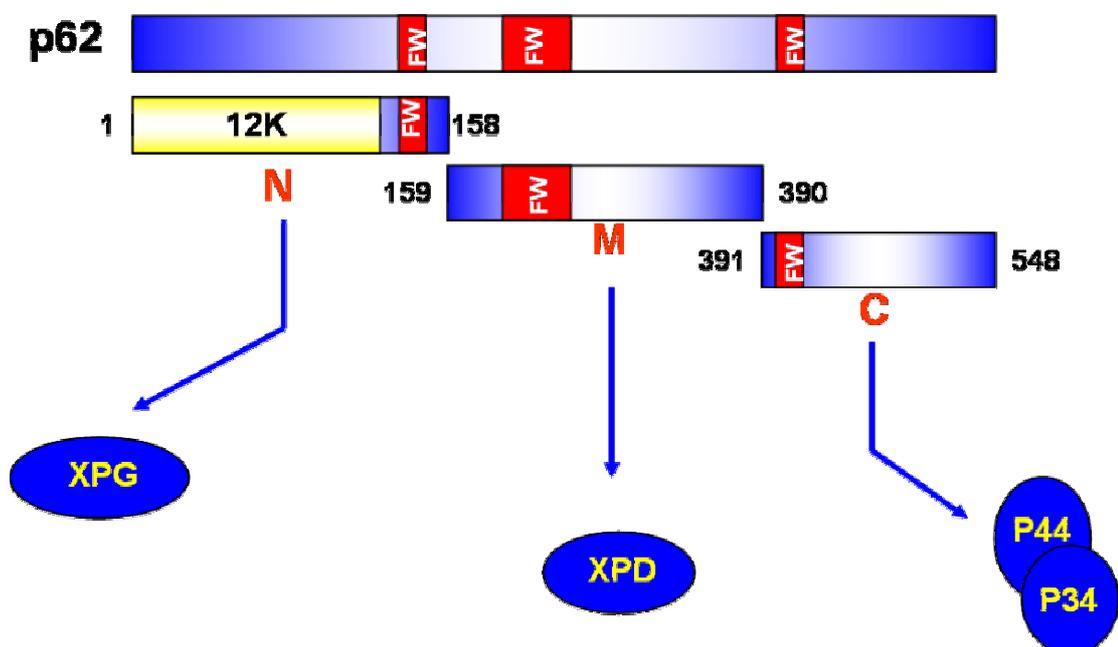


Figure 22 : Représentation schématique des domaines de p62.

Les motifs FW de p62 sont indiqués en rouge. Le domaine N-terminal de 12 kDa est représenté en jaune. XPG, XPD et le complexe p34/p44 interagissent avec le domaine N-terminal (1-158), médian (159-390) et C-terminal (391-538), respectivement..

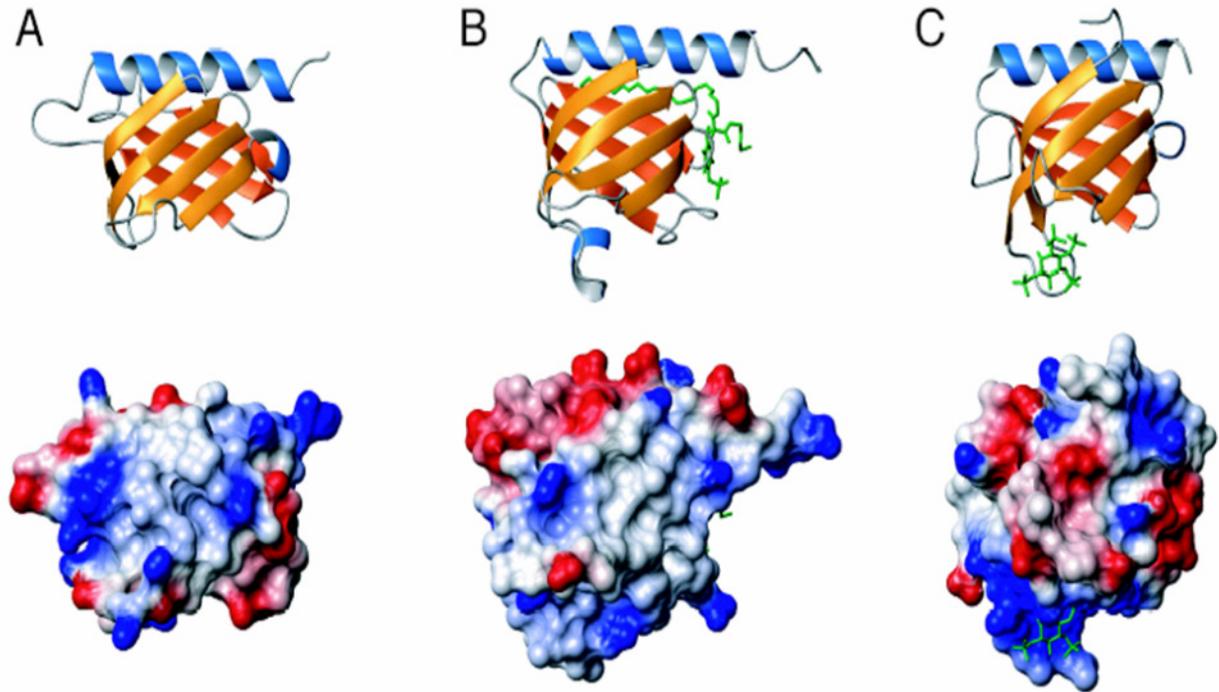


Figure 23 : Homologie structurale entre p62(1-109) et le domaine PTB ou PH.

A-Structure RMN du domaine amino-terminal de p62 révélant un repliement de type PH/PTB.

B-Structure du domaine PH (pour *Pleckstrin Homology*) de Dapp1/Phish en complexe avec 1,3,4,5-tetrakisphosphate.

C-Structure du domaine PTB (pour *Protein Tyrosine Binding*) du récepteur à l'insuline 1 en complexe avec le phospholipide du récepteur de l'interleukin 4.

La surface électrostatique de chaque protéine est indiquée en dessous de la protéine correspondante (les résidus positifs en bleu et négatifs en rouge) (Article soumis).

1.3- La sous unité p52

Il s'agit d'un polypeptide de 436 acides aminés et d'un poids moléculaire de 52 kDa. C'est la dernière sous-unité clonée de TFIIH (Marinoni et al., 1997). Son homologue de levure *tfb2* est nécessaire à la viabilité de la cellule et une délétion carboxy-terminale aboutit à une sensibilité au rayonnement ultraviolet (Feaver et al., 1997). Nous avons reproduit cette mutation dans le système reconstitué de baculovirus, et avons montré que cette mutation affecte l'ancrage et la fonction de l'hélicase XPB. En effet, cette mutation induit une instabilité de XPB, qui est à l'origine de la diminution de l'activité de transcription et de réparation de TFIIH (Jawhari et al., 2002a). p52 possède deux domaines d'interaction avec XPB ainsi qu'un domaine sensible aux protéases (Figure 24).

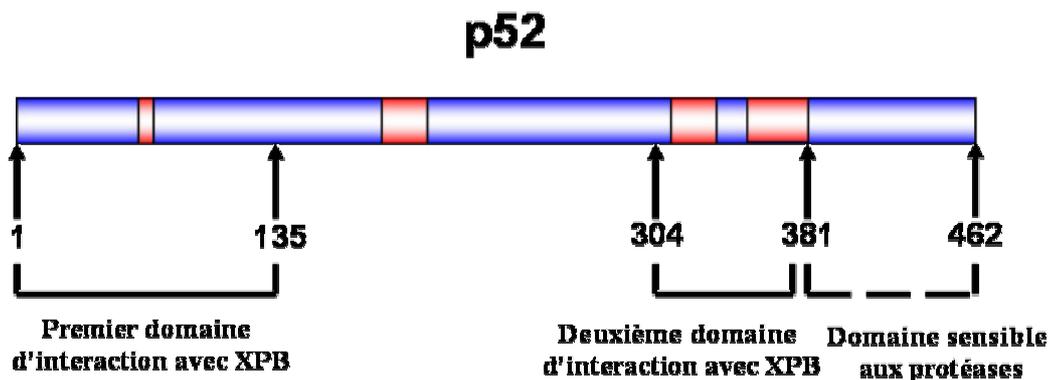


Figure 24 : Représentation schématique de p52.

En rouge sont représentées les régions conservées de l'homme à la levure. p52 possède deux domaines d'interaction à XPB. Le domaine C-terminal est sensible aux protéases. (Jawhari et al., 2002a).

1.4- La sous unité p44

p44 est un polypeptide de 395 acides aminés pour un poids moléculaire de 44 kDa (Humbert et al., 1994). Il existe deux copies de gène *p44* dans le génome humain, une copie localisée dans la région télomérique et l'autre dans la région centromérique du même chromosome 5 (Humbert et al., 1994). Son homologue de levure *Ssl1* (*Suppressor of Stem Loop 1*) est capable de stimuler la traduction d'un ARN possédant une structure en épingle à cheveux en début de séquence (Yoon et al., 1992). Des

mutations de ce gène chez la levure sont responsable d'une hypersensibilité au UV et une déficience de l'activité transcriptionnelle (Wang et al., 1995). Au sein de TFIIH, p44 est au centre d'un réseau d'interaction notamment avec p34 impliquant son domaine carboxy-terminale (Fribourg et al., 2001). La structure du domaine C-terminal de p44(321-395) a été résolue par résonance magnétique nucléaire et révèle un domaine *ring-finger* non canonique (*ring finger like*) avec deux motifs de liaison au zinc : un de type C4 et un de type CX2CX2-4FCADCD, et une grande homologie avec la protéine kinase C (Fribourg & al., 2000) (Figure 25). Des mutations ponctuelles dans ce domaine carboxy-terminal altèrent l'activité de TFIIH lors de l'échappée du promoteur (*promoter escape*) (Tremeau-Bravard et al., 2001). Le domaine amino-terminal de p44 est impliqué dans une interaction avec le domaine carboxy-terminal de l'hélicase XPD (Coin et al., 1999; Seroz et al., 2000). Il s'agit d'une interaction fonctionnelle puisqu'elle est à l'origine d'une stimulation de l'activité hélicase de XPD (Coin et al., 1999; Seroz et al., 2000). Par ailleurs, Ssl1/p44 interagit avec tfb1/p62 chez la levure et cela semble être indispensable à la viabilité de la cellule (Matsui et al., 1995). La même interaction a été décrite chez l'homme suite à des expériences de co-expression et co-purification (Jawhari et al., 2002b).

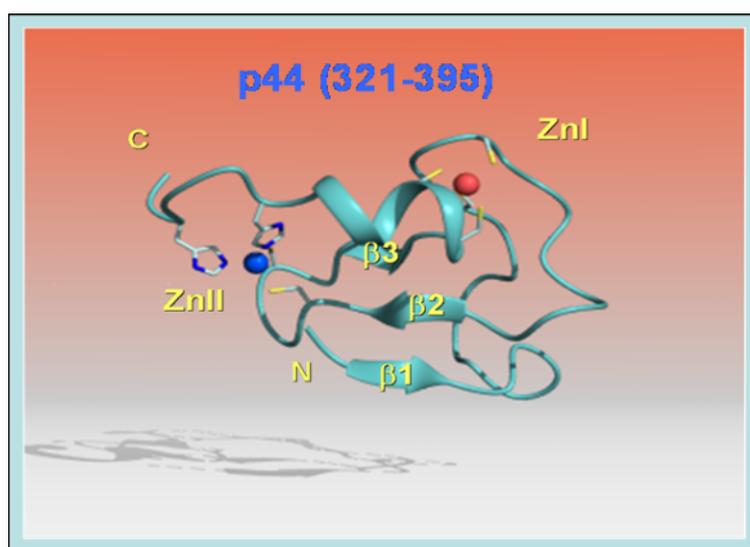


Figure 25 : Structure RMN de p44 (321-395).
En rouge et bleu sont représentés les ions de Zn^{2+} I et II, respectivement (Fribourg et al., 2000).

1.5- La sous unité p34

p34 est un polypeptide de 303 acides aminés et d'un poids moléculaire de 34 kDa (Humbert et al., 1994). Son homologue chez la levure *tfb4* est indispensable pour la viabilité de la cellule (Feaver et al., 1997). Il possède un domaine de liaison au zinc correspondant au motif de type C4 de p44. Il a été montré que *tfb4* est indispensable aussi bien lors de la transcription que lors de la réparation d'ADN par N.E.R (Feaver et al., 1997). Le domaine amino-terminal de p34 est impliqué dans une interaction avec le domaine carboxy-terminal de p44 (voir chapitre-V résultats p34/p44). Récemment, nous avons découvert qu'une délétion du domaine carboxy-terminal de p34 provoque une absence d'incorporation de p52 (Voir chapitre autres résultats). Ceci renforce le rôle architectural de p34 au sein du core-TFIIH.

2- Les sous unités du CAK

Le CAK (*Cyclin dependant kinase activating kinase*) est un sous-complexe stable qui peut exister sous deux formes *in vivo* : sous forme libre ou associé au core-TFIIH (Drapkin et al., 1996; Rossignol et al., 1997). Il est composé de *cdk7*, cyclin H et Mat1.

2.1- La sous-unité Cdk7

Cdk7 est un polypeptide de 346 acides aminés et d'un poids moléculaire de 40 kDa (Tassan et al., 1994). Son homologue chez la levure est Kin 28 (Simon & al., 1986) et MO15 chez le Xénope (Fisher and Morgan, 1994). Cdk7 est une kinase avec un domaine de liaison à l'ATP de type GXGKT, et possède une boucle T phosphorylable caractéristique des kinases (Figure 29A). Au sein de CAK, l'activité kinase de *cdk7* est régulée par la cycline H (Nigg, 1996). Cdk7 est capable de phosphoryler le domaine carboxy-terminal de l'ARN polymérase II (Roy et al., 1994; Schneider et al., 1998) et aussi phosphoryler d'autres protéines de la famille des récepteurs nucléaires (Bastien et al., 2000; Chen et al., 2000; Keriél et al., 2002; Rochette-Egly et al., 1997).

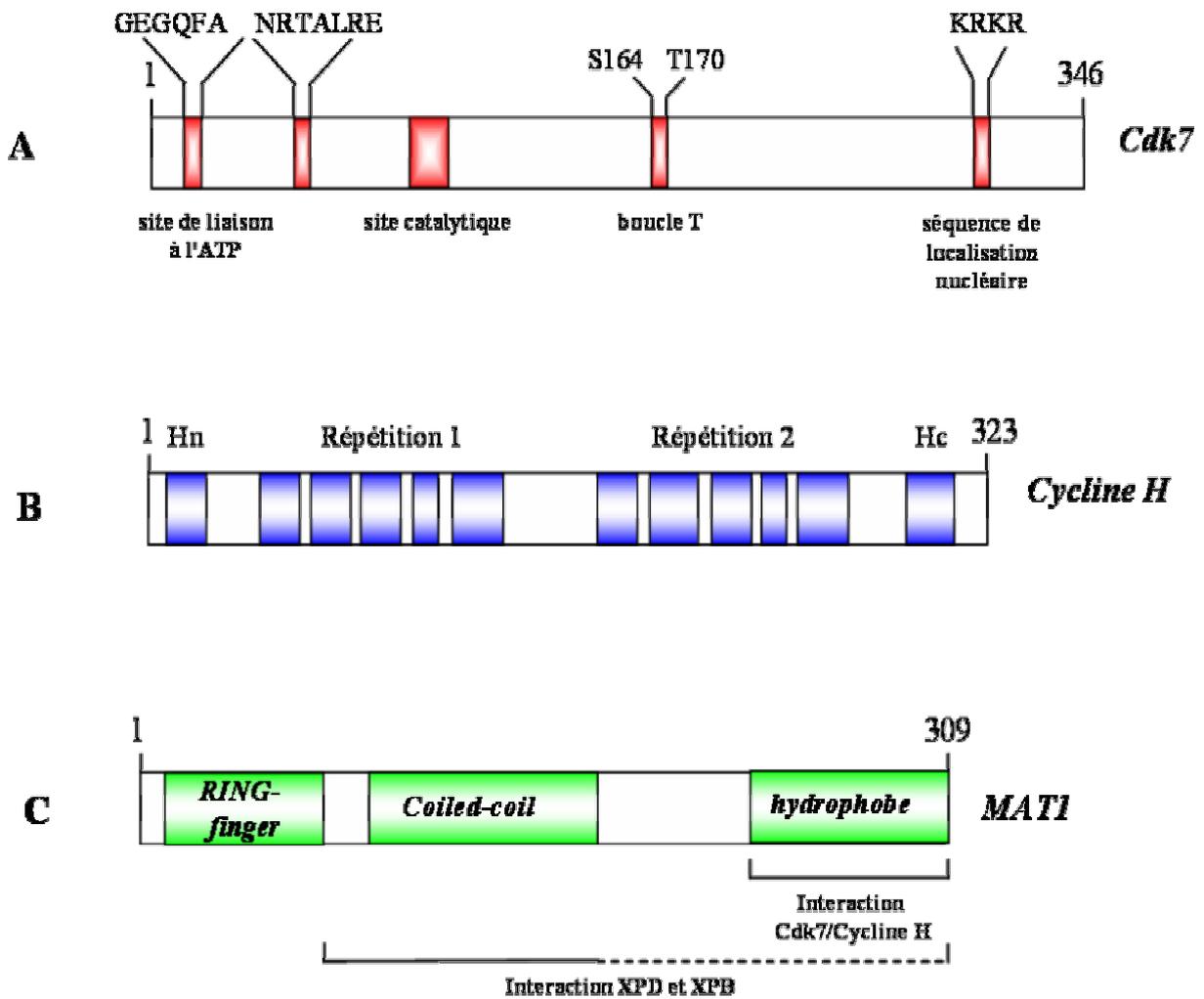


Figure 26: Représentation schématique des sous-unités du CAK.

A-Pour *cdk7*, les carrés rouge représentent le site de fixation de l'ATP, le site catalytique, la boucle T et la séquence de localisation nucléaire.

B-Pour la cycline H, les carrés bleu représentent les hélices α . Les répétitions 1 et 2 contiennent chacune 5 hélices et adoptent un repliement caractéristique du type *cyclin fold*. Hn et Hc correspondent à des hélices amino et carboxy-terminales, respectivement.

C-Pour MAT1, les carrés verts représentent le *ring-finger*, le *coiled-coil* et le domaine hydrophobe. Les zones d'interactions avec le complexe binaire *cdk7/cyclinH*, XPD ou XPB sont indiquées.

2.2- La sous-unité Cyclin H

La cycline H est un polypeptide de 323 acides aminés (Makela et al., 1994), d'un poids moléculaire de 38 kDa. La cycline H contient deux domaines répétés d'une centaine d'acides aminés qui forment deux séries symétriques de 5 hélices α (Figure 26B). Il s'agit d'une boîte cycline adoptant un repliement de type « *cyclin fold* » caractéristique de toutes les cyclines (Brown et al., 1995; Russo et al., 1996). La structure cristallographique de la cycline H a été résolue à 2.6 Å et a révélé une grande similitude avec une autre cycline, la cycline A, montrant une bonne conservation du repliement canonique « *cyclin fold* » (Andersen et al., 1997; Andersen et al., 1996; Kim et al., 1996) (Figure 27A). De plus la cyclin A est capable d'activer cdk7 en présence de Mat1 et la cycline H peut interagir avec cdk2, ce qui confirme l'homologie fonctionnelle suggérée par cette structure. La structure du complexe binaire cdk2/cyclinA montre que la cycline A interagit avec cdk2 par l'intermédiaire des 2 lobes de la cycline A (Jeffrey et al., 1995) (Figure 27B).

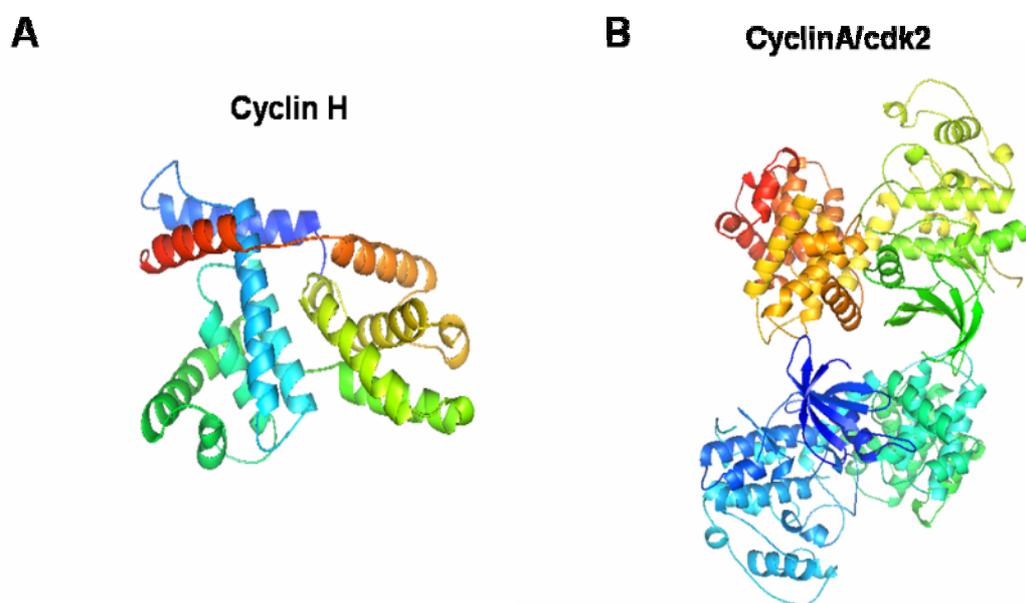


Figure 27 : Structure cristallographique de la cycline H et du complexe cycline A/cdk2.

A-Structure de la cycline H (Andersen et al., 1997).

B-Structure du complexe cyclineA/cdk2 (Jeffrey et al., 1995).

2.3- La sous-unité Mat1

Mat1 (pour *Ménage-à-Trois 1*) est un polypeptide de 312 acides aminés (Tassan et al., 1995; Yee et al., 1995), correspondant à une protéine de 32 kDa. Mat1 est organisé en trois domaines : un domaine *ring-finger* dans sa partie amino-terminale, un domaine *coiled-coil* dans sa partie médiane et un domaine hydrophobe en carboxy-terminal (Busso et al., 2000; Tassan et al., 1995) (Figure 26C). Le domaine hydrophobe sert à recruter le complexe binaire cdk7/cyclinH. Mat1 ancre le CAK au core TFIIH par l'intermédiaire des hélicases XPB et XPD (Busso et al., 2000). Mat1 régule l'activité kinase de cdk7 et notamment stimule sa capacité à phosphoryler *in vitro* un peptide mimant le domaine carboxy-terminale de l'ARN polymérase II (Rossignol et al., 1997) ou encore permet à la kinase de phosphoryler les facteurs p53 et Oct1 (Inamoto et al., 1997; Ko et al., 1997). Quant au domaine amino-terminal, la structure a été résolue par résonance magnétique nucléaire et a révélée un repliement entrecroisé caractéristique des *ring-fingers* (Gervais et al., 2001). Il se compose de deux hélices α et de trois feuillets β correspondant à un repliement $\beta\beta\alpha\beta$ caractéristique des *ring-fingers* et possède une hélice α supplémentaire (Figure 28). Deux sites de liaisons au zinc ont été décrits et plusieurs résidus conservés ont été identifiés et notamment : 4 résidus hydrophobes (Leu 19, Leu 21, Met 22 et Val 23), 4 résidus chargés positivement et localisés à la surface du domaine (Lys 20, Arg 54, Lys 55 et Arg 59) et une tyrosine strictement conservée située à la surface du domaine (Figure 28).

Ring finger de MAT1

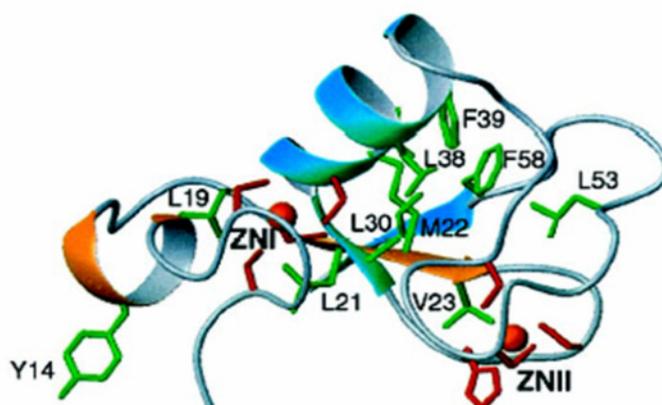


Figure 28 : Structure RMN du domaine *ring-finger* de MAT 1

Les ions Zn^{2+} sont indiqués par des sphères rouges. Les résidus qui interagissent avec les ions Zn^{2+} sont représentés en rouge et les résidus conservés en vert (Gervais et al., 2001).

3- La sous-unité p80/XPD

XPD est un polypeptide de 761 acides aminés et d'un poids moléculaire de 80 kDa. Des gènes codant pour des homologues de XPD ont été caractérisés chez *Schizosaccharmyces. pombe* (Murray et al., 1992), chez le hamster (Kirchner et al., 1994) et chez le poisson *Xiphophorus maculatus* (Della Coletta et al., 1995). XPD possède une activité hélicase de polarité (5'-3') opposée à celle de XPB, dépendante de l'ATP (Schaeffer et al., 1994; Sung et al., 1993). XPD a été caractérisée de la même manière que XPB, c'est-à-dire pour sa capacité à compléter les cellules de malades du groupe *Xeroderma pigmentosum* de groupe D (Wang et al., 1994). Le rôle de XPD dans la transcription semble être architectural, puisque des mutations dans son site ATPase ou hélicase n'affectent presque pas l'activité transcriptionnelle de TFIIH *in vitro* (Coin et al., 1999; Tirode et al., 1999) et *in vivo* (Winkler et al., 2000). Cependant, de telles mutations semblent être dramatiques pour l'activité de réparation de TFIIH ; ce qui suggère fortement que XPD est l'hélicase indispensable pour la réparation alors que XPB est indispensable pour la transcription (Coin et al., 1998; Sung et al., 1996). Néanmoins, la délétion totale du gène est létale aussi bien chez la levure que chez la souris (de Boer et al., 1998), suggérant que le processus de transcription nécessite la présence de XPD au sein de TFIIH.

Au sein de TFIIH, XPD est associée à la sous-unité p44 et cette association stimule son activité hélicase de 10 fois (Coin et al., 1998) faisant de p44 la sous-unité régulatrice de XPD au sein du core. La sous unité Mat1 peut aussi interagir avec XPD (Busso et al., 2000), et servirait de sous unité charnière entre la core-TFIIH et la CAK. Il a été d'ailleurs montré que XPD peut appartenir soit au complexe core TFIIH soit au complexe CAK (Rossignol et al., 1997).

De la même manière que *XPB*, des mutations dans le gène *XPB* peuvent être à l'origine de maladies génétiques rares comme le *Xeroderma pigmentosum*, le syndrome de Cockayne et la trichothiodystrophie (voir chapitre IV), qui se caractérisent par une déficience dans les fonctions de transcription et/ou de réparation de l'ADN. Des mutations dans la partie N-terminale de XPD peuvent être à l'origine d'un problème dans l'ancrage du CAK au core, aboutissant ainsi à une déficience dans la phosphorylation du récepteur nucléaire RAR α (Keriel et al., 2002).

Un modèle structural de XPD a été établi sur la base de comparaison de séquence et de prédiction de structure secondaire avec l'hélicase bactérienne UvrB (Figure 29). D'après ce modèle, XPD présente une grande homologie structurale avec UvrB. XPD possède 7 motifs hélicases bien conservés, un site de fixation de l'ATP et est organisé en deux domaines hélicases amino et carboxy-terminal (Bienstock et al., 2003).

Une étude par mutagenèse de XPD a montré l'existence de deux types de mutation : des mutations qui affectent l'activité hélicase de XPD et par conséquent la réparation par NER, ainsi que des mutations qui affectent l'interaction p44/XPD et par conséquent la transcription. Ce deuxième type de mutation serait responsable de la trichothiodystrophie (Dubaele et al., 2003).

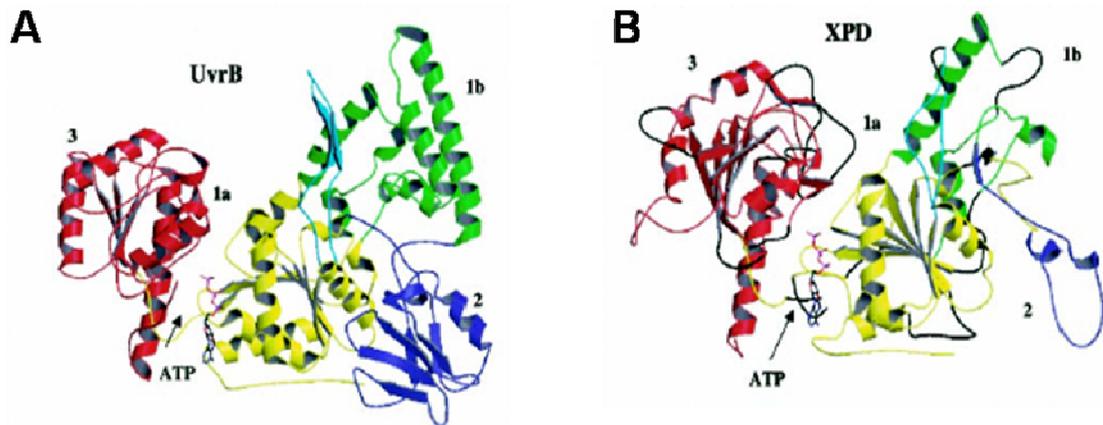


Figure 29 : Structure de Uvr-B et modèle de XPD

A- Structure cristallographique de Uvr-B (Theis et al., 1999).

B- Modèle de XPD basé sur la structure d'Uvr-B. Les couleurs des différents domaines Traduisent l'homologie structurale des deux protéines (Bienstock et al., 2003).

Chapitre IV - TFIIH à l'interface de processus cellulaires fondamentaux

1- Rôle de TFIIH dans la transcription de base

TFIIH, comme son nom l'indique, est un facteur générale de transcription des gènes de classe II (Gerard et al., 1991). Récemment, un rôle de TFIIH dans la transcription des gènes de classe I codant pour les ARNs ribosomales, a été démontré (Iben et al., 2002). Les activités hélicases de sens opposés des sous-unités XPB et XPD ainsi que l'activité kinase de cdk7, permettent à TFIIH de jouer un rôle complexe dans la transcription et notamment dans l'ouverture du promoteur et le départ de l'ARN polymérase II (Coin and Egly, 1998; Egly, 2001).

1.1- Rôle de TFIIH dans l'ouverture du promoteur

Les étapes précoces de la transcription et notamment l'ouverture du promoteur sur 11 paires de bases de -9 à +2 de part et d'autre du site d'initiation (*promoter opening*), requièrent l'ATP comme source d'énergie (Dvir et al., 1996; Holstege et al., 1996; Kugel and Goodrich, 1998; Timmers, 1994). L'hydrolyse de l'ATP est assurée par les hélicases de TFIIH. Afin d'étudier plus précisément les mécanismes moléculaires de TFIIH, un système reconstitué permettant l'expression des différentes sous-unités ainsi que différentes versions mutées des hélicases XPB et/ou XPD a été mis au point en utilisant les cellules d'insectes infectées par des baculovirus recombinant (Coin et al., 1999; Tirode et al., 1999). Ces études ont montré que XPB est indispensable pour l'étape d'ouverture du promoteur et que XPD stimule de manière indirecte cet évènement, laissant ainsi l'exclusivité de l'ouverture du promoteur pour XPB (Bradsher et al., 2000; Coin et al., 1999; Moreland et al., 1999; Tirode et al., 1999; Winkler et al., 2000).

Une approche utilisant la méthode de pontage chimique ADN-protéine appliquée à l'étude de l'initiation de la transcription a montré que XPB interagit avec l'ADN en aval du site d'initiation +1 jusqu'à la position +38 (Douziech et al., 2000; Kim et al.,

2000). De plus, d'autres études montrent l'importance de la région +23 à +35 pour l'initiation de la transcription.

1.2- Rôle de TFIIH dans le départ du promoteur

Le passage de l'initiation à l'élongation est appelé « échappée du promoteur » (*promoter escape*). La phosphorylation du domaine carboxy-terminal de l'ARN polymérase II par la kinase de TFIIH est l'évènement clé de cette étape de la transcription (Yan and Gralla, 1999). Cette phosphorylation provoque un changement de conformation du complexe d'initiation aboutissant sur un re-largage de certains facteurs d'initiation. Il a été montré que TBP et TFIIE ont une interaction préférentielle avec la forme non-phosphorylée de l'ARN polymérase II et sont donc probablement décrochés du promoteur lors du *promoter escape* (Usheva et al., 1992). Les sous-unités p44 et XPB de TFIIH semblent être importants pour l'échappée du promoteur (Fukuda et al., 2002; Tremeau-Bravard et al., 2001).

Un rôle de TFIIH dans la synthèse d'ARN ribosomal *in vivo* et *in vitro*, a été établi (Iben et al., 2002). En effet, co-localisé au niveau du nucléole en complexe avec l'ARN polymérase I, TFIIH est plus particulièrement impliqué dans l'étape post-initiation de la transcription (Iben et al., 2002). Ceci a été confirmé par l'existence du complexe TFIIH/ARN polymérase I/CSB *in vivo* (Bradsher et al., 2002).

2- Rôle de TFIIH dans la transcription activée

La découverte, du rôle de la kinase de TFIIH pour la phosphorylation de certains récepteurs nucléaires a constitué la première preuve de la régulation de la fonction des récepteurs nucléaires par un facteur de transcription de base. En effet, il a été démontré que cdk7 phosphoryle les récepteurs de l'acide rétinoïque (RAR α et RAR γ) et le récepteur aux oestrogènes (ER α). Cette phosphorylation est indispensable pour la fonction de transactivation de ces récepteurs (Bastien et al., 2000; Chen et al., 2000; Rochette-Egly et al., 1997). Il a été montré récemment que certaines mutations dans l'hélicase XPD de certains patients atteints de *Xeroderma pigmentosum* ou *trichothiodystrophie*, sont à l'origine d'une diminution de l'activité de phosphorylation de TFIIH vis-à-vis de RAR α à la suite d'une déstabilisation de l'architecture du complexe TFIIH et notamment de l'ancrage du sous-complexe CAK au core TFIIH (Keriel et al., 2002).

3- Rôle de TFIIH dans la réparation par Excision-Resynthèse de Nucléotides

3.1- Généralités

L'ADN, support de l'information génétique nécessaire au développement et au bon fonctionnement de chaque organisme, est la cible d'agents génotoxiques externes (comme les radiations ionisantes, les UV et les agents chimiques) et internes comme les radicaux libres, provoquant des lésions altérant l'intégrité de l'information génétique. Ces lésions peuvent être à l'origine de nombreuses maladies génétiques, de carcinogenèse, de vieillissement et de mort cellulaire.

C'est justement pour préserver ce capital génétique, que la cellule a mis en place des systèmes de détection et de réparations de ces lésions (Figure 30).

Il existe 4 mécanismes de réparation :

- Le mécanisme de **réparation par excision de base** permettant la réparation des bases modifiées par oxydation, alkylation ou méthylation.
- Le mécanisme de **réparation des mésappariements** permettant la réparation des mésappariements A-G et T-C.
- Le mécanisme de **recombinaison homologue** permettant la réparation des cassures double brins et les pontages inter-brins.
- Le mécanisme **d'Excision-Resynthèse de Nucléotide (NER)** permettant essentiellement la réparation des lésions induisant une forte torsion de la double hélice d'ADN comme les adduits et les attaques des rayonnement UV ou de quelques hydrocarbures.

Nous traiterons essentiellement la réparation par excision resynthèse de nucléotides et nous verrons le rôle que joue le facteur TFIIH dans ce processus ainsi que dans un processus similaire de couplage entre la transcription et la réparation appelé TCR (pour *Transcription Coupled Repair*).

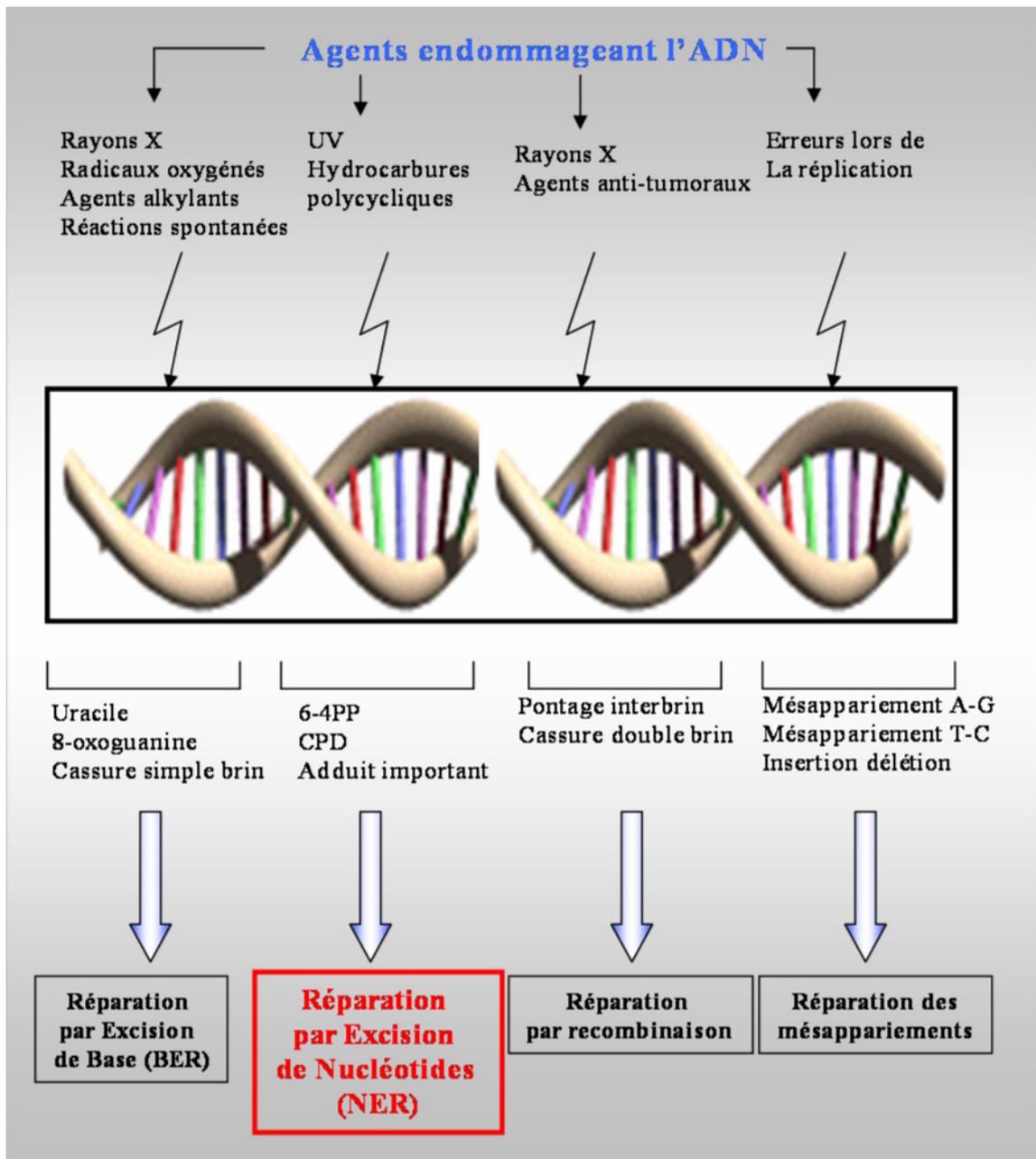


Figure 30 : Mécanismes de réparation de l'ADN (Hoeijmakers, 2001).

3.2- Le mécanisme du NER

Ce processus cellulaire requiert une vingtaine de protéines agissant dans un ordre précis dont les principales étapes sont :

- La reconnaissance de la lésion au niveau de l'ADN.
- L'ouverture de l'ADN autour de la lésion.
- L'incision de part et d'autre de la lésion.
- L'excision complète du brin lésé.
- La resynthèse et la ligation de l'ADN.

Les facteurs de réparation impliqués dans le NER sont : XPA, l'hétéro-trimère RPA, le complexe XPC/hHR23B, le complexe TFIIH, les deux endonucléases XPG et l'hétéro-dimère ERCC1-XPF. L'ADN polymérase resynthétise la partie manquante après élimination de la lésion avec l'aide de RPC et PCNA. L'ADN ligase permet la formation de la liaison phospho-diester de l'ADN néo-synthétisé. (Table 4 et Figure 31).

