

Université de Montréal

**Mécanismes de transcription par l'ARN polymérase II : Étude
structure-fonction du site catalytique et rôles des facteurs de
transcription TFIIA, TFIIE et TFIIIF**

par

Marie-France Langelier

Programme de Biochimie
Faculté de Médecine

Thèse présentée à la Faculté des études supérieures
en vue de l'obtention du grade
Philosophiae Doctor (Ph.D.)
en biochimie

Janvier 2005

Copyright Marie-France Langelier, 2005



W
4
U58
2005
V.082

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

Cette thèse intitulée :
Mécanismes de transcription par l'ARN polymérase II : Étude structure-fonction du site
catalytique et rôles des facteurs de transcription TFIIA, TFIIE et TFIIIF

Présentée par :
Marie-France Langelier

a été évaluée par un jury composé des personnes suivantes :

Muriel Aubry, présidente rapporteur et
représentante du doyen de la FES

Benoit Coulombe, directeur de recherche

François Robert, membre du jury

Luc Gaudreau, examinateur externe

RÉSUMÉ

L'ARN polymérase II (ARN pol II) est l'enzyme qui catalyse la synthèse de tous les ARN messagers (ARNm) cellulaires chez les eucaryotes, au cours du processus transcriptionnel. L'ARN pol II ne peut initier la transcription de manière spécifique à un gène sans l'aide d'un groupe de facteurs accessoires, les facteurs généraux de transcription (FGT) TBP (TFIID), TFIIB, TFIIA, TFIIE, TFIIF et TFIIH. Ces facteurs recrutent l'ARN pol II au promoteur d'un gène et assistent l'enzyme lors de l'ouverture de la double-hélice d'ADN de -9 à +2, l'initiation de la transcription au site +1 et le dégagement du promoteur. Bien que plusieurs études aient été faites sur les FGT, notre connaissance de leur mécanisme d'action demeure incomplète.

Par le passé, des études de photo-pontage protéine-ADN ont permis de montrer que l'ADN du promoteur est enroulé autour de l'ARN pol II dans un complexe d'initiation comprenant TBP, TFIIB, TFIIE et TFIIF. Ces études ont aussi permis de positionner les FGT par rapport à l'ADN dans ce même complexe. Dans cette thèse, nous avons utilisé la technique de photo-pontage pour localiser TFIIA dans le complexe d'initiation afin de mieux comprendre son rôle à cette étape de la transcription. Ainsi, nous avons pu montrer que TFIIA occupe des positions auxquelles sont aussi retrouvés TFIIF, TFIIE et l'ARN pol II, suggérant une interaction possible entre ces facteurs. Par chromatographie d'affinité, nous avons montré que TFIIA est capable de lier TFIIF et TFIIE et avons identifié les sous-unités impliquées dans ces interactions. Nous avons défini deux domaines de la sous-unité RAP74 de TFIIF, responsables de l'association avec TFIIA et avons démontré la capacité de ce dernier à stimuler l'initiation via RAP74 et la sous-unité β de TFIIE.

Le rôle de TFIIE α lors de l'initiation a aussi été étudié. En effet, la technique de photo-pontage protéine-ADN a permis de localiser la plus grande sous-unité de TFIIE à la position -10 du promoteur, juste en amont de la région qui forme la bulle transcriptionnelle. Nous avons pu observer que TFIIE α stimule l'initiation de façon plus

importante en présence d'une matrice présentant une ouverture artificielle de 2 nucléotides entre -11 et -1. Nos résultats suggèrent que TFIIE α facilite l'ouverture du promoteur en contactant l'ADN en amont de la bulle de transcription, lorsque celle-ci est partiellement ouverte durant les étapes précédant l'initiation.

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 différents modèles pour expliquer la catalyse de formation du lien phosphodiester, l'entrée des nucléotides au site actif et la translocation des acides nucléiques pendant la polymérisation. De plus, des fonctions ont été suggérées pour plusieurs éléments structuraux des sous-unités catalytiques de l'enzyme, Rpb1 et Rpb2, se retrouvant près de la région de l'hybride ADN-ARN et donc près du site actif. Afin de déterminer les rôles de ces éléments dans le processus transcriptionnel, nous avons procédé à la mutagenèse de ceux-ci et à la caractérisation des mutants obtenus à l'aide d'essais fonctionnels. Cette étude a permis de déterminer que les boucles *switch 3* de Rpb2 et les boucles *rudder* et *zipper* de Rpb1 sont impliquées dans la formation du complexe prétranscriptionnel au promoteur. De plus, nous avons identifié un rôle pour la boucle *fork loop 1* de Rpb2, à l'étape d'initiation de la transcription. Finalement, nous montrons que l'acide aminé strictement conservé E791 de Rpb2 est impliqué dans la liaison d'un ion métallique et le chargement du nucléotide au site catalytique pour la polymérisation.

Mots clés : Expression génétique, facteurs généraux de transcription, synthèse des ARN messagers, structure cristallographique, mécanisme catalytique, mutagenèse dirigée, essais fonctionnels, photo-pontage protéine-ADN.

ABSTRACT

RNA polymerase II (RNAPII) synthesizes all messenger RNA (mRNA) in eukaryotic cells. In order to initiate gene-specific transcription at a promoter, RNAPII requires the help of the General Transcription Factors (GTFs) TBP (TFIID), TFIIB, TFIIA, TFIIIE, TFIIF and TFIIH. These factors recruit RNAPII to a gene's promoter and assist the enzyme during opening of the DNA helix between -9 and +2, transcription initiation at the +1 site and promoter escape. Although many studies have been carried out on the GTFs, our knowledge of their precise roles remains incomplete.

Previous protein-DNA photo-crosslinking studies have shown that promoter DNA is wrapped around RNA pol II in a preinitiation complex also including TBP, TFIIB, TFIIF and TFIIIE. These studies have allowed positioning of the GTFs relative to DNA in this complex. In this thesis, we have used the photo-crosslinking technology to localize TFIIA in the initiation complex in order to understand its role at this stage of the transcription process. We have been able to localize TFIIA at positions also occupied by TFIIF, TFIIIE and RNAPII, suggesting a possible interaction between these factors. Using affinity chromatography, we have shown that TFIIA is able to bind to both TFIIF and TFIIIE and have identified the subunits implicated in this interaction. We have identified two domains within the RAP74 subunit of TFIIF involved in the association with TFIIA. Finally we have shown that TFIIA stimulates transcription initiation through its interaction with RAP74 and the β subunit of TFIIIE.

The role of TFIIIE α in transcription initiation was also studied. Protein-DNA photo-crosslinking has allowed us to localize the largest subunit of TFIIIE near nucleotide -10, a position immediately upstream of the transcription bubble in the open complex. Increased stimulation of transcription initiation was obtained using a DNA template artificially melted in the -11/-1 region. Our results suggest that TFIIIE α facilitates open complex formation through its interaction with the upstream end of the partially opened transcription bubble.

Important advances were made in understanding the catalytic mechanism of RNAPII with the publication of high resolution crystallographic structures of the enzyme, either alone, or during transcript elongation on a DNA template. These structures have allowed proposing different models to explain phosphodiester bond synthesis, nucleotide entry to the active site and nucleic acid translocation during polymerization. Furthermore, functions were suggested for a number of structural elements encoded Rpb1 and Rpb2 and found in the region of the DNA-RNA hybrid region of the active site. In order to address experimentally the roles of these elements in the transcription process, we have used site-directed mutagenesis to design various mutants that were then characterized using different functional assays. We have been able to determine that the Rpb2 *switch 3* loop and the Rpb1 *rudder* and *zipper* loops are necessary for preinitiation complex formation on promoter DNA. We have also described a role for Rpb2 *fork loop 1* at the initiation stage of transcription. Finally, we have shown that the strictly conserved Rpb2 amino acid E791 is implicated in binding of a metal ion and in nucleotide loading in the active site for polymerization.

Key words : Gene expression, general transcription factors, transcription initiation, RNA polymerase II, mRNA synthesis, crystallographic structure, catalytic mechanisms, site-directed mutagenesis, functional assays, protein-DNA photo-crosslinking

TABLE DES MATIÈRES

RÉSUMÉ DE LA THÈSE.....	iii
ABSTRACT.....	v
TABLE DES MATIÈRES.....	vii
LISTE DES TABLEAUX.....	x
LISTE DES FIGURES.....	xi
LISTE DES ABRÉVIATIONS.....	xv
CHAPITRE 1	
INTRODUCTION	
1.1 La régulation de la transcription.....	2
1.1.1 La chromatine.....	2
1.1.1.1 Les complexes de remodelage de la chromatine dépendants de l'ATP...3	
1.1.1.2 La modification covalente des queues des histones.....	5
1.1.2 Activateurs.....	7
1.1.3 Répresseurs.....	9
1.1.4 Le complexe médiateur et le concept d'holoenzyme.....	11
1.1.5 Les séquences d'ADN régulatrices.....	16
1.2 Les étapes de la transcription.....	17
1.2.1 Le promoteur basal.....	18
1.2.2 Les facteurs généraux de transcription et l'initiation transcriptionnelle....	21
1.2.2.1 TFIID.....	21
1.2.2.2 TFIIB.....	25
1.2.2.3 TFIIF.....	26
1.2.2.4 TFIIE.....	27
1.2.2.5 TFIIH.....	29

1.2.2.6 TFIIA.....	32
1.2.3 L'élongation de l'ARNm et la terminaison de la transcription.....	33
1.2.3.1 Les facteurs d'élongation.....	34
1.2.3.1.1 Les facteurs d'élongation qui modulent l'activité de l'ARN polymérase II.....	34
1.2.3.1.1.1 TFIIF.....	34
1.2.3.1.1.2 TFIIS.....	35
1.2.3.1.1.3 Autres facteurs d'élongation modulant l'activité de l'ARN polymérase II.....	36
1.2.3.1.2 Les facteurs d'élongation qui modifient la structure de la chromatine.....	37
1.2.3.2 Le rôle du CTD dans le cycle transcriptionnel.....	38
1.2.3.3 La terminaison de la transcription.....	40
1.3 L'ARN polymérase II : structure et mécanisme catalytique.....	41
1.3.1 L'ARN polymérase II.....	41
1.3.2 Structure de l'ARN polymérase II.....	43
1.3.2.1 La structure de l'ARN polymérase II à 10 sous-unités.....	43
1.3.2.2 La structure de l'ARN polymérase II à 12 sous-unités.....	52
1.3.2.3 ARN polymérase II-TFIIS.....	54
1.3.2.4 ARN polymérase II-TFIIB.....	55
1.3.2.5 ARN polymérase II-TFIIF.....	56
1.3.3. Le mécanisme catalytique.....	57
1.4 Objectifs de recherche.....	63

CHAPITRE 2 :

INTERACTIONS STRUCTURALES ET FONCTIONNELLES DU FACTEUR (TF) IIA AVEC TFIIE ET TFIIF DANS L'INITIATION DE LA TRANSCRIPTION PAR L'ARN POLYMÉRASE II

2.1 Préambule.....	64
2.2 Article.....	65

CHAPITRE 3 :

RÔLES CENTRAUX POUR LA PINCE DE L'ARN POLYMÉRASE II ET TFIIE DANS L'INITIATION, RÉVÉLÉS PAR PHOTO-PONTAGE PROTÉINE-ADN D'UN COMPLEXE PRÉTRANSCRIPTIONNEL

3.1 Préambule.....	72
3.2 Article.....	73

CHAPITRE 4 :

RPAP1, UNE NOUVELLE PROTÉINE ASSOCIÉE À L'ARN POLYMÉRASE II PURIFIÉE À PARTIR DE SOUS-UNITÉS RECOMBINANTES SAUVAGES ET MUTANTES

4.1 Préambule.....	84
4.2 Article.....	86

CHAPITRE 5 :

UNE ANALYSE MUTATIONNELLE DU CENTRE CATALYTIQUE DE L'ARN POLYMERASE II

5.1 Préambule.....	103
5.2 Article	104

CHAPITRE 6

DISCUSSION

6.1 Localisation de TFIIA dans le complexe prétranscriptionnel et rôles de TFIIIF et TFIIE dans l'initiation.....	138
6.2 Étude structure-fonction de l'ARN polymérase II : justification de la stratégie expérimentale choisie.....	142
6.3 Structure et mécanisme catalytique de l'ARN polymérase II : au-delà de la cristallographie.....	143
6.4 Les mutants présentant des défauts d'assemblage.....	146
6.5 Un répertoire des mutations affectant la fonction des ARN polymérases à sous-unités multiples, cartographiées sur la structure de l'enzyme.....	148

CONCLUSIONS ET PERSPECTIVES.....	150
----------------------------------	-----

BIBLIOGRAPHIE.....	153
--------------------	-----

LISTE DES TABLEAUX

CHAPITRE I

I. Les facteurs généraux de transcription humains.....	17
II. Les sous-unités de l'ARN polymérase II.....	42
III. Les éléments structuraux du site catalytique de l'ARN polymérase II.....	50

CHAPITRE II

I. Comparaison entre les rôles prédits par les structures cristallographiques et les rôles déterminés par les essais biochimiques pour les éléments structuraux du site catalytique de l'ARN polymérase II.....	144
---	-----

LISTE DES FIGURES

CHAPITRE I

1. Structure de TBP sur la boîte TATA.....	22
2. Structure cristallographique de l'ARN polymérase II à 10 sous-unités.....	45
3. Les canaux de l'ARN polymérase II.....	47
4. Conservation des ARN polymérases.....	48
5. Contacts des éléments structuraux de l'ARN polymérase II avec l'hybride ADN-ARN.....	49
6. Structure de l'ARN polymérase II à 12 sous-unités.....	53
7. Localisation des facteurs de clivage TFIIS et Gre B sur l'ARN polymérase.....	55
8. Le complexe ARN polymérase II-TFIIB.....	56
9. Structure du complexe ARN polymérase II-TFIIF, déterminée par microscopie électronique.....	57
10. Modèle proposé par Kornberg pour expliquer le mécanisme de translocation des acides nucléiques durant la polymérisation.....	59
11. Le modèle de translocation des acides nucléiques dirigée par le NTP arrivant proposé par Burton.....	60
12. La translocation des acides nucléiques chez les ARN polymérases à une sous-unité.....	62

CHAPITRE II

1. Photo-cross-linking of rTFIIA on the AdMLP.....	67
2. Interactions of natural TFIIA with components of the basal transcription machinery.....	68
3. Interactions of TFIIA α/β and TFIIA γ with RAP74,TFIIE56, and TFIIE34.....	68
4. Interactions of TFIIA α/β with RAP74 deletion mutants.....	69
5. Association of RAP74 with a TBP-TFIIA-promoter complex.....	69
6. Stimulation of basal transcription by TFIIA.....	70

CHAPITRE III

1. Photoprobes used for in-gel protein-DNA photo-cross-linking.....	75
2. In-gel photo-cross-linking of the RNAP II preinitiation complex.....	76
3. Summary of our in-gel photo-cross-linking results with a TBP-TFIIB-TFIIF RNAPII-TFIIE promoter complex.....	77
4. In-gel photo-cross-linking data.....	78
5. TFIIE α stimulates transcription initiation.....	80
6. Topological model of the RNAP II preinitiation complex.....	81

CHAPITRE IV

1. TAP of human transcription factors.....	89
2. Purification of TAP-tagged human RNAPII complex.....	90
3. Regions of homology between <i>Homo sapiens</i> RPAP1, <i>D. melanogaster</i> CG32104-PB, and <i>S. cerevisiae</i> Ydr527wp/RBA50, as determined by BLAST analyses.....	92
4. TAP-tagged human RNAPII assembles on promoter DNA both <i>in vitro</i> and <i>in vivo</i> and binds to acetylated histones.....	93
5. TAP-tagged human RNAPII complex can initiate transcription <i>in vitro</i>	94
6. Purification and functional analysis of TAP-tagged human RNAPII with a mutation in fork loop 1.....	96
7. Purification and functional analysis of TAP-tagged human RNAPII with a mutation in switch 3.....	98
8. Association of mutant forms of the RNAPII complex with chromatin <i>in vivo</i>	99

CHAPITRE V

1. Summary of the different RNAP II mutants analyzed.....	131
2. RNAP II mutants of the zipper and the rudder are defective in preinitiation complex formation.....	132

3. Purification of wild type and metal B mutant RNAP II and characterization of their ability to form a pre-initiation complex.....133
4. Mutant MB E791A is sensitive to NTP concentrations during polymerization.....134
5. Mutant MB E791A is sensitive to MgCl₂ concentrations during both NTP polymerization and transcript cleavage. TFIIS stimulates transcription of both wild type RNAP II and mutant MB E791A at low NTP concentrations.....135
6. Association of wild type and mutated RNAP II with chromatin *in vivo*.....136
7. Proposed models for the NTP polymerization reaction.....137

CHAPITRE VI

1. Localisation de TFIIA dans le complexe de préinitiation de l'ARN polymérase II...141

LISTE DES ABRÉVIATIONS

6-AU	6-azauracil
ADN	Acide désoxyribonucléique
ADA	Adaptor
ADP	Adénosine diphosphate
ARC	<i>Activator-recruited factor</i>
ARN	Acide ribonucléique
ARNm	ARN messenger
ARN pol	ARN polymérase
ARNr	ARN ribosomaux
ARNsn	<i>Small nuclear RNA</i>
ARNt	ARN de transfert
ATP	Adénosine triphosphate
ATPase	ATP hydrolase
BRE	<i>TFIIB recognition element</i>
BRCA	<i>Breast and ovarian tumor suppressor protein</i>
CAK	<i>Cyclin-dependant kinase activating kinase</i>
CBP	<i>CREB-binding protein et Calmodulin-binding peptide</i>
CBD	<i>Chitin-binding domain</i>
CCD	<i>Charge clustered region</i>
Cdc	<i>Cell division cycle</i>
CDK	<i>Cyclin dependant kinase</i>
CHD	<i>Chromodomain SWI/SNF-like helicase/ATPase domain and DNA binding domain</i>
ChIP	Immunoprécipitation de la chromatine
CHRAC	<i>Chromatin accessibility complex</i>
CoTC	<i>Cotranscriptional cleavage</i>
CRSP	<i>Cofactor required for Sp1 activation</i>
CS	Syndrome de Cockayne

CTD	<i>Carboxy-terminal domain</i>
CTDK-1	<i>CTD kinase 1</i>
C-terminal	Carboxy-terminal
CPF	<i>Cleavage/polyadenylation factor</i>
CPSF	<i>Cleavage and polyadenylation specificity factor</i>
CstF	<i>Cleavage stimulatory factor</i>
CTP	Cytosine triphosphate
dNTP	Désoxyribonucléotide triphosphate
DMT	<i>DNA methyltransferase</i>
DPE	<i>Downstream promoter element</i>
DRIP	<i>Vitamin-D-receptor interacting proteins</i>
DRB	5,6-dichloro-1-b-D-ribofuranosylbenzimidazole
DSE	<i>Downstream sequence element</i>
DSIF	<i>DRB-sensitivity inducing factor</i>
ELL	Eleven-nineteen lysine-rich in leukemia
Fcp1	<i>TFIIF-associated CTD phosphatase</i>
FGT	Facteurs généraux de transcription
FIR	<i>FBP interaction repressor</i>
GAL	Galactose
GCN	<i>General control of amino acid synthesis</i>
GTP	Guanosine triphosphate
HAT	<i>Histone acétyltransférase</i>
His	Histidine
HBX	<i>Hepatitis B virus X protein</i>
HDAC	<i>Histone déacétylase</i>
HMG	<i>High mobility group</i>
HMT	Histone méthyltransférase
INF	Interferon
Inr	Initiateur
ISWI	<i>Imitation of SWI</i>
LCR	<i>Locus control region</i>

MAT	Ménage à trois
Med	Médiateur
MudPIT	<i>Multidimensional protein identification technology</i>
MTE	<i>Motif ten element</i>
NAT	<i>Negative regulator of activated transcription</i>
NC	<i>Negative cofactor</i>
N-CoR	<i>Nuclear receptor corepressor</i>
NELF	<i>Negative elongation factor</i>
NER	<i>Nucleotide excision repair</i>
NF	<i>Nuclear factor</i>
N-terminal	Amino-terminal
NTP	Ribonucléotide triphosphate
NuRD	<i>Nucleosome remodeling factor</i>
ORC	<i>Origin recognition complex</i>
p	Polypeptide
PAP	Poly(A) polymérase
pb	Paire de base
PCAF	<i>p300/CBP-associated factor</i>
PC	<i>Positive cofactor</i>
PABP II	<i>Poly(A) binding protein II</i>
PPAR	<i>Peroxisome proliferator-activated receptor</i>
P-TEFb	<i>Positive transcription elongation factor b</i>
RAP	<i>RNA polymerase II-associated protein</i>
RAR	<i>Retinoic acid receptor</i>
Rpb	RNA polymerase B
RSC	<i>Remodel the structure of chromatin</i>
RXR	<i>Retinoid X receptor</i>
SAGA	SPT-ADA-GCN5 acétyltransférase
SANT	SWI3, ADA2, N-CoR, TFIIB
Scp1	<i>Small CTD-phosphatase 1</i>
Ser	Sérine

SMCC	<i>SRB-Med-containing cofactor</i>
SNF	<i>Sucrose non fermenting</i>
Spt	<i>Supressor of Ty insertion</i>
Srb	<i>Supressor of RNA polymerase B</i>
SREBP-1	<i>Sterol response element binding protein</i>
STAGA	<i>Supressor of Ty insertion-TAF-Gcn5-acetyltransferase</i>
Strep	Streptavidine
SWI	<i>Switching mating-type</i>
TAC	<i>TBP-TFIIA-containing complex</i>
TAF	<i>TBP-associated factor</i>
TAP	<i>Tandem affinity purification</i>
TBP	<i>TATA binding protein</i>
TCR	<i>Transcription coupled repair</i>
TF	<i>Transcription factor</i>
TFTC	<i>TBP-free TAF-containing complex</i>
TR	<i>Thyroid hormone receptor</i>
TRF	<i>TBP related factor</i>
TRAP	<i>Thyroid hormone receptor</i>
TTD	Trichothiodystrophie
UAS	<i>Upstream activating sequence</i>
URS	<i>Upstream repressing sequence</i>
UTP	Uridine triphosphate
UV	Ultraviolet
VDR	<i>Vitamin D receptors</i>
VIH	Virus de l'immunodéficience acquise
VP16	<i>Herpes simplex virion protein 16</i>
XP	<i>Xerodema pigmentosum</i>

REMERCIEMENTS

Je remercie Benoit Coulombe de m'avoir accueillie dans son laboratoire en 1998 pour mon premier stage en recherche, puis de m'avoir permis de réaliser ma maîtrise et mon doctorat sous sa direction. Ainsi, j'ai pu acquérir une formation de très grande qualité grâce à ses précieux conseils, à son encadrement rigoureux et à sa grande disponibilité. Benoit m'a aussi donné la chance de participer à de nombreux congrès internationaux, ainsi qu'à un stage dans un laboratoire en France.

Je remercie Diane Forget pour tous ses conseils et suggestions, pour son support technique, pour le rôle très important qu'elle joue quotidiennement dans l'organisation du laboratoire, pour ses nombreux encouragements et pour sa foi inébranlable dans ma capacité à réussir.

Je remercie tous les membres du laboratoire que j'ai côtoyé durant ces années qui m'ont aidé dans l'avancement de mon projet tant par des conseils techniques que par des discussions scientifiques. Je les remercie aussi pour tous les encouragements et pour les moments inoubliables passés ensemble.

Je remercie ma famille et mes amis pour leurs encouragements, leur grande compréhension et leur soutien.

CHAPITRE 1

INTRODUCTION

Les ARN polymérasés (ARN pol) dépendantes de l'ADN sont des enzymes qui catalysent la formation d'un ARN à partir d'une matrice d'ADN. Chez les procaryotes, une seule ARN pol synthétise tous les types d'ARN cellulaires. Chez les organismes eucaryotes, il existe cependant trois types d'ARN polymérasés, les ARN pol I, II et III, ayant des fonctions spécifiques. L'ARN pol I est responsable de la synthèse des ARN ribosomiaux (ARNr), l'ARN pol II, des ARN messagers (ARNm) et des ARNsn (*small nuclear RNA*) et l'ARN pol III, des ARN de transfert (ARNt) et de l'ARNr 5s.

L'ARNm constitue un code pour la synthèse d'une protéine, par le mécanisme de la traduction cellulaire. Il existe dans la cellule eucaryote primitive de levure environ 5500 gènes codant pour des protéines (Blandin *et al.*, 2000). Chez l'humain, le nombre passe à 20 000 (International Human Genome Sequencing Consortium, 2004). La transcription de ces gènes est un processus qui doit être hautement régulé pour permettre à la cellule de répondre à son environnement, de se spécialiser dans une fonction particulière au sein d'un organisme et de se diviser adéquatement. Pour ce faire, l'ARN pol II nécessite l'aide d'une pléiade de protéines et de séquences d'ADN régulatrices, qui interviennent dans le processus transcriptionnel. Il existe donc des protéines qui modulent l'organisation de l'ADN, qui, dans la cellule eucaryote, se retrouve sous forme de chromatine. D'autres protéines, les activateurs transcriptionnels, stimulent la transcription en se liant à des séquences régulatrices spécifiques sur l'ADN et à des coactivateurs. Les répresseurs et les corépresseurs inhibent l'expression de gènes par des mécanismes variés. Les facteurs généraux de transcription (FGT) TFIID, TFIIB, TFIIA, TFIIF, TFIIE et TFIIH permettent le recrutement de l'ARN pol II au promoteur d'un gène et l'initiation de la synthèse de l'ARNm au site +1. Finalement, plusieurs facteurs modulent l'élongation du transcrit et la terminaison de la transcription.

La publication de structures cristallographiques à haute résolution de l'ARN pol bactérienne et de l'ARN pol II de levure, seules ou en complexe avec l'ADN ou des facteurs de régulation, a permis de proposer différentes hypothèses et modèles pour expliquer le fonctionnement de l'enzyme. Cependant, peu d'expériences biochimiques ont été réalisées à ce jour afin de confirmer ou d'infirmer ces derniers. Dans cette thèse, nous nous sommes intéressés aux rôles des FGT dans le processus transcriptionnel, plus particulièrement à ceux de TFIIA, TFIIF et TFIIIE, ainsi qu'au mécanisme catalytique de l'ARN pol II en relation avec sa structure. L'introduction présentée traitera donc de la régulation de la transcription par l'ARN pol II, des différentes étapes du processus transcriptionnel et du rôle des FGT dans certaines de ces étapes, de la structure de l'ARN pol II et de son mécanisme catalytique.

1.1 La régulation de la transcription

1.1.1 La chromatine

Dans le noyau eucaryote, l'ADN est retrouvé sous une forme condensée : la chromatine. La fibre de chromatine renferme différents niveaux de compaction de l'ADN. À la base on retrouve le nucléosome, formé d'histones (H2A, H2B, H3 et H4) (revu par Horn et Peterson, 2002). Les histones sont des protéines basiques qui contiennent deux domaines fonctionnels : un domaine appelé *histone-fold*, qui supporte les interactions histone-histone et histone-ADN et un domaine N-terminal riche en lysine. Le domaine N-terminal ainsi que la région C-terminale des histones peuvent être le site de nombreuses modifications post-traductionnelles (principalement l'acétylation, la méthylation, la phosphorylation et l'ubiquitination). La structure cristallographique d'un nucléosome a révélé que celui-ci était constitué de 146 paires de base (pb) d'ADN enroulées 1,7 fois autour d'un octamère des histones H2A, H2B, H3 et H4, chacune présente en 2 copies (Luger *et al.*, 1997). Les queues N-terminales des histones sont projetées vers l'extérieur de la structure où elles peuvent interagir avec des protéines régulatrices, avec l'ADN de liaison (ADN *linker*), situé entre deux nucléosomes, ainsi qu'avec l'ADN des nucléosomes avoisinants. Ces interactions entre nucléosomes vont

permettre la formation d'une structure plus compacte appelée la fibre de chromatine de 30 nm (revu par Woodcock et Dimitrov, 2001). L'histone H1 ou histone *linker*, qui interagit avec chaque nucléosome, va stabiliser la structure de la fibre chromatinienne. Les fibres de chromatine peuvent encore interagir entre elles pour former une structure plus condensée de 100 à 300 nm pendant la métaphase.

Un niveau aussi élevé de compaction est réfractaire aux différents processus biologiques de l'ADN, tels la réplication, la réparation et la transcription, constituant ainsi un mode de régulation pour ces derniers. Deux familles d'enzymes existent dans la cellule qui, en réponse à des signaux d'activation ou de répression, vont modifier la structure de la chromatine afin de favoriser ou réprimer la transcription de gènes cibles : les complexes de remodelage de la chromatine dépendants de l'ATP et les complexes qui modifient les histones de façon covalente.

1.1.1.1 Les complexes de remodelage de la chromatine dépendants de l'ATP

Les complexes de remodelage de la chromatine utilisent l'énergie de l'ATP afin de modifier les interactions histones-ADN (revu par Lusser et Kadonaga, 2003). Tous les complexes de remodelage possèdent une sous-unité ATPase faisant partie de la famille *Snf2-like* (Eisen *et al.*, 1995). Cette famille est subdivisée en sous-familles en fonction de la présence de domaines protéiques spécifiques, distincts du motif ATPase (Lusser et Kadonaga, 2003). Par exemple, tous les membres de la sous-famille SNF2, incluant les sous-unités ATPase des complexes SWI/SNF, RSC et hSWI/SNF, possèdent un bromodomaine. Les membres de la sous-famille ISWI (RSF, hCHRAC, NoCR) possèdent tous un domaine SANT (SWI3, ADA2, N-CoR, TFIIB), alors que les membres de la famille CHD1 (NuRD, Mi-2) contiennent un chromodomaine et un domaine de liaison à l'ADN. Le rôle de ces domaines protéiques dans la régulation transcriptionnelle sera discuté plus loin.

L'un des complexes de remodelage de la chromatine les mieux caractérisés est le complexe SWI/SNF (Martens et Winston, 2003). Le complexe SWI/SNF a été identifié

génétiqnement chez la levure comme étant impliqué dans la régulation de la transcription du gène de l'endonucléase HO, qui joue un rôle dans le processus de changement de type sexuel (SWI : *switching mating-type*) et dans la transcription du gène *SUC2* (SNF : *sucrose non-fermenting*). Plus tard, il a été démontré que SWI/SNF était impliqué dans la régulation d'un nombre restreint de gènes chez la levure (environ 5 %) en conditions de croissance sur milieu riche, soit en activant ou en réprimant la transcription (Holstege *et al.*, 1998; Martens et Winston, 2003). Chez l'humain, hSWI/SNF est requis pour la transcription de groupes de gènes, incluant les gènes de réponse au choc thermique (de, I *et al.*, 2000), aux récepteurs stéroïdiens (Fryer et Archer, 1998) et les gènes impliqués dans la progression du cycle cellulaire (Zhang *et al.*, 2000). Le complexe RSC, identifié chez la levure, semble jouer un rôle plus général dans la régulation transcriptionnelle. Des expériences d'immunoprécipitation de la chromatine (ChIP) ont montré qu'il était recruté sur plusieurs promoteurs de l'ARN pol III de façon basale et sur certains promoteurs de l'ARN pol II par des activateurs ou des répresseurs (Ng *et al.*, 2002).

Plusieurs modèles ont été proposés afin d'expliquer comment les complexes de remodelage de la chromatine modifient les interactions histones-ADN de façon à moduler l'activité transcriptionnelle (Saha *et al.*, 2002; Aoyagi et Hayes, 2002; Whitehouse *et al.*, 2003). Des expériences biochimiques semblent indiquer qu'il existerait des mécanismes différents de remodelage selon les familles de complexes (Lusser et Kadonaga, 2003). Par exemple, un modèle de translocation a été proposé pour les complexes de remodelage faisant partie de la famille ISWI, qui participent à l'assemblage des nucléosomes et augmentent la stabilité de la chromatine (Ito *et al.*, 1997; LeRoy *et al.*, 2000; Gelbart *et al.*, 2001). Dans ce modèle, le complexe de remodelage défait les contacts entre les histones et l'ADN en se déplaçant sur la chromatine. Après le passage du complexe, les contacts sont rétablis de façon à ce que le nucléosome se retrouve à une position différente de sa position initiale. En ce qui concerne les complexes de la sous-famille SNF2 qui, dans les essais biochimiques altèrent la structure du nucléosome (Imbalzano *et al.*, 1996; Jaskelioff *et al.*, 2000; Lorch *et al.*, 2001), différents modèles ont été suggérés, dont un modèle de transfert latéral de l'ADN (Lusser et Kadonaga, 2003). Dans ce modèle, le complexe de remodelage catalyse le déplacement d'un tour d'ADN

nucléosomal sur l'octamère d'histone de façon latérale, alors que le tour suivant est déplacé du nucléosome. Ceci conduit à l'exposition d'une nouvelle portion d'ADN.

1.1.1.2 La modification covalente des queues des histones

Les queues N-terminales des histones sont sujettes à des modifications covalentes au niveau de certains de leurs résidus (revu par Khorasanizadeh, 2004). Les histones peuvent être acétylées (lysines), méthylées (lysines, arginines), phosphorylées (sérines, thréonines), ubiquitinées (lysines), sumoylées (lysines) ou ADP-ribosylées (acides glutamiques). L'acétylation des histones par les histones acétyltransférases (HAT) est la mieux caractérisée de ces modifications. Il existe plusieurs familles de HAT, dont les familles GCN5/PCAF, CBP/p300 et MYST, définies selon l'homologie de séquence de ses membres et ayant des spécificités différentes pour les résidus et les histones acétylées (Marmorstein, 2001). Les premières études semblaient montrer que les membres de la famille GCN5/PCAF fonctionnaient comme coactivateur pour un certain nombre de gènes seulement (Holstege *et al.*, 1998; Lee *et al.*, 2000). Une étude plus récente de localisation à l'échelle génomique montre cependant que GCN5 et Esa1 de la famille MYST se retrouvent sur les éléments régulateurs de la plupart des gènes activement transcrits (Robert *et al.*, 2004). Tous les HAT caractérisées à ce jour sont retrouvées sous la forme de grands complexes *in vivo*. Par exemple, Gcn5 de levure se retrouve dans les complexes SAGA et Ada (Grant *et al.*, 1997) et PCAF fait partie d'un complexe *SAGA-like* (Ogryzko *et al.*, 1998). La formation de ces différents complexes incluant les HAT module la spécificité d'acétylation de ces derniers.

L'acétylation des lysines est une modification réversible grâce aux histones déacétylases (HDAC). Les HAT et les HDAC interagissent avec différents facteurs de transcription pour permettre leur recrutement au promoteur (Featherstone, 2002). Au promoteur de l'INF- β , les facteurs de transcription NF- κ B, IRF1 et l'hétérodimère ATF2/c-Jun se lient de façon coopérative à leurs sites de reconnaissance sur l'*enhancer*, qui est exempt de nucléosome, en présence de la protéine architecturale HMG1(Y) pour former un grand complexe appelé l'*enhanceosome*. L'*enhanceosome* constitue une

nouvelle surface d'interactions qui permet le recrutement d'un complexe HAT contenant GCN5. Ce dernier acétyle le nucléosome, qui masque la boîte TATA et le site d'initiation du promoteur permettant le recrutement de la machinerie transcriptionnelle basale (l'ARN pol II et les FGT) (Merika et Thanos, 2001; Merika *et al.*, 1998).

Les autres modifications des histones ont différents rôles à jouer dans la cellule. La méthylation des lysines par les histones méthyltransférases (HMT) est importante pour le processus de transfert épigénétique (Lachner *et al.*, 2003). Au cours de ce processus, des régions chromosomiques sont marquées par la méthylation afin de conserver un patron spécifique d'activité transcriptionnelle suite à la division cellulaire. Les histones méthylées sont ensuite reconnues et liées par des protéines impliquées dans la condensation de la chromatine et la répression. La méthylation des lysines peut aussi être liée à l'activation transcriptionnelle (Zhang et Reinberg, 2001). L'ubiquitination des lysines et la méthylation des arginines sont des modifications qui corrént avec l'activation de la transcription (Zhang, 2003; Zhang et Reinberg, 2001), alors que la sumoylation de l'histone H4 est importante pour la répression transcriptionnelle (Shiio et Eisenman, 2003). Finalement, la phosphorylation des sérines est impliquée dans la condensation des chromosomes durant la mitose et est aussi liée à l'activation transcriptionnelle. Plusieurs enzymes ont été identifiées comme étant capables d'inverser ces modifications (Khorasanizadeh, 2004). Ubp8, une sous-unité du complexe SAGA, catalyse le retrait de l'ubiquitine des lysines de l'histone H2B. Des phosphatases de la famille PP1 (protein phosphatase 1) permettent la déphosphorylation des sérines de l'histone H3. Récemment, une première histone déméthylase a été identifiée. Cette enzyme appelée LSD1 déméthyle spécifiquement la lysine 4 des histones H3, réprimant ainsi la transcription de gènes cibles (Shi *et al.*, 2004).

Différents modèles tentent d'expliquer comment les modifications covalentes des histones influencent la structure de la chromatine de façon à permettre la transcription. Les premiers modèles proposaient que l'annulation d'une charge positive sur une lysine par l'acétylation diminuerait la stabilité de l'interaction des queues N-terminales avec l'ADN internucléosomal et avec l'ADN des nucléosomes voisins permettant une

ouverture de la chromatine. Un modèle plus récent propose que les modifications post-traductionnelles spécifiques des histones constituent un code (*histone code model*) qui est ensuite reconnu par d'autres facteurs, comme des enzymes remodelant la chromatine ou des facteurs généraux de transcription, conduisant à l'initiation de la transcription (Strahl et Allis, 2000; Agalioti *et al.*, 2002). Par exemple, au promoteur de l'INF- β , suite au recrutement de GCN5 par l'*enhanceosome* et à l'acétylation du nucléosome masquant la boîte TATA et le site d'initiation (voir ci-haut), un complexe CBP-ARN pol II holoenzyme est recruté. Le complexe de remodelage de la chromatine SWI/SNF arrive ensuite au promoteur, via des interactions avec CBP et les queues des histones acétylées (Agalioti *et al.*, 2000). SWI/SNF remodèle la chromatine permettant la liaison de TFIID à la boîte TATA. Ce dernier induit une courbure dans la boîte TATA qui entraîne le glissement du nucléosome vers une position plus en aval, libérant le site +1 et permettant l'initiation (Lomvardas et Thanos, 2001). Il a été démontré que l'acétylation des histones du nucléosome à ce promoteur par GCN5 constituait un code reconnu par SWI/SNF et TFIID, essentiel à leur recrutement. (Agalioti *et al.*, 2002).

L'existence de bromodomaines, de chromodomaines et de domaines SANT, capables d'interagir respectivement avec des histones acétylées, méthylées et non-acétylées, sur un grand nombre de protéines associées à la chromatine supporte l'hypothèse du code des histones (Marmorstein et Berger, 2001; Jacobson *et al.*, 2000; Yu *et al.*, 2003). En effet, tous les complexes de remodelage de la chromatine contiennent une sous-unité possédant l'un ou l'autre de ces domaines. Les bromodomaines sont aussi retrouvés chez plusieurs coactivateurs dont GCN5, p300/CBP et TAFII250 (Marmorstein et Berger, 2001).

1.1.2 Activateurs

Les activateurs, aussi appelés facteurs de transcription liant l'ADN, sont des protéines modulaires constituées de différents domaines ayant des fonctions spécifiques dans le processus d'activation (revu par Triezenberg, 1995; Remenyi *et al.*, 2004). Tout d'abord, ils possèdent un domaine de liaison à l'ADN qui leur permet de reconnaître et de

lier un site spécifique sur ce dernier. La reconnaissance du site spécifique peut se faire en collaboration avec d'autres activateurs, qui interagissent entre eux via des domaines de coopérativité. Le domaine d'activation permet quant à lui le recrutement de la machinerie transcriptionnelle au promoteur, soit via des interactions directes avec les FGT, soit via des interactions avec des coactivateurs. Dans certains cas, les domaines de liaison à l'ADN et d'activation se retrouvent sur la même protéine, comme dans le cas des activateurs GAL4 et GCN4. Dans d'autres cas, ces domaines se retrouvent sur des protéines différentes, comme dans le cas du complexe Oct-1, où la protéine Oct-1 lie l'ADN et les protéines OCA-B et VP16 agissent comme coactivateurs en permettant le recrutement de la machinerie transcriptionnelle au promoteur.

Il existe chez la cellule eucaryote environ 20 000 gènes, dont l'expression doit être finement contrôlée par un nombre restreint de régulateurs transcriptionnels (2000-3000). Pour ce faire, les activateurs utilisent un mode de contrôle combinatoire (Remenyi *et al.*, 2004). Par exemple, les récepteurs nucléaires PPAR (*peroxisome proliferator-activated receptors*), VDR (*vitamin D receptors*), TR (*thyroid hormone receptors*) et RAR (*retinoic acid receptors*) se lient à leur site de reconnaissance sur l'ADN sous forme d'hétérodimère avec la protéine RXR (*retinoid X receptor*). Le site de reconnaissance de ces facteurs de transcription est constitué de deux répétitions directes, l'une liée par un récepteur nucléaire spécifique et l'autre par le partenaire d'interaction RXR. La spécificité de reconnaissance se fait grâce à l'espacement entre les répétitions directes : PPAR, VDR, TR et RAR se lient préférentiellement à des sites espacés respectivement de 1, 3, 4 et 5 nucléotides (Glass, 1994). Dans cet exemple, une famille de facteurs de transcription est capable de reconnaître un grand nombre de sites de régulation différents à l'aide d'un seul partenaire d'interaction.

Un autre exemple est celui des membres de la famille des facteurs POU. Les facteurs POU peuvent homodimériser et hétérodimériser sur des sites spécifiques sur l'ADN. Les facteurs POU Oct-1 et Oct-2 homodimérisent sur les éléments de réponse PORE et MORE, retrouvés sur des gènes spécifiquement exprimés dans les cellules B. La liaison de PORE par l'homodimère d'Oct-1 ou d'Oct-2 permet le recrutement du

coactivateur OBF-1, alors que l'organisation des mêmes homodimères sur MORE bloque le site de liaison au coactivateur, empêchant l'interaction avec ce dernier (Tomilin *et al.*, 2000).

1.1.3 Répresseurs

Les répresseurs et corépresseurs transcriptionnels répriment la transcription via des mécanismes variés, qui peuvent être classés en trois catégories. Les répresseurs peuvent inhiber la machinerie transcriptionnelle basale, contrecarrer la fonction d'un activateur ou encore participer au remodelage et à la compaction de la chromatine (revu par Gaston et Jayaraman, 2003).

Les répresseurs qui inhibent la transcription basale ont différents modes d'action. La protéine Mot1, qui avait été identifiée initialement comme un répresseur général de la transcription, est maintenant reconnue comme étant capable d'avoir un effet positif ou négatif sur la transcription, spécifique à certains gènes (Dasgupta *et al.*, 2002). Mot1 s'associe au complexe TBP-boîte TATA et peut dissocier TBP de l'ADN, une activité dépendante de l'ATP (Gumbs *et al.*, 2003; Auble *et al.*, 1994). *In vitro*, Mot1 inhibe la transcription et cette inhibition peut être contrée par TFIIA (Auble *et al.*, 1994). *In vivo*, Mot1 est capable de déplacer un TBP seul sur l'ADN, mais pas un TBP en complexe avec d'autres protéines, d'une localisation génomique inappropriée (Geisberg *et al.*, 2002). D'autres répresseurs bloquent l'interaction entre certains facteurs généraux de transcription, réprimant ainsi la transcription. La protéine NC2 lie TBP et empêche son interaction avec TFIIA (Meisterernst et Roeder, 1991). MDM2 s'associe à TFIIE et TBP et interfère avec la transcription basale (Thut *et al.*, 1995). Le répresseur N-CoR, qui bloque l'activation de la transcription par les récepteurs nucléaires lorsque ceux-ci ne sont pas liés à leur ligand, interagit avec TAF32, TFIIB et TAF70, inhibant l'interaction entre TFIIB et TAF32 (Muscat *et al.*, 1998).

D'autres répresseurs régulent la transcription en contrecarrant la fonction d'activateurs, soit en réduisant la stabilité de la protéine régulatrice, soit en contrôlant sa

localisation cellulaire ou encore en inhibant son interaction avec l'ADN ou avec un coactivateur. Par exemple, MDM2 permet l'ubiquitination de p53 et favorise ainsi son exportation du noyau vers le cytoplasme pour la dégradation par le protéasome (Inoue *et al.*, 2001). Le facteur de transcription NF κ B est régulé par la protéine I κ B, qui masque son signal de localisation nucléaire, l'empêchant ainsi d'être importé au noyau et d'agir comme activateur (Verma *et al.*, 1995). La liaison entre I κ B et NF κ B est elle-même régulée par des kinases et des enzymes d'ubiquitination qui vont conduire à la dégradation du répresseur. Chez la levure, la protéine Tup1 réprime la transcription, entre autres, en bloquant l'interaction entre Srb7, un composant de l'holoenzyme de l'ARN Pol II, et le coactivateur Med6. Comme la liaison entre Srb7 et Med6 est essentielle pour la fonction de plusieurs activateurs, Tup1 a donc pour effet de réprimer l'expression des gènes régulés par ces derniers (Gromoller et Lehming, 2000). Plusieurs études montrent que Tup 1 inhibe aussi la transcription en se liant à des histones hypoacétylées et par le recrutement de HDAC (Zhang et Reese, 2004).

Le dernier groupe de répresseurs inhibe la transcription en participant au remodelage et la compaction de la chromatine. Tel que mentionné plus haut, la structure de la chromatine joue un rôle important dans la régulation de l'expression génétique. Jusqu'à tout récemment, il était accepté que les régions du génome où la chromatine est la moins condensée, régions appelées euchromatine, étaient les plus propices à l'activité transcriptionnelle. Inversement, les régions d'ADN hautement condensées, appelées hétérochromatine, contiennent peu de gènes activement transcrits. Toutefois, une étude récente montre que la distinction entre l'hétérochromatine et l'euchromatine n'est pas aussi simple (Gilbert *et al.*, 2004). En effet, il a été démontré que des régions de chromatine compacte se retrouvent à la fois dans l'hétérochromatine et l'euchromatine. De plus, bien que les régions de chromatine dites ouvertes soient les régions qui contiennent le plus grand nombre de gènes, ces régions ne corrèlent pas toujours avec la plus grande activité transcriptionnelle. L'étude montre aussi que des gènes actifs se retrouvent dans les régions de chromatine compacte. Les processus de compaction et de décompaction de la chromatine, bien qu'ils jouent des rôles importants dans la régulation transcriptionnelle, ne dictent pas toujours le niveau d'expression des gènes.

La méthylation des dinucléotides CpG par les ADN méthyltransférases (DMT) est importante pour le processus de répression transcriptionnelle. L'hétérochromatine est généralement hautement méthylée au niveau des CpG, alors que l'euchromatine contient à la fois des régions très méthylées et d'autres non-méthylées. La méthylation des CpG inhibe la transcription de différentes façons. Par exemple, la méthylation peut bloquer la liaison d'activateurs en modifiant la structure de la chromatine ou en recrutant des protéines liant les méthyl-CpG (*methyl-CpG-binding proteins*) (Gaston et Fried, 1995). En plus de compétitionner avec la liaison des activateurs à l'ADN, les protéines liant les méthyl-CpG peuvent, dans certains cas, recruter des HDAC qui vont à leur tour modifier la structure de la chromatine (Nan *et al.*, 1998; Zhang *et al.*, 1999b).

La famille des protéines Sir est un bon exemple de protéines qui répriment la transcription en modifiant la structure de la chromatine (Gasser et Cockell, 2001). Les protéines Sir sont requises pour la répression au niveau des télomères (*telomeric silencing*) et pour le choix du type sexuel chez la levure *S. cerevisiae*. Ces protéines, dont certaines ont une affinité pour les histones H3 et H4 déacétylées (Sir3p et Sir4p), homodimérisent et hétérodimérisent sur l'ADN pour former un état hautement réprimé qui se propage le long de la fibre de chromatine. Au niveau des télomères, les protéines Sir3p et Sir4p sont recrutées par Rap1p, une protéine liant l'ADN, et le complexe Ku70/80. Au niveau du loci du changement de type sexuel chez la levure, la protéine ORC recrute Sir1p, qui elle participe au recrutement de Sir4P. Sir3p et Sir4p se lient directement aux queues N-terminales des histones et Sir4p participe au recrutement de Sir2, une histone déacétylase qui va modifier la structure de la chromatine.

1.1.4 Le complexe médiateur et le concept d'holoenzyme

Le complexe médiateur a d'abord été identifié chez la levure comme un complexe coactivateur, nécessaire pour supporter l'activation par un activateur transcriptionnel (revu par Myers et Kornberg, 2000). En effet, l'ARN pol II et les FGT seuls, ne peuvent répondre aux activateurs. Une première étude a permis d'identifier le complexe médiateur

comme étant capable de contrer la capacité d'interférence d'un activateur sur un autre *in vivo* (Kelleher, III *et al.*, 1990; Flanagan *et al.*, 1991). Cette interférence avait été attribuée à la séquestration d'un facteur présent en quantité limitante dans les cellules. Lorsque l'expérience de séquestration a été reproduite avec des extraits nucléaires de levure, ni l'ARN pol II, ni les facteurs généraux de transcription, ajoutés individuellement en excès, n'ont pu contrer l'effet de l'activateur (Flanagan *et al.*, 1991; Kelleher, III *et al.*, 1990). En revanche, une fraction nucléaire impure ajoutée en excès a pu relever l'interférence. L'activité présente dans cette fraction a été appelée médiateur et il a été proposé que celui-ci serve d'adaptateur entre un activateur et la machinerie transcriptionnelle basale.

D'autres études menées en parallèle ont aussi conduit à l'identification du médiateur et à la proposition du concept de l'holoenzyme (Koleske et Young, 1994; Thompson *et al.*, 1993; Kim *et al.*, 1994). Une première étude avait pour but la purification des protéines Srb (*Suppressor of RNA polymerase B*) de la levure. Les protéines Srb avaient été identifiées, lors d'un criblage génétique chez *S. cerevisiae*, comme étant des supprimeurs dominants (Srb2, Srb4, Srb5, Srb6) ou récessifs (Srb7 à Srb11) des phénotypes de troncation du domaine carboxy-terminal (CTD) de la sous-unité Rpb1 de l'ARN pol II (Nonet et Young, 1989). Le CTD de Rpb1 est composé d'un heptapeptide (YSPTSPS) répété de 26 fois chez la levure à 52 fois chez l'humain (Nonet *et al.*, 1987). Chez la levure, une troncation complète du CTD est létale, alors que des délétions partielles confèrent des phénotypes de sensibilité au froid (*cold-sensitive*), supprimés par des mutations dans les gènes *Srb*. La tentative de purification des Srb par chromatographie conventionnelle a permis d'isoler un grand complexe contenant non seulement ces dernières, mais aussi l'ARN pol II et certains facteurs généraux de transcription (Koleske et Young, 1994; Thompson *et al.*, 1993). Ce complexe, qui avait la capacité d'initier la transcription de façon spécifique et de répondre à des activateurs lorsque supplémenté avec les FGT manquants purifiés, a été appelé holoenzyme (Koleske et Young, 1994). En incubant l'holoenzyme avec des anticorps dirigés contre le CTD, on a pu séparer l'holoenzyme en différents sous-complexes : l'un contenant l'ARN pol II et

les FGT et l'autre contenant les Srb (Kim *et al.*, 1994). Ce dernier sous-complexe était en fait le médiateur, le complexe capable de supporter l'activation.

Chez *S. cerevisiae*, le médiateur est un complexe composé de plus de 20 sous-unités (revu par Myers et Kornberg, 2000; Lee et Young, 2000; Lewis et Reinberg, 2003), dont la nomenclature a été récemment révisée afin de l'uniformiser et de permettre une comparaison entre les médiateurs des différentes espèces (Bourbon *et al.*, 2004). Il contient des protéines Srb, Srb2 (MED20), 4 (MED17), 5 (MED18) et 6 (MED22), qui participent à l'activation transcriptionnelle. Par exemple, la protéine Srb4 est une cible directe de l'activateur GAL4 (Koh *et al.*, 1998). Srb4 et Srb6 sont codés par des gènes essentiels et sont requis pour la transcription de la plupart des gènes de classe II (transcrits par l'ARN pol II) (Holstege *et al.*, 1998; Thompson *et al.*, 1993; Thompson et Young, 1995). D'autres protéines Srb, Srb8 (MED12), 9 (MED9), 10 (CDK8) et 11 (Cyclin C) sont impliquées dans la répression de groupes de gènes particuliers. Srb10, une kinase cycline-dépendante et sa partenaire cycline Srb11 forment un complexe capable de phosphoryler le CTD de l'ARN pol II avant l'initiation, l'empêchant d'être recrutée au promoteur (Hengartner *et al.*, 1998). En condition de croissance exponentielle sur milieu riche, Srb10 est impliqué dans la répression de gènes qui codent pour des protéines impliquées dans les changements morphologiques en réponse au stress et au manque de nutriments (Holstege *et al.*, 1998). Le médiateur de levure contient aussi des protéines Med, dont la protéine Med6 (MED6), requise pour l'expression d'environ 10% des gènes (Holstege *et al.*, 1998). Med6 est aussi requis pour l'activation par VP16 (Lee *et al.*, 1997).

Chez les mammifères, plusieurs médiateurs à composition variable ont été purifiés (Myers et Kornberg, 2000; Lewis et Reinberg, 2003). Les médiateurs de mammifères incluent les très grands complexes de 1 à 2 MDa TRAP/SMCC (*thyroid hormone receptor/SRB-Med-containing cofactor*) (Malik et Roeder, 2000), ARC (*activator-recruited factor*) (Naar *et al.*, 1999), DRIP (*vitamin D receptor-interacting proteins*) (Rachez *et al.*, 1999), NAT (*negative regulator of activated transcription*) (Sun *et al.*, 1998), les médiateurs murins (Sato *et al.*, 2003; Jiang *et al.*, 1998), et les complexes plus

petits entre 500 et 700 kDa PC2 (*positive cofactor 2*) (Malik et Roeder, 2000) et CRSP (*cofactor required for Sp1 activation*) (Taatjes *et al.*, 2002).

Un des complexes les mieux caractérisés est le complexe TRAP/SMCC. Le complexe SMCC a d'abord été isolé à partir de cellules humaines en ajoutant une étiquette permettant la purification, sur les homologues humains de Srb7, Srb10 ou Srb11. Le complexe purifié était capable de supporter l'activation par GAL4 *in vitro* (Gu *et al.*, 1999). Le complexe SMCC contient plusieurs protéines Srb et Med. Le complexe TRAP a été isolé de manière indépendante en fonction de sa capacité à supporter l'activation par le récepteur de l'hormone thyroïdienne, lorsque celui-ci est lié à son ligand (Fondell *et al.*, 1999). Il s'est avéré par la suite que TRAP et SMCC étaient des complexes identiques (Ito *et al.*, 1999). De façon similaire, le complexe ARC a été identifié comme un cofacteur requis pour l'activation par Sp1 et SREBP-1a (Naar *et al.*, 1999), puis le complexe DRIP comme un complexe liant le récepteur de la vitamine D (Rachez *et al.*, 1998). Il est maintenant reconnu que ces deux complexes sont identiques et interagissent avec un certain nombre d'activateurs et de récepteurs nucléaires hormonaux (Naar *et al.*, 1999; Rachez *et al.*, 1999). Plusieurs sous-unités de ARC/DRIP se retrouvent aussi dans le complexe CRSP, d'abord identifié comme coactivateur pour le facteur de transcription Sp1 (Ryu *et al.*, 1999). Finalement, le complexe NAT, qui contient les homologues de Srb10 et Srb11, est capable de réprimer l'activation de la transcription *in vitro* (Sun *et al.*, 1998).

Une étude récente a permis d'identifier une série de protéines consensus, présentes dans tous les complexes médiateurs de mammifères (Sato *et al.*, 2004). En ajoutant une étiquette sur 6 sous-unités, préalablement identifiées comme faisant partie des différents complexes médiateurs, les auteurs ont purifié des complexes dont la composition a ensuite été étudiée à l'aide de la technique MudPIT (*multidimensional protein identification technology*) (Washburn *et al.*, 2001), une technique de spectrométrie de masse très sensible. L'étude a révélé que 37 protéines étaient communes à tous les complexes déjà connus. L'hétérogénéité recensée précédemment dans la littérature entre les différents complexes était donc due surtout à une sensibilité moindre

des techniques d'identification des protéines utilisées. Tel que décrit précédemment, il semble tout de même exister des formes variées de médiateurs qui diffèrent par la présence ou l'absence du module CDK8/Cyclin C/TRAP240/TRAP230. Ce module correspondrait au module Srb9, 10, 11 et 12 chez la levure (Taatjes *et al.*, 2002; Samuelsen *et al.*, 2003; Malik et Roeder, 2000). Des évidences montrent que les médiateurs qui contiennent le module CDK8 sont moins efficaces dans leur fonction de coactivateur (Taatjes *et al.*, 2002). De plus, ces médiateurs pourraient dans certains cas réprimer la transcription (Kuchin *et al.*, 1995; Hengartner *et al.*, 1998; Song *et al.*, 1996).

La purification initiale du médiateur en complexe avec l'ARN pol II et certains FGT a conduit à la proposition du concept de l'holoenzyme tel que décrit plus haut (Koleske et Young, 1994; Thompson *et al.*, 1993; Kim *et al.*, 1994). De nombreuses études plus récentes semblent toutefois montrer que le médiateur n'existerait pas sous forme d'holoenzyme avec la polymérase (Featherstone, 2002). Par exemple, il a été démontré par des expériences d'immunoprécipitation de la chromatine (ChIP) que le médiateur était recruté au niveau des séquences activatrices, appelées UAS (*upstream activating sequence*), et non au niveau du promoteur basal des gènes *GAL in vivo*, où la polymérase et les FGT sont recrutés (Kuras *et al.*, 2003). De plus, le recrutement du médiateur est indépendant du recrutement de ces derniers et de la présence d'un promoteur basal fonctionnel (Kuras *et al.*, 2003). Une autre étude a montré qu'au promoteur *GALI*, le médiateur serait recruté plus rapidement que la pol II et les FGT suite à l'induction de la transcription du gène (Bryant et Ptashne, 2003). De plus, il a été démontré qu'il y avait une différence temporelle et spatiale dans le recrutement du médiateur et de la polymérase au promoteur *HO* (Bhoite *et al.*, 2001; Cosma, 2002). Chez la drosophile, le médiateur co-localise avec le facteur de transcription *heat shock* à l'élément de réponse de ce dernier, suivant l'induction à la chaleur, indépendamment de la présence d'un promoteur basal (Park *et al.*, 2001). Dans le même ordre d'idée, le récepteur de l'hormone thyroïdienne et le récepteur de la vitamine D sont isolés seuls en conditions non inductrices et en complexe avec le médiateur en présence de leurs ligands respectifs (Fondell *et al.*, 1996; Malik et Roeder, 2000).

Toutes ces études suggèrent que, bien que le médiateur puisse interagir avec le complexe de l'ARN pol II et des FGT pour permettre l'activation, cette interaction est transitoire. Le médiateur ne se retrouverait donc pas sous la forme d'holoenzyme avec l'ARN pol II, mais serait plutôt recruté indépendamment au niveau des régions régulatrices par les activateurs.

1.1.5 Les séquences d'ADN régulatrices

Les séquences d'ADN régulatrices sont des éléments qui jouent un rôle de premier ordre dans le contrôle de l'expression génétique. Les séquences liées par des activateurs transcriptionnels sont appelées *enhancers* (revu par Blackwood et Kadonaga, 1998; Mitchell et Tjian, 1989; Struhl, 1995) ou UAS (*upstream activating sequence*). Les *enhancers* regroupent souvent plusieurs sites de liaison à des activateurs et influencent la transcription indépendamment de leur orientation et de la distance qui les sépare du promoteur basal. Les UAS sont des *enhancers* retrouvés entre autres chez la levure. Ils se retrouvent relativement près du site d'initiation de la transcription et ne fonctionnent qu'en amont des gènes.

Les éléments liés par des répresseurs transcriptionnels, qui reconnaissent une séquence spécifique, sont appelés URS (*upstream repressing sequence*). Ces répresseurs liant l'ADN peuvent inhiber la transcription en interférant avec la fonction d'un activateur, en recrutant des complexes qui modifient la structure de la chromatine ou en bloquant le recrutement de la machinerie transcriptionnelle au promoteur (Hanna-Rose et Hansen, 1996; Gaston et Jayaraman, 2003). Les *silencers* sont définis comme des éléments qui répriment la transcription indépendamment de leur orientation et de leur position (Ogbourne et Antalis, 1998). Les LCR (*locus control region*) sont des séquences qui modulent la transcription en influençant la structure de la chromatine sur une grande région d'ADN. Ce faisant, ils maintiennent l'accessibilité de ces régions aux facteurs de transcription (Grosveld, 1999). Les *insulators* sont des éléments qui isolent l'effet des *enhancers* et des *silencers* sur des régions particulières, empêchant ces derniers de moduler la transcription d'autres gènes de manière indésirable (Fourel *et al.*, 2004).

1.2 Les étapes de la transcription

L'ARN pol II seule ne peut reconnaître un promoteur et initier la transcription de manière spécifique au site d'initiation, appelé le site +1. Pour ce faire, la polymérase a besoin de l'aide des facteurs généraux de transcription (FGT) TFIID, TFIIB, TFIIF, TFIIE et TFIIH (tableau I). Suite à la reconnaissance du promoteur, la double hélice d'ADN est ouverte dans la région -9 à +2 (par rapport au site +1) pour former la bulle de transcription, ce qui permet à la polymérase d'initier la synthèse du transcrit. L'ARN pol II passe ensuite par un cycle d'initiations avortées au cours duquel de courts transcrits entre 2 et 10 nucléotides sont synthétisés puis relâchés. La polymérase dégage ensuite le promoteur et entre dans l'étape d'élongation du transcrit. Finalement, le transcrit est polyadénylé et la transcription se termine par le recyclage de la polymérase pour permettre à celle-ci d'initier un nouveau cycle. Ces différentes étapes du processus transcriptionnel, ainsi que les rôles des FGT et d'autres facteurs dans celles-ci seront discutés dans la section qui suit.

Tableau I. Les facteurs généraux de transcription humains

Facteur	Sous-unité	Poids moléculaire (kDa)	Caractéristiques
TFIID	TBP	38	Lie la boîte TATA, induit une courbure de 80° dans l'ADN, commun aux trois ARN pol
	TAF1 à 15	250 à 15	Lient l'Inr et le DPE et impliqués dans la coactivation et la corépression
TFIIB		35	Lie TBP et stabilise son interaction avec la boîte TATA, lie le BRE, est impliqué dans le choix du site d'initiation
TFIIA	α	37	Lie TBP et augmente son affinité pour la boîte TATA, coactivation, anti-
	β	19	

	γ	13	répression, α et β résultent d'un clivage protéolytique d'un précurseur
TFIIF	RAP74	58	Recrutement de l'ARN pol II au promoteur, enroulement du promoteur,
	RAP30	28	stimule l'élongation, participe au dégagement du promoteur
TFIIE	α	56	Recrutement de TFIIF, stimule activités kinases et hélicases de TFIIF, impliqué dans l'ouverture du promoteur
	β	34	indépendamment de TFIIF, impliqué dans le dégagement du promoteur via la stimulation de l'activité kinase de TFIIF
TFIIF	XPB	89	Hélicase 3'-5', ouverture du promoteur, dégagement du promoteur, essentielle pour la transcription et le NER
	XPB	80	Hélicase 5'-3', essentielle pour le NER, stimule la transcription
	p62	62	NER, est la cible d'activateurs
	p52	52	NER
	p44	44	NER, stimule activité de XPB
	cycline H	37	Fait partie de CAK, phosphorylation du CTD
	p34	34	NER
	Mat1	32	Fait partie de CAK, interagit avec cyclin H et Cdk7
	Cdk7	40	Fait partie de CAK, phosphorylation du CTD
	TFB5	8	Recrutement de TFIIF au promoteur

1.2.1 Le promoteur basal

Les promoteurs basaux peuvent être constitués des différents éléments suivants : la boîte TATA, l'initiateur (Inr), le DPE (*downstream promoter element*), le BRE (*TFIIB recognition element*) et le MTE (*motif ten element*) (revu par Smale et Kadonaga, 2003) (Lim *et al.*, 2004).

La boîte TATA est un élément ayant la séquence consensus TATAA/TAAG/A (Patikoglou *et al.*, 1999), localisé entre -25 et -30 par rapport au site +1 chez la plupart des eucaryotes. Chez la levure *S. cerevisiae* cependant, la boîte TATA est localisée entre -40 et -120 par rapport au site d'initiation (Struhl, 1989). La boîte TATA est reconnue et liée par la protéine TBP (*TATA binding protein*). Bien que la séquence consensus TATAAG soit la séquence reconnue avec le plus d'affinité par TBP (Wong et Bateman, 1994), celui-ci est tout de même capable de lier une grande diversité de séquences riches en A/T (Hahn *et al.*, 1989; Patikoglou *et al.*, 1999). Des études récentes montrent d'ailleurs qu'un nombre restreint de gènes contiennent des boîtes TATA conformes à la séquence consensus. Chez la drosophile, il a été démontré que sur 1941 promoteurs étudiés, 33% contenaient une boîte TATA (Ohler *et al.*, 2002). Chez l'humain, les boîtes TATA ont été trouvées sur 32% des 1031 promoteurs testés (Suzuki *et al.*, 2001). Plus récemment, il a été démontré que seulement 20 % des gènes chez la levure contenaient des éléments TATA, fidèles à la séquence consensus (Basehoar *et al.*, 2004). Étonnamment, il s'est avéré que ces gènes étaient en grande majorité des gènes de réponse au stress, hautement régulés, qui utilisent préférentiellement SAGA comme coactivateur. Au contraire, l'étude a révélé que les gènes ne contenant pas de boîte TATA utilisent plutôt TFIID pour la coactivation (Basehoar *et al.*, 2004). TFIID est un complexe formé de TBP, associé à un groupe de protéines appelées TAF (*TBP associated factors*), dont le rôle sera décrit dans une section suivante (1.2.2.1).

L'initiateur (Inr) est un élément du promoteur basal qui chevauche le site +1, ayant la séquence consensus suivante : Py Py A(+1) N T/A Py Py (Javahery *et al.*, 1994; Lo et Smale, 1996). L'Inr favorise l'initiation précise au site +1 et contribue à la force des promoteurs contenant une boîte TATA (Hahn *et al.*, 1985; Nagawa et Fink, 1985; Corden *et al.*, 1980; Tokunaga *et al.*, 1984). Ces deux éléments fonctionnent en synergie lorsqu'ils sont séparés de 25 à 30 paires de bases (pb), mais indépendamment lorsque la distance les séparant excède 30 pb (O'Shea-Greenfield et Smale, 1992). Lorsqu'ils sont distancés de 15 à 20 pb la synergie est conservée mais le choix du site d'initiation est dicté par la boîte TATA, c'est-à-dire qu'il se retrouve à 25 pb en aval de celle-ci (O'Shea-

Greenfield et Smale, 1992). Au niveau de certains promoteurs qui ne contiennent pas de boîte TATA (*TATA-less promoters*), mais contiennent un Inr, la liaison de TBP à l'ADN dans la région -25 à -30 est quand même essentielle (Martinez *et al.*, 1995). Dans ces cas, TBP peut lier de façon plus ou moins stable une séquence qui diverge considérablement de la séquence consensus. Pour d'autres promoteurs sans boîte TATA, la liaison de TBP à l'ADN n'est pas nécessaire pour l'initiation de la transcription (Martinez *et al.*, 1995).

L'Inr est reconnu, entre autres, par certains TAF de TFIID. TAF_{II}250 (TAF1) et TAF_{II}150 (TAF2) sont capables de lier l'Inr (Chalkley et Verrijzer, 1999) et dans un essai reconstitué de transcription supportent l'activité de ce dernier en présence de TBP (Verrijzer *et al.*, 1995). Bien qu'il soit considéré que l'ARN pol II ne puisse reconnaître un promoteur et initier la transcription de façon spécifique sans l'aide des FGT, Carcamo et collaborateurs ont montrés que l'enzyme seule possède une certaine capacité de reconnaissance de l'Inr (Carcamo *et al.*, 1991). De plus, en absence de TAF, l'ARN pol II, TBP, TFIIB et TFIIF peuvent former un complexe stable sur un promoteur sans boîte TATA et contenant un Inr (Weis et Reinberg, 1997).

Le DPE (*downstream promoter element*) est un élément retrouvé dans plusieurs promoteurs ne contenant pas de boîte TATA, qui est lié par TFIID (Kutach et Kadonaga, 2000). Cet élément est localisé de +28 à +32 et sa séquence consensus est la suivante : A/G G A/T C/T G/A/C (Smale et Kadonaga, 2003). Les promoteurs contenant un DPE contiennent aussi un Inr. Sur ces promoteurs, des mutations au niveau du DPE ou de l'Inr entraînent la perte de liaison de TFIID et conséquemment la perte d'activité transcriptionnelle (Burke et Kadonaga, 1996). TFIID se lie donc de manière coopérative au DPE et à l'Inr. La distance entre ces deux éléments est très importante puisqu'une insertion ou une délétion d'une seule pb conduit à une diminution considérable de la liaison de TFIID et de l'activité du promoteur (Kutach et Kadonaga, 2000).

L'élément BRE (*TFIIB recognition element*) est localisé immédiatement en amont de la boîte TATA et est reconnu, comme son nom l'indique, par TFIIB (Nikolov *et al.*, 1995; Lagrange *et al.*, 1998). Ce motif de séquence consensus G/C G/C G/A C G C C

(Lagrange *et al.*, 1998) n'a pas été retrouvé chez la levure et chez les plantes à ce jour. Une première étude a montré que la liaison de TFIIB au BRE stimule la transcription dans un essai *in vitro* reconstitué avec les FGT et l'ARN pol II (Lagrange *et al.*, 1998). Cependant, une autre étude a montré que le BRE réprime plutôt la transcription basale dans des essais *in vitro* avec extraits nucléaires et *in vivo*, dans des essais de transfection (Evans *et al.*, 2001). Cette répression peut être contrée par des activateurs, ce qui conduit à une stimulation plus grande de la transcription en présence du BRE qu'en absence de celui-ci.

Récemment, un nouvel élément du promoteur a été mis en évidence: le MTE (*motif ten element*) (Lim *et al.*, 2004). Le MTE a été découvert lors d'une analyse informatique des séquences promotrices chez la drosophile (Ohler *et al.*, 2002). Il fait partie de 4 motifs surreprésentés, localisés près du site d'initiation de la transcription. Les 3 autres motifs identifiés sont la boîte TATA, l'Inr et le DPE. Le MTE, localisé de +18 à +27, est important pour la transcription basale et fonctionne avec l'Inr. En revanche, il fonctionne indépendamment de la boîte TATA et du DPE et la perte d'activité transcriptionnelle, due à des mutations au niveau de ces derniers, peut être compensée par son ajout. En présence des éléments TATA et DPE, il stimule la transcription de manière synergique.

1.2.2 Les facteurs généraux de transcription et l'initiation transcriptionnelle

1.2.2.1 TFIID

TFIID est un complexe d'environ 750 kDa composé de TBP (*TATA binding protein*) et d'environ 14 TAF (*TBP-associated factors*) (revu par Hampsey, 1998; Davidson, 2003). TBP (38 kDa) est une protéine, qui en terme de structure, est formée de deux modules : un module C-terminal et un module N-terminal. Le module C-terminal, composé de deux répétitions directes, est très conservé entre les espèces (80% d'identité entre la levure et l'humain) et sert à la reconnaissance de la boîte TATA. Le module C-terminal de TBP a la forme d'une selle, qui en se liant à l'élément TATA sur l'ADN

induit une courbure de 80° dans celui-ci (Kim *et al.*, 1993a; Kim *et al.*, 1993b) (figure 1). C'est aussi au niveau de ce module que se fait l'interaction avec les facteurs TFIIA, TFIIB et les TAF.

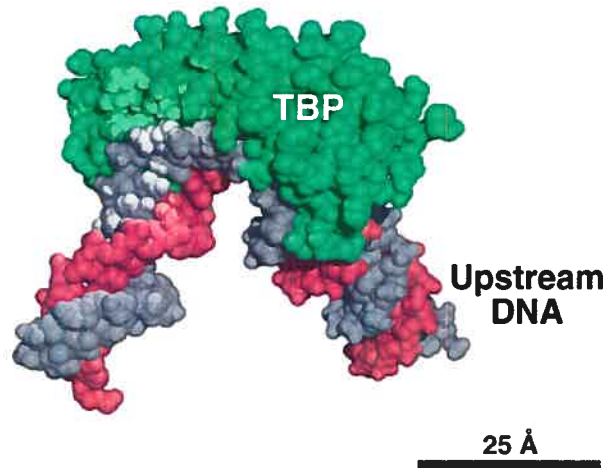


Figure 1 : Structure de TBP sur la boîte TATA. La liaison de TBP (vert) à la boîte TATA induit une courbure de 80° dans l'ADN (gris et rose). Figure tirée de Hahn, 2004.

Le module N-terminal est variable en terme de séquence et de longueur, mais est tout de même bien conservé entre les vertébrés et est caractérisé par la présence d'une région contenant une série de glutamines. Chez l'humain, il existe une variation allélique dans le nombre de répétitions et les individus normaux ont entre 28 et 42 glutamines. Un nombre de répétitions entre 47 et 56 conduit à des maladies neurodégénératives comme l'ataxie cérébrale (Fujigasaki *et al.*, 2001). Le module N-terminal, sans pour autant lier directement la boîte TATA, influence l'association de TBP à l'ADN. Il a été observé que le TBP complet (N et C-terminal) se lie à la boîte TATA en deux étapes. Premièrement, un complexe instable dans lequel l'ADN n'est pas courbé est formé, puis, la courbure est induite. Au contraire, le C-terminal de TBP seul lie l'ADN en une étape, formant simultanément le complexe courbé (Zhao et Herr, 2002). Les auteurs de l'étude ont suggéré que le N-terminal inhibe la formation de la courbure de 80° dans l'ADN, mais

qu'un changement de conformation de TBP, possiblement induit par la liaison de TFIIB, peut bloquer cette inhibition.

