

Université de Montréal

**Analyse de la localisation génomique et identification de nouvelles
fonctions des sous-unités Rpb4/Rpb7 de l'ARN polymérase II et des
facteurs TFIIF, TFIIIS et UBR5**

**par
Marilena Cojocaru**

**Programme de biochimie
Faculté de médecine**

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

**Juillet, 2010
©, Marilena Cojocaru, 2010**

**Université de Montréal
Faculté des études supérieures et postdoctorales**

Cette thèse intitulée:

**Analyse de la localisation génomique et identification de nouvelles
fonctions des sous-unités Rpb4/Rpb7 de l'ARN polymérase II et des
facteurs TFIIF, TFIIS et UBR5**

présentée par: Marilena Cojocar

a été évaluée par un jury composé des personnes suivantes:

**Dr Muriel Aubry
président-rapporteur**

**Dr Benoit Coulombe
directeur de recherche**

**Dr François Robert
co-directeur de recherche**

**Dr Éric Lécuyer
membre du jury**

**Dr Luc Gaudreau
examineur externe**

RÉSUMÉ

Grâce à un grand nombre d'études biochimiques, génétiques et structurales effectuées dans les dernières années, des avancements considérables ont été réalisés et une nouvelle vision du processus par lequel la machinerie transcriptionnelle de l'ARN polymérase II (Pol II) décode l'information génétique a émergé. De nouveaux indices ont été apportés sur la diversité des mécanismes de régulation de la transcription, ainsi que sur le rôle des facteurs généraux de transcription (GTFs) dans cette diversification. Les travaux présentés dans cette thèse amènent de nouvelles connaissances sur le rôle des GTFs humains dans la régulation des différentes étapes de la transcription.

Dans la première partie de la thèse, nous avons analysé la fonction de la Pol II et des GTFs humains, en examinant de façon systématique leur localisation génomique. Les patrons obtenus par immunoprécipitation de la chromatine (ChIP) des versions de GTFs portant une étiquette TAP (*Tandem-Affinity Purification*) indiquent de nouvelles fonctions *in vivo* pour certains composants de cette machinerie et pour des éléments structuraux de la Pol II. Nos résultats suggèrent que TFIIF et l'hétérodimère Rpb4–Rpb7 ont une fonction spécifique pendant l'étape d'élongation transcriptionnelle *in vivo*. De plus, notre étude amène une première image globale de la fonction des GTFs pendant la réaction transcriptionnelle dans des cellules mammifères vivantes.

Deuxièmement, nous avons identifié une nouvelle fonction de TFIIS dans la régulation de CDK9, la sous-unité kinase du facteur P-TEFb (*Positive Transcription Elongation Factor b*). Nous avons identifié deux nouveaux partenaires d'interaction pour TFIIS, soit CDK9 et la E3 ubiquitine ligase UBR5. Nous montrons que UBR5 catalyse l'ubiquitination de CDK9 *in vitro*. De plus, la polyubiquitination de CDK9 dans des cellules humaines est dépendante de UBR5 et TFIIS. Nous montrons aussi que UBR5, CDK9 and TFIIS co-localisent le long du gène γ fibrinogen (γ FBG) et que la surexpression de TFIIS augmente les niveaux d'occupation par CDK9 de

régions spécifiques de ce gène, de façon dépendante de UBR5. Nous proposons que TFIIS a une nouvelle fonction dans la transition entre les étapes d'initiation et d'élongation transcriptionnelle, en régulant la stabilité des complexes CDK9-Pol II pendant les étapes précoces de la transcription.

Mots clés : immuno-précipitation de la chromatine, facteurs généraux de transcription, tandem-affinity purification, ARN polymérase II, hétérodimère Rpb4-Rpb7, facteur de transcription IIF (TFIIF), facteur de transcription IIS (TFIIS), ubiquitine ligase UBR5, Positive-Transcription Elongation Factor b (P-TEFb), ubiquitination de CDK9.

ABSTRACT

Biochemical, genetic and structural studies made over the last years bring a new view on the RNA polymerase II (Pol II) machinery and the process by which it decodes the genetic information. They provided new insights into the diversity of the transcriptional regulation mechanisms, and on the role played by the general transcription factors (GTFs). The studies presented in this thesis provide new evidence on the role of human GTFs in the regulation of different stages of transcription.

In the first part of the thesis, we investigated the function of the human Pol II and GTFs in living cells, by systematically analyzing their genomic location. The location profiles obtained by chromatin immunoprecipitation (ChIP) of TAP (tandem-affinity purification) tagged versions of these factors indicate new *in vivo* functions for several components of this machinery, and for structural elements of the Pol II. These results suggest that TFIIF and the heterodimer Rpb4–Rpb7 have a specific function during the elongation stage *in vivo*. Additionally, our study offers for the first time a general picture of GTFs function during the Pol II transcription reaction in live mammalian cells, and provides a framework to uncover new regulatory hubs.

Secondly, we report on the identification of a new function of the factor TFIIS in the regulation of CDK9, the kinase subunit of the Positive Transcription Elongation Factor b (P-TEFb). We identify two interaction partners for TFIIS, namely CDK9 and the E3 ubiquitin ligase UBR5. We show that UBR5 catalyzes the ubiquitination of CDK9 *in vitro*. Moreover, the polyubiquitination of CDK9 in human cells is dependent upon both UBR5 and TFIIS, and does not signal its degradation. We also

show that UBR5, CDK9 and TFIIS co-localize along specific regions of the γ fibrinogen (γ FBG) gene, and that the overexpression of TFIIS increases the occupancy of CDK9 along this gene in a UBR5 dependant manner. We propose a new function of TFIIS in the transition between initiation and elongation stages, by regulating the stability of the early CDK9-Pol II transcribing complexes.

Key words: chromatin immunoprecipitation, general transcription factors, tandem-affinity purification, RNA polymerase II, Rpb4–Rpb7 heterodimer, transcription factor IIF (TFIIF), transcription factor IIS (TFIIS), UBR5 ubiquitin ligase, Positive Transcription Elongation Factor b (P-TEFb), CDK9 ubiquitination.

TABLE DE MATIÈRES

Résumé	iii
Abstract	v
Table des matières	vii
Liste des figures	x
Liste des abréviations	xv
Dédicace	xviii
Remerciements	xix
Chapitre 1	
Introduction	1
1.1 Considérations générales.....	2
1.2 L'ARN polymerase II.....	3
1.2.1 Le mécanisme catalytique et les éléments structuraux de l'ARN Pol II.....	5
1.2.2 Structure de l'ARN Pol II en complexe avec des facteurs généraux de transcription.....	15
1.3 Les étapes de la transcription et les facteurs généraux de transcription.....	19
1.3.1 L'initiation de la transcription et les facteurs généraux de transcription impliqués	20
1.3.1.1 TFIID.....	24
1.3.1.2 TFIIB.....	30
1.3.1.3 TFIIF.....	32
1.3.1.4 TFIIE.....	34
1.3.1.5 TFIIH.....	36
1.3.1.6 TFIIA.....	39

1.3.1.7 Le complexe Médiateur.....	41
1.3.2 L'élongation.....	48
1.3.2.1 Les facteurs qui modulent les pauses de l'ARN Pol II.....	53
1.3.2.1.1 TFIIF.....	53
1.3.2.1.2 P-TEFb.....	54
1.3.2.2 Les facteurs qui régulent la polymérase arrêtée.....	57
1.3.2.2.1 TFIIS.....	59
1.3.3 La terminaison.....	65
1.4 La régulation de la transcription.....	67
1.4.1 La chromatine.....	67
1.4.2 Les activateurs et les répresseurs de la transcription.....	71
1.4.3 Les modifications post-traductionnelles des facteurs impliqués dans la transcription.....	75
1.4.3.1 L'ubiquitination et la E3 ubiquitine ligase UBR5	77
1.4.4 D'autres mécanismes de régulation de la transcription.....	81
1.5 Objectifs de recherche	83

Chapitre 2

LOCALISATION GÉNOMIQUE DE LA MACHINERIE TRANSCRIPTIONNELLE DE L'ARN POL II: IDENTIFICATION DU ROLE DE TFIIF ET RPB7 DANS LES ÉTAPES INITIALES ET TARDIVES DE LA TRANSCRIPTION

2.1 Préambule.....	85
2.2 Article	86

Chapitre 3

RPAPI, UNE NOUVELLE PROTÉINE ASSOCIÉE À L'ARN POL II PURIFIÉE À PARTIR DE SOUS-UNITÉS RECOMBINANTES DE TYPE SAUVAGE ET MUTANT

3.1 Préambule.....	127
3.2 Article	128

Chapitre 4

LE FACTEUR TFIIS COOPÈRE AVEC LA LIGASE DE TYPE E3 UBR5 AFIN D'UBIQUITINER LA SOUS-UNITÉ CDK9 DU FACTEUR *POSITIVE-TRANSCRIPTION ELONGATION FACTOR b*

4.1 Préambule.....	185
4.2 Article	186

Chapitre 5

Discussion

5.1 Localisation génomique de la machinerie transcriptionnelle de l'ARN Pol II.....	224
5.2 Le facteur TFIIS et ses nouveaux partenaires d'interaction.....	234
5.2 Identification de nouvelles fonctions pour le facteur TFIIS et pour l'ubiquitine ligase UBR5.....	235

Chapitre 6

Conclusions et perspectives.....241

Références.....244

Annexe I.....302

LISTE DES FIGURES

CHAPITRE 1

1. Modèle du mécanisme d'incorporation des nucléotides dans le transcrit d'ARN.....8
2. Représentation des différents éléments structuraux localisés près du site catalytique de l'ARN polymérase II en élongation.....10
3. Modèle structural de l'ARN polymérase II complète.....11
4. Les modifications post-traductionnelles du domaine C-terminal (CTD) de l'ARN polymérase II et son rôle pendant les étapes de la transcription.....14
5. Représentation des domaines structuraux de TFIIB.....16
6. Modèle du complexe ARN polymérase II-TFIIS.....18
7. Les principales étapes de la transcription par l'ARN polymérase II.....20
8. Représentation schématique de l'organisation structurale de TFIID chez la levure.....26
9. Modèle structural du complexe Médiateur-ARN Polymérase II-ADN.....45
10. Modèle pour l'équilibre entre l'élongation et les processus alternatifs.....50

11. Les domaines fonctionnels de TFIIS et l'alignement des séquences des homologues chez l'humain et chez la levure	60
12. Modèle du mécanisme par lequel les facteurs négatifs d'élongation DSIF et NELF inhibent TFIIS.....	64
13. Modèle de la régulation par le positionnement des nucléosomes des deux classes des gènes, de maintenance ou régulés.....	70
14. Représentation schématique des domaines fonctionnels de UBR5.....	80

CHAPITRE 2

1. TAP of the RNAPII general transcription machinery.....	119
2. Occupancy of the regions proximal to the TIS and further downstream TRs of selected human genes by GTF and RNAPII subunits obtained by ChIP.....	120
3. Occupancy of the regions proximal to the TIS and further downstream TRs of selected human genes by the TFIIF subunit Rap30 and the RNAPII subunit Rpb11 as determined by ChIP.....	121
4. Occupancy of the regions proximal to the TIS and further downstream TRs of selected human genes by Rpb7 as determined by ChIP.....	122

Figures supplémentaires

1. Network of protein complexes involving the RNAPII general transcription machinery.....124
2. Genomic position of PCR products.....125
3. Occupancy of (i) regions proximal to the transcriptional initiation site (TIS) and (ii) further downstream transcribed regions (TRs) by subunits of RNAPII and GTFs analysed by chromatin immunoprecipitation in order to confirm the results obtained with a first subunit.....126

CHAPITRE 3

1. TAP of human transcription factors.....176
2. Purification of TAP-tagged human RNAPII complex.....177
3. Regions of homology between *Homo sapiens* RPAP1, *D. melanogaster* CG32104-PB, and *S. cerevisiae* YDR527wp/RBA50, as determined by BLAST.....178
4. TAP-tagged human RNAPII assembles on promoter both *in vivo* and *in vitro* and binds to acetylated histones.....179
5. TAP-tagged human RNAPII complex can initiate transcription *in vitro*.....180
6. Purification and functional analysis of TAP-tagged human RNAPII with a mutation in fork loop 1.....181

7. Purification and functional analysis of TAP-tagged human RNAPII with a mutation in switch 3.....183
8. Association of mutants forms of the RNAPII complex with chromatin *in vivo*.....184

CHAPITRE 4

1. TFIIIS interacts with the E3 ligase UBR5 and the P-TEFb kinase CDK9.....217
2. UBR5 ubiquitinates CDK9 *in vitro*.....218
3. TFIIIS and UBR5 overexpression affect the ubiquitination of CDK9 *in vivo*.....219
4. TFIIIS or UBR5 knockdown decreases the ubiquitination of CDK9 *in vivo*, and UBR5 is required for the effect of TFIIIS on this ubiquitination event.....220
5. UBR5 knockdown does not affect the stability of CDK9.....221
6. TFIIIS, CDK9, and UBR5 are recruited to the γ FBG gene, and TFIIIS increases the occupancy of this gene by CDK9 in a UBR5-dependent manner.222

CHAPITRE 5

1. Modèle décrivant le rôle fonctionnel de l'interaction TFIIS-UBR5-CDK9....239

Annexe I

1. Analyse du rôle de Rpb7 dans la formation des ARNs de RPL3, U83b et U43.....302
2. Analyse comparative de TFIIS.1 et TFIIS.....303

LISTE DES ABRÉVIATIONS

6-AU	6-azauracil
ADN	Acide désoxyribonucléique
ARN	Acide ribonucléique
ARNm	ARN messenger
ARNr	ARN ribosomaux
ATP	Adénosine triphosphate
BRE	<i>TFIIB Recognition Element</i>
CAK	<i>Cyclin-dependent kinase Activating Kinase</i>
CBP	<i>Calmodulin-Binding Peptide</i>
Cdc	<i>Cell division cycle</i>
CDK	<i>Cyclin Dependent Kinase</i>
ChIP	Immunoprécipitation de la chromatine
CTD	<i>Carboxy-Terminal Domain</i>
CPSF	<i>Cleavage and Polyadenylation Specificity Factor</i>
dNTP	Désoxyribonucléotide triphosphate
DPE	<i>Downstream Promoter Element</i>
DRB	5,6-dichloro-1-b-D-ribofuranosylbenzimidazole
DSE	<i>Downstream Sequence Element</i>
DSIF	<i>DRB-Sensitivity Inducing Factor</i>
ELL	<i>Eleven-nineteen Lysine-rich in Leukemia</i>
Fcp1	<i>TFIIF-associated CTD phosphatase</i>
GTFs	Facteurs généraux de transcription
HAT	Histone acétyltransférase
HDAC	Histone déacétylase
HECT	Domaine d'homologie à la E6AP C-terminal
HMG	<i>High mobility group</i>
Inr	Initiateur
miARNs	micro ARNs
LCR	<i>Locus Control Region</i>

MAT	Ménage à trois
Med	Médiateur
MTE	<i>Motif Ten Element</i>
NELF	<i>Negative Elongation Factor</i>
NER	<i>Nucleotide Excision Repair</i>
PABP	<i>Poly(A) Binding Protein</i>
pb	Paire de bases
PC	<i>Positive Cofactor</i>
pCAF	<i>p300/CBP-Associated Factor</i>
PIC	<i>Preinitiation Complex</i>
Pol II	ARN polymerase II
p-TEFb	<i>Positive-Transcription Elongation Factor b</i>
RAR	<i>Retinoic Acid Receptor</i>
Rpb	<i>RNA polymerase B</i>
snoARNs	<i>small nucleolar RNAs</i>
RXR	<i>Retinoic X Receptor</i>
SAGA	SPT-ADA-GCN5 acétyltransférase
SCP1	<i>Small CTD-Phosphatase 1</i>
Ser	Sérine
Spt	<i>Suppressor of Ty insertion</i>
Srb	<i>Suppressor of RNA polymerase B</i>
STAGA	<i>Suppressor of Ty insertion-TAF-Gcn5-Acetyltransferase</i>
TBP	<i>TATA-box Binding Protein</i>
TAF	<i>TBP Associated Factor</i>
TAP	<i>Tandem Affinity Purification</i>
TCR	<i>Transcription Coupled Repair</i>
TFTC	<i>TBP-Free TAF-Containing Complex</i>
TRF	<i>TBP Related Factor</i>
UAS	<i>Upstream Activating Sequence</i>
UV	Radiations ultraviolet
VDR	<i>Vitamine D Receptors</i>

XP

Xeroderma pigmentosum

*À mes parents,
Ioana et Tudor*

REMERCIEMENTS

Je remercie mon directeur de recherches, le Dr Benoit Coulombe de m'avoir accueilli dans son laboratoire, pour ses précieux conseils, son support continu, et sa grande disponibilité. Benoit m'a donné la chance de travailler sur un projet de recherche d'un haut intérêt scientifique et son encadrement rigoureux m'a permis de réaliser chaque étape de ce projet.

Je remercie les membres du laboratoire avec lesquels j'ai collaboré et qui m'ont aidé par des conseils techniques et par des discussions scientifiques tout au long de mon doctorat.

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

CHAPITRE I

INTRODUCTION

1.1 Considérations générales

L'ARN Pol II est l'enzyme responsable de la synthèse de tous les ARNm, soit les ARN qui codent pour des protéines. Elle code également pour des petits ARN tels que les miARN (micro ARN), les snoARN ("small nucleolar" ARN) et certains snARN ("small nuclear" ARN) (U1,U2. etc.). Chez les organismes eucaryotes il existe deux autres ARN polymérase, chacune ayant des fonctions spécifiques. L'ARN polymérase I est responsable de la synthèse des grands ARN ribosomiaux (ARNr) tandis que l'ARN Pol III est impliquée dans la synthèse des ARN de transfert (ARNt), de l'ARNr 5S et de certains snARN tels que U6 et 7SK. Chez les organismes procaryotes il existe une seule ARN polymérase qui est responsable de la synthèse de tous les ARN, alors que chez les plantes il existe cinq ARN polymérase. Les ARN polymérase IV et V chez les plantes synthétisent des ARN qui joueraient un rôle dans le processus d'inhibition de l'expression génique par un mécanisme de déméthylation des cytosines dans le contexte de la chromatine (Matzke et al., 2009).

Le processus transcriptionnel a été étudié extensivement : ses étapes spécifiques ont été caractérisées, des systèmes complètement reconstitués avec des facteurs purifiés ont été utilisés pour mettre à l'épreuve notre compréhension à son égard, les fonctions dans ce processus d'un grand nombre de facteurs et complexes protéiques ont été caractérisées et, finalement, le rôle de différents éléments structuraux de l'ARN Pol II a été raffiné par des analyses cristallographiques à haute résolution. Cependant, les études de la transcription dans des cellules vivantes, dans des conditions physiologiques ou pathologiques, amènent continuellement des informations qui montrent que notre compréhension du processus transcriptionnel reste fragmentaire. Des aspects tels que la coordination globale de la transcription en réponse à différents stimuli et l'interconnexion avec d'autres processus biologiques sont encore peu caractérisés.

Plusieurs complexes protéiques ont été impliqués dans la transcription par la L'ARN Pol II dans le contexte des nucléosomes. La machinerie appelée générale est composée de l'ARN Pol II et des facteurs généraux de transcription (GTFs) : TFIIA, TFIIB, TFIID, TFIIE, TFIIF et TFIIH. Elle est nécessaire et suffisante pour la transcription basale à partir de matrices d'ADN purifiées (Roeder, 1996). La transcription dépendante d'activateurs dans des systèmes reconstitués requiert en plus des co-activateurs tel que le complexe Médiateur, qui établissent un pont entre les protéines régulatrices et la machinerie générale de transcription. (Malik and Roeder, 2000). La transcription efficace à travers la barrière des nucléosomes nécessite d'autres facteurs tels que des facteurs de remodelage de la chromatine dépendants de l'ATP (Neely et al., 2002), des facteurs impliqués dans la modification post-traductionnelle des histones par acétylation, méthylation, ubiquitination (Vaquero et al., 2003), etc. Le nombre des facteurs impliqués dans la transcription augmente continuellement et il a été estimé que 6% des gènes qui codent pour des protéines représentent des gènes encodant des facteurs de transcription (Tupler et al., 2001).

1.2 L'ARN Pol II

L'ARN polymérase II est formée de 12 sous-unités appelées Rpb1 à Rpb12. Les sous-unités Rpb4 et Rpb7 forment un hétérodimère qui peut dissocier de la polymérase (Choder, 2004; Orlicky et al., 2001). Les ARN polymerases I et III ont 14 et 17 sous-unités respectivement, les sous-unités supplémentaires étant greffées sur la structure de base à 10 sous-unités. Le site catalytique de l'ARN Pol II est formé par les sous-unités Rpb1 et Rpb2. Ces deux sous-unités sont homologues aux sous-unités β et β' de la polymérase bactérienne. Les sous-unités de l'ARN Pol II sont structurellement et fonctionnellement conservées entre les espèces eucaryotes alors que la polymérase bactérienne possède quelques sous-unités telles que $\alpha_2\beta\beta'$, ω et le facteur σ (Cramer, 2002). Le degré d'homologie de séquence varie en fonction de l'espèce. Par exemple, entre la sous-unité Rpb1 humaine et celle d'autres espèces eucaryotes il y a 50% d'homologie avec Rpb1 de *Plasmodium*

falciparum, 56% avec Rpb1 de *Saccharomyces cerevisiae* et 99.9% avec celle de *Mus musculus* ou de *Bos taurus*. Les sous-unités Rpb5, 6, 8, 10 et 12 sont communes aux trois ARN polymérase I, II et III (Roeder, 1996). Les autres sous-unités qui sont spécifiques à l'ARN Pol II présentent cependant différents degrés d'homologie avec des sous-unités des autres polymérase.

Toutes les sous-unités de l'ARN Pol II sont encodées par des gènes essentiels sauf Rpb4 et Rpb9 (Pillai et al., 2001; Ziegler et al., 2003). Cependant, Rpb4 devient essentiel dans certaines conditions, tels que le stress thermique, une restriction nutritionnelle ou une exposition à l'éthanol (Miyao et al., 2001; Pillai et al., 2003; Woychik and Young, 1989) et le mutant $\Delta rpb4$ chez la levure présente une sensibilité au 6-azauracil (Verma-Gaur et al., 2008). La délétion de Rpb9 quand à elle, confère une sensibilité au 6-azauracil et aux inhibiteurs de l'assemblage des microtubules tels que le nocodazole et le benomyl (Hemming et al., 2000; Malagon et al., 2004). Ces observations suggèrent que malgré que les sous-unités Rpb4 et Rpb9 ne soient pas nécessaires pour la transcription globale en conditions normales, elles jouent des rôles spécifiques dans des conditions déterminées et permettent une régulation de la transcription dans ces conditions. La sensibilité au 6-azauracil est généralement associé à des défauts d'élongation et en effet, dans le cas de Rpb9, une étude par mutagenèse de différents domaines de Rpb9 a montré que cette sous-unité est impliquée dans la régulation de l'élongation (Hemming et al., 2000). Les sous-unités essentielles jouent des rôles fondamentaux pour l'activité transcriptionnelle, par exemple les sous-unités Rpb3 et Rpb11 forment un hétérodimère qui initierait l'assemblage de la polymérase (Kimura and Ishihama, 2000).

Chez la levure toutes les sous-unités sont codées par des gènes uniques. Chez l'humain une diversification de certaines sous-unités a été observée. Par exemple, la sous-unité Rpb11 est codée par une famille de gènes (Grandemange et al., 2001). Les isoformes de Rpb11 présentent des domaines C-terminaux divergents, cependant ceci n'empêche pas la formation du hétérodimère Rpb3-Rpb11. Il a été proposé que cette diversification spécifique à l'humain permet de nouvelles relations entre la transcription et d'autres processus biologiques (Benga et al., 2005).

Les sous-unités Rpb4 et Rpb7 des espèces eucaryotes sont homologues du point de vue structural et fonctionnel avec RpoE et RpoF chez les archaea, ce qui suggère que l'hétérodimère formé par ces deux sous-unités a évolué avant que la séparation des archaea et des eucaryotes ne se produise (Choder, 2004). Des paralogues de ces deux sous-unités ont été identifiés dans la composition des trois types de ARN polymérase eucaryotes, ce qui supporte l'hypothèse qu'ils ont évolués des mêmes molécules ancestrales. Alors que la séquence de Rpb7 est hautement conservée à travers les espèces, celle de Rpb4 est divergente, cette sous-unité étant la moins conservée des sous-unités de la polymérase. En effet, il a été proposé que Rpb4 a servi de véhicule afin d'introduire de nouveaux domaines régulateurs dans la structure de la polymérase à travers l'évolution (Choder, 2004; Siaut et al., 2003).

1.2.1 Le mécanisme catalytique et les éléments structuraux de l'ARN Pol II.

Des études structurales combinées à des essais de transcription *in vitro* ont été utilisés afin de comprendre les différentes étapes de la transcription du point de vue mécanistique. Les premières études cristallographiques de l'ARN Pol II ont été effectuées chez la levure en utilisant l'enzyme formée de dix sous-unités (Fu et al, 1999, Cramer et al, 2001). Cette forme ne contient pas les sous-unités Rpb4/Rpb7 qui sont sous-stoechiométriques comparativement aux autres sous-unités en conditions de croissance exponentielle (Choder, 2004; Rosenheck and Choder, 1998). Dans ces conditions, autour de 20% des molécules d'ARN Pol II contiendraient ces deux sous-unités, alors qu'en phase stationnaire, l'hétérodimère Rpb4/Rpb7 est associé à 100% des molécules d'ARN Pol II. La structure cristallographique de la polymérase complète, avec ses 12 sous-unités, a été obtenue avec des cristaux purifiés des cellules en phase stationnaire après une restriction nutritionnelle (Jensen et al., 1998). Deux autres approches ont été utilisées afin d'obtenir des cristaux de l'ARN Pol II à 12 sous-unités : la première a utilisé une version portant une étiquette TAP (*Tandem Affinity Purification*) de Rpb4 afin de purifier la polymérase (Bushnell and Kornberg, 2003a), alors que la

deuxième a utilisé des cristaux purifiés d'une lignée $\Delta rpb4$ qui ont été incubés par la suite avec les sous-unités Rpb4 et Rpb7 recombinantes (Armache et al., 2003; Armache et al., 2005).

Les études cristallographiques à haute résolution ont révélé des domaines et des surfaces d'interaction, ce qui a permis de prédire des fonctions spécifiques pour ces éléments structuraux. Ainsi, au centre de l'enzyme se trouvent les sous-unités Rpb1 et Rpb2 qui forment un canal chargé positivement, dont le rôle est d'accueillir les ribonucléotides. Au fond de ce canal se trouve le centre catalytique formé par des résidus acides strictement conservés de Rpb1 et Rpb2. Le mécanisme catalytique est commun à tous les polymérases (Steitz, 1998) et implique deux ions métalliques (Mg^{2+}): le métal A est coordonné de façon stable par des résidus aspartate et lie la région 3' de l'ARN, alors que le métal B lie le NTP qui sera incorporé dans la chaîne d'ARN. Afin de décrire le processus d'incorporation des NTPs, deux modèles principaux ont été proposés sur la base des données structurales. Dans le premier modèle le NTP entrant lierait dans un site de pré-insertion de la forme inactive, en conformation ouverte du centre catalytique, ce qui induirait un changement de conformation de la boucle appelée 'trigger' et la fermeture subséquente du centre actif. La liaison du NTP au site d'insertion et la catalyse se produiraient dans cette conformation (Kettenberger et al., 2004). Dans un deuxième modèle, le NTP lierait tout d'abord dans un site d'entrée du pore situé au dessous du centre catalytique et dans une deuxième étape il serait orienté par le métal B directement dans le site d'insertion (Westover et al., 2004b).

Le premier modèle a été renforcé par des études plus récentes de l'ARN Pol II eucaryote (Wang et al., 2006) ou bactérienne (Vassylyev et al., 2007) et le mécanisme proposé explique aussi la sélection du NTP correct (Brueckner et al., 2009). En liant d'abord au site de pré-insertion, les NTPs interagissent avec la matrice d'ADN (Fig. 1). Ceci permettrait de discriminer entre les NTPs entrants, ainsi qu'entre les NTPs et les désoxy-NTPs pour prévenir la synthèse d'ADN. Le changement de conformation de la boucle 'trigger' en position fermée se fait en

parallèle avec la translocation du NTP au site d'insertion et ceci mènerait à une conformation de la polymérase qui serait apte pour la catalyse. L'incorporation du nucléotide correct au site d'insertion (site +1) libère un groupement pyrophosphate. La libération de ce groupement faciliterait le mouvement de la boucle trigger et la boucle de pontage ce qui permettrait que la matrice d'ADN glisse dans la position intermédiaire entre les états de pré- et de post-translocation. Le relâchement et le retour des deux boucles à la conformation initiale droite permettrait le glissement des acides nucléiques dans la position de post-translocation et libérerait le site +1 pour un nouveau cycle d'incorporation de NTPs.

Des études structurales et biochimiques de l'ARN Pol II bactérienne (Bar-Nahum et al., 2005; Tuske et al., 2005) ou eucaryote (Brueckner and Cramer, 2008) en élongation et en complexe avec la drogue α -amanitine ou avec différents antibiotiques viennent à l'appui de ce mécanisme de translocation et montrent qu'il a été conservé à travers l'évolution. Des modifications similaires dans la structure des boucles de pontage et trigger ont été observées chez plusieurs espèces. Ces deux boucles coopèrent et génèrent une structure analogue à une roue à rochet (*ratchet*) qui permet la translocation unidirectionnelle de l'hybride ADN-ARN. L'oscillation Brownienne entre les états de pré- et post-translocation impliquerait un état intermédiaire stabilisé par la α -amanitine dans lequel le NTP lié au site +1 agit comme un cliquet (*pawl*) qui stabiliserait l'état de post-translocation pendant que la boucle trigger s'enfonce et permet le glissage de l'hybride vers cet état (Fig 1). Ce mécanisme est conservé et pourrait être divisé en deux étapes : dans la première la matrice d'ADN en aval glisse jusqu'à ce que le nucléotide suivant (registre +2) arrive au site de pré-insertion et dans la deuxième, ce nucléotide subi une rotation à 90^0 jusqu'à la position d'insertion dans le site +1.

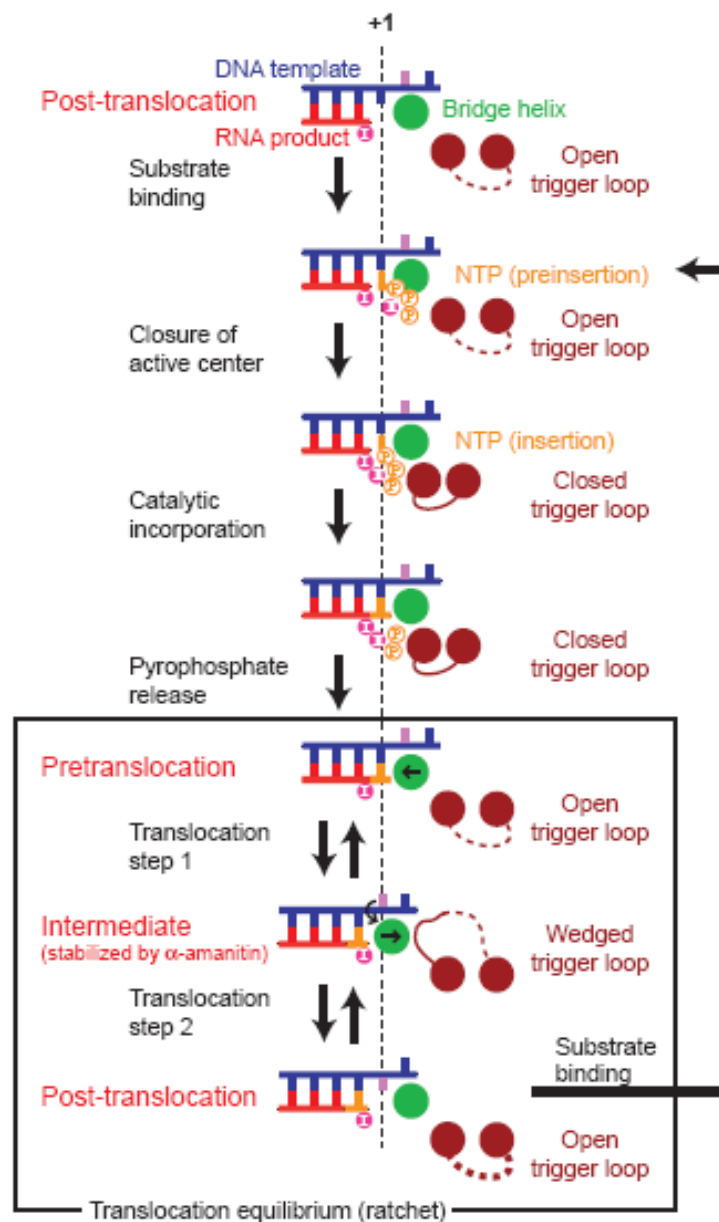


Figure 1. Modèle du mécanisme d'incorporation des nucléotides dans le transcrit d'ARN. Le cadre en bas de l'image représente le mécanisme de roue à rochet (*ratchet*) impliquant l'équilibre entre les états de pré- et post-translocation et une conformation intermédiaire dans laquelle le nucléotide matrice subit une rotation dans le site +1. Tiré de Brueckner and Cramer, 2008.

Des études cristallographiques de l'ARN Pol II en complexe avec des acides nucléiques ont contribué à élucider le mécanisme d'élongation. Des fonctions spécifiques ont été proposées pour plusieurs boucles de Rpb1 ou Rpb2 qui sont situées près du centre catalytique. La première structure cristallographique de l'ARN Pol II en complexe avec des acides nucléiques a utilisé une matrice d'ADN qui, ayant une extension simple-brin en 3', supporte une initiation indépendante du promoteur (Gnatt et al., 2001). Cette analyse a montré le site d'entrée de la partie aval de l'ADN dans le canal principal de la polymérase et a mis en évidence qu'un hybride ADN-ARN de 8 à 9 pb forme la bulle transcriptionnelle au niveau du site catalytique.

Deux modèles ont été suggérés : dans le premier l'ADN est ouvert avant d'entrer dans le canal principal et uniquement le brin matrice entrerait par la suite dans ce canal, alors que dans le deuxième un module mobile de la polymérase appelé "pince" devrait être en position ouverte pour permettre à l'ADN double brin d'entrer dans le canal et atteindre le site catalytique. La "boucle gâchette" ("trigger") a été impliquée dans la translocation des acides nucléiques durant la polymérisation puisque des études cristallographiques du complexe d'élongation ont montré qu'elle adopte deux conformations distinctes, droite ou courbée (Gnatt et al., 2001; Zhang et al., 1999) (Fig. 1). Pour d'autres boucles (Fig. 2), telles que la "boucle d'embranchement" ("fork loop"), le "gouvernail" ("rudder") et le "capuchon" ("hood") il a été suggéré qu'elles contribueraient à la séparation de l'ARN du brin d'ADN matrice, ce qui maintiendrait un hybride ADN-ARN de 8 à 9 pb (Westover et al., 2004a). Les boucles d'embranchement 1 et 2 seraient impliquées dans le maintien de l'ouverture de la bulle transcriptionnelle en aval et dans la restriction de la longueur de l'hybride ADN-ARN, respectivement (Cramer et al., 2001; Gnatt et al., 2001; Westover et al., 2004a). Le maintien de l'ouverture de la bulle transcriptionnelle en amont ainsi que le rôle de guider le transcrit d'ARN vers le canal de sortie est attribué à la boucle appelée "fermeture à glissière" ("zipper") (Cramer et al., 2001; Gnatt et al., 2001).

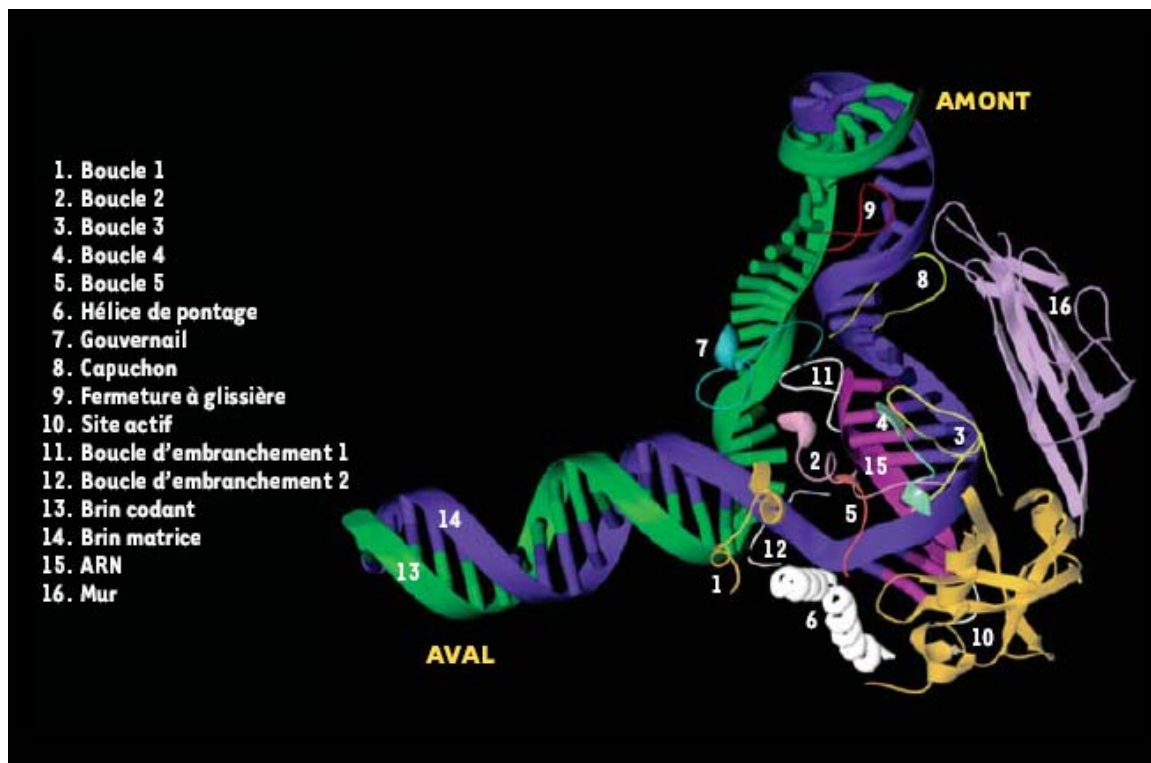


Figure 2. Représentation des différents éléments structuraux localisés près du site catalytique de l'ARN Pol II en élongation. L'hybride ADN-ARN, ainsi que les éléments structuraux de la polymérase sont numérotés et identifiés dans la liste située à gauche. Figure tirée de (Langelier et al., 2002).

Les études cristallographiques à haute résolution de la polymérase complète ont révélé la position exacte des sous-unités Rpb4/Rpb7 (Armache et al., 2003; Armache et al., 2005). Cet hétérodimère forme une structure qui ressort de la surface de la polymérase et est située entre la pince et le domaine C-terminal (CTD) de la sous-unité Rpb1 (Fig. 3). La partie N-terminale de Rpb7 forme une structure appelée “pointe” ou “tip” qui s'insère dans une fente formée à la surface de la polymérase par les sous-unités Rpb1, Rpb6 et Rpb2 (Armache et al., 2003). À partir des études comparatives des structures de la polymérase à 10 sous-unités et celle de la polymérase complète il a été proposé que l'association de Rpb4/Rpb7 à la polymérase n'est possible que lorsque la pince mobile est en position fermée

(Armache et al., 2003; Cramer et al., 2001). À l'inverse, la "pince" ne peut adopter la position ouverte qu'après la dissociation de Rpb4/Rpb7.

L'association de Rpb4/Rpb7 provoquerait plusieurs changements de conformation de la polymérase : les boucles "switch" 1 à 3 situées à la base de la pince deviennent ordonnées, ce qui générerait une surface pour la liaison de l'ADN (Cramer et al., 2001; Gnatt et al., 2001); le domaine appelé "dock" est agrandi ce qui favoriserait l'interaction avec les facteurs d'initiation qui lient ce domaine, tels que TFIIB, TFIIF et TBP (Jensen et al., 1998; Killeen et al., 1992) et, de plus, en liant les trois sous-unités Rpb1, Rpb2 et Rpb6 l'hétérodimère Rpb4/Rpb7 contribuerait à la stabilisation de l'assemblage de la polymérase (Armache et al., 2003). Parmi les caractéristiques communes à tous les orthologues et paralogues de Rpb4/Rpb7, les principales sont leur capacité d'interagir l'un avec l'autre (Sampath et al., 2003; Siaux et al., 2003) ainsi que celle de lier l'ARN (Orlicky et al., 2001). Des analyses récentes par FRET ou par pontage de l'ARN ont montré que le transcrit d'ARN en élongation longe la voie de sortie générée par Rpb4 et Rpb7 (Chen et al., 2009a; Ujvari and Luse, 2006a) et non pas celle de Rpb8 tel que proposé par les études cristallographiques de la polymérase complète.

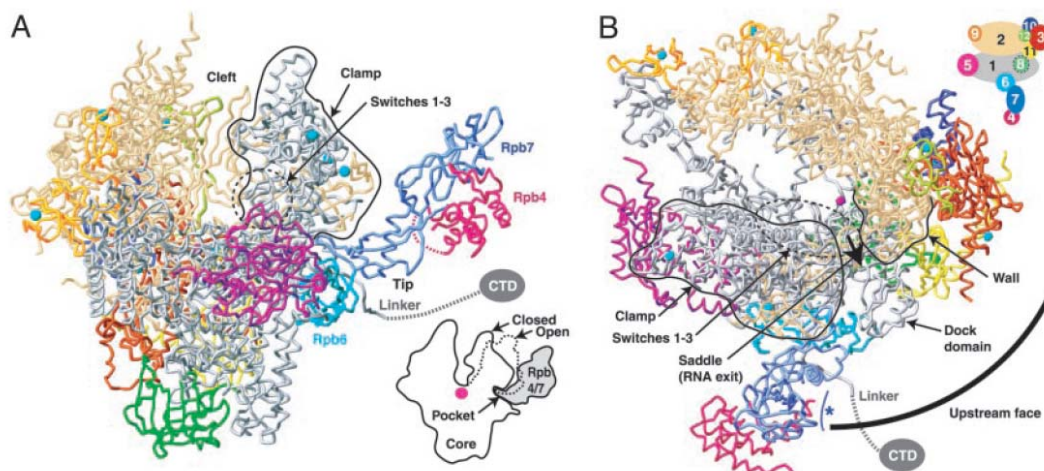


Figure 3. Modèle structural de l'ARN Pol II complète. A) Vue de l'ARN Pol II à 12 sous-unités montrant la pince mobile qui ferme lorsque les sous-unités dissociables

Rpb4/Rpb7 lient la polymérase, mais qui reste en position ouverte en leur absence. Les boucles “switch” 1-3 situées à la base de la pince sont indiquées. Le code des couleurs est celui présenté dans le schéma B. B) Vue amont de l'ARN polymérase complète. Le segment de demi-cercle est centré sur le centre catalytique et il a un rayon équivalent à la distance minimale entre la boîte TATA et le site d'initiation de la transcription (25 pb). Le canal de sortie de l'ARN est indiqué et l'astérisque marque un possible site d'interaction de Rpb7 avec l'ARN. Le code des couleurs pour les sous-unités est indiqué dans le schéma à droite. Tiré de Armache et al, 2003.

Une structure importante de l'ARN Pol II est le domaine C-terminal (CTD) de la sous-unité Rpb1. Ce domaine sert de plateforme pour l'assemblage de différents facteurs pendant le passage de la polymérase à travers les stades de la transcription (Fig. 4) (Egloff and Murphy, 2008a). Le CTD est formé d'un héptapeptide répété un nombre variable de fois en fonction de l'espèce. L'analyse comparative de la séquence du CTD chez différentes espèces a mis en évidence une corrélation entre le nombre de répétitions et la complexité du génome. Ainsi, il y a 26 répétitions de cet heptapeptide chez la levure et 52 chez l'humain. Le patron des modifications post-traductionnelles du CTD est dépendent de l'état de la polymérase et de l'étape de la transcription : l'ARN polymerase II libre, avant d'être recrutée sur les promoteurs, possède un CTD qui est hypophosphorylé mais qui peut être glycosylé (Kelly et al., 1993; Phatnani and Greenleaf, 2006); dans le complexe de préinitiation le CTD est phosphorylé sur la Ser⁵, alors que pendant l'élongation la Ser² est aussi phosphorylée (Phatnani and Greenleaf, 2006). Différentes kinases sont impliquées dans la phosphorylation de ces résidus : la sous-unité CDK7 (*Cycline-Dependent Kinase 7*) de TFIIH est responsable de la phosphorylation de la Ser⁵ ainsi que la Ser⁷ au promoteur (Boeing et al., 2009; Glover-Cutter et al., 2009; Komarnitsky et al., 2000); CDK9, la sous-unité kinase de P-TEFb, phosphoryle la Ser² (Peterlin and Price, 2006), alors que la sous-unité CDK8 du complexe Médiateur mène à la formation d'une polymérase hyperphosphorylée qui serait inactive en transcription (Palancade and Bensaude, 2003). La polymérase est recyclée par

déphosphorylation, différentes phosphatases ciblant des résidus spécifiques. Ainsi, FCP1 (*TFIIF-associated CTD Phosphatase*) est responsable de la déphosphorylation de la Ser² (Meinhart et al., 2005), la Ser⁵ est déphosphorylée par SCP1 (*Small CTD-Phosphatase 1*) (Yeo et al., 2003), par la sous-unité Ssu72 de CPF (*Cleavage/Polyadenylation Factor*) (Krishnamurthy et al., 2004), ainsi que par Rtr1 (Mosley et al., 2009). Aucune phosphatase n'est encore connue pour la Ser⁷.

Une étude récente (Lolli, 2009) a montré que la forme non-phosphorylée du CTD lie l'ADN et que le complexe CTD-ADN est reconnu et phosphorylé par CDK7, mais pas par CDK9. Lorsque phosphorylé par CDK7, le CTD dissocie de l'ADN et permet ainsi l'initiation de la transcription. Le CTD est reconnu par différents facteurs en fonction des résidus qui sont phosphorylés. Par exemple, la forme phosphorylée sur la Ser⁵ est reconnue par les enzymes qui ajoutent une coiffe à l'extrémité 5' de l'ARNm synthétisé, alors que la forme phosphorylée sur les sérines 2 et 5 est reconnue par les facteurs d'élongation, d'épissage et de polyadénylation (Fig. 4) (Peterlin and Price, 2006). La Ser⁷ est reconnue par le complexe Intégrateur au niveau des gènes qui codent pour des snRNA, cette interaction étant importante pour la formation des extrémités 3' des snRNA (Egloff et al., 2007).

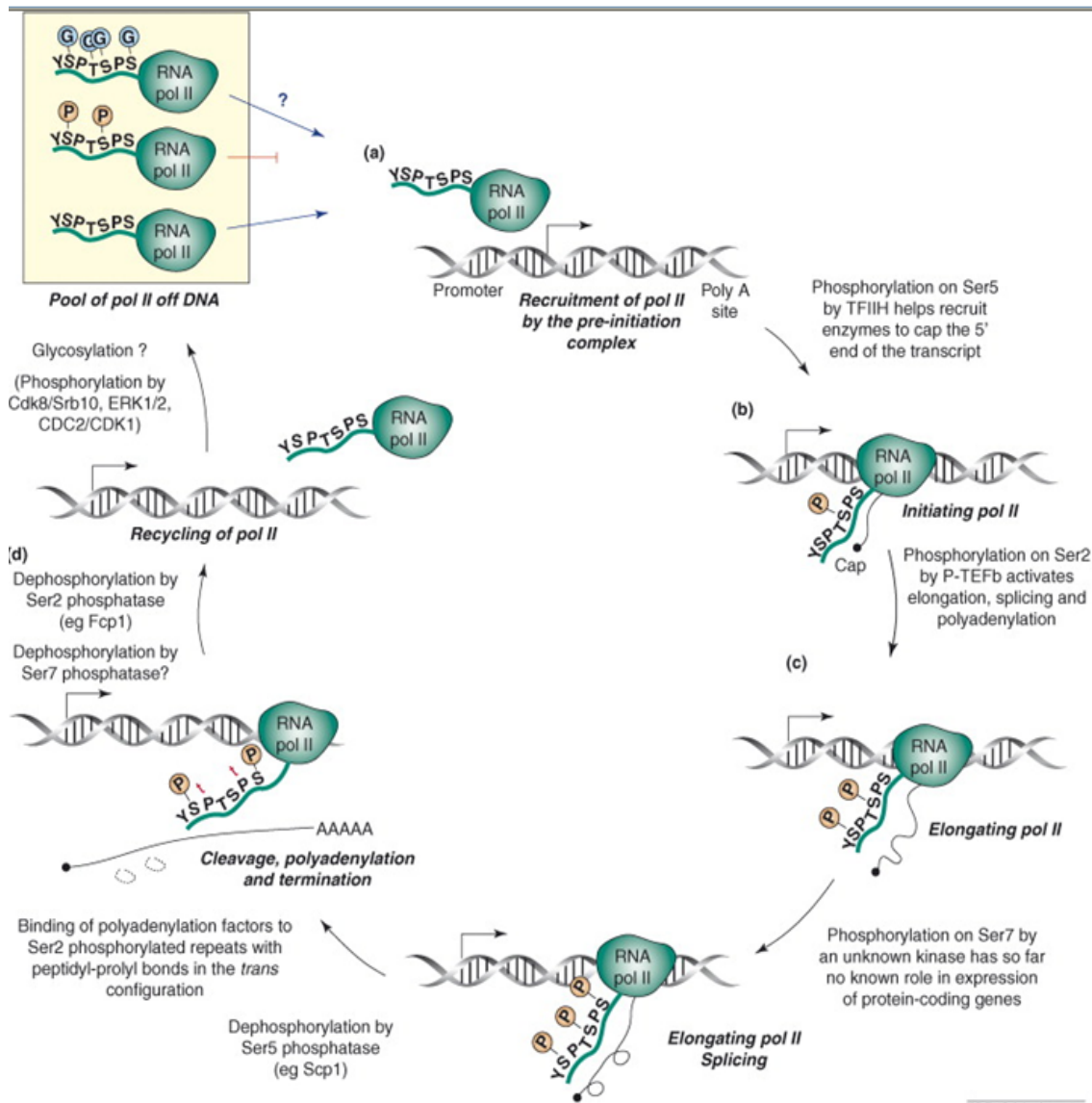


Figure 4. Les modifications post-traductionnelles du domaine CTD de l'ARN Pol II et son rôle pendant les étapes de la transcription. L'ARN Pol II non-phosphorylée est recrutée dans le complexe de pré-initiation de la transcription. La phosphorylation du CTD par: TFIIF sur la Ser⁵ mène au recrutement des enzymes qui ajoutent une coiffe à l'extrémité 5' du transcrit et par P-TEFb sur la Ser² active l'élongation, l'épissage et la polyadenylation. Des étapes successives de déphosphorylation par différentes phosphatases contribuent au recyclage de l'ARN pol II. La forme déphosphorylée ou glycosylée entrerait dans de nouveaux cycles de transcription alors que la forme hyperphosphorylée suite à l'action de différentes

kinases telles que CDK8, ERK1/2 constitue une réserve de polymérase inactive. Tiré de Egloff et Murphy, 2008.

1.2.2 Structure de l'ARN Pol II en complexe avec des facteurs généraux de transcription

L'analyse structurale de l'ARN polymérase en complexe avec différents facteurs de transcription combinée à des études biochimiques fournissent des modèles du processus transcriptionnel et proposent des fonctions spécifiques pour ces facteurs. Ainsi, la structure du complexe de l'ARN Pol II avec le facteur TFIIB publiée récemment (Kostrewa et al., 2009) confirme les résultats des études biochimiques et génétiques effectuées dans les deux dernières décennies (Deng and Roberts, 2007) et propose de nouvelles fonctions de TFIIB dans l'ouverture de la bulle transcriptionnelle et dans la formation de la conformation fermée ou ouverte de la pince.

Le rôle dans la sélection du site d'initiation a été attribué à un domaine du TFIIB appelé "reader" (Fig. 5) puisqu'il approche le site actif de la polymérase et aide dans la reconnaissance du site +1 de façon analogue à un scanner. Ce rôle avait été attribué par une étude antérieure (Bushnell et al., 2004) au domaine "finger" qui approche le site catalytique en entrant par le canal de sortie de l'ARNm. Des études par mutagenèse confirment les nouvelles fonctions proposées (Hawkes and Roberts, 1999; Kostrewa et al., 2009). Les analyses structurales et biochimiques du complexe ARN Pol II-TFIIB convergent sur un modèle cohérent de l'initiation de la transcription et de la transition entre l'initiation et l'élongation (Kostrewa et al., 2009). TFIIB lie par son domaine N-terminal qui contient un motif ruban de zinc (*zinc ribbon*) le domaine "dock" de l'ARN Pol II, tandis que par son domaine C-terminal lie l'ADN et le recrute ainsi sur la polymérase. Le domaine "core" de TFIIB contribue au positionnement du promoteur sur le centre actif. L'ouverture de l'ADN autour de 20 pb en aval de la boîte TATA est favorisée par le domaine

“linker”. Suite à ce phénomène, le brin matrice descend dans le canal de la polymérase, le domaine “reader” stabilise la bulle transcriptionnelle et aide la sélection du site d’initiation de la transcription. La synthèse de transcrits plus longs que 7 nucléotides mène à l’éjection de TFIIB à cause des interférences spatiales entre le transcrit et les domaines “linker” et “reader” de TFIIB, ce qui permet la formation du complexe l’élargissement.

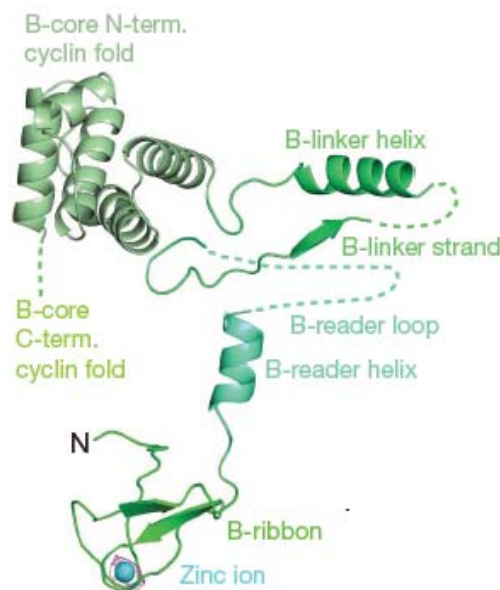


Figure 5. Représentation des domaines structuraux de TFIIB incluant les domaines nouvellement identifiés, “linker” et “reader”. Tiré de (Kostrewa et al., 2009).

Un autre complexe de l’ARN Pol II qui a été étudié est celui avec le facteur TFIIS (Fig. 6). TFIIS est organisé en trois domaines, I à III. Tandis que les domaines II et III sont nécessaires pour l’activité de stimulation de la polymérase en élargissement et sont conservés entre les espèces, le domaine N-terminal (I) n’est pas nécessaire pour cette activité et sa séquence est plutôt divergente entre les espèces (Awrey et al., 1998; Nakanishi et al., 1995). Les co-cristaux ARN Pol II-TFIIS contenaient une version de TFIIS ayant les domaines II et III qui sont nécessaires pour la liaison à

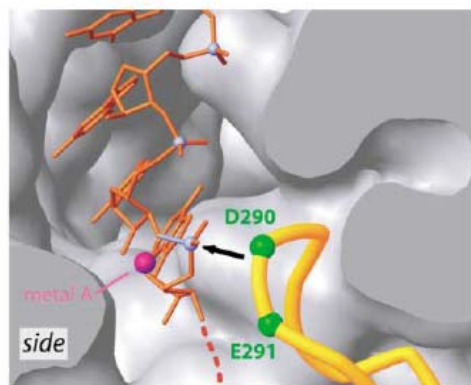
l'ADN et pour l'activité de stimulation du clivage de l'ARN, respectivement (Kettenberger et al., 2003; Kettenberger et al., 2004). Le domaine II lie la surface de la polymérase au niveau de la mâchoire formée par les sous-unités Rpb1/Rpb9. La région "linker" entre les domaines II et III s'insère dans le canal d'entrée de l'ADN de façon à ce que le domaine III arrive au niveau du site catalytique où il positionne deux acides aminés strictement conservés d'une boucle β -harpin de ce domaine à la proximité du métal A. Ces deux acides aminés approchent la chaîne d'ARN. Un mécanisme à deux ions métalliques serait commun aux deux activités catalytiques de la polymérase, de polymérisation et de clivage du transcrit. La différence entre ces deux activités est que pour la polymérisation le métal B est coordonné par le nucléotide NTP qui sert de substrat, alors que pour le clivage ce métal ainsi qu'une molécule d'eau sont coordonnés par un des deux acides aminés conservés situés sur la boucle en épingle à cheveux de type β de TFIIS. La molécule d'eau est nécessaire pour l'attaque nucléophile du lien phosphodiester et son clivage (Fig. 6A).

Lorsque la polymérase rencontre des obstacles pour la transcription, elle recule sur la matrice d'ADN et l'extrémité 3' du transcrit est dirigée dans le pore qui est déjà occupé par le domaine III de TFIIS (Fig. 6B). Dans cette situation, l'extrémité 3' de l'ARN n'est plus alignée sur le site actif et pour que la transcription puisse reprendre, la portion du transcrit qui est sortie dans le pore doit être clivée. De cette façon une nouvelle extrémité 3' de l'ARN est générée et cette extrémité sera alignée et disponible pour la polymérisation.

L'analyse comparative des structures de l'ARN polymérase libre et celle de son complexe avec TFIIS a révélé que TFIIS induit des changements de conformation du site actif de la polymérase. TFIIS lie des résidus de Rpb1, ce qui induit des changements de conformation qui se propageraient à travers les boucles "gâchette" ("trigger"), de "pontage" ("bridge helix"), les boucles 1 et 2 ("switch" 1 et 2) et finalement la "boucle d'embranchement 2" ("fork loop"). Ces changements de conformations résulteraient dans l'élargissement du site de liaison de l'hybride

ADN-ARN à la polymérase ce qui implique que la force de contact hybride-polymérase est modifiée. Le repositionnement de l'hybride dans le centre actif serait favorisé de cette façon et ceci faciliterait le clivage du transcrit (Kettenberger et al., 2003).

A



B

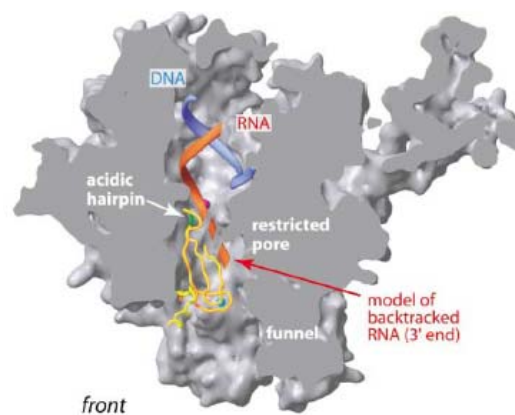


Figure 6. Modèle du complexe ARN Pol II-TFIIS. A) Détail du centre actif de la polymérase avec l'ARN et le domaine III de TFIIS. Les deux acides aminés invariants sont indiqués. La flèche noire indique la direction de l'attaque nucléophile sur le lien phosphodiester qui sera clivé. B) Image transversale du modèle du complexe de la polymérase avec la matrice d'ADN, l'ARN et le domaine III de TFIIS. Le modèle montre le transcrit qui entre dans le pore occupé par TFIIS lorsque la polymérase recule sur la matrice d'ADN. Figure tirée de (Kettenberger et al., 2003).

1.3 Les étapes de la transcription par l'ARN Pol II et les facteurs généraux de transcription

La transcription par l'ARN Pol II se déroule en trois grandes étapes, soit l'initiation, l'élongation et la terminaison. Dans la première étape, la machinerie transcriptionnelle générale et l'ARN Pol II elle-même sont recrutées au promoteur pour former le complexe de préinitiation de la transcription (PIC). Ce complexe se trouve tout d'abord dans une conformation inactive et, afin que la transcription débute, un changement de conformation doit se produire. Ceci permet l'ouverture de l'ADN sur 11 à 15 pb autour du promoteur, le glissement du brin matrice dans le centre actif et la synthèse du premier lien phosphodiester de l'ARN, tel que présenté dans la section précédente. Avant d'entrer en élongation productive, la polymérase synthétise plusieurs courtes séquences d'ARN qui sont relâchées dans des cycles d'initiations avortées. Lorsqu'un transcrit de plus de 30 pb est synthétisé un nouveau changement de conformation permet à la polymérase de dégager le promoteur alors que les GTFs, restent liés aux promoteurs (par exemple TBP) ou sont relâchés. Les facteurs qui restent fixés sur le promoteur forment un échafaudage pour des cycles de réinitiation (Fig. 7). D'autres facteurs tels que des facteurs d'élongation, de coiffage du transcrit ou d'épissage peuvent voyager avec la polymérase en élongation. Lorsque les sites de terminaison sont rencontrés, la polymérase et le transcrit polyadénylé sont relâchés (Hahn, 2004; Sikorski and Buratowski, 2009).

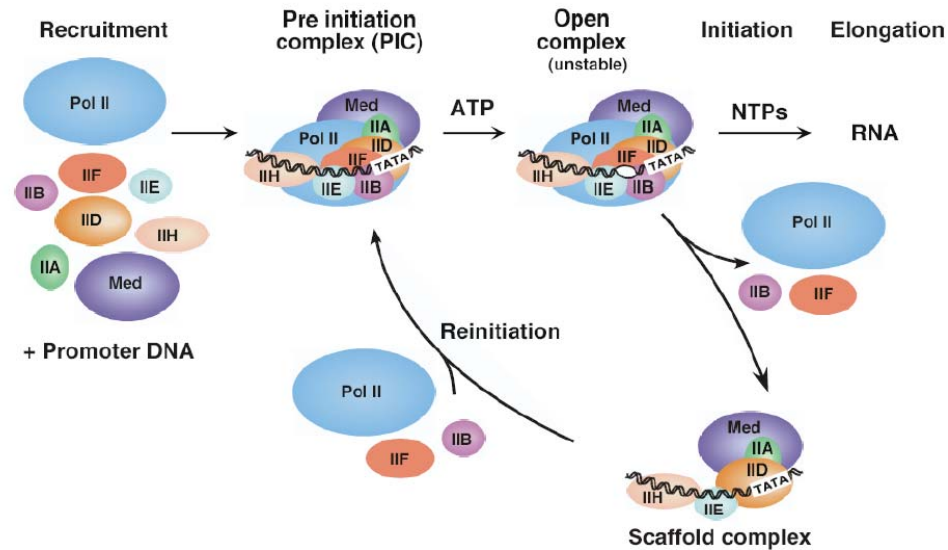


Figure 7. Les principales étapes de la transcription par l'ARN Pol II. Tirée de (Hahn, 2004)

1.3.1 L'initiation de la transcription et les facteurs généraux de transcription impliqués

Chez les organismes eucaryotes, les promoteurs constituent des plateformes d'assemblage du PIC dont la fonction est de sélectionner le site d'initiation de la transcription (TSS). La formation du PIC débute avec la liaison du facteur TFIID à des séquences spécifiques telles que la boîte TATA, l'initiateur (INR) ou DPR (*Downstream Promoter Element*) suivi du recrutement des autres GTFs et de l'ARN Pol II, soit de façon séquentielle, soit sous forme d'holoenzyme déjà pré-assemblée. Pour la transcription basale, la présence des GTFs et de la polymérase est suffisante pour l'assemblage du PIC aux promoteurs, alors que pour la transcription régulée différents co-facteurs sont nécessaires pour que les signaux amenés par les activateurs soient transmis à la machinerie générale.

Les étapes de la transcription ne sont pas régulées par des facteurs uniques mais plutôt par des combinaisons variées de facteurs ubiquitaires ou spécifiques à différents types cellulaires. La capacité d'un organisme d'utiliser un nombre limité de protéines dans la régulation des aspects variés est basée sur un mécanisme appelé contrôle combinatoire (Britten and Davidson, 1969). Ainsi, la modulation de la transcription de façon spécifique au gène ou au type cellulaire se fait par le recrutement d'un cofacteur ou d'une combinaison de plusieurs cofacteurs. Les principales catégories de co-facteurs sont : les facteurs associés à TBP(TAFs), le complexe Médiateur ainsi que des cofacteurs positifs (PC1 à PC4) ou négatifs (NC1/HMGB1). Un sommaire des fonctions des facteurs impliqués dans l'initiation de la transcription est présenté dans le Tableau 1, à la fin de cette section (pages 44-45).

Quoique les sous-unités de l'ARN Pol II et les GTFs soient fortement conservés dans l'évolution, les séquences des promoteurs sont variées (Muller et al., 2007). Ceci suggère que les complexes d'initiation de la transcription ont des compositions aussi variées, et, implicitement, que les GTFs ne sont pas requis pour la transcription de tous les gènes (Sikorski and Buratowski, 2009). Des études à l'échelle du génome qui montrent que les GTFs ne sont localisés que sur une fraction variable de promoteurs (Huisinga and Pugh, 2004a; Zhu et al., 2006) confirment cette hypothèse et redéfinissent la notion de facteurs généraux de transcription comme étant des facteurs nécessaires pour la transcription d'un grand nombre de gènes. D'autres observations restreignent encore plus cette notion de "facteur général de transcription": les GTFs peuvent être exprimés que dans certains types cellulaires et chez certains organismes ils peuvent être absents (par exemple, l'expression des facteurs de la famille de TBP) (Muller and Tora, 2004).

L'assemblage de complexes de préinitiation de composition variable en GTFs sur les promoteurs de différents gènes est proposé comme un nouveau niveau de régulation de la transcription (Muller et al., 2007; Muller and Tora, 2004). Le dogme de l'initiation par un groupe constant de GTFs qui reconnaissent des

séquences consensus du promoteur telle que la boîte TATA a été infirmé par de nombreuses observations. Ainsi, la boîte TATA n'est présente que dans environ 10% (Bajic et al., 2006) à 20% des promoteurs (Sandelin et al., 2007); l'assemblage de TBP n'est pas absolument nécessaire pour l'initiation et n'est pas nécessairement un indicateur de la transcription active; finalement, plusieurs membres de la famille de TBP peuvent lier des promoteurs spécifiques et initier la transcription (Bartfai et al., 2004; Chong et al., 2005; Dantonel et al., 1999; Muller et al., 2007). Plutôt que des séquences universelles pour tous les promoteurs reconnus par des PICs de composition constante, les promoteurs ont des séquences très variables qui sont reconnues de façon combinatoire par des PICs de composition spécifique. Dans ce sens, il a été démontré récemment par une nouvelle technique de spectroscopie de fluorescence combinée à une puce de protéines, que le complexe ternaire TBP-TFIIA-TFIIB lie un groupe de séquences différent de celui reconnu par chaque facteur individuellement (Bonham et al., 2009).

Le promoteur basal a été initialement défini comme étant la séquence d'ADN minimale qui est suffisante pour que l'initiation de la transcription sur une matrice d'ADN se fasse à un niveau basal à partir d'un site d'initiation bien défini et unique (Smale, 1994). Des études *in vitro* utilisant des facteurs purifiés et des séquences d'ADN variées ont identifié des séquences consensus pour lesquelles ces facteurs ont des affinités et auxquelles ils lient spécifiquement. Ainsi TBP lie la séquence consensus TATAA/TAAG/A (boîte TATA) située à une distance de 25-30pb par rapport au site d'initiation de la transcription (Patikoglou et al., 1999). La séquence consensus appelée "initiateur" (Inr) puisqu'elle inclut le site +1 est reconnue par différentes TAFs (*TBP Associated Factors*) (Chalkley and Verrijzer, 1999). Un autre élément identifié dans les promoteurs qui possèdent un "initiateur" est la séquence DPE (*Downstream Promoter Element*), ces deux séquences étant liées de façon coopérative par TFIID (Kutach and Kadonaga, 2000). TFIIB fait des contacts avec des séquences spécifiques situées immédiatement en amont et en aval de la boîte TATA. Ces séquences consensus sont appelées BRE(*u*) ou (*d*) pour "TFIIB Recognition Element" (*upstream* ou *downstream*) et sont reconnues par des

domaines distincts de TFIIB (Deng and Roberts, 2006). Deux autres éléments ont été identifiés plus récemment: le MTE (*Motif Ten Element*) qui fonctionne de façon synergique avec “initiateur”, mais aussi avec TATA et DPE (Lim et al., 2004); le XCPE2 (*X Core Promoter Element 2*) qui joue un rôle dans l’assemblage du PIC à des promoteurs dépourvus de boîte TATA et qui présentent de nombreux sites d’initiation de la transcription (Anish et al., 2009). Ces séquences consensus ont le rôle de positionner le site d’initiation de la transcription, d’établir la direction de la transcription et elles contribuent également à la force du promoteur (Juven-Gershon et al., 2006).

Des mutations des séquences consensus du promoteur ont été associées à des maladies spécifiques. Par exemple, des mutations de la boîte TATA qui mènent à la disparition ou à la formation d’une telle séquence, ou qui modifient la puissance du promoteur seraient associées à un risque accru de thalassémie (mutations des boîtes TATA des gènes globine), de cancers hépatocellulaires (mutations de la boîte TATA du gène Interleukine 1B), pulmonaires (mutations de la boîte TATA du gène ou P450 2A6) oraux (mutations de la boîte TATA du gène UGT 1A7), et d’hypertension artérielle (mutations des boîtes TATA des gènes pour les récepteurs β de l’estrogène) (Savinkova et al., 2009).

Des nouvelles techniques qui permettent l’identification de sites d’initiation de la transcription à grande échelle (par exemple CAGE ou “*CAP Analysis of Gene Expression*”), ainsi que des approches bioinformatiques, ont montré qu’à l’échelle du génome la présence de ces sites consensus dans la séquence des promoteurs est loin d’être universelle (Sandelin et al., 2007; Trinklein et al., 2003; Yamashita et al., 2006). Deux classes majeures de promoteurs ont été identifiées : dans la première les promoteurs possèdent une région restreinte d’initiation de la transcription, présentent une boîte TATA et une régulation serrée de l’expression de façon spécifique au tissu ou au stade développemental, alors que les promoteurs de la deuxième présentent plusieurs sites d’initiation de la transcription distribués sur des

distances variables, n'ont pas de boîte TATA mais plutôt des ilots CpG et sont le plus souvent associés à des gènes de maintenance.

Il a été évalué que chez l'humain et chez la souris, autour de 50% des gènes utilisent des promoteurs alternatifs qui peuvent être à des distances allant jusqu'à des centaines de kilobases (Baek et al., 2007; Takeda et al., 2007). Un exemple est le gène impliqué dans l'hématopoïèse RUNX (*Runt-related transcription factor 1*) dont les promoteurs alternatifs sont à plus de 160 kb (Levanon et al., 2003). L'utilisation des promoteurs alternatifs peut générer des ARNm qui ont des régions 5'UTR (*5'Untranslated Region*) différentes mais le même premier exon, ce qui permet une régulation différentielle de la traduction. Les caractéristiques de la séquence de ces 5'UTR, telles que les structures secondaires et les sites d'entrée pour les ribosomes, modulent l'efficacité de traduction de la protéine. Ceci signifie que l'usage des promoteurs alternatifs ne contribue pas qu'à la régulation de la transcription mais aussi à celle de la traduction (Fossat et al., 2005). Lorsque les promoteurs alternatifs sont séparés par des exons les isoformes générées ont des séquences différentes et peuvent avoir des fonctions différentes et même antagonistes. Un exemple est le gène LEF1 (*Lymphoid Enhancer Factor 1*) qui, en utilisant deux promoteurs alternatifs, produit une isoforme longue et une courte. L'isoforme longue recrute la β -caténine à des gènes régulés par la voie WNT, alors que l'isoforme courte ne peut pas interagir avec la β -caténine et inhibe l'activation de ces gènes (Arce et al., 2006).

1.3.1.1 TFIID

TFIID désigne le complexe formé par TBP (*TATA Binding Protein*) et 14 TAFs (*TBP Associated Factors*). TBP est impliqué dans la transcription par les trois types d'ARN polymérases nucléaires et son rôle le mieux caractérisé est d'initier la formation du PIC. Son recrutement aux promoteurs est régulé par les facteurs auxquels il est associé, tels que les TAFs ou les répresseurs Mot1 et NC2 dans le cas

de la transcription par l'ARN Pol II (Pugh, 2000) et le CF (*Core Factor*) ou UAF (*Upstream Activator Factor*) pour la transcription par l'ARN Pol I (Steffan et al., 1996). Les TAFs ne sont pas présents uniquement dans TFIID, ils font également partie d'autres complexes, tels que les complexes SAGA (*Spt5-Ada-Gcn5-Acetyltransferase*) chez la levure et STAGA (*Spt3-TAF9-GCN5L*) chez l'humain et chez *Drosophila*. De nombreuses analyses biochimiques ou génétiques ont identifié les orthologues des TAFs chez plusieurs espèces et une nomenclature unifiée a été émise (Tora, 2002).

TFIID joue trois rôles indépendants dans la transcription: il lie les promoteurs qui ont ou pas une séquence TATA, il interagit avec des activateurs pour augmenter l'assemblage du PIC et il possède des activités enzymatiques par lesquelles il modifie la chromatine ainsi que d'autres facteurs impliqués dans la transcription.

Des études *in vitro* ont suggéré que TFIID forme un échafaudage qui est nécessaire pour la réinitiation, cependant des études *in vivo* suggèrent un autre modèle. Ce modèle est basé sur l'observation que l'atténuation de l'expression de TBP ou de différentes TAFs dans des cellules embryonnaires murines ou humaines mènent à l'arrêt de la prolifération et à l'apoptose, alors que dans des cellules adultes post-mitotiques l'atténuation de l'expression de ces facteurs n'induit pas l'apoptose (Martianov et al., 2002; Metzger et al., 1999; Mohan, Jr. et al., 2003). Dans ce dernier cas, le seul effet observé au niveau transcriptionnel est le défaut dans l'activation de l'expression des gènes spécifiques au tissu, alors qu'aucun effet sur l'expression des gènes qui étaient déjà actifs dans les cellules embryonnaires n'a pas observé. Ceci suggère que TFIID est important pour l'activation de la transcription pendant les différents stades du développement ou dans des cellules adultes, mais pas pour la transcription des gènes qui sont déjà actifs, ou en d'autres mots pour la réinitiation (Tatarakis et al., 2008).

Une seule molécule de TAF1, TAF2 et TAF7 est présente dans le complexe TFIID, alors que deux ou même plusieurs molécules des autres sous-unités sont présentes

(Leurent et al., 2004). Les analyses structurales du complexe sont limitées à cause de l'hétérogénéité reliée aux réarrangements structuraux. Les informations fournies par des études à faible résolution se limitent à la forme globale du complexe et au positionnement des sous-unités (Fig. 8). TFIID présente trois lobes qui forment une structure en forme de selle et cette structure est conservée entre les espèces (Andel, III et al., 1999; Brand et al., 1999). TBP se trouve dans la partie centrale de la selle. Il a été observé par des études de microscopie électronique que le complexe a une forme dynamique qui peut osciller entre une conformation "ouverte" et "fermée", ce qui modifie la spécificité de reconnaissance du promoteur (Liu et al., 2008). La sous-unité TAF2 peut être absente ou présente à des concentrations molaires faibles dans les préparations de TFIID, et sa présence a un impact important sur la structure de TFIID (Grob et al., 2006). Une étude récente sur la structure des complexes TFIID avec chacun des activateurs p53, Sp1 et c-Jun a montré que l'interaction avec ces activateurs induit des changements structuraux sur des régions restreintes et conservées de TFIID et suggère que ces changements très précis transmettent l'information amenée par les activateurs. (Liu et al., 2009).

