



HAL
open science

Étude intégrative du rôle de deux sous unités essentielles du Médiateur de la transcription dans la mise en place des complexes de pré-initiation

Thomas Eychenne

► **To cite this version:**

Thomas Eychenne. Étude intégrative du rôle de deux sous unités essentielles du Médiateur de la transcription dans la mise en place des complexes de pré-initiation. Biologie moléculaire. Université Paris-Saclay, 2016. Français. NNT : 2016SACLS266 . tel-01558266

HAL Id: tel-01558266

<https://theses.hal.science/tel-01558266>

Submitted on 7 Jul 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

NNT : 2016SACLS266

Thèse de doctorat
de
L'Université Paris-Saclay
préparée à
L'Université Paris Sud

Ecole Doctorale n° 577
Structure et Dynamique des Systèmes Vivants
Spécialité de doctorat : Sciences de la Vie et de la Santé

Par

M. Thomas EYCHENNE

Etude intégrative du rôle de deux sous unités essentielles du Médiateur de la transcription
dans la mise en place des complexes de pré-initiation.

Thèse présentée et soutenue à Gif sur Yvette, le 22 septembre 2016 :

Composition du Jury :

Monsieur Capy Pierre
Monsieur Davidson Irwin
Monsieur Libri Domenico
Monsieur Gazin Claude
Monsieur Werner Michel
Madame Soutourina Julie

Professeur de l'Université Paris Sud
Directeur de Recherche IGBMC Strasbourg
Directeur de Recherche Institut Jacques Monod
Chargé de Recherche CNRS
Directeur de Recherche CEA
Directrice de Recherche CEA

Président
Rapporteur
Rapporteur
Examineur
Directeur de thèse
Co-directrice de thèse

*« Pour se frayer un sentier nouveau,
Il faut être capable de s'égarer »
Jean Rostand*

Remerciements

Me voici arrivé au moment d'écrire cette page de remerciement qui, au final, est l'une des plus difficiles à écrire de ce manuscrit ! Après pas loin de 5 ans passés dans le laboratoire, j'ai eu l'opportunité de vivre beaucoup de choses et de rencontrer beaucoup de monde, j'aimerais n'oublier personne.

Je vais d'abord remercier les membres de mon jury de thèse : Domenico Libri, Irwin Davidson, Pierre Capy et Claude Gazin qui ont accepté avec gentillesse, d'évaluer et de juger mon travail afin de sanctionner ces années de travail d'un succès. J'ajoute une mention particulière pour Claude Gazin qui a, en plus de cela, accepté d'être mon tuteur tout au long de mon parcours. Il a, durant ces années, pu suivre mon travail pour lequel il a toujours été positif et enthousiaste. Ces réunions de comité de thèse ont toujours abouti à des nouvelles pistes à creuser.

Je remercie également Julie et Michel de m'avoir accueilli dans le laboratoire qui a subi beaucoup de changements durant mes années de présence en étant d'abord celui de Michel puis celui de Julie. Merci à toi Michel de m'avoir accepté pour faire mon stage de M2 et de pouvoir poursuivre par une thèse. Merci d'avoir été mon directeur de thèse et, malgré tes prises de galon successives, d'avoir su trouver du temps quand cela était nécessaire pour que nous puissions discuter de mes résultats. Merci à toi Julie d'avoir accepté de devenir, à très juste titre, ma co-directrice de thèse après toute l'aide que tu as pu m'apporter tout au long de ma thèse. Merci pour ta disponibilité permanente et pour toutes les discussions que nous avons pu avoir. Quand je suis arrivé, je ne connaissais pas le travail avec la levure et tu m'as appris bon nombre de méthodes avec la rigueur et l'efficacité qui te caractérisent. Merci à vous deux, Michel et Julie, de m'avoir permis de bénéficier d'une grande liberté et d'une autonomie totale qui m'ont permis d'explorer autant d'aspect de mon projet de thèse. Je garderai un souvenir exceptionnel de mon passage dans ce laboratoire qui m'aura permis de m'épanouir, de pouvoir avoir la chance de voyager pour aller chaque année à des congrès mais surtout de continuer à aimer la recherche malgré les difficultés qui peuvent en découler.

Je remercie également tous mes collègues du laboratoire que j'ai eu l'opportunité de côtoyer. Merci Claire, qui a réalisé les cribles pour obtenir les mutants sur lesquels tout mon travail a reposé et avec qui j'ai passé d'excellents moments. Tu faisais partie avec moi du « Club » des usagers de la Francilienne qui peuvent mettre 2 heures pour venir au labo ! Merci aux « post-doc du bureau d'en face ». Camille avec qui il a été agréable d'essayer d'aller courir (certes que 2 fois seulement !) sur le centre du CEA et de discuter bio-info et autre. Merci à Nuno, qui a déblayé le terrain par son travail pour que nous puissions faire les manips de biochimie pour mon projet, mais aussi pour ton aide et tes conseils. Grâce à toi, même après ton départ la Déesse du labo a veillé sur nous ! Merci aussi à Sylvie pour son implication dans la vie du labo. Un grand merci à Fanny, ma collègue de derrière mon écran qui a toujours su apporter du pétillant à la vie du labo et qui m'a initié à la bio-informatique. Merci à Jérémie, du « labo des CNRS », nos conversations ont toujours été du meilleur goût ! Les soirées mémorables avec Fanny et Jérémie nous ont permis de tester de nouveaux styles vestimentaires mais surtout de décompresser et de vivre de bons moments en dehors de la thèse.

Arrivé à la moitié de ma thèse, de nouveaux visages sont arrivés. Merci à Adrien, avec qui il a toujours été agréable d'échanger. Merci aussi à Diya, ma nouvelle collègue de derrière mon écran, héritière de la gestion des enzymes de restrictions et dorénavant seule représentante des thésards du labo, les échanges avec toi ont toujours été agréables ! Bon courage pour la fin de ta thèse ! Elodie, dont le passage a été assez rapide mais fort sympathique. Cyril, que je n'ai pas pu côtoyer longtemps mais avec qui il a été plaisant de travailler. Liza, qui m'a permis de parfaire mon anglais mais surtout qui a permis d'apporter une dimension supplémentaire à mon travail par ses analyses de bio-informatique. Merci aussi à Pierre Nicolas qui a travaillé avec Liza pour développer ces analyses. Merci à Olivier A, le super bio-informaticien qui a mis au point beaucoup d'outils indispensables pour faire nos analyses de ChIP-seq. Mes visites à Evry pour travailler avec toi ont toujours été plaisantes. Merci aussi de m'avoir accueilli à Evry pour une bonne partie de la rédaction de ce manuscrit. Merci à Nathalie, arrivée en remplacement de Claire, qui m'a apporté de l'aide pour certaines manips mais surtout pour toutes les doses de café que je t'ai « emprunté » ! Merci aussi à Bénédicte, qui nous a rejoint récemment. Cela a été un plaisir de bénéficier de ton expertise des manips et grâce à ton aide, nous avons pu ajouter l'aspect *in vitro* de mon travail. Tu devrais maintenant poursuivre une partie de mon travail et je suis sûr que tu en feras des choses super. Merci aussi à Claire P et Arach qui ont rejoint le laboratoire il y a peu. Merci à Guillaume pour tes conseils pour mes plusieurs tentatives avec les cellules humaines mais aussi

pour les pauses cafés partagées avec toi. Merci également à Cindy et Anna pour leur aide dans cette partie « cellules humaines ». Merci à David, Virginie et Laila, de la plateforme de spectrométrie de masse pour leur investissement dans l'analyse de mes manips de purification du Médiateur.

Je n'oublierais bien sûr pas de remercier Catherine et Chantal puis Monique, les secrétaires du service. Votre travail indispensable a toujours été d'une très grande efficacité. A chaque déplacement, tout a toujours été géré d'une main de maître ! Merci aussi à Manu, qui a géré la commande de tous les produits et réactifs dont j'ai pu avoir besoin. Je remercie toutes les personnes du service que j'ai eu la chance de rencontrer durant toutes ces années : Thaïz, Eugène, Matthieu, Sylvain, Regina, Irina, Jean-Marc, Pauline, Véronique, Arlette, Cristelle, Rosa, Olivier, Joël, Christine, Linda, Marie-Claude, Marie, Anne, Agathe, Céline... et sans doute certains que j'oublie...

Merci à Jeanne, qui a été présente pour moi pendant cette période difficile qu'est la thèse et qui m'a apporté le soutien nécessaire. Merci pour ton implication dans l'organisation de la soutenance, je crois qu'on s'est tous bien régalez grâce à toi.

Merci à mes parents mais aussi à Patrice, qui ont su m'élever et me faire devenir celui que je suis aujourd'hui et me mener jusqu'ici. Merci à mes frères aussi. Nicolas, nous sommes 2 Docteurs Eychenne maintenant ! Plus largement je remercie toute ma famille. J'ai une pensée particulière pour mes grands-parents René et Henriette, partis trop tôt et à qui j'aurais été très fier de montrer ce travail.

Table des matières

Remerciements	5
Table des matières	9
Table des illustrations	13
Liste des abréviations.....	17
Avant-Propos.....	21
Introduction	23
Chapitre I : La transcription des gènes chez les eucaryotes.....	25
I. L'ARN polymérase II.	26
1) Composition de l'ARN polymérase II et conservation chez les eucaryotes.....	27
a) Conservation des ARN polymérases.....	27
b) Composition des ARN polymérases eucaryotes.....	29
2) Structure de l'ARN polymérase II et fonction de ses sous-unités.....	30
a) Structure tridimensionnelle de la Pol II.....	30
b) Fonction des différentes sous-unités de la Pol II.....	31
3) Le Domaine Carboxy-Terminal de l'ARN polymérase II.....	34
II. La transcription par l'ARN polymérase II.....	36
1) Structure des gènes de classe II.....	36
a) La boîte TATA.....	37
b) L'élément de reconnaissance de TFIIB (BRE).....	38
c) La séquence initiateur (Inr).....	39
d) L'élément promoteur aval (DPE).....	39
2) Les facteurs généraux de la transcription.	40
a) TFIID.....	41
b) TFIIA.....	42
c) TFIIB.....	43
d) TFIIF.....	44
e) TFIIE.....	45
f) TFIIH.....	45
3) L'initiation de la transcription.	46
a) La mise en place du complexe de pré-initiation.....	46
b) La transition entre l'initiation et l'élongation.....	50
4) L'élongation de la transcription par la Pol II.....	51
a) La pause à proximité du promoteur.....	52
b) L'élongation productive.....	54

5)	La terminaison de la transcription et sa ré-initiation.	54
a)	La terminaison.	54
b)	La ré-initiation de la transcription.	55
6)	La dynamique de la transcription.	55
III.	La transcription pervasive.	59
IV.	La transcription dans le contexte chromatinien.	61
1)	Structure de la chromatine.	62
2)	La modification des histones.	63
3)	Les variants d'histones.	65
4)	Le remodelage de la chromatine.	65
V.	La régulation de la transcription par l'ARN polymérase II.	66
1)	Les éléments de régulation des promoteurs.	67
2)	Les facteurs de transcription.	68
a)	Les activateurs de la transcription.	68
b)	Les répresseurs de la transcription.	69
3)	Les co-activateurs.	69
a)	Les complexes agissant sur la mise en place du PIC.	70
b)	Les remodeleurs de la chromatine.	73
Chapitre II : Le Médiateur de la transcription par l'ARN polymérase II.		77
I.	Le Médiateur de <i>Saccharomyces cerevisiae</i>	77
1)	Découverte du Médiateur.	77
2)	Composition et organisation du Médiateur.	78
3)	Structure du Médiateur.	79
II.	Le Médiateur chez les mammifères.	83
1)	Découverte du Médiateur des mammifères.	83
2)	Composition du Médiateur des mammifères.	84
3)	Conservation du Médiateur au cours de l'évolution.	85
III.	Rôle du Médiateur dans la régulation de l'expression des gènes.	87
1)	Interaction entre le Médiateur et les activateurs spécifiques.	87
2)	Interaction entre le Médiateur et la Pol II.	91
3)	Interactions entre le Médiateur et les facteurs généraux de la transcription.	93
a)	Médiateur et TFIID.	93
b)	Médiateur et TFIIA.	94
c)	Médiateur et TFIIIB.	94
d)	Médiateur et TFIIIF.	94
e)	Médiateur et TFIIIE.	95
f)	Médiateur et TFIIH.	95
4)	Interaction du Médiateur avec les co-activateurs.	95

5)	Interaction du Médiateur avec d'autres protéines impliquées dans la transcription.	97
a)	Interaction entre le Médiateur et la cohésine.....	97
b)	Interaction entre le Médiateur et TFIIIS.	97
6)	Rôle du Médiateur dans l'assemblage du PIC <i>in vivo</i>	98
7)	Le Médiateur en tant que facteur général de la transcription.	101
IV.	La régulation négative de la transcription par le Médiateur.	102
V.	Rôles du Médiateur au-delà de l'initiation de la transcription.....	103
1)	Rôle du Médiateur dans les étapes de la transcription <i>post</i> -assemblage du PIC.....	103
2)	Rôle du Médiateur dans la ré-initiation de la transcription.	104
3)	Rôle du Médiateur dans le maintien des télomères.	105
4)	Rôle du Médiateur dans la réparation de l'ADN.	105
VI.	Rôles du Médiateur dans des processus du développement.....	107
VII.	Pathologies associées au Médiateur.	109
1)	Maladies du développement.....	109
2)	Implication du Médiateur dans les cancers humains.	110
	Résultats et discussion	115
	Partie I : Etude intégrative du rôle de la sous-unité Med10 dans l'initiation de la transcription....	117
I.	Contexte de travail.	117
II.	Article	121
	Partie II : Etude du rôle de la sous-unité essentielle Med7.....	123
I.	Contexte de travail.	123
II.	Résultats.....	123
1)	Collection de mutants thermosensibles de Med7.....	123
2)	Interactions entre Med7 et ses partenaires au sein du Médiateur.....	126
3)	Analyse de l'intégrité du Médiateur dans les contextes mutants.	127
4)	Effet des mutations de Med7 sur l'occupation des composants du PIC.	128
	Conclusions et perspectives	133
I.	Collection de mutants thermosensibles de deux sous-unités du module du milieu du Médiateur.....	135
II.	Analyse de l'intégrité du Médiateur dans les mutants de ses sous-unités.	136
III.	Interaction entre le Médiateur et TFIIB.....	136
IV.	Rôle du Médiateur dans l'assemblage du PIC.	138
1)	Effet des mutations de sous-unités du module du milieu sur la formation du PIC.	138
2)	Lien entre le Médiateur et l'architecture des promoteurs.....	140
3)	Effet gène-spécifique des mutations de la sous-unité Med7 : étude du phénotype de floculation.	141
4)	Rôle des différentes sous-unités essentielles dans la formation du PIC <i>in vivo</i>	142
V.	Analyse structurale du Médiateur.....	142

VI. Conservation du rôle de Med7, Med10 et du Médiateur en général chez les mammifères.	143
Méthodologie.....	147
Chapitre I : Purification et analyse du Médiateur par spectrométrie de masse.	149
I) Principe.....	150
II) Protocole de la purification.	152
Chapitre II : Analyse de la formation du PIC <i>in vitro</i>	157
I) Principe.....	157
II) Protocole.	158
1) Purification des activateurs.	158
2) Préparation de la matrice d'ADN.....	158
3) Préparation des extraits nucléaires de levure.....	158
4) Formation du PIC <i>in vitro</i>	162
Chapitre III : Analyses bioinformatiques des données de CHIP-seq.	165
I) Principe du CHIP-seq.....	165
II) Traitement des données issues du séquençage.....	165
III) Analyse des données.	170
1) Analyse des effets globaux des mutations du Médiateur sur la formation du PIC.	170
2) Analyse de l'effet gène-spécifique des mutations du Médiateur sur la formation du PIC : cas de l'architecture des promoteurs.....	171
Bibliographie	173
Annexes.....	209
Chapitre I : Publications.	211

Table des illustrations

Figure 1: Déroulement d'un cycle de transcription. Schéma simplifié d'un cycle de transcription par la Pol II. (Adapté depuis Orphanides and Reinberg 2002).	26
Figure 2 : Comparaison des structures des ARN polymérases bactérienne (A), d'archée (B) et de la Pol II eucaryote (C).	28
Figure 3 : Exemple de structure atomique de l'ARN polymérase II de <i>S. cerevisiae</i> obtenue par diffraction des rayons (D'après Armache et al. 2005).	31
Figure 4: Répartition et évolution des modifications du CTD au cours de la transcription d'un gène visualisé par CHIP chez la levure. (D'après Eick and Geyer 2013b).	35
Figure 5 : Représentation schématique de la structure d'un promoteur cœur et des protéines interagissant avec les différents domaines. (Thomas and Chiang 2006)	37
Figure 6: Schéma représentant les différents domaines structuraux de la protéine TFIIB (Sainsbury, Niesser, and Cramer 2013).	43
Figure 7 : Schéma représentant l'assemblage séquentiel et linéaire du PIC <i>in vitro</i> . (d'après Sainsbury, Bernecky, and Cramer 2015).	47
Figure 8 : Modèles structuraux du PIC dans les conformations fermées, ouvertes et en début de transcription (Plaschka et al. 2016).	49
Figure 9: Schéma représentatif de l'ensemble de la transcription par l'ARN polymérase II (d'après Koch et al. 2008).	51
Figure 10: Schéma représentant les trois modèles proposés pour l'établissement des pauses de la Pol II (d'après Kwak and Lis 2013)	53
Figure 11: Schéma explicatif de la méthode FRAP (d'après Phair and Misteli 2001)	56
Figure 12: Schéma explicatif de la méthode du CHIP compétitif (d'après Lickwar, Mueller, and Lieb 2013)	57
Figure 13: Schéma représentant la méthode de marquage MS2/PP7 des ARNs (D'après Lenstra and Larson 2016).	58
Figure 14: Schéma représentant l'organisation de l'ADN sous forme de chromatine. (d'après Griffiths et al., 2004, http://www.mun.ca/).	61
Figure 15 : Schéma représentant la régulation de l'initiation de la transcription par l'ARN polymérase II chez <i>S. cerevisiae</i>	66
Figure 16: Composition et organisation modulaire du Médiateur de <i>S. cerevisiae</i> (d'après Guglielmi et al. 2004).	78
Figure 17: Carte des interactions entre les sous-unités du Médiateur (Guglielmi et al. 2004)	79
Figure 18: Modèle structural du module de tête du Médiateur (Larivière et al. 2012).	80
Figure 19: Modèle structural du module du milieu du Médiateur (d'après Larivière et al. 2013).	81
Figure 20 : Modèle structural du Médiateur de <i>S. cerevisiae</i> obtenu par cryo-EM (d'après Tsai et al. 2014).	82
Figure 21: Modèle structural du Médiateur proposé par modélisation intégrative (d'après Robinson et al. 2015)	83
Figure 22: Composition en sous-unités du Médiateur des eucaryotes pluricellulaires (Malik and Roeder 2010).	84
Figure 23: Conservation des sous-unités du Médiateur chez différents eucaryotes (d'après Bourbon 2008).	86
Figure 24: Comparaison des structures 3D des Médiateurs humain et de levure observés par microscopie électronique (Tsai et al. 2014).	87
Figure 25: Modèle structural obtenu par cryo-microscopie électronique du PIC partiel composé de la Pol II, TFIIA, TFIIB, TBP, TFIIF et du Médiateur cœur (d'après Plaschka et al. 2015b).	92

Figure 26: Schéma représentant les interactions entre le Médiateur et certains co-activateurs (d'après Ansari and Morse 2013)	96
Figure 27: Schéma montrant l'activation de la transcription <i>in vivo</i> et les différentes interventions du Médiateur dans ce processus (d'après Malik and Roeder 2010)	99
Figure 28: Modèle d'assemblage branché du PIC influencé par le Médiateur (Esnault et al. 2008; Soutourina et al. 2011; Eyboulet et al. 2015).....	100
Figure 29: Schéma de la voie de réparation de l'ADN par excision de nucléotides (Boiteux and Jinks-Robertson 2013).....	106
Figure 30: Phénotype thermosensible des mutants de Med7 obtenus par PCR mutagène aléatoire (à gauche) et par mutagénèse dirigée (à droite).....	124
Figure 31: Structure de la sous-unité Med7 obtenue par cristallographie et diffraction des rayons X. (Baumli, Hoepfner, and Cramer 2005)	125
Figure 32: Phénotype de floculation de certains mutants de Med7.	126
Figure 33: Analyse par double-hybride des interactions entre les versions mutées de Med7 et les sous-unités Med9, Med10 et Med21. Les couleurs indiquent des mutants dont le profil d'interaction est similaire.	126
Figure 34: Analyse par co-immunoprécipitation de l'intégrité du Médiateur dans le cas des mutants sélectionnés pour nos études.	128
Figure 35: Niveau d'occupation de la Pol II sur les promoteurs et ORFs des gènes <i>ADH1</i> , <i>PYK1</i> et <i>PMA1</i> mesuré par ChIP pour les mutants <i>med7-5</i> , <i>101</i> , <i>131</i> , <i>173</i>	129
Figure 36: Niveau d'occupation de TFIIH sur les promoteurs des gènes <i>ADH1</i> , <i>PYK1</i> et <i>PMA1</i> mesuré par ChIP pour les mutants <i>med7-5</i> , <i>101</i> , <i>131</i> , <i>173</i>	129
Figure 37: Niveau d'occupation de TFIID (TBP) sur les promoteurs des gènes <i>ADH1</i> , <i>PYK1</i> et <i>PMA1</i> mesuré par ChIP pour les mutants <i>med7-5</i> , <i>101</i> , <i>131</i> , <i>173</i>	130
Figure 38 : Niveau d'occupation des modules de queue (Med15) et de tête (Med17) du Médiateur sur les promoteurs des gènes <i>ADH1</i> , <i>PYK1</i> et <i>PMA1</i> mesuré par ChIP pour les mutants <i>med7-5</i> , <i>101</i> , <i>131</i> et <i>173</i>	130
Figure 39 : Position sur la protéine des résidus mutés dans les différents mutants de TFIIB utilisés dans notre étude (Sainsbury, Niesser, and Cramer 2013).	137
Figure 40: Schéma représentant un modèle de régulation type "interrupteur" de l'expression du gène <i>FLO11</i> (d'après Bumgarner et al. 2009).....	141
Figure 41: Alignement des séquences protéiques de Med10 chez l'Homme et la levure <i>S. cerevisiae</i> (Bourbon 2008).	143
Figure 42: Test de prolifération des lignées cellulaires HeLa sauvages ou exprimant stablement <i>Med10-WT</i> ou <i>med10-196</i>	144
Figure 43: Alignement des séquences protéiques de Med7 chez l'Homme et la levure <i>S. cerevisiae</i> . (D'après Bourbon 2008)	145
Figure 44: Schéma explicatif de la démarche expérimentale mise en œuvre pour l'étude de la composition du Médiateur par spectrométrie de masse.....	152
Figure 45 : Schéma de principe de l'expérience d'assemblage du PIC <i>in vitro</i>	157
Figure 46 : Visualisation sur un explorateur de génome des signaux de ChIP-seq de la Pol II et des pics détectés par l'analyse de <i>Peak calling</i>	168
Figure 47 : Exemple de nuage de point représentant les densités de lecture sur les régions codante dans le cas d'un ChIP de la Pol II.	170
Figure 48 : Exemple de représentation logarithmique des occupations de la Pol II dans les souches sauvage et mutante de Med10 et représentation des trois groupes de gènes définis.....	171
Figure 49 : Exemple de représentations graphiques des analyses de données de ChIP-seq en termes d'occupation des nucléosomes, de présence de boîtes TATA et de dynamique des nucléosomes pour les trois groupes de gènes définis pour la Pol II.	172

Tableau 1: Composition en sous-unités des ARN polymérase bactériennes, archées et eucaryotes (d'après Werner and Grohmann 2011).	27
Tableau 2 : Récapitulatif des fonctions des sous-unités de la Pol II.	32
Tableau 3: Rôle et composition en sous-unités des facteurs généraux de la transcription (d'après Thomas and Chiang 2006).	40
Tableau 4: Tableau récapitulatif des modifications d'histones. (D'après Li, Carey, and Workman 2007)	64
Tableau 5: Tableau récapitulatif des différents variant d'histones, leur localisation et leur effet sur la transcription. (D'après Li, Carey, and Workman 2007).	65
Tableau 6 : Nomenclature unifiée pour les TAFs (Tora 2002).	71
Tableau 7: Tableau récapitulatif des sous-unités de SAGA chez la levure et chez l'Homme (d'après Rodríguez-Navarro 2009).	72
Tableau 8: Tableau récapitulatif des classes de remodeleurs de la chromatine et leur représentant chez la levure, la drosophile et l'Humain (d'après Clapier and Cairns 2009).	73
Tableau 9: Tableau récapitulatif des interactions connues entre différentes sous-unités du Médiateur et des facteurs de transcription. (Adapté depuis Sohail Malik and Roeder 2010; Ansari and Morse 2013)	90
Tableau 10: Tableau récapitulatif des fonctions connues des sous-unités du Médiateur dans des processus physiologiques et développementaux (d'après Yin and Wang 2014)	108
Tableau 11: Implication de sous-unités du Médiateur dans des pathologies humaines du développement (d'après Spaeth, Kim, and Boyer 2011).	110
Tableau 12: Implications de sous-unités du Médiateur dans ces cancers et causes génétiques associées (d'après Spaeth, Kim, and Boyer 2011; Schiano et al. 2014).	113
Tableau 13: Tableau récapitulatif des résidus mutés dans les différents mutants de Med7.	124
Tableau 14: Tableau récapitulatif des facteurs de normalisation appliqués à nos données de ChIP-seq dans l'étude du mutant <i>med10-196</i>	167

Liste des abréviations.

3C	<i>Chromatin conformation capture</i>
5FOA	<i>5-Fluoroorotic acid</i>
acetylCoA	Acétyl co-enzyme A
AF2	<i>Activation Function 2</i>
AR	<i>Androgen Receptor</i>
ARNm	ARN messenger
ARNnc	ARN non-codant
ARNr	ARN ribosomal
ARNt	ARN de transfert
ATP	<i>Adénosine tri-phosphate</i>
BRE	<i>Site de reconnaissance de TFIIIB, TFIIIB recognition element</i>
CBP	<i>Calmodulin Binding Protein</i>
Cdk	<i>cyclin-dependant kinase</i>
CF	<i>cleavage factor</i>
CHD	<i>Chromodomain helicase DNA binding</i>
ChIP	Immunoprécipitation de la chromatine
ChIP-exo	<i>ChIP suivie d'une digestion de l'ADN par l'exonucléase I</i>
ChIP-seq	ChIP suivie d'un séquençage à haut-débit
CID	<i>CTD interacting domain</i>
co-IP	co-immunoprécipitation
CPD	<i>Cyclobutane-pyrimidine dimer</i>
CPF	<i>cleavage and polyadenylation factor</i>
CPSF	<i>cleavage and polyadenylation specificity factor</i>
CRC	<i>Colorectal Cancer</i>
CRISPR/Cas9	<i>Clustered Regularly Intespaced Short Palindromic Repeats</i>
CRSP	<i>Cofactor Required for Sp1 Activation</i>
cryo-EM	<i>cryo Microscopie Electronique</i>
CS	<i>Syndrôme de Cokayne</i>
CstF	<i>cleavage stimulatory factor</i>
CTD	domaine Carboxy terminal, <i>carboxy terminal domain</i>
cTFIIH	<i>core-TFIIH</i>
CUT	<i>cryptic unstable transcript</i>
Da	Dalton
DBD	<i>DNA-binding domain</i>
DCE	<i>downstream core element</i>
DO	Densité Optique
DPE	<i>downstream promoter element</i>

DRIP	<i>Vitamine D Receptor Interacting Proteins</i>
DSIF	<i>DRB-sensitivity inducing factor</i>
DTT	<i>DiThiothréitol</i>
DUB	<i>Dé-ubiquitine</i>
EDTA	<i>Ethylène Diamine Tetra-acetic Acid</i>
EGR1	<i>Early Growth Response 1</i>
EII	<i>Eleven nineteen leucine-rich leukemia 1</i>
enChIP	<i>engineered-DNA binding molecule-mediated ChIP</i>
ENCODE	<i>encyclopedia of DNA Element</i>
ER	<i>Estrogen Receptor</i>
FACT	<i>facilitates chromatin transcription</i>
FRAP	<i>fluorescence recovery after photobleaching</i>
GFP	<i>green fluorescent protein</i>
GGR	<i>Global genome repair</i>
GR	<i>Glucocorticoid Receptor</i>
gRNA	<i>guide RNA</i>
GST	<i>Glutathione-S-Transferase</i>
GTF	<i>Facteur général de la transcription, general transcription factor</i>
HAT	<i>Histone acétyl-transférase</i>
HDAC	<i>Histone deacetylase</i>
HNF4	<i>Hepatocyte Nuclear Factor 4</i>
HSA	<i>helicase- SANT</i>
HSP	<i>Heat Shock protein</i>
HTH	<i>Hélice-tour-Hélice</i>
IEG	<i>Immediate Early Gene</i>
INO80	<i>Inositol requiring 80</i>
Inr	<i>Element initiateur de la transcription</i>
ISWI	<i>Imitation SWI</i>
LC	<i>Liquid chromatography</i>
lncRNA	<i>long non-coding RNA</i>
MACS	<i>Model-based Analysis of ChIP-Seq</i>
Mb	<i>Méga Base</i>
MNase	<i>Nucléase Micrococcale</i>
MS	<i>Mass spectrometry</i>
MS/MS	<i>Spectrométrie de Masse en tandem</i>
MTE	<i>motif ten element</i>
MudPit	<i>Multi Dimensionnal Protein Identification Technology</i>
N	<i>Nucléotide</i>
NAT	<i>Negative Regulator of Activated Transription</i>
NDR	<i>nucleosome depleted region</i>
NELF	<i>negative elongation factor</i>

NER	<i>Réparation de l'ADN par excision de nucléotides, nucleotide excision repair</i>
NET-seq	<i>native elongating transcript</i>
NFR	<i>Nucleosome free region</i>
NNS	<i>Nrd1-Nap3-Sen1</i>
NR	<i>Nuclear Receptor</i>
NTP	<i>Nucléotide tri-phosphate</i>
P-TEFb	<i>positive transcription elongation factor</i>
PASR	<i>promoter associated small RNA</i>
PC2	<i>Positive Cofactor 2</i>
PIC	<i>Complexe de pré-initiation, preinitiation complex</i>
Pol	<i>ARN polymérase</i>
Pol II	<i>ARN polymérase II</i>
poly(A)	<i>Poly-adenylation</i>
PPAR	<i>Peroxisome Proliferator Activated Receptor</i>
PROMPT	<i>promoter upstream transcript</i>
PSA	<i>Prostatic Specific Antigene</i>
Py	<i>Pyrimidine</i>
RAR	<i>Récepteur à la Vitamine A, retinoic acid receptor</i>
RSC	<i>Remodel structure of Chromatin</i>
SAGA	<i>Spt-Ada-Gcn5 acetyltransferase</i>
SALSA	<i>SAGA-like Spt8 absent</i>
SDS-PAGE	<i>SDS-Polyacrylamide Gel Electrophoresis</i>
SEC	<i>Super Elongation Complex</i>
shRNA	<i>short hairpin RNA</i>
SILAC	<i>Stable Isotope Labeling with Amino Acids in Cell culture</i>
siRNA	<i>small interfering RNA</i>
SLIK	<i>SAGA-like</i>
SMCC	<i>SRB Mediator Containg Cofactor</i>
snARN	<i>small nuclear RNA</i>
snoRNA	<i>small nucleolar RNA</i>
SPT	<i>Single Particle Tracking</i>
SRB	<i>Suppressor of RNA Polymerase B truncation</i>
SREBP1a	<i>Smad2/3/4 and sterol Regulatory Element Binding Protein 1 a</i>
SUT	<i>stable unannotated transcript</i>
Swi/Snf	<i>Switch defecting/Sucrose Non-Fermenting</i>
TA	<i>transcription activator</i>
TAF	<i>TBP associated factor</i>
TAND	<i>Taf1-N-terminal domain</i>
TAP	<i>Tandem Affinity Purification</i>
TARS	<i>terminator associated small RNA</i>
TBP	<i>TATA-box binding protein</i>

TCR	<i>Transcription-coupled repair</i>
TEV	<i>Tobacco Etch Virus</i>
TF	<i>transcription factor</i>
TPE	<i>Telomeres Position Effect</i>
TR	<i>transcription repressor</i>
TRAP	<i>Thyroid-hormone Receptor Associated Protein</i>
ts	<i>Thermo-sensible</i>
TSS	<i>Site de démarrage de la transcription, transcription start site</i>
TTD	<i>Trichothiodystrophie</i>
TTS	<i>Site de terminaison de la transcription, transcription terminaison site</i>
UAS	<i>Upstream activating sequence</i>
URS	<i>Upstream repressive sequence</i>
UTR	<i>Untranslated Region</i>
VDR	<i>Vitamine D Receptor</i>
XP	<i>Xéoderma Pigmentosum</i>
XUT	<i>Xrn1-sensitive unstable transcript</i>
YPD	<i>Yeast-extract Peptone Dextrose</i>

Avant-Propos

Le Médiateur de la transcription a été découvert il y a maintenant plus de vingt ans. Il a été mis en évidence d'abord chez la levure *Saccharomyces cerevisiae* puis sa présence a été étendue à l'ensemble des eucaryotes. Ce complexe protéique est essentiel pour l'expression des gènes grâce à son rôle de co-régulateur dans l'initiation de la transcription. Au cours des deux dernières décennies, il a fait l'objet d'un grand nombre d'études et les mécanismes dans lesquels il est impliqué sont à présent plus clairs. Néanmoins, il reste un certain nombre de questions sans réponses notamment sur sa fonction au niveau moléculaire ou encore quant à sa structure tridimensionnelle complète. Le Médiateur se voit également attribuer des rôles dépassant le cadre de la transcription, suggérant qu'il est impliqué dans un grand nombre de processus nucléaires. Enfin, ces dernières années, des pathologies telles que des cancers ou des maladies mentales ont été associées à des défauts d'expression ou des mutations de certaines sous-unités du Médiateur. Cela en fait un objet d'étude de choix pouvant, dans l'avenir, devenir une cible thérapeutique potentielle.

Notre laboratoire a participé par ses travaux à une partie des découvertes sur le Médiateur. En effet, nous avons caractérisé, par un crible double-hybride, les différentes interactions qui peuvent exister au sein du complexe et avec d'autres protéines nucléaires. Ce crible a permis de mettre en évidence aussi bien les contacts intrinsèques au complexe que des contacts avec d'autres complexes nucléaires. Deux de ces interactions ont été étudiées de façon fonctionnelle à l'aide de mutants thermosensibles (ts) de sous-unités essentielles. Les travaux du laboratoire ont permis de montrer que le Médiateur a un rôle essentiel dans le recrutement de la machinerie de transcription mais également de montrer, pour la première fois, un lien fonctionnel opéré par le Médiateur entre la transcription et la réparation des dommages à l'ADN.

Le succès de ces études a incité le laboratoire à poursuivre la caractérisation des sous-unités essentielles du Médiateur. Tirant parti de cette expérience, nous avons donc généré des mutants thermosensibles de plusieurs sous-unités essentielles. Mon travail de thèse porte sur deux d'entre elles, Med7 et Med10, appartenant au module du milieu du Médiateur en vue de caractériser sa fonction qui était alors mal définie. Lors de mon arrivée au laboratoire, une collection de 15 mutants avait été obtenue pour Med7 et un mutant avait été obtenu pour Med10.

J'ai commencé par une caractérisation biochimique de complexes contenant des sous-unités mutantes et démontré, par co-immunoprécipitation, que les mutations n'affectaient pas la stabilité du Médiateur. J'ai aussi développé une méthode inédite au laboratoire de purification du Médiateur afin de l'analyser par spectrométrie de masse pour augmenter la résolution de l'analyse de la stabilité et de l'assemblage du complexe. J'ai ensuite étudié, par immunoprécipitation de la chromatine (ChIP), la localisation des facteurs de transcription sur l'ADN sur des gènes modèles puis j'ai étendu l'analyse à l'échelle du génome par l'utilisation de la méthode du ChIP-sequencing (ChIP-seq). Cette étude a permis de montrer pour la première fois un rôle important du module du milieu du Médiateur sur la localisation génomique de l'ensemble des facteurs généraux de la transcription. L'analyse détaillée des données génomiques a permis également de montrer que le Médiateur pouvait influencer de façon différentielle l'assemblage du complexe de pré-initiation de la transcription (PIC) en fonction de l'architecture des promoteurs. Enfin, nous avons mis en évidence un contact fonctionnel entre le Médiateur et le facteur général de la transcription TFIIB, déterminant pour son recrutement au sein du PIC.

Ce manuscrit présentera l'introduction de l'état des connaissances au sujet de la transcription et du Médiateur. Les résultats de cette étude seront ensuite présentés sous forme d'une publication scientifique sur le rôle de la sous-unité Med10. Les résultats sur le rôle de Med7 seront ensuite présentés. Enfin, une partie discutera et mettra en perspective les résultats puis une dernière partie détaillera les méthodes et analyses mises en œuvres dans ce projet.

Introduction

Chapitre I : La transcription des gènes chez les eucaryotes

L'expression des gènes est un processus fondamental du vivant. En effet, l'état physiologique d'une cellule ou d'un organisme dépend des gènes qui sont exprimés à un instant t . Afin de répondre aux contraintes de son environnement ou aux signaux physiologiques, les cellules doivent pouvoir moduler l'expression de leurs gènes. De plus, l'expression des gènes est un processus central dans la différenciation et la spécialisation des cellules. C'est un processus très dynamique et très hautement régulé. La transcription est la première étape de l'expression des gènes. Il s'agit d'un mécanisme très complexe faisant intervenir, chez les eucaryotes, un très grand nombre de protéines et donc possédant différents niveaux de régulation. Au centre de la fonction transcriptionnelle on retrouve les ARN polymérases (Pol), enzymes utilisant l'ADN comme substrat pour synthétiser les ARNs. Chez les eucaryotes, il existe 5 ARN polymérases qui transcrivent chacune une classe de gènes différente. Les ARN polymérase I et III transcrivent les ARNs de transfert (ARNt) et ribosomiaux (ARNr), nécessaires à la traduction des protéines ainsi que des ARNs impliqués dans l'épissage des ARNs messagers (Werner, Thuriaux, and Soutourina 2009). Deux autres ARN polymérases, les ARN polymérases IV et V, ont été mises en évidence chez les plantes mais ne sont pas retrouvées chez les autres eucaryotes. Elles sont dispensables et leur rôle semble être limité à la transcription des petits ARNs impliqués dans le phénomène de *silencing* des gènes (Haag and Pikaard 2011). Pour finir, l'ARN polymérase II (Pol II) assure la transcription des ARNs messagers (ARNm) nécessaires à la synthèse des protéines et un grand nombre d'ARNs non-codants (ARNnc) dont la plupart ont des rôles régulateurs. Dans cette étude, nous nous intéressons uniquement à la transcription des gènes de classe II réalisée par la Pol II.

Le cycle de transcription (Figure 1), comporte trois étapes, chacune pouvant être régulée. La première étape est l'initiation. Elle consiste en le recrutement d'un activateur, ou facteur de transcription, sur une séquence d'ADN spécifique. Cet activateur va permettre, à l'aide de co-régulateurs, le recrutement d'un ensemble de facteurs nommés facteur généraux de la transcription (GTF) et de la Pol II. Ce recrutement prend place sur une séquence spécifique appelée promoteur afin de former le complexe de pré-initiation (PIC). L'élongation constitue la deuxième étape du cycle de la transcription, pendant laquelle la Pol II parcourt le gène transcrit en synthétisant l'ARN. Enfin la troisième étape est la terminaison qui aboutit à la séparation de la Pol II de l'ADN et la libération de l'ARN *néo*-synthétisé. Ce découpage en étapes est artificiel car elles sont en réalité très liées et ne sont pas nécessairement indépendantes.

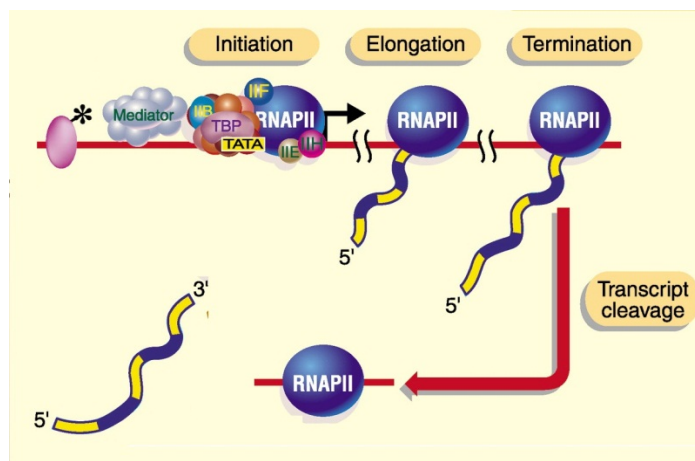


Figure 1: Déroulement d'un cycle de transcription. Schéma simplifié d'un cycle de transcription par la Pol II. (Adapté depuis Orphanides and Reinberg 2002).

L'initiation correspond à l'assemblage du PIC, l'élongation à la synthèse de l'ARN et la terminaison à la libération de l'ARN et de la Pol II.

Cette introduction décrira la transcription par la Pol II et sa machinerie. Un grand nombre d'études parmi celles présentées ont été réalisées chez la levure *S. cerevisiae*. Il s'agit d'un modèle de choix pour l'étude de la transcription chez les eucaryotes puisque les processus de la transcription sont très conservés au cours de l'évolution. La petite taille de son génome (environ 12 Mb) et le fait que celui-ci soit entièrement séquencé depuis 1996 et donc très bien annoté (Goffeau et al. 1996) en font un modèle très puissant pour les expériences de génomique fonctionnelle. De plus, *S. cerevisiae* est très facile à manipuler en laboratoire et permet l'obtention de mutants thermosensibles de gènes essentiels grâce à la grande facilité de modification de son génome, par recombinaison, et de la disponibilité d'outils tels que les vecteurs navettes et les marqueurs d'auxotrophie. Enfin, il est possible d'obtenir rapidement de grandes quantités de cellules ce qui est très utile pour réaliser des expériences de biochimie.

I. L'ARN polymérase II.

Au cœur du dogme central de la biologie moléculaire se trouvent la transmission et l'expression de l'information génétique. Celle-ci se fait par le biais de l'ARN. La synthèse des ARNs est accomplie lors de la transcription par les ARN polymérases. La découverte de cette activité enzymatique remonte à 1959 par la mise en évidence de la synthèse d'ARN utilisant une matrice d'ADN (Weiss and Gladstone 1959). Les études biochimiques ultérieures concernant les enzymes responsables de cette activité chez les mammifères ont permis d'isoler trois ARN polymérases (Roeder and Rutter 1969). Leur nom a été donné en fonction de l'ordre de sortie de la colonne de chromatographie. Elles ont alors été appelées ARN polymérase (Pol) A, B et C (aujourd'hui I, II et III, respectivement). Il aura néanmoins fallu attendre les études menées par la suite par les laboratoires de P. Chambon et R. Roeder pour connaître la spécificité enzymatique de ces trois enzymes. Ces études basées sur la sensibilité inhibitrice sur ces enzymes d'un antibiotique, l' α -amanitine, ont permis de montrer leur spécificité (Gniazdowski et al. 1970; Kedingner et al. 1970; Weinmann and Roeder 1974). Ainsi, il a été montré que l'ARN polymérase II, sensible aux faibles concentrations d' α -amanitine, est l'enzyme synthétisant les ARNm. Pol I et Pol III ont été associées à la transcription des ARNs ribosomiques et de transferts. La Pol II est un complexe enzymatique de 0.5 MDa comprenant 12 sous-

unités chez la levure et l'Homme. Les 12 sous-unités sont nommées Rpb1-12, selon l'ordre décroissant de leur masse moléculaire (Young 1991).

1) Composition de l'ARN polymérase II et conservation chez les eucaryotes.

a) Conservation des ARN polymérases.

Dans le monde du vivant, la transmission de l'information génétique passe par l'ARN. Il apparaît dès lors que la conservation du mécanisme de la transcription au cours de l'évolution implique une conservation des enzymes accomplissant cette fonction. Ainsi, dans les trois domaines du vivant que sont les bactéries, les archées et les eucaryotes, il existe des ARN polymérases. Contrairement aux eucaryotes, les archées et les bactéries ne possèdent qu'une seule ARN polymérase pour la transcription de l'ensemble de leurs gènes. Il semble que l'ancêtre commun aux bactéries, archées et eucaryotes possédait une ARN polymérase très proche de celle des bactéries actuelles à savoir une enzyme composée de 5 sous-unités α , α II, β , β' et ω . Au cours de l'évolution des sous-unités supplémentaires sont apparues ainsi que des protéines partenaires appelées facteurs de transcription, apportant un niveau de complexité supplémentaire et permettant une régulation plus complexe (Werner and Grohmann 2011). Cependant, ces enzymes présentent toutes un cœur structural commun. Ce cœur structural est à la base de l'activité enzymatique et est responsable des mécanismes moléculaires très proches au sein de toutes ces ARN polymérases. Ce socle commun est composé d'un cœur catalytique, unique composant de l'ARN polymérase bactérienne (Tableau 1).

<i>Bactéries</i>	<i>Archées</i>	<i>Eucaryotes</i>				
		Pol II	Pol III	Pol I	Pol IV	Pol V
β'	Rpo1 (RpoA)	Rpb1	Rpc160	Rpa190	NRPD1	NRPE1
β	Rpo2 (RpoB)	Rpb2	Rpc128	Rpa135	NRPD/E2	NRPD/E2
αI	Rpo3 (RpoD)	Rpb3	Rpc40	Rpc40	RPB3	RPB3
αII	Rpo11 (RpoL)	Rpb11	Rpc19	Rpc19	RPB11	RPB11
ω	Rpo6 (RpoK)	Rpb6	Rpb6	Rpb6	RPB6	RPB6
	Rpo5 (RpoH)	Rpb5	Rpb5	Rpb5	RPB5	NRPE5
	Rpo8 (RpoG)	Rpb8	Rpb8	Rpb8	RPB8	RPB8
	Rpo10 (RpoN)	Rpb10	Rpb10	Rpb10	RPB10	RPB10
	Rpo12 (RpoP)	Rpb12	Rpb12	Rpb12	RPB12	RPB12
	Rpo4 (RpoF)	Rpb4	Rpc17	Rpa14	NRPD/E4	NRPD/E4
	Rpo7 (RpoE)	Rpb7	Rpc25	Rpa43	NRPD7	NRPE7
		Rpb9	Rpc11	Rpa12	NRPD9b	
	Rpo13					
			Rpc31	Rpa34		
			Rpc34	Rpa49		
			Rpc37			
			Rpc53			
			Rpc82			

Tableau 1: Composition en sous-unités des ARN polymérases bactériennes, archées et eucaryotes (d'après Werner and Grohmann 2011).

Les ARN polymérases d'archées et les trois ARN polymérases eucaryotes contiennent des homologues des sous-unités de la polymérase bactérienne, composant leur cœur catalytique. Ainsi, les sous-unités Rpb1, Rpb2, Rpb3 et Rpb6 et leurs homologues chez les archées (Rpo1, 2, 3 et 6) comportent des similarités de séquences avec les sous-unités β' , β , α , et ω (Tan et al. 2000; Minakhin et al. 2001). La sous-unité Rpb11 possède également un domaine de similitude avec α (Ulmasov, Larkin, and Guilfoyle 1996).

En plus de cette homologie de séquence, il apparaît que les sous-unités Rpb1 et Rpb2 ont des fonctions similaires à celles de β' et β , toutes deux responsables de la plus grande part de l'activité catalytique de l'enzyme. Rpb1 tout comme β' a la capacité de se lier à l'ADN alors que Rpb2 lie le substrat nucléotidique de la même manière que β (Hampsey 1998). Enfin, Rpb3 présente aussi des similitudes fonctionnelles avec la sous-unité α bactérienne. En effet des mutants de Rpb3 ou α aboutissent à des défauts d'assemblage des deux ARN polymérases (Hampsey 1998).

Bien qu'il existe des similitudes, la conservation des séquences primaires des sous-unités des ARN polymérase des trois règnes est limitée. Il semblerait que la conservation fonctionnelle soit plutôt le reflet d'une conservation structurale (Figure 2).

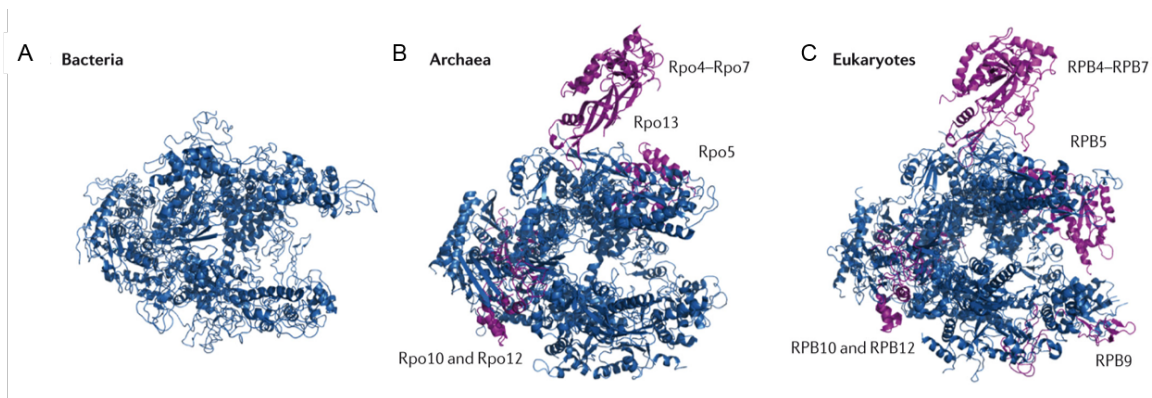


Figure 2 : Comparaison des structures des ARN polymérases bactérienne (A), d'archée (B) et de la Pol II eucaryote (C).

Les sous-unités en bleu sont les sous-unités universelles conservées entre les 3 règnes. Les sous-unités en magenta sont les sous-unités uniques aux polymérases d'archées et eucaryotes. (Werner and Grohmann 2011)

Ainsi, les deux plus grandes sous-unités de chacune de ces ARN polymérases (β , β' , Rpo1 et Rpo2, et Rpb1 et Rpb2) contiennent toutes un motif structural appelé tonneau- β -double-psi formant le site actif de ces enzymes (Werner and Grohmann 2011).

Toutes ces ARN polymérases contiennent une plateforme d'assemblage (Werner and Weinzierl, 2002). Chez les archées et les eucaryotes, cette plateforme est composée d'un complexe stable formé des deux hétérodimères Rpb10-Rpb12 (Rpo10-Rpo12 chez les archées) et Rpb3-Rpb11 (Rpo3-Rpo11 chez les archées). Le sous-complexe Rpb3-Rpb11 semble d'ailleurs être homologue à celui formé par les sous-unités α chez les bactéries. Les sous-unités supplémentaires présentes chez les eucaryotes et archées ont permis l'apparition de nouveaux éléments structuraux. On retrouve, par exemple, un complexe Rpb4-7 (Rpo4-7) appelé domaine tige, ou *stalk*, situé en périphérie de l'enzyme et capable de se dissocier du complexe principal (Grohmann and Werner 2011). Les deux sous-unités Rpb5 et Rpb8 (Rpo5 et Rpo8) sont quant à elles impliquées dans la séparation des brins de l'ADN transcrit et dans l'initiation précoce (Grünberg et al. 2010; Kang et al. 2006).

Outre ces similitudes, il existe des sous-unités spécifiques à chaque règne du vivant. En effet, la sous-unité Rpo13 des archées n'est retrouvée que chez ces dernières et n'a pas d'équivalent chez les eucaryotes. Sa fonction reste à l'heure actuelle obscure mais elle pourrait avoir un rôle dans l'initiation de la transcription similaire à celui de certains facteurs généraux de la transcription eucaryotes (Korkhin et al. 2009). La sous-unité Rpb9 et ses paralogues des Pol I et Pol III (Rpc11 et Rpc12, respectivement) sont exclusives aux eucaryotes. Elle est impliquée dans le contact entre la

polymérase et le GTF TFIIF, nécessaire à la détermination du site d'initiation de la transcription (Ziegler et al. 2003).

b) Composition des ARN polymérases eucaryotes.

Les ARN polymérases I, II et III sont bien conservées chez l'ensemble des eucaryotes. En effet, dans le cas de la Pol II, des études de complémentations ont permis de montrer que l'expression de sous-unités humaines chez la levure pouvaient se substituer à leur équivalent soit partiellement (Rpb4, 7 et 9) soit totalement (Rpb6, 8, 10, 12) (McKune et al. 1995; Khazak et al. 1998; Shpakovski et al. 1995).

Les trois ARN polymérases eucaryotes ont une composition semblable de 12 sous-unités cœurs qui sont soit homologues, soit identiques (Tableau 1). Notamment, les trois enzymes partagent 5 sous-unités : Rpb5, 6, 8, 10 et 12 (Werner, Thuriaux, and Soutourina 2009). Le domaine tige, composé de Rpb4-7 pour la Pol II, est présent dans les ARN Pol I et III mais avec une conservation de séquence réduite. Sa fonction régulatrice en demeure néanmoins similaire (Peyroche et al. 2002; Siat et al. 2003; Sadhale and Woychik 1994).

Les Pol I et III possèdent des sous-unités supplémentaires par rapport à la Pol II. Ainsi les deux sous-unités Rpc37 et Rpc53 appartenant à la Pol III et leurs paralogues Rpa34 et Rpa49 retrouvés dans la Pol I n'ont aucun équivalent dans la Pol II. Ces protéines possèdent des similitudes avec TFIIF, l'un des facteurs généraux de la transcription de la Pol II (voir Chapitre I.II.1) (Kuhn et al. 2007). De plus, la Pol III possède trois sous-unités spécifiques (Rpc31, Rpc34 et Rpc82) qui sont impliquées dans l'initiation de la transcription (Wang and Roeder 1997). Les deux sous-unités Rpc34 et Rpc82 de la Pol III forment un sous-complexe qui comporte une homologie avec TFIIE, un autre GTF de la Pol II. Il est intéressant de noter que l'interaction entre la Pol II et TFIIE se situe topologiquement dans la même région que celle entre Rpc34 et Rpc82 dans le cas de la Pol III (Carter and Drouin 2010; Wu et al. 2012). Cela suggère un mécanisme évolutif au cours duquel l'incorporation des sous-unités additionnelles de la Pol III proviendrait de l'utilisation permanente des GTFs TFIIE et TFIIF.

A l'heure actuelle, la structure atomique complète des ARN polymérases humaines n'est pas résolue à l'exception de celle du sous-complexe Rpb4-7. La conservation de séquence de la Pol II entre la levure *S. cerevisiae* et l'Homme est de 53%. Cela suggère donc encore ici une conservation structurale. Très récemment le laboratoire de P. Cramer a résolu la structure de l'ARN polymérase II bovine (Bernecky et al. 2016). Cette Pol II provenant d'un mammifère montre de très grandes similitudes avec celle de la levure *S. cerevisiae*. Les structures atomiques de la Pol II de *S. cerevisiae* et de *Schizosaccharomyces pombe* sont disponibles. Il est à noter aussi que les mécanismes d'initiation de la transcription chez l'Homme et chez *S. pombe* sont très similaires, notamment dans la structure de leur promoteurs (Spåhr et al. 2009). L'intérêt de l'étude des mécanismes de transcription chez *S. cerevisiae* et *S. pombe* est alors très grand pour la compréhension de la transcription chez l'Homme.

2) Structure de l'ARN polymérase II et fonction de ses sous-unités.

a) Structure tridimensionnelle de la Pol II.

La structure moléculaire à l'échelle atomique de l'ARN polymérase II a été résolue dans les années 2000 par le laboratoire de R. Kornberg, prix Nobel de chimie pour ses travaux sur les mécanismes de la transcription eucaryote. La première structure obtenue (Cramer et al. 2000) ne faisait apparaître que 10 sous-unités, Rpb4 et Rpb7 étant absentes. Ceci est expliqué par le fait que ce dimère est trouvé en quantité *sub*-stoechiométrique et qu'il peut se dissocier du « cœur » que composent les 10 autres sous-unités. Ce travail a néanmoins permis la première visualisation de la Pol II et a été le support de toutes les autres études structurales de la Pol II qui ont suivies (Cramer, Bushnell, and Kornberg 2001; Armache, Kettenberger, and Cramer 2003; Armache et al. 2005; Bushnell and Kornberg 2003; Bushnell et al. 2004). Toutes ces données structurales ont permis une avancée majeure dans la compréhension des mécanismes de la transcription.

Les structures de la Pol II ont permis de montrer que celle-ci est composée d'une masse centrale formée par les deux plus grandes sous-unités (Rpb1 et Rpb2) (Cramer 2004; Cramer et al. 2000) (Figure 3). Cette masse forme un sillon (*cleft*), chargé positivement, par lequel l'ADN s'introduit. Au fond de ce sillon se trouve le site actif de l'enzyme marqué par la présence d'un ion magnésium. L'ADN entrant passe par une structure formée de deux éléments mobiles appelés mâchoires (*jaws*). Dans le cœur du sillon se trouve une structure appelée mur (*wall*) qui bloque la progression de l'ADN et de l'ARN les forçant à une courbure de près de 90°. Cette courbure oriente le duplex ADN-ARN vers un pore en forme d'entonnoir inversé (*funnel*). L'ARN *néo*-synthétisé s'échappe de l'enzyme par ce pore. Sur l'un des côtés du sillon on retrouve une structure mobile en forme de pince (*stalk*) qui se trouve en position ouverte en absence d'ADN. Cette pince se referme en allant jusqu'au mur en présence du duplex ADN-ARN afin de le maintenir en place. Ce changement de conformation permet le passage de l'état inactif (initiation) à l'état actif (élongation). L'analyse de la distribution des charges électrostatiques de surface sur la Pol II révèle une charge globalement négative à l'exception des structures interne (sillon, site actif, mur et pince). Cela est consistant avec le fait que ces structures internes forment le « chemin » emprunté par l'ADN et l'ARN, globalement chargés négativement (Cramer 2004; Cramer et al. 2000)

Plus tard, l'obtention de la structure complète de la Pol II (Bushnell and Kornberg 2003) a permis de situer le dimère Rpb4-Rpb7 à la base de la pince, à proximité de la cavité d'où sort l'ARN transcrit.

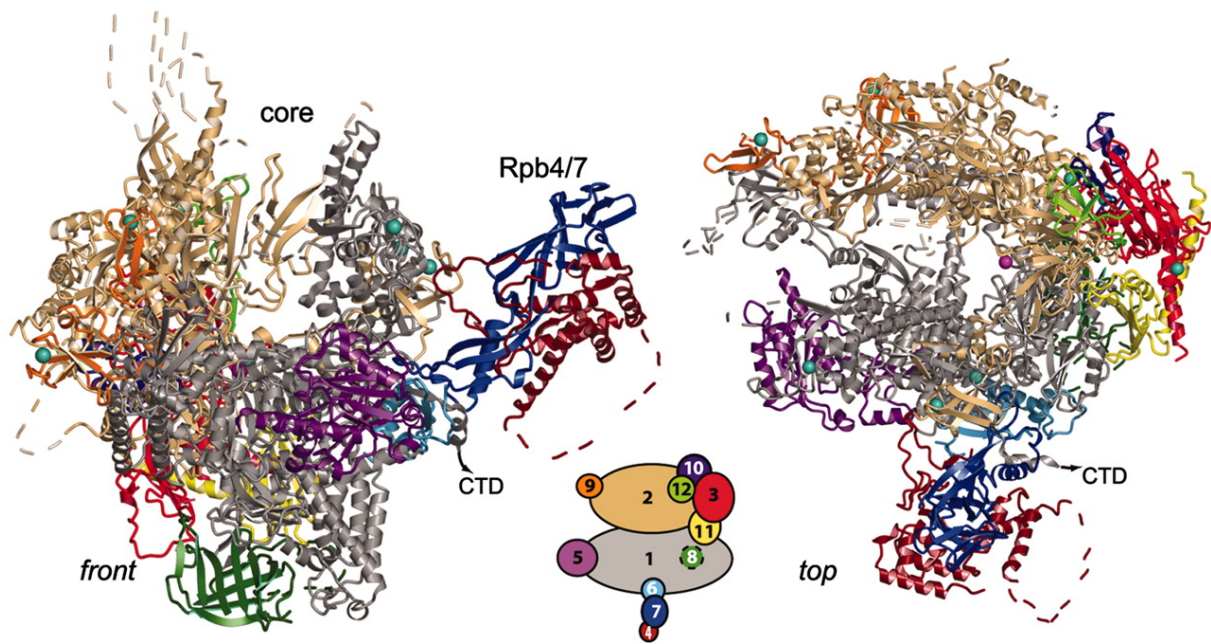


Figure 3 : Exemple de structure atomique de l'ARN polymérase II de *S. cerevisiae* obtenue par diffraction des rayons (D'après Armache et al. 2005).

La structure est visible selon deux orientations : vu de face (à gauche) ou vue de dessus (à droite). Les sous-unités sont représentées selon les couleurs indiquées sur le schéma au centre. La localisation du domaine carboxy-terminal (CTD) ainsi que celle des 8 ions zinc (sphères cyan) et de l'ion magnésium du site actif (sphère rose) est également indiquée.

b) Fonction des différentes sous-unités de la Pol II

Bien que la visualisation de la structure soit primordiale pour comprendre le fonctionnement global de la Pol II, le rôle de chaque sous-unité n'est pas totalement éclairé par ces données. Un certain nombre d'études biochimiques et génétiques ont été menées pour élucider ces questions (Tableau 2).

En plus de leur rôle primordial dans l'activité catalytique de l'enzyme, les sous-unités Rpb1 et Rpb2 ont un rôle dans la détermination du site de démarrage de la transcription (Berroteran, Ware, and Hampsey 1994). En effet, il a été montré que des mutations de ces sous-unités provoquent des défauts dans la détermination du site d'initiation de la transcription (*transcription start site*, TSS). Ce rôle dans le choix du site de démarrage de la transcription est probablement en partie dû à l'interaction de Rpb1 avec TFIIB, facteur général de la transcription impliqué dans le choix du TSS. De plus, un lien fonctionnel et une interaction directe entre Rpb1 et le facteur d'élongation TFIIIS ont été montrés. Cela confère à Rpb1 un rôle dans l'élongation de la transcription (Archambault, Drebot, et al. 1992; Archambault, Lacroute, et al. 1992).

La sous-unité Rpb3, quant à elle, possède un rôle similaire à celui des sous-unités α bactériennes à savoir un rôle dans l'assemblage de l'enzyme par son effet stabilisateur. Elle joue également un rôle dans la transcription activée (Tan et al. 2000; Hampsey 1998). De plus, nous avons montré au laboratoire que Rpb3 interagit avec le complexe Médiateur, co-activateur de la transcription, et que cette interaction est nécessaire à la transcription de l'ensemble du génome (Soutourina et al. 2011) (voir chapitre II).

<i>Nom de la sous-unité</i>	<i>Masse (kDa)</i>	<i>Fonctions</i>
<i>Rpb1</i>	220	Rôle dans la sélection du TSS Rôle dans l'élongation Rôle dans l'assemblage de la Pol II Porte le CTD et le site actif de l'enzyme Interagit avec TFIIIB et TFIIIS
<i>Rpb2</i>	140	Rôle dans la sélection du TSS Rôle dans l'élongation Rôle dans l'assemblage de la Pol II
<i>Rpb3</i>	45	Analogue de la sous-unité α chez <i>E. coli</i> Rôle dans l'assemblage de la Pol II Impliquée dans la réponse aux activateurs Interagit avec le Médiateur
<i>Rpb4</i>	25	Impliquée dans l'initiation de la transcription Rôle dans l'élongation Rôle dans la maturation des ARN Rôle dans le métabolisme des ARN Rôle dans l'activation de la transcription Rôle dans la réparation de l'ADN
<i>Rpb5</i>	25	Rôle dans l'activation de la transcription Interagit avec TFIIIB
<i>Rpb6</i>	18	Impliquée dans l'élongation de la transcription. Maintien du sous-complexe Rpb4-7
<i>Rpb7</i>	19	Impliquée dans l'initiation de la transcription. Rôle dans l'élongation Rôle dans la maturation des ARN Rôle dans le métabolisme des ARN
<i>Rpb8</i>	16	Rôle dans l'assemblage de la Pol II
<i>Rpb9</i>	14	Impliquée dans la sélection du site d'initiation Contrôle la fidélité de la transcription Interagit avec TFIIIE et TFIIIF
<i>Rpb10</i>	8	Consolide l'architecture de la Pol II
<i>Rpb11</i>	14	Rôle dans l'assemblage de la Pol II
<i>Rpb12</i>	8	Consolide l'architecture de la Pol II

Tableau 2 : Récapitulatif des fonctions des sous-unités de la Pol II.

Les sous-unités essentielles sont indiquées en rouge.

La structure de la Pol II a révélé que celle-ci s'organise en deux modules : un cœur catalytique composé de 10 sous-unités et un sous-complexe Rpb4-7 pouvant se dissocier du cœur. Ce sous-module présente un rôle très important dans la fonction de la Pol II. Les sous-unités Rpb4 et Rpb7 ont un rôle fonctionnel commun puisque la purification de la Pol II dans une souche de levure $\Delta rpb4$ aboutie à une enzyme dépourvue de la sous-unité Rpb7 (Edwards et al. 1991). La caractérisation de cette souche a permis de montrer que, *in vitro*, cet hétérodimère n'était pas requis pour l'élongation de la transcription mais indispensable pour l'initiation. L'absence de Rpb4-7 n'empêche cependant pas le recrutement de la Pol II au promoteur. Cela suggère donc que la nécessité de ce sous-module dans l'initiation se trouve dans des étapes *post* recrutement de la Pol II. En ce sens, il a été montré que Rpb4-7 possède des domaines de liaison à l'ADN et à l'ARN. L'interaction entre ces sous-unités et l'ADN permettrait donc de stabiliser la Pol II sur le promoteur, une fois celle-ci recrutée (Orlicky et al. 2001).

Ces études laissent entendre que ce module serait lié à la Pol II durant l'initiation et qu'il se dissocierait lors de l'élongation.

Il semble cependant qu'il en soit autrement *in vivo*. En effet, la localisation génomique de la Pol II a montré que l'on peut détecter la présence de Rpb4-7 non seulement sur les régions promotrices, comme suggéré par les données *in vitro*, mais aussi sur les régions transcrites (Jasiak et al. 2008; Verma-Gaur et al. 2008). Le niveau d'occupation de ces deux sous-unités et leur profil sont très similaires à celui de Rpb3. Cela suggère donc que ce module accompagne le cœur de la Pol II durant l'élongation. De plus, une interaction entre Rpb7 et Nrd1 a également été mise en évidence (Mitsuzawa, Kanda, and Ishihama 2003). Nrd1 est une protéine qui fait partie de la machinerie de maturation des ARNs, cela indique un rôle probable de Rpb7 dans ce processus.

Des études ont proposé un rôle cytoplasmique de Rpb4-7, notamment dans l'export et le métabolisme des ARNs, en particulier leur dégradation (Lotan et al. 2005; Lotan et al. 2007; Farago et al. 2003). Ces études sont néanmoins controversées. En effet, le laboratoire de P. Cramer suggère que la diminution de la dégradation de ARNs provoquée par la délétion de *RPB4* est plutôt due à un effet tampon qui vise à allonger les temps de demi-vie des ARNs pour compenser le défaut transcriptionnel causé par l'absence de Rpb4 (Schulz et al. 2014).

Comme expliqué plus tôt, la résolution de la structure complète de la Pol II (Bushnell and Kornberg 2003) et les analyses de cette structure en solution par microscopie électronique (Craighead, Chang, and Asturias 2002) ont permis de situer le dimère Rpb4-7 à la base de la pince. Ce positionnement permet d'imaginer un rôle de ce dimère dans la régulation de cette pince mobile. En effet, les modèles de structures montrent que si la pince est en position ouverte, il y a un encombrement stérique important entre celle-ci et Rpb7 provoquant une collision. Dans cette configuration, Rpb4-7 ne peut se lier à la Pol II que lorsque la pince est fermée et inversement, la pince ne peut s'ouvrir que lorsque Rpb4-7 est détaché du cœur de la Pol II (Armache, Kettenberger, and Cramer 2003). Il existe alors deux possibilités pour l'initiation de la transcription. La première est la dissociation complète de Rpb4-7 qui laisse la pince ouverte et qui autorise l'ADN à pénétrer dans le sillon et jusqu'au site actif. Rpb4-7 se lie ensuite une fois la pince fermée. Cette possibilité est à nuancer par le fait que la dissociation du complexe Rpb4-7 n'a jamais été observée *in vivo*. La deuxième possibilité consiste à imaginer que le module Rpb4-7 est en permanence lié à la Pol II. La pince est alors fermée et empêche l'ADN d'entrer. L'entrée de la matrice d'ADN se ferait alors après dénaturation où le brin transcrit, assez fin pour passer la pince, peut se glisser jusqu'au site actif. Cette deuxième possibilité semble plutôt en accord avec les conformations observées en solution puisque jusqu'à maintenant, aucune conformation ou position de la pince laissant la possibilité à de l'ADN double brin de rentrer n'a été observée (Craighead, Chang, and Asturias 2002). C'est également le modèle retenu par le laboratoire de P. Cramer (Cheung and Cramer 2012).

Les sous-unités Rpb9 et Rpb4, sont toutes deux non essentielles pour la croissance en milieu standard. Cependant, la modification des conditions de croissances (température, changement de source de carbone...) les rendent indispensables pour la croissance (Hampsey 1998).

Concernant Rpb9, la délétion de cette sous-unité mène à une diminution de la fidélité de la transcription *in vitro* (Walmacq et al. 2009). Ce phénotype est dû à une augmentation des *mis-incorporations* des dNTP au cours de l'élongation. Deux explications sont avancées pour expliquer ce défaut : l'une au niveau de l'incorporation du nucléotide, l'autre au niveau de la correction d'épreuve (*proofreading*). Il semblerait que Rpb9 participe à la régulation de l'entrée des nucléotides au sein du site actif de l'enzyme, régulant ainsi l'activité catalytique de l'enzyme. L'autre explication réside dans un mécanisme lié à TFIIIS. En effet, Rpb9 interagit avec TFIIIS, facteur d'élongation ayant une activité de correction au cours de l'élongation de la transcription.

Des mutants de Rpb9 aboutissent à des défauts de définition du TSS (Hekmatpanah and Young 1991; Hull, McKune, and Woychik 1995). Une interaction entre Rpb9 et le GTF TFIIIF, nécessaire à l'assemblage des PICs et au choix du TSS, a été suggérée. En ce sens, l'interaction entre la Pol II et TFIIIF est abolie *in vitro* en l'absence de Rpb9 (Ziegler et al. 2003). Des expériences menées par le laboratoire de S. Hahn ont permis de confirmer le contact direct entre Rpb9 et TFIIIF (Eichner et al. 2010). De plus, Rpb9 interagit avec Tfa1, une des sous-unités du facteur général de la transcription TFIIIE (van Mullem

et al. 2002). Ce GTF est requis pour l'assemblage des PICs et au recrutement de TFIIH, autre GTF nécessaire au passage à l'étape d'élongation. Enfin, un rôle de Rpb9 a été suggéré dans la réparation de l'ADN par excision de nucléotide couplée à la transcription (Li et al., 2002).

Il est très intéressant de constater ici, que les sous-unités Rpb4 et Rpb9, pourtant indispensables, ont un certain nombre de rôles important dans la fonction de la Pol II. Les autres sous-unités dont le rôle n'a pas été présenté en détail ici ont pour la plupart un rôle de structuration de la Pol II ou dans son assemblage.

3) Le Domaine Carboxy-Terminal de l'ARN polymérase II.

L'ARN polymérase II possède une caractéristique unique qui la différencie des autres ARN polymérases. En effet, sa plus grande sous-unité, Rpb1, possède une structure très particulière à son extrémité carboxy-terminale. Ce domaine est appelé CTD (*Carboxy-Terminal Domain*) et constitue une véritable plateforme de signalisation et de régulation de la Pol II. Sa nature non structurée rend sa localisation exacte impossible sur la structure de la Pol II. Néanmoins, sa présence est suggérée dans une région située à proximité du pore d'où l'ARN sort de l'enzyme lui conférant un emplacement privilégié pour interagir avec de nombreux partenaires.

Le CTD est constitué d'un motif répété de 7 résidus : Tyr-Ser-Pro-Thr-Ser-Pro-Ser. Cet heptapeptide est très conservé chez les eucaryotes même si le nombre de répétitions varie selon les espèces. On retrouve, par exemple, 24 répétitions chez *S. cerevisiae* contre 52 chez l'Homme (Eick and Geyer 2013). Il est intéressant de noter que la complexité de l'organisme n'est pas corrélée avec le nombre de répétitions bien que les eucaryotes pluricellulaires ont tendance à posséder un CTD plus long (Eick and Geyer 2013). L'heptapeptide présenté plus haut est une séquence consensus puisque chez certaines espèces, le motif retrouvé est différent comme, par exemple, chez les amibes *Mastigamoeba* et les parasites *Plasmodium* (Stiller and Hall 2002). Chez d'autres organismes, comme les protistes parasites du genre *Trypanosoma* ou *Leishmania*, il n'existe pas de motifs identifiables au niveau du CTD (Eick and Geyer 2013; Stiller and Hall 2002).

Le CTD est un domaine essentiel à la viabilité cellulaire (Nonet, Sweetser, and Young 1987; Allison et al. 1988). Néanmoins, la délétion d'un nombre restreint de répétitions ne génère pas de phénotypes observables. Il faut avoir supprimé au moins la moitié ou plus des répétitions pour voir des effets sur la croissance chez la levure et les mammifères. Ainsi, le nombre minimum de répétitions nécessaire à la viabilité chez *S. cerevisiae* est de 11 (Thompson et al. 1993; Bartolomei et al. 1988). Au-delà de l'importance du nombre de répétitions, il est intéressant de noter que la nature des résidus qui composent ce motif est essentielle. En effet, des mutants de substitutions des sérines 2 et 5 en alanine ou glutamate sont létales chez la levure *S. cerevisiae* (West and Corden 1995).

Si l'on s'intéresse maintenant à la composition du motif, il est surprenant de noter sa relativement faible complexité puisqu'il n'est composé que de 4 résidus d'acides aminés différents. Néanmoins, leur nature physico-chimique est particulièrement intéressante. En effet, on retrouve dans ce motif deux résidus prolines. Ce résidu particulier, lorsqu'il se trouve dans une chaîne peptidique, peut se trouver dans deux conformations différentes (*cis* ou *trans*) conférant ainsi une conformation de la chaîne variable en générant des courbures. Il a été montré que dans le motif du CTD, on peut retrouver les deux conformations (Meinhart et al. 2005). L'action d'enzymes du type prolyl-isomérase permet de moduler la conformation de ces résidus. Il existe donc, étant donné le nombre de deux prolines et de deux conformations possibles, 4 combinaisons de conformations du motif. De plus, bien que la structure du CTD ne puisse pas être déterminée sur le modèle de la Pol II complète, des expériences de dichroïsme circulaire et de RMN ont permis de suggérer la présence de coudes β (β turns) dus à la présence des résidus prolines (Bienkiewicz et al. 2000; Meinhart et al. 2005).

L'autre caractéristique essentielle de la composition du motif de l'heptapeptide est la présence des trois résidus portant une fonction hydroxyle. Ces résidus forment un espace hydrophile fortement propice à des modifications *post*-traductionnelle par estérification telles que la phosphorylation. De la

même manière que pour la conformation des prolines, il y a là encore de nombreuses possibilités de combinaison puisque les résidus ne sont pas forcément modifiés en même temps. Cela laisse entrevoir un potentiel de régulation immense de telle façon que l'on peut parler de « code » du CTD (Buratowski 2003).

Ainsi, le CTD est phosphorylable au niveau des sérines 2, 5 et 7 et l'on peut retrouver le CTD dans un état phosphorylé ou non *in vivo*. On parle de forme IIA pour la forme de la Pol II non-phosphorylée et de forme IIO pour sa forme phosphorylée. (Eick and Geyer 2013; Hampsey 1998). Au moins 4 kinases catalysant ces réactions ont été identifiées. Chez *S. cerevisiae*, il s'agit de Kin28, Cdk8, Bur1 et Ctk1 (Eick and Geyer 2013). Ces quatre kinases appartiennent à la famille des CDK (*cyclin-dependent kinase*), c'est à dire qu'elles sont couplées et régulées par des cyclines. Kin28, couplée à la cycline 1 (Ccl1), fait partie du complexe TFIIH, facteur général de la transcription (Valay et al. 1996). Bur1, et sa cycline Bur2, font partie du complexe P-TEFb, facteur d'élongation de la transcription (Keogh, Podolny, and Buratowski 2003). Concomitamment, il existe des phosphatases capables d'enlever ces modifications afin de contrôler l'état de phosphorylation de la Pol II. La balance entre les deux états est dynamique et constitue un point de contrôle important dans la régulation de la transcription.

Des expériences d'immunoprécipitation de la chromatine (ChIP) ont permis l'analyse des profils de phosphorylation du CTD sur des gènes en cours de transcription (Figure 4). La Pol II est recrutée sous sa forme IIA, non phosphorylée, sur le promoteur. L'assemblage du PIC permet le recrutement de TFIIH et donc de Kin28 qui va phosphoryler le CTD au niveau de la Ser5 et de la Ser7 (Hengartner et al. 1998; Akhtar et al. 2009). C'est cette modification qui permet la transition entre l'initiation et l'élongation. La kinase Cdk8, appartenant au Médiateur, est également connue pour phosphoryler la Ser5 *in vitro* mais avec cette fois un rôle répresseur (Voir Chapitre II.IV) (Hengartner et al. 1998). Rapidement après le début de l'élongation, la phosphatase Ssu72 retire la phosphorylation de la Ser5 et le complexe P-TEFb *via* sa kinase Bur1, phosphoryle la Ser2 (Ni et al. 2004). La phosphorylation de la Ser7 est quant à elle persistante tout au long de la transcription (Tietjen et al. 2010). Il faut attendre que la Pol II parvienne au niveau du site de terminaison de la transcription (*transcription terminaison site*, TTS) pour que les Ser2 et Ser7 soient dephosphorylées par les phosphatases Fcp1 et Ssu72, respectivement (Majello and Napolitano 2001).

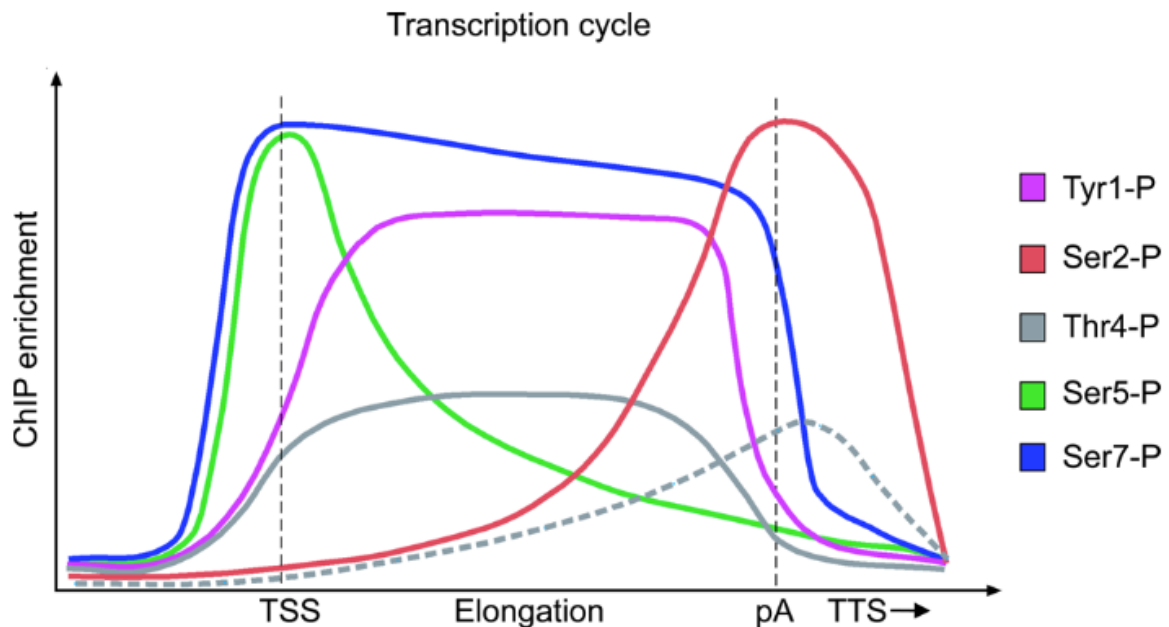


Figure 4: Répartition et évolution des modifications du CTD au cours de la transcription d'un gène visualisé par ChIP chez la levure. (D'après Eick and Geyer 2013b).

TSS : site de démarrage de la transcription, pA : site de poly-adénylation, TTS : site de terminaison de la transcription.

Ces modifications du CTD ont un effet non seulement sur la transcription à mais également sur d'autres processus couplés à celle-ci comme la maturation de l'ARN et des modifications de la chromatine. En effet, la phosphorylation différentielle des résidus va faire varier l'affinité de certaines protéines pour le CTD, permettant de moduler leur recrutement en fonction du cycle de transcription. Les protéines partenaires ont pour la plupart un domaine CID (*CTD-interacting domain*) permettant leur interaction avec le CTD (Meinhart et al. 2005). L'idée d'un couplage des événements de transcription et de maturation de l'ARN transcrit a longtemps été suggérée et est maintenant admise. Le CTD joue un rôle important dans ce couplage dont voici quelques exemples. L'ARN-guanilyltransférase (Cgt1) et l'ARN-méthyltransférase interagissent avec le CTD phosphorylé (McCracken et al. 1997; Cho et al. 1997). Ces deux enzymes font partie des protéines impliquées dans l'ajout de la coiffe sur l'ARN naissant. Elles ont une affinité pour la Ser5, et sont donc recrutées dès le début de la phase d'initiation. D'autres études ont montrées une interaction entre le CTD phosphorylé et des protéines impliquées dans l'épissage, en particulier les protéines de la famille SR (*Serine/Arginine-rich*) (Yuryev et al. 1996; Mortillaro et al. 1996). Le CTD possède aussi un rôle dans les événements couplés à la transcription tardive. Il a été montré que celui-ci pouvait interagir avec CPSF (*cleavage and polyadenylation specificity factor*) et CstF (*Cleavage stimulatory factor*), deux protéines impliquées dans l'ajout de la queue poly-adénylée et dans le clivage de l'ARN en 3' lors de la terminaison (McCracken et al. 1997). Enfin, les phosphorylations sur les Ser5 et Ser2 permettent le recrutement des histones méthyl-transférase Set1 et Set2. Ces enzymes vont permettre la modification des histones au cours de la transcription par ajout de groupements méthyls (Hampsey and Reinberg 2003). Ces modifications de la chromatine en cours d'élongation pourraient servir à « marquer » le gène transcrit afin de jouer un rôle dans la ré-initiation mais également à intervenir dans le recrutement des protéines impliquées dans les processus de maturation de l'ARN (Hampsey and Reinberg 2003).

Il est désormais bien établi que les phosphorylations des sérines ne sont pas les seules modifications *post*-traductionnelles du CTD. En effet, les tyrosines 1 (Tyr1) et les thréonine 4 (Thr4) peuvent elles aussi être phosphorylées (Eick and Geyer 2013). On retrouve ainsi un assez haut niveau de phosphorylation de Tyr1 et Thr4 au cours de l'élongation. Ces deux modifications auraient un rôle protecteur contre une terminaison prématurée en inhibant le recrutement des facteurs de terminaison (Hsin et al. 2011; Meinhart et al. 2005). Les Ser2 et Ser5 peuvent, en plus de la phosphorylation, être sujettes à des glycosylations. Ces résidus ne pouvant recevoir qu'une sorte de modification à la fois, la glycosylation est proposée comme étant une protection contre une phosphorylation de la Pol II avant son recrutement au promoteur (Ranuncolo et al. 2012). Enfin, le motif du CTD n'est pas nécessairement identique tout du long des répétitions. Sur l'extrémité C-terminale, on retrouve à la place du résidu Ser7 un résidu lysine ou arginine. Chez les mammifères il a été montré que ce résidu pouvait être méthylé (Sims et al. 2011). Cette méthylation aboutirait à une inhibition de la transcription de certains petits ARN nucléaires au rôle régulateur.

L'ensemble de ces modifications *post*-traductionnelles et la modulation de la conformation des résidus prolines du CTD sont donc d'une grande complexité et offrent de nombreuses possibilités de régulation de la transcription.

II. La transcription par l'ARN polymérase II.

1) Structure des gènes de classe II.

Les gènes transcrits par l'ARN polymérase II sont appelés gènes de classe II. Leur transcription commence par le recrutement des protéines de la machinerie de la transcription au niveau d'une

portion de l'ADN appelée promoteur. Un promoteur peut être considéré comme constitué de deux parties distinctes : une partie cœur (*core promoter*) et une partie régulatrice qui peut être distale ou proximale. Le promoteur cœur est l'élément minimal requis pour une initiation de la transcription *in vitro*. L'étude des séquences présentes dans les promoteurs cœurs a permis de mettre en évidence 7 éléments récurrents dans ces régions (Figure 5) (Thomas and Chiang 2006).

On retrouve ainsi l'élément appelé boîte-TATA (*TATA-box*), au côté du motif BRE (*TFIIB-recognition element*) et de la séquence Initiatrice appelée *Inr*. En aval du site de démarrage de la transcription se situe le DCE (*downstream core element*) lui-même composé du motif à 10 éléments MTE (*motif ten element*) et de l'élément promoteur aval DPE (*downstream promoter element*). Ces motifs vont être liés par des protéines spécifiques des GTFs TFIID et TFIIB.

Ces éléments ne sont pas systématiquement présents tous à la fois dans tous les promoteurs, permettant une modulation fine de la régulation de l'expression des gènes. Ces motifs correspondent à des séquences consensus. Chez la levure *S. cerevisiae*, à l'exception de la boîte-TATA, ces motifs n'existent pas sous cette forme consensus. Il existe cependant des séquences similaires.

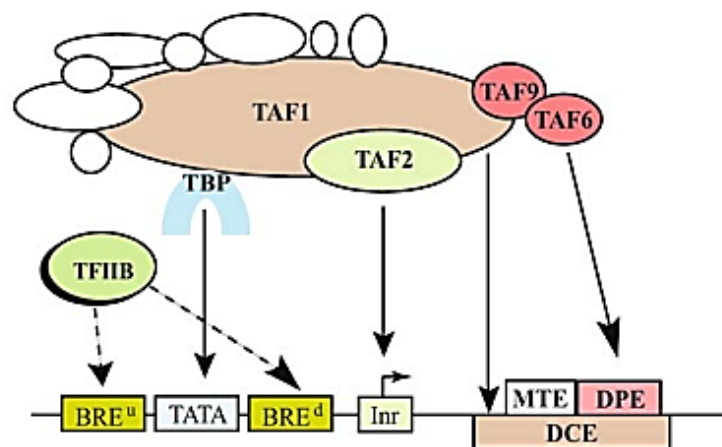


Figure 5 : Représentation schématique de la structure d'un promoteur cœur et des protéines interagissant avec les différents domaines. (Thomas and Chiang 2006)

a) La boîte TATA.

La boîte TATA est le premier élément du promoteur cœur qui a été identifié dans les gènes de classe II (Breathnach and Chambon 1981). L'étude d'un ensemble de régions promotrices de différents eucaryotes a permis la mise en évidence d'une séquence TATAAA en amont de la région codante, environ 25-30 pb avant le site de démarrage de la transcription (Smale and Kadonaga 2003). Des études de mutagenèse sur cette séquence ont montrées le rôle clé de ce motif pour l'initiation de la transcription *in vitro* (Wasylyk et al. 1980). Par la suite, il a été montré que cet élément était reconnu et lié par une protéine appelée TBP (*TATA-binding protein*). C'est le recrutement de la TBP qui initie l'assemblage du PIC *in vitro* (Klein and Struhl 1994; Smale and Kadonaga 2003).

La boîte TATA est située à environ 25-30 pb du site de démarrage de la transcription chez la majorité des eucaryotes. Chez la levure *S. cerevisiae*, elle est retrouvée 40 à 120 pb avant le TSS (Smale and Kadonaga 2003). Cela suggère donc des mécanismes différents dans la définition du TSS après le recrutement du PIC. Pour la majorité des eucaryotes, le PIC se forme au niveau de la boîte TATA et le TSS est défini par sa position 25 à 30 pb plus loin. Chez *S. cerevisiae*, il est proposé que le PIC soit là encore assemblé sur la boîte TATA mais que la Pol II se déplace ensuite afin de « scanner » l'ADN

jusqu'à un TSS (Giardina and Lis 1993; Kuehner and Brow 2006).

Malgré son importance dans l'assemblage du PIC, il est très frappant de noter que chez la levure, seuls environ 20% des gènes possèdent une boîte-TATA correspondant au consensus (Basehoar, Zanton, and Pugh 2004). On peut ainsi parler de promoteur *TATA-containing* et *TATA-less*. Cette situation se retrouve chez l'Homme puisqu'une boîte TATA canonique n'est présente que dans 10% des gènes (Gershenzon and Ioshikhes 2005). Néanmoins, la TBP est toujours recrutée sur les promoteurs, que la boîte TATA canonique soit présente ou absente (Kuras and Struhl 1999; Li et al. 1999).

Partant de cette observation, il a été montré qu'il existe deux voies d'assemblage des PICs, toutes deux nécessitant la présence de la TBP. L'une est dépendante de TFIID, et correspond aux promoteurs dépourvus de boîte TATA. Les gènes concernés sont dans l'ensemble des gènes constitutivement exprimés comme les gènes codants pour les protéines ribosomales (Mencía et al. 2002). L'autre voie est quant à elle dépendante de SAGA (Spt-Ada-Gcn5 acétyltransférase), un complexe co-activateur de la transcription contenant la TBP et ayant en plus une activité sur l'organisation de la chromatine (voir Chapitre I.V). Les gènes régulés par SAGA possèdent une boîte TATA canonique et codent pour des gènes de réponse au stress (Basehoar, Zanton, and Pugh 2004). Ces deux voies ne sont cependant pas exclusives. De plus, une étude récente a montré que SAGA était requis pour la transcription de l'ensemble du génome à la fois chez l'Homme et chez la levure (Bonnet et al. 2014). Cette séparation en deux voies d'assemblages des PICs, l'une TFIID-dépendante et l'autre SAGA-dépendante, doit être nuancée. En effet, la dépendance sur l'assemblage du PIC vis à vis de l'un ou l'autre de ces complexes semble plutôt liée à leur capacité à faciliter le recrutement de la TBP sur l'un ou l'autre des types de promoteurs. Dans le cas des gènes où TFIID est supposé être responsable du recrutement de la TBP, SAGA est quand même recruté et agit probablement par ses autres activités de co-activateurs (histone acétyl-transférase).

Récemment, le laboratoire de F. Pugh a analysé la distribution génomique à haute résolution des facteurs généraux de la transcription (Rhee and Pugh 2012). L'étude utilise la méthode de ChIP-exo, permettant l'analyse de la liaison d'une protéine à l'ADN à la résolution d'un nucléotide. Ces travaux ont permis une compréhension plus fine des mécanismes d'assemblage du PIC selon que le promoteur contienne ou non une boîte TATA canonique. Tout d'abord, l'analyse des sites de liaison de TFIID a permis de montrer que dans 99% des cas, ce GTF était lié à une séquence TATA dégénérée de deux ou un nucléotide par rapport à la boîte TATA consensus. Les auteurs proposent donc la notion d'éléments *TATA-like* en remplacement de *TATA-less* bien qu'on ne puisse pas définir de séquence consensus pour ces éléments. Ensuite, il apparaît que dans le cas des gènes dépourvus de boîte TATA canonique, le recrutement de TFIID se fait de manière coopérative entre l'élément *TATA-like* et le nucléosome +1 (le premier nucléosome situé après le TSS). Cela contraint la Pol II à scanner le promoteur à la recherche du TSS dans un espace restreint. Cette observation est cohérente avec le fait que les TSS des gènes sans boîtes TATA soient plus près de l'élément *TATA-like* que celui des gènes possédant une boîte TATA. Les gènes possédant une boîte TATA montrent eux un comportement différent. En effet, SAGA va permettre le recrutement de la TBP sur la boîte TATA et rentre en compétition avec le nucléosome +1. La présence de SAGA entraîne une déplétion de celui-ci levant un frein à la recherche du TSS pour le PIC puisque l'espace disponible n'est pas restreint. Cette étape étant limitante et la diminution de la présence du nucléosome +1 rendant l'initiation plus productive, ce mécanisme confère à ces gènes une grande flexibilité d'expression.

b) L'élément de reconnaissance de TFIIB (BRE).

L'élément de reconnaissance de TFIIB (*TFIIB recognition element*, BRE) n'est pas retrouvé chez la levure sous une forme canonique mais est présent chez les autres eucaryotes examinés. Il est intéressant de noter que c'est l'unique motif du promoteur cœur qui est reconnu par un autre GTF que TFIID. L'idée de la présence d'un tel motif vient de l'observation que sur les structures du complexe

TFIIB-TBP-TATA, TFIIB interagit avec l'ADN en amont et en aval du motif TATA (Lee and Hahn 1995). La séquence consensus du BRE est 5'-G/C-G/C-G/A-C-G-C-C-3' (Lagrange et al. 1998). Cet élément peut être retrouvé en amont ou en aval de la boîte TATA (Deng and Roberts 2005), on parle alors d'*upstream* BRE (BREu) et de *downstream* BRE (BREd). Lors de son identification, il a d'abord été associé à une activation de la transcription basale *in vitro* (Lagrange et al. 1998). Des études ultérieures ont montré un rôle négatif de BRE dans la transcription basale *in vivo* mais que cet effet est aboli lors de la présence d'un activateur (Evans, Fairley, and Roberts 2001). L'élément BRE possède une double fonction apportant une possibilité de régulation. Il faut noter que l'élément BRE est reconnu par un domaine particulier de TFIIB formé d'une structure Hélice-Tour-Hélice (HTH), absente chez la levure et les plantes, suggérant que TFIIB ne reconnaît pas de BRE chez ces organismes (Smale and Kadonaga 2003).

c) La séquence initiatrice (Inr).

L'analyse de certains promoteurs a révélé que le TSS est souvent composé d'une adénine au niveau du site de démarrage (+1) précédée d'une cytosine toutes deux entourée de quelques pyrimidines (Smale and Kadonaga 2003). Des mutations de ces régions provoquent des défauts dans la transcription mais aussi dans la définition du TSS (Grosschedl and Birnstiel 1980). L'idée de l'existence d'une séquence Inr définissant le TSS et a été démontrée. Elle est suffisante à l'initiation de la transcription *in vitro* et *in vivo* (Smale and Baltimore 1989). La séquence Inr proposée est Py-Py-A(+1)-N-T/A-Py-Py où Py signifie pyrimidine et N peut être n'importe quel nucléotide (Smale and Kadonaga 2003). TFIID est capable de lier l'Inr *via* ses sous-unités Taf1 et Taf2 (Kaufmann and Smale 1994). L'interaction de TFIID avec l'Inr peut se substituer à une boîte TATA ou bien agir de façon synergique dans le cas où le promoteur contient les deux éléments (Weis and Reinberg 1997). Pour finir, l'Inr est un motif spécifique aux promoteurs des eucaryotes unicellulaires et n'est par conséquent pas retrouvé chez *S. cerevisiae*, renforçant le fait que les TSS chez cette espèce n'est pas défini de façon précise.

d) L'élément promoteur aval (DPE).

L'élément promoteur aval (*downstream promoter element*, DPE) a été identifié comme un motif du promoteur nécessaire à la liaison de TFIID purifié sur certains promoteurs dépourvus de boîte TATA. Le DPE peut néanmoins exister dans les promoteurs des gènes *TATA-like*. Il se situe à +28-+32 nucléotides de l'adénine +1 de l'Inr et est composé de la séquence consensus A/G-G-A/T-C-G-T-G (Smale and Kadonaga 2003). Cet élément est par contre toujours accompagné d'un élément Inr. Des études ont montré que dans les promoteurs ne contenant que ces deux éléments, des mutations dans l'un ou l'autre, ou encore une modification de l'espacement entre eux, menait à une forte diminution du recrutement de TFIID et de la transcription (Burke and Kadonaga 1996; Kutach and Kadonaga 2000). Ces travaux proposent donc que les éléments Inr et DPE peuvent agir comme un promoteur cœur en l'absence de boîte TATA. D'un point de vue mécanistique, le recrutement de TFIID sur cet élément DPE se fait *via* les sous unités Taf6 et Taf9 de TFIID (Theisen, Lim, and Kadonaga 2010). Ces éléments DPE consensus ne sont retrouvés que chez les eucaryotes pluricellulaires (Smale and Kadonaga 2003).

2) Les facteurs généraux de la transcription.

Le promoteur et ses structures vont permettre le recrutement de la machinerie de transcription. Les ARN polymérases, dont la Pol II, peuvent se lier à l'ADN de façon non spécifique mais sont incapables de reconnaître les promoteurs et d'y initier la transcription. Des expériences de transcription *in vitro* à partir de Pol II purifiée complétée par des fractions cellulaires de protéines brutes a permis de proposer l'existence de facteurs additionnels permettant la transcription (Weil et al. 1979). Le fractionnement d'un extrait nucléaire a permis de sélectionner quatre fractions nommées A, B, C et D dont trois sont nécessaires à l'initiation de la transcription (Matsui et al. 1980; Samuels, Fire, and Sharp 1982). Les fractions A et D ont été appelées respectivement TFIIA et TFIID. La fraction C a été ensuite sous-fractionnée pour mettre en évidence les facteurs TFIIB, TFIIE, TFIIIF et TFIIH (Sawadogo and Roeder 1985). Ces protéines ont été alors nommées facteurs généraux de la transcription. Ces GTFs, conservés chez tous les eucaryotes, ont été extensivement étudiés afin de déterminer leur composition et de comprendre (Tableau 3). Additionnés à la Pol II, ces GTFs sont les composants nécessaires à la formation du PIC *in vitro*. Ces GTFs vont permettre le positionnement de la Pol II sur le promoteur.

Des expériences de biochimie ont permis de proposer un modèle d'assemblage du PIC et de préciser l'ordre d'incorporation des GTFs et leur rôle (Buratowski et al. 1989). Dans la suite, les GTFs seront décrits dans l'ordre de leur incorporation *in vitro*. Une description plus détaillée du PIC et de son assemblage sera présentée ultérieurement (Voir Chapitre I.II.3).

<i>Facteur général</i>	<i>Nombre de sous-unités</i>	<i>Fonctions</i>
<i>TFIIA</i>	2/3	Stabilise la liaison entre la TBP et l'ADN Rôle dans la réponse aux activateurs
<i>TFIIB</i>	1	Stabilise la liaison entre la TBP et l'ADN Participe au choix du TSS Participe au recrutement du complexe Pol II/TFIIIF
<i>TFIID</i>	14/15	Se lie <i>via</i> sa sous-unité TBP au promoteur Induit une courbure de l'ADN Permet le recrutement de TFIIA et TFIIB
<i>TFIIE</i>	2	Recrutement de TFIIH Active l'activité kinase de TFIIH Impliqué dans l'échappement du promoteur
<i>TFIIIF</i>	2/3	Lié à la Pol II, il permet son recrutement Recrute TFIIE et TFIIH Rôle avec TFIIB dans le choix du TSS Participe à l'étape d'élongation
<i>TFIIH</i>	10	Possède une activité hélicase permettant l'ouverture de la bulle de transcription Possède une activité ATPase nécessaire à l'échappement du promoteur Possède une activité kinase phosphorylant le CTD Rôle dans la réparation de l'ADN

Tableau 3: Rôle et composition en sous-unités des facteurs généraux de la transcription (d'après Thomas and Chiang 2006).

a) TFIID.