TBP a longtemps été considéré comme un facteur universel, c'est-à-dire requis pour la transcription de tous les gènes des trois ARN polymérase eucaryotes (Hampsey, 1998). En effet, TBP s'associe à des TAF_I pour former le complexe SL1 et à des TAF_{III} pour former le complexe TFIIB, impliqués dans la transcription des gènes par l'ARN pol I et III respectivement (Davidson, 2003). L'universalité de TBP a été remise en question par la découverte de facteurs similaires à TBP, les TRF (*TBP related factors*). Le premier TRF, TRF1, a été identifié chez la drosophile et demeure, à ce jour unique à cet organisme (Crowley *et al.*, 1993). TRF1 est homologue au module C-terminal de TBP et est exprimé chez l'embryon, ainsi qu'au niveau du système nerveux chez l'adulte et au niveau des cellules germinales mâles. Des études biochimiques ont montré que l'immunodéplétion de TRF1 dans des extraits nucléaires d'embryons de drosophile conduit à l'inhibition de la transcription par la pol III. En revanche, l'immunodéplétion de TBP n'a pas d'effet sur l'expression des gènes transcrits par la pol III, indiquant que TRF1 remplace TBP dans cette fonction chez certains tissus de la drosophile (Takada *et al.*, 2000).

TRF1 est aussi impliqué dans la transcription d'au moins un gène de classe II (Holmes et Tjian, 2000). Chez la drosophile, le gène *Tudor* est transcrit à partir de deux promoteurs. TRF1 participe à la transcription au niveau du promoteur le plus en amont, alors que TBP se lie au promoteur en aval. Bien que TRF1 soit capable de lier une boîte TATA consensus (Hansen *et al.*, 1997), au promoteur *Tudor*, il lie un élément riche en TC. TRF1 a donc la capacité de se lier à des promoteurs sans boîte TATA.

TRF2, qui présente une homologie de 40% avec le module C-terminal de TBP, est retrouvé chez plusieurs espèces eucaryotes (Dantonel *et al.*, 1999). La séquence consensus reconnue par TRF2 n'a pas encore été déterminée, mais on sait qu'il ne lie pas les boîtes TATA. Cependant, TRF2 comme TRF1, est capable de s'associer avec TFIIA et TFIIB (Rabenstein *et al.*, 1999). Tout comme TBP, TRF2 est requis pour la

transcription de gènes spécifiques durant l'embryogénèse (Veenstra *et al.*, 2000). Finalement, TRF3 est homologue à 96% avec TBP au niveau du module C-terminal (Persengiev *et al.*, 2003). TRF3 est présent chez la plupart des vertébrés, mais absent chez la drosophile et chez le ver *C. elegans*. TRF3 est exprimé dans tous les tissus humains et murins étudiés. Bien que très homologue à TBP, TRF3 se retrouve dans un complexe autre que TFIID, de poids moléculaire beaucoup plus faible que celui-ci. (Persengiev *et al.*, 2003).

Les TAF_{II} sont des protéines qui s'associent à TBP pour former le complexe TFIID. À ce jour, 13 TAF_{II} conservés au cours de l'évolution ont été identifiés chez différentes espèces étudiées (revu par Sanders et Weil, 2000). Une nomenclature uniformisée a été publiée récemment qui nomme les TAF_{II} de TAF1 à TAF15 par ordre décroissant de poids moléculaire (Tora, 2002). Chez la levure, tous les TAF à l'exception de TAF14 sont essentiels à la viabilité cellulaire (Matangkasombut *et al.*, 2004). Plusieurs TAF contiennent un motif *histone fold* retrouvé sur les histones, qui leur permet d'interagir entre eux et de former un octamère semblable à l'octamère d'histone (Selleck *et al.*, 2001). La structure générale de TFIID, telle que révélée par microscopie électronique, a la forme d'une pince à plusieurs lobes (Leurent *et al.*, 2002; Leurent *et al.*, 2004). TBP est localisé dans le creux de la pince où il peut interagir avec l'ADN. Un certain nombre de TAF se retrouvent dans plusieurs complexes autres que TFIID. Par exemple, un groupe de TAF (TAF5, 6, 9, 10 et 12) fait partie du complexe de levure SAGA, un complexe ayant une activité HAT (Grant *et al.*, 1998). D'autres TAF se retrouvent dans les complexes humains PCAF et STAGA (Ogryzko *et al.*, 1998). Finalement, il existe un complexe appelé TFTC (*TBP-free TAF-containing complex*), qui contient des TAF mais pas TBP (Wieczorek *et al.*, 1998). Ce complexe a une activité HAT et est capable de remplacer TFIID dans un essai de transcription *in vitro*, suggérant qu'il contient une protéine pouvant remplir la fonction de TBP.

Les TAF sont impliqués dans la reconnaissance de l'initiateur et du DPE, tel que décrit à la section précédente (1.2.1), et dans la coactivation. Plusieurs études ont montré l'existence d'interaction entre les TAF et des activateurs *in vitro* (Sauer et Tjian, 1997).

Des études ont aussi montré que certains TAF de TFIID étaient requis pour l'activation (Sauer *et al.*, 1995b; Sauer *et al.*, 1995a). Cependant, dans d'autres systèmes *in vitro*, les TAF ne sont pas requis et l'activation dépend plutôt du médiateur (Fondell *et al.*, 1999; Wu *et al.*, 1998; Oelgeschlager *et al.*, 1998). Finalement, dans certains cas, les TAF et le médiateur peuvent coopérer pour obtenir une synergie des niveaux d'activation (Naar *et al.*, 1999).

1.2.2.2 TFIIB

TFIIB est un facteur qui interagit avec TBP, TFIIF et l'ARN pol II, facilitant le recrutement de celle-ci au promoteur (Buratowski *et al.*, 1989). TFIIB est constitué de deux domaines fonctionnels. Le domaine C-terminal interagit avec TBP et contient un motif hélice-tour-hélice qui lie les séquences 5' et 3' entourant la boîte TATA (Nikolov *et al.*, 1995; Reese, 2003). Certains promoteurs contiennent un élément BRE, localisé en 5' de la boîte TATA. L'interaction de TFIIB avec le BRE au niveau de ces promoteurs stabilise le complexe TBP-TFIIB sur l'ADN (Qureshi et Jackson, 1998; Wolner et Gralla, 2001). Le domaine N-terminal de TFIIB contient un motif ruban de zinc (*zinc ribbon*) qui lie l'ARN pol II et TFIIF et permet leur entrée dans le complexe pré-transcriptionnel (Ha *et al.*, 1993; Pardee *et al.*, 1998). TFIIB est aussi impliqué dans le choix du site d'initiation de la transcription. Plusieurs mutations de TFIIB ont été identifiées et caractérisées, qui ont pour effet de modifier le site où l'ARN pol II initie la transcription (Berroteran *et al.*, 1994; Pinto *et al.*, 1994; Pinto *et al.*, 1992).

Les modules C et N-terminaux de TFIIB, qui sont séparés par une région flexible hautement chargée appelée CCD (*charge clustered domain*), peuvent interagir entre eux pour former un TFIIB de conformation dite fermée. Le CCD module l'interaction entre les domaines C et N-terminaux de la protéine (Hawkes *et al.*, 2000; Fairley *et al.*, 2002; Zheng *et al.*, 2004). Cette interaction semble réguler les différentes activités de TFIIB en transcription (Elsby et Roberts, 2004). En effet, il a été démontré que l'activateur VP16 qui contacte TFIIB est capable de détruire l'interaction entre les deux modules, induisant un changement de conformation au niveau de TFIIB (Roberts et Green, 1994). Des

mutations dans le CCD, qui induisent une augmentation de la liaison entre les deux modules, nuisent au recrutement de TFIIB par des activateurs et à l'activation *in vitro* et *in vivo* (Hawkes *et al.*, 2000; Glossop *et al.*, 2004). Un changement de conformation se produit aussi au niveau de TFIIB, suite à l'association au complexe TBP-ADN (Fairley *et al.*, 2002). La conformation adoptée varie selon le fait que le promoteur possède un élément BRE ou non. Finalement, il a aussi été montré que des mutants qui modifient le choix du site d'initiation s'assemblent au promoteur avec une conformation différente du type sauvage (Fairley *et al.*, 2002). Ces études montrent donc que la conformation de TFIIB joue un rôle important dans l'activation transcriptionnelle, la formation du complexe d'initiation au promoteur et la détermination du site d'initiation de la transcription.

1.2.2.3 TFIIF

TFIIF est composé de deux sous-unités, RAP 30 (26 kDa) et RAP 74 (58 kDa). Ces deux protéines ont d'abord été identifiées grâce à leur capacité d'interaction avec l'ARN pol II d'où le nom de RAP pour *RNA polymerase II associated protein* (Sopta *et al.*, 1985). Il a ensuite été démontré que les sous-unités RAP30 et RAP74 interagissaient entre elles et étaient requises pour l'initiation de la transcription (Flores *et al.*, 1989; Burton *et al.*, 1988). TFIIF présente des homologies de séquence et de fonction avec le facteur σ , nécessaire à la reconnaissance du promoteur par l'ARN pol bactérienne. En effet, les deux sous-unités ont des homologies de séquences avec certaines régions de σ (Garrett *et al.*, 1992; McCracken et Greenblatt, 1991; Sopta *et al.*, 1989; Sun et Hampsey, 1995; Yonaha *et al.*, 1993). De plus, tout comme σ avec l'ARN pol bactérienne, TFIIF a une forte affinité pour l'ARN pol II (Sopta *et al.*, 1985), supprime la liaison non-spécifique de celle-ci à l'ADN, via la sous-unité RAP30 (Killeen et Greenblatt, 1992), et favorise l'arrivée de la polymérase au promoteur (Killeen *et al.*, 1992; Conaway *et al.*, 1991).

TFIIF joue un rôle important dans l'organisation du complexe pré-transcriptionnel. Il a été démontré, par des expériences de photo-pontage protéine-ADN,

que la sous-unité RAP74 induisait un changement important dans la topologie du promoteur au niveau d'un complexe d'initiation de la transcription comprenant TBP, TFIIB, TFIIF, TFIIE et l'ARN pol II (Robert *et al.*, 1998) (revu par Coulombe et Burton, 1999). TFIIF entre dans le complexe sous la forme d'un hétérotétramère de forme $\alpha_2\beta_2$, dans lequel deux molécules de RAP74 interagissent entre elles via une région appelée HIR1 (*homomeric interaction region*). L'entrée de RAP74 dans le complexe d'initiation entraîne la formation d'une deuxième courbure dans l'ADN, près du site +1 (la première courbure étant induite par TBP au niveau de la boîte TATA) permettant un enroulement serré du promoteur autour de l'ARN pol II (Robert *et al.*, 1998). Les mutants de RAP74, qui ne sont pas capables de former l'hétérotétramère et de supporter l'enroulement, présentent des défauts d'initiation (Ren *et al.*, 1999). L'hypothèse d'un enroulement du promoteur autour du complexe d'initiation de la pol II est supportée par des images de microscopie électronique de l'ARN pol II avec les FGT (Forget *et al.*, 1997) et de la l'ARN pol bactérienne (Rivetti *et al.*, 1999). TFIIF est aussi impliqué dans l'étape d'élongation de l'ARNm, un rôle qui sera discuté dans une section suivante (1.2.3.1.1.1).

1.2.2.4 TFIIE

TFIIE, tout comme TFIIF est un hétérotétramère de type $\alpha_2\beta_2$, formé des sous-unités TFIIE α (56 kDa) et TFIIE β (34 kDa) (Ohkuma *et al.*, 1990). TFIIE interagit avec plusieurs composantes du complexe d'initiation. TFIIE α lie TBP, TFIIE β , RAP74, TFIIH et l'ARN pol II (Maxon *et al.*, 1994; Ohkuma *et al.*, 1995). TFIIE β homodimérise et lie RAP30, TFIIE α , TFIIB et l'ARN pol II (Okamoto *et al.*, 1998; Maxon *et al.*, 1994). Les rôles de TFIIE en transcription sont en grande partie reliés à TFIIH. TFIIE est essentiel au recrutement de TFIIH dans le complexe transcriptionnel (Flores *et al.*, 1992) et stimule ses activités kinase et hélicases (Ohkuma et Roeder, 1994). L'activité kinase de TFIIH permet la phosphorylation du CTD de l'ARN pol II, étape nécessaire au dégagement du promoteur, alors que les activités hélicases, qui requièrent l'hydrolyse de l'ATP, permettent l'ouverture de la bulle de transcription de -9 à +2.

TFIIE est donc impliqué dans l'étape du dégagement du promoteur, qui est l'étape de transition entre l'initiation de la transcription et l'élongation de l'ARNm, via la stimulation de la phosphorylation du CTD par TFIIH (Goodrich et Tjian, 1994; Yamamoto *et al.*, 2001; Kumar *et al.*, 1998; Watanabe *et al.*, 2003). En effet, des mutants de TFIIE β qui affectent la transition entre l'initiation et l'élongation, sont incapables de stimuler la phosphorylation du CTD (Watanabe *et al.*, 2003). TFIIE possède aussi des activités indépendantes de TFIIH, notamment dans l'ouverture du promoteur. TFIIE, TFIIH et l'hydrolyse de l'ATP ne sont pas requis pour la transcription à partir d'une matrice d'ADN sous-enroulée (*negative supercoiling*) (Goodrich et Tjian, 1994; Parvin *et al.*, 1994). Ceci est dû au fait qu'une matrice sous-enroulée présente moins que 10,1 paires de base par tour. L'énergie requise pour ouvrir le promoteur sur ce type de matrice est donc beaucoup moins grande. Il a été démontré que sur une matrice sous-enroulée, TFIIE stimule la transcription en absence de TFIIH et que le degré de stimulation corrèle avec la stabilité de l'hélice d'ADN (Timmers, 1994; Holstege *et al.*, 1995). Par exemple, dans des conditions où le promoteur est moins susceptible d'être ouvert (concentration ionique élevée, température faible), la transcription est plus dépendante de TFIIE (Holstege *et al.*, 1995). D'autres études ont montré que l'utilisation de matrices pré-ouvertes dans la région de la bulle transcriptionnelle permet de contourner la dépendance en TFIIE et TFIIH pour la transcription (Pan et Greenblatt, 1994; Holstege *et al.*, 1995; Holstege *et al.*, 1996). Chez la levure, des mutants de TFIIE α sensibles au refroidissement (*cold-sensitive*) ont été caractérisés (Kuldell et Buratowski, 1997; Sakurai *et al.*, 1997). Ces mutants, qui présentent une délétion de la partie C-terminale de TFIIE α , sont incapables de croître et de supporter la transcription d'un certain nombre de gènes à 16⁰ C. Les auteurs de l'étude ont suggéré que cette partie de la protéine jouerait un rôle important en transcription à faible température, possiblement dans l'ouverture du promoteur, dans des conditions qui sont stabilisantes pour l'hélice d'ADN (Kuldell et Buratowski, 1997). Finalement, TFIIE β est capable de lier l'ADN simple-brin (Okamoto *et al.*, 1998) et double-brin (Yokomori *et al.*, 1998; Okuda *et al.*, 2000), renforçant l'idée que TFIIE soit impliqué dans l'ouverture de la bulle de transcription, autrement que par la stimulation des activités hélicases de TFIIH.

1.2.2.5 TFIIH

TFIIH est un facteur composé de neuf sous-unités, incluant deux sous-complexes : le noyau (*core* TFIIH) et le complexe CAK (*cyclin-activating kinase complex*) (revu par Zurita et Merino, 2003). Le noyau de TFIIH comprend les sous-unités XPB (*xeroderma pigmentosum B*), p62, p52, p44 et p34. Le complexe CAK inclut Cdk7 (*cyclin-dependent kinase 7*), cycline H et MAT1 (ménage à trois). La protéine XPD (*xeroderma pigmentosum D*) lie le noyau à CAK. TFIIH est impliqué dans trois processus cellulaires : la réparation de l'ADN, le contrôle du cycle cellulaire et la transcription. Au niveau de la réparation, TFIIH participe au NER (*nucleotide excision repair*) qui permet d'éliminer certains types de dommages à l'ADN, tels que ceux induits par l'exposition aux rayons ultraviolets. Il participe aussi au TCR (*transcription coupled repair*), un processus qui couple la réparation de l'ADN à la transcription (Le Page *et al.*, 2000). Le complexe CAK est important pour le contrôle du cycle cellulaire durant la transition de la phase G2 à M (Larochelle *et al.*, 1998).

En transcription, TFIIH est impliqué à plusieurs niveaux. Les sous-unités XPB et XPD possèdent des activités hélicases 3'-5' et 5'-3', respectivement, qui requièrent l'hydrolyse de l'ATP (Drapkin *et al.*, 1994; Schaeffer *et al.*, 1993; Schaeffer *et al.*, 1994). Ces activités hélicases servent à ouvrir le promoteur de -9 à +2 pour former la bulle transcriptionnelle et permettre à la polymérase d'initier la synthèse du transcrit au site +1. XPB et XPD sont des hélicases qui fonctionnent sur un substrat d'ADN simple-brin. Il a été proposé que l'enroulement serré de l'ADN autour de l'ARN pol II, induit une torsion dans l'ADN près du site +1, qui pourrait fournir une portion d'ADN simple-brin suffisante pour être reconnue par les hélicases de TFIIH (Robert *et al.*, 1998; Coulombe et Burton, 1999). À l'aide de la technique de photo-pontage protéine-ADN, XPB et XPD ont été localisés dans la région -9 à +2 du promoteur, là où elles participent à la formation de la bulle transcriptionnelle (Douziech *et al.*, 2000). Un modèle alternatif a aussi été proposé par un deuxième groupe utilisant une méthode de photo-pontage incluant des étapes de lavages au sarkosyl (Kim *et al.*, 2000). Dans ces conditions, les hélicases ont pu

être photo-pontées uniquement en aval du site +1 et non dans la région de la bulle transcriptionnelle. Ainsi, les auteurs ont proposé que TFIIH participerait à l'ouverture du promoteur en induisant une rotation de l'ADN en aval par rapport à un site en amont fixé par des interactions protéines-ADN, générant une torsion suffisante pour former la bulle de transcription.

Alors que l'hélicase XPB est essentielle aux processus de transcription et de réparation, XPD est quant à elle requise seulement pour la réparation, bien qu'elle stimule l'ouverture du promoteur et la transcription (Tirode *et al.*, 1999). Des mutations au niveau de XPB et XPD sont responsables de trois types de syndromes, qui causent une photosensibilité extrême chez l'humain : le *xeroderma pigmentosum* (XP), le syndrome de Cockayne (CS) et la trichothiodystrophie (TTD). La photosensibilité causée par le syndrome XP est due à des défauts dans le processus NER, alors que celle des syndromes CS et TTD provient de défauts dans le TCR et la transcription (Giglia-Mari *et al.*, 2004).

Suite à l'ouverture de la bulle de transcription de -9 à +2 par les hélicases de TFIIH, l'ARN pol II peut catalyser la formation du premier lien phosphodiester entre les ribonucléotides (NTP) +1 et +2. La transcription progresse ensuite et la bulle s'étend conséquemment en aval, d'une paire de base de plus que le dernier NTP ajouté (Holstege *et al.*, 1997). Suite à l'ajout du quatrième NTP à l'ARNm, un changement de conformation se produit au niveau du complexe transcriptionnel, qui conduit à la stabilisation du complexe ouvert. Suite à cette étape, TFIIH n'est plus requis pour le maintien de la bulle de transcription. Entre +1 et +10, l'ARN pol II passe par un cycle d'initiations avortées (*abortive initiation*) au cours duquel elle relâche de courts transcrits de 2 à 10 nucléotides. Une autre transition importante se produit à la position +11 : la bulle de transcription se referme de -9 à +2 et l'ARN pol II peut passer en élongation (Holstege *et al.*, 1997). Cette transition entre l'initiation et l'élongation est appelée le dégagement du promoteur.

TFIIH est aussi impliqué dans le dégagement du promoteur via le complexe CAK (Cdk7, cycline H et MAT1), qui possède une activité kinase (Feaver *et al.*, 1994). Dans la

cellule, CAK peut se retrouver seul ou en complexe avec TFIIH, ce qui influence sa préférence pour le substrat phosphorylé. CAK seul phosphoryle préférentiellement Cdk2, impliqué dans la régulation du cycle cellulaire, alors que CAK en complexe avec TFIIH phosphoryle le CTD de l'ARN pol II (Rossignol *et al.*, 1997). Le CTD peut être phosphorylé au niveau des sérines 2 (Ser 2) et 5 (Ser 5) de l'heptapeptide YSPTSPS. L'état de phosphorylation du CTD varie durant les différentes étapes du processus transcriptionnel. L'ARN pol II est recrutée au promoteur sous la forme hypophosphorylée, appelée forme IIa (Lu *et al.*, 1991). Suite à la formation du complexe pré-transcriptionnel, Cdk7 phosphoryle le CTD au niveau de la Ser 5 et convertit l'ARN pol II en une forme hyperphosphorylée, la forme IIo (Akoulitchev *et al.*, 1995). Cette phosphorylation du CTD par TFIIH permet à l'ARN pol II de dégager le promoteur, possiblement en permettant le relâchement de certains FGT, qui restent au promoteur (Pokholok *et al.*, 2002). En effet, il a été démontré que TBP et TFIIE interagissaient préférentiellement avec l'ARN pol IIa (Usheva *et al.*, 1992; Maxon *et al.*, 1994). D'autres kinases qui phosphorylent le CTD de l'ARN pol II, participant ainsi à la régulation de la transcription, seront décrites dans une section suivante (1.2.3.2).

TFIIH interagit aussi avec plusieurs régulateurs transcriptionnels tels que VP16 (*Herpes simplex virion protein*), FIR (*FBP interaction repressor*), HBX (*hepatitis B virus X protein*) et BRCA1 (*breast and ovarian tumor suppressor protein 1*) (Zurita et Merino, 2003). De plus, TFIIH, via Cdk 7, phosphoryle des récepteurs nucléaires hormonaux. Par exemple, il a été démontré dans des cellules de patients XP et TTD, qu'une mutation dans la partie C-terminale de XPD conduit à une diminution de la phosphorylation de certains récepteurs nucléaires AF-1 par TFIIH et conséquemment à une diminution de la capacité d'activation (Keriel *et al.*, 2002).

Récemment, une dixième sous-unité de TFIIH, TFB5, a été identifiée chez la levure et chez l'humain (Giglia-Mari *et al.*, 2004; Ranish *et al.*, 2004). Chez la levure, TFB5 est recruté au promoteur *in vivo* et *in vitro* et est requis pour la transcription basale et activée des gènes *GAL*. Il est aussi nécessaire au recrutement de TFIIH au promoteur. Chez l'humain, des mutations dans TFB5 causent le TTD groupe A. Jusqu'à présent,

aucune mutation dans les sous-unités de TFIIH n'avait permis d'expliquer le syndrome du TTD de ce groupe (Giglia-Mari *et al.*, 2004).

1.2.2.6 TFIIA

TFIIA est un facteur qui lie TBP et augmente son affinité pour la boîte TATA *in vitro* (Yokomori *et al.*, 1994). TFIIA en s'associant à TBP contacte simultanément l'ADN de façon non-spécifique en amont et au niveau de l'élément TATA (Tan *et al.*, 1996; Geiger *et al.*, 1996). *In vitro*, TFIIA n'est pas requis pour la transcription basale en présence de TBP (Ozer *et al.*, 1994; Sun *et al.*, 1994; Yokomori *et al.*, 1994; Ma *et al.*, 1996; Coulombe *et al.*, 1992; Cortes *et al.*, 1992). Cependant, lorsque TFIID est utilisé au lieu de TBP seul, la transcription devient dépendante en TFIIA (Sun *et al.*, 1994) et moins le système est pur, plus la dépendance est importante. Cette différence a été attribuée à la capacité de TFIIA à déplacer des répresseurs liant TBP, présents dans TFIID (Chicca *et al.*, 1998; Mermelstein *et al.*, 1996; Merino *et al.*, 1993; Kokubo *et al.*, 1998). Par exemple, TFIIA compétitionne avec TAF1 de levure pour la liaison à TBP. TAF1 interagit avec TBP et inhibe sa capacité à lier la boîte TATA (Kokubo *et al.*, 1998).

Chez la levure, il a été démontré par l'utilisation d'une souche ayant un allèle mutant de TBP, incapable d'interagir avec TFIIA, que l'absence de liaison TBP-TFIIA affectait l'expression d'environ 3% des gènes (Stargell *et al.*, 2000). Dans cette étude, la capacité de plusieurs activateurs à stimuler la transcription était aussi affectée. D'autres études *in vitro* et *in vivo* montrent que TFIIA est essentiel à la fonction de plusieurs activateurs, tels Sp1, VP16, Zta et NTF1 (Yokomori *et al.*, 1994; Stargell et Struhl, 1995; Lieberman *et al.*, 1997; Lieberman, 1994; Kobayashi *et al.*, 1995).

TFIIA humain est composé de trois sous-unités, TFIIA α (35 kDa), β (21 kDa) et γ (12 kDa) (Ma *et al.*, 1993; DeJong et Roeder, 1993; Ozer *et al.*, 1994; DeJong *et al.*, 1995). Les sous-unités α et β sont codées par le gène *TOA1*, alors que la sous-unité γ est codée par le gène *TOA2*. Chez la levure, il n'existe que deux sous-unités, TOA1 et TOA2

(Ranish *et al.*, 1992). La partie N-terminale de TOA1 est homologue à une région de α , alors que la partie C-terminale est homologue à une région de β (Ma *et al.*, 1993; DeJong et Roeder, 1993). TOA2 est homologue à γ (Ozer *et al.*, 1994; DeJong *et al.*, 1995).

Les sous-unités α et β humaines, sont produites par clivage post-traductionnel d'un précurseur (Hoiby *et al.*, 2004). Il a été démontré que le clivage de TFIIA le cible pour l'ubiquitination et la dégradation par le protéasome (Hoiby *et al.*, 2004). Cette dégradation permettrait possiblement d'assurer un contrôle du niveau de TFIIA cellulaire et de réguler ainsi la transcription. À l'origine, seules les formes clivées de α et β avaient été détectées dans les extraits cellulaires en association avec γ . Plus récemment, la forme non-clivée TFIIA $\alpha\beta$ en complexe avec γ et TBP a été isolée à partir de cellules de carcinome embryonnaire (Mitsiou et Stunnenberg, 2000; Mitsiou et Stunnenberg, 2003). Ce complexe, qui a été nommé TAC (*TBP-TFIIA-containing complex*), s'associe à la chromatine *in vivo* et est impliqué dans l'activation transcriptionnelle chez ces cellules.

1.2.3 L'élongation de l'ARNm et la terminaison de la transcription

Suite à l'initiation de la transcription et à la synthèse des 10 premiers liens phosphodiester, l'ARN pol II dégage le promoteur pour entrer dans le mode de l'élongation productive. À cette étape, un échange entre les facteurs généraux de transcription et les facteurs d'élongation se produit (Pokholok *et al.*, 2002). Des expériences de ChIP chez la levure ont montré que les facteurs généraux de transcription TBP, TFIIB, TFIIE, TFIIF, TFIIH, ainsi que le médiateur étaient localisés uniquement aux promoteurs de gènes exprimés, mais pas dans la région transcrite. Au contraire, les facteurs d'élongation TFIIS, Spt5 et le complexe Paf1/Cdc73 ont été retrouvés aux promoteurs et dans la région codante de ces gènes (Pokholok *et al.*, 2002).

1.2.3.1 Les facteurs d'élongation

Il existe plusieurs facteurs qui régulent l'élongation via différents mécanismes (revu par Sims, III *et al.*, 2004). Certains facteurs modulent l'activité de l'ARN pol II en influençant les pauses et les arrêts. D'autres participent au remodelage de la chromatine et à la modification covalente des histones durant l'élongation.

1.2.3.1.1 Les facteurs d'élongation qui modulent l'activité de l'ARN pol II

1.2.3.1.1.1 TFIIF

TFIIF possède un rôle non seulement au niveau de l'initiation de la transcription, mais aussi à l'étape d'élongation de l'ARNm. Plusieurs études ont montré que TFIIF stimule la vitesse d'élongation et supprime les pauses transitoires faites par la polymérase (Flores *et al.*, 1989; Price *et al.*, 1989; Bengal *et al.*, 1991; Izban et Luse, 1992a; Tan *et al.*, 1994). Les deux sous-unités de TFIIF sont impliquées dans cette activité (Tan *et al.*, 1994; Lei *et al.*, 1998). De plus, TFIIF joue un rôle dans le dégagement du promoteur, puisqu'il supprime les initiations avortées et contribue avec TFIIFH à prévenir les arrêts des complexes d'élongation précoces (Yan *et al.*, 1999). Tel que mentionné précédemment, des expériences de ChIP ont montré que TFIIF, ainsi que les autres FGT, étaient localisés uniquement aux promoteurs des gènes *in vivo* et non dans la région transcrite (Pokholok *et al.*, 2002). Il a été proposé que TFIIF s'associerait à la pol II en élongation de manière transitoire dans la région codante afin de supprimer les pauses, mais qu'il ne voyagerait pas de manière continue avec celle-ci (Zawel *et al.*, 1995; Sims, III *et al.*, 2004). En support à l'hypothèse d'un rôle de TFIIF en élongation, celui-ci interagit avec plusieurs facteurs d'élongation dont Spt5 (Lindstrom *et al.*, 2003), DSIF (Lindstrom *et al.*, 2003) et des composants du complexe Paf (Shi *et al.*, 1997).

1.2.3.1.1.2 TFIIS

TFIIS est un facteur d'élongation qui aide l'ARN pol II à sortir des sites d'arrêts (revu par Fish et Kane, 2002; Conaway *et al.*, 2003). Chez la levure, le gène codant pour TFIIS, *PPR2*, n'est pas essentiel à la viabilité, mais son inactivation confère une sensibilité à la drogue 6-azauracil (6-AU) (Nakanishi *et al.*, 1992; Williams et Kane, 1996). Le 6-AU réduit la concentration intracellulaire de GTP et UTP (Exinger et Lacroute, 1992). Des expériences *in vitro* ont montré que la réduction de la concentration de nucléotides augmentait la fréquence des arrêts et diminuait l'efficacité d'élongation (Izban et Luse, 1992b; Reines, 1992; Reines *et al.*, 1992). Ainsi, la sensibilité au 6-AU reflète souvent un défaut en élongation (Riles *et al.*, 2004).

L'arrêt est défini comme une interruption de la transcription, de laquelle la polymérase ne peut sortir sans l'aide de facteurs auxiliaires. Le modèle actuel propose que lorsque l'ARN pol possède en son site catalytique un hybride ADN-ARN instable, l'enzyme a tendance à reculer sur la matrice (*backtracking*) jusqu'à l'atteinte d'un registre plus stable et à s'arrêter (Nudler *et al.*, 1997). Afin de sortir de l'arrêt et de poursuivre la transcription, elle doit cliver l'extrémité 3' de l'ARNm. La pause, quant à elle, peut être résolue sans l'aide de facteurs supplémentaires et des études montrent qu'elle ne passerait pas par une étape de *backtracking* et de clivage de l'ARNm (Erie, 2002; Neuman *et al.*, 2003).

L'activité de clivage de l'extrémité 3' de l'ARNm est une activité endonucléotidique, intrinsèque à l'ARN polymérase (Krummel et Chamberlin, 1989; Surratt *et al.*, 1991; Fish et Kane, 2002; Wind et Reines, 2000). Cette activité est grandement stimulée par les facteurs Gre A et Gre B chez les procaryotes (Borukhov *et al.*, 1992; Borukhov *et al.*, 1993) et TFIIS chez les eucaryotes (Izban et Luse, 1992b; Reines, 1992; Wang et Hawley, 1993). La structure cristallographique de TFIIS en complexe avec l'ARN pol II a permis de mieux comprendre comment TFIIS stimule le clivage du transcrit (Kettenberger *et al.*, 2003). TFIIS entre dans le site catalytique via le pore et y insère deux acides aminés acides, conservés entre les espèces (Jeon *et al.*,

1994), qui se retrouvent près de la localisation proposée pour un des ions métalliques qui forment le site catalytique (métal B) (Kettenberger *et al.*, 2003). Il a été suggéré que les deux acides aminés acides participent à la coordination de ce dernier (Kettenberger *et al.*, 2003). Un mécanisme à deux ions métalliques, métal A et B, a été proposé pour les activités de polymérisation et de clivage par les ARN polymérases (Steitz, 1998; Opalka *et al.*, 2003; Sosunov *et al.*, 2003; Sosunova *et al.*, 2003; Cramer *et al.*, 2001; Gnatt *et al.*, 2001). Ce mécanisme sera décrit de façon plus détaillée à la section 1.3.3.

Des études récentes de cinétique, faites avec un essai qui mesure avec précision la vitesse de formation de liens phosphodiester spécifiques, ont permis de décrire avec plus de détails les rôles de TFIIF et TFIIS dans la modulation de l'activité de l'ARN pol II lors de l'élongation du transcrit (Zhang *et al.*, 2003; Zhang et Burton, 2004). Ces études montrent que TFIIF augmente la vitesse d'élongation de la polymérase, en stabilisant l'état post-transloqué du complexe transcriptionnel (Zhang et Burton, 2004). La translocation des acides nucléiques, qui permet à l'enzyme de se déplacer sur la matrice d'ADN, sera discutée à la section 1.3.3. TFIIS, quant à lui, n'a pas d'effet direct sur la vitesse de l'enzyme. Cependant, il stimule l'élongation en supprimant les pauses transitoires. Ensemble, les deux facteurs ont un effet combinatoire pour la suppression de ces dernières (Zhang *et al.*, 2003).

1.2.3.1.1.3 Autres facteurs d'élongation modulant l'activité de l'ARN pol II

Le complexe Elongin et le facteur ELL stimulent aussi l'élongation en supprimant les pauses dans la transcription (Bradsher *et al.*, 1993a; Bradsher *et al.*, 1993b; Shilatifard *et al.*, 1996). Les facteurs d'élongation P-TEFb (*positive transcription elongation factor b*), DSIF (*DRB-sensitivity-inducing factor*) et NELF (*negative elongation factor*) ont été identifiés grâce à leur capacité à induire la sensibilité au DRB (5,6-dichloro-1-b-D-ribofuranosylbenzimidazole) (Wada *et al.*, 1998a). Le traitement au DRB de cellules de mammifère affecte l'élongation, conduisant à la formation de transcrits tronqués (Chodosh *et al.*, 1989; Yamaguchi *et al.*, 1998). Dans un essai reconstitué avec des

facteurs purifiés, la transcription n'est pas sensible au DRB. Les facteurs énumérés ci-dessus lui confèrent cette sensibilité (Wada *et al.*, 1998b).

P-TEFb est un facteur d'élongation qui phosphoryle le CTD de l'ARN pol II au niveau de la Ser 2 (Marshall et Dahmus, 2000; Shim *et al.*, 2002; Ni *et al.*, 2004). P-TEFb co-localise avec l'ARN pol II hypophosphorylée en pause au promoteur (Lis *et al.*, 2000) et contrecarre les pauses induites par DSIF/NELF (Sims, III *et al.*, 2004). DSIF est un complexe composé de Spt4 et Spt5, les homologues humains des protéines portant le même nom, identifiées à l'origine chez *S. cerevisiae* (Winston *et al.*, 1984). Spt4 et 5 interagissent avec l'ARN pol II (Wada *et al.*, 1998a) et jouent un rôle positif en élongation (Rondon *et al.*, 2004; Bourgeois *et al.*, 2002). Par exemple, Spt4 relève les pauses de la polymérase induites par le facteur de remodelage de la chromatine Isw1p (Morillon *et al.*, 2003). Spt5, quant à lui, coopère avec TAT pour prévenir la terminaison prématurée et les pauses en élongation tardive en transcription des gènes du VIH (Bourgeois *et al.*, 2002). NELF est un complexe ayant un effet négatif sur l'élongation en favorisant les pauses transcriptionnelles en présence de DSIF (Yamaguchi *et al.*, 1999). P-TEFb peut contrer l'effet de NELF en phosphorylant le CTD de l'ARN pol II et la sous-unité Spt5 de DSIF. Ceci entraîne la dissociation de NELF du complexe d'élongation en pause (Yamaguchi *et al.*, 1999; Kim et Sharp, 2001; Ivanov *et al.*, 2000). Il a été proposé que les pauses induites par DSIF/NELF permettent le recrutement de l'enzyme de coiffage et conséquemment le coiffage de l'ARNm en 5' du gène. La phosphorylation du CTD et de DSIF par P-TEFb entraînerait le départ de NELF et le retour en élongation productive (Sims, III *et al.*, 2004).

1.2.3.1.2 Les facteurs d'élongation qui modifient la structure de la chromatine

La présence de nucléosomes sur l'ADN *in vivo* peut constituer un blocage majeur à l'élongation par l'ARN pol II. Certains complexes de remodelage et plusieurs protéines modifiant de façon covalente la queue des histones participent à la régulation de l'élongation en modulant la structure de la chromatine. Par exemple, le complexe FACT (*facilitates chromatin transcription*) a été isolé grâce à sa capacité à stimuler la

transcription sur une matrice de chromatine (Orphanides *et al.*, 1998). FACT est hautement conservé entre les espèces eucaryotes et joue un rôle après l'initiation (Belotserkovskaya et Reinberg, 2004). Chez la levure, le complexe est formé des sous-unités Spt16/Cdc68 et Pob3. Des mutants de Spt16 sont sensibles au 6-AU, suggérant un rôle en élongation. De plus, Spt16 interagit génétiquement et physiquement avec plusieurs autres facteurs d'élongation (Squazzo *et al.*, 2002; Krogan *et al.*, 2002; Sims, III *et al.*, 2004). De nombreuses études supportent un modèle dans lequel FACT facilite l'élongation sur la chromatine en déstabilisant la structure du nucléosome et en enlevant spécifiquement un dimère de H2A-H2B (Kireeva *et al.*, 2002; Orphanides *et al.*, 1999; Belotserkovskaya *et al.*, 2003). Ceci permettrait à la polymérase de traverser le nucléosome. FACT est aussi impliqué dans le rétablissement de la structure chromatinienne, suite à la transcription (Belotserkovskaya *et al.*, 2003; Schwabish et Struhl, 2004).

La méthyltransférase Set1 de levure catalyse le transfert de groupements méthyl sur la lysine 4 (K4) des histones H3 (Briggs *et al.*, 2001). Set1 s'associe préférentiellement avec l'ARN pol II phosphorylée en Ser5 et induit la triméthylation de H3-K4 sur les gènes activement transcrits (Ng *et al.*, 2003). Il a été démontré que Isw1p, une protéine impliquée dans le remodelage de la chromatine et qui régule différentes étapes de la transcription dont l'élongation, reconnaît les histones H3-K4 méthylées (Morillon *et al.*, 2003; Santos-Rosa *et al.*, 2003). De plus, Isw1p et Set1 sont requis pour la distribution adéquate de l'ARN pol II sur certains gènes et le recrutement de facteurs de clivage et de polyadénylation (Santos-Rosa *et al.*, 2003), suggérant un rôle dans la régulation de l'élongation pour ces deux protéines, via une modification de la structure de la chromatine.

1.2.3.2 Le rôle du CTD dans le cycle transcriptionnel

Le CTD de Rpb1 joue un rôle très important dans l'élongation et dans le processus transcriptionnel. Tel que mentionné précédemment, le CTD est phosphorylé par TFIIF au niveau de la Ser 5 suite à la formation du complexe pré-transcriptionnel.

Cette phosphorylation de la Ser 5 permet le recrutement de l'enzyme de coiffage de l'ARNm (Cho *et al.*, 1997; Komarnitsky *et al.*, 2000). Chez la levure, la kinase CTDK-1 (*CTD kinase 1*) phosphoryle le CTD sur la Ser 2 durant l'élongation (Cho *et al.*, 2001), permettant l'association des facteurs de clivage et de polyadénylation de l'ARNm au complexe d'élongation dans la partie 3' du gène (Kim *et al.*, 2004a; Ahn *et al.*, 2004). CTDK-1 est composé de trois sous-unités (CTK1 à CTK3). CTK1, la sous-unité qui possède l'activité kinase, est homologue à la kinase cycline-dépendante Cdk9, une sous-unité du facteur d'élongation de mammifère P-TEFb, qui phosphoryle aussi le CTD (Price, 2000).

L'ARN pol II doit être déphosphorylée à la fin du processus transcriptionnel, afin de lui permettre d'être recrutée de nouveau au promoteur et d'entreprendre un autre cycle de transcription. La phosphatase Fcp1 (*TFIIF-associated CTD phosphatase*) a été clonée dans des expériences de double-hybride chez la levure, grâce à sa capacité d'interagir avec la sous-unité RAP74 de TFIIF (Archambault *et al.*, 1997; Archambault *et al.*, 1998b). Fcp1 déphosphoryle le CTD au niveau des Ser 2, contrecarrant ainsi l'activité de CTK1 (Cho *et al.*, 2001). L'activité de Fcp1 est stimulée par TFIIF et TFIIB inhibe cette stimulation (Chambers *et al.*, 1995). Fcp1 se lie à l'ARN pol II via les sous-unités Rpb4 et Rpb7 (Kamenski *et al.*, 2004). Une autre phosphatase, Scp1 (*small CTD-phosphatase 1*) déphosphoryle le CTD au niveau de la Ser 5 (Yeo *et al.*, 2003). Tout comme Fcp1, l'activité de Scp1 est stimulée par TFIIF (Yeo *et al.*, 2003). Finalement, la protéine Ssu72 déphosphoryle aussi les Ser 5 du CTD (Krishnamurthy *et al.*, 2004). Ssu72 fait partie du complexe de clivage et de polyadénylation des ARNm CPF (*cleavage/polyadenylation factor*). Comme la fonction de Ssu72 en 3' des gènes est indépendante de son activité phosphatase, il a été proposé que celui-ci jouerait deux rôles en transcription : le premier dans la déphosphorylation du CTD pour le recyclage de la polymérase et le second comme facteur essentiel au clivage du transcrit et à la terminaison de la transcription (Krishnamurthy *et al.*, 2004).

1.2.3.3 La terminaison de la transcription

La formation de l'extrémité 3' du transcrit passe par le clivage et la polyadénylation de celui-ci. Le site de clivage de l'ARNm se situe au niveau d'un dinucléotide CA, entre la séquence hautement conservée AAUAAA et l'élément DSE (*downstream sequence element*) (revu par Proudfoot *et al.*, 2002). Deux complexes reconnaissent ces séquences présentes dans l'ARNm transcrits. Le complexe CPSF (*cleavage and polyadenylation specificity factor*) interagit avec la séquence AAUAAA, alors que CstF (*cleavage stimulatory factor*) lie le DSE. L'association de CPSF et CstF à l'ARNm est coopérative puisque la liaison de l'un augmente l'affinité de l'autre pour son site (Zhao *et al.*, 1999). Les facteurs CF I (*cleavage factor I*), CF II (*cleavage factor II*), ainsi que la poly(A) polymérase (PAP) sont aussi essentiels pour le clivage de l'ARNm. Suite au clivage, la PAP, en présence de CPSF, ajoute la queue de poly(A). La protéine PABP II (*poly(A) binding protein II*) se lie à la chaîne de poly(A) et augmente la processivité de la PAP (Wahle et Ruegsegger, 1999).

Le mécanisme de terminaison de la transcription par l'ARN pol II est peu connu. La terminaison implique un relâchement du transcrit préalablement clivé et un relâchement de l'enzyme de la matrice d'ADN. Elle est dépendante d'un signal de polyadénylation fonctionnel et la force du signal dicte l'efficacité de la terminaison (Osheim *et al.*, 1999). D'autres facteurs sont importants pour la terminaison suite au clivage du transcrit. Par exemple, il a été démontré que l'homologue du facteur humain CstF-64 chez la levure est requis pour la terminaison, indépendamment de sa fonction dans le clivage du transcrit. De plus, le coactivateur PC4 interagit avec le domaine C-terminal de CstF-64 et influence la terminaison (Calvo et Manley, 2001). Finalement, des séquences en aval du signal de polyadénylation jouent un rôle important dans la terminaison. Une séquence identifiée sur le gène de la β -globine, située entre 900 et 1600 pb en aval du site de polyadénylation, induit une activité de clivage cotranscriptionnel du transcrit qui a été nommée CoTC (*cotranscriptional cleavage*). Cette activité est complètement indépendante du clivage et de la polyadénylation de l'ARNm, qui se passent beaucoup plus en amont, et est requise pour la terminaison (Dye et Proudfoot,

2001). Le CoTC implique une activité d'auto-clivage de l'ARNm (Teixeira *et al.*, 2004). Récemment, il a été démontré que les exonucléases Rat1 chez la levure et Xrn2 chez les mammifères participent à une dégradation 5'-3' de la portion d'ARN aval, générée suite au clivage du transcrit. Cette activité est nécessaire à la terminaison de la transcription (Kim *et al.*, 2004b; West *et al.*, 2004). Il a été proposé que l'activité d'auto-clivage de l'ARN CoTC, au gène de la β -globine, agit comme précurseur de la terminaison en fournissant une extrémité 5', qui sert ensuite de substrat à l'exonucléase Xrn2 (Teixeira *et al.*, 2004; Tollervey, 2004; West *et al.*, 2004).

1.3 L'ARN polymérase II : structure et mécanisme catalytique

1.3.1 L'ARN polymérase II

L'ARN pol II est composée de 12 sous-unités, appelées Rpb1 à Rpb12, de poids moléculaires variant entre 10 et 220 kDa (tableau II) (revu par Young, 1991; Coulombe et Burton, 1999; Lee et Young, 2000). L'ARN pol II est hautement conservée entre les espèces eucaryotes. Par exemple, il existe 70% d'homologie et 50% d'identité entre Rpb1 de levure et Rpb1 humain. Les deux plus grandes sous-unités, Rpb1 et Rpb2, sont les sous-unités catalytiques de l'enzyme. Elles sont homologues aux sous-unités β' et β de l'ARN pol bactérienne, composée de quatre sous-unité ($\alpha_2\beta\beta'$) et du facteur de reconnaissance des promoteurs σ . Rpb3 et Rpb11 présentent des homologies avec α et forment un hétérodimère *in vitro*, semblable à l'homodimère formé par α lors de l'assemblage de l'ARN pol bactérienne (Larkin et Guilfoyle, 1997). Les sous-unités Rpb5, 6, 8, 10 et 12 sont communes aux trois ARN pol. Plusieurs sous-unités non-identiques, dont les sous-unités catalytiques, sont tout de même homologues entre les ARN pol I, II et III. Bien que les rôles des sous-unités communes ont été peu étudiés, il est connu que Rpb5 lie la protéine virale HBx (*Hepatitis B virus X protein*), qui active la transcription (Cheong *et al.*, 1995). Rpb6 est impliqué dans l'assemblage et stabilise le complexe de 12 sous-unités (Minakhin *et al.*, 2001). De plus, elle contribue à l'association de Rpb4 et Rpb7 à l'ARN pol II (Tan *et al.*, 2003).

Tableau II. Les sous-unités de l'ARN polymérase II

Sous-unité	Poids moléculaire (humain) (kDa)	Essentielle (levure)	Caractéristiques
Rpb1	220	+	catalytique, CTD
Rpb2	140	+	catalytique
Rpb3	33	+	assemblage
Rpb4	16	-	sous-complexe avec Rpb7
Rpb5	25	+	commune
Rpb6	14.4	+	commune, assemblage
Rpb7	19	+	sous-complexe avec Rpb4
Rpb8	17	+	commune
Rpb9	14.5	-	élongation, choix du site d'initiation
Rpb10	7.6	+	commune
Rpb11	13.3	+	assemblage
Rpb12	7	+	commune

Chez la levure, tous les gènes codant pour les sous-unités de l'ARN pol II sont essentiels, à l'exception des gènes codant pour Rpb4 et Rpb9. La délétion de Rpb9 confère une sensibilité au 6-azauracil et produit un phénotype synthétique lorsque combinée avec une délétion du gène codant pour TFIIS (Hemming *et al.*, 2000). De plus, il a été démontré que Rpb9 joue un rôle en élongation en stimulant, avec TFIIS, la sortie de l'ARN pol II des sites d'arrêts (Awrey *et al.*, 1997). Rpb9 est aussi important pour le choix du site d'initiation de la transcription (Hull *et al.*, 1995; Sun *et al.*, 1996). La délétion de Rpb9 dans les cellules conduit à un déplacement du site d'initiation, en amont du site normalement utilisé par la polymérase. Ce déplacement serait dû à une interaction déficiente entre TFIIF et l'ARN pol II Δ Rpb9 (Ziegler *et al.*, 2003). Finalement, Rpb9 est impliqué au niveau du processus TCR (*transcription coupled repair*) dans une voie indépendante de la protéine de réparation Rad26 (Li et Smerdon, 2002).

La délétion de Rpb4 confère, quant à elle, un phénotype de sensibilité à la chaleur (Woychik et Young, 1989; Giaever *et al.*, 2002). Des expériences d'analyse d'expression sur micropuces ont permis de montrer que 98% des gènes ne sont plus exprimés dans la lignée Δ rpb4 à la température restrictive. À la température permissive, les gènes

constitutifs sont transcrits normalement, alors que l'expression des gènes inductibles est presque complètement abolie (Miyao *et al.*, 2001). L'exposition à certains stress (éthanol, milieu pauvre en nutriments) au niveau de ces cellules conduit à une suppression de l'expression de la plupart des gènes (Choder et Young, 1993; Sheffer *et al.*, 1999). Chez la levure, Rpb4 et Rpb7 forment un sous-complexe dissociable du reste de l'ARN pol II (Edwards *et al.*, 1991). Quatre-vingt pourcent de l'ARN pol II purifiée à partir de cellules en croissance exponentielle ne contient pas le sous-complexe. Au contraire, en phase stationnaire, la majorité de l'ARN pol II isolée contient Rpb4 et Rpb7 (Choder et Young, 1993). *In vitro*, l'ARN pol II Δ Rpb4/Rpb7 est incapable d'initier la transcription de manière spécifique à un promoteur, bien qu'elle soit active sur une matrice ne contenant pas de promoteur (Edwards *et al.*, 1991; Orlicky *et al.*, 2001). Rpb4/7 ne sont cependant pas requis pour la liaison au promoteur, mais bien pour l'étape subséquente d'initiation de la transcription (Orlicky *et al.*, 2001).

Plusieurs hypothèses ont été avancées pour tenter d'expliquer ces résultats, ainsi que le rôle de Rpb4 et Rpb7 dans le processus transcriptionnel (Choder, 2004). Rpb4 et 7 pourraient être importants pour prévenir un changement de conformation de l'ARN pol II en conditions de stress. Rpb4/7 pourraient être nécessaires à la transcription d'un nombre restreint de promoteurs. Finalement, le sous-complexe pourrait interagir avec le reste de l'ARN pol II seulement durant certaines étapes du cycle de la transcription. Rpb4 est aussi impliqué dans des étapes post-transcriptionnelles du processus de l'expression génétique. En effet, il a été démontré que Rpb4 est essentiel pour l'exportation des ARNm au cytoplasme en conditions de stress (Farago *et al.*, 2003). Cette fonction de Rpb4 est indépendante de son rôle en transcription.

1.3.2 Structure de l'ARN polymérase II

1.3.2.1 La structure de l'ARN polymérase II à 10 sous-unités

La publication de structures cristallographiques à haute résolution de l'ARN pol II de levure a permis de proposer plusieurs hypothèses sur le fonctionnement de cette

enzyme et sur le mécanisme transcriptionnel (Gnatt *et al.*, 2001; Cramer *et al.*, 2001; Bushnell et Kornberg, 2003; Bushnell *et al.*, 2002; Bushnell *et al.*, 2004; Westover *et al.*, 2004a; Westover *et al.*, 2004b; Kettenberger *et al.*, 2003; Armache *et al.*, 2003). Les structures de l'ARN pol II à 10 sous-unités (Δ Rpb4/7), seule (Cramer *et al.*, 2001) ou en élongation sur l'ADN (Gnatt *et al.*, 2001; Westover *et al.*, 2004a), ont montré que Rpb1 et Rpb2 forment le cœur de l'enzyme au centre duquel se retrouve un canal chargé positivement (figures 2 et 3) (revu par Langelier *et al.*, 2002). Ce canal, qui accueille les acides nucléiques, est appelé *cleft*, sillon ou canal principal. Un domaine protéique de Rpb2, le mur ou *wall*, en ferme l'extrémité. Au fond du sillon, se retrouve le site catalytique de l'enzyme, formé de résidus acides de Rpb1 et Rpb2 strictement conservés, qui coordonnent deux ions métalliques, les métaux A et B.

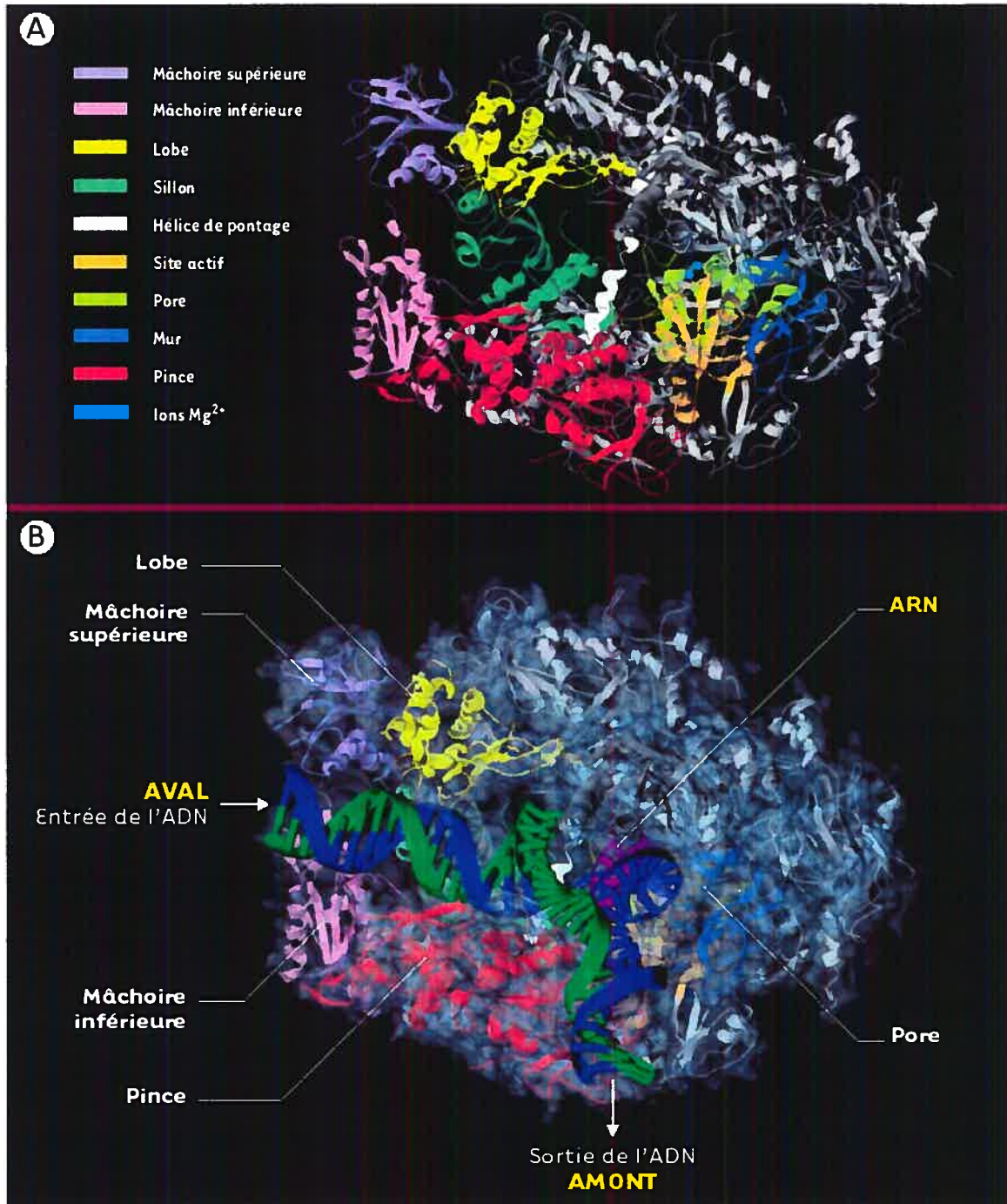


Figure 2 : Structure cristallographique de l'ARN polymérase II à 10 sous-unités. A) Enzyme libre à une résolution de 2.8 Å. Les principaux domaines ont été colorés selon le code indiqué. B) Enzyme en élongation sur une matrice d'ADN à une résolution de 3.3 Å. Le code de couleur est le même que celui présenté en A. Figure tirée de Langelier *et al.*, 2002.

Les acides nucléiques sont retenus dans le canal principal par plusieurs domaines protéiques, dont les mâchoires supérieures et inférieures (*upper and lower jaws*), qui contactent l'ADN en aval, et la pince (*clamp*). La structure de l'ARN pol II seule a révélé que la pince est dans une conformation dite ouverte, prête à accueillir l'ADN dans le canal principal. Au contraire, au niveau de la structure de la polymérase en élongation, la pince est dans une conformation fermée. Ceci a conduit les auteurs à proposer que la pince est un élément mobile qui se referme sur les acides nucléiques pour stabiliser le complexe d'élongation (Cramer *et al.*, 2001; Gnatt *et al.*, 2001). Cependant, une structure plus récente de l'ARN pol II comprenant les 12 sous-unités a entraîné une modification de ce modèle (Armache *et al.*, 2003; Bushnell et Kornberg, 2003) (section 1.3.2.1). Un deuxième canal, appelé canal secondaire ou pore, a la forme d'un entonnoir inversé (*funnel*), qui s'ouvre dans le canal principal, tout près du site actif. Il a été proposé que le nucléotide (NTP) arrivant, ainsi que le pyrophosphate relâché, circulent par le pore. Le pore accommoderait aussi l'extrémité 3' du transcrit lorsque l'ARN pol II recule sur la matrice d'ADN, lors des arrêts transcriptionnels (Cramer *et al.*, 2001; Gnatt *et al.*, 2001). Le canal qui permet la sortie de l'ARNm se retrouve, quant à lui, entre le mur de Rpb2 et la pince de Rpb1 (figure 3).

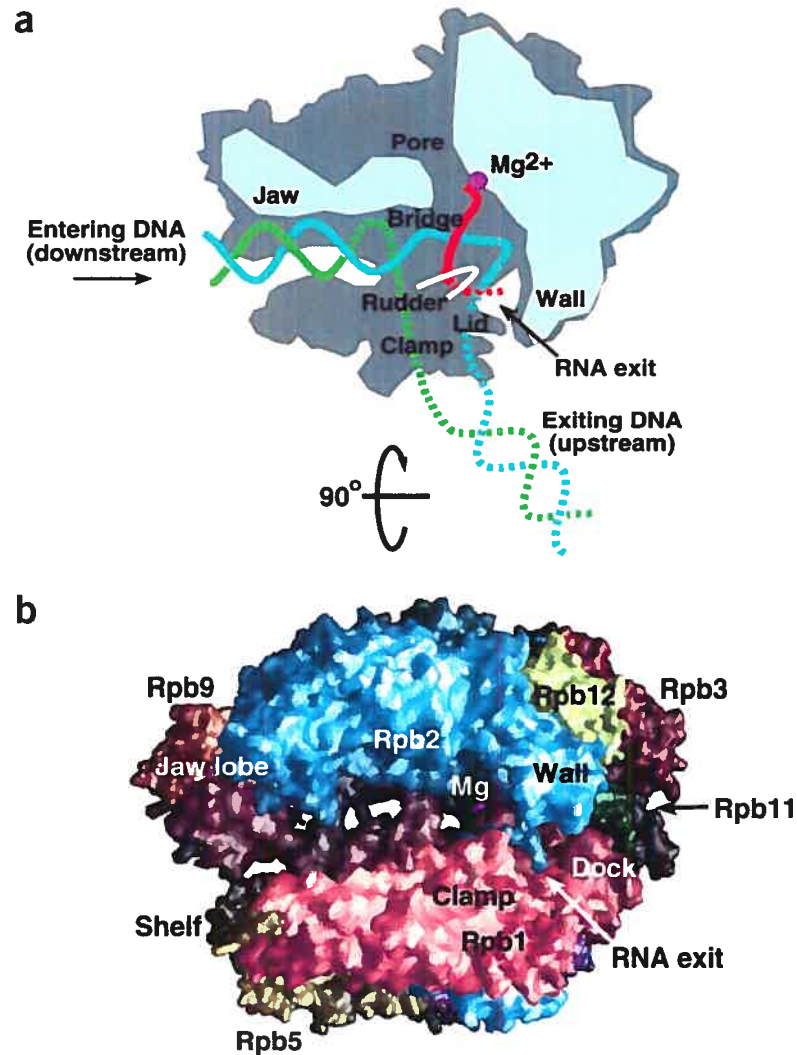


Figure 3 : Les canaux de l'ARN polymérase II. A) Le positionnement des acides nucléiques par rapport aux trois canaux de l'ARN pol II est représenté sous forme schématique. Dans cette représentation, l'ARN pol II a été scindée en deux pour montrer l'intérieur du site catalytique. Le brin matrice de l'ADN est indiqué en bleu; le brin transcrit, en vert; l'ARNm, en rouge et l'ion Mg^{2+} , en mauve. B) Vue de l'ARN pol II à 10 sous-unités montrant le canal principal ou sillon. Les sous-unités visibles sont colorées avec différentes couleurs et sont identifiées. Figure tirée de Hahn, 2004.

Tel que mentionné précédemment, la conservation de l'ARN pol II entre les espèces eucaryotes est grande, avec 53% d'identité entre l'humain et la levure, répartie sur toute la surface (figure 4A) (Cramer *et al.*, 2001). Lorsque l'on compare la séquence de l'ARN pol II de levure avec l'ARN polymérase bactérienne, on note que l'homologie est localisée près du site catalytique (figure 4C). L'homologie de structure, déterminée

par comparaison avec la structure de l'ARN pol bactérienne (Zhang *et al.*, 1999a), est encore plus étendue, si bien que seule la surface périphérique diffère (figure 4D). C'est cette surface qui est le plus souvent impliquée dans des interactions avec des protéines régulatrices, spécifiques aux procaryotes et aux eucaryotes (Cramer *et al.*, 2001).

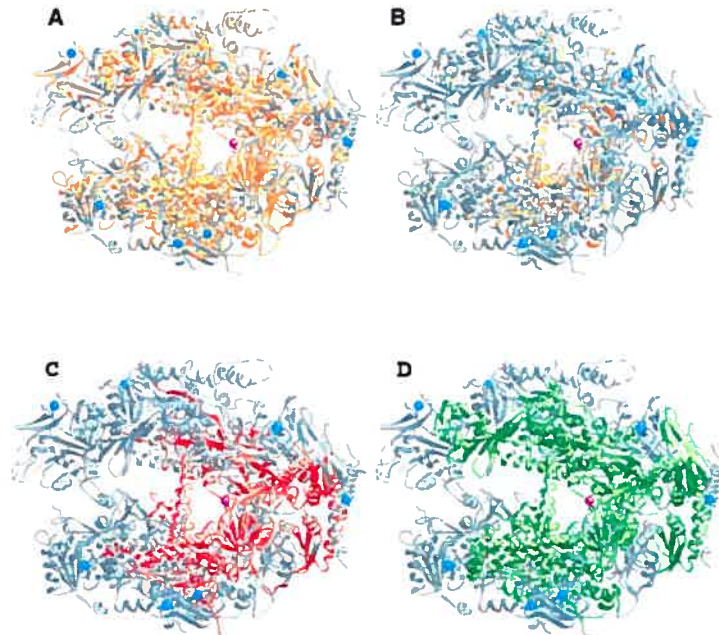


Figure 4 : Conservation des ARN polymérases. A) Les résidus identiques entre l'ARN pol II de levure et l'ARN pol II humaine sont colorés en orange. B) Les résidus identiques entre l'ARN pol II de levure et l'ARN pol de *E. coli* sont indiqués en orange. C) Les régions d'homologies de séquence entre les deux sous-unités catalytiques de l'ARN pol II de levure et de l'ARN pol de *E. coli* sont colorées en rouge. D) Les régions structurellement homologues entre l'ARN pol II de levure et l'ARN pol bactérienne sont montrées en vert. Figure tirée de Cramer *et al.*, 2001.

Les structures de l'ARN pol II en élongation, composées de l'enzyme liée à une matrice d'ADN et à un ARNm, ont révélé la présence d'un hybride ADN-ARN de 8 à 9 paires de bases dans le canal principal (Gnatt *et al.*, 2001; Westover *et al.*, 2004a). Ces structures ont aussi permis de montrer l'existence de boucles et d'hélices de Rpb1 et Rpb2, localisées près du site catalytique, dans la région de l'hybride. Ces éléments ont été nommés selon leur localisation ou leur fonction présumée dans le processus transcriptionnel (figure 5 et tableau III). Par exemple, le *bridge helix* ou hélice de pontage est une hélice de Rpb1, conservée entre les espèces eucaryotes, qui traverse le canal

principal et qui se retrouve près du site +1. Comme deux conformations distinctes de l'hélice ont été cristallographiées dans des structures différentes, l'une dans laquelle l'hélice de pontage est droite (Cramer *et al.*, 2001; Gnatt *et al.*, 2001; Westover *et al.*, 2004a) et l'autre dans laquelle elle est courbée (Zhang *et al.*, 1999a), il a été proposé que cette dernière oscille entre deux conformations pour induire la translocation des acides nucléiques durant la polymérisation (Gnatt *et al.*, 2001).

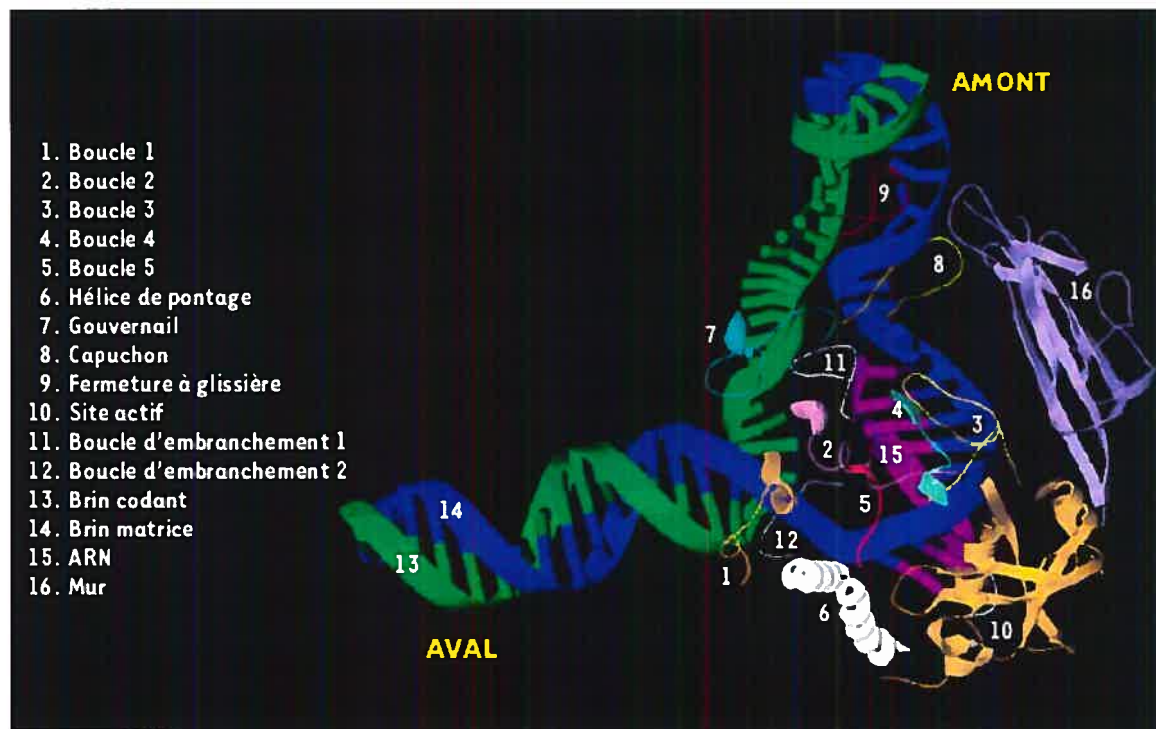


Figure 5 : Contacts des éléments structuraux de l'ARN polymérase II avec l'hybride ADN-ARN. Les domaines, boucles et hélices de Rpb1 et Rpb2 se retrouvant dans la région de l'hybride ADN-ARN sont numérotés et identifiés sur la figure. L'ADN matrice est coloré en bleu ; l'ADN transcrit, en vert et l'ARNm, en rose. Figure tirée de Langelier *et al.*, 2002.

La localisation des boucles *rudder* (gouvernail), *lid* (capuchon) et *fork loop 1* (boucle d'embranchement 1), suggère que celles-ci participent à la séparation de l'ARN de la matrice d'ADN, afin de maintenir un hybride de 8 à 9 pb (Westover *et al.*, 2004a). Le *rudder* contacte le brin d'ADN matrice aux positions -9, -10 et -11 et pourrait ainsi prévenir la réassociation de l'ARN. Chez l'ARN pol bactérienne, une étude fonctionnelle du *rudder* a cependant montré que cette boucle est impliquée dans l'initiation en

supportant l'ouverture de la bulle transcriptionnelle dans sa partie aval, ainsi que dans la stabilité du complexe d'élongation, mais pas dans le maintien de la longueur de l'hybride ADN-ARN (Kuznedelov *et al.*, 2002a). Le *lid* interagit avec l'ARNm aux positions -8, -9 et -10 et servirait possiblement à le guider vers le canal de sortie. La boucle *fork loop 1*, qui lie les bases -5, -6 et -7 de l'ARN, préviendrait la formation d'un hybride étendu. La boucle *fork loop 2* (boucle d'embranchement 2) pourrait être impliquée dans le maintien de l'extrémité aval de la bulle de transcription, alors que des rôles de dissociation de l'hybride, sortie de l'ARNm et maintien de l'extrémité amont de la bulle ont été proposés pour le *zipper* (fermeture à glissière).

Tableau III. Les éléments structuraux du site catalytique de l'ARN polymérase II

Sous-unité	Domaine	Position humain	% d'homologie levure-humain	Rôles proposés
Rpb1 (aa 1-1970)	Métal A	495,497,499	100	Coordination du Mg^{2+} -A ^{1,2} pour la catalyse (polymérisation)
	<i>Bridge helix</i>	834-868	97	Translocation des acides nucléiques ^{1,2}
	<i>rudder</i>	321-334	90	Séparation de l'hybride ADN-ARN ³
	<i>lid</i>	262-274	70	Séparation de l'hybride ADN-ARN, maintien de la bulle (aval) ^{3,4}
	<i>zipper</i>	45-55	70	Séparation de l'hybride ADN-ARN, sortie de l'ARNm et maintien de la bulle (amont) ^{1,2}
	<i>Switch 1</i>	1414-1436	87	Liaison du brin matrice d'ADN au niveau de l'hybride ADN-ARN, maintien de la bulle (aval) ^{4,5}
	<i>Switch 2</i>	342-361	100	Liaison du brin matrice d'ADN au niveau de l'hybride ADN-ARN, maintien de la bulle (aval) ^{4,5}
	<i>Switch 5</i>	1463-1464	100	Couplage fermeture de la pince à la présence d'un hybride ADN-ARN dans le sillon central ^{1,2}
Rpb2 (aa 1-1174)	Métal B	791-792	100	Coordination du Mg^{2+} -B ^{1,2} pour la catalyse (polymérisation)
	<i>Fork loop 1</i>	453-465	85	Séparation de l'hybride ADN-ARN ³

	<i>Fork loop 2</i>	487-499	100	Maintien de la bulle (aval) ^{1,2}
	<i>Flap</i>	874-888	47	Reconnaissance du promoteur ^{6,7} et des pauses (procaryote) ⁷
	<i>Switch 3</i>	1064-1086	95	Liaison du brin matrice d'ADN au niveau de l'hybride ADN-ARN ⁵
	<i>Switch 4</i>	1108-1115	86	Couplage fermeture de la pince à la présence d'un hybride ADN-ARN dans le sillon central ^{1,2}

1. (Cramer *et al.*, 2001), 2. (Gnatt *et al.*, 2001), 3. (Westover *et al.*, 2004a), 4. Kettenberger, 2004 à ajouter
5. (Armache *et al.*, 2003), 6. (Toulokhonov *et al.*, 2001), 7. (Toulokhonov et Landick, 2003), 8. (Kuznedelov *et al.*, 2002b)

La boucle *flap* est située au niveau du domaine *wall* et se retrouve près du canal de sortie de l'ARNm. Chez l'ARN pol bactérienne, il a été montré que le *flap* interagit avec les structures secondaires d'ARN *hairpins*, formées lorsque l'ARNm en voie de synthèse sort du canal (Toulokhonov *et al.*, 2001; Toulokhonov et Landick, 2003). Cette interaction est essentielle à l'induction des pauses transcriptionnelles par les structures *hairpins* (Toulokhonov *et al.*, 2001; Toulokhonov et Landick, 2003), un mécanisme de régulation courant chez la pol bactérienne (Mooney *et al.*, 1998). Le *flap* est aussi impliqué dans un changement de conformation du facteur σ bactérien, nécessaire à la reconnaissance du promoteur (Kuznedelov *et al.*, 2002b). En effet, la plupart des promoteurs procaryotes sont composés de deux éléments, les éléments -10 et -35, liés par σ . L'élément -10 est reconnu par la région 2.4 de σ , alors que l'élément -35 l'est par la région 4.2 (Gross *et al.*, 1998). La boucle *flap*, qui lie la région 4, induit l'éloignement des domaines 2 et 4 pour permettre à cette dernière de s'associer à l'élément -35 pour la reconnaissance du promoteur (Kuznedelov *et al.*, 2002b). Aucune étude n'a été faite à ce jour pour savoir si les rôles du *flap* étaient conservés chez l'ARN pol II.

Cinq boucles de Rpb1 et Rpb2 situées près de la pince ont été appelées les *switches* (boucles 1 à 5 de la figure 5). Ces boucles ont des conformations différentes entre la structure de l'ARN pol II libre (Cramer *et al.*, 2001) et la structure en élongation (Gnatt *et al.*, 2001). Les *switches* 1, 2 et 3 sont désordonnées dans le cristal de l'ARN pol II seule, mais ordonnées dans celui de l'enzyme en élongation, où elles contactent les

acides nucléiques au niveau de l'hybride ADN-ARN. Il a donc été proposé que les *switches* servent à coupler la fermeture de la pince à la présence d'un hybride ADN-ARN dans le canal principal (Gnatt *et al.*, 2001). Cependant, suite à la résolution de deux structures cristallographiques de l'ARN pol II de levure comprenant les douze sous-unités, de nouveaux rôles ont été suggérés pour la pince et les *switches* dans le mécanisme transcriptionnel.