3.2.1- La reconnaissance de la lésion

La reconnaissance de la lésion est l'étape initiatrice du processus de réparation de l'ADN. Le facteur XPA a été la première protéine dont la capacité à reconnaître l'ADN endommagé a été décrite. La protéine XPC possède une affinité limitée pour l'ADN lésé. Cependant, associée à hHR23B pour former le complexe XPC/hHR23B, XPC est capable de reconnaître la lésion au niveau des distorsions de la double hélice d'ADN. De plus les changements de conformation dans la zone endommagée, facilitent le recrutement des autres facteurs, notamment le facteur TFIIH (Sugasawa et al., 1998; Wakasugi and Sancar, 1999). Un autre facteur UV-DDB (*UV-damaged DNA binding factor*), qui est un hétéro-dimère composé des sous-unités p48 et p127, jouerait un rôle important dans l'identification de certaines lésions de l'ADN dans un contexte chromatinien. Il a été notamment démontré que p48 peut être associé aux Histone Acétyl Transférases (HATs), suggérant un rôle dans l'augmentation de l'accessibilité des dommages au sein de la chromatine (Datta et al., 2001). Cependant,

UV-DDB semble ne pas être essentielle pour l'initiation du mécanisme de NER (Aboussekhra et al., 1995).

<i>Facteurs de NER</i>	<i>Fonction(s)</i>
XPC-hHR23B	Détection des lésions et recrutement des autres facteurs (spécifique GG-NER)
CSA, CSB	Déplacement de l'ARN polymérase II de la région lésée (spécifique TC-NER)
TFIIH	Catalyse l'ouverture de l'ADN autour de la lésion et facilite l'assemblage du complexe de réparation
XPA	Fixe les lésions de l'ADN et facilite l'assemblage du complexe de réparation
RPA	Stabilise l'ouverture de l'ADN et positionne les nucléases
XPG	Catalyse l'incision 3' et stabilise le complexe de réparation
ERCC1-XPF	Catalyse l'incision 5'

Table 4 : Les facteurs impliqués dans la réparation par excision re-synthèse de nucléotides (NER).

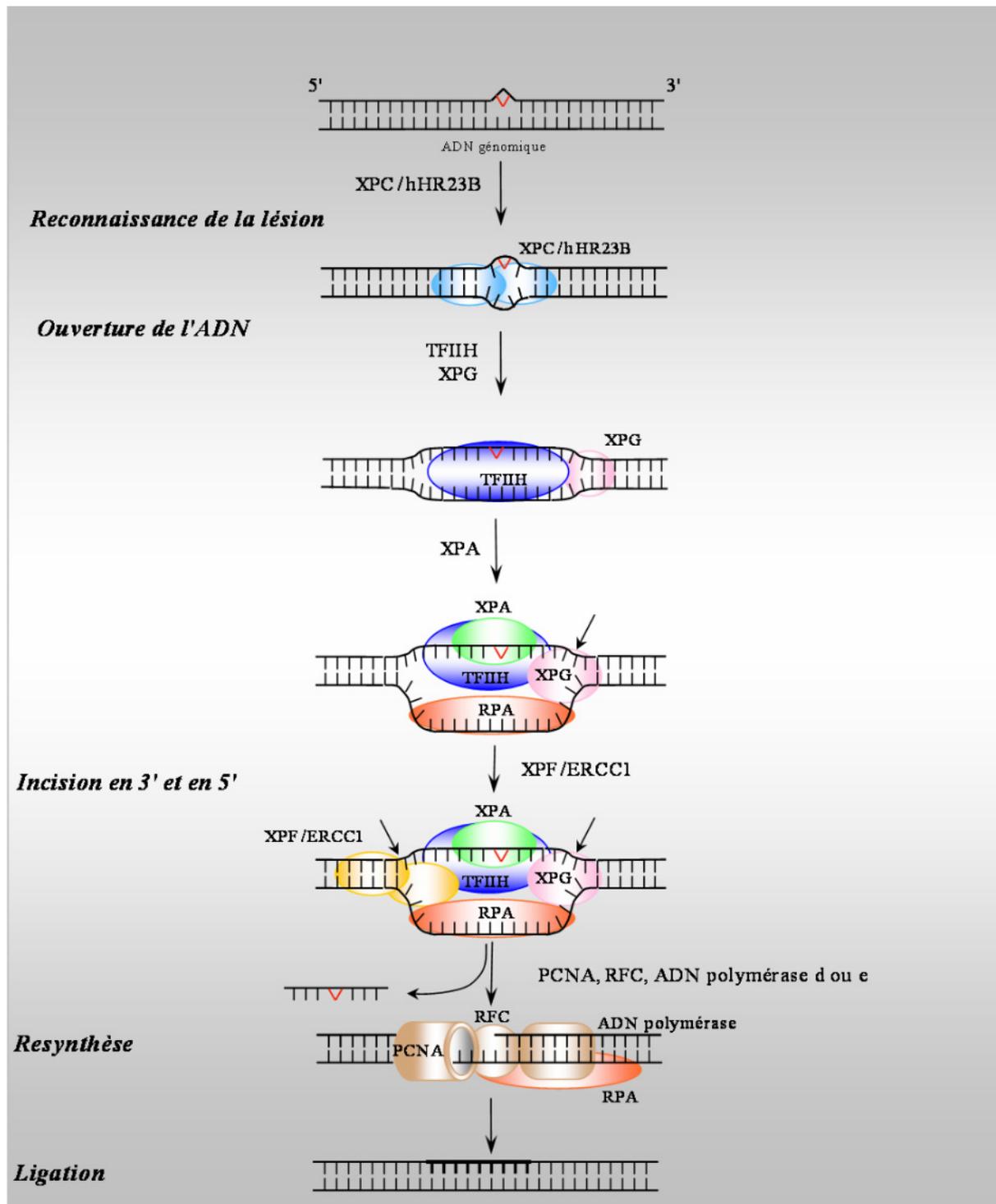


Figure 31 : Représentation de la réaction de réparation par excision-re-synthèse de nucléotides (NER) (de Laat et al., 1999).

3.2.2- L'ouverture autour de la lésion

Cette étape requiert les deux activités hélicases ATP-dépendante de polarité opposés de XPB et XPD du facteur TFIIH (Evans et al., 1997). La fixation de TFIIH, permettant l'ouverture de l'ADN, semble être stabilisée par le complexe RPA et le facteur XPA (de Laat et al., 1998; de Laat et al., 1999).

3.2.3- La double incision

L'ADN ouvert autour de la lésion est incisé en 3' et en 5' grâce aux deux endonucléases XPG et ERCC1/XPF, respectivement. L'incision est opérée spécifiquement au niveau des jonctions ADN simple brin/double brin (O'Donovan et al., 1994). Un ordre dans cette incision a été proposé :

Tout d'abord, XPG cliverait la lésion en 3' indépendamment de XPF. TFIIH, interagissant avec XPG, pourrait participer à son bon positionnement au sein du complexe de réparation (Habraken et al., 1996; Iyer et al., 1996). Enfin, XPF clive en 5' une fois que le clivage de 3' a été opéré par XPG. Le facteur RPA coordonnerait la spécificité du clivage de ERCC1/XPF (de Laat et al., 1998).

3.2.4- La resynthèse et la ligation d'ADN

Il s'agit de la dernière étape du processus de NER. Le fragment excisé est re-synthétisé par les ADN polymérase δ et ϵ . RPA semble être impliqué dans cette étape en restant lié à l'ADN simple brin. Le facteur RFC reconnaît l'ADN en son extrémité 3' et facilite l'ancrage de PCNA, permettant le positionnement des polymérase. RFC et PCNA augmentent la processivité des polymérase. Enfin, la ligase I, fixe le brin néo-synthétisé et l'incorpore au reste du duplex de l'ADN aboutissant à un ADN réparé.

3.3- Le rôle de TFIIH dans le mécanisme NER

Un rôle possible de TFIIH dans la reconnaissance de la lésion a été proposé à la suite de l'observation de l'augmentation de l'affinité de XPD et de RAD3 pour les lésions induites sur l'ADN par les UV (Sung et al., 1994).

Le facteur TFIIH, est le seul facteur de réparation par NER à posséder des activités hélicases dépendantes de l'ATP. TFIIH est impliqué directement dans l'étape de reconnaissance-incision de la réparation par NER, étant donné qu'elle est dépendante de l'hydrolyse de l'ATP (Evans et al., 1997). Les activités hélicases de sens opposé XPB et XPD sont responsables de l'ouverture de l'ADN de part et d'autre de la lésion. Des expériences de réparation *in vitro* sur un substrat d'ADN lésé pré-ouvert, en absence de TFIIH est à l'origine d'un problème dans la réaction d'excision (Mu et al., 1997). Ces résultats montrent que le rôle de TFIIH ne se limite pas seulement à ouvrir l'ADN autour de la lésion, mais semble être plus complexe que cela. En effet, TFIIH est capable d'interagir avec XPA (Park et al., 1995), XPC (Drapkin et al., 1994), XPG (Iyer et al., 1996) et XPF/ERCC1 (Evans et al., 1997). TFIIH semble stimuler les activités de XPG et du complexe XPF/ERCC1, suggérant un rôle probable de TFIIH dans le bon positionnement des endonucléases autour de la lésion. Par ailleurs, il a été montré que des mutations dans l'hélicase XPD provoquent une abolition totale de l'incision en 3' et en 5' *in vitro* (Winkler et al., 2000), qu'une délétion carboxy-terminale de p52 chez l'homme et la levure était à l'origine d'une sensibilité accrue aux rayonnements UV, accompagnée d'un défaut de réparation d'ADN (Feaver et al., 1999; Jawhari et al., 2002). Les souches de levure mutées dans les gènes codant pour *tfb1/p62* et pour *Ssl11/p44*, présentent une sensibilité accrue aux UV (Matsui et al., 1995; Wang et al., 1995; Yoon et al., 1992). Concernant le rôle du CAK dans le processus de NER, il a été montré que le complexe kinasique n'était pas nécessaire à ce mécanisme chez la levure (Sung et al., 1996; Svejstrup et al., 1995) et qu'il pouvait avoir un rôle inhibiteur pour le NER chez l'homme (Araujo et al., 2000).

3.4- TFIIH et la réparation couplée à la transcription

Le mécanisme de NER décrit ci-dessus est appelé réparation globale du génome GGR (pour *Global Genome repair*). Il concerne les lésions situées sur la totalité du génome. Toutefois, il existe un autre mécanisme de réparation pour les lésions situées sur les brins codant des gènes activement transcrits, il s'agit du mécanisme de NER couplé à la transcription TCR (pour *transcription coupled repair*) (Figure 32). Le TCR est un mécanisme relativement bien conservé de la bactérie à l'homme, puisqu'une protéine TRCF (pour *transcription-repair coupling factor*) chez *E. coli* reconnaît et déplace l'ARN polymérase, bloquée au niveau d'une lésion de l'ADN transcrit et recrute ainsi la machinerie de réparation (Selby and Sancar, 1993). Chez les eucaryotes, CSB (*Cockayne syndrome Group B*) est une protéine qui joue un rôle similaire à celui de TRCF. En effet, comme TRCF, CSB possède une activité ATPase et des motifs hélicases (Guzder et al., 1996; van Gool et al., 1994). De plus, CSB semble être associée à l'ARN polymérase lors de l'élongation, alors qu'elle ne possède pas d'activité hélicase et n'est pas capable de déplacer l'ARN Pol II (Selby and Sancar, 1997).

Le GGR et le TCR sont identiques à l'exception de la reconnaissance de l'ADN, qui est assurée par XPC/hHR23B et l'ARN polymérase II, respectivement. En effet, pour le GGR le complexe XPC/hHR23B reconnaît la lésion comme cela a été décrit précédemment, alors que pour le TCR, c'est l'ARN Pol II elle-même une fois bloquée au niveau de la lésion, qui sert de signal pour le recrutement de la machinerie de réparation de l'ADN.

Ce phénomène de couplage entre la transcription et la réparation suppose que TFIIH soit dynamique et soit capable de participer aux deux mécanismes. En fait, des études d'immuno-localisation de TFIIH *in vivo* ont montré d'abord que TFIIH est distribué de manière homogène dans le noyau et peut se déplacer librement pour interagir avec l'ARN polymérase I et II pendant 25 et 6 secondes, respectivement (Figure 33). Une irradiation UV est capable d'induire une relocalisation de TFIIH au niveau du site de lésion et de l'immobiliser pendant 4 minutes (Hoogstraten et al., 2002).

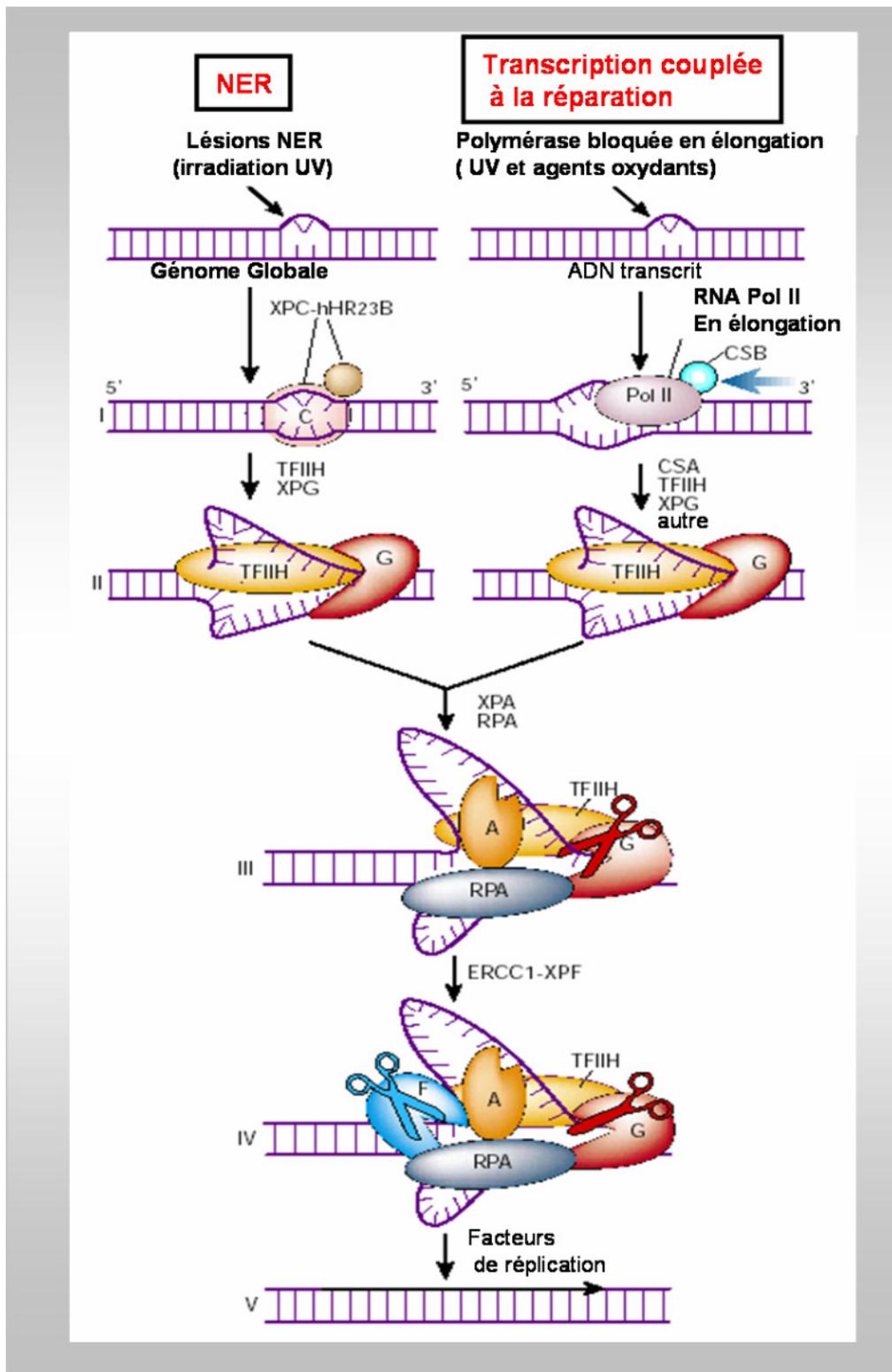


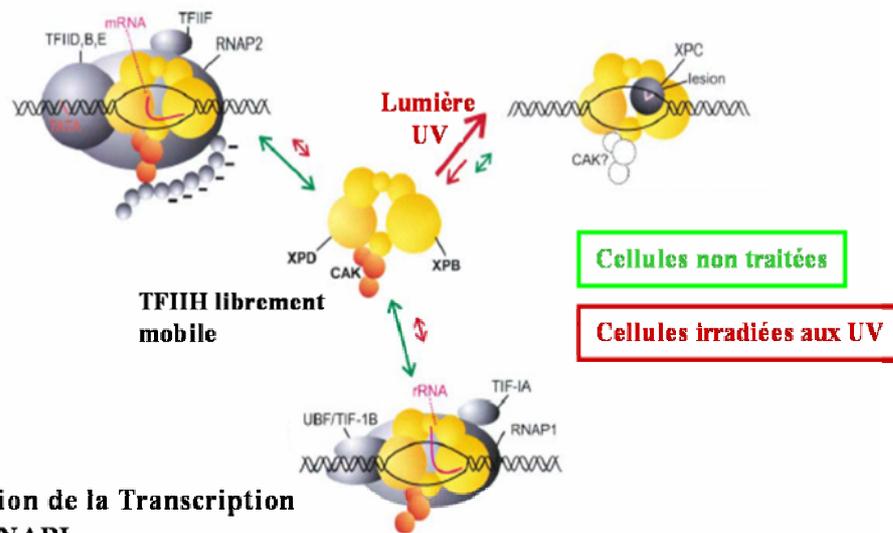
Figure 32: Modèle de réparation de l'ADN du génome globale par NER (Nucleotide Excision Repair) couplée au TCR (Transcription coupled repair). (Hoeijmakers, 2001).

**Initiation de la Transcription
par RNAPII.**

Fixation 2-10 secondes

**Réparation par excision
re-synthèse de nucléotide**

Fixation 3-4 minutes



**Initiation de la Transcription
par RNAPI.**

Fixation 25 secondes

Figure 33 : Modèle proposé pour le couplage de la transcription par l'ARN polymérase I et II ainsi que le NER, par le facteur TFIIH.

Les flèches vertes représentent l'équilibre de TFIIH en conditions normales. Après lésions induites aux UV, un déplacement rapide des complexes de TFIIH vers le processus de NER, est mis en évidence (flèches rouges) (Hoogstraten et al., 2002)

4- Rôle de TFIIH dans le cycle cellulaire

La présence de la kinase cdk7 au sein du facteur TFIIH laisse à penser que TFIIH joue un rôle dans la régulation du cycle cellulaire. En fait, il a été démontré que le complexe CAK (contenant cdk7, la cycline H et Mat1) pouvait se trouver *in vivo* sous forme libre ou associé aux sous-unités du core formant le complexe TFIIH (Rossignol et al., 1997). Cdk7 phosphoryle des éléments de la machinerie transcriptionnelle (comme le domaine carboxy-terminal de l'ARN polymérase II ou encore les récepteurs nucléaires), quand il est associé à TFIIH, alors que, sous forme libre, au sein du CAK, cdk7 phosphoryle les Cdks, contrôlant ainsi la progression du cycle cellulaire. Indirectement, TFIIH déplacerait l'équilibre vers un CAK associé diminuant ainsi la population de CAK libre et donc l'effet régulateur du CAK sur le cycle cellulaire. Chez la drosophile, il a été montré récemment que XPD régule négativement le rôle de cdk7 au cours du cycle cellulaire. En effet, un excès de XPD semble « titrer » l'activité du CAK en diminuant la quantité de CAK libre et par conséquent la phosphorylation des boucles T des Cdks (Chen et al., 2003).

Par ailleurs, une interaction a été décrite entre le suppresseur de tumeur p53 et TFIIH. En effet, le domaine carboxy-terminal de p53 est capable d'interagir avec p62, avec le domaine amino-terminal de XPD et avec le motif hélicase III de XPB (Leveillard et al., 1996; Wang et al., 1995; Xiao et al., 1994). L'interaction de p53 inhibe les activités hélicases XPB et XPD *in vitro*. p53 interagit aussi avec la cycline H et de ce fait inhibe l'activité kinase de cdk7 vis-à-vis de cdk2 ou d'un peptide mimant le domaine carboxy-terminal de l'ARN polymérase II (Schneider et al., 1998). De plus, cdk7 est capable de phosphoryler p53 augmentant ainsi sa capacité à lier l'ADN (Lu et al., 1997). Par ailleurs, il a été montré que p53 permet l'expression de l'inhibiteur des cdks p21^{Cip1} bloquant la progression du cycle cellulaire à la transition G1-S en réponse à des agents génotoxiques (Agarwal et al., 1998). Il est donc possible que dans un contexte de lésion de l'ADN, TFIIH phosphoryle p53, stimulant son activité de liaison à l'ADN et activerait ainsi l'expression de p21^{Cip1} aboutissant à un arrêt de la progression cellulaire et à un début de réparation. Ce modèle est très peu probable si on considère que p53 inhibe les activités hélicases de TFIIH et que ces dernières sont essentielles pour une réparation optimale.

La protéine p16^{Ink4a} a été caractérisée comme étant capable d'inhiber la transcription de base en diminuant la capacité de TFIIH à phosphoryler le CTD de l'ARN polymérase II, ce qui a pour effet d'empêcher la transition entre l'initiation et l'élongation, permettant le contrôle de la transcription au cours du cycle cellulaire.

5- Rôle de TFIIH dans l'apoptose

Les cellules mutées dans les gènes codant pour les hélicases XPB et XPD sont incapables d'induire l'apoptose dépendante de p53. Cependant, une transfection transitoire des gènes *xpb* et *xpd* permet aux cellules de malades de rentrer en apoptose (Robles et al., 1999; Wang et al., 1996), montrant une implication de TFIIH dans le processus de mort cellulaire programmée. Le mécanisme moléculaire mis en jeu dans l'implication de TFIIH dans l'apoptose n'a pas encore été élucidé à ce jour.

6-TFIIH et d'autres phénomènes

Un certain nombre de travaux récents implique TFIIH dans d'autres phénomènes rares et peu caractérisés. En effet, TFIIH a été isolé en association avec un ARN non-codant, participant à la reconnaissance du site d'épissage 5', U1 snRNA (pour *U1 small nuclear RNA*). Cette interaction se fait via la cyclin H et semble avoir une signification biologique, puisque U1 snRNA est connu pour stimuler la formation de la première liaison phospho-diester par l'ARN polymérase II (Kwek et al., 2002).

Par ailleurs, des relations entre les sous-unités du facteur de transcription TFIIH et le protéasome 26S ont été mises en évidence. En effet, l'hélicase p89/XPB interagit avec SUG1 (Weeda et al., 1997) et la sous-unité p44 présente une homologie de séquence avec S5a (Aravind and Ponting, 1998) ; les protéines SUG1 et S5a étant des sous-unités du protéasome 26S. De plus, p62 présente une homologie de séquence avec la protéine de levure dos2p impliquée dans la machinerie d'ubiquitination (Doerks et al., 2002). Ce qui constitue un argument supplémentaire pour un lien entre TFIIH et la machinerie de dégradation.

7- Maladies génétiques rares associées à TFIIH

Des mutations dans les gènes codant pour les deux hélicases XPB et XPD de TFIIH sont à l'origine de trois maladies génétiques autosomiques récessives rares qui sont des syndromes de transcription et/ou de réparation. Il s'agit de *Xeroderma pigmentosum* (XP), du syndrome de Cockayne (CS) et de la trichotiodystrophie (TTD). Les mutations répertoriées semblent être plus nombreuses pour XPD et ceci est probablement lié au caractère létal des mutations de XPB (Figure 34). Les seuls mutants répertoriés pour XPB sont positionnés dans le domaine N ou C-terminal, alors que ceux de XPD sont distribués sur toute la protéine (Figure 34).

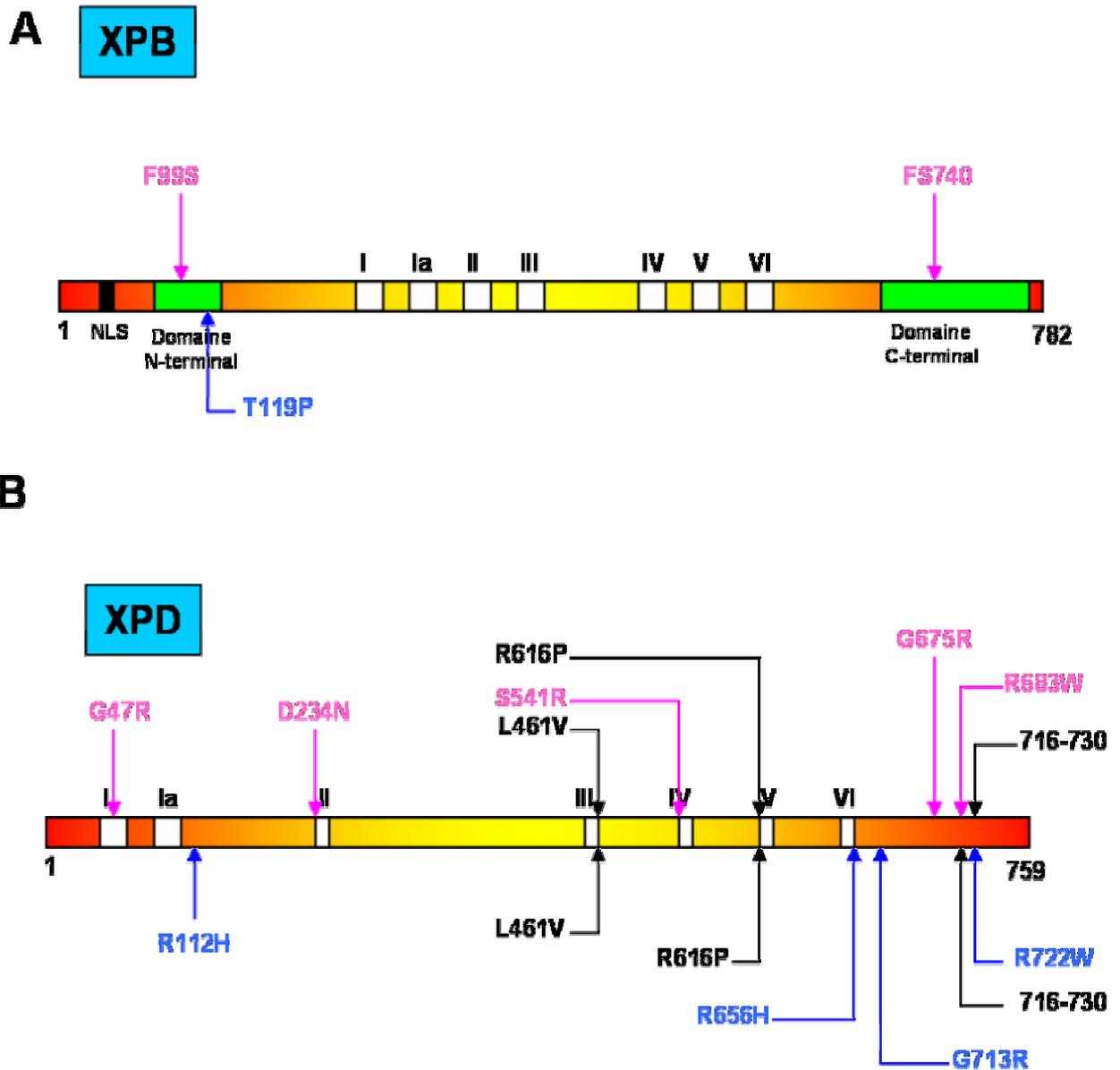


Figure 34 : Représentation de différentes mutations affectant les hélicases XPD et XPB.

A- Mutations répertoriées dans la protéine XPB.

B- Mutations répertoriées dans la protéine XPD.

Les mutations XP, TTD et XP/TTD sont indiquées en rose, bleu et noir, respectivement. Les carrés blancs représentent les domaines hélicases et les carrés verts correspondent aux domaines N et C-terminal de XPB (Gorbalenya and Koonin, 1991).

7.1- Le Xeroderma pigmentosum

L'origine du nom de cette maladie provient de ces symptômes associés (du grec *xeros*, *dermos* et *pigmentosum* pour assèchement, peau et pigmentation, respectivement). La fréquence d'apparition de cette maladie est d'environ un pour un million. L'espérance de vie d'un malade XP est de 30 ans. Les patients *Xeroderma pigmentosum* (XP) présentent également une sensibilité accrue aux rayonnements ultraviolet, une prédisposition au cancer de la peau (2000 fois supérieur à la normale), voire une dégénérescence neurologique (Bootsma et al., 1995; Chu and Mayne, 1996) (Tables 5 et 6, Figure 35). Des retards de croissance, une immaturité sexuelle ainsi que des anomalies dégénératives neurologiques associées à un retard mental et une microcéphalie peuvent être également associées à cette maladie (Chu and Mayne, 1996).

Outre les rayonnements UV, les carcinogènes chimiques, comme les dérivés du benzène ou les produits de pyrolyse du tryptophane, sont autant de facteurs à risque pour les patients XP. Cette maladie est considérée comme une maladie génétique résultant d'une déficience dans le système de réparation par Excision Resynthèse de Nucléotide (NER). Il existe 8 protéines capables de compléter les cellules de patients XP et donc en corriger la déficience. Il s'agit de facteurs de réparation d'ADN par NER correspondant aux groupes de complémentation de XP-A à XP-G. La fréquence relative et les syndromes associés diffèrent d'un groupe de complémentation à un autre (Table 6). En effet, les formes les plus sévères du XP sont associées à XP-A, alors que le groupe le moins affecté correspond à XP-E. De plus, dans les cellules XP-C, la réparation se fait normalement sur les gènes activement transcrits (Tyrrell and Amaudruz, 1987), alors que la réparation couplée à la transcription (TCR) et la réparation globale du génome (GGR) sont affectées chez tous les autres groupes de complémentation.



Figure 35 : Patients *Xeroderma pigmentosum*.

A- Assèchement de la peau.

B- Lésions et Carcinomes.

(Photos, Daniel Wallach)

Il existe un groupe de complémentation XP-V (*Xeroderma pigmentosum* variant) présentant un phénotype similaire aux autres groupes à l'exception près qu'il n'implique pas un défaut de réparation proprement dit mais plutôt un défaut de re-synthèse de l'ADN. En fait, dans les cellules XP-V, c'est la fonction de l'ADN polymérase η qui est affectée (Masutani et al., 1999). Cette polymérase permet la réplication de l'ADN au travers d'une lésion, et constitue ainsi un moyen de réparation post-réplivative de l'ADN (McDonald et al., 1999). De ce fait, chez les patients XP-V, les mutations s'accumulent dans les cellules.

Le très faible nombre de mutation *xpb* par rapport à *xpd* est probablement dû à leur caractère létal, certainement à cause de l'importance de XPB dans le processus de transcription. Ces mutations dans *xpb* et *xpd* n'aboutissent pas seulement à des défauts de réparation. En effet, ces mutations peuvent être à l'origine d'autres maladies, le syndrome de Cockayne et la trichothiodystrophie (Table 5).

Symptômes	Syndromes			
	XP	XP/ICS	CS	TTD
Cutanés				
Photosensibilité	++	++	+	+/-
Anomalies pigmentaires	++	+	-	-
Cheveux et ongles cassant	-	-	-	+
Ichtyose	-	-	-	+
Cancer de la peau	++	+	-	-
Développementaux				
Retard de croissance	-	+	+	+
Immaturité sexuelle	-	+	+	+
Neurologiques				
Neurodégénérescence	+	+	+	-
Neurodysmyélination	-	+	+	+
Dégénération mentale	+/-	+	+	+
Ataxie	+/-	+	+	+/-
Microcéphalie	+/-	+	+	+/-
Cellulaires				
UDS (% contrôle)	5-50	5-50	normale	5-50
Sensibilité aux UV	++	++	+	+/-

XP : xeroderma pigmentosum; CS : syndrome de Cockayne; TTD : trichothiodystrophie

Table 5 : Symptômes associés aux syndromes : *Xeroderma pigmentosum* (XP), *trichothiodystrophie* (TTD) et *syndrome de Cockayne* (CS).
(UDS : Unscheduled DNA synthesis, réparation d'ADN mesurée *in vivo*) (Hoeijmakers, 2001).

Groupe de complémentation	Fréquence relative	Réparation (NER)			Sensibilité aux UV	Clinique		Syndromes associés
		GGR	TCR	UDS		Cancers cutanés	Anomalie neurologique	
XP-A	élevée	-	+	<5%	+++	+	++ ^a	XP
XP-B	très rare	-	+	10-40%	++	+/-	++/+ ^b	XP/ICS, TTD
XP-C	élevée	-	+	15- 30%	+	+	-	XP
XP-D	intermédiaire	-	-	15- 30%	++	+/-	++/± ^b	XP, XP/ICS, TTD
XP-E	rare	-	?	>50%	±	+/-	-	XP
XP-F	rare	-	-	15-30%	+	+/-	++/± ^a	XP, XP/ICS
CS-A	intermédiaire	+	-	100%	+	-	++ ^b	CS
CS-B	élevée	+	-	100%	+	-	++ ^b	CS
TTD-A	très rare	-	-	10%	+	-	+ ^b	TTD

GGR : répartition globale du génome; TCR : réparation couplée à la transcription; UDS: Unscheduled DNA synthesis (synthèse d'ADN non programmée); ^a : neurodégénérescence, ^b: démyélinisation.

Table 6 : Groupes de complémentation associés aux maladies liées à un défaut du système NER (Hoeijmakers, 2001)..

7.2- Le syndrome de Cockayne

Le syndrome de Cockayne (CS) décrit par le généticien du même nom est caractérisé par un retard de croissance, un retard mental, des troubles neurologiques, oculaires (cataracte, rétinopathie pigmentaire) et auditifs (surdit  progressive) ainsi qu'une microc phalie associ e   une d my linisation (Friedberg, 1996). Les patients CS pr sentent  galement des anomalies du squelette : un faci s en « t te d'oiseau » et une ost oporose chez les patients les plus  g s (Tables 5 et 6).

L'esp rance de vie chez les patients CS est de 12 ans. La cause de la mort est souvent en relation avec une pneumonie ou des troubles respiratoires (Nance and Berry, 1992). Les patients CS sont photosensibles mais ne pr sentent ni d'anomalie de pigmentation ni de pr disposition aux cancers cutan s. Le syndrome de Cockayne est li    un d faut de la r paration coupl e   la transcription (TCR) et non pas   la r paration globale du g nome (Venema et al., 1990). Cinq prot ines sont capable de compl menter les cellules de malades CS : CSA, CSB, XPB, XPD et XPG (Broughton et al., 1995; Nospikel et al., 1997; Vermeulen et al., 1993; Vermeulen et al., 1994). Des mutations des g nes CSA et CSB sont   l'origine d'un syndrome de Cockayne typique dit « pur » alors que des mutations de XPB, XPD ou XPG aboutissent sur un syndrome qui combine les ph notypes de patients XP et CS (XP/CS).

Le fait que les sous-unit s XPB et XPD de TFIIH puissent  tre   l'origine du syndrome de Cockayne et que certaines prot ines comme CSA ou CSB puissent interagir avec TFIIH, renforce l'id e selon laquelle TFIIH serait impliqu  dans la r paration coupl e   la transcription (TCR) (Henning et al., 1995; Iyer et al., 1996).

7.3- La Trichothiodystrophie

La trichothiodystrophie (TTD) est une maladie génétique autosomique récessive dont le phénotype le plus marquant est les cheveux cassants formant des nœuds et ayant des extrémités en brosse (Table 5 et 6). Des études biochimiques ont mis en évidence un déficit en protéines soufrées de faible poids moléculaire (Gillespie and Marshall, 1983). Les patients TTD développent également une microcéphalie, un retard mental, des troubles neurologiques associés à une hypomyélinisation, ichtyose, hypogonadisme et nanisme cachectique. La moitié des patients présente une photosensibilité aux UV, mais pas de prédisposition aux cancers de la peau. Les cellules TTD réparent de 4 à 10 fois moins bien l'ADN que les cellules saines. Les hélicases XPD et XPB de TFIIH constituent deux groupes de complémentation TTD (Stefanini et al., 1986; Vermeulen et al., 1994). Il existe un troisième groupe de complémentation TTD-A dont la protéine correspondante n'a pas encore été isolée. En fait, la micro-injection d'ADNc des sous-unités isolées de TFIIH ne rétablit pas une activité de réparation normale. Cependant, l'injection du facteur TFIIH purifié est capable de corriger la déficience en réparation (Vermeulen et al., 1994), montrant que TTD-A a un rapport avec TFIIH. En effet, il a été montré qu'une baisse de la concentration en TFIIH dans les cellules TTD-A est à l'origine de la maladie (Vermeulen et al., 2000). De plus, l'injection des sous-unités isolées de TFIIH n'a pas permis de corriger la déficience (Vermeulen et al., 2000). Actuellement, on pense que le facteur TTD-A est une protéine qui stabiliserait TFIIH et qui évite sa dégradation (Bergmann and Egly, 2001).

Chapitre V - Résultats

Expression de protéines fusionnées au peptide FLAG dans les cellules d'insectes. Application pour le facteur multi-protéique de transcription et réparation TFIIH

Le facteur de transcription-réparation TFIIH a été purifié pour la première fois à partir d'extrait de cellules HeLa. Il s'agissait d'une purification en 6 étapes, utilisant les chromatographies : héparine, DEAE (Di-Ethyl Amino Ethyl), Sulfopropyl, Phényl, hydroxyapatite et enfin DEAE, avant l'obtention d'un complexe TFIIH pur et actif en transcription (Gerard et al., 1991). La lourdeur de la purification de ce facteur a amené certains auteurs à générer une lignée cellulaire stable exprimant la sous unité XPB double fusionnée en ses extrémités amino et carboxy-terminales à un peptide hexa-histidine et hémagglutinine, respectivement (Winkler et al., 2000). Ce système d'immuno-purification a permis de simplifier la préparation de TFIIH en vue d'une étude fonctionnelle. La volonté de caractériser les différentes sous-unités de TFIIH, et comprendre le rôle précis de chacune dans les fonctions de TFIIH, a été un argument supplémentaire en faveur de la mise en place d'un système de reconstitution de TFIIH, en utilisant notamment le système d'expression de baculovirus. La production de baculovirus recombinant exprimant des mutants de XPB, XPD et cdk7, a permis d'assigner un rôle à XPB dans l'ouverture de l'ADN autour du promoteur, un rôle stimulateur de XPD dans la transcription et enfin un rôle de protéine kinase capable de phosphoryler le CTD de l'ARN polymérase II, pour cdk7 (Tirode et al., 1999). Ce même système a été très utile pour la reproduction de mutations de malades *xeroderma pigmentosum*, *trichothiodystrophie* ou *syndrome de Cockayne* et pour l'étude de l'effet de ces mutations sur les fonctions de TFIIH (Coin et al., 1999).

De tels systèmes n'étant pas très adaptés pour l'étude des interactions des sous-unités du core-TFIIH et pour la recherche de sous-complexes stables, nous avons généré des baculovirus recombinants exprimant toutes les sous unités du core-TFIIH

fusionnées au peptide Flag en N-terminal, permettant ainsi de s'affranchir des anticorps monoclonaux dont le pouvoir immuno-précipitant est variable d'un anticorps à un autre. De plus, l'étude en microscopie électronique de TFIIH recombinant entier et/ou des différents sous-complexes, nécessite un échantillon homogène. C'est la raison pour laquelle, nous avons mis en place une stratégie d'immuno-purification en deux étapes (*his-tag* et Flag).

