

Université de Montréal

Étude fonctionnelle des domaines structuraux des sous-unités
catalytiques de l'ARN polymérase II humaine

par

Dania Baali

Département de biochimie

Faculté de médecine

Mémoire présenté à la Faculté des études supérieures
en vue de l'obtention du grade de M.Sc.
en biochimie

Juillet, 2005

© Dania Baali, 2005



w

4

U58

2005

v.153

AVIS

L'auteur a autorisé l'Université de Montréal à reproduire et diffuser, en totalité ou en partie, par quelque moyen que ce soit et sur quelque support que ce soit, et exclusivement à des fins non lucratives d'enseignement et de recherche, des copies de ce mémoire ou de cette thèse.

L'auteur et les coauteurs le cas échéant conservent la propriété du droit d'auteur et des droits moraux qui protègent ce document. Ni la thèse ou le mémoire, ni des extraits substantiels de ce document, ne doivent être imprimés ou autrement reproduits sans l'autorisation de l'auteur.

Afin de se conformer à la Loi canadienne sur la protection des renseignements personnels, quelques formulaires secondaires, coordonnées ou signatures intégrées au texte ont pu être enlevés de ce document. Bien que cela ait pu affecter la pagination, il n'y a aucun contenu manquant.

NOTICE

The author of this thesis or dissertation has granted a nonexclusive license allowing Université de Montréal to reproduce and publish the document, in part or in whole, and in any format, solely for noncommercial educational and research purposes.

The author and co-authors if applicable retain copyright ownership and moral rights in this document. Neither the whole thesis or dissertation, nor substantial extracts from it, may be printed or otherwise reproduced without the author's permission.

In compliance with the Canadian Privacy Act some supporting forms, contact information or signatures may have been removed from the document. While this may affect the document page count, it does not represent any loss of content from the document.

Université de Montréal
Faculté des études supérieures

Ce mémoire intitulé :
Étude fonctionnelle des domaines structuraux des sous-unités
catalytiques de l'ARN polymérase II humaine

présenté par :
Dania Baali

a été évalué par un jury composé des personnes suivantes :

Jacques Archambault, IRCM
président-rapporteur

Benoît Coulombe, IRCM
directeur de recherche

Marc Therrien, IRIC
membre du jury

RÉSUMÉ

L'ARN polymérase II (ARN pol II) est l'enzyme qui catalyse la synthèse de tous les ARN messagers (ARNm) cellulaires. Avec l'assistance des facteurs généraux de transcription (FGTs) TFIIA, TFIIB, TFIID, TFIIE, TFIIF et TFIIH, l'ARN pol II reconnaît et lie les promoteurs géniques, ouvre la double-hélice d'ADN dans la région du site d'initiation +1, catalyse la polymérisation des rNTPs en utilisant l'un des brins d'ADN comme matrice, et se déplace le long de l'unité génique en allongeant l'ARN en synthèse. La compréhension du mécanisme catalytique de l'ARN pol II a connu une grande avancée suite à la publication de structures cristallographiques à haute résolution de l'enzyme seule ou en élongation d'un ARNm sur une matrice d'ADN. Ces structures ont permis de proposer des fonctions pour plusieurs éléments structuraux des sous-unités catalytiques de l'enzyme, Rpb1 et Rpb2. Notre laboratoire étudie la relation entre la structure et la fonction des composantes de la machinerie transcriptionnelle au cours des diverses étapes de la transcription. Afin de déterminer les rôles de certains éléments structuraux dans le processus transcriptionnel, nous avons procédé à leur mutagenèse et à la caractérisation des mutants obtenus à l'aide d'essais fonctionnels. Cette étude a permis de déterminer que dans plusieurs cas, les mutations affectent le repliement de la sous-unité catalytique et la rendent incapable de s'assembler et de former le complexe de l'ARN pol II à 12 sous-unités. Par contre, parmi les mutants capables de s'assembler et de former correctement le complexe de polymérase à 12 sous-unités, la plupart sont bloqués aux étapes précoces du mécanisme de transcription. De plus, nous montrons que la boucle *switch 2* est impliquée dans la spécificité de l'ARN pol II à se lier à l'ADN dans le complexe pré-transcriptionnel. Finalement, nous avons identifié un rôle pour la boucle *zipper* au niveau de l'ouverture de la bulle de transcription et de la séparation des brins d'ADN en amont du site actif.

Mots Clés : Transcription, ARNP II, facteurs généraux de transcription, complexe de pré-initiation, TAP, protéomique

ABSTRACT

RNA polymerase II (RNAP II) is the enzyme responsible for the synthesis of all the messenger RNA (mRNA) in the cell. In conjunction with the general transcription factors (GTFs) TFIIA, TFIIB, TFIID, TFIIE, TFIIIF and TFIIH, RNAP II recognizes and binds to gene promoters, opens the DNA double-helix at the initiation site +1, catalyzes the polymerization of ribonucleotides using one of the DNA strands as a template, and moves along the gene unit as the nascent mRNA is being elongated. The understanding of the catalytic mechanism of RNAP II has greatly improved after the publication of high resolution crystal structures of a free enzyme or one that is elongating an mRNA on a DNA template. These structures have suggested functions for many structural elements of the catalytic subunits of the enzyme, Rpb1 and Rpb2. Our laboratory studies the relationship between the structure and the function of the components of the transcriptional machinery during the various stages of the transcription reaction. In order to determine the roles of some structural elements in the transcription process, we have mutated them and characterized the resulting mutants in functional assays. This study provided evidence that in many cases, mutations affect the folding of the catalytic subunits and render them incapable of assembling into 12 subunit RNAP II complex. However, among the mutants capable of assembly and formation of the RNAP II complex, most show an early defect prior to transcription initiation. Furthermore, we show that the *Switch 2* loop is implicated in the binding specificity of the RNAP II for DNA in the pre-initiation complex. Finally, we have identified a role for the *zipper* loop in the opening of the transcription bubble and the separation of the DNA strands upstream of the active site.

Key Words : Transcription, RNA polymerase II, general transcription factors, pre-initiation complex, TAP-tag, proteomics

TABLE DES MATIÈRES

RÉSUMÉ	ii
ABSTRACT	iii
TABLES DES MATIÈRES	iv
LISTE DES TABLEAUX	vi
LISTE DES FIGURES	vii
LISTE DES ABRÉVIATIONS	ix
REMERCIEMENTS	xi
CHAPITRE 1 : INTRODUCTION	1
1.1 Structure d'un gène de classe II	1
1.2 Promoteur et séquences régulatrices	2
1.2.1 Les éléments du promoteur basal	2
1.2.2 Éléments régulateurs	3
1.2.3 La chromatine et son remodelage	4
1.2.3.1 Le complexe SAGA	7
1.2.3.2 Le Médiateur	9
1.3 Les facteurs généraux de transcription	12
1.3.1 TFIID	12
1.3.1.1 TBP	13
1.3.1.2 Les facteurs associés à TBP (TAFs)	14
1.3.1.3 Protéines apparentées à TBP (TRFs)	16
1.3.1.4 Structure du facteur TFIID	17
1.3.2 TFIIA	18
1.3.3 TFIIB	21
1.3.4 TFIIF	24
1.3.5 TFIIE	26
1.3.6 TFIIH	27
1.3.7 TFIIS	30
1.4 Étapes de la réaction transcriptionnelle	33
1.4.1 La formation du complexe de pré-initiation (PIC)	34
1.4.2 Formation du complexe ouvert, initiation et dégagement du promoteur	38
1.4.3 Élongation	39
1.4.3.1 Fidélité, spécificité et translocation	41
1.4.3.2 Modèle d'élongation	42
1.4.4 Terminaison, recyclage de la polymérase, réinitiation et régulation	43
1.5 Architecture de l'ARN pol II	45

1.5.1	Réaction de polymérisation	50
1.5.2	Structure du site actif de l'ARN pol II	51
1.5.3	Un seul site actif pour deux activités	53
1.5.4	Ouverture de la bulle de transcription et chemin de sortie de l'ARN	54
1.5.5	Le sous-complexe Rpb4-Rpb7	58
1.5.6	Le domaine C-terminal de Rpb1	59
1.5.7	Cristallographies de l'ARN pol II avec les FGTs	61
1.6	Objectifs et hypothèses	64
CHAPITRE 2 : MATÉRIELS ET MÉTHODES		67
2.1	Construction des plasmides codant pour la sous-unité Rpb1 de l'ARN pol II fusionnée à l'étiquette TAP	67
2.2	Le système d'induction	69
2.3	Transfection stable de cellules EcR	70
2.4	Purification TAP des complexes contenant l'ARN pol II humaine	72
2.5	Identification des protéines par spectrométrie de masse	75
2.6	Purification de facteurs généraux de transcription	75
2.7	Essai de transcription	75
2.8	Essai de retardement sur gel	77
2.9	Essai d'élongation	77
CHAPITRE 3 : RÉSULTATS		79
3.1	Purification par affinité d'un complexe contenant l'ARN pol II humaine fusionnée à l'étiquette TAP	79
3.2	Le complexe de l'ARN pol II fusionné à l'étiquette TAP peut être recruté à un promoteur <i>in vitro</i>	82
3.3	Le complexe de l'ARN pol II fusionné à l'étiquette TAP permet l'initiation correcte de la transcription <i>in vitro</i>	83
3.4	Purification et analyse fonctionnelle de variantes de l'ARN pol II humaine générée par la mutagenèse dirigée des régions <i>lid</i> et <i>zipper</i>	83
3.5	Purification et analyse fonctionnelle d'une variante humaine de l'ARN pol II générée par la mutagenèse dirigée des régions <i>switch</i> 1 et 2 de Rpb1	88
CHAPITRE 4 : DISCUSSION		90
4.1	Purification simple et efficace de l'ARN pol II humaine dans des conditions natives	90
4.2	Analyse structure-fonction du centre catalytique de l'ARN pol II	91
CONCLUSION ET PERSPECTIVES		97
BIBLIOGRAPHIE		99

LISTE DES TABLEAUX

Tableau I :	Les FGTs et les coactivateurs reliés à l'ARN pol II de <i>S. cerevisiae</i> ...	36
Tableau II :	Facteurs d'élongation associés à l'ARN pol II	40
Tableau III :	Noms des mutants et sommaire des expériences réalisées sur les complexes d'ARN pol II humaine portant l'étiquette TAP sur les sous-unités Rpb1 et Rpb2 mutées	86

LISTE DES FIGURES

Figure 1: Éléments du promoteur basal	3
Figure 2: Structure du nucléosome	5
Figure 3: Structure du complexe SAGA	9
Figure 4: Structure modulaire du Médiateur	12
Figure 5: Structure de TBP	14
Figure 6: Une représentation schématique de TFIIB	22
Figure 7: Organisation des domaines des deux plus grosses sous-unités de TFIIF chez la levure	25
Figure 8: Caractéristiques structurales de TFIIE humain	26
Figure 9: Un modèle pour le mécanisme d'ouverture du promoteur par TFIIF	29
Figure 10: Topologie conservée des complexes ARN polymérase avec les facteurs de clivages respectifs	32
Figure 11: Étapes de la transcription par l'ARN pol II	35
Figure 12: Mécanismes d'initiation et de ré-initiation de l'ARN pol II	44
Figure 13: Architecture des sous-unités de l'ARN polymérase	45
Figure 14: Structure cristallographique de l'ARN polymérase II	48
Figure 15: Modèle du site actif « ajustable » de l'ARN pol II	51
Figure 16: Structure de l'ARN pol II et son interaction avec le sous-complexe Rpb4/Rpb7	53
Figure 17: Structure de l'ARN pol II complète	54
Figure 18: Conservation des compartiments du sillon de l'ARN pol II	56
Figure 19: Construction des plasmides codant pour la sous-unité Rpb1 de l'ARN pol II fusionnée à l'étiquette TAP	68
Figure 20: Système d'induction à l'ecdysone	69
Figure 21: Méthode de purification par affinité en tandem (TAP)	71
Figure 22: Protocole de la méthode TAP	74
Figure 23: Essais fonctionnels	76
Figure 24: Purification du complexe de l'ARN pol II fusionné à TAP	80
Figure 25: Rpb2, Rpb4, Rpb7, Rpb11, TFIIB et RAP30 fusionnés à l'étiquette TAP ont mené à la purification d'un complexe à 17 sous-unités	82

Figure 26: Le complexe de l'ARN pol II fusionné à l'étiquette TAP peut être recruté à un promoteur <i>in vitro</i>	84
Figure 27: Essai de transcription <i>in vitro</i>	84
Figure 28: Purification et analyses des sous-unités du complexe d'ARN pol II humain portant l'étiquette TAP sur Rpb1 et les mutations Sw2-348 et Zip	87
Figure 29: Essai d'élongation	88
Figure 30: Séparation de l'hybride ARN-ADN et voie de sortie de l'ARN	93

LISTE DES ABRÉVIATIONS

AdML	<i>Adenovirus major late promoter</i>
Ada	<i>Adaptor</i>
ADN	Acide déoxyribonucléique
ARN	Acide ribonucléique
ARN pol II	ARN polymérase II
ARNm	ARN messenger
ATP	Adénosine triphosphate
BRE	<i>TFIIB recognition element</i>
CBP	<i>Calmodulin-binding protein</i>
CCD	<i>Charged cluster domain</i>
CDK	<i>Cyclin-dependent kinase</i>
CTD	<i>C-terminal domain</i>
DPE	<i>Downstream promoter element</i>
ELL	<i>Eleven-nineteen lysine-rich leukemia</i>
FACT	<i>Facilitates chromatin transcription</i>
FGT	Facteur général de transcription
FRET	<i>Fluorescence resonance energy transfer</i>
H2	Histone 2
HAT	Histone acétyltransférase
HDAC	<i>Histone deacetylase complex</i>
HFD	<i>Histone-fold domain</i>
HSV	<i>Herpes simplex virus</i>
IgG	Immuno-globuline G
Inr	<i>Initiator</i>
KDa	Kilodalton
MTE	<i>Motif ten element</i>
NER	<i>Nucleotide excision repair</i>
PAGE	<i>Polyacrylamide gel electrophoresis</i>
Pb	Paire de bases
PCAF	<i>p300/CBP-associated factor</i>

PIC	<i>Pre-initiation complex</i>
P-TEFb	<i>Positive-transcriptional elongation factor b</i>
Pu	Purine
Py	Pyrimidine
RAP	<i>RNA polymerase II-associated protein</i>
Rpb	<i>RNA polymerase B</i>
SAGA	<i>Spt/Ada/Gcn5 acetyltransferase complex</i>
SC	Syndrome de Cockayne
SDS	Sodium dodecyl sulphate
SWI	<i>Mating-type switching</i>
TAF	<i>TBP-associated factor</i>
TAP	<i>Tandem affinity peptide</i>
TBP	<i>TATA box-binding protein</i>
TCR	<i>Transcription-coupled repair</i>
TEV	<i>Tobacco etch virus</i>
TFII	<i>RNA polymerase II transcription factor</i>
TFTC	<i>TBP-free TAF_{II}-containing</i>
TLF	<i>TBP-like factor</i>
TRF	<i>TBP-related factor</i>
TTD	Trichothiodystrophie
VP16	<i>Viral protein 16</i>
XP	Xeroderma Pigmentosum

REMERCIEMENTS

Le présent projet de maîtrise a bénéficié du soutien de plusieurs personnes qu'il me fait plaisir de souligner ici. Je tiens à remercier tout particulièrement mon directeur de recherche, Benoit Coulombe, de m'avoir permis de parfaire mes connaissances dans son laboratoire, et aussi pour son encadrement, sa disponibilité et son soutien. Je tiens également à remercier Marie-France Langelier pour mon apprentissage pratique, son aide et ses conseils, ainsi que tous les membres du laboratoire pour l'aide, l'appui et le support moral qu'ils m'ont accordés tout au long de ma maîtrise. Finalement, un remerciement infini à mes parents Hana et Yasser pour leur assistance soutenue et leur encouragement de manière indéfectible, ainsi qu'à mes sœurs et frères Lama, Roula, Ammar et Yahya; et à mon mari, Assad, qui est une source d'encouragement et de motivation pour moi.

Chapitre 1 : Introduction

L'ARN polymérase II (ARN pol II) est l'enzyme responsable de la transcription des gènes codant pour les protéines dans les cellules eucaryotes. Les dernières évaluations proposent que les cellules humaines contiennent 20 à 25 mille gènes. Certains de ces gènes sont exprimés de façon constitutive, comme les gènes responsables des fonctions métaboliques communes à toutes les cellules. D'autres ne sont exprimés que lorsque la cellule entre dans une certaine voie de différenciation ou lorsqu'elle atteint son état de différenciation terminal, tandis que certains sont exprimés en réponse aux conditions environnementales de la cellule. Une régulation pointue de la transcription par l'ARN pol II est donc un élément important pour la maturation et le développement cellulaire, et est crucial pour la réponse aux stimuli et autres signaux externes. Plusieurs mécanismes permettent aux cellules eucaryotes de réguler l'expression des gènes, soit en modulant le niveau de transcription des gènes, en altérant le taux de maturation des chaînes d'ARNm dans le noyau, en modifiant la stabilité ou le degré de dégradation des ARNm, ou en contrôlant l'efficacité avec laquelle les ribosomes traduisent les chaînes d'ARNm en polypeptides.

1.1 Structure d'un gène de classe II

Les gènes de classe II codent pour les protéines et possèdent de façon générale, les éléments suivants : des exons (qui en majorité codent pour des séquences polypeptidiques), des introns (régions internes enlevées de l'ARNm durant la maturation), un site d'initiation de la transcription, un promoteur (le promoteur basal situé dans les 50 paires de bases (pb) en amont du site d'initiation et un promoteur plus élargi qui peut s'étendre à plus de 200 pb) et des régions régulatrices éloignées telles que des « enhancers » et des « silencers ». Les gènes adjacents sont souvent séparés par une région isolante, « insulator », qui prévient la propagation des signaux d'activation ou de répression d'un gène à son entourage.

1.2 Promoteur et séquences régulatrices

Une portion importante de l'information régulant le niveau d'expression des gènes est codée par le promoteur et les séquences régulatrices éloignées (Smale 1997, Butler 2002, Hochheimer 2003). Le promoteur basal dirige l'initiation adéquate de la transcription et possède typiquement une longueur de ~50 nucléotides. Plusieurs sous-régions fonctionnelles, nommées éléments du promoteur, sont caractéristiques. Cependant, ces éléments ne sont pas présents universellement dans tous les promoteurs de base et la plupart des promoteurs contiennent un ou plusieurs de ces éléments. Ces éléments représentent des sites de liaisons pour des sous-unités de la machinerie transcriptionnelle et servent à orienter de manière asymétrique la transcription unidirectionnelle à partir du promoteur.

1.2.1 Les éléments du promoteur basal

Parmi les éléments du promoteur basal, on retrouve la boîte TATA, l'initiateur (Inr), l'élément de réponse du facteur TFIIB (BRE) et l'élément en aval du promoteur (DPE) (figure 1). La boîte TATA (Goldberg 1979) est située ~26-31 nucléotides en amont du site d'initiation et possède une séquence consensus de TATA(A/T)A(A/T). Elle représente le site de reconnaissance et de liaison de la protéine TBP. L'élément BRE possède une séquence consensus CGCC et est situé immédiatement en amont de près de 12% des boîtes TATA (Burke 1996). L'Inr renferme le site d'initiation de la transcription (+1) et possède une séquence consensus PyPyAN(T/A)PyPy chez l'humain, où l'adénosine est typiquement désignée +1. Le DPE (Burke 1996) est centré autour de la position +30 et possède une séquence consensus PuG(A/T)CGTG. Le facteur général de transcription (FGT) TFIID lie coopérativement les motifs de l'Inr et du DPE (Kaufmann 1994, Purnell 1994, Verrijzer 1994, Chalkley 1999, Smale 2003).

Il est fort probable que plusieurs éléments du promoteur basal demeurent à découvrir. Une analyse statistique de 205 promoteurs de *Drosophila* a d'ailleurs suggéré que ~31% des promoteurs ne contiennent ni boîte TATA ni motif DPE (Kutach 2000). Plus

récemment, une analyse informatisée de plus de 2000 promoteurs de *Drosophila* a identifié un motif sur-représenté comme un nouvel élément du promoteur basal (Ohler 2002). Cette séquence, nommée MTE (pour « motif ten element ») et située ~20-30 nucléotides en aval du site d'initiation de la transcription, peut induire la transcription par l'ARN pol II en synergie avec l'Inr mais indépendamment de la boîte TATA ou de l'élément DPE (Lim 2004).

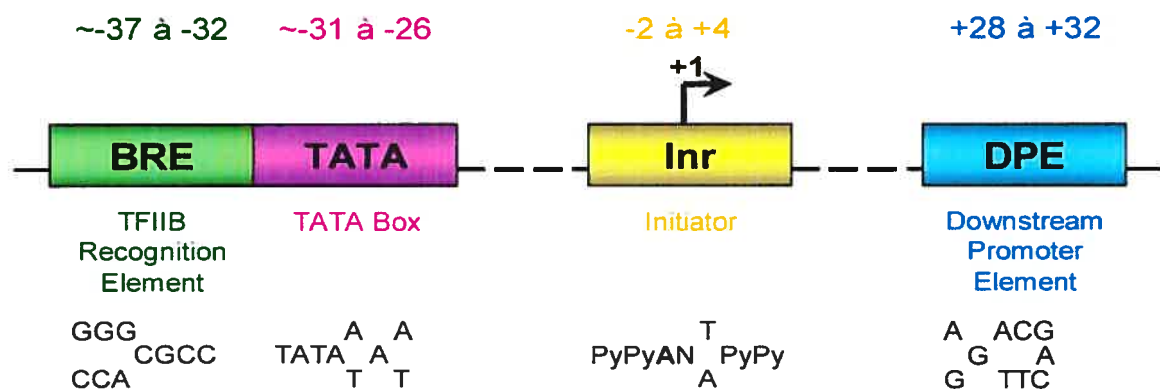


Figure 1 : Éléments du promoteur basal

Certains des motifs du promoteur basal qui participent à la transcription par l'ARN pol II sont illustrés. Chacun de ces éléments se retrouve seulement chez un sous-ensemble de promoteurs et les promoteurs spécifiques peuvent ainsi contenir quelques uns ou tous ces motifs. Le BRE est une extension en amont de certaines boîtes TATA. Le DPE requière la présence d'un Inr et est situé précisément à +28 à +32 par rapport au nucléotide A+1 de l'Inr.

Image adapté de: Butler 2002

1.2.2 Éléments régulateurs

Certains facteurs de transcription se lient à des régions d'ADN situées à des milliers de pb du gène qu'ils contrôlent, et leur liaison peut augmenter ou diminuer le niveau de transcription du gène. Les « enhancers » et les « silencers » sont des éléments d'ADN qui peuvent être situés en amont, en aval ou même à l'intérieur du gène qu'ils contrôlent.

Les protéines liant les « enhancers », grâce à leurs domaines d'activation, peuvent stimuler la transcription de gènes à des milliers de pb par la formation d'une boucle dans l'ADN (Knight 1991). Par contre, lorsque des facteurs de transcription lient les « silencers », une condensation de la chromatine réprime l'expression du gène en question.

Étant donné que les enhancers et silencers agissent à de très grandes distances, la cellule doit utiliser un moyen d'empêcher leur action sur certains autres gènes qui pourraient être situés sur la même région du chromosome. Un élément d'ADN, nommé « insulator », répond à ce besoin. Les insulators sont des séquences d'ADN aussi petites que 42 pb situées entre les enhancers et les promoteurs ou entre les silencers et les promoteurs de gènes ou groupes de gènes adjacents. Lorsque liés par la protéine correspondante, leur fonction consiste à empêcher un gène d'être influencé par l'activation ou la répression de ses voisins. Par exemple, les insulators des vertébrés sont actifs lorsque liés par une protéine à 11 doigt de zinc nommée CTCF (pour « CCCTC binding factor », une séquence de nucléotides découverte dans la plupart des insulators) (Fillipova 2005).

1.2.3 La chromatine et son remodelage

Le matériel génétique des eucaryotes est empaqueté dans un complexe protéique d'histones pour former la chromatine, une matrice sur laquelle a lieu la transcription, la réplication, la recombinaison, ainsi que la réparation. La chromatine est organisée en régions plus ou moins condensées, nommées respectivement hétérochromatine et euchromatine, dispersées à travers le génome. Elle devient empaquetée en chromosomes individuels durant la division cellulaire. L'hétérochromatine est inaccessible aux facteurs protéiques et n'est pas transcrite. Cependant, les régions d'euchromatine sont plus accessibles et l'expression d'un gène requiert la perturbation sélective de la structure d'hétérochromatine par des enzymes de modification et de remodelage de la chromatine (Workman 1998).

L'unité de base de la chromatine est le nucléosome (figure 2), qui consiste en près de 165 pb d'ADN enroulé en 2 tours autour d'un octamère d'histones, composé de dimères de H2A, H2B, H3 et H4. L'ADN enroulé autour de la surface de l'octamère d'histones n'est accessible aux protéines régulatrices que du côté opposé au nucléosome ou jusqu'à ce que la chromatine soit remodelée. Les extrémités N-terminales des histones H3 et H4 sont accessibles et certains enzymes modifient ces « queues » pour promouvoir le déplacement des nucléosomes et le déroulement de l'ADN.

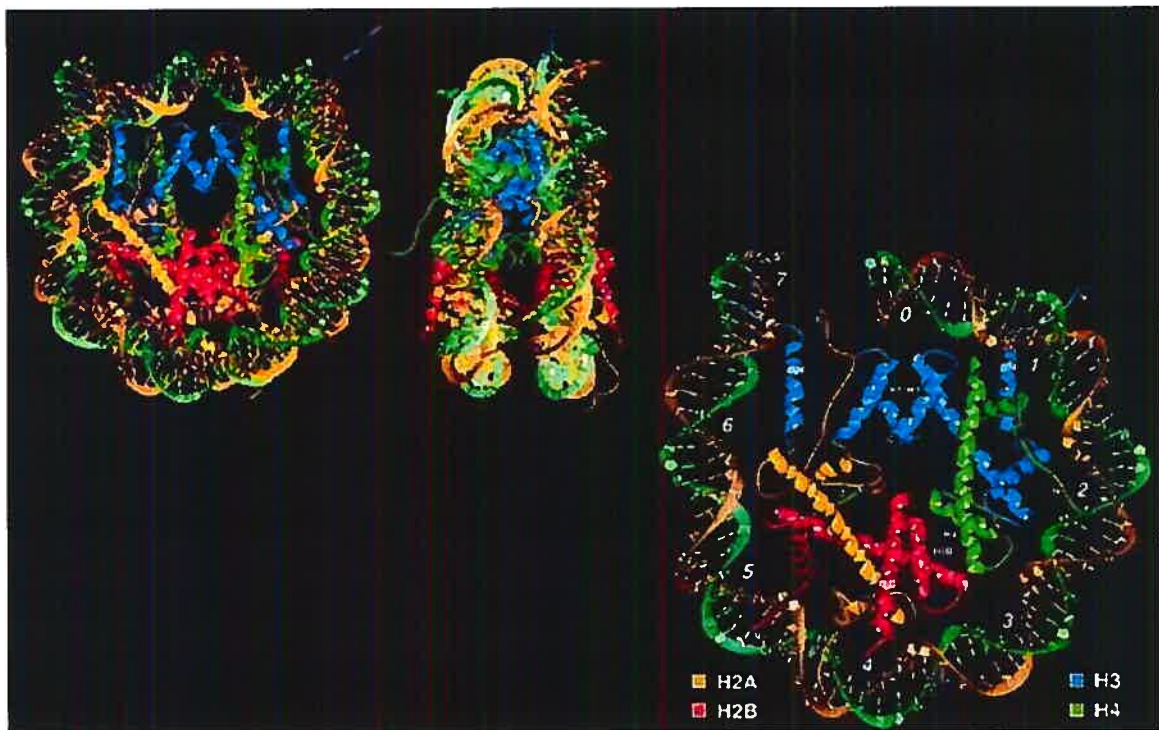


Figure 2 : Structure du nucléosome.

4 paires de monomères protéiques forment l'octamère d'histone. Chaque nucléosome est composé d'un octamère central, d'une histone H1 additionnelle et de l'ADN liant un nucléosome à son voisin (~54 pb). L'histone H1 lie le nucléosome central ainsi que l'ADN entre deux nucléosomes. L'ADN s'enroule vers la gauche autour de l'octamère et ces tours sont topologiquement équivalents à un super-enroulement négatif.

Image tirée de: Luger 1997

Certaines lysines sont acétylées, méthylées ou ubiquitinées, certaines arginines sont méthylées et des sérines sont phosphorylées (Davie 1998). La sumoylation, processus par lequel une molécule SUMO (pour « small ubiquitin-related modifier ») est ajoutée sur certaines lysines de la queue de l'histone H4, régule la répression par des interactions avec d'autres protéines de liaison et de modification de la chromatine (Shiio 2003). Toutes ces modifications forment un assortiment de marqueurs de l'état local du matériel génétique appelé le code d'histones. Certaines modifications entraînent des effets sur l'expression des gènes, par exemple, les queues d'histones sont acétylées dans les régions d'euchromatine qui sont actives en transcription. L'ajout du groupement acétyl permet de neutraliser les charges positives des lysines et ainsi réduire les fortes interactions entre les nucléosomes et l'ADN chargé négativement. La méthylation peut aussi neutraliser la charge des lysines, mais son effet est ambivalent : la méthylation de la Lys-4 de H3 est associée aux gènes actifs tandis que la méthylation de la Lys-9 de H3 est associée aux gènes inactifs. Ainsi, ces modifications ne font pas qu'affecter les interactions du nucléosome avec l'ADN, mais permettent la reconnaissance de ces régions par des coactivateurs et des facteurs impliqués dans la régulation de la transcription. D'ailleurs plusieurs domaines protéiques sont connus pour lier les groupements acétyl ou méthyl. Par exemple, les protéines possédant des bromodomaines interagissent avec les queues d'histones acétylées tandis que les protéines possédant des chromodomaines interagissent avec les queues d'histones méthylées.

L'ADN des eucaryotes est donc ordinairement réfractaire à la transcription dû à son organisation en nucléosomes (Kornberg 1999b). L'octamère d'histone, autour duquel l'ADN est enroulé, interfère avec le mécanisme de transcription, et les nucléosomes non-acétylés servent donc de répresseurs transcriptionnels. Ils aident à assurer l'inactivité de tous les gènes eucaryotes, sauf ceux dont la transcription a été stimulée par des mécanismes spécifiques de régulation positive (Boeger 2005). Les promoteurs actifs sont associés à des histones modifiés soit par acétylation, méthylation ou phosphorylation (Wu 2001a). L'équipe du Dr Roger Kornberg a récemment démontré que l'activation transcriptionnelle ne produit pas une forme altérée de la chromatine, mais qu'elle représente plutôt un processus dynamique dans lequel les nucléosomes sont

continuellement enlevés des promoteurs activés et sont réassemblés (Boeger 2003). Une étude ultérieure avait suggéré que les nucléosomes glissent et s'éloignent de la boîte TATA et du site d'initiation (Lomvardas 2001) mais il s'avère plus exact de penser que les nucléosomes sont enlevés par désassemblage (Boeger 2004) impliquant des complexes de remodelage de la chromatine, tels que SWI/SNF (Cairns 1994, Peterson 1995) et RSC (Cairns 1996). RSC pourrait désassembler le nucléosome en transférant l'octamère d'histone sur une autre molécule d'ADN (Lorch 1999).

La transcription dépend de la présence d'activateurs qui recrutent d'autres facteurs pour contrer la répression due à la chromatine. La réponse aux activateurs requiert la présence de complexes coactivateurs, parmi lesquels se trouvent le complexe SAGA et le Médiateur. La découverte des coactivateurs est venue entre autres suite à l'observation que les systèmes *in vitro* ne supportent pas la transcription activateur-dépendente dans un système ne contenant que la machinerie transcriptionnelle basale et des activateurs. De ce fait, le rôle des coactivateurs pourrait être de faire le pont entre les activateurs et la machinerie transcriptionnelle. Les coactivateurs sont donc des complexes multi-protéiques recrutés aux promoteurs par des activateurs gène-spécifiques pour faciliter l'initiation de la transcription en intégrant les signaux des activateurs et répresseurs liés aux éléments régulateurs du/des gènes en question. Le moment de leur recrutement varie selon la composition et l'arrangement des éléments basals et des sites de liaisons des activateurs sur les promoteurs. Certains coactivateurs, tel que le complexe SAGA, possèdent une activité histone acétyltransférase (HAT) qui permet de faciliter l'accès de la machinerie transcriptionnelle au promoteur ainsi que les interactions protéine-protéine en neutralisant les charges protéiques. D'autres coactivateurs sont responsables du remodelage de la chromatine qui facilite également l'accès de la machinerie transcriptionnelle au promoteur. Le Médiateur, quant à lui, relie l'activateur directement au promoteur basal et aux FGTs (Lewis 2003).

1.2.3.1 Le complexe SAGA

La régulation transcriptionnelle chez les eucaryotes est liée à la dynamique de la chromatine. Le complexe HAT de levure SAGA (Spt-Ada-Gcn5) est un cofacteur multi-protéique de la transcription par l'ARN pol II. SAGA a d'abord été identifié comme un complexe HAT (Grant 1997). Ce complexe de 1.8 MDa comprend des produits de plusieurs classes de gènes, notamment, les protéines Ada (Ada1, Ada2, Ada3, Gcn5 et Ada5), l'ensemble de protéines Spt reliées à TBP, un sous-ensemble des facteurs associés à TBP (TAFs) et le produit du gène essentiel TRA1 (Martinez 2002). Gcn5 est la sous-unité catalytique du complexe et est régulées par Ada2 et Ada3 (Balasubramanian 2002, Sterner 1999). Plusieurs homologues des complexes HAT Gcn5 ont été caractérisés chez la Drosophile et l'humain tels que TFTC (pour « TBP-free TAF-containing complex »), PCAF (pour « p300- et CBP-associated factor ») et le complexe STAGA, tous possédant Gcn5, les protéines Ada, Spt, TAF ainsi que des protéines semblables aux TAFs, en plus de TRRAP, un homologue de Tra1, (pour « transformation / transcription domain-associated protein ») (Martinez 2002).

La structure du complexe SAGA, telle que révélée par microscopie électronique à une résolution de 30Å, a mené à une grande avancée dans la compréhension de la fonction du complexe (Wu 2002). Le complexe SAGA est formé de 5 domaines modulaires avec un diamètre de 70 à 100Å (figure 3). La superposition des structures de SAGA de levure et de TFTC humain indique qu'en plus de leur composition de sous-unités semblables, les deux complexes possèdent un niveau élevé de conservation structurelle en taille et en forme (Wu 2002, Brand 1999a). Chaque domaine du complexe SAGA possède une fonction spécialisée et distincte. Le domaine I contient Tra1 qui forme la surface d'interaction avec l'activateur. Les domaines II, III et IV contiennent la protéine TAF5 ainsi que plusieurs TAFs possédant des motifs HFD (pour « histone-fold domain ») repliés de manière semblable aux histones. TAF5 semble jouer un rôle central dans l'architecture de SAGA, comparable à son rôle dans le facteur TFIID (qui sera discuté plus en détail dans les prochaines sections). En plus de l'activité HAT qui y est localisée, deux sous-unités possédant des bromodomaines, Gcn5 et Spt7, ont été détectées au domaine III. Le domaine V représente le module interagissant avec la protéine TBP et sa structure semble être variable et plus flexible; la sous-unité Spt8 n'étant pas présente chez

tous les complexes SAGA, elle pourrait s'associer de manière dynamique comme sous-unité régulatrice (Pray-Grant 2002, Sterner 2002).

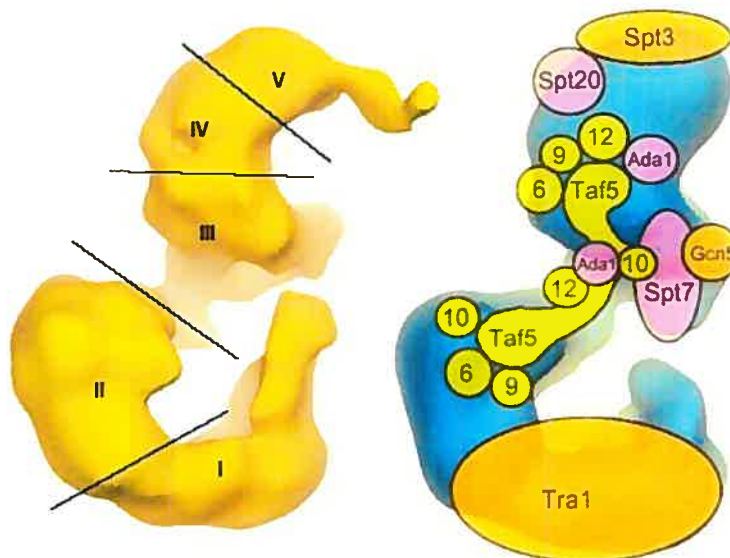


Figure 3 : Structure du complexe SAGA

Représentation tri-dimensionnelle de la surface du complexe SAGA avec le domaine flexible. Le complexe mesure $\sim 18 \times 28$ nm et comprend 5 domaines séparés par les barres noires.

Images tirées de: Timmers 2004, Wu 2004

1.2.3.2 Le Médiateur

L'initiation de la transcription par l'ARN pol II se divise en deux étapes. D'abord, l'activité de certains complexes de modification et de remodelage de la chromatine est requise pour supprimer la répression induite par l'état condensé (Urnov 2002). Ensuite, un PIC contenant la polymérase et les FGTs est recruté par les activateurs aux séquences de reconnaissance et s'assemble au promoteur. Cependant, les interactions directes entre les différents activateurs et les FGTs ne semblent pas suffisantes pour l'activation de la transcription. Le Médiateur a été identifié comme un complexe qui intègre les signaux venant des activateurs et répresseurs liés sur les régions régulatrices d'un gène et les relie aux FGTs situés au promoteur basal. Le Médiateur a d'abord été identifié chez la levure,

puis chez les eucaryotes supérieurs, et sa présence est requise pour l'expression basale et régulée des gènes de classe II (Flanagan 1991, Kim 1994a, Koleske 1994, Lee 1999, Mittler 2001).

Les premières indications expérimentales de la présence d'un complexe Médiateur humain datent de 1996 lorsque le groupe de Roeder (Fondell 1996) a isolé un complexe multiprotéique coactivateur du récepteur de l'hormone thyroïdienne T3 (nommé TRAP pour « Thyroid hormone Receptor coActivator Protein ») qui a été identifié plus tard comme étant le Médiateur humain. L'implication du Médiateur dans la transcription activée ou réprimée des cellules mammifères est reflétée par la multitude d'études par lesquelles ce complexe a été identifié. Le complexe TRAP est nécessaire à l'activation de la transcription *in vitro* à partir d'un promoteur contenant l'élément de réponse de T3 (Fondell 1996). Un autre complexe protéique interagissant avec le récepteur de la vitamine D3 (VDR), DRIP (pour « vitamin D3 Receptor Interacting Proteins »), est requis pour l'activité transcriptionnelle de VDR *in vitro* (Rachez 1998). ARC (pour « activator-recruited cofactor ») a été identifié comme un complexe qui augmente l'activation transcriptionnelle par SREBP-1a, VP16 et la sous-unité p65 de NF-kappaB sur des matrices avec chromatine (Naar 1999). La caractérisation des complexes TRAP, DRIP, ARC et du Médiateur humain a révélé qu'ils possèdent tous une composition identique de sous-unités (Rachez 1999, Malik 2000a). Un ensemble de complexes similaires, qui semblent correspondre à un sous-module du Médiateur, a également été isolé et requiert la présence de plusieurs cofacteurs (Malik 2000a). Deux des plus petits complexes, le CRSP (pour « cofactor required for Sp1 ») et le Médiateur de souris ont été isolés par fractionnement biochimique (Jiang 1998, Ryu 1999a) et aident à l'activation transcriptionnelle seulement en présence des TAFs (Malik 2000a, Ryu 1999b). Un autre petit complexe Médiateur, nommé NAT pour « negative regulator of activated transcription », démontre quant à lui un effet d'inhibition sur la transcription *in vitro* (Sun 1998).

La composition du Médiateur de levure est clairement distincte des complexes semblables chez d'autres eucaryotes. Initialement, 24 sous-unités avaient été purifiées et

identifiées comme complexe Médiateur requis pour la réponse de l'ARN pol II aux activateurs gènes-spécifiques (Kim 1994b, Thompson 1993). Par contre, seules huit des sous-unités du Médiateur possèdent des homologues hautement conservés chez la souris et l'humain (Bourbon 2004) et ce faible taux de conservation au niveau de la séquence primaire a mené à l'idée que le complexe Médiateur humain varie en structure et en fonction du complexe de levure. Une autre interprétation concevable est que le Médiateur fonctionne comme interface entre des protéine régulatrices qui évoluent rapidement et la machinerie basale conservée de transcription (Balciunas 2003). Ainsi, les 8 protéines du Médiateur « basal » conservé se retrouvent chez toutes les cellules eucaryotes et sont responsables d'interagir avec la machinerie transcriptionnelle; par contre, les sous-unités responsables de l'interaction avec les activateurs et répresseurs spécifiques aux gènes sont moins conservées. Selon les études structurales par microscopie électronique, les Médiateurs isolés et purifiés à partir de cellules de levure et de souris possèdent une structure compacte. Par contre, lorsque l'ARN pol II est présente, ces Médiateurs adoptent une forme étendue et enlacent la polymérase globulaire. La structure étendue révèle 4 sous-modules distincts du Médiateur, dont une région Cdk8, une région tête, un centre et une région queue (figure 4) (Bourbon 2004). Le module Cdk8 possède une activité protéine kinase et est impliqué dans la répression transcriptionnelle, notamment par phosphorylation de plusieurs activateurs transcriptionnels les rendant inactifs (Chi 2001). Par exemple, le complexe CAK (pour « cdk-activating kinase) de TFIIH qui phosphoryle normalement le CTD de la polymérase et aide à l'initiation et au dégagement du promoteur (Holstege 1998, Liao 1995), peut être phosphorylé et inhibé par Cdk8. Les contacts directs sont formés entre l'ARN pol II et les régions tête et centre du Médiateur. En effet, ces régions contiennent toutes les sous-unités conservées et essentielles du complexe Médiateur, et ces sous-unités sont positionnées sur la surface interagissant avec la polymérase. Les sous-unités non-conservées de la tête et du centre, ainsi que la plus grosse partie du Médiateur, la région étendue de la queue, ne contactent pas la polymérase (Asturias 1999) et sont placées vers l'extérieur pour interagir et recevoir les signaux des protéines régulatrices (Bourbon 2004). Tel que prédit, la région queue du complexe humain diffère significativement dans sa structure du complexe de levure et pourrait contenir des sous-

unités impliquées dans l'interaction avec les activateurs et répresseurs, telles que TRAP220, TRAP230 et TRAP240 (Malik 2000b).