1.3.2.1 La structure de l'ARN polymérase II à 12 sous-unités

La purification à homogénéité de l'ARN pol II de levure comprenant les 12 sous-unités, dans le but d'obtenir un cristal, est difficile à réaliser à cause de la présence du sous-complexe Rpb4/7 en quantité sous-stoichiométrique au sein de l'enzyme. Deux groupes ont utilisé des stratégies différentes afin d'obtenir une structure cristallographique à haute résolution de l'enzyme complète (Bushnell et Kornberg, 2003; Armache *et al.*, 2003; Armache *et al.*, 2004). Dans un premier cas, l'ARN pol II a été purifiée à partir d'une lignée cellulaire dans laquelle Rpb4 portait une étiquette TAP (*Tandem Affinity Purification*), permettant sa purification (Bushnell et Kornberg, 2003). Dans le second cas, l'ARN pol II à 10 sous-unités a été purifiée à partir d'une lignée $\Delta rpb4$, puis incubée avec Rpb4 et Rpb7 recombinants, exprimés et purifiés en bactérie (Armache *et al.*, 2003; Armache *et al.*, 2004).

Les deux structures publiées ont révélé sensiblement les mêmes choses. Rpb4 et Rpb7 s'attachent à l'ARN pol II à la base de la pince via un domaine de Rpb7 (figure 6). La surface d'interaction est restreinte, expliquant la dissociabilité du complexe. Rpb4 et Rpb7 se retrouvent ainsi près du canal de sortie de l'ARNm, ce qui est consistant avec un rôle possible de Rpb4 dans l'exportation du messager au cytoplasme (Farago *et al.*, 2003). Le sous-complexe est aussi localisé près du CTD de Rpb1, où il pourrait participer au recrutement de Fcp1 (Kimura *et al.*, 2002). Dans les deux structures de l'ARN pol II à 12 sous-unités, la pince se trouve dans une conformation fermée, similaire à la conformation retrouvée au niveau de l'enzyme à 10 sous-unités en élongation (Gnatt *et al.*, 2001). Dans cette conformation, le chargement direct de l'ADN dans le canal

principal est impossible. Ceci invalide le modèle mentionné précédemment, où l'ARN pol II lie l'ADN en conformation ouverte, puis la pince se referme sur les acides nucléiques pour l'élongation. Un premier modèle alternatif propose que Rpb4/7 se dissocieraient du complexe pour permettre l'ouverture de la pince et la liaison à l'ADN (Armache *et al.*, 2003). Suite à la formation de la bulle de transcription, le sous-complexe pourrait se réassocier, induisant la fermeture de la pince. Un deuxième modèle suggère que l'ADN serait d'abord chargé sur la pince fermée (Armache *et al.*, 2003; Bushnell et Kornberg, 2003). Suite à l'ouverture du promoteur par TFIIF, l'ADN descendrait dans le canal principal où il pourrait interagir avec les *switches*. Plus précisément, il a été proposé que les *switches* 1 à 3 forment une région d'interaction pour le brin matrice d'ADN dans la région de l'hybride, lorsque ce brin descend dans le canal (Armache *et al.*, 2003).

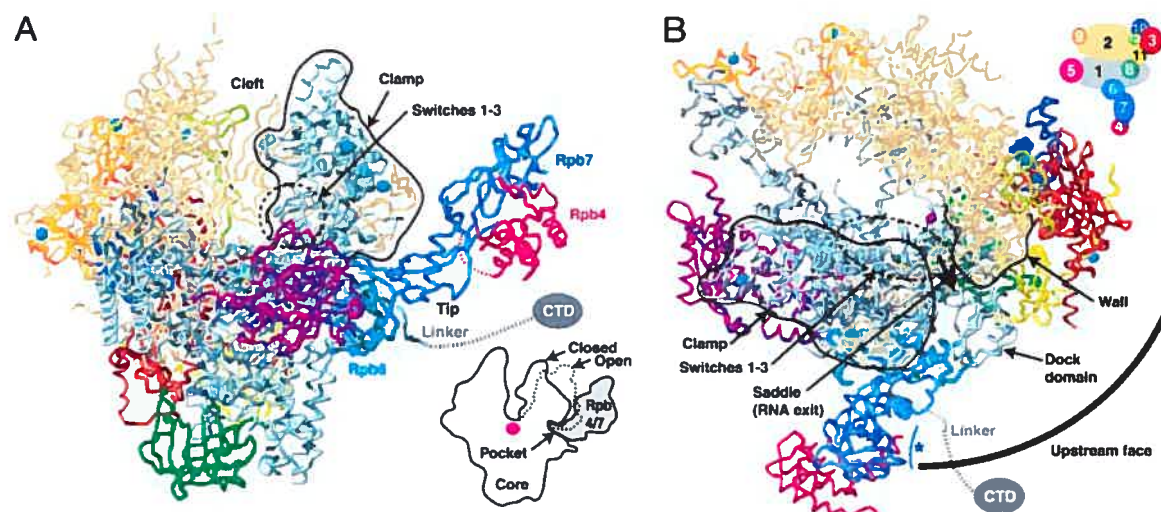


Figure 6 : Structure de l'ARN polymérase II à 12 sous-unités. A) Vue de l'ARN pol II à 12 sous-unités montrant la pince qui ferme le canal principal en présence de Rpb4 et Rpb7. Le schéma à droite de la figure montre les positions ouverte (tel que retrouvée dans la structure à 10 sous-unités) et fermée de la pince. Les sous-unités sont colorées selon le code indiqué sur le schéma en B. B) Vue de l'ARN pol II montrant la localisation des sous-unités Rpb4 et Rpb7 au bas de la pince, près du canal de sortie de l'ARNm. Figure tirée de Armache *et al.*, 2003.

1.3.2.3 ARN polymérase II-TFIIS

Une co-cristallographie de l'ARN pol II et du facteur TFIIS a été réalisée, permettant de mieux comprendre le rôle de celui-ci dans le clivage et l'élongation transcriptionnelle (Kettenberger *et al.*, 2003). Ce co-cristal a été produit en incubant le cristal de l'ARN pol II à 12 sous-unités avec TFIIS. La structure résultante montre que TFIIS est localisé à la surface de l'ARN pol II au niveau de l'entonnoir, où se retrouve l'extrémité 3' du transcrit lors des pauses. Une boucle de TFIIS, ayant un motif ruban de zinc (*zinc ribbon*), entre complètement dans le site catalytique par le pore (figure 7A). À l'extrémité de cette boucle se retrouvent deux acides aminés conservés essentiels à l'activité de TFIIS (Jeon *et al.*, 1994), qui sont localisés près du site de liaison du deuxième ion métallique, le métal B, et qui pourraient participer à sa coordination. Ainsi, TFIIS stabiliserait cet ion dans le site catalytique et stimulerait l'activité de clivage de l'ARN pol II, favorisant la sortie des sites d'arrêts et conséquemment l'élongation. TFIIS induit aussi des changements de conformation importants dans la structure de la polymérase, particulièrement au niveau du *bridge helix*, des *switches* 1 et 2 et du *fork loop* 2. Ces changements au cœur du site catalytique, dans des domaines impliqués dans la liaison à l'hybride ADN-ARN, pourraient servir à repositionner les acides nucléiques pour la réaction de clivage. En effet, une récente structure de TFIIS avec l'ARN pol II en élongation a révélé que celui-ci induit un réalignement de l'extrémité 3' de l'ARNm (Kettenberger *et al.*, 2004). Ce réalignement pourrait positionner le transcrit pour le clivage par la polymérase lors des arrêts. Des images de microscopie électronique de l'ARN pol bactérienne avec le facteur de clivage Gre B, ont montré une grande similarité avec le complexe ARN pol II-TFIIS. Gre B, qui est un homologue fonctionnel de TFIIS, se lie à la polymérase bactérienne via le pore et insère une boucle acide dans le site catalytique (figure 7B).

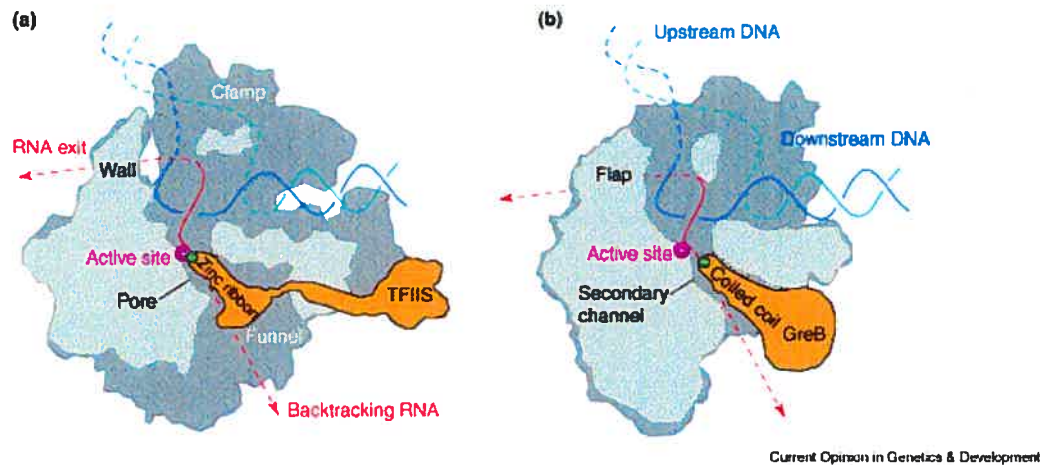


Figure 7 : Localisation des facteurs de clivage TFIIS et Gre B sur l'ARN polymérase. A) L'ARN pol II a été scindée en deux pour montrer l'intérieur du site catalytique et la localisation de TFIIS par rapport à celui-ci, telle que révélée par la cristallographie ARN pol II-TFIIS (Kettenberger *et al.*, 2003). Le brin matrice de l'ADN est indiqué en bleu; le brin transcrit, en vert et l'ARNm, en rouge. B) Localisation de Gre B par rapport à l'ARN polymérase bactérienne, telle que déterminée par les images de microscopie électronique (Opalka *et al.*, 2003). Le code de couleur est le même qu'en A. Figure tirée de Cramer, 2004.

1.3.2.4 ARN polymérase II-TFIIB

Un complexe formé de TFIIB lié à l'ARN pol II a été cristallographié (figure 8) (Bushnell *et al.*, 2004). Dans cette structure, TFIIB est localisé sur le domaine *dock* de la polymérase. Un module de TFIIB appelé *finger domain* entre par le canal de sortie de l'ARNm, jusqu'au site catalytique où il se retrouve près du site de liaison de l'hybride, expliquant le rôle de TFIIB dans le choix du site d'initiation. Plusieurs résidus du *finger* ont d'ailleurs été identifiés par mutagenèse comme étant importants pour la sélection de ce site (Pardee *et al.*, 1998; Hawkes et Roberts, 1999; Bell et Jackson, 2000). Comme TFIIB occupe le canal de sortie de l'ARNm, il a été proposé que lorsque ce dernier atteint plus de 9 nucléotides (la longueur de l'hybride), il participe au relâchement de TFIIB, nécessaire au dégagement du promoteur par l'ARN pol II (Cramer 2004).

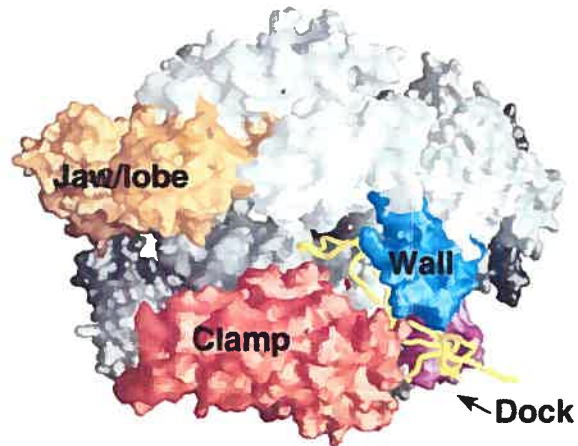


Figure 8 : Le complexe ARN polymérase II-TFIIB. La partie N-terminale de TFIIB (beige) est localisée sur la structure de l'ARN pol II, tel que déterminé par cristallographie (Bushnell *et al.*, 2004). Les principaux domaines de l'ARN pol II sont identifiés. Figure tirée de Hahn, 2004.

1.3.2.5 ARN polymérase II-TFIIF

Des images de microscopie électronique de l'ARN pol II liée à TFIIF ont permis de voir que la surface d'interaction entre ces deux protéines est très étendue (Chung *et al.*, 2003) (figure 9). Tfg1, l'homologue de RAP74 chez la levure, se lie à la pol II sur la pince et au niveau de Rpb4 et Rpb7. Tfg2, l'homologue de RAP30, fait plusieurs contacts au niveau du sillon central (*cleft*). Comme le positionnement de TFIIF est similaire à celui du facteur σ bactérien, il a été proposé que TFIIF serait un homologue structural de σ .

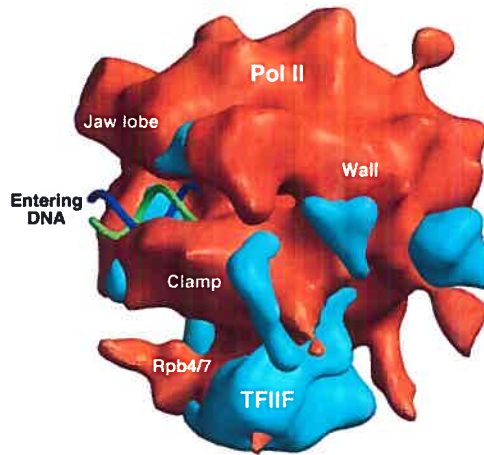


Figure 9 : Structure du complexe ARN polymérase II-TFIIF, déterminée par microscopie électronique. L'ARN pol II est colorée en orange et TFIIF, en bleu. Les principaux domaines de l'enzyme sont indiqués. L'ADN qui entre dans le site catalytique (ADN aval) a été modélisé sur la structure dérivée des images de microscopie électronique (Chung *et al.*, 2003). Figure tirée de Hahn, 2004.

1.3.3 Le mécanisme catalytique

Un mécanisme à deux ions métalliques a été proposé pour toutes les polymérases par Steitz en 1998 (Steitz, 1998). Dans ce mécanisme, un premier cation, le métal A, diminue l'affinité de l'oxygène en 3' pour l'hydrogène et facilite ainsi l'attaque du O⁻ sur le phosphate α en 5' du nucléotide arrivant. Le deuxième cation, appelé métal B, favorise le départ du pyrophosphate. Chez les ARN pol eucaryotes et procaryotes, le métal A est un ion Mg²⁺ coordonné par trois aspartates strictement conservés de Rpb1/ β' , contenus dans le motif NADFDGD (Cramer *et al.*, 2001; Gnatt *et al.*, 2001; Zhang *et al.*, 1999a). Le métal B, quant à lui, n'a pas été localisé de façon définitive dans les premières cristallographies des ARN pol procaryotes et eucaryotes (Cramer *et al.*, 2001; Gnatt *et al.*, 2001; Zhang *et al.*, 1999a). En effet, dans la cristallographie de l'ARN pol II de levure à 10 sous-unités, le métal B a été observé au voisinage de deux résidus acides conservés de Rpb2 (E836 et D837) et d'un résidu conservé de Rpb1 (D481) (Cramer *et al.*, 2001; Gnatt *et al.*, 2001). Cependant, il se retrouvait trop éloigné de ceux-ci pour permettre la coordination. De plus, comme le Mg²⁺-B était présent en quantité sous-

stochiométrique au sein de l'enzyme, il a été proposé que l'ion métallique B entre au site catalytique par le pore avec le nucléotide arrivant, qui lui fournit une coordination supplémentaire pour le stabiliser près des résidus E836 et D837 de Rpb2 et D481 de Rpb1 (Cramer *et al.*, 2001). Suite à la catalyse, le pyrophosphate relâché quitte le site actif par le pore, entraînant avec lui le métal B.

Un mécanisme à deux ions métalliques a aussi été proposé pour l'activité de clivage de l'ARN pol II lors des arrêts transcriptionnels. Cette activité requiert, en effet, la présence de cations divalents (Reines *et al.*, 1992). De plus, la résolution de structures de l'ARN pol II liée à TFIIS et de l'ARN pol bactérienne liée à Gre B suggère que les deux facteurs de clivage fonctionnent en fournissant des résidus acides pour la coordination d'un second ion métallique au site catalytique (Kettenberger *et al.*, 2003; Sosunov *et al.*, 2003; Sosunova *et al.*, 2003; Opalka *et al.*, 2003).

Récemment une nouvelle structure de l'ARN pol II de levure sur une matrice d'ADN et d'ARN a permis la localisation du second ion métallique (Westover *et al.*, 2004b). Dans cette structure, l'ARN de 9 nucléotides possède un terminateur de chaîne en 3'. En incubant l'ARN pol II avec cette matrice en présence d'un nucléotide complémentaire au dNTP placé au site +1, on obtient une structure où le NTP arrivant est positionné au site A (*addition site*), près pour la formation du lien phosphodiester et l'incorporation dans la chaîne d'ARN. Dans cet état, le Mg^{2+} -B est coordonné par les phosphates β et γ du NTP, les résidus D481 et D483 de Rpb1 et D837 de Rpb2. Lorsque l'ARN pol II a été incubée avec l'un ou l'autre des trois nucléotides non-complémentaires au site +1, il a été observé que celui-ci occupait un autre site dans une orientation très différente, la base et le sucre étant dirigés vers le pore. Ce deuxième site a été appelé le site E (*entry site*) et il a été proposé qu'il servait à accueillir le NTP arrivant, avant que celui-ci ne subisse une rotation pour aller se positionner au site A (Westover *et al.*, 2004b). Dans ce modèle, le NTP entre par le pore et suite à la formation du lien phosphodiester, le *bridge helix* se courbe pour permettre la translocation des acides nucléiques, puis reprend sa position initiale, libérant ainsi le site A pour l'arrivée du

prochain nucléotide (Gnatt *et al.*, 2001; Cramer *et al.*, 2001; Westover *et al.*, 2004a; Westover *et al.*, 2004b) (figure 10).

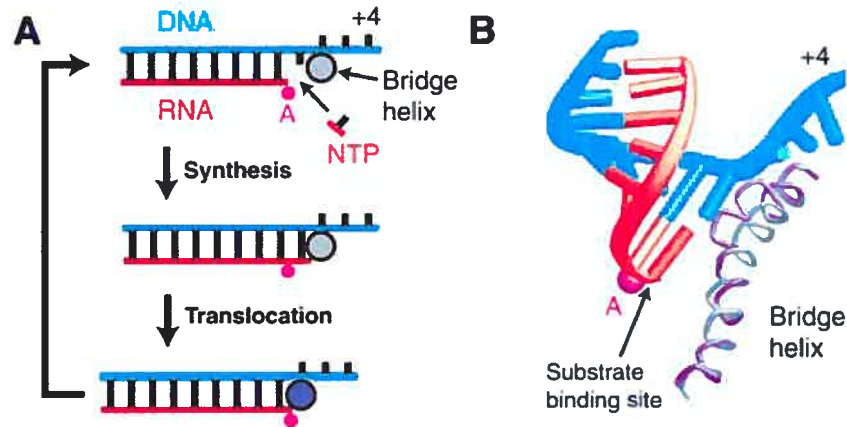


Figure 10 : Modèle proposé par Kornberg pour expliquer le mécanisme de translocation des acides nucléiques durant la polymérisation. A) Représentation schématique du cycle d'addition des nucléotides. Lorsque le *bridge helix* est en conformation droite (grise), le site d'addition (A) est libre et peut accueillir le NTP arrivant au site catalytique. Suite à la synthèse du lien phosphodiester, le *bridge helix* se courbe vers les acides nucléiques (mauve) et induit la translocation d'une position. En revenant à la position droite, le site A est de nouveau libéré pour un autre cycle. B) Superposition du *bridge helix* en conformation courbée, tel que retrouvé dans la structure de la pol bactérienne (mauve) (Zhang *et al.*, 1999a), sur le *bridge helix* de l'ARN pol II (gris) en conformation droite (Gnatt *et al.*, 2001; Cramer *et al.*, 2001). Figure tirée de Gnatt *et al.*, 2001.

Cependant un autre modèle a été proposé pour expliquer l'arrivée des nucléotides au site catalytique et la translocation des acides nucléiques durant la polymérisation. Ce modèle, le *NTP-driven translocation model*, basé sur des études cinétiques, suggère qu'il existe un pré-appariement du NTP arrivant avec le dNTP à $i+2$ (Nedialkov *et al.*, 2003) (figure 11). Le pré-appariement induit le mouvement du *bridge helix*, qui permet la translocation du nucléotide jusqu'au site A pour la catalyse. La sélection du nucléotide complémentaire se ferait donc avant l'entrée au site catalytique et seul un pré-appariement correct pourrait induire la translocation. Ce modèle implique aussi que l'entrée des NTP ne se fait pas par le pore, mais bien par le canal principal. Le *bridge helix* constitue, en effet, un obstacle au passage d'un NTP arrivant par le pore vers le

dNTP à $i+2$, localisé dans le canal principal. Le pore servirait à la sortie du pyrophosphate-métal B relâché.

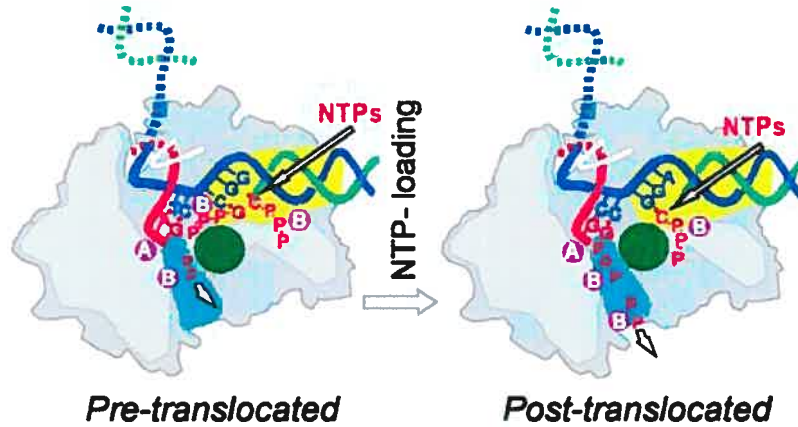


Figure 11 : Le modèle de translocation des acides nucléiques dirigée par le NTP arrivant proposé par Burton. Le NTP (rouge) arrive au site actif par le canal principal (jaune) couplé à l'ion Mg^{2+} -B (mauve) et s'apparie à l'ADN à la position $i+2$. Un appariement correct induit le mouvement du *bridge helix* (vert) qui conduit le NTP au site d'addition pour la catalyse. Suite à la synthèse du lien phosphodiester, le *bridge helix* reprend sa conformation initiale et le pyrophosphate libéré quitte le site actif par le pore avec l'ion Mg^{2+} -B. Figure tirée de Zhang et Burton, 2004.

Afin de discriminer entre les deux modèles, une cristallographie semblable à celle publiée par Westover et col. (Westover *et al.*, 2004b), pourrait être réalisée. Dans cette structure, l'ARN pol II liée à la matrice d'ADN-ARN décrite précédemment devrait être incubée avec les 4 NTP ayant des terminateurs de chaîne. De cette façon le NTP complémentaire au dNTP $i+1$ irait se placer au site A et on pourrait voir si les autres nucléotides se positionnent à $i+2$, $i+3$, et les sites suivant, ce qui irait dans le sens du modèle *NTP-driven translocation* de Burton (Nedialkov *et al.*, 2003). Au contraire, s'ils occupent seulement le site E, ceci irait dans le sens d'un modèle proposé par Kornberg où le NTP entre par le pore et subit une rotation du site E au site A (Westover *et al.*, 2004b).

La cristallographie de l'ARN pol II liée à la drogue alpha-amanitine, qui inhibe l'initiation et l'élongation de la transcription (Rudd et Luse, 1996; Chafin *et al.*, 1995) a permis de montrer que celle-ci se lie dans le site catalytique de l'enzyme, très près du

bridge helix, sans toutefois bloquer le site A (Bushnell *et al.*, 2002). Il a été proposé que l'alpha-amanitine nuit au mouvement du *bridge helix*, bloquant ainsi la translocation. Des études cinétiques ont par la suite confirmé cette hypothèse (Gong *et al.*, 2004).

Les ARN pol à une sous-unité, comme celle du phage T7, ne sont pas homologues en terme de séquence aux ARN pol à sous-unités multiples procaryotes et eucaryotes. Elles présentent quand même quelques éléments de structures similaires à ces dernières (revu par Landick, 2004). Par exemple, l'ARN pol T7 possède un sillon central, qui accueille les acides nucléiques. Elle utilise aussi un mécanisme à deux ions métalliques pour la catalyse. Cependant chez l'ARN pol T7, les deux ions métalliques entrent avec le NTP arrivant et le métal A participe à la discrimination entre NTP et dNTP (Temiakov *et al.*, 2004). L'entrée des NTP se fait dans un axe perpendiculaire à l'hybride (voir figure 3A), ce qui se rapproche du modèle d'entrée des NTP par le canal principal et non par le pore, proposé pour les ARN pol à sous-unités multiples (Nedialkov *et al.*, 2003).

L'ARN pol T7 possède une hélice 'O', dans le site catalytique, qui aurait des fonctions dans la translocation des acides nucléiques semblables à celles du *bridge helix*. L'hélice 'O' contient une tyrosine (Tyr 639), qui joue un rôle crucial dans la translocation en passant d'une conformation ouverte à une conformation fermée durant le processus. Le mécanisme se passe en 4 étapes (figure 13) (Yin et Steitz, 2004; Temiakov *et al.*, 2004; Landick, 2004). Tout d'abord le NTP arrivant se positionne dans un site de pré-insertion où il s'apparie avec le dNTP au site i+1 de la matrice d'ADN. À cette position, la Tyr 639 interagit avec le métal A qui lie le 2'OH du nucléotide arrivant, permettant l'exclusion des dNTP et la sélection d'un NTP complémentaire. La Tyr 639 se déplace ensuite pour prendre une conformation dite fermée, permettant l'entrée du NTP au site S (site substrat, qui correspond au site A dans les modèles discutés précédemment). Suite à la catalyse et à la libération du pyrophosphate, l'hélice reprend sa conformation ouverte, ce qui entraîne le déplacement du NTP incorporé au site P (site produit) et la libération du site S pour l'arrivée du nucléotide suivant. Ce modèle ressemble à celui proposé par Burton et collaborateurs pour l'ARN pol II (Nedialkov *et al.*, 2003) dans le sens où il existe un pré-appariement du NTP arrivant avant l'entrée au site où se fera la catalyse

(site A ou S). Dans le modèle de Burton et collaborateurs, ainsi que dans celui proposé pour l'ARN pol T7, le pré-appariement sert de mode de sélection du substrat. Cependant, dans le cas de l'ARN pol II, c'est le pré-appariement correct qui induit la courbure du *bridge helix* et le déplacement du NTP au site A pour la catalyse. Au contraire, pour la pol T7, le pré-appariement correct induit le changement de conformation de l'hélice 'O', qui permet l'entrée au site A, mais c'est la catalyse qui conduit à la translocation des acides nucléiques, lorsque l'hélice 'O' reprend sa position initiale ouverte, libérant ainsi le site S.

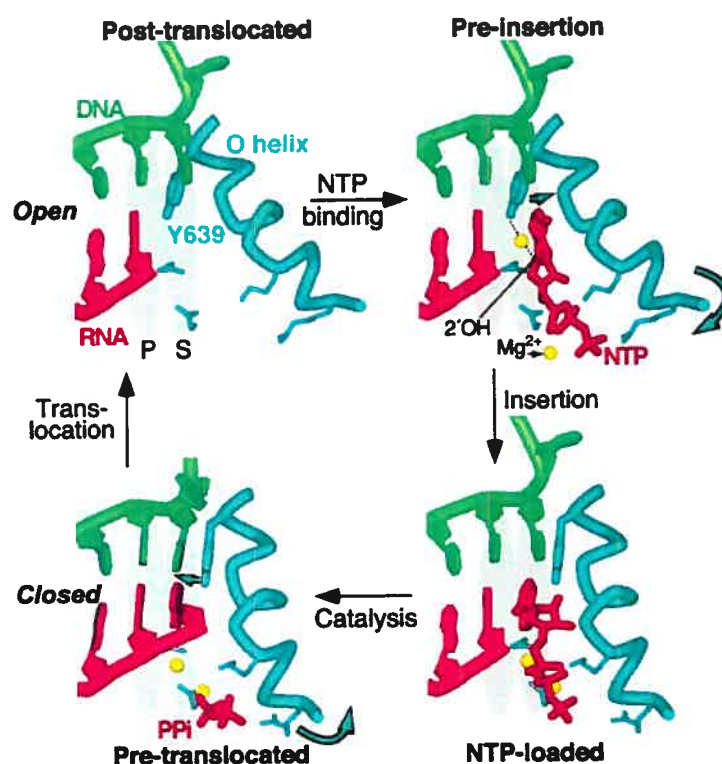


Figure 12 : La translocation des acides nucléiques chez les ARN polymérases à une sous-unité. Les quatre étapes du processus de translocation sont montrées. L'hélice 'O' est colorée en turquoise; l'ADN matrice, en vert; l'ARNm et le NTP arrivant, en rouge; les ions métalliques, en jaune. La position de la tyrosine 639 (Y639) est indiquée. Figure tirée de Landick, 2004.

1.4 Objectifs de recherche

Le premier objectif de cette thèse était d'étudier les rôles de certains FGT de l'ARN pol II dans le processus d'initiation de la transcription, en relation avec l'organisation du complexe transcriptionnel. Plus précisément, la localisation du facteur TFIIA dans ce complexe suggérait des interactions possibles avec plusieurs autres FGT. À l'aide de la technique de chromatographie d'affinité, nous avons pu déterminer que le facteur TFIIA interagit avec la sous-unité RAP74 de TFIIF et les sous-unités α et β de TFIIIE. Le domaine de TFIIF impliqué dans cette interaction a été identifié et nous montrons que TFIIA stimule l'initiation de la transcription via son interaction avec RAP74 et TFIIIE β . Nous avons aussi déterminé que TFIIIE α , qui fait contact avec l'ADN dans la région de la bulle transcriptionnelle, est impliqué dans l'ouverture du promoteur.

Le second objectif de cette thèse était d'étudier le mécanisme catalytique de l'enzyme ARN pol II en regard avec sa structure. Particulièrement, suite à la publication de structures cristallographiques à haute résolution de l'ARN pol II de levure, il a été proposé que certains domaines et motifs de l'enzyme joueraient des rôles critiques dans le mécanisme transcriptionnel. Nous avons donc entrepris une étude de mutagenèse dirigée de l'ARN pol II humaine, qui ciblait ces structures, suivie d'une caractérisation biochimique des mutants obtenus. Pour ce faire nous avons mis au point un système permettant l'expression dans les cellules humaines des sous-unités Rpb1 et Rpb2 sauvages ou mutantes portant une étiquette TAP et la purification des ARN pol II assemblées dans ces cellules. Les ARN pol II sauvages ou mutantes purifiées à l'aide de cette méthode ont ensuite pu être testées dans différents essais fonctionnels. Ce travail nous a permis de mieux comprendre le rôle de plusieurs domaines de Rpb1 et Rpb2 dans le processus de transcription, notamment au niveau des boucles *fork loop 1*, *switch 3*, *rudder* et *zipper*. De plus, il nous a permis d'identifier l'acide aminé E791 comme étant impliqué dans la liaison du métal B pour l'activité de clivage et de polymérisation de l'ARN pol II.

CHAPITRE 2

INTERACTIONS STRUCTURALES ET FONCTIONNELLES DU FACTEUR (TF) IIA AVEC TFIIE ET TFIIF DANS L'INITIATION DE LA TRANSCRIPTION PAR L'ARN POLYMÉRASE II

2.1 Préambule

Ma contribution à l'article présenté au chapitre 2 est la suivante. J'ai identifié par chromatographie d'affinité les sous-unités de TFIIA responsables de l'interaction avec RAP74, TFIIE56 (α) et TFIIE34 (β) (figure 3). J'ai purifié tous les mutants de RAP74 étudiés et je les ai utilisés en chromatographie d'affinité afin de déterminer le domaine d'interaction à TFIIA (figure 4). Finalement, j'ai réalisé les essais d'initiation (figure 6B et C) et j'ai rédigé le manuscrit avec Dr Benoit Coulombe.

Diane Forget a effectué les essais de photo-pontages protéine-ADN (figure 1) et de transcription (figure 6A). Andrés Rojas a déterminé les interactions de TFIIA avec les différents facteurs généraux de transcription (figure 2) et Yanie Porlier a fait l'expérience de retard sur gel (figure 5).

2.2 Article

Langelier, M.-F., Forget, D., Rojas, A., Porlier, Y., Burton, Z.F., Coulombe B. (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(42):38652-7.

Structural and Functional Interactions of Transcription Factor (TF) IIA with TFIIE and TFIIIF in Transcription Initiation by RNA Polymerase II*

Received for publication, July 9, 2001
Published, JBC Papers in Press, August 16, 2001, DOI 10.1074/jbc.M106422200

Marie-France Langelier^{†§¶}, Diane Forget[‡], Andrés Rojas[‡], Yanie Porlier^{‡§||}, Zachary F. Burton^{**}, and Benoit Coulombe^{‡§¶‡‡}

From the [‡]Laboratory of Gene Transcription, Institut de Recherches Cliniques de Montréal, 110 Avenue des Pins Ouest, Montréal, Québec H2W 1R7, Canada, the [§]Département de biochimie, Université de Montréal, Canada, and the ^{**}Department of Biochemistry, Michigan State University, East Lansing, Michigan 48824

A topological model for transcription initiation by RNA polymerase II (RNAPII) has recently been proposed. This model stipulates that wrapping of the promoter DNA around RNAPII and the general initiation factors TBP, TFIIB, TFIIE, TFIIIF and TFIIH induces a torsional strain in the DNA double helix that facilitates strand separation and open complex formation. In this report, we show that TFIIA, a factor previously shown to both stimulate basal transcription and have co-activator functions, is located near the cross-point of the DNA loop where it can interact with TBP, TFIIE56, TFIIE34, and the RNAPII-associated protein (RAP) 74. In addition, we demonstrate that TFIIA can stimulate basal transcription by stimulating the functions of both TFIIE34 and RAP74 during the initiation step of the transcription reaction. These results provide novel insights into mechanisms of TFIIA function.

Initiation of transcription by RNA polymerase II (RNAPII)¹ proceeds through the formation of a preinitiation complex containing RNAPII and the general transcription factors (TF) TBP (the TATA box-binding protein of TFIID), TFIIB, TFIIE, TFIIIF, and TFIIH on promoter DNA (reviewed in Refs. 1 and 2). The first step in preinitiation complex assembly is the recognition of the TATA element of the promoter by TBP. The binding of TBP to the TATA box induces a DNA bend of ~90° (3, 4). TFIIB can associate with the TBP-promoter complex (5). Mammalian TFIIIF, which is composed of the subunits RAP74 and RAP30, directly binds to RNAPII and has been shown to participate in recruitment of the enzyme to the preinitiation complex (6, 7). TFIIE, which is also composed of two subunits called TFIIE56 and TFIIE34, is involved in the melting of promoter DNA at the

transcription initiation site through a mechanism that is ATP-independent (8, 9). Finally, TFIIH, which has kinase and helicase activities, mediates the ATP-dependent melting of the promoter DNA in the region of the initiation site and is involved in the transition between the initiation and elongation states of the complex (10–15).

Recent results describing both the structure of the basal transcription machinery and the topological organization of the preinitiation complex have considerably improved our understanding of transcription initiation mechanisms. Determination of the atomic structure of yeast RNAPII at 2.8 angstroms resolution and that of elongating yeast RNAPII at 3.3 angstroms by Kornberg and co-workers (16, 17) has revealed key features of both the interaction between the enzyme and template DNA and the basis of its catalytic activity. Analysis of the molecular organization of the preinitiation complex using site-specific protein-DNA photo-cross-linking has provided insights on the topology of the preinitiation complex containing RNAPII and the general transcription factors (18–26). Recently, we have proposed a topological model, the DNA wrapping model, which describes transcription initiation by RNAPII (23, 26, 27). This model accounts for our photo-cross-linking data and several additional data obtained in various laboratories. The DNA wrapping model stipulates that a role for the general transcription factors is to help in the wrapping of the promoter DNA in the preinitiation complex in such a way that a torsional strain is progressively developed upstream of the transcription start site and results in the partial unwinding of the DNA helix. This region of unwound DNA is used as a substrate by the single-stranded DNA helicases of TFIIH that catalyze open complex formation (26).

First isolated as a general transcription factor, TFIIA has a rather controversial role in transcription initiation. Human TFIIA is composed of three subunits: α (35 kDa), β (19 kDa), and γ (12 kDa) (28–31). The α and β subunits are encoded by the same gene and are produced by posttranslational cleavage of a precursor (29, 30). In yeast, TFIIA is composed of only two subunits encoded by two different genes, *TOA1* and *TOA2* (32). The N-terminal part of the polypeptide produced from the *TOA1* gene is homologous to the human α subunit, and the C-terminal part is homologous to the β subunit (29, 30). *TOA2* encodes a polypeptide homologous to γ (28, 31). The posttranslational cleavage of human α/β has been demonstrated to be non-essential because wild type activity can be recovered with uncleaved recombinant α/β and γ renatured together (33). TFIIA is not essential for basal transcription *in vitro*, but it has been shown to stimulate basal transcription in a variety of systems (28, 33–37). TFIIA binds TBP and increases the affinity of TBP for the TATA box (36–38). TFIIA can displace

* This work was supported by grants from the Canadian Institutes for Health Research and the Cancer Research Society Inc. (to B. C.) and the National Institutes of Health (to Z. F. B.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

¶ Holds a studentship from the Natural Sciences and Engineering Research Council of Canada.

|| Holds a studentship from the Natural Sciences and Engineering Research Council of Canada.

‡‡ Senior scholar from the Fonds de recherche en santé du Québec. To whom correspondence should be addressed. Tel.: 514-987-5662; Fax: 514-987-5663; [REDACTED]

¹ The abbreviations used are: RNAPII, RNA polymerase II; TF, transcription factor; TBP, TATA box-binding protein; RAP, RNA polymerase II-associated protein; nTFIIA, natural human TFIIA; rTFIIA, recombinant human TFIIA; PAGE, polyacrylamide gel electrophoresis; BSA, bovine serum albumin; AdMLP, adenovirus major late promoter; nt, nucleotide(s).

certain repressors, including Dr1-DRAP1/NC2, topoisomerase 1, HMG1, and Mot-1, from the TFIID complex, indicating that TFIIA is involved in antirepression (39–43). Human TFIIA also plays a role in activated transcription, being required for the functioning of some activators (28, 29, 31, 33, 34, 44–47). For example, TFIIA binds to the activator Zta and mediates its stimulation of TFIID binding to the TATA box (28). Similarly, TFIIA enhances the activation of transcription by the activators Sp1, VP16, and NTF1 (34). The activators VP16 and Zta, which bind TFIIA, stimulate the assembly of a TFIIA-TFIID-promoter complex, consistent with the roles *in vivo* of these factors in activated transcription (46). The function of TFIIA in transcriptional antirepression and activation have been separated and is associated with distinct subunits of the factor (35). Subunits β and γ are essential for antirepression, whereas α is not. Conversely all three subunits are required for activation.

Previous photo-cross-linking experiments performed with a TBP-TFIIA-promoter complex have revealed that TFIIA makes promoter contacts both in the region of the TATA box and upstream of it (18, 20). We have now determined the position of TFIIA in a preinitiation complex assembled in the presence of TBP, TFIIB, TFIIE, TFIIF, and RNAPII. Our results indicate that TFIIA makes promoter contacts not only on the TATA box and upstream of it in the -40 region, as is observed in the TBP-TFIIA-promoter complex, but also in the $+26$ region. Given the two extreme promoter positions approached by TFIIA and the small size of TFIIA (67 kDa), these results suggest that TFIIA is located near the cross-point of the wrapped DNA structure, where it can simultaneously contact nucleotides -40 and $+26$. TFIIF, TFIIE, and RNAPII also cross-link to the -40 and $+26$ positions (23, 26), suggesting that TFIIA may directly interact with these factors. We report here that TFIIA directly interacts with RAP74, TFIIE56, and TFIIE34 in addition to the previously determined interaction with TBP. Furthermore, we use an abortive initiation assay to provide evidence that the stimulatory effect of TFIIA on basal transcription is exerted through a stimulation of the activity of RAP74 and TFIIE34 at the initiation stage of transcript formation.

EXPERIMENTAL PROCEDURES

Protein Factors—Recombinant yeast TBP (48), human TFIIB (49), human RAP30 (50), human RAP74 (wild type and deletion mutants) (50), human TFIIE56 and TFIIE34 (51–53), and calf thymus RNAPII (54) were prepared as previously described. Natural human TFIIA (nTFIIA) was partly purified using protein-affinity chromatography with immobilized TBP as we described (36). Recombinant human TFIIA (rTFIIA) was produced from uncleaved α/β and γ subunits that carried histidine tags (28). The two polypeptides were independently purified on Ni-NTA agarose columns (Qiagen) under denaturing conditions and were either used individually as TFIIA α/β and TFIIA γ or renatured together to produce rTFIIA.

Protein-DNA Photo-cross-linking—The synthesis of the photoreactive nucleotide N_3R -dUMP, the preparation of the probes, and the conditions for binding reactions were as described (55). Two photoprobes containing the modified nucleotide at positions $-39/-40$ and $+26$ were used. For each probe, the concentration of poly(dI-dC) in the binding reactions was optimized to favor specific over nonspecific binding. A typical reaction with all the factors contained 200 ng each of TBP, TFIIB, RAP30, RAP74, TFIIE56, TFIIE34, rTFIIA, and RNAPII. UV irradiation, nuclease treatment, and SDS-PAGE analysis of radio-labeled photo-cross-linking products were performed as described previously (55).

Protein-Protein Interactions—Protein-protein interactions were analyzed essentially as we previously described (22). RAP74 wt, RAP74 deletion mutants, RAP30, TFIIE34, TFIIE56, RNAPII, TFIIB, TBP, and bovine serum albumin (BSA) were immobilized on Affi-gel 10 (Bio-Rad) at a concentration of 1 to 5 mg/ml resin. Microcolumns were made with $\sim 20 \mu\text{l}$ of this resin. A volume of $50 \mu\text{l}$ of nTFIIA (Fig. 2), recombinant TFIIA α/β , recombinant TFIIA γ (Fig. 3), or rTFIIA (Fig. 4), which contained 100 ng of nTFIIA or 200 ng of TFIIA α/β , TFIIA γ or

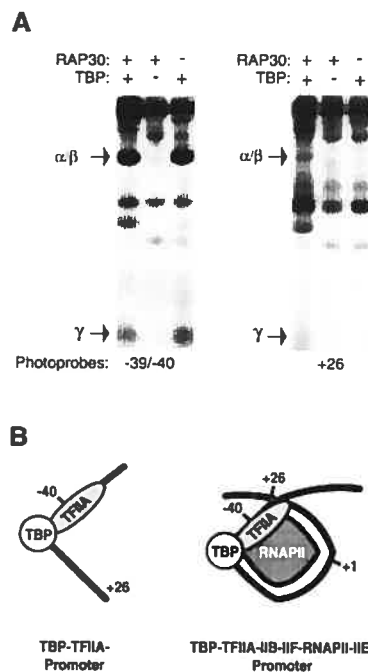


FIG. 1. Photo-cross-linking of rTFIIA on the AdMLP. *A*, photo-cross-linking experiments with photoprobes $-39/-40$ and $+26$ were performed in the presence of rTFIIA (α/β and γ), TBP, TFIIB, TFIIF (RAP74 and RAP30), TFIIE (p56 and p34), and RNAPII. The specificity of the cross-linking signals was assessed by comparing reactions performed with all the factors to ones performed in the absence of TBP. Under our reaction conditions, the absence of TBP has the same effect as the use of a photoprobe with a mutation in the TATA element (TATAAAA to TAGAGAA) (55). The position of TFIIA α/β and TFIIA γ are indicated. *B*, schematic representation of promoter contacts by TFIIA in the context of the TBP-TFIIA-promoter complex and the TBP-TFIIA-TFIIB-TFIIF-RNAPII-TFIIE-promoter complex in which the DNA adopts a wrapped structure. Positions $-39/-40$, $+26$ and $+1$ are indicated. Only TBP, TFIIA, and RNAPII are represented to simplify the diagram.

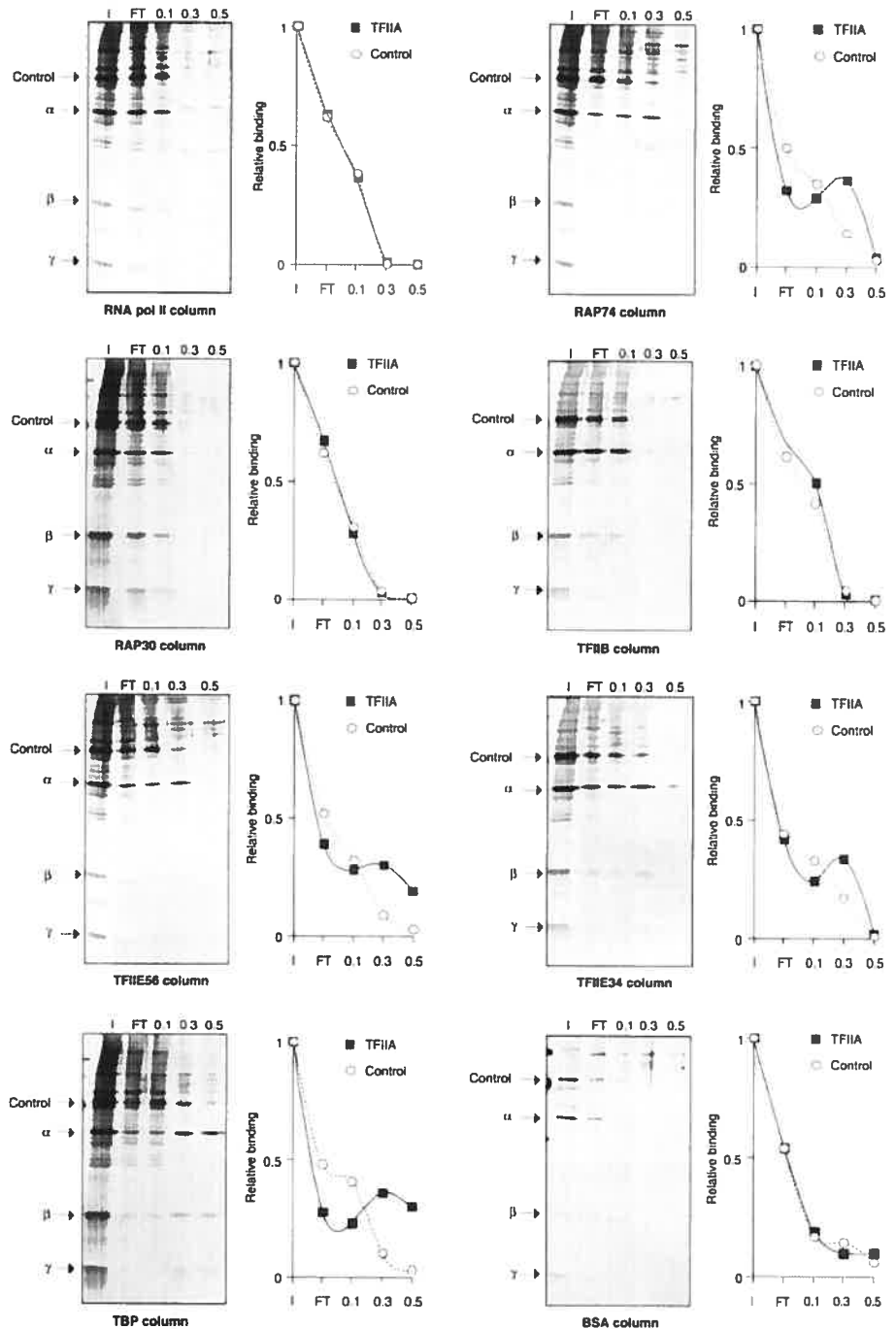
rTFIIA, was then loaded on the different columns. The flowthrough was collected and the columns successively eluted with $50 \mu\text{l}$ of ACB buffer (10 mM Hepes, pH 7.9, 0.2 mM EDTA, 20% glycerol and 1 mM dithiothreitol) containing 0.1 M, 0.3 M, and 0.5 M NaCl. An aliquot of the input, the flowthrough, and the various salt elutions were analyzed on SDS-PAGE and revealed by silver (Fig. 2 and Fig. 4) or zinc staining (Fig. 3). The intensity of the bands was evaluated using the UN-SCAN-IT software. It was considered that TFIIA, or one of its subunits, was binding to a particular column when the intensity of the 0.3 M salt band was higher than the intensity of both the 0.1 M band and the flowthrough band. In contrast, when the intensity of the 0.3 M band was lower than the 0.1 M or the flowthrough band, we considered that TFIIA was not binding to the column. The specificity of TFIIA binding was assessed by comparing the binding of the TFIIA subunits to the binding of a contaminant polypeptide of the nTFIIA preparation (Fig. 2).

Gel Mobility Shift Assay—Plasmid DNA containing the adenovirus major late promoter (AdMLP) was digested with the restriction enzymes *Bam*HI and *Dra*I, and the 110-base pair fragment containing the promoter was filled-in using the Klenow fragment of DNA polymerase in the presence of [α - ^{32}P]dGTP. Gel mobility shift assays were performed as described previously (56). Complexes were assembled using highly purified TBP (20 ng), TFIIA (120 ng), and, when indicated, RAP74-(1–517) (640 ng), (1–136) (160 ng), (1–75) (91 ng), (363–409) (58 ng), and (363–444) (107 ng).

Transcription Assay—Transcription assays were performed as described previously (57). TBP (120 ng), TFIIB (120 ng), RAP30 (120 ng), RAP74 (260 ng), TFIIE34 (160 ng), TFIIE56 (240 ng), RNAPII (660 ng), and various amounts of TFIIA were incubated with 500 ng of the supercoiled DNA template containing the AdMLP from nucleotides -50 to $+10$ fused to a G-less cassette. Under these conditions a 391-nt run-off transcript is produced.

Abortive Initiation Assay—Templates were prepared by annealing two 80-base pair DNA oligonucleotides carrying the strands of the

FIG. 2. Interactions of natural TFIIA with components of the basal transcription machinery. Protein-affinity chromatography was performed using microcolumns containing immobilized RNAPII, RAP74, RAP30, TFIIIE56, TFIIIE34, TFIIIB, TBP, and BSA. A volume of 50 μ l of nTFIIA (100 ng of α) was chromatographed through each column. The flowthrough was collected in each case. The columns were eluted with 50 μ l of buffer containing increasing amounts of NaCl (e.g. 0.1, 0.3, and 0.5 M). The fractions were analyzed using SDS-PAGE and compared with the input (I). The positions of the nTFIIA subunits and a contaminant polypeptide of the TFIIA preparation that served as negative control are indicated. In each case, a diagram showing the relative intensities of both TFIIA α and a contaminant band (negative control) in the various fractions is shown.



AdMLP from -45 to +35 as described (58). Typically, 12 ng of the double-stranded DNA template were incubated for 60 min at 30 °C with TBP (60 ng), TFIIIB (30 ng), RAP30 (30 ng), RNAPII (165 ng), and, when indicated, RAP74 (65 ng), TFIIIE34 (40 ng), TFIIIE56 (60 ng), and various amounts of TFIIA in 20 μ l of a reaction mixture containing 125 μ M ATP, 125 μ M CTP, 1.7 μ M UTP, 2.5 μ M $[\alpha\text{-}^{32}\text{P}]\text{UTP}$, 1.25 mM MgCl_2 , 0.5 mM EGTA, and 125 units/ml RNase inhibitor. The synthesis was stopped by incubating at 68 °C for 3 min, and the reaction mixture then cooled on ice for 5 min. Calf intestine alkaline phosphatase (8 U, CIP) and 2 μ l of 2 \times CIP buffer (500 mM Tris, pH 8.9, 1 mM EDTA) were added to 10 μ l of the transcription reaction, and the resulting solution was incubated for 20 min at 37 °C to reduce the background caused by free radiolabeled nucleotides. The CIP reaction was stopped by adding 3 μ l of loading buffer (50% glycerol, 200 mM EDTA, 0.05% bromphenol blue). Transcripts were analyzed on a 23% polyacrylamide denaturing gel containing 7 M urea and were quantitated using a PhosphorImager (Molecular Dynamics).

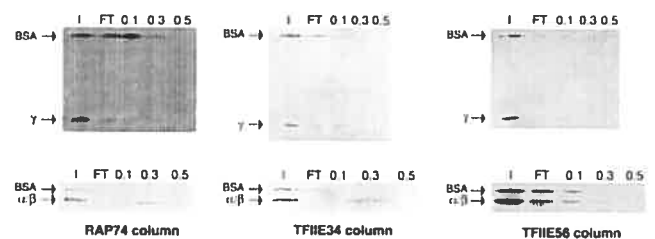


FIG. 3. Interactions of TFIIA α/β and TFIIA γ with RAP74, TFIIIE56, and TFIIIE34. Protein-affinity chromatography was performed using microcolumns containing immobilized RAP74, TFIIIE34, and TFIIIE56. A volume of 50 μ l containing 200 ng of TFIIA α/β and TFIIA γ was chromatographed through each column. Fractions were collected and analyzed as in Fig. 2. The positions of α/β , γ , and BSA (as an internal negative control) are indicated.

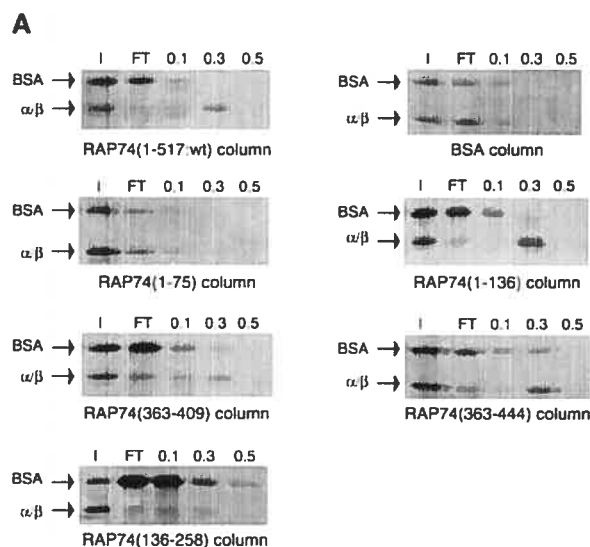


FIG. 4. Interactions of TFIIA α/β with RAP74 deletion mutants. *A*, protein affinity chromatography was performed using columns containing immobilized RAP74 fragments (wild type or deletions mutants). A volume of 50 μ l containing 200 ng of TFIIA α/β was chromatographed through each column. Fractions were collected and analyzed as in Fig. 2. The positions of TFIIA α/β and BSA (as an internal negative control) are indicated. *B*, summary of the interactions between the RAP74 deletion mutants and TFIIA. The various domains of RAP74 are indicated in the *top part*. The two TFIIA-binding domains of RAP74 are deduced from our analysis.

RESULTS AND DISCUSSION

TFIIA Is Located Near the Cross-point of the Wrapped DNA Structure in the Initiation Complex—We have previously shown that TFIIA assembled with TBP on the AdMLP (*e.g.* TBP-TFIIA-promoter complex) cross-linked to positions -31/-29, -25/-30, -39/-40, and -42 (18). We have now analyzed the position of TFIIA in a preinitiation complex composed of TBP, TFIIB, RAP74, RAP30, TFIIE56, TFIIE34, and RNAPII using site-specific protein-DNA photo-cross-linking. Both TFIIA α/β and TFIIA γ cross-linked in the region of the TATA box to photoprobe -31/-29 (data not shown), upstream of it to photoprobe -39/-40 (Fig. 1A, *left panel*) and downstream of it to photoprobe +26 (Fig. 1A, *right panel*). The cross-linking of TFIIA to positions -39/-40 required the presence of TBP but not that of RAP30 (Fig. 1A), TFIIB, or RNAPII (data not shown). The cross-linking of TFIIA to position +26, however, required the presence of TBP and RAP30 (Fig. 1A, *right panel*),

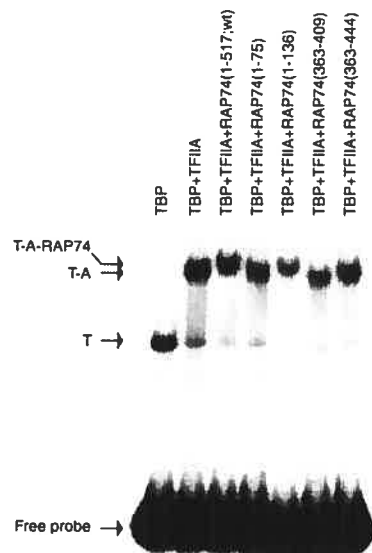


FIG. 5. Association of RAP74 with a TBP-TFIIA-promoter complex. Gel mobility shift assays were performed using a radiolabeled DNA fragment comprising the AdMLP in the presence of TBP alone, TBP, and TFIIA and TBP, TFIIA and various fragments of RAP74 (RAP74-(1-517-wt), RAP74-(1-75), RAP74-(1-136), RAP74-(363-409), and RAP74-(363-444)). The positions of the TBP (T), TBP-TFIIA (T-A), and TBP-TFIIA-RAP74 (T-A-RAP74) complexes and that of the free probe are indicated.

as well as that of TFIIB and RNAPII (data not shown). These results indicate that promoter contacts by TFIIA in the -39/-40 region do not necessitate assembly of a preinitiation complex containing TFIIB, TFIIF, and RNAPII (*e.g.* a TBP-TFIIA-promoter complex is sufficient), whereas the promoter contact by TFIIA in the +26 region requires the assembly of a preinitiation complex (*e.g.* a TBP-TFIIA-TFIIB-TFIIF-RNAPII-TFIIE-promoter complex) in which promoter DNA adopts a wrapped structure (see Fig. 1B for a schematic representation).

The cross-linking of TFIIA to promoter regions downstream of the transcription initiation site (+10 to +30) was not unexpected. According to the DNA wrapping model, the DNA helices upstream of TATA (-40/-60 region) and downstream the initiation site (+10/+30 region) are juxtaposed in space. Our results provide additional support for the notion that promoter DNA is wrapped in the initiation complex and indicate that TFIIA is localized near the cross-point of the wrapped DNA structure.

TFIIA Directly Interacts with TBP, RAP74, TFIIE56, and TFIIE34—In the context of a preinitiation complex assembled with TFIIA, TBP, TFIIB, TFIIE, TFIIF, and RNAPII, we obtained cross-linking of TFIIA to photoprobes that are also cross-linked by other components of the complex. More specifically TFIIE34, RAP74, and RAP30 cross-link to photoprobe +26, while Rpb2, TFIIE34, RAP74, and RAP30 cross-link to photoprobe -39/-40 (23, 26). These observations indicate that TFIIA is in close proximity to these factors in the preinitiation complex, suggesting that TFIIA could directly interact with TFIIE, TFIIF, and RNAPII. To test this hypothesis, nTFIIA was chromatographed over different affinity columns containing immobilized RAP74, RAP30, TFIIE56, TFIIE34, TFIIB, and RNAPII. Columns containing TBP and BSA were used as positive and negative controls, respectively because TBP has been shown to interact with TFIIA (36, 37). The flowthrough was collected in each case and the columns were successively eluted with buffer containing 0.1 M, 0.3 M, and 0.5 M NaCl. The various fractions were analyzed by SDS-PAGE. Natural TFIIA

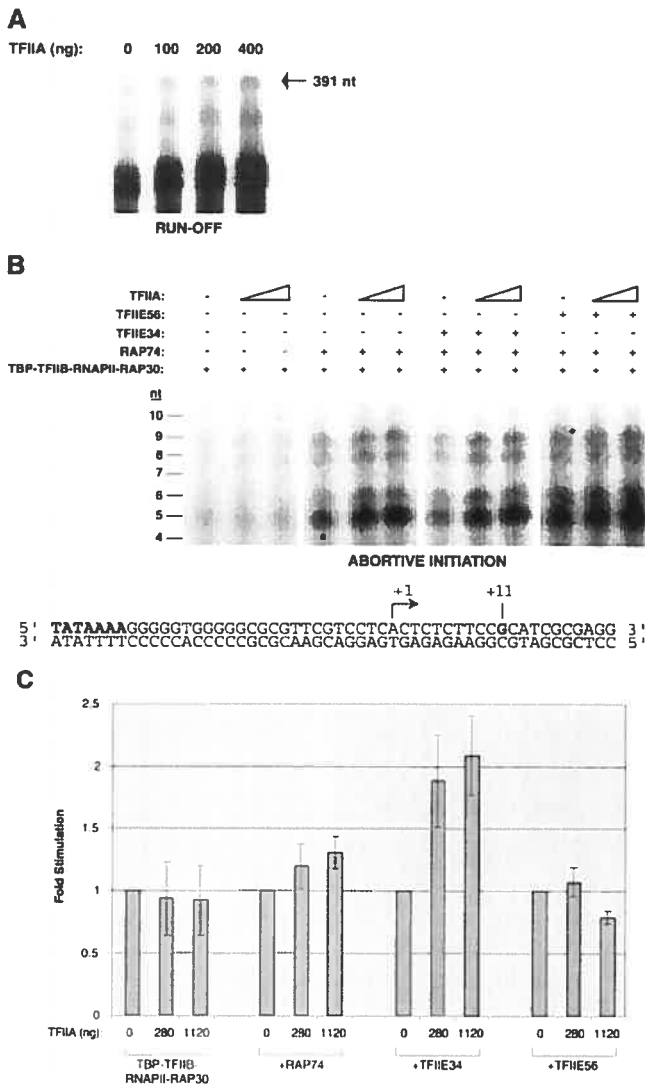


FIG. 6. Stimulation of basal transcription by TFIIA. *A*, run-off transcription assays were performed on a supercoiled template carrying the AdMLP using TBP, TFIIIB, TFIIIE, TFIIIF, and RNAPII in either the absence or the presence of increasing amounts of TFIIA (100, 200, and 400 ng). The position of the accurately initiated transcript (391 nt) is indicated. *B*, abortive initiation assays were performed on a synthetic double-stranded oligonucleotide carrying the AdMLP using TBP, TFIIIB, RAP30, and RNAPII in either the absence or the presence of RAP74 alone, RAP74 and TFIIIE34, and RAP74 and TFIIIE56. Increasing amounts of TFIIA were added to the reactions. The positions of the abortive transcripts (4–10 nt) are indicated. *C*, quantification of the stimulatory effect of TFIIA on abortive initiation. The intensity of the bands corresponding to the abortive transcripts from 4–6 experiments were quantitated using a PhosphorImager. The measured intensities were used to calculate the ratios of amount of transcript produced in the presence of TFIIA to that produced in its absence in each case (Fold Stimulation).

(α , β , and γ subunits) was retained on the TFIIIE56, TFIIIE34, RAP74, and TBP columns but not on the TFIIIB, RAP30, RNAPII, and BSA columns (Fig. 2). Contaminant bands of the TFIIA fraction were visible in the flowthrough of all the columns. The binding of TFIIA to the affinity columns is most easily visualized by examining the elution of the β and γ subunits.

To further characterize these interactions, we next used the individual subunits of TFIIA in our affinity chromatography experiments. TFIIA α/β and TFIIA γ were individually chromatographed on affinity columns containing immobilized

RAP74, TFIIIE34 and TFIIIE56. BSA was added to the input as an internal negative control. Fig. 3 shows that TFIIA α/β , but not TFIIA γ , interacts with RAP74 and TFIIIE34. TFIIA α/β did not bind to the TFIIIE56 column, whereas TFIIA γ was retarded on the TFIIIE56 column, suggesting a weak interaction. This finding is surprising in view of the observation that nTFIIA binds strongly to the TFIIIE56 column. Perhaps this is due to the association of TFIIA α/β and TFIIA γ resulting in a conformational change that favors binding to TFIIIE56.

Two groups have previously reported interactions between TFIIA and TFIIIE. Both used recombinant TFIIA, not the natural protein. In agreement with our results, Yamamoto *et al.* (59) obtained binding of TFIIIE56 to TFIIA γ , and Yokomori *et al.* (60) obtained binding of TFIIIE34 to TFIIA α/β . However, in contrast to both our data and that of Yokomori *et al.* (60), Yamamoto *et al.* (59) observed a weak binding of TFIIIE34 to TFIIA γ but not to TFIIA α/β . The use of different binding assays may have caused this apparent discrepancy. Consistent with the interaction we observe, Yokomori *et al.* (60) have also shown that TFIIIE interacts with the TFIIA-TBP complex on promoter DNA.

RAP74 Contains Two Distinct TFIIA-binding Domains—To determine the domain or domains of RAP74 responsible for the interaction with TFIIA, a series of RAP74 deletion mutants were immobilized on different affinity columns and rTFIIA (α/β and γ renatured together) chromatographed through each column. BSA was added to the input as an internal control, and a BSA column served as a negative control. Fig. 4A shows some representative data, and a summary is presented in Fig. 4B. All the N-terminal fragments of RAP74, except for RAP74-(1–75), which contains only the first 75 amino acids of the polypeptide, were bound by rTFIIA (Fig. 4A). Because rTFIIA did bind to RAP74-(1–136) but not to RAP74-(1–75), our results define a first domain of interaction between these two proteins that encompasses amino acids 76–136. We next used the C-terminal fragments of RAP74 in our affinity chromatography experiments. RAP74-(207–517), RAP74-(358–517), and RAP74-(407–517) were all observed to bind to RAP74, indicating the existence of a second interacting domain. To delineate this domain more precisely, two additional mutants, RAP74-(363–444) and RAP74-(363–409), were used. TFIIA bound well to RAP74-(363–444) but not to RAP74-(363–409) (Fig. 4A). These results define a second TFIIA-interacting domain of RAP74 located between amino acids 410 and 444. The existence of an additional putative TFIIA-binding domain in the central RAP74 region was ruled out by passing rTFIIA on columns containing RAP74-(136–258) and RAP74-(258–356). In each case, rTFIIA was not retained on the column. The two TFIIA-interacting domains of RAP74 that we identified are both localized in conserved regions of the protein, domain 76–136 being localized in conserved region I and domain 410–444 in conserved region III (see Fig. 4B for a schematic representation).

To assess whether or not RAP74 can associate with TFIIA on promoter DNA, we performed gel mobility shift experiments. RAP74 fragments carrying the TFIIA-binding domains (*e.g.* RAP74-(1–517;wt), RAP74-(1–136), and RAP74-(363–444)) but not fragments lacking the TFIIA-binding regions (*e.g.* RAP74-(1–75) and RAP74-(363–409)) gel shifted a TBP-TFIIA-promoter complex (Fig. 5). These results indicate that RAP74, through both its TFIIA-binding domains, can associate with TFIIA on promoter DNA.

TFIIA Stimulates the Activity of TFIIIE34 and RAP74 in Transcription Initiation—Several reports have shown that TFIIA can stimulate basal transcription by RNAPII *in vitro*, but is not essential for the basal transcription reaction (28, 33–37). For example, recombinant TFIIA increased the forma-

tion of a 391-nt run-off transcript from a supercoiled template carrying the AdMLP fused to a G-less cassette in the presence of TBP, TFIIB, TFIIE, TFIIF, and core RNAPII (Fig. 6A). To test whether or not TFIIA acts on the initiation step of the transcription reaction, we developed an abortive initiation assay in which an 80-base pair double-stranded oligonucleotide carrying the AdMLP was used to drive transcription initiation in the presence of RNAPII and the general transcription factors. In this assay only abortive transcripts of 2 to 10 nt in length can be synthesized because the transcription reaction is performed in the absence of GTP on a template that harbors a G at position +11 (see Fig. 6B). Abortive initiation was shown to be promoter-specific because the use of DNA templates with mutations in the AdMLP that impair transcription in run-off assays did not support synthesis of abortive transcripts.² Under our reaction conditions, abortive initiation minimally requires the presence of TBP, TFIIB, RAP30, and RNAPII (Fig. 6). With this minimal set of factors, the addition of TFIIA does not stimulate the formation of abortive transcripts (Fig. 6, B and C). When either RAP74 alone, or RAP74 and TFIIE34 are added to the system, the formation of abortive transcripts is increased, indicating that these two factors are involved in transcriptional initiation (Fig. 6, data not shown, and Refs. 58, 61, 62). The addition of TFIIA to reactions containing either RAP74 alone, or RAP74 and TFIIE34, in addition to TBP, TFIIB, RAP30, and RNAPII, had a stimulatory effect (Fig. 6, B and C), indicating that TFIIA can stimulate the initiation stage of the transcription reaction. Because TFIIA stimulates transcription initiation only when RAP74 and TFIIE34 are present in the reaction, our results suggest that TFIIA enhances the functions of TFIIF and TFIIE to stimulate the basal transcription reaction.

The role of TFIIA in RNAPII transcription is not completely resolved. However, several reports suggest that an important function of this factor is to act as a co-factor in transcriptional activation (28, 29, 33, 34, 44–47). In this paper, we establish the existence of structural and functional interactions between TFIIA and both TFIIE and TFIIF within the initiation complex. TFIIE and TFIIF have been shown to be involved in the melting of promoter DNA near the initiation site during open complex formation (9, 58). Considering that an activator such as Gal4-VP16, whose full activity requires the presence of TFIIA, can stimulate promoter melting near the initiation site (63), it is possible to envision an activation mechanism in which TFIIA functions as a bridge between the activator proteins and the general transcription factors TFIIE and TFIIF. This connection may help to explain how upstream activators influence and stimulate the biochemical events occurring at the transcription start site.

Acknowledgments—We thank the members of our laboratories for helpful discussions, Diane Bourque for artwork, and Will Home for critical reading of the manuscript.

REFERENCES

- Orphanides, G., Lagrange, T., and Reinberg, D. (1996) *Genes Dev.* **10**, 2657–2683
- Hampsey, M. (1998) *Microbiol. Mol. Biol. Rev.* **62**, 465–503
- Kim, Y., Geiger, J. H., Hahn, S., and Sigler, P. B. (1993) *Nature* **365**, 512–520
- Kim, J. L., Nikolov, D. B., and Burley, S. K. (1993) *Nature* **365**, 520–527
- Maldonado, E., Ha, I., Cortes, P., Weis, L., and Reinberg, D. (1990) *Mol. Cell. Biol.* **10**, 6335–6347
- Sopta, M., Carthew, R. W., and Greenblatt, J. (1985) *J. Biol. Chem.* **260**, 10353–10360
- Conaway, R. C., Garrett, K. P., Hanley, J. P., and Conaway, J. W. (1991) *Proc. Natl. Acad. Sci. U. S. A.* **88**, 6205–6209
- Flores, O., Lu, H., and Reinberg, D. (1992) *J. Biol. Chem.* **267**, 2786–2793
- Holstege, F. C., Tantin, D., Carey, M., van der Vliet, P. C., and Timmers, H. T. (1995) *EMBO J.* **14**, 810–819
- Schaeffer, L., Roy, R., Humbert, S., Moncollin, V., Vermeulen, W., Hoeijmakers, J. H., Chambon, P., and Egly, J. M. (1993) *Science* **260**, 58–63
- Drapkin, R., Reardon, J. T., Ansari, A., Huang, J. C., Zawel, L., Ahn, K., Sancar, A., and Reinberg, D. (1994) *Nature* **368**, 769–772
- Lu, H., Flores, O., Weinmann, R., and Reinberg, D. (1991) *Proc. Natl. Acad. Sci. U. S. A.* **88**, 10004–10008
- Serizawa, H., Conaway, R. C., and Conaway, J. W. (1992) *Proc. Natl. Acad. Sci. U. S. A.* **89**, 7476–7480
- Dvir, A., Conaway, R. C., and Conaway, J. W. (1997) *Proc. Natl. Acad. Sci. U. S. A.* **94**, 9006–9010
- Yan, Q., Moreland, R. J., Conaway, J. W., and Conaway, R. C. (1999) *J. Biol. Chem.* **274**, 35668–35675
- Gnatt, A. L., Cramer, P., Fu, J., Bushnell, D. A., and Kornberg, R. D. (2001) *Science* **292**, 1876–1882
- Cramer, P., Bushnell, D. A., and Kornberg, R. D. (2001) *Science* **292**, 1863–1876
- Coulombe, B., Li, J., and Greenblatt, J. (1994) *J. Biol. Chem.* **269**, 19962–19967
- Robert, F., Forget, D., Li, J., Greenblatt, J., and Coulombe, B. (1996) *J. Biol. Chem.* **271**, 8517–8520
- Lagrange, T., Kim, T. K., Orphanides, G., Ebricht, Y. W., Ebricht, R. H., and Reinberg, D. (1996) *Proc. Natl. Acad. Sci. U. S. A.* **93**, 10620–10625
- Forget, D., Robert, F., Grondin, G., Burton, Z. F., Greenblatt, J., and Coulombe, B. (1997) *Proc. Natl. Acad. Sci. U. S. A.* **94**, 7150–7155
- Kim, T. K., Lagrange, T., Wang, Y. H., Griffith, J. D., Reinberg, D., and Ebricht, R. H. (1997) *Proc. Natl. Acad. Sci. U. S. A.* **94**, 12268–12273
- Robert, F., Douziech, M., Forget, D., Egly, J. M., Greenblatt, J., Burton, Z. F., and Coulombe, B. (1998) *Mol. Cell* **2**, 341–351
- Douziech, M., Forget, D., Greenblatt, J., and Coulombe, B. (1999) *J. Biol. Chem.* **274**, 19868–19873
- Kim, T. K., Ebricht, R. H., and Reinberg, D. (2000) *Science* **288**, 1418–1421
- Douziech, M., Coin, F., Chipoulet, J. M., Arai, Y., Ohkuma, Y., Egly, J. M., and Coulombe, B. (2000) *Mol. Cell. Biol.* **20**, 8168–8177
- Coulombe, B., and Burton, Z. F. (1999) *Microbiol. Mol. Biol. Rev.* **63**, 457–478
- Ozer, J., Moore, P. A., Bolden, A. H., Lee, A., Rosen, C. A., and Lieberman, P. M. (1994) *Genes Dev.* **8**, 2324–2335
- Ma, D., Watanabe, H., Mermelstein, F., Admon, A., Oguri, K., Sun, X., Wada, T., Imai, T., Shiroya, T., and Reinberg, D. (1993) *Genes Dev.* **7**, 2246–2257
- DeJong, J., and Roeder, R. G. (1993) *Genes Dev.* **7**, 2220–2234
- DeJong, J., Bernstein, R., and Roeder, R. G. (1995) *Proc. Natl. Acad. Sci. U. S. A.* **92**, 3313–3317
- Ranish, J. A., Lane, W. S., and Hahn, S. (1992) *Science* **255**, 1127–1129
- Sun, X., Ma, D., Sheldon, M., Yeung, K., and Reinberg, D. (1994) *Genes Dev.* **8**, 2336–2348
- Yokomori, K., Zeidler, M. P., Chen, J. L., Verrijzer, C. P., Mlodzik, M., and Tjian, R. (1994) *Genes Dev.* **8**, 2313–2323
- Ma, D., Olave, I., Merino, A., and Reinberg, D. (1996) *Proc. Natl. Acad. Sci. U. S. A.* **93**, 6583–6588
- Coulombe, B., Killeen, M., Liljelund, P., Honda, B., Xiao, H., Ingles, C. J., and Greenblatt, J. (1992) *Gene Expr.* **2**, 99–110
- Cortes, P., Flores, O., and Reinberg, D. (1992) *Mol. Cell. Biol.* **12**, 413–421
- Buratowski, S., Hahn, S., Guarente, L., and Sharp, P. A. (1989) *Cell* **56**, 549–561
- Auble, D. T., Hansen, K. E., Mueller, C. G., Lane, W. S., Thorner, J., and Hahn, S. (1994) *Genes Dev.* **8**, 1920–1934
- Ge, H., and Roeder, R. G. (1994) *J. Biol. Chem.* **269**, 17136–17140
- Inostroza, J. A., Mermelstein, F. H., Ha, I., Lane, W. S., and Reinberg, D. (1992) *Cell* **70**, 477–489
- Meisterernst, M., Roy, A. L., Lieu, H. M., and Roeder, R. G. (1991) *Cell* **66**, 981–993
- Merino, A., Madden, K. R., Lane, W. S., Champoux, J. J., and Reinberg, D. (1993) *Nature* **365**, 227–232
- Wang, W., Gralla, J. D., and Carey, M. (1992) *Genes Dev.* **6**, 1716–1727
- Kang, J. J., Auble, D. T., Ranish, J. A., and Hahn, S. (1995) *Mol. Cell. Biol.* **15**, 1234–1243
- Kobayashi, N., Boyer, T. G., and Berk, A. J. (1995) *Mol. Cell. Biol.* **15**, 6465–6473
- Chi, T., and Carey, M. (1996) *Genes Dev.* **10**, 2540–2550
- Ingles, C. J., Shales, M., Cress, W. D., Triezenberg, S. J., and Greenblatt, J. (1991) *Nature* **351**, 588–590
- Ha, I., Lane, W. S., and Reinberg, D. (1991) *Nature* **352**, 689–695
- Finkelstein, A., Kostrub, C. F., Li, J., Chavez, D. P., Wang, B. Q., Fang, S. M., Greenblatt, J., and Burton, Z. F. (1992) *Nature* **355**, 464–467
- Ohkuma, Y., Sumimoto, H., Hoffmann, A., Shimasaki, S., Horikoshi, M., and Roeder, R. G. (1991) *Nature* **354**, 398–401
- Sumimoto, H., Ohkuma, Y., Sinn, E., Kato, H., Shimasaki, S., Horikoshi, M., and Roeder, R. G. (1991) *Nature* **354**, 401–404
- Peterson, M. G., Inostroza, J., Maxon, M. E., Flores, O., Admon, A., Reinberg, D., and Tjian, R. (1991) *Nature* **354**, 369–373
- Hodo, H. G., III, and Blatti, S. P. (1977) *Biochem.* **16**, 1334–2343
- Robert, F., and Coulombe, B. (2001) *Methods Mol. Biol.* **148**, 383–393
- Wolner, B. S., and Gralla, J. D. (2000) *Mol. Cell. Biol.* **20**, 3608–3615
- Burton, Z. F., Ortolan, L. G., and Greenblatt, J. (1986) *EMBO J.* **5**, 2923–2930
- Pan, G., and Greenblatt, J. (1994) *J. Biol. Chem.* **269**, 30101–30104
- Yamamoto, S., Watanabe, Y., van der Spek, P. J., Watanabe, T., Fujimoto, H., Hanaoka, F., and Ohkuma, Y. (2001) *Mol. Cell. Biol.* **21**, 1–15
- Yokomori, K., Verrijzer, P. C., and Tjian, R. (1998) *Proc. Natl. Acad. Sci. U. S. A.* **95**, 6722–6727
- Lei, L., Ren, D., and Burton, Z. F. B. (1999) *Mol. Cell. Biol.* **19**, 8372–8382
- Holstege, F. C. P., van der Vliet, P. C., and Timmers, M. H. T. (1996) *EMBO J.* **15**, 1666–1677
- Wang, W., Carey, M., and Gralla, J. D. (1992) *Science* **255**, 450–453

² M. F. Langelier, Y. Porlier, and B. Coulombe, manuscript in preparation.