TFIID est ainsi le premier GTF à se lier au promoteur selon les modèles d'assemblages *in vitro*. Il constitue ainsi le « noyau » autour duquel le PIC se forme. Il est composé de la TBP et de plusieurs TBP-associated factors (TAFs). Chez la levure *S. cerevisiae*, il existe 14 TAFs dont 13 sont conservés chez l'ensemble des eucaryotes (Thomas and Chiang 2006). Chez plusieurs organismes, TFIID a été montré comme interagissant avec la boîte TATA (Sawadogo and Roeder 1985; Buratowski et al. 1988; Cavallini et al. 1988). Il a alors été montré que la TBP, la protéine suffisante à la reconnaissance de la boîte TATA présenté plus tôt, faisait partie de TFIID (Horikoshi et al. 1989). La TBP humaine et de levure sont interchangeables, indiquant une parfaite conservation de fonction (Buratowski et al. 1988). Les TAFs ont été découvertes plus tard. En effet, après la découverte de la TBP et sa capacité à lier la boîte TATA, s'est posé la question de savoir si cette protéine était la seule responsable de la fonction de TFIID. Dans une étude de transcription *in vitro* stimulée par un activateur nommé Sp1, les auteurs ont montré que l'activation ne fonctionnait que lorsque l'on utilisait une purification partielle de TFIID. L'utilisation de TBP recombinante seule ne permettait pas la réponse à l'activateur (Pugh and Tjian 1990). Cette expérience montre donc que la présence des TAFs est nécessaire à la fonction de TFIID. Elle suggère également un rôle co-activateur de TFIID. En plus de leur rôle de co-activateur, certaines TAFs ont un rôle dans la liaison avec des éléments de promoteur cœur tels que l'Inr et le DPE comme mentionné précédemment.

L'organisation structurale de TFIID ainsi que la façon dont il interagit avec l'ADN ne sont pas totalement claires. En effet, la structure atomique n'est pas connue. Il n'existe que des structures de certains domaines de sous-unités de TFIID. En effet, sa grande taille (environ 1MDa) et le grand nombre de sous-unités rendent la purification de TFIID à partir de protéines recombinante difficile. Il faut donc produire TFIID à partir du complexe endogène, ce qui limite les rendements. Cependant, il existe des structures du complexe par cryo-microscopie électronique (cryo-EM) (Andel et al. 1999; Bieniossek et al. 2013). Ces études proposent un modèle de TFIID en forme de fer à cheval organisé en 3 lobes asymétriques. Cette organisation serait flexible et permettrait une interaction avec les différents éléments du promoteur, même s'ils sont espacés. Il faut également noter une forte conservation de structure entre les complexes de levure et humains (Sainsbury, Bernecky, and Cramer 2015).

Concernant son organisation spatiale, TFIID est composé d'un cœur symétrique formé par des dimères de 5 TAFs (Taf4, 5, 6, 9 et 12). Ce cœur permet l'assemblage des autres TAFs (Taf8 et 10) et de la TBP, apportant ainsi l'asymétrie. Les structures partielles des TAFs obtenues montrent des domaines similaires à ceux retrouvés dans les histones (*histone fold*), permettant l'interaction entre les TAFs (Thuault et al. 2002).

Très récemment, le laboratoire d'E. Nogales a publié une structure du complexe TFIID humain lié à un promoteur à une résolution de 7-16 Angström (Louder et al. 2016). Dans cette étude, les auteurs ont purifié TFIID lié à un promoteur dit *super-core*, à savoir un promoteur contenant l'ensemble des motifs connus (TATA, Inr, MTE et DPE). Cela permet d'avoir une affinité améliorée de TFIID au promoteur et de s'affranchir de l'hétérogénéité des complexes purifiés afin d'améliorer la résolution des analyses de microscopie. Les modèles proposent la localisation de la TBP et de TFIIA sur la boîte TATA dans un des lobes (lobe A) de TFIID. TFIIA a un rôle de pont entre les lobes A et B. Le lobe C quant à lui est lié à l'élément promoteur aval. Dans le lobe C, on trouve les Taf1, 2 et 7, responsables de la liaison au MTE et DPE. Ce lobe est rattaché au reste de TFIID par un dimère formé par Taf6. TFIID peut être trouvé dans deux conformations différentes : l'une lorsqu'il est lié à l'ADN, l'autre lorsqu'il est libre. Dans sa forme libre, le domaine N-terminal de Taf1 (*Taf1 N-terminal domain*, TAND) interagit avec la TBP. Ce domaine TAND présente une homologie avec TFIIA et s'insère dans un sillon de TBP, sur le même domaine où TFIIA interagit avec la TBP. Il y a donc une compétition avec TFIIA. En l'absence de TFIIA, l'interaction de Taf1 avec la TBP déstabilise l'assemblage sur la boîte TATA. Dans le modèle de structure obtenu, on note que lorsque TFIID est lié à l'ADN, la sous-unité Taf1 se retrouve à l'opposé de TBP, levant ainsi son rôle inhibiteur. Les modèles structuraux suggèrent le lien entre le changement de conformation et le recrutement de TFIID au promoteur.

Concernant la TBP, une structure 3D obtenue par cristallographie montre la liaison de celle-ci à la boîte TATA (Kim et al. 1993). Sur cette structure, on observe que la partie concave de la TBP interagit avec le sillon majeur de l'ADN comprenant les bases de la boîte TATA. Le motif riche en A et T forme une surface hydrophobe, complémentaire avec la partie de TBP en interaction. La liaison entre TBP et l'ADN se fait principalement par des interactions électrostatiques entre la protéine et le squelette phosphaté de l'ADN et non pas uniquement avec les bases du motif TATA. Le peu de contacts spécifiques avec les bases de l'ADN et TBP explique probablement le fait que celle-ci puisse aussi lier les éléments *TATA-like*. Etant donné sa faible spécificité pour sa liaison à l'ADN, la TBP peut se lier à des séquences d'ADN riche en A et T qui ne sont pas des promoteurs et ainsi mener à une transcription parasite. Des études suggèrent une régulation par une dimérisation de la TBP libre prévenant sa liaison à l'ADN. Le contact entre le domaine TAND de Taf1 et la TBP évoqué plus haut est aussi un mécanisme proposé (Kato et al. 1994; Chasman et al. 1993).

Enfin, lorsque la TBP est liée au promoteur on observe une torsion d'environ 90° de l'ADN au niveau de la liaison. Cette déformation joue un rôle dans le recrutement des autres GTFs, notamment TFIIB, et dans l'orientation du sens de démarrage de la transcription (Nikolov et al. 1996).

TFIID peut intervenir en tant que co-activateur *via* les TAFs. Son rôle en ce sens sera expliqué plus tard dans ce chapitre.

b) TFIIA.

Le facteur général TFIIA est composé de trois sous-unités chez l'Homme (α , β et γ) et de deux sous-unités chez la levure (Toa1 et Toa2). Chez l'Homme, un seul gène, homologue à Toa1 code pour un peptide $\alpha\beta$ clivé ensuite en deux sous-unités distinctes (DeJong and Roeder 1993; Ranish and Hahn 1991). Les deux sous-unités chez la levure sont essentielles, montrant ici un rôle important de ce GTF. TFIIA présente la capacité de se lier à la boîte TATA également mais aussi d'interagir avec la TBP. Son rôle principal est de stabiliser le complexe TBP-TATA (Imbalzano, Zaret, and Kingston 1994). C'est ainsi qu'un complexe TFIIA-TBP-TATA a été cristallisé chez *S. cerevisiae* (Tan et al. 1996) mais aussi chez l'Homme (Bleichenbacher, Tan, and Richmond 2003). La comparaison des structures de ces deux espèces montre une grande similarité. TFIIA permet de stabiliser la conformation de TFIID lié à la boîte TATA et par conséquent le complexe TFIID-TATA-TFIIA (Louder et al. 2016).

De façon intéressante, TFIIA a été montré comme dispensable pour une initiation correcte de la transcription mais comme important dans la transcription activée (Kang et al. 1995; Ma et al. 1996). Ces observations ont abouti à proposer un rôle ambivalent pour TFIIA, à savoir de GTF et de co-activateur. L'activité co-activatrice de TFIIA fonctionne de deux manières. Il présente une activité anti-répressive et une activité co-activatrice à proprement parler.

Le rôle antirepresseur a été mis en évidence dans une expérience de transcription activée *in vitro*. Dans cette expérience, lorsque l'ajout des GTFs est fait sous forme de protéines recombinantes, TFIIA n'est pas nécessaire à l'activation. Lorsque les GTFs sont incorporés sous forme de fractions hétérogènes (contenant d'autres protéines que le GTF purifié), TFIIA devient indispensable à l'activation (Wu 1998). Cela suggère donc que TFIIA contrecarre l'effet represser de protéines inhibitrices de la transcription présentes dans les différentes fraction utilisées. Ce rôle est probablement dû à l'effet stabilisateur du complexe TBP-TATA par TFIIA.

Concernant le rôle co-activateur, TFIIA a été montré comme interagissant avec un certain nombre de protéines activatrices, co-activatrices et des composant du PIC. Par exemple, TFIIA interagit avec l'activateur Gal4-VP16 (Dion and Coulombe 2003). Il a par ailleurs été montré un rôle de TFIIA dans la formation du complexe TFIID-TATA en réponse à l'activateur Rap1 (Papai et al. 2010). Plusieurs études ont montrées des interactions entre TFIIA et des TAFs appartenant à TFIID ou SAGA (Lieberman and Berk 1994; Thomas and Chiang 2006).

Enfin, comme mentionné précédemment, une interaction entre Taf1 et la TBP au niveau de son domaine de liaison à l'ADN a été montrée comme inhibitrice de l'initiation. Il existe une compétition

entre TFIIA et Taf1 pour cette interaction qui permet d'annuler l'inhibition de la reconnaissance de la boîte TATA par TFIID (Kokubo et al. 1998). La présence de TFIIA permet donc de façon générale de stabiliser la TBP sur la boîte TATA formant ainsi une plateforme d'assemblage du PIC.

c) TFIIB.

TFIIB est le seul GTF composé d'une unique protéine. Chez la levure *S. cerevisiae*, le polypeptide de 38kDa est encodé par le gène essentiel *SUA7*. Les structures de TFIIB de différentes espèces ont été obtenues et présentent de grandes similarités (Deng and Roberts 2007). Au niveau N-terminal, on retrouve une structure en doigt de zinc (*B-ribbon*), qui est la région la mieux conservée. Suivent ensuite un domaine *B-reader* et *B-linker*. Enfin, se trouve à la suite un domaine nommé *B-core*, divisé en une partie N et C terminale, comportant une structure dite *cyclin fold* (Figure 6) (Sainsbury, Niesser, and Cramer 2013; Kostrewa et al. 2009).



Figure 6: Schéma représentant les différents domaines structuraux de la protéine TFIIB (Sainsbury, Niesser, and Cramer 2013).

Le trait noir indique les domaines en contact avec la Pol II.

TFIIB est recruté suite à la liaison de TFIID ou de la TBP sur le promoteur et est nécessaire au recrutement de la Pol II (Buratowski et al. 1989). La présence de TFIIB sur ce complexe dit ternaire permet la stabilisation de celui-ci. Ce recrutement est dû à l'interaction possible de TFIIB avec la TBP mais aussi avec la boîte TATA. Un modèle structural du complexe TBP-TATA-TFIIB montre une interaction entre TFIIB et la TBP par le dessous du complexe TBP-ADN (Nikolov et al. 1995). TFIIB est également en contact avec la sous-unité Taf9 (Goodrich et al. 1993). Grâce à la courbure induite par la fixation de la TBP, TFIIB peut également interagir avec l'ADN à la fois en amont ou en aval de la boîte TATA, au niveau des motifs BRE (Nikolov et al. 1995). De plus, dans l'étude de la distribution génomique des GTFs par ChIP-exo faite par le laboratoire de F. Pugh, les auteurs ont montrés que TFIIB présentait quatre points d'interactions à l'ADN *in vivo* (Rhee and Pugh 2012). Deux d'entre eux sont situés de part et d'autre de la boîte TATA à environ 20 pb. Les deux autres points sont eux situés plus loin, à proximité du TSS. Ces observations sont en accord avec les données structurales obtenues à partir de plusieurs structures. La liaison de TFIIB de façon asymétrique peut expliquer également l'orientation correcte du PIC pour une initiation dans la bonne direction. L'ensemble des interactions entre TFIIB et l'ADN ou TFIID/TBP se font *via* le domaine *B-core* (Kostrewa et al. 2009).

En sus de son rôle stabilisateur, TFIIB va permettre de promouvoir la suite de l'assemblage du PIC. En effet, en plus de son interaction avec la TBP et Taf9, TFIIB est engagé dans des contacts avec les deux sous-unités de TFIIF et la Pol II permettant ainsi leur recrutement concomitant dans le PIC. Le contact entre TFIIB et TFIIF se fait par le domaine N-terminal de TFIIB (Ha et al. 1993).

La co-crystallisation de la Pol II et de TFIIB a permis de comprendre la topologie de ces interactions et de proposer des mécanismes d'action (Sainsbury, Niesser, and Cramer 2013). Ainsi, le *B-ribbon* est en interaction avec le *dock* de la Pol II. Le *B-core* N-terminal interagit avec le mur. Le *B-reader* contacte le tunnel de sortie de l'ARN et le sillon. Enfin le *B-linker* est en interaction avec la pince. L'interaction entre TFIIB et le sillon est proposée comme un mécanisme de fermeture de celui-ci. Le contact avec le *dock* induit un changement conformationnel de la Pol II. Cette réorganisation permet à son tour le

contact de TFIIB avec le mur. TFIIF se situe également à cette position, au niveau du mur. L'interaction de TFIIB avec le mur suggère un rôle dans la stabilisation de l'interaction de TFIIF et la Pol II. Le changement de conformation s'accompagne d'un remodelage du site actif augmentant l'activité catalytique de l'enzyme. La liaison de TFIIB avec la Pol II est donc importante pour permettre son recrutement, sa stabilité mais aussi pour rendre ce complexe actif.

L'un des rôles connus et important de TFIIB est celui de la détermination du TSS. Une démonstration de ce rôle a été faite dans une analyse génétique cherchant à identifier des suppresseurs d'un défaut de traduction du gène codant pour le cytochrome C1 (*CYC1*) (Pinto, Ware, and Hampsey 1992). Ce suppresseur, un mutant de *sua7*, induisait un changement de TSS rétablissant un cadre de lecture correct pour la traduction de *Cyc1*. D'autres études notamment chez l'Homme ont montrées un rôle similaire (Deng and Roberts 2007). Le laboratoire de M. Hampsey a obtenu par la suite une collection de mutants de TFIIB induisant pour un grand nombre d'entre eux, un décalage du TSS (Pinto et al. 1994). Les résidus mutés se trouvent dans le domaine *B-reader*. De nouveau, les données structurales proposent un mécanisme associé au rôle de TFIIB dans le choix du TSS. En effet, dans le co-cristal TFIIB-Pol II, on observe la présence du domaine *B-reader* au sein du sillon. L'un des résidus de ce domaine (R64) est en contact avec la séquence Inr suggérant une reconnaissance de ce motif (Sainsbury, Niesser, and Cramer 2013).

d) TFIIF.

TFIIF est composé de trois sous-unités chez la levure (Tfg1, 2 et 3) et de deux chez l'Homme (Rap74 et Rap30) (Henry et al. 1994; Flores et al. 1988). Chez la levure, Tfg1 et Tfg2 sont les analogues de Rap74 et de Rap30 et sont essentielles à la viabilité (Thomas and Chiang 2006). Initialement, les protéines Rap74 et Rap30 ont été identifiées comme faisant partie des protéines interagissantes de la Pol II et nécessaires à une initiation correcte de la transcription (Burton et al. 1988; Burton, Ortolan, and Greenblatt 1986). De façon consistante avec ces observations, il a ensuite été montré que TFIIF est recruté au sein du PIC accompagné de la Pol II (Buratowski et al. 1989). Il faut noter également que TFIIF présente une analogie fonctionnelle, et dans une moindre mesure structurale, avec les facteurs σ bactériens (Hampsey 1998). La liaison entre la Pol II et TFIIF est très stable si bien que 50% de la Pol II isolée depuis un extrait nucléaire de levure sont retrouvés complexés à TFIIF. Cette association présente un intérêt biologique puisqu'elle renforce la spécificité et l'efficacité de la transcription en prévenant les recrutements de la Pol II non-spécifiques sur de l'ADN ne contenant pas de promoteur (Conaway et al. 1991).

Le rôle principal de TFIIF dans l'assemblage du PIC est de permettre le recrutement de la Pol II par son contact avec celle-ci. En plus de cela, il permet de stabiliser le PIC. En effet, TFIIF est également impliqué dans des contacts avec les autres GTFs comme TBP et TFIIB. Cela augmente ainsi le réseau d'interactions entre toutes ces protéines et donc la stabilité de l'assemblage. Il est suggéré également que l'incorporation de TFIIF induit un changement de la topologie du promoteur en permettant à l'ADN de s'enrouler autour de la Pol II. Cela permet la formation d'un complexe TBP-TFIIB-Pol II-TFIIF stable (Robert et al. 1998).

Récemment, une analyse structurale chez l'Homme a proposé un rôle selon lequel TFIIF faciliterait l'insertion et la stabilisation de l'ADN au sein du sillon de la Pol II (He et al. 2013).

Tout comme TFIIB, TFIIF joue un rôle dans la sélection du TSS (Ghazy et al. 2004). Cela peut s'expliquer par la proximité de TFIIB et TFIIF du site actif de la Pol II. De plus des mutations de TFIIF ont été montrées comme suppresseurs de mutants de TFIIB induisant un décalage du TSS (Freire-Picos et al. 2005; Sun and Hampsey 1995).

Enfin, TFIIF est impliqué dans les étapes précoces de l'élongation et dans la résolution des pauses en collaborant avec TFIIS, facteur impliqué dans l'élongation de la transcription (Zhang, Zobeck, and Burton 2005; Zhang and Burton 2004).

e) TFIIIE.

TFIIIE est le cinquième GTF recruté dans la formation du PIC *in vitro*. Chez la levure *S. cerevisiae*, il est formé de deux sous-unités Tfa1 et Tfa2, toutes deux essentielles à la viabilité (Hampsey 1998). Ce GTF interagit à la fois avec TFIIH, les deux sous-unités de TFIIIF et la forme non-phosphorylée de la Pol II (Flores et al. 1988; Maxon, Goodrich, and Tjian 1994).

Des expériences de pontage (*crosslink*) ont permis de situer l'interaction de TFIIIE avec la Pol II et l'ADN (Robert et al. 1996; Forget et al. 2004; Chen, Warfield, and Hahn 2007). Celui-ci se trouve juste avant le TSS et contacte la Pol II *via* Rpb1, à proximité de la pince. Ces études suggèrent un rôle de TFIIIE dans la stabilisation de l'ouverture de la bulle de transcription en liant le brin non transcrit. De plus, TFIIIE présente la capacité de se lier à de l'ADN simple brin (Kuldell and Buratowski 1997) renforçant ainsi l'idée d'un rôle dans la dénaturation de l'ADN requise pour l'initiation.

Enfin, comme mentionné plus haut, TFIIIE interagit avec TFIIH. Ce contact va permettre son recrutement mais pas seulement. En effet, TFIIIE a été montré comme pouvant stimuler les activités enzymatiques de TFIIH, à savoir l'activité CTD-kinase, l'activité ATPase et l'activité hélicase nécessaire à l'ouverture de la bulle de transcription (Ohkuma et al. 1995; Yoshiaki Ohkuma and Roeder 1994; Serizawa, Conaway, and Conaway 1994). Cela suggère donc un rôle de TFIIIE dans la transition entre l'initiation et l'élongation.

f) TFIIH.

TFIIH est le dernier facteur général de la transcription à être incorporé dans le modèle d'assemblage *in vitro*. Ce facteur est composé de 10 sous-unités organisées en deux modules : un module cœur (cTFIIH) composé de 7 sous-unités et un module kinase (TFIIK) composé des 3 sous-unités restantes (Gibbons et al. 2012). Des modèles structuraux ont permis de montrer une grande conservation entre l'Homme et la levure (Schultz et al. 2000; Chang and Kornberg 2000). TFIIH s'organise sous forme d'un anneau correspondant au module cœur et pouvant interagir avec l'ADN, sur lequel repose le module kinase. La jonction entre les deux modules correspond à l'emplacement de Rad3/XPD.

Comme évoqué dans le paragraphe précédent, TFIIH comporte trois activités enzymatiques. Une activité kinase, portée par la sous-unité Kin28 (Cdk7 chez l'Homme) permet la phosphorylation du CTD de la Pol II sur la serine 5. Cette modification est associée au passage en élongation de la transcription (Ohkuma et al. 1995). TFIIH possède également une activité hélicase, portée par les deux sous-unités Ssl2/XPB et Rad3/XPD. L'activité hélicase permet l'ouverture d'un ADN double brin. Elle est donc nécessaire à l'ouverture de la bulle de transcription. Il apparaît néanmoins que seule Ssl2/XPB soit responsable de l'ouverture de la bulle de transcription et il est proposé que celle-ci agisse plutôt en tant que translocase, pour déplacer l'ADN (Douziech et al. 2000; Grünberg, Warfield, and Hahn 2012). L'activité ATPase, quant à elle, est portée par ces deux mêmes hélicases. Cette activité constitue une source énergétique nécessaire à l'ouverture de la bulle de transcription.

Il est également établi que TFIIH interagit avec de nombreux activateurs de la transcription, lui conférant ainsi un rôle important dans l'activation et la régulation de la transcription (Zurita and Merino 2003). L'interaction avec les activateurs peut augmenter le recrutement de TFIIH au sein du PIC. Ce GTF peut également agir sur l'activation de la transcription en modifiant les activateurs pour moduler leur activité. Il a notamment été montré que TFIIH, *via* sa kinase, pouvait phosphoryler le domaine d'activation du récepteur nucléaire à l'acide rétinoïque (*retinoic acid receptor*, RAR). Cela augmente la réponse transcriptionnelle liée à ce récepteur nucléaire (Rochette-Egly et al. 1997; Bastien

et al. 2000).

Enfin, en plus de son rôle dans la transcription, TFIIH est un facteur impliqué dans la réparation de l'ADN. Il est impliqué dans la voie de réparation des dommages à l'ADN par excision de nucléotides (*nucleotide excision repair*, NER). Cette voie de réparation repose sur la coupure en amont et en aval du squelette phosphaté de l'ADN pour extraire un fragment contenant la lésion qui sera ensuite re-synthétisé. TFIIH joue un rôle dans cette voie par ses deux sous-unités hélicases Ssl2/XPB et Rad3/XPD (Compe and Egly 2012). Il est très important de noter que chez l'Homme, des mutations dans des sous-unités de TFIIH provoquent des pathologies très lourdes telles que Xeroderma pigmentosum (XP), la trichotiodystrophie (TTD) ou le syndrome de Cockayne (CS). Ces maladies provoquent une sensibilité extrême aux rayonnements ultra-violet (UV), des prédispositions importantes à des cancers et des maladies mentales. Il semblerait que ces pathologies ne peuvent être totalement expliquées par le défaut de réparation de l'ADN, montrant ainsi le rôle primordial de TFIIH dans la transcription (Compe and Egly 2012).

3) L'initiation de la transcription.

a) La mise en place du complexe de pré-initiation.

Le recrutement des GTFs et de la Pol II sur le promoteur sous forme de PIC constitue la première étape de la transcription. Des expériences de biochimie ont permis de préciser le rôle et l'ordre d'assemblage des GTFs dans la formation du PIC *in vitro* (Buratowski et al. 1989). Les auteurs ont ajouté sur une matrice d'ADN les différents GTFs en faisant varier l'ordre d'addition. Ils ont ainsi pu isoler 5 complexes intermédiaires. Cela a mené à proposer un modèle d'assemblage linéaire et séquentiel du PIC (Figure 7). Selon ce modèle, l'assemblage commence par la liaison de TFIID sur le promoteur *via* sa sous-unité TBP. TFIIA et TFIIB sont ensuite recrutés sur ce complexe TFIID-ADN afin de le stabiliser. La Pol II, accompagnée de TFIIF est ensuite incorporée, suivie de TFIIIE et de TFIIH.

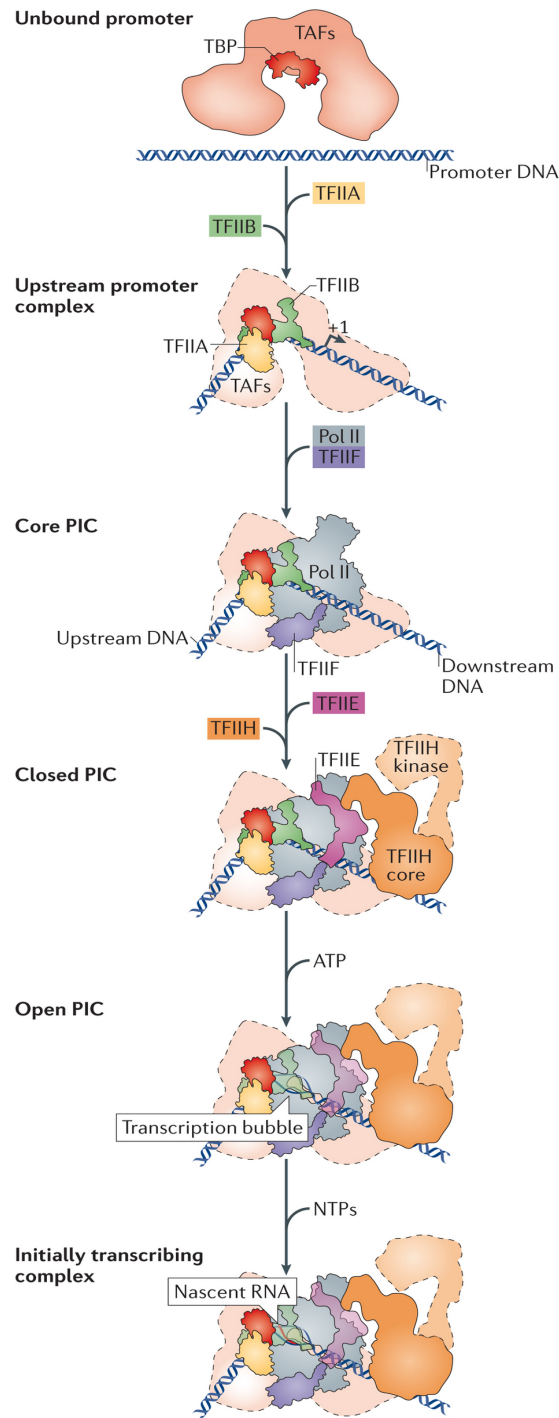


Figure 7 : Schéma représentant l'assemblage séquentiel et linéaire du PIC *in vitro*. (d'après Sainsbury, Bernecky, and Cramer 2015)

Le PIC est un complexe gigantesque comprenant plus de 50 polypeptides. Leur incorporation est donc un processus extrêmement complexe faisant appel à un réseau d'interaction tout aussi complexe déjà évoqué dans la présentation des différents GTFs. Les études structurales de PIC permettent une meilleure compréhension des mécanismes moléculaires responsables de cet assemblage en permettant la visualisation de l'ensemble de ces protéines. Dans l'une de ces études, (He et al. 2013) les auteurs ont étudié par microscopie électronique différents intermédiaires de la formation du PIC humain en fonction de leur ordre d'assemblage. Ils ont assemblé le PIC par ajout successif des GTFs purifiés sur une matrice d'ADN contenant un promoteur. Le premier complexe observé est celui comprenant l'ADN, TBP, TFIIA, TFIIB et la Pol II. De manière intéressante, sur ce premier intermédiaire, la molécule d'ADN n'est pas visible. Cela suggère que celle-ci n'occupe pas une position unique dans ce modèle. De plus, on constate que la Pol II n'est en contact avec aucun autre GTF que TFIIB et qu'elle ne contacte pas l'ADN. Les auteurs ont ensuite incorporé TFIIF pour former un autre intermédiaire. Cet ajout a pour effet de bloquer cette fois l'ADN dans une position définie. Cela suggère donc un rôle de TFIIF dans le modelage du PIC afin de le faire passer dans une conformation fixée. L'ajout subséquent de TFIIE au complexe montre une localisation de TFIIE/TFIIF à la surface de la Pol II, proche du site actif et du sous-complexe Rpb4/7. La présence de ces deux facteurs permet donc une stabilisation du PIC en formation mais aussi un placement correct de l'ADN dans le site actif de la Pol II. Enfin, lorsque TFIIH est ajouté au PIC, on observe que celui-ci peut interagir à la fois avec TFIIE à la base de la pince mais également avec l'ADN. La sous-unité XPB se place elle en aval du PIC. Cette étude fait aussi la comparaison entre les PICs fermés et ouverts. Pour cela, la matrice d'ADN utilisée contient une partie où les bases ne sont plus complémentaires afin de créer des mésappariements et donc une hélice ouverte. Cela permet la mise en évidence des transitions structurales observées lors de l'ouverture de l'ADN dans la bulle de transcription. Le modèle structural obtenu permet aux auteurs de proposer une action de XPB en tant que translocase plutôt qu'en tant qu'hélicase. Cette observation est consistante avec ce que le laboratoire de S. Hahn propose (Grünberg, Warfield, and Hahn 2012). En effet, il est suggéré que XPB induit un mouvement de l'ADN qui permet l'insertion de 10pb dans le sillon de la Pol II. Cela provoque un super-enroulement de l'ADN qui induit son ouverture. TFIIB et TFIIF rentrent en contact avec l'ADN pour maintenir cet état ouvert.

Au même moment, le laboratoire de R. Kornberg a également proposé un modèle structural de PIC (Murakami et al. 2013). Ce modèle, obtenu chez la levure *S. cerevisiae*, consiste en l'analyse du complexe entièrement formé, contrairement à celui d'E. Nogales (He et al. 2013) qui analyse différents intermédiaires. Ces deux modèles se contredisent en particulier sur les positionnements de TFIIE, TFIIF et TFIIH mais également sur la position de l'ADN par rapport à la Pol II. Dans le modèle de R. Kornberg, il est proposé que l'assemblage du PIC se fait uniquement en présence des GTFs sur le promoteur. La Pol II serait incorporée ensuite à l'aide de TFIIF qui aurait un rôle de chaperone. Dans le modèle proposé par cette étude, la Pol II n'est pas en contact avec l'ADN au cours de l'assemblage mais uniquement avec les GTFs. Le promoteur est lié par TBP, TFIIB et TFIIE en amont et par TFIIH en aval. Entre les deux se trouverait l'ADN libre pouvant subir l'action des hélicases et ainsi l'ouverture de la bulle de transcription et l'insertion de l'ADN dans le sillon de la Pol II. Les GTFs stabilisent ensuite ce complexe.

L'étude de R. Kornberg combine à la fois des données structurales et des données de pontages chimiques (*crosslinks*) couplés à la spectrométrie de masse. Cette méthode permet l'analyse des interactions protéine-protéine et donc d'affiner les modèles reconstruits. Néanmoins, la force de l'étude du laboratoire d'E. Nogales repose dans le fait qu'elle s'intéresse aux intermédiaires d'assemblage, permettant la visualisation des changements se produisant sur le PIC. De plus, elle propose une visualisation des changements d'états entre les formes ouvertes et fermées du PIC, conforme à des observations mécanistiques concernant le rôle de XPB.

De nouvelles structures de PIC à la fois chez la levure et chez l'Homme ont été obtenues très récemment (He et al. 2016; Plaschka et al. 2016a). Le travail d'E. Nogales dans cette nouvelle structure propose cette fois un modèle de PIC formé entièrement, sans les intermédiaires. Le complexe est composé de l'ensemble des GTFs et de TFIIS. Il est analysé par cryo-microscopie électronique à une résolution proche de la résolution atomique. Les auteurs, en faisant varier les matrices d'ADN,

proposent les conformations fermée, ouverte et en début de transcription. Les résultats proposés ici sont en accord avec leur précédent modèle. L'une des nouveautés de ce modèle est la visualisation précise de TFIIIE grâce à l'amélioration de la résolution. Sur ce modèle, TFIIIE se situe à proximité du dimère Rpb4-7 de la Pol II, en accord avec les données existantes. Il est également en contact avec l'ADN, avant le TSS. TFIIIE semble avoir une structure très flexible. La comparaison des états ouverts et fermés du PIC met en évidence un changement de conformation de TFIIIE lors de la transition entre les deux états. Ce changement de conformation semble nécessaire au chargement de l'ADN dans le sillon de la Pol II, qui est ensuite stabilisé par la nouvelle conformation prise par TFIIIE.

Le complexe de *S. cerevisiae* étudié dans le travail du laboratoire de P. Cramer (Plaschka et al. 2016) est un PIC composé de tous les GTFs, à l'exception de TFIIH, dans l'état ouvert et fermé également. Le PIC est préparé en présence du complexe Médiateur, bien que sa structure soit exclue du modèle proposé. Dans la reconstruction de leur modèle, les auteurs ont inclus des données structurales existantes de TFIIH. Les résultats obtenus dans ce modèle sont en accord avec le modèle humain, notamment quant à la localisation de TFIIIE et TFIIIF (Figure 8). Le modèle d'ouverture de la bulle de transcription semble là aussi lié à l'activité translocase de TFIIH. Les différences observées entre cette nouvelle structure de PIC chez la levure et la précédente faite par le laboratoire de R. Kornberg peuvent trouver une explication dans le fait que dans cette nouvelle étude, le PIC a été préparé en présence du Médiateur. La présence du Médiateur peut ainsi autoriser des conformations différentes du PIC. De plus, les matrices d'ADN utilisées sont également différentes et peuvent avoir un effet différentiel.

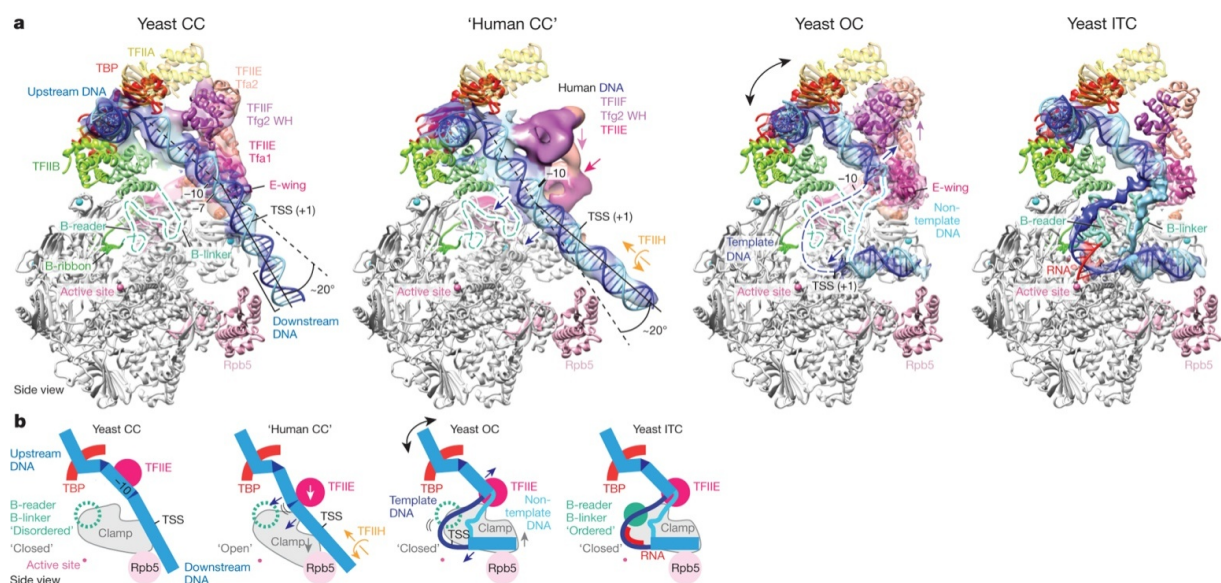


Figure 8 : Modèles structuraux du PIC dans les conformations fermées, ouvertes et en début de transcription (Plaschka et al. 2016).

Sur cette figure, les modèles structuraux du PIC (a) dans les états fermés (chez la levure et chez l'Homme), ouvert et en début de transcription (chez la levure) sont représentés. Les mouvements de l'ADN et de certaines protéines du PIC lors des changements de conformation sont indiqués et représentés schématiquement (b).

Il est difficile de déterminer lequel de ces modèles est le plus proche de la réalité sachant que dans ces études structurales, l'analyse n'est faite que sur une partie des images acquises, pouvant privilégier une conformation plutôt qu'une autre. De plus, les conditions expérimentales de formation du PIC ne sont pas identiques dans ces trois études. Une conformation particulière peut ainsi être privilégiée par sa stabilité dans les conditions utilisées. Enfin, il est possible que selon la nature et la composition du

promoteur, les mécanismes puissent varier. Il semble cependant que les deux derniers modèles proposés soient plus près de la réalité sachant qu'ils sont en accord avec des données biochimiques et mécanistiques, notamment quant à l'activité translocase de TFIIH ainsi qu'au positionnement de TFIIIE et TFIIIF.

Malgré tout, l'ensemble de ces études apporte des éclairages quant aux mécanismes moléculaires permettant le recrutement et l'assemblages des différents composants du PIC cohérents avec les modèles observés *in vitro*.

L'assemblage du PIC *in vivo* est quant à lui plus compliqué puisqu'il fait intervenir des co-activateurs, agissant à différents moments dans le processus. Notre laboratoire a ainsi pu proposer un modèle selon lequel il existe différents ordres d'arrivées des constituants des PICs, formant ainsi des intermédiaires d'assemblages pouvant différer du modèle linéaire présenté (Esnault et al. 2008; Soutourina et al. 2011; Eyboulet et al. 2015). L'analyse de la composition des PICs par CHIP-exo a montré que de façon générale, les PICs sont composés de façon homogène, les différences étant surtout liées au niveau d'occupation, en accord avec le niveau d'expression du gène transcrit (Rhee and Pugh 2012). L'assemblage du PIC *in vivo* sera abordé plus en détail dans le chapitre II.

b) La transition entre l'initiation et l'élongation.

A la suite de la formation du PIC, le cycle de transcription est poursuivi par l'élongation. La transition entre ces deux étapes est marquée par des modifications dans le PIC et sur la Pol II. Cela a pour effet de permettre l'échappement du promoteur. En effet, le CTD de la Pol II est phosphorylé sur la Ser5 par la sous-unité kinase de TFIIH. Cette modification aboutit à la désolidarisation de la Pol II et de la TBP et TFIIIE (Usheva et al. 1992). Ensuite, il a été montré que l'élongation ne peut fonctionner que si les deux brins d'ADN sont séparés (Holstege, Fiedler, and Timmers 1997). Là encore, cette transition est médiée par TFIIH avec sa sous-unité Ssl2/XPB (Guzder et al. 1994; Moreland et al. 1999). Cela a pour effet l'ouverture de la bulle de transcription, qui reste fixe en aval mais qui s'allonge dans le sens de l'élongation (Pal, Ponticelli, and Luse 2005).

S'ensuit alors le début de l'élongation. Les premières réactions sont généralement abortives. C'est à dire que l'ARN produit ne comporte que quelques nucléotides (5-10) et est immédiatement libéré. Pour palier à cet effet, la machinerie de la transcription fait appel à TFIIH à l'aide d'un mécanisme ATP dépendant. En effet, des tests de transcription *in vitro* utilisant un analogue inhibiteur compétitif de l'ATP ont montrés des défauts dans l'échappement du promoteur. L'ajout de TFIIH au contraire stimule cet échappement (Dvir, Conaway, and Conaway 1996; Dvir, Conaway, and Conaway 1997). TFIIIE joue aussi un rôle dans l'élongation précoce en stimulant l'activité de TFIIH (Lin and Gralla 2005). Enfin, TFIIIF est aussi impliqué dans la limitation de la transcription abortive en augmentant la vitesse d'introduction des nucléotides dans le site actif de la Pol II (Yan et al. 1999). Une fois que le transcrit dépasse une longueur de 10 nucléotides, la bulle de transcription se déplace et est considérée comme échappée du promoteur : c'est le début de l'élongation processive.

4) L'élongation de la transcription par la Pol II

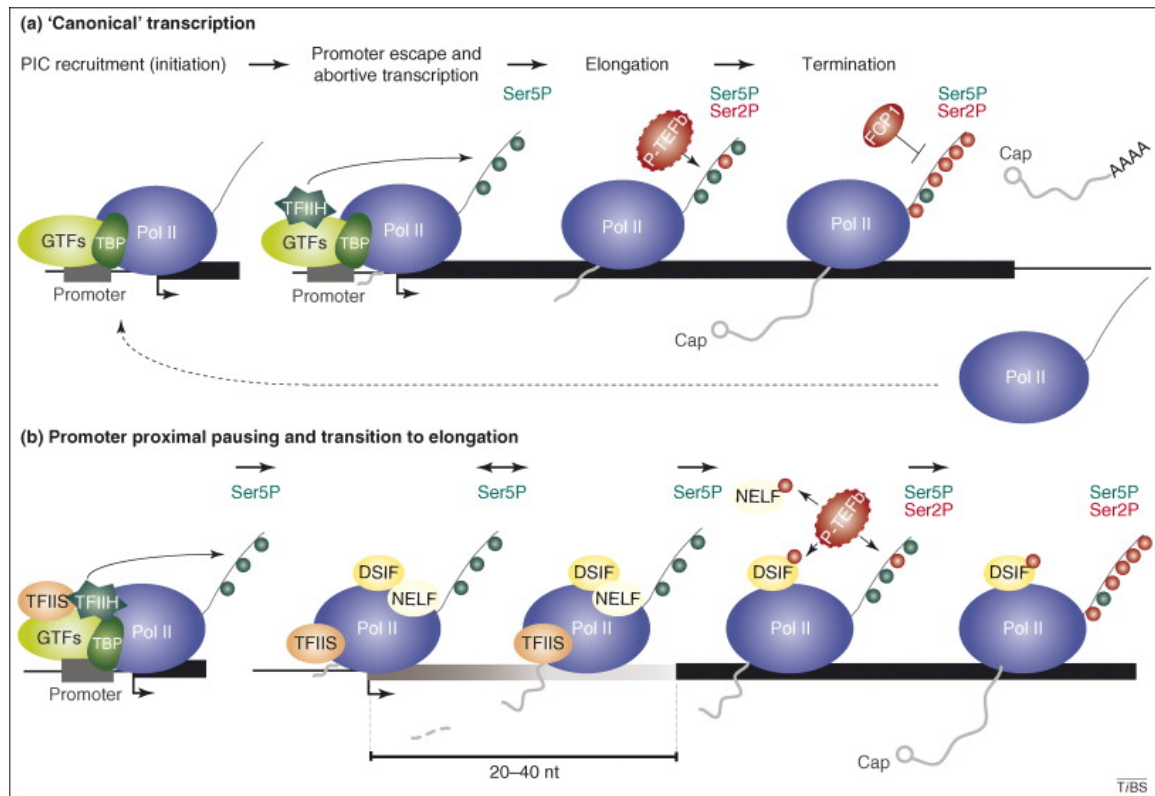


Figure 9: Schéma représentatif de l'ensemble de la transcription par l'ARN polymérase II (d'après Koch et al. 2008)

(a) La transcription est initiée par la formation du PIC. La kinase de TFIIF phosphoryle la sérine 5 du CTD, facilitant l'échappée du promoteur. L'élongation est initiée par la phosphorylation de la sérine 2 par la kinase Cdk9 du facteur d'élongation P-TEFb. La phosphorylation de la sérine 5 décroît, tandis que la phosphorylation de la sérine 2 augmente le long du gène. Une fois la Pol II arrivée au site de terminaison, la sérine 2 du CTD est déphosphorylée. L'ARN coiffé et polyadénylé est relargué, la Pol II peut être recyclée.

(b) La Pol II est engagée en transcription et produit des transcrits de 20 à 40 nucléotides. Cependant, seule la sérine 5 du CTD est phosphorylée. Ce processus de pause est régulé par les protéines DSIF, NELF et TFIIS. La transition nécessite le recrutement de P-TEFb, qui phosphoryle la sérine 2 du CTD, et les facteurs NELF et DSIF. Lorsque NELF se dissocie, l'élongation productive peut être amorcée.

Après l'initiation de la transcription et l'échappement du promoteur, la Pol II entre en élongation (Figure 9). Cette étape du cycle de transcription peut être séparée en deux phases : une phase d'élongation précoce et une phase d'élongation productive. Durant la phase d'élongation précoce, la Pol II peut se retrouver en état de pause à proximité du promoteur (Kwak and Lis 2013). Ce phénomène est variable selon les gènes et les organismes considérés.

a) La pause à proximité du promoteur.

La distribution génomique de la Pol II a été analysée par ChIP dans plusieurs organismes. L'une des premières études chez la drosophile a montré un résultat intrigant à savoir une accumulation de la Pol II à proximité du TSS (Muse et al. 2007). Cette accumulation est due à un processus de pause qui se déroule à environ 20-60 nt après TSS (Kwak and Lis 2013). Deux complexes protéiques sont impliqués dans ce mécanisme, DSIF (*DRB-sensitivity inducing factor*) et NELF (*negative elongation factor*) (Tadashi Wada et al. 1998; Yamaguchi et al. 1999). Le complexe DSIF est composé de deux facteurs d'élongation appelés Spt4 et Spt5, tous deux conservés de la levure à l'Homme (Hartzog et al. 1998). NELF quant à lui est constitué de 5 sous-unités appelées NELF-A, B, C, D et E (Narita et al. 2003), absentes chez la levure. Ces deux facteurs présentent la capacité de se lier à la Pol II en pause mais également à l'ARN naissant. Les mécanismes moléculaires induisant cette pause ne sont pas très clairs mais il est proposé que DSIF puisse se lier à l'ARN naissant lorsque celui-ci est encore relativement court (environ 20 nt). Cette liaison permet ensuite le recrutement de NELF et probablement d'autres facteurs pour induire la pause (Missra and Gilmour 2010).

De façon intéressante, l'étude de la distribution de la Pol II sur le génome de la levure montre un profil particulier. En effet, globalement, la Pol II est distribuée de façon uniforme le long du corps des gènes, sans accumulation au niveau du TSS (Steinmetz et al. 2006). Cela suggère l'absence de la pause où tout du moins une résolution très rapide de ce mécanisme.

Chez les mammifères et la drosophile, la distribution de la Pol II montre un enrichissement près du TSS et une faible occupation sur le corps des gènes. Il a été montré que ces signaux d'occupation de Pol II à proximité du promoteur correspondent à des complexes compétents pour l'élongation et non pas à des PICs (Core, Waterfall, and Lis 2008). Concomitamment, on retrouve également à proximité de ces signaux de Pol II des marques d'histones spécifiques des gènes actifs (Guenther et al. 2007). Ces enrichissements correspondent donc bien à des complexes d'élongations en pause. Il faut noter que ce processus de pause est très important puisqu'il concerne presque la moitié des gènes indépendamment du fait qu'ils soient inductibles ou constitutifs (Core, Waterfall, and Lis 2008). Néanmoins, il a été montré que les gènes de réponse à certains stress ou des gènes impliqués dans les processus développementaux étaient plus concernés (Muse et al. 2007; Zeitlinger et al. 2007; Core and Lis 2008).

Les mécanismes impliqués dans ces pauses ont été le sujet de nombreuses études afin de comprendre leur rôle physiologique. A l'heure actuelle, il est proposé que le site de pause soit défini selon trois modèles non exclusifs : un modèle cinétique, un modèle de barrière et un modèle d'interaction (Figure 10). Le modèle cinétique propose que le recrutement des facteurs de pause soit en compétition avec la vitesse d'élongation de la Pol II. Ainsi, selon le type de gène et la nature des bases à transcrire, l'énergie nécessaire à l'élongation est différente et module la vitesse d'élongation. Sur les gènes plus lents à transcrire, les facteurs de pause sont recrutés plus facilement. Le modèle de barrière propose que la Pol II en élongation peut rencontrer un obstacle physique, un nucléosome, sur la chromatine perturbant ainsi sa progression. Enfin, le modèle d'interaction propose que la Pol II puisse interagir avec des protéines induisant sa pause (Kwak and Lis 2013).

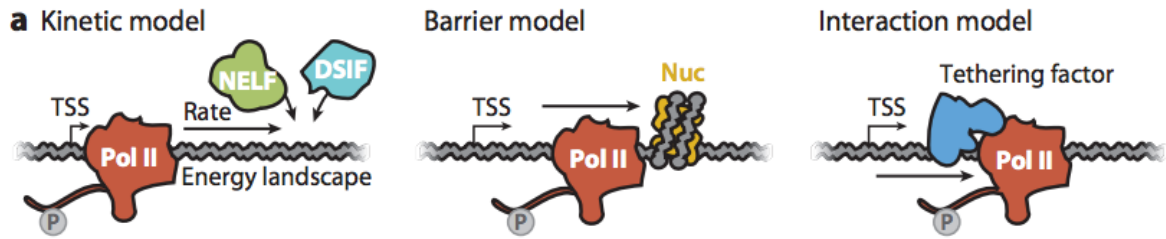


Figure 10: Schéma représentant les trois modèles proposés pour l'établissement des pauses de la Pol II (d'après Kwak and Lis 2013)

La pause de la Pol II se trouve à des positions différentes selon le gène considéré. L'analyse de la région promotrice des gènes montre que les éléments régulateurs du promoteur (Inr, DPE ..) ont une influence sur le site de pause. Ainsi, les gènes contenant des éléments promoteurs forts (le plus proche du consensus) tendent à présenter une pause de la Pol II proche du promoteur alors que les gènes avec des éléments promoteurs plus faible présentent une pause plus proche du nucléosome +1 (Kwak et al. 2013). Il est donc proposé un mécanisme dirigé par le promoteur ou un mécanisme dirigé par les nucléosomes.

Ces pauses ont un rôle physiologique. Elles constituent d'abord un niveau de régulation transcriptionnel supplémentaire. En effet, elles permettent de limiter le taux de transcription d'un gène ou au contraire, une activation plus rapide, l'initiation n'étant plus une étape limitante. Il apparaît que ces pauses contribuent également à permettre au gène transcrit de rester dans un état ouvert et actif (Jonkers and Lis 2015).

Une étude récente a permis de montrer que les pauses étaient relativement stables dans le temps. Il est proposé qu'elles offrent une fenêtre temporelle permettant l'intégration de signaux régulateurs par la Pol II et un point de contrôle avant l'élongation productive (Henriques et al. 2013). Enfin, les pauses pourraient être impliquées dans l'ajout de la coiffe en 3' de l'ARN naissant et constituer un contrôle et un pré-requis pour l'élongation productive (Pei, Schwer, and Shuman 2003; Kwak and Lis 2013).

La résolution des pauses est le mécanisme permettant le passage à l'élongation productive. Elle implique l'action d'un complexe protéique appelé P-TEFb (*positive transcription elongation factor*). P-TEFb est un hétérodimère composé de la cycline C et de la kinase Cdk9. Son action consiste en la phosphorylation de la Ser2 du CTD de la Pol II et en la phosphorylation de DSIF et NELF. Ces modifications vont avoir pour conséquence la dissociation de NELF. DSIF reste associé avec la Pol II et devient, sous sa forme phosphorylée, un facteur d'élongation productif (Wada et al. 1998). La modification du CTD, quant à elle, permet le recrutement d'autres facteurs d'élongation mais aussi de complexes protéiques impliqués dans la maturation de l'ARN (Kwak and Lis 2013).

Lorsque la Pol II entre en état de pause, elle subit un mouvement de recul sur l'ADN. Cela aboutit, au sein du site actif, à un décalage de l'extrémité 3' de l'ARN transcrit avec la matrice d'ADN. Pour remédier à cela, le facteur d'élongation TFIIIS va intervenir en stimulant l'activité nucléase de la Pol II afin de cliver l'ARN. Cela a pour conséquence le réaligement de l'ARN avec l'ADN (Adelman et al. 2005).

b) L'élongation productive.

Lorsque la Pol II entre dans la phase d'élongation productive, la bulle de transcription se déplace le long de l'ADN qui s'ouvre au fur et à mesure. La polymérisation de l'ARN se fait par incorporation des NTPs (nucléotides tri-phosphate), formation de la liaison phospho-diester et libération de phosphate. L'ARN synthétisé forme un duplex ARN-ADN d'environ 8 nt, jouant un rôle important dans la stabilité du complexe d'élongation (Kireeva et al. 2000).

La vitesse d'élongation de la transcription est très variable en fonction du gène considéré. En effet, le complexe d'élongation rencontre également durant sa progression des obstacles comme des nucléosomes, distribués de façon différente selon les gènes. Un travail sur l'analyse *in vivo* des transcrits naissants par NET-seq (*native elongating transcript*) a permis de montrer, par exemple, que la Pol II passe autant de temps en élongation qu'en pause (Churchman and Weissman 2011). Des études ont mis en évidence des facteurs impliqués dans le déplacement des nucléosomes pour limiter ces barrières. Il s'agit, par exemple, des facteurs FACT (*facilitates chromatin transcription*) ou Spt6, tout deux possédant la capacité de se lier aux histones (Kwak and Lis 2013). Ces protéines voyagent avec la Pol II en élongation et vont permettre le déplacement des nucléosomes bloquant la progression de la transcription. D'autres protéines sont requises pour stimuler l'activité de la Pol II et ainsi limiter les pauses. Il s'agit par exemple de TFIIF, ELL (*eleven-nineteen lysine-rich leukemia 1*) ou les élonguines. TFIIS intervient encore pour résoudre les éventuelles pauses au cours de l'élongation. Chez les mammifères, l'élongation ne se déroule pas à la même vitesse tout au long du gène. Au début du gène, elle est d'environ 0,5kb par minute alors qu'elle peut atteindre 5-5kb par minutes après 15kb (Jonkers and Lis 2015). Il est proposé que cela soit dû au fait que les modifications de la Pol II (phosphorylation) et des complexes associés soit graduelles et cumulatives au cours de la progression. Par exemple, la phosphorylation de la Ser2 du CTD atteint son maximum après quelques kilo-bases (Heidemann et al. 2013). Cette augmentation graduelle aurait un sens biologique puisqu'elle pourrait permettre le recrutement des facteurs impliqués dans les processus co-transcriptionnels tels que l'épissage.

Une fois que la bulle de transcription arrive à l'extrémité 3' du gène à transcrire, la Pol II ralentit et peut marquer une pause afin de permettre la terminaison de la transcription.

5) La terminaison de la transcription et sa ré-initiation.

a) La terminaison.

La terminaison est la dernière étape du cycle de transcription. Elle est importante puisqu'elle permet la définition de la fin du transcrit. Cette étape est aussi très importante car elle empêche l'interférence avec les gènes situés en aval.

La terminaison prend place lorsque l'ARN polymérase, l'ADN et l'ARN sont libérés. C'est un processus complexe et pas totalement compris à ce jour. Un fait bien connu néanmoins est que cette terminaison est dépendante de la poly-adénylation (poly(A)) de l'ARN *néo*-synthétisé (Birse et al. 1998; Osheim, Proudfoot, and Beyer 1999). L'arrivée sur le signal de poly(A) engendre un ralentissement voire une pause du complexe d'élongation et provoque sa destabilisation. Chez la levure *S. cerevisiae*, la terminaison de la transcription des ARN messagers fait intervenir les complexes CPF (*cleavage and polyadenylation factor*) et CF (*cleavage factor*) (Porrua and Libri 2015). Ils sont recrutés par une interaction entre la sous-unité Pcf11 et la Ser5 phosphorylée du CTD. Des interactions avec des séquences spécifiques du transcrit interviennent également. La sous-unité Ysh1 de CPF possède une

activité endonucléase qui permet le clivage du transcrit au niveau du site de poly(A). Le transcrit libéré est aussitôt pris en charge par la poly(A)-polymérase Pap1 qui synthétise la queue poly(A). L'ARN sera ensuite exporté dans le cytoplasme.

Chez *S. cerevisiae*, la Pol II continue à transcrire l'ADN après le clivage pendant environ 150pb. La libération de la Pol II est à l'heure actuelle en débat. Deux mécanismes sont proposés. Un modèle allostérique propose que la liaison des facteurs de terminaison engendre un changement de conformation de la Pol II qui mène à la perte des facteurs d'élongation et à la dissociation de la Pol II. L'autre modèle, appelé torpedo, suggère que le complexe de terminaison permet le recrutement d'une 5'-3' exonucléase (Rat1) capable de dégrader le transcrit jusqu'à la Pol II menant à sa dissociation. (Kim et al. 2004; Porrua and Libri 2015; Richard and Manley 2009). Un modèle unifiant le processus allostérique et torpedo a été aussi proposé (Luo, Johnson, and Bentley 2006). Basé sur des expériences *in vitro* les auteurs proposent que Rat1 est recruté en même temps que le complexe de terminaison et que la dégradation se déroule en même temps que le changement allostérique libérant ainsi la Pol II.

Un autre mode de terminaison a été mis en évidence pour les ARNs non codants (ARNnc). En effet, ceux-ci ne possèdent pas de poly(A) et ne peuvent pas, par conséquent, utiliser les mêmes mécanismes. Pour la terminaison de leur transcription, ces ARNnc vont être directement reconnus par un complexe NNS (Nrd1-Nab3-Sen1). Ce complexe est recruté par une interaction avec la Ser2 phosphorylée du CTD. Il va permettre le clivage du transcrit et la dissociation de la Pol II (Lykke-Andersen and Jensen 2007; Porrua and Libri 2015).

Les mécanismes de terminaison chez les mammifères sont aussi peu clairs à ce jour mais la conservation des protéines impliquées suggère des mécanismes similaires.

b) La ré-initiation de la transcription.

Après l'initiation de la transcription et le départ en élongation de la Pol II, il a été observé *in vitro*, que TFIIA, TFIIID, TFIIH et le Médiateur restent sur le promoteur. Cet assemblage forme une plateforme (*scaffold*) pouvant faciliter la ré-initiation (voir chapitre II). Ces observations n'ont pas encore été faites *in vivo* mais cette plateforme pourrait exister et permettre un recrutement plus rapide des GTFs manquant pour initier un nouveau cycle de transcription (Liu et al. 2004; Yudkovsky, Ranish, and Hahn 2000). Parallèlement, la Pol II subit des modifications telles que les déphosphorylations de son CTD avant de pouvoir de nouveau entrer dans un cycle de transcription (Yeo et al. 2003).

6) La dynamique de la transcription.

La transcription, comme tous les processus cellulaires, est réalisée dans un contexte obéissant à des lois chimiques, thermodynamiques et cinétiques. De ce fait, c'est un processus possédant une dynamique intrinsèque très grande. De plus, chaque étape du cycle de la transcription possède un caractère dynamique ajoutant de la complexité sur le signal transcriptionnel final. La phase d'initiation est dépendante de la reconnaissance du promoteur par les protéines de la machinerie de la transcription. L'élongation est un processus dynamique dont la vitesse varie. Les méthodes utilisées pour étudier la transcription *in vitro* (approches biochimiques) ou *in vivo* (ChIP) ne permettent pas la mesure de cette dynamique. En effet, elles ne permettent que la visualisation d'un état fixé, à un instant donné. De plus, les méthodes d'investigation *in vivo* sont réalisées sur des populations de cellules. Il est difficile de visualiser les comportements individuels des cellules et donc d'en connaître la variabilité. Cependant, ces dernières années, les progrès techniques ont

permis de mettre au point des méthodes pour investiguer de façon dynamique des processus cellulaires. On parle de biologie quantitative ou systémique. L'un des principaux progrès est celui des méthodes d'imagerie par microscopie photonique à fluorescence. Des méthodes à haute-résolution ont été mises au point permettant la visualisation par exemple de molécules uniques dans des cellules uniques.

L'une des premières questions adressée avec ces méthodes est celle de savoir quel est le comportement des facteurs de transcription dans le noyau et comment ils pouvaient reconnaître leur site de liaison sur l'ADN. Pour cela, la méthode du FRAP (*fluorescence recovery after photobleaching*) a été très informative (Hager, McNally, and Misteli 2009). Cette méthode repose sur un principe assez simple à savoir que les fluorophores, lorsqu'ils sont soumis à une forte exposition lumineuse, perdent leur capacité de fluorescence. Ce phénomène est appelé photoblanchiment (*photobleaching*). La protéine d'intérêt est fusionnée à une protéine fluorescente (par exemple la GFP, *green fluorescent protein*). Sous l'objectif du microscope, on soumet à l'aide d'un laser une portion étroite du noyau à un photoblanchiment. Une zone non fluorescente apparaît ainsi (Figure 11). L'analyse en fonction du temps la réapparition de la fluorescence dans cette parcelle du noyau permet de déduire les propriétés cinétiques de la protéine considérée. Celles-ci sont directement liées à la mobilité de la protéine d'intérêt par sa diffusion au sein du noyau.

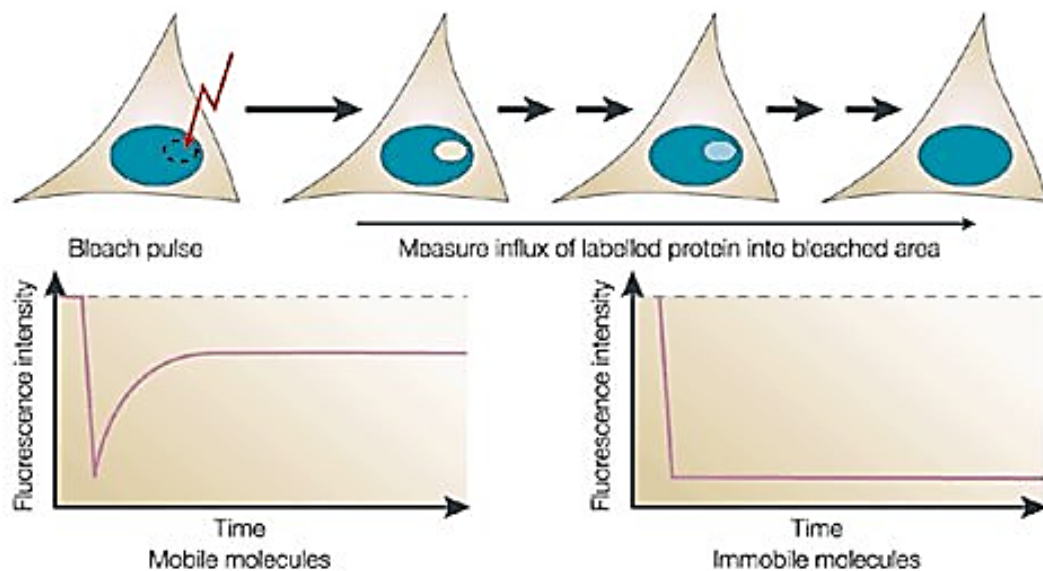


Figure 11: Schéma explicatif de la méthode FRAP (d'après Phair and Misteli 2001)

Sur la base de cette méthode, il a ainsi été montré que les facteurs de transcription sont très mobiles au sein du noyaux et se déplacent dans les trois dimensions (Phair et al. 2004; Sprague et al. 2004; Stenoien et al. 2001). Ce phénomène de diffusion est aléatoire et passif. Etant donné ces déplacements aléatoires, les protéines nucléaires rencontrent régulièrement des obstacles tels que d'autres protéines ou la chromatine. Il a été observé que les facteurs de transcription se lient très souvent et de façon transitoire et non spécifique sur la chromatine jusqu'à ce qu'ils « trouvent » leur site de liaison spécifique. Une fois liées à leur site spécifique, le temps de résidence devient plus long. La structure de la chromatine (positionnement des nucléosomes, boucles) permet de faciliter cette recherche des sites spécifiques en modulant l'accessibilité. Les expériences de FRAP ont également permis de mesurer le temps de liaison des facteurs de transcription. Il apparaît que ce temps est plutôt court, entre quelques millisecondes et environ 100 secondes (Farla et al. 2004; Hinow et al. 2006). Néanmoins, la mesure de ce temps est dépendante de modélisations mathématiques permettant l'interprétation des données de FRAP. Ces modèles, à l'heure actuelle toujours en développement, ne permettent pas une très grande précision. Une autre méthode

permet de mesurer la dynamique de liaison d'une protéine à l'ADN *in vivo*. Il s'agit de la méthode du ChIP compétitif (Figure 11). Pour cela, la protéine d'intérêt est présente en deux copies, chacune portant une étiquette différente. L'une des copie est exprimée de façon constitutive, l'autre de façon inducible. Après induction, on réalise une expérience de ChIP contre les deux étiquettes à différents temps. Cela permet d'analyser la compétition de liaison des deux versions et donc de mesurer le temps moyen de remplacement et d'occupation de la protéine d'interêt sur un *locus* particulier (Lickwar, Mueller, and Lieb 2013).

General strain design and characterization

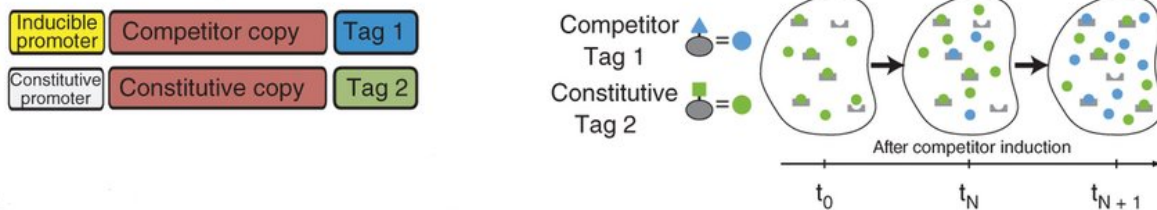


Figure 12: Schéma explicatif de la méthode du ChIP compétitif (d'après Lickwar, Mueller, and Lieb 2013)

Le ChIP compétitif a été utilisé pour analyser la dynamique de certains facteurs de transcription tels que Gal4 et Rap1 chez la levure *S. cerevisiae* (Nalley, Johnston, and Kodadek 2006; Lickwar et al. 2012). Cette méthode a permis de montrer que, sur certains *loci*, les facteurs peuvent présenter des temps de résidence assez longs (500s en moyenne). Dans le cas de Rap1, il existe une corrélation positive entre les temps de résidence et l'expression des gènes et négative avec l'occupation des nucléosomes, suggérant une compétition entre Rap1 et les nucléosomes. Ces observations sont cohérentes avec le fait que la structure de la chromatine influence la recherche des sites de liaison par les facteurs de transcription.

La formation du PIC a été également étudiée par des expériences de FRAP. Les mobilités de la TBP, de TFIIB, de Taf1 et de la Pol II ont été examinées (Sprouse et al. 2008). Les courbes de réapparition de fluorescence sont différentes pour ces facteurs. Cela suggère qu'il n'y a pas de pré-assemblage des GTFs avant la liaison au promoteur et qu'ils se présentent au promoteur à des temps différents et forment aléatoirement des PICs complets lorsque tous sont présents en même temps. Ce comportement est également suggéré dans une étude chez l'Homme par FRAP de la dynamique de la Pol II (Darzacq et al. 2007). Dans cette étude, les résultats ont conduit à proposer un modèle selon lequel seule 1/90 des ARN polymérases II arrive à former un PIC compétent pour l'élongation. Il semble que, *in vivo*, le PIC soit formé par des collisions aléatoires des différents GTFs et de la Pol II au promoteur et que cela soit intrinsèquement inefficace. L'ensemble de ces données montre que l'assemblage du PIC est un processus stochastique, c'est à dire aléatoire. Le temps de résidence individuel d'un facteur peut, s'il est régulé, améliorer la probabilité de formation d'un PIC complet. Ainsi, par exemple, la présence de structures de la chromatine favorable peut améliorer la liaison d'un facteur et ainsi augmenter la probabilité du recrutement des autres facteurs. Ces expériences permettent d'entrevoir comment la dynamique peut aussi agir sur la régulation spécifique de l'expression d'un gène.

Une méthode d'imagerie permettant la visualisation de la synthèse d'un ARN de façon dynamique et en temps réel a également été développée. Cette méthode repose sur l'insertion au sein de l'ARN transcrit de séquence appelées MS2 ou PP7, issues de bactériophages, permettant la formation de boucles type épingle à cheveux (*hairpins*). Des protéines de ces mêmes bactériophages capables de se lier à ces boucles sont fusionnées à un rapporteur fluorescent type GFP et exprimées de façon constitutives (Lenstra and Larson 2016). Lorsque l'ARN hybride est synthétisé, les protéines fluorescentes se lient à celui-ci et on peut alors observer la transcription en temps réel (Figure 13).

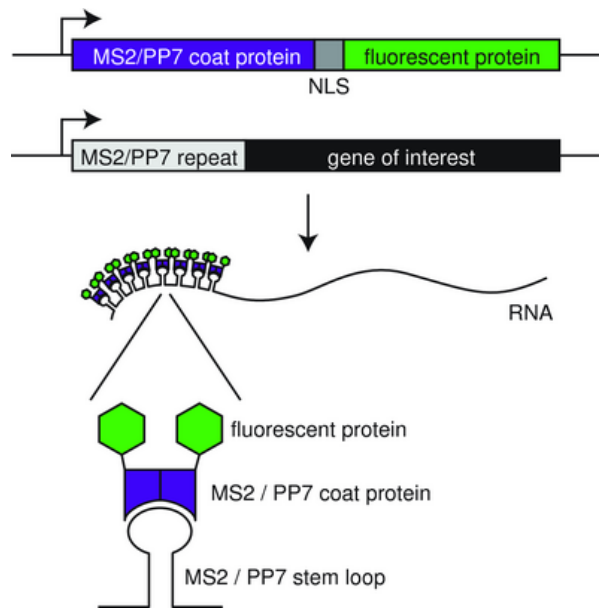


Figure 13: Schéma représentant la méthode de marquage MS2/PP7 des ARNs (D'après Lenstra and Larson 2016)

A l'aide de cette méthode, il a été possible de mesurer *in vivo* la vitesse d'élongation de la transcription. Ainsi, chez la levure *S. cerevisiae*, il a été calculé une vitesse d'élongation située entre 0,84 kb/min et 3,66 kb/min (Hocine et al. 2013). Cependant, les auteurs ont constaté une forte variabilité de ces vitesses entre les cellules. D'autres études ont mesuré des valeurs de vitesses d'élongation différentes sur d'autres gènes. Cette grande variabilité peut être indicative de vitesse réellement différente mais peut aussi refléter la complexité des systèmes de mesures. De ce fait, les méthodes d'analyses en cellules uniques permettent d'analyser la variabilité des vitesses de transcription et son donc d'un intérêt grandissant (Lenstra et al. 2016).

Enfin, l'utilisation de ces méthodes d'analyses sur des cellules uniques a permis de constater que, dans une même population, on retrouve des niveaux de transcription extrêmement variables. On parle de bruit génétique (Blake et al. 2003; Raj et al. 2006). Il semble que cette hétérogénéité d'expression soit due au fait que la transcription se déroule sous forme de pulsation (*burst*), durant lesquelles plusieurs initiations productives se succèdent rapidement. Cette notion de *bursting* est cohérente avec l'hypothèse selon laquelle l'initiation de la transcription soit un phénomène stochastique.

L'analyse de la dynamique de la transcription est un champ d'investigation assez récent. A l'heure actuelle, les méthodes utilisées sont réalisées *in vivo* mais avec des systèmes encore artificiels (ajout de protéines étiquetées, multiples répétitions des motifs PP7/MS2) pouvant induire des biais. Elles feront probablement l'objet de nouveaux développements technologiques et améliorations. Elles ont cependant permis d'appréhender la transcription d'un point de vue dynamique.

III. La transcription pervasive.

Les récents progrès des méthodes d'analyse par séquençage à haut-débit ont permis l'investigation profonde des transcriptomes (l'ensemble des transcrits produits par une cellule). Les études précédentes se limitaient à l'analyse des transcrits correspondants aux gènes connus et annotés. Il était connu depuis plusieurs années que les ARN polymérases pouvaient transcrire des ARNs non codant (ARNnc) stables tels que les snARN (*small nuclear RNAs*), les snoRNA (*small nucleolar RNA*) et les ARNt (ARN de transfert). Plus tard, avec le développement des technologies, il a été découvert que des transcrits correspondants à des régions intergéniques supposées silencieuses existent dans les cellules. L'idée de transcription pervasive a ainsi été proposée. Ce terme de transcription pervasive se réfère au fait que la transcription peut intervenir partout sur le génome sans correspondre nécessairement à un gène classique et annoté (Jacquier 2009). Ainsi, lorsque l'on analyse le transcriptome de *S. cerevisiae*, on observe qu'environ 85% du génome est transcrit (David et al. 2006; David, Clauder-Münster, and Steinmetz 2014). Bien que pas totalement comprise à ce jour, cette transcription pervasive semble avoir un rôle dans le fonctionnement des cellules et offre un champ d'investigation pour la compréhension de la transcription (Jensen, Jacquier, and Libri 2013).

Ces transcrits peuvent provenir de régions intergéniques (Nagalakshmi et al. 2008) ou bien du corps des gènes, dans les régions codantes. Les transcrits provenant des séquences codantes peuvent être produits dans le sens du gène ou de façon anti-sens (Carninci, Kasukawa, et al., FANTOM Consortium and RIKEN Genome Exploration Research Group and Genome Science Group (Genome Network Project Core Group) 2005; Katayama, Tomaru, Kasukawa, et al., RIKEN Genome Exploration Research Group, Genome Science Group (Genome Network Project Core Group), and FANTOM Consortium 2005). Les transcrits anti-sens ont été mis en évidence chez la levure en étudiant des mutants des facteurs d'élongation Spt6 et Spt16. Ces protéines sont nécessaires au rétablissement de l'organisation des nucléosomes après le passage de la Pol II au cours de l'élongation. Lorsque ceux-ci sont déficients, l'ADN pauvre en nucléosomes permet l'assemblage d'un PIC sur des promoteurs dits cryptiques au cœur de la région codante, aboutissant à une transcription parasite intragénique (Kaplan, Laprade, and Winston 2003).

Plus tard, des études de transcriptome de souches de levures déficientes pour certaines voies de dégradation des ARNs ont ajouté des preuves supplémentaires à l'existence de la transcription pervasive (Jensen, Jacquier, and Libri 2013). Ainsi, la délétion de Rrp6, la sous-unité catalytique 3'-5' exonucléase du complexe exosome de la levure, provoque une accumulation de transcrit cryptiques appelés CUTs (*cryptic unstable transcripts*). Ces transcrits sont dégradés rapidement en temps normal et ne sont pas présents de façon suffisantes dans les conditions sauvages pour être détectés (Wyers et al. 2005; Davis and Ares 2006; Houalla et al. 2006). Des analyses globales ultérieures des CUTs (Neil et al. 2009; Xu et al. 2009) ont montré que certains d'entre eux étaient moins sensibles à la dégradation par Rrp6 et ont été nommés SUTs (*stable unannotated transcript*). Enfin, la délétion d'une autre exoribonucléase appelée Xrn1 a permis la mise en évidence d'un troisième type de transcrits cryptiques appelés XUTs (*Xrn1-sensitive unstable transcript*) (van Dijk et al. 2011).

La transcription pervasive existe également chez les mammifères. Les analyses à grande échelle telles que celles menées dans le projet ENCODE (*ENCyclopedia Of DNA Element*¹) ont montrées qu'environ 75% du génome est transcrit (Djebali et al. 2012). Dans le cas des mammifères, les ARNs non codants sont séparés en deux catégories selon leur taille. Les ARNs de plus de 200nt sont ainsi appelés longs ARN non-codants (lncRNA) (Jensen, Jacquier, and Libri 2013). Leur transcription est généralement associée à la transcription des ARNm. Les petits transcrits pervasifs sont quant à eux issus de la transcription des régions 5' et 3' des gènes et sont appelés PASRs

¹ ENCODE est un projet international de recherche publique lancé en 2003 proposant de définir les éléments fonctionnels du génome humain afin de comprendre le rôle de l'ADN non codant, majoritaire dans le génome.

(*promoter-associated small-RNA*) et TARS (*terminator-associated small-RNA*) (Kapranov et al. 2007). Ces ARNs semblent être des ARNs naissants dont la transcription a cessé rapidement et dont la dégradation est protégée par la Pol II pausée près du TSS. Des transcrits divergents appelés PROMPTs (*PROMoter upstream Transcripts*) ont également été identifiés chez les mammifères (Preker et al. 2008). Tout comme les CUTs chez la levure, ils sont sensibles à la dégradation donc comparables. Les lncRNAs sont eux plutôt comparables aux XUTs et SUTs.

L'existence de cette transcription pervasive soulève un grand nombre de questions. Comment se déroule-t-elle et comment est-elle contrôlée ? Présente-t-elle un rôle fonctionnel ?

Concernant les mécanismes de transcription des ARN cryptiques chez la levure, il a été montré qu'ils sont transcrits par la Pol II. Les CUTs et SUTs sont majoritairement transcrits à partir de séquences se trouvant dans des régions dépourvues de nucléosomes (*nucleosomes depleted regions*, NDRs) situées en 5' ou en 3' des gènes transcrits. De façon surprenante, un grand nombre d'entre eux proviennent d'une transcription divergente (Jensen, Jacquier, and Libri 2013). Cette observation indique que les promoteurs des gènes sont intrinsèquement bidirectionnels. Leur polarité pour la transcription d'un ARNm est due à l'instabilité d'un des deux transcrits dont l'élongation est rapidement stoppée et suivie de sa dégradation (Core, Waterfall, and Lis 2008; Preker et al. 2008; Schulz et al. 2013; Seila et al. 2008). La transcription de ces ARNs divergents repose sur l'assemblage de deux PICs différents. En effet, la mutation d'une boîte TATA d'un gène perturbe la transcription de ce gène mais augmente également la transcription divergente, indiquant la compétition entre la formation des deux PICs (Neil et al. 2009). L'analyse de la distribution des GTFs sur l'ensemble du génome révèle également la présence de deux PICs sur les promoteurs à partir desquels les CUTs sont transcrits (Rhee and Pugh 2012). Néanmoins, il semble que les deux PICs formés dans ces cas soient régulés de la même manière, par le même activateur (Neil et al. 2009).

Une hypothèse concernant la possibilité pour une cellule de transcrire ces ARNs non codants divergents est que la Pol II fonctionne selon un mécanisme d'essai/erreur. Elle va initier une transcription qui peut être non fonctionnelle avant de transcrire un ARN correct (Berretta and Morillon 2009). Néanmoins plusieurs études proposent une fonction à cette transcription pervasive. Il a été observé que les conditions de croissance ou le stade de développement d'un organisme pouvaient induire un changement du profil d'expression des XUTs, laissant imaginer un rôle fonctionnel (van Dijk et al. 2011; Stolc et al. 2004). La transcription pervasive pourrait avoir un effet régulateur sur l'expression de certains gènes en agissant par interférence transcriptionnelle. C'est, par exemple, le cas pour l'ARNnc *SRG1*, dont la transcription interfère avec l'assemblage du PIC sur le promoteur adjacent du gène *SER3* réprimant ainsi sa transcription (Martens, Laprade, and Winston 2004). Enfin, il est proposé que la transcription pervasive ait un intérêt évolutif. En effet, la transcription d'ARNnc pourrait permettre une alternative aux mutations ponctuelles dans la sélection naturelle. Les transcrits cryptiques pourraient ainsi devenir des gènes potentiels ou trouver une fonction. Certains ARNnc sont traduits en peptides qui pourraient, d'un point de vue évolutif, procurer un avantage pour la cellule qui l'exprime (Tudek, Candelli, and Libri 2015; Ingolia et al. 2014).

En dépit de ces quelques exemples et des hypothèses avancées quant au rôle fonctionnel de la transcription pervasive, celle-ci nécessite un contrôle très strict. En effet, elle peut avoir un rôle délétère si elle devient trop importante. Elle peut avoir pour conséquence une interférence indésirable sur la transcription, perturbant ainsi l'expression des gènes. Une activité transcriptionnelle trop élevée pourrait avoir un effet de titration sur la Pol II, les NTPs et les protéines impliquées dans le transport des ARNs, dommageable pour la transcription des gènes qui doivent l'être. La transcription pervasive pourrait également mener à une instabilité du génome (Preker et al. 2008). Pour toutes ces raisons, la transcription pervasive est contrôlée. Le premier point de contrôle est l'initiation de cette transcription. Ainsi, des facteurs de remodelage de la chromatine tels que Spt6 et FACT ou Iswi2 et l'histone déacétylase Rpd3 (Kaplan, Laprade, and Winston 2003) contrôlent le niveau de nucléosomes afin d'éviter une initiation non désirée. Le second niveau de contrôle concerne la stabilité de ces transcrits. Dans un grand nombre de cas,

la cellule peut faire appel à un mécanisme de terminaison couplé à la dégradation. Chez la levure, il s'agit de la voie de terminaison NNS, présentée plus haut.

IV. La transcription dans le contexte chromatinien.

Dans les cellules eucaryotes, la double hélice d'ADN est retrouvée associée à des complexes protéiques formant une structure appelée chromatine (Figure 14). La fonction première de cette organisation est la compaction des très longues chaînes d'acides nucléiques afin de réduire drastiquement l'espace qu'elles occupent. La chromatine sert également à la régulation d'autres processus impliquant l'ADN tels que la transcription, la réplication, la recombinaison et la réparation en modulant l'accès à l'ADN. Il est donc important de connaître la biologie de cette organisation chromatinienne puisque c'est dans ce contexte que se déroule la transcription *in vivo*.

La chromatine est composée de nucléosomes, des complexes multi-protéiques formés d'histones. Il existe chez les eucaryotes 5 histones différentes, nommées H1, H2A, H2B, H3 et H4. Un nucléosome est formé de 8 histones organisées en un tétramère (H3-H4)₂ et un tétramère (H2A-H2B)₂. 146 paires de base d'ADN sont enroulées autour d'un nucléosome. L'histone H1 est quant à elle liée aux nucléosomes afin de stabiliser leur structure et moduler le niveau de compaction de la chromatine en rapprochant plusieurs nucléosomes (Kornberg and Lorch 1999). La longue chaîne d'ADN est enroulé autour de cette succession de nucléosomes souvent comparée à un collier de perle.

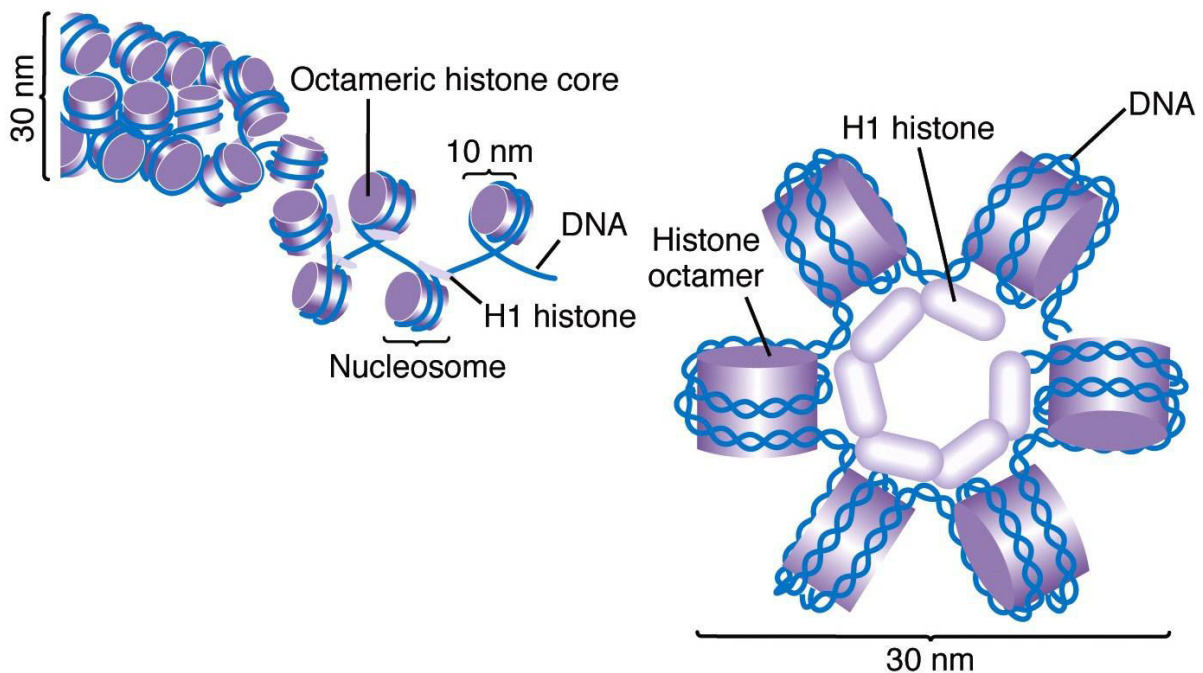


Figure 14: Schéma représentant l'organisation de l'ADN sous forme de chromatine. (d'après Griffiths et al., 2004, <http://www.mun.ca/>).

1) Structure de la chromatine.

La présence de la structure chromatinienne limite fortement l'accès à l'ADN pour les protéines, en particulier celles impliquées dans la transcription. Cependant, la chromatine peut exister sous plusieurs formes en fonction de son niveau de compaction. Lorsque la compaction est plus faible, on parle d'euchromatine et au contraire, d'hétérochromatine lorsque le niveau de compaction est grand. Les régions d'euchromatine peu denses en nucléosomes sont plus favorables à l'accès à l'ADN et contiennent préférentiellement les promoteurs (Lee et al. 2004; Sekinger, Moqtaderi, and Struhl 2005).

Des études ont été menées afin de déterminer la position génomique des nucléosomes. Ces travaux reposent sur l'utilisation du ChIP sur des nucléosomes isolés. La chromatine est traitée par une nucléase bactérienne (nucléase micrococciale, MNase) permettant la digestion de l'ADN non protégé par les nucléosomes. Le laboratoire de F. Pugh a ainsi compilé et analysé les données obtenues par plusieurs laboratoires pour définir une cartographie des nucléosomes sur le génome de *S. cerevisiae* (Jiang and Pugh 2009). Ils ont également utilisé la méthode de ChIP-exo pour analyser cette distribution (Rhee et al. 2014). Des études similaires ont été faites chez l'Homme (Valouev et al. 2011; Schones et al. 2008). Ces études ont révélé une organisation générale de la chromatine sur le génome. Au niveau du promoteur, on retrouve une région dépourvue de nucléosomes (*nucleosome-depleted region*, NDR) d'environ 200pb. Cette région est flanquée en 5' par le nucléosome -1 et en 3' par le nucléosome +1 (Jiang and Pugh 2009; Yuan et al. 2005). Entre ces deux nucléosomes se trouve le TSS. À la suite du nucléosome +1 on retrouve un enchaînement régulier de nucléosomes séparés par une distance relativement régulière et homogène, on parle de nucléosomes en phase. Lorsque l'on s'éloigne du début du gène, cette distribution devient plus hétérogène et aléatoire, on parle alors de distribution floue (*fuzzy*). À la fin du gène, on retrouve une nouvelle NDR, pouvant être liée au processus de terminaison de la transcription ou au promoteur du gène suivant.

La présence de la NDR en 5' du gène a un rôle très important pour la transcription puisque c'est sur cette partie du gène que le PIC se forme. Il semblerait que l'absence de nucléosomes sur les promoteurs soit favorisée par la présence de plusieurs adénine et thymine, enrichies dans ces régions (Yuan et al. 2005; Mavrigh et al. 2008; Anderson and Widom 2001). L'intervention de protéines remodelleuses de la chromatine est également responsable du maintien et de la modulation des NDR. Comme évoqué précédemment (Chapitre I.II.1.a), le nucléosome +1 pourrait avoir un rôle important dans la mise en place du PIC et dans la détermination du TSS des gènes (Rhee and Pugh 2012). Les gènes sans boîte TATA canonique recourant préférentiellement à TFIID vont utiliser ce nucléosome +1 pour restreindre la position du TSS. Les gènes dépendants plutôt de SAGA vont eux préférentiellement favoriser l'éviction du nucléosome +1. Une étude a d'ailleurs été menée pour étudier la dynamique des nucléosomes chez *S. cerevisiae* par une approche de ChIP compétitif (Dion et al. 2007). Ces travaux ont permis de montrer que les nucléosomes au niveau des régions promotrices sont beaucoup plus souvent renouvelés que dans les régions codantes. De plus, sur les régions promotrices de certains gènes, les nucléosomes +1 et -1 sont particulièrement dynamiques et présentent des niveaux de renouvellement élevés. Les auteurs parlent de nucléosomes *hot*. Ces taux de renouvellements pourraient être une façon d'apporter de la flexibilité dans les régions promotrices afin de rendre ces gènes plus souples à exprimer.

2) La modification des histones.

La structure du nucléosome a été résolue par cristallographie et diffraction des rayons X en 1997 (Luger et al. 1997). Le premier éclairage obtenu par cette structure a été de voir comment la molécule d'ADN, dont le squelette de phosphate est globalement chargé négativement, pouvait être compacté en dépit des répulsions électrostatiques au sein de la double hélice. Les histones sont globalement basiques et, au pH physiologique, chargées positivement. La structure montre comment les charges électrostatiques expliquent les interactions entre les histones et l'ADN. Une seconde observation très intéressante est le fait que les histones sont principalement trouvées sous forme globulaire et structurées à l'exception de leur extrémités. Ces extrémités, appelées queues, ne sont pas structurées et présentent la caractéristique de passer au travers des tours fait par l'ADN et d'être ainsi exposées à l'extérieur du nucléosome. Cette exposition les rend sujettes à des possibles modifications *post*-traductionnelles ainsi qu'à des interactions avec d'autres protéines (Kouzarides 2007). Ainsi, il est établi que les histones peuvent être méthylées, acétylées, SUMOylées, ubiquitinylées, ADP-ribosylées et phosphorylées (Tableau 4). On parle de marques d'histones. La séquence de ces queues d'histones pourrait permettre la modification de 60 résidus. En tenant compte du fait que la méthylation des lysines ou des arginines peut être multiple (di- ou tri-méthylation), le nombre de combinaisons de modifications possible est gigantesque. Bien qu'il faille noter que ces modifications n'interviennent pas toutes en même temps, cela laisse entrevoir la complexité et l'énorme potentiel régulateur qu'elles peuvent procurer. En effet, ces marques d'histones peuvent permettre le recrutement de protéines effectrices sur des régions spécifiques du génome et moduler leur fonction.

Les modifications des histones vont permettre la modification de leurs propriétés physico-chimiques et ainsi moduler la fonction de la chromatine. Ainsi, la phosphorylation et l'acétylation vont, par exemple, neutraliser une partie de la charge positive de l'histone et entraîner un relâchement de l'interaction entre l'ADN et les histones. L'ajout de marques d'histone permet également de moduler les interactions et le recrutement de protéines impliquées dans les processus nucléaires.

Comme indiqué plus haut, la chromatine existe sous forme d'hétérochromatine et d'euchromatine. Ainsi, les zones d'hétérochromatine, plutôt inactives, sont associées à de faibles niveaux d'acétylation mais à des hauts niveaux de méthylations de certains résidus (H3K9, H3K27 et H4K20) (Kouzarides 2007).

Les modifications d'histones peuvent être associées au processus de transcription. Certaines sont associées à une transcription active comme l'acétylation des histones H3/H4 et les méthylations des lysines K4 et K36 de l'histone H3 (Li, Carey, and Workman 2007). Les modifications menant à une répression sont celles de l'hétérochromatine. Il est également montré que les différents éléments d'un gène présentent des marques d'histones différentes. Les promoteurs sont plutôt enrichis en H3K4me3, H3K9ac et H3K14ac alors que les régions codantes des gènes transcrits sont caractérisées par la marque H3K36me3.

<i>Modification</i>	<i>Histone</i>	<i>Position</i>	<i>Enzyme</i>	<i>Rôle dans la transcription</i>	
<i>Méthylation</i>			<i>S. cerevisiae</i>	Mammifères	
	H3	K4	Set1	MLL, ALL-1 Set9/7 ALR-1/2 ALR, Set1	Activation
		K9	n/a	Suv39h, G9a Eu HMTase I ESET SETBD1	Répression/Activation
		K27		E(Z)	Répression
		K36	Set2	HYPB Smyd2 NSD1	Recrutement de Rpd3S pour réprimer l'initiation interne
		K79	Dot1	Dot1L	Activation
	H4	K20		PR-Set7 Set8	<i>silencing</i>
<i>Arg Méthylation</i>	H3	R2		CARM1	Activation
		R17		CARM1	Activation
		R26		CARM1	Activation
<i>Phosphorylation</i>	H4	R3		PRMT1	Activation
	H3	S10	Snf1		Activation
<i>Ubiquitination</i>	H2B	K120/123	Rad6/Bre1	UbcH6 RNF20/40	Activation
	H2A	K119		hPRC1L	Repression
<i>Acétylation</i>	H3	K56			Activation
	H4	K16	Sas2/NuA4	hMOF	Activation
	Htz1	K14	NuA4,SAGA		Activation

Tableau 4: Tableau récapitulatif des modifications d'histones. (D'après Li, Carey, and Workman 2007)

Les effets connus sur l'activité transcriptionnelle sont indiqué dans la colonne « Rôle dans la transcription ».

L'effet que procurent les modifications d'histones sur la transcription n'est pas forcément connu de façon stricte mais plusieurs mécanismes peuvent être proposés. Certaines modifications pourraient, par exemple, favoriser l'éviction du nucléosome +1 afin de faciliter l'assemblage du PIC et l'initiation de la transcription (Li, Carey, and Workman 2007). D'autres modifications peuvent favoriser le recrutement de protéines modulant la structure de la chromatine (remodeleurs). Enfin, la modification d'histones sur la région transcrite peut permettre de limiter l'accès à l'ADN pour limiter la transcription pervasive, par exemple.

3) Les variants d'histones.

Les histones présentées plus haut ont longtemps été considérées comme universelles. Cependant, il est apparu que certaines formes variantes de ces protéines existent (Kamakaka and Biggins 2005)(Tableau 5). La différence majeure entre ces variants et les histones canoniques est que celles-ci ne sont exprimées et incorporées au sein de la chromatine qu'en dehors de la phase S du cycle cellulaire. C'est-à-dire en dehors de la phase de réplication de l'ADN. Les variations de séquence avec les histones canoniques peuvent se trouver dans la queue de celles-ci mais aussi dans le domaine globulaire (Li, Carey, and Workman 2007). L'incorporation de ces variants au sein de la chromatine va modifier sa structure ou la possibilité de modification et ainsi impacter la transcription.

<i>Histone</i>	<i>Variant</i>	<i>Rôle sur la transcription</i>	<i>Localisation</i>
<i>H3</i>	H3.3	Activation de la transcription	Région transcrite
<i>H2A</i>	macroH2A	Inactivation du chromosome X	Chromosome X inactivé
	H2AZ	Activation/ répression de la transcription	Promoteur, hétérochromatine
	H2ABbd	Activation de la transcription	Chromosome X actif et autosomes
	H2A.X	Répression	Canonique, distribué sur tout le génome

Tableau 5: Tableau récapitulatif des différents variant d'histones, leur localisation et leur effet sur la transcription. (D'après Li, Carey, and Workman 2007).

4) Le remodelage de la chromatine.

Les modifications d'histones ne sont pas les seuls facteurs à pouvoir moduler la structure de la chromatine. Certains complexes protéiques peuvent agir sur cette structure en modifiant la position des nucléosomes sur l'ADN ou encore les interactions entre les nucléosomes et l'ADN. On parle de remodeleurs de la chromatine. Ces complexes utilisent l'ATP comme source énergétique pour modifier les contacts entre l'ADN et les histones et ainsi favoriser leur déplacement sur l'ADN ou bien même leur éviction (Vignali et al. 2000)

Ils permettent donc de moduler l'accessibilité de l'ADN aux protéines liées à la transcription. Cette activité biologique est retrouvée dans des complexes co-activateurs. Pour cette raison, ils feront l'objet d'une description plus détaillée dans la partie de ce chapitre consacrée aux co-activateurs.

V. La régulation de la transcription par l'ARN polymérase II.

La transcription des gènes est l'un des processus cellulaires les plus régulés. En effet, la modulation de l'expression des gènes en fonction de l'environnement ou du type cellulaire implique un contrôle très fin. La transcription n'est bien entendu pas le seul processus régulant le niveau d'une fonction biologique. La stabilité de l'ARN, la traduction de la protéine, sa stabilité ou son activité sont elles aussi sujettes à des régulations. Néanmoins, transcription, en général, et son initiation en particulier, constitue le premier point de contrôle.

La régulation de la transcription peut intervenir à différents niveaux, au cours des étapes du cycle de la transcription. Cette partie se concentrera sur la régulation de l'initiation de la transcription (Figure 15), bien que d'autres régulations existent.

Chez la levure *S. cerevisiae*, les signaux extra- et intracellulaires qui aboutissent à une réponse transcriptionnelle vont permettre le recrutement de protéines sur des séquences régulatrices activatrices (*upstream activating sequences*, UAS) ou répressives (*upstream repressing sequences*, URS). Ces séquences sont des éléments des promoteurs situés à distance du promoteur cœur. Les protéines qui y sont recrutées s'appellent facteurs de transcription (*transcription factor*, TFs) et peuvent être facteurs d'activation (*transcription activator*, TAs) ou facteur répresseur (*transcription repressor*, TRs). Ces TFs vont interagir avec des co-activateurs ou co-represseurs afin de transmettre les signaux à la machinerie de transcription en stimulant ou diminuant son activité.

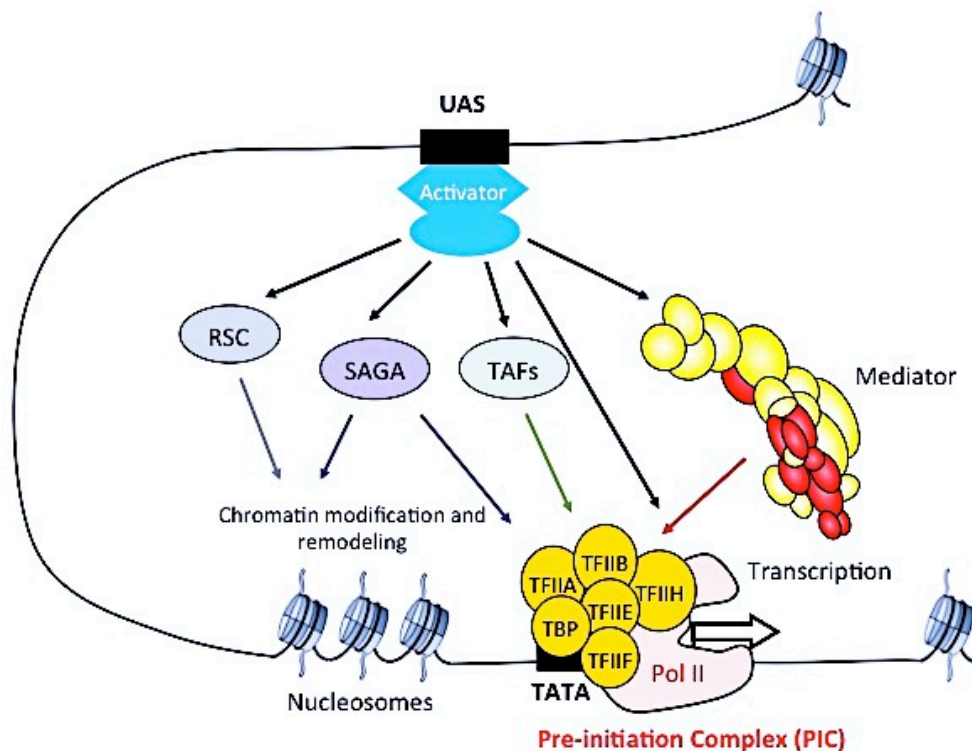


Figure 15 : Schéma représentant la régulation de l'initiation de la transcription par l'ARN polymérase II chez *S. cerevisiae*.