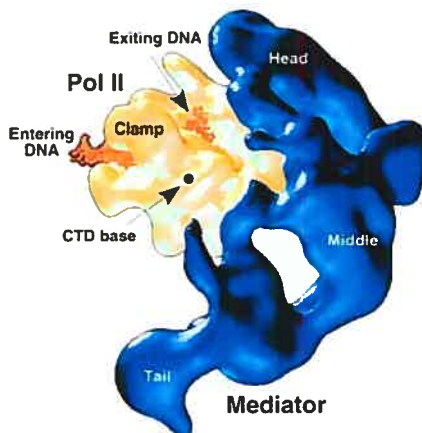


Figure 4 : Structure modulaire du Médiateur

La structure modulaire du complexe Médiateur facilite des interactions multiples avec les facteurs de régulation (représentés par VP16 et E1A) et permet les transitions entre les différentes formes du Médiateur.

Image tirée de: Blazek 2005

1.3 Les facteurs généraux de transcription (FGTs)

La reconnaissance du promoteur basal par la machinerie transcriptionnelle est essentielle pour le positionnement et l'assemblage adéquat d'un PIC contenant l'ARN pol II et les FGTs TFIIB, TFIID, TFIIE, TFIIF et TFIIH (Buratowski 1994, Buratowski 1989, Kornberg 1999a). La transcription par l'ARN pol II commence avec la liaison de facteurs régulateurs sur un élément d'ADN près du site d'initiation (tableau I, page 36). Ces facteurs permettent l'accès de la machinerie transcriptionnelle au promoteur basal soit par le recrutement de complexes de remodelage ou de modification de la chromatine, soit par des interactions avec l'ARN pol II ou avec le complexe Médiateur.

1.3.1 TFIID

1.3.1.1 TBP

Le facteur TBP (« TATA-binding protein ») lie spécifiquement l'élément de la boîte TATA du promoteur basal (Sawadogo 1985). Plusieurs études cristallographiques ont été effectuées sur cette protéine, soit seule ou intégrée aux complexes TBP-ADN, TBP-ADN-TFIIA et TBP-ADN-TFIIB (Nikolov 1997). TBP est une protéine bipartite, avec un domaine C-terminal très conservé et une région N-terminale variable en longueur et en séquence, et qui n'est pas requise pour la transcription chez certaines lignées de levure (Burley 1996). Le domaine C-terminal est conservé à près de 80% entre la levure et l'humain et adopte une structure en selle (figure 5) avec une face concave hydrophobe interagissant avec l'ADN et une face convexe qui interagit avec les autres facteurs impliqués dans la transcription (Burley 1996, Liu 1998, Kamada 2001). La structure cristallographique du complexe TBP-ADN démontre que TBP se lie au sillon mineur de l'élément TATA, désappariant ainsi près du tiers d'un tour et introduisant une courbure prononcée de 80Å dans l'ADN en direction du sillon majeur (Kim 1994a, Kim 1993). Cette courbure est médiée par l'insertion de 2 phénylalanines entre les 2 premières et les 2 dernières pb de la séquence TATAAA. Le sillon mineur est ainsi élargi et s'ajuste à la surface concave de la selle. L'enroulement particulier de l'ADN autour du nucléosome module l'orientation du sillon mineur de l'élément TATA, ce qui affecte la capacité de TBP à se lier dans le contexte du nucléosome (Imbalzano 1994). La courbure de l'ADN pourrait servir à rapprocher l'activateur du promoteur basal et du PIC, ou à stabiliser l'assemblage de ce dernier (Burley 1996).

La protéine TBP est un monomère de 27 kDa qui a d'abord été identifié à partir d'extraits de cellules de mammifères fractionnés pour avoir une activité indispensable, requise avec d'autres fractions, pour entreprendre l'initiation précise de la transcription par l'ARN pol II *in vitro* (Matsui 1980, Samuels 1982). Contrairement au complexe partiellement purifié, la protéine TBP recombinante est cependant incapable de supporter la transcription activée (Pugh 1990). À l'aide d'anticorps anti-TBP, plusieurs polypeptides associés à TBP, nommés TAFs, ont été purifiés et identifiés (Dymlacht 1991, Tanese 1991). Il a été découvert également que TBP participe à l'initiation de la transcription

des 3 ARN pol et que ce sont ces protéines associées qui confèrent aux complexes contenant TBP la spécificité pour chaque système de polymérase (Hernandez 1993).

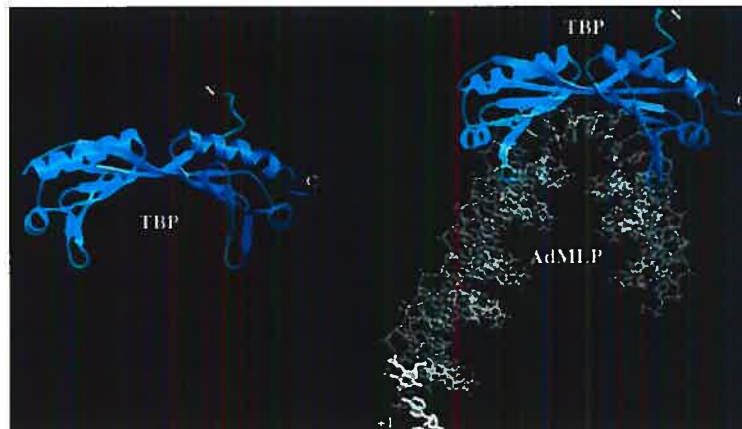


Figure 5 : Structure de TBP

Structures en ruban de TBP seul et en complexe avec l'ADN, TFIIB et TFIIA

Image tirée de: Nikolov 1997

1.3.1.2 Les facteurs associés à TBP (TAFs)

Présentement, 13 polypeptides ont été isolés dans la famille des facteurs TAFs (pour « TBP-associated factors »), associés à TBP chez l'humain, la Drosophile et la levure (Tora 2002, Matangasombut 2004). L'analyse des séquences primaires des TAFs identifiés a révélé un niveau de conservation élevé. Plusieurs motifs se trouvent présents dans les homologues des 3 espèces étudiées, tels que la répétition WD40 chez TAF5 et le domaine de repliement des histones présents chez plus de la moitié des TAFs. Par contre, certains motifs se retrouvent seulement chez les TAFs d'eucaryotes supérieurs. Par exemple, lorsque comparé au TAF_{II}250 humain, celui de la levure est homologue dans sa partie N-terminale (Reese 1994) mais manque un domaine C-terminal acide et deux copies du motif bromodomaine qui interagit avec les queues d'histones acétylées (Marmorstein 2001). Certains TAFs sont présents chez plusieurs complexes liés à la transcription tandis que d'autres sont spécifiques au facteur TFIID, notamment TAF1, qui possède une activité HAT (Martinez 2002).

La découverte de domaines structuraux partagés entre les histones et les TAFs a été faite à partir d'observation de similarités de séquences importantes entre TAF9 et TAF6 de la *Drosophile*, et le domaine de repliement des histones H3 et H4 respectivement. Une structure co-cristallographique de ces deux domaines montre qu'ils interagissent et forment un hétérotetramère, semblable à celui de H3 et H4 (Xie 1996). L'identification d'une homologie à H2B chez TAF12 (Hoffmann 1996) et d'une homologie à H2A chez TAF4 (Gangloff 2000) a mené à l'hypothèse que ces 4 TAFs peuvent former un complexe semblable à l'octamère d'histones et ayant des propriétés moléculaires apparentées (Matangkasombut 2004, Selleck 2001). Ainsi, l'octamère formé par les TAFs pourrait également enrouler le promoteur d'ADN d'une manière semblable au nucléosome (Hoffmann 1997, Oelgeschlager 1996).

Une autre diversification de l'activité des TAFs est observée chez la protéine humaine BTAF1 et son homologue chez la levure, Mot1p (Pereira 2003). Ces deux protéines appartiennent à la famille d'ATPase semblable à SNF2 et possèdent un domaine ATPase à leur extrémité C-terminale (van der Knaap 1997, Chicca 1998). Une proportion importante de TBP est associée avec chacune de ces protéines régulatrices pour former un complexe nommé B-TFIID (Poon 1994, Timmers 1994). Malgré que B-TFIID soit aussi efficace que TFIID pour supporter la transcription *in vitro*, plusieurs différences sont observées : la transcription médiée par B-TFIID ne répond pas à un certain nombre d'activateurs, B-TFIID démontre une activité ATPase et ne forme pas un complexe stable avec l'ADN (Timmers 1994). La plus importante activité des protéines BTAF1 et Mot1p est de dissocier TBP de l'élément TATA en utilisant l'énergie de l'hydrolyse d'ATP, expliquant ainsi leur rôle dans l'inhibition transcriptionnelle (Auble 1994, Gumbs 2003). Par contre, d'autres résultats indiquent que BTAF1 et Mot1p utilisent cette activité pour réguler positivement la transcription en mobilisant TBP sur le promoteur d'ADN (Andrau 2002, Geisberg 2002). Des modèles récents proposent que Mot1p, comme moteur à ATPase, se sert de l'ADN en amont de la boîte TATA comme levier pour dissocier TBP de l'ADN (Darst 2001, Darst 2003, Gumbs 2003).

Un sous-ensemble des protéines TAFs se trouve présent dans plusieurs complexes reliés à la transcription, tels que les complexes HAT SAGA et P/CAF chez la levure et l'humain respectivement (Grant 1998, Ogryzko 1998). Un autre complexe HAT purifié à partir de cellules humaines, contenant 7 polypeptides TAFs différents mais sans TBP, a été isolé et nommé TFTC (pour « TBP-free TAF-containing complex ») (Wieczorek 1998). Ce complexe est capable de remplacer TFIID dans un essai de transcription *in vitro*, indiquant que, soit les TAFs ont la capacité de reconnaître le promoteur basal et d'initier la transcription sans TBP, soit que le complexe contient un facteur pouvant se substituer à TBP (Brand 2001, Matangkasombut 2004). De plus, la régulation des 2 complexes pourrait être mécaniquement différente puisque TFIID se localise à l'élément TATA tandis que SAGA se lie sur des sites d'activation plus en amont (Larschan 2001, Bhaumik 2001).

1.3.1.3 Protéines apparentées à TBP (TRFs)

Des variantes de TBP ont été caractérisées chez les eucaryotes multicellulaires et représentent des formes tissus- ou stade-spécifiques. En effet, 3 protéines TRF1 (pour « TBP-related factor 1 »), TLF (pour « TBP-like factor »)/TRF2 et TRF3 ont été identifiées et caractérisées (Dantonel 1999, Crowley 1993, Rabenstein 1999, Teichmann 1999, Moore 1999, Hansen 1997, Persengiev 2003). L'homologie entre ces protéines est limitée au domaine C-terminal conservé. TRF2 présente 60% d'homologie et 40% d'identité avec TBP, mais TRF1 et TRF3 sont encore plus proches de TBP (Davidson 2004). Chez tous ces facteurs, les sites d'interaction avec les facteurs TFIIA et TFIIB sont conservés, et TRF1 et TRF3 se fixent sur les séquences TATA. En revanche, TRF2 ne fixe pas les éléments TATA (Dantonel 1999) et aucun site de fixation de TRF2 n'est actuellement connu.

TRF1 existe seulement chez la Drosophile où il est exprimé entre autres pendant l'embryogenèse puis, de façon spécifique, dans le système nerveux central et les cellules germinales chez l'adulte (Crowley 1993, Hansen 1997, Holmes 2000). La fonction de TRF2 a été étudiée *in vivo* chez plusieurs organismes et est nécessaire à l'expression d'un

sous-ensemble de gènes lors de l'initiation de la transcription dans l'embryon précoce et à la progression du cycle cellulaire (Dantonel 2000, Kaltenbach 2000a, Veenstra 2000, Muller 2001). Dans certaines cellules lymphoïdes de poulet, TRF2 est cytoplasmique et est importé dans le noyau uniquement lors de la phase G2 du cycle cellulaire ou après l'exposition au stress (UV) (Shimada 2003). Chez les mammifères, TRF2 a une fonction beaucoup plus spécialisée et son expression coïncide avec l'augmentation de l'expression d'autres facteurs de transcription, tels que l'ARN pol II et TFIIB (Martianov 2002, Zhang 2001). De plus, TRF2 pourrait être impliqué dans le maintien de l'intégrité de l'hétérochromatine centromérique dans les spermatides (Martianov 2002). La fonction de TRF3 est moins caractérisée que TRF1 et TRF2. Son domaine C-terminal possède une forte homologie avec TBP et se lie à l'élément de la boîte TATA (Persengiev 2003). Dans les cas où TBP est inactivé, un ralentissement du cycle cellulaire et une augmentation de l'apoptose ont été observés dans plusieurs expériences (Um 2001, Johnson 2003). Il est possible que TBP soit nécessaire à l'activation *de novo* de la transcription, mais que la réinitiation soit supportée par TRF3 (Oelgeschlager 2002).

1.3.1.4 Structure du facteur TFIID

Le facteur TFIID est composé de plusieurs sous-unités (Albright 2000, Burley 1996) dont au moins 13 sont conservées à travers l'évolution (Gangloff 2001). Des études structurales des sous-complexes TBP et TAFs permettent d'élucider l'organisation et le mécanisme d'action de TFIID. Une structure à basse résolution du complexe entier a été obtenue par microscopie électronique (Andel, III 1999, Brand 1999b). Le complexe adopte une structure en « selle de cheval », formant 3-4 lobes liés par des régions flexibles. Il a été suggéré que la surface concave représente le site de liaison à l'ADN et que l'ADN est accommodé dans le canal central. TBP a été localisé au lobe central, à l'endroit où TFIIA et TFIIB se lient également. Une détermination plus précise des positions des sous-unités individuelles sera essentielle pour comprendre la fonction du complexe TFIID.

Le complexe TFIID est maintenu par plusieurs interactions protéine-protéine de nature TBP-TAF et TAF-TAF (Burley 1996, Gangloff 2001). Il a été suggéré que la plus grosse sous-unité, TAF1, joue un rôle de protéine d'échafaudage pour le complexe puisqu'elle interagit directement avec TBP et plusieurs autres TAFs. Tel que mentionné plus tôt, plusieurs TAFs contiennent un motif de repliement homologue aux protéines histones (« histone fold domain ») et qui permet des contacts serrés et importants entre les histones et l'ADN (Luger 1997). La distribution spatiale de TAF5, TAF6 et TAF10 chez SAGA et TFIID est semblable et a été utilisée pour aligner les deux complexes et comparer leurs fonctions (Leurent 2002). L'alignement suggère que le sillon de 4 nm de largeur chez TFIID, impliqué dans la liaison de l'ADN, est semblable au sillon formé par les domaines II, III et IV de SAGA. Par contre, la localisation de leur activité HAT et des facteurs portant les bromodomains diffère, suggérant que les deux complexes interagissent avec la chromatine de manière différente (Leurent 2002).

1.3.2 TFIIA

D'abord isolé comme FGT, TFIIA détient un rôle plutôt controversé dans l'initiation de la transcription. Le facteur TFIIA humain est composé de 3 sous-unités : α (35 kDa), β (19 kDa) et γ (12 kDa) (Ozer 1994, DeJong 1993). Les sous-unités α et β sont codées par le même gène et produites par clivage post-traductionnel d'un précurseur (DeJong 1993). Chez la levure, TFIIA est composé de 2 sous-unités codées par 2 gènes différents, *TOA1* et *TOA2* (Ranish 1996). La région N-terminale du polypeptide produit par *TOA1* est homologue à la sous-unité α humaine, et la région C-terminale est homologue à la sous-unité β , tandis que *TOA2* code pour un polypeptide homologue à γ (Ozer 1994, DeJong 1995). Le clivage post-traductionnel du précurseur humain α/β a été prouvé non-essentiel à son activité (Sun 1994).

TFIIA n'est pas essentiel dans la réaction de transcription basale *in vitro*, mais semble jouer un rôle stimulateur dans plusieurs systèmes *in vivo* (Ozer 1994, Ma 1996, Sun 1994, Coulombe 1992, Cortes 1992). TFIIA lie TBP, augmente son affinité pour la boîte

TATA et stabilise le complexe ADN-TFIID (Coulombe 1992, Cortes 1992, Buratowski 1989, Pugh 2000). TFIIA peut également déplacer certains répresseurs, incluant Dr1-DRAP1/NC2, topoisomérase 1, HMG1 et la sous-unité Mot-1 du complexe TFIID, indiquant que TFIIA est impliqué dans l'anti-répression (Auble 1994, Inostroza 1992, Meisterernst 1991, Merino 1993). Le facteur TFIIA humain joue également un rôle dans la transcription activée, puisqu'il est requis pour le fonctionnement de certains activateurs (Ozer 1994, Ma 1996, DeJong 1995, Sun 1994, Wang 1992, Kang 1995, Chi 1996). Par exemple, TFIIA lie l'activateur Zta et médie son rôle stimulateur dans la liaison de TFIID à la boîte TATA (Ozer 1994). De la même façon, TFIIA stimule l'activation de la transcription par les activateurs Sp1, VP16 et NTF1 (Kobayashi 1995). Les activateurs liant TFIIA, VP16 et Zta, stimulent l'assemblage d'un complexe TFIIA-TFIID-promoteur, ce qui concorde avec leurs rôles *in vivo* dans la transcription activée (Kobayashi 1995). La fonction de TFIIA dans l'anti-répression et l'activation a été divisée et associée avec les sous-unités distinctes du facteur (Ma 1996). Les sous-unités β et γ sont essentielles pour l'anti-répression, tandis que α est essentielle pour l'activation.

Des études de photo-pontage démontrent que TFIIA contacte l'ADN du promoteur non seulement sur la boîte TATA et en amont à la position -40, mais également dans la région +26 dans le contexte d'un PIC (Langelier 2001). Étant donné la petite taille de TFIIA (67kDa) et les deux positions extrêmes du promoteur contactées par TFIIA, il a été proposé que TFIIA se situe près du croisement de l'ADN lorsqu'enroulé autour de la polymérase (Langelier 2001).

Le rôle de TFIIA dans la transcription n'est pas complètement défini mais les études mentionnées plus haut suggèrent qu'il agit comme co-facteur dans l'activation transcriptionnelle. TFIIA interagit directement avec TBP, RAP74, TFIIE56 et TFIIE34, et stimule la transcription basale dans un essai *in vitro* (Langelier 2001). Donc, considérant qu'un activateur tel que Gal4-VP16, dont l'activité requiert la présence de TFIIA, puisse stimuler l'ouverture du promoteur près du site d'initiation (Wang 1992), il

est possible d'envisager un mécanisme d'activation dans lequel TFIIA agirait comme un pont entre les protéines activatrices et les facteurs TFIIE et TFIIIF (Langelier 2001).

TFIIA lie TBP et facilite sa liaison à l'ADN via plusieurs mécanismes distincts et active ainsi la transcription TBP-dépendente des gènes de classe II. TFIIA facilite la stabilisation de TBP sur la boîte TATA en induisant un changement conformationnel chez TBP (Ranish 1999). De plus, TFIIA agit comme anti-répresseur contre les co-répresseurs qui inhibent la liaison de TBP à l'ADN et à TFIIB (Kim 1995). TBP forme un homodimère inactif en solution, et la dissociation de ce dimère est induite par une interaction directe avec TFIIA (Coleman 1995, Coleman 1997, Taggart 1996, Jackson-Fisher 1999). Puisque certains activateurs transcriptionnels dépendent des interactions TFIIA-TBP et TFIIA-activateurs (Damania 1998, Ozer 1998), TFIIA pourrait être responsable de la sélectivité du promoteur par la machinerie transcriptionnelle grâce à des interactions avec les facteurs régulateurs et TBP.

Comme mentionné plus tôt, les eucaryotes supérieurs possèdent des familles de gènes reliées à TBP, telles que les TRFs. Il a été démontré que des mutations au niveau de TRF2, ou TLP (pour « TBP-like protein ») chez plusieurs espèces causent des anomalies lors du développement accompagnées de baisse du niveau d'expression de certains gènes (Dantonel 2000, Kaltenbach 2000b, Martianov 2002, Muller 2001, Veenstra 2000). Des études chez le poulet ont démontré que les TLP en solution ou liés au promoteur fonctionnent comme des activateurs (Ohbayashi 1999) et affectent les niveaux de transcription de certains gènes reliés au cycle cellulaire (Shimada 2003). La protéine TLP est structurellement similaire à TBP et possède une fonction dans la régulation transcriptionnelle. TLP lie TFIIA *in vivo* mais la fonction de TFIIA sur TLP n'est pas connue. Il a été démontré que TLP est associé à TFIIA et TFIIB, qu'il possède une plus grande affinité pour TFIIA que TBP et que le complexe TLP-TFIIA est beaucoup plus stable que le complexe TBP-TFIIA (Nakadai 2004). Les TLP forment des dimères et des trimères plus stables que le dimère TBP mais ces multimérisations peuvent toujours être inhibées par TFIIA. La localisation de TLP est majoritairement cytoplasmique et est dépendante de sa capacité à lier TFIIA (Nakadai 2004). En effet, lorsque le niveau

nucléaire de TLP est augmenté (Shimada 2003, Martianov 2002) par un mécanisme indépendant de TFIIA, TLP s'associe à TFIIA et inhibe la transcription TBP-dépendente. Lorsque les cellules ne requièrent plus de TLP pour leur activité, les TLP sont exportés du noyau par une association à TFIIA de manière à ne pas s'accumuler et affecter la transcription des gènes cibles.

1.3.3 TFIIIB

L'initiation de la transcription des gènes de classe II requière l'assemblage du PIC sur la boîte TATA du promoteur. TFIIIB joue un rôle crucial dans ce processus en faisant le pont entre TBP et l'ARN pol II. De plus, TFIIIB est la cible d'activateurs transcriptionnels qui augmentent le taux de formation du PIC et assurent sa stabilité. TFIIIB possède une activité acétyltransférase et peut s'auto-acétyler pour réguler la transcription (Choi 2003).

TFIIIB est composé de deux domaines distincts par leur repliement et leur fonction (figure 6A) : un domaine C-terminal replié de manière semblable à deux cyclines (ou « cyclin-like folds ») qui interagit avec la région C-terminale de TBP et l'ADN en amont et en aval de la boîte TATA ; et un domaine N-terminal contenant un doigt de zinc interagissant avec la polymérase, ainsi qu'un domaine hautement conservé de résidus chargés qui modulent la fonction de TFIIIB, le domaine CCD (pour « charged cluster domain ») (Bagby 1995, Nikolov 1995, Zhu 1996). Modulées par le domaine CCD, les extrémités N- et C-terminales de TFIIIB peuvent interagir de manière intramoléculaire et varier entre un état fermé et un état ouvert de la protéine (figure 6B) (Hawkes 2000, Fairley 2002, Zheng 2004). Cette interaction intramoléculaire a suscité beaucoup d'intérêt grâce à son potentiel de constituer un mécanisme de contrôle de la transcription (Reese 2003). Des analyses FRET ont démontré qu'un changement conformationnel chez TFIIIB se produit lors de son assemblage avec le complexe TBP-promoteur et que la fréquence de ce changement augmente en présence d'un activateur (Zheng 2004). TFIIIB peut effectuer des contacts séquence-spécifiques avec l'ADN du promoteur via un motif

hélice-tour-hélice dans son domaine C-terminal qui stabilise le complexe TFIIB-TBP-promoteur (Lagrange 1998). Ces contacts impliquent l'élément de reconnaissance de TFIIB (BRE) situé en amont de la boîte TATA, ainsi qu'une séquence du promoteur en aval de la séquence TATA (Fairley 2002). Le domaine d'activation de la protéine VP16 du virus de l'herpès simplex HSV défait l'interaction avec le BRE en induisant un changement conformationnel de TFIIB (Roberts 1994, Evans 2001). Cependant, les mutants du domaine CCD qui démontrent un défaut d'activation sont capables de s'assembler normalement au promoteur et de former le PIC contenant l'ARN pol II (Glossop 2004).

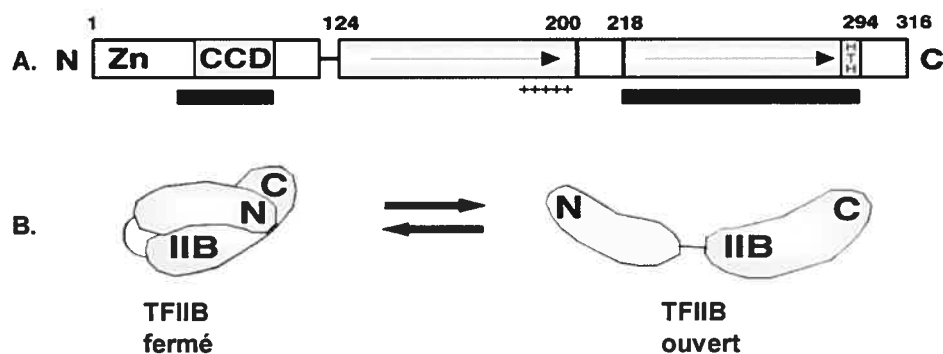


Figure 6 : Une représentation schématique de TFIIB

A. Représentation linéaire de TFIIB. Les deux répétitions directes imparfaites sont montrés par des flèches. Le domaine à doigt de zinc (Zn), le domaine CCD, la région basique (+++++) et le motif hélice-tour-hélice (HTH) sont également indiqués. Les traits pleins indiquent les régions de TFIIB qui sont impliquées dans les interaction intramoléculaires.

B. Les domaines N- et C-terminaux de TFIIB interagissent pour former un équilibre entre l'état ouvert et l'état fermé de la molécule.

Ainsi, le domaine CCD de TFIIB joue un rôle central dans la sélection du site d'initiation de la transcription. Des études de mutagenèse suggèrent que les déterminants du CCD qui servent à moduler l'activation de la transcription et la sélection du site d'initiation se chevauchent, mais sont distinctes (Hawkes 1999). Une mutation ponctuelle peut alors causer une translocation du site d'initiation en amont ou en aval, ou même diminuer la transcription de manière dépendante de la séquence de l'Inr (Fairley 2002). L'analyse cristallographique du complexe TFIIB-ARN pol II révèle que le domaine CCD forme une structure en forme de doigt qui pénètre au site actif de la polymérase à partir du canal de sortie de l'ARN (Bushnell 2004). Il a été postulé que ce domaine en forme de boucle, ou doigt, interagit avec l'ADN au site actif et que sa présence stabilise la liaison du court hybride ADN-ARN à l'ARN pol II (Gnatt 2001, Westover 2004). Ces résultats aident à expliquer l'importance du domaine CCD dans la sélection du site d'initiation et la force du promoteur. Le domaine en doigt de zinc interagit, quant à lui, avec le domaine « dock » de l'ARN pol II. L'interaction de ces deux domaines de TFIIB est donc forte et indépendante d'autres facteurs. Par contre, la localisation des domaines « linker » et « core » dépend du contexte du PIC, puisqu'ils ne sont pas ordonnés dans la structure du complexe ARN pol II-TFIIB mais deviennent ordonnés en présence des autres composantes du PIC, telles que le promoteur (Asturias 2004).

Des études de photo-pontage site-spécifiques ont révélé des interactions considérables entre les domaines CCD, « linker » et « core » de TFIIB et les sous-unités Tfg1 et Tfg2 de TFIIF (Chen 2004). L'interaction du « core » de TFIIB avec Tfg2 près du sillon menant au site actif de l'ARN pol II coïncide avec la localisation de Tgf2 dans la structure du complexe ARN pol II-TFIIF obtenu par cryo-microscopie électronique (Chung 2003). Cette interaction est particulièrement intéressante puisque Tfg2 apparaît comme l'homologue structural du facteur bactérien σ et joue un rôle dans l'orientation de l'ADN du promoteur au site actif. Des interactions entre les domaines CCD et « linker » de TFIIB avec Tfg1 ont également été détectées au site actif, là où les analyses par cryo-microscopie électronique n'avaient détecté aucune densité de Tfg1. Ces résultats suggèrent que TFIIF est entièrement ordonné en présence des autres composantes du PIC (Chen 2004, Asturias 2004).

1.3.4 TFIIF

TFIIF est un hétérotetramère $\alpha_2\beta_2$ formé par les sous-unités RAP74 et RAP30 associées à l'ARN pol II (RAP pour « RNA pol II-associated protein ») (Conaway 1989, Flores 1990, Wang 1994). L'organisation linéaire des 2 plus grosses sous-unités de TFIIF (Tfg1 levure/RAP74 mammifère et Tfg2 levure/RAP30) est montrée à la figure 7A (Chung 2003, Ghazy 2004). Les sous-unités RAP74 et RAP30 sont organisées de manière semblable, ayant des régions N- et C-terminales repliées en domaine globulaire et une région centrale chargée et désordonnée. TFIIF lie l'ARN pol II en solution et aide son recrutement au promoteur en réduisant son affinité pour les sites d'ADN non-spécifiques et en augmentant son affinité pour le PIC (Conaway 1991, Flores 1991, Killeen 1992). Dans le PIC formé sur le promoteur majeur tardif d'adénovirus, RAP74 forme certains contacts spécifiques entre l'ADN et RAP30, TFIIE et l'ARN pol II (Forget 1997, Robert 1996, Robert 1998) et de ce fait, aide l'ADN à s'enrouler autour de la polymérase et des FGTs (figure 7B) (Robert 1998, Forget 2004).

Plusieurs études ont démontré que TFIIF peut stimuler l'activité de Fcp1, une phosphatase du domaine C-terminal de la sous-unité Rpb1 de la polymérase (Chambers 1995). De plus, il a été suggéré que TFIIF stimule l'élongation en supprimant les pauses transcriptionnelles (Bengal 1991, Izbán 1992a, Tan 1994). TFIIF est un facteur général d'initiation et d'élongation qui, *in vitro*, stimule l'élongation 5 à 10 fois, atteignant des taux sur des matrices sans chromatine semblables aux taux observés *in vivo*, où les matrices sont complexées en chromatine (Ren 1999, Lei 1999, Renner 2001). TFIIF stimule la réaction de polymérisation et accélère les étapes plus lentes qui suivent normalement la polymérisation, notamment la translocation et le relâchement du pyrophosphate (Nedialkov 2003). TFIIF aide donc le complexe d'élongation à s'engager dans la réaction de synthèse et stabilise ce complexe après la translocation (Zhang 2004). TFIIF lie la surface externe de l'ARN pol II, resserre la pince qui retient l'hybride ARN-ADN et optimise les conditions de polymérisation durant l'élongation (Zhang 2003). De plus, le facteur d'élongation TFIIS (voir plus bas) coopère avec TFIIF pour coordonner

les réactions de synthèse d'ARN et d'arrêt, menant à l'accélération des taux d'entrée et de sortie des pauses (Zhang 2003).

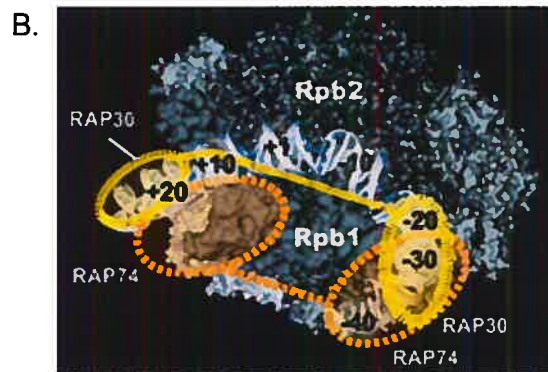
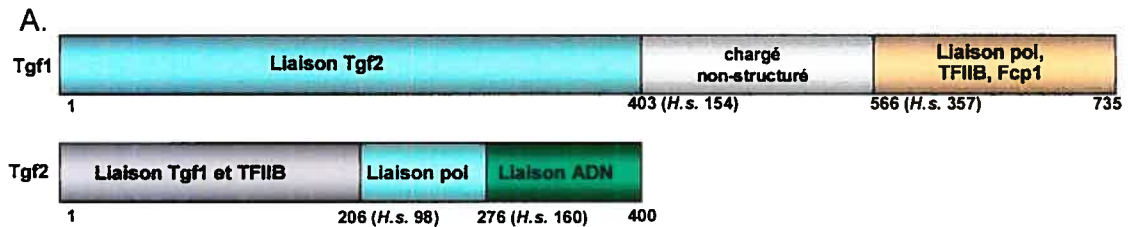


Figure 7 : Organisation des domaines des deux plus grosses sous-unités de TFIIF chez la levure

A. Tgf1 et Tgf2, homologues de Rap74 et Rap30 respectivement chez l'humain, sont organisés de manière semblable, avec des domaines N- et C-terminaux globulaires

B. Localisation des sous-unités de TFIIF dans le PIC selon des résultats de photopontage. Deux régions distinctes du promoteur ont été pontées par chaque sous-unité (cercles pointillés). Malgré que ces résultats suggèrent que deux copies de TFIIF sont présentes, la possibilité que ce facteur possède au moins deux domaines reliés par un "linker" désordonné ne peut être niée.

Image tirée de: Forget 2004

1.3.5 TFIIE

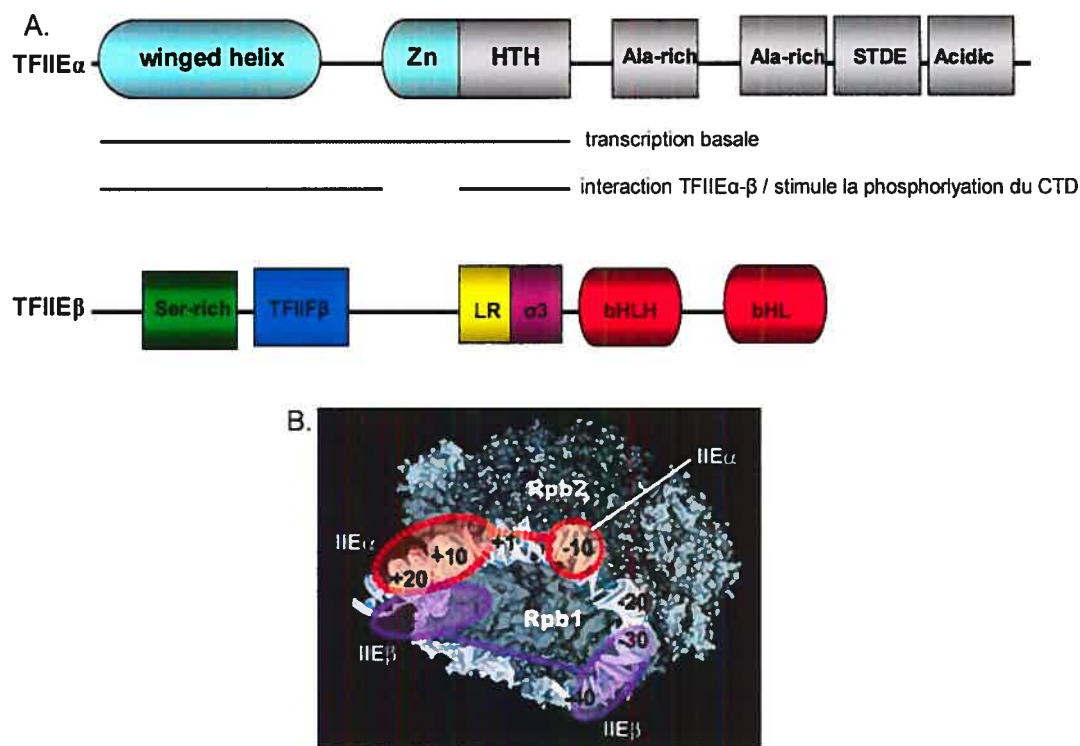


Figure 8 : Caractéristiques structurales de TFIIE humain

A. TFIIE α possède une région en doigt de zinc (Zn), un motif hélice-tour-hélice (HTH), une région riche en alanines (Ala), une région riche en sérines, thréonines, aspartates et glutamates (STDE) et une région acide. Les régions de TFIIE α requises pour la transcription basale, l'interaction avec TFIIE et la stimulation du domaine CTD kinase sont indiquées par les lignes noires. TFIIE β possède une région riche en sérines, une région semblable à la région de TFIIF (RAP30) capable de lier l'ADN double brin, un motif de leucines répété, une région semblable au sous-domaine 3 du facteur σ bactérien, un motif basique d'hélice-boucle-hélice et un motif semblable à ce dernier mais omettant la deuxième hélice.

B. Localisation des sous-unités de TFIIE dans le PIC selon des résultats de photopontage. Deux régions distinctes du promoteur ont été obtenues pour chaque sous-unité (cercles pointillés). Malgré que ces résultats suggèrent que deux copies de TFIIE sont présentes, la possibilité que ce facteur possède au moins deux domaines reliés par un "linker" désordonné ne peut être niée.

Image tirée de: Forget 2004

La facteur TFIIE humain est un hétérotétramère constitué de 2 sous-unités α (57 kDa, 439 acides aminés) et de 2 sous-unités β (34 kDa, 291 acides aminés) (Ohkuma 1990, Ohkuma 1991, Sumimoto 1991, Peterson 1991) qui possèdent plusieurs séquences et motifs structuraux caractéristiques (figure 8A). Malgré que TFIIE ait été caractérisé fonctionnellement, peu d'informations sont disponibles sur sa structure, surtout au niveau de TFIIE α en raison de sa faible solubilité (Ohkuma 1995, Okamoto 1998, Okuda 2000).

La région essentielle à la transcription basale et à la croissance cellulaire est située dans la moitié N-terminale de TFIIE α qui contient également tous les motifs structuraux tels qu'une région riche en leucines et un domaine en doigt de zinc (Ohkuma 1995, Ohkuma 1991). Cette moitié N-terminale démontre une grande similarité avec les homologues de TFIIE α d'archéobactéries récemment identifiés, mais ces derniers ne possèdent pas de région correspondant à la moitié C-terminale de TFIIE α (Bell 2001). La région acide près de l'extrémité C-terminale est la seule à lier directement TFIIH et à posséder un effet stimulateur sur la transcription basale. La sous-unité TFIIE β possède, quant à elle, 4 régions fonctionnelles (figure 8A), notamment, une région N-terminale riche en serines, un noyau central, une structure hélice-boucle-hélice basique et une structure hélice-boucle C-terminale basique (Okamoto 1998, Okuda 2000, Sumimoto 1991). La région centrale lie l'ADN double brin et possède une structure tridimensionnelle semblable à une tête de fourchette (Okuda 2000). TFIIE β lie également l'ADN simple brin via sa région d'hélice-boucle C-terminale basique qui représente également le site de liaison de TFIIIB et RAP30 (figure 8B) (Okamoto 1998, Forget 2004).

Durant l'initiation, TFIIE recrute TFIIH au PIC et stimule ses activités enzymatiques (Lu 1992), puis, lors de la transition de l'initiation à l'élongation, TFIIE joue un rôle direct dans le dégagement du promoteur en régulant les activités kinase et hélicase de TFIIH (discuté plus bas) (Ohkuma 1994, Drapkin 1994).

1.3.6 TFIIH

TFIIH se distingue parmi les autres facteurs de transcription par sa grande taille, ses activités catalytiques et ses multiples rôles fonctionnels (Svejstrup 1996). Composé de 10 sous-unités et d'une masse totale de près de 500 kDa, TFIIH est comparable à l'ARN pol II en terme de taille et de complexité (Giglia-Mari 2004). Les deux plus grosses sous-unités de TFIIH, nommées Ssl2 et Rad3 chez la levure, et XPB et XPD chez l'humain, sont des ATPases et hélicases ADN-dépendentes essentielles pour l'ouverture du promoteur au site actif de la polymérase (Schaeffer 1993, Schaeffer 1994, Guzder 1995). XPB est une hélicase 3'→5' et XPD est une hélicase 5'→3' (Dvir 1996, Moreland 1999, Winkler 2000). XPB joue un rôle important lors de l'ouverture du promoteur et l'initiation, tandis que les deux hélicases contribuent ensemble au dégagement du promoteur. Des analyses structurales tridimensionnelles récentes du complexe TFIIH de l'humain et de la levure ont révélé que 6 sous-unités (XPB, XPD, p62, p52, p44 et p34) forment un complexe basal, ou « core », et adoptent une structure en forme d'anneau. Le centre de l'anneau possède des dimensions lui permettant d'accueillir le passage de l'ADN double brin, mais il n'existe pas de preuve appuyant ce modèle (figure 9) (Chang 2000, Schultz 2000). Trois autres sous-unités (Cdk7, cycline H et MAT1) forment une structure sphérique posée sur cet anneau par la liaison à XPD (Coin 1998, Schultz 2000). Deux des petites sous-unités forment donc une paire de CDK-cycline, où CDK7 dépend de la cycline H pour phosphoryler le domaine C-terminal de l'ARN pol II durant la transition de l'initiation à l'élongation (Feaver 1994, Lu 1992). Les 10 sous-unités de TFIIH sont conservées en terme de structure et de séquence d'acides aminés de la levure à l'humain (Schultz 2000, Chang 2000, Svejstrup 1996). De plus, un complexe intact et complet, nommé « holo » TFIIH a été isolé à partir des deux organismes (Svejstrup 1995, Conaway 1989, Gerard 1991, Flores 1992). En plus de son rôle dans la transcription, la paire cdk-cycline détient, chez les humains, un rôle différent qui consiste à activer d'autres protéines Cdk régulatrices du cycle cellulaire (Tassan 1995). Le complexe basal de TFIIH, quant à lui, fait partie d'un « réparosome » d'ADN responsable de la réparation des erreurs par excision de nucléotides (Svejstrup 1995, Feaver 2000).



Figure 9 : Un modèle pour le mécanisme d'ouverture du promoteur par TFIID

TFIID est en vert et sa sous-unité p89/XPB est en orange.

Image tirée de: Douziech 2000

Des mutations héréditaires affectant le facteur de transcription et de réparation TFIID ont été associées à 3 syndromes photo-hypersensibles, notamment le Xeroderma Pigmentosum (XP), le XP combiné au syndrome de Cockayne (SC) (une maladie neuro-développementale) et la Trichothiodystrophie (TTD) (une maladie semblable au syndrome de Cockayne) (Lehmann 2001). TFIID joue un rôle central dans la transcription des ARN pol I et II, la réparation par excision de nucléotides (NER) et la réparation couplée à la transcription (TCR) (Egly 2001). Le NER corrige un grand nombre de lésions qui induisent la distorsion ou la déformation de l'hélice d'ADN, incluant les lésions induites par les rayons ultra-violet (Hoeijmakers 2001). Le processus de TCR est une voie de réparation qui élimine les blessures classiques réparées par le NER en plus des lésions oxydatives d'ADN, mais ce mécanisme de réparation se limite au brin d'ADN transcrit (de Boer 2000, Le Page 2000). Les caractéristiques photosensibles du XP sont causées par un mécanisme NER défectif, tandis que les caractéristiques du SC et de TTD sont attribuées à une fonction de TCR ou de transcription défectueuses. La plupart des individus touchés par les syndromes associés à TFIID possèdent une mutation chez XPD, ou dans certains cas, chez XPB (Botta 2002). Un troisième gène identifié récemment, TFB5, semble responsable du TTD et est associé à TFIID en tant que dixième sous-unité du complexe (Giglia-Mari 2004).