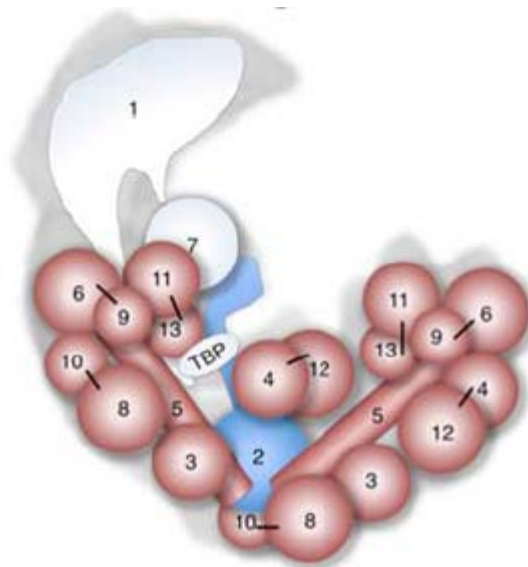


Figure 8. Représentation schématique de l'organisation structurale de TFIID chez la levure, indiquant la position approximative des sous-unités (TAF1 à TAF13 et

TBP). La dimension des sphères est proportionnelle à la masse moléculaire des protéines et les lignes noires indiquent les interactions protéine-protéine documentées (Tiré de Cler et al., 2009).

TFIID est le premier composant du PIC qui lie le promoteur, autant dans le cas d'un recrutement séquentiel des facteurs d'initiation, que lorsque la polymérase est recrutée sous forme d'holoenzyme. Les interactions de TFIID fait avec le promoteur impliquent les facteurs TBP et TAFs. Le domaine C-terminal de TBP présente deux répétitions directes qui sont hautement conservées et c'est par cette région que TBP lie les promoteurs (Kim et al., 1993a). Ce domaine a la forme d'une selle et il lie l'ADN dans le sillon mineur, ce qui induit une courbure de 80° de l'ADN vers le sillon majeur (Kim et al., 1993a). Les interactions avec les facteurs TFIIA, TFIIB et les TAFs impliquent aussi le domaine C-terminal. Pour ce qui est de son domaine N-terminal, malgré qu'il ne lie pas l'ADN directement, il aurait le rôle de moduler la capacité de TBP à lier les promoteurs, en induisant une conformation qui inhiberait la formation de la courbure de l'ADN (Zhao and Herr, 2002). Cette région est hautement divergente entre les espèces, cependant une région avec plusieurs répétitions de glutamine est conservée. Une expansion de cette région à un nombre supérieur de 42 répétitions est souvent associée à des maladies neurodégénératives, telle que l'ataxie spinocérébrale (Nakamura et al., 2001).

TBP n'est pas l'équivalent fonctionnel de TFIID, la présence du complexe TFIID purifié étant nécessaire pour l'initiation de la transcription par des activateurs tel que Sp1, tandis que le polypeptide TBP purifié n'est pas suffisant pour cette activation (Pugh and Tjian, 1990). Il a été démontré que les activateurs interagissent avec les TAFs et ainsi ils recrutent TFIID aux promoteurs (Garbett et al., 2007; Reeves and Hahn, 2005). TBP ne lie pas les séquences du promoteur avec une grande spécificité, cependant plusieurs types d'interactions diminuent son recrutement non-spécifique sur l'ADN, telles que l'homodimérisation de TBP (Jackson-Fisher et al., 1999) ainsi que l'interaction compétitive avec différents facteurs au niveau du domaine de TBP qui interagit avec l'ADN (Kokubo et al., 1998). Par exemple,

TAF1 interagit avec des résidus lysine et arginine hautement conservés du domaine de TBP qui lie la boîte TATA et au facteur TFIIA inhibant ainsi ces deux interactions de TBP (Kokubo et al., 1998). Aussi, l'orientation dans laquelle il lie la boîte TATA n'est pas dictée par les interactions TBP-ADN mais plutôt par d'autres facteurs avec lesquels TBP interagit, tels que TFIIB et TFIIA (Pugh, 1996).

En plus de TBP, les organismes eucaryotes multicellulaires possèdent au moins un des membres de la famille de TBP, appelés TRFs pour "*TBP Related Factors*". Chez la levure et les vers, il n'existe pas d'homologue pour ces facteurs. Un premier membre de cette classe, TRF1 présente une expression restreinte à certains types cellulaires (par exemple les neurones, les gonades) (Holmes and Tjian, 2000) et il cible des gènes transcrits par l'ARN Pol II et III (Isogai et al., 2007; Takada et al., 2000). Quoique ses surfaces de liaison à l'ADN présentent un grand degré d'homologie avec celles de TBP, TRF1 lie une série distincte de gènes, au niveau de régions du promoteur qui sont riches en TC (Pugh, 2000). TRF1 interagit, tout comme TBP, avec un TAF qui fonctionne spécifiquement dans la transcription par l'ARN Pol III, la BRF. Par contre, aux promoteurs des gènes transcrits par l'ARN Pol II, il s'associe à des TAFs spécifiques tels que les TAFs neuronaux (Hansen et al., 1997). Il existe encore deux autres TRFs. TRF2 présente moins de similarité avec TBP et une liaison à l'ADN n'a pas été montrée, les surfaces de liaison à l'ADN n'étant pas conservées chez ce membre de la famille TRF. Il a été proposé qu'il joue plutôt un rôle dans la répression de la transcription, puisqu'il entrerait en compétition avec TBP pour la liaison de TFIIA (Teichmann et al., 1999).

TFIID fait de multiples contacts avec les promoteurs. En plus de l'association entre TBP et la boîte TATA, les TAFs font des contacts additionnels avec l'ADN. Par exemple, TAF1 et TAF2 reconnaissent la séquence Inr, alors que TAF6 et TAF9 font des contacts avec la séquence DPE. Ces contacts additionnels sont requis pour l'initiation à des promoteurs qui n'ont pas de boîte TATA, où les TAFs sont essentiels (Huisinga and Pugh, 2004b). Plusieurs TAFs ont un domaine appelé "histone fold" qui contribue à la reconnaissance des séquences spécifiques du

promoteur ainsi qu'à l'intégrité du complexe TFIID. Les TAFs qui possèdent ce domaine forment une structure octamérique similaire aux histones (Gangloff et al., 2001) qui augmenterait la reconnaissance spécifique du promoteur, mais aussi les contacts non-spécifiques (Shao et al., 2005).

TAF1 présente de nombreuses activités enzymatiques par lesquelles il est impliqué dans la modification post-traductionnelle des histones. Ainsi, il possède des activités: 1) histone acétyltransférase (HAT) pour les histones H3 et H4 (Mizzen et al., 1996); 2) kinase pour la phosphorylation de H2B (Maile et al., 2004); 3) d'activation et de conjugaison de l'ubiquitine pour H1 (Pham and Sauer, 2000). Ces activités enzymatiques sont importantes pour la modulation de la transcription dans le contexte de la chromatine. De plus, TAF1 peut aussi modifier d'autres facteurs tels que TFIIE β (acétylation) (Imhof et al., 1997) et Rap74, PC4 et TFIIA β (phosphorylation) (Dikstein et al., 1996a; Solow et al., 2001).

Des interactions directes et spécifiques avec des activateurs ont été caractérisées pour de nombreux TAFs (Burley and Roeder, 1996). De cette façon, les TAFs agissent comme des co-activateurs qui font un pont entre les activateurs et la machinerie transcriptionnelle générale. L'hypothèse que TFIID est requis pour la transcription activée de façon générale a été émise à partir de l'observation qu'en utilisant des systèmes de transcription *in vitro* TFIID est nécessaire pour la transcription activée, alors que TBP ne suffit pas. Cependant, plusieurs études *in vivo* ont montré que chez la levure, juste une partie des gènes dépendent de TAFs pour la transcription (Kuras et al., 2000; Li et al., 2000). Même si pas tous les gènes dépendent des TAFs pour leur transcription, il a été montré par des études à grande échelle chez l'humain et chez la levure qu'une majorité de gènes dépendent de certains TAFs, telles que TAF1 et TAF9 (Huisinga and Pugh, 2004a). Les gènes dépendants des TAFs sont surtout des gènes de maintenance (qui sont des gènes exprimés de façon constitutive dans tous les types cellulaires), qui ne présentent pas de boîte TATA (Mencia et al., 2002).

Des paralogues des TAFs ont été décrits chez plusieurs espèces, allant de la levure jusqu'à l'humain. Ces paralogues présentent des patrons d'expression spécifiques aux types cellulaires ou à un stade de développement précis, ce qui suggère qu'ils ne sont pas des composants universels du PIC et qu'ils ont plutôt le rôle de régulation des différents programmes transcriptionnels. Par exemple, les paralogues TAF10 et TAF10b sont exprimés de façon différente pendant les stades de l'embryogenèse (Georgieva et al., 2000). D'autres paralogues des TAFs ne sont exprimés que dans certains types cellulaires, un exemple étant TAF4b, dont l'expression est spécifique aux lymphocytes B et aux gonades (Dikstein et al., 1996b; Freiman et al., 2002).

Plusieurs complexes qui contiennent des TAFs, mais pas de TBP ont été identifiés chez plusieurs espèces. Dans ces complexes, une composition variée en TAFs est accompagnée de sous-unités qui présentent des activités différentes allant des acétyltransférases (Struhl and Moqtaderi, 1998) aux désubiquitinasés (Daniel et al., 2004; Zhao et al., 2008). Quelques exemples sont les complexes SAGA (*Spt5-Ada-Gcn5-Acetyltransferase*) et SLIK (*SAGA-Like Complex*) chez la levure, TFTC (*TBP-Free TAF_{II} containing Complex*) et STAGA (*Spt3-TAF9-GCN5L*) chez l'humain et *Drosophila*. Il a été démontré par des essais transcriptionnels *in vitro* utilisant des matrices d'ADN variées que TFTC peut reconnaître des promoteurs qui ont, ou qui n'ont pas une boîte TATA. Les autres complexes qui contiennent des TAFs ont été impliqués dans la transcription activée dans le contexte de la chromatine (Muller and Tora, 2004), une exception étant le complexe PRC1 (*Polycomb Repressive Complex 1*) qui a plutôt des fonctions de répression de la transcription (Lund and van, 2004).

1.3.1.2 TFIIB

TFIIB joue plusieurs rôles dans l'initiation de la transcription, tels que la reconnaissance du promoteur, l'ouverture de la bulle transcriptionnelle, la sélection du site d'initiation de la transcription, ainsi que l'activation. Il a aussi été impliqué dans la transition entre l'initiation et l'élongation et dans la formation de boucles

entre les promoteurs et les terminateurs des gènes actifs chez la levure (Hahn, 2004; Kostrewa et al., 2009; Orphanides et al., 1996). Des études par mutagenèse ont montré que les différentes fonctions dans l'initiation sont indépendantes les unes des autres (Fairley et al., 2002; Hawkes and Roberts, 1999).

Il a été démontré que TFIIB joue un rôle de cofacteur dans la catalyse de la polymérisation par l'ARN Pol II puisqu'il coordonne un ion de Mg^{+2} par deux résidus aspartate de son domaine appelé “*finger*” (Tran and Gralla, 2008). Le même domaine aurait le rôle d'empêcher la transition entre l'initiation et l'élongation avant que le transcrit atteigne entre 6 et 16 nucléotides. En effet, TFIIB aurait une fonction de contrôle de qualité, puisqu'une transition à l'élongation productive ne pourrait se faire qu'en sa présence et non pas à des régions génomiques où la polymérase serait recrutée non-spécifiquement en l'absence de TFIIB (Kostrewa et al., 2009; Tran and Gralla, 2008).

TFIIB est conservé entre les espèces et présente une homologie de séquence et de structure avec le facteur TFB (*Transcription Factor B*) chez les archaea, ainsi qu'avec le facteur BRF (*TFIIB-Related Facteur*) qui est impliqué dans la transcription par l'ARN Pol III chez les eucaryotes (Hampsey, 1998; Orphanides et al., 1996).

TFIIB interagit au niveau des promoteurs à la fois avec l'ADN, les facteurs TBP, TFIIF et l'ARN Pol II elle-même. Les domaines de TFIIB qui sont importants pour son interaction avec la polymérase ont été présentés dans la section 1.2.2, qui décrit l'analyse cristallographique du complexe TFIIB-ARN Pol II (Kostrewa et al., 2009).

TFIIB lie l'ADN des promoteurs au niveau des séquences consensus BRE (upstream/downstream) qui flanquent la boîte TATA et cette interaction stabilise le complexe TFIIB-TBP-ADN (Orphanides et al., 1996). Le domaine de liaison à l'ADN se trouve dans la région C-terminale de TFIIB qui contient un motif hélice-tour-hélice, alors que l'interaction avec l'ARN Pol II et TFIIF se fait au niveau du

domaine N-terminal qui contient un motif ruban de zinc (Hahn, 2004; Reese, 2003). L'interaction TFIIB-ADN est modulée par des activateurs tels que Gal4, VP16 par un mécanisme qui implique la destruction des interactions intramoléculaires entre les domaines C et N-terminaux de TFIIB. Ceci mène à un changement de conformation de TFIIB et l'exposition du domaine de liaison à l'ADN (Hawkes et al., 2000). Des mutations dans le domaine à doigts de zinc de TFIIB qui favorisent les interactions intramoléculaires sont également associées à un défaut dans le recrutement du PIC dépendant des activateurs, ainsi qu'une activation défectueuse *in vivo* et *in vitro* (Elsby and Roberts, 2004; Glossop et al., 2004). L'interaction physique entre différents activateurs et TFIIB a aussi été mise en évidence dans le contexte du phénomène de formation de boucles ou "looping" des gènes, où TFIIB lie autant les promoteurs que les régions 3' des gènes (El et al., 2009; Singh and Hampsey, 2007).

1.3.1.3 TFIIF

Le facteur TFIIF a comme rôle principal d'augmenter l'affinité de la l'ARN Pol II pour le complexe TBP-TFIIB-promoteur. Approximativement 50% de l'ARN Pol II dans des extraits nucléaires de levure est associé avec TFIIF, cette forme étant apte pour l'initiation et la réinitiation de la transcription (Rani et al., 2004). En effet, l'interaction TFIIF-ARN Pol II favorise le recrutement de la polymérase par le complexe TBP-TFIIB-ADN et augmente la stabilité du complexe formé, en augmentant l'affinité de la polymérase pour ce complexe. Le changement topologique induit à l'ADN par TFIIF, ainsi que l'enroulement subséquent de l'ADN autour du PIC contribuent à une augmentation de la résistance à la répression induite par des facteurs négatifs de transcription qui ciblent le PIC (Hou et al., 2000).

TFIIF joue également un rôle dans la sélection du site d'initiation (Fairley et al., 2002). Ssu1 ou Tfg1, l'homologue de Rap74 chez la levure, a été isolé comme supprimeur des défauts dans la sélection du site d'initiation de la transcription

lorsque TFIIB ou Rpb1 sont mutés. Il a été proposé que le domaine de dimérisation de TFIIF soit situé près du domaine “*finger*” de TFIIB et du centre catalytique de la polymérase, de façon à ce que l’interface d’interaction TFIIF-TFIIB-polymérase affecte la sélection du site d’initiation de la transcription (Freire-Picos et al., 2005). Des expériences de photo-pontage protéine-protéine combinées aux essais de clivage par le radical hydroxyle, ainsi que par mutagenèse de la polymérase ont montré que le domaine de dimérisation de TFIIF interagit avec le lobe formé par la sous-unité Rpb2 et avec des domaines de Rpb9. Ceci pourrait expliquer pourquoi des mutations de Rpb9, sous-unité située loin du centre actif de la polymérase affectent la sélection du site d’initiation de la transcription (Chen et al., 2007). TFIIF a aussi un rôle dans le dégagement du promoteur en diminuant la fréquence des initiations avortées et en favorisant l’activité de TFIIF dans l’inhibition des arrêts prématurés de la polymérase pendant les étapes précoces de l’élongation (Yan et al., 1999). Une étude par mutagenèse de Tfg1 a montré que celui-ci a le rôle de stabiliser l’hybride ADN-ARN dans le centre catalytique de la polymérase et de favoriser la formation du premier lien phosphodiester (Khapersky et al., 2008).

Chez l’humain le facteur TFIIF est formé de deux sous-unités, Rap30 et Rap74 (*RNA polymerase associated protein*). Chez la levure il est composé de trois sous-unités, Tfg1, Tfg2 et Tfg3. Tfg1 et Tfg2 sont les homologues de Rap74 et Rap30, respectivement et ils sont essentiels. La troisième sous-unité, Tfg3 présente un grand degré d’homologie avec la sous-unité TAF14 de TFIID et n’est pas essentielle (Cairns et al., 1996). Rap30 possède deux régions d’homologie de séquence avec les facteurs σ , qui sont impliqués dans la reconnaissance du promoteur chez la bactérie. La première de ces régions est située dans le domaine central et est impliquée dans l’interaction avec l’ARN Pol II, tandis que la deuxième est située dans le domaine C-terminal et possède des sites cryptiques d’interaction avec l’ADN (Sopta et al., 1989; Tan et al., 1994). Dans le domaine N-terminal se trouve une région d’interaction avec Rap74 et une autre pour l’interaction avec TFIIB (Fang and Burton, 1996). Rap74 possède trois régions fonctionnelles qui interagissent avec de nombreux facteurs : Rap30, l’ARN Pol II, TFIIB, TAF1,

TFIIA et la phosphatase FCP1 (Archambault et al., 1998; Fang and Burton, 1996; Langelier et al., 2001; Ruppert and Tjian, 1995). La région centrale de Rap74 est exposée et désordonnée dans le PIC (Yong et al., 1998). Cette région peut être phosphorylée par TAF1 mais aussi par autophosphorylation, due à l'activité sérine/thréonine kinase de Rap74. L'autophosphorylation de Rap74 mène à une diminution de l'activité transcriptionnelle de la polymérase en élongation dans des essais de type "run-off" (Rossignol et al., 1999). Le domaine C-terminal de Rap74 possède des sites cryptiques d'interaction avec l'ADN qui ont un motif hélice boucle hélice (HTH ou "*helix-turn-helix*"). Ce domaine contribue à l'interaction de TFIIIF avec l'ADN et l'agrandi en la prolongeant en amont de la boîte TATA et en aval de la séquence Inr, tel qu'observé par des expériences de photo-pontage protéines-ADN (Forget et al., 1997; Robert et al., 1998). Le recrutement de Rap74 dans le PIC induit une deuxième courbure, après celle induite par TBP, ce qui mène à l'enroulement du promoteur autour du PIC (Robert et al., 1998).

1.3.1.4 TFIIIE

TFIIIE facilite le recrutement de TFIIH au promoteur et stimule les activités hélicase et kinase de TFIIH, étant essentiel pour l'ouverture du promoteur. Dans l'assemblage séquentiel du PIC, TFIIIE est recruté après la formation du complexe TBP-TFIIIB-RNA polymérase II-TFIIIF sur le promoteur. Son recrutement se fait par des interactions directes avec des activateurs tels que les facteurs p53 chez l'humain (Okuda et al., 2008), Krüppel et Abdominal B chez *Drosophila* (Okamoto et al., 1998; Zhu and Kuziora, 1996). La formation du PIC qui contient TFIIIE favorise à son tour le recrutement du facteur TFIIH. TFIIIE lie l'ADN du promoteur autour de 10 pb en amont du site d'initiation de la transcription, qui correspond au point d'ouverture de la bulle transcriptionnelle (Forget et al., 2004a). En effet, l'utilisation de matrices pré-ouvertes élimine la dépendance de TFIIIE et TFIIH pour l'initiation de la transcription (Holstege et al., 1996; Pan and Greenblatt, 1994). TFIIIE stimule les activités enzymatiques de TFIIH, soit les activités ATPase, kinase du CTD de l'ARN Pol II et hélicase de l'ADN (Kumar et al., 1998; Watanabe et al., 2003;

Yamamoto et al., 2001). Ces activités sont impliquées dans l'ouverture de la bulle transcriptionnelle et la transition de l'initiation à l'élongation. Il a été montré que TFIIE a aussi des activités transcriptionnelles indépendantes de TFIIH. Par exemple, TFIIE stimule la transcription en absence de TFIIH lorsqu'une matrice d'ADN sous-enroulée ("negatively supercoiled"), qui a moins de 10,1 pb par tour d'hélice a été utilisée (Holstege et al., 1995; Timmers, 1994). L'énergie requise pour ouvrir ce type de matrice est inférieure à celle nécessaire pour des matrices normalement enroulées, de façon à ce que l'ouverture des matrices sous-enroulées ne dépende pas de l'activité ATPase de TFIIH (Parvin et al., 1994). De plus, dans des conditions qui sont défavorables à l'ouverture du promoteur, telles qu'une force ionique élevée ou une température basse, la transcription est d'avantage dépendante de TFIIE, ce qui appuie l'hypothèse que TFIIE a un rôle dans l'ouverture du promoteur en absence TFIIH (Holstege et al., 1995)

Ce facteur possède deux sous-unités, α (56 kDa) et β (34 kDa) qui forment un hétérotétramère $\alpha_2\beta_2$ (Sumimoto et al., 1991). Les deux sous-unités présentent une homologie de séquence avec les facteurs σ bactériens. D'autres domaines fonctionnels présents dans chacune des deux sous-unités sont: une séquence qui stimule l'activité kinase de TFIIH située dans la région N-terminale de chacune des sous-unités et un domaine HTH qui lie l'ADN double brin (Okamoto et al., 1998; Okuda et al., 2000). Le domaine C-terminal de TFIIE β a deux régions basiques : la première est représentée par un domaine HTH qui fait des contacts avec TFIIE α et la deuxième par un domaine hélice qui interagit avec l'ARN Pol II, l'ADN simple brin, TFIIF (la sous-unité Rap30), TFIIB et le facteur de transcription Krüppel chez *Drosophila* (Okamoto et al., 1998). Le domaine C-terminal de TFIIE α n'est pas essentiel chez la levure et fait des interactions avec TFIIH qui favorisent le recrutement de ce dernier dans le complexe de préinitiation de la transcription (Kuldell and Buratowski, 1997; Ohkuma et al., 1991; Ohkuma et al., 1995). Plusieurs autres domaines de TFIIE α ont été identifiés par l'analyse de sa séquence, le des mieux caractérisé étant un motif à doigt de Zn conservé entre les espèces qui

est essentiel pour la transcription (Okuda et al., 2004) et dont la fonction serait d'augmenter l'affinité de TBP pour la boîte TATA (Yokomori et al., 1998).

1.3.1.5 TFIIH

TFIIH joue trois rôles dans le processus transcriptionnel : ATPase dépendante de l'ADN (Conaway and Conaway, 1989; Roy et al., 1994), hélicase dépendante d'ATP (Schaeffer et al., 1993; Serizawa et al., 1993) et kinase pour la CTD de l'ARN Pol II (Feaver et al., 1991; Lu et al., 1992). En plus de ces activités enzymatiques dans la transcription, certaines sous-unités de TFIIH ont aussi des fonctions dans la réparation de l'ADN de type NER (*Nucleotide Excision Repair*) (Coin et al., 2008) ainsi que TCR (*Transcription Coupled Repair*) (Sarker et al., 2005). La dixième sous-unité de TFIIH, TFB5, qui a été identifiée il y a quelques années seulement, a été impliquée dans la réparation de l'ADN (Giglia-Mari et al., 2004; Ranish et al., 2004). Lorsque dissocié de TFIIH, le module CAK est impliqué dans la transition entre les phases G2 et M (métaphase) (Larochelle et al., 1998), alors que TFIIH complet est impliqué dans la transition entre les phases G1 et S (synthèse) du cycle cellulaire (Matsuno et al., 2007).

Le facteur TFIIH est formé de 10 sous-unités regroupées dans 2 modules fonctionnels : le noyau (ou "core") et le module CAK (*Cyclin-Activating Kinase*) (Ranish et al., 2004; Zurita and Merino, 2003). Les sous-unités du noyau sont XPB (*Xeroderma Pigmentosum complementation group B*), p62, p52, p44 et p34. Le module CAK est constitué de CDK7 (*Cyclin-Dependent Kinase 7*), Cycline H et Mat1 (*Ménage à trois*). La connexion entre les deux modules de TFIIH est assurée par la sous-unité XPD (*Xeroderma Pigmentosum complementation group D*).

L'activité ATPase de TFIIH est nécessaire autant pour l'initiation que pour l'étape ultérieure de dégagement du promoteur. L'hydrolyse de l'ATP par TFIIH est requise pour l'ouverture du promoteur et pour la formation du premier lien phosphodiester (Holstege et al., 1998). En absence de TFIIH la polymérase entre

dans des cycles d'initiations avortés et elle ne peut dégager le promoteur pour entrer en élongation productive qu'après l'ajout de TFIIH et d'ATP (Dvir et al., 1997; Kumar et al., 1998). TFIIH possède une activité hélicase bidirectionnelle, 3'-5' par la sous-unité XPB et 5'-3' par XPD (Schaeffer et al., 1993; Schaeffer et al., 1994). Les deux activités nécessitent l'hydrolyse de l'ATP. Tandis que l'activité hélicase de XPB est requise autant pour la transcription que pour la réparation de l'ADN, celle de XPD est uniquement nécessaire pour la réparation de l'ADN (Zurita and Merino, 2003). Deux modèles ont été proposés afin d'expliquer l'implication des activités hélicase de TFIIH dans l'ouverture de la bulle transcriptionnelle. Le premier est basé sur des études de photo-pontage protéine-ADN qui ont révélé que XPB est localisé dans la région -9 et +2 par rapport au site d'initiation de la transcription (+1) (Douziech et al., 2000). Ce modèle propose que l'activité hélicase de TFIIH participe à l'ouverture de la bulle transcriptionnelle autant en amont qu'en aval du site +1. Le deuxième modèle est basé sur des analyses de photo-pontage couplées à des lavages au sarkosyl, qui ont mis en évidence que XPB n'est localisé qu'en aval du site +1 (Kim et al., 2000). Ce modèle propose que TFIIH n'agisse pas comme une hélicase conventionnelle, mais plutôt il induirait une rotation de l'ADN en aval par rapport à un site fixé par des interactions protéine-ADN situé en amont. Cette rotation induirait une torsion suffisante dans l'ADN pour ouvrir le promoteur. Des études par mutagenèse du XPB suggèrent aussi que l'activité hélicase de TFIIH est nécessaire pour l'ouverture du promoteur autant *in vivo* que *in vitro* (Guzman and Lis, 1999; Tirode et al., 1999). À l'opposé, une autre étude qui a utilisé des mutants de XPB qui inactivent sélectivement l'activité hélicase sans affecter son activité ATPase, suggère que l'activité hélicase n'est pas impliquée dans l'ouverture du promoteur, puisque ces mutants ne présentent que des défauts dans le dégagement du promoteur (Lin et al., 2005). Dans le même sens, une étude par photo-pontage a rapporté que XPB ne localise pas dans la région de la bulle transcriptionnelle (Holstege et al., 1997).

La sous-unité CDK7 confère au facteur TFIIH une activité kinase pour le CTD de l'ARN Pol II, qui est importante pour le dégagement du promoteur (Hampsey,

1998). L'activité kinase de CDK7 est régulée par les autres sous-unités du complexe CAK (cycline H et MAT1) ainsi que par TFIIE, XPD, le Médiateur et le petit ARN nucléaire U1 (Keriel et al., 2002; O'Gorman et al., 2006; Svejstrup et al., 1996). De plus, l'interaction entre TFIIH et des activateurs peut stimuler son activité kinase (Di Lello et al., 2006). La phosphorylation du CTD de l'ARN Pol II par CDK7 se fait sur le résidu sérine 5, modification qui est reconnue par les enzymes de coiffage de l'ARNm (Egloff and Murphy, 2008a; Komarnitsky et al., 2000; Pei et al., 2001). Il a été montré récemment que CDK7 est aussi responsable de la phosphorylation du résidu sérine 7 du CTD de la polymérase et que cette modification est importante pour l'accumulation et l'entrée en pause de la polymérase dans les régions situées en aval, proximales au promoteur (Glover-Cutter et al., 2009). Dans la même étude, les mutants qui inhibent l'activité kinase de CDK7 réduisent le recrutement du facteur négatif d'élongation NELF dans ces régions. De plus, ces mutants modifient le patron des modifications post-traductionnelles des histones le long des régions transcrites, ce qui suggère un rôle de CDK7 dans l'élongation. CDK7 phosphoryle aussi d'autres protéines, telles que p53, ER α (*Estrogen Receptor α*), RAR (*Retinoic Acid Receptor α*) (Chen et al., 2000; Lu et al., 1997; Rochette-Egly et al., 1997), etc.

Plusieurs sous-unités de TFIIH ont été impliquées dans la réparation de l'ADN par la voie NER : XPB, XPD, p52, p44 et p62 (Jawhari et al., 2002; Schaeffer et al., 1994; Wang et al., 1995). Des mutations de XPB ou XPD sont responsables de trois maladies génétiques multisystémiques : xeroderma pigmentosum (XP), trichothiodystrophie (TTD) et le syndrome de Cockayne (CS). Ces maladies sont dues à des défauts dans la réparation de l'ADN par NER (XP) ou par TCR (TTD et CS) et se manifestent par une photosensitivité sévère, une neurodégénérescence ou un syndrome de type progéria (Brooks et al., 2008; Kamenisch and Berneburg, 2009). Des défauts dans la réparation des dommages de l'ADN ne constituent pas le seul mécanisme de ces maladies. Par exemple, dans certaines formes de TTD, des mutations de XPD seraient plutôt responsables de la diminution de son activité comme co-activateur transcriptionnel pour l'hormone thyroïdienne et une altération du

patron d'expression des gènes reliés à la myélinisation dans les cellules gliales (Compe et al., 2007).

Récemment, il a été montré que la sous-unité p44 possède une activité E3 ubiquitine-ligase. Des mutations de son domaine RING ont annulé cette activité. Cependant, un rôle de TFIIF dans l'ubiquitination de l'ARN Pol II ou d'autres facteurs impliqués dans la réparation de l'ADN (Takagi et al., 2005) et une contribution au dégagement de la matrice transcriptionnelle demeurent spéculatifs. Par contre, un comportement inhabituel a été récemment démontré pour la sous-unité XPD, qui serait capable de contourner des obstacles tels que des protéines liées à l'ADN simple brin sans qu'aucune protéine ne soit éjectée de l'ADN (Honda et al., 2009).

1.3.1.6 TFIIA

TFIIA a deux types de fonctions dans la transcription, celle d'anti-répresseur et de co-activateur. TFIIA stabilise l'interaction TBP-ADN en faisant des contacts avec l'ADN en amont de la boîte TATA ainsi que par des interactions avec TBP sur la surface opposée à celle qui lie TFIIB (Coulombe et al., 1994; Geiger et al., 1996; Lagrange et al., 1996; Tan et al., 1996). TFIIA n'est pas requis pour la transcription basale ou activée lorsque des facteurs de transcription purifiés sont utilisés (Coulombe et al., 1992; Ma et al., 1996; Ozer et al., 1994; Wu and Chiang, 1998). Cependant, TFIIA stimule autant la transcription basale que celle activée lorsque TFIID est utilisé à la place de TBP (Orphanides et al., 1996; Warfield et al., 2004). Ceci a amené l'hypothèse que TFIIA aide à contrer l'activité inhibitrice des répresseurs présents dans les fractions cellulaires partiellement purifiées (Mermelstein et al., 1996; Thomas and Chiang, 2006). Cette hypothèse a été confirmée par des études démontrant que TFIIA entre en compétition avec des répresseurs tels que NC2, BTAF1 et HMGA1 pour la liaison à TBP et empêche l'effet inhibiteur de ces facteurs sur l'interaction TBP-ADN (Auble and Hahn, 1993; Ge and Roeder, 1994; Xie et al., 2000).

TFIIA joue un rôle central dans l'initiation de la transcription puisqu'il interagit physiquement autant avec des facteurs généraux de transcription (ex. TBP, plusieurs TAFs, TFIIF, TFIIE), qu'avec des activateurs (ex. Gal4-VP16, Zta) et des co-facteurs transcriptionnels (ex. PC4, HMG2) (revu par (Pugh, 2000)). L'un des mécanismes proposés pour son activité comme co-activateur serait que TFIIA induit des changements de conformation de TFIID qui devient apte à interagir avec l'ARN Pol II sous forme de holoenzyme (Chi and Carey, 1996). Un autre mécanisme est basé sur des études de photo-pontage protéine-ADN et de mutagenèse. Ce mécanisme propose que la fonction de TFIIA comme co-activateur serait de stimuler et de stabiliser l'interaction TFIID-ADN dans le cadre du complexe activateur-TFIIA-TFIID-ADN (Dion and Coulombe, 2003; Kobayashi et al., 1995; Ranish et al., 1999). Ceci faciliterait l'étape limitante de reconnaissance du promoteur. De plus, TFIIA stimule des étapes qui suivent la formation du PIC au promoteur puisqu'il stimule l'activité de TFIIE et de TFIIF (Langelier et al., 2001). Cependant, TFIIA ne serait requis que pour la transcription d'une fraction des gènes *in vivo*, puisque dans une souche de levure qui possède un TBP muté et inapte à former l'interaction TBP-TFIIA, l'expression de seulement 3% de gènes est affectée (Stargell et al., 2000). Une préférence de TFIIA pour des promoteurs ayant des séquences BRE a été récemment rapporté par des études d'immunoprécipitation de la chromatine couplées à une diminution de l'expression de TFIIA (Deng et al., 2009). Aussi, TFIIA présente plutôt une activité d'inhibition de la transcription, tel qu'observé par le même groupe lorsqu'une séquence sauvage ou mutante du BRE est utilisée afin de tester l'activité des promoteurs après une diminution de l'expression de TFIIA.

L'activité transcriptionnelle de TFIIA est régulée par des modifications post-traductionnelles, telle qu'une phosphorylation par TAF1, ainsi qu'une acétylation par p300 qui favorisent l'interaction TFIIA-TBP (Mitsiou and Stunnenberg, 2003; Solow et al., 2001)

Chez les eucaryotes supérieurs, TFIIA est constitué de trois sous-unités : TFIIA α , TFIIA β et TFIIA γ . Les sous-unités TFIIA α et TFIIA β sont codées par le gène TFIIA $\alpha\beta$ sous forme d'un précurseur commun à partir duquel les deux sous-unités sont générées par clivage protéolytique (DeJong and Roeder, 1993). Si initialement on croyait que seulement les sous-unités indépendantes peuvent former le facteur TFIIA, ultérieurement il a été montré que le précurseur TFIIA $\alpha\beta$ peut aussi s'associer à la sous-unité TFIIA γ et à TBP (Mitsiou and Stunnenberg, 2000). La sous-unité TFIIA γ est codée par le gène unique TFIIA γ . Chez la levure TFIIA ne possède que deux sous-unités qui sont codées par TOA1 et TOA2. TFIIA $\alpha\beta$ présente un grand degré d'homologie avec les régions N et C-terminales de TOA1 (DeJong and Roeder, 1993; Ma et al., 1993). Un paralogue de TFIIA $\alpha\beta$, ALF, a été identifié dans des types cellulaires spécifiques tels que les testicules chez l'humain ou les oocytes chez *Xenopus laevis* et il a été démontré que ce paralogue lie TFIIA γ et stabilise la liaison de TBP aux promoteurs (Han et al., 2003; Ozer et al., 2000).

1.3.1.7 Le complexe Médiateur

Le complexe Médiateur a fait l'objet de plusieurs controverses autant sur sa composition et la conservation de ses sous-unités entre les espèces, que sur sa fonction dans la transcription comme co-activateur ou plutôt comme facteur général de transcription. Initialement, le complexe Médiateur a été isolé chez la levure comme co-activateur nécessaire pour supporter des niveaux élevés de transcription en présence des activateurs (revu par (Myers and Kornberg, 2000)). Cependant, des études ultérieures ont démontré que le complexe Médiateur peut stimuler également la transcription basale dans des systèmes reconstitués avec des facteurs généraux purifiés et en absence d'activateurs (Kornberg, 2005). Ces observations ont mené au débat à savoir si le complexe Médiateur est un co-activateur ou un facteur général de transcription. Des études à l'échelle du génome chez la levure ont montré qu'il est localisé en amont de la plupart des gènes actifs et même de certains gènes qui ne sont pas transcrits activement, ce qui appuie la fonction comme GTF (Andrau et al., 2006; Zhu et al., 2006). De plus, Médiateur est nécessaire pour le recrutement de la

polymérase aux promoteurs des gènes exprimés de façon constitutive, puisqu'un mutant sensible au froid de la sous-unité essentielle Srb4 démontre une diminution marquée de l'association de la polymérase à ces promoteurs (Ansari et al., 2009). À l'opposé, ce complexe n'est pas requis pour la transcription basale dans des systèmes purifiés comme le sont les GTFs (à l'exception de TFIIA), quoi qu'il puisse la stimuler, tel que mentionné plus haut (Kornberg, 2005).

Des indices biochimiques directs sur l'activité de coactivateur du Médiateur ont été obtenus en utilisant un système reconstitué avec les GTFs, l'ARN Pol II et les co-facteurs transcriptionnels hautement purifiés (Wu et al., 2003). Il a été montré dans ce système que le Médiateur stimule autant le recrutement de TBP ou TFIID aux promoteurs ainsi que celui de l'ARN Pol II après que l'échafaudage de facteurs d'initiation soit formé. Ainsi, lorsque le système reconstitué est saturé en TBP ou TFIID, le Médiateur augmente spécifiquement le recrutement de la polymérase suite à l'activation mais pas celui des autres GTFs. Ceci montre que le Médiateur fonctionne à une étape subséquente à l'association de TBP ou TFIID au promoteur. À des concentrations basses de TBP ou TFIID, le Médiateur augmente aussi le recrutement de ces deux facteurs en présence de l'activateur.

Plusieurs mécanismes ont été proposés pour expliquer l'activité stimulatrice de Médiateur sur la transcription basale. De façon similaire à ce qui a été observé dans le cas de la transcription activée, le Médiateur peut stabiliser des complexes intermédiaires du PIC. Dans ce sens, des mutants conditionnels des sous-unités essentielles du Médiateur ont diminué le recrutement de la polymérase aux promoteurs de gènes exprimés de façon constitutive mais pas celui des GTFs tels que TFIIE ou TFIIH (Esnault et al., 2008). Un autre mécanisme implique la stimulation de l'activité kinase de TFIIH par laquelle le CTD de la polymérase est phosphorylé lors de l'initiation, ce qui facilite le dégagement du promoteur (Kim et al., 1994). La capacité de stabiliser un échafaudage de GTFs qui restent associés aux promoteurs après l'initiation constitue un mécanisme par lequel le Médiateur

stimule la réinitiation de la transcription (Reeves and Hahn, 2003; Yudkovsky et al., 2000).

Plusieurs sous-unités du Médiateur ont des activités enzymatiques : CDK8 est une kinase pour le CTD de l'ARN Pol II mais aussi pour d'autres facteurs tels que la sous-unité de cycline H de TFIIH et pour PC4 (Akoulitchev et al., 2000); la sous-unité MED5 possède une activité HAT pour les histones H3 soit sous une forme libre ou dans le contexte de la chromatine (Lorch et al., 2000) et finalement la sous-unité MED8 s'associe à un complexe qui présente une activité de type E3 ubiquitine-ligase et peut polyubiquitiner l'ARN Pol II *in vitro*; cependant aucun substrat naturel n'a été encore identifié (Brower et al., 2002).

Le complexe est formé d'environ 20 sous-unités chez la levure (Boube et al., 2002), alors qu'autour de 30 sous-unités ont été identifiées dans plusieurs complexes de type Médiateur chez les mammifères (Myers and Kornberg, 2000). Les différences de composition entre les complexes de type Médiateur chez les mammifères ont été attribuées à l'existence de formes de ce complexe ayant des fonctions distinctes mais aussi à des limitations d'ordre technique, comme par exemple une sensibilité insuffisante, qui auraient empêché l'identification de toutes les sous-unités (Conaway et al., 2005). Les complexes de type Médiateur sont regroupés en deux catégories : (i) les complexes à haut poids moléculaire (1-2 MDa) tels que TRAP (*Thyroid hormone Receptor-Associated Proteins*), DRIP (*Vitamin D Receptor-Interacting Proteins*), NAT (*Negative regulator of Activated Transcription*) et ARC (*Activator-Recruited Complex*) et (ii) les complexes de petit poids moléculaire (500-700kDa) tels que Médiateur, PC2 (*Positive Coactivator 2*) et CRSP (*Cofactor Required for the activation of Sp1*) (revu par (Conaway et al., 2005; Malik and Roeder, 2005)).

Structurellement, le complexe Médiateur est formé de trois modules appelés "tête", "noyau" et "queue". Un quatrième module appelé "kinase" est présent dans certaines formes de Médiateur et inclut des sous-unités kinase (CDK8, CDK11) et

cycline (Cyclin C). Les sous-unités du Médiateur ont été identifiées pour la première fois chez la levure, comme étant des supprimeurs des phénotypes de sensibilité au froid chez les mutants avec une délétion du domaine CTD de l'ARN Pol II et elles ont été appelées Srb (*Suppressor of RNA polymérase B*). Une nomenclature unifiée a été émise et les sous-unités ont été appelées MED1 à MED31. Cette unification est basée sur l'homologie de séquence et de fonction à travers les espèces, malgré que certaines des sous-unités présentes chez la levure (MED2, MED3 et MED5) ne soient pas conservées chez les eucaryotes supérieurs (Boube et al., 2002; Brower et al., 2002; Conaway et al., 2005; Sato et al., 2003).

Le module "tête" interagit avec l'ARN Pol II et représente une surface majeure d'encrage pour le recrutement de celle-ci (Fig. 9). Il a été montré que les activateurs induisent des changements de conformation de ce module, et ceci pourrait favoriser l'interaction avec la polymérase (Lee et al., 1999; Taatjes et al., 2002). Des contacts additionnels des deux autres modules "noyau" et "queue" avec les activateurs et la polymérase favoriseraient encore plus le recrutement de cette dernière (Taatjes et al., 2002). Le module kinase est dissociable et est trouvé uniquement dans des formes de Médiateur de grand poids moléculaire. Ce module présente une activité kinase par laquelle il peut phosphoryler le CTD de la polymérase sur les résidus Ser² et Ser⁵ (Borggreffe et al., 2002). Ces modifications inhibent le recrutement de la polymérase dans le PIC, menant à la conclusion que le module kinase joue un rôle dans la répression transcriptionnelle. Cependant, des études récentes ont montré que CDK8 joue plutôt un rôle de co-activateur et est nécessaire pour l'activation de gènes tel que p21 (Donner et al., 2007). De plus, CDK8 et l'activité kinase ne sont pas nécessaires pour la répression de la transcription et l'effet d'inhibition de la transcription serait dû simplement à l'association compétitive du module kinase et de l'ARN Pol II sur la même surface du Médiateur (Knuesel et al., 2009).

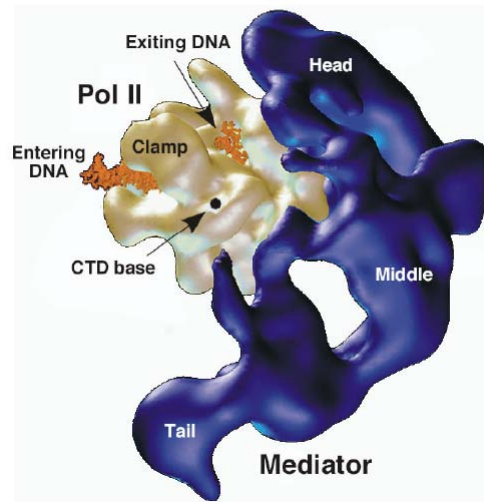


Figure 9. Modèle structural du complexe Médiateur-ARN Polymérase II-ADN. Les modules appelés “tête” (“head”), “noyau” (“middle”) et “queue” (“tail”) du Médiateur (bleu) sont représentés. L’ARN polymérase II (doré) est représentée avec la matrice d’ADN amont et aval. Tiré de Hahn (2004).

Tableau 1

Sommaire des fonctions des facteurs impliqués dans l'assemblage du PIC

Facteurs	Fonctions
L'ARN Pol II	Enzyme responsable de la synthèse des ARNm et de petits ARN non-codants (miARN, snoARN). Possède 12 sous-unités.
TFIIA	Co-activateur. Interagit avec des activateurs et des facteurs de la machinerie transcriptionnelle basale; neutralise la répression transcriptionnelle par des cofacteurs négatifs tels que NC2. Possède 2-3 sous-unités dépendamment de l'espèce.
TFIIB	Impliqué dans la sélection du promoteur; stabilise l'interaction de TFIID avec le promoteur; facilite le recrutement du complexe ARN Pol II/TFIIF au promoteur. Constitué d'un seul peptide.
TFIID	Initie la formation du PIC soit par la liaison du TBP (TATA-box binding protein) aux séquences TATA soit des TAFs (TBP Associated Factors) à d'autres séquences du promoteur; co-activateur. Possède 14 sous-unités, incluant TBP et différentes TAFs.
TFIIE	Facilite le recrutement de TFIIH au promoteur; stimule les activités hélicase et kinase de TFIIH; essentiel pour l'ouverture du promoteur, Possède 2 sous-unités.
TFIIF	S'associe à la L'ARN Pol II avant le recrutement aux promoteurs; augmente l'affinité de l'ARN Pol II pour le complexe TBP-TFIIB-promoteur; stimule l'enroulement de l'ADN du promoteur autour de la polymérase; requis pour le recrutement de TFIIE/TFIIH aux promoteurs; facilite la sélection du site d'initiation de la transcription ainsi que le dégagement du promoteur après l'initiation. Possède 2-3 sous-unités dépendamment d'espèce.

- TFIIH** ATPase/hélicase nécessaire pour l'ouverture ainsi que pour le dégagement du promoteur; activité hélicase impliquée dans la TCR (*Transcription Coupled Repair*), activité kinase impliquée dans la phosphorylation du CTD de la L'ARN Pol II; favorise la transition entre l'initiation et l'élongation. Possède 10 sous-unités.
- TFIIS** Requis pour la formation du PIC et pour le recrutement de tous les facteurs d'initiation, à l'exception de TBP.
- Médiator** Constitue un pont entre les activateurs et la machinerie transcriptionnelle basale; stimule la transcription basale et activée. Possède 20 à 30 sous-unités.
- SAGA** Interagit avec des histones H3, TBP et des activateurs; activité histone acétyltransférase, déubiquitinase. Possède 20 sous-unités
- Mot1** Peut avoir des effets positifs ou négatifs sur la transcription; induit la dissociation du complexe TBP/ADN de manière dépendante d'ATP. Possède une sous-unité.
- NC2** Peut avoir des effets positifs ou négatifs sur la transcription; lie le complexe TBP/ADN et inhibe l'assemblage du PIC. Possède deux sous-unités.

1.3.2 L'élongation

Quoique le recrutement de l'ARN Pol II aux promoteurs soit une étape limitante pour la transcription, les étapes subséquentes telles que le dégagement du promoteur et l'élongation sont également étroitement régulées (Pelechano et al., 2009; Wade and Struhl, 2008). Des facteurs d'élongation variés ont des fonctions spécifiques en fonction du contexte transcriptionnel et leur activité est hautement coordonnée avec celle des facteurs impliqués dans des processus tels que le coiffage, l'épissage, la polyadénylation et l'exportation des transcrits. Ce haut degré de coordination a amené l'hypothèse que l'étape d'élongation de la transcription joue un rôle central parmi les processus nucléaires, pour lesquelles il servirait d'«organisateur» (Perales and Bentley, 2009).

La transition entre l'initiation et l'élongation est vue comme une étape distincte de la transcription, cependant elle n'est pas facilement mise en évidence dans beaucoup de systèmes expérimentaux (Cheng and Price, 2007; Uptain et al., 1997). Le début de la transcription se caractérise par une instabilité manifestée par la formation de courts transcrits qui sont relâchés et une tendance de la polymérase à glisser latéralement et de s'arrêter (Pal and Luse, 2003). Ceci indique qu'un complexe d'élongation de la transcription (TEC) stable n'est pas formé lors de la polymérisation des premiers nucléotides. Cette instabilité persiste jusqu'à ce que le transcrit atteigne environ 20 nucléotides et que les facteurs d'initiation, à l'exception de TFIIF (voir plus bas) perdent le contact avec la polymérase. Les analyses structurales des co-cristaux ARN Pol II-TFIIB amènent une explication possible de ce phénomène. Ces analyses suggèrent que les transcrits plus longs que 7 nucléotides mènent à une interférence stérique avec des domaines de TFIIB et, afin qu'une élongation productive soit possible, le facteur TFIIB doit être relâché (Bushnell et al., 2004; Kostrewa et al., 2009). Après l'éjection de TFIIB, des transcrits plus longs peuvent être accommodés au sein de la polymérase, et lorsqu'ils dépassent 8 nucléotides ils sont dirigés dans le canal de sortie de la

polymérase. Les interactions du transcrit avec ce canal contribueraient à un changement de conformation de la polymérase qui impliquerait la fermeture de la pince mobile sur le duplex ADN-ARN, stabilisant ainsi le complexe en élongation (Gnatt et al., 2001; Pal and Luse, 2003).

L'élongation processive consiste en l'ajout de nucléotides tel que dicté par la matrice d'ADN, dans une réaction de polymérisation qui se trouve en compétition avec plusieurs processus alternatifs. Ces processus peuvent conduire à l'édition du transcrit, la pyrophosphorolyse, la pause, l'arrêt à la terminaison de la transcription (Fig 10A). L'entrée dans un de ces processus est favorisée du point de vue énergétique en fonction de la séquence du brin matrice d'ADN, des interactions de la polymérase avec l'ADN ou avec le transcrit, ainsi qu'avec des facteurs d'élongation (Greive and von Hippel, 2005).

Dans les dernières années, des techniques d'analyse des molécules uniques, comme par exemple la microscopie à force atomique (AFM) et la résonance de transfert d'énergie entre molécules fluorescentes (FRET), ont permis de mettre en évidence des phénomènes difficilement observables dans des populations hétérogènes de molécules, telles qu'utilisées dans les analyses biochimiques. Ainsi, il a été montré que, lors des premières étapes de l'élongation, le brin d'ADN matrice est plié dans le centre actif puisque la polymérase tire la partie en aval de l'ADN, ce qui mène à l'excision de dinucléotides du transcrit. L'énergie accumulée de cette façon par la polymérase serait utilisée pour dégager le promoteur et pour entrer en élongation productive (Kapanidis et al., 2006; Revyakin et al., 2006). D'autres phénomènes reliés à l'élongation, observés par ce type de méthodes, incluent des modifications ultrastructurales telle que la courbure induite par la polymérase en élongation à l'ADN (Dame et al., 2003) (Fig. 10B), ainsi que l'existence de sites de pause de durée très courte, appelées pauses élémentaires, qui interrompent fréquemment la transcription. Ce dernier phénomène a été observé même dans des gènes qui n'ont pas de sites de pause définis, et il a été suggéré qu'il servirait de précurseur pour des pauses plus longues qui régulent l'élongation (Artsimovitch and Landick, 2000).

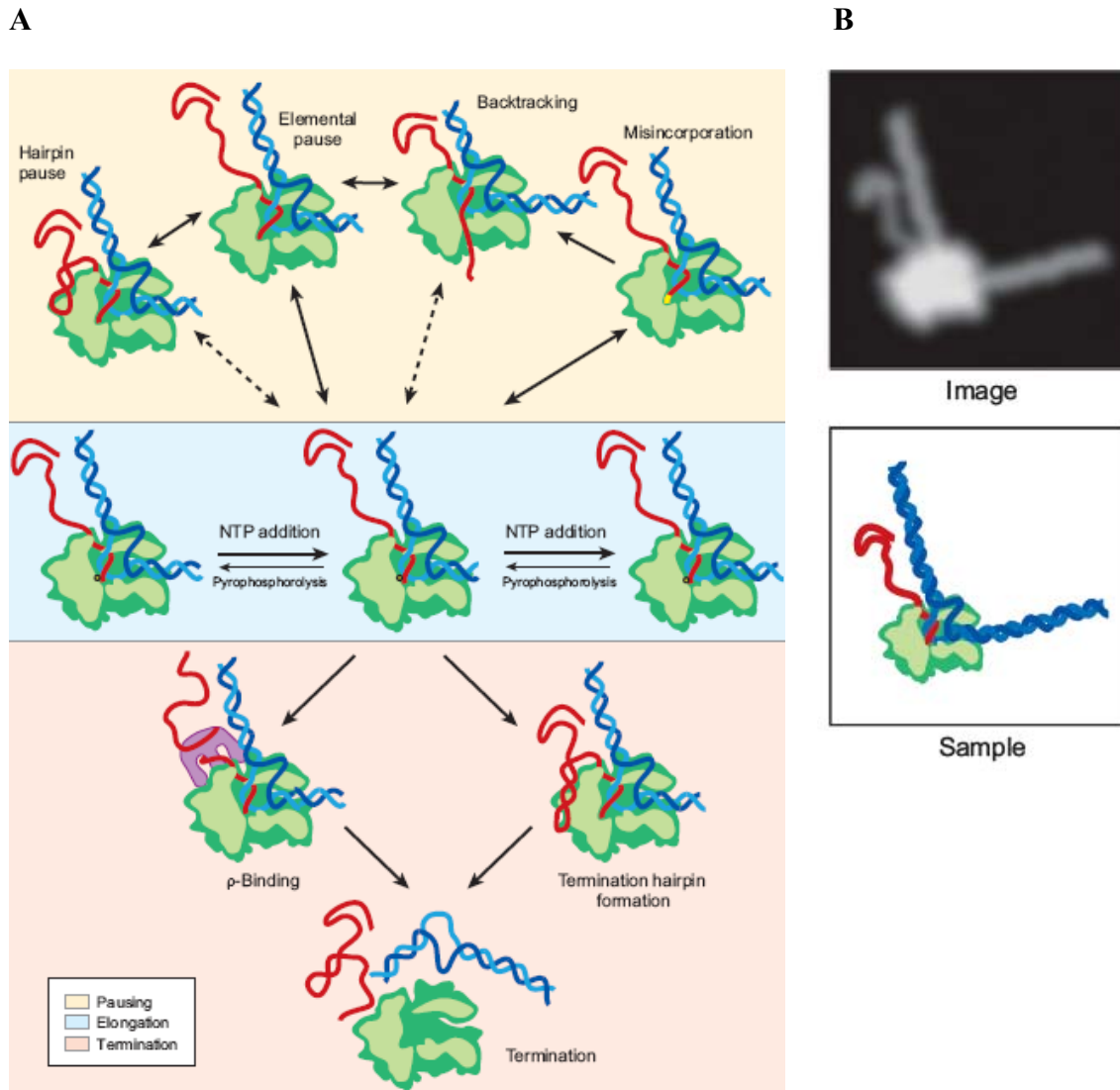


Figure 10. Modèle pour l'équilibre entre l'élongation et les processus alternatifs. A) L'ajout de NTPs lors de l'élongation est en équilibre avec le processus inverse de pyrophosphorolyse (centre). Les états alternatifs de pause (haut) peuvent être directement en équilibre avec l'état d'élongation ou peuvent résulter de l'état de pause élémentaire. Ces états de pause peuvent être stabilisés par des structures secondaires du transcrit d'ARN ou lorsque la polymérase est reculée sur la matrice d'ADN. Lorsqu'un nucléotide est mal apparié avec la matrice d'ADN, la polymérase peut reculer sur la matrice. La terminaison (bas) se produit lorsque la polymérase rencontre des séquences qui génèrent des structures en épingle à

cheveux dans le transcrit, ou, chez les procaryotes, lorsque le facteur Rho s'associe au transcrit d'ARN. Code de couleurs: ARN Pol II - vert, ADN – bleu, ARN-rouge, nucléotide mal apparié – jaune. B) Image d'un complexe d'élongation actif (TEC) formé par une seule molécule d'ARN Pol II, d'ADN et d'ARN. Haut : image obtenue par microscopie AFM (*Atomic Force Microscopy*) qui montre la courbure induite à l'ADN par la polymérase en élongation à l'ADN. L'image à deux dimensions représente les modifications de position d'un scanner à lumière laser. Ce scanner passe sur une surface plate sur laquelle un complexe TEC est déposé. Bas : schéma avec le code de couleurs comme dans A). Tiré de Herbert et al., 2008.

Par le processus d'édition, un nucléotide mal apparié avec la matrice d'ADN est excisé du transcrit, alors que la pyrophosphorolyse est la réaction inverse de la polymérisation, dans laquelle le dernier nucléotide incorporé dans le transcrit est excisé. D'autres états inaptes pour l'élongation sont atteints lorsque la polymérase est arrêtée, reculée sur la matrice ou simplement en état de pause (Fig. 10A). Deux types de pauses peuvent être distinguées : de courte durée ou élémentaires, lorsque la polymérase arrête l'incorporation de nucléotides dans le transcrit mais elle peut reprendre la polymérisation par elle-même; et des pauses stabilisées, ou de longue durée lorsque la polymérase est reculée sur la matrice ou lorsque l'ARN forme des structures secondaires (Herbert et al., 2008). Ces états alternatifs auraient le rôle de moduler la production des ARN et de permettre le recrutement dans le TEC de facteurs régulateurs (Artsimovitch and Landick, 2002; Ring et al., 1996).

Il a été montré que les pauses de courte durée se produisent fréquemment et avec une probabilité assez constante à chaque position donnée sur la matrice et il a été proposé qu'elles impliquent plutôt un réarrangement structural de la polymérase qu'un reculement sur la matrice (Neuman et al., 2003). Une étude cinétique des molécules de polymérase et d'ARN transcrits, marqués *in vivo* et mesurés par FRAP (*Photoactivation and Recovery of Fluorescence After Photobleaching*) en temps réel pendant leur synthèse suggère que la polymérase est plus processive que précédemment évalué (70 bases par seconde comparativement à environ 30 bases

par seconde) mais elle fait des pauses fréquentes pendant l'élongation (Darzacq et al., 2007). Plusieurs facteurs qui régulent l'état de pause ou la vitesse de transcription (“*transcriptional rate*”) ont été caractérisés : TFIIF, P-TEFb (*Positive-Transcription Elongation Factor b*), DSIF (*DRB Sensitivity-Inducing Factor*), NELF (*Negative Elongation Factor*), ELL et les Élongines (revu par (Sims, III et al., 2004a)).

Dans les dernières années un nombre croissant d'études ont montré que la régulation de l'élongation est un mécanisme majeur de régulation de la transcription, qui est beaucoup plus courant que l'on croyait initialement (revu par (Core and Lis, 2008; Margaritis and Holstege, 2008). Le recrutement de l'ARN Pol II était considéré auparavant comme la principale étape limitante de la transcription (Ptashne and Gann, 1997). Dans ce sens, il a été observé que chez la levure, la présence de l'ARN Pol II dans la région 5' des gènes est corrélée avec leur niveau d'expression, ce qui indique une régulation au niveau du recrutement (Robert et al., 2004). La présence de la polymérase engagée dans l'élongation et arrêtée dans les régions en aval, proximales au promoteur, était vue comme étant exceptionnelle (ex. le gène *HSP70* chez la *D. mélanogaster*) (Rougvie and Lis, 1988). Cependant, l'analyse à grande échelle de la localisation génomique de l'ARN Pol II combinée aux données d'expression génique a mis en évidence la présence de la polymérase dans les régions 5' de 10 à 30% de gènes, même ceux qui n'étaient pas exprimés. Ce type d'observation a été fait chez la levure en phase stationnaire (Radonjic et al., 2005) ainsi que chez l'humain et chez *D. mélanogaster*, où des complexes de préinitiation associés dans les régions 5' d'un grand nombre de gènes inactifs a été rapporté (Guenther et al., 2007; Kim et al., 2005a; Muse et al., 2007). Ces études suggèrent que la régulation de la transcription de ces gènes se fait à un stade ultérieur au recrutement du PIC. La cartographie par un traitement au permanganate identifiant la présence de la bulle transcriptionnelle en aval du promoteur (Muse et al., 2007; Zeitlinger et al., 2007), ainsi que l'étude des modifications post-traductionnelles des histones mettant en évidence la présence de marques d'initiation (ex: la triméthylation des lysines 4 sur les histones H3) (Guenther et al., 2007), appuient la

conclusion que la polymérase arrêtée dans les régions 5' des gènes est engagée en transcription. Une étude récente à haute résolution de la localisation génomique de l'ARN Pol II chez la levure suggère que le recrutement de la polymérase constitue une étape limitante pour la transcription d'un grand nombre de gènes (17%) (Venters and Pugh, 2009). Cette étude révèle aussi que, dans un nombre équivalent de gènes (18%), ce serait les étapes initiales d'élongation et l'arrêt de la polymérase immédiatement en aval du promoteur qui détermineraient les niveaux d'expression.

1.3.2.1 Les facteurs qui modulent les pauses de l'ARN Pol II

1.3.2.1.1 TFIIIF

Des études biochimiques *in vitro* ont montré que le rôle du facteur TFIIIF dans la transcription n'est pas limité à l'initiation, mais il est également impliqué dans l'élongation. Des analyses *in vitro* des complexes d'élongation fonctionnels formés sur des matrices d'ADN immobilisées ont montré que TFIIIF est présent dans les complexes ternaires formés par la polymérase en élongation, l'ADN et le transcrit d'ARN (Cheng and Price, 2007; Zawel et al., 1995). Des études de compétition de matrices *in vitro*, dans lesquelles l'une des matrices d'ADN a été pré-incubée avec l'ensemble complet de GTFs alors que l'autre a été incubée avec les mêmes facteurs, mais sans TFIIIF, ont montré que TFIIIF est requis pour la reprise de l'élongation lorsque la polymérase est arrêtée et qu'il peut se dissocier et se réassocier avec la polymérase en élongation (Zawel et al., 1995). L'analyse par immunoprécipitation de la chromatine de quelques régions transcrites chez la levure n'a pas permis de détecter la présence *in vivo* de TFIIIF dans les complexes d'élongation (Krogan et al., 2002; Pokholok et al., 2002a). Cependant, étant donné qu'un nombre très limité de régions transcrites ont été étudiées et prenant en considération que les études *in vitro* ont suggéré que TFIIIF ne voyage pas avec la polymérase en élongation, mais plutôt il se dissocie et se réassocie aux sites de

pause, une conclusion sur la fonction de TFIIF lors de l'étape d'élongation *in vivo* ne peut pas être formulée.

En utilisant un système reconstitué *in vitro* avec des facteurs hautement purifiés, il a été montré que TFIIF et NELF fonctionnent de façon compétitive dans le contrôle de l'élongation (Renner et al., 2001). Il a été proposé que TFIIF et NELF lient le même domaine de l'ARN Pol II et, afin que TFIIF puisse interagir avec la polymérase, NELF doit être relâché. Ce modèle expliquerait pourquoi le complexe d'élongation est résistant à l'action stimulatrice de TFIIF avant que pTEFb ne phosphoryle NELF et favorise ainsi sa dissociation de la polymérase (Cheng and Price, 2007).

1.3.2.1.2 P-TEFb

P-TEFb (*Positive-Transcription Elongation Factor b*) est un complexe formé par la kinase CDK9 (*Cyclin-Dependent Kinase 9*) et une cycline. Chez l'humain, CDK9 peut lier une des cyclines T1, T2a, T2b ou K (Cyc T1, Cyc T2 ou CycK). alors que chez *Drosophila* il lie une seule cycline (CycT) (revu par (Peterlin and Price, 2006)). CDK9 est une serine/thréonine kinase de la famille cdc2 qui présente un grand degré d'homologie avec d'autres kinases dépendantes de cyclines, telles que cdc2, cdk2, cdk3 et cdk5 (De et al., 1997; Grana et al., 1994). Cependant CDK9 n'a pas de fonction dans le cycle cellulaire et a été classifiée comme membre de la famille de kinases cdc2 sur la base de son homologie avec le motif PSTAIRE de cdc2 (Grana et al., 1994). Aussi, bien que CDK9 soit exprimé de façon ubiquitaire, il a été rapporté que dans certaines cellules différenciées, comme par exemple les neurones et les lymphocytes, ses niveaux d'expression sont plus élevés, ce qui suggère des fonctions spécifiques dans ces cellules (Marshall and Grana, 2006).

Suite à l'initiation transcriptionnelle, l'ARN Pol II a tendance à s'arrêter à une distance entre 20 et 40 nucléotides en aval du promoteur. Des études *in vitro* ont

apporté des indices sur les facteurs qui contribuent à ce phénomène, soit les séquences d'ADN, la longueur du transcrit d'ARN (Pal et al., 2001; Ujvari et al., 2002), ainsi que la présence des facteurs d'initiation, tel TFIIB, qui ne sont pas encore relâchés (Kostrewa et al., 2009; Tran and Gralla, 2008). Il a été proposé que ces facteurs contribuent à un changement de conformation de la polymérase qui lui permettrait d'entrer en élongation productive, tel que présenté au début de cette section. Les facteurs négatifs d'élongation DSIF et NELF auraient le rôle de stabiliser la polymérase à ces sites d'arrêt (Peterlin and Price, 2006). Ces deux facteurs s'associent physiquement à la polymérase et la présence de chacun est requise pour que leur action inhibitrice se manifeste (Renner et al., 2001).

P-TEFb est nécessaire pour l'expression de la majorité des gènes, à l'exception de certains gènes courts et qui n'ont pas d'introns, tels que les gènes qui codent pour le snARN U2 ou l'histone H2B (Medlin et al., 2005). En effet, l'inhibition *in vivo* de l'activité de P-TEFb par des agents pharmacologiques tels que le flavopiridol mène à une diminution globale de la transcription (Chao and Price, 2001). L'analogue d'ATP 5,6-dichloro-1- β -D-ribofuranosylbenzimidazole (DRB) inhibe la transition entre les étapes d'initiation et d'élongation productive de la transcription et il a été montré par des essais de type "run-off" que cet effet est dû à l'inhibition de P-TEFb et non pas de la polymérase elle-même (Marshall and Price, 1995). Dans la même étude, il a été démontré que P-TEFb est un facteur limitant pour l'entrée de la polymérase en élongation productive. Cependant, des essais d'élongation *in vitro* ont montré qu'il n'affecte pas la vitesse d'élongation et son effet sur la transcription est plutôt dû à la régulation de l'activité d'autres facteurs (Renner et al., 2001).

P-TEFb régule l'étape d'élongation de la transcription en phosphorylant le CTD de l'ARN Pol II sur la sérine 2 (Garriga and Grana, 2004) de même que les sous-unités Spt5 et RD des facteurs DSIF et NELF, respectivement (Fujinaga et al., 2004; Kim and Sharp, 2001). Ces modifications mènent au relâchement de l'effet inhibiteur des facteurs DSIF et NELF sur l'élongation, étant donné qu'une fois phosphorylé, DSIF est converti d'un facteur négatif en un facteur positif d'élongation (Yamada et al.,

2006), alors que NELF est dissocié de la polymérase (Fujinaga et al., 2004; Peterlin and Price, 2006). Chez l'humain, c'est la sous-unité CDK9 qui phosphoryle les trois cibles mentionnées. Chez la levure, deux homologues fonctionnels de CDK9 sont impliqués : Ctk1 (*CTD kinase1*) et Bur2 phosphorylent le CTD de la polymérase (Keogh et al., 2003; Patturajan et al., 1999), alors que la kinase Bur2 est responsable de la phosphorylation de la sous-unité Spt5 de DSIF (Zhou et al., 2009). NELF n'a pas d'homologue chez la levure. La phosphorylation de la sérine 2 du CTD de la polymérase permet le couplage de la transcription avec les processus de coiffage et de polyadénylation des ARNm (Ahn et al., 2004; Ni et al., 2004). Ctk1 a un rôle dans la transition entre l'initiation et l'élongation. Il a été proposé qu'il régule la dissociation des facteurs généraux de transcription qui restent au promoteur pour former un échafaudage de la polymérase qui entre en élongation. Des essais de transcription *in vitro* avec des matrices immobilisées, combinés à des essais d'immunoprécipitation de la chromatine ont montré chez une souche *Ctk1Δ* que les niveaux des facteurs détectés normalement au promoteur dans l'échafaudage sont diminués et qu'ils sont plutôt localisés tout au long des gènes, tout comme la polymérase (Ahn et al., 2009).

Plus récemment, un rôle de P-TEFb dans l'expression des gènes inductibles par des cytokines a été décrit. Le mécanisme de régulation de la transcription pour ces gènes diffère de celui des gènes de maintenance: l'ARN Pol II est recrutée dans des complexes de transcription par des facteurs de transcription inductibles et n'est pas détectée dans leurs régions 5' avant l'induction par des cytokines (revu par (Brasier, 2008)). Ceci indique qu'avant l'induction, la polymérase n'est pas déjà recrutée au promoteur ou en pause dans les régions proximales de celui-ci, un état dans lequel elle serait stimulée par P-TEFb pour entrer en élongation. L'induction de l'expression des gènes cibles par les voies TNF (*Tumor Necrotic Factor*), NFκB (*Nuclear Factor Kappa-B*) ou IL-6 (Interleukin 6)-Jak(*Janus kinase 3*), STAT3 (*Signal Transducer and Activator of Transcription 3*) implique l'interaction dans le nucléoplasme des facteurs NFκB ou STAT3 activés et pTEFb (Hou et al., 2007; Nowak et al., 2008). Ces interactions sont importantes afin de cibler CDK9 à des

gènes inductibles par ces voies. Ainsi, l'expression des gènes γ FBG (*gamma fibrinogen*) et p21waf est dépendante de P-TEFb, tel que démontré par le atténuation de l'expression de CDK9 par siRNA. De plus, dans le cas du gène γ FBG, le recrutement de l'ARN Pol II suite à l'induction par un traitement à l'IL-6 dépend de la présence de CDK9 actif sur ce gène. Cette dépendance a été démontrée en utilisant un inhibiteur de l'activité kinase de CDK9, le flavopiridol, et en analysant l'impact sur la localisation de la polymérase le long du gène γ FBG par immunoprécipitation de la chromatine (Hou et al., 2007).

Le recrutement de P-TEFb dans les complexes transcriptionnels se fait par plusieurs mécanismes : par des activateurs liés à l'ARN, comme l'activateur Tat du virus HIV (*human immunodeficiency virus*) lié à la séquence TAR (*trans-acting response element*) du transcrite de HIV-1 (Wei et al., 1998); par des activateurs liés à l'ADN, par exemple c-Myc (Kanazawa et al., 2003), VP16 (Kurosu and Peterlin, 2004), des récepteurs nucléaires (Lee et al., 2001), NF- κ B (Barboric et al., 2001), etc; par des co-activateurs comme CIITA (*class II transactivator*) (Kanazawa et al., 2000); et finalement, par des activateurs liés à la chromatine, par exemple Brd4 (*bromodomain containing 4*) qui lie les histones acétylées (Jang et al., 2005). Ces interactions sont régulées entre autre par des modifications post-traductionnelles; par exemple, il a été montré que l'ubiquitination de VP16 augmente son affinité pour la sous-unité Cycline T1 de P-TEFb (Kurosu and Peterlin, 2004).

Brd4 possède deux domaines qui reconnaissent les histones acétylées (bromodomains) et c'est par ces domaines qu'il interagit également avec P-TEFb (Yang et al., 2005). Ainsi, des modifications post-traductionnelles au niveau de la chromatine mènent à l'activation de l'élongation par une cascade d'interactions. Par exemple, suite à un seul événement de phosphorylation au niveau de l'enhancer du gène FOSL1, une cascade d'interactions mène au recrutement de P-TEFb et à l'activation de l'élongation. La stimulation des cellules avec du sérum active la kinase PIM1 qui phosphoryle la sérine 10 des histones H3. Cette forme des histones H3 est reconnue par la histone-acétyltransférase MOF qui acétyle la lysine 16 de

l'histone H4; les histones H4 acétylées sont reconnues par Brd4 qui recrute à son tour le facteur P-TEFb (Zippo et al., 2009). Il a été montré que Brd4 est requis pour l'expression des gènes importants pour la transition entre les phases G1 et S du cycle cellulaire par un mécanisme qui implique le recrutement de CDK9 (Mochizuki et al., 2008), ce qui suggère que P-TEFb a un rôle dans la progression du cycle cellulaire même si son niveau d'expression ne varie pas pendant les différentes phases du cycle cellulaire (Garriga et al., 2003).

En plus de son rôle dans la transcription, P-TEFb joue aussi un rôle dans l'export des ARNm et la traduction, puisque Ctk1, sa sous-unité kinase chez la levure, interagit avec le complexe TREX (*Transcription and mRNA Export*) et les polyribosomes, phosphoryle la protéine ribosomale Rsp2 et est requise pour la traduction efficace (Rother and Strasser, 2007).

L'activité de P-TEFb est régulée par des mécanismes multiples (revu par (Peterlin and Price, 2006). Un mécanisme majeur consiste en sa séquestration dans un complexe formé avec le petit ARN 7SK et les protéines Hexim1/2 et LARP7. Sous cette forme, l'activité kinase de CDK9 est inhibée (Byers et al., 2005; Krueger et al., 2008; Michels et al., 2004; Yang et al., 2001). Dans les cellules, il existe un équilibre entre cette forme inactive en complexe avec 7SK, Hexim1/2 et LARP7 et la forme active dans laquelle CDK9 lie une cycline T (Krueger et al., 2008; Markert et al., 2008). Dans certains types cellulaires, CDK9 se trouve en complexe avec des protéines "heat shock", comme dans les cellules sanguines mononucléaires quiescentes où la cycline T1 n'est que faiblement exprimée (Rice and Herrmann, 2003). CDK9 est également régulé par des modifications post-traductionnelles : une phosphorylation de la thréonine 186 située dans la boucle T est nécessaire pour l'activité kinase de CDK9 (Li et al., 2005); une autophosphorylation de la région C-terminale est requise pour l'association avec l'activateur Tat et la séquence TAR (Ammosova et al., 2005). Une acétylation de CDK9 a été décrite avec des effets opposés en fonction des acétyltransférases impliquées : GNAT, P/CAF et GCN5 acétylent les lysines 44 et 48, ce qui inhibe l'activité kinase de CDK9 (Sabo et al.,

2008), alors que p300/CBP acétyle la lysine 44, ce qui mène à une augmentation de son activité kinase (Fu et al., 2007). Une ubiquitination de CDK9 a également été observée et l'effet de cette modification dépend du niveau d'expression de CDK9. Ce point sera discuté au Chapitre 4.

CDK9 est impliqué dans de nombreux processus physiologiques tels que la croissance cellulaire et la différenciation, mais aussi dans une multitude de processus pathologiques allant de l'hypertrophie du myocarde à différents types de cancer (revu par (Wang and Fischer, 2008). Des inhibiteurs de CDK9, tels que le flavopiridol et la roscovitine, ont montré un effet clinique prometteur dans certains types de cancers (par exemple la leucémie lymphocytaire chronique) et comme traitement antiviral du HIV-1, bien qu'ils ne soient pas entièrement spécifiques et qu'ils présentent une inhibition croisée d'autres CDKs (Christian et al., 2007). L'effet anti-tumoral observé serait dû à l'inhibition de la transcription des gènes anti-apoptotiques par l'inhibition de l'activité kinase de CDK9 pour le CTD de l'ARN Pol II (Chen et al., 2005). Une fenêtre thérapeutique étroite reste tout de même un obstacle majeur pour l'utilisation thérapeutique de ces inhibiteurs, étant donné que P-TEFb est important pour la transcription d'une majorité de gènes.

1.3.2.2 Les facteurs qui régulent la polymérase arrêtée

1.3.2.2.1 TFIIIS

Lorsque la polymérase est en état d'arrêt ou reculée, malgré qu'elle reste active du point de vue catalytique, elle ne peut reprendre la polymérisation qu'en présence de facteurs d'élongation (Conaway et al., 2000) comme le facteur TFIIIS. En effet, celui-ci stimule l'activité endonucléolytique de la polymérase et favorise l'excision de l'extrémité en 3' du transcrit qui n'est plus alignée sur le centre catalytique, comme déjà présenté dans la section 1.2.2.