CHAPITRE 3

RÔLES CENTRAUX POUR LA PINCE DE L'ARN POLYMÉRASE II ET TFIIE DANS L'INITIATION, RÉVÉLÉS PAR PHOTO-PONTAGE PROTÉINE-ADN D'UN COMPLEXE PRÉTRANSCRIPTIONNEL

3.1 Préambule

J'ai contribué à l'article du chapitre 3 en réalisant les expériences d'initiation présentées à la figure 5. Dans cette figure, des essais d'initiation ont été faits sur une série de matrices ayant des préouvertures à différentes positions dans la région de la bulle de transcription. Ces essais ont permis de montrer, conjointement avec le résultat de photo-pontage de TFIIE α à -10, que TFIIE joue un rôle dans l'initiation en stimulant l'ouverture du promoteur.

Diane Forget a mis au point la méthode de photo-pontage en gel utilisée dans cet article et a réalisé, en collaboration avec Cynthia Thérien, toutes les expériences effectuées avec cette technique (figure 1A, 2, 3 et 4). Vincent Trinh a produit la simulation moléculaire de la figure 1B et a participé à la réalisation du modèle (figure 6).

3.2 Article

Forget, D., Langelier, M.-F., Thérien, C., Trinh, V. and Coulombe, B. (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**(3):1122-31

Photo-Cross-Linking of a Purified Preinitiation Complex Reveals Central Roles for the RNA Polymerase II Mobile Clamp and TFIIE in Initiation Mechanisms

Diane Forget, Marie-France Langelier, Cynthia Thérien, Vincent Trinh, and Benoit Coulombe*

Laboratory of Gene Transcription, Institut de Recherches Cliniques de Montréal, Montréal, Québec, Canada H2W 1R7

Received 24 September 2003/Accepted 3 November 2003

The topological organization of a TATA binding protein-TFIIB-TFIIF-RNA polymerase II (RNAP II)-TFIIE-promoter complex was analyzed using site-specific protein-DNA photo-cross-linking of gel-purified complexes. The cross-linking results for the subunits of RNAP II were used to determine the path of promoter DNA against the structure of the enzyme. The results indicate that promoter DNA wraps around the mobile clamp of RNAP II. Cross-linking of TFIIF and TFIIE both upstream of the TATA element and downstream of the transcription start site suggests that both factors associate with the RNAP II mobile clamp. TFIIE α closely approaches promoter DNA at nucleotide -10, a position immediately upstream of the transcription bubble in the open complex. Increased stimulation of transcription initiation by TFIIE α is obtained when the DNA template is artificially premelted in the -11/-1 region, suggesting that TFIIE α facilitates open complex formation, possibly through its interaction with the upstream end of the partially opened transcription bubble. These results support the central roles of the mobile clamp of RNAP II and TFIIE in transcription initiation.

RNA polymerase II (RNAP II) is the 12-subunit enzyme that synthesizes all mammalian mRNA (5, 58). Biochemical analyses have revealed that the transcription reaction involves a number of successive steps that lead to the formation of a pre-mRNA (19, 38). In the first step, RNAP II locates promoter DNA and positions its catalytic center near the transcriptional initiation site of the promoter DNA. This first step proceeds through the formation of a preinitiation complex that contains, in addition to RNAP II, a number of general initiation factors, including TATA binding protein (TBP), TFIIB, TFIIF, TFIIE, and TFIIH. Assembly of the preinitiation complex requires specific binding of some general initiation factors to core promoter elements, such as binding of TBP to the TATA box (3), TFIIB to the TFIIB recognition element (29), and TAFs of TFIID to Inr (50) and the downstream promoter element (2). Formation of the preinitiation complex is accompanied by topological changes, including bending and wrapping of promoter DNA around the protein core of the complex (13, 28, 30, 36, 42, 45). Recent evidence indicates that DNA wrapping during transcription initiation is required for the accurate positioning of the initiation site near the enzyme catalytic center and the induction of topological constraints essential for promoter melting (8, 42). In the second step, promoter DNA between nucleotides (nt) -9 and +2 is melted in such a way that the template strand becomes accessible to NTP polymerization (21, 23, 26, 57). Both TFIIF and TFIIE were shown to participate in promoter melting (22, 39). Full open-complex formation requires the action of the ATP-de-

pendent helicase activity of TFIIH (21, 23). Two distinct single-stranded DNA helicases, XPB/p89 and XPD/p80, have been identified as components of TFIIH (46–48, 53, 56). A major role in promoter DNA melting prior to initiation was attributed to XPB/p89 (27, 54). During the third step, RNAP II enters a cycle of abortive initiation events in which the enzyme synthesizes many short 2- to 10-nt transcripts (21). A structural transition within the initiation complex, including a remodeling of the transcription bubble, occurs when RNAP II reaches register 11 and enters the processive phase of the transcription reaction (21). TFIIH is responsible for the melting of the template DNA during promoter escape (16, 32, 51) and for the phosphorylation of the RNAP II carboxy-terminal domain (10, 31, 49). Formation of the mRNA is completed through transcript elongation and is followed by termination of the transcription reaction.

The availability of crystal structures for both eukaryotic (6, 7, 14, 15) and prokaryotic (4, 33, 34, 55, 59) RNAPs has been invaluable for the understanding of the many molecular features of the transcription reaction. For example, the structure of elongating RNAP II has revealed the position of the RNA-DNA duplex formed during the transcription reaction (15). The available structures support a model in which the DNA enters the enzyme through a channel formed by a pair of “jaws” before accessing a deep cleft, at the bottom of which is buried the active site with its Mg²⁺ ions; the DNA then turns by ~90° along a wall, where the upstream end exits the enzyme (15). Recently, the crystal structure of the *Thermus aquaticus* RNAP holoenzyme bound to a fork junction DNA fragment was resolved (33). The structure shows that the DNA lies across one face of the holoenzyme, completely outside the active-site channel. A similar high-resolution structure remains to be obtained for RNAP II.

* Corresponding author. Mailing address: Laboratory of Gene Transcription, Institut de Recherches Cliniques de Montréal, 110 Ave. des Pins Ouest, Montréal, Québec, Canada H2W 1R7. Phone: (514) 987-5662. Fax: (514) 987-5663. [REDACTED]

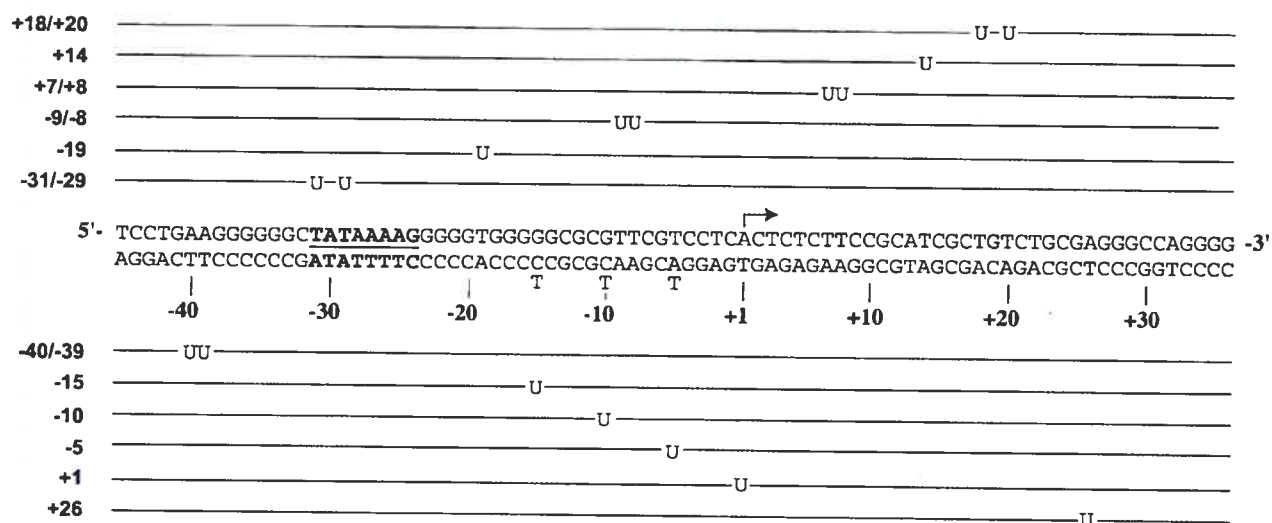
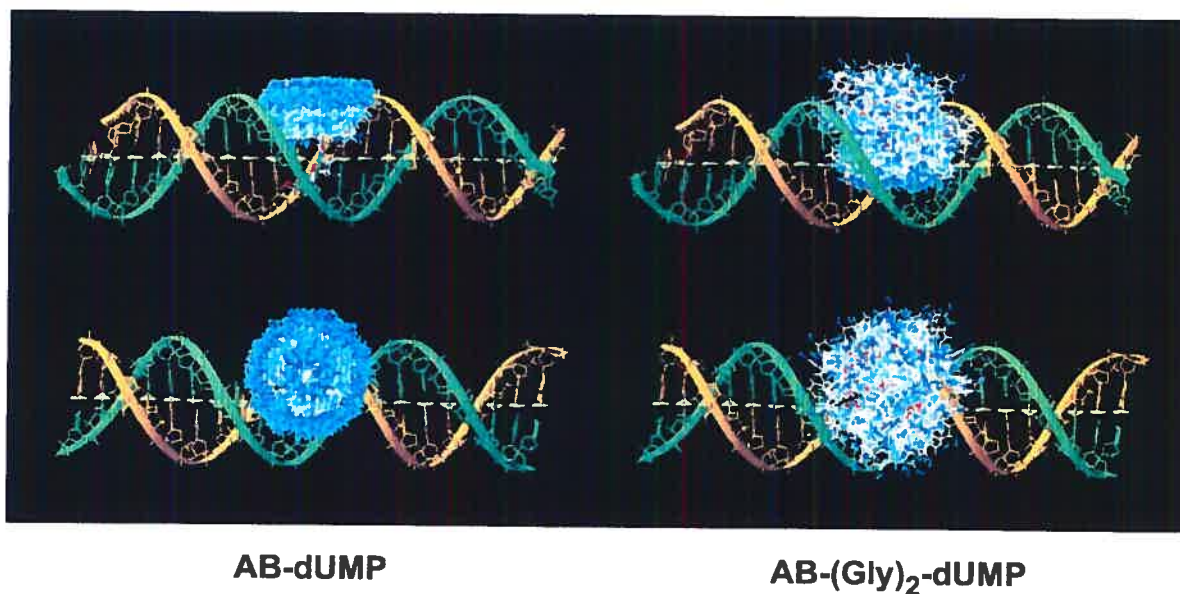
A**B**

FIG. 1. Photoprobes used for in-gel protein-DNA photo-cross-linking. (A) The 12 photoprobes used in our analysis are schematically shown. The position of the photonucleotide AB-dUMP or AB-(Gly)₂-dUMP (U) along the adenovirus major late promoter is indicated (positions -15, -10, and -5) in each case and is used to name each photoprobe. Mutations (N to T) used in some of our photoprobes are indicated. The positions of the TATA element (underlined and boldface) and initiation site (arrow) are indicated. (B) Low-energy conformers of AB-dUMP and AB-(Gly)₂-dUMP in the DNA major groove. Molecular simulation was used to predict the various possible conformers in each case. The photoreactive nitrene is shown in blue.

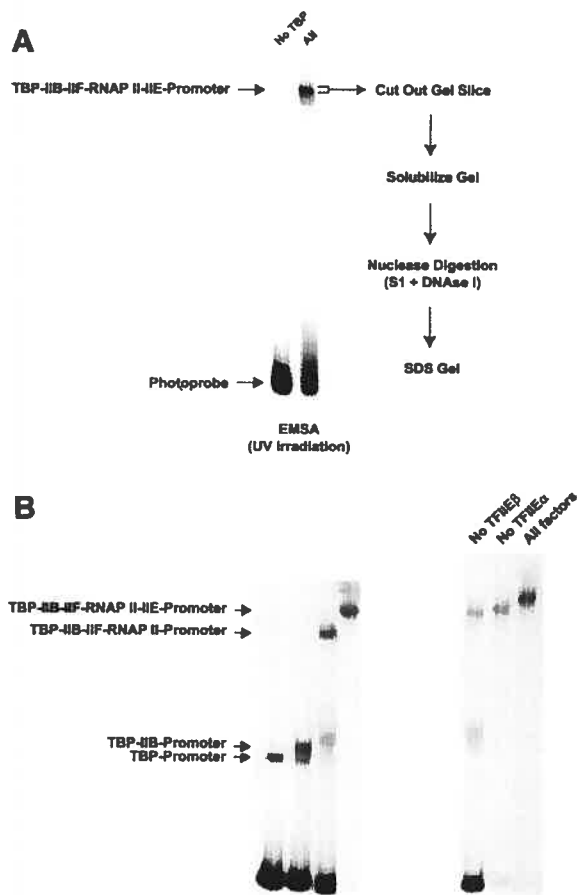


FIG. 2. In-gel photo-cross-linking of the RNAP II preinitiation complex. (A) Schematic representation of the in-gel photo-cross-linking procedure. Preinitiation complexes were assembled by mixing each photoprobe with RNAP II and the general transcription factors. The TBP-IIB-IIF-RNAP II-IIE-promoter complex was electrophoresed in native polyacrylamide-*N,N'*-bisacryloylcystamine gels and irradiated with UV directly in the gels. Gel slices containing the cross-linked complexes were recovered, solubilized, and treated with nucleases. The resulting cross-linked products were then analyzed by SDS-PAGE. EMSA, electrophoretic mobility shift assay. (B) Gel mobility shift experiments showing the various intermediate complexes assembled using subsets of the general initiation factors and RNAP II. The positions of the various complexes are shown.

Over the years, site-specific protein-DNA photo-cross-linking has been used to analyze the topological organization of RNAP II complexes bound to DNA. The results indicate that the RNAP II preinitiation complex forms a very compact structure in which the DNA is tightly bent and wrapped against the core of the complex (13, 45). Our proposed model, named the DNA-wrapping model, describes basal transcription mechanisms (5). However, two puzzling features stand out from our cross-linking results: (i) some subunits of the complex cross-link to a very extensive region of the promoter, and (ii) some promoter positions can cross-link many different polypeptides. One possible explanation for these observations stems from the fact that our cross-linking procedure was performed on protein-DNA complexes that were not purified after assembly, possibly allowing the formation of a heterogeneous subpopu-

lation of minor complexes in addition to the main complex. These subcomplexes may result from nonspecific interactions between components of the transcription machinery and promoter DNA. In order to circumvent this possible caveat and to improve the resolution of our photo-cross-linking procedure, we have developed an in-gel protein-DNA photo-cross-linking method in which protein-DNA complexes are separated using native gel electrophoresis prior to their UV irradiation and extraction from the gel. We report here the use of this method to study cross-linking of the transcription machinery assembled into a TBP-TFIIB-TFIIF-RNAP II-TFII E-promoter complex. As expected, this in-gel photo-cross-linking methodology increases the resolution by allowing more precise localization of the various factors along the promoter DNA. Our results suggest key roles both for the mobile clamp of RNAP II and for TFII E during the preinitiation stage of transcription.

MATERIALS AND METHODS

Protein factors. Recombinant yeast TBP (25), human TFIIB (18), RAP30 (11), RAP74 (11), TFII E α and TFII E β (37, 41, 52), and calf thymus RNAP II (20) were prepared as previously described.

In-gel photo-cross-linking. The photoreactive nucleotides azidobenzoic (AB)-dUTP and AB-(Gly)₂-dUTP were synthesized as described by Bartholomew and colleagues and Persinger and Bartholomew (1, 40). The in-gel photo-cross-linking procedure was performed as recently described (12). Briefly, various photoprobes placing one or two photonucleotides in juxtaposition to one (or a few) radiolabeled nucleotide at a specific position along promoter DNA were prepared enzymatically and gel purified. Each probe was incubated for 30 min at 37°C with highly purified RNAP II and recombinant general initiation factors (see below and figure legends for details) and then loaded on a polyacrylamide-*N,N'*-bisacryloylcystamine native gel (35). The gels were run in a cold room and irradiated with UV light (254 nm), and the complexes were localized using a PhosphorImager (Molecular Dynamics). Gel slices containing the complexes were excised and solubilized with dithiothreitol (0.4 M final concentration). The cross-linked complexes were then treated with S1 nuclease and DNase I prior to analysis by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) as previously described (44). Immunoprecipitation of the cross-linking products used specific antibodies directed against the various subunits of the general transcription factors as previously described (44).

Initiation assay. Abortive initiation assays were done essentially as previously described (30, 39) with the following modifications. Double-stranded oligonucleotide templates (12 ng) carrying the adenovirus major late promoter from -45 to +35 were incubated for 60 min at 30°C with TBP (60 ng), TFIIB (30 ng), RAP74 (65 ng), RAP30 (30 ng), RNAP II (165 ng), TFII E β (40 ng), and, when indicated, TFII E α (60 ng) in 20 μ l of a reaction mixture containing 750 μ M ATP, 750 μ M CTP, 10 μ M UTP, 2.5 μ M [α -³²P]UTP, 12.5 mM MgCl₂, 3 mM EGTA, and 0.84 U of RNase inhibitor/ml. After the reaction was stopped, the mixture was treated with calf intestine alkaline phosphatase (8 U) in order to reduce the background caused by free radiolabeled nucleotides. Transcripts were analyzed on a 23% polyacrylamide denaturing gel containing 7 M urea and quantified using a PhosphorImager.

Computer modeling. Prediction of low-energy AB-dUMP and AB-(Gly)₂-dUMP conformers was performed using the Sybyl Molecular Modeling Package (version 6.4; Tripos Associates, St. Louis, Mo.). The models (see Fig. 6) were built by importing the RNAP II (accession code, 1150) and TBP-TFIIB-DNA (accession code, 1V0l) PDB files from SWISS-PDB version 3.7 (17) and VMD version 1.7.2 (24) into the Cinema 4D XL Modeling Package (version 6.3; Maxon Computer, Thousand Oaks, Calif.).

RESULTS

Photo-cross-linking of gel-purified RNAP II preinitiation complexes. To analyze the molecular organization of the RNAP II preinitiation complex, we developed a procedure for photo-cross-linking proteins to specific sites along promoter DNA following purification of the protein-DNA complexes by electrophoresis in native gels. A total of 12 different photo-

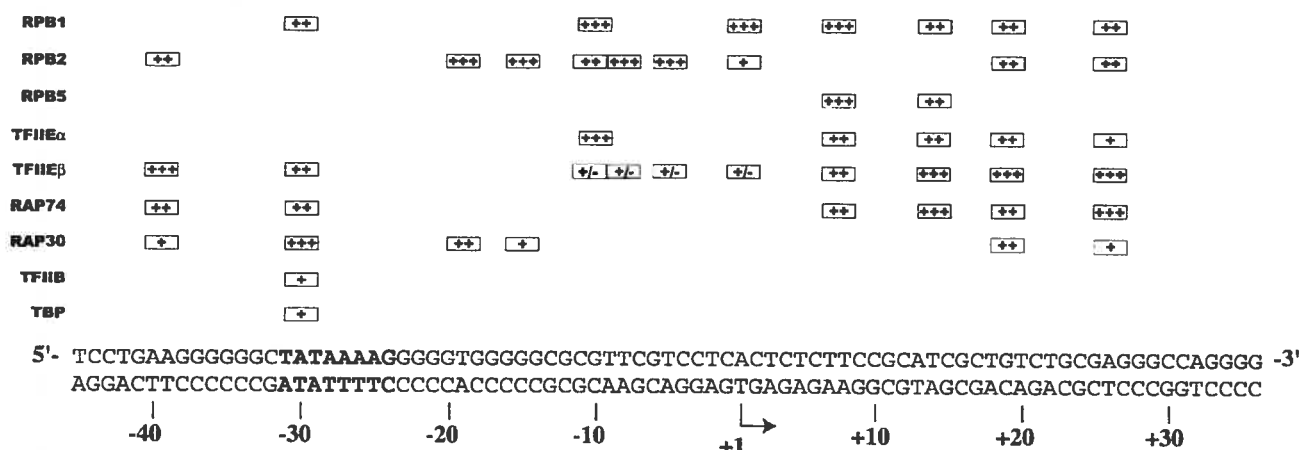


FIG. 3. Summary of our in-gel photo-cross-linking results with a TBP-TFIIB-TFIIF-RNAP II-TFIIEx-promoter complex. The boxes indicate the positions where each individual factor was found to cross-link to promoter DNA. +++, ++, and + represent the average relative intensities of the cross-linking signals obtained for each photoprobe using AB-dUMP. For example, the cross-linking of RPB2 at position -19 is more efficient than that of RAP30. The \pm signs for the cross-linking of TFIIEb to photoprobes -10, -9/-8, -5, and +1 indicate the uncertainty concerning their specificities, as nonspecific complexes made of multiple molecules of TFIIEb were formed on these photoprobes in the absence of TBP (see the text for details). The use of AB-(Gly)₂-dUTP in our photoprobes led to cross-linking results that are qualitatively similar to those using AB-dUMP. The positions of the TATA element (boldface) and initiation site (arrow) are indicated.

probes that place one or two photoreactive nucleotides in juxtaposition to one or a few radiolabeled nucleotides at specific positions along the adenovirus major late promoter (Fig. 1A) were used to analyze the promoter contacts by a TBP-TFIIB-TFIIF-RNAP II-TFIIEx complex. Recombinant TBP, TFIIB, RAP74, RAP30, TFIIEx, TFIIEb, and highly purified RNAP II were mixed with the various photoprobes, and the resulting complexes were separated on native polyacrylamide gels (Fig. 2A). The gels were irradiated with UV light, and gel slices containing the protein-DNA complexes were excised. Using this procedure, any intermediate complex that may have formed was disregarded. After dissolution of the gel fragments, the samples were treated with nucleases and the cross-linked polypeptides were then analyzed by SDS-PAGE. The cross-linked polypeptides were visualized by autoradiography of the gels and identified according to their molecular weights. The identities of the cross-linked polypeptides were confirmed either by immunoprecipitation of the cross-linking products with specific antibodies or by using tagged factors in the cross-linking reactions. The results of at least four independent experiments using each individual photoprobe were compiled to produce Fig. 3, which shows the cross-links that were reproducibly obtained. In each case, a control reaction lacking TBP was assembled to ensure promoter binding specificity (Fig. 2A shows an example). We also performed gel mobility shift experiments using subsets of the general initiation factors and RNAP II to ensure that our conditions for complex assembly allowed the formation of a bona fide TBP-TFIIB-TFIIF-RNAP II-TFIIEx-promoter complex (Fig. 2B).

Our photo-cross-linking experiments used two different photoreactive nucleotides with side chains of various lengths, AB-dUMP (10.0 Å) and AB-(Gly)₂-dUMP (18.6 Å). Both photoreactive nucleotides cross-linked the same set of polypeptides at all tested positions (Fig. 3 shows a summary). Molecular simulations indicated that the photoreactive nucleotides AB-

dUMP and AB-(Gly)₂-dUMP can adopt many different conformations as a result of the rotation of their aryl groups around the axes of some chemical bonds. The results of the simulations indicated that the reactive nitrene of AB-dUMP can probe an extensive space (~20 by 20 by 10 Å) in the major groove of the DNA helix (Fig. 1B). The simulations also indicated that only a slightly larger space is probed by the nitrene of AB-(Gly)₂-dUMP than by that of AB-dUMP. This result may explain the similarity of the cross-linking data obtained with both photonucleotides. The flexibility of both AB-dUMP and AB-(Gly)₂-dUMP indicates that photoprobes containing a single photonucleotide could theoretically react with multiple polypeptides that approach the space occupied by the reactive nitrene. This situation is exacerbated for those photoprobes containing two photonucleotides.

Promoter contacts by the RNAP II subunits RPB1, RPB2, and RPB5. The in-gel site-specific protein-DNA photo-cross-linking results are summarized in Fig. 3, and some representative data are shown in Fig. 4. RPB1 cross-linked to seven positions from nt -31/-29 to +26, RPB2 cross-linked to nine positions from nt -40/-39 to +26, and RPB5 cross-linked to two positions from nt +7/+8 to +14 (Fig. 4B). Notably, RPB2 cross-linked to the promoter at six consecutive positions between -19 and +1, indicating that this promoter region closely approaches the second-largest RNAP II subunit. Between +7/+8 and +14, the promoter cross-linked to both RPB1 and RPB5, while further downstream, between +18/+20 and +26, cross-links were obtained with both RPB1 and RPB2. RPB1 cross-linked to the TATA box (-31/-29) and RPB2 upstream of TATA at -40/-39. The region between -10 and +1, where promoter melting occurs in the presence of TFIIB and hydrolyzable ATP, makes only seven cross-links with the TBP-TFIIB-TFIIF-RNAP II-TFIIEx complex, four of which are with RPB2.

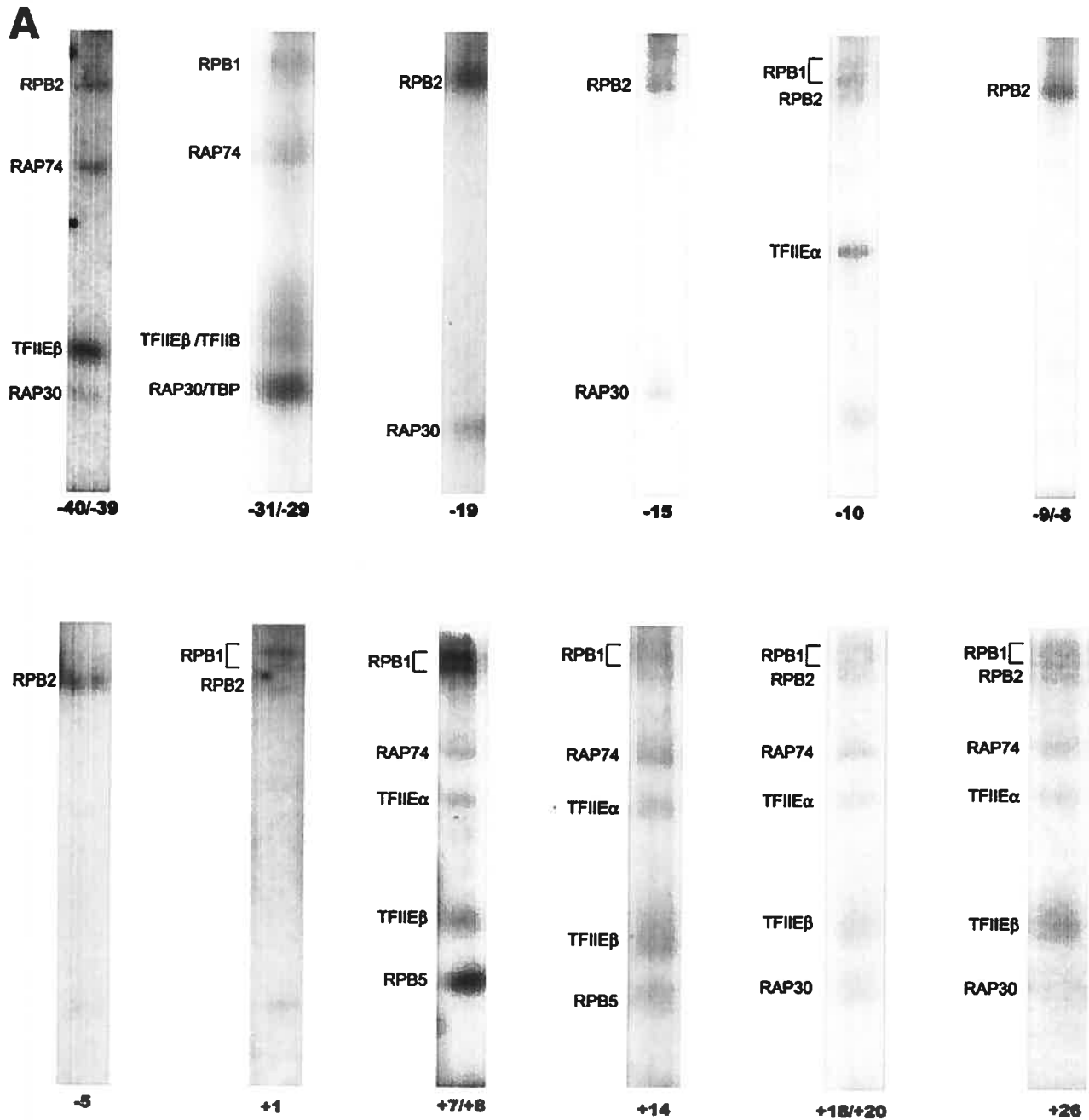


FIG. 4. In-gel photo-cross-linking data. (A) Examples of the results obtained in our photo-cross-linking experiments with photoprobes prepared using AB-dUMP. The position of each cross-linked polypeptide is indicated. (B) Gel showing that the polypeptide that specifically cross-links at positions +7/+8 and +14 comigrates with the RPB5 subunit of RNAP II as identified using mass spectrometry. (C and D) RAP30 (C) and TFIIIE β (D) cross-link to photoprobes -40/-39 and +18/+20. The identities of the cross-linking bands were confirmed by immunoprecipitation of cross-linked polypeptides with antibodies (Ab) raised against RAP30 and TFIIIE β . +, present; -, absent.

Promoter contacts by TFIIIF subunits. Our results, shown in Fig. 3 and 4, reveal a limited number of cross-links by TFIIIF subunits along promoter DNA. RAP74 strongly cross-linked to four positions between +7/+8 and +26 and cross-linked more weakly to TATA and upstream of it. RAP30 cross-linked to four positions between -40/-39 and -15. Surprisingly, additional cross-links by RAP30 were observed at positions +18/

+20 and +26. The cross-linking of RAP30 to nt -40/-39 and +18/+20, two positions \sim 60 bp apart, has been confirmed by immunoprecipitation of cross-linked RAP30 using an antibody directed against RAP30 (Fig. 4C). Our photo-cross-linking data for TFIIIF subunits indicates that the location of RAP74 is centered downstream of +1 whereas that of RAP30 is centered on the TATA box and immediately downstream of it.

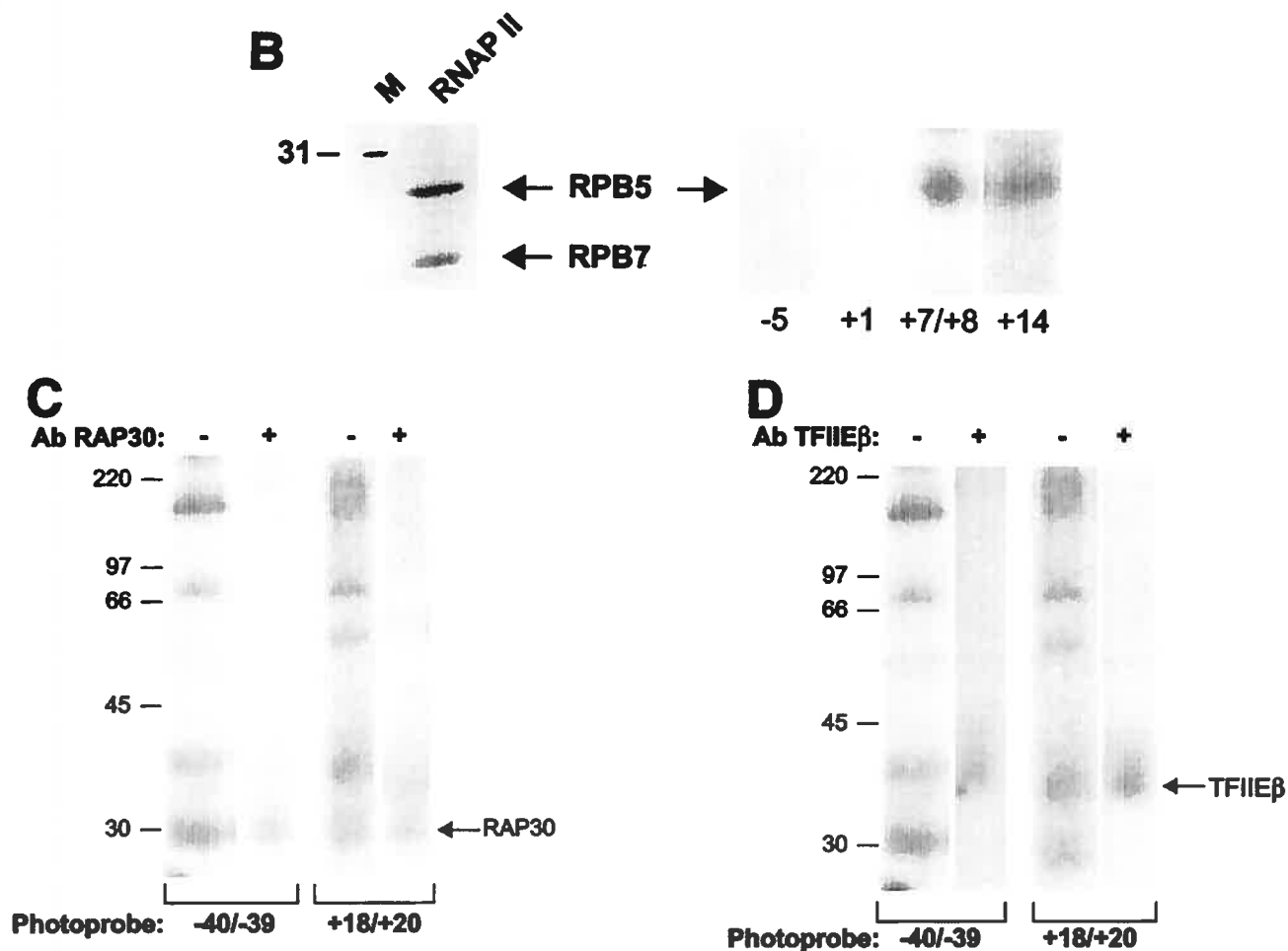


FIG. 4—Continued.

Additional contacts by RAP74 upstream of TATA and by RAP30 downstream of +1 may be the result of tight bending and wrapping of the promoter DNA in the complex and/or of the presence of two RAP74-RAP30 dimers in the preinitiation complex (see Discussion).

Promoter contacts by TFIIIE subunits. The small subunit of TFIIIE, TFIIIE β , cross-linked specifically to four positions downstream of +1 between +7/+8 and +26 and to two positions upstream of TATA between -40/-39 and -31/-29 (see below). The cross-linking of TFIIIE β to nt -40/-39 and +18/+20, two positions ~60 bp apart, has been confirmed by immunoprecipitation of the cross-linked polypeptide using an antibody raised against TFIIIE β (Fig. 4D). Notably, TFIIIE β cross-linked to the same photoprobes as the RAP74 subunit of TFIIIF, suggesting that these two subunits are closely positioned in the preinitiation complex. TFIIIE α also cross-linked to the promoter downstream of +1 between +7/+8 and +26; the cross-link to +26 was weak. Notably, TFIIIE α cross-linked to nucleotide -10. These results indicate that TFIIIE α is the only general initiation factor that closely approaches the promoter in the region where DNA melting is to occur in the open complex.

We also obtained cross-linking of TFIIIE β to photoprobes

-10, -9/-8, -5, and +1. However, and for a reason that remains to be determined, all four photoprobes supported the efficient formation of a low-mobility complex in the absence of TBP (data not shown). When these nonspecific complexes, assembled in the absence of TBP, were UV irradiated, gel purified, and processed through our cross-linking procedure, we obtained strong signals corresponding to TFIIIE β , as if these nonspecific complexes contained an aggregate made of multiple molecules of TFIIIE β . Because we obtained strong nonspecific cross-linking of TFIIIE β in the absence of TBP, we cannot be certain that the weak TFIIIE β cross-linking signals obtained in the presence of TBP at positions -10, -9/-8, -5, and +1 are specific. For this reason, the cross-linking of TFIIIE β between -10 and +1 is considered tentative (Fig. 3).

Role of TFIIIE α in transcription initiation. TFIIIE has previously been shown to stimulate promoter melting *in vitro* (22, 23, 39). The cross-linking of TFIIIE α to nucleotide -10 at the upstream end of the transcription bubble in the open complex supports this role. If the contact by TFIIIE α at position -10 is important to promoter melting, we would predict that TFIIIE α plays a key role during transcription initiation and affects transcription *in vitro* when using templates carrying short single-stranded (premelted) DNA regions. To test this hypothesis, we

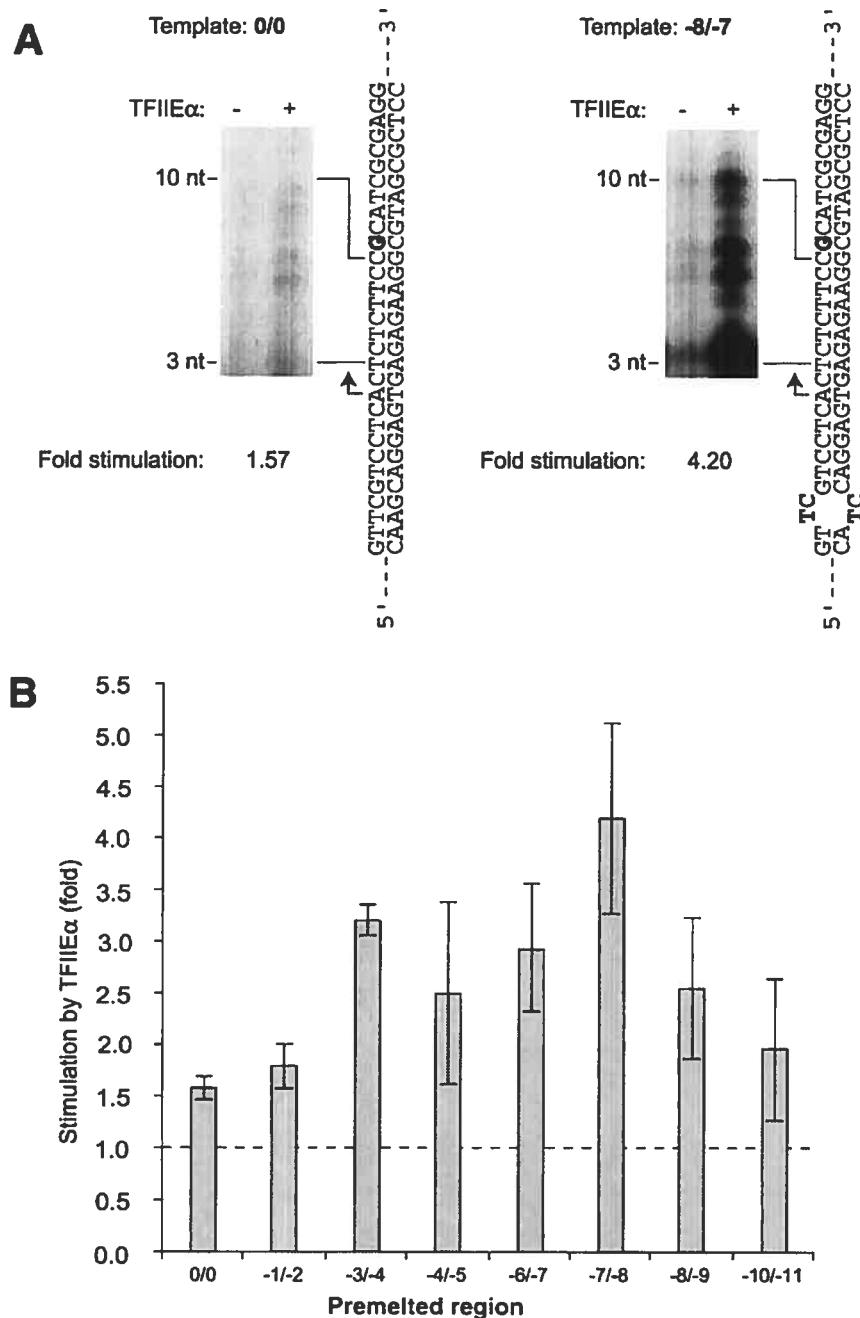


FIG. 5. TFIIEx stimulates transcription initiation. (A) All double-stranded (0/0) and artificially premelted (-8/-7) DNA templates were used in an abortive initiation assay in either the presence (+) or the absence (-) of TFIIEx. The positions of the abortive transcripts are indicated (3 to 10 nt). Stimulation levels obtained through the addition of TFIIEx are indicated. The positions of the first G residue (boldface) and the initiation site (arrows) are indicated. (B) The position of the 2-bp premelted region along the promoter DNA influences the stimulation of transcription by TFIIEx. Templates lacking (0/0) or carrying (-1/-2, -3/-4, -4/-5, -6/-7, -7/-8, -8/-9, and -10/-11) a 2-nt premelted region were used. Stimulation by TFIIEx is indicated. The error bars indicate standard deviations.

used an *in vitro* initiation assay that minimally requires the presence of TBP, TFIIB, TFIIF, and RNAP II (Fig. 5). A short double-stranded oligonucleotide containing the adenovirus major late promoter has been used as a template and was shown to support accurately initiated transcription *in vitro* (30). Compared to a reaction containing TBP, TFIIB, TFIIF,

RNAP II, and TFIIE β , the number of initiation events was slightly increased upon addition of TFIIE α (Fig. 5A, left). Notably, the presence of a 2-nt single-stranded DNA region at -8/-7 resulted in strong stimulation of initiation by TFIIE α (Fig. 5A, right). Alternating this premelted region between -11 and -1 revealed that maximal stimulation occurred in the

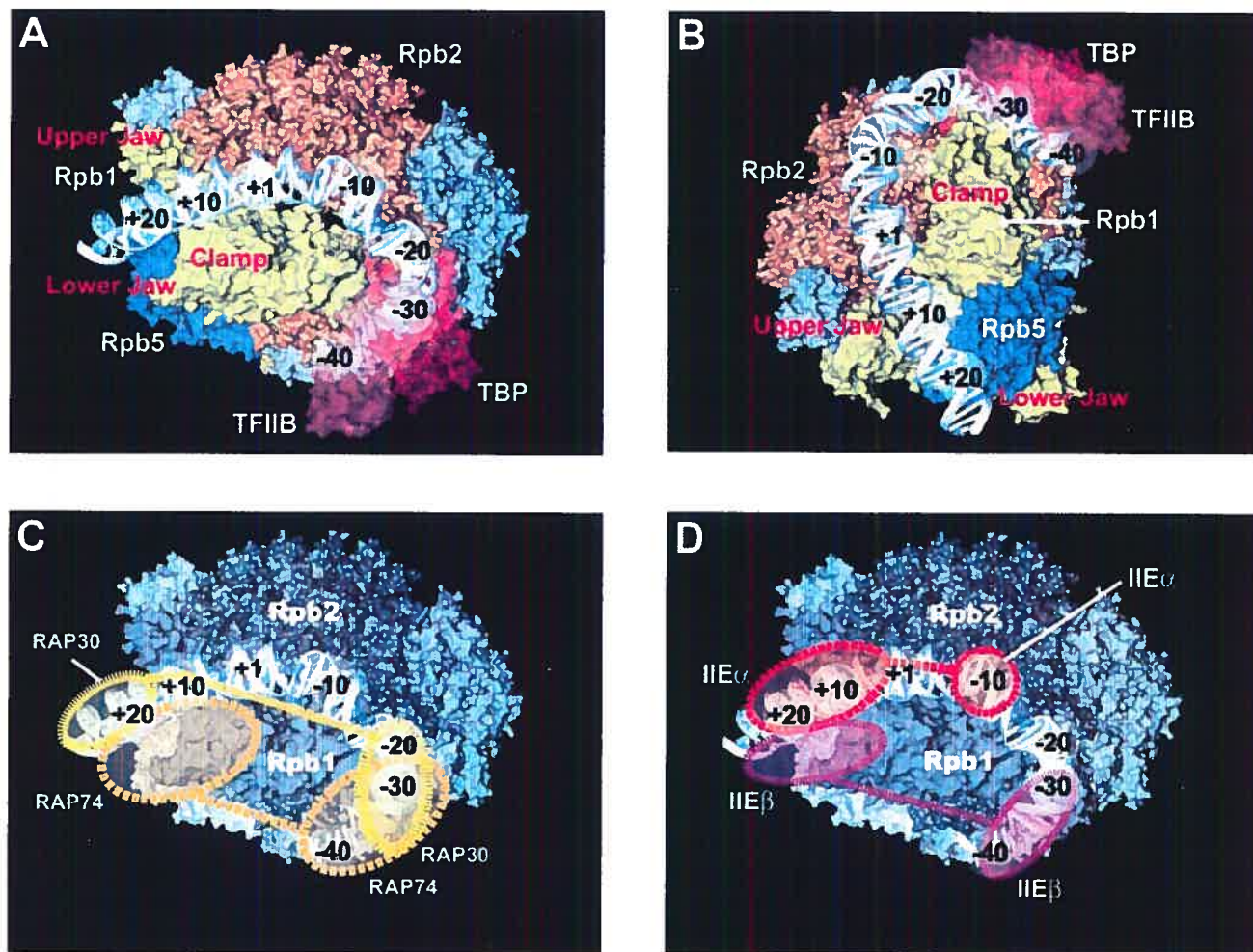


FIG. 6. Topological model of the RNAP II preinitiation complex. (A and B) Top (A) and side (B) of the proposed trajectory of the promoter DNA around the mobile clamp of RNAP II. The path of promoter DNA has been drawn to account for our in-gel photo-cross-linking results (see the text). (C and D) Locations of TFIIF (C) and TFIIE (D) subunits according to our photo-cross-linking results. For each subunit, we obtained cross-linking to two distinct promoter regions (dashed circles). Although our results suggest that two copies each of TFIIE and TFIIF subunits are present within the complex, we cannot rule out the possibility that these factors have two (or more) distinct domains held together by an unfolded linker.

upstream half of the region that was to form the transcription bubble prior to initiation and where TFIIE α cross-links to the promoter (Fig. 5B). These results support the notion that TFIIE α , which closely approaches promoter DNA at position -10 , is important in that it stabilizes promoter melting and stimulates the initiation of transcription.

DISCUSSION

Our previous photo-cross-linking data, along with our electron microscopy studies, provided evidence to support promoter DNA bending and wrapping in the preinitiation complex (13, 45). This model has been supported by a variety of results from biochemical and structural analyses, both in prokaryotes and in eukaryotes (9, 42, 43, 55; see reference 5 for a review). Our previous photo-cross-linking studies were limited due to the fact that the protein-DNA complexes submitted to UV cross-linking had not been purified after assembly, leading

to the possibility that intermediate subcomplexes could contribute to nonrelevant cross-links of some factors to promoter positions. To address this issue and with the goal of increasing the resolution of our topological analysis, we have developed a method allowing us to cross-link protein-DNA complexes after native gel purification of the relevant complexes. Overall, the results indicate that the use of the in-gel photo-cross-linking procedure (i) only slightly reduced the total number of promoter cross-links by the transcription machinery, especially those of TFIIF and TFIIE in the vicinity of the transcription initiation site, and (ii) allowed the cross-linking of TFIIE α in the absence of TFIIF (compare Fig. 3 with reference 8, Fig. 1). Because the TBP-TFIIB-TFIIF-RNAP II-TFIIE-promoter complex studied here was run through a gel prior to protein-DNA photo-cross-linking and because such a complex does not assemble in the absence of TBP (or when a photoprobe with a mutated TATA box is used; data not shown), we are confident that the nonspecific interactions are eliminated. Notably, the

length of the promoter region that is contacted by RNAP II, TFIIF, and TFIIE also remains basically unchanged compared to our previous results. More specifically, RNAP II cross-links to the DNA from -40 to $+26$, a 66-bp region that corresponds to 225 Å of B-form DNA. This expanse is significantly larger than the largest span of RNAP II, which is ~ 140 Å. Furthermore, as we observed in our previous experiments, RAP74, RAP30, and TFIIE β come into close contact with the promoter both upstream of TATA and downstream of $+1$. Together, these results indicate that promoter DNA is wrapped in the preinitiation complex (Fig. 6 shows a model), fully supporting our previous conclusions.

In the model shown in Fig. 6, the trajectory of the promoter DNA has been drawn to best account for our in-gel cross-linking results. Because the TBP-TFIIB-TFIIF-RNAP II-TFIIE-promoter complex is likely to be an intermediate between R_{Pc} (closed complex) and R_{Po} (open complex), we do not expect the promoter DNA to be fully melted in the -9 -to- $+2$ region, with the template strand entering the active site of the enzyme. Open-complex formation, which is obtained in the presence of TFIIF and ATP, is expected to produce an important conformational change in the complex, bringing the template strand of the promoter deep into the active-site cleft. In our topological model, the downstream end of the promoter follows the channel between the two "jaws," where it cross-links to RPB1, RPB2, and RPB5. Between $+1$ and -20 , the DNA closely approaches RPB2 before making a right turn in the region of the TATA element, where TBP and TFIIB also cross-link weakly, to maintain close proximity to RPB1. Tight bending redirects the promoter toward a domain of RPB2 in the -40 region. Only this path of the promoter DNA can easily explain the various contacts observed by the RNAP II subunits. As shown in Fig. 6B, the deduced path of the promoter DNA produces a left-handed loop against the mobile clamp of RNAP II, suggesting a role for the clamp structure in preinitiation mechanisms. Such a left-handed wrapping around the clamp is expected to relax the DNA helix, as opposed to a right-handed wrapping that would induce compaction of the DNA helix. Relaxation of promoter DNA may play a role in producing a partial single-stranded DNA region necessary for the action of the single-stranded helicase of TFIIF during open-complex formation.

Cross-linking of TFIIF and TFIIE both upstream of the TATA element and downstream of the initiation site indicates that both factors are associated with the mobile clamp of RNAP II. Tight wrapping of the promoter DNA around the polymerase clamp may be responsible for contacts by these two factors to promoter positions >60 bp away from each other (e.g., simultaneous promoter contacts at positions -40 and $+26$). Alternatively, or additionally, the putative presence of two molecules of each subunit of both TFIIF and TFIIE may help to explain the distal promoter contacts. Both TFIIF and TFIIE were shown to exist as $\alpha_2\beta_2$ heterotetramers in solution. In support of this idea that two dimers of TFIIF are associated within the preinitiation complex, we have recently purified an RNAP II-containing complex using TAP-tagged RAP30 that contains both a tagged and an endogenous RAP30 molecule (C. Jeronimo, M.-F. Langelier, M. Zeghouf, M. Cojocar, D. Baali, D. Forget, S. Mnaimneh, A. P. Davierwala, J. Pootoolal, M. Chandy, V. Canadien, B. K. Beattie, D. P. Richards, J. L.

Workman, J. R. Hughes, J. Greenblatt, and B. Coulombe, submitted for publication).

The cross-link of TFIIE α to nucleotide -10 is of interest. TFIIE α contains a DNA binding activity that has been anticipated as being important to promoter melting. The region carrying the DNA binding activity is homologous to region 2 of prokaryotic sigma factors (37). Notably, and in full support of our results, region 2 of σA was found to contact position -10 in the RNAP holoenzyme-promoter complex (33). Stimulation of transcription initiation by TFIIE α , a phenomenon that is enhanced when the promoter is artificially premelted in the upstream portion of the melted region of the open complex, suggests that the -10 contact by TFIIE α stabilizes promoter melting, thereby stimulating initiation. Taken together, our results provide novel insights into the fine molecular details of transcriptional initiation mechanisms.

ACKNOWLEDGMENTS

We are grateful to the members of our laboratory for helpful discussions. We thank Bryan Wilkes for his invaluable help in producing the model in Fig. 1B, Diane Bourque for artwork, and Julie Edwards for critical reading of the manuscript. We also thank members of the laboratory of Yvan Guindon for help with chemical procedures.

This work is supported by a grant from the Canadian Institutes for Health Research. M.-F.L. and V.T. hold studentships from the Natural Sciences and Engineering Research Council of Canada and the Fonds de la recherche en santé du Québec. B.C. is a senior scholar from the Fonds de la Recherche en Santé du Québec.

REFERENCES

1. Bartholomew, B., G. A. Kassavetis, B. R. Braun, and E. P. Geiduschek. 1990. The subunit structure of *Saccharomyces cerevisiae* transcription factor IIIc probed with a novel photocrosslinking reagent. *EMBO J.* 9:2197-2205.
2. 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.
3. Burley, S. K., and R. G. Roeder. 1996. Biochemistry and structural biology of transcription factor IID (TFIID). *Annu. Rev. Biochem.* 65:769-799.
4. Campbell, E. A., O. Muzzin, M. Chlenov, J. L. Sun, C. A. Olson, O. Weinman, M. L. Trester-Zedlitz, and S. A. Darst. 2002. Structure of the bacterial RNA polymerase promoter specificity sigma subunit. *Mol. Cell* 9:527-539.
5. Coulombe, B., and Z. F. Burton. 1999. DNA bending and wrapping around RNA polymerase: a "revolutionary" model describing transcriptional mechanisms. *Microbiol. Mol. Biol. Rev.* 63:457-478.
6. 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.
7. 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.
8. Douziach, 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.
9. 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.
10. Feaver, W. J., O. Gileadi, Y. Li, and R. D. Kornberg. 1991. CTD kinase associated with yeast RNA polymerase II initiation factor b. *Cell* 67:1223-1230.
11. Finkelstein, A., C. F. Kostrub, J. Li, D. P. Chavez, B. Q. Wang, S. M. Fang, J. Greenblatt, and Z. F. Burton. 1992. A cDNA encoding RAP74, a general initiation factor for transcription by RNA polymerase II. *Nature* 355:464-467.
12. Forget, D., and B. Coulombe. 2003. Topology of the RNA polymerase II transcription initiation complex: site-specific protein-DNA photo-cross-linking of purified complexes. *Methods Enzymol.* 370:701-712.
13. 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. USA* 94:7150-7155.
14. 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.
15. 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.
 16. Goodrich, J. A., and R. Tjian. 1994. Transcription factors IIE and IIH and ATP hydrolysis direct promoter clearance by RNA polymerase II. *Cell* 77:145–156.
 17. Guex, N., and M. C. Peitsch. 1997. SWISS-MODEL and the Swiss-Pdb-Viewer: an environment for comparative protein modeling. *Electrophoresis* 18:2714–2723.
 18. Ha, I., W. S. Lane, and D. Reinberg. 1991. Cloning of a human gene encoding the general transcription initiation factor IIB. *Nature* 352:689–695.
 19. Hampsey, M. 1998. Molecular genetics of the RNA polymerase II general transcriptional machinery. *Microbiol. Mol. Biol. Rev.* 62:465–503.
 20. Hodo, H. G., III, and S. P. Blatti. 1977. Purification using polyethylenimine precipitation and low molecular weight subunit analyses of calf thymus and wheat germ DNA-dependent RNA polymerase II. *Biochemistry* 16:2334–2343.
 21. 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.
 22. Holstege, F. C., D. Tantin, M. Carey, P. C. van der Vliet, and H. T. Timmers. 1995. The requirement for the basal transcription factor IIE is determined by the helical stability of promoter DNA. *EMBO J.* 14:810–819.
 23. Holstege, F. C., P. C. van der Vliet, and H. T. Timmers. 1996. Opening of an RNA polymerase II promoter occurs in two distinct steps and requires the basal transcription factors IIE and IIH. *EMBO J.* 15:1666–1677.
 24. Humphrey, W., A. Dalke, and K. Schulten. 1996. VMD: visual molecular dynamics. *J. Mol. Graph.* 14:33–38.
 25. Ingles, C. J., M. Shales, W. D. Cress, S. J. Triezenberg, and J. Greenblatt. 1991. Reduced binding of TFIID to transcriptionally compromised mutants of VP16. *Nature* 351:588–590.
 26. Jiang, Y., S. J. Triezenberg, and J. D. Gralla. 1994. Defective transcriptional activation by diverse VP16 mutants associated with a common inability to form open promoter complexes. *J. Biol. Chem.* 269:5505–5508.
 27. Kim, T. K., R. H. Ebricht, and D. Reinberg. 2000. Mechanism of ATP-dependent promoter melting by transcription factor IIH. *Science* 288:1418–1422.
 28. Kim, T. K., T. Lagrange, Y. H. Wang, J. D. Griffith, D. Reinberg, and R. H. Ebricht. 1997. Trajectory of DNA in the RNA polymerase II transcription preinitiation complex. *Proc. Natl. Acad. Sci. USA* 94:12268–12273.
 29. Lagrange, T., A. N. Kapanidis, H. Tang, D. Reinberg, and R. H. Ebricht. 1998. New core promoter element in RNA polymerase II-dependent transcription: sequence-specific DNA binding by transcription factor IIB. *Genes Dev.* 12:34–44.
 30. Langelier, M. F., D. Forget, A. Rojas, Y. Portier, 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.
 31. Lu, H., O. Flores, R. Weinmann, and D. Reinberg. 1991. The nonphosphorylated form of RNA polymerase II preferentially associates with the preinitiation complex. *Proc. Natl. Acad. Sci. USA* 88:10004–10008.
 32. 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.
 33. 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.
 34. Murakami, K. S., S. Masuda, and S. A. Darst. 2002. Structural basis of transcription initiation: RNA polymerase holoenzyme at 4 Å resolution. *Science* 296:1280–1284.
 35. Naryshkin, N., Y. Kim, Q. Dong, and R. H. Ebricht. 2001. Site-specific protein-DNA photocrosslinking. Analysis of bacterial transcription initiation complexes. *Methods Mol. Biol.* 148:337–361.
 36. Oelgeschlager, T., C. M. Chiang, and R. G. Roeder. 1996. Topology and reorganization of a human TFIID-promoter complex. *Nature* 382:735–738.
 37. 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.
 38. Orphanides, G., T. Lagrange, and D. Reinberg. 1996. The general transcription factors of RNA polymerase II. *Genes Dev.* 10:2657–2683.
 39. Pan, G., and J. Greenblatt. 1994. Initiation of transcription by RNA polymerase II is limited by melting of the promoter DNA in the region immediately upstream of the initiation site. *J. Biol. Chem.* 269:30101–30104.
 40. Persinger, J., and B. Bartholomew. 2001. Site-directed DNA photoaffinity labeling of RNA polymerase III transcription complexes. *Methods Mol. Biol.* 148:363–381.
 41. 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.
 42. 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.
 43. Rivetti, C., M. Guthold, and C. Bustamante. 1999. Wrapping of DNA around the *E. coli* RNA polymerase open promoter complex. *EMBO J.* 18:4464–4475.
 44. Robert, F., and B. Coulombe. 2001. Use of site-specific protein-DNA photocrosslinking to analyze the molecular organization of the RNA polymerase II initiation complex. *Methods Mol. Biol.* 148:383–393.
 45. 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.
 46. Roy, R., J. P. Adamczewski, T. Seroz, W. Vermeulen, J. P. Tassin, L. Schaeffer, E. A. Nigg, J. H. Hoeijmakers, and J. M. Egly. 1994. The MO15 cell cycle kinase is associated with the TFIIH transcription-DNA repair factor. *Cell* 79:1093–1101.
 47. 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.
 48. 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.
 49. Serizawa, H., R. C. Conaway, and J. W. Conaway. 1992. A carboxyl-terminal domain kinase associated with RNA polymerase II transcription factor delta from rat liver. *Proc. Natl. Acad. Sci. USA* 89:7476–7480.
 50. Smale, S. T., and D. Baltimore. 1989. The “initiator” as a transcription control element. *Cell* 57:103–113.
 51. Spangler, L., X. Wang, J. W. Conaway, R. C. Conaway, and A. Dvir. 2001. TFIIH action in transcription initiation and promoter escape requires distinct regions of downstream promoter DNA. *Proc. Natl. Acad. Sci. USA* 98:5544–5549.
 52. 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.
 53. 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.
 54. 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.
 55. Vassilyev, D. G., S. Sekine, O. Laptchenko, 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.
 56. Wang, Z., S. Buratowski, J. Q. Svejstrup, W. J. Feaver, X. Wu, R. D. Kornberg, T. F. Donahue, and E. C. Friedberg. 1995. The yeast TFB1 and SSL1 genes, which encode subunits of transcription factor IIH, are required for nucleotide excision repair and RNA polymerase II transcription. *Mol. Cell. Biol.* 15:2288–2293.
 57. Yan, M., and J. D. Gralla. 1997. Multiple ATP-dependent steps in RNA polymerase II promoter melting and initiation. *EMBO J.* 16:7457–7467.
 58. Young, R. A. 1991. RNA polymerase II. *Annu. Rev. Biochem.* 60:689–715.
 59. Zhang, G., E. A. Campbell, L. Minakhin, C. Richter, K. Severinov, and S. A. Darst. 1999. Crystal structure of *Thermus aquaticus* core RNA polymerase at 3.3 Å resolution. *Cell* 98:811–824.

CHAPITRE 4

RPAPI, UNE NOUVELLE PROTÉINE ASSOCIÉE À L'ARN POLYMÉRASE II PURIFIÉE À PARTIR DE SOUS-UNITÉS RECOMBINANTES SAUVAGES ET MUTANTES

4.1 Préambule

Dans cet article pour lequel je suis co-auteure principale, j'ai purifié l'ARN polymérase II portant une étiquette TAP sur la sous-unité Rpb2 (Rpb2-TAP) et j'ai ainsi participé à l'identification des composants du complexe étudié (figure 2B). J'ai participé à la construction de la lignée Rpb11-TAP et j'ai montré pour la première fois que l'ARN polymérase II purifiée avec la méthode TAP était active en transcription *in vitro*. La figure montrée dans l'article (figure 5A) est cependant celle de Célia Jeronimo. J'ai aussi montré que le complexe purifié pouvait transcrire en absence de TFIIB et TFIIF recombinant, indiquant que ces derniers étaient des constituants fonctionnels du complexe (figure 5B). J'ai réalisé le clonage, l'établissement des lignées stables et la purification des deux mutants de l'ARN polymérase II étudiés (figure 6B). Finalement, j'ai fait tous les essais permettant la caractérisation de ces mutants (figures 6 et 7) à l'exception des expériences d'immunoprécipitation de la chromatine (figure 8). J'ai participé à la rédaction du manuscrit avec Célia Jeronimo et le Dr Benoit Coulombe.

Célia Jeronimo a réalisé les expériences présentées aux figures 1B et 2 (la figure 2B a été faite en collaboration avec Dania Baali (Rpb4 et Rpb7) et moi (Rpb2)). La figure 3A a été réalisée par Célia Jeronimo et Dominique Bergeron et la figure 3B, par le groupe du Dr Timothy Hughes, incluant Sanie Mnaimneh, Jeff Pootoolal et Armaity Davierwala. La figure 4A a été faite par Diane Forget, la figure 4B et 8 par Marilena Cojocar et la figure 4C par Mark Chandy, étudiant chez le Dr Jerry Workman. Veronica Canadien, Bryan Beattie et Dawn Richards ont effectué la spectrométrie de masse, qui a permis

l'identification des constituants du complexe (figure 2B). Mahel Zeghouf a réalisé le clonage et l'établissement de la lignée cellulaire stable pour la lignée Rpb11-TAP.

4.2 Article

Jeronimo, C.*, Langelier, M.-F.*, Zeghouf, M., Cojocaru, M., Bergeron, D., Baali, D., Forget, D., Mnaimneh, S., Davierwala, A.P., Pootoolal, J., Chandy, M., Canadien, V., Beattie, B.K., Richards, D.P., Workman, J.L., Hughes, T.R., Greenblatt, J. and Coulombe, B. (2004) RPAP1, a novel human RNA polymerase II-associated protein affinity purified using recombinant wild type and mutated polymerase subunits, **Mol. Cell. Biol.**, *24*(16):7043-7058.

* Ces deux auteures ont contribué de façon équivalente à cet article.

RPAP1, a Novel Human RNA Polymerase II-Associated Protein Affinity Purified with Recombinant Wild-Type and Mutated Polymerase Subunits

Célia Jeronimo,^{1†} Marie-France Langelier,^{1†} Mahel Zeghouf,² Marilena Cojocar,¹ Dominique Bergeron,¹ Dania Baali,¹ Diane Forget,¹ Sanie Mnaimneh,² Armaity P. Davierwala,² Jeff Pootoolal,² Mark Chandy,^{3,4} Veronica Canadien,⁵ Bryan K. Beattie,⁵ Dawn P. Richards,⁵ Jerry L. Workman,³ Timothy R. Hughes,² Jack Greenblatt,² and Benoit Coulombe^{1*}

Laboratory of Gene Transcription, Institut de Recherches Cliniques de Montréal, Montréal, Québec H2W 1R7,¹ Banting and Best Department of Medical Research, University of Toronto, Toronto, Ontario M5G 1L6,² and Affinium Pharmaceuticals, Toronto, Ontario M5J 1V6,⁵ Canada; Stowers Institute for Medical Research, Kansas City, Missouri 64110³; and Pennsylvania State University College of Medicine, Hershey, Pennsylvania 17033⁴

Received 24 December 2003/Returned for modification 9 February 2004/Accepted 14 May 2004

We have programmed human cells to express physiological levels of recombinant RNA polymerase II (RNAPII) subunits carrying tandem affinity purification (TAP) tags. Double-affinity chromatography allowed for the simple and efficient isolation of a complex containing all 12 RNAPII subunits, the general transcription factors TFIIB and TFIIF, the RNAPII phosphatase Fcp1, and a novel 153-kDa polypeptide of unknown function that we named RNAPII-associated protein 1 (RPAP1). The TAP-tagged RNAPII complex is functionally active both *in vitro* and *in vivo*. A role for RPAP1 in RNAPII transcription was established by shutting off the synthesis of Ydr527wp, a *Saccharomyces cerevisiae* protein homologous to RPAP1, and demonstrating that changes in global gene expression were similar to those caused by the loss of the yeast RNAPII subunit Rpb11. We also used TAP-tagged Rpb2 with mutations in fork loop 1 and switch 3, two structural elements located strategically within the active center, to start addressing the roles of these elements in the interaction of the enzyme with the template DNA during the transcription reaction.

RNA polymerase II (RNAPII) is the multisubunit enzyme that synthesizes all mRNA precursors in eukaryotes. RNAPII is highly conserved among species, and in humans, RNAPII consists of 12 subunits, named Rpb1 to Rpb12 (16, 88). The two largest subunits, Rpb1 (220 kDa) and Rpb2 (140 kDa), form the enzyme's catalytic center and are homologous to the β' and β subunits of bacterial RNAP, respectively. Five subunits, Rpb5, Rpb6, Rpb8, Rpb10, and Rpb12, are also found in RNAPI and RNAPIII. Rpb3 and Rpb11 are homologous to the α_2 homodimer involved in bacterial RNAP assembly. Rpb9 was attributed a role in elongation through its action at DNA arrest sites (3). In *Saccharomyces cerevisiae*, Rpb4 and Rpb7 form a subcomplex that can dissociate from the enzyme upon changes in environmental conditions. Under active growth conditions, most yeast RNAPII molecules do not contain the Rpb4-Rpb7 dimer, which primarily associates during the stationary phase or following stress (11). Functional studies of human RNAPII have been limited due to the lack of appropriate systems for purifying variant forms of the human enzyme.

The availability of crystal structures of both yeast (17, 18, 26, 30) and bacterial (8, 57, 58, 84, 89) RNAPs has been invaluable

for understanding many of the molecular features of the transcription reaction. For example, the structure of elongating RNAPII has revealed the positioning of the RNA-DNA duplex during the transcription reaction (30). The structures available support a model in which the DNA enters the enzyme through a channel formed by a pair of "jaws" before accessing a deep cleft, at the bottom of which is buried the Mg^{2+} -ion-containing active site; the DNA then turns by about 90° along a wall where the upstream end exits the enzyme (30). Many loops and helices either directly contact or closely approach the RNA-DNA duplex, thus suggesting putative functions for these structural elements in the transcription reaction. Again, the lack of an efficient method for purifying variant forms of RNAPII with mutations within specific structural elements has precluded any detailed structure-function analysis of human RNAPII.

The activity of RNAPII at the various stages of the transcription reaction is regulated by several proteins, including the general transcription factors (14, 32, 65). TATA-binding protein (TBP), the TATA box-binding subunit of TFIID, recognizes the TATA box of promoters, and its binding induces a bend of approximately 80° in the DNA (45, 46). TFIIB associates with the TBP-promoter complex and participates in RNAPII recruitment (54). TFIIF, composed of subunits RAP74 and RAP30, binds directly to RNAPII and also helps to recruit the enzyme to the promoter (15, 80). TFIIF can induce further bending and wrapping of the promoter DNA around RNAPII during preinitiation complex formation (72).

* Corresponding author. Mailing address: Laboratory of Gene Transcription, Institut de Recherches Cliniques de Montréal, 110 avenue des Pins Ouest, Montréal, Québec, Canada H2W 1R7. Phone: (514) 987-5662. Fax: (514) 987-5663. [REDACTED]

† C.J. and M.-F.L. contributed equally to this work.

Both TFIIE and TFIIH were shown to participate in promoter melting at the transcription initiation site (38, 66). TFIIE, which is also composed of two subunits, TFIIE56/ α and TFIIE34/ β , stimulates the helicase and kinase activities of TFIIH (53, 62, 63). The helicase activity of TFIIH is required to fully open the promoter DNA between nucleotides (nt) -9 and $+2$ (37, 39, 44, 83), and its kinase activity phosphorylates the carboxy-terminal domain (CTD) of Rpb1 (21, 53, 75–77). The phosphorylation of the CTD on Ser5 by TFIIH has been shown to be required for the transition from initiation to early elongation and allows for the recruitment of the capping enzyme (35, 73). In yeast, the Ctk1 kinase phosphorylates the CTD on Ser2 at a later stage of the elongation phase (10), leading to the recruitment of mRNA processing factors (78). The CTD phosphatase Fcp1 is able to dephosphorylate the Rpb1 CTD on Ser2 in order to recycle RNAPII for reinitiation (10, 48, 49).

Unlike the case for all of the general transcription factors and the prokaryotic RNAP, it has been impossible to date to purify functionally active forms of mammalian RNAPII reconstituted from its recombinant subunits. We now report the purification of a functionally active human RNAPII by the use of doubly tagged subunits expressed in mammalian cells. Using the TAP procedure (71), we isolated a human RNAPII-containing complex which comprises the 12 subunits of core RNAPII, TFIIB, TFIIF, Fcp1, and RPAP1, a novel polypeptide of unknown function. The tagged RNAPII complex is functionally active both *in vitro* and *in vivo*. We capitalized on this purification procedure to generate two variant forms of human RNAPII by using TAP-tagged Rpb2. The first mutant has a deletion in fork loop 1, a structure that is expected to participate in maintaining the open state of the transcription bubble in elongating RNAPII (18, 30) or during RNA-DNA strand separation (85). The second mutant carries a triple alanine substitution in switch 3, a loop that directly contacts the DNA template within the DNA-RNA hybrid. Our results allowed us to detail the roles of fork loop 1 and switch 3 in transcriptional mechanisms both *in vitro* and *in vivo*.