Résumé de l'article

Le facteur de transcription/réparation TFIIH est utilisé comme système modèle d'application de l'expression de protéines fusionnées au FLAG dans les cellules d'insectes, pour des études structurales et fonctionnelles. Dans ce présent travail, nous avons construit des baculovirus recombinants exprimant les quatre sous-unités du core-TFIIH (p34, p44, p52 et p62), fusionnées en N-terminal au peptide Flag. En utilisant ces virus recombinants, nous avons établi un protocole basé sur une chromatographie d'affinité par le Flag, permettant une analyse systématique des interactions pour l'étude de complexes multi-protéiques. Une stratégie de purification en deux étapes (FLAG et *his-tag*) a été développée pour l'identification et la purification de sous-complexes stables de TFIIH. Cette procédure simple de purification a permis l'obtention d'un TFIIH recombinant contenant l'ensemble des sous-unités. Il est actif en transcription et est structuralement analogue au TFIIH endogène observé par microscopie électronique et analyse d'images

[Signalement bibliographique ajouté par : ULP – SCD – Service des thèses électroniques]

Expression of FLAG Fusion Proteins in Insect Cells: Application to the Multi-subunit Transcription/DNA Repair Factor TFIID

Anass Jawhari, Muriel Uhring, Corinne Crucifix, Sébastien Fribourg, Patrick Schultz, Arnaud Poterszman, Jean Marc Egly and Dino Moras

Protein Expression and Purification, 2002, Vol. 24, N°3, Pages 513-523

Pages 513-523 :

- La publication présentée ici dans la thèse est soumise à des droits détenus par un éditeur commercial.
- Pour les utilisateurs ULP, il est possible de consulter cette publication sur le site de l'éditeur :

<http://dx.doi.org/10.1006/pep.2001.1597>

- Il est également possible de consulter la thèse sous sa forme papier ou d'en faire une demande via le service de prêt entre bibliothèques (PEB), auprès du Service Commun de Documentation de l'ULP: peb.sciences@scd-ulp.u-strasbg.fr

p52 régule la fonction de XPB au sein de TFIIH

Depuis 10 ans, l'équipe de Jean Marc Egly (IGBMC), a caractérisé systématiquement la fonction de plusieurs sous-unités de TFIIH. La fonction des sous-unités p52, p34 et p62 reste assez mal connue. Dans l'objectif de mieux comprendre la contribution de p52 dans les fonctions de TFIIH, nous avons initié ce travail.

p52 a été la dernière sous-unité découverte de TFIIH. Il s'agit d'un polypeptide de 436 acides aminés et d'un poids moléculaire de 52 kDa. p52 avec p34, p44, p62 et XPB forment le core-TFIIH. Des expériences de transcription et de réparation *in vitro* et de micro-injections *in vivo* ont montré l'implication directe de cette sous unité, nouvellement identifiée, dans la transcription et la réparation d'ADN par excision re-synthèse de nucléotide (NER) (Marinoni et al., 1997).

Chez la levure, Tfb2 l'homologue de p52 humain (63.9 % de similarité et 39.7 % d'identité de séquence) est essentiel pour la viabilité de la cellule, alors qu'une délétion des ces 60 derniers aminoacides n'est pas létale. Une telle mutation est à l'origine d'un dysfonctionnement dans le système de NER et une sensibilité importante aux rayonnements ultraviolets (Feaver et al., 1997).

Dans le but d'explorer les effets d'une telle mutation sur l'activité de transcription de TFIIH humain (qui n'avait pas pu être testé par Feaver & collaborateurs) et pour mieux comprendre les bases moléculaires du dysfonctionnement du NER, nous avons initié ce travail.

Nous avons d'abord étudié l'effet d'une délétion carboxy-terminale de la sous unité p52 sur l'architecture de TFIIH, en utilisant le système reconstitué de baculovirus. Nous avons pu montrer que cette mutation affecte de manière spécifique l'ancrage de XPB au sein de TFIIH. Au même moment, Jean Philippe Lainé de l'équipe de Jean Marc Egly avait montré qu'une telle mutation causait une importante réduction de l'activité de transcription dans des tests *in vitro*. Ensemble, nous avons étudié la relation architecture-fonction d'une telle délétion. En effet, nous avons pu établir que cette mutation induit une instabilité de XPB, qui est à l'origine de la diminution de l'activité de transcription et de réparation de TFIIH, montrant d'une part que le

système de levure et celui de l'homme sont très proche et d'autre part que p52 joue un rôle dans l'ancrage de XPB au complexe TFIIH.

Résumé de l'article

Pour une meilleure compréhension du facteur de transcription/réparation TFIIH, nous avons cherché à préciser le rôle de la sous-unité p52 dans la fonction de TFIIH. En utilisant le système reconstitué de transcription ou de réparation d'ADN par excision re-synthèse de nucléotides *in vitro*, nous avons montré qu'une délétion du domaine carboxy-terminal de p52 cause une importante réduction des activités de transcription et de réparation de TFIIH. Cette mutation, empêche l'ouverture correcte du promoteur, et ne semble pas affecter les autres activités enzymatiques de TFIIH. De plus, nous avons montré que p52 entier conditionne l'ancrage de XPB au sein de TFIIH, expliquant l'origine des problèmes de transcription et de réparation du mutant de p52. Nous avons ensuite montré, que ces deux sous unités interagissent physiquement, et avons localisé les domaines impliqués dans cette interaction. L'ensemble de nos résultats, montre que la sous-unité p52 régule la fonction de XPB à travers une interaction directe, de la même manière que p44 régule l'activité hélicase de XPD.

[Signalement bibliographique ajouté par : ULP – SCD – Service des thèses électroniques]

p52 Mediates XPB Function within the Transcription/Repair Factor TFIID

Anass Jawhari, Jean-Philippe Lainé, Sandy Dubaele, Valérie Lamour, Arnaud Poterszman, Frédéric Coin, Dino Moras, and Jean-Marc Egly

The Journal of Biological Chemistry, 2002, Vol. 277, N° 35, Pages 31761-31767

Pages 31761-31767 :

- La publication présentée ici dans la thèse est soumise à des droits détenus par un éditeur commercial.
- Pour les utilisateurs ULP, il est possible de consulter cette publication sur le site de l'éditeur :

<http://www.jbc.org/cgi/content/abstract/M203792200v1>

- Il est également possible de consulter la thèse sous sa forme papier ou d'en faire une demande via le service de prêt entre bibliothèques (PEB), auprès du Service Commun de Documentation de l'ULP: peb.sciences@scd-ulp.u-strasbg.fr

Structure d'un domaine de p62 révèle un repliement du type PTB/PH, indispensable pour la réparation de l'ADN.

p62 est la première sous-unité clonée de TFIIF (Fischer et al., 1992). Son homologue de levure Tfb1 est indispensable pour la transcription des gènes codant pour les protéines et pour la réparation de l'ADN par excision re-synthèse de nucléotides (Wang et al., 1995). Sa contribution exacte à ces deux fonctions de TFIIF reste jusqu'à présent pas bien comprise. Cependant, un certain nombre de travaux semblent désigner p62 comme la cible de plusieurs protéines partenaires de TFIIF. En effet, dans un contexte de réparation d'ADN, p62 participe à la reconnaissance de XPC/HR23B (Yokoi et al., 2000) et dans un contexte de transcription, l'homologue de levure de p62 (Tfb1), interagit avec la sous-unité α du facteur de transcription TFIIE (Uetz et al., 2000). p62 est aussi la cible d'un ensemble de régulateurs de transcription, tels que VP16, p53, E2F-1, HBx, et EBNA 2 (Chen et al., 2003; Leveillard et al., 1996; Tong et al., 1995; Truant et al., 1995; Vandell and Kouzarides, 1999; Xiao et al., 1994). Le récepteur humain à l'œstrogène ER α interagit avec TFIIF, via p62 (Chen et al., 2000). p62, étant une protéine à l'interface d'un réseau d'interaction, et dans le but de mieux comprendre sa fonction, nous avons voulu résoudre sa structure. Nous avons essayé de l'exprimer dans *E.coli*, sans succès. Nous avons alors envisagé d'isoler ses domaines. Une étude de conservation de séquences, a suggéré une organisation en trois domaines : amino-terminal, médian et carboxy-terminal. Par une étude de protéolyse, nous avons pu isoler un domaine amino-terminal stable de 12 kDa (voir article 4). Nous avons établi un protocole d'expression et purification, aboutissant sur des quantités importantes de protéine pure et de bonne qualité. Nous avons résolu sa structure par résonance magnétique nucléaire. Cette structure a révélé un repliement du type PTB/PH (*Phosphotyrosine Binding /Pleckstrin Homology*). Une telle homologie structurale ne pouvait pas être prédite par homologie de séquences. Nous avons pu reconstituer un complexe TFIIF ne contenant pas ce domaine sans affecter l'incorporation des autres sous-unités au complexe. Cette délétion n'affecte pas la transcription. Cependant, elle cause une diminution très importante de

l'activité NER de TFIIH. Une interaction directe avec l'endonucléase XPG a été établie et pourrait être à l'origine du dysfonctionnement de la réparation de l'ADN. Ainsi, nous avons pu montrer que le domaine amino-terminal de p62 est un PH/PTB, qui régule l'activité NER, probablement en interagissant avec XPG.

Résumé de l'article

TFIIH est un complexe multi-protéique impliqué dans deux processus cellulaires importants : la réparation de l'ADN par excision re-synthèse de nucléotide et la transcription. Nous avons mené une étude structurale de p62 humain, la sous unité du core-TFIIH, qui participe à la reconnaissance de multiples partenaires de TFIIH, dont le facteur de transcription TFIIE et un certain nombre de régulateur de la transcription. La structure de ce domaine a été résolue par résonance magnétique nucléaire et a révélée la présence de sept feuillets- β « *sandwich* » suivis d'une hélice carboxy-terminale. Ce repliement est caractéristique des domaines PH (*Pleckstrin Homology*) et PTB (*PhosphoTyrosine Binding*). Nous montrons qu'une délétion du domaine amino-terminal de p62(1-108) n'affecte pas l'intégrité de TFIIH et n'a aucun effet sur l'activité de transcription de TFIIH. Cependant, elle provoque une diminution très importante de l'activité de réparation par NER de TFIIH. Une interaction directe avec l'endonucléase XPG a été établie et pourrait être à l'origine du dysfonctionnement de la réparation de l'ADN.

[Signalement bibliographique ajouté par : ULP – SCD – Service des thèses électroniques]

TFIIH contains a PH domain involved in DNA nucleotide excision repair

Virginie Gervais*, Valérie Lamour*, **Anass Jawhari***, Florent Frindel, Emeric Wasielewski, Sandy Dubaele, Jean-Marc Egly, Jean-Claude Thierry, Bruno Kieffer & Arnaud Poterszman

Nature Structural & Molecular Biology, 2004, Vol. 11, Pages 616 - 622

*These authors contributed equally to this work.

Pages 616 – 622 :

- La publication présentée ici dans la thèse est soumise à des droits détenus par un éditeur commercial.
- Pour les utilisateurs ULP, il est possible de consulter cette publication sur le site de l'éditeur :

<http://www.nature.com/nsmb/journal/v11/n7/full/nsmb782.html>

- Il est également possible de consulter la thèse sous sa forme papier ou d'en faire une demande via le service de prêt entre bibliothèques (PEB), auprès du Service Commun de Documentation de l'ULP: peb.sciences@scd-ulp.u-strasbg.fr

Utilisation de protéolyse ménagée combinée à la spectrométrie de masse pour l'identification de domaines stables de p62.

Les protéines sont généralement composées de régions structurées et de région plus flexibles. Très souvent, les régions les plus structurées correspondent à ce que l'on appelle communément des domaines, correspondant à une fonction particulière. Les régions flexibles peuvent adopter plusieurs conformations, sont sensibles aux protéases et correspondent très souvent à des jonctions inter-domaines. Les analyses de conservation de séquences ainsi que les prédictions de structures secondaires, permettent de proposer une définition de domaines au sein d'une protéine multi-domaines. Ces prédictions doivent être confirmés par des données expérimentales et notamment des expériences de protéolyse ménagée, et des caractérisations biophysiques. La définition de domaine est une tâche relativement difficile, mais néanmoins nécessaire pour l'obtention d'un échantillon de bonne qualité (homogénéité, solubilité, monodispersité et stabilité), en vue d'une étude structurale. C'est dans l'objectif de définir des domaines structuraux de p62, que ce travail a été initié.

p62 est une sous unité du core-TFIID. C'est la première sous-unité découverte (Fischer et al., 1992). Son rôle biologique reste assez mal connu, même si un certain nombre de données impliquent p62 dans un réseau d'interaction avec plusieurs partenaires de TFIID. C'est le cas du facteur de transcription TFIIE, du facteur de réparation XPG, des régulateurs de transcription VP16, p53, E2F-1, HBx, EBNA2 et FBP, et du récepteur à l'œstrogène ER α . (Chen et al., 2000; Iyer et al., 1996; Vandel and Kouzarides, 1999; Xiao et al., 1994). Nous avons utilisé une approche systématique combinant la protéolyse et la spectrométrie de masse qui a permis d'identifier deux domaines stables de p62 : le premier a été identifié suite à une protéolyse spontanée et correspond à p62(1-109) et le deuxième a été identifié par protéolyse ménagée et correspond à une région très conservée contenant le motif FWxx $\Phi\Phi$ (Φ : Tyr ou Phe).

Résumé de l'article

TFIIH est un facteur multi-protéique qui joue un rôle central dans la transcription et la réparation de l'ADN par excision- resynthèse de nucléotides. p62 est une sous-unité du core-TFIIH qui interagit avec VP16, p53, Erα et E2F1 dans un contexte de transcription et avec XPG dans un contexte de réparation. Des expériences de protéolyse ménagée couplée à la spectrométrie de masse ont permis de définir des domaines structuraux de p62. Trois modules ont été identifiés, le premier correspond à p62(1-108) et a été isolé par protéolyse spontanée, les deux autres correspondent à p62(109-171) et p62(186-240) et contiennent un motif conservé de dipeptide FW (Phénylalanine, tryptophane). p62(186-240) correspond à la région la plus conservée contenant la motif FWxxΦΦ (Φ : tyrosine et phénylalanine), trouvé dans d'autres familles de protéines. La caractérisation biophysique basée sur des expériences de fluorescence indique que ce domaine est replié correctement et correspond à une entité structurale.

[Signalement bibliographique ajouté par : ULP – SCD – Service des thèses électroniques]

Domain Architecture of the p62 Subunit from the Human Transcription/Repair Factor TFIIF Deduced by Limited Proteolysis and Mass Spectrometry Analysis

Anass Jawhari, Stéphanie Boussert, Valérie Lamour, R. Andrew Atkinson, Bruno Kieffer, Olivier Poch, Noelle Potier, Alain van Dorsselaer, Dino Moras, and Arnaud Poterszman

Biochemistry, 2004, Vol. 43, Pages 14420-14430

Pages 14420-14430 :

- La publication présentée ici dans la thèse est soumise à des droits détenus par un éditeur commercial.
- Pour les utilisateurs ULP, il est possible de consulter cette publication sur le site de l'éditeur :

<http://pubs.acs.org/cgi-bin/article.cgi/bichaw/2004/43/i45/html/bi048884c.html>

- Il est également possible de consulter la thèse sous sa forme papier ou d'en faire une demande via le service de prêt entre bibliothèques (PEB), auprès du Service Commun de Documentation de l'ULP: peb.sciences@scd-ulp.u-strasbg.fr

Etude structurale du facteur de transcription TFIIE

L'initiation de la transcription des gènes codant pour les protéines est un processus complexe, qui nécessite l'action de l'ARN polymérase II et six facteurs de transcriptions de base : TFIIA, TFIIB, TFIID, TFIIE, TFIIF et TFIIH (Orphanides et al., 1996 ; Roeder et al., 1996). Ces facteurs généraux s'assemblent de manière séquentielle au niveau du promoteur pour permettre l'initiation de la transcription (Hampsey, 1998). Le facteur TFIIE rejoint le complexe de pré-initiation de la transcription (PIC) après l'ARN polymérase II et avant le facteur TFIIH (Buratowski et al., 1989; Flores et al., 1992). TFIIE interagit avec la forme non phosphorylée de l'ARN polymérase II, avec TFIIF et avec TFIIH (Flores et al., 1989; Maxon et al., 1994). Au sein du PIC, TFIIE recrute TFIIH et stimule sa capacité à phosphoryler le domaine carboxy-terminal de l'ARN polymérase II (Lu et al., 1992).

Pour mieux comprendre l'assemblage de TFIIE avec TFIIH, qui constitue la dernière étape de l'initiation de la transcription, nous avons envisagé de résoudre la structure du complexe TFIIE/TFIIH par microscopie électronique. Pour ce faire, nous avons décidé d'abord de résoudre la structure de TFIIE seul et de caractériser l'interaction TFIIE/TFIIH avant de s'intéresser au complexe.

Nous avons co-exprimé et purifié TFIIE humain. Nous avons ensuite essayé d'étudier son comportement en solution. Par des expériences de pontage chimique et d'ultra-centrifugation analytique, nous avons pu montrer que TFIIE se comporte en solution comme un dimère $\alpha\beta$ et non pas comme un tétramère $\alpha_2\beta_2$ tel qu'il est admis dans la littérature (Inostroza et al., 1991; Ohkuma et al., 1990). La structure de TFIIE a été résolue en microscopie électronique. Elle a révélée une organisation en domaines distincts.

Résumé de l'article

TFIIE est un facteur de transcription de base, connu pour interagir avec TFIIH et stimuler sa capacité à phosphoryler le CTD de l'ARN polymérase II. Dans le but d'étudier la structure de ce facteur, nous l'avons surproduit dans *E.coli*, purifié et étudié son comportement en solution. Ce complexe existe en une seule espèce en gel natif. Des expériences de pontage chimique ainsi que d'ultracentrifugation analytique montrent que TFIIE existe en solution sous forme de dimère $\alpha\beta$, contrairement à ce qui est admis jusque là. De plus la structure de TFIIE a été résolue par microscopie électronique et analyse d'images à 17Å. Elle a révélée un TFIIE organisé en domaines distincts.

**Oligomeric state and molecular structure of human
transcription factor TFIIIE**

To be submitted to *EMBO.J*

Jawhari, A^{1*}. Uhring, M^{1*}. De carlo, S^{1,2}. Moras, D¹. Schultz, P^{¶1,1,2} & Poterszman, A¹.

Keywords: electron microscopy, TFIIIE, image analysis, immunolabelling, 3-D model, oligomeric state, co-immunopurification, crosslinking and ultra-centrifugation.

¹ Department of Structural Biology and Genomics, Institut de Génétique et de Biologie Moléculaire et Cellulaire, CNRS/INSERM/ULP, Illkirch, France

² Ecole Supérieure de Biotechnologie de Strasbourg, Pôle API, Illkirch, France

* These authors contribute equally to this work

¶To whom correspondence should be addressed :

Patrick Schultz Department of Structural Biology and Genomics, Institut de Génétique et de Biologie Moléculaire et Cellulaire, CNRS/INSERM/ULP, 1, rue Laurent Fries, BP10142, F-67404 Illkirch, France. Tel.: +33 3 90 24 4800; Fax: +33 3 88 65 3201; E-mail: patrick.schultz@igbmc.u-strasbg.fr

Chapitre VI – Etude cristallographique du complexe p34/p44 humain

I- Expression et Purification du complexe p34/p44

1-Définition du complexe stable

L'utilisation du système reconstitué de TFIID dans les cellules d'insecte (Tirode et al., 1999) a permis une étude d'interaction des sous-unités de TFIID (Jawhari et al., 2002; Tirode et al., 1999). Ce système a servi à mettre en évidence une interaction directe entre p34 et p44 (Fribourg et al., 2001). Cette interaction avait déjà été décrite chez la levure (Feaver et al., 2000; Iyer et al., 1996). Des tests de co-expression systématique dans *E.coli* de différents domaines de p34 avec différents domaines de p44 ont permis de localiser les zones d'interaction (Fribourg et al., 2001). En effet, p34 (1-233) (p34N) et p44 (331-395) (p44C) sont à l'interface de l'interaction des deux sous-unités. De plus, seule la co-expression de p34N avec le fragment C-terminal de p44 permet d'obtenir un complexe soluble et stable, alors que ce même p34N exprimé seul est complètement insoluble.

2-Expression

Les ADNc de p34(1-233) et p44(331-395) ont été clonés dans les vecteurs pET15b (Ampicilline) et pACYC2 (Chloramphénicol), respectivement (Figure 36). p34N est exprimé en fusion avec un peptide *his-tag* en N-terminal et p44C est exprimé sans aucune fusion. Une double transformation est réalisée dans la souche bactérienne BL21 (DE3). 6 litres de LB en fioles ou en incubateur sontensemencés par une préculture poussée à partir de bactéries fraîchement transformées, en présence d'ampicilline et de chloramphénicol, à 37°C. L'expression est induite par l'addition d'IPTG à 0.8mM final dans les milieux de culture dont la température a été abaissée à 18°C et stabilisée pendant 1 heure. L'induction se poursuit à 18°C pendant la nuit.

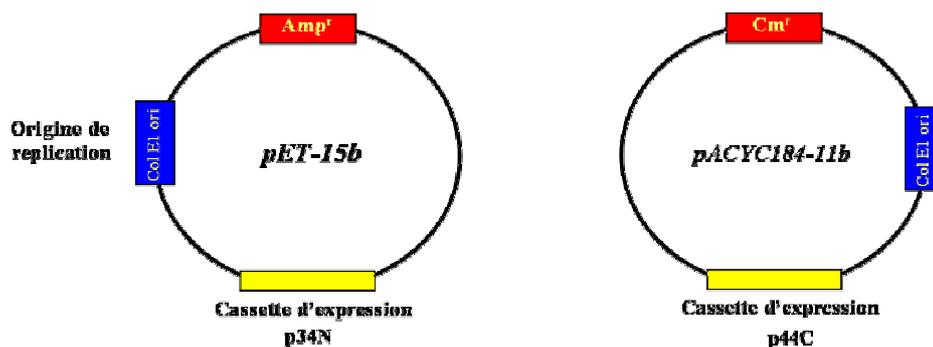


Figure 36 : Les vecteurs utilisés pour la co-expression.

Les vecteurs pET-15b et pACYC184-11b possèdent deux origines de réplication différentes et deux résistances aux antibiotiques différentes, l'ampicilline et le chloramphénicol, respectivement.

3-Purification

La purification du complexe p34N/p44C commence par une lyse cellulaire dans le tampon A (20mM Tris HCl, 250 mM NaCl, 1 mM β -mercaptoéthanol), contenant 1 mg de lysozyme. L'échantillon est soniqué durant 10 minutes de temps effectif par des impulsions d'ultra-sons toutes les 2.5 secondes à 25% d'amplitude à 4°C, en limitant la température à 12°C (Sonificateur Bioblock). L'échantillon est ensuite centrifugé pendant 1h à 40000 G et à 4°C. La purification de ce complexe s'effectue en trois étapes : d'abord une chromatographie d'affinité au peptide hexa-histidine, ensuite une chromatographie hydrophobe (TSK-Phényl) et enfin une chromatographie de filtration sur gel. Ceci est indiqué dans le schéma de purification (Figure 37).

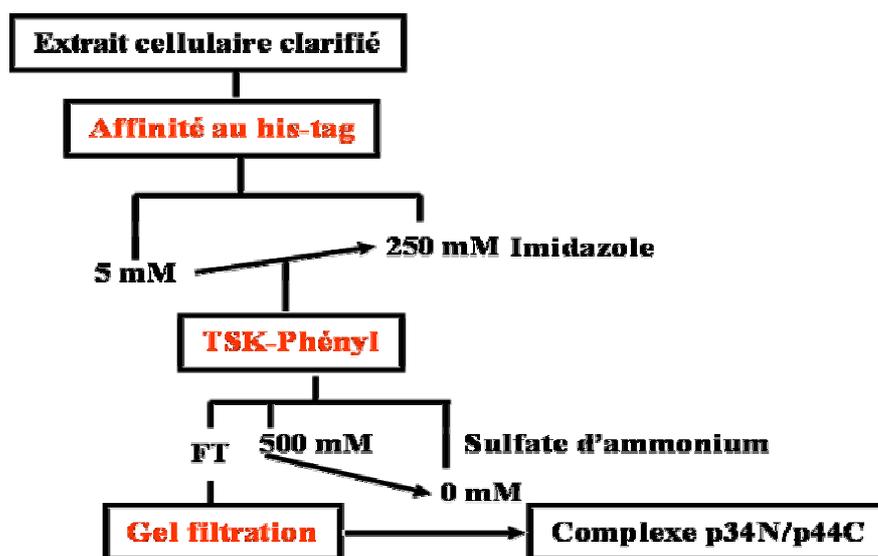


Figure 37 : Schéma de purification du complexe p34N/p44C.

3.1-Chromatographie d'affinité

La protéine p34N est fusionnée au niveau de sa partie amino-terminale à une séquence poly-histidine. La première étape de purification du complexe est réalisée sur une colonne d'affinité au cobalt (Talon, Clontech). Après lyse cellulaire et centrifugation de l'extrait brute, le surnageant est incubé pendant 45 minutes sous agitation, à 4°C avec un équivalent de 5 ml de résine pour 6 litres de culture. Après centrifugation 1000 G pendant 5 minutes à 4°C, le surnageant équivalent à l'effluent est prélevé et la résine est coulée sur un support de chromatographie de basse pression. La résine est lavée avec 20 volumes de tampon A (20 mM Tris HCl, 250 mM NaCl, 1 mM β -mercaptoéthanol) seul, ensuite par 5 volumes du même tampon en présence de 5 mM d'imidazole et enfin la protéine est éluée par une solution contenant le tampon A en présence de 250 mM imidazole. Le gel dénaturant montre deux espèces majoritaires correspondant à la taille de p34N (27 kDa) et p44C (7 kDa) (Figure 38). Les fractions issues de la première étape de purification, contiennent en majorité le p34N et p44C, mais comportent aussi un ensemble de contaminants. Il est alors évident qu'une seule étape de purification ne suffit pas pour obtenir des fractions très pures.

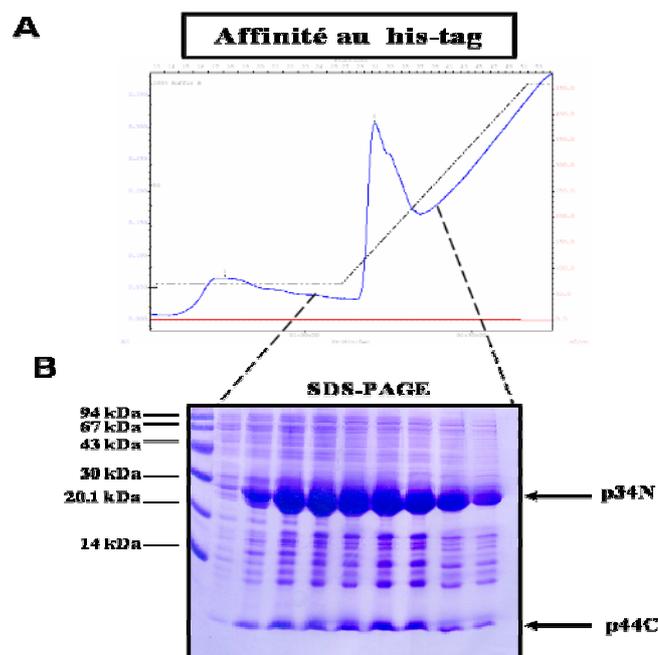


Figure 38 : Analyse des résultats de la chromatographie d'affinité au *his-tag* réalisée sur le complexe p34N/p44C.

A- Chromatogramme de l'affinité au *his-tag*.

B- Analyse des résultats par SDS-PAGE 15%.

3.2-Chromatographie sur colonne d'interactions hydrophobes

La fraction d'élution de la colonne d'affinité est additionnée d'un volume de sulfate d'ammonium 4 M, pH 7.4 jusqu'à atteindre une concentration finale de 500 mM en sulfate d'ammonium. L'échantillon est injecté sur une colonne de 1.5 ml de résine Phényl 650S (Tosohaas) équilibrée dans le tampon B (20 mM Tris-HCl pH 7.8, 50 mM KCl, 500 mM de sulfate d'ammonium, 0.5 mM DTT, 0.05 % CHAPS). La colonne est lavée avec le tampon A, puis l'élution de la protéine est effectuée par un gradient décroissant de sulfate d'ammonium. Il est intéressant de noter que les fractions de ce complexe sont nettement moins contaminées qu'en sortie de chromatographie d'affinité (Figure 39). Les fractions contenant le complexe sont rassemblées, l'extension *his-tag* est digérée par 0.25 U de thrombine bovine par mg de protéine à 4°C durant une nuit.

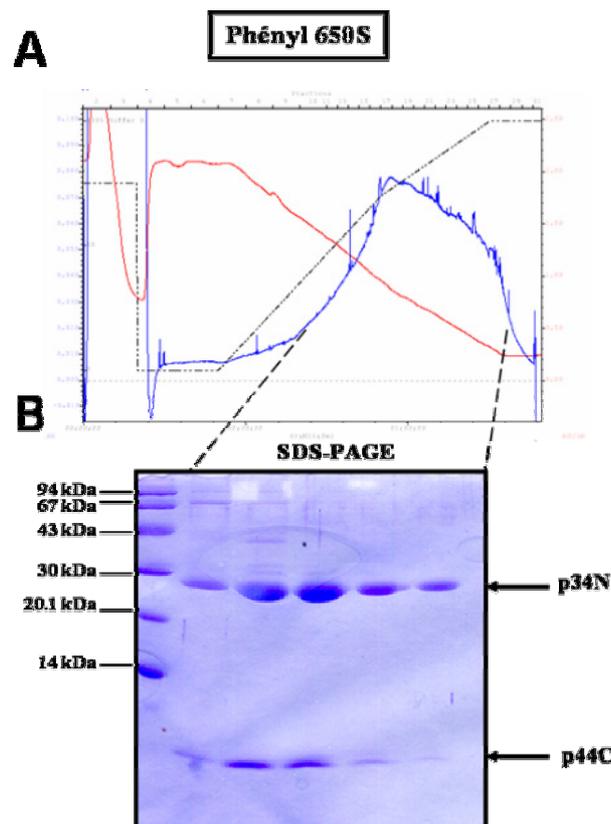


Figure 39 : Analyse des résultats de la chromatographie d'interaction hydrophobe (Phényl 650S) de p34N/p44C.

A- Chromatogramme de la Phényl 650S.

B- Analyse des résultats par SDS-PAGE 15%..

3.3-Chromatographie de filtration sur gel

Le complexe est élué à un volume correspondant à une protéine non agrégée avec un épaulement à la base du pic (Figure 40). Cet épaulement correspond à un polymère de p34N seule que l'on retrouve à chaque purification.

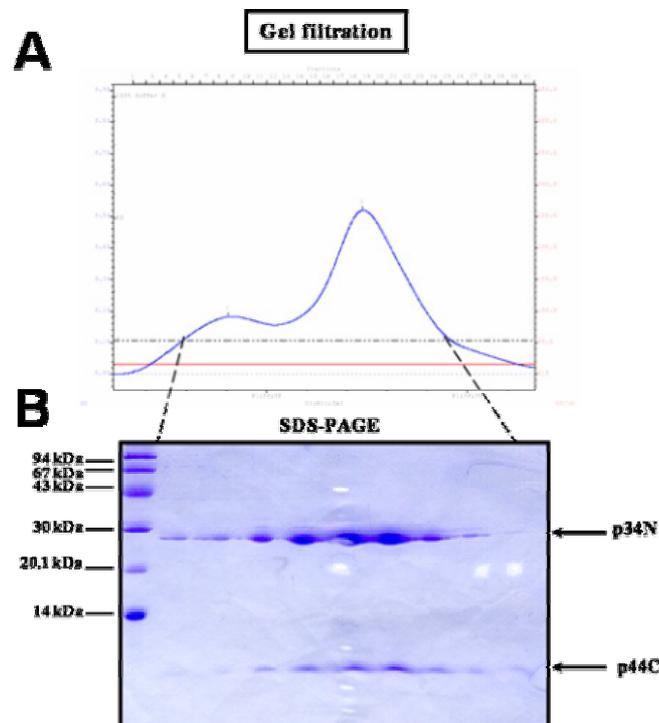


Figure 40 : Analyse des résultats de la chromatographie de filtration sur gel S200 16/60

A- Chromatogramme de gel filtration montrant un pic de complexe p34N/p44C et un épaulement correspondant à un polymère de p34N.

B- Analyse des résultats par SDS-PAGE 15%.

3.4-Rendement de purification

Le rendement de purification est de 25 mg de protéine pure pour 6 litres de culture. Il a été amélioré sensiblement par l'utilisation du milieu 2X LB à la place de LB simple pour les cultures. L'utilisation de l'incubateur à 12 litres a été très efficace également, puisqu'elle a permis de multiplier par un facteur 4 la biomasse ainsi que la quantité de protéine pure.

	Quantité de protéine (mg)	Taux de purification (%)
Extrait clarifié	800	-
Colonne d'affinité au his-tag	70	9
Colonne d'interaction hydrophobe	40	57
Colonne de gel filtration	25	62

Table 7 : Rendement de purification du complexe p34N/p44C à partir de 6 litres de culture.

4- Caractérisation du complexe p34N/p44C

Après purification, le complexe p34(1-233)/p44(331-395) a été caractérisé ensuite par électrophorèse en condition native (Figure 41) et par spectrométrie de masse, en conditions natives et dénaturantes (Figure 42).

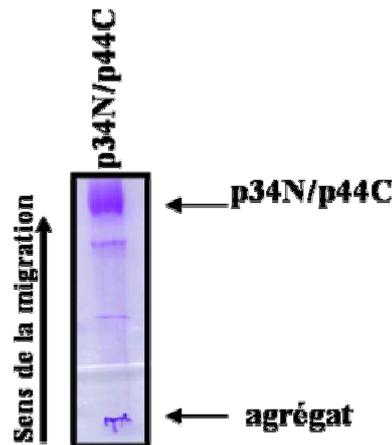
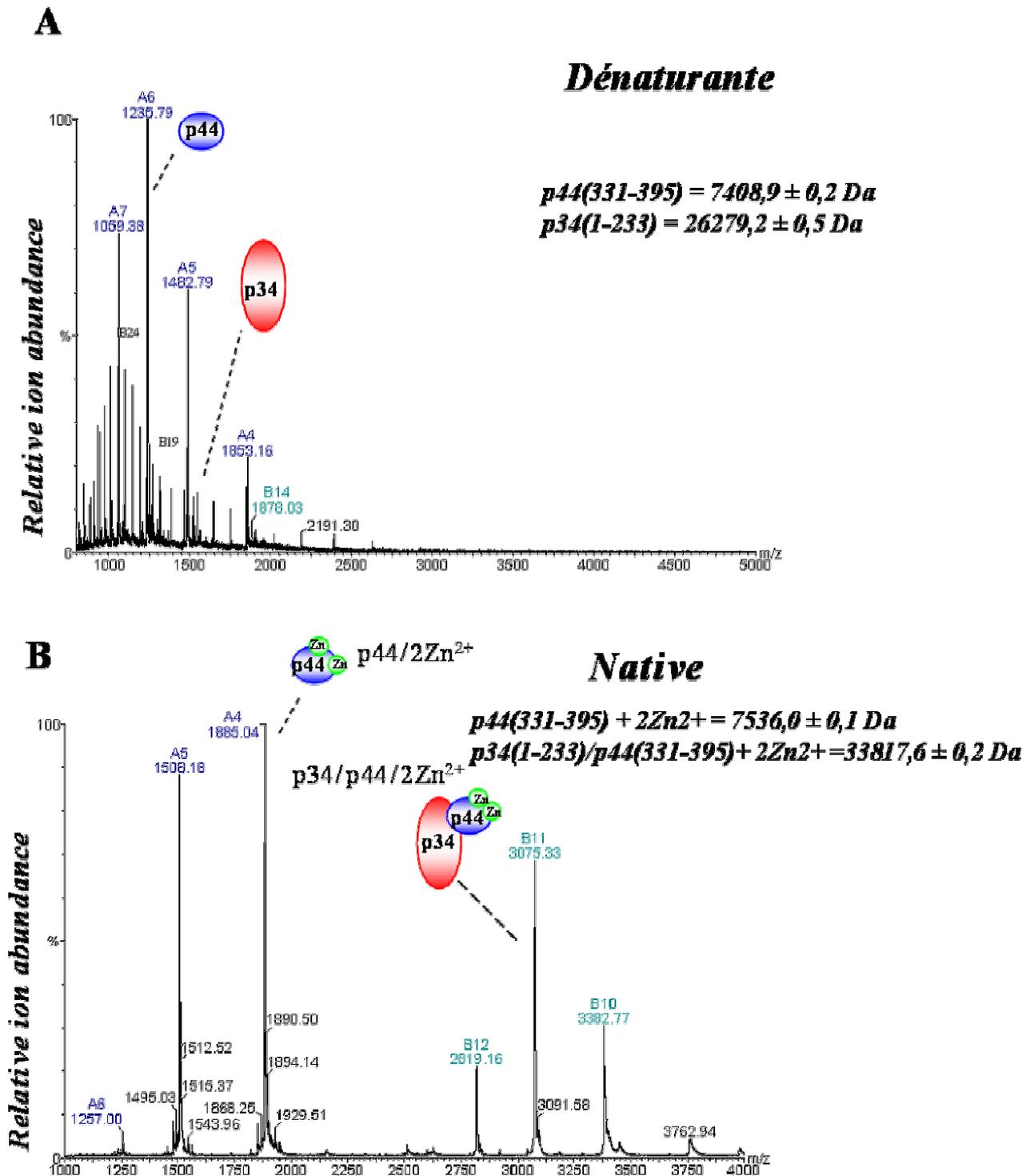


Figure 41 : gel natif (8-25) du complexe p34N/p44C.

Le gel natif montre l'existence d'une bande correspondant à seule population de complexe p34N/p44C. Cependant, cette bande semble un peu diffuse, ce qui pourrait être interprété par l'existence de zones flexibles au sein du complexe. De plus, la présence d'un agrégat de protéine est mis en évidence et doit être due à l'existence de polymère de p34N seul, ce qui a été observé en gel filtration (Figure 40). La spectrométrie de masse en conditions dénaturantes montre l'existence de deux protéines de masse 26284.2 Da et 7817.3 Da (Figure 42), ce qui correspond aux masses respectives de p34(1-233) et p44(331-395). L'analyse en conditions natives, montre l'existence de deux espèces : la première a une masse de 34227.3 Da correspondant à la masse d'un hétéro-dimère plus deux ions Zinc et la deuxième de masse 7536.0 Da correspondant à la masse de p44(331-395) plus 2 ions Zinc.



Ayant un complexe de qualité et de quantité suffisante (2 ml à 10 mg/ml), nous avons envisagé d'effectuer un criblage de conditions de cristallisation.

II- Cristallisation du complexe p34/p44

1- Généralités

Un cristal est un arrangement fini, tridimensionnel et périodique de molécules, constituant un réseau dont la maille est l'unité élémentaire. Des opérations de symétrie particulières déterminent l'agencement des molécules à l'intérieur de la maille. L'onde diffusée par le cristal correspond à la somme des ondes diffusées par chaque électron de toutes les mailles. Les rayons X seront diffusés à partir du cristal dans des directions privilégiées de l'espace, du fait de la périodicité des molécules dans le cristal. Il s'agit dans ce cas de la diffraction des rayons X, par opposition à la diffusion dans le cas d'une solution dans laquelle les molécules sont aléatoirement orientées. La diffraction aux rayons X dans un cristal est assimilée à la réflexion du rayon incident sur les plans réticulaires, l'onde réfléchi étant non nulle uniquement pour les plans satisfaisant la relation de Brag :

$$2d \sin \theta = n\lambda$$

d est la distance inter-réticulaire, λ la longueur d'onde et θ l'angle entre le rayon incident et le plan réticulaire.

L'image de diffraction du réseau réel cristallin est le réseau réciproque. Les dimensions de l'espace réciproques (\AA^{-1}) sont inversement proportionnelles aux dimensions de l'espace direct (\AA).

Le calcul de la densité électronique par synthèse de Fourier requiert la connaissance des phases, les intensités seules n'étant pas suffisantes. Si la densité électronique est de qualité suffisante, la connaissance de la séquence de la protéine doit permettre la construction manuelle du modèle tridimensionnel. Ce modèle est ensuite affiné en imposant des contraintes cristallographiques et stéréochimiques (chiralité, longueur de liaison..etc).

2- Introduction à la cristallisation

2.1- Les étapes de la cristallisation

La détermination de la structure tridimensionnelle d'une macromolécule biologique par cristallographie des rayons X nécessite l'obtention de monocristaux de taille et de qualité suffisantes. Un cristal est un empilement de molécules régulièrement agencées dans les trois directions de l'espace. Et c'est cet empilement qui détermine la qualité du cristal. En effet, plus l'ordre dans le cristal est parfait, plus le cristal est de qualité et plus son pouvoir de diffraction est important. La cohésion de l'édifice cristallin est assurée par diverses forces intermoléculaires de nature non covalente: interaction de types Van der Waals, liaisons hydrogènes et interactions ioniques. Les cristaux de macromolécules biologiques contiennent de 30 à 70 % de solvant situé dans des canaux. Ces canaux permettent aux petites molécules (atomes lourds, ligands, agents cryoprotectants, etc..) de diffuser à l'intérieur du cristal.