Un des phénotypes les plus évidents du syndrome de TTD est la réduction dramatique des niveaux cellulaires de TFIID (~30%) (Giglia-Mari 2004). Une surexpression de TFB5 augmente de manière transitoire la quantité de complexes TFIID (mais pas des ARNm

encodant les différentes sous-unités) dans les cellules mutées ou de type sauvage, indiquant que TFB5 est impliqué dans la régulation post-transcriptionnelle. Ces résultats suggèrent un modèle où TFB5 joue un rôle d'échaffaudage pour l'assemblage et le maintien du complexe TFIIH, et son absence pourrait affecter la longévité du complexe et ses activités (Giglia-Mari 2004). Malgré que plusieurs parallèles sur la composition, la structure et la fonction de TFIIH peuvent être tirées entre la levure et l'humain, des différences dans la contribution relative de TFB5 sont observées. Chez la levure, l'absence de TFB5 affecte l'efficacité de transcription et de réparation par NER tandis que chez l'humain, les mutations de TFB5 n'affectent pas sérieusement la transcription (Tirode 1999). De plus, l'absence de TFB5 chez l'humain affecte la stabilité et les niveaux de TFIIH tandis que chez la levure, son absence ne réduit pas les niveaux de TFIIH. L'effet sévère d'une mutation de TFB5 sur la fonction du NER suggère que le NER requiert une plus grande concentration de TFIIH que la transcription. En effet, des études dans les cellules vivantes démontrent que TFIIH participe plus longtemps au cycle de réparation par NER qu'au cycle de transcription (Hoogstraten 2002), expliquant la hausse du niveau de TFIIH lors d'une réparation par NER.

Plusieurs groupes ont démontré le recrutement de TFIIH durant le cycle de transcription (Pokholok 2002). Des études d'immunoprécipitation de la chromatine démontrent que TFIIH lie préférentiellement la région du promoteur et très peu la région codante des gènes, et son profil de distribution ne semble pas être modifié en présence de produits endommageant l'ADN (Heo 2004). TFIIH ne voyage donc pas nécessairement avec l'ARN pol II en élongation même lorsque l'ADN est endommagé, il est probablement recruté à l'ARN pol II arrêtée seulement lorsqu'il y a une demande pour l'assemblage du complexe de réparation (Lee 2002).

1.3.7 TFIS

TFIS est un facteur général d'élongation connu pour son effet sur les complexes d'élongation pausés ou arrêtés (Reines 1992, Wang 1993, Izban 1993, Wind 2000).

TFIIS est un analogue fonctionnel des facteurs de transcription eubactériens Gre (Kulich 2001). Durant l'élongation de l'ARNm, la polymérase peut rencontrer des séquences d'ADN qui causent le mouvement inverse de l'enzyme. De plus, l'ARN pol II est capable de détecter l'incorporation inappropriée de nucléotides à la chaîne d'ARN naissante et de corriger ces erreurs (« proofreading »). Une telle erreur engendre une paire de base mal appariée dans l'hybride ADN-ARN, et la distorsion induite de l'hybride déstabilise le complexe d'élongation et favorise un mouvement de l'ARN pol II vers l'arrière (Erie 2002). Durant cette marche arrière, ou « backtracking », des éléments structuraux de l'ARN pol II doivent maintenir les deux brins d'ADN ouverts en amont de la bulle de transcription. La « bridge helix » désapparie la matrice d'ADN et l'extrémité 3' de l'ARN durant la marche arrière, et l'extrémité 3' de l'ARN désapparié quitte le site actif à travers un pore qui traverse l'enzyme, ce qui peut mener à une pause et/ou à un arrêt transcriptionnel (Armache 2003).

Une pause est définie comme un blocage temporaire de l'élongation duquel l'ARN pol II peut échapper elle-même sans l'aide de facteurs accessoires. Normalement, l'hybride ADN-ARN prévient la marche arrière en gardant un registre de la position du dernier nucléotide transcrit (Nudler 1997). Cependant, certains sites dans l'ADN causent la déstabilisation de l'hybride, ce qui induit le recul de la polymérase et une pause transcriptionnelle (Palangat 2001). Durant l'arrêt transcriptionnel, l'ARN pol II effectue également une marche arrière, mais elle ne peut reprendre la transcription sans l'aide de TFIIS pour cliver le transcrit. La sortie de l'état d'arrêt requiert donc le clivage de l'ARN avec l'aide de TFIIS (Guo 1998, Rudd 1994, Fish 2002, Gu 1995a). L'ARN pol II possède une faible activité nucléase 3'→5' intrinsèque qui est très fortement stimulée par TFIIS et qui permet le clivage de dinucléotides ou de chaînes de 7 à 9 ribonucléotides à partir de l'extrémité 3' de complexes d'élongation pausés ou arrêtés respectivement (Gu 1995, Wang 1993). *In vitro*, le clivage induit par TFIIS augmente la capacité de correction, ou « proofreading », de l'ARN pol II, ce qui augmente la fidélité de transcription (Thomas 1998). L'ARN pol II isolée comporte un niveau assez élevé de mauvaise incorporation de nucléotides durant l'élongation, estimé à 1 par 10^3 à 10^5 (Uptain 1997). Le groupe de Hawley a appliqué un protocole forçant l'ARN pol II à

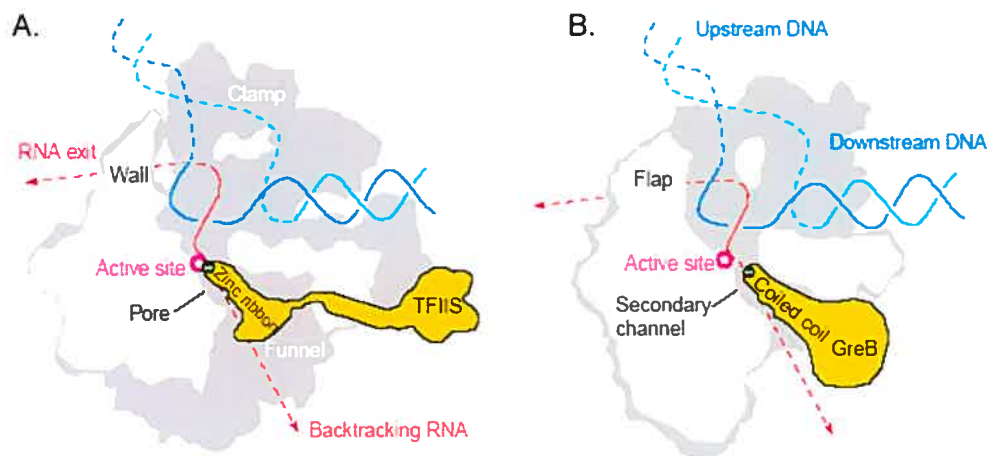


Figure 10 : Topologie conservée des complexes ARN polymérase avec les facteurs de clivages respectifs

Les directions de sortie de l'ARN durant la synthèse et durant la marche arrière sont indiquées par les flèches rouges.

A. Schématisation de l'ARN pol II en complexe avec TFIIIS.

B. Schéma correspondant à l'ARN polymérase bactérienne en complexe avec GreB.

L'ADN codant (bleu) et l'ARN (rouge) sont placés selon la structure du complexe d'élongation (Gnatt 2001). La sphère rose démarque la localisation du métal A.

Image tirée de: Cramer 2004

incorporer de l'IMP durant l'élongation. Ces auteurs ont démontré que l'addition de TFIIIS a grandement réduit l'incorporation de l'IMP et que le ribonucléotide mal apparié est enlevé efficacement de l'extrémité 3' de la chaîne d'ARN naissante (Thomas 1998).

TFIIIS a été le premier FGT purifié à partir des essais de synthèse de longs ARNm *in vitro* (Sekimizu 1976). Plus tard, des essais d'immunoprécipitation de la chromatine ont démontré qu'il pouvait s'associer avec la polymérase en transcription *in vivo* (Pokholok 2002). TFIIIS est un facteur modulaire qui comprend un domaine I N-terminal, un domaine II central et un domaine III C-terminal. Le domaine I, peu conservé, forme une unité de 4 hélices et n'est pas requis pour l'activité de TFIIIS (Booth 2000). Le domaine II forme une unité de 3 hélices (Morin 1996, Olmsted 1998), et le domaine III adopte un repliement en doigt de zinc avec une tige-boucle (Olmsted 1998, Qian 1993). Le

domaine II et la séquence entre II et III sont requis pour la liaison à la polymérase, tandis que le domaine III est essentiel à la stimulation du clivage de l'ARN.

Récemment, les rôles des facteurs TFIIS et GreB ont été éclaircis par des études structurales de ces facteurs liés à leurs polymérases respectives (figure 10) (Opalka 2003, Kettenberger 2003). TFIIS et Gre pénètrent le pore, ou canal secondaire, de la polymérase avec un ion Mg^{2+} dans le site actif. L'ion Mg^{2+} retenu par TFIIS ou GreB collabore avec un second ion Mg^{2+} , nommé métal A, à l'intérieur du site actif pour cliver la chaîne d'ARN naissante (Zhang 2003). Le domaine III atteint et complémente le site actif à l'aide d'une tige-boucle β hautement conservée et ayant deux résidus acides à proximité de l'ion métallique A de la polymérase et se rapprochant de la structure sucre-phosphate de l'ARN. Le domaine III de TFIIS ne bloque pas entièrement le pore dans lequel il pénètre, mais plutôt le restreint (Kettenberger 2003), laissant suffisamment d'espace pour l'entrée des NTP, tel que requis pour la polymérisation en présence de TFIIS (Horikoshi 1984). De plus, la liaison du domaine III de TFIIS induit le repliement d'une boucle de Rpb1 correspondant à la « trigger loop » bactérienne (Vassylyev 2002, Kettenberger 2003). TFIIS effectue des contacts potentiels près de cette boucle qui, lorsque repliée, referme le pore. Les changements structuraux de la « trigger loop » sont transmis à l'hélice nommée « bridge helix ». TFIIS forme des liens avec Rpb1 causant le mouvement vers le haut de la moitié C-terminale de la « bridge helix » et le déplacement vers l'extérieur des boucles *switch 1* et *2* (discutées plus loin dans le texte), causant ainsi l'élargissement du site de liaison de l'hybride ARN-ADN. Donc, en plus de compléter le centre catalytique de la polymérase, la liaison de TFIIS joue un rôle allostérique et réarrange la conformation du site actif et de l'environnement de l'hybride ARN-ADN.

1.4 Étapes de la réaction transcriptionnelle

La transcription des gènes de classe II peut être séparée en six étapes distinctes : la formation du PIC, la formation du complexe ouvert (ouverture de la bulle

transcriptionnelle), l'initiation, le dégagement du promoteur, l'élongation et la terminaison.

1.4.1 La formation du complexe de pré-initiation (PIC)

La reconnaissance du promoteur basal par la machinerie transcriptionnelle est essentielle pour le positionnement et l'assemblage adéquat d'un PIC contenant l'ARN pol II et les FGTs TFIIB, TFIID, TFIIE, TFIIF et TFIIH (Buratowski 1994, Buratowski 1989, Kornberg 1999a). La formation du PIC commence avec la liaison de facteurs régulateurs sur un élément d'ADN près du site d'initiation (figure 11, tableau I). Ces facteurs peuvent agir directement en interagissant avec la machinerie transcriptionnelle, ou indirectement, en recrutant des coactivateurs qui modifient ou remodelent la chromatine.

TFIID lie d'abord la boîte TATA via sa sous-unité TBP. TFIIA reconnaît TBP, stabilise le duplexe ADN-TBP et stimule le recrutement de TFIIB (Pugh 2000, Dion 2003). TBP, une molécule symétrique, requiert la présence d'un autre élément pour déterminer l'orientation de la transcription. TFIIB en se liant à TBP via l'élément BRE, dirige la directionnalité de la transcription en induisant l'assemblage antisymétrique du complexe et spécifie le site d'initiation (Buratowski 1989, Lagrange 1998). Ainsi, la liaison spécifique de la machinerie transcriptionnelle au promoteur basal résulte de la coopération entre au moins deux des FGTs, et les combinaisons peuvent varier entre différents promoteurs. TFIIA et TFIIB peuvent donc lier le complexe TBP-ADN simultanément et de manière synergétique et le stabiliser. TFIIF lie directement l'ARN pol II en solution et la recrute au promoteur (Killeen 1992). De plus, TFIIF induit l'enroulement du promoteur autour de la polymérase durant la formation du PIC (Robert 1998). TFIIE lie indépendamment l'ARN pol II et recrute le complexe TFIIH.

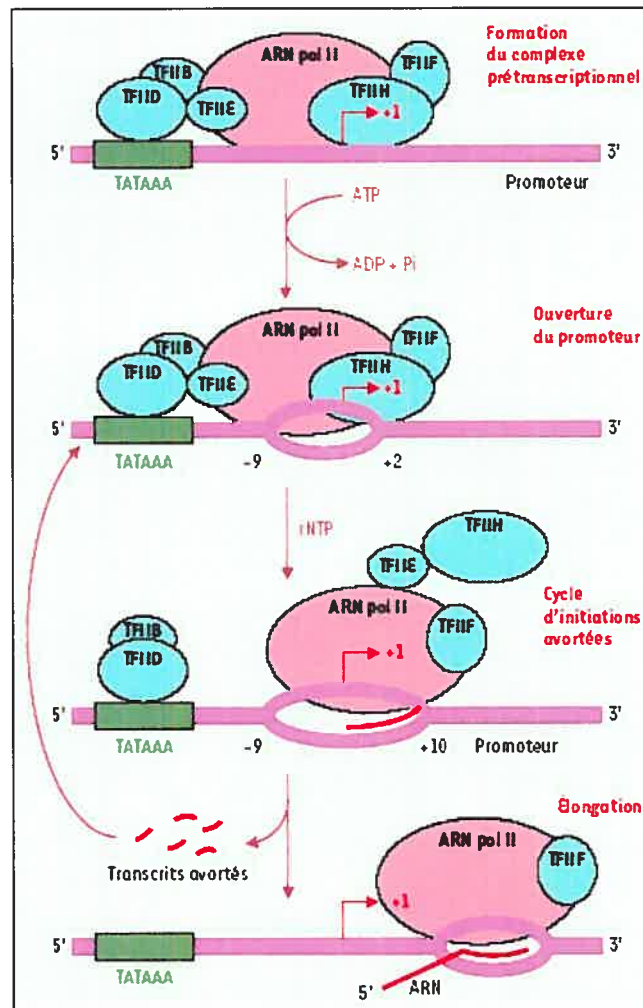


Figure 11 : Étapes de la transcription par l'ARN pol II

L'ARN pol II et les FGTs s'assemblent d'abord sur le promoteur pour former le PIC. TFIID catalyse l'ouverture du promoteur entre -9 et +2 par rapport au site d'initiation (+1) pour former la bulle transcriptionnelle, qui s'étend en aval au fur et à mesure que la transcription progresse. Jusqu'à la position +10, l'ARN pol II effectue un cycle d'initiations avortées. À la position +11, le promoteur se referme entre -9 et +2 et le complexe entre alors dans un mode d'élongation productive. La bulle transcriptionnelle reste alors ouverte entre -9 et +2 par rapport à la position du dernier rNTP ajouté.

Image tirée de: Langelier 2002

Tableau I:Les FGTs et les coactivateurs reliés à l'ARN pol II de *S. cerevisiae*

Facteur		Nb. de s.-u.	Fonction
TFIIA		2	<ul style="list-style-type: none"> • Stabilise TBP et la liaison TFIID-ADN • Bloque les inhibiteurs de transcription • Régule positivement et négativement certains gènes
TFIIB		1	<ul style="list-style-type: none"> • Lie TBP, l'ARN pol II et le promoteur d'ADN • Aide à fixer le site d'initiation
TFIID	TBP	1	<ul style="list-style-type: none"> • Lie l'élément TATA et déforme l'ADN au promoteur • Plateforme pour l'assemblage de TFIIB, TFIIA et les TAFs
	TAFs	14	<ul style="list-style-type: none"> • Lient l'Inr et le DPE • Représentent la cible des facteurs de régulation
Médiateur		24	<ul style="list-style-type: none"> • Se lie coopérativement avec l'ARN pol II • Activités kinase et acétyltransférase • Stimule la transcription basale et activée • Cible des facteur de régulation
TFIIF*		3	<ul style="list-style-type: none"> • Lie l'ARN pol II et est impliqué dans son recrutement au PIC ainsi que dans la formation du complexe ouvert • Lie Fcp1 • Stimule l'enroulement de l'ADN du promoteur autour de l'ARN pol II
TFIIE		2	<ul style="list-style-type: none"> • Lie le promoteur à proximité du site d'initiation • Aide à ouvrir ou stabiliser la bulle de transcription dans le complexe ouvert
TFIIH		10	<ul style="list-style-type: none"> • Possède plusieurs fonctions dans la transcription et la réparation de l'ADN • Activité kinase et deux activités hélicases • Essentiel pour la formation du complexe ouvert
SAGA	TAF	5	<ul style="list-style-type: none"> • Fonction inconnue
	Spt, Ada, Sgfs	9	<ul style="list-style-type: none"> • Organisent la structure du complexe SAGA • Interagissent avec TBP, TFIIA et Gcn5
	Gcn5	1	<ul style="list-style-type: none"> • HAT
	Tra1	1	<ul style="list-style-type: none"> • Sous-unité essentielle, cible des activateurs • La troisième plus grosse protéine chez la levure. Composante du complexe NuA HAT
	Ubp8	1	<ul style="list-style-type: none"> • Ligase d'ubiquitine

* La levure possède une sous-unité additionnelle par rapport à l'humain

Chez la polymérase bactérienne, plusieurs facteurs σ existent et sont utilisés dans des circonstances différentes. Les promoteurs bactériens sont reconnus par un seul facteur σ à la fois, tandis que l'initiation par l'ARN pol II requiert les 5 FGTs et au moins 15 autres polypeptides (Kornberg 1999a). Les facteurs σ participent à deux fonctions distinctes, notamment à reconnaître les régions -10 et -35 du promoteur et à faciliter l'ouverture de la bulle de transcription autour du site d'initiation (Gross 1998). Le rôle de la reconnaissance du promoteur implique non seulement la liaison au promoteur, mais également la capacité des facteurs σ à défaire les complexes ADN-ARN polymérase non-spécifiques. Chez les eucaryotes, TFIIF est également capable de défaire ces complexes non-spécifiques (Conaway 1990, Killeen 1992, Tan 1995) et sa sous-unité RAP30 constitue clairement un homologue des facteurs σ (Henry 1994, Sopta 1989).

Une étude de photo-pontage site-spécifique ADN-protéine des complexes d'ARN pol II liés à l'ADN indique que le PIC forme une structure très compacte dans laquelle l'ADN est replié et enroulé de manière serrée autour du noyau du complexe (Forget 2003, Robert 1998). L'ARN pol II se retrouve liée par photo-pontage à l'ADN de la position -40 à +26, une région de 66 pb correspondant à 225Å d'ADN de la forme B. Cette région est significativement plus étendue que la plus grande dimension de l'ARN pol II qui correspond à ~140Å (Forget 1997). Le chemin qu'emprunte l'ADN du promoteur à la surface de la polymérase a été déduit comme un enroulement droit autour de la pince mobile de l'ARN pol II. Un tel enroulement vers la droite détend l'hélice d'ADN, à l'opposé d'un enroulement vers la gauche qui induirait une compaction de l'hélice. Le déroulement de l'hélice d'ADN du promoteur pourrait jouer un rôle dans la production d'une région partielle d'ADN simple-brin nécessaire à l'action de l'hélicase simple-brin de TFIIH dans la formation du complexe ouvert (Forget 2004). Le photo-pontage de TFIIF et TFIIE en amont de l'élément TATA et en aval du site d'initiation indique que les deux facteurs sont associés avec la pince mobile. L'enroulement serré du promoteur autour de la pince serait responsable des contacts de ces deux facteurs à plus de 60 pb de distance. De plus, la présence de deux molécules de chacune des sous-unités de TFIIF et TFIIE ($\alpha_2\beta_2$) expliquerait l'étendue des contacts. TFIIE α , homologue à la région de liaison de l'ADN du facteur σ procaryote (Ohkuma 1991), possède une activité de liaison

à l'ADN importante lors de l'ouverture du promoteur et a été liée par photo-pontage au nucléotide -10 (Forget 2004).

1.4.2 Formation du complexe ouvert, initiation et dégagement du promoteur

Suite à l'assemblage du PIC, la machinerie transcriptionnelle sépare les brins d'ADN dans la région du site d'initiation sur environ 11 pb pour former la bulle de transcription. Cette ouverture découle de l'activité hélicase ATP-dépendante de TFIIF qui est requise pour l'ouverture des brins d'ADN entre les nucléotides -9 et +2, le site +1 étant le site d'initiation de la transcription (Guzman 1999). L'état actif requière l'hydrolyse de molécules d'ATP jusqu'à la position +4 puisque l'ajout d'ATP γ S, un analogue d'ATP non-hydrolysable, maintient le complexe dans un état fermé (Conaway 1989, Holstege 1997). Le complexe transcriptionnel demeure fonctionnellement instable jusqu'à ce que le transcrit naissant atteigne une longueur de près de 15 nucléotides. 3 étapes de transition doivent être franchies avant de parvenir à l'élongation stable (Dvir 2002). D'abord, la synthèse de l'ARN commence avec une phase abortive durant laquelle plusieurs courtes molécules d'ARN sont relâchées puisqu'elles ne sont pas liées de manière stable au complexe de transcription. Après la formation de 3 liens phosphodiesters, le complexe ouvert devient insensible à la présence de ATP γ S (Holstege 1997) et l'ARN demeure associé au complexe de transcription (Kugel 2002). Lorsque le transcrit atteint une longueur de quatre nucléotides, il y a une baisse du niveau de transcription abortive, et cette transition est nommée « escape commitment », ou engagement pour l'échappement (Kugel 2000). Une deuxième barrière doit être franchie lorsque l'ARN atteint une longueur de 10 nucléotides. L'ajout du onzième ribonucléotide à l'ARNm en synthèse cause un changement conformationnel dans lequel la bulle de transcription se repositionne de façon à ce que le registre +1 soit au centre de la nouvelle bulle formée. Cette étape irréversible implique la fermeture soudaine de l'hélice d'ADN entre les position -9 et +2 et le complexe devient alors processif. Finalement, la troisième transition est démontrée par la présence continue du cofacteur ATP et du facteur TFIIF jusqu'à ce que l'ARN soit d'une longueur de 15 nucléotides. Le passage réussi du complexe transcriptionnel à travers ces étapes est appelé

« dégagement du promoteur », et représente l'étape limitante de la transcription (Kugel 1998).

1.4.3 Élongation

La vitesse d'élongation *in vivo* atteint 1200 à 2000 nucléotides par minute, tandis qu'*in vitro*, elle atteint seulement 100 à 300 nucléotides par minute sur des matrices d'ADN nu, en plus d'être freinée par des sites de pauses très fréquents (Conaway 1999). Dans les deux cas, l'élongation efficace doit surmonter plusieurs obstacles intrinsèques à la matrice d'ADN chromatinisée ou à la séquence de certains éléments, tel que les répétitions de thymines. Il existe 3 obstacles principaux à l'élongation, notamment, les pauses, les arrêts et la terminaison (Uptain 1997, Shilatifard 2003). La plupart des facteurs d'élongation identifiés servent à contrôler ou surmonter ces obstacles (tableau II). Les facteurs d'élongation connus sont regroupés en 3 catégories fonctionnelles. La première est composée des facteurs impliqués dans l'arrêt séquence-dépendent ou induit par les intercalants, et inclut la famille des facteurs TFIIIS et PTEF-b (pour « positive transcription elongation factor b »). Comme mentionné plus tôt, TFIIIS induit l'élongation efficace par l'ARN pol II en empêchant l'arrêt prématuré dû à une séquence d'ADN spécifique, en plus de réactiver une polymérase arrêtée. PTEF-b est un facteur hétérodimérique composé de la cycline T et de Cdk9, responsable de phosphoryler la sérine 2 du CTD et dont l'activité kinase est contrebalancée par l'action de la phosphatase Fcp1 (Cho 1999, Cho 2001, Archambault 1997). La deuxième classe de facteurs régule le taux d'élongation à travers la chromatine et inclut le facteur FACT (pour « facilitates chromatin transcription »), tandis que la troisième classe augmente le taux catalytique d'élongation en altérant le K_m et le V_{max} de l'ARN pol II, et inclut TFIIF et les familles de facteurs Elongins et ELL (Elmendorf 2001, Izban 1992b). Les fonctions de chacun des facteurs sont mentionnées plus en détails dans le tableau II.

Tableau II : Facteurs d'élongation associés à l'ARN pol II (adapté de Sim III 2004)

Facteurs d'élongation	Sous-unités	Actif / Passif	Propriétés
TFIIF	RAP30, RAP74	Actif	Essentiel pour la formation du PIC, aide à sortir des pauses, module TFIIIS
Elongins	Elongin A, B, C	Actif	Aide à sortir des pauses, rôle dans l'ubiquitination
ELL	EAP 20, 30, 45	Actif	Aide à sortir des pauses
DSIF	Spt 4, 5	Actif	Stimule l'élongation, supprime la terminaison précoce de la transcription, stimule le coiffage
NELF	NELF -A, -B, -C, -D, -E	Actif	Pause l'ARN pol II pour permettre le coiffage temporel
CSB		Actif	Stimule l'élongation, module TFIIIS, joue un rôle dans la rescousse de la polymérase au point de dommage à l'ADN et dans la réparation par excision couplée à la transcription
FCP1		Actif	Stimule l'élongation, recycle l'ARN pol II en déphosphorylant les Ser5 du CTD, rôle dans le coiffage
TFIIS		Actif	Stimule le clivage par l'ARN pol II du transcrit naissant pour contrer l'arrêt
Spt6		Actif	Stimule l'élongation, module la structure de la chromatine, activité chaperonne pour les histones
HDAg		Actif	Stimule l'élongation, lie l'ARN pol II, déplace NELF, fonctionnellement distinct de TFIIF
19S Proteasome	Rpt4, Sug1 (Rpt6), autres	Actif	Recruté par la mono-ubiquitination de H2B, impliqué dans la méthylation de H3-K4
P-TEFb*	Cdk9, cyclin T	Passif	Délivre de la pause induite par NELF, phosphoryle Ser2 du CTD et DSIF (Spt5)
Ssu72		Passif	Déphosphorylation des Ser5 du CTD, rôle dans la maturation 3' des transcrits
SWI/SNF	BRG1, autres	Passif	Remodelage de la chromatine ATP-dépendant
lsw1p (lsw1b)	lsw1p, lsw2p, loc4p	Passif	Régule la phosphorylation des Ser2 et Ser5, la méthylation de H3-K4 et H3-K36, et le recrutement des facteurs de maturation 3' du transcrit
Chd1		Passif	Interagit génétiquement et physiquement avec les facteurs d'élongation, localisé dans les régions codantes
FACT	Spt16	Passif**	Facilite l'élongation à travers la chromatine, régule la structure de la chromatine, activité chaperonne pour les histones
Set1		Passif	Méthylation de H3-K4, localisé dans les régions de promoteur et des régions codantes
Set2		Passif	Méthylation de H3-K6, localisé dans les régions codantes
Paf	Paf1, Cdc73, Ctr9, Leo1, Rtf1	Passif	Module la mono-ubiquitination de H2B et la méthylation de H3-K3, recrute Set1 et Set2, impliqué dans la modulation de la maturation de l'ARNm
THO	Hpr1, Tho2, Mft1, Thp2	Passif	Requis pour la transcription de long transcrits, ou ceux contenant un niveau élevé de GC
TREX	THO, Sub2, Yra1	Passif	Relie l'élongation à l'épissage et l'export des pre-ARNm, surveillance, interaction avec Spt6
lsw1/Spn1		Passif	S'associe à Spt6, localisé dans les régions codantes

* Ctk1 et Bur1 sont des homologues de Cdk9, la sous-unité catalytique de P-TEFb

** Affecte le niveau de transcription de l'ARN pol II sur des matrices de chromatine

1.4.3.1 Fidélité, spécificité et translocation

La fidélité de transcription est définie comme étant la capacité de l'ARN pol II d'incorporer correctement les nucléotides complémentaires aux bases du brin matrice de l'ADN. Cette fidélité dépend du bon positionnement du NTP entrant pour optimiser les interactions Watson-Crick entre le NTP et la base codante de la matrice.

Une autre caractéristique importante de l'ARN pol II est sa spécificité pour la synthèse de chaînes d'ARN plutôt que des chaînes d'ADN. Cette spécificité peut être accomplie par au moins trois mécanismes possibles. Premièrement, le groupe discriminatoire 2'-OH du NTP entrant peut former des liens hydrogènes avec certains résidus conservés de l'ARN pol II (Cramer 2001, Gnatt 2001). Deuxièmement, les groupes 2'-OH des quelques derniers nucléotides incorporés dans la chaîne d'ARN naissante forment directement des liens hydrogènes avec des résidus de l'ARN pol II, ce qui stabilise les complexes qui incorporent des ribonucleotides. Finalement, le centre catalytique de l'ARN pol II est complémentaire à la conformation spécifique de l'hybride ADN-ARN.

L'ARN pol II doit lier fortement l'ADN et l'ARN pour créer un complexe d'élongation stable et processif. En même temps, elle doit permettre la translocation précise des acides nucléiques pour se déplacer sur la matrice d'ADN à une vitesse considérable de plusieurs centaines de nucléotides par minute. Les différentes structures cristallographiques de l'ARN pol II ont donné plusieurs indices pour comprendre cette translocation (ces structures, ainsi que le mécanisme de transcription, seront discutés en détail plus tard). Tout d'abord, les interactions avec les acides nucléiques se font via le squelette de sucre et de phosphate, ce qui évite les interactions base-spécifiques qui interfèreraient avec la translocation. De plus, plusieurs charges positives de type « long-range » forment une enveloppe autour des acides nucléiques à une distance allant jusqu'à 8Å (Thomas 1998). De telles interactions électrostatiques permettent une liaison forte des acides nucléiques sans restreindre leur mouvement. Finalement, la translocation est accompagnée de changements conformationnels chez l'ARN pol II des régions entourant les acides

nucléiques. Un exemple de ces mouvements est observé sur une hélice située près du site catalytique, nommée « bridge helix », et sera discuté plus tard.

1.4.3.2 Modèle d'élongation

Le groupe de Burton a récemment proposé (Nedialkov 2003) un modèle de translocation de l'ARN pol II humaine induit par la présence des NTP. Leur méthode consistait à suivre la progression de molécules d'ARN pol II humaines pausées en élongation sur une séquence d'ADN et observer la sortie des pauses et l'élongation. La sortie des pauses requiert seulement la translocation, le relâchement du pyrophosphate étant complété durant la pause, tandis que l'élongation processive requiert la translocation et le relâchement du pyrophosphate. Ils ont démontré que ces deux étapes dépendaient de la présence du prochain NTP entrant. Ils ont également noté que les deux étapes sont régulées différemment par les facteurs d'élongation TFIIF et l'antigène de l'hépatite δ . L'antigène de l'hépatite δ (HDAg) est un facteur d'élongation viral de l'ARN pol II codé par le virus de l'hépatite δ (HDV) (Lai 1995). Son rôle dans l'élongation semble important du côté clinique puisque le virus HDV coexiste et complique l'infection chronique et sévère du virus humain de l'hépatite B. Le FGT TFIIF, tel que mentionné plus tôt, joue un rôle critique lors de la phase de dégagement du promoteur (Yan 1999, Dvir 2002) et stimule l'élongation par l'ARN pol II 5 à 10 fois *in vitro* en supprimant les pauses transcriptionnelles (Renner 2001, Tan 1994, Lei 1999, Bengal 1991, Izban 1992a, Price 1989). Leurs résultats suggèrent que TFIIF induit un resserrement de la pince sur l'hybride ARN-ADN et une géométrie optimale pour la catalyse au site actif. Ainsi, lorsque la pince est serrée, la « bridge helix » s'enfonce sur l'hybride et le stabilise. Advenant une pause, le complexe d'élongation tend à se trouver à l'état pré-transloqué et le prochain NTP codant est requis pour stimuler la translocation et pousser le complexe hors de la pause. Cependant, dans le cas du facteur HDAg, leurs études suggèrent qu'il peut desserrer la prise de la pince et faciliter la translocation de l'hybride. De ce fait, l'élongation devient moins dépendante du chargement du prochain NTP codant puisque le NTP entrant n'est plus requis pour pousser le complexe hors de la pause (Nedialkov 2003). Le modèle de Burton confirme également qu'à des concentrations limitantes de

NTP, la translocation au site actif est l'étape critique et limitante de la synthèse d'ARN et qu'elle est induite par la liaison du prochain NTP codé par la matrice.

En se basant sur les résultats de structure, plusieurs équipes ont suggéré que le pore était la route principale des NTP à charger (Gnatt 2001, Cramer 2001, Vassylyev 2002, Gnatt 2002, Zhang 1999b), mais les études de cinétique de l'équipe de Burton sont en désaccord avec cette hypothèse. Ils suggèrent plutôt que durant la synthèse processive, les NTPs sont chargés par le canal principal, avant l'étape de translocation. Durant l'état pré-transloqué, l'hybride ADN-ARN semble bloquer l'accès des NTP à travers le pore secondaire, mais après la translocation, par exemple suite à une pause transcriptionnelle ou une translocation NTP-indépendante, la base d'ADN $n+1$ semble dirigée vers le pore pour le chargement du nucléotide (Zhang 2004).

1.4.4 Terminaison, recyclage de la polymérase, réinitiation et régulation

La réaction de terminaison est couplée à une réaction de modification de l'extrémité 3' de la chaîne d'ARN. La plupart des précurseurs d'ARNm eucaryotes sont clivés à un site spécifique dans la région 3' non-traduite avant d'être polyadénylés. La polyadénylation est un processus biochimique par lequel une chaîne de plusieurs adénines (200 à 250 en moyenne) est ajoutée de manière covalente à l'extrémité 3' des ARNm et sert à la stabilité du messager et son exportation du noyau (Hirose 1998). La terminaison de la transcription est également couplée à la déphosphorylation du CTD de l'ARN pol II. Cette déphosphorylation est requise pour la réinitiation de la transcription, puisque l'ARN pol II peut seulement joindre le complexe d'initiation lorsque son CTD est déphosphorylé. La CTD phosphatase Fcp1 joue un rôle important dans ce recyclage (Cho 1999, Kobor 1999) et c'est la sous-unité Rpb4 de la polymérase qui recrute Fcp1 à proximité du CTD (Kimura 2002, Armache 2003).

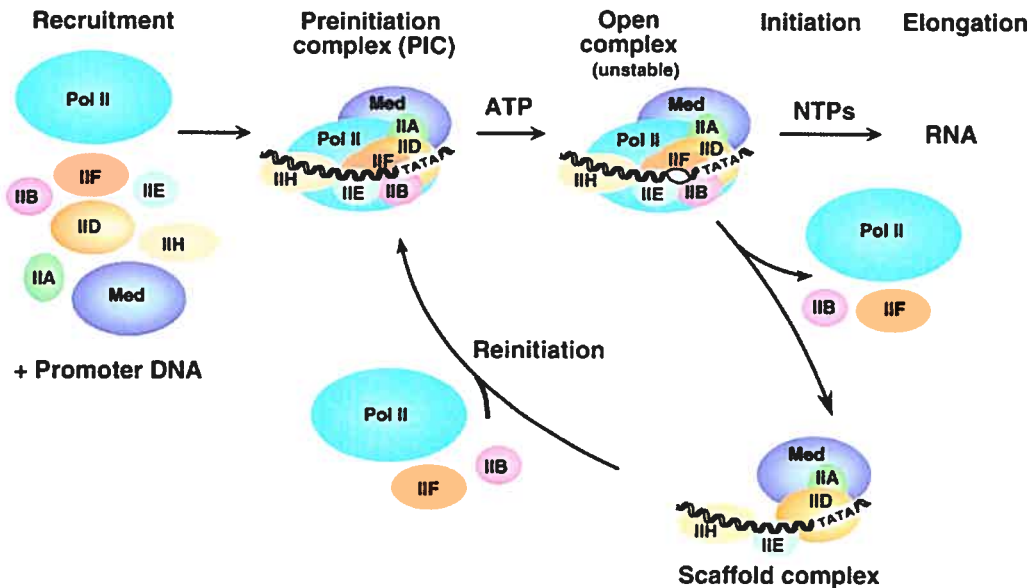


Figure 12 : Mécanismes d'initiation et de ré-initiation de l'ARN pol II

À la suite de l'initiation, un complexe d'échafaudage est laissé sur le promoteur pour permettre la ré-initiation de la transcription.

Image tirée de: Hahn 2004

La réinitiation de la transcription se produit par un mécanisme différent de celui de l'initiation (figure 12) (Hahn 1998). Suite à l'initiation et au dégagement du promoteur, une partie de la machinerie transcriptionnelle est laissée sur la région promotrice et devient un échafaud pour l'assemblage d'un nouveau complexe d'initiation. Cet échafaud comprend TFIIA, TFIID, TFIIE, TFIIH et un complexe Médiateur, et peut être stabilisé par des activateurs transcriptionnels, tel que Gal4-VP16 (Yudkovsky 2000). Le complexe ARN pol II-TFIIF peut alors être recruté pour d'autres cycles de transcription (Rani 2004). La réinitiation, ainsi que l'initiation et l'élongation, représentent des cibles importantes pour la régulation de l'ARN pol II étant donné que cette dernière est considérée le point ultime des voies de signalisation et de transduction.

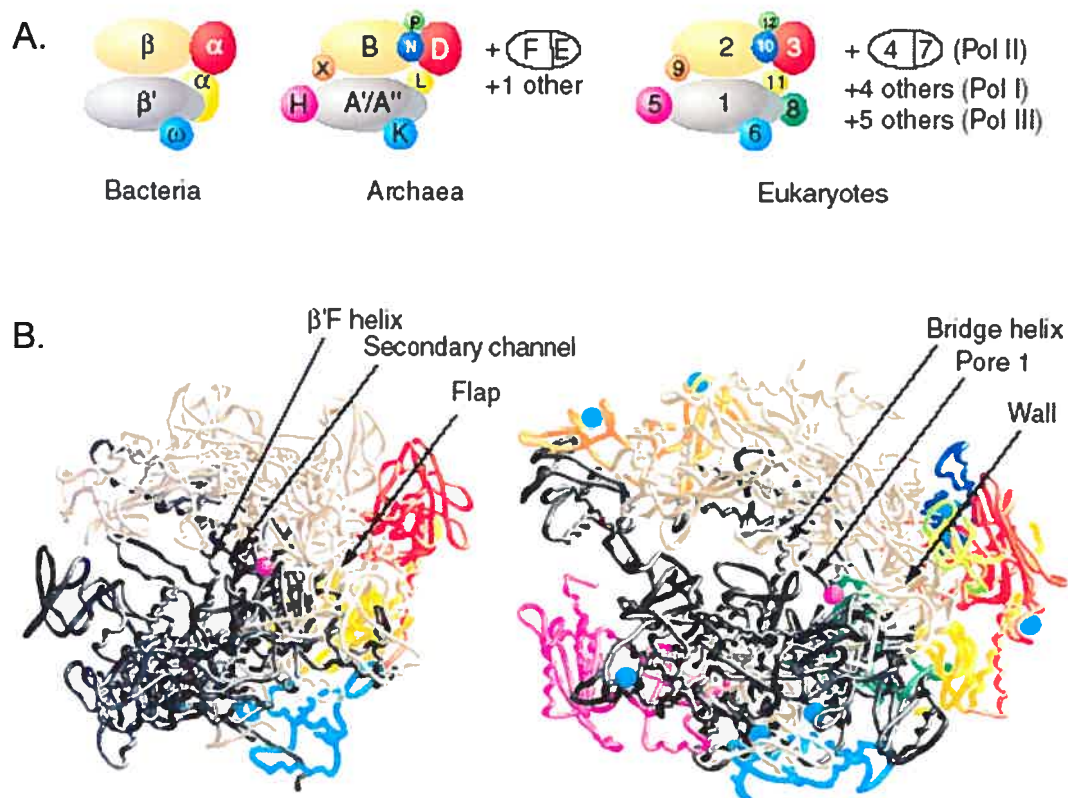


Figure 13 : Architecture des sous-unités de l'ARN polymérase

A. Présentation schématique des sous-unités des polymérases bactérienne, archée et eucaryote. Les sous-unité correspondantes possèdent les mêmes couleurs.

B. Structure tri-dimensionnelle de la polymérase de bactérie *T. aquaticus* et de levure *S. cerevisiae*. Le métal A du site actif est illustré en rose et les ions de zinc sont en bleu.

Image tirée de: Cramer 2002

1.5 Architecture de l'ARN pol II

Les structures cristallographiques des ARN polymérases bactérienne et de levure montrent des éléments protéiques hautement similaires dans la région du site actif, suggérant que le mécanisme de formation de la chaîne d'ARN est conservé (figure 13) (Cramer 2001, Zhang 1999). L'ARN pol II comprend 12 sous-unités, nommées Rpb1 à Rpb12, d'un poids total de plus de 500 kDa. Les deux plus grandes sous-unités, Rpb1 (220 kDa) et Rpb2 (140 kDa) forment le site catalytique de l'enzyme et sont respectivement homologues aux sous-unités β' et β de l'ARN polymérase bactérienne.

Cinq sous-unités, notamment Rpb5, Rpb6, Rpb8, Rpb10 et Rpb12, de poids moléculaire variant entre 7 et 28 kDa, se retrouvent également chez les ARN polymérase I (responsable de la synthèse des ARNr) et III (responsable de la synthèse des ARNt et ARNr 5S). Rpb3 et Rpb11 sont homologues à l'homodimère α_2 impliqué dans l'assemblage de l'ARN polymérase bactérienne, tandis que Rpb9 est impliquée dans l'élongation à travers les sites d'arrêts. Chez *S. cerevisiae*, Rpb4 et Rpb7 forment un sous-complexe qui peut se dissocier du reste de l'enzyme selon les conditions environnementales (Orlicky 2001) (voir plus bas). Rpb7 est essentiel à la survie et la croissance de la cellule (McKune 1993, Mitsuzawa 2003), tandis que Rpb4 est dispensable dans des conditions de cultures optimales mais essentiel dans certaines conditions extrêmes (Woychik 1989, Sheffer 1999, Choder 1993, Farago 2003, Miyao 2001). Lorsque la cellule est en croissance active, la plupart des complexes d'ARN pol II ne contiennent pas le dimère Rpb4-Rpb7, ce dernier tend plutôt à s'associer durant la phase stationnaire ou à la suite d'un stress (Khazak 1998).

Les études structurales à haute résolution requièrent de grandes quantités des protéines purifiées qui ne peuvent pas être obtenues par surexpression dû à la complexité de l'enzyme. Les études fonctionnelles sur l'ARN pol II humaine étaient restreintes dû au manque de système de purification approprié des différentes formes de l'enzyme humaine. Des rôles généraux avaient été attribués à certaines sous-unités mais leur implication précise dans les différentes étapes de la transcription restait floue jusqu'à la détermination, par Roger Kornberg et ses collaborateurs, de la structure d'une enzyme contenant 10 des 12 sous-unités à une résolution de 2.8Å et de celle de l'enzyme en élongation à 3.3Å, n'incluant pas le dimère Rpb4-Rpb7 (figure 14) (Cramer 2001, Gnatt 2001). Ces structures ont permis d'explorer le mécanisme transcriptionnel en permettant de voir, à l'échelle atomique, les chaînes latérales des acides aminés de l'enzyme libre et de positionner le duplex ARN-ADN par rapport aux différents éléments du site catalytique.