MATERIALS AND METHODS

Protein factors and antibodies. Recombinant yeast TBP (42), human TFIIB (31), RAP74 and RAP30 (22), TFIIE56/ α and TFIIE34/ β (64, 67, 81), highly purified TFIIH (20, 27), and calf thymus RNAPII (36) were purified as previously described. Antibodies raised against Rpb1 (N-20) were obtained from Santa Cruz Biotechnology.

TAP of human RNAPII-containing complex. The cDNAs encoding human Rpb2 (kindly provided by Marc Vigneron, IGBMC, Strasbourg, France) (1), Rpb4 (Invitrogen; accession number BE883306), Rpb7 (Invitrogen; accession number BE256210), Rpb11 (Invitrogen; accession number AA814184), TFIIB (31), and RAP30 (22) were cloned into the mammalian expression vector pMZI (61) in such a way that each polypeptide carries a tandem affinity purification (TAP) tag at its C terminus (71). Stable cell lines carrying these plasmids in EcR-293 cells (Invitrogen) were produced by transfection, using the calcium phosphate method. Induction for 24 h with 1 to 3 μ M ponasterone A, an ecdysone analog, was used to express the TAP-tagged subunits of RNAPII, TFIIB, and RAP30 at near-physiological levels. The TAP-tagged complexes were then purified as previously described (71). The TAP eluates were concentrated and stored at -80°C until use.

Protein identification. The TAP eluates were run in sodium dodecyl sulfate (SDS) gels and silver stained. The protein bands were excised from the gels, reduced, alkylated, and subjected to in-gel tryptic digestion as previously described (50). The resulting tryptic peptides were purified and identified by matrix-assisted laser desorption/ionization-time of flight (MALDI-TOF) mass spectrometry (50).

DNA microarray analysis of Tet-promoter mutants. Tet-promoter alleles were constructed as previously described (86). Mutant and isogenic wild-type cultures were grown in parallel in SC medium with 10 μ g of doxycycline/ml for a total of 22 h. Dilutions were made at 16 h to ensure that mutant and wild-type optical densities matched as closely as possible to 10^7 cells/ml at harvest. The cells were pelleted by a 2-min centrifugation in a room-temperature table-top centrifuge and then frozen immediately in liquid nitrogen. Total RNAs were prepared by hot acidic phenol-chloroform extraction followed by ethanol precipitation. Poly(A)⁺ RNA purification and cDNA labeling were performed as described previously (55). Microarrays were composed of 70-mer oligonucleotides, each specific to a different yeast gene (Operon Technologies), which were spotted onto poly-L-lysine-coated slides. Spotting buffer, blocking, hybridization, and washing were according to a procedure described by Hegde and collaborators (34). Each array was normalized by using grid-by-grid lowess smoothing (87). All measurements were taken in fluor-reversed pairs [i.e., each time a mutant was analyzed, it was hybridized to two arrays, and data were combined by averaging the normalized log(ratio)].

ChIP. Chromatin immunoprecipitation (ChIP) assays were performed as previously described (60, 82), with the following minor modifications. Approximately 2×10^8 cells were cross-linked by the use of formaldehyde, and chromatin fragments of an average size of 0.3 to 0.5 kbp were prepared by four sonication treatments (Fisher Sonic Dismembrator; 20-s burst at 50% maximum power). For the immunoprecipitation step, immunoglobulin G (IgG)-Sepharose beads (Amersham Biosciences) were first blocked with a 5-mg/ml solution of bovine serum albumin in phosphate-buffered saline and then incubated with the sample for 1 h. After reversal of the cross-links, PCR analyses were performed with both the immunoprecipitated and input DNAs. Primer sets designed for amplifying both the promoter and the 5' untranslated region of the CCNA1 gene were used.

For the experiments shown in Fig. 8, relative occupancy was evaluated by quantitative PCR (Q-PCR) as previously described (5), with minor modifications. The quantification of the target DNA in the immunoprecipitated sample was carried out by generating a standard curve with a fivefold dilution series of the nonimmunoprecipitated sample (input DNA) for each couple of primer sets and DNA. The apparent immunoprecipitation efficiency was calculated by dividing the amount of the DNA fragments spanning the promoter or the coding region of the specified genes by the amount of the DNA fragment from the control region in the immunoprecipitated sample. The following cell lines were analyzed: Rpb2 wt, Rpb2 fork1 Δ 458-459, Rpb2 sw3-1078, and Rpb2 fork1 Δ 453-465. Q-PCR was performed by using the Mx4000 Multiplex Quantitative PCR system (Stratagene). The PCR protocol was applied according to the manufacturer's specifications (Qiagen). Primers were designed for both the promoter and transcribed regions of the GNB2L1 (positions -200 to $+55$ and $+998$ to $+1165$, respectively) and FTL (positions -157 to $+33$ and $+955$ to $+1176$, respectively) genes and also for an internal sequence of the expressed sequence tag located at chr17:49926700-49927999 as a control. Oligonucleotide sequences are available upon request. All amplified DNA fragments were unique, as determined by a BLAST search on the April 2003 assembly version of the UCSC Genome Browser (<http://www.genome.ucsc.edu>), by a melting curve analysis, and by agarose gel electrophoresis (data not shown).

Gel mobility shift assays. Gel mobility shift assays were performed as previously described (52). Complexes were assembled by using highly purified TBP (80 ng), TFIIB (60 ng), RAP30 (320 ng), RAP74 (640 ng), TFIIE34 (160 ng), TFIIE56 (320 ng), and calf thymus RNAPII or the TAP-tagged human RNAPII complex (the amounts of RNAPII used in the various experiments are detailed in the figure legends).

Histone tail peptide binding assay. Histone tail peptide binding reactions were performed to determine whether the complex binds specifically to acetylated histone tails. For each reaction, 10 ng of biotinylated H3 peptides was bound to streptavidin-conjugated magnetic beads at 25°C for 30 min. The beads were then incubated with 1 nM RNAPII complex (Rpb11 TAP purified or calf thymus) for 1 h at 30°C . The beads were pelleted with a magnetic particle concentrator (MPC-S; Dynal) and the supernatant was removed. After being washed twice, the beads were resuspended in SDS-polyacrylamide gel electrophoresis (SDS-PAGE) lysis buffer and subjected to SDS-PAGE analysis. The complex was detected by immunoblotting with an anti-Rpb1 antibody (N-20; Santa Cruz).

Transcription assay. Transcription reactions were performed as described previously (6). TBP (30 ng), TFIIB (30 ng), RAP30 (30 ng), RAP74 (65 ng), TFIIE34 (40 ng), TFIIE56 (60 ng), TFIIH (4 ng), and calf thymus RNAPII or the TAP-tagged human RNAPII complex (the amounts of RNAPII used for the various experiments are detailed in the figure legends) were incubated with 500 ng of linearized DNA template containing the adenovirus major late (AdML) promoter from nt -50 to $+10$ fused to a G-less cassette. Under these conditions, a 391-nt runoff transcript was produced.

Abortive initiation assay. Abortive initiation assays were performed as described previously (52), with the following modifications. The DNA template (12 ng), either premelted in the -9 to $+2$ region or fully double-stranded, was incubated with TBP (60 ng), TFIIB (30 ng), RAP30 (30 ng), RAP74 (65 ng), TFIIE34 (40 ng), TFIIE56 (60 ng), and the TAP-tagged human RNAPII complex (the amounts of RNAPII used for the various experiments are detailed in the figure legends) in a 20- μ l reaction mixture containing 750 μ M ATP, 750 μ M CTP, 10 μ M UTP, 2.5 μ Ci of [α - 32 P]UTP, 12.5 mM MgCl₂, 3 mM EGTA, and 0.84 U of RNase inhibitor/ml. The transcripts were analyzed in a 23% polyacrylamide denaturing gel containing 7 M urea.

Elongation assay. The template was prepared by annealing two complementary oligonucleotides of random sequences, one of which carried an extension of 15 CMP as described previously (66). Typically, the template (9.3 ng) was incubated with 88 ng of TAP-tagged human RNAPII Rpb2 wt or 440 ng of TAP-tagged human RNAPII Rpb2 fork I (Δ 458-459)-TAP, and the reactions were performed as described previously (66). The transcripts were analyzed in 18% polyacrylamide-9% urea denaturing gels.

Nucleotide sequence accession number. The RPAP1 name and sequence were registered in the HUGO database under accession number NP_056355.

RESULTS

TAP of human RNAPII-containing complex. We have adapted the TAP procedure (69, 71), originally developed to isolate protein complexes from yeast, for the purification of human RNAPII and general transcription factors. Native conditions were maintained during the TAP procedure through the use of buffers with a low salt concentration (0.1 M Na⁺) and a neutral pH. As shown in Fig. 1A, which summarizes the TAP procedure, full-length cDNAs encoding polypeptides of interest were cloned in an ecdysone-inducible mammalian expression vector (61). The vector was engineered to encode polypeptides carrying the following at the C terminus: a TAP tag consisting of two copies of the IgG-binding domain of *Staphylococcus aureus* protein A, a cleavage site for the tobacco etch virus protease, and a calmodulin-binding peptide (CBP) (71). The tagged polypeptides were expressed in EcR-293 cells at near-physiological levels after a 24-h induction with the ecdysone analog ponasterone A. The concentration of ponasterone A required to obtain near-physiological levels of expression of the tagged polypeptides was determined in a pilot experiment in which we compared the levels of TAP-tagged polypeptides with the endogenous polypeptide by Western blotting (Fig. 1B). Whole-cell extracts were prepared from the ponasterone A-induced cell lines and were used in the double-affinity chromatography procedure, which consisted of successive steps with IgG-Sepharose and calmodulin beads (71). The eluates were frozen and analyzed by different assays.

We first expressed the Rpb11 subunit of RNAPII in EcR-293 cells and used the lysate of ponasterone A-induced cells for double-affinity purification. The eluate was subjected to SDS-PAGE and the bands were identified by mass spectrometry. A total of 17 polypeptides that are not present in mock-induced cells and that represent major components of the eluate were identified (Fig. 2A). The other major bands that can be visualized on the gel have all been identified and correspond to proteins that bind nonspecifically to our affinity columns (data not shown). The complex contained the 12 subunits of RNAPII, TFIIB, both subunits of TFIIF (RAP30 and RAP74), Fcp1, the TFIIF-associated CTD phosphatase, and a novel human polypeptide that we named RPAP1.

To validate the existence of a unique 17-subunit complex containing RNAPII, we tagged other subunits of the complex,

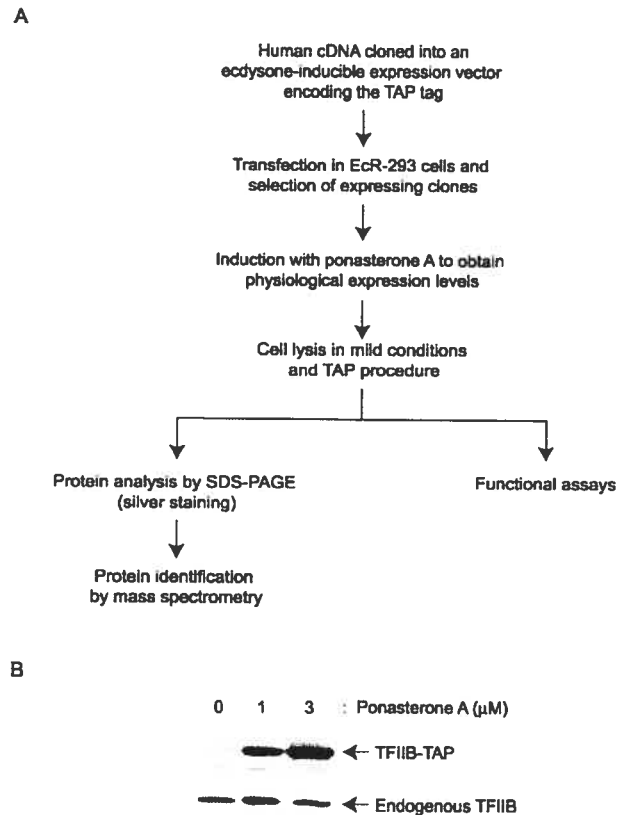
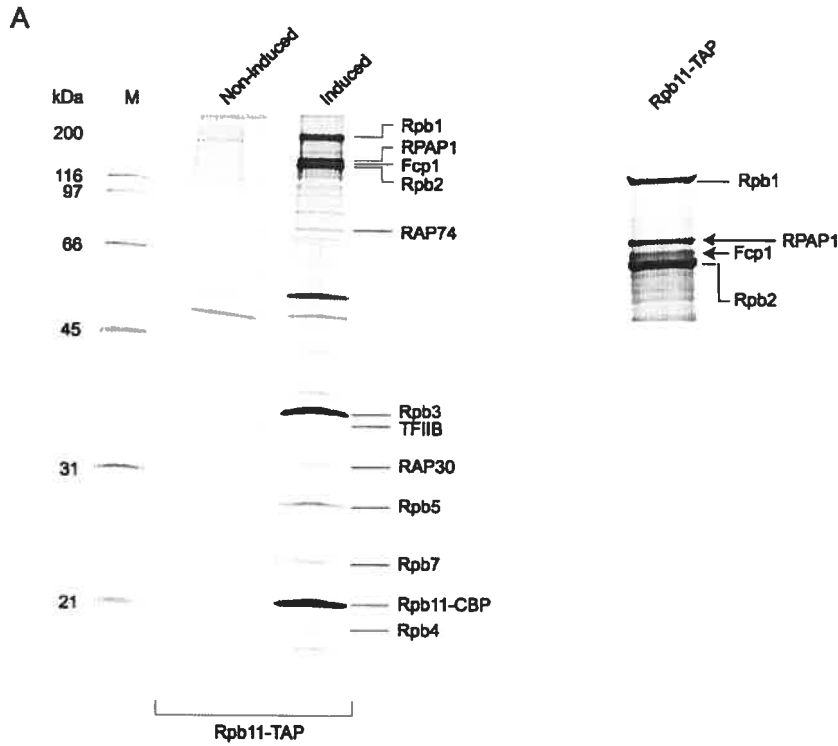


FIG. 1. TAP of human transcription factors. (A) Overview of the TAP procedure. (B) Data from a pilot experiment comparing the expression levels of a TAP-tagged polypeptide and its endogenous counterpart upon induction with ponasterone A. TFIIB-TAP expression in EcR-293 cells after induction with different concentrations of ponasterone A (0, 1, and 3 μ M) for 24 h was compared to that of endogenous TFIIB by Western blotting with an antibody raised against TFIIB.

including Rpb2, Rpb4, Rpb7, TFIIB, and RAP30, expressed them in stably transfected EcR-293 cells, and used them as baits in the TAP procedure. The eluates were analyzed by silver staining of SDS-PAGE gels, Western blotting, and/or mass spectrometry (Fig. 2B). The results indicated that all 17 polypeptides are bona fide components of the human RNAPII complex. Notably, the eluate prepared with TAP-tagged RAP30 contained both tagged RAP30 with its CBP residual domain and nontagged RAP30 (Fig. 2C). The presence of RAP30 lacking the CBP domain was confirmed by Western blotting (data not shown). These results indicate that two molecules of RAP30 are present within the complex, suggesting that TFIIF is associated with RNAPII as an $\alpha_2\beta_2$ heterotetramer (see Discussion). Finally, our Western blot analysis with an antibody specifically directed against the N-terminal part of Rpb1 (N-20) revealed that the tagged RNAPII complex contains the hypophosphorylated form of RNAPII (IIA form) (Fig. 2D).

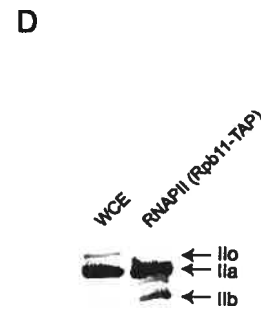
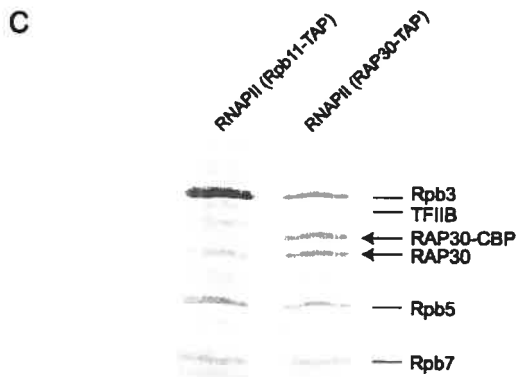
RPAP1 is structurally and functionally related to human RNAPII. RPAP1 is a previously identified human gene product (DKFZP727M111 protein) of unknown function. BLAST se-



B

Cell lines expressing TAP-tagged polypeptides

Polypeptides	Rpb11	Rpb2	Rpb4	Rpb7	TFIIIB	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
TFIIIB	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



quence searches against publicly available databases revealed highly significant alignments over the full length of the RPAP1 sequence with proteins of unknown function from *Mus musculus* (accession no. BAC65787), *Rattus norvegicus* (accession no. XP_230480), and to a lesser extent, *Drosophila melanogaster* (accession no. NP_648573 and CG32104-PB). In addition, these proteins show significant similarities to the *Saccharomyces cerevisiae* Ydr527wp protein (accession no. NP_010816), which is required for cell viability and was shown to bind the Rpb10 subunit of RNAPII in a two-hybrid screen (28, 43). The positions of the different alignments are shown in Fig. 3A. Notably, no particular motif or domain could be predicted with high confidence on the RPAP1 sequence. However, in view of the participation of RPAP1 in the RNAPII complex, an IMPALA analysis against the BLOCK database (74) revealed a very weak similarity ($E = 0.23$) between the bromodomain-containing protein 4 (Brd4) and the RPAP1 region spanning amino acids 196 to 312 (see below). We also performed a sequence-structure comparison analysis by using the structure-guided alignment program FUGUE against the HOMSTRAD database (56). The result suggests that RPAP1 belongs to the ARM repeat protein superfamily (12). Indeed, the RPAP1 region spanning amino acids 508 to 960 shows an alignment with ARM with a Z score of 4.08 (corresponding to 95% confidence). Interestingly, both Ydr527wp and the RPAP1 *D. melanogaster* homologue were also predicted, as reported in the *Saccharomyces* Genome and FlyBase databases, respectively, to belong to the ARM repeat superfamily (<http://genome-www4.stanford.edu/cgi-bin/SGD/locus.pl?locus=YDR527W> and <http://flybase.bio.indiana.edu/bin/fbidq.html?FBgn0052104>).

We created a Tet-promoter allele of the yeast *YDR527w* gene, turned off gene expression by adding doxycycline to the growth medium for 24 h, and compared gene expression in this mutant to expression patterns from 121 other Tet-promoter mutants affecting genes implicated in diverse biological processes (S. Mnaimneh, A. P. Davierwala, J. Haynes, J. Moffat, W. T. Peng, W. Zhang, X. Yang, J. Pootoolal, G. Chua, A. Lopez, M. Trochesset, D. Morse, N. J. Krogan, S. L. Hiley, Z. Li, Q. Morris, J. Grigull, N. Mitsakakis, C. J. Roberts, J. F. Greenblatt, C. Boone, C. A. Kaiser, B. J. Andrews, and T. R. Hughes, submitted for publication). Among all of these, the effect of the *tet-YDR527w* allele correlated most strongly with that of *tet-RPB11* (Fig. 3B). A clustering analysis did not reveal any specific group of functionally related transcripts that were up- or down-regulated uniquely in these two mutants, supporting the idea that the expression pattern observed is a conse-

quence of an overall perturbation in transcription rather than a disruption of any individual pathway.

To further define the interaction between RPAP1 and RNAPII, we cloned the human RPAP1 cDNA and derived a human cell line that can be induced to express TAP-tagged RPAP1. An analysis of the TAP-tagged RPAP1 eluate revealed the presence of two RNAPII subunits, Rpb2 and Rpb3 (Fig. 3A, right panel). These results further demonstrate the physical interaction between RPAP1 and RNAPII. Recent TAP-tagging experiments in yeast indicated that Ydr527wp associates with Rpb2, Rpb3, and Rpb11 (33). Because Ydr527wp and RPAP1 are both physically and functionally related, the 50-kDa yeast protein encoded by the *YDR527W* gene was named RBA50 (the name was registered in the *Saccharomyces* Genome Database).

The TAP-tagged RNAPII complex can be recruited to promoters both in vitro and in vivo and can bind to acetylated histone H3. In order to determine whether the TAP-tagged RNAPII complex is functional, we first verified its ability to enter a preinitiation complex in vitro. The electrophoretic mobility shift assay (EMSA) results in Fig. 4A show that the human RNAPII complex can form a preinitiation complex in the presence of TBP, TFIIB, TFIIF, and TFIIE on a radiolabeled probe carrying the AdML promoter. Compared to an equivalent preinitiation complex assembled with highly purified calf thymus RNAPII, the TAP-tagged RNAPII complex had a slightly lower mobility. Because our calf thymus RNAPII contains neither Fcp1 nor RPAP1 (data not shown), our results indicate that the preinitiation complex assembled with the TAP-tagged human RNAPII complex includes at least one of these two polypeptides (Fcp1 and/or RPAP1).

ChIP experiments were used to determine whether the tagged RNAPII complex can also be recruited to promoters in vivo. As shown in Fig. 4B, DNA fragments containing the *CCNA1* (cyclin A) promoter (positions -50 to +50) were enriched more efficiently than DNA fragments containing an upstream region (positions -1000 to -900) of the same gene when IgG-Sepharose beads were used to bind the protein A component of the TAP tag. Our results demonstrate that TAP-tagged RNAPII is specifically recruited to the promoter of actively transcribed genes in vivo.

As mentioned above, BLAST searches against publicly available databases revealed weak similarities between the amino acid sequence of RPAP1 and a variety of proteins and domains, including Brd4. Because Brd4 was recently shown to specifically interact with histones H3 and H4 through acety-

FIG. 2. Purification of TAP-tagged human RNAPII complex. (A) Whole-cell extracts prepared from induced or noninduced EcR-293 cells programmed to express TAP-tagged Rpb11 (Rpb11-TAP) were purified by the TAP procedure, and the eluates were analyzed by SDS-PAGE. Gel slices containing the most abundant polypeptides were excised, digested with trypsin, and analyzed by peptide mass fingerprinting using MALDI-TOF analysis. The positions of core RNAPII subunits, including Rpb11 carrying the residual CBP domain (Rpb11-CBP), TFIIB, RAP74, RAP30, Fcp1, and RPAP1, a human gene product (DKFZP727M111 protein) of unknown function, are indicated. (B) TAP tagging of Rpb2, Rpb4, Rpb7, Rpb11, TFIIB, and RAP30 allowed for purification of the same 17-polypeptide RNAPII complex. The components of the complex were identified according to their molecular weights in SDS gels stained with silver (SS), their peptide mass fingerprints by MALDI-TOF analysis (MS), and their immunoreactivities in Western blots (WB). (C) SDS gel showing that both tagged (RAP30-CBP) and nontagged RAP30 are present in the RNAPII complex purified with TAP-tagged RAP30. Only nontagged RAP30 is found in the complex purified with TAP-tagged Rpb11. (D) The TAP-tagged RNAPII complex contains a hypophosphorylated CTD. The N-20 antibody, which specifically recognizes the N terminus of Rpb1, was used to detect both the hypophosphorylated (IIa) and hyperphosphorylated (IIo) forms of RNAPII in the whole-cell extract (WCE) and the Rpb11-TAP eluate. A truncated form of Rpb1 lacking the CTD (IIb) was also detected.

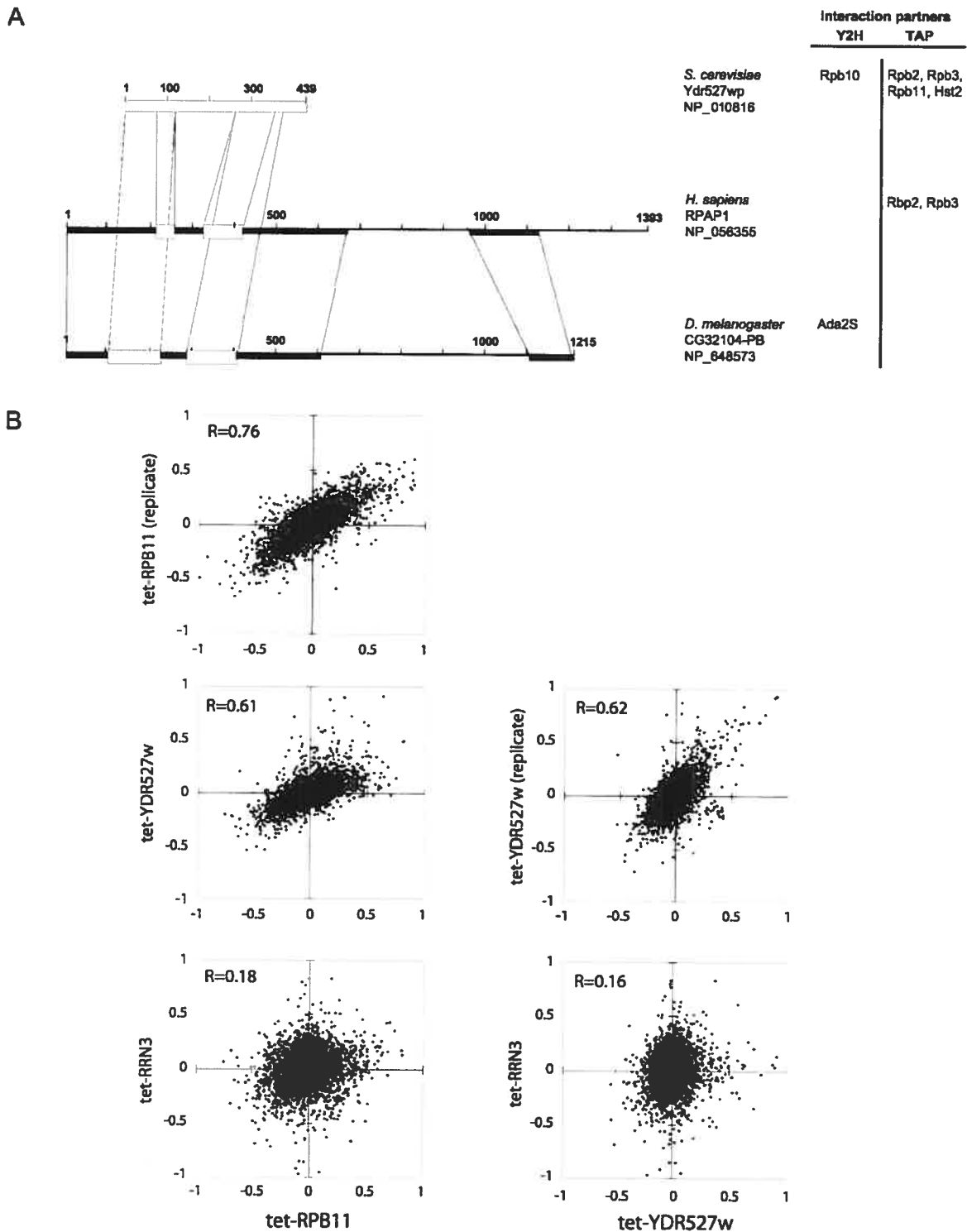


FIG. 3. (A) Regions of homology between *Homo sapiens* RPAP1, *D. melanogaster* CG32104-PB, and *S. cerevisiae* Ydr527wp/RBA50, as determined by BLAST analyses. Boxes represent regions with significant alignments. Black boxes, E values ranging from 8×10^{-37} to 1×10^{-10} ; open boxes, E values ranging from 4×10^{-5} to 0.1. The *M. musculus* and *R. norvegicus* orthologues, with 80% identity with the *H. sapiens* RPAP1 over the full length, were not presented. Reported protein-protein interactions of RPAP1 (this article) and Ydr527wp/RBA50 (33) with transcription factors are indicated on the right. (B) Correlations among mRNA abundance profiles for Tet-promoter mutants of *RPB11*, *YDR527w*, and *RRN3*, measured by using oligonucleotide microarrays. Each mutant was compared to an isogenic control without the *tetO7* promoter; $\log(\text{ratios})$ are plotted. Tet-*RPB11* and Tet-*YDR527w* were assayed twice, with the replicate cultures being grown, extracted, and assayed on different dates. The correlation of Tet-*RPB11* and Tet-*YDR527w* to a Tet-promoter mutant of *RRN3*, a gene involved in polymerase I transcription, is shown for contrast.

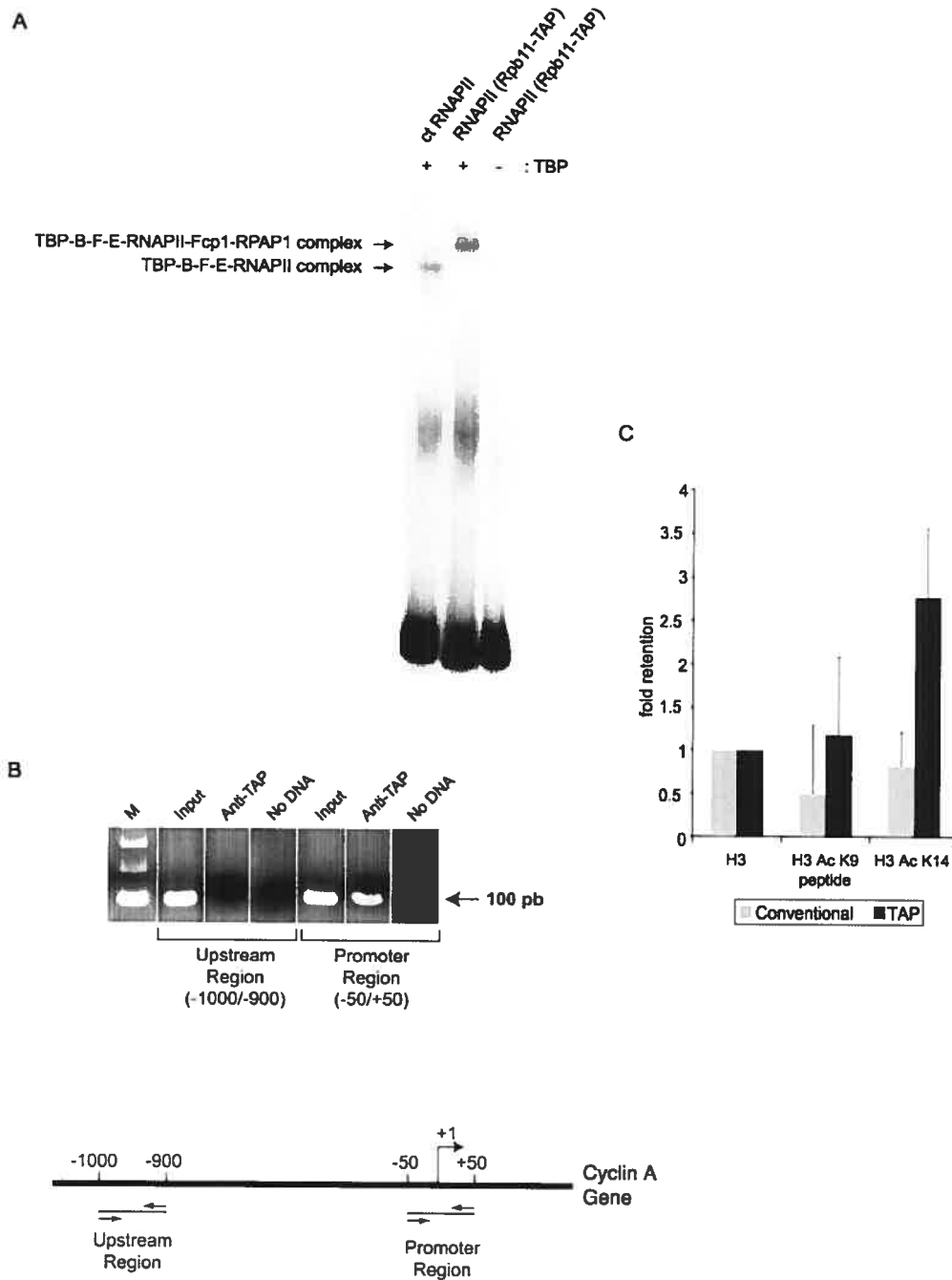


FIG. 4. TAP-tagged human RNAPII assembles on promoter DNA both in vitro and in vivo and binds to acetylated histones. (A) EMSA performed with 150 ng of purified calf thymus (ct) RNAPII or TAP-tagged human RNAPII in the presence of TBP, TFIIB, TFIIF, and TFIIE. A control reaction assembled in the absence of TBP was included. A radiolabeled DNA fragment containing the AdML promoter was used as a probe. (B) ChIP experiments showing that TAP-tagged RNAPII is specifically recruited to the cyclin A promoter in vivo. PCR amplification using sets of primers specific to chromosomal regions either encompassing (-50 to +50) or located upstream of (-1000 to -900) the transcription start site was used on DNA fragments enriched with IgG beads (anti-TAP). The input lanes correspond to DNA that was not subjected to immunoprecipitation. A PCR control that lacks DNA (no DNA) was included in each case. (C) Peptide binding experiment showing that TAP-tagged RNAPII is retained on acetylated histones. The results of three independent experiments were analyzed with ImageQuant software, corrected for the background, and shown relative to unmodified H3. Conventionally purified RNAPII was used as a negative control. The TAP-tagged RNAPII-bound histone tails acetylated at lysine 14 more efficiently than at lysine 9.

lated lysine residues (19), we tested the ability of the complex to bind to acetylated histones. As shown in Fig. 4C, the TAP-tagged RNAPII complex specifically bound to acetylated histone H3 *in vitro*, while conventionally purified RNAPII was not retained on acetylated histone H3. TAP-tagged RNAPII bound histone tails acetylated at lysine 14 more efficiently than those acetylated at lysine 9. This result suggests that RPAP1 may recruit RNAPII to chromatin through its interaction with acetylated histones.

The TAP-tagged RNAPII complex supports accurately initiated transcription *in vitro*. TAP-tagged RNAPII was tested for the ability to initiate transcription *in vitro* by use of a transcription system reconstituted with purified general initiation factors. Purified calf thymus RNAPII or purified TAP-tagged human RNAPII was incubated with a DNA template containing the AdML promoter in the presence of purified TBP, TFIIB, TFIIF, TFIIIE, and TFIIH. The formation of a 391-nt transcript indicated that the tagged RNAPII complex is competent for promoter-selective transcription *in vitro* (Fig. 5A).

When the general initiation factors were individually omitted from the transcription reaction (Fig. 5B), transcription was abolished except in the case of recombinant TFIIB or TFIIF. These data fully support our conclusion that TFIIB and TFIIF are components of the tagged RNAPII complex and establish that they are functionally active within this complex.

Purification and functional analysis of a human RNAPII variant with a site-directed mutation in fork loop 1. As part of an effort to understand the fine molecular details of the transcription reaction, we used the TAP procedure to generate and isolate RNAPII variants with site-directed mutations in structural elements predicted to have roles in the transcription reaction according to the crystal structure of the yeast enzyme. Because the Rpb2 element known as fork loop 1 (Fig. 6A) is highly conserved from yeast to humans, absent from bacterial RNAP, and proposed to be involved in maintaining the transcription bubble (17, 30) or in RNA-DNA strand separation (85), we first chose to modify this element. In yeast RNAPII, fork loop 1 amino acids Lys471 and Arg476 appear to contact the RNA phosphates at positions -5 , -6 , and -7 in the hybrid region. Deletion of the whole 13-amino-acid loop totally abolished the ability of Rpb2 to assemble with the other subunits of RNAPII (data not shown), probably because the deletion destabilized Rpb2, which then became misfolded. A smaller deletion of two amino acids located near the center of the loop ($\Delta 458-459$), which corresponded to amino acids Lys471 and Ala472 in the yeast enzyme, was successfully used to purify the 17-subunit RNAPII complex (Fig. 6B). This mutant was named Rpb2 fork1 $\Delta 458-459$. The silver-stained SDS gel in Fig. 6B shows that the relative abundance of some components of the tagged RNAPII complex varied between the wild type and the mutant, indicating that the mutation may have slightly affected complex assembly. By quantifying the relative intensity of each subunit, we established that the most important difference concerned Rpb1, which was at the very most fivefold less abundant in the mutant than in the wild-type complex.

In order to elucidate the function of fork loop 1 in transcription, we compared Rpb2 fork1 $\Delta 458-459$ with wild-type RNAPII in several different assays. The assays shown in Fig. 6C to E used fivefold larger amounts of the mutant complex than of

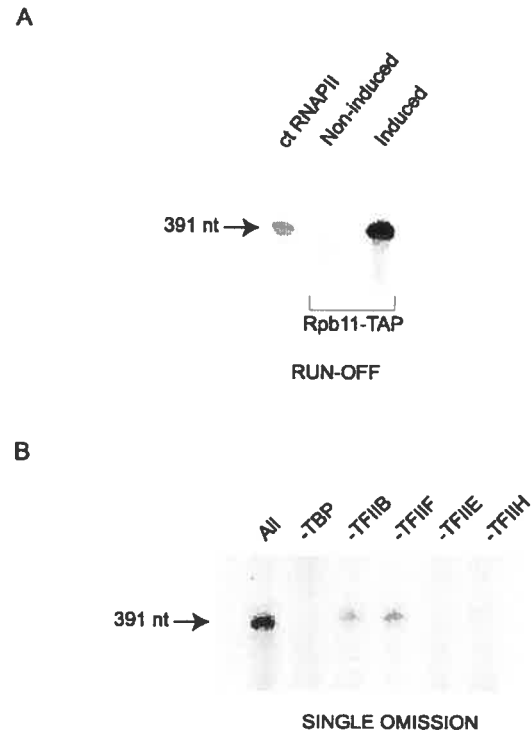


FIG. 5. TAP-tagged human RNAPII complex can initiate transcription *in vitro*. (A) *In vitro* transcription reactions were reconstituted by using 165 ng of purified calf thymus (ct) RNAPII or TAP-tagged human RNAPII in the presence of TBP, TFIIB, TFIIF, TFIIIE, and TFIIH. A control reaction performed with an eluate from noninduced cells was included. The linearized DNA template carries the AdML promoter and directs the synthesis of a 391-nt transcript. (B) Single omission assay in which each general transcription factor was omitted from the reconstituted system (all) described for panel A.

the wild type to ensure that the amount of the mutant complex was not underestimated (see above). As shown in Fig. 6C, Rpb2 fork1 $\Delta 458-459$ was defective in runoff transcription assays when we used the AdML promoter in the presence of all of the general initiation factors.

In order to delineate with more precision the molecular defect that impairs the transcriptional activity of Rpb2 fork1 $\Delta 458-459$ *in vitro*, we tested its activity in a promoter-independent assay. As described previously (29, 79), RNAPII can transcribe from a template carrying a 3' C extension (C-tailed template) in the absence of the general initiation factors. The template used for Fig. 6D carried a 15-C extension on the 3' strand. The transcribed strand was devoid of dCMP except at positions $+25$ to $+27$ ($+1$ corresponds to the single-stranded-double-stranded junction). In the absence of CTP, wild-type RNAPII produced a transcript of 35 nt. In the presence of CTP, the major transcript was 60 nt in length, indicating that the polymerase starts transcription 10 nt upstream of the double-stranded junction. In contrast to the RNAPII complex carrying wild-type Rpb2, Rpb2 fork1 $\Delta 458-459$ was inactive in the C-tailed template elongation assay.

Figure 6E shows that Rpb2 fork1 $\Delta 458-459$ was also defective in an abortive initiation assay carried out with a short double-stranded oligonucleotide carrying the AdML promoter

in the presence of the general initiation factors, but in the absence of GTP. A G nucleotide strategically located at +11 in the template blocked transcription at position +10 (Fig. 6E). When we used a fully double-stranded template (0/0), Rpb2 fork1 Δ 458-459 did not produce abortive transcripts. We rationalized that if Rpb2 fork1 Δ 458-459 was defective in promoter melting, a premelted template would counteract the defect in the mutant. The premelted template ($-9/+2$), which carried an unpaired DNA region from -9 to $+2$ that mimicked the fully melted promoter, supported the synthesis of abortive transcripts, although with a poor efficiency compared to that of the wild type. This finding indicated that a mutation in fork loop 1 does not simply affect promoter melting and its maintenance, but rather affects the interaction of RNAPII with the melted promoter region.

We finally assessed the ability of Rpb2 fork1 Δ 458-459 to assemble into a preinitiation complex on promoter-containing DNA. EMSA experiments showed that Rpb2 fork1 Δ 458-459 can enter a preinitiation complex with the general initiation factors as efficiently as wild-type RNAPII (Fig. 6F).

Human RNAPII variant with a site-directed mutation in switch 3. The structure named switch 3 is the only loop of Rpb2 located near the DNA-RNA hybrid that is ordered in the crystal structure of elongating yeast RNAPII (30). Switch 3 makes direct contact with the template DNA near positions -5 (arginine 1122) and -2 (arginine 1129) (Fig. 6A). Since R1129 is juxtaposed with one of the Rpb2 domains that mediate the interaction with Rpb1, we reasoned that disruption of this amino acid would likely affect the assembly of RNAPII and decided to turn our attention to R1122. We produced a mutant in which amino acids R1122, S1123, and R1124 were replaced with alanine residues. In human Rpb2, these amino acids correspond to R1078, S1079, and R1080. This mutant was named Rpb2 sw3-1078.

Figure 7A shows that TAP-tagged Rpb2 with the sw3-1078 mutation is as efficient as the wild type at forming the 17-subunit RNAPII complex. However, EMSA experiments (Fig. 7B) indicated that purified Rpb2 sw3-1078 is not efficient at entering a preinitiation complex on promoter DNA in the presence of the general initiation factors. Increasing the concentration of Rpb2 sw3-1078 by a factor of 6 compared to the wild type was necessary in order to achieve efficient formation of a preinitiation complex. As expected on the basis of the EMSA results, Rpb2 sw3-1078 did not support transcription *in vitro* (Fig. 7C). Increasing the concentration of Rpb2 sw3-1078 by a factor of 6 compared to the wild type led to the production of runoff transcripts, but the transcription efficiency remained slightly lower than that of the wild type, indicating that the mutation in switch 3 not only affects the ability of Rpb2 sw3-1078 to assemble into a preinitiation complex, but also, to a lesser extent, affects its capacity to support transcription.

Mutants Rpb2 fork1 Δ 458-459 and Rpb2 sw3-1078 are defective *in vivo*. To assess the defects of Rpb2 fork1 Δ 458-459 and Rpb2 sw3-1078 *in vivo*, we used ChIP experiments to evaluate their association with both the promoter and the transcription unit of two genes that are actively transcribed in human cells, the guanine nucleotide binding protein beta polypeptide 2-like 1 (GNB2L1) gene and the ferritin light polypeptide (FTL) gene. As shown in Fig. 8, Rpb2 fork1 Δ 458-459 is recruited to both promoters with an efficiency similar to

that of a wild-type TAP-tagged RNAPII complex, but it is not found in association with the transcription units. In contrast, Rpb2 sw3-1078 is not enriched in the promoter regions or the transcribed regions. The mutant Rpb2 fork1 Δ 453-465, which carries a complete deletion of the 13-amino-acid loop and does not assemble into a bona fide 17-subunit RNAPII complex, was used as a negative control that was enriched at neither the promoter nor the transcription unit of either the GNB2L1 or FTL gene. These *in vivo* results fully support our *in vitro* studies indicating that fork loop 1 is essential for transcription but is not involved in preinitiation complex assembly, whereas switch 3 is required for efficient preinitiation complex formation.

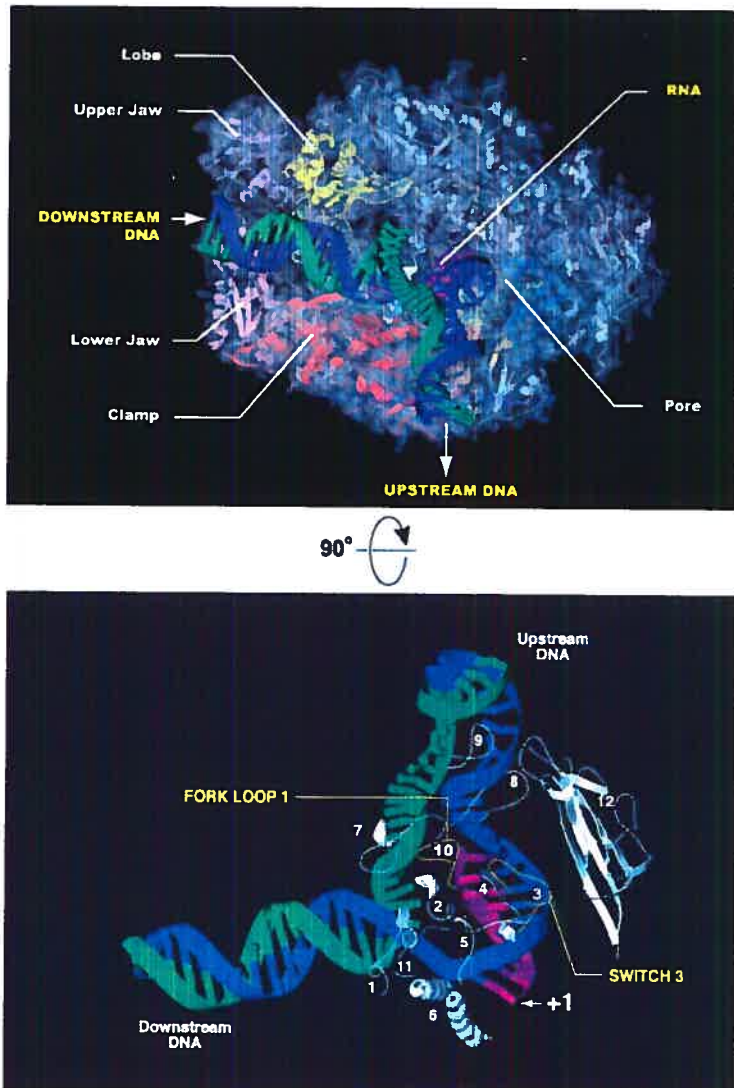
DISCUSSION

Simple, efficient purification of human RNAPII under native conditions. Classical approaches for the purification of mammalian RNAPII and its general transcription factors have typically used many different chromatography steps, including ion-exchange chromatography. Eluting these columns with buffers containing high salt concentrations most certainly affected the polypeptide composition of the purified complexes. For example, regulatory subunits, which are often less tightly associated with the enzyme they regulate, may have been lost during the purification procedure. Seraphin and colleagues (71) developed the TAP procedure to allow for the purification of protein complexes under native conditions. Two affinity chromatography steps are used. Elution from an IgG column is performed by a proteolytic treatment with the tobacco etch virus protease while elution from a calmodulin column uses EGTA as a chelating agent. For this paper, we adapted the TAP procedure for the purification of protein complexes from human cells, using an ecdysone-inducible expression system that allows for the expression of the TAP-tagged polypeptides at near-physiological levels. The expression levels were modulated by adjusting the concentration of the inducer, ponasterone A (Fig. 1B). By preventing overexpression of the tagged polypeptides, we limited the formation of nonspecific interactions.

We report here the use of the TAP procedure with human cells for the easy and efficient purification of a complex containing the 12 subunits of RNAPII, both subunits of TFIIF, TFIIB, Fcp1, and RPAP1, a novel polypeptide of unknown function. This 17-subunit complex was purified through the alternate TAP tagging of six of its subunits, demonstrating that all of the components are part of a unique complex. Because (i) the TAP-tagged polypeptides are expressed at near-physiological levels in stably transfected human cells, (ii) protein extraction is performed under gentle conditions, and (iii) the affinity purification steps are achieved under native conditions, our results indicate that this TAP-tagged RNAPII complex is the most abundant soluble, stable form of RNAPII in human cells. If a mediator complex (70) is associated with RNAPII in human cell extracts, either its concentration is much lower than that of RNAPII or its association with RNAPII is too weak to survive double affinity purification.

The presence of the TAP tag at the C termini of subunits of the RNAPII-containing complex does not alter its activity. The TAP-tagged RNAPII complex is fully active *in vitro*, as it can

A



B

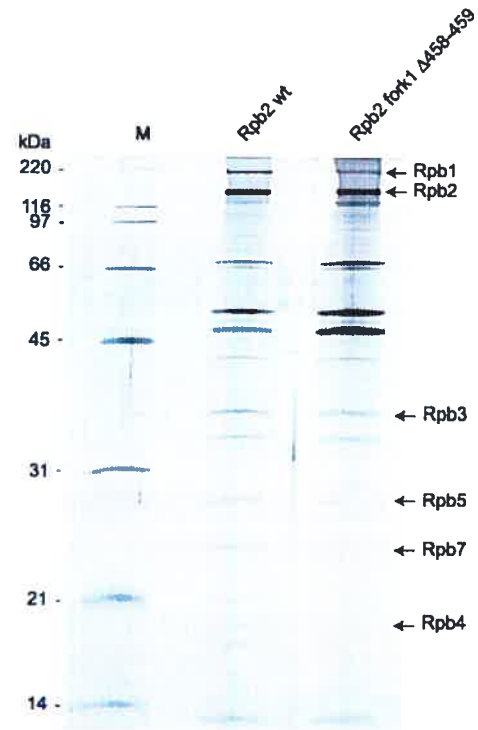
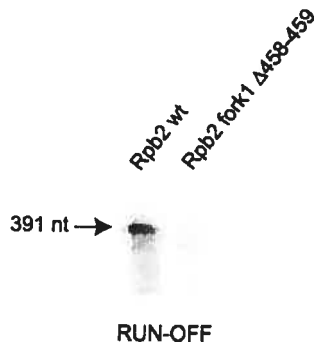


FIG. 6. Purification and functional analysis of TAP-tagged human RNAPII with a mutation in fork loop 1. (A) (Top) Model showing elongating yeast RNAPII (30). The template DNA strand (blue) from -3 to $+10$ and the 9-nt RNA (pink) were placed according to the crystal structure. The remainder of the template strand and the coding strand (green) were modeled in Cinema 4D, and their exact positions are speculative. (Bottom) Simplified view of yeast RNAPII catalytic center. Different domains of Rpb1 and Rpb2 located near the DNA-RNA hybrid are shown. Numbers: 1 to 5, switches; 6, bridge helix; 7, rudder; 8, lid; 9, zipper; 10, fork loop 1; 11, fork loop 2; 12, wall. The parts of the lid (amino acids 250 to 258) and the rudder (amino acids 315 to 320) that are absent in the PDB files corresponding to elongating RNAPII were reconstructed based on the crystal structure of free yeast RNAPII (18). The missing parts of fork loop 1 (amino acids 468 to 476) and fork loop 2 (amino acids 503 to 508) were modeled in Cinema 4D, and their exact positions are speculative. (B) Silver-stained SDS gel of wild type TAP-tagged RNAPII (Rpb2 wt) and a mutant of human RNAPII carrying a two-amino-acid deletion in the Rpb2 fork loop 1 (Rpb2 fork1 $\Delta 458-459$) domain. The positions of some RNAPII subunits and molecular size markers are indicated. (C) In vitro transcription reactions (runoff) contained either 24 ng of wild-type RNAPII (Rpb2 wt) or 120 ng of Rpb2 fork1 $\Delta 458-459$ in the presence of TBP, TFIIB, TFIIF, TFIIE, and TFIIH. (D) Elongation assays were performed with a C-tailed template carrying a 15-C extension in the absence of general transcription factors. The coding strand lacks CMP except in the $+25$ -to- $+27$ region. The positions of transcripts produced by RNAPII in the presence (+CTP; 60 nt) or absence (-CTP; 35 nt) of CTP are indicated. The reactions contained 88 ng of Rpb2 wt and 440 ng of Rpb2 fork1 $\Delta 458-459$. (E) Abortive initiation assays were performed with 100 ng of Rpb2 wt and 500 ng of Rpb2 fork1 $\Delta 458-459$ in the presence of TBP, TFIIB, TFIIF, and TFIIE on closed (0/0) or premelted ($-9/+2$) templates in the absence of GTP. The templates are schematized and the 3- to 10-nt abortive transcripts are indicated. (F) EMSAs were performed with the AdML promoter in the presence of TBP, TFIIB, TFIIF, and TFIIE. Two hundred nanograms of Rpb2 wt and 500 ng of Rpb2 fork 1 $\Delta 58-459$ were used in the reactions.

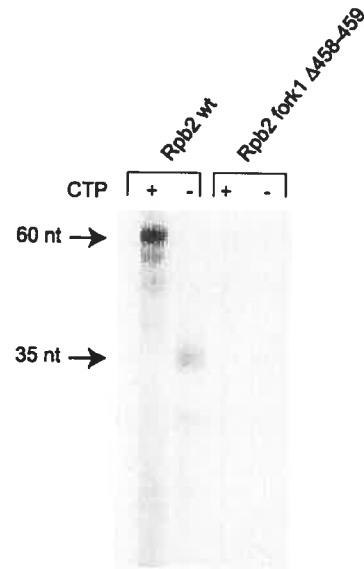
form a preinitiation complex with the general initiation factors on the AdML promoter, accurately initiate transcription, and elongate the transcripts (Fig. 4 and 5). In addition, both the TFIIB and TFIIF components of the complex are functionally

active in transcription reactions. Notably, TAP-tagged RNAPII can be recruited to transcriptionally active promoters in vivo, an indication that the purified human RNAPII complex is biologically active.

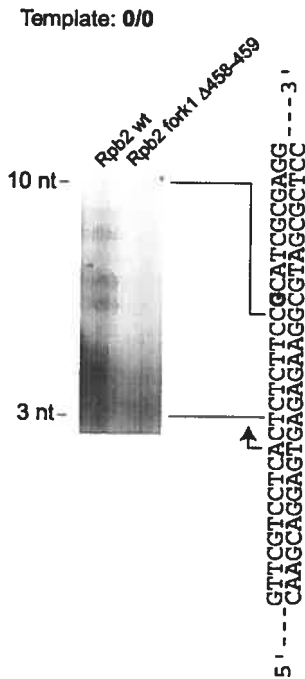
C



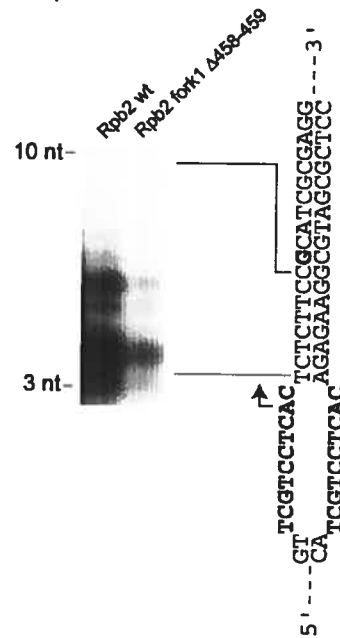
D



E



Template: -9/+2



F

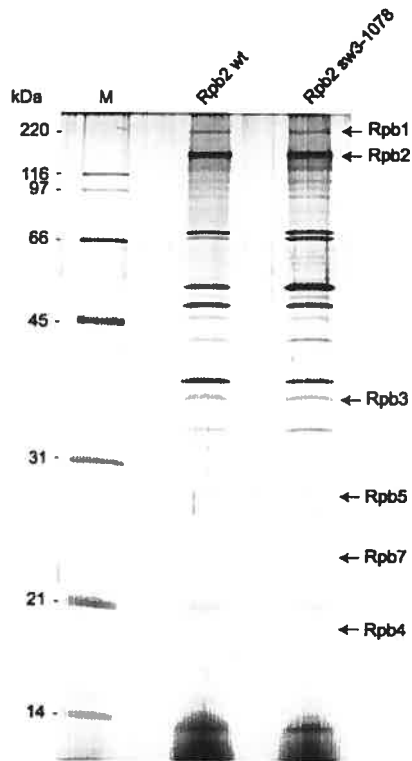


FIG. 6—Continued.

Five different regulatory subunits are stably associated with soluble RNAPII in human cells. The soluble form of RNAPII is expected to be the form of the enzyme that is recruited to promoters upon activation signals. In support of this hypothesis, the CTD of Rpb1 in the TAP-tagged RNAPII complex is in the hypophosphorylated state, the form of the enzyme that

is recruited to promoter DNA. A possible function of Fcp1 in this complex could be to maintain the CTD in this unphosphorylated state, as previously proposed (47). Our EMSAs suggested that Fcp1 can enter the preinitiation complex with the RNAPII complex. The entry of Fcp1 with RNAPII at the promoter is supported by the finding by Buratowski and col-

A



B



C

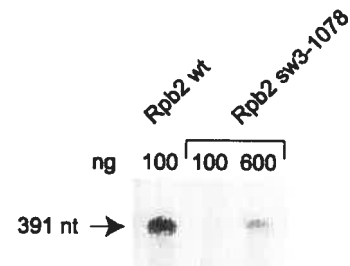


FIG. 7. Purification and functional analysis of TAP-tagged human RNAPII with a mutation in switch 3. (A) Silver-stained SDS gel of wild-type TAP-tagged RNAPII (Rpb2 wt) and a mutant of human RNAPII carrying a triple alanine substitution in the Rpb2 switch 3 (Rpb2 sw3-1078) domain. (B) EMSAs were performed with the AdML promoter with TFIIB, TFIIF, and TFIIE in either the presence or the absence of TBP. The amounts used for the wild type and the mutant are indicated. (C) In vitro transcription reactions (runoff) contained different amounts of Rpb2 wt and Rpb2 sw3-1078 in the presence of TBP, TFIIB, TFIIF, TFIIE, and TFIIF.

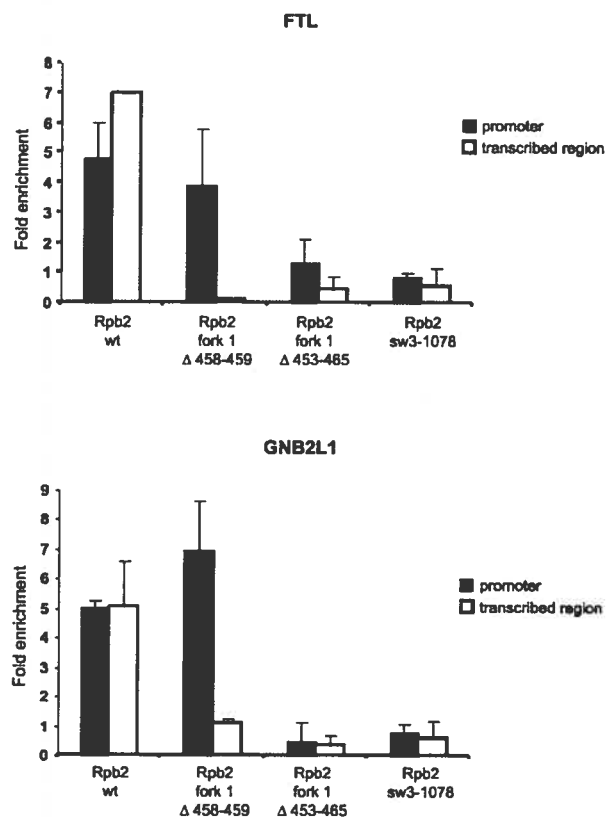


FIG. 8. Association of mutant forms of the RNAPII complex with chromatin in vivo. The fold enrichment over the control region of both the promoter and the transcribed region of the *FTL* and *GNB2L1* genes after the immunoprecipitation of chromatin fragments was quantified by Q-PCR. Enrichment was normalized to an intergenic region of chromosome 17 (see Materials and Methods). Data points represent the means of four experiments, including two independent ChIP assays.

leagues that the CTD phosphatase colocalizes with RNAPII at active promoters in yeast (10). Therefore, the role of Fcp1 in maintaining RNAPII in the hypophosphorylated state may be required until RNAPII becomes processive and clears the promoter.

Both TFIIB and TFIIF have previously been shown to regulate the recruitment of RNAPII to promoter DNA (15, 54, 80). TFIIB binds to TBP and forms a TBP-TFIIB-promoter complex (4, 59). TFIIB binds directly to a DNA motif named the TFIIB recognition element, a core element found in several RNAPII promoters (51). TFIIF binds to both RNAPII and TFIIB and is also involved in RNAPII recruitment to core promoters (15, 54, 80). In addition, TFIIF was shown to regulate the contacts of RNAPII with promoter DNA before transcription initiation (20, 25, 72). TFIIF plays a central role in promoter bending and wrapping against the mobile clamp of RNAPII, a process that is essential for transcription initiation (24). The function of TFIIF in promoter wrapping and transcriptional initiation requires a homomeric interaction region of RAP74, called HIR-1, that maintains TFIIF as a $\alpha_2\beta_2$ heterotetramer (72). Our results showing that two molecules of RAP30 are present within the tagged RNAPII complex sup-

port the notion that TFIIF works as a $\alpha_2\beta_2$ heterotetramer in vivo (13, 23).

Notably, both TFIIF and TFIIB can also bind to Fcp1 in a competitive manner and with opposite effects, with TFIIF stimulating and TFIIB inhibiting the phosphatase activity (9). Fcp1 may not efficiently dephosphorylate RNAPII at Ser2 of the CTD in the region downstream of the promoter (10) because TFIIF appears not to be present in this complex (50, 68). Clearly, TFIIB, TFIIF, and Fcp1 participate in an important network of regulatory interactions during RNAPII transcription.

The TAP-tagged RNAPII complex also contains at least one polypeptide that can specifically bind to acetylated histones in vitro. One putative candidate is RPAP1. Whether or not RPAP1 contains a bromodomain remains unclear because the definitive identification of a bromodomain requires structural data. An essential yeast polypeptide with regions of homology with RPAP1 was previously shown to interact with the Rpb10 subunit of RNAPII (28, 43) and also affected global gene expression in the same way as Rpb11 in our Tet-promoter mutant experiments (Fig. 3). These findings support the notion that the association of RPAP1 with RNAPII is physiologically relevant, although its exact role remains to be elucidated.

Structure-guided alignments indicated that RPAP1 could belong to the ARM repeat protein superfamily (12). ARM repeat proteins contain tandem copies of a degenerate 40-amino-acid sequence motif that forms a conserved three-dimensional structure. A single ARM repeat consists of three α helices. Such repeats are found in a variety of proteins, including the mammalian armadillo homologs β -catenin and importin α (40). β -catenin is mainly detected in adherens junctions where it links the membrane-anchored cadherin to actin filament. Cytoplasmic β -catenin is degraded by the ubiquitin-proteasome system. However, when Wnt signaling is activated, cytoplasmic β -catenin is stabilized. The increased level of β -catenin leads to its nuclear translocation, where it affects transcriptional activation. Although it was reported to be a cytoplasmic protein (41), Ydr527wp/RBA50 is found in complex with three RNAPII subunits and appears to affect transcription. Similar to β -catenin, Ydr527wp/RBA50 may have a dual cellular location and may be translocated to the nucleus.

Proteins that were previously shown to interact with RNAPII, such as the mediator and various elongation factors, were not found to be major components of our TAP-tagged RNAPII eluate. This is likely because these RNAPII-interacting proteins are associated with RNAPII in the insoluble fraction and, consequently, are mainly lost during the purification procedure. However, these complexes may represent minor forms of RNAPII complexes in our TAP-tagged eluates.

Structure-function analysis of the human RNAPII active center. The crystal structure of eukaryotic RNAPII was used to predict the function of many structural elements found in the active center. For example, some loops and helices that either directly contact or closely approach the DNA-RNA duplex in the active site of the elongating enzyme were suggested to regulate the formation and maintenance of both the transcription bubble and the RNA-DNA hybrid. In order to begin assessing the functions of these structural elements, we expressed TAP-tagged Rpb2 mutants carrying site-directed mutations and purified the mutant forms of RNAPII. Our data

indicate that the assembly and/or stability of the RNAPII complex is affected by mutations in the active site. For example, the full deletion of fork loop 1 led to defects in the assembly of the 17-subunit RNAPII complex. However, a shorter mutant with a deletion of only two amino acids from Rpb2 was found to assemble correctly in solution *in vivo* (Fig. 6) and allowed us to analyze the function of this element during the transcription reaction. The Rpb2 fork1 Δ 458-459 mutant was found to enter a preinitiation complex with the general initiation factors with an efficiency similar to that of the wild-type complex, but it was incapable of transcription initiation and elongation *in vitro*. When we used a premelted template in which the transcription bubble was fully open between positions -9 and $+2$ in our initiation assay, Rpb2 fork1 Δ 458-459 was able to initiate transcript formation, but not very efficiently. These findings indicate that there may be a defect at the level of the interaction between fork loop 1 and the melted DNA at very early stages of the transcription reaction. For example, the two-amino-acid-shorter fork loop 1 in the RNAPII mutant may be unable to contact one of the DNA strands in the transcription bubble, leading to incorrect positioning of the DNA template in the active site and impairing the ability of the enzyme to initiate transcription efficiently. Since this mutant supported transcription initiation much less efficiently even when a fully open template was used, our results do not support a direct role for fork loop 1 in promoter melting.

Given the recent crystal structure of elongating yeast RNAPII showing an interaction between Lys 471 (Lys 458 in human RNAPII) and the RNA around positions -5 to -7 (85), an alternative possibility to explain the defect of Rpb2 fork 1 Δ 458-459 is that during transcription initiation, the mobile fork loop 1 makes crucial contacts with the first arriving ribonucleotides, making it essential for phosphodiester bond formation. According to this speculative view, fork loop 1 would keep contact with the RNA as transcription progresses and be displaced until it reaches the position seen in the crystal structure, where it would participate in RNA-DNA strand separation.

The five switches located at the bottom of the cleft of RNAPII were first proposed to be involved in coupling the closure of the mobile clamp to the presence of a DNA-RNA hybrid (18, 30). This hypothesis was based on the fact that most of the switches are disordered in the crystal structure of the 10-subunit RNAPII complex in solution (e.g., lacking the Rpb4-Rpb7 dimer), whereas they are ordered in the elongating RNAPII structure. The ordering of the switches coincided with the movement of the clamp, which was in an open state in the structure of the free enzyme to permit the entry of DNA into the cleft and in a closed state in the structure of the RNAPII-DNA-RNA ternary complex to ensure the stability of the elongating complex. Since some switches, including switch 3, directly contact the DNA in the DNA-RNA hybrid region, they were proposed to be involved in sensing the presence of nascent RNA and coupling it to the closure of the clamp on the DNA. However, the recent crystal structures of the 12-subunit RNAPII complex (2, 7) revealed that the clamp is in a closed state in the presence of the Rpb4-Rpb7 dimer. These results led the authors to propose that the DNA is first loaded onto the clamp far from the active site in a way that is similar to that of the bacterial RNA polymerase holoenzyme-promoter DNA

complex (57) and that it reaches the cleft only after DNA melting has occurred. In this situation, switch 3 could only contact DNA after the opening of the promoter and entry of the DNA into the cleft.

Our results indicate that switch 3 is important for the assembly of RNAPII into a preinitiation complex on promoter DNA. Mutating amino acids 1078 to 1080 clearly reduced the ability of human RNAPII to form a preinitiation complex in the presence of the general transcription factors both *in vitro* and *in vivo*. *In vitro*, sixfold larger amounts of Rpb2 sw3-1078 were necessary to form a complex in gel shift experiments. Our results suggest that switch 3 is necessary for the accurate interaction of the clamp with the promoter DNA before opening of the transcription bubble and that this is the primary defect in our switch 3 mutant. Because the structure of the complete polymerase shows that a duplex DNA cannot bind switch 3 before melting, it is likely that the effect of the switch 3 mutation is either allosteric or mediated by interactions with a general initiation factor such as TFIIB or TFIIF. The importance of the interaction of switch 3 with the template DNA at positions -2 and -5 after promoter melting might be reflected by the fact that the transcription reactions carried out in the presence of sixfold larger amounts of the sw3-1078 mutant did not yield as many transcripts as did the wild type. This may also suggest that the preinitiation complexes formed with the sw3-1078 mutant are not fully functional, presumably because the interactions of switch 3 with the DNA template in the DNA-RNA hybrid region are impaired.

ACKNOWLEDGMENTS

We are grateful to the members of our laboratory for helpful discussions. We thank Diane Bourque for artwork and Julie Edwards for critical reading of the manuscript. We also thank Takahiro Nagase from the Kazusa DNA Research Institute for kindly providing the cDNA encoding RPAP1.

This work was supported by grants from the Canadian Institutes for Health Research (to B.C.), Genome Canada (to B.C., J.G., and T.R.H.), Genome Québec (to B.C.), the Ontario Genomics Institute (to J.G. and T.R.H.), and the Natural Sciences and Engineering Research Council of Canada (to T.R.H.). C.J. holds a studentship from the Canadian Institutes for Health Research and M.F.L. is supported by the Natural Sciences and Engineering Research Council of Canada and the Fonds Québécois de la Recherche sur la Nature et les Technologies. A.P.D. is supported by a C. H. Best Postdoctoral Fellowship. B.C. is a senior scholar from the Fonds de la Recherche en Santé du Québec.