TFIIS possède trois domaines fonctionnels (Fig. 11). Le domaine I, situé dans la région N-terminal, a une structure de ruban à quatre hélices. Ce domaine est divergeant entre les espèces, et n'est requis ni pour l'activité de stimulation de la polymérase ni pour l'interaction de TFIIS avec celle-ci (Booth et al., 2000). Cependant, il est essentiel à la stimulation de l'initiation (Kim et al., 2007) et il permet l'interaction de TFIIS avec l'ARN Pol II sous forme d'holoenzyme (Pan et al., 1997). Le domaine II forme un ruban à trois hélices, alors que le domaine III présente un motif à doigts de zinc qui a une structure secondaire en épingle à cheveux de type β . Entre les domaines II et III se trouve un domaine appelé "linker" qui a une structure désordonnée. Le "linker" et le domaine II sont importants pour l'interaction avec la polymérase. Le domaine III est essentiel pour la stimulation de l'activité de clivage de la polymérase. (Awrey et al., 1998; Olmsted et al., 1998)

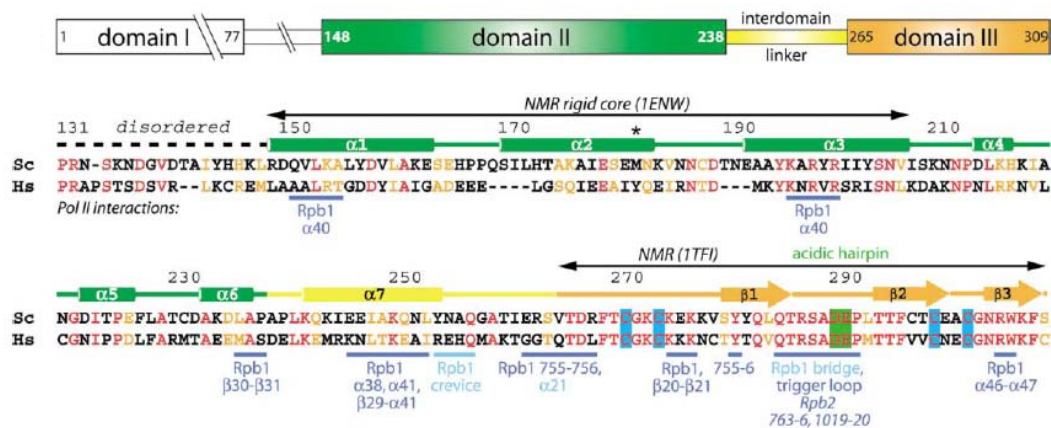


Figure 11. Les domaines fonctionnels de TFIIS et l'alignement des séquences des homologues chez l'humain et chez la levure. Les structures secondaires sont indiquées en haut de la séquence et les acides aminés identiques ou conservés sont indiqués respectivement en rouge ou en orange. Les quatre résidus cystéine qui coordonnent un atome de zinc sont représentés sur un fond bleu et les deux acides aminés invariants D290 et E291 sur un fond vert. Les séquences impliquées dans les interactions avec les différents éléments structuraux de l'ARN Pol II sont soulignées. Tiré de Kettenberger et al (2003).

Le gène TFIIIS chez la levure, appelé *PPR2*, a été cloné pour la première fois comme gène régulateur de la voie de biosynthèse de la pyrimidine (Hubert et al., 1983). Ultérieurement, il a été identifié comme gène codant pour une protéine impliquée dans le processus de transfert des brins d'ADN dans la recombinaison des homologues pendant la méiose et appelé *DST1 (DNA strand transfer 1)* (Clark et al., 1991). Cependant, aucune fonction spécifique dans ce processus n'a encore été rapportée.

Tandis que TFIIIS est exprimé de façon ubiquitaire, des protéines appartenant à la famille de TFIIIS sont exprimés dans un nombre limité de tissus, ce qui suggère un rôle tissulaire spécifique. Le gène TFIIIS.1 appartenant à cette famille possède des introns alors que TFIIIS est un gène sans intron. Il a été proposé que TFIIIS.1 serait le gène ancestral à partir duquel TFIIIS a évolué (Weaver and Kane, 1997).

L'ARN Pol II possède une activité endonucléolitique intrinsèque qui est défavorisée en conditions normales, mais qui peut être stimulée par un pH élevé ou par des facteurs tels que GreA et GreB chez la bactérie (Weilbaecher et al., 2003). Ces facteurs sont les homologues fonctionnels et non structuraux du facteur TFIIIS chez les eucaryotes (Kettenberger et al., 2003). Autant TFIIIS que les facteurs Gre s'insèrent dans le canal secondaire de la polymérase, appelé aussi "pore". Ils modifient l'activité catalytique de la polymérase en contribuant à la coordination du métal B par deux acides aminés strictement conservés (Kettenberger et al., 2003; Opalka et al., 2003; Sosunova et al., 2003). Des homologues fonctionnels des facteurs Gre bactériens ont été identifiés chez toutes les espèces eucaryotes, ainsi que chez les *archaea*, ce qui suggère que leur fonction est importante pour la cellule. Cependant, ces facteurs ne sont pas essentiels chez la plupart des espèces incluant les eucaryotes inférieurs car leur délétion n'est pas létale. La délétion du *DST1* génère des phénotypes de sensibilité au froid ou au 6-azauracil chez *S. cerevisiae* (Lennon, III et al., 1998). La sensibilité au 6-azauracil est considérée comme étant spécifique aux défauts d'élongation, puisqu'il inhibe l'orotidylate

décarboxylase qui est l'enzyme responsable de la synthèse de GTP et UTP (Exinger and Lacroute, 1992). De la même manière, dans des essais d'élongation *in vitro*, il a été observé qu'une concentration basse en nucléotides mène à l'arrêt de la polymérase (Sims, III et al., 2004b). De plus, il a été montré que TFIIS régule l'expression *in vitro* du gène *SSMI* qui contient des sites d'arrêt de la transcription, en stimulant la reprise de l'élongation à ces sites d'arrêt. Puisque la protéine codée par ce gène est impliquée dans la résistance au 6-azauracil, il a été suggéré que TFIIS a aussi un rôle dans cette résistance (Shimoaraiso et al., 2000).

La délétion de TFIIS chez la souris est létale car elle mène au blocage de la différenciation des érythroblastes pendant le développement embryonnaire (Ito et al., 2006). Il a été proposé qu'une augmentation de l'apoptose, due à une diminution de l'expression du gène qui code pour le facteur anti-apoptotique *Bcl-xL*, en serait la cause (Nagata et al., 2009). En effet, ce gène a un site d'arrêt de la transcription dans son exon 2 et TFIIS permettrait la transcription au-delà de ce site.

TFIIS stimule la transcription en diminuant les pauses et en rendant la polymérase plus processive, c'est-à-dire qu'il augmente le nombre de transcrits par nombre d'initiations. Cependant, il n'affecte pas la vitesse de transcription (Bengal et al., 1991; Reinberg and Roeder, 1987). La polymérase en élongation peut rencontrer différents obstacles : des séquences d'ADN spécifiques tels que des régions riches en AT qui induisent une courbure dans l'ADN, des protéines ou d'autres molécules qui lient l'ADN, des bases modifiées (Izban and Luse, 1992; Kerppola and Kane, 1990; Ueno et al., 1992). Lorsqu'elle ne peut plus transloquer dans la direction de la transcription, la polymérase a tendance à reculer sur la matrice et s'arrêter, situation dans laquelle l'ajout de TFIIS stimule le clivage de fragments de 7 à 9 nucléotides et la reprise de l'élongation (Gu and Reines, 1995). TFIIS peut stimuler le clivage du transcrit autant par des polymérases en pause que arrêtées, avec la libération de fragments de transcrit plus courts lorsque la polymérase est en pause (Shilatifard et al., 2003). Une nouvelle fonction de TFIIS a été proposée par Xiong et

Burton, dans l'isomérisation de la polymérase et l'éjection du NTP au site +1 lorsque la translocation est bloquée (Xiong and Burton, 2007).

Des interactions génétiques entre TFIIIS et des kinases pour le CTD de la polymérase (Ctk1, Kin28 ou Bur1) ont été observées. Il a été proposé que la perturbation de la fonction du CTD augmente la fréquence des arrêts de la polymérase *in vivo* et, par conséquent, une dépendance accrue à TFIIIS est observée (Lindstrom and Hartzog, 2001). Il a également été observé que les défauts de transcription des mutants de TFIIIS sont augmentés de façon synergique quand ils sont combinés à des mutations des sous-unités de la polymérase ou des autres facteurs d'élongations tels que FACT et Elongator (revu par (Shilatfard et al., 2003)). Il a été proposé que la sous-unité Rpb9 de l'ARN Pol II augmente la capacité de TFIIIS à stimuler le clivage des transcrits par la polymérase (Koyama et al., 2007).

Tel que décrit au début de cette section, pour un grand nombre de gènes, la transition entre l'initiation et l'élongation est une étape importante de régulation de la transcription. En effet, une fois que l'ARN Pol II libère le promoteur elle a tendance à entrer en pause après la synthèse de courts transcrits. Les pauses sont stabilisées par les facteurs négatifs DSIF et NELF. En utilisant un système reconstitué *in vitro* avec des complexes d'élongation arrêtés sur des matrices avec des sites de pause de différentes forces, Palangat et ses collaborateurs ont montré que DSIF et NELF inhibent l'activité stimulatrice de TFIIIS sur l'élongation (Palangat et al., 2005). Basé sur des essais de compétition, ils établissent que cet effet est dû à l'inhibition de l'interaction de TFIIIS avec la polymérase et ceci indépendamment de la longueur du transcrit. Cette conclusion est appuyée par plusieurs observations, entre autres en augmentant la concentration de TFIIIS l'effet inhibiteur de DSIF/NELF sur la polymérase est annulé. Un modèle est proposé (Fig. 12) dans lequel TFIIIS ne peut interagir et stimuler la polymérase que suite à l'action de P-TEFb sur le complexe d'élongation arrêté et suite à la phosphorylation des facteurs DSIF et NELF. Ainsi, la phosphorylation de NELF mène à sa

dissociation du complexe d'élongation, situation dans laquelle TFIIS peut lier la polymérase, stimuler le clivage du transcrit et la reprise de l'élongation.

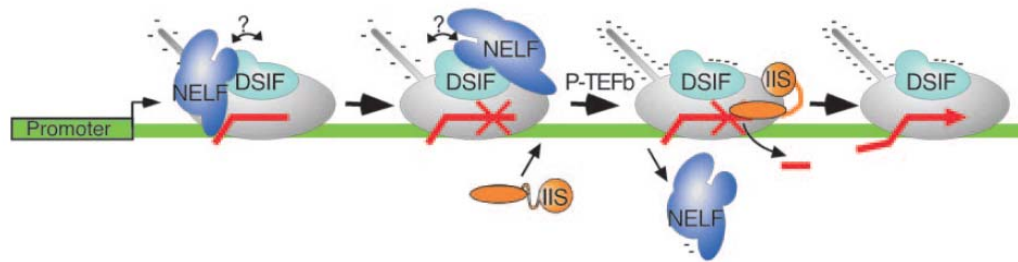


Figure 12. Modèle du mécanisme par lequel les facteurs négatifs d'élongation DSIF et NELF inhibent TFIIS. DSIF recrute NELF dans le complexe d'élongation formé suite au dégagement du promoteur et NELF lie la polymérase possiblement à plus d'un site (la flèche bidirectionnelle). La polymérase est en pause dans les régions proximales au promoteur aussi longtemps que DSIF et NELF y sont associés. TFIIS ne peut lier la polymérase que lorsque NELF est dissocié suite à sa phosphorylation par P-TEFb. TFIIS stimule la polymérase qui clive le transcrit et reprend l'élongation productive. Tiré de (Palangat et al., 2005)

Lorsque des nucléotides sont mal appariés dans le site +1, le facteur TFIIS stimule l'activité d'édition dans des systèmes reconstitués *in vitro* (Jeon and Agarwal, 1996). Cependant, des études génétiques chez la levure ont mené à des résultats contradictoires. Des essais de fidélité de la transcription *in vivo* n'ont pas montré un effet significatif de TFIIS sur les erreurs de transcription (Nesser et al., 2006; Shaw et al., 2002), alors que des mutations du gène *DST1* diminuant la capacité de TFIIS à stimuler le clivage des ARN ont aussi diminué la fidélité de la transcription (Koyama et al., 2007). De plus, Koyama et ses collaborateurs ont rapporté qu'une double mutation de *DST1* et *Rpb9* entraîne une diminution encore plus marquée de la fidélité de la transcription que le mutant *DST1*.

Dans les dernières années, plusieurs études ont mis en évidence le rôle de TFIIS dans l'initiation de la transcription (revu dans (Sikorski and Buratowski, 2009)). Les approches classiques pour l'étude des facteurs d'initiation de la transcription ont échoué à identifier cette fonction. Ceci est probablement dû au fait que TFIIS n'affecte pas l'initiation via des interactions avec des séquences d'ADN spécifiques dans des régions régulatrices (Shilatifard et al., 2003). Un criblage par spectrométrie de masse a mis en évidence la présence de TFIIS dans les complexes de préinitiation (Kim et al., 2007). Dans la même étude, il a été montré que TFIIS est requis pour la formation du PIC et qu'en son absence le recrutement de tous les facteurs d'initiation, à l'exception de TBP est diminué. L'étude du recrutement du PIC *in vivo*, par immunoprécipitation de chromatine, montre que la délétion du TFIIS affecte le recrutement de TBP et de la polymérase au promoteur du gène *GALI* mais pas celui des co-activateurs SAGA et Médiateur (Prather et al., 2005). De plus, la délétion de *DSTI* est synthétique létale avec celle des sous-unités des complexes *Médiateur* et *Swi/Snf* ce qui vient à l'appui d'un rôle de TFIIS dans l'activation de la transcription (Davie and Kane, 2000; Malagon et al., 2004).

1.3.3 La terminaison

Chez les eucaryotes les gènes se succèdent au long des deux brins d'ADN et la distance moyenne entre deux gènes adjacents n'est que quelques centaines de nucléotides. Dépendamment de l'orientation des gènes, un promoteur peut être adjacent ou même superposé à des séquences spécifiques de terminaison. Ceci implique un contrôle très étroit de la terminaison de la transcription, puisque si la polymérase engagée sur un des gènes continue la lecture au-delà des séquences de terminaison, le promoteur du gène suivant sera bloqué si les gènes sont dans le même sens. Si les gènes sont dans le sens opposé, un ARN anti-sens ou une collision des polymérases peuvent en résulter (Callen et al., 2004; Shearwin et al., 2005).

Chez les eucaryotes, la terminaison de la transcription est couplée à la formation de l'extrémité 3' du transcrit et débute avec la reconnaissance de séquences spécifiques du transcrit. Le mécanisme de formation de ces extrémités est différent pour les gènes qui codent pour des ARNm de ceux qui codent pour des snoARN. Dans le cas des ARNm, le signal de polyadénylation est reconnu par le complexe formé du facteur de clivage et de polyadénylation (CPF) et des facteurs de clivage IA and IB (CFIA et CFIB). Ce complexe orchestre le clivage endonucléolytique et la polyadénylation des extrémités 3'. Après le clivage du transcrit, la polymérase se dissocie de l'ADN et le petit fragment d'ARN résultant du clivage est dégradé. Deux modèles ont été proposés pour décrire le couplage entre la polyadénylation et la terminaison (Buratowski, 2005). Dans le modèle "allostérique", le complexe de transcription subit un changement de conformation une fois qu'il dépasse le signal de polyadénylation et, conséquemment, les facteurs d'élongation sont remplacés par des facteurs de terminaison. Dans le deuxième modèle, appelé "torpedo", suite au clivage par le complexe de polyadénylation, une nouvelle extrémité 5' est générée et celle-ci constitue le substrat pour la nucléase Rat1/Xrn2. Cette nucléase digère le fragment d'ARN que la polymérase transcrit au-delà du site polyadénylation et déclenche la dissociation de la polymérase de l'ADN (revu par (Bentley, 2005) et (Lykke-Andersen and Jensen, 2007)). Un facteur important pour le couplage entre la polyadénylation et la terminaison est PCF11p qui interagit avec l'ARNm et le domaine C-terminal de la sous-unité Rpb1 de l'ARN Pol II. En utilisant un système d'élongation *in vitro* avec des facteurs purifiés, il a été démontré qu'il provoque la désintégration du complexe d'élongation (Zhang et al., 2005b).

La terminaison de la transcription des petits ARN nucléolaires (snoARN) se fait par un processus différent chez la levure et chez les vertébrés. Chez les vertébrés, la plupart de ces petits ARN sont codés dans les introns des gènes des ARNm et ils sont générés par clivage endonucléolytique à partir des ARN pré-messagers (Kiss et al., 2006). Chez la levure, les snoARN sont codés par des gènes indépendants et ils ont leurs propres promoteurs et séquences de terminaison (Kim et al., 2006; Kiss, 2006). Dans ce dernier cas, le complexe responsable de la terminaison est formé par

les protéines Nrd1 et Nab3 qui fonctionnent de façon analogue au complexe de polyadénylation et clivage, c'est à dire en liant à la fois le transcrit à des sites spécifiques appelés sites de reconnaissance pour Nrd1, ainsi que le domaine CTD de la polymérase (Steinmetz and Brow, 1998). Les extrémités générées par le clivage sont par la suite digérées par le complexe exosome des ARN, processus stimulé par le complexe TRAM qui fournit un site d'entrée sur l'ARN (LaCava et al., 2005).

Une corrélation étroite entre la phosphorylation du CTD de la polymérase et la polyadénylation est appuyée par plusieurs indices. Une délétion ou une inhibition de CDK9 ou de son homologue Ctk1 génère des défauts de polyadénylation (Ahn et al., 2004; Ni et al., 2004). De plus, la phosphorylation du CTD sur la sérine 2 augmente l'affinité du facteur Pcf11 pour le CTD de la polymérase (Licatalosi et al., 2002).

Chez les bactéries et les phages, la terminaison est dépendante du facteur Rho. Il s'agit d'une protéine avec des fonctions ATPase et hélicase. La terminaison dépend également des éléments spécifiques riches en cytosines situés autant en amont, dans les régions codantes, qu'en aval des gènes (Ciampi, 2006). Lorsque Rho lie l'ARNm sa fonction ATPase est activée. Ceci fournit l'énergie nécessaire pour qu'il glisse le long de l'ARNm jusqu'à ce qu'il rencontre la polymérase. Son activité hélicase est responsable de la dissociation du transcrit de la polymérase. Chez les procaryotes, la terminaison se produit aussi à des sites qui génèrent des structures secondaires stables en épingle à cheveux dans le transcrit et qui sont suivis par une queue de polyU. Ces régions favorisent la terminaison puisque les hybrides ADN-ARN formés sont instables (von Hippel and Yager, 1992).

1.4 La régulation de la transcription

1.4.1 La chromatine

La transcription est un processus complexe qui est hautement régulé. Un premier niveau de régulation est constitué par la compaction de l'ADN dans la chromatine :

la restriction physique de l'accès à l'ADN de l'ARN Pol II et des facteurs associés maintient les gènes à l'état inactif. La chromatine est formée de structures appelées nucléosomes qui sont constitués d'ADN et d'histones. Un nucléosome est constitué d'une séquence de 147 pb qui s'enroule deux fois autour d'un octamère d'histones. Les domaines N-terminaux des histones dépassent de la surface des nucléosomes et jouent un rôle important dans la compaction des nucléosomes en structures supérieures. Ces domaines subissent des modifications post-traductionnelles telles que l'acétylation, la méthylation, la phosphorylation, l'ubiquitination, la poly (ADP-ribosyl)ation, etc. Des combinaisons spécifiques de modifications post-traductionnelles des histones signalent soit l'activation soit l'inactivation des gènes (Sims, III and Reinberg, 2008). Lors de l'activation, la chromatine des régions régulatrices ainsi que celle située au long des gènes doit être remodelée afin de permettre l'accès de la machinerie transcriptionnelle à l'ADN. Les principales enzymes impliquées dans le remodelage de la chromatine font partie de deux classes : les histone acétyltransférases et les enzymes de remodelage de la chromatine dépendantes de l'ATP (ex SWI/SNF, INO 80) (Clapier and Cairns, 2009). Lorsqu'elles sont recrutées sur la chromatine, ces enzymes remodelent et repositionnent les nucléosomes de façon à ce que les activateurs et la machinerie transcriptionnelle soient recrutés. À l'opposé, l'inactivation des gènes est associée à des modifications spécifiques telles que la méthylation de l'ADN, ainsi qu'à un environnement répressif de la chromatine qui est généré par des histone desacétylases et des enzymes de remodelage de la chromatine dépendantes de l'ATP.

Les enzymes de remodelage de la chromatine dépendantes de l'ATP sont classées en trois catégories : la famille ISWI constituée des complexes qui aident l'organisation et l'assemblage des nucléosomes, la famille SWI/SNF dont les membres ont pour fonction de déplacer ou d'éjecter les nucléosomes et finalement la famille SWR1 qui se spécialise dans la reconstruction des nucléosomes par une insertion de la variante d'histone H2A.Z (revu par (Cairns, 2009)).

L'organisation des nucléosomes dans la région des promoteurs a un impact majeur sur la façon dont un gène est régulé (Fig. 13). Les gènes de maintenance, qui sont exprimés de façon constitutive, ont dans leurs promoteurs une région déplétée de nucléosomes (NDR pour "*nucleosome depleted region*") ou "*ouverte*". Cette région est caractérisée par des séquences riches en poly (A:T) qui défavorisent la formation des nucléosomes et est liée en permanence par des activateurs. Les régions NDR sont flanquées par des nucléosomes strictement positionnés par des séquences de dinucléotides AA-TT qui favorisent une courbure dans la structure de l'ADN et la formation des nucléosomes, appelés "-1" ou "+1". Ces promoteurs sont généralement dépourvus de boîte TATA.

Les gènes régulés ont des promoteurs "*couverts*" par des nucléosomes positionnés faiblement, qui couvrent les sites d'initiation de la transcription et la plupart de sites de liaison des activateurs (BS). Un site de liaison d'un activateur 'pionnier' reste tout de même accessible, et la reconnaissance de ce site par l'activateur déclenche la cascade de modifications qui mène à l'activation de la transcription. L'expression de ces gènes est dépendante des complexes de remodelage de la chromatine qui mènent à l'exposition des autres sites de liaison pour différents activateurs. Contrairement aux promoteurs des gènes de maintenance, les promoteurs des gènes régulés ont une boîte TATA située à des distances variables du TSS (revu par Cairns,2009).

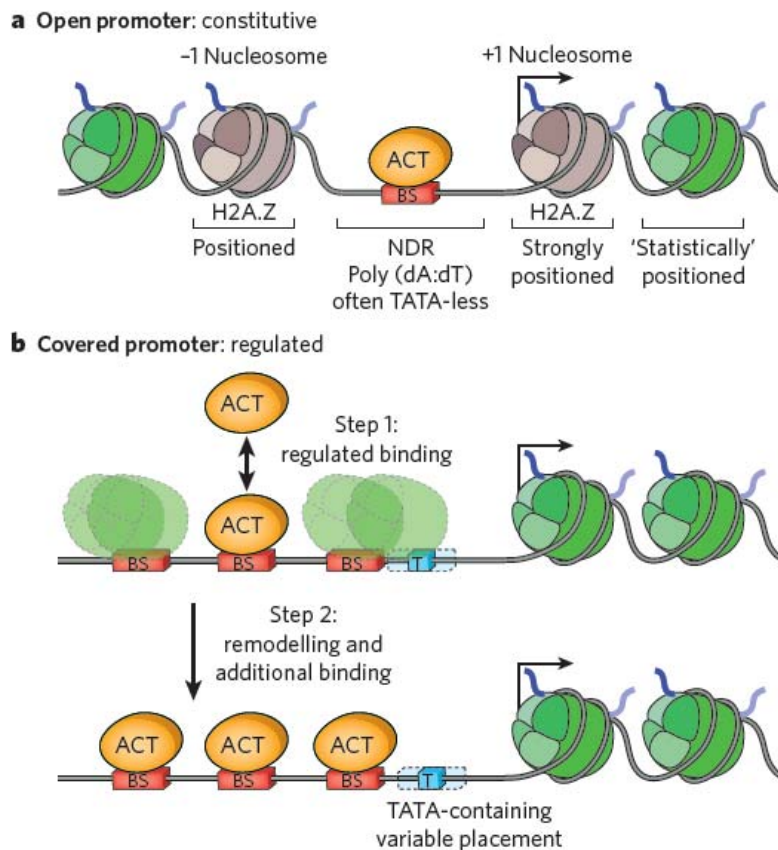


Figure 13. Modèle de la régulation par le positionnement des nucléosomes des deux classes de gènes, de maintenance ou régulés. Les gènes de maintenance(a) se caractérisent par des promoteurs “*ouverts*”, dépourvus de nucléosomes (NDR “nucleosomes depleted region”) dans la région proximale au site d’initiation de la transcription (TSS). Les gènes régulés (b) ont de promoteurs “*couverts*”, un nucléosome régulateur étant positionné dans la région proximale au TSS en état de répression transcriptionnelle. Un premier activateur (ACT) interagit avec des séquences de liaison (BS) spécifiques ce qui permet le recrutement des enzymes de remodelage de la chromatine. Suite à l’activité de remodelage de la chromatine, d’autres séquences sont exposées et peuvent être reconnues par des activateurs. Tiré de (Cairns, 2009).

1.4.2 Les activateurs et les répresseurs de la transcription

Le séquençage complet du génome humain nous fournit la base pour l'étape prochaine, connaître comment les 20 000 à 25 000 gènes codant pour des protéines sont régulés, quels sont les éléments régulateurs pour chacun d'entre eux et comment la coordination de leur expression est réalisée. Les patrons d'expression génique spécifiques à différents états physiologiques et pathologiques excèdent largement le nombre de facteurs qui lient l'ADN identifiés jusqu'à présent (~2000) (revu par (Maston et al., 2006)). Le contrôle combinatoire par la liaison simultanée de plusieurs facteurs régulateurs à des sites situés à différentes distances des promoteurs assure une régulation à la fois précise et diversifiée de l'expression des gènes (revu par (Remenyi et al., 2004)). Les facteurs qui lient l'ADN constituent en effet l'interface entre l'information génétique et la machinerie transcriptionnelle qui la décode en réponse à des signaux cellulaires transmis par ces facteurs.

Les activateurs sont des protéines modulaires qui ont des domaines typiques tels qu'un domaine de liaison à l'ADN par lequel des séquences spécifiques d'ADN sont reconnues, ainsi qu'un domaine d'activation qui est requis pour l'activation de la transcription. De plus, des domaines de coopérativité assurent l'interaction entre plusieurs activateurs qui peuvent collaborer pour activer la transcription. Les activateurs sont classés en fonction de leur domaine de liaison à l'ADN, chaque classe liant des séquences d'ADN spécifiques. Parmi les différentes classes d'activateurs, on peut citer ceux qui possèdent un domaine de liaison à l'ADN de type *forkhead*, *homeobox*, à doigt de zinc riche en cystéine, PIT-ONC-UNC (POU), bHLH (*basic helix-loop-helix*) et bZIP (*basic leucin zipper*). Certains activateurs n'appartiennent à aucune de ces classes alors que les membres de certaines classes reconnaissent des séquences d'ADN très divergeantes, puisque les acides aminés importants pour la reconnaissance de l'ADN sont variables d'un membre à l'autre (revu par (Kadonaga, 2004; Pabo and Sauer, 1992)).

Les séquences d'ADN qui sont reconnues par des activateurs sont courtes (6 à 12 pb) et la spécificité de liaison est dictée par quelques bases de ces séquences (4 à 6 pb). Puisque ces séquences sont dégénérées, la spécificité de liaison résulte de la composition spécifique des complexes formés par homo ou hétérodimérisation des activateurs. La séquence d'ADN liée peut jouer des rôles régulateurs très précis. Par exemple, l'affinité de liaison des morphogènes à l'ADN est dépendante de la séquence d'ADN. Ceci a un rôle déterminant afin de générer des patrons d'expression génique corrélés aux gradients des morphogènes pendant le développement (Papatsenko and Levine, 2005). Aussi, les conformations des activateurs changent en fonction des séquences précises liées, ce qui modifie leur capacité régulatrice (ex. les récepteurs nucléaires des hormones stéroïdes) (Geserick et al., 2005).

Les domaines d'activation recrutent la machinerie générale de transcription soit directement soit par l'entremise de coactivateurs. Par ces interactions, les activateurs peuvent stimuler différentes étapes de la transcription. Un des mécanismes d'activation implique le recrutement du PIC par l'interaction avec un ou plusieurs facteurs généraux de transcription (revu par (Ptashne and Gann, 1997)). Les activateurs peuvent aussi stimuler des étapes subséquentes à la formation du PIC comme la réinitiation et l'élongation (Lee and Young, 2000) ou en recrutant des facteurs impliqués dans le remodelage de la chromatine (Lemon and Tjian, 2000). Un exemple pour ce dernier type de régulation est le recrutement du complexe de remodelage de la chromatine Swi/Snf par l'activateur MyoD. Par exemple, l'activation du gène de la myogénine implique la collaboration entre l'activateur ubiquitaire PBX1 qui est lié de façon constitutive à ses séquences régulatrices et l'activateur spécifique au tissu musculaire MyoD, qui lui, attire le complexe Swi/SNF (de la Serna et al., 2005). Certains activateurs stimulent spécifiquement une seule étape de la transcription, alors que d'autres peuvent en stimuler plusieurs. Ainsi, Sp1 et CTF stimulent l'initiation, l'activateur Tat de HIV1 stimule l'élongation, alors que VP16, p53 et E2F1 stimulent l'initiation et l'élongation (Blau et al., 1996).

Une propriété remarquable des activateurs est leur capacité d'activer la transcription de façon synergique, c'est-à-dire que l'amplitude de l'activation d'un groupe d'activateurs dépasse celle qui est obtenue par l'addition des effets individuels. La synergie est observée entre plusieurs molécules d'un même activateur ou des activateurs différents. Le mécanisme exact n'a pas encore été élucidé, cependant il a été proposé qu'il résulte des interactions qui se produisent après le recrutement des activateurs sur l'ADN (Maston et al., 2006). Des complexes stables d'activateurs appelés "*enhancesomes*" stimulent la transcription de façon synergique au niveau de l'activation et non pas de l'interaction avec l'ADN. Par exemple, l'activation du gène IFN β implique le recrutement du coactivateur CBP/p300 par les domaines d'activation de l'*enhancesome* formé par NF-kappa B, IRF1, ATF2/c-Jun et HMG I. Chacun des domaines d'activation de ces facteurs est essentiel pour que l'effet synergique se produise, puisque la délétion de n'importe lequel d'entre eux diminue le recrutement de CBP/p300 (Merika and Thanos, 2001).

La régulation transcriptionnelle représente l'équilibre entre l'action des activateurs et des répresseurs et, en conditions spécifiques, une des deux voies est favorisée (revu par (Lee and Young, 2000)). Des indices sur cet équilibre ont été amenés par des études génétiques qui ont montré que des défauts des activateurs sont compensés par des défauts de répresseurs et vice versa. Par exemple, les phénotypes de sensibilité au froid des mutants des sous-unités du Médiateur sont contrecarrés par des mutations des régulateurs négatifs NC2 ou Not (Lee et al., 1998). De plus, dans certaines conditions, des activateurs peuvent agir comme des répresseurs en recrutant des cofacteurs différents. Par exemple, l'activateur VDIR (*VDR Interacting Repressor*) de la famille bHLH, lie directement les éléments de réponse au récepteur de la vitamine D, recrute le coactivateur p300 et active l'expression du gène α hydrolase en réponse aux signaux activateurs transmis par PKA (*Protein Kinase A*). Lorsque l'hétérodimère VDR/RXR (*Vitamin D Receptor/Retinoid X Receptor*) interagit avec VDIR, le coactivateur p300 est dissocié et le corepresseur

HDAC est recruté par VDR/RXR. Ceci mène à la répression transcriptionnelle du gène (Murayama et al., 2004).

Deux classes majeures de répresseurs transcriptionnels ont été identifiées : généraux et spécifiques au gène. Les répresseurs généraux comme MOT1 et NC2 interagissent avec TBP et provoquent sa dissociation de l'ADN ou empêchent le recrutement de l'ARN Pol II sous forme d'holoenzyme par TBP lié à l'ADN (Auble, 2009; Darst et al., 2003). Les répresseurs spécifiques aux gènes interagissent avec les activateurs ou entrent en compétition avec ceux-ci pour la liaison à l'ADN, comme par exemple le répresseur Acr1 et l'activateur ATF/CREB qui se lient aux mêmes séquences d'ADN (Garcia-Gimeno and Struhl, 2000; Vincent and Struhl, 1992). D'autres répresseurs de cette catégorie comme Ssn6-Tup1 sont recrutés par des activateurs et ils recrutent à leur tour des déacétylases de histones qui établissent un environnement chromatinien répressif (Malave and Dent, 2006). Les protéines du groupe Polycomb, qui sont responsables de la répression transcriptionnelle de centaines de gènes chez les eucaryotes supérieurs sont recrutées par des facteurs qui lient l'ADN (ex. PHO, YY1, OCT4) mais aussi par des ARN non-codants comme un motif d'ARN de *XIST* (*X-Inactivative Specific Transcript*). Il a été proposé que l'activité des répresseurs transcriptionnels est due à l'inhibition de l'étape d'élongation causée par l'activité d'ubiquitine ligase pour les histones H2A des protéines PRC1 (*Polycomb Repressive Complex 1*) et par la compaction de la chromatine (revu par (Simon and Kingston, 2009)).

Les déacétylases des histones sont largement impliquées dans la répression transcriptionnelle. Des enzymes avec ce type d'activité, telles que mSin3 et NuRD, sont recrutées par des facteurs qui lient l'ADN ou par des coactivateurs comme SMRT, Groucho, N-Cor (revu par (Hildmann et al., 2007)). mSin3 interagit aussi avec des protéines qui reconnaissent des histones méthylées, comme MeCP2 (Kokura et al., 2001). Chez la levure, le complexe formé par les déacétylases des histones *Sir2*, *Sir3* et *Sir4* est impliqué dans la répression transcriptionnelle des télomères (Grozinger and Schreiber, 2002). SirT1 est un membre de la famille des

sirtuines Sir2 qui présente une activité déacétylase dépendante de NAD^+ (Imai et al., 2000) et qui régule l'équilibre entre l'euchromatique et l'hétérochromatique en fonction de l'état métabolique de la cellule. Le rôle de SirT1 dans la régulation de la transcription est accompli par plusieurs mécanismes (revu par (Zhang and Kraus, 2009)). La répression transcriptionnelle implique non seulement son activité de déacétylase des histones mais aussi le recrutement des enzymes avec une activité méthylase des histones et l'ADN. De plus, SirT1 désacétyle des facteurs de transcription (ex. p53, $\text{ER}\alpha$, FOXO, etc) et des cofacteurs (ex. PGC1), ce qui peut avoir un impact positif ou négatif sur la transcription (Feige et al., 2008).

1.4.3 Les modifications post-traductionnelles des facteurs impliqués dans la transcription

De nombreuses modifications post-traductionnelles (MPT) des protéines, incluant celles impliquées dans la transcription, ont été décrites comme étant importantes afin de définir les fonctions précises de ces protéines dans différents contextes biologiques. Parmi les MPT les plus étudiées, on peut citer : l'acétylation, la méthylation, la phosphorylation, l'ubiquitination, la sumoylation, la neddylation et la poly(ADP) ribosylation. Les modifications des histones montrent bien l'importance fonctionnelle des MPT. Il a ainsi été proposé qu'elles contribuent à la diversification fonctionnelle pendant l'évolution et, dans certains cas, à des changements héréditaires qui n'impliquent pas des changements de séquence de l'ADN (Sims, III and Reinberg, 2008; Turner, 2007). Les MPT génèrent des surfaces chimiques nouvelles qui sont reconnues par des domaines protéiques distincts, ce qui déclenche des cascades effectrices. Basé sur les données expérimentales accumulées, le rôle attribué aux MPT des protéines est de spécifier la fonction de ces protéines en réponse à des signaux de l'environnement cellulaire. Par contre, les MPT ne peuvent pas être vues comme génératrices d'un "code des protéines" (incluant les histones) de façon analogue au "code génétique" par lequel un triplet de nucléotides spécifie un seul aminoacide indépendamment du contexte

cellulaire (Sims, III and Reinberg, 2008). Il existe de nombreux exemples de MPT qui conduisent à des effets différents en fonction du contexte cellulaire.

Un exemple lié à la transcription est l'ARN Pol II elle-même et ses MPT. Le patron de phosphorylation du domaine C-terminal de sa plus grande sous-unité est reconnu par des facteurs spécifiques aux différentes étapes de la transcription. Par exemple, la phosphorylation de la Ser⁵ lors de l'initiation de la transcription est reconnue par l'enzyme de coiffage des ARNm. Puisque le CTD est situé près du canal de sortie des ARNm, ceci assure leur protection par l'ajout d'une coiffe dès leur sortie à la surface de la polymérase (Fabrega et al., 2003). De plus, la Ser⁵ phosphorylée recrute ou facilite le recrutement des facteurs impliqués dans l'étape d'élongation, comme Paf1 (Qiu et al., 2006) et CDK9 (Viladevall et al., 2009). Cependant, cette modification du CTD ne signifie pas que la polymérase entrera nécessairement en élongation, mais plutôt, qu'elle pourra aussi entrer en pause ou terminer la transcription dépendamment du contexte. Par exemple, lors de la transcription des séquences initiales des snoRNAs ou des petits ARN CUT (*Cryptic Unstable Transcripts*), la phosphorylation du CTD sur la Ser⁵ contribue au recrutement du complexe Nrd1-Nab3-Sen1 et à la terminaison de la transcription (Vasiljeva et al., 2008). Puisque ce complexe se lie à l'ARN de façon dépendante de séquence, la terminaison n'est observée que pour les gènes qui ont les séquences de reconnaissance appropriées dans leur régions 5' (Buratowski, 2009; Kim et al., 2006). Cet exemple montre que la phosphorylation de la Ser⁵ du CTD a des fonctions spécifiques dépendamment des séquences qui sont transcrites.

Des facteurs comme TFIID et le Médiateur fonctionnent comme des canaux à travers lesquels l'information est transmise avec la contribution de MPT (revu par (Bjorklund and Gustafsson, 2005a; Cler et al., 2009)). Ainsi, plusieurs sous-unités de TFIID possèdent des domaines qui reconnaissent des MPT spécifiques: TAF1 a un double bromodomaine par lequel il reconnaît des histones acétylées, TAF3 a un homéodomaine qui reconnaît les histones méthylées (H3K4me3) et TAF5 possède un domaine WD40 qui peut lier des histones méthylées (de préférence H3K4me3)

(Cler et al., 2009). D'un autre coté, TFIID possède de nombreuses activités enzymatiques par lesquelles il acétyle, phosphoryle ou transfère des molécules d'ubiquitine sur des histones et des facteurs de transcription comme TFIIE et TFIIA (Imhof et al., 1997; Maile et al., 2004; Mizzen et al., 1996; Pham and Sauer, 2000; Solow et al., 2001). De plus, l'activité de TFIID est régulée par sumoylation de TAF5, une modification qui empêche l'interaction de TFIID avec le promoteur (Boyer-Guittaut et al., 2005).

Dans le Chapitre 4, je présenterai des analyses par lesquelles la E3 ubiquitine ligase UBR5 (*Ubiquitin protein ligase E3 component n-Recognin 5*) est identifiée comme facteur associé à des complexes transcriptionnels et qu'elle est impliquée dans la régulation par ubiquitination de CDK9. Afin de pouvoir mettre en perspective les résultats de cette étude, dans la section ci-dessous je présenterai brièvement ce type de MPT ainsi que les informations disponibles dans la littérature sur UBR5.

1.4.3.1 L'ubiquitination et la E3 ubiquitine ligase UBR5

Chez les espèces eucaryotes, la modification des protéines par ubiquitination joue un rôle important dans une multitude de processus allant de la division cellulaire, la présentation d'antigènes jusqu'à la transcription. La fonction biochimique la mieux caractérisée de cette modification est son rôle de substrat pour la reconnaissance par le protéasome, menant ainsi à la dégradation ciblée des protéines. Cependant, d'autres fonctions pour l'ubiquitination ont été identifiées dans des processus comme la signalisation cellulaire, le transport des protéines, l'endocytose, et la localisation subcellulaire (Hershko, 2005; Hershko and Ciechanover, 1998; Rotin and Kumar, 2009). Une séquence de trois réactions mène à l'ubiquitination des protéines. Dans une première étape, une enzyme activatrice E1 active une molécule d'ubiquitine de façon dépendante de l'ATP en formant un lien thioester entre le groupement carboxyle C-terminal de l'ubiquitine et un résidu cystéine de l'enzyme. L'ubiquitine activée est transférée dans une deuxième étape sur un résidu cystéine de l'enzyme de conjugaison E2 et, finalement, une enzyme E3 ubiquitine ligase

facilite la formation d'un lien isopeptide entre le groupement carboxyle C-terminal de la molécule d'ubiquitine activée et un résidu lysine du substrat ou la molécule d'ubiquitine précédente (Goldberg, 2003). Une chaîne de polyubiquitines est formée lorsqu'une des sept lysines d'une molécule d'ubiquitine est liée à la glycine C-terminale d'une autre molécule d'ubiquitine. Les chaînes de polyubiquitines les mieux caractérisées jusqu'à présent utilisent les lysines K48 ou K63. Les chaînes de polyubiquitines qui utilisent la lysine K48 ciblent le plus souvent la protéine pour la dégradation, alors que celles qui utilisent la lysine K63 régulent les fonctions cellulaires des protéines, leur localisation cellulaire et les interactions protéine-protéine (Bernassola et al., 2008). L'ubiquitination est une modification réversible et plusieurs déubiquitinases (DUB) ont été caractérisées comme étant des enzymes capables de cliver les molécules d'ubiquitine (Reyes-Turcu et al., 2009).

Les ubiquitine ligases sont des enzymes hautement spécifiques qui transfèrent des molécules d'ubiquitine sur d'autres protéines en reconnaissant des structures caractéristiques, appelées degrons. Ceux-ci sont constitués de séquences riches en phosphosérine ou en hydroxyproline, ainsi que des acides aminés N-terminaux basiques (Tasaki et al., 2005). Les E3 ubiquitine ligases sont classées en fonction de l'homologie de leur domaine de liaison à l'enzyme E2 en trois sous-familles: les E3 ligases qui ont un domaine d'homologie à la E6AP C-terminal (HECT), celles qui ont un domaine RING et finalement celles qui possèdent une boîte U. Les propriétés de ces E3 ligases sont différentes. Ainsi, celles qui possèdent un domaine RING servent d'échafaudage pour le transfert de l'ubiquitine de l'enzyme E2 au substrat. En contraste avec cette famille, les E3 ubiquitine ligases à domaine HECT possèdent une activité catalytique intrinsèque. Le transfert de l'ubiquitine par cette famille d'enzymes comporte trois étapes : la liaison à la E2 ligase, le transfert de l'ubiquitine sur la E3 par la formation d'un intermédiaire ubiquitin-thioester avec le résidu cystéine catalytique situé dans la partie C-terminale du domaine HECT, et le transfert de l'ubiquitine sur le substrat (Bernassola et al., 2008). Le transfert de l'ubiquitine de E2 à E3 est initié par l'attaque nucléophile du lien ubiquitine-thioester de E2 par le résidu cystéine catalytique de la E3 ligase. Chez l'humain, il

existe une vingtaine de protéines à domaine HECT dont certaines interagissent avec la E2 UbcH5, alors que d'autres interagissent avec UbcH7 (Schwarz et al., 1998)

UBR5 est une E3 ubiquitin ligase à domaine HECT qui est aussi connue sous le nom de EDD1 (*E3 identified by Differential Display*). Ce gène code pour l'orthologue humain du gène suppresseur de tumeur "*hyperplastic discs*" (HYD) chez *D. melanogaster*. Il a été classé initialement parmi les E3 ubiquitin ligases suite à l'identification du domaine catalytique HECT situé dans sa région C-terminale (Callaghan et al., 1998). Ce même groupe a montré que la protéine traduite *in vitro* lie de façon spécifique l'ubiquitine, alors que la version ayant une mutation ponctuelle de la cystéine catalytique dans le domaine HECT perd cette capacité.

Plusieurs domaines fonctionnels de UBR5 ont été identifiés ultérieurement (Fig. 14), suggèrent que UBR5 est une protéine qui lie l'ubiquitine et qu'elle est impliquée dans des processus d'ubiquitination. Ainsi, UBR5 présente dans sa région N-terminale un domaine UBA (*Ubiquitin-Associated*) et il a été montré par mutagenèse dirigée que ce domaine lie l'ubiquitine, deux de ses résidus étant requis : Val¹⁹⁶ et Leu²²⁴ (Kozlov et al., 2007). Le domaine PABP-C est impliqué dans la spécificité de liaison aux substrats qui sont régulés par ubiquitination. La structure cristallographique de ce domaine et des essais d'interaction *in vitro* ont montré que le domaine PABP-C de UBR5 lie la protéine Paip2 (*PABP-interacting protein 2*) (Deo et al., 2001). Paip2 interagit et régule l'activité de la protéine PABP qui possède tout comme UBR5 un domaine PABP-C. L'interaction entre Paip2 et PABP se fait aussi par le domaine PABP-C, de façon analogue à l'interaction entre Paip2 et UBR5. Cette analogie d'interaction a une signification fonctionnelle importante. Il a été démontré que UBR5 lie et ubiquitine Paip2 en fonction des niveaux d'expression de son autre partenaire d'interaction, la protéine PABP : lorsque les niveaux de PABP sont diminués par atténuation de l'expression, Paip2 lie UBR5, ce qui mène à son ubiquitination et sa dégradation par le protéasome (Yoshida et al., 2006). En contrôlant les niveaux de la protéine Paip2, UBR5 est

impliqué dans la régulation de l'initiation de la traduction des ARNm, processus dans lequel Paip2 joue un rôle central (Khaleghpour et al., 2001; Svitkin et al., 2001; Yoshida et al., 2006).

UBR5 possède trois domaines putatifs de localisation nucléaire (NLS ou “*Nuclear Localization Sequence*”). Le premier est situé dans la région centrale et n'est pas fonctionnel pour l'import nucléaire. Les deux autres sont situés dans la région N-terminale et leur présence simultanée est requise pour l'interaction avec l'importin α , tel qu'observé lors d'essais d'interaction *in vitro* (Henderson et al., 2002). En accord avec cette observation, l'analyse de la localisation cellulaire de UBR5 a indiqué une localisation principalement nucléaire.

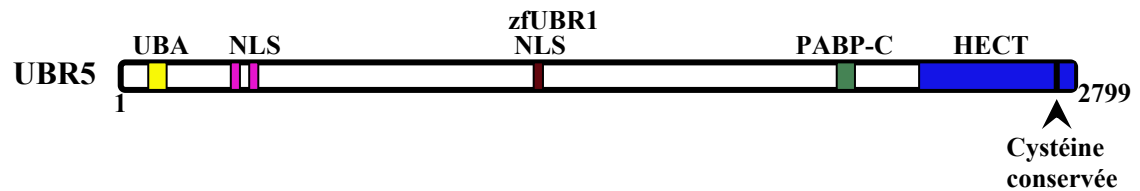


Figure 14. Représentation schématique des domaines fonctionnels de UBR5. Les domaines UBA (*ubiquitin-associated*), trois NLS putatifs (*nuclear localization sequence*), zfUBR1 (domaine d'homologie avec le motif à doigts de zinc de la N-recogin) PABP-C (poly(A)-binding protein) et HECT sont représentés. La cystéine catalytique est indiquée. Adapté de (Henderson et al., 2002)

Des homologues de UBR5 ont été identifiés chez les espèces eucaryotes supérieurs de *C. elegans* jusqu'à l'humain, avec une identité de séquence variable. Par exemple, entre les homologues chez l'humain et *D. melanogaster*, il y a 46% d'identité de séquence. L'homologue chez *D. melanogaster*, appelé HYD a été impliquée dans la terminaison de la prolifération. Un rôle de suppresseur de tumeur a été proposé pour cette protéine car des mutations spécifiques ont entraîné des défauts de développement et des phénotypes tumoraux (Mansfield et al., 1994). Un

rôle dans la progression tumorale a été avancé par plusieurs études chez l'humain qui ont montrées que UBR5 est fréquemment surexprimé dans différents types de cancers, tels que les cancers ovarien, du sein, les carcinomes hépatocellulaires et d'autres (Chin et al., 2007a; Clancy et al., 2003; O'Brien et al., 2008).

UBR5 est aussi impliqué dans la réponse cellulaire aux dommages de l'ADN et un premier indice a été son interaction avec la protéine TopBP1, qui a une fonction connue dans la réponse aux bris double brin de l'ADN (*Topoisomerase IIbeta-Binding Protein*) (Honda et al., 2002; Makiniemi et al., 2001; Yamane et al., 2002). En conditions normales, cette interaction mène à l'ubiquitination de TopBP1 et sa dégradation protéosomale. Suite à l'induction des dommages à l'ADN, TopBP1 est phosphorylée, modification qui inhibe l'ubiquitination par UBR5. UBR5 interagit également avec d'autres protéines impliquées dans la réponse aux dommages de l'ADN, telle que CIB/KIP (*Calcium and Integrin-Binding Protein*), qui à son tour interagit avec des protéines avec des fonctions connues dans ce processus, comme la protéine kinase dépendante d'ADN (Wu and Lieber, 1997), PLK-1 et PLK3 (*Polo-Like Kinase-1 et 3*) (Winkles and Alberts, 2005). Une interaction directe a été rapportée entre UBR5 et CHK2 (*Checkpoint Kinase 2*), une kinase qui phosphoryle de nombreux substrats en réponse aux dommages de l'ADN et induit l'arrêt du cycle cellulaire (Henderson et al., 2006). Cette interaction stimule l'activité kinase de CHK2, cependant on ne sait pas si une ubiquitination de CHK2 est impliquée.

La délétion de HYD est létale chez *Drosophila* dans les stades de développement précoces. Chez la souris, l'ablation génique de UBR5 est létale, causant des défauts dans le développement extraembryonnaire, alors que les hétérozygotes générés par ciblage du gène UBR5 ne démontrent pas de phénotype (Saunders et al., 2004).

1.4.4 D'autres mécanismes de régulation de la transcription

En plus des mécanismes bien connus de régulation de la transcription par des activateurs, des répresseurs et le remodelage ou la modification de la chromatine, la

transcription des ARN non-codants (ARNnc) représente un mécanisme de régulation qui est largement répandu (Berretta and Morillon, 2009; Faulkner and Carninci, 2009). L'interférence transcriptionnelle chez la levure et la transcription bidirectionnelle à partir de séquences qui chevauchent les promoteurs des ARNm autant chez la levure que chez les mammifères ont été décrits comme mécanismes de régulation de la transcription par des petits ARNnc (Carninci, 2009). Environ 50% des gènes codants pour des ARNm ont dans leur région 3'UTR des séquences qui codent pour des petits ARNnc et le rôle proposé de ces ARN serait de maintenir un état euchromatique et de permettre des niveaux élevés de transcription (Kapranov et al., 2007). D'autres petits ARN, appelés CUT (*Cryptic Unstable Transcripts*), puisqu'ils sont dégradés rapidement après leur synthèse, jouent un rôle dans la répression de la transcription en recrutant localement des histone-desacétylases (Camblong et al., 2007).

Des nouvelles techniques moléculaires et d'imagerie cellulaire ont mis en évidence un autre mécanisme de régulation de la transcription par des structures tridimensionnelles dynamiques impliquant la formation de boucles intra et inter-chromosomiques (Phillips and Corces, 2009). Ainsi, la chromatine est condensée et décondensée, et forme des territoires tridimensionnelles qui sont transcrits de façon coordonnée sur de grandes distances. Ces territoires fonctionnels seraient établis afin d'utiliser des ressources communes dans les compartiments nucléaires. La transcription régulée pendant le développement, l'expression monoallélique, l'inactivation transcriptionnelle sur des longues distances, comme par exemple l'inactivation du chromosome X, sont des exemples de régulation coordonnée de la transcription sur des territoires tridimensionnels plutôt que de façon linéaire, à de loci des gènes uniques. Quelques facteurs impliqués dans la formation de ces structures supérieures d'organisation de la chromatine sont les facteurs CTCF (*CCCTC-binding Factor*), HP1 (*Heterochomatin Protéin 1*) et les complexes polycomb ainsi que des éléments en *cis* comme les MAR (*Matrix Attachment Regions*) (Galande et al., 2007; Kheradmand et al., 2009; Phillips and Corces, 2009; Wegel and Shaw, 2005).

1.5 Objectifs de recherche

Mon premier objectif a été d'étudier de façon systématique l'implication des facteurs généraux de transcription humains dans les étapes de transcription *in vivo*, de façon à ce que les stades précoces (formation du PIC, initiation, dégagement du promoteur) soient distingués de l'étape d'élongation. Nous avons également étudié la localisation génomique de plusieurs sous-unités de l'ARN Pol II, de type sauvage ou mutant. Aucune autre étude systématique chez l'humain n'a été effectuée et la fonction spécifique des GTFs dans ces étapes était basée sur des études *in vitro* ou sur des études fragmentaires *in vivo*, dans la plupart des cas chez d'autres espèces. Afin d'effectuer cette étude, nous avons utilisé une collection de GTFs portant une étiquette TAP et une analyse de leur localisation génomique par immunoprécipitation de la chromatine (ChIP). Dans le but de faire la distinction entre les étapes précoces et tardives de la transcription, nous avons étudié des gènes qui présentent une divergence minimale de leurs sites d'initiation de la transcription (TIS). Les régions génomiques analysées ont été choisies de façon à ce qu'elles incluent le TIS ou qu'elles soient situées en aval, dans les régions transcrites de ces gènes. Cette étude nous a permis d'obtenir pour la première fois une image globale de la fonction de la machinerie transcriptionnelle de l'ARN Pol II humaine *in vivo* et a révélé des nouvelles fonctions *in vivo* pour le facteur TFIIF, ainsi que pour la sous-unité Rpb7 de l'ARN Pol II.

Mon deuxième objectif a été d'identifier de nouveaux partenaires d'interaction et des nouvelles fonctions pour le facteur TFIIS dans le contexte de ces nouvelles interactions. Suite à l'identification de l'ubiquitine-ligase UBR5 comme nouveau partenaire d'interaction de TFIIS, notre but a été de trouver la protéine substrat de cette ubiquitine ligase et d'identifier le rôle fonctionnel de cette interaction. Plusieurs essais biochimiques (ubiquitination, luciférase, d'immunoprécipitation de la chromatine et une atténuation de l'expression par des siRNAs) nous ont permis d'identifier CDK9, la sous-unité kinase de P-TEFb, comme substrat et d'établir que

sa polyubiquitination n'as pas pour rôle le ciblage de celle-ci pour la dégradation. Nous avons établi que TFIIIS augmente l'association de CDK9 le long du gène γ Fibrinogen de façon dépendante de UBR5.

CHAPITRE 2

LOCALISATION GÉNOMIQUE DE LA MACHINERIE TRANSCRIPTIONNELLE DE L'ARN POL II : IDENTIFICATION DU ROLE DE TFIIF ET RPB7 DANS LES ÉTAPES INITIALES ET TARDIVES DE LA TRANSCRIPTION

2.1 Préambule

Ma contribution à cet article consiste en: l'analyse par immunoprécipitation de la chromatine (ChIP) de la localisation génomique des sous-unités des GTFs et de la polymérase (Figures 2, 3 et 4 ainsi que la figure supplémentaire 3); la sélection à partir des bases des données DBTSS et MGC des gènes à analyser par ChIP, choisis puisqu'ils présentent une divergence minimale des TIS ("*Transcriptional Initiation Site*") alternatifs (en collaboration avec Pierre Côte). J'ai écrit l'article en proportion de 75% et j'ai fait les Figures 2, 3 et 4 et les Figures supplémentaires 2 et 3.

Le clonage, l'établissement des lignées stables exprimant les différentes sous-unités de l'ARN Pol II et des GTFs portant une étiquette TAP ainsi que la purification par la méthode TAP ont été réalisés par Diane Forget, Annie Bouchard et Célia Jérónimo (Figure 1A et 1B et la Figure supplémentaire 1). La lignée stable Rpb11-TAP a été générée dans le laboratoire du Dr Jack Greenblatt. Celia Jeronimo a effectué également les tests de transcription *in vitro* en utilisant des GTFs portant une étiquette TAP (Figure 1D).

Dominique Bergeron et Guy Poirier ont contribué à l'analyse par spectrométrie de masse des complexes purifiés par la méthode TAP présentés dans la Figure 1A et 1B et dans la Figure supplémentaire 1.

2.2 Article

Cojocaru M, Jeronimo C, Forget D, Bouchard A, Bergeron D, Côte P, Poirier GG, Greenblatt J, Coulombe B. Genomic location of the human RNA polymerase II general machinery: evidence for a role of TFIIF and Rpb7 at both early and late stages of transcription. **Biochem J.** 2008 Jan 1;409(1):139-47.

**Genomic location of the human RNA polymerase II general machinery:
Evidence for a role of TFIIF and Rpb7 at both early and late stages of
transcription**

Marilena Cojocaru ^{*}, Célia Jeronimo ^{*}, Diane Forget ^{*}, Annie Bouchard ^{*},
Dominique Bergeron ^{*}, Pierre Côte ^{*}, Guy G. Poirier [†], Jack Greenblatt [‡] and
Benoit Coulombe ^{*}

^{*} Institut de recherches cliniques de Montréal, Montréal, Québec, Canada;

[†] Centre Hospitalier Universitaire de Québec, Université Laval, Québec, Québec,
Canada;

[‡] Banting and Best Department of Medical Research, University of Toronto,
Toronto, Ontario, Canada.

Corresponding author:

Benoit Coulombe

Laboratory of Gene Transcription and Proteomics

Institut de recherches cliniques de Montréal

110 avenue des Pins Ouest

Montréal, Québec, Canada, H2W 1R7

Tel.: (514) 987-5662, Fax: (514) 987-5663,

Short Title:

Genomic location of the human general transcription machinery

ABSTRACT

The functions ascribed to the mammalian general transcription factors (GTFs) during the various stages of the RNA polymerase II (RNAPII) transcription reaction are based largely on *in vitro* studies. To gain insight as to the functions of the GTFs in living cells, we have analyzed the genomic location of several human GTF and RNAPII subunits carrying a Tandem-Affinity Purification (TAP) tag. ChIP experiments using anti-tag beads (TAP-ChIP) allowed the systematic localization of the tagged factors. Enrichment of regions located close to the transcriptional initiation site (TIS) *versus* further downstream transcribed regions (TRs) of 9 human genes, selected for the minimal divergence of their alternative TIS, was analyzed by Q-PCR. We show that, in contrast to reports using the yeast system, human TFIIF associates both with regions proximal to the TIS and with further downstream TRs, indicating an *in vivo* function in elongation for this GTF. Unexpectedly, we found that the Rpb7 subunit of RNAPII, known to be required only for the initiation phase of transcription, remains associated with the polymerase during early elongation. Moreover, ChIP experiments conducted under stress conditions suggest that Rpb7 is involved in the stabilisation of transcribing polymerase molecules, from initiation to late elongation stages. Together, our results provide for the first time a general picture of GTF function during the RNAPII transcription reaction in live mammalian cells and show that TFIIF and Rpb7 are involved in both early and late transcriptional stages.

Keywords: RNA polymerase II, general transcription factors, genomic location, chromatin immunoprecipitation, tandem-affinity purification.

Abbreviations used: ChIP, chromatin immunoprecipitation; EcR 293, human embryonic kidney cells stably transfected with the vector pVgRXR, expressing upon induction the heterodimeric ecdysone receptor; GTF, general transcription factor; LC-MS/MS, tandem mass spectrometry; MALDI-ToF, matrix-assisted laser desorption ionization-time of flight mass spectrometry; RNAPII, RNA polymerase II; TAP, tandem affinity purification; TIS, transcriptional initiation site; TR, transcribed region.

INTRODUCTION

Since their first purification two decades ago, the set of mammalian factors required for transcription by RNA polymerase II (RNAPII) have drawn a great deal of attention. First believed to be required for the transcription of all class II genes, it is now generally accepted that the so-called general transcription factors (GTFs) TFIIA, TFIIB, TFIID, TFIIE, TFIIIF, TFIIH and TFIIS/TCEA1 form the core of the machinery that participates in the transcription of a large fraction of class II genes (reviewed in ref. (Hahn, 2004)). Using this wide definition, the TFIID/STAGA/TFIIIC/PCAF and Mediator complexes can also be classified as GTFs because they also participate in the transcription of a large proportion of mammalian genes (Bjorklund and Gustafsson, 2005b; Muller and Tora, 2004).

However, studies in yeast have revealed that some GTFs are not required for the transcription of a significant number of genes (Fan et al., 2006; Holstege et al., 1998). Similar studies in mammalian cells are still lacking.

Many reports have focused on defining the specific function of the GTFs during transcription. Biochemical analyses have revealed that the transcription reaction involves a number of successive steps that lead to the formation of a pre-mRNA (Hahn, 2004). In the first step, RNAPII positions its catalytic center near the transcriptional initiation site (TIS) of promoter DNA. This first step is preceded by the formation of a preinitiation complex that contains, in addition to RNAPII, a number of general initiation factors including TFIID (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, TFIIB to the BRE, and TAFs of TFIID to the Inr and DPE. Both TFIIA and the Mediator complex have been shown to stimulate transcriptional initiation in reconstituted systems. Formation of the preinitiation complex is accompanied by topological changes including bending and wrapping of promoter DNA around the protein core of the complex (Forget et al., 2004a). In the second step, promoter DNA is melted between nucleotides -9 to +2 in such a way that the template strand becomes accessible for NTP polymerization. Both TFIIF and TFIIE were shown to participate in promoter melting (Pan and Greenblatt, 1994). Two distinct single-stranded DNA helicases, XPB/ERCC3 and XPD/ERCC2, have been identified as components of TFIIH. XPB/ERCC3 was attributed a major role in promoter DNA melting prior to initiation (Bradsher et al., 2000). During the third step, RNAPII

enters a cycle of abortive initiation events wherein the enzyme synthesizes many short 2-10 nt transcripts. TFIIH is responsible for the melting of the template DNA during promoter escape and for the phosphorylation of the C-terminal domain (CTD) of the Rpb1 subunit of RNAPII (Zurita and Merino, 2003). Crystallographic structures of RNAPII suggest that the dissociable heterodimer formed by Rpb4 and Rpb7 stabilizes the RNAPII complex in a clamp-closed conformation which is believed to favour efficient initiation (Armache et al., 2003). This view is consistent with the hypothesis that the Rpb4/Rpb7 heterodimer is dispensable for transcript elongation (Gnatt et al., 2001), as RNAPII adopts the same conformation when bound to DNA even in the absence of Rpb4/Rpb7. Formation of the mRNA is completed through transcript elongation and termination of the transcription reaction. While some reports have proposed a role for both TFIIE and TFIIH in transcript elongation (Blau et al., 1996; Mason and Struhl, 2003), a large body of information has accumulated to support a role for TFIIIF and TFIIS/TCEA1 after promoter clearance (reviewed in(Sims, III et al., 2004a)). TFIIS/TCEA1 can rescue paused or arrested elongation complexes through stimulation of an intrinsic endoribonuclease activity that allows transcript cleavage by backtracked RNAPII (Wang and Hawley, 1993). TFIIIF was shown to stimulate the rate of NTP addition by RNAPII and to modulate the activity of TFIIS/TCEA1 (Zhang and Burton, 2004). However, most of these studies have been performed using *in vitro* systems and the information on the roles of the various GTFs *in vivo* remains fragmentary.

To gain insight into the function of the GTFs in live mammalian cells and determine whether specific GTFs act at early *versus* late stages of the transcription

reaction *in vivo*, we have affinity tagged many components of the RNAPII transcription machinery, expressed the tagged polypeptides at physiological levels in human cells using an inducible system and analyzed their genomic location by chromatin immunoprecipitation (ChIP) coupled to Q-PCR detection of the enriched regions. The location of RNAPII, TFIIA, TFIIB, TFIIE, TFIIF, TFIID/STAGA/TFTC/PCAF, TFIIH, TFIIS/TCEA1 and the Mediator was analyzed both in the region proximal to TIS and further downstream in TRs of 9 human genes that were selected according to (i) their active transcriptional status in human 293 cells, (ii) the minimal divergence of their TIS and (iii) their belonging to various functional classes according to gene ontology (GO) terms. Our results shed light on the function of the GTFs in living cells by showing that although all the factors, including TFIIS/TCEA1, can associate with regions proximal to the TIS, only TFIIF, TFIIS/TCEA1, the Mediator and STAGA/TFTC/PCAF complexes and the heterodimer Rpb4/Rpb7 are detected with RNAPII in TRs situated further downstream along actively transcribed genes.

EXPERIMENTAL

Generation of cell lines expressing tagged polypeptides

Tandem Affinity Purification (TAP)-tagged GTF and RNAPII subunits were cloned and expressed in human EcR-293 cells as previously described (Jeronimo et al., 2004). The following C-terminal TAP-tagged polypeptides were analyzed: RNAPII (Rpb2, Rpb4, Rpb7 and Rpb11), TFIIF (RAP30 and RAP74), TFIIB,

TFIIA (the γ subunit and the $\alpha\beta$ precursor), TFIIE (TFIIE34 and TFIIE56), TFIIH (XPB/ERCC3 and CDK7), TFIID/STAGA/TFTC/PCAF (TAF10 and TAF13), the Mediator (SRB7 and TRFP) and TFIIS/TCEA1. Near physiological expression levels were obtained by inducing the cells for 24 to 48 hrs with 3-6 μM Ponasterone A (Invitrogen). In the case of TFIIS/TCEA1 two conditions were used in parallel: the cells were either exposed to a total dose of 12.5 J/m^2 of UVC radiations and allowed to recover for 20 min. at 37°C or unexposed, with a total of 48 hrs of culture in both cases. For the heat shock the cells were incubated for 1 hr at 42°C in parallel with matched controls.

Affinity purification and identification of proteins by mass spectrometry

Protein affinity purification was performed as we previously described (Jeronimo et al., 2004) (detailed protocols are available on our Web site at <http://www.ircm.qc.ca/microsites/hupi/en/733.html>). The TAP eluates were run on SDS gels, stained with silver or Sypro Ruby (Bio-Rad) and gel slices were excised and digested with trypsin as previously described (Jeronimo et al., 2004). The resulting tryptic peptides were purified and identified by either matrix-assisted laser desorption ionization-time of flight (MALDI-ToF) mass spectrometry (Krogan et al., 2002) or tandem mass spectrometry (LC-MS/MS) with microcapillary reversed-phase high-pressure liquid chromatography coupled to an LCQ DecaXP (ThermoFinnigan) or LTQ (ThermoElectron) quadrupole ion trap mass spectrometer with a nanospray interface. Resulting peptide MS/MS spectra were interpreted using the MASCOT (Matrix Science) software and searched against proteins in either the

National Center for Biotechnology Information (NCBI) non-redundant protein database or the Uniref protein database (Bairoch et al., 2005).

Western blots and antibodies

Proteins from TAP eluates were separated by SDS-PAGE and transferred onto a Polyvinylidene fluoride (PVDF) membrane and probed with the appropriate antibodies. The horseradish peroxidase-conjugated secondary antibodies were purchased from GE Healthcare. Signals were detected using the ECL system (GE Healthcare).

Transcription assay

In vitro transcription reactions were performed as previously described (Jeronimo et al., 2004) using highly purified calf thymus RNAPII and recombinant TBP, TFIIB, TFIIE, TFIIIF and TFIIH. RNAPII, TBP, TFIIE and TFIIH were independently omitted from the reconstituted system and replaced by the eluate of affinity purified Rpb11-TAP, TFIIA $\alpha\beta$ -TAP, TFIIE56-TAP and XPB-TAP. Control reactions in which an eluate from non-induced cells is added were included in each case.

Chromatin immunoprecipitation

TAP-ChIP experiments were performed as previously described (Jeronimo et al., 2004). Briefly, 10^8 cells were cross-linked with formaldehyde at a final concentration of 1%. A cross-link time of 10 min. at room temperature was

generally allowed before adding glycine to a final concentration of 125 mM. A reduced cross-link time of 1 to 2 min. was necessary in the case of the cells expressing the TAP-tagged TFIIIE56, TAF10, TAF13 and SRB7. Chromatin was sonicated to obtain fragments of an average length of 450 bp. After immunoprecipitation with immunoglobulin G (IgG) beads, washing and elution, the cross-links were reversed at 65°C, and the DNA was purified. The enrichment of regions proximal to the TIS and further downstream TRs for the 9 genes was evaluated by Q-PCR as previously described (Jeronimo et al., 2004), using two control regions: an internal sequence of an expressed sequence tag located at chr17:49926700-49927999 and a sequence located in a 1.330 kbp gene-less region of the 13q21.33 band (Nobrega et al., 2003). To ensure that the enrichment obtained was specific to the TAP-tagged factors, the immunoprecipitated DNA from the untagged EcR 293 parental cells was systematically analyzed in parallel. These control experiments allowed us to estimate the background level, which ranged between 0.8- and 1.4-fold enrichment over control regions. In pilot experiments (data not shown), we found that treating the samples with proteinase K at a final concentration of 0.2 µg/µl during the cross-link reversal greatly decreased non-specific enrichment of transcribed regions, which can result at this step, as reported previously (Nagy et al., 2003).

The genes to be analyzed were selected using the databases DBTSS (DataBase of Transcription Start Site, April 2003 version, http://dbtss.hgc.jp/samp_home.html) and MGC (Mammalian Gene Collection, Mars 2003 version, <http://mgc.nci.nih.gov/>). The selection criteria was the minimal

divergence of their alternative TIS, empirically defined as the number of full-length cDNAs situated at a distance of less than 40 bp from the 5' ends provided by RefSeq. Primers were designed by using the Primer3 algorithm (see Table 1 for the sequence of the primers and Supplementary Data, Fig. 2 for the genomic location of the PCR products, determined on the UCSC Genome Browser (Hsu et al., 2005)). The fragments to be analyzed have been chosen based on three criteria: (i) to include the TIS or a position situated at a distance of at least 1000bp from the TIS; (ii) to rank between those with the best score generated by the primer design algorithm; and, (iii) to be unique. All PCR-amplified fragments are unique as verified by BLAT on the April 2003 assembly version of the UCSC Genome Browser (<http://genome.ucsc.edu/>), by Q-PCR melting curve analysis and by agarose gel electrophoresis (data not shown).

In this study, we consider that a region is occupied by a given factor when: (i) reproducible and statistically significant higher enrichment values are obtained for that region compared to control regions and (ii) low background signals are obtained for the same region in a parallel control ChIP experiments conducted with the parental cells in which the tagged factor is not expressed.

RESULTS AND DISCUSSION

TAP-tagged components of the RNAPII machinery are functional

The specific function of the various GTFs that assist mammalian RNAPII during the successive stages of the transcription reaction is largely based on *in vitro*

studies (reviewed in ref. (Hahn, 2004)). Although this biochemical data has proved invaluable to understanding transcription at a fine molecular level, the function of the various GTFs in living mammalian cells remains mainly unaddressed. We reasoned that a systematic analysis of the genomic location of many individual components of the general transcription machinery would help to understand the roles of the GTFs *in vivo*. We took advantage of our expanding collection of cell lines programmed to express, upon induction, physiological levels of polypeptides carrying a TAP tag, including four subunits of RNAPII, two subunits of each TFIIA, TFIIIE, TFIIIF, TFIID/STAGA/TFTC/PCAF, TFIIF, TFIIF, TFIID/STAGA/TFTC/PCAF, TFIIF and the Mediator as well as TFIIB and TFIIS/TCEA1.

To ensure that the tagged polypeptides expressed at physiological levels in human cells behave properly (*e.g.* the tag does not impair their activity), we performed two lines of experiments. First, we affinity purified all the tagged proteins in native conditions, analyzed the eluates using mass spectrometry and defined their network of interactions (see Fig. 1A for an example and Supplementary Data, Fig. 1 for details). Consistent with our previously published results (Jeronimo et al., 2004), RNAPII is linked to TFIIB, TFIIIF, FCP1/CTDP1 and the RNAPII-associated protein RPAP1 *in vivo*. Both RNAPII and TFIIIF are connected to the Mediator complex, which itself is linked to TFIID/STAGA/TFTC/PCAF subunits. TFIID/STAGA/TFTC/PCAF associates with TFIIA and TBP-related factors (TRFs) *in vivo*. Tagged TFIIS/TCEA1 co-purifies with RNAPII subunits, as previously described (Fish and Kane, 2002). Surprisingly, however, the RNAPII molecules that interact with TFIIS/TCEA1 appear to be

mainly in the hypophosphorylated Ila form (Fig. 1B); purifications from whole-cell lysates prepared in the presence of phosphatase inhibitors indicated that this finding is not merely the consequence of CTD dephosphorylation during the preparation of the cell extracts (data not shown). Because RNAPII with a hypophosphorylated Rpb1 C-terminal domain (CTD) represents the form of the enzyme that is recruited to promoter DNA prior to transcription initiation, our results support the notion that TFIIS/TCEA1 can associate with RNAPII at promoters *in vivo* (see below). Also surprising is the finding that tagged Rpb11 co-purifies with TFIIS.1/TCEA2 (see Supplementary Data, Fig. 1). Comparison of the amino acid sequence of TFIIS/TCEA1 with TFIIS.1/TCEA2 shows that the C-terminal region, the portion of the polypeptide known to be necessary and sufficient for activity *in vitro* (Nakanishi et al., 1995), is highly conserved (Fig. 1C). A previous study has shown that the amino-terminal region of human TFIIS/TCEA1 interacts with RNAPII holoenzyme, whereas its C-terminal portion binds free RNAPII (Pan et al., 1997). To our knowledge this is the first time that the protein encoded by the TFIIS.1/TCEA2 gene is purified and shown to associate with RNAPII. Notably, both TFIIE and TFIIH are only weakly linked to each other and to other GTFs in our network (Supplementary Data, Fig. 1). Both CDK7 and XPB/ERCC3 co-purify with the DNA repair factor ERCC5/XPG; XPB/ERCC3 co-purifies with BCR (Fig. 1A), the polypeptide forming the chimeric BCR-ABL protein in the Philadelphia chromosome. ERCC5/XPG connects TFIIH to TFIID/STAGA/TFTC/PCAF (Supplementary Data, Fig. 1). Together, these results indicate that the tagged factors have the ability

to associate with their endogenous interaction partners *in vivo* and reveal new aspects of the TFIIS-RNAPII interaction in live mammalian cells.

Secondly, the ability of the purified GTFs and RNAPII to support transcription was assessed. The GTFs and RNAPII affinity purified using their tagged subunits were able to replace their highly-purified counterparts in *in vitro* transcription assays (Fig. 1D). Together, the results of the proteomic analysis and the biochemical assays support the notion that the tagged polypeptides associate with their cognate partners to form functionally active transcription factors in mammalian cells.

TAP-ChIP as a systematic method for the localization of proteins along mammalian genomic DNA

The genomic location analysis of transcription factors by chromatin immunoprecipitation in a systematic fashion is known to be difficult mainly because of the requirement of antibodies able to efficiently immunoprecipitate proteins after a cross-linking step. TAP-ChIP experiments used IgG beads, which target the Protein A moiety of the TAP tag, to pull-down the TAP-tagged polypeptides covalently cross-linked to chromatin fragments. This method allowed us to systematically localize many components of the RNAPII general transcription machinery. Pilot experiments have revealed that the second affinity purification step (*e.g.* calmodulin beads) does not significantly improve the enrichment values of DNA fragments in our conditions (data not shown). After purification, enriched DNA sequences were analyzed using Q-PCR with primer sets designed for

detecting both regions proximal to the TIS and further downstream TRs of 9 genes actively transcribed in 293 cells.

The ability to discriminate between the location of a protein either in regions close to the TIS *versus* further downstream TRs of a gene demands that: (i) the TIS be known with accuracy and (ii) alternative, distant TIS not be used. To meet these requirements, we developed a procedure that allowed us to select genes having well characterized TIS. Two databases reporting on experimentally determined human TISs using various methodologies were compared for the minimal divergence of alternative TISs. Among the selected genes, 9 were chosen for (i) their belonging to various GO classes, (ii) their activity in human 293 cells as verified by quantitative RT-PCR (data not shown) and (iii) our ability to design primer sets that are efficient in Q-PCR detection.

In this study, the regions to be analysed by Q-PCR were selected in such a way that early *versus* late transcription events can be monitored. A distinction between early steps of the reaction (*e.g.* promoter binding, first-bond formation and promoter clearance) is difficult to address principally because of the limitation imposed by the resolution of the ChIP technique which depends on the size of the sonicated fragments (around 450 bp in our case). In addition, it is conceivable that the efficiency of bringing down a promoter fragment would be much higher when a protein is associated with a tightly bound, engaged polymerase as compared to a loosely bound polymerase in a preinitiation complex.