La cristallisation est un phénomène qui se déroule en trois étapes successives :

2.1.1-La sursaturation.

La sursaturation est un état métastable où la concentration de soluté est supérieure à sa solubilité. La concentration saline permet d'atteindre cet état de sursaturation (Figure 43).

2.1.2-La nucléation.

L'état de sursaturation favorise l'apparition de germes cristallins correspondant au franchissement d'une barrière énergétique, c'est la nucléation. Une forte sursaturation et une grande solubilité favorisent la nucléation. Lorsque la sursaturation est trop forte, la phase solide est constituée d'un précipité amorphe (Figure 43).

2.1.3-La croissance cristalline

La croissance cristalline se fait par fixation progressive des macromolécules dans le réseau tridimensionnel du cristal en formation. Il est très courant d'avoir une

apparition de germes sans croissance cristalline, ou que cette croissance soit trop rapide donnant ainsi lieu à la formation d'amas de cristaux de très mauvaise qualité. Cette étape reste limitante, car il ne suffit pas d'avoir des germes cristallins, il faut obtenir des cristaux monocristallins de taille macroscopique suffisante et de bonne qualité intrinsèque. Le pouvoir de diffraction d'un cristal est directement lié à l'ordre d'agencement des macromolécules. Plus l'ordre est parfait, plus le pouvoir de diffraction est meilleur.

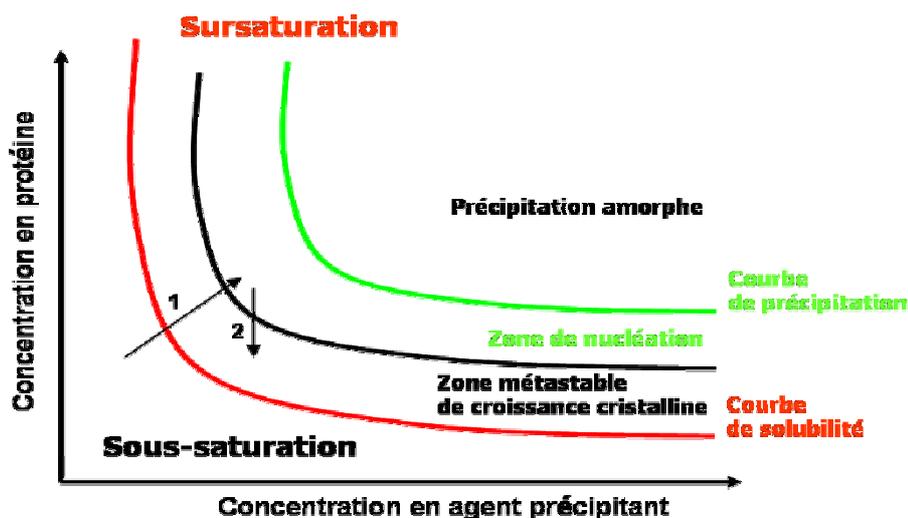


Figure 43 : Diagramme de phases.

Les étapes 1 et 2 correspondent à la nucléation et la croissance cristalline, respectivement.

2.2- Les paramètres influençant la cristallisation

Les paramètres influençant la cristallisation sont de deux types, physico-chimiques et propre à la macromolécule elle-même:

2.2.1- Les paramètres physico-chimiques

Les paramètres physico-chimiques sont les suivants: la présence d'agents précipitants de nature ionique ou non ioniques et à différentes concentrations, la température de cristallisation ainsi que le pH sont autant de facteurs agissant sur la solubilité de la macromolécule et par conséquent sur sa cristallogénèse.

A basse concentration saline, la solubilité des protéines croît avec la force ionique. L'effet des cations est alors prépondérant et les molécules de protéines ont tendance à s'associer voire s'agréger pour contrebalancer ces déficits électrostatiques. Ainsi, les

protéines ont tendance à s'exclure de la phase solide. Ce phénomène appelé le « *salting-in* » peut permettre la cristallisation.

Cependant, à haute concentration saline, la solubilité des protéines diminue si la force ionique augmente. La couche de solvation de la protéine est perturbée à cause des ions présents en solution et aboutit à une désolvation des protéines favorisant ainsi l'association voire l'agrégation de la protéine pour minimiser les surfaces hydrophobes exposées au solvant. Ce phénomène connu sous le nom de « *salting-out* » est très utilisé pour la cristallisation.

Il faut noter par ailleurs l'utilisation courante d'autres composants tels que les polymères organiques comme le polyéthylène glycol (PEG), certains alcools comme le MPD, ainsi que quelques détergents.

2.2.2- Les paramètres liés à la macromolécule

L'échantillon doit remplir aussi bien des critères quantitatifs que qualitatifs. En effet, il doit être très concentré, très pure et très homogène. Il existe des méthodes biophysiques permettant d'évaluer la qualité de l'échantillon : La spectrométrie de masse en conditions natives et dénaturantes permet d'avoir la masse exacte de la protéine et l'état d'homogénéité de l'échantillon. L'électrophorèse en conditions natives ainsi que la diffusion dynamique de la lumière DLS (*Dynamic Light Scattering*) permet d'évaluer l'homogénéité et la monodispersité de la protéine, avant la cristallisation. En principe, plus la protéine est homogène plus sa cristallisation est possible.

2.3- Les techniques de cristallisation

Il existe plusieurs méthodes de cristallisation : la cristallisation en « *batch* », la diffusion de vapeur, la cristallisation par dialyse et l'ensemencement.

2.3.1- La cristallisation en batch

La protéine en sursaturation est placée dans les conditions finales de cristallisation dans lesquelles l'apparition de microcristaux est susceptible de se produire (Figure

44). En « *microbatch* », le volume de protéine (2 à 10 μl) est placé sous huile pour empêcher l'évaporation de la goutte.

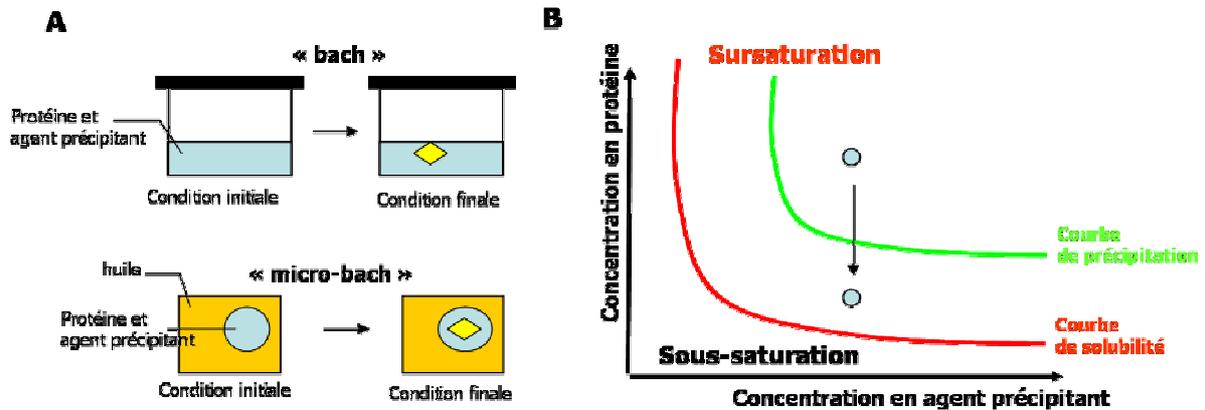


Figure 44 : Représentation schématique de la cristallisation « *en batch* ».

A-Représentation des conditions initiales et finales.

B-Diagramme de phase.

2.3.2-La cristallisation par diffusion de vapeur

Par cette méthode, la sursaturation est atteinte par échange-diffusion de solvants volatils (principalement de l'eau) dans une enceinte scellée, permettant ainsi l'apparition de germes cristallins. Dans cette enceinte sont placées, sans être en contact, deux solutions : (1) une solution de protéine plus agent précipitant et (2) une solution « de réservoir » contenant l'agent précipitant. Une goutte de la solution (1) est placée soit sur une lamelle sillonnée à l'envers (technique de la goutte suspendue), soit à l'endroit (technique de la goutte assise) (Figure 45).

Le volume du réservoir est généralement de 500 à 1000 μl , et celui de la goutte est de 1 à 50 μl .

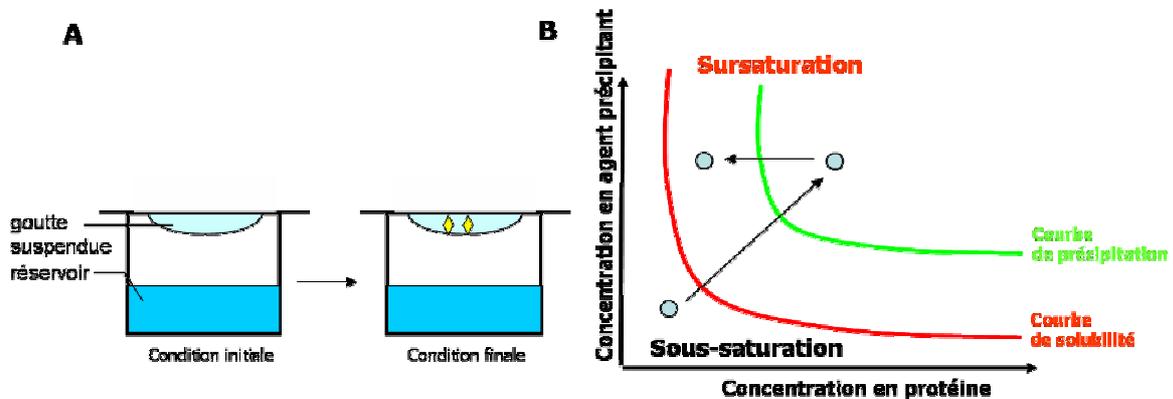


Figure 45 : Représentation schématique de la cristallisation par diffusion de vapeur.

A-Représentation des conditions initiales et finales

B-Diagramme de phase.

2.3.3-La cristallisation par dialyse

Par cette technique, la concentration de la protéine reste identique, mais la diffusion de l'agent précipitant à travers une membrane perméable permet à la protéine d'atteindre un état de sursaturation favorisant l'apparition de germes cristallins (Figure 46).

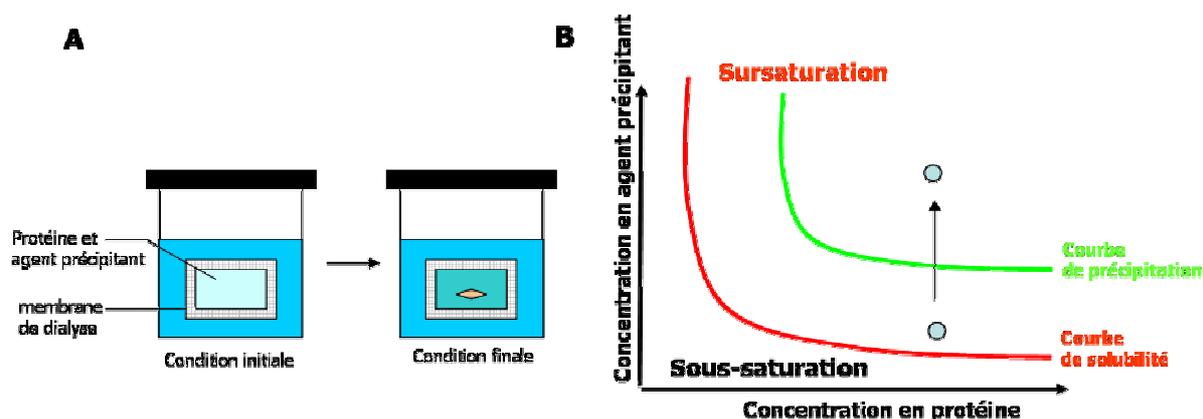


Figure 46 : Représentation schématique de la cristallisation par dialyse.

A-Protocole opératoire,
B-Diagramme de phase.

2.3.4-L'ensemencement

La technique d'ensemencement permet de régler le problème d'une trop forte nucléation. Des cristaux sont prélevés puis lavés dans une solution en sous-saturation (le but étant de nettoyer la surface du cristal), puis replacés dans une solution permettant la croissance des cristaux et non la nucléation. Lorsque le germe est microscopique, on parle de micro-ensemencement. L'ensemencement est aussi fait à partir de petits cristaux pour augmenter leur taille, on parle alors de macro-ensemencement (Figure 47).

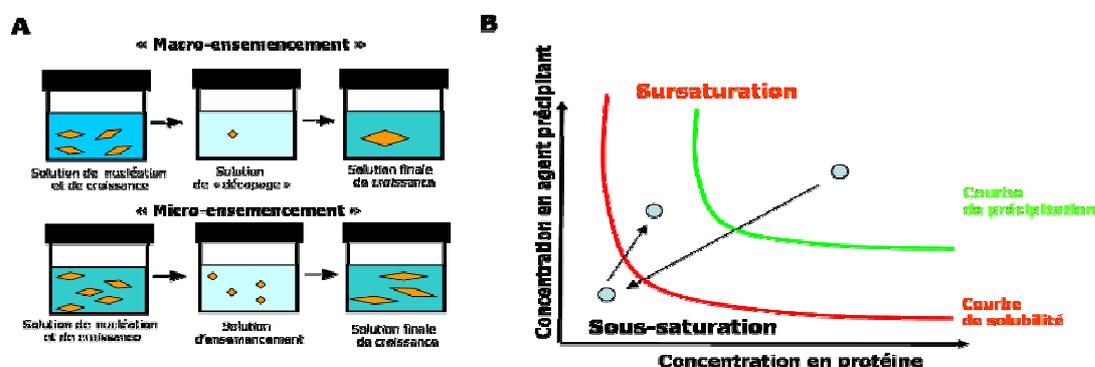


Figure 47 : Représentation schématique de la cristallisation par ensemencement.

A-Mode opératoire.
B-Diagramme de phase.

2.4-Criblage de conditions de cristallisation

La cristallisation étant dépendante de multiples paramètres aussi bien physico-chimiques (pH, température, agents précipitants) que propres à la molécule considérée (concentration, stabilité, homogénéité, hydrophobicité), il est *à priori* impossible de prévoir les conditions de cristallisations d'une macromolécule donnée. Ainsi, il est très courant de procéder à un criblage systématique de divers agents précipitants et additifs, diverses températures et différents pH, procédé relativement empirique, même si certaines études de cristallogenèse tentent d'établir des règles de cristallogenèse (Ducruix et Giégé., 2000). Il existe une panoplie de conditions initiales de cristallisations commercialisées par la société *Hampton Research* dans les kits de cristallisation Hampton screen I, II, Membfac, Natrix, PEG-ion, ammonium sulfate, ..etc. Le principe de ces conditions de criblage (*screens*) est la couverture globale d'un espace de cristallisation multidimensionnel dans un nombre limité de conditions (Plan factoriel). Ces conditions combinent plusieurs paramètres (agents précipitants, sels de nature et de concentrations différentes ainsi que différents pH).

Un certain nombre de « screens » ont été développé dans le laboratoire et concernent essentiellement la famille des récepteurs nucléaires.

Ces premiers criblages peuvent donner soit des cristaux soit des indications sur le comportement de la protéine vis-à-vis du type d'agent précipitant utilisé, du pH optimal, de la concentration de protéine et de la température optimale. Ces informations sont ensuite utilisées, et un criblage plus précis autour de ces conditions est effectué de manière à obtenir des cristaux ou d'améliorer les microcristaux obtenus.

2.5-La cryoprotection

Avant de soumettre le cristal à des radiations aux rayons X, il est monté dans un capillaire à 4°C. Cependant, cette technique cause des dommages dus à l'échauffement et à la génération de radicaux libres. C'est la raison pour laquelle, la technique de cryoprotection a été développée. Cette technique consiste à refroidir le cristal à une température de -170°C, en l'ayant trempé au préalable dans une solution

de cryoprotectant. Ceci permet d'éviter la formation de glace cristalline qui serait dommageable pour le pouvoir diffractant du cristal.

Pour le cas du complexe p34N/p44C nous avons procédé de la façon suivante : nous avons testé les cristaux en capillaire, à 4°C. Les clichés pris ont servi de référence. Ensuite, nous avons testé différents agents cryo-protectants à différentes concentrations dans la liqueur mère. Quand le cliché de diffraction révèle la présence de glace cristalline, visualisée par des anneaux de diffraction à une résolution de 3.6 Å, la concentration en agents cryoprotectant est augmentée. Les agents cryoprotectants les plus communément utilisés sont le glycérol, l'éthylène glycol et le PEG 400. L'isopropanol, les alcools de petites tailles, les sucres, et l'huile peuvent également être utilisés comme cryoprotectants. Il est préférable que la solution mère de la goutte soit elle-même cryo-protectante, ce qui permet la congélation directe des cristaux, réduisant ainsi leurs manipulations.

3-Cristallisation de complexe p34/p44

3.1-cristallogénèse

Des essais de cristallisation du complexe p34N/p44C ont été effectués à partir d'une concentration supérieure à 5 mg/ml. La protéine est conditionnée dans un tampon Tris-HCl 20 mM, pH 7.5, 250 mM NaCl et 2 mM β-mercaptoéthanol. Lors de la recherche des conditions de cristallisations, nous avons criblé plus de 200 conditions de cristallisation différentes à partir de « *screens* » commercialisés par la société Hampton. Les acides carboxyliques et notamment l'acétate de magnésium semblaient être de bons agents précipitants. De la même manière, les PEG en général et plus particulièrement le PEG 3350 donnaient des précipités cristallins.

En effet, les meilleurs précipités ont été obtenus suite à l'utilisation du *kit PEG-ion*, notamment les conditions 5 et 25, qui correspondent à 0.2 M MgCl₂, 20 % PEG 3350 et 0.2 M acétate de Mg, 20 % PEG 3350, respectivement. Cependant, il s'agissait de microcristaux, non exploitables, mais témoignant néanmoins d'un phénomène d'ordre certain. Nous avons donc entrepris un affinement de la condition 5, en faisant varier le pH, la concentration de PEG et la concentration en acétate de magnésium. Nous avons ainsi obtenu des cristaux de forme hexagonale, de taille

plus grande et donc totalement exploitable pour des tests ou des enregistrements au synchrotron. Ces cristaux ont été obtenus dans 9% PEG 3350, 50 mM Hepes pH 7.5 et 0.15 M Acétate de magnésium (Figure 48).

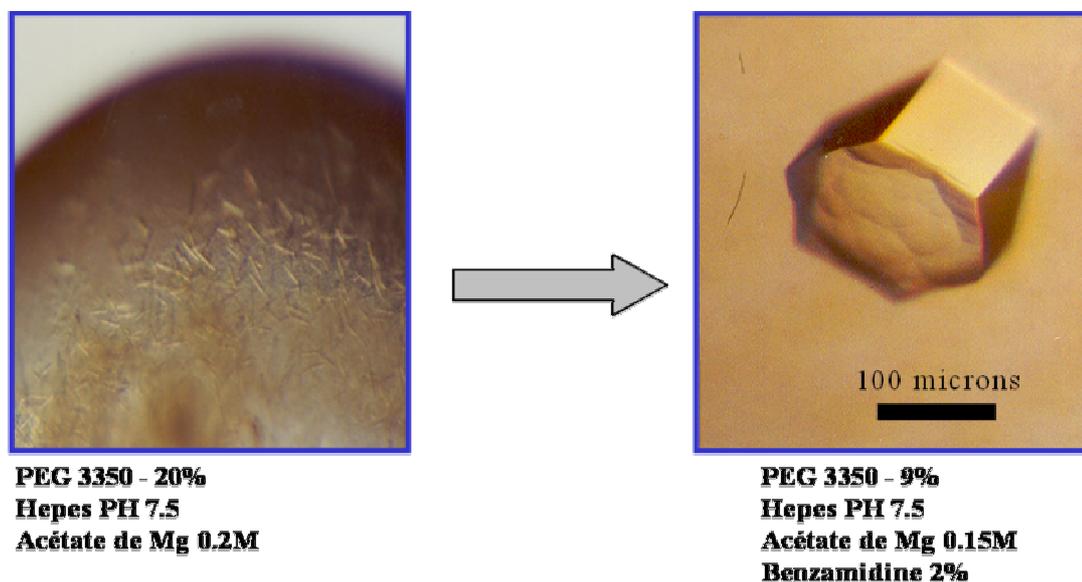


Figure 48 : Amélioration de la taille de cristaux en affinant les conditions de cristallisation.

Nous avons essayé de savoir l'effet des cations divalents sur la cristallisation. Pour cela, nous avons essayé de remplacer l'acétate de magnésium dans la même condition décrite précédemment, par le chlorure de magnésium, le sulfate de magnésium, le chlorure de calcium et le chlorure de manganèse. Tous ces ions ont permis d'obtenir des cristaux de qualité relativement médiocre. Néanmoins, seul le magnésium permet d'obtenir des cristaux qui diffractent.

Par ailleurs, nous avons effectué des essais de cristallisation à des concentrations différentes de protéine et avons constaté qu'à 2.5 mg/ml, il est possible d'obtenir des cristaux, mais leur taille est 10 fois plus petite que ceux obtenus à 5 mg/ml. A des concentrations au dessus de 5 mg/ml, la protéine a tendance à précipiter dans la goutte, aboutissant parfois à des cristaux complètement enfouis dans « une peau » de précipité, ce qui rend l'extraction et la congélation des cristaux très difficile. Nous avons établi que 5 mg/ml est la concentration optimale pour la cristallisation.

L'addition de 2% de benzamidine a contribué à l'augmentation de la taille des cristaux. L'utilisation d'analogues de cet additif tels que le p-amino dihydrochlorobenzamidine, le benzamidine-oxime, le 4-amino-benzamidine, le thio-

benzamidine, le chlorure de guanidine, le benzaldéhyde, le bromo-benzal et le méthyl-benzylamine, n'a permis l'obtention d'aucun cristal. Ceci montre l'importance de cet additif, sans l'expliquer.

L'utilisation du micro et macro-ensemencement pour améliorer l'aspect macroscopique des cristaux a permis d'obtenir des cristaux hexagonaux avec une surface plus nette. Cependant, la qualité de diffraction de ces cristaux demeurait inchangée.

3.2- Caractérisation des cristaux

Les cristaux obtenus ont été caractérisés d'abord par coloration en utilisant le colorant commercial *ISIT^R*, qui permet de colorer les cristaux de protéine en bleu. Une autre manière de caractériser ces cristaux consiste à les récupérer, les laver, les dissoudre dans l'acétate d'ammonium 50 mM et les analyser par spectrométrie de masse (Potier et al., 2000). Dans le cas des cristaux de p34N/p44C, nous obtenons une masse correspondant à la masse du complexe de départ ce qui indique que les cristaux obtenus sont bien des cristaux de p34N/p44C (Annexe 7).

3.3- Congélation

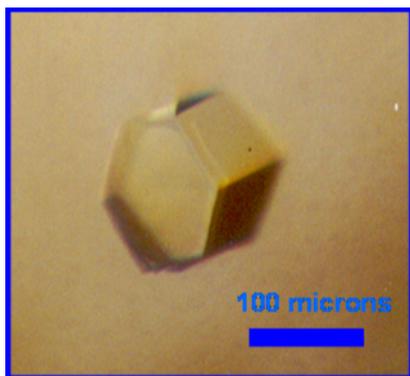
Avant de soumettre le cristal à des radiations aux rayons X, il a été monté dans un capillaire et irradié à 4°C. La limite de diffraction obtenue est de 3.5 Å, mais une telle irradiation cause des dommages importants au cristal, et rend quasiment impossible d'enregistrer un espace complet en utilisant un seul cristal. C'est la raison pour laquelle nous avons envisagé la cryoprotection des cristaux avant de les refroidir à -170°C, et de tester quatre agents cryoprotectants : l'Éthylène Glycol, le propane-diol, le PEG 400 et le glucose. Nous avons choisi de congeler les cristaux à 25% de PEG 400 et 25% d'Éthylène-glycol, puisque les meilleurs résultats ont été obtenus dans ces conditions. Pour éviter de transférer les cristaux de la liqueur mère à la solution cryoprotectante, nous avons essayé de cristalliser la protéine dans une solution mère déjà cryoprotectante. Malheureusement, les cristaux obtenus étaient beaucoup plus petits. De plus, ces cristaux étaient généralement entourés de « peau » de précipité, ce qui rendait leur manipulation très difficile.

3.4- Collecte de données

Les cristaux de taille 150x100x100 microns (Figure 49) ont été congelés dans 25% d'Éthylène-glycol, ont été testés et enregistrés au synchrotron sur BM30. La figure 50 montre un cliché de diffraction de ce cristal.

Le temps d'exposition était de 300s, avec une oscillation d'un degré. Les statistiques de ce jeu de données sont présentées dans la table 8.

Le groupe d'espace est P6122 ou P6522, avec des paramètres de maille de $a=b=80$, $c=380$. Le jeu de données a un facteur d'accord de 3.5%. Le paramètre de maille de 380 est trop important pour un complexe de 70 kDa.



P6122 or P6522
a=b=80, c=380

Figure 49: Cristal de p34N/p44C enregistré à l'ESRF sur BM 30.

<i>Limit</i>	<i>Å</i>	<i>I</i>	<i>error</i>	<i>Stat</i>	<i>Chi2</i>	<i>R-fac_t</i>	<i>R-fac_c</i>
25.00	8.57	60568.0	1097.3	712.8	1.477	0.028	0.030
8.57	6.85	21096.4	422.3	276.0	1.200	0.038	0.040
6.85	6.00	9413.9	328.5	203.8	1.002	0.043	0.045
6.00	5.45	9292.1	369.6	211.8	0.900	0.034	0.037
5.45	5.07	9843.5	430.1	240.5	0.858	0.036	0.038
5.07	4.77	10534.8	483.8	248.3	0.829	0.037	0.041
4.77	4.53	10811.1	506.6	267.9	0.804	0.038	0.041
4.53	4.34	9868.1	488.1	277.6	0.878	0.043	0.045
4.34	4.17	7302.1	371.2	221.8	0.898	0.047	0.049
4.17	4.03	6179.9	333.8	211.1	0.870	0.052	0.054
4.03	3.90	4959.7	314.1	226.7	1.131	0.073	0.085
3.90	3.79	4353.7	308.5	234.9	1.251	0.088	0.122
3.79	3.69	3213.4	284.7	216.1	1.131	0.101	0.085
3.69	3.60	3420.0	361.6	285.6	1.658	0.130	0.245
toutes réflexions		11611.8	427.3	273.1	1.129	0.039	0.035

Table 8 : Statistiques du jeu de données enregistré à l'ESRF sur BM30.

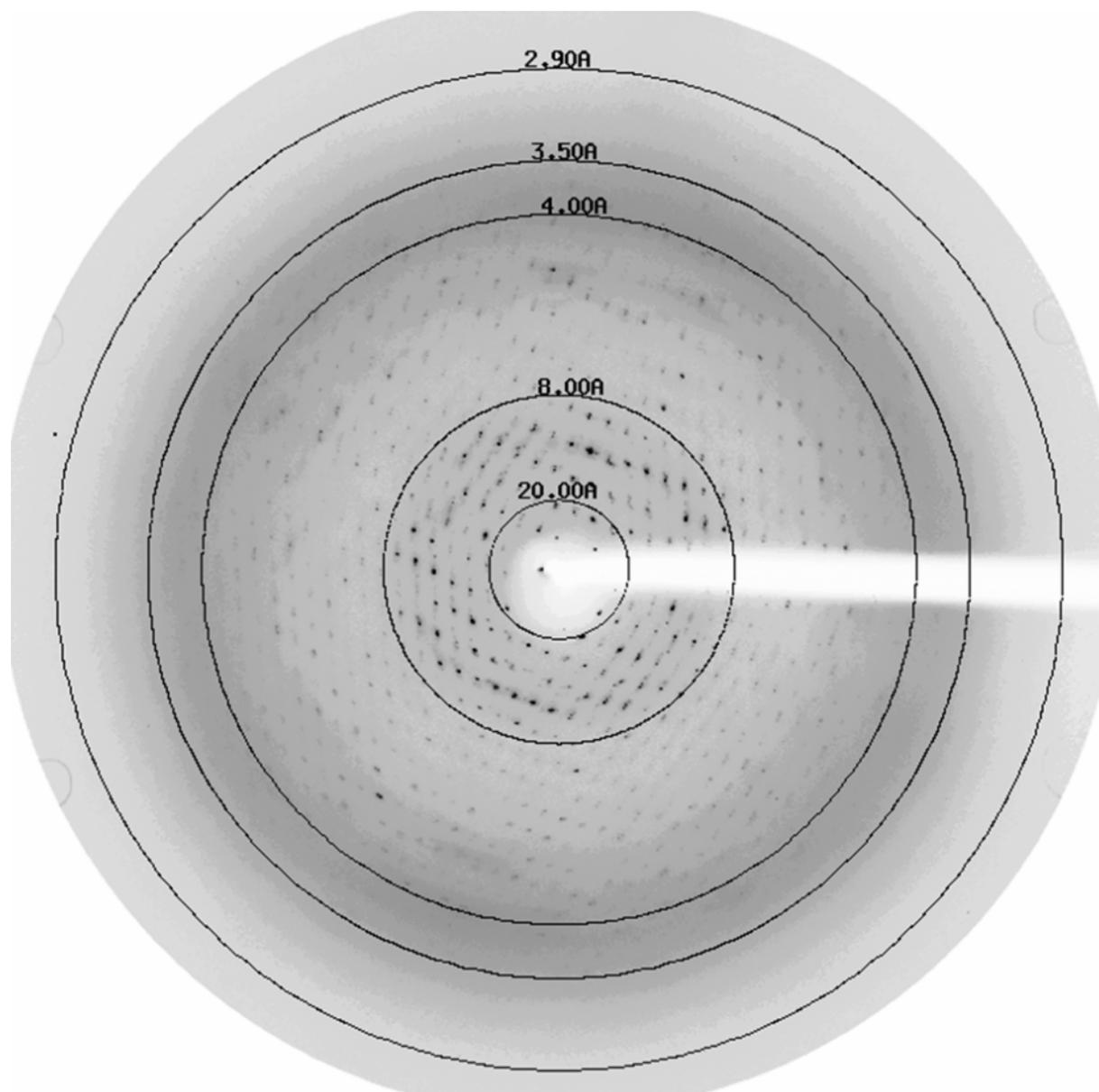


Figure 50: Cliché de diffraction d'un cristal de p34N/p44C enregistré au synchrotron sur BM30.

3.5- phasage

Le remplacement isomorphe multiple (MIR). Cela consiste à introduire dans le cristal de macromolécule native, par diffusion ou par co-cristallisation, des atomes lourds capables de diffuser fortement les rayons X. La position des atomes lourds dans la maille permet de calculer leur contribution aux amplitudes des rayons X diffractés et d'en extrapoler l'information de la phase.

Dans le cas des cristaux de p34N/p44C, nous avons criblé une dizaine de dérivés lourds par spectrométrie de masse et avons ainsi éliminé les dérivés lourds trop réactifs, qui dénaturent le complexe p34N/p44C comme le K_2PtCl_6 ainsi que les dérivés lourds non réactifs comme le $HgBr_2$ (Annexe 4). Nous avons gardé 2 dérivés lourds, un covalent et un non-covalent : le $Sm(SO_4)_3$ et le thiomersal (Ethylmercure thio salicylate) (Annexes 5 et 6).

Pratiquement, nous avons diffusé les atomes lourds en injectant 1 μ l de solution de dérivés lourds à différentes concentrations dans la goutte de 3 μ l où se trouve le cristal d'intérêt. Le cristal est ensuite congelé après 30 min, 1h00, 2H00 et 4H00.

Au synchrotron nous testions les cristaux les plus longtemps diffusés par les dérivés lourds, et dans le cas où le cristal diffracte encore, nous enregistrons un jeu de données. Dans le cas contraire, nous testions des cristaux moins exposés aux dérivés lourds. Tous les jeux de données enregistrés sont traités et mis à l'échelle avec le jeu de données des cristaux natifs. Malheureusement aucune substitution n'a été observée.

Nous avons essayé d'utiliser le pouvoir de diffusion du Zinc, puisque le complexe p34N/p44C possède au moins 2 Zincs. Aucun signal n'a été obtenu.

Remplacer naturellement la méthionine par la sélénométhionine en cultivant les bactéries dans un milieu minimum riche en sélénométhionine est une méthode très utilisée, que nous avons essayée intensivement sans succès. Deux approches ont été utilisées. La première consistait à utiliser la souche classique BL21(DE3) et inhiber la voie de biosynthèse de la méthionine. Elle a permis de synthétiser le complexe p34N/p44C avec un taux de substitution très faible (2 Sélénométhionines/10). La deuxième approche a consisté à utiliser une souche bactérienne auxotrophe à la méthionine B834, et n'a permis aucune synthèse protéique.

Nous avons également essayé de produire le complexe p34N/p44C, en exprimant p34N et p44C dans le même vecteur polycistronique. Ceci ne nous a permis d'obtenir que peu de protéine (moins de 1 mg pour 6 litres de culture) souvent sous forme protéolysée.

Le remplacement moléculaire. Cette méthode permet la détermination de phases initiales pour une molécule dont la structure est inconnue à partir d'une molécule « parente ». La molécule modèle de structure connue est placée et orientée dans la maille par l'utilisation conjointe de fonctions de rotation et de translation (Rossmann, 1990). Le remplacement moléculaire est une méthode de phasage que nous n'avons pas pu envisager étant donné qu'aucune structure d'un homologue du complexe p34N/p44C n'a été résolue et n'est disponible à la PDB (*Protein Data Bank*).

En résumé, la mauvaise qualité de diffraction des cristaux du complexe p34N/p44C, ainsi que la difficulté de phasage auxquelles nous étions confrontés, nous ont amené à envisager d'autres solutions.

4- Autres solutions

4.1- Mutagenèse

Dans le but d'améliorer la qualité de diffraction des cristaux de p34N/p44C, nous avons entrepris des expériences de mutagenèse sur p34N et p44C. En effet, l'extrémité carboxy-terminale de p44 n'étant pas bien conservée, nous avons généré deux versions tronquées de p44(331-395) sauvage : p44(331-381) et p44(331-390). Par ailleurs, certaines cystéines de p34N n'étant pas conservées, nous les avons mutées en serine, c'est le cas des cystéines 37, 142, 194 et 207. La majorité de ces mutants posait des problèmes d'agrégation. Les deux complexes que nous avons réussi à purifier étaient composés du p34N sauvage et de versions tronquées de p44C : p44(331-381) et p44(331-390). Malheureusement, ces deux complexes ne cristallisaient ni dans les conditions du sauvage ni dans un criblage plus large.

4.2- Purification

Lors de la purification initiale du complexe p34N par affinité au *his-tag* (Protocole A- Figure 51), nous avons remarqué que p34N avait tendance à être sur-représenté et à former des agrégats en gel filtration. C'est la raison pour laquelle, nous avons utilisé la fusion GST sur p44C pour purifier le complexe (Protocole B- Figure 51). Le profil de gel filtration obtenu est bien meilleur. En effet, aucun épaulement n'a été observé, contrairement au complexe purifié par affinité au *his-tag* (Figure 51). De plus, le complexe purifié grâce au protocole B a un poids moléculaire plus petit que celui purifié avec le protocole A. En effet, la courbe de calibration permet de calculer la masse moléculaire de l'espèce A et B correspondant à 89 et 37 kDa (Figure 51 D). Ceci a été confirmé par des expériences de diffusion de lumière (Table 9). Le gel natif réalisé sur le complexe p34N/p44C purifié des deux manières, montre que l'utilisation de la fusion GST sur p44C pour la purification, permet l'obtention d'un complexe de meilleur qualité ne contenant aucun agrégat (Figure 51C).

Des essais de cristallisation ont été effectués sur ce complexe nouvellement purifié et ont permis d'obtenir des cristaux sous forme de baguettes fines de petite taille dans les mêmes conditions que les cristaux dont nous avons enregistré les données auparavant (HEPES pH 7.5, 10 % PEG 3350, 0.15 M d'acétate de Mg). Un criblage plus large, de conditions de cristallisation, a permis d'obtenir des micro-cristaux dans les conditions suivantes : Tris hydrochloride pH 8.5, 8 % de PEG 8000 (Condition 36, *Screen Hampton I*) ; Tris K citrate monohydrate pH 8.3, 20% de PEG 3350 (Condition 47, *Screen Hampton PEG-ion*) et Tris HCl pH 7.5, 8 % de PEG 6000, 0.2 M d'acétate d'ammonium (Condition 10, *Screen LBD maison*) (Figure 52). Malheureusement, toutes les tentatives d'amélioration de la taille de ces cristaux ont été vaines.

	Rh (nm)	MM (kDa)
p34N/p44C (A)	4.00	85.8
p34N/p44C (B)	3.06	44.1

Table 9 : Résumé des résultats de diffusion de lumière (DLS) de deux complexes p34N/p44C, purifiés différemment. Le protocole de purification A correspond à une étape d'affinité au peptide hexa-histidine suivie d'une filtration sur gel. Le protocole de purification B correspond à une étape d'affinité au *his-tag* suivie d'une étape d'affinité à la GST et enfin une filtration sur gel. Ces étapes sont décrites dans la figure 54. Rh est le rayon hydrodynamique et MM correspond à la masse moléculaire calculée à partir du Rh et en considérant que les protéines soient globulaires.

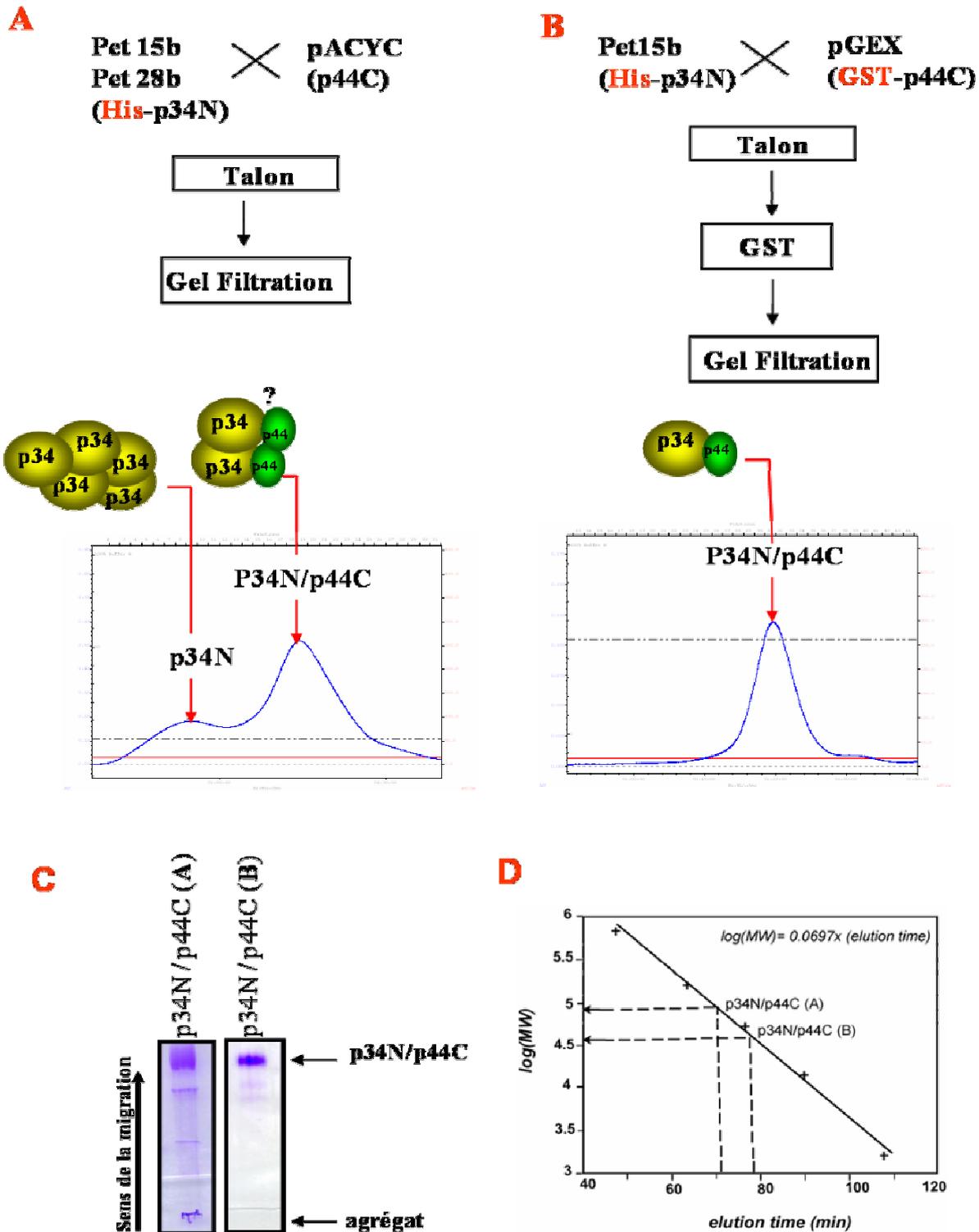


Figure 51 : Comparaison du comportement du complexe p34N/p44C en gel filtration en fonction de l'affinité utilisée.