REFERENCES

1. Acker, J., M. Wintzerith, M. Vigneron, and C. Kedinger. 1992. Primary structure of the second largest subunit of human RNA polymerase II (or B). *J. Mol. Biol.* **226**:1295-1299.
2. Armache, K. J., H. Kettenberger, and P. Cramer. 2003. Architecture of initiation-competent 12-subunit RNA polymerase II. *Proc. Natl. Acad. Sci. USA* **100**:6964-6968.
3. Awrey, D. E., R. G. Weilbaecher, S. A. Hemming, S. M. Orlicky, C. M. Kane, and A. M. Edwards. 1997. Transcription elongation through DNA arrest sites. A multistep process involving both RNA polymerase II subunit RPB9 and TFIIS. *J. Biol. Chem.* **272**:14747-14754.
4. 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.
5. Bryant, G. O., and M. Ptashne. 2003. Independent recruitment *in vivo* by Gal4 of two complexes required for transcription. *Mol. Cell* **11**:1301-1309.
6. Burton, Z. F., L. G. Ortolan, and J. Greenblatt. 1986. Proteins that bind to RNA polymerase II are required for accurate initiation of transcription at the adenovirus 2 major late promoter. *EMBO J.* **5**:2923-2930.
7. Bushnell, D. A., and R. D. Kornberg. 2003. Complete, 12-subunit RNA

- polymerase II at 4.1-Å resolution: implications for the initiation of transcription. *Proc. Natl. Acad. Sci. USA* 100:6969–6973.
8. Campbell, E. A., O. Muzzini, M. Chlenov, J. L. Sun, C. A. Olson, O. Weinman, M. L. Trester-Zedlitz, and S. A. Darst. 2002. Structure of the bacterial RNA polymerase promoter specificity sigma subunit. *Mol. Cell* 9:527–539.
 9. 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.
 10. 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.
 11. 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.
 12. Coates, J. C. 2003. Armadillo repeat proteins: beyond the animal kingdom. *Trends Cell Biol.* 13:463–471.
 13. 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.
 14. 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.
 15. 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. USA* 88:6205–6209.
 16. Coulombe, B., and Z. F. Burton. 1999. DNA bending and wrapping around RNA polymerase: a “revolutionary” model describing transcriptional mechanisms. *Microbiol. Mol. Biol. Rev.* 63:457–478.
 17. 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.
 18. Cramer, P., D. A. Bushnell, and R. D. Kornberg. 2001. Structural basis of transcription: RNA polymerase II at 2.8 Å resolution. *Science* 292:1863–1876.
 19. Dey, A., F. Chitsaz, A. Abbasi, T. Misteli, and K. Ozato. 2003. The double bromodomain protein Brd4 binds to acetylated chromatin during interphase and mitosis. *Proc. Natl. Acad. Sci. USA* 100:8758–8763.
 20. 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 IIH revealed by protein-DNA photo-cross-linking. *Mol. Cell. Biol.* 20:8168–8177.
 21. Feaver, W. J., O. Gileadi, Y. Li, and R. D. Kornberg. 1991. CTD kinase associated with yeast RNA polymerase II initiation factor b. *Cell* 67:1223–1230.
 22. Finkelstein, A., C. F. Kostrub, J. Li, D. P. Chavez, B. Q. Wang, S. M. Fang, J. Greenblatt, and Z. F. Burton. 1992. A cDNA encoding RAP74, a general initiation factor for transcription by RNA polymerase II. *Nature* 355:464–467.
 23. 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.
 24. 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.
 25. 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. USA* 94:7150–7155.
 26. 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.
 27. 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.
 28. Giaever, G., A. M. Chu, L. Ni, C. Connelly, L. Riles, S. Veronneau, S. Dow, A. Lucau-Danila, K. Anderson, B. Andre, A. P. Arkin, A. Astromoff, M. El Bakkoury, R. Bangham, R. Benito, S. Brachat, S. Campanaro, M. Curtiss, K. Davis, A. Deutschbauer, K. D. Entian, P. Flaherty, F. Foury, D. J. Garfinkel, M. Gerstein, D. Gotte, U. Guldener, J. H. Hegemann, S. Hempel, Z. Herman, D. F. Jaramillo, D. E. Kelly, S. L. Kelly, P. Kottler, D. LaBonte, D. C. Lamb, N. Lan, H. Liang, H. Liao, L. Liu, C. Luo, M. Lussier, R. Mao, P. Menard, S. L. Ooi, J. L. Revuelta, C. J. Roberts, M. Rose, P. Ross-Macdonald, B. Scherens, G. Schimmack, B. Shafer, D. D. Shoemaker, S. Sookhai-Mahadeo, R. K. Storms, J. N. Strathern, G. Valle, M. Voet, G. Volckaert, C. Y. Wang, T. R. Ward, J. Wilhelm, E. A. Wenzler, Y. Yang, G. Yen, E. Youngman, K. Yu, H. Bussey, J. D. Boeke, M. Snyder, P. Philippsen, R. W. Davis, and M. Johnston. 2002. Functional profiling of the *Saccharomyces cerevisiae* genome. *Nature* 418:387–391.
 29. Gnatt, A., J. Fu, and R. D. Kornberg. 1997. Formation and crystallization of yeast RNA polymerase II elongation complexes. *J. Biol. Chem.* 272:30799–30805.
 30. 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.
 31. Ha, I., W. S. Lane, and D. Reinberg. 1991. Cloning of a human gene encoding the general transcription initiation factor IIB. *Nature* 352:689–695.
 32. Hampsey, M. 1998. Molecular genetics of the RNA polymerase II general transcriptional machinery. *Microbiol. Mol. Biol. Rev.* 62:465–503.
 33. Hazbun, T. R., L. Malmstrom, S. Anderson, B. J. Graczyk, B. Fox, M. Riffle, B. A. Sundin, J. D. Aranda, W. H. McDonald, C. H. Chiu, B. E. Snysman, P. Bradley, E. G. Muller, S. Fields, D. Baker, J. R. Yates III, and T. N. Davis. 2003. Assigning function to yeast proteins by integration of technologies. *Mol. Cell* 12:1353–1365.
 34. Hegde, P., R. Qi, K. Abernathy, C. Gay, S. Dharap, R. Gaspard, J. E. Hughes, E. Snesrud, N. Lee, and J. Quackenbush. 2000. A concise guide to cDNA microarray analysis. *BioTechniques* 29:548–556.
 35. Ho, C. K., and S. Shuman. 1999. Distinct roles for CTD Ser-2 and Ser-5 phosphorylation in the recruitment and allosteric activation of mammalian mRNA capping enzyme. *Mol. Cell* 3:405–411.
 36. Hodo, H. G., III, and S. P. Blattli. 1977. Purification using polyethylenimine precipitation and low molecular weight subunit analyses of calf thymus and wheat germ DNA-dependent RNA polymerase II. *Biochemistry* 16:2334–2343.
 37. 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.
 38. Holstege, F. C., D. Tantin, M. Carey, P. C. van der Vliet, and H. T. Timmers. 1995. The requirement for the basal transcription factor IIE is determined by the helical stability of promoter DNA. *EMBO J.* 14:810–819.
 39. Holstege, F. C., P. C. van der Vliet, and H. T. Timmers. 1996. Opening of an RNA polymerase II promoter occurs in two distinct steps and requires the basal transcription factors IIE and IIH. *EMBO J.* 15:1666–1677.
 40. Huber, A. H., W. J. Nelson, and W. I. Weis. 1997. Three-dimensional structure of the armadillo repeat region of beta-catenin. *Cell* 90:871–882.
 41. Huh, W. K., J. V. Falvo, L. C. Gerke, A. S. Carroll, R. W. Howson, J. S. Weissman, and E. K. O’Shea. 2003. Global analysis of protein localization in budding yeast. *Nature* 425:686–691.
 42. Ingles, C. J., M. Shales, W. D. Cress, S. J. Triezenberg, and J. Greenblatt. 1991. Reduced binding of TFIID to transcriptionally compromised mutants of VP16. *Nature* 351:588–590.
 43. Ito, T., T. Chiba, R. Ozawa, M. Yoshida, M. Hattori, and Y. Sakaki. 2001. A comprehensive two-hybrid analysis to explore the yeast protein interactome. *Proc. Natl. Acad. Sci. USA* 98:4569–4574.
 44. Jiang, Y., S. J. Triezenberg, and J. D. Gralla. 1994. Defective transcriptional activation by diverse VP16 mutants associated with a common inability to form open promoter complexes. *J. Biol. Chem.* 269:5505–5508.
 45. 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.
 46. Kim, Y., J. H. Geiger, S. Hahn, and P. B. Sigler. 1993. Crystal structure of a yeast TBP/TATA-box complex. *Nature* 365:512–520.
 47. 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.
 48. 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.
 49. Kobor, M. S., L. D. Simon, J. Omichinski, G. Zhong, J. Archambault, and J. Greenblatt. 2000. A motif shared by TFIIF and TFIIB mediates their interaction with the RNA polymerase II carboxy-terminal domain phosphatase Fcp1p in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* 20:7438–7449.
 50. Krogan, N. J., M. Kim, S. H. Ahn, G. Zhong, M. S. Kobor, G. Cagney, A. Emili, A. Shilatifard, S. Buratowski, and J. F. Greenblatt. 2002. RNA polymerase II elongation factors of *Saccharomyces cerevisiae*: a targeted proteomics approach. *Mol. Cell. Biol.* 22:6979–6992.
 51. Lagrange, T., A. N. Kapanidis, H. Tang, D. Reinberg, and R. H. Ebricht. 1998. New core promoter element in RNA polymerase II-dependent transcription: sequence-specific DNA binding by transcription factor IIB. *Genes Dev.* 12:34–44.
 52. 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 TFIIF in transcription initiation by RNA polymerase II. *J. Biol. Chem.* 276:38652–38657.
 53. Lu, H., L. Zawel, L. Fisher, J. M. Egly, and D. Reinberg. 1992. Human general transcription factor IIH phosphorylates the C-terminal domain of RNA polymerase II. *Nature* 358:641–645.
 54. Maldonado, E., I. Ha, P. Cortes, L. Weis, and D. Reinberg. 1990. Factors involved in specific transcription by mammalian RNA polymerase II: role of

- transcription factors IIA, IID, and IIB during formation of a transcription-competent complex. *Mol. Cell. Biol.* **10**:6335–6347.
55. Marton, M. J., J. L. DeRisi, H. A. Bennett, V. R. Iyer, M. R. Meyer, C. J. Roberts, R. Stoughton, J. Burchard, D. Slade, H. Dai, D. E. Bassett, Jr., L. H. Hartwell, P. O. Brown, and S. H. Friend. 1998. Drug target validation and identification of secondary drug target effects using DNA microarrays. *Nat. Med.* **4**:1293–1301.
 56. Mizuguchi, K., C. M. Deane, T. L. Blundell, and J. P. Overington. 1998. HOMSTRAD: a database of protein structure alignments for homologous families. *Protein Sci.* **7**:2469–2471.
 57. 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.
 58. Murakami, K. S., S. Masuda, and S. A. Darst. 2002. Structural basis of transcription initiation: RNA polymerase holoenzyme at 4 Å resolution. *Science* **296**:1280–1284.
 59. 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.
 60. Nissen, R. M., and K. R. Yamamoto. 2000. The glucocorticoid receptor inhibits NFkappaB by interfering with serine-2 phosphorylation of the RNA polymerase II carboxy-terminal domain. *Genes Dev.* **14**:2314–2329.
 61. No, D., T. P. Yao, and R. M. Evans. 1996. Ecdysone-inducible gene expression in mammalian cells and transgenic mice. *Proc. Natl. Acad. Sci. USA* **93**:3346–3351.
 62. Ohkuma, Y., S. Hashimoto, C. K. Wang, M. Horikoshi, and R. G. Roeder. 1995. Analysis of the role of TFIIE in basal transcription and TFIIF-mediated carboxy-terminal domain phosphorylation through structure-function studies of TFIIE- α . *Mol. Cell. Biol.* **15**:4856–4866.
 63. Ohkuma, Y., and R. G. Roeder. 1994. Regulation of TFIIF ATPase and kinase activities by TFIIE during active initiation complex formation. *Nature* **368**:160–163.
 64. 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.
 65. Orphanides, G., T. Lagrange, and D. Reinberg. 1996. The general transcription factors of RNA polymerase II. *Genes Dev.* **10**:2657–2683.
 66. Pan, G., and J. Greenblatt. 1994. Initiation of transcription by RNA polymerase II is limited by melting of the promoter DNA in the region immediately upstream of the initiation site. *J. Biol. Chem.* **269**:30101–30104.
 67. 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 TFIIE. *Nature* **354**:369–373.
 68. 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.
 69. Puig, O., F. Caspari, G. Rigaut, B. Rutz, E. Bouveret, E. Bragado-Nilsson, M. Wilm, and B. Seraphin. 2001. The tandem affinity purification (TAP) method: a general procedure of protein complex purification. *Methods* **24**:218–229.
 70. Rachez, C., and L. P. Freedman. 2001. Mediator complexes and transcription. *Curr. Opin. Cell Biol.* **13**:274–280.
 71. 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.
 72. 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.
 73. Rodriguez, C. R., E. J. Cho, M. C. Keogh, C. L. Moore, A. L. Greenleaf, and S. Buratowski. 2000. Kin28, the TFIIF-associated carboxy-terminal domain kinase, facilitates the recruitment of mRNA processing machinery to RNA polymerase II. *Mol. Cell. Biol.* **20**:104–112.
 74. Schaffer, A. A., Y. I. Wolf, C. P. Ponting, E. V. Koonin, L. Aravind, and S. F. Altschul. 1999. IMPALA: matching a protein sequence against a collection of PSI-BLAST-constructed position-specific score matrices. *Bioinformatics* **15**:1000–1011.
 75. Serizawa, H., R. C. Conaway, and J. W. Conaway. 1992. A carboxyl-terminal-domain kinase associated with RNA polymerase II transcription factor delta from rat liver. *Proc. Natl. Acad. Sci. USA* **89**:7476–7480.
 76. Serizawa, H., T. P. Makela, J. W. Conaway, R. C. Conaway, R. A. Weinberg, and R. A. Young. 1995. Association of Cdk-activating kinase subunits with transcription factor TFIIF. *Nature* **374**:280–282.
 77. Shiekhattar, R., F. Mermelstein, R. P. Fisher, R. Drapkin, B. Dynlacht, H. C. Wessling, D. O. Morgan, and D. Reinberg. 1995. Cdk-activating kinase complex is a component of human transcription factor TFIIF. *Nature* **374**:283–287.
 78. Skaar, D. A., and A. L. Greenleaf. 2002. The RNA polymerase II CTD kinase CTDK-I affects pre-mRNA 3' cleavage/polyadenylation through the processing component Pti1p. *Mol. Cell* **10**:1429–1439.
 79. Sluder, A. E., D. H. Price, and A. L. Greenleaf. 1988. Elongation by *Drosophila* RNA polymerase II. Transcription of 3'-extended DNA templates. *J. Biol. Chem.* **263**:9917–9925.
 80. Sopta, M., R. W. Carthew, and J. Greenblatt. 1985. Isolation of three proteins that bind to mammalian RNA polymerase II. *J. Biol. Chem.* **260**:10353–10360.
 81. 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.
 82. Takahashi, Y., J. B. Rayman, and B. D. Dynlacht. 2000. Analysis of promoter binding by the E2F and pRB families in vivo: distinct E2F proteins mediate activation and repression. *Genes Dev.* **14**:804–816.
 83. Tirole, F., D. Busso, F. Coin, and J. M. Egly. 1999. Reconstitution of the transcription factor TFIIF: assignment of functions for the three enzymatic subunits, XPB, XPD, and cdk7. *Mol. Cell* **3**:87–95.
 84. 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.
 85. 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.
 86. Wu, L. F., T. R. Hughes, A. P. Davierwala, M. D. Robinson, R. Stoughton, and S. J. Altschul. 2002. Large-scale prediction of *Saccharomyces cerevisiae* gene function using overlapping transcriptional clusters. *Nat. Genet.* **31**:255–265.
 87. Yang, Y. H., and T. Speed. 2002. Design issues for cDNA microarray experiments. *Nat. Rev. Genet.* **3**:579–588.
 88. Young, R. A. 1991. RNA polymerase II. *Annu. Rev. Biochem.* **60**:689–715.
 89. Zhang, G., E. A. Campbell, L. Minakhin, C. Richter, K. Severinov, and S. A. Darst. 1999. Crystal structure of *Thermus aquaticus* core RNA polymerase at 3.3 Å resolution. *Cell* **98**:811–824.

CHAPITRE 5

UNE ANALYSE MUTATIONNELLE DU CENTRE CATALYTIQUE DE L'ARN POLYMÉRASE II

5.1 Préambule

Dans l'article présenté au chapitre 5, j'ai effectué le clonage, l'établissement des lignées stables et la purification pour tous les mutants de l'ARN pol II présentés, ainsi que le type sauvage Rpb1-TAP, en collaboration avec Dania Baali. J'ai effectué la purification de l'ARN pol II de type sauvage Rpb2-TAP. J'ai fait tous les essais présentés dans l'article (figure 1, 2A, 3, 4, 5, 6). J'ai participé intellectuellement à la conception du modèle, qui a été réalisé par Vincent Trinh (figure 7). Ce dernier a aussi fait le modèle de la figure 2B et a participé au clonage du mutant de levure (figure 4). J'ai rédigé l'article avec Dr Benoit Coulombe.

5.2 Article

Langelier, M.-F., Baali, D., Trinh, V., Greenblatt, J., Archambault, J. and Coulombe, B. (2004) A Mutational Analysis of the Human RNA Polymerase II Catalytic Center, article en préparation.

A Mutational Analysis of the Human RNA Polymerase II Catalytic Center

Marie-France Langelier¹, Dania Baali¹, Vincent Trinh¹, Jack Greenblatt², Jacques Archambault³ and Benoit Coulombe^{1*}

¹: Laboratory of Gene Transcription and

³: Laboratory of Molecular Virology

Institut de recherches cliniques de Montréal


110 avenue des Pins Ouest

Montréal (Québec) Canada

H2W 1R7

Tel. : (514) 987-5662

Fax : (514) 987-5663



²: Banting and Best Department of Medical Research

University of Toronto

Toronto, Ontario, M5G 1L6, Canada

* Corresponding Author

Running Title

Mutational analysis of the RNA pol II catalytic center

SUMMARY

We have affinity purified and characterized human RNA polymerase (RNAP) II mutants having deletions or substitutions in specific structural elements of the enzyme's catalytic center. Mutants of fork loop 1, switch 3, the zipper and the rudder, four loops predicted to play essential roles during transcript elongation, are impaired in the formation of an active preinitiation complex on promoter DNA, indicative of a defect early in the transcription reaction prior to transcription initiation. Analysis of an RNAP II mutant with the E791A substitution in Rpb2 revealed a role for this residue in binding metal B (Mg^{2+} -B) during both NTP polymerization and cleavage of backtracked transcripts. In light of recent crystallographic data reported on coordination of the NTP- Mg^{2+} by yeast Rpb2 D837, equivalent to human D792, both in the "addition" (A) and the "entry" (E) sites, our results define an additional NTP- Mg^{2+} interaction site, the "pre-loading" (P) site, closer to the enzyme's pore.

INTRODUCTION

RNA polymerase (RNAP) II, the enzyme that synthesizes mRNA in mammals, is composed of 12 subunits, termed Rpb1 to Rpb12 (Coulombe et Burton, 1999; Young, 1991). Crystal structures of yeast RNAP II (Cramer *et al.*, 2001; Gnatt *et al.*, 2001; Westover *et al.*, 2004a; Bushnell et Kornberg, 2003; Armache *et al.*, 2003) have confirmed that the 2 largest subunits, Rpb1 and Rpb2, form the enzyme's catalytic center. These structures also revealed the presence of a cleft at the surface of the enzyme which buries this active center. The cleft is formed, on one side, by a mobile "clamp" which is connected to the protein by a series of 5 "switch" regions. In the elongation complex, the DNA enters the cleft through a pair of "jaws" and the DNA helix is opened before the active center. The template strand binds to switches 1 and 2 and the "bridge" helix. Beyond the bridge, the template strand and the nascent RNA form a 9-bp duplex. Beneath the active site and in line with the bridge is found a "pore" or secondary channel; the space formed by the cleft at the surface of the protein being the main channel. The secondary channel widens towards the bottom of RNAP II, forming an inverted "funnel". In the past, both channels have been proposed as privileged routes for NTP entry, and pyrophosphate exit (Cramer *et al.*, 2001; Gnatt *et al.*, 2001; Nedialkov *et al.*, 2003).

A number of loops in Rpb1 and Rpb2 are located near the DNA-RNA hybrid region. These loops were named according to their position and to their presumed role in transcription. The "rudder" was proposed to be involved in DNA-RNA strand separation (Westover *et al.*, 2004a), the "zipper" in separating the template and non-template DNA strands (Gnatt *et al.*, 2001) and the "fork" loop 1 in maintenance of the upstream part of the bubble (Gnatt *et al.*, 2001) and/or in DNA-RNA strand separation (Westover *et al.*, 2004a). However, the lack of biochemical data greatly limits our understanding of the roles of these loops in the transcription reaction.

DNA and RNAPs catalyze the nucleophilic attack of the 3' hydroxyl of a first nucleotide on the α -phosphate of a second incoming NTP, leading to phosphodiester bond formation and the release of a pyrophosphate molecule. Based on crystal structures

of T7 DNA polymerase (Doublet et al., 1998) and *B. stearothermophilus* DNA polymerase (Kiefer et al., 1998), a two metal ion mechanism has been proposed (Steitz, 1998; Landick, 2004). Through this mechanism, a first Mg^{2+} ion (metal A) lowers the affinity of the 3' oxygen for the hydrogen thereby facilitating the O^- attack on the 5' α -phosphate. The second Mg^{2+} ion (metal B) facilitates the release of the pyrophosphate. High resolution crystal structures of yeast RNAP II revealed the position of the two Mg^{2+} ions in the catalytic site of the enzyme (Cramer et al., 2001; Gnatt et al., 2001; Westover et al., 2004b). In the yeast RNAP II structure, metal A is coordinated by three aspartate residues of Rpb1, D481, D483 and D485, contained within the strictly conserved motif NADFDGD (Cramer et al., 2001; Gnatt et al., 2001; Westover et al., 2004b). Metal B was recently localized near residue D481 and D483 of Rpb1 as well as the highly conserved residue D837 of Rpb2 (Westover et al., 2004b). Previous studies have suggested, however, that the strictly conserved residue E836 is also implicated in coordination of metal B in the active site (Cramer et al., 2001; Gnatt et al., 2001). A number of studies have also proposed that metal B would necessitate the presence of the incoming NTP for its complete coordination required for the polymerization process (Cramer et al., 2001; Gnatt et al., 2001; Westover et al., 2004b; Nedialkov et al., 2003). However, biochemical data in support of a role for E836, D837 and NTP as metal B coordinating residues is currently lacking.

In addition to their polymerization activity, DNA-dependant RNAPs can also catalyze the 3' endonucleolytic cleavage of transcripts to stimulate elongation (Conaway et al., 2000; Shilatifard et al., 2003). First discovered in *E. coli* (Krummel et Chamberlin, 1989; Surratt *et al.*, 1991), this cleavage activity takes place when RNAP has backtracked at pause and arrest sites (Wind et Reines, 2000; Fish et Kane, 2002). Weakly associated DNA-RNA hybrids can induce backtracking of the enzyme to a more stable register (Nudler et al., 1997). To resume elongation from this more stable position, RNAP needs to cleave the 3' end of the transcript that now extrudes from the catalytic site through the pore. This intrinsic nuclease activity is stimulated at basic pH, strongly suggesting a hydrolytic mechanism rather than pyrophosphorolysis (Weilbaecher et al., 2003). The 3' cleavage activity is also enhanced by specific factors, GreA and GreB in bacteria

(Borukhov et al., 1992; Borukhov et al., 1993) and TFIIS in eukaryotes (Wind et Reines, 2000; Fish et Kane, 2002). TFIIS was first identified as an elongation factor that helps RNAP II to elongate through arrest sites *in vitro* (Reinberg et Roeder, 1987; Reines *et al.*, 1989). TFIIS greatly stimulates the intrinsic nuclease activity of RNAP II (Izban et Luse, 1992b; Reines, 1992; Wang et Hawley, 1993). The crystal structure of TFIIS in a complex with RNAP II (Kettenberger et al., 2003) provided insight into the mechanism by which this factor enhances transcript cleavage. Specifically, it revealed that two conserved acidic residues in TFIIS essential for its activity (Jeon et al., 1994), are located in the vicinity of the polymerase Mg^{2+} -B and could participate in its coordination. Therefore it was suggested that a two metal ion mechanism is involved in the 3' endonucleolytic activity of RNAP II. A two metal ion mechanism is consistent with the requirement for divalent cations for the cleavage activity (Reines, 1992). In addition, recent studies on GreB, the prokaryotic homologue of TFIIS, have also provided evidence for a two metal ion mechanism for the 3' transcript cleavage activity of bacterial RNAP (Opalka et al., 2003; Sosunova et al., 2003).

In order to investigate the role of the loops of Rpb1 and Rpb2 strategically located within the catalytic center, we have created and characterized a series of mutants carrying deletions or substitutions in these structural elements. We have affinity purified wild type and mutant RNAP II from human cells through a Tandem Affinity Peptide (TAP) tag placed at the C-terminal end of either Rpb1 or Rpb2. We found that among the mutants that were able to correctly assemble into a 12-subunit core complex, most were blocked in the early stages of the transcription mechanism. We have also studied mutant RNAP II having alanine substitutions in Rpb2 amino acids proposed to coordinate the second Mg^{2+} ion, namely residues E791 and D792, equivalent to yeast E836 and D837. We found that the D792A substitution in Rpb2 impairs the assembly of RNAP II, thereby preventing the study of its effect during transcription. In contrast, the E791A substitution does not affect assembly of the enzyme but rather makes its polymerization activity more sensitive to low NTP concentrations and both its polymerization and cleavage activities more sensitive to low Mg^{2+} concentrations. Accordingly, we found that the equivalent substitution in yeast Rpb2, E836A, confers sensitivity to low nucleotide concentrations *in*

vivo. Our results demonstrate for the first time that amino acid E791 (E836 in yeast) participates in binding of metal B and is important both for the polymerization and cleavage reactions. They also support the hypothesis that this second Mg^{2+} ion enters the catalytic center with the incoming NTP. We propose, for the polymerization reaction, that Rpb2 E791 interacts with the NTP- Mg^{2+} complex and participates in its loading to the active site prior to catalysis.

RESULTS AND DISCUSSION

Mutations in loops of the RNAP II active center affect enzyme assembly and formation of the preinitiation complex

We have programmed human 293 cells to express, upon induction with the ecdysone analog ponasterone A, physiological levels of either wild-type or mutated RNAP II subunits carrying a TAP tag and used TAP purification to isolate variant forms of RNAP II from cell extracts (Jeronimo *et al.*, 2004). Purified complexes were visualized by silver-staining of SDS gels. Figure 1 lists all the Rpb1 and Rpb2 loops and motifs that were targeted in our analysis. Of a total of 25 mutants tested, 16 were defective in supporting the stable assembly of the core 12-subunit RNAP II. Strikingly, mutations affecting Rpb2 appear to affect more often the assembly and/or stability of the enzyme than mutations in Rpb1. Deletions of two amino-acids and triple alanine substitutions in some loops of Rpb2 led to the inability of the polymerase to assemble into a stable complex. In contrast, 12-subunit RNAP II was purified from TAP-tagged Rpb1 mutants carrying large deletions of the zipper and the rudder and extensive substitutions in switch 2. Because the mutations impairing the assembly of stable RNAP II are not in domains directly involved, based on the crystallographic data, in subunit-subunit interactions (Cramer *et al.*, 2001; Gnatt *et al.*, 2001; Westover *et al.*, 2004a), we suggest that they affect the folding of the mutated subunit itself, preventing appropriate association with its partner subunits.

Four mutants targeting loops of Rpb1 and Rpb2 were found to assemble correctly *in vivo*, but present major defects at early stages of the transcription process *in vitro*. Mutants zip Δ 45-55, rud Δ 321-334 and the previously characterized sw3 R1078A S1079A R1080A are defective in the formation of a preinitiation complex on promoter DNA in the presence of the general initiation factors (Fig. 2A). Not surprisingly, these 3 mutants, but not the wild type enzyme, were unable to initiate transcription *in vitro* ((Jeronimo *et al.*, 2004) and data not shown). They were also unable to support transcription from a C-tailed template in an assay that bypasses the requirement for the general initiation factors ((Jeronimo *et al.*, 2004) and data not shown). The fourth mutant, fork1 Δ 458-459, was as active as wild type in forming a preinitiation complex on promoter DNA, but was unable to initiate transcription (Jeronimo *et al.*, 2004).

A possible explanation for the failure of RNAP II mutants in switch 3, the zipper and the rudder to enter a preinitiation complex with the general initiation factors on promoter DNA could be that these mutant proteins are unable to interact with DNA properly. This would also account for their defect in transcribing from a C-tailed template. Alternatively, it is possible that the mutants are unable to interact with one of the general initiation factors involved in the formation of the preinitiation complex. This alternative explanation would imply, however, that the mutants have a second defect in elongation in order to explain their inability to transcribe in the general initiation factor-independent C-tailed template assay. Recently, the finger domain of TFIIB was shown to enter, through the RNA exit channel, deep in the polymerase catalytic site near fork loop 2 ((Chen et Hahn, 2004; Bushnell *et al.*, 2004); see Fig. 2B). It is therefore possible that one (or some) of our RNAP II mutants is deficient in its interaction with TFIIB, thereby preventing the formation of a preinitiation complex.

RNAP II mutant MB E791A is sensitive to low NTP concentrations

The Rpb2 mutants MB E791A, MB D792A, and MB E791A/D792A were designed to experimentally assess whether these highly conserved residues are involved in the coordination of Mg^{2+} -B in the active center, as predicted by early crystallographic

data (Cramer et al., 2001; Gnatt et al., 2001). Only MB E791A assembled correctly in the cell extract (Fig. 3A) and was able to form a preinitiation complex with the general initiation factors on promoter DNA *in vitro* (Fig. 3B). Both MB D792A and MB E791A/D792A lacked some of the small RNAP II subunits. As expected from their assembly defect, both MB D792A and MB E791A/D792A were also defective in run-off transcription reactions (data not shown). In contrast, the correctly assembled mutant MB E791A was able to synthesize an accurately initiated 391-nt run-off transcript, albeit less efficiently than the wild type enzyme (Fig. 4A).

Previous reports have shown that the substitution equivalent to E791A in *E. coli* (Lee et al., 1991) and in the archae *M. jannaschii* (Werner et Weinzierl, 2002) impaired the ability of these enzymes to transcribe at low NTP concentrations. As shown in Figure 4A, the human mutant MB E791A is also sensitive to low NTP concentrations *in vitro*. Increasing the concentration of non-radioactive UTP in the assay partially alleviated the failure of MB E791A to produce run-off transcripts, but had no effect on the wild type enzyme. Because the same amount of radioactive UTP (0.08 μM , 0.25 μCi) was used in all of the assays, increasing the concentration of unlabeled UTP from 2.5 μM UTP to 10 μM resulted in less incorporation of the labeled NTP in the 391-nt transcript synthesized by the wild type enzyme, as indicated by the lower intensity of that band. Further increasing the concentration of non-radioactive UTP above 40 μM led to a decrease in signal for both the wild type and the mutant enzyme.

To confirm the sensitivity of MB E791A to NTP concentration, we immobilized the biotinylated template on streptavidin coated beads. The template used contained the AdML promoter which directed the synthesis of a 24-nt transcript in the absence of GTP as it contains a G residue at position +24 (see Fig. 4B for a schematic representation). Wild type RNAP II or MB E791A were pre-assembled with the GTF on the immobilized template. Transcription was allowed to proceed for 30 minutes in the presence of a mix of NTP (Mix 1). Mix 1 contained 100 μM ATP, CTP and 5 μCi of [α - ^{32}P] UTP (0.08 μM) and allowed the synthesis of transcripts of various lengths (Fig. 4B, lane 1 and 6). Under these conditions, the wild type enzyme was able to produce predominantly the 24-nt

transcript. In contrast, the majority of the MB E791A enzyme was stalled at around +10, indicating that the mutant enzyme is impaired in its ability to transcribe at limiting NTP concentrations. Increasing the concentration of non-radioactive NTP in Mix 2 stimulated production of the 24-nt transcript by the mutant polymerase (compare lanes 7 to 10 to lanes 2 to 5). At 0.64 μ M of NTP, the mutant was as active as the wild type enzyme.

Figure 4C shows that the equivalent mutant in *S. cerevisiae*, Rpb2 E836A, is sensitive to 6-Azauracil (6-AU) *in vivo*. Growth of yeast cells in the presence of 6-AU, an inhibitor of IMP-dehydrogenase, leads to a depletion of the intracellular pool of GTP and UTP (Exinger et Lacroute, 1992). As a result, most yeast mutants with a defect in transcriptional elongation show a growth defect on 6-AU containing medium (Riles et al., 2004). This was found to be the case for the yeast RNAP II mutant E836A, which showed severely reduced growth in the presence of 30 μ g/ml 6-AU, as compared to WT cells (Fig. 4C). We found that the 6-AU sensitivity imposed by the E836A substitution in Rpb2 was more pronounced than that associated with the previously reported *rpo21-18* or *rpo21-24* mutation in Rpb1 (data not shown), which affect the binding of TFIIS to RNAP II (Archambault *et al.*, 1992b). These results suggest that the yeast E836A RNAP II is sensitive to low nucleotide concentrations *in vivo*, consistent with the *in vitro* transcription results obtained for the human RNAP II counterpart (E791A). In the absence of 6-AU (left panel), the yeast strain carrying the E836A mutant *rpb2* allele did not present any growth defect, suggesting that the E836 side chain is not essential under these conditions.

MB E791A is more sensitive than wild type RNAP II to MgCl₂ concentrations both during NTP polymerization and transcript cleavage

To test the hypothesis that E836/E791 is involved in the coordination of the second Mg²⁺ ion, we looked at the ability of MB E791A to transcribe at different MgCl₂ concentrations. Figure 5A and 5B show that MB E791A is more sensitive to low MgCl₂ concentrations than wild type when assayed for the production of the 24-nt transcript

from immobilized templates. These results strongly suggest that Rpb2 amino acid E791 is involved in the binding of metal B during NTP polymerization.

Since a two metal ion mechanism was also proposed for the cleavage activity of RNAPs (Jeon et al., 1994; Opalka et al., 2003; Sosunova et al., 2003), we tested the ability of the mutant MB E791A to cleave the 3' end of a transcript. In the experiments shown in Figure 5C and 5D, wild type and mutant RNAP II were stalled at G+24 using our transcription assay on immobilized templates, with a Mix 2 containing 100 μ M ATP, CTP and UTP. Complexes were washed and incubated at pH 9.5 in the absence of NTP for 1 hour. It has been shown that in a stalled complex, RNAP II is able to cleave the 3' end of a transcript at a basic pH in the absence of TFIIS (Weilbaecher et al., 2003). At pH 8, intrinsic cleavage is almost inexistent, but increasing the pH to 9.5 stimulates this RNAP II activity. The cleavage activity of RNAP II can also be greatly stimulated by the addition of TFIIS. Figure 5C shows that MB E791A is able to cleave the 3' end of a transcript as efficiently as the wild type. At pH 9.5, the wild type and the mutant cleave the RNA at two major sites, one at +22 and the other at +18. The addition of TFIIS stimulates both the wild type and the mutant to cleave, favoring the first (+22) site.

Previous studies using an *E. coli* mutant RNAP (E813A/D814A) analogous to our double mutant MB E791A/D792A (E813A/D814A in *E. coli*) have suggested that these two residues are not involved in binding of metal B in the catalytic site (Sosunov et al., 2003). Specifically, using a cleavage assay similar to the one used in Figure 5C, Goldbarb and colleagues have shown that the ability of the mutant enzyme to cleave the 3' end of a transcript was equivalent to that of the wild type enzyme at all $MgCl_2$ concentrations tested. We therefore determined the effect of decreasing the $MgCl_2$ concentration on the cleavage activity of MB E791A at pH 9.5 (Fig. 5D). We measured the disappearance of the 24-nt band for the wild type and the mutant at different $MgCl_2$ concentrations (x axis). The y axis represents the ratio of the intensity of the cleaved band to the uncleaved band. In contrast to Sosunov et al. (Sosunov et al., 2003) the graph shows that at low $MgCl_2$ concentrations, MB E791A is less active than wild type. At higher $MgCl_2$ concentrations (5 mM), the mutant is as efficient as the wild type.

Together, these results show that Rpb2 amino acid E791 is involved in binding of Mg^{2+} in the active site of RNAP II for both the polymerization and cleavage reactions. The discrepancy between our results and those of Goldfarb and colleagues might come from the fact that $MgCl_2$ concentrations lower than 5 mM were not tested by Sosunov et al. (Sosunov et al., 2003). In the model presented in their paper, Sosunov et al. do not implicate residue E813 in coordination of metal B for the cleavage reaction. However, they implicate D814, but only at high pH as a result of a salt-bond breakage between Rpb2 amino acids R1106 and D814. In contrast, our results suggest an implication of E791 (E813) in binding of Mg^{2+} in the active site for the cleavage reaction.

TFIIS stimulates transcription by both wild type and MB E791A RNAP II at limiting NTP concentrations *in vitro*

In yeast, the TFIIS gene (*PPR2*) is not essential for viability (Williams et Kane, 1996; Nakanishi *et al.*, 1992). However, knocking out the gene confers sensitivity to 6-AU (Williams et Kane, 1996; Nakanishi *et al.*, 1992). *In vitro* experiments have shown that reducing the concentration of NTPs causes RNAP II to pause and overall elongate less efficiently (Izban et Luse, 1992b; Reines, 1992; Reines *et al.*, 1992). Since MB E791A is sensitive to low nucleotide concentrations, we decided to test the effect of TFIIS on transcription by this mutant *in vitro*. Figure 5E shows that at 10 μ M UTP, TFIIS stimulates synthesis of the 391-nt transcript by MB E791A, but has no effect on transcription by the wild type enzyme. Reducing UTP concentrations from 2.5 to 0.625 μ M causes a decrease in the synthesis of the run-off transcript by the wild type enzyme and causes the enzyme to arrest at different sites (Fig. 5F). At these limiting concentrations of NTP, TFIIS now stimulates transcription by the wild type RNAP II. These *in vitro* results confirm that TFIIS helps RNAP II to transcribe at low NTP concentrations, as suggested by previous *in vivo* experiments (Williams et Kane, 1996; Nakanishi *et al.*, 1992). We hypothesize that TFIIS helps RNAP II to resume arrest at low NTP concentrations by stimulating its 3' transcript cleavage activity, possibly by stabilizing the binding of metal B in the active site.

Metal B mutants are defective for transcription *in vivo*

In order to determine the transcriptional defects of the three human metal B mutants *in vivo*, ChIP experiments were performed using IgG beads to precipitate the different TAP-tagged RNAP II bound to chromatin (Fig. 6). Quantitative PCR was used to determine the enrichment of wild type and mutant RNAP II on the promoter and coding region of 3 actively transcribed genes relative to an untranscribed region. The 3 transcribed genes used in these studies were: the ferritin light polypeptide gene (*FTL*), the guanine nucleotide binding protein beta polypeptide 2-like 1 gene (*GNB2L1*) and the enolase 1 (alpha) gene (*ENO1*).

As we previously described (Jeronimo *et al.*, 2004), wild type TAP-tagged RNAP II was found at the promoter and in the transcribed region of the 3 genes tested, in substantially greater amounts than in the negative control (mock-transfected cell line). Consistent with the observation that MB E791A is able to form a pre-initiation complex *in vitro*, this mutant was present at the promoter of *FTL*, *GNB2L1* and *ENO1* genes *in vivo*. Our results indicate that MB E791A is also located in the coding region of all 3 genes, albeit at a lower level than wild type. For the *FTL* and *GNB2L1* genes, the MB E791A mutation had a greater effect on the density of RNAP II in the coding region than at the promoter. For the *ENO1* gene, the mutant is similarly affected in both regions. Taken together, our ChIP results indicate that the E791A mutation has a more profound impact on transcript elongation than promoter recruitment *in vivo*.

We propose that a similar situation takes place in yeast cells. In the yeast strain carrying the E836A mutation, the RNAP II mutant is probably altered in its elongation efficiency on some genes, but to an extent that does not affect growth under normal conditions. The addition of the drug 6-AU has a dramatic effect on growth of the mutant since it further reduces the altered enzyme's capacity to elongate by lowering the concentration of nucleotides in the cell (see Fig. 4C). The finding that mutant MB E791A is less efficiently recruited to promoters and less present in the transcribed regions *in vivo* cannot be explained solely by the *in vitro* defects presented in this article. We speculate

that, *in vivo*, other factors that regulate transcription interact with the mutated polymerase and affect its ability to associate with promoters and to transcribe.

Molecular modeling of the active center of RNAP II

Recently, Kornberg and colleagues (Westover et al., 2004b) have published the crystal structure of yeast RNAP II bound to a DNA-RNA duplex and a nucleotide positioned at the $i+1$ site, ready for bond formation. In this structure, Mg^{2+} -B was found to be located in proximity to Rpb1 residues D481 and D483 and Rpb2 D837. On this basis, the authors proposed that these amino acids are implicated in Mg^{2+} -B coordination at the active site. At this step of the polymerization process, Rpb2 E836 is located too far from Mg^{2+} -B to interact with it. The resolution of a RNAP II structure with a mismatched nucleotide revealed the existence of a second binding site for the incoming NTP, called the “entry” or E site. It was suggested that the NTP- Mg^{2+} complex enters the active site through the pore and binds to the E site in an inverted orientation, followed by a rotation and subsequent binding to the so-called “addition” or A site.

To account for our results showing that the mutant E791A is sensitive to NTP and $MgCl_2$ concentrations during the polymerization reaction, we propose that Rpb2 E836/791 is implicated in loading the NTP- Mg^{2+} into the E site (Fig.7A). If so, the sensitivity to NTP concentrations in the polymerization assay would be explained by the fact that, at low NTP concentrations, the number of NTP- Mg^{2+} that enters the pore is reduced, affecting the ability of the mutant to load it into the active site. The interaction of E836/791 and the metal B is predicted to be transient, as the crystallographic data has shown that when the E site is occupied by a NTP, metal B is located too far from E836 to be coordinated by it. In support of our hypothesis, amino acid E836 is the most accessible acidic residue in the vicinity of the active site encountered by the incoming NTP- Mg^{2+} , when it enters from the pore. We have called this new position the “pre-loading” (P) site (panel 1). In this state, metal B is located at a distance suitable for binding by Rpb2 E836 and D837 and the β and γ phosphates of the NTP. The NTP- Mg^{2+} is next loaded in the E site at a position where metal B can interact with Rpb2 D837, Rpb1 D481 and D483 and

the NTP β and γ phosphates, as described (Westover et al., 2004b) (panel 2). The nucleotide is next flipped and positioned at the A site for catalysis (panel 3).

In a recent series of papers, Burton and colleagues have proposed that the incoming NTPs access the active center through the main channel (Nedialkov *et al.*, 2003; Zhang *et al.*, 2003; Zhang et Burton, 2004). Furthermore, these authors have proposed an NTP-driven translocation model in which the incoming NTP pairs with the DNA base at the $i+2$ site. According to this model, this pairing is necessary to induce translocation of the NTP at $i+1$ site (A site) for catalysis and allows substrate selection prior to arrival into the catalytic site. Our data is also consistent with the idea that the NTPs access the catalytic center through the main channel. This would imply, as depicted in Figure 7B, that translocation of the NTP-Mg²⁺ from the $i+2$ site (panel 1) to the A site (panel 3) transit through an interaction with the E836 (panel 2). The interaction of E836 with NTP-Mg²⁺ would play a crucial role in loading the NTP or retaining it in the active site for catalysis. If this model is correct, the position of the E site would be more consistent with it being a site for expulsion and subsequent exit of a mismatched nucleotide.

The mutant MB E791A is also sensitive to MgCl₂ concentration in the cleavage reaction, suggesting that E791 is implicated in binding of metal B for this activity of the polymerase. The locations of E836 and the second Mg²⁺ ion revealed by the crystal structures of elongating RNAPII published by Kornberg and colleagues is not consistent with a coordination of metal B by E836 (Westover et al., 2004b). Therefore, we propose that a reconfiguration of the active site is necessary for the cleavage activity of the backtracked RNAP II and places E836 and metal B at a distance suitable for interaction.

EXPERIMENTAL PROCEDURES

Protein factors

Recombinant yeast TBP (Ingles et al., 1991), human TFIIB (Ha et al., 1991), RAP74 and RAP30 (Finkelstein et al., 1992), TFIIE56/ α and TFIIE34/ β (Ohkuma et al., 1991; Sumimoto et al., 1991; Peterson et al., 1991), TFIIH (Gerard et al., 1991; Douziech et al., 2000), calf thymus RNAP II (Hodo, III et Blatti, 1977) and GST-TFIIS (Robert et al., 2002) were purified as previously described.

Tandem affinity purification of wild type and mutant RNAP II

TAP-tagged wild type and mutant RNAP II subunits were expressed in human EcR-293 cells and purified as previously described (Rigaut *et al.*, 1999; Jeronimo *et al.*, 2004).

Gel mobility shift assay

Gel mobility shift assays were performed as previously described (Langelier et al., 2001) with the following modifications. Complexes were assembled using highly purified TBP (40 ng), TFIIB (120 ng), RAP30 (200 ng), RAP74 (400 ng), TFIIE34 (80 ng), TFIIE56 (110 ng) and wild type or mutant RNAP II (450 ng).

Run-off transcription assay

Transcription reactions were performed as described previously (Robert et al., 2002). TBP (30 ng), TFIIB (30 ng), RAP30 (30 ng), RAP74 (65 ng), TFIIE34 (40 ng), TFIIE56 (60 ng), TFIIH (4 ng) and wild type or mutant RNAP II (165 ng) were incubated with 500 ng of linearized DNA template containing the adenovirus major late (AdML) promoter from nucleotides -50 to +10 fused to a G-less cassette. Reactions were

performed at various concentrations of NTP specified in the figure legends. Under these conditions a 391-nt run-off transcript is produced.

Transcription assay on immobilized template

A biotinylated template containing the AdML promoter from nucleotides -50 to +10 fused to a G-less cassette and modified to contain a G at position +24 was immobilized on Dynabeads (Dynal, Oslo, Norway) according to the manufacturers' recommendations. The biotinylated template (225 ng) was pre-incubated with TBP (15 ng), TFIIB (15 ng), RAP30 (15 ng), RAP74 (30 ng), TFIIE34 (20 ng), TFIIE56 (30 ng), TFIIH (2 ng) and RNAP II (83 ng) for 30 minutes at 30°C in transcription buffer (12 mM Hepes pH 7.9, 50 mM KCl, 0.1 mM EDTA, 0.1 mM EGTA, 10 % glycerol, 1 mM DTT). Mix 1 was added (100 μM ATP, 100 μM CTP and 5 μCi of [α -³²P] UTP (0.08 μM), 3 mM EGTA, 12.5 mM MgCl₂, 1 U/μl RNase inhibitor) and transcription was allowed to proceed for 30 minutes at 30°C. Beads were then washed sequentially with 10 volumes of 2X transcription buffer containing 0.3% sarkosyl and 1mg/ml BSA, 5 volumes of 2X transcription buffer containing 1mg/ml BSA and 5 volumes of 1X transcription buffer containing 1mg/ml BSA. Beads were resuspended in 1X transcription buffer and incubated for 30 minutes at 30°C with Mix 2 (variable concentrations of ATP, CTP and UTP, 3 mM EGTA, 12.5 mM MgCl₂, 1 U/μl RNase inhibitor; see figure legends). Transcription was stopped by adding 200 μl of stop solution (0.1 M sodium acetate, 0.5% SDS, 2 mM EDTA, 100 μg/ml tRNA). After phenol/chloroform extraction and ethanol precipitation, transcripts were analyzed on an 18% polyacrylamide, 9% urea denaturing gel.

Yeast strains and manipulations

The *S. cerevisiae* strain yBC-1, which conditionally expresses Rpb2 under the control of the Gal1 promoter was constructed in two steps. First, strain W303-1A (*MATa can1-100 his3-11,15 leu2-3,112 trp1-1 ura3-1 ade2-1*) was transformed with the plasmid

p*GAL-RPB2 (URA3)*. Second, the chromosomal *RPO22* allele of this merodiploid strain was disrupted with *LEU2*. The resulting strain, yBC-1, is dependent on galactose for growth. To test the phenotype associated with the rpo22 E836A mutation, yBC-1 was transformed with a vector (*TRP1, CEN, ARS*) containing either the mutant *RPB2* E836A allele, or the wild-type *RPO22* allele as a control, and the transformants were streaked onto solid, synthetic glucose-medium lacking or containing 6-AU (30 µg/ml) and grown at 30°C for 5 days. Yeast transformations were performed as previously described (Ito et al., 1983). Details on the construction of these plasmids will be provided upon request.

Cleavage assay

RNAP II complexes stalled at nucleotide +24 were formed as described above with Mix 2 containing 100 µM of non-radioactive ATP, CTP and UTP, 3 mM EGTA, 12.5 mM MgCl₂, 1 U/µl RNase inhibitor. Beads were washed 3 times with 5 volumes of 1X transcription buffer containing 1 mg/ml BSA. Beads were resuspended either in 20 µl of cleavage buffer pH 9.5 (70 mM CAPS-KOH pH 9.5, 100 mM KCl, 5mM MgCl₂, 5% glycerol, 1 mM DTT) and incubated for one hour at 30⁰, or in 20 µl of TFIIS containing buffer pH 8 (70 mM Tris pH 8, 100 mM KCl, 5mM MgCl₂, 5% glycerol, 1 mM DTT, 0.42 ng/µl TFIIS) and incubated for 5 minutes at 30⁰C. In the experiments shown in Figure 5D, various concentrations of MgCl₂ were used in the cleavage buffer pH 9.5 as indicated. Cleavage reactions were stopped by the addition of 200 µl of stop solution and transcripts were processed as described above.

Chromatin Immunoprecipitation

Chromatin immunoprecipitation (ChIP) assays were performed as previously described (Nissen et Yamamoto, 2000; Takahashi *et al.*, 2000) with the modifications implemented by Jeronimo et al. (Jeronimo *et al.*, 2004). Cell lines expressing the following RNAP II mutants were analyzed: wt RNAP II, MB E791A/D792A, MB E791A and MB D792A. A mock-transfected cell line was used as a negative control. Quantitative PCR (Q-PCR) experiments were performed as previously described (Bryant

et Ptashne, 2003). Primer pairs were designed for both the promoter and transcribed regions of *GNB2L1* (-200 to +55 and +998 to +1165 respectively), *FTL* (-157 to +33 and +955 to +1176 respectively) and *ENO1* (-90 to +91 and +1143 to +1292 respectively) genes. The position of the promoter and transcribed regions was selected based on the major transcription start site listed in the DataBase of Transcription Start Site (DBTSS, April 2003 version, http://dbtss.hgc.jp/samp_home.html) and the Mammalian Gene Collection (MGC, March 2003 version, <http://mgc.nci.nih.gov/>). Oligonucleotide sequences will be made available upon request. Relative immunoprecipitation efficiency was calculated by dividing the amount of DNA fragments spanning the promoter or coding region of the specified genes by the amount of DNA fragments from a conserved gene-less region located in the 13q21.33 band (Nobrega et al., 2003) in the immunoprecipitated sample. Data points originate from two independent sets of duplicate Q-PCR reactions using chromatin obtained from three independent immunoprecipitations, giving a total of twelve data points for each immunoprecipitation.

Molecular Modeling

Modeling was performed using the crystallographic structure of elongating yeast RNAP II (Westover et al., 2004b; Gnatt et al., 2001) and the program Swiss PDP viewer.

ACKNOWLEDGEMENTS

We are grateful to the members of our laboratory for helpful discussions and comments on this paper. We thank Diane Bourque for artwork and Julie Edwards for critical reading of the manuscript. This work is supported by grants from the Canadian Institutes for Health Research (to B.C.), Genome Canada (to B.C., J.G.), Genome Québec (to B.C.). M.F.L. holds a studentship from the Fonds québécois de la recherche sur la nature et les technologies and V.T from the Fonds de la recherche en santé du Québec. B.C. is a senior scholar from the Fonds de la recherche en santé du Québec.

Reference List

Archambault,J., Lacroute,F., Ruet,A., and Friesen,J.D. (1992). Genetic interaction between transcription elongation factor TFIIS and RNA polymerase II. *Mol Cell Biol* *12*, 4142-4152.

Armache,K.J., Kettenberger,H., and Cramer,P. (2003). Architecture of initiation-competent 12-subunit RNA polymerase II. *Proc. Natl. Acad. Sci. U. S. A* *100*, 6964-6968.

Borukhov,S., Polyakov,A., Nikiforov,V., and Goldfarb,A. (1992). GreA protein: a transcription elongation factor from *Escherichia coli*. *Proc. Natl. Acad. Sci. U. S. A* *89*, 8899-8902.

Borukhov,S., Sagitov,V., and Goldfarb,A. (1993). Transcript cleavage factors from *E. coli*. *Cell* *72*, 459-466.

Bryant,G.O. and Ptashne,M. (2003). Independent recruitment in vivo by gal4 of two complexes required for transcription. *Mol. Cell* *11*, 1301-1309.

Bushnell,D.A. and Kornberg,R.D. (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., Westover,K.D., Davis,R.E., and Kornberg,R.D. (2004). Structural basis of transcription: an RNA polymerase II-TFIIB cocrystal at 4.5 Angstroms. *Science* *303*, 983-988.

Chen,H.T. and Hahn,S. (2004). Mapping the location of TFIIB within the RNA polymerase II transcription preinitiation complex: a model for the structure of the PIC. *Cell* *119*, 169-180.

Conaway,J.W., Shilatifard,A., Dvir,A., and Conaway,R.C. (2000). Control of elongation by RNA polymerase II. *Trends Biochem. Sci.* *25*, 375-380.

Coulombe,B. and Burton,Z.F. (1999). DNA bending and wrapping around RNA polymerase: a "revolutionary" model describing transcriptional mechanisms. *Microbiol. Mol. Biol. Rev.* *63*, 457-478.

Cramer,P., Bushnell,D.A., and Kornberg,R.D. (2001). Structural basis of transcription: RNA polymerase II at 2.8 angstrom resolution. *Science* *292*, 1863-1876.

Doublet,S., Tabor,S., Long,A.M., Richardson,C.C., and Ellenberger,T. (1998). Crystal structure of a bacteriophage T7 DNA replication complex at 2.2 A resolution. *Nature* *391*, 251-258.

Douziech,M., Coin,F., Chipoulet,J.M., Arai,Y., Ohkuma,Y., Egly,J.M., and Coulombe,B. (2000). Mechanism of promoter melting by the xeroderma pigmentosum complementation group B helicase of transcription factor IIH revealed by protein-DNA photo-cross-linking. *Mol. Cell Biol.* 20, 8168-8177.

Exinger,F. and Lacroute,F. (1992). 6-Azauracil Inhibition of Gtp Biosynthesis in *Saccharomyces-Cerevisiae*. *Current Genetics* 22, 9-11.

Finkelstein,A., Kostrub,C.F., Li,J., Chavez,D.P., Wang,B.Q., Fang,S.M., Greenblatt,J., and Burton,Z.F. (1992). A cDNA encoding RAP74, a general initiation factor for transcription by RNA polymerase II. *Nature* 355, 464-467.

Fish,R.N. and Kane,C.M. (2002). Promoting elongation with transcript cleavage stimulatory factors. *Biochim. Biophys. Acta* 1577, 287-307.

Gerard,M., Fischer,L., Moncollin,V., Chipoulet,J.M., Chambon,P., and Egly,J.M. (1991). Purification and interaction properties of the human RNA polymerase B(II) general transcription factor BTF2. *J. Biol. Chem.* 266, 20940-20945.

Gnatt,A.L., Cramer,P., Fu,J., Bushnell,D.A., and Kornberg,R.D. (2001). Structural basis of transcription: an RNA polymerase II elongation complex at 3.3 Å resolution. *Science* 292, 1876-1882.

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

Hodo,H.G., III and Blattl,S.P. (1977). Purification using polyethylenimine precipitation and low molecular weight subunit analyses of calf thymus and wheat germ DNA-dependent RNA polymerase II. *Biochemistry* 16, 2334-2343.

Ingles,C.J., Shales,M., Cress,W.D., Triezenberg,S.J., and Greenblatt,J. (1991). Reduced binding of TFIID to transcriptionally compromised mutants of VP16. *Nature* 351, 588-590.

Ito,H., Fukuda,Y., Murata,K., and Kimura,A. (1983). Transformation of intact yeast cells treated with alkali cations. *J. Bacteriol.* 153, 163-168.

Izban,M.G. and Luse,D.S. (1992). 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.

Jeon,C., Yoon,H., and Agarwal,K. (1994). The transcription factor TFIIS zinc ribbon dipeptide Asp-Glu is critical for stimulation of elongation and RNA cleavage by RNA polymerase II. *Proc. Natl. Acad. Sci. U. S. A* 91, 9106-9110.

Jeronimo,C., Langelier,M.F., Zeghouf,M., Cojocar,M., Bergeron,D., Baali,D., Forget,D., Mnaimneh,S., Davierwala,A.P., Pootoolal,J., Chandy,M., Canadien,V., Beattie,B.K., Richards,D.P., Workman,J.L., Hughes,T.R., Greenblatt,J., and

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, 7043-7058.

Kettenberger,H., Armache,K.J., and Cramer,P. (2003). Architecture of the RNA polymerase II-TFIIS complex and implications for mRNA cleavage. *Cell* 114, 347-357.

Kiefer,J.R., Mao,C., Braman,J.C., and Beese,L.S. (1998). Visualizing DNA replication in a catalytically active *Bacillus* DNA polymerase crystal. *Nature* 391, 304-307.

Krummel,B. and Chamberlin,M.J. (1989). RNA chain initiation by *Escherichia coli* RNA polymerase. Structural transitions of the enzyme in early ternary complexes. *Biochemistry* 28, 7829-7842.

Landick,R. (2004). Active-site dynamics in RNA polymerases. *Cell* 116, 351-353.

Langelier,M.F., Forget,D., Rojas,A., Porlier,Y., Burton,Z.F., and Coulombe,B. (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.

Lee,J., Kashlev,M., Borukhov,S., and Goldfarb,A. (1991). A β Subunit Mutation Disrupting the Catalytic Function of *Escherichia coli* RNA Polymerase. *PNAS* 88, 6018-6022.

Nakanishi,T., Nakano,A., Nomura,K., Sekimizu,K., and Natori,S. (1992). Purification, gene cloning, and gene disruption of the transcription elongation factor S-II in *Saccharomyces cerevisiae*. *J. Biol. Chem.* 267, 13200-13204.

Nedialkov,Y.A., Gong,X.Q., Hovde,S.L., Yamaguchi,Y., Handa,H., Geiger,J.H., Yan,H., and Burton,Z.F. (2003). NTP-driven Translocation by Human RNA Polymerase II. *J. Biol. Chem.* 278, 18303-18312.

Nissen,R.M. and Yamamoto,K.R. (2000). The glucocorticoid receptor inhibits NF κ B by interfering with serine-2 phosphorylation of the RNA polymerase II carboxy-terminal domain. *Genes Dev.* 14, 2314-2329.

Nobrega,M.A., Ovcharenko,I., Afzal,V., and Rubin,E.M. (2003). Scanning human gene deserts for long-range enhancers. *Science* 302, 413.

Nudler,E., Mustaev,A., Lukhtanov,E., and Goldfarb,A. (1997). The RNA-DNA hybrid maintains the register of transcription by preventing backtracking of RNA polymerase. *Cell* 89, 33-41.

Ohkuma,Y., Sumimoto,H., Hoffmann,A., Shimasaki,S., Horikoshi,M., and Roeder,R.G. (1991). Structural motifs and potential sigma homologies in the large subunit of human general transcription factor TFIIE. *Nature* 354, 398-401.

Opalka,N., Chlenov,M., Chacon,P., Rice,W.J., Wriggers,W., and Darst,S.A. (2003). Structure and function of the transcription elongation factor GreB bound to bacterial RNA polymerase. *Cell* 114, 335-345.

Peterson,M.G., Inostroza,J., Maxon,M.E., Flores,O., Admon,A., Reinberg,D., and Tjian,R. (1991). Structure and functional properties of human general transcription factor IIE. *Nature* 354, 369-373.

Reinberg,D. and Roeder,R.G. (1987). Factors involved in specific transcription by mammalian RNA polymerase II. Transcription factor IIS stimulates elongation of RNA chains. *J. Biol. Chem.* 262, 3331-3337.

Reines,D. (1992). Elongation factor-dependent transcript shortening by template-engaged RNA polymerase II. *J. Biol. Chem.* 267, 3795-3800.

Reines,D., Chamberlin,M.J., and Kane,C.M. (1989). Transcription elongation factor SII (TFIIS) enables RNA polymerase II to elongate through a block to transcription in a human gene in vitro. *J. Biol. Chem.* 264, 10799-10809.

Reines,D., Ghanouni,P., Li,Q.Q., and Mote,J., Jr. (1992). The RNA polymerase II elongation complex. Factor-dependent transcription elongation involves nascent RNA cleavage. *J. Biol. Chem.* 267, 15516-15522.

Rigaut,G., Shevchenko,A., Rutz,B., Wilm,M., Mann,M., and Seraphin,B. (1999). A generic protein purification method for protein complex characterization and proteome exploration. *Nat. Biotechnol.* 17, 1030-1032.

Riles,L., Shaw,R.J., Johnston,M., and Reines,D. (2004). Large-scale screening of yeast mutants for sensitivity to the IMP dehydrogenase inhibitor 6-azauracil. *Yeast* 21, 241-248.

Robert,F., Blanchette,M., Maes,O., Chabot,B., and Coulombe,B. (2002). A human RNA polymerase II-containing complex associated with factors necessary for spliceosome assembly. *J. Biol. Chem.* 277, 9302-9306.

Shilatifard,A., Conaway,R.C., and Conaway,J.W. (2003). The RNA polymerase II elongation complex. *Annu. Rev. Biochem.* 72, 693-715.

Sosunov,V., Sosunova,E., Mustaev,A., Bass,I., Nikiforov,V., and Goldfarb,A. (2003). Unified two-metal mechanism of RNA synthesis and degradation by RNA polymerase. *EMBO J.* 22, 2234-2244.

Sosunova,E., Sosunov,V., Kozlov,M., Nikiforov,V., Goldfarb,A., and Mustaev,A. (2003). Donation of catalytic residues to RNA polymerase active center by transcription factor Gre. *Proc. Natl. Acad. Sci. U. S. A* 100, 15469-15474.

Steitz,T.A. (1998). A mechanism for all polymerases. *Nature* 391, 231-232.

- Sumimoto,H., Ohkuma,Y., Sinn,E., Kato,H., Shimasaki,S., Horikoshi,M., and Roeder,R.G. (1991). Conserved sequence motifs in the small subunit of human general transcription factor TFIIE. *Nature* 354, 401-404.
- Surratt,C.K., Milan,S.C., and Chamberlin,M.J. (1991). Spontaneous cleavage of RNA in ternary complexes of Escherichia coli RNA polymerase and its significance for the mechanism of transcription. *Proc. Natl. Acad. Sci. U. S. A* 88, 7983-7987.
- Takahashi,Y., Rayman,J.B., and Dynlacht,B.D. (2000). Analysis of promoter binding by the E2F and pRB families in vivo: distinct E2F proteins mediate activation and repression. *Genes Dev.* 14, 804-816.
- Wang,D. and Hawley,D.K. (1993). Identification of a 3'-->5' exonuclease activity associated with human RNA polymerase II. *Proc. Natl. Acad. Sci. U. S. A* 90, 843-847.
- Weilbaecher,R.G., Awrey,D.E., Edwards,A.M., and Kane,C.M. (2003). Intrinsic transcript cleavage in yeast RNA polymerase II elongation complexes. *J. Biol. Chem.* 278, 24189-24199.
- Werner,F. and Weinzierl,R.O. (2002). A Recombinant RNA Polymerase II-like Enzyme Capable of Promoter-Specific Transcription. *Mol. Cell* 10, 635-646.
- Westover,K.D., Bushnell,D.A., and Kornberg,R.D. (2004a). Structural basis of transcription: separation of RNA from DNA by RNA polymerase II. *Science* 303, 1014-1016.
- Westover,K.D., Bushnell,D.A., and Kornberg,R.D. (2004b). Structural Basis of Transcription; Nucleotide Selection by Rotation in the RNA Polymerase II Active Center. *Cell* 119, 481-489.
- Williams,L.A. and Kane,C.M. (1996). Isolation and characterization of the Schizosaccharomyces pombe gene encoding transcript elongation factor TFIIS. *Yeast* 12, 227-236.
- Wind,M. and Reines,D. (2000). Transcription elongation factor SII. *Bioessays* 22, 327-336.
- Young,R.A. (1991). RNA polymerase II. *Annu. Rev. Biochem.* 60, 689-715.
- Zhang,C. and Burton,Z.F. (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., Yan,H., and Burton,Z.F. (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.

FIGURE LEGENDS

Figure 1. Summary of the different RNAP II mutants analyzed. The structural domains in Rpb1 and Rpb2 that were targeted by mutagenesis are indicated as well as a description of the different mutations. In total, 25 different RNAP II mutants have been generated, affinity purified and assayed for the assembly of the 12-subunit core enzyme, preinitiation complex (PIC) formation, abortive initiation, elongation on a C-tailed template and promoter-dependent run-off transcription. ND: Not Determined. NA: Not Applicable.

Figure 2. RNAP II mutants of the zipper and the rudder are defective in preinitiation complex formation. (A) The zip $\Delta 45-55$ and rud $\Delta 321-334$ are defective in forming a preinitiation complex on promoter DNA in the presence of the general initiation factors. EMSAs were performed using either wild type RNAP II or zip $\Delta 45-55$ and rud $\Delta 321-334$ in the presence of TBP, TFIIB, TFIIF and TFIIE on the AdML promoter. A control reaction assembled in the absence of TBP is included in each case. The position of the TBP-IIB-IIF-IIE-RNAP II-promoter complex is indicated. (B) Top panel. Crystal structure of yeast RNAPII in complex with the DNA-RNA hybrid (Westover et al., 2004a). The TFIIB structure (residues 17 to 119 from pdb 1R5U) (Bushnell et al., 2004) was modeled in the RNAPII structure 1SFO (Westover et al., 2004a). Middle panel. Specific elements of Rpb2 are shown in red. Subunits Rpb1 and Rpb3 to Rpb12 were omitted. Bottom panel. Specific elements of Rpb1 are shown in red. Subunits Rpb2 to Rpb12 were omitted. The view has been rotated 90° for clarity.

Figure 3. Purification of wild type and metal B mutant RNAP II and characterization of their ability to form a pre-initiation complex. (A) The TAP purification procedure was used to isolate wild type (wt) and mutant (MB E791A,D792A; MB E791A; MB D792A) RNAP II. The purified RNAP II were visualized on a SDS gel stained with silver. The position of molecular weight makers and RNAP II subunits are indicated. (B) An EMSA was performed using either wild type RNAP II or MB E791A in the presence of TBP, TFIIB, TFIIF and TFIIE. A control reaction assembled in the absence of TBP is included

in each case. The position of the TBP-IIB-IIF-IIE-RNAP II-promoter complex is indicated.

Figure 4. Mutant MB E791A is sensitive to NTP concentrations during polymerization. (A) *In vitro* transcription reactions were reconstituted with either wild type RNAP II or MB E791A in the presence of TBP, TFIIB, TFIIF, TFIIE and TFIIH. The linearized AdML template directs the synthesis of a 391-nt transcript. The reaction contained 750 μ M ATP and CTP, 2.5 μ Ci of [α - 32 P] UTP (0.04 μ M) and increasing amounts of non-radioactive UTP, as indicated in the figure. (B) *In vitro* transcription reactions on immobilized template. A scheme of the transcription procedure is presented. Transcriptional complexes were assembled with wild type RNAP II or MB E791 mutant in the presence of TBP, TFIIB, TFIIF, TFIIE and TFIIH on the immobilized template. The AdML template directs the synthesis of a 24-nt transcript in the absence of GTP. (C) The *S. cerevisiae* strain yBC-1 (RPB2::LEU2, pGAL-RPB2) carrying either the wild-type or mutant E836A *RPO22* allele on a plasmid (*TRP1*, *CEN*, *ARS*) were streaked onto solid, synthetic glucose-medium lacking or containing 6-AU (30 μ g/ml) and grown at 30°C for 5 days.

Figure 5. Mutant MB E791A is sensitive to MgCl₂ concentrations during both NTP polymerization and transcript cleavage. TFIIS stimulates transcription of both wild type RNAP II and mutant MB E791A at low NTP concentrations. (A) *In vitro* transcription reactions on immobilized template at different MgCl₂ concentrations in Mix 1. Reactions were stopped 30 minutes after addition of Mix1. (B) The intensity of the 24 nt-transcript band obtained at each MgCl₂ concentration was quantified using a phosphorimager and normalized to the intensity of the band obtained at 9.33 mM MgCl₂. These relative values are reported on the y axis of the graph and are the mean of three independent experiments. (C) Cleavage reactions were performed as described in Figure 4B except that Mix 2 contained 100 μ M ATP, 100 μ M CTP and 100 μ M UTP. After 30 minutes incubation with Mix 2, complexes were washed and incubated for one hour at pH 9.5 or for 5 minutes at pH 8 in the presence of 8.2 ng of TFIIS. (D) The intensity of the 24 nt-transcript band was quantified using a phosphorimager from cleavage reactions

performed for one hour at pH 9.5 in a buffer containing the indicated concentrations of MgCl_2 . The intensity of each band was divided by the intensity of the uncleaved band (reaction stopped after 30 minutes of incubation with Mix 2) and reported on the y axis of the graph. The ratios calculated are the mean of three independent experiments. (E) *In vitro* transcription reactions were reconstituted using wild type or mutant RNAP II in the presence of TBP, TFIIB, TFIIF, TFIIIE, TFIIH and various amounts of TFIIS where indicated. The reaction contained 750 μM ATP and CTP, 10 μM UTP and 2.5 μCi of [α - ^{32}P] UTP (0.04 μM). (F) Reactions were performed as in (E) for the wild type RNAP II with the exception that reactions contained various amounts of non-radioactive UTP and 4.1 ng of TFIIS, where indicated.

Figure 6. Association of wild type and mutated RNAP II with chromatin *in vivo*. Fold enrichment over the control region of both the promoter and the transcribed region of *FTL*, *GNB2L1* and *ENO1* genes following immunoprecipitation of chromatin fragments was determined using Q-PCR. Enrichment was normalized to a conserved gene-less region located in the 13q21.33 band (see Materials and Methods). Data points originate from two independent sets of duplicate Q-PCR reactions using chromatin obtained from three independent immunoprecipitations, giving a total of twelve data points for each immunoprecipitation. The negative control is a mock transfected cell line.

Figure 7. Proposed models for the NTP polymerization reaction. (A) The NTP-Mg(B) accessing the active center from the pore first binds to a “Pre-entry” (P) site where it can contact Rpb2 residue E836/E791. From the P site, the NTP-Mg(B) moves to the E site before rotating to access the A site where it is paired to the template DNA. (B) The incoming NTP-Mg(B) can alternatively enter the catalytic center through the enzyme’s main channel and associate with E836/E791 in the P site. The subsequent steps are as in (A). The exact location of the P site is speculative. See text for details. Incoming NTP: pink, Rpb1 D481, D483, D485, bridge helix: green, Rpb2 E836: yellow, Rpb2 D837: orange, RNA: magenta, DNA: blue, Metal A and B: red.