Les structures disponibles ont permis l'identification de domaines qui ont été nommés selon leur fonction présumée et leur localisation et qui sont souvent composés de parties

de différentes sous-unités (figure 14). Huit ions de zinc se retrouvent près de la surface de l'enzyme et servent à stabiliser le complexe. Les sous-unités Rpb1 et Rpb9 lient chacune deux ions de zinc, tandis que les sous-unités Rpb2, Rpb3, Rpb10 et Rpb12 en lient un seul. Les deux plus grosses sous-unités de l'ARN pol II, Rpb1 et Rpb2, forment la masse centrale du noyau ainsi qu'un sillon (« cleft ») chargé positivement (Cramer 2000, Cramer 2001). Le centre catalytique est formé entre une pince mobile (« clamp »), une « bridge helix » (hélice de pontage) qui traverse le *cleft* et un mur, ou « wall », qui bloque l'extrémité du *cleft* et fait courber l'ADN de 90° vers le haut. Une paire de mâchoires, ou « jaws », borde l'entrée du *cleft*, la mâchoire supérieure formée par Rpb9 et une partie de Rpb1, et la mâchoire inférieure chargée positivement formée par Rpb5 et Rpb1. Au fond du *cleft* se trouvent deux ions Mg^{2+} qui indiquent l'emplacement du site actif. L'hybride prend place dans ce sillon, où sont localisées la majorité des charges positives de l'enzyme, entre la « bridge helix » de Rpb1 qui traverse le *cleft* et le *wall* de Rpb2 (Cramer 2001). Au-dessus du site catalytique se trouve un canal qui servirait, lors de l'élongation normale, à l'entrée des rNTP. Au fond du *cleft* se trouve un pore, ou canal secondaire, qui s'étend jusqu'à l'arrière de l'enzyme comme une entonnoir renversé et qui servirait au relâchement des pyrophosphates et à l'entrée des rNTP lorsque l'enzyme est à l'état post-transloqué (Nedialkov 2003). La bordure du pore comprend une boucle de Rpb1 liant un ion Mg^{2+} (métal A) via 3 résidus d'aspartates D481, D483 et D485, faisant partie d'un motif NADFDGD strictement conservé (Cramer 2001, Gnatt 2001, Westover 2004b). Un second ion métallique, le métal B, est localisé près des résidus D481 et D483 de Rpb1, ainsi que des résidus E836 et D837 de Rpb2 hautement conservés (Cramer 2001, Gnatt 2001, Westover 2004b) et semble nécessiter la présence d'un NTP pour sa coordination adéquate (Cramer 2001, Gnatt 2001, Westover 2004b, Nedialkov 2003).

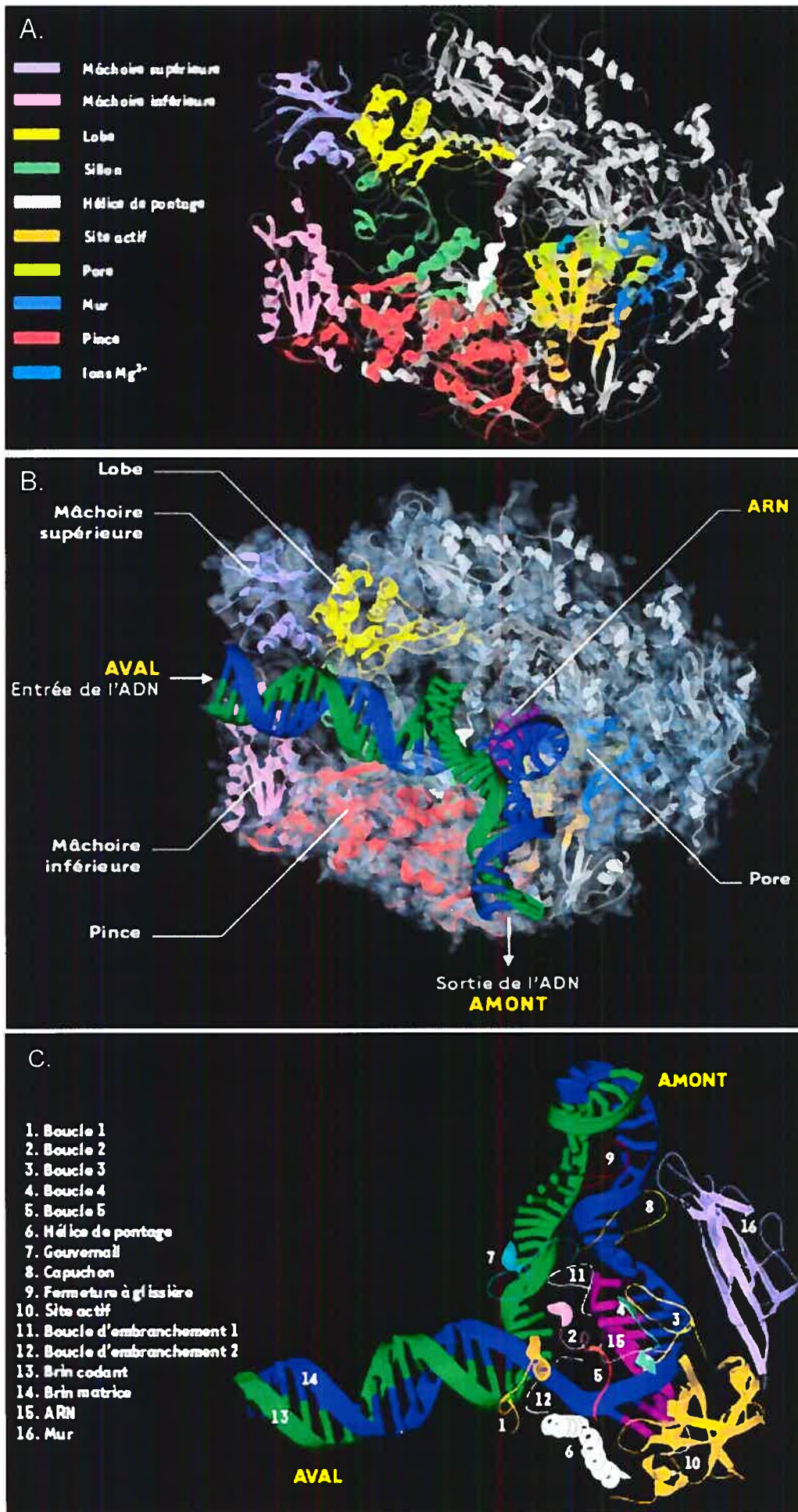


Figure 14 : Structure cristallographique de l'ARN polymérase II

A. Enzyme libre à une résolution de 2,8 Å. 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 pore localisé à l'arrière de l'enzyme permettrait la sortie de l'extrémité 3' libre de l'enzyme.

B. Enzyme en élongation à une résolution de 3,3 Å. La pince formée de parties de Rpb1 et Rpb2 effectue une rotation de 30° qui maintient la matrice d'ADN et le transcrit au sein du complexe transcriptionnel. L'extrapolation du trajet probable des acides nucléiques au sein de l'enzyme montre que l'ADN double brin situé en aval fait contact avec la mâchoire inférieure puis passe entre le lobe de Rpb2 et une partie de la pince formée par Rpb1. Des flèches indiquent les sites d'entrée et de sortie de l'ADN.

C. Contacts ARN pol II-ADN-ARN dans la région de la bulle de transcription. Plusieurs domaines des sous-unités Rpb1 et Rpb2 font contact avec l'ADN et l'ARN au niveau de la bulle transcriptionnelle en formant un réseau complexe d'interactions qui crée une zone hautement complémentaire à la liaison des acides nucléiques. Cette zone contribue à la stabilité du complexe transcriptionnel et donc à l'efficacité de l'élongation par la polymérase. Les boucles 1 à 5 sont nommées dans le texte "Switch".

Image tirée de: Langelier 2002

1.5.1 Réaction de polymérisation

Notre compréhension des mécanismes d'action de l'ARN pol II a également été révolutionnée par la publication des structures cristallographiques de la polymérase et d'un complexe en élongation ARN pol II-ADN-ARN (figure 14, bas) (Cramer 2001, Cramer 2000, Gnatt 2001, Fu 1999). La différence de conformation du site actif entre la polymérase en élongation et l'enzyme libre a été utilisée pour proposer un mécanisme d'addition du nucléotide à la chaîne d'ARN, demandant la participation de deux ions magnésium, et conservé à travers les différentes ARN et ADN polymérases de plusieurs organismes (figure 15A). Le métal A diminue l'affinité de l'oxygène 3' à l'hydrogène, facilitant ainsi l'attaque nucléophile du O⁻ sur le phosphate α en 5', ce qui mène à la formation d'un lien phosphodiester. Le métal B facilite le relâchement de la molécule de pyrophosphate et les deux ions Mg²⁺ stabilisent la structure et la charge de l'état de transition (Doublie 1998).

La résolution de la structure de l'ARN pol II avec un nucléotide mal-apparié a dévoilé l'existence de deux sites de liaison des NTP entrants. Il a été suggéré que le complexe NTP-Mg²⁺ (qui représente le métal B) pénètre d'abord au site actif à travers le pore et se lie au site d'entrée (E) dans une orientation inversée. Par la suite, le NTP subit une rotation et se trouve au site d'addition (A) pour la catalyse du lien phosphodiester (Kettenberger 2004b), et cette translocation serait coordonnée par le résidu E836 de Rpb2 (Langelier 2005). De plus, les acides aminés de la polymérase le long du site de liaison des NTP et du site actif sont hautement conservés à travers les espèces et les ARN polymérases nucléaires. Cette observation suggère que le mécanisme de liaison, de sélection et d'incorporation des NTP est semblable chez tous ces enzymes (Kettenberger 2004b). En effet, la discrimination pour les NTP par rapport aux dNTP chez la polymérase bactérienne implique le même résidu que celui correspondant à N479 de la levure (Svetlov 2004).

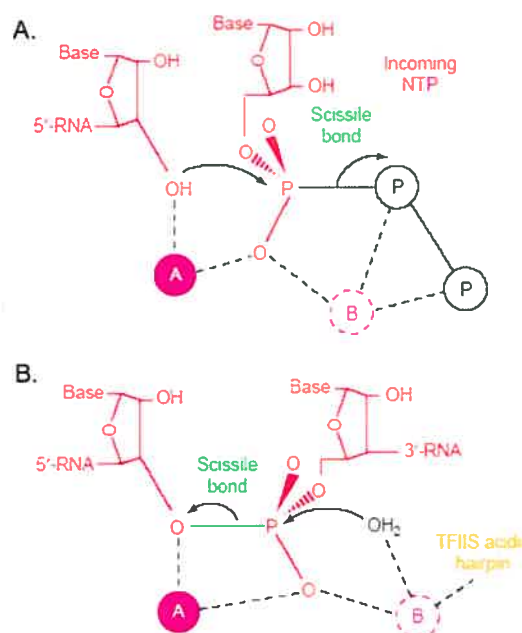


Figure 15 : Modèle du site actif « ajustable » de l'ARN pol II

A. Mécanisme présumé de l'incorporation de nucléotide durant la polymérisation de l'ARN.

B. Mécanisme présumé du clivage hydrolytique de l'ARN stimulé par TFIIS. Une caractéristique clé du modèle est qu'un métal B interagit soit avec le NTP entrant (pour la polymérisation), soit avec la tige-boucle acide de TFIIS et une molécule d'eau nucléophile (clivage)

Image tirée de: Cramer 2004

1.5.2 Structure du site actif de l'ARN pol II

Au centre du noyau de l'ARN pol II en élongation, l'ADN et l'ARN forment un hybride de 9 pb au-dessus du métal A et du pore, et la *clamp* adopte un état fermé (Gnatt 2001). À la base de la *clamp*, du côté extérieur, s'étend une région « linker » d'une centaine de pb reliant la région CTD de Rpb1. Cette région est désordonnée dans toutes les structures cristallographiques. Dans la structure de la polymérase en élongation, la position de la *clamp* mobile restreint le diamètre du *cleft*, trop étroit pour accommoder le passage linéaire de l'ADN double-brin du promoteur. Par contre, dans celle de l'enzyme libre, la *clamp* est distancée du reste de la polymérase, laissant un espace qui pourrait, en principe, accommoder l'ADN double-brin. Un mécanisme d'interaction de la polymérase

avec son promoteur a donc été proposé, dans lequel le mouvement de la pince mènerait l'enzyme de la conformation en initiation à celle en élongation et expliquerait la nature processive de l'enzyme, en fermant la pince sur l'ADN (Cramer 2000). Toutes les structures sur lesquelles ce modèle est basé comportaient la forme de l'ARN pol II à 10 sous-unités, compétente en élongation. Les sous-unités Rpb4 et Rpb7, essentielles pour l'initiation dirigée, interféraient avec la cristallisation et se trouvaient absentes des structures étudiées. Récemment, la structure cristallographique de la polymérase à 12 sous-unités a été élucidée et contredit cette hypothèse en montrant la *clamp* dans son état fermé, en absence d'ADN. Il a été proposé que l'ADN se situe plutôt sur le dessus de la *clamp* lors de la formation du PIC, et que lorsque le promoteur s'ouvre, l'ADN descend alors au fond du *cleft* (Armache 2003, Bushnell 2003). La *clamp* se retrouve dans un état fermé et son mouvement est restreint par le complexe Rpb4/Rpb7 qui forme une cale en dessous (figure 16). Cette observation suggère que la *clamp* est fermée durant l'initiation et que le duplexe d'ADN du promoteur est initialement lié à l'extérieur du *cleft*. Après l'ouverture de l'ADN, la matrice simple-brin pourrait glisser à l'intérieur du *cleft* et lier le site actif pour initier la synthèse de l'ARN. L'ouverture étendue de la *clamp* pour le passage du duplexe d'ADN (Cramer 2001) nécessiterait la dissociation de Rpb4/Rpb7 et semble plutôt improbable (Cramer 2004).

Plusieurs boucles et hélices se rapprochent ou entrent en contact avec le duplexe ARN-ADN, suggérant qu'elles pourraient jouer un rôle important dans la réaction de transcription (figure 14c) (Langelier 2002). Les structures de la polymérase libre et en association avec le duplexe ARN-ADN démontrent que l'ARN pol II possède une *clamp* rigide mais mobile qui oscille de 30Å entre un état ouvert, lorsque l'enzyme est libre en solution, et un état fermé lorsque l'enzyme est en transcription. Dans la structure du complexe d'élongation, la *clamp* lie la matrice d'ADN en amont et au centre de l'hybride ARN-ADN. À la base de cette *clamp*, une série de 5 boucles, appelées *switches*, subissent des changements conformationnels en couplant la rotation de la *clamp* à la présence d'un hybride ARN-ADN dans le *cleft*, et relie la *clamp* au reste de l'ARN pol II. Les boucles *switch* 1 et 2 sont très peu ordonnées dans la structure de l'enzyme en solution, mais deviennent ordonnées dans la structure du complexe en élongation. Cette

stabilisation est induite par la liaison des *switches* à l'ADN en duplexe avec l'ARN et celui plus en aval de l'hybride. La *clamp* est stabilisée par 3 ions de zinc et est formée par l'extrémité C-terminale de Rpb2 ainsi que de deux régions de Rpb1, l'une localisée à l'extrémité N-terminale et l'autre près du domaine C-terminal (CTD) de Rpb1.

1.5.3 Un seul site actif pour deux activités

Les études structurales et biochimiques ont identifié un mécanisme unifié comportant deux ions métalliques dans un simple site actif ajustable entre la polymérisation et le clivage de l'ARN chez toutes les ARN polymérase nucléaires (Kettenberger 2003, Sosunov 2003, Opalka 2003). Selon ce modèle, la polymérisation et le clivage de l'ARN requièrent le métal A, mais c'est la coordination différentielle du métal B qui alterne l'activité de la polymérase entre la polymérisation et le clivage (figure 15). Pour la polymérisation, le métal B lie les phosphates du NTP entrant, tandis que pour le clivage stimulé de l'ARN, la tige-boucle acide de TFIIIS positionne le métal B ainsi qu'une molécule d'eau nucléophile au site actif. Lors du clivage en absence de TFIIIS, le métal B est lié par un nucléotide additionnel non-attaché mais situé dans le pore (Sosunov 2003). Il est à noter que ces mécanismes demandent d'être vérifiés par des études additionnelles tel que des structures à haute résolution.

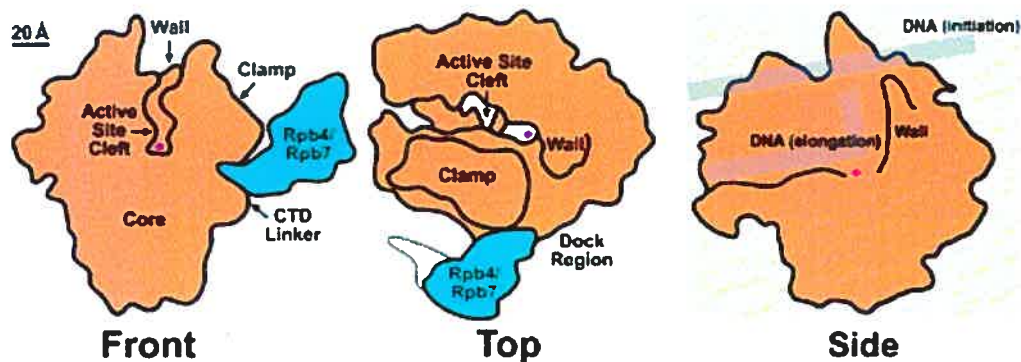


Figure 16 : Structure de l'ARN pol II et son interaction avec le sous-complexe Rpb4/Rpb7

Image tirée de: Asturias 2004

1.5.4 Ouverture de la bulle de transcription et chemin de sortie de l'ARN

Trois boucles, nommées gouvernail ou « rudder », capuchon ou « lid », et fermeture à glissière ou « zipper », s'étendent de la *clamp* et jouent un rôle dans la dissociation de l'hybride, l'ouverture du côté amont de la bulle de transcription et la sortie de l'ARN. Plus précisément, la boucle *zipper* semble jouer un rôle dans la séparation des brins d'ADN codant et non-codants en amont du site actif. Les boucles d'embranchement *fork1* et *fork2* maintiennent la bulle de transcription en amont et en aval, respectivement. Ainsi, l'ADN en aval est désapparié, ou déroulé, avant d'entrer au site actif de la polymérase, et est ré-apparié en amont lorsqu'il sort de la polymérase (Erie 2002, Shilatifard 2003). L'ADN entre donc via la paire de *jaws* et s'étend à travers le *cleft* en direction du site actif (figure 17). L'hybride surgit du site actif le long du *wall* à angle droit, tandis que l'extrémité 3' de la chaîne d'ARN croissante s'étend par-dessus le pore à travers lequel les nucléotides peuvent entrer. En amont du point de séparation des brins d'ADN, la matrice positionne une base dans le site actif qui dirige la synthèse d'ARN. L'hybride ADN-ARN forme plusieurs contacts avec la polymérase (Gnatt 2001) et

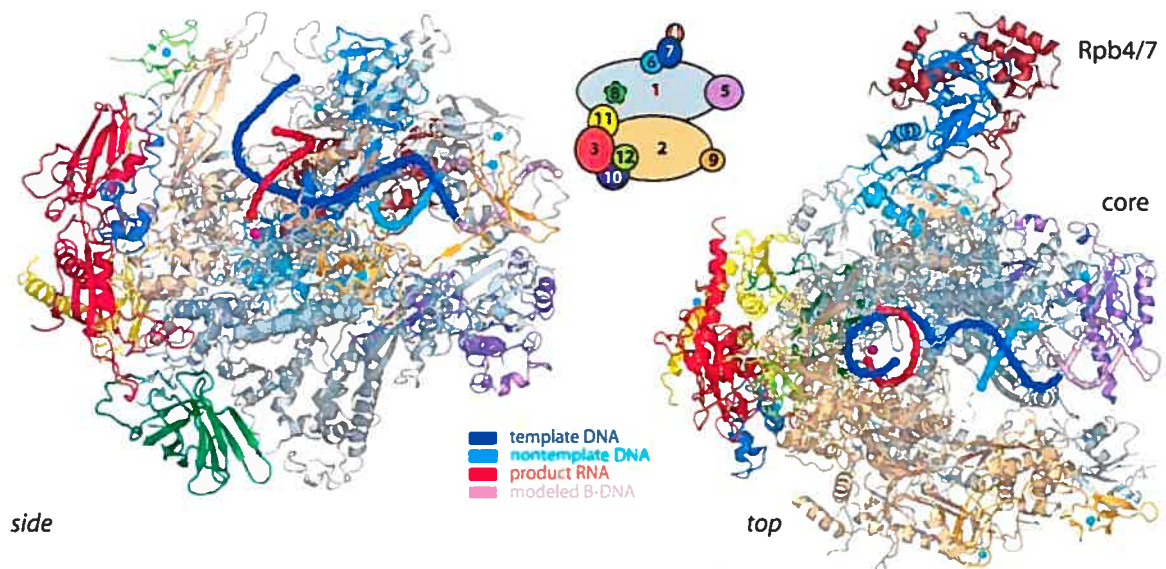


Figure 17 : Structure de l'ARN pol II complète

La structure contient les 12 sous-unités de l'ARN pol II ainsi que l'ARN codé et l'ADN de la matrice (la région mauve représente la structure prédite de l'ADN avant son entrée au site actif).

Image tirée de: Kettenberger 2004

contribut ainsi la stabilité et la processivité au complexe d'élongation (Kireeva 2000).

En amont de l'hybride, la polymérase continue de contacter l'ARN et l'ADN de l'ADN matrice, et ces contacts pourraient être responsables de la séparation des brins et la sortie adéquate de l'ARN (Kettenberger 2004b). Suite à la polymérisation, l'ARN se sépare de l'hybride et entre dans un canal de sortie localisé à la surface de l'enzyme. Les boucles *rudder* et *lid* participent à la séparation de l'ARN de la matrice d'ADN. Le canal de sortie de l'ARNm est localisé tout près d'un domaine en forme de quai nommé « dock » et du dernier résidu ordonné de Rpb1, c'est-à-dire juste avant son domaine CTD. Le *lid* et la *saddle*, une structure en forme de selle située entre le *wall* et la *clamp* (Cramer 2001), forment un tunnel étroit de sortie de l'ARN qui permet seulement le passage d'un brin simple (figure 18) (Westover 2004a). L'extrémité 5' de l'ARN est située à l'entrée de ce tunnel, et son extension à travers le tunnel mènerait à deux canaux potentiels qui lieraient l'ARN et se projettent jusqu'au domaine *dock* de l'ARN pol II (figure 18) (Cramer 2000, Cramer 2001).

Une étude a récemment proposé un modèle d'ouverture de la bulle transcriptionnelle suite à l'observation de paires de bases d'ADN en aval du site actif (Kettenberger 2004b). Contrairement à un modèle où plusieurs molécules de NTPs s'enligneraient avant le site actif par jumelage aux bases d'ADN complémentaires (Zhang 2004), cette équipe a proposé que l'ADN des positions +5 à +9 est essentiellement de la forme canonique B, mais que le brin codant dévie de cette conformation des positions +2 à +4. Cette déviation semble être due à 3 résidus de la boucle *switch 2* (R326, K330 et R337 chez la levure) chargés positivement qui tirent le brin codant loin de l'axe du duplexe, et par 3 résidus de *switch 1* (E1403, E1404 et E1407 chez la levure) chargés négativement qui repoussent le brin d'ADN. Ainsi, une distortion et une déstabilisation du duplexe d'ADN entrant, induites par l'enzyme, pourraient conduire au désappariement des brins. De plus, la boucle *fork2* bloque stériquement la liaison du duplexe, interfère avec le brin non-codant en amont de la position +3 et prévient la réassociation des brins séparés (Kettenberger 2004b). Il est à noter que lors de l'initiation, l'activité hélicase de TFIIF joue un rôle essentiel dans l'ouverture de la bulle de transcription. Les hélicases agissent

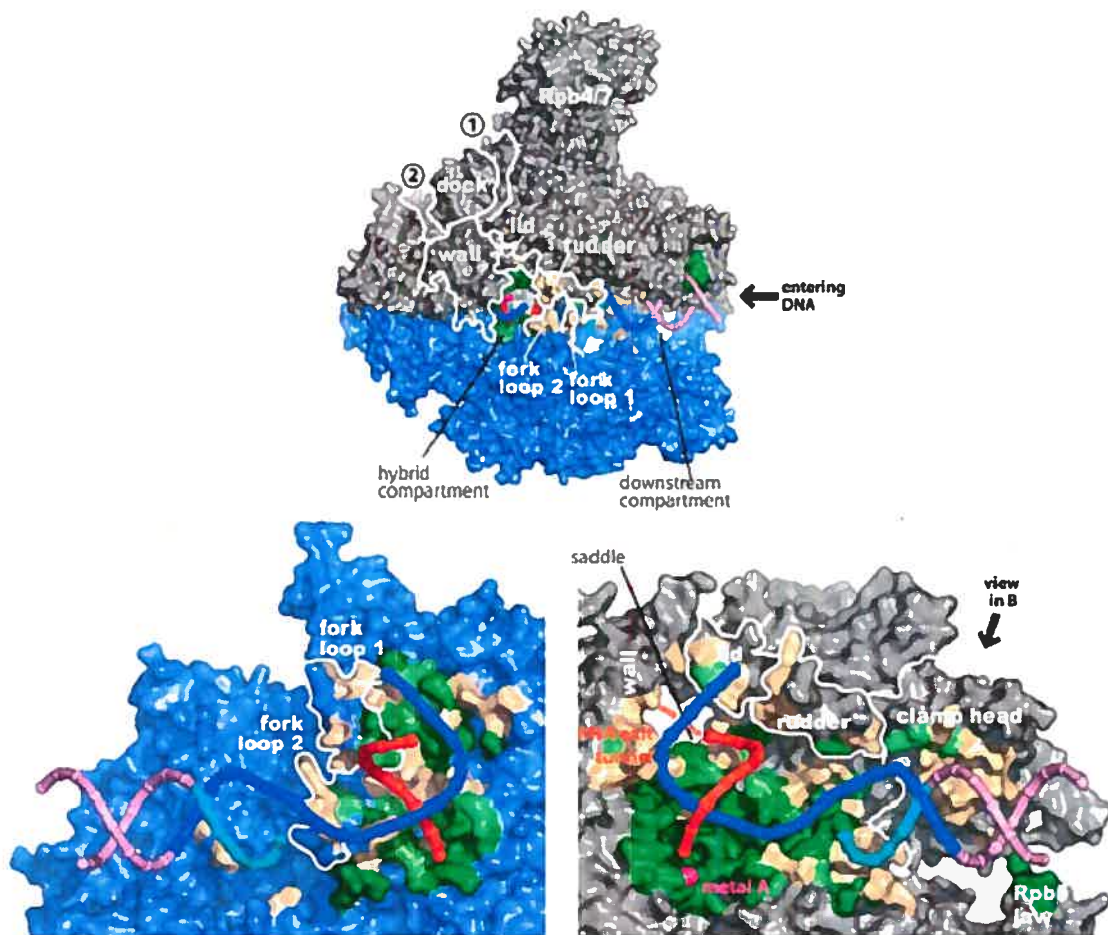


Figure 18 : Conservation des compartiments du sillon de l'ARN pol II

Les boucles de la polymérase qui partitionnent le sillon sont illustrées. Les canaux de sortie de l'ARN proposés précédemment sont identifiés par 1 et 2. Le site actif est ouvert en deux moitiés dans les modèles du bas et les résidus de la surface moléculaire allant jusqu'à 8Å des acides nucléiques sont en beige. Les résidus invariables et conservés entre les ARN pol I, II et III sont en vert foncé et pâle respectivement.

Image tirée de: Kettenberger 2004

en déstabilisant les brins d'acides nucléiques par le mouvement ATP-dépendant de deux domaines qui interagissent avec des régions simple- et double-brins d'acides nucléiques. En se servant de ce mécanisme, l'hélicase XPB de TFIIH se lie au promoteur d'ADN et initie l'ouverture de la bulle en induisant une distorsion dans l'ADN autour du site d'initiation (Woychik 2002). C'est ensuite au tour des boucles de la polymérase à jouer un rôle dans le maintien de la bulle transcriptionnelle lors des étapes d'élongation.

Après l'initiation, la chaîne d'ARN grandissante pourrait déclencher des transitions structurelles qui mènent à la formation d'un complexe d'élongation stable. Tel que mentionné plus tôt, cette chaîne d'ARN grandissante déplace le facteur d'initiation TFIIB, qui se lie au domaine *dock* (Chen 2003) et s'étend au dessus de la *saddle* et jusqu'au site de l'hybride (Bushnell 2004, Chen 2004). La transition entre un PIC fermé à un complexe ouvert inclut le mouvement de l'ADN, qui se trouve à l'intérieur du sillon durant l'élongation mais qui est suspendu au-dessus de ce dernier dans le complexe d'initiation (Armache 2003, Bushnell 2003, Craighead 2002). L'ADN en amont se trouve lié à plusieurs facteurs durant l'initiation mais demeure mobile durant l'élongation (Chen 2004). De plus, la transition est accompagnée de l'enfouissement total de l'hybride ADN-ARN dans le sillon, un point marquant des complexes d'élongation. Deux boucles qui s'étendent de chaque côté du sillon, le *rudder* et la boucle *fork1*, sont mobiles dans la polymérase libre (Cramer 2001) mais deviennent ordonnées et se touchent dans le complexe d'élongation, créant deux éléments dans le *cleft* pour maintenir la bulle de transcription (Westover 2004a). Étant donné l'importance du *lid*, des boucles *fork1* et *fork2* et du *rudder* pour la séparation des acides nucléiques, il est surprenant que ces boucles ne soient pas conservées chez les polymérases I et III, et varient même en longueur. Ces boucles seraient simplement responsables d'un cloisonnement topologique passif du *cleft* de l'ARN pol II (Kettenberger 2004b).

La transition entre l'initiation et l'élongation a été structurellement caractérisée pour la polymérase à sous-unité simple du bactériophage T7, qui ne requiert pas de facteurs d'initiation. Une grande région de cette polymérase se réarrange durant cette transition pour créer un site pour la liaison de l'hybride et un tunnel de sortie pour la chaîne d'ARN synthétisée (Steitz 2004, Tahirov 2002, Yin 2002). Chez l'ARN pol II, le site de liaison de l'hybride et le tunnel de sortie sont déjà définis dans la forme de l'enzyme compétente à l'initiation, mais sont partiellement occupés durant l'initiation avec TFIIB (Bushnell 2004) et possiblement avec TFIIE (Forget 2004, Kim 2000, Meinhart 2003) et TFIIF (Chung 2003). Ces facteurs pourraient tous contribuer à la stabilisation d'un complexe de transcription, mais doivent libérer le site de liaison de l'hybride et le tunnel de sortie de l'ARN durant la formation d'un complexe d'élongation stable. Malgré que les

complexes d'initiation des polymérase de bactériophage et d'eucaryotes sont très différents, l'arrangement des acides nucléiques dans les complexes d'élongation demeure semblable, incluant la longueur de l'hybride ADN-ARN, la courbure de 90° des bases de la matrice aux positions +2 et +1 et la position relative de l'hybride par rapport à l'ADN en aval (Kettenberger 2004b).

1.5.5 Le sous-complexe Rpb4-Rpb7

Le sous-complexe formé par les sous-unités Rpb4 et Rpb7 interagit avec le reste de l'ARN pol II via deux petites interfaces, la plus grande étant la pointe de Rpb7. Cette pointe, formée de deux boucles, interagit avec 5 régions qui forment une « poche » dans l'ARN pol II, formée de 3 régions de Rpb1, une région de Rpb2 et une région de Rpb6 (Armache 2003, Bushnell 2003). L'importance de la région de Rpb6, qui forme une hélice α conservée, a été démontrée biochimiquement alors que des mutants à cette région deviennent incapables de lier l'hétérodimère Rpb4/Rpb7 (Tan 2003). Ce point de contact majeur entre Rpb7 et la poche de la polymérase n'implique pas Rpb4, ce qui explique comment Rpb7 interagit avec la polymérase indépendamment de Rpb4 (Sheffer 1999). Un autre point de contact se trouve dans la région N-terminale de Rpb4 qui interagit avec l'extrémité N-terminale de Rpb1, stabilisant d'avantage l'interaction (Bushnell 2003). À l'exception de ces contacts, l'hétérodimère est projeté vers l'extérieur de la polymérase et est accessible à d'autres facteurs. De plus, Rpb4/Rpb7 est situé à un endroit stratégique sur la surface de l'ARN pol II, près de la sortie du transcrit et adjacent à la région « linker » liant le CTD (Armache 2003, Bushnell 2003, Chung 2003). Cette localisation semble importante puisque Rpb7 interagit avec ESS1, une prolyl isomérase qui lie le CTD et régule l'élongation et la terminaison (Wu 2001b).

Le sous-complexe Rpb4/Rpb7 se retrouve en quantité sous-stoichiométrique par rapport aux autres sous-unités de la polymérase de *S. cerevisiae* (Kolodziej 1991), et tend à se dissocier du complexe (Edwards 1991). Ceci peut s'expliquer par le fait que l'interaction de l'hétérodimère avec la polymérase varie selon les conditions environnantes. Ainsi, chez les cellules en prolifération exponentielle, seulement ~20% des polymérase de *S.*

cerevisiae contiennent Rpb4/Rpb7 (Choder 1993). Par contre, dans les cellules privées de nutriments et qui entrent dans la phase stationnaire, Rpb4/Rpb7 lie la polymérase de manière stœchiométrique (Choder 1993). Il a été proposé que ni Rpb4 ni Rpb7 ne sont sujettes à la phosphorylation (Kolodziej 1991) mais qu'une modification encore inconnue de la polymérase soit responsable de l'interaction de l'hétérodimère dans les situations de déprivation. Plusieurs explications ont été proposées par rapport à cette interaction : premièrement, le sous-complexe Rpb4/Rpb7 pourrait protéger l'ARN pol II en évitant certains changements conformationnels néfastes pouvant survenir durant les périodes de déprivation à long-terme (Choder 1993). Deuxièmement, l'interaction de Rpb4/Rpb7 avec la polymérase est promoteur-spécifique (Khazak 1998, Na 2003) et l'hétérodimère serait recruté dans seulement 20% des initiations dans les cellules en prolifération normale. Troisièmement, l'hétérodimère et la polymérase pourraient interagir seulement à certaines étapes spécifiques du cycle transcriptionnel (Choder 2004). Il est à noter que les homologues de Rpb4/Rpb7 chez *S. pombe* et l'humain s'associent de manière stœchiométrique à leurs polymérases respectives durant les phases de croissance exponentielle (Sakurai 1999, Khazak 1998) et que les différences mécanistiques entre les espèces sont encore méconnues.

1.5.6 Le domaine C-terminal de Rpb1

Le domaine C-terminal (CTD) de Rpb1 est unique à l'ARN pol II et est composé, chez l'humain, de 52 répétitions en tandem d'un heptapeptide Tyr-Ser-Pro-Thr-Ser-Pro-Ser. Cette séquence consensus est hautement conservée à travers les organismes mais le nombre de répétitions semble varier d'un organisme à un autre. La séquence répétée dévie de plus en plus de la séquence consensus à l'approche de l'extrémité C-terminale (Fong 2001). Ces deux observations ont mené à la suggestion que différentes régions du CTD pourraient avoir des fonctions spécifiques. Le CTD s'étend de la base de la pince mobile à l'extérieur de la polymérase et devient désordonné dans les structures cristallographiques après quelques résidus seulement. Si cette région adoptait une conformation complètement étendue, l'extrémité C-terminale aurait une longueur de 1000Å, en d'autres termes, 7 fois le diamètre de l'ARN pol II. Le CTD pourrait en

principe atteindre n'importe quel endroit sur la surface, mais les études démontrent qu'un CTD non-phosphorylé adopte plutôt un état compacté à proximité de la surface de l'enzyme.

Le CTD sert de plate-forme pour l'assemblage des facteurs régulant l'initiation, l'élongation, la modification de l'ARNm et la terminaison, couplant ainsi les différentes étapes de transcription aux étapes de maturation de l'ARNm. 5 des 7 acides aminés de la séquence répétée peuvent en principe être phosphorylés, et c'est par un patron différentiel de phosphorylation que certaines protéines sont recrutées à certains moments de la transcription (Dahmus 1996, O'Brien 1994). L'ARN pol II avec un CTD hypophosphorylé est recrutée au promoteur pour la formation du PIC, et devient phosphorylée sur ses sérines 5 durant l'initiation par la kinase cycline-dépendante CDK7/cycline H (sous-unité du facteur TFIIF). Lors de la transition vers l'élongation, l'ARN pol II est phosphorylée sur ses sérines 2 par la kinase CDK9/cyclines T et K (sous-unité du facteur d'élongation PTEF-b), ce qui entraîne le recrutement à la polymérase en élongation des facteurs de polyadénylation d'ARN et les facteurs de terminaison (Komarnitsky 2000). La phosphorylation de la polymérase non-liée à l'ADN par des kinases telles que ERK et CDK8/cycline C sert plutôt à diminuer la transcription en inhibant la formation du PIC.

D'un point de vue métabolique, les phosphatases du CTD servent donc à recycler les ARN pol II. En effet, les phosphatases déphosphorylent le CTD pour permettre son incorporation adéquate et efficace au PIC et son association au médiateur qui se fait exclusivement via un CTD non-phosphorylé. Fcp1 est la phosphatase responsable de déphosphoryler des sérines 2 (Kobor 1999), et Ssu72, une sous-unité du facteur de clivage et de polyadénylation (CPSF), serait responsable de déphosphoryler les sérines 5 (Krishnamurthy 2004). Fcp1 est la première composante d'une nouvelle famille de phosphatases identifiées ne portant aucune similarité de séquence avec les autres phosphatases, à l'exception d'un motif de signature DXDX(T/V) appartenant à la super-famille des phototransférases et des phosphohydrolases (Collet 1998). La séquence de Fcp1 comprend deux régions conservées. La région N-terminale contenant le motif

FCPH (pour « Fcp1 homology ») (Archambault 1997) inclut le motif de signature DXDX(T/V) avec des résidus importants pour la catalyse, et la région C-terminale contenant le domaine BRCT (pour « breast cancer protein-related carboxy-terminal ») lie le CTD phosphorylé (Yu 2003, Kamenski 2004). Fcp1 est conservé à travers les eucaryotes et est essentiel pour la viabilité cellulaire (Kobor 1999). La déphosphorylation du CTD par Fcp1 peut faciliter le recyclage de la forme hyperphosphorylée de la polymérase pour un nouveau cycle de transcription (Cho 1999). Fcp1 déphosphoryle la polymérase en solution (Chambers 1995, Cho 1999) et lorsqu'elle est associée au complexe d'élongation (Cho 1999, Lehmann 2000). Cette activité phosphatase est stimulée par TFIIF et est inhibée par TFIIB (Chambers 1995). Ensemble, ces résultats indiquent que Fcp1 est un régulateur central du cycle de transcription, mais sa structure et mécanisme d'action demeurent peu compris. Récemment, une deuxième phosphatase, Scp1, a été identifiée chez les eucaryotes supérieurs (Yeo 2003) et sa structure cristallographique a été résolue (Kamenski 2004). Scp1 démontre une grande homologie au domaine catalytique FCPH de Fcp1 mais ne contient pas le domaine BRCT. La structure démontre un repliement central et un centre actif semblables à ceux des phosphotransférases et phosphohydrolases qui partagent le motif de signature DXDX(T/V) avec Fcp1/Scp1. Le premier aspartate du motif subit une phosphorylation assistée par un métal durant la catalyse, résultant en un phospho-aspartate intermédiaire (Kamenski 2004). La spécificité de la réaction semble découler de la liaison du CTD à une poche hydrophobe entre le site actif et le domaine d'insertion unique à Fcp1/Scp1. De plus, la spécificité pourrait provenir du recrutement de la phosphatase près du CTD via le sous-complexe Rpb4/Rpb7, requis pour la liaison de Fcp1 *in vitro*, tel que mentionné plus tôt.

1.5.7 Cristallographies de l'ARN pol II avec les FGTs

La structure cristallographique du facteur TFIIB en association avec l'ARN pol II a démontré que le ruban de zinc de TFIIB lie le domaine *dock* de l'ARN pol II (Bushnell 2004), suggéré auparavant de lier les facteurs d'initiation (Cramer 2001). Le ruban de zinc ne contacte pas l'ADN du promoteur (Chen 2003), ce qui renforcerait le chargement

de l'ADN au dessus du sillon. Par contre, après l'ouverture de l'ADN, TFIIB semble contacter les deux brins de l'ADN dans la bulle de transcription, tel que suggéré par les études de pontage d'ADN aux protéines du complexe transcriptionnel d'archéobactérie (Renfrow 2004). La forme cristallographique démontre qu'une extension de TFIIB atteint le site de liaison de l'hybride (Cramer 2004). Le ruban de zinc pourrait donc stabiliser la présence d'un court transcrit dans le complexe de transcription pendant l'initiation. L'ARN en synthèse entrerait éventuellement en contact avec TFIIB et pourrait forcer le déplacement et le relâchement de TFIIB et ainsi faciliter le dégagement du promoteur (Chen 2003). De la même manière, le facteur d'initiation bactérien σ pénètre le site de liaison de l'hybride dans la polymérase (Murakami 2002, Vassylyev 2002), suggérant que des parties de σ et de TFIIB possèdent des fonctions analogues.

Le chargement du promoteur sur l'ARN pol II requiert la participation du facteur TFIIF (Killeen 1992, Tyree 1993). Une étude récente de microscopie électronique d'un complexe ARN pol II-TFIIF a démontré que la deuxième plus grosse sous-unité de TFIIF s'étend au-dessus et le long du sillon de la polymérase, de manière semblable à l'interaction de son homologue par séquence, le facteur σ de la polymérase bactérienne (Chung 2003). Certaines régions de TFIIF, attribuées à sa plus grosse sous-unité, lient le complexe Rpb4/Rpb7 (Chung 2003) et cette interaction est observée également chez les homologues des polymérases I et III (Ferri 2000, Juo 2003). Ainsi, Rpb4/Rpb7 relie la polymérase aux facteurs d'initiation, et contribuent à la reconnaissance des promoteurs (Cramer 2004).

La comparaison entre la structure du complexe d'ARN pol II et celle de la polymérase bactérienne liée au facteur σ et au promoteur (Murakami 2002) démontre une ressemblance au niveau de la topologie générale du chargement du promoteur sur les polymérases. Des domaines de TFIIF et TFIIB lient la polymérase à des régions liées également par le facteur σ et possèdent apparemment des fonctions semblables. TBP, TFIIB et des domaines de TFIIF s'assemblent à la surface d'entrée de l'ARN de la polymérase, là où l'ADN du promoteur se situe dans les complexes bactériens. Certains mécanismes régulateurs peuvent aussi être conservés entre les ARN pol II et la

polymérase bactérienne. La microscopie électronique démontre que le complexe co-activateur Médiateur se lie près de l'hétérodimère Rbp3-Rpb11 de la polymérase, et que la surface correspondante sur la polymérase bactérienne est importante pour la régulation transcriptionnelle (Davis 2002).