A- Purification du complexe p34N/p44C par affinité au *his-tag* fusionné à p34N.

Le chromatogramme de gel filtration montre un épaulement correspondant à un agrégat de p34N.

B- Purification du complexe p34N/p44C par affinité au *his-tag* (*his-p34N*) et ensuite par affinité au glutathion (GST-p44C).

C- Gel natif du complexe p34N/p44C purifié de deux manières A et B.

D- Calibration de gel filtration S200 16/60 par des marqueurs *biorad* et indication des temps d'éluions de p34N/p44C ainsi que le logarithme de leur masse moléculaire.

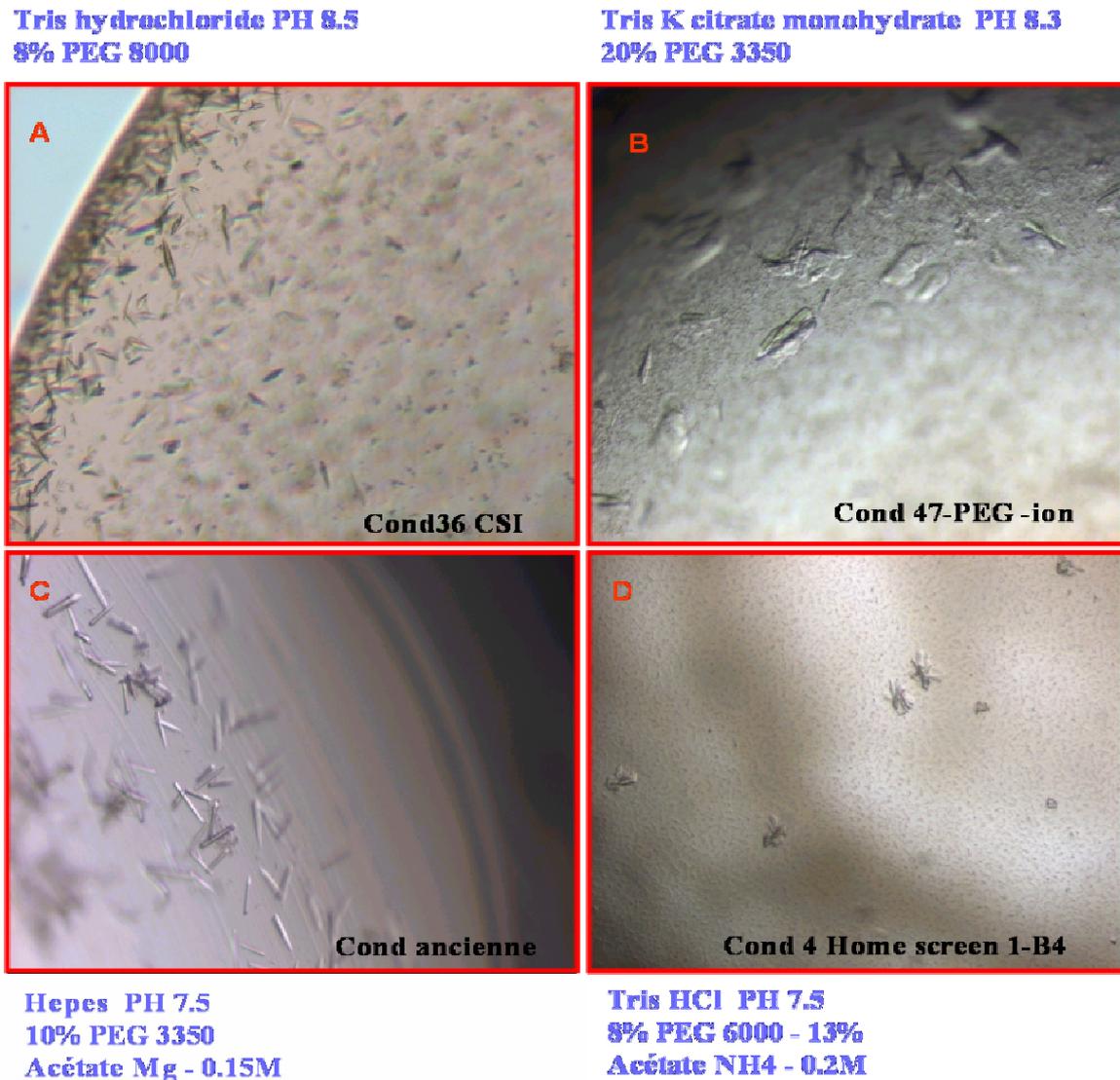


Figure 52 : Conditions de cristallisation du complexe p34N/p44C purifié suivant le protocole B.

A- Condition 36, Hampton screen I. B- Condition 47, Hampton screen PEG-ion.
C- Condition des anciens cristaux. D- Condition 4, Screen « LBD maison ».

4.3- Vers un complexe plus grand

La mauvaise qualité de diffraction des cristaux de p34N/p44C pourrait être due au fait que ce complexe soit non physiologique, qu'il soit un peu désordonné dans le cristal. C'est la raison pour laquelle, nous avons pensé que l'ajout d'un partenaire à ce complexe, pourrait le stabiliser et augmenter son ordre dans le cristal, augmentant ainsi son pouvoir de diffraction. Nous savons d'après nos études d'architecture, réalisées en utilisant le système de reconstitution de TFIID dans les cellules d'insecte, que p34, p44 et p62 sont capables de former un complexe stable. Nous avons essayé

de savoir d'une part quel domaine de p62 interagit avec le complexe p34/p44, et d'autre part quel domaine de p44 est important pour former le complexe p34/p44/p62. La Figure 53, montre l'existence d'un complexe minimal p34/p44C/p62C.

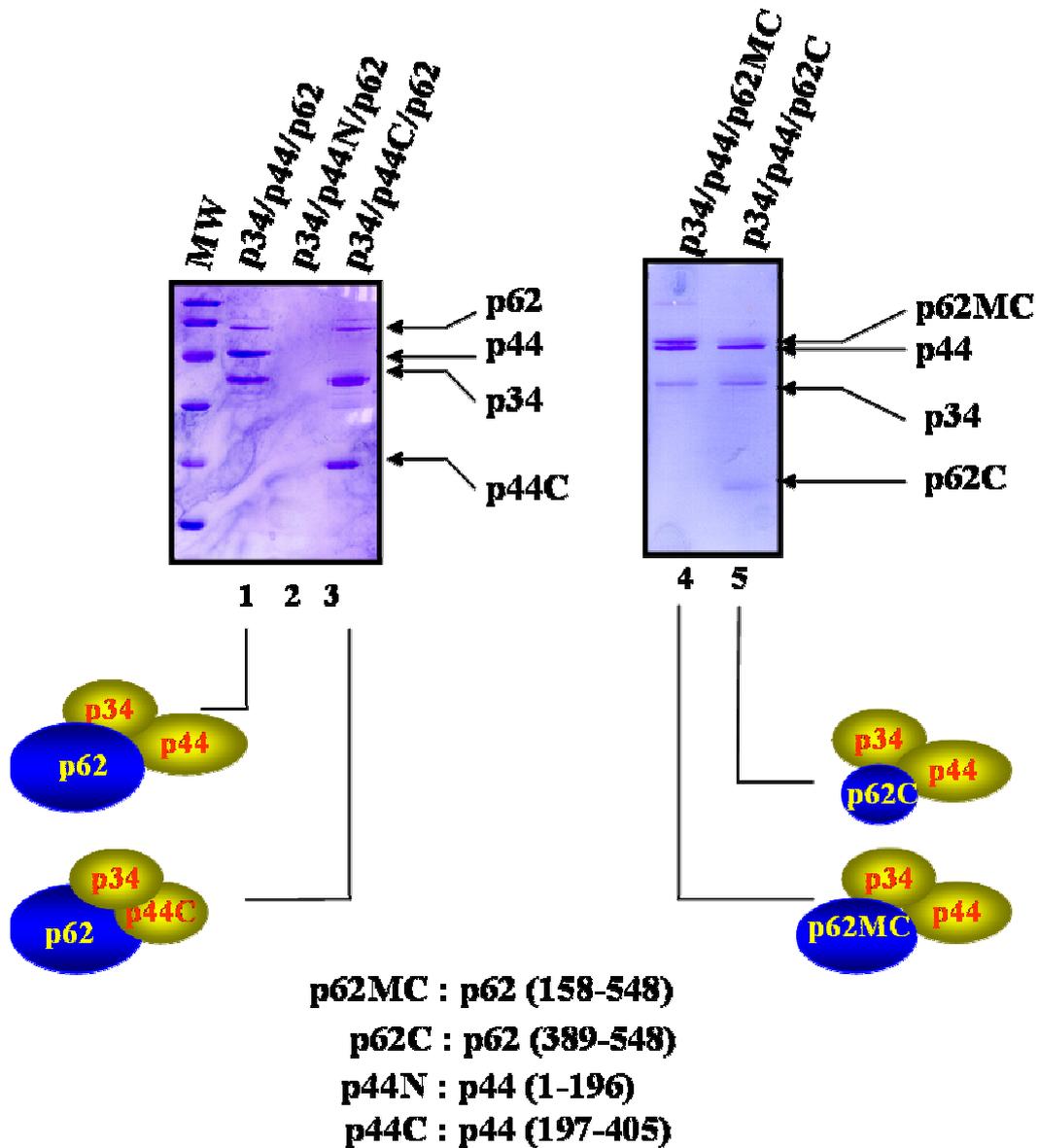


Figure 53 : Etude architecturale du complexe p34/p44/p62.

Les complexes sont purifiés par immuno-affinité à partir d'extraits de cellules d'insecte, infectées par une combinaison de différents baculovirus recombinants.

Les pistes 1, 2, 3, 4 et 5 correspondent au complexes p34/p44/p62, p34/p44(1-196)/p62, p34/p44(197-405)/p62, p34/p44/p62(158-548) et p34/p44/p62(389-548).

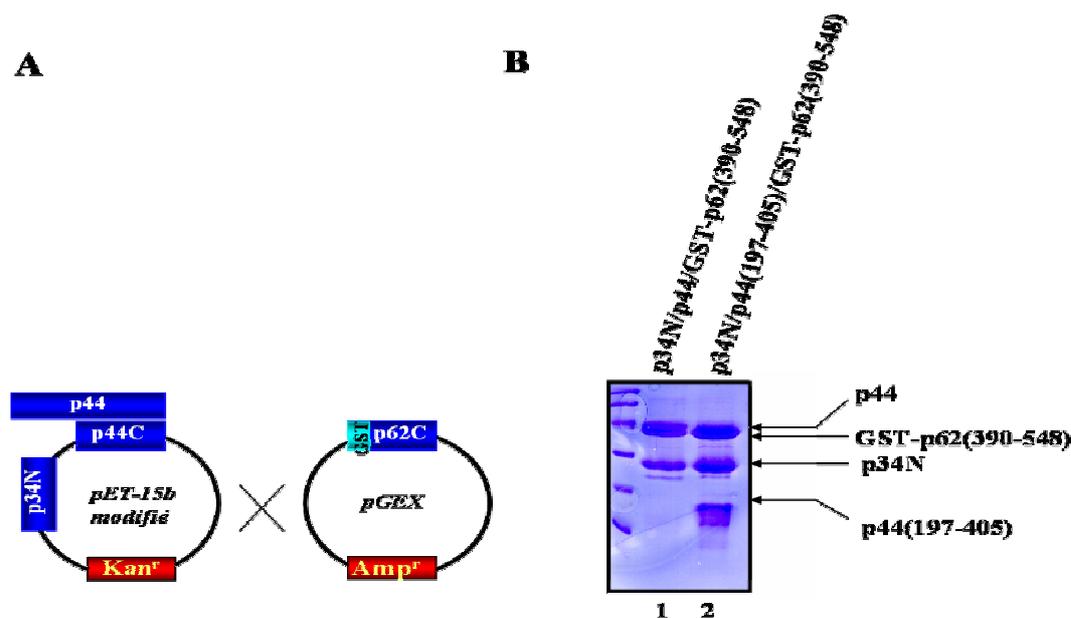


Figure 54 : Expression du complexe ternaire minimal p34N/p44C/p62C dans *E. coli*.
 A- représentation des vecteurs d'expression utilisés pour la co-expression du complexe ternaire.
 B- Les complexes p34/p44/GST-p62(390-548) et p34/p44(197-405)/GST-p62(390-548), ont été surproduits dans *E. coli* (piste 1 et 2, respectivement).

Le complexe minimal p34N/p44C/p62C a été surproduit dans *E. coli* en exprimant d'une part p34N et p44C dans le même vecteur polycistronique pET15b modifié et d'autre part p62C en fusion à la GST (Figure 54). Malheureusement après coupure de la GST, nous aboutissions sur une quantité faible de complexe ternaire (1 mg de complexe pur pour 6 litres de culture) ne permettant pas d'effectuer des essais de cristallisation. A ce stade, nous pensons que l'expression de ce complexe doit être encore optimisée, de manière à obtenir un complexe en plus grande quantité, en criblant d'abord des versions plus ou moins grande de p44C et en exprimant ensuite les trois protéines dans le même vecteur polycistronique.

4.4- Changement d'espèces

Après toutes les tentatives d'amélioration de la qualité de diffraction des cristaux du complexe p34N/p44C sans grand succès, et sachant que p34 et p44 sont relativement bien conservés de la levure à l'homme, nous nous sommes proposés de cribler la cristallisation de ce même complexe, chez d'autres espèces. Nous avons choisi cinq espèces différentes : *Saccharomyces cerevisiae*, *Encephalitozoon cuniculi*, *Arabidopsis thaliana*, *Danio rerio* et *Mus musculus*.

Le clonage des cinq ADNc codant pour p34N et p44C est en cours de réalisation, et devrait nous permettre d'obtenir des cristaux de meilleure qualité de diffraction.

Chapitre VII - Autres résultats

1- Architecture du core-TFIID.

Nous avons mis en évidence (Chapitre résultats, étude cristallographique de p34/p44) l'existence d'un complexe ternaire stable composé de p34N/p44C/p62C. Dans l'objectif d'élargir ce complexe, nous avons cherché à connaître la contribution du domaine carboxy-terminal de p34 à l'architecture du core-TFIID. Nous avons utilisé le système reconstitué de baculovirus pour exprimer le core-TFIID et une version tronquée ne possédant pas le domaine carboxy-terminal de p34. Les deux complexes ont été purifiés par immuno-affinité au peptide Flag. La présence des différentes sous-unités est confirmée par coloration au bleu de Coomassie et par western blot après SDS-PAGE. Le premier complexe purifié contient toutes les sous-unités du core-TFIID (p34, p44, p52, p62, XPD et XPB), alors que le complexe tronqué contient toutes les sous-unités à l'exception de p52 (Figure 55). En effet si on compare la piste 3 et 4, on peut voir une diminution très importante de la bande de p52 chez le mutant par rapport au sauvage. Nous vérifions que p52 est exprimé de la même manière dans les extraits bruts correspondant aux deux complexes.

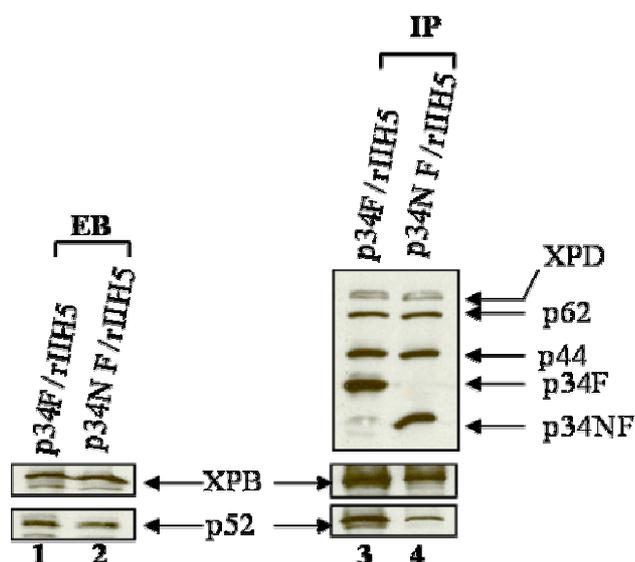


Figure 55 : Importance du domaine C-terminal de p34 pour l'ancrage de p52 au core-TFIID.

EB et IP correspondent à l'extrait brut et à l'immuno-purification, respectivement.

Les pistes 1 et 3 correspondent au complexe sauvage recombinant de 6 sous-unités (p34F/rIIH5).

Les pistes 2 et 4 correspondent au complexe recombinant de 6 sous-unités tronquées en C-terminal de p34 (p34NF/rIIH5). Les protéines sont détectées par western blot, en utilisant des anticorps monoclonaux.

Ce résultat montre que le domaine C-terminal de p34 est indispensable pour l'ancrage de p52 au core-TFIIH, ce qui suggère une interaction directe entre p34C-terminal et p52. Il reste maintenant à préciser le domaine de p52 impliqué dans cette interaction. Nos résultats préliminaires semblent désigner le domaine N-terminal de p52 dans cette interaction. Ces résultats suggèrent un complexe core-TFIIH minimal composé de p34 entier/p44C/p62C/p52N, qui pourrait être une cible potentielle pour la cristallisation.

2- Fonction du domaine PH/PTB de p62

Nous avons vu dans l'article 3 que la structure de p62(1-109) a révélé une homologie avec la famille de PH/PTB (*Pleckstrin Homology/Phospho-Tyrosine Binding domain*). La famille des protéines PH est essentiellement impliquée dans des interactions protéines/protéines, et aussi dans des interactions avec les lipides. Nous avons montré que le domaine PH/PTB de p62 n'est pas impliqué dans l'architecture de TFIIH, qu'il est indispensable pour la réparation de l'ADN par NER dans des tests *in vitro* et qu'il est capable d'interagir directement avec l'endonucléase XPG (article 3).

Nous avons voulu de savoir si p62(1-109) était capable de fixer des lipides. Pour ce faire, nous avons incubé p62(1-109) fusionné à la GST sur une membrane comportant 16 lipides différents (PIP STRIPS™, société *ECHELON BIOSCIENCES INC*). Après lavage, la protéine fixée aux lipides est révélée par *dot-blot*, en utilisant un anticorps monoclonal dirigé contre un épitope de la GST. La GST seule a été utilisée comme contrôle négatif.

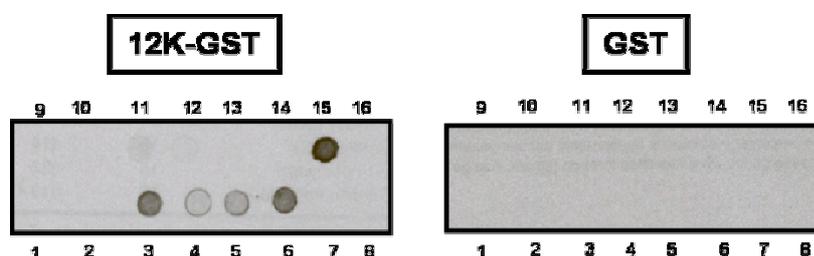


Figure 56 : Interaction entre le domaine 12K de p62 avec des lipides.

(1) Acide phosphatidique, (2) Phosphatidyl-choline, (3) Phosphatidyl-inositol, (4) Phosphatidyl 3 inositol, (5) Phosphatidyl-4P-inositol, (6) Phosphatidyl-5P-inositol, (7) Phosphatidyléthanolamine, (8) Phosphatidyl-choline, (9) Sphingosine-1P, (10) Phosphatidyl-3,4diP inositol, (11) Phosphatidyl-3,5diP inositol, (12) Phosphatidyl-4,5 diP inositol, (13), Phosphatidyl-3,4,5 triP inositol, (14) Phosphatidyl-amine, (15) Phosphatidyl-serine, (16) Blanc. Les protéines sont détectées par western blot en utilisant un anticorps anti-GST. Les expériences ont été faites de la même manière avec une fusion GST de p62(1-109) et avec la GST seule.

La figure 56 montre que p62(1-109) fusionné à la GST fixe fortement le phosphatidyl-sérine, le phosphatidyl-inositol et le phosphatidyl-5P-inositol. Le phosphatidyl-3P-inositol et le phosphatidyl-4P-inositol sont fixés plus faiblement, alors que les autres lipides ne sont pas du tout fixés. Ces interactions sont spécifiques à p62(1-109), la GST seule n'exhibant aucun signal.

Ce résultat est étonnant dans la mesure où aucun rôle de lipides dans les fonctions de TFIIH n'a été établi jusque là. La signification biologique de ce type d'interaction reste à définir.

3- Autre fonction du PH/PTB de p62

Nous avons vu dans l'article 4, que p62 contient un motif FW (Phe-Trp) conservé dans sa partie médiane. Des expériences de protéolyse ménagée et de spectrométrie de masse ont permis de définir un domaine stable de p62 contenant ce motif. Nous avons supposé que ce domaine pourrait être impliqué dans une interaction protéine-protéine. Un criblage d'interaction de ce domaine avec toutes les autres sous-unités du core-TFIIH a été effectué et a permis de montrer une interaction avec l'hélicase XPD. Le domaine médian (18K ou p62(159-390)), le domaine stable contenant le motif FW (8K ou p62(186-240)), le domaine PH/PTB de p62 (12K ou p62(1-109)) tous fusionnés à la GST ainsi que la GST seule, ont été exprimés et purifiés par affinité au Glutathion.

L'expérience consistait à mélanger les protéines fusionnées à la GST, après purification et dialyse, avec un extrait brut de cellules d'insectes contenant XPD et à les co-purifier par affinité au Glutathion. Après analyse par western blot (Figure 57), Le domaine médian, ainsi que domaine minimal comportant le motif FW, sont capables de co-purifier l'hélicase XPD, ce qui n'est le cas ni de la GST seule ni du domaine N-terminal de p62 (12K). Ce résultat montre que le domaine minimal médian (8K ou p62(186-240)) est suffisant pour interagir avec XPD. Comme ce domaine minimal contient une région conservée correspondant à un motif FW, on peut imaginer que p62 contacte XPD via ce motif. La résolution de la structure RMN de ce domaine permettra d'avoir des précisions sur la localisation de ce motif et de mieux comprendre le mode de reconnaissance p62/XPD.

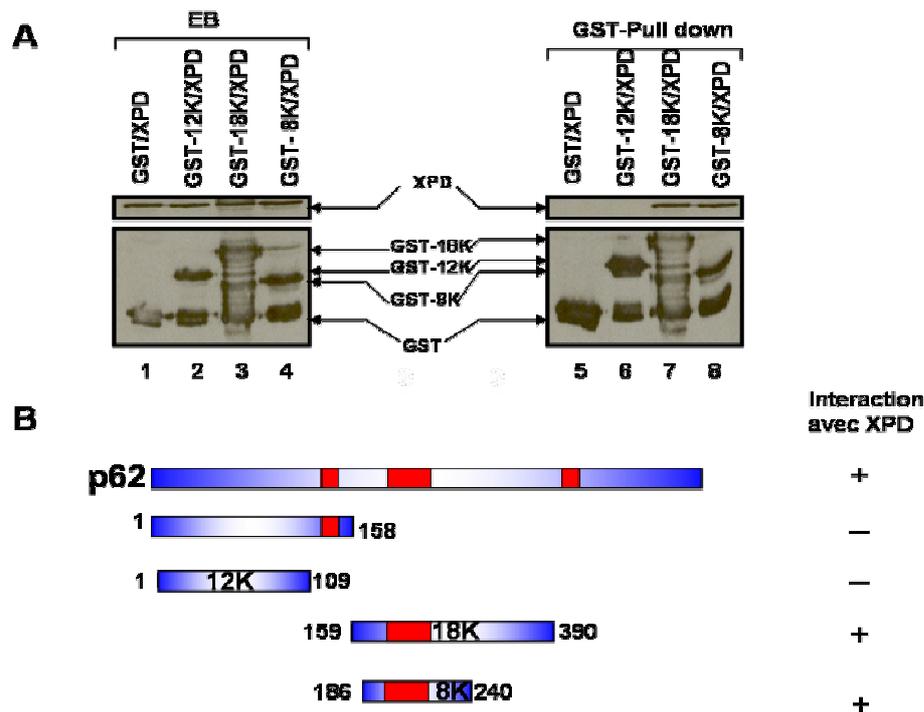


Figure 57 : Interaction entre le domaine minimal médian de p62 et XPD.

- A- Western blot utilisant des anticorps monoclonaux dirigés contre la GST et XPD. Les pistes 1, 2, 3 et 4 correspondent au mélange de XPD avec la GST, le 12K-GST, le 18K-GST et le 8K-GST, respectivement. Les pistes 5, 6, 7 et 8 correspondent à l'éluat de l'affinité GST des complexes 1, 2, 3 et 4, respectivement. EB et GST-pull down correspondent à l'extrait brut et l'éluat de l'affinité à la GST, respectivement.
- B- Représentation schématique des domaines de p62 qui interagissent avec XPD. En rouge sont indiqués les motifs FWxxΦΦ.

4- Etude de TFIIE

4.1-Protéolyse ménagée de TFIIE

Au cours de sa purification, TFIIE posait des problèmes de stabilité. Pour définir les domaines à l'interface d'interactions entre les sous-unités α et β , nous avons mis en place un protocole de protéolyse ménagée. Pratiquement, 10 μ l de TFIIE purifié (1.5 mg/ml) sont protéolysés en présence de 2 μ l de solution de chymotrypsine à différentes concentrations. Le mélange est chargé sur un gel dénaturant 12 % et analysé par électrophorèse en conditions dénaturantes. La figure 58 montre que TFIIE est encore stable à 0.01 mg/ml de protéase. La bande correspondant à la sous-unité α disparaît totalement pour générer un fragment de 43 kDa. Le séquençage d'Edman de ce fragment indique une digestion de la protéine en C-terminal. Ce résultat suggère une flexibilité de ce domaine par rapport au reste du complexe.

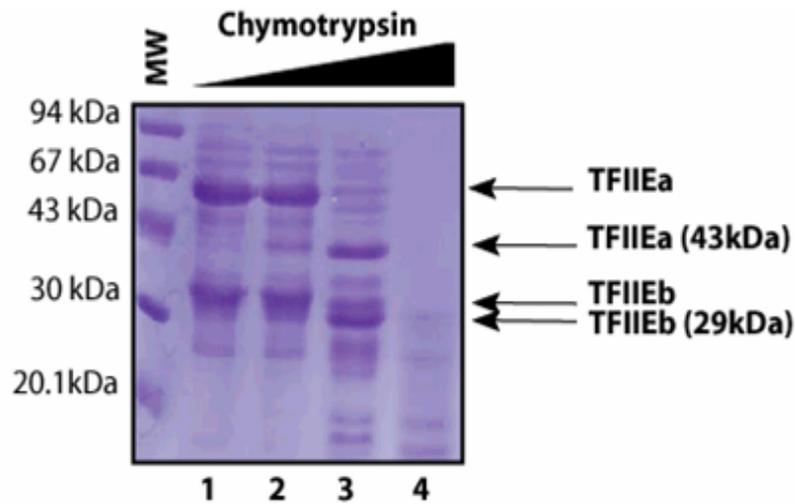


Figure 58 : Protéolyse ménagée du facteur TFIIIE.

10 μ l de TFIIIE purifié (1.5 mg/ml) sont incubés avec 2 μ l de solution de chymotrypsine à différentes concentrations (Piste 1, 2, 3 et 4 correspondant à 0 mg/ml, 0.01 mg/ml, 0.1 mg/ml et 1 mg/ml de chymotrypsine) à température ambiante pendant 1 heure. Le mélange est chargé sur gel dénaturant 12% et analysé par électrophorèse et coloration au bleu de coomassie.

4.2-Interaction entre TFIIIE et TFIIH

Pour mieux comprendre l'assemblage de TFIIIE avec TFIIH, qui constitue la dernière étape de l'initiation de la transcription, nous avons étudié l'interaction des deux complexes. Nous avons d'abord utilisé le système d'expression de baculovirus pour produire des sous-complexes de TFIIH à 2, 3, 4, 5, 6 et 9 sous-unités. TFIIIE en fusion à la GST au niveau N-terminal de sa sous-unité α , a été purifié par affinité au glutathion et par gel filtration. La capacité de chaque sous-complexe de TFIIH à interagir avec TFIIIE a été testée par « GST-pull down ». Pour chaque expérience nous vérifions l'expression des protéines dans l'extrait brut et nous nous assurons que les complexes ne se fixent pas de manière non spécifique à la résine GST. Les résultats ont été analysés par western blot et indiquent que les complexes comportant 3, 4, 5, 6 et 9 sous-unités sont co-purifiés par la GST-TFIIIE, ce qui n'est pas le cas du complexe à 2 sous-unités (p34/p44) (Figure 59). Ceci indique que TFIIH interagit avec TFIIIE non pas via p34 et p44 mais via p62.

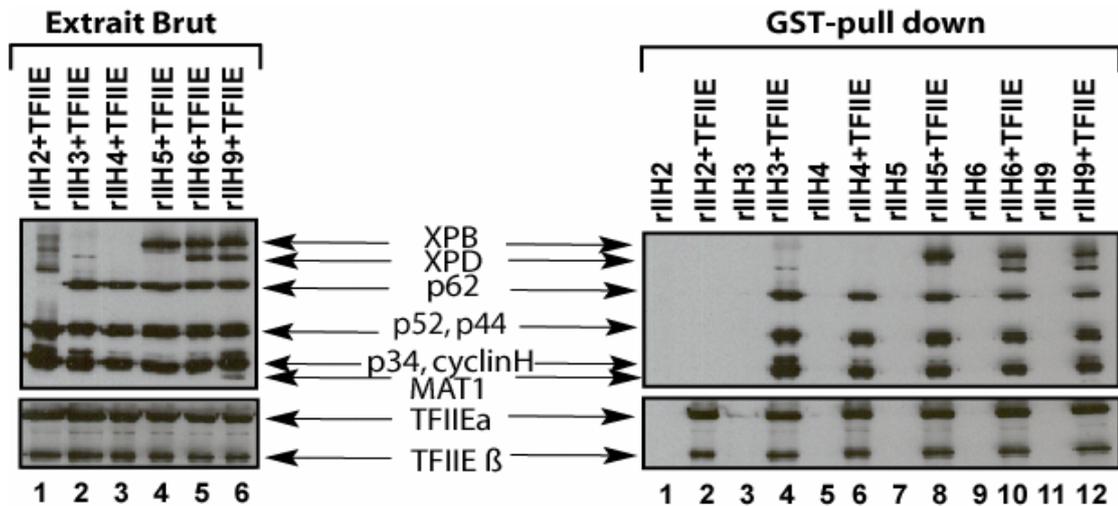


Figure 59 : p62 est suffisant pour interagir avec TFIIE.

Interaction des différents sous-complexes de TFIIH avec TFIIE. TFIIE est purifié par affinité à la GST et gel filtration. Pratiquement, 0.5 mg de GST-TFIIE est incubé pendant 1h00 à 4°C avec chaque 20 ml d'extrait brute correspondant à chaque sous-complexe de TFIIH (à partir d'une culture de cellules d'insecte de 250 ml) et 50 µl de résine GST. Après lavage avec 20 volumes du tampon A (20 mM tris HCl, 250 mM NaCl, 2 mM β-mercaptoéthanol) et 5 volumes du tampon A contenant 0.1% de Np40, les complexes sont élués par addition de glutathion 50 mM. Pour chaque réaction un témoin négatif est utilisé, il s'agit d'incuber les sous-complexes avec la résine GST en absence de GST-TFIIE. L'ensemble des résultats est analysé par western blot, en utilisant un anticorps monoclonal pour chaque protéine. rIIH2, rIIH3, rIIH4, rIIH5, rIIH6 et rIIH9 correspondent à p34/p44, p34/p44/p62, p34/p44/p62/p52, p34/p44/p52/p62/XPB, p34/p44/p52/p62/XPB/XPB et TFIIH entier.

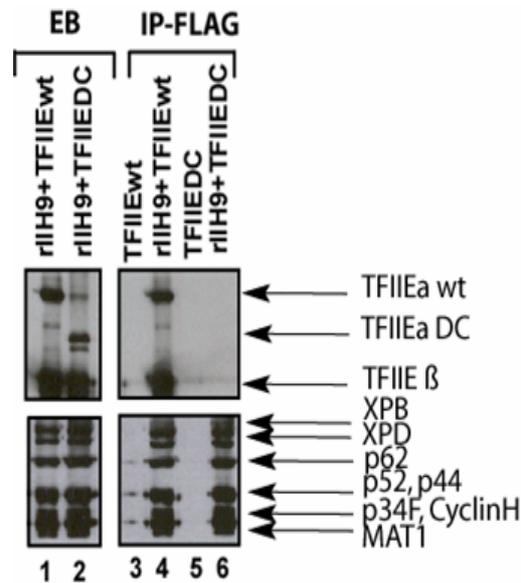


Figure 60 : Le domaine c-terminal de TFIIEα est suffisant pour interagir avec TFIIH.

A-Interaction d'une forme tronquée de TFIIE avec TFIIH. TFIIEwt et TFIIE DC correspond à TFIIE sauvage et TFIIEα (1-349)/β, respectivement. Les deux formes de TFIIE ont été purifiées par affinité au *his-tag* et gel filtration. Pratiquement, 0.5 mg de TFIIE sauvage (TFIIEwt) ou TFIIE tronqué (TFIIE DC) est incubé pendant 1h00 à 4°C avec 20 ml d'extrait brute de TFIIH entier recombinant (à partir d'une culture de cellules d'insecte de 250 ml) et 100 µl de résine FLAG-M2. Le TFIIH recombinant est fusionné à un peptide Flag en N-terminal de p34. Après lavage avec 10 volumes du tampon A (20 mM tris HCl, 250 mM NaCl, 2 mM β-mercaptoéthanol) et 5 volumes du tampon A contenant 0.1% de Np40, les complexes sont élués par addition de peptide compétiteur Flag. Pour chaque réaction un témoin négatif est utilisé, il s'agit d'incuber les deux complexes de TFIIE avec la résine Flag en absence de TFIIH. L'ensemble des résultats est analysé par western blot, en utilisant des anticorps monoclonaux pour chaque protéine. La piste 1 et 2 correspondent au mélange des complexes avant immuno-purification par le Flag. Les pistes 4 et 6 correspondent à l'éluat. Et les piste 3 et 5 correspondent aux témoins négatifs. EB et IP correspond à l'extrait brut et l'éluat de l'immuno-purification, respectivement; rIIH9 correspond au TFIIH entier recombinant.

Par ailleurs, nous avons généré une version tronquée de TFIIE (TFIIEDC correspondant à TFIIE α (1-349)/ β) et avons testé sa capacité à interagir avec TFIIH entier exprimé dans les cellules d'insecte. Pratiquement, TFIIE sauvage (TFIIEwt) et TFIIE tronqué (TFIIEDC) ont été purifiés par affinité au *his-tag* et par gel filtration, ensuite ils ont été incubés avec l'extrait brut de TFIIH entier contenant un Flag sur la sous-unité p34. Après l'immuno-purification au Flag, et un lavage extensif, les complexes sont élués par compétition au peptide Flag. Les résultats, analysés par western blot, montrent que TFIIE sauvage est co-purifié avec TFIIH, ce qui n'est pas le cas du TFIIE tronqué (Figure 60).

Ceci indique que TFIIE interagit avec TFIIH via le domaine C-terminal de la sous-unité α . Les résultats de protéolyse ménagée suggèrent que le domaine C-terminal de TFIIE α est flexible. Ce même domaine est responsable du recrutement de TFIIH. On peut penser qu'il s'agit d'un domaine adaptateur entre les deux complexes qui pourrait être stabilisé par l'interaction avec TFIIH.

Chapitre VIII - Discussion et perspectives

1-Expression et purification de TFIIH

Une étude structurale nécessite l'obtention d'un échantillon de qualité et de quantité suffisantes. La purification du TFIIH endogène à partir des cellules HeLa et du core-TFIIH recombinant à partir des cellules d'insecte, ont permis de résoudre la structure de TFIIH par microscopie électronique à basse résolution (38Å). Dans l'objectif d'augmenter la résolution, nous avons essayé d'améliorer la purification du TFIIH recombinant. Pour ce faire, des baculovirus recombinants exprimant les quatre sous-unités du core-TFIIH, fusionnées en N-terminal au peptide FLAG, ont été générés. Une procédure simple de purification par immuno-affinité au peptide Flag a été établie et a permis de produire un TFIIH recombinant entier, actif en transcription et structuralement similaire au TFIIH endogène.

Du point de vue quantitatif, 100µg de TFIIH endogène pur sont obtenus à partir de 100 litres de culture de cellules HeLa. Quant au TFIIH recombinant, 10µg sont obtenus à partir d'un litre de culture de cellules d'insecte Sf9. Cependant, ces quantités de protéine restent trop faibles pour permettre de cribler des conditions de cristallisation. Au-delà de ces considérations quantitatives, l'échantillon de TFIIH endogène pose un certain nombre de problèmes d'ordre qualitatif. En effet, il est assez peu homogène et a tendance à s'agréger. Nous pensions améliorer sa qualité en utilisant un TFIIH recombinant purifié par immuno-affinité, il n'en est rien. La qualité de l'échantillon est à reconsidérer en optimisant les deux paramètres suivants :

Purification

L'utilisation d'autres « tag » ou protéines de fusion comme la GST, la thyrodoxine, la CBP (*Calmoduline Binding Protein*) ou le tag-HA (*hémagglutinine*) seules ou en combinaison par exemple avec le Flag ou le *his-tag*, pourrait être une alternative pour la purification d'un TFIIH homogène et stoechiométrique. Il faut noter que, sur un gel dénaturant, les sous-unités du CAK (cdk7, cycline H et Mat-1) paraissent souvent sous représentées par rapport aux sous-unités du core. De plus la densité électronique calculée et attribuée au CAK est trop faible par rapport à sa masse moléculaire. C'est la

raison pour laquelle, il est plus prudent d'essayer d'améliorer la résolution de TFIIH en travaillant sur le core-TFIIH recombinant (complexe à 5 ou à 6) de manière à éviter ce problème de sous-représentation du CAK.

De manière générale, l'optimisation de la purification n'a de sens que si l'expression est optimale.

Expression

L'expression est certainement le paramètre le plus limitant pour l'obtention d'un TFIIH de bonne qualité. Dans le système baculovirus, notre expérience a montré que la bonne co-expression des sous-unités de TFIIH est directement liée à la combinaison des virus utilisée. Une autre combinaison de virus a été utilisée par Fukuda et collaborateurs. Trois vecteurs baculovirus pAcAB3 ont servi à l'expression de toutes les sous-unités de TFIIH (Fukuda et al., 2001). Le premier vecteur exprimait *cdk7*, cycline H et Mat-1. Le deuxième exprimait XPB, p34 et p52. Le troisième exprimait XPD, p44 et p62. Ce TFIIH purifié par double affinité au Flag et *his-tag* semble stoechiométrique en comparaison au notre. Il convient de reproduire cette combinaison de virus de manière à évaluer la qualité de l'échantillon par microscopie électronique.

Par ailleurs, l'établissement d'une souche stable de cellule HeLa contenant une version de TFIIH avec un « *tag* » particulier, pourrait également être une alternative à l'amélioration de l'expression de TFIIH. Ceci a déjà été fait et a permis d'obtenir un TFIIH fonctionnel (Winkler et al., 2000).

L'ensemble des optimisations discutées ici concerne un TFIIH humain. Envisager de changer d'organisme pourrait être une solution pour l'amélioration de la qualité de l'échantillon. Le groupe de RD. Kornberg a réussi à obtenir des cristaux 2D du core-TFIIH de levure (Chang and Kornberg, 2000), avec un échantillon de très grande qualité. Essayer de faire de même avec un complexe plus grand en utilisant le core-TFIIH comme point de départ, nous paraît être une alternative intéressante. Il existe des souches modifiées de levure, contenant un *TAP-tag* (*Tandem Affinity Purification-tag*) sur certaines sous-unités de TFIIH. La purification du TFIIH à partir de ces souches de levure, permettrait d'évaluer sa qualité.