	Structural element	Mutant	12-subunit assembly	PIC formation	Abortive initiation	Elongation on C-tailed template	Promoter dependent transcription
Rpb1	Lid	lid Δ265-271	-	NA	NA	NA	NA
	Rudder	rud Δ321-334	+	-	-	-	ND
	Switch 1	sw S1431A F1432A E1433A	+	ND	ND	ND	ND
	Switch 2	sw2 K346A E347A G348A R349A V350A R351A	+	ND	ND	ND	ND
	Zipper	zip Δ45-55	+	-	-	-	-
Rpb2	Flap Loop	flap Δ876-887	+	+	ND	ND	ND
		flap Δ880-881	+	+	ND	ND	ND
	ForkL Loop 1	fork 1 Δ453-465	-	NA	NA	NA	NA
		fork 1 Δ455-464	-	NA	NA	NA	NA
		fork 1 Δ458-459	+	+	-	-	-
		fork 1 W453A G454A D455A	-	NA	NA	NA	NA
		fork 1 R463A G465A	-	NA	NA	NA	NA
	Fork Loop 2	fork 2 Δ487-499	-	NA	NA	NA	NA
		fork 2 Δ493-494	-	NA	NA	NA	NA
		fork 2 G490A R491A D492A	-	NA	NA	NA	NA
		fork 2 K497A P498A R499A	-	NA	NA	NA	NA
	Metal B	MB E791A D792A	-	-	NA	NA	-
		MB E791A	+	+	+	+	+/-
		MB D792A	-	-	NA	NA	-
	Switch 3	sw3 Δ1064-1085	-	NA	NA	NA	NA
		sw 3 R1078A S1079A R1080A	+	-	-	-	-
		sw 3 L1084A R1085A F1086A	-	NA	NA	NA	NA
	Switch4	sw4 Δ1109-1114	-	NA	NA	NA	NA
		sw 4 Δ1111-1112	-	NA	NA	NA	NA
		sw 4 S1111A D1112A P1113A	-	NA	NA	NA	NA

Figure 1

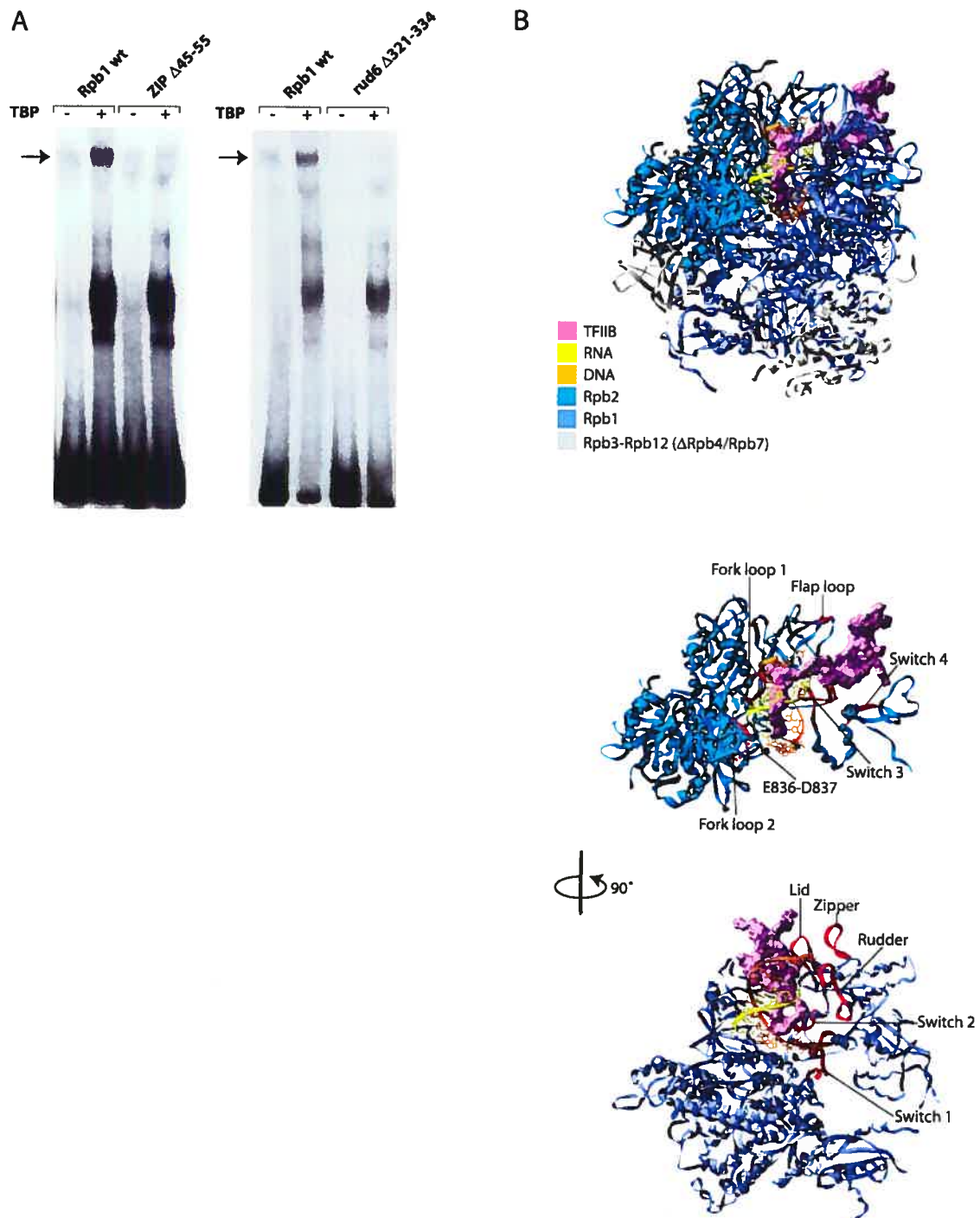
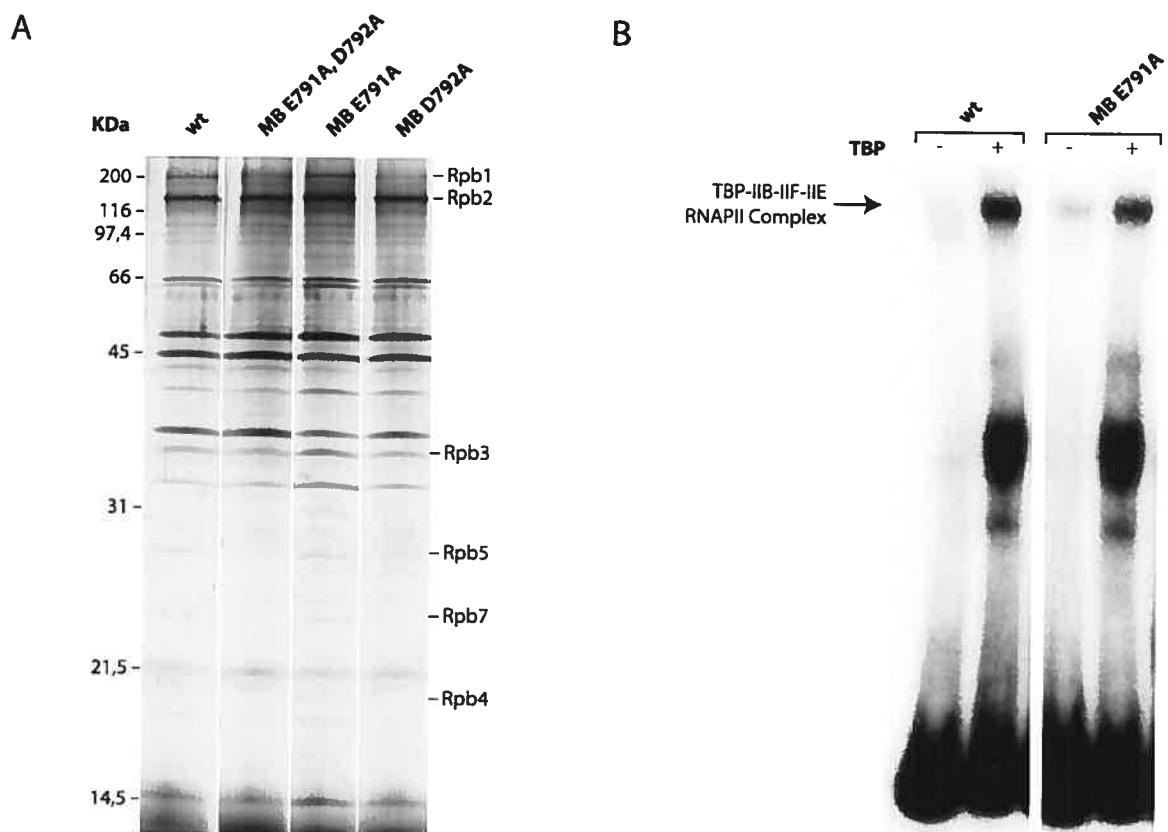


Figure 2

**Figure 3**

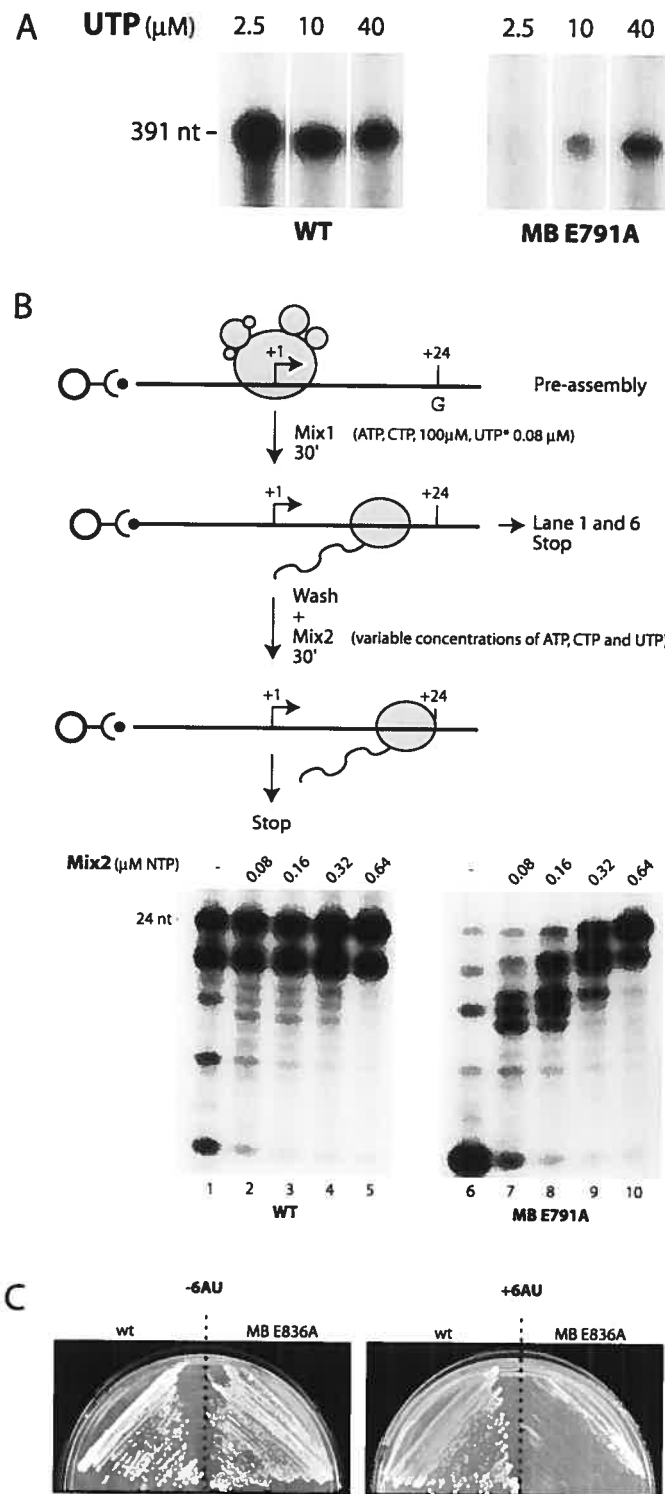


Figure 4

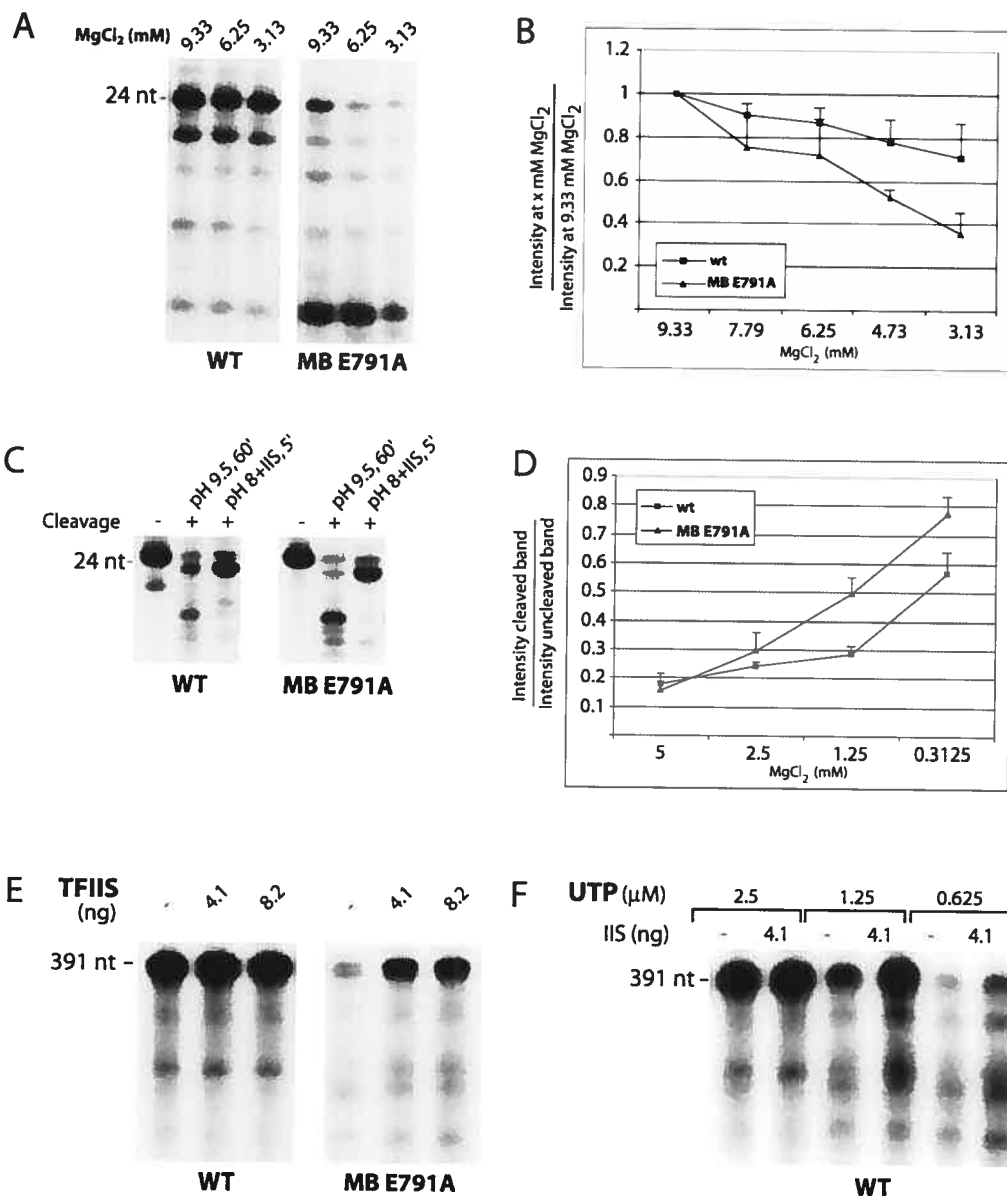
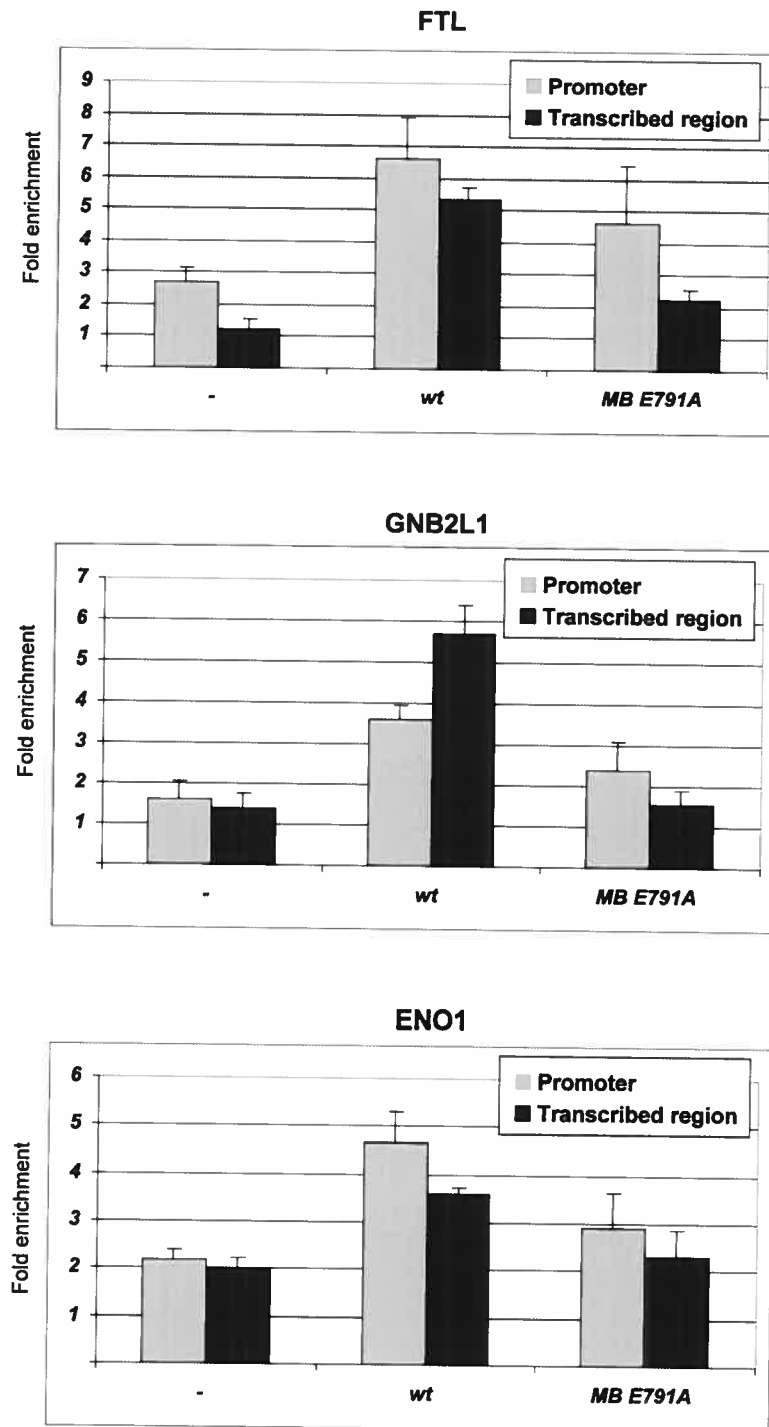


Figure 5

**Figure 6**

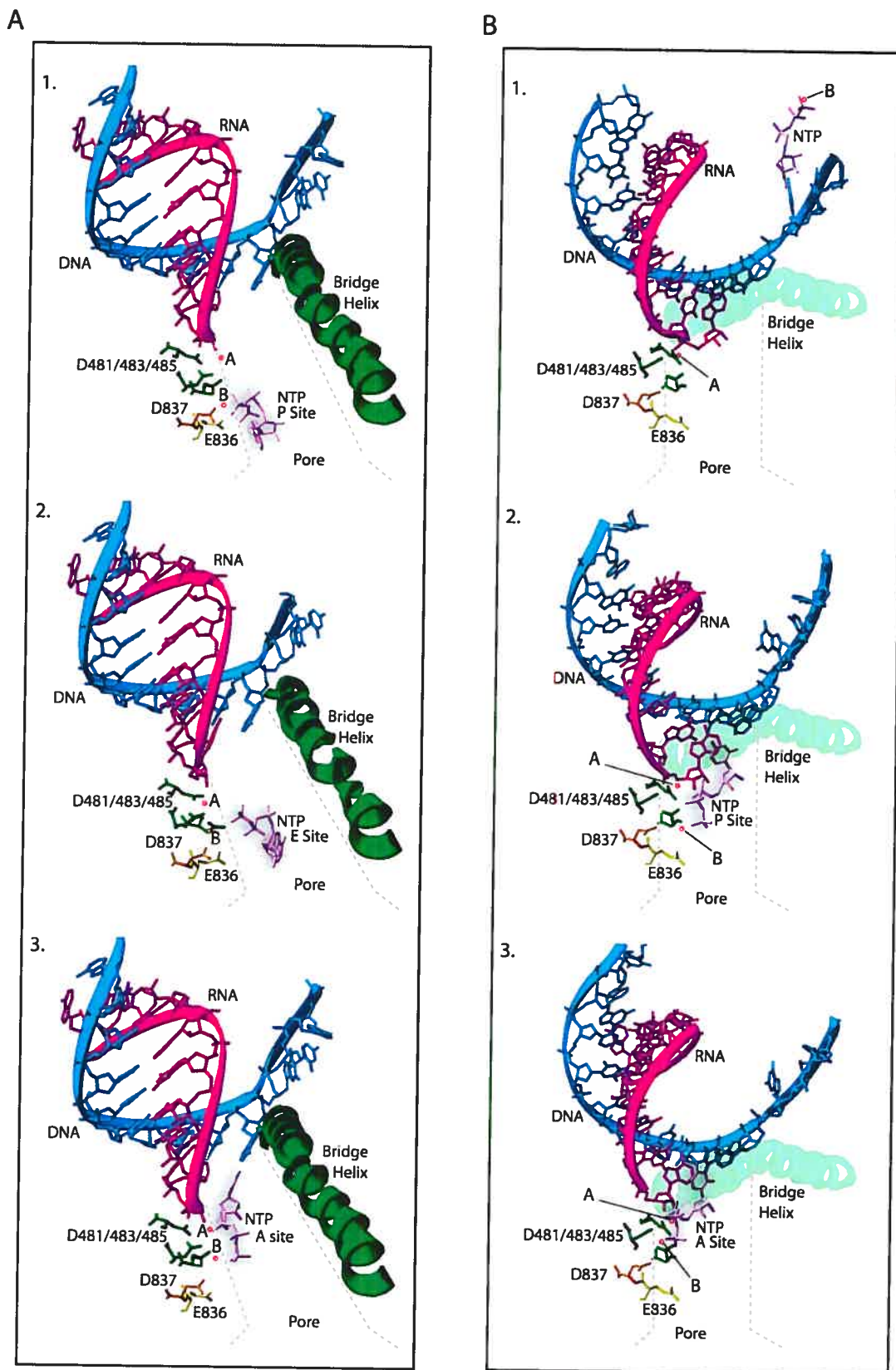


Figure 7

CHAPITRE 6

DISCUSSION

Une discussion des résultats ayant déjà été présentée dans chacun des articles de cette thèse, le chapitre suivant tentera d'aller au-delà de celles-ci. Ainsi, dans cette section, une localisation est proposée pour TFIIA dans le nouveau modèle du complexe prétranscriptionnel présenté par Forget et col. (Forget *et al.*, 2004). Cette localisation permet de proposer des hypothèses pour expliquer comment TFIIA arrive à stimuler l'initiation via TFIIF et TFIIE. Le rôle de ces derniers dans l'étape d'initiation est aussi discuté. Dans un deuxième temps, le choix de la stratégie expérimentale utilisée pour l'analyse structure-fonction de l'ARN pol II est justifié. Ensuite, une réflexion sur pourquoi les études biochimiques fonctionnelles sont essentielles et complémentaires à la résolution de structures cristallographiques pour la compréhension du mécanisme catalytique de la transcription est présentée. Des raisons pouvant expliquer les défauts d'assemblage de plusieurs mutants analysés sont proposées. Finalement, une série d'observations, qui ont été faites suite à la cartographie sur la structure de l'enzyme de tous les mutants d'ARN polymérase à sous-unités multiples étudiés à ce jour, est présentée.

6.1 Localisation de TFIIA dans le complexe prétranscriptionnel et rôles de TFIIF et TFIIE dans l'initiation

L'utilisation de la technique de photo-pontage protéine-ADN a permis au laboratoire du Dr Benoit Coulombe, au cours des dernières années, d'étudier le complexe d'initiation de la transcription de l'ARN pol II. Cette technique a permis non seulement le positionnement des FGT et de la pol II dans le complexe, mais aussi de proposer le

modèle d'enroulement du promoteur autour de l'ARN pol II (Robert *et al.*, 1996; Robert *et al.*, 1998; Forget *et al.*, 1997; Douziech *et al.*, 2000; Coulombe, 1999).

Selon ce modèle, les régions -40 et +26 du promoteur se retrouvent juxtaposées lorsqu'un complexe d'initiation comprenant TBP, TFIIB, TFIIF, TFIIE et l'ARN pol II s'assemble sur l'ADN. Le facteur TFIIA a été localisé dans la région de la boîte TATA et en amont de celle-ci dans des études précédentes de photo-pontage sur un complexe TBP-TFIIA-ADN (Coulombe *et al.*, 1994; Lagrange *et al.*, 1996). Dans l'article présenté au chapitre 2 (Langelier *et al.*, 2001), nous avons déterminé la localisation de TFIIA dans un complexe d'initiation TBP-TFIIA-TFIIB-TFIIF-TFIIE-ARN pol II aux positions -40 et +26 du promoteur. Alors que le photo-pontage à -40 est dépendant de la présence de TBP, celui à +26 est dépendant à la fois de TBP et de RAP30, indiquant que ce dernier est permis uniquement lorsque le promoteur est enroulé autour de la polymérase. TFIIF, TFIIE et l'ARN pol II étant aussi localisés à ces mêmes positions dans le complexe d'initiation, nous avons étudié la possibilité d'une interaction entre TFIIA et ces différents facteurs. Par chromatographie d'affinité, nous avons pu déterminer que TFIIA interagissait avec RAP74, TFIIE α (TFIIE34) et TFIIE β (TFIIE56). Nous avons aussi déterminé deux domaines de RAP74 impliqués dans l'interaction avec TFIIA et avons montré que ce dernier stimulait l'initiation via son association avec RAP74 et TFIIE β .

La principale critique concernant la technique de photo-pontage protéine-ADN classique est qu'elle est faite sur des complexes d'initiation non purifiés, qui pourraient donc inclure un certain nombre de sous-complexes non-fonctionnels. La présence de ceux-ci peut nuire à l'interprétation des résultats de photo-pontage. Afin de contrecarrer à ce problème et d'augmenter la résolution de la méthode, le laboratoire a développé une nouvelle technique, décrite dans l'article présenté au chapitre 3, où le photo-pontage est réalisé sur des complexes d'initiation préalablement purifiés en gel natif (Forget *et al.*, 2004). Quelques différences ont été observées par rapport aux résultats obtenus avec l'ancienne méthode pour certains composants du complexe. Notamment, des photo-pontages ont été obtenus pour la sous-unité Rpb5 et quelques contacts de RAP74 ont été perdus dans la région -15 à -5. À la lumière de la structure cristallographique à haute

résolution de l'ARN pol II en élongation (Gnatt *et al.*, 2001), un nouveau modèle du complexe d'initiation a pu être proposé. Dans ce modèle, l'ADN en aval est localisé sur la pince. Entre les positions +10 et +20, il fait des contacts avec les domaines *upper* et *lower jaws*. La position +1 se retrouve près du site catalytique de l'enzyme, mais le brin d'ADN matrice n'y accéderait que suite à l'ouverture du promoteur, tel que proposé par la cristallographie à 12 sous-unités de la pol II (Armache *et al.*, 2003; Bushnell et Kornberg, 2003). L'ADN en amont s'enroule ensuite autour de la pince en prenant une courbure dans la région de la boîte TATA, en accord avec la cristallographie de TBP sur l'ADN (Kim *et al.*, 1993b; Kim *et al.*, 1993a). Les deux sous-unités de TFIIF et TFIIE ont été positionnées dans ce modèle.

Il est possible d'intégrer les résultats présentés dans l'article Langelier et col. en 2001, au modèle présenté par Forget et col. en 2004. En effet, en plaçant la sous-unité α/β de TFIIA sur la pince dans le modèle (voir figure 1), on peut expliquer les contacts de celui-ci avec les positions -40 et +26. De plus, à cette position, α/β peut interagir avec RAP74, TFIIE β et TBP, tel que montré par les chromatographies d'affinité. Le positionnement de γ , indiqué à la figure 1, permet d'expliquer le photo-pontage à -40 et l'interaction avec TFIIE α . La sous-unité γ se retrouve cependant un peu loin de la position +26. Le photo-pontage de γ à cette position est d'ailleurs beaucoup plus faible que celui de α/β à cette même position.

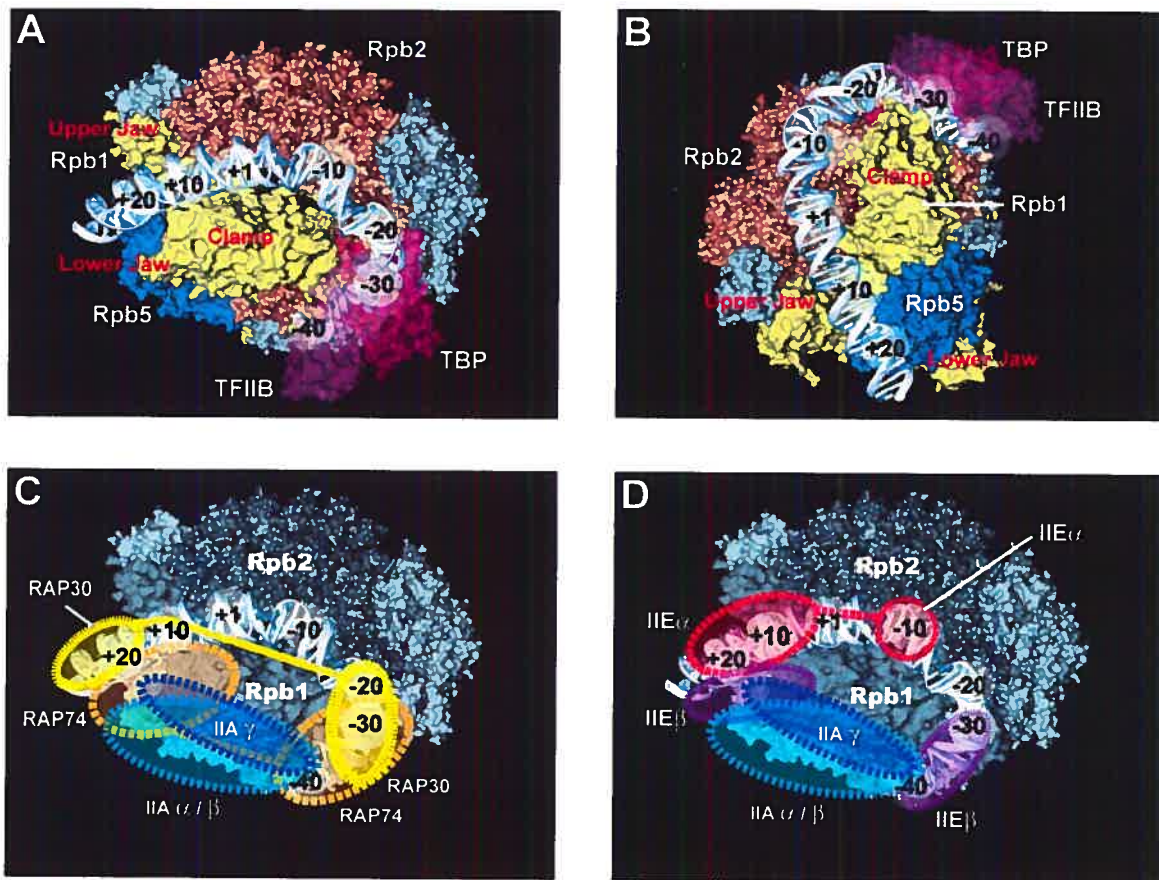


Figure 1 : Localisation de TFIIA dans le complexe de préinitiation de l'ARN polymérase II. Les sous-unités de TFIIA α/β et γ ont été placées dans le modèle du complexe de préinitiation présenté au chapitre III (figure 6). Figure modifiée de Forget *et al.*, 2004.

La localisation de TFIIA dans le modèle présenté à la figure 1 nous permet de proposer des rôles pour celui-ci dans l'initiation de la transcription. On peut penser que TFIIA, en liant TFIIF et TFIIE dans le complexe d'initiation, stabilise l'interaction de ces derniers avec le complexe et favorise leur positionnement correct. En conséquence, TFIIA pourrait avoir un effet stimulateur sur l'initiation de la transcription basale via RAP74 et TFIIE β , tel que montré par Langelier et col., 2001. Différents rôles ont été proposés pour TFIIF et TFIIE dans l'initiation, qui ont été décrits précédemment dans cet ouvrage. Entre autres, TFIIF, et particulièrement la sous-unité RAP74, est nécessaire à l'enroulement du promoteur. Un enroulement gauche autour de la pince de l'ARN pol II peut induire une relaxation dans l'ADN et produire une portion simple-brin dans l'hélice,

reconnue par les hélicases de TFIIH qui ouvrent la bulle de transcription. De cette façon, une stabilisation de l'association de TFIIF au complexe par TFIIA pourrait stimuler l'initiation en favorisant la formation de la bulle transcriptionnelle.

L'article de Forget et col., 2004 montre que TFIIEx joue lui aussi un rôle important dans l'ouverture du promoteur. Les études de photo-pontage ont montré que TFIIEx fait contact avec l'ADN dans la région +5 à +25, mais aussi au niveau de la région -10, qui correspond à la partie amont de la bulle transcriptionnelle. De plus, dans des essais d'initiation, TFIIEx stimule l'initiation sur une matrice préouverte entre -11 et -1. La stimulation est particulièrement grande lorsque la préouverture est localisée dans la partie amont de la bulle. Nous proposons donc que TFIIEx, grâce à un contact avec l'ADN près de la position -10, stabilise le promoteur ouvert et stimule l'initiation. On peut envisager qu'une préouverture artificielle comme celle utilisée dans l'article, mime un intermédiaire de la réaction de formation de la bulle par TFIIH. L'intermédiaire simple-brin serait lié par TFIIEx pour stabiliser l'ouverture. Alternativement, la préouverture artificielle pourrait créer, conjointement avec la relaxation de l'ADN induite par l'enroulement du promoteur, une vaste portion d'ADN simple-brin, stabilisée par TFIIEx. Des études ont d'ailleurs démontré que TFIIEx était impliqué dans l'ouverture du promoteur d'une manière indépendante de TFIIH (Timmers, 1994; Holstege *et al.*, 1995). À la lumière de ceci, TFIIA en contactant TFIIEx pourrait le stabiliser dans le complexe d'initiation. Comme TFIIEx favorise l'ouverture du promoteur, TFIIA peut donc ainsi stimuler l'initiation de la transcription.

6.2 Étude structure-fonction de l'ARN polymérase II : justification de la stratégie expérimentale choisie

Les articles présentés aux chapitres 4 (Jeronimo *et al.*, 2004) et 5 (Langelier *et al.*, 2005) présentent l'étude de mutants de l'ARN pol II humaine. Cette étude a été basée sur les différentes structures cristallographiques de l'ARN pol II à haute résolution. Ces structures ont permis de proposer une série d'hypothèses sur les détails du mécanisme

catalytique de l'enzyme, ainsi que sur les rôles de certaines boucles, hélices et motifs dans le processus transcriptionnel. Afin de tester ces hypothèses, nous avons entrepris une vaste étude structure-fonction de l'ARN pol II humaine, en procédant à la mutagenèse de plusieurs domaines de Rpb1 et Rpb2 et à la caractérisation biochimique des mutants obtenus.

Pour ce faire, nous avons choisi d'utiliser une méthode qui permet d'induire l'expression, à des niveaux physiologiques, de la sous-unité Rpb1 ou Rpb2 sauvage ou mutante dans des cellules humaines. La sous-unité porte à son extrémité C-terminal une étiquette TAP (*Tandem affinity purification*), qui permet la purification à l'aide de la méthode portant le même nom (Rigaut *et al.*, 1999). Différentes considérations justifient le choix de cette stratégie. Premièrement, la technique choisie devait conduire à la purification de complexes fonctionnels. Avec la méthode TAP, qui est entièrement réalisée à faible concentration ionique, nous avons pu obtenir de tels complexes, pouvant ensuite être testés dans différents essais biochimiques. Deuxièmement, la stratégie utilisée devait minimiser la perte des partenaires d'interactions. Cette dernière considération était importante, puisqu'une étude protéomique était aussi en cours sur l'ARN pol II (travail réalisé par Célia Jeronimo). Le choix d'un système d'expression inductible et facilement contrôlable a aussi été fait dans ce sens, afin de réduire la présence d'interactions non-physiologiques. Finalement, comme les complexes purifiés étaient destinés à être testés dans des essais fonctionnels nécessitant une grande quantité de protéine, l'efficacité de purification était aussi importante. La méthode TAP a été comparée à plusieurs autres méthodes utilisant des étiquettes (FLAG, Strep, His, CBD) et a donné la meilleure efficacité pour la purification de protéines présentes à faible concentration dans un extrait (Puig *et al.*, 2001).

6.3 Structure et mécanisme catalytique de l'ARN polymérase II : au-delà de la cristallographie

Suite à la caractérisation biochimique des mutants de l'ARN pol II obtenus nous avons pu proposer des rôles pour les différents éléments structuraux étudiés. Il est

intéressant de comparer les conclusions obtenues aux fonctions qui avaient été attribuées initialement à partir des cristallographies (tableau I). Alors que des rôles dans le maintien de la bulle de transcription et la séparation de l'hybride ADN-ARN avaient été suggérés pour plusieurs domaines, la plupart des mutants étudiés sont bloqués avant l'initiation. Plusieurs d'entre eux sont mêmes incapables de former le complexe prétranscriptionnel au promoteur. Il est possible que ces boucles possèdent effectivement les rôles mentionnés ci-haut, une fois l'initiation commencée. En effet, le positionnement de plusieurs boucles a été déterminé dans un complexe en élongation (Gnatt *et al.*, 2001; Westover *et al.*, 2004a), puisqu'au niveau de l'enzyme seule celles-ci étaient désordonnées (Cramer *et al.*, 2001). Notre étude montre cependant que ces boucles jouent un rôle bien avant les étapes proposées, qui n'avait pas été mis en évidence dans les structures cristallographiques. La cristallographie des complexes d'initiation ouverts et fermés permettra de suggérer de nouveaux rôles pour ces éléments.

Tableau I. Comparaison entre les rôles prédits par les structures cristallographiques et les rôles déterminés par les essais biochimiques pour les éléments structuraux du site catalytique de l'ARN polymérase II

Sous-unité	Domaine	Rôles prédits	Rôles déterminés
Rpb1	<i>Zipper</i>	Séparation de l'hybride ADN-ARN, sortie de l'ARNm et maintien de la bulle (amont) ^{1,2}	Formation du complexe prétranscriptionnel au promoteur, liaison à l'ADN
Rpb1	<i>Rudder</i>	Séparation de l'hybride ADN-ARN ³	Formation du complexe prétranscriptionnel au promoteur, liaison à l'ADN
Rpb2	Métal B	Coordination du Mg ²⁺ -B ^{1,2} pour la catalyse (polymérisation)	Coordination du Mg ²⁺ -B nécessaire au chargement du NTP au site catalytique pour la polymérisation. Coordination du Mg ²⁺ -B pour le clivage
Rpb2	<i>Fork loop 1</i>	Séparation de l'hybride	Rôle après l'ouverture

		ADN-ARN ³	du promoteur et avant l'initiation
Rpb2	Switch 3	Liaison du brin matrice d'ADN au niveau de l'hybride ADN-ARN ⁴	Formation du complexe prétranscriptionnel au promoteur, liaison à l'ADN

1. (Cramer *et al.*, 2001), 2. (Gnatt *et al.*, 2001), 3. (Westover *et al.*, 2004a), 4. (Armache *et al.*, 2003)

Le cas de l'identification des acides aminés impliqués dans la coordination du métal B pour la catalyse est similaire. Les premières structures cristallographiques de l'ARN pol II suggéraient que les résidus E836 et D837 de Rpb2, D481 de Rpb1 et le NTP coordonnaient le Mg²⁺-B (Cramer *et al.*, 2001; Gnatt *et al.*, 2001). Une cristallographie plus récente a cependant montré la localisation de l'ion métallique B à une position différente de celle prédite. Dans cette cristallographie, les résidus E836 et D837 n'occupent plus la même position, si bien que le métal B ne peut plus être coordonné par E836 pour la réaction de polymérisation (Westover *et al.*, 2004b). Nos résultats biochimiques nous ont cependant montré que le E836 est important pour la liaison du Mg²⁺ pour la réaction de polymérisation. Le modèle le plus récent n'impliquant pas le résidu de Rpb2 E836 dans la coordination à aucune des étapes du mécanisme catalytique (Westover *et al.*, 2004b), nous avons pu raffiner ce modèle en proposant une interaction transitoire entre E836 et le métal B, essentielle au chargement du NTP au site catalytique.

Au niveau du mécanisme de translocation des acides nucléiques, des études biochimiques ne semblent pas confirmer complètement les modèles proposés par les cristallographes. En effet, les études de cinétique suggèrent l'existence d'un préappariement du NTP arrivant avec la matrice d'ADN à i+2, comme mode de sélection du NTP avant son entrée au site catalytique (Nedialkov *et al.*, 2003). Dans ce modèle le NTP entre par le canal principal. Au contraire, dans le modèle suggéré par les cristallographes, le NTP entre par le pore et se positionne à un premier site à partir duquel il subit une rotation pour aller se placer au site d'addition pour la catalyse (Westover *et al.*, 2004b). Tel que discuté précédemment (voir section 1.3.3 de

l'introduction), la cristallographie d'autres états du processus transcriptionnel pourra peut-être nous aider à discriminer entre ces deux possibilités.

On peut donc conclure que, bien que la résolution de structures cristallographiques soit essentielle et nous permette de proposer un grand nombre d'hypothèses sur le fonctionnement de l'enzyme ARN polymérase, les études biochimiques sont nécessaires afin de confirmer ou d'infirmer ces modèles.

6.4 Les mutants présentant des défauts d'assemblage

Au cours de l'analyse structure-fonction de l'ARN pol II présentée aux chapitres IV et V, nous avons créé de nombreux mutants de délétions et de substitutions, qui ont par la suite montré d'importants défauts d'assemblage (voir figure 1, chapitre VI). Alors que de grandes délétions et substitutions ont été faites dans plusieurs éléments structuraux de Rpb1 sans altérer l'assemblage, plusieurs domaines de Rpb2 n'ont pu être étudiés à cause d'une incapacité des mutants à former le complexe à 12 sous-unités. Dans le cas de certains domaines, même de courtes délétions de 2 acides aminés et des substitutions de 3 résidus, ont conduit à la purification d'enzymes incomplètes (boucles *switch 4* et *fork loop 2*).

On peut retourner à la structure de l'ARN pol II pour tenter d'expliquer ces défauts d'assemblage. Il est important de noter que plusieurs des éléments structuraux étudiés étaient incomplets dans les premières cristallographies à haute résolution (Cramer *et al.*, 2001; Gnatt *et al.*, 2001) à partir desquelles a été planifiée la mutagenèse. Par exemple, dans le cas des boucles *lid*, *rudder* et *fork loop 1*, seuls le début et la fin des boucles étaient disponibles dans les premières structures, alors que les boucles complètes ont pu être localisées dans une cristallographie plus récente (Westover *et al.*, 2004a).

L'examen de la localisation des éléments *switch 3* et *4* permet de voir que ceux-ci sont localisés dans une région de Rpb2 située à la base de la pince. Ces éléments sont donc impliqués dans de nombreuses interactions avec Rpb1 et avec d'autres domaines de

Rpb2. Il est donc facile de croire que de petites délétions et des substitutions dans ces *switches* peuvent avoir un effet très important sur l'organisation de la structure de Rpb2 et sur l'interaction avec les autres sous-unités, en particulier Rpb1. Au contraire, des boucles comme le *zipper* et le *rudder* sont des boucles qui sont dirigées dans le canal principal, complètement libres d'interactions avec d'autres domaines de Rpb1 ou Rpb2. Conséquemment, ces boucles ont pu être délétées complètement, les mutants en résultant ne présentant aucun problème d'assemblage.

La localisation de *fork loop 1* dans la structure de l'ARN pol II montre que son résidu hydrophobe W466 (W453 chez l'humain) semble important pour le maintien du repliement correct de Rpb2. En effet, l'acide aminé W466 pointe vers l'extérieur de la boucle et vers un cœur hydrophobe formé par deux hélices de Rpb2, l'une localisée aux positions 200 à 205 de la séquence et l'autre aux positions 410 à 420. Le rapprochement de ces deux domaines éloignés pour former un centre hydrophobe peut être affecté par la substitution du W466 ou par son déplacement, dans le cas de mutants de délétion. Ainsi, seul le mutant de *fork loop 1* ayant une délétion de 2 acides aminés ($\Delta 458-459$, chez l'humain) a conduit à la purification d'une enzyme complète. Les deux mutants présentant des délétions plus grandes et deux autres mutants ayant des substitutions n'ont pu s'assembler correctement.

Dans le cas d'une simple substitution comme celui du mutant D792A, potentiellement impliqué dans la coordination de l'ion métal B au site catalytique, il peut être difficile à première vue de comprendre le défaut d'assemblage. Dans la structure de l'ARN pol II de levure, on peut cependant observer que l'équivalent du D792, le D837 est impliqué dans un pont salin avec le résidu R1020, faisant partie d'un autre domaine. L'hypothèse que ce lien est important pour un repliement adéquat de Rpb2, lui permettant d'interagir avec les autres sous-unités, peut être faite. Le D837 se retrouve aussi près de plusieurs résidus de Rpb1 au site actif, sa substitution peut donc potentiellement entraîner des modifications au niveau de l'interaction entre les deux sous-unités catalytiques.

6.5 Un répertoire des mutations affectant la fonction des ARN polymérase à sous-unités multiples, cartographiées sur la structure de l'enzyme

L'étude présentée dans cette thèse est une analyse fonctionnelle de domaines particuliers de l'enzyme ARN pol II, qui avaient été mis en évidence par les différentes structures cristallographiques. Au cours des 15 dernières années plusieurs études de mutagenèse ont été faites autant chez l'ARN pol bactérienne que chez les ARN pol I, II et III eucaryotes. La plupart de ces études ont été réalisées avant la publication des structures à haute résolution et donc, les conclusions tirées ne tenaient pas compte de la localisation de ces mutations dans la structure de l'enzyme. Il peut être intéressant de répertorier l'ensemble de ces mutations et de les positionner dans les structures de l'ARN pol procaryote et eucaryote. Ce travail a été réalisé dans le laboratoire du Dr Benoit Coulombe par Vincent Trinh et a permis de faire une série d'observations très intéressantes (Trinh *et al.*, 2005).

Par exemple, presque tous les mutants de l'ARN pol II qui présentent une sensibilité au 6AU pouvant être supprimée par une surexpression de TFIIS sont localisés sur le trajet de ce dernier sur la polymérase (Powell et Reines, 1996; Archambault *et al.*, 1992c; Archambault *et al.*, 1992a), tel que déterminé par la co-cristallographie (Kettenberger *et al.*, 2003). On peut penser que ces mutants sont déficients dans leur capacité d'interagir avec le facteur de clivage et d'élongation. Au contraire, les mutants dont la sensibilité au 6AU n'est pas contrecarrée par TFIIS sont dispersés ailleurs sur la structure (Lennon, III *et al.*, 1998; Chen et Hampsey, 2004; Archambault *et al.*, 1992c; Powell et Reines, 1996; Dieci *et al.*, 1995). De plus, cette étude a permis de définir un domaine de Rpb2 situé dans la voûte du canal principal à côté du lobe et près de l'hybride ADN-ARN, impliqué dans l'élongation de l'ARNm. En effet, plusieurs mutations qui affectent l'élongation, tant chez les procaryotes (Sosunov *et al.*, 2003; Lee *et al.*, 1991; Zaychikov *et al.*, 1996; Bartlett *et al.*, 1998; Heisler *et al.*, 1993; Sagitov *et al.*, 1993) que chez les eucaryotes (Chen et Hampsey, 2004; Archambault *et al.*, 1992c; Scafe *et al.*, 1990a; Berroteran *et al.*, 1994; Chen *et al.*, 1996; Dieci *et al.*, 1995) (Chen *et al.*, 1993;

Thuillier *et al.*, 1996; Archambault *et al.*, 1990; Drebot *et al.*, 1993; Archambault *et al.*, 1992a; Scafe *et al.*, 1990b), se retrouvent au niveau de ce domaine. Des supprimeurs de mutants d'élongation (Rozenfeld et Thuriaux, 2001; Archambault *et al.*, 1998a) et des mutants présentant à la fois une efficacité d'élongation et de terminaison altérée (Jin et Gross, 1991; Bobkova *et al.*, 1999; Rozenfeld et Thuriaux, 2001; Weilbaecher *et al.*, 1994; Tavormina *et al.*, 1996a; Severinov *et al.*, 1995; Tavormina *et al.*, 1996b) y sont aussi localisés.

Chez les eucaryotes, un groupe de mutants défectueux dans le choix du site d'initiation de la transcription se retrouve dans le domaine lobe de Rpb2 (Hekmatpanah et Young, 1991; Chen et Hampsey, 2004). Cette région, qui fait contact avec l'ADN dans la partie aval, serait donc possiblement impliquée dans le positionnement correct des acides nucléiques pour l'initiation. Il est intéressant de constater que ces mutations se retrouvent au voisinage du site d'interaction de Rpb9 avec Rpb2, puisque des études ont montré que la sous-unité Rpb9 est impliquée dans le choix du site +1 (Hull *et al.*, 1995; Sun *et al.*, 1996). Une boucle de Rpb1, le *trigger loop*, situé près du *bridge helix* entre les acides aminés 1080 et 1096 (levure) serait aussi importante pour ce choix, puisque deux mutants ont été localisés sur cet élément structural, l'un au début et l'autre à la fin de la boucle (Hekmatpanah et Young, 1991).

Le travail de cartographie a aussi permis de déterminer que la lysine 1065 de la sous-unité β de l'ARN pol d'*E.coli*, substituée chez plusieurs mutants ayant des défauts dans le dégagement du promoteur (Mustaev *et al.*, 1991; Kashlev *et al.*, 1990; Sagitov *et al.*, 1993), se retrouve à proximité de l'extrémité 3' de l'ARNm. Ainsi, ce résidu pourrait interagir avec l'ARN et contribuer à sa stabilisation dans le site catalytique, favorisant le passage en élongation productive. Au contraire, une mutation de la lysine 1065 pourrait conduire à une déstabilisation de l'ARNm naissant entraînant son relâchement et le maintien de l'ARN pol dans le cycle d'initiations abortives. Il serait intéressant de voir si cette lysine qui est strictement conservée joue le même rôle en transcription chez les eucaryotes.

CONCLUSIONS ET PERSPECTIVES

Le premier objectif de cette thèse était d'étudier le rôle de certains des FGT à l'étape d'initiation en relation avec leur position dans l'organisation du complexe prétranscriptionnel. Nous avons atteint cet objectif en montrant que TFIIA occupe une position dans le complexe d'initiation qui lui permet d'interagir avec TFIIF et TFIIE. Ces interactions, entre TFIIA et les sous-unités RAP74 de TFIIF et β de TFIIE, entraînent une stimulation de l'initiation lorsque TFIIA est ajouté à la réaction transcriptionnelle. La stimulation par TFIIA passe possiblement par une stabilisation de TFIIF et TFIIE dans le complexe, leur permettant de mieux jouer leurs rôles en initiation.

L'implication de TFIIE α dans l'étape d'ouverture du promoteur a aussi été démontrée dans cette thèse à l'aide de la technique de photo-pontage protéine-ADN et d'essais d'initiation. La première a permis de localiser TFIIE α à la position -10 du promoteur, juste en amont de la région où la bulle de transcription se forme. À l'aide d'essais d'initiation, nous avons montré que TFIIE α stimule l'initiation de façon plus importante sur des matrices préouvertes de 2 nucléotides entre -11 et -1. La stimulation était particulièrement grande avec des préouvertures dans la partie amont de la bulle de transcription. Nos résultats suggèrent que TFIIE α favorise l'ouverture du promoteur en contactant l'ADN dans la région -10. Il est donc possible que TFIIA stimule l'initiation en stabilisant, entre autres, TFIIE dans le complexe transcriptionnel, facilitant ainsi son rôle dans l'ouverture du promoteur.

Les expériences de photo-pontage protéine-ADN présentées au chapitre III (Forget *et al.*, 2004) ont été réalisées à l'aide de la technique dite « en gel », sur des complexes purifiés. Combinées à la structure à haute résolution de l'ARN pol II de levure en élongation, ces expériences ont permis de proposer un modèle pour l'organisation du complexe prétranscriptionnel. Par ailleurs, la localisation de TFIIA dans le complexe d'initiation au chapitre 2 (Langelier *et al.*, 2001) a été faite avec la méthode de photo-pontage classique, c'est-à-dire en solution, sur des complexes non-purifiés. Il serait donc

intéressant de reprendre les expériences de localisation de TFIIA dans le complexe de préinitiation, avec la technique de photo-pontage en gel. Ceci nous permettrait de voir si TFIIA est photo-ponté aux positions précédemment identifiées et à de nouvelles positions le long du promoteur. Ainsi on pourrait confirmer ou infirmer le positionnement proposé à la figure 1 de la discussion (chapitre VI) près de TFIIF et TFIIE, qui supporte un rôle pour TFIIA dans la stabilisation de ces derniers dans le complexe.

Il serait aussi intéressant d'utiliser la méthode de photo-pontage en gel et d'étudier l'effet de l'ajout de TFIIH sur l'organisation du complexe prétranscriptionnel. Les expériences pourraient être faites en absence et en présence d'ATP, qui dans le dernier cas permettrait l'étude d'un complexe où la bulle de transcription est ouverte. Ainsi il serait possible de voir comment la localisation de TFIIE α change dans ces différentes conditions, ce qui nous permettrait de mieux comprendre son rôle dans l'ouverture du promoteur. Éventuellement, des complexes d'initiation et d'élongation pourraient être étudiés. L'étude du trajet de l'ADN sur l'ARN pol II dans ces différents contextes pourrait nous permettre de confirmer les modèles basés sur les structures cristallographiques de la polymérase et de voir comment l'ensemble des FGT sont organisés dans ces complexes.

Le deuxième objectif de cette thèse était d'étudier les rôles des divers éléments structuraux de Rpb1 et Rpb2, localisés dans la région du site catalytique de l'ARN pol II, qui avaient été mis en évidence par les cristallographies à haute résolution de l'enzyme. Grâce à une stratégie de mutagenèse dirigée combinée à l'étude fonctionnelle des mutants obtenus, nous avons pu définir des rôles pour plusieurs de ces éléments, dont les boucles *switch 3*, *rudder*, *zipper*, *fork loop 1* et le résidu E791 de Rpb2. Plusieurs mutants étudiés présentaient des défauts dans la formation du complexe prétranscriptionnel au promoteur. Éventuellement, il serait très intéressant de vérifier si ces mutants sont déficients dans une interaction avec un des FGT ou s'ils sont tout simplement incapables de lier l'ADN.

Plusieurs éléments structuraux ciblés dans cette étude n'ont pu être étudiés à cause de leur incapacité à supporter l'assemblage des 12 sous-unités de l'ARN pol II. Dans la

poursuite de ce projet, il serait pertinent, en se basant sur les plus récentes structures cristallographiques, de planifier de nouveaux mutants présentant des substitutions de quelques résidus seulement, qui potentiellement auraient moins d'effet sur l'intégrité de la polymérase. De plus, il serait intéressant de continuer l'analyse de plusieurs autres éléments structuraux de Rpb1 et Rpb2, qui n'ont pas été étudiés dû à des limitations de temps. Par exemple, la caractérisation d'une série de mutants de substitutions du *bridge helix* à l'aide d'expériences de cinétiques d'élongation pourra nous permettre d'analyser le rôle de celui-ci dans la translocation des acides nucléiques. Il serait aussi pertinent de tester des mutants de la boucle *flap* dans divers essais fonctionnels. Chez les procaryotes, la boucle flap est impliquée, entre autres, dans la reconnaissance des structures d'ARN *hairpins* qui induisent des pauses. Chez les eucaryotes, l'induction de pauses dans la transcription par des structures d'ARN est un phénomène beaucoup moins étudié, qui a pourtant lieu lors de la transcription au promoteur HIV-1. Il serait donc très intéressant de tester des mutants du domaine *flap* pour la reconnaissance de la pause à ce promoteur et ainsi de voir si ce mécanisme de régulation de l'élongation est conservé entre les procaryotes et les eucaryotes.

BIBLIOGRAPHIE

- Agalioti,T., Chen,G., et Thanos,D. (2002). Deciphering the transcriptional histone acetylation code for a human gene. *Cell 111*, 381-392.
- Agalioti,T., Lomvardas,S., Parekh,B., Yie,J., Maniatis,T., et Thanos,D. (2000). Ordered recruitment of chromatin modifying and general transcription factors to the IFN-beta promoter. *Cell 103*, 667-678.
- Ahn,S.H., Kim,M., et Buratowski,S. (2004). Phosphorylation of serine 2 within the RNA polymerase II C-terminal domain couples transcription and 3' end processing. *Mol. Cell 13*, 67-76.
- Akoulitchev,S., Makela,T.P., Weinberg,R.A., et Reinberg,D. (1995). Requirement for TFIID kinase activity in transcription by RNA polymerase II. *Nature 377*, 557-560.
- Aoyagi,S. et Hayes,J.J. (2002). hSWI/SNF-Catalyzed Nucleosome Sliding Does Not Occur Solely via a Twist-Diffusion Mechanism. *Mol. Cell Biol. 22*, 7484-7490.
- Archambault,J., Chambers,R.S., Kobor,M.S., Ho,Y., Cartier,M., Bolotin,D., Andrews,B., Kane,C.M., et Greenblatt,J. (1997). An essential component of a C-terminal domain phosphatase that interacts with transcription factor IIF in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. U. S. A 94*, 14300-14305.
- Archambault,J., Drebot,M.A., Stone,J.C., et Friesen,J.D. (1992a). Isolation and phenotypic analysis of conditional-lethal, linker-insertion mutations in the gene encoding the largest subunit of RNA polymerase II in *Saccharomyces cerevisiae*. *Mol. Gen. Genet. 232*, 408-414.
- Archambault,J., Jansma,D.B., Kawasoe,J.H., Arndt,K.T., Greenblatt,J., et Friesen,J.D. (1998a). Stimulation of transcription by mutations affecting conserved regions of RNA polymerase II. *J. Bacteriol. 180*, 2590-2598.
- Archambault,J., Lacroute,F., Ruet,A., et Friesen,J.D. (1992b). Genetic interaction between transcription elongation factor TFIIS and RNA polymerase II. *Mol Cell Biol 12*, 4142-4152.
- Archambault,J., Lacroute,F., Ruet,A., et Friesen,J.D. (1992c). Genetic interaction between transcription elongation factor TFIIS and RNA polymerase II. *Mol. Cell Biol. 12*, 4142-4152.
- Archambault,J., Pan,G., Dahmus,G.K., Cartier,M., Marshall,N., Zhang,S., Dahmus,M.E., et Greenblatt,J. (1998b). FCP1, the RAP74-interacting subunit of a human protein phosphatase that dephosphorylates the carboxyl-terminal domain of RNA polymerase II. *J. Biol. Chem. 273*, 27593-27601.

Archambault, J., Schappert, K.T., et Friesen, J.D. (1990). A suppressor of an RNA polymerase II mutation of *Saccharomyces cerevisiae* encodes a subunit common to RNA polymerases I, II, and III. *Mol Cell Biol* *10*, 6123-6131.

Armache, K.J., Kettenberger, H., et Cramer, P. (2003). Architecture of initiation-competent 12-subunit RNA polymerase II. *Proc. Natl. Acad. Sci. U. S. A* *100*, 6964-6968.

Armache, K.J., Mitterweger, S., Meinhart, A., et Cramer, P. (2004). Structures of complete RNA polymerase II and its subcomplex Rpb4/7. *J. Biol Chem.*

Auble, D.T., Hansen, K.E., Mueller, C.G., Lane, W.S., Thorner, J., et Hahn, S. (1994). Mot1, a global repressor of RNA polymerase II transcription, inhibits TBP binding to DNA by an ATP-dependent mechanism. *Genes Dev.* *8*, 1920-1934.

Awrey, D.E., Weilbaecher, R.G., Hemming, S.A., Orlicky, S.M., Kane, C.M., et Edwards, A.M. (1997). Transcription elongation through DNA arrest sites. A multistep process involving both RNA polymerase II subunit RPB9 and TFIIS. *J. Biol. Chem.* *272*, 14747-14754.

Bartlett, M.S., Gaal, T., Ross, W., et Gourse, R.L. (1998). RNA polymerase mutants that destabilize RNA polymerase-promoter complexes alter NTP-sensing by *rrm P1* promoters. *J. Mol Biol* *279*, 331-345.

Basehoar, A.D., Zanton, S.J., et Pugh, B.F. (2004). Identification and distinct regulation of yeast TATA box-containing genes. *Cell* *116*, 699-709.

Bell, S.D. et Jackson, S.P. (2000). The role of transcription factor B in transcription initiation and promoter clearance in the archaeon *Sulfolobus acidocaldarius*. *J. Biol. Chem.* *275*, 12934-12940.

Belotserkovskaya, R., Oh, S., Bondarenko, V.A., Orphanides, G., Studitsky, V.M., et Reinberg, D. (2003). FACT facilitates transcription-dependent nucleosome alteration. *Science* *301*, 1090-1093.

Belotserkovskaya, R. et Reinberg, D. (2004). Facts about FACT and transcript elongation through chromatin. *Curr. Opin. Genet. Dev.* *14*, 139-146.

Bengal, E., Flores, O., Krauskopf, A., Reinberg, D., et Aloni, Y. (1991). Role of the mammalian transcription factors IIF, IIS, and IIX during elongation by RNA polymerase II. *Mol. Cell Biol.* *11*, 1195-1206.

Berroteran, R.W., Ware, D.E., et Hampsey, M. (1994). The *sua8* suppressors of *Saccharomyces cerevisiae* encode replacements of conserved residues within the largest subunit of RNA polymerase II and affect transcription start site selection similarly to *sua7* (TFIIB) mutations. *Mol. Cell Biol.* *14*, 226-237.

Bhoite, L.T., Yu, Y., et Stillman, D.J. (2001). The Swi5 activator recruits the Mediator complex to the HO promoter without RNA polymerase II. *Genes Dev.* *15*, 2457-2469.

Blackwood,E.M. et Kadonaga,J.T. (1998). Going the distance: a current view of enhancer action. *Science* 281, 61-63.

Blandin,G., Durrens,P., Tekaiia,F., Aigle,M., Bolotin-Fukuhara,M., Bon,E., Casaregola,S., de Montigny,J., Gaillardin,C., Lepingle,A., Llorente,B., Malpertuy,A., Neuveglise,C., Ozier-Kalogeropoulos,O., Perrin,A., Potier,S., Souciet,J., Talla,E., Toffano-Nioche,C., Wesolowski-Louvel,M., Marck,C., et Dujon,B. (2000). Genomic exploration of the hemiascomycetous yeasts: 4. The genome of *Saccharomyces cerevisiae* revisited. *FEBS Lett.* 487, 31-36.

Bobkova,E.V., Habib,N., Alexander,G., et Hall,B.D. (1999). Mutational analysis of the hydrolytic activity of yeast RNA polymerase III. *J. Biol. Chem.* 274, 21342-21348.

Borukhov,S., Polyakov,A., Nikiforov,V., et Goldfarb,A. (1992). GreA protein: a transcription elongation factor from *Escherichia coli*. *Proc. Natl. Acad. Sci. U. S. A* 89, 8899-8902.

Borukhov,S., Sagitov,V., et Goldfarb,A. (1993). Transcript cleavage factors from *E. coli*. *Cell* 72, 459-466.

Bourbon,H.M., Aguilera,A., Ansari,A.Z., Asturias,F.J., Berk,A.J., Bjorklund,S., Blackwell,T.K., Borggreffe,T., Carey,M., Carlson,M., Conaway,J.W., Conaway,R.C., Emmons,S.W., Fondell,J.D., Freedman,L.P., Fukasawa,T., Gustafsson,C.M., Han,M., He,X., Herman,P.K., Hinnebusch,A.G., Holmberg,S., Holstege,F.C., Jaehning,J.A., Kim,Y.J., Kuras,L., Leutz,A., Lis,J.T., Meisterernest,M., Naar,A.M., Nasmyth,K., Parvin,J.D., Ptashne,M., Reinberg,D., Ronne,H., Sadowski,I., Sakurai,H., Sipiczki,M., Sternberg,P.W., Stillman,D.J., Strich,R., Struhl,K., Svejstrup,J.Q., Tuck,S., Winston,F., Roeder,R.G., et Kornberg,R.D. (2004). A unified nomenclature for protein subunits of mediator complexes linking transcriptional regulators to RNA polymerase II. *Mol Cell* 14, 553-557.

Bourgeois,C.F., Kim,Y.K., Churcher,M.J., West,M.J., et Karn,J. (2002). Spt5 cooperates with human immunodeficiency virus type 1 tat by preventing premature RNA release at terminator sequences. *Mol. Cell Biol.* 22, 1079-1093.

Bradsher,J.N., Jackson,K.W., Conaway,R.C., et Conaway,J.W. (1993a). RNA polymerase II transcription factor SIII. I. Identification, purification, and properties. *J. Biol. Chem.* 268, 25587-25593.

Bradsher,J.N., Tan,S., McLaury,H.J., Conaway,J.W., et Conaway,R.C. (1993b). RNA polymerase II transcription factor SIII. II. Functional properties and role in RNA chain elongation. *J. Biol. Chem.* 268, 25594-25603.

Briggs,S.D., Bryk,M., Strahl,B.D., Cheung,W.L., Davie,J.K., Dent,S.Y., Winston,F., et Allis,C.D. (2001). Histone H3 lysine 4 methylation is mediated by Set1 and required for cell growth and rDNA silencing in *Saccharomyces cerevisiae*. *Genes Dev.* 15, 3286-3295.

- Bryant,G.O. et Ptashne,M. (2003). Independent recruitment in vivo by gal4 of two complexes required for transcription. *Mol. Cell* 11, 1301-1309.
- Buratowski,S., Hahn,S., Guarente,L., et Sharp,P.A. (1989). Five intermediate complexes in transcription initiation by RNA polymerase II. *Cell* 56, 549-561.
- Burke,T.W. et Kadonaga,J.T. (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.
- Burton,Z.F., Killeen,M., Sopta,M., Ortolan,L.G., et Greenblatt,J. (1988). RAP30/74: a general initiation factor that binds to RNA polymerase II. *Mol. Cell Biol.* 8, 1602-1613.
- Bushnell,D.A., Cramer,P., et Kornberg,R.D. (2002). Structural basis of transcription: alpha-amanitin-RNA polymerase II cocrystal at 2.8 Å resolution. *Proc. Natl. Acad. Sci. U. S. A* 99, 1218-1222.
- Bushnell,D.A. et Kornberg,R.D. (2003). Complete, 12-subunit RNA polymerase II at 4.1-Å resolution: Implications for the initiation of transcription. *Proc. Natl. Acad. Sci. U. S. A* 100, 6969-6973.
- Bushnell,D.A., Westover,K.D., Davis,R.E., et Kornberg,R.D. (2004). Structural basis of transcription: an RNA polymerase II-TFIIB cocrystal at 4.5 Å. *Science* 303, 983-988.
- Calvo,O. et Manley,J.L. (2001). Evolutionarily conserved interaction between CstF-64 and PC4 links transcription, polyadenylation, and termination. *Mol Cell* 7, 1013-1023.
- Carcamo,J., Buckbinder,L., et Reinberg,D. (1991). The initiator directs the assembly of a transcription factor IID-dependent transcription complex. *Proc. Natl. Acad. Sci. U. S. A* 88, 8052-8056.
- Chafin,D.R., Guo,H., et Price,D.H. (1995). Action of alpha-amanitin during pyrophosphorolysis and elongation by RNA polymerase II. *J. Biol. Chem.* 270, 19114-19119.
- Chalkley,G.E. et Verrijzer,C.P. (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., Wang,B.Q., Burton,Z.F., et Dahmus,M.E. (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.
- Chen,B.S. et Hampsey,M. (2004). Functional interaction between TFIIB and the Rpb2 subunit of RNA polymerase II: implications for the mechanism of transcription initiation. *Mol. Cell Biol.* 24, 3983-3991.

Chen,H.T. et Hahn,S. (2004). Mapping the location of TFIIB within the RNA polymerase II transcription preinitiation complex: a model for the structure of the PIC. *Cell* *119*, 169-180.

Chen,Y., Chafin,D., Price,D.H., et Greenleaf,A.L. (1996). *Drosophila* RNA polymerase II mutants that affect transcription elongation. *J. Biol. Chem.* *271*, 5993-5999.

Chen,Y., Weeks,J., Mortin,M.A., et Greenleaf,A.L. (1993). Mapping mutations in genes encoding the two large subunits of *Drosophila* RNA polymerase II defines domains essential for basic transcription functions and for proper expression of developmental genes. *Mol. Cell Biol.* *13*, 4214-4222.

Cheong,J.H., Yi,M., Lin,Y., et Murakami,S. (1995). Human RPB5, a subunit shared by eukaryotic nuclear RNA polymerases, binds human hepatitis B virus X protein and may play a role in X transactivation. *EMBO J.* *14*, 143-150.

Chicca,J.J., Auble,D.T., et Pugh,B.F. (1998). Cloning and biochemical characterization of TAF-172, a human homolog of yeast Mot1. *Mol. Cell Biol.* *18*, 1701-1710.

Cho,E.J., Kobor,M.S., Kim,M., Greenblatt,J., et Buratowski,S. (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,E.J., Takagi,T., Moore,C.R., et Buratowski,S. (1997). mRNA capping enzyme is recruited to the transcription complex by phosphorylation of the RNA polymerase II carboxy-terminal domain. *Genes Dev.* *11*, 3319-3326.

Choder,M. (2004). Rpb4 and Rpb7: subunits of RNA polymerase II and beyond. *Trends Biochem. Sci.* *29*, 674-681.

Choder,M. et Young,R.A. (1993). A portion of RNA polymerase II molecules has a component essential for stress responses and stress survival. *Mol. Cell Biol.* *13*, 6984-6991.

Chodosh,L.A., Fire,A., Samuels,M., et Sharp,P.A. (1989). 5,6-Dichloro-1-beta-D-ribofuranosylbenzimidazole inhibits transcription elongation by RNA polymerase II in vitro. *J. Biol. Chem.* *264*, 2250-2257.

Chung,W.H., Craighead,J.L., Chang,W.H., Ezeokonkwo,C., Bareket-Samish,A., Kornberg,R.D., et Asturias,F.J. (2003). RNA Polymerase II/TFIIF Structure and Conserved Organization of the Initiation Complex. *Mol. Cell* *12*, 1003-1013.

Conaway,J.W., Shilatifard,A., Dvir,A., et Conaway,R.C. (2000). Control of elongation by RNA polymerase II. *Trends Biochem. Sci.* *25*, 375-380.

Conaway,R.C., Garrett,K.P., Hanley,J.P., et Conaway,J.W. (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.

Conaway,R.C., Kong,S.E., et Conaway,J.W. (2003). TFIIS and GreB: two like-minded transcription elongation factors with sticky fingers. *Cell* *114*, 272-274.

Corden,J., Wasylyk,B., Buchwalder,A., Sassone-Corsi,P., Kedinger,C., et Chambon,P. (1980). Promoter sequences of eukaryotic protein-coding genes. *Science* *209*, 1406-1414.

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

Cosma,M. (2002). Ordered recruitment. Gene-specific mechanism of transcription activation. *Mol. Cell* *10*, 227.

Coulombe,B. (1999). DNA wrapping in transcription initiation by RNA polymerase II. *Biochem. Cell Biol.* *77*, 257-264.

Coulombe,B. et Burton,Z.F. (1999). DNA bending and wrapping around RNA polymerase: a "revolutionary" model describing transcriptional mechanisms. *Microbiol. Mol. Biol. Rev.* *63*, 457-478.

Coulombe,B., Killeen,M., Liljelund,P., Honda,B., Xiao,H., Ingles,C.J., et Greenblatt,J. (1992). Identification of three mammalian proteins that bind to the yeast TATA box protein TFIID. *Gene Expr.* *2*, 99-110.

Coulombe,B., Li,J., et Greenblatt,J. (1994). Topological localization of the human transcription factors IIA, IIB, TATA box-binding protein, and RNA polymerase II-associated protein 30 on a class II promoter. *J. Biol. Chem.* *269*, 19962-19967.

Cramer,P., Bushnell,D.A., et Kornberg,R.D. (2001). Structural basis of transcription: RNA polymerase II at 2.8 angstrom resolution. *Science* *292*, 1863-1876.

Crowley,T.E., Hoey,T., Liu,J.K., Jan,Y.N., Jan,L.Y., et Tjian,R. (1993). A new factor related to TATA-binding protein has highly restricted expression patterns in *Drosophila*. *Nature* *361*, 557-561.

Dantoni,J.C., Wurtz,J.M., Poch,O., Moras,D., et Tora,L. (1999). The TBP-like factor: an alternative transcription factor in metazoa? *Trends Biochem. Sci.* *24*, 335-339.

Dasgupta,A., Darst,R.P., Martin,K.J., Afshari,C.A., et Auble,D.T. (2002). Mot1 activates and represses transcription by direct, ATPase-dependent mechanisms. *Proc. Natl. Acad. Sci. U. S. A* *99*, 2666-2671.

Davidson,I. (2003). The genetics of TBP and TBP-related factors. *Trends Biochem. Sci.* *28*, 391-398.

de, L.S., I, Carlson, K.A., Hill, D.A., Guidi, C.J., Stephenson, R.O., Sif, S., Kingston, R.E., et Imbalzano, A.N. (2000). Mammalian SWI-SNF complexes contribute to activation of the hsp70 gene. *Mol. Cell Biol.* 20, 2839-2851.

DeJong, J., Bernstein, R., et Roeder, R.G. (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. et Roeder, R.G. (1993). A single cDNA, hTFIIA/alpha, encodes both the p35 and p19 subunits of human TFIIA. *Genes Dev.* 7, 2220-2234.

Dieci, G., Hermann-Le Denmat, S., Lukhtanov, E., Thuriaux, P., Werner, M., et Sentenac, A. (1995). A universally conserved region of the largest subunit participates in the active site of RNA polymerase III. *EMBO J.* 14, 3766-3776.

Doublet, S., Tabor, S., Long, A.M., Richardson, C.C., et Ellenberger, T. (1998). Crystal structure of a bacteriophage T7 DNA replication complex at 2.2 Å resolution. *Nature* 391, 251-258.

Douziech, M., Coin, F., Chipoulet, J.M., Arai, Y., Ohkuma, Y., Egly, J.M., et Coulombe, B. (2000). Mechanism of promoter melting by the xeroderma pigmentosum complementation group B helicase of transcription factor IIIH revealed by protein-DNA photo-cross-linking. *Mol. Cell Biol.* 20, 8168-8177.

Drapkin, R., Reardon, J.T., Ansari, A., Huang, J.C., Zawel, L., Ahn, K., Sancar, A., et Reinberg, D. (1994). Dual role of TFIIH in DNA excision repair and in transcription by RNA polymerase II. *Nature* 368, 769-772.

Drebot, M.A., Johnston, G.C., Friesen, J.D., et Singer, R.A. (1993). An impaired RNA polymerase II activity in *Saccharomyces cerevisiae* causes cell-cycle inhibition at START. *Mol Gen. Genet.* 241, 327-334.

Dye, M.J. et Proudfoot, N.J. (2001). Multiple transcript cleavage precedes polymerase release in termination by RNA polymerase II. *Cell* 105, 669-681.

Edwards, A.M., Kane, C.M., Young, R.A., et Kornberg, R.D. (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.

Eisen, J.A., Sweder, K.S., et Hanawalt, P.C. (1995). Evolution of the SNF2 family of proteins: subfamilies with distinct sequences and functions. *Nucleic Acids Res.* 23, 2715-2723.

Elsby, L.M. et Roberts, S.G. (2004). The role of TFIIB conformation in transcriptional regulation. *Biochem. Soc. Trans.* 32, 1098-1099.

- 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., Fairley,J.A., et Roberts,S.G. (2001). Activator-mediated disruption of sequence-specific DNA contacts by the general transcription factor TFIIB. *Genes Dev.* 15, 2945-2949.
- Exinger,F. et Lacroute,F. (1992). 6-Azaauracil Inhibition of Gtp Biosynthesis in *Saccharomyces-Cerevisiae*. *Current Genetics* 22, 9-11.
- Fairley,J.A., Evans,R., Hawkes,N.A., et Roberts,S.G. (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., Nahari,T., Hammel,C., Cole,C.N., et Choder,M. (2003). Rpb4p, a subunit of RNA polymerase II, mediates mRNA export during stress. *Mol Biol Cell* 14, 2744-2755.
- Featherstone,M. (2002). Coactivators in transcription initiation: here are your orders. *Curr. Opin. Genet. Dev.* 12, 149-155.
- Feaver,W.J., Svejstrup,J.Q., Henry,N.L., et Kornberg,R.D. (1994). Relationship of CDK-activating kinase and RNA polymerase II CTD kinase TFIIH/TFIIK. *Cell* 79, 1103-1109.
- Finkelstein,A., Kostrub,C.F., Li,J., Chavez,D.P., Wang,B.Q., Fang,S.M., Greenblatt,J., et Burton,Z.F. (1992). A cDNA encoding RAP74, a general initiation factor for transcription by RNA polymerase II. *Nature* 355, 464-467.
- Fish,R.N. et Kane,C.M. (2002). Promoting elongation with transcript cleavage stimulatory factors. *Biochim. Biophys. Acta* 1577, 287-307.
- Flanagan,P.M., Kelleher,R.J., III, Sayre,M.H., Tschochner,H., et Kornberg,R.D. (1991). A mediator required for activation of RNA polymerase II transcription in vitro. *Nature* 350, 436-438.
- Flores,O., Lu,H., et Reinberg,D. (1992). Factors involved in specific transcription by mammalian RNA polymerase II. Identification and characterization of factor IIIH. *J. Biol. Chem.* 267, 2786-2793.
- Flores,O., Maldonado,E., et Reinberg,D. (1989). Factors involved in specific transcription by mammalian RNA polymerase II. Factors IIE and IIF independently interact with RNA polymerase II. *J. Biol. Chem.* 264, 8913-8921.
- Fondell,J.D., Ge,H., et Roeder,R.G. (1996). Ligand induction of a transcriptionally active thyroid hormone receptor coactivator complex. *Proc. Natl. Acad. Sci. U. S. A* 93, 8329-8333.

- Fondell, J.D., Guermah, M., Malik, S., et Roeder, R.G. (1999). Thyroid hormone receptor-associated proteins and general positive cofactors mediate thyroid hormone receptor function in the absence of the TATA box-binding protein-associated factors of TFIID. *Proc. Natl. Acad. Sci. U. S. A* *96*, 1959-1964.
- Forget, D., Langelier, M.F., Therien, C., Trinh, V., et Coulombe, B. (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., Robert, F., Grondin, G., Burton, Z.F., Greenblatt, J., et Coulombe, B. (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.
- Fourel, G., Magdinier, F., et Gilson, E. (2004). Insulator dynamics and the setting of chromatin domains. *Bioessays* *26*, 523-532.
- Fryer, C.J. et Archer, T.K. (1998). Chromatin remodelling by the glucocorticoid receptor requires the BRG1 complex. *Nature* *393*, 88-91.
- Fujigasaki, H., Martin, J.J., De Deyn, P.P., Camuzat, A., Deffond, D., Stevanin, G., Dermaut, B., Van Broeckhoven, C., Durr, A., et Brice, A. (2001). CAG repeat expansion in the TATA box-binding protein gene causes autosomal dominant cerebellar ataxia. *Brain* *124*, 1939-1947.
- Garrett, K.P., Serizawa, H., Hanley, J.P., Bradsher, J.N., Tsuboi, A., Arai, N., Yokota, T., Arai, K., Conaway, R.C., et Conaway, J.W. (1992). The carboxyl terminus of RAP30 is similar in sequence to region 4 of bacterial sigma factors and is required for function. *J. Biol. Chem.* *267*, 23942-23949.
- Gasser, S.M. et Cockell, M.M. (2001). The molecular biology of the SIR proteins. *Gene* *279*, 1-16.
- Gaston, K. et Fried, M. (1995). CpG methylation and the binding of YY1 and ETS proteins to the Surf-1/Surf-2 bidirectional promoter. *Gene* *157*, 257-259.
- Gaston, K. et Jayaraman, P.S. (2003). Transcriptional repression in eukaryotes: repressors and repression mechanisms. *Cell Mol Life Sci.* *60*, 721-741.
- Geiger, J.H., Hahn, S., Lee, S., et Sigler, P.B. (1996). Crystal structure of the yeast TFIIA/TBP/DNA complex. *Science* *272*, 830-836.
- Geisberg, J.V., Moqtaderi, Z., Kuras, L., et Struhl, K. (2002). Mot1 associates with transcriptionally active promoters and inhibits association of NC2 in *Saccharomyces cerevisiae*. *Mol. Cell Biol.* *22*, 8122-8134.
- Gelbart, M.E., Rechsteiner, T., Richmond, T.J., et Tsukiyama, T. (2001). Interactions of Isw2 chromatin remodeling complex with nucleosomal arrays: analyses using recombinant yeast histones and immobilized templates. *Mol. Cell Biol.* *21*, 2098-2106.

Gerard,M., Fischer,L., Moncollin,V., Chipoulet,J.M., Chambon,P., et Egly,J.M. (1991). Purification and interaction properties of the human RNA polymerase B(II) general transcription factor BTF2. *J. Biol. Chem.* 266, 20940-20945.

Giaever,G., Chu,A.M., Ni,L., Connelly,C., Riles,L., Veronneau,S., Dow,S., Lucau-Danila,A., Anderson,K., Andre,B., Arkin,A.P., Astromoff,A., El Bakkoury,M., Bangham,R., Benito,R., Brachat,S., Campanaro,S., Curtiss,M., Davis,K., Deutschbauer,A., Entian,K.D., Flaherty,P., Foury,F., Garfinkel,D.J., Gerstein,M., Gotte,D., Guldener,U., Hegemann,J.H., Hempel,S., Herman,Z., Jaramillo,D.F., Kelly,D.E., Kelly,S.L., Kotter,P., LaBonte,D., Lamb,D.C., Lan,N., Liang,H., Liao,H., Liu,L., Luo,C., Lussier,M., Mao,R., Menard,P., Ooi,S.L., Revuelta,J.L., Roberts,C.J., Rose,M., Ross-Macdonald,P., Scherens,B., Schimmack,G., Shafer,B., Shoemaker,D.D., Sookhai-Mahadeo,S., Storms,R.K., Strathern,J.N., Valle,G., Voet,M., Volckaert,G., Wang,C.Y., Ward,T.R., Wilhelmy,J., Winzeler,E.A., Yang,Y., Yen,G., Youngman,E., Yu,K., Bussey,H., Boeke,J.D., Snyder,M., Philippsen,P., Davis,R.W., et Johnston,M. (2002). Functional profiling of the *Saccharomyces cerevisiae* genome. *Nature* 418, 387-391.