1) Les éléments de régulation des promoteurs.

Les éléments de régulation sont des séquences d'ADN permettant le recrutement et la liaison de protéines régulatrices. Bien que leur rôle soit similaire, leur nature est différente entre la levure et les eucaryotes pluricellulaires

Chez la levure, les éléments de régulation sont appelés UAS pour les séquences activatrices et URS pour les séquences répressives. On retrouve ces séquences quasi-exclusivement en amont du promoteur cœur et dans des régions dépourvues de nucléosomes (Hahn and Young 2011; Xue et al. 2004). De façon générale, ces U(A/R)S se situent à quelques centaines de paires de bases du TSS et perdent de leur efficacité en fonction de la distance (Dobi and Winston 2007).

Chez les eucaryotes pluricellulaires, les éléments de régulation activateurs sont appelés *enhancers*. Contrairement à la situation chez *S. cerevisiae*, ces séquences peuvent être situées en amont ou en aval de la région promotrice cœur (S. Hahn and Young 2011). Ils peuvent, par exemple, être situés au sein même du gène transcrit (Blackwood and Kadonaga 1998)

La différence majeure avec la levure réside dans la distance à laquelle les éléments régulateurs se trouvent du promoteur. En effet, ils sont en général situés à plusieurs kilo-bases du promoteur. Pour ne citer qu'un exemple, l'*enhancer* d'un gène de drosophile appelé *cut* se situe à 85kb du promoteur (Blackwood and Kadonaga 1998). Lorsque le facteur de transcription se lie à un élément de régulation, son rôle est de transmettre un signal, généralement par l'intermédiaire d'un co-activateur, à la machinerie de la Pol II situé au promoteur cœur. Etant donné les grandes distances qui les séparent, cela implique une organisation topologique permettant ces contacts. Cette organisation nécessite, par exemple, la formation de boucles de l'ADN rapprochant deux régions. Une méthode d'analyse de la conformation spatiale de l'ADN appelée 3C (*chromosome conformation capture*) a été développée il y a quelques années (Dekker et al. 2002). Cette méthode permet la mise en évidence d'interactions physiques entre deux *loci* au sein du noyau. Des expériences appliquant cette méthode ont permis la mise en évidence de telles boucles et de protéines impliquées dans leur formation (Jing et al. 2008; Vakoc et al. 2005). Les facteurs de transcription ne sont pas les seules protéines impliquées dans ce phénomène puisque des protéines architecturales de l'ADN telles que la cohésine sont requises pour moduler la topologie de la chromatine (Plank and Dean 2014). Les progrès des méthodes de séquençage à haut-débit ont apportés une forte contribution à la compréhension de ces mécanismes de conformation de l'ADN. En effet, la combinaison de la méthode de 3C avec le séquençage à haut débit a permis de développer la méthode de Hi-C. En utilisant cette approche, il est possible d'analyser la structure topologique en 3 dimensions au sein du noyau (Dixon et al. 2012; Lieberman-Aiden et al. 2009). L'analyse de ces données a permis la mise en évidence d'un grand nombre de contacts entre des éléments régulateurs et des promoteurs avec parfois des distances très longues. Cela suggère donc que l'organisation spatiale de la chromatine pourrait avoir un rôle important dans la régulation de l'expression des gènes.

Les *enhancers* contiennent des séquences permettant la liaison de facteurs de transcription. Souvent, ces régions contiennent plusieurs sites de liaison d'un ou plusieurs facteurs de transcription qui peuvent agir de façon coopérative ou redondante (Plank and Dean 2014). On retrouve sur ces régions des marques d'histones caractéristiques telles que des méthylations de l'histone H3 sur la lysine 4. Le laboratoire de R. Young a montré l'existence de régions particulières appelées *super-enhancers* (Lovén et al. 2013; Whyte et al. 2013). Ces régions, contiennent une forte occupation de facteurs de transcription et du co-activateur Médiateur, et contrôlent l'expression de plusieurs gènes cibles en même temps. Leur activité est très importante dans le contrôle du programme transcriptionnel au cours du développement.

Les séquences régulatrices, chez la levure ou chez les autres eucaryotes, consistent en une séquence d'ADN reconnue par un facteur de transcription. Les facteurs de transcription reconnaissent pour la plupart un motif consensus. Il est possible de définir ces motifs par l'analyse

des profils de liaison d'un facteur de transcription par les méthodes de CHIP-seq ou CHIP-exo ainsi que par des outils bioinformatiques permettant la prédiction de ces séquences (Starick et al. 2015; Bardet et al. 2013; Boeva 2016).

2) Les facteurs de transcription.

Les facteurs de transcription représentent une grande famille de protéines. En effet, la base de données YEASTRACT qui recense les protéines de cette famille chez la levure *S. cerevisiae* compte à ce jour près de 200 gènes codant pour des TFs (Teixeira et al. 2014). Cela fait de cette famille l'une des plus grande classe de protéines (Hahn and Young 2011). Chez les autres eucaryotes, la famille des facteurs de transcription est également nombreuse.

Bien que certains d'entre eux puissent trouver leur activité de facteur de transcription dans leur liaison avec un autre facteur de transcription, la caractéristique majeure des TFs est leur capacité à se lier à l'ADN en reconnaissant les éléments régulateurs. Pour cette raison, ils sont structurés autour d'un domaine de liaison à l'ADN (*DNA-binding domain*, DBD). Il existe plusieurs types de DBD qui vont permettre la classification des TFs en fonction de la nature de ce domaine (Hahn and Young 2011). On distingue ainsi trois classes principales de TFs : Les DBD stabilisés par un atome de zinc (*zinc-finger*, *GATA-finger*), les DBD structurés en Hélice-Tour-Hélice (HTH) et les DBD dits *leucine-zipper* (Luscombe et al. 2000; Hahn and Young 2011).

Comme expliqué précédemment, il est possible par les méthodes de CHIP et ses dérivées, d'analyser les interactions entre les protéines et l'ADN ainsi que la localisation de ces interactions. Il ressort de ces analyses que les TFs se lient à l'ADN de façon préférentielle dans des régions pauvres en nucléosomes (Badis et al. 2008) et qu'ils tendent à maintenir cet état. Ces méthodes permettent aussi d'analyser les sites de fixation des TFs et par la suite de déterminer les gènes qu'ils régulent ainsi que leurs potentiels partenaires (Zhu et al. 2009).

a) Les activateurs de la transcription.

Les activateurs de la transcription sont des facteurs de transcription dont le rôle est, comme leur nom l'indique, d'augmenter l'activité transcriptionnelle. Ils sont composés du DBD et d'au moins un domaine d'activation (Triezenberg 1995). Lorsqu'ils sont fixés sur l'UAS ou l'*enhancer*, ces activateurs vont avoir pour rôle le recrutement des co-activateurs. Par leur intermédiaire, ils vont pouvoir stimuler le recrutement de la machinerie de la transcription. Chez la levure, deux activateurs sont particulièrement étudiés. Il s'agit de Gal4 et Gcn4. Il a été montré qu'ils stimulent le recrutement de SAGA, NuA4, Swi/Snf, TFIID et du Médiateur (Fishburn, Mohibullah, and Hahn 2005; Reeves and Hahn 2005). De manière intéressante, un activateur peut stimuler l'expression de plusieurs gènes en même temps, formant ainsi un réseau de régulation. Le TF Hsf1 est un bon exemple puisqu'il va permettre la régulation simultanée de gènes de réponse à un stress en activant la transcription des gènes HSP (*heat shock protein*) (Hahn and Thiele 2004). Chez les mammifères, certains activateurs sont même proposés comme étant des régulateurs « maitres » (*master regulators*) pouvant réguler tout un ensemble de gènes associés à un programme transcriptionnel. C'est le cas, par exemple, d'Oct4 dans les cellules souches embryonnaires de souris. Celui-ci joue un rôle majeur dans le maintien de l'état de cellule souche en établissant un réseau régulateur sur un *super-enhancer* (Whyte et al. 2013).

L'activité de ces activateurs est elle aussi régulée. En effet, ils n'agissent pas continuellement sur leurs gènes cibles et sont contrôlés de différentes façons afin de permettre la réponse à des *stimuli* particuliers. Cette régulation peut intervenir au niveau de l'expression du TF, modulée en fonction de la physiologie (Hahn and Young 2011). La localisation cellulaire des TFs est aussi un moyen de régulation. En effet, certains TFs peuvent être séquestrés dans le cytoplasme et subir une translocation nucléaire en réponse à un signal (Komeili and O'Shea 2000). Les TFs sont également sujets à des modifications *post*-traductionnelles telles des phosphorylations ou des ubiquitinations pouvant réguler leur activité (Salghetti et al. 2001; Hahn and Young 2011). Ils sont en effet souvent cibles de voies de signalisation cellulaire qui aboutissent à ces modifications. Enfin, des changements de conformation soit par liaison avec un ligand soit par liaison avec un autre facteur permettent également la régulation des TFs. C'est le cas, par exemple, des récepteurs nucléaires dont l'activité dépend de leur liaison avec leur ligand.

b) Les répresseurs de la transcription.

Au contraire des activateurs de la transcription, certains TFs vont avoir un effet négatif sur l'expression de gènes. On définit un répresseur comme un TF dont la délétion ou la mutation a pour effet une augmentation de la transcription de leurs gènes cible. Ils peuvent agir de trois façons. (1) en limitant l'accès à l'ADN en renforçant la structure de la chromatine ou en perturbant l'assemblage du PIC. C'est le cas du complexe Ssn6-Tup1, conservé de la levure à l'Homme, qui interagit avec les histones (Smith and Johnson 2000). (2) en agissant sur la mise en place du PIC directement. C'est le cas des répresseurs généraux Mot1 et NC2 qui vont perturber la liaison de la TBP sur l'ADN, et ainsi déstabiliser la formation du PIC (Goppelt et al. 1996; Auble et al. 1994). (3) en interagissant directement avec des activateurs, bloquant ainsi leur liaison à l'ADN. Les répresseurs Hsp90 et Gal80 ont, par exemple, ce rôle en se liant avec Hsf1 et Gal4 (Sil et al. 1999; Zou et al. 1998). Il est également intéressant de noter que certains facteurs de transcription peuvent avoir un rôle à la fois activateur et répresseur. C'est le cas de Rap1 par exemple (Bendjennat and Weil 2008)

3) Les co-activateurs.

Les co-activateurs sont des complexes protéiques présents chez tous les eucaryotes. Ils agissent après la liaison à l'ADN des activateurs pour favoriser la transcription de leurs gènes cibles. Ils vont agir soit directement sur le PIC et sa formation soit en modifiant la structure de la chromatine pour favoriser un environnement propice (Näär, Lemon, and Tjian 2001). Contrairement aux GTFs, ils sont dispensables à l'activité basale de la Pol II *in vitro*. Ils sont cependant nécessaires aux mécanismes d'activation d'un grand nombre de gènes *in vivo*.

a) Les complexes agissant sur la mise en place du PIC.

Parmi les complexes co-activateurs agissant directement sur la formation du PIC, on retrouve TFIID et SAGA. Ils possèdent tous les deux cette activité grâce à la présence de TAFs, dont cinq sont communes aux deux co-activateurs.

Un troisième complexe agit aussi sur l'assemblage du PIC. Il s'agit du Médiateur, un complexe conservé chez l'ensemble des eucaryotes faisant un lien physique entre les activateurs et le PIC. Il fera l'objet d'une description détaillée dans le chapitre II de ce manuscrit.

- TFIID.

TFIID est composé de la TBP et de 14 TAFs (*TBP-associated factors*) (Tableau 6). Comme indiqué dans la présentation de TFIID, les TAFs ont un rôle dans la reconnaissance du promoteur en stimulant et stabilisant l'assemblage du complexe TBP-TATA. TFIID est majoritairement requis pour la transcription des gènes dépourvus de boîte TATA canonique. Les TAFs renforcent l'affinité de TFIID pour le promoteur en interagissant avec les éléments Inr et DPE, guidant ainsi le recrutement de la TBP sur l'élément *TATA-like*. Néanmoins, des études ont suggéré que les TAFs pourraient être impliqués dans la transcription activée en ayant un rôle adaptateur entre la machinerie de la Pol II et les activateurs (Albright and Tjian 2000; Näär, Lemon, and Tjian 2001). Il a été aussi montré par des études fonctionnelles que Taf4 était impliqué dans la régulation des gènes impliqués dans la différenciation cellulaire des hépatocytes ou des cellules souches embryonnaire (Alpern et al. 2014; Langer et al. 2016). Des travaux sur la protéine Taf7 ont montré son implication dans la transcription régulée par le facteur c-Jun (Munz et al. 2003). Ces études montrent donc que les TAFs peuvent contacter des activateurs spécifiques permettant de recruter TFIID. La composition en TAFs de TFIID chez la levure est constante. Ce n'est pas le cas chez les mammifères. En effet, sa composition peut varier selon le type cellulaire et ainsi déterminer un programme transcriptionnel défini par des interactions avec des activateurs particuliers (Kolthur-Seetharam, Martianov, and Davidson 2008). Enfin, Taf1 possède des activités enzymatiques histone acétyl-transférase (HAT) et kinase (Mizzen et al. 1996) ainsi que deux domaines de liaison aux histones acétylées (bromodomains). Ces deux caractéristiques sont communes avec les complexes co-activateurs remodelateurs de la chromatine.

<i>Nom</i>	<i>H. sapiens</i>	<i>S. cerevisiae</i>
<i>TAF1</i>	TAFII250	Taf145/130
<i>TAF2</i>	TAFII150	Taf150 ou TSM1
<i>TAF3</i>	TAFII140	Taf47
<i>TAF4</i>	TAFII130/135	Taf48 ou MPT1
<i>TAF4b</i>	TAFII105	
<i>TAF5</i>	TAFII100	Taf90
<i>TAF5L</i>	PAF65b	
<i>TAF6</i>	TAFII80	Taf60
<i>TAF6L</i>	PAF65	
<i>TAF7</i>	TAFII55	Taf67
<i>TAF7L</i>	TAF2Q	
<i>TAF8</i>	(BAB71460)	Taf65
<i>TAF9</i>	TAFII32/31	Taf17
<i>TAF9L</i>	TAFII31L	
<i>TAF10</i>	TAFII30	Taf25
<i>TAF11</i>	TAFII28	Taf40
<i>TAF12</i>	TAFII20/15	Taf61/68
<i>TAF13</i>	TAFII18	Taf19 ou FUN81
<i>TAF14</i>		Taf30
<i>TAF15</i>	TAFII68	

Tableau 6 : Nomenclature unifiée pour les TAFs (Tora 2002).

Les TAFs communes aux complexes TFIID et SAGA sont indiquées en rouge.

- SAGA.

Le complexe SAGA est particulièrement impliqué dans la régulation de la transcription fortement régulée des gènes impliqués dans la réponse au stress et contenant majoritairement une boîte TATA canonique. Il semble cependant que SAGA soit impliqué dans la régulation de l'ensemble des gènes (Bonnet et al. 2014). En plus de son rôle dans le recrutement de la TBP au promoteur, SAGA possède d'autres fonctions. En effet, grâce à des activités enzymatiques, ce complexe agit sur la chromatine afin d'en modifier les propriétés. Il est composé de 21 sous-unités (Tableau 6) largement conservées des familles Ada, Spt et TAF (Rodríguez-Navarro 2009).

Ce complexe est organisé en cinq modules. Le module Ada contient les protéines du même nom. Ces protéines ont été identifiées initialement dans un crible d'interaction génétique avec l'activateur Gcn4 et l'activateur chimérique Gal4-VP16 comme étant des protéines adaptatrices en favorisant la réponse à ces activateurs (Marcus et al. 1994). Ceci indique que le complexe peut être directement recruté sur le promoteur par des activateurs. Le module HAT (histone acétyltransférase) contient la sous-unité Gcn5 permettant l'acétylation de l'histone H3. Cette modification a pour effet une ouverture de la chromatine, favorisant la transcription (Owen et al. 2000). Le module DUB (déubiquitynase) quant à lui contient la sous-unité Ubp8 (USP22 chez l'Homme) qui va catalyser le retrait de l'ubiquitine sur l'histone H2B. Le retrait de cette marque d'histone augmente l'activation de la transcription (Daniel et al. 2004). Enfin, les modules Spt et TAF, contenant les protéines du même nom, vont avoir un rôle dans la stabilité de SAGA mais aussi dans son recrutement au promoteur. En effet, les protéines Spt3 et Spt8 interagissent avec la TBP

et favorisent la formation du PIC (Sermwittayawong and Tan 2006; Mohibullah and Hahn 2008; Laprade, Rose, and Winston 2007). Les TAFs jouent ce rôle en reconnaissant les éléments du promoteur. Enfin, le module TAF contient la sous-unité Tra1 présentant la capacité d'interagir avec des activateurs pour stimuler le recrutement de SAGA (Brown et al. 2001).

<i>Modules de SAGA</i>	<i>S. cerevisiae</i>	<i>H. sapiens</i>
<i>HAT</i>	Gcn5	GCN5
<i>Ada</i>	Ada1	ADA1
	Ada2	ADA2
	Ada3	ADA3
<i>Spt</i>	Spt8	ND
	Spt20	SPT20
	Spt7	SPT7
	Spt3	SPT3
<i>Taf</i>	Taf5	TAF5L
	Taf6	TAF6L
	Taf9	TAF9
	Taf10	TAF10
	Taf12	TAF12
	Tra1	TRRAP
	Sgf29	SGF29
<i>DUBm</i>	Ubp8	USP22
	Sgf11	ATXN7L3
	Sus1	ENY2
	Sgf73	ATXN7
	Chd1	ND
	Rtg2	ND

Tableau 7: Tableau récapitulatif des sous-unités de SAGA chez la levure et chez l'Homme (d'après Rodríguez-Navarro 2009).

De manière intéressante, le rôle de SAGA ne se limite pas à ces fonctions dans l'initiation de la transcription. Des études récentes ont montré qu'il était impliqué dans l'étape d'élongation. En effet, l'analyse par CHIP de la localisation de SAGA sur la chromatine a montré qu'il se trouvait sur la séquence codante et pas seulement sur les promoteurs (Govind et al. 2007). Cette présence est dépend de la phosphorylation du CTD sur la Ser5, montrant que SAGA est associé à la forme active de la Pol II. Ces observations suggèrent que le recrutement de SAGA sur les régions transcrites permette l'acétylation des *via* sa sous-unité Gcn5. Cette modification favoriserait l'éviction des nucléosomes au cours de l'élongation afin d'augmenter son efficacité (Govind et al. 2007).

SAGA est également impliqué dans l'export des ARNs *néo-synthétisés*. Sa sous-unité Sus1 interagit avec les complexes protéiques des pores nucléaires, orientant ainsi les *loci* transcrits vers le lieu d'export (Cabal et al. 2006). De plus, un rôle de SAGA dans la réparation de l'ADN a été mis en évidence (Guo et al. 2011).

Enfin, il est important de noter qu'il existe des complexes *SAGA-like*. Chez la levure on retrouve par exemple le complexe SAGA mais aussi SLIK (*SAGA-like complex*) autrement appelé SALSA (*SAGA*

altered Spt8 absent) (Pray-Grant et al. 2002; Sterner, Belotserkovskaya, and Berger 2002). Ces complexes ont des rôles semblables à celui de SAGA.

b) Les remodeleurs de la chromatine.

L'accès par les facteurs de transcription et de la machinerie aux séquences régulatrices sur les promoteurs est un aspect crucial de la régulation de la transcription. La structure de la chromatine est importante et peut être modulée en modifiant les histones mais aussi en modifiant leur position ou leur contact avec l'ADN. Les complexes remodeleurs de la chromatine permettent le déplacement des nucléosomes en utilisant l'hydrolyse de l'ATP comme source d'énergie. Ils ont donc comme caractère commun des domaines ATPase, entourés par d'autres domaines. Les remodeleurs sont classés en quatre familles différentes en fonction des domaines ATPase qu'ils contiennent : SWI/SNF, ISWI, CHD et INO80 (Clapier and Cairns 2009). Ces complexes sont très conservés de la levure à l'Homme (Tableau 8). Les remodeleurs ne sont bien sûr pas exclusivement réservés à la régulation de la transcription mais interviennent dans tous les processus liés à la chromatine (réplication, réparation de l'ADN...).

Famille	Organisme		
	Levure	Drosophile	Humain
SWI/SNF	SWI/SNF - RSC	BAP - PBAP	BAF - PBAF
ISWI	ISWI1a- ISWI1b –ISWI2	NURF – CHRAC - ACF	NURF – CHRAC - ACF
CHD	CHD1	CHD1 – Mi2/NuRD	CHD1 -NuRD
INO80	INO80 – SWR1	Pho – dINO80- Tip60	INO80-SRCAP-TRRAP/TIP60

Tableau 8: Tableau récapitulatif des classes de remodeleurs de la chromatine et leur représentant chez la levure, la drosophile et l'Humain (d'après Clapier and Cairns 2009).

- SWI/SNF

Les gènes codant pour les protéines formant les remodeleurs de la famille SWI/SNF (*switch defecting/sucrose non-fermenting*) a été découverts lors d'analyse de mutants affectant le changement de type sexuel (*mating type switch*) ou la croissance en glucose chez la levure. Une étude génétique de suppresseurs de mutations dans ces gènes a révélé des gènes codants pour des histones, suggérant le lien fonctionnel avec la chromatine (Mohrmann and Verrijzer 2005). Plus tard, deux complexes de près de 2MDa ont été purifiés, représentant les deux membres de cette famille (Cairns et al. 1994; Cairns et al. 1996). Il s'agit de SWI/SNF (BAF chez l'humain) et de RSC (*Remodel the Structure of Chromatin*, PBAF chez l'Homme). Ces deux complexes sont composés d'une dizaine de sous-unités pour le premier et d'environ 15 sous-unités pour le second. Certaines des sous-unités sont très conservés chez les eucaryotes (Mohrmann and Verrijzer 2005). SWI/SNF et RSC sont construits autour d'un cœur catalytique proche et partagent certaines sous-unités (Cairns et al. 1998). D'autres sont homologues entre les deux complexes. D'un point de vue structural, ce qui définit ces deux complexes comme appartenant à la famille SWI/SNF est la structure des sous-unités ATPase. Elles possèdent un domaine HSA (*helicase-SANT*), un domaine

post-HSA et un bromodomaine en C-terminal, permettant l'interaction avec les nucléosomes (Clapier and Cairns 2009). Leur activité ATPase est stimulée par la liaison à l'ADN, permise par le domaine hélicase (Laurent, Treich, and Carlson 1993).

Malgré ces similarités, les deux complexes présentent des différences. En effet, les gènes ciblés sont différents. SWI/SNF ne cible qu'environ 5% des gènes (Holstege et al. 1998) et est indispensable pour la viabilité cellulaire. Au contraire, RSC est beaucoup plus abondant, essentiel à la viabilité et plus ubiquitaire quant aux gènes qu'il régule (Cairns et al. 1996). Il est notamment impliqué dans l'expression de gènes de réponse aux stress, l'activation de la transcription, la ségrégation des chromosomes et la transcription des gènes de classe III (Hsu et al. 2003; Mohrmann and Verrijzer 2005).

D'un point de vue mécanistique, la façon dont les remodeleurs de la famille SWI/SNF interviennent pour réguler la transcription a été le sujet de beaucoup d'études. Ainsi, il a été montré chez la levure que ces complexes peuvent interagir directement avec les activateurs de la transcription afin de permettre leur recrutement et ainsi agir sur la chromatine pour faciliter l'initiation (Natarajan et al. 1999; Yudkovsky et al. 1999). Il a été également montré chez l'Homme qu'ils pouvaient interagir avec certains récepteurs nucléaires (Wallberg et al. 2000) ou d'autres facteurs de transcription tels que C/EBP ou EKLf (Armstrong, Bieker, and Emerson 1998; Kowenz-Leutz and Leutz 1999). Une fois recrutés, ils vont permettre le glissement et/ou l'éjection des nucléosomes.

- ISWI.

Les remodeleurs de la famille ISWI ont été identifiés pour la première fois chez la drosophile. Des protéines similaires ont été ensuite identifiées chez les autres eucaryotes (Corona and Tamkun 2004). Chez la levure, l'Homme et la drosophile, on retrouve trois complexes ISWI de 2 à 4 sous-unités (Clapier and Cairns 2009). Les sous-unités ATPase ont une très grande homologie avec celles de la famille SWI/SNF et sont donc nommées ISWI pour *Imitation SWI*. Les sous-unités ATPase de cette famille possèdent un domaine SANT et SLIDE capables de se lier à la queue non modifiée des histones et à l'ADN, respectivement (Boyer, Latek, and Peterson 2004; Corona and Tamkun 2004).

L'activité de ces complexes ISWI a été initialement montrée *in vitro* dans l'activation de la transcription sur la chromatine en permettant le glissement des nucléosomes (Mizuguchi et al. 1997; Owen-Hughes and Workman 1996). Cette activité a ensuite été montrée *in vivo* chez la drosophile. ISWI est essentiel à la viabilité et nécessaire à l'expression des gènes (Tsukiyama et al. 1995; Deuring et al. 2000). De manière intéressante, ISWI est également impliqué dans la répression de la transcription en permettant la maintenance d'un état de la chromatine défavorable à la transcription dans les régions d'hétérochromatine (Goldmark et al. 2000; Deuring et al. 2000). Enfin, chez la levure, ISWI joue un rôle dans les étapes *post* initiation en intervenant dans la terminaison et dans la régulation de la structure de la chromatine dans la région 3' du gène transcrit (Alén et al. 2002).

- CHD

Les protéines CHD (*chromodomain, helicase, DNA-binding*) constituent une classe de remodeleurs de la chromatine très conservée de la levure à l'Homme. La fonction précise de ces protéines est relativement mal caractérisée (Marfella and Imbalzano 2007). Les protéines de cette famille possèdent un domaine ATPase dérivé de SNF2 et deux motifs chromodomaine en tandem dans la région N-terminale. Certaines protéines Chd agissent seules, d'autres forment des complexes (Clapier and Cairns 2009).

- INO80

La dernière famille de remodeleurs de la chromatine est la famille des INO80 (*inositol requiring 80*). Elle contient plus de 10 protéines dont le complexe SWR1-*related* (Clapier and Cairns 2009). Un des complexes a été initialement purifié chez la levure *S. cerevisiae* (Shen et al. 2000). La

caractéristique principale des ATPases de cette classe de protéine est l'insertion d'un autre domaine au milieu du domaine catalytique. Cette insertion permet la liaison de certaines protéines partenaires dont des hélicases. Concernant leur rôle, les INO80 ont été montrées comme activant la transcription et la réparation de l'ADN (Clapier and Cairns 2009). Enfin, il est intéressant de noter que le complexe SWR1 remodèle la chromatine de manière particulière puisqu'il permet le retrait de dimères d'histones H2A-H2B canoniques et les remplace par le variant H2A.Z-H2B (Kobor et al. 2004; Gaku Mizuguchi et al. 2004).

Chapitre II : Le Médiateur de la transcription par l'ARN polymérase II.

Dans un système de transcription *in vitro*, la Pol II en présence de ses facteurs généraux de la transcription peut assurer le niveau basal de la transcription. Dans les systèmes reconstitués, l'ajout d'un activateur aux GTFs et Pol II purifiés ne permet pas l'activation de la transcription. C'est cette observation qui est à la base de la découverte des co-activateurs, dont le Médiateur fait partie. Ce co-régulateur, spécifique aux eucaryotes intervient dans la transmission des signaux des activateurs jusqu'à la machinerie de base de la transcription.

Ce chapitre présentera le complexe Médiateur chez la levure et chez les mammifères ainsi que son rôle dans la régulation de la transcription. Nous verrons ensuite les rôles du Médiateur au-delà de l'initiation de la transcription ainsi que les pathologies pouvant lui être associées.

I. Le Médiateur de *Saccharomyces cerevisiae*.

1) Découverte du Médiateur.

Le Médiateur de la transcription par l'ARN polymérase II a été découvert il y a maintenant près de 25 ans par le laboratoire de R. Kornberg. L'idée de son existence a fait suite à l'observation d'un phénomène d'interférence entre les activateurs (Kelleher, Flanagan, and Kornberg 1990). Dans une expérience de transcription *in vitro* en présence d'un activateur et d'un extrait nucléaire, l'ajout d'un deuxième activateur ne pouvant pas se lier au promoteur supprime l'activation. Cela suggère que ce deuxième activateur non spécifique a un effet de titration d'un facteur supplémentaire requis pour la transcription. Lorsque l'on ajoute une fraction protéique nucléaire de *S. cerevisiae* à ce système, l'activité transcriptionnelle est restaurée. Ces observations ont permis de montrer l'existence d'un facteur, le Médiateur, capable de stimuler la transcription activée *in vitro* en faisant un lien entre l'activateur et la machinerie de transcription (Flanagan et al. 1991).

En parallèle, des études génétiques ont permis de confirmer cette idée. Des travaux de recherche ont été menés afin d'identifier des gènes suppresseurs de mutations de troncatures du CTD de la Pol II. Un groupe de gènes, appelé *SRB* (*Suppressor of RNA polymerase B*) a été mis en évidence. La caractérisation biochimique des protéines codées par ces gènes a mené à la purification d'un complexe capable de se lier au CTD et de stimuler la transcription *in vitro* (Hengartner et al. 1995; Nonet and Young 1989; Thompson et al. 1993). Ce complexe a été depuis identifié comme étant le Médiateur. Plus tard, le Médiateur a été purifié sous forme d'un complexe d'environ 20 sous-unités, capable de stimuler la transcription activée mais également la transcription basale *in vitro* (Kim et al. 1994).

La présence du Médiateur a été depuis observée chez d'autres espèces telles que la drosophile, le ver *Caenorhabditis elegans*, la plante *Arabidopsis thaliana* et les mammifères (Park, Gim, et al. 2001; Kwon et al. 1999; Bäckström et al. 2007).

2) Composition et organisation du Médiateur.

La première purification du Médiateur après sa découverte a permis d'isoler un complexe de 20 sous-unités (Kim et al. 1994). Cette étude a mis en évidence que ce complexe contenait les protéines codées par les gènes *SRB* mais également d'autres protéines déjà identifiées comme impliquées dans la transcription. Des études plus approfondies basées sur des méthodes de purification différentes et d'analyses par spectrométrie de masse ont été ensuite réalisées afin de déterminer sa composition exacte chez la levure (Liu et al. 2001; Gustafsson et al. 1998; Han et al. 1999). Il est apparu que le Médiateur est composé de 24 sous-unités. Une dernière sous-unité (Med31) a ensuite été découverte comme faisant partie du Médiateur (Guglielmi et al. 2004; Linder and Gustafsson 2004), portant à 25 le nombre de sous-unités du Médiateur. Des expériences de biochimie dans lesquelles le complexe est dissocié par de l'urée ont suggéré l'existence d'une organisation modulaire du Médiateur (Kang et al. 2001). Ainsi, le Médiateur est organisé en un module de tête, de milieu, de queue et un module kinase (Figure 16). Ces modules ont été caractérisés par des méthodes biochimiques et de microscopie électronique. Le module de queue, composé de Med2, 3, 5, 14, 15 et 16, apparaît être le module impliqué dans la liaison du Médiateur avec les activateurs (Park et al. 2000; Han et al. 1999). Les modules de milieu (Med1, 4, 7, 9, 10, 21 et 31) et de tête (Med6, 8, 11, 17, 18, 19, 20 et 22), sont quant à eux impliqués dans des contacts avec la Pol II et la machinerie de transcription (Davis et al. 2002). Enfin, le module kinase, composé de Med12 et 13, Cdk8 et CycC, est plus labile et n'est par conséquent, pas toujours associé au Médiateur. Il comprend un couple Cycline/Kinase capable de phosphoryler le CTD. Il est plutôt associé à la répression de la transcription (Liao et al. 1995; Hengartner et al. 1998; Chi et al. 2001). Tous ces modules ont pu ensuite être assemblés à partir de protéines recombinantes (Koschubs et al. 2010; Takagi et al. 2006). A l'heure actuelle, on ne connaît pas bien les voies d'assemblages du Médiateur *in vivo* menant au complexe entier. Il semble néanmoins que le complexe soit stable à l'exception du module kinase, parfois retrouvé seul *in vivo* (Poss, Ebmeier, and Taatjes 2013).

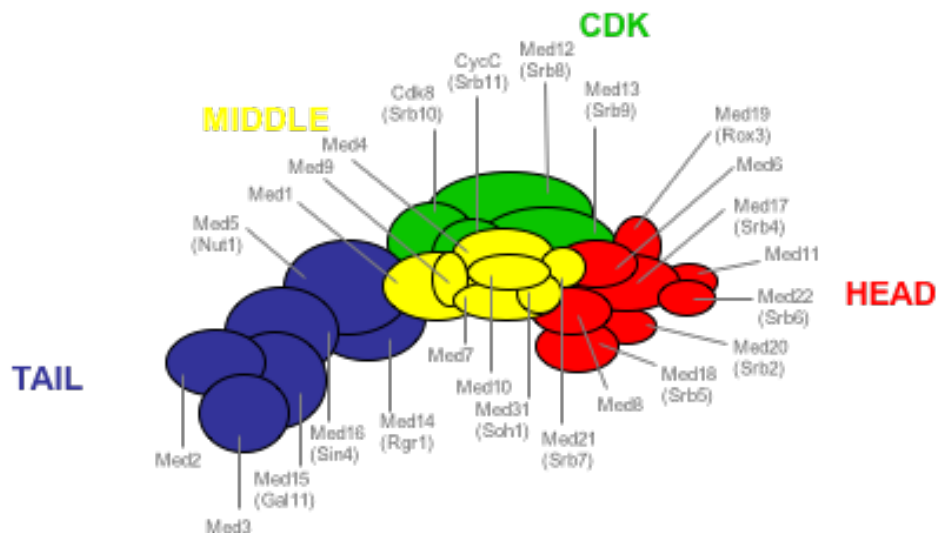


Figure 16: Composition et organisation modulaire du Médiateur de *S. cerevisiae* (d'après Guglielmi et al. 2004)

Des expériences de double-hybride ainsi que de co-immunoprécipitations (co-IP) ont permis de préciser les interactions existantes au sein du Médiateur (Guglielmi et al. 2004). Cela a permis une vue des contacts qui existent entre les modules (Figure 17). Ainsi, il apparaît que le module de tête interagit avec le module de milieu *via* le contact entre les sous-unités Med17 et Med21. Le module de queue peut se replier sur celui du milieu grâce à une interaction entre Med3 et Med21. La sous-unité Med14, qui se trouve à l'interface entre le module du milieu et la queue, joue un rôle important dans l'organisation structurale du complexe puisque sa délétion entraîne une dissociation du complexe (Li et al. 1995). Enfin, l'association du module Kinase au reste du Médiateur pourrait se faire par le contact entre Med13 et Med17.

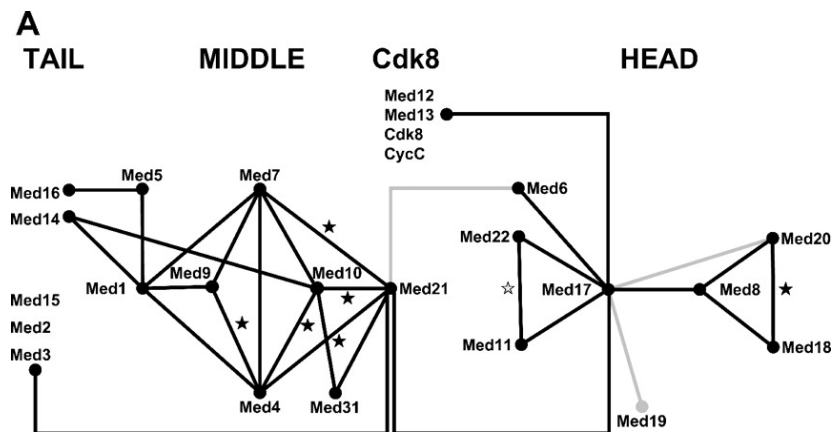


Figure 17: Carte des interactions entre les sous-unités du Médiateur (Guglielmi et al. 2004)

Pour produire cette carte d'interaction, les données de plusieurs études ont été combinées (Guglielmi et al. 2004; Ito et al. 2001; Uetz et al. 2000). Les traits gris correspondent à des interactions non observées dans le criblage issu de l'étude de Guglielmi et al. Les étoiles noires correspondent à des interactions entre les sous-unités du module du milieu homologues chez la drosophile. L'étoile grise correspond à une interaction entre Med11 et Med22 conservée chez le ver *C. elegans*.

3) Structure du Médiateur.

Les premières données structurales concernant le Médiateur ont été obtenues par microscopie électronique (Dotson et al. 2000). La première de ces études, en utilisant des mutants de certaines sous-unités, a permis de confirmer l'organisation modulaire. En effet, la délétion de Med16 abouti, par exemple, a une perte de densité électronique au niveau de la localisation supposée du module de queue. Il a également été montré que le Médiateur pouvait se trouver sous deux conformations. Lorsqu'il est isolé, il adopte une conformation relativement compacte et fermée. Le module de queue semble replié sur celui du milieu. Lorsqu'il est lié à la Pol II ou à un activateur, le Médiateur adopte une conformation plus allongée, ouverte et propice aux interactions avec les facteurs de la transcription. Ce changement de conformation apparaît très important pour le rôle du Médiateur dans l'activation de la transcription.

En raison de la très grande taille du Médiateur et sa grande flexibilité, à l'heure actuelle, aucune structure cristallographique complète du complexe n'a pu être obtenue. Néanmoins, les structures atomiques de certaines sous-unités et de sous-complexes ainsi que du module de tête entier sont disponibles. Ainsi, les structures connues sont celles de la cycline C, de Med1, du domaine KIX de Med15, du sous-complexe Med7/Med21, du sous-complexe Med18/Med20, du sous-complexe Med8C/Med18/med20, du sous-complexe Med11/Med22 et du sous-complexe Med7N/Med31 (Larivière, Seizl, and Cramer 2012). De manière intéressante, la structure du sous-complexe Med7/Med21 a révélé la présence d'une région charnière flexible, pouvant en partie

expliquer la possibilité de changements de conformation du Médiateur lorsqu'il se lie à la Pol II ou un activateur (Baumli, Hoepfner, and Cramer 2005). L'analyse du sous-complexe Med8/Med18/Med20 a montré qu'il était important dans la mise en place du PIC par une interaction directe entre Med8 et la TBP. La sous-unité Med8 est également importante pour la structure du module de tête (Larivière et al. 2008). La structure du sous-module Med11/Med22 a révélé la présence d'une interface conservée pour l'interaction avec TFIID (Seizl et al. 2011). Par ailleurs, une interaction directe entre la sous-unité Rad3 de TFIID et Med11 a été montrée par notre laboratoire (Esnault et al. 2008).

La structure atomique du module de tête entier a été obtenue par des laboratoires différents. Deux d'entre elles chez *S. cerevisiae* dont une en interaction avec le CTD (Robinson et al. 2012; Imasaki et al. 2011). La troisième structure a été obtenue chez *S. pombe* (Larivière et al. 2012) et prend en compte la sous-unité Med6. La résolution est également meilleure et permet la visualisation de régions de certaines sous-unités absentes dans les autres analyses. Le modèle obtenu chez *S. pombe* a ainsi permis de raffiner la structure du Médiateur de *S. cerevisiae*. Le module de tête est constitué de trois domaines distincts (Med6-8-11-22, Med17 et Med18-20), centrés sur un ensemble d'hélices appartenant à cinq sous-unités différentes. Cet ensemble hélicoïdal assure la stabilité des sous-unités de la tête et fournit des sites de liaison pour les GTFs et la Pol II. Med17 occupe une place centrale dans le module. Il est proposé une structure ressemblant à une tête de crocodile (Figure 18) contenant 8 éléments distincts. Parmi ces éléments, 4 sont mobiles. Il s'agit de l'épaule (*shoulder*), du doigt (*finger*), de la mâchoire mobile (*moveable jaw*) et du nez (*nose*). Le sous-module Med18-Med20 forme la mâchoire mobile par rapport au reste de la structure et sa flexibilité permettrait l'interaction avec la machinerie de Pol II. Concernant la liaison avec le CTD, il est proposé que celui-ci soit en interaction avec Med17, Med6 et Med8 (Robinson et al. 2012)

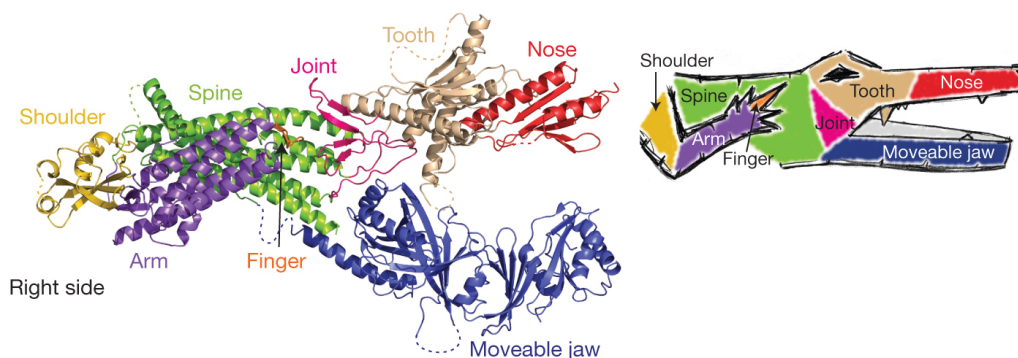


Figure 18: Modèle structural du module de tête du Médiateur (Larivière et al. 2012).

Concernant le module de milieu, le laboratoire de P. Cramer a utilisé des méthodes de pontage chimique (*crosslink*) couplés à la spectrométrie de masse pour étudier sa structure (Larivière et al. 2013). Cette méthode permet de mettre en évidence les résidus en contact dans une interaction protéine-protéine. Le *crosslink* étant réalisé avec une molécule dont la taille est connue, cela permet de déterminer la distance de l'interaction entre les deux résidus. En utilisant ces données, on peut ainsi fixer des contraintes topologiques afin de reconstruire un modèle structural se basant sur les données existantes de structure et sur des modélisations. Par cette approche, les auteurs ont pu proposer un modèle structural du module du milieu contenant 6 sous-unités, Med1 étant manquante (Figure 19). On observe sur ce modèle deux régions similaires formées par les dimères Med7/Med21 et Med9/Med4. Ces deux dimères sont flexibles et forment une région charnière. Ce modèle est cohérent avec l'hypothèse selon laquelle ce module jouerait un rôle important dans les transitions conformationnelles. De plus, il suggère une structure étendue longue de 180 Å, en accord avec les observations faites par microscopie électronique. L'organisation du module offre

une grande surface d'interaction corroborant l'idée que la surface d'interaction entre la machinerie de la Pol II et le Médiateur est étendue. Enfin, l'implication de résidus conservés dans les interactions visibles sur ce modèle suggère une bonne conservation de cette structure chez l'Homme.

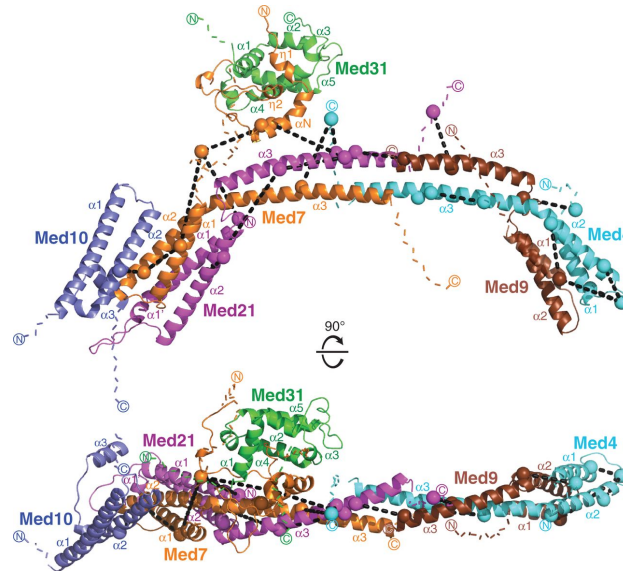


Figure 19: Modèle structural du module du milieu du Médiateur (d'après Larivière et al. 2013).

Les laboratoires des F. Asturias et G. Cai ont réalisés récemment des études structurales du Médiateur par microscopie électronique (Tsai et al. 2014; Wang et al. 2014). Les nouveaux modèles proposés (L'un d'entre eux est présenté en Figure 20) à des résolutions d'environ 20 Å indiquent que ceux obtenus précédemment (Cai et al. 2009), à des résolutions inférieures, étaient incorrects. Cela a permis un positionnement plus précis des modules aboutissant à une réorganisation complète par rapport aux modèles précédents. Le module de tête préalablement positionné à la base de la structure, surmonté par les modules du milieu et de queue, se retrouve désormais adjacent au module du milieu, la base correspondant finalement au module de queue.

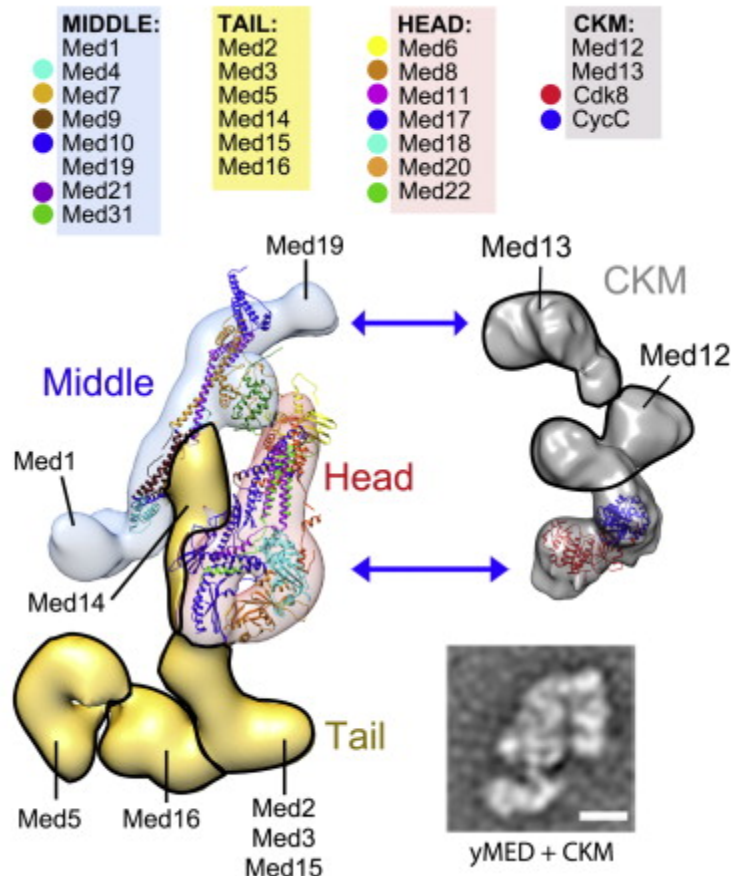


Figure 20 : Modèle structural du Médiateur de *S. cerevisiae* obtenu par cryo-EM (d'après Tsai et al. 2014).

La position des différents modules du Médiateur est indiquée. Les structures atomiques connues sont incluses dans le modèle. Le modèle du module kinase est également proposé et la localisation de son interaction avec le coeur du Médiateur est indiquée.

Récemment, le laboratoire de R. Kornberg a proposé un modèle structural du Médiateur entier, basé sur des expériences de *crosslink* couplé à la spectrométrie de masse et de microscopie électronique. L'intégration de ces données avec les données structurales existantes, notamment du module de tête, a permis la reconstruction d'un modèle du complexe entier. (Robinson et al. 2015). Cette étude offre pour la première fois un modèle structural du Médiateur entier avec dans certaines régions une résolution quasi-atomique (Figure 21). Le modèle proposé montre une validation de presque toutes les interactions observées jusqu'à maintenant par double hybride ou co-IP à l'exception de trois d'entre elles (Med3-Med21, Med1-Med7, and Med1-Med5). Les auteurs expliquent ces différences par le fait que ces interactions puissent avoir lieu dans une autre conformation du Médiateur. Cependant, malgré ces différences, la présence de presque toutes les interactions ainsi que la correspondance avec les données de microscopie électronique déjà existantes valide le modèle obtenu. L'analyse du module du milieu dans ce modèle révèle la présence de deux dimères (Med7C-Med21 et Med4-Med9) formant une plateforme d'assemblage pour les autres sous-unités. Les sous-unités Med10, Med19 et Med31 interagissent avec Med7-Med21 alors que Med1 interagit avec Med4-Med9. Cela est en accord avec le modèle du module du milieu proposé par P. Cramer. Cette étude a permis également de modéliser le module de queue. Il apparaît que les sous-unités interagissent par l'intermédiaire de super-hélices (*coiled-coil*). Ce module forme une grande plateforme d'interaction, cohérente avec son rôle dans la liaison à des activateurs. Enfin, l'une des principales nouveautés de cette étude est d'avoir proposé un rôle architectural unique pour Med17 et Med14. En effet, on observe que l'extrémité N-

terminale de Med17 interagit avec le module du milieu. Cette interaction est essentielle pour la cohésion entre le module de tête et du milieu. Ce résultat est cohérent avec le fait que des tronçatures de ce domaine sont létales ou encore pour expliquer le phénotype thermosensible d'un mutant *med17-138* dont la tête est dissociée du reste du Médiateur (Holstege et al. 1998). Concernant Med14, les analyses de *crosslink* montrent des contacts entre elle et des sous-unités tout au long du complexe. Le modèle situe ainsi Med14 tout au long du Médiateur, dans les trois modules. L'extrémité C-terminale de Med14 contacte le module de queue. Dans le module du milieu Med14 interagit avec le dimère Med4-Med9. Dans le module de tête Med14 et Med17 interagissent également. Ces observations confirment les conclusions d'une étude récente selon laquelle la sous-unité Med14 est essentielle à la formation d'un Médiateur actif lorsqu'il est reconstitué *in vitro* (Cevher et al. 2014).

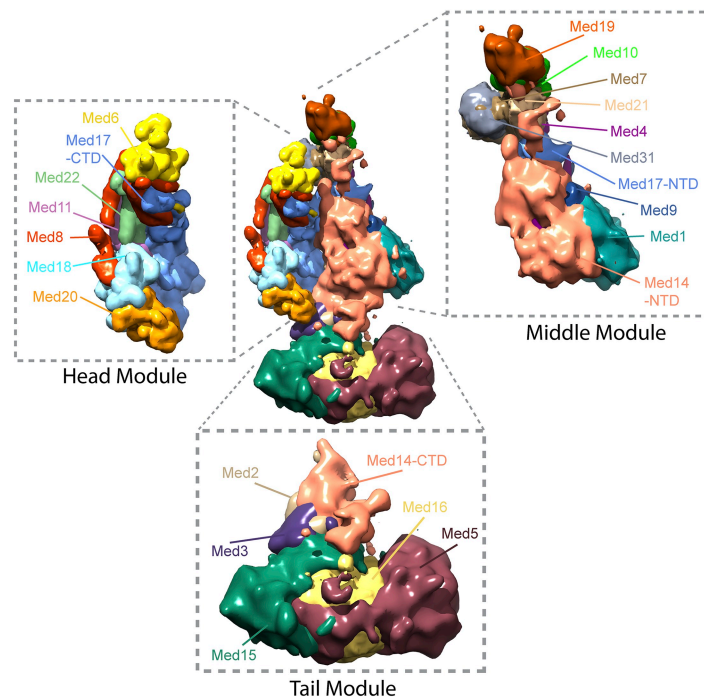


Figure 21: Modèle structural du Médiateur proposé par modélisation intégrative (d'après Robinson et al. 2015)

II. Le Médiateur chez les mammifères.

1) Découverte du Médiateur des mammifères.

L'idée de l'existence du Médiateur chez les mammifères est née lors de la mise en évidence par le laboratoire de R. Roeder d'un complexe aux propriétés similaires à celui du Médiateur de *S. cerevisiae* (Fondell, Ge, and Roeder 1996). Ce complexe a été co-purifié avec le récepteur aux hormones thyroïdiennes dans des cellules humaines HeLa. Il est pour cette raison appelé TRAP (*thyroid hormone receptor associated protein*). TRAP est nécessaire à l'activation de la transcription par le récepteur aux hormones thyroïdiennes dans un système *in vitro*.

A la suite de cela, plusieurs laboratoires ont purifié des complexes possédant les mêmes

propriétés. Ils ont pour la plupart été co-purifiés avec un activateur de la transcription. Ainsi, les complexes CRSP (*cofactor required for Sp1 activation*), PC2 (*positive cofactor 2*), DRIP (*vitamin D receptor-interacting proteins*), SMCC (*suppressor of RNA polymerase B Mediator containing cofactor*) et NAT (*negative regulator of activated transcription*) ont été mis en évidence à partir de cellules HeLa (Gu et al. 1999; Malik et al. 2000; Rachez et al. 1998; Ryu et al. 1999; Sun et al. 1998). Enfin, des complexes pouvant interagir avec différents domaines d'activation de la transcription ont été mis en évidence, il s'agit des complexes ARC et Médiateur humain (Boyer et al. 1999; Näär et al. 1999).

2) Composition du Médiateur des mammifères.

Les différentes purifications présentées au-dessus ont laissé penser qu'il existait plusieurs Médiateurs chez les mammifères. L'analyse des protéines qui composent ces complexes a permis l'identification d'une trentaine de sous-unités dont beaucoup présentent des similitudes avec les sous-unités du Médiateur de la levure (Figure 22).

L'investigation biochimique des différents complexes ayant des activités du type Médiateurs montre qu'il existe des complexes de hauts poids moléculaires (plus de 1MDa, TRAP/SMCC, DRIP et ARC) et des plus petits (de 500 à 700 kDa, PC2, CRSP et NAT). Les gros complexes contiennent le module kinase ou une partie, les petits complexes en sont dépourvus (Conaway et al. 2005). Il a également été montré que ces différents Médiateurs ne présentent pas systématiquement la même composition en sous-unités. Cette hétérogénéité pourrait être expliquée par la régulation de l'expression des gènes tissu spécifique chez les mammifères, ou bien par des protocoles différents de purification (Malik and Roeder 2005; Conaway et al. 2005).

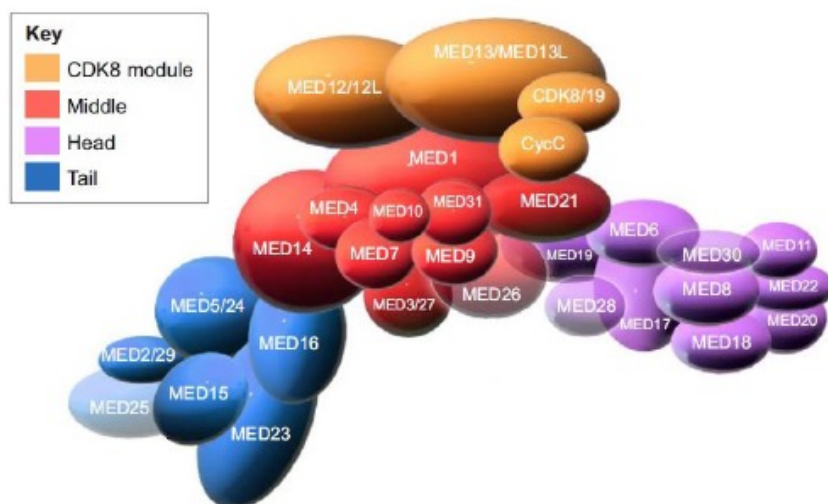


Figure 22: Composition en sous-unités du Médiateur des eucaryotes pluricellulaires (Malik and Roeder 2010).

Néanmoins, une étude protéomique menée par le laboratoire de R. Conaway et J. Conaway a apporté des arguments pour conforter l'idée qu'il existe un seul et unique Médiateur humain de composition uniforme. Les différences de composition seraient la conséquence de pertes de sous-unités durant les purifications et/ou à un manque de sensibilité dans la détection. La méthode

utilisée ici est appelée MudPIT (*MU*lti*D*imensional *P*rotein *I*dentification *T*echnology). Elle consiste en une analyse par spectrométrie de masse d'un mélange de protéines sans passage par une étape d'électrophorèse en deux dimensions. Cette électrophorèse en deux dimensions présente une résolution limitée pouvant rendre certaines protéines indétectables. Dans la méthode de MudPIT, cette étape est remplacée par une séparation en deux phases sur colonne de chromatographie (Washburn, Wolters, and Yates 2001). Cette technique a été utilisée sur des Médiateurs purifiés *via* les sous-unités Med9, Med10, Med19, Med26, Med28 ou Med29. Au cours de cette analyse toutes les sous-unités des complexes préalablement identifiés ont été détectées (Sato et al. 2004). Cela permet ainsi de proposer un Médiateur humain composé de 30 sous-unités. Il ne peut cependant pas être totalement exclu que dans certains types cellulaires, certaines sous-unités peuvent varier au sein du Médiateur. Notamment, la sous-unité Med26 présente la caractéristique de ne pas toujours être présente dans le complexe et de moduler, en fonction de sa présence, l'interaction avec la Pol II (Paoletti et al. 2006).

3) Conservation du Médiateur au cours de l'évolution.

Lorsque le Médiateur et ses sous-unités ont été identifiés chez *S. cerevisiae*, des alignements de séquences ont été réalisés afin de trouver des homologues chez l'Homme. Ces études n'ont permis d'identifier des homologues que pour 8 sous-unités à savoir Med14, Med4, Med6, Med7, Med10, Med17, Med21, Cdk8, CycC et Med31 (Rachez and Freedman 2001). Cela laissait entrevoir que les Médiateurs de levure et humain étaient très différents. Cependant, la découverte des Médiateurs humains et d'autres espèces a permis d'améliorer les analyses de génomique comparative. Grâce à des méthodes d'alignements de séquences multiples des sous-unités du Médiateur, il a été possible d'identifier de nombreuses régions similaires entre les sous-unités des Médiateurs humains et de levure (Boube et al. 2002; Bourbon 2008). Ainsi, 22 sous-unités parmi les 25 du Médiateur de la levure ont été associées avec un homologue chez l'Homme. Seules les sous-unités Med2, Med3 et Med5 n'ont pas d'homologies dans ces études. Il est intéressant de noter que ces trois sous-unités appartiennent au module de queue et sont supposées interagir avec les activateurs spécifiques. La diversification des activateurs chez les autres eucaryotes a probablement menée à une évolution plus importante de ces sous-unités. Il semble cependant que ces trois sous-unités puissent être homologues des sous-unités Med29, Med27 et Med24, respectivement. Les travaux de H.M Bourbon ont permis de montrer que la quasi-totalité des sous-unités du Médiateur chez *S. cerevisiae* ont un homologue chez les métazoaires (Figure 23). Il est proposé que les sous-unités supplémentaires chez les métazoaires soient apparues au cours de l'évolution, permettant des interactions avec des activateurs plus diversifiés, augmentant ainsi les possibilités de régulation lors de l'activation de la transcription.

L'existence de ces homologies a abouti à la définition d'une nomenclature commune des sous-unités du Médiateur pour toutes les espèces (Bourbon et al. 2004).

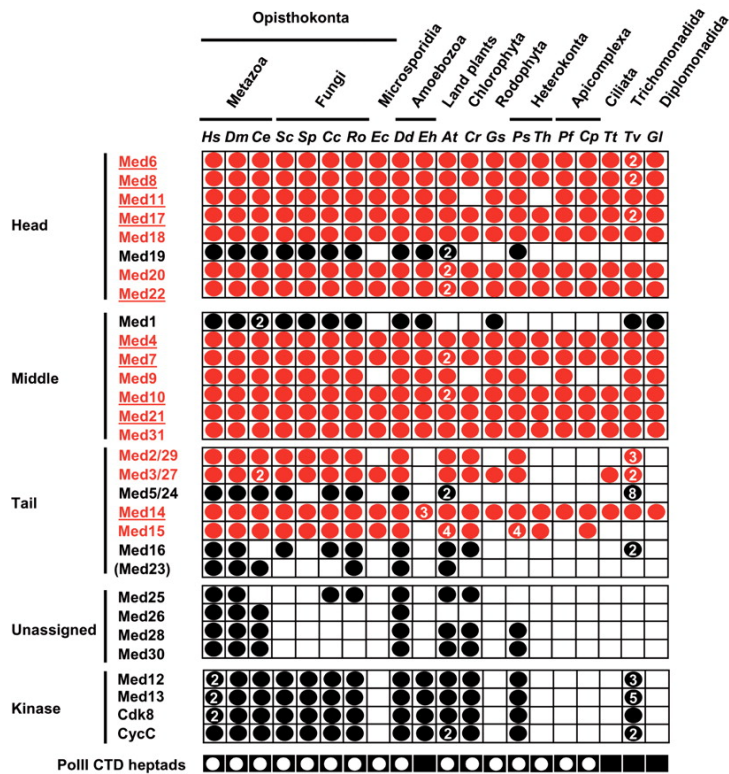


Figure 23: Conservation des sous-unités du Médiateur chez différents eucaryotes (d'après Bourbon 2008).

Les sous-unités sont groupées selon le module auquel elles appartiennent conformément au modèle décrit chez *S. cerevisiae*. Les sous-unités essentielles à la viabilité cellulaire sont soulignées. Les points indiquent qu'au moins un homologue est présent dans le génome analysé, si plusieurs paralogues ont été détectés, leur nombre est indiqué. Les sous-unités les plus conservées sont indiquées en rouge. Celles-ci incluent les sous-unités essentielles pour la viabilité cellulaire chez la levure. Hs : *H. sapiens*, Dm : *D. melanogaster*, Ce : *C. elegans*, Sc : *S. cerevisiae*, Sp : *S. pombe*, Cc : *C. cinereus*, Ro : *R. oryzae*, Ec : *E. cuniculi*, Dd : *D. discoideum*, Eh : *E. histolytica*, At : *A. thaliana*, Cr : *C. reinhardtii*, Gs : *G. sulphuraria*, Ps : *P. sojae*, Th : *T. pseudonana*, Pf : *P. falciparum*, Cp : *C. parvum*, Tt : *T. thermophila*, Tv : *T. vaginalis*, Gl : *G. lamblia*.

La conservation structurale a été également le sujet de plusieurs études (Cai et al. 2009; Tsai et al. 2014). La visualisation par microscopie électronique du Médiateur humain et de levure permet la comparaison des deux structures (Figure 24). Il apparaît au travers de ces études que l'architecture globale du complexe ainsi que la position des modules est conservée. De plus, les formes des modules de tête et du milieu sont très similaires. Cela est cohérent puisque ces deux modules sont impliqués dans les contacts avec la machinerie de transcription, conservé elle aussi. La structure du Médiateur semble donc conservée entre la levure et l'Homme. La ressemblance des structures, malgré les différences de séquences entre les espèces peut s'expliquer par une bonne conservation des structures secondaires des sous-unités du Médiateur.

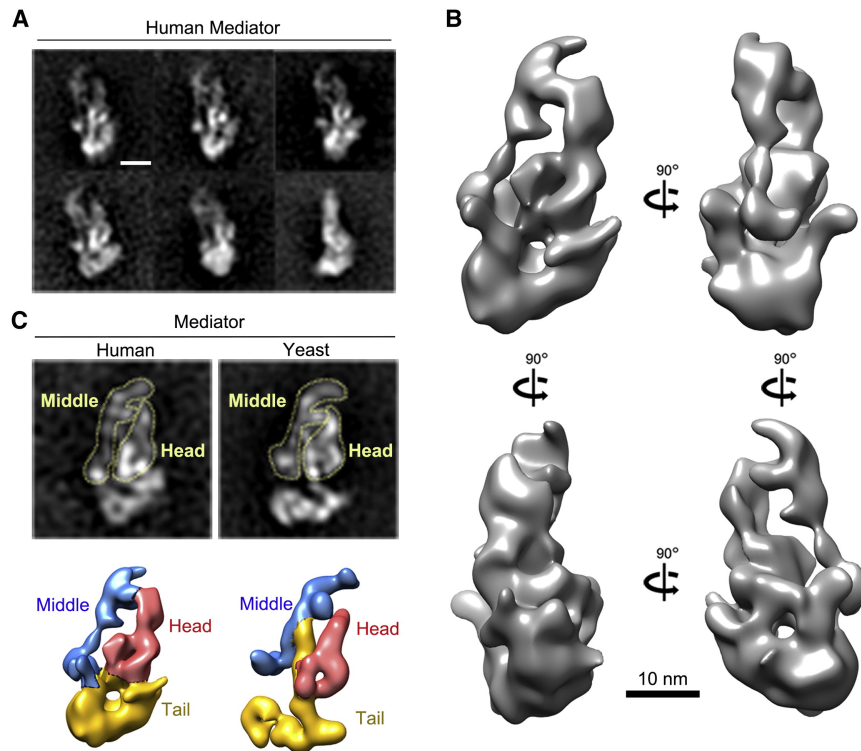


Figure 24: Comparaison des structures 3D des Médiateurs humain et de levure observés par microscopie électronique (Tsai et al. 2014).

(A et C) Images des Médiateurs de levure et humain obtenus par microscopie électronique (EM) et comparaison des structures résultantes. (B) Modèle 3D du Médiateur humain calculé à partir des images EM à la résolution estimée de 3 nm.

III. Rôle du Médiateur dans la régulation de l'expression des gènes.

La caractérisation du Médiateur a permis de définir un certain nombre de ses fonctions. Sa fonction première est l'activation de la transcription en intégrant les signaux de régulation des activateurs vers la machinerie de la Pol II. Il est capable d'interagir aussi bien avec les activateurs, les GTFs et la Pol II. Il présente donc un rôle central dans la régulation de la transcription.

1) Interaction entre le Médiateur et les activateurs spécifiques.

Le recrutement d'un activateur sur une région régulatrice (UAS ou *enhancer*) est le premier pas de la transcription activée en induisant un signal positif sur la transcription. Le Médiateur va agir comme un relais physique pour transmettre ce signal à la machinerie de la Pol II. De ce fait, il présente la capacité d'interagir avec un grand nombre de ces activateurs (Tableau 9). Les premières expériences de génétique et de biochimie chez *S. cerevisiae* ont permis de proposer l'implication du module de queue dans les interactions avec les activateurs. En effet, la délétion ou la mutation de gènes du module de queue tels que Med15, Med3 et Med2 aboutissent à une

incapacité d'activer la transcription en réponse aux activateurs Gal4 et Gcn4. Le niveau basal de transcription reste lui inchangé (Myers et al. 1999; Lee et al. 1999). Inversement, l'inactivation des sous-unités n'appartenant pas au module de queue mène à une diminution de la transcription activée et basale. L'interaction avec l'activateur n'est cependant pas perdue. Cela indique donc un rôle très important du module de queue dans la liaison avec les activateurs. Il a été ensuite montré *in vitro* que les activateurs Gal4, Gcn4 et VP16 peuvent interagir avec Med15 (Park et al. 2000). D'un point de vue structural, Med15 possède un domaine conservé appelé domaine KIX, nécessaire pour l'interaction avec certains activateurs (Thakur et al. 2009). Néanmoins, les sous-unités du module de queue ne sont pas les seules à pouvoir interagir avec les activateurs. Certaines sous-unités des modules de tête et du milieu peuvent interagir avec certains activateurs. C'est le cas, par exemple, de Med17 qui présente la capacité de se lier à Gal4 *in vitro* (Koh et al. 1998).

Chez les mammifères, le Médiateur a été découvert par co-purification avec des récepteurs nucléaires (NR). Ces récepteurs, qui appartiennent à une famille d'activateur, sont recrutés sur l'ADN après liaison avec un ligand. Ces observations suggèrent l'existence d'une interaction entre les activateurs et le Médiateur. Il est ainsi établi que le Médiateur peut interagir avec des récepteurs nucléaires tels que VDR (*vitamin D receptor*), PPAR (*peroxisome proliferator-activated receptors*), HNF4 (*hepatocyte nuclear factor 4*), GR (*glucocorticoid receptor*) et ER (*estrogen receptor*) (Rachez et al. 1999; Ge et al. 2002; Malik et al. 2002; Hittelman et al. 1999; Kang et al. 2002). L'analyse de ces interactions a permis de montrer l'implication de Med1 dans ces contacts. Med1 est une sous-unité du module du milieu mais néanmoins proche de la jonction entre les module de queue et du milieu. Cette protéine possède des domaines LxxLL (Leucine, X, X, Leucine, Leucine), capables d'interagir avec le domaine AF2 (*activation function 2*), caractéristique des récepteurs nucléaires. Cette interaction permet *in vitro* le recrutement du Médiateur entier (Malik et al. 2004; Ansari and Morse 2013). Cependant, l'interaction entre certains récepteurs nucléaires et le Médiateur est indépendante de Med1. Il a ainsi été montré que Med14 interagit avec PPAR γ , HNF4 et le récepteur aux glucocorticoïdes (Grøntved et al. 2010; Malik et al. 2002; Hittelman et al. 1999).

D'autres sous-unités que Med1 et Med14 interagissent avec des activateurs, différents des récepteurs nucléaires, chez les mammifères. Ainsi, Med15 est impliqué dans la liaison avec le facteur de transcription SREBP-1a (*Smad2/3/4 and sterol regulatory element binding protein 1a*). Cette interaction se fait par l'intermédiaire du domaine KIX conservé de Med15 (Yang et al. 2006). Deux sous-unités absentes chez la levure *S. cerevisiae*, Med23 et Med25, interagissent également avec des activateurs (Ansari and Morse 2013). Il est intéressant de noter que toutes ces sous-unités appartiennent au module de queue ou en sont très proches, Med14 étant impliqué dans la liaison entre les modules de queue et du milieu. Cela suggère donc un rôle clé du module de queue dans la liaison entre le Médiateur et les activateurs chez les mammifères, comme chez la levure.

Le module kinase est lui aussi impliqué dans la réponse aux activateurs. En effet, chez la drosophile, les sous-unités Med12 et Med13 interviennent dans la voie de signalisation Wnt importante pour le développement embryonnaire, *via* une interaction avec la β -caténine (Carrera et al. 2008).

Les activateurs spécifiques vont promouvoir l'expression de gènes particuliers, impliqués dans des fonctions cellulaires déterminées. L'interaction d'une ou plusieurs sous-unités du Médiateur avec ces régulateurs les impliquent dans des processus physiologiques. Par exemple, chez l'Homme, Med25 en interagissant avec HNF4, est impliqué dans le métabolisme des xénobiotiques et des lipides (Rana et al. 2011). L'interaction entre Med15 et SREBP1 implique le Médiateur dans l'homéostasie du cholestérol (Yang et al. 2006). Chez l'Homme, Med13 est impliqué dans le contrôle de l'homéostasie du glucose en régulant l'expression des gènes MURR1 et MEF2, impliqués dans le métabolisme du glucose (Amoasii et al. 2016).

Comme évoqué plus tôt, il existe dans les cellules des eucaryotes pluricellulaires de facteurs de transcription maîtres (*master regulator*). Ces facteurs contrôlent l'expression d'un grand nombre de gènes en même temps et sont donc impliqués dans le maintien ou l'établissement d'un

programme transcriptionnel particulier. C'est le cas, par exemple, de GATA1, régulateur maître de l'érythropoïèse. Il a été montré que Med1 pouvait interagir avec GATA1 (Stumpf et al. 2006). La sous-unité Med25 est quant à elle impliquée dans la chondrogenèse par une interaction avec Sox9 (Nakamura et al. 2011). Les *super-enhancers* montrent un enrichissement important du Médiateur et des régulateurs maîtres, suggérant que les interactions entre eux permettent l'établissement de ces régions particulières (Whyte et al. 2013). Il apparaît que des modifications même légères de l'occupation du Médiateur ou des régulateurs maîtres dans les *super-enhancers* changent fortement l'expression des gènes cibles et le programme transcriptionnel. Il est donc proposé que le Médiateur, en favorisant la formation de ces régions régulatrices, soit un coordinateur maître (*master coordinator*) du développement cellulaire.

Enfin, en plus de permettre son recrutement, la liaison d'un activateur avec le Médiateur promeut des changements conformationnels. Des études ont montré que l'interaction de différents activateurs avec le Médiateur ne va pas aboutir à la même conformation du complexe (Taatjes et al. 2002). Cela permet de proposer que ce mécanisme pourrait expliquer une part de la spécificité de réponse du Médiateur pour un activateur particulier. La différence de conformation est également impliquée dans le recrutement d'autres co-activateurs. Il a ainsi été observé par des méthodes protéomiques (MudPIT) que les protéines associées au Médiateur en interaction avec SREBP étaient différentes de celles liées au Médiateur en interaction avec l'activateur VP16, par exemple (Ebmeier and Taatjes 2010). Ainsi, les signaux de régulation des différentes voies de signalisation convergent vers le Médiateur à partir des activateurs. Le Médiateur pourrait alors intégrer ces signaux et prendre une conformation particulière, aboutissant à une réponse transcriptionnelle spécifique, notamment en recrutant les co-activateurs requis *via* sa conformation. Le recrutement des autres co-activateurs ne dépend probablement pas uniquement de la conformation et peut être coopératif avec celui du Médiateur. Malgré cette nuance, on peut aisément appréhender le rôle central du Médiateur dans l'intégration des signaux régulateurs de la transcription et imaginer son implication dans des pathologies (Voir Chapitre II.VII).

<i>Sous-unité</i>	<i>Facteur de transcription</i>	<i>Organisme</i>	<i>Fonction</i>
<i>Med1</i>	TR	<i>H. sapiens</i>	Récepteur de l'hormone thyroïdienne
	VDR	<i>H. sapiens</i>	Récepteur de la vitamine D
	HNF4	<i>H. sapiens</i>	Développement du foie
	ER	<i>H. sapiens</i>	Récepteurs aux œstrogènes
	P53	<i>H. sapiens</i>	Régulation du cycle cellulaire, suppresseur de tumeur
	GR	<i>H. sapiens</i>	Récepteur aux glucocorticoïdes
	STAT2	<i>H. sapiens</i>	Croissance, survie et différenciation en réponse à l'interféron
	GATA1	<i>M. musculus</i>	Régulateur majeur de l'érythropoïèse
<i>Med12</i>	PPAR γ	<i>M. musculus</i>	Métabolisme des lipides
	REST	<i>H. sapiens</i>	Différenciation neuronale
	Sox9	<i>H. sapiens</i>	Régulateur majeur de la chondrogenèse
	Nanog β -caténine	<i>H. sapiens</i> <i>D. melanogaster</i>	Régulateur majeur de la pluripotence des cellules ES Développement embryonnaire dépendant de la voie Wnt
<i>Med14</i>	PPAR γ	<i>H. sapiens</i>	Métabolisme des lipides
	HNF4	<i>H. sapiens</i>	Développement du foie
	GR	<i>H. sapiens</i>	Récepteur aux glucocorticoïdes
<i>Med15</i>	Gal4	<i>S. cerevisiae</i>	Réponse au galactose
	Gcn4	<i>S. cerevisiae</i>	Réponse à une carence en acides aminés
	Oaf1	<i>S. cerevisiae</i>	β -oxydation des acides gras
	Msn2	<i>S. cerevisiae</i>	Réponse au stress
	Pdr1	<i>S. cerevisiae</i>	Régulateur majeur pour les gènes de résistance aux drogues
	SREBP	<i>H. sapiens</i>	Homéostasie des lipides
	SMAD2-3	<i>H. sapiens</i>	Développement du mésenchyme en réponse à TGF β
	SBP-1	<i>C. elegans</i>	Homologue de SREBP1, impliqué dans le métabolisme des lipides
<i>Med17</i>	NHR-49	<i>C. elegans</i>	Homologue de HNF4, impliqué dans le métabolisme des lipides et la survie cellulaire
	VP16	<i>HSV</i>	TF dérivé du HSV
	P53	<i>H. sapiens</i>	Régulation du cycle cellulaire, suppresseur de tumeur
<i>Med23</i>	HSF	<i>D. melanogaster</i>	Réponse à un choc thermique
	E1A	<i>H. sapiens</i>	Gène exprimé durant la réplication des adénovirus
	ELK1	<i>H. sapiens</i>	Adipogenèse insulino-dépendante
<i>Med25</i>	C/EBP	<i>H. sapiens</i>	Prolifération cellulaire, croissance et différenciation
	VP16	<i>HSV</i>	TF dérivé du HSV
	IE62	<i>H. sapiens</i>	TF associé au tégument des virions du VZV
	HNF4	<i>H. sapiens</i>	Développement du foie
	Sox9	<i>H. sapiens</i>	Régulateur majeur de la chondrogenèse
	PEA3	<i>H. sapiens</i>	Croissance cellulaire et différenciation

Tableau 9: Tableau récapitulatif des interactions connues entre différentes sous-unités du Médiateur et des facteurs de transcription. (Adapté depuis Sohail Malik and Roeder 2010; Ansari and Morse 2013)

2) Interaction entre le Médiateur et la Pol II.

Les premiers travaux de purifications du Médiateur chez la levure *S. cerevisiae* ont montré qu'il pouvait exister sous deux formes, l'une libre et l'autre associée à la Pol II (Lee et al. 1999; Thompson et al. 1993). Le complexe formé par le Médiateur et la Pol II a été initialement appelé holoenzyme. Des expériences de co-immunoprécipitation et de *crosslink* par le formaldéhyde sont venues confirmer cette interaction (Tardiff, Abruzzi, and Rosbash 2007). Bien que ces expériences montrent une interaction entre la Pol II et le Médiateur, seules les observations par microscopie électronique ont permis de laisser entrevoir comment s'organise ce contact (Davis et al. 2002). Il a ainsi été proposé un modèle où les sous-unités Rpb3 et Rpb11 de la Pol II sont face au Médiateur. La Pol II interagit avec le Médiateur au niveau des modules de tête et de milieu. Cependant, deux modèles obtenus ont proposé une orientation différente de la Pol II d'environ 90° (Davis et al. 2002; Cai et al. 2009). Ces différences pourraient s'expliquer par le fait de l'absence de GTFs et d'activateur dans cette structure, ceux-ci pouvant participer à fixer une conformation plus stable. Les difficultés d'interprétation des images pourraient également être en cause dans ces différences d'observation. La dernière étude en date du Médiateur par microscopie électronique faite par le laboratoire de F. Asturias a proposé un nouveau modèle d'assemblage du Médiateur et de la Pol II (Tsai et al. 2014). Ce modèle impliquerait plusieurs étapes. La première étape serait la liaison du CTD de la Pol II dans la jointure entre les modules de tête et de milieu. Cette liaison provoquerait de grands changements conformationnels aboutissant au déplacement du module de milieu pour libérer l'espace où la Pol II se fixe. Ce modèle est en accord avec la structure du module de tête en interaction avec le CTD proposé par le laboratoire de R. Kornberg (Robinson et al. 2012) et des analyses *in vitro* à partir du Médiateur purifié de *S. cerevisiae* ont montré qu'en absence du CTD de la Pol II, le complexe Med-Pol II ne peut pas se former (Myers et al. 1998). La liaison du Médiateur avec un activateur pourrait provoquer une rotation du module de queue facilitant là encore le placement de la Pol II.

A partir des données de microscopie électronique, il a été proposé que les sous-unités Rpb1, Rpb2, Rpb3, Rpb6 et Rpb11 sont en contact avec les modules de tête et de milieu (Davis et al. 2002). Les modèles plus récents ont aussi suggéré une interaction de Med18 et Med20 avec les sous-unités Rpb4 et Rpb7 et avec la pince de la Pol II (Cai et al. 2009). Plus tard, une étude menée par notre laboratoire faisant appel à des expériences de photo-*crosslink in vivo* a mis en évidence une interaction directe entre Rpb3 et Med17 (Soutourina et al. 2011). Cette interaction a ensuite été caractérisée par des expériences de génomique fonctionnelle et a été montrée comme nécessaire au recrutement de la Pol II sur l'ensemble des gènes de levure. De plus, la conservation des sous-unités Med17 et Rpb3 au cours de l'évolution laisse penser que ce contact est conservé chez tous les eucaryotes.

Récemment, le laboratoire de P. Cramer a proposé un modèle structural de PIC partiel comprenant la Pol II, les GTFs (exceptés TFIIF et TFIIE) ainsi qu'un Médiateur cœur (cMed) comprenant les modules de tête et milieu (Plaschka et al. 2015). Les auteurs proposent une structure obtenue par cryo-microscopie électronique et d'expériences de *crosslink* (Figure 25). Le positionnement de la Pol II sur ce modèle est encore différent de celui proposé jusqu'alors (Davis et al. 2002; Cai et al. 2009). Il est proposé que le Médiateur soit positionné sur la Pol II au niveau de la tige Rpb4-Rpb7. Il existe, d'après les auteurs, 3 interfaces d'interactions entre le Médiateur et la Pol II. L'une entre Med18-Med20 et le dimère Rpb3-Rpb11. La deuxième interface est celle formée entre Rpb4-Rpb7 et Med8. La troisième interface est moins précise et implique le module de milieu. Ces données sont en contradiction avec les résultats obtenus par notre laboratoire puisque ce modèle ne visualise pas le contact que nous avons mis en évidence entre Rpb3 et Med17. Il est important de noter que cette interaction a été observée *in vivo* et qu'elle est fonctionnelle et nécessaire pour le recrutement de la Pol II sur l'ensemble des gènes de classe II.

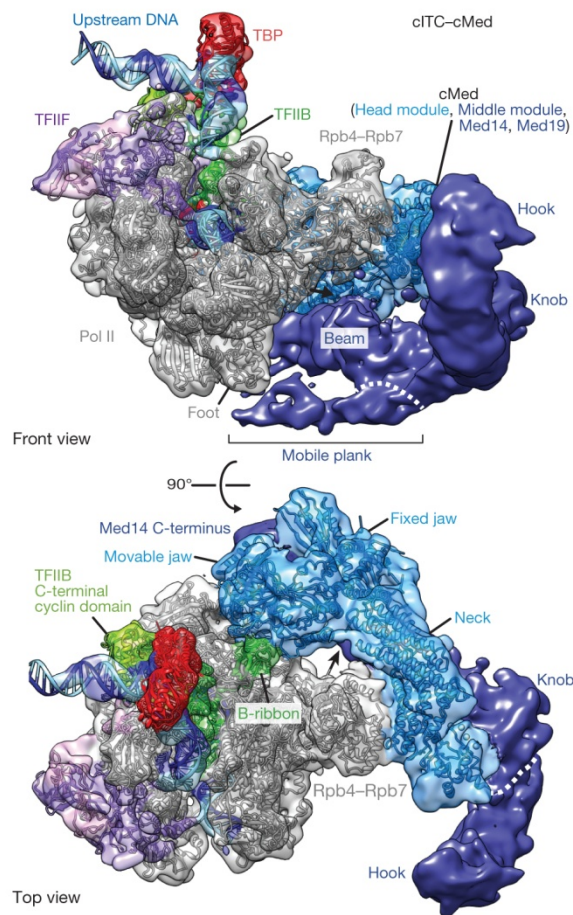


Figure 25: Modèle structural obtenu par cryo-microscopie électronique du PIC partiel composé de la Pol II, TFIIA, TFIIB, TBP, TFIIF et du Médiateur cœur (d'après Plaschka et al. 2015b).

Le module du milieu du Médiateur est coloré en violet, le Module de tête est coloré en bleu, la Pol II est colorée en gris.

Dans le modèle de Médiateur entier proposé par le laboratoire de R. Kornberg présenté plus tôt (Robinson et al. 2015), les auteurs ont proposé un modèle d'interaction entre la Pol II et le Médiateur. Leurs résultats ne sont pas totalement en accord avec ceux du laboratoire de P. Cramer. En effet, une partie des *crosslinks* entre le Médiateur et la Pol II ne correspond pas à ceux obtenus dans le modèle de PIC partiel. Dans le modèle de P. Cramer, l'absence du module de queue et de la sous-unité Med1 peut expliquer les différences observées. En effet, les données de microscopie électronique montrent que le module de queue joue un rôle important dans la réorganisation du Médiateur lorsqu'il se lie à la Pol II. De plus, les activateurs, absents du modèle, pourraient aussi être une source de flexibilité et expliquer ces différences entre les modèles. Il est aussi possible que certaines conformations proposées ne soient pas totalement celles qui existent *in vivo*. La purification du Médiateur en dehors du contexte cellulaire favorise probablement la stabilisation de conformations particulières selon les conditions expérimentales utilisées. De plus, la très grande flexibilité du Médiateur explique probablement le fait que ces analyses structurales ne donnent pas les mêmes conclusions. Enfin, les modèles obtenus à partir des expériences de microscopie électronique sont basés sur l'analyse d'un faible pourcentage des images acquises et pourraient ne représenter qu'une partie des conformations. Bien que les modèles s'améliorent et soient pour certains de plus en plus cohérents avec des données biochimiques, il faudra probablement d'autres analyses et améliorations des protocoles de préparation des échantillons pour fournir un modèle structural plus robuste. Malgré tout, ces modèles et les données

biochimiques et fonctionnelles illustrent bien le rôle essentiel de l'interaction entre le Médiateur et la Pol II dans l'activation de la transcription.

3) Interactions entre le Médiateur et les facteurs généraux de la transcription.

Le Médiateur est impliqué dans la transcription activée mais est également capable de stimuler la transcription basale *in vitro*. Ainsi, il est établi que le Médiateur stimule la formation du PIC et interagit avec certains GTFs (Takagi and Kornberg 2006).

a) Médiateur et TFIID.

TFIID est le premier GTF à intervenir dans l'assemblage du PIC *in vitro*. Un certain nombre d'expériences de biochimie et de génétique ont rapidement suggéré un lien entre le Médiateur et TFIID. En effet, dans des expériences d'assemblage de PIC *in vitro* à partir d'extraits nucléaires humain, il a été montré que les recrutements de TFIID et du Médiateur sont interdépendants à la fois dans des conditions activées et basales (Johnson et al. 2002; Baek et al. 2002). Ces observations suggèrent un rôle synergique entre le Médiateur et TFIID, chacun facilitant le recrutement de l'autre. Des expériences réalisées chez la levure *S. cerevisiae* ont permis de montrer des résultats semblables et de suggérer une interaction entre le Médiateur et TFIID (Koleske et al. 1992; Takahashi, Kasahara, and Kokubo 2009).

L'existence d'une interaction directe entre le Médiateur et TFIID a été observée par le laboratoire de R. Conaway et J. Conaway. En effet, l'analyse protéomique par MudPIT des protéines interagissant avec Med26 et des expériences biochimiques ont révélées un contact direct avec TFIID chez l'Homme (Takahashi et al. 2011). Ce contact n'est pas requis pour le recrutement de TFIID mais important pour réguler la transition de l'interaction entre Med26 et d'autres protéines impliquées dans l'élongation. La résolution de la structure du sous-complexe Med8-Med18-Med20 et sa caractérisation fonctionnelle ont permis de mettre en évidence un contact direct entre Med8 et la sous-unité TBP de TFIID chez la levure (Larivière et al. 2006). Il semblerait néanmoins que ce contact implique plus généralement le sous-module Med8-Med18-Med20. Il est proposé que cette liaison avec la TBP nécessite le contact entre le Médiateur et un activateur. Cela permettrait un changement de conformation plus propice à l'interaction avec la TBP. La modélisation de l'interaction entre la TBP et le Médiateur à partir des données de microscopie électronique est cohérente avec ces interactions puisque l'on observe un placement de la TBP à l'interface des sous-modules Med11-Med17-Med22 et Med6-Med8-Med18 (Cai et al. 2010). De plus, dans cette étude du module de tête par microscopie électronique, on observe que celui-ci est flexible et peut être dans différentes conformations. L'ajout de TBP induit une transition significative dans une conformation plutôt ouverte. Il est donc suggéré que la liaison du module de tête à la TBP induit un changement conformationnel pouvant favoriser ensuite le contact avec la Pol II.

b) Médiateur et TFIIA.

A ce jour, aucune interaction directe n'a été mise en évidence entre le Médiateur et TFIIA. Il a néanmoins été montré que le Médiateur pouvait former un complexe stable avec TFIIID et TFIIA (DAMed), nécessaire à la transcription activée *in vitro* (Johnson and Carey 2003). Ce complexe est également compétent pour permettre la formation d'un PIC même en l'absence d'activateur. Il est proposé qu'il forme une plateforme permettant le recrutement des autres GTFs et de la Pol II.

c) Médiateur et TFIIIB.

Des expériences de biochimie ont permis de montrer que les modules de tête et de milieu, lorsqu'ils sont isolés du reste du complexe, chez *S. cerevisiae*, peuvent interagir avec TFIIIB *in vitro* (Kang et al. 2001). La structure du complexe Pol II-TFIIIB étant connue (Sainsbury, Niesser, and Cramer 2013), il a été possible de modéliser sa présence au sein du complexe Pol II-Médiateur. Il apparaît que TFIIIB se situe dans l'espace situé entre le Médiateur et la Pol II, à proximité des modules du milieu et de tête (Cai et al. 2010). Les travaux de P. Cramer concernant la structure du PIC partiel présenté plus tôt proposent une interaction entre Med18-Med20 et le *B-ribbon* de TFIIIB (Plaschka et al. 2015). Ce modèle ne montre cependant pas de contacts avec le module du milieu, comme suggéré par les données biochimiques et les modélisations.

Enfin, le Médiateur facilite le recrutement de TFIIIB au sein du PIC et stimule de ce fait l'activation de la transcription *in vitro*. Dans ces expériences de transcription, l'ajout d'un excès de TFIIIB peut même contrebalancer l'absence de Médiateur humain (Baek, Kang, and Roeder 2006)

d) Médiateur et TFIIIF.

La Pol II et TFIIIF forment un complexe stable et sont recrutés ensemble au sein du PIC, selon le modèle d'assemblage *in vitro*. La démonstration d'une interaction entre le Médiateur et TFIIIF n'a jamais été faite. Néanmoins, il est très intéressant de noter que TFIIIF permet de stabiliser l'orientation de la Pol II lorsqu'elle est liée à un complexe Médiateur-activateur (Bernecky et al. 2011). De plus, il a été observé que la Pol II ne pouvait se lier au Médiateur de façon stable qu'en présence de TFIIIF (Takagi et al. 2006). Cette observation est d'ailleurs une explication au fait que dans les modèles de structures du complexe Pol II-Médiateur obtenus par microscopie électronique chez la levure, différentes orientations de la Pol II puissent être trouvées. En effet, dans ces structures, TFIIIF est absent de l'assemblage et ne fixe pas une conformation stable. Ces données suggèrent qu'il existe peut-être un contact entre le Médiateur et TFIIIF n'ayant pas encore été identifié. Une autre explication serait que la liaison de TFIIIF sur la Pol II peut induire une conformation de la Pol II favorisant le contact avec le Médiateur et régulant l'orientation de cette interaction.

e) Médiateur et TFIIIE.

Des études de biochimie et de génétique ont mis en évidence une interaction entre la sous-unité Med15 du Médiateur et les deux sous-unités de TFIIIE (Sakurai et al. 1996). Il est également démontré que, *in vitro*, le module du milieu du Médiateur isolé interagit avec la grande sous-unité de TFIIIE chez *S. cerevisiae* (Kang et al. 2001). Ces contacts sont probablement impliqués dans le fait que le Médiateur facilite le recrutement de TFIIIE dans des expériences de transcription *in vitro* (Baek, Kang, and Roeder 2006). De plus, il est connu que TFIIIE et TFIIH interagissent entre eux (Maxon, Goodrich, and Tjian 1994). Le Médiateur pourrait ainsi favoriser les recrutements de TFIIIE et TFIIH. Notre laboratoire a par ailleurs montré que le Médiateur peut influencer le recrutement de TFIIIE dans le PIC *in vivo* (Esnault et al. 2008). Il a cependant été montré que lorsque l'on utilise certains activateurs, TFIIIE peut être dispensable. Dans cette situation, TFIIH est donc recruté indépendamment par le Médiateur (Hiroshi Sakurai and Fukasawa 2003).

f) Médiateur et TFIIH.

Dès la découverte du Médiateur, il a été établi que celui-ci pouvait se lier au CTD de la Pol II. TFIIH étant le GTF responsable de la phosphorylation du CTD, permettant le passage en élongation, un lien entre le Médiateur et TFIIH a été rapidement suggéré. En accord avec cette hypothèse, des expériences de biochimie ont montré que le Médiateur était capable de stimuler la phosphorylation du CTD (Kim et al. 1994). Enfin, le Médiateur par son rôle d'échafaudage est capable de stimuler le recrutement et la stabilisation de TFIIH au sein du PIC (Guidi et al. 2004; Nair, Kim, and Myers 2005). Plus tard, notre laboratoire a mis en évidence par des expériences de double-hybride et de co-immunoprécipitation un contact entre Med11, sous-unité de la tête, et Rad3, sous-unité du module cœur de TFIIH (Esnault et al. 2008). De façon intéressante, nous avons montré que ce contact est fonctionnel puisque des mutations de Med11 affectent l'occupation de TFIIH *in vivo* et ont une influence sur les voies d'assemblages du PIC (Voir Chapitre II.III.6). Des études structurales accomplies par le laboratoire de P. Cramer sont venues confirmer ce résultat (Seizl et al. 2011). Enfin, récemment, une étude *in vitro* chez l'Homme a montré que la liaison du Médiateur à l'activateur p53 induit un changement conformationnel important pour l'activation de la Pol II et de TFIIH. Ceci confirme le rôle du Médiateur dans la stimulation de la phosphorylation du CTD (Meyer et al. 2010).

4) Interaction du Médiateur avec les co-activateurs.

En plus des interactions avec les facteurs généraux de la transcription, le Médiateur est capable d'interagir avec certains co-activateurs (Figure 26). Les co-activateurs peuvent être catégorisés en deux types : ceux agissant sur la chromatine pour en modifier les propriétés et ceux interagissant avec la machinerie de transcription pour en stimuler le fonctionnement.

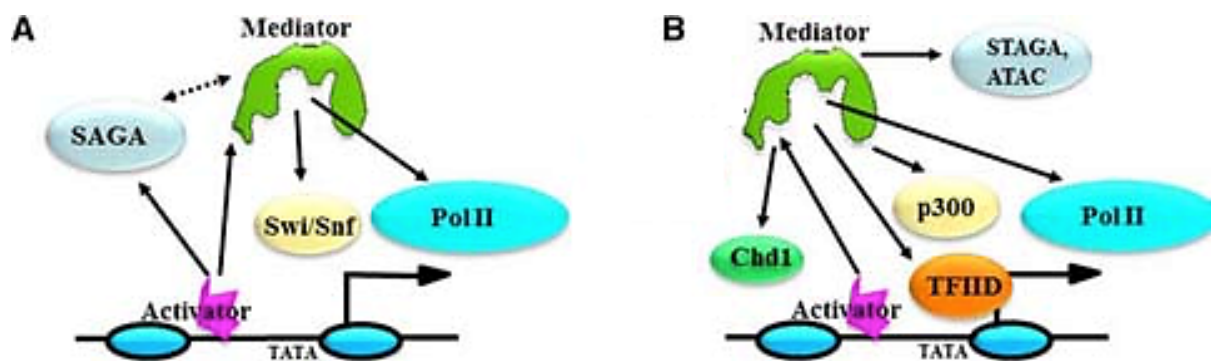


Figure 26: Schéma représentant les interactions entre le Médiateur et certains co-activateurs (d'après Ansari and Morse 2013)

(A) interactions chez *S. cerevisiae* (B) interactions chez l'Homme.

Parmi les co-activateurs agissant sur la chromatine, le Médiateur est impliqué dans des interactions avec SAGA et Swi/Snf chez la levure *S. cerevisiae* (Yoon et al. 2003). En effet, il a été montré *in vivo* par des expériences de CHIP que le recrutement de Swi/Snf par l'activateur Gcn4 était dépendant de la présence du Médiateur, en particulier des sous-unités Med15, Med2 et Med19, et de SAGA. Il est intéressant de noter que sur certains gènes, SAGA et le Médiateur sont recrutés de façon coopérative et synergique par ce même activateur (Qiu et al. 2005). Il a été montré par ailleurs que le Médiateur pouvait recruter Swi/Snf indépendamment de la présence d'un activateur. Dans une expérience où Med15 est fusionné au domaine de liaison à l'ADN de l'activateur Gal4, on observe un recrutement de Swi/Snf (Lemieux and Gaudreau 2004). Il n'a cependant pas à ce jour été montré d'interaction entre les sous-unités du Médiateur et de Swi/Snf.

Chez les métazoaires, plusieurs études ont été menées pour établir des liens entre le Médiateur et les co-activateurs. Par exemple, il a été montré une interaction physique et fonctionnelle entre le Médiateur et STAGA, un variant de SAGA possédant lui aussi une activité histone acétyl-transférase (HAT) (Liu et al. 2008). Les deux complexes interagissent *in vitro*. Lors de l'activation par l'oncoprotéine Myc, STAGA recrute le Médiateur afin de promouvoir la prolifération cellulaire. De la même manière, un recrutement coopératif du Médiateur et de p300 a été mis en évidence en réponse à l'activation par HNF4 (Malik et al. 2002). La protéine p300 permet l'acétylation des histones H3 et H4, marques associées à une transcription active. Ces résultats ont suggéré un lien fonctionnel entre le Médiateur et p300. Plus tard, une interaction directe entre p300 et le Médiateur a été montrée *in vitro*. Dans un système reconstitué activé par Gal4-VP16, le Médiateur et p300 sont recrutés de façon coopérative sur une matrice chromatinienne. L'accès au promoteur est masqué dans ces conditions. Les auteurs ont rajouté de l'acétyl-CoA (acétyl-CoEnzyme A) au système pour permettre l'apport d'une source de fonction acétyl. On observe alors une acétylation des histones par p300, et ainsi un relâchement de la structure chromatinienne. TFIIID peut alors être recruté et la transcription activée (Black et al. 2006).

De la même manière, un complexe entre le Médiateur et l'HAT ATAC a été observé chez la souris et montré comme impliqué dans la régulation d'un groupe de gènes (Krebs et al. 2010). Chez la souris également, des travaux récents ont montré une interaction entre le Médiateur et le remodeleur de la chromatine Chd1 (Lin et al. 2011). Cette interaction se fait spécifiquement par la sous-unité Med1. En effet, la diminution d'expression de Med1 empêche l'interaction du Médiateur avec Chd1 alors que la diminution de Med23 n'a pas d'effet. Chd1 appartient à la famille des remodeleurs à chromodomaines. Cela signifie qu'il peut interagir avec les histones et en particulier avec la marque d'histone H3K4me3. Pour cette raison, Chd1 est enrichi au niveau des promoteurs actifs, où cette marque est particulièrement présente. Les résultats de cette étude

montrent que bien que Chd1 puisse être recruté par la présence de la marque H3K4me3, sa présence et son activité sont stimulées par son recrutement *via* le Médiateur. De façon cohérente, chez la drosophile, la délétion des chromodomaines de Chd1 ne perturbent pas son recrutement sur les promoteurs des gènes transcrits (Morettini et al. 2011). Ces observations montrent que le Médiateur, en plus de son rôle de plateforme pour la formation du PIC, peut aussi coordonner le recrutement des autres co-activateurs nécessaires à l'activation de la transcription.

5) Interaction du Médiateur avec d'autres protéines impliquées dans la transcription.

a) Interaction entre le Médiateur et la cohésine.

Les régions régulatrices de l'ADN situées sur le promoteur distal doivent être mises en contact avec les régions proximales abritant le promoteur cœur où se déroule l'assemblage du PIC. Pour cela, l'ADN forme des boucles (*DNA looping*). Des protéines architecturales de la chromatine interviennent pour moduler la topologie de la chromatine. La cohésine est un complexe conservé chez les eucaryotes qui est impliqué dans la formation de ces boucles (Dorsett 2011). Un lien entre le Médiateur et la cohésine a été suggéré. En effet, la sous-unité Med1 est particulièrement impliquée dans la transcription activée par les récepteurs nucléaires. Il a été montré que l'extinction de l'expression de Med1, en plus d'affecter la transcription des gènes régulés par les récepteurs nucléaires, affecte également la présence de boucles réunissant les *enhancers* et les promoteurs (Park et al. 2005). Le laboratoire de R. Young a analysé par ChIP l'occupation de Médiateur dans les cellules ES de souris. Cette étude montre une présence de celui-ci à la fois sur les *enhancers* et sur les régions promotrices proximales. Cette observation confirme le fait que le Médiateur présente la capacité physique de faire le lien entre les activateurs et la machinerie de Pol II (Kagey et al. 2010). De façon intéressante, dans cette même étude, l'analyse du profil de liaison de la cohésine montre une colocalisation avec le Médiateur sur les *enhancers*. Les auteurs ont également évalué la fonctionnalité de cette co-occupation en inhibant l'expression de la cohésine ou du Médiateur. On observe en conséquence une diminution de l'expression de certains gènes. Cela montre donc que l'expression de certains gènes dépend de la présence à la fois du Médiateur et de la cohésine. Par ailleurs, l'interaction entre ces deux complexes a été montrée par des expériences de co-IP. Enfin, des expériences de 3C ont permis d'observer que seuls les gènes liés par le Médiateur et la cohésine présentaient des boucles reliant leurs *enhancers* à leurs promoteurs. L'ensemble de ces résultats montrent que le Médiateur, par son interaction avec la cohésine, peut réguler l'expression de gènes cible en permettant la formation de boucles.

b) Interaction entre le Médiateur et TFIIIS.

Comme indiqué plus tôt, TFIIIS est un facteur d'élongation qui interagit avec la Pol II. Son rôle dans l'étape d'élongation est de résoudre les pauses de la Pol II en permettant la reprise de la transcription. Pour cela, il stimule l'activité de clivage de la Pol II sur l'ARN naissant pour qu'il puisse

retrouver une position correcte dans le site catalytique de la Pol II.

Chez la levure, une interaction entre un domaine situé dans la partie C-terminale de TFIIIS (codé par le gène *DST1*) et le Médiateur a été montré par des expériences de double-hybride. Ce domaine n'est pas impliqué dans la liaison à la Pol II (Wery et al. 2004). De plus, des études d'interactions génétiques chez *S. cerevisiae* ont montré que les délétions conjointes de *dst1* et de *soh1* (codant pour la sous-unité Med31 du Médiateur) aboutissent à un phénotype synthétique létal (Malagon et al. 2004). Cela suggère ainsi un lien fonctionnel entre ces deux protéines. Plus tard, des expériences de recrutement du PIC *in vivo* sur le gène inductible *GAL1* ont montré que TFIIIS était recruté au niveau du promoteur (Prather, Larschan, and Winston 2005). La délétion de TFIIIS dans cette étude montre une diminution du recrutement des composants du PIC. Notre laboratoire a ensuite montré que TFIIIS pouvait être recruté sur les promoteurs de certains gènes (*ADH1*, *VTC3* et *MET17*) de façon indépendante de la Pol II (Guglielmi et al. 2007). Nos résultats suggèrent que le Médiateur, *via* sa sous-unité Med31, coopère avec TFIIIS pour favoriser le recrutement de la Pol II. Le domaine de liaison à la Pol II de TFIIIS est requis pour cette activité tandis que son domaine de stimulation du clivage est dispensable. En coopération avec le Médiateur, TFIIIS pourrait stimuler le recrutement de la Pol II ou sa stabilisation grâce à son domaine d'interaction avec l'enzyme. Au même moment, le laboratoire de J. Ranish a montré de façon cohérente, grâce à des expériences *in vitro*, que TFIIIS était un composant du complexe de pré-initiation (Kim et al. 2007).

6) Rôle du Médiateur dans l'assemblage du PIC *in vivo*.

La mise en place du PIC a été très étudiée *in vitro* et on connaît aujourd'hui bien l'ordre d'assemblage des GTFs ainsi que les mécanismes moléculaires impliqués. Cependant, *in vivo*, ce processus est plus complexe puisqu'il faut tenir compte que la majorité des gènes voient leur expression régulée par des activateurs et qu'elle implique plusieurs co-activateurs dont le Médiateur. Néanmoins, les différentes interactions impliquant le Médiateur présentées ici laissent entrevoir comment il peut avoir un rôle dans l'activation de la transcription et sur l'assemblage du PIC. En effet, suite à la liaison de l'activateur sur la région régulatrice, des co-activateurs dont des remodeleurs de la chromatine peuvent intervenir pour modifier la composition en nucléosomes. Le Médiateur peut être ensuite recruté *via* ses interactions avec l'activateur et les autres co-activateurs. En subissant une transition de conformation suite à cette liaison, il peut stimuler le recrutement des facteurs généraux grâce à ses interactions avec ceux-ci et ainsi favoriser la formation du PIC et le stabiliser (Figure 27). La méthode du ChIP, en permettant l'analyse de la liaison des protéines à l'ADN *in vivo*, a apporté un grand nombre d'avancées permettant de proposer ce modèle. En effet, il a été possible grâce à ces expériences de montrer à quel moment le Médiateur intervient dans la mise en place du PIC et de quelle manière il peut l'influencer. Les gènes du type *GAL* chez *S. cerevisiae* sont un bon objet d'étude pour ces expériences. En effet, il s'agit de gènes inductibles par le galactose dont la régulation a été étudiée en détail. Dans une culture en absence de galactose, ces gènes ne sont pas exprimés. L'ajout de ce sucre comme source de carbone va permettre l'activation de ces gènes. Il devient alors possible de suivre par ChIP l'arrivée des différentes protéines de la machinerie de la transcription en fonction du temps après l'ajout de galactose. Ainsi, il a été montré qu'après l'induction, on observe rapidement la liaison de l'activateur Gal4 sur l'UAS qui va à son tour recruter le co-activateur SAGA puis le Médiateur. Suite à cela, ces deux co-activateurs vont permettre le recrutement de la Pol II et des GTFs (Bryant and Ptashne 2003). D'autres études de ChIP ont montrées que le Médiateur était recruté sur les régions régulatrices et ce de façon indépendante de la Pol II (Park, Werner, et al. 2001; Kuras,

Borggreve, and Kornberg 2003). Son recrutement permet ensuite l'assemblage du PIC.

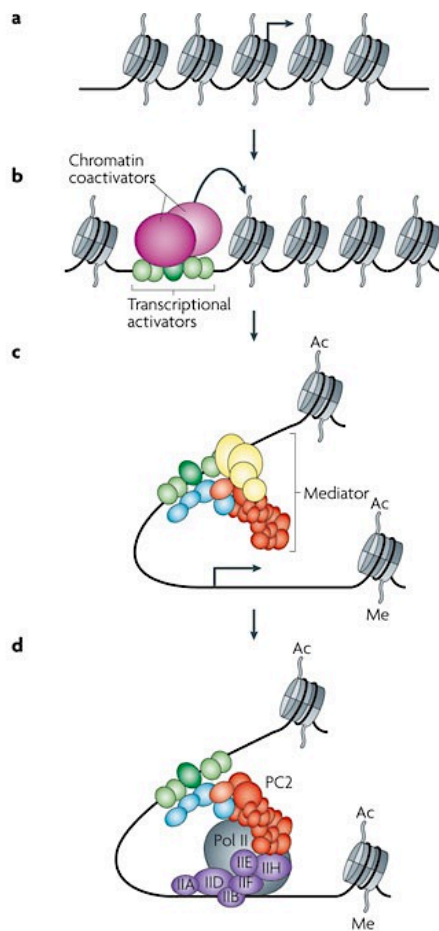


Figure 27: Schéma montrant l'activation de la transcription *in vivo* et les différentes interventions du Médiateur dans ce processus (d'après Malik and Roeder 2010)

(a): Le promoteur du gène à transcrire se trouve dans une structure chromatinienne dense, fermée à la transcription.

(b): un activateur spécifique est recruté sur la séquence activatrice. L'activateur recrute des co-activateurs permettant le remodelage de la chromatine et le dépôt de marques d'histones favorables à la transcription afin d'établir une région libre de nucléosomes.

(c): Le Médiateur est recruté par ses interactions avec l'activateur.

(d): Le Médiateur recrute la machinerie de Pol II afin de permettre la formation du PIC et sa stabilisation.

Les travaux de notre laboratoire ont permis de montrer que le Médiateur pouvait orchestrer la mise en place du PIC de façon différente de celle observé *in vitro*. Tout d'abord, comme évoqué plus tôt, nous avons identifié une interaction entre Rad3, sous-unité du module coeur de TFIIF, et Med11, sous-unité du module de tête du Médiateur (Esnault et al. 2008). L'analyse de mutants thermosensibles de Med11 a révélé une diminution de l'occupation de TFIIF, indiquant un rôle important du Médiateur dans le recrutement de TFIIF. En effet, certaines mutations de Med11 induisent un effet différentiel pour certains gènes sur le recrutement des modules coeur et kinase de TFIIF indépendamment de TFIIE et de la Pol II. Ces données ont permis de mettre en évidence un rôle du Médiateur dans le recrutement indépendant de TFIIF, TFIIE et la Pol II. Cela nous a aussi permis de montrer un rôle du Médiateur dans la stabilisation des modules coeurs et kinase de TFIIF. Nous avons également investigué le rôle de Med17, une autre sous-unité essentielle du module de tête en caractérisant des mutants thermosensibles (Eyboulet et al. 2015). L'étude de ces mutants a permis d'observer un recrutement indépendant des deux modules de TFIIF mais également de la TBP.

Ces deux études nous ont permis d'observer *in vivo* des intermédiaires dans l'assemblage du PIC que l'on n'observe pas *in vitro*. Cela nous a mené à proposer un modèle d'assemblage branché du PIC *in vivo* (Figure 28), contrairement au modèle linéaire et séquentiel observé *in vitro*. L'assemblage de ces différents intermédiaires est orchestré par le Médiateur par un réseau d'interaction entre ses sous-unités et les GTFs.

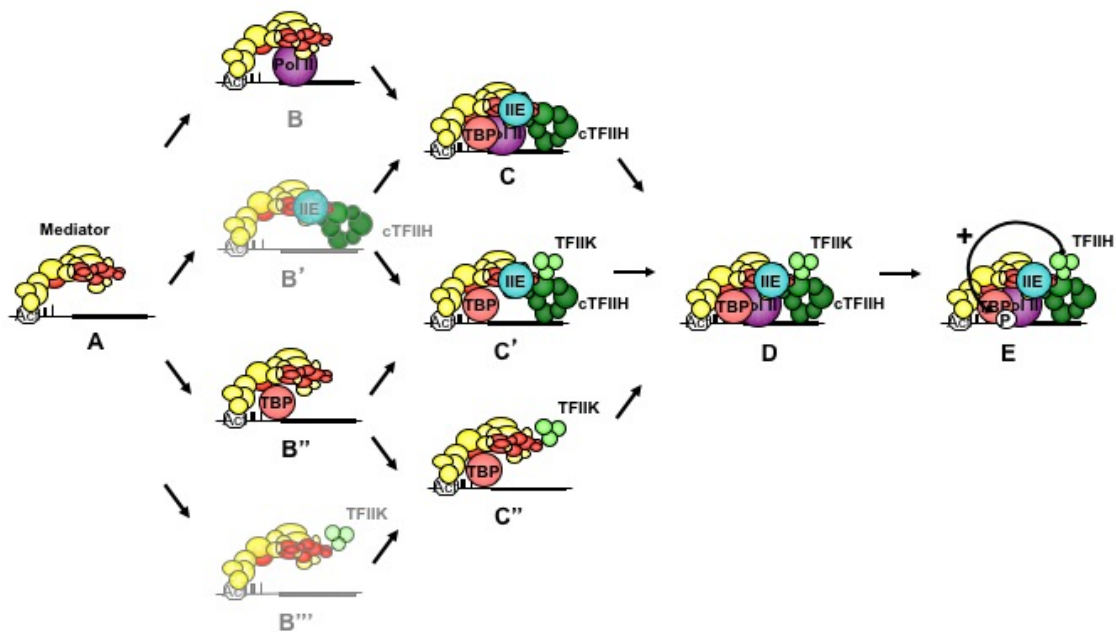


Figure 28: Modèle d'assemblage branché du PIC influencé par le Médiateur (Esnault et al. 2008; Soutourina et al. 2011; Eyboulet et al. 2015)

Les intermédiaires mis en évidence par nos études sont en couleurs vives, les intermédiaires suggérés sont en couleurs transparentes.

Le Médiateur étant impliqué dans la mise en place du PIC, sa localisation sur la chromatine pourrait être attendue sur la région promotrice des gènes. Comme mentionné, plus haut, les analyses de CHIP montrent la présence du Médiateur sur les éléments régulateurs, indépendamment de la Pol II et des GTFs. Les études de localisation génomique du Médiateur par CHIP-chip et CHIP-seq chez *S. cerevisiae* (Andrau et al. 2006; Eyboulet et al. 2015) ont montré une localisation globale sur les UAS. De plus, ces études ont mis en évidence que le Médiateur était recruté accompagné de son module Kinase même si sa présence dans le complexe n'est probablement que transitoire. La majorité des études s'accordent à proposer que le Médiateur soit principalement présent sur les UAS mais certains proposent néanmoins une localisation de celui-ci sur les promoteurs et également sur les régions codantes (Kuras, Borggreffe, and Kornberg 2003; Andrau et al. 2006). Concernant cette présence sur les régions codantes, cela a laissé imaginer un rôle potentiel du Médiateur dans les étapes d'élongation. Néanmoins, nos études de CHIP-seq ont montré que ces signaux d'occupations ne sont pas spécifiques. En effet, ces signaux sont majoritairement trouvés sur des gènes les plus activement transcrits et on peut les observer dans les contrôles négatifs de nos expériences (CHIP avec un anticorps non relevant ou CHIP avec la chromatine d'une souche ne portant pas l'étiquette contre lequel l'anticorps est dirigé). Nous pensons que ces signaux artéfactuels pourraient être dus à une structure très ouverte de la chromatine des gènes activement transcrits et ne sont pas spécifiques. D'autres laboratoires ont également observé la présence de ces artéfacts (Park et al. 2013; Teytelman et al. 2013; Jeronimo and Robert 2014). Par contre, pour ce qui est de la présence du Médiateur sur le promoteur, les travaux des laboratoires de F. Robert et de K. Struhl ont apporté un élément de réponse, en étudiant un mutant de Kin28 sensible à une drogue inhibitrice de son activité kinase (Jeronimo and Robert 2014; Wong, Jin, and Struhl 2014). Lorsque les cellules sont traitées par cette drogue, la phosphorylation du CTD sur Ser5 est pas inhibée et le départ en élongation est ralenti. On observe

dans cette situation une accumulation du Médiateur sur la région promotrice. Les auteurs proposent à partir de ces études que le Médiateur est préférentiellement recruté sur les UAS mais qu'il peut s'associer de façon transitoire au promoteur. Cette liaison est dépendante de la phosphorylation du CTD et donc de l'activité kinase de Kin28. Dans un contexte sauvage, la transition entre l'initiation et l'élongation médiée par Kin28 semble être un processus rapide, expliquant pourquoi l'association du Médiateur avec les promoteurs n'est pas facilement observée.

7) Le Médiateur en tant que facteur général de la transcription.

Les facteurs généraux de la transcription sont définis par leur caractère indispensable au recrutement de la Pol II sur le promoteur et pour ainsi permettre la transcription basale *in vitro*. Des études ont montrées que le Médiateur était également capable de stimuler cette activité basale à la fois chez la levure *S. cerevisiae* et chez l'Homme (Baek, Kang, and Roeder 2006; Takagi and Kornberg 2006). Les méthodes biochimiques de purification des protéines ainsi que la transcription *in vitro* ont été largement utilisées pour étudier la transcription basale. Néanmoins, *in vivo*, l'étude de cette transcription est difficile car il n'est pas possible de s'assurer de l'absence de tout activateur sous le contrôle duquel se trouverait un promoteur. Récemment, le laboratoire de L. Kuras a mis au point un système pour l'étudier (Lacombe et al. 2013). Ce système repose sur le recrutement artificiel de la machinerie de base de la transcription *in vivo* sans l'utilisation d'un activateur (*activator bypass*). Pour cela, les auteurs ont fusionné le facteur général de la transcription TFIIB à un domaine d'un facteur de transcription humain (RFX) n'ayant pas d'activité chez la levure. Ils ont ensuite inséré le site de recrutement de RFX en amont du promoteur du gène *MET17*, pour permettre le recrutement de TFIIB à cette position. Par cette méthode, il a été observé que le Médiateur est recruté, indépendamment de la présence d'un activateur, mais par ses capacités à lier les GTFs et la Pol II *in vivo*. Cela est également un argument suggérant un lien étroit entre TFIIB et le Médiateur. Ces résultats montrent ainsi que le Médiateur est un composant de la machinerie de base de la transcription *in vivo* et lui confère l'une des caractéristiques d'un facteur général de la transcription.

L'autre caractéristique des GTFs est le fait qu'ils soient requis pour la transcription de l'ensemble des gènes. Une étude fonctionnelle étudiant l'effet de mutations des gènes essentiels chez *S. cerevisiae* sur l'expression globale du transcriptome a montré le rôle essentiel du Médiateur. En effet, des mutations de *Med17* aboutissent à un effet négatif sur la transcription du même niveau que des mutations de la sous-unité Rpb1 de la Pol II (Holstege et al. 1998; Thompson and Young 1995). Ceci est également consistant avec l'étude de notre laboratoire ayant identifié le contact entre la Pol II et le Médiateur puisque la perte de cette interaction provoque une diminution globale de l'occupation des gènes par la Pol II et donc de l'activité transcriptionnelle (Soutourina et al. 2011). Enfin, il a été montré que le Médiateur était présent à la fois sur les régions promotrices des gènes inductibles et constitutifs (Ansari, He, and Morse 2009). L'ensemble de ces données procurent au Médiateur les caractéristiques d'un facteur général de la transcription. Bien qu'il ne soit pas considéré strictement comme un GTF, le Médiateur est un co-activateur impliqué de manière quasi-universelle dans l'expression des gènes *in vivo*.

IV. La régulation négative de la transcription par le Médiateur.

Le Médiateur est associé à l'activation de la transcription. Cependant, les analyses fonctionnelles de mutations de certaines sous-unités du Médiateur ont montré une augmentation de la transcription de certains gènes, indiquant ainsi une régulation négative par le Médiateur (Holstege et al. 1998; Koschubs et al. 2009; van de Peppel et al. 2005). Parmi ces sous-unités, on retrouve celles du module kinase. Le module kinase est composé des sous-unités Med12, Med13 et de la *cycline-dependant kinase* Cdk8 accompagné de la cycline C (CycC). La mutation abolissant l'activité kinase de Cdk8 abouti à une sur-expression de 3% des gènes, montrant ainsi son rôle represser (Holstege et al. 1998). Des mutations ou délétions des autres sous-unités ont montré également un rôle represser. Les mécanismes impliquant ce module dans la répression de la transcription ne sont pas totalement élucidés à ce jour mais quelques études ont apportés des indications. Ainsi, il a été observé que la kinase Cdk8 était capable de phosphoryler le CTD de la Pol II sur les Ser2 et Ser5 avant l'assemblage du PIC (Hengartner et al. 1998). Cela a pour effet une repression de l'initiation de la transcription. Un deuxième mécanisme, cette fois indépendant de l'activité kinase est aussi proposé. Tout d'abord, la purification du Médiateur montre que l'on peut le trouver dans deux états différents : l'un avec le module Kinase, l'autre sans (Liu et al. 2001; Sato et al. 2004). La présence ou l'absence de ce module est impliqué dans la régulation négative par le Médiateur. En effet, des expériences de microscopie électronique et de biochimie ont montré que la liaison du module kinase et de la Pol II sur le cœur du Médiateur sont mutuellement exclusives (Elmlund et al. 2006; Knuesel et al. 2009). Cette interaction entre le module cœur et le module kinase semble donc contrôler la capacité du Médiateur à recruter la Pol II. Il a été également observé que Cdk8 peut phosphoryler la sous-unité Med2 (van de Peppel et al. 2005). Cette modification induit une repression d'un ensemble de gènes impliqués dans la réponse aux carences en fer. Med2 faisant partie du module de queue, impliqué dans la liaison aux activateurs, il est possible que sa phosphorylation régule son interaction avec le/les facteurs de transcription contrôlant ces gènes. Chez les mammifères, il a été aussi montré que Cdk8 peut phosphoryler la cycline H de TFIIH et ainsi inhiber l'activité de TFIIH pour la phosphorylation du CTD (Akoulitchev, Chuikov, and Reinberg 2000).

Enfin, récemment il a été proposé que le module kinase puisse permettre l'autorégulation du Médiateur. La sous-unité Med3 peut en effet être phosphorylée par Cdk8. Cette modification permet l'interaction de Med3 avec l'ubiquitine ligase Grr1, aboutissant à sa dégradation (Gonzalez et al. 2014). La perte de cette sous-unité induit une diminution de l'activité co-activatrice du Médiateur sur certains gènes.

L'ensemble de ces données indique quelques mécanismes expliquant le rôle négatif du module Kinase sur la transcription. Cependant, des études ont montrées que le module kinase pouvait également stimuler la transcription. *In vitro*, le module kinase est important pour le recrutement de la Pol II dans la réponse à l'activation par le récepteur de l'hormone thyroïdienne (Belakavadi and Fondell 2010). Dans un autre cas, il a été rapporté que le module kinase est impliqué dans la régulation positive des gènes de réponse au serum chez l'Homme (Donner et al. 2010). Les auteurs proposent que le module kinase serait recruté après l'initiation afin de stabiliser le recrutement de P-TEFb et son assemblage avec la Pol II afin de permettre un passage en élongation rapide.

Récemment, le laboratoire de D. Taatjes à exploré, grâce à des méthodes de protéomique quantitative, les substrats des kinases Cdk8 et un homologue très proche Cdk19 (Poss et al. 2016). Les auteurs ont pour cela utilisé un inhibiteur spécifique de ces deux kinases et analysés le protéome des cellules quand ces kinases sont actives ou inactives. Cette étude a révélée que ces deux protéines phosphorylent environ 64 cibles. La plupart d'entre elles sont soit des sous-unités du Médiateur lui même, soit des facteurs impliqués dans la transcription, suggérant la possibilité d'une autorégulation de l'activité du Médiateur. La plupart des autres cibles sont des facteurs de

transcription. Il est ici donc proposé que les kinases Cdk8 et Cdk19 soient des protéines régulatrices des régulateurs de la transcription. Néanmoins, l'inhibition des activités kinase de Cdk8 et Cdk19 ne modifie l'expression que d'un petit ensemble de gènes. Cela montre que l'effet régulateur du module kinase ne repose pas uniquement sur l'activité enzymatique mais aussi sur la présence des autres sous-unités en permettant des interactions avec des partenaires régulateurs.

Le module kinase possède donc un double rôle represseur et activateur dans le Médiateur, dans un nombre restreint de gène. Cette dualité de fonction du module Kinase semble donc être une façon pour la cellule de moduler de façon spécifique et fine l'expression de certains gènes en faisant appel à ces mécanismes.

V. Rôles du Médiateur au-delà de l'initiation de la transcription.

Notre laboratoire a réalisé un crible double-hybride entre les sous-unités du Médiateur et une banque génomique afin de détecter les protéines interagissant avec le complexe. Ce crible a révélé un grand nombre de contacts entre le Médiateur et des protéines nucléaires dont la fonction n'est pas liée à l'initiation de la transcription. Cela laisse entrevoir un rôle du Médiateur dans des processus nucléaires autres que l'initiation de la transcription. En ce sens, des études ont permis de montrer que le Médiateur est impliqué dans les étapes *post*-recrutement du PIC et dans la maturation des ARNm, mais également dans d'autres processus tels que le maintien des télomères. Nous avons également récemment identifié un rôle du Médiateur dans la réparation de l'ADN.

1) Rôle du Médiateur dans les étapes de la transcription *post*-assemblage du PIC.

Chez les animaux, après l'initiation et le passage en élongation, la Pol II se retrouve en état de pause au niveau du promoteur proximal sur un grand nombre de gènes. La sortie de cet état fait appel au complexe de superélongation (SEC, *super elongation complex*) contenant les protéines de la famille ELL/EAF. Une étude chez l'Homme a montré un lien entre le Médiateur et le recrutement de ces facteurs (Takahashi et al. 2011). Ces travaux ont montré que la sous-unité Med26 interagit avec TFIID lors de l'assemblage du PIC et cette liaison est échangée avec une interaction avec ELL/EAF pour recruter le SEC à proximité de la Pol II. Le Médiateur est également impliqué dans les étapes d'élongation précoces par son lien avec TFIIIS. En effet, des expériences *in vitro* de transcription sur des matrices de chromatine ont montrées que le recrutement de TFIIIS par le Médiateur était important pour stimuler l'activité de la Pol II afin de passer outre la barrière du nucléosome +1 (Nock et al. 2012).

Comme évoqué plus tôt, le module kinase pourrait être également impliqué dans la stimulation du processus d'élongation. En effet, l'absence de Cdk8 est corrélée à une élongation ralentie et à une diminution de la phosphorylation des Ser2 et Ser5. Il est ainsi proposé que ce module permette le recrutement du facteur d'élongation P-TEFb contenant la kinase Cdk9 pour phosphoryler le CTD (Donner et al. 2010). De la même manière, chez l'Homme, Cdk8 stimule la sortie des pauses de la Pol II des gènes régulés par HIF1A en permettant le recrutement de P-TEFb et de SEC (Galbraith et al. 2013). Chez la souris il a été rapporté que des cellules délétées pour Med23 présentent un défaut d'expression du gène Med23-dépendant *EGR1*. L'analyse des profils

de liaison du Médiateur et de la Pol II sur le promoteur sont inchangés dans cette situation. Seules les occupations de P-TEFb et de la Pol II sur la région codante sont diminuées. Cela montre que la délétion de Med23 n'affecte que la phase de sortie de pause et d'élongation, en empêchant le recrutement de P-TEFb. Il a par ailleurs été montré dans cette étude une interaction entre Med23 et P-TEFb (Wang et al. 2013). L'ensemble de ces exemples montrent que le Médiateur, notamment le module kinase et la sous-unité Med23, peut avoir un rôle dans les étapes d'élongation précoces.

Chez la levure *S. cerevisiae*, il a été proposé que le Médiateur pouvait être impliqué dans l'étape de terminaison sur les deux gènes *CHA1* et *INO1*. En effet, une expérience de ChIP a localisé la sous-unité Med18 au niveau du terminateur de ces deux gènes (Mukundan and Ansari 2011). Lorsque *MED18* est délété, on observe un défaut de recrutement des facteurs de terminaison ainsi qu'une accumulation de la Pol II au niveau du terminateur. Il s'agit cependant que d'un exemple sur deux gènes inductibles et il n'est pas rapporté si ce mécanisme apparaît sur d'autres gènes. De plus, ici, les auteurs n'analysent que la présence de Med18 et pas d'autres sous-unités du Médiateur. Il n'est donc pas possible de savoir si le Médiateur entier intervient dans ce cas, sachant que les études à l'échelle du génome ne montrent pas d'occupation du Médiateur sur les terminateurs des gènes.

Enfin, des études par spectrométrie de masse ont montré un lien entre le Médiateur et des facteurs de maturation de l'ARN. En particulier, Med23 interagit avec HnRNP L, une protéine nécessaire à l'épissage des ARNs chez les mammifères. En plus de son interaction, Med23 régule l'épissage de certains gènes par HnRNP L. Les auteurs proposent que Med23 puisse contacter ce facteur de maturation depuis le promoteur (Huang et al. 2012).

2) Rôle du Médiateur dans la ré-initiation de la transcription.

L'analyse des profils d'occupation du Médiateur sur le génome montre qu'il n'est pas présent sur les régions codantes, à l'exception de quelques artéfacts présentés plus tôt. Partant de cette observation, il est donc très peu probable que le Médiateur suive la Pol II lorsque celle-ci s'échappe du promoteur. Le laboratoire de S. Hahn a développé un système *in vitro* afin d'étudier la composition du PIC après l'initiation (Yudkovsky, Ranish, and Hahn 2000). L'expérience est basée sur l'assemblage d'un complexe de pré-initiation sur une matrice d'ADN immobilisée contenant un promoteur et un site de liaison d'un activateur. Cette matrice est incubée en présence d'un activateur et d'un extrait nucléaire pour permettre l'assemblage du PIC. Les NTPs sont ensuite ajoutés au système pour permettre l'élongation. Les auteurs ont pu alors observer que le départ en élongation aboutissait à un départ de la Pol II, de TFIIB et de TFIIIF. De façon étonnante, les autres GTFs et le Médiateur restent présents sur le promoteur, stabilisés par l'activateur. Lorsque l'on rajoute sur ce « PIC partiel » un deuxième extrait nucléaire mais cette fois provenant d'une souche mutante pour les différents facteurs restés sur le promoteur, on observe que l'initiation est de nouveau possible. Les auteurs proposent ainsi que ces facteurs forment une plateforme (*scaffold*) pour permettre la ré-initiation. Cette plateforme permet une ré-initiation rapide aboutissant à des niveaux de transcription plus élevés que lors de la première initiation. L'existence de cette plateforme n'a cependant été démontrée qu'*in vitro* et son existence *in vivo* n'a jamais été mise en évidence. L'existence du phénomène de *bursting* transcriptionnel est néanmoins cohérent avec la possibilité d'une ré-initiation rapide et efficace de la transcription.

Une autre étude par cette méthode a été menée en utilisant des mutants nuls pour les sous-unités *med3* et *med16* (Reeves and Hahn 2003). On observe dans ces mutants que l'association du Médiateur avec la Pol II est impactée mais que la première initiation n'est pas affectée. Les ré-

initiations ultérieures sont par contre beaucoup plus faibles. Cela montre l'importance du Médiateur dans ces étapes de ré-initiation.

3) Rôle du Médiateur dans le maintien des télomères.

Les télomères sont les séquences d'ADN formant les extrémités des chromosomes. Leur rôle principal est la protection des chromosomes contre la dégradation par les exonucléases nucléaires qui peuvent digérer les extrémités d'ADN. Une digestion excessive menerait à la perte progressive des séquences des chromosomes et donc des gènes. Ainsi, la longueur des télomères est variable selon le stade de vieillissement des cellules. Les cellules vieillissantes tendent à posséder des télomères plus courts. Chez les eucaryotes, ces télomères sont constitués d'un grand nombre de répétitions d'un motif particulier. Chez la levure *S. cerevisiae*, on trouve entre 300 et 500 répétitions du motif [(TG₁₋₃)_n] (Walmsley et al. 1984) ainsi qu'une extension 3' simple brin d'une séquence riche en G, d'une longueur variable selon la phase du cycle cellulaire (Wellinger, Wolf, and Zakian 1993). Les gènes situés à proximité des télomères (régions *sub*-télomériques) ont une régulation particulière. En effet, ils peuvent être réprimés (*silencing*) de façon réversible selon un mécanisme appelé *telomeres position effect* (TPE). Ce phénomène est conservé chez tous les eucaryotes mais les mécanismes restent à l'heure actuelle peu compris (Ottaviani, Gilson, and Magdinier 2008). Il est néanmoins connu que la structure de la chromatine, notamment certaines marques d'histones, régulent le TPE. Deux enzymes sont importantes dans ce processus. Il s'agit de Sas2, une histone acétyl-transférase et de Sir2, une histone déacétylase (HDAC). Les activités antagonistes de ces deux enzymes permettent l'établissement d'un gradient d'acétylation des histones dans les régions *sub*-télomériques (Dang et al. 2009). Ce gradient définit la position de la chromatine active et inactive. Le niveau d'acétylation régulé par ces deux enzymes influence fortement la durée de vie des cellules. La chromatine active dans ces régions tend à induire un raccourcissement des télomères et ainsi à une diminution de la longévité cellulaire. Une étude récente a montré que le Médiateur était présent sur les régions hétérochromatiques *sub*-télomériques (Zhu et al. 2011). Lorsque la sous-unité Med5 est délétée, on observe une diminution de l'occupation du Médiateur sur ces régions ainsi qu'une augmentation de l'acétylation. Le TPE diminue alors puisque la répression de la transcription induite par l'acétylation est réduite. Une autre étude a montré que cette présence du Médiateur requiert le module de queue puisque seules des modifications de ce module induisent un effet (Peng and Zhou 2012). Il est proposé que le Médiateur peut, par sa présence dans les régions *sub*-télomériques, réguler la balance entre les enzymes Sir2 et Sas2. Le Médiateur peut donc réguler le TPE et ainsi la maintenance des structures télomériques. Le mécanisme d'action du Médiateur n'est pas encore connu. On ne sait pas, par exemple, comment il est recruté sur ces régions ni le lien fonctionnel avec Sas2 et Sir2.

4) Rôle du Médiateur dans la réparation de l'ADN.

Le crible double-hybride réalisé par notre laboratoire entre les sous-unités du Médiateur et la banque génomique de *S. cerevisiae* nous a permis d'identifier une interaction entre la sous-unité Med17 et la protéine Rad2 (Eyboulet et al. 2013). Rad2 est une endonucléase impliquée dans la réparation des dommages à l'ADN dans la voie de réparation par excision de nucléotides (*nucleotide excision repair*, NER). Cette voie de réparation NER est utilisée en réponse aux

dommages provoqués, par exemple, par les rayonnements UV. Ce type de rayonnements conduit à l'apparition de lésions de type dimère de pyrimidine (*cyclobutane pyrimidin dimers*, CPD ou 6,4 *photoproducts*). Une machinerie spécifique pour ce type de lésions intervient alors pour exciser le brin d'ADN contenant la lésion. Il existe deux voies d'action de la réparation par NER. L'une concerne l'intervention sur le génome entier (*global genome repair*, GGR). L'autre est couplée à la transcription (*transcription coupled repair*, TCR). Les deux voies diffèrent dans la façon dont la lésion est reconnue. Dans le cas de la TCR, la lésion est repérée lorsque l'ARN polymérase en élongation est bloquée par celle-ci. Les deux voies convergent ensuite. Rad2 intervient dans les deux voies en permettant la coupure du brin d'ADN en 3' de la lésion (Figure 29). Un rôle non enzymatique de Rad2 dans la reconnaissance des dommages au cours de la TCR est également proposé (Sarker et al. 2005). L'interaction entre le Médiateur et Rad2 a été confirmée par co-IP. Nous avons ensuite analysé le profil d'occupation de Rad2 par ChIP-seq et montré que Rad2 co-occupe les mêmes régions que le Médiateur, à savoir les UAS. Il y a même une forte corrélation de leur liaison sur l'ensemble du génome. Nous avons analysé les potentiels phénotype d'une délétion de Rad2. Seule une sensibilité aux UV a été observée, suggérant que l'interaction entre le Médiateur et Rad2 ne joue pas de rôle dans la transcription. De plus, nous avons vérifié que cette délétion n'engendrait aucune diminution de l'occupation de la Pol II et du niveau d'expression de certains gènes. Ensuite, nous avons utilisé notre collection de mutant de Med17 afin de regarder quels pouvaient être leurs effets sur cette interaction. De façon intéressante, nous avons observé que certains d'entre eux présentaient une sensibilité aux UV. De façon concomitante, ces mutants de Med17 UV-sensibles présentent une perte de l'interaction avec Rad2 et montrent une diminution de l'occupation de cette protéine sur l'ADN. Nous avons proposé que le Médiateur pouvait, *via* son interaction avec Rad2, induire son recrutement sur le promoteur. Rad2 peut ensuite suivre la Pol II pour faciliter la réparation d'éventuels dommages à l'ADN sur les gènes transcrits. Nos études actuelles sur le lien entre le Médiateur et la réparation de l'ADN nous indiquent que Rad2 n'est pas la seule protéine de la machinerie NER interagissant avec le Médiateur. De plus, nous avons montré par co-IP l'existence d'une interaction entre le Médiateur et XPG, l'homologue humain de Rad2, est conservée dans les cellules humaines. Une étude récente a également confirmé ce lien fonctionnel entre le Médiateur et Rad2/XPG chez l'Homme (Kikuchi et al. 2015).

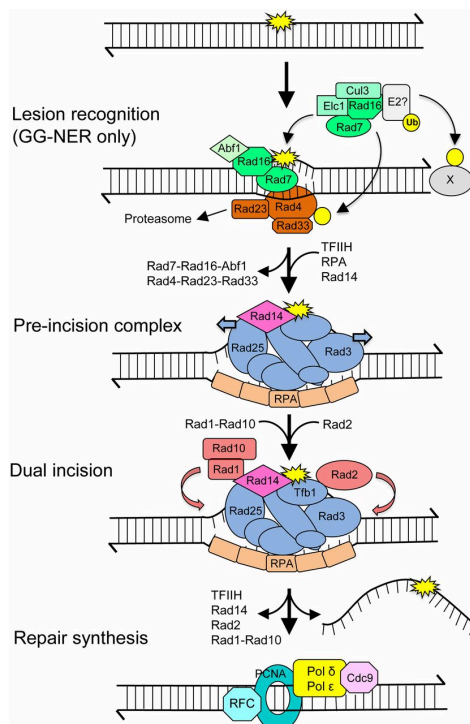


Figure 29: Schéma de la voie de réparation de l'ADN par excision de nucléotides (Boiteux and Jinks-Robertson 2013).

VI. Rôles du Médiateur dans des processus du développement.

Du fait de ses nombreuses interactions avec des activateurs spécifiques, certaines sous-unités du Médiateur sont impliquées dans des processus physiologiques chez les eucaryotes pluricellulaires et particulièrement dans le développement. A la différence de *S. cerevisiae*, il est très difficile d'introduire des mutations sur des gènes de façon ciblée chez ces organismes. Le développement récent des méthodes d'édition du génome avec l'utilisation du système CRISPR/Cas9 (*clustered regularly interspaced short palindromic repeats*) devrait permettre très rapidement de pouvoir insérer de façon aisée de mutations spécifiques dans ces organismes afin d'en étudier l'effet. En attendant, la majorité des études du rôle des sous-unités du Médiateur chez ces les animaux et les plantes repose sur la modulation de l'expression de la protéine d'intérêt en baissant ou augmentant son expression. Cela peut être fait par ARN interférence dans le cas des pertes de fonction ou par transfection du gène codant pour la protéine pour la surexprimer. Chez la souris, relativement propice à la recombinaison homologue, il est possible d'ingénierer des animaux dans lesquels le gène est délété (*knock out*, KO). L'utilisation de ces méthodes de génomique fonctionnelle a permis de montrer le rôle d'un certain nombre de sous-unités du Médiateur dans la physiologie chez l'Homme, la souris, la drosophile mais aussi le poisson zèbre et le ver *C. elegans*.

Ainsi, chez *C. elegans*, des expériences d'ARN interférence inactivant Med7, Med10 ou Med6 ont montré un rôle essentiel de ces sous-unités dans les étapes de développement précoce puisqu'elle régulent des gènes essentiels à l'embryogénèse (Kwon et al. 1999). De la même manière, des souris invalidées pour la sous-unité Med21 meurent au stade blastocyste, avant que les cellules souches embryonnaires n'aient pu apparaître (Tudor et al. 1999). D'autres souris KO pour des sous-unités du Médiateur ont ensuite été étudiées (Tableau 10). Toutes ces invalidations génétiques provoquent une létalité embryonnaire, bien que le stade auquel elles provoquent la mort varie (Yin and Wang 2014). Ainsi, l'invalidation de Med1 provoque la mort de l'embryon à 11,5 jours (E11.5). On observe sur ces embryons des malformations cardiaques, un défaut du développement neuronal, une nécrose hépatique et un défaut d'hématopoïèse (Ito et al. 2000). La délétion de *MED12* induit une mort à E7.5 avant la fin de la gastrulation (Rocha et al. 2010). Les souris KO pour la sous-unité Med23 meurent entre E9 et E10.5, en inhibant la transcription régulée par *Egr1* (*early growth response 1*) provoquant un défaut du système circulatoire (Balamotis et al. 2009). La délétion de *MED24* provoque une sévère hypoplasie cardiaque et des défauts du développement du système nerveux, provoquant la mort embryonnaire aux alentours de E10.5 (Ito et al. 2002). Enfin, les souris *CDK8* KO meurent avant même leur implantation intra-utérine (Westerling, Kuuluvainen, and Mäkelä 2007). Ces études d'invalidations montrent que les différentes sous-unités analysées ont un effet différent dans le développement, en impliquant des voies de régulation génétique différentes. Elles montrent également le rôle prépondérant du Médiateur dans les processus développementaux. L'implication de Med1 dans la formation des *super-enhancers* présentés plus haut est également importante dans le développement. En effet, ces supers structures régulatrices contrôlent le programme transcriptionnel des cellules souches embryonnaires et permettent donc de réguler la pluripotence (Whyte et al. 2013).

Les travaux menés chez la drosophile ont également permis de mettre en évidence le rôle de certaines sous-unités du Médiateur dans le développement. Par exemple, il a été montré que le module Cdk8 est important dans la régulation du développement de la patte, de l'œil et des organes sensoriels ainsi que dans la morphogénèse des segments (Boube et al. 2000; Loncle et al. 2007). Les sous-unités Med15 et Med31 ont elles été associées au développement des ailes et à la définition de l'axe antéro-postérieur, respectivement (Terriente-Félix et al. 2010; Bosveld et al. 2008).

<i>Sous-unité</i>	<i>Phénotype observé/implication</i>	<i>Organisme</i>
<i>Med1</i>	KO létal à E11,5 Adipogénèse, érythropoïèse Différenciation des glandes mammaires Déférenciation des cellules luminaire Maintient de l'état de pluripotence des cellules ES	Souris Souris Souris Souris Souris
<i>Med10</i>	Développement embryonnaire, Médiation de la voie Wnt	Poisson zèbre
<i>Med12</i>	KO létal à E7,5 Maitient des cellules ES Différenciation des neurones Développement rénal et sensoriel, développement du cerveau, de l'endoderme et du cartilage Développement de la rétine et des ailes Développement de la vulve Division des cellules T	Souris Souris Souris Poisson zèbre Drosophile <i>C. elegans</i> <i>C. elegans</i>
<i>Med13</i>	Developpement de la rétine et des ailes Développement de la vulve Divison des cellules T Développement du cœur et du cerveau	Drosophile <i>C. elegans</i> <i>C. elegans</i> Homme
<i>Med14</i>	Adipogénèse	Souris
<i>Med15</i>	Développement des ailes Développement du mésoderme Métabolisme des lipides Développement du cœur	Drosophile Xenope Homme Homme
<i>Med19</i>	Développement des neurones Développement des oligodendrocytes	Souris Homme
<i>Med21</i>	KO létal au stade blastocyste Différenciation des kératinocytes	Souris Souris
<i>Med23</i>	KO létal à E9-10,5 Adipogénèse Différenciation du muscle lisse Développement de la vulve	Souris Souris Souris <i>C. elegans</i>
<i>Med24</i>	KO létal à E8-10,5 Différenciation des cellules de la rétine Adipogénèse	Souris Poisson zèbre Homme
<i>Med25</i>	Chondrogénèse Sécrétion d'insuline Métabolisme des xénobiotiques et lipides Développement du système nerveux	Souris Homme Homme Homme
<i>Med26</i>	Différenciation neuronale	Souris
<i>Med27</i>	Différenciation des cellules de la rétine	Poisson zèbre
<i>Med28</i>	Différenciation du muscle lisse	Souris
<i>Med31</i>	KO létal à E16,5-18,5 Chondrogénèse Formation de l'axe antéropostérieur	Souris Souris Drosophile
<i>Cdk8</i>	KO létal à E2,5-3 Développement de l'œil Réponse immunitaire Différenciation cellulaire	Souris Drosophile Homme Homme

Tableau 10:Tableau récapitulatif des fonctions connues des sous-unités du Médiateur dans des processus physiologiques et développementaux (d'après Yin and Wang 2014)

Chez le poisson zèbre, la sous-unité Med12 est impliquée dans le développement cérébral, la chondrogénèse, le développement des yeux, des reins et de l'endoderme (Hong et al. 2005; Rau, Fischer, and Neumann 2006; Shin et al. 2008).

L'interaction des sous-unités du Médiateur avec des activateurs spécifiques implique le complexe dans des processus physiologiques tels que le métabolisme des acides gras, du cholestérol et des xénobiotiques (Yin and Wang 2014). Il est également impliqué dans la différenciation et la prolifération cellulaire. Ainsi Med1 et Med21 ont un rôle stimulateur sur la différenciation des kératinocytes ainsi que leur prolifération. Au contraire, Med10 possède un rôle inhibiteur (Oda et al. 2007; Oda et al. 2010). Med1 est également impliquée dans la différenciation des myélocytes au cours de l'hématopoïèse (Urahama et al. 2005).

VII. Pathologies associées au Médiateur.

Le Médiateur occupe une place centrale dans la régulation de l'expression des gènes impliqués dans les processus développementaux et la prolifération cellulaire. Il apparaît alors que des anomalies dans la fonction du Médiateur peuvent avoir de graves conséquences sur la physiologie et causer des pathologies. Bien que longtemps suspectée, l'association du Médiateur à des pathologies humaines n'est qu'assez récente, grâce aux progrès du séquençage et du génotypage. Ainsi, des défauts d'expression ou des mutations de certaines sous-unités du Médiateur ont pu être liées à des maladies du développement ou des cancers.

1) Maladies du développement.

La régulation des processus du développement est une fonction importante du Médiateur. Il a été expliqué plus haut (Tableau 9) que le Médiateur était impliqué dans le développement notamment du système nerveux central et périphérique. Des pathologies du développement du système nerveux ont ainsi été associées à des défauts du Médiateur (Tableau 11).

Ainsi, des mutations de la sous-unité Med12 causent les syndromes d'Opitz-Kaveggia et de Lujan (Risheg et al. 2007; Schwartz et al. 2007). Ces syndromes provoquent notamment des retards mentaux, une petite taille, une macrocéphalie et une dysmorphie des membres. Il a ensuite été montré que ces mutations de Med12 affectent une interaction entre le Médiateur et un répresseur de la différenciation neuronale appelé REST. Cela aboutit à l'expression anormale de gènes neuronaux induisant le syndrome. La sous-unité Med12 est également liée à des pathologies psychiatriques telles que la schizophrénie et des psychoses (Spaeth, Kim, and Boyer 2011). Un autre syndrome dû à des défauts du développement du système nerveux est lié au Médiateur. Il s'agit de la maladie de Charcot-Marie-Tooth. Cette pathologie, une neuropathie sensorielle et motrice, peut être causée par des mutations de plusieurs gènes, dont le gène *MED25* (Leal et al. 2009). Il semblerait que la mutation identifiée perturbe l'interaction entre Med25 et des kinases possédant un domaine SH3 (Abl et Lck) importante dans le développement du système nerveux sensoriel et périphérique.

Une mutation (S371P) de la sous-unité Med17 cause une atrophie cérébrale infantile (Kaufmann et al. 2010). Cette mutation non-sens semble affecter l'expression de gènes impliqués dans le développement des oligodendrocytes. Lorsque la mutation est transposée chez la levure en mutant le résidu équivalent (M504P), on observe un phénotype thermosensible à 37°C. Dans notre laboratoire, nous avons investigué l'effet de cette mutation chez la levure. Nous avons

observé qu'elle affecte le contact entre Med17 et la Pol II et provoque une diminution globale de l'occupation de la Pol II (Eyboulet et al. 2015).

Le syndrome de DiGeorge, est une maladie congénitale parmi les plus communes (Spaeth, Kim, and Boyer 2011). Elle cause des phénotypes très divers pouvant toucher plusieurs organes. On retrouve, par exemple, comme signes cliniques des défauts cardiaques, des anomalies palatines, des dysmorphies faciales, des insuffisances immunitaires et des maladies neuro-psychiatriques (schizophrénie, psychoses). Ce syndrome est dû à des délétions chromosomiques dont celle du gène *MED15* (Berti et al. 2001). Il semble que la maladie trouve son origine dans les étapes très précoces de migration et de différenciation des cellules au cours de la formation de la crête neurale à l'état embryonnaire. La sous-unité Med15 du Médiateur y joue alors un rôle important dans les voies de signalisation du développement, notamment la voie SMAD2/3. Enfin, une mutation de la sous-unité Med23 a récemment été associée à une déficience intellectuelle non syndromique. La mutation identifiée (R617Q), diminue l'expression des gènes de réponse précoce au sérum (IEG, *immediate early genes*) c-Jun et c-Fos. Ces gènes sont impliqués dans le développement cérébral, l'apprentissage et la construction de la mémoire à long terme (Hashimoto et al. 2011).

Enfin, des défauts du Médiateur peuvent causer des pathologies du développement aussi au niveau cardiovasculaire. Comme évoqué, le syndrome de DiGeorge peut provoquer des défauts cardiaques. Des mutations de la sous-unité *MED13L* provoquent un syndrome de transposition des grandes artères (TGA). Cette pathologie est caractérisée par une inversion de la position des artères pulmonaires et de l'aorte au niveau du cœur (Muncke et al. 2003). L'incidence de cette pathologie est de 5 à 7% des cas de maladie congénitales du cœur. Cela en fait la malformation cardiaque cyanogène du nourrisson la plus fréquente.

<i>Sous-unité</i>	<i>Pathologie</i>	<i>Cause génétique</i>
<i>Med12</i>	Syndrome d'Opitz-Kaveggia (syndrome FG) Syndrome de Lujan Syndrome d'Ohdo Schizophrénie/psychose	Mutation R971W Mutation N1007S Mutation faux-sens Polymorphisme
<i>Med13L</i>	Transposition des grosses artères	Mutations E251G, R1872H, D2023G, translocation chromosomique
<i>Med15</i>	Syndrome de DiGeorge	Délétion chromosomique
<i>Med17</i>	Atrophie cérébrale infantile	Mutation L371P
<i>Med23</i>	Déficit intellectuel non syndromique	Mutation R617Q
<i>Med25</i>	Maladie de Charcot-Marie-Tooth	Mutation A335V
<i>Cdk19</i>	Décollement de rétine congénital, microcéphalie, déficience intellectuelle	Inversion chromosomique précicentrique

Tableau 11: Implication de sous-unités du Médiateur dans des pathologies humaines du développement (d'après Spaeth, Kim, and Boyer 2011).

Les degrés d'implications des sous-unités du Médiateur dans ces pathologies sont variables.

2) Implication du Médiateur dans les cancers humains.

Les cancers sont des pathologies dont l'une des caractéristiques principales est une dérégulation de la transcription. Partant de ce constat, il n'est pas surprenant que le Médiateur, en tant qu'intermédiaire clé dans la régulation de la transcription, puisse être impliqué dans ces maladies. Ces dernières années, un grand nombre de types de cancers ont pu être associés à des défauts du Médiateur (Tableau 12). Certaines associations entre le cancer et le Médiateur sont

seulement corrélatives. C'est-à-dire que l'on observe seulement un changement d'expression/une mutation d'une sous-unité du Médiateur dans la maladie. Cela n'indique pas que ces changements soient la cause de la maladie. Par exemple, très récemment, une étude à large échelle à partir de bases de données de transcriptomes provenant de différents types de tumeurs a été menée. Ce travail révèle que toutes les sous-unités du Médiateur présentent une variation d'expression (positive ou négative) dans certaines tumeurs par rapport à un tissu sain (Syring et al. 2016). Il apparaît même que l'expression des sous-unités puisse varier selon le stade de progression de la maladie (tumeur précoce, métastases). Cela ne permet néanmoins pas de connaître les mécanismes moléculaires à l'origine de la maladie ou bien même le lien de causalité entre ces changements d'expression et la maladie. Malgré cela, il existe un certain nombre de cas répertoriés et étudiés dans la littérature dans lesquels les sous-unités du Médiateur sont directement impliquées dans des cancers.

Le cancer de la prostate est l'un des cancers les plus courants chez l'homme et une cause importante de mortalité. La survie et la prolifération des cellules de cancers de la prostate est assurée par les gènes régulés par le récepteur nucléaire aux androgènes (*androgène receptor*, AR). Lors de la description du Médiateur des mammifères, il a été montré que celui-ci était important pour la régulation de la transcription par les récepteurs nucléaires, notamment l'AR. La sous-unité Med1 est une cible directe de ce récepteur. Une étude a ainsi montré que dans 50% des cancers de la prostate, on observe une surexpression de Med1 (Vijayvargia, May, and Fondell 2007). Cette surexpression n'est pas due à une amplification du gène *MED1* comme il peut être retrouvé dans d'autres cancers mais à une augmentation de l'expression. Il est avancé que cela peut être le résultat d'une phosphorylation excessive de Med1 par une voie de signalisation qui stabilise la protéine (Spaeth, Kim, and Boyer 2011). Les auteurs de cette étude ont ensuite utilisé l'ARN interférence pour diminuer l'expression des sous-unités Med1 et Med17 dans des cellules de cancer de la prostate afin d'observer l'effet sur la prolifération. Ils ont ainsi constaté que la prolifération androgène-dépendante et indépendante était réduite et que l'apoptose augmentait. De plus, l'inhibition de l'expression de Med1 et Med17 entraîne une diminution de l'expression de marqueurs spécifiques du cancer de la prostate tels que l'antigène prostatique spécifique (*prostatic specific antigen*, PSA) et la peptidase dérivée de la kallikreine (*kallikrein-related peptidase*). Cela montre que Med1 est requis pour la prolifération et la survie de cellules cancéreuses dans le cancer de la prostate. Bien qu'il semble que dans ces cancers, Med1 rende la prolifération AR-dépendante trop importante, le mécanisme précis de son rôle oncogène n'est pas totalement connu. Cela laisse néanmoins imaginer que Med1 pourrait être une cible thérapeutique potentielle de ce type de cancer. D'autres études ont rapportées un rôle d'autres sous-unités du Médiateur dans le cancer de la prostate. C'est le cas par exemple de Med24, Med14 et Med19, toutes trois impliquées dans la régulation de l'expression des gènes AR-dépendants (Wang et al. 2002; Cui et al. 2011).

Le cancer du sein est le cancer les plus fréquents chez les femmes. Il est notamment l'une des premières cause de mortalité par cancer chez les femmes (Jemal et al. 2011). Dans un grand nombre de cas, ces cancers sont considérés comme hormono-dépendants. Ces cancers sont dépendants de la prolifération induite par l'expression des gènes régulés par le récepteur aux oestrogènes (ER). La sous-unité Med1 est là encore impliquée dans la régulation de l'expression des gènes cibles de ER. Il a ainsi été montré que dans 50% des cancers du sein, Med1 est surexprimé (Zhu et al. 1999). Dans 50% de ces cas, donc 25% du total, on observe une amplification chromosomique du gène *MED1*, aboutissant à cette surexpression.

D'autres cancers du sein ont la caractéristique de surexprimer la protéine Her2. C'est le cas d'environ 30% des cancers du sein, qui sont généralement plus agressifs, invasifs et résistants aux chimiothérapies. L'expression de Her2 est régulée positivement principalement par le facteur de transcription ESX surexprimé dans ces cancers. Il a été montré qu'il existe une interaction entre Med23 et ESX (Asada et al. 2002). L'inhibition de cette interaction par un peptide masquant

l'interface d'interaction a pour effet une très nette diminution de la prolifération des cellules cancéreuses et augmente leur mortalité. Par ces deux exemples, on voit encore comment le Médiateur et ses sous-unités pourraient être cibles de traitements contre le cancer.

Enfin, un dernier exemple d'implication du Médiateur dans les cancer est celui du rôle de Cdk8 dans les cancers colo-rectaux (*colorectal cancer*, CRC). Le CRC est le deuxième cancer le plus fréquent chez la femme et le troisième chez l'homme (Jemal et al. 2011). Dans un grand nombre de ces cancers, on retrouve une activation constitutive de la voie de signalisation Wnt/ β -caténine, une voie impliquée dans le développement. Lorsqu'elle est trop active, cette signalisation aboutit à une prolifération anormale des cellules. Un crible d'ARN interférence pour identifier les gènes impliqués dans la régulation de la β -caténine et la prolifération des cellules du CRC a été réalisé (Firestein et al. 2008). Parmi les gènes identifiés, on retrouve *CDK8*, la kinase du Médiateur. Cdk8 est très souvent surexprimé dans les CRC. La réduction de son expression induit une diminution de la prolifération des cellules. Il semble donc que Cdk8, *via* son activité kinase est très impliqué dans la régulation de la voie de la β -caténine. Il est même aujourd'hui admis que Cdk8 est un oncogène.

Comme expliqué précédemment, le Médiateur est impliqué dans la formation des structures régulatrices de la chromatine appelées *super-enhancer*. La diminution, même faible, de l'occupation du Médiateur sur ces régions aboutit à une forte diminution de l'expression des gènes contrôlés par ces *super-enhancers*. Ainsi, dans le cas de cancers où le programme transcriptionnel pourrait être contrôlé par un *super-enhancer*, le ciblage thérapeutique d'une sous-unité du Médiateur pourrait être d'un grand intérêt. Cela laisse entrevoir la façon dont le Médiateur pourrait, dans le futur, être une cible thérapeutique spécifique, malgré son rôle ubiquitaire dans la régulation de la transcription.

<i>Sous-unité</i>	<i>Cancer associé</i>	<i>Cause génétique</i>
<i>Med1</i>	Cancer de la prostate et du sein	Surexpression
	Mélanome	Surexpression
	Cancer du poumon et du colon	Sous-expression
<i>Med4</i>	Cancer cervical	Sous-expression
	Rétinoblastome	Sous-expression
<i>Med12</i>	Léiomyome	Mutation de l'intron 1 et exon 2
	Cancer du sein	-
	Cancer du colon	Mutation de l'exon 2
	Cancer du poumon	Sous-expression
	Cancer de la prostate	Mutation
<i>Med13</i>	Cancer du cerveau	Motif d'édition altéré
	Cancer du sein	Amplification du gène
	Rétinoblastome	-
<i>Med14</i>	Cancer du sein	-
	Cancer de la prostate	Surexpression
<i>Med15</i>	Cancer du sein	-
	Cancer de la prostate	Hyperméthylation
<i>Med17</i>	Cancer de la prostate	Surexpression
<i>Med19</i>	Cancer de la vessie, du sein, du colon, du poumon, de l'estomac, de l'ovaire, de la prostate, du pancréas, de la langue	Surexpression
	Ostéosarcome	Surexpression
	Mélanome	Délétion chromosomique
<i>Med23</i>	Cancer du sein	-
	Cancer de l'estomac	Mutation
	Phéochromocytome	Sous-expression
	Cancer du poumon	Surexpression
	Cancer de la prostate et du sein	Surexpression
<i>Med24</i>	Cancer de la prostate et du sein	Surexpression
<i>Med28</i>	Cancer du sein, du colon et de la prostate	Surexpression
<i>Med29</i>	Cancer du pancréas	Surexpression
<i>Med30</i>	Cancer du sein	Surexpression
<i>Cdk8</i>	Cancer du colon	Surexpression
	Mélanome	Surexpression
<i>CycC</i>	Cancer du colon	Surexpression
	Carcinome hépatocellulaire	Surexpression
	Leucémie	Délétion chromosomique
	Ostéosarcome	Sous-expression
<i>Cdk19</i>	Myélome	Surexpression
	Ostéosarcome	Surexpression

Tableau 12: Implications de sous-unités du Médiateur dans ces cancers et causes génétiques associées (d'après Spaeth, Kim, and Boyer 2011; Schiano et al. 2014).

Les degrés d'implications des sous-unités du Médiateur dans ces pathologies sont variables.

Résultats et discussion

Partie I : Etude intégrative du rôle de la sous-unité Med10 dans l'initiation de la transcription.

I. Contexte de travail.

La partie introductive a permis d'entrevoir comment se déroule la transcription. Son initiation est bien caractérisée *in vitro*. Le modèle retenu est celui d'un assemblage séquentiel et linéaire des différents facteurs généraux de la transcription et de la Pol II pour former le complexe de pré-initiation sur le promoteur du gène à transcrire. Cet assemblage débute par le recrutement de TFIID, suivi de TFIIA et TFIIIB formant un complexe stable servant de plateforme au recrutement du reste du PIC. La Pol II accompagnée de TFIIIF est ensuite incorporée. Enfin, TFIIIE et TFIIH sont recrutés pour terminer l'assemblage du PIC (Buratowski et al. 1989). Ce modèle ne tient cependant pas compte des co-activateurs, dont le Médiateur, alors que ceux-ci influencent la formation du PIC *in vivo*.

Le Médiateur a été l'objet d'un grand nombre d'études ayant permis d'éclairer certaines de ses fonctions. Notamment, il est nécessaire, *in vitro*, à l'activation de la transcription en réponse aux activateurs spécifiques, il augmente la transcription basale et stimule l'activité kinase de TFIIH (Kim et al. 1994). Il a également été établi que le Médiateur interagit avec de nombreux partenaires de la machinerie de transcription. Ainsi, des contacts entre le Médiateur et la TBP, TFIIH et la Pol II ont été montrés à la fois *in vivo* et *in vitro* (Larivière et al. 2006; Esnault et al. 2008; Soutourina et al. 2011). Des expériences *in vitro* ont également montré des contacts entre TFIIIE et le Médiateur (Sakurai et al. 1996; Sakurai and Fukasawa 2003). Enfin, des études *in vitro* ont proposé une interaction entre TFIIIB et le module du milieu du Médiateur (Kang et al. 2001). Toutes ces données montrent que le Médiateur peut influencer l'assemblage du PIC en favorisant le recrutement ou la stabilisation des GTFs et de la Pol II par le biais de ce réseau d'interactions. Le Médiateur étant conservé chez tous les eucaryotes, il occupe une place centrale dans la régulation de la transcription dans toutes les espèces. Il est important de noter que le Médiateur est impliqué dans des pathologies humaines. La compréhension détaillée de son rôle pourrait donc permettre dans le futur d'en faire une cible thérapeutique.

Néanmoins, on ne connaît pas exactement tous les mécanismes moléculaires mis en jeu entre le Médiateur et ses partenaires dans le contrôle de l'expression de gènes. Cela est en grande partie dû à la complexité du Médiateur et en particulier à la présence de sous-unités essentielles, au nombre de 10 parmi les 25 sous-unités chez *S. cerevisiae*. Mon projet de thèse s'inscrit dans un projet plus large du laboratoire à savoir la compréhension des mécanismes moléculaires impliquant le Médiateur dans l'activation de la transcription à l'échelle du génome grâce à l'étude de ses sous-unités essentielles. En particulier, nous souhaiterions comprendre comment les interactions protéine-protéine au sein du Médiateur et entre le Médiateur et ses nombreux partenaires permettent d'intégrer les signaux de régulation.

Dans notre laboratoire, à l'aide d'un crible double-hybride, nous avons pu décrire le réseau d'interactions au sein du Médiateur (Guglielmi et al. 2004). Ce crible a permis d'identifier également des contacts entre certaines sous-unités du Médiateur et d'autres protéines nucléaires. L'une de ces interactions est celle entre Med11, une sous-unité essentielle du module de tête, et

Rad3, une sous-unité de TFIIF. Ce contact a été caractérisé en étudiant des mutants conditionnels de Med11. Ce travail a permis de montrer que le Médiateur favorise de façon indépendante le recrutement de la Pol II, de TFIIE et de TFIIF ainsi que la stabilisation du module kinase de TFIIF (TFIIFK) (Esnault et al. 2008). Notre laboratoire a également caractérisé le rôle de Med17, une autre sous-unité essentielle du module de tête. Il a ainsi été montré que Med17 interagit directement avec la Pol II (Soutourina et al. 2011) et que ce contact était requis pour la transcription de l'ensemble des gènes. L'analyse de mutants conditionnels de Med17 a également été menée. Ce travail a permis de montrer que le Médiateur pouvait orchestrer des voies multiples d'assemblages du PIC (Eyboulet et al. 2015). Ces trois études sur des sous-unités essentielles du Médiateur nous ont mené à proposer un modèle ramifié d'assemblage du PIC, différant fortement du modèle séquentiel proposé *in vitro* en absence de co-activateur (Figure 28).

Forts de ces succès, nous voulions poursuivre nos études des sous-unités essentielles du Médiateur. Mon projet de thèse a consisté en la caractérisation de la sous-unité essentielle Med10, appartenant au module du milieu dont on ne connaît pas bien le rôle. Le module du milieu est composé des sous-unités Med1, 4, 7, 9, 10, 14, 21 et 31. La majorité des connaissances existantes sont basées sur des études structurales et biochimiques. On sait notamment que le module du milieu, tout comme le module de tête, est engagé dans les contacts entre le Médiateur et le PIC (Davis et al. 2002; Plaschka et al. 2015; Kang et al. 2001). Il est également connu que le Médiateur est un complexe présentant une structure très flexible. Il peut, par exemple, changer de conformation suite à sa liaison avec des activateurs. Ces changements de conformations semblent très importants dans le fonctionnement du Médiateur. Le module du milieu est particulièrement impliqué dans ces transitions de conformation en formant une région charnière (Larivière et al. 2013; Robinson et al. 2015). Quant à la sous-unité Med10, elle est impliquée dans un grand nombre d'interactions au sein du module du milieu. Elle est aussi impliquée dans l'embryogénèse chez le poisson zèbre (Lin et al. 2007). Outre ces données, on ne connaît pas exactement le rôle fonctionnel de Med10 et du module du milieu, en particulier *in vivo*.

Afin d'étudier le rôle de la sous-unité Med10, nous avons appliqué la démarche mise en place pour l'étude de Med11 et Med17, à savoir l'obtention de mutants thermosensibles. Il existe désormais d'autres méthodes pour étudier des protéines essentielles chez la levure. Il s'agit des méthodes d'*anchor away* (Haruki, Nishikawa, and Laemmli 2008) et de dégron inductible (*auxin degron*) (Nishimura et al. 2009). Toutes deux sont basées sur la délétion conditionnelle de la protéine d'intérêt, soit en l'excluant du noyau soit en induisant sa dégradation. Néanmoins, l'utilisation de mutants thermosensibles (ts) présente l'avantage de pouvoir maintenir la protéine présente au sein du complexe et ne perturber qu'un aspect de sa fonction, par exemple en perturbant des interactions spécifiques et/ou sa structure. L'obtention de ces mutants ts se fait par mutagenèse aléatoire sur le gène codant pour la sous-unité d'intérêt. A mon arrivée au laboratoire, cette méthode nous avait permis d'obtenir un mutant ts du gène *MED10* que nous avons nommé *med10-196*. En séquençant le gène muté, nous avons constaté que la protéine présentait les substitutions : L53S, I79T, E82D et N108I.

La carte d'interaction entre les sous-unités du Médiateur (Guglielmi et al. 2004) indique que Med10 est situé au cœur d'un réseau d'interactions dans le module du milieu. Nous avons donc en tout premier lieu analysé par double-hybride les interactions entre Med10 et ses partenaires connus au sein du Médiateur, dans les contextes sauvage et mutant. Cela nous a permis de constater que les mutations de *med10-196* ne perturbait aucune interaction de Med10 avec ses partenaires, exceptée celle avec Med14.

L'intérêt de notre approche par des mutants ts comparée à celles de délétions conditionnelles est de pouvoir étudier un complexe entier si les mutations n'affectent pas la stabilité ou l'assemblage du complexe. Nous avons donc souhaité vérifier, en analysant la composition du Médiateur, que la mutation *med10-196* n'affectait pas sa stabilité. Pour cela, nous avons mis en place une méthode inédite au laboratoire. Nous avons purifié le Médiateur dans les contextes sauvage et mutant et analysé leur composition par spectrométrie de masse (MS). Ces analyses ont révélé que toutes les sous-unités des modules de tête, milieu et queue identifiées dans le

Médiateur de la souche sauvage étaient présentes dans le Médiateur mutant, indiquant que la mutation *med10-196* n'affecte pas l'assemblage ou la stabilité du complexe.

L'intégrité du complexe étant vérifiée, nous avons réalisé une caractérisation fonctionnelle du mutant *med10-196*. Nous souhaitions savoir si la mutation *med10-196* perturbait l'interaction entre le Médiateur et les GTFs. Pour cela, nous avons réalisé des expériences de co-IP. Nous avons constaté que les GTFs TFIIA et TFIIB co-immunoprécipitent avec le Médiateur. De manière intéressante, nous avons remarqué que dans le cas du mutant *med10-196*, l'interaction entre le Médiateur et TFIIB était diminuée contrairement à celle avec TFIIA, restée inchangée. L'interaction entre le Médiateur et TFIIB laissait imaginer un lien fonctionnel entre le Médiateur et ce GTF. Nous avons combiné la mutation *med10-196* avec différents mutants de TFIIB afin d'observer les phénotypes résultants. Nous avons ainsi mis en évidence des phénotypes synthétiques létaux pour trois mutants de TFIIB lorsqu'ils étaient combinés avec *med10-196*. Cela a permis de confirmer le lien fonctionnel, et cette fois *in vivo*, entre TFIIB et le Médiateur.

Nous souhaitions préciser par quelle(s) sous-unité(s) du Médiateur l'interaction avec TFIIB se faisait. Pour cela, nous avons réalisé des expériences de double-hybride entre les sous-unités du module du milieu et TFIIB. Nous avons ainsi constaté un contact entre Med14 et TFIIB et dans une moindre mesure entre Med10 et TFIIB. Nous avons également testé la combinaison des mutations de TFIIB avec des mutations de Med14. Là encore, nous avons observé des phénotypes synthétiques létaux. La létalité observée a été beaucoup plus massive qu'avec le mutant *med10-196*. Cela est cohérent avec le fait que l'interaction entre le Médiateur et TFIIB se fasse par l'intermédiaire de Med14. Ces données nous ont ainsi permis de montrer *in vivo* un lien fonctionnel fort entre le module du milieu du Médiateur et TFIIB, en identifiant un nouveau contact entre Med14 et TFIIB.

Afin de connaître le rôle de Med10 et du module du milieu dans la formation du PIC, nous avons réalisé des expériences de ChIP. Nous avons analysé l'occupation de l'ensemble des GTFs, de la Pol II et du Médiateur sur la chromatine dans les contextes *MED10* et *med10-196* sur trois gènes modèles. Nous avons observé que dans le cas du mutant, l'occupation de tous les GTFs, à l'exception de TFIID et TFIIA, était diminuée, indiquant que leur occupation dans les PICs *in vivo* est indépendante du Médiateur et des autres GTFs. L'occupation de la Pol II était également fortement diminuée. Le Médiateur quant à lui présentait un effet différentiel selon le gène considéré. Nous avons souhaité analyser l'effet de cette mutation sur la formation du PIC sur l'ensemble du génome. Pour cela, nous avons réalisé des expériences de ChIP-seq. Cette analyse a montré que la mutation n'affectait pas l'occupation de TFIIA et TFIID. La Pol II et les autres GTFs sont eux affectés négativement sur l'ensemble du génome. On a pu observer que TFIIB et la Pol II étaient les protéines dont l'occupation était la plus diminuée. Le Médiateur quant à lui voit son occupation diminuée mais de manière moins importante que les GTFs. Ces résultats nous ont permis de montrer que la sous-unité Med10 et le module du milieu ont un rôle important dans la formation du PIC sur l'ensemble du génome. Ensuite, nous avons souhaité analyser comment la mutation *med10-196* affecte la formation du PIC *in vitro* par une approche biochimique. Ces expériences, en utilisant une approche développée par S. Hahn consistant en l'assemblage du PIC sur une matrice d'ADN immobilisée, ont été mise au point pour la première fois au laboratoire. Elles sont très intéressantes puisqu'elles permettent de mettre en évidence des effets des mutations du Médiateur sur l'assemblage du PIC selon une autre approche que celles utilisées *in vivo*. Ces expériences nous ont permis de confirmer nos observations faites *in vivo* puisque l'incorporation de TFIIB était là encore très affectée, contrairement à celle de TFIIA.

Nous avons souhaité analyser nos données de ChIP-seq plus en profondeur afin de savoir si la mutation *med10-196* pouvait avoir des effets gène-spécifique. En ce sens, nous avons intégré dans nos analyses bioinformatiques des données existante concernant l'architecture des promoteurs. Pour cela, nous avons, pour chaque protéine analysée par ChIP-seq, défini trois groupes de gènes selon l'effet de la mutation *med10-196* sur leur occupation. Nous avons ainsi défini le groupe des 25% des gènes les plus affectés, les 25% les moins affectés et les gènes dont le comportement suit de la tendance générale. Pour chacun de ces groupes et pour toutes les protéines, nous avons analysé l'occupation en nucléosomes et la présence des éléments TATA (Rhee and Pugh 2012) ainsi

que la dynamique des nucléosomes +1 et -1 (Dion et al. 2007). Nous avons pu ainsi montrer que les gènes les plus affectés par la mutation de Med10 pour l'occupation de la Pol II, TFIIB et TFIID étaient majoritairement des gènes possédant une boîte TATA canonique. Ces gènes montrent de façon générale un niveau d'occupation de nucléosomes plus faible ainsi qu'une dynamique plus importante des nucléosomes +1 et -1. Lorsque l'on s'intéresse au Médiateur, on constate que les gènes les plus affectés présentent un comportement totalement inversé, à savoir qu'ils sont dépourvus de boîte TATA et plus enrichis en nucléosomes. Cette analyse nous a ainsi permis de montrer que le Médiateur avait une influence différente sur l'assemblage du PIC selon l'architecture du promoteur des gènes.

Par le travail mené ici, nous avons montré un rôle général du module du milieu du Médiateur, et en particulier de la sous-unité Med10, dans la formation du PIC à la fois *in vitro* et *in vivo*. Nous avons pu mettre en évidence un contact et un lien fonctionnel entre le Médiateur et TFIIB. Ce contact semble important pour la formation du PIC. De plus, nous avons mis en évidence pour la première fois une relation entre le Médiateur et la formation du PIC en fonction de l'architecture du promoteur.

II. Article

Functional interplay between Mediator and TFIIB in preinitiation complex assembly in relation to promoter architecture

Thomas Eychenne,^{1,2} Elizaveta Novikova,^{1,2} Marie-Bénédicte Barrault,^{1,2} Olivier Alibert,³ Claire Boschiero,^{1,2} Nuno Peixeiro,^{1,2} David Cornu,⁴ Virginie Redeker,^{4,5} Laurent Kuras,¹ Pierre Nicolas,⁶ Michel Werner,^{1,2} and Julie Soutourina^{1,2}

¹Institute for Integrative Biology of the Cell (I2BC), Commissariat à l'Énergie Atomique (CEA), Centre National de la Recherche Scientifique (CNRS), Université Paris Sud, Université Paris Saclay, F-91198 Gif-sur-Yvette Cedex, France; ²Institut de Biologie et de Technologies de Saclay (IBITECS), CEA, F-91191 Gif-sur-Yvette Cedex, France; ³Laboratoire d'Exploration Fonctionnelle des Génomes (LEFG), Institut de Radiobiologie Cellulaire et Moléculaire (IRCM), CEA, Genopole G2, F-91057 Evry Cedex, France; ⁴Service d'Identification et de Caractérisation des Protéines par Spectrométrie de masse (SICaPS), CNRS, F-91198 Gif-sur-Yvette Cedex, France; ⁵Paris-Saclay Institute of Neuroscience (Neuro-PSI), CNRS, F-91198 Gif-sur-Yvette Cedex, France; ⁶Mathématiques et Informatique Appliquées du Génome à l'Environnement (MaIAGE), Institut National de la Recherche Agronomique (INRA), Université Paris-Saclay, 78350 Jouy-en-Josas, France

Mediator is a large coregulator complex conserved from yeast to humans and involved in many human diseases, including cancers. Together with general transcription factors, it stimulates preinitiation complex (PIC) formation and activates RNA polymerase II (Pol II) transcription. In this study, we analyzed how Mediator acts in PIC assembly using in vivo, in vitro, and in silico approaches. We revealed an essential function of the Mediator middle module exerted through its Med10 subunit, implicating a key interaction between Mediator and TFIIB. We showed that this Mediator–TFIIB link has a global role on PIC assembly genome-wide. Moreover, the amplitude of Mediator's effect on PIC formation is gene-dependent and is related to the promoter architecture in terms of TATA elements, nucleosome occupancy, and dynamics. This study thus provides mechanistic insights into the coordinated function of Mediator and TFIIB in PIC assembly in different chromatin contexts.

[*Keywords:* Mediator; RNA polymerase II transcription; TFIIB; preinitiation complex; *Saccharomyces cerevisiae*; promoter architecture]

Supplemental material is available for this article.

Received June 17, 2016; revised version accepted September 12, 2016.

In eukaryotes, RNA polymerase II (Pol II) transcription machinery is responsible for the synthesis of mRNAs and several classes of noncoding RNAs. This machinery is composed of a huge assembly of nuclear proteins. Regulation of Pol II transcription starts with the binding to gene regulatory elements of specific transcription factors (TFs), which then recruit coregulator complexes (coactivators and corepressors), and proceeds with the binding to promoter DNA of basal components, including Pol II and general TFs (GTFs). A coordinated action of all of these components in the context of promoter chromatin through a complex interaction network is required for regulated transcription of each gene.

Mediator of transcriptional regulation is one of the coregulator complexes serving as a bridge between DNA-binding TFs and the basal Pol II transcription machinery (Kornberg 2005). Together with the GTFs (TFIIA, TFIIB, TFIID, TFIIIE, TFIIF, and TFIIFH), this large multiprotein complex conserved in all eukaryotes promotes the assembly of preinitiation complexes (PICs) and activates transcription (Flanagan et al. 1991; Kim et al. 1994; Ranish et al. 1999). Mediator is essential for life in yeast and is generally required for Pol II recruitment and transcription (Thompson and Young 1995; Holstege et al. 1998; Soutourina et al. 2011). It also stimulates the phosphorylation

Corresponding authors: julie.soutourina@cea.fr, michel.werner@cea.fr
Article published online ahead of print. Article and publication date are online at <http://www.genesdev.org/cgi/doi/10.1101/gad.285775.116>.

© 2016 Eychenne et al. This article is distributed exclusively by Cold Spring Harbor Laboratory Press for the first six months after the full-issue publication date (see <http://genesdev.cshlp.org/site/misc/terms.xhtml>). After six months, it is available under a Creative Commons License (Attribution-NonCommercial 4.0 International), as described at <http://creativecommons.org/licenses/by-nc/4.0/>.

of the Pol II C-terminal domain (CTD) by TFIIF (Kim et al. 1994; Jiang et al. 1998). Mediator cooperates with GTFs to stabilize the PIC, in particular through interactions with TFIID, TFIIF, and Pol II (Johnson et al. 2002; Esnault et al. 2008; Soutourina et al. 2011; Eyboullet et al. 2015). A functional interplay between Mediator and TFIIB was also suggested based on in vitro experiments with human Mediator (Baek et al. 2006). In yeast, reconstituted sub-complexes were shown to interact with TFIIB in vitro (Kang et al. 2001). Artificial tethering of TFIIB to the promoter led to Mediator recruitment and PIC assembly in vivo (Lacombe et al. 2013). TFIIB, encoded by the *SUA7* gene in yeast, is one of the GTFs important for Pol II recruitment that determines the accuracy of start site selection (Pinto et al. 1992; Liu et al. 2010; Sainsbury et al. 2013).

The Mediator complex is composed of 25 subunits in yeast and up to 30 subunits in mammals (Malik and Roeder 2010). It has a modular organization with head, middle, and tail modules that constitute the core Mediator and the Cdk8 kinase module (Kornberg 2005). In yeast, 10 Mediator subunits are essential for cell viability. The contribution to the function of the complex of most Mediator subunits remains to be elucidated. In addition, how a single complex can be involved in ensuring tight transcriptional regulation of hundreds or thousands of genes through different specific interactions with TFs is largely unknown. The importance of the complex is highlighted by the fact that many Mediator subunits are involved in human diseases like neurodevelopmental pathologies or cancers (Kaufmann et al. 2010; Hashimoto et al. 2011; Spaeth et al. 2011; Schiano et al. 2014).

The function of Mediator middle module subunits has been very poorly studied. Based on protein cross-linking, a model for the Mediator middle module was proposed (Larivière et al. 2013). A cryo-electron microscopic model of a part of the Mediator complex within a partial PIC was also reported recently (Plaschka et al. 2015). Med10 is one of the essential and most conserved Mediator subunits belonging to the middle module. Only one conditional mutant in the yeast Med10 subunit (also known as Nut2), *med10-1ts* (N2D L61S L64P), was previously isolated (Han et al. 1999). Med10 is engaged in a number of contacts within the complex and interacts with the Med14 subunit that is central in Mediator architecture. The spatial organization of the Mediator modules was revised recently (Tsai et al. 2014; Wang et al. 2014). Med14 links the middle and tail module and, according to a recent model of entire yeast Mediator, makes extensive contacts with subunits from all three core modules (Robinson et al. 2015). A reconstitution of the active human core Mediator complex revealed a critical role of the Med14 subunit in Mediator architecture and function (Cevher et al. 2014).

Regulation of Pol II transcription occurs in the context of promoter chromatin. Most of the Pol II transcribed genes are preceded by a 5' nucleosome-free promoter region (NFR) on which the PICs assemble. The -1 nucleosome, located on the upstream side of the promoter NFR, can potentially control access to regulatory sequences. The +1 nucleosome is the first nucleosome en-

countered by the transcription machinery, and its function could differ depending on the nature of the TATA element (TATA box or TATA-like element) present on the promoter (Rhee and Pugh 2012). Promoters containing TATA-like elements are more dependent on the TFIID complex and may rely more on NFR-adjacent nucleosomes for PIC assembly. In addition, histone turnover was reported to be rapid on promoters (Dion et al. 2007), but the function of nucleosome dynamics in PIC assembly remains unknown.

In this study, we selected a conditional mutant in the Med10 Mediator middle subunit and analyzed the effect of *med10* mutation on PIC assembly on a genomic scale in vivo and in vitro. We show that this *med10* mutation does not have a major effect on Mediator stability but specifically modifies a contact of this subunit with Med14. We demonstrate that Mediator interaction with TFIIB was decreased in the *med10* mutant and that this interaction occurs via Med14 and, to a lesser extent, Med10. ChIP-seq (chromatin immunoprecipitation [ChIP] combined with high-throughput sequencing) analysis showed that genome-wide occupancy of PIC components was differentially affected in the *med10* mutant. The most pronounced decreases were observed for TFIIB and Pol II. TFIIA and TFIID occupancies were unmodified. Our genome-wide analysis also revealed that the impact of *med10* mutation on PIC formation and transcription is gene-dependent and is correlated with the promoter architecture.

Results

The conditional med10 mutant in the Mediator middle module

To enhance our understanding of the mechanisms in which Mediator is implicated and to clarify the role of the middle module, we selected a conditional mutant in the Med10 subunit, *med10-196*, by random mutagenesis using error-prone PCR followed by a screen for a temperature-sensitive growth phenotype as described previously for the Med11 and Med17 subunits (Supplemental Material; Esnault et al. 2008; Soutourina et al. 2011). The four residues mutated in *med10-196* (L53S, I79T, E82D, and N108I) were localized within the conserved domains according to multiple sequence alignments and secondary structure features (Bourbon 2008) and corresponded to identical (for L53S, E82D, and N108I) or functionally close (I79T) amino acids in the human Med10 protein (Fig. 1A). *med10-196* showed a strong temperature-sensitive phenotype and did not grow at 37°C (Fig. 1B). The individual mutations, when separated by site-directed mutagenesis, displayed normal growth phenotypes (data not shown). Hence, only the *med10-196* mutant was used for further analysis.

Med10 is engaged in multiple contacts within the Mediator complex that can be revealed using the yeast two-hybrid system (Guglielmi et al. 2004). We tested the ability of Med10-196 to interact with Med4, Med7, Med14, Med21, and Med31 (Fig. 1C). All tested interactions

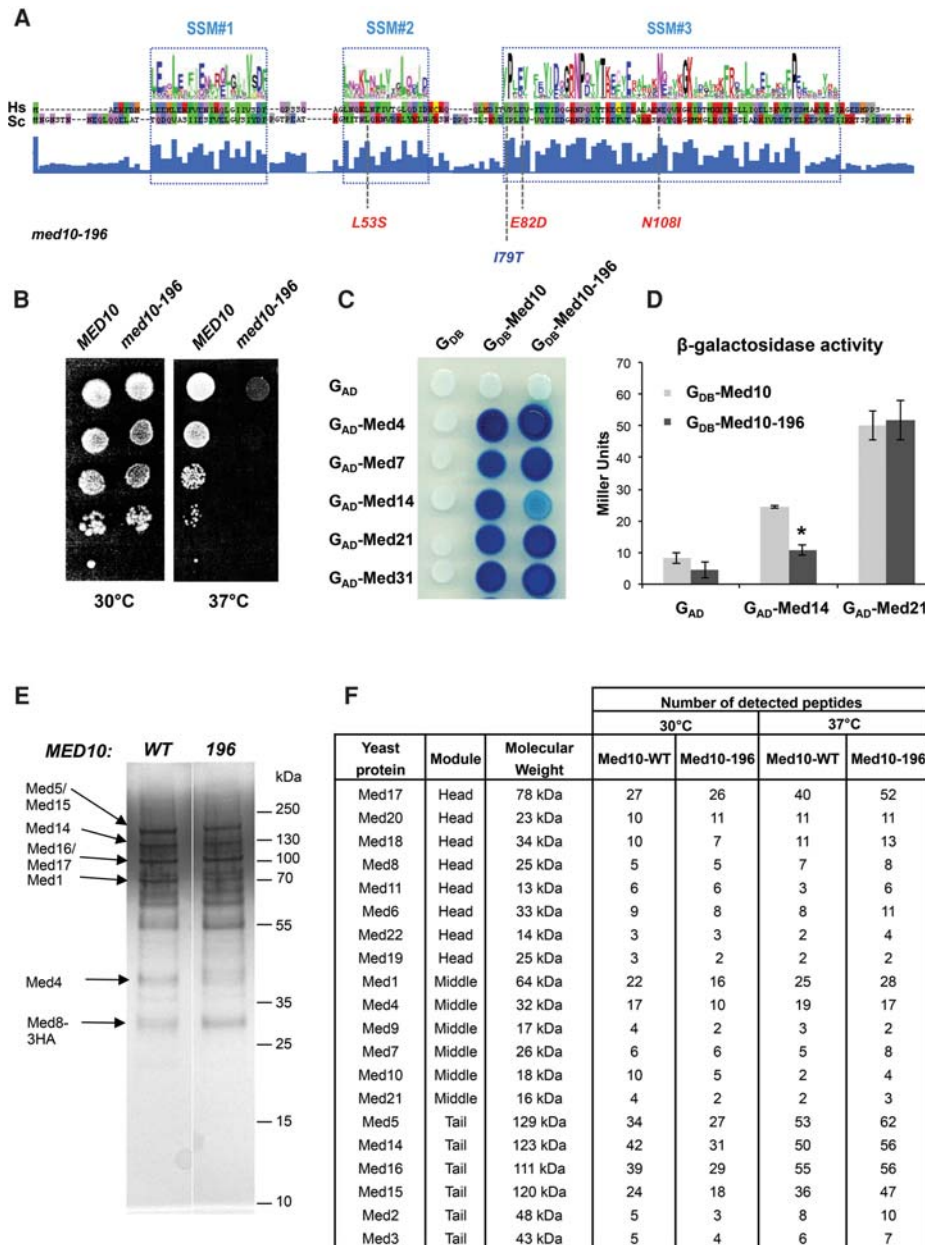


Figure 1. Phenotypes of the *med10-196* mutant. (A) Location of the mutations in *med10-196*. The Med10 conserved domains 1, 2, and 3 (signature-specific motif [SSM]) from *Saccharomyces cerevisiae* (Sc) and *Homo sapiens* (Hs) with the consensus generated by WebLogo were adapted from Bourbon (2008). The mutated residues corresponding to identical (for L53S, E82D, and N108I) or functionally close (I79T) amino acids in the human Med10 protein are indicated in red and blue, respectively. (B) Thermosensitive growth phenotype of the *med10* mutant. Cultures of wild-type and mutant *med10* yeast strains were serially diluted, spotted on YPD agar plates, and incubated for 3 d at permissive (30°C) or nonpermissive (37°C) temperatures. (C) Two-hybrid interactions of the Med10-196 protein with its partners. Wild-type or mutant Med10 was fused to the Gal4 DNA-binding domain (G_{DB} -Med10), and Med4, Med7, Med14, Med21, and Med31 were fused to the Gal4 activation domain (G_{AD} -Med4, G_{AD} -Med7, G_{AD} -Med14, G_{AD} -Med21, and G_{AD} -Med31). (D) Quantitative analysis of two-hybrid interactions between Med10 and Med14. Wild-type or mutant Med10 was fused to the Gal4 DNA-binding domain (G_{DB} -Med10), and Med14 or Med21 was fused to the Gal4 activation domain (G_{AD} -Med14 and G_{AD} -Med21). β -Galactosidase was assayed according to the Miller method, as described in the Supplemental Material. The mean values and standard deviation (indicated by error bars) of three independent experiments are shown. Asterisk represents a significant difference between the wild type and the mutant at P -value < 0.002 in a Student's t -test. (E) Silver stain SDS-PAGE analysis of purified Mediator complex from wild-type and *med10* mutant strains. Cells were grown at 30°C, and core Mediator complex was purified as described in the Supplemental Material. Mediator subunits with low molecular weight were not detectable by silver staining after SDS-PAGE. (F) Mass spectrometry analysis of Mediator integrity in the mutant grown at 30°C or transferred to 37°C. Core Mediator complex containing head, middle, and tail modules was purified from the *med10-196* mutant and a wild-type strain. For each Mediator subunit identified by mass spectrometry analysis in wild-type and the *med10* mutant, the number of identified peptides is indicated.

were similar between Med10-196 and the wild-type protein except for an interaction with the Med14 subunit that was significantly decreased in the mutant (Fig. 1C,D).

Since Mediator mutations could lead to dissociation of Mediator subunits, impairing Mediator function, we analyzed the effect of *med10-196* on the integrity of the core Mediator. Mediator was purified from *med10-196* and a wild-type strain (Fig. 1E) and was analyzed by mass spectrometry as described previously (Fig. 1F; Eyboulet et al. 2015). All 20 Mediator subunits detectable in a wild-type strain by this approach were present in the Mediator complex purified from the *med10* mutant grown at 30°C (Fig. 1F, 30°C) or after being transferred to 37°C for 90 min (Fig. 1F, 37°C).

med10 mutation affects Mediator-TFIIB interaction

To explore the molecular mechanisms of PIC assembly involving the Med10 Mediator subunit, we analyzed potential interactions between Mediator and PIC components by coimmunoprecipitation (co-IP) experiments. We started by investigating potential Mediator-TFIIB contact with whole-yeast extracts. Mediator was immunoprecipitated

via the HA-tagged Med5 Mediator subunit with an anti-HA antibody (Fig. 2A). Using TFIIB-specific antibodies, we showed that TFIIB coimmunoprecipitated with Mediator. We then examined whether this contact was modified in the *med10* mutant. The results demonstrated that Mediator interaction with TFIIB was reduced in *med10-196* transferred to 37°C compared with the wild type. In contrast, the Mediator interaction with TFIIA remained unchanged by *med10* mutation compared with the wild-type strain (Supplemental Fig. S1).

To investigate the functional interplay between Mediator and TFIIB, we combined 16 specific *sua7* (TFIIB) mutations (Wu et al. 1999) with *med10-196* and showed that three of them had synthetic phenotypes (Fig. 2B; Supplemental Fig. S2A). *sua7-34* (L52P) was co-lethal with the *med10-196* mutation, and *sua7-11* (L136P) and *sua7-36* (S53P) *med10-196* had a slow-growth phenotype at 30°C. On their own, these three synthetic TFIIB mutations do not lead to a start site shift and confer cold- and heat-sensitive phenotypes (Wu et al. 1999). The mutated residues in *sua7-34* and *sua7-36* localize to the N-terminal zinc ribbon domain (B ribbon), and L136P of *sua7-11* is situated in the first part of the core domain. The allele-specific

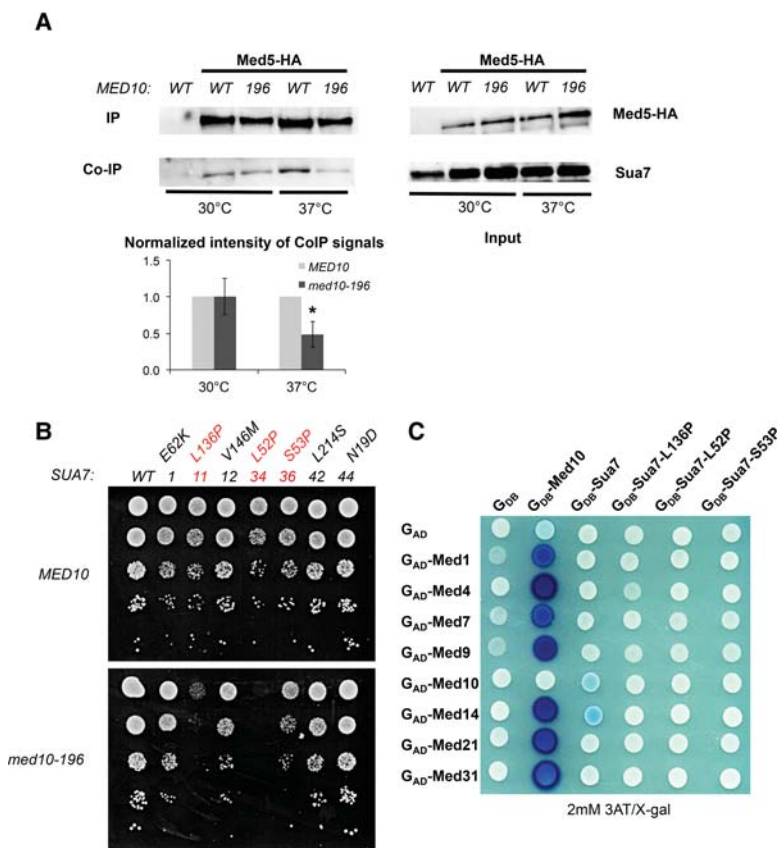


Figure 2. Functional interaction between Mediator and TFIIB. (A) Co-IP between Mediator and TFIIB in the *med10-196* mutant compared with the wild-type strain. Wild-type and *med10* mutant strains carrying a Med5-HA tag were grown to exponential phase at 30°C or transferred for 90 min to 37°C. Mediator was immunoprecipitated (IP) through Med5-HA from crude extracts (Input) of wild-type and mutant strains using magnetic beads coupled to anti-HA antibodies. *MED10* strain carrying a nontagged Mediator subunit was used as a negative control. Coimmunoprecipitated TFIIB was detected by Western blotting using anti-TFIIB antibodies. The intensity of immune staining for coimmunoprecipitated TFIIB signals relative to the wild type was normalized against immunoprecipitation signals and is displayed in the *bottom* panel. The mean values and standard deviation (indicated by error bars) of three independent experiments are shown. The asterisk represents a significant difference between the wild type and the mutant at P -value < 0.05 in a Student's t -test. (B) Specific *sua7* mutants have synthetic phenotypes in combination with *med10-196*. The strain deleted for *med10* and complemented by a *TRP1* plasmid carrying *MED10* or *med10-196* and also deleted for *sua7* and complemented by a *URA3* plasmid carrying *SUA7* was transformed by the *HIS3* plasmids carrying either a wild-type or mutated version of *SUA7*. Transformants were serially diluted, spotted on 5-FOA-containing agar plates to counterselect the wild-type *SUA7*-bearing plasmid (see the Supplemental Material), and incubated for 3 d at 30°C. The *sua7* mutants [*sua7-11* (L136P), *sua7-34* (L52P), and *sua7-36* (S53P)] showing

synthetic phenotypes with *med10-196* are indicated in red. (C) A two-hybrid interaction between Mediator and Sua7 is decreased with Sua7 mutants. Wild-type or mutant Sua7 was fused to the Gal4 DNA-binding domain (G_{DB}-Sua7), and Med1, Med4, Med7, Med9, Med10, Med14, Med21, and Med31 were fused to the Gal4 activation domain (G_{AD}-Med1, G_{AD}-Med4, G_{AD}-Med7, G_{AD}-Med9, G_{AD}-Med10, G_{AD}-Med14, G_{AD}-Med21, and G_{AD}-Med31). Med10 two-hybrid interactions with Mediator subunits are shown as positive controls.

synthetic phenotypes between the *med10-196* and *sua7* mutations are consistent with the physiological importance of Mediator–TFIIB interaction.

To characterize Mediator middle–TFIIB interaction, Sua7 fused to the Gal4 DNA-binding domain was used in a two-hybrid assay with Mediator middle subunits (Med1, Med4, Med7, Med9, Med10, Med14, Med21, and Med31) fused to the Gal4 activation domain (Fig. 2C; Supplemental Fig. S2B,C). Interactions between Med10 and other Mediator subunits were used as positive controls. We showed that the Med14 Mediator subunit interacts with Sua7 in a two-hybrid system. The Med10 subunit also showed a two-hybrid interaction with TFIIB but with a lower intensity. In addition, these interactions were lost with Sua7-11, Sua7-34, and Sua7-36 mutants, indicating their specificity. Interestingly, the only contact within the Mediator middle module decreased in *med10-196* was a two-hybrid interaction with the Med14 subunit (Fig. 1C,D), suggesting that the *med10* mutation might alter Med10–Med14 contact, leading to reduced Mediator–TFIIB interaction.

We then investigated the functional interplay between the Mediator Med14 subunit and TFIIB by combining specific *sua7* (TFIIB) mutations (Wu et al. 1999) with C-terminal truncations in *MED14* that have temperature-sensitive phenotypes (*med14-752*, *med14-686*, and *med14-483* with a stop codon at position 752, 686, or 483, respectively) (Supplemental Fig. S3A). It should be noted that further C-terminal truncation of *MED14* to position 471 led to a lethal phenotype (data not shown). We observed allele-specific synthetic phenotypes between the *med14* and *sua7* mutations that are consistent with the physiological importance of Mediator–TFIIB interaction (Supplemental Fig. S3B).

Taken together, these results show that Mediator interacts with TFIIB via the Med14 and Med10 Mediator subunits and that this contact is important for a functional interplay between Mediator and the GTF.

med10 mutation differentially affects chromatin occupancy of PIC components

To investigate the molecular consequences of *med10* mutation on PIC assembly in vivo, we determined the chromatin occupancy of Mediator, Pol II, and all GTFs on the constitutively expressed genes *ADH1*, *PYK1*, and *PMA1*. The association of Mediator was measured by ChIP experiments with three different Mediator subunits: Med5-HA (tail), Med15-HA (tail), and Med17-HA (head) (Fig. 3A–C). The Mediator occupancy of the promoter regions remained unchanged or decreased slightly in *med10-196* except for the *PMA1* promoter region showing a decrease in Med5 Mediator subunit occupancy and, to a greater extent, Med17 Mediator subunit occupancy. Pol II (Rpb1) association with the promoter and transcribed regions of class II genes was reduced compared with the wild type (Fig. 3D).

To assess the ability of *med10* mutation to influence the occupancy of GTFs, ChIP was performed against TFIIA (Toa2 subunit), TFIIB (Sua7), TFIID (TBP), TFIIE

(Tfa2 subunit), TFIIF (Tfg1 subunit), and the two TFIIH modules TFIIH core (Rad3 subunit) and TFIHK kinase module (Kin28 subunit) (Fig. 3E–K). Interestingly, the most pronounced effect of *med10-196* was observed for TFIIB (Fig. 3E), which is consistent with the decreased Mediator–TFIIB interaction in this mutant. Promoter region occupancy of TFIIA and TBP remained largely unchanged (Fig. 3F,G). In contrast, *med10* mutation led to a decrease in the occupancy of all other GTFs, including the TFIIF, TFIIE, and TFIIH modules, to an extent depending on the GTF (Fig. 3H–K). These ChIP results indicated that a mutant in the Mediator middle module has a pronounced effect on the recruitment and/or stability of specific PIC components in vivo.

med10-196 has a global effect on PIC assembly in vivo

To extend the PIC assembly analysis to the whole yeast genome, ChIP-seq experiments were performed for Mediator (Med15 and Med17 subunits), Pol II (Rpb1), and all GTFs (TFIIA [Toa2], TFIIB [Sua7], TFIID [TBP and Taf1], TFIIE [Tfa2], TFIIF [Tfg1], and two TFIIH modules: TFIIH core [Rad3] and TFIHK kinase module [Kin28]). Input DNA and DNA from ChIP with an untagged strain were used as negative controls. Supplemental Figure S4A shows an example of ChIP-seq density distributions for all proteins on a selected class II gene in wild-type strains after subtraction of the normalized control of an untagged strain as described previously (Eyboulet et al. 2013, 2015). As expected, a metagene analysis around the transcription start sites (TSSs) in wild-type strains showed that Pol II is distributed inside transcribed regions and that Mediator (Med15 and Med17 subunits) is located upstream of TSSs on regulatory regions (Supplemental Fig. S4B, top graph). The ChIP-seq distributions of GTFs overlapped with each other close to the TSS with a maximum of peak densities depending on the protein (Supplemental Fig. S4B, bottom graph). A similar analysis performed with the *med10* mutant did not indicate any change in Mediator, GTFs, or Pol II peak positions (data not shown).