La société *cellzome* met à disposition sur son site (<http://yeast.cellzome.com>) la composition des complexes de levure immuno-purifiés avec le système *TAP-tag*. Les sous-unités de TFIIH de levure appartiennent à plusieurs complexes, dont le plus représentatif est ID-131. Il contient Tfb1/p62, Tfb3/Mat1, Tfb4/p34, Ssl1/p44 et Kin28/Cdk7. Les homologues de levure de p52, XPB, XPD et la cycline H n'ont été répertoriés dans aucun complexe. A titre comparatif, pour l'ARN polymérase II, 9 sous-unités sur 12 ont été détectées dans le complexe ID-145. Ceci traduit certainement une différence de stabilité entre la polymérase et TFIIH. Du point de vue physiologique, cette différence de stabilité peut trouver une explication dans le fait que TFIIH puisse se trouver soit au sein du complexe de pré-initiation de la transcription (PIC), soit au sein d'un complexe de réparation, alors que l'ARN polymérase II appartient uniquement au PIC. TFIIH serait alors flexible et permettrait des interactions dans différents contextes biologiques. Si cette hypothèse est vraie, son expression dans un complexe plus large de transcription ou de réparation, devrait améliorer sa stabilité. Dans ce sens, une étude structurale par microscopie électronique de complexe TFIIH/TFIIE, ADN/TFIIE/TFIIH, XPG/TFIIH ou encore ADN/XPC-HR23B/TFIIH est envisageable. Cette étude nécessite au préalable une caractérisation biochimique des conditions de formation de ces complexes. L'utilisation de la technologie de résonance plasmonique de surface (*Biacore*) nous a permis de trouver la concentration en sel adéquate pour la fixation de TFIIE à l'ADN simple brin (fixation à 50mM et non pas à 150mM de NaCl). Elle devrait par ailleurs nous permettre de déterminer les conditions de formation des complexes. Une fois ces conditions établies, l'homogénéité de l'échantillon devrait être améliorée et l'analyse par microscopie électronique simplifiée.

Quant à la cristallisation de TFIIH entier, elle a été limitée par la difficulté de préparer un TFIIH pur et très concentré. Les études d'architecture du core-TFIIH nous ont permis de sur-exprimer un complexe stable : p34N/p44C. Les cristaux obtenus ayant un pouvoir de diffraction limité, nous nous sommes proposés de cribler sa cristallisation chez d'autres organismes, de manière à améliorer la qualité de diffraction. Par ailleurs, nous avons pu former un complexe minimal stable du core-TFIIH : p34N/p44C/p62C. Ce complexe était sous-représenté en p62C. C'est la raison pour laquelle nous nous sommes proposés d'abord de tester différentes constructions de p62C et ensuite

d'exprimer les trois protéines dans un même vecteur polycistronique. Ce complexe une fois caractérisé, représenterait une cible potentielle pour la cristallisation, et permettrait d'apporter des informations structurales sur l'organisation du core-minimal de TFIIH. L'utilisation de la protéolyse ménagée couplée à la spectrométrie de masse, nous a permis de délimiter deux domaines stables de p62. Une telle stratégie permettra sans doute l'identification au sein de TFIIH de nouvelles cibles pour la cristallographie et la RMN.

2- Etude architecturale de TFIIH

Depuis la découverte du facteur TFIIH, un effort a été fourni pour l'étude de son architecture. Le premier réseau d'interactions des sous-unités de TFIIH a été proposé en 1996 (Iyer et al., 1996). La table 10 représente l'ensemble des interactions établies entre les sous-unités de TFIIH à ce jour.

Chez le levure, d'une part des expériences de double hybride ont montré d'une part que Tfb2 interagit avec Rad 25 (Feaver et al., 2000) et d'autre part l'étude d'un mutant de Tfb1 a montré que Tfb1 et Ssl1 interagissent directement (Matsui et al., 1995).

Chez l'Homme, p34 est le partenaire de p44 (Fribourg et al., 2001). p44 interagit aussi avec XPD et stimule son activité hélicase (Coin et al., 1998; Seroz et al., 2000). Par ailleurs, p44 est capable d'interagir avec p62 (Jawhari et al., 2002b; Matsui et al., 1995). Quant à la sous-unité Mat1, il a été montré qu'elle interagit avec la cycline H, cdk7, XPB et XPD (Busso et al., 2000). Un certain nombre de domaines d'interaction ont été localisés, c'est le cas du domaine C-terminal de XPD qui contacte le domaine N-terminal de p44, ou encore du domaine amino-terminal de XPB qui interagit avec les domaines N-terminal et médian de p52 (Coin et al., 1998; Jawhari et al., 2002a).

	P34	P44	P52	P62	XPB	XPB	Cdk7	CyclinH	MAT1
P34	-	+	-		-	-	-	-	-
P44	+	-	-	+	+	-	-	-	-
P52	+	-	-	+		+	-	-	-
P62	-	+	+	-	-	-	-	-	-
XPB	-	+	-	+	-	+	-	-	+
XPB	-	-	+	-	+	-	-	-	+
Cdk7	-	-	-	-	-	-	-	+	+
CyclinH	-	-	-	-	-	-	+	-	+
MAT1	-	-	-	-	+	+	+	+	-

Table 10 : Résumé des interactions entre les différentes sous-unités de TFIIH.

En plus et moins sont indiquées les interactions caractérisées et non caractérisées, respectivement..

En plus et en gras, sont indiquées les interactions identifiées au laboratoire.

Nous avons pu apporter plus de précisions à cette étude, en montrant d'abord l'existence d'un complexe stable p34N/p44C, et ensuite en établissant que le domaine C-terminal de p62 est capable de former un complexe ternaire avec p34N/p44C. Par ailleurs nos résultats préliminaires suggèrent d'une part une interaction entre le domaine C-terminal de p34 et p52 et d'autre part une interaction directe entre les deux hélicases XPB et XPD (voir chapitre VI « autres résultats »).

Une étude d'immuno-localisation de TFIIH a permis de situer les sous-unités cdk7, p44 et XPD dans la structure du TFIIH entier. p44 et XPD ont été localisés au niveau de la structure en anneau et cdk7 au niveau de la protubérance du CAK. La structure du core-TFIIH de levure a montré l'existence de 3 régions distinctes indiquées en rouge, bleu et vert (Figure 61) correspondant à Rad 3 (XPB), Ssl1 (p44) et Tfb (Tfb1/p62, Tfb2/p52 et Tfb3/p34), respectivement. L'ensemble de ces informations structurales, ainsi que les données d'interaction nous ont permis de proposer un modèle d'organisation architecturale de TFIIH (Figure 61B)

Néanmoins, ce modèle doit être confirmé par la localisation des sous-unités restantes (p34, p52, p62 et XPB) par microscopie électronique.

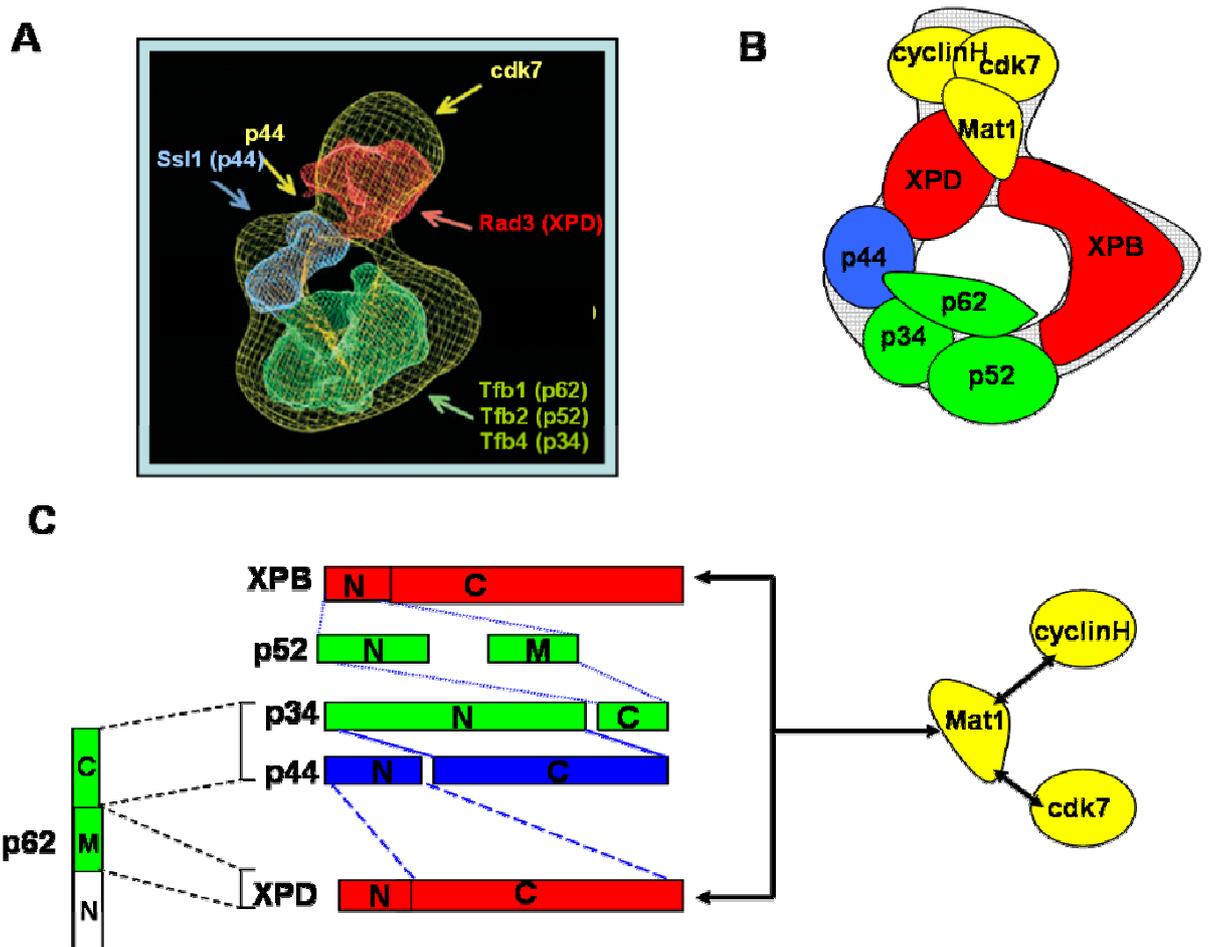


Figure 61 : Organisation moléculaire de TFIIH.

- A- Superposition des structures moléculaires de TFIIH humain et de levure (Chang and Kornberg, 2000; Schultz et al., 2000).
- B- Modèle de l'organisation des sous-unités de TFIIH, tenant compte de la structure moléculaire humaine et de levure ainsi que des données d'interactions.
- C- Représentation schématique des interactions entre les domaines des différentes sous-unités de TFIIH en se basant sur nos résultats et ceux qui sont publiés.

3- p52 régule la fonction de XPB au sein de TFIIH

TFIIH possède trois activités enzymatiques : deux activités hélicases et une activité kinase catalysées par XPB, XPD et cdk7, respectivement (Coin and Egly, 1998; Egly, 2001). L'activité de cdk7 dépend de la présence de la cycline H et de MAT1 d'une part et de l'état de phosphorylation de la boucle T d'autre part (Riedl and Egly, 2000). De la même manière, p44 interagit avec XPD et stimule son activité hélicase (Coin et al., 1998). En effet, il a été démontré que certaines mutations dans le domaine carboxy-terminal de XPD pouvaient être à l'origine d'une perte d'interaction avec p44, aboutissant à une

absence de stimulation de l'activité hélicase, ce qui explique le problème de réparation de l'ADN observé chez certains patients XP-D. Des mutations ponctuelles dans le domaine amino-terminal de p44 provoquent le même type de problème de réparation, car l'interaction XPD/p44 est affectée (Seroz et al., 2000). Chez certains malades XP-D, une déstabilisation du complexe TFIIH avec une diminution de l'ancrage du complexe CAK/XPD a été observée. Ceci se manifeste par une baisse de la phosphorylation et de la transactivation de certains récepteurs nucléaires (Keriel et al., 2002).

Quant à p52, son rôle au sein de TFIIH est jusque là assez mal connu. En effet, p52 a été la dernière sous-unité clonée de TFIIH. Des expériences de transcription et de réparation *in vitro* ont montré l'implication directe de cette sous unité dans la transcription et la réparation de l'ADN par excision et resynthèse de nucléotide (NER) (Marinoni et al., 1997). Chez la levure, l'homologue de p52 humain (63.9% de similarité et 39.7 % d'identité de séquence), Tfb2, est essentiel pour la viabilité de la cellule. Bien qu'une délétion de ces 60 derniers aminoacides ne soit pas létale, elle est cependant à l'origine d'un dysfonctionnement dans le système de NER et d'une sensibilité accrue aux rayonnements ultraviolets (Feaver et al., 1997).

Nous avons montré que p52 était le partenaire privilégié de l'hélicase XPB. Nous avons utilisé le système reconstitué de baculovirus (Tirode et al., 1999) pour reproduire un TFIIH recombinant délété de son domaine carboxy-terminal de p52. Ce complexe mutant a été testé en transcription et réparation *in vitro*. Il présente une activité de transcription et de réparation très réduite. Nous avons montré que cette mutation empêche l'ouverture correcte du promoteur, mais elle n'a aucun effet sur les autres activités enzymatiques de TFIIH (ATPase, hélicase et kinase). Nous avons mis en évidence que p52 entier conditionne l'ancrage de XPB au sein de TFIIH, expliquant ainsi les effets d'une telle délétion sur l'activité de transcription et de réparation. En effet, XPB est incapable d'interagir avec le complexe p34/p44/p62 indépendamment de p52. Nous avons montré que p52 et XPB interagissent directement, ce qui est en corrélation avec les données de double hybride qui impliquent Tfb2 dans une interaction avec Rad 25, les homologues respectifs de p52 et XPB (Feaver et al., 2000). De plus la labilité de XPB, mise en évidence pour ce mutant de p52 n'est pas spécifique à

l'homme, puisque son homologue de levure, Rad 25, est connu pour être aussi labile (Chang and Kornberg, 2000).

Deux domaines de p52 impliqués dans l'interaction ont été localisés (Figure 62) et correspondent à p52(1-135) (1^{er} domaine) et p52(304-381) (2^{ème} domaine). Le fait que le domaine p52(381-462) n'interagisse pas avec XPB pourrait être lié à son positionnement à la surface du complexe TFIIH, comme semble le suggérer nos expériences de protéolyse ménagée.

Le 1^{er} domaine est nécessaire mais pas suffisant pour un bon ancrage de XPB. Le 2^{ème} domaine se situe dans une région très conservée de l'homme à la levure et permet l'interaction et la stabilisation de XPB. Il doit probablement permettre à XPB d'adopter une bonne conformation d'abord au sein de TFIIH et peut-être au sein des complexes de transcription et de réparation. Une analyse précise des séquences récentes des homologues de p52 révèle une très grande homologie entre le premier et deuxième domaine de p52 (Dr Odile Lecompte et Dr Olivier Poch, communication personnelle), suggérant une duplication du domaine de reconnaissance de XPB.

Chez certains malades XP-B/TTD (*Xeroderma pigmentosum-B/ trichothiodystrophie*), une mutation dans les 40 derniers résidus de XPB est à l'origine du même phénotype que celui du mutant p52, c'est-à-dire un problème d'ouverture du promoteur, lié à un mauvais positionnement de XPB sur l'ADN (Coin et al., 1999; Douziech et al., 2000). Nos résultats et ceux obtenus à partir de ces malades XP-B/TTD soulignent l'importance de la conformation de XPB pour l'activité de TFIIH.

Des mutations des hélicases XPD et XPB ont été répertoriées chez les malades depuis des années. Cependant aucune mutation de p52 n'a été recensée à ce jour. On peut imaginer que ce type de mutation soit létal, ce qui rend impossible son recensement. Dans ce sens, faire une souris *Knock-out* du domaine C-terminal de p52 permettrait de voir l'effet d'une telle mutation *in vivo*.

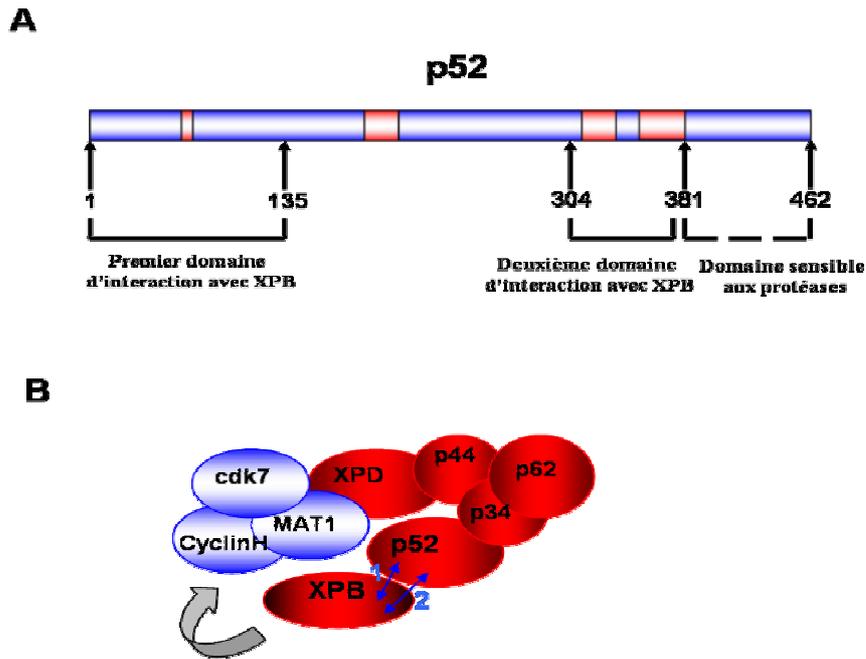


Figure 62 : Interaction entre p52 et XPB.

- A- Représentation schématique des domaines de p52. En rouge sont indiquées les régions très conservées entre l'homme et la levure.
- B- Représentation schématique de l'ancrage de XPB au complexe TFIID, via les deux points d'ancrage de p52.

Du point de vue moléculaire, nous avons montré qu'au sein du TFIID tronqué dans la partie C-terminale de p52, XPB est très labile. En effet, à haute concentration de NaCl, XPB est complètement absent en fin de purification alors qu'à basse concentration de NaCl, il semble être incorporé correctement. En revanche, le complexe n'est pas capable d'ouvrir le promoteur correctement.

De manière indirecte, nous avons préparé un TFIID qui ne comporte pas XPB ou en comporte dans une mauvaise conformation. Il serait alors intéressant de résoudre la structure de ce complexe dans ces deux conditions par microscopie électronique. La comparaison de ces deux structures avec celle du complexe sauvage devrait nous permettre de localiser XPB, dans le cas où l'absence de XPB n'affecte pas la structure de TFIID. De plus, cette comparaison permettrait de visualiser la bonne et mauvaise conformation de XPB au sein de TFIID. La même approche a été utilisée pour caractériser un mutant de XPD dont l'interaction p44/XPD était altérée. Après analyse des images en microscopie électronique, deux formes majoritaires ont été observées, une forme ouverte et une forme fermée, correspondant au TFIID mutant et sauvage, respectivement (Dr Patrick Schultz, communication personnelle) (Figure 63).

Cette forme ouverte serait une forme inactive engendrée par l'absence d'interaction entre p44 et XPD et pourrait être une explication moléculaire de la maladie.

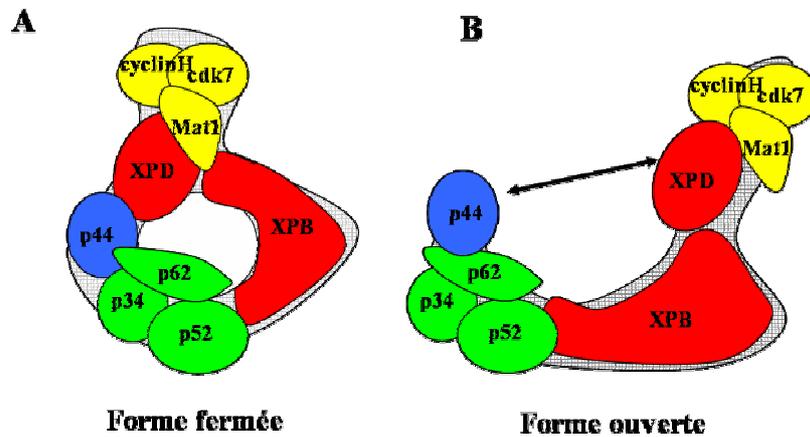


Figure 63 : Représentation schématique de l'architecture de TFIIH sauvage et muté en XPD.

A- Forme fermée de TFIIH (sauvage).

B- Forme ouverte de TFIIH (mutant XPD affectant l'interaction entre p44 et XPD).

Nous avons montré que p52 et XPB forment un complexe stable à 250 mM NaCl. Les domaines impliqués dans cette interaction ont été localisés. Cristalliser et résoudre la structure cristallographique de ce complexe binaire stable, nous permettra de mieux comprendre comment les domaines de p52 interagissent avec le domaine N-terminal de XPB, mettant en évidence de manière générale le mode de reconnaissance entre l'hélicase et sa protéine régulatrice.

Une analyse fine des séquences d'un agent infectieux *Gardia. intestinalis*, montre de manière surprenante que seule XPB et le domaine C-terminal de p52 existe chez cet organisme (Dr odile Lecomte et Dr olivier Poch, communication personnelle). Ce résultat suggère que *Gardia* a conservé la sous-unité clé de TFIIH, XPB et son régulateur p52, absolument indispensable pour la transcription. Nous postulons que chez cet organisme, la transcription est plus importante que le NER, puisqu'elle n'a pas conservé les facteurs de réparation XPG et XPC. En effet, nous pensons que le NER est un mécanisme très spécifique, dont tous les organismes n'ont pas exclusivement besoin, ce qui n'est pas le cas du mécanisme universel et indispensable qu'est la transcription.

4-Etude structure-fonction de p62

Une étude de séquence de p62 suggère que cette sous-unité est organisée en trois domaines : un domaine amino-terminal (p62N), un domaine médian (p62M) et un domaine carboxy-terminal (p62C) (Figure 64). Nos données d'interaction impliquent p62C dans une interaction avec le complexe p34/p44, suggérant un rôle architectural pour ce domaine. P62M comporte la région la plus conservée correspondant à un motif FW (Phe,Trp) nommé BSD (*BTF2-like transcription factors, synapse-associated and Dos2-like protein*) (Doerks et al., 2002). Nous avons montré que ce domaine interagit avec XPD au sein de TFIIH, probablement via ce motif. La structure du domaine minimal de p62M de 8 kDa définie par protéolyse ménagée et spectrométrie de masse, est en cours de résolution par RMN et devrait apporter des réponses concernant le mode de reconnaissance de XPD par p62.

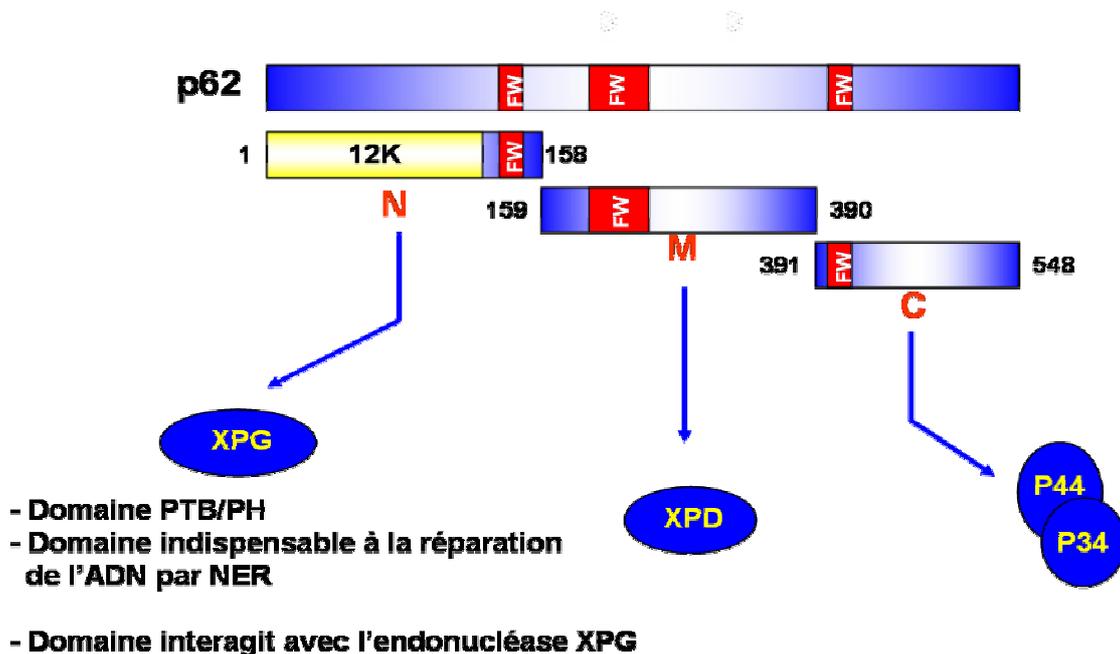


Figure 64 : Représentation schématique des domaines de p62.

Les carrés rouges représentent les motifs FW de p62. Le domaine N-terminal de 12 kDa est représenté en jaune.

XPG, XPD et le complexe p34/p44 interagissent avec le domaine N-terminal (1-158), médian (159-390) et C-terminal (391-538), respectivement.

La structure du domaine N-terminal a été résolue par résonance magnétique nucléaire et a révélée la présence de sept feuillet- β « sandwich » suivis d'une hélice carboxy-terminale. Ce repliement est caractéristique des domaines PH (*Pleckstrin Homology*) et PTB (*PhosphoTyrosine Binding*). Une délétion de ce domaine amino-terminal de p62(1-109) n'affecte pas l'intégrité de TFIIF et n'a aucun effet sur son activité de transcription. Cependant, elle provoque une diminution très importante de son activité de réparation par excision re-synthèse de nucléotide (NER). Une interaction directe de ce domaine avec l'endonucléase XPG a été établie et pourrait être à l'origine du dysfonctionnement de la réparation par NER.

Le domaine PH/PTB de p62 est aussi capable d'interagir de manière spécifique avec certains lipides comme le phosphatidyl-sérine, le phosphatidyl-inositol et le phosphatidyl-5P-inositol. La signification biologique de cette interaction reste à définir. Un modèle de fonctionnement de l'ARN polymérase II *in vivo* a été proposé (Cook, 1999). Il suggère qu'elle est immobilisée au niveau d'une structure nucléaire plus large qui constitue une unité de transcription locale appelée « *transcription factorie* », sur laquelle l'ADN vient se glisser pour être transcrit. Contrairement au modèle de l'assemblage séquentiel qui propose un assemblage et une dissociation des facteurs à chaque cycle de transcription, ce modèle constitue une économie pour la cellule puisqu'une fois les facteurs de transcription assemblés, ils seront utilisés pour la transcription. Dans ce sens, on peut imaginer que le domaine PH/PTB de p62 puisse servir à TFIIF pour se fixer au niveau de la matrice nucléoprotéique en association avec les autres facteurs de transcription.

Par ailleurs, un certain nombre de travaux désignent p62 comme la cible de plusieurs protéines partenaires de TFIIF. En effet, dans un contexte de réparation d'ADN, p62 participe à la reconnaissance de XPC/HR23B (Yokoi et al., 2000) et de l'endonucléase XPG.

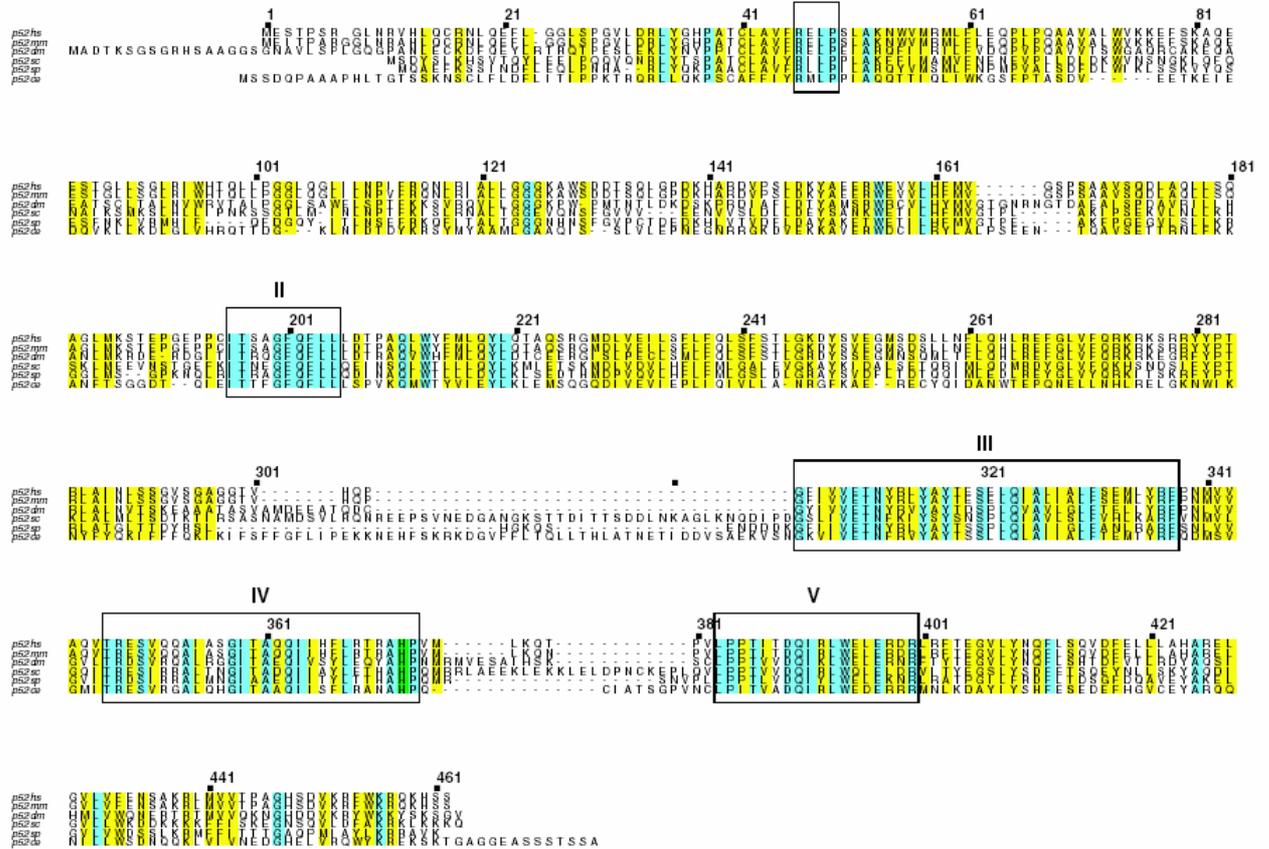
Dans un contexte de transcription, l'homologue de levure de p62 (Tfb1) interagit avec la sous-unité α du facteur de transcription TFIIE (Uetz et al., 2000). p62 est aussi la cible d'un ensemble de régulateurs de transcription, tels que VP16, p53, E2F-1, HBx, et EBNA 2 (Leveillard et al., 1996; Tong et al., 1995; Truant et al., 1995; Vandel and Kouzarides, 1999; Xiao et al., 1994). Le récepteur humain à l'œstrogène ER α interagit avec TFIIF, via

p62 (Chen et al., 2000). Nos données suggèrent que le domaine médian et carboxy-terminal ont essentiellement un rôle architectural et que le domaine amino-terminal serait plutôt impliqué dans des interactions protéine-protéine. Il a été d'ailleurs proposé que ER α , interagit avec TFIIH via p62 et en particulier grâce à une hélice LLQQLL, faisant intervenir un mécanisme de reconnaissance du type récepteur-coactivateur (Chen et al., 2000). La structure de p62(1-109) permet de localiser cette hélice, qui est plaquée contre le « sandwich » de feuillets- β et semble inaccessible pour l'interaction avec ER α . Des expériences de relaxation en RMN ont montré la rigidité de cette hélice, mettant en doute le mécanisme de reconnaissance proposé. p62(1-109) est une protéine très basique, et une grande partie de ces partenaires possèdent un domaine acide. Il est alors possible que l'interaction soit de nature polaire. Dans ce cas, il est très difficile de parler de spécificité d'interaction et ainsi de rôle de régulateur de p62 au sein de TFIIH.

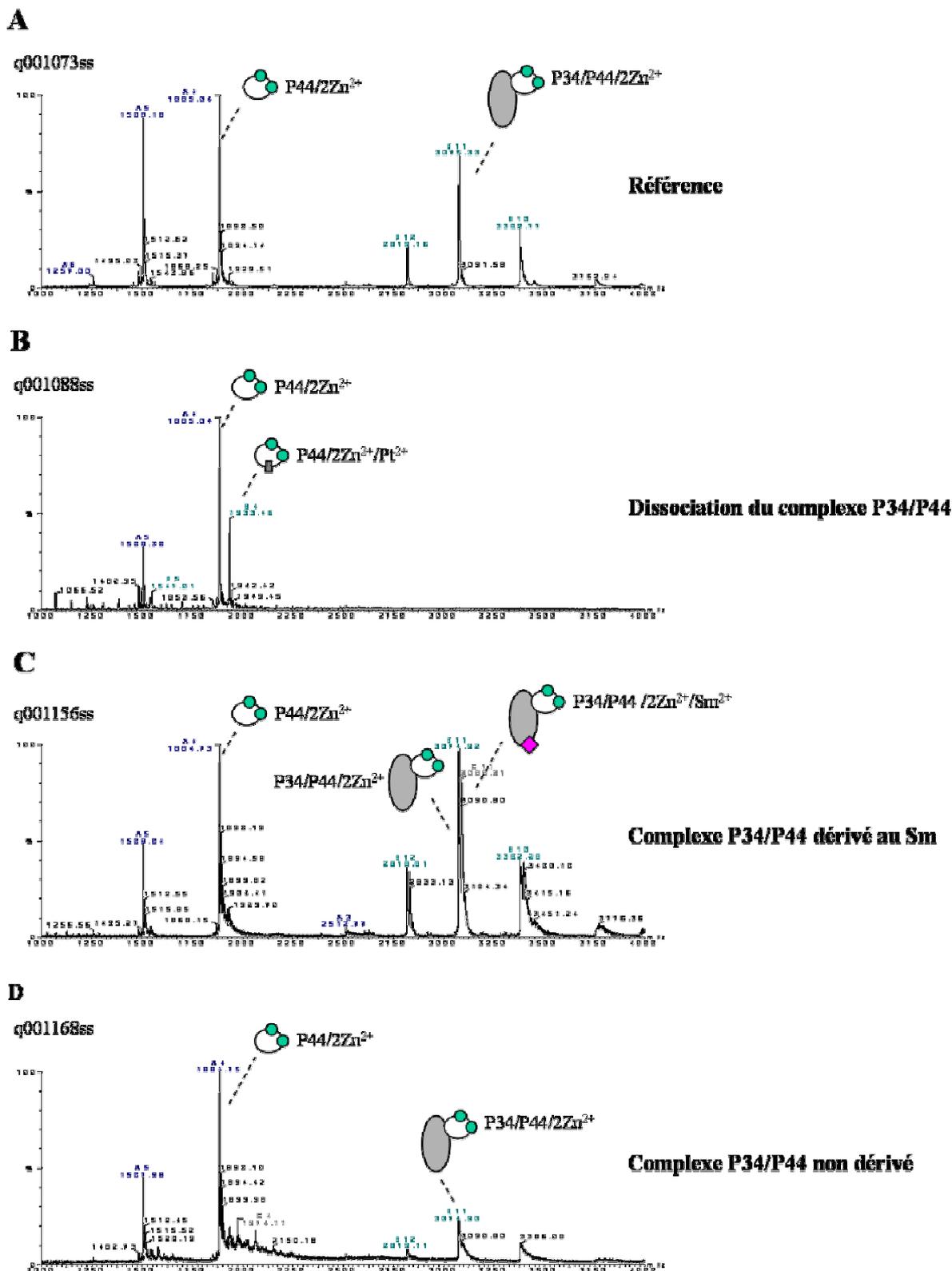
5-Etude du facteur de TFIIE

Dans le but d'étudier la structure du facteur de transcription de base TFIIE, nous l'avons surproduit dans *E.coli*, purifié et étudié son comportement en solution. Par gel natif, nous avons vérifié que ce complexe existe en une seule espèce. Des expériences de pontage chimique et d'ultracentrifugation analytique montrent que TFIIE existe en solution sous forme de dimère $\alpha\beta$, contrairement à ce qui est généralement admis dans la littérature. En effet, le groupe de RG. Roeder avait isolé et caractérisé TFIIE par gel filtration. En se basant sur l'élution de TFIIE en gel filtration, ce groupe a proposé que TFIIE est un tétramère $\alpha_2\beta_2$ (Ohkuma et al., 1990). Une étude en microscopie électronique de cristaux 2D du complexe ARN pol II-TFIIE de levure a suggéré que TFIIE peut exister sous forme d'un dimère $\alpha\beta$. Nous avons résolu la structure de TFIIE par microscopie électronique et analyse d'images à 17Å. Elle a révélé une organisation de TFIIE en domaines distincts. TFIIE est connu pour interagir avec TFIIH et stimuler sa capacité à phosphoryler le CTD de l'ARN polymérase II. Pour mieux comprendre les bases moléculaires de cette interaction, nous avons commencé par localiser les zones de contact des deux complexes. p62 semble suffisant pour recruter TFIIE, et le domaine C-terminal de TFIIE α est impliqué dans cette interaction. Ce même domaine C-terminal est sensible aux protéases et pourrait être stabilisé par l'interaction avec TFIIH. Une étude de localisation des extrémités de TFIIE α est en cours. Elle permettra de cartographier TFIIE et assigner ces domaines distincts. Faire la structure de TFIIE/TFIIH par microscopie électronique nécessite d'établir les conditions d'association des deux complexes. Pour ce faire, une caractérisation précise de cette interaction pourrait être effectuée par résonance plasmonique de surface (*Biacore*) notamment. Il n'est pas impossible que l'interaction des deux stabilise le complexe TFIIE/TFIIH, et permette d'augmenter la résolution de TFIIH. Cette structure contribuera sans doute à une meilleure compréhension de l'assemblage de ces deux facteurs. Ceci constituera les bases moléculaires de la dernière étape de l'initiation de la transcription. De la même manière, on peut imaginer former des complexes plus larges notamment ADN/TFIIE/TFIIH ou encore ADN/ARN pol II/TFIIE/TFIIH permettant de proposer un modèle précis de l'initiation de la transcription des gènes codant pour les protéines.

p52

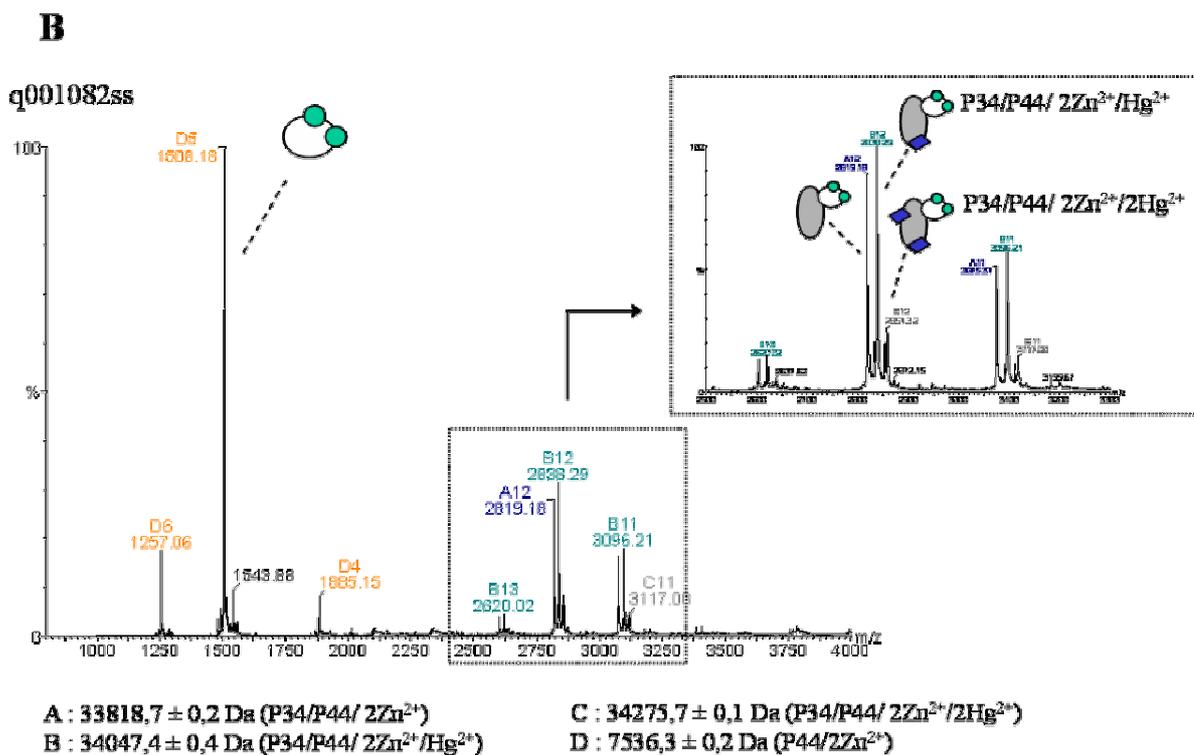
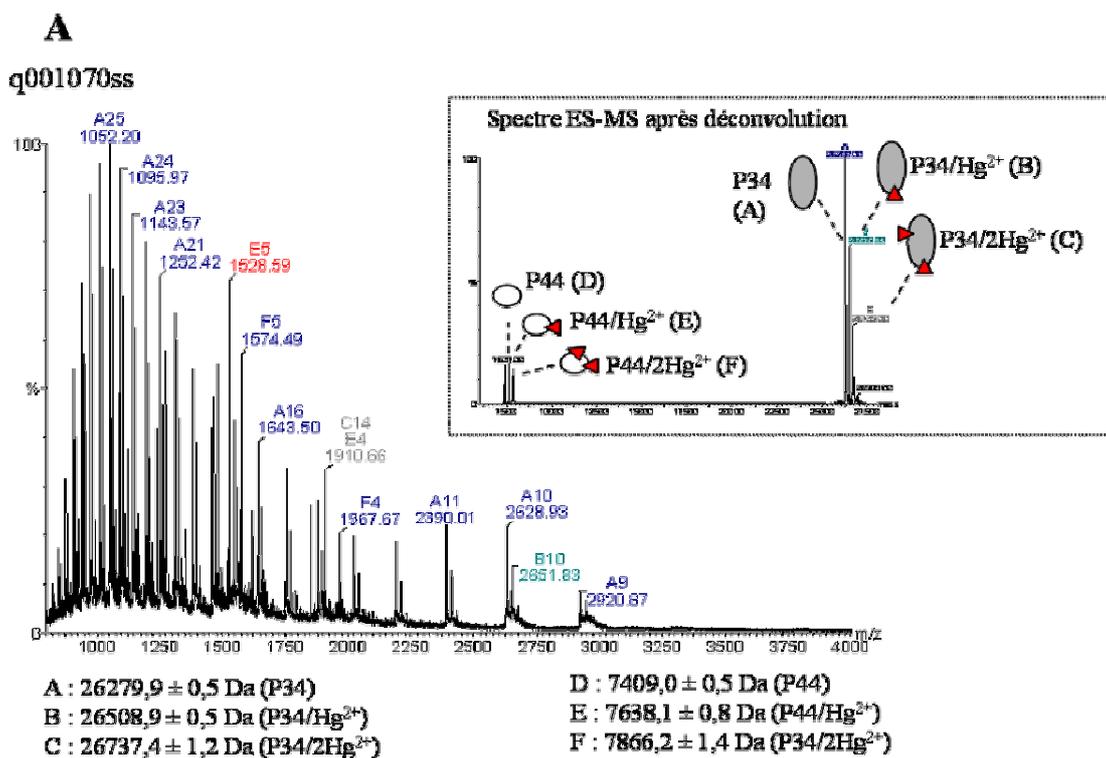


Annexe 2 : Alignement de séquences de la sous-unité p52.