The results of our TAP-ChIP experiments (Fig. 2) indicate that, as expected, TFIIA, TFIIB, TFIID, TFIIE, TFIIF, TFIIH, and the Mediator assemble with

RNAPII in regions proximal to the TIS of transcribed genes in live mammalian cells. The analysis of a second subunit confirmed the results obtained for these factors (see Supplementary Data, Fig. 3). Moreover, when N-tagged factors or specific antibodies were used in validation experiments, we obtained similar location profiles as when using the C-tagged version (data not shown). The lack of detection of initiation factors in some regions close to the TIS may reflect a low rate of initiation events (*e.g.* under the detection limit of the ChIP procedure used here).

Our results also indicate that the transcription elongation factor TFIIS/TCEA1 is not stably recruited along the tested genes under normal growth conditions (data not shown). Notably, TFIIS/TCEA1 was detected at a number of regions close to the TIS and in further downstream TRs after cell exposure to UVC radiations (Fig. 2), suggesting a role for TFIIS/TCEA1 both at early and late stages of transcription. Cyclobutane pyrimidine dimers, the major DNA lesions induced by UVC radiations, are known to arrest the polymerase and to induce the recruitment of TFIIS/TCEA1 *in vitro* (Donahue et al., 1994). Whether the association of TFIIS/TCEA1 with genomic DNA is simply a consequence of an increased number of stalled elongating RNAPII molecules along damaged DNA or the signature of an active role for this factor in transcription coupled DNA repair (TCR), as suggested by genetic studies in yeast (Wong and Ingles, 2001), is not clear. A recent study in human cells showed that TFIIS is recruited to lesion-arrested RNAPII via the CSA and CSB proteins, suggesting a non-specific role of TFIIS in TCR (Fousteri et al., 2006). Although the location of TFIIS/TCEA1 at regions close to the TIS does not necessarily imply that it has a function in transcription initiation, the finding that

TFIIS/TCEA1 also associates with the hypophosphorylated form of RNAPII (see Fig. 1B) fully supports the conclusion that it plays a role in very early stages of transcription, as recently proposed in yeast TFIIS/TCEA1 (Prather et al., 2005).

Two TAFs, TAF10 and TAF13, were detected along our tested genomic regions with very similar occupancy profiles of regions close to the TIS but distinct profiles in further downstream TRs (Fig. 2). Both TAFs were detected at half of the regions close to the TIS analyzed, indicating that the complex recruited is TFIID, TAF13 being a TFIID-specific subunit (Martinez, 2002). Of note, at some regions close to the TIS (3 out of 8), neither TAF10 nor TAF13 were detected, which is in agreement with a previous report showing that these TAFs are required for the expression of only a fraction of the yeast genes (Huisinga and Pugh, 2004a). Our results indicate that TAF10, a component of both TFIID and the STAGA/TFTC/PCAF complexes (reviewed in ref. (Martinez, 2002)), is recruited to transcribed regions (see Fig. 2 and data not shown) and suggest a role for this TAF during transcriptional elongation of specific genes. The observation that TAF13 is not recruited to these regions suggests the TAF10-containing complex present at these locations is STAGA, TFTC or PCAF. While the function of SAGA at different stages of the transcription reaction is supported by a number of reports (see ref. (Timmers and Tora, 2005) for a recent review), its human homologue complexes STAGA/TFTC/PCAF are only poorly characterized. In support to our observation, the STAGA complex has been implicated in diverse transcription-coupled processes such as chromatin modification, pre-mRNA splicing and DNA repair (Martinez et al., 2001).

The Mediator subunit SRB7 was detected at a fraction of the regions close to the TIS that we have tested (Fig. 2), which is in agreement with data reported by Gromoller *et al.* (Gromoller and Lehming, 2000) showing that Srb7p is essential for the activation of a subset of genes and with recent observations in yeast that Mediator subunits are recruited to the promoter of highly transcribed genes (Andrau *et al.*, 2006; Zhu *et al.*, 2006). Unexpectedly, SRB7 was also localized in one transcribed region, suggesting a role in late stages of transcription for this coactivator.

TFIIF is involved in post-initiation stages of transcription *in vivo*

As shown in Figure 3, our results indicate that the TFIIF subunit RAP30 co-localizes with RNAPII in many regions located close to the TIS but also in further downstream TRs, demonstrating for the first time a role for this factor in transcriptional elongation *in vivo*. The analysis of the other TFIIF subunit RAP74 confirmed this observation (see Supplementary Data, Fig. 3). This result contrasts with *in vivo* data in yeast (Krogan *et al.*, 2002), but fully supports previous *in vitro* results (reviewed in (Hahn, 2004; Hampsey, 1998; Orphanides *et al.*, 1996; Woychik and Hampsey, 2002)) showing that TFIIF acts at both the initiation and elongation stages of the transcription reaction and those of genetic screens indicating a complex interdependence between TFIIF and the elongation factor TFIIIS/TCEA1 (Fish *et al.*, 2006).

It is interesting to note the gene-specific differences in the relative occupancy of TFIIF and RNAPII (Fig. 3): while the RNAPII subunit Rpb11 showed

higher occupancy values than the TFIIF subunit RAP30 at a majority of the regions located close to the TIS, the relative occupancy ratio is inverted in the case of the POLR2E gene where RAP30 showed a higher signal than Rpb11. Furthermore, RNAPII/TFIIF occupancy ratios significantly differ between regions close to the TIS and those situated further downstream in TRs as, for half of the downstream TRs occupied by RAP30, this polypeptide showed higher occupancy signals than Rpb11. This finding may reflect the presence of transcription complexes with different polypeptide composition at these positions, affecting the accessibility of the tag and/or the cross-linking efficiency. This is supported by a recent paper showing that RNAPII stalled close to the promoter *in vitro* can not respond to TFIIF, and enter in productive elongation only after the Positive Transcription Elongation Factor b (P-TEFb) has functioned, at which point it begins to elongate and move rapidly away from the promoter (Cheng and Price, 2007). Another possible explanation is that TFIIF is preferentially recruited on certain sequences in order to support elongation and to suppress pausing of the polymerase as previously suggested (Zhang and Burton, 2004).

Rpb7 associates with both the regions proximal to the TIS and with further downstream TRs of active genes

Previous reports showed that Rpb7 is required for the initiation but not the elongation stage of transcription *in vitro* (Orlicky et al., 2001; Rosenheck and Choder, 1998). *In vivo* studies provided evidence that human Rpb7 can interact with transcriptional activators such as the oncogenic fusion protein EAD-Flt1 (Zhou and

Lee, 2001) and regulate transcription of genes such as VEGF (Na et al., 2003). Nevertheless, no previous study addressed the direct implication of this RNAPII subunit at distinct stages of transcription *in vivo*.

ChIP experiments using a TAP-tagged version of Rpb7 show that it associates with regions close to the TIS and, surprisingly, with further downstream TRs as well (Fig. 4). As Rpb4 and Rpb7 are known to form a heterodimer which can dissociate from the polymerase (reviewed in ref. (Choder, 2004)), we generated and used cell lines expressing the TAP-tagged version of Rpb4 in order to confirm the Rpb7 results. Indeed, Rpb4 was also detected in regions close to the TIS and in further downstream TRs (see Supplementary Data, Fig. 3). Neither Rpb7 nor Rpb4 were shown to have a direct function in transcription elongation to date (see ref. (Choder, 2004) for a recent review).

Because the Rpb4/Rpb7 dimer was previously suggested to have a role in the transcriptional response to stress, we next addressed the possible implication of Rpb7 in heat shock. As shown in Figure 4C, and similarly to the core RNAPII subunit Rpb11, Rpb7 occupies all the regions close to the TIS after a heat shock of 1 hr at 42°C. We noted a significant increase (40% or higher) in the enrichment of most regions close to the TIS (*e.g.* 7 out of 8) by tagged Rpb7 when comparing heat-shock to non heat-shock conditions. Similar results were obtained with the tagged version of Rpb4, for which we observed a significant increase (>50%) in the enrichment of regions close to the TIS after a heat shock (data not shown). Notably, even at regions close to the TIS showing a decreased Rpb11 occupancy after heat shock (Fig. 4C, for example SFRS2), a higher Rpb7 occupancy is observed.

Conversely, in the case of POLR2E, a marked decrease in Rpb11 occupancy is accompanied by a lower Rpb7 occupancy after heat shock. This observation argues in favour of a function for the Rpb4/Rpb7 heterodimer in the stabilization of RNAPII during the early stages of transcription under stress conditions.

The same TRs were occupied by Rpb7 before and after heat shock, with similar or increased enrichments, suggesting that the requirement for this subunit during the elongation stage of transcription is maintained in stress conditions (Fig. 4D). Of note, Rpb7 and Rpb4 were both detected in one transcribed region which was not occupied by these polypeptides in normal conditions (Fig. 4B, 4D and data not shown). In order to address in more detail the requirement of these subunits for transcriptional elongation under stress conditions, we designed primer sets along the entire transcribed region of two constitutive heat-shock genes, known to present increased expression levels in response to a heat shock: enolase 1 (ENO1) and the heat shock 70kDa protein 8 (HSPA8). The change in the occupancy pattern observed for the RNAPII core subunit Rpb11 is in agreement with an increased expression of these genes in response to a heat shock, presenting higher occupancy values compared to normal conditions (Fig 4) and becoming detectable in downstream TRs which are not occupied under normal condition (Fig. 4E and 4F). In contrast to the core subunit Rpb11, Rpb7 occupied only the TRs proximal to the TIS under normal conditions (up to around +1200) of both genes, while Rpb11 is detected much further downstream in the same conditions (Fig. 4E and 4F). This does not seem to be related to a detection limitation in the case of Rpb7 since occupancy values for these two RNAPII subunits are either similar or Rpb7 presents

even higher occupancy values than Rpb11 in the TRs proximal to the TIS of these two genes. Interestingly, after heat shock, both Rpb7 and Rpb11 are detected down to the end of both TRs. Moreover, Rpb7 is detected at the 3' extremity of the HSPA8 gene, which is in agreement with the hypothesis of its involvement in the formation of 3' termini of the mRNA (Mitsuzawa et al., 2003). We cannot rule out the possibility that Rpb7 remains associated with the elongating polymerase all along the genes under normal conditions, but if this is the case, its conformation or its position in the elongation complex must be different from those under heat shock. Our results showing that Rpb4 is recruited to TRs (Supplementary Data, Fig. 3) confirm that the Rpb4/Rpb7 heterodimer has a function during transcriptional elongation. It has been previously shown that Rpb4 genetically interacts with the elongation factor TFIIS/TCEA1 and with the RNAPII subunit Rpb9, which also plays a role in elongation (Awrey et al., 1997; Wery et al., 2004). Moreover, it has been shown that yeast Rpb4 is a regulator of transcription-coupled DNA repair pathways (Li and Smerdon, 2002). Taken together, our results suggest that the Rpb4/Rpb7 heterodimer is involved in both the initiation and elongation stages of transcription, having a specific role in stress conditions such as heat shock.

In conclusion, we used a systematic affinity tagging procedure coupled to ChIP experiments and gene-specific Q-PCR detection of the enriched genomic regions to determine the location of components of the general transcription machinery along active class II genes in human cells. Based on the location of the various polypeptides, our results help define their function during the transcription reaction in living mammalian cells.

ACKNOWLEDGEMENTS

We are grateful to the members of our laboratory, François Robert and members of his laboratory and Jacques Archambault for helpful discussions and comments on this manuscript. We are also grateful to Jean-Marc Egly (IGBMC, Strasbourg, France) for antibodies. 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, Genome Canada and Genome Québec. M.C. and C.J. hold studentships from the Canadian Institutes for Health Research and the Fonds de recherche en santé du Québec.

REFERENCES

- 1 Hahn, S. (2004) Structure and mechanism of the RNA polymerase II transcription machinery. *Nat.Struct.Mol Biol* **11**, 394-403
- 2 Muller, F. and Tora, L. (2004) The multicoloured world of promoter recognition complexes. *EMBO J.* **23**, 2-8
- 3 Bjorklund, S. and Gustafsson, C. M. (2005) The yeast Mediator complex and its regulation. *Trends Biochem.Sci.* **30**, 240-244
- 4 Holstege, F. C., Jennings, E. G., Wyrick, J. J., Lee, T. I., Hengartner, C. J., Green, M. R., Golub, T. R., Lander, E. S., and Young, R. A. (1998) Dissecting the regulatory circuitry of a eukaryotic genome. *Cell* **95**, 717-728
- 5 Fan, X., Chou, D. M., and Struhl, K. (2006) Activator-specific recruitment of Mediator in vivo. *Nat.Struct.Mol.Biol.* **13**, 117-120
- 6 Forget, D., Langelier, M. F., Therien, 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**, 1122-1131
- 7 Pan, G. and 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
- 8 Bradsher, J., Coin, F., and Egly, J. M. (2000) Distinct roles for the helicases of TFIIH in transcript initiation and promoter escape. *J.Biol.Chem.* **275**, 2532-2538
- 9 Zurita, M. and Merino, C. (2003) The transcriptional complexity of the TFIIH complex. *Trends Genet.* **19**, 578-584
- 10 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
- 11 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
- 12 Mason, P. B. and Struhl, K. (2003) The FACT complex travels with elongating RNA polymerase II and is important for the fidelity of transcriptional initiation in vivo. *Mol.Cell Biol.* **23**, 8323-8333

- 13 Blau, J., Xiao, H., McCracken, S., O'Hare, P., Greenblatt, J., and Bentley, D. (1996) Three functional classes of transcriptional activation domain. *Mol.Cell Biol.* **16**, 2044-2055
- 14 Sims, R. J., III, Belotserkovskaya, R., and Reinberg, D. (2004) Elongation by RNA polymerase II: the short and long of it. *Genes Dev.* **18**, 2437-2468
- 15 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
- 16 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
- 17 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
- 18 Krogan, N. J., Kim, M., Ahn, S. H., Zhong, G., Kobor, M. S., Cagney, G., Emili, A., Shilatifard, A., Buratowski, S., and Greenblatt, J. F. (2002) RNA polymerase II elongation factors of *Saccharomyces cerevisiae*: a targeted proteomics approach. *Mol.Cell Biol.* **22**, 6979-6992
- 19 Bairoch, A., Apweiler, R., Wu, C. H., Barker, W. C., Boeckmann, B., Ferro, S., Gasteiger, E., Huang, H., Lopez, R., Magrane, M., Martin, M. J., Natale, D. A., O'Donovan, C., Redaschi, N., and Yeh, L. S. (2005) The Universal Protein Resource (UniProt). *Nucleic Acids Res.* **33**, D154-D159
- 20 Nobrega, M. A., Ovcharenko, I., Afzal, V., and Rubin, E. M. (2003) Scanning human gene deserts for long-range enhancers. *Science* **302**, 413
- 21 Nagy, P. L., Cleary, M. L., Brown, P. O., and Lieb, J. D. (2003) Genomewide demarcation of RNA polymerase II transcription units revealed by physical fractionation of chromatin. *Proc.Natl.Acad.Sci.U.S.A* **100**, 6364-6369
- 22 Hsu, F., Pringle, T. H., Kuhn, R. M., Karolchik, D., Diekhans, M., Haussler, D., and Kent, W. J. (2005) The UCSC Proteome Browser. *Nucleic Acids Res.* **33**, D454-D458
- 23 Fish, R. N. and Kane, C. M. (2002) Promoting elongation with transcript cleavage stimulatory factors. *Biochim.Biophys.Acta* **1577**, 287-307
- 24 Nakanishi, T., Shimoaraiso, M., Kubo, T., and Natori, S. (1995) Structure-function relationship of yeast S-II in terms of stimulation of RNA polymerase

- II, arrest relief, and suppression of 6-azauracil sensitivity. *J.Biol.Chem.* **270**, 8991-8995
- 25 Pan, G., Aso, T., and Greenblatt, J. (1997) Interaction of elongation factors TFIIS and elongin A with a human RNA polymerase II holoenzyme capable of promoter-specific initiation and responsive to transcriptional activators. *J.Biol.Chem.* **272**, 24563-24571
 - 26 Donahue, B. A., Yin, S., Taylor, J. S., Reines, D., and Hanawalt, P. C. (1994) Transcript cleavage by RNA polymerase II arrested by a cyclobutane pyrimidine dimer in the DNA template. *Proc.Natl.Acad.Sci.U.S.A* **91**, 8502-8506
 - 27 Wong, J. M. and Ingles, C. J. (2001) A compromised yeast RNA polymerase II enhances UV sensitivity in the absence of global genome nucleotide excision repair. *Mol Gen.Genet.* **264**, 842-851
 - 28 Fousteri, M., Vermeulen, W., van Zeeland, A. A., and Mullenders, L. H. (2006) Cockayne syndrome A and B proteins differentially regulate recruitment of chromatin remodeling and repair factors to stalled RNA polymerase II in vivo. *Mol.Cell* **23**, 471-482
 - 29 Prather, D. M., Larschan, E., and Winston, F. (2005) Evidence that the elongation factor TFIIS plays a role in transcription initiation at GAL1 in *Saccharomyces cerevisiae*. *Mol Cell Biol* **25**, 2650-2659
 - 30 Martinez, E. (2002) Multi-protein complexes in eukaryotic gene transcription. *Plant Mol Biol* **50**, 925-947
 - 31 Huisinga, K. L. and Pugh, B. F. (2004) A genome-wide housekeeping role for TFIID and a highly regulated stress-related role for SAGA in *Saccharomyces cerevisiae*. *Mol.Cell* **13**, 573-585
 - 32 Timmers, H. T. and Tora, L. (2005) SAGA unveiled. *Trends Biochem.Sci.* **30**, 7-10
 - 33 Martinez, E., Palhan, V. B., Tjernberg, A., Lymar, E. S., Gamper, A. M., Kundu, T. K., Chait, B. T., and Roeder, R. G. (2001) Human staga complex is a chromatin-acetylating transcription coactivator that interacts with pre-mrna splicing and dna damage-binding factors in vivo. *Mol.Cell Biol.* **21**, 6782-6795
 - 34 Gromoller, A. and Lehming, N. (2000) Srb7p is essential for the activation of a subset of genes. *FEBS Lett.* **484**, 48-54
 - 35 Zhu, X., Wiren, M., Sinha, I., Rasmussen, N. N., Linder, T., Holmberg, S., Ekwall, K., and Gustafsson, C. M. (2006) Genome-wide occupancy profile of

- mediator and the srb8-11 module reveals interactions with coding regions. *Mol.Cell* **22**, 169-178
- 36 Andrau, J. C., van de, P. L., Lijnzaad, P., Bijma, T., Koerkamp, M. G., van de, P. J., Werner, M., and Holstege, F. C. (2006) Genome-Wide Location of the Coactivator Mediator: Binding without Activation and Transient Cdk8 Interaction on DNA. *Mol.Cell* **22**, 179-192
 - 37 Woychik, N. A. and Hampsey, M. (2002) The RNA Polymerase II Machinery. Structure Illuminates Function. *Cell* **108**, 453-463
 - 38 Orphanides, G., Lagrange, T., and Reinberg, D. (1996) The general transcription factors of RNA polymerase II. *Genes Dev.* **10**, 2657-2683
 - 39 Hampsey, M. (1998) Molecular genetics of the RNA polymerase II general transcriptional machinery. *Microbiol.Mol.Biol.Rev.* **62**, 465-503
 - 40 Fish, R. N., Ammerman, M. L., Davie, J. K., Lu, B. F., Pham, C., Howe, L., Ponticelli, A. S., and Kane, C. M. (2006) Genetic interactions between TFIIF and TFIIS. *Genetics* **173**, 1871-1884
 - 41 Cheng, B. and Price, D. H. (2007) Properties of RNA polymerase II elongation complexes before and after the P-TEFb-mediated transition into productive elongation. *J.Biol Chem.* **282**, 21901-21912
 - 42 Orlicky, S. M., Tran, P. T., Sayre, M. H., and 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
 - 43 Rosenheck, S. and Choder, M. (1998) Rpb4, a subunit of RNA polymerase II, enables the enzyme to transcribe at temperature extremes in vitro. *J.Bacteriol.* **180**, 6187-6192
 - 44 Zhou, H. and Lee, K. A. (2001) An hsRBP4/7-dependent yeast assay for trans-activation by the EWS oncogene. *Oncogene* **20**, 1519-1524
 - 45 Na, X., Duan, H. O., Messing, E. M., Schoen, S. R., Ryan, C. K., di Sant'Agnese, P. A., Golemis, E. A., and Wu, G. (2003) Identification of the RNA polymerase II subunit hsRBP7 as a novel target of the von Hippel-Lindau protein. *EMBO J.* **22**, 4249-4259
 - 46 Choder, M. (2004) Rpb4 and Rpb7: subunits of RNA polymerase II and beyond. *Trends Biochem.Sci.* **29**, 674-681
 - 47 Mitsuzawa, H., Kanda, E., and Ishihama, A. (2003) Rpb7 subunit of RNA polymerase II interacts with an RNA-binding protein involved in processing of transcripts. *Nucleic Acids Res.* **31**, 4696-4701

- 48 Wery, M., Shematorova, E., Van Driessche, B., Vandenhoute, J., Thuriaux, P., and Van, M., V (2004) Members of the SAGA and Mediator complexes are partners of the transcription elongation factor TFIIS. *EMBO J.* **23**, 4232-4242
- 49 Awrey, D. E., Weilbaecher, R. G., Hemming, S. A., Orlicky, S. M., Kane, C. M., and Edwards, A. M. (1997) Transcription elongation through DNA arrest sites. A multistep process involving both RNA polymerase II subunit RPB9 and TFIIS. *J.Biol.Chem.* **272**, 14747-14754
- 50 Li, S. and Smerdon, M. J. (2002) Rpb4 and Rpb9 mediate subpathways of transcription-coupled DNA repair in *Saccharomyces cerevisiae*. *EMBO J.* **21**, 5921-5929

Table 1. Sequence of the primers used to analyze the location of various polypeptides

The position of the PCR products relative to the assigned TIS is indicated for each primer set.

Official gene symbol	Gene length (bp)	Primer sequence	PCR product position
POLA2	35000	TIS - Fw aaagcaaggggaagggtt	-61 +123
		TIS - Re gaatggagggagcagaaat	
		Trx - Fw ggtctgtgcaaatgaatc	
		Trx - Re atgactccttacctgctctac	+988 +1199
GTF2F2	163565	TIS - Fw ttctcagttatgctgacc	-238 0
		TIS - Re ttacctccagaactg	
		Trx - Fw ctaagaggtcttctgtcg	
		Trx - Re actattctgggtatgacagg	+978 +1186
FTL	1559	TIS - Fw gctgagactcctatgtgct	-157 +33
		TIS - Re acactgttgaagcaagagac	
		Trx - Fw tatagaagccagctgaagat	
		Trx - Re gtgaaatgaggctctgaa	+957 +1176
TIMM17A	15141	TIS - Fw aaccaatgctcatagacctt	-80 +84
		TIS - Re gcaagagaaatgcaagagac	
		Trx - Fw caggacctaaataactactctg	
		Trx - Re tattacagtgtgctgagctctac	+910 +1094
GNB2L1	6964	TIS - Fw gaatgtgctgtttcagagt	-200 +55
		TIS - Re atggcttagagaactagca	
		Trx - Fw aaaggggtctgtatttctg	
		Trx - Re ttaactagagatcggttcc	+998 +1165
SFRS2	3215	TIS - Fw gccaatcagaaggtttcat	-86 +17
		TIS - Re gcacctgagtaacaactgg	
		Trx - Fw gaaggtccaagccaagt	
		Trx - Re gattcccagacattacat	+1005 +1194
POLR2E	7170	TIS - Fw aactgccgctctgtaag	-161 +90
		TIS - Re atgatggctctgaggatt	
		Trx - Fw gagatagggttctattctg	
		Trx - Re gaccagtatgatcctgagag	+812 +1093
ENO1	17717	TIS - Fw ggtgagggaaatgagtgac	-3 +151
		TIS - Re accgaggtgaacgtaaag	
		Trx 1 - Fw taggccaagaaggatgat	
		Trx 1 - Re gaattagggacacggtaaat	+1140 +1289
		Trx 2 - Fw gaacaagacctcaattgcta	
		Trx 2 - Re agcactggactaaactagg	+3428 +3702
		Trx 3 - Fw gagaattgtgaaactcttc	
		Trx 3 - Re gtgactcacagatgggac	+5373 +5558
		Trx 4 - Fw gcacaagtttagagggttta	
		Trx 4 - Re cagctccttcaattct	+17106 +17321
HSPA8	4643	TIS - Fw ctgtgattgggtctgta	-141 +88
		TIS - Re aactcttgagcagaggtt	
		Trx 1 - Fw acctggagtcattgtagta	
		Trx 1 - Re aataccattatccctgtcaa	+1175 +1463
		Trx 2 - Fw ctgaaatctggataacgtaggag	
		Trx 2 - Re ctctccctgtatttactggac	+2026 +2324
		Trx 3 - Fw gtaccatttggatgcaagttc	
		Trx 3 - Re tacagctcttcttctactgat	+3265 +3549
		Trx 4 - Fw gtcagggagaaagaagggttatta	
		Trx 4 - Re atgtgtggaacaatgctacatctac	+4152 +4439

FIGURE LEGENDS**Figure 1 Tandem affinity purification (TAP) of the RNAPII general transcription machinery**

(A) SYPRO-stained SDS gels showing the CDK7- and XPB/ERCC3-TAP eluates. Although the identity of most visible bands is known, only interaction partners that were validated through examination of the literature are shown. Gene symbols and common names are provided. Tagged polypeptides are indicated by an asterisk. Western blots validating the presence of the TFIIF subunits as well as XPG and BCR in both CDK7- and XPB/ERCC3-TAP eluates are shown. (B) TAP-tagged TFIIS/TCEA1 co-purifies with Rpb1 carrying a hypophosphorylated C-terminal domain (CTD). The N-20 antibody, which detects both the hypophosphorylated (IIa) and hyperphosphorylated (IIo) forms of RNAPII in a whole cell extract (WCE) by being directed to the N-terminal part of Rpb1, revealed the presence of hypophosphorylated RNAPII in both the Rpb11- and TFIIS/TCEA1-TAP eluates. (C) Amino acid sequence alignment of the TFIIS/TCEA1 and TFIIS.1/TCEA2 proteins. The peptides sequences obtained by MS/MS that identify the TFIIS.1/TCEA2 protein in the Rpb11-TAP eluate are boxed. (D) *In vitro* transcription reactions were reconstituted using calf thymus RNAPII in the presence of the classically purified general transcription factors TBP, TFIIB, TFIIF and TFIIF. The linearized DNA template carries the AdML promoter and directs the synthesis of a 391-nt transcript. The TAP-tagged GTFs (TFIIA $\alpha\beta$ -TAP, TFIIE56-TAP and XPB-TAP) and RNAPII (Rpb11-TAP) were used to replace their cognate

highly-purified counterparts after their omission from the reconstituted system (All). A control reaction performed with an eluate obtained from non-induced cells is included.

Figure 2 Occupancy of regions proximal to the transcriptional initiation site (TIS) and further downstream transcribed regions (TRs) of selected human genes by the GTF and RNAPII subunits obtained by chromatin immunoprecipitation

Fold enrichment of regions located close to the TIS (A) and further downstream TRs (B) over control regions are represented for each TAP tagged factor, including subunits of RNAPII (Rpb11) and the GTFs TFIIA (TFIIA $\alpha\beta$), TFIIB, TFIIE (TFIIE56), TFIIH (XPB), TFIID/STAGA (TAF10 and TAF13), the Mediator (SRB7) and TFIIS/TCEA1. Color code: Black: >5-fold enrichment; Grey: significantly enriched with less than 5-fold enrichment values; White: not significantly enriched; X: not determined. The number of regions situated close to the TIS and further downstream TRs occupied by each factor is indicated as a ratio (observed/total tested). The results shown for TFIIS/TCEA1 were obtained by the analysis of cells cross-linked after an irradiation with a total dose of 12.5 J/m² of UVC (TFIISexp). Student's *t*-tests were performed to determine the significance of the enrichment values obtained in four experiments, including at least two independent ChIP assays.

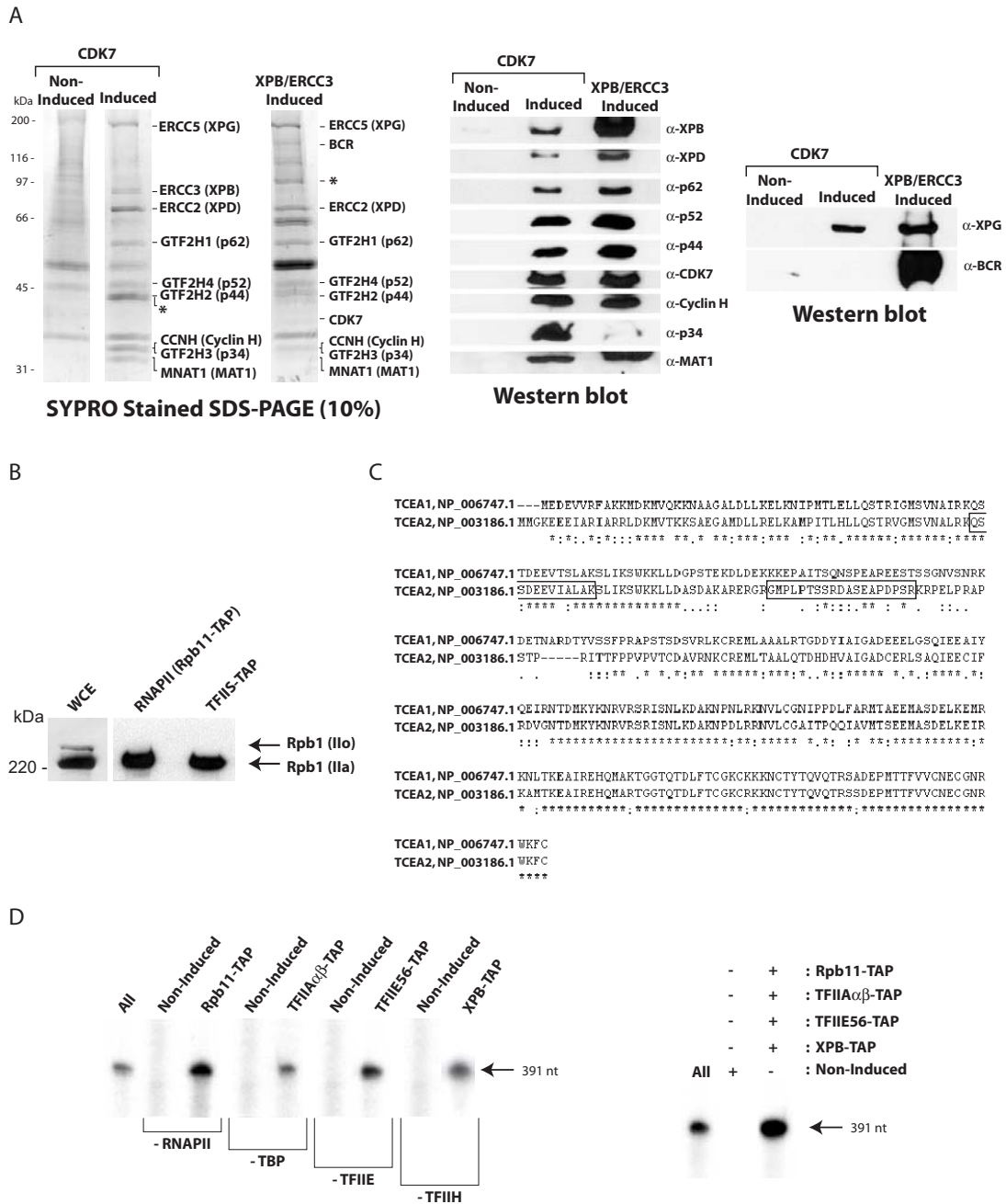
Figure 3 Occupancy of regions proximal to the transcriptional initiation site (TIS) and further downstream transcribed regions (TRs) of selected human genes by the the TFIIIF subunit RAP30 and the RNAPII subunit Rpb11 as determined by chromatin immunoprecipitation

Fold enrichment of regions situated close to the TIS (A) and further downstream TRs (B) over control regions are represented for 9 selected genes. Student's *t*-tests were performed to determine the significance of the enrichment values obtained in four experiments, including at least two independent ChIP assays (* $P < 0.05$, ** $P \leq 0.001$).

Figure 4 Occupancy of regions proximal to the transcriptional initiation site (TIS) and further downstream transcribed regions (TRs) of selected human genes by Rpb7 as determined by chromatin immunoprecipitation

(A, B) Fold enrichment of regions situated close to the TIS (A) and further downstream TRs (B) over control regions are represented for Rpb7 and Rpb11 when cells are cultured in normal conditions. (C, D) Fold enrichment of regions situated close to the TIS (C) and further downstream TRs (D) after a heat shock (1 hr at 42°C) (HS) are represented for Rpb7 and Rpb11. (E, F) The distribution along the genes ENO1 (E) and HSPA8 (F) is represented for Rpb7 and Rpb11 cultured in normal conditions or after a heat shock (HS). Color code: Black: >10-fold enrichment; Dark grey: between 5- and 10-fold enrichment; Light grey: significantly enriched with less than 5-fold enrichment values; White: not significantly enriched. Student's *t*-tests were performed to determine the significance of the enrichment

values obtained in four experiments, including at least two independent ChIP assays
(* $P < 0.05$, ** $P \leq 0.001$).



		TIS							Downstream TR											
		FTL	GNB2L1	EN1	POLA2	SFRS2	TIMM17A	POLR2E	HSP8A	FTL	GNB2L1	EN1	POLA2	SFRS2	TIMM17A	POLR2E	HSP8A	GTF2F2		
RNAPII	Rpb11	■	■	■	■	■	■	■	■	8/8	■	■	■	■	■	■	■	■	■	8/9
TFIIA	TFIIA $\alpha\beta$	■	■	■	■	■	■	■	■	4/7										0/8
TFIIB		■	■	■	■	■	■	■	■	7/8										0/9
TFIIE	E56	■	■	■	■	■	■	■	■	3/7										0/8
TFIIH	XPB	■	■	■	■	■	■	■	■	1/7										6/9
STAGA	TAF10	■	■	■	■	■	■	■	■	4/8	■	■	■	■	■	■	■	■	■	1/9
TFIID	TAF13	■	■	■	■	■	■	■	■	5/8										0/9
Mediator	SRB7	■	■	■	■	■	■	■	■	5/8	■	■	■	■	■	■	■	■	■	1/9
TFIIS	TFIISexp	■	■	■	■	■	■	■	■	3/7	■	■	■	■	■	■	■	■	■	3/8

Figure 2

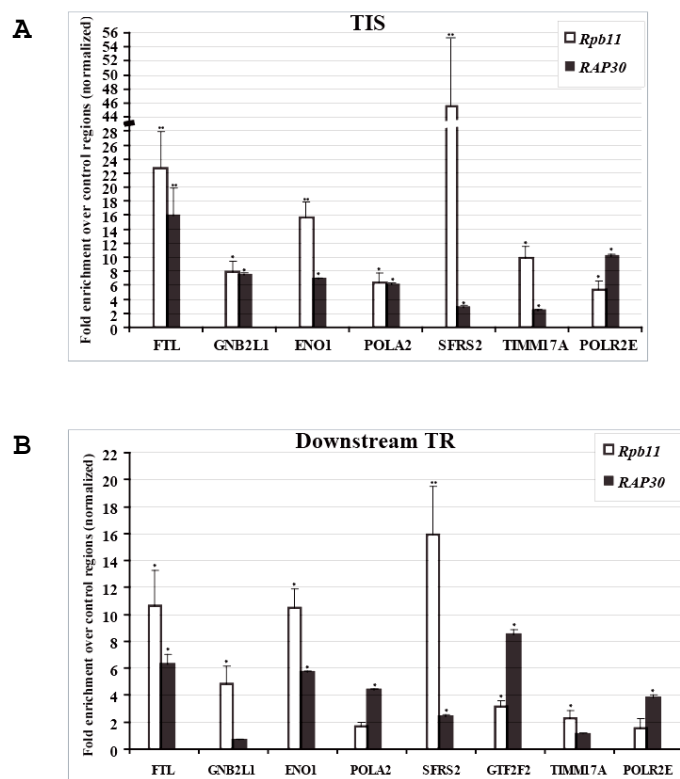


Figure 3

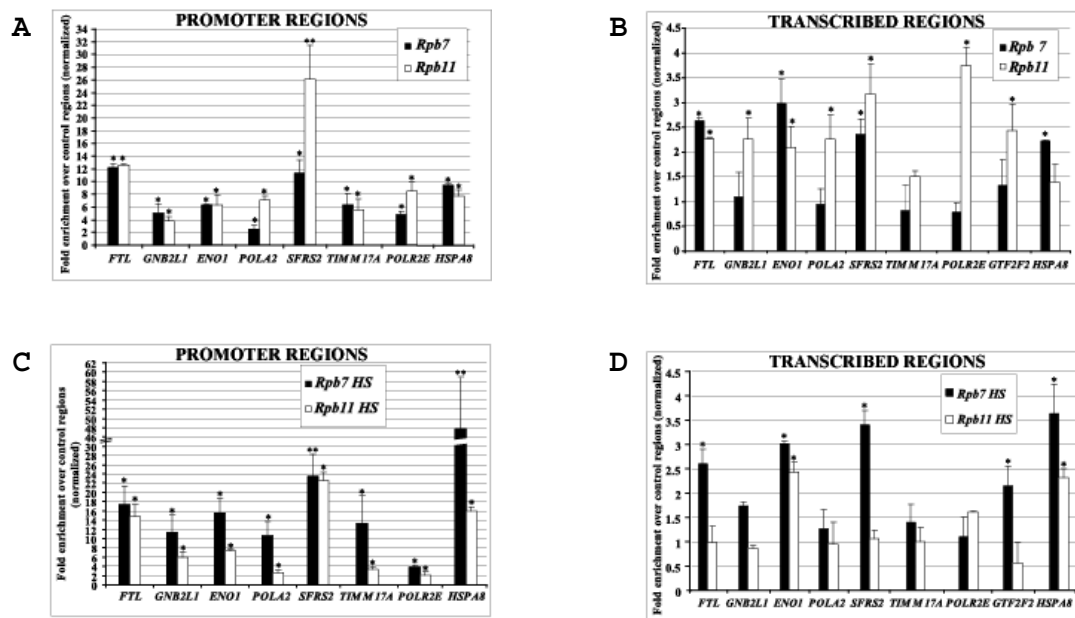


Figure 4

SUPPLEMENTARY DATA

Figure 1 Network of protein complexes involving the RNAPII general transcription machinery

Overview of a network showing the polypeptides found in association with the affinity tagged GTF and RNAPII subunits. The interactions are represented as directional edges extending from the tagged protein (red circles) using the web-based software VisANT (<http://visant.bu.edu>).

Figure 2 Genomic position of PCR products

The position of the PCR products along genomic DNA is represented on the March 2006 version of the UCSC Genome Browser database at <http://genome.ucsc.edu>.

Figure 3 Occupancy of (i) regions proximal to the transcriptional initiation site (TIS) and (ii) further downstream transcribed regions (TRs) by subunits of RNAPII and GTFs analysed by chromatin immunoprecipitation in order to confirm the results obtained with a first subunit

Color code: Black: >5-fold enrichment; Grey: significantly enriched with less than 5-fold enrichment values; White: not significantly enriched. Student's *t*-tests were performed to determine the significance of the enrichment values obtained in four experiments, including at least two independent ChIP assays.

Cojocaru et al., Supplementary Data

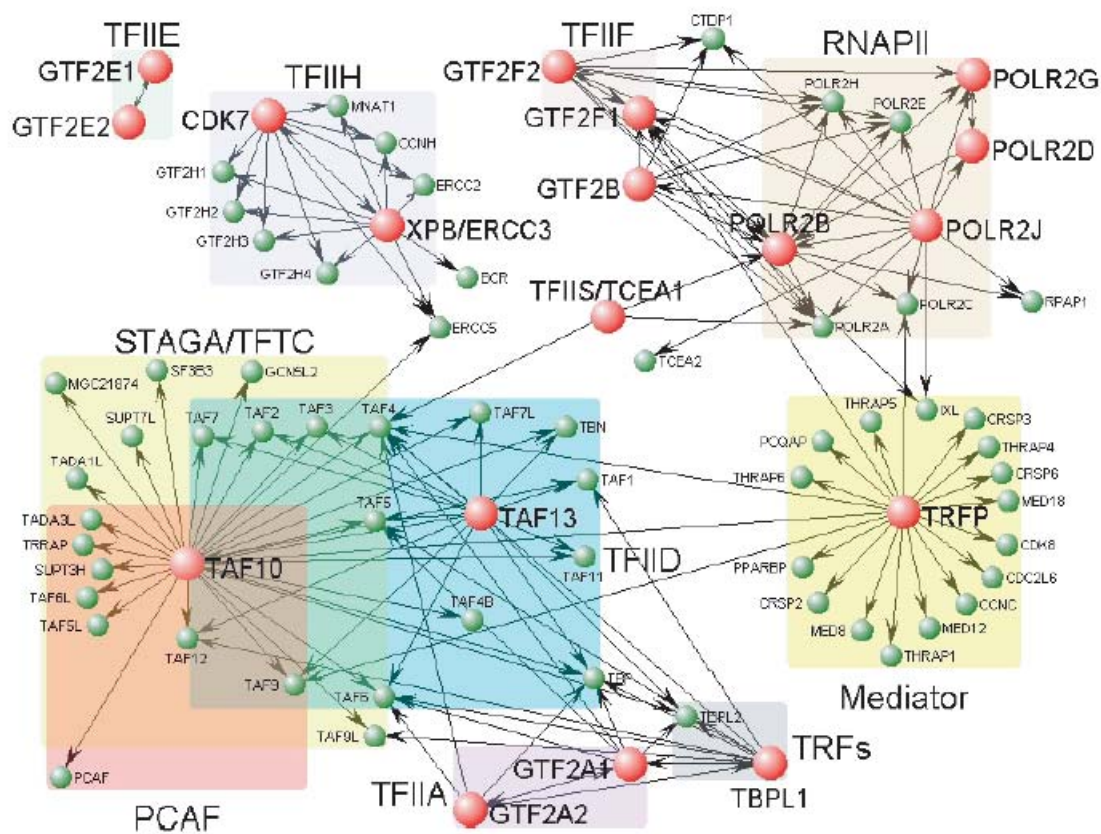


Figure 1

Cojocaru et al., Supplementary data

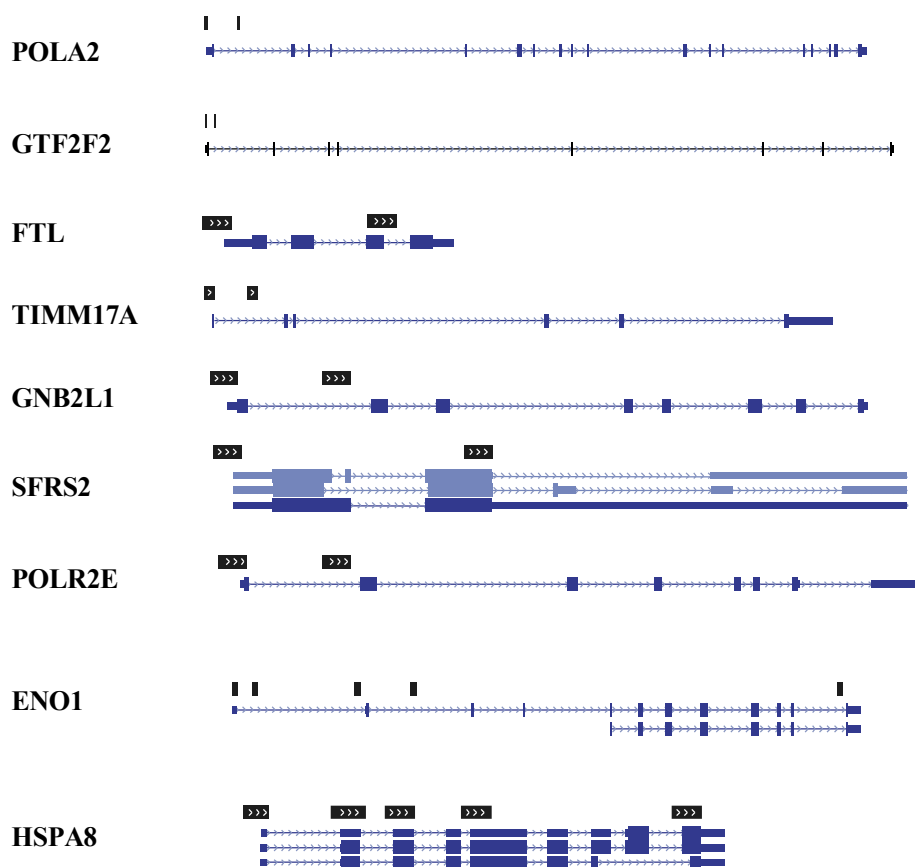
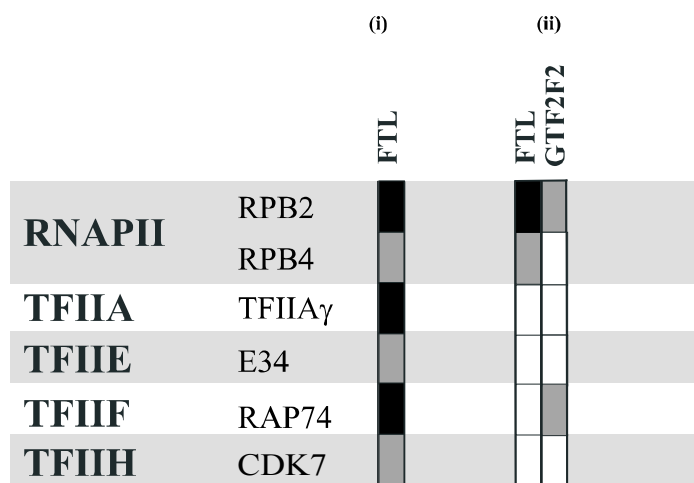


Figure 2

**Figure 3**

CHAPITRE 3

RPAPI, UNE NOUVELLE PROTÉINE ASSOCIÉE À L'ARN POL II PURIFIÉE À PARTIR DE SOUS-UNITÉS RECOMBINANTES DE TYPE SAUVAGE ET MUTANT

3.1 Préambule

Pour cet article, où je suis co-auteure, j'ai effectué les essais qui ont permis de caractériser le comportement des mutants de la sous-unité Rpb2 *in vivo*, par immunoprécipitation de la chromatine (Figure 4B et 8).

Marie-France Langelier a purifié la sous-unité Rpb2 sauvage ou mutante de l'ARN Pol II par la méthode TAP. Elle a également effectué le clonage, l'établissement des lignées stables ainsi que les essais biochimiques *in vitro* pour ces protéines recombinantes. (Figures 2B, 5B, 6 et 7)

Célia Jeronimo a purifié la sous-unité Rpb11 par la méthode TAP et a effectué les expériences présentées dans les Figures 1A, 2 (la Figure 2B est réalisée en collaboration avec Dania Baali et Marie-France Langelier) et 3B (en collaboration avec Dominique Bergeron).

Mahel Zeghouf a réalisé le clonage et l'établissement de la lignée stable Rpb11-TAP, dans le laboratoire du Dr Jack Greenblatt.

La Figure 3B a été réalisée par le groupe du Dr Timothy Hughes : Sanie Mnaimneh, Jeff Pootoolal et Armaity Davierwala. La Figure 4A a été réalisée par Diane Forget, 4C par Mark Chandy, étudiant chez le Dr Jerry Workman. Veronica Canadien; Bryan Beattie et Dawn Richards ont effectué la spectrométrie de masse pour la caractérisation du complexe présenté dans la Figure 2B.

3.2 Article

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

* Ces deux auteurs ont une contribution équivalente à cet article.

**RPAP1, a novel human RNA polymerase II-associated protein affinity purified
using recombinant wild type and mutated polymerase subunits**

Célia Jeronimo^{1#}, Marie-France Langelier^{1#}, Mahel Zeghouf², Marilena Cojocaru¹,
Dominique Bergeron¹, Dania Baali¹, Diane Forget¹, Sanie Mnaimneh², Armaity P.
Davierwala², Jeff Pootoolal², Mark Chandy^{3,4}, Veronica Canadien⁵, Bryan K.
Beattie⁵, Dawn P. Richards⁵, Jerry L. Workman³, Timothy R. Hughes², Jack
Greenblatt² and Benoit Coulombe^{1*}

¹ : Laboratory of Gene Transcription
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

³ : Stowers Institute for Medical Research
Kansas City, MO 64110, USA

⁴ : Pennsylvania State University College of Medicine
Hershey, PA 17033, USA

⁵ : Affinium Pharmaceuticals
100 University Avenue, 10th floor, Toronto, Ontario, M5J 1V6, Canada

These authors have contributed equally to this work

* Corresponding Author

Running Title:

RPAP1, a novel RNA polymerase II-associated protein

ABSTRACT

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 RNA Polymerase II-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 synthesis of Ydr527wp, a yeast protein homologous to RPAP1, and demonstrating that changes in global gene expression were similar to those caused by 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 their roles in the interaction of the enzyme with the template DNA during the transcription reaction.

INTRODUCTION

RNA polymerase (RNAP) II is the multisubunit enzyme that synthesizes all mRNA precursors in eukaryotes. RNAPII is highly conserved among species and, in humans, RNAPII consists of twelve subunits, namely Rpb1 to Rpb12 (Coulombe and Burton, 1999; Young, 1991). 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 (Awrey et al., 1997). In yeast, Rpb4 and Rpb7 form a subcomplex that can dissociate from the enzyme upon changes in environmental conditions. In active growth conditions, most yeast RNAPII molecules do not contain the Rpb4-Rpb7 dimer, which primarily associates during the stationary phase or following stress (Choder and Young, 1993). 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 (Cramer et al., 2000; Cramer et al., 2001; Fu et al., 1999; Gnatt et al., 2001) and bacterial (Campbell et al., 2002; Murakami et al., 2002a; Murakami et al., 2002b; Vassylyev et al., 2002; Zhang et al., 1999) RNAP 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 (Gnatt et al., 2001). 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 (Gnatt et al., 2001). 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 having 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 a number of proteins, including the general transcription factors (Conaway et al., 2000; Hampsey, 1998; Orphanides et al., 1996). 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 (Kim et al., 1993a; Kim et al., 1993b). TFIIB associates with the TBP-promoter complex and participates in RNAPII recruitment (Maldonado et al., 1990). TFIIF, composed of subunits RAP74 and RAP30, binds directly to RNAPII and also helps to recruit the enzyme to the promoter (Conaway et al., 1991; Sopta et al., 1985). TFIIF can induce further bending and wrapping of the promoter DNA around RNAPII during preinitiation complex formation (Robert et al., 1998). Both TFIIE and TFIIH were shown to participate in promoter melting at the transcription initiation site (Holstege et al.,

1995; Pan and Greenblatt, 1994). TFIIE, a factor also composed of two subunits, TFIIE56/ α and TFIIE34/ β , stimulates the helicase and kinase activities of TFIIH (Lu et al., 1992; Ohkuma et al., 1995; Ohkuma and Roeder, 1994). The helicase activity of TFIIH is required to fully open promoter DNA between nucleotides -9 and +2 (Holstege et al., 1996; Holstege et al., 1997; Jiang et al., 1994; Tirode et al., 1999) and its kinase activity phosphorylates the carboxy-terminal domain (CTD) of Rpb1 (Feaver et al., 1991; Lu et al., 1992; Serizawa et al., 1992; Serizawa et al., 1995; Shiekhattar et al., 1995). 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 (Ho and Shuman, 1999; Rodriguez et al., 2000). In yeast, the Ctk1 kinase phosphorylates the CTD on Ser2 at a later stage of the elongation phase (Cho et al., 2001), leading to the recruitment of mRNA processing factors (Skaar and Greenleaf, 2002). The CTD phosphatase, Fcp1, is able to dephosphorylate the Rpb1 CTD on Ser2 in order to recycle RNAPII for reinitiation (Cho et al., 2001; Kobor et al., 1999; Kobor et al., 2000).

Unlike all 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 on the purification of a functionally active human RNAPII using doubly tagged subunits expressed in mammalian cells. Using the TAP procedure (Rigaut et al., 1999), we isolated a human RNAPII-containing complex, which comprises the twelve 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 using TAP tagged Rpb2. The first mutant carries a deletion in fork loop 1, a structure expected to participate in maintaining the open state of the transcription bubble in elongating RNAPII (Cramer et al., 2001; Gnatt et al., 2001) or in RNA-DNA strand separation (Westover et al., 2004a). 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 (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; Peterson et al., 1991; Sumimoto et al., 1991), and highly purified TFIIH (Douziech et al., 2000; Gerard et al., 1991) and calf thymus RNAPII (Hodo, III and Blatti, 1977) were purified as previously described. Antibodies raised against Rpb1 (N-20) were obtained from Santa Cruz Biotechnology.

Tandem affinity purification of a human RNAPII-containing complex

The cDNAs encoding human Rpb2 (kindly provided by Marc Vigneron, IGBMC, Strasbourg, France) (Acker et al., 1992), Rpb4 (Invitrogen; accession number: BE883306), Rpb7 (Invitrogen; accession number: BE256210), Rpb11

(Invitrogen; accession number: AA814184), TFIIB (Ha et al., 1991) and RAP30 (Finkelstein et al., 1992) were cloned into the mammalian expression vector pMZI (No et al., 1996) in such a way that each polypeptide carries a TAP-tag at its C-terminus (Rigaut et al., 1999). Stable cell lines carrying these plasmids in EcR-293 cells (Invitrogen) were produced by transfection using the calcium phosphate method. Induction for 24 hours with 1-3 μM 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 (Rigaut et al., 1999). The TAP eluates were concentrated and stored at -80°C until ready to use.

Protein identification

The TAP eluates were run on 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 (Krogan et al., 2002). The resulting tryptic peptides were purified and identified by matrix-assisted laser desorption ionization-time of flight (MALDI-TOF) mass spectrometry (Krogan et al., 2002).

DNA microarray analysis of tet-promoter mutants

Tet-promoter alleles were constructed as previously described (Wu et al., 2002). Mutant and isogenic wild type cultures were grown in parallel in SC medium with 10 $\mu\text{g}/\text{ml}$ doxycycline for a total of 22 hours. Dilutions were made at 16 hours to ensure that mutant and wild type ODs matched as closely as possible to 10^7 cells/ml

at harvest. Cells were pelleted by 2 minute centrifugation in a room-temperature table-top centrifuge and frozen immediately in liquid nitrogen. Total RNA was prepared by hot acidic phenol:chloroform extraction followed by ethanol precipitation. Poly-A⁺ RNA purification and cDNA labeling were as described previously (Marton et al., 1998). 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 followed a procedure described by Hegde and collaborators (Hegde et al., 2000). Each array was normalized using grid-by-grid lowess smoothing (Yang and Speed, 2002). All measurements were taken in fluor-reversed pairs (i.e., each time a mutant was analyzed, it was hybridized to two arrays, combined by averaging the normalized log(ratio)).

Chromatin Immunoprecipitation

Chromatin immunoprecipitation assays were performed as previously described (Nissen and Yamamoto, 2000; Takahashi et al., 2000) with the following minor modifications. Approximately 2×10^8 cells were cross-linked using formaldehyde and chromatin fragments of an average size of 0.3-0.5 kbp were prepared using 4 sonication treatments (Fisher Sonic Dismembrator, 20 sec burst at 50% maximum power). For the immunoprecipitation step, the IgG-sepharose beads (Amersham Biosciences) were first blocked using a 5 mg/ml solution of BSA in PBS and then incubated with the sample for 1 hr. After reversal of the cross-links, the PCR analyses were performed using both the immunoprecipitated and input

DNA. Primer sets designed for amplifying both the promoter and the 5' untranslated region of the CCNA1 gene were used.

In the experiments of Figure 8, relative occupancy was evaluated by quantitative PCR (Q-PCR) as already described (Bryant and Ptashne, 2003) with minor modifications. The quantification of the target DNA in the immunoprecipitated sample was carried out by generating a standard curve with a 5 fold dilution series of the non-immunoprecipitated sample (input DNA) for each couple of primer sets and DNA. 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 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 GNB2L1 (-200 to +55 and +998 to +1165 respectively) and FTL (-157 to +33 and +955 to +1176 respectively) genes, and also for an internal sequence of the EST located at chr17:49926700-49927999 as a control. Oligonucleotide sequences are available on request. All amplified DNA fragments are unique as determined by BLAST on the April 2003 assembly version of the UCSC Genome Browser (<http://www.genome.ucsc.edu>), by performing a melting curve analysis and by agarose gel electrophoresis (data not shown).

Gel Mobility Shift Assay

Gel mobility shift assays were performed as we previously described (Langelier et al., 2001). Complexes were assembled 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 (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 magnetic beads at 25°C for 30 min. The beads were then incubated with 1 nM of RNAPII complex (Rpb11 TAP purified or calf thymus) for 1 h at 30°C. Using a magnetic particle concentrator (Dynal, MPC-S), the beads were pelleted and the supernatant was removed. After washing twice, the beads were resuspended in SDS-PAGE lysis buffer and subjected to SDS PAGE analysis. The complex was detected by immunoblot with an anti-Rpb1 antibody (Santa Cruz, N-20).

Transcription Assay

Transcription reactions were performed as described previously (Burton et al., 1986). 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 (amounts of RNAPII used in 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 nucleotides -50 to +10 fused to a G-less cassette. Under these conditions a 391-nt run-off transcript is produced.

Abortive Initiation Assay

Abortive initiation assays were performed as we described previously (Langelier et al., 2001) with the following modifications. The DNA template (12 ng), either premelted in the -9 to +2 region or fully double-stranded, were 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 (amounts of RNAPII used in the various experiments are detailed in the figure legends) in a 20 ul reaction mixture containing 750 μ M ATP, 750 μ M CTP, 10 μ M UTP, 2,5 μ Ci [α -³²P] UTP, 12.5 mM MgCl₂, 3 mM EGTA and 0.84 units/ml RNase inhibitor. Transcripts were analyzed on a 23% polyacrylamide denaturing gel containing 7 M urea.

Elongation assay

The template was prepared by annealing two complementary oligonucleotides of random sequence, one carrying an extension of 15 CMP as described previously (Pan and Greenblatt, 1994). 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 fork1 (Δ 458-459)-TAP and the reactions were performed as described (Pan and Greenblatt, 1994). Transcripts were analyzed on 18% polyacrylamide, 9% urea denaturing gels.

RESULTS

Tandem affinity purification of a human RNAPII-containing complex

We have adapted the tandem affinity purification (TAP) procedure (Puig et al., 2001; Rigaut et al., 1999), originally developed to isolate protein complexes from yeast, for the purification of human RNAPII and general transcription factors. Native conditions are maintained during the TAP procedure through the use of buffers having a low salt concentration (0.1 M Na⁺) and a neutral pH. As shown in Figure 1A, which summarizes the TAP procedure, full length cDNAs encoding polypeptides of interest were cloned in an ecdysone-inducible mammalian expression vector (No et al., 1996). The vector was engineered to encode polypeptides carrying, at their C-terminus, a TAP tag consisting of two copies of the immunoglobulin G (IgG)-binding domain of *Staphylococcus aureus* protein A, a cleavage site for the tobacco etch virus (TEV) protease, and a calmodulin-binding peptide (CBP) (Rigaut et al., 1999). The tagged polypeptides were expressed in EcR-293 cells at near physiological levels following a 24 hour induction using the ecdysone analog, ponasterone A. The concentration of ponasterone A required for obtaining 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 (see Fig. 1B).

Whole cell extracts were prepared from the ponasterone A-induced cell lines and used in the double affinity chromatography procedure, which consisted of successive steps on IgG-sepharose and calmodulin beads (Rigaut et al., 1999). The eluates were frozen and analyzed using 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 submitted to SDS-PAGE and the bands were identified using 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 non-specifically to our affinity columns (data not shown). The complex contained the twelve subunits of RNAPII, TFIIB, both subunits of TFIIF (RAP30 and RAP74), Fcp1, the TFIIF-associating CTD phosphatase, and a novel human polypeptide that we named RPAP1 (accession number NP_056355). The name was registered in the HUGO database.

To validate the existence of a unique 17-subunit complex containing RNAPII, other subunits of the complex, including Rpb2, Rpb4, Rpb7, TFIIB and RAP30, were tagged, expressed in stably-transfected EcR-293 cells and used as baits in the TAP procedure. The eluates were analyzed using silver-stained SDS 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 using TAP-tagged RAP30 contained both a tagged RAP30 with its CBP residual domain and non-tagged RAP30 (Fig. 2C). The presence of RAP30 lacking the CBP domain was confirmed by Western blotting (data not shown). These results indicated 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 using 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 RNA polymerase II

RPAP1 is a previously identified human gene product (DKFZP727M111 protein) of unknown function. BLAST sequence searches against publicly available databases revealed highly significant alignments over the full length of the RPAP1 sequence and proteins of unknown function from *Mus musculus* (BAC65787), *Rattus norvegicus* (XP_230480) and, to a lesser extent, *Drosophila melanogaster* (NP_648573, CG32104-PB). In addition, these proteins show significant similarities to the *Saccharomyces cerevisiae* Ydr527wp protein (NP_101816), a protein required for cell viability that was shown to bind the Rpb10 subunit of RNAPII using a two-hybrid screen (Giaever et al., 2002; Ito et al., 2001). The positions of the different alignments are shown in Figure 3A. Notably, no particular motif or domain could be predicted with high confidence on the RPAP1 sequence. However, and in view of the participation of RPAP1 to the RNAPII complex, the IMPALA

analysis against the BLOCK database (Schaffer et al., 1999) revealed a very weak similarity ($E=0.23$) between the bromodomain-containing protein 4 (Brd4) and the RPAP1 region spanning amino acid position 196 to 312 (see next section). We have also performed a sequence-structure comparison analysis using the structure-guided alignment program FUGUE against the HOMSTRAD database (Mizuguchi et al., 1998). The result suggests that RPAP1 belongs to the ARM repeat protein superfamily (Coates, 2003). Indeed, the RPAP1 region spanning amino acid 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 SGD and FlyBase databases respectively, to belong to the ARM repeat superfamily (SGD: <http://genome-www4.stanford.edu/cgi-bin/SGD/locus.pl?locus=YDR527W>, FlyBase: <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 hours, and compared gene expression in this mutant to expression patterns from 121 other tet-promoter mutants affecting genes implicated in diverse biological processes (Mnaimneh et al., 2004). Among all of these, the effect of the *tet-YDR527w* allele correlated most strongly with that of *tet-RPB11* (Fig. 3B). 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 consequence of an overall perturbation in transcription rather than any individual pathway.

To further define the interaction between RPAP1 and RNAPII, we have cloned the human RPAP1 cDNA and derived a human cell line that can be induced to express TAP-tagged RPAP1. Analysis of the TAP-tagged RPAP1 eluate revealed the presence of two RNAPII subunits, Rpb2 and Rpb3 (Fig. 3A, right part). 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 (Hazbun et al., 2003). Because Ydr527wp and RPAP1 are both physically and functionally related, the yeast 50 kDa 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 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 gel mobility shift (EMSA) experiment in Figure 4A shows that the human RNAPII complex can form a preinitiation complex in the presence of TBP, TFIIB, TFIIF and TFIIIE on a radiolabeled probe carrying the adenovirus major late (AdML) promoter. As compared to an equivalent preinitiation complex assembled with highly purified calf thymus RNAPII, the TAP-tagged RNAPII complex formed a complex of

slightly lower mobility. Because our calf thymus RNAPII contains neither Fcp1 nor RPAP1 (data not shown), our results indicate that the preinitiation complex assembled using the TAP-tagged human RNAPII complex includes at least one of these two polypeptides (Fcp1 and/or RPAP1).

Chromatin immunoprecipitation (ChIP) experiments were used to determine whether the tagged RNAPII complex can also be recruited to promoters *in vivo*. As shown in Figure 4B, DNA fragments containing the CCNA1 (Cyclin A) promoter (-50 to +50) were enriched more efficiently than DNA fragments containing an upstream region (-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 the bromodomain-containing protein 4 (Brd4). Because Brd4 was recently shown to specifically interact with histones H3 and H4 through acetylated lysine residues (Dey et al., 2003), we tested the ability of the complex to bind to acetylated histones. As shown in Figure 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. The TAP-tagged RNAPII bound histone tails acetylated on lysine 14 more efficiently

than on 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 its ability to initiate transcription *in vitro* using a transcription system reconstituted with the purified general initiation factors. Purified calf thymus RNAPII or purified TAP-tagged human RNAPII were incubated with a DNA template containing the AdML promoter in the presence of purified TBP, TFIIB, TFIIF, TFIIE 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” (see Fig. 6A) is highly conserved from yeast to human, absent in bacterial RNAP, and proposed to be involved in maintaining the transcription bubble (Cramer et al., 2000; Gnatt et al., 2001) or in RNA-DNA strand separation (Westover et al., 2004a), 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 became misfolded. A smaller deletion of two amino acids located near the center of the loop (Δ 458-459), which correspond 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 Δ 458-459. The silver stained SDS gel in Figure 6B shows that the relative abundance of some components of the tagged RNAPII complex varied between wild type and 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 is at the very most 5-fold less abundant in the mutant as compared to the wild type complex.

In order to elucidate the function of fork loop 1 in transcription, we compared Rpb2 fork1 Δ 458-459 with wild type RNAPII in a number of different

assays. The assays shown in Figure 6 (C-E) used 5-fold amounts of the mutant complex as compared to wild type to ensure that the amount of the mutant complex was not underestimated (see above). As shown in Figure 6C, Rpb2 fork1 Δ 458-459 was defective in run-off transcription assays when we used the AdML promoter in the presence of all the general initiation factors.

In order to delineate with more precision the molecular defect that impairs the transcriptional activity of Rpb2 fork1 Δ 458-459 *in vitro*, we tested its activity using a promoter-independent assay. As described previously (Gnatt et al., 1997; Sluder et al., 1988), RNAPII can transcribe in the absence of the general initiation factors from a template carrying a 3' C extension (C-tailed template). The template used in Figure 6D carries a 15 C extension on the 3' strand. The transcribed strand is 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 nucleotides. In the presence of CTP the major transcript was 60 nucleotides in length, indicating that the polymerase starts transcription 10 nucleotides upstream of the double-stranded junction. In contrast to RNAPII carrying a wild type Rpb2, Rpb2 fork1 Δ 458-459 was inactive in the C-tailed template elongation assay.

Figure 6E shows that Rpb2 fork1 Δ 458-459 is also defective in an abortive initiation assay carried out using 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 strategically located at + 11 in the template blocks transcription at +10 (see 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 is defective in promoter melting, a pre-melted template would counteract the defect in the mutant. The premelted template (-9/+2), which carries an unpaired DNA region from -9 to +2 that mimics the fully melted promoter, supported the synthesis of abortive transcripts, although with poor efficiency as compared to wild type. This finding indicated that a mutation in fork loop 1 does not simply affect promoter melting and its maintenance, but rather the interaction of RNAPII with the melted promoter region.

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

A 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 (Gnatt et al., 2001). Switch 3 makes direct contact with template DNA near -5 (arginine 1122) and -2 (arginine 1129) (see Fig. 6A). Since R1129 is juxtaposed to 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, R1124 are replaced with alanines. 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 wild type in forming the 17-subunit RNAPII complex. However, EMSA experiments (Fig. 7B) indicated that purified Rpb2 sw3-1078 is not efficient in 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 as compared to wild type was necessary in order to achieve efficient formation of a pre-initiation 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 as compared to wild type led to the production of run-off transcripts, but the transcription efficiency remained slightly lower than wild type, indicating that the mutation in switch 3 does not only affect the ability of sw3-1078 to assemble into a pre-initiation complex but also, to a lesser extent, its capacity to support transcription.

Mutants Rpb2 fork1 Δ 458-459 and Rpb2 sw3-1078 are defective *in vivo*

To assess the defect 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 gene (GNB2L1) and

the ferritin light polypeptide gene (FTL). As shown in Figure 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 is not found in association with the transcription units. In contrast, sw3-1078 is neither enriched in the promoter regions nor in 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 is enriched at neither the promoter nor the transcription unit of either the GNB2L1 or the FTL genes. These *in vivo* results fully support our *in vitro* studies indicating that fork loop 1 is essential for transcription but is not involved in pre-initiation complex assembly, whereas switch 3 is required for efficient pre-initiation complex formation.

DISCUSSION

Simple, efficient purification of human RNAPII in 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 (Rigaut et al., 1999) have developed the TAP procedure for allowing the purification of protein complexes in native conditions. Two affinity chromatography steps are used. Elution from the

IgG column is performed through a proteolytic treatment with the TEV protease while the elution from the calmodulin column uses EGTA as a chelating agent. In this paper we have 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. Expression levels were modulated by adjusting the concentration of the inducer, ponasterone A (see Fig. 1B). By preventing the overexpression of the tagged polypeptides, the formation of non-specific interactions was limited.

We report here on the use of the TAP procedure in human cells for the easy and efficient purification of a complex containing the twelve 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 6 of its subunits, demonstrating that all the components are part of a unique complex. Because *(i)* the TAP-tagged polypeptides are expressed near physiological levels in stably transfected human cells, *(ii)* protein extraction is performed in gentle conditions and *(iii)* the affinity purification steps are achieved in 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 (Rachez and Freedman, 2001) 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-terminus 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 form a preinitiation complex with the general initiation factors on the AdML promoter, accurately initiate transcription and elongate the transcripts (see Figs. 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.

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 it was previously proposed (Kimura et al., 2002). Our EMSA experiments suggest 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 colleagues that the CTD phosphatase co-localizes with RNAPII at active promoters in yeast (Cho et al., 2001). 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 (Conaway et al., 1991; Maldonado et al., 1990; Sopta et al., 1985). TFIIB binds to TBP and forms a TBP-TFIIB-promoter complex (Bagby et al., 1995; Nikolov et al., 1995). TFIIB binds directly to a DNA motif named the TFIIB-recognition element (BRE), a core element found in a number of RNAPII promoters (Lagrange et al., 1998). TFIIF binds to both RNAPII and TFIIB and is also involved in RNAPII recruitment to core promoters (Conaway et al., 1991; Maldonado et al., 1990; Sopta et al., 1985). In addition, TFIIF was shown to regulate the contacts of RNAPII with promoter DNA before transcription initiation (Douziech et al., 2000; Forget et al., 1997; Robert et al., 1998). TFIIF plays a central role in promoter bending and wrapping against the mobile clamp of RNAPII, a process that is essential to transcription initiation (Forget et al., 2004b). 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 (Robert et al., 1998). Our results showing that two molecules of RAP30 are present within the tagged RNAPII complex support the notion that TFIIF works as a $\alpha_2\beta_2$ heterotetramer *in vivo* (Conaway and Conaway, 1989; Flores et al., 1990).

Notably, both TFIIF and TFIIB can also bind to Fcp1 in a competitive manner and with opposite effects, TFIIF stimulating and TFIIB inhibiting the phosphatase activity (Chambers et al., 1995). Fcp1 may not efficiently

dephosphorylate RNAPII on Ser2 of the CTD in the region downstream of the promoter (Cho et al., 2001) because TFIIF appears not to be present in this complex (Krogan et al., 2002; Pokholok et al., 2002b). 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 (Giaever et al., 2002; Ito et al., 2001) and it also affects 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, whereas its exact role remains to be elucidated.

Structure-guided alignments indicated that RPAP1 could belong to the ARM repeat protein superfamily (Coates, 2003). Armadillo (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 (Huber et al., 1997). 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, the cytoplasmic β -catenin is stabilized. Increased level of β -catenin leads to its nuclear translocation where it affects transcriptional activation. Although reported as a cytoplasmic protein (Huh et al., 2003), Ydr527wp/RBA50 is found in complex with three RNAPII subunits and appears to affect transcription. Similarly to β -catenin, Ydr527wp/RBA50 may have a dual cellular location and be translocated to the nucleus.

Proteins previously shown to interact with RNAPII, such as the mediator and various elongation factors were not found as 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, mainly lost in our 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 have expressed TAP-tagged Rpb2 carrying site-directed mutations and purified the mutant forms of RNAPII. Our data indicate that the assembly and/or the stability of the RNAPII complex are 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 mutation that removes only 2 amino acids in Rpb2 was found to assemble correctly in solution *in vivo* (see 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 pre-initiation complex with the general initiation factors with efficiency similar to that of the wild type complex, but was incapable of transcription initiation and elongation *in vitro*. When we used a premelted template in which the transcription bubble was fully open between -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 might 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 2 a.a. shorter fork loop 1 in the RNAPII mutant may be unable to contact one of the DNA strands in the transcription bubble, leading to an incorrect positioning of the DNA template in the active site and impairing the ability of the enzyme to initiate transcription efficiently. Since this mutant supports 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) and the RNA around positions -5 to -7 (Westover et al., 2004a), an alternative possibility to explain the defect of Rpb2 fork 1 Δ 458-459 would be 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 to RNA-DNA strand separation.

The 5 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 (Cramer et al., 2001; Gnatt et al., 2001). This hypothesis was based on the fact that most of the switches are disordered in the crystal structure of the 10-subunit RNAPII 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 the 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. As 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 (Armache et al., 2003; Bushnell and Kornberg, 2003b) 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 on the clamp far from the active site in a way similar to the bacterial RNA polymerase holoenzyme-promoter DNA complex (Murakami et al., 2002a), and 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 assembly of RNAPII into a pre-initiation complex on promoter DNA. Mutating amino acids 1078 to 1080 clearly reduced the ability of human RNAPII to form a pre-initiation complex in the presence of the general transcription factors both *in vitro* and *in vivo*. *In vitro*, 6-fold higher 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 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 6-fold higher amounts of the sw3-1078 mutant did not yield as many transcripts as wild type. This may also

suggest that the pre-initiation 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.

ACKNOWLEDGEMENTS

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 is 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. from 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.

Reference List

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.U.S.A.*

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-A resolution: Implications for the initiation of transcription. *Proc.Natl.Acad.Sci.U.S.A.*
8. 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.
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.U.S.A* 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 angstrom 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.U.S.A* 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 IIIH 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.U.S.A* 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. Kotter, 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. Wilhelmy, E. A. Winzeler, 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. Snynsman, 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* 29:548-4, 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. 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.
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.U.S.A* 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. Eubright. 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 IIIH 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. Mnaimneh, S., A. P. Davierwala, J. Haynes, J. Moffat, W. T. Peng, W. Zhang, X. Yang, J. Pootoolal, G. Chua, A. Lopez, M. Trocheset, ., 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. 2004. Compendium II: Exploration of essential gene functions via titratable promoter alleles. - Submitted.
 58. 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.
 59. Murakami, K. S., S. Masuda, and S. A. Darst. 2002. Structural basis of transcription initiation: RNA polymerase holoenzyme at 4 a resolution. *Science* 296:1280-1284.
 60. 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.
 61. 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.
 62. No, D., T. P. Yao, and R. M. Evans. 1996. Ecdysone-inducible gene expression in mammalian cells and transgenic mice. *Proc.Natl.Acad.Sci.U.S.A* 93:3346-3351.
 63. Ohkuma, Y., S. Hashimoto, C. K. Wang, M. Horikoshi, and R. G. Roeder. 1995. Analysis of the role of TFIIE in basal transcription and TFIIH-mediated carboxy-terminal domain phosphorylation through structure-function studies of TFIIE-alpha. *Mol.Cell Biol.* 15:4856-4866.
 64. Ohkuma, Y. and R. G. Roeder. 1994. Regulation of TFIIH ATPase and kinase activities by TFIIE during active initiation complex formation. *Nature* 368:160-163.
 65. 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.

66. Orphanides, G., T. Lagrange, and D. Reinberg. 1996. The general transcription factors of RNA polymerase II. *Genes Dev.* 10:2657-2683.
67. 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.
68. 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.
69. 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.
70. Puig, O., F. Caspary, 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.
71. Rachez, C. and L. P. Freedman. 2001. Mediator complexes and transcription. *Curr.Opin.Cell Biol.* 13:274-280.
72. 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.
73. 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.
74. Rodriguez, C. R., E. J. Cho, M. C. Keogh, C. L. Moore, A. L. Greenleaf, and S. Buratowski. 2000. Kin28, the TFIIH-associated carboxy-terminal domain kinase, facilitates the recruitment of mRNA processing machinery to RNA polymerase II. *Mol.Cell Biol.* 20:104-112.
75. 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.
76. 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.U.S.A* 89:7476-7480.
77. 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 TFIIH. *Nature* 374:280-282.