Suite à l'assemblage du complexe ARN pol II-TFIIF-TFIIB-TBP-ADN, l'ADN du promoteur est désapparié avec l'aide de TFIIE et TFIIH. Les mécanismes correspondants sont encore peu compris mais impliquent des changements structuraux importants. Cette idée est renforcée par l'observation que la polymérase à sous-unité unique du bactériophage T7 subit un changement de repliement dramatique durant la transition de l'initiation à l'élongation (Yin 2002). Une étude structurale récente démontre que la région N-terminale de la grosse sous-unité de TFIIE, conservée et essentielle, forme un domaine en hélice-tour-hélice qui recrute TFIIE à la polymérase (Meinhart 2003). Ce motif se trouve souvent chez les protéines liant l'ADN mais dans ce cas, c'est plutôt la petite sous-unité de TFIIE qui contacte l'ADN (Okuda 2000). Le facteur TFIIF, quant à lui, possède deux de ces motifs, un liant l'ADN (Kamada 2001) et l'autre liant le domaine en hélice C-terminale de la phosphatase Fcp1 (Kamada 2003, Nguyen BD 2003).

Conclusion

Une des plus importantes contributions des travaux présentés durant les dernières années est que les gros complexes multi-protéiques impliqués dans la transcription ressemblent à des casses-têtes formés par des pièces flexibles, dans lesquels l'image est compréhensible seulement lorsque toutes les composantes sont présentes. Cette perception devra guider la conception et l'interprétation des études structurales futures de la machinerie transcriptionnelle des eucaryotes.

1.6 Objectifs et hypothèses

L'objectif général poursuivi par notre équipe est d'analyser l'organisation moléculaire de l'ARN pol II humaine. Notre laboratoire étudie la relation entre la structure et la fonction des composantes de la machinerie transcriptionnelle au cours des diverses étapes de la transcription. Notre objectif ultime est de construire un modèle moléculaire détaillé du processus transcriptionnel, une tâche compliquée par le grand nombre de polypeptides impliqués dans la transcription et par l'aspect dynamique de la réaction transcriptionnelle. De plus, nous cherchons à comprendre comment la transcription est régulée par la séquence et la structure de la matrice d'ADN, ainsi que par les activateurs transcriptionnels dont plusieurs jouent un rôle important dans la croissance et la différenciation cellulaire, ainsi que dans l'établissement de maladies comme le cancer et les infections virales.

Au centre du complexe polypeptidique responsable de la synthèse de l'ARN se trouve un site catalytique hautement régulé, conservé à travers les espèces et formé par les deux plus grosses sous-unités de l'ARN pol II. L'objectif du présent projet consiste à investiguer le rôle de différentes boucles de la sous-unité Rpb1 qui se localisent de manière stratégique aux environs du centre catalytique de la polymérase. Nous avons émis l'hypothèse que certaines structures ou acides aminés conservés jouent un rôle spécifique dans une ou plusieurs des multiples étapes de la transcription. Nous avons donc procédé à la création de plusieurs mutations dirigées et observé que, parmi les mutants capables de s'assembler et de former correctement une enzyme à 12 sous-unités, la plupart sont bloqués aux étapes précoces du mécanisme de transcription.

Ainsi, une série de mutants portant des mutations ou des substitutions à l'intérieur des éléments structuraux entourant le site actif ont été créés et caractérisés. L'ARN pol II de type sauvage et ses formes mutantes ont été purifiées par affinité à partir de cellules humaines à l'aide de l'étiquette TAP (pour « Tandem Affinity Peptide ») placée à l'extrémité C-terminale de Rbp1, Rpb4 ou de Rpb7.

Deux polymérases mutantes ayant des substitutions d'alanine dans les acides aminés des domaines *switch* 1 et 2 de Rpb1, proposés pour entrer en contact direct avec les nucléotides, ont été étudiées. En effet, la boucle *switch* 1, dont la structure montre un contact avec la matrice d'ADN près de la localisation de l'hybride avec l'ARN et ses résidus S1431, F1432 et E1433, équivalents aux résidus S1401, F1402 et E1403 de la levure, ont été substitués simultanément par des alanines. Nos résultats montrent que cette mutation interfère probablement avec le repliement de la sous-unité Rpb1, l'empêchant de s'assembler à la polymérase tel que démontré par la purification par chromatographie d'affinité migrée sur gel de polyacrylamide dénaturant et révélé au nitrate d'argent. La boucle *switch* 2, quant à elle, entre en contact avec l'hybride ADN-ARN dans le site actif ; ses résidus K346, E347, G348, R349, V350 et R351 (équivalents aux résidus K332, E333, G334, R335, I336 et R337 de la levure) ont été substitués simultanément par des alanines. Cette mutation n'a pas interféré avec le repliement de la sous-unité Rpb1, qui a pu s'assembler en polymérase, mais a empêché la polymérase de former correctement le PIC sur le promoteur et d'entreprendre la transcription dans un système *in vitro* sur le promoteur majeur tardif d'adénovirus.

Deux autres polymérases mutantes ayant des délétions de plusieurs acides aminés des domaines *lid* et *zipper* de Rpb1 ont également été étudiées. Les boucles *lid* et *zipper* sont hautement conservées chez les polymérases bactériennes, de levure et humaine et s'étendent de la base du *clamp* pour participer à la séparation de l'hybride ADN-ARN et maintenir la partie en amont de la bulle de transcription dans le site actif. La délétion partielle du *lid* détruit complètement la capacité de Rpb1 à s'assembler avec les autres sous-unités de l'ARN pol II. Par contre, la délétion complète du *zipper* a seulement affecté légèrement l'assemblage de l'enzyme, surtout au niveau des sous-unités Rpb4 et Rpb7. D'autres expériences *in vitro* ont révélé que cette délétion empêche le PIC non seulement de s'assembler sur un promoteur et d'entreprendre la transcription mais prévient également l'élongation indépendante des étapes d'initiation.

Ces résultats confirment que les boucles conservées du site catalytique jouent un rôle crucial dans la formation d'un complexe d'ARN pol II stable, l'assemblage au promoteur

et la processivité de la transcription. L'intégration des résultats de ce projet permettra de décrire un modèle moléculaire plus détaillé de la transcription par l'ARN pol II humaine.

Chapitre 2 : Matériels et méthodes

2.1 Construction des plasmides codant pour la sous-unité Rpb1 de l'ARN pol II fusionnée à l'étiquette TAP

Le clonage est fréquemment utilisé pour étudier la structure et la fonction des gènes ainsi que pour modifier leur niveau d'expression. De façon générale, les clones sont construits par ligation d'un vecteur et d'un fragment d'insert de type sauvage ou portant une mutation. Les constructions sont ensuite transformées dans des cellules bactériennes compétentes, cultivées et sélectionnées par résistance à certains antibiotiques. Le plasmide d'ADN est isolé à partir de ces bactéries, clivé par des enzymes de restriction et analysé sur gel d'agarose par électrophorèse. Les plasmides sont séquencés pour assurer la présence de l'insert avec la mutation désirée.

Le cDNA codant pour la sous-unité Rpb1 humaine (Invitrogen) est clonée dans le vecteur d'expression pMZI (No 1996) de manière à fusionner son extrémité C-terminale à l'étiquette TAP (pour « Tandem Affinity Peptide ») (discuté en détail plus bas). Pour les clones mutants, le cDNA de type sauvage est d'abord cloné dans un petit vecteur d'expression pBS (« plasmid Blue Script » de Stratagene) de 2.9 kb au site NheI. La mutagenèse dirigée est obtenue par PCR inverse en utilisant des amorces phosphorylées à leur extrémité 5' et qui contiennent les mutations désirées (figure 19). Le produit d'amplification qui consiste en la séquence de la sous-unité mutée et du vecteur, doit être ligué, puis transformé dans les cellules *E. coli* compétentes XL-1.

Les bactéries ayant acquis le plasmide d'intérêt sont sélectionnées par le gène de résistance à l'ampicilline porté sur le vecteur pBS. La ligation par la « T4 DNA ligase » (Life) catalyse la formation du lien phosphodiester par la condensation du phosphate en 5' et du groupe hydroxyl en 3'. Le plasmide de différents clones est extrait (Miniprep de Qiagen), linéarisé par digestion avec l'enzyme de restriction à NheI (Amersham) et migré sur gel d'agarose. Les clones positifs montrent deux bandes, l'une à 6 kb (codant pour Rpb1) et l'autre à 3 kb (pour le vecteur). Un second PCR est réalisé sur les clones

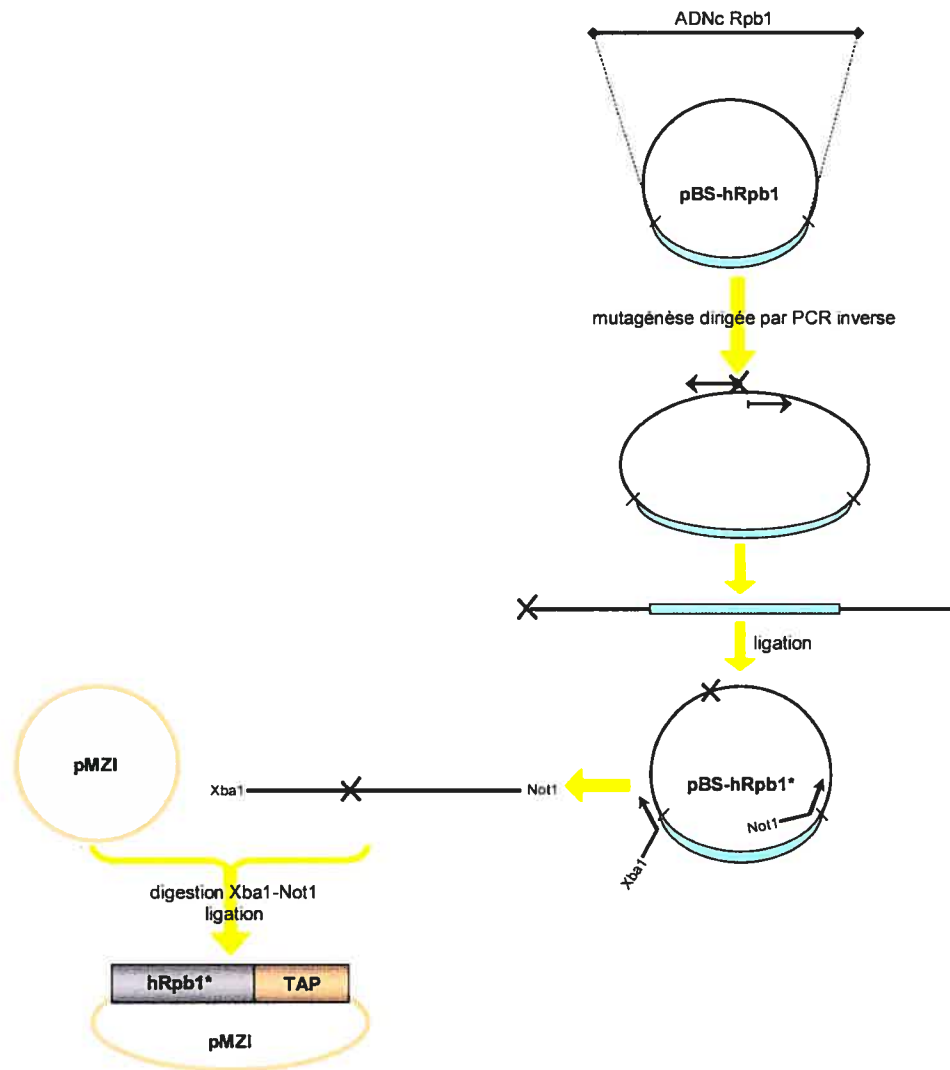


Figure 19 : Construction des plasmides codant pour la sous-unité Rpb1 de l'ARN pol II fusionnée à l'étiquette TAP.

positifs afin d'insérer les sites de restriction XbaI à l'extrémité 5' et NotI à l'extrémité 3'. Ces sites permettent le sous-clonage de Rpb1 dans le plasmide pMZI entre un promoteur inductible et la séquence codant pour l'étiquette TAP. L'amorce introduisant NotI modifie aussi le codon Stop de la séquence Rpb1 ce qui joint l'étiquette TAP à l'extrémité C-terminale de Rpb1. L'insert et le vecteur sont digérés, ligés et transformés dans les bactéries compétentes. Plusieurs clones sont sélectionnés par le gène de résistance à l'ampicilline, et leur plasmide extrait est digéré par l'un des deux enzymes de restrictions utilisés lors du clonage. Les clones positifs montrent un bande de 12 kb (6 kb appartenant au vecteur pMZI et 6 kb appartenant à l'insert) et sont séquencés de nouveau.

2.2 Le système d'induction

Le système d'induction (Invitrogen 25-0243) est basé sur la capacité de l'hormone d'insecte 20-OH ecdysone d'activer l'expression de gènes via son propre récepteur (No 1996). La technique utilise un hétérodimère formé d'une part par le récepteur d'ecdysone (EcR dérivé de la *Drosophila* et modifié pour contenir le domaine de transactivation VP16), et d'autre part, par le récepteur X rétinoïde (RXR dérivé des cellules mammifères). En présence du ligand, l'ecdysone ou un de ses analogues la muristérone A ou la ponastérose A, l'hétérodimère lie l'élément de réponse à l'ecdysone modifié (RE). Cette liaison de l'hétérodimère à RE active la transcription. Puisque les cellules mammifères ne répondent pas normalement à l'ecdysone ou à l'un de ses analogues et ne contiennent pas le récepteur d'ecdysone, les niveaux de transcription basale demeurent absents, ou très bas (figure 20).

EcR 293

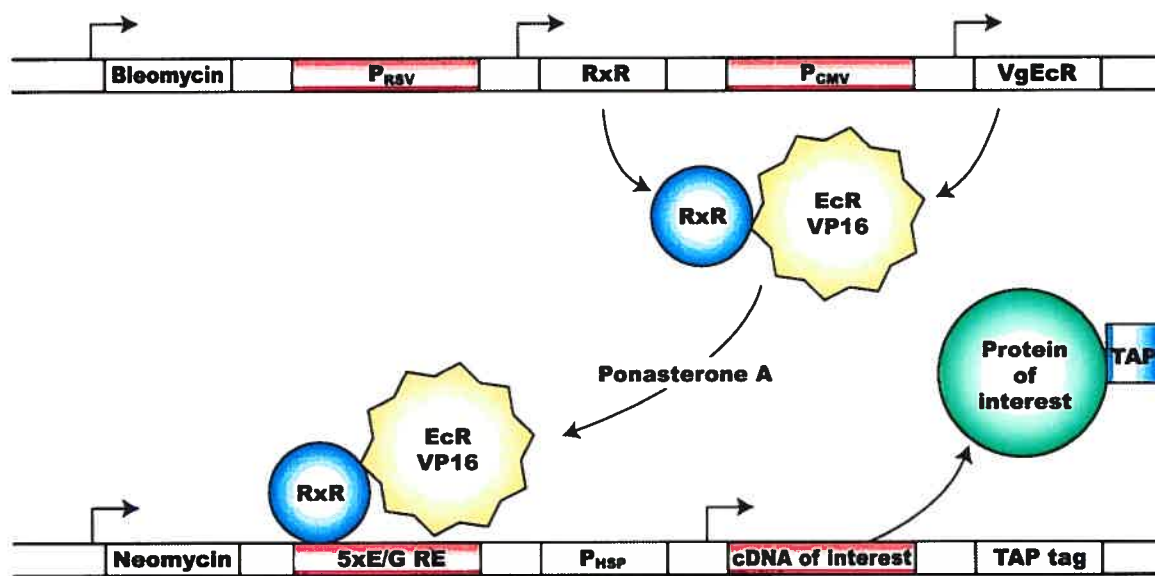


Figure 20 : Système d'induction à l'ecdysone (Voir l'explication dans le texte)

Image tirée de: Catalogue Invitrogen 20-0243

Les deux sous-unités d'un récepteur d'ecdysone de *Drosophila* sont donc exprimées constitutivement à partir d'un vecteur d'expression stable, EcR 293. Le promoteur répondant à l'ecdysone (p-HSP), qui régule ultimement l'expression du gène cloné, est situé sur un deuxième vecteur d'expression inductible. La lignée cellulaire utilisée a été transfectée avec le vecteur régulateur EcR 293 qui contient le gène de résistance Bléomycine pour faciliter la sélection des cellules EcR-293 transfectées de manière stable. Le vecteur d'expression inductible contenant 5 éléments de réponse modifiés (5x E/G RE) en amont d'un promoteur minimal, suivi du gène d'intérêt, en plus des gènes de résistance Néomycine et Ampicilline est ensuite transfecté. En présence de l'inducteur, le récepteur d'ecdysone fonctionnel lie le promoteur répondant à l'ecdysone et active l'expression du gène d'intérêt. Ce système assure l'expression régulée et sans fuite de la protéine d'intérêt. La ponastérone A a été choisie dû à sa plus grande disponibilité par rapport à l'ecdysone et à son efficacité semblable. La concentration requise pour obtenir un niveau d'expression du gène d'intérêt semblable au niveau physiologique avait été déterminé précédemment en comparant les niveaux du polypeptide fusionné à l'étiquette TAP et le polypeptide endogène par immunobuvardage (figure 21b).

2.3 Transfection stable de cellules EcR

Deux types de transfection permettent l'introduction d'ADN dans une cellule, la méthode transitoire est la plus rapide car l'ADN n'est pas intégré dans le génome des cellules hôtes mais le nombre de plasmides qui entrent au noyau ne peut être contrôlé et ne se retrouve dans le noyau que pour quelques jours seulement. La méthode stable dépend de l'intégration de l'ADN dans le génome des cellules. Un marqueur de résistance aux antibiotiques permet la sélection des cellules contenant le transgène d'intérêt suivant deux semaines de culture *in vitro* en présence d'antibiotique.

L'ADN présenté comme un précipité de phosphate de calcium ($CaPO_4$)-ADN aux cellules en culture améliore grandement l'efficacité de transfection. Cette méthode est une technique simple et peu coûteuse qui peut servir pour les transfactions stables et

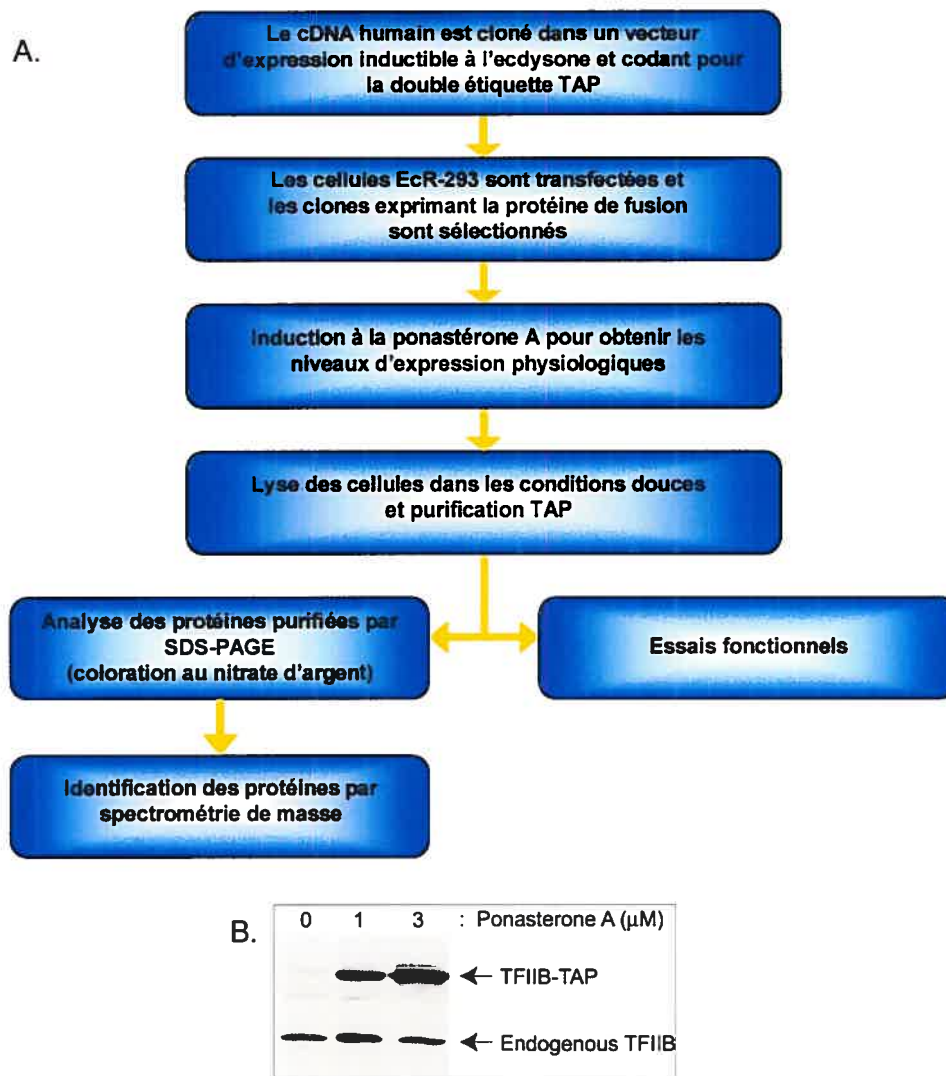


Figure 21 : Méthode de purification par affinité en tandem (TAP)

A. Schéma général de la procédure TAP

B. Résultat d'une expérience pilote comparant les niveaux d'expression du polypeptide TFIIB fusionné à TAP avec les niveaux endogènes après l'induction à la ponastérone A. L'expression de TFIIB-TAP dans des cellules EcR-293 après induction de 24 heures à l'aide de différentes concentrations de ponastérone A (0, 1 et 3 μM) est comparé au niveau de TFIIB endogène par immuno-buvardage avec un anticorps contre TFIIB (Image tirée de: Jeronimo 2004)

transitoires (Graham 1973). 8 μg d'ADN, obtenu par la trousse de purification de Qiagen (« QIAfilter Plasmid Maxi Kit »), sont incubés dans un volume de 1 mL en présence de 122 mM CaCl_2 et 50% de tampon HeBS 2X (1.5 mM $\text{Na}_2\text{HPO}_4 \cdot 7\text{H}_2\text{O}$, 274 mM NaCl, 54.5 mM HEPES; pH 7.0) durant 10 min. à température ambiante. La suspension est

ensuite ajoutée goutte par goutte à une couche de cellules EcR-293 adhérentes et confluentes à 60% dans un pétri de 100 mm et les cellules sont incubées 12 à 14 heures dans un incubateur humidifié à 37°C et 5% de CO₂. L'incubation est d'abord réalisée dans un milieu non-sélectif pour permettre l'intégration du plasmide d'intérêt, les cellules sont ensuite transférées dans un milieu adéquat pour sélectionner les cellules ayant intégrées le plasmide (milieu DMEM de GIBCO, 10% sérum (« Fetal Bovine Serum » de GIBCO), 2 mM glutamine, 30 µg/ml bléocine et 300 µg/ml généticine).

Les cellules sont cultivées dans le milieu sélectif durant 10 à 14 jours, avec des changements de milieu fréquents pour éliminer les cellules mortes et les débris, jusqu'à ce que des colonies distinctes soient visualisées. Les colonies individuelles sont récupérées et transférées dans des plaques à plusieurs puits pour isoler les clones tout en les faisant pousser dans un milieu sélectif. Une fraction de chaque colonie est assujettie à l'induction de l'expression du gène d'intérêt par la ponastérone A (Invitrogen). Ces cellules sont lysées dans un tampon de lyse par congélation à -80°C et décongélation (tampon 0.1% Triton X-100, 0.1 M Tris-HCl pH 8.0), et les différents échantillons sont analysés sur gel de polyacrylamide 12.5% en conditions dénaturantes (SDS-Page) et testées par la méthode d'immunobuvardage afin de détecter l'expression de la protéine d'intérêt. L'immunobuvardage a été effectué par transfert de protéine de type « Western » sur une membrane de PVDF (polyvinylidène difluorure de Millipore). Le premier anticorps utilisé est l'anticorps Pol II (C21) sc-900 (Santa Cruz) contre le domaine CTD de la sous-unité Rpb1 de l'ARN pol II et le deuxième anticorps est un anti-IgG de lapin conjugué à la HRP (Horseradish peroxidase) (Amersham Pharmacia Biotech). Les cellules exprimant le gène d'intérêt de manière stable peuvent être congelées dans l'azote liquide pour utilisation future.

2.4 Purification TAP des complexes contenant l'ARN pol II humaine

La méthode de purification via l'étiquette TAP (Tandem Affinity Peptide) demande la fusion de l'étiquette à une protéine d'intérêt et l'introduction de la construction dans une cellule hôte (Rigaut 1991). Pour des résultats optimaux, l'expression de la protéine de

fusion a été maintenue près des niveaux physiologiques puisqu'une surexpression induit la protéine à former des associations non-spécifiques. Les cellules sont lysées (Xiao 1994) et la protéine de fusion est isolée, avec ses partenaires protéiques associés, par deux étapes de purification par affinité et élution. La purification TAP permet d'isoler des complexes protéiques actifs et complets puisqu'elle est effectuée sous les conditions physiologiques natives, à faible concentration de sels et à l'aide d'étiquettes à forte spécificité et affinité (Rigaut 1991, Jeronimo 2004).

L'étiquette TAP consiste en deux domaines de liaison IgG de la protéine A de *Staphylococcus aureus* et un peptide liant la calmoduline (CBP pour « calmodulin binding protein »), séparés par un site de clivage de la protéase TEV (figure 22A). La protéine A lie fortement une matrice d'IgG, nécessitant l'utilisation de la protéase TEV pour éluer les protéines dans les conditions natives (figure 22B). L'éluat de cette première étape de purification par affinité est incubé par la suite avec des billes de calmoduline en présence de calcium. Après plusieurs lavages pour enlever les contaminants et la protéase TEV, le matériel lié aux billes est relâché sous des conditions douces à l'aide d'EGTA qui chélate le calcium et permet aux billes de relâcher les complexes protéiques.

À partir de cellules transfectées exprimant correctement la protéine d'intérêt, des extraits cellulaires sont obtenus après broyage des cellules dans le tampon A (contenant 10 mM d'HEPES pH 7.9, 1.5 mM MgCl₂, 10 mM KCl, 0.5 mM DTT, 0.5 mM AEBSF (Bioshop) et un inhibiteur de protéase sans EDTA (Roche)), suivi du tampon B (contenant 50 mM d'HEPES pH 7.9, 1.5 mM MgCl₂, 0.5 mM DTT, 0.5 mM AEBSF, 1.26 M d'acétate de potassium et 75% de glycérol). Après une ultracentrifugation pour enlever les membranes et précipités cellulaires, l'extrait est dialysé pendant environ 16 heures dans le tampon de dialyse 100G contenant 10 mM HEPES pH 7.9, 0.1 mM EDTA, 0.1 mM

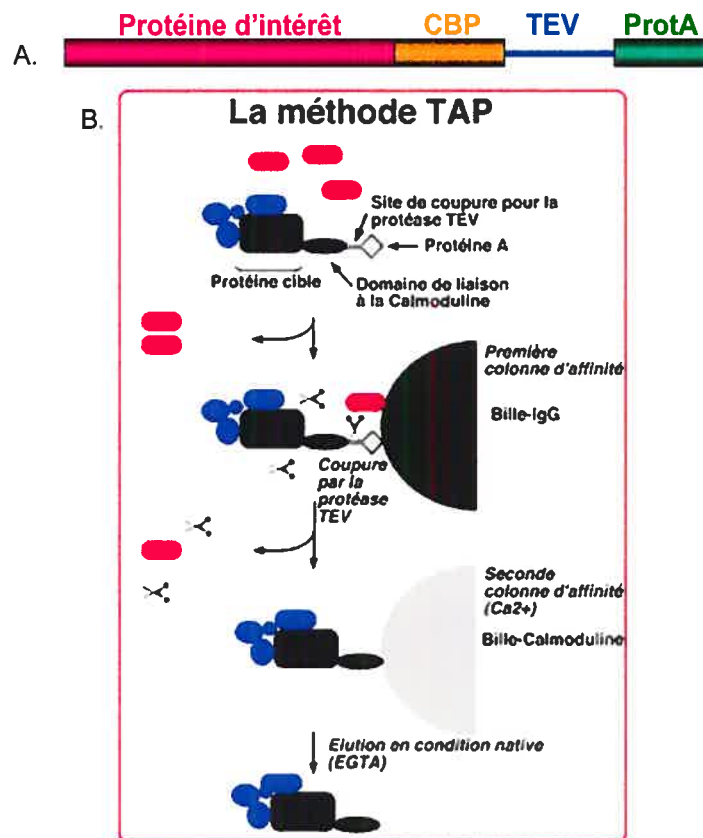


Figure 22 : Protocole de la méthode TAP

A. Représentation schématique de l'étiquette TAP fusionnée à une protéine d'intérêt à son extrémité C-terminale

B. Survol de la stratégie de la purification TAP

Images adaptées de: Puig 2001

DTT, 0.1 mM d'acétate de potassium et 10% de glycérol et est exposé à des billes d'IgG-Sépharose 6 (Amersham) durant une heure. Les protéines liées aux billes sont relâchés par incubation avec la protéase du virus Tobacco etch (TEV de Invitrogen) dans 10 mM Tris-HCl pH 8, 100 mM NaCl, 0.1% Triton X-100, 0.5 mM EDTA, 10% glycérol et 1 mM DTT. Les protéines éluées sont ensuite passées à travers une deuxième colonne d'affinité contenant les billes de calmoduline-Sépharose 4B (Stratagene) en présence de 1.25 mM de calcium pendant deux heures. Les complexes protéiques sont élués à l'aide de 10 mM Tris-HCl (pH 8), 100 mM NaCl, 2 mM EGTA, 10 mM β -mercaptoethanol, 1 mM imidazole et 10% glycérol, et concentrés soit par lyophilization pour les essais non-

fonctionnels, soit par dialyse contre du polyéthylène glycol (PEG-800 de Fisher Scientific) pour les essais fonctionnels. Toutes les étapes de la purification sont effectuées à 4°C ou sur la glace pour prévenir la dénaturation des protéines ou des complexes isolés.

2.5 Identification des protéines par spectrométrie de masse

Les éluats de la purification TAP sont migrés sur gel dénaturant contenant le sulphate dodecyl sodium (SDS) et révélés au nitrate d'argent. L'analyse préalable du gel permet de vérifier la composition grossière du complexe isolé; ensuite, les bandes de peptides sont découpées des gels, réduites, alkalisées et digérées « in-gel » avec la trypsine tel que décrit précédemment (Krogan 2002). Les bandes résultantes sont identifiées par spectrométrie de masse.

2.6 Purification de facteurs généraux de transcription

Les protéines γ TBPr (Ingles 1991), hTFIIBr (Ha 1991), RAP74 et RAP30 (Finkelstein 1992), TFIIE56/ α et TFIIE34/ β (Ohkuma 1991, Peterson 1991, Sumimoto 1991), TFIIH (Douziech 2000, Gerard 1997) et l'ARN pol II de thymus de veau (Hodo 1977) ont été préparées tel que décrit précédemment.

2.7 Essais de transcription

Les essais de transcription (figure 23A) (Burton 1986) consistent à incuber TBP (30 ng), TFIIB (30 ng), RAP30 (30 ng), RAP74 (65 ng), TFIIE34 (40 ng), TFIIE56(60 ng), TFIIH (4 ng) et 165 ng d'ARN pol II de thymus de veau ou humaine portant l'étiquette TAP avec 500 ng de matrice d'ADN linéarisé, contenant le promoteur d'adénovirus majeur tardif (AdML) de -50 à +10 nucléotides, et fusionné à une cassette sans guanine (« G-less ») (Sawadogo 1985). À l'ajout des ribonucléotides ATP, CTP et UTP et de MgCl₂ comme cofacteur, un transcrit de 391 nucléotides est produit.

La réaction de 20 μL , contenant 20 mM d'HEPES, pH 7.9, 50 mM KCl, 10% glycérol, 0.1 mM EDTA, 5 mM β -mercaptoethanol, 200 $\mu\text{g/ml}$ BSA (albumine de sérum de bœuf), et 8 mM MgCl_2 , est initié avec l'addition de 750 μM ATP, 750 μM CTP, 10 μM UTP et 2.5 μCi [α - ^{32}P]UTP et est incubée 30 min. à 30°C. La réaction est ensuite arrêtée avec l'addition de 20 μL d'une solution Stop contenant 0.1 M NaAc pH5.0, 0.5% SDS, 2 mM EDTA et 100 $\mu\text{g/ml}$ ARNt de levure. Les chaînes d'ARN sont ensuite précipitées à l'aide d'éthanol et analysées sur un gel dénaturant de 4% de polyacrylamide (acrylamide : bisacrylamide = 20:1) contenant 8.3 M d'urée, migré à 300V durant 1 heure 30 min. et séché 2 heures à 78°C.

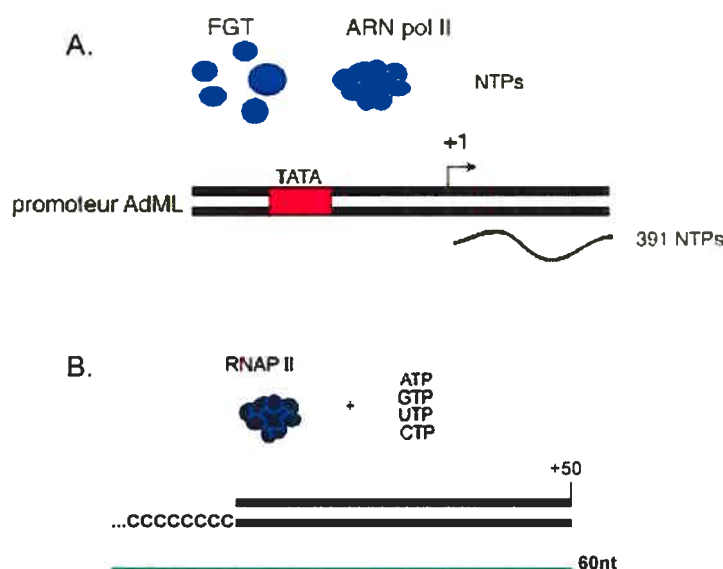


Figure 23 : Essais fonctionnels

A. Schéma d'essai de transcription *in vitro*. Les FGTs et l'ARN pol II sont incubés en présence de matrice d'ADN linéarisée contenant le promoteur d'adénovirus majeur tardif (AdML) de -50 à +10 nucléotides, et fusionné à une cassette sans guanine; un transcrit de 391 nucléotides est produit.

B. Schéma d'essai d'élongation *in vitro*. L'ARN pol II est incubée avec une matrice d'ADN formée de 2 brins appariés et laissant une extrémité 3' libre. La polymérase se lie à l'extrémité libre et entreprend l'élongation sans l'aide des FGTs.

2.8 Essai de retardement sur gel

L'essai de retardement sur gel sert à vérifier la formation d'un PIC contenant l'ARN pol II et les FGTs à un promoteur (Langelier 2001). Un plasmide d'ADN contenant le promoteur d'adénovirus majeur tardif a été digéré à l'aide des enzymes de restriction *Bam*HI et *Dra*I, et le fragment de 110 pb contenant le promoteur a été marquée à l'aide du fragment Klenow de l'ADN polymérase d'*E. Coli* en présence de [α - 32 P]dGTP pour produire une sonde. Les protéines hautement purifiés TBP (40 ng), TFIIB (60 ng), RAP30 (320 ng), RAP74 (640 ng), TFIIE34 (160 ng), TFIIE56 (320 ng) et 150 ng d'ARN pol II de thymus de veau, ou 450 ng d'ARN pol II humaine portant l'étiquette TAP ont été incubés avec 3000 cpm de la sonde marquée et 0.25 ng d'oligonucléotide compétiteur dIdC. La réaction de 20 μ L contenant 12 mM d'HEPES, 6 mM KCl, 0.12 μ M EDTA, 5.72 μ M β -mercaptoéthanol, 8 mM MgCl₂, 1 μ g BSA et 24% de glycérol, a été incubée 30 min. à 30°C, suivi de 5 min. sur la glace. Les complexes ont ensuite été passés sur un gel refroidi de 4.5% polyacrylamide (40:1 ratio d'acrylamide à bisacrylamide), 2 mM MgCl₂, 0.5X Tris-borate-EDTA (TBE), dans un tampon 0.5X TBE, 2 mM MgCl₂, également refroidi, durant 1 heure à 400V et à 4°C. Le gel a été séché et visualisé par Phosphor-Imager (Molecular Dynamics Inc.) ainsi que par exposition sur un film Biomax MS (Kodac Inc.).

2.9 Essai d'élongation

L'essai d'élongation utilise un duplexe d'ADN portant une extension simple-brin en 3' reconnu et transcrit par l'ARN pol II qui permet la transcription en absence de FGTs et en évitant les étapes limitantes de la reconnaissance du promoteur et de l'ouverture de la bulle de transcription (Kadesch 1982). La matrice a été préparée en hybridant 2 oligonucléotides de séquences complémentaires, dont l'un contient une extension de 15 CMP, ou cytidine monophosphate (Pan 1994). Les brins transcrits et non-transcrits (ratio de 1:1.1) ont été incubés dans une réaction d'hybridation de 12 μ L contenant 25 mM Tris-Cl, pH 8.0, 50 mM KCl, et 8 mM MgCl₂, chauffée à 90°C pendant 10 min., et refroidie à température ambiante pendant 5 heures. Durant l'essai d'élongation (figure

23B) (Pan 1994), 10 ng de matrice ont été incubés avec 165 ng d'ARN pol II de thymus de veau ou humaine portant l'étiquette TAP. La réaction a été initiée avec l'addition de 750 μM ATP, 750 μM CTP, 750 μM GTP, 10 μM UTP et 3.75 μCi de $[\alpha\text{-}^{32}\text{P}]\text{UTP}$ et est incubée 30 min. à 30°C. La réaction est ensuite arrêtée avec l'addition de 20 μL d'une solution Stop contenant 0.1 M NaAc pH 5.0, 0.5% SDS, 2 mM EDTA et 100 $\mu\text{g/ml}$ ARNt de levure. Les chaînes d'ARN sont ensuite précipitées à l'aide d'éthanol et analysées sur un gel dénaturant de 18% de polyacrylamide (acrylamide:bisacrylamide = 40:2) contenant 9% d'urée et migré à 400V durant 3 heures 30 min..

Chapitre 3 : Résultats

Parmi les résultats présentés dans ce chapitre, j'ai entièrement fait les figures 24B, 26, 27, 28 et 29. Dans la figure 25, les sections Rpb4 et Rpb7 ont également été faites par moi tandis que le reste est une collaboration avec Célia Jérónimo et Marie-France Langelier. La figure 24A a été faite par Célia Jérónimo, et le tableau III est une collaboration avec Marie-France Langelier. Certains de ces résultats ont été publiés dans :

Jeronimo C, Langelier MF, Zeghouf M, Cojocar M, Bergeron D, Baali D, Forget D, Mnaimneh S, Davierwala AP, Pootoolal J, Chandy M, Canadien V, Beattie BK, Richards DP, Workman JL, Hughes TR, Greenblatt J, Coulombe B. 2004. RPAP1, a novel human RNA polymerase II-associated protein affinity purified with recombinant wild-type and mutated polymerase subunits. *Mol Cell Biol.* 24(16): 7043-58.

Langelier MF, Baali D, Trinh V, Greenblatt J, Archambault J, Coulombe B. 2005. The highly conserved glutamic acid 791 of Rpb2 is involved in the binding of NTP and Mg(B) in the active center of human RNA polymerase II. *Nucleic Acids Res.* 33(8): 2629-39.

Dans la première publication, j'ai cloné et purifié les sous-unités Rpb4 et Rpb7, ainsi que le mutant de Rpb2, sw3-1078, par la méthode TAP. Dans le second, j'ai cloné et purifié les trois mutants de Rpb2 utilisés lors de l'étude, notamment MB E791A-D792A, MB E791A et MB D792A.

3.1 Purification par affinité d'un complexe contenant l'ARN pol II humaine fusionnée à l'étiquette TAP

Notre laboratoire a adapté la procédure TAP (Puig 2001, Rigaut 1999), développée originellement pour isoler des complexes protéiques à partir de cellules de levure, pour la purification de l'ARN pol II humaine et des FGTs. Durant la procédure TAP, les conditions natives sont maintenues grâce à l'utilisation de tampons à basse concentration de sel (0.1 M Na⁺) et à pH neutre. Tel que démontré dans la figure 21A, les cDNA codants pour les polypeptides d'intérêts portent la double étiquette TAP à leur extrémité C-terminale et sont clonés dans un vecteur d'expression mammifère inductible à

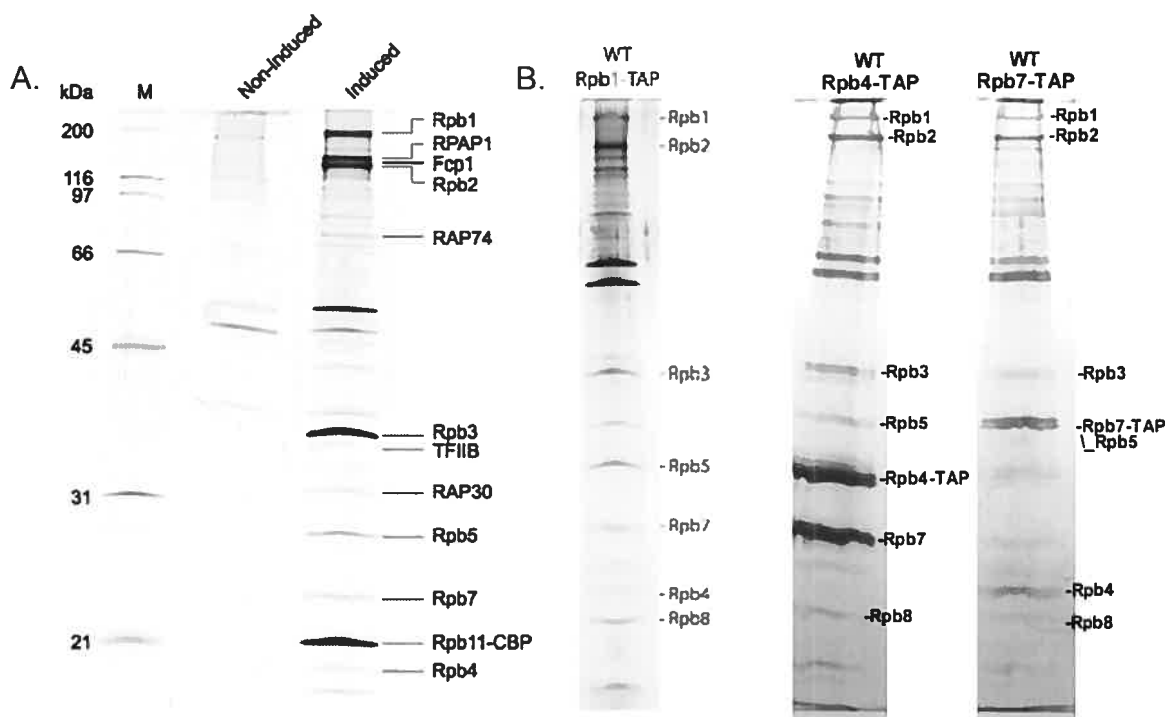


Figure 24 : Purification du complexe de l'ARN pol II fusionné à TAP

A. Les lysats cellulaires préparés à partir de cellules EcR-293, induites et non-induites, programmées pour exprimer la sous-unité Rpb11-TAP ont été purifiés par la méthode TAP et les éluats ont été analysés par SDS-PAGE. Les bandes de gel contenant les polypeptides les plus abondants ont été découpés, digérés à la trypsine et analysés par MALDI-TOF. Les sous-unités identifiées sont indiquées.