Annexe 4 : Etude de la dérivation du complexe p34N/p44C par spectrométrie de masse en conditions non dénaturantes.

- A- Spectre ES-MS du complexe p34N/p44C seul.
- B- Spectre ES-MS du complexe p34N/p44C avec 10 équivalents en K_2PtCl_4 .
- C- Spectre ES-MS du complexe p34N/p44C avec 10 équivalents en $SmCl_3$
- D- Spectre ES-MS du complexe p34N/p44C avec 10 équivalents en $HgBr_2$

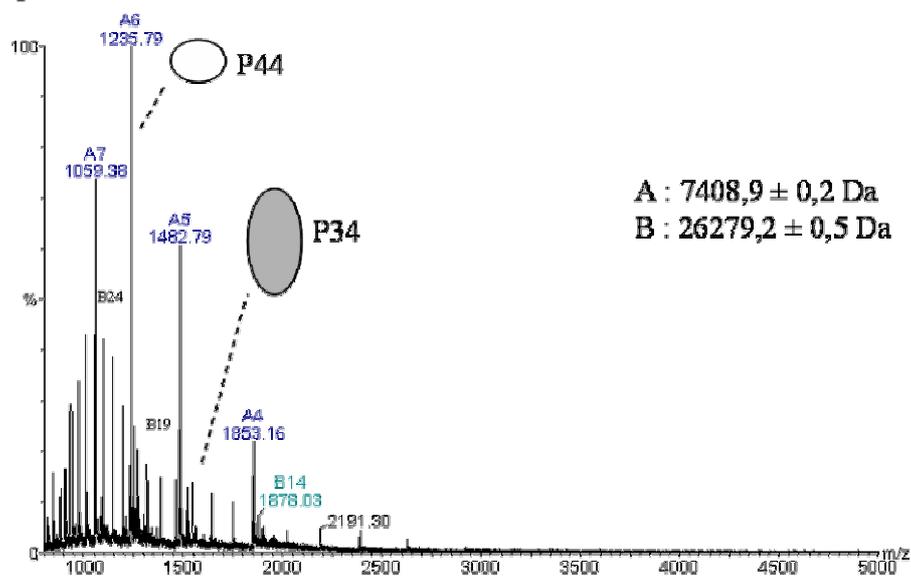


Annexe 5 : Dérivation covalente du complexe p34N/p44C par le mercure.

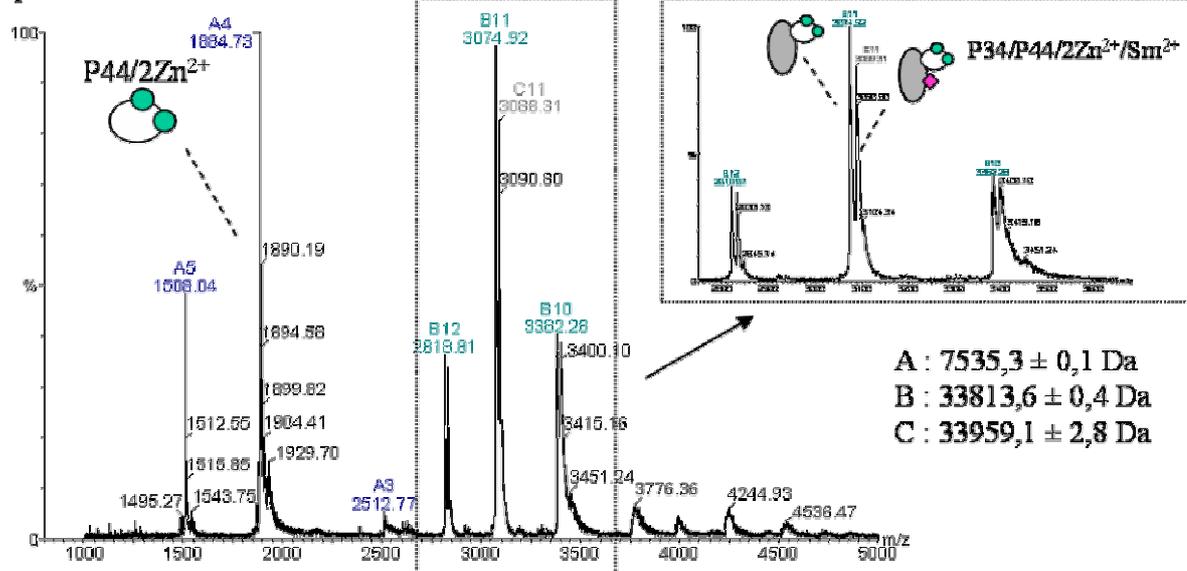
- A- Spectre ES-MS du complexe p34N/p44C avec 10 équivalents en thiomersal en conditions dénaturantes
- B- Spectre ES-MS du complexe p34N/p44C avec 10 équivalents en thiomersal en conditions non dénaturantes.

A

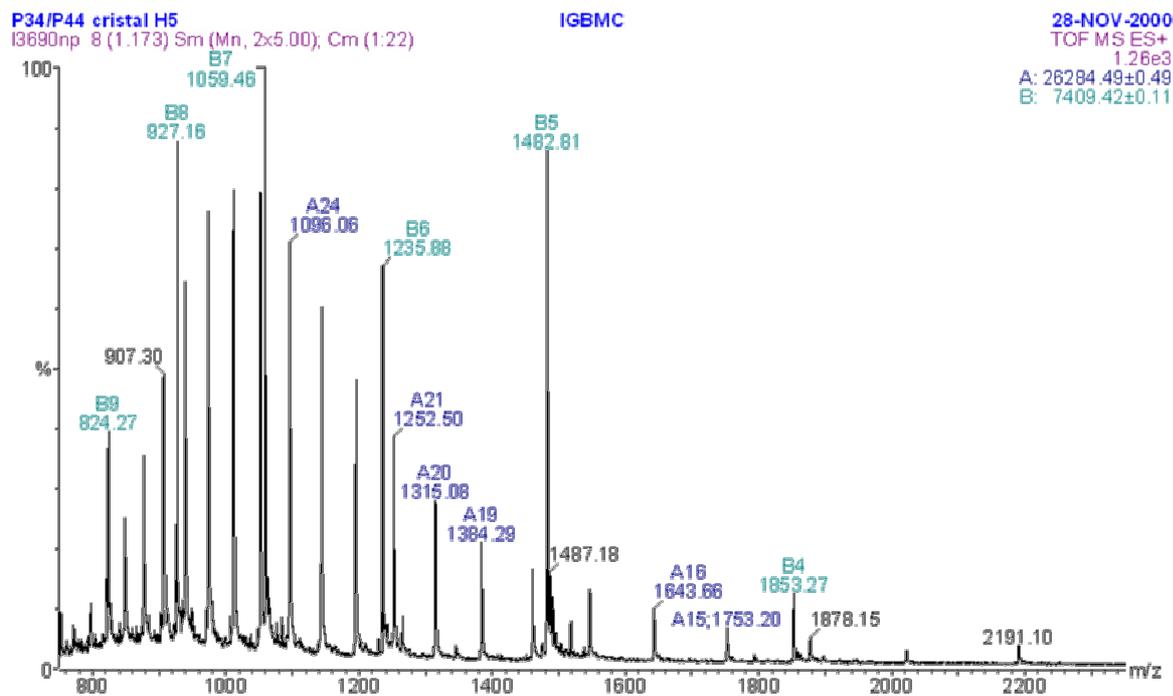
q001179ss

**B**

q001156ss

**Annexe 6 : Dérivation non-covalente du complexe p34N/p44C par le samarium.**

- C- Spectre ES-MS du complexe p34N/p44C avec 10 équivalents en $\text{Sm}(\text{SO}_4)_3$ en conditions dénaturantes
- D- Spectre ES-MS du complexe p34N/p44C avec 10 équivalents en $\text{Sm}(\text{SO}_4)_3$ en conditions non dénaturantes.



Annexe 7 : Spectre ES-MS des cristaux de p34N/p44C.

Les cristaux ont été lavés puis fondus dans une solution d'acétate d'ammonium à 50 mM.

- Aboussekhra, A., Biggerstaff, M., Shivji, M. K., Vilpo, J. A., Moncollin, V., Podust, V. N., Protic, M., Hubscher, U., Egly, J. M., and Wood, R. D. (1995). Mammalian DNA nucleotide excision repair reconstituted with purified protein components. *Cell* 80, 859-868.
- Agarwal, M. L., Taylor, W. R., Chernov, M. V., Chernova, O. B., and Stark, G. R. (1998). The p53 network. *J Biol Chem* 273, 1-4.
- Allfrey, V. G., Pogo, B. G., Littau, V. C., Gershey, E. L., and Mirsky, A. E. (1968). Histone acetylation in insect chromosomes. *Science* 159, 314-316.
- Allison, L. A., Wong, J. K., Fitzpatrick, V. D., Moyle, M., and Ingles, C. J. (1988). The C-terminal domain of the largest subunit of RNA polymerase II of *Saccharomyces cerevisiae*, *Drosophila melanogaster*, and mammals: a conserved structure with an essential function. *Mol Cell Biol* 8, 321-329.
- 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.
- Andersen, G., Busso, D., Poterszman, A., Hwang, J. R., Wurtz, J. M., Ripp, R., Thierry, J. C., Egly, J. M., and Moras, D. (1997). The structure of cyclin H: common mode of kinase activation and specific features. *Embo J* 16, 958-967.
- Andersen, G., Poterszman, A., Egly, J. M., Moras, D., and Thierry, J. C. (1996). The crystal structure of human cyclin H. *FEBS Lett* 397, 65-69.
- Araujo, S. J., Tirode, F., Coin, F., Pospiech, H., Syvaioja, J. E., Stucki, M., Hubscher, U., Egly, J. M., and Wood, R. D. (2000). Nucleotide excision repair of DNA with recombinant human proteins: definition of the minimal set of factors, active forms of TFIIH, and modulation by CAK. *Genes Dev* 14, 349-359.
- Aravind, L., and Ponting, C. P. (1998). Homologues of 26S proteasome subunits are regulators of transcription and translation. *Protein Sci* 7, 1250-1254.
- Arents, G., Burlingame, R. W., Wang, B. C., Love, W. E., and Moudrianakis, E. N. (1991). The nucleosomal core histone octamer at 3.1 Å resolution: a tripartite protein assembly and a left-handed superhelix. *Proc Natl Acad Sci U S A* 88, 10148-10152.
- Armache, K. J., Kettenberger, H., and Cramer, P. (2003). Architecture of initiation-competent 12-subunit RNA polymerase II. *Proc Natl Acad Sci U S A* 100, 6964-6968.
- Aso, T., Conaway, J. W., and Conaway, R. C. (1995). The RNA polymerase II elongation complex. *Faseb J* 9, 1419-1428.
- Awrey, D. E., Weilbaeher, R. G., Hemming, S. A., Orlicky, S. M., Kane, C. M., and Edwards, A. M. (1997). Transcription elongation through DNA arrest sites. A multistep process involving both RNA polymerase II subunit RPB9 and TFIIIS. *J Biol Chem* 272, 14747-14754.

- Bartolomei, M. S., Halden, N. F., Cullen, C. R., and Corden, J. L. (1988). Genetic analysis of the repetitive carboxyl-terminal domain of the largest subunit of mouse RNA polymerase II. *Mol Cell Biol* 8, 330-339.
- Bastien, J., Adam-Stitah, S., Riedl, T., Egly, J. M., Chambon, P., and 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, 21896-21904.
- Berger, S. L. (2002). Histone modifications in transcriptional regulation. *Curr Opin Genet Dev* 12, 142-148.
- Bergmann, E., and Egly, J. M. (2001). Trichothiodystrophy, a transcription syndrome. *Trends Genet* 17, 279-286.
- Bienstock, R. J., Skorvaga, M., Mandavilli, B. S., and Van Houten, B. (2003). Structural and functional characterization of the human DNA repair helicase XPD by comparative molecular modeling and site-directed mutagenesis of the bacterial repair protein UvrB. *J Biol Chem* 278, 5309-5316.
- Birck, C., Poch, O., Romier, C., Ruff, M., Mengus, G., Lavigne, A. C., Davidson, I., and 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.
- Bootsma, D., Weeda, G., Vermeulen, W., van Vuuren, H., Troelstra, C., van der Spek, P., and Hoeijmakers, J. (1995). Nucleotide excision repair syndromes: molecular basis and clinical symptoms. *Philos Trans R Soc Lond B Biol Sci* 347, 75-81.
- Bradsher, J., Auriol, J., Proietti de Santis, L., Iben, S., Vonesch, J. L., Grummt, I., and Egly, J. M. (2002). CSB is a component of RNA pol I transcription. *Mol Cell* 10, 819-829.
- Bradsher, J., Coin, F., and Egly, J. M. (2000). Distinct roles for the helicases of TFIIF in transcript initiation and promoter escape. *J Biol Chem* 275, 2532-2538.
- Bradsher, J. N., Jackson, K. W., Conaway, R. C., and Conaway, J. W. (1993). RNA polymerase II transcription factor SIII. I. Identification, purification, and properties. *J Biol Chem* 268, 25587-25593.
- Brand, M., Leurent, C., Mallouh, V., Tora, L., and Schultz, P. (1999). Three-dimensional structures of the TAFII-containing complexes TFIID and TIFC. *Science* 286, 2151-2153.
- Broughton, B. C., Thompson, A. F., Harcourt, S. A., Vermeulen, W., Hoeijmakers, J. H., Botta, E., Stefanini, M., King, M. D., Weber, C. A., Cole, J., and et al. (1995). Molecular and cellular analysis of the DNA repair defect in a patient in xeroderma pigmentosum complementation group D who has the clinical features of xeroderma pigmentosum and Cockayne syndrome. *Am J Hum Genet* 56, 167-174.
- Brown, N. R., Noble, M. E., Endicott, J. A., Garman, E. F., Wakatsuki, S., Mitchell, E., Rasmussen, B., Hunt, T., and Johnson, L. N. (1995). The crystal structure of cyclin A. *Structure* 3, 1235-1247.

- Bryant, G. O., Martel, L. S., Burley, S. K., and Berk, A. J. (1996). Radical mutations reveal TATA-box binding protein surfaces required for activated transcription in vivo. *Genes Dev* *10*, 2491-2504.
- Burke, T. W., and Kadonaga, J. T. (1997). The downstream core promoter element, DPE, is conserved from *Drosophila* to humans and is recognized by TAFII60 of *Drosophila*. *Genes Dev* *11*, 3020-3031.
- Burley, S. K., and Roeder, R. G. (1996). Biochemistry and structural biology of transcription factor IID (TFIID). *Annu Rev Biochem* *65*, 769-799.
- Bushnell, D. A., and Kornberg, R. D. (2003). Complete, 12-subunit RNA polymerase II at 4.1-Å resolution: implications for the initiation of transcription. *Proc Natl Acad Sci U S A* *100*, 6969-6973.
- Busso, D., Keriel, A., Sandrock, B., Poterszman, A., Gileadi, O., and Egly, J. M. (2000). Distinct regions of MAT1 regulate cdk7 kinase and TFIIF transcription activities. *J Biol Chem* *275*, 22815-22823.
- Bustin, M., and Reeves, R. (1996). High-mobility-group chromosomal proteins: architectural components that facilitate chromatin function. *Prog Nucleic Acid Res Mol Biol* *54*, 35-100.
- Cadena, D. L., and Dahmus, M. E. (1987). Messenger RNA synthesis in mammalian cells is catalyzed by the phosphorylated form of RNA polymerase II. *J Biol Chem* *262*, 12468-12474.
- Castro, J., Merino, C., and Zurita, M. (2002). Molecular characterization and developmental expression of the TFIIF factor p62 gene from *Drosophila melanogaster*: effects on the UV light sensitivity of a p62 mutant fly. *DNA Repair (Amst)* *1*, 359-368.
- Chambers, R. S., and Dahmus, M. E. (1994). Purification and characterization of a phosphatase from HeLa cells which dephosphorylates the C-terminal domain of RNA polymerase II. *J Biol Chem* *269*, 26243-26248.
- Chambers, R. S., and Kane, C. M. (1996). Purification and characterization of an RNA polymerase II phosphatase from yeast. *J Biol Chem* *271*, 24498-24504.
- Chang, W. H., and Kornberg, R. D. (2000). Electron crystal structure of the transcription factor and DNA repair complex, core TFIIF. *Cell* *102*, 609-613.
- Chao, D. M., Gadbois, E. L., Murray, P. J., Anderson, S. F., Sonu, M. S., Parvin, J. D., and Young, R. A. (1996). A mammalian SRB protein associated with an RNA polymerase II holoenzyme. *Nature* *380*, 82-85.
- Chasman, D. I., Flaherty, K. M., Sharp, P. A., and Kornberg, R. D. (1993). Crystal structure of yeast TATA-binding protein and model for interaction with DNA. *Proc Natl Acad Sci U S A* *90*, 8174-8178.
- Chen, D., Riedl, T., Washbrook, E., Pace, P. E., Coombes, R. C., Egly, J. M., and Ali, S. (2000). Activation of estrogen receptor alpha by S118 phosphorylation involves a ligand-dependent interaction with TFIIF and participation of CDK7. *Mol Cell* *6*, 127-137.

- Chen, J., Larochelle, S., Li, X., and Suter, B. (2003). Xpd/Ercc2 regulates CAK activity and mitotic progression. *Nature* 424, 228-232.
- Chesnut, J. D., Stephens, J. H., and Dahmus, M. E. (1992). The interaction of RNA polymerase II with the adenovirus-2 major late promoter is precluded by phosphorylation of the C-terminal domain of subunit IIa. *J Biol Chem* 267, 10500-10506.
- Cho, H., Kim, T. K., Mancebo, H., Lane, W. S., Flores, O., and Reinberg, D. (1999). A protein phosphatase functions to recycle RNA polymerase II. *Genes Dev* 13, 1540-1552.
- Chu, G., and Mayne, L. (1996). Xeroderma pigmentosum, Cockayne syndrome and trichothiodystrophy: do the genes explain the diseases? *Trends Genet* 12, 187-192.
- Coin, F., Bergmann, E., Tremeau-Bravard, A., and Egly, J. M. (1999). Mutations in XPB and XPD helicases found in xeroderma pigmentosum patients impair the transcription function of TFIIH. *Embo J* 18, 1357-1366.
- Coin, F., and Egly, J. M. (1998). Ten years of TFIIH. *Cold Spring Harb Symp Quant Biol* 63, 105-110.
- Coin, F., Marinoni, J. C., Rodolfo, C., Fribourg, S., Pedrini, A. M., and Egly, J. M. (1998). Mutations in the XPD helicase gene result in XP and TTD phenotypes, preventing interaction between XPD and the p44 subunit of TFIIH. *Nat Genet* 20, 184-188.
- Conaway, J. W., Shilatifard, A., Dvir, A., and Conaway, R. C. (2000). Control of elongation by RNA polymerase II. *Trends Biochem Sci* 25, 375-380.
- Conaway, R. C., Kong, S. E., and Conaway, J. W. (2003). TFIIIS and GreB: two like-minded transcription elongation factors with sticky fingers. *Cell* 114, 272-274.
- Cook, P. R. (1999). The organization of replication and transcription. *Science* 284, 1790-1795.
- Corden, J. L., Cadena, D. L., Ahearn, J. M., Jr., and Dahmus, M. E. (1985). A unique structure at the carboxyl terminus of the largest subunit of eukaryotic RNA polymerase II. *Proc Natl Acad Sci U S A* 82, 7934-7938.
- Cote, J., Quinn, J., Workman, J. L., and Peterson, C. L. (1994). Stimulation of GAL4 derivative binding to nucleosomal DNA by the yeast SWI/SNF complex. *Science* 265, 53-60.
- Craighead, J. L., Chang, W. H., and Asturias, F. J. (2002). Structure of yeast RNA polymerase II in solution: implications for enzyme regulation and interaction with promoter DNA. *Structure (Camb)* 10, 1117-1125.
- Cramer, P., Bushnell, D. A., and Kornberg, R. D. (2001). Structural basis of transcription: RNA polymerase II at 2.8 angstrom resolution. *Science* 292, 1863-1876.
- Dahmus, M. E. (1981). Phosphorylation of eukaryotic DNA-dependent RNA polymerase. Identification of calf thymus RNA polymerase subunits phosphorylated by two purified protein kinases, correlation with in vivo sites of phosphorylation in HeLa cell RNA polymerase II. *J Biol Chem* 256, 3332-3339.

- Datta, A., Bagchi, S., Nag, A., Shiyonov, P., Adami, G. R., Yoon, T., and Raychaudhuri, P. (2001). The p48 subunit of the damaged-DNA binding protein DDB associates with the CBP/p300 family of histone acetyltransferase. *Mutat Res* 486, 89-97.
- Davis, J. A., Takagi, Y., Kornberg, R. D., and Asturias, F. A. (2002). Structure of the yeast RNA polymerase II holoenzyme: Mediator conformation and polymerase interaction. *Mol Cell* 10, 409-415.
- de Boer, J., Donker, I., de Wit, J., Hoeijmakers, J. H., and Weeda, G. (1998). Disruption of the mouse xeroderma pigmentosum group D DNA repair/basal transcription gene results in preimplantation lethality. *Cancer Res* 58, 89-94.
- de Laat, W. L., Appeldoorn, E., Sugasawa, K., Weterings, E., Jaspers, N. G., and Hoeijmakers, J. H. (1998). DNA-binding polarity of human replication protein A positions nucleases in nucleotide excision repair. *Genes Dev* 12, 2598-2609.
- de Laat, W. L., Jaspers, N. G., and Hoeijmakers, J. H. (1999). Molecular mechanism of nucleotide excision repair. *Genes Dev* 13, 768-785.
- DeJong, J., and Roeder, R. G. (1993). A single cDNA, hTFIIA/alpha, encodes both the p35 and p19 subunits of human TFIIA. *Genes Dev* 7, 2220-2234.
- Della Coletta, L., Rolig, R. L., Fossey, S., Morizot, D. C., Nairn, R. S., and Walter, R. B. (1995). Characterization of the Xiphophorus fish (Teleostei: Poeciliidae) ERCC2/XPD locus. *Genomics* 26, 70-76.
- Deroo, B. J., and Archer, T. K. (2001). Glucocorticoid receptor-mediated chromatin remodeling in vivo. *Oncogene* 20, 3039-3046.
- Doerks, T., Huber, S., Buchner, E., and Bork, P. (2002). BSD: a novel domain in transcription factors and synapse-associated proteins. *Trends Biochem Sci* 27, 168-170.
- Dotson, M. R., Yuan, C. X., Roeder, R. G., Myers, L. C., Gustafsson, C. M., Jiang, Y. W., Li, Y., Kornberg, R. D., and Asturias, F. J. (2000). Structural organization of yeast and mammalian mediator complexes. *Proc Natl Acad Sci U S A* 97, 14307-14310.
- Douziech, M., Coin, F., Chipoulet, J. M., Arai, Y., Ohkuma, Y., Egly, J. M., and Coulombe, B. (2000). Mechanism of promoter melting by the xeroderma pigmentosum complementation group B helicase of transcription factor IIIH revealed by protein-DNA photo-cross-linking. *Mol Cell Biol* 20, 8168-8177.
- Drapkin, R., Le Roy, G., Cho, H., Akoulitchev, S., and Reinberg, D. (1996). Human cyclin-dependent kinase-activating kinase exists in three distinct complexes. *Proc Natl Acad Sci U S A* 93, 6488-6493.
- Drapkin, R., Reardon, J. T., Ansari, A., Huang, J. C., Zawel, L., Ahn, K., Sancar, A., and Reinberg, D. (1994). Dual role of TFIIH in DNA excision repair and in transcription by RNA polymerase II. *Nature* 368, 769-772.

Dubaele, S., Proietti De Santis, L., Bienstock, R. J., Keriél, A., Stefanini, M., Van Houten, B., and Egly, J. M. (2003). Basal transcription defect discriminates between xeroderma pigmentosum and trichothiodystrophy in XPD patients. *Mol Cell* 11, 1635-1646.

Dvir, A., Garrett, K. P., Chalut, C., Egly, J. M., Conaway, J. W., and Conaway, R. C. (1996). A role for ATP and TFIIF in activation of the RNA polymerase II preinitiation complex prior to transcription initiation. *J Biol Chem* 271, 7245-7248.

Egly, J. M. (2001). The 14th Datta Lecture. TFIIF: from transcription to clinic. *FEBS Lett* 498, 124-128.

Evans, E., Moggs, J. G., Hwang, J. R., Egly, J. M., and Wood, R. D. (1997). Mechanism of open complex and dual incision formation by human nucleotide excision repair factors. *Embo J* 16, 6559-6573.

Feaver, W. J., Henry, N. L., Wang, Z., Wu, X., Svejstrup, J. Q., Bushnell, D. A., Friedberg, E. C., and Kornberg, R. D. (1997). Genes for Tfb2, Tfb3, and Tfb4 subunits of yeast transcription/repair factor IIF. Homology to human cyclin-dependent kinase activating kinase and IIF subunits. *J Biol Chem* 272, 19319-19327.

Feaver, W. J., Huang, W., and Friedberg, E. C. (1999). The TFB4 subunit of yeast TFIIF is required for both nucleotide excision repair and RNA polymerase II transcription. *J Biol Chem* 274, 29564-29567.

Feaver, W. J., Huang, W., Gileadi, O., Myers, L., Gustafsson, C. M., Kornberg, R. D., and Friedberg, E. C. (2000). Subunit interactions in yeast transcription/repair factor TFIIF. Requirement for Tfb3 subunit in nucleotide excision repair. *J Biol Chem* 275, 5941-5946.

Fischer, L., Gerard, M., Chalut, C., Lutz, Y., Humbert, S., Kanno, M., Chambon, P., and Egly, J. M. (1992). Cloning of the 62-kilodalton component of basic transcription factor BTF2. *Science* 257, 1392-1395.

Fisher, R. P., and Morgan, D. O. (1994). A novel cyclin associates with MO15/CDK7 to form the CDK-activating kinase. *Cell* 78, 713-724.

Flanagan, P. M., Kelleher, R. J., 3rd, Sayre, M. H., Tschochner, H., and Kornberg, R. D. (1991). A mediator required for activation of RNA polymerase II transcription in vitro. *Nature* 350, 436-438.

Flores, O., Lu, H., and Reinberg, D. (1992). Factors involved in specific transcription by mammalian RNA polymerase II. Identification and characterization of factor IIF. *J Biol Chem* 267, 2786-2793.

Flores, O., Maldonado, E., Burton, Z., Greenblatt, J., and Reinberg, D. (1988). Factors involved in specific transcription by mammalian RNA polymerase II. RNA polymerase II-associating protein 30 is an essential component of transcription factor IIF. *J Biol Chem* 263, 10812-10816.

- Flores, O., Maldonado, E., and Reinberg, D. (1989). Factors involved in specific transcription by mammalian RNA polymerase II. Factors IIE and IIF independently interact with RNA polymerase II. *J Biol Chem* 264, 8913-8921.
- Fribourg, S., Kellenberger, E., Rogniaux, H., Poterszman, A., Van Dorsselaer, A., Thierry, J. C., Egly, J. M., Moras, D., and Kieffer, B. (2000). Structural characterization of the cysteine-rich domain of TFIIF p44 subunit. *J Biol Chem* 275, 31963-31971.
- Fribourg, S., Romier, C., Werten, S., Gangloff, Y. G., Poterszman, A., and Moras, D. (2001). Dissecting the interaction network of multiprotein complexes by pairwise coexpression of subunits in *E. coli*. *J Mol Biol* 306, 363-373.
- Friedberg, E. C. (1996). Cockayne syndrome--a primary defect in DNA repair, transcription, both or neither? *Bioessays* 18, 731-738.
- Fukuda, A., Nogi, Y., and Hisatake, K. (2002). The regulatory role for the ERCC3 helicase of general transcription factor TFIIF during promoter escape in transcriptional activation. *Proc Natl Acad Sci U S A* 99, 1206-1211.
- Fukuda, A., Yamauchi, J., Wu, S. Y., Chiang, C. M., Muramatsu, M., and Hisatake, K. (2001). Reconstitution of recombinant TFIIF that can mediate activator-dependent transcription. *Genes Cells* 6, 707-719.
- Gangloff, Y. G., Sanders, S. L., Romier, C., Kirschner, D., Weil, P. A., Tora, L., and Davidson, I. (2001). 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, 1841-1853.
- Geiger, J. H., Hahn, S., Lee, S., and Sigler, P. B. (1996). Crystal structure of the yeast TFIIF/TBP/DNA complex. *Science* 272, 830-836.
- Gerard, M., Fischer, L., Moncollin, V., Chipoulet, J. M., Chambon, P., and Egly, J. M. (1991). Purification and interaction properties of the human RNA polymerase B(II) general transcription factor BTF2. *J Biol Chem* 266, 20940-20945.
- Gervais, V., Busso, D., Wasielewski, E., Poterszman, A., Egly, J. M., Thierry, J. C., and Kieffer, B. (2001). Solution structure of the N-terminal domain of the human TFIIF MAT1 subunit: new insights into the RING finger family. *J Biol Chem* 276, 7457-7464.
- Gileadi, O., Feaver, W. J., and Kornberg, R. D. (1992). Cloning of a subunit of yeast RNA polymerase II transcription factor b and CTD kinase. *Science* 257, 1389-1392.
- Gillespie, J. M., and Marshall, R. C. (1983). A comparison of the proteins of normal and trichothiodystrophic human hair. *J Invest Dermatol* 80, 195-202.
- Gnatt, A. L., Cramer, P., Fu, J., Bushnell, D. A., and Kornberg, R. D. (2001). Structural basis of transcription: an RNA polymerase II elongation complex at 3.3 Å resolution. *Science* 292, 1876-1882.

- Gorbalenya, A. E., and Koonin, E. V. (1991). Endonuclease (R) subunits of type-I and type-III restriction-modification enzymes contain a helicase-like domain. *FEBS Lett* 291, 277-281.
- Gu, W., Malik, S., Ito, M., Yuan, C. X., Fondell, J. D., Zhang, X., Martinez, E., Qin, J., and Roeder, R. G. (1999). A novel human SRB/MED-containing cofactor complex, SMCC, involved in transcription regulation. *Mol Cell* 3, 97-108.
- Guzder, S. N., Habraken, Y., Sung, P., Prakash, L., and Prakash, S. (1996). RAD26, the yeast homolog of human Cockayne's syndrome group B gene, encodes a DNA-dependent ATPase. *J Biol Chem* 271, 18314-18317.
- Habraken, Y., Sung, P., Prakash, S., and Prakash, L. (1996). Transcription factor TFIIH and DNA endonuclease Rad2 constitute yeast nucleotide excision repair factor 3: implications for nucleotide excision repair and Cockayne syndrome. *Proc Natl Acad Sci U S A* 93, 10718-10722.
- Hampsey, M. (1998). Molecular genetics of the RNA polymerase II general transcriptional machinery. *Microbiol Mol Biol Rev* 62, 465-503.
- Hassan, A. H., Neely, K. E., and Workman, J. L. (2001). Histone acetyltransferase complexes stabilize swi/snf binding to promoter nucleosomes. *Cell* 104, 817-827.
- Henning, K. A., Li, L., Iyer, N., McDaniel, L. D., Reagan, M. S., Legerski, R., Schultz, R. A., Stefanini, M., Lehmann, A. R., Mayne, L. V., and et al. (1995). The Cockayne syndrome group A gene encodes a WD repeat protein that interacts with CSB protein and a subunit of RNA polymerase II TFIIH. *Cell* 82, 555-564.
- Hoeijmakers, J. H. (2001). Genome maintenance mechanisms for preventing cancer. *Nature* 411, 366-374.
- Hoffman, A., Sinn, E., Yamamoto, T., Wang, J., Roy, A., Horikoshi, M., and Roeder, R. G. (1990). Highly conserved core domain and unique N terminus with presumptive regulatory motifs in a human TATA factor (TFIID). *Nature* 346, 387-390.
- Holstege, F. C., Fiedler, U., and Timmers, H. T. (1997). Three transitions in the RNA polymerase II transcription complex during initiation. *Embo J* 16, 7468-7480.
- Holstege, F. C., Jennings, E. G., Wyrick, J. J., Lee, T. I., Hengartner, C. J., Green, M. R., Golub, T. R., Lander, E. S., and Young, R. A. (1998). Dissecting the regulatory circuitry of a eukaryotic genome. *Cell* 95, 717-728.
- Holstege, F. C., van der Vliet, P. C., and Timmers, H. T. (1996). Opening of an RNA polymerase II promoter occurs in two distinct steps and requires the basal transcription factors IIE and IIH. *Embo J* 15, 1666-1677.
- Hoogstraten, D., Nigg, A. L., Heath, H., Mullenders, L. H., van Driel, R., Hoeijmakers, J. H., Vermeulen, W., and Houtsmuller, A. B. (2002). Rapid switching of TFIIH between RNA polymerase I and II transcription and DNA repair in vivo. *Mol Cell* 10, 1163-1174.

- Humbert, S., van Vuuren, H., Lutz, Y., Hoeijmakers, J. H., Egly, J. M., and Moncollin, V. (1994). p44 and p34 subunits of the BTF2/TFIIH transcription factor have homologies with SSL1, a yeast protein involved in DNA repair. *Embo J* 13, 2393-2398.
- Hwang, J. R., Moncollin, V., Vermeulen, W., Seroz, T., van Vuuren, H., Hoeijmakers, J. H., and Egly, J. M. (1996). A 3' --> 5' XPB helicase defect in repair/transcription factor TFIIH of xeroderma pigmentosum group B affects both DNA repair and transcription. *J Biol Chem* 271, 15898-15904.
- Iben, S., Tschochner, H., Bier, M., Hoogstraten, D., Hozak, P., Egly, J. M., and Grummt, I. (2002). TFIIH plays an essential role in RNA polymerase I transcription. *Cell* 109, 297-306.
- Inamoto, S., Segil, N., Pan, Z. Q., Kimura, M., and Roeder, R. G. (1997). The cyclin-dependent kinase-activating kinase (CAK) assembly factor, MAT1, targets and enhances CAK activity on the POU domains of octamer transcription factors. *J Biol Chem* 272, 29852-29858.
- Inostroza, J., Flores, O., and Reinberg, D. (1991). Factors involved in specific transcription by mammalian RNA polymerase II. Purification and functional analysis of general transcription factor IIE. *J Biol Chem* 266, 9304-9308.
- Inostroza, J. A., Mermelstein, F. H., Ha, I., Lane, W. S., and Reinberg, D. (1992). Dr1, a TATA-binding protein-associated phosphoprotein and inhibitor of class II gene transcription. *Cell* 70, 477-489.
- Ito, M., Yuan, C. X., Malik, S., Gu, W., Fondell, J. D., Yamamura, S., Fu, Z. Y., Zhang, X., Qin, J., and Roeder, R. G. (1999). Identity between TRAP and SMCC complexes indicates novel pathways for the function of nuclear receptors and diverse mammalian activators. *Mol Cell* 3, 361-370.
- Iyer, N., Reagan, M. S., Wu, K. J., Canagarajah, B., and Friedberg, E. C. (1996). Interactions involving the human RNA polymerase II transcription/nucleotide excision repair complex TFIIH, the nucleotide excision repair protein XPG, and Cockayne syndrome group B (CSB) protein. *Biochemistry* 35, 2157-2167.
- Javahery, R., Khachi, A., Lo, K., Zenzie-Gregory, B., and Smale, S. T. (1994). DNA sequence requirements for transcriptional initiator activity in mammalian cells. *Mol Cell Biol* 14, 116-127.
- Jawhari, A., Laine, J. P., Dubaele, S., Lamour, V., Poterszman, A., Coin, F., Moras, D., and Egly, J. M. (2002a). p52 Mediates XPB function within the transcription/repair factor TFIIH. *J Biol Chem* 277, 31761-31767.
- Jawhari, A., Uhring, M., Crucifix, C., Fribourg, S., Schultz, P., Poterszman, A., Egly, J. M., and Moras, D. (2002b). Expression of FLAG fusion proteins in insect cells: application to the multi-subunit transcription/DNA repair factor TFIIH. *Protein Expr Purif* 24, 513-523.
- Jeffrey, P. D., Russo, A. A., Polyak, K., Gibbs, E., Hurwitz, J., Massague, J., and Pavletich, N. P. (1995). Mechanism of CDK activation revealed by the structure of a cyclinA-CDK2 complex. *Nature* 376, 313-320.