78. 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 TFIID. *Nature* 374:283-287.
79. 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.
80. 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.
81. 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.
82. 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.
83. 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.
84. Tirode, F., D. Busso, F. Coin, and J. M. Egly. 1999. Reconstitution of the transcription factor TFIID: assignment of functions for the three enzymatic subunits, XPB, XPD, and cdk7. *Mol.Cell* 3:87-95.
85. Vassylyev, D. G., S. Sekine, O. Laptenko, J. Lee, M. N. Vassylyeva, S. Borukhov, and S. Yokoyama. 2002. Crystal structure of a bacterial RNA polymerase holoenzyme at 2.6 Å resolution. *Nature* 417:712-719.
86. 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.
87. Wu, L. F., T. R. Hughes, A. P. Davierwala, M. D. Robinson, R. Stoughton, and S. J. Altschuler. 2002. Large-scale prediction of *Saccharomyces cerevisiae* gene function using overlapping transcriptional clusters. *Nat.Genet.* 31:255-265.
88. Yang, Y. H. and T. Speed. 2002. Design issues for cDNA microarray experiments. *Nat.Rev.Genet.* 3:579-588.
89. Young, R. A. 1991. RNA polymerase II. *Annu.Rev.Biochem.* 60:689-715.

90. 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.

FIGURE LEGENDS

Figure 1. Tandem affinity purification (TAP) of human transcription factors. (A) Overview of the TAP procedure. (B) Data from a pilot experiment comparing the expression level 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 hours was compared to endogenous TFIIB using Western blotting with an antibody raised against TFIIB.

Figure 2. Purification of a TAP-tagged human RNAPII complex. (A) Whole cell extracts prepared from induced or non-induced EcR-293 cells programmed to express TAP-tagged Rpb11 (Rpb11-TAP) were purified using the TAP procedure and the eluates 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 position 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 the purification of the same 17-polypeptide RNAPII complex. The components of the complex were identified according to their molecular weight in SDS gels stained with silver (SS), peptide mass fingerprinting using MALDI-TOF analysis (MS) and immunoreactivity in Western blots (WB).

(C) An SDS gel showing that both tagged (RAP30-CBP) and non-tagged RAP30 are present in the RNAPII complex purified using TAP-tagged RAP30. Only non-tagged RAP30 is found in the complex purified using TAP-tagged Rpb11. (D) The TAP-tagged RNAPII complex contains a hypophosphorylated CTD. The N-20 antibody that 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.

Figure 3. (A) Regions of homology between *H. sapiens* RPAP1, *D. melanogaster* CG32104-PB and *S. cerevisiae* Ydr527wp/RBA50 as determined using BLAST analyses. Boxes represent regions with significant alignments. Black boxes: E values ranging from $8e-37$ to $1e-10$. Open boxes: E values ranging from $4e-05$ to 0.1. The *M. musculus* and *R. norvegicus* orthologues, harboring 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 (Hazbun et al., 2003) with transcription factors are indicated on the right part of the figure. (B) Correlations among mRNA abundance profiles in tet-promoter mutants of *RPB11*, *YDR527w*, and *RRN3*, measured using oligonucleotide microarrays. Each mutant was compared to an isogenic control without the tetO7 promoter; log(ratios) are plotted. Tet-*RPB11* and tet-*YDR527w* were assayed twice, with the replicate cultures grown, extracted, and assayed on different dates. Correlation of tet-*RPB11*

and tet-*YDR527* to a tet-promoter mutant of *RRN3*, a gene involved in Pol I transcription, is shown for contrast.

Figure 4. TAP-tagged human RNAPII assembles on promoter DNA both *in vitro* and *in vivo* and binds to acetylated histones. (A) An EMSA experiment was performed using 150 ng of both 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 is included. A radiolabeled DNA fragment containing the AdML promoter is used as a probe. (B) Chromatin immunoprecipitation 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 (-1000 to -900) of the transcription start site were used on DNA fragments enriched using IgG beads (anti-TAP). The input lanes correspond to DNA that was not submitted to immunoprecipitation. A PCR control that lacks DNA (no DNA) was included in each case. (C) A peptide binding experiment shows that TAP tagged RNAPII is retained on acetylated histones. The results of three independent experiments were analysed using Image Quant corrected for background and shown relative to unmodified H3. Conventionally purified RNAPII was used as a negative control. The TAP-tagged RNAPII bound histone tails acetylated on lysine 14 more efficiently than on lysine 9.

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

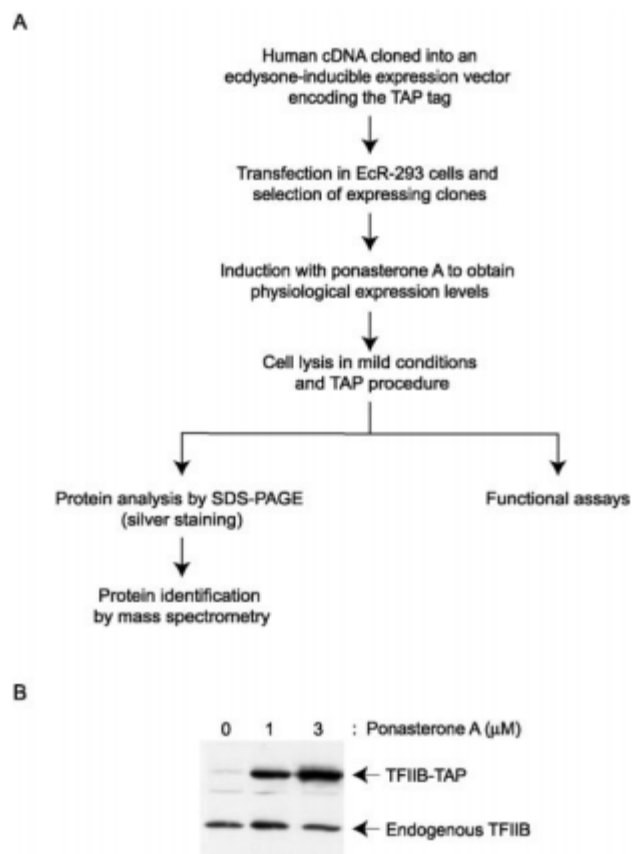
Figure 6. Purification and functional analysis of a TAP-tagged human RNAPII with a mutation in fork loop 1. (A) Top panel. A model showing elongating yeast RNAPII (Gnatt et al., 2001). The template DNA strand (blue) from -3 to +10 and the 9-nt RNA (pink) were placed according to the crystal structure. The remaining of the template strand and the coding strand (green) were modeled in Cinema 4D and their exact position is speculative. Bottom panel. Simplified view of the yeast RNAPII catalytic center. Different domains of Rpb1 and Rpb2 located near the DNA-RNA hybrid are shown. 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 (a.a. 250-258) and the rudder (a.a. 315-320) that are absent in the PDB files corresponding to elongating RNAPII were reconstructed based on the crystal structure of free yeast RNAPII (Cramer et al., 2001). Missing parts of fork loop 1 (a.a. 468-476) and fork loop 2 (a.a. 503-508) were modeled in Cinema 4D and their exact position is speculative. (B) Silver-stained SDS gel of wild type TAP-tagged RNAPII (Rpb2 wt)

and a mutant of human RNAPII carrying a two amino acids deletion in the Rpb2 fork loop 1 (Rpb2 fork1 Δ 458-459) domain. The position of some RNAPII subunits and molecular weight markers are indicated. (C) *In vitro* transcription reactions (run-off) contained either 24 ng of wild type RNAPII (Rpb2 wt) or 120 ng of Rpb2 fork1 Δ 458-459 in the presence of TBP, TFIIB, TFIIF, TFIIE and TFIIH. (D) Elongation assays were performed using a C-tailed template carrying a 15-C extension in the absence of the 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 of CTP (-CTP; 35 nt) are indicated. The reactions contained 88 ng of Rpb2 wt and 440 ng of Rpb2 fork1 Δ 458-459. (E) Abortive initiation assays were performed with 100 ng of Rpb2 wt and 500 ng of Rpb2 fork1 Δ 458-459 in the presence of TBP, TFIIB, TFIIF and TFIIE on closed (0/0) or pre-melted (-9/+2) templates in the absence of GTP. Templates are schematized and the 3 to 10 nucleotide abortive transcripts are indicated. (F) EMSA experiments were performed on AdML promoter in the presence of TBP, TFIIB, TFIIF and TFIIE. 200 ng of Rpb2 wt and 500 ng of Rpb2 fork 1 Δ 58-459 were used in the reactions.

Figure 7. Purification and functional analysis of a 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) EMSA experiments were performed on AdML promoter with TFIIB, TFIIF and TFIIE in either the

presence or the absence of TBP. Amounts used for the wild type and the mutant are indicated. (C) *In vitro* transcription reactions (run-off) contained different amounts of Rpb2 wt and Rpb2 sw3-1078 in the presence of TBP, TFIIB, TFIIF, TFIIE and TFIID.

Figure 8. Association of mutant forms of the RNAPII complex with chromatin *in vivo*. Fold enrichment over the control region of both the promoter and the transcribed region of FTL and GNB2L1 genes following 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 mean of four experiments, including two independent chromatin immunoprecipitation assays.

**Figure 1**

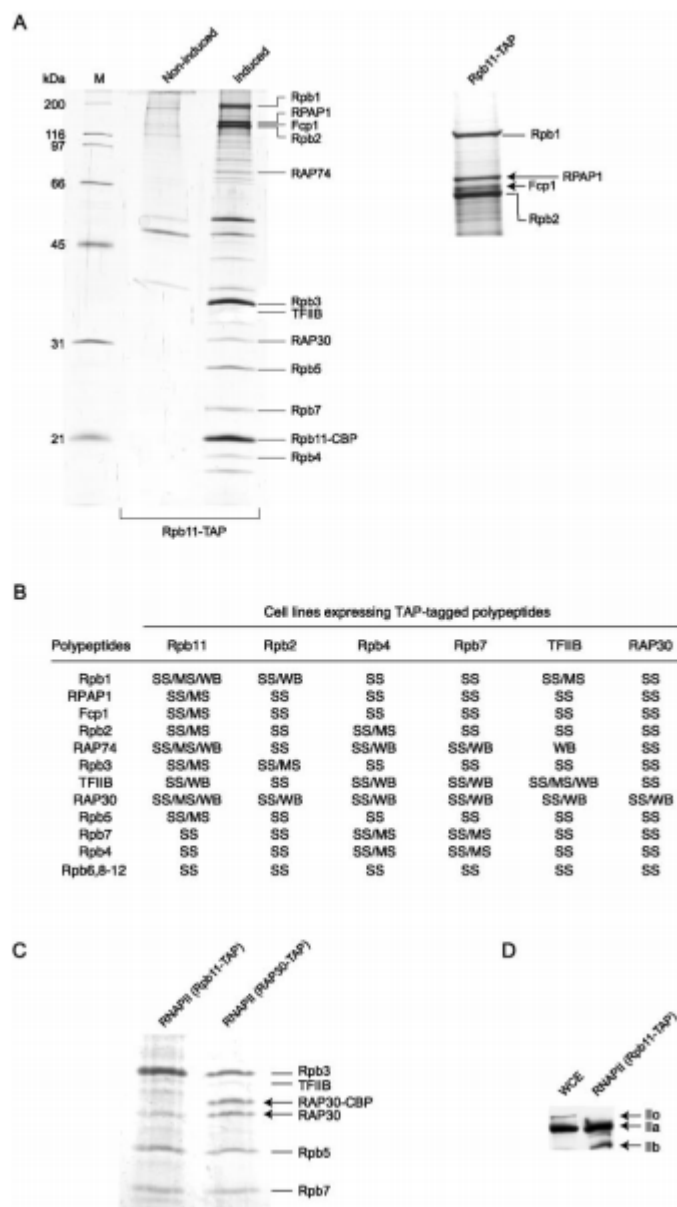


Figure 2

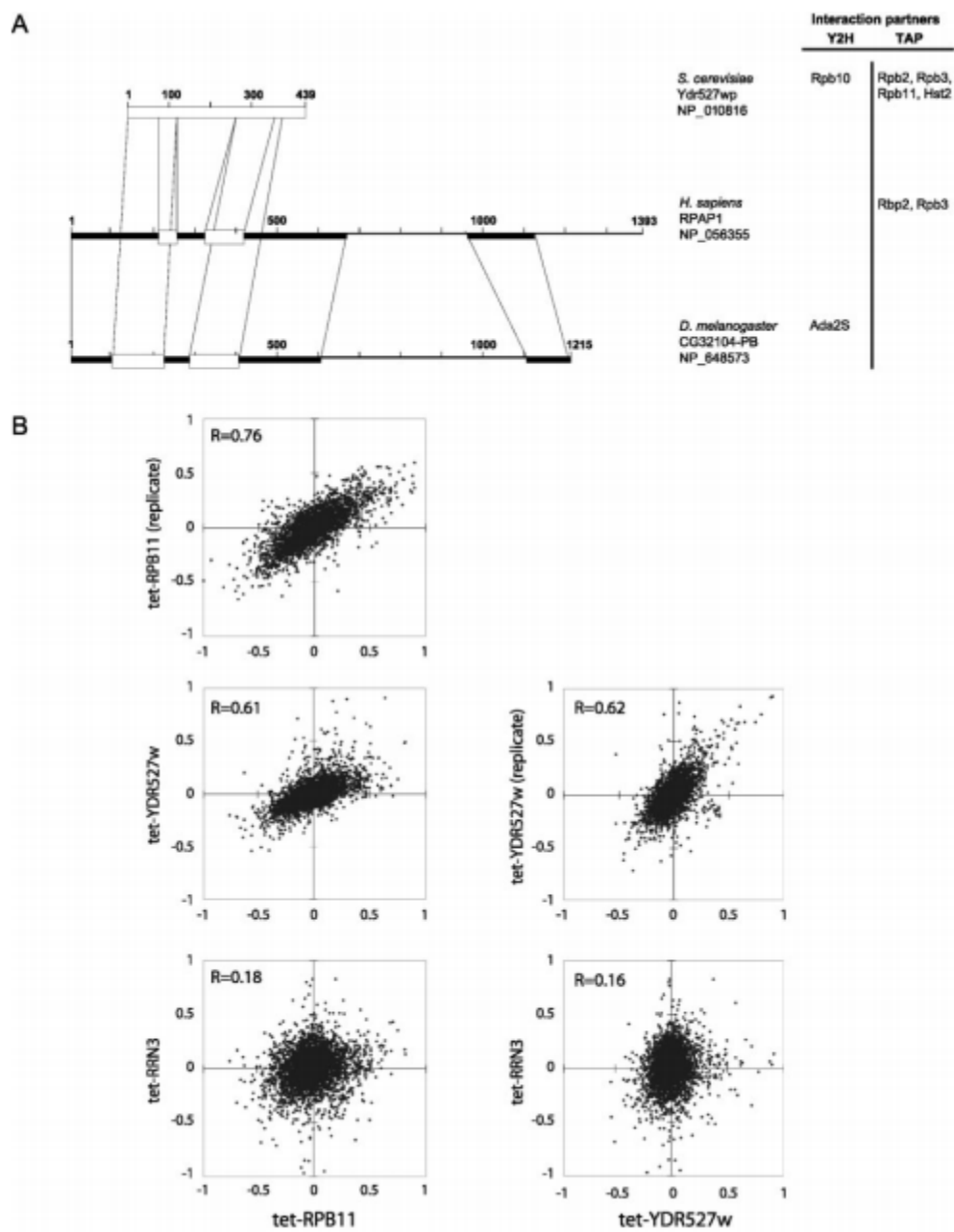


Figure 3

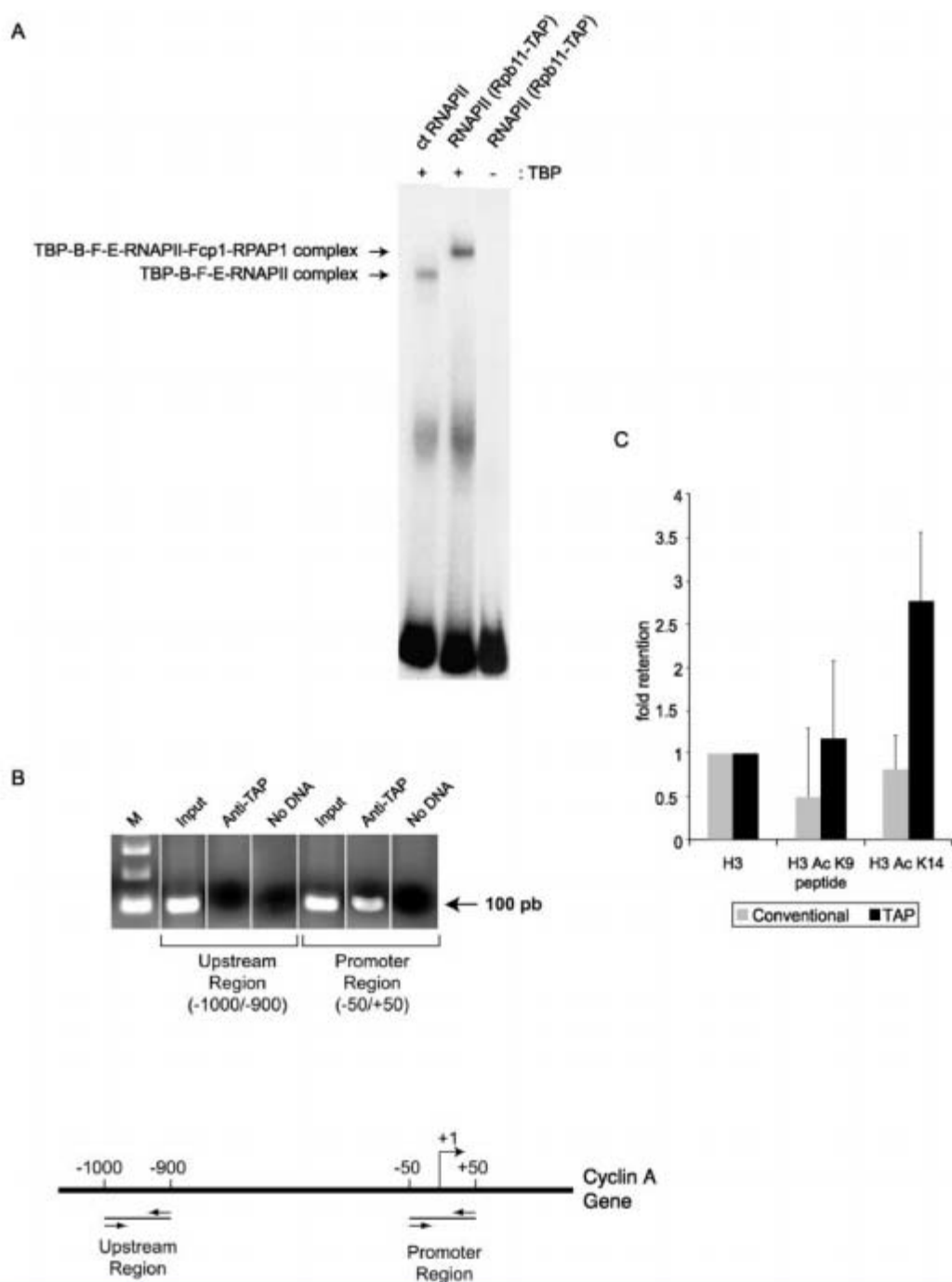
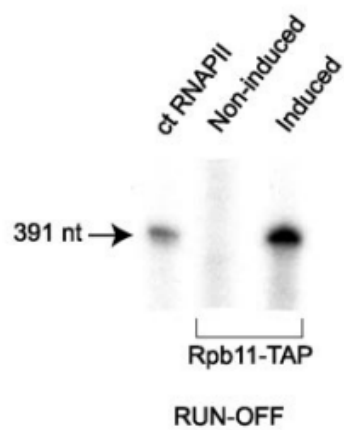


Figure 4

A



B

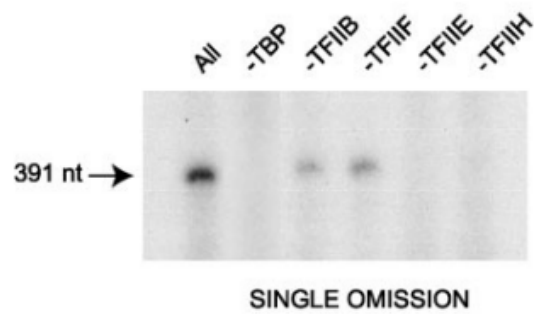
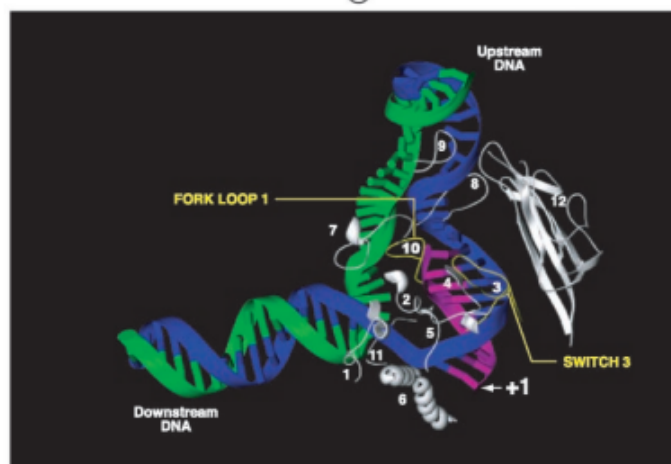
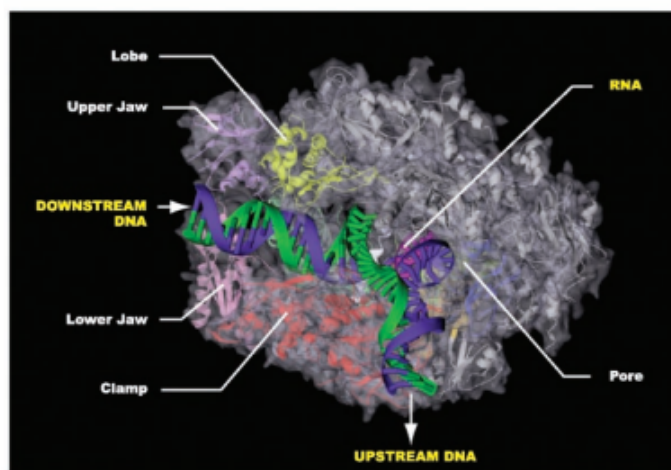


Figure 5

A



B

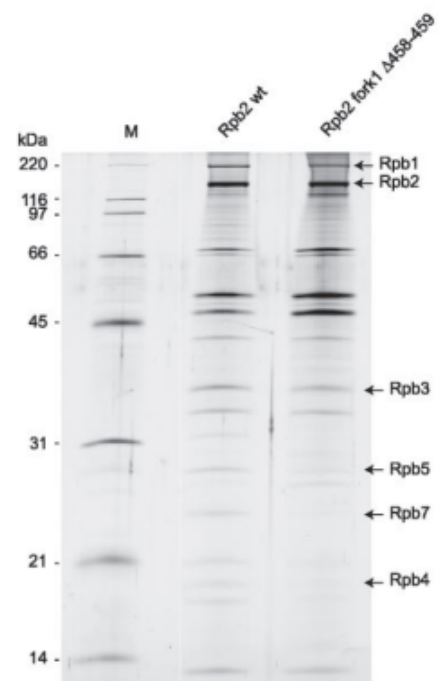


Figure 6

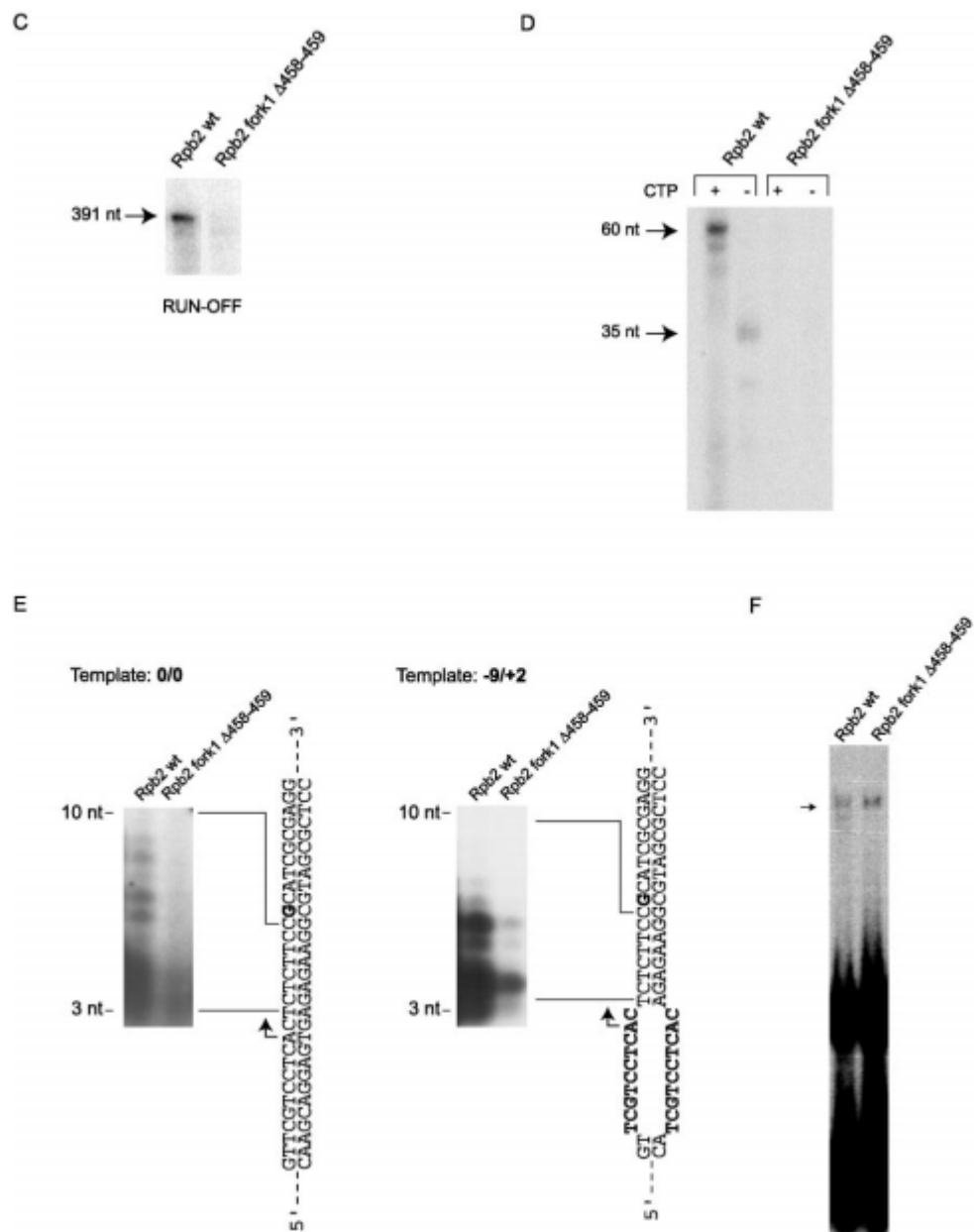


Figure 6 - continued

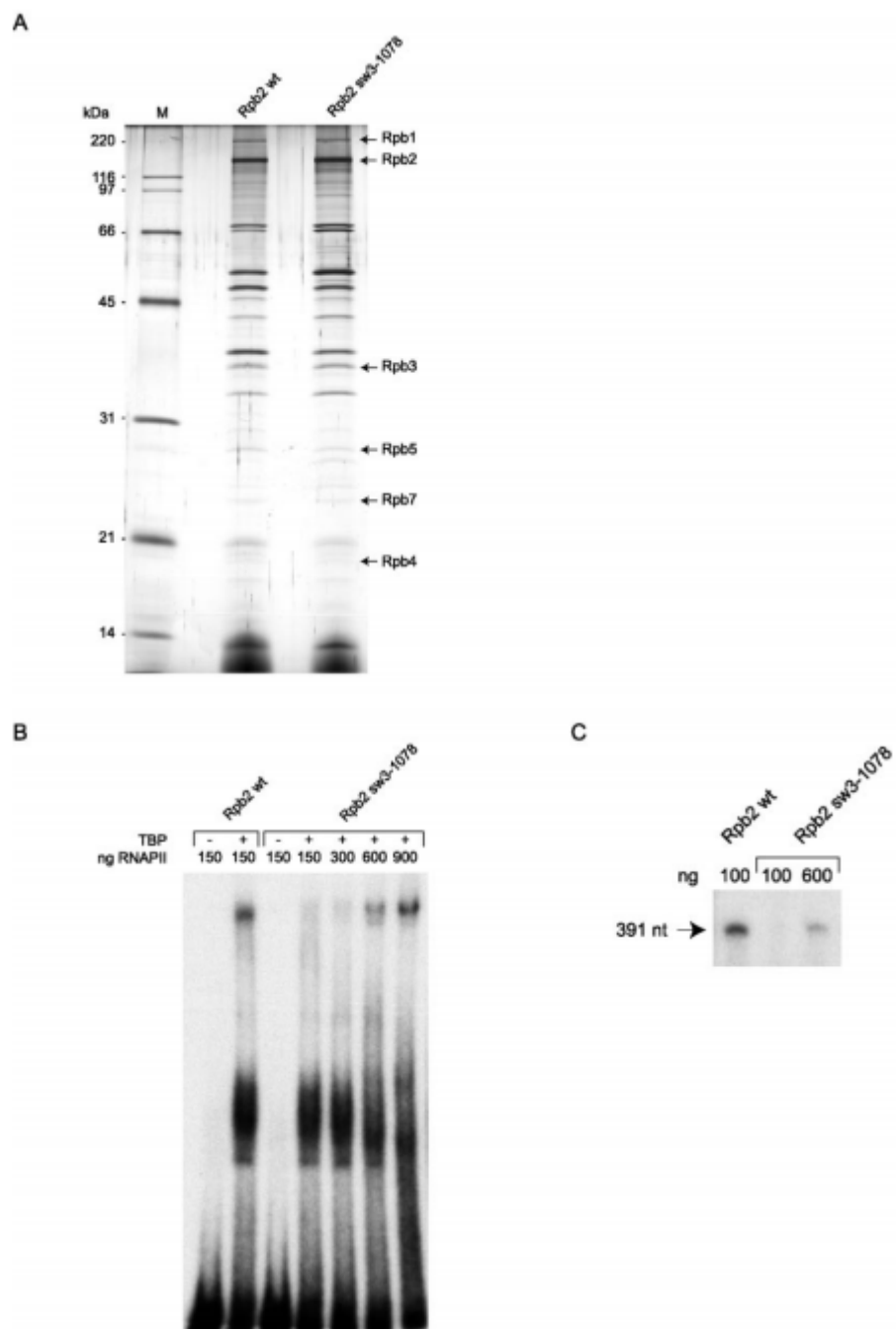


Figure 7

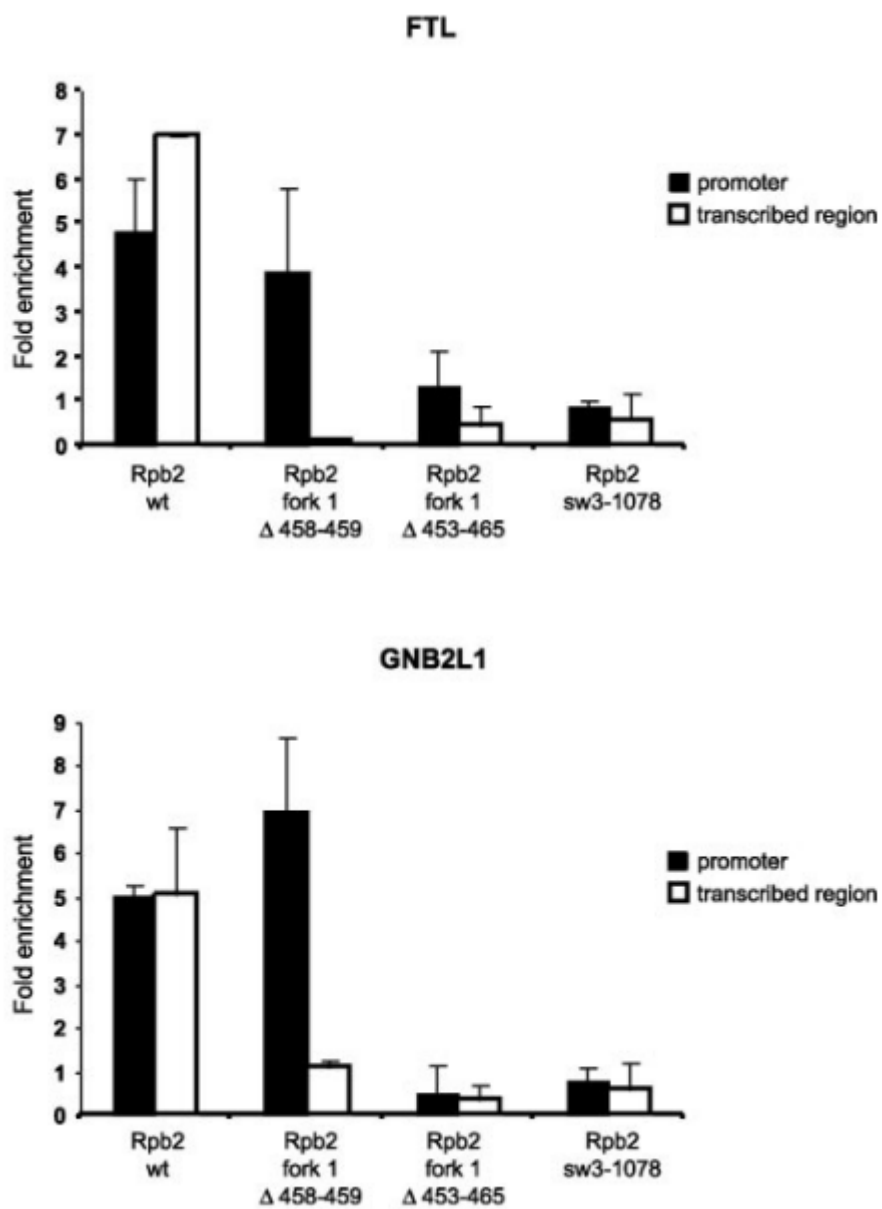


Figure 8

CHAPITRE 4

LE FACTEUR TFIIS COOPÈRE AVEC LA LIGASE DE TYPE E3 UBR5 AFIN D'UBIQUITINER LA SOUS-UNITÉ CDK9 DU FACTEUR *POSITIVE-TRANSCRIPTION ELONGATION* *FACTOR b*

4.1 Préambule

Pour cet article, j'ai effectué le clonage de TFIIS-Flag, TFIIS.1-Flag, UBR5-Flag, TFIIS.1-GST, CDK9-Rluc. J'ai établi les lignées stables TFIIS-Flag et UBR5-Flag. J'ai effectué les expériences d'immunoprécipitation et de "pull down" présentées dans la Figure 1B, 1C et 1D. J'ai fait les essais d'ubiquitination *in vitro*, *in vivo* et d'atténuation de l'expression par des siRNA présentés dans les Figures 2, 3 et 4. J'ai effectué les essais luciférase et d'atténuation de l'expression par siRNA présentés dans la Figure 5. J'ai également effectué les analyses par CHIP et d'atténuation de l'expression par siRNA présentées dans les Figures 6 et 7. J'ai écrit l'article en proportion d'environ 85 % et j'ai fait les Figures.

Annie Bouchard a effectué le clonage et l'établissement des lignées stables de TFIIS-TAP et de TFIIS.1-TAP. Elle a également purifié par la méthode TAP les complexes formés par ces protéines recombinantes (Figure 1A).

Philippe Cloutier a contribué aux expériences d'ubiquitination *in vitro* présentées dans la Figure 2, en effectuant le clonage et l'établissement de la lignée stable de UBR5-TAP. La protéine CDK9 portant une étiquette His, qui a été utilisée pour ces mêmes expériences d'ubiquitination *in vitro*, a été purifiée par Jeff Cooper et Katayoun Varzavand dans le laboratoire du Dr David Price.

4.2 Article

Cojocaru M, Bouchard A, Cloutier P, Cooper J, Varzavand K, Price D, and Coulombe B, Transcription factor IIS cooperates with the E3 ligase UBR5 to ubiquitinate the CDK9 subunit of the positive transcription elongation factor b, accepté pour publication dans Journal of Biological Chemistry.

Transcription factor IIS cooperates with the E3 ligase UBR5 to ubiquitinate the CDK9 subunit of the positive transcription elongation factor b

Marilena Cojocaru¹, Annie Bouchard¹, Philippe Cloutier¹, Jeff J. Cooper², Katayoun Varzavand², David H. Price², and Benoit Coulombe¹

From Institut de recherches cliniques de Montréal, Montréal (Québec) Canada¹, and Biochemistry Department, University of Iowa, Iowa City, Iowa 52242, USA²

Running head: Ubiquitination of CDK9 by the E3 ligase UBR5

Address correspondence to: Benoit Coulombe, Gene Transcription and Proteomics Laboratory, Institut de recherches cliniques de Montréal, 110 avenue des Pins Ouest, Montréal (Québec), Canada H2W 1R7, Tel.: (514) 987-5662, Fax: (514) 987-5663

Abstract

Elongation of transcription by mammalian RNA polymerase II (RNAPII) is regulated by specific factors, including transcription factor IIS (TFIIS) and positive transcription elongation factor b (P-TEFb). We show that the E3 ubiquitin ligase UBR5 associates with the CDK9 subunit of P-TEFb to mediate its polyubiquitination in human cells. TFIIS also binds UBR5 to stimulate CDK9 polyubiquitination. Co-localization of UBR5, CDK9, and TFIIS along specific regions of the γ fibrinogen (γ FBG) gene indicates that a ternary complex involving these factors participates in the transcriptional regulation of this gene. In support of this notion, overexpression of TFIIS not only modifies the ubiquitination pattern of CDK9 *in vivo*, but also increases the association of CDK9 with various regions of the γ FBG gene. Notably, the TFIIS-mediated increase in CDK9 loading is obtained during both basal and activated transcription of the γ FBG gene. This increased CDK9 binding is paralleled by an increase in the recruitment of RNAPII along the γ FBG gene and the phosphorylation of the RPB1 CTD on Ser2, a known target of CDK9. Together, these results identify UBR5 as a novel E3 ligase that regulates transcription, and define an additional function of TFIIS in the regulation of CDK9.

Introduction

TFIIS is the first general transcription factor reported to associate with RNAPII (Sekimizu et al., 1976). The functions of TFIIS in transcription have been extensively studied, and while early evidence pointed to a specific role in elongation, more recent studies have revealed the implication of TFIIS in early transcriptional stages such as formation of the preinitiation complex (Kim et al., 2007) and release of the polymerase from elongation complexes paused close to the promoter (Adelman et al., 2005). The architecture of TFIIS is modular and comprises three domains with distinct functions: domain III is essential for the stimulation of elongation; domain II and the flexible linker between domains II and III are required for binding to RNAPII; and domain I is required for the interaction with the Mediator and other initiation factors (Awrey et al., 1998; Kettenberger et al., 2003; Wery et al., 2004). The most well-characterized function of TFIIS in elongation consists in the rescue of backtracked RNAPII through stimulation of its intrinsic endonucleolytic activity (Nickels and Hochschild, 2004). In support of TFIIS's function in both early and late stages of transcription, chromatin immunoprecipitation studies in human cells (Cojocaru et al., 2008) and lower eukaryotes (Kim et al., 2007; Prather et al., 2005) have revealed the association of TFIIS along the entire sequence of tested genes, including promoters, and have shown that this factor is required for the recruitment of polymerase molecules at specific promoters.

TFIIS participates in important regulatory mechanisms that evolved from lower to higher eukaryotes. For instance, while yeast TFIIS is not essential for growth, TFIIS knockout mice die during embryonic development (Ito et al., 2006). The regulation of post-initiation stages of transcription was generally believed to be involved in the control of a small subset of genes (e.g., heat shock response genes in *Drosophila*, *c-myc*, *c-fos*, HIV-1) (Saunders et al., 2006). However, genome-wide studies in different systems suggest that this regulatory mechanism is much more common than initially thought (Margaritis and Holstege, 2008), given that paused polymerase molecules were identified along a majority of transcribed genes.

Because TFIIS stimulates the release of polymerase from promoter-proximal paused complexes, the regulatory potential of TFIIS could be much wider than previously believed.

The transcription elongation factor P-TEFb also stimulates the release of promoter-proximal paused RNAPII (Fujinaga et al., 2004; Ni et al., 2008; Yamada et al., 2006). This activity is mediated by P-TEFb's kinase subunit, CDK9, which phosphorylates multiple targets, including the C-terminal domain (CTD) of the Rpb1 subunit of RNAPII and the negative elongation factors DSIF and NELF (Peterlin and Price, 2006). These phosphorylation events abolish the inhibitory effect of DSIF and NELF (Fujinaga et al., 2004; Ping and Rana, 2001) and allow the phosphorylated CTD to become a platform for the recruitment of stimulatory elongation factors (Egloff and Murphy, 2008b). These concerted actions cause the release of the polymerase from its paused state and its entry into productive elongation. Notably, it has been shown that DSIF and NELF not only inhibit RNAPII, but also negatively affect the activity of TFIIS *in vitro*, both negative factors being required (Palangat et al., 2005). These data suggest that a self-regulatory loop could be formed with positive and negative elongation factors inhibiting each other; P-TEFb counteracts DSIF and NELF, which in turn inhibits TFIIS. While DSIF and NELF have been shown to function as negative elongation factors only when both proteins are present, no direct link between TFIIS and P-TEFb has been reported yet.

In this article, we report on the identification of a new interaction partner for TFIIS and CDK9, namely, the E3 ubiquitin ligase UBR5. We show that UBR5 catalyzes the ubiquitination of CDK9 *in vitro*, and that both UBR5 and TFIIS are essential for this ubiquitination event in human cells. We also show that overexpression of TFIIS increases the occupancy of CDK9 along the γ FBG gene, an event that is accompanied by an increased loading of RNAPII and an increased phosphorylation of the Ser2 residue of its RPB1 CTD. Moreover, we provide direct evidence that the effect of TFIIS on the ubiquitination of CDK9 and its occupancy level along the γ FBG gene is dependent upon the UBR5 ubiquitin ligase.

Experimental Procedures

Protein affinity purification coupled with mass spectrometry- Human EcR 293 cells stably transfected were induced with 3 μ M Ponasterone A (Invitrogen) to express TAP-tagged versions of TFIIS and TFIIS.1. Whole-cell extracts were subjected to TAP purification, and the eluates were run on SDS-PAGE gels, as previously described (Jeronimo et al., 2004). Non-induced cells were used as TAP purification controls. After silver or Sypro Ruby (Bio-Rad) staining, the bands were excised and digested with trypsin. The tryptic peptides were purified and identified by tandem mass spectrometry, as previously described (Jeronimo et al., 2007). In order to discriminate between specific and spurious interactions, we used a computational method that we have previously described (Cloutier et al., 2009; Jeronimo et al., 2007). This method uses an algorithm that assigns an interaction reliability (IR) score to each protein interaction based on parameters such as the strength of the MS score and the local topology of the network (e.g., number of common interaction partners and the identification of the interaction in reciprocal purifications). The graph in Fig. 1A shows the interaction partners of TFIIS with an IR score higher than 0.73. We estimate that above this threshold, the specificity is around 88%, and the rate of false positives is lower than 12%. By setting this threshold value of IR, we evaluate that the sensitivity of our method to detect relevant interactions approaches 72%, which suggests a rate of 28% false negatives (see Cloutier et al., 2009, for details on the computational method used to evaluate the confidence of our interaction networks).

Chemicals and immunoreagents- The following antibodies were used: monoclonal anti-TFIIS (Transduction Laboratories), M2 monoclonal anti-FLAG (Sigma), polyclonal anti-CDK9 (C-20), anti-UBR5 (N-19), RNA Pol II (N-20 and C-21), anti-ubiquitin (FL-76), anti-CAND1 (A-13) (Santa Cruz Biotechnologies), monoclonal 6X His-tag (ab18184) (Abcam), and monoclonal anti phospho-Ser² C-terminal domain (CTD) Pol II (H5) (Convance). An UBR5 antiserum produced in rabbit was kindly provided by Robert L. Southerland (Sydney, Australia) and used for immunoprecipitation assays. Ubiquitin, ubiquitin activating enzyme (UBE1),

ubiquitin conjugating enzyme (UbcH5b), and ubiquitin aldehyde (Ub-H) were purchased from BostonBiochem; recombinant human IL-6 was purchased from Invitrogen; dissucinimidyl glutarate (DSG) was purchased from Pierce; N-ethylmaleimide (NEM), adenosine 5'- (β,γ -imido) triphosphate tetralithium (AMP-PNP), and cycloheximide were purchased from Sigma; and lactacystin was purchased from Cayman Chemical.

Plasmid constructions- PCR amplification was used to clone cDNAs for TFIIS, TFIIS.1, and UBR5 into the mammalian expression vector pMZI (Zeghouf et al., 2004) carrying a TAP tag at its C terminus. From this template, we subcloned TFIIS into pCMV-3Tag-1A vector (Startagene) encoding three copies of the FLAG epitope at the N terminus of the insert. GST-TFIIS was expressed from the pGEX-4T1 plasmid (GE Healthcare) in the *Escherichia coli* BL21 Star (DE3) strain (Invitrogen). The pCMV-Tag2B-UBR5 construct, which expresses UBR5 with a FLAG epitope at its N terminus, was obtained from Robert L. Southerland (Sydney, Australia). The cDNA for CDK9 was cloned by PCR amplification into pRLuc-C2 (h) and pRLuc-N2(h) vectors (BioSignal Packard) encoding the Renilla luciferase protein. As normalizing vector in luciferase assays, we used the pGL3 plasmid (Promega) encoding the Firefly luciferase (Fluc). The pMT123 vector encoding HA-ubiquitin polypeptide was kindly provided by Sylvain Meloche (IRIC, Université de Montréal, Canada).

Cell lines and culture conditions- Stable human embryonic kidney Ecr 293 cells (derived from HEK 293) were produced as previously described (Jeronimo et al., 2004), and grown in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% FBS, 2 mM L-glutamine, 30 μ g/ml Bleocin (Calbiochem), and 300 μ g/ml G418 (Invitrogen). Human HeLa and HepG2 cells were maintained in Eagle's minimal essential medium (EMEM) supplemented with 10% FBS and 2 mM L-glutamine. Stable HepG2 cell lines were obtained by transfection with the FuGENE reagent (Roche) and a selection of stable clones with 600 μ g/ml of G418 (Invitrogen). For IL-6 induction, the cells were serum starved for at least 16 h, and IL-6 was added at a final concentration of 10 ng/ml for 30 min.

GST pull-down and immunoprecipitation assays- TFIS-GST was purified by affinity chromatography on glutathione-Sepharose 4B (GE Healthcare Biosciences) according to the manufacturer's instructions. Sepharose-bound TFIS-GST was incubated with HEK 293 whole-cell extracts (1 mg) prepared under native conditions (Rigaut et al., 1999). After thorough washing of the beads, the bound proteins were eluted at the indicated ionic strength, resolved on SDS-PAGE, and immunoblotted with indicated antibodies. For immunoprecipitation experiments, 0.5-1 mg of HEK 293 whole-cell extracts prepared under native conditions were incubated with indicated antibodies, and the immunocomplexes were captured on protein A-Sepharose beads. After extensive washing with IPP100G (10 mM Tris, 100 mM NaCl, 0.1% Triton X100), the bound proteins were eluted, resolved on SDS-PAGE, and immunoblotted with the indicated antibodies.

siRNA knockdown experiments- Pools of ON-TARGET plus Smartpool oligonucleotides (four oligonucleotides) targeting TFIS, UBR5, and, as a negative control, a mix of four oligonucleotides that do not target any known human gene (Dharmacon) were used. HeLA and HepG2 cells cultured in 6-well dishes were transfected using Oligofectamine (Invitrogen) and siPORT NeoFX transfection agent (Ambion), respectively, according to manufacturers' instructions. For efficient knockdown, we used a double transfection procedure: after the first transfection, the cells were incubated at 37° C for 24 h and then retransfected using the same protocol. To determine the knockdown efficiencies, the cell lysates obtained after 48, 72, and 96 h of culture were resolved on SDS-PAGE gels and tested by immunoblotting with relevant antibodies. Since the most efficient knockdown was obtained after 48 h (data not shown), these conditions were used for all siRNA knockdown experiments.

In vitro ubiquitination assay- Full-length CDK9 carrying a C-terminal 6-His tag was obtained by baculovirus expression along with cyclin T1 (1-290), followed by Ni-NTA and Mono S purification. UBR5 was purified by the TAP procedure from a human EcR 293 cell line stably expressing full-length UBR5 with a C-terminal TAP-tag. After the second step of the TAP purification, the calmoduline beads containing immunoprecipitated UBR5 were washed twice with cold ubiquitination

buffer (20 mM Tris-HCl pH 7.5, 50 mM NaCl, 1 mM MgCl₂, 1 mM DTT, 0.5% Nonidet P-40, and protease inhibitor cocktail), then washed twice with 50 mM Tris-HCl pH 7.5 and 1 mM DTT. The reaction mixture (60 µl) for *in vitro* ubiquitination assays contained E1 (150 ng), E2 (UbcH5b, 500 ng), purified CDK9-His protein (2.5 µg), AMP-PNP (2 mM), ubiquitin aldehyde (5 µM), and 10% of the pellet containing affinity purified UBR5. The assay was performed in ubiquitination buffer: 50 mM Tris-HCl (pH 7.5), 5 mM MgCl₂, and 1 mM DTT. Following pre-incubation at 37°C for 5 minutes, ubiquitin (500 ng) was added to start the reaction. The mixtures were incubated for 55 min with gentle agitation every 10 min. The reaction was stopped by adding SDS-loading buffer and boiling for 10 min. After the isolation of beads by centrifugation (1 min at 1000 RPM), the supernatant was loaded on 6% SDS-PAGE gels. The ubiquitination reactions were analysed by Western blotting.

In vivo ubiquitination assays- Transient transfection of HeLa cells with vectors driving the expression of TFIIS-FLAG or the empty vector, together with a ubiquitin-HA expression construct, was performed with calcium phosphate in 100 mm dishes. Lactacystin (12 µM) was added for 5 h. After 48 h of culture, the cells were lysed for 1 h at 4°C in lysis buffer containing 1% NP-40, 10 mM Tris-HCl pH 7.4, 150 mM NaCl, 2 mM EDTA, 10% glycerol, 2 mM N-ethylmaleimide (NEM), and protease inhibitor cocktail (Roche), and supplemented with 6 M urea. The soluble lysate was diluted to 0.3 M urea with lysis buffer supplemented with 1 mM DTT, and immunoprecipitated at 4°C overnight with a CDK9 antibody (10 µg). The immunocomplexes were collected with Protein A Sepharose beads, and after extensive washing, they were eluted with reducing SDS sample buffer. Eluates and a sample (5%) of the whole-cell extracts were tested by immunoblotting with the indicated antibodies, using a standard procedure. The ubiquitination in HepG2 cells stably expressing UBR5-FLAG or TFIIS-FLAG was assessed using the same procedure.

To analyze the effect of TFIIS or UBR5 knockdown on the ubiquitination of CDK9, HeLa cells cultured in 6-well plates were transfected with siRNAs and co-transfected with the DNA plasmids and siRNAs, as indicated in the figures, using

oligofectamine. After 24 h, the cells were transfected a second time with siRNAs. Final concentrations of siRNAs were 25 and 50nM, as indicated. After a total culture time of 48h, the cells were lysed and subjected to ubiquitination assays and to parallel verification of the knockdown efficiency by immunoblotting with relevant antibodies.

Luciferase assay- HeLa cells were transfected directly into white 96-well plates with plasmids expressing N- or C-terminal CDK9-Renilla luciferase fusion proteins and with a normalizing plasmid expressing the Firefly luciferase (Fluc) using Oligofectamine (Invitrogen). UBR5-targeting and control siRNAs were co-transfected with the DNA plasmids, and after 24 h, a second transfection of siRNAs was performed. After a total of 48 h, the Renilla and Firefly luciferase activities were measured with the Dual-Glo Luciferase system (Promega), according to manufacturer's instructions. *De novo* protein synthesis was inhibited by adding cycloheximide at a final concentration of 100 µg/ml, and the luciferase substrates were added at the indicated time points.

Chromatin immunoprecipitation- Two-step ChIP assays were performed as previously described (Nowak et al., 2005). HepG2 cells were cultured in 100-mm dishes, and $4-6 \times 10^6$ cells were cross-linked at room temperature in two steps. First, DSG was added for 45 min. for protein-protein cross-linking, and second, protein-DNA cross-linking was performed by adding formaldehyde for 10 min. After washing with PBS, the cells were lysed, and nuclei were resuspended in SDS lysis buffer and sonicated on ice 12 times using 15 s pulses (Fisher Scientific Model 100 Sonic Dismembrator). The average length of DNA fragments was 350 bp. The immunoprecipitation of equal amounts of DNA was performed overnight at 4°C in ChIP dilution buffer. We used 4 µg of the indicated antibodies, or 10 µg in the case of the H5 antibody. The immunoprecipitated complexes were collected with Protein A Sepharose (GE Healthcare) for CDK9 (C20) and RNAPII (N20) antibodies, and with Dynabeads PanMouse IgG or rat anti-mouse IgM (Invitrogen) for anti-FLAG (M2) and anti-RNAPII phospho-Ser2-CTD form (H5) antibodies, respectively. After washing and elution, the cross-links were reversed at 65°C for 4 h, and the DNA was purified by phenol/chloroform extraction. The quantification of the

immunoprecipitated DNA fragments was performed by quantitative PCR (QPCR), as described (Jeronimo et al., 2004). The primers used are described elsewhere (Hou et al., 2007). For siRNA knockdowns coupled with ChIP experiments, HepG2 cells were cultured in 100-mm dishes, a double transfection of siRNAs was performed with siPORT NeoFX (Ambion), and after a total culture time of 48 h, the cells were subjected to the two-step ChIP assay as described above. For ChIP-re-ChIP experiments, complexes were washed with washing buffer containing a protease inhibitor cocktail (Roche Diagnostics), eluted by incubation for 30 min. at 37°C in Re-ChIP buffer (0.5 mM DTT, 1% Triton X-100, 2 mM EDTA, 150 mM NaCl, 20 mM Tris-HCl, pH 8, protease inhibitor cocktail), and subjected again to the ChIP procedure.

RESULTS

TFIIS interacts with the E3 ubiquitin ligase UBR5 and CDK9. To map the space of the human interactome that contains the transcription elongation factor TFIIS (official gene symbol: TCEA1) and to identify novel interaction partners that may reveal new functions for TFIIS, we used affinity purification of TAP-tagged TFIIS coupled with mass spectrometry (AP-MS) to identify the interaction partners. High confidence interactions were selected computationally, as previously described (Cloutier et al., 2009; Jeronimo et al., 2007). As expected, TFIIS was found to interact with RNAPII and components of the RNA processing machinery (Fig. 1A). The interaction graph in Fig. 1A reveals the existence of two proteins that have not been previously reported to interact with TFIIS, namely, TFIIS.1 (official gene symbol: TCEA2) and UBR5. TFIIS.1 is a TFIIS-related protein that is expressed in specific tissues, with high level of expression in the testis (Umehara et al., 1995). Our functional analysis of TFIIS.1 revealed that, similarly to TFIIS, this protein stimulates transcriptional elongation *in vitro* and presents the same localization pattern in human cells (data not shown), suggesting similar roles *in vivo*. Based on these results, we decided not to proceed with further analysis of the TFIIS-TFIIS.1 interaction. UBR5 is a member of the HECT-domain family of E3 ubiquitin ligases

(Callaghan et al., 1998). The presence of UBR5 in the TFIIS-TAP eluate was confirmed by Western blotting (data not shown). To confirm the TFIIS-UBR5 interaction, we performed *in vitro* pull-down experiments using a TFIIS-GST fusion protein coupled with glutathione beads and a HEK 293 whole-cell extract. Immunoblotting of the eluate with anti-UBR5 antibody revealed the presence of UBR5 (Fig. 1B).

According to the literature and our unpublished data, it was unlikely that TFIIS and RNAPII represent targets for UBR5 ubiquitination (data not shown). Recent reports, however, have indicated that CDK9, the kinase subunit of transcription elongation factor P-TEFb, which also interacts with RNAPII, is ubiquitinated in human cells (Barboric et al., 2005; Kiernan et al., 2001). We tested the hypothesis of UBR5 being involved in CDK9 ubiquitination, possibly through an interaction with TFIIS and/or RNAPII. We first assessed whether CDK9 can interact with TFIIS and UBR5 by probing the TFIIS-GST eluate for the presence of CDK9 by Western blotting. Figure 1B shows that both UBR5 and CDK9, but not a control protein, are present in the TFIIS-GST pull-down eluate. CAND1 (cullin-associated and neddylation-dissociated 1) was used as a control because this protein was identified as an interaction partner of TFIIS.1 but not of TFIIS in our affinity purification experiments.

The constitutive interaction between these three proteins was further confirmed using immunoprecipitation with anti-CDK9 (Fig. 1C) and anti-UBR5 antibodies (Fig. 1D). These results indicate that TFIIS, UBR5, and CDK9 interact together to form a complex *in vivo*. Of note, the Rpb1 subunit of RNAPII is also found in the anti-UBR5 immunoprecipitate (Fig. 1D). This result suggests that UBR5 is indeed part of a complex also containing RNAPII, and confirms the result of our systematic AP-MS analysis (Fig. 1A).

UBR5 mediates the ubiquitination of CDK9 in vitro. Since a previous report indicated that endogenous CDK9 molecules are ubiquitinated (Barboric et al., 2005), it was legitimate to speculate that its interaction partner, UBR5, may be involved in this modification. To test this hypothesis, we first performed *in vitro* ubiquitination assays using TAP affinity purified UBR5 and CDK9 purified from

insect cells as a substrate. When the complete set of components was used in the ubiquitination reaction, a robust polyubiquitination of CDK9 was observed, with the appearance of slower migrating high molecular weight species (Fig 2; Lanes 3, 6, and 8). These forms contained ubiquitinated CDK9 molecules, as they were detected with both antibodies directed towards CDK9 (anti-tag and anti-CDK9), as well as with the anti-ubiquitin antibody. When one of the components was omitted (E1, E2, CDK9, or UBR5) no ubiquitination was observed. These results show that UBR5 directly catalyzes the ubiquitination of CDK9 *in vitro*.

TFIIS and UBR5 are required for the ubiquitination of CDK9 in vivo. We next performed *in vivo* ubiquitination assays in extracts from cells in which either UBR5 or TFIIS were individually overexpressed or depleted. HeLA cells were co-transfected with constructs driving the expression of TFIIS carrying a FLAG epitope (TFIIS-FLAG) and an ubiquitin-HA peptide. Co-transfections of the empty vector and the ubiquitin-HA construct were used as controls. The cells were treated with the proteasome inhibitor lactacystin, and the whole-cell extracts were incubated with an anti-CDK9 antibody. After CDK9 was isolated, the eluates were resolved on SDS-PAGE and immunoblotted with anti-ubiquitin, anti-CDK9, or anti-FLAG antibodies. The overexpression of TFIIS markedly increased the smear of high molecular weight proteins as compared to the controls (Fig. 3A, right panel). This effect was specific to CDK9 since no significant modification in the general ubiquitination pattern was observed in the whole-cell extracts (Fig. 3A, left panel). Immunoblotting with an anti-CDK9 antibody shows that the high molecular weight species are polyubiquitinated CDK9 molecules (Fig. 3B). Moreover, these species appear specifically in the eluates from the CDK9 immunoprecipitates but not in the eluates from mock immunoprecipitation (Fig. 3C). Interestingly, proteasome inhibition by lactacystin does not affect the stability of the polyubiquitinated CDK9 species, while a stabilization of the polyubiquitinated proteins is observed in the whole-cell extracts using the same treatment (Fig. 3C).

To assess whether UBR5 is involved in the ubiquitination of CDK9 *in vivo*, we performed ubiquitination assays using extracts from HepG2 cells stably transfected with a construct expressing UBR5-FLAG. As controls, we used extracts

prepared in parallel from a stable HepG2 cell line expressing TFIIS-FLAG or from parental HepG2 cells. Figure 3D shows that overexpression of UBR5 specifically increased the smear of high molecular weight CDK9 species.

To confirm the role of UBR5 and TFIIS in CDK9 ubiquitination, siRNA treatments were used to knockdown each protein individually in HeLa cells (Fig. 4B and 4D, respectively), and the whole-cell extracts were subjected to CDK9 immunoprecipitation and *in vivo* ubiquitination assays, as described above. The knockdown of each of TFIIS or UBR5 proteins reduced the smear of polyubiquitinated CDK9 species in a dose-dependent fashion compared to control non-targeting siRNAs (Fig. 4A and 4C, respectively). Together, these results indicate that both UBR5 and TFIIS are involved in CDK9 ubiquitination *in vivo*.

UBR5 is required for the effect of TFIIS on CDK9 ubiquitination in vivo. Since both TFIIS and P-TEFb stimulate transcriptional elongation by RNAPII, we hypothesized that TFIIS may potentiate the activity of P-TEFb by recruiting the E3 ubiquitin ligase UBR5 to facilitate CDK9 ubiquitination. To test this idea, we performed *in vivo* ubiquitination assays in HeLa cells co-transfected with constructs driving the expression of TFIIS-FLAG and siRNAs targeting UBR5. As shown in Figure 4E, TFIIS overexpression increased the accumulation of high molecular weight CDK9 species as compared to transfection with the empty vector when a non-targeting control siRNA was co-transfected. However, co-transfection with the UBR5-targeting siRNA almost completely abolished this effect on CDK9 ubiquitination. This suggests that the effect of TFIIS on CDK9 ubiquitination is related, to a significant extent, to the activity of UBR5. The residual ubiquitination of CDK9 may be due to the incomplete depletion of UBR5.

The knockdown of UBR5 does not affect the stability of CDK9. While endogenous CDK9 is a long-lived protein, with a half-life ($t_{1/2}$) between 4 and 7 h depending on the cell type, the exogenously overexpressed CDK9 fusion protein was reported to degrade faster, with a $t_{1/2}$ dependent on the level of its overexpression (Garriga et al., 2003). Since we have shown that the ubiquitination of CDK9 is dependent on UBR5, we reasoned that if the role of this modification is to target CDK9 for proteosomal degradation, then the knockdown of UBR5 should have a detectable

effect on the stability of this kinase. We used a luciferase-based reporter assay combined with siRNA treatment to evaluate this hypothesis. Constructs encoding N- or C-terminal CDK9-Renilla luciferase fusion proteins were co-transfected into HeLa cells with a normalizing plasmid expressing the Firefly luciferase (Fluc) and with UBR5-targeting or control siRNAs. After a double-transfection procedure, *de novo* protein synthesis was blocked by incubating the cells with cycloheximide for 2 to 4 h, and the luciferase activity was measured with a dual luciferase assay (Gagnon et al., 2009). As a control, we also performed the same set of experiments with the Renilla luciferase plasmid with no attached CDK9 moiety. The results indicate that, in our conditions, the overexpressed CDK9 has a half-life of less than 2 h (Fig. 5, upper panel), while the control protein Renilla luciferase (Rluc) was relatively stable over the duration of the experiment (Fig. 5, lower panel). In cells in which UBR5 was knocked down, the kinetics of CDK9 degradation was similar to that of cells transfected with control non-targeting siRNA or mock-treated cells (Fig. 5, upper panel). More specifically, the luciferase activity of the CDK9-Rluc fusion protein, normalized to that of Fluc in each sample, presents similar values at all time-points and for all tested conditions. Similar results were obtained when the RLuc, fused to the N-terminal (Fig. 5, upper panel) or C-terminal (data not shown) ends of CDK9 was analyzed. These results suggest that UBR5 is not involved in regulating CDK9 stability, and that the ubiquitination event mediated by this E3 ligase does not target CDK9 for degradation by the proteasome.

TFIIS increases CDK9 occupancy along the γ FBG gene in a UBR5-dependent manner, and activates its transcription. It was shown previously that the ubiquitination of CDK9 facilitates transactivation of HIV-1 genome transcription by Tat (Barboric et al., 2005), suggesting that CDK9 ubiquitination may have a stimulatory effect on transcription. One possibility is that TFIIS promotes the activity of this kinase by mediating its ubiquitination by UBR5. To investigate this hypothesis, we used as a model the γ FBG gene, which has been shown to be regulated through the recruitment of CDK9 upon IL-6 induction of HepG2 hepatocarcinoma cells (Hou et al., 2007). ChIP assays were performed in HepG2 cells with a two- step cross-linking procedure. Parental HepG2 cells, or a stable

HepG2 cell line expressing the TFIIS-FLAG protein, were used for this ChIP procedure. The positions of the primers used to analyze the recruitment of different proteins along the γ FBG gene are shown in Figure 6A. Figure 6B shows a significant increase in CDK9 occupancy levels along the γ FBG gene upon IL-6 induction, as reported previously (Hou et al., 2007). Overexpression of TFIIS resulted in increased CDK9 occupancy levels along the sequence of this gene in basal conditions, with a further increase after IL-6 induction (Fig. 6B). This finding suggests that TFIIS stimulates the association of CDK9 with the γ FBG gene chromatin.

To assess whether this effect is due to the increased recruitment of TFIIS itself on the γ FBG gene, we performed ChIP assays with an anti-FLAG antibody in the HepG2 cells stably expressing the FLAG-TFIIS protein. As shown in Figure 6C, TFIIS is detected both in the region of the TATA box and in downstream transcribed regions of the γ FBG gene, with higher enrichment values after IL-6 induction. The low TFIIS ChIP signals in the IL-6-RE1 region, both in basal conditions and after IL-6 induction, suggest that the effect of TFIIS on CDK9 occupancy in this region may be the result of modifications induced to CDK9 before its recruitment to the preinitiation complex. A similar set of experiments was performed in HepG2 cells stably expressing a form of UBR5 carrying a FLAG-tag. The location profile of UBR5 along the γ FBG gene was similar to that observed for TFIIS (Fig. 6D). This supports the idea that TFIIS mediates the effect of UBR5 on CDK9 via a physical association of all three proteins in the same genomic regions. As for TFIIS, the UBR5 signals observed in the IL-6-RE1 region were low (Fig. 6D), supporting the notion that the modifications of CDK9 by UBR5 may take place both before and after the recruitment of CDK9 to the chromatin.

To further define whether the effect of TFIIS on CDK9 loading along the γ FBG gene requires UBR5, we assessed the impact of UBR5 knockdown (Fig. 6E, lower panel) on CDK9 occupancy in cells that overexpress TFIIS. The results show a decreased CDK9 occupancy of the promoter and exon 5 regions upon UBR5 depletion, as compared to a control siRNA treatment (Fig. 6E, upper panel). The effect of UBR5 knockdown on the CDK9 occupancy at the TATA box and exon 5

regions cannot be explained by a lower loading of CDK9 in the upstream IL6_RE1, as no significant effect was observed in this upstream region. Rather, this finding suggests a local, on-chromatin event. This conclusion is supported by the presence of the three proteins, CDK9, TFIIS, and UBR5, on TATA box and exon 5 regions, where an effect of UBR5 knockdown on the CDK9 occupancy is observed, paralleled by the absence of TFIIS and UBR5 in the IL6_RE1 region where no significant effect of UBR5 knockdown is observed. Taken together, these results argue in favor of a model in which the E3 ubiquitin ligase UBR5 mediates the activity of TFIIS on CDK9 in a mechanism that stimulates the loading of CDK9 along transcribed regions.

To investigate whether the effect of TFIIS on the recruitment of CDK9 has a functional consequence on transcription, we performed ChIP assays with two RNAPII antibodies. One recognizes all forms of RNAPII (referred to here as total RNAPII) (Fig. 6F) and one specifically recognizes phospho-Ser2 on the CTD of the large subunit (Fig. 6G). Ser2 residues are targets of CDK9 and RNAPII containing phospho-Ser2 is found in the 3' region of genes (Rahl et al., 2010). We analyzed both the parental HepG2 and TFIIS overexpressing cells. Total RNAPII loading on the γ FBG gene is significantly increased upon induction. Notably, the overexpression of TFIIS increased the recruitment of total RNAPII to the promoter both in basal and induced conditions, suggesting that TFIIS increases the *de novo* recruitment of RNAPII. This conclusion is supported by the finding that TFIIS and RNAPII have similar location patterns on the γ FBG promoter in cells overexpressing TFIIS both in basal conditions and after IL-6 induction (Fig. 6C and 6F). Figure 6G shows that the occupancy level of Phospho-Ser2-CTD RNAPII is increased both in promoter and downstream transcribed regions upon IL-6 induction of parental HepG2 cells, which is again in agreement with previous results (Hou et al, 2007). Of note, overexpression of TFIIS markedly increased the occupancy level of this form of polymerase along the promoter and downstream transcribed regions, and the induction with IL-6 further increased the enrichment values. These results support the notion that the effect of TFIIS on the association of CDK9 with γ FBG

gene is functionally relevant (see discussion). Notably, neither TFIIS nor the Phospho-Ser2-CTD RNAPII was detected in the IL6_RE1 region.

Finally, in order to determine if TFIIS, UBR5 and CDK9 belong to the same protein complex on the γ FBG gene, we performed ChIP-re-ChIP experiments in cells stably expressing FLAG-tagged TFIIS or UBR5 proteins, by using anti-CDK9 and anti-FLAG antibodies in sequential immunoprecipitations. Figure 6H shows that, when promoter-bound CDK9 is immunoprecipitated, both TFIIS and UBR5 are co-immunoprecipitated, as a significantly increased enrichment is observed in cells overexpressing TFIIS-FLAG or UBR5-FLAG, but not in parental HepG2 cells. Nevertheless, no significant enrichment was observed at the IL6_RE1, suggesting, on one hand, that TFIIS and UBR5 are not associated with CDK9 over this region, and, on the other hand, that the signal obtained on the promoter region is specific. No co-localization signal for the three proteins was observed in transcribed regions downstream from the promoter (data not shown).

DISCUSSION

In this study, we provide evidence that TFIIS is involved in the ubiquitination of CDK9 (the kinase subunit of P-TEFb) by recruiting the E3 ubiquitin ligase UBR5. Our results indicate that UBR5 catalyzes the polyubiquitination of CDK9, and this modification does not lead to the degradation of CDK9. Rather, TFIIS increases the loading of CDK9 over a cytokine inducible gene in a UBR5-dependent fashion.

Independent reports showing that TFIIS and P-TEFb are required for the entry of RNAPII into productive elongation (Adelman et al., 2005; Ni et al., 2008) are in agreement with our findings. Our study provides a link between these two factors and reveals a new function of TFIIS in recruiting the E3 ubiquitin ligase UBR5 and thereby potentiating the loading of CDK9 on chromatin. UBR5 is a HECT domain E3 ubiquitin ligase involved in various processes, including the transactivation of progesterone and miocardin-controlled transcription (Henderson et al., 2002; Hymowitz and Wertz, 2010), the DNA damage response (Henderson et

al., 2006) and, on a larger scale, the initiation, maintenance, and/or termination of cell proliferation (Mansfield et al., 1994). UBR5 amplification and overexpression was reported to be common in different types of cancers, such as ovarian, breast, and hepatocellular carcinoma (Chin et al., 2007b; Clancy et al., 2003). Interestingly, UBR5 has been shown to interact with and activate the DNA damage checkpoint kinase CHK2 (Henderson et al., 2006). Our finding that endogenous UBR5 interacts under normal conditions with both TFIIS (Fig. 1A, 1B and 1D) and CDK9 (Fig. 1C and 1D) suggests that UBR5 participates in general RNAPII transcription mechanisms. Indeed, the chromatin immunoprecipitation of a FLAG-tagged version of UBR5 shows its association with transcribed regions in normal conditions (Fig. 6D), a finding that is further supported by the identification of RNAPII itself in immunoprecipitates of endogenous UBR5 (Fig. 1D).

Previous studies also pointed to a possible link between TFIIS and P-TEFb. TFIIS-deficient yeast cells were shown to present an increased dependence on BUR1/BUR2 or CTK1, the yeast orthologues of P-TEFb (Lindstrom and Hartzog, 2001). Also, by using a combination of defined and crude *in vitro* transcription systems, a study conducted in humans concluded in the existence of a factor facilitating the function of P-TEFb in transcriptional elongation (Cheng and Price, 2007). Notably, whereas TFIIS, a non-essential gene in yeast, evolved to gain essential functions in higher eukaryotes, UBR5 does not have a yeast homologue and appeared later during evolution.

CDK9 is known to be ubiquitinated *in vivo*, and the level of its expression was suggested to be determinant for the fate of the modified protein. When overexpressed, CDK9 is ubiquitinated and rapidly degraded (Kiernan et al., 2001); however, the endogenous protein is long-lived (Garriga et al., 2003), and its ubiquitination may rather modulate its function in transcription by RNAPII (Barboric et al., 2005; Kiernan et al., 2001). Our results show that TFIIS is involved in the ubiquitination of endogenous CDK9 by recruiting the E3 ubiquitin ligase UBR5. Four lines of evidence support this conclusion. Firstly, the interaction partner of TFIIS, UBR5, directly catalyzes the ubiquitination of CDK9 in an *in vitro* ubiquitination assay (Fig. 2). Secondly, when overexpressed, both TFIIS and UBR5

increase the smear of polyubiquitinated CDK9 species in *in vivo* ubiquitination assays (Fig. 3). Thirdly, the knockdown of each of these proteins decreases the levels of ubiquitination of endogenous CDK9 (Fig 4A and 4C). Fourthly, the knockdown of UBR5 abolishes the capacity of TFIIS to increase the smear of polyubiquitinated CDK9 species (Fig. 4E). The appearance of new bands of polyubiquitinated species of CDK9 upon overexpression of TFIIS or UBR5 may indicate the use of new sites of ubiquitination on CDK9, or alternatively, the generation of intermediate polyubiquitinated species on the same site. On the other hand, dual ubiquitin ligase/deubiquitylase activities have been reported, as for example in the case of A20, an essential enzyme for attenuating NF- κ B signaling (Hymowitz and Wertz, 2010). Based on our results, we cannot exclude that, when overexpressed, UBR5 may rather act as a deubiquitylase, leading to the appearance of intermediate polyubiquitinated species of CDK9. This may be relevant for the case of tumors presenting an overexpression of UBR5. Notwithstanding, *in vivo* ubiquitination experiments after a knockdown of UBR5 suggest that physiological levels of UBR5 are necessary for the ubiquitination of CDK9. These aspects are currently under further investigation.

In a previous report, the E3 ubiquitin ligase SCF^{SKP2} (SKP1-Cul1-F-box-protein ubiquitin ligase, S-phase-kinase-associated protein 2) was shown to ubiquitinate CDK9 (Barboric et al., 2005). Because we did not identify this ubiquitin ligase as a TFIIS interaction partner in our systematic AP-MS analysis, it is possible that different ubiquitination pathways participate in the ubiquitination of CDK9. The context could specify the ubiquitination pathway; for example, SCF^{SKP2} was shown to mediate the ubiquitination of CDK9 in the context of Tat transactivation of HIV1 transcription (Barboric et al., 2005) or to serve as a signal for proteosomal degradation when CDK9 is overexpressed (Kiernan et al., 2001). As a precedent, it has been shown that p27 is efficiently ubiquitinated by SCF^{SKP2} only when it forms a complex with a CDK (cyclin-dependent kinase) but not when it is in the free, non-complexed form (Lacy et al., 2004). Our results showing that the ubiquitination of endogenous CDK9 is dependent upon TFIIS and UBR5

suggest a role for this modification as a downstream event in processes mediated by TFIIS.