B. Sont présentés également les éluats préparés à partir de cellules programmées pour exprimer Rpb1-TAP, Rpb4-TAP et Rpb7-TAP et purifiés par la méthode TAP. Seules les sous-unités de l'ARN pol II sont indiquées.

l'ecdysone (No 1996). Ces polypeptides sont exprimés dans des cellules EcR-293 à des niveaux physiologiques après induction avec un analogue de l'ecdysone, la ponastérone A. La concentration de ponastérone A requise pour obtenir les niveaux d'expression physiologiques des polypeptides d'intérêt a été déterminée lors d'une expérience qui comparait les niveaux des polypeptides fusionnés avec les polypeptides endogènes par immuno-buvardage (voir figure 21B). Les lysats cellulaires ont été préparés à partir des lignées cellulaires induites à la ponastérone A et passés à travers deux colonnes de chromatographie d'affinité successives constituées de billes de IgG-Sépharose et de

calmoduline (Rigaut 1999). Les éluats ont été congelés et analysés par différentes méthodes.

Plusieurs sous-unités de l'ARN pol II, notamment Rpb1, Rpb2, Rpb4, Rpb7 et Rpb11, ont été exprimés individuellement dans des cellules EcR-293 et les lysats des cellules induites à la ponastérone A ont été utilisés pour la purification par double affinité. Les éluats ont été migrés à travers un gel SDS-PAGE (pour « polyacrylamide gel electrophoresis » dénaturant) et les bandes colorées au nitrate d'argent ont été identifiées par spectrométrie de masse. Comme contrôle négatif, les lysats de cellules transfectées mais non induites à la ponastérone A ont également été purifiés par double affinité et migrés et révélés sur le gel SDS-PAGE. Plusieurs bandes majeures sont détectées sur le gel chez les éluats des cellules induites et des contrôles négatifs. Cependant, un total de 17 polypeptides, absents dans les cellules contrôles et formant la majorité des composantes des éluats ont été identifiés (figure 24). Les autres bandes majeures qui se trouvent dans les deux éluats ont également été identifiées et correspondent à des protéines qui lient les colonnes d'affinité de manière non-spécifique. Dans les éluats des cellules induites, le complexe contient les 12 sous-unités de l'ARN pol II, TFIIB, les deux sous-unités de TFIIF (RAP30 et RAP74), Fcp1 (la phosphatase du CTD associée à TFIIF) ainsi qu'un polypeptide humain nouvellement identifié, RPAP1 (Jeronimo 2004). Pour valider l'existence d'un complexe unique à 17 sous-unités, TFIIB et RAP30 ont été fusionné à l'étiquette TAP, exprimés dans des cellules EcR-293 transfectées de manière stable et purifiés par la méthode TAP. Les éluats ont été analysés par gel SDS-PAGE révélé au nitrate d'argent, immuno-buvarge (Western) et/ou spectrométrie de masse (figure 25). Ces résultats indiquent que les 17 polypeptides sont des composantes du complexe de l'ARN pol II humaine (Jeronimo 2004). Dans les cas où l'ARN pol II a été purifiée à partir de l'étiquette TAP lié à l'extrémité C-terminale de Rpb4 ou Rpb7 (figure 24), une hausse importante de la concentration relative des deux sous-unités est observée et confirme que les deux polypeptides forment un sous-complexe qui se rattache à l'ensemble des sous-unités, mais qui existe également à l'état non-lié à l'ARN pol II dans les cellules transfectées.

Cell lines expressing TAP-tagged polypeptides						
Polypeptides	Rpb11	Rpb2	Rpb4	Rpb7	TFIIB	RAP30
Rpb1	SS/MS/WB	SS/WB	SS	SS	SS/MS	SS
RPAP1	SS/MS	SS	SS	SS	SS	SS
Fcp1	SS/MS	SS	SS	SS	SS	SS
Rpb2	SS/MS	SS	SS/MS	SS	SS	SS
RAP74	SS/MS/WB	SS	SS/WB	SS/WB	WB	SS
Rpb3	SS/MS	SS/MS	SS	SS	SS	SS
TFIIB	SS/WB	SS	SS/WB	SS/WB	SS/MS/WB	SS
RAP30	SS/MS/WB	SS/WB	SS/WB	SS/WB	SS/WB	SS/WB
Rpb5	SS/MS	SS	SS	SS	SS	SS
Rpb7	SS	SS	SS/MS	SS/MS	SS	SS
Rpb4	SS	SS	SS/MS	SS/MS	SS	SS
Rpb6,8-12	SS	SS	SS	SS	SS	SS

Figure 25 : Rpb2, Rpb4, Rpb7, Rpb11, TFIIB et RAP30 fusionnés à l'étiquette TAP ont mené à la purification d'un complexe à 17 sous-unités.

Les composantes du complexe ont été identifiées selon leur poids moléculaire dans les gels colorés au nitrate d'argent (SS), leur empreinte par analyse MALDI-TOF (MS), et leur immuno-réactivités (WB).

Image tirée de: Jeronimo 2004

3.2 Le complexe de l'ARN pol II fusionné à l'étiquette TAP peut être recruté à un promoteur *in vitro*

Pour déterminer si le complexe de l'ARN pol II fusionné à l'étiquette TAP purifié est fonctionnel, son habileté à former le PIC sur un promoteur *in vitro* a été testée dans un essai de retardement sur gel. L'essai de retardement sur gel sert à déterminer les interactions ADN-protéines et est basé sur le principe que les complexes de protéines et d'ADN migrent à travers un gel de polyacrylamide natif plus lentement que les fragments d'ADN libre ou double-brin. L'essai est réalisé en incubant d'abord les protéines avec une sonde d'ADN marquée au [α - 32 P] contenant le site de liaison du complexe protéique et les produits de la réaction sont analysés sur un gel de polyacrylamide non-dénaturant. La spécificité de liaison entre le complexe protéique et la sonde est établie par compétition avec des fragments d'ADN ou d'oligonucléotides contenant des séquences d'ADN non-relées à celle de la sonde. Dans ces expériences, l'oligonucléotide dIdC a été utilisé comme compétiteur à la sonde. Cette spécificité est validée en comparant des

complexes assemblés avec tous les facteurs FGTs avec des complexes assemblés en l'absence de TBP. L'absence de TBP prévient l'assemblage du PIC sur la boîte TATA. La figure 26A démontre que le complexe d'ARN pol II humaine peut former un PIC en présence de TBP, TFIIB, TFIIF et TFIIIE sur une sonde radio-marquée contenant le promoteur d'adénovirus majeur tardif. Tel que discuté précédemment, un assemblage fait en absence de TBP ne supporte pas la formation du PIC.

3.3 Le complexe de l'ARN pol II fusionné à l'étiquette TAP permet l'initiation correcte de la transcription *in vitro*

L'ARN pol II fusionnée à l'étiquette TAP a été testée pour sa capacité d'initier la transcription *in vitro* dans un système reconstitué avec les FGTs purifiés. L'ARN pol II humaine fusionnée à l'étiquette TAP et celle du thymus de veau ont été incubées avec une matrice d'ADN contenant le promoteur d'adénovirus majeur tardif en présence des facteurs purifiés TBP, TFIIB, TFIIF, TFIIIE et TFIIH. La matrice consiste en une séquence d'ADN linéarisée contenant le promoteur et une région transcrite de 391 nucléotides de long (figure 27A). La formation d'un transcrit de 391 nucléotides indique que la polymérase fusionnée à TAP est compétente pour une transcription sélective au promoteur *in vitro* (figure 27B).

3.4 Purification et analyse fonctionnelle de variantes de l'ARN pol II humaine générée par la mutagenèse dirigée des régions *lid* et *zipper*

Afin de comprendre les détails moléculaires de la réaction de transcription, la méthode TAP a été utilisée pour isoler des variants de la polymérase mutés à certains éléments qui, selon la structure cristallographique de l'enzyme de levure, jouent certains rôles présumés dans la réaction de transcription. Le tableau III représente l'ensemble des mutants étudiés et, pour chacun, les essais qui ont été effectués. Les boucles *lid* et *zipper* de Rpb1 sont hautement conservées chez les polymérases bactériennes, de levure et humaine. Elles se projettent à partir de la base de la *clamp* dans le *cleft* pour participer à la séparation de

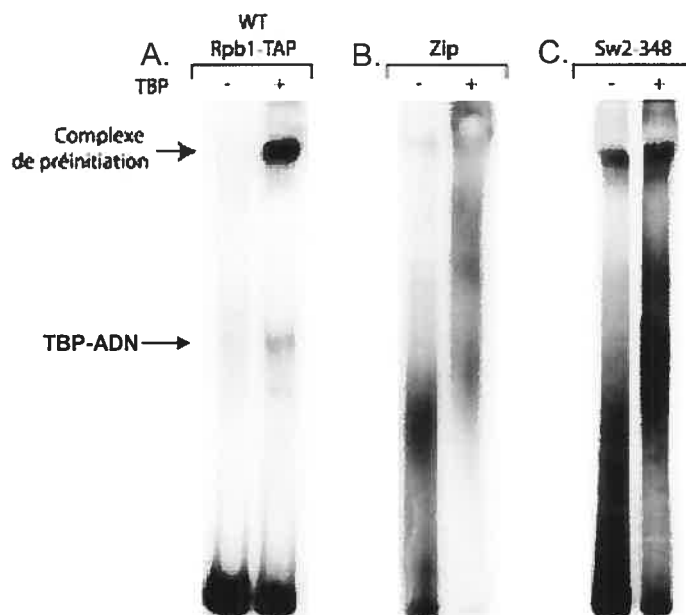


Figure 26 : Le complexe de l'ARN pol II fusionné à l'étiquette TAP peut être recruté à un promoteur *in vitro*.

L'essai de retardement sur gel (EMSA) a été effectué avec 450 ng d'ARN pol II fusionnée à TAP, sauvage (A) ou mutante (B, Zip; C, Sw2-348), en présence de TBP, TFIIB, TFIIF et TFIIE. Une réaction contrôle assemblée en absence de TBP est présentée. La sonde utilisée est formée d'un fragment d'ADN radio-marqué contenant le promoteur AdML. La position du PIC est indiquée.

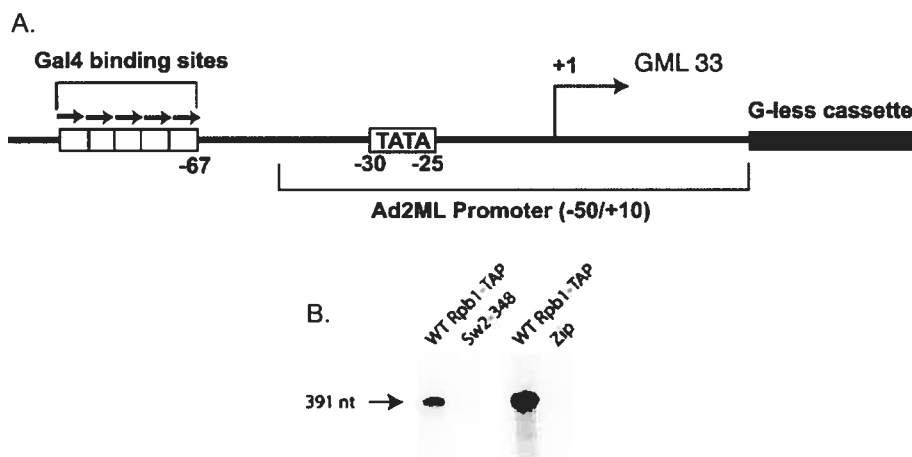


Figure 27 : Essai de transcription *in vitro*

A. Représentation schématisée de la matrice d'ADN linéaire utilisée pour les essais de transcription *in vitro*. Cinq sites de liaison de Gal4 répétés en tandem sont situés à 37 paires de bases en amont de la boîte TATA.

B. Les réactions de transcription *in vitro* contiennent 165 ng d'ARN pol II de type sauvage (WT Rpb1-TAP) ou mutantes (Sw2-348 et Zip) en présence de TBP, TFIIB, TFIIF, TFIIE et TFIIH. La position du transcrit de 391 nt est indiquée.

l'hybride ADN-ARN et maintenir le côté amont de la bulle de transcription dans le site actif (voir section 1.5.2). À l'exception de la phénylalanine 252 (chez la levure), les résidus de la boucle *lid* n'entrent pas en contact avec les acides nucléiques séparés et sont peu conservés à travers les 3 ARN polymérases nucléaires. Par conséquent, le *lid* jouerait probablement un rôle stérique dans la séparation des brins ADN-ARN (Kettenberger 2004b). De plus, le *lid* et la *saddle* entre le *wall* et la *clamp* (Cramer 2001) forment un tunnel étroit de sortie de l'ARN qui permet seulement la sortie d'un simple brin (figure 18). Chez la levure comme chez l'humain, la boucle correspondant au *lid* s'étend sur 13 acides aminés et une délétion de 7 résidus a été effectuée pour ne garder que les 3 premiers et les 3 derniers acides aminés et donner au mutant une boucle plus courte (mutant Lid6 – tableau III). Le raccourcissement de la boucle *lid* détruit complètement la capacité de Rpb1 à s'assembler avec les autres sous-unités de l'ARN pol II (figure 28), probablement parce que la délétion déstabilise Rpb1, la rendant incapable de se replier correctement et d'interagir avec ces partenaires.

La boucle correspondant au *zipper* s'étend sur 11 acides aminés et un mutant portant la délétion complète de la boucle a été créé (mutant Zip) (tableau III). Le mutant a été utilisé pour purifier avec succès tous les éléments du complexe de l'ARN pol II. Le gel révélé au nitrate d'argent dans la figure 28 montre que l'abondance relative de quelques protéines du complexe varient entre le mutant et la polymérase de type sauvage, indiquant que la délétion n'a affecté que légèrement l'assemblage du complexe, surtout au niveau des sous-unités Rpb4 et Rpb7. Pour élucider la fonction du *zipper* dans la transcription, nous avons comparé le mutant Zip avec l'ARN pol II sauvage dans différents essais. Nous avons tenté d'établir la capacité du mutant Zip à s'assembler dans le PIC sur un ADN contenant le promoteur d'adénovirus. L'essai de retardement sur gel démontre que le complexe mutant Zip est incapable de s'assembler sur le promoteur d'ADN et de former le PIC (figure 26B). La figure 27B montre que le mutant Zip est inactif dans les essais de transcription *in vitro* avec le promoteur d'AdML et en présence de tous les FGTs. Pour démarquer avec plus de précision le défaut moléculaire qui affecte l'activité transcriptionnelle *in vitro* du mutant Zip, son activité dans un essai d'élongation, indépendant du promoteur a également été testée.

Tableau III :

Noms des mutants et sommaire des expériences réalisées sur les complexes d'ARN pol II humaine portant l'étiquette TAP sur les sous-unités Rpb1 ou Rpb2 mutées

	Domaine structurel	Mutant	Assemblage du complexe à 12 sous-unités	Formation du PIC	Transcription promoteur-dépendente	Essai d'élongation
Rpb1	Lid	lid Δ 265-271	-	NA	NA	NA
	Rudder	rud Δ 321-334	+	-	ND	-
	Switch 1	sw1 S1431A F1432A E1433A	+	ND	ND	ND
	Switch 2	sw2 K346A E347A G348A R349A V350A R351A	+	+	ND	ND
	Zipper	zip Δ 45-55	+	-	-	-
Rpb2	Flap loop	flap Δ 876-887	+	+	ND	ND
		flap Δ 880-881	+	+	ND	ND
	Fork loop1	fork 1 Δ 453-465	-	NA	NA	NA
		fork 1 Δ 455-464	-	NA	NA	NA
		fork 1 Δ 458-459	+	+	-	-
		fork 1 W453A G454A D455A	-	NA	NA	NA
		fork 1 R463A G465A	-	NA	NA	NA
	Fork loop2	fork 2 Δ 487-499	-	NA	NA	NA
		fork 2 Δ 493-494	-	NA	NA	NA
		fork 2 G490A R491A D492A	-	NA	NA	NA
		fork 2 K497A P498A R499A	-	NA	NA	NA
	Métal B	MB E791A D792A	-	-	-	NA
		MB E791A	+	+	-/+	+
		MB D792A	-	-	-	NA
	Switch 3	sw3 Δ 1064-1085	-	NA	NA	NA
		sw3 R1078A S1079A R1080A	+	-	-	-
		sw3 L1084A R1085A F1086A	-	NA	NA	NA
	Switch 4	sw4 Δ 1109-1114	-	NA	NA	NA
		sw4 Δ 1111-1112	-	NA	NA	NA
		sw4 S1111A D1112A P1113A	-	NA	NA	NA

ND: Non-Déterminé, NA: Non-Applicable

L'essai d'élongation est basé sur l'observation que l'ARN pol II purifiée possède une plus grande affinité pour les matrices de type homopolymère simple-brin (Dezelee 1974). Lorsqu'une extension simple-brin est attachée en 3' à un duplexe d'ADN, l'ARN pol II s'y attache préférentiellement et initie la synthèse de chaînes d'ARN (figure 23B); il est donc possible de réaliser un essai quantitatif pour étudier la transcription en évitant les étapes limitantes de la reconnaissance du promoteur et de l'ouverture de la bulle de transcription. En d'autres termes, cet essai permet de mesurer la capacité catalytique de transcription des ARN pol II purifiées indépendamment des FGTs (Kadesch 1982). La matrice utilisée dans la figure 29 porte une extension de 15 dC sur le brin 3'. L'ARN pol II sauvage produit des transcrits de 60 nt de long, indiquant que la polymérase initie la transcription 10 nt en amont de la jonction double-brin. Par contre, le complexe d'ARN pol II portant la délétion Zip est inactif dans l'essai d'élongation (figure 29).

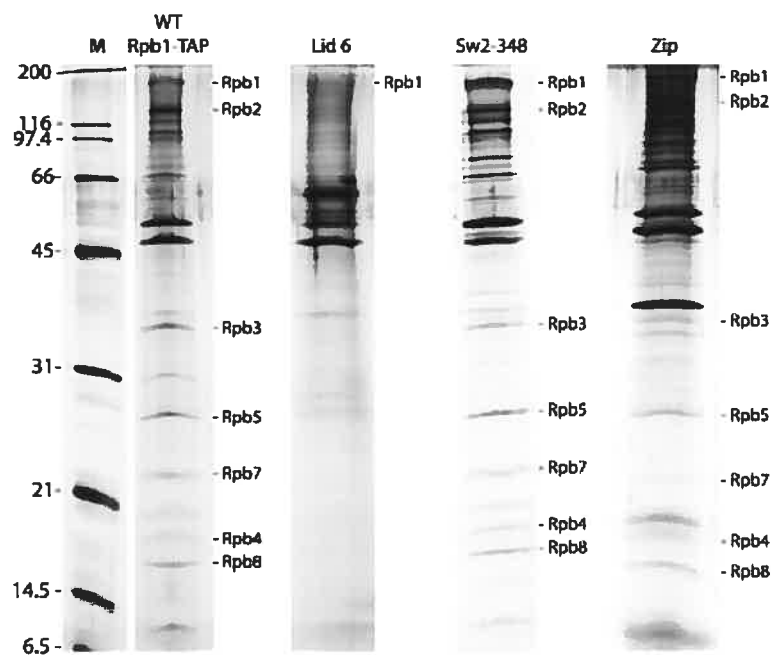


Figure 28 : Purification et identification des sous-unités du complexe d'ARN pol II humain portant l'étiquette TAP sur Rpb1 et les mutations Sw2-348 et Zip.

Le gel de polyacrylamide SDS coloré au nitrate d'argent montre d'abord la sous-unité Rpb1 sauvage portant l'étiquette TAP et certains des mutants illustrés dans le tableau 3. Le mutant Lid6 démontre un problème au niveau de l'assemblage de l'ARN pol II. Les purifications ont été effectuées au minimum 3 fois et les résultats sont reproductibles.

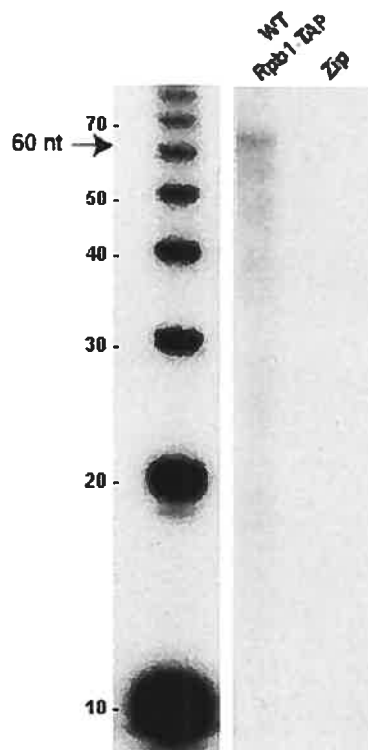


Figure 29 : Essai d'élongation

L'expérience a été effectuée sur une matrice portant un extension de 15 C simple-brin sur son extrémité 3' et en l'absence des FGTs. La taille attendue des transcrits est de 60 nt et les réactions contiennent 165 ng d'ARN pol II (sauvage ou mutante).

3.5 Purification et analyse fonctionnelle d'une variante humaine de l'ARN pol II générée par la mutagenèse dirigée des régions *switch* 1 et 2 de Rpb1

Les boucles *switch* 1 et 2 sont très peu ordonnées dans la structure de l'enzyme en solution mais deviennent ordonnées dans la structure du complexe en élongation. Cette disposition est induite par la liaison des *switches* à l'ADN en duplexe avec l'ARN et celui plus en aval de l'hybride. Ainsi, la boucle *switch* 1 entre en contact direct avec la matrice d'ADN près de la localisation de l'hybride avec l'ARN (figure 14C). Puisque la boucle est située à un endroit clé par rapport au *clamp* et au site actif, une délétion de la boucle, au complet ou en partie, affecterait probablement le repliement de Rpb1 et l'assemblage de l'ARN pol II. Parmi les 22 acides aminés de cette boucle, 3 sont conservés en terme de séquence et de position relative chez les polymérases de bactéries, levure et humains.

Un mutant dans lequel ces acides aminés, notamment S1401, F1402 et E1403, sont substitués par des résidus d'alanine (tableau III) a été créé. Chez la sous-unité Rpb1 humaine, ces acides aminés correspondent à S1431, F1432 et E1433. Le mutant a été nommé Sw1-1431. La mutation Sw1-1431 prévient probablement le repliement adéquat du peptide Rpb1 fusionné à TAP et, de ce fait, l'assemblage du complexe transcriptionnel (figure non-présentée).

La boucle *switch 2* entre également en contact avec l'hybride ADN-ARN dans le site actif (figure 14C). Pour les mêmes raisons mentionnées précédemment, une stratégie de mutagenèse dirigée a été empruntée plutôt qu'une délétion. Un mutant dans lequel les acides aminés K332, E333, G334, R335, I336 et R337 ont été remplacés par des résidus d'alanine a été créé. Chez la sous-unité Rpb1 humaine, ces acides aminés correspondent à K346, E347, G348, R349, V350 et R351. Le mutant a été nommé Sw2-348 (tableau III). La figure 28 montre que la sous-unité Rpb1 portant la mutation Sw2-348 est aussi efficace que celle du type sauvage à s'assembler en solution. Par contre, l'essai de retardement sur gel (figure 26) indique que le mutant Sw2-348 n'est pas aussi spécifique que le type sauvage pour entrer dans le PIC sur un promoteur en présence des FGTs. Le signal de retard sur gel étalé (« smear ») en présence de TBP à la figure 26C pourrait représenter des complexes transitoires instables. De plus, tel qu'on s'attendrait d'après les résultats de retard sur gel, le mutant Sw2-348 est incapable de supporter la transcription *in vitro* (figure 27B).

Chapitre 4 : Discussion

4.1 Purification simple et efficace de l'ARN pol II humaine dans des conditions natives

Les approches classiques pour la purification de l'ARN pol II humaine utilisent typiquement plusieurs étapes de chromatographie, incluant la chromatographie d'échange d'ions. L'élution de ces colonnes à l'aide de tampons à haute concentration de sels modifie la composition polypeptidique des complexes purifiés. Par exemple, les protéines régulatrices, souvent liées faiblement à l'enzyme qu'elles régulent, peuvent être perdues durant ces procédures de purification. Le groupe de Séraphin (Rigaut 1999) a développé la procédure TAP pour permettre la purification de complexes protéiques dans les conditions natives. Pour réaliser cette étude, nous avons adapté la procédure TAP pour la purification de complexes protéiques exprimés au niveau physiologique à partir de cellules humaines. Ainsi, en empêchant la surexpression des protéines fusionnées à TAP, la technique permet de limiter la formation d'interactions protéiques non-spécifiques.

Notre laboratoire a démontré l'utilisation de la méthode TAP sur des cellules humaines pour la purification simple et efficace du complexe transcriptionnel contenant les 12 sous-unités de l'ARN pol II ainsi que les deux sous-unités de TFIIF, TFIIB, la phosphatase Fcp1 et une protéine de fonction inconnue, RPAP1 (Jeronimo 2004a, figure 24A). Dans un effort pour comprendre si les 17 polypeptides font partie du même complexe, plusieurs d'entre elles ont été fusionnées à la double-étiquette TAP (figure 24B). La purification ayant donné les mêmes résultats, nous avons démontré que tous ces polypeptides font partie d'un unique complexe. De plus, trois raisons nous permettent de croire que ce complexe d'ARN pol II fusionné à TAP est la forme la plus abondante, soluble et stable dans les cellules humaines. Premièrement, les polypeptides fusionnés à TAP sont exprimés au niveau physiologique dans les cellules transfectées de manière stable. Deuxièmement, l'extraction des protéines est réalisée dans des conditions douces, à basse concentration de sels; et finalement, les étapes de purification par affinité sont

accomplies dans des conditions natives. Si le complexe médiateur se retrouve associé (Rachez 2001) à l'ARN pol II dans les extraits de cellules humaines (voir section 1.2.3.2), il n'apparaît pas sur les gels colorés au nitrate d'argent possiblement à cause de deux raisons : soit que sa concentration est beaucoup moins élevée que celle de l'ARN pol II, soit que son association avec la polymérase est trop faible pour survivre à la purification par double affinité.

La présence de l'étiquette TAP à l'extrémité C-terminale des différentes sous-unités de l'ARN pol II n'affecte pas leur activité. Les complexes d'ARN pol II sauvages fusionnés à TAP sont parfaitement actifs *in vitro*, puisqu'ils sont capables de former un PIC avec les FGTs sur un promoteur d'adénovirus (figure 26), initier adéquatement la transcription (figure 27B) et allonger des transcrits (figure 29).

4.2 Analyse structure-fonction du centre catalytique de l'ARN pol II

La structure cristallographique de l'ARN pol II eucaryote a été utilisée pour prédire la fonction de plusieurs éléments structuraux dans le site actif. Par exemple, plusieurs boucles et hélices qui lient directement ou se rapprochent du duplexe ARN-ADN au site actif de l'enzyme en élongation ont été proposées pour réguler la formation et maintenir la bulle de transcription et l'hybride ARN-ADN. Le tableau III énumère tous les motifs et les boucles des sous-unités Rpb1 et Rpb2 qui ont été ciblés par notre analyse. Une série de 25 mutants a été créée par délétions partielles ou substitutions par des alanines. Les substitutions par alanines permettent d'éliminer les chaînes latérales au-delà du carbone β , tout en gardant une charge neutre. De plus, elles sont plus adéquates que les substitutions par des glycines qui pourraient avoir un effet déstabilisant sur la structure secondaire de la protéine. Nos résultats indiquent que les mutations au site actif affectent l'assemblage et/ou la stabilité du complexe d'ARN pol II à 12 sous-unités puisque 16 parmi ces mutants présentent des défauts d'assemblage. Étant donné que les mutations affectant l'assemblage de l'ARN pol II ne se situent pas, selon les structures cristallographiques, dans des domaines directement impliqués dans les interactions entre les différentes sous-unités (Cramer 2001, Gnatt 2001, Westover 2004a), il est probable

que ces mutations affectent plutôt le repliement de la sous-unité mutée, l'empêchant de s'associer correctement avec les sous-unités adjacentes.

La mutation au domaine *switch* 1 se trouve parmi celles qui affectent le repliement de Rpb1. Ainsi, des 22 acides aminés formant la boucle *switch* 1 à la base de la *clamp*, 3 sont conservés à travers les espèces et ont été remplacés par des alanines aux positions 1431, 1432 et 1433. Ce mutant, nommé Sw1-1431, a mené à des défauts dans l'assemblage du complexe de l'ARN pol II à 17 sous-unités. La délétion de 7 des 13 acides aminés de la boucle *lid* a également mené à certains défauts dans l'assemblage du complexe d'ARN pol II (figures non-présentées). Une étude structurale récente a démontré que 3 boucles, dont le *lid* (en plus du *rudder* et de la boucle *fork1*), qui étaient désordonnées et ne pouvait être résolues dans toutes les structures cristallographiques antérieures, devenaient ordonnées dans la structure de la polymérase en transcription à l'état post-transloqué (Westover 2004a). Les 3 boucles joueraient un rôle clé dans la séparation de l'hybride ADN-ARN (figure 30). La mise en ordre de ces boucles prouve leur interaction entre elles et avec l'ARN et l'ADN. Le *lid* servirait de cale (« wedge ») pour pousser l'ARN et l'ADN à se séparer et interagirait avec les résidus -8, -9 et -10 de l'ARN. Cette boucle servirait de barrière pour maintenir la séparation des brins et guider l'ARN à travers la voie de sortie. Le résidu de Rpb1 à la pointe de la cale, Phe 252 chez la levure, rompt le lien ARN-ADN à la position -10 tout en contactant la base d'ADN, son extrémité aromatique étant perpendiculaire à la base d'ADN (Westover 2004). Chez l'humain, ce résidu correspond à Met 265, un autre acide aminé de la classe hydrophobe. Le résidu Phe 264 (correspondant à His 277 chez l'humain) pourrait contacter l'ADN de manière semblable à la position -10 ou -11.

En plus d'interagir avec l'ARN et l'ADN, le *lid* interagit avec le *rudder* qui lui-même interagit avec la boucle *fork1* (Westover 2004a). Il a été proposé qu'ensemble, ces boucles contribuent au maintien de l'amont de la bulle de transcription. De plus, elles se trouvent à proximité de la *saddle* dans la polymérase (voir figures 18 et 30B), entre la *clamp* et le *wall*, où le *lid* interagit avec d'autres éléments pour former un arc (ou « arch ») au dessus de la *saddle* (figure 30B). L'ARN quitte la région du site catalytique

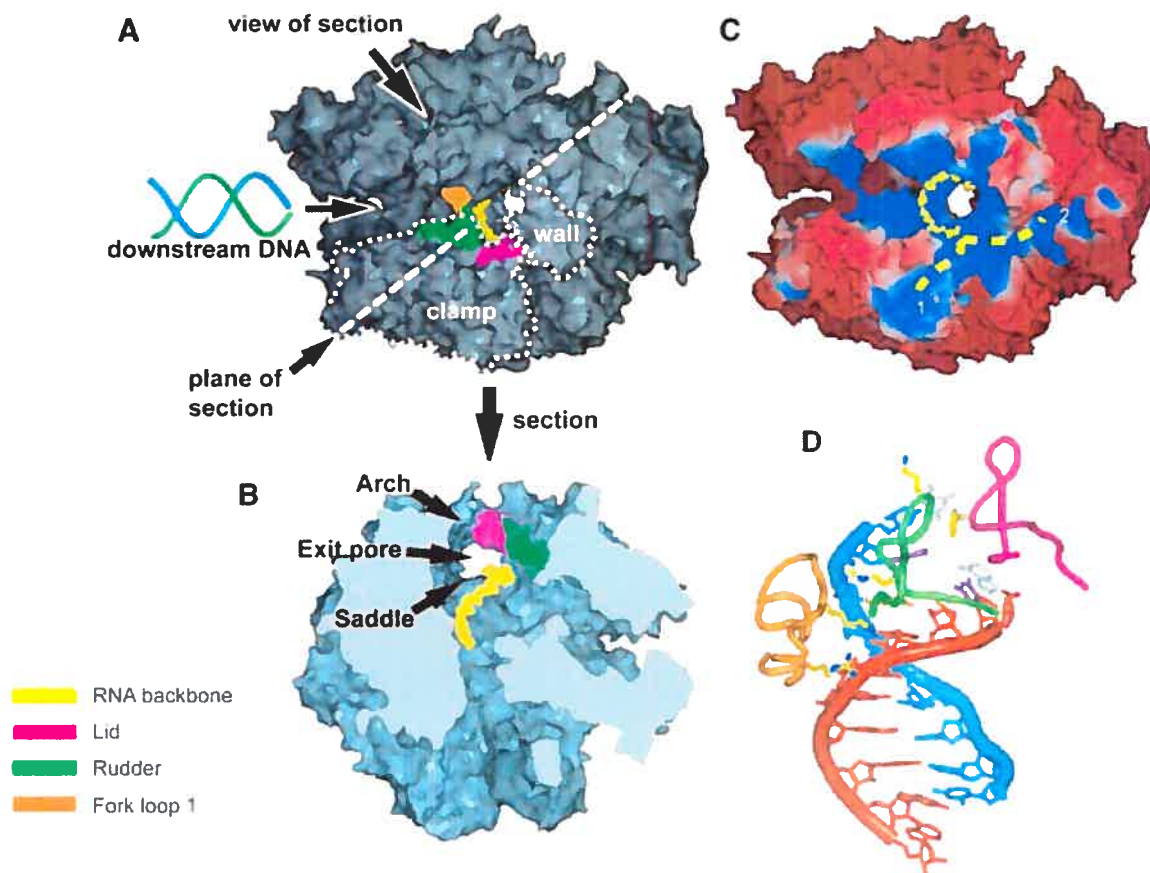


Figure 30 : Séparation de l'hybride ARN-ADN et voie de sortie de l'ARN

A. Le complexe en élongation est illustré avec la localisation de la pince et du mur en pointillé.

B. Vue sectionnée du complexe en élongation pour laisser voir l'arc au-dessus de la selle et le pore de sortie à travers lequel passe l'ARN après séparation de la matrice d'ADN.

C. Deux voies de sortie de l'ARN possibles. Le structure est colorée selon le potentiel électrostatique de la surface (négatif en rouge, neutre en blanc et positif en bleu (les domaines mur, rudder et lid ont été omis pour permettre de mieux visualiser les trajectoires 1 et 2 de l'ARN)).

D. Quelques résidus des boucles Lid, rudder et embranchement 1 interagissent avec l'ARN et l'ADN.

Image tirée de: Westover 2004

à travers un pore de sortie en dessous de l'*arch*, tandis que l'ADN quitte au-dessus de l'*arch*, empêchant la possibilité de réassociation. Au-delà de la *saddle*, deux voies de

sortie de l'ARN avaient été identifiées selon les distributions des charges à la surface de l'ARN pol II (figure 30C) (Westover 2004). La voie 1, un sillon chargé positivement traversant la base de la pince, mène à la sous-unité Rpb7 qui contient un repliement retrouvé chez les protéines liant les chaînes d'ARN simple-brin (Bushnell 2003). La voie 2, un sillon chargé positivement traversant le côté arrière de la polymérase, mène à Rpb8, qui contient également des motifs pour l'interaction possible avec des acides nucléiques simple-brin (Krapp 1998). Il a été proposé que l'ARN devait effectuer un tournant aigu pour atteindre la voie 1 et que seule la voie 2 est chargée positivement chez la bactérie (Westover 2004). L'ARN suivant la voie 2 passe d'abord en dessous du domaine *lid*. Or, le réseau d'interactions complexes dont fait partie le domaine *lid* explique sa nécessité lors du repliement de Rpb1 et de son assemblage dans le complexe de la polymérase et justifie le défaut d'assemblage du complexe lors d'un raccourcissement de la boucle.

Deux autres mutants semblaient ne pas affecter le repliement de Rpb1 et la formation du complexe en solution. En effet, le remplacement de 6 acides aminés, notamment K346, E347, G348, R349, V350 et R351, de la boucle *switch 2* (mutant Sw2-348) par des alanines n'a pas affecté l'assemblage du complexe en solution (figure 28), ce qui a permis d'analyser la fonction de cette boucle dans la réaction de transcription. Le mutant Sw2-348 semble pouvoir former le PIC avec tous les FGTs avec une efficacité moindre au complexe contenant la polymérase de type sauvage. Par contre, le mutant forme un complexe aussi spécifique dans sa liaison au promoteur et reproductible dans le cas où TBP est omis de la réaction. Or, tous les résultats d'assemblage du PIC démontrent que TBP est le premier facteur à s'assembler spécifiquement sur le promoteur de la sonde étudiée, que sa présence est requise pour la formation d'un complexe spécifique et qu'en son absence, les FGTs et la polymérase ne s'assemblent pas au promoteur.

Les 5 *switches* situées au bas de la pince mobile ont d'abord été impliquées dans le jumelage de la fermeture de la pince à la présence de l'hybride ADN-ARN (Cramer 2001, Gnatt 2001). Cette hypothèse a été basée sur le fait que ces boucles sont désordonnées dans la structure cristallographique de l'ARN pol II à 10 sous-unités en

solution (sans le dimère Rpb4-Rpb7), mais se retrouvent ordonnées dans l'enzyme en élongation. Le réarrangement de ces boucles coïncide avec le mouvement de la pince de l'état ouvert dans l'enzyme libre, qui était suggéré de permettre l'entrée de l'ADN dans le sillon, à l'état fermé dans la polymérase liée à l'hybride ADN-ARN pour assurer la stabilité du complexe en élongation. Puisque les boucles entrent en contact direct avec l'ADN dans la région de l'hybride, elles pourraient être responsable de détecter la présence d'un ARN naissant et de jumeler la fermeture de la pince sur l'ADN. Par contre, les structures cristallographiques récentes du complexe d'ARN pol II à 12 sous-unités (Armache 2003, Bushnell 2003) ont révélé que la pince est fermée en présence du dimère Rpb4-Rpb7. Ces résultats ont permis aux auteurs de proposer que l'ADN est d'abord chargé sur la pince loin du site actif de manière semblable à la polymérase bactérienne (Murakami 2002) et atteint le sillon après que la bulle de transcription ait été ouverte. Dans ce cas, les boucles *switches* entreraient en contact avec l'ADN seulement après l'ouverture du promoteur et l'entrée de l'ADN dans le sillon. Les résultats de l'étude structurale par mutagenèse indiquent que *switch 2* est important pour l'assemblage spécifique du PIC sur un promoteur d'ADN. L'alanine est un acide aminé hydrophobe et le site de la mutation de Sw2-348 se trouve près du *cleft*, un domaine de la polymérase chargé positivement. Or, le remplacement de 6 acides aminés par des alanines aurait clairement affecté la spécificité et l'activité de l'ARN pol II à former le PIC ou d'entreprendre la transcription en présence des FGTs *in vitro*.

L'autre mutant qui semblait ne pas affecter le repliement de Rpb1 et l'assemblage du complexe en solution est celui portant la délétion complète de la boucle *zipper*. À première vue, cette mutation n'a pas affecté l'assemblage du complexe en solution (figure 28), mais en examinant de plus près, on réalise que les sous-unités Rpb4 et Rpb7 se retrouvent en quantités plus faibles chez le mutant Zip que chez la polymérase de type sauvage. Cette diminution est attendue puisque Rpb4 interagit avec une région à l'extrémité N-terminale de Rpb1 située à proximité du Zipper. Nous avons tout de même poursuivi l'analyse de la fonction de cette boucle dans la réaction de transcription. Le mutant Zip semble incapable de former le PIC avec tous les FGTs. Par ce fait même, il est incapable de mener une réaction de transcription reconstituée *in vitro* sur le promoteur

AdML. De plus, dans l'essai d'élongation, réalisé indépendamment du promoteur et en absence de FGTs, le mutant s'est prouvé incapable de produire des transcrits.

Il a été suggéré que la boucle *zipper* joue un rôle dans la séparation des brins codant et non-codant de l'ADN. Ces résultats supportent l'hypothèse que la délétion de la boucle *zipper* mènerait à un défaut au niveau de l'ouverture de la bulle de transcription et la séparation des brins d'ADN en amont du site actif. La polymérase serait incapable de contacter l'ADN à cet endroit clé, menant au mauvais positionnement de la matrice d'ADN dans le site actif et corrompant la capacité de l'enzyme à initier la transcription efficacement. Cette interprétation justifierait le défaut au niveau de l'élongation à partir d'une matrice étendue à son extrémité 3'. Une autre alternative serait que le mutant est incapable d'interagir avec l'un des FGTs impliqués dans la formation du PIC. En effet, il a été observé que le domaine en forme de doigt du facteur TFIIB pénètre au centre catalytique de la polymérase à travers le canal de sortie de l'ARN (Chen 2004, Bushnell 2004). Lors de l'ouverture de la bulle de transcription, une région d'ADN désappariée est retenue transitoirement par une interaction du brin non-codant avec RAP30 et une interaction du brin codant avec le domaine en doigt de zinc de TFIIB et le site actif de la polymérase. C'est à cet endroit que l'ADN du promoteur est d'abord plié, dû à la flexibilité de la région simple-brin, avant de descendre dans le *cleft*. C'est également à cet endroit que l'ADN interagit vraisemblablement pour la première fois avec le *cleft* du site actif de la polymérase et avec les structures protéiques qui définissent la bulle de transcription, notamment le *rudder* et le *zipper* à l'extrémité amont, et la boucle *fork2* à l'extrémité avale (Gnatt 2001, Cramer 2001, Bushnell 2004). La délétion du *zipper* aurait donc un impact important au niveau de l'interaction avec TFIIB, de la formation du complexe ouvert et de l'initiation (figures 26B, 27B et 28). Cette explication impliquerait par contre que le mutant possède un deuxième défaut au niveau de l'élongation expliquant son incapacité de transcrire dans l'essai d'élongation indépendant des facteurs généraux (figure 29).

Conclusion et perspectives

Les travaux réalisés dans le laboratoire du Dr Benoit Coulombe (Jeronimo 2004, Langelier 2005) ont démontré l'importance des études biochimiques dans l'analyse structure-fonction de l'ARN pol II. Malgré que les structures cristallographiques donnent un aperçu du mécanisme transcriptionnel en se basant sur le positionnement des éléments protéiques de l'ARN pol II et des facteurs qui y sont associés, les dissections biochimiques et génétiques élaborées demeurent un moyen essentiel pour construire un modèle moléculaire détaillé du processus transcriptionnel, ainsi que de comprendre les mécanismes exacts de la régulation et de la synthèse des ARNm.

Les objectifs de mon travail de maîtrise étaient de purifier, grâce à la méthode TAP, une forme de l'ARN pol II exprimée au niveau physiologique dans les cellules humaines et active dans les essais de transcription *in vitro*. Un autre objectif était de démontrer que les différentes boucles de la sous-unité Rpb1 qui se localisent de manière stratégique aux environs du centre catalytique de la polymérase jouent un rôle crucial dans la formation d'un complexe d'ARN pol II stable en solution, l'assemblage au promoteur (PIC) et la processivité de la transcription.