- Juo, Z. S., Kassavetis, G. A., Wang, J., Geiduschek, E. P., and Sigler, P. B. (2003). Crystal structure of a transcription factor IIIB core interface ternary complex. *Nature* *422*, 534-539.
- Kamada, K., Shu, F., Chen, H., Malik, S., Stelzer, G., Roeder, R. G., Meisterernst, M., and Burley, S. K. (2001). Crystal structure of negative cofactor 2 recognizing the TBP-DNA transcription complex. *Cell* *106*, 71-81.
- Kelleher, R. J., 3rd, Flanagan, P. M., and Kornberg, R. D. (1990). A novel mediator between activator proteins and the RNA polymerase II transcription apparatus. *Cell* *61*, 1209-1215.
- Keriel, A., Stary, A., Sarasin, A., Rochette-Egly, C., and Egly, J. M. (2002). XPD mutations prevent TFIIH-dependent transactivation by nuclear receptors and phosphorylation of RARalpha. *Cell* *109*, 125-135.
- Kettenberger, H., Armache, K. J., and Cramer, P. (2003). Architecture of the RNA polymerase II-TFIIS complex and implications for mRNA cleavage. *Cell* *114*, 347-357.
- Khrapunov, S., Pastor, N., and Brenowitz, M. (2002). Solution structural studies of the *Saccharomyces cerevisiae* TATA binding protein (TBP). *Biochemistry* *41*, 9559-9571.
- Kim, J. L., and Burley, S. K. (1994). 1.9 Å resolution refined structure of TBP recognizing the minor groove of TATAAAAG. *Nat Struct Biol* *1*, 638-653.
- Kim, J. L., Nikolov, D. B., and Burley, S. K. (1993). Co-crystal structure of TBP recognizing the minor groove of a TATA element. *Nature* *365*, 520-527.
- Kim, K. K., Chamberlin, H. M., Morgan, D. O., and Kim, S. H. (1996). Three-dimensional structure of human cyclin H, a positive regulator of the CDK-activating kinase. *Nat Struct Biol* *3*, 849-855.
- Kim, T. K., Ebright, R. H., and Reinberg, D. (2000). Mechanism of ATP-dependent promoter melting by transcription factor IIH. *Science* *288*, 1418-1422.
- Kim, W. Y., and Dahmus, M. E. (1989). The major late promoter of adenovirus-2 is accurately transcribed by RNA polymerases IIO, IIA, and IIB. *J Biol Chem* *264*, 3169-3176.
- Kirchner, J. M., Salazar, E. P., Lamerdin, J. E., Montgomery, M. A., Carrano, A. V., and Weber, C. A. (1994). Cloning and molecular characterization of the Chinese hamster ERCC2 nucleotide excision repair gene. *Genomics* *23*, 592-599.
- Ko, L. J., Shieh, S. Y., Chen, X., Jayaraman, L., Tamai, K., Taya, Y., Prives, C., and Pan, Z. Q. (1997). p53 is phosphorylated by CDK7-cyclin H in a p36MAT1-dependent manner. *Mol Cell Biol* *17*, 7220-7229.
- Kobor, M. S., Archambault, J., Lester, W., Holstege, F. C., Gileadi, O., Jansma, D. B., Jennings, E. G., Kouyoumdjian, F., Davidson, A. R., Young, R. A., and Greenblatt, J. (1999). An unusual eukaryotic protein phosphatase required for transcription by RNA polymerase II and CTD dephosphorylation in *S. cerevisiae*. *Mol Cell* *4*, 55-62.

- Koleske, A. J., and Young, R. A. (1995). The RNA polymerase II holoenzyme and its implications for gene regulation. *Trends Biochem Sci* 20, 113-116.
- Kugel, J. F., and Goodrich, J. A. (1998). Promoter escape limits the rate of RNA polymerase II transcription and is enhanced by TFIIE, TFIIH, and ATP on negatively supercoiled DNA. *Proc Natl Acad Sci U S A* 95, 9232-9237.
- Kuras, L., and Struhl, K. (1999). Binding of TBP to promoters in vivo is stimulated by activators and requires Pol II holoenzyme. *Nature* 399, 609-613.
- Kwek, K. Y., Murphy, S., Furger, A., Thomas, B., O'Gorman, W., Kimura, H., Proudfoot, N. J., and Akoulitchev, A. (2002). U1 snRNA associates with TFIIH and regulates transcriptional initiation. *Nat Struct Biol* 9, 800-805.
- Lagrange, T., Kapanidis, A. N., Tang, H., Reinberg, D., and Ebright, R. H. (1998). New core promoter element in RNA polymerase II-dependent transcription: sequence-specific DNA binding by transcription factor IIB. *Genes Dev* 12, 34-44.
- Laybourn, P. J., and Dahmus, M. E. (1990). Phosphorylation of RNA polymerase IIA occurs subsequent to interaction with the promoter and before the initiation of transcription. *J Biol Chem* 265, 13165-13173.
- Lee, M., and Struhl, K. (2001). Multiple functions of the nonconserved N-terminal domain of yeast TATA-binding protein. *Genetics* 158, 87-93.
- Lee, M. S., and Garrard, W. T. (1992). Uncoupling gene activity from chromatin structure: promoter mutations can inactivate transcription of the yeast HSP82 gene without eliminating nucleosome-free regions. *Proc Natl Acad Sci U S A* 89, 9166-9170.
- Lei, L., Ren, D., and Burton, Z. F. (1999). The RAP74 subunit of human transcription factor IIF has similar roles in initiation and elongation. *Mol Cell Biol* 19, 8372-8382.
- 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.
- Leuther, K. K., Bushnell, D. A., and Kornberg, R. D. (1996). Two-dimensional crystallography of TFIIIB- and IIE-RNA polymerase II complexes: implications for start site selection and initiation complex formation. *Cell* 85, 773-779.
- Leveillard, T., Andera, L., Bissonnette, N., Schaeffer, L., Bracco, L., Egly, J. M., and Wasylyk, B. (1996). Functional interactions between p53 and the TFIIH complex are affected by tumour-associated mutations. *Embo J* 15, 1615-1624.
- Li, X. Y., Virbasius, A., Zhu, X., and Green, M. R. (1999). Enhancement of TBP binding by activators and general transcription factors. *Nature* 399, 605-609.
- Lin, Y. S., Ha, I., Maldonado, E., Reinberg, D., and Green, M. R. (1991). Binding of general transcription factor TFIIIB to an acidic activating region. *Nature* 353, 569-571.

- Littlefield, O., Korkhin, Y., and Sigler, P. B. (1999). The structural basis for the oriented assembly of a TBP/TFB/promoter complex. *Proc Natl Acad Sci U S A* *96*, 13668-13673.
- Liu, D., Ishima, R., Tong, K. I., Bagby, S., Kokubo, T., Muhandiram, D. R., Kay, L. E., Nakatani, Y., and 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.
- Lorch, Y., Zhang, M., and Kornberg, R. D. (1999). Histone octamer transfer by a chromatin-remodeling complex. *Cell* *96*, 389-392.
- Lu, H., Fisher, R. P., Bailey, P., and Levine, A. J. (1997). The CDK7-cycH-p36 complex of transcription factor IIH phosphorylates p53, enhancing its sequence-specific DNA binding activity in vitro. *Mol Cell Biol* *17*, 5923-5934.
- Lu, H., Flores, O., Weinmann, R., and Reinberg, D. (1991). The nonphosphorylated form of RNA polymerase II preferentially associates with the preinitiation complex. *Proc Natl Acad Sci U S A* *88*, 10004-10008.
- Luger, K., Mader, A. W., Richmond, R. K., Sargent, D. F., and Richmond, T. J. (1997). Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* *389*, 251-260.
- Ma, L., Siemssen, E. D., Noteborn, H. M., and van der Eb, A. J. (1994). The xeroderma pigmentosum group B protein ERCC3 produced in the baculovirus system exhibits DNA helicase activity. *Nucleic Acids Res* *22*, 4095-4102.
- Makela, T. P., Tassan, J. P., Nigg, E. A., Frutiger, S., Hughes, G. J., and Weinberg, R. A. (1994). A cyclin associated with the CDK-activating kinase MO15. *Nature* *371*, 254-257.
- Maldonado, E., Shiekhatar, R., Sheldon, M., Cho, H., Drapkin, R., Rickert, P., Lees, E., Anderson, C. W., Linn, S., and Reinberg, D. (1996). A human RNA polymerase II complex associated with SRB and DNA-repair proteins. *Nature* *381*, 86-89.
- Marinoni, J. C., Roy, R., Vermeulen, W., Miniou, P., Lutz, Y., Weeda, G., Seroz, T., Gomez, D. M., Hoeijmakers, J. H., and Egly, J. M. (1997). Cloning and characterization of p52, the fifth subunit of the core of the transcription/DNA repair factor TFIIH. *Embo J* *16*, 1093-1102.
- Marshall, N. F., Peng, J., Xie, Z., and Price, D. H. (1996). Control of RNA polymerase II elongation potential by a novel carboxyl-terminal domain kinase. *J Biol Chem* *271*, 27176-27183.
- Martin, M. L., Lieberman, P. M., and Curran, T. (1996). Fos-Jun dimerization promotes interaction of the basic region with TFIIIE-34 and TFIIIF. *Mol Cell Biol* *16*, 2110-2118.
- Masutani, C., Kusumoto, R., Yamada, A., Dohmae, N., Yokoi, M., Yuasa, M., Araki, M., Iwai, S., Takio, K., and Hanaoka, F. (1999). The XPV (xeroderma pigmentosum variant) gene encodes human DNA polymerase ϵ . *Nature* *399*, 700-704.
- Matsui, P., DePaulo, J., and Buratowski, S. (1995). An interaction between the Tfb1 and Ssl1 subunits of yeast TFIIH correlates with DNA repair activity. *Nucleic Acids Res* *23*, 767-772.

- Matsui, T., Segall, J., Weil, P. A., and Roeder, R. G. (1980). Multiple factors required for accurate initiation of transcription by purified RNA polymerase II. *J Biol Chem* 255, 11992-11996.
- Maxon, M. E., Goodrich, J. A., and Tjian, R. (1994). Transcription factor IIE binds preferentially to RNA polymerase IIa and recruits TFIIF: a model for promoter clearance. *Genes Dev* 8, 515-524.
- McDonald, J. P., Ropic-Otrin, V., Epstein, J. A., Broughton, B. C., Wang, X., Lehmann, A. R., Wolgemuth, D. J., and Woodgate, R. (1999). Novel human and mouse homologs of *Saccharomyces cerevisiae* DNA polymerase ϵ . *Genomics* 60, 20-30.
- Meisterernst, M., and Roeder, R. G. (1991). Family of proteins that interact with TFIID and regulate promoter activity. *Cell* 67, 557-567.
- Mizzen, C. A., and Allis, C. D. (1998). Linking histone acetylation to transcriptional regulation. *Cell Mol Life Sci* 54, 6-20.
- Mizzen, C. A., Yang, X. J., Kokubo, T., Brownell, J. E., Bannister, A. J., Owen-Hughes, T., Workman, J., Wang, L., Berger, S. L., Kouzarides, T., *et al.* (1996). The TAF(II)250 subunit of TFIID has histone acetyltransferase activity. *Cell* 87, 1261-1270.
- Moreland, R. J., Tirode, F., Yan, Q., Conaway, J. W., Egly, J. M., and Conaway, R. C. (1999). A role for the TFIIF XPB DNA helicase in promoter escape by RNA polymerase II. *J Biol Chem* 274, 22127-22130.
- Mu, D., Wakasugi, M., Hsu, D. S., and Sancar, A. (1997). Characterization of reaction intermediates of human excision repair nuclease. *J Biol Chem* 272, 28971-28979.
- Murray, J. M., Doe, C. L., Schenk, P., Carr, A. M., Lehmann, A. R., and Watts, F. Z. (1992). Cloning and characterisation of the *S. pombe* rad15 gene, a homologue to the *S. cerevisiae* RAD3 and human ERCC2 genes. *Nucleic Acids Res* 20, 2673-2678.
- Nance, M. A., and Berry, S. A. (1992). Cockayne syndrome: review of 140 cases. *Am J Med Genet* 42, 68-84.
- Neish, A. S., Anderson, S. F., Schlegel, B. P., Wei, W., and Parvin, J. D. (1998). Factors associated with the mammalian RNA polymerase II holoenzyme. *Nucleic Acids Res* 26, 847-853.
- Nelson, D. A., Ferris, R. C., Zhang, D. E., and Ferez, C. R. (1986). The beta-globin domain in immature chicken erythrocytes: enhanced solubility is coincident with histone hyperacetylation. *Nucleic Acids Res* 14, 1667-1682.
- Nigg, E. A. (1996). Cyclin-dependent kinase 7: at the cross-roads of transcription, DNA repair and cell cycle control? *Curr Opin Cell Biol* 8, 312-317.
- Nikolov, D. B., Chen, H., Halay, E. D., Hoffman, A., Roeder, R. G., and Burley, S. K. (1996). Crystal structure of a human TATA box-binding protein/TATA element complex. *Proc Natl Acad Sci U S A* 93, 4862-4867.

- Nikolov, D. B., Chen, H., Halay, E. D., Usheva, A. A., Hisatake, K., Lee, D. K., Roeder, R. G., and Burley, S. K. (1995). Crystal structure of a TFIIB-TBP-TATA-element ternary complex. *Nature* *377*, 119-128.
- Nikolov, D. B., Hu, S. H., Lin, J., Gasch, A., Hoffmann, A., Horikoshi, M., Chua, N. H., Roeder, R. G., and Burley, S. K. (1992). Crystal structure of TFIID TATA-box binding protein. *Nature* *360*, 40-46.
- Nogales, E. (2000). Recent structural insights into transcription preinitiation complexes. *J Cell Sci* *113 Pt 24*, 4391-4397.
- Nonet, M. L., and Young, R. A. (1989). Intragenic and extragenic suppressors of mutations in the heptapeptide repeat domain of *Saccharomyces cerevisiae* RNA polymerase II. *Genetics* *123*, 715-724.
- Nouspikel, T., Lalle, P., Leadon, S. A., Cooper, P. K., and Clarkson, S. G. (1997). A common mutational pattern in Cockayne syndrome patients from xeroderma pigmentosum group G: implications for a second XPG function. *Proc Natl Acad Sci U S A* *94*, 3116-3121.
- Nudler, E., Goldfarb, A., and Kashlev, M. (1994). Discontinuous mechanism of transcription elongation. *Science* *265*, 793-796.
- O'Brien, T., and Tjian, R. (1998). Functional analysis of the human TAFII250 N-terminal kinase domain. *Mol Cell* *1*, 905-911.
- O'Donovan, A., Davies, A. A., Moggs, J. G., West, S. C., and Wood, R. D. (1994). XPG endonuclease makes the 3' incision in human DNA nucleotide excision repair. *Nature* *371*, 432-435.
- Ogryzko, V. V., Schiltz, R. L., Russanova, V., Howard, B. H., and Nakatani, Y. (1996). The transcriptional coactivators p300 and CBP are histone acetyltransferases. *Cell* *87*, 953-959.
- Ohkuma, Y., Sumimoto, H., Horikoshi, M., and Roeder, R. G. (1990). Factors involved in specific transcription by mammalian RNA polymerase II: purification and characterization of general transcription factor TFIIE. *Proc Natl Acad Sci U S A* *87*, 9163-9167.
- Okuda, M., Watanabe, Y., Okamura, H., Hanaoka, F., Ohkuma, Y., and Nishimura, Y. (2000). Structure of the central core domain of TFIIEbeta with a novel double-stranded DNA-binding surface. *Embo J* *19*, 1346-1356.
- Oliva, R., Bazett-Jones, D. P., Locklear, L., and Dixon, G. H. (1990). Histone hyperacetylation can induce unfolding of the nucleosome core particle. *Nucleic Acids Res* *18*, 2739-2747.
- Opalka, N., Chlenov, M., Chacon, P., Rice, W. J., Wriggers, W., and Darst, S. A. (2003). Structure and function of the transcription elongation factor GreB bound to bacterial RNA polymerase. *Cell* *114*, 335-345.
- Orphanides, G., Lagrange, T., and Reinberg, D. (1996). The general transcription factors of RNA polymerase II. *Genes Dev* *10*, 2657-2683.

- Orphanides, G., LeRoy, G., Chang, C. H., Luse, D. S., and Reinberg, D. (1998). FACT, a factor that facilitates transcript elongation through nucleosomes. *Cell* 92, 105-116.
- Owen-Hughes, T., and Workman, J. L. (1994). Experimental analysis of chromatin function in transcription control. *Crit Rev Eukaryot Gene Expr* 4, 403-441.
- Ozer, J., Moore, P. A., Bolden, A. H., Lee, A., Rosen, C. A., and Lieberman, P. M. (1994). Molecular cloning of the small (gamma) subunit of human TFIIA reveals functions critical for activated transcription. *Genes Dev* 8, 2324-2335.
- Park, C. H., Mu, D., Reardon, J. T., and Sancar, A. (1995). The general transcription-repair factor TFIIH is recruited to the excision repair complex by the XPA protein independent of the TFIIIE transcription factor. *J Biol Chem* 270, 4896-4902.
- Perez-Howard, G. M., Weil, P. A., and Beechem, J. M. (1995). Yeast TATA binding protein interaction with DNA: fluorescence determination of oligomeric state, equilibrium binding, on-rate, and dissociation kinetics. *Biochemistry* 34, 8005-8017.
- Peterson, C. L., and Workman, J. L. (2000). Promoter targeting and chromatin remodeling by the SWI/SNF complex. *Curr Opin Genet Dev* 10, 187-192.
- Potier, N., Lamour, V., Poterszman, A., Thierry, J. C., Moras, D., and Van Dorsselaer, A. (2000). Characterization of crystal content by ESI-MS and MALDI-MS. *Acta Crystallogr D Biol Crystallogr* 56 Pt 12, 1583-1590.
- Pugh, B. F., and Tjian, R. (1990). Mechanism of transcriptional activation by Sp1: evidence for coactivators. *Cell* 61, 1187-1197.
- Ranish, J. A., and Hahn, S. (1996). Transcription: basal factors and activation. *Curr Opin Genet Dev* 6, 151-158.
- Rea, S., Eisenhaber, F., O'Carroll, D., Strahl, B. D., Sun, Z. W., Schmid, M., Opravil, S., Mechtler, K., Ponting, C. P., Allis, C. D., and Jenuwein, T. (2000). Regulation of chromatin structure by site-specific histone H3 methyltransferases. *Nature* 406, 593-599.
- Reese, J. C., Zhang, Z., and Kurpad, H. (2000). Identification of a yeast transcription factor IID subunit, TSG2/TAF48. *J Biol Chem* 275, 17391-17398.
- Reines, D., Conaway, J. W., and Conaway, R. C. (1996). The RNA polymerase II general elongation factors. *Trends Biochem Sci* 21, 351-355.
- Richmond, T. J., and Davey, C. A. (2003). The structure of DNA in the nucleosome core. *Nature* 423, 145-150.
- Riedl, T., and Egly, J. M. (2000). Phosphorylation in transcription: the CTD and more. *Gene Expr* 9, 3-13.
- Roberts, S. G., and Green, M. R. (1994). Activator-induced conformational change in general transcription factor TFIIB. *Nature* 371, 717-720.

- Roberts, S. G., Ha, I., Maldonado, E., Reinberg, D., and Green, M. R. (1993). Interaction between an acidic activator and transcription factor TFIIB is required for transcriptional activation. *Nature* *363*, 741-744.
- Robles, A. I., Wang, X. W., and Harris, C. C. (1999). Drug-induced apoptosis is delayed and reduced in XPD lymphoblastoid cell lines: possible role of TFIIH in p53-mediated apoptotic cell death. *Oncogene* *18*, 4681-4688.
- Rochette-Egly, C., Adam, S., Rossignol, M., Egly, J. M., and Chambon, P. (1997). Stimulation of RAR alpha activation function AF-1 through binding to the general transcription factor TFIIH and phosphorylation by CDK7. *Cell* *90*, 97-107.
- Rossignol, M., Kolb-Cheynel, I., and Egly, J. M. (1997). Substrate specificity of the cdk-activating kinase (CAK) is altered upon association with TFIIH. *Embo J* *16*, 1628-1637.
- Rossmann, M. G. (1990). The molecular replacement method. *Acta Crystallogr A* *46 (Pt 2)*, 73-82.
- Roy, R., Adamczewski, J. P., Seroz, T., Vermeulen, W., Tassan, J. P., Schaeffer, L., Nigg, E. A., Hoeijmakers, J. H., and Egly, J. M. (1994). The MO15 cell cycle kinase is associated with the TFIIH transcription-DNA repair factor. *Cell* *79*, 1093-1101.
- Rudd, M. D., Izban, M. G., and Luse, D. S. (1994). The active site of RNA polymerase II participates in transcript cleavage within arrested ternary complexes. *Proc Natl Acad Sci U S A* *91*, 8057-8061.
- Ruppert, S., Wang, E. H., and Tjian, R. (1993). Cloning and expression of human TAFII250: a TBP-associated factor implicated in cell-cycle regulation. *Nature* *362*, 175-179.
- Russo, A. A., Jeffrey, P. D., and Pavletich, N. P. (1996). Structural basis of cyclin-dependent kinase activation by phosphorylation. *Nat Struct Biol* *3*, 696-700.
- Schaeffer, L., Moncollin, V., Roy, R., Staub, A., Mezzina, M., Sarasin, A., Weeda, G., Hoeijmakers, J. H., and Egly, J. M. (1994). The ERCC2/DNA repair protein is associated with the class II BTF2/TFIIH transcription factor. *Embo J* *13*, 2388-2392.
- Schaeffer, L., Roy, R., Humbert, S., Moncollin, V., Vermeulen, W., Hoeijmakers, J. H., Chambon, P., and Egly, J. M. (1993). DNA repair helicase: a component of BTF2 (TFIIH) basic transcription factor. *Science* *260*, 58-63.
- Schneider, E., Montenarh, M., and Wagner, P. (1998). Regulation of CAK kinase activity by p53. *Oncogene* *17*, 2733-2741.
- Schultz, P., Fribourg, S., Poterszman, A., Mallouh, V., Moras, D., and Egly, J. M. (2000). Molecular structure of human TFIIH. *Cell* *102*, 599-607.
- Selby, C. P., and Sancar, A. (1993). Molecular mechanism of transcription-repair coupling. *Science* *260*, 53-58.

- Selby, C. P., and Sancar, A. (1997). Human transcription-repair coupling factor CSB/ERCC6 is a DNA-stimulated ATPase but is not a helicase and does not disrupt the ternary transcription complex of stalled RNA polymerase II. *J Biol Chem* 272, 1885-1890.
- Selleck, W., Howley, R., Fang, Q., Podolny, V., Fried, M. G., Buratowski, S., and Tan, S. (2001). A histone fold TAF octamer within the yeast TFIID transcriptional coactivator. *Nat Struct Biol* 8, 695-700.
- Seroz, T., Perez, C., Bergmann, E., Bradsher, J., and Egly, J. M. (2000). p44/SSL1, the regulatory subunit of the XPD/RAD3 helicase, plays a crucial role in the transcriptional activity of TFIIH. *J Biol Chem* 275, 33260-33266.
- Sobel, R. E., Cook, R. G., Perry, C. A., Annunziato, A. T., and Allis, C. D. (1995). Conservation of deposition-related acetylation sites in newly synthesized histones H3 and H4. *Proc Natl Acad Sci U S A* 92, 1237-1241.
- Stargell, L. A., and Struhl, K. (1995). The TBP-TFIIA interaction in the response to acidic activators in vivo. *Science* 269, 75-78.
- Stefanini, M., Lagomarsini, P., Arlett, C. F., Marinoni, S., Borrone, C., Crovato, F., Trevisan, G., Cordone, G., and Nuzzo, F. (1986). Xeroderma pigmentosum (complementation group D) mutation is present in patients affected by trichothiodystrophy with photosensitivity. *Hum Genet* 74, 107-112.
- Sterner, D. E., and Berger, S. L. (2000). Acetylation of histones and transcription-related factors. *Microbiol Mol Biol Rev* 64, 435-459.
- Strahl, B. D., and Allis, C. D. (2000). The language of covalent histone modifications. *Nature* 403, 41-45.
- Struhl, K. (1998). Histone acetylation and transcriptional regulatory mechanisms. *Genes Dev* 12, 599-606.
- Studitsky, V. M., Clark, D. J., and Felsenfeld, G. (1994). A histone octamer can step around a transcribing polymerase without leaving the template. *Cell* 76, 371-382.
- Sugasawa, K., Ng, J. M., Masutani, C., Iwai, S., van der Spek, P. J., Eker, A. P., Hanaoka, F., Bootsma, D., and Hoeijmakers, J. H. (1998). Xeroderma pigmentosum group C protein complex is the initiator of global genome nucleotide excision repair. *Mol Cell* 2, 223-232.
- Sung, P., Bailly, V., Weber, C., Thompson, L. H., Prakash, L., and Prakash, S. (1993). Human xeroderma pigmentosum group D gene encodes a DNA helicase. *Nature* 365, 852-855.
- Sung, P., Guzder, S. N., Prakash, L., and Prakash, S. (1996). Reconstitution of TFIIH and requirement of its DNA helicase subunits, Rad3 and Rad25, in the incision step of nucleotide excision repair. *J Biol Chem* 271, 10821-10826.
- Sung, P., Watkins, J. F., Prakash, L., and Prakash, S. (1994). Negative superhelicity promotes ATP-dependent binding of yeast RAD3 protein to ultraviolet-damaged DNA. *J Biol Chem* 269, 8303-8308.

- Svejstrup, J. Q., Wang, Z., Feaver, W. J., Wu, X., Bushnell, D. A., Donahue, T. F., Friedberg, E. C., and Kornberg, R. D. (1995). Different forms of TFIIH for transcription and DNA repair: holo-TFIIH and a nucleotide excision repairosome. *Cell* 80, 21-28.
- Tan, S., Hunziker, Y., Sargent, D. F., and Richmond, T. J. (1996). Crystal structure of a yeast TFIIA/TBP/DNA complex. *Nature* 381, 127-151.
- Tansey, W. P., and Herr, W. (1995). The ability to associate with activation domains in vitro is not required for the TATA box-binding protein to support activated transcription in vivo. *Proc Natl Acad Sci U S A* 92, 10550-10554.
- Tassan, J. P., Jaquenoud, M., Fry, A. M., Frutiger, S., Hughes, G. J., and Nigg, E. A. (1995). In vitro assembly of a functional human CDK7-cyclin H complex requires MAT1, a novel 36 kDa RING finger protein. *Embo J* 14, 5608-5617.
- Tassan, J. P., Schultz, S. J., Bartek, J., and Nigg, E. A. (1994). Cell cycle analysis of the activity, subcellular localization, and subunit composition of human CAK (CDK-activating kinase). *J Cell Biol* 127, 467-478.
- Theis, K., Chen, P. J., Skorvaga, M., Van Houten, B., and Kisker, C. (1999). Crystal structure of UvrB, a DNA helicase adapted for nucleotide excision repair. *Embo J* 18, 6899-6907.
- Timmers, H. T. (1994). Transcription initiation by RNA polymerase II does not require hydrolysis of the beta-gamma phosphoanhydride bond of ATP. *Embo J* 13, 391-399.
- Tirode, F., Busso, D., Coin, F., and Egly, J. M. (1999). Reconstitution of the transcription factor TFIIH: assignment of functions for the three enzymatic subunits, XPB, XPD, and cdk7. *Mol Cell* 3, 87-95.
- Tong, X., Drapkin, R., Reinberg, D., and Kieff, E. (1995). The 62- and 80-kDa subunits of transcription factor IIIH mediate the interaction with Epstein-Barr virus nuclear protein 2. *Proc Natl Acad Sci U S A* 92, 3259-3263.
- Tremeau-Bravard, A., Perez, C., and Egly, J. M. (2001). A role of the C-terminal part of p44 in the promoter escape activity of transcription factor IIIH. *J Biol Chem* 276, 27693-27697.
- Truant, R., Antunovic, J., Greenblatt, J., Prives, C., and Cromlish, J. A. (1995). Direct interaction of the hepatitis B virus HBx protein with p53 leads to inhibition by HBx of p53 response element-directed transactivation. *J Virol* 69, 1851-1859.
- Tsai, F. T., and Sigler, P. B. (2000). Structural basis of preinitiation complex assembly on human pol II promoters. *Embo J* 19, 25-36.
- Tyrrell, R. M., and Amaudruz, F. (1987). Evidence for two independent pathways of biologically effective excision repair from its rate and extent in cells cultured from sun-sensitive humans. *Cancer Res* 47, 3725-3728.
- Uetz, P., Giot, L., Cagney, G., Mansfield, T. A., Judson, R. S., Knight, J. R., Lockshon, D., Narayan, V., Srinivasan, M., Pochart, P., *et al.* (2000). A comprehensive analysis of protein-protein interactions in *Saccharomyces cerevisiae*. *Nature* 403, 623-627.

- Usheva, A., Maldonado, E., Goldring, A., Lu, H., Houbavi, C., Reinberg, D., and Aloni, Y. (1992). Specific interaction between the nonphosphorylated form of RNA polymerase II and the TATA-binding protein. *Cell* 69, 871-881.
- van Gool, A. J., Verhage, R., Swagemakers, S. M., van de Putte, P., Brouwer, J., Troelstra, C., Bootsma, D., and Hoeijmakers, J. H. (1994). RAD26, the functional *S. cerevisiae* homolog of the Cockayne syndrome B gene ERCC6. *Embo J* 13, 5361-5369.
- van Holde, K. E., Lohr, D. E., and Robert, C. (1992). What happens to nucleosomes during transcription? *J Biol Chem* 267, 2837-2840.
- Vandel, L., and Kouzarides, T. (1999). Residues phosphorylated by TFIID are required for E2F-1 degradation during S-phase. *Embo J* 18, 4280-4291.
- Venema, J., Mullenders, L. H., Natarajan, A. T., van Zeeland, A. A., and Mayne, L. V. (1990). The genetic defect in Cockayne syndrome is associated with a defect in repair of UV-induced DNA damage in transcriptionally active DNA. *Proc Natl Acad Sci U S A* 87, 4707-4711.
- Vermeulen, W., Bergmann, E., Auriol, J., Rademakers, S., Frit, P., Appeldoorn, E., Hoeijmakers, J. H., and Egly, J. M. (2000). Sublimiting concentration of TFIID transcription/DNA repair factor causes TTD-A trichothiodystrophy disorder. *Nat Genet* 26, 307-313.
- Vermeulen, W., Jaeken, J., Jaspers, N. G., Bootsma, D., and Hoeijmakers, J. H. (1993). Xeroderma pigmentosum complementation group G associated with Cockayne syndrome. *Am J Hum Genet* 53, 185-192.
- Vermeulen, W., Scott, R. J., Rodgers, S., Muller, H. J., Cole, J., Arlett, C. F., Kleijer, W. J., Bootsma, D., Hoeijmakers, J. H., and Weeda, G. (1994). Clinical heterogeneity within xeroderma pigmentosum associated with mutations in the DNA repair and transcription gene ERCC3. *Am J Hum Genet* 54, 191-200.
- Verrijzer, C. P., and Tjian, R. (1996). TAFs mediate transcriptional activation and promoter selectivity. *Trends Biochem Sci* 21, 338-342.
- Vettese-Dadey, M., Grant, P. A., Hebbes, T. R., Crane-Robinson, C., Allis, C. D., and Workman, J. L. (1996). Acetylation of histone H4 plays a primary role in enhancing transcription factor binding to nucleosomal DNA in vitro. *Embo J* 15, 2508-2518.
- Vignali, M., Steger, D. J., Neely, K. E., and Workman, J. L. (2000). Distribution of acetylated histones resulting from Gal4-VP16 recruitment of SAGA and NuA4 complexes. *Embo J* 19, 2629-2640.
- Wakasugi, M., and Sancar, A. (1999). Order of assembly of human DNA repair excision nuclease. *J Biol Chem* 274, 18759-18768.
- Wang, E. H., and Tjian, R. (1994). Promoter-selective transcriptional defect in cell cycle mutant ts13 rescued by hTAFII250. *Science* 263, 811-814.

- Wang, X. W., Vermeulen, W., Coursen, J. D., Gibson, M., Lupold, S. E., Forrester, K., Xu, G., Elmore, L., Yeh, H., Hoeijmakers, J. H., and Harris, C. C. (1996). The XPB and XPD DNA helicases are components of the p53-mediated apoptosis pathway. *Genes Dev* *10*, 1219-1232.
- Wang, Z., Buratowski, S., Svejstrup, J. Q., Feaver, W. J., Wu, X., Kornberg, R. D., Donahue, T. F., and Friedberg, E. C. (1995). The yeast TFB1 and SSL1 genes, which encode subunits of transcription factor IIIH, are required for nucleotide excision repair and RNA polymerase II transcription. *Mol Cell Biol* *15*, 2288-2293.
- Wang, Z., Svejstrup, J. Q., Feaver, W. J., Wu, X., Kornberg, R. D., and Friedberg, E. C. (1994). Transcription factor b (TFIIH) is required during nucleotide-excision repair in yeast. *Nature* *368*, 74-76.
- Weeda, G., Rossignol, M., Fraser, R. A., Winkler, G. S., Vermeulen, W., van 't Veer, L. J., Ma, L., Hoeijmakers, J. H., and Egly, J. M. (1997). The XPB subunit of repair/transcription factor TFIIH directly interacts with SUG1, a subunit of the 26S proteasome and putative transcription factor. *Nucleic Acids Res* *25*, 2274-2283.
- Weeda, G., van Ham, R. C., Masurel, R., Westerveld, A., Odijk, H., de Wit, J., Bootsma, D., van der Eb, A. J., and Hoeijmakers, J. H. (1990a). Molecular cloning and biological characterization of the human excision repair gene ERCC-3. *Mol Cell Biol* *10*, 2570-2581.
- Weeda, G., van Ham, R. C., Vermeulen, W., Bootsma, D., van der Eb, A. J., and Hoeijmakers, J. H. (1990b). A presumed DNA helicase encoded by ERCC-3 is involved in the human repair disorders xeroderma pigmentosum and Cockayne's syndrome. *Cell* *62*, 777-791.
- Werten, S., Mitschler, A., Romier, C., Gangloff, Y. G., Thuault, S., Davidson, I., and 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.
- Wilson, C. J., Chao, D. M., Imbalzano, A. N., Schnitzler, G. R., Kingston, R. E., and Young, R. A. (1996). RNA polymerase II holoenzyme contains SWI/SNF regulators involved in chromatin remodeling. *Cell* *84*, 235-244.
- Winkler, G. S., Araujo, S. J., Fiedler, U., Vermeulen, W., Coin, F., Egly, J. M., Hoeijmakers, J. H., Wood, R. D., Timmers, H. T., and Weeda, G. (2000). TFIIH with inactive XPD helicase functions in transcription initiation but is defective in DNA repair. *J Biol Chem* *275*, 4258-4266.
- Xiao, H., Pearson, A., Coulombe, B., Truant, R., Zhang, S., Regier, J. L., Triezenberg, S. J., Reinberg, D., Flores, O., Ingles, C. J., and et al. (1994). Binding of basal transcription factor TFIIH to the acidic activation domains of VP16 and p53. *Mol Cell Biol* *14*, 7013-7024.
- Xie, X., Kokubo, T., Cohen, S. L., Mirza, U. A., Hoffmann, A., Chait, B. T., Roeder, R. G., Nakatani, Y., and Burley, S. K. (1996). Structural similarity between TAFs and the heterotetrameric core of the histone octamer. *Nature* *380*, 316-322.

- Xie, Z., and Price, D. (1997). Drosophila factor 2, an RNA polymerase II transcript release factor, has DNA-dependent ATPase activity. *J Biol Chem* 272, 31902-31907.
- Yamaguchi, Y., Takagi, T., Wada, T., Yano, K., Furuya, A., Sugimoto, S., Hasegawa, J., and Handa, H. (1999). NELF, a multisubunit complex containing RD, cooperates with DSIF to repress RNA polymerase II elongation. *Cell* 97, 41-51.
- Yan, M., and Gralla, J. D. (1999). The use of ATP and initiating nucleotides during postrecruitment steps at the activated adenovirus E4 promoter. *J Biol Chem* 274, 34819-34824.
- Yee, A., Nichols, M. A., Wu, L., Hall, F. L., Kobayashi, R., and Xiong, Y. (1995). Molecular cloning of CDK7-associated human MAT1, a cyclin-dependent kinase-activating kinase (CAK) assembly factor. *Cancer Res* 55, 6058-6062.
- Yokoi, M., Masutani, C., Maekawa, T., Sugasawa, K., Ohkuma, Y., and Hanaoka, F. (2000). The xeroderma pigmentosum group C protein complex XPC-HR23B plays an important role in the recruitment of transcription factor IIIH to damaged DNA. *J Biol Chem* 275, 9870-9875.
- Yoon, H., Miller, S. P., Pabich, E. K., and Donahue, T. F. (1992). SSL1, a suppressor of a HIS4 5'-UTR stem-loop mutation, is essential for translation initiation and affects UV resistance in yeast. *Genes Dev* 6, 2463-2477.
- Young, R. A. (1991). RNA polymerase II. *Annu Rev Biochem* 60, 689-715.
- Zawel, L., and Reinberg, D. (1995). Common themes in assembly and function of eukaryotic transcription complexes. *Annu Rev Biochem* 64, 533-561.
- Zehring, W. A., Lee, J. M., Weeks, J. R., Jokerst, R. S., and Greenleaf, A. L. (1988). The C-terminal repeat domain of RNA polymerase II largest subunit is essential in vivo but is not required for accurate transcription initiation in vitro. *Proc Natl Acad Sci U S A* 85, 3698-3702.
- Zhang, J., and Corden, J. L. (1991a). Identification of phosphorylation sites in the repetitive carboxyl-terminal domain of the mouse RNA polymerase II largest subunit. *J Biol Chem* 266, 2290-2296.
- Zhang, J., and Corden, J. L. (1991b). Phosphorylation causes a conformational change in the carboxyl-terminal domain of the mouse RNA polymerase II largest subunit. *J Biol Chem* 266, 2297-2302.

Liste des publications

Jawhari, A. Uhring, M. Crucifix, C. Schultz, P. Poterszman, Egly, J.M & Moras, D (2002) Expression of FLAG fusion proteins in insect cells: Application to the multi-subunit transcription/DNA repair factor TFIIH. *Protein Expression and Purification*, 2002, Vol. 24.

Jawhari, A. Lainé, J.P. Dubaele, S. Lamour, V. Coin, F. Poterszman, A. Moras, D & Egly, J.M. (2002) p52 mediates XPB function within the Transcription/Repair Factor TFIIH. *The Journal of Biological Chemistry*, 2002, Vol. 277, No. 35.

¹Gervais, V. ¹Lamour, V. ¹**Jawhari, A.** Frindel F, Wasielewski E, Dubaele S, Egly JM, Thierry JC, Kieffer B, Poterszman A. TFIIH contains a PH domain involved in DNA nucleotide excision repair. *Nature Structural & Molecular Biology*, 2004, Jul;11(7):616-22. (¹ co-premier auteurs).

Jawhari, A. Boussert S, Lamour V, Atkinson RA, Kieffer B, Poch O, Potier N, van Dorsselaer A, Moras D, Poterszman A. Domain architecture of the p62 subunit from the human transcription/repair factor TFIIH deduced by limited proteolysis and mass spectrometry analysis. *Biochemistry*. 2004 Nov 16;43(45):14420-30

Jawhari, A., Uhring, M., De carlo, S. Moras, D. Schultz, P & Poterszman, A. Oligomeric state and molecular structure of human TFIIH. *EMBO. J, soumis*.

Posters :

Jawhari, A., Gervais. V, Lamour. V, Uhring. M, Schultz. P, Egly. JM, Poterszman. A & Moras. D. *Journée Rencontre des Boursiers de l'ARC, 14 Octobre 2003.*

Boussert. S., **Jawhari. A.**, Poterszman. A, Potier. N & Van Dorsselaer. A. *Spectrométrie de masse et analyse protéomique, Septembre, 2003.*

Jawhari, A., Lainé, J.P., Uhring, M., Poterszman, A., Egly, J.M. et Moras, D. *5th Transcription meeting, Heidelberg, Août 2002.*

Jawhari, A., Lainé, J.P., Uhring, M., Poterszman, A., Egly, J.M. et Moras, D. *NMR and Crystallography School, ERICE Sicile, Juin 2001.*

Fribourg, S., **Jawhari, A.**, Kellengerger, E., Rogniaux, H., Potersman, A., Schultz, P., Egly, J.M & Moras, D. *EMBL transcription meeting, Août, 2000.*