To investigate how general the effects of *med10-196* on PIC assembly in vivo were, the genome-wide occupancy of Mediator, Pol II, and all GTFs was analyzed using ChIP-seq experiments in the mutant to compare them with the wild-type strains. To detect the potential global effects of Mediator mutations on PIC component occupancy, we performed the normalization step relative to quantitative PCR (qPCR) data on a set of selected regions as described previously (Eyboulet et al. 2015). Regression analysis of PIC component binding in *med10-196* versus wild type was systematically performed.

Genome-wide Mediator occupancy was analyzed for the Med15 tail subunit and the Med17 head subunit. As described above, ChIP results on selected class II gene promoters showed no effect or a slight decrease in Mediator occupancy in the *med10-196* mutant, except for the *PMA1* promoter region with a decrease in Med5 Mediator occupancy and, to a greater extent, Med17 Mediator occupancy. Genome-wide analysis of the *med10-196* mutant revealed a global 1.4-fold decrease (the slope of regression

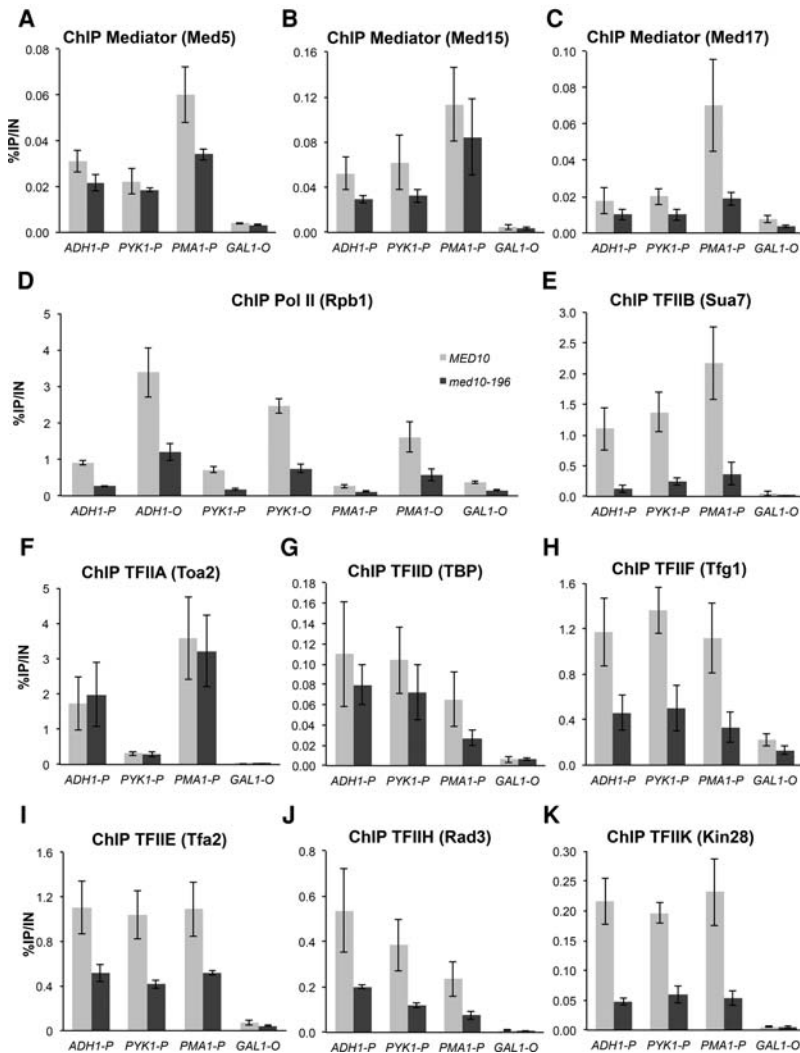


Figure 3. ChIP analysis of PIC components in *med10-196*. Cells were grown to exponential phase at 30°C on YPD medium and then transferred for 90 min to 37°C. The immunoprecipitated protein is indicated together with the complex to which it belongs: Mediator tail (Med5) (A), Mediator tail (Med15) (B), Mediator head (Med17) (C), Pol II (Rpb1) (D), TFIIB (Sua7) (E), TFIIA (Toa2) (F), TFIID (TBP) (G), TFIIF (Tfg1) (H), TFIIE (Tfa2) (I), TFIIH core module (Rad3) (J), and TFIIH kinase module (Kin28) (K). All proteins, except Rpb1, were tagged with HA or TAP (see the Supplemental Material). Immunoprecipitated DNA was amplified with primers corresponding to *ADH1*, *PMA1* and *PYK1* ORF (O) or promoters (P). P1 primers listed in Supplemental Table S3 located close to upstream regulatory regions were used for Mediator ChIP experiments, and P2 primers located close to core promoters were used for Pol II and GTF ChIP experiments. *GAL1* ORF was used as a negative control, since it is repressed in glucose-supplemented rich medium. The mean values and standard deviation (indicated by error bars) of three independent experiments are shown.

line, shown in red in Fig. 4A, is equal to 0.70) in Med15 Mediator occupancy with a high correlation coefficient (R^2 is equal to 0.92) (Fig. 4A). The effect of *med10* mutation on Med17 Mediator occupancy was more pronounced, with a global 1.8-fold decrease suggesting some differences in the chromatin association of Mediator head and tail modules in *med10-196* (Fig. 4B).

Similar to the ChIP results on selected gene promoters, Pol II and TFIIB occupancies were the most affected by the *med10* mutation. A large global decrease was observed for genome-wide binding of Pol II (3.8-fold) and TFIIB (4.3-fold) in *med10-196* (Fig. 4C,D). In contrast, genome-wide occupancy of two other PIC components, TFIIA and TFIID (TBP and Taf1 subunits), remained largely unchanged compared with the wild type (slopes are equal to 0.90, 1.04, and 0.95, and R^2 is equal to 0.8, 0.86, and 0.94, respectively) (Fig. 4E–G). A global decrease in *med10-196* was also observed for TFIIF, TFIIE, TFIIH core, and TFIHK occupancy to varying extents depending on the GTF (1.8-fold, 1.4-fold, 3.1-fold, and twofold decrease, respectively) (Fig. 4H–K). The genome-wide analysis of the *med10* mutant showed a key role of the

Mediator middle module in PIC assembly in vivo and demonstrated that *med10* mutation can lead to a moderate effect on Mediator occupancy accompanied by a large decrease in Pol II and TFIIB occupancy and a decrease in several PIC components. TFIIA and TFIID behaved in a radically different manner without any change in *med10-196* compared with the wild type. Our results also indicate an independent behavior for PIC components, including TFIIE, core and kinase TFIIH modules, or TFIIF and Pol II.

Gene-specific effects of med10-196 mutation on PIC assembly in vivo

Our previous analysis of the impact of conditional mutations in the Med17 Mediator subunit on genome-wide PIC assembly showed a global nature of the effects but also suggested that some effects could still be gene-specific (Eyboulet et al. 2015). The high correlation coefficients obtained when comparing the genome-wide occupancy of the different PIC components between *med10-196* and the wild type highlight the global effects of the mutation

(Fig. 4). To address the question of the gene-specific effects, we analyzed the variations of the ratio across genes between the normalized ChIP-seq tag densities computed in the mutant and the wild type. For each measured GTF and Mediator subunit, tag densities in the mutant versus the wild type were plotted in \log_2 scale, such as a constant ratio translates into an alignment of the points along a line parallel to the diagonal ($Y = X$) whose intercept corresponds to the \log_2 ratio (Fig. 5A–C).

We first investigated the amplitude of the *med10* mutation effects on the different PIC component occupancy. For each PIC component, we classified the genes into three groups based on the amplitude of the *med10* mutation effects: the upper quartile group (ratios in the highest 25%) (Fig. 5, in red), the interquartile group (general trend) (Fig. 5, in gray), and the lower quartile group (ratios in the lowest 25%) (Fig. 5, in blue) as described in the Materials and Methods (Fig. 5; Supplemental Fig. S5). Since the general effect on occupancy was negative for most PIC elements, the lower quartile group corresponded to the most impacted genes. However, for TFIIA and TFIIID (TBP and Taf1), the lower and upper quartiles distinguished genes that are impacted negatively and positively.

Next, to investigate the functional interplay between promoter chromatin architecture and PIC assembly, we compared the nucleosome occupancy profiles from Rhee and Pugh (2012) between the three groups of genes built to reflect the impact of *med10-196* on each PIC component (Fig. 5D–F; Supplemental Fig. S5). These groups did not correlate with changes in nucleosome positioning. However, for GTFs such as TFIIIB, TFIIH, TFIIK, and Pol II, the groups correlated with the occupancy for nucleosomes on TSS-surrounding regions; namely, the most affected genes (lower quartile groups) showed low nucleosome occupancy, whereas the least affected genes (upper quartile groups) displayed high nucleosome occupancy (Fig. 5D,E; Supplemental Fig. S5). An opposite trend was observed for changes in Mediator occupancy between the mutant and the wild type: The least affected genes (upper quartile groups) showed lower nucleosome occupancy than the most affected genes (lower quartile groups) (Fig. 5F).

Next, we determined how TATA-box-containing genes (Rhee and Pugh 2012) were distributed between the groups (Fig. 5G–I; Supplemental Fig. S5). The genes classified as the most affected by *med10-196* for TFIIIB and Pol II occupancy contained a TATA box significantly more often than those classified as the least affected (Fig. 5G,H). Again, the situation was opposite for the impact of the mutation on Mediator occupancy: The genes classified as the least affected contained a TATA box more often than the genes classified as the most affected (Fig. 5I). The dynamics of nucleosome -1 and $+1$ as defined by Dion et al. (2007) follow the TATA-box enrichment pattern (Fig. 5G–I; Supplemental Fig. S5). Both the most affected genes for TFIIIB and Pol II occupancy and the least affected genes for Mediator occupancy exhibited promoter regions with more dynamic NFR-adjacent nucleosomes (hot nucleosome -1 and $+1$).

Patterns of nucleosome occupancy, TATA-box presence, and nucleosome dynamics suggest that some level

of correlation exists between the impacts of *med10-196* mutation on the different PIC components measured at the gene level. To describe these correlations more directly, the three groups of genes reflecting the impact of *med10-196* on each PIC component were visualized in heat maps with genes ordered by TFIIIB groups (Fig. 5J), Pol II groups (Fig. 5K), and Med17 groups (Fig. 5L). For each PIC component, the group of more affected genes is colored in blue, the general trend group genes are in gray, and the group of less affected genes is in red. These heat maps allow the observation of a positive correlation between the impact of the mutation on TFIIIB, TFIIH, TFIIK, and Pol II occupancy and a negative correlation between the impact on these GTFs and on Mediator subunits. To quantify the observed relationships, we performed a Spearman correlation analysis (Fig. 5M). The impacts of *med10-196* were the most intercorrelated for TFIIIB, TFIIH, TFIIK, and Pol II. The impact on TFIIIF also showed some correlation with this set of four PIC components. Although TBP occupancy was globally unchanged in the *med10* mutant, slight gene-specific variations also tended to follow the variations observed for the four most correlated PIC elements but simultaneously showed some positive correlation with the impact on TFIIIE. In contrast, the measured impact of *med10* mutation on Mediator occupancy exhibited slight negative correlations with most other PIC components.

In conclusion, the impact of Mediator *med10* mutation on PIC formation and transcription has amplitude, which is gene-dependent and correlated with the promoter architecture in terms of TATA elements, nucleosome occupancy, and dynamics.

med10 mutation affects PIC assembly in vitro

To investigate the role of the essential Med10 subunit of Mediator on PIC assembly in vitro, we performed biochemical experiments of PIC formation on immobilized DNA templates (Ranish et al. 1999). Nuclear extracts were prepared from the *med10-196* mutant and the wild-type strain as described previously (Ranish et al. 1999). The amounts of PIC components, including Mediator (Med14 and Med17 subunits), Pol II (Rpb3 subunit), TFIIIB (Sua7), and TFIIA (Toa1 subunit), were analyzed by Western blotting with *med10-196* and wild-type nuclear extracts (Fig. 6, lanes 1,2, Input). Histone H3 was used as a loading control. All PIC components tested were unchanged in the *med10* mutant compared with the wild-type extract. We then investigated the impact of *med10-196* mutation on the PIC formation in vitro using an immobilized template system. The PICs were assembled from Med10 wild-type or Med10-196 nuclear extracts in the absence or presence of Gal4-Gcn4 activator and analyzed by Western blotting (Fig. 6, lanes 3–10, PIC). As expected, the activator stimulated the recruitment of the PIC components tested, including Mediator, Pol II, TFIIIB, and TFIIA, using nuclear extract prepared from the wild-type strain (Fig. 6, lanes 3,4). Under the conditions of the assay, TFIIH bound to the immobilized template to a similar extent in the presence and absence of the activator

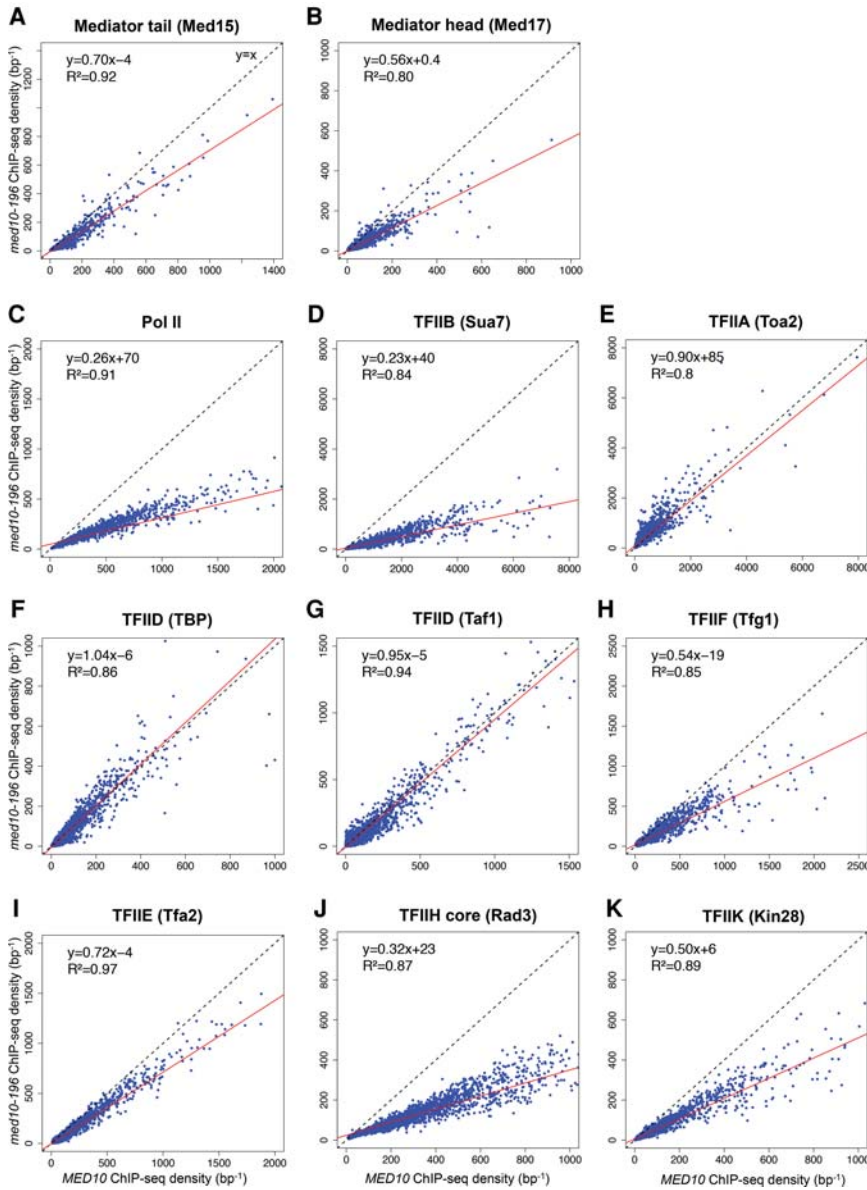


Figure 4. Genome-wide effects of the *med10-196* mutation on PIC component distribution. Cells were grown at 30°C in YPD medium and then transferred for 90 min to 37°C. The ChIP-seq densities of sequence tags in Mediator tail (Med15) (A), Mediator head (Med17) (B), TFIIB (Sua7) (D), TFIIA (Toa2) (E), TFIID (TBP) (F), TFIID (Taf1) (G), TFIIF (Tfg1) (H), TFIIE (Tfa2) (I), TFIIH core (Rad3) (J), and TFIIK (Kin28) (K) were calculated for promoter regions of Pol II transcribed genes. (C) The densities of sequence tags in the Pol II ChIP-seq experiments were calculated for the Pol II transcribed genes. Tag densities were normalized relative to qPCR data on a set of selected genes. Each point on the plot corresponds to one promoter region or one ORF. Promoter regions correspond to intergenic regions in tandem or in divergent orientation, excluding intergenic regions encompassing Pol III transcribed genes. In all, 2694 intergenic regions corresponding to 3303 Pol II-enriched genes were used for these analyses. A linear regression (red line) for ChIP-seq density in the mutant versus ChIP-seq density in wild type and an R^2 correlation coefficient are indicated.

(data not shown), which likely reflects the nonspecific DNA-binding activities of this factor, consistent with previous observations (Baek et al. 2006). This situation has precluded the analysis of the *med10* mutation's effect on TFIIH recruitment in vitro. In agreement with our ChIP-seq results, the recruitment of Mediator, Pol II, and TFIIB was defective in the *med10-196* nuclear extract (Fig. 6, lane 8) compared with the wild-type extract (Fig. 6, lane 4). Similar to the situation that we observed for TFIIA genome-wide binding, the level of this factor in PIC assembled in vitro was unchanged in the *med10* mutant extract. Addition of increasing amounts of purified wild-type core Mediator stimulated the recruitment of each of the affected PIC components (Fig. 6, lanes 9,10). The results show that a functional Med10 Mediator subunit is required for the stable recruitment of Mediator, Pol II, and TFIIB, suggesting a role for this Mediator subunit in efficient PIC assembly and/or stability.

Discussion

In this study, an integrative analysis of a conditional mutant in the Med10 Mediator middle module subunit was performed combining in vivo, in vitro, and in silico approaches, giving mechanistic insights into our understanding of Mediator function. We show that the Mediator middle module has a general role in PIC formation and that a functional interplay between Mediator and TFIIB is important for PIC component assembly and Pol II transcription. (1) We demonstrated an essential role of Mediator in TFIIB recruitment and/or stabilization during PIC formation in vivo and in vitro. We identified Mediator interactions with this GTF via Med14 and Med10 subunits using a two-hybrid approach. *med10* mutation specifically modified a contact of this subunit with Med14 and led to a decreased Mediator-TFIIB interaction, demonstrating a functional link between Mediator and TFIIB in

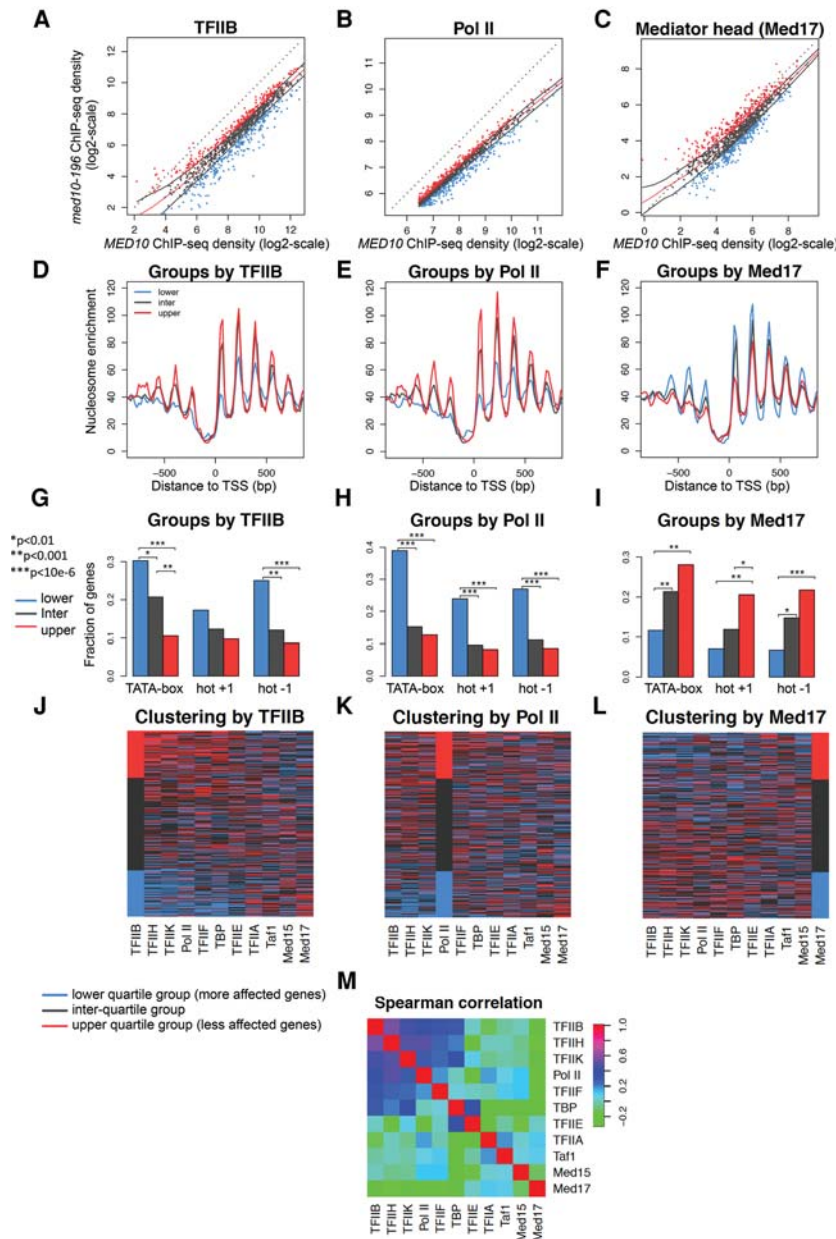


Figure 5. Clustering analysis of genome-wide PIC occupancy ratios between *med10-196* and the wild type. (A–C) ChIP-seq tag densities in the mutant versus the wild type were plotted as in Figure 4 but with log₂ scales. To aggregate data at the gene level, divergent genes with double peaks for GTFs were excluded from the analysis as well as intergenic regions encompassing Pol III transcribed genes and centromeric regions. The genes with the lowest Pol II occupancy (lowest 25%) were also excluded from the analysis. For each PIC component, three groups of genes were defined according to the occupancy ratios between the *med10-196* mutant and the wild type: the lowest 25% (lower quartile group in blue), the highest 25% (upper quartile group in red), and the ratios between 25% and 75% (interquartile group in gray). The red line corresponds to the median trend, and black lines correspond to the first and third quartiles of the data. Blue points correspond to the lowest 25% of mutant/wild-type values, and gray points correspond to values between 25% and 75% of the mutant/wild type. The red points correspond to the 25% of the genes that have the highest mutant/wild-type ratio. Tag densities were calculated as described in the legend for Figure 4 and were analyzed for the TFIIIB (A,D,G,J), Pol II (B,E,H,K), and Mediator (Med17) groups (C,F,I,L). (D,E,F) The groups determined by TFIIIB (A), Pol II (B), and the Mediator Med17 subunit (C) were analyzed for nucleosome occupancy in a 1600-base-pair (bp) window centered on the TSS. *P*-values determined by Wilcoxon test for the differences between the gene groups for the maximum values of nucleosome occupancy on the region between 0 and 100 bp relative to the TSS were as follows: For TFIIIB groups, lower quartile versus interquartile was 3×10^{-6} , interquartile versus upper quartile was $<10 \times 10^{-12}$, and lower quartile versus upper quartile was $<10 \times 10^{-12}$; for Pol II groups, lower quartile versus interquartile was 6×10^{-4} , interquartile versus upper quartile was 6×10^{-3} , and lower quartile versus upper quartile was 7×10^{-6} ; and for Med17 groups, interquartile versus upper quartile was 0.002, and lower quartile versus upper quartile

was 6×10^{-5} . (G–I) The groups determined by TFIIIB (A), Pol II (B), and the Mediator Med17 subunit (C) were analyzed for the presence of the TATA box for dynamic (hot) nucleosomes –1 and +1. *P*-values determined by Fisher test are indicated by asterisks with a value key in the left panel. (J–L) The heat maps clustered by the TFIIIB groups (J), Pol II groups (K), and Med17 groups (L) summarize the group distribution according to the occupancy ratios between the mutant and the wild type for each GTF, Mediator subunit, and Pol II. For each PIC component, lower quartile group genes are colored in blue, interquartile group genes are gray, and upper quartile group genes are red. The genes are ordered in the heat maps inside each group by interquartile range (IQR) score. (M) Pair-wise Spearman correlations between mutant and wild-type ratios for each PIC component were calculated. The correlated PIC components (>0.48) were TFIIIB, TFIIH, TFIIK, and Pol II.

vivo. (2) Our results reveal a general function of the Med10 Mediator subunit and the Mediator middle module in PIC assembly and/or stability in vivo and in vitro. *med10* mutation has a differential effect on PIC component occupancy, with a very large global decrease in Pol II and TFIIIB. (3) Our genome-wide analysis revealed that the range of the effects of *med10* mutation on PIC formation and transcription is gene-specific and is correlated with the pro-

motor architecture, including the presence of TATA elements, the level of nucleosome occupancy, and nucleosome dynamics.

Functional link of Mediator with TFIIIB

Previously, a functional cooperation between Mediator and TFIIIB was suggested in yeast and humans (Kang

et al. 2001; Baek et al. 2006; Lacombe et al. 2013). Our present work supports the existence of a strong functional link between Mediator and TFIIB through contacts between the middle module and the GTF. We identified Med14 and Med10 as the subunits that interact with TFIIB, as indicated by two-hybrid interactions and their loss in specific TFIIB mutants. Moreover, Mediator–TFIIB co-IP decreased in *med10-196*, and allele-specific synthetic phenotypes were identified between *sua7* and *med10-196* or *med14* mutations. These observations are in line with the physiological importance of Mediator–TFIIB interaction. Interestingly, the *sua7-36* (S53P) mutant that had a synthetic phenotype with *med10-196* was shown previously to be defective for activation of specific genes (Wu and Hampsey 1999). The only contact within the Mediator middle decreased in *med10-196* was a two-hybrid interaction with the Med14 subunit, but all Mediator core subunits, including Med14, were identified by mass spectrometry analysis of purified Med10-196 Mediator. We noted that TFIIB was the PIC component whose recruitment was the most affected by the *med10* mutation in biochemical experiments with immobilized templates and in our genome-wide location analysis of PIC formation. Based on our results, we propose that the Med10 mutation specifically alters the Med10–Med14 contact, introducing conformational or functional changes in the Mediator middle module that lead to reduced Mediator interaction with TFIIB, destabilization of the GTFs within the PIC, and reduced Pol II recruitment and transcription.

Although our data show that Med14 and Med10 interact with TFIIB, we do not exclude the possibility that oth-

er Mediator subunits might be contacting this GTF. A recent cryo-electron microscopic model of a partial PIC, including a part of Mediator, suggests that Med18–Med20 subunits of the Mediator head bind the TFIIB B-ribbon domain (Plaschka et al. 2015). However, this model does not precisely define the positions of the Med10 and Med14 subunits, and the PIC complex is partial and stabilized by a cross-link with BS3, precluding any conclusion about the positioning of Med14–TFIIB and Med10–TFIIB contacts. In addition, an earlier biochemical study with reconstituted Mediator modules showed a pull-down between TFIIB and the Med9/10 subcomplex (Kang et al. 2001). It should also be noted that in vivo PIC assembly is a dynamic process involving different contacts acting at specific steps and that Mediator is particularly dynamic and flexible in this process.

Both Mediator and TFIIB are interacting with Pol II and are important for the enzyme recruitment. The PIC assembly culminates in Pol II recruitment, and therefore it is not surprising that many PIC components contact Pol II, the main player of transcription. We could not exclude that, in addition to Mediator–TFIIB contact, other interactions within the PIC could be involved in *med10* mutation effects.

Global role of the Mediator middle module in PIC formation

This study demonstrates a general function of the Med10 Mediator subunit and Mediator middle module in PIC assembly and/or stability in vivo and in vitro. We showed that the behaviors of TFIIB, TFIIF, TFIK, and Pol II were correlated on individual promoters when Med10 was mutated. TFIID and TFIIA were essentially unaffected, suggesting that their recruitment can be independent from that of the other GTFs or that these components are stabilized in the *med10* mutant.

Based on in vitro reconstitution studies, a model for PIC assembly in a linear sequence was initially proposed with TFIID recruitment as a first component, followed by TFIIA and TFIIB arrival, Pol II recruitment in association with TFIIF, and completion of PIC assembly by TFIIF and TFIIE (Buratowski et al. 1989; Ranish and Hahn 1996). This model, which does not include Mediator, cannot explain our in vivo observations that suggest an independent behavior for PIC components that are supposed to be incorporated together such as TFIIE, core and kinase TFIIF modules, or TFIIF and Pol II. The Med10 mutation led to a large decrease in Pol II occupancy, but the effect on TFIIF, which is thought to be incorporated to the PIC together with Pol II, was more moderate. According to a linear model, TFIIE and TFIIF finalize the PIC assembly and depend on the incorporation of Pol II. However, the ranges of the *med10* mutation effects differ considerably. Recent electron microscopy-based structural models of human and yeast PIC have provided important information on transcription initiation mechanisms (He et al. 2013, 2016; Murakami et al. 2013, 2015; Plaschka et al. 2015, 2016). However, they are currently incomplete (Plaschka et al. 2015), and the dynamics of PIC assembly are

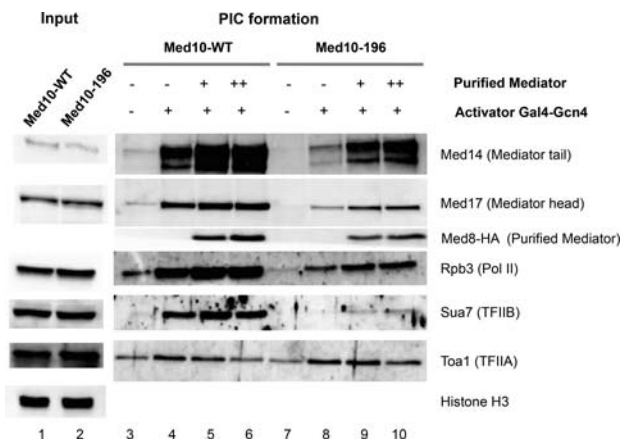


Figure 6. In vitro PIC formation in the *med10-196* mutant. (Lanes 1,2) Nuclear extracts were prepared from the wild-type and *med10-196* mutant strains and analyzed by Western blotting (Input). Histone H3 was used as a loading control. PIC assembly assay was performed by incubating the indicated nuclear extracts with the *HIS4* immobilized template for 40 min in the absence (lanes 3,7) or presence (lanes 4–6,8–10) of Gal4-Gcn4 activator as described in the Supplemental Material. Increasing amounts of purified core Mediator were added as indicated (lanes 5,6 for Med10 wild-type extract; lanes 9,10 for Med10-196 extract). Western blotting for Med8-HA was used as a control for the purified Mediator addition.

unavailable. Our functional *in vivo* study thus complements structural studies to investigate different aspects of PIC assembly on promoters.

Previously, our analysis of mutants in the Med11 and Med17 Mediator head subunits showed that Mediator stabilizes TFIIF kinase and TFIIF core modules independently and selectively contributes to TBP recruitment or stabilization on the chromatin and that a direct Mediator–Pol II interaction is generally required for Pol II recruitment and transcription (Esnault et al. 2008; Soutourina et al. 2011; Eyboulet et al. 2015). Based on our results, we proposed a model for PIC assembly through multiple pathways (Esnault et al. 2008) and suggested that Mediator independently orchestrates multiple steps of PIC assembly *in vivo* (Eyboulet et al. 2015). In the present work, we analyzed genome-wide occupancy of all PIC components and showed that mutations in the Med10 Mediator middle module that affect Mediator–TFIIIB contact led to a pronounced decrease in TFIIIB (coordinated with some but not all PIC components) and a global impact on Pol II recruitment and transcription, demonstrating a role of Mediator in PIC assembly on a genomic scale at the step of TFIIIB binding.

Mediator impact on PIC assembly related to promoter architecture

Mediator serves as a bridge between specific TFs and the general Pol II transcriptional machinery, acting in gene-specific transcriptional regulation. Despite the global nature of the Mediator effects on PIC assembly, the range of these effects could be gene-specific, as suggested by previous Med17 mutant analysis (Eyboulet et al. 2015). Here, we directly addressed the question of a global versus a gene-specific nature of Mediator impact on PIC formation by performing in-depth analysis of our ChIP-seq data of all PIC components in the Med10 mutant compared with the wild type. We determined how the genes most or least affected for one PIC component are distributed for Mediator, GTFs, and Pol II. The range of the effects of Med10 mutation is not simply due to the experimental variability. The genes the most and the least affected by Mediator mutation were, in general, specific for each PIC component but show a common tendency (correlation) for several factors like TFIIIB, TFIIF, TFIIF, and Pol II.

Since the PIC assembly occurs in a chromatin context, we examined the nucleosomal organization of the promoter regions in different groups for Med10 mutation impact. This analysis revealed a connection between the nucleosome occupancy around the TSS and the impact of Med10 Mediator mutation on TFIIIB, TFIIF, TFIIF, and Pol II occupancy, with the most affected genes showing significantly lower nucleosome occupancy on TSS-surrounding regions. It should be noted that it was not the case for TFIIA, TFIID (Taf1), TFIIE, and TFIIF (Supplemental Fig. S5). Interestingly, the opposite tendency was observed for the Mediator subunits, since the most impacted genes tended to have higher nucleosome occupancy. Previously, in the wild-type context, a mode of PIC and nucleosome assembly on promoter regions was

proposed to be TATA element-dependent with a competitive assembly for TATA-box-containing promoters and a cooperative assembly for promoters with TATA-like elements (Rhee and Pugh 2012). Our work demonstrates that the Med10 Mediator subunit mutation has an impact on PIC assembly in a TATA element-dependent and nucleosome-dependent manner. It should be noted that the recruitment or stability of TBP that recognizes the TATA elements remained globally unchanged in the *med10* Mediator mutant as well as the Taf1 TFIID-specific subunit. Interestingly, the dynamics of NFR-adjacent nucleosomes -1 and $+1$ (Dion et al. 2007) follow the tendency similar to that of the TATA-box presence. For the mutant to wild-type occupancy ratios of Mediator subunits, the opposite situation occurs. A generally modest effect on Mediator subunit occupancy is accompanied by a higher nucleosome occupancy, depletion for TATA-box-containing promoters, and low dynamics of -1 and $+1$ nucleosomes for the most affected groups. It should be noted that Mediator is recruited by specific TFs on upstream regulatory elements that could have chromatin organization different from that of core promoter elements. Our results suggest that changes in Mediator occupancy that are the consequence of the Med10 Mediator mutation have a reverse relationship with promoter architecture compared with the core PIC components (TFIIIB, TFIIF, TFIIF, and Pol II). We also suggest that the transcriptional effects of Med10 Mediator mutation are related to the function of Mediator in coordination with TFIIIB and other GTFs, leading to coordinated changes in Pol II occupancy and transcription. We envision that Mediator occupancy does not completely reflect its *in vivo* function, suggesting the importance of evaluating the dynamics of Mediator binding in future studies.

Taken together, our results suggest a functional interplay between the PIC assembly mechanisms coordinated by Mediator and the promoter architecture. The TATA-box-containing genes are thought to be characterized by a greater plasticity and flexibility in their expression. Here we show that Mediator has a global impact on the PIC assembly and transcription and that the TATA-box-containing genes with lower nucleosome occupancy and dynamic -1 and $+1$ nucleosomes are the most influenced by Med10 Mediator mutation. From a mechanistic point of view, this work suggests how functional interplay between Mediator, TFIIIB, other GTFs, and the promoter architecture leads to gene-specific transcription. It will help to understand the global to gene-specific rules governing transcription regulation. An integrative analysis of PIC assembly mechanisms using mutants in different Mediator subunits will likely permit future modeling of the PIC assembly pathways at each promoter and prediction of the Mediator-related regulation rules.

Materials and methods

Strains and plasmids

All *Saccharomyces cerevisiae* strains are described in Supplemental Table S1. All plasmids are listed in Supplemental Table

S2. The oligonucleotides used in this study are in Supplemental Table S3.

ChIP and ChIP-seq

ChIP experiments were performed as described previously (Ghavi-Helm et al. 2008). Cell cultures (100 mL) were grown to exponential phase in YPD medium at 30°C, transferred for 90 min to a shaking incubator set at 37°C to allow gradual warming, and cross-linked with 1% formaldehyde for 10 min. The 3HA-tagged proteins were immunoprecipitated with 12CA5 antibody and Pol II with 8WG16 anti-CTD antibody (Covance) bound to IgG magnetic beads (Dynabead). TAP-tagged proteins were immunoprecipitated with IgG magnetic beads (Dynabeads). ChIP-seq experiments were performed as described previously (Eyboulet et al. 2013). Chromatin preparation for ChIP-seq experiments was performed as described previously for conventional ChIP, except that an additional sonication step with a Bioruptor (Diagenode) (six cycles of 40 sec with medium intensity setting) was included to generate DNA fragments of ~200-base-pair (bp) mean size. DNA sequencing of 40-nucleotide tags was performed on a GA-IIx, Hi-Seq, or Next-Seq sequencer using the procedures recommended by the manufacturer (Illumina). Input DNA and DNA from ChIP with an untagged strain were used as negative controls. The ChIP-seq data have been deposited to Array Express under accession number E-MTAB-4607.

Data analysis

ChIP-seq data were analyzed as described previously (Eyboulet et al. 2013, 2015). The number of mappable tags for each ChIP-seq experiment is in Supplemental Table S4. To compare read counts in wild-type and *med10* mutant ChIP-seq data, a count of reads was determined on promoter regions of Pol II transcribed genes for all GTFs and Mediator and on Pol II transcribed gene ORFs for Pol II. For ChIP-seq data analysis, promoter regions were defined as corresponding intergenic regions in tandem or in divergent orientation on the yeast genome. To consider only Pol II transcribed gene promoters, intergenic regions encompassing Pol III transcribed genes were excluded. Read numbers were normalized relative to qPCR data on a set of selected regions (Supplemental Table S5) as described previously (Eyboulet et al. 2015). The median of normalization coefficients between ChIP-qPCR and read ratios used for ChIP-seq data normalization is in Supplemental Table S6.

To aggregate data at the gene level (Fig. 5; Supplemental Fig. S5), tag density (per 1-bp bin) was averaged across whole intergenic regions, corresponding to the regions between tandem genes, except for Pol II, where the averaging was performed over the ORFs. Divergent genes with double peaks for GTFs were excluded from the analysis as well as centromeric regions. R software version 3.2.2 was used for computational analysis. In order to avoid potential biases that could arise from low occupancy values associated with low transcription levels, the genes showing the lowest Pol II occupancy (the first quartile) were excluded from the analyses shown in Figure 5 and Supplemental Figure S5, leading to 1065 remaining genes.

To investigate the amplitude of the *med10* mutation effects on the binding of a particular GTF, Mediator subunit, or Pol II for each of our measurements, the genes were divided into three groups: The lower quartile group with the lowest 25% of mutant versus wild-type ratios comprises 267 genes, and the upper quartile group with the highest 25% comprises another 267 genes, while the interquartile group with the values between 25% and 75% comprises 531 genes. These groups for TFIIB, Pol II, and Me-

diator (Med17) are illustrated in Figure 5, A–C. As the trend of mutant versus wild-type values was not always perfectly linear and in order to correct for this nonlinearity, we developed a script that estimates locally (i.e., for a given x) the median and the interquartile range (IQR) to compute a score obtained as $[y - \text{median}(x)] / \text{IQR}(x)$, where x and y denote the values for a particular gene in the wild type and the mutant, respectively. In practice, for each x , we applied the linear quantile regression implemented in the Quantreg package (version 5.19) to obtain the local median and IQR after weighting the points according to their distance to x with a normal kernel and a bandwidth equal to one-fifth of the total span of the data on the X -axis. The collections of the genes most influenced by the mutation and the least influenced by mutation were defined with this score.

The promoter architecture was determined for each of the three groups. The presence of the TATA boxes and nucleosome enrichment profiles in a 1600-bp window centered on the TSS were taken from Rhee and Pugh (2012). The dynamics of the nucleosomes (Dion et al. 2007) were determined for +1 and –1 nucleosomes, and only the significantly dynamic (hot) nucleosomes were included in the analysis. To determine the significance of the results for the genes containing the TATA box and hot –1 and hot +1 nucleosomes, we performed a Fisher test, while, for the difference between nucleosome enrichment, the P -values were obtained by a Wilcoxon test applied to the maximum value obtained in the 100 bp downstream from the TSS for each gene. Both Fisher test and Wilcoxon test are from the R base package stats.

Spearman correlations between two sets of mutant versus wild-type ratio values on the genes in the GTFs, Mediator subunits, or Pol II were computed and visualized using the heat maps produced with heatmap3 package version 1.1.1.

Acknowledgments

We thank M. Hampsey for *sua7* mutant plasmids; S. Hahn for immobilized template plasmids; F.J. Asturias for the CA001 strain; C. Cibot for Med14 truncation mutants; M.-C. Nevers and H. Volland for anti-Med14 and anti-Med17 antibodies; C. Thermes, E. Van Dijk, and Y. Jaszczyszyn for performing the high-throughput sequencing of ChIP samples; the Service de Pharmacologie et Immunoanalyse (Commissariat à l'Énergie Atomique/Saclay) for monoclonal antibodies; and A. Goldar, C. Gazin, and V. Fromion for fruitful discussions. This work has benefited from the facilities and expertise of the high-throughput sequencing and the Service d'Identification et de Caractérisation des Protéines par Spectrométrie de Masse platforms of the Institute for Integrative Biology of the Cell, Centre National de la Recherche Scientifique. This work was supported by the Agence Nationale de la Recherche (ANR-08-BLAN-0229, ANR 11 BSV8 020 01, and ANR-14-CE10-0012-01) and the Fondation ARC (SL 2201130607079). T.E. was supported by the Fondation pour la Recherche Médicale (FDT20150531985).

References

- Baek HJ, Kang YK, Roeder RG. 2006. Human Mediator enhances basal transcription by facilitating recruitment of transcription factor IIB during preinitiation complex assembly. *J Biol Chem* **281**: 15172–15181.
- Bourbon HM. 2008. Comparative genomics supports a deep evolutionary origin for the large, four-module transcriptional mediator complex. *Nucleic Acids Res* **36**: 3993–4008.
- Buratowski S, Hahn S, Guarente L, Sharp PA. 1989. Five intermediate complexes in transcription initiation by RNA polymerase II. *Cell* **56**: 549–561.

- Cevher MA, Shi Y, Li D, Chait BT, Malik S, Roeder RG. 2014. Reconstitution of active human core Mediator complex reveals a critical role of the MED14 subunit. *Nat Struct Mol Biol* **21**: 1028–1034.
- Dion MF, Kaplan T, Kim M, Buratowski S, Friedman N, Rando OJ. 2007. Dynamics of replication-independent histone turnover in budding yeast. *Science* **315**: 1405–1408.
- Esnault C, Ghavi-Helm Y, Brun S, Soutourina J, Van Berkum N, Boschiero C, Holstege F, Werner M. 2008. Mediator-dependent recruitment of TFIIF modules in preinitiation complex. *Mol Cell* **31**: 337–346.
- Eyboulet F, Cibot C, Eychenne T, Neil H, Alibert O, Werner M, Soutourina J. 2013. Mediator links transcription and DNA repair by facilitating Rad2/XPG recruitment. *Genes Dev* **27**: 2549–2562.
- Eyboulet F, Wydau-Demattis S, Eychenne T, Alibert O, Neil H, Boschiero C, Nevers MC, Volland H, Cornu D, Redeker V, et al. 2015. Mediator independently orchestrates multiple steps of preinitiation complex assembly in vivo. *Nucleic Acids Res* **43**: 9214–9231.
- Flanagan PM, Kelleher RJ III, Sayre MH, Tschochner H, Kornberg RD. 1991. A mediator required for activation of RNA polymerase II transcription in vitro. *Nature* **350**: 436–438.
- Ghavi-Helm Y, Michaut M, Acker J, Aude JC, Thuriaux P, Werner M, Soutourina J. 2008. Genome-wide location analysis reveals a role of TFIIS in RNA polymerase III transcription. *Genes Dev* **22**: 1934–1947.
- Guglielmi B, van Berkum NL, Klapholz B, Bijma T, Boube M, Boschiero C, Bourbon HM, Holstege FC, Werner M. 2004. A high resolution protein interaction map of the yeast Mediator complex. *Nucleic Acids Res* **32**: 5379–5391.
- Han SJ, Lee YC, Gim BS, Ryu GH, Park SJ, Lane WS, Kim YJ. 1999. Activator-specific requirement of yeast mediator proteins for RNA polymerase II transcriptional activation. *Mol Cell Biol* **19**: 979–988.
- Hashimoto S, Boissel S, Zarhrate M, Rio M, Munnich A, Egly JM, Colleaux L. 2011. MED23 mutation links intellectual disability to dysregulation of immediate early gene expression. *Science* **333**: 1161–1163.
- He Y, Fang J, Taatjes DJ, Nogales E. 2013. Structural visualization of key steps in human transcription initiation. *Nature* **495**: 481–486.
- He Y, Yan C, Fang J, Inouye C, Tjian R, Ivanov I, Nogales E. 2016. Near-atomic resolution visualization of human transcription promoter opening. *Nature* **533**: 359–365.
- Holstege FC, Jennings EG, Wyrick JJ, Lee TI, Hengartner CJ, Green MR, Golub TR, Lander ES, Young RA. 1998. Dissecting the regulatory circuitry of a eukaryotic genome. *Cell* **95**: 717–728.
- Jiang YW, Veschambre P, Erdjument-Bromage H, Tempst P, Conaway JW, Conaway RC, Kornberg RD. 1998. Mammalian mediator of transcriptional regulation and its possible role as an end-point of signal transduction pathways. *Proc Natl Acad Sci* **95**: 8538–8543.
- Johnson KM, Wang J, Smallwood A, Arayata C, Carey M. 2002. TFIID and human mediator coactivator complexes assemble cooperatively on promoter DNA. *Genes Dev* **16**: 1852–1863.
- Kang JS, Kim SH, Hwang MS, Han SJ, Lee YC, Kim YJ. 2001. The structural and functional organization of the yeast mediator complex. *J Biol Chem* **276**: 42003–42010.
- Kaufmann R, Straussberg R, Mandel H, Fattal-Valevski A, Ben-Zeev B, Naamati A, Shaag A, Zenvirt S, Konen O, Mimouni-Bloch A, et al. 2010. Infantile cerebral and cerebellar atrophy is associated with a mutation in the MED17 subunit of the transcription preinitiation mediator complex. *Am J Hum Genet* **87**: 667–670.
- Kim YJ, Bjorklund S, Li Y, Sayre MH, Kornberg RD. 1994. A multiprotein mediator of transcriptional activation and its interaction with the C-terminal repeat domain of RNA polymerase II. *Cell* **77**: 599–608.
- Kornberg RD. 2005. Mediator and the mechanism of transcriptional activation. *Trends Biochem Sci* **30**: 235–239.
- Lacombe T, Poh SL, Barbey R, Kuras L. 2013. Mediator is an intrinsic component of the basal RNA polymerase II machinery in vivo. *Nucleic Acids Res* **41**: 9651–9662.
- Larivière L, Plaschka C, Seizl M, Petrotchenko EV, Wenzek L, Borchers CH, Cramer P. 2013. Model of the Mediator middle module based on protein cross-linking. *Nucleic Acids Res* **41**: 9266–9273.
- Liu X, Bushnell DA, Wang D, Calero G, Kornberg RD. 2010. Structure of an RNA polymerase II–TFIIB complex and the transcription initiation mechanism. *Science* **327**: 206–209.
- Malik S, Roeder RG. 2010. The metazoan Mediator co-activator complex as an integrative hub for transcriptional regulation. *Nat Rev Genet* **11**: 761–772.
- Murakami K, Elmlund H, Kalisman N, Bushnell DA, Adams CM, Azubel M, Elmlund D, Levi-Kalisman Y, Liu X, Gibbons BJ, et al. 2013. Architecture of an RNA polymerase II transcription pre-initiation complex. *Science* **342**: 1238724.
- Murakami K, Tsai KL, Kalisman N, Bushnell DA, Asturias FJ, Kornberg RD. 2015. Structure of an RNA polymerase II preinitiation complex. *Proc Natl Acad Sci* **112**: 13543–13548.
- Pinto I, Ware DE, Hampsey M. 1992. The yeast SUA7 gene encodes a homolog of human transcription factor TFIIB and is required for normal start site selection in vivo. *Cell* **68**: 977–988.
- Plaschka C, Larivière L, Wenzek L, Seizl M, Hemann M, Tegunov D, Petrotchenko EV, Borchers CH, Baumeister W, Herzog F, et al. 2015. Architecture of the RNA polymerase II–Mediator core initiation complex. *Nature* **518**: 376–380.
- Plaschka C, Hantsche M, Dienemann C, Burzinski C, Plitzko J, Cramer P. 2016. Transcription initiation complex structures elucidate DNA opening. *Nature* **533**: 353–358.
- Ranish JA, Hahn S. 1996. Transcription: basal factors and activation. *Curr Opin Genet Dev* **6**: 151–158.
- Ranish JA, Yudkovsky N, Hahn S. 1999. Intermediates in formation and activity of the RNA polymerase II preinitiation complex: holoenzyme recruitment and a postrecruitment role for the TATA box and TFIIB. *Genes Dev* **13**: 49–63.
- Rhee HS, Pugh BF. 2012. Genome-wide structure and organization of eukaryotic pre-initiation complexes. *Nature* **483**: 295–301.
- Robinson PJ, Trnka MJ, Pellarin R, Greenberg CH, Bushnell DA, Davis R, Burlingame AL, Sali A, Kornberg RD. 2015. Molecular architecture of the yeast Mediator complex. *Elife* **4**: e08719.
- Sainsbury S, Niesser J, Cramer P. 2013. Structure and function of the initially transcribing RNA polymerase II–TFIIB complex. *Nature* **493**: 437–440.
- Schiano C, Casamassimi A, Rienzo M, de Nigris F, Sommese L, Napoli C. 2014. Involvement of Mediator complex in malignancy. *Biochim Biophys Acta* **1845**: 66–83.
- Soutourina J, Wydau S, Ambroise Y, Boschiero C, Werner M. 2011. Direct interaction of RNA polymerase II and mediator required for transcription in vivo. *Science* **331**: 1451–1454.
- Spaeth JM, Kim NH, Boyer TG. 2011. Mediator and human disease. *Semin Cell Dev Biol* **22**: 776–787.
- Thompson CM, Young RA. 1995. General requirement for RNA polymerase II holoenzymes in vivo. *Proc Natl Acad Sci* **92**: 4587–4590.

Eychenne et al.

- Tsai KL, Tomomori-Sato C, Sato S, Conaway RC, Conaway JW, Asturias FJ. 2014. Subunit architecture and functional modular rearrangements of the transcriptional mediator complex. *Cell* **157**: 1430–1444.
- Wang X, Sun Q, Ding Z, Ji J, Wang J, Kong X, Yang J, Cai G. 2014. Redefining the modular organization of the core Mediator complex. *Cell Res* **24**: 796–808.
- Wu WH, Hampsey M. 1999. An activation-specific role for transcription factor TFIIB in vivo. *Proc Natl Acad Sci* **96**: 2764–2769.
- Wu WH, Pinto I, Chen BS, Hampsey M. 1999. Mutational analysis of yeast TFIIB. A functional relationship between Ssu72 and Sub1/Tsp1 defined by allele-specific interactions with TFIIB. *Genetics* **153**: 643–652.



Functional interplay between Mediator and TFIIB in preinitiation complex assembly in relation to promoter architecture

Thomas Eychenne, Elizaveta Novikova, Marie-Bénédicte Barrault, et al.

Genes Dev. published online September 29, 2016

Access the most recent version at doi:[10.1101/gad.285775.116](https://doi.org/10.1101/gad.285775.116)

Supplemental Material <http://genesdev.cshlp.org/content/suppl/2016/09/29/gad.285775.116.DC1.html>

P<P Published online September 29, 2016 in advance of the print journal.

Creative Commons License This article is distributed exclusively by Cold Spring Harbor Laboratory Press for the first six months after the full-issue publication date (see <http://genesdev.cshlp.org/site/misc/terms.xhtml>). After six months, it is available under a Creative Commons License (Attribution-NonCommercial 4.0 International), as described at <http://creativecommons.org/licenses/by-nc/4.0/>.

Email Alerting Service Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](#).

To subscribe to *Genes & Development* go to:
<http://genesdev.cshlp.org/subscriptions>

Supplemental Materials

Functional interplay between Mediator and TFIIB in preinitiation complex assembly in relation to promoter architecture

Thomas Eychenne, Elizaveta Novikova, Marie-Bénédicte Barrault, Olivier Alibert, Claire Boschiero, Nuno Peixeiro, David Cornu, Virginie Redeker, Laurent Kuras, Pierre Nicolas, Michel Werner and Julie Soutourina

Supplemental Materials and Methods

Yeast strains

All *S. cerevisiae* strains are listed in Supplemental Table S1. *MED10* or *MED14* were deleted in YPH499 (*MATa ura3-52 his3-Δ200 ade2-101uaa trp1-Δ63 lys2-801uag leu2-Δ1*) complemented with pVV208-MED10 or pVV208-MED14, respectively, and replaced by a KanMX6 marker using the standard one-step method (Longtine et al. 1998). *med10* conditional mutant was obtained as previously described for *med17* mutants (Soutourina et al. 2011). We mutagenized *MED10* by error-prone PCR and introduced the mutated fragment in a pVV204 vector bearing the *TRP1* marker by gap-repair in the strain carrying a *med10* deletion complemented by wild-type *MED10* borne on a *URA3* vector (pVV208-MED10). After chasing the wild-type *MED10* allele on 5-FOA, colonies growing at 30°C were replica plated at 37°C to identify thermosensitive colonies. The plasmids were isolated, re-transformed in *med10Δ*/pVV208-MED10 strain and sequenced. Med5-HA, Med15-HA, Med17-HA, Rad3-HA and Kin28-HA and Taf1-HA strains carrying C-terminal HA-tagged version of Med5, Med15 and Med17 Mediator subunits, Rad3 and Kin28 TFIIB subunits and Taf1 were constructed by inserting 3HA epitopes followed by His3MX6 marker using the standard one-step methods. The HA-TBP strain expressing N-terminal HA-tagged version of TBP was obtained by inserting 3HA epitopes preceded by *LEU2* marker. Toa2-TAP, TFIIB-TAP, Tfg1-TAP, Tfa2-TAP strains carrying C-terminal TAP-tagged versions of Toa2, TFIIB,

Tfg1 and Tfa2 were constructed by inserting TAP cassette followed by His3MX6 marker using the one-step standard method.

Double-tagged strains carrying Toa2-Myc and Med5-HA version of Toa2 and Med5 were constructed by first inserting 3HA epitope followed by His3MX6 marker and by inserting 13Myc epitope followed by a *TRP1* marker using standard one step protocol. These double-tagged strains were then transformed by pAG415-MED10 or *med10* mutant plasmid. Med10 mutants and a WT strain for Mediator purification were derived from protease deficient CA001 strain (*MATa leu2 trp1 ura3-52 prb1-1122 pep4-3 prc1-407 gal2 MED22::10His-TEV-ProteinA::KITRP1 MED8-PreSci-3HA::KanR*) carrying C-terminal HA-tagged version of Med8 and C-terminal 10xHis-TEV-ProteinA-tagged version of Med22 (Cai et al. 2009). *MED10* was deleted in CA001 complemented with pVV208-MED10 and replaced by a natNT2 marker from pFA6a-natNT2 (Janke et al. 2004) using the standard one-step method, then transformed by pAG415-GPD-MED10 or *med10*-mutant plasmids. Finally, *MED10 URA3* plasmid was chased on 5-FOA medium.

For synthetic phenotype tests between *med10-196* and *sua7* mutants, a strain was deleted for *med10* by introduction of an *ADE2* cassette and complemented by a centromeric plasmid carrying *MED10* or *med10-196* and a *TRP1* marker. This strain was deleted for *sua7* by insertion of a KanMx6 cassette and complemented by a plasmid carrying *SUA7* and a *URA3* marker. This double-deleted strain was transformed by the plasmids carrying either WT or mutated version of *SUA7* with a *HIS3* marker (Wu et al. 1999). Transformants were spotted on medium containing 5-FOA to chase *SUA7 URA3* plasmid.

Unless otherwise stated, yeast strains were grown at 30°C in YPD rich medium containing 2% of glucose.

Plasmid constructions and cloning

All plasmids are listed in Supplemental Table S2. All cloning experiments were done using the Gateway Invitrogen cloning method. Wild-type *MED10* or *MED14* genes were amplified from YPH499 genomic DNA using oligonucleotides matching the gene sequence initiation codon and following codons for the 5' forward primer and the stop codon and preceding codons for the 3' reverse primer. The oligonucleotides were flanked with *attB1* or *attB2* sequences, respectively. For *med14* C-terminal truncations (*med14-483*, *686*, *752*), the corresponding fragments were amplified by the same procedure with appropriate oligonucleotides matching the gene sequence initiation codon and following codons for the 5' forward primer and the stop codon at amino acid position 483, 686 or 752, respectively, and preceding codons for the 3' reverse primer. The amplified sequence was cloned into pDONR201 (Invitrogen) using standard BP reaction. The recombinant plasmid was sequence verified. The cloned sequence was then transferred into pVV208 (CEN URA3), pVV204 (CEN TRP1) (Van Mullem et al. 2003), or pAG415-GPD vector (Alberti et al. 2007) by the LR reaction.

Two hybrid assay

Two-hybrid experiments were done as described previously (Werner et al. 1993). Briefly, the Y190 yeast strain was co-transformed by pVV213 and pVV212 carrying the bait protein fused with Gal4-DNA-Binding-Domain and the prey fused to Gal4-Activating-Domain, respectively. The activation of the *lacZ* or *HIS3* reporters was tested using an X-Gal overlay plate assay or 3AT plate assay.

For quantitative analysis, β -galactosidase was assayed according to the Miller method using Yeast β -galactosidase assay kit (Thermo Scientific) with *o*-nitrophenyl- β -D-galactopyranoside (ONPG) as a substrate, following the manufacturer's instructions.

Alternatively, when indicated, a more sensitive β -galactosidase assay using chlorophenol red- β -galactopyranoside (CPRG, Roche) as a substrate was performed. Briefly, cells from 1.5 ml of culture grown to mid-log phase were harvested by centrifugation, washed with 1 ml of buffer A (10 mM HEPES, pH 7.3, 150 mM NaCl, 50 mM L-Aspartate, 1% BSA, 0.05% Tween 20) and resuspended in the 300 μ l of the same buffer. The OD₆₀₀ was measured at this stage. 100 μ l of cell suspension were transferred to a microfuge tube and flash-frozen in liquid nitrogen. Frozen pellet was immediately thawed in a 37°C water bath. The freeze and thaw cycle was repeated twice to break the cells. 700 μ l of buffer B (2.23 mM of CPRG in Buffer A) were added, mixed thoroughly and incubated at room temperature overnight. The reaction was stopped by addition of 500 μ l of 3 mM ZnCl₂. After centrifugation to pellet cell debris, absorbance at 570nm was measured. Miller units were calculated according to (Miller 1972). The mean values and standard deviation (indicated by error bars) of three independent experiments were calculated.

Coimmunoprecipitation experiments

Whole yeast extracts were prepared from 100 ml of cells grown to exponential phase in YPD medium at 30°C. When indicated, the cells were then transferred for 90 min to 37°C. Whole yeast extract preparation, immunoprecipitation (IP) in IP buffer (50 mM HEPES, pH 7.5, 100 mM NaCl, 20% glycerol, 1 mM dithiothreitol, 0.5 mM EDTA, 0.05% NP-40) supplemented with a protease inhibitor cocktail (Complete, Roche) and 1 mM PMSF and western blotting were performed as described previously (Soutourina et al. 2006).

In some cases, whole cell extract were extracted by disrupting yeast cells in lysis buffer (50 mM Tris-HCl [pH 7.5], 15% glycerol, 5mM MgCl₂) in Mikro-Dismembrator S (Sartorius Stedim Biotech S.A.). After centrifugation at 13,000 \times g for 20 min, the resulting supernatant plus 100 mM NaCl and Igepal 0.05% was used as input. HA-tagged proteins were

immunoprecipitated using HA isolation kit (Miltenyi Biotec), following the manufacturer's instructions. All washes were performed using the lysis buffer.

The 12CA5 anti-HA antibodies were used against HA-tagged proteins; the 9E10 monoclonal antibody was used against Myc-tagged proteins; rabbit-polyclonal anti-Sua7p (Abcam) was used against Sua7; rabbit polyclonal anti-Med14 and anti-Med17 antibodies against corresponding N-terminal peptides were used to detect Med14 and Med17 Mediator subunits, respectively.

The relative intensity of immune staining was quantified using ImageJ version 1.46r. The intensity of immune staining for coimmunoprecipitated TFIIB signals relative to the wild-type was normalized against immunoprecipitation signals. The mean values and standard deviation (indicated by error bars) of three independent experiments were calculated.

Mediator purification

Mediator purification was done as previously described (Eyboulet et al. 2015). *MED10* WT and *med10-196* strains were grown in 10 liters of 2X YPD overnight to late exponential phase at 30°C or transferred for 90 min to 37°C.

Mass spectrometry

To analyze Mediator integrity in *med10* mutant, the purified Mediator complex was analyzed by mass spectrometry. A short migration of proteins from the Mediator complex was performed in NuPage 4–12% gel (Life Technologies), about 1 cm below the stacking gel. After coomassie blue staining, the 1 cm band was cut, and proteins were digested overnight using trypsin. Tryptic peptides were extracted with 60% acetonitrile and 0.1% (v/v) formic acid, vacuum dried and resuspended in 0.1% (v/v) formic acid prior to nanoLC–MS/MS mass spectrometry analyses.

The nanoLC–MS/MS analyses were performed with the Triple-TOF 4600 mass spectrometer (AB Sciex, Framingham, MA, USA) coupled to the nanoRSLC ultra performance liquid chromatography (UPLC) system (Thermo Scientific) equipped with a trap column (Acclaim PepMap100C18, 75 μm i.d. \times 2 cm, 3 μm) and an analytical column (Acclaim PepMapRSLCC18, 75 μm i.d. \times 25 cm, 2 μm , 100 \AA). Peptide separation was performed with a 5–35% solvent B gradient for 40 min. Solvent A was 0.1% formic acid in water, and solvent B was 0.1% formic acid in 100% acetonitrile. The nanoLC–MS/MS experiments were conducted using the data-dependent acquisition method by selecting the 20 most intense precursors for CID fragmentation with Q1 quadrupole set at low resolution for better sensitivity. Raw data were processed using MS Data Converter software (AB Sciex) for generating .mgf data files. Protein identification was performed using the MASCOT database search engine (Matrix science, London, UK) against the Swissprot database (release 2016_05) with trypsin specificity (with 2 missed cleavages), carbamidomethylation of cysteines and oxidation of methionines set as fixed and variable modifications, respectively. Peptide and fragment tolerance were respectively set at 20 ppm and 0.05 Da. Results of three replicates were merged. Only proteins with at least two unique peptides with mascot ion scores above identity threshold (25) at less than 1% FDR were considered.

In vitro assay of preinitiation complex assembly

The *in vitro* analysis of preinitiation complex was done as previously described (Ranish et al. 1999). Briefly, a biotin-labeled-PCR-generated DNA fragment was used as a template and was bound to Streptavidin beads (Dynabeads). This matrix was incubated with purified Gal4-Gcn4 activator and a nuclear extract from WT or mutant strains. Activator was purified as described and nuclear extract was prepared with standard procedure (Ranish et al. 1999; Reeves and Hahn 2003). After assembly on the DNA matrix, the reactions were washed

and eluted from the beads by restriction enzyme digestion (PstI, Invitrogen). Condition without DNA was used as a control. Condition with DNA matrix but without activator was used as a control of basal transcription. When indicated, purified Mediator was added to the reaction. Eluates were analyzed by western blotting. Rabbit polyclonal anti-Med14 and anti-Med17 antibodies were used against corresponding N-terminal peptides to detect Med14 and Med17 Mediator subunits, respectively. 12CA5 monoclonal antibody was used to detect HA-tagged protein. Mouse monoclonal antibody 8WG16 and 1Y26 were used to detect Pol II subunits Rpb1 and Rpb3, respectively. Rabbit-polyclonal anti-Sua7p (Abcam) was used to detect Sua7, rabbit-polyclonal anti-Toa1 (Abcam) was used to detect Toa1, rabbit-polyclonal antibody anti-Kin28 (Biolegend) was used to detect Kin28. Rabbit polyclonal antibody against H3 histone (Abcam) was used as a loading control of nuclear extract.

Supplemental Table S1. Yeast strains

Name	Tagged or Mutant protein	Genotype	Reference
YPH499	WT	<i>MATa ura3-52 his3-Δ200 ade2-101uaa trp1-Δ63 lys2-801uag leu2-Δ1</i>	(Sikorski and Hieter 1989)
Y4987	WT	<i>MATa ura3-52 his3-Δ200 ade2-101uaa trp1-Δ63 lys2-801uag leu2-Δ1 med10::KanMX6//MED10 CEN URA3</i>	This work
Y5466	WT	<i>MATa ura3-52 his3-Δ200 ade2-101uaa trp1-Δ63 lys2-801uag leu2-Δ1 med10::KanMX6//MED10 CEN TRP1</i>	This work
Y5467	med10-196	<i>MATa ura3-52 his3-Δ200 ade2-101uaa trp1-Δ63 lys2-801uag leu2-Δ1 med10::KanMX6//med10-196 CEN TRP1</i>	This work
Y5801	Med5-HA	<i>MATa ura3-52 his3-Δ200 ade2-101uaa trp1-Δ63 lys2-801uag leu2-Δ1 med10::KanMX6 MED5::3HA::HIS3 // MED10 CEN URA3</i>	This work
Y5802	Med15-HA	<i>MATa ura3-52 his3-Δ200 ade2-101uaa trp1-Δ63 lys2-801uag leu2-Δ1 med10::KanMX6 MED15::3HA::HIS3 // MED10 CEN URA3</i>	This work
Y5803	Med17-HA	<i>MATa ura3-52 his3-Δ200 ade2-101uaa trp1-Δ63 lys2-801uag leu2-Δ1 med10::KanMX6 MED17::3HA::HIS3 // MED10 CEN URA3</i>	This work
Y5804	Rad3-HA	<i>MATa ura3-52 his3-Δ200 ade2-101uaa trp1-Δ63 lys2-801uag leu2-Δ1 med10::KanMX6 RAD3::3HA::HIS3 // MED10 CEN URA3</i>	This work
Y5805	Kin28-HA	<i>MATa ura3-52 his3-Δ200 ade2-101uaa trp1-Δ63 lys2-801uag leu2-Δ1 med10::KanMX6 KIN28::3HA::HIS3 // MED10 CEN URA3</i>	This work

Y5806	HA-TBP	<i>MATa ura3-52 his3-Δ200 ade2-101uaa trp1-Δ63 lys2-801uag leu2-Δ1 med10::KanMX6 LEU2::3HA::SPT15 // MED10 CEN URA3</i>	This work
Y6581	Taf1-HA	<i>MATa ura3-52 his3-Δ200 ade2-101uaa trp1-Δ63 lys2-801uag leu2-Δ1 med10::KanMX6 TAF1::3HA::HIS3 // MED10 CEN URA3</i>	This work
Y5859	Med5-HA	<i>MATa ura3-52 his3-Δ200 ade2-101uaa trp1-Δ63 lys2-801uag leu2-Δ1 med10::KanMX6 MED5::3HA::HIS3 // MED10 CEN TRP1</i>	This work
Y5860	Med5-HA <i>med10-196</i>	<i>MATa ura3-52 his3-Δ200 ade2-101uaa trp1-Δ63 lys2-801uag leu2-Δ1 med10::KanMX6 MED5::3HA::HIS3 // med10-196 CEN TRP1</i>	This work
Y5861	Med15-HA	<i>MATa ura3-52 his3-Δ200 ade2-101uaa trp1-Δ63 lys2-801uag leu2-Δ1 med10::KanMX6 MED15::3HA::HIS3 // MED10 CEN TRP1</i>	This work
Y5862	Med15-HA <i>med10-196</i>	<i>MATa ura3-52 his3-Δ200 ade2-101uaa trp1-Δ63 lys2-801uag leu2-Δ1 med10::KanMX6 MED15::3HA::HIS3 // med10-196 CEN TRP1</i>	This work
Y5863	Med17-HA	<i>MATa ura3-52 his3-Δ200 ade2-101uaa trp1-Δ63 lys2-801uag leu2-Δ1 med10::KanMX6 MED17::3HA::HIS3 // MED10 CEN TRP1</i>	This work
Y5864	Med15-HA <i>med10-196</i>	<i>MATa ura3-52 his3-Δ200 ade2-101uaa trp1-Δ63 lys2-801uag leu2-Δ1 med10::KanMX6 MED17::3HA::HIS3 // med10-196 CEN TRP1</i>	This work
Y5865	Rad3-HA	<i>MATa ura3-52 his3-Δ200 ade2-101uaa trp1-Δ63 lys2-801uag leu2-Δ1 med10::KanMX6 RAD3::3HA::HIS3 // MED10 CEN TRP1</i>	This work
Y5866	Rad3-HA <i>med10-196</i>	<i>MATa ura3-52 his3-Δ200 ade2-101uaa trp1-Δ63 lys2-801uag leu2-Δ1 med10::KanMX6 RAD3::3HA::HIS3 // med10-196 CEN TRP1</i>	This work
Y5867	Kin28-HA	<i>MATa ura3-52 his3-Δ200 ade2-101uaa trp1-Δ63 lys2-801uag leu2-Δ1 med10::KanMX6 KIN28::3HA::HIS3 // MED10 CEN TRP1</i>	This work
Y5868	Kin28-HA <i>med10-196</i>	<i>MATa ura3-52 his3-Δ200 ade2-101uaa trp1-Δ63 lys2-801uag leu2-Δ1 med10::KanMX6 KIN28::3HA::HIS3 // med10-196 CEN TRP1</i>	This work
Y5869	HA-TBP	<i>MATa ura3-52 his3-Δ200 ade2-101uaa trp1-Δ63 lys2-801uag leu2-Δ1 med10::KanMX6 LEU2::3HA::SPT15 // MED10 CEN TRP1</i>	This work
Y5870	HA-TBP <i>med10-196</i>	<i>MATa ura3-52 his3-Δ200 ade2-101uaa trp1-Δ63 lys2-801uag leu2-Δ1 med10::KanMX6 LEU2::3HA::SPT15 // med10-196 CEN TRP1</i>	This work
Y6596	Taf1-HA	<i>MATa ura3-52 his3-Δ200 ade2-101uaa trp1-Δ63 lys2-801uag leu2-Δ1 med10::KanMX6 TAF1::3HA::HIS3 // MED10 CEN TRP1</i>	This work
Y6597	Taf1-HA <i>med10-196</i>	<i>MATa ura3-52 his3-Δ200 ade2-101uaa trp1-Δ63 lys2-801uag leu2-Δ1 med10::KanMX6 TAF1::3HA::HIS3 // MED10 CEN TRP1</i>	This work
CA001	Med22-10HisTEVProteinA Med8-HA	<i>MATa leu2 trp1 ura3-52 prb1-1122 pep4-3 prc1-407 gal2 MED22 ::10His-TEV-ProteinA::KITRP1 MED8-PreSci-3HA::KanR</i>	(Cai et al. 2009)
Y6464	Med22-10HisTEVProteinA Med8-HA <i>MED10</i>	<i>MATa leu2 trp1 ura3-52 prb1-1122 pep4-3 prc1-407 gal2 MED22 ::10His-TEV-ProteinA::KITRP1 MED8-PreSci-3HA::KanR // MED10 CEN URA3</i>	This work
Y6491	Med22-10HisTEVProteinA Med8-HA <i>MED10</i>	<i>MATa leu2 trp1 ura3-52 prb1-1122 pep4-3 prc1-407 gal2 MED22 ::10His-TEV-ProteinA::KITRP1 MED8-PreSci-3HA::KanR // MED10 CEN LEU</i>	This work

Y6492	Med22-10HisTEVProteinA Med8-HA <i>med10-196</i>	<i>MATa leu2 trp1 ura3-52 prb1-1122 pep4-3 prc1-407 gal2 MED22 ::10His-TEV-ProteinA::KITRP1 MED8-PreSci-3HA::KanR // med10-196 CEN LEU</i>	This work
Y6826	Toa2-TAP	<i>MATa ura3-52 his3-Δ200 ade2-101uaa trp1-Δ63 lys2-801uag leu2-Δ1 med10::KanMX6 TOA2::TAP::HIS3 // MED10 CEN URA3</i>	This work
Y6600	Tfg1-TAP	<i>MATa ura3-52 his3-Δ200 ade2-101uaa trp1-Δ63 lys2-801uag leu2-Δ1 med10::KanMX6 TFG1::TAP::HIS3 // MED10 CEN URA3</i>	This work
Y6603	Tfa2-TAP	<i>MATa ura3-52 his3-Δ200 ade2-101uaa trp1-Δ63 lys2-801uag leu2-Δ1 med10::KanMX6 TFA2::TAP::HIS3 // MED10 CEN URA3</i>	This work
Y6493	TFIIB-TAP	<i>MATa ura3-52 his3-Δ200 ade2-101uaa trp1-Δ63 lys2-801uag leu2-Δ1 med10::KanMX6 SUA7::TAP::HIS3 // MED10 CEN URA3</i>	This work
Y6665	Toa2-TAP <i>MED10</i>	<i>MATa ura3-52 his3-Δ200 ade2-101uaa trp1-Δ63 lys2-801uag leu2-Δ1 med10::KanMX6 TOA2::TAP::HIS3 // MED10 CEN TRP1</i>	This work
Y6666	Toa2-TAP <i>med10-196</i>	<i>MATa ura3-52 his3-Δ200 ade2-101uaa trp1-Δ63 lys2-801uag leu2-Δ1 med10::KanMX6 TOA2::TAP::HIS3 // med10-196 CEN TRP1</i>	This work
Y6608	Tfg1-TAP <i>MED10</i>	<i>MATa ura3-52 his3-Δ200 ade2-101uaa trp1-Δ63 lys2-801uag leu2-Δ1 med10::KanMX6 TFG1::TAP::HIS3 // MED10 CEN TRP1</i>	This work
Y6609	Tfg1-TAP <i>med10-196</i>	<i>MATa ura3-52 his3-Δ200 ade2-101uaa trp1-Δ63 lys2-801uag leu2-Δ1 med10::KanMX6 TFG1::TAP::HIS3 // med10-196 CEN TRP1</i>	This work
Y6606	Tfa2-TAP <i>MED10</i>	<i>MATa ura3-52 his3-Δ200 ade2-101uaa trp1-Δ63 lys2-801uag leu2-Δ1 med10::KanMX6 TFA2::TAP::HIS3 // MED10 CEN TRP1</i>	This work
Y6607	Tfa2-TAP <i>med10-196</i>	<i>MATa ura3-52 his3-Δ200 ade2-101uaa trp1-Δ63 lys2-801uag leu2-Δ1 med10::KanMX6 TFA2::TAP::HIS3 // med10-196 CEN TRP1</i>	This work
Y6667	TFIIB-TAP <i>MED10</i>	<i>MATa ura3-52 his3-Δ200 ade2-101uaa trp1-Δ63 lys2-801uag leu2-Δ1 med10::KanMX6 SUA7::TAP::HIS3 // MED10 CEN TRP1</i>	This work
Y6668	TFIIB-TAP <i>med10-196</i>	<i>MATa ura3-52 his3-Δ200 ade2-101uaa trp1-Δ63 lys2-801uag leu2-Δ1 med10::KanMX6 SUA7::TAP::HIS3 // med10-196 CEN TRP1</i>	This work
BY4741	WT	<i>MATa uraD0 his3D1 lysD0 leu2D0</i>	(Brachmann et al. 1998)
Y6809	WT	<i>MATa uraD0 his3D1 lysD0 leu2D0 med10::KanMX6 // MED10 CEN URA3</i>	This work
Y6810	Rad3-HA <i>MED10</i>	<i>MATa uraD0 his3D1 lysD0 leu2D0 med10::KanMX6 RAD3::3HA::HIS3Mx6 // MED10 CEN URA3</i>	This work
Y6829	Rad3-HA <i>MED10</i>	<i>MATa uraD0 his3D1 lysD0 leu2D0 med10::KanMX6 RAD3::3HA::HIS3Mx6 // MED10 CEN LEU2</i>	This work
Y6830	Rad3-HA <i>med10-196</i>	<i>MATa uraD0 his3D1 lysD0 leu2D0 med10::KanMX6 RAD3::3HA::HIS3Mx6 // med10-196 CEN LEU2</i>	This work
Y6972	Toa2-Myc Med5-HA	<i>MATa ura3-52 his3-Δ200 ade2-101uaa trp1-Δ63 lys2-801uag leu2-Δ1 med10::KanMX6 MED5::3HA::HIS3 TOA2::13Myc::TRP1 // MED10 CEN URA3</i>	This work
Y6991	Toa2-Myc Med5-HA <i>MED10</i>	<i>MATa ura3-52 his3-Δ200 ade2-101uaa trp1-Δ63 lys2-801uag leu2-Δ1 med10::KanMX6 MED5::3HA::HIS3 TOA2::13Myc::TRP1 //</i>	This work

MED10 CEN LEU2

Y6992	Toa2-Myc Med5-HA <i>med10-196</i>	<i>MATa ura3-52 his3-Δ200 ade2-101uaa trp1-Δ63 lys2-801uag leu2-Δ1 med10::KanMX6 MED5::3HA::HIS3 TOA2::13Myc::TRP1 // med10-196 CEN LEU2</i>	This work
Y6864	<i>MED10 WT</i>	<i>MATa ura3-52 his3-Δ200 ade2-101uaa trp1-Δ63 lys2-801uag leu2-Δ1 med10::ADE2 sua7::KanMX6// MED10 CEN TRP1 SUA7 CEN URA3</i>	This work
Y6865	<i>med10-196</i>	<i>MATa ura3-52 his3-Δ200 ade2-101uaa trp1-Δ63 lys2-801uag leu2-Δ1 med10::ADE2 sua7::KanMX6// med10-196 CEN TRP1 SUA7 CEN URA3</i>	This work
Y5373	<i>MED14 WT</i>	<i>MATa ura3-52 his3-Δ200 ade2-101uaa trp1-Δ63 lys2-801uag leu2-Δ1 med14::KanMX6// MED14 CEN URA3</i>	This work
Y5669	<i>MED14 WT</i>	<i>MATa ura3-52 his3-Δ200 ade2-101uaa trp1-Δ63 lys2-801uag leu2-Δ1 med14::KanMX6// MED14 CEN TRP1</i>	This work
Y5665	<i>med14-483</i>	<i>MATa ura3-52 his3-Δ200 ade2-101uaa trp1-Δ63 lys2-801uag leu2-Δ1 med14::KanMX6// med14-Δ483 CEN TRP1</i>	This work
Y5666	<i>med14-686</i>	<i>MATa ura3-52 his3-Δ200 ade2-101uaa trp1-Δ63 lys2-801uag leu2-Δ1 med14::KanMX6//med14-Δ686 CEN TRP1</i>	This work
Y5668	<i>med14-752</i>	<i>MATa ura3-52 his3-Δ200 ade2-101uaa trp1-Δ63 lys2-801uag leu2-Δ1 med14::KanMX6// med14-Δ752 CEN TRP1</i>	This work
Y7057	<i>MED14 WT</i>	<i>MATa ura3-52 his3-Δ200 ade2-101uaa trp1-Δ63 lys2-801uag leu2-Δ1 med14::ADE2 sua7::KanMX6// MED14 CEN TRP1 SUA7 CEN URA3</i>	This work
Y7058	<i>med14-483</i>	<i>MATa ura3-52 his3-Δ200 ade2-101uaa trp1-Δ63 lys2-801uag leu2-Δ1 med14::ADE2 sua7::KanMX6// med14-Δ483 CEN TRP1 SUA7 CEN URA3</i>	This work
Y7059	<i>med14-686</i>	<i>MATa ura3-52 his3-Δ200 ade2-101uaa trp1-Δ63 lys2-801uag leu2-Δ1 med14::ADE2 sua7::KanMX6//med14-Δ686 CEN TRP1 SUA7 CEN URA3</i>	This work
Y7060	<i>med14-752</i>	<i>MATa ura3-52 his3-Δ200 ade2-101uaa trp1-Δ63 lys2-801uag leu2-Δ1 med14::ADE2 sua7::KanMX6// med14-Δ752 CEN TRP1 SUA7 CEN URA3</i>	This work

* Plasmid descriptions are separated from chromosomal genotype by //.

Supplemental Table S2. Plasmids

Name	Genotype	Reference
pVV208-MED10	<i>Amp CEN URA3 MED10</i>	This work
pVV204-MED10	<i>Amp CEN TRP1 MED10</i>	This work
pVV204-med10-196	<i>Amp CEN TRP1 med10-196</i>	This work
pAG-GPD-MED10	<i>Amp CEN LEU2 MED10</i>	This work
pAG-GPD-med10-196	<i>Amp CEN LEU2 med10-196</i>	This work
pVV212	<i>Amp 2μ TRP1 pADHI GAL4-BD tADHI</i>	
pVV213	<i>Amp 2μ LEU2 pADHI GAL4-AD tADHI</i>	
pVV213-MED1	<i>Amp 2μ LEU2 pADHI GAL4-AD tADHI MED1</i>	This work
pVV213-MED4	<i>Amp 2μ LEU2 pADHI GAL4-AD tADHI MED4</i>	This work
pVV213-MED7	<i>Amp 2μ LEU2 pADHI GAL4-AD tADHI MED7</i>	This work
pVV213-MED9	<i>Amp 2μ LEU2 pADHI GAL4-AD tADHI MED9</i>	This work
pVV213-MED10	<i>Amp 2μ LEU2 pADHI GAL4-AD tADHI MED10</i>	This work
pVV213-med10-196	<i>Amp 2μ LEU2 pADHI GAL4-AD tADHI med10-196</i>	This work
pVV213-MED14	<i>Amp 2μ LEU2 pADHI GAL4-AD tADHI MED14</i>	This work
pVV213-MED21	<i>Amp 2μ LEU2 pADHI GAL4-AD tADHI MED21</i>	This work
pVV213-MED31	<i>Amp 2μ LEU2 pADHI GAL4-AD tADHI MED31</i>	This work
pVV212-SUA7	<i>Amp 2μ TRP1 pADHI GAL4-BD tADHI SUA7</i>	This work
pVV212-sua7-11	<i>Amp 2μ TRP1 pADHI GAL4-BD tADHI sua7-11</i>	This work
pVV212-sua7-34	<i>Amp 2μ TRP1 pADHI GAL4-BD tADHI sua7-34</i>	This work
pVV212-sua7-36	<i>Amp 2μ TRP1 pADHI GAL4-BD tADHI sua7-36</i>	This work
pVV208-MED14	<i>Amp CEN URA3 MED14</i>	This work
pVV204-MED14	<i>Amp CEN TRP1 MED14</i>	This work
pVV204-med14-483	<i>Amp CEN TRP1 med14-483</i>	This work
pVV204-med14-686	<i>Amp CEN TRP1 med14-686</i>	This work
pVV204-med14-752	<i>Amp CEN TRP1 med14-752</i>	This work
pDW11	<i>Amp CEN SUA7 URA3</i>	This work
pM299	<i>Amp CEN SUA7 HIS3</i>	(Wu et al. 1999)
pM326	<i>Amp CEN sua7-1 HIS3</i>	(Wu et al. 1999)
pM365	<i>Amp CEN sua7-8 HIS3</i>	(Wu et al. 1999)
pM376	<i>Amp CEN sua7-11 HIS3</i>	(Wu et al. 1999)
pM377	<i>Amp CEN sua7-12 HIS3</i>	(Wu et al. 1999)
pM378	<i>Amp CEN sua7-13 HIS3</i>	(Wu et al. 1999)
pM392	<i>Amp CEN sua7-18 HIS3</i>	(Wu et al. 1999)
pM401	<i>Amp CEN sua7-24 HIS3</i>	(Wu et al. 1999)
pM404	<i>Amp CEN sua7-28 HIS3</i>	(Wu et al. 1999)
pM415	<i>Amp CEN sua7-20 HIS3</i>	(Wu et al. 1999)
pM502	<i>Amp CEN sua7-31 HIS3</i>	(Wu et al. 1999)
pM505	<i>Amp CEN sua7-34 HIS3</i>	(Wu et al. 1999)
pM506	<i>Amp CEN sua7-35 HIS3</i>	(Wu et al. 1999)
pM507	<i>Amp CEN sua7-36 HIS3</i>	(Wu et al. 1999)
pM512	<i>Amp CEN sua7-41 HIS3</i>	(Wu et al. 1999)
pM513	<i>Amp CEN sua7-42 HIS3</i>	(Wu et al. 1999)
pM515	<i>Amp CEN sua7-44 HIS3</i>	(Wu et al. 1999)

Supplemental Table S3. Oligonucleotides

Name	Forward	Reverse
ATP1-O	TCTTGCAGTCGGTGATGGTA	TATCGGAACCGAAAAGAACG
ALD6-P	AGCCGACAAAAGAAAAACGA	CGTTCCCAAGAGGAGATCAA
ALD6-O	CCCATTTGGTCTTTGACGAT	GCCTTGAAAGCAGCCAATAG
SRM1-P	GGCAAGGCAGGATGAATAAA	ATTGGTGGCGACTGTTCTTT
QCR6-O	GATGACGATAACGAGCAGCA	TACCCTCCTCCGTGTTCTTG
PIL1-O	GATTCTTTGGGAAGGGTGGT	ATCGTCATCGTTTTCCAACC
MGR1-P	TGCATGGCATCGTCTTTTAC	GTATGGTCGGTGGAATTGCT
RPA34-P	CATCAGGGCCAATCAATTCT	TTTCGACTTCAGCATTGCAC
PSA1-P	CTGCACCCGATCCTTCTTAC	TTCTGTTTTCGCGTTCCTCA
ALD6-P	ATGCGATATAGCACCGACCA	CTTGACCTCGAACGGTGTTT
ACT1-P	TCCACGTCTCTTGCATAAA	GGTTTGAGTAGAAAGGGGAAGG
APA1-P	GAGCGGCCTGAAATACTGTC	AACGTTCCGAAACAGGACAC
PIM1-P	GCCACCGGAAGGAAATAAGT	GTTCAAGCACCTCGAAAACC
STI1-P	CCAAAAGTCTGCTCCCAAAT	TGCAGCGTTACCTTGTTGTT
YIP5-P	CAACGTCTACGCTCAAGGT	CCCTCGAGATCATCGTCAAT
PSA1-P1	CTGCACCCGATCCTTCTTAC	TTCTGTTTTCGCGTTCCTCA
PSA1-P2	GGAGCCACCACGTTTATTTT	GTTTACAAGAGCGCCACGAG
PRB1-P	CCGGGTTAGCAGAGTAGCAG	CTCTTTATCGCGGGTGTGTT
HSP42-P	GGGAGGCCTCTGTGAAGTTA	GCCTGAACGTGTCCCTATGT
HSP42-O	CACCTGGTAACGCAAGAACA	CATCGTTCGCTCACCTACAA
HSP30-P1	ATCCCGATCCCGACTCTTAT	GGCCAATTAGGGAGAGGAAA
HSP30-P2	ATTTTGTTGGCCATTTTCCA	CGGGATATGGCTTTGCTTAC
HSP30-O	CAAGGTGTGATATGCCAACG	ATAGCCTCACCGTCTGGTTG
HSP150-P1	GGAACACTTGAAGTCTAACGACA	AAGCAGAGGCAACCAAAGTC
HSP150-P2	ATTATCCTGGCCGCTAATCC	GTGGCAAAGCAGTGAGATTG
HSP104-P1	CCCATCTCAAAGAACTGCAA	GGAACAAGTGACAAAGGAACG
HSP104-O	GGTACCACGGCAGGTTATGT	CCGTCACTCAAACATTTGTAGC
HSP12-O	CAAGGGTGTCTTCCAAGGTG	CGACCGGAAACATATTCGAC
ADH1-P1	ATAGGCGCATGCAACTTCTT	CATCAGCTCTGGAACAACGA
ADH1-P2	TTCCTTCATTCACGCACACT	AGGGAACGAGAACAATGACG
ADH1-O	GGGTATTGACGGTGGTGAAG	AAACGTTGATGACACCGTGA
PYK1-P1	CGCACCGTCACAAAGTGTT	TGGGAAGGAAAGGAAATCAC
PYK1-P2	CCTTTCCTTCCCATATGATGC	ACTTTGAAAGGGGACCATGA
PYK1-O	TGCTTTGAGAAAGGCTGGTT	TCTGATTTCTGGACCCTTGG
PMA1-O	GGTTTTGGTCATTGCCACTT	ACGGCCATAGTGGTGGTAAC
PMA1-P1	AACAAACCCGGTCTCGAAG	GAAGTGCCGCATTAGGAAAT
PMA1-P2	GATGGTGGGTACCGCTTATG	TTGGTGTTATAGGAAAGAAAGAG
GAL1-P	ACGCTTAACTGCTCATTGCT	TGTTCCGAGCAGTGCGGCGC
GAL1-O1	AAAGAACTTGCACCGGAAA	GGCCCATATTCGTTTAAACA
GAL1-O2	ACATTTCCACACCCTGGAAC	TTCTTCGCGAGAACAATTCA
PMA1-O2	TCTCCAAAGCCGTTAAATG	CCGTTTATAGCACCGAAGTT

Supplemental Table S4. Total number of mapped reads for ChIP-seq experiments

Sample	Million of mapped reads
INPUT-MED10	13.6
INPUT-med10-196	12.8
NT	1.9
PolII-MED10	9.0
PolII-med10-196	10.0
TFIIB-MED10	15.9
TFIIB-med10-196	8.8
Rad3-MED10	6.4
Rad3-med10-196	7.7
Kin28-MED10	2.5
Kin28-med10-196	3.4
Med15-MED10	3.5
Med15-med10-196	4.0
Med17-MED10	4.3
Med17-med10-196	4.5
TBP-MED10	5.7
TBP-med10-196	3.4
Toa2-MED10	10.2
Toa2-med10-196	10.5
Tfa2-MED10	5.5
Tfa2-med10-196	5.1
Tfg1-MED10	6.5
Tfg1-med10-196	3.5
Taf1-MED10	5.1
Taf1-med10-196	3.8

Supplemental Table S5. Regions used for ChIP-seq data normalization

Protein	Regions
Pol II	ADH1-O, ATP1-O, ALD6-O, GAL1-O2, HSP150-O, HSP30-O, HSP42-O,UTP20-O, PMA1-O, PYK1-O, PHO84-O
Kin28, Rad3	ADH1-P2, GAL1-O2, HSP104-P1, HSP30-P1, HSP42-P, MGR1-P, PMA1-P2, PYK1-P2, MGR1-P RPA34-P, SRM1-P
TBP	ADH1-P2, GAL1-O2, PIM1-P, PMA1-P2,PRB1-P PYK1-P2, RPA34-P,STI1-P, YIP5-P, PSA1-P2
Med15	ADH1-P1, ARG3-P1, FLC1-P, GAL1-O1, PMA1-P1, PYK1-P1, HSP42-P, HSP150-P2, HSP30-P2, PSA1-P1, PRB1-P
Med17	ADH1-P1, ARG3-P1, FLC1-P, GAL1-O1, PMA1-P1, PYK1-P1, HSP42-P, HSP150-P2, HSP30-P2, PSA1-P1, PRB1-P
TFIIB	ADH1-P, ALD6-P, APA1-P, ATP1-P, GAL1-P, PMA1-P, PRB1-P, PSA1-P, PYK1-P, RPA34-P, SRM1-P, STI1-P
Tfg1	ADH1-P, ALD6-P, APA1-P, ATP1-P, GAL1-P, PMA1-P, PRB1-P, PSA1-P,PYK1-P, RPA34-P, SRM1-P, STI1-P
Tfa2	ADH1-P, ALD6-P, APA1-P, ATP1-P, GAL1-P, PMA1-P, PRB1-P, PSA1-P, PYK1-P, RPA34-P, SRM1-P, STI1-P
Toa2	ADH1-P, ALD6-P, APA1-P, ATP1-P, GAL1-P, PMA1-P, PRB1-P, PSA1-P, PYK1-P, RPA34-P, SRM1-P, STI1-P
Taf1	HSP150-P1, ALD6-P, YIP5-P, PIM1-P, HSP12-P, PSA1-P, ACT1-P, QCR6-P, PRB1-P, PMA1-P, ADH1-P, GAL1-O2

Supplemental Table S6. Normalization coefficients for ChIP-seq data

Protein	Mutant <i>med10-196</i>
Pol II	0.45
Med15	0.46
Med17	0.45
TFIIB	0.57
Toa2	0.97
Tfg1	1.2
Tfa2	0.66
Rad3	0.42
Kin28	0.48
TBP	1.93
Taf1	1.04

Supplemental Figures legends

Supplemental Figure S1. Coimmunoprecipitation between Mediator and TFIIA in *med10-196* mutant compared to the wild-type.

Wild-type and *med10* mutant strains carrying a Med5-HA tag and a Toa2-Myc tag were grown to exponential phase at 30°C or transferred to 37°C for 90 min. Mediator was immunoprecipitated (IP) through Med5-HA from crude extracts (Input) of wild-type and mutant strains using magnetic beads coupled to anti-HA antibodies. Control immunoprecipitation with IgG magnetic beads only was also performed (IgG). Coimmunoprecipitated Toa2-Myc was detected by western blotting using anti-Myc antibodies.

Supplemental Figure S2. Functional interaction between Mediator and TFIIIB.

(A) Specific *sua7* mutants have synthetic phenotypes in combination with *med10-196*.

The strain deleted for *med10* and complemented by a *TRP1* plasmid carrying *MED10* or *med10-196*, and also deleted for *sua7* and complemented by a *URA3* plasmid carrying *SUA7*, was transformed by the *HIS3* plasmids carrying either WT or mutated version of *SUA7*. Transformants were serially diluted, spotted on 5-FOA-containing agar plates to counterselect the WT *SUA7*-bearing plasmid (*see* Supplemental Materials and Methods), and incubated at 30°C for 3 days, or on YPD agar plates and incubated at 30°C for 3 days. The *sua7* mutants (*sua7-11* (*L136P*), *-34* (*L52P*) and *36* (*S53P*)) showing synthetic phenotypes with *med10-196* are indicated in red.

(B) Two-hybrid interaction between Med10 and Med14 Mediator subunits and Sua7. Wild-type or mutant Sua7 was fused to the Gal4 DNA-binding domain (G_{DB} -Sua7, Sua7- L136P), and Med10, 14, 21 were fused to the Gal4 activation domain (G_{AD} -Med10, 14, 21). Three independent clones for G_{DB} -Sua7 were included to allow evaluation of the reproducibility of the experiments.

(C) Quantitative analysis of two-hybrid interactions between Med10 and Med14 Mediator subunits and TFIIB. Wild-type or mutant Sua7 was fused to the Gal4 DNA-binding domain (G_{DB} -Sua7, Sua7-L136P), and Med10, 14, 21 were fused to the Gal4 activation domain (G_{AD} -Med10, 14, 21). A control corresponds to the empty vectors (G_{DB} , G_{AD}). β -galactosidase was assayed using CPRG as a substrate as described in Supplemental Materials and Methods. The values were normalized to the control. The mean values and standard deviation (indicated by error bars) of three independent experiments are shown. Asterisk represents a significant difference compared to the control for Med10-Sua7 and Med14-Sua7 at p-value <0.005 in a Student t-test. The background level was represented by a dotted line.

Supplemental Figure S3. Specific *sua7* mutants are lethal with *med14* mutants.

(A) Thermo-sensitive growth phenotype of the *med14* mutants. *med14-752*, *686* and *483* correspond to C-terminal truncations with a stop codon at position 752, 686 or 483, respectively. Cultures of WT and mutant *med14* yeast strains were serially diluted, spotted on YPD agar plates and incubated for 3 days at permissive (30°C) or non-permissive (37°C) temperatures.

(B) Specific *sua7* mutants are lethal with *med14* mutants.

The strain deleted for *med14* and complemented by a *TRP1* plasmid carrying *MED14* or *med14* mutants, and also deleted for *sua7* and complemented by a *URA3* plasmid carrying *SUA7*, was transformed by the *HIS3* plasmids carrying either WT or mutated version of *SUA7*. Transformants were serially diluted, spotted on 5-FOA-containing agar plates to counterselect the WT *SUA7*-bearing plasmid, and incubated at 30°C for 3 days, or on YPD agar plates and incubated at 30°C for 3 days.

Supplemental Figure S4. ChIP-seq profiles of preinitiation complex components on yeast genome.

Cells were grown at 30°C in YPD medium and then transferred for 90 minutes to 37°C.

(A) Example of ChIP-seq profiles of PIC components on chromosome VII from 397,758 to 400,300 for the wild-type and *med10-196* mutant. Densities of sequence tags were assessed from ChIP-seq experiments and displayed using the IGB yeast genome browser. Input DNA and DNA from ChIP with an untagged strain were used as negative controls. Densities of sequence tags were displayed after subtraction of the normalized control of an untagged strain.

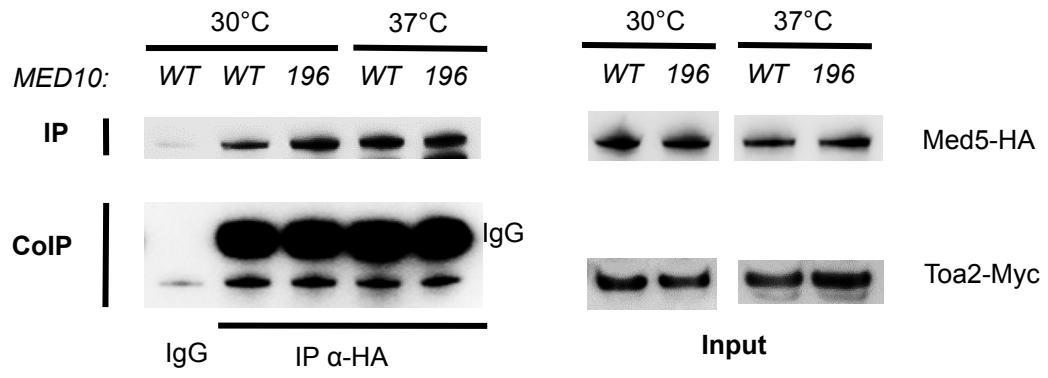
(B) Distribution of PIC components ChIP-seq densities around TSS. Pol II, Mediator (Med15 and Med17 subunits) and TFIIB profiles are displayed on the upper graph. TFIIA, B, D (TBP and Taf1), E, F, H and K profiles are shown on the lower graph. The tag density values for TFIIB were rescaled (x0.5) as indicated. Intergenic regions encompassing Pol III-transcribed genes and divergent genes were excluded. The tag density was determined for each protein in a 1600-bp window centered on the TSS. Mean tag density for each nucleotide position was then calculated and plotted over the window.