L'observation des bandes colorées au nitrate d'argent et les analyses de spectrométrie de masse ont permis de conclure que le complexe purifié contient toutes les sous-unités de l'ARN pol II. De plus, les essais de retard sur gel, de transcription et d'élongation *in vitro* ont démontré que l'ARN pol II fusionnée à l'étiquette TAP et purifiée par double chromatographie d'affinité est capable de s'assembler au promoteur et de supporter la transcription spécifique et processive d'un transcrit d'ARN. Ces résultats indiquent que l'ajout de l'étiquette TAP aux sous-unités étudiées n'affectent pas leur fonction. Nous proposons alors que cette technique peut être employée pour la purification simple et efficace de différentes formes mutantes de l'ARN pol II dans l'effort d'étudier les structures et acides aminés conservés entourant le site actif.

Nous avons donc procédé à la mutagenèse dirigée à certaines boucles localisées de manière stratégique autour du site catalytique de l'ARN pol II. Dans plusieurs cas, les mutations affectent probablement le repliement de la sous-unité fusionnée à l'étiquette TAP et la rendent incapable de s'assembler et de former le complexe de l'ARN pol II à 12 sous-unités. Par contre, parmi les mutants capables de s'assembler et de former correctement le complexe de polymérase à 12 sous-unités, la plupart sont bloqués aux étapes précoces du mécanisme de transcription. Dans le cas de la boucle *switch 2*, le remplacement par des acides aminés hydrophobes d'une région rapprochée du cleft chargé positivement affecterait la spécificité de l'ARN pol II à se lier à l'ADN dans le contexte du PIC. Par contre, la délétion totale de la boucle *zipper* mènerait à un défaut au niveau de l'ouverture de la bulle de transcription et de la séparation des brins d'ADN en amont du site actif. La polymérase semble incapable de contacter l'ADN à cet endroit et de le positionner adéquatement dans le site actif afin d'initier la transcription.

Les étapes futures envisagées pour la continuité de ce projet impliquent la construction d'autres mutants au niveau des boucles et hélices impliquées dans les étapes d'élongation tel que la translocation, le *backtracking* et le *proofreading*. Ceci pourrait compléter le modèle moléculaire et vérifier les hypothèses émises suite à l'observation des structures cristallographiques. De plus, il serait intéressant d'approfondir les expériences sur certains mutants déjà en cours afin de confirmer nos conclusions, par exemple, un essai de transcription *in vitro* avec un promoteur ouvert devrait corriger le défaut du mutant *zipper*. Finalement, il serait intéressant de comparer les effets des mutants du site catalytique de la polymérase à des mutants des domaines importants chez les FGTs. Ceci pourrait nous permettre d'observer la complémentarité du site actif déjà décrite dans la littérature et de confirmer la fonction de chacun des domaines.

Bibliographie

- Albert, T., J. Mautner, J. O. Funk, K. Hortnagel, A. Pullner, and D. Eick.** 1997. Nucleosomal structures of c-myc promoters with transcriptionally engaged RNA polymerase II. *Mol.Cell Biol.* **17**:4363-4371.
- Albright, S. R. and R. Tjian.** 2000. TAFs revisited: more data reveal new twists and confirm old ideas. *Gene* **242**:1-13.
- Andel, F., III, A. G. Ladurner, C. Inouye, R. Tjian, and E. Nogales.** 1999. Three-dimensional structure of the human TFIID-IIA-IIB complex. *Science* **286**:2153-2156.
- Andrau, J. C., C. J. Van Oevelen, H. A. Van Teeffelen, P. A. Weil, F. C. Holstege, and H. T. Timmers.** 2002. Mot1p is essential for TBP recruitment to selected promoters during in vivo gene activation. *EMBO J.* **21**:5173-5183.
- Archambault, J., D. B. Jansma, J. H. Kawasoe, K. T. Arndt, J. Greenblatt, and J. D. Friesen.** 1998. Stimulation of transcription by mutations affecting conserved regions of RNA polymerase II. *J.Bacteriol.* **180**:2590-2598.
- Archambault J., R.S. Chambers, M.S. Kobor, Y. Ho, M. Cartier, D. Bolotin, B. Andrews, C.M. Kane and J. Greenblatt.** 1997. *Proc Natl Acad Sci U S A.* **(26)**:14300-5.
- Armache, K. J., H. Kettenberger, and P. Cramer.** 2003. Architecture of initiation-competent 12-subunit RNA polymerase II. *Proc.Natl.Acad.Sci.U.S.A* **100**:6964-6968.
- Asturias, F. J.** 2004. RNA polymerase II structure, and organization of the preinitiation complex. *Curr.Opin.Struct.Biol.* **14**:121-129.
- Asturias, F. J. and J. L. Craighead.** 2003. RNA polymerase II at initiation. *Proc.Natl.Acad.Sci.U.S.A* **100**:6893-6895.
- Asturias, F. J., Y. W. Jiang, L. C. Myers, C. M. Gustafsson, and R. D. Kornberg.** 1999. Conserved structures of mediator and RNA polymerase II holoenzyme. *Science* **283**:985-987.
- Auble, D. T., K. E. Hansen, C. G. Mueller, W. S. Lane, J. Thorner, and S. Hahn.** 1994. Mot1, a global repressor of RNA polymerase II transcription, inhibits TBP binding to DNA by an ATP-dependent mechanism. *Genes Dev.* **8**:1920-1934.
- Bagby, S., S. Kim, E. Maldonado, K. I. Tong, D. Reinberg, and M. Ikura.** 1995. Solution structure of the C-terminal core domain of human TFIIB: similarity to cyclin A and interaction with TATA-binding protein. *Cell* **82**:857-867.

- Balasubramanian, R., M. G. Pray-Grant, W. Selleck, P. A. Grant, and S. Tan.** 2002. Role of the Ada2 and Ada3 transcriptional coactivators in histone acetylation. *J.Biol.Chem.* **277**:7989-7995.
- Balciunas, D., M. Hallberg, S. Bjorklund, and H. Ronne.** 2003. Functional interactions within yeast mediator and evidence of differential subunit modifications. *J.Biol.Chem.* **278**:3831-3839.
- Bell, B., E. Scheer, and L. Tora.** 2001. Identification of hTAF(II)80delta Links Apoptotic Signaling Pathways to Transcription Factor TFIID Function. *Mol.Cell* **8**:591-600.
- Bengal, E., O. Flores, A. Krauskopf, D. Reinberg, and Y. Aloni.** 1991. Role of the mammalian transcription factors IIF, IIS, and IIX during elongation by RNA polymerase II. *Mol.Cell Biol.* **11**:1195-1206.
- Bhaumik, S. R. and M. R. Green.** 2001. SAGA is an essential in vivo target of the yeast acidic activator Gal4p. *Genes Dev.* **15**:1935-1945.
- Bhaumik, S. R. and M. R. Green.** 2002. Differential Requirement of SAGA Components for Recruitment of TATA-Box-Binding Protein to Promoters In Vivo. *Mol.Cell Biol.* **22**:7365-7371.
- Boeger, H., D. A. Bushnell, R. Davis, J. Griesenbeck, Y. Lorch, J. S. Strattan, K. D. Westover, and R. D. Kornberg.** 2005. Structural basis of eukaryotic gene transcription. *FEBS Lett.* **579**:899-903.
- Boeger, H., J. Griesenbeck, J. S. Strattan, and R. D. Kornberg.** 2003. Nucleosomes unfold completely at a transcriptionally active promoter. *Mol.Cell* **11**:1587-1598.
- Boeger, H., J. Griesenbeck, J. S. Strattan, and R. D. Kornberg.** 2004. Removal of promoter nucleosomes by disassembly rather than sliding in vivo. *Mol Cell* **14**:667-673.
- Booth, V., C. M. Koth, A. M. Edwards, and C. H. Arrowsmith.** 2000. Structure of a conserved domain common to the transcription factors TFIIS, elongin A, and CRSP70. *J.Biol Chem.* **275**:31266-31268.
- Botta, E., T. Nardo, A. R. Lehmann, J. M. Egly, A. M. Pedrini, and M. Stefanini.** 2002. Reduced level of the repair/transcription factor TFIH in trichothiodystrophy. *Hum.Mol.Genet.* **11**:2919-2928.
- Bourbon, H. M., A. Aguilera, A. Z. Ansari, F. J. Asturias, A. J. Berk, S. Bjorklund, T. K. Blackwell, T. Borggreffe, M. Carey, M. Carlson, J. W. Conaway, R. C. Conaway, S. W. Emmons, J. D. Fondell, L. P. Freedman, T. Fukasawa, C. M. Gustafsson, M. Han, X. He, P. K. Herman, A. G. Hinnebusch, S. Holmberg, F. C. Holstege, J. A. Jaehning, Y. J. Kim, L. Kuras, A. Leutz, J. T. Lis, M. Meisterernest, A. M. Naar, K. Nasmyth, J. D. Parvin, M. Ptashne, D.**

- Reinberg, H. Ronne, I. Sadowski, H. Sakurai, M. Sipiczki, P. W. Sternberg, D. J. Stillman, R. Strich, K. Struhl, J. Q. Svejstrup, S. Tuck, F. Winston, R. G. Roeder, and R. D. Kornberg.** 2004. A unified nomenclature for protein subunits of mediator complexes linking transcriptional regulators to RNA polymerase II. *Mol Cell* **14**:553-557.
- Brand, M., C. Leurent, V. Mallouh, L. Tora, and P. Schultz.** 1999a. Three-dimensional structures of the TAFII-containing complexes TFIID and TFTC. *Science* **286**:2151-2153.
- Brand, M., J. G. Moggs, M. Oulad-Abdelghani, F. Lejeune, F. J. Dilworth, J. Stevenin, G. Almouzni, and L. Tora.** 2001. UV-damaged DNA-binding protein in the TFTC complex links DNA damage recognition to nucleosome acetylation. *EMBO J.* **20**:3187-3196.
- Brand, M., K. Yamamoto, A. Staub, and L. Tora.** 1999b. Identification of TATA-binding protein-free TAFII-containing complex subunits suggests a role in nucleosome acetylation and signal transduction. *J.Biol.Chem.* **274**:18285-18289.
- Buratowski, S.** 1994. The basics of basal transcription by RNA polymerase II. *Cell* **77**:1-3.
- Buratowski, S., S. Hahn, L. Guarente, and P. A. Sharp.** 1989. Five intermediate complexes in transcription initiation by RNA polymerase II. *Cell* **56**:549-561.
- Burke, T. W. and J. T. Kadonaga.** 1996. *Drosophila* TFIID binds to a conserved downstream basal promoter element that is present in many TATA-box-deficient promoters. *Genes Dev.* **10**:711-724.
- Burke, T. W. and J. T. Kadonaga.** 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 R. G. Roeder.** 1996. Biochemistry and structural biology of transcription factor IID (TFIID). *Annu.Rev.Biochem.* **65**:769-799.
- Bushnell, D. A. and R. D. Kornberg.** 2003. Complete, 12-subunit RNA polymerase II at 4.1-A resolution: Implications for the initiation of transcription. *Proc.Natl.Acad.Sci.U.S.A* **100**:6969-6973.
- Bushnell, D. A., K. D. Westover, R. E. Davis, and R. D. Kornberg.** 2004. Structural basis of transcription: an RNA polymerase II-TFIIB cocystal at 4.5 Angstroms. *Science* **303**:983-988.
- Butler, J. E. and J. T. Kadonaga.** 2002. The RNA polymerase II core promoter: a key component in the regulation of gene expression. *Genes Dev.* **16**:2583-2592.

- Cairns, B. R., Y. J. Kim, M. H. Sayre, B. C. Laurent, and R. D. Kornberg.** 1994. A multisubunit complex containing the SWI1/ADR6, SWI2/SNF2, SWI3, SNF5, and SNF6 gene products isolated from yeast. *Proc.Natl.Acad.Sci.U.S.A* **91**:1950-1954.
- Cairns, B. R., Y. Lorch, Y. Li, M. Zhang, L. Lacomis, H. Erdjument-Bromage, P. Tempst, J. Du, B. Laurent, and R. D. Kornberg.** 1996. RSC, an essential, abundant chromatin-remodeling complex. *Cell* **87**:1249-1260.
- Chalkley, G. E. and C. P. Verrijzer.** 1999. DNA binding site selection by RNA polymerase II TAFs: a TAF(II)250-TAF(II)150 complex recognizes the initiator. *EMBO J.* **18**:4835-4845.
- Chambers, R. S., B. Q. Wang, Z. F. Burton, and M. E. Dahmus.** 1995. The activity of COOH-terminal domain phosphatase is regulated by a docking site on RNA polymerase II and by the general transcription factors IIF and IIB. *J.Biol.Chem.* **270**:14962-14969.
- Chang, W. H. and R. D. Kornberg.** 2000. Electron crystal structure of the transcription factor and DNA repair complex, core TFIID. *Cell* **102**:609-613.
- Chen, B. S. and M. Hampsey.** 2004. Functional interaction between TFIIB and the Rpb2 subunit of RNA polymerase II: implications for the mechanism of transcription initiation. *Mol.Cell Biol.* **24**:3983-3991.
- Chi, T. and M. Carey.** 1996. Assembly of the isomerized TFIIA--TFIID--TATA ternary complex is necessary and sufficient for gene activation. *Genes Dev.* **10**:2540-2550.
- Chi, Y., M. J. Huddleston, X. Zhang, R. A. Young, R. S. Annan, S. A. Carr, and R. J. Deshaies.** 2001. Negative regulation of Gcn4 and Msn2 transcription factors by Srb10 cyclin-dependent kinase. *Genes Dev.* **15**:1078-1092.
- Chicca, J. J., D. T. Auble, and B. F. Pugh.** 1998. Cloning and biochemical characterization of TAF-172, a human homolog of yeast Mot1. *Mol.Cell Biol.* **18**:1701-1710.
- Cho, E. J., M. S. Kobor, M. Kim, J. Greenblatt, and S. Buratowski.** 2001. Opposing effects of Ctk1 kinase and Fcp1 phosphatase at Ser 2 of the RNA polymerase II C-terminal domain. *Genes Dev.* **15**:3319-3329.
- Cho, H., T. K. Kim, H. Mancebo, W. S. Lane, O. Flores, and D. Reinberg.** 1999. A protein phosphatase functions to recycle RNA polymerase II. *Genes Dev.* **13**:1540-1552.
- Choder, M.** 2004. Rpb4 and Rpb7: subunits of RNA polymerase II and beyond. *Trends Biochem.Sci.* **29**:674-681.

- Choder, M. and R. A. Young.** 1993. A portion of RNA polymerase II molecules has a component essential for stress responses and stress survival. *Mol.Cell Biol.* **13**:6984-6991.
- Choi, C. H., M. Hiromura, and A. Usheva.** 2003. Transcription factor IIB acetylates itself to regulate transcription. *Nature* **424**:965-969.
- Chung, W. H., J. L. Craighead, W. H. Chang, C. Ezeokonkwo, A. Bareket-Samish, R. D. Kornberg, and F. J. Asturias.** 2003. RNA Polymerase II/TFIIF Structure and Conserved Organization of the Initiation Complex. *Mol.Cell* **12**:1003-1013.
- Coin, F. and J. M. Egly.** 1998. Ten years of TFIIF. *Cold Spring Harb.Symp.Quant.Biol.* **63**:105-110.
- Coleman, R. A. and B. F. Pugh.** 1997. Slow dimer dissociation of the TATA binding protein dictates the kinetics of DNA binding. *Proc.Natl.Acad.Sci.U.S.A* **94**:7221-7226.
- Coleman, R. A., A. K. Taggart, L. R. Benjamin, and B. F. Pugh.** 1995. Dimerization of the TATA binding protein. *J.Biol.Chem.* **270**:13842-13849.
- Conaway, J. W. and R. C. Conaway.** 1989. A multisubunit transcription factor essential for accurate initiation by RNA polymerase II. *J.Biol.Chem.* **264**:2357-2362.
- Conaway, J. W. and R. C. Conaway.** 1990. An RNA polymerase II transcription factor shares functional properties with Escherichia coli sigma 70. *Science* **248**:1550-1553.
- Conaway, J. W. and R. C. Conaway.** 1999. Transcription elongation and human disease. *Annu.Rev.Biochem.* **68**:301-319.
- Conaway, J. W., A. Shilatifard, A. Dvir, and R. C. Conaway.** 2000. Control of elongation by RNA polymerase II. *Trends Biochem.Sci.* **25**:375-380.
- Conaway, R. C., K. P. Garrett, J. P. Hanley, and J. W. Conaway.** 1991. Mechanism of promoter selection by RNA polymerase II: mammalian transcription factors alpha and beta gamma promote entry of polymerase into the preinitiation complex. *Proc.Natl.Acad.Sci.U.S.A* **88**:6205-6209.
- Cortes, P., O. Flores, and D. Reinberg.** 1992. Factors involved in specific transcription by mammalian RNA polymerase II: purification and analysis of transcription factor IIA and identification of transcription factor IIJ. *Mol.Cell Biol.* **12**:413-421.
- Coulombe, B.** 1999. DNA wrapping in transcription initiation by RNA polymerase II. *Biochem.Cell Biol.* **77**:257-264.

- Coulombe, B., M. Killeen, P. Liljelund, B. Honda, H. Xiao, C. J. Ingles, and J. Greenblatt.** 1992. Identification of three mammalian proteins that bind to the yeast TATA box protein TFIID. *Gene Expr.* **2**:99-110.
- Cramer, P., D. A. Bushnell, J. Fu, A. L. Gnatt, B. Maier-Davis, N. E. Thompson, R. R. Burgess, A. M. Edwards, P. R. David, and R. D. Kornberg.** 2000. Architecture of RNA polymerase II and implications for the transcription mechanism. *Science* **288**:640-649.
- Cramer, P., D. A. Bushnell, and R. D. Kornberg.** 2001. Structural basis of transcription: RNA polymerase II at 2.8 angstrom resolution. *Science* **292**:1863-1876.
- Crowley, T. E., T. Hoey, J. K. Liu, Y. N. Jan, L. Y. Jan, and R. Tjian.** 1993. A new factor related to TATA-binding protein has highly restricted expression patterns in *Drosophila*. *Nature* **361**:557-561.
- Dahmus, M. E.** 1996. Reversible phosphorylation of the C-terminal domain of RNA polymerase II. *J.Biol.Chem.* **271**:19009-19012.
- Damania, B., P. Lieberman, and J. C. Alwine.** 1998. Simian virus 40 large T antigen stabilizes the TATA-binding protein-TFIIA complex on the TATA element. *Mol.Cell Biol.* **18**:3926-3935.
- Dantanel, J. C., S. Quintin, L. Lakatos, M. Labouesse, and L. Tora.** 2000. TBP-like factor is required for embryonic RNA polymerase II transcription in *C. elegans* 84. *Mol.Cell* **6**:715-722.
- Dantanel, J. C., J. M. Wurtz, O. Poch, D. Moras, and L. Tora.** 1999. The TBP-like factor: an alternative transcription factor in metazoa? *Trends Biochem.Sci.* **24**:335-339.
- Darst, R. P., A. Dasgupta, C. Zhu, J. Y. Hsu, A. Vroom, T. Muldrow, and D. T. Auble.** 2003. Mot1 Regulates the DNA Binding Activity of Free TATA-binding Protein in an ATP-dependent Manner. *J.Biol.Chem.* **278**:13216-13226.
- Darst, R. P., D. Wang, and D. T. Auble.** 2001. MOT1-catalyzed TBP-DNA disruption: uncoupling DNA conformational change and role of upstream DNA. *EMBO J.* **20**:2028-2040.
- Davidson, I., I. Martianov, and S. Viville.** 2004. [TBP, a universal transcription factor?]. *Med.Sci.(Paris)* **20**:575-579.
- Davie, J. R.** 1998. Covalent modifications of histones: expression from chromatin templates. *Curr.Opin.Genet.Dev.* **8**:173-178.
- de Boer, J. and J. H. Hoeijmakers.** 2000. Nucleotide excision repair and human syndromes. *Carcinogenesis* **21**:453-460.

- DeJong, J., R. Bernstein, and R. G. Roeder.** 1995. Human general transcription factor TFIIA: characterization of a cDNA encoding the small subunit and requirement for basal and activated transcription. *Proc.Natl.Acad.Sci.U.S.A* **92**:3313-3317.
- DeJong, J. and R. G. Roeder.** 1993. A single cDNA, hTFIIA/alpha, encodes both the p35 and p19 subunits of human TFIIA. *Genes Dev.* **7**:2220-2234.
- Dezelee, S., A. Sentenac, and P. Fromageot.** 1974. Role of deoxyribonucleic acid-ribonucleic acid hybrids in eukaryotes. Study of the template requirements of yeast ribonucleic acid polymerases and nature of the ribonucleic acid product. *J.Biol Chem.* **249**:5971-5977.
- Dion, V. and B. Coulombe.** 2003. Interactions of a DNA-bound Transcriptional Activator with the TBP-TFIIA-TFIIB-Promoter Quaternary Complex. *J.Biol.Chem.* **278**:11495-11501.
- Doublet, S., S. Tabor, A. M. Long, C. C. Richardson, and T. Ellenberger.** 1998. Crystal structure of a bacteriophage T7 DNA replication complex at 2.2 Å resolution. *Nature* **391**:251-258.
- Douziech, M., F. Coin, J. M. Chipoulet, Y. Arai, Y. Ohkuma, J. M. Egly, and B. Coulombe.** 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., A. Sancar, and D. Reinberg.** 1994. Where transcription meets repair. *Cell* **77**:9-12.
- Dvir, A.** 2002. Promoter escape by RNA polymerase II. *Biochim.Biophys.Acta* **1577**:208-223.
- Dvir, A., J. W. Conaway, and R. C. Conaway.** 2001. Mechanism of transcription initiation and promoter escape by RNA polymerase II. *Curr.Opin.Genet.Dev.* **11**:209-214.
- Dvir, A., R. C. Conaway, and J. W. Conaway.** 1996. Promoter escape by RNA polymerase II. A role for an ATP cofactor in suppression of arrest by polymerase at promoter-proximal sites. *J.Biol.Chem.* **271**:23352-23356.
- Dvir, A., S. Tan, J. W. Conaway, and R. C. Conaway.** 1997. Promoter escape by RNA polymerase II. Formation of an escape-competent transcriptional intermediate is a prerequisite for exit of polymerase from the promoter. *J.Biol.Chem.* **272**:28175-28178.
- Dynlacht, B. D., T. Hoey, and R. Tjian.** 1991. Isolation of coactivators associated with the TATA-binding protein that mediate transcriptional activation. *Cell* **66**:563-576.

- Edwards, A. M., C. M. Kane, R. A. Young, and R. D. Kornberg.** 1991. Two dissociable subunits of yeast RNA polymerase II stimulate the initiation of transcription at a promoter in vitro. *J.Biol.Chem.* **266**:71-75.
- Egly, J. M.** 2001. The 14th Datta Lecture. TFIIH: from transcription to clinic. *FEBS Lett.* **498**:124-128.
- Elmendorf, B. J., A. Shilatifard, Q. Yan, J. W. Conaway, and R. C. Conaway.** 2001. Transcription factors TFIIF, ELL, and Elongin negatively regulate SII-induced nascent transcript cleavage by non-arrested RNA polymerase II elongation intermediates. *J.Biol.Chem.* **276**:23109-23114.
- Erie, D. A.** 2002. The many conformational states of RNA polymerase elongation complexes and their roles in the regulation of transcription. *Biochim.Biophys.Acta* **1577**:224-239.
- Evans, R., J. A. Fairley, and S. G. Roberts.** 2001. Activator-mediated disruption of sequence-specific DNA contacts by the general transcription factor TFIIB. *Genes Dev.* **15**:2945-2949.
- Fairley, J. A., R. Evans, N. A. Hawkes, and S. G. Roberts.** 2002. Core promoter-dependent TFIIB conformation and a role for TFIIB conformation in transcription start site selection. *Mol.Cell Biol.* **22**:6697-6705.
- Farago, M., T. Nahari, C. Hammel, C. N. Cole, and M. Choder.** 2003. Rpb4p, a subunit of RNA polymerase II, mediates mRNA export during stress. *Mol Biol Cell* **14**:2744-2755.
- Feaver, W. J., W. Huang, O. Gileadi, L. Myers, C. M. Gustafsson, R. D. Kornberg, and E. C. Friedberg.** 2000. Subunit interactions in yeast transcription/repair factor TFIIH. Requirement for Tfb3 subunit in nucleotide excision repair. *J.Biol.Chem.* **275**:5941-5946.
- Feaver, W. J., J. Q. Svejstrup, N. L. Henry, and R. D. Kornberg.** 1994. Relationship of CDK-activating kinase and RNA polymerase II CTD kinase TFIIH/TFIIK. *Cell* **79**:1103-1109.
- Fiedler, U. and H. T. Marc Timmers.** 2000. Peeling by binding or twisting by cranking: models for promoter opening and transcription initiation by RNA polymerase II. *Bioessays* **22**:316-326.
- Fillipova, G. N., M. K. Cheng, J. M. Moore, J. P. Truong, Y. J. Hu, D. K. Nguyen, K. D. Tsuchiya, and C. M. Distche.** 2005. Boundaries between chromosomal domains of X inactivation and escape bind CTCF and lack CpG methylation during early development. *Dev.Cell* **8**:31-42.
- Fish, R. N. and C. M. Kane.** 2002. Promoting elongation with transcript cleavage stimulatory factors. *Biochim.Biophys.Acta* **1577**:287-307.

- Flanagan, P. M., R. J. Kelleher, III, M. H. Sayre, H. Tschochner, and R. D. Kornberg.** 1991. A mediator required for activation of RNA polymerase II transcription in vitro. *Nature* **350**:436-438.
- Flores, O., I. Ha, and D. Reinberg.** 1990. Factors involved in specific transcription by mammalian RNA polymerase II. Purification and subunit composition of transcription factor IIF. *J.Biol.Chem.* **265**:5629-5634.
- Flores, O., H. Lu, M. Killeen, J. Greenblatt, Z. F. Burton, and D. Reinberg.** 1991. The small subunit of transcription factor IIF recruits RNA polymerase II into the preinitiation complex. *Proc.Natl.Acad.Sci.U.S.A* **88**:9999-10003.
- Flores, O., H. Lu, and D. Reinberg.** 1992. Factors involved in specific transcription by mammalian RNA polymerase II. Identification and characterization of factor IIIH. *J.Biol.Chem.* **267**:2786-2793.
- Fondell, J. D., H. Ge, and R. G. Roeder.** 1996. Ligand induction of a transcriptionally active thyroid hormone receptor coactivator complex. *Proc.Natl.Acad.Sci.U.S.A* **93**:8329-8333.
- Fong, N. and D. L. Bentley.** 2001. Capping, splicing, and 3' processing are independently stimulated by RNA polymerase II: different functions for different segments of the CTD. *Genes Dev.* **15**:1783-1795.
- Forget, D. and B. Coulombe.** 2003. Site-specific protein-DNA photo-cross-linking of purified complexes: Topology of the RNA polymerase II transcription initiation complex, p. 701-712. *In* S. L. Adhya and S. Garges (ed.), *RNA Polymerase and Associated Factors, Part C*, vol. 370. Elsevier Academic Press, San Diego, Calif.
- Forget, D., M. F. Langelier, C. Therien, V. Trinh, and B. Coulombe.** 2004. Photo-cross-linking of a purified preinitiation complex reveals central roles for the RNA polymerase II mobile clamp and TFIIE in initiation mechanisms. *Mol.Cell Biol.* **24**:1122-1131.
- Forget, D., F. Robert, G. Grondin, Z. F. Burton, J. Greenblatt, and B. Coulombe.** 1997. RAP74 induces promoter contacts by RNA polymerase II upstream and downstream of a DNA bend centered on the TATA box. *Proc.Natl.Acad.Sci.U.S.A* **94**:7150-7155.
- Fu, J., A. L. Gnatt, D. A. Bushnell, G. J. Jensen, N. E. Thompson, R. R. Burgess, P. R. David, and R. D. Kornberg.** 1999. Yeast RNA polymerase II at 5 Å resolution. *Cell* **98**:799-810.
- Gangloff, Y. G., J. C. Pointud, S. Thuault, L. Carre, C. Romier, S. Muratoglu, M. Brand, L. Tora, J. L. Couderc, and I. Davidson.** 2001. The TFIID components human TAF(II)140 and Drosophila BIP2 (TAF(II)155) are novel metazoan homologues of yeast TAF(II)47 containing a histone fold and a PHD finger. *Mol.Cell Biol.* **21**:5109-5121.

- Gangloff, Y. G., S. Werten, C. Romier, L. Carre, O. Poch, D. Moras, and I. Davidson.** 2000. The human TFIID components TAF(II)135 and TAF(II)20 and the yeast SAGA components ADA1 and TAF(II)68 heterodimerize to form histone-like pairs. *Mol.Cell Biol.* **20**:340-351.
- Geisberg, J. V., Z. Moqtaderi, L. Kuras, and K. Struhl.** 2002. Mot1 associates with transcriptionally active promoters and inhibits association of NC2 in *Saccharomyces cerevisiae*. *Mol.Cell Biol.* **22**:8122-8134.
- Gerard, M., L. Fischer, V. Moncollin, J. M. Chipoulet, P. Chambon, and J. M. Egly.** 1991. Purification and interaction properties of the human RNA polymerase B(II) general transcription factor BTF2. *J.Biol.Chem.* **266**:20940-20945.
- Ghazy, M. A., S. A. Brodie, M. L. Ammerman, L. M. Ziegler, and A. S. Ponticelli.** 2004. Amino acid substitutions in yeast TFIIF confer upstream shifts in transcription initiation and altered interaction with RNA polymerase II. *Mol Cell Biol* **24**:10975-10985.
- Giglia-Mari, G., F. Coin, J. A. Ranish, D. Hoogstraten, A. Theil, N. Wijgers, N. G. Jaspers, A. Raams, M. Argentini, P. J. van der Spek, E. Botta, M. Stefanini, J. M. Egly, R. Aebersold, J. H. Hoeijmakers, and W. Vermeulen.** 2004. A new, tenth subunit of TFIIH is responsible for the DNA repair syndrome trichothiodystrophy group A. *Nat.Genet.* **36**:714-719.
- Glossop, J. A., T. R. Dafforn, and S. G. Roberts.** 2004. A conformational change in TFIIB is required for activator-mediated assembly of the preinitiation complex. *Nucleic Acids Res.* **32**:1829-1835.
- Gnatt, A.** 2002. Elongation by RNA polymerase II: structure-function relationship. *Biochim.Biophys.Acta* **1577**:175.
- Gnatt, A. L., P. Cramer, J. Fu, D. A. Bushnell, and R. D. Kornberg.** 2001. Structural basis of transcription: an RNA polymerase II elongation complex at 3.3 Å resolution. *Science* **292**:1876-1882.
- Goldberg, G., T. Zarucki-Schulz, P. Caldwell, H. Weissbach, and N. Brot.** 1979. Regulation of the in vitro synthesis of *E. coli* ribosomal protein L12. *Biochem.Biophys.Res.Comm.* **91**:1453-1461.
- Grant, P. A., L. Duggan, J. Cote, S. M. Roberts, J. E. Brownell, R. Candau, R. Ohba, T. Owen-Hughes, C. D. Allis, F. Winston, S. L. Berger, and J. L. Workman.** 1997. Yeast Gcn5 functions in two multisubunit complexes to acetylate nucleosomal histones: characterization of an Ada complex and the SAGA (Spt/Ada) complex. *Genes Dev.* **11**:1640-1650.
- Grant, P. A., D. Schieltz, M. G. Pray-Grant, D. J. Steger, J. C. Reese, J. R. Yates, III, and J. L. Workman.** 1998. A subset of TAF(II)s are integral components of

the SAGA complex required for nucleosome acetylation and transcriptional stimulation. *Cell* **94**:45-53.

- Gross, C. A., C. Chan, A. Dombroski, T. Gruber, M. Sharp, J. Tupy, and B. Young.** 1998. The functional and regulatory roles of sigma factors in transcription. *Cold Spring Harb.Symp.Quant.Biol* **63**:141-155.
- Gu, W. and D. Reines.** 1995a. Identification of a decay in transcription potential that results in elongation factor dependence of RNA polymerase II. *J.Biol.Chem.* **270**:11238-11244.
- Gu, W. and D. Reines.** 1995b. Variation in the size of nascent RNA cleavage products as a function of transcript length and elongation competence. *J.Biol.Chem.* **270**:30441-30447.
- Gu, X. and W. F. Marzluff.** 1996. 3' Processing and termination of mouse histone transcripts synthesized in vitro by RNA polymerase II. *Nucleic Acids Res.* **24**:3797-3805.
- Gumbs, O. H., A. M. Campbell, and P. A. Weil.** 2003. High-affinity DNA binding by a Mot1p-TBP complex: implications for TAF-independent transcription. *EMBO J.* **22**:3131-3141.
- Guo, Y. and J. D. Gralla.** 1998. Promoter opening via a DNA fork junction binding activity. *Proc.Natl.Acad.Sci.U.S.A* **95**:11655-11660.
- Guzder, S. N., Y. Habraken, P. Sung, L. Prakash, and S. Prakash.** 1995. Reconstitution of yeast nucleotide excision repair with purified Rad proteins, replication protein A, and transcription factor TFIIH. *J.Biol.Chem.* **270**:12973-12976.
- Guzman, E. and J. T. Lis.** 1999. Transcription factor TFIIH is required for promoter melting in vivo. *Mol.Cell Biol.* **19**:5652-5658.
- Hahn, S.** 1998. Activation and the role of reinitiation in the control of transcription by RNA polymerase II. *Cold Spring Harb.Symp.Quant.Biol.* **63**:181-188.
- Hansen, S. K., S. Takada, R. H. Jacobson, J. T. Lis, and R. Tjian.** 1997. Transcription properties of a cell type-specific TATA-binding protein, TRF. *Cell* **91**:71-83.
- Hawkes, N. A., R. Evans, and S. G. Roberts.** 2000. The conformation of the transcription factor TFIIB modulates the response to transcriptional activators in vivo. *Curr.Biol.* **10**:273-276.
- Hawkes, N. A. and S. G. Roberts.** 1999. The role of human TFIIB in transcription start site selection in vitro and in vivo. *J.Biol.Chem.* **274**:14337-14343.

- Henry, N. L., A. M. Campbell, W. J. Feaver, D. Poon, P. A. Weil, and R. D. Kornberg.** 1994. TFIIF-TAF-RNA polymerase II connection. *Genes Dev.* **8**:2868-2878.
- Heo, J. H., S. J. Jeong, J. W. Seol, H. J. Kim, J. W. Han, H. W. Lee, and E. J. Cho.** 2004. Differential regulation of gene expression by RNA polymerase II in response to DNA damage. *Biochem.Biophys.Res.Comm.* **325**:892-898.
- Hernandez, N.** 1993. TBP, a universal eukaryotic transcription factor? *Genes Dev.* **7**:1291-1308.
- Hirose, Y. and J. L. Manley.** 1998. RNA polymerase II is an essential mRNA polyadenylation factor. *Nature* **395**:93-96.
- Hochheimer, A. and R. Tjian.** 2003. Diversified transcription initiation complexes expand promoter selectivity and tissue-specific gene expression. *Genes Dev.* **17**:1309-1320.
- Hoeijmakers, J. H.** 2001. From xeroderma pigmentosum to the biological clock contributions of Dirk Bootsma to human genetics. *Mutat.Res.* **485**:43-59.
- Hoffmann, A., C. M. Chiang, T. Oelgeschlager, X. Xie, S. K. Burley, Y. Nakatani, and R. G. Roeder.** 1996. A histone octamer-like structure within TFIID. *Nature* **380**:356-359.
- Hoffmann, A., T. Oelgeschlager, and R. G. Roeder.** 1997. Considerations of transcriptional control mechanisms: do TFIID-core promoter complexes recapitulate nucleosome-like functions? *Proc.Natl.Acad.Sci.U.S.A* **94**:8928-8935.
- Holmes, M. C. and R. Tjian.** 2000. Promoter-selective properties of the TBP-related factor TRF1. *Science* **288**:867-870.
- Holstege, F. C., U. Fiedler, and H. T. Timmers.** 1997. Three transitions in the RNA polymerase II transcription complex during initiation. *EMBO J.* **16**:7468-7480.
- Holstege, F. C., E. G. Jennings, J. J. Wyrick, T. I. Lee, C. J. Hengartner, M. R. Green, T. R. Golub, E. S. Lander, and R. A. Young.** 1998. Dissecting the regulatory circuitry of a eukaryotic genome. *Cell* **95**:717-728.
- Hoogstraten, D., A. L. Nigg, H. Heath, L. H. Mullenders, R. van Driel, J. H. Hoeijmakers, W. Vermeulen, and A. B. Houtsmuller.** 2002. Rapid switching of TFIID between RNA polymerase I and II transcription and DNA repair in vivo. *Mol.Cell* **10**:1163-1174.
- Horikoshi, M., K. Sekimizu, and S. Natori.** 1984. Analysis of the stimulatory factor of RNA polymerase II in the initiation and elongation complex. *J.Biol Chem.* **259**:608-611.

- Huisinga, K. L. and B. F. Pugh.** 2004. A genome-wide housekeeping role for TFIID and a highly regulated stress-related role for SAGA in *Saccharomyces cerevisiae*. *Mol.Cell* **13**:573-585.
- Imbalzano, A. N., K. S. Zaret, and R. E. Kingston.** 1994. Transcription factor (TF) IIB and TFIIA can independently increase the affinity of the TATA-binding protein for DNA. *J.Biol.Chem.* **269**:8280-8286.
- Inostroza, J. A., F. H. Mermelstein, I. Ha, W. S. Lane, and D. Reinberg.** 1992. Dr1, a TATA-binding protein-associated phosphoprotein and inhibitor of class II gene transcription. *Cell* **70**:477-489.
- Izban, M. G. and D. S. Luse.** 1992a. Factor-stimulated RNA polymerase II transcribes at physiological elongation rates on naked DNA but very poorly on chromatin templates. *J.Biol.Chem.* **267**:13647-13655.
- Izban, M. G. and D. S. Luse.** 1992b. The RNA polymerase II ternary complex cleaves the nascent transcript in a 3'----5' direction in the presence of elongation factor SII. *Genes Dev.* **6**:1342-1356.
- Izban, M. G. and D. S. Luse.** 1993. SII-facilitated transcript cleavage in RNA polymerase II complexes stalled early after initiation occurs in primarily dinucleotide increments. *J.Biol.Chem.* **268**:12864-12873.
- Jackson-Fisher, A. J., C. Chitikila, M. Mitra, and B. F. Pugh.** 1999. A role for TBP dimerization in preventing unregulated gene expression. *Mol.Cell* **3**:717-727.
- Jeon, C. and K. Agarwal.** 1996. Fidelity of RNA polymerase II transcription controlled by elongation factor TFIIS. *Proc.Natl.Acad.Sci.U.S.A* **93**:13677-13682.
- Jeronimo, C., M. F. Langelier, M. Zeghouf, M. Cojocaru, D. Bergeron, D. Baali, D. Forget, S. Mnaimneh, A. P. Davierwala, J. Pootoolal, M. Chandy, V. Canadien, B. K. Beattie, D. P. Richards, J. L. Workman, T. R. Hughes, J. Greenblatt, and B. Coulombe.** 2004a. RPAP1, a novel human RNA polymerase II-associated protein affinity purified with recombinant wild-type and mutated polymerase subunits. *Mol.Cell Biol.* **24**:7043-7058.
- Jiang, Y. W., P. Veschambre, H. Erdjument-Bromage, P. Tempst, J. W. Conaway, R. C. Conaway, and R. D. Kornberg.** 1998. Mammalian mediator of transcriptional regulation and its possible role as an end-point of signal transduction pathways. *Proc.Natl.Acad.Sci.U.S.A* **95**:8538-8543.
- Johnson, S. A., L. Dubeau, M. Kawalek, A. Dervan, A. H. Schonthal, C. V. Dang, and D. L. Johnson.** 2003. Increased expression of TATA-binding protein, the central transcription factor, can contribute to oncogenesis. *Mol.Cell Biol.* **23**:3043-3051.

- Kadesch, T. R. and M. J. Chamberlin.** 1982. Studies of in vitro transcription by calf thymus RNA polymerase II using a novel duplex DNA template. *J.Biol.Chem.* **257**:5286-5295.
- Kaltenbach, L., M. A. Horner, J. H. Rothman, and S. E. Mango.** 2000b. The TBP-like factor CeTLF is required to activate RNA polymerase II transcription during *C. elegans* embryogenesis. *Mol.Cell* **6**:705-713.
- Kaltenbach, L., M. A. Horner, J. H. Rothman, and S. E. Mango.** 2000a. The TBP-like factor CeTLF is required to activate RNA polymerase II transcription during *C. elegans* embryogenesis. *Mol.Cell* **6**:705-713.
- Kamada, K., F. Shu, H. Chen, S. Malik, G. Stelzer, R. G. Roeder, M. Meisterernst, and S. K. Burley.** 2001. Crystal structure of negative cofactor 2 recognizing the TBP-DNA transcription complex. *Cell* **106**:71-81.
- Kamenski, T., S. Heilmeier, A. Meinhart, and P. Cramer.** 2004. Structure and mechanism of RNA polymerase II CTD phosphatases. *Mol Cell* **15**:399-407.
- Kang, J. J., D. T. Auble, J. A. Ranish, and S. Hahn.** 1995. Analysis of the yeast transcription factor TFIIA: distinct functional regions and a polymerase II-specific role in basal and activated transcription. *Mol.Cell Biol.* **15**:1234-1243.
- Kaufmann, J. and S. T. Smale.** 1994. Direct recognition of initiator elements by a component of the transcription factor IID complex. *Genes Dev.* **8**:821-829.
- Kettenberger, H., K. J. Armache, and P. Cramer.** 2003. Architecture of the RNA polymerase II-TFIIS complex and implications for mRNA cleavage. *Cell* **114**:347-357.
- Khazak, V., J. Estojak, H. Cho, J. Majors, G. Sonoda, J. R. Testa, and E. A. Golemis.** 1998. Analysis of the interaction of the novel RNA polymerase II (pol II) subunit hsRBP4 with its partner hsRBP7 and with pol II. *Mol.Cell Biol.* **18**:1935-1945.
- Killeen, M., B. Coulombe, and J. Greenblatt.** 1992. Recombinant TBP, transcription factor IIB, and RAP30 are sufficient for promoter recognition by mammalian RNA polymerase II. *J.Biol.Chem.* **267**:9463-9466.
- Kim, J., S. Klooster, and D. J. Shapiro.** 1995. Intrinsically bent DNA in a eukaryotic transcription factor recognition sequence potentiates transcription activation. *J.Biol.Chem.* **270**:1282-1288.
- Kim, J. L., D. B. Nikolov, and S. K. Burley.** 1993. Co-crystal structure of TBP recognizing the minor groove of a TATA element. *Nature* **365**:520-527.