Giglia-Mari,G., Coin,F., Ranish,J.A., Hoogstraten,D., Theil,A., Wijgers,N., Jaspers,N.G., Raams,A., Argentini,M., van der Spek,P.J., Botta,E., Stefanini,M., Egly,J.M., Aebersold,R., Hoeijmakers,J.H., et Vermeulen,W. (2004). A new, tenth subunit of TFIID is responsible for the DNA repair syndrome trichothiodystrophy group A. *Nat. Genet.* 36, 714-719.

Gilbert,N., Boyle,S., Fiegler,H., Woodfine,K., Carter,N.P., et Bickmore,W.A. (2004). Chromatin architecture of the human genome: gene-rich domains are enriched in open chromatin fibers. *Cell* 118, 555-566.

Glass,C.K. (1994). Differential recognition of target genes by nuclear receptor monomers, dimers, and heterodimers. *Endocr. Rev.* 15, 391-407.

Glossop,J.A., Dafforn,T.R., et Roberts,S.G. (2004). A conformational change in TFIIB is required for activator-mediated assembly of the preinitiation complex. *Nucleic Acids Res.* 32, 1829-1835.

Gnatt,A.L., Cramer,P., Fu,J., Bushnell,D.A., et Kornberg,R.D. (2001). Structural basis of transcription: an RNA polymerase II elongation complex at 3.3 Å resolution. *Science* 292, 1876-1882.

Gong,X.Q., Nedialkov,Y.A., et Burton,Z.F. (2004). Alpha-amanitin blocks translocation by human RNA polymerase II. *J. Biol Chem.* 279, 27422-27427.

Goodrich,J.A. et Tjian,R. (1994). Transcription factors IIE and IIH and ATP hydrolysis direct promoter clearance by RNA polymerase II. *Cell* 77, 145-156.

Grant,P.A., Duggan,L., Cote,J., Roberts,S.M., Brownell,J.E., Candau,R., Ohba,R., Owen-Hughes,T., Allis,C.D., Winston,F., Berger,S.L., et Workman,J.L. (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., Schieltz,D., Pray-Grant,M.G., Steger,D.J., Reese,J.C., Yates,J.R., III, et Workman,J.L. (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.

Gromoller,A. et Lehming,N. (2000). Srb7p is a physical and physiological target of Tup1p. *EMBO J.* *19*, 6845-6852.

Gross,C.A., Chan,C., Dombroski,A., Gruber,T., Sharp,M., Tupy,J., et Young,B. (1998). The functional and regulatory roles of sigma factors in transcription. *Cold Spring Harb. Symp. Quant. Biol.* *63*, 141-155.

Grosveld,F. (1999). Activation by locus control regions? *Curr. Opin. Genet. Dev.* *9*, 152-157.

Gu,W., Malik,S., Ito,M., Yuan,C.X., Fondell,J.D., Zhang,X., Martinez,E., Qin,J., et Roeder,R.G. (1999). A novel human SRB/MED-containing cofactor complex, SMCC, involved in transcription regulation. *Mol. Cell* *3*, 97-108.

Gumbs,O.H., Campbell,A.M., et Weil,P.A. (2003). High-affinity DNA binding by a Mot1p-TBP complex: implications for TAF-independent transcription. *EMBO J.* *22*, 3131-3141.

Ha,I., Lane,W.S., et Reinberg,D. (1991). Cloning of a human gene encoding the general transcription initiation factor IIB. *Nature* *352*, 689-695.

Ha,I., Roberts,S., Maldonado,E., Sun,X., Kim,L.U., Green,M., et Reinberg,D. (1993). Multiple functional domains of human transcription factor IIB: distinct interactions with two general transcription factors and RNA polymerase II. *Genes Dev.* *7*, 1021-1032.

Hahn,S., Buratowski,S., Sharp,P.A., et Guarente,L. (1989). Yeast TATA-binding protein TFIID binds to TATA elements with both consensus and nonconsensus DNA sequences. *Proc. Natl. Acad. Sci. U. S. A* *86*, 5718-5722.

Hahn,S., Hoar,E.T., et Guarente,L. (1985). Each of three "TATA elements" specifies a subset of the transcription initiation sites at the CYC-1 promoter of *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. U. S. A* *82*, 8562-8566.

Hampsey,M. (1998). Molecular genetics of the RNA polymerase II general transcriptional machinery. *Microbiol. Mol. Biol. Rev.* *62*, 465-503.

Hanna-Rose,W. et Hansen,U. (1996). Active repression mechanisms of eukaryotic transcription repressors. *Trends Genet.* *12*, 229-234.

- Hansen,S.K., Takada,S., Jacobson,R.H., Lis,J.T., et Tjian,R. (1997). Transcription properties of a cell type-specific TATA-binding protein, TRF. *Cell* 91, 71-83.
- Hawkes,N.A., Evans,R., et Roberts,S.G. (2000). The conformation of the transcription factor TFIIB modulates the response to transcriptional activators in vivo. *Curr. Biol.* 10, 273-276.
- Hawkes,N.A. et Roberts,S.G. (1999). The role of human TFIIB in transcription start site selection in vitro and in vivo. *J. Biol. Chem.* 274, 14337-14343.
- Heisler,L.M., Suzuki,H., Landick,R., et Gross,C.A. (1993). Four contiguous amino acids define the target for streptolydigin resistance in the beta subunit of Escherichia coli RNA polymerase. *J. Biol Chem.* 268, 25369-25375.
- Hekmatpanah,D.S. et Young,R.A. (1991). Mutations in a conserved region of RNA polymerase II influence the accuracy of mRNA start site selection. *Mol. Cell Biol.* 11, 5781-5791.
- Hemming,S.A., Jansma,D.B., Macgregor,P.F., Goryachev,A., Friesen,J.D., et Edwards,A.M. (2000). RNA polymerase II subunit Rpb9 regulates transcription elongation in vivo. *J. Biol. Chem.* 275, 35506-35511.
- Hengartner,C.J., Myer,V.E., Liao,S.M., Wilson,C.J., Koh,S.S., et Young,R.A. (1998). Temporal regulation of RNA polymerase II by Srb10 and Kin28 cyclin-dependent kinases. *Mol. Cell* 2, 43-53.
- Hodo,H.G., III et Blatti,S.P. (1977). Purification using polyethylenimine precipitation and low molecular weight subunit analyses of calf thymus and wheat germ DNA-dependent RNA polymerase II. *Biochemistry* 16, 2334-2343.
- Hoiby,T., Mitsiou,D.J., Zhou,H.Q., Erdjument-Bromage,H., Tempst,P., et Stunnenberg,H.G. (2004). Cleavage and proteasome-mediated degradation of the basal transcription factor TFIIA
2. *EMBO J.* 23, 3083-3091.
- Holmes,M.C. et Tjian,R. (2000). Promoter-selective properties of the TBP-related factor TRF1. *Science* 288, 867-870.
- Holstege,F.C., Fiedler,U., et Timmers,H.T. (1997). Three transitions in the RNA polymerase II transcription complex during initiation. *EMBO J.* 16, 7468-7480.
- Holstege,F.C., Jennings,E.G., Wyrick,J.J., Lee,T.I., Hengartner,C.J., Green,M.R., Golub,T.R., Lander,E.S., et Young,R.A. (1998). Dissecting the regulatory circuitry of a eukaryotic genome. *Cell* 95, 717-728.
- Holstege,F.C., Tantin,D., Carey,M., van der Vliet,P.C., et Timmers,H.T. (1995). The requirement for the basal transcription factor IIE is determined by the helical stability of promoter DNA. *EMBO J.* 14, 810-819.

- Holstege,F.C., van der Vliet,P.C., et Timmers,H.T. (1996). Opening of an RNA polymerase II promoter occurs in two distinct steps and requires the basal transcription factors IIE and IIH. *EMBO J.* *15*, 1666-1677.
- Horn,P.J. et Peterson,C.L. (2002). MOLECULAR BIOLOGY: Chromatin Higher Order Folding--Wrapping up Transcription. *Science* *297*, 1824-1827.
- Hull,M.W., McKune,K., et Woychik,N.A. (1995). RNA polymerase II subunit RPB9 is required for accurate start site selection. *Genes Dev.* *9*, 481-490.
- Imbalzano,A.N., Schnitzler,G.R., et Kingston,R.E. (1996). Nucleosome disruption by human SWI/SNF is maintained in the absence of continued ATP hydrolysis. *J. Biol. Chem.* *271*, 20726-20733.
- Ingles,C.J., Shales,M., Cress,W.D., Triezenberg,S.J., et Greenblatt,J. (1991). Reduced binding of TFIID to transcriptionally compromised mutants of VP16. *Nature* *351*, 588-590.
- Inoue,T., Geyer,R.K., Howard,D., Yu,Z.K., et Maki,C.G. (2001). MDM2 can promote the ubiquitination, nuclear export, and degradation of p53 in the absence of direct binding. *J. Biol Chem.* *276*, 45255-45260.
- International Human Genome Sequencing Consortium (2004). Finishing the euchromatic sequence of the human genome. *Nature* *431*, 931-945.
- Ito,H., Fukuda,Y., Murata,K., et Kimura,A. (1983). Transformation of intact yeast cells treated with alkali cations. *J. Bacteriol.* *153*, 163-168.
- Ito,M., Yuan,C.X., Malik,S., Gu,W., Fondell,J.D., Yamamura,S., Fu,Z.Y., Zhang,X., Qin,J., et Roeder,R.G. (1999). Identity between TRAP and SMCC complexes indicates novel pathways for the function of nuclear receptors and diverse mammalian activators. *Mol. Cell* *3*, 361-370.
- Ito,T., Bulger,M., Pazin,M.J., Kobayashi,R., et Kadonaga,J.T. (1997). ACF, an ISWI-containing and ATP-utilizing chromatin assembly and remodeling factor. *Cell* *90*, 145-155.
- Ivanov,D., Kwak,Y.T., Guo,J., et Gaynor,R.B. (2000). Domains in the SPT5 protein that modulate its transcriptional regulatory properties. *Mol Cell Biol* *20*, 2970-2983.
- Izban,M.G. et Luse,D.S. (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. et Luse,D.S. (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.

Jacobson,R.H., Ladurner,A.G., King,D.S., et Tjian,R. (2000). Structure and function of a human TAFII250 double bromodomain module. *Science* 288, 1422-1425.

Jaskelioff,M., Gavin,I.M., Peterson,C.L., et Logie,C. (2000). SWI-SNF-mediated nucleosome remodeling: role of histone octamer mobility in the persistence of the remodeled state. *Mol. Cell Biol.* 20, 3058-3068.

Javahery,R., Khachi,A., Lo,K., Zenzie-Gregory,B., et Smale,S.T. (1994). DNA sequence requirements for transcriptional initiator activity in mammalian cells. *Mol. Cell Biol.* 14, 116-127.

Jeon,C., Yoon,H., et Agarwal,K. (1994). The transcription factor TFIIS zinc ribbon dipeptide Asp-Glu is critical for stimulation of elongation and RNA cleavage by RNA polymerase II. *Proc. Natl. Acad. Sci. U. S. A* 91, 9106-9110.

Jeronimo,C., Langelier,M.F., Zeghouf,M., Cojocaru,M., Bergeron,D., Baali,D., Forget,D., Mnaimneh,S., Davierwala,A.P., Pootoolal,J., Chandy,M., Canadien,V., Beattie,B.K., Richards,D.P., Workman,J.L., Hughes,T.R., Greenblatt,J., et Coulombe,B. (2004). RPAP1, a novel human RNA polymerase II-associated protein affinity purified with recombinant wild-type and mutated polymerase subunits 149. *Mol Cell Biol* 24, 7043-7058.

Jiang,Y.W., Veschambre,P., Erdjument-Bromage,H., Tempst,P., Conaway,J.W., Conaway,R.C., et Kornberg,R.D. (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.

Jin,D.J. et Gross,C.A. (1991). RpoB8, a rifampicin-resistant termination-proficient RNA polymerase, has an increased Km for purine nucleotides during transcription elongation. *J. Biol. Chem.* 266, 14478-14485.

Kamenski,T., Heilmeyer,S., Meinhart,A., et Cramer,P. (2004). Structure and mechanism of RNA polymerase II CTD phosphatases. *Mol Cell* 15, 399-407.

Kashlev,M., Lee,J., Zalenskaya,K., Nikiforov,V., et Goldfarb,A. (1990). Blocking of the initiation-to-elongation transition by a transdominant RNA polymerase mutation. *Science* 248, 1006-1009.

Kelleher,R.J., III, Flanagan,P.M., et Kornberg,R.D. (1990). A novel mediator between activator proteins and the RNA polymerase II transcription apparatus. *Cell* 61, 1209-1215.

Keriel,A., Stary,A., Sarasin,A., Rochette-Egly,C., et Egly,J.M. (2002). XPD Mutations Prevent TFIID-Dependent Transactivation by Nuclear Receptors and Phosphorylation of RARalpha. *Cell* 109, 125-135.

Kettenberger,H., Armache,K.J., et Cramer,P. (2003). Architecture of the RNA polymerase II-TFIIS complex and implications for mRNA cleavage. *Cell* 114, 347-357.

- Kettenberger,H., Armache,K.J., et Cramer,P. (2004). Complete RNA Polymerase II Elongation Complex Structure and Its Interactions with NTP and TFIIS. *Mol Cell* 16, 955-965.
- Khorasanizadeh,S. (2004). The nucleosome: from genomic organization to genomic regulation. *Cell* 116, 259-272.
- Kiefer,J.R., Mao,C., Braman,J.C., et Beese,L.S. (1998). Visualizing DNA replication in a catalytically active *Bacillus* DNA polymerase crystal. *Nature* 391, 304-307.
- Killeen,M., Coulombe,B., et Greenblatt,J. (1992). Recombinant TBP, transcription factor IIB, and RAP30 are sufficient for promoter recognition by mammalian RNA polymerase II. *J. Biol. Chem.* 267, 9463-9466.
- Killeen,M.T. et Greenblatt,J.F. (1992). The general transcription factor RAP30 binds to RNA polymerase II and prevents it from binding nonspecifically to DNA. *Mol. Cell Biol.* 12, 30-37.
- Kim,J.B. et Sharp,P.A. (2001). Positive transcription elongation factor B phosphorylates hSPT5 and RNA polymerase II carboxyl-terminal domain independently of cyclin-dependent kinase-activating kinase. *J. Biol. Chem.* 276, 12317-12323.
- Kim,J.L., Nikolov,D.B., et Burley,S.K. (1993a). Co-crystal structure of TBP recognizing the minor groove of a TATA element. *Nature* 365, 520-527.
- Kim,M., Ahn,S.H., Krogan,N.J., Greenblatt,J.F., et Buratowski,S. (2004a). Transitions in RNA polymerase II elongation complexes at the 3' ends of genes. *EMBO J.* 23, 354-364.
- Kim,M., Krogan,N.J., Vasiljeva,L., Rando,O.J., Nedeia,E., Greenblatt,J.F., et Buratowski,S. (2004b). The yeast Rat1 exonuclease promotes transcription termination by RNA polymerase II. *Nature* 432, 517-522.
- Kim,T.K., Ebricht,R.H., et Reinberg,D. (2000). Mechanism of ATP-dependent promoter melting by transcription factor IIIH. *Science* 288, 1418-1422.
- Kim,Y., Geiger,J.H., Hahn,S., et Sigler,P.B. (1993b). Crystal structure of a yeast TBP/TATA-box complex. *Nature* 365, 512-520.
- Kim,Y.J., Bjorklund,S., Li,Y., Sayre,M.H., et Kornberg,R.D. (1994). 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., Suzuki,H., et Ishihama,A. (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.

- Kireeva, M.L., Walter, W., Tchernajenko, V., Bondarenko, V., Kashlev, M., et Studitsky, V.M. (2002). Nucleosome Remodeling Induced by RNA Polymerase II. Loss of the H2A/H2B Dimer during Transcription. *Mol. Cell* 9, 541-552.
- Kobayashi, N., Boyer, T.G., et Berk, A.J. (1995). A class of activation domains interacts directly with TFIIA and stimulates TFIIA-TFIID-promoter complex assembly. *Mol. Cell Biol.* 15, 6465-6473.
- Koh, S.S., Ansari, A.Z., Ptashne, M., et Young, R.A. (1998). An activator target in the RNA polymerase II holoenzyme. *Mol. Cell* 1, 895-904.
- Kokubo, T., Swanson, M.J., Nishikawa, J.I., Hinnebusch, A.G., et Nakatani, Y. (1998). The yeast TAF145 inhibitory domain and TFIIA competitively bind to TATA-binding protein. *Mol. Cell Biol.* 18, 1003-1012.
- Koleske, A.J. et Young, R.A. (1994). An RNA polymerase II holoenzyme responsive to activators. *Nature* 368, 466-469.
- Komarnitsky, P., Cho, E.J., et Buratowski, S. (2000). Different phosphorylated forms of RNA polymerase II and associated mRNA processing factors during transcription. *Genes Dev.* 14, 2452-2460.
- Krishnamurthy, S., He, X., Reyes-Reyes, M., Moore, C., et Hampsey, M. (2004). Ssu72 Is an RNA polymerase II CTD phosphatase. *Mol Cell* 14, 387-394.
- Krogan, N.J., Kim, M., Ahn, S.H., Zhong, G., Kobor, M.S., Cagney, G., Emili, A., Shilatifard, A., Buratowski, S., et Greenblatt, J.F. (2002). RNA polymerase II elongation factors of *Saccharomyces cerevisiae*: a targeted proteomics approach. *Mol. Cell Biol.* 22, 6979-6992.
- Krummel, B. et Chamberlin, M.J. (1989). RNA chain initiation by *Escherichia coli* RNA polymerase. Structural transitions of the enzyme in early ternary complexes. *Biochemistry* 28, 7829-7842.
- Kuchin, S., Yeghiayan, P., et Carlson, M. (1995). Cyclin-dependent protein kinase and cyclin homologs SSN3 and SSN8 contribute to transcriptional control in yeast. *Proc. Natl. Acad. Sci. U. S. A* 92, 4006-4010.
- Kuldell, N.H. et Buratowski, S. (1997). Genetic analysis of the large subunit of yeast transcription factor IIE reveals two regions with distinct functions. *Mol. Cell Biol.* 17, 5288-5298.
- Kumar, K.P., Akoulitchev, S., et Reinberg, D. (1998). Promoter-proximal stalling results from the inability to recruit transcription factor IIH to the transcription complex and is a regulated event. *Proc. Natl. Acad. Sci. U. S. A* 95, 9767-9772.
- Kuras, L., Borggreffe, T., et Kornberg, R.D. (2003). Association of the Mediator complex with enhancers of active genes. *Proc. Natl. Acad. Sci. U. S. A* 100, 13887-13891.

Kutach, A.K. et Kadonaga, J.T. (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.

Kuznedelov, K., Korzheva, N., Mustaev, A., et Severinov, K. (2002a). Structure-based analysis of RNA polymerase function: the largest subunit's rudder contributes critically to elongation complex stability and is not involved in the maintenance of RNA-DNA hybrid length. *EMBO J.* *21*, 1369-1378.

Kuznedelov, K., Minakhin, L., Niedziela-Majka, A., Dove, S.L., Rogulja, D., Nickels, B.E., Hochschild, A., Heyduk, T., et Severinov, K. (2002b). A role for interaction of the RNA polymerase flap domain with the sigma subunit in promoter recognition. *Science* *295*, 855-857.

Lachner, M., O'Sullivan, R.J., et Jenuwein, T. (2003). An epigenetic road map for histone lysine methylation. *J. Cell Sci.* *116*, 2117-2124.

Lagrange, T., Kapanidis, A.N., Tang, H., Reinberg, D., et Ebright, R.H. (1998). New core promoter element in RNA polymerase II-dependent transcription: sequence-specific DNA binding by transcription factor IIB. *Genes Dev.* *12*, 34-44.

Lagrange, T., Kim, T.K., Orphanides, G., Ebright, Y.W., Ebright, R.H., et Reinberg, D. (1996). High-resolution mapping of nucleoprotein complexes by site-specific protein-DNA photocrosslinking: organization of the human TBP-TFIIA-TFIIB-DNA quaternary complex. *Proc. Natl. Acad. Sci. U. S. A* *93*, 10620-10625.

Landick, R. (2004). Active-site dynamics in RNA polymerases. *Cell* *116*, 351-353.

Langelier, M.F., Baali, D., Trinh, V., Greenblatt, J., Archambault, J., et Coulombe, B. (2005). A Mutational Analysis of the Human RNA Polymerase II Catalytic Center. Article en préparation.

Langelier, M.F., Forget, D., Rojas, A., Porlier, Y., Burton, Z.F., et Coulombe, B. (2001). Structural and functional interactions of transcription factor (TF) IIA with TFIIE and TFIIF in transcription initiation by RNA polymerase II. *J. Biol. Chem.* *276*, 38652-38657.

Langelier, M.F., Trinh, V., et Coulombe, B. (2002). Gros plan sur l'ARN polymérase II. *Medecine Sciences* *18*, 211-216.

Larkin, R.M. et Guilfoyle, T.J. (1997). Reconstitution of yeast and *Arabidopsis* RNA polymerase alpha-like subunit heterodimers. *J. Biol Chem.* *272*, 12824-12830.

Larochelle, S., Pandur, J., Fisher, R.P., Salz, H.K., et Suter, B. (1998). Cdk7 is essential for mitosis and for in vivo Cdk-activating kinase activity. *Genes Dev.* *12*, 370-381.

Le Page,F., Kwoh,E.E., Avrutskaya,A., Gentil,A., Leadon,S.A., Sarasin,A., et Cooper,P.K. (2000). Transcription-coupled repair of 8-oxoguanine: requirement for XPG, TFIID, and CSB and implications for Cockayne syndrome. *Cell* 101, 159-171.

Lee,J., Kashlev,M., Borukhov,S., et Goldfarb,A. (1991). A β Subunit Mutation Disrupting the Catalytic Function of Escherichia coli RNA Polymerase. *PNAS* 88, 6018-6022.

Lee,T.I., Causton,H.C., Holstege,F.C., Shen,W.C., Hannett,N., Jennings,E.G., Winston,F., Green,M.R., et Young,R.A. (2000). Redundant roles for the TFIID and SAGA complexes in global transcription. *Nature* 405, 701-704.

Lee,T.I. et Young,R.A. (2000). Transcription of eukaryotic protein-coding genes. *Annu. Rev. Genet.* 34, 77-137.

Lee,Y.C., Min,S., Gim,B.S., et Kim,Y.J. (1997). A transcriptional mediator protein that is required for activation of many RNA polymerase II promoters and is conserved from yeast to humans. *Mol. Cell Biol.* 17, 4622-4632.

Lei,L., Ren,D., Finkelstein,A., et Burton,Z.F. (1998). Functions of the N- and C-terminal domains of human RAP74 in transcriptional initiation, elongation, and recycling of RNA polymerase II. *Mol. Cell Biol.* 18, 2130-2142.

Lennon,J.C., III, Wind,M., Saunders,L., Hock,M.B., et Reines,D. (1998). Mutations in RNA polymerase II and elongation factor SII severely reduce mRNA levels in *Saccharomyces cerevisiae*. *Mol. Cell Biol.* 18, 5771-5779.

LeRoy,G., Loyola,A., Lane,W.S., et Reinberg,D. (2000). Purification and characterization of a human factor that assembles and remodels chromatin. *J. Biol. Chem.* 275, 14787-14790.

Leurent,C., Sanders,S., Ruhlmann,C., Mallouh,V., Weil,P.A., Kirschner,D.B., Tora,L., et Schultz,P. (2002). Mapping histone fold TAFs within yeast TFIID. *EMBO J.* 21, 3424-3433.

Leurent,C., Sanders,S.L., Demeny,M.A., Garbett,K.A., Ruhlmann,C., Weil,P.A., Tora,L., et Schultz,P. (2004). Mapping key functional sites within yeast TFIID. *EMBO J.* 23, 719-727.

Lewis,B.A. et Reinberg,D. (2003). The mediator coactivator complex: functional and physical roles in transcriptional regulation. *J. Cell Sci.* 116, 3667-3675.

Li,S. et Smerdon,M.J. (2002). Rpb4 and Rpb9 mediate subpathways of transcription-coupled DNA repair in *Saccharomyces cerevisiae*. *EMBO J.* 21, 5921-5929.

Lieberman,P. (1994). Identification of functional targets of the Zta transcriptional activator by formation of stable preinitiation complex intermediates. *Mol. Cell Biol.* 14, 8365-8375.

Lieberman,P.M., Ozer,J., et Gursel,D.B. (1997). Requirement for transcription factor IIA (TFIIA)-TFIID recruitment by an activator depends on promoter structure and template competition. *Mol. Cell Biol.* 17, 6624-6632.

Lim,C.Y., Santoso,B., Boulay,T., Dong,E., Ohler,U., et Kadonaga,J.T. (2004). The MTE, a new core promoter element for transcription by RNA polymerase II. *Genes Dev.* 18, 1606-1617.

Lindstrom,D.L., Squazzo,S.L., Muster,N., Burckin,T.A., Wachter,K.C., Emigh,C.A., McCleery,J.A., Yates,J.R., III, et Hartzog,G.A. (2003). Dual roles for Spt5 in pre-mRNA processing and transcription elongation revealed by identification of Spt5-associated proteins. *Mol. Cell Biol.* 23, 1368-1378.

Lis,J.T., Mason,P., Peng,J., Price,D.H., et Werner,J. (2000). P-TEFb kinase recruitment and function at heat shock loci. *Genes Dev.* 14, 792-803.

Lo,K. et Smale,S.T. (1996). Generality of a functional initiator consensus sequence. *Gene* 182, 13-22.

Lomvardas,S. et Thanos,D. (2001). Nucleosome sliding via TBP DNA binding in vivo. *Cell* 106, 685-696.

Lorch,Y., Zhang,M., et Kornberg,R.D. (2001). RSC unravels the nucleosome. *Mol. Cell* 7, 89-95.

Lu,H., Flores,O., Weinmann,R., et Reinberg,D. (1991). The nonphosphorylated form of RNA polymerase II preferentially associates with the preinitiation complex. *Proc. Natl. Acad. Sci. U. S. A* 88, 10004-10008.

Luger,K., Mader,A.W., Richmond,R.K., Sargent,D.F., et Richmond,T.J. (1997). Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* 389, 251-260.

Lusser,A. et Kadonaga,J.T. (2003). Chromatin remodeling by ATP-dependent molecular machines. *Bioessays* 25, 1192-1200.

Ma,D., Olave,I., Merino,A., et Reinberg,D. (1996). Separation of the transcriptional coactivator and antirepression functions of transcription factor IIA. *Proc. Natl. Acad. Sci. U. S. A* 93, 6583-6588.

Ma,D., Watanabe,H., Mermelstein,F., Admon,A., Oguri,K., Sun,X., Wada,T., Imai,T., Shiroya,T., Reinberg,D., et . (1993). Isolation of a cDNA encoding the largest subunit of TFIIA reveals functions important for activated transcription. *Genes Dev.* 7, 2246-2257.

Malik,S. et Roeder,R.G. (2000). Transcriptional regulation through Mediator-like coactivators in yeast and metazoan cells. *Trends Biochem. Sci.* 25, 277-283.

Marmorstein,R. (2001). Structure of histone acetyltransferases. *J. Mol. Biol.* 311, 433-444.

- Marmorstein,R. et Berger,S.L. (2001). Structure and function of bromodomains in chromatin-regulating complexes. *Gene* 272, 1-9.
- Marshall,N.F. et Dahmus,M.E. (2000). C-terminal domain phosphatase sensitivity of RNA polymerase II in early elongation complexes on the HIV-1 and adenovirus 2 major late templates. *J. Biol. Chem.* 275, 32430-32437.
- Martens,J.A. et Winston,F. (2003). Recent advances in understanding chromatin remodeling by Swi/Snf complexes. *Curr. Opin. Genet. Dev.* 13, 136-142.
- Martinez,E., Zhou,Q., L'Etoile,N.D., Oelgeschlager,T., Berk,A.J., et Roeder,R.G. (1995). Core promoter-specific function of a mutant transcription factor TFIID defective in TATA-box binding. *Proc. Natl. Acad. Sci. U. S. A* 92, 11864-11868.
- Matangkasombut,O., Auty,R., et Buratowski,S. (2004). Structure and function of the TFIID complex. *Proteins in Eukaryotic Transcription* 67, 67-92.
- Maxon,M.E., Goodrich,J.A., et Tjian,R. (1994). Transcription factor IIE binds preferentially to RNA polymerase IIa and recruits TFIIF: a model for promoter clearance. *Genes Dev.* 8, 515-524.
- McCracken,S. et Greenblatt,J. (1991). Related RNA polymerase-binding regions in human RAP30/74 and *Escherichia coli* sigma 70. *Science* 253, 900-902.
- Meisterernst,M. et Roeder,R.G. (1991). Family of proteins that interact with TFIID and regulate promoter activity. *Cell* 67, 557-567.
- Merika,M. et Thanos,D. (2001). Enhanceosomes. *Curr. Opin. Genet. Dev.* 11, 205-208.
- Merika,M., Williams,A.J., Chen,G., Collins,T., et Thanos,D. (1998). Recruitment of CBP/p300 by the IFN beta enhanceosome is required for synergistic activation of transcription. *Mol. Cell* 1, 277-287.
- Merino,A., Madden,K.R., Lane,W.S., Champoux,J.J., et Reinberg,D. (1993). DNA topoisomerase I is involved in both repression and activation of transcription. *Nature* 365, 227-232.
- Mermelstein,F., Yeung,K., Cao,J., Inostroza,J.A., Erdjument-Bromage,H., Egelson,K., Landsman,D., Levitt,P., Tempst,P., et Reinberg,D. (1996). Requirement of a corepressor for Dr1-mediated repression of transcription. *Genes Dev.* 10, 1033-1048.
- Minakhin,L., Bhagat,S., Brunning,A., Campbell,E.A., Darst,S.A., Ebright,R.H., et Severinov,K. (2001). Bacterial RNA polymerase subunit omega and eukaryotic RNA polymerase subunit RPB6 are sequence, structural, and functional homologs and promote RNA polymerase assembly. *Proc. Natl. Acad. Sci. U. S. A* 98, 892-897.
- Mitchell,P.J. et Tjian,R. (1989). Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins. *Science* 245, 371-378.

- Mitsiou,D.J. et Stunnenberg,H.G. (2000). TAC, a TBP-sans-TAFs complex containing the unprocessed TFIIA α precursor and the TFIIA γ subunit. *Mol. Cell* 6, 527-537.
- Mitsiou,D.J. et Stunnenberg,H.G. (2003). p300 is involved in formation of the TBP-TFIIA-containing basal transcription complex, TAC. *EMBO J.* 22, 4501-4511.
- Miyao,T., Barnett,J.D., et Woychik,N.A. (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.
- Mooney,R.A., Artsimovitch,I., et Landick,R. (1998). Information processing by RNA polymerase: recognition of regulatory signals during RNA chain elongation. *J. Bacteriol.* 180, 3265-3275.
- Morillon,A., Karabetsov,N., O'Sullivan,J., Kent,N., Proudfoot,N., et Mellor,J. (2003). Isw1 chromatin remodeling ATPase coordinates transcription elongation and termination by RNA polymerase II. *Cell* 115, 425-435.
- Muscat,G.E., Burke,L.J., et Downes,M. (1998). The corepressor N-CoR and its variants RIP13a and RIP13Delta1 directly interact with the basal transcription factors TFIIB, TAFII32 and TAFII70. *Nucleic Acids Res.* 26, 2899-2907.
- Mustaev,A., Kashlev,M., Lee,J.Y., Polyakov,A., Lebedev,A., Zalenskaya,K., Grachev,M., Goldfarb,A., et Nikiforov,V. (1991). Mapping of the priming substrate contacts in the active center of Escherichia coli RNA polymerase. *J. Biol. Chem.* 266, 23927-23931.
- Myers,L.C. et Kornberg,R.D. (2000). Mediator of transcriptional regulation. *Annu. Rev. Biochem.* 69, 729-749.
- Naar,A.M., Beurang,P.A., Zhou,S., Abraham,S., Solomon,W., et Tjian,R. (1999). Composite co-activator ACR mediates chromatin-directed transcriptional activation. *Nature* 398, 828-832.
- Nagawa,F. et Fink,G.R. (1985). The relationship between the "TATA" sequence and transcription initiation sites at the HIS4 gene of Saccharomyces cerevisiae. *Proc. Natl. Acad. Sci. U. S. A* 82, 8557-8561.
- Nakanishi,T., Nakano,A., Nomura,K., Sekimizu,K., et Natori,S. (1992). Purification, gene cloning, and gene disruption of the transcription elongation factor S-II in Saccharomyces cerevisiae. *J. Biol. Chem.* 267, 13200-13204.
- Nan,X., Ng,H.H., Johnson,C.A., Laherty,C.D., Turner,B.M., Eisenman,R.N., et Bird,A. (1998). Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. *Nature* 393, 386-389.

Nedialkov, Y.A., Gong, X.Q., Hovde, S.L., Yamaguchi, Y., Handa, H., Geiger, J.H., Yan, H., et Burton, Z.F. (2003). NTP-driven Translocation by Human RNA Polymerase II. *J. Biol. Chem.* 278, 18303-18312.

Neuman, K.C., Abbondanzieri, E.A., Landick, R., Gelles, J., et Block, S.M. (2003). Ubiquitous transcriptional pausing is independent of RNA polymerase backtracking. *Cell* 115, 437-447.

Ng, H.H., Robert, F., Young, R.A., et Struhl, K. (2002). Genome-wide location and regulated recruitment of the RSC nucleosome-remodeling complex. *Genes Dev.* 16, 806-819.

Ng, H.H., Robert, F., Young, R.A., et Struhl, K. (2003). Targeted recruitment of Set1 histone methylase by elongating Pol II provides a localized mark and memory of recent transcriptional activity. *Mol. Cell* 11, 709-719.

Ni, Z., Schwartz, B.E., Werner, J., Suarez, J.R., et Lis, J.T. (2004). Coordination of transcription, RNA processing, and surveillance by P-TEFb kinase on heat shock genes. *Mol. Cell* 13, 55-65.

Nikolov, D.B., Chen, H., Halay, E.D., Usheva, A.A., Hisatake, K., Lee, D.K., Roeder, R.G., et Burley, S.K. (1995). Crystal structure of a TFIIB-TBP-TATA-element ternary complex. *Nature* 377, 119-128.

Nissen, R.M. et Yamamoto, K.R. (2000). The glucocorticoid receptor inhibits NFkappaB by interfering with serine-2 phosphorylation of the RNA polymerase II carboxy-terminal domain. *Genes Dev.* 14, 2314-2329.

Nobrega, M.A., Ovcharenko, I., Afzal, V., et Rubin, E.M. (2003). Scanning human gene deserts for long-range enhancers. *Science* 302, 413.

Nonet, M., Sweetser, D., et Young, R.A. (1987). Functional redundancy and structural polymorphism in the large subunit of RNA polymerase II. *Cell* 50, 909-915.

Nonet, M.L. et Young, R.A. (1989). Intragenic and extragenic suppressors of mutations in the heptapeptide repeat domain of *Saccharomyces cerevisiae* RNA polymerase II. *Genetics* 123, 715-724.

Nudler, E., Mustaev, A., Lukhtanov, E., et Goldfarb, A. (1997). The RNA-DNA hybrid maintains the register of transcription by preventing backtracking of RNA polymerase. *Cell* 89, 33-41.

O'Shea-Greenfield, A. et Smale, S.T. (1992). Roles of TATA and initiator elements in determining the start site location and direction of RNA polymerase II transcription. *J. Biol. Chem.* 267, 1391-1402.

Oelgeschlager, T., Tao, Y., Kang, Y.K., et Roeder, R.G. (1998). Transcription activation via enhanced preinitiation complex assembly in a human cell-free system lacking TAFII. *Mol. Cell* 1, 925-931.

Ogbourne, S. et Antalis, T.M. (1998). Transcriptional control and the role of silencers in transcriptional regulation in eukaryotes. *Biochem. J.* 331 (Pt 1), 1-14.

Ogryzko, V.V., Kotani, T., Zhang, X., Schiltz, R.L., Howard, T., Yang, X.J., Howard, B.H., Qin, J., et Nakatani, Y. (1998). Histone-like TAFs within the PCAF histone acetylase complex. *Cell* 94, 35-44.

Ohkuma, Y., Hashimoto, S., Wang, C.K., Horikoshi, M., et Roeder, R.G. (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. et Roeder, R.G. (1994). Regulation of TFIIH ATPase and kinase activities by TFIIE during active initiation complex formation. *Nature* 368, 160-163.

Ohkuma, Y., Sumimoto, H., Hoffmann, A., Shimasaki, S., Horikoshi, M., et Roeder, R.G. (1991). Structural motifs and potential sigma homologies in the large subunit of human general transcription factor TFIIE. *Nature* 354, 398-401.

Ohkuma, Y., Sumimoto, H., Horikoshi, M., et Roeder, R.G. (1990). Factors involved in specific transcription by mammalian RNA polymerase II: purification and characterization of general transcription factor TFIIE. *Proc. Natl. Acad. Sci. U. S. A.* 87, 9163-9167.

Ohler, U., Liao, G.C., Niemann, H., et Rubin, G.M. (2002). Computational analysis of core promoters in the Drosophila genome. *Genome Biol* 3, RESEARCH0087.

Okamoto, T., Yamamoto, S., Watanabe, Y., Ohta, T., Hanaoka, F., Roeder, R.G., et Ohkuma, Y. (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., Watanabe, Y., Okamura, H., Hanaoka, F., Ohkuma, Y., et Nishimura, Y. (2000). Structure of the central core domain of TFIIEbeta with a novel double-stranded DNA-binding surface. *EMBO J.* 19, 1346-1356.

Opalka, N., Chlenov, M., Chacon, P., Rice, W.J., Wriggers, W., et Darst, S.A. (2003). Structure and function of the transcription elongation factor GreB bound to bacterial RNA polymerase. *Cell* 114, 335-345.

Orlicky, S.M., Tran, P.T., Sayre, M.H., et Edwards, A.M. (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.

Orphanides,G., LeRoy,G., Chang,C.H., Luse,D.S., et Reinberg,D. (1998). FACT, a factor that facilitates transcript elongation through nucleosomes. *Cell* 92, 105-116.

Orphanides,G., Wu,W.H., Lane,W.S., Hampsey,M., et Reinberg,D. (1999). The chromatin-specific transcription elongation factor FACT comprises human SPT16 and SSRP1 proteins. *Nature* 400, 284-288.

Osheim,Y.N., Proudfoot,N.J., et Beyer,A.L. (1999). EM visualization of transcription by RNA polymerase II: downstream termination requires a poly(A) signal but not transcript cleavage. *Mol. Cell* 3, 379-387.

Ozer,J., Moore,P.A., Bolden,A.H., Lee,A., Rosen,C.A., et Lieberman,P.M. (1994). Molecular cloning of the small (γ) subunit of human TFIIA reveals functions critical for activated transcription. *Genes Dev.* 8, 2324-2335.

Pan,G. et Greenblatt,J. (1994). Initiation of transcription by RNA polymerase II is limited by melting of the promoter DNA in the region immediately upstream of the initiation site. *J. Biol. Chem.* 269, 30101-30104.

Pardee,T.S., Bangur,C.S., et Ponticelli,A.S. (1998). The N-terminal region of yeast TFIIB contains two adjacent functional domains involved in stable RNA polymerase II binding and transcription start site selection. *J. Biol. Chem.* 273, 17859-17864.

Park,J.M., Gim,B.S., Kim,J.M., Yoon,J.H., Kim,H.S., Kang,J.G., et Kim,Y.J. (2001). *Drosophila* Mediator complex is broadly utilized by diverse gene-specific transcription factors at different types of core promoters. *Mol. Cell Biol.* 21, 2312-2323.

Parvin,J.D., Shykind,B.M., Meyers,R.E., Kim,J., et Sharp,P.A. (1994). Multiple sets of basal factors initiate transcription by RNA polymerase II. *J. Biol. Chem.* 269, 18414-18421.

Patikoglou,G.A., Kim,J.L., Sun,L., Yang,S.H., Kodadek,T., et Burley,S.K. (1999). TATA element recognition by the TATA box-binding protein has been conserved throughout evolution. *Genes Dev.* 13, 3217-3230.

Persengiev,S.P., Zhu,X., Dixit,B.L., Maston,G.A., Kittler,E.L., et Green,M.R. (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,M.G., Inostroza,J., Maxon,M.E., Flores,O., Admon,A., Reinberg,D., et Tjian,R. (1991). Structure and functional properties of human general transcription factor IIE. *Nature* 354, 369-373.

Pinto,I., Ware,D.E., et Hampsey,M. (1992). The yeast SUA7 gene encodes a homolog of human transcription factor TFIIB and is required for normal start site selection in vivo. *Cell* 68, 977-988.

- Pinto, I., Wu, W.H., Na, J.G., et Hampsey, M. (1994). Characterization of sua7 mutations defines a domain of TFIIB involved in transcription start site selection in yeast. *J. Biol. Chem.* *269*, 30569-30573.
- Pokholok, D.K., Hannett, N.M., et Young, R.A. (2002). Exchange of RNA polymerase II initiation and elongation factors during gene expression in vivo. *Mol Cell* *9*, 799-809.
- Powell, W. et Reines, D. (1996). Mutations in the second largest subunit of RNA polymerase II cause 6-azauracil sensitivity in yeast and increased transcriptional arrest in vitro. *J. Biol. Chem.* *271*, 6866-6873.
- Price, D.H. (2000). P-TEFb, a cyclin-dependent kinase controlling elongation by RNA polymerase II. *Mol. Cell Biol.* *20*, 2629-2634.
- Price, D.H., Sluder, A.E., et Greenleaf, A.L. (1989). Dynamic interaction between a Drosophila transcription factor and RNA polymerase II. *Mol. Cell Biol.* *9*, 1465-1475.
- Proudfoot, N.J., Furger, A., et Dye, M.J. (2002). Integrating mRNA processing with transcription. *Cell* *108*, 501-512.
- Puig, O., Caspary, F., Rigaut, G., Rutz, B., Bouveret, E., Bragado-Nilsson, E., Wilm, M., et Seraphin, B. (2001). The tandem affinity purification (TAP) method: a general procedure of protein complex purification. *Methods* *24*, 218-229.
- Qureshi, S.A. et Jackson, S.P. (1998). Sequence-specific DNA binding by the *S. shibatae* TFIIB homolog, TFB, and its effect on promoter strength. *Mol. Cell* *1*, 389-400.
- Rabenstein, M.D., Zhou, S., Lis, J.T., et Tjian, R. (1999). TATA box-binding protein (TBP)-related factor 2 (TRF2), a third member of the TBP family. *Proc. Natl. Acad. Sci. U. S. A* *96*, 4791-4796.
- Rachez, C., Lemon, B.D., Suldan, Z., Bromleigh, V., Gamble, M., Naar, A.M., Erdjument-Bromage, H., Tempst, P., et Freedman, L.P. (1999). Ligand-dependent transcription activation by nuclear receptors requires the DRIP complex. *Nature* *398*, 824-828.
- Rachez, C., Suldan, Z., Ward, J., Chang, C.P., Burakov, D., Erdjument-Bromage, H., Tempst, P., et Freedman, L.P. (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.
- Ranish, J.A., Hahn, S., Lu, Y., Yi, E.C., Li, X.J., Eng, J., et Aebersold, R. (2004). Identification of TFB5, a new component of general transcription and DNA repair factor IIIH. *Nat. Genet.* *36*, 707-713.
- Ranish, J.A., Lane, W.S., et Hahn, S. (1992). Isolation of two genes that encode subunits of the yeast transcription factor IIA. *Science* *255*, 1127-1129.
- Reese, J.C. (2003). Basal transcription factors. *Curr. Opin. Genet. Dev.* *13*, 114-118.

- Reinberg,D. et Roeder,R.G. (1987). Factors involved in specific transcription by mammalian RNA polymerase II. Transcription factor IIS stimulates elongation of RNA chains. *J. Biol. Chem.* *262*, 3331-3337.
- Reines,D. (1992). Elongation factor-dependent transcript shortening by template-engaged RNA polymerase II. *J. Biol. Chem.* *267*, 3795-3800.
- Reines,D., Chamberlin,M.J., et Kane,C.M. (1989). Transcription elongation factor SII (TFIIS) enables RNA polymerase II to elongate through a block to transcription in a human gene in vitro. *J. Biol. Chem.* *264*, 10799-10809.
- Reines,D., Ghanouni,P., Li,Q.Q., et Mote,J., Jr. (1992). The RNA polymerase II elongation complex. Factor-dependent transcription elongation involves nascent RNA cleavage. *J. Biol. Chem.* *267*, 15516-15522.
- Remenyi,A., Scholer,H.R., et Wilmanns,M. (2004). Combinatorial control of gene expression. *Nat. Struct. Mol Biol* *11*, 812-815.
- Ren,D., Lei,L., et Burton,Z.F. (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.
- Rigaut,G., Shevchenko,A., Rutz,B., Wilm,M., Mann,M., et Seraphin,B. (1999). A generic protein purification method for protein complex characterization and proteome exploration. *Nat. Biotechnol.* *17*, 1030-1032.
- Riles,L., Shaw,R.J., Johnston,M., et Reines,D. (2004). Large-scale screening of yeast mutants for sensitivity to the IMP dehydrogenase inhibitor 6-azauracil. *Yeast* *21*, 241-248.
- Rivetti,C., Guthold,M., et Bustamante,C. (1999). Wrapping of DNA around the E.coli RNA polymerase open promoter complex. *EMBO J.* *18*, 4464-4475.
- Robert,F., Blanchette,M., Maes,O., Chabot,B., et Coulombe,B. (2002). A human RNA polymerase II-containing complex associated with factors necessary for spliceosome assembly. *J. Biol. Chem.* *277*, 9302-9306.
- Robert,F., Douziech,M., Forget,D., Egly,J.M., Greenblatt,J., Burton,Z.F., et Coulombe,B. (1998). Wrapping of promoter DNA around the RNA polymerase II initiation complex induced by TFIIF. *Mol. Cell* *2*, 341-351.
- Robert,F., Forget,D., Li,J., Greenblatt,J., et Coulombe,B. (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.
- Robert,F., Pokholok,D.K., Hannett,N.M., Rinaldi,N.J., Chandy,M., Rolfe,A., Workman,J.L., Gifford,D.K., et Young,R.A. (2004). Global Position and Recruitment of HATs and HDACs in the Yeast Genome. *Mol Cell* *16*, 199-209.

- Roberts,S.G. et Green,M.R. (1994). Activator-induced conformational change in general transcription factor TFIIB. *Nature* 371, 717-720.
- Rondon,A.G., Gallardo,M., Garcia-Rubio,M., et Aguilera,A. (2004). Molecular evidence indicating that the yeast PAF complex is required for transcription elongation. *EMBO Rep.* 5, 47-53.
- Rossignol,M., Kolb-Cheynel,I., et Egly,J.M. (1997). Substrate specificity of the cdk-activating kinase (CAK) is altered upon association with TFIID. *EMBO J.* 16, 1628-1637.
- Rozenfeld,S. et Thuriaux,P. (2001). A genetic look at the active site of RNA polymerase III. *EMBO Rep.* 2, 598-603.
- Rudd,M.D. et Luse,D.S. (1996). Amanitin greatly reduces the rate of transcription by RNA polymerase II ternary complexes but fails to inhibit some transcript cleavage modes. *J. Biol Chem.* 271, 21549-21558.
- Ryu,S., Zhou,S., Ladurner,A.G., et Tjian,R. (1999). The transcriptional cofactor complex CRSP is required for activity of the enhancer-binding protein Sp1. *Nature* 397, 446-450.
- Sagitov,V., Nikiforov,V., et Goldfarb,A. (1993). Dominant lethal mutations near the 5' substrate binding site affect RNA polymerase propagation. *J. Biol Chem.* 268, 2195-2202.
- Saha,A., Wittmeyer,J., et Cairns,B.R. (2002). Chromatin remodeling by RSC involves ATP-dependent DNA translocation. *Genes Dev.* 16, 2120-2134.
- Sakurai,H., Ohishi,T., et Fukasawa,T. (1997). Promoter structure-dependent functioning of the general transcription factor IIE in *Saccharomyces cerevisiae*. *J. Biol. Chem.* 272, 15936-15942.
- Samuelson,C.O., Baraznenok,V., Khorosjutina,O., Spahr,H., Kieselbach,T., Holmberg,S., et Gustafsson,C.M. (2003). TRAP230/ARC240 and TRAP240/ARC250 Mediator subunits are functionally conserved through evolution. *Proc. Natl. Acad. Sci. U. S. A* 100, 6422-6427.
- Sanders,S.L. et Weil,P.A. (2000). Identification of two novel TAF subunits of the yeast *Saccharomyces cerevisiae* TFIID complex. *J. Biol. Chem.* 275, 13895-13900.
- Santos-Rosa,H., Schneider,R., Bernstein,B.E., Karabetsou,N., Morillon,A., Weise,C., Schreiber,S.L., Mellor,J., et Kouzarides,T. (2003). Methylation of histone H3 K4 mediates association of the Isw1p ATPase with chromatin. *Mol Cell* 12, 1325-1332.
- Sato,S., Tomomori-Sato,C., Banks,C.A., Parmely,T.J., Sorokina,I., Brower,C.S., Conaway,R.C., et Conaway,J.W. (2003). A mammalian homolog of *Drosophila melanogaster* transcriptional coactivator intersex is a subunit of the mammalian Mediator complex. *J. Biol Chem.* 278, 49671-49674.

Sato,S., Tomomori-Sato,C., Parmely,T.J., Florens,L., Zybaylov,B., Swanson,S.K., Banks,C.A., Jin,J., Cai,Y., Washburn,M.P., Conaway,J.W., et Conaway,R.C. (2004). A set of consensus mammalian mediator subunits identified by multidimensional protein identification technology. *Mol Cell* *14*, 685-691.

Sauer,F., Hansen,S.K., et Tjian,R. (1995a). DNA template and activator-coactivator requirements for transcriptional synergism by *Drosophila bicoid*. *Science* *270*, 1825-1828.

Sauer,F., Hansen,S.K., et Tjian,R. (1995b). Multiple TAFIIs directing synergistic activation of transcription. *Science* *270*, 1783-1788.

Sauer,F. et Tjian,R. (1997). Mechanisms of transcriptional activation: differences and similarities between yeast, *Drosophila*, and man. *Curr. Opin. Genet. Dev.* *7*, 176-181.

Scafe,C., Martin,C., Nonet,M., Podos,S., Okamura,S., et Young,R.A. (1990a). Conditional mutations occur predominantly in highly conserved residues of RNA polymerase II subunits. *Mol. Cell Biol.* *10*, 1270-1275.

Scafe,C., Nonet,M., et Young,R.A. (1990b). RNA polymerase II mutants defective in transcription of a subset of genes. *Mol Cell Biol* *10*, 1010-1016.

Schaeffer,L., Moncollin,V., Roy,R., Staub,A., Mezzina,M., Sarasin,A., Weeda,G., Hoeijmakers,J.H., et Egly,J.M. (1994). The ERCC2/DNA repair protein is associated with the class II BTF2/TFIIH transcription factor. *EMBO J.* *13*, 2388-2392.

Schaeffer,L., Roy,R., Humbert,S., Moncollin,V., Vermeulen,W., Hoeijmakers,J.H., Chambon,P., et Egly,J.M. (1993). DNA repair helicase: a component of BTF2 (TFIIH) basic transcription factor. *Science* *260*, 58-63.

Schwabish,M.A. et Struhl,K. (2004). Evidence for Eviction and Rapid Deposition of Histones upon Transcriptional Elongation by RNA Polymerase II. *Mol Cell Biol* *24*, 10111-10117.

Selleck,W., Howley,R., Fang,Q., Podolny,V., Fried,M.G., Buratowski,S., et Tan,S. (2001). A histone fold TAF octamer within the yeast TFIID transcriptional coactivator. *Nat. Struct. Biol.* *8*, 695-700.

Severinov,K., Markov,D., Severinova,E., Nikiforov,V., Landick,R., Darst,S.A., et Goldfarb,A. (1995). Streptolydigin-resistant mutants in an evolutionarily conserved region of the beta' subunit of *Escherichia coli* RNA polymerase. *J. Biol Chem.* *270*, 23926-23929.

Sheffer,A., Varon,M., et Choder,M. (1999). Rpb7 can interact with RNA polymerase II and support transcription during some stresses independently of Rpb4. *Mol. Cell Biol.* *19*, 2672-2680.

Shi,X., Chang,M., Wolf,A.J., Chang,C.H., Frazer-Abel,A.A., Wade,P.A., Burton,Z.F., et Jaehning,J.A. (1997). Cdc73p and Paf1p are found in a novel RNA polymerase II-containing complex distinct from the Srbp-containing holoenzyme. *Mol. Cell Biol.* *17*, 1160-1169.

Shi,Y., Lan,F., Matson,C., Mulligan,P., Whetstine,J.R., Cole,P.A., Casero,R.A., et Shi,Y. (2004). Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. *Cell* *119*, 941-953.

Shiio,Y. et Eisenman,R.N. (2003). Histone sumoylation is associated with transcriptional repression. *Proc. Natl. Acad. Sci. U. S. A* *100*, 13225-13230.

Shilatifard,A., Conaway,R.C., et Conaway,J.W. (2003). The RNA polymerase II elongation complex. *Annu. Rev. Biochem.* *72*, 693-715.

Shilatifard,A., Lane,W.S., Jackson,K.W., Conaway,R.C., et Conaway,J.W. (1996). An RNA polymerase II elongation factor encoded by the human ELL gene. *Science* *271*, 1873-1876.

Shim,E.Y., Walker,A.K., Shi,Y., et Blackwell,T.K. (2002). CDK-9/cyclin T (P-TEFb) is required in two postinitiation pathways for transcription in the *C. elegans* embryo. *Genes Dev.* *16*, 2135-2146.

Sims,R.J., III, Belotserkovskaya,R., et Reinberg,D. (2004). Elongation by RNA polymerase II: the short and long of it. *Genes Dev.* *18*, 2437-2468.

Smale,S.T. et Kadonaga,J.T. (2003). The RNA polymerase II core promoter. *Annu. Rev. Biochem.* *72*, 449-479.

Song,W., Treich,I., Qian,N., Kuchin,S., et Carlson,M. (1996). SSN genes that affect transcriptional repression in *Saccharomyces cerevisiae* encode SIN4, ROX3, and SRB proteins associated with RNA polymerase II. *Mol. Cell Biol.* *16*, 115-120.

Sopta,M., Burton,Z.F., et Greenblatt,J. (1989). Structure and associated DNA-helicase activity of a general transcription initiation factor that binds to RNA polymerase II. *Nature* *341*, 410-414.

Sopta,M., Carthew,R.W., et Greenblatt,J. (1985). Isolation of three proteins that bind to mammalian RNA polymerase II. *J. Biol. Chem.* *260*, 10353-10360.

Sosunov,V., Sosunova,E., Mustaev,A., Bass,I., Nikiforov,V., et Goldfarb,A. (2003). Unified two-metal mechanism of RNA synthesis and degradation by RNA polymerase. *EMBO J.* *22*, 2234-2244.

Sosunova,E., Sosunov,V., Kozlov,M., Nikiforov,V., Goldfarb,A., et Mustaev,A. (2003). Donation of catalytic residues to RNA polymerase active center by transcription factor Gre. *Proc. Natl. Acad. Sci. U. S. A* *100*, 15469-15474.

Squazzo,S.L., Costa,P.J., Lindstrom,D.L., Kumer,K.E., Simic,R., Jennings,J.L., Link,A.J., Arndt,K.M., et Hartzog,G.A. (2002). The Paf1 complex physically and functionally associates with transcription elongation factors in vivo. *EMBO J.* 21, 1764-1774.

Stargell,L.A., Moqtaderi,Z., Dorris,D.R., Ogg,R.C., et Struhl,K. (2000). TFIIA has activator-dependent and core promoter functions in vivo. *J. Biol. Chem.* 275, 12374-12380.

Stargell,L.A. et Struhl,K. (1995). The TBP-TFIIA interaction in the response to acidic activators in vivo. *Science* 269, 75-78.

Steitz,T.A. (1998). A mechanism for all polymerases. *Nature* 391, 231-232.

Strahl,B.D. et Allis,C.D. (2000). The language of covalent histone modifications. *Nature* 403, 41-45.

Struhl,K. (1989). Molecular mechanisms of transcriptional regulation in yeast. *Annu. Rev. Biochem.* 58, 1051-1077.

Struhl,K. (1995). Yeast transcriptional regulatory mechanisms. *Annu. Rev. Genet.* 29, 651-674.

Sumimoto,H., Ohkuma,Y., Sinn,E., Kato,H., Shimasaki,S., Horikoshi,M., et Roeder,R.G. (1991). Conserved sequence motifs in the small subunit of human general transcription factor TFIIE. *Nature* 354, 401-404.

Sun,X., Ma,D., Sheldon,M., Yeung,K., et Reinberg,D. (1994). Reconstitution of human TFIIA activity from recombinant polypeptides: a role in TFIID-mediated transcription. *Genes Dev.* 8, 2336-2348.

Sun,X., Zhang,Y., Cho,H., Rickert,P., Lees,E., Lane,W., et Reinberg,D. (1998). NAT, a human complex containing Srb polypeptides that functions as a negative regulator of activated transcription. *Mol. Cell* 2, 213-222.

Sun,Z.W. et Hampsey,M. (1995). Identification of the gene (SSU71/TFG1) encoding the largest subunit of transcription factor TFIIF as a suppressor of a TFIIB mutation in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. U. S. A* 92, 3127-3131.

Sun,Z.W., Tessmer,A., et Hampsey,M. (1996). Functional interaction between TFIIB and the Rpb9 (Ssu73) subunit of RNA polymerase II in *Saccharomyces cerevisiae*. *Nucleic Acids Res.* 24, 2560-2566.

Surratt,C.K., Milan,S.C., et Chamberlin,M.J. (1991). Spontaneous cleavage of RNA in ternary complexes of *Escherichia coli* RNA polymerase and its significance for the mechanism of transcription. *Proc. Natl. Acad. Sci. U. S. A* 88, 7983-7987.

Suzuki,Y., Tsunoda,T., Sese,J., Taira,H., Mizushima-Sugano,J., Hata,H., Ota,T., Isogai,T., Tanaka,T., Nakamura,Y., Suyama,A., Sakaki,Y., Morishita,S., Okubo,K., et Sugano,S. (2001). Identification and characterization of the potential promoter regions of 1031 kinds of human genes. *Genome Res.* *11*, 677-684.

Taatjes,D.J., Naar,A.M., Andel,I.F., Nogales,E., et Tjian,R. (2002). Structure, Function, and Activator-Induced Conformations of the CRSP Coactivator. *Science* *295*, 1058-1062.

Takada,S., Lis,J.T., Zhou,S., et Tjian,R. (2000). A TRF1:BRF complex directs *Drosophila* RNA polymerase III transcription. *Cell* *101*, 459-469.

Takahashi,Y., Rayman,J.B., et Dynlacht,B.D. (2000). Analysis of promoter binding by the E2F and pRB families in vivo: distinct E2F proteins mediate activation and repression. *Genes Dev.* *14*, 804-816.

Tan,Q., Prysak,M.H., et Woychik,N.A. (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., Aso,T., Conaway,R.C., et Conaway,J.W. (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., Hunziker,Y., Sargent,D.F., et Richmond,T.J. (1996). Crystal structure of a yeast TFIIA/TBP/DNA complex. *Nature* *381*, 127-151.

Tavormina,P.L., Landick,R., et Gross,C.A. (1996a). Isolation, purification, and in vitro characterization of recessive-lethal-mutant RNA polymerases from *Escherichia coli*. *J. Bacteriol.* *178*, 5263-5271.

Tavormina,P.L., Reznikoff,W.S., et Gross,C.A. (1996b). Identifying interacting regions in the beta subunit of *Escherichia coli* RNA polymerase. *J. Mol Biol* *258*, 213-223.

Teixeira,A., Tahiri-Alaoui,A., West,S., Thomas,B., Ramadass,A., Martianov,I., Dye,M., James,W., Proudfoot,N.J., et Akoulitchev,A. (2004). Autocatalytic RNA cleavage in the human beta-globin pre-mRNA promotes transcription termination. *Nature* *432*, 526-530.

Temiakov,D., Patlan,V., Anikin,M., McAllister,W.T., Yokoyama,S., et Vassilyev,D.G. (2004). Structural basis for substrate selection by t7 RNA polymerase. *Cell* *116*, 381-391.

Thompson,C.M., Koleske,A.J., Chao,D.M., et Young,R.A. (1993). A multisubunit complex associated with the RNA polymerase II CTD and TATA-binding protein in yeast. *Cell* *73*, 1361-1375.

Thompson,C.M. et Young,R.A. (1995). General requirement for RNA polymerase II holoenzymes in vivo. *Proc. Natl. Acad. Sci. U. S. A* *92*, 4587-4590.

Thuillier,V., Brun,I., Sentenac,A., et Werner,M. (1996). Mutations in the alpha-amanitin conserved domain of the largest subunit of yeast RNA polymerase III affect pausing, RNA cleavage and transcriptional transitions. *EMBO J.* *15*, 618-629.

Thut,C.J., Chen,J.L., Klemm,R., et Tjian,R. (1995). p53 transcriptional activation mediated by coactivators TAFII40 and TAFII60. *Science* *267*, 100-104.

Timmers,H.T. (1994). Transcription initiation by RNA polymerase II does not require hydrolysis of the beta-gamma phosphoanhydride bond of ATP. *EMBO J.* *13*, 391-399.

Tirode,F., Busso,D., Coin,F., et Egly,J.M. (1999). Reconstitution of the transcription factor TFIIF: assignment of functions for the three enzymatic subunits, XPB, XPD, and cdk7. *Mol. Cell* *3*, 87-95.

Tokunaga,K., Hirose,S., et Suzuki,Y. (1984). In monkey COS cells only the TATA box and the cap site region are required for faithful and efficient initiation of the fibroin gene transcription. *Nucleic Acids Res.* *12*, 1543-1558.

Tollervey,D. (2004). Molecular biology: Termination by torpedo. *Nature* *432*, 456-457.

Tomilin,A., Remenyi,A., Lins,K., Bak,H., Leidel,S., Vriend,G., Wilmanns,M., et Scholer,H.R. (2000). Synergism with the coactivator OBF-1 (OCA-B, BOB-1) is mediated by a specific POU dimer configuration. *Cell* *103*, 853-864.

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.

Toulokhonov,I., Artsimovitch,I., et Landick,R. (2001). Allosteric control of RNA polymerase by a site that contacts nascent RNA hairpins. *Science* *292*, 730-733.

Toulokhonov,I. et Landick,R. (2003). The flap domain is required for pause RNA hairpin inhibition of catalysis by RNA polymerase and can modulate intrinsic termination. *Mol Cell* *12*, 1125-1136.

Trizeenberg,S.J. (1995). Structure and function of transcriptional activation domains. *Curr. Opin. Genet. Dev.* *5*, 190-196.

Trinh,V., Langelier,M.F., et Coulombe,B. (2005). A repertoire of mutations affecting the function of multisubunit RNA polymerases. Article en préparation.

Usheva,A., Maldonado,E., Goldring,A., Lu,H., Houbavi,C., Reinberg,D., et Aloni,Y. (1992). Specific interaction between the nonphosphorylated form of RNA polymerase II and the TATA-binding protein. *Cell* *69*, 871-881.

Veenstra,G.J., Weeks,D.L., et Wolffe,A.P. (2000). Distinct roles for TBP and TBP-like factor in early embryonic gene transcription in *Xenopus*. *Science* *290*, 2312-2315.

Verma, I.M., Stevenson, J.K., Schwarz, E.M., Van Antwerp, D., et Miyamoto, S. (1995). Rel/NF-kappa B/I kappa B family: intimate tales of association and dissociation. *Genes Dev.* *9*, 2723-2735.

Verrijzer, C.P., Chen, J.L., Yokomori, K., et Tjian, R. (1995). Binding of TAFs to core elements directs promoter selectivity by RNA polymerase II. *Cell* *81*, 1115-1125.

Wada, T., Takagi, T., Yamaguchi, Y., Ferdous, A., Imai, T., Hirose, S., Sugimoto, S., Yano, K., Hartzog, G.A., Winston, F., Buratowski, S., et Handa, H. (1998a). DSIF, a novel transcription elongation factor that regulates RNA polymerase II processivity, is composed of human Spt4 and Spt5 homologs. *Genes Dev.* *12*, 343-356.

Wada, T., Takagi, T., Yamaguchi, Y., Watanabe, D., et Handa, H. (1998b). Evidence that P-TEFb alleviates the negative effect of DSIF on RNA polymerase II-dependent transcription in vitro. *EMBO J.* *17*, 7395-7403.

Wahle, E. et Ruegsegger, U. (1999). 3'-End processing of pre-mRNA in eukaryotes. *FEMS Microbiol. Rev.* *23*, 277-295.

Wang, D. et Hawley, D.K. (1993). Identification of a 3'→5' exonuclease activity associated with human RNA polymerase II. *Proc. Natl. Acad. Sci. U. S. A* *90*, 843-847.

Washburn, M.P., Wolters, D., et Yates, J.R., III (2001). Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nat. Biotechnol.* *19*, 242-247.

Watanabe, T., Hayashi, K., Tanaka, A., Furumoto, T., Hanaoka, F., et Ohkuma, Y. (2003). The carboxy terminus of the small subunit of TFIIE regulates the transition from transcription initiation to elongation by RNA polymerase II. *Mol. Cell Biol.* *23*, 2914-2926.

Weilbaecher, R., Hebron, C., Feng, G., et Landick, R. (1994). Termination-altering amino acid substitutions in the beta' subunit of Escherichia coli RNA polymerase identify regions involved in RNA chain elongation. *Genes Dev.* *8*, 2913-2927.

Weilbaecher, R.G., Awrey, D.E., Edwards, A.M., et Kane, C.M. (2003). Intrinsic transcript cleavage in yeast RNA polymerase II elongation complexes. *J. Biol. Chem.* *278*, 24189-24199.

Weis, L. et Reinberg, D. (1997). Accurate positioning of RNA polymerase II on a natural TATA-less promoter is independent of TATA-binding-protein-associated factors and initiator-binding proteins. *Mol. Cell Biol.* *17*, 2973-2984.

Werner, F. et Weinzierl, R.O. (2002). A Recombinant RNA Polymerase II-like Enzyme Capable of Promoter-Specific Transcription. *Mol. Cell* *10*, 635-646.

- West,S., Gromak,N., et Proudfoot,N.J. (2004). Human 5' --> 3' exonuclease Xrn2 promotes transcription termination at co-transcriptional cleavage sites. *Nature* 432, 522-525.
- Westover,K.D., Bushnell,D.A., et Kornberg,R.D. (2004a). Structural basis of transcription: separation of RNA from DNA by RNA polymerase II. *Science* 303, 1014-1016.
- Westover,K.D., Bushnell,D.A., et Kornberg,R.D. (2004b). Structural Basis of Transcription; Nucleotide Selection by Rotation in the RNA Polymerase II Active Center. *Cell* 119, 481-489.
- Whitehouse,I., Stockdale,C., Flaus,A., Szczelkun,M.D., et Owen-Hughes,T. (2003). Evidence for DNA translocation by the ISWI chromatin-remodeling enzyme. *Mol Cell Biol* 23, 1935-1945.
- Wieczorek,E., Brand,M., Jacq,X., et Tora,L. (1998). Function of TAF(II)-containing complex without TBP in transcription by RNA polymerase II. *Nature* 393, 187-191.
- Williams,L.A. et Kane,C.M. (1996). Isolation and characterization of the *Schizosaccharomyces pombe* gene encoding transcript elongation factor TFIIS. *Yeast* 12, 227-236.
- Wind,M. et Reines,D. (2000). Transcription elongation factor SII. *Bioessays* 22, 327-336.
- Winston,F., Chaleff,D.T., Valent,B., et Fink,G.R. (1984). Mutations affecting Ty-mediated expression of the HIS4 gene of *Saccharomyces cerevisiae*. *Genetics* 107, 179-197.
- Wolner,B.S. et Gralla,J.D. (2001). TATA-flanking sequences influence the rate and stability of TATA-binding protein and TFIIB binding. *J. Biol. Chem.* 276, 6260-6266.
- Wong,J.M. et Bateman,E. (1994). TBP-DNA interactions in the minor groove discriminate between A:T and T:A base pairs. *Nucleic Acids Res.* 22, 1890-1896.
- Woodcock,C.L. et Dimitrov,S. (2001). Higher-order structure of chromatin and chromosomes. *Curr. Opin. Genet. Dev.* 11, 130-135.
- Woychik,N.A. et Young,R.A. (1989). RNA polymerase II subunit RPB4 is essential for high- and low-temperature yeast cell growth. *Mol. Cell Biol.* 9, 2854-2859.
- Wu,S.Y., Kershner,E., et Chiang,C.M. (1998). TAFII-independent activation mediated by human TBP in the presence of the positive cofactor PC4. *EMBO J.* 17, 4478-4490.
- Yamaguchi,Y., Takagi,T., Wada,T., Yano,K., Furuya,A., Sugimoto,S., Hasegawa,J., et Handa,H. (1999). NELF, a multisubunit complex containing RD, cooperates with DSIF to repress RNA polymerase II elongation. *Cell* 97, 41-51.

- Yamaguchi, Y., Wada, T., et Handa, H. (1998). Interplay between positive and negative elongation factors: drawing a new view of DRB. *Genes Cells* 3, 9-15.
- Yamamoto, S., Watanabe, Y., van der Spek, P.J., Watanabe, T., Fujimoto, H., Hanaoka, F., et Ohkuma, Y. (2001). Studies of nematode TFIIE function reveal a link between Ser-5 phosphorylation of RNA polymerase II and the transition from transcription initiation to elongation. *Mol. Cell Biol.* 21, 1-15.
- Yan, Q., Moreland, R.J., Conaway, J.W., et Conaway, R.C. (1999). Dual roles for transcription factor IIF in promoter escape by RNA polymerase II. *J. Biol. Chem.* 274, 35668-35675.
- Yeo, M., Lin, P.S., Dahmus, M.E., et Gill, G.N. (2003). A novel RNA polymerase II C-terminal domain phosphatase that preferentially dephosphorylates serine 5. *J. Biol. Chem.* 278, 26078-26085.
- Yin, Y.W. et Steitz, T.A. (2004). The structural mechanism of translocation and helicase activity in t7 RNA polymerase. *Cell* 116, 393-404.
- Yokomori, K., Verrijzer, C.P., et Tjian, R. (1998). An interplay between TATA box-binding protein and transcription factors IIE and IIA modulates DNA binding and transcription. *Proc. Natl. Acad. Sci. U. S. A* 95, 6722-6727.
- Yokomori, K., Zeidler, M.P., Chen, J.L., Verrijzer, C.P., Mlodzik, M., et Tjian, R. (1994). *Drosophila* TFIIA directs cooperative DNA binding with TBP and mediates transcriptional activation. *Genes Dev.* 8, 2313-2323.
- Yonaha, M., Aso, T., Kobayashi, Y., Vasavada, H., Yasukochi, Y., Weissman, S.M., et Kitajima, S. (1993). Domain structure of a human general transcription initiation factor, TFIIF. *Nucleic Acids Res.* 21, 273-279.
- Young, R.A. (1991). RNA polymerase II. *Annu. Rev. Biochem.* 60, 689-715.
- Yu, J., Li, Y., Ishizuka, T., Guenther, M.G., et Lazar, M.A. (2003). A SANT motif in the SMRT corepressor interprets the histone code and promotes histone deacetylation. *EMBO J.* 22, 3403-3410.
- Zawel, L., Kumar, K.P., et Reinberg, D. (1995). Recycling of the general transcription factors during RNA polymerase II transcription. *Genes Dev.* 9, 1479-1490.
- Zaychikov, E., Martin, E., Denissova, L., Kozlov, M., Markovtsov, V., Kashlev, M., Heumann, H., Nikiforov, V., Goldfarb, A., et Mustaev, A. (1996). Mapping of catalytic residues in the RNA polymerase active center. *Science* 273, 107-109.
- Zhang, C. et Burton, Z.F. (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., Yan,H., et Burton,Z.F. (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,G., Campbell,E.A., Minakhin,L., Richter,C., Severinov,K., et Darst,S.A. (1999a). Crystal structure of *Thermus aquaticus* core RNA polymerase at 3.3 Å resolution. *Cell* *98*, 811-824.

Zhang,H.S., Gavin,M., Dahiya,A., Postigo,A.A., Ma,D., Luo,R.X., Harbour,J.W., et Dean,D.C. (2000). Exit from G1 and S phase of the cell cycle is regulated by repressor complexes containing HDAC-Rb-hSWI/SNF and Rb-hSWI/SNF. *Cell* *101*, 79-89.

Zhang,Y. (2003). Transcriptional regulation by histone ubiquitination and deubiquitination. *Genes Dev.* *17*, 2733-2740.

Zhang,Y., Ng,H.H., Erdjument-Bromage,H., Tempst,P., Bird,A., et Reinberg,D. (1999b). Analysis of the NuRD subunits reveals a histone deacetylase core complex and a connection with DNA methylation. *Genes Dev.* *13*, 1924-1935.

Zhang,Y. et Reinberg,D. (2001). Transcription regulation by histone methylation: interplay between different covalent modifications of the core histone tails. *Genes Dev.* *15*, 2343-2360.

Zhang,Z.J. et Reese,J.C. (2004). Redundant mechanisms are used by Ssn6-Tup1 in repressing chromosomal gene transcription in *Saccharomyces cerevisiae*. *J. Biol. Chem.* *279*, 39240-39250.

Zhao,J., Hyman,L., et Moore,C. (1999). Formation of mRNA 3' ends in eukaryotes: mechanism, regulation, and interrelationships with other steps in mRNA synthesis. *Microbiol. Mol Biol Rev.* *63*, 405-445.

Zhao,X. et Herr,W. (2002). A Regulated Two-Step Mechanism of TBP Binding to DNA. A Solvent-Exposed Surface of TBP Inhibits TATA Box Recognition. *Cell* *108*, 615-627.

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

Ziegler,L.M., Khapersky,D.A., Ammerman,M.L., et Ponticelli,A.S. (2003). Yeast RNA Polymerase II Lacking the Rpb9 Subunit Is Impaired for Interaction with Transcription Factor IIF. *J. Biol. Chem.* *278*, 48950-48956.

Zurita,M. et Merino,C. (2003). The transcriptional complexity of the TFIID complex. *Trends Genet.* *19*, 578-584.