We used a sensitive luciferase assay (Gagnon et al., 2009) to evaluate whether the ubiquitination of CDK9 mediated by UBR5 has the role of triggering its degradation. Our results show that this is not the case, since the knockdown of UBR5 has no effect on the kinetics of CDK9 degradation (Fig. 5, upper panel), suggesting that the depletion of UBR5 does not stabilize CDK9. This finding further suggests that CDK9 ubiquitination by UBR5 does not act in targeting CDK9 for degradation when it is overexpressed. Moreover, the endogenous CDK9 protein was not stabilized by the knockdown of UBR5 or that of TFIIS (Fig. 4B and 4D), implying that when CDK9 is expressed at physiological levels, its ubiquitination by UBR5 has functions other than triggering its degradation, as in the case of the overexpressed protein. In support of this notion, the inhibition of proteosomal degradation by lactacystin did not affect the level of polyubiquitinated species of CDK9, while a parallel stabilization of the general polyubiquitinated proteins is observed in the whole-cell extracts from the same cells (Fig. 3C). Based on these results, it is tempting to speculate that polyubiquitinated CDK9 is recycled by deubiquitination, which would be in agreement with endogenous CDK9 being a long-lived protein, despite its ubiquitination in normal conditions. Alternatively, the polyubiquitin chains attached to CDK9 molecules could be edited, leading to a later degradation, as it has been reported for other proteins, such as Rip1 and IRAK1 (Newton et al., 2008).

Experiments using pharmacological inhibitors or depletion by a siRNA treatment showed that CDK9 is generally required for RNAPII transcription (Chao and Price, 2001; Rahl et al., 2010; Shim et al., 2002). Its inhibition may affect transcription in two different ways. First, an accumulation of initiated RNAPII molecules, but paused in the promoter-proximal region of the transcription unit, was detected in the case of the *Hsp70* gene in *Drosophila melanogaster* (Ni et al., 2008). Second, an inability to initiate transcription was observed for several cytokine inducible genes, such as those induced by IL-6 and TNF (Brasier, 2008). We took advantage of an IL-6 inducible system in order to further investigate the effect of

TFIIS and UBR5 on CDK9. We elected to use the γ FBG gene as a model because its expression upon IL-6 induction has been shown to be dependent upon P-TEFb (Hou et al., 2007), providing a framework for the study of both initiation and post-initiation stages of transcription.

The activation of the γ FBG gene by IL-6 involves STAT3 activation by the JAK/Tyk tyrosine kinase pathway (Heinrich et al., 1998), which provokes the nuclear translocation of STAT3, the formation of a complex with CDK9, and their recruitment to the IL-6 response elements of the γ FBG gene (Hou et al., 2007). By using antibodies detecting either all forms of RNAPII (total RNAPII) or the phospho-Ser2 CTD form, we showed that the enzyme is located both on the promoter and downstream transcribed regions only after IL-6 induction (Fig 6F and 6D). This finding suggests that upon IL-6 treatment, the transcription machinery is recruited to the promoter of this gene with no RNAPII molecules engaged in transcription or paused in the promoter-proximal region in the basal state. Although we do not know whether the polymerase is prone to stall in the promoter-proximal regions of the γ FBG gene upon IL-6 induction, our results indicate that TFIIS is recruited to these regions. Its occupancy profile, paralleling that of UBR5, shows a similar increase in the promoter and downstream transcribed regions upon IL-6 induction, suggesting that TFIIS plays a role in both the initiation and elongation stages of transcription of this gene (Fig 6C and 6D).

Our ChIP experiments show that TFIIS is able to enhance the expression of the γ FBG gene in basal conditions. When TFIIS is overexpressed, the occupancy of CDK9 along the promoter and downstream transcribed regions presents a significant increase compared to parental uninduced cells, and this is paralleled by a significant increase of the total RNAPII and of its phospho-Ser2 CTD form (Fig 6B, 6F, and 6G). This finding suggests that TFIIS somehow mimics the events that follow STAT3 activation by IL-6. Remarkably, the occupancy of CDK9 is significantly increased on the upstream IL6_RE1 element by TFIIS overexpression, even though TFIIS does not itself occupy this region. Moreover, the polymerase occupancy is not increased concomitantly on the IL6_RE1 region, an effect that mimics RNAPII occupancy upon IL-6 induction, arguing against a non-specific effect of TFIIS on

CDK9 occupancy. Based on this observation we propose that TFIIS associates with CDK9 before its recruitment to DNA and induces modifications that promote the binding of CDK9 to the chromatin. Of note, concomitant TFIIS overexpression and IL-6 induction do not have an additive effect, since the occupancy levels of CDK9 on the IL6_RE1 region under these conditions are similar to those observed after an IL-6 treatment of parental cells. This observation could indicate that the modification induced by IL-6 is of the “on-off” type and that TFIIS is able to activate this switch without a further capacity to increase CDK9 recruitment.

The ubiquitination of CDK9 was shown to have both activating and inhibitory effects depending on the transcriptional context. It can facilitate the formation of a ternary complex, comprising P-TEFb, the Tat activator, and the TAR RNA motif on the HIV-1 LTR (Barboric et al., 2005), but it can also have a negative impact on transcription activated by CIITA (major histocompatibility complex class II transactivator) (Kiernan et al., 2001). Since an association between CDK9, TFIIS, and UBR5 was observed in the soluble fraction of cell extracts (Fig. 1), and since the ubiquitination of CDK9 is dependent upon these two proteins (Fig. 4), it is tempting to speculate that this modification promotes CDK9 recruitment to the chromatin template. However, the knockdown of UBR5 in cells overexpressing TFIIS does not lead to a significant decrease of CDK9 occupancy at the IL6_RE1 (Fig. 6E), which rather suggests that neither the association with UBR5 nor the ubiquitination are essential for the CDK9 recruitment step. Notwithstanding, a significant decrease of CDK9 occupancy upon UBR5 knockdown is observed on, and downstream of, the promoter, which is apparently not caused by a reduced recruitment on the upstream regions (Fig. 6E). In agreement with these results, ChIP-re-ChIP experiments show that TFIIS, UBR5 and CDK9 belong to the same protein complex and co-localize on the promoter but not on the upstream IL6_RE1 (Fig. 6H). Of note, the effect of UBR5 on CDK9 occupancy is lost further downstream of the promoter, despite the fact that UBR5 remains physically associated with chromatin all along the gene. Together, these observations suggest that UBR5 is important for the transcriptional stages following initiation, such as the transition from initiation to early elongation. We speculate that the recruitment of

UBR5 potentiated by TFIIS, and the subsequent modifications induced to CDK9, favor the stability of the CDK9-RNAPII elongating complex, and thereby promote RNAPII phosphorylation on Ser2 of the CTD of its Rpb1 subunit. This model is supported by the observation that TFIIS increases the loading along the promoter and downstream transcribed regions not only of total RNAPII and CDK9, but also of phospho-Ser2 CTD RNAPII. Further investigations should elucidate whether the ubiquitination of CDK9 by UBR5 on the chromatin is required for this transcriptional effect.

CDK9 is widely implicated in regulating the elongation phase of transcription by RNAPII and plays a key role in a multitude of biological processes, such as cell differentiation, proliferation, immune response, and inflammation. Its activity is tightly regulated by mechanisms involving (i) a dynamic release from and capture by inhibitory (7SK/HEXIM) and activating (phosphorylated transcription factors) complexes; (ii) regulated expression of its two isoforms; and (iii) posttranslational modifications, including phosphorylation and ubiquitination (Brasier, 2008; Peterlin and Price, 2006; Wang and Fischer, 2008). Taken together, our results depict a new mechanism of CDK9 regulation by the general transcription factor TFIIS. We show that this mechanism involves the ubiquitination of CDK9, most likely in the soluble cellular fraction, through the recruitment of the E3 ubiquitin ligase UBR5. We propose that TFIIS coordinates the initiation of transcription with an increased stability of the early CDK9(P-TEFb)-RNAPII transcribing complexes, which is dependent upon UBR5. Further investigation should elucidate whether the function of TFIIS in the regulation of CDK9 is general or, rather, specific to cytokine inducible genes.

REFERENCES

1. Sekimizu, K., Kobayashi, N., Mizuno, D., and Natori, S. (1976) *Biochemistry* **15**, 5064-5070
2. Kim, B., Nesvizhskii, A. I., Rani, P. G., Hahn, S., Aebersold, R., and Ranish, J. A. (2007) *Proc. Natl. Acad. Sci. U. S. A* **104**, 16068-16073
3. Adelman, K., Marr, M. T., Werner, J., Saunders, A., Ni, Z. Y., Andrulis, E. D., and Lis, J. T. (2005) *Molecular Cell* **17**, 103-112
4. Awrey, D. E., Shimasaki, N., Koth, C., Weilbaecher, R., Olmsted, V., Kazanis, S., Shan, X., Arellano, J., Arrowsmith, C. H., Kane, C. M., and Edwards, A. M. (1998) *J. Biol Chem.* **273**, 22595-22605
5. Wery, M., Shematorova, E., Van Driessche, B., Vandenhoute, J., Thuriaux, P., and Van, M., V (2004) *EMBO J.* **23**, 4232-4242
6. Kettenberger, H., Armache, K. J., and Cramer, P. (2003) *Cell* **114**, 347-357
7. Nickels, B. E. and Hochschild, A. (2004) *Cell* **118**, 281-284
8. Cojocar, M., Jeronimo, C., Forget, D., Bouchard, A., Bergeron, D., Cote, P., Poirier, G. G., Greenblatt, J., and Coulombe, B. (2008) *Biochem. J.* **409**, 139-147
9. Prather, D. M., Larschan, E., and Winston, F. (2005) *Mol Cell Biol* **25**, 2650-2659
10. Ito, T., Arimitsu, N., Takeuchi, M., Kawamura, N., Nagata, M., Saso, K., Akimitsu, N., Hamamoto, H., Natori, S., Miyajima, A., and Sekimizu, K. (2006) *Mol Cell Biol* **26**, 3194-3203
11. Saunders, A., Core, L. J., and Lis, J. T. (2006) *Nat. Rev. Mol Cell Biol* **7**, 557-567
12. Margaritis, T. and Holstege, F. C. (2008) *Cell* **133**, 581-584
13. Ni, Z., Saunders, A., Fuda, N. J., Yao, J., Suarez, J. R., Webb, W. W., and Lis, J. T. (2008) *Mol Cell Biol* **28**, 1161-1170
14. Fujinaga, K., Irwin, D., Huang, Y., Taube, R., Kurosu, T., and Peterlin, B. M. (2004) *Mol. Cell Biol.* **24**, 787-795

15. Yamada, T., Yamaguchi, Y., Inukai, N., Okamoto, S., Mura, T., and Handa, H. (2006) *Mol Cell* **21**, 227-237
16. Peterlin, B. M. and Price, D. H. (2006) *Mol. Cell* **23**, 297-305
17. Ping, Y. H. and Rana, T. M. (2001) *J. Biol. Chem.* **276**, 12951-12958
18. Egloff, S. and Murphy, S. (2008) *Trends Genet.* **24**, 280-288
19. Palangat, M., Renner, D. B., Price, D. H., and Landick, R. (2005) *Proc. Natl. Acad. Sci. U. S. A*
20. Jeronimo, C., Forget, D., Bouchard, A., Li, Q., Chua, G., Poitras, C., Therien, C., Bergeron, D., Bourassa, S., Greenblatt, J., Chabot, B., Poirier, G. G., Hughes, T. R., Blanchette, M., Price, D. H., and Coulombe, B. (2007) *Mol Cell* **27**, 262-274
21. Cloutier, P., Al Khoury, R., Lavalleye-Adam, M., Faubert, D., Jiang, H., Poitras, C., Bouchard, A., Forget, D., Blanchette, M., and Coulombe, B. (2009) *Methods* **48**, 381-386
22. Umehara, T., Kida, S., Yamamoto, T., and Horikoshi, M. (1995) *Gene* **167**, 297-302
23. Callaghan, M. J., Russell, A. J., Woollatt, E., Sutherland, G. R., Sutherland, R. L., and Watts, C. K. (1998) *Oncogene* **17**, 3479-3491
24. Kiernan, R. E., Emiliani, S., Nakayama, K., Castro, A., Labbe, J. C., Lorca, T., Nakayama, K. K., and Benkirane, M. (2001) *Mol Cell Biol* **21**, 7956-7970
25. Barboric, M., Zhang, F., Besenicar, M., Plemenitas, A., and Peterlin, B. M. (2005) *J. Virol.* **79**, 11135-11141
26. Garriga, J., Bhattacharya, S., Calbo, J., Marshall, R. M., Truongcao, M., Haines, D. S., and Grana, X. (2003) *Mol Cell Biol* **23**, 5165-5173
27. Gagnon, D., Joubert, S., Senechal, H., Fradet-Turcotte, A., Torre, S., and Archambault, J. (2009) *J. Virol.* **83**, 4127-4139
28. Hou, T., Ray, S., and Brasier, A. R. (2007) *J. Biol Chem.* **282**, 37091-37102
29. Rahl, P. B., Lin, C. Y., Seila, A. C., Flynn, R. A., McCuine, S., Burge, C. B., Sharp, P. A., and Young, R. A. (2010) *Cell* **141**, 432-445

30. Henderson, M. J., Russell, A. J., Hird, S., Munoz, M., Clancy, J. L., Lehrbach, G. M., Calanni, S. T., Jans, D. A., Sutherland, R. L., and Watts, C. K. (2002) *J. Biol Chem.* **277**, 26468-26478
31. Hymowitz, S. G. and Wertz, I. E. (2010) *Nat. Rev. Cancer* **10**, 332-341
32. Henderson, M. J., Munoz, M. A., Saunders, D. N., Clancy, J. L., Russell, A. J., Williams, B., Pappin, D., Khanna, K. K., Jackson, S. P., Sutherland, R. L., and Watts, C. K. (2006) *J. Biol Chem.* **281**, 39990-40000
33. Mansfield, E., Hersperger, E., Biggs, J., and Shearn, A. (1994) *Dev. Biol* **165**, 507-526
34. Clancy, J. L., Henderson, M. J., Russell, A. J., Anderson, D. W., Bova, R. J., Campbell, I. G., Choong, D. Y., Macdonald, G. A., Mann, G. J., Nolan, T., Brady, G., Olopade, O. I., Woollatt, E., Davies, M. J., Segara, D., Hacker, N. F., Henshall, S. M., Sutherland, R. L., and Watts, C. K. (2003) *Oncogene* **22**, 5070-5081
35. Chin, S. F., Teschendorff, A. E., Marioni, J. C., Wang, Y., Barbosa-Morais, N. L., Thorne, N. P., Costa, J. L., Pinder, S. E., van de Wiel, M. A., Green, A. R., Ellis, I. O., Porter, P. L., Tavare, S., Brenton, J. D., Ylstra, B., and Caldas, C. (2007) *Genome Biol.* **8**, R215
36. Lindstrom, D. L. and Hartzog, G. A. (2001) *Genetics* **159**, 487-497
37. Cheng, B. and Price, D. H. (2007) *J. Biol Chem.* **282**, 21901-21912
38. Lacy, E. R., Filippov, I., Lewis, W. S., Otieno, S., Xiao, L., Weiss, S., Hengst, L., and Kriwacki, R. W. (2004) *Nat. Struct. Mol Biol* **11**, 358-364
39. Newton, K., Matsumoto, M. L., Wertz, I. E., Kirkpatrick, D. S., Lill, J. R., Tan, J., Dugger, D., Gordon, N., Sidhu, S. S., Fellouse, F. A., Komuves, L., French, D. M., Ferrando, R. E., Lam, C., Compaan, D., Yu, C., Bosanac, I., Hymowitz, S. G., Kelley, R. F., and Dixit, V. M. (2008) *Cell* **134**, 668-678
40. Chao, S. H. and Price, D. H. (2001) *J. Biol. Chem.* **276**, 31793-31799
41. Shim, E. Y., Walker, A. K., Shi, Y., and Blackwell, T. K. (2002) *Genes Dev.* **16**, 2135-2146
42. Brasier, A. R. (2008) *Cell Cycle* **7**, 2661-2666
43. Heinrich, P. C., Behrmann, I., Muller-Newen, G., Schaper, F., and Graeve, L. (1998) *Biochem. J.* **334 (Pt 2)**, 297-314

44. Wang, S. and Fischer, P. M. (2008) *Trends Pharmacol. Sci.* **29**, 302-313
45. 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) *Mol Cell Biol* **24**, 7043-7058
46. Zeghouf, M., Li, J., Butland, G., Borkowska, A., Canadien, V., Richards, D., Beattie, B., Emili, A., and Greenblatt, J. F. (2004) *J. Proteome. Res.* **3**, 463-468
47. Rigaut, G., Shevchenko, A., Rutz, B., Wilm, M., Mann, M., and Seraphin, B. (1999) *Nat. Biotechnol.* **17**, 1030-1032
48. Nowak, D. E., Tian, B., and Brasier, A. R. (2005) *Biotechniques* **39**, 715-725

FOOTNOTES

*We are grateful to the members of our laboratory and Jacques Archambault for helpful discussions and comments. We thank Denis Faubert (IRCM Proteomics Discovery Platform) for mass spectrometry analysis. This work is supported by grants from the Canadian Institutes for Health Research (CIHR), Genome Canada, Génome Québec, the Natural Sciences and Engineering Research Council of Canada (NSERC), and the Canadian Foundation for Innovation (CFI). MC was awarded studentships from the Canadian Institutes of Health Research (CIHR) and the Fonds de recherche en santé du Québec (FRSQ).

FIGURE LEGENDS

Fig. 1. TFIIS interacts with the E3 ligase UBR5 and the P-TEFb kinase CDK9. (A) High confidence interaction network for human TFIIS. The graph shows part of a network of high-confidence interactions obtained using affinity purification (*in vivo* pull-down) coupled with mass spectrometry (AP-MS), with 77 TAP-tagged baits targeting many transcription and RNA processing factors (see Cloutier et al., 2009). (B) *In vitro* pull-down using TFIIS-GST. *In vitro* pull-down experiments used HEK

293 whole-cell extracts with TFIIS-GST bound to glutathione-sepharose beads or with control glutathione-sepharose beads. The eluates and input (5%) were immunoblotted using antibodies directed against UBR5, CDK9, or the control protein CAND1. (C, D) Immunoprecipitation using anti-CDK9 (C) and anti-UBR5 (D) antibodies. Immunoprecipitation experiments used HEK 293 whole-cell extracts with an antibody raised against CDK9 (C) or UBR5 (D) bound to protein A beads or with mock protein A beads. The eluates and input (5%) were run on SDS gels and immunoblotted using antibodies against UBR5, TFIIS, CDK9, and Rpb1.

Fig. 2. UBR5 ubiquitinates CDK9 *in vitro*. Ubiquitination of baculovirus-expressed CDK9 carrying a His-tag was obtained in the presence of the enzymes E1, E2 (UbcH5b), and E3 (affinity purified UBR5). The reactions were incubated with ubiquitin, ubiquitin aldehyde, and AMP-PNP, as described in Experimental Procedures. Higher molecular weight ubiquitinated forms of CDK9-His were detected by immunoblotting with anti-His or anti-CDK9 antibodies (left panel). To confirm that the higher molecular weight bands contained ubiquitinated forms of CDK9, replicate reactions were analyzed by immunoblotting with an anti-ubiquitin antibody (right panel).

Fig. 3. TFIIS and UBR5 overexpression affect the ubiquitination of CDK9 *in vivo*. (A) HeLa cells transfected with a vector driving the expression of FLAG-tagged TFIIS or the empty vector, together with a construct driving the expression of ubiquitin-HA, were treated with the proteasome inhibitor lactacystin and lysed. The whole-cell extract (5% input) (left panel) or eluates from anti-CDK9 immunoprecipitation with the C20 antibody (right panel) were run on SDS gels and immunoblotted with anti-ubiquitin or anti-FLAG antibodies (B) Immuno-detection of CDK9 in the anti-CDK9 immunoprecipitate obtained as described in (A). (C) HeLa cells treated with lactacystin or mock-treated were lysed. The lysates were immunoprecipitated using anti-CDK9 antibody (C20) bound to protein A beads or with mock protein A beads. The eluates and the input (5%) were run on SDS gels and immunoblotted using anti-ubiquitin and anti-CDK9 antibodies. (D) HepG2 cells

stably transfected with constructs driving the expression of UBR5-FLAG or TFIIS-FLAG were subjected to *in vivo* ubiquitination as described above. Parental HepG2 cells treated in parallel were used as a control. The eluates of CDK9 immunoprecipitation (C20) (right panel) and the input (5%) (left panel) were immunoblotted using anti-ubiquitin, anti-CDK9, anti-FLAG, and anti-UBR5 antibodies.

Fig. 4. TFIIS or UBR5 knockdown decreases the ubiquitination of CDK9 *in vivo*, and UBR5 is required for the effect of TFIIS on this ubiquitination event. (A) HeLa cells were transfected with the indicated concentrations of siRNAs targeting TFIIS or control non-targeting siRNAs, or mock transfected. Cells were treated with lactacystin, lysed, and the lysate was used in immunoprecipitation experiments with an anti-CDK9 antibody (C20). The eluates were run on SDS gels and immunoblotted with anti-ubiquitin or anti-CDK9 antibodies. (B) Comparison of knockdown efficiencies resulting from treatment with siRNAs targeting TFIIS or control non-targeting siRNAs. (C) HeLa cells treated with the indicated concentrations of siRNAs targeting UBR5 or non-targeting siRNAs were assayed for *in vivo* ubiquitination (as in A). (D) Comparison of knockdown efficiencies resulting from treatment with siRNAs targeting UBR5 or control non-targeting siRNAs. (E) HeLa cells treated with siRNAs targeting UBR5, non-targeting siRNAs (50nM), or mock-treated were transfected with a construct driving the expression of TFIIS-FLAG, or with an empty vector. An *in vivo* ubiquitination assay (as in A) was performed with the transfected cells. The lysates were used for anti-CDK9 immunoprecipitation, and the eluates (left panel) or the input (5%) (right panel) were run on SDS gels and immunoblotted with anti-ubiquitin or anti-CDK9 antibodies. The expression of the FLAG-tagged protein was monitored by immunodetection with an anti-FLAG antibody.

Fig. 5. UBR5 knockdown does not affect the stability of CDK9. HeLa cells were co-transfected with constructs driving the expression of Renilla luciferase (RLuc) fused to the N-terminal region of CDK9, and with Firefly luciferase (FLuc) as an

internal normalization control, together with siRNAs targeting UBR5 or control non-targeting siRNAs. Cycloheximide was added to culture medium for 0, 2, and 4 h to block overall protein synthesis, and luciferase activities were measured (upper panel). Parallel control experiments using the Rluc construct alone were conducted (lower panel). Each value represents the average of 4 assays, including 2 independent experiments.

Fig. 6. TFIIS, CDK9, and UBR5 are recruited to the γ FBG gene, and TFIIS increases the occupancy of this gene by CDK9 in a UBR5-dependent manner. (A) Schematic representation of the γ FBG gene and the position of primers used to amplify specific regions (IL6_RE1, TATA box, exon 5 and exon 7). (B) Parental or TFIIS-FLAG overexpressing HepG2 cells were subjected to a two-step ChIP assay prior to or following treatment with IL-6. After cross-linking, cell lysis and sonication, the nucleoprotein complexes were immunoprecipitated with antibodies raised against CDK9 (C20). The occupancy levels were evaluated by quantitative PCR with primers along the γ FBG gene, as indicated in (A). (C) TFIIS-FLAG-overexpressing HepG2 cells were used in ChIP experiments (as in (B)), prior to or following treatment with IL-6, with an anti-FLAG (M2) antibody at the immunoprecipitation step. (D) UBR5-FLAG-overexpressing HepG2 cells were used in ChIP experiments (as in (B)), prior to or following treatment with IL-6, with the anti-FLAG (M2) antibody at the immunoprecipitation step. (E) TFIIS-FLAG-overexpressing HepG2 cells were treated with either siRNAs targeting UBR5 or control non-targeting siRNAs. Cells treated with IL-6 were used in ChIP experiments with an anti-CDK9 antibody (C20) (upper panel). The knockdown efficiency is shown (lower panel). (F) The N20 antibody that recognizes all forms of RNAPII (total) was used in ChIP experiments (as in (B)), prior to or following treatment with IL-6. (G) An antibody that specifically recognizes the RNAPII phospho-Ser2-CTD form (H5) was used for the ChIP experiments. (H) Cells overexpressing TFIIS-FLAG, UBR5-FLAG or parental HepG2 were treated with IL-6 and subjected to the ChIP-re-ChIP procedure with anti-CDK9 followed by anti-FLAG antibodies. The enrichment values for the IL6_RE1 and TATA box

regions are shown. Data are presented as means \pm SD of at least three values including two independent experiments. Student *t* tests were performed, and the levels of significance are indicated: *, *p* value <0.05; **, *p* value <0.01.

Figure 1

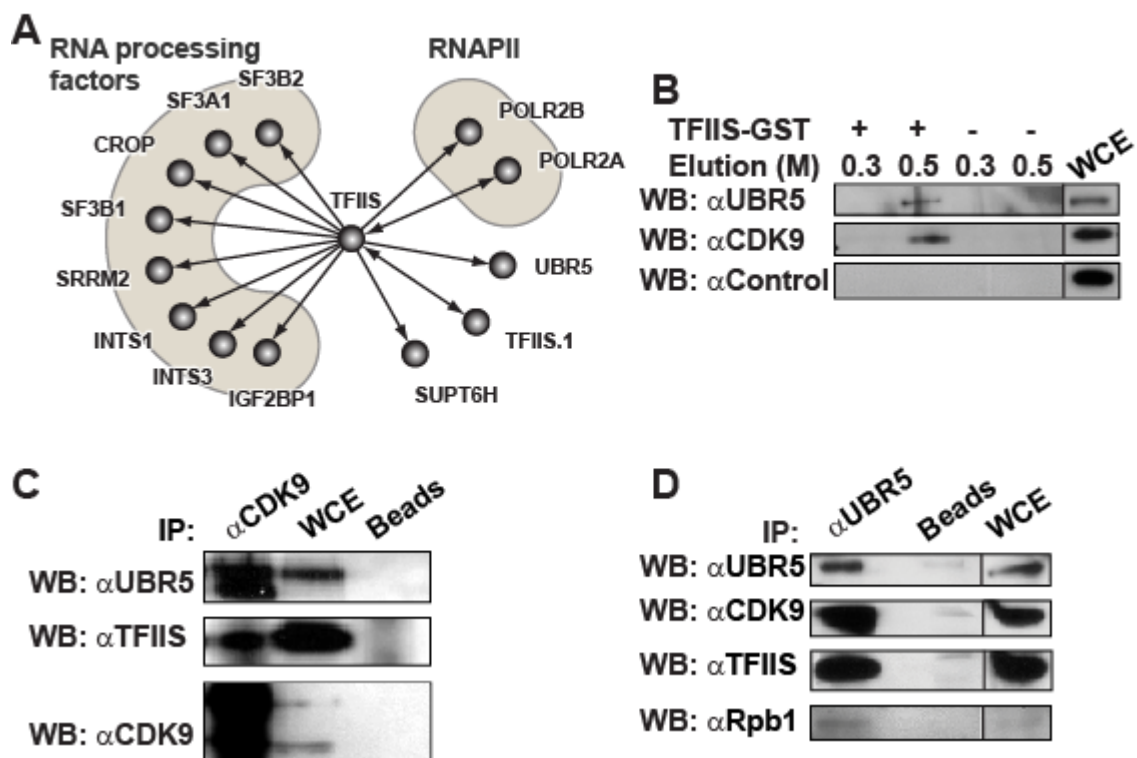


Figure 2

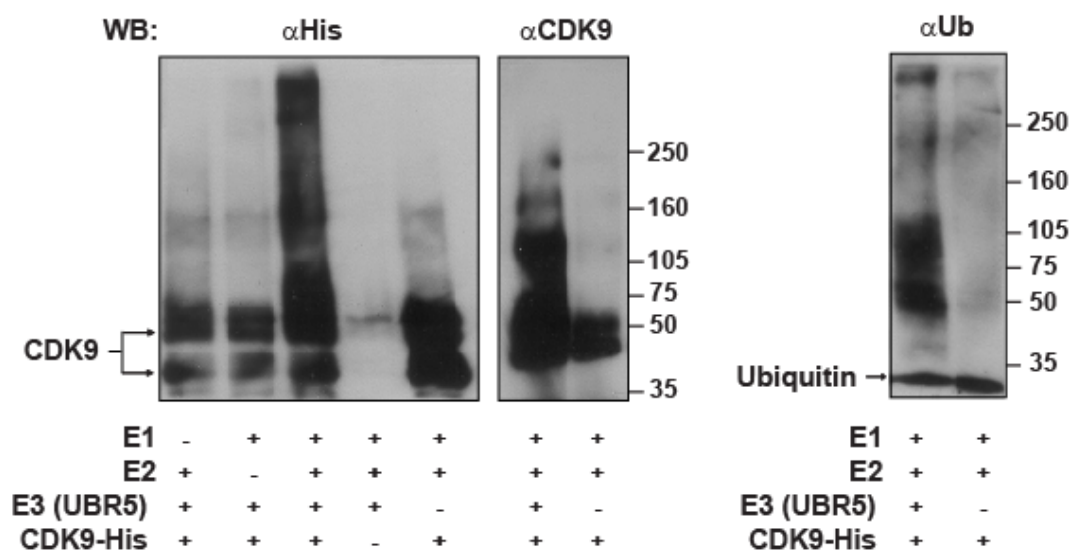


Figure 3

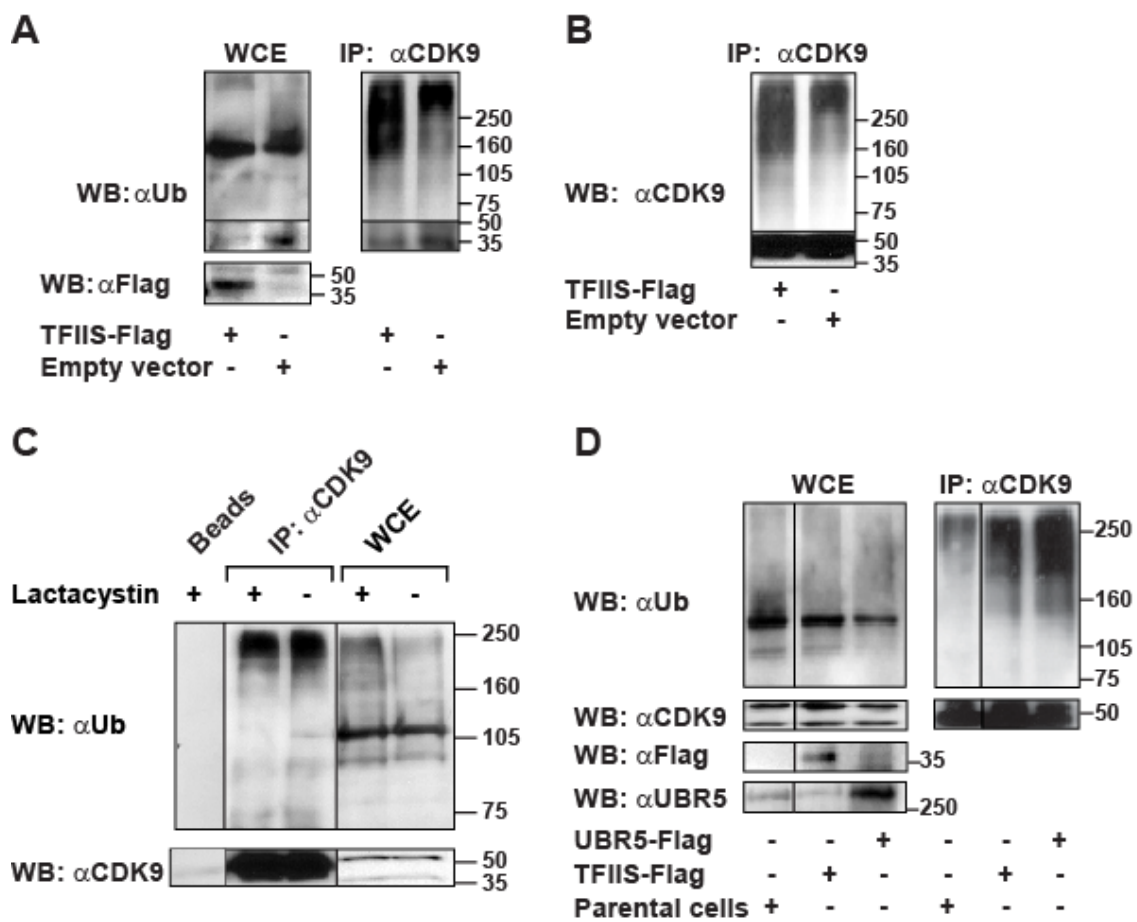


Figure 4

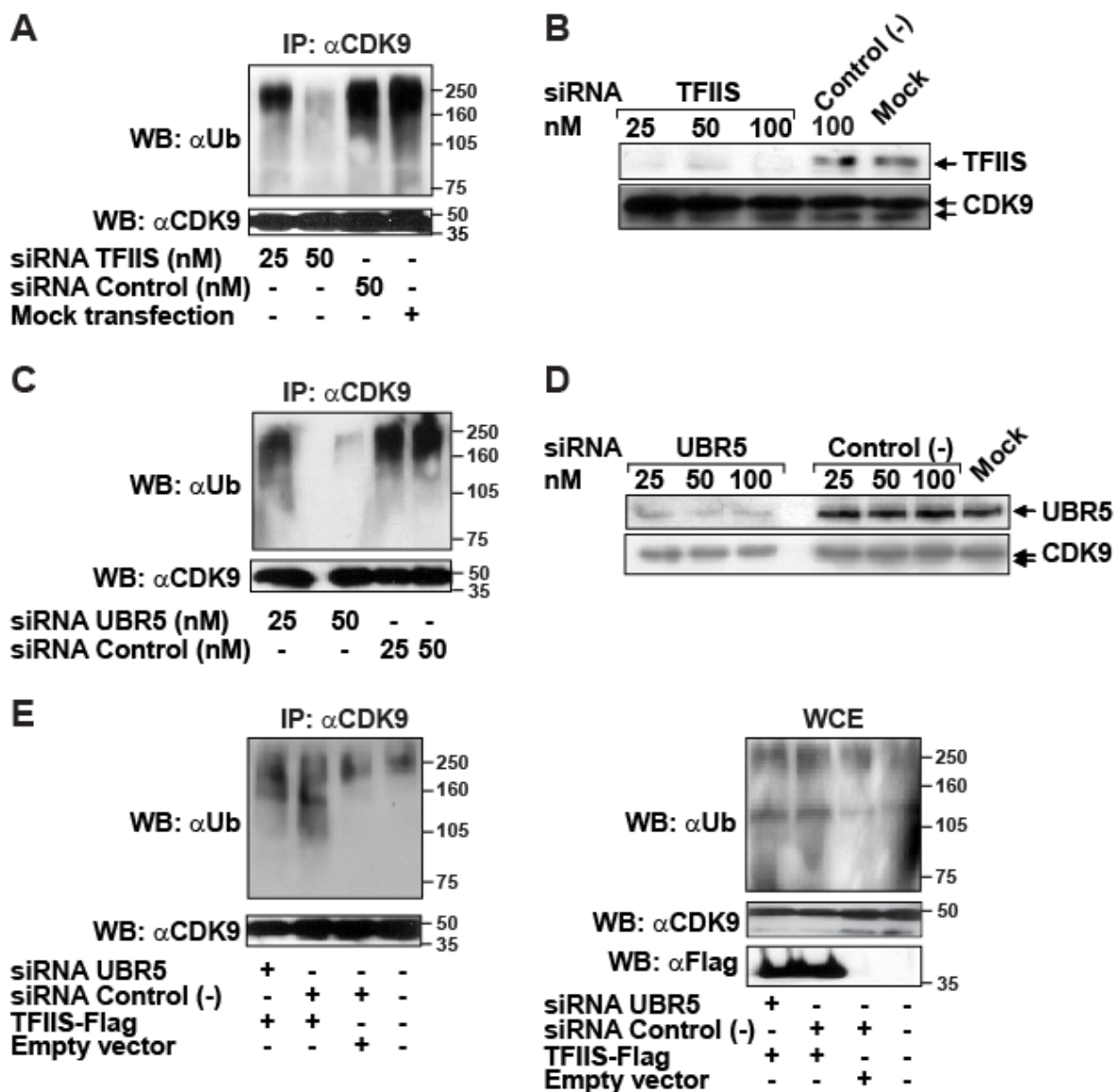


Figure 5

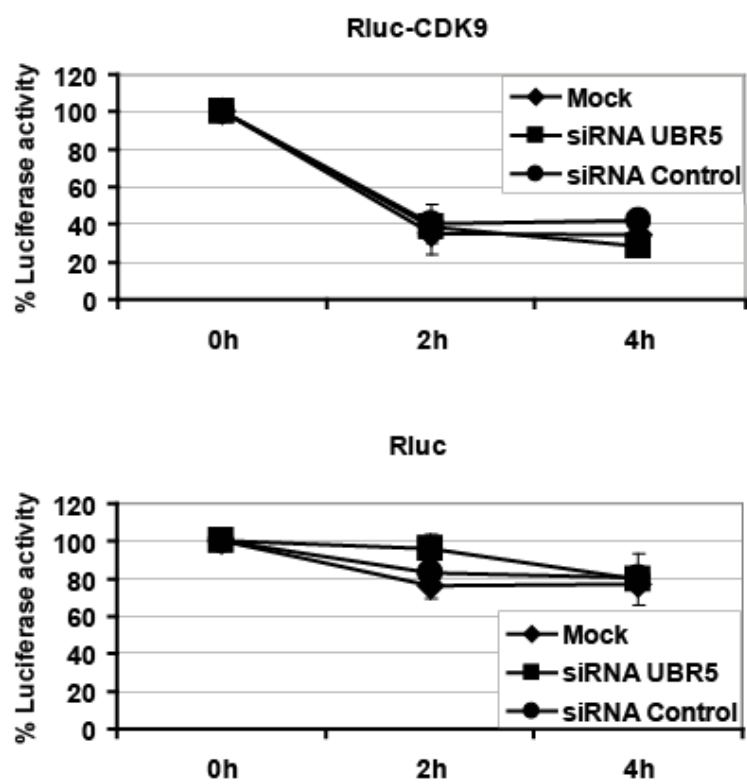


Figure 6

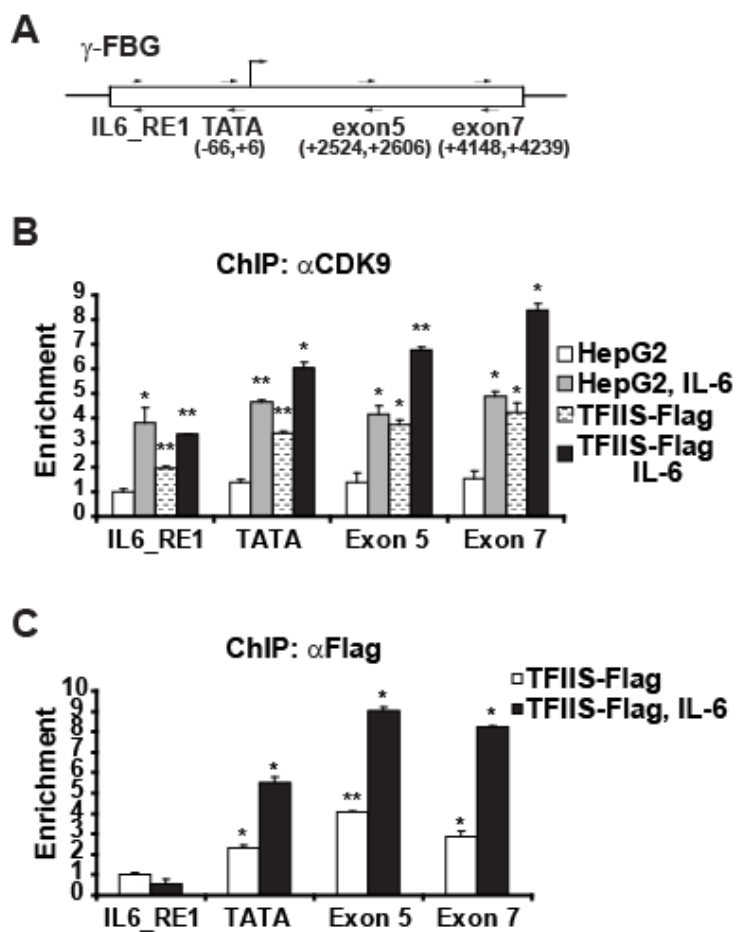
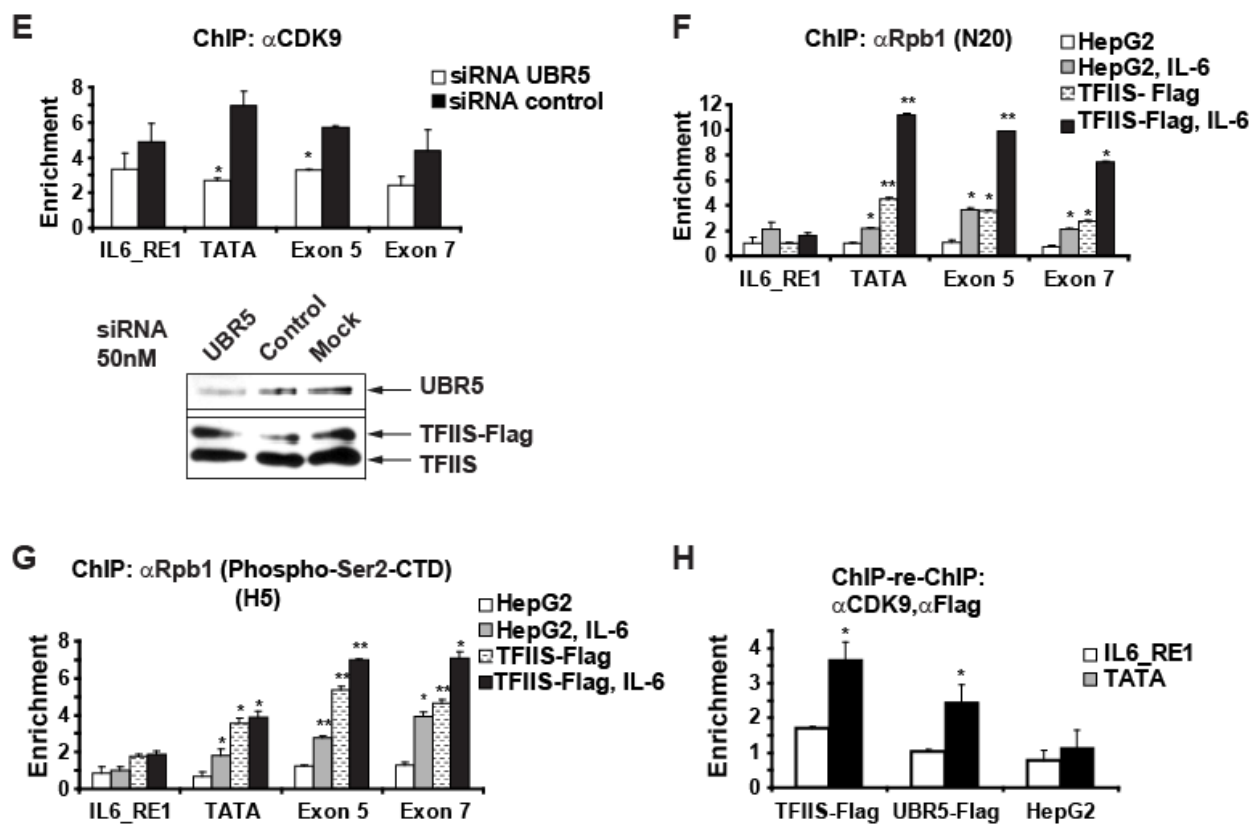


Figure 6 - Continued



CHAPITRE 5

DISCUSSION

Une discussion des résultats ayant déjà été présentée dans les articles, je développerai dans cette section les aspects qui n'ont été abordés que sommairement et je discuterai également des points qui n'ont pas été présentés. Je viserai de préciser la contribution amenée à l'avancement des connaissances dans le domaine de la transcription par l'ARN Pol II.

5.1 Localisation génomique de la machinerie transcriptionnelle de l'ARN Pol II

Tel que présenté dans le premier chapitre, un nombre croissant d'indices appuient la notion que les GTFs ne constituent pas un groupe constant qui s'associent à l'ARN Pol II de façon universelle lors de la transcription (Buratowski, 2009; Muller et al., 2007). Il a été proposé qu'ils soient impliqués dans la régulation de la transcription et que des complexes de composition variable en GTFs reconnaissent des séquences spécifiques aux gènes. De plus, plusieurs étapes de la transcription sont régulées tel que suggéré par des études de la distribution de la polymérase le long des gènes : la régulation au niveau du recrutement du PIC serait reflétée par l'absence de la polymérase ou une distribution uniforme de celle-ci sur les gènes, alors qu'une régulation des étapes qui suivent le recrutement se traduiraient par une accumulation de la polymérase sur le gène (revu par (Fuda et al., 2009)). Les indices sur le rôle des GTFs dans la régulation de ces étapes de la transcription *in vivo* restent très limités et les fonctions qui leur sont attribuées sont encore largement basées sur des études *in vitro*. Dans ce contexte, le premier objectif de recherche a été d'étudier systématiquement la localisation génomique des GTFs humains de pour permettre de distinguer entre leur implication dans les phases précoces de la transcription (par

exemple le recrutement du PIC, le dégagement du promoteur) et leur implication dans l'élongation.

L'approche choisie a été d'utiliser un système d'expression inductible des facteurs de transcription portant une étiquette TAP qui permet d'un coté une expression à des niveaux quasi-physiologiques et, d'un autre coté, d'utiliser une procédure d'immunoprécipitation standardisée avec des anticorps reconnaissant l'étiquette.

Une analyse systématique de la localisation génomique d'un grand nombre de protéines est dépendante d'un nombre de facteurs d'ordre biologique mais aussi technique. Parmi ces facteurs, on peut énumérer une interaction transitoire de certaines protéines avec l'ADN, une capacité de détection des interactions dépendante des séquences des protéines et d'ADN impliquées, ainsi que des complexes protéiques spécifiques formés sur différentes séquences génomiques différentes ou dans différents types cellulaires (Aleksic and Russell, 2009; Aparicio et al., 2004; Collas, 2010; Gilchrist et al., 2009). Des limitations techniques, telles que la disponibilité d'anticorps spécifiques et efficaces pour reconnaître chacune des protéines liée à l'ADN, ne sont pas négligeables dans le contexte d'une analyse de ce type. Prenant en considération ces facteurs, une analyse systématique par immunoprécipitation de la chromatine ne peut pas se réduire à une standardisation de chacune de ses étapes, mais implique aussi que des conditions optimales doivent être trouvées pour chaque protéine, tout en limitant les sources de variation. Dans notre étude, le paramètre qui nous a permis d'optimiser le ratio signal-bruit de fond a été le temps de pontage par le formaldéhyde, un temps variable étant utilisé pour les différents polypeptides. Il est connu que des facteurs tels que la composition en acides aminés ou la distance spécifique entre l'ADN et le polypeptide étudié dans le contexte des complexes transcriptionnels ont un impact majeur sur l'efficacité de pontage (Gilchrist et al., 2006; Nagy et al., 2003; Siomin et al., 1973).

Les contrôles les plus largement utilisés sont l'utilisation d'un contrôle sans anticorps ou d'un anticorps contrôle préférentiellement du même isotype pour l'étape

d'immunoprécipitation, ainsi que l'analyse de l'enrichissement des régions génomiques d'intérêt versus des régions génomiques non-ciblées (Collas, 2010).

Dans notre étude, nous avons utilisé plusieurs contrôles, pour les raisons suivantes :

- une analyse parallèle de la chromatine obtenue des cellules parentales, qui n'expriment donc pas de facteur portant une étiquette TAP, en utilisant pour l'immunoprécipitation le même anticorps, minimise les faux positifs qui pourraient résulter d'une immunoprécipitation non-spécifique;
- l'analyse de deux sous-unités des facteurs à plusieurs sous-unités permet d'assurer que les signaux observés sont spécifiques;
- l'utilisation de contrôles internes pour chaque expérience permet d'évaluer si une protéine lie spécifiquement les régions d'intérêt. Trois contrôles internes différents ont été utilisés dans nos analyses par ChIP : des séquences situées en amont du promoteur analysé, une région génomique connue comme étant dépourvue de gènes située dans la bande 13q21.33 (Nobrega et al., 2003) et un EST (*Expressed Sequence Tag*) sur le chromosome 17 (chr17:49926700–49927999).

De plus, l'analyse par ChIP des mutants de différents éléments structuraux de l'ARN Pol II prouve que notre approche est efficace, d'un côté, pour détecter des interactions protéine-ADN réelles et, d'un autre côté, pour faire la distinction entre les étapes précoces et tardives de la transcription. Tel que présenté dans le Chapitre 3, les différents mutants de la sous-unité Rpb2 de la polymérase ont été caractérisés en détail et leur comportement dans des essais de transcription *in vitro* concorde parfaitement avec ce qu'on a observé dans l'analyse par ChIP. Ainsi, le mutant de la boucle d'embranchement 1 ayant une délétion des acides aminés 458-459 (Rpb2 Fork 1 Δ 458-459), qui s'assemble dans le complexe de préinitiation tel que démontré par des essais de retard de mobilité sur gel de type EMSA (*Electromobility Shift Assays*) mais qui n'est pas actif dans des tests d'élongation *in vitro*, est détecté uniquement sur les promoteurs mais pas en aval, dans les régions transcrites de deux gènes analysés par ChIP. À l'opposé, la sous-unité Rpb2 sauvage est détectée quant à elle le long des deux gènes, ce qui prouve que l'absence de

signal du mutant dans les régions transcrites n'est pas due à une limitation de détection qui pourrait être reliée à une faible densité des molécules de polymérase en élongation. De plus, le mutant ayant une triple mutation de la boucle switch 3 (Rpb2 sw3-1078) qui ne s'assemble pas dans le complexe de préinitiation dans des tests EMSA et qui est également inactif pour transcrire *in vitro*, n'est pas détecté sur aucune région des gènes testés par CHIP. Dans leur ensemble, les essais de transcription *in vitro* des mutants de la polymérase viennent à l'appui des profils de localisation observés *in vivo*.

Le système d'expression inductible des facteurs portant une étiquette TAP nous a permis de contrôler les niveaux d'expression afin d'éviter des interactions protéine-protéine et protéine-ADN non-physiologiques. Cependant, ceci implique que dans la population de cellules analysées par CHIP, les séquences d'ADN peuvent avoir été liées par la protéine endogène ou celle exprimée de façon ectopique. Par conséquent, l'absence du signal lorsqu'on analyse la protéine étiquetée ne peut pas exclure que la protéine endogène lie la séquence analysée. Un test pilote comparant l'anticorps anti-tag avec des anticorps spécifiques pour la sous-unité Rpb11 de la polymérase ainsi que pour le facteur TFIIS a montré que dans notre système d'analyse, les patrons de localisation génomique des protéines endogènes et celles exprimées de façon ectopique sont similaires (données non montrées). De plus, une bonne corrélation entre les profils de localisation génomique de deux sous-unités différentes d'un même facteur exprimées de façon ectopique (Chapitre 2, Fig. 2 et Fig. supplémentaire 3) renforcent nos conclusions et suggèrent que les patrons observés reflètent bien le comportement des protéines endogènes.

Afin d'analyser la fonction *in vivo* des différents facteurs de la machinerie transcriptionnelle de l'ARN Pol II et de faire la distinction entre leur implication dans les étapes précoces ou tardives de la transcription, nous avons choisi d'analyser leur localisation génomique le long des gènes dont le promoteur a été bien caractérisé. Des études à l'échelle du génome en utilisant des séquences pleine longueur des ADNc afin de déterminer la position des sites d'initiation de la

transcription ont montré que la majorité des gènes de mammifères possèdent une région d'initiation de 50 à 100pb plutôt qu'un site unique d'initiation et que la plupart possèdent plusieurs régions de ce type situées à des distances variables (revu par (Davuluri et al., 2008; Sandelin et al., 2007). Des analyses phylogénétiques, ainsi qu'avec des gènes rapporteurs, ont indiqué l'utilisation de promoteurs alternatifs dans un pourcentage allant jusqu'à 50% de gènes chez l'humain et chez la souris (Baek et al., 2007; Kimura et al., 2006; Sun et al., 2006; Takeda et al., 2007). L'utilisation différentielle de ces promoteurs alternatifs dans des types cellulaires différents, ou dans un seul type cellulaire dans des conditions environnementales différentes, contribue à la diversification fonctionnelle des gènes. De plus, l'utilisation des promoteurs alternatifs situés dans les régions codantes des gènes est un mécanisme fréquemment utilisé pour produire des ARNm tronqués, non-codants (Kim et al., 2005b; Tsuritani et al., 2007). L'importance de ce phénomène est illustrée par des facteurs qui ont une fonction répressive sur l'utilisation de ces promoteurs alternatifs, comme par exemple Chd1, une enzyme de remodelage de la chromatine et le complexe Paf1 chez la levure (Chu et al., 2007; Quan and Hartzog, 2010).

Nous avons choisi d'analyser des gènes qui n'utilisent pas ce type de régulation mais qui ont plutôt une seule région d'initiation. Ceci assure que les résultats obtenus pour des régions de promoteurs alternatifs actifs dans la population de cellules analysées par ChIP ne soient pas attribués à des régions transcrites. Par ailleurs, l'identification de la polymérase, ainsi qu'au moins deux facteurs d'initiation sur chacun des promoteurs analysés, montre qu'il s'agit effectivement des promoteurs utilisés. Notre approche permet une distinction non-ambigüe entre le rôle dans les étapes précoces et tardives de la transcription. Une telle distinction dans le cas des gènes à promoteurs multiples nécessiterait comme point de départ une analyse qui permettrait de savoir lesquels des promoteurs alternatifs sont actifs. L'utilisation du ChIP-seq ou ChIP-on-CHIP tiling pour analyser des marqueurs de l'initiation comme des modifications de la chromatine et la forme de la polymérase ayant le CTD phosphorylé sur la sérine 5 pourraient répondre à cette question.

Notre approche nous a permis de mettre en évidence pour la première fois la fonction de la sous-unité Rpb7 dans l'élongation *in vivo* et cette observation a été confirmée par l'analyse de la sous-unité Rpb4, avec laquelle elle forme un hétérodimère dissociable de l'ARN Pol II. Plusieurs études publiées ultérieurement ont amené des indices qui appuient le rôle du hétérodimère Rpb4/Rpb7 dans l'élongation. Ainsi, une analyse à l'échelle du génome chez la levure a montré que le patron de localisation génomique de Rpb7 est hautement corrélé avec celui de la sous-unité Rpb3 qui fait partie du noyau de la polymérase (Jasiak et al., 2008). Un même type d'analyse qui comparait la localisation de Rpb4 avec celle de Rpb3 a mené à la même conclusion (Verma-Gaur et al., 2008). Ces études appuient notre conclusion que, *in vivo*, l'hétérodimère Rpb4/Rpb7 s'associe à la polymérase en élongation et qu'il n'est pas impliqué uniquement dans l'étape d'initiation, tel que proposé par des études *in vitro* (Edwards et al., 1991).

Sur la base du profil de localisation observé en condition de stress thermique, nous avons proposé que l'hétérodimère Rpb4/Rpb7 a pour rôle de stabiliser la polymérase en élongation. En accord avec cette hypothèse, il a été rapporté que chez la levure, des mutants ayant une délétion de Rpb4 présentent un défaut global de la transcription par l'ARN Pol II à des températures élevées (Maillet et al., 1999; Miyao et al., 2001). De plus, ces mutants présentent une sensibilité au 6-azauracil, phénotype spécifique aux défauts d'élongation, ainsi qu'une transcription défectueuse des gènes longs et riches en GC (Verma-Gaur et al., 2008). Si des études initiales *in vitro* ont montré que Rpb4/Rpb7 ont un rôle dans l'initiation, mais sont dispensables pour l'étape d'élongation (Edwards et al., 1991; Orlicky et al., 2001), il a été montré récemment, que les sous-unités F/E, les homologues de Rpb4/Rpb7 chez les *archaea*, stimulent l'élongation *in vitro* (Hirtreiter et al., 2010). L'ARN Pol II des *archaea* est considérée comme un bon modèle d'étude de l'ARN Pol II eucaryote. Dans ce système, les homologues de Rpb4/Rpb7 stimulent la vitesse de la polymérase en élongation et diminuent les pauses. L'interaction de l'hétérodimère Rpb4/Rpb7 avec l'ARN, mais aussi le changement de conformation

tel que la fermeture de la pince mobile induite à la polymérase ont été proposés par les auteurs pour expliquer ces activités. Nos résultats sont en accord avec ceux de ces études et montrent que la fonction de l'hétérodimère Rpb4/Rpb7 dans l'élongation *in vivo* est conservée chez l'humain.

La localisation de Rpb7 dans la région 3' du gène HSPA8 (Chapitre 2, Fig. 4) a soulevé la question de savoir si cette sous-unité a un rôle dans la formation des extrémités 3' des ARN, tel qu'il avait été suggéré également par des études chez la levure (Mitsuzawa et al., 2003; Runner et al., 2008). Rpb7 possède un domaine OB-fold de liaison aux acides nucléiques simple-brin par lequel il interagit avec l'ARN transcrit qui émerge de la polymérase (Chen et al., 2009b; Meka et al., 2005; Ujvari and Luse, 2006b). Il interagit aussi avec le facteur de terminaison Nrd1 qui a un rôle dans la formation des extrémités 3' des petits ARN nucléaires et nucléolaires (Mitsuzawa et al., 2003). De plus, des mutants *rpb4Δ* chez *S. cerevisiae* présentent une diminution des niveaux d'occupation des régions 3' des gènes qui codent pour des ARNm ou des snoARN par l'ARN Pol II, ainsi que par des facteurs impliqués dans la formation des extrémités 3' des ARN comme Rna14 et Rna15 (Runner et al., 2008).

Une analyse par Q-PCR effectuée tel que décrit dans le Chapitre 2 a montré que Rpb7, ainsi que la sous-unité Rpb11 du noyau de la polymérase occupent d'autres gènes, tel que RPL3, et ceci autant en conditions normales qu'après un choc thermique (Annexe I, Fig. 1A, page 232). Les deux sous-unités de la polymérase sont présentes dans la région 3' de ce gène. RPL3 code pour un ARNm et, de plus, des snoARN variés (U43, U82, U83a, U83b) sont codés à partir de ses introns (Annexe I, schéma situé en bas dans la figure 1A, page 232). Nous avons émis l'hypothèse que si Rpb7 joue un rôle dans la formation des extrémités 3' de ces ARNm ou snoARN, alors une déplétion de Rpb7 par des ARN d'interférence mènerait à des défauts dans ce processus, avec l'apparition des ARN plus longs, non-clivés. Une atténuation de l'expression efficace de Rpb7 a été effectuée (Annexe I, Fig. 1B, page 232) et les ARN totaux extraits de ces cellules ou de cellules

contrôles ont été testés par Northern blot avec des sondes qui reconnaissent les ARN de RPL3, U43 ainsi que U83A. La déplétion de Rpb7 n'a pas eu d'effet détectable dans la formation de ces ARN, puisque seulement des ARN de la même taille que ceux présents dans les extraits des cellules contrôles ont été formés (Fig. 1C). Ceci suggère que Rpb7 n'a pas un rôle essentiel dans le processus de formation des extrémités 3' de ces ARN. Il y a trois possibilités qui pourraient expliquer cette observation : (i) cette fonction n'a pas été conservée chez l'humain; (ii) Rpb7 joue un rôle redondant dans ce processus avec d'autres facteurs qui restent à identifier (iii) son rôle possible n'est pas général et l'étude d'autres ARN transcrits par l'ARN Pol II pourrait révéler une dépendance de cette sous-unité. La présence de Rpb7 dans la région 3' des gènes pourrait indiquer que l'hétérodimère Rpb4/Rpb7 a un rôle dans la stabilisation de la polymérase tel que suggéré par l'analyse en conditions de choc thermique (Fig. 4, Chapitre 2), sans qu'il ait une fonction spécifique dans la formation des extrémités 3' des ARN. Alternativement, ce profil de localisation pourrait être relié à son rôle dans l'export des ARN. Il a été montré chez la levure, que l'hétérodimère Rpb4/Rpb7 voyage entre le noyau et le cytoplasme et est impliqué dans la dégradation cytoplasmique des ARNm de façon dépendante de la transcription, puisque l'export du complexe Rpb4/Rpb7-ARN est dépendant de l'association correcte de l'hétérodimère à la polymérase en transcription (Goler-Baron et al., 2008; Selitrennik et al., 2006). De plus, la régulation de l'expression des snoARN se fait surtout au niveau post-transcriptionnel (Nabavi et al., 2008).

L'étude du facteur TFIIF par immunoprécipitation de la chromatine nous a permis de mettre en évidence pour la première fois sa localisation *in vivo* autant aux promoteurs que sur des régions transcrites situées en aval du promoteur. Son rôle dans l'initiation de la transcription autant *in vitro* que *in vivo* avait déjà été caractérisé, tel que décrit dans le Chapitre 1. De nombreuses études ont montré qu'il est impliqué dans des étapes distinctes de l'initiation comme la sélection du site d'initiation et le recrutement de la polymérase dans le PIC, ainsi que dans la formation du premier lien phosphodiester et le dégagement du promoteur (Chen et

al., 2007; Fairley et al., 2002; Freire-Picos et al., 2005; Khapersky et al., 2008; Yan et al., 1999). En accord avec ces études, notre analyse a montré que TFIIF s'associe aux promoteurs des gènes analysés. Nos données sont également en accord avec celles des études qui ont identifié *in vivo* l'association de TFIIF sur quelques promoteurs chez la levure (Krogan et al., 2002; Pokholok et al., 2002a). Ces deux dernières études n'ont cependant pas détecté TFIIF dans les régions transcrites de ces mêmes gènes. En contraste, notre étude montre que chez l'humain TFIIF s'associe à des régions transcrites en aval des promoteurs. Cette divergence pourrait avoir plusieurs explications : (i) TFIIF ne reste pas associé à la polymérase après l'initiation mais il s'associe aux sites de pause et se dissocie par la suite de la polymérase en élongation, tel qu'il a été observé dans des systèmes de transcription *in vitro* (Zawel et al., 1995); par conséquent, l'analyse d'un nombre trop restreint de régions transcrites pourrait ne pas détecter TFIIF; (ii) des particularités de la transcription chez les eucaryotes supérieurs pourraient nécessiter une association plus fréquente de TFIIF à la polymérase en élongation chez ces espèces en comparaison à la levure. Ainsi, il a été montré que la densité des nucléosomes sur les exons est plus élevée que sur les introns, ce qui ralentit la polymérase et favorise le recrutement co-transcriptionnel des facteurs d'épissage (de la Mata et al., 2003; Schwartz et al., 2009). Le rôle de TFIIF d'augmenter la vitesse de transcription *in vitro* est bien caractérisé. Des études *in vitro* de la cinétique de synthèse des transcrits par l'ARN Pol II ont montré que TFIIF a pour rôle de stimuler la translocation en aval et qu'il est requis pour que la formation du lien phosphodiester soit complétée (Xiong and Burton, 2007; Zhang et al., 2005a; Zhang and Burton, 2004). De plus, il stimule la transcription à travers des obstacles tels que des lésions oxydatives de l'ADN (Charlet-Berguerand et al., 2006). Nos résultats montrent que TFIIF s'associe à la polymérase en élongation et ceci pourrait indiquer qu'il a pour rôle d'augmenter la vitesse de transcription *in vivo* à des sites où la polymérase est ralentie par différents obstacles tels que des nucléosomes ou des lésions de l'ADN.

Nos résultats suggèrent que, tout comme les complexes d'initiation de la transcription, les complexes d'élongation n'ont pas une composition constante étant

donné que des facteurs comme TFIIF, TFIIIS ainsi que la sous-unité Rpb7 de la polymérase n'ont été détectés que sur une fraction des régions transcrites analysées. Ceci n'est pas dû à une limitation de détection étant donné que, dans la majorité des cas testés, l'analyse d'une deuxième sous-unité a confirmé les résultats obtenus avec la première sous-unité (Chapitre 2, Fig. suppl. 3 et données non-montrées). De plus, des profils de localisation spécifiques et différents de ceux de la polymérase elle-même sont observés dans les régions transcrites alors que des profils très similaires sont obtenus aux promoteurs (Chapitre 2, Fig. 4). D'autres études ont aussi proposé qu'un échange dynamique des facteurs d'élongation se produit le long des gènes et que ce phénomène assure une adaptation continue de la polymérase à l'environnement chromatinien ainsi qu'une réponse adéquate à différents stress ou à des signaux cellulaires (revu pas (Saunders et al., 2006; Svejstrup, 2007).

Dans son ensemble, notre étude a amené la première image globale sur l'implication spécifique des facteurs généraux de transcription humains associés à l'ARN Pol II dans les étapes précoces et tardives de la transcription *in vivo*. Cette étude a mis en évidence pour la première fois la fonction *in vivo* dans l'étape d'élongation des sous-unités Rpb4/Rpb7 de la polymérase, ainsi que du facteur TFIIF. Nos résultats sont en accord avec ceux des études récentes qui appuient l'hypothèse selon laquelle le contrôle de l'élongation est un mécanisme important de régulation de la transcription (Gilchrist et al., 2009; Margaritis and Holstege, 2008; Muse et al., 2007). Ils sont également en accord avec d'autres études sur des facteurs connus pour avoir des fonctions uniquement dans une étape de la transcription, montrant que ces facteurs sont en effet impliqués dans plusieurs étapes (Donner et al., 2010; Ghavi-Helm et al., 2008; Guglielmi et al., 2007; Kim et al., 2007; Zhu et al., 2006). Trois facteurs parmi ceux que nous avons étudiés, soit TFIIIS, Médiateur et TAF10 illustrent par leur patron de localisation cette multifonctionnalité des facteurs généraux de transcription associés à l'ARN Pol II.

5.2 Le facteur TFIIIS et ses nouveaux partenaires d'interaction.

Au cours des dernières années, plusieurs études ont montré que le facteur TFIIIS possède des fonctions autres que celle de stimuler l'ARN Pol II pendant l'étape d'élongation. Il aurait un rôle dans la formation du complexe de préinitiation, la stimulation de l'ARN polymérase arrêtée dans les régions proximales aux promoteurs après l'initiation, ainsi que dans la transcription par l'ARN Pol III (Adelman et al., 2005; Ghavi-Helm et al., 2008; Guglielmi et al., 2007; Prather et al., 2005). Nous avons raisonné qu'en identifiant ses partenaires d'interaction, d'autres fonctions pourraient être mises en évidence. Tel que décrit dans le Chapitre 1, des études cristallographiques, biochimiques et génétiques ont montré que le domaine N-terminal de TFIIIS n'est pas impliqué dans l'interaction avec la polymérase ou dans la stimulation de l'activité endonucléolytique de cette dernière, mais plutôt dans la formation des complexes de préinitiation (Booth et al., 2000; Kettenberger et al., 2003; Kim et al., 2007; Pan et al., 1997). Ce domaine est divergeant entre les espèces et il a été proposé qu'il ait servi à l'acquisition de nouvelles fonctions à travers l'évolution.

Notre analyse effectuée par chromatographie d'affinité du facteur TFIIIS portant une étiquette TAP a identifié des nouveaux partenaires d'interaction tels que TFIIIS.1 et UBR5, en plus des partenaires attendus tels que des sous-unités de l'ARN Pol II (Chapitre 4, Fig. 1A). L'interaction avec TFIIIS.1 a été confirmée par l'analyse réciproque, le facteur TFIIIS étant aussi identifié comme partenaire d'interaction de TFIIIS.1-TAP suite à la purification par chromatographie d'affinité de ce dernier. TFIIIS.1 fait partie de la famille de TFIIIS mais sa fonction dans la transcription est très peu caractérisée. La seule étude analysant le rôle de TFIIIS.1 a montré qu'il stimule la transactivation par l'activateur GRIP1tau et augmente l'expression d'un gène rapporteur (Nakata et al., 2004). Cependant, aucune étude n'avait cherché si TFIIIS.1 possède une activité de stimulation de la transcription *in vitro*. Étant donné que TFIIIS.1 et TFIIIS présentent un haut degré d'homologie de séquence dans la région C-terminale et centrale mais une région N-terminale divergente (Chapitre 2,

Fig. 1C), nous avons testé si TFIIS.1 stimule la transcription *in vitro* de la même façon que TFIIS. Il était connu que TFIIS stimulait la transcription dans des conditions de concentration basses en nucléotides, quand la polymérase a tendance à entrer en pause et être moins efficace en élongation (Langelier et al., 2005). Des essais de type “run-off” ont été effectués en utilisant un système reconstitué avec des facteurs recombinants et comme matrice le promoteur AdML (*Adenovirus Major Late*) fusionné à une cassette dépourvue de guanine (Annexe I, Fig. 2A, page 233). On a observé que, comparativement au contrôle, TFIIS.1 stimule la transcription en conditions de concentrations basses en nucléotides et il est presque aussi efficace que TFIIS pour produire le transcrit de 391 nucléotides. Ceci montre que TFIIS.1 peut lier les molécules de polymérase qui sont en pause transcriptionnelle et stimule la reprise de l'élongation. Suite à ces résultats, nous avons analysé la localisation génomique de TFIIS.1-TAP de la même façon que décrite dans le Chapitre 2. TFIIS.1 n'a été détecté sur aucun des gènes testés (données non-montrées) et ceci autant en conditions normales que suite à une exposition à des radiations ultraviolettes de type UVC. Afin de tester si une séquestration dans d'autres compartiments cellulaires pourrait expliquer ces résultats, nous avons analysé la localisation cellulaire de la protéine de fusion TFIIS.1-GFP (*green fluorescent protein*) ainsi que celle de TFIIS. Les deux protéines ont présenté une localisation nucléaire autant en conditions normales que suite à une exposition à des radiations ultraviolettes de type UVC (Annexe I, Fig. 2B, page 233). Ces résultats suggèrent que si TFIIS.1 a une fonction de stimulation de la polymérase de façon similaire qu'observé *in vitro*, alors cette activité devrait être restreinte à certains gènes cibles ou à des conditions spécifiques autres que celles analysées.

5.3 Identification de nouvelles fonctions pour le facteur TFIIS et pour l'ubiquitine ligase UBR5

Les résultats présentés dans le Chapitre 4 montrent que TFIIS a une nouvelle fonction, celle de favoriser l'ubiquitination de CDK9 en recrutant l'ubiquitine ligase

de type E3 UBR5 (Chapitre 4, Fig. 1, Fig. 3 et Fig. 4). CDK9 est une protéine hautement régulée, plusieurs mécanismes étant impliqués dans la modulation de son activité, entre autres par son recrutement dans des complexes régulateurs ainsi que par des modifications post- traductionnelles, tel que décrit dans le Chapitre 1. Nos résultats apportent des indices sur un nouveau mécanisme de régulation de CDK9 par son association dans un complexe avec TFIIS et UBR5, ce qui favorise son ubiquitination. De plus, nos résultats montrent que cette modification n'est pas impliquée dans la dégradation de CDK9 (Chapitre 4, Fig. 5).

Nos résultats indiquent aussi que TFIIS augmente le recrutement de CDK9 ainsi que celui de l'ARN Pol II sur le gène γ Fibrinogen dont l'expression est dépendante de CDK9 (Chapitre 4, Fig 7A, 7C et 7D). Puisque cet effet est observé même dans des conditions basales, quand normalement ce gène n'est pas exprimé, ces résultats suggèrent que TFIIS est impliqué dans l'activation de la transcription de ce gène, en accord avec d'autres études qui ont montré que TFIIS est impliqué dans l'initiation de la transcription et qu'il est requis pour la formation du complexe de préinitiation (Guglielmi et al., 2007; Prather et al., 2005). De plus, nos résultats montrent que l'effet de TFIIS sur les niveaux d'occupation par CDK9 du promoteur et des régions transcrites en aval du promoteur dépend de UBR5 (Chapitre 4, Fig 7B). L'association physique de TFIIS et de UBR5 avec ces régions suggèrent que l'effet de TFIIS sur l'occupation par CDK9 de ces mêmes régions résulte des interactions qui ont lieu sur la chromatine. Ces interactions pourraient impliquer l'ubiquitination sur la chromatine de CDK9 par UBR5 ou alternativement, un changement de conformation induit lorsque TFIIS et UBR5 s'associent aux complexes transcriptionnels, favorisant l'association de CDK9. Des études cristallographiques ont montré que le facteur TFIIS induit des changements de conformation significatifs à la polymérase elle-même, lorsque son domaine C-terminal s'insère dans le centre catalytique de celle-ci (Kettenberger et al., 2003). Cependant, il est peu probable que ce changement de conformation soit le seul événement dont dépend l'association de CDK9 aux complexes transcriptionnels en élongation, puisque UBR5 est aussi requis, tel que mentionné plus haut.

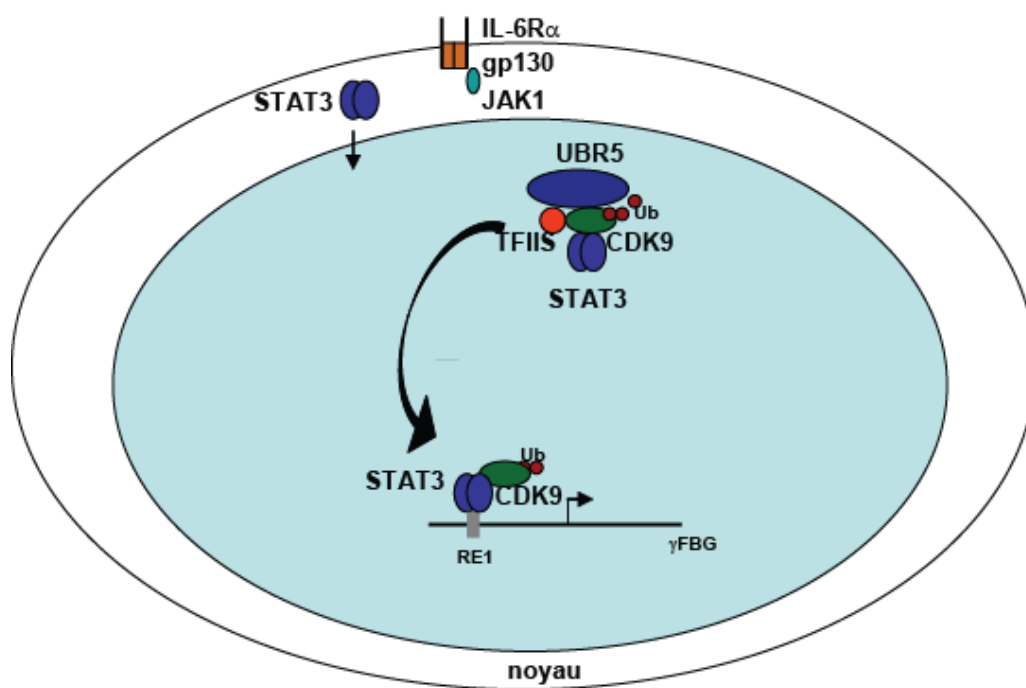
Les expériences de co-immunoprecipitation de protéines endogènes et par pull-down en utilisant la protéine de fusion TFIIS-GST montrent que ces trois protéines interagissent directement ou indirectement et forment un complexe autant *in vivo* que *in vitro*. Pour ce qui est des interactions rapportées précédemment pour les protéines impliquées, il a été montré par exemple que l'interaction de CDK9 avec une autre ubiquitine ligase de type E3, SCF^{SKP2}, se fait par l'intermédiaire de la Cycline T1, le partenaire de CDK9 avec lequel il forme un complexe actif en transcription. Une interaction directe entre UBR5 et une autre kinase, CHK2 a été démontrée chez l'humain, cependant il n'a pas été montré que cette interaction résulte en une ubiquitination de CDK9. L'analyse par chromatographie d'affinité de la protéine de fusion CDK9-TAP effectuée dans notre laboratoire n'a pas identifié les protéines UBR5 ou TFIIS comme partenaires d'interaction majeurs alors que la Cycline T1 et les protéines Hexim1/2 l'ont été (Jeronimo et al., 2007). Ceci suggère que TFIIS, UBR5 et CDK9 ne forment pas un complexe majeur, d'abondance comparable au complexe de CDK9 avec la Cycline T1 ou avec les protéines Hexim 1/2 et le petit ARN 7SK. Alternativement, l'interaction UBR5-CDK9 pourrait se faire à travers la Cycline T1, de manière similaire à l'interaction CDK9- SCF^{SKP2}.