Supplemental Figure S5. Clustering analysis of genome-wide PIC occupancy ratios between *med10-196* and the wild type.

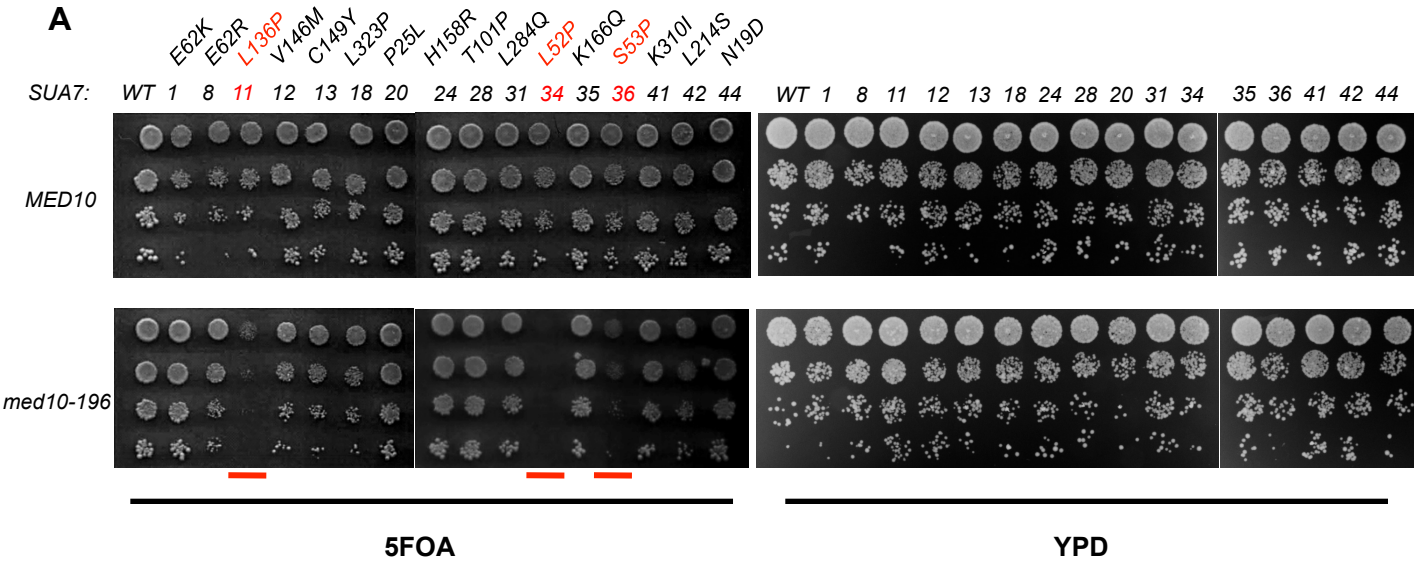
Clustering analysis was performed for TFIIH, TFIK, Med15 Mediator, TFIIA, TFIID (TBP), TFIID (Taf1), TFIIE and TFIIIF as described in **Figure 5** legend and Materials and Methods. The groups determined by indicated PIC component (left graph) were analyzed for the nucleosome occupancy in a 1600-bp window centered on the TSS (central graph) and for the presence of the TATA-box, for dynamic (hot) nucleosomes -1 and +1 (right graph).

Supplemental references

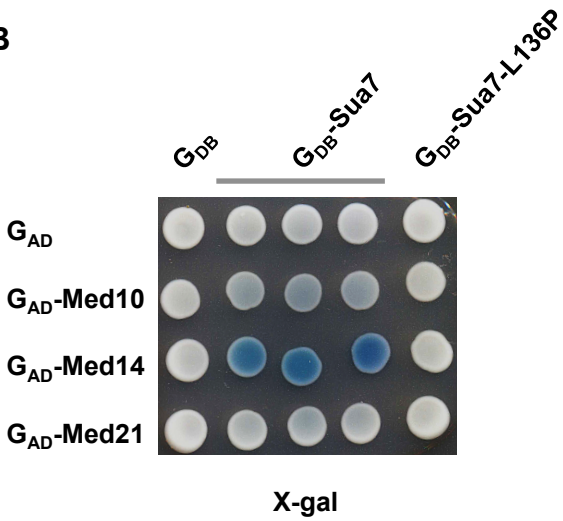
- Alberti S, Gitler AD, Lindquist S. 2007. A suite of Gateway cloning vectors for high-throughput genetic analysis in *Saccharomyces cerevisiae*. *Yeast* **24**: 913-919.
- Brachmann CB, Davies A, Cost GJ, Caputo E, Li J, Hieter P, Boeke JD. 1998. Designer deletion strains derived from *Saccharomyces cerevisiae* S288C: a useful set of strains and plasmids for PCR-mediated gene disruption and other applications. *Yeast* **14**: 115-132.
- Cai G, Imasaki T, Takagi Y, Asturias FJ. 2009. Mediator structural conservation and implications for the regulation mechanism. *Structure* **17**: 559-567.
- Eyboulet F, Wydau-Dematteis S, Eychenne T, Alibert O, Neil H, Boschiero C, Nevers MC, Volland H, Cornu D, Redeker V et al. 2015. Mediator independently orchestrates multiple steps of preinitiation complex assembly in vivo. *Nucleic Acids Res* **43**: 9214-9231.
- Janke C, Magiera MM, Rathfelder N, Taxis C, Reber S, Maekawa H, Moreno-Borchart A, Doenges G, Schwob E, Schiebel E et al. 2004. A versatile toolbox for PCR-based tagging of yeast genes: new fluorescent proteins, more markers and promoter substitution cassettes. *Yeast* **21**: 947-962.
- Longtine MS, McKenzie A, 3rd, Demarini DJ, Shah NG, Wach A, Brachat A, Philippsen P, Pringle JR. 1998. Additional modules for versatile and economical PCR-based gene deletion and modification in *Saccharomyces cerevisiae*. *Yeast* **14**: 953-961.
- Miller JH. 1972. *Experiments in Molecular Genetics*. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
- Ranish JA, Yudkovsky N, Hahn S. 1999. Intermediates in formation and activity of the RNA polymerase II preinitiation complex: holoenzyme recruitment and a postrecruitment role for the TATA box and TFIIB. *Genes Dev* **13**: 49-63.
- Reeves WM, 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.
- Sikorski RS, Hieter P. 1989. A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in *Saccharomyces cerevisiae*. *Genetics* **122**: 19-27.
- Soutourina J, Bordas-Le Floch V, Gendrel G, Flores A, Ducrot C, Dumay-Odelot H, Soularue P, Navarro F, Cairns BR, Lefebvre O et al. 2006. Rsc4 connects the chromatin remodeler RSC to RNA polymerases. *Mol Cell Biol* **26**: 4920-4933.
- Soutourina J, Wydau S, Ambroise Y, Boschiero C, Werner M. 2011. Direct interaction of RNA polymerase II and mediator required for transcription in vivo. *Science* **331**: 1451-1454.
- Van Mullem V, Wery M, De Bolle X, Vandenhoute J. 2003. Construction of a set of *Saccharomyces cerevisiae* vectors designed for recombinational cloning. *Yeast* **20**: 739-746.
- Werner M, Chaussivert N, Willis IM, Sentenac A. 1993. Interaction between a complex of RNA polymerase III subunits and the 70-kDa component of transcription factor IIIB. *J Biol Chem* **268**: 20721-20724.
- Wu WH, Pinto I, Chen BS, Hampsey M. 1999. Mutational analysis of yeast TFIIB. A functional relationship between Ssu72 and Sub1/Tsp1 defined by allele-specific interactions with TFIIB. *Genetics* **153**: 643-652.



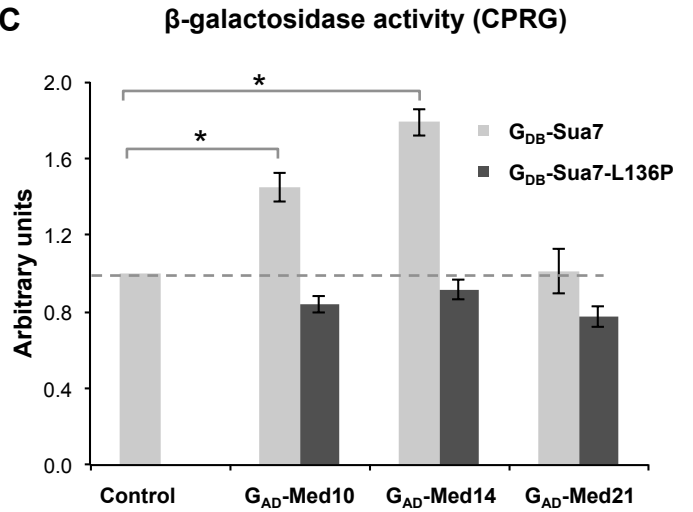
A



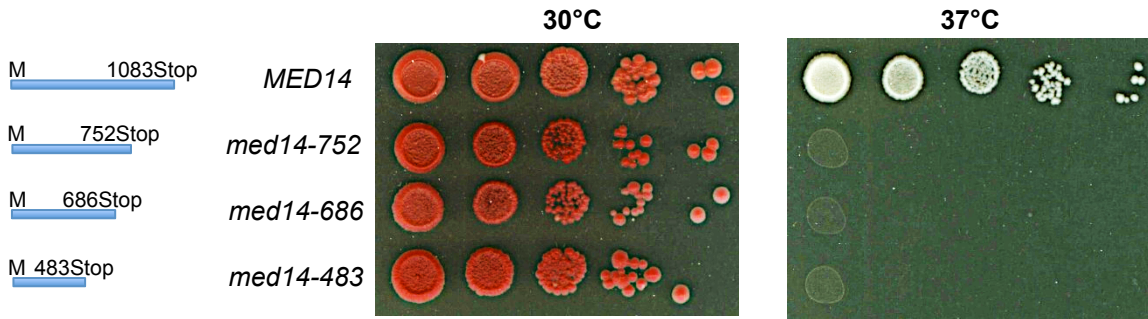
B



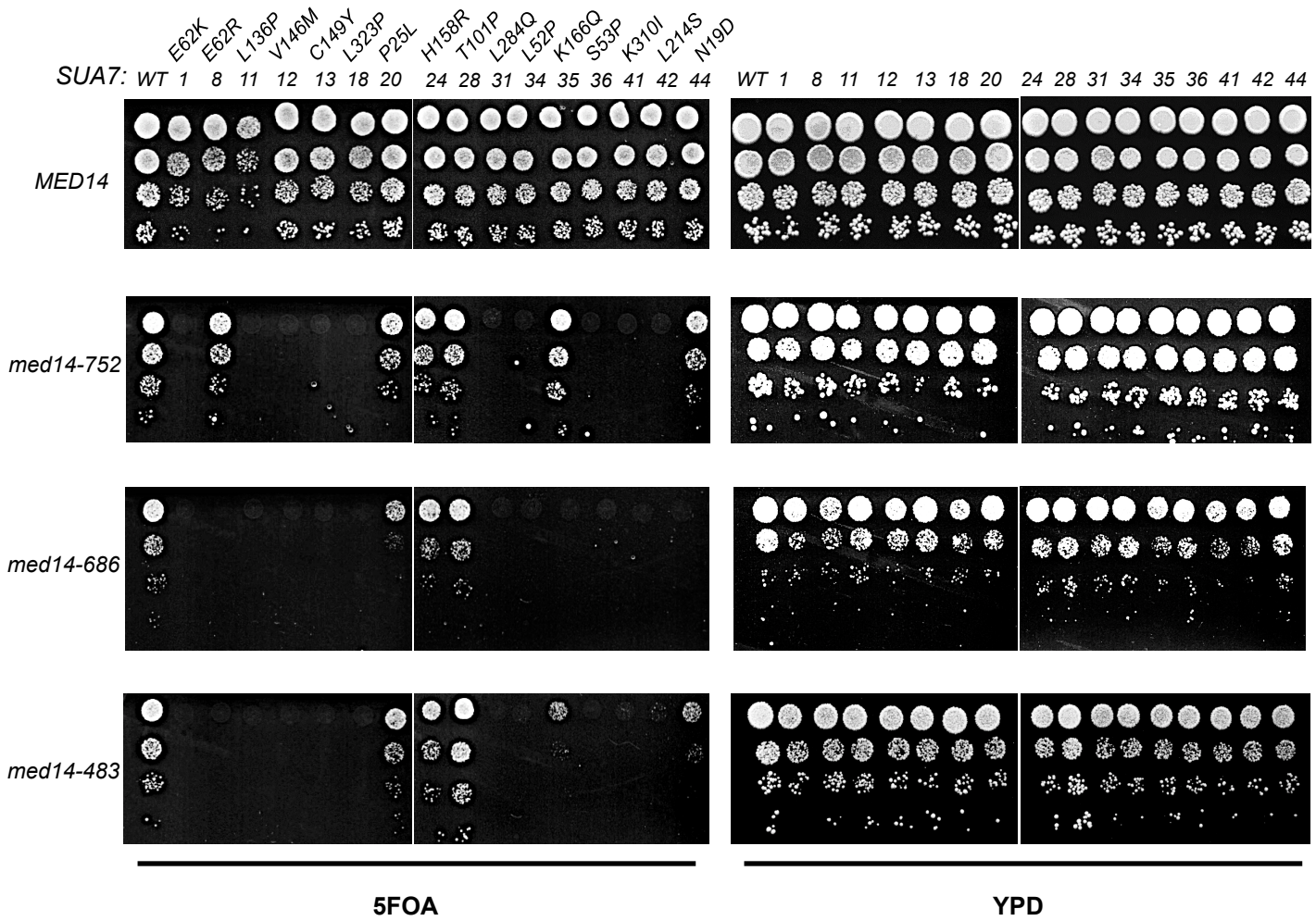
C



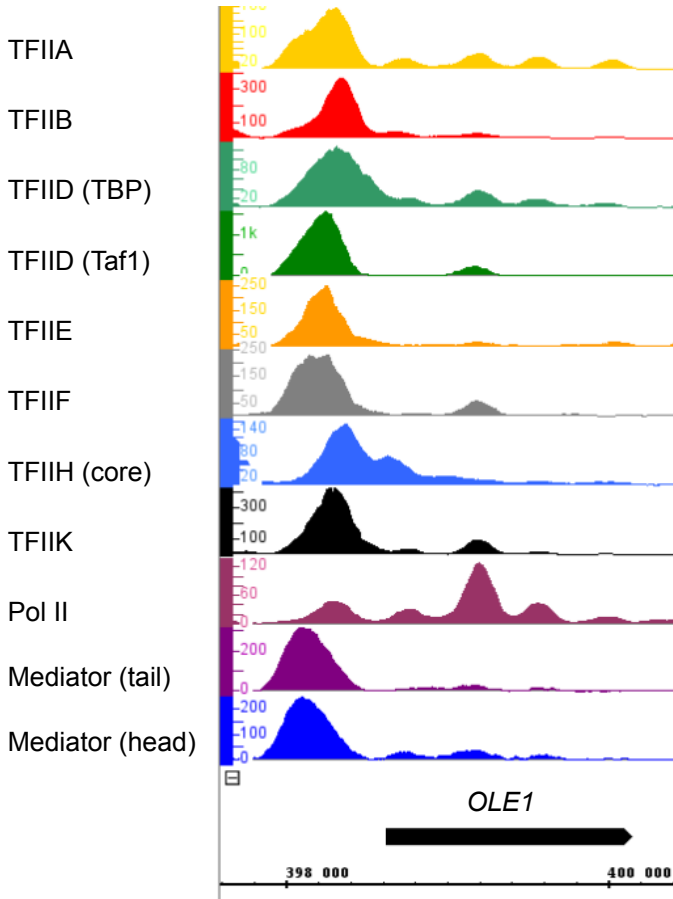
A



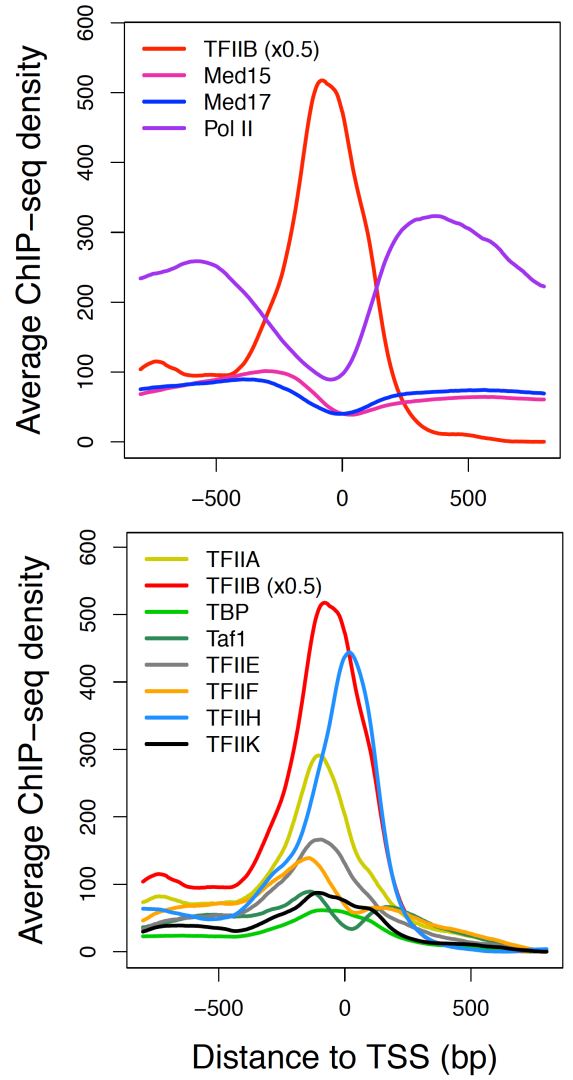
B

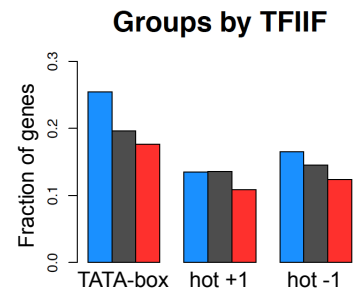
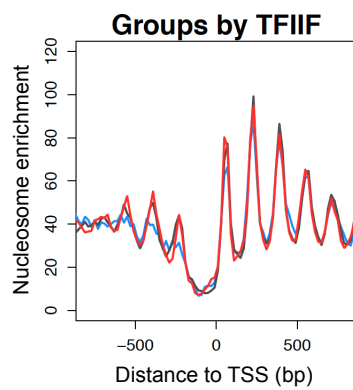
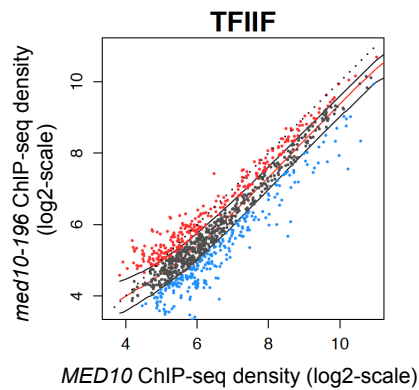
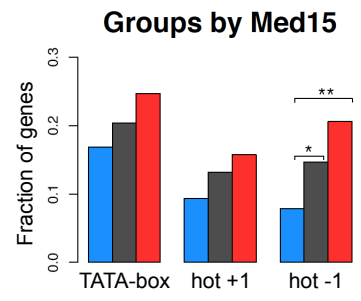
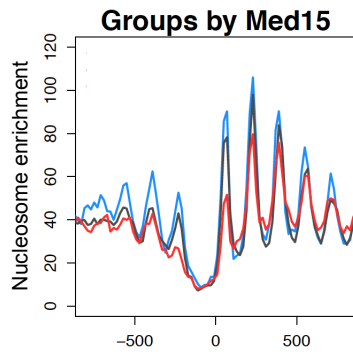
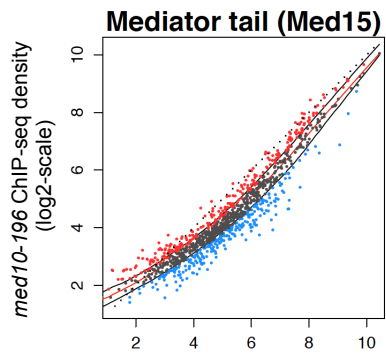
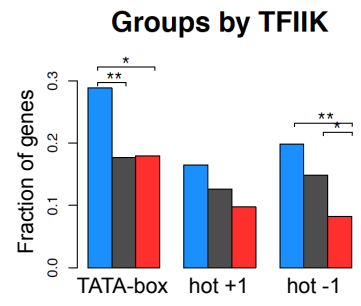
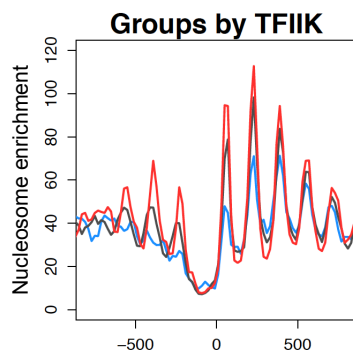
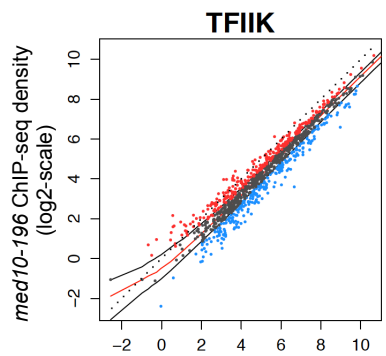
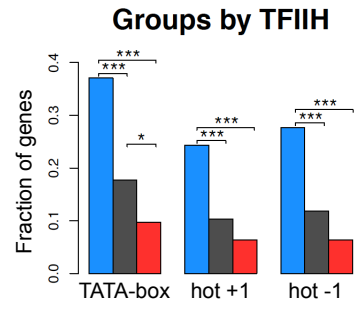
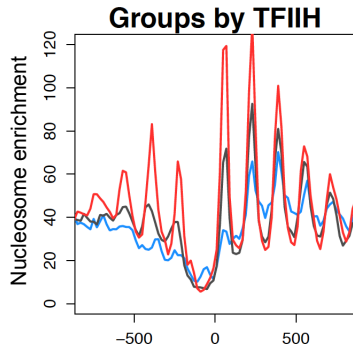
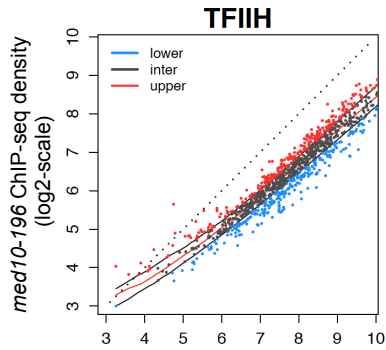


A

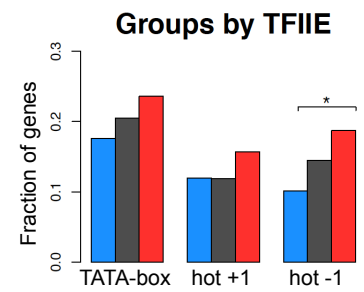
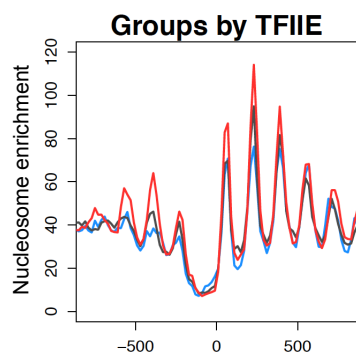
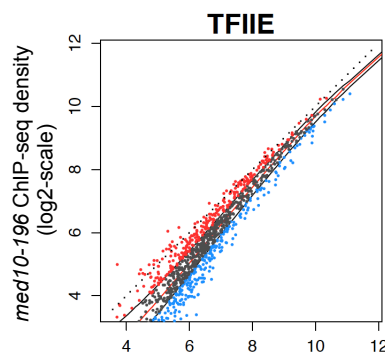
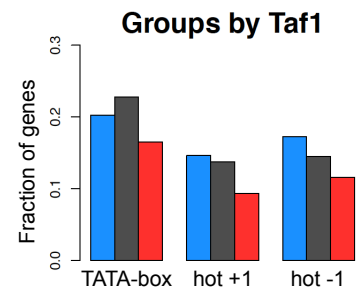
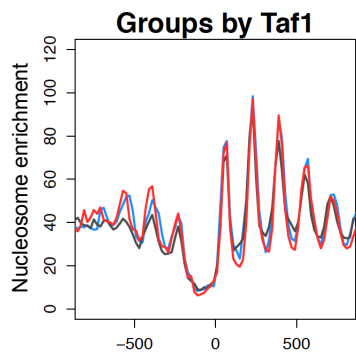
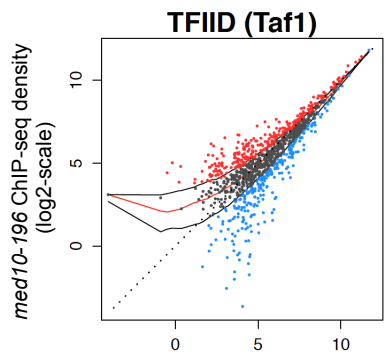
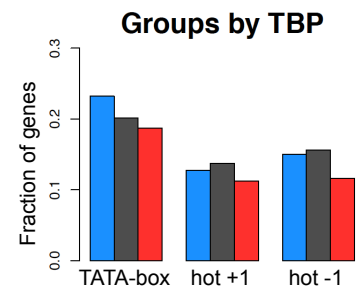
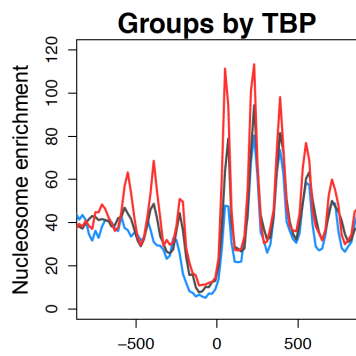
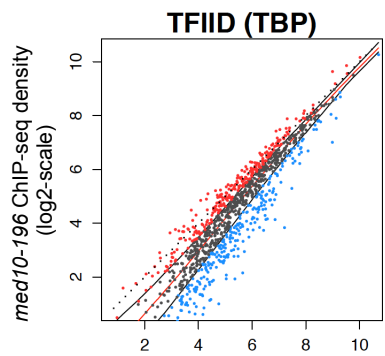
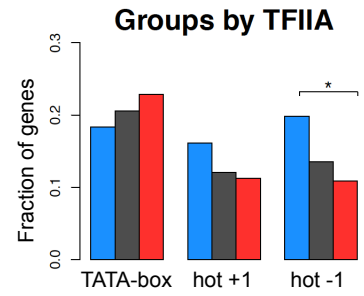
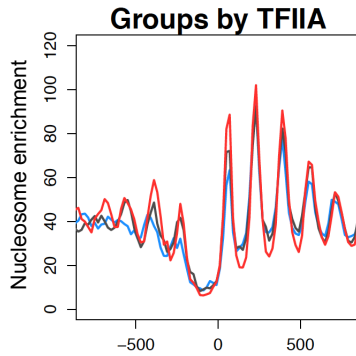
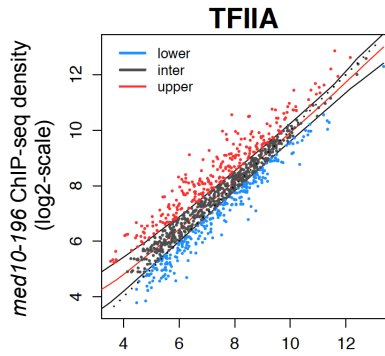


B





➔ (continued)



MED10 ChIP-seq density (log2-scale)

Distance to TSS (bp)

Partie II : Etude du rôle de la sous-unité essentielle Med7.

I. Contexte de travail.

Dans le cadre de la démarche entreprise par notre laboratoire pour caractériser les sous-unités essentielles du Médiateur, et en particulier de celles du module du milieu, nous avons initié l'étude de la sous-unité Med7. Une autre partie de mon projet de thèse a consisté en la caractérisation de mutants thermosensibles de Med7, selon les mêmes approches que pour l'étude de Med10 expliqué plus tôt. Nous avons d'abord étudié les interactions entre Med7 et ses partenaires au sein du Médiateur par double-hybride, vérifié l'intégrité du complexe puis analysé l'effet des mutations sur l'assemblage du PIC. A mon arrivée au laboratoire, une collection de 15 mutants thermosensibles de Med7 avait été obtenue et ces expériences ont été initiées pour certains d'entre eux.

II. Résultats.

1) Collection de mutants thermosensibles de Med7.

Afin d'obtenir des mutants thermosensibles de la sous-unité Med7, nous avons mis en œuvre la méthode de PCR mutagène aléatoire utilisée par notre laboratoire dans nos travaux précédents (Esnault et al. 2008; Soutourina et al. 2011; Eyboulet et al. 2015). Le criblage des clones obtenus nous a permis d'isoler initialement 11 mutants, portant tous plusieurs résidus mutés. L'analyse des séquences des différents mutants nous a permis d'observer que certains résidus mutés étaient présents dans plusieurs mutants (D63G, L126P, L133P et S173P). Nous avons décidé d'obtenir par mutagenèse dirigée les mutants ponctuels affectant ces positions. De manière intéressante, ces mutants présentent un phénotype thermosensible. Cela nous a donc permis d'obtenir une collection de 15 mutants thermosensibles (Figure 30). Les positions des résidus mutés pour ces mutants sont indiqués dans le tableau 13.

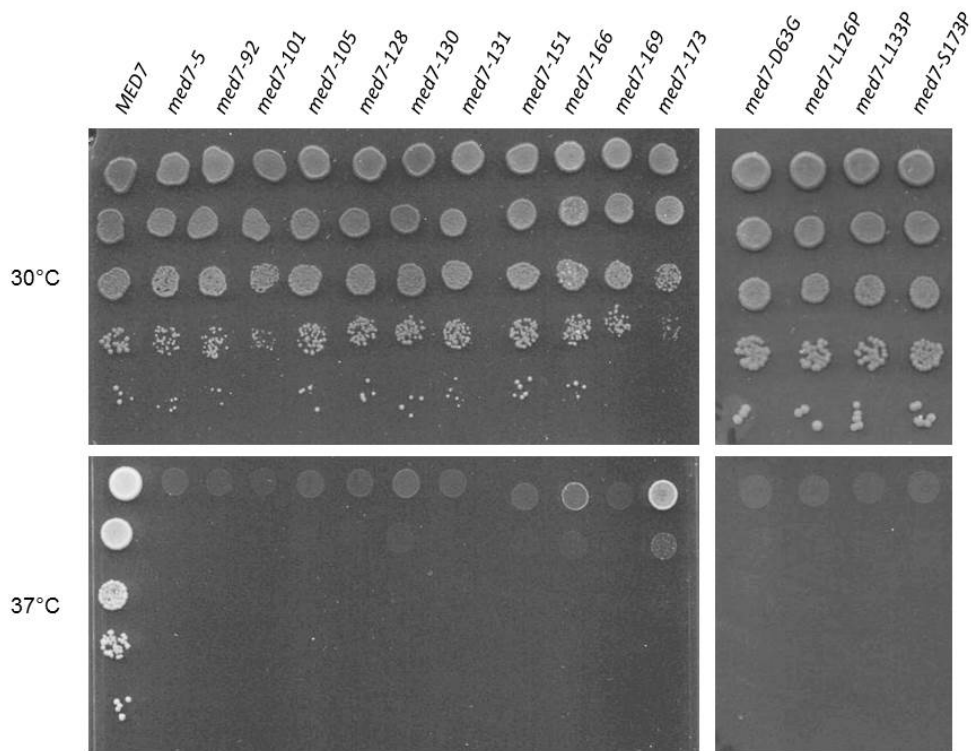


Figure 30: Phénotype thermosensible des mutants de Med7 obtenus par PCR mutagène aléatoire (à gauche) et par mutagenèse dirigée (à droite).

<i>Mutant</i>	<i>Résidus mutés</i>
<i>med7-5</i>	K38R, Q44R, L131S, S173P
<i>med7-92</i>	E164G, S173P, L203H
<i>med7-101</i>	N7I, Y34H, Q75L, T98A, K103E, K174N, H160L, I192V, K197N, L203P
<i>med7-105</i>	N27Y, D63G, L133P
<i>med7-128</i>	N140S, T153A, L147P, E179G
<i>med7-130</i>	N111D, L133P, K202E, Q215R
<i>med7-131</i>	D63G, L123S, L133P, M145V
<i>med7-151</i>	L126P, E191A
<i>med7-166</i>	L127Q, Q169L, L147P
<i>med7-169</i>	L126P, R152G, Q181R
<i>med7-173</i>	E8D, P90S
<i>med7-D63G</i>	D63G
<i>med7-L126P</i>	L126P
<i>med7-L133P</i>	L133P
<i>med7-S173P</i>	S173P

Tableau 13: Tableau récapitulatif des résidus mutés dans les différents mutants de Med7

La structure tridimensionnelle partielle (résidus 102 à 205) de Med7 a été obtenue par cristallographie (Baumli, Hoepfner, and Cramer 2005). Nous avons situé certains résidus mutés compris dans ce fragment de Med7 sur cette structure (Figure 31). On observe qu'ils sont répartis tout au long de la structure de la protéine, organisée en trois hélices α . Un grand nombre des mutations retrouvées dans nos mutants correspondent à des substitutions en proline, notamment trois des mutations ponctuelles isolées (L126P, L133P et S173P). La présence de résidus proline dans les chaînes d'acides aminés induit des courbures. En particulier, on sait que ces prolines peuvent déstabiliser les structures en hélices. Il est possible que ces mutations puissent déstabiliser la structure de la protéine et ainsi expliquer le phénotype ts. Cependant, la répartition des mutations tout au long de la structure est propice à modifier des contacts éventuels avec d'autres partenaires, d'autant plus que les résidus semblent être exposés à la surface de la protéine. La perte éventuelle de contacts pourrait elle aussi être responsable des phénotypes ts.

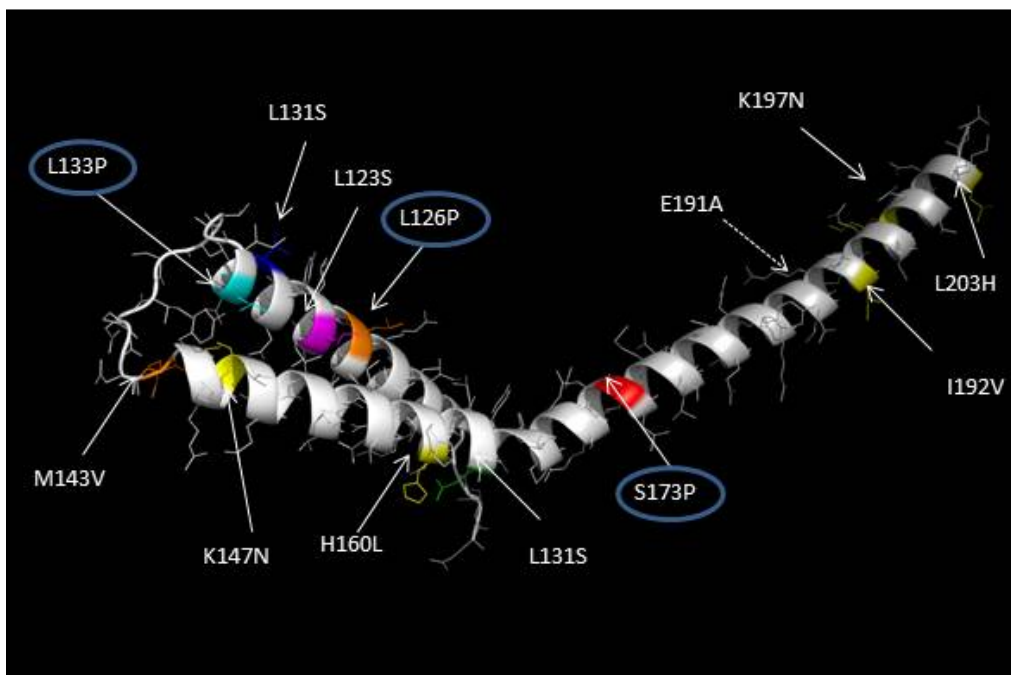


Figure 31: Structure de la sous-unité Med7 obtenue par cristallographie et diffraction des rayons X. (Baumli, Hoepfner, and Cramer 2005)

La structure représentée ici correspond aux résidus 102 à 205 de Med7. Les résidus mutés dans les mutants de notre collection sont indiqués par des flèches lorsque cela est possible. Les positions mutées dans trois des quatre mutants ponctuels isolés sont entourées en bleu.

Certains mutants de Med7 que nous avons obtenus présentent un phénotype supplémentaire à la thermosensibilité. En effet, ils présentent un phénotype de floculation (Figure 32). Ce phénotype est en temps normal un phénotype que l'on retrouve dans les souches de levure sauvages d'isolats naturels. Les modèles de levures de laboratoire, en particulier la souche S288c dont dérivent plusieurs contextes génétiques, possèdent une mutation dans le gène *FLO8*, abolissant ce phénotype. Ces souches ont été sélectionnées afin de faciliter les manipulations en laboratoire. Le phénotype de floculation implique un grand nombre de gènes concernant l'adhésion cellulaire. Certaines mutations de Med7 semblent donc avoir un effet sur l'expression de ces gènes. Il est intéressant de noter que ce phénotype est déjà visible à la température permissive de 30°C.

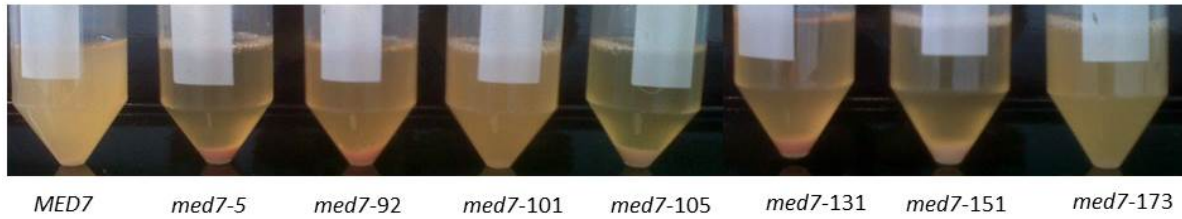


Figure 32: Phénotype de floculation de certains mutants de Med7.

Les cultures ont été menées en fin de phase exponentielle de croissance en milieu riche (YPD) à la température permissive de 30°C. La floculation est observée après avoir laissé les cultures au repos pendant 10 minutes.

2) Interactions entre Med7 et ses partenaires au sein du Médiateur

Nous avons observé l'effet des mutations de Med7 sur ses interactions au sein du Médiateur. Pour cela, nous avons analysé par double-hybride les interactions entre les versions mutées de Med7 et trois de ses partenaires : Med9, Med10 et Med21 (Figure 33).

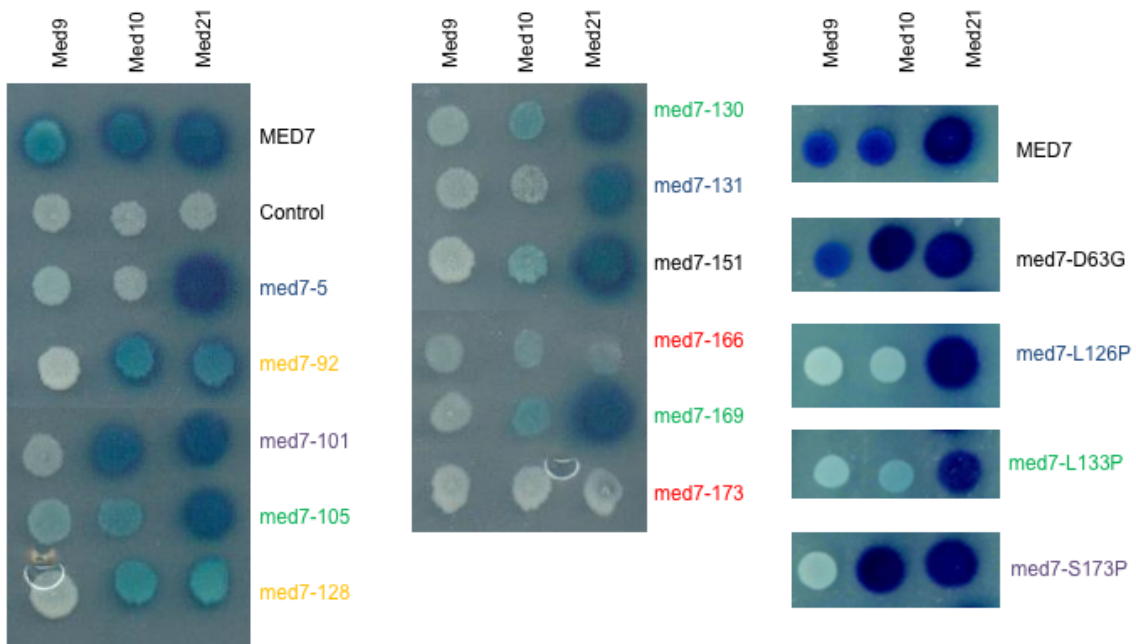


Figure 33: Analyse par double-hybride des interactions entre les versions mutées de Med7 et les sous-unités Med9, Med10 et Med21. Les couleurs indiquent des mutants dont le profil d'interaction est similaire.

Cette analyse montre que les différentes mutations n'ont pas le même effet sur les interactions entre Med7 et ses partenaires. En effet, on remarque que dans le cas des mutants *med7-5*, *med7-131* et *med7-L126P* (groupe bleu), seule l'interaction avec Med21 est préservée, les interactions

avec Med9 et Med10 étant abolies. Concernant les mutants *med7-92* et *med7-128* (groupe jaune), on observe que l'interaction avec Med9 est perdue et que celles avec Med10 et Med21 est légèrement diminuée. Les mutants *med7-105*, *med7-130*, *med7-151*, *med7-169* et *med7-L133P* (groupe vert) présentent une perte d'interaction avec Med9, une diminution de l'interaction avec Med10 et une interaction inchangée avec Med21. Les mutants *med7-173* et *med7-166* (groupe rouge) montrent une perte des interactions avec les trois sous-unités testées. Dans le cas des mutants *med7-101* et *med7-S173P* (groupe violet), seule l'interaction avec Med9 est perdue. Enfin, le mutant *med7-D63G* présente un profil assez similaire à celui de *MED7-WT*. Les différences au niveau des profils d'interaction peuvent probablement être expliquées par le fait que les résidus mutés dans les différents mutants sont répartis le long de la protéine, perturbant ainsi probablement de façon différentielle les contacts de Med7 avec ses partenaires.

Nous avons, sur cette base, sélectionné plusieurs mutants que nous souhaitons analyser plus en détail. Pour cela, nous avons choisi au moins un mutant de chaque groupe, et privilégié les mutants dont le nombre de mutations est le plus faible. Nous avons ainsi choisi les mutants *med7-5*, *med7-101*, *med7-131*, *med7-173* ainsi que les mutants *med7-L126P*, *med7-L133P* et *med7-S173P*.

3) Analyse de l'intégrité du Médiateur dans les contextes mutants.

Dans notre démarche d'étude des mutants ts du Médiateur, nous souhaitons étudier des mutants n'affectant pas la stabilité ou l'assemblage du Médiateur. Nous avons donc analysé son intégrité dans les contextes mutants par co-immunoprécipitation. Pour cela, nous avons immunoprécipité la sous-unité Med5 (module de queue) portant une étiquette HA et vérifié que les sous-unités Med14 (module de queue/milieu) et Med17 (module de tête) co-immunoprécipitent. Cela nous permet ainsi de vérifier la présence d'au moins une sous-unité de chacun des modules principaux du Médiateur. On peut ainsi observer que dans les mutants que nous avons sélectionné, l'intégrité du complexe ne semble pas affectée par les mutations de Med7 (Figure 34).

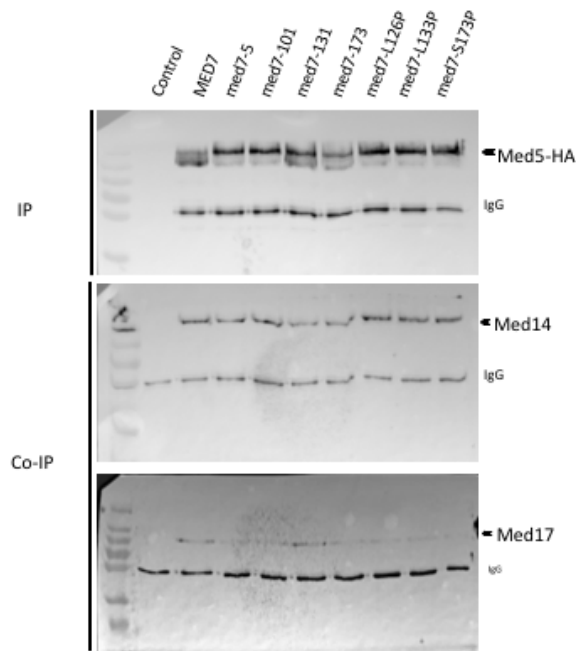


Figure 34: Analyse par co-immunoprécipitation de l'intégrité du Médiateur dans le cas des mutants sélectionnés pour nos études.

Détection par western blot de la sous-unité Med5-HA utilisée pour l'immunoprécipitation, et de Med17 et Med14. Le contrôle consiste en une immunoprécipitation en présence d'un anticorps non spécifique de l'étiquette HA. La bande IgG correspond à la chaîne lourde de l'anticorps utilisé pour l'IP.

Néanmoins, nous avons constaté que les mutants ponctuels *med7-L162P*, *L133P* et *S173P* présentaient un taux considérable de réversion/suppression de leur phénotype ts. Cette propriété représente un problème puisque nous ne pouvons pas être certains de travailler sur une population homogène de cellules et introduit une grande source de variabilité. Nous avons donc décidé de ne pas poursuivre l'étude de ces mutants.

4) Effet des mutations de Med7 sur l'occupation des composants du PIC.

L'intégrité du Médiateur dans les mutants de Med7 étant vérifiée, nous avons initié la caractérisation fonctionnelle de ces mutants. Pour cela nous avons analysé par CHIP l'occupation de la Pol II, du Médiateur, de TFIID cœur (Rad3), TFIID kinase (Kin28) et de la TBP sur trois gènes modèles (*ADH1*, *PYK1*, *PMA1*). Le gène *GAL1* étant réprimé dans une culture en présence de glucose, il a été utilisé comme contrôle négatif. S'agissant de mutants thermonsensibles, pour induire le phénotype, il faut placer les cellules (*shift*) à la température non permissive (37°C). Pour cela, nous avons réalisé une première expérience de CHIP après un *shift* à 37°C de 90 minutes. Nous n'avons pas observé d'effet sur l'occupation de la Pol II. Nous avons alors réalisé un suivi de la croissance en culture liquide des différents mutants à 37°C. Cela nous a permis d'observer que la croissance des mutants n'était affectée qu'après environ 8 heures de *shift*. Nous avons ainsi retenu ce temps de culture à 37°C avant de procéder à nos expériences de CHIP.

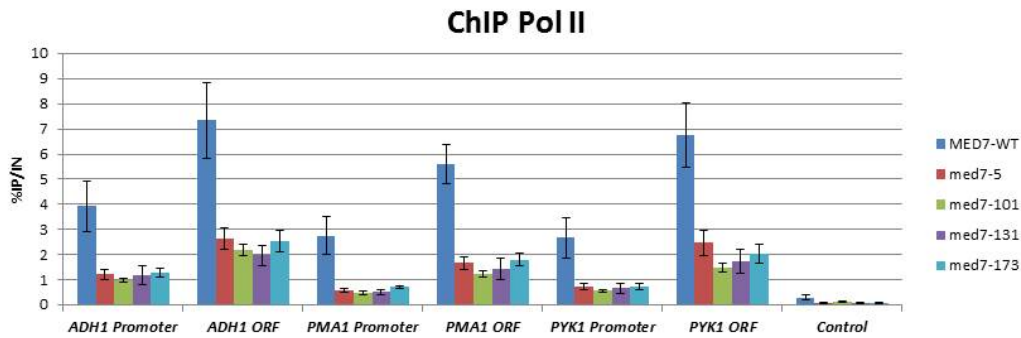


Figure 35: Niveau d'occupation de la Pol II sur les promoteurs et ORFs des gènes *ADH1*, *PYK1* et *PMA1* mesuré par ChIP pour les mutants *med7-5*, *101*, *131*, *173*.

Les ChIP sont réalisés sur trois cultures indépendantes pour chaque souche, après un passage de 8h à 37°C.

Nous avons ainsi pu observer que pour tous les mutants sélectionnés, le niveau d'occupation de la Pol II est diminué d'environ 50%, aussi bien sur les promoteurs que sur les ORFs (Figure 35). Cela montre que les mutations affectent l'activité transcriptionnelle, du moins sur ces trois gènes modèles.

Lorsque l'on s'intéresse à TFIIH, on observe que l'occupation de son module cœur (Rad3) est diminuée sur les trois gènes étudiés et ce pour l'ensemble des mutants (Figure 36A). On observe que tous les mutants présentent un effet du même ordre de grandeur à savoir une diminution d'un facteur trois environ. L'occupation du module kinase (Kin28) est elle aussi fortement affectée (Figure 36B). Pour les mutants *med7-5*, *101*, *131* et *173*, on observe une diminution d'occupation d'environ 4 fois sur les gènes *PMA1* et *PYK1* et d'environ 5 fois sur le gène *ADH1*. Ces résultats montrent que les mutations ont un effet différent sur les deux modules de TFIIH. Ceci confirme de nouveau des observations que nous avons faites dans l'étude des mutants de Med11, Med17 (Esnault et al. 2008; Eyboulet et al. 2015) et dans notre étude de Med10, à savoir que le Médiateur peut influencer de façon indépendante l'occupation des deux modules de TFIIH. Cependant, l'effet des mutations de Med7 est inverse de celle qui touchent Med11.

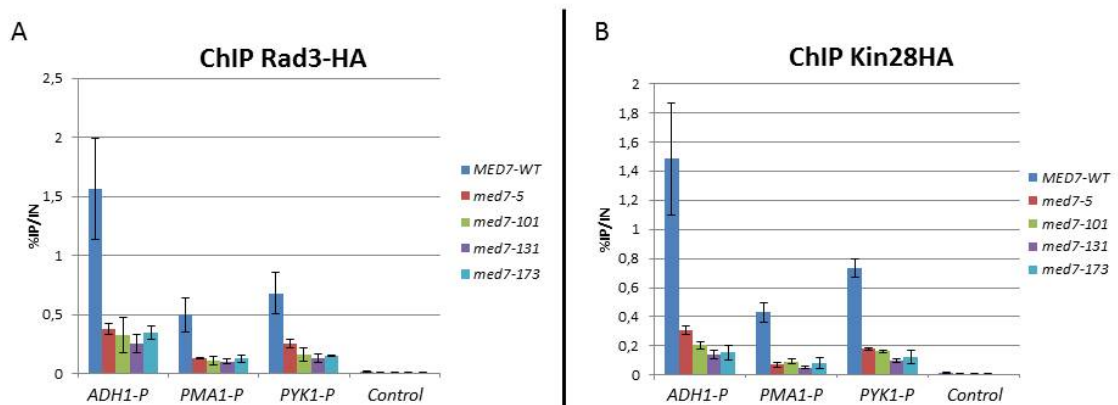


Figure 36: Niveau d'occupation de TFIIH sur les promoteurs des gènes *ADH1*, *PYK1* et *PMA1* mesuré par ChIP pour les mutants *med7-5*, *101*, *131*, *173*.

Les ChIP sont réalisés sur trois cultures indépendantes pour chaque souche, après un passage de 8h à 37°C. L'occupation de TFIIH est observée pour ses deux modules. A: Module cœur (Rad3) B: Module Kinase (Kin28).

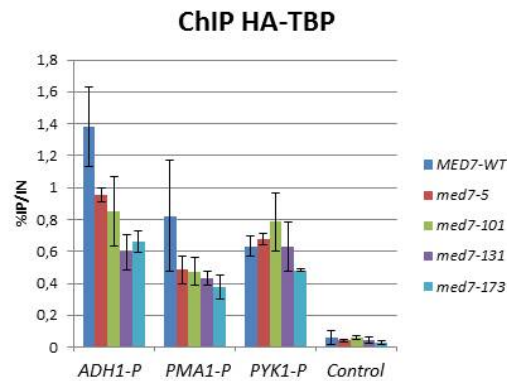


Figure 37: Niveau d'occupation de TFIID (TBP) sur les promoteurs des gènes *ADH1*, *PYK1* et *PMA1* mesuré par ChIP pour les mutants *med7-5*, *101*, *131*, *173*.

Les ChIP sont réalisés sur trois cultures indépendantes pour chaque souche, après un passage de 8h à 37°C.

Nous souhaitons analyser également l'occupation de TFIID, par l'étude de la sous-unité TBP (Figure 37). On peut constater que pour les 4 mutants, les mutations ont un effet différentiel selon le gène analysé. En effet, sur les promoteurs des gènes *ADH1* et *PMA1*, on observe une légère diminution de l'occupation de la TBP. Sur le promoteur de *PYK1*, il semble que l'occupation de la TBP ne soit pas affectée, excepté pour le mutant *med7-173* où l'on peut entrevoir une légère diminution. On observe que pour ces trois mutants, la tendance est à la diminution d'occupation de la TBP sur les gènes *PMA1* et surtout *ADH1*. Les effets sont par contre moindres que pour la Pol II et TFIIF.

La TBP se comporte donc de la même manière que dans le mutant *med10-196*, à savoir une tendance à la diminution. Cependant, lorsque nous avons analysé sa distribution génomique, nous avons constaté que globalement, son niveau d'occupation ne variait pas. Il sera donc très intéressant d'analyser l'occupation de la TBP au niveau du génome dans nos mutants de Med7.

Enfin, nous avons analysé l'occupation du Médiateur sur les régions promotrices des trois gènes modèles *ADH1*, *PYK1* et *PMA1*. Nous avons mesuré l'occupation du module de queue (Med15) et du module de tête (Med17) (Figure38).

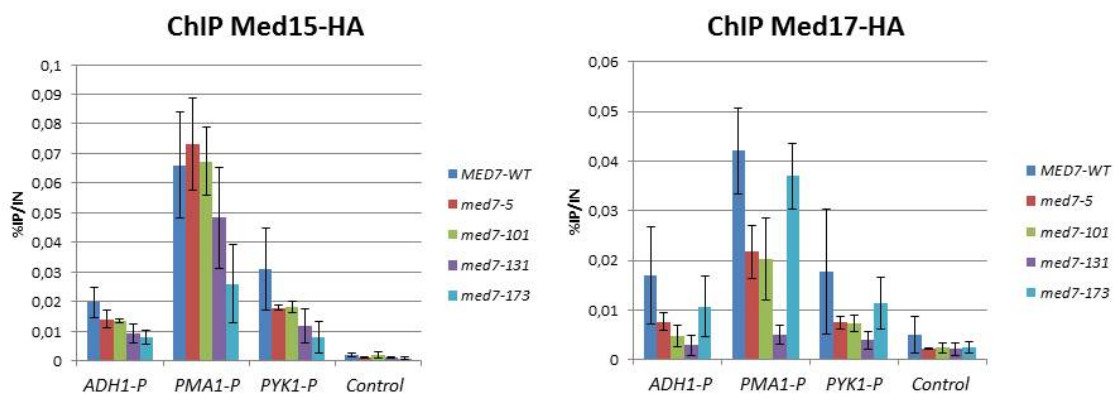


Figure 38 : Niveau d'occupation des modules de queue (Med15) et de tête (Med17) du Médiateur sur les promoteurs des gènes *ADH1*, *PYK1* et *PMA1* mesuré par ChIP pour les mutants *med7-5*, *101*, *131* et *173*.

Les ChIP sont réalisés sur trois cultures indépendantes pour chaque souche, après un passage de 8h à 37°C.

L'analyse par CHIP de l'occupation du Médiateur révèle que l'occupation du module de queue ne semble pas ou du moins très peu affectée pour l'ensemble des mutants. Dans le cas du mutant *med7-173*, il semble néanmoins qu'il y ait une légère diminution de l'occupation de Med15. Lorsque l'on s'intéresse au module de tête, on constate un comportement différent. On observe, en particulier, que dans le cas des mutants *med7-5*, *101* et *131*, l'occupation de Med17 est diminuée sur les promoteurs des trois gènes étudiés, et ce de façon très nette sur le gène *PMA1*. Le mutant *med7-131* est très affecté pour l'occupation de Med17. De façon inverse à ce que nous observons pour Med15, l'occupation de Med17 n'est pas affectée dans le mutant *med7-173*. Ces résultats semblent indiquer que dans le cas des mutations *med7-5*, *101* et *131*, les modules de tête et de queue du Médiateur se comportent de façon indépendante. Il est également possible que les mutations destabilisent l'organisation des modules de façon différentielle.

Ces résultats montrent que nous avons isolé des mutants de Med7 qui ont un effet sur l'assemblage du PIC. Afin de savoir si ces effets sont généralisés à l'ensemble du génome, nous souhaitons étendre l'analyse en réalisant des expériences de CHIP-seq. Initialement, nous avons décidé de faire ces expériences sur les mutants *med7-5*, *med7-L126P*, *med7-L133P* et *med7-S173P*. En effet, nous voulions privilégier des mutants dont le nombre de mutations était le plus faible possible. Nous avons conservé le mutant *med7-5* car celui-ci est l'un des mutants dont le phénotype de floculation est le plus marqué, et donc intéressant pour l'investigation des causes de ce phénotype. Nous avons exclu les mutants ponctuels *med7-L126P*, *med7-L133P* et *med7-S173P* de nos analyses en raison de leur fort taux de reversion.

A ce jour, nous avons réalisé les expériences de CHIP-seq pour le mutant *med7-5*. Nous disposons des données pour la Pol II, TFIIH (Rad3 et Kin28), la TBP et le Médiateur (Med15 et Med17). Ces données sont en cours d'analyse.

Nous avons défini en parallèle deux autres mutants que nous souhaiterions ajouter à nos analyses. Il s'agit des mutants *med7-105* et *med7-151*. La sélection des mutants *med7-5*, *med7-105* et *med7-151* nous permet d'étudier des mutants où l'on retrouve les mutations L126P, L133P et S173P, associées à d'autres mutations. Après avoir étudié l'intégrité du Médiateur pour les mutants *med7-105* et *med7-151*, nous réaliserons les expériences de CHIP et de CHIP-seq.

Conclusions et perspectives

L'étude réalisée au cours de ma thèse nous a permis de caractériser la sous-unité Med10 en utilisant des approches *in vitro*, *in vivo* et *in silico*. Nous avons ainsi mis en évidence un lien fonctionnel *in vivo* entre le Médiateur et TFIIB. Nous avons analysé ce lien et montré une interaction jusqu'alors inconnue entre la sous-unité Med14 et TFIIB. Nous avons également montré que le module du milieu, et plus particulièrement la sous-unité Med10, présente un rôle crucial pour l'assemblage du PIC et ce de façon globale. L'analyse de données de CHIP-seq que nous avons obtenues nous a permis de révéler que l'effet du Médiateur sur l'assemblage du PIC était variable selon l'architecture des promoteurs des gènes considérés.

D'autre part, le travail initié sur l'étude de la sous-unité Med7 n'en est qu'à ses débuts. Néanmoins, nous avons déjà pu observer des résultats intéressants et cela nous laisse déjà entrevoir des perspectives pour la suite.

I. Collection de mutants thermosensibles de deux sous-unités du module du milieu du Médiateur.

La mutagenèse aléatoire effectuée sur le gène *MED10* nous a permis d'obtenir un seul mutant. Nous avons essayé de séparer les différents résidus mutés parmi les 4 substitutions observées dans le mutant *med10-196*. Les mutants ponctuels obtenus n'ont pas montrés de phénotypes thermosensibles. Nous avons également essayé d'obtenir les différentes combinaisons de double-mutations. Là encore, aucun phénotype ts n'a été observé. Il apparaît que c'est la combinaison de ces 4 mutations qui provoque ce phénotype. A l'heure actuelle, on ne connaît pas la structure tridimensionnelle de Med10, en tout les cas pas à l'échelle atomique. Il est donc impossible d'observer quels pourraient être les effets des mutations sur la structure. Cependant, les résidus mutés dans *med10-196* sont parmi les plus conservés de Med10. Il est raisonnable d'imaginer qu'ils puissent être impliqués dans l'établissement d'une structure qui s'en trouve destabilisée par ces mutations. Les quatre résidus mutés dans notre mutant pourraient être des résidus clés dans la structure de la sous-unité.

Notre collection de mutants de la sous-unité Med7 comporte 15 mutants thermosensibles. Comme évoqué plus tôt, les mutants *med7-D63G*, *med7-L126P*, *med7-L133P* et *med7-S173P* présentent un phénotype ts à 37°C mais montrent un fort taux de réversion/suppression. La perte de ce phénotype ts signifie l'apparition de clones révertants. Cette capacité peut provenir de l'apparition de mutations de gènes suppresseurs des mutations de Med7. Ces suppresseurs peuvent être intragéniques, c'est-à-dire une mutation aléatoire de Med7 supprimant le phénotype. S'agissant de mutants-ponctuels, il peut s'agir tout simplement d'une réversion de la mutation. Les suppresseurs extragéniques sont, eux, des mutations d'un autre gène que *MED7* qui abolissent le phénotype ts. Dans le cas où la suppression serait due à l'apparition de mutations extragéniques, il serait intéressant de connaître les gènes concernés. Cela pourrait permettre de mettre en évidence des liens fonctionnels avec d'autres protéines.

II. Analyse de l'intégrité du Médiateur dans les mutants de ses sous-unités.

Durant ce travail de thèse, nous avons analysé la composition en terme de sous-unités du Médiateur dans les contextes sauvages et mutants. J'ai pu mettre au point une approche de purification du Médiateur suivie de son analyse par spectrométrie de masse. Cette méthode a permis d'améliorer notre résolution. En effet, dans notre approche systématique d'étude de mutants du Médiateur (Eyboulet et al. 2015; Esnault et al. 2008; Soutourina et al. 2011), l'une des conditions pour investiguer ces mutants est que ceux-ci n'affectent que la fonction du complexe, sans altérer l'assemblage ou la stabilité. Jusqu'alors, nous analysions l'intégrité du complexe par des méthodes de co-immunoprécipitation. Nous étions limités à l'observation de la présence de seulement quelques sous-unités, du fait qu'il existe peu d'anticorps efficaces pour la détection des sous-unités du Médiateur. Il est possible de construire des souches portant des étiquettes pour visualiser les différentes sous-unités mais la très grande taille du complexe (25 sous-unités) rendrait ce travail fastidieux. La nouvelle méthode nous permet désormais d'observer la présence de 20 sous-unités du Médiateur cœur (sans le module kinase) en une seule expérience. Seule la sous-unité Med31 n'est pas détectée dans nos conditions. Cela est dû à sa faible masse moléculaire, mais également sa composition en acides aminés peu propice à sa fragmentation, rendant sa détection difficile par spectrométrie de masse.

Néanmoins, en tant que telle, notre approche ne permet que de visualiser la présence ou l'absence d'une sous-unité, elle ne renseigne pas sur la stoechiométrie du complexe. Nous souhaiterions ainsi continuer le développement de cette méthode pour la rendre quantitative. Nous essayons de mettre au point des conditions d'analyse par spectrométrie de masse pour quantifier l'abondance des différentes sous-unités. La méthode repose sur l'analyse de plusieurs réplicats biologiques à l'aide d'outils bioinformatiques de quantification sans marquages tel qu'il est possible de le faire avec le logiciel MaxQuant (Cox and Mann 2008). Nous pourrions également mettre au point une méthode d'analyse par spectrométrie de masse quantitative par marquage isotopique comme avec la méthode de SILAC (*stable isotope labeling with amino acids in cell culture*) (Ong, Foster, and Mann 2003) qui pourrait être appliquée à nos travaux.

La mise au point de ces approches nous permettra par la suite d'analyser de façon exhaustive et quantitative l'intégrité du Médiateur issu de souches mutantes pour la sous-unité Med7 mais aussi dans les mutants d'autres sous-unités essentielles dans nos projets futurs.

III. Interaction entre le Médiateur et TFIIB.

Notre étude du mutant de Med10 nous a permis de découvrir un nouveau contact entre le Médiateur et TFIIB, par le biais de la sous-unité Med14. Nous avons pu montrer que ce contact est fonctionnel. En effet, la combinaison de mutations de Med10 et Med14 avec des mutations de TFIIB présente des phénotypes synthétiques létaux. De plus, ce contact est diminué dans le mutant *med10-196* ce qui conduit à des conséquences fortes sur l'assemblage du PIC. Ce lien entre le Médiateur était déjà suggéré dans la littérature (Kang et al. 2001; Lacombe et al. 2013; Baek, Kang, and Roeder 2006). Cependant, notre étude a montré pour la première fois, à notre connaissance, ce lien *in vivo* et a permis d'identifier les sous-unités impliquées.

Notre approche par double hybride nous a permis de révéler l'existence du contact entre Med14 et TFIIB. Cette méthode ne permet pas de prouver de façon stricte que l'interaction entre les deux partenaires testés est directe. Il pourrait être intéressant de le montrer. Il pourrait, par exemple, être envisagé une approche *in vitro* du type GST-pull down. Habituellement, ces

expériences sont réalisées avec des protéines recombinantes afin de vérifier si celles-ci peuvent interagir entre elles. Dans le cas du Médiateur, il est difficile d'envisager cette expérience en tant que telle. En effet, la plupart des sous-unités du Médiateur ne sont pas solubles lorsqu'elles sont exprimées seules, sans leurs partenaires directs (Koschubs et al. 2010). Il faudrait ainsi réaliser une expérience durant laquelle on mettrait en présence le Médiateur entier ou le module du milieu purifié et la protéine TFIIB recombinante afin de montrer que l'interaction a toujours lieu, sans intermédiaire. L'ajout de molécule permettant des pontages chimiques à la réaction pourrait même permettre de confirmer le contact direct avec Med14. Cette information serait très utile mais la démonstration d'une interaction directe *in vivo*, dans un contexte physiologique serait encore plus rigoureuse. La méthode de *photo-crosslink* utilisée par notre laboratoire pour montrer l'interaction entre Med17 et Rpb3 pourrait être envisagée (Soutourina et al. 2011).

A l'heure actuelle, aucun modèle structural de la sous-unité Med14 n'est disponible. Il ne nous est donc pas possible d'utiliser des données existantes pour prédire les domaines d'interaction entre Med14 et TFIIB. Il serait intéressant d'analyser par double hybride l'interaction entre ces deux protéines en utilisant des fragments des deux partenaires afin de préciser les domaines impliqués dans le contact. Lorsque l'on analyse la répartition des résidus mutés dans la protéine TFIIB dans les différents mutants qui sont coléaux avec les mutants de Med10 et Med14 (Figure 39), on constate qu'ils sont répartis sur toute la protéine, dans les différents domaines structuraux. A ce stade, il n'est pas possible de prédire quel domaine en particulier interagit avec Med14. Il est néanmoins possible que certaines des mutations de TFIIB affectent la structure tertiaire de la protéine et par conséquent l'interface d'interaction avec Med14. Il sera ainsi utile d'analyser à l'aide de prédictions de structures, l'effet des mutations de TFIIB sur sa conformation générale. Parallèlement, il serait intéressant de segmenter Med14 en plusieurs domaines afin d'identifier les domaines en interaction avec TFIIB, en utilisant une approche par double-hybride. La connaissance de la topologie de l'interaction entre Med14 et TFIIB pourra être utile pour raffiner les futurs modèles structuraux du Médiateur et du PIC.

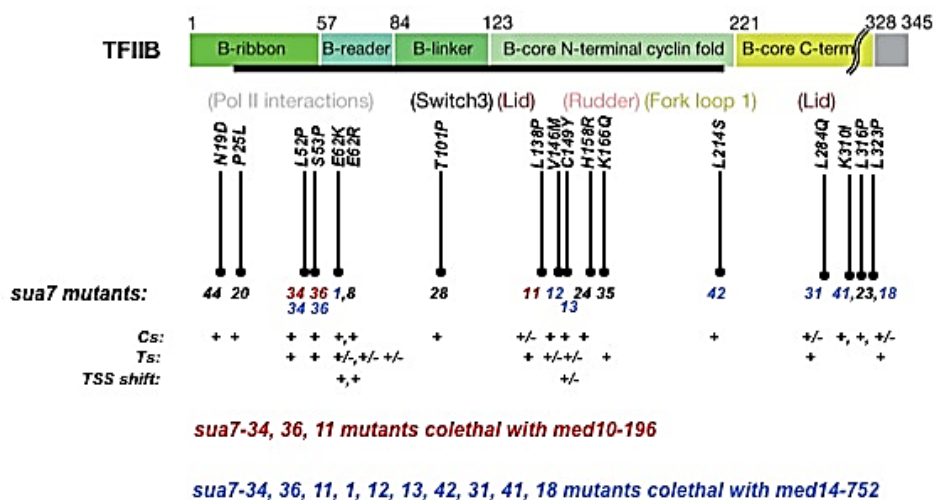


Figure 39 : Position sur la protéine des résidus mutés dans les différents mutants de TFIIB utilisés dans notre étude (Sainsbury, Niesser, and Cramer 2013).

Les phénotypes associés aux différents mutants de TFIIB (issus de Wu et al. 1999) que nous avons utilisé en combinaison avec les mutants de Med10 et Med14 sont indiqués. Cs : cold-sensitive, Ts : temperature-sensitive TSS shift : mutants qui provoquent un décalage du site d'initiation de la transcription

IV. Rôle du Médiateur dans l'assemblage du PIC.

1) Effet des mutations de sous-unités du module du milieu sur la formation du PIC.

L'analyse par CHIP de l'assemblage du PIC *in vivo* a montré que la mutation *med10-196* affecte fortement l'assemblage du PIC sur l'ensemble du génome. Cela nous a permis de montrer un rôle important de Med10 et du module du milieu dans la formation du PIC *in vivo*. De façon intéressante, le facteur présentant la plus forte diminution de son occupation est TFIIB. Ceci est cohérent avec le fait que l'interaction entre le Médiateur et TFIIB est diminuée dans le mutant *med10-196*. Les occupations de TFIID et TFIIA ne sont quant à elle pas affectées. Nous avons pu observer que la présence de TFIIH était affecté de façon importante également. Ces deux modules n'ont cependant pas le même comportement. Ceci est cohérent avec nos études précédentes ayant montré que le cœur de TFIIH et TFIIK pouvaient être recrutés indépendamment (Esnault et al. 2008; Eyboulet et al. 2015). TFIIE et TFIIF sont eux aussi affectés mais avec un effet intermédiaire en comparaison à TFIIB et TFIIH. Ces données montrent que l'assemblage du PIC *in vivo* ne correspond pas à la séquence linéaire proposée *in vitro*, de la même manière que dans nos études des sous-unités Med17 et Med11. Par exemple, dans le mutant *med10-196*, on observe que l'effet sur l'occupation de TFIIF n'est pas de la même intensité que celui sur la Pol II. Dans le modèle linéaire *in vitro*, TFIIF et la Pol II sont incorporés ensemble au sein du PIC. Nos données de CHIP-seq suggèrent un découplage des recrutements de TFIIF et de la Pol II.

Concernant la sous-unité Med7, les premiers résultats que nous avons obtenus par CHIP nous montrent que nous avons isolés des mutants thermosensibles qui perturbent l'assemblage du PIC. Nous avons notamment observé que les mutants analysés présentaient une forte diminution de l'occupation de la Pol II, montrant un effet sur l'activité transcriptionnelle. Nous avons pu constater une diminution de l'occupation de TFIIH au niveau de ses deux modules (cœur et kinase), et ce, de manière indépendante. Concernant le comportement du Médiateur nous avons constaté que l'effet des mutations était variable selon le module du Médiateur observé. La poursuite de nos analyses de CHIP-seq nous permettra d'investiguer l'effet des différentes mutations sur l'ensemble du génome. L'apparition d'un phénotype de floculation chez certains des mutants peut laisser penser que nous pourrions observer des effets gène spécifiques en perturbant l'expression de gènes particuliers. Enfin, nous souhaiterions, tout comme dans l'étude de la sous-unité Med10, étendre notre analyse à tous les GTFs. Une attention particulière sera portée sur TFIIB. En effet, une mutation dans le module du milieu pourrait aboutir à une destabilisation de l'interaction entre Med14 et TFIIB que nous avons mis en évidence. Ceci renforcerait l'idée de l'importance du module du milieu dans l'incorporation de TFIIB au sein du PIC. L'obtention de toutes ces données de CHIP-seq nous permettra probablement de mieux connaître le rôle de Med7 dans l'assemblage du PIC et apportera une meilleure compréhension du rôle du module du milieu dans la fonction du Médiateur.

Dans nos analyses, l'effet des mutations des sous-unités du Médiateur est observé par CHIP. Cette méthode fournit des instantanés de la liaison des facteurs étudiés. Le Médiateur a un rôle stabilisateur du PIC que les mutations pourraient perturber. Il pourrait être intéressant d'analyser la dynamique des différents GTFs. Dans le cas où certaines des mutations du Médiateur affecteraient la stabilisation du PIC, on peut s'attendre à ce que les différents GTFs présentent une dynamique différente. Une analyse par FRAP (Phair et al. 2004) ou par suivi de particule unique (*single particle tracking*) (Basu et al. 2016) de plusieurs sous-unités des GTFs dans les contextes

sauvage et mutant pourrait révéler ce genre de défaut.

L'information donnée par l'expérience de ChIP est l'occupation d'une protéine. On ne peut pas parler de recrutement. Nous pourrions étudier le recrutement à proprement parler des composants du PIC. Pour cela, l'analyse de la mise en place du PIC sur le gène *GAL1* pourrait être très utile. Comme évoqué dans la partie introductive, ce gène n'est activé qu'en présence de galactose dans le milieu de culture. Après ajout de galactose, on peut ainsi suivre en fonction du temps l'arrivée des composants du PIC. Dernièrement, une méthode appelée enChIP (*engineered-DNA binding molecule-mediated ChIP*), a été mise au point pour isoler de façon spécifique un *locus* particulier (Fujita and Fujii 2013). Cette méthode repose sur l'utilisation de la protéine Cas9. Cette protéine fait partie de la machinerie du système CRISPR/Cas9. Lorsqu'elle est co-exprimée avec un ARN guide (gRNA), elle présente la capacité de se lier à l'ADN sur une séquence spécifique complémentaire du gRNA. Dans les conditions physiologiques, le rôle de la Cas9 est d'induire une coupure de l'ADN. Dans la méthode enChIP, la Cas9 utilisée est inactive catalytiquement. Elle ne peut ainsi qu'être recrutée sur un *locus* et on réalise un ChIP dirigé contre la Cas9 afin de purifier uniquement le *locus* ciblé par le gRNA. Il devient alors possible de purifier les protéines présentes sur ce *locus*. Dans notre cas, il pourrait être envisagé de cibler la Cas9 en amont du promoteur du gène *GAL1*. On pourrait ensuite réaliser des enChIP à différents temps après induction par le galactose. Une analyse par spectrométrie de masse quantitative des protéines recrutées sur le promoteur en fonction du temps dans les contextes sauvage et mutant pourrait nous permettre de mettre en évidence des défauts de recrutement des différents GTFs. On pourrait ainsi vérifier si l'ordre d'arrivée des GTFs est perturbé dans nos différents mutants. Dans le cas du mutant *med10-196*, on pourrait vérifier si la mutation affecte le recrutement de TFIIB ou seulement sa stabilisation. Cette méthode pourrait être appliquée pour d'autres types de gènes inductibles, régulés par des activateurs différents. Les effets observés pourraient être variable selon les activateurs et ainsi montrer éclairer le fonctionnement du Médiateur sur le recrutement du PIC de façon gène-spécifique.

Nous avons, au cours de mon projet de thèse, réalisé une analyse *in vitro* de l'effet de la mutation de Med10 sur la formation du PIC. Cette méthode repose sur l'incubation d'un extrait nucléaire de levure (d'une souche sauvage ou mutante) avec avec une matrice d'ADN contenant un promoteur, immobilisée sur des billes magnétiques, en présence ou en absence d'un activateur purifié. Au terme de cette incubation, le PIC est assemblé et sa composition en protéines est analysée. Cette expérience nous a permis de confirmer l'observation faite *in vivo*, à savoir un défaut d'activation de la transcription. On a également observé que le défaut d'incorporation de TFIIB était aussi visible *in vitro*. L'utilisation de cette approche est un progrès par rapport à nos études précédentes puisqu'elle permet la confirmation de nos résultats selon deux approches différentes. Néanmoins, à ce jour, notre méthode d'analyse *in vitro* est limitée en résolution. En effet, nous analysons actuellement la composition du PIC par western blot, limitant ainsi la visualisation à quelques protéines en raison du faible nombre d'anticorps disponibles. Nous sommes actuellement en train de mettre au point une méthode d'analyse par spectrométrie de masse quantitative pour améliorer la résolution de la méthode en permettant potentiellement la visualisation simultanée de l'ensemble des protéines impliquées dans la formation du PIC. Cette méthode, lorsqu'elle sera au point, nous permettra d'étudier tous nos mutants thermosensibles du Médiateur afin d'analyser leur effet sur le recrutement des composants du PIC et sur la réponse à l'activation.

L'approche utilisée pour l'assemblage du PIC *in vitro* peut également être utilisée pour analyser l'activité transcriptionnelle. En effet, l'ajout de NTP au système autorise la Pol II à transcrire la matrice d'ADN. L'analyse des ARN transcrits pourrait nous permettre de mettre en évidence des défauts d'activité transcriptionnelle ou dans les phases de transition entre l'initiation et l'élongation. Cette approche est à la base des études qui ont mis en évidence l'existence *in vitro* d'une plateforme permettant la ré-initiation de la transcription (Yudkovsky, Ranish, and Hahn 2000). On pourrait ainsi utiliser ces expériences pour vérifier si nos différents mutants

thermosensibles du Médiateur affectent la ré-initiation de la transcription.

2) Lien entre le Médiateur et l'architecture des promoteurs.

L'analyse des données de ChIP-seq réalisées au cours de notre étude nous a permis de mettre en évidence un lien entre le Médiateur et la formation du PIC dans le contexte chromatinien. En effet, l'amplitude de l'effet de la mutation *med10-196* sur certains facteurs est liée à l'architecture du promoteur des gènes en termes d'élément TATA, de composition en nucléosomes et de dynamique des nucléosomes +1 et -1. Dans le cas de TFIIB, TFIID et la Pol II, les gènes les plus affectés sont ceux enrichis en boîte TATA. Ils sont majoritairement plus pauvres en nucléosomes sur leur promoteurs et leurs nucléosomes +1 et -1 sont plus dynamiques. Un mécanisme d'assemblage du PIC en fonction de l'architecture des promoteurs en terme d'éléments TATA a été proposé (Rhee and Pugh 2012). Il est proposé que les gènes contenant une boîte TATA voient le PIC assemblé en compétition avec le nucléosome +1 et que pour les promoteurs contenant un élément TATA-like, l'assemblage du PIC soit coopératif avec le nucléosome +1. Nos résultats montrent que le Médiateur peut influencer l'assemblage du PIC probablement en influençant ces voies d'assemblage. Pour le Médiateur, de manière intrigante, la tendance est inversée. En effet, les gènes les plus affectés pour l'occupation du Médiateur sont majoritairement des gènes dont le promoteur est dépourvu de boîte TATA et enrichis en terme de présence de nucléosomes. Ce résultat est intéressant car il suggère que l'occupation du Médiateur ne reflète pas complètement sa fonction *in vivo*. Il serait utile d'investiguer la dynamique du complexe afin de vérifier la corrélation qui pourrait exister entre son temps de résidence sur les régions régulatrices et l'activité transcriptionnelle. Actuellement, au laboratoire, nous mettons au point des expériences de ChIP-compétitive afin d'analyser la dynamique du Médiateur de façon globale.

Chez la levure *S. cerevisiae*, environ 20% des gènes possèdent une boîte TATA canonique (Rhee and Pugh 2012). Ces gènes sont considérés comme dépendant de SAGA pour leur expression, notamment pour le recrutement de la TBP. Certains travaux *in vitro* ont montré qu'il pouvait exister un recrutement coopératif du Médiateur et de SAGA suite à la présence d'un activateur (Qiu et al. 2005). L'observation du fait que les gènes les plus affectés pour certains facteurs soient des gènes possédant une boîte TATA pourrait être dû à un défaut de présence de SAGA. La mutation *med10-196* pourrait perturber une interaction potentielle entre SAGA et le Médiateur et ainsi perturber leur coopération. Nous pourrions analyser par co-IP la possibilité d'une interaction entre SAGA et le Médiateur dans le mutant *med10-196* pour éventuellement mettre en évidence des défauts. Une analyse par ChIP-seq de la distribution de SAGA dans le contexte mutant serait également très intéressante puisque l'on pourrait éventuellement corréler l'effet de la mutation *med10-196* sur son occupation avec l'effet sur les GTFs. Ceci permettrait de mettre à jour un lien fonctionnel entre le Médiateur et SAGA dans l'expression des gènes en fonction de l'architecture de leur promoteurs.

3) Effet gène-spécifique des mutations de la sous-unité Med7 : étude du phénotype de floculation.

Comme nous avons pu le voir plus tôt, certains des mutants de Med7 que nous avons obtenus présentent un phénotype particulier de floculation. Ce phénotype implique l'expression d'un réseau complexe de gènes, notamment les gènes de la famille *FLO*. Ces gènes, à l'exception de *FLO8* qui code pour un facteur de transcription, codent pour des protéines de surface de la famille des lectines. Le phénotype de floculation est important dans les applications industrielles faisant appel à la levure telle que la fabrication d'alcools. Ces gènes sont également impliqués dans la formation des biofilms, la croissance invasive dans l'agar et la croissance pseudo-hyphale (Guo et al. 2000; Reynolds and Fink 2001). Ces phénotypes sont impliqués dans la pathogénie des champignons parasites des eucaryotes pluricellulaires. La compréhension de leur expression est donc importante.

Des mutations de sous-unités du Médiateur ont déjà été rapportées comme provoquant des phénotypes de floculation. Dans le cadre d'une collaboration avec le laboratoire de Laurent Kuras, une collection de mutants de Med21 a été obtenue. Certains de ces mutants présentaient eux aussi un phénotype de floculation (données non publiées). On observe aussi ces phénotypes lors de la délétion de Cdk8 qui a pour effet une très forte augmentation de l'expression de *FLO11* (Holstege et al. 1998; Linder et al. 2008).

En partant de cette dernière observation, il pourrait être intéressant d'analyser par co-IP la relation entre le module kinase et le cœur du Médiateur dans ces mutants qui floculent. On peut, par exemple, s'attendre à une perte de ce module. Néanmoins, comme celui-ci n'est pas toujours présent sur le complexe, il pourrait être judicieux d'analyser par CHIP la présence des modules Cdk8 et du cœur sur les promoteurs des gènes *FLO*. Les mutations pourraient avoir pour effet d'induire la dissociation du module kinase sur ces gènes et ainsi lever un rôle inhibiteur.

Un mécanisme de régulation a été proposé pour le gène *FLO11* (Figure 40) (Bumgarner et al. 2009).

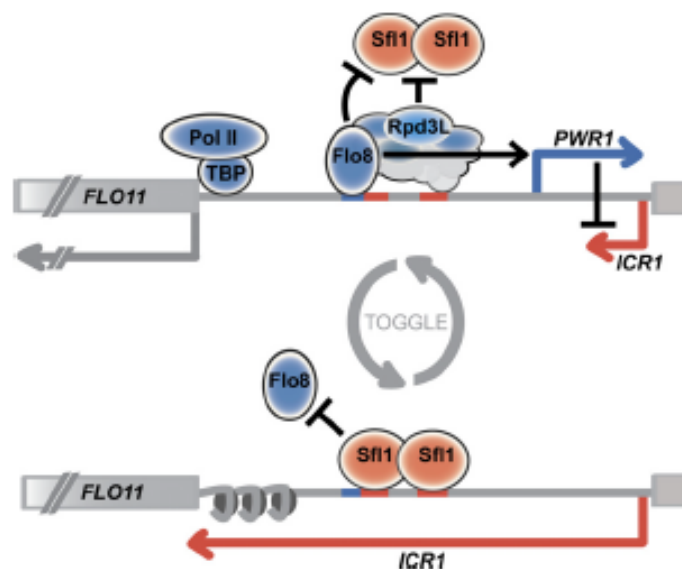


Figure 40: Schéma représentant un modèle de régulation type "interrupteur" de l'expression du gène *FLO11* (d'après Bumgarner et al. 2009).

Dans ce modèle, l'expression de *FLO11* est régulée par l'expression de deux long ncRNA : *PWR1* et *ICR1*. L'expression de ces lncRNA est elle-même régulée par deux facteurs de transcription différents, Flo8 et Sfl1. Lorsque Flo8 est recruté, il permet le recrutement d'un remodelleur de la chromatine Rpd3L et du PIC pour la transcription du gène *FLO11* et du lncRNA divergent *PWR1*. Lorsque le gène *FLO11* est réprimé, le facteur de transcription Sfl1 est recruté au promoteur. Il empêche la liaison de Flo8. On observe alors la transcription convergente du lncRNA *ICR1* inhibant l'expression de *FLO11* par interférence transcriptionnelle.

Ce modèle ne prend néanmoins pas en compte le Médiateur dans la régulation de la transcription de *FLO11*. Il doit intervenir dans ce système et il est possible que les mutations de Med7 puissent être impliquées dans son dérèglement. Il pourrait ainsi être intéressant d'étudier la régulation de l'expression des lncRNA dans les contextes mutants de Med7 afin de détecter un rôle potentiel du Médiateur. D'une manière plus générale, ce système pourrait être utilisé pour étudier le rôle du Médiateur dans l'expression d'ARN non codants.

L'investigation de l'implication du Médiateur dans le phénotype de floculation pourrait apporter des informations quant au rôle du complexe dans la régulation de gènes particuliers.

4) Rôle des différentes sous-unités essentielles dans la formation du PIC *in vivo*.

L'étude des sous-unités Med7 et Med10 réalisée ici s'inscrit dans la démarche du laboratoire d'analyser le rôle des sous-unités essentielles du Médiateur. L'analyse de la sous-unité Med7 est en cours et l'étude du rôle d'autres sous-unités essentielles sera poursuivie. Les données obtenues pourront être intégrées avec celles des études de Med11, Med17 et Med10. Nous espérons, à l'aide de toutes ces données à l'échelle du génome, pouvoir réaliser des analyses bioinformatiques qui permettront de comprendre de façon détaillée le rôle du Médiateur dans la formation du PIC et dans la régulation de la transcription. Nous espérons, à terme, pouvoir extraire de ces analyses les règles gouvernant la mise en place du PIC *in vivo* et la contribution du Médiateur dans ce processus pour chaque gène *S. cerevisiae*.

V. Analyse structurale du Médiateur.

La structure tridimensionnelle complète et à l'échelle atomique du Médiateur est à l'heure actuelle inconnue. Néanmoins, le Médiateur possède une structure flexible, lui permettant des changements de conformations. On sait, par exemple, qu'il peut changer de conformation suite à sa liaison à des activateurs (Taatjes et al. 2002). Le module du milieu possède une structure très flexible, pouvant jouer le rôle de charnière dans ces transitions conformationnelles (Larivière et al. 2013). Il est donc envisageable que la mutation *med10-196* ainsi que les mutations de Med7 puissent perturber la structure du module du milieu et ainsi la flexibilité intrinsèque du Médiateur. La sous-unité Med7 est impliquée dans la flexibilité du module du milieu (Baumli, Hoepfner, and Cramer 2005). Des perturbations structurales pourraient expliquer, par exemple, le défaut d'interaction entre Med14 et TFIIB observé dans le mutant *med10-196*. La mutation pourrait aboutir à une conformation peu propice au contact entre Med14 et TFIIB. De façon plus générale, un défaut de conformation pourrait aussi expliquer les effets sur la formation du PIC.

Nous avons envisagé d'analyser par microscopie à force atomique (AFM) le Médiateur issu du mutant *med10-196*, en présence et en l'absence d'un activateur, afin de montrer d'éventuels défauts dans les changements de conformations en comparaison au Médiateur sauvage. Ces expériences sont toujours en cours, en collaboration avec Olivier Piétrement de l'Institut Gustave Roussy. Actuellement, les méthodes de cryo-microscopie électronique sont en plein essor. Ces méthodes permettent aujourd'hui d'atteindre des résolutions quasi-atomiques dans certains cas. Nous pourrions utiliser ces approches afin de réaliser nos analyses structurales. Nous améliorerions nettement la résolution en comparaison à l'AFM. Ces études de structures pourraient également être menées dans le cas de mutants d'autres sous-unités du Médiateur, afin d'analyser leurs effets éventuels sur la topologie du complexe.

VI. Conservation du rôle de Med7, Med10 et du Médiateur en général chez les mammifères.

Étant donné son rôle central dans l'activation de la transcription, le Médiateur est impliqué dans de nombreuses pathologies humaines et notamment dans différents types de cancers. L'étude du Médiateur est donc indispensable afin de décrypter les mécanismes pouvant mener à ces pathologies et pouvoir éventuellement les traiter.

La comparaison des séquences primaires des sous-unités du Médiateur entre les eucaryotes montre une conservation limitée pour certaines sous-unités. Néanmoins, la prédiction de structures secondaires révèle des motifs conservés entre la levure *S. cerevisiae* et l'Homme (Bourbon 2008). Les mutations présentes dans le mutant *med10-196* concernent des résidus conservés chez l'Homme (Figure 41). Ce constat a également été fait pour certains mutants de Med17 que nous avons obtenus.

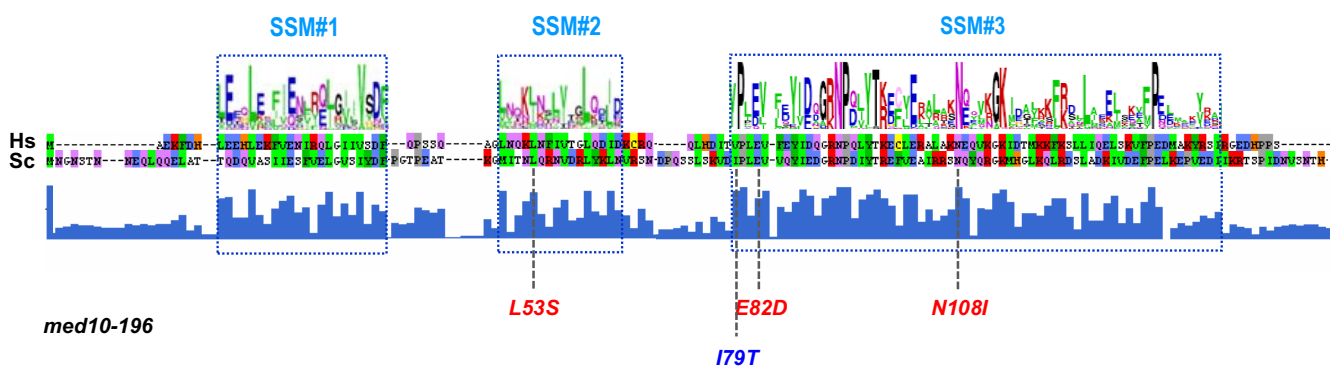


Figure 41: Alignement des séquences protéiques de Med10 chez l'Homme et la levure *S. cerevisiae* (Bourbon 2008).

Les mutations de *med10-196* sont indiquées sous les séquences. Hs : *Homo sapiens*, Sc : *Saccharomyces cerevisiae*, SSM : *sequence signature motif*.

Nous aimerions ainsi étendre nos travaux à l'étude de ces sous-unités essentielles du Médiateur chez l'Homme. Pour cela, il faut obtenir par mutagenèse dirigée les mutants des sous-unités humaines afin de les exprimer dans des lignées cellulaires. Concernant Med10, j'ai initié cette étude durant mon projet de thèse. J'ai établi des lignées stables de cellules HeLa exprimant les mutations correspondantes à *med10-196*. Pour cela, nous avons commencé par cloner le gène *MED10* sauvage et mutagénisé dans un vecteur retroviral bicistronique (pREV) permettant la co-expression de Med10 et d'un marqueur IL2Ra. Ce marqueur est une protéine de surface, récepteur

d'une cytokine. Ce vecteur a été utilisé pour générer des particules rétrovirales afin de transformer des cellules HeLa. En utilisant des billes magnétiques anti-IL2Ra, nous avons pu isoler les cellules ayant intégré la construction et ainsi obtenir une population de cellules exprimant les versions sauvage et mutante de Med10.

Sur cette lignée, nous avons ensuite réalisé des tests de prolifération (Figure 42). On peut constater dans ces expériences que l'expression de la version mutante de Med10 à elle seule ne semble pas provoquer de phénotype visible sur la prolifération. Cela suggère que la version endogène de Med10 est suffisante pour une viabilité normale. Par contre, lorsque l'on diminue l'expression de la sous-unité Med10 endogène à l'aide d'un siRNA (*small interfering RNA*) ciblant la partie 3' UTR (*untranslated region*) de *MED10*, on observe une diminution d'environ 25% de la prolifération des cellules exprimant la version mutée de Med10. Ce résultat est tout à fait préliminaire puisque la validation de l'efficacité du siRNA dirigé contre *MED10* n'a pas été faite et l'expérience nécessite d'être reproduite. Malgré tout, ce résultat suggère que la mutation *med10-196* semble induire un phénotype de ralentissement de la croissance également chez l'Homme.

Malheureusement, nous avons remarqué ensuite que cette approche n'était peut être pas la plus adaptée pour notre étude. En effet, après un certain nombre de générations, ces cellules HeLa n'exprimaient plus de façon suffisante la version exogène sauvage ou mutante de Med10. Le marqueur utilisé (IL2Ra) ne permet pas le maintien d'une pression de sélection sur ces lignées cellulaires, comme cela peut être fait avec des marqueurs de résistance à un antibiotique. Cela implique la nécessité de retrier systématiquement les cellules avant chaque expérience pour travailler avec une lignée exprimant de façon homogène la construction.

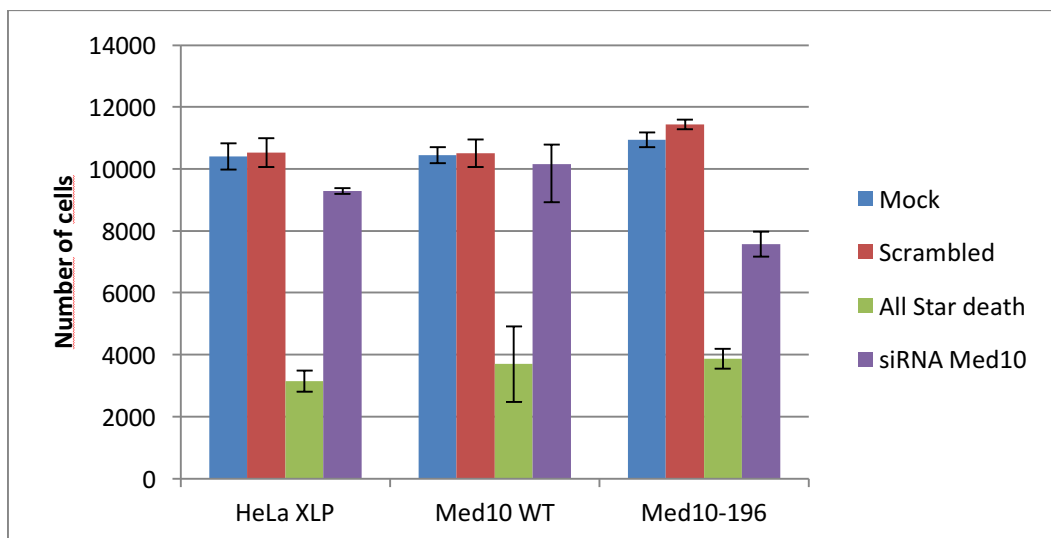


Figure 42: Test de prolifération des lignées cellulaires HeLa sauvages ou exprimant stablement *Med10-WT* ou *med10-196*.

Mock signifie transfection contrôle négatif, sans siRNA. *Scrambled* signifie transfection d'un siRNA sans cible. *All Star death* est un siRNA contrôle positif de transfection en induisant la mort cellulaire. *siRNA Med10* est la condition de transfection d'un siRNA dirigé contre *Med10*.

Nous avons donc décidé de mettre en place une lignée stable mais cette fois en utilisant un marqueur de résistance à la Néomycine pour pouvoir maintenir une pression de sélection sur les lignées stables obtenues. Les versions sauvage et mutante de Med10 ont ainsi été clonées dans le vecteur pIRES-Neo, permettant d'effectuer des transfections stables dans des cellules. La transfection d'un siRNA dirigé contre la version endogène est coûteuse et apporte de la variabilité aux différentes expériences. Nous avons donc opté pour une stratégie d'expression stable d'un

shRNA (*short hairpin RNA*) dans ces lignées exprimant la version mutée de Med10. Cela permettra l'obtention de lignées exprimant la version mutée de Med10 dans un contexte où l'expression de la version endogène sauvage est quasi nulle. Il peut également être envisagé l'utilisation de la technologie CRISPR/Cas9 pour déléter totalement la version endogène de Med10. Cette technologie pourrait même permettre en théorie d'obtenir la version mutante de Med10 au *locus* naturel du gène.

Les lignées stables obtenues seront le support d'études de prolifération afin de vérifier si l'on peut reproduire les premiers résultats. Nous analyserons ensuite ces lignées de façon phénotypique (forme, temps de génération, différentes conditions de croissance ou de stress). Nous réaliserons ensuite des études de CHIP et de CHIP-seq afin de vérifier si la mutation de Med10 induit des effets sur la composition du PIC chez l'Homme également.

La comparaison de la séquence de Med7 entre l'Homme et *S. cerevisiae* montre, tout comme pour Med10, la conservation des certaines régions (Figure 43). On remarque que les mutations que nous avons isolées en mutants ponctuels chez la levure (*med7-L126P*, *med7-L133P* et *med7-S173P*) se trouvent toutes les trois dans une région conservée. Nous avons initié la même approche que celle utilisée pour la sous-unité Med10 afin de transposer ces mutations dans des lignées cellulaires HeLa. Les résultats que nous obtiendrons pourront nous permettre de mieux comprendre la fonction du Médiateur chez l'Homme. Etant donné son implication dans des pathologies, la compréhension détaillée du Médiateur humain dans la régulation de l'expression des gènes pourrait permettre d'en faire une cible thérapeutique dans le futur.