- Kim, T. K., S. Hashimoto, R. J. Kelleher, III, P. M. Flanagan, R. D. Kornberg, M. Horikoshi, and R. G. Roeder.** 1994a. Effects of activation-defective TBP mutations on transcription initiation in yeast. *Nature* **369**:252-255.
- Kim, T. K. and R. G. Roeder.** 1997. Critical role of the second stirrup region of the TATA-binding protein for transcriptional activation both in yeast and human. *J.Biol.Chem.* **272**:7540-7545.
- Kim, Y. J., S. Bjorklund, Y. Li, M. H. Sayre, and R. D. Kornberg.** 1994b. A multiprotein mediator of transcriptional activation and its interaction with the C-terminal repeat domain of RNA polymerase II. *Cell* **77**:599-608.
- Kimura, M., H. Sakurai, and A. Ishihama.** 2001. Intracellular contents and assembly states of all 12 subunits of the RNA polymerase II in the fission yeast *Schizosaccharomyces pombe*. *Eur.J.Biochem.* **268**:612-619.
- Kimura, M., H. Suzuki, and A. Ishihama.** 2002. Formation of a Carboxy-Terminal Domain Phosphatase (Fcp1)/TFIIF/RNA Polymerase II (pol II) Complex in *Schizosaccharomyces pombe* Involves Direct Interaction between Fcp1 and the Rpb4 Subunit of pol II. *Mol.Cell Biol.* **22**:1577-1588.
- Knight, J. D., R. Li, and M. Botchan.** 1991. The Activation Domain of the Bovine Papillomavirus-E2 Protein Mediates Association of Dna-Bound Dimers to Form Dna Loops. *Proceedings of the National Academy of Sciences of the United States of America* **88**:3204-3208.
- Kobayashi, N., T. G. Boyer, and A. J. Berk.** 1995. A class of activation domains interacts directly with TFIIA and stimulates TFIIA-TFIID-promoter complex assembly. *Mol.Cell Biol.* **15**:6465-6473.
- Kobor, M. S., J. Archambault, W. Lester, F. C. Holstege, O. Gileadi, D. B. Jansma, E. G. Jennings, F. Kouyoumdjian, A. R. Davidson, R. A. Young, and J. Greenblatt.** 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 R. A. Young.** 1994. An RNA polymerase II holoenzyme responsive to activators. *Nature* **368**:466-469.
- Kolodziej, P. A. and R. A. Young.** 1991. Mutations in the three largest subunits of yeast RNA polymerase II that affect enzyme assembly. *Mol.Cell Biol.* **11**:4669-4678.
- Komarnitsky, P., E. J. Cho, and S. Buratowski.** 2000. Different phosphorylated forms of RNA polymerase II and associated mRNA processing factors during transcription. *Genes Dev.* **14**:2452-2460.
- Kornberg, R. D.** 1999a. Eukaryotic transcriptional control. *Trends Cell Biol.* **9**:M46-M49.

- Kornberg, R. D. and Y. Lorch.** 1999b. Chromatin-modifying and -remodeling complexes. *Curr.Opin.Genet.Dev.* **9**:148-151.
- Krishnamurthy S, He X, Reyes-Reyes M, Moore C and Hampsey M.** 2004. Ssu72 Is an RNA polymerase II CTD phosphatase. *Mol Cell.* **14**(3):387-94.
- Kugel, J. F. and J. A. Goodrich.** 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.
- Kugel, J. F. and J. A. Goodrich.** 2000. A kinetic model for the early steps of RNA synthesis by human RNA polymerase II. *J.Biol.Chem.* **275**:40483-40491.
- Kugel, J. F. and J. A. Goodrich.** 2002. Translocation after synthesis of a four-nucleotide RNA commits RNA polymerase II to promoter escape. *Mol.Cell Biol.* **22**:762-773.
- Kulich, D. and K. Struhl.** 2001. TFIIS enhances transcriptional elongation through an artificial arrest site in vivo. *Mol.Cell Biol.* **21**:4162-4168.
- Kutach, A. K. and J. T. Kadonaga.** 2000. The downstream promoter element DPE appears to be as widely used as the TATA box in *Drosophila* core promoters. *Mol.Cell Biol.* **20**:4754-4764.
- Lagrange, T., A. N. Kapanidis, H. Tang, D. Reinberg, and R. H. Ebright.** 1998. New core promoter element in RNA polymerase II-dependent transcription: sequence-specific DNA binding by transcription factor IIB. *Genes Dev.* **12**:34-44.
- Lai, M. M.** 1995. The molecular biology of hepatitis delta virus. *Annu.Rev.Biochem.* **64**:259-286.
- Langelier, M. F., D. Forget, A. Rojas, Y. Porlier, Z. F. Burton, and B. Coulombe.** 2001. Structural and functional interactions of transcription factor (TF) IIA with TFIIE and TFIIIF in transcription initiation by RNA polymerase II. *J.Biol.Chem.* **276**:38652-38657.
- Langelier, M. F., V. Trinh, and B. Coulombe.** 2002. Gros plan sur l'ARN polymérase II. *Medecine Sciences* **18**:211-216.
- Larschan, E. and F. Winston.** 2001. The *S. cerevisiae* SAGA complex functions in vivo as a coactivator for transcriptional activation by Gal4. *Genes Dev.* **15**:1946-1956.
- Le Page, F., E. E. Kwoh, A. Avrutskaya, A. Gentil, S. A. Leadon, A. Sarasin, and P. K. Cooper.** 2000. Transcription-coupled repair of 8-oxoguanine: requirement for XPG, TFIIH, and CSB and implications for Cockayne syndrome. *Cell* **101**:159-171.

- Lee, C. H., M. R. Murphy, J. S. Lee, and J. H. Chung.** 1999. Targeting a SWI/SNF-related chromatin remodeling complex to the beta-globin promoter in erythroid cells. *Proc.Natl.Acad.Sci.U.S.A* **96**:12311-12315.
- Lee, T. I., N. J. Rinaldi, F. Robert, D. T. Odom, Z. Bar-Joseph, G. K. Gerber, N. M. Hannett, C. T. Harbison, C. M. Thompson, I. Simon, J. Zeitlinger, E. G. Jennings, H. L. Murray, D. B. Gordon, B. Ren, J. J. Wyrick, J. B. Tagne, T. L. Volkert, E. Fraenkel, D. K. Gifford, and R. A. Young.** 2002. Transcriptional regulatory networks in *Saccharomyces cerevisiae*. *Science* **298**:799-804.
- Lehmann, A. R.** 2001. The xeroderma pigmentosum group D (XPD) gene: one gene, two functions, three diseases. *Genes Dev.* **15**:15-23.
- Lei, L., D. Ren, and Z. F. Burton.** 1999. The RAP74 subunit of human transcription factor IIF has similar roles in initiation and elongation. *Mol.Cell Biol.* **19**:8372-8382.
- Leurent, C., S. Sanders, C. Ruhlmann, V. Mallouh, P. A. Weil, D. B. Kirschner, L. Tora, and P. Schultz.** 2002. Mapping histone fold TAFs within yeast TFIID. *EMBO J.* **21**:3424-3433.
- Li, S. and M. J. Smerdon.** 2002. Rpb4 and Rpb9 mediate subpathways of transcription-coupled DNA repair in *Saccharomyces cerevisiae*. *EMBO J.* **21**:5921-5929.
- Liao, S. M., J. Zhang, D. A. Jeffery, A. J. Koleske, C. M. Thompson, D. M. Chao, M. Viljoen, H. J. van Vuuren, and R. A. Young.** 1995. A kinase-cyclin pair in the RNA polymerase II holoenzyme. *Nature* **374**:193-196.
- Lim, C. Y., B. Santoso, T. Boulay, E. Dong, U. Ohler, and J. T. Kadonaga.** 2004. The MTE, a new core promoter element for transcription by RNA polymerase II. *Genes Dev.* **18**:1606-1617.
- Liu, D., R. Ishima, K. I. Tong, S. Bagby, T. Kokubo, D. R. Muhandiram, L. E. Kay, Y. Nakatani, and M. Ikura.** 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.
- Lomvardas, S. and D. Thanos.** 2001. Nucleosome sliding via TBP DNA binding in vivo. *Cell* **106**:685-696.
- Lorch, Y., M. Zhang, and R. D. Kornberg.** 1999. Histone octamer transfer by a chromatin-remodeling complex. *Cell* **96**:389-392.
- Lu, H., L. Zawel, L. Fisher, J. M. Egly, and D. Reinberg.** 1992. Human general transcription factor IIIH phosphorylates the C-terminal domain of RNA polymerase II. *Nature* **358**:641-645.

- Luger, K., A. W. Mader, R. K. Richmond, D. F. Sargent, and T. J. Richmond.** 1997. Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* **389**:251-260.
- Ma, D., I. Olave, A. Merino, and D. Reinberg.** 1996. Separation of the transcriptional coactivator and antirepression functions of transcription factor IIA. *Proc.Natl.Acad.Sci.U.S.A* **93**:6583-6588.
- Maillet, I., J. M. Buhler, A. Sentenac, and J. Labarre.** 1999. Rpb4p is necessary for RNA polymerase II activity at high temperature. *J.Biol.Chem.* **274**:22586-22590.
- Malik, S., W. Gu, W. Wu, J. Qin, and R. G. Roeder.** 2000a. The USA-derived transcriptional coactivator PC2 is a submodule of TRAP/SMCC and acts synergistically with other PCs. *Mol.Cell* **5**:753-760.
- Malik, S. and R. G. Roeder.** 2000b. Transcriptional regulation through Mediator-like coactivators in yeast and metazoan cells. *Trends Biochem.Sci.* **25**:277-283.
- Marmorstein, R. and S. L. Berger.** 2001. Structure and function of bromodomains in chromatin-regulating complexes. *Gene* **272**:1-9.
- Martianov, I., S. Viville, and I. Davidson.** 2002. RNA polymerase II transcription in murine cells lacking the TATA binding protein. *Science* **298**:1036-1039.
- Martinez, E.** 2002. Multi-protein complexes in eukaryotic gene transcription. *Plant Mol Biol* **50**:925-947.
- Matangkasombut, O., R. Auty, and S. Buratowski.** 2004. Structure and function of the TFIID complex. *Proteins in Eukaryotic Transcription* **67**:67-92.
- Matsui, T., J. Segall, P. A. Weil, and R. G. Roeder.** 1980. Multiple factors required for accurate initiation of transcription by purified RNA polymerase II. *J.Biol.Chem.* **255**:11992-11996.
- McKune, K., K. L. Richards, A. M. Edwards, R. A. Young, and N. A. Woychik.** 1993. RPB7, one of two dissociable subunits of yeast RNA polymerase II, is essential for cell viability
136. *Yeast* **9**:295-299.
- Meisterernst, M. and R. G. Roeder.** 1991. Family of proteins that interact with TFIID and regulate promoter activity. *Cell* **67**:557-567.
- Merino, A., K. R. Madden, W. S. Lane, J. J. Champoux, and D. Reinberg.** 1993. DNA topoisomerase I is involved in both repression and activation of transcription. *Nature* **365**:227-232.
- Mitsuzawa, H. and A. Ishihama.** 2003. RNA polymerase II transcription apparatus in *Schizosaccharomyces pombe*. *Curr.Genet.*

- Mittler, G., E. Kremmer, H. T. Timmers, and M. Meisterernst.** 2001. Novel critical role of a human Mediator complex for basal RNA polymerase II transcription. *EMBO Rep.* **2**:808-813.
- Miyao, T., J. D. Barnett, and N. A. Woychik.** 2001. Deletion of the RNA Polymerase Subunit RPB4 Acts as a Global, Not Stress-specific, Shut-off Switch for RNA Polymerase II Transcription at High Temperatures. *J.Biol.Chem.* **276**:46408-46413.
- Moore, P. A., J. Ozer, M. Salunek, G. Jan, D. Zerby, S. Campbell, and P. M. Lieberman.** 1999. A human TATA binding protein-related protein with altered DNA binding specificity inhibits transcription from multiple promoters and activators. *Mol.Cell Biol.* **19**:7610-7620.
- Moreland, R. J., F. Tirode, Q. Yan, J. W. Conaway, J. M. Egly, and R. C. Conaway.** 1999. A role for the TFIIH XPB DNA helicase in promoter escape by RNA polymerase II. *J.Biol.Chem.* **274**:22127-22130.
- Morin, P. E., D. E. Awrey, A. M. Edwards, and C. H. Arrowsmith.** 1996. Elongation factor TFIIS contains three structural domains: solution structure of domain II. *Proc.Natl.Acad.Sci.U.S.A* **93**:10604-10608.
- Muller, H., A. P. Bracken, R. Vernell, M. C. Moroni, F. Christians, E. Grassilli, E. Prosperini, E. Vigo, J. D. Oliner, and K. Helin.** 2001. E2Fs regulate the expression of genes involved in differentiation, development, proliferation, and apoptosis. *Genes Dev.* **15**:267-285.
- Murakami, K. S., S. Masuda, E. A. Campbell, O. Muzzin, and S. A. Darst.** 2002. Structural Basis of Transcription Initiation: An RNA Polymerase Holoenzyme-DNA Complex. *Science* **296**:1285-1290.
- Na, X., H. O. Duan, E. M. Messing, S. R. Schoen, C. K. Ryan, P. A. di Sant'Agnes, E. A. Golemis, and G. Wu.** 2003. Identification of the RNA polymerase II subunit hsRPB7 as a novel target of the von Hippel-Lindau protein. *EMBO J.* **22**:4249-4259.
- Naar, A. M., P. A. Beurang, S. Zhou, S. Abraham, W. Solomon, and R. Tjian.** 1999. Composite co-activator ACR mediates chromatin-directed transcriptional activation. *Nature* **398**:828-832.
- Nakadai, T., M. Shimada, D. Shima, H. Handa, and T. A. Tamura.** 2004. Specific interaction with transcription factor IIA and localization of the mammalian TATA-binding protein-like protein (TLP/TRF2/TLF). *J.Biol Chem.* **279**:7447-7455.
- Nedialkov, Y. A., X. Q. Gong, S. L. Hovde, Y. Yamaguchi, H. Handa, J. H. Geiger, H. Yan, and Z. F. Burton.** 2003. NTP-driven Translocation by Human RNA Polymerase II. *J.Biol.Chem.* **278**:18303-18312.

- Nikolov, D. B. and S. K. Burley.** 1997. RNA polymerase II transcription initiation: a structural view. *Proc.Natl.Acad.Sci.U.S.A* **94**:15-22.
- Nikolov, D. B., H. Chen, E. D. Halay, A. A. Usheva, K. Hisatake, D. K. Lee, R. G. Roeder, and S. K. Burley.** 1995. Crystal structure of a TFIIB-TBP-TATA-element ternary complex. *Nature* **377**:119-128.
- No, D., T. P. Yao, and R. M. Evans.** 1996. Ecdysone-inducible gene expression in mammalian cells and transgenic mice. *Proc.Natl.Acad.Sci.U.S.A* **93**:3346-3351.
- Nudler, E., A. Mustaev, E. Lukhtanov, and A. Goldfarb.** 1997. The RNA-DNA hybrid maintains the register of transcription by preventing backtracking of RNA polymerase. *Cell* **89**:33-41.
- O'Brien, T., S. Hardin, A. Greenleaf, and J. T. Lis.** 1994. Phosphorylation of RNA polymerase II C-terminal domain and transcriptional elongation. *Nature* **370**:75-77.
- Oelgeschlager, T.** 2002. Regulation of RNA polymerase II activity by CTD phosphorylation and cell cycle control. *J.Cell Physiol* **190**:160-169.
- Oelgeschlager, T., C. M. Chiang, and R. G. Roeder.** 1996. Topology and reorganization of a human TFIID-promoter complex. *Nature* **382**:735-738.
- Ogryzko, V. V., T. Kotani, X. Zhang, R. L. Schiltz, T. Howard, X. J. Yang, B. H. Howard, J. Qin, and Y. Nakatani.** 1998. Histone-like TAFs within the PCAF histone acetylase complex. *Cell* **94**:35-44.
- Ohbayashi, T., Y. Makino, and T. A. Tamura.** 1999. Identification of a mouse TBP-like protein (TLP) distantly related to the drosophila TBP-related factor. *Nucleic Acids Res.* **27**:750-755.
- Ohkuma, Y., S. Hashimoto, C. K. Wang, M. Horikoshi, and R. G. Roeder.** 1995. Analysis of the role of TFIIE in basal transcription and TFIIH-mediated carboxy-terminal domain phosphorylation through structure-function studies of TFIIE-alpha. *Mol.Cell Biol.* **15**:4856-4866.
- Ohkuma, Y. and R. G. Roeder.** 1994. Regulation of TFIIH ATPase and kinase activities by TFIIE during active initiation complex formation. *Nature* **368**:160-163.
- Ohkuma, Y., H. Sumimoto, A. Hoffmann, S. Shimasaki, M. Horikoshi, and R. G. Roeder.** 1991. Structural motifs and potential sigma homologies in the large subunit of human general transcription factor TFIIE. *Nature* **354**:398-401.
- Ohkuma, Y., H. Sumimoto, M. Horikoshi, and R. G. Roeder.** 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.
- Ohler, U., G. C. Liao, H. Niemann, and G. M. Rubin.** 2002. Computational analysis of core promoters in the *Drosophila* genome. *Genome Biol* **3**:RESEARCH0087.
- Okamoto, T., S. Yamamoto, Y. Watanabe, T. Ohta, F. Hanaoka, R. G. Roeder, and Y. Ohkuma.** 1998. Analysis of the role of TFIIE in transcriptional regulation through structure-function studies of the TFIIEbeta subunit. *J.Biol.Chem.* **273**:19866-19876.
- Okuda, M., Y. Watanabe, H. Okamura, F. Hanaoka, Y. Ohkuma, and Y. Nishimura.** 2000. Structure of the central core domain of TFIIEbeta with a novel double-stranded DNA-binding surface. *EMBO J.* **19**:1346-1356.
- Olmsted, V. K., D. E. Awrey, C. Koth, X. Shan, P. E. Morin, S. Kazanis, A. M. Edwards, and C. H. Arrowsmith.** 1998. Yeast transcript elongation factor (TFIIS), structure and function. I: NMR structural analysis of the minimal transcriptionally active region. *J.Biol Chem.* **273**:22589-22594.
- Opalka, N., M. Chlenov, P. Chacon, W. J. Rice, W. Wriggers, and S. A. Darst.** 2003. Structure and function of the transcription elongation factor GreB bound to bacterial RNA polymerase. *Cell* **114**:335-345.
- Orlicky, S. M., P. T. Tran, M. H. Sayre, and A. M. Edwards.** 2001. Dissociable Rpb4-Rpb7 subassembly of rna polymerase II binds to single-strand nucleic acid and mediates a post-recruitment step in transcription initiation. *J.Biol.Chem.* **276**:10097-10102.
- Ozer, J., K. Mitsouras, D. Zerby, M. Carey, and P. M. Lieberman.** 1998. Transcription factor IIA derepresses TATA-binding protein (TBP)-associated factor inhibition of TBP-DNA binding. *J.Biol.Chem.* **273**:14293-14300.
- Ozer, J., P. A. Moore, A. H. Bolden, A. Lee, C. A. Rosen, and P. M. Lieberman.** 1994. Molecular cloning of the small (gamma) subunit of human TFIIA reveals functions critical for activated transcription. *Genes Dev.* **8**:2324-2335.
- Pal, M. and D. S. Luse.** 2002. Strong natural pausing by RNA polymerase II within 10 bases of transcription start may result in repeated slippage and reextension of the nascent RNA. *Mol.Cell Biol.* **22**:30-40.
- Pal, M., D. McKean, and D. S. Luse.** 2001. Promoter clearance by RNA polymerase II is an extended, multistep process strongly affected by sequence. *Mol.Cell Biol.* **21**:5815-5825.
- Palangat, M. and R. Landick.** 2001. Roles of RNA:DNA hybrid stability, RNA structure, and active site conformation in pausing by human RNA polymerase II. *J.Mol.Biol.* **311**:265-282.

- Pereira, L. A., M. P. Klejman, and H. T. Timmers.** 2003. Roles for BTAF1 and Mot1p in dynamics of TATA-binding protein and regulation of RNA polymerase II transcription. *Gene* **315**:1-13.
- Persengiev, S. P., X. Zhu, B. L. Dixit, G. A. Maston, E. L. Kittler, and M. R. Green.** 2003. TRF3, a TATA-box-binding protein-related factor, is vertebrate-specific and widely expressed. *Proc.Natl.Acad.Sci.U.S.A* **100**:14887-14891.
- Peterson, C. L. and J. W. Tamkun.** 1995. The SWI-SNF complex: a chromatin remodeling machine? *Trends Biochem.Sci.* **20**:143-146.
- Peterson, M. G., J. Inostroza, M. E. Maxon, O. Flores, A. Admon, D. Reinberg, and R. Tjian.** 1991. Structure and functional properties of human general transcription factor IIE. *Nature* **354**:369-373.
- Pillai, B., V. Sampath, N. Sharma, and P. Sadhale.** 2001. Rpb4, a non-essential subunit of core RNA polymerase II of *Saccharomyces cerevisiae* is important for activated transcription of a subset of genes. *J.Biol.Chem.* **276**:30641-30647.
- Pokholok, D. K., N. M. Hannett, and R. A. Young.** 2002. Exchange of RNA polymerase II initiation and elongation factors during gene expression in vivo. *Mol Cell* **9**:799-809.
- Poon, D., A. M. Campbell, Y. Bai, and P. A. Weil.** 1994. Yeast Taf170 is encoded by MOT1 and exists in a TATA box-binding protein (TBP)-TBP-associated factor complex distinct from transcription factor IID. *J.Biol.Chem.* **269**:23135-23140.
- Pray-Grant, M. G., D. Schieltz, S. J. McMahon, J. M. Wood, E. L. Kennedy, R. G. Cook, J. L. Workman, J. R. Yates, III, and P. A. Grant.** 2002. The novel SLIK histone acetyltransferase complex functions in the yeast retrograde response pathway. *Mol.Cell Biol.* **22**:8774-8786.
- Price, D. H., A. E. Sluder, and A. L. Greenleaf.** 1989. Dynamic interaction between a *Drosophila* transcription factor and RNA polymerase II. *Mol.Cell Biol.* **9**:1465-1475.
- Pugh, B. F.** 2000. Control of gene expression through regulation of the TATA-binding protein. *Gene* **255**:1-14.
- Pugh, B. F. and R. Tjian.** 1990. Mechanism of transcriptional activation by Sp1: evidence for coactivators. *Cell* **61**:1187-1197.
- Purnell, B. A., P. A. Emanuel, and D. S. Gilmour.** 1994. TFIID sequence recognition of the initiator and sequences farther downstream in *Drosophila* class II genes. *Genes Dev.* **8**:830-842.
- Qian, X., S. N. Gozani, H. Yoon, C. J. Jeon, K. Agarwal, and M. A. Weiss.** 1993. Novel zinc finger motif in the basal transcriptional machinery: three-dimensional

NMR studies of the nucleic acid binding domain of transcriptional elongation factor TFIIS. *Biochemistry* **32**:9944-9959.

- Rabenstein, M. D., S. Zhou, J. T. Lis, and R. Tjian.** 1999. TATA box-binding protein (TBP)-related factor 2 (TRF2), a third member of the TBP family. *Proc.Natl.Acad.Sci.U.S.A* **96**:4791-4796.
- Rachez, C., B. D. Lemon, Z. Suldan, V. Bromleigh, M. Gamble, A. M. Naar, H. Erdjument-Bromage, P. Tempst, and L. P. Freedman.** 1999. Ligand-dependent transcription activation by nuclear receptors requires the DRIP complex. *Nature* **398**:824-828.
- Rachez, C., Z. Suldan, J. Ward, C. P. Chang, D. Burakov, H. Erdjument-Bromage, P. Tempst, and L. P. Freedman.** 1998. A novel protein complex that interacts with the vitamin D3 receptor in a ligand-dependent manner and enhances VDR transactivation in a cell-free system. *Genes Dev.* **12**:1787-1800.
- Rani, P. G., J. A. Ranish, and S. Hahn.** 2004. RNA polymerase II (Pol II)-TFIIF and Pol II-mediator complexes: the major stable Pol II complexes and their activity in transcription initiation and reinitiation. *Mol.Cell Biol.* **24**:1709-1720.
- Ranish, J. A. and S. Hahn.** 1996. Transcription: basal factors and activation. *Curr.Opin.Genet.Dev.* **6**:151-158.
- Ranish, J. A., N. Yudkovsky, and S. Hahn.** 1999. Intermediates in formation and activity of the RNA polymerase II preinitiation complex: holoenzyme recruitment and a postrecruitment role for the TATA box and TFIIB. *Genes Dev.* **13**:49-63.
- Reese, J. C.** 2003. Basal transcription factors. *Curr.Opin.Genet.Dev.* **13**:114-118.
- Reese, J. C., L. Apone, S. S. Walker, L. A. Griffin, and M. R. Green.** 1994. Yeast TAFIIS in a multisubunit complex required for activated transcription. *Nature* **371**:523-527.
- Reines, D.** 1992. Elongation factor-dependent transcript shortening by template-engaged RNA polymerase II. *J.Biol.Chem.* **267**:3795-3800.
- Ren, D., L. Lei, and Z. F. Burton.** 1999. A region within the RAP74 subunit of human transcription factor IIF is critical for initiation but dispensable for complex assembly. *Mol.Cell Biol.* **19**:7377-7387.
- Renner, D. B., Y. Yamaguchi, T. Wada, H. Handa, and D. H. Price.** 2001. A Highly Purified RNA Polymerase II Elongation Control System. *J.Biol.Chem.* **276**:42601-42609.
- Rigaut, G., A. Shevchenko, B. Rutz, M. Wilm, M. Mann, and B. Seraphin.** 1999. A generic protein purification method for protein complex characterization and proteome exploration. *Nat.Biotechnol.* **17**:1030-1032.

- Rivetti, C., S. Codeluppi, G. Dieci, and C. Bustamante.** 2003. Visualizing RNA extrusion and DNA wrapping in transcription elongation complexes of bacterial and eukaryotic RNA polymerases. *J.Mol.Biol.* **326**:1413-1426.
- Robert, F., M. Douziech, D. Forget, J. M. Egly, J. Greenblatt, Z. F. Burton, and B. Coulombe.** 1998. Wrapping of promoter DNA around the RNA polymerase II initiation complex induced by TFIIF. *Mol.Cell* **2**:341-351.
- Robert, F., D. Forget, J. Li, J. Greenblatt, and B. Coulombe.** 1996. Localization of subunits of transcription factors IIE and IIF immediately upstream of the transcriptional initiation site of the adenovirus major late promoter. *J.Biol.Chem.* **271**:8517-8520.
- Roberts, S. G. and M. R. Green.** 1994. Activator-induced conformational change in general transcription factor TFIIB. *Nature* **371**:717-720.
- Rudd, M. D., M. G. Izban, and D. S. Luse.** 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.
- Ryu, S. and R. Tjian.** 1999a. Purification of transcription cofactor complex CRSP. *Proc.Natl.Acad.Sci.U.S.A* **96**:7137-7142.
- Ryu, S., S. Zhou, A. G. Ladurner, and R. Tjian.** 1999b. The transcriptional cofactor complex CRSP is required for activity of the enhancer-binding protein Sp1. *Nature* **397**:446-450.
- Sakurai, H., H. Mitsuzawa, M. Kimura, and A. Ishihama.** 1999. The Rpb4 subunit of fission yeast *Schizosaccharomyces pombe* RNA polymerase II is essential for cell viability and similar in structure to the corresponding subunits of higher eukaryotes. *Mol.Cell Biol.* **19**:7511-7518.
- Samkurashvili, I. and D. S. Luse.** 1996. Translocation and transcriptional arrest during transcript elongation by RNA polymerase II. *J.Biol.Chem.* **271**:23495-23505.
- Samuels, M., A. Fire, and P. A. Sharp.** 1982. Separation and characterization of factors mediating accurate transcription by RNA polymerase II. *J.Biol.Chem.* **257**:14419-14427.
- Sawadogo, M. and R. G. Roeder.** 1985. Interaction of a gene-specific transcription factor with the adenovirus major late promoter upstream of the TATA box region. *Cell* **43**:165-175.
- Schaeffer, L., V. Moncollin, R. Roy, A. Staub, M. Mezzina, A. Sarasin, G. Weeda, J. H. Hoeijmakers, and J. M. Egly.** 1994. The ERCC2/DNA repair protein is associated with the class II BTF2/TFIIH transcription factor. *EMBO J.* **13**:2388-2392.

- Schaeffer, L., R. Roy, S. Humbert, V. Moncollin, W. Vermeulen, J. H. Hoeijmakers, P. Chambon, and J. M. Egly.** 1993. DNA repair helicase: a component of BTF2 (TFIIH) basic transcription factor. *Science* **260**:58-63.
- Schultz, P., S. Fribourg, A. Poterszman, V. Mallouh, D. Moras, and J. M. Egly.** 2000. Molecular structure of human TFIIH. *Cell* **102**:599-607.
- Sekimizu, K., N. Kobayashi, D. Mizuno, and S. Natori.** 1976. Purification of a factor from Ehrlich ascites tumor cells specifically stimulating RNA polymerase II. *Biochemistry* **15**:5064-5070.
- Selleck, W., R. Howley, Q. Fang, V. Podolny, M. G. Fried, S. Buratowski, and S. Tan.** 2001. A histone fold TAF octamer within the yeast TFIID transcriptional coactivator. *Nat.Struct.Biol.* **8**:695-700.
- Sheffer, A., M. Varon, and M. Choder.** 1999. Rpb7 can interact with RNA polymerase II and support transcription during some stresses independently of Rpb4. *Mol.Cell Biol.* **19**:2672-2680.
- Shiio, Y. and R. N. Eisenman.** 2003. Histone sumoylation is associated with transcriptional repression. *Proc.Natl.Acad.Sci.U.S.A* **100**:13225-13230.
- Shilatifard, A., R. C. Conaway, and J. W. Conaway.** 2003. The RNA polymerase II elongation complex. *Annu.Rev.Biochem.* **72**:693-715.
- Shimada, M., T. Nakadai, T. A. Tamura, T. E. Adamson, D. H. Price, T. E. Adamson, and D. H. Price.** 2003. TATA-Binding Protein-Like Protein (TLP/TRF2/TLF) Negatively Regulates Cell Cycle Progression and Is Required for the Stress-Mediated G(2) Checkpoint
Cotranscriptional Processing of *Drosophila* Histone mRNAs. *Mol.Cell Biol.* **23**:4107-4120.
- Singh, M. V., C. E. Bland, and P. A. Weil.** 2004. Molecular and genetic characterization of a Taf1p domain essential for yeast TFIID assembly. *Mol Cell Biol* **24**:4929-4942.
- Smale, S. T.** 1997. Transcription initiation from TATA-less promoters within eukaryotic protein-coding genes. *Biochim.Biophys.Acta* **1351**:73-88.
- Smale, S. T., A. Jain, J. Kaufmann, K. H. Emami, K. Lo, and I. P. Garraway.** 1998. The initiator element: a paradigm for core promoter heterogeneity within metazoan protein-coding genes. *Cold Spring Harb.Symp.Quant.Biol.* **63**:21-31.
- Smale, S. T. and J. T. Kadonaga.** 2003. The RNA polymerase II core promoter. *Annu.Rev.Biochem.* **72**:449-479.

- Sopta, M., Z. F. Burton, and J. Greenblatt.** 1989. Structure and associated DNA-helicase activity of a general transcription initiation factor that binds to RNA polymerase II. *Nature* **341**:410-414.
- Sterner, D. E., P. A. Grant, S. M. Roberts, L. J. Duggan, R. Belotserkovskaya, L. A. Pacella, F. Winston, J. L. Workman, and S. L. Berger.** 1999. Functional organization of the yeast SAGA complex: distinct components involved in structural integrity, nucleosome acetylation, and TATA-binding protein interaction. *Mol.Cell Biol.* **19**:86-98.
- Sterner, D. E., X. Wang, M. H. Bloom, G. M. Simon, and S. L. Berger.** 2002. The SANT Domain of Ada2 Is Required for Normal Acetylation of Histones by the Yeast SAGA Complex. *J.Biol.Chem.* **277**:8178-8186.
- Sumimoto, H., Y. Ohkuma, E. Sinn, H. Kato, S. Shimasaki, M. Horikoshi, and R. G. Roeder.** 1991. Conserved sequence motifs in the small subunit of human general transcription factor TFIIE. *Nature* **354**:401-404.
- Sun, X., D. Ma, M. Sheldon, K. Yeung, and D. Reinberg.** 1994. Reconstitution of human TFIIA activity from recombinant polypeptides: a role in TFIID-mediated transcription. *Genes Dev.* **8**:2336-2348.
- Sun, X., Y. Zhang, H. Cho, P. Rickert, E. Lees, W. Lane, and D. Reinberg.** 1998. NAT, a human complex containing Srb polypeptides that functions as a negative regulator of activated transcription. *Mol.Cell* **2**:213-222.
- Svejstrup, J. Q., P. Vichi, and J. M. Egly.** 1996. The multiple roles of transcription/repair factor TFIIF. *Trends Biochem.Sci.* **21**:346-350.
- Svejstrup, J. Q., Z. Wang, W. J. Feaver, X. Wu, D. A. Bushnell, T. F. Donahue, E. C. Friedberg, and R. D. Kornberg.** 1995. Different forms of TFIIF for transcription and DNA repair: holo-TFIIF and a nucleotide excision repairosome. *Cell* **80**:21-28.
- Taggart, A. K. and B. F. Pugh.** 1996. Dimerization of TFIID when not bound to DNA. *Science* **272**:1331-1333.
- Tan, Q., K. L. Linask, R. H. Ebricht, and N. A. Woychik.** 2000. Activation mutants in yeast RNA polymerase II subunit RPB3 provide evidence for a structurally conserved surface required for activation in eukaryotes and bacteria. *Genes Dev.* **14**:339-348.
- Tan, Q., M. H. Prysak, and N. A. Woychik.** 2003. Loss of the Rpb4/Rpb7 Subcomplex in a Mutant Form of the Rpb6 Subunit Shared by RNA Polymerases I, II, and III. *Mol.Cell Biol.* **23**:3329-3338.

- Tan, S., T. Aso, R. C. Conaway, and J. W. Conaway.** 1994. Roles for both the RAP30 and RAP74 subunits of transcription factor IIF in transcription initiation and elongation by RNA polymerase II. *J.Biol.Chem.* **269**:25684-25691.
- Tan, S., R. C. Conaway, and J. W. Conaway.** 1995. Dissection of transcription factor TFIIF functional domains required for initiation and elongation. *Proc.Natl.Acad.Sci.U.S.A* **92**:6042-6046.
- Tanese, N., B. F. Pugh, and R. Tjian.** 1991. Coactivators for a proline-rich activator purified from the multisubunit human TFIID complex. *Genes Dev.* **5**:2212-2224.
- Tassan, J. P., M. Jaquenoud, P. Leopold, S. J. Schultz, and E. A. Nigg.** 1995. Identification of human cyclin-dependent kinase 8, a putative protein kinase partner for cyclin C. *Proc.Natl.Acad.Sci.U.S.A* **92**:8871-8875.
- Teichmann, M., Z. Wang, E. Martinez, A. Tjernberg, D. Zhang, F. Vollmer, B. T. Chait, and R. G. Roeder.** 1999. Human TATA-binding protein-related factor-2 (hTRF2) stably associates with hTFIIA in HeLa cells. *Proc.Natl.Acad.Sci.U.S.A* **96**:13720-13725.
- Thomas, M. J., A. A. Platas, and D. K. Hawley.** 1998. Transcriptional fidelity and proofreading by RNA polymerase II. *Cell* **93**:627-637.
- Thompson, C. M., A. J. Koleske, D. M. Chao, and R. A. Young.** 1993. A multisubunit complex associated with the RNA polymerase II CTD and TATA-binding protein in yeast. *Cell* **73**:1361-1375.
- 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., D. Busso, F. Coin, and J. M. Egly.** 1999. Reconstitution of the transcription factor TFIIH: assignment of functions for the three enzymatic subunits, XPB, XPD, and cdk7. *Mol.Cell* **3**:87-95.
- Tora, L.** 2002. A unified nomenclature for TATA box binding protein (TBP)-associated factors (TAFs) involved in RNA polymerase II transcription. *Genes Dev.* **16**:673-675.
- Um, M., J. Yamauchi, S. Kato, and J. L. Manley.** 2001. Heterozygous disruption of the TATA-binding protein gene in DT40 cells causes reduced cdc25B phosphatase expression and delayed mitosis. *Mol.Cell Biol.* **21**:2435-2448.
- Uptain, S. M., C. M. Kane, and M. J. Chamberlin.** 1997. Basic mechanisms of transcript elongation and its regulation. *Annu.Rev.Biochem.* **66**:117-172.

- Urnov, F. D., E. J. Rebar, A. Reik, and P. P. Pandolfi.** 2002. Designed transcription factors as structural, functional and therapeutic probes of chromatin in vivo: Fourth in review series on chromatin dynamics. *EMBO Rep.* **3**:610-615.
- van der Knaap, J. A., J. W. Borst, P. C. van der Vliet, R. Gentz, and H. T. Timmers.** 1997. Cloning of the cDNA for the TATA-binding protein-associated factor II170 subunit of transcription factor B-TFIID reveals homology to global transcription regulators in yeast and *Drosophila*. *Proc.Natl.Acad.Sci.U.S.A* **94**:11827-11832.
- Vassilyev, D. G., S. Sekine, O. Laptenko, J. Lee, M. N. Vassilyeva, S. Borukhov, and S. Yokoyama.** 2002. Crystal structure of a bacterial RNA polymerase holoenzyme at 2.6 Å resolution. *Nature* **417**:712-719.
- Veenstra, G. J., D. L. Weeks, and A. P. Wolffe.** 2000. Distinct roles for TBP and TBP-like factor in early embryonic gene transcription in *Xenopus*. *Science* **290**:2312-2315.
- Verrijzer, C. P., K. Yokomori, J. L. Chen, and R. Tjian.** 1994. *Drosophila* TAFII150: similarity to yeast gene TSM-1 and specific binding to core promoter DNA. *Science* **264**:933-941.
- Wang, D. and D. K. Hawley.** 1993. Identification of a 3'→5' exonuclease activity associated with human RNA polymerase II. *Proc.Natl.Acad.Sci.U.S.A* **90**:843-847.
- Wang, E. H. and R. Tjian.** 1994. Promoter-selective transcriptional defect in cell cycle mutant ts13 rescued by hTAFII250. *Science* **263**:811-814.
- Wang, W., M. Carey, and J. D. Gralla.** 1992. Polymerase II promoter activation: closed complex formation and ATP-driven start site opening. *Science* **255**:450-453.
- Weber, J. A., D. J. Taxman, Q. Lu, and D. S. Gilmour.** 1997. Molecular architecture of the hsp70 promoter after deletion of the TATA box or the upstream regulation region. *Mol.Cell Biol.* **17**:3799-3808.
- Westover, K. D., D. A. Bushnell, and R. D. Kornberg.** 2004. Structural basis of transcription: separation of RNA from DNA by RNA polymerase II. *Science* **303**:1014-1016.
- Wieczorek, E., M. Brand, X. Jacq, and L. Tora.** 1998. Function of TAF(II)-containing complex without TBP in transcription by RNA polymerase II. *Nature* **393**:187-191.
- Wind, M. and D. Reines.** 2000. Transcription elongation factor SII. *Bioessays* **22**:327-336.
- Winkler, G. S., S. J. Araujo, U. Fiedler, W. Vermeulen, F. Coin, J. M. Egly, J. H. Hoeijmakers, R. D. Wood, H. T. Timmers, and G. Weeda.** 2000. TFIIF with

inactive XPD helicase functions in transcription initiation but is defective in DNA repair. *J.Biol.Chem.* **275**:4258-4266.

- Workman, J. L. and R. E. Kingston.** 1998. Alteration of nucleosome structure as a mechanism of transcriptional regulation. *Annu.Rev.Biochem.* **67**:545-579.
- Woychik, N. A. and M. Hampsey.** 2002. The RNA Polymerase II Machinery. Structure Illuminates Function. *Cell* **108**:453-463.
- Woychik, N. A. and R. A. Young.** 1989. RNA polymerase II subunit RPB4 is essential for high- and low-temperature yeast cell growth. *Mol.Cell Biol.* **9**:2854-2859.
- Wu, J., K. M. Parkhurst, R. M. Powell, and L. J. Parkhurst.** 2001a. DNA sequence-dependent differences in TATA-binding protein-induced DNA bending in solution are highly sensitive to osmolytes. *J.Biol.Chem.* **276**:14623-14627.
- Wu, P. Y. and F. Winston.** 2002. Analysis of Spt7 Function in the *Saccharomyces cerevisiae* SAGA Coactivator Complex. *Mol.Cell Biol.* **22**:5367-5379.
- Wu, X., A. Chang, M. Sudol, and S. D. Hanes.** 2001b. Genetic interactions between the ESS1 prolyl-isomerase and the RSP5 ubiquitin ligase reveal opposing effects on RNA polymerase II function. *Curr.Genet.* **40**:234-242.
- Xie, X., T. Kokubo, S. L. Cohen, U. A. Mirza, A. Hoffmann, B. T. Chait, R. G. Roeder, Y. Nakatani, and S. K. Burley.** 1996. Structural similarity between TAFs and the heterotetrameric core of the histone octamer. *Nature* **380**:316-322.
- Yan, Q., R. J. Moreland, J. W. Conaway, and R. C. Conaway.** 1999. Dual roles for transcription factor IIF in promoter escape by RNA polymerase II. *J.Biol.Chem.* **274**:35668-35675.
- Yeo M, Lin PS, Dahmus ME and Gill GN.** 2003. A novel RNA polymerase II C-terminal domain phosphatase that preferentially dephosphorylates serine 5. *J Biol Chem.* **278**(28):26078-85.
- Yu X, Chini CC, He M, Mer G and Chen J.** 2003. The BRCT domain is a phospho-protein binding domain. *Science.* **302**(5645):639-42.
- Yudkovsky, N., J. A. Ranish, and S. Hahn.** 2000. A transcription reinitiation intermediate that is stabilized by activator. *Nature* **408**:225-229.
- Zhang, C. and Z. F. Burton.** 2004. Transcription factors IIF and IIS and nucleoside triphosphate substrates as dynamic probes of the human RNA polymerase II mechanism. *J.Mol Biol* **342**:1085-1099.
- Zhang, C., H. Yan, and Z. F. Burton.** 2003. Combinatorial Control of Human RNA Polymerase II (RNAP II) Pausing and Transcript Cleavage by Transcription

Factor IIF, Hepatitis {delta} Antigen, and Stimulatory Factor II. *J.Biol.Chem.* **278**:50101-50111.

Zhang, D. and M. Rosbash. 1999a. Identification of eight proteins that cross-link to pre-mRNA in the yeast commitment complex. *Genes Dev.* **13**:581-592.

Zhang, G., E. A. Campbell, L. Minakhin, C. Richter, K. Severinov, and S. A. Darst. 1999b. Crystal structure of *Thermus aquaticus* core RNA polymerase at 3.3 Å resolution. *Cell* **98**:811-824.

Zhang, W. H., R. Srihari, R. N. Day, and F. Schaufele. 2001. CCAAT/enhancer-binding protein alpha alters histone H3 acetylation at large subnuclear domains. *J.Biol.Chem.* **276**:40373-40376.

Zheng, L., K. P. Hoeflich, L. M. Elsby, M. Ghosh, S. G. Roberts, and M. Ikura. 2004. FRET evidence for a conformational change in TFIIB upon TBP-DNA binding. *Eur.J.Biochem.* **271**:792-800.

Zhu, A. and M. A. Kuziora. 1996. Homeodomain interaction with the beta subunit of the general transcription factor TFIIE. *J.Biol.Chem.* **271**:20993-20996.