L'implication de l'ubiquitine ligase de type E3 UBR5 dans la transcription par l'ARN Pol II a été mise en évidence par une étude précédente montrant que cette ligase joue un rôle positif dans la transactivation par le récepteur de la progestérone (Henderson et al., 2002). Nos résultats présentés dans le Chapitre 4 (Fig. 6B et Fig. 7B) montrent que UBR5 est impliqué dans la transition entre les étapes d'initiation et d'élongation de la transcription et qu'il est requis pour que CDK9 s'associe aux complexes transcriptionnels actifs. Le recrutement physique de UBR5, en parallèle avec celui de CDK9 et de la polymérase elle-même, appuie cette conclusion (Chapitre 4, Fig. 6B et Fig. 7A, 7C et 7D).

La Figure 1 (Chapitre 5) présente un modèle basé sur nos résultats qui décrit la signification fonctionnelle de l'interaction TFIIS-UBR5-CDK9. Dans ce modèle,

l'interaction de ces trois protéines dans la fraction cellulaire soluble favorise l'ubiquitination de CDK9 par UBR5 et l'interaction avec le partenaire connu de CDK9, l'activateur STAT3, suite à l'induction par l'Interleukine 6 (Fig. 1A). Cette interaction mène à la phosphorylation de STAT3 par CDK9 et au recrutement de ces deux protéines à l'élément de réponse à l'Interleukine 6 situé en amont du gène γ Fibrinogène. TFIIS, UBR5 et l'ARN Pol II sont recrutés ensuite au niveau du promoteur. La phosphorylation du CTD de l'ARN Pol II par CDK9 favorise l'entrée en élongation. Nous proposons que le recrutement de UBR5 et de TFIIS favorise la stabilité du complexe CDK9-RNA Pol II au niveau du promoteur et promeut la modification par phosphorylation de cette dernière (Fig. 1B). Ceci favoriserait finalement la transition entre l'initiation et l'élongation.

A



B

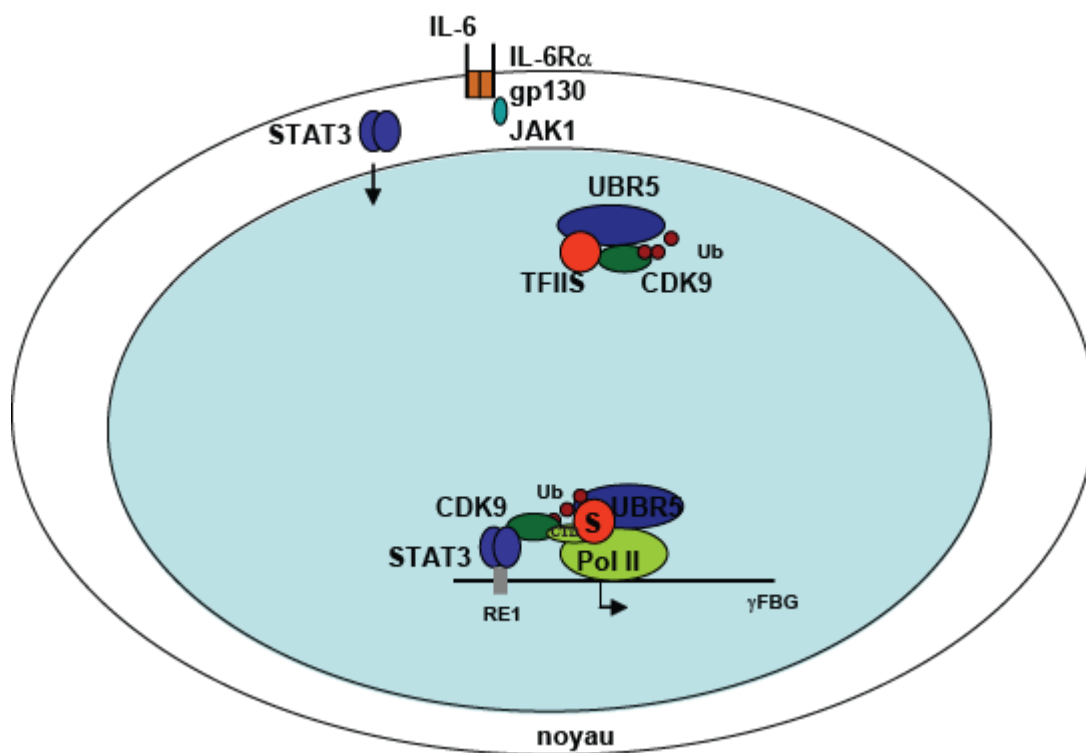


Figure 1. Modèle décrivant le rôle fonctionnel de l'interaction TFIIS-UBR5-CDK9

Dans l'ensemble, nos résultats suggèrent que TFIIIS est impliqué dans la régulation de CDK9 autant dans la fraction cellulaire soluble que sur la chromatine et que cette fonction implique le recrutement de l'ubiquitine ligase de type E3 UBR5. De plus, à partir de ces résultats, nous avons proposé une nouvelle fonction pour TFIIIS, celle de coordonner l'initiation de la transcription avec l'entrée en élongation productive.

CONCLUSIONS ET PERSPECTIVES

Le premier objectif a été d'étudier systématiquement la localisation génomique des GTFs et de l'ARN Pol II humains afin de faire la distinction entre leur implication dans les étapes précoces et tardives de la transcription. Dans ce but, on a utilisé comme approche l'immunoprécipitation de la chromatine et une collection de facteurs portant une étiquette TAP, afin d'effectuer une analyse systématique. L'analyse d'un groupe de gènes qui utilisent des promoteurs uniques nous a permis d'évaluer de façon spécifique la localisation *in vivo* des GTFs au niveau des promoteurs ou dans des régions transcrites situées en aval. Notre étude a amené une première image d'ensemble sur la fonction *in vivo* des GTFs chez l'humain et nous a permis de mettre en évidence le rôle dans l'élongation des sous-unités Rpb4/Rpb7 ainsi que du facteur TFIIIF.

Dans une prochaine étape, il serait intéressant d'étudier la localisation de cet ensemble de facteurs à l'échelle du génome humain et de combiner cette analyse avec celle des marqueurs de l'initiation de la transcription tels que des modifications de la chromatine et la forme de la polymérase ayant le CTD phosphorylé sur la sérine 5. Ce type d'analyse pourrait nous apporter des indices sur la fonction globale des GTFs dans les étapes de la transcription *in vivo* et ainsi de savoir si, chez l'humain, l'utilisation des groupes distincts de GTFs constitue un mécanisme majeur de régulation de la transcription. Des gènes à promoteurs multiples pourraient également être analysés par cette approche. Une méthode efficace d'étiquetage des protéines endogènes humaines par le "knock in" dans le locus du gène endogène ciblé a été rapportée (Zhang et al., 2008). Par cette technique, la compétition entre les versions étiquetées et endogènes des protéines pour les mêmes sites génomiques est éliminée et une détection plus efficace des protéines étiquetées par Chip-on-CHIP devient possible. Une telle approche permettrait une analyse systématique de la localisation des GTFs à l'échelle du génome et pourrait répondre

à de multiples questions. Par exemple, en utilisant plusieurs types cellulaires, il serait possible d'identifier des rôles spécifiques des GTFs humains selon le type cellulaire, tel que suggéré par des études à petite échelle. Une analyse du patron de localisation à l'échelle du génome dans différentes conditions de stress pourrait indiquer si les GTFs ont des rôles dans la réponse transcriptionnelle spécifique à ces types de stress. La déplétion de TFIIF ou des sous-unités Rpb4/Rpb7 et l'analyse de l'impact sur la distribution de la polymérase à l'échelle du génome serait un moyen de savoir si leur rôle dans l'élongation chez l'humain est général ou plutôt spécifique à différents types de gènes ou régions transcrites. Une telle étude serait d'un grand intérêt, étant donné qu'une régulation aberrante de la transcription par la machinerie transcriptionnelle générale a été impliquée dans de pathologies diverses allant de différents types de cancers à des maladies chroniques comme le diabète et les affections inflammatoires chroniques (revu par (Davuluri et al., 2008).

Le deuxième objectif a été de trouver de nouveaux partenaires d'interaction et de nouvelles fonctions pour le facteur TFIIS. Cet objectif a été atteint en utilisant comme point de départ la purification par chromatographie affinité en tandem (TAP) de TFIIS et l'identification des partenaires d'interaction par spectrométrie de masse. Cette approche a révélé la E3 ubiquitine ligase UBR5 comme nouveau partenaire d'interaction. Nous avons également démontré que TFIIS et UBR5 interagissent avec la sous-unité kinase de P-TEFb, CDK9, et sont impliqués dans son ubiquitination *in vivo*. Des études d'ubiquitination *in vitro* ont montré que UBR5 catalyse le transfert d'ubiquitine sur CDK9 avec la formation des espèces polyubiquitinées. Cette modification par polyubiquitination de CDK9 n'est pas un signal de dégradation, tel que nous avons pu l'observer par des essais de type luciférase. Nous avons démontré que TFIIS, UBR5 et CDK9 co-localisent sur le gène γ fibrinogène et que TFIIS régule la distribution de CDK9 sur ce gène de façon dépendante de UBR5.

Plusieurs aspects liés à la nouvelle fonction de TFIIS dans la régulation de CDK9 restent à étudier dans la poursuite de ce projet. Premièrement, il serait pertinent

d'effectuer une étude détaillée afin de caractériser les domaines impliqués dans les interactions qui mènent à la formation du complexe TFIIS-UBR5-CDK9. Deuxièmement, une analyse par spectrométrie de masse permettra d'identifier des sites d'ubiquitination de CDK9 par UBR5. Un troisième point qui pourrait être adressé est le mécanisme par lequel TFIIS et UBR5 régulent CDK9 au niveau de la chromatine. En analysant la distribution génomique de CDK9 dans des cellules qui expriment une version de UBR5 avec une mutation de la cystéine catalytique pourrait indiquer si une ubiquitination de CDK9 sur la chromatine est impliquée dans sa régulation.

Dans l'ensemble, nos résultats amènent pour la première fois une image globale de la fonction des GTFs dans la transcription par l'ARN Pol II *in vivo* et fournissent des indices sur de nouvelles fonctions des sous-unités Rpb4/Rpb7 de la polymérase et de plusieurs facteurs tels que TFIIF, TFIIS et UBR5. Ces nouvelles données contribuent à la compréhension mécanistique du processus transcriptionnel par l'ARN Pol II dans les cellules vivantes et identifient de nouveaux rôles pour des composants de la machinerie générale de transcription dans des étapes régulatrices cruciales de l'expression génique. Nos résultats appuient la nouvelle vue qui suggère que la machinerie transcriptionnelle générale de l'ARN Pol II constitue une voie active de diversification dans la régulation de l'expression génique, plutôt qu'une simple interface invariable entre les activateurs/répresseurs et l'ARN Pol II. Le potentiel régulateur de cette machinerie reste encore peu exploré, malgré le nombre croissant d'indices sur son implication dans de pathologies variées incluant des maladies cardiaques, différents types de cancers ainsi que des infections virales. Une meilleure compréhension et intégration des différents niveaux régulateurs de l'expression génique nous conduiront à de nouvelles approches thérapeutiques.

REFERENCES

1. Acker, J., Wintzerith, M., Vigneron, M., and Kedinger, C. (1992). Primary structure of the second largest subunit of human RNA polymerase II (or B). *J. Mol. Biol.* 226, 1295-1299.
2. Adelman, K., Marr, M.T., Werner, J., Saunders, A., Ni, Z.Y., Andrulis, E.D., and Lis, J.T. (2005). Efficient release from promoter-proximal stall sites requires transcript cleavage factor TFIIS. *Molecular Cell* 17, 103-112.
3. Ahn, S.H., Keogh, M.C., and Buratowski, S. (2009). Ctk1 promotes dissociation of basal transcription factors from elongating RNA polymerase II. *EMBO J.* 28, 205-212.
4. Ahn, S.H., Kim, M., and 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.
5. Akoulitchev, S., Chuikov, S., and Reinberg, D. (2000). TFIID is negatively regulated by cdk8-containing mediator complexes. *Nature* 407, 102-106.
6. Aleksic, J., and Russell, S. (2009). ChIPing away at the genome: the new frontier travel guide. *Mol. Biosyst.*
7. Ammosova, T., Washington, K., Debebe, Z., Brady, J., and Nekhai, S. (2005). Dephosphorylation of CDK9 by protein phosphatase 2A and protein phosphatase-1 in Tat-activated HIV-1 transcription. *Retrovirology.* 2, 47.
8. Andel, F., III, Ladurner, A.G., Inouye, C., Tjian, R., and Nogales, E. (1999). Three-dimensional structure of the human TFIID-IIA-IIB complex. *Science* 286, 2153-2156.
9. Andrau, J.C., van de, P.L., Lijnzaad, P., Bijma, T., Koerkamp, M.G., van de, P.J., Werner, M., and Holstege, F.C. (2006). Genome-Wide Location of the Coactivator Mediator: Binding without Activation and Transient Cdk8 Interaction on DNA. *Mol. Cell* 22, 179-192.
10. Anish, R., Hossain, M.B., Jacobson, R.H., and Takada, S. (2009). Characterization of transcription from TATA-less promoters: identification of a new core promoter element XCPE2 and analysis of factor requirements. *PLoS. One.* 4, e5103.
11. Ansari, S.A., He, Q., and Morse, R.H. (2009). Mediator complex association with constitutively transcribed genes in yeast. *Proc. Natl. Acad. Sci. U. S. A* 106, 16734-16739.

12. Aparicio, O., Geisberg, J.V., and Struhl, K. (2004). Chromatin immunoprecipitation for determining the association of proteins with specific genomic sequences in vivo. *Curr. Protoc. Cell Biol.* Chapter 17, Unit.
13. Arce, L., Yokoyama, N.N., and Waterman, M.L. (2006). Diversity of LEF/TCF action in development and disease. *Oncogene* 25, 7492-7504.
14. Archambault, J., Pan, G., Dahmus, G.K., Cartier, M., Marshall, N., Zhang, S., Dahmus, M.E., and Greenblatt, J. (1998). 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.
15. 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.
16. Armache, K.J., Mitterweger, S., Meinhart, A., and Cramer, P. (2005). Structures of complete RNA polymerase II and its subcomplex, Rpb4/7. *J. Biol Chem.* 280, 7131-7134.
17. Artsimovitch, I., and Landick, R. (2000). Pausing by bacterial RNA polymerase is mediated by mechanistically distinct classes of signals. *Proc. Natl. Acad. Sci. U. S. A* 97, 7090-7095.
18. Artsimovitch, I., and Landick, R. (2002). The transcriptional regulator RfaH stimulates RNA chain synthesis after recruitment to elongation complexes by the exposed nontemplate DNA strand. *Cell* 109, 193-203.
19. Auble, D.T. (2009). The dynamic personality of TATA-binding protein. *Trends Biochem. Sci.* 34, 49-52.
20. Auble, D.T., and Hahn, S. (1993). An ATP-dependent inhibitor of TBP binding to DNA. *Genes Dev.* 7, 844-856.
21. Awrey, D.E., Shimasaki, N., Koth, C., Weilbaecher, R., Olmsted, V., Kazanis, S., Shan, X., Arellano, J., Arrowsmith, C.H., Kane, C.M., and Edwards, A.M. (1998). Yeast transcript elongation factor (TFIIS), structure and function. II: RNA polymerase binding, transcript cleavage, and read-through. *J. Biol Chem.* 273, 22595-22605.
22. Awrey, D.E., Weilbaecher, R.G., Hemming, S.A., Orlicky, S.M., Kane, C.M., and Edwards, A.M. (1997). Transcription elongation through DNA arrest sites. A multistep process involving both RNA

- polymerase II subunit RPB9 and TFIIS. *J. Biol. Chem.* 272, 14747-14754.
23. Baek, D., Davis, C., Ewing, B., Gordon, D., and Green, P. (2007). Characterization and predictive discovery of evolutionarily conserved mammalian alternative promoters. *Genome Res.* 17, 145-155.
 24. Bagby, S., Kim, S., Maldonado, E., Tong, K.I., Reinberg, D., and Ikura, M. (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.
 25. Bairoch, A., Apweiler, R., Wu, C.H., Barker, W.C., Boeckmann, B., Ferro, S., Gasteiger, E., Huang, H., Lopez, R., Magrane, M., Martin, M.J., Natale, D.A., O'Donovan, C., Redaschi, N., and Yeh, L.S. (2005). The Universal Protein Resource (UniProt). *Nucleic Acids Res.* 33, D154-D159.
 26. Bajic, V.B., Tan, S.L., Christoffels, A., Schonbach, C., Lipovich, L., Yang, L., Hofmann, O., Kruger, A., Hide, W., Kai, C., Kawai, J., Hume, D.A., Carninci, P., and Hayashizaki, Y. (2006). Mice and men: their promoter properties. *PLoS. Genet.* 2, e54.
 27. Bar-Nahum, G., Epshtein, V., Ruckenstein, A.E., Rafikov, R., Mustaev, A., and Nudler, E. (2005). A ratchet mechanism of transcription elongation and its control. *Cell* 120, 183-193.
 28. Barboric, M., Nissen, R.M., Kanazawa, S., Jabrane-Ferrat, N., and Peterlin, B.M. (2001). NF-kappaB binds P-TEFb to stimulate transcriptional elongation by RNA polymerase II. *Mol. Cell* 8, 327-337.
 29. Barboric, M., Zhang, F., Besenicar, M., Plemenitas, A., and Peterlin, B.M. (2005). Ubiquitylation of Cdk9 by Skp2 facilitates optimal Tat transactivation. *J. Virol.* 79, 11135-11141.
 30. Bartfai, R., Balduf, C., Hilton, T., Rathmann, Y., Hadzhiev, Y., Tora, L., Orban, L., and Muller, F. (2004). TBP2, a vertebrate-specific member of the TBP family, is required in embryonic development of zebrafish. *Curr. Biol* 14, 593-598.
 31. Benga, W.J., Grandemange, S., Shpakovski, G.V., Shematorova, E.K., Kedinger, C., and Vigneron, M. (2005). Distinct regions of RPB11 are required for heterodimerization with RPB3 in human and yeast RNA polymerase II. *Nucleic Acids Res.* 33, 3582-3590.

32. Bengal, E., Flores, O., Krauskopf, A., Reinberg, D., and 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.
33. Bentley, D.L. (2005). Rules of engagement: co-transcriptional recruitment of pre-mRNA processing factors. *Curr. Opin. Cell Biol* 17, 251-256.
34. Bernassola, F., Karin, M., Ciechanover, A., and Melino, G. (2008). The HECT family of E3 ubiquitin ligases: multiple players in cancer development. *Cancer Cell* 14, 10-21.
35. Berretta, J., and Morillon, A. (2009). Pervasive transcription constitutes a new level of eukaryotic genome regulation. *EMBO Rep.* 10, 973-982.
36. Bjorklund, S., and Gustafsson, C.M. (2005a). The yeast Mediator complex and its regulation. *Trends Biochem. Sci.* 30, 240-244.
37. Bjorklund, S., and Gustafsson, C.M. (2005b). The yeast Mediator complex and its regulation. *Trends Biochem. Sci.* 30, 240-244.
38. Blau, J., Xiao, H., McCracken, S., O'Hare, P., Greenblatt, J., and Bentley, D. (1996). Three functional classes of transcriptional activation domain. *Mol. Cell Biol.* 16, 2044-2055.
39. Boeing, S., Rigault, C., Heidemann, M., Eick, D., and Meisterernst, M. (2009). RNAPII CTD SER-7 phosphorylation is established in a mediator-dependent fashion. *J. Biol. Chem.*
40. Bonham, A.J., Neumann, T., Tirrell, M., and Reich, N.O. (2009). Tracking transcription factor complexes on DNA using total internal reflectance fluorescence protein binding microarrays. *Nucleic Acids Res.* 37, e94.
41. Booth, V., Koth, C.M., Edwards, A.M., and Arrowsmith, C.H. (2000). Structure of a conserved domain common to the transcription factors TFIIIS, elongin A, and CRSP70. *J. Biol Chem.* 275, 31266-31268.
42. Borggreffe, T., Davis, R., Erdjument-Bromage, H., Tempst, P., and Kornberg, R.D. (2002). A complex of the Srb8, -9, -10, and -11 transcriptional regulatory proteins from yeast. *J. Biol. Chem.* 277, 44202-44207.
43. Boube, M., Joulia, L., Cribbs, D.L., and Bourbon, H.M. (2002). Evidence for a Mediator of RNA Polymerase II Transcriptional Regulation Conserved from Yeast to Man. *Cell* 110, 143-151.

44. Boyer-Guittaut, M., Birsoy, K., Potel, C., Elliott, G., Jaffray, E., Desterro, J.M., Hay, R.T., and Oelgeschlager, T. (2005). SUMO-1 modification of human transcription factor (TF) IID complex subunits: inhibition of TFIID promoter-binding activity through SUMO-1 modification of hsTAF5. *J. Biol. Chem.* 280, 9937-9945.
45. Bradsher, J., Coin, F., and Egly, J.M. (2000). Distinct roles for the helicases of TFIID in transcript initiation and promoter escape. *J. Biol. Chem.* 275, 2532-2538.
46. Brand, M., Leurent, C., Mallouh, V., Tora, L., and Schultz, P. (1999). Three-dimensional structures of the TAFII-containing complexes TFIID and TFTC. *Science* 286, 2151-2153.
47. Brasier, A.R. (2008). Expanding role of cyclin dependent kinases in cytokine inducible gene expression. *Cell Cycle* 7, 2661-2666.
48. Britten, R.J., and Davidson, E.H. (1969). Gene regulation for higher cells: a theory. *Science* 165, 349-357.
49. Brooks, P.J., Cheng, T.F., and Cooper, L. (2008). Do all of the neurologic diseases in patients with DNA repair gene mutations result from the accumulation of DNA damage? *DNA Repair (Amst)* 7, 834-848.
50. Brower, C.S., Sato, S., Tomomori-Sato, C., Kamura, T., Pause, A., Stearman, R., Klausner, R.D., Malik, S., Lane, W.S., Sorokina, I., Roeder, R.G., Conaway, J.W., and Conaway, R.C. (2002). Mammalian mediator subunit mMED8 is an Elongin BC-interacting protein that can assemble with Cul2 and Rbx1 to reconstitute a ubiquitin ligase. *Proc. Natl. Acad. Sci. U. S. A* 99, 10353-10358.
51. Brueckner, F., and Cramer, P. (2008). Structural basis of transcription inhibition by alpha-amanitin and implications for RNA polymerase II translocation. *Nat. Struct. Mol. Biol.* 15, 811-818.
52. Brueckner, F., Ortiz, J., and Cramer, P. (2009). A movie of the RNA polymerase nucleotide addition cycle. *Curr. Opin. Struct. Biol.* 19, 294-299.
53. Bryant, G.O., and Ptashne, M. (2003). Independent recruitment in vivo by gal4 of two complexes required for transcription. *Mol. Cell* 11, 1301-1309.
54. Buratowski, S. (2005). Connections between mRNA 3' end processing and transcription termination. *Curr. Opin. Cell Biol* 17, 257-261.
55. Buratowski, S. (2009). Progression through the RNA polymerase II CTD cycle. *Mol. Cell* 36, 541-546.

56. Burley, S.K., and Roeder, R.G. (1996). Biochemistry and structural biology of transcription factor IID (TFIID). *Annu. Rev. Biochem.* 65, 769-799.
57. Burton, Z.F., Ortolan, L.G., and Greenblatt, J. (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.
58. Bushnell, D.A., and Kornberg, R.D. (2003a). 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.
59. Bushnell, D.A., and Kornberg, R.D. (2003b). Complete, 12-subunit RNA polymerase II at 4.1-A resolution: Implications for the initiation of transcription. *Proc. Natl. Acad. Sci. U. S. A.*
60. 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.
61. Byers, S.A., Price, J.P., Cooper, J.J., Li, Q., and Price, D.H. (2005). HEXIM2, a HEXIM1-related protein, regulates positive transcription elongation factor b through association with 7SK. *J. Biol Chem.* 280, 16360-16367.
62. Cairns, B.R. (2009). The logic of chromatin architecture and remodelling at promoters. *Nature* 461, 193-198.
63. Cairns, B.R., Henry, N.L., and Kornberg, R.D. (1996). TFG/TAF30/ANC1, a component of the yeast SWI/SNF complex that is similar to the leukemogenic proteins ENL and AF-9. *Mol. Cell Biol.* 16, 3308-3316.
64. Callaghan, M.J., Russell, A.J., Woollatt, E., Sutherland, G.R., Sutherland, R.L., and Watts, C.K. (1998). Identification of a human HECT family protein with homology to the *Drosophila* tumor suppressor gene hyperplastic discs. *Oncogene* 17, 3479-3491.
65. Callen, B.P., Shearwin, K.E., and Egan, J.B. (2004). Transcriptional interference between convergent promoters caused by elongation over the promoter. *Mol Cell* 14, 647-656.
66. Camblong, J., Iglesias, N., Fickentscher, C., Dieppois, G., and Stutz, F. (2007). Antisense RNA stabilization induces transcriptional gene silencing via histone deacetylation in *S. cerevisiae*. *Cell* 131, 706-717.

67. Campbell, E.A., Muzzin, O., Chlenov, M., Sun, J.L., Olson, C.A., Weinman, O., Trester-Zedlitz, M.L., and Darst, S.A. (2002). Structure of the Bacterial RNA Polymerase Promoter Specificity sigma Subunit. *Mol. Cell* 9, 527-539.
68. Carninci, P. (2009). Molecular biology: The long and short of RNAs. *Nature* 457, 974-975.
69. Chalkley, G.E., and 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.
70. Chambers, R.S., Wang, B.Q., Burton, Z.F., and 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.
71. Chao, S.H., and Price, D.H. (2001). Flavopiridol inactivates P-TEFb and blocks most RNA polymerase II transcription in vivo. *J. Biol. Chem.* 276, 31793-31799.
72. Charlet-Berguerand, N., Feuerhahn, S., Kong, S.E., Ziserman, H., Conaway, J.W., Conaway, R., and Egly, J.M. (2006). RNA polymerase II bypass of oxidative DNA damage is regulated by transcription elongation factors. *EMBO J.* 25, 5481-5491.
73. Chen, C.Y., Chang, C.C., Yen, C.F., Chiu, M.T., and Chang, W.H. (2009a). Mapping RNA exit channel on transcribing RNA polymerase II by FRET analysis. *Proc. Natl. Acad. Sci. U. S. A* 106, 127-132.
74. Chen, C.Y., Chang, C.C., Yen, C.F., Chiu, M.T., and Chang, W.H. (2009b). Mapping RNA exit channel on transcribing RNA polymerase II by FRET analysis. *Proc. Natl. Acad. Sci. U. S. A* 106, 127-132.
75. Chen, D., Riedl, T., Washbrook, E., Pace, P.E., Coombes, R.C., Egly, J.M., and Ali, S. (2000). Activation of estrogen receptor alpha by S118 phosphorylation involves a ligand-dependent interaction with TFIIF and participation of CDK7. *Mol. Cell* 6, 127-137.
76. Chen, H.T., Warfield, L., and Hahn, S. (2007). The positions of TFIIF and TFIIE in the RNA polymerase II transcription preinitiation complex. *Nat. Struct. Mol. Biol.* 14, 696-703.
77. Chen, R., Keating, M.J., Gandhi, V., and Plunkett, W. (2005). Transcription inhibition by flavopiridol: mechanism of chronic lymphocytic leukemia cell death. *Blood* 106, 2513-2519.

78. Cheng, B., and Price, D.H. (2007). Properties of RNA polymerase II elongation complexes before and after the P-TEFb-mediated transition into productive elongation. *J. Biol Chem.* 282, 21901-21912.
79. Chi, T., and Carey, M. (1996). Assembly of the isomerized TFIIA--TFIID--TATA ternary complex is necessary and sufficient for gene activation. *Genes Dev.* 10, 2540-2550.
80. Chin, S.F., Teschendorff, A.E., Marioni, J.C., Wang, Y., Barbosa-Morais, N.L., Thorne, N.P., Costa, J.L., Pinder, S.E., van de Wiel, M.A., Green, A.R., Ellis, I.O., Porter, P.L., Tavaré, S., Brenton, J.D., Ylstra, B., and Caldas, C. (2007a). High-resolution aCGH and expression profiling identifies a novel genomic subtype of ER negative breast cancer. *Genome Biol.* 8, R215.
81. Chin, S.F., Teschendorff, A.E., Marioni, J.C., Wang, Y., Barbosa-Morais, N.L., Thorne, N.P., Costa, J.L., Pinder, S.E., van de Wiel, M.A., Green, A.R., Ellis, I.O., Porter, P.L., Tavaré, S., Brenton, J.D., Ylstra, B., and Caldas, C. (2007b). High-resolution aCGH and expression profiling identifies a novel genomic subtype of ER negative breast cancer. *Genome Biol.* 8, R215.
82. Cho, E.J., Kobor, M.S., Kim, M., Greenblatt, J., and 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.
83. Choder, M. (2004). Rpb4 and Rpb7: subunits of RNA polymerase II and beyond. *Trends Biochem. Sci.* 29, 674-681.
84. Choder, M., and 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.
85. Chong, J.A., Moran, M.M., Teichmann, M., Kaczmarek, J.S., Roeder, R., and Clapham, D.E. (2005). TATA-binding protein (TBP)-like factor (TLF) is a functional regulator of transcription: reciprocal regulation of the neurofibromatosis type 1 and c-fos genes by TLF/TRF2 and TBP. *Mol. Cell Biol.* 25, 2632-2643.
86. Christian, B.A., Grever, M.R., Byrd, J.C., and Lin, T.S. (2007). Flavopiridol in the treatment of chronic lymphocytic leukemia. *Curr. Opin. Oncol.* 19, 573-578.
87. Chu, Y., Simic, R., Warner, M.H., Arndt, K.M., and Prelich, G. (2007). Regulation of histone modification and cryptic transcription by the Bur1 and Paf1 complexes. *EMBO J.* 26, 4646-4656.

88. Ciampi, M.S. (2006). Rho-dependent terminators and transcription termination. *Microbiology* 152, 2515-2528.
89. Clancy, J.L., Henderson, M.J., Russell, A.J., Anderson, D.W., Bova, R.J., Campbell, I.G., Choong, D.Y., Macdonald, G.A., Mann, G.J., Nolan, T., Brady, G., Olopade, O.I., Woollatt, E., Davies, M.J., Segara, D., Hacker, N.F., Henshall, S.M., Sutherland, R.L., and Watts, C.K. (2003). EDD, the human orthologue of the hyperplastic discs tumour suppressor gene, is amplified and overexpressed in cancer. *Oncogene* 22, 5070-5081.
90. Clapier, C.R., and Cairns, B.R. (2009). The biology of chromatin remodeling complexes. *Annu. Rev. Biochem.* 78, 273-304.
91. Clark, A.B., Dykstra, C.C., and Sugino, A. (1991). Isolation, DNA sequence, and regulation of a *Saccharomyces cerevisiae* gene that encodes DNA strand transfer protein alpha. *Mol. Cell Biol.* 11, 2576-2582.
92. Cler, E., Papai, G., Schultz, P., and Davidson, I. (2009). Recent advances in understanding the structure and function of general transcription factor TFIID. *Cell Mol. Life Sci.* 66, 2123-2134.
93. Cloutier, P., Al Khoury, R., Lavalley-Adam, M., Faubert, D., Jiang, H., Poitras, C., Bouchard, A., Forget, D., Blanchette, M., and Coulombe, B. (2009). High-resolution mapping of the protein interaction network for the human transcription machinery and affinity purification of RNA polymerase II-associated complexes. *Methods* 48, 381-386.
94. Coates, J.C. (2003). Armadillo repeat proteins: beyond the animal kingdom. *Trends Cell Biol.* 13, 463-471.
95. Coin, F., Oksenysh, V., Mocquet, V., Groh, S., Blattner, C., and Egly, J.M. (2008). Nucleotide excision repair driven by the dissociation of CAK from TFIIH. *Mol. Cell* 31, 9-20.
96. Cojocaru, M., Jeronimo, C., Forget, D., Bouchard, A., Bergeron, D., Cote, P., Poirier, G.G., Greenblatt, J., and Coulombe, B. (2008). Genomic location of the human RNA polymerase II general machinery: evidence for a role of TFIIF and Rpb7 at both early and late stages of transcription. *Biochem. J.* 409, 139-147.
97. Collas, P. (2010). The Current State of Chromatin Immunoprecipitation. *Mol. Biotechnol.*
98. Compe, E., Malerba, M., Soler, L., Marescaux, J., Borrelli, E., and Egly, J.M. (2007). Neurological defects in trichothiodystrophy reveal a coactivator function of TFIIH. *Nat. Neurosci.* 10, 1414-1422.

99. Conaway, J.W., and Conaway, R.C. (1989). A multisubunit transcription factor essential for accurate initiation by RNA polymerase II. *J. Biol. Chem.* 264, 2357-2362.
100. 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.
101. Conaway, R.C., Garrett, K.P., Hanley, J.P., and 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.
102. Conaway, R.C., Sato, S., Tomomori-Sato, C., Yao, T., and Conaway, J.W. (2005). The mammalian Mediator complex and its role in transcriptional regulation. *Trends Biochem. Sci.* 30, 250-255.
103. Core, L.J., and Lis, J.T. (2008). Transcription regulation through promoter-proximal pausing of RNA polymerase II. *Science* 319, 1791-1792.
104. 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.
105. Coulombe, B., Killeen, M., Liljelund, P., Honda, B., Xiao, H., Ingles, C.J., and Greenblatt, J. (1992). Identification of three mammalian proteins that bind to the yeast TATA box protein TFIID. *Gene Expr.* 2, 99-110.
106. Coulombe, B., Li, J., and 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.
107. Cramer, P. (2002). Common structural features of nucleic acid polymerases. *Bioessays* 24, 724-729.
108. Cramer, P., Bushnell, D.A., Fu, J., Gnatt, A.L., Maier-Davis, B., Thompson, N.E., Burgess, R.R., Edwards, A.M., David, P.R., and Kornberg, R.D. (2000). Architecture of RNA polymerase II and implications for the transcription mechanism. *Science* 288, 640-649.
109. 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.

110. Dame, R.T., Wyman, C., and Goosen, N. (2003). Insights into the regulation of transcription by scanning force microscopy. *J. Microsc.* 212, 244-253.
111. Daniel, J.A., Torok, M.S., Sun, Z.W., Schieltz, D., Allis, C.D., Yates, J.R., III, and Grant, P.A. (2004). Deubiquitination of Histone H2B by a Yeast Acetyltransferase Complex Regulates Transcription. *J. Biol. Chem.* 279, 1867-1871.
112. Dantanel, J.C., Wurtz, J.M., Poch, O., Moras, D., and Tora, L. (1999). The TBP-like factor: an alternative transcription factor in metazoa? *Trends Biochem. Sci.* 24, 335-339.
113. Darst, R.P., Dasgupta, A., Zhu, C., Hsu, J.Y., Vroom, A., Muldrow, T., and Auble, D.T. (2003). Mot1 Regulates the DNA Binding Activity of Free TATA-binding Protein in an ATP-dependent Manner. *J. Biol. Chem.* 278, 13216-13226.
114. Darzacq, X., Shav-Tal, Y., de, T., V, Brody, Y., Shenoy, S.M., Phair, R.D., and Singer, R.H. (2007). In vivo dynamics of RNA polymerase II transcription. *Nat. Struct. Mol. Biol.* 14, 796-806.
115. Davie, J.K., and Kane, C.M. (2000). Genetic interactions between TFIIS and the Swi-Snf chromatin-remodeling complex. *Mol. Cell Biol.* 20, 5960-5973.
116. Davuluri, R.V., Suzuki, Y., Sugano, S., Plass, C., and Huang, T.H. (2008). The functional consequences of alternative promoter use in mammalian genomes. *Trends Genet.* 24, 167-177.
117. de la Mata, M., Alonso, C.R., Kadener, S., Fededa, J.P., Blaustein, M., Pelisch, F., Cramer, P., Bentley, D., and Kornblihtt, A.R. (2003). A slow RNA polymerase II affects alternative splicing in vivo. *Mol. Cell* 12, 525-532.
118. de la Serna, I.L., Ohkawa, Y., Berkes, C.A., Bergstrom, D.A., Dacwag, C.S., Tapscott, S.J., and Imbalzano, A.N. (2005). MyoD targets chromatin remodeling complexes to the myogenin locus prior to forming a stable DNA-bound complex. *Mol. Cell Biol.* 25, 3997-4009.
119. De, L.A., Esposito, V., Baldi, A., Claudio, P.P., Fu, Y., Caputi, M., Pisano, M.M., Baldi, F., and Giordano, A. (1997). CDC2-related kinase PITALRE phosphorylates pRb exclusively on serine and is widely expressed in human tissues. *J. Cell Physiol* 172, 265-273.
120. DeJong, J., and Roeder, R.G. (1993). A single cDNA, hTFIIA/alpha, encodes both the p35 and p19 subunits of human TFIIA. *Genes Dev.* 7, 2220-2234.

121. Deng, W., Malecova, B., Oelgeschlager, T., and Roberts, S.G. (2009). TFIIB recognition elements control the TFIIA-NC2 axis in transcriptional regulation. *Mol. Cell Biol.* 29, 1389-1400.
122. Deng, W., and Roberts, S.G. (2006). Core promoter elements recognized by transcription factor IIB. *Biochem. Soc. Trans.* 34, 1051-1053.
123. Deng, W., and Roberts, S.G. (2007). TFIIB and the regulation of transcription by RNA polymerase II. *Chromosoma* 116, 417-429.
124. Deo, R.C., Sonenberg, N., and Burley, S.K. (2001). X-ray structure of the human hyperplastic discs protein: an ortholog of the C-terminal domain of poly(A)-binding protein. *Proc. Natl. Acad. Sci. U. S. A* 98, 4414-4419.
125. Dey, A., Chitsaz, F., Abbasi, A., Misteli, T., and Ozato, K. (2003). The double bromodomain protein Brd4 binds to acetylated chromatin during interphase and mitosis. *Proc. Natl. Acad. Sci. U. S. A* 100, 8758-8763.
126. Di Lello, P., Jenkins, L.M., Jones, T.N., Nguyen, B.D., Hara, T., Yamaguchi, H., Dikeakos, J.D., Appella, E., Legault, P., and Omichinski, J.G. (2006). Structure of the Tfb1/p53 complex: Insights into the interaction between the p62/Tfb1 subunit of TFIIF and the activation domain of p53. *Mol Cell* 22, 731-740.
127. Dikstein, R., Ruppert, S., and Tjian, R. (1996a). TAFII250 is a bipartite protein kinase that phosphorylates the base transcription factor RAP74. *Cell* 84, 781-790.
128. Dikstein, R., Zhou, S., and Tjian, R. (1996b). Human TAFII 105 is a cell type-specific TFIID subunit related to hTAFII130. *Cell* 87, 137-146.
129. Dion, V., and Coulombe, B. (2003). Interactions of a DNA-bound transcriptional activator with the TBP-TFIIA-TFIIB-promoter quaternary complex. *J. Biol Chem.* 278, 11495-11501.
130. Donahue, B.A., Yin, S., Taylor, J.S., Reines, D., and Hanawalt, P.C. (1994). Transcript cleavage by RNA polymerase II arrested by a cyclobutane pyrimidine dimer in the DNA template. *Proc. Natl. Acad. Sci. U. S. A* 91, 8502-8506.
131. Donner, A.J., Ebmeier, C.C., Taatjes, D.J., and Espinosa, J.M. (2010). CDK8 is a positive regulator of transcriptional elongation within the serum response network. *Nat. Struct. Mol. Biol.* 17, 194-201.

132. Donner, A.J., Szostek, S., Hoover, J.M., and Espinosa, J.M. (2007). CDK8 is a stimulus-specific positive coregulator of p53 target genes. *Mol. Cell* 27, 121-133.
133. Douziech, M., Coin, F., Chipoulet, J.M., Arai, Y., Ohkuma, Y., Egly, J.M., and Coulombe, B. (2000). Mechanism of promoter melting by the xeroderma pigmentosum complementation group B helicase of transcription factor IIIH revealed by protein-DNA photo-cross-linking. *Mol. Cell Biol.* 20, 8168-8177.
134. Dvir, A., Conaway, R.C., and Conaway, J.W. (1997). A role for TFIIH in controlling the activity of early RNA polymerase II elongation complexes. *Proc. Natl. Acad. Sci. U. S. A* 94, 9006-9010.
135. Edwards, A.M., Kane, C.M., Young, R.A., and 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.
136. Egloff, S., and Murphy, S. (2008a). Cracking the RNA polymerase II CTD code. *Trends Genet.* 24, 280-288.
137. Egloff, S., and Murphy, S. (2008b). Cracking the RNA polymerase II CTD code. *Trends Genet.* 24, 280-288.
138. Egloff, S., O'Reilly, D., Chapman, R.D., Taylor, A., Tanzhaus, K., Pitts, L., Eick, D., and Murphy, S. (2007). Serine-7 of the RNA polymerase II CTD is specifically required for snRNA gene expression. *Science* 318, 1777-1779.
139. El, K.B., Medler, S., Raghunayakula, S., and Ansari, A. (2009). Gene looping is conferred by activator-dependent interaction of transcription initiation and termination machineries. *J. Biol. Chem.* 284, 25015-25025.
140. Elsby, L.M., and Roberts, S.G. (2004). The role of TFIIB conformation in transcriptional regulation. *Biochem. Soc. Trans.* 32, 1098-1099.
141. Esnault, C., Ghavi-Helm, Y., Brun, S., Soutourina, J., Van, B.N., Boschiero, C., Holstege, F., and Werner, M. (2008). Mediator-dependent recruitment of TFIIH modules in preinitiation complex. *Mol. Cell* 31, 337-346.
142. Exinger, F., and Lacroute, F. (1992). 6-Azauracil Inhibition of Gtp Biosynthesis in *Saccharomyces-Cerevisiae*. *Current Genetics* 22, 9-11.

143. Fabrega, C., Shen, V., Shuman, S., and Lima, C.D. (2003). Structure of an mRNA capping enzyme bound to the phosphorylated carboxy-terminal domain of RNA polymerase II. *Mol. Cell* 11, 1549-1561.
144. Fairley, J.A., Evans, R., Hawkes, N.A., and 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.
145. Fan, X., Chou, D.M., and Struhl, K. (2006). Activator-specific recruitment of Mediator in vivo. *Nat. Struct. Mol. Biol.* 13, 117-120.
146. Fang, S.M., and Burton, Z.F. (1996). RNA polymerase II-associated protein (RAP) 74 binds transcription factor (TF) IIB and blocks TFIIB-RAP30 binding. *J. Biol. Chem.* 271, 11703-11709.
147. Faulkner, G.J., and Carninci, P. (2009). Altruistic functions for selfish DNA. *Cell Cycle* 8, 2895-2900.
148. Feaver, W.J., Gileadi, O., Li, Y., and Kornberg, R.D. (1991). CTD kinase associated with yeast RNA polymerase II initiation factor b. *Cell* 67, 1223-1230.
149. Feige, J.N., Lagouge, M., Canto, C., Strehle, A., Houten, S.M., Milne, J.C., Lambert, P.D., Matak, C., Elliott, P.J., and Auwerx, J. (2008). Specific SIRT1 activation mimics low energy levels and protects against diet-induced metabolic disorders by enhancing fat oxidation. *Cell Metab* 8, 347-358.
150. 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.
151. Fish, R.N., Ammerman, M.L., Davie, J.K., Lu, B.F., Pham, C., Howe, L., Ponticelli, A.S., and Kane, C.M. (2006). Genetic interactions between TFIIF and TFIIS. *Genetics* 173, 1871-1884.
152. Fish, R.N., and Kane, C.M. (2002). Promoting elongation with transcript cleavage stimulatory factors. *Biochim. Biophys. Acta* 1577, 287-307.
153. Flores, O., Ha, I., and Reinberg, D. (1990). Factors involved in specific transcription by mammalian RNA polymerase II. Purification and subunit composition of transcription factor IIF. *J. Biol. Chem.* 265, 5629-5634.
154. Forget, D., Langelier, M.F., Therien, C., Trinh, V., and Coulombe, B. (2004a). 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.
155. Forget, D., Langelier, M.F., Therien, C., Trinh, V., and Coulombe, B. (2004b). 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.
 156. Forget, D., Robert, F., Grondin, G., Burton, Z.F., Greenblatt, J., and 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.
 157. Fossat, N., Courtois, V., Chatelain, G., Brun, G., and Lamonerie, T. (2005). Alternative usage of Otx2 promoters during mouse development. *Dev. Dyn.* 233, 154-160.
 158. Fousteri, M., Vermeulen, W., van Zeeland, A.A., and Mullenders, L.H. (2006). Cockayne syndrome A and B proteins differentially regulate recruitment of chromatin remodeling and repair factors to stalled RNA polymerase II in vivo. *Mol. Cell* 23, 471-482.
 159. Freiman, R.N., Albright, S.R., Chu, L.E., Zheng, S., Liang, H.E., Sha, W.C., and Tjian, R. (2002). Redundant role of tissue-selective TAF(II)105 in B lymphocytes. *Mol. Cell Biol.* 22, 6564-6572.
 160. Freire-Picos, M.A., Krishnamurthy, S., Sun, Z.W., and Hampsey, M. (2005). Evidence that the Tfg1/Tfg2 dimer interface of TFIIF lies near the active center of the RNA polymerase II initiation complex. *Nucleic Acids Res.* 33, 5045-5052.
 161. Fu, J., Gnatt, A.L., Bushnell, D.A., Jensen, G.J., Thompson, N.E., Burgess, R.R., David, P.R., and Kornberg, R.D. (1999). Yeast RNA polymerase II at 5 Å resolution. *Cell* 98, 799-810.
 162. Fu, J., Yoon, H.G., Qin, J., and Wong, J. (2007). Regulation of P-TEFb elongation complex activity by CDK9 acetylation. *Mol. Cell Biol.* 27, 4641-4651.
 163. Fuda, N.J., Ardehali, M.B., and Lis, J.T. (2009). Defining mechanisms that regulate RNA polymerase II transcription in vivo. *Nature* 461, 186-192.
 164. Fujinaga, K., Irwin, D., Huang, Y., Taube, R., Kurosu, T., and Peterlin, B.M. (2004). Dynamics of human immunodeficiency virus transcription: P-TEFb phosphorylates RD and dissociates negative effectors from the transactivation response element. *Mol. Cell Biol.* 24, 787-795.

165. Gagnon, D., Joubert, S., Senechal, H., Fradet-Turcotte, A., Torre, S., and Archambault, J. (2009). Proteasomal degradation of the papillomavirus E2 protein is inhibited by overexpression of bromodomain-containing protein 4. *J. Virol.* 83, 4127-4139.
166. Galande, S., Purbey, P.K., Notani, D., and Kumar, P.P. (2007). The third dimension of gene regulation: organization of dynamic chromatin loopscape by SATB1. *Curr. Opin. Genet. Dev.* 17, 408-414.
167. Gangloff, Y.G., Romier, C., Thuault, S., Werten, S., and Davidson, I. (2001). The histone fold is a key structural motif of transcription factor TFIID. *Trends Biochem. Sci.* 26, 250-257.
168. Garbett, K.A., Tripathi, M.K., Cencki, B., Layer, J.H., and Weil, P.A. (2007). Yeast TFIID serves as a coactivator for Rap1p by direct protein-protein interaction. *Mol Cell Biol* 27, 297-311.
169. Garcia-Gimeno, M.A., and Struhl, K. (2000). Aca1 and Aca2, ATF/CREB activators in *Saccharomyces cerevisiae*, are important for carbon source utilization but not the response to stress. *Mol. Cell Biol.* 20, 4340-4349.
170. Garriga, J., Bhattacharya, S., Calbo, J., Marshall, R.M., Truongcao, M., Haines, D.S., and Grana, X. (2003). CDK9 is constitutively expressed throughout the cell cycle, and its steady-state expression is independent of SKP2. *Mol Cell Biol* 23, 5165-5173.
171. Garriga, J., and Grana, X. (2004). Cellular control of gene expression by T-type cyclin/CDK9 complexes. *Gene* 337, 15-23.
172. Ge, H., and Roeder, R.G. (1994). The high mobility group protein HMG1 can reversibly inhibit class II gene transcription by interaction with the TATA-binding protein. *J. Biol. Chem.* 269, 17136-17140.
173. Geiger, J.H., Hahn, S., Lee, S., and Sigler, P.B. (1996). Crystal structure of the yeast TFIIA/TBP/DNA complex. *Science* 272, 830-836.
174. Georgieva, S., Kirschner, D.B., Jagla, T., Nabirochkina, E., Hanke, S., Schenkel, H., de Lorenzo, C., Sinha, P., Jagla, K., Mechler, B., and Tora, L. (2000). Two novel *Drosophila* TAF(II)s have homology with human TAF(II)30 and are differentially regulated during development. *Mol. Cell Biol.* 20, 1639-1648.
175. 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.

176. Geserick, C., Meyer, H.A., and Haendler, B. (2005). The role of DNA response elements as allosteric modulators of steroid receptor function. *Mol. Cell Endocrinol.* 236, 1-7.
177. Ghavi-Helm, Y., Michaut, M., Acker, J., Aude, J.C., Thuriaux, P., Werner, M., and Soutourina, J. (2008). Genome-wide location analysis reveals a role of TFIIS in RNA polymerase III transcription. *Genes Dev.* 22, 1934-1947.
178. 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., and Johnston, M. (2002). Functional profiling of the *Saccharomyces cerevisiae* genome. *Nature* 418, 387-391.
179. 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., and Vermeulen, W. (2004). A new, tenth subunit of TFIIF is responsible for the DNA repair syndrome trichothiodystrophy group A. *Nat. Genet.* 36, 714-719.
180. Gilchrist, A., Au, C.E., Hiding, J., Bell, A.W., Fernandez-Rodriguez, J., Lesimple, S., Nagaya, H., Roy, L., Gosline, S.J., Hallett, M., Paiement, J., Kearney, R.E., Nilsson, T., and Bergeron, J.J. (2006). Quantitative proteomics analysis of the secretory pathway. *Cell* 127, 1265-1281.
181. Gilchrist, D.A., Fargo, D.C., and Adelman, K. (2009). Using ChIP-chip and ChIP-seq to study the regulation of gene expression: genome-wide localization studies reveal widespread regulation of transcription elongation. *Methods* 48, 398-408.
182. Glossop, J.A., Dafforn, T.R., and 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.

183. Glover-Cutter, K., Larochelle, S., Erickson, B., Zhang, C., Shokat, K., Fisher, R.P., and Bentley, D.L. (2009). TFIIF-associated Cdk7 kinase functions in phosphorylation of C-terminal domain Ser7 residues, promoter-proximal pausing, and termination by RNA polymerase II. *Mol. Cell Biol.* 29, 5455-5464.
184. Gnatt, A., Fu, J., and Kornberg, R.D. (1997). Formation and crystallization of yeast RNA polymerase II elongation complexes. *J. Biol. Chem.* 272, 30799-30805.
185. 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.
186. Goldberg, A.L. (2003). Protein degradation and protection against misfolded or damaged proteins. *Nature* 426, 895-899.
187. Goler-Baron, V., Selitrennik, M., Barkai, O., Haimovich, G., Lotan, R., and Choder, M. (2008). Transcription in the nucleus and mRNA decay in the cytoplasm are coupled processes. *Genes Dev.* 22, 2022-2027.
188. Grana, X., De, L.A., Sang, N., Fu, Y., Claudio, P.P., Rosenblatt, J., Morgan, D.O., and Giordano, A. (1994). PITALRE, a nuclear CDC2-related protein kinase that phosphorylates the retinoblastoma protein in vitro. *Proc. Natl. Acad. Sci. U. S. A* 91, 3834-3838.
189. Grandemange, S., Schaller, S., Yamano, S., Du, M.S., Shpakovski, G.V., Mattei, M.G., Kedinger, C., and Vigneron, M. (2001). A human RNA polymerase II subunit is encoded by a recently generated multigene family. *BMC. Mol. Biol.* 2, 14.
190. Greive, S.J., and von Hippel, P.H. (2005). Thinking quantitatively about transcriptional regulation. *Nat. Rev. Mol. Cell Biol.* 6, 221-232.
191. Grob, P., Cruse, M.J., Inouye, C., Peris, M., Penczek, P.A., Tjian, R., and Nogales, E. (2006). Cryo-electron microscopy studies of human TFIID: conformational breathing in the integration of gene regulatory cues. *Structure.* 14, 511-520.
192. Gromoller, A., and Lehming, N. (2000). Srb7p is essential for the activation of a subset of genes. *FEBS Lett.* 484, 48-54.
193. Grozinger, C.M., and Schreiber, S.L. (2002). Deacetylase enzymes: biological functions and the use of small-molecule inhibitors. *Chem. Biol.* 9, 3-16.

194. Gu, W., and Reines, D. (1995). Variation in the size of nascent RNA cleavage products as a function of transcript length and elongation competence. *J. Biol. Chem.* 270, 30441-30447.
195. Guenther, M.G., Levine, S.S., Boyer, L.A., Jaenisch, R., and Young, R.A. (2007). A chromatin landmark and transcription initiation at most promoters in human cells. *Cell* 130, 77-88.
196. Guglielmi, B., Soutourina, J., Esnault, C., and Werner, M. (2007). TFIIS elongation factor and Mediator act in conjunction during transcription initiation in vivo. *Proc. Natl. Acad. Sci. U. S. A* 104, 16062-16067.
197. Guzman, E., and Lis, J.T. (1999). Transcription factor TFIID is required for promoter melting in vivo. *Mol. Cell Biol.* 19, 5652-5658.
198. 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.
199. Hahn, S. (2004). Structure and mechanism of the RNA polymerase II transcription machinery. *Nat. Struct. Mol Biol* 11, 394-403.
200. Hampsey, M. (1998). Molecular genetics of the RNA polymerase II general transcriptional machinery. *Microbiol. Mol. Biol. Rev.* 62, 465-503.
201. Han, S., Xie, W., Hammes, S.R., and DeJong, J. (2003). Expression of the germ cell-specific transcription factor ALF in *Xenopus* oocytes compensates for translational inactivation of the somatic factor TFIIA. *J. Biol. Chem.* 278, 45586-45593.
202. Hansen, S.K., Takada, S., Jacobson, R.H., Lis, J.T., and Tjian, R. (1997). Transcription properties of a cell type-specific TATA-binding protein, TRF. *Cell* 91, 71-83.
203. Hawkes, N.A., Evans, R., and Roberts, S.G. (2000). The conformation of the transcription factor TFIIB modulates the response to transcriptional activators in vivo. *Curr. Biol.* 10, 273-276.
204. Hawkes, N.A., and 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.
205. Hazbun, T.R., Malmstrom, L., Anderson, S., Graczyk, B.J., Fox, B., Riffle, M., Sundin, B.A., Aranda, J.D., McDonald, W.H., Chiu, C.H., Snysman, B.E., Bradley, P., Muller, E.G., Fields, S., Baker, D., Yates, J.R., III, and Davis, T.N. (2003). Assigning function to yeast proteins by integration of technologies. *Mol. Cell* 12, 1353-1365.

206. Hegde, P., Qi, R., Abernathy, K., Gay, C., Dharap, S., Gaspard, R., Hughes, J.E., Snesrud, E., Lee, N., and Quackenbush, J. (2000). A concise guide to cDNA microarray analysis. *Biotechniques* 29, 548-4, 556.
207. Heinrich, P.C., Behrmann, I., Muller-Newen, G., Schaper, F., and Graeve, L. (1998). Interleukin-6-type cytokine signalling through the gp130/Jak/STAT pathway. *Biochem. J.* 334 (Pt 2), 297-314.
208. Hemming, S.A., Jansma, D.B., Macgregor, P.F., Goryachev, A., Friesen, J.D., and Edwards, A.M. (2000). RNA polymerase II subunit Rpb9 regulates transcription elongation in vivo. *J. Biol. Chem.* 275, 35506-35511.
209. Henderson, M.J., Munoz, M.A., Saunders, D.N., Clancy, J.L., Russell, A.J., Williams, B., Pappin, D., Khanna, K.K., Jackson, S.P., Sutherland, R.L., and Watts, C.K. (2006). EDD mediates DNA damage-induced activation of CHK2. *J. Biol Chem.* 281, 39990-40000.
210. Henderson, M.J., Russell, A.J., Hird, S., Munoz, M., Clancy, J.L., Lehrbach, G.M., Calanni, S.T., Jans, D.A., Sutherland, R.L., and Watts, C.K. (2002). EDD, the human hyperplastic discs protein, has a role in progesterone receptor coactivation and potential involvement in DNA damage response. *J. Biol Chem.* 277, 26468-26478.
211. Herbert, K.M., Greenleaf, W.J., and Block, S.M. (2008). Single-molecule studies of RNA polymerase: motoring along. *Annu. Rev. Biochem.* 77, 149-176.
212. Hershko, A. (2005). The ubiquitin system for protein degradation and some of its roles in the control of the cell division cycle. *Cell Death. Differ.* 12, 1191-1197.
213. Hershko, A., and Ciechanover, A. (1998). The ubiquitin system. *Annu. Rev. Biochem.* 67, 425-479.
214. Hildmann, C., Riester, D., and Schwienhorst, A. (2007). Histone deacetylases--an important class of cellular regulators with a variety of functions. *Appl. Microbiol. Biotechnol.* 75, 487-497.
215. Hirtreiter, A., Grohmann, D., and Werner, F. (2010). Molecular mechanisms of RNA polymerase--the F/E (RPB4/7) complex is required for high processivity in vitro. *Nucleic Acids Res.* 38, 585-596.
216. Ho, C.K., and Shuman, S. (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.

217. Hodo, H.G., III, and 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.
218. Holmes, M.C., and Tjian, R. (2000). Promoter-selective properties of the TBP-related factor TRF1. *Science* 288, 867-870.
219. Holstege, F.C., Fiedler, U., and Timmers, H.T. (1997). Three transitions in the RNA polymerase II transcription complex during initiation. *EMBO J.* 16, 7468-7480.
220. Holstege, F.C., Jennings, E.G., Wyrick, J.J., Lee, T.I., Hengartner, C.J., Green, M.R., Golub, T.R., Lander, E.S., and Young, R.A. (1998). Dissecting the regulatory circuitry of a eukaryotic genome. *Cell* 95, 717-728.
221. Holstege, F.C., Tantin, D., Carey, M., van der Vliet, P.C., and 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.
222. Holstege, F.C., van der Vliet, P.C., and Timmers, H.T. (1996). Opening of an RNA polymerase II promoter occurs in two distinct steps and requires the basal transcription factors IIE and IIH. *EMBO J.* 15, 1666-1677.
223. Honda, M., Park, J., Pugh, R.A., Ha, T., and Spies, M. (2009). Single-molecule analysis reveals differential effect of ssDNA-binding proteins on DNA translocation by XPD helicase. *Mol. Cell* 35, 694-703.
224. Honda, Y., Tojo, M., Matsuzaki, K., Anan, T., Matsumoto, M., Ando, M., Saya, H., and Nakao, M. (2002). Cooperation of HECT-domain ubiquitin ligase hHYD and DNA topoisomerase II-binding protein for DNA damage response. *J. Biol. Chem.* 277, 3599-3605.
225. Hou, S.Y., Wu, S.Y., Zhou, T., Thomas, M.C., and Chiang, C.M. (2000). Alleviation of human papillomavirus E2-mediated transcriptional repression via formation of a TATA binding protein (or TFIID)-TFIIB-RNA polymerase II-TFIIF preinitiation complex. *Mol. Cell Biol.* 20, 113-125.
226. Hou, T., Ray, S., and Brasier, A.R. (2007). The functional role of an interleukin 6-inducible CDK9/STAT3 complex in human gamma-fibrinogen gene expression. *J. Biol. Chem.* 282, 37091-37102.

227. Hsu, F., Pringle, T.H., Kuhn, R.M., Karolchik, D., Diekhans, M., Haussler, D., and Kent, W.J. (2005). The UCSC Proteome Browser. *Nucleic Acids Res.* 33, D454-D458.
228. Huber, A.H., Nelson, W.J., and Weis, W.I. (1997). Three-dimensional structure of the armadillo repeat region of beta-catenin. *Cell* 90, 871-882.
229. Hubert, J.C., Guyonvarch, A., Kammerer, B., Exinger, F., Liljelund, P., and Lacroute, F. (1983). Complete sequence of a eukaryotic regulatory gene. *EMBO J.* 2, 2071-2073.
230. Huh, W.K., Falvo, J.V., Gerke, L.C., Carroll, A.S., Howson, R.W., Weissman, J.S., and O'Shea, E.K. (2003). Global analysis of protein localization in budding yeast. *Nature* 425, 686-691.
231. Huisinga, K.L., and Pugh, B.F. (2004a). A genome-wide housekeeping role for TFIID and a highly regulated stress-related role for SAGA in *Saccharomyces cerevisiae*. *Mol. Cell* 13, 573-585.
232. Huisinga, K.L., and Pugh, B.F. (2004b). A genome-wide housekeeping role for TFIID and a highly regulated stress-related role for SAGA in *Saccharomyces cerevisiae*. *Mol. Cell* 13, 573-585.
233. Hymowitz, S.G., and Wertz, I.E. (2010). A20: from ubiquitin editing to tumour suppression. *Nat. Rev. Cancer* 10, 332-341.
234. Imai, S., Armstrong, C.M., Kaerberlein, M., and Guarente, L. (2000). Transcriptional silencing and longevity protein Sir2 is an NAD-dependent histone deacetylase. *Nature* 403, 795-800.
235. Imhof, A., Yang, X.J., Ogryzko, V.V., Nakatani, Y., Wolffe, A.P., and Ge, H. (1997). Acetylation of general transcription factors by histone acetyltransferases. *Curr. Biol.* 7, 689-692.
236. 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.
237. Isogai, Y., Takada, S., Tjian, R., and Keles, S. (2007). Novel TRF1/BRF target genes revealed by genome-wide analysis of *Drosophila* Pol III transcription. *EMBO J.* 26, 79-89.
238. Ito, T., Arimitsu, N., Takeuchi, M., Kawamura, N., Nagata, M., Saso, K., Akimitsu, N., Hamamoto, H., Natori, S., Miyajima, A., and Sekimizu, K. (2006). Transcription elongation factor S-II is required for definitive hematopoiesis. *Mol Cell Biol* 26, 3194-3203.

239. Ito, T., Chiba, T., Ozawa, R., Yoshida, M., Hattori, M., and Sakaki, Y. (2001). A comprehensive two-hybrid analysis to explore the yeast protein interactome. *Proc. Natl. Acad. Sci. U. S. A* 98, 4569-4574.
240. Izban, M.G., and Luse, D.S. (1992). 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.
241. Jackson-Fisher, A.J., Chitikila, C., Mitra, M., and Pugh, B.F. (1999). A role for TBP dimerization in preventing unregulated gene expression. *Mol. Cell* 3, 717-727.
242. Jang, M.K., Mochizuki, K., Zhou, M., Jeong, H.S., Brady, J.N., and Ozato, K. (2005). The bromodomain protein Brd4 is a positive regulatory component of P-TEFb and stimulates RNA polymerase II-dependent transcription. *Mol Cell* 19, 523-534.
243. Jasiak, A.J., Hartmann, H., Karakasili, E., Kalocsay, M., Flatley, A., Kremmer, E., Strasser, K., Martin, D.E., Soding, J., and Cramer, P. (2008). Genome-associated RNA polymerase II includes the dissociable Rpb4/7 subcomplex. *J. Biol. Chem.* 283, 26423-26427.
244. Jawhari, A., Laine, J.P., Dubaele, S., Lamour, V., Poterszman, A., Coin, F., Moras, D., and Egly, J.M. (2002). p52 Mediates XPB Function within the Transcription/Repair Factor TFIIH. *J. Biol. Chem.* 277, 31761-31767.
245. Jensen, G.J., Meredith, G., Bushnell, D.A., and Kornberg, R.D. (1998). Structure of wild-type yeast RNA polymerase II and location of Rpb4 and Rpb7. *EMBO J.* 17, 2353-2358.
246. Jeon, C., and Agarwal, K. (1996). Fidelity of RNA polymerase II transcription controlled by elongation factor TFIIS. *Proc. Natl. Acad. Sci. U. S. A* 93, 13677-13682.
247. Jeronimo, C., Forget, D., Bouchard, A., Li, Q., Chua, G., Poitras, C., Therien, C., Bergeron, D., Bourassa, S., Greenblatt, J., Chabot, B., Poirier, G.G., Hughes, T.R., Blanchette, M., Price, D.H., and Coulombe, B. (2007). Systematic analysis of the protein interaction network for the human transcription machinery reveals the identity of the 7SK capping enzyme. *Mol Cell* 27, 262-274.
248. 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.
249. Jiang, Y., Triezenberg, S.J., and Gralla, J.D. (1994). Defective transcriptional activation by diverse VP16 mutants associated with a common inability to form open promoter complexes. *J. Biol. Chem.* 269, 5505-5508.
 250. Juven-Gershon, T., Hsu, J.Y., and Kadonaga, J.T. (2006). Perspectives on the RNA polymerase II core promoter. *Biochem. Soc. Trans.* 34, 1047-1050.
 251. Kadonaga, J.T. (2004). Regulation of RNA polymerase II transcription by sequence-specific DNA binding factors. *Cell* 116, 247-257.
 252. Kamenisch, Y., and Berneburg, M. (2009). Progeroid syndromes and UV-induced oxidative DNA damage. *J. Investig. Dermatol. Symp. Proc.* 14, 8-14.
 253. Kanazawa, S., Okamoto, T., and Peterlin, B.M. (2000). Tat competes with CIITA for the binding to P-TEFb and blocks the expression of MHC class II genes in HIV infection. *Immunity*. 12, 61-70.
 254. Kanazawa, S., Soucek, L., Evan, G., Okamoto, T., and Peterlin, B.M. (2003). c-Myc recruits P-TEFb for transcription, cellular proliferation and apoptosis. *Oncogene* 22, 5707-5711.
 255. Kapanidis, A.N., Margeat, E., Ho, S.O., Kortkhonjia, E., Weiss, S., and Ebright, R.H. (2006). Initial transcription by RNA polymerase proceeds through a DNA-scrunching mechanism. *Science* 314, 1144-1147.
 256. Kapranov, P., Cheng, J., Dike, S., Nix, D.A., Dutttagupta, R., Willingham, A.T., Stadler, P.F., Hertel, J., Hackermuller, J., Hofacker, I.L., Bell, I., Cheung, E., Drenkow, J., Dumais, E., Patel, S., Helt, G., Ganesh, M., Ghosh, S., Piccolboni, A., Sementchenko, V., Tammana, H., and Gingeras, T.R. (2007). RNA maps reveal new RNA classes and a possible function for pervasive transcription. *Science* 316, 1484-1488.
 257. Kelly, W.G., Dahmus, M.E., and Hart, G.W. (1993). RNA polymerase II is a glycoprotein. Modification of the COOH-terminal domain by O-GlcNAc. *J. Biol. Chem.* 268, 10416-10424.
 258. Keogh, M.C., Podolny, V., and Buratowski, S. (2003). Bur1 kinase is required for efficient transcription elongation by RNA polymerase II. *Mol. Cell Biol.* 23, 7005-7018.

259. Keriel, A., Stary, A., Sarasin, A., Rochette-Egly, C., and Egly, J.M. (2002). XPD Mutations Prevent TFIID-Dependent Transactivation by Nuclear Receptors and Phosphorylation of RARalpha. *Cell* 109, 125-135.
260. Kerppola, T.K., and Kane, C.M. (1990). Analysis of the signals for transcription termination by purified RNA polymerase II. *Biochemistry* 29, 269-278.
261. 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.
262. Kettenberger, H., Armache, K.J., and Cramer, P. (2004). Complete RNA Polymerase II Elongation Complex Structure and Its Interactions with NTP and TFIIS. *Mol Cell* 16, 955-965.
263. Khaleghpour, K., Svitkin, Y.V., Craig, A.W., DeMaria, C.T., Deo, R.C., Burley, S.K., and Sonenberg, N. (2001). Translational repression by a novel partner of human poly(A) binding protein, Paip2. *Mol. Cell* 7, 205-216.
264. Khapersky, D.A., Ammerman, M.L., Majovski, R.C., and Ponticelli, A.S. (2008). Functions of *Saccharomyces cerevisiae* TFIIF during transcription start site utilization. *Mol. Cell Biol.* 28, 3757-3766.
265. Kheradmand, K.S., Solaimani, K.P., Farahbakhshian, E., Pourfarzad, F., von, L.M., and Verrijzer, C.P. (2009). EZH2-dependent chromatin looping controls INK4a and INK4b, but not ARF, during human progenitor cell differentiation and cellular senescence. *Epigenetics. Chromatin.* 2, 16.
266. Kiernan, R.E., Emiliani, S., Nakayama, K., Castro, A., Labbe, J.C., Lorca, T., Nakayama, K.K., and Benkirane, M. (2001). Interaction between cyclin T1 and SCF(SKP2) targets CDK9 for ubiquitination and degradation by the proteasome. *Mol Cell Biol* 21, 7956-7970.
267. Killeen, M., Coulombe, B., and 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.
268. Kim, B., Nesvizhskii, A.I., Rani, P.G., Hahn, S., Aebersold, R., and Ranish, J.A. (2007). The transcription elongation factor TFIIS is a component of RNA polymerase II preinitiation complexes. *Proc. Natl. Acad. Sci. U. S. A* 104, 16068-16073.

269. Kim, J.B., and 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.
270. Kim, J.L., Nikolov, D.B., and Burley, S.K. (1993a). Co-crystal structure of TBP recognizing the minor groove of a TATA element. *Nature* 365, 520-527.
271. Kim, M., Vasiljeva, L., Rando, O.J., Zhelkovsky, A., Moore, C., and Buratowski, S. (2006). Distinct pathways for snoRNA and mRNA termination. *Mol Cell* 24, 723-734.
272. Kim, T.H., Barrera, L.O., Zheng, M., Qu, C., Singer, M.A., Richmond, T.A., Wu, Y., Green, R.D., and Ren, B. (2005a). A high-resolution map of active promoters in the human genome. *Nature* 436, 876-880.
273. Kim, T.H., Barrera, L.O., Zheng, M., Qu, C., Singer, M.A., Richmond, T.A., Wu, Y., Green, R.D., and Ren, B. (2005b). A high-resolution map of active promoters in the human genome. *Nature* 436, 876-880.
274. Kim, T.K., Ebricht, R.H., and Reinberg, D. (2000). Mechanism of ATP-dependent promoter melting by transcription factor IIH. *Science* 288, 1418-1422.
275. Kim, Y., Geiger, J.H., Hahn, S., and Sigler, P.B. (1993b). Crystal structure of a yeast TBP/TATA-box complex. *Nature* 365, 512-520.
276. Kim, Y.J., Bjorklund, S., Li, Y., Sayre, M.H., and 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.
277. Kimura, K., Wakamatsu, A., Suzuki, Y., Ota, T., Nishikawa, T., Yamashita, R., Yamamoto, J., Sekine, M., Tsuritani, K., Wakaguri, H., Ishii, S., Sugiyama, T., Saito, K., Isono, Y., Irie, R., Kushida, N., Yoneyama, T., Otsuka, R., Kanda, K., Yokoi, T., Kondo, H., Wagatsuma, M., Murakawa, K., Ishida, S., Ishibashi, T., Takahashi-Fujii, A., Tanase, T., Nagai, K., Kikuchi, H., Nakai, K., Isogai, T., and Sugano, S. (2006). Diversification of transcriptional modulation: large-scale identification and characterization of putative alternative promoters of human genes. *Genome Res.* 16, 55-65.
278. Kimura, M., and Ishihama, A. (2000). Involvement of multiple subunit-subunit contacts in the assembly of RNA polymerase II. *Nucleic Acids Res.* 28, 952-959.

279. Kimura, M., Suzuki, H., and 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.
280. Kiss, T. (2006). SnoRNP biogenesis meets Pre-mRNA splicing. *Mol Cell* 23, 775-776.
281. Kiss, T., Fayet, E., Jady, B.E., Richard, P., and Weber, M. (2006). Biogenesis and intranuclear trafficking of human box C/D and H/ACA RNPs. *Cold Spring Harb. Symp. Quant. Biol.* 71, 407-417.
282. Knuesel, M.T., Meyer, K.D., Bernecky, C., and Taatjes, D.J. (2009). The human CDK8 subcomplex is a molecular switch that controls Mediator coactivator function. *Genes Dev.* 23, 439-451.
283. Kobayashi, N., Boyer, T.G., and 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.
284. Kobor, M.S., Archambault, J., Lester, W., Holstege, F.C., Gileadi, O., Jansma, D.B., Jennings, E.G., Kouyoumdjian, F., Davidson, A.R., Young, R.A., and Greenblatt, J. (1999). An unusual eukaryotic protein phosphatase required for transcription by RNA polymerase II and CTD dephosphorylation in *S. cerevisiae*. *Mol. Cell* 4, 55-62.
285. Kobor, M.S., Simon, L.D., Omichinski, J., Zhong, G., Archambault, J., and Greenblatt, J. (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.
286. Kokubo, T., Swanson, M.J., Nishikawa, J.I., Hinnebusch, A.G., and Nakatani, Y. (1998). The yeast TAF145 inhibitory domain and TFIIA competitively bind to TATA-binding protein. *Mol. Cell Biol.* 18, 1003-1012.
287. Kokura, K., Kaul, S.C., Wadhwa, R., Nomura, T., Khan, M.M., Shinagawa, T., Yasukawa, T., Colmenares, C., and Ishii, S. (2001). The Ski protein family is required for MeCP2-mediated transcriptional repression. *J. Biol. Chem.* 276, 34115-34121.
288. Komarnitsky, P., Cho, E.J., and Buratowski, S. (2000). Different phosphorylated forms of RNA polymerase II and associated mRNA processing factors during transcription. *Genes Dev.* 14, 2452-2460.

289. Kornberg, R.D. (2005). Mediator and the mechanism of transcriptional activation. *Trends Biochem. Sci.* 30, 235-239.
290. Kostrewa, D., Zeller, M.E., Armache, K.J., Seizl, M., Leike, K., Thomm, M., and Cramer, P. (2009). RNA polymerase II-TFIIB structure and mechanism of transcription initiation. *Nature*.
291. Koyama, H., Ito, T., Nakanishi, T., and Sekimizu, K. (2007). Stimulation of RNA polymerase II transcript cleavage activity contributes to maintain transcriptional fidelity in yeast. *Genes Cells* 12, 547-559.
292. Kozlov, G., Nguyen, L., Lin, T., De, C.G., Park, M., and Gehring, K. (2007). Structural basis of ubiquitin recognition by the ubiquitin-associated (UBA) domain of the ubiquitin ligase EDD. *J. Biol. Chem.* 282, 35787-35795.
293. Krishnamurthy, S., He, X., Reyes-Reyes, M., Moore, C., and Hampsey, M. (2004). Ssu72 Is an RNA polymerase II CTD phosphatase. *Mol Cell* 14, 387-394.
294. Krogan, N.J., Kim, M., Ahn, S.H., Zhong, G., Kobor, M.S., Cagney, G., Emili, A., Shilatifard, A., Buratowski, S., and Greenblatt, J.F. (2002). RNA polymerase II elongation factors of *Saccharomyces cerevisiae*: a targeted proteomics approach. *Mol. Cell Biol.* 22, 6979-6992.
295. Krueger, B.J., Jeronimo, C., Roy, B.B., Bouchard, A., Barrandon, C., Byers, S.A., Searcey, C.E., Cooper, J.J., Bensaude, O., Cohen, E.A., Coulombe, B., and Price, D.H. (2008). LARP7 is a stable component of the 7SK snRNP while P-TEFb, HEXIM1 and hnRNP A1 are reversibly associated. *Nucleic Acids Res.* 36, 2219-2229.
296. Kuldell, N.H., and 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.
297. Kumar, K.P., Akoulitchev, S., and 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.
298. Kuras, L., Kosa, P., Mencia, M., and Struhl, K. (2000). TAF-Containing and TAF-independent forms of transcriptionally active TBP in vivo. *Science* 288, 1244-1248.
299. Kurosu, T., and Peterlin, B.M. (2004). VP16 and ubiquitin; binding of P-TEFb via its activation domain and ubiquitin facilitates elongation of transcription of target genes. *Curr. Biol* 14, 1112-1116.