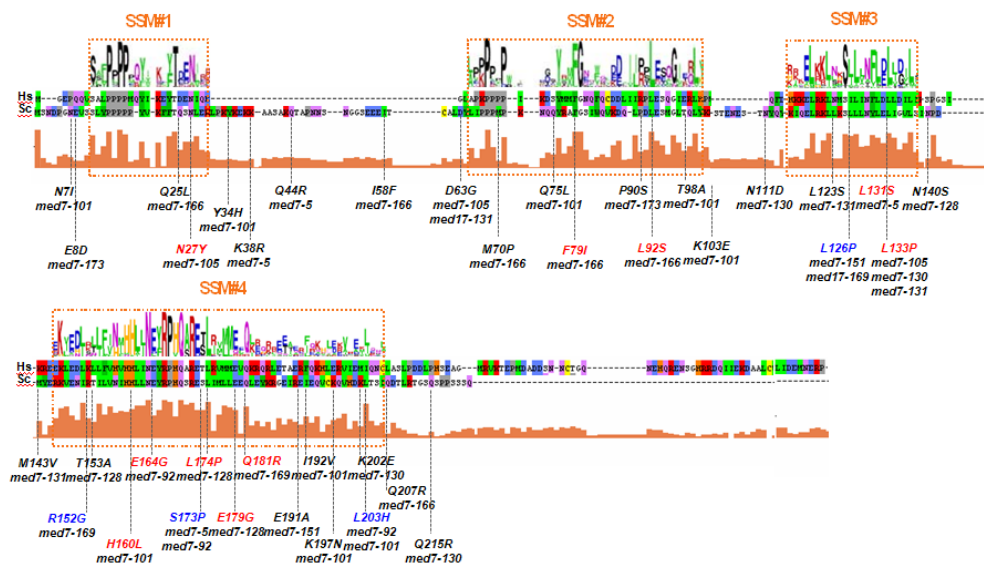


Figure 43: Alignement des séquences protéiques de Med7 chez l'Homme et la levure *S. cerevisiae*. (D'après Bourbon 2008)

Méthodologie

Chapitre I : Purification et analyse du Médiateur par spectrométrie de masse.

Dans notre laboratoire, les études des sous-unités essentielles du Médiateur reposent sur la caractérisation de mutants thermosensibles. Ces mutants présentent l'intérêt, contrairement aux délétions, de maintenir la protéine présente au sein du Médiateur. Les mutations perturbent ainsi seulement la fonction de la protéine, en affectant des interactions spécifique et/ou sa structure. Néanmoins, ces mutations peuvent également destabiliser le Médiateur et ainsi aboutir à la perte de son intégrité. C'est, par exemple, ce qui se produit dans le cas de la sous-unité Med17 avec le mutant *med17-138* dont la mutation engendre la dissociation du module de tête (Holstege et al. 1998). La perte de l'intégrité du Médiateur pour les mutants que nous étudions est un phénomène que nous souhaitons exclure. En effet, la dissociation du complexe conduit à la perte totale de la fonction du Médiateur et a de très grande probabilité d'affecter la transcription de façon massive. Les conclusions tirées de l'analyse du mutant en question seraient alors probablement indirectes et indépendantes du rôle de la sous-unité étudiée. Pour cette raison, dans notre démarche expérimentale, après l'obtention des mutants thermosensibles des sous-unités du Médiateur, nous analysons systématiquement la stabilité du complexe avant toute étude fonctionnelle. Jusqu'alors, nous analysions l'intégrité du complexe par des méthodes de co-immunoprécipitation. Nous étions limité à la visualisation de la présence de seulement quelques sous-unités, cela étant dû au fait qu'il n'existe pas ou peu d'anticorps efficaces pour la détection des sous-unités du Médiateur.

Pour cette raison, nous avons décidé de mettre au point une méthode pour analyser, par spectrométrie de masse, la composition du Médiateur dans les contextes sauvage et mutants pour les différentes sous-unités. La spectrométrie de masse est une méthode analytique permettant d'identifier des espèces chimiques en fonction de leur masse moléculaire. En biologie moléculaire, l'utilisation de la spectrométrie de masse a été développée ces dernières années pour la caractérisation des protéines. Notamment, la méthode de spectrométrie de masse en tandem (MS/MS) est très utilisée dans les approches de protéomique. Par ces approches, on cherche à caractériser les protéines présentes dans un mélange plus ou moins complexe. Les protéines sont dans un premier temps digérées par une protéase (très souvent la trypsine) pour générer des peptides. Ces peptides sont ensuite soumis à une étape de chromatographie en phase liquide à l'aide d'une colonne (LC, *liquid chromatography*). Cette étape permet de séparer les peptides en fonction de leur propriétés physico-chimique (taille, hydrophobie...). A la sortie de la colonne, les peptides sont injectés dans le spectromètre de masse afin d'être analysés par MS/MS. Au cours d'une analyse MS/MS, les molécules sont soumises à deux analyses successives. Dans le cas de peptides, la première analyse mesure la masse du peptide. Au terme de cette mesure, le peptide est fragmenté puis les masses de ses produits de fragmentation sont analysées. Chaque peptide ayant une composition donnée, les profils de fragmentation ainsi que de masses sont uniques à chaque peptides. Connaissant la séquence d'une protéine, il est possible de déterminer les peptides obtenus par digestion tryptique, la masse de ces peptides et ses produits de fragmentation. La projection des spectres de MS/MS sur des bases de données de protéines permet donc d'attribuer à chaque peptide analysé, la protéine dont il est issu.

Initialement, nous souhaitions utiliser l'analyse par spectrométrie de masse à l'issue de la réaction d'immunoprécipitation. L'inconvénient de cette approche est qu'elle fait appel à l'utilisation d'un anticorps, ajouté en excès, pour isoler le Médiateur. Cet anticorps est ainsi présent dans le mélange final à analyser par MS/MS. Etant donné qu'il est ajouté en excès, le signal

obtenu pour sa détection est très supérieur à celui des autres protéines. La détection des sous-unités du Médiateur risquait, dans ces conditions, d'être masquée par ce signal. Nous avons donc opté pour une stratégie de purification biochimique du Médiateur.

I) Principe.

Les propriétés physico-chimiques des protéines permettent de les isoler à partir d'un extrait cellulaire en utilisant différentes méthodes biochimiques. L'ajout d'étiquettes par fusion sur une protéine permet d'utiliser des méthodes de purification par affinité. Ainsi, une méthode appelée TAP (*tandem affinity purification*) a été mise au point il y a quelques années (Rigaut et al. 1999). Cette méthode repose sur la fusion d'une étiquette particulière (TAP-tag) sur une protéine d'intérêt. Ce TAP-tag contient deux étiquettes différentes, permettant une purification par affinité en deux étapes. La première étiquette est un fragment de la protéine A, une protéine bactérienne ayant une très forte affinité pour les anticorps, en particulier les IgG. On utilise ainsi une première colonne d'affinité composée de résine couverte d'IgG. Les deux étiquettes sont séparées par un site spécifique de clivage par une endo-protéase, généralement la TEV protéase (*tobacco etch virus protease*). La digestion par la TEV permet ainsi l'élution de la première colonne. La deuxième étiquette, est généralement composée de la protéine CBP (*calmodulin binding protein*), ayant une forte affinité pour une protéine appelée calmoduline. On réalise alors une deuxième purification par affinité à l'aide d'une résine couverte de calmoduline. La liaison entre la CBP et la calmoduline est dépendante de la présence de calcium. La modification de la concentration en calcium permet d'éluer la deuxième colonne et donc de récupérer la protéine purifiée. Les conditions de purification utilisées dans la méthode TAP sont très peu stringentes et sont donc propices au maintien des interactions entre les protéines et à la purification des protéines dans des conditions natives. C'est donc une méthode très utilisée pour la purification de complexes protéiques.

Dans notre étude, nous avons adapté un protocole de purification du Médiateur déjà publié par le laboratoire de F. Asturias (Cai et al. 2009). Ce protocole a été utilisé pour purifier le Médiateur pour une analyse par microscopie électronique. Il consiste en une purification à l'aide d'un TAP-tag modifié. Dans ce cas, la deuxième étiquette est composée d'une répétition de 10 résidus Histidine, ayant une forte affinité pour le nickel. On peut ainsi utiliser une résine couverte de nickel, et éluer à l'aide d'imidazole, une molécule proche de l'histidine rentrant en compétition avec l'étiquette.

Afin d'appliquer cette méthode à notre projet, nous nous sommes procuré la souche de levure CA001 (*MATa leu2 trp1 ura3-52 prb1-1122 pep4-3 prc1-407 gal2 MED22::10His-TEV-ProteinA::KITRP1 MED8-PreSci-3HA::KanR*) auprès du laboratoire de F. Asturias. Cette souche de *S. cerevisiae* présentant la caractéristique d'être déficiente pour certaines protéases endogènes, la rendant plus propice à des purifications de protéines. Son contexte génétique est dérivé de la souche S288c, utilisée dans nos travaux. Elle comporte l'ajout du TAP-tag modifié sur la sous-unité Med22, une sous-unité du module de tête et d'une étiquette HA sur la sous-unité Med8, elle aussi faisant partie du module de tête.

La démarche mise en place commence par la construction de la souche afin d'obtenir nos mutants dans ce contexte génétique. Nous présenterons ici (Figure 44) la méthode appliquée pour le cas de l'étude du mutant *med10-196*, mais la démarche est applicable pour toute autre sous-unité. Pour cela, nous avons commencé par sous-cloner les gènes *MED10* et *med10-196* dans un vecteur pAG-GPD-LEU (Alberti, Gitler, and Lindquist 2007). Ce vecteur est un plasmide permettant l'utilisation de la technologie Gateway®, basé sur des clonages entre vecteurs à l'aide de recombinaison homologue *in vitro*. Par l'utilisation de cette technologie, et selon les protocoles du

fournisseur, nous avons réalisé une réaction LR pour obtenir le vecteur pAG-GPD-MED10 et pAG-GPD-med10-196. Ces plasmides permettent l'expression des deux versions Med10 (sauvage ou mutante), sous contrôle d'un promoteur constitutif.

La souche CA001 est en premier lieu transformée par le plasmide pVV208-MED10-URA3. On procède ensuite, dans cette souche transformée, à la délétion au *locus* du gène MED10 à l'aide d'une cassette de résistance à la Nourséothricine (Janke et al. 2004). Cette souche obtenue est ensuite transformée par les plasmides pAG-MED10 ou pAG-med10-196. Certains des clones obtenus sont ensuite mis en culture sur des boîtes de Pétri contenant du 5FOA (5-*fluoro-orotic acid*), pour permettre la chasse du plasmide pVV208-MED10-URA. Les clones obtenus sont ensuite testés pour leur phénotype thermosensible à 37°C.

Les souches CA001 exprimant la version sauvage ou mutée de Med10 sont ensuite utilisées pour la purification du Médiateur. Pour cela, 10L de culture sont réalisés dans du milieu riche 2X YPD (*yeast-extract peptone dextrose*). L'utilisation d'un tel milieu permet l'obtention d'une plus grande densité de cellules. La culture est menée jusqu'en fin de phase exponentielle. Les cellules sont récupérées par centrifugation. Le culot cellulaire est lysé à l'aide d'une presse de Eaton. L'extrait protéique brut ainsi obtenu est clarifié par une ultracentrifugation, puis soumis à une précipitation sélective au sulfate d'ammonium. Lors de cette précipitation, on isole la fraction des protéines comprise entre 30 et 55% de saturation de sulfate d'ammonium. Cette fraction de protéine est reprise dans un tampon et incubée en présence de résine IgG. Après différents lavages, la résine est incubée en présence de TEV protéase afin d'éluer le Médiateur lié à la résine. Cet éluat est ensuite soumis à une incubation avec une résine de nickel (Ni-NTA). Après lavage, le Médiateur est élué.

Avant toute analyse par MS, nous contrôlons la présence de certaines sous-unités du Médiateur afin de vérifier l'efficacité de la purification. Pour cela, une partie de l'échantillon est analysée par western blot pour vérifier la présence des sous-unités Med14, Med17 et Med8-HA. Nous préparons ensuite l'échantillon pour l'analyse MS. La présence de certains composés chimiques, notamment des détergents, n'est pas compatible avec des analyses par MS. Pour cette raison, nous ne pouvons fournir nos échantillons en solution. Nous déposons ainsi nos échantillons sur un gel d'acrylamide afin de réaliser une migration par SDS-PAGE (*SDS-polyacrylamide gel electrophoresis*). On réalise une migration très courte, seulement pour faire « rentrer » les protéines dans le gel, sans séparation. Cela permet l'élimination des molécules chimiques pouvant interférer avec l'analyse MS. Ce gel est ensuite traité par la plateforme de spectrométrie de masse de l'institut de biologie intégrative de la cellule (i2BC) de Gif sur Yvette. La bande du gel correspondant à l'échantillon est découpée. Les protéines sont digérées par la trypsinne directement dans le gel. Les peptides obtenus sont extraits du gel et soumis à une analyse LC-MS/MS. À la suite de cette analyse on obtient une liste des peptides identifiés dans les échantillons.

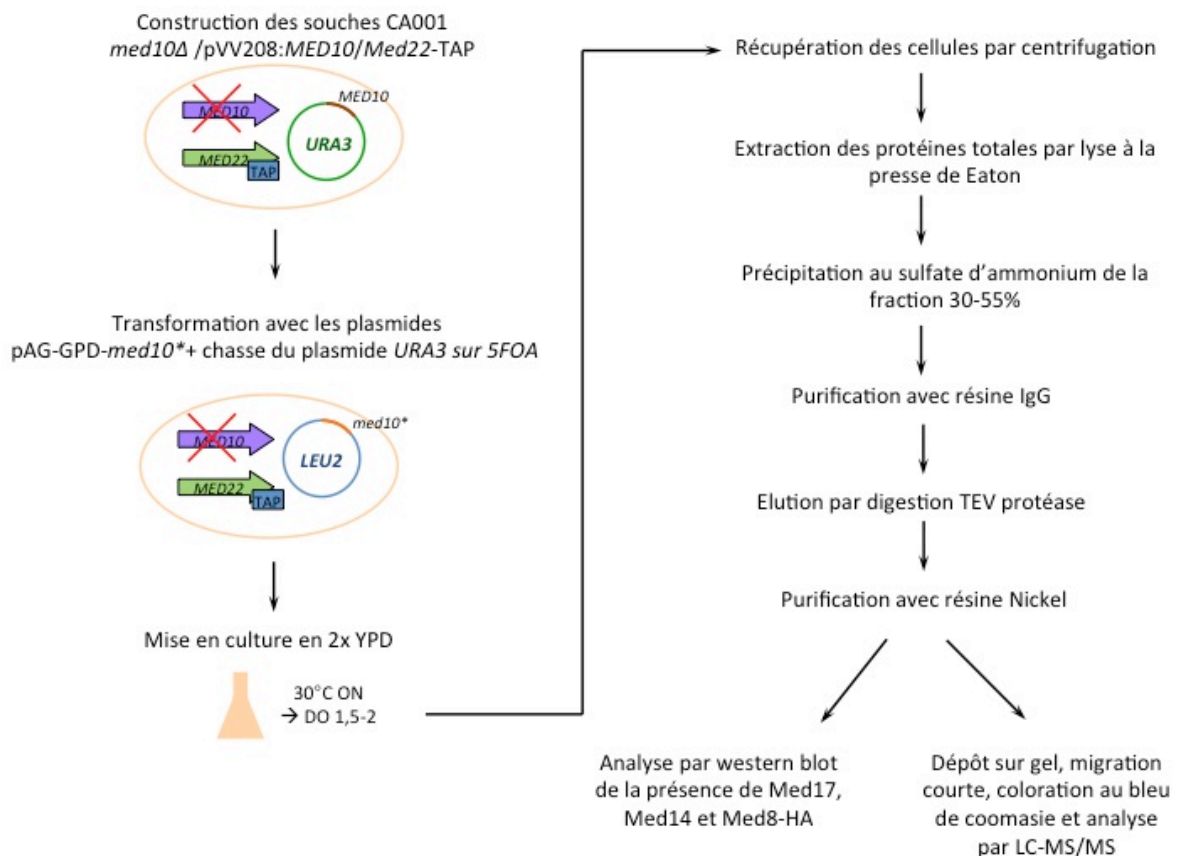


Figure 44: Schéma explicatif de la démarche expérimentale mise en œuvre pour l'étude de la composition du Médiateur par spectrométrie de masse.

II) Protocole de la purification.

Le protocole a été mis au point pour un volume de culture de 10L. Il peut être réalisé pour 2L de culture en adaptant les proportions de réactifs à toutes les étapes.

- Préparation de la souche :

La souche CA001 à partir de laquelle la purification est réalisée est sortie de collection et striée sur une boîte YPD solide. Avant de commencer le protocole de purification, vérifier le phénotype thermosensible dans le cas de mutants.

- Culture :

Afin d'augmenter le nombre de cellules pour la purification, le milieu de culture utilisé est du YPD deux fois concentré. Le glucose est ainsi à une concentration de 4% final. A ces concentrations, la stérilisation par autoclave peut dégrader le sucre. Il est donc préparé séparément et ajouté de

façon stérile après passage à l'autoclave du milieu.

- Préparation du milieu de culture : Pour 10L de milieu, peser 200g de *Yeast extract* et 400g de Peptone. Dissoudre dans 9L d'eau distillée. Une fois dissous, répartir 900mL de milieu dans 10 Fernbach. Stériliser par passage à l'autoclave. Parallèlement, peser 400g de glucose et le dissoudre dans 1L d'eau distillée (Attention, la dissolution est longue !!). Après dissolution, filtrer de façon stérile la solution de glucose à 40%. Après autoclave, ajouter 100mL de glucose 40% à chaque Fernbach sous une hotte stérile.
- Au jour 0,ensemencer une colonie de la souche CA001 dans un volume de 50mL de milieu 2X YPD, pour une préculture sur la nuit.
- Au jour 1, dans l'après-midi,ensemencer les Fernbach avec 5mL de préculture saturée. Incuber à 30°C sur la nuit. (Avant d'ensemencer, vérifier au microscope que la préculture n'est pas contaminée par des bactéries, c'est un milieu riche, favorable aux contaminations).
- Au jour 3 : Vérifier la densité optique. Typiquement, on atteint une DO d'environ 2, en fin de phase exponentielle, début de phase stationnaire. Vérifier l'absence de contaminations.

Centrifuger les cultures dans des pots de centrifugation d'un litre, à 3000g pendant 10 minutes. Laver les culots avec de l'eau distillée puis centrifuger 10 minutes à 3000g. Eliminer l'eau et peser le culot. Habituellement, on obtient environ 15-20g de cellules par litre de culture en fin de phase exponentielle. Le culot peut être congelé à cette étape et gardé à -80°C.

- Préparation de l'extrait total de protéine :

Note : Pour lyser les cellules, nous utilisons une presse de Eaton. La presse doit être à -80°C au moment de la lyse. Penser, quelques jours avant l'expérience, à enrouler la presse et le piston dans du papier aluminium et à les placer au congélateur à -80°C.

Matériel nécessaire : Azote liquide, Un becher en verre (500mL), un bécher en plastique (500mL), des spatules et des gants de protections contre le froid.

- Avant de procéder à lyse, sortir la presse et le piston du congélateur. Placer le piston en immersion dans de l'azote liquide pendant 2-3 minutes. Visser ensuite le piston sur la tête de la machine. Placer le bloc de la presse sous le piston. Aligner le piston sur l'orifice du bloc. Faire descendre le piston jusqu'en bas du bloc. Fixer avec les vis le bloc à l'aide des 4 pattes métalliques. Régler la pression d'utilisation de l'appareil sur 18 tonnes.
- Resuspendre le culot de cellules dans le tampon de resuspension (1mL par gramme de cellules) et garder dans la glace.
- Congeler la suspension de cellules avec de l'azote. Utiliser un bécher en verre. (Attention, manier avec précaution le bécher, l'ajout d'azote le rend fragile). Ajouter l'azote par petite quantité en remuant avec une spatule jusqu'à obtention d'un extrait congelé, ressemblant à du pop-corn.
- Ajouter l'extrait congelé dans l'orifice de la presse. Ajouter, si besoin, un peu d'azote liquide dans l'orifice pour terminer la congélation des cellules.
- Placer un bécher EN PLASTIQUE sous la presse (Note : l'extrait lysé sortant de sous la presse peut être expulsé violemment. Il faut donc utiliser un bécher en plastique car le verre risque de se casser. De plus, ne JAMAIS placer ses mains sous la presse pendant la lyse, risque de blessure).
- Actionner la descente du piston jusqu'en bas. Le lysat est expulsé sous la presse dans le bécher. Il prend l'apparence de « spaghetti ».
- Renouveler jusqu'à épuisement de la suspension de cellules.

- Récupérer le lysat à l'aide d'une spatule et le placer dans des tubes Falcon. Placer ces tubes à -80°C.

- Précipitation au sulfate d'ammonium

- Sortir les lysats du congélateur et laisser décongeler dans la glace.
- Placer ces lysats dans des tubes pour rotor 45Ti (Beckman-Coulter) pour ultracentrifugation 1h à 40000 rpm.
- Récupérer le surnageant, mesurer son volume et placer dans un pot de centrifugation de 250mL
- Peser la quantité de sulfate d'ammonium nécessaire à l'obtention d'une saturation de 30%. (164g pour un litre)
- En chambre froide, sous agitation, ajouter la poudre de sulfate d'ammonium petit à petit.
- Laisser agiter doucement pendant 1h.
- Centrifuger à 3000g pendant 45 minutes.
- Récupérer le surnageant dans un pot de centrifugation de 250mL propre.
- Peser la quantité de sulfate d'ammonium nécessaire à l'obtention d'une saturation de 55% (148g pour un litre).
- En chambre froide, sous agitation, ajouter la poudre de sulfate d'ammonium petit à petit.
- Laisser agiter doucement sur la nuit.
- Le matin, centrifuger à 3000g pendant 45 minutes.
- Jeter le surnageant avec précaution et conserver le précipitat dans la glace.

- Purification sur colonne IgG

- Resuspendre le précipitat de protéine dans 20mL de tampon IgG binding buffer.
- Répartir la suspension dans deux tubes Falcon de 15mL (2*10mL).
- Dans un tube Falcon, laver 3 fois avec le tampon IgG binding buffer 1mL de résine IgG (IgG Sepharose 6 Fast Flow, GE Healthcare).
- Ajouter 500µL de résine dans chaque tube contenant l'extrait de protéine.
- Mettre sur une roue rotative pendant 3 heures en chambre froide.
- Transférer l'incubation dans une colonne vide de 5mL (ThermoScientific) et laisser le liquide s'écouler par gravité
- Laver la colonne avec 50mL de Washing buffer I, par gravité.
- Laver la colonne avec 50mL de Washing buffer II, par gravité.
- Laver la colonne avec environ 40mL de TEV digestion buffer (Sans inhibiteur de protéase)
- Préparer 2mL de TEV digestion buffer dans lequel on ajoute 100 U de TEV protéase (AcTEV™ Protease, ThermoScientific) et ajouter dans la colonne.
- Laisser la digestion sur la nuit en chambre froide à 4°C.
- Le matin, éluer la colonne avec 2 mL de TEV digestion buffer dans un tube Falcon de 15mL.

- Purification sur colonne Nickel

- Diluer 6-8 fois l'éluat de la première colonne dans du tampon Ni-NTA Binding buffer.
- Dans un tube Falcon, laver 3 fois avec le tampon Ni-NTA Binding buffer 1mL de résine Nickel (Ni Sepharose High Performance, GE Healthcare).
- Ajouter 1mL de résine à l'éluat.

- Mettre sur une roue rotative pendant 3 heures en chambre froide.
 - Transférer l'incubation dans une colonne vide de 5mL (ThermoScientific) et laisser le liquide s'écouler par gravité.
 - Laver la colonne avec 10mL de Ni-NTA wash buffer, par gravité.
 - Eluer avec 1mL de Ni-NTA elution buffer.
 - Ajouter 100µL de glycérol, aliquoter et conserver à -80°C.
- Analyse par western blot

Déposer 3 fois 10µL sur un gel SDS-PAGE et analyser la présence des sous-unités Med17, Med14 et Med8-HA.

- Dépôt sur gel pour analyse par MS

Note : Ne préparer le gel que le jour où il sera possible d'aller déposer le gel à la plateforme de spectrométrie de masse. Prendre le maximum de précautions pour éviter les contamination par des kératines (gants, laver les récipients, éviter les poussières etc.)

- Déposer 20µL de la purification dans un gel pré-coulé 4-12% (NuPAGE™ Novex™ 4-12% Bis-Tris Protein Gels, ThermoScientific)
- Faire migrer environ 5 minutes à 180V, 400mA. Arrêter la migration quand le bleu de charge est entièrement entré dans le gel. Pour indication, charger un marqueur de poids moléculaire.
- Laver le gel 3 fois 5 minutes dans de l'eau distillée
- Colorer le gel avec du bleu de coomasie (Imperial™ Protein Stain, ThermoScientific). Si utilisation d'un autre bleu, s'assurer de sa compatibilité avec la MS.
- Rincer 3 fois avec de l'eau distillée

Tampons et milieu utilisés pour ce protocole :

Toutes les solutions sont filtrées et gardées à 4°C. Quand indiqué, l'inhibiteur de protéase (Complete Protease Inhibitor, Roche) est ajouté extemporanément.

- 2X YPD : pour 1 litre 20g de yeast extract, 40g de peptone, 4% de glucose
- Solution de 40% glucose : 400g de D-Glucose par litre. Filtrer stérilement
- *Resuspension Buffer* : 250mM Tris pH7.6, 300mM sulfate ammonium, 2mM EDTA, 30% glycérol, 5mM 2-Mercapthoethanol, Inhibiteur de protéase
- *IgG binding buffer* : 25mM Phosphate pH7.5, 200mM acétate de potassium, 1mM EDTA, 10% glycérol, 5mM 2-Mercapthoethanol, Inhibiteur de protéase
- *IgG Wash buffer I* : 50mM Phosphate pH7.5, 500mM sulfate ammonium, 1mM EDTA, 10% glycérol, 5mM 2-Mercapthoethanol, Inhibiteur de protéase
- *IgG Wash buffer II* : 50mM Phosphate pH7.5, 100mM sulfate ammonium, 1mM EDTA, 10% glycérol, 5mM 2-Mercapthoethanol, Inhibiteur de protéase
- *TEV digestion buffer* : 50mM Phosphate pH7.5, 100mM acétate de potassium, 1mM EDTA, 10% glycérol, 2-Mercapthoethanol
- *Ni-NTA binding buffer* : 50mM Phosphate pH7.5, 0,01% NP40, 10% glycérol, 500mM acétate de potassium, 10mM imidazole, Inhibiteur de protéase
- *Ni-NTA wash buffer* : 50mM Phosphate pH7.5, 0,01% NP40, 10% glycérol, 500mM acétate de potassium, 40mM imidazole, Inhibiteur de protéase
- *Ni-NTA elution buffer* : 50mM Phosphate pH7.5, 0,01% NP40, 10% glycérol, 500mM acétate de potassium, 200mM imidazole, Inhibiteur de protéase

Chapitre II : Analyse de la formation du PIC *in vitro*

I) Principe.

Durant mon projet de thèse, nous avons souhaité caractériser les effets du mutant *med10-196* sur l'assemblage du PIC. Nous avons utilisé des approches *in vivo* en analysant par ChIP l'occupation des différents GTFs. De plus, pour la première fois dans notre laboratoire, nous avons utilisé des approches *in vitro* pour caractériser le mutant *med10-196*. Nous avons utilisé une méthode développée et publiée par le laboratoire de S. Hahn (Ranish, Yudkovsky, and Hahn 1999). Cette méthode repose sur l'assemblage d'un PIC sur une matrice d'ADN immobilisée (Figure 45). Cette matrice d'ADN est composée d'un promoteur, d'un site de liaison d'un activateur (Gal4) et des premières paires de base du gène *HIS4*. En amont du promoteur, l'extrémité 5' du brin d'ADN est biotinyllé. Cela permet de lier la molécule d'ADN sur des billes magnétiques couplées à la streptavidine, une protéine se liant avec la biotine avec une forte affinité. On peut alors incuber cette matrice avec un extrait nucléaire de levure, provenant d'une souche sauvage ou mutante, en l'absence ou en présence d'un activateur purifié. Cela permet l'assemblage du PIC. A proximité de l'extrémité de l'ADN liée aux billes, se trouve un site de restriction (*Pst*I) permettant, par digestion, l'éluion du PIC formé. Les protéines peuvent ainsi être isolées et analysées par western blot ou spectrométrie de masse.

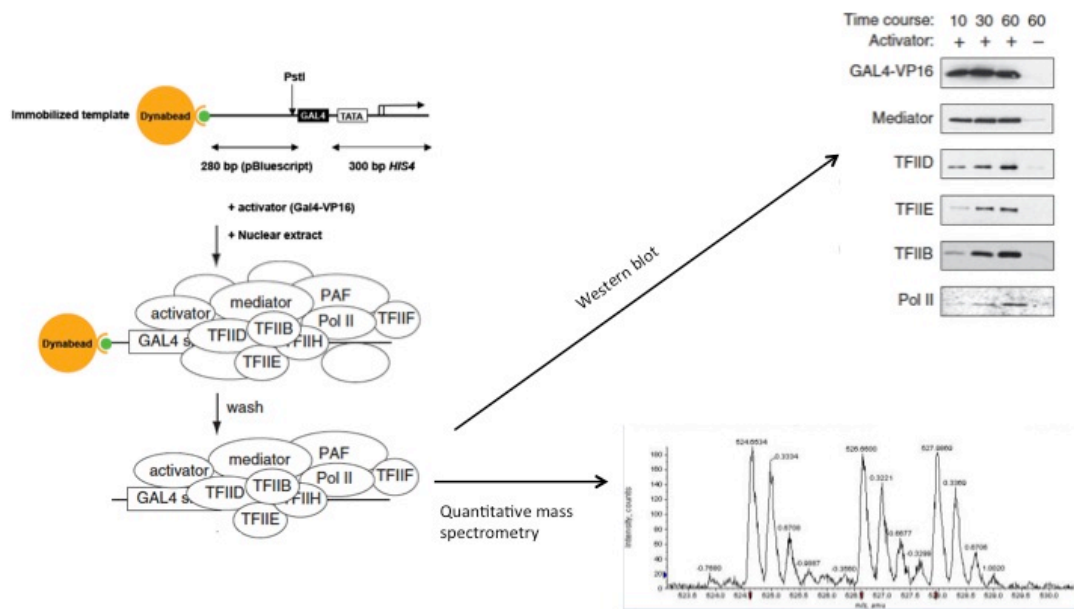


Figure 45 : Schéma de principe de l'expérience d'assemblage du PIC *in vitro*.

II) Protocole.

1) Purification des activateurs.

L'expérience d'assemblage du PIC *in vitro* est utilisée pour d'étudier la transcription basale en permettant la formation du PIC sur le promoteur de la matrice d'ADN immobilisé. Elle permet également d'étudier la transcription activée grâce au site de liaison pour un activateur. Sur la matrice d'ADN que nous utilisons, se trouve un site de liaison pour l'activateur Gal4. Nous disposons de la possibilité d'utiliser trois activateurs différents, tous les trois étant basés sur la fusion du domaine de liaison à l'ADN de Gal4. Ce domaine de liaison à l'ADN de Gal4 est fusionné soit à VP16 (Sadowski et al. 1988), un domaine d'activation issu d'une protéine virale, soit au domaine d'activation de Gcn4, un activateur naturel de *S. cerevisiae* ou fusionné à AH, un domaine d'activation synthétique (Wang, Gralla, and Carey 1992). Nous avons obtenu de la part de S. Hahn les plasmides dans lesquels ces constructions sont clonées en vue de purifier les protéines recombinantes dans *Escherichia coli*. Nous disposons ainsi du plasmide pET21A contenant la construction 6xHis-Gal4(1-147)-Gcn4(1-150), du plasmide pJL2 contenant la construction Gal4(1-147)-VP16 et du plasmide pJL2 contenant la construction Gal4(1-147)-AH.

Ces activateurs ont été purifiés selon les méthodes et protocoles développés par S. Hahn (Ranish, Yudkovsky, and Hahn 1999; Reeves and Hahn 2003).

2) Préparation de la matrice d'ADN.

La matrice d'ADN utilisée dans l'expérience de formation de PIC *in vitro* est un fragment d'ADN contenant un promoteur, le site de liaison de Gal4 ainsi que les premières bases du gène *HIS4* (Figure 45). Cette matrice est préparée par une réaction de PCR utilisant le plasmide pSH515 contenant les 144pb du promoteur cœur du gène *HIS4* et le site de liaison de Gal4 en amont. L'oligonucléotide utilisé en 5' est nommé p965 et sa séquence est Btn-TAATGCAGCTGGCAGCAGG, Btn signifiant biotinylé. L'oligonucléotide en 3' peut lui être variable, afin de déterminer la taille de la matrice que nous souhaitons utiliser. Dans notre cas, nous avons utilisé l'oligonucléotide pNot dont la séquence est GGCCGCTCTAGCTGCATTAATG. Cette amorce nous permet d'obtenir la matrice contenant seulement le promoteur entier, sans les premières bases du gène *HIS4*.

3) Préparation des extraits nucléaires de levure.

Les extraits nucléaires utilisés dans nos expériences ont été préparés selon le protocole publié par S. Hahn (Ranish, Yudkovsky, and Hahn 1999). Nous avons utilisé 4L de culture de levure en milieu YPD 3% de glucose. Les cultures sont réalisées dans 4 fois un litre dans des Fernbach.

Au jour 0, ensemercer chaque Fernach à une densité optique à 600nm de 0.01 à partir d'une préculture de 50mL. Mettre en culture à 30°C sur la nuit.

Au jour 1, mesurer la densité optique, la culture doit être en phase exponentielle (DO entre

0.9-1.2).

- Préparation de l'extrait nucléaire :

La préparation de l'extrait se déroule sur deux jours. Toutes les centrifugations sont réalisées à 4°C.

Jour1 :

- Centrifuger les cellules dans des pots de centrifugation d'un litre pendant 10 minutes à 3000g.
- Eliminer le milieu de culture et peser le culot de cellules. La quantité de cellule attendue est d'environ 50 g pour 4 L de culture.
- Reprendre le culot de cellules dans 35mL de Tris pH7.5 à 50mM et 30mM DTT. Bien décoller le culot du pot à l'aide d'une pipette ou d'un grattoir à cellules.
- Transférer la suspension dans un pot de centrifugation de 500mL et incubé à 30°C pendant 15 minutes
- Pendant ce temps, préparer la solution de Zymolyase (Zymolyase 100T, Amsbio). Peser la quantité nécessaire pour une concentration de 6 mg/mL de Zymolyase dans du Tris pH7.5 50mM, 1mM DTT, 2X d'inhibiteur de protéases.
- Centrifuger l'incubation de cellules à 3000g pendant 10 minutes.
- Resuspendre le culot dans 20mL de YPD/S
- Ajouter la Zymolyase. Utiliser 18mg (3mL)
- Incuber à 30°C et vérifier la formation de sphéroplastes toutes les 15 minutes. Pour cela, dans des cuves de spectromètre, mettre 1 mL de SDS 1% et 10 µL de suspension de cellules. Mesurer la DO initiale à 600nm. Attendre jusqu'à atteindre une diminution de 80% de la DO initiale. Cela peut prendre entre 30 minutes et 2h. Si la formation de sphéroplastes est très lente, après une heure, rajouter 3mL supplémentaires de Zymolyase 100T. Attention, des cellules dont la croissance a été trop haute peuvent ne jamais donner de sphéroplastes.
- Après l'obtention de 80% de sphéroplastes, ajouter 100mL de YPD/S à température ambiante et centrifuger 15 minutes à 3000g. Jeter le surnageant.
- Resuspendre le culot dans 250 mL de YPD/S à température ambiante. Incuber pendant 30 minutes à 30°C.
- Centrifuger à 3000g pendant 12 minutes . Jeter le surnageant.
- Resuspendre dans 200 mL de YPD/S froid (4°C).
- Centrifuger à 3000g pendant 12 minutes . Jeter le surnageant.
- Resuspendre dans 200 mL de YPD/S froid (4°C).
- Centrifuger à 3000g pendant 12 minutes . Jeter le surnageant.
- Resuspendre dans 250 mL de Sorbitol à 1M (4°C).
- Centrifuger à 3000g pendant 12 minutes . Jeter le surnageant.
- Resuspendre dans 100 mL de tampon A à 4°C.
- Transférer dans un homogénéiseur Dounce. Faire trois allers-retours de piston.
- Transférer le lysat dans un pot de centrifugation de 250 mL et centrifuger à 6300g pendant 8 minutes. Récupérer le surnageant et le transférer dans un pot de 250 mL propre. Répéter 4 fois.
- Après la 4^{ème} centrifugation, récupérer le surnageant et le transférer dans un pot de 250 mL propre. Centrifuger à 25300 g pendant 30 minutes. Cela permet de récupérer les noyaux dans le culot.
- Eliminer le surnageant et resuspendre les noyaux dans 10 mL de tampon B. Congeler dans de l'azote liquide et conserver à -80°C.

Jour 2 :

- Décongeler la suspension de noyaux dans la glace. Mesurer le volume et ajouter la quantité nécessaire de sulfate d'ammonium 3M pour une concentration finale de 0.5M. Mélanger et placer en chambre froide sur une roue pendant 30 minutes.
 - Transférer l'incubation dans un tube de centrifugation pour rotor 50.2 Ti (Beckman-Coulter). Centrifuger à 140000g pendant 90 minutes.
 - Récupérer le surnageant avec précaution en évitant le culot et transférer dans un tube 50.2 Ti propre. Ajouter un petit barreau magnétique.
 - Ajouter 0.35g/mL de sulfate d'ammonium. Incuber 30 minutes sous agitation douce en chambre froide.
 - Retirer le barreau magnétique et centrifuger à 18000g pendant 20 minutes.
 - Jeter le surnageant et centrifuger 4 minutes à 18000g. Jeter tout reste de surnageant.
 - Resuspendre le culot dans du tampon C. Le volume à ajouter est entre 400 µL et 1.5 mL.
 - Dialyser l'extrait contre du tampon C supplémenté de 75mM de sulfate d'ammonium pendant 5 heures au minimum.
 - Aliquoter l'extrait et doser la quantité de protéines. La concentration attendue se situe entre 25 et 50 mg/mL.
-
- Tampons et solutions pour la préparation des extraits nucléaires (Les volumes sont donnés pour 3 extraits).
-
- 50 mM Tris, 7.5 250 mL : 1.5 g Tris dans 250 mL H₂O, ajuster le pH à 7.5 avec du HCl.
 - YPD 3% par litre : 10 g Yeast extract, 20 g Bactopeptone, 30 g Glucose. Ajouter un litre d'eau. Autoclaver.
 - YPD/S 2.5 litres, conservé à température ambiante (YPD with 1M sorbitol) : 25g Yeast extract, 50g Bactopeptone, 50g glucose, 455g Sorbitol, ajouter 2.5 litres d'eau au total. Filtrer stérilement.
 - YPD/S 1.5 litres, conservé à 4°C : 15g Yeast extract, 30g Bactopeptone, 30g Dextrose, 273g Sorbitol, ajouter 1.5 litres d'eau au total. Filtrer stérilement.
 - 1M Sorbitol conservé à 4°C : 273g Sorbitol, ajouter 1.5 litres d'eau au total. Filtrer stérilement

- Tampon A, à filtrer et conserver à 4°C. Pour 650mL

18% Ficoll 400 (or polysucrose 400)	117g Ficoll
10 mM Tris 7.5	6.5 ml 1M Tris 7.5
20 mM K Acetate	13 ml 1M K Acetate
5 mM Mg Acetate	3.25 ml 1M Mg Acetate
1 mM EDTA	2.6 ml 0.25M EDTA
0.5 mM Spermidine	82 mg Spermidine
0.15 mM Spermine	61 µL 1.6M Spermine

3 mM DTT et les inhibiteurs de protéase sont ajoutés juste avant utilisation.

- Tampon B, à filtrer et à conserver à 4°C, Pour 250mL

100 mM Tris acetate pH 7.9	3g Tris
50 mM potassium acetate	12.5 ml 1M K Acetate
10 mM MgSO4	2.5 ml 1M MgSO4
20% glycerol	50 ml glycerol
2 mM EDTA	2 ml 0.25M EDTA

3 mM DTT et les inhibiteurs de protéase sont ajoutés juste avant utilisation.

- Tampon C, à filtrer et à conserver à 4°C, Pour 250mL.

20 mM HEPES 7.6	1.19g HEPES
10 mM MgSO4	2.5 ml 1M MgSO4
1 mM EGTA	1 ml 0.25 M EGTA
20% glycerol	50 ml glycerol

Ajuster le pH avec KOH.

3 mM DTT et les inhibiteurs de protéase sont ajoutés juste avant utilisation.

- Tampon C+ 75mM de sulfate d'ammonium, à filtrer et à conserver à 4°C, Pour 1.5L

20 mM HEPES 7.6	7.14g HEPES
10 mM MgSO4	15 ml 1M MgSO4
1 mM EGTA	6 ml 0.25 M EGTA
20% glycerol	300 ml glycerol
75 mM Ammonium Sulfate	14.8g AmSO4

Ajuster le pH avec KOH

3 mM DTT et les inhibiteurs de protéase sont ajoutés juste avant utilisation

4) Formation du PIC *in vitro*.

La réaction de formation du PIC *in vitro* est réalisée selon le protocole mis au point par le laboratoire de S. Hahn (Ranish, Yudkovsky, and Hahn 1999). Généralement, nous réalisons l'expérience selon trois conditions différentes. L'une sans ADN, comme contrôle négatif pour tenir compte des liaisons non-spécifiques des protéines sur les billes, une autre avec l'ADN mais sans présence d'activateur, comme contrôle du niveau basal de formation du PIC et la dernière avec l'ADN et en présence d'activateur pour la condition de transcription activée.

- Préparation des billes magnétiques :

Les billes utilisées (Dynabeads® M-280 Streptavidin, ThermoScientific) sont des billes magnétiques couvertes de streptavidine, ayant une forte affinité pour la biotine. L'incubation des billes avec la matrice d'ADN biotinylée permet la liaison de celle-ci sur les billes.

- Prélever 5 µL de billes par réaction et placer dans un tube Eppendorff de 2 mL.
- Laver les billes deux fois avec 400 µL de TE+1M de NaCl.
- Resuspendre les billes à une concentration de 10mg/mL dans du TE+ 1M NaCl + 0.003% NP-40.
- Ajouter 8.6ng d'ADN par µg de billes.
- Incuber à 21°C pendant 30 minutes dans un ThermoMixer Eppendorf avec une agitation de 30 secondes toutes les 2 minutes à 600 rpm.
- Laver les billes une fois dans du TE + 1M NaCl.
- Resuspendre les billes dans du tampon de blocage (1 µL/µg de billes).
- Incuber 15 minutes à température ambiante dans un ThermoMixer Eppendorf avec une agitation de 30 secondes toutes les 2 minutes à 600 rpm.

- Laver les billes trois fois dans 400 μL de tampon de transcription.
- Resuspendre les billes dans du tampon de transcription à une concentration de 10 mg/mL.
- Utiliser les billes dans la journée et les garder dans la glace jusqu'à utilisation.

- Réaction de formation du PIC :

Les réactions de formation de PIC sont réalisées dans un volume final de 100 μL .

- Par réaction, préparer 40 μL de milieu réactionnel contenant :
 - 20 μL de tampon de transcription 5X
 - 12 μL de phosphocréatine à 12mM
 - 2.52 μL de DTT à 0.1M
 - 400ng de créatine phosphokinase, 0.4 μL à 2 $\mu\text{g}/\mu\text{L}$
 - 0.5 μL de NP-40 à 10%
 - 4.58 μL d'eau distillée.
- Par réaction, ajouter 360 μg d'extrait nucléaire dans les 40 μL de milieu réactionnel, compléter avec du tampon de transcription 1x pour obtenir un volume de 100 μL après l'ajout du volume de billes+activateur.
- Incuber pendant 10 minutes à température ambiante.
- Centrifuger les réaction à 7500xg pendant 2 minutes, à 4°C. transférer les surnageants dans un nouveau tube. Ajouter toute protéines recombinantes à cette étape.
- Ajouter 1 μg d'ADN génomique de *E. coli* à chaque réaction
- Ajouter 120 ng de Gal4-AH, ou 96 ng de Gal4-VP16 ou Gal4-Gcn4 pour 5 μL de billes préparées avec l'ADN. Incuber 10 minutes à température ambiante en agitant ponctuellement pour éviter que les billes ne sédimentent.
- A chaque mix réactionnel+extrait nucléaire, ajouter 5 μL de bille + le volume d'activateur ajouté.
- Incuber les réaction pendant 40 minutes dans un ThermoMixer Eppendorf avec une agitation de 30 secondes toutes les 2 minutes à 600 rpm.
- Laver trois fois les réactions avec 400 μL de tampon de lavage.
- Eluer les échantillons par digestion avec l'enzyme PstI, dans le tampon prévu pour l'enzyme. Utiliser 60 U par réaction. Incuber 30 minutes à 37°C. Note : utiliser l'enzyme Invitrogen, dont la teneur en BSA est plus réduite que celle de certains fournisseur. La BSA peut être gênante dans la détection des protéines et pour la spectrométrie de masse. Essayer de réaliser l'éluion dans le volume le plus petit possible (typiquement 30 μL), pour concentrer les échantillons.
- Récupérer le surnageant. Ajouter le volume nécessaire de bleu de charge pour SDS-PAGE. Analyser par western blot et conserver le reste à -20°C.

Tampons utilisés.

Tampon de blocage :

- 1X Tampon de transcription (pH 7.6)
 - 60mg/ml caséine
 - 5mg/ml poly vinyl pyrrolidone
 - 2.5mM DTT
- Equilibrer le pH à 8 avec KOH
Garder à -20° C en aliquots de 1mL

5x Acetate transcription buffer

- 500 mM Acetate de Potassium, pH 7.6
 - 100 mM HEPES, pH 7.6
 - 5 mM EDTA
 - 25 mM MgOAc
- Garder à -20°C

Tampon de lavage :

- 1x Tampon de transcription (pH 7.6)
 - 0.05% NP-40
 - 2.5mM DTT
- Préparer juste avant utilisation

ADN génomique de *E. coli* digéré par HaeIII :

- Préparer de l'ADN génomique de *E. coli* à partir d'une souche non transformée et 50mL de culture (extraction au phénol-chloroforme). Digérer l'ADN génomique avec l'enzyme HaeIII sur la nuit. Purifier au phénol-chloroforme et précipiter à l'éthanol. Reprendre dans du TE à une concentration d'environ 1mg/mL. Analyser sur gel d'agarose et stocker à -20°C.

Creatine phospho kinase (sigma C-3755 165 units/mg) :

- 2.0 mg/ml in HA + 0.1 M d'acétate de potassium
- Stocker en aliquots de 12 µL à -80°C.

Phospho creatine (64 mg/mL) :

- 128 mg phospho creatine
 - 2 mL TE pH 7.5
 - Ajouter 4 µL d'acide acétique 1M.
 - Vérifier le pH avec du papier pH. Doit être d'environ 7.5
- Aliquoter et stocker à -80°C

Chapitre III : Analyses bioinformatiques des données de ChIP-seq.

I) Principe du ChIP-seq.

L'immunoprécipitation de chromatine (ChIP) est une méthode très utilisée pour l'analyse de la liaison d'une protéine à l'ADN. Elle consiste en l'enrichissement, par immunoprécipitation d'une protéine, des fragments d'ADN liés à cette protéine. Pour cela, les cellules sont traitées par du formaldéhyde, une molécule qui permet d'établir des pontages chimiques covalents (*crosslink*) entre les protéines et les acides nucléiques. La chromatine est ensuite préparée, puis fragmentée, généralement par l'utilisation d'ultrasons (sonication), en fragments de 1kb à 200pb. On réalise une immunoprécipitation de la protéine d'intérêt liée à l'ADN à l'aide d'un anticorps. Après différents lavages, on réverte le *crosslink* par chauffage à 65°C et on digère les protéines à l'aide de protéases. L'ADN est ensuite purifié et analysé. On peut analyser l'ADN enrichi de façon ciblée (par PCR et PCR quantitative) ou de façon globale par hybridation sur une puce à ADN (ChIP-on-chip) ou séquençage à haut-débit (ChIP-seq).

Dans ce travail de thèse, nous avons utilisé le ChIP-seq pour analyser et comparer l'occupation des différents composants du PIC au niveau génomique, entre une souche sauvage et un souche mutante. Après avoir réalisé des ChIP dirigés contre l'ensemble des GTFs, la Pol II et deux sous-unités du Médiateur, l'ADN a été purifié et séquençé par la plateforme de séquençage à haut-débit de l'institut de biologie intégrative de la cellule (i2BC) de Gif sur Yvette. Les méthodes de séquençage à haut-débit, ou de nouvelle génération (*Next Generation Sequencing*, NGS), ont subies de développements technologiques très importants ces dernières années. Elles sont devenues très puissantes pour l'analyse profonde de l'ADN et de l'ARN. Elles génèrent cependant un volume de données très grand, dont les analyses requièrent des compétences spécifiques de bioinformatique. Cette partie présentera quelques aspects des analyses de données de ChIP-seq réalisées au cours de mon projet de thèse. La majorité des analyses a été faite sous Unix (système CentOS, Environnement Gnome) directement en lignes de commande Bash. Les représentations graphiques ont été effectuées dans le logiciel R version 2.14.0. Dans la suite, différents formats de fichier seront évoqués. La description des types de fichiers et des données qu'ils contiennent peut être trouvée sur le site de l'université de Californie Santa Cruz à l'adresse suivante <http://genome.ucsc.edu/FAQ/FAQformat>.

II) Traitement des données issues du séquençage.

Le résultat donné par le séquenceur consiste en des lectures (*reads*) qui renferment la séquence du fragment séquençé. Ces lectures sont contenues dans un fichier de format .fastq. Ce type de fichier, en plus de la séquence des lectures, contient des scores de qualités associés à

chacunes d'elles. Dans notre analyse réalisée au cours de ma thèse, nous avons suivis la démarche d'analyse des données de CHIP-seq mise en place dans l'étude des mutants de Med17 (Eyboulet et al. 2015). La première étape est de filtrer les fichiers .fastq afin de supprimer les adaptateurs à l'aide du logiciel Cutadapt version 1.3 (Martin 2011). Ensuite, nous réalisons une analyse de la qualité des lectures. Cela est fait par l'utilisation du logiciel fastQC (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc>). Ce programme permet d'analyser les séquences d'un fichier .fastq selon différents critères. Son utilisation met en évidence des problèmes ayant pu apparaître au cours du séquençage, comme, par exemple, une sur-amplification d'une séquence particulière. Cela nous permet de nous assurer de la qualité des données avec lesquelles nous travaillons.

Suite à cela, nous cherchons à identifier, pour chaque lecture issue du séquençage, la ou les localisations sur le génome de référence. On parle d'alignement. Plusieurs logiciels existent, pour faire cet alignement. Nous avons choisi d'utiliser Bowtie version 0.12.7 (Langmead et al. 2009), un outil d'alignement de lectures de petites tailles (quelques dizaines de nucléotides) sur de larges génomes. L'alignement des lectures est réalisé suivant une combinaison de différentes options permettant de déterminer quels alignements sont autorisés et doivent être conservés. Dans notre étude, nous avons utilisé la commande suivante avec les paramètres :

```
bowtie -a --best --strata -v2 -m1 /genomes/SacCer/Bowtie/SacCer3 Fichier.fastq >
Alignement.sam
```

-a : Bowtie reporte tous les alignements valides.

--best – strata : Bowtie reporte le meilleur alignement satisfaisant les critères imposés.

-v2 : Mode « v » pour ne pas tenir compte des valeurs de qualité des lectures, valeur « 2 » pour autoriser jusqu'à 2 discordances.

-m1 : Seules les lectures avec un positionnement unique sur le génome sont retenues.

/genomes/SacCer/Bowtie/SacCer3 : Chemin d'accès au génome de référence, ici de *S. cerevisiae* v.3 (avril 2011), indexé pour Bowtie.

En résumé, la combinaison de ces différents paramètres permet de reporter les lectures s'alignant de manière unique, avec 0, 1 ou 2 mésappariements sur le génome de *S. cerevisiae*. Le meilleur alignement, contenant le plus faible nombre de mésappariement sera retenu. A la sortie de ce traitement, on obtient un fichier au format .sam. Ce type de fichier contient la séquence de la lecture associée à sa localisation sur le génome. C'est un format très volumineux. Nous avons décidé d'utiliser le logiciel SAMtools version 0.1.16 (Li et al. 2009) pour convertir les fichiers .sam en fichiers .bam. Les fichiers .bam renferment la même information mais de façon compressée.

L'intérêt de la méthode de CHIP-seq est de connaître la localisation de la liaison d'une protéine sur la chromatine mais également le niveau d'occupation de cette protéine. Le niveau d'occupation est directement lié à la quantité d'ADN et donc au nombre de lecture que l'on retrouve sur une région particulière du génome. Cette étape est réalisée par la création d'un fichier de densité WIG ou BEDGraph qui décompte les lectures sur des intervalles définis. Pour cela, nous avons utilisé le logiciel BedTools version 2.15.0 (Quinlan and Hall 2010). Brièvement, nous avons tout d'abord utilisé ce logiciel pour convertir les fichiers .bam en fichiers .bed. Ce format renferme les coordonnées d'une lecture sur le génome, sans la séquence de celle-ci. Les fichiers .bed sont ensuite utilisés par Bedtools pour générer les fichiers de densité à l'aide de la fonction CoverageBed. En sortie, nous obtenons un fichier de format BEDgraph. Les fichiers .WIG et BEDGraph sont très volumineux. Afin de fluidifier la visualisation dans l'explorateur de génome, ces fichiers sont convertis au format bigWIG (.bw) à l'aide de BedTools.

L'objectif de nos travaux est de comparer les occupations des composants du PIC entre des

souches sauvages et mutantes. Le nombre total de lectures alignées n'est pas le même d'un échantillon à l'autre, ce qui peut amener à considérer un enrichissement sur les gènes de certains échantillons simplement parce qu'ils contiennent globalement plus de lectures que les autres. De plus, des biais peuvent subvenir lors des différents traitements subis par les échantillons lors de leur préparation. Ainsi, pour pouvoir effectuer une comparaison quantitative il est nécessaire de normaliser les données. Il existe à l'heure actuelle certaines méthodes de normalisation des données de séquençage. Néanmoins, celles-ci ne sont pas compatibles si l'effet des mutations aboutit à une diminution globale de l'occupation d'une des protéines étudiées. Le rôle central du Médiateur sur la transcription laisse facilement imaginer que les effets de mutations de ses sous-unités pourraient avoir des effets globaux. Partant de cet *a priori*, nous avons décidé de nous baser sur les valeurs de qPCR pour comparer nos données. Pour cela, pour chaque protéine, nous avons sélectionné une dizaine de régions (ORF ou promoteurs), sur lesquelles nous avons mesuré les signaux d'occupation par qPCR. Nous avons ensuite calculé le rapport R_{qpcr} de niveau d'occupation entre la souche mutante ($V_{qpcrMutant}$) et la souche sauvage (V_{qpcrWT}) dans ces différentes régions ($R_{qpcr} = V_{qpcrMutant} / V_{qpcrWT}$). Nous avons analysé la densité de lecture sur ces mêmes régions dans les données de CHIP-seq. Cela nous a permis de calculer le rapport $R_{CHIP-seq}$ entre les niveaux d'occupation dans le mutant et le sauvage ($R_{CHIP-seq} = V_{CHIP-seq-Mutant} / V_{CHIP-seq-WT}$). Nous avons ensuite calculé le rapport Q entre les rapports R_{qpcr} et $R_{CHIP-seq}$, ($Q = R_{qpcr} / R_{CHIP-seq}$). Nous avons déterminé la médiane de ce rapport Q pour les 10 régions que nous avons choisi d'analyser. Cette valeur correspond au facteur de normalisation que nous avons appliqué sur les valeurs de densité obtenues dans les données pour la souche mutante.

<i>Protéine</i>	<i>Facteur de normalisation</i>
<i>Pol II</i>	0.45
<i>Med15</i>	0.46
<i>Med17</i>	0.45
<i>TFIIB</i>	0.57
<i>Toa2</i>	0.97
<i>Tfg1</i>	1.2
<i>Tfa2</i>	0.66
<i>Rad3</i>	0.42
<i>Kin28</i>	0.48
<i>TBP</i>	1.93
<i>Taf1</i>	1.04

Tableau 14: Tableau récapitulatif des facteurs de normalisation appliqués à nos données de CHIP-seq dans l'étude du mutant *med10-196*.

Après avoir appliqué ces facteurs de normalisation, nous obtenons un fichier de densité au format BigWIG pour chaque protéine étudiée, pour la souche sauvage et mutante.

L'étape suivante a été la définition des régions à analyser. Nous sommes partis du principe que les gènes étant enrichis pour la Pol II étaient des gènes actifs. Nous souhaitions analyser la formation du PIC sur ces gènes. Nous avons donc décidé de déterminer quels étaient les gènes enrichis pour la présence de Pol II. Cette détermination est basée sur une analyse de *Peak-calling*.

Cette analyse permet de déterminer si la densité de lectures d'une région particulière dans un échantillon est significativement enrichie par rapport à un contrôle. Dans notre cas, le contrôle est l'ADN total (ou input). Si cela est le cas, on considère que la région est enrichie et que la protéine est liée. Pour cette analyse nous avons utilisé le logiciel MACS2 (*Model-based Analysis of ChIP-Seq*) version 2.0.10 (<https://github.com/taoliu/MACS>). Cette analyse par MACS2 nous a permis de définir 5062 régions enrichies pour la Pol II avec un seuil d'erreur (*False discovery rate, FDR*) de 5%. Afin de pouvoir interpréter biologiquement les résultats de ChIP-seq, les pics identifiés doivent être annotés, c'est à dire qu'on doit pouvoir leur attribuer le gène ou la région promotrice au sein de laquelle ils se trouvent. De plus, on constate sur l'exemple présenté en Figure 46 que plusieurs pics peuvent être détectés au sein d'une même région codante. En utilisant l'outil BEDtools avec sa fonction *intersect*, nous avons pu définir à quel gène correspondent les différents pics et attribuer un gène à plusieurs pics se trouvant dans une région codante.

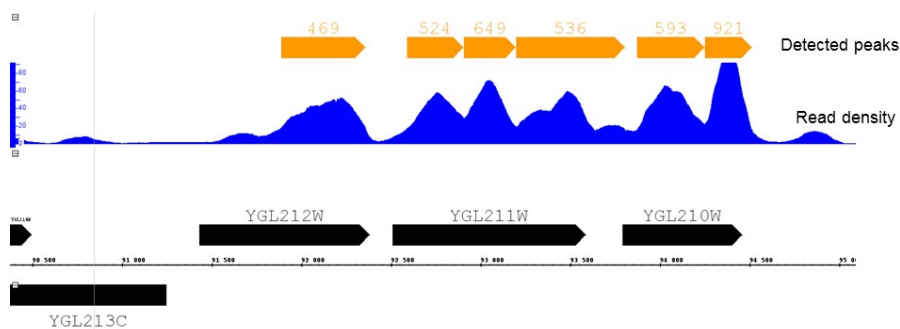


Figure 46 : Visualisation sur un explorateur de génome des signaux de ChIP-seq de la Pol II et des pics détectés par l'analyse de Peak calling

La commande utilisée est :

```
intersectBed -wa -a RégionsCodantes_S.cerevisiae.bed -b PeaksPolIIDétectés.bed | sort -u > GenesEnrichisEnPolII.bed
```

-wa : Pour chaque chevauchement, on garde la région contenue dans le fichier A.

-a : Fichier A au format BED, à comparer au fichier B.

-b : Fichier B au format BED.

sort -u : Permet de supprimer les doublons.

Le fichier d'annotation a été obtenu par les données recueillies depuis la base de donnée SGD (*Saccharomyces genome database*).

http://downloads.yeastgenome.org/curation/chromosomal_feature/saccharomyces_cerevisiae.gff

Cette commande permet de détecter les intersections existantes entre deux fichiers au format bed. Dans notre cas, on compare les pics de Pol II détectés par MACS2 avec la liste des régions codantes de *S cerevisiae*. L'ajout de l'option **-wa** permet d'obtenir en sortie, le fichier correspondant aux régions codantes issues du fichier d'annotation dans lesquelles un pic est chevauchant. Cela nous a permis de définir 3303 régions codantes significativement enrichies en Pol II et de définir les régions sur lesquelles nous allons analyser les densités. Les autres protéines analysées sont des protéines du PIC, dont la liaison se situe sur le promoteur. Nous avons ainsi défini un fichier regroupant les 2694 régions promotrices associées au 3303 gènes.

Pour pouvoir comparer et quantifier les niveaux d'occupation des différentes protéines étudiées, il faut pouvoir extraire les données de densités sur des régions d'intérêt. Pour cela, nous avons utilisé un script développé par Olivier Alibert de l'IRCM d'Evry.

Ce script se nomme get_tag_density.

La commande utilisée est :

```
get_tag_density -s -f IP-Prot.bw Fichier_d'annotation_des_régions_d'intérêt.bed >
densité_de_lecture.tsv
```

-s : Tiens compte du brin pour orienter les données de densité.

-f : Fichier de densité de la protéine d'intérêt.

TSS.bed : Fichier BED contenant les coordonnées des régions pour lesquelles on souhaite récupérer les données de densité.

En sortie de cette commande, on récupère un tableau de données contenant pour chaque région spécifiée dans le fichier d'annotation :

```
chr1 170911 172511 YAR018C 0 -
442;444;447;450;451;451;451;451;451;451;451;451;451;451;451;
451;451;451;451;451;451;451;451;451;451;451;452;452;453;453;
453;453;454;454;454;456;457;458;458;458;460;460;461;461;461;
461;463;463;463;463;466;466;466;466;466;466;466;466;466;466;
466;466;466;466;466;466;466;466;466;466;466;466;466;466;
466;466;466;466;466;466;466;466;466;466;466;466;466;466;
466;466;466;466;466;466;466;466;469;469;468;469;469;469;
469;469;470;470;470;470;469;469;469;469;469;469;469;469;469;
468;467;467;467;467;464;462;460;460;456;450;445;441;435;431;
431;430;427;427;427;427;427;427;427;427;427;427;427;427;
425;419;419;413;411;408;408;408;408;407;407;407;406;406;405;...
449.98 56847 372 470, 171301;171302;171303... 128486
```

chr1 170911 172511 YAR018C - : Informations extraites du fichiers BED indiquant les coordonnées de la région.

442;444;447... : Densité de la protéine par position sur l'ensemble de la région définie en entrée

449.98 : moyenne des densités sur la région.

56847 : somme des densités sur la région.

372 470 : Densité minimale et maximale.

171301;171302;171303... : Positions génomiques où la densité est maximale.

128486 : Moyenne des positions de densité maximale.

Ainsi, pour chaque protéines étudiées, nous avons extrait de cette façon les densités moyennes (colone 7) de lectures. Pour la Pol II, nous avons récupéré les densités sur les régions codantes des gènes enrichis, pour les GTFs, sur les régions intergéniques.

III) Analyse des données.

Les fichiers obtenus suite à l'extraction des densités sont le support des analyses que nous réalisons.

1) Analyse des effets globaux des mutations du Médiateur sur la formation du PIC.

Afin d'investiguer les effets des mutations sur la formation du PIC de façon globale, nous analysons, à l'aide du logiciel R, la corrélation qui peut exister entre les occupations des différentes protéines étudiées dans les contextes sauvage et mutant. Pour cela, nous traçons le nuage de point des densités moyennes de lectures, sur les régions définies préalablement, dans la souche sauvage *versus* la souche mutante. Nous calculons ensuite la regression linéaire existante au sein de ce nuage de point. La comparaison de la pente de la droite de regression avec la droite $y=x$ nous permet de montrer la tendance générale de l'effet de la mutation sur l'occupation du facteur considéré. La figure 47 montre un exemple de représentation graphique de cette analyse.

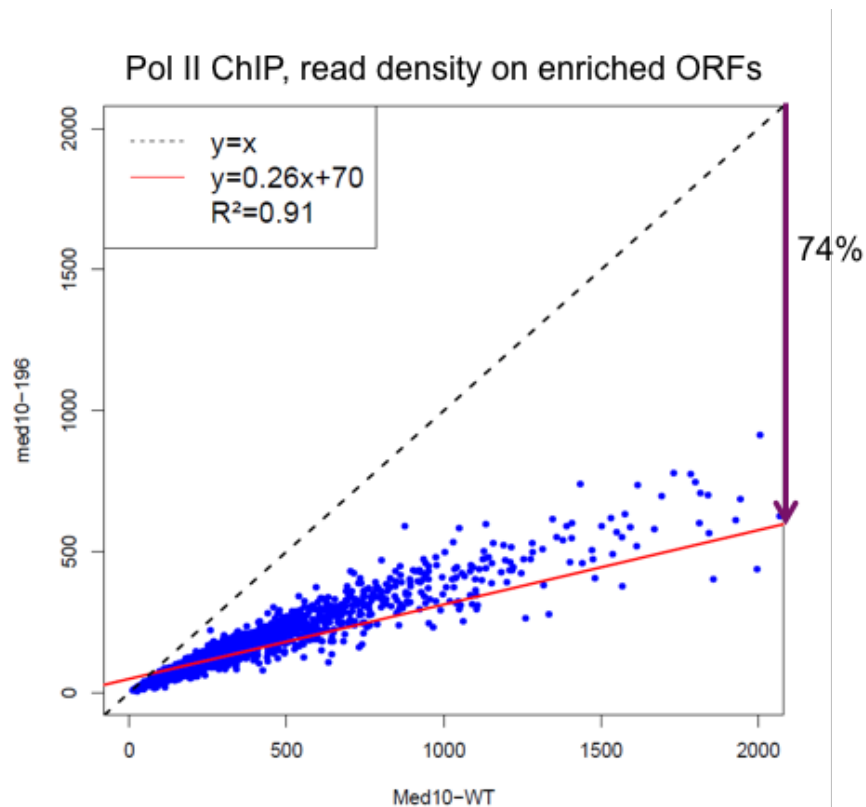


Figure 47 : Exemple de nuage de point représentant les densités de lecture sur les régions codante dans le cas d'un ChIP de la Pol II.

2) Analyse de l'effet gène-spécifique des mutations du Médiateur sur la formation du PIC : cas de l'architecture des promoteurs.

Dans nos travaux sur l'étude de la mutation *med10-196*, nous avons souhaité savoir si la mutation pouvait influencer de manière spécifique la formation du PIC. *In vivo*, la transcription se déroule dans un contexte chromatinien. De plus il est connu que la composition en terme de nucléosome et la présence de boîte TATA au sein d'un promoteur peut influencer la façon dont le PIC se forme. Pour cette raison, nous avons souhaité analyser l'effet de la mutation sur les gènes en prenant en compte la composition en terme d'éléments TATA, d'occupation en nucléosome et de dynamique des nucléosomes dans les promoteurs. Nous avons pour cela utilisé des données publiées. Pour l'occupation des nucléosomes et les éléments TATA, nous avons utilisé les données issues de l'analyse du laboratoire de F. Pugh (Rhee and Pugh, 2012). Concernant la dynamique des nucléosomes, nous avons utilisé les données du laboratoire d'O. Rando (Dion et al., 2007).

Afin d'investiguer un effet-gène spécifique potentiel, la première étape a consisté en un filtrage des gènes. En effet, dans le cas où des gènes sont divergents, la compacité du génome de *S. cerevisiae* et la résolution de nos expériences de CHIP-seq ne permet pas de déterminer la présence de deux PICs sur ces régions. Pour cette raison, nous avons exclu les régions intergéniques correspondant aux gènes divergents et n'avons conservés que les régions intergéniques des gènes en tandem. La TBP se liant également sur les promoteurs des gènes de classe III, dont les promoteurs peuvent être situés au sein de régions codantes ou à proximité des promoteurs de classe II, nous avons décidé d'exclure les gènes contenant ou se trouvant à proximité d'un gène de classe III. Enfin, pour s'affranchir de la variabilité due à des faibles signaux d'occupation, nous avons décidé de ne conserver que les gènes les plus exprimés. Pour cela, nous avons exclu le premier quartile des gènes sélectionnés, dont les signaux d'occupation de la Pol II sont les plus faibles. Après tout ces filtrages, nous avons obtenu un ensemble de 1065 gènes sur lesquels nous avons menés nos analyses.

Afin d'investiguer l'effet gène-spécifique, nous avons décidé, pour chaque facteur étudié, de définir trois groupes. Le groupe des 25% des gènes les plus affectés par la mutation (bas-quartile, 267 gènes), le groupe des 25% des gènes les moins affectés (haut-quartile, 267 gènes) et les 50% des gènes dont le comportement suit la tendance générale (interquartile, 531 gènes). Pour cela, nous avons d'abord appliqué une transformation logarithmique (\log_2) aux données d'occupation. La tendance de la régression entre les valeurs d'occupation entre le sauvage et le mutant n'étant pas toujours totalement linéaire, nous avons utilisé un script, développé par Pierre Nicolas de l'INRA de Jouy en Josas, permettant de déterminer un score à partir duquel les groupes sont définis. L'application de cette analyse pour chaque facteur nous a permis de définir pour chaque facteur, les trois groupes de gènes pour chaque facteur étudié (Figure 48).

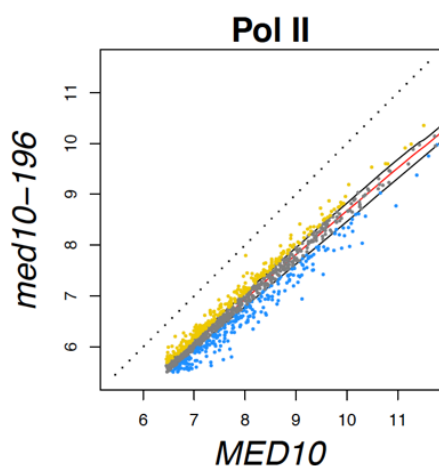


Figure 48 : Exemple de représentation logarithmique des occupations de la Pol II dans les souches sauvage et mutante de *Med10* et représentation des trois groupes de gènes définis.

Les gènes représentés en jaune correspondent au haut-quartile (25% des gènes les moins affectés par la mutation), les gènes représentés en bleu correspondent au bas-quartile (25% des gènes les plus affectés par la mutation) et les gènes représentés en gris correspondent à l'interquartile.

Une fois ces trois groupes définis pour l'ensemble des facteurs, nous avons analysé l'architecture du promoteur des gènes pour l'ensemble de ces groupes. Pour cela, nous avons analysé la présence ou l'absence de boîte-TATA canonique. Nous avons aussi défini les profils d'occupation des nucléosomes sur une fenêtre de 1600pb, centrée sur le TSS. Enfin, nous avons analysé la dynamique des nucléosomes +1 et -1 au sein des promoteur. Seuls les nucléosomes significativement dynamiques (*hot*) ont été inclus dans l'analyse. La figure 49 présente un exemple des résultats observés par cette analyse.

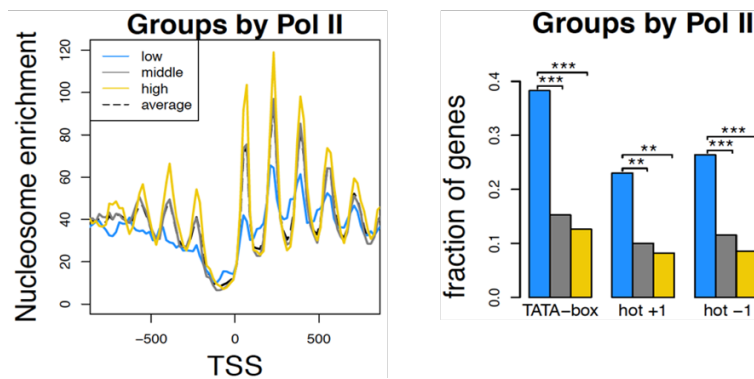


Figure 49 : Exemple de représentations graphiques des analyses de données de ChIP-seq en termes d'occupation des nucléosomes, de présence de boîtes TATA et de dynamique des nucléosomes pour les trois groupes de gènes définis pour la Pol II.

Les gènes représentés en jaune correspondent au haut-quartile (25% des gènes les moins affectés par la mutation), les gènes représentés en bleu correspondent au bas-quartile (25% des gènes les plus affectés par la mutation) et les gènes représentés en gris correspondent à l'interquartile.

Afin de déterminer si la significativité des différences observées, nous avons réalisé des tests statistiques. Dans le cas des données concernant les boîtes TATA et les nucléosomes *hot*, nous avons réalisé des tests de Fischer. Les différences sont considérées comme significatives lorsque la valeur-p (*p-value*) est inférieure à 0.05. Concernant l'occupation des nucléosomes, nous avons utilisé un test de Wilcoxon. Ce test a été réalisé en utilisant la valeur maximum d'occupation en nucléosomes à une distance de 100pb après le TSS. De la même manière, les différences sont considérées comme significatives lorsque la valeur-p (*p-value*) est inférieure à 0.05. Ces tests ont été réalisés à l'aide du logiciel R et sa fonction de base *stats*.

Enfin, afin de quantifier les relations et l'interdépendance entre les effets de la mutation *med10-196* sur les différents GTFs, nous avons réalisé des corrélations de Spearman entre les ratios de valeurs du mutant *versus* la souche sauvage pour tous les facteurs étudiés. Ces corrélations ont été réalisées à l'aide du logiciel R et sa fonction de base *stats*. Les données sont visualisées à l'aide de distributions par intensité (*heatmap*) produites à l'aide du logiciel R et de son extension *heatmap3* version 1.1.1.

Bibliographie

- Adelman, K, M. T. Marr, J. Werner, A. Saunders, Z. Ni, E. D. Andrulis, and J. T. Lis. 2005. "Efficient Release from Promoter-Proximal Stall Sites Requires Transcript Cleavage Factor TFIIIS." *Molecular Cell* 17 (1): 103–12. doi:10.1016/j.molcel.2004.11.028.
- Akhtar, M. S., M. Heidemann, J. R. Tietjen, D. W. Zhang, R. D. Chapman, D. Eick, and A. Z. Ansari. 2009. "TFIIH Kinase Places Bivalent Marks on the Carboxy-Terminal Domain of RNA Polymerase II." *Molecular Cell* 34 (3): 387–93. doi:10.1016/j.molcel.2009.04.016.
- Akoulitchev, S., S. Chuikov, and D. Reinberg. 2000. "TFIIH Is Negatively Regulated by cdk8-Containing Mediator Complexes." *Nature* 407 (6800): 102–6. doi:10.1038/35024111.
- Alberti, S., A. D. Gitler, and S. Lindquist. 2007. "A Suite of Gateway® Cloning Vectors for High-Throughput Genetic Analysis in *Saccharomyces Cerevisiae*." *Yeast (Chichester, England)* 24 (10): 913–19. doi:10.1002/yea.1502.
- Albright, S.R, and R. Tjian. 2000. "TAFs Revisited: More Data Reveal New Twists and Confirm Old Ideas." *Gene* 242 (1–2): 1–13. doi:10.1016/S0378-1119(99)00495-3.
- Alén, C., N. A. Kent, H. S. Jones, J. O’Sullivan, A. Aranda, and N. J. Proudfoot. 2002. "A Role for Chromatin Remodeling in Transcriptional Termination by RNA Polymerase II." *Molecular Cell* 10 (6): 1441–52.
- Allison, L. A., J. K. Wong, V. D. Fitzpatrick, M. Moyle, and C. J. Ingles. 1988. "The C-Terminal Domain of the Largest Subunit of RNA Polymerase II of *Saccharomyces Cerevisiae*, *Drosophila Melanogaster*, and Mammals: A Conserved Structure with an Essential Function." *Molecular and Cellular Biology* 8 (1): 321–29. doi:10.1128/MCB.8.1.321.
- Alpern, D., D. Langer, B. Ballester, S. Le Gras, C. Romier, G. Mengus, and I. Davidson. 2014. "TAF4, a Subunit of Transcription Factor II D, Directs Promoter Occupancy of Nuclear Receptor HNF4A during Post-Natal Hepatocyte Differentiation." *eLife* 3: e03613. doi:10.7554/eLife.03613.
- Amoasii, L., W. Holland, E. Sanchez-Ortiz, K.K. Baskin, M. Pearson, S. C. Burgess, B. R. Nelson, R. Bassel-Duby, and E. N. Olson. 2016. "A MED13-Dependent Skeletal Muscle Gene Program Controls Systemic Glucose Homeostasis and Hepatic Metabolism." *Genes & Development* 30 (4): 434–46. doi:10.1101/gad.273128.115.
- Andel, F., A. G. Ladurner, C. Inouye, R. Tjian, and E. Nogales. 1999. "Three-Dimensional Structure of the Human TFIIID-IIA-IIIB Complex." *Science* 286 (5447): 2153–56. doi:10.1126/science.286.5447.2153.
- Anderson, J. D., and J. Widom. 2001. "Poly(dA-dT) Promoter Elements Increase the Equilibrium Accessibility of Nucleosomal DNA Target Sites." *Molecular and Cellular Biology* 21 (11): 3830–39. doi:10.1128/MCB.21.11.3830-3839.2001.
- Andrau, J.-C., L. van de Pasch, P. Lijnzaad, T. Bijma, M. Groot Koerkamp, J. van de Peppel, M. Werner, and F. C. P. Holstege. 2006. "Genome-Wide Location of the Coactivator Mediator: Binding without Activation and Transient Cdk8 Interaction on DNA." *Molecular Cell* 22 (2): 179–92. doi:10.1016/j.molcel.2006.03.023.
- Ansari, S. A., Q. He, and R. H. Morse. 2009. "Mediator Complex Association with Constitutively Transcribed Genes in Yeast." *Proceedings of the National Academy of Sciences of the United States of America* 106 (39): 16734–39. doi:10.1073/pnas.0905103106.
- Ansari, S. A., and R. H. Morse. 2013. "Mechanisms of Mediator Complex Action in Transcriptional Activation." *Cellular and Molecular Life Sciences* 70 (15): 2743–56. doi:10.1007/s00018-013-1265-9.
- Archambault, J., M. A. Drebot, J. C. Stone, and J. D. Friesen. 1992. "Isolation and Phenotypic Analysis of Conditional-Lethal, Linker-Insertion Mutations in the Gene Encoding the Largest Subunit of RNA Polymerase II in *Saccharomyces Cerevisiae*." *Molecular & General Genetics: MGG* 232 (3): 408–14.
- Archambault, J., F. Lacroute, A. Ruet, and J. D. Friesen. 1992. "Genetic Interaction between Transcription Elongation Factor TFIIIS and RNA Polymerase II." *Molecular and Cellular Biology* 12 (9): 4142–52. doi:10.1128/MCB.12.9.4142.
- Armache, K.-J., H. Kettenberger, and P. Cramer. 2003. "Architecture of Initiation-Competent 12-Subunit RNA Polymerase II." *Proceedings of the National Academy of Sciences* 100 (12): 6964–68. doi:10.1073/pnas.1030608100.

Armache, K.-J., S. Mitterweger, A. Meinhart, and P. Cramer. 2005. "Structures of Complete RNA Polymerase II and Its Subcomplex, Rpb4/7." *Journal of Biological Chemistry* 280 (8): 7131–34. doi:10.1074/jbc.M413038200.

Armstrong, J. A., J. J. Bieker, and B. M. Emerson. 1998. "A SWI/SNF-Related Chromatin Remodeling Complex, E-RC1, Is Required for Tissue-Specific Transcriptional Regulation by EKLF in Vitro." *Cell* 95 (1): 93–104.

Asada, S., Y. Choi, M. Yamada, S.-C. Wang, M.-C. Hung, J. Qin, and M. Uesugi. 2002. "External Control of Her2 Expression and Cancer Cell Growth by Targeting a Ras-Linked Coactivator." *Proceedings of the National Academy of Sciences of the United States of America* 99 (20): 12747–52. doi:10.1073/pnas.202162199.

Auble, D. T., K. E. Hansen, C. G. Mueller, W. S. Lane, J. Thorner, and S. Hahn. 1994. "Mot1, a Global Repressor of RNA Polymerase II Transcription, Inhibits TBP Binding to DNA by an ATP-Dependent Mechanism." *Genes & Development* 8 (16): 1920–34.

Bäckström, S., N. Elfving, R. Nilsson, G. Wingsle, and S. Björklund. 2007. "Purification of a Plant Mediator from Arabidopsis Thaliana Identifies PFT1 as the Med25 Subunit." *Molecular Cell* 26 (5): 717–29. doi:10.1016/j.molcel.2007.05.007.

Badis, G., E. T. Chan, H. van Bakel, L. Pena-Castillo, D. Tillo, K. Tsui, C.D. Carlson, et al. 2008. "A Library of Yeast Transcription Factor Motifs Reveals a Widespread Function for Rsc3 in Targeting Nucleosome Exclusion at Promoters." *Molecular Cell* 32 (6): 878–87. doi:10.1016/j.molcel.2008.11.020.

Baek, H. J., Y. K. Kang, and R. G. Roeder. 2006. "Human Mediator Enhances Basal Transcription by Facilitating Recruitment of Transcription Factor IIB during Preinitiation Complex Assembly." *The Journal of Biological Chemistry* 281 (22): 15172–81. doi:10.1074/jbc.M601983200.

Baek, H. J., S. Malik, J. Qin, and R. G. Roeder. 2002. "Requirement of TRAP/Mediator for Both Activator-Independent and Activator-Dependent Transcription in Conjunction with TFIID-Associated TAFs." *Molecular and Cellular Biology* 22 (8): 2842–52. doi:10.1128/MCB.22.8.2842-2852.2002.

Balamotis, M.I. A., M. A. Pennella, J. L. Stevens, B. Wasyluk, A. S. Belmont, and A. J. Berk. 2009. "Complexity in Transcription Control at the Activation Domain-Mediator Interface." *Science Signaling* 2 (69): ra20. doi:10.1126/scisignal.1164302.

Bardet, A. F., J. Steinmann, S. Bafna, J. A. Knoblich, J. Zeitlinger, and A. Stark. 2013. "Identification of Transcription Factor Binding Sites from ChIP-Seq Data at High Resolution." *Bioinformatics (Oxford, England)* 29 (21): 2705–13. doi:10.1093/bioinformatics/btt470.

Bartolomei, M. S., N. F. Halden, C. R. Cullen, and J. L. Corden. 1988. "Genetic Analysis of the Repetitive Carboxyl-Terminal Domain of the Largest Subunit of Mouse RNA Polymerase II." *Molecular and Cellular Biology* 8 (1): 330–39. doi:10.1128/MCB.8.1.330.

Basehoar, A. D., S. J Zanton, and B.F. Pugh. 2004. "Identification and Distinct Regulation of Yeast TATA Box-Containing Genes." *Cell* 116 (5): 699–709. doi:10.1016/S0092-8674(04)00205-3.

Bastien, J., S. A. Stitah, T. Riedl, J.-M. Egly, P. Chambon, and C. Rochette-Egly. 2000. "TFIIH Interacts with the Retinoic Acid Receptor Γ and Phosphorylates Its AF-1-Activating Domain through cdk7." *Journal of Biological Chemistry* 275 (29): 21896–904. doi:10.1074/jbc.M001985200.

Basu, S., Y. L. Tan, E. J. R. Taylor, E. D. Laue, and S. F. Lee. 2016. "Studying the Dynamics of Chromatin-Binding Proteins in Mammalian Cells Using Single-Molecule Localisation Microscopy." *Methods in Molecular Biology (Clifton, N.J.)* 1431: 235–63. doi:10.1007/978-1-4939-3631-1_17.

Baumli, S., S. Hoepfner, and P. Cramer. 2005. "A Conserved Mediator Hinge Revealed in the Structure of the MED7.MED21 (Med7.Srb7) Heterodimer." *The Journal of Biological Chemistry* 280 (18): 18171–78. doi:10.1074/jbc.M413466200.

Belakavadi, M., and J. D. Fondell. 2010. "Cyclin-Dependent Kinase 8 Positively Cooperates with Mediator to Promote Thyroid Hormone Receptor-Dependent Transcriptional Activation." *Molecular and Cellular Biology* 30 (10): 2437–48. doi:10.1128/MCB.01541-09.

Bendjennat, M., and P. A. Weil. 2008. "The Transcriptional Repressor Activator Protein Rap1p Is a Direct Regulator of TATA-Binding Protein." *The Journal of Biological Chemistry* 283 (13): 8699–8710. doi:10.1074/jbc.M709436200.

- Bernecky, C., P. Grob, C. C. Ebmeier, E. Nogales, and D. J. Taatjes. 2011. "Molecular Architecture of the Human Mediator-RNA Polymerase II-TFIIF Assembly." *PLoS Biology* 9 (3): e1000603. doi:10.1371/journal.pbio.1000603.
- Bernecky, C., F. Herzog, W. Baumeister, J. M. Plitzko, and P. Cramer. 2016. "Structure of Transcribing Mammalian RNA Polymerase II." *Nature* 529 (7587): 551–54. doi:10.1038/nature16482.
- Berretta, J., and A. Morillon. 2009. "Pervasive Transcription Constitutes a New Level of Eukaryotic Genome Regulation." *EMBO Reports* 10 (9): 973–82. doi:10.1038/embor.2009.181.
- Berroteran, R. W., D. E. Ware, and M. Hampsey. 1994. "The sua8 Suppressors of *Saccharomyces Cerevisiae* Encode Replacements of Conserved Residues within the Largest Subunit of RNA Polymerase II and Affect Transcription Start Site Selection Similarly to sua7 (TFIIB) Mutations." *Molecular and Cellular Biology* 14 (1): 226–37. doi:10.1128/MCB.14.1.226.
- Berti, L., G. Mittler, G. K. Przemec, G. Stelzer, B. Günzler, F. Amati, E. Conti, et al. 2001. "Isolation and Characterization of a Novel Gene from the DiGeorge Chromosomal Region That Encodes for a Mediator Subunit." *Genomics* 74 (3): 320–32. doi:10.1006/geno.2001.6566.
- Bieniossek, C., G. Papai, C. Schaffitzel, F. Garzoni, M. Chaillet, E. Scheer, P. Papadopoulos, L. Tora, P. Schultz, and I. Berger. 2013. "The Architecture of Human General Transcription Factor TFIID Core Complex." *Nature* 493 (7434): 699–702. doi:10.1038/nature11791.
- Bienkiewicz, E. A., A. -Y. M. Woody, and R. W. Woody. 2000. "Conformation of the RNA Polymerase II C-Terminal Domain: Circular Dichroism of Long and Short fragments1." *Journal of Molecular Biology* 297 (1): 119–33. doi:10.1006/jmbi.2000.3545.
- Birse, C. E., L. Minvielle-Sebastia, B. A. Lee, W. Keller, and N. J. Proudfoot. 1998. "Coupling Termination of Transcription to Messenger RNA Maturation in Yeast." *Science (New York, N.Y.)* 280 (5361): 298–301.
- Blackwood, E. M., and J. T. Kadonaga. 1998. "Going the Distance: A Current View of Enhancer Action." *Science (New York, N.Y.)* 281 (5373): 60–63.
- Blake, W. J., M. Kærn, C. R. Cantor, and J. J. Collins. 2003. "Noise in Eukaryotic Gene Expression." *Nature* 422 (6932): 633–37. doi:10.1038/nature01546.
- Bleichenbacher, M., S. Tan, and T. J. Richmond. 2003. "Novel Interactions Between the Components of Human and Yeast TFIIA/TBP/DNA Complexes." *Journal of Molecular Biology* 332 (4): 783–93. doi:10.1016/S0022-2836(03)00887-8.
- Boeva, V.. 2016. "Analysis of Genomic Sequence Motifs for Deciphering Transcription Factor Binding and Transcriptional Regulation in Eukaryotic Cells." *Bioinformatics and Computational Biology*, 24. doi:10.3389/fgene.2016.00024.
- Boiteux, S., and S. Jinks-Robertson. 2013. "DNA Repair Mechanisms and the Bypass of DNA Damage in *Saccharomyces Cerevisiae*." *Genetics* 193 (4): 1025–64. doi:10.1534/genetics.112.145219.
- Bonnet, J., C-Y. Wang, T. Baptista, S. D. Vincent, W-C. Hsiao, M. Stierle, C-F. Kao, L. Tora, and D. Devys. 2014. "The SAGA Coactivator Complex Acts on the Whole Transcribed Genome and Is Required for RNA Polymerase II Transcription." *Genes & Development* 28 (18): 1999–2012. doi:10.1101/gad.250225.114.
- Bosveld, F., S. van Hoek, and O. C. M. Sibon. 2008. "Establishment of Cell Fate during Early *Drosophila* Embryogenesis Requires Transcriptional Mediator Subunit dMED31." *Developmental Biology* 313 (2): 802–13. doi:10.1016/j.ydbio.2007.11.019.
- Boube, M., C. Faucher, L. Joulia, D. L. Cribbs, and H. M. Bourbon. 2000. "Drosophila Homologs of Transcriptional Mediator Complex Subunits Are Required for Adult Cell and Segment Identity Specification." *Genes & Development* 14 (22): 2906–17.
- Boube, M., L. Joulia, D.L. Cribbs, and H-M. Bourbon. 2002. "Evidence for a Mediator of RNA Polymerase II Transcriptional Regulation Conserved from Yeast to Man." *Cell* 110 (2): 143–51.
- Bourbon, H-M. 2008. "Comparative Genomics Supports a Deep Evolutionary Origin for the Large, Four-Module Transcriptional Mediator Complex." *Nucleic Acids Research* 36 (12): 3993–4008. doi:10.1093/nar/gkn349.

Bourbon, H-M., A. Aguilera, A. Z. Ansari, F. J. Asturias, A. J. Berk, S. Bjorklund, T. K. Blackwell, et al. 2004. "A Unified Nomenclature for Protein Subunits of Mediator Complexes Linking Transcriptional Regulators to RNA Polymerase II." *Molecular Cell* 14 (5): 553–57. doi:10.1016/j.molcel.2004.05.011.

Boyer, L. A., R. R. Latek, and C. L. Peterson. 2004. "The SANT Domain: A Unique Histone-Tail-Binding Module?" *Nature Reviews. Molecular Cell Biology* 5 (2): 158–63. doi:10.1038/nrm1314.

Boyer, T. G., M. E. Martin, E. Lees, R. P. Ricciardi, and A. J. Berk. 1999. "Mammalian Srb/Mediator Complex Is Targeted by Adenovirus E1A Protein." *Nature* 399 (6733): 276–79. doi:10.1038/20466.

Breathnach, R., and P. Chambon. 1981. "Organization and Expression of Eucaryotic Split Genes Coding for Proteins." *Annual Review of Biochemistry* 50: 349–83. doi:10.1146/annurev.bi.50.070181.002025.

Brown, C. E., L. Howe, K. Sousa, S. C. Alley, M. J. Carrozza, S. Tan, and J. L. Workman. 2001. "Recruitment of HAT Complexes by Direct Activator Interactions with the ATM-Related Tra1 Subunit." *Science (New York, N.Y.)* 292 (5525): 2333–37. doi:10.1126/science.1060214.

Bryant, G. O., and M. Ptashne. 2003. "Independent Recruitment in Vivo by Gal4 of Two Complexes Required for Transcription." *Molecular Cell* 11 (5): 1301–9.

Bumgarner, S.L., R.D. Dowell, P. Grisafi, D. K. Gifford, and G. R. Fink. 2009. "Toggle Involving Cis-Interfering Noncoding RNAs Controls Variegated Gene Expression in Yeast." *Proceedings of the National Academy of Sciences of the United States of America* 106 (43): 18321–26. doi:10.1073/pnas.0909641106.

Buratowski, S. 2003. "The CTD Code." *Nature Structural & Molecular Biology* 10 (9): 679–80. doi:10.1038/nsb0903-679.

Buratowski, S., S. Hahn, L. Guarente, and P.A. Sharp. 1989. "Five Intermediate Complexes in Transcription Initiation by RNA Polymerase II." *Cell* 56 (4): 549–61. doi:10.1016/0092-8674(89)90578-8.

Buratowski, S., S. Hahn, P. A. Sharp, and L. Guarente. 1988. "Function of a Yeast TATA Element-Binding Protein in a Mammalian Transcription System." *Nature* 334 (6177): 37–42. doi:10.1038/334037a0.

Burke, T. W., and J. T. Kadonaga. 1996. "Drosophila TFIID Binds to a Conserved Downstream Basal Promoter Element That Is Present in Many TATA-Box-Deficient Promoters." *Genes & Development* 10 (6): 711–24. doi:10.1101/gad.10.6.711.

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." *The EMBO Journal* 5 (11): 2923–30.

Burton, Z. F., M. Killeen, M. Sopta, L. G. Ortolan, and J. Greenblatt. 1988. "RAP30/74: A General Initiation Factor That Binds to RNA Polymerase II." *Molecular and Cellular Biology* 8 (4): 1602–13. doi:10.1128/MCB.8.4.1602.

Bushnell, D. A., and R. D. Kornberg. 2003. "Complete, 12-Subunit RNA Polymerase II at 4.1-Å Resolution: Implications for the Initiation of Transcription." *Proceedings of the National Academy of Sciences* 100 (12): 6969–73. doi:10.1073/pnas.1130601100.

Bushnell, D. A., K. D. Westover, R. E. Davis, and R. D. Kornberg. 2004. "Structural Basis of Transcription: An RNA Polymerase II-TFIIB Cocystal at 4.5 Angstroms." *Science* 303 (5660): 983–88. doi:10.1126/science.1090838.

Cabal, G. G., A. Genovesio, S. Rodriguez-Navarro, C. Zimmer, O. Gadal, A. Lesne, H. Buc, et al. 2006. "SAGA Interacting Factors Confine Sub-Diffusion of Transcribed Genes to the Nuclear Envelope." *Nature* 441 (7094): 770–73. doi:10.1038/nature04752.

Cai, G., T. Imasaki, Y. Takagi, and F. J. Asturias. 2009. "Mediator Structural Conservation and Implications for the Regulation Mechanism." *Structure (London, England: 1993)* 17 (4): 559–67. doi:10.1016/j.str.2009.01.016.

Cai, G., T. Imasaki, K. Yamada, F. Cardelli, Y. Takagi, and F. J. Asturias. 2010. "Mediator Head Module Structure and Functional Interactions." *Nature Structural & Molecular Biology* 17 (3): 273–79. doi:10.1038/nsmb.1757.

Cairns, B. R., Y. J. Kim, M. H. Sayre, B. C. Laurent, and R. D. Kornberg. 1994. "A Multisubunit Complex Containing the SWI1/ADR6, SWI2/SNF2, SWI3, SNF5, and SNF6 Gene Products Isolated from Yeast." *Proceedings of the National Academy of Sciences of the United States of America* 91 (5): 1950–54.

Cairns, B. R., Y. Lorch, Y. Li, M. Zhang, L. Lacomis, H. Erdjument-Bromage, P. Tempst, J. Du, B. Laurent, and R. D. Kornberg. 1996. "RSC, an Essential, Abundant Chromatin-Remodeling Complex." *Cell* 87 (7): 1249–60.

Cairns, B. R., H. Erdjument-Bromage, P. Tempst, F. Winston, and R. D. Kornberg. 1998. "Two Actin-Related Proteins Are Shared Functional Components of the Chromatin-Remodeling Complexes RSC and SWI/SNF." *Molecular Cell* 2 (5): 639–51. doi:10.1016/S1097-2765(00)80162-8.

Carninci, P., T. Kasukawa, S. Katayama, J. Gough, M. C. Frith, N. Maeda, R. Oyama, et al., FANTOM Consortium and RIKEN Genome Exploration Research Group and Genome Science Group (Genome Network Project Core Group). 2005. "The Transcriptional Landscape of the Mammalian Genome." *Science (New York, N.Y.)* 309 (5740): 1559–63. doi:10.1126/science.1112014.

Carrera, I., F. Janody, N. Leeds, F. Duvéau, and J. E. Treisman. 2008. "Pygopus Activates Wingless Target Gene Transcription through the Mediator Complex Subunits Med12 and Med13." *Proceedings of the National Academy of Sciences of the United States of America* 105 (18): 6644–49. doi:10.1073/pnas.0709749105.

Carter, R., and G. Drouin. 2010. "The Increase in the Number of Subunits in Eukaryotic RNA Polymerase III Relative to RNA Polymerase II Is due to the Permanent Recruitment of General Transcription Factors." *Molecular Biology and Evolution* 27 (5): 1035–43. doi:10.1093/molbev/msp316.

Cavallini, B., J. Huet, J.-L. Plassat, A. Sentenac, and J.-M. Egly. 1988. "A Yeast Activity Can Substitute for the HeLa Cell TATA Box Factor." *Nature* 334 (6177): 77–80. doi:10.1038/334077a0.

Cevher, M. A., Y. Shi, D. Li, B. T. Chait, S. Malik, and R. G. Roeder. 2014. "Reconstitution of Active Human Core Mediator Complex Reveals a Critical Role of the MED14 Subunit." *Nature Structural & Molecular Biology* 21 (12): 1028–34. doi:10.1038/nsmb.2914.

Chang, W.-H., and R. D. Kornberg. 2000. "Electron Crystal Structure of the Transcription Factor and DNA Repair Complex, Core TFIID." *Cell* 102 (5): 609–13. doi:10.1016/S0092-8674(00)00083-0.

Chasman, D. I., K. M. Flaherty, P. A. Sharp, and R. D. Kornberg. 1993. "Crystal Structure of Yeast TATA-Binding Protein and Model for Interaction with DNA." *Proceedings of the National Academy of Sciences of the United States of America* 90 (17): 8174–78.

Chen, H.-T., L. Warfield, and S. Hahn. 2007. "The Positions of TFIIF and TFIIE in the RNA Polymerase II Transcription Preinitiation Complex." *Nature Structural & Molecular Biology* 14 (8): 696–703. doi:10.1038/nsmb1272.

Cheung, A. C. M., and P. Cramer. 2012. "A Movie of RNA Polymerase II Transcription." *Cell* 149 (7): 1431–37. doi:10.1016/j.cell.2012.06.006.

Chi, Y., M. J. Huddleston, X. Zhang, R. A. Young, R. S. Annan, S. A. Carr, and R. J. Deshaies. 2001. "Negative Regulation of Gcn4 and Msn2 Transcription Factors by Srb10 Cyclin-Dependent Kinase." *Genes & Development* 15 (9): 1078–92. doi:10.1101/gad.867501.

Cho, E.-J., T. Takagi, C. R. Moore, and S. Buratowski. 1997. "mRNA Capping Enzyme Is Recruited to the Transcription Complex by Phosphorylation of the RNA Polymerase II Carboxy-Terminal Domain." *Genes & Development* 11 (24): 3319–26. doi:10.1101/gad.11.24.3319.

Churchman, L. S. and J. S. Weissman. 2011. "Nascent Transcript Sequencing Visualizes Transcription at Nucleotide Resolution." *Nature* 469 (7330): 368–73. doi:10.1038/nature09652.

Clapier, C.R., and B. R. Cairns. 2009. "The Biology of Chromatin Remodeling Complexes." *Annual Review of Biochemistry* 78: 273–304. doi:10.1146/annurev.biochem.77.062706.153223.

Compe, E., and J.-M. Egly. 2012. "TFIID: When Transcription Met DNA Repair." *Nature Reviews Molecular Cell Biology* 13 (6): 343–54. doi:10.1038/nrm3350.

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." *Proceedings of the National Academy of Sciences of the United States of America* 88 (14): 6205–9.

Conaway, R. C., S. Sato, C. Tomomori-Sato, T. Yao, and J. W. Conaway. 2005. "The Mammalian Mediator Complex and Its Role in Transcriptional Regulation." *Trends in Biochemical Sciences* 30 (5): 250–55. doi:10.1016/j.tibs.2005.03.002.

Core, L. J., and J. T. Lis. 2008. "Transcription Regulation Through Promoter-Proximal Pausing of RNA Polymerase II." *Science* 319 (5871): 1791–92. doi:10.1126/science.1150843.

Core, L. J., J.J. Waterfall, and J. T. Lis. 2008. "Nascent RNA Sequencing Reveals Widespread Pausing and Divergent Initiation at Human Promoters." *Science* 322 (5909): 1845–48. doi:10.1126/science.1162228.

Corona, D. F. V., and J.W. Tamkun. 2004. "Multiple Roles for ISWI in Transcription, Chromosome Organization and DNA Replication." *Biochimica Et Biophysica Acta* 1677 (1-3): 113–19. doi:10.1016/j.bbaexp.2003.09.018.

Cox, J., and M. Mann. 2008. "MaxQuant Enables High Peptide Identification Rates, Individualized P.p.b.-Range Mass Accuracies and Proteome-Wide Protein Quantification." *Nature Biotechnology* 26 (12): 1367–72. doi:10.1038/nbt.1511.

Craighead, J. L, W-H. Chang, and F. J Asturias. 2002. "Structure of Yeast RNA Polymerase II in Solution: Implications for Enzyme Regulation and Interaction with Promoter DNA." *Structure* 10 (8): 1117–25. doi:10.1016/S0969-2126(02)00813-4.

Cramer, P. 2004. "RNA Polymerase II Structure: From Core to Functional Complexes." *Current Opinion in Genetics & Development* 14 (2): 218–26. doi:10.1016/j.gde.2004.01.003.

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 Cramer, P., D. A. Bushnell, and R. D. Kornberg. 2001. "Structural Basis of Transcription: RNA Polymerase II at 2.8 Ångstrom Resolution." *Science* 292 (5523): 1863–76. doi:10.1126/science.1059493.

Cui, X., D. Xu, C. Lv, F. Qu, J. He, M. Chen, Y. Liu, et al. 2011. "Suppression of MED19 Expression by shRNA Induces Inhibition of Cell Proliferation and Tumorigenesis in Human Prostate Cancer Cells." *BMB Reports* 44 (8): 547–52.

Dang, W., K. K. Steffen, R. Perry, J.A. Dorsey, F.B. Johnson, A. Shilatifard, M. Kaeberlein, B. K. Kennedy, and S. L. Berger. 2009. "Histone H4 Lysine 16 Acetylation Regulates Cellular Lifespan." *Nature* 459 (7248): 802–7. doi:10.1038/nature08085.

Daniel, J.y A., M. S. Torok, Z-W. Sun, D. Schieltz, C. D. Allis, J. R. Yates, and P. A. Grant. 2004. "Deubiquitination of Histone H2B by a Yeast Acetyltransferase Complex Regulates Transcription." *The Journal of Biological Chemistry* 279 (3): 1867–71. doi:10.1074/jbc.C300494200.

Darzacq, X., Y. Shav-Tal, V. de Turreis, Y. Brody, S. M. Shenoy, R. D. Phair, and R H. Singer. 2007. "In Vivo Dynamics of RNA Polymerase II Transcription." *Nature Structural & Molecular Biology* 14 (9): 796–806. doi:10.1038/nsmb1280.

David, L., S. Clauder-Münster, and L. M. Steinmetz. 2014. "High-Density Tiling Microarray Analysis of the Full Transcriptional Activity of Yeast." *Methods in Molecular Biology (Clifton, N.J.)* 1205: 257–73. doi:10.1007/978-1-4939-1363-3_16.

David, L., W. Huber, M. Granovskaia, J.Toedling, C. J. Palm, L. Bofkin, T. Jones, R. W. Davis, and L. M. Steinmetz. 2006. "A High-Resolution Map of Transcription in the Yeast Genome." *Proceedings of the National Academy of Sciences of the United States of America* 103 (14): 5320–25. doi:10.1073/pnas.0601091103.

Davis, C. A., and Manuel A.. 2006. "Accumulation of Unstable Promoter-Associated Transcripts upon Loss of the Nuclear Exosome Subunit Rrp6p in *Saccharomyces Cerevisiae*." *Proceedings of the National Academy of Sciences of the United States of America* 103 (9): 3262–67. doi:10.1073/pnas.0507783103.

Davis, J. A, Y. Takagi, R. D Kornberg, and Francisco J Asturias. 2002. "Structure of the Yeast RNA Polymerase II Holoenzyme: Mediator Conformation and Polymerase Interaction." *Molecular Cell* 10 (2): 409–15. doi:10.1016/S1097-2765(02)00598-1.

DeJong, J., and R. G. Roeder. 1993. "A Single cDNA, hTFIIA/alpha, Encodes Both the p35 and p19 Subunits of Human TFIIA." *Genes & Development* 7 (11): 2220–34. doi:10.1101/gad.7.11.2220.

Dekker, J., K. Rippe, M. Dekker, and N. Kleckner. 2002. "Capturing Chromosome Conformation." *Science (New York, N.Y.)* 295 (5558): 1306–11. doi:10.1126/science.1067799.

Deng, W., and S. G. E. Roberts. 2005. "A Core Promoter Element Downstream of the TATA Box That Is Recognized by TFIIB." *Genes & Development* 19 (20): 2418–23. doi:10.1101/gad.342405.

———. 2007. “TFIIB and the Regulation of Transcription by RNA Polymerase II.” *Chromosoma* 116 (5): 417–29. doi:10.1007/s00412-007-0113-9.

Deuring, R., L. Fanti, J. A. Armstrong, M. Sarte, O. Papoulas, M. Prestel, G. Daubresse, et al. 2000. “The ISWI Chromatin-Remodeling Protein Is Required for Gene Expression and the Maintenance of Higher Order Chromatin Structure in Vivo.” *Molecular Cell* 5 (2): 355–65.

Dion, M. F., T. Kaplan, M. Kim, S. Buratowski, N. Friedman, and O. J. Rando. 2007. “Dynamics of Replication-Independent Histone Turnover in Budding Yeast.” *Science (New York, N.Y.)* 315 (5817): 1405–8. doi:10.1126/science.1134053.

Dion, V., and B. Coulombe. 2003. “Interactions of a DNA-Bound Transcriptional Activator with the TBP-TFIIA-TFIIB-Promoter Quaternary Complex.” *Journal of Biological Chemistry* 278 (13): 11495–501. doi:10.1074/jbc.M211938200.

Dixon, J. R., S. Selvaraj, F. Yue, A. Kim, Y. Li, Y. Shen, M. Hu, J. S. Liu, and B. Ren. 2012. “Topological Domains in Mammalian Genomes Identified by Analysis of Chromatin Interactions.” *Nature* 485 (7398): 376–80. doi:10.1038/nature11082.

Djebali, S., C. A. Davis, A. Merkel, A. Dobin, T. Lassmann, A. Mortazavi, A. Tanzer, et al. 2012. “Landscape of Transcription in Human Cells.” *Nature* 489 (7414): 101–8. doi:10.1038/nature11233.

Dobi, K. C., and F. Winston. 2007. “Analysis of Transcriptional Activation at a Distance in *Saccharomyces Cerevisiae*.” *Molecular and Cellular Biology* 27 (15): 5575–86. doi:10.1128/MCB.00459-07.

Donner, A. J., C. C. Ebmeier, D. J. Taatjes, and J. M. Espinosa. 2010. “CDK8 Is a Positive Regulator of Transcriptional Elongation within the Serum Response Network.” *Nature Structural & Molecular Biology* 17 (2): 194–201. doi:10.1038/nsmb.1752.

Dorsett, D. 2011. “Cohesin: Genomic Insights into Controlling Gene Transcription and Development.” *Current Opinion in Genetics & Development* 21 (2): 199–206. doi:10.1016/j.gde.2011.01.018.

Dotson, M. R., C. X. Yuan, R. G. Roeder, L. C. Myers, C. M. Gustafsson, Y. W. Jiang, Y. Li, R. D. Kornberg, and F. J. Asturias. 2000. “Structural Organization of Yeast and Mammalian Mediator Complexes.” *Proceedings of the National Academy of Sciences* 97 (26): 14307–10. doi:10.1073/pnas.260489497.

Douziech, M., F. Coin, J-M. Chipoulet, Y. Arai, Y. Ohkuma, J-M. Egly, and B. Coulombe. 2000. “Mechanism of Promoter Melting by the Xeroderma Pigmentosum Complementation Group B Helicase of Transcription Factor IIH Revealed by Protein-DNA Photo-Cross-Linking.” *Molecular and Cellular Biology* 20 (21): 8168–77. doi:10.1128/MCB.20.21.8168-8177.2000.

Dvir, A., R. C. Conaway, and J. W. Conaway. 1996. “Promoter Escape by RNA Polymerase II A ROLE FOR AN ATP COFACTOR IN SUPPRESSION OF ARREST BY POLYMERASE AT PROMOTER-PROXIMAL SITES.” *Journal of Biological Chemistry* 271 (38): 23352–56. doi:10.1074/jbc.271.38.23352.

———. 1997. “A Role for TFIIF in Controlling the Activity of Early RNA Polymerase II Elongation Complexes.” *Proceedings of the National Academy of Sciences* 94 (17): 9006–10.

Ebmeier, C. C., and D. J. Taatjes. 2010. “Activator-Mediator Binding Regulates Mediator-Cofactor Interactions.” *Proceedings of the National Academy of Sciences of the United States of America* 107 (25): 11283–88. doi:10.1073/pnas.0914215107.

Edwards, A. M., C. M. Kane, R. A. Young, and R. D. Kornberg. 1991. “Two Dissociable Subunits of Yeast RNA Polymerase II Stimulate the Initiation of Transcription at a Promoter in Vitro.” *Journal of Biological Chemistry* 266 (1): 71–75.

Eichner, J., H-T. Chen, L. Warfield, and S. Hahn. 2010. “Position of the General Transcription Factor TFIIF within the RNA Polymerase II Transcription Preinitiation Complex.” *The EMBO Journal* 29 (4): 706–16. doi:10.1038/emboj.2009.386.

Eick, D., and M. Geyer. 2013. “The RNA Polymerase II Carboxy-Terminal Domain (CTD) Code.” *Chemical Reviews* 113 (11): 8456–90. doi:10.1021/cr400071f.

Elmlund, H., V. Baraznenok, M. Lindahl, C. O. Samuelsen, P. J. B. Koeck, S. Holmberg, H. Hebert, and C. M. Gustafsson. 2006. “The Cyclin-Dependent Kinase 8 Module Sterically Blocks Mediator Interactions with RNA Polymerase II.” *Proceedings of the National Academy of Sciences of the United States of America* 103 (43): 15788–93. doi:10.1073/pnas.0607483103.

Esnault, C., Y. Ghavi-Helm, S. Brun, J. Soutourina, N. Van Berkum, C. Boschiero, F. Holstege, and M. Werner. 2008. "Mediator-Dependent Recruitment of TFIID Modules in Preinitiation Complex." *Molecular Cell* 31 (3): 337–46. doi:10.1016/j.molcel.2008.06.021.

Evans, Rachel, Jennifer A. Fairley, and Stefan G. E. Roberts. 2001. "Activator-Mediated Disruption of Sequence-Specific DNA Contacts by the General Transcription Factor TFIIB." *Genes & Development* 15 (22): 2945–49. doi:10.1101/gad.206901.

Eyboulet, F., C. Cibot, T. Eychenne, H. Neil, O. Alibert, M. Werner, and J. Soutourina. 2013. "Mediator Links Transcription and DNA Repair by Facilitating Rad2/XPG Recruitment." *Genes & Development* 27 (23): 2549–62. doi:10.1101/gad.225813.113.

Eyboulet, F., S. Wydau-Dematteis, T. Eychenne, O. Alibert, H. Neil, C. Boschiero, M-C. Nevers, et al. 2015. "Mediator Independently Orchestrates Multiple Steps of Preinitiation Complex Assembly in Vivo." *Nucleic Acids Research* 43 (19): 9214–31. doi:10.1093/nar/gkv782.

Farago, M., T. Nahari, C. Hammel, C. N. Cole, and M. Choder. 2003. "Rpb4p, a Subunit of RNA Polymerase II, Mediates mRNA Export during Stress." *Molecular Biology of the Cell* 14 (7): 2744–55. doi:10.1091/mbc.E02-11-0740.

Farla, P., R. Hersmus, B. Geverts, P. O. Mari, A. L. Nigg, H. J. Dubbink, J. Trapman, and A. B. Houtsmuller. 2004. "The Androgen Receptor Ligand-Binding Domain Stabilizes DNA Binding in Living Cells." *Journal of Structural Biology* 147 (1): 50–61. doi:10.1016/j.jsb.2004.01.002.

Firestein, R., A. J. Bass, S. Y. Kim, I. F. Dunn, S. J. Silver, I. Guney, E. Freed, et al. 2008. "CDK8 Is a Colorectal Cancer Oncogene That Regulates Beta-Catenin Activity." *Nature* 455 (7212): 547–51. doi:10.1038/nature07179.

Fishburn, J., N. Mohibullah, and S. Hahn. 2005. "Function of a Eukaryotic Transcription Activator during the Transcription Cycle." *Molecular Cell* 18 (3): 369–78. doi:10.1016/j.molcel.2005.03.029.

Flanagan, P. M., R. J. Kelleher, M. H. Sayre, H. Tschochner, and R. D. Kornberg. 1991. "A Mediator Required for Activation of RNA Polymerase II Transcription in Vitro." *Nature* 350 (6317): 436–38. doi:10.1038/350436a0.

Flores, O., E. Maldonado, Z. Burton, J. Greenblatt, and D. Reinberg. 1988. "Factors Involved in Specific Transcription by Mammalian RNA Polymerase II. RNA Polymerase II-Associating Protein 30 Is an Essential Component of Transcription Factor IIF." *Journal of Biological Chemistry* 263 (22): 10812–16.

Fondell, J. D., H. Ge, and R. G. Roeder. 1996. "Ligand Induction of a Transcriptionally Active Thyroid Hormone Receptor Coactivator Complex." *Proceedings of the National Academy of Sciences of the United States of America* 93 (16): 8329–33.

Forget, D., M-F. Langelier, C. Thérien, 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." *Molecular and Cellular Biology* 24 (3): 1122–31. doi:10.1128/MCB.24.3.1122-1131.2004.

Freire-Picos, M. A., S. Krishnamurthy, Z-W. Sun, and M. Hampsey. 2005. "Evidence That the Tfg1/Tfg2 Dimer Interface of TFIIF Lies near the Active Center of the RNA Polymerase II Initiation Complex." *Nucleic Acids Research* 33 (16): 5045–52. doi:10.1093/nar/gki825.

Fujita, T., and H. Fujii. 2013. "Efficient Isolation of Specific Genomic Regions and Identification of Associated Proteins by Engineered DNA-Binding Molecule-Mediated Chromatin Immunoprecipitation (enChIP) Using CRISPR." *Biochemical and Biophysical Research Communications* 439 (1): 132–36. doi:10.1016/j.bbrc.2013.08.013.

Galbraith, M. D., M. A. Allen, C. L. Bensard, X. Wang, M. K. Schwinn, B. Qin, H. W. Long, et al. 2013. "HIF1A Employs CDK8-Mediator to Stimulate RNAPII Elongation in Response to Hypoxia." *Cell* 153 (6): 1327–39. doi:10.1016/j.cell.2013.04.048.

Ge, K., M. Guermah, C-X. Yuan, M. Ito, A. E. Wallberg, B. M. Spiegelman, and R. G. Roeder. 2002. "Transcription Coactivator TRAP220 Is Required for PPAR Gamma 2-Stimulated Adipogenesis." *Nature* 417 (6888): 563–67. doi:10.1038/417563a.

Gershenson, N. I., and I. P. Ioshikhes. 2005. "Synergy of Human Pol II Core Promoter Elements Revealed by Statistical Sequence Analysis." *Bioinformatics* 21 (8): 1295–1300. doi:10.1093/bioinformatics/bti172.

Ghazy, M. A., S. A. Brodie, M. L. Ammerman, L. M. Ziegler, and A. S. Ponticelli. 2004. "Amino Acid Substitutions in Yeast TFIIF Confer Upstream Shifts in Transcription Initiation and Altered Interaction with RNA Polymerase II." *Molecular and Cellular Biology* 24 (24): 10975–85. doi:10.1128/MCB.24.24.10975-10985.2004.

Giardina, C., and J. T. Lis. 1993. "DNA Melting on Yeast RNA Polymerase II Promoters." *Science* 261 (5122): 759–62. doi:10.1126/science.8342041.

Gibbons, B. J., E. J. Brignole, M. Azubel, K. Murakami, N. R. Voss, D. A. Bushnell, F. J. Asturias, and R. D. Kornberg. 2012. "Subunit Architecture of General Transcription Factor TFIIF." *Proceedings of the National Academy of Sciences* 109 (6): 1949–54. doi:10.1073/pnas.1105266109.

Gniazdowski, M., J. L. Mandel, F. Gissinger, C. Keding, and P. Chambon. 1970. "Calf Thymus RNA Polymerases Exhibit Template Specificity." *Biochemical and Biophysical Research Communications* 38 (6): 1033–40. doi:10.1016/0006-291X(70)90343-8.

Goffeau, A., B. G. Barrell, H. Bussey, R. W. Davis, B. Dujon, H. Feldmann, F. Galibert, et al. 1996. "Life with 6000 Genes." *Science* 274 (5287): 546–67. doi:10.1126/science.274.5287.546.

Goldmark, J. P., T. G. Fazio, P. W. Estep, G. M. Church, and T. Tsukiyama. 2000. "The Isw2 Chromatin Remodeling Complex Represses Early Meiotic Genes upon Recruitment by Ume6p." *Cell* 103 (3): 423–33.

Gonzalez, D., N. Hamidi, R. Del Sol, J. J. Benschop, T. Nancy, C. Li, L. Francis, et al. 2014. "Suppression of Mediator Is Regulated by Cdk8-Dependent Grr1 Turnover of the Med3 Coactivator." *Proceedings of the National Academy of Sciences of the United States of America* 111 (7): 2500–2505. doi:10.1073/pnas.1307525111.

Goodrich, J. A., T. Hoey, C. J. Thut, A. Admon, and R. Tjian. 1993. "Drosophila TAFII40 Interacts with Both a VP16 Activation Domain and the Basal Transcription Factor TFIIB." *Cell* 75 (3): 519–30. doi:10.1016/0092-8674(93)90386-5.

Goppelt, A., G. Stelzer, F. Lottspeich, and M. Meisterernst. 1996. "A Mechanism for Repression of Class II Gene Transcription through Specific Binding of NC2 to TBP-Promoter Complexes via Heterodimeric Histone Fold Domains." *The EMBO Journal* 15 (12): 3105–16.

Govind, C. K., F. Zhang, H. Qiu, K. Hofmeyer, and A. G. Hinnebusch. 2007. "Gcn5 Promotes Acetylation, Eviction, and Methylation of Nucleosomes in Transcribed Coding Regions." *Molecular Cell* 25 (1): 31–42. doi:10.1016/j.molcel.2006.11.020.

Grohmann, D., and F. Werner. 2011. "Cycling through Transcription with the RNA Polymerase F/E (RPB4/7) Complex: Structure, Function and Evolution of Archaeal RNA Polymerase." *Research in Microbiology, Archaea and the Tree of Life*, 162 (1): 10–18. doi:10.1016/j.resmic.2010.09.002.

Grøntved, L., M. S. Madsen, M. Boergesen, R. G. Roeder, and S. Mandrup. 2010. "MED14 Tethers Mediator to the N-Terminal Domain of Peroxisome Proliferator-Activated Receptor Gamma and Is Required for Full Transcriptional Activity and Adipogenesis." *Molecular and Cellular Biology* 30 (9): 2155–69. doi:10.1128/MCB.01238-09.

Grosschedl, R., and M. L. Birnstiel. 1980. "Identification of Regulatory Sequences in the Prelude Sequences of an H2A Histone Gene by the Study of Specific Deletion Mutants in Vivo." *Proceedings of the National Academy of Sciences of the United States of America* 77 (3): 1432–36.

Grünberg, S., C. Reich, M. E. Zeller, M. S. Bartlett, and M. Thomm. 2010. "Rearrangement of the RNA Polymerase Subunit H and the Lower Jaw in Archaeal Elongation Complexes." *Nucleic Acids Research* 38 (6): 1950–63. doi:10.1093/nar/gkp1190.

Grünberg, S., L. Warfield, and S. Hahn. 2012. "Architecture of the RNA Polymerase II Preinitiation Complex and Mechanism of ATP-Dependent Promoter Opening." *Nature Structural & Molecular Biology* 19 (8): 788–96. doi:10.1038/nsmb.2334.

Gu, W., S. Malik, M. Ito, C. X. Yuan, J. D. Fondell, X. Zhang, E. Martinez, J. Qin, and R. G. Roeder. 1999. "A Novel Human SRB/MED-Containing Cofactor Complex, SMCC, Involved in Transcription Regulation." *Molecular Cell* 3 (1): 97–108.

Guenther, M. G., S. S. Levine, L. A. Boyer, R. Jaenisch, and R. A. Young. 2007. "A Chromatin Landmark and Transcription Initiation at Most Promoters in Human Cells." *Cell* 130 (1): 77–88. doi:10.1016/j.cell.2007.05.042.

- Guglielmi, B., J. Soutourina, C. Esnault, and M. Werner. 2007. "TFIIS Elongation Factor and Mediator Act in Conjunction during Transcription Initiation in Vivo." *Proceedings of the National Academy of Sciences of the United States of America* 104 (41): 16062–67. doi:10.1073/pnas.0704534104.
- Guglielmi, B., N.L. van Berkum, B. Klapholz, T. Bijma, M. Boube, C. Boschiero, H-M. Bourbon, F. C. P. Holstege, and M. Werner. 2004. "A High Resolution Protein Interaction Map of the Yeast Mediator Complex." *Nucleic Acids Research* 32 (18): 5379–91. doi:10.1093/nar/gkh878.
- Guidi, B. W., G. Bjornsdottir, D. C. Hopkins, L. Lacomis, H. Erdjument-Bromage, P. Tempst, and L. C. Myers. 2004. "Mutual Targeting of Mediator and the TFIID Kinase Kin28." *The Journal of Biological Chemistry* 279 (28): 29114–20. doi:10.1074/jbc.M404426200.
- Guo, B., C. A. Styles, Q. Feng, and G. R. Fink. 2000. "A Saccharomyces Gene Family Involved in Invasive Growth, Cell–cell Adhesion, and Mating." *Proceedings of the National Academy of Sciences of the United States of America* 97 (22): 12158–63.
- Guo, R., J. Chen, D. L. Mitchell, and D. G. Johnson. 2011. "GCN5 and E2F1 Stimulate Nucleotide Excision Repair by Promoting H3K9 Acetylation at Sites of Damage." *Nucleic Acids Research* 39 (4): 1390–97. doi:10.1093/nar/gkq983.
- Gustafsson, C. M., L. C. Myers, J. Beve, H. Spåhr, M. Lui, H. Erdjument-Bromage, P. Tempst, and R. D. Kornberg. 1998. "Identification of New Mediator Subunits in the RNA Polymerase II Holoenzyme from Saccharomyces Cerevisiae." *Journal of Biological Chemistry* 273 (47): 30851–54. doi:10.1074/jbc.273.47.30851.
- Ha, I., S. Roberts, E. Maldonado, X. Sun, L. U. Kim, M. Green, and D. Reinberg. 1993. "Multiple Functional Domains of Human Transcription Factor IIB: Distinct Interactions with Two General Transcription Factors and RNA Polymerase II." *Genes & Development* 7 (6): 1021–32. doi:10.1101/gad.7.6.1021.
- Haag, J. R., and C. S. Pikaard. 2011. "Multisubunit RNA Polymerases IV and V: Purveyors of Non-Coding RNA for Plant Gene Silencing." *Nature Reviews Molecular Cell Biology* 12 (8): 483–92. doi:10.1038/nrm3152.
- Hager, G. L., J. G. McNally, and T. Misteli. 2009. "Transcription Dynamics." *Molecular Cell* 35 (6): 741–53. doi:10.1016/j.molcel.2009.09.005.
- Hahn, J-S., and D. J. Thiele. 2004. "Activation of the Saccharomyces Cerevisiae Heat Shock Transcription Factor under Glucose Starvation Conditions by Snf1 Protein Kinase." *The Journal of Biological Chemistry* 279 (7): 5169–76. doi:10.1074/jbc.M311005200.
- Hahn, S., and E. T. Young. 2011. "Transcriptional Regulation in Saccharomyces Cerevisiae: Transcription Factor Regulation and Function, Mechanisms of Initiation, and Roles of Activators and Coactivators." *Genetics* 189 (3): 705–36. doi:10.1534/genetics.111.127019.
- Hampsey, M. 1998. "Molecular Genetics of the RNA Polymerase II General Transcriptional Machinery." *Microbiology and Molecular Biology Reviews* 62 (2): 465–503.
- Hampsey, M., and D. Reinberg. 2003. "Tails of Intrigue." *Cell* 113 (4): 429–32. doi:10.1016/S0092-8674(03)00360-X.
- Han, S. J., Y. C. Lee, B. S. Gim, G-H. Ryu, S. J. Park, W. S. Lane, and Y-J. Kim. 1999. "Activator-Specific Requirement of Yeast Mediator Proteins for RNA Polymerase II Transcriptional Activation." *Molecular and Cellular Biology* 19 (2): 979–88. doi:10.1128/MCB.19.2.979.
- Hartzog, G. A., T. Wada, H. Handa, and F. Winston. 1998. "Evidence That Spt4, Spt5, and Spt6 Control Transcription Elongation by RNA Polymerase II in Saccharomyces Cerevisiae." *Genes & Development* 12 (3): 357–69.
- Haruki, H., J. Nishikawa, and U. K. Laemmli. 2008. "The Anchor-Away Technique: Rapid, Conditional Establishment of Yeast Mutant Phenotypes." *Molecular Cell* 31 (6): 925–32. doi:10.1016/j.molcel.2008.07.020.
- Hashimoto, S., S. Boissel, M. Zarhrate, M. Rio, A. Munnich, J-M. Egly, and L. Colleaux. 2011. "MED23 Mutation Links Intellectual Disability to Dysregulation of Immediate Early Gene Expression." *Science (New York, N.Y.)* 333 (6046): 1161–63. doi:10.1126/science.1206638.
- He, Y., J. Fang, D. J. Taatjes, and E. Nogales. 2013. "Structural Visualization of Key Steps in Human Transcription Initiation." *Nature* 495 (7442): 481–86. doi:10.1038/nature11991.

- He, Y, C. Yan, J. Fang, C. Inouye, R. Tjian, I. Ivanov, and E. Nogales. 2016. "Near-Atomic Resolution Visualization of Human Transcription Promoter Opening." *Nature* 533 (7603): 359–65. doi:10.1038/nature17970.
- Heidemann, M., C. Hintermair, K. Voß, and D. Eick. 2013. "Dynamic Phosphorylation Patterns of RNA Polymerase II CTD during Transcription." *Biochimica Et Biophysica Acta* 1829 (1): 55–62. doi:10.1016/j.bbagr.2012.08.013.
- Hekmatpanah, D. S., and R. A. Young. 1991. "Mutations in a Conserved Region of RNA Polymerase II Influence the Accuracy of mRNA Start Site Selection." *Molecular and Cellular Biology* 11 (11): 5781–91. doi:10.1128/MCB.11.11.5781.
- Hengartner, C. J., C. M. Thompson, J. Zhang, D. M. Chao, S. M. Liao, A. J. Koleske, S. Okamura, and R. A. Young. 1995. "Association of an Activator with an RNA Polymerase II Holoenzyme." *Genes & Development* 9 (8): 897–910.
- Hengartner, C. J., V. E. Myer, S-M. Liao, C. J. Wilson, S. S. Koh, and R. A. Young. 1998. "Temporal Regulation of RNA Polymerase II by Srb10 and Kin28 Cyclin-Dependent Kinases." *Molecular Cell* 2 (1): 43–53. doi:10.1016/S1097-2765(00)80112-4.
- Henriques, T., D. A. Gilchrist, S. Nechaev, M. Bern, G. W. Muse, A. Burkholder, D. C. Fargo, and K. Adelman. 2013. "Stable Pausing by RNA Polymerase II Provides an Opportunity to Target and Integrate Regulatory Signals." *Molecular Cell* 52 (4): 517–28. doi:10.1016/j.molcel.2013.10.001.
- Henry, N. L., A. M. Campbell, W. J. Feaver, D. Poon, P. A. Weil, and R. D. Kornberg. 1994. "TFIIIF-TAF-RNA Polymerase II Connection." *Genes & Development* 8 (23): 2868–78. doi:10.1101/gad.8.23.2868.
- Hinow, P., C. E. Rogers, C. E. Barbieri, J. A. Pietenpol, A. K. Kenworthy, and E. DiBenedetto. 2006. "The DNA Binding Activity of p53 Displays Reaction-Diffusion Kinetics." *Biophysical Journal* 91 (1): 330–42. doi:10.1529/biophysj.105.078303.
- Hittelman, A. B., D. Burakov, J. A. Iñiguez-Lluhí, L. P. Freedman, and M. J. Garabedian. 1999. "Differential Regulation of Glucocorticoid Receptor Transcriptional Activation via AF-1-Associated Proteins." *The EMBO Journal* 18 (19): 5380–88. doi:10.1093/emboj/18.19.5380.
- Hocine, S., P. Raymond, D. Zenklusen, J. A. Chao, and R. H. Singer. 2013. "Single-Molecule Analysis of Gene Expression Using Two-Color RNA Labeling in Live Yeast." *Nature Methods* 10 (2): 119–21. doi:10.1038/nmeth.2305.
- Holstege, F. C., U. Fiedler, and H. T. Timmers. 1997. "Three Transitions in the RNA Polymerase II Transcription Complex during Initiation." *The EMBO Journal* 16 (24): 7468–80. doi:10.1093/emboj/16.24.7468.
- Holstege, F. C., E. G. Jennings, J. J. Wyrick, T. I. Lee, C. J. Hengartner, M. R. Green, T. R. Golub, E. S. Lander, and R. A. Young. 1998a. "Dissecting the Regulatory Circuitry of a Eukaryotic Genome." *Cell* 95 (5): 717–28.
- Hong, S-K., C. E. Haldin, N. D. Lawson, B. M. Weinstein, I. B. Dawid, and N. A. Hukriede. 2005. "The Zebrafish *kohtalo/trap230* Gene Is Required for the Development of the Brain, Neural Crest, and Pronephric Kidney." *Proceedings of the National Academy of Sciences of the United States of America* 102 (51): 18473–78. doi:10.1073/pnas.0509457102.
- Horikoshi, M., C. K. Wang, H. Fujii, J. A. Cromlish, P. A. Weil, and R. G. Roeder. 1989. "Cloning and Structure of a Yeast Gene Encoding a General Transcription Initiation Factor TFIID That Binds to the TATA Box." *Nature* 341 (6240): 299–303. doi:10.1038/341299a0.
- Houalla, R., F. Devaux, A. Fatica, J. Kufel, D. Barrass, C. Torchet, and D. Tollervey. 2006. "Microarray Detection of Novel Nuclear RNA Substrates for the Exosome." *Yeast (Chichester, England)* 23 (6): 439–54. doi:10.1002/yea.1369.
- Hsin, J-P., A. Sheth, and J. L. Manley. 2011. "RNAP II CTD Phosphorylated on Threonine-4 Is Required for Histone mRNA 3' End Processing." *Science* 334 (6056): 683–86. doi:10.1126/science.1206034.
- Hsu, J-M., J. Huang, P. B. Meluh, and B. C. Laurent. 2003. "The Yeast RSC Chromatin-Remodeling Complex Is Required for Kinetochores Function in Chromosome Segregation." *Molecular and Cellular Biology* 23 (9): 3202–15. doi:10.1128/MCB.23.9.3202-3215.2003.

Huang, Y., W. Li, X. Yao, Q.-J. Lin, J.-W. Yin, Y. Liang, M. Heiner, B. Tian, J. Hui, and G. Wang. 2012. "Mediator Complex Regulates Alternative mRNA Processing via the MED23 Subunit." *Molecular Cell* 45 (4): 459–69. doi:10.1016/j.molcel.2011.12.022.

Hull, M. W., K. McKune, and N. A. Woychik. 1995. "RNA Polymerase II Subunit RPB9 Is Required for Accurate Start Site Selection." *Genes & Development* 9 (4): 481–90. doi:10.1101/gad.9.4.481.

Imasaki, T., G. Calero, G. Cai, K.-L. Tsai, K. Yamada, F. Cardelli, H. Erdjument-Bromage, et al. 2011. "Architecture of the Mediator Head Module." *Nature* 475 (7355): 240–43. doi:10.1038/nature10162.

Imbalzano, A. N., K. S. Zaret, and R. E. Kingston. 1994. "Transcription Factor (TF) IIB and TFIIA Can Independently Increase the Affinity of the TATA-Binding Protein for DNA." *Journal of Biological Chemistry* 269 (11): 8280–86.

Ingolia, N. T., G. A. Brar, N. Stern-Ginossar, M. S. Harris, G. J. S. Talhouarne, S. E. Jackson, M.R. Wills, and J. S. Weissman. 2014. "Ribosome Profiling Reveals Pervasive Translation outside of Annotated Protein-Coding Genes." *Cell Reports* 8 (5): 1365–79. doi:10.1016/j.celrep.2014.07.045.

Ito, M., C. X. Yuan, H. J. Okano, R. B. Darnell, and R. G. Roeder. 2000. "Involvement of the TRAP220 Component of the TRAP/SMCC Coactivator Complex in Embryonic Development and Thyroid Hormone Action." *Molecular Cell* 5 (4): 683–93.

Ito, M., H. J. Okano, R. B. Darnell, and R. G. Roeder. 2002. "The TRAP100 Component of the TRAP/Mediator Complex Is Essential in Broad Transcriptional Events and Development." *The EMBO Journal* 21 (13): 3464–75. doi:10.1093/emboj/cdf348.

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." *Proceedings of the National Academy of Sciences* 98 (8): 4569–74. doi:10.1073/pnas.061034498.

Jacquier, A.. 2009. "The Complex Eukaryotic Transcriptome: Unexpected Pervasive Transcription and Novel Small RNAs." *Nature Reviews. Genetics* 10 (12): 833–44. doi:10.1038/nrg2683.

Janke, C., M.M. Magiera, N. Rathfelder, C. Taxis, S. Reber, H. Maekawa, A. Moreno-Borchart, et al. 2004. "A Versatile Toolbox for PCR-Based Tagging of Yeast Genes: New Fluorescent Proteins, More Markers and Promoter Substitution Cassettes." *Yeast (Chichester, England)* 21 (11): 947–62. doi:10.1002/yea.1142.

Jasiak, A. J., H. Hartmann, E. Karakasili, M. Kalocsay, A. Flatley, E. Kremmer, K. Strässer, D. E. Martin, J. Söding, and P. Cramer. 2008. "Genome-Associated RNA Polymerase II Includes the Dissociable Rpb4/7 Subcomplex." *Journal of Biological Chemistry* 283 (39): 26423–27. doi:10.1074/jbc.M803237200.

Jemal, A., F. Bray, M. M. Center, J. Ferlay, E. Ward, and D. Forman. 2011. "Global Cancer Statistics." *CA: A Cancer Journal for Clinicians* 61 (2): 69–90. doi:10.3322/caac.20107.

Jensen, T. H., A. Jacquier, and D. Libri. 2013. "Dealing with Pervasive Transcription." *Molecular Cell* 52 (4): 473–84. doi:10.1016/j.molcel.2013.10.032.

Jeronimo, C., and F. Robert. 2014. "Kin28 Regulates the Transient Association of Mediator with Core Promoters." *Nature Structural & Molecular Biology* 21 (5): 449–55. doi:10.1038/nsmb.2810.

Jiang, C., and B. F. Pugh. 2009. "A Compiled and Systematic Reference Map of Nucleosome Positions across the *Saccharomyces Cerevisiae* Genome." *Genome Biology* 10 (10): R109. doi:10.1186/gb-2009-10-10-r109.

———. 2009b. "Nucleosome Positioning and Gene Regulation: Advances through Genomics." *Nature Reviews. Genetics* 10 (3): 161–72. doi:10.1038/nrg2522.

Jing, H., C. R. Vakoc, L. Ying, S. Mandat, H. Wang, X. Zheng, and G. A. Blobel. 2008. "Exchange of GATA Factors Mediates Transitions in Looped Chromatin Organization at a Developmentally Regulated Gene Locus." *Molecular Cell* 29 (2): 232–42. doi:10.1016/j.molcel.2007.11.020.

Johnson, K. M., and M. Carey. 2003. "Assembly of a mediator/TFIID/TFIIA Complex Bypasses the Need for an Activator." *Current Biology: CB* 13 (9): 772–77.

Johnson, K. M., J. Wang, A. Smallwood, C. Arayata, and M. Carey. 2002. "TFIID and Human Mediator Coactivator Complexes Assemble Cooperatively on Promoter DNA." *Genes & Development* 16 (14): 1852–63. doi:10.1101/gad.995702.

Jonkers, I., and J. T. Lis. 2015. "Getting up to Speed with Transcription Elongation by RNA Polymerase II." *Nature Reviews. Molecular Cell Biology* 16 (3): 167–77. doi:10.1038/nrm3953.

Kagey, M. H., J. J. Newman, S. Bilodeau, Y. Zhan, D. A. Orlando, N. L. van Berkum, C. C. Ebmeier, et al. 2010. "Mediator and Cohesin Connect Gene Expression and Chromatin Architecture." *Nature* 467 (7314): 430–35. doi:10.1038/nature09380.

Kamakaka, R. T., and S. Biggins. 2005. "Histone Variants: Deviants?" *Genes & Development* 19 (3): 295–310. doi:10.1101/gad.1272805.

Kang, J. J., D. T. Auble, J. A. Ranish, and S. Hahn. 1995. "Analysis of the Yeast Transcription Factor TFIIA: Distinct Functional Regions and a Polymerase II-Specific Role in Basal and Activated Transcription." *Molecular and Cellular Biology* 15 (3): 1234–43. doi:10.1128/MCB.15.3.1234.

Kang, J. S., S.H. Kim, M. S. Hwang, S. J. Han, Y. Chul Lee, and Y.-Joon Kim. 2001. "The Structural and Functional Organization of the Yeast Mediator Complex." *Journal of Biological Chemistry* 276 (45): 42003–10. doi:10.1074/jbc.M105961200.

Kang, X., Y. Hu, Y. Li, X. Guo, X. Jiang, L. Lai, B. Xia, and C. Jin. 2006. "Structural, Biochemical, and Dynamic Characterizations of the hRPB8 Subunit of Human RNA Polymerases." *Journal of Biological Chemistry* 281 (26): 18216–26. doi:10.1074/jbc.M513241200.

Kang, Y. K., M. Guermah, C-X. Yuan, and R. G. Roeder. 2002. "The TRAP/Mediator Coactivator Complex Interacts Directly with Estrogen Receptors Alpha and Beta through the TRAP220 Subunit and Directly Enhances Estrogen Receptor Function in Vitro." *Proceedings of the National Academy of Sciences of the United States of America* 99 (5): 2642–47. doi:10.1073/pnas.261715899.

Kaplan, C. D., L. Laprade, and F. Winston. 2003. "Transcription Elongation Factors Repress Transcription Initiation from Cryptic Sites." *Science (New York, N.Y.)* 301 (5636): 1096–99. doi:10.1126/science.1087374.

Kapranov, P., J. Cheng, S. Dike, D. A. Nix, R. Dutttagupta, A. T. Willingham, P. F. Stadler, et al. 2007. "RNA Maps Reveal New RNA Classes and a Possible Function for Pervasive Transcription." *Science (New York, N.Y.)* 316 (5830): 1484–88. doi:10.1126/science.1138341.

Katayama, S., Y. Tomaru, T. Kasukawa, K. Waki, M. Nakanishi, M. Nakamura, H. Nishida, et al., RIKEN Genome Exploration Research Group, Genome Science Group (Genome Network Project Core Group), and FANTOM Consortium. 2005. "Antisense Transcription in the Mammalian Transcriptome." *Science (New York, N.Y.)* 309 (5740): 1564–66. doi:10.1126/science.1112009.

Kato, K., Y. Makino, T. Kishimoto, J. Yamauchi, S. Kato, M. Muramatsu, and T. Tamura. 1994. "Multimerization of the Mouse TATA-Binding Protein (TBP) Driven by Its C-Terminal Conserved Domain." *Nucleic Acids Research* 22 (7): 1179–85.

Kaufmann, J., and S. T. Smale. 1994. "Direct Recognition of Initiator Elements by a Component of the Transcription Factor IID Complex." *Genes & Development* 8 (7): 821–29. doi:10.1101/gad.8.7.821.

Kaufmann, R., R. Straussberg, H. Mandel, A. Fattal-Valevski, B. Ben-Zeev, A. Naamati, A. Shaag, et al. 2010. "Infantile Cerebral and Cerebellar Atrophy Is Associated with a Mutation in the MED17 Subunit of the Transcription Preinitiation Mediator Complex." *American Journal of Human Genetics* 87 (5): 667–70. doi:10.1016/j.ajhg.2010.09.016.

Kedinger, C., M. Gniazdowski, J. L. Mandel, F. Gissinger, and P. Chambon. 1970. "Alpha-Amanitin: A Specific Inhibitor of One of Two DNA-Dependent RNA Polymerase Activities from Calf Thymus." *Biochemical and Biophysical Research Communications* 38 (1): 165–71.

Kelleher, R. J., P. M. Flanagan, and R. D. Kornberg. 1990. "A Novel Mediator between Activator Proteins and the RNA Polymerase II Transcription Apparatus." *Cell* 61 (7): 1209–15.

Keogh, M-C., V. Podolny, and S. Buratowski. 2003. "Bur1 Kinase Is Required for Efficient Transcription Elongation by RNA Polymerase II." *Molecular and Cellular Biology* 23 (19): 7005–18. doi:10.1128/MCB.23.19.7005-7018.2003.

Khazak, V., J. Estojak, H. Cho, J. Majors, G. Sonoda, J. R. Testa, and E. A. Golemis. 1998. "Analysis of the Interaction of the Novel RNA Polymerase II (pol II) Subunit hRPB4 with Its Partner hRPB7 and with Pol II." *Molecular and Cellular Biology* 18 (4): 1935–45. doi:10.1128/MCB.18.4.1935.

Kikuchi, Y., H. Umemura, S. Nishitani, S. Iida, R. Fukasawa, H. Hayashi, Y. Hirose, A. Tanaka, K. Sugawara, and Y. Ohkuma. 2015. "Human Mediator MED17 Subunit Plays Essential Roles in Gene Regulation by Associating with the Transcription and DNA Repair Machineries." *Genes to Cells: Devoted to Molecular & Cellular Mechanisms* 20 (3): 191–202. doi:10.1111/gtc.12210.

- Kim, B., A. I. Nesvizhskii, P. G. Rani, S. Hahn, R. Aebersold, and J. A. Ranish. 2007. "The Transcription Elongation Factor TFIIS Is a Component of RNA Polymerase II Preinitiation Complexes." *Proceedings of the National Academy of Sciences of the United States of America* 104 (41): 16068–73. doi:10.1073/pnas.0704573104.
- Kim, M., N. J. Krogan, L. Vasiljeva, O. J. Rando, E. Nedeá, J. F. Greenblatt, and S. Buratowski. 2004. "The Yeast Rat1 Exonuclease Promotes Transcription Termination by RNA Polymerase II." *Nature* 432 (7016): 517–22. doi:10.1038/nature03041.
- Kim, Y. J., S. Björklund, Y. Li, M. H. Sayre, and R. D. Kornberg. 1994a. "A Multiprotein Mediator of Transcriptional Activation and Its Interaction with the C-Terminal Repeat Domain of RNA Polymerase II." *Cell* 77 (4): 599–608.
- Kim, Y., J. H. Geiger, S. Hahn, and P. B. Sigler. 1993. "Crystal Structure of a Yeast TBP/TATA-Box Complex." *Nature* 365 (6446): 512–20. doi:10.1038/365512a0.
- Kireeva, M. L., N. Komissarova, D. S. Waugh, and M. Kashlev. 2000. "The 8-Nucleotide-Long RNA:DNA Hybrid Is a Primary Stability Determinant of the RNA Polymerase II Elongation Complex." *The Journal of Biological Chemistry* 275 (9): 6530–36.
- Klein, C., and K. Struhl. 1994. "Increased Recruitment of TATA-Binding Protein to the Promoter by Transcriptional Activation Domains in Vivo." *Science* 266 (5183): 280–82. doi:10.1126/science.7939664.
- Knuesel, M. T., K. D. Meyer, C. Bernecky, and D. J. Taatjes. 2009. "The Human CDK8 Subcomplex Is a Molecular Switch That Controls Mediator Coactivator Function." *Genes & Development* 23 (4): 439–51. doi:10.1101/gad.1767009.
- Kobor, M. S., S. Venkatasubrahmanyam, M. D. Meneghini, J. W. Gin, J. L. Jennings, A. J. Link, H. D. Madhani, and J. Rine. 2004. "A Protein Complex Containing the Conserved Swi2/Snf2-Related ATPase Swr1p Deposits Histone Variant H2A.Z into Euchromatin." *PLoS Biology* 2 (5): E131. doi:10.1371/journal.pbio.0020131.
- Koch, F., F. Jourquin, P. Ferrier, and J-C. Andrau. 2008. "Genome-Wide RNA Polymerase II: Not Genes Only!" *Trends in Biochemical Sciences* 33 (6): 265–73. doi:10.1016/j.tibs.2008.04.006.
- Koh, S. S., A. Z. Ansari, M. Ptashne, and R. A. Young. 1998. "An Activator Target in the RNA Polymerase II Holoenzyme." *Molecular Cell* 1 (6): 895–904.
- Kokubo, T., M.J. Swanson, J-I. Nishikawa, A. G. Hinnebusch, and Y. Nakatani. 1998. "The Yeast TAF145 Inhibitory Domain and TFIIA Competitively Bind to TATA-Binding Protein." *Molecular and Cellular Biology* 18 (2): 1003–12. doi:10.1128/MCB.18.2.1003.
- Koleske, A. J., S. Buratowski, M. Nonet, and R. A. Young. 1992. "A Novel Transcription Factor Reveals a Functional Link between the RNA Polymerase II CTD and TFIID." *Cell* 69 (5): 883–94.
- Kolthur-Seetharam, Ullas, Igor Martjanov, and Irwin Davidson. 2008. "Specialization of the General Transcriptional Machinery in Male Germ Cells." *Cell Cycle (Georgetown, Tex.)* 7 (22): 3493–98. doi:10.4161/cc.7.22.6976.
- Komeili, A., and E. K O'Shea. 2000. "Nuclear Transport and Transcription." *Current Opinion in Cell Biology* 12 (3): 355–60. doi:10.1016/S0955-0674(00)00100-9.
- Korkhin, Y., U. M. Unligil, O. Littlefield, P. J. Nelson, D. I. Stuart, P. B. Sigler, S. D. Bell, and N. G. A. Abrescia. 2009. "Evolution of Complex RNA Polymerases: The Complete Archaeal RNA Polymerase Structure." *PLOS Biol* 7 (5): e1000102. doi:10.1371/journal.pbio.1000102.
- Kornberg, R. D., and Y. Lorch. 1999. "Twenty-Five Years of the Nucleosome, Fundamental Particle of the Eukaryote Chromosome." *Cell* 98 (3): 285–94. doi:10.1016/S0092-8674(00)81958-3.
- Koschubs, T., K. Lorenzen, S. Baumli, S. Sandström, A. J. R. Heck, and P. Cramer. 2010. "Preparation and Topology of the Mediator Middle Module." *Nucleic Acids Research* 38 (10): 3186–95. doi:10.1093/nar/gkq029.
- Koschubs, T., M. Seizl, L. Larivière, F. Kurth, S. Baumli, D. E. Martin, and P. Cramer. 2009. "Identification, Structure, and Functional Requirement of the Mediator Submodule Med7N/31." *The EMBO Journal* 28 (1): 69–80. doi:10.1038/emboj.2008.254.

Kostrewa, D., M. E. Zeller, K-J. Armache, M. Seizl, K. Leike, M. Thomm, and P. Cramer. 2009. "RNA Polymerase II-TFIIB Structure and Mechanism of Transcription Initiation." *Nature* 462 (7271): 323–30. doi:10.1038/nature08548.

Kouzarides, T. 2007. "Chromatin Modifications and Their Function." *Cell* 128 (4): 693–705. doi:10.1016/j.cell.2007.02.005.

Kowenz-Leutz, E., and A. Leutz. 1999. "A C/EBP Beta Isoform Recruits the SWI/SNF Complex to Activate Myeloid Genes." *Molecular Cell* 4 (5): 735–43.

Krebs, A. R., J. Demmers, K. Karmodiya, N-C. Chang, A. C. Chang, and L. Tora. 2010. "ATAC and Mediator Coactivators Form a Stable Complex and Regulate a Set of Non-Coding RNA Genes." *EMBO Reports* 11 (7): 541–47. doi:10.1038/embor.2010.75.

Kuehner, J. N., and D. A. Brow. 2006. "Quantitative Analysis of in Vivo Initiator Selection by Yeast RNA Polymerase II Supports a Scanning Model." *Journal of Biological Chemistry* 281 (20): 14119–28. doi:10.1074/jbc.M601937200.

Kuhn, C-D., S. R. Geiger, S. Baumli, M. Gartmann, J. Gerber, S. Jennebach, T. Mielke, H. Tschochner, R. Beckmann, and P. Cramer. 2007. "Functional Architecture of RNA Polymerase I." *Cell* 131 (7): 1260–72. doi:10.1016/j.cell.2007.10.051.

Kuldell, N. H., and S. Buratowski. 1997. "Genetic Analysis of the Large Subunit of Yeast Transcription Factor IIE Reveals Two Regions with Distinct Functions." *Molecular and Cellular Biology* 17 (9): 5288–98. doi:10.1128/MCB.17.9.5288.

Kuras, L., T. Borggreffe, and R. D. Kornberg. 2003. "Association of the Mediator Complex with Enhancers of Active Genes." *Proceedings of the National Academy of Sciences of the United States of America* 100 (24): 13887–91. doi:10.1073/pnas.2036346100.

Kuras, L., and K. Struhl. 1999. "Binding of TBP to Promoters in Vivo Is Stimulated by Activators and Requires Pol II Holoenzyme." *Nature* 399 (6736): 609–13. doi:10.1038/21239.

Kutach, A. K., and J. T. Kadonaga. 2000. "The Downstream Promoter Element DPE Appears To Be as Widely Used as the TATA Box in Drosophila Core Promoters." *Molecular and Cellular Biology* 20 (13): 4754–64. doi:10.1128/MCB.20.13.4754-4764.2000.

Kwak, H., N. J. Fuda, L. J. Core, and J. T. Lis. 2013. "Precise Maps of RNA Polymerase Reveal How Promoters Direct Initiation and Pausing." *Science (New York, N.Y.)* 339 (6122): 950–53. doi:10.1126/science.1229386.

Kwak, H., and J. T. Lis. 2013. "Control of Transcriptional Elongation." *Annual Review of Genetics* 47: 483–508. doi:10.1146/annurev-genet-110711-155440.

Kwon, J. Y., J. M. Park, B. S. Gim, S. J. Han, J. Lee, and Y. J. Kim. 1999. "Caenorhabditis Elegans Mediator Complexes Are Required for Developmental-Specific Transcriptional Activation." *Proceedings of the National Academy of Sciences of the United States of America* 96 (26): 14990–95.

Lacombe, T., S.L. Poh, R. Barbey, and L. Kuras. 2013. "Mediator Is an Intrinsic Component of the Basal RNA Polymerase II Machinery in Vivo." *Nucleic Acids Research* 41 (21): 9651–62. doi:10.1093/nar/gkt701.

Lagrange, T., A. N. Kapanidis, H. Tang, D. Reinberg, and R. H. Ebright. 1998. "New Core Promoter Element in RNA Polymerase II-Dependent Transcription: Sequence-Specific DNA Binding by Transcription Factor IIB." *Genes & Development* 12 (1): 34–44. doi:10.1101/gad.12.1.34.

Langer, D., I. Martianov, D. Alpern, M. Rhinn, C. Keime, P. Dollé, G. Mengus, and I. Davidson. 2016. "Essential Role of the TFIID Subunit TAF4 in Murine Embryogenesis and Embryonic Stem Cell Differentiation." *Nature Communications* 7: 11063. doi:10.1038/ncomms11063.

Langmead, B., C. Trapnell, M. Pop, and S. L. Salzberg. 2009. "Ultrafast and Memory-Efficient Alignment of Short DNA Sequences to the Human Genome." *Genome Biology* 10 (3): R25. doi:10.1186/gb-2009-10-3-r25.

Laprade, L., D. Rose, and F. Winston. 2007. "Characterization of New Spt3 and TATA-Binding Protein Mutants of Saccharomyces Cerevisiae: Spt3 TBP Allele-Specific Interactions and Bypass of Spt8." *Genetics* 177 (4): 2007–17. doi:10.1534/genetics.107.081976.

Larivière, L., S. Geiger, S. Hoepfner, S. Röther, K. Strässer, and P. Cramer. 2006. "Structure and TBP Binding of the Mediator Head Subcomplex Med8-Med18-Med20." *Nature Structural & Molecular Biology* 13 (10): 895–901. doi:10.1038/nsmb1143.

Larivière, L., C. Plaschka, M. Seizl, E. V. Petrotchenko, L. Wenzek, C. H. Borchers, and P. Cramer. 2013. "Model of the Mediator Middle Module Based on Protein Cross-Linking." *Nucleic Acids Research*, August, gkt704. doi:10.1093/nar/gkt704.

Larivière, L., C. Plaschka, M. Seizl, L. Wenzek, F. Kurth, and P. Cramer. 2012. "Structure of the Mediator Head Module." *Nature* 492 (7429): 448–51. doi:10.1038/nature11670.

Larivière, L., M. Seizl, and P. Cramer. 2012. "A Structural Perspective on Mediator Function." *Current Opinion in Cell Biology* 24 (3): 305–13. doi:10.1016/j.ceb.2012.01.007.

Larivière, L., M. Seizl, S. van Wageningen, S. Röther, L. van de Pasch, H. Feldmann, K. Sträßer, S. Hahn, F. C. P. Holstege, and P. Cramer. 2008. "Structure–system Correlation Identifies a Gene Regulatory Mediator Submodule." *Genes & Development* 22 (7): 872–77. doi:10.1101/gad.465108.

Laurent, B. C., I. Treich, and M. Carlson. 1993. "The Yeast SNF2/SWI2 Protein Has DNA-Stimulated ATPase Activity Required for Transcriptional Activation." *Genes & Development* 7 (4): 583–91.

Leal, A., K. Huehne, F. Bauer, H. Sticht, P. Berger, U. Suter, B. Morera, et al. 2009. "Identification of the Variant Ala335Val of MED25 as Responsible for CMT2B2: Molecular Data, Functional Studies of the SH3 Recognition Motif and Correlation between Wild-Type MED25 and PMP22 RNA Levels in CMT1A Animal Models." *Neurogenetics* 10 (4): 275–87. doi:10.1007/s10048-009-0183-3.

Lee, C-K, Y. Shibata, B. Rao, B. D. Strahl, and J. D. Lieb. 2004. "Evidence for Nucleosome Depletion at Active Regulatory Regions Genome-Wide." *Nature Genetics* 36 (8): 900–905. doi:10.1038/ng1400.

Lee, S., and S. Hahn. 1995. "Model for Binding of Transcription Factor TFIIB to the TBP-DNA Complex." *Nature* 376 (6541): 609–12. doi:10.1038/376609a0.

Lee, Y. C., J. M. Park, S. Min, S. J. Han, and Y. J. Kim. 1999. "An Activator Binding Module of Yeast RNA Polymerase II Holoenzyme." *Molecular and Cellular Biology* 19 (4): 2967–76.

Lemieux, K., and L. Gaudreau. 2004. "Targeting of Swi/Snf to the Yeast GAL1 UAS G Requires the Mediator, TAF IIs, and RNA Polymerase II." *The EMBO Journal* 23 (20): 4040–50. doi:10.1038/sj.emboj.7600416.

Lenstra, T. L., and D. R. Larson. 2016. "Single-Molecule mRNA Detection in Live Yeast." *Current Protocols in Molecular Biology / Edited by Frederick M. Ausubel ... [et Al.]* 113: 14.24.1–14.24.15. doi:10.1002/0471142727.mb1424s113.

Lenstra, T. L., J. Rodriguez, H. Chen, and D. R. Larson. 2016. "Transcription Dynamics in Living Cells." *Annual Review of Biophysics*, April. doi:10.1146/annurev-biophys-062215-010838.

Li, B., M. Carey, and J. L. Workman. 2007. "The Role of Chromatin during Transcription." *Cell* 128 (4): 707–19. doi:10.1016/j.cell.2007.01.015.

Li, H., B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin, and 1000 Genome Project Data Processing Subgroup. 2009. "The Sequence Alignment/Map Format and SAMtools." *Bioinformatics (Oxford, England)* 25 (16): 2078–79. doi:10.1093/bioinformatics/btp352.

Li, S. 2002. "Rpb4 and Rpb9 Mediate Subpathways of Transcription-Coupled DNA Repair in *Saccharomyces Cerevisiae*." *The EMBO Journal* 21 (21): 5921–29. doi:10.1093/emboj/cdf589.

Li, X-Y., A. Virbasius, X. Zhu, and M. R. Green. 1999. "Enhancement of TBP Binding by Activators and General Transcription Factors." *Nature* 399 (6736): 605–9. doi:10.1038/21232.

Li, Y., S. Bjorklund, Y. W. Jiang, Y. J. Kim, W. S. Lane, D. J. Stillman, and R. D. Kornberg. 1995. "Yeast Global Transcriptional Regulators Sin4 and Rgr1 Are Components of Mediator complex/RNA Polymerase II Holoenzyme." *Proceedings of the National Academy of Sciences of the United States of America* 92 (24): 10864–68.

Liao, S-M., J. Zhang, D. A. Jeffery, A. J. Koleske, C. M. Thompson, D. M. Chao, M. Viljoen, H. J. J. van Vuuren, and R. A. Young. 1995. "A Kinase–cyclin Pair in the RNA Polymerase II Holoenzyme." *Nature* 374 (6518): 193–96. doi:10.1038/374193a0.

Lickwar, C. R., F. Mueller, S. E. Hanlon, J. G. McNally, and J. D. Lieb. 2012. "Genome-Wide Protein-DNA Binding Dynamics Suggest a Molecular Clutch for Transcription Factor Function." *Nature* 484 (7393): 251–55. doi:10.1038/nature10985.

Lickwar, C. R., F. Mueller, and J. D. Lieb. 2013. "Genome-Wide Measurement of Protein-DNA Binding Dynamics Using Competition ChIP." *Nature Protocols* 8 (7): 1337–53. doi:10.1038/nprot.2013.077.

Lieberman, P. M., and A. J. Berk. 1994. "A Mechanism for TAFs in Transcriptional Activation: Activation Domain Enhancement of TFIID-TFIIA--Promoter DNA Complex Formation." *Genes & Development* 8 (9): 995–1005. doi:10.1101/gad.8.9.995.

Lieberman-Aiden, E., N. L. van Berkum, L. Williams, M. Imakaev, T. Ragozcy, A. Telling, I. Amit, et al. 2009. "Comprehensive Mapping of Long-Range Interactions Reveals Folding Principles of the Human Genome." *Science (New York, N.Y.)* 326 (5950): 289–93. doi:10.1126/science.1181369.

Lin, J. J., L. W. Lehmann, G. Bonora, R. Sridharan, A. A. Vashisht, N. Tran, K. Plath, J. A. Wohlschlegel, and M. Carey. 2011. "Mediator Coordinates PIC Assembly with Recruitment of CHD1." *Genes & Development* 25 (20): 2198–2209. doi:10.1101/gad.17554711.

Lin, X., L. Rinaldo, A. F. Fazly, and X. Xu. 2007. "Depletion of Med10 Enhances Wnt and Suppresses Nodal Signaling during Zebrafish Embryogenesis." *Developmental Biology* 303 (2): 536–48. doi:10.1016/j.ydbio.2006.11.034.

Lin, Y. C., and J. D. Gralla. 2005. "Stimulation of the XPB ATP-Dependent Helicase by the Beta Subunit of TFIIIE." *Nucleic Acids Research* 33 (9): 3072–81. doi:10.1093/nar/gki623.

Linder, T., and C. M. Gustafsson. 2004. "The Soh1/MED31 Protein Is an Ancient Component of Schizosaccharomyces Pombe and Saccharomyces Cerevisiae Mediator." *The Journal of Biological Chemistry* 279 (47): 49455–59. doi:10.1074/jbc.M409046200.

Linder, T., N. N. Rasmussen, C. O. Samuelsen, E. Chatzidaki, V. Baraznenok, J. Beve, P. Henriksen, C. M. Gustafsson, and S. Holmberg. 2008. "Two Conserved Modules of Schizosaccharomyces Pombe Mediator Regulate Distinct Cellular Pathways." *Nucleic Acids Research* 36 (8): 2489–2504. doi:10.1093/nar/gkn070.

Liu, X., M. Vorontchikhina, Y-L. Wang, F. Faiola, and E. Martinez. 2008. "STAGA Recruits Mediator to the MYC Oncoprotein to Stimulate Transcription and Cell Proliferation." *Molecular and Cellular Biology* 28 (1): 108–21. doi:10.1128/MCB.01402-07.

Liu, Y., J. A. Ranish, R. Aebersold, and S. Hahn. 2001. "Yeast Nuclear Extract Contains Two Major Forms of RNA Polymerase II Mediator Complexes." *The Journal of Biological Chemistry* 276 (10): 7169–75.

Liu, Y., C. Kung, J. Fishburn, A. Z. Ansari, K. M. Shokat, and S. Hahn. 2004. "Two Cyclin-Dependent Kinases Promote RNA Polymerase II Transcription and Formation of the Scaffold Complex." *Molecular and Cellular Biology* 24 (4): 1721–35.

Loncle, N., M. Boube, L. Joulia, C. Boschiero, M. Werner, D. L. Cribbs, and H-M. Bourbon. 2007. "Distinct Roles for Mediator Cdk8 Module Subunits in Drosophila Development." *The EMBO Journal* 26 (4): 1045–54. doi:10.1038/sj.emboj.7601566.

Lotan, R., V. Goler-Baron, L. Harel-Sharvit, L. Duek, D. Melamed, and M. Choder. 2005. "The RNA Polymerase II Subunit Rpb4p Mediates Decay of a Specific Class of mRNAs." *Genes & Development* 19 (24): 3004–16. doi:10.1101/gad.353205.

Lotan, R., V. Goler-Baron, L. Duek, G. Haimovich, and M. Choder. 2007. "The Rpb7p Subunit of Yeast RNA Polymerase II Plays Roles in the Two Major Cytoplasmic mRNA Decay Mechanisms." *The Journal of Cell Biology* 178 (7): 1133–43. doi:10.1083/jcb.200701165.

Louder, R. K., Y. He, J. R. López-Blanco, J. Fang, P. Chacón, and E. Nogales. 2016. "Structure of Promoter-Bound TFIID and Model of Human Pre-Initiation Complex Assembly." *Nature* 531 (7596): 604–9. doi:10.1038/nature17394.

Lovén, J., H. A Hoke, C. Y Lin, A. Lau, D. A Orlando, C. R Vakoc, J. E Bradner, T. I. Lee, and R. A Young. 2013. "Selective Inhibition of Tumor Oncogenes by Disruption of Super-Enhancers." *Cell* 153 (2): 320–34. doi:10.1016/j.cell.2013.03.036.

Luger, K., A. W. Mäder, R. K. Richmond, D. F. Sargent, and T. J. Richmond. 1997. "Crystal Structure of the Nucleosome Core Particle at 2.8 Å Resolution." *Nature* 389 (6648): 251–60. doi:10.1038/38444.

Luo, W., A. W. Johnson, and D. L. Bentley. 2006. "The Role of Rat1 in Coupling mRNA 3'-End Processing to Transcription Termination: Implications for a Unified Allosteric-Torpedo Model." *Genes & Development* 20 (8): 954–65. doi:10.1101/gad.1409106.

- Luscombe, N. M., S. E. Austin, H. M. Berman, and J. M. Thornton. 2000. "An Overview of the Structures of Protein-DNA Complexes." *Genome Biology* 1 (1): REVIEWS001. doi:10.1186/gb-2000-1-1-reviews001.
- Lykke-Andersen, S., and T. Heick Jensen. 2007. "Overlapping Pathways Dictate Termination of RNA Polymerase II Transcription." *Biochimie* 89 (10): 1177–82. doi:10.1016/j.biochi.2007.05.007.
- Ma, D., I. Olave, A. Merino, and D. Reinberg. 1996. "Separation of the Transcriptional Coactivator and Antirepression Functions of Transcription Factor IIA." *Proceedings of the National Academy of Sciences of the United States of America* 93 (13): 6583–88.
- Majello, B., and G. Napolitano. 2001. "Control of RNA Polymerase II Activity by Dedicated CTD Kinases and Phosphatases." *Frontiers in Bioscience: A Journal and Virtual Library* 6 (October): D1358–1368.
- Malagon, F., A. H. Tong, B. K. Shafer, and J. N. Strathern. 2004. "Genetic Interactions of DST1 in *Saccharomyces Cerevisiae* Suggest a Role of TFIIIS in the Initiation-Elongation Transition." *Genetics* 166 (3): 1215–27.
- Malik, S., W. Gu, W. Wu, J. Qin, and R. G. Roeder. 2000. "The USA-Derived Transcriptional Coactivator PC2 Is a Submodule of TRAP/SMCC and Acts Synergistically with Other PCs." *Molecular Cell* 5 (4): 753–60.
- Malik, S., M. Guermah, C.-X. Yuan, W. Wu, S. Yamamura, and R. G. Roeder. 2004. "Structural and Functional Organization of TRAP220, the TRAP/mediator Subunit That Is Targeted by Nuclear Receptors." *Molecular and Cellular Biology* 24 (18): 8244–54. doi:10.1128/MCB.24.18.8244-8254.2004.
- Malik, S., and R. G. Roeder. 2005. "Dynamic Regulation of Pol II Transcription by the Mammalian Mediator Complex." *Trends in Biochemical Sciences* 30 (5): 256–63. doi:10.1016/j.tibs.2005.03.009.
- . 2010. "The Metazoan Mediator Co-Activator Complex as an Integrative Hub for Transcriptional Regulation." *Nature Reviews. Genetics* 11 (11): 761–72. doi:10.1038/nrg2901.
- Malik, S., A. E. Wallberg, Y. K. Kang, and R. G. Roeder. 2002. "TRAP/SMCC/mediator-Dependent Transcriptional Activation from DNA and Chromatin Templates by Orphan Nuclear Receptor Hepatocyte Nuclear Factor 4." *Molecular and Cellular Biology* 22 (15): 5626–37.
- Marcus, G. A., N. Silverman, S. L. Berger, J. Horiuchi, and L. Guarente. 1994. "Functional Similarity and Physical Association between GCN5 and ADA2: Putative Transcriptional Adaptors." *The EMBO Journal* 13 (20): 4807–15.
- Marfella, C. G. A., and A. N. Imbalzano. 2007. "The Chd Family of Chromatin Remodelers." *Mutation Research* 618 (1-2): 30–40. doi:10.1016/j.mrfmmm.2006.07.012.
- Martens, J. A., L. Laprade, and F. Winston. 2004. "Intergenic Transcription Is Required to Repress the *Saccharomyces Cerevisiae* SER3 Gene." *Nature* 429 (6991): 571–74. doi:10.1038/nature02538.
- Martin, M.I. 2011. "Cutadapt Removes Adapter Sequences from High-Throughput Sequencing Reads." *EMBnet.journal* 17 (1): pp. 10–12.
- Matsui, T., J. Segall, P. A. Weil, and R. G. Roeder. 1980. "Multiple Factors Required for Accurate Initiation of Transcription by Purified RNA Polymerase II." *Journal of Biological Chemistry* 255 (24): 11992–96.
- Mavrich, T. N., I. P. Ioshikhes, B. J. Venters, C. Jiang, L. P. Tomsho, J. Qi, S. C. Schuster, I. Albert, and B. F. Pugh. 2008. "A Barrier Nucleosome Model for Statistical Positioning of Nucleosomes throughout the Yeast Genome." *Genome Research* 18 (7): 1073–83. doi:10.1101/gr.078261.108.
- Maxon, M. E., J. A. Goodrich, and R. Tjian. 1994a. "Transcription Factor IIE Binds Preferentially to RNA Polymerase Ila and Recruits TFIIH: A Model for Promoter Clearance." *Genes & Development* 8 (5): 515–24. doi:10.1101/gad.8.5.515.
- McCracken, S., N. Fong, E. Rosonina, K. Yankulov, G. Brothers, D. Siderovski, A. Hessel, et al. 1997. "5'-Capping Enzymes Are Targeted to Pre-mRNA by Binding to the Phosphorylated Carboxy-Terminal Domain of RNA Polymerase II." *Genes & Development* 11 (24): 3306–18. doi:10.1101/gad.11.24.3306.
- McCracken, S., N. Fong, K. Yankulov, S. Ballantyne, G. Pan, J. Greenblatt, S. D. Patterson, M. Wickens, and D. L. Bentley. 1997. "The C-Terminal Domain of RNA Polymerase II Couples mRNA Processing to Transcription." *Nature* 385 (6614): 357–61. doi:10.1038/385357a0.

- McKune, K., P. A. Moore, M. W. Hull, and N. A. Woychik. 1995. "Six Human RNA Polymerase Subunits Functionally Substitute for Their Yeast Counterparts." *Molecular and Cellular Biology* 15 (12): 6895–6900. doi:10.1128/MCB.15.12.6895.
- Meinhart, A., T. Kamenski, S. Hoepfner, S. Baumli, and P. Cramer. 2005. "A Structural Perspective of CTD Function." *Genes & Development* 19 (12): 1401–15. doi:10.1101/gad.1318105.
- Mencía, M., Z. Moqtaderi, J. V. Geisberg, L. Kuras, and K. Struhl. 2002. "Activator-Specific Recruitment of TFIID and Regulation of Ribosomal Protein Genes in Yeast." *Molecular Cell* 9 (4): 823–33. doi:10.1016/S1097-2765(02)00490-2.
- Meyer, K. D., S.-C. Lin, C. Bernecky, Y. Gao, and D. J. Taatjes. 2010. "p53 Activates Transcription by Directing Structural Shifts in Mediator." *Nature Structural & Molecular Biology* 17 (6): 753–60. doi:10.1038/nsmb.1816.
- Minakhin, L., S. Bhagat, A. Brunning, E. A. Campbell, S. A. Darst, R. H. Ebright, and K. Severinov. 2001. "Bacterial RNA Polymerase Subunit Ω and Eukaryotic RNA Polymerase Subunit RPB6 Are Sequence, Structural, and Functional Homologs and Promote RNA Polymerase Assembly." *Proceedings of the National Academy of Sciences* 98 (3): 892–97. doi:10.1073/pnas.98.3.892.
- Missra, A., and D. S. Gilmour. 2010. "Interactions between DSIF (DRB Sensitivity Inducing Factor), NELF (negative Elongation Factor), and the Drosophila RNA Polymerase II Transcription Elongation Complex." *Proceedings of the National Academy of Sciences* 107 (25): 11301–6. doi:10.1073/pnas.1000681107.
- Mitsuzawa, H., E. Kanda, and A. Ishihama. 2003. "Rpb7 Subunit of RNA Polymerase II Interacts with an RNA-binding Protein Involved in Processing of Transcripts." *Nucleic Acids Research* 31 (16): 4696–4701. doi:10.1093/nar/gkg688.
- Mizuguchi, G., T. Tsukiyama, J. Wisniewski, and C. Wu. 1997. "Role of Nucleosome Remodeling Factor NURF in Transcriptional Activation of Chromatin." *Molecular Cell* 1 (1): 141–50.
- Mizuguchi, G., X. Shen, J. Landry, W.-H. Wu, S. Sen, and C. Wu. 2004. "ATP-Driven Exchange of Histone H2AZ Variant Catalyzed by SWR1 Chromatin Remodeling Complex." *Science (New York, N.Y.)* 303 (5656): 343–48. doi:10.1126/science.1090701.
- Mizzen, C. A., X. J. Yang, T. Kokubo, J. E. Brownell, A. J. Bannister, T. Owen-Hughes, J. Workman, et al. 1996. "The TAF(II)250 Subunit of TFIID Has Histone Acetyltransferase Activity." *Cell* 87 (7): 1261–70.
- Mohibullah, N., and S. Hahn. 2008. "Site-Specific Cross-Linking of TBP in Vivo and in Vitro Reveals a Direct Functional Interaction with the SAGA Subunit Spt3." *Genes & Development* 22 (21): 2994–3006. doi:10.1101/gad.1724408.
- Mohrmann, L., and C. P. Verrijzer. 2005. "Composition and Functional Specificity of SWI2/SNF2 Class Chromatin Remodeling Complexes." *Biochimica et Biophysica Acta (BBA) - Gene Structure and Expression* 1681 (2–3): 59–73. doi:10.1016/j.bbaexp.2004.10.005.
- Morettini, S., M. Tribus, A. Zeilner, J. Sebald, B. Campo-Fernandez, G. Scheran, H. Wörle, V. Podhraski, D. V. Fyodorov, and A. Lusser. 2011. "The Chromodomains of CHD1 Are Critical for Enzymatic Activity but Less Important for Chromatin Localization." *Nucleic Acids Research* 39 (8): 3103–15. doi:10.1093/nar/gkq1298.
- Mortillaro, M. J., B. J. Blencowe, X. Wei, H. Nakayasu, L. Du, S. L. Warren, P. A. Sharp, and R. Berezney. 1996. "A Hyperphosphorylated Form of the Large Subunit of RNA Polymerase II Is Associated with Splicing Complexes and the Nuclear Matrix." *Proceedings of the National Academy of Sciences of the United States of America* 93 (16): 8253–57.
- Mukundan, B., and A. Ansari. 2011. "Novel Role for Mediator Complex Subunit Srb5/Med18 in Termination of Transcription." *The Journal of Biological Chemistry* 286 (43): 37053–57. doi:10.1074/jbc.C111.295915.
- Mullem, V. Van, M. Wery, M. Werner, J. Vandehaute, and P. Thuriaux. 2002. "The Rpb9 Subunit of RNA Polymerase II Binds Transcription Factor TFIIE and Interferes with the SAGA and Elongator Histone Acetyltransferases." *Journal of Biological Chemistry* 277 (12): 10220–25. doi:10.1074/jbc.M107207200.
- Muncke, N., C. Jung, H. Rüdiger, H. Ulmer, R. Roeth, A. Hubert, E. Goldmuntz, et al. 2003. "Missense Mutations and Gene Interruption in PROSIT240, a Novel TRAP240-like Gene, in Patients with

Congenital Heart Defect (transposition of the Great Arteries)." *Circulation* 108 (23): 2843–50. doi:10.1161/01.CIR.0000103684.77636.CD.

Munz, C., E. Psichari, D. Mandilis, A-C. Lavigne, M. Spiliotaki, T. Oehler, I. Davidson, L. Tora, P. Angel, and A. Pintzas. 2003. "TAF7 (TAFII55) Plays a Role in the Transcription Activation by c-Jun." *The Journal of Biological Chemistry* 278 (24): 21510–16. doi:10.1074/jbc.M212764200.

Murakami, K., H. Elmlund, N. Kalisman, D. A. Bushnell, C. M. Adams, M. Azubel, D. Elmlund, et al. 2013. "Architecture of an RNA Polymerase II Transcription Pre-Initiation Complex." *Science* 342 (6159): 1238724. doi:10.1126/science.1238724.

Muse, G. W., D. A. Gilchrist, S. Nechaev, R. Shah, J. S. Parker, S. F. Grissom, J. Zeitlinger, and K. Adelman. 2007. "RNA Polymerase Is Poised for Activation across the Genome." *Nature Genetics* 39 (12): 1507–11. doi:10.1038/ng.2007.21.

Myers, L. C., C. M. Gustafsson, D. A. Bushnell, M. Lui, H. Erdjument-Bromage, P. Tempst, and R. D. Kornberg. 1998. "The Med Proteins of Yeast and Their Function through the RNA Polymerase II Carboxy-Terminal Domain." *Genes & Development* 12 (1): 45–54.

Myers, L. C., C. M. Gustafsson, K. C. Hayashibara, P. O. Brown, and R. D. Kornberg. 1999. "Mediator Protein Mutations That Selectively Abolish Activated Transcription." *Proceedings of the National Academy of Sciences of the United States of America* 96 (1): 67–72.

Näär, A. M., P. A. Beurang, S. Zhou, S. Abraham, W. Solomon, and R. Tjian. 1999. "Composite Co-Activator ARC Mediates Chromatin-Directed Transcriptional Activation." *Nature* 398 (6730): 828–32. doi:10.1038/19789.

Näär, A. M., B. D. Lemon, and R. Tjian. 2001. "Transcriptional Coactivator Complexes." *Annual Review of Biochemistry* 70: 475–501. doi:10.1146/annurev.biochem.70.1.475.

Nagalakshmi, U., Z. Wang, K. Waern, C. Shou, D. Raha, M. Gerstein, and M. Snyder. 2008. "The Transcriptional Landscape of the Yeast Genome Defined by RNA Sequencing." *Science (New York, N.Y.)* 320 (5881): 1344–49. doi:10.1126/science.1158441.

Nair, D., Y. Kim, and L. C. Myers. 2005. "Mediator and TFIID Govern Carboxyl-Terminal Domain-Dependent Transcription in Yeast Extracts." *The Journal of Biological Chemistry* 280 (40): 33739–48. doi:10.1074/jbc.M506067200.

Nakamura, Y., K. Yamamoto, X. He, B. Otsuki, Y. Kim, H. Murao, T. Soeda, et al. 2011. "Wwp2 Is Essential for Palatogenesis Mediated by the Interaction between Sox9 and Mediator Subunit 25." *Nature Communications* 2: 251. doi:10.1038/ncomms1242.

Nalley, K., S. A. Johnston, and T. Kodadek. 2006. "Proteolytic Turnover of the Gal4 Transcription Factor Is Not Required for Function in Vivo." *Nature* 442 (7106): 1054–57. doi:10.1038/nature05067.

Narita, T., Y. Yamaguchi, K. Yano, S. Sugimoto, S. Chanarat, T. Wada, D-K. Kim, et al. 2003. "Human Transcription Elongation Factor NELF: Identification of Novel Subunits and Reconstitution of the Functionally Active Complex." *Molecular and Cellular Biology* 23 (6): 1863–73. doi:10.1128/MCB.23.6.1863-1873.2003.

Natarajan, K., B. M. Jackson, H. Zhou, F. Winston, and A. G. Hinnebusch. 1999. "Transcriptional Activation by Gcn4p Involves Independent Interactions with the SWI/SNF Complex and the SRB/mediator." *Molecular Cell* 4 (4): 657–64.

Neil, H., C. Malabat, Y. d'Aubenton-Carafa, Z. Xu, L. M. Steinmetz, and A. Jacquier. 2009. "Widespread Bidirectional Promoters Are the Major Source of Cryptic Transcripts in Yeast." *Nature* 457 (7232): 1038–42. doi:10.1038/nature07747.

Ni, Z., B. E. Schwartz, J. Werner, J-R. Suarez, and J. T. Lis. 2004. "Coordination of Transcription, RNA Processing, and Surveillance by P-TEFb Kinase on Heat Shock Genes." *Molecular Cell* 13 (1): 55–65. doi:10.1016/S1097-2765(03)00526-4.

Nikolov, D. B., H. Chen, E. D. Halay, A. Hoffman, R. G. Roeder, and S. K. Burley. 1996. "Crystal Structure of a Human TATA Box-Binding protein/TATA Element Complex." *Proceedings of the National Academy of Sciences of the United States of America* 93 (10): 4862–67.

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 (6545): 119–28. doi:10.1038/377119a0.

Nishimura, K., T. Fukagawa, H. Takisawa, T. Kakimoto, and M. Kanemaki. 2009. "An Auxin-Based Degron System for the Rapid Depletion of Proteins in Nonplant Cells." *Nature Methods* 6 (12): 917–22. doi:10.1038/nmeth.1401.

Nock, A., J. M. Ascano, M. J. Barrero, and S. Malik. 2012. "Mediator-Regulated Transcription through the +1 Nucleosome." *Molecular Cell* 48 (6): 837–48. doi:10.1016/j.molcel.2012.10.009.

Nonet, M. L., and R. A. Young. 1989. "Intragenic and Extragenic Suppressors of Mutations in the Heptapeptide Repeat Domain of *Saccharomyces Cerevisiae* RNA Polymerase II." *Genetics* 123 (4): 715–24.

Nonet, M., D. Sweetser, and R. A. Young. 1987. "Functional Redundancy and Structural Polymorphism in the Large Subunit of RNA Polymerase II." *Cell* 50 (6): 909–15. doi:10.1016/0092-8674(87)90517-4.

Oda, Y., R. J. Chalkley, A. L. Burlingame, and D. D. Bikle. 2010. "The Transcriptional Coactivator DRIP/mediator Complex Is Involved in Vitamin D Receptor Function and Regulates Keratinocyte Proliferation and Differentiation." *The Journal of Investigative Dermatology* 130 (10): 2377–88. doi:10.1038/jid.2010.148.

Oda, Y., M. H. Ishikawa, N. P. Hawker, Q-C. Yun, and D. D. Bikle. 2007. "Differential Role of Two VDR Coactivators, DRIP205 and SRC-3, in Keratinocyte Proliferation and Differentiation." *The Journal of Steroid Biochemistry and Molecular Biology* 103 (3-5): 776–80. doi:10.1016/j.jsbmb.2006.12.069.

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." *Molecular and Cellular Biology* 15 (9): 4856–66.

Ohkuma, Y., and R. G. Roeder. 1994. "Regulation of TFIIH ATPase and Kinase Activities by TFIIE during Active Initiation Complex Formation." *Nature* 368 (6467): 160–63. doi:10.1038/368160a0.

Ong, S-E, L. J. Foster, and M. Mann. 2003. "Mass Spectrometric-Based Approaches in Quantitative Proteomics." *Methods (San Diego, Calif.)* 29 (2): 124–30.

Orlicky, S. M., P. T. Tran, M. H. Sayre, and A. M. Edwards. 2001. "Dissociable Rpb4-Rpb7 Subassembly of RNA Polymerase II Binds to Single-Strand Nucleic Acid and Mediates a Post-Recruitment Step in Transcription Initiation." *Journal of Biological Chemistry* 276 (13): 10097–102. doi:10.1074/jbc.M003165200.

Orphanides, G., and D. Reinberg. 2002. "A Unified Theory of Gene Expression." *Cell* 108 (4): 439–51. doi:10.1016/S0092-8674(02)00655-4.

Osheim, Y. N., N. J. Proudfoot, and A. L. Beyer. 1999. "EM Visualization of Transcription by RNA Polymerase II: Downstream Termination Requires a poly(A) Signal but Not Transcript Cleavage." *Molecular Cell* 3 (3): 379–87.

Ottaviani, A., E. Gilson, and F. Magdinier. 2008. "Telomeric Position Effect: From the Yeast Paradigm to Human Pathologies?" *Biochimie* 90 (1): 93–107. doi:10.1016/j.biochi.2007.07.022.

Owen, D. J., P. Ornaghi, J. C. Yang, N. Lowe, P. R. Evans, P. Ballario, D. Neuhaus, P. Filetici, and A. A. Travers. 2000. "The Structural Basis for the Recognition of Acetylated Histone H4 by the Bromodomain of Histone Acetyltransferase *gcn5p*." *The EMBO Journal* 19 (22): 6141–49. doi:10.1093/emboj/19.22.6141.

Owen-Hughes, T., and J. L. Workman. 1996. "Remodeling the Chromatin Structure of a Nucleosome Array by Transcription Factor-Targeted Trans-Displacement of Histones." *The EMBO Journal* 15 (17): 4702–12.

Pal, M., A. S. Ponticelli, and D. S. Luse. 2005. "The Role of the Transcription Bubble and TFIIB in Promoter Clearance by RNA Polymerase II." *Molecular Cell* 19 (1): 101–10. doi:10.1016/j.molcel.2005.05.024.

Paoletti, A. C., T. J. Parmely, C. Tomomori-Sato, S. Sato, D. Zhu, R. C. Conaway, J. W. Conaway, L. Florens, and M. P. Washburn. 2006. "Quantitative Proteomic Analysis of Distinct Mammalian Mediator Complexes Using Normalized Spectral Abundance Factors." *Proceedings of the National Academy of Sciences of the United States of America* 103 (50): 18928–33. doi:10.1073/pnas.0606379103.

Papai, G., M. K. Tripathi, C. Ruhlmann, J. H. Layer, P. A. Weil, and P. Schultz. 2010. "TFIIA and the Transactivator Rap1 Cooperate to Commit TFIID for Transcription Initiation." *Nature* 465 (7300): 956–60. doi:10.1038/nature09080.

- Park, D., Y. Lee, G. Bhupindersingh, and V. R. Iyer. 2013. "Widespread Misinterpretable ChIP-Seq Bias in Yeast." *PloS One* 8 (12): e83506. doi:10.1371/journal.pone.0083506.
- Park, J. M., B. S. Gim, J. M. Kim, J. H. Yoon, H. S. Kim, J. G. Kang, and Y. J. Kim. 2001. "Drosophila Mediator Complex Is Broadly Utilized by Diverse Gene-Specific Transcription Factors at Different Types of Core Promoters." *Molecular and Cellular Biology* 21 (7): 2312–23. doi:10.1128/MCB.21.7.2312-2323.2001.
- Park, J. M., J. Werner, J. M. Kim, J. T. Lis, and Y. J. Kim. 2001. "Mediator, Not Holoenzyme, Is Directly Recruited to the Heat Shock Promoter by HSF upon Heat Shock." *Molecular Cell* 8 (1): 9–19.
- Park, Jin Mo, Hye-Suk Kim, Sang Jun Han, Moon-Sun Hwang, Young Chul Lee, and Young-Joon Kim. 2000. "In Vivo Requirement of Activator-Specific Binding Targets of Mediator." *Molecular and Cellular Biology* 20 (23): 8709–19. doi:10.1128/MCB.20.23.8709-8719.2000.
- Park, S. W., G. Li, Y-P. Lin, M. J. Barrero, K. Ge, R. G. Roeder, and L-N. Wei. 2005. "Thyroid Hormone-Induced Juxtaposition of Regulatory Elements/factors and Chromatin Remodeling of Crabp1 Dependent on MED1/TRAP220." *Molecular Cell* 19 (5): 643–53. doi:10.1016/j.molcel.2005.08.008.
- Pei, Y., B. Schwer, and S. Shuman. 2003. "Interactions between Fission Yeast Cdk9, Its Cyclin Partner Pch1, and mRNA Capping Enzyme Pct1 Suggest an Elongation Checkpoint for mRNA Quality Control." *The Journal of Biological Chemistry* 278 (9): 7180–88. doi:10.1074/jbc.M211713200.
- Peng, J., and J-Q. Zhou. 2012. "The Tail-Module of Yeast Mediator Complex Is Required for Telomere Heterochromatin Maintenance." *Nucleic Acids Research* 40 (2): 581–93. doi:10.1093/nar/gkr757.
- Peyroche, G., E. Levillain, M. Siaux, I. Callebaut, P. Schultz, A. Sentenac, M. Riva, and C. Carles. 2002. "The A14–A43 Heterodimer Subunit in Yeast RNA Pol I and Their Relationship to Rpb4–Rpb7 Pol II Subunits." *Proceedings of the National Academy of Sciences* 99 (23): 14670–75. doi:10.1073/pnas.232580799.
- Phair, R. D., and T. Misteli. 2001. "Kinetic Modelling Approaches to in Vivo Imaging." *Nature Reviews Molecular Cell Biology* 2 (12): 898–907. doi:10.1038/35103000.
- Phair, R. D., P. Scaffidi, C. Elbi, J. Vecerová, A. Dey, K. Ozato, D. T. Brown, G. Hager, M. Bustin, and T. Misteli. 2004. "Global Nature of Dynamic Protein-Chromatin Interactions in Vivo: Three-Dimensional Genome Scanning and Dynamic Interaction Networks of Chromatin Proteins." *Molecular and Cellular Biology* 24 (14): 6393–6402. doi:10.1128/MCB.24.14.6393-6402.2004.
- Pinto, I., W. H. Wu, J. G. Na, and M. Hampsey. 1994. "Characterization of sua7 Mutations Defines a Domain of TFIIB Involved in Transcription Start Site Selection in Yeast." *Journal of Biological Chemistry* 269 (48): 30569–73.
- Pinto, I., D. E. Ware, and M. Hampsey. 1992. "The Yeast SUA7 Gene Encodes a Homolog of Human Transcription Factor TFIIB and Is Required for Normal Start Site Selection in Vivo." *Cell* 68 (5): 977–88. doi:10.1016/0092-8674(92)90040-J.
- Plank, J. L., and A. Dean. 2014. "Enhancer Function: Mechanistic and Genome-Wide Insights Come Together." *Molecular Cell* 55 (1): 5–14. doi:10.1016/j.molcel.2014.06.015.
- Plaschka, C., M. Hantsche, C. Dienemann, C. Burzinski, J. Plitzko, and P. Cramer. 2016. "Transcription Initiation Complex Structures Elucidate DNA Opening." *Nature* 533 (7603): 353–58. doi:10.1038/nature17990.
- Plaschka, C., L. Larivière, L. Wenzek, M. Seizl, M. Hemann, D. Tegunov, E. V. Petrotchenko, et al. 2015. "Architecture of the RNA Polymerase II-Mediator Core Initiation Complex." *Nature* 518 (7539): 376–80. doi:10.1038/nature14229.
- Porrua, O., and D. Libri. 2015. "Transcription Termination and the Control of the Transcriptome: Why, Where and How to Stop." *Nature Reviews. Molecular Cell Biology* 16 (3): 190–202. doi:10.1038/nrm3943.
- Poss, Z. C., C. C. Ebmeier, A. T. Odell, A. Tangpeerachaikul, T. Lee, H. E. Pelish, M. D. Shair, R. D. Dowell, W. M. Old, and D. J. Taatjes. 2016. "Identification of Mediator Kinase Substrates in Human Cells Using Cortistatin A and Quantitative Phosphoproteomics." *Cell Reports* 15 (2): 436–50. doi:10.1016/j.celrep.2016.03.030.

- Poss, Z. C., C. C. Ebmeier, and D. J. Taatjes. 2013. "The Mediator Complex and Transcription Regulation." *Critical Reviews in Biochemistry and Molecular Biology* 48 (6): 575–608. doi:10.3109/10409238.2013.840259.
- Prather, D. M., E. Larschan, and F. Winston. 2005. "Evidence That the Elongation Factor TFIIS Plays a Role in Transcription Initiation at GAL1 in *Saccharomyces Cerevisiae*." *Molecular and Cellular Biology* 25 (7): 2650–59. doi:10.1128/MCB.25.7.2650-2659.2005.
- Pray-Grant, M. G., D. Schieltz, S. J. McMahon, J. M. Wood, E. L. Kennedy, R. G. Cook, J. L. Workman, J. R. Yates, and P. A. Grant. 2002. "The Novel SLIK Histone Acetyltransferase Complex Functions in the Yeast Retrograde Response Pathway." *Molecular and Cellular Biology* 22 (24): 8774–86.
- Preker, P., J. Nielsen, S. Kammler, S. Lykke-Andersen, M. S. Christensen, C. K. Mapendano, M. H. Schierup, and T. H. Jensen. 2008. "RNA Exosome Depletion Reveals Transcription Upstream of Active Human Promoters." *Science (New York, N.Y.)* 322 (5909): 1851–54. doi:10.1126/science.1164096.
- Pugh, B. F., and R. Tjian. 1990. "Mechanism of Transcriptional Activation by Sp1: Evidence for Coactivators." *Cell* 61 (7): 1187–97. doi:10.1016/0092-8674(90)90683-6.
- Qiu, H., C. Hu, F. Zhang, G. Jiunn Hwang, M. J. Swanson, C. Boonchird, and A. G. Hinnebusch. 2005. "Interdependent Recruitment of SAGA and Srb Mediator by Transcriptional Activator Gcn4p." *Molecular and Cellular Biology* 25 (9): 3461–74. doi:10.1128/MCB.25.9.3461-3474.2005.
- Quinlan, A. R., and I. M. Hall. 2010. "BEDTools: A Flexible Suite of Utilities for Comparing Genomic Features." *Bioinformatics (Oxford, England)* 26 (6): 841–42. doi:10.1093/bioinformatics/btq033.
- Rachez, C., and L. P. Freedman. 2001. "Mediator Complexes and Transcription." *Current Opinion in Cell Biology* 13 (3): 274–80.
- Rachez, C., B. D. Lemon, Z. Suldan, V. Bromleigh, M. Gamble, A. M. Näär, H. Erdjument-Bromage, P. Tempst, and L. P. Freedman. 1999. "Ligand-Dependent Transcription Activation by Nuclear Receptors Requires the DRIP Complex." *Nature* 398 (6730): 824–28. doi:10.1038/19783.
- Rachez, C., Z. Suldan, J. Ward, C. P. Chang, D. Burakov, H. Erdjument-Bromage, P. Tempst, and L. P. Freedman. 1998. "A Novel Protein Complex That Interacts with the Vitamin D3 Receptor in a Ligand-Dependent Manner and Enhances VDR Transactivation in a Cell-Free System." *Genes & Development* 12 (12): 1787–1800.
- Raj, A., C. S. Peskin, D. Tranchina, D. Y. Vargas, and S. Tyagi. 2006. "Stochastic mRNA Synthesis in Mammalian Cells." *PLoS Biology* 4 (10): e309. doi:10.1371/journal.pbio.0040309.
- Rana, R., S. Surapureddi, W. Kam, S. Ferguson, and J. A. Goldstein. 2011. "Med25 Is Required for RNA Polymerase II Recruitment to Specific Promoters, Thus Regulating Xenobiotic and Lipid Metabolism in Human Liver." *Molecular and Cellular Biology* 31 (3): 466–81. doi:10.1128/MCB.00847-10.
- Ranish, J. A., and S. Hahn. 1991. "The Yeast General Transcription Factor TFIIA Is Composed of Two Polypeptide Subunits." *Journal of Biological Chemistry* 266 (29): 19320–27.
- Ranish, J. A., N. Yudkovsky, and S. Hahn. 1999. "Intermediates in Formation and Activity of the RNA Polymerase II Preinitiation Complex: Holoenzyme Recruitment and a Postrecruitment Role for the TATA Box and TFIIB." *Genes & Development* 13 (1): 49–63.
- Ranuncolo, S. M., S. Ghosh, J. A. Hanover, G. W. Hart, and B. A. Lewis. 2012. "Evidence of the Involvement of O-GlcNAc-Modified Human RNA Polymerase II CTD in Transcription in Vitro and in Vivo." *Journal of Biological Chemistry* 287 (28): 23549–61. doi:10.1074/jbc.M111.330910.
- Rau, M. J., S. Fischer, and C. J. Neumann. 2006. "Zebrafish Trap230/Med12 Is Required as a Coactivator for Sox9-Dependent Neural Crest, Cartilage and Ear Development." *Developmental Biology* 296 (1): 83–93. doi:10.1016/j.ydbio.2006.04.437.
- Reeves, W. M., and S. Hahn. 2003. "Activator-Independent Functions of the Yeast Mediator sin4 Complex in Preinitiation Complex Formation and Transcription Reinitiation." *Molecular and Cellular Biology* 23 (1): 349–58.
- . 2005. "Targets of the Gal4 Transcription Activator in Functional Transcription Complexes." *Molecular and Cellular Biology* 25 (20): 9092–9102. doi:10.1128/MCB.25.20.9092-9102.2005.
- Reynolds, T. B., and G. R. Fink. 2001. "Bakers' Yeast, a Model for Fungal Biofilm Formation." *Science* 291 (5505): 878–81. doi:10.1126/science.291.5505.878.

Rhee, H. S., A. R. Bataille, L. Zhang, and B. F. Pugh. 2014. "Subnucleosomal Structures and Nucleosome Asymmetry across a Genome." *Cell* 159 (6): 1377–88. doi:10.1016/j.cell.2014.10.054.

Rhee, H. S., and B. F. Pugh. 2012. "Genome-Wide Structure and Organization of Eukaryotic Pre-Initiation Complexes." *Nature* 483 (7389): 295–301. doi:10.1038/nature10799.

Richard, P., and J. L. Manley. 2009. "Transcription Termination by Nuclear RNA Polymerases." *Genes & Development* 23 (11): 1247–69. doi:10.1101/gad.1792809.

Rigaut, G., A. Shevchenko, B. Rutz, M. Wilm, M. Mann, and B. Séraphin. 1999. "A Generic Protein Purification Method for Protein Complex Characterization and Proteome Exploration." *Nature Biotechnology* 17 (10): 1030–32. doi:10.1038/13732.

Risheg, H., J. M. Graham, R. D. Clark, R. C. Rogers, J. M. Opitz, J.B. Moeschler, A. P. Peiffer, et al. 2007. "A Recurrent Mutation in MED12 Leading to R961W Causes Opitz-Kaveggia Syndrome." *Nature Genetics* 39 (4): 451–53. doi:10.1038/ng1992.

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." *Molecular Cell* 2 (3): 341–51. doi:10.1016/S1097-2765(00)80278-6.

Robert, F., D. Forget, J. Li, J. Greenblatt, and B. Coulombe. 1996. "Localization of Subunits of Transcription Factors IIE and IIF Immediately Upstream of the Transcriptional Initiation Site of the Adenovirus Major Late Promoter." *Journal of Biological Chemistry* 271 (15): 8517–20. doi:10.1074/jbc.271.15.8517.

Robinson, P. J. J., D. A. Bushnell, M. J. Trnka, A. L. Burlingame, and R. D. Kornberg. 2012. "Structure of the Mediator Head Module Bound to the Carboxy-Terminal Domain of RNA Polymerase II." *Proceedings of the National Academy of Sciences of the United States of America* 109 (44): 17931–35. doi:10.1073/pnas.1215241109.

Robinson, P. J., M. J. Trnka, R. Pellarin, C. H. Greenberg, D. A. Bushnell, R. Davis, A. L. Burlingame, A. Sali, and R. D. Kornberg. 2015. "Molecular Architecture of the Yeast Mediator Complex." *eLife* 4 (September): e08719. doi:10.7554/eLife.08719.

Rocha, P. P., M. Scholze, W. Bleiss, and H. Schrewe. 2010. "Med12 Is Essential for Early Mouse Development and for Canonical Wnt and Wnt/PCP Signaling." *Development (Cambridge, England)* 137 (16): 2723–31. doi:10.1242/dev.053660.

Rochette-Egly, C., S. Adam, M. Rossignol, J. M. Egly, and P. Chambon. 1997. "Stimulation of RAR Alpha Activation Function AF-1 through Binding to the General Transcription Factor TFIIF and Phosphorylation by CDK7." *Cell* 90 (1): 97–107.

Rodríguez-Navarro, S.. 2009. "Insights into SAGA Function during Gene Expression." *EMBO Reports* 10 (8): 843–50. doi:10.1038/embor.2009.168.

Roeder, R. G., and W. J. Rutter. 1969. "Multiple Forms of DNA-Dependent RNA Polymerase in Eukaryotic Organisms." *Nature* 224 (5216): 234–37.

Ryu, S., S. Zhou, A. G. Ladurner, and R. Tjian. 1999. "The Transcriptional Cofactor Complex CRSP Is Required for Activity of the Enhancer-Binding Protein Sp1." *Nature* 397 (6718): 446–50. doi:10.1038/17141.

Sadhale, P. P., and N. A. Woychik. 1994. "C25, an Essential RNA Polymerase III Subunit Related to the RNA Polymerase II Subunit RPB7." *Molecular and Cellular Biology* 14 (9): 6164–70. doi:10.1128/MCB.14.9.6164.

Sadowski, I., J. Ma, S. Triezenberg, and M. Ptashne. 1988. "GAL4-VP16 Is an Unusually Potent Transcriptional Activator." *Nature* 335 (6190): 563–64. doi:10.1038/335563a0.

Sainsbury, S., C. Bernecky, and P. Cramer. 2015. "Structural Basis of Transcription Initiation by RNA Polymerase II." *Nature Reviews Molecular Cell Biology* 16 (3): 129–43. doi:10.1038/nrm3952.

Sainsbury, S., J. Niesser, and P. Cramer. 2013. "Structure and Function of the Initially Transcribing RNA Polymerase II-TFIIB Complex." *Nature* 493 (7432): 437–40. doi:10.1038/nature11715.

Sakurai, H., Y. J. Kim, T. Ohishi, R. D. Kornberg, and T. Fukasawa. 1996. "The Yeast GAL11 Protein Binds to the Transcription Factor IIE