300. Kutach, A.K., and 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.
301. LaCava, J., Houseley, J., Saveanu, C., Petfalski, E., Thompson, E., Jacquier, A., and Tollervey, D. (2005). RNA degradation by the exosome is promoted by a nuclear polyadenylation complex. *Cell* 121, 713-724.
302. Lacy, E.R., Filippov, I., Lewis, W.S., Otieno, S., Xiao, L., Weiss, S., Hengst, L., and Kriwacki, R.W. (2004). p27 binds cyclin-CDK complexes through a sequential mechanism involving binding-induced protein folding. *Nat. Struct. Mol Biol* 11, 358-364.
303. Lagrange, T., Kapanidis, A.N., Tang, H., Reinberg, D., and Ebright, R.H. (1998). New core promoter element in RNA polymerase II-dependent transcription: sequence-specific DNA binding by transcription factor IIB. *Genes Dev.* 12, 34-44.
304. Lagrange, T., Kim, T.K., Orphanides, G., Ebright, Y.W., Ebright, R.H., and 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.
305. Langelier, M.F., Baali, D., Trinh, V., Greenblatt, J., Archambault, J., and Coulombe, B. (2005). The highly conserved glutamic acid 791 of Rpb2 is involved in binding of NTP and Mg(B) in the active center of human RNA polymerase II. *Nucleic Acids Res.* 33, 2629-2639.
306. 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.
307. Langelier, M.F., Trinh, V., and Coulombe, B. (2002). Gros plan sur l'ARN polymérase II. *Medecine Sciences* 18, 211-216.
308. Larochelle, S., Pandur, J., Fisher, R.P., Salz, H.K., and Suter, B. (1998). Cdk7 is essential for mitosis and for in vivo Cdk-activating kinase activity. *Genes Dev.* 12, 370-381.
309. Lee, D.K., Duan, H.O., and Chang, C. (2001). Androgen receptor interacts with the positive elongation factor P-TEFb and enhances the efficiency of transcriptional elongation. *J. Biol. Chem.* 276, 9978-9984.
310. Lee, T.I., Wyrick, J.J., Koh, S.S., Jennings, E.G., Gadbois, E.L., and Young, R.A. (1998). Interplay of positive and negative regulators in

- transcription initiation by RNA polymerase II holoenzyme. *Mol. Cell Biol.* 18, 4455-4462.
311. Lee, T.I., and Young, R.A. (2000). Transcription of eukaryotic protein-coding genes. *Annu. Rev. Genet.* 34, 77-137.
 312. Lee, Y.C., Park, J.M., Min, S., Han, S.J., and Kim, Y.J. (1999). An activator binding module of yeast RNA polymerase II holoenzyme. *Mol. Cell Biol.* 19, 2967-2976.
 313. Lemon, B., and Tjian, R. (2000). Orchestrated response: a symphony of transcription factors for gene control. *Genes Dev.* 14, 2551-2569.
 314. Lennon, J.C., III, Wind, M., Saunders, L., Hock, M.B., and 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.
 315. Leurent, C., Sanders, S.L., Demeny, M.A., Garbett, K.A., Ruhlmann, C., Weil, P.A., Tora, L., and Schultz, P. (2004). Mapping key functional sites within yeast TFIID. *EMBO J.* 23, 719-727.
 316. Levanon, D., Glusman, G., Bettoun, D., Ben-Asher, E., Negreanu, V., Bernstein, Y., Harris-Cerruti, C., Brenner, O., Eilam, R., Lotem, J., Fainaru, O., Goldenberg, D., Pozner, A., Woolf, E., Xiao, C., Yarmus, M., and Groner, Y. (2003). Phylogenesis and regulated expression of the RUNT domain transcription factors RUNX1 and RUNX3. *Blood Cells Mol. Dis.* 30, 161-163.
 317. Li, Q., Price, J.P., Byers, S.A., Cheng, D., Peng, J., and Price, D.H. (2005). Analysis of the large inactive P-TEFb complex indicates that it contains one 7SK molecule, a dimer of HEXIM1 or HEXIM2, and two P-TEFb molecules containing Cdk9 phosphorylated at threonine 186. *J. Biol Chem.* 280, 28819-28826.
 318. Li, S., and Smerdon, M.J. (2002). Rpb4 and Rpb9 mediate subpathways of transcription-coupled DNA repair in *Saccharomyces cerevisiae*. *EMBO J.* 21, 5921-5929.
 319. Li, X.Y., Bhaumik, S.R., and Green, M.R. (2000). Distinct classes of yeast promoters revealed by differential TAF recruitment. *Science* 288, 1242-1244.
 320. Licatalosi, D.D., Geiger, G., Minet, M., Schroeder, S., Cilli, K., McNeil, J.B., and Bentley, D.L. (2002). Functional Interaction of Yeast Pre-mRNA 3' End Processing Factors with RNA Polymerase II. *Mol. Cell* 9, 1101-1111.

321. Lim, C.Y., Santoso, B., Boulay, T., Dong, E., Ohler, U., and Kadonaga, J.T. (2004). The MTE, a new core promoter element for transcription by RNA polymerase II. *Genes Dev.* 18, 1606-1617.
322. Lin, Y.C., Choi, W.S., and Gralla, J.D. (2005). TFIIH XPB mutants suggest a unified bacterial-like mechanism for promoter opening but not escape. *Nat. Struct. Mol Biol.*
323. Lindstrom, D.L., and Hartzog, G.A. (2001). Genetic interactions of Spt4-Spt5 and TFIIIS with the RNA polymerase II CTD and CTD modifying enzymes in *Saccharomyces cerevisiae*. *Genetics* 159, 487-497.
324. Liu, W.L., Coleman, R.A., Grob, P., King, D.S., Florens, L., Washburn, M.P., Geles, K.G., Yang, J.L., Ramey, V., Nogales, E., and Tjian, R. (2008). Structural changes in TAF4b-TFIID correlate with promoter selectivity. *Mol. Cell* 29, 81-91.
325. Liu, W.L., Coleman, R.A., Ma, E., Grob, P., Yang, J.L., Zhang, Y., Dailey, G., Nogales, E., and Tjian, R. (2009). Structures of three distinct activator-TFIID complexes. *Genes Dev.* 23, 1510-1521.
326. Lolli, G. (2009). Binding to DNA of the RNA-polymerase II C-terminal domain allows discrimination between Cdk7 and Cdk9 phosphorylation. *Nucleic Acids Res.* 37, 1260-1268.
327. Lorch, Y., Beve, J., Gustafsson, C.M., Myers, L.C., and Kornberg, R.D. (2000). Mediator-nucleosome interaction. *Mol. Cell* 6, 197-201.
328. Lu, H., Fisher, R.P., Bailey, P., and Levine, A.J. (1997). The CDK7-cycH-p36 complex of transcription factor IIIH phosphorylates p53, enhancing its sequence-specific DNA binding activity in vitro. *Mol. Cell Biol.* 17, 5923-5934.
329. Lu, H., Zawel, L., Fisher, L., Egly, J.M., and Reinberg, D. (1992). Human general transcription factor IIIH phosphorylates the C-terminal domain of RNA polymerase II. *Nature* 358, 641-645.
330. Lund, A.H., and van, L.M. (2004). Polycomb complexes and silencing mechanisms. *Curr. Opin. Cell Biol.* 16, 239-246.
331. Lykke-Andersen, S., and Jensen, T.H. (2007). Overlapping pathways dictate termination of RNA polymerase II transcription. *Biochimie* 89, 1177-1182.
332. Ma, D., Olave, I., Merino, A., and 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.
333. Ma, D., Watanabe, H., Mermelstein, F., Admon, A., Oguri, K., Sun, X., Wada, T., Imai, T., Shiroya, T., Reinberg, D., and . (1993). Isolation of a cDNA encoding the largest subunit of TFIIA reveals functions important for activated transcription. *Genes Dev.* 7, 2246-2257.
334. Maile, T., Kwoczyński, S., Katzenberger, R.J., Wassarman, D.A., and Sauer, F. (2004). TAF1 activates transcription by phosphorylation of serine 33 in histone H2B. *Science* 304, 1010-1014.
335. Maillet, I., Buhler, J.M., Sentenac, A., and Labarre, J. (1999). Rpb4p is necessary for RNA polymerase II activity at high temperature. *J. Biol. Chem.* 274, 22586-22590.
336. Makiniemi, M., Hillukkala, T., Tuusa, J., Reini, K., Vaara, M., Huang, D., Pospiech, H., Majuri, I., Westerling, T., Makela, T.P., and Syvaöja, J.E. (2001). BRCT domain-containing protein TopBP1 functions in DNA replication and damage response. *J. Biol. Chem.* 276, 30399-30406.
337. Malagon, F., Tong, A.H., Shafer, B.K., and Strathern, J.N. (2004). Genetic interactions of DST1 in *Saccharomyces cerevisiae* suggest a role of TFIIS in the initiation-elongation transition. *Genetics* 166, 1215-1227.
338. Malave, T.M., and Dent, S.Y. (2006). Transcriptional repression by Tup1-Ssn6. *Biochem. Cell Biol.* 84, 437-443.
339. Maldonado, E., Ha, I., Cortes, P., Weis, L., and Reinberg, D. (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.
340. Malik, S., and Roeder, R.G. (2000). Transcriptional regulation through Mediator-like coactivators in yeast and metazoan cells. *Trends Biochem. Sci.* 25, 277-283.
341. Malik, S., and Roeder, R.G. (2005). Dynamic regulation of pol II transcription by the mammalian Mediator complex. *Trends Biochem. Sci.* 30, 256-263.
342. Mansfield, E., Hersperger, E., Biggs, J., and Shearn, A. (1994). Genetic and molecular analysis of hyperplastic discs, a gene whose product is required for regulation of cell proliferation in *Drosophila melanogaster* imaginal discs and germ cells. *Dev. Biol.* 165, 507-526.

343. Margaritis, T., and Holstege, F.C. (2008). Poised RNA polymerase II gives pause for thought. *Cell* 133, 581-584.
344. Markert, A., Grimm, M., Martinez, J., Wiesner, J., Meyerhans, A., Meyuhas, O., Sickmann, A., and Fischer, U. (2008). The La-related protein LARP7 is a component of the 7SK ribonucleoprotein and affects transcription of cellular and viral polymerase II genes. *EMBO Rep.* 9, 569-575.
345. Marshall, N.F., and Price, D.H. (1995). Purification of P-TEFb, a transcription factor required for the transition into productive elongation. *J. Biol. Chem.* 270, 12335-12338.
346. Marshall, R.M., and Grana, X. (2006). Mechanisms controlling CDK9 activity. *Front Biosci.* 11, 2598-2613.
347. Martianov, I., Viville, S., and Davidson, I. (2002). RNA polymerase II transcription in murine cells lacking the TATA binding protein. *Science* 298, 1036-1039.
348. Martinez, E. (2002). Multi-protein complexes in eukaryotic gene transcription. *Plant Mol Biol* 50, 925-947.
349. Martinez, E., Palhan, V.B., Tjernberg, A., Lymar, E.S., Gamper, A.M., Kundu, T.K., Chait, B.T., and Roeder, R.G. (2001). Human staga complex is a chromatin-acetylating transcription coactivator that interacts with pre-mrna splicing and dna damage-binding factors in vivo. *Mol. Cell Biol.* 21, 6782-6795.
350. Marton, M.J., DeRisi, J.L., Bennett, H.A., Iyer, V.R., Meyer, M.R., Roberts, C.J., Stoughton, R., Burchard, J., Slade, D., Dai, H., Bassett, D.E., Jr., Hartwell, L.H., Brown, P.O., and Friend, S.H. (1998). Drug target validation and identification of secondary drug target effects using DNA microarrays. *Nat. Med.* 4, 1293-1301.
351. Mason, P.B., and Struhl, K. (2003). The FACT complex travels with elongating RNA polymerase II and is important for the fidelity of transcriptional initiation in vivo. *Mol. Cell Biol.* 23, 8323-8333.
352. Maston, G.A., Evans, S.K., and Green, M.R. (2006). Transcriptional regulatory elements in the human genome. *Annu. Rev. Genomics Hum. Genet.* 7, 29-59.
353. Matsuno, M., Kose, H., Okabe, M., and Hiromi, Y. (2007). TFIIH controls developmentally-regulated cell cycle progression as a holocomplex. *Genes Cells* 12, 1289-1300.

354. Matzke, M., Kanno, T., Daxinger, L., Huettel, B., and Matzke, A.J. (2009). RNA-mediated chromatin-based silencing in plants. *Curr. Opin. Cell Biol.* 21, 367-376.
355. Medlin, J., Scurry, A., Taylor, A., Zhang, F., Peterlin, B.M., and Murphy, S. (2005). P-TEFb is not an essential elongation factor for the intronless human U2 snRNA and histone H2b genes. *EMBO J.* 24, 4154-4165.
356. Meinhart, A., Kamenski, T., Hoepfner, S., Baumli, S., and Cramer, P. (2005). A structural perspective of CTD function. *Genes Dev.* 19, 1401-1415.
357. Meka, H., Werner, F., Cordell, S.C., Onesti, S., and Brick, P. (2005). Crystal structure and RNA binding of the Rpb4/Rpb7 subunits of human RNA polymerase II. *Nucleic Acids Res.* 33, 6435-6444.
358. Mencia, M., Moqtaderi, Z., Geisberg, J.V., Kuras, L., and Struhl, K. (2002). Activator-Specific Recruitment of TFIID and Regulation of Ribosomal Protein Genes in Yeast. *Mol. Cell* 9, 823-833.
359. Merika, M., and Thanos, D. (2001). Enhanceosomes. *Curr. Opin. Genet. Dev.* 11, 205-208.
360. Mermelstein, F., Yeung, K., Cao, J., Inostroza, J.A., Erdjument-Bromage, H., Egelson, K., Landsman, D., Levitt, P., Tempst, P., and Reinberg, D. (1996). Requirement of a corepressor for Dr1-mediated repression of transcription. *Genes Dev.* 10, 1033-1048.
361. Metzger, D., Scheer, E., Soldatov, A., and Tora, L. (1999). Mammalian TAF(II)30 is required for cell cycle progression and specific cellular differentiation programmes. *EMBO J.* 18, 4823-4834.
362. Michels, A.A., Fraldi, A., Li, Q., Adamson, T.E., Bonnet, F., Nguyen, V.T., Sedore, S.C., Price, J.P., Price, D.H., Lania, L., and Bensaude, O. (2004). Binding of the 7SK snRNA turns the HEXIM1 protein into a P-TEFb (CDK9/cyclin T) inhibitor. *EMBO J.* 23, 2608-2619.
363. Mitsiou, D.J., and Stunnenberg, H.G. (2000). TAC, a TBP-sans-TAFs complex containing the unprocessed TFIIAalpha precursor and the TFIIAgamma subunit. *Mol. Cell* 6, 527-537.
364. Mitsiou, D.J., and Stunnenberg, H.G. (2003). p300 is involved in formation of the TBP-TFIIA-containing basal transcription complex, TAC. *EMBO J.* 22, 4501-4511.
365. Mitsuzawa, H., Kanda, E., and Ishihama, A. (2003). Rpb7 subunit of RNA polymerase II interacts with an RNA-binding protein involved in processing of transcripts. *Nucleic Acids Res.* 31, 4696-4701.

366. Miyao, T., Barnett, J.D., and 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.
367. Mizuguchi, K., Deane, C.M., Blundell, T.L., and Overington, J.P. (1998). HOMSTRAD: a database of protein structure alignments for homologous families. *Protein Sci.* 7, 2469-2471.
368. Mizzen, C.A., Yang, X.J., Kokubo, T., Brownell, J.E., Bannister, A.J., Owen-Hughes, T., Workman, J., Wang, L., Berger, S.L., Kouzarides, T., Nakatani, Y., and Allis, C.D. (1996). The TAF(II)250 subunit of TFIID has histone acetyltransferase activity. *Cell* 87, 1261-1270.
369. Mnaimneh, S., Davierwala, A.P., Haynes, J., Moffat, J., Peng, W.T., Zhang, W., Yang, X., Pootoolal, J., Chua, G., Lopez, A., Trochesset, M., Morse, D., Krogan, N.J., Hiley, S.L., Li, Z., Morris, Q., Grigull, J., Mitsakakis, N., Roberts, C.J., Greenblatt, J.F., Boone, C., Kaiser, C.A., Andrews, B.J., and Hughes, T.R. (2004). Compendium II: Exploration of essential gene functions via titratable promoter alleles. - Submitted.
370. Mochizuki, K., Nishiyama, A., Jang, M.K., Dey, A., Ghosh, A., Tamura, T., Natsume, H., Yao, H., and Ozato, K. (2008). The bromodomain protein Brd4 stimulates G1 gene transcription and promotes progression to S phase. *J. Biol. Chem.* 283, 9040-9048.
371. Mohan, W.S., Jr., Scheer, E., Wendling, O., Metzger, D., and Tora, L. (2003). TAF10 (TAF(II)30) is necessary for TFIID stability and early embryogenesis in mice. *Mol. Cell Biol.* 23, 4307-4318.
372. Mosley, A.L., Pattenden, S.G., Carey, M., Venkatesh, S., Gilmore, J.M., Florens, L., Workman, J.L., and Washburn, M.P. (2009). Rtr1 is a CTD phosphatase that regulates RNA polymerase II during the transition from serine 5 to serine 2 phosphorylation. *Mol. Cell* 34, 168-178.
373. Muller, F., Demeny, M.A., and Tora, L. (2007). New Problems in RNA Polymerase II Transcription Initiation: Matching the Diversity of Core Promoters with a Variety of Promoter Recognition Factors. *J. Biol Chem.* 282, 14685-14689.
374. Muller, F., and Tora, L. (2004). The multicoloured world of promoter recognition complexes. *EMBO J.* 23, 2-8.
375. Murakami, K.S., Masuda, S., Campbell, E.A., Muzzin, O., and Darst, S.A. (2002a). Structural Basis of Transcription Initiation: An RNA Polymerase Holoenzyme-DNA Complex. *Science* 296, 1285-1290.

376. Murakami, K.S., Masuda, S., and Darst, S.A. (2002b). Structural basis of transcription initiation: RNA polymerase holoenzyme at 4 Å resolution. *Science* 296, 1280-1284.
377. Murayama, A., Kim, M.S., Yanagisawa, J., Takeyama, K., and Kato, S. (2004). Transrepression by a liganded nuclear receptor via a bHLH activator through co-regulator switching. *EMBO J.* 23, 1598-1608.
378. Muse, G.W., Gilchrist, D.A., Nechaev, S., Shah, R., Parker, J.S., Grissom, S.F., Zeitlinger, J., and Adelman, K. (2007). RNA polymerase is poised for activation across the genome. *Nat. Genet.* 39, 1507-1511.
379. Myers, L.C., and Kornberg, R.D. (2000). Mediator of transcriptional regulation. *Annu. Rev. Biochem.* 69, 729-749.
380. Na, X., Duan, H.O., Messing, E.M., Schoen, S.R., Ryan, C.K., di Sant'Agnes, P.A., Golemis, E.A., and Wu, G. (2003). Identification of the RNA polymerase II subunit hsRPB7 as a novel target of the von Hippel-Lindau protein. *EMBO J.* 22, 4249-4259.
381. Nabavi, S., Nellimarla, S., and Nazar, R.N. (2008). Post-transcriptional regulation of the U3 small nucleolar RNA. *J. Biol. Chem.* 283, 21404-21410.
382. Nagata, M., Ito, T., Arimitsu, N., Koyama, H., and Sekimizu, K. (2009). Transcription arrest relief by S-II/TFIIS during gene expression in erythroblast differentiation. *Genes Cells* 14, 371-380.
383. Nagy, P.L., Cleary, M.L., Brown, P.O., and Lieb, J.D. (2003). Genomewide demarcation of RNA polymerase II transcription units revealed by physical fractionation of chromatin. *Proc. Natl. Acad. Sci. U. S. A* 100, 6364-6369.
384. Nakamura, K., Jeong, S.Y., Uchihara, T., Anno, M., Nagashima, K., Nagashima, T., Ikeda, S., Tsuji, S., and Kanazawa, I. (2001). SCA17, a novel autosomal dominant cerebellar ataxia caused by an expanded polyglutamine in TATA-binding protein. *Hum. Mol. Genet.* 10, 1441-1448.
385. Nakanishi, T., Shimoaraiso, M., Kubo, T., and Natori, S. (1995). Structure-function relationship of yeast S-II in terms of stimulation of RNA polymerase II, arrest relief, and suppression of 6-azauracil sensitivity. *J. Biol. Chem.* 270, 8991-8995.
386. Nakata, A., Ito, T., Nagata, M., Hori, S., and Sekimizu, K. (2004). GRIP1tau, a novel PDZ domain-containing transcriptional activator, cooperates with the testis-specific transcription elongation factor SII-T1. *Genes Cells* 9, 1125-1135.

387. Neely, K.E., Hassan, A.H., Brown, C.E., Howe, L., and Workman, J.L. (2002). Transcription Activator Interactions with Multiple SWI/SNF Subunits. *Mol. Cell Biol.* 22, 1615-1625.
388. Nesser, N.K., Peterson, D.O., and Hawley, D.K. (2006). RNA polymerase II subunit Rpb9 is important for transcriptional fidelity in vivo. *Proc. Natl. Acad. Sci. U. S. A* 103, 3268-3273.
389. Neuman, K.C., Abbondanzieri, E.A., Landick, R., Gelles, J., and Block, S.M. (2003). Ubiquitous transcriptional pausing is independent of RNA polymerase backtracking. *Cell* 115, 437-447.
390. Newton, K., Matsumoto, M.L., Wertz, I.E., Kirkpatrick, D.S., Lill, J.R., Tan, J., Dugger, D., Gordon, N., Sidhu, S.S., Fellouse, F.A., Komuves, L., French, D.M., Ferrando, R.E., Lam, C., Compaan, D., Yu, C., Bosanac, I., Hymowitz, S.G., Kelley, R.F., and Dixit, V.M. (2008). Ubiquitin chain editing revealed by polyubiquitin linkage-specific antibodies. *Cell* 134, 668-678.
391. Ni, Z., Saunders, A., Fuda, N.J., Yao, J., Suarez, J.R., Webb, W.W., and Lis, J.T. (2008). P-TEFb is critical for the maturation of RNA polymerase II into productive elongation in vivo. *Mol Cell Biol* 28, 1161-1170.
392. Ni, Z., Schwartz, B.E., Werner, J., Suarez, J.R., and Lis, J.T. (2004). Coordination of transcription, RNA processing, and surveillance by P-TEFb kinase on heat shock genes. *Mol. Cell* 13, 55-65.
393. Nickels, B.E., and Hochschild, A. (2004). Regulation of RNA polymerase through the secondary channel. *Cell* 118, 281-284.
394. Nikolov, D.B., Chen, H., Halay, E.D., Usheva, A.A., Hisatake, K., Lee, D.K., Roeder, R.G., and Burley, S.K. (1995). Crystal structure of a TFIIB-TBP-TATA-element ternary complex. *Nature* 377, 119-128.
395. Nissen, R.M., and 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.
396. No, D., Yao, T.P., and Evans, R.M. (1996). Ecdysone-inducible gene expression in mammalian cells and transgenic mice. *Proc. Natl. Acad. Sci. U. S. A* 93, 3346-3351.
397. Nobrega, M.A., Ovcharenko, I., Afzal, V., and Rubin, E.M. (2003). Scanning human gene deserts for long-range enhancers. *Science* 302, 413.

398. Nowak, D.E., Tian, B., and Brasier, A.R. (2005). Two-step cross-linking method for identification of NF-kappaB gene network by chromatin immunoprecipitation. *Biotechniques* 39, 715-725.
399. Nowak, D.E., Tian, B., Jamaluddin, M., Boldogh, I., Vergara, L.A., Choudhary, S., and Brasier, A.R. (2008). RelA Ser276 phosphorylation is required for activation of a subset of NF-kappaB-dependent genes by recruiting cyclin-dependent kinase 9/cyclin T1 complexes. *Mol. Cell Biol.* 28, 3623-3638.
400. O'Brien, P.M., Davies, M.J., Scurry, J.P., Smith, A.N., Barton, C.A., Henderson, M.J., Saunders, D.N., Gloss, B.S., Patterson, K.I., Clancy, J.L., Heinzelmann-Schwarz, V.A., Murali, R., Scolyer, R.A., Zeng, Y., Williams, E.D., Scurr, L., Defazio, A., Quinn, D.I., Watts, C.K., Hacker, N.F., Henshall, S.M., and Sutherland, R.L. (2008). The E3 ubiquitin ligase EDD is an adverse prognostic factor for serous epithelial ovarian cancer and modulates cisplatin resistance in vitro. *Br. J. Cancer* 98, 1085-1093.
401. O'Gorman, W., Kwek, K.Y., Thomas, B., and Akoulitchev, A. (2006). Non-coding RNA in transcription initiation. *Biochem. Soc. Symp.*, 131-140.
402. Ohkuma, Y., Hashimoto, S., Wang, C.K., Horikoshi, M., and 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.
403. Ohkuma, Y., and Roeder, R.G. (1994). Regulation of TFIIH ATPase and kinase activities by TFIIE during active initiation complex formation. *Nature* 368, 160-163.
404. 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.
405. Okamoto, T., Yamamoto, S., Watanabe, Y., Ohta, T., Hanaoka, F., Roeder, R.G., and 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.
406. Okuda, M., Tanaka, A., Arai, Y., Satoh, M., Okamura, H., Nagadoi, A., Hanaoka, F., Ohkuma, Y., and Nishimura, Y. (2004). A novel zinc finger structure in the large subunit of human general transcription factor TFIIE. *J. Biol. Chem.* 279, 51395-51403.

407. Okuda, M., Tanaka, A., Satoh, M., Mizuta, S., Takazawa, M., Ohkuma, Y., and Nishimura, Y. (2008). Structural insight into the TFIIE-TFIIF interaction: TFIIE and p53 share the binding region on TFIIF. *EMBO J.* 27, 1161-1171.
408. Okuda, M., Watanabe, Y., Okamura, H., Hanaoka, F., Ohkuma, Y., and Nishimura, Y. (2000). Structure of the central core domain of TFIIEbeta with a novel double-stranded DNA-binding surface. *EMBO J.* 19, 1346-1356.
409. Olmsted, V.K., Awrey, D.E., Koth, C., Shan, X., Morin, P.E., Kazanis, S., Edwards, A.M., and Arrowsmith, C.H. (1998). Yeast transcript elongation factor (TFIIS), structure and function. I: NMR structural analysis of the minimal transcriptionally active region. *J. Biol. Chem.* 273, 22589-22594.
410. 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.
411. Orlicky, S.M., Tran, P.T., Sayre, M.H., and 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.
412. Orphanides, G., Lagrange, T., and Reinberg, D. (1996). The general transcription factors of RNA polymerase II. *Genes Dev.* 10, 2657-2683.
413. Ozer, J., Moore, P.A., Bolden, A.H., Lee, A., Rosen, C.A., and Lieberman, P.M. (1994). Molecular cloning of the small (gamma) subunit of human TFIIA reveals functions critical for activated transcription. *Genes Dev.* 8, 2324-2335.
414. Ozer, J., Moore, P.A., and Lieberman, P.M. (2000). A testis-specific transcription factor IIA (TFIIAtau) stimulates TATA-binding protein-DNA binding and transcription activation. *J. Biol. Chem.* 275, 122-128.
415. Pabo, C.O., and Sauer, R.T. (1992). Transcription factors: structural families and principles of DNA recognition. *Annu. Rev. Biochem.* 61, 1053-1095.
416. Pal, M., and Luse, D.S. (2003). The initiation-elongation transition: Lateral mobility of RNA in RNA polymerase II complexes is greatly reduced at +8/+9 and absent by +23. *Proc. Natl. Acad. Sci. U. S. A* 100, 5700-5705.

417. Pal, M., McKean, D., and Luse, D.S. (2001). Promoter clearance by RNA polymerase II is an extended, multistep process strongly affected by sequence. *Mol. Cell Biol.* 21, 5815-5825.
418. Palancade, B., and Bensaude, O. (2003). Investigating RNA polymerase II carboxyl-terminal domain (CTD) phosphorylation. *Eur. J. Biochem.* 270, 3859-3870.
419. Palangat, M., Renner, D.B., Price, D.H., and Landick, R. (2005). A negative elongation factor for human RNA polymerase II inhibits the anti-arrest transcript-cleavage factor TFIIS. *Proc. Natl. Acad. Sci. U. S. A.*
420. Pan, G., Aso, T., and Greenblatt, J. (1997). Interaction of elongation factors TFIIS and elongin A with a human RNA polymerase II holoenzyme capable of promoter-specific initiation and responsive to transcriptional activators. *J. Biol. Chem.* 272, 24563-24571.
421. Pan, G., and 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.
422. Papatsenko, D., and Levine, M. (2005). Quantitative analysis of binding motifs mediating diverse spatial readouts of the Dorsal gradient in the *Drosophila* embryo. *Proc. Natl. Acad. Sci. U. S. A.* 102, 4966-4971.
423. Parvin, J.D., Shykind, B.M., Meyers, R.E., Kim, J., and Sharp, P.A. (1994). Multiple sets of basal factors initiate transcription by RNA polymerase II. *J. Biol. Chem.* 269, 18414-18421.
424. Patikoglou, G.A., Kim, J.L., Sun, L., Yang, S.H., Kodadek, T., and Burley, S.K. (1999). TATA element recognition by the TATA box-binding protein has been conserved throughout evolution. *Genes Dev.* 13, 3217-3230.
425. Patturajan, M., Conrad, N.K., Bregman, D.B., and Corden, J.L. (1999). Yeast carboxyl-terminal domain kinase I positively and negatively regulates RNA polymerase II carboxyl-terminal domain phosphorylation. *J. Biol. Chem.* 274, 27823-27828.
426. Pei, Y., Hausmann, S., Ho, C.K., Schwer, B., and Shuman, S. (2001). The length, phosphorylation state, and primary structure of the RNA polymerase II carboxyl-terminal domain dictate interactions with mRNA capping enzymes. *J. Biol. Chem.* 276, 28075-28082.

427. Pelechano, V., Jimeno-Gonzalez, S., Rodriguez-Gil, A., Garcia-Martinez, J., Perez-Ortin, J.E., and Chavez, S. (2009). Regulon-specific control of transcription elongation across the yeast genome. *PLoS. Genet.* 5, e1000614.
428. Perales, R., and Bentley, D. (2009). "Cotranscriptionality": the transcription elongation complex as a nexus for nuclear transactions. *Mol. Cell* 36, 178-191.
429. Peterlin, B.M., and Price, D.H. (2006). Controlling the elongation phase of transcription with P-TEFb. *Mol. Cell* 23, 297-305.
430. 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.
431. Pham, A.D., and Sauer, F. (2000). Ubiquitin-activating/conjugating activity of TAFII250, a mediator of activation of gene expression in *Drosophila*. *Science* 289, 2357-2360.
432. Phatnani, H.P., and Greenleaf, A.L. (2006). Phosphorylation and functions of the RNA polymerase II CTD. *Genes Dev.* 20, 2922-2936.
433. Phillips, J.E., and Corces, V.G. (2009). CTCF: master weaver of the genome. *Cell* 137, 1194-1211.
434. Pillai, B., Sampath, V., Sharma, N., and Sadhale, P. (2001). Rpb4, a non-essential subunit of core RNA polymerase II of *Saccharomyces cerevisiae* is important for activated transcription of a subset of genes. *J. Biol. Chem.* 276, 30641-30647.
435. Pillai, B., Verma, J., Abraham, A., Francis, P., Kumar, Y., Tatu, U., Brahmachari, S.K., and Sadhale, P.P. (2003). Whole Genome Expression Profiles of Yeast RNA Polymerase II Core Subunit, Rpb4, in Stress and Nonstress Conditions. *J. Biol. Chem.* 278, 3339-3346.
436. Ping, Y.H., and Rana, T.M. (2001). DSIF and NELF interact with RNA polymerase II elongation complex and HIV-1 Tat stimulates P-TEFb-mediated phosphorylation of RNA polymerase II and DSIF during transcription elongation. *J. Biol. Chem.* 276, 12951-12958.
437. Pokholok, D.K., Hannett, N.M., and Young, R.A. (2002a). Exchange of RNA polymerase II initiation and elongation factors during gene expression in vivo. *Mol Cell* 9, 799-809.

438. Pokholok, D.K., Hannett, N.M., and Young, R.A. (2002b). Exchange of RNA Polymerase II Initiation and Elongation Factors during Gene Expression In Vivo. *Mol. Cell* 9, 799-809.
439. Prather, D.M., Larschan, E., and Winston, F. (2005). Evidence that the elongation factor TFIIS plays a role in transcription initiation at GAL1 in *Saccharomyces cerevisiae*. *Mol Cell Biol* 25, 2650-2659.
440. Ptashne, M., and Gann, A. (1997). Transcriptional activation by recruitment. *Nature* 386, 569-577.
441. Pugh, B.F. (1996). Mechanisms of transcription complex assembly. *Curr. Opin. Cell Biol.* 8, 303-311.
442. Pugh, B.F. (2000). Control of gene expression through regulation of the TATA-binding protein. *Gene* 255, 1-14.
443. Pugh, B.F., and Tjian, R. (1990). Mechanism of transcriptional activation by Sp1: evidence for coactivators. *Cell* 61, 1187-1197.
444. Puig, O., Caspary, F., Rigaut, G., Rutz, B., Bouveret, E., Bragado-Nilsson, E., Wilm, M., and Seraphin, B. (2001). The tandem affinity purification (TAP) method: a general procedure of protein complex purification. *Methods* 24, 218-229.
445. Qiu, H., Hu, C., Wong, C.M., and Hinnebusch, A.G. (2006). The Spt4p subunit of yeast DSIF stimulates association of the Paf1 complex with elongating RNA polymerase II. *Mol. Cell Biol.* 26, 3135-3148.
446. Quan, T.K., and Hartzog, G.A. (2010). Histone H3K4 and K36 Methylation, Chd1 and Rpd3S Oppose the Functions of *Saccharomyces cerevisiae* Spt4-Spt5 in Transcription. *Genetics* 184, 321-334.
447. Rachez, C., and Freedman, L.P. (2001). Mediator complexes and transcription. *Curr. Opin. Cell Biol.* 13, 274-280.
448. Radonjic, M., Andrau, J.C., Lijnzaad, P., Kemmeren, P., Kockelkorn, T.T., van Leenen, D., van Berkum, N.L., and Holstege, F.C. (2005). Genome-wide analyses reveal RNA polymerase II located upstream of genes poised for rapid response upon *S. cerevisiae* stationary phase exit. *Mol Cell* 18, 171-183.
449. Rahl, P.B., Lin, C.Y., Seila, A.C., Flynn, R.A., McCuine, S., Burge, C.B., Sharp, P.A., and Young, R.A. (2010). c-Myc regulates transcriptional pause release. *Cell* 141, 432-445.
450. Rani, P.G., Ranish, J.A., and Hahn, S. (2004). RNA polymerase II (Pol II)-TFIIF and Pol II-mediator complexes: the major stable Pol II

- complexes and their activity in transcription initiation and reinitiation. *Mol. Cell Biol.* 24, 1709-1720.
451. Ranish, J.A., Hahn, S., Lu, Y., Yi, E.C., Li, X.J., Eng, J., and Aebersold, R. (2004). Identification of TFB5, a new component of general transcription and DNA repair factor IIIH. *Nat. Genet.* 36, 707-713.
 452. Ranish, J.A., Yudkovsky, N., and Hahn, S. (1999). Intermediates in formation and activity of the RNA polymerase II preinitiation complex: holoenzyme recruitment and a postrecruitment role for the TATA box and TFIIB. *Genes Dev.* 13, 49-63.
 453. Reese, J.C. (2003). Basal transcription factors. *Curr. Opin. Genet. Dev.* 13, 114-118.
 454. Reeves, W.M., and Hahn, S. (2003). Activator-independent functions of the yeast mediator sin4 complex in preinitiation complex formation and transcription reinitiation. *Mol. Cell Biol.* 23, 349-358.
 455. Reeves, W.M., and Hahn, S. (2005). Targets of the Gal4 transcription activator in functional transcription complexes. *Mol Cell Biol* 25, 9092-9102.
 456. 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.
 457. Remenyi, A., Scholer, H.R., and Wilmanns, M. (2004). Combinatorial control of gene expression. *Nat. Struct. Mol Biol* 11, 812-815.
 458. Renner, D.B., Yamaguchi, Y., Wada, T., Handa, H., and Price, D.H. (2001). A Highly Purified RNA Polymerase II Elongation Control System. *J. Biol. Chem.* 276, 42601-42609.
 459. Revyakin, A., Liu, C., Ebright, R.H., and Strick, T.R. (2006). Abortive initiation and productive initiation by RNA polymerase involve DNA scrunching. *Science* 314, 1139-1143.
 460. Reyes-Turcu, F.E., Ventii, K.H., and Wilkinson, K.D. (2009). Regulation and cellular roles of ubiquitin-specific deubiquitinating enzymes. *Annu. Rev. Biochem.* 78, 363-397.
 461. Rice, A.P., and Herrmann, C.H. (2003). Regulation of TAK/P-TEFb in CD4+ T lymphocytes and macrophages. *Curr. HIV. Res.* 1, 395-404.
 462. 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.
463. Ring, B.Z., Yarnell, W.S., and Roberts, J.W. (1996). Function of *E. coli* RNA polymerase sigma factor sigma 70 in promoter-proximal pausing. *Cell* 86, 485-493.
464. Robert, F., Douziech, M., Forget, D., Egly, J.M., Greenblatt, J., Burton, Z.F., and Coulombe, B. (1998). Wrapping of promoter DNA around the RNA polymerase II initiation complex induced by TFIIF. *Mol. Cell* 2, 341-351.
465. Robert, F., Pokholok, D.K., Hannett, N.M., Rinaldi, N.J., Chandy, M., Rolfe, A., Workman, J.L., Gifford, D.K., and Young, R.A. (2004). Global Position and Recruitment of HATs and HDACs in the Yeast Genome. *Mol Cell* 16, 199-209.
466. Rochette-Egly, C., Adam, S., Rossignol, M., Egly, J.M., and Chambon, P. (1997). Stimulation of RAR alpha activation function AF-1 through binding to the general transcription factor TFIIF and phosphorylation by CDK7. *Cell* 90, 97-107.
467. Rodriguez, C.R., Cho, E.J., Keogh, M.C., Moore, C.L., Greenleaf, A.L., and Buratowski, S. (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.
468. Roeder, R.G. (1996). The role of general initiation factors in transcription by RNA polymerase II. *Trends Biochem. Sci.* 21, 327-335.
469. Rosenheck, S., and Choder, M. (1998). Rpb4, a subunit of RNA polymerase II, enables the enzyme to transcribe at temperature extremes in vitro. *J. Bacteriol.* 180, 6187-6192.
470. Rossignol, M., Keriell, A., Staub, A., and Egly, J.M. (1999). Kinase activity and phosphorylation of the largest subunit of TFIIF transcription factor. *J. Biol. Chem.* 274, 22387-22392.
471. Rother, S., and Strasser, K. (2007). The RNA polymerase II CTD kinase Ctk1 functions in translation elongation. *Genes Dev.* 21, 1409-1421.
472. Rotin, D., and Kumar, S. (2009). Physiological functions of the HECT family of ubiquitin ligases. *Nat. Rev. Mol. Cell Biol.* 10, 398-409.
473. Rougvie, A.E., and Lis, J.T. (1988). The RNA polymerase II molecule at the 5' end of the uninduced hsp70 gene of *D. melanogaster* is transcriptionally engaged. *Cell* 54, 795-804.

474. Roy, R., Schaeffer, L., Humbert, S., Vermeulen, W., Weeda, G., and Egly, J.M. (1994). The DNA-dependent ATPase activity associated with the class II basic transcription factor BTF2/TFIIH. *J. Biol. Chem.* 269, 9826-9832.
475. Runner, V.M., Podolny, V., and Buratowski, S. (2008). The Rpb4 subunit of RNA polymerase II contributes to cotranscriptional recruitment of 3' processing factors. *Mol. Cell Biol.* 28, 1883-1891.
476. Ruppert, S., and Tjian, R. (1995). Human TAFII250 interacts with RAP74: implications for RNA polymerase II initiation. *Genes Dev.* 9, 2747-2755.
477. Sabo, A., Lusic, M., Cereseto, A., and Giacca, M. (2008). Acetylation of conserved lysines in the catalytic core of cyclin-dependent kinase 9 inhibits kinase activity and regulates transcription. *Mol. Cell Biol.* 28, 2201-2212.
478. Sampath, V., Rekha, N., Srinivasan, N., and Sadhale, P. (2003). The conserved and non-conserved regions of Rpb4 are involved in multiple phenotypes in *Saccharomyces cerevisiae*. *J. Biol. Chem.* 278, 51566-51576.
479. Sandelin, A., Carninci, P., Lenhard, B., Ponjavic, J., Hayashizaki, Y., and Hume, D.A. (2007). Mammalian RNA polymerase II core promoters: insights from genome-wide studies. *Nat. Rev. Genet.* 8, 424-436.
480. Sarker, A.H., Tsutakawa, S.E., Kostek, S., Ng, C., Shin, D.S., Peris, M., Campeau, E., Tainer, J.A., Nogales, E., and Cooper, P.K. (2005). Recognition of RNA Polymerase II and Transcription Bubbles by XPG, CSB, and TFIIH: Insights for Transcription-Coupled Repair and Cockayne Syndrome. *Mol Cell* 20, 187-198.
481. Sato, S., Tomomori-Sato, C., Banks, C.A., Sorokina, I., Parmely, T.J., Kong, S.E., Jin, J., Cai, Y., Lane, W.S., Brower, C.S., Conaway, R.C., and Conaway, J.W. (2003). Identification of Mammalian Mediator Subunits with Similarities to Yeast Mediator Subunits Srb5, Srb6, Med11, and Rox3. *J. Biol. Chem.* 278, 15123-15127.
482. Saunders, A., Core, L.J., and Lis, J.T. (2006). Breaking barriers to transcription elongation. *Nat. Rev. Mol Cell Biol* 7, 557-567.
483. Saunders, D.N., Hird, S.L., Withington, S.L., Dunwoodie, S.L., Henderson, M.J., Biben, C., Sutherland, R.L., Ormandy, C.J., and Watts, C.K. (2004). Edd, the murine hyperplastic disc gene, is essential for yolk sac vascularization and chorioallantoic fusion. *Mol. Cell Biol.* 24, 7225-7234.

484. Savinkova, L.K., Ponomarenko, M.P., Ponomarenko, P.M., Drachkova, I.A., Lysova, M.V., Arshinova, T.V., and Kolchanov, N.A. (2009). TATA box polymorphisms in human gene promoters and associated hereditary pathologies. *Biochemistry (Mosc.)* 74, 117-129.
485. Schaeffer, L., Moncollin, V., Roy, R., Staub, A., Mezzina, M., Sarasin, A., Weeda, G., Hoeijmakers, J.H., and Egly, J.M. (1994). The ERCC2/DNA repair protein is associated with the class II BTF2/TFIIH transcription factor. *EMBO J.* 13, 2388-2392.
486. Schaeffer, L., Roy, R., Humbert, S., Moncollin, V., Vermeulen, W., Hoeijmakers, J.H., Chambon, P., and Egly, J.M. (1993). DNA repair helicase: a component of BTF2 (TFIIH) basic transcription factor. *Science* 260, 58-63.
487. Schaffer, A.A., Wolf, Y.I., Ponting, C.P., Koonin, E.V., Aravind, L., and Altschul, S.F. (1999). IMPALA: matching a protein sequence against a collection of PSI-BLAST-constructed position-specific score matrices. *Bioinformatics.* 15, 1000-1011.
488. Schwartz, S., Meshorer, E., and Ast, G. (2009). Chromatin organization marks exon-intron structure. *Nat. Struct. Mol. Biol.* 16, 990-995.
489. Schwarz, S.E., Rosa, J.L., and Scheffner, M. (1998). Characterization of human hect domain family members and their interaction with UbcH5 and UbcH7. *J. Biol. Chem.* 273, 12148-12154.
490. Sekimizu, K., Kobayashi, N., Mizuno, D., and Natori, S. (1976). Purification of a factor from Ehrlich ascites tumor cells specifically stimulating RNA polymerase II. *Biochemistry* 15, 5064-5070.
491. Selitrennik, M., Duek, L., Lotan, R., and Choder, M. (2006). Nucleocytoplasmic shuttling of the Rpb4p and Rpb7p subunits of *Saccharomyces cerevisiae* RNA polymerase II by two pathways. *Eukaryot. Cell* 5, 2092-2103.
492. Serizawa, H., Conaway, R.C., and Conaway, J.W. (1992). A carboxyl-terminal-domain kinase associated with RNA polymerase II transcription factor delta from rat liver. *Proc. Natl. Acad. Sci. U. S. A* 89, 7476-7480.
493. Serizawa, H., Conaway, R.C., and Conaway, J.W. (1993). Multifunctional RNA polymerase II initiation factor delta from rat liver. Relationship between carboxyl-terminal domain kinase, ATPase, and DNA helicase activities. *J. Biol. Chem.* 268, 17300-17308.

494. Serizawa, H., Makela, T.P., Conaway, J.W., Conaway, R.C., Weinberg, R.A., and Young, R.A. (1995). Association of Cdk-activating kinase subunits with transcription factor TFIID. *Nature* 374, 280-282.
495. Shao, H., Revach, M., Moshonov, S., Tzuman, Y., Gazit, K., Albeck, S., Unger, T., and Dikstein, R. (2005). Core promoter binding by histone-like TAF complexes. *Mol Cell Biol* 25, 206-219.
496. Shaw, R.J., Bonawitz, N.D., and Reines, D. (2002). Use of an in vivo reporter assay to test for transcriptional and translational fidelity in yeast. *J. Biol. Chem.* 277, 24420-24426.
497. Shearwin, K.E., Callen, B.P., and Egan, J.B. (2005). Transcriptional interference--a crash course. *Trends Genet.* 21, 339-345.
498. Shiekhatar, R., Mermelstein, F., Fisher, R.P., Drapkin, R., Dynlacht, B., Wessling, H.C., Morgan, D.O., and Reinberg, D. (1995). Cdk-activating kinase complex is a component of human transcription factor TFIID. *Nature* 374, 283-287.
499. Shilatifard, A., Conaway, R.C., and Conaway, J.W. (2003). The RNA polymerase II elongation complex. *Annu. Rev. Biochem.* 72, 693-715.
500. Shim, E.Y., Walker, A.K., Shi, Y., and 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.
501. Shimoaraiso, M., Nakanishi, T., Kubo, T., and Natori, S. (2000). Transcription elongation factor S-II confers yeast resistance to 6-azauracil by enhancing expression of the SSM1 gene. *J. Biol. Chem.* 275, 29623-29627.
502. Siaut, M., Zaros, C., Levivier, E., Ferri, M.L., Court, M., Werner, M., Callebaut, I., Thuriaux, P., Sentenac, A., and Conesa, C. (2003). An Rpb4/Rpb7-like complex in yeast RNA polymerase III contains the orthologue of mammalian CGRP-RCP. *Mol. Cell Biol.* 23, 195-205.
503. Sikorski, T.W., and Buratowski, S. (2009). The basal initiation machinery: beyond the general transcription factors. *Curr. Opin. Cell Biol.* 21, 344-351.
504. Simon, J.A., and Kingston, R.E. (2009). Mechanisms of polycomb gene silencing: knowns and unknowns. *Nat. Rev. Mol. Cell Biol.* 10, 697-708.

505. Sims, R.J., III, Belotserkovskaya, R., and Reinberg, D. (2004a). Elongation by RNA polymerase II: the short and long of it. *Genes Dev.* 18, 2437-2468.
506. Sims, R.J., III, Mandal, S.S., and Reinberg, D. (2004b). Recent highlights of RNA-polymerase-II-mediated transcription. *Curr. Opin. Cell Biol.* 16, 263-271.
507. Sims, R.J., III, and Reinberg, D. (2008). Is there a code embedded in proteins that is based on post-translational modifications? *Nat. Rev. Mol. Cell Biol.* 9, 815-820.
508. Singh, B.N., and Hampsey, M. (2007). A transcription-independent role for TFIIB in gene looping. *Mol. Cell* 27, 806-816.
509. Siomin, Y.A., Simonov, V.V., and Poverenny, A.M. (1973). The reaction of formaldehyde with deoxynucleotides and DNA in the presence of amino acids and lysine-rich histone. *Biochim. Biophys. Acta* 331, 27-32.
510. Skaar, D.A., and Greenleaf, A.L. (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.
511. Sluder, A.E., Price, D.H., and Greenleaf, A.L. (1988). Elongation by *Drosophila* RNA polymerase II. Transcription of 3'-extended DNA templates. *J. Biol. Chem.* 263, 9917-9925.
512. Smale, S.T. (1994). Core promoter architecture for eukaryotic protein-coding genes. In *Transcription: mechanisms and regulation*, Conaway, R.C., and Conaway, J.W., eds. (New York, NY: Raven Press), pp. 63-81.
513. Solow, S., Salunek, M., Ryan, R., and Lieberman, P.M. (2001). Taf(II) 250 phosphorylates human transcription factor IIA on serine residues important for TBP binding and transcription activity. *J. Biol. Chem.* 276, 15886-15892.
514. Sopta, M., Burton, Z.F., and 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.
515. Sopta, M., Carthew, R.W., and Greenblatt, J. (1985). Isolation of three proteins that bind to mammalian RNA polymerase II. *J. Biol. Chem.* 260, 10353-10360.
516. 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.
517. Stargell, L.A., Moqtaderi, Z., Dorris, D.R., Ogg, R.C., and Struhl, K. (2000). TFIIA has activator-dependent and core promoter functions in vivo. *J. Biol. Chem.* 275, 12374-12380.
 518. Steffan, J.S., Keys, D.A., Dodd, J.A., and Nomura, M. (1996). The role of TBP in rDNA transcription by RNA polymerase I in *Saccharomyces cerevisiae*: TBP is required for upstream activation factor-dependent recruitment of core factor. *Genes Dev.* 10, 2551-2563.
 519. Steinmetz, E.J., and Brow, D.A. (1998). Control of pre-mRNA accumulation by the essential yeast protein Nrd1 requires high-affinity transcript binding and a domain implicated in RNA polymerase II association. *Proc. Natl. Acad. Sci. U. S. A* 95, 6699-6704.
 520. Steitz, T.A. (1998). A mechanism for all polymerases. *Nature* 391, 231-232.
 521. Struhl, K., and Moqtaderi, Z. (1998). The TAFs in the HAT. *Cell* 94, 1-4.
 522. 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.
 523. Sun, H., Palaniswamy, S.K., Pohar, T.T., Jin, V.X., Huang, T.H., and Davuluri, R.V. (2006). MPromDb: an integrated resource for annotation and visualization of mammalian gene promoters and ChIP-chip experimental data. *Nucleic Acids Res.* 34, D98-103.
 524. Svejstrup, J.Q. (2007). Contending with transcriptional arrest during RNAPII transcript elongation. *Trends Biochem. Sci.* 32, 165-171.
 525. Svejstrup, J.Q., Feaver, W.J., and Kornberg, R.D. (1996). Subunits of yeast RNA polymerase II transcription factor TFIIF encoded by the CCL1 gene. *J. Biol. Chem.* 271, 643-645.
 526. Svitkin, Y.V., Imataka, H., Khaleghpour, K., Kahvejian, A., Liebig, H.D., and Sonenberg, N. (2001). Poly(A)-binding protein interaction with eIF4G stimulates picornavirus IRES-dependent translation. *RNA.* 7, 1743-1752.
 527. Taatjes, D.J., Naar, A.M., Andel, I.F., Nogales, E., and Tjian, R. (2002). Structure, Function, and Activator-Induced Conformations of the CRSP Coactivator. *Science* 295, 1058-1062.

528. Takada, S., Lis, J.T., Zhou, S., and Tjian, R. (2000). A TRF1:BRF complex directs *Drosophila* RNA polymerase III transcription. *Cell* 101, 459-469.
529. Takagi, Y., Masuda, C.A., Chang, W.H., Komori, H., Wang, D., Hunter, T., Joazeiro, C.A., and Kornberg, R.D. (2005). Ubiquitin ligase activity of TFIIF and the transcriptional response to DNA damage. *Mol Cell* 18, 237-243.
530. 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.
531. Takeda, J., Suzuki, Y., Nakao, M., Kuroda, T., Sugano, S., Gojobori, T., and Imanishi, T. (2007). H-DBAS: alternative splicing database of completely sequenced and manually annotated full-length cDNAs based on H-Invitational. *Nucleic Acids Res.* 35, D104-D109.
532. Tan, S., Garrett, K.P., Conaway, R.C., and Conaway, J.W. (1994). Cryptic DNA-binding domain in the C terminus of RNA polymerase II general transcription factor RAP30. *Proc. Natl. Acad. Sci. U. S. A* 91, 9808-9812.
533. Tan, S., Hunziker, Y., Sargent, D.F., and Richmond, T.J. (1996). Crystal structure of a yeast TFIIA/TBP/DNA complex. *Nature* 381, 127-151.
534. Tasaki, T., Mulder, L.C., Iwamatsu, A., Lee, M.J., Davydov, I.V., Varshavsky, A., Muesing, M., and Kwon, Y.T. (2005). A family of mammalian E3 ubiquitin ligases that contain the UBR box motif and recognize N-degrons. *Mol. Cell Biol.* 25, 7120-7136.
535. Tatarakis, A., Margaritis, T., Martinez-Jimenez, C.P., Kouskouti, A., Mohan, W.S., Haroniti, A., Kafetzopoulos, D., Tora, L., and Talianidis, I. (2008). Dominant and redundant functions of TFIID involved in the regulation of hepatic genes. *Mol. Cell* 31, 531-543.
536. Teichmann, M., Wang, Z., Martinez, E., Tjernberg, A., Zhang, D., Vollmer, F., Chait, B.T., and Roeder, R.G. (1999). Human TATA-binding protein-related factor-2 (hTRF2) stably associates with hTFIIA in HeLa cells. *Proc. Natl. Acad. Sci. U. S. A* 96, 13720-13725.
537. Thomas, M.C., and Chiang, C.M. (2006). The general transcription machinery and general cofactors. *Crit Rev. Biochem. Mol. Biol.* 41, 105-178.
538. 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.

539. Timmers, H.T., and Tora, L. (2005). SAGA unveiled. *Trends Biochem. Sci.* 30, 7-10.
540. Tirode, F., Busso, D., Coin, F., and Egly, J.M. (1999). Reconstitution of the transcription factor TFIIH: assignment of functions for the three enzymatic subunits, XPB, XPD, and cdk7. *Mol. Cell* 3, 87-95.
541. 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.
542. Tran, K., and Gralla, J.D. (2008). Control of the timing of promoter escape and RNA catalysis by the transcription factor IIb fingertip. *J. Biol. Chem.* 283, 15665-15671.
543. Trinklein, N.D., Aldred, S.J., Saldanha, A.J., and Myers, R.M. (2003). Identification and functional analysis of human transcriptional promoters. *Genome Res.* 13, 308-312.
544. Tsuritani, K., Irie, T., Yamashita, R., Sakakibara, Y., Wakaguri, H., Kanai, A., Mizushima-Sugano, J., Sugano, S., Nakai, K., and Suzuki, Y. (2007). Distinct class of putative "non-conserved" promoters in humans: comparative studies of alternative promoters of human and mouse genes. *Genome Res.* 17, 1005-1014.
545. Tupler, R., Perini, G., and Green, M.R. (2001). Expressing the human genome. *Nature* 409, 832-833.
546. Turner, B.M. (2007). Defining an epigenetic code. *Nat. Cell Biol.* 9, 2-6.
547. Tuske, S., Sarafianos, S.G., Wang, X., Hudson, B., Sineva, E., Mukhopadhyay, J., Birktoft, J.J., Leroy, O., Ismail, S., Clark, A.D., Jr., Dharia, C., Napoli, A., Laptenko, O., Lee, J., Borukhov, S., Ebright, R.H., and Arnold, E. (2005). Inhibition of bacterial RNA polymerase by streptolydigin: stabilization of a straight-bridge-helix active-center conformation. *Cell* 122, 541-552.
548. Ueno, A., Baek, K., Jeon, C., and Agarwal, K. (1992). Netropsin specifically enhances RNA polymerase II termination at terminator sites in vitro. *Proc. Natl. Acad. Sci. U. S. A* 89, 3676-3680.
549. Ujvari, A., and Luse, D.S. (2006a). RNA emerging from the active site of RNA polymerase II interacts with the Rpb7 subunit. *Nat. Struct. Mol Biol* 13, 49-54.
550. Ujvari, A., and Luse, D.S. (2006b). RNA emerging from the active site of RNA polymerase II interacts with the Rpb7 subunit. *Nat. Struct. Mol Biol* 13, 49-54.

551. Ujvari, A., Pal, M., and Luse, D.S. (2002). RNA Polymerase II Transcription Complexes May Become Arrested If the Nascent RNA Is Shortened to Less than 50 Nucleotides. *J. Biol. Chem.* 277, 32527-32537.
552. Umehara, T., Kida, S., Yamamoto, T., and Horikoshi, M. (1995). Isolation and characterization of a cDNA encoding a new type of human transcription elongation factor S-II. *Gene* 167, 297-302.
553. Uptain, S.M., Kane, C.M., and Chamberlin, M.J. (1997). Basic mechanisms of transcript elongation and its regulation. *Annu. Rev. Biochem.* 66, 117-172.
554. Vaquero, A., Loyola, A., and Reinberg, D. (2003). The constantly changing face of chromatin. *Sci. Aging Knowledge Environ.* 2003, RE4.
555. Vasiljeva, L., Kim, M., Mutschler, H., Buratowski, S., and Meinhart, A. (2008). The Nrd1-Nab3-Sen1 termination complex interacts with the Ser5-phosphorylated RNA polymerase II C-terminal domain. *Nat. Struct. Mol. Biol.* 15, 795-804.
556. Vassylyev, D.G., Sekine, S., Laptenko, O., Lee, J., Vassylyeva, M.N., Borukhov, S., and Yokoyama, S. (2002). Crystal structure of a bacterial RNA polymerase holoenzyme at 2.6 Å resolution. *Nature* 417, 712-719.
557. Vassylyev, D.G., Vassylyeva, M.N., Zhang, J., Palangat, M., Artsimovitch, I., and Landick, R. (2007). Structural basis for substrate loading in bacterial RNA polymerase. *Nature* 448, 163-168.
558. Venters, B.J., and Pugh, B.F. (2009). A canonical promoter organization of the transcription machinery and its regulators in the *Saccharomyces* genome. *Genome Res.* 19, 360-371.
559. Verma-Gaur, J., Rao, S.N., Taya, T., and Sadhale, P. (2008). Genomewide recruitment analysis of Rpb4, a subunit of polymerase II in *Saccharomyces cerevisiae*, reveals its involvement in transcription elongation. *Eukaryot. Cell* 7, 1009-1018.
560. Viladevall, L., St Amour, C.V., Rosebrock, A., Schneider, S., Zhang, C., Allen, J.J., Shokat, K.M., Schwer, B., Leatherwood, J.K., and Fisher, R.P. (2009). TFIIF and P-TEFb coordinate transcription with capping enzyme recruitment at specific genes in fission yeast. *Mol. Cell* 33, 738-751.
561. Vincent, A.C., and Struhl, K. (1992). ACR1, a yeast ATF/CREB repressor. *Mol. Cell Biol.* 12, 5394-5405.

562. von Hippel, P.H., and Yager, T.D. (1992). The elongation-termination decision in transcription. *Science* 255, 809-812.
563. Wade, J.T., and Struhl, K. (2008). The transition from transcriptional initiation to elongation. *Curr. Opin. Genet. Dev.* 18, 130-136.
564. Wang, D., Bushnell, D.A., Westover, K.D., Kaplan, C.D., and Kornberg, R.D. (2006). Structural basis of transcription: role of the trigger loop in substrate specificity and catalysis. *Cell* 127, 941-954.
565. 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.
566. Wang, S., and Fischer, P.M. (2008). Cyclin-dependent kinase 9: a key transcriptional regulator and potential drug target in oncology, virology and cardiology. *Trends Pharmacol. Sci.* 29, 302-313.
567. Wang, Z., Buratowski, S., Svejstrup, J.Q., Feaver, W.J., Wu, X., Kornberg, R.D., Donahue, T.F., and Friedberg, E.C. (1995). The yeast TFB1 and SSL1 genes, which encode subunits of transcription factor IIIH, are required for nucleotide excision repair and RNA polymerase II transcription. *Mol. Cell Biol.* 15, 2288-2293.
568. Warfield, L., Ranish, J.A., and Hahn, S. (2004). Positive and negative functions of the SAGA complex mediated through interaction of Spt8 with TBP and the N-terminal domain of TFIIA. *Genes Dev.* 18, 1022-1034.
569. Watanabe, T., Hayashi, K., Tanaka, A., Furumoto, T., Hanaoka, F., and 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.
570. Weaver, Z.A., and Kane, C.M. (1997). Genomic characterization of a testis-specific TFIIS (TCEA2) gene. *Genomics* 46, 516-519.
571. Wegel, E., and Shaw, P. (2005). Gene activation and deactivation related changes in the three-dimensional structure of chromatin. *Chromosoma* 114, 331-337.
572. Wei, P., Garber, M.E., Fang, S.M., Fischer, W.H., and Jones, K.A. (1998). A novel CDK9-associated C-type cyclin interacts directly with HIV-1 Tat and mediates its high-affinity, loop-specific binding to TAR RNA. *Cell* 92, 451-462.

573. 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.
574. Wery, M., Shematorova, E., Van Driessche, B., Vandenhoute, J., Thuriaux, P., and Van, M., V. (2004). Members of the SAGA and Mediator complexes are partners of the transcription elongation factor TFIIS. *EMBO J.* 23, 4232-4242.
575. 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.
576. 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.
577. Winkles, J.A., and Alberts, G.F. (2005). Differential regulation of polo-like kinase 1, 2, 3, and 4 gene expression in mammalian cells and tissues. *Oncogene* 24, 260-266.
578. Wong, J.M., and Ingles, C.J. (2001). A compromised yeast RNA polymerase II enhances UV sensitivity in the absence of global genome nucleotide excision repair. *Mol Gen. Genet.* 264, 842-851.
579. Woychik, N.A., and Hampsey, M. (2002). The RNA Polymerase II Machinery. *Structure Illuminates Function.* *Cell* 108, 453-463.
580. Woychik, N.A., and 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.
581. Wu, L.F., Hughes, T.R., Davierwala, A.P., Robinson, M.D., Stoughton, R., and Altschuler, S.J. (2002). Large-scale prediction of *Saccharomyces cerevisiae* gene function using overlapping transcriptional clusters. *Nat. Genet.* 31, 255-265.
582. Wu, S.Y., and Chiang, C.M. (1998). Properties of PC4 and an RNA polymerase II complex in directing activated and basal transcription in vitro. *J. Biol. Chem.* 273, 12492-12498.
583. Wu, S.Y., Zhou, T., and Chiang, C.M. (2003). Human mediator enhances activator-facilitated recruitment of RNA polymerase II and promoter recognition by TATA-binding protein (TBP) independently of TBP-associated factors. *Mol. Cell Biol.* 23, 6229-6242.
584. Wu, X., and Lieber, M.R. (1997). Interaction between DNA-dependent protein kinase and a novel protein, KIP. *Mutat. Res.* 385, 13-20.

585. Xie, J., Collart, M., Lemaire, M., Stelzer, G., and Meisterernst, M. (2000). A single point mutation in TFIIA suppresses NC2 requirement in vivo. *EMBO J.* 19, 672-682.
586. Xiong, Y., and Burton, Z.F. (2007). A tunable ratchet driving human RNA polymerase II translocation adjusted by accurately templated nucleoside triphosphates loaded at downstream sites and by elongation factors. *J. Biol. Chem.* 282, 36582-36592.
587. Yamada, T., Yamaguchi, Y., Inukai, N., Okamoto, S., Mura, T., and Handa, H. (2006). P-TEFb-mediated phosphorylation of hSpt5 C-terminal repeats is critical for processive transcription elongation. *Mol Cell* 21, 227-237.
588. Yamamoto, S., Watanabe, Y., van der Spek, P.J., Watanabe, T., Fujimoto, H., Hanaoka, F., and 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.
589. Yamane, K., Wu, X., and Chen, J. (2002). A DNA damage-regulated BRCT-containing protein, TopBP1, is required for cell survival. *Mol. Cell Biol.* 22, 555-566.
590. Yamashita, R., Suzuki, Y., Wakaguri, H., Tsuritani, K., Nakai, K., and Sugano, S. (2006). DBTSS: DataBase of Human Transcription Start Sites, progress report 2006. *Nucleic Acids Res.* 34, D86-D89.
591. Yan, Q., Moreland, R.J., Conaway, J.W., and Conaway, R.C. (1999). Dual roles for transcription factor IIF in promoter escape by RNA polymerase II. *J. Biol. Chem.* 274, 35668-35675.
592. Yang, Y.H., and Speed, T. (2002). Design issues for cDNA microarray experiments. *Nat. Rev. Genet.* 3, 579-588.
593. Yang, Z., Yik, J.H., Chen, R., He, N., Jang, M.K., Ozato, K., and Zhou, Q. (2005). Recruitment of P-TEFb for stimulation of transcriptional elongation by the bromodomain protein Brd4. *Mol Cell* 19, 535-545.
594. Yang, Z., Zhu, Q., Luo, K., and Zhou, Q. (2001). The 7SK small nuclear RNA inhibits the CDK9/cyclin T1 kinase to control transcription. *Nature* 414, 317-322.
595. Yeo, M., Lin, P.S., Dahmus, M.E., and Gill, G.N. (2003). A novel RNA polymerase II C-terminal domain phosphatase that preferentially dephosphorylates serine 5. *J. Biol Chem.* 278, 26078-26085.

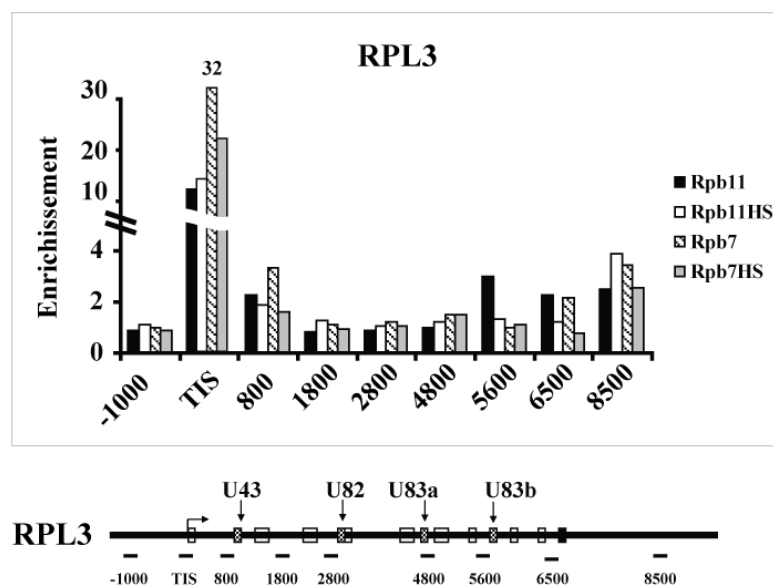
596. Yokomori, K., Verrijzer, C.P., and Tjian, R. (1998). An interplay between TATA box-binding protein and transcription factors IIE and IIA modulates DNA binding and transcription. *Proc. Natl. Acad. Sci. U. S. A* 95, 6722-6727.
597. Yong, C., Mitsuyasu, H., Chun, Z., Oshiro, S., Hamasaki, N., and Kitajima, S. (1998). Structure of the human transcription factor TFIIF revealed by limited proteolysis with trypsin. *FEBS Lett.* 435, 191-194.
598. Yoshida, M., Yoshida, K., Kozlov, G., Lim, N.S., De, C.G., Pang, Z., Berlanga, J.J., Kahvejian, A., Gehring, K., Wing, S.S., and Sonenberg, N. (2006). Poly(A) binding protein (PABP) homeostasis is mediated by the stability of its inhibitor, Paip2. *EMBO J.* 25, 1934-1944.
599. Young, R.A. (1991). RNA polymerase II. *Annu. Rev. Biochem.* 60, 689-715.
600. Yudkovsky, N., Ranish, J.A., and Hahn, S. (2000). A transcription reinitiation intermediate that is stabilized by activator. *Nature* 408, 225-229.
601. Zawel, L., Kumar, K.P., and Reinberg, D. (1995). Recycling of the general transcription factors during RNA polymerase II transcription. *Genes Dev.* 9, 1479-1490.
602. Zeghouf, M., Li, J., Butland, G., Borkowska, A., Canadien, V., Richards, D., Beattie, B., Emili, A., and Greenblatt, J.F. (2004). Sequential Peptide Affinity (SPA) system for the identification of mammalian and bacterial protein complexes. *J. Proteome. Res.* 3, 463-468.
603. Zeitlinger, J., Stark, A., Kellis, M., Hong, J.W., Nechaev, S., Adelman, K., Levine, M., and Young, R.A. (2007). RNA polymerase stalling at developmental control genes in the *Drosophila melanogaster* embryo. *Nat. Genet.* 39, 1512-1516.
604. 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.
605. Zhang, C., Zobeck, K.L., and Burton, Z.F. (2005a). Human RNA polymerase II elongation in slow motion: role of the TFIIF RAP74 alpha1 helix in nucleoside triphosphate-driven translocation. *Mol Cell Biol* 25, 3583-3595.
606. Zhang, G., Campbell, E.A., Minakhin, L., Richter, C., Severinov, K., and Darst, S.A. (1999). Crystal structure of *Thermus aquaticus* core RNA polymerase at 3.3 Å resolution. *Cell* 98, 811-824.

607. Zhang, T., and Kraus, W.L. (2009). SIRT1-dependent regulation of chromatin and transcription: Linking NAD(+) metabolism and signaling to the control of cellular functions. *Biochim. Biophys. Acta*.
608. Zhang, X., Guo, C., Chen, Y., Shulha, H.P., Schnetz, M.P., LaFramboise, T., Bartels, C.F., Markowitz, S., Weng, Z., Scacheri, P.C., and Wang, Z. (2008). Epitope tagging of endogenous proteins for genome-wide ChIP-chip studies. *Nat. Methods* 5, 163-165.
609. Zhang, Z., Fu, J., and Gilmour, D.S. (2005b). CTD-dependent dismantling of the RNA polymerase II elongation complex by the pre-mRNA 3'-end processing factor, Pcf11. *Genes Dev.* 19, 1572-1580.
610. Zhao, X., and 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.
611. Zhao, Y., Lang, G., Ito, S., Bonnet, J., Metzger, E., Sawatsubashi, S., Suzuki, E., Le, G., X, Stunnenberg, H.G., Krasnov, A., Georgieva, S.G., Schule, R., Takeyama, K., Kato, S., Tora, L., and Devys, D. (2008). A TFTC/STAGA module mediates histone H2A and H2B deubiquitination, coactivates nuclear receptors, and counteracts heterochromatin silencing. *Mol. Cell* 29, 92-101.
612. Zhou, H., and Lee, K.A. (2001). An hsRPB4/7-dependent yeast assay for trans-activation by the EWS oncogene. *Oncogene* 20, 1519-1524.
613. Zhou, K., Kuo, W.H., Fillingham, J., and Greenblatt, J.F. (2009). Control of transcriptional elongation and cotranscriptional histone modification by the yeast BUR kinase substrate Spt5. *Proc. Natl. Acad. Sci. U. S. A* 106, 6956-6961.
614. Zhu, A., and Kuziora, M.A. (1996). Homeodomain interaction with the beta subunit of the general transcription factor TFIIE. *J. Biol. Chem.* 271, 20993-20996.
615. Zhu, X., Wiren, M., Sinha, I., Rasmussen, N.N., Linder, T., Holmberg, S., Ekwall, K., and Gustafsson, C.M. (2006). Genome-wide occupancy profile of mediator and the srb8-11 module reveals interactions with coding regions. *Mol. Cell* 22, 169-178.
616. Ziegler, L.M., Khapersky, D.A., Ammerman, M.L., and 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.

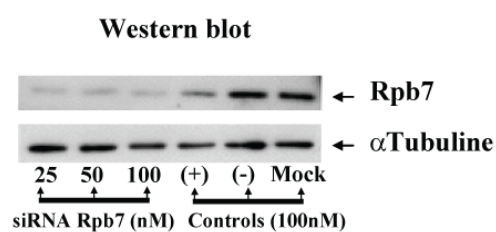
617. Zippo, A., Serafini, R., Rocchigiani, M., Pennacchini, S., Krepelova, A., and Oliviero, S. (2009). Histone crosstalk between H3S10ph and H4K16ac generates a histone code that mediates transcription elongation. *Cell* 138, 1122-1136.
618. Zurita, M., and Merino, C. (2003). The transcriptional complexity of the TFIID complex. *Trends Genet.* 19, 578-584.

Annexe I

A



B



C

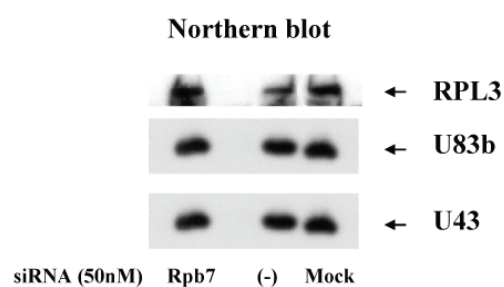
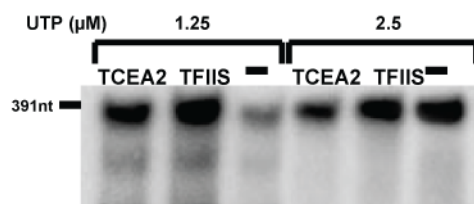


Figure 1. Analyse du rôle de Rpb7 dans la formation des ARN de RPL3, U83b et U43. A) Analyse par ChIP-Q-PCR de la localisation des sous-unités Rpb7 et Rpb11 portant une étiquette TAP au long du gène RPL3, effectuée en conditions normales ou après une exposition des cellules à un choc thermique (HS), tel que décrit dans le Chapitre 2. En bas : le schéma du gène RPL3 et la position des oligos utilisés sont représentées. B) Efficacité du knock-down du Rpb7 avec des ARN d'interférence,

analysée comme décrit dans le Chapitre 4. Les concentrations des différents ARN d'interférence sont indiquées, incluant les contrôles qui n'ont pas de cible (-) ou qui ciblent un autre ARN (MEPCE) ou une transfection mock. C) Analyse par Northern blot de l'effet de l'atténuation de l'expression de Rpb7 sur la formation des ARN de RPL3, U83b et U43. Des sondes d'ARN reconnaissant les ARN indiqués ont été utilisées.

A



B

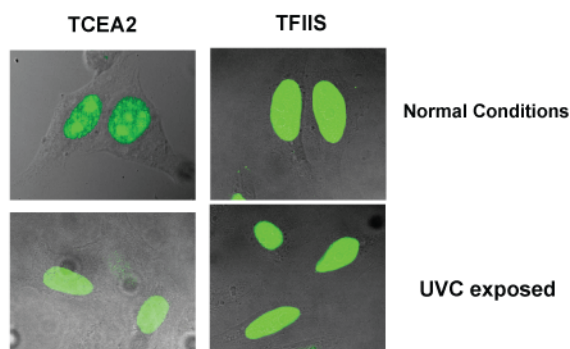


Figure 2. Analyse comparative de TFIS.1 et TFIS. (A) L'activité transcriptionnelle *in vitro* de TFIS.1 analysée par des essais de type "run-off". Le système reconstitué avec des facteurs recombinants TBP, TFIIB, Rap30, Rap74, TFIE34, TFIE56 (30 ng chacun), TFIIH (4 ng), RNAPII (165 ng) et une matrice d'ADN linéaire (500 ng) constituée par le promoteur AdML (-10 à +10) fusionné à une cassette dépourvue de guanine a été incubé avec des NTPs incluant différentes

concentrations de UTP non-réactif tel qu'indiqué dans l'image et les facteurs recombinants TFIIS.1 ou TFIIS (4 ng). Comme contrôle, le même système reconstitué sans ces deux derniers facteurs a été utilisé. Un produit "run-off" de 391 nt est transcrit dans ce système. (B) Localisation cellulaire de TFIIS.1-GFP ou TFIIS-GFP analysée dans des cellules HeLa en conditions normales ou après une exposition des cellules à des radiations UVC (12.5J/m²) et un temps de récupération de 20min. à 37C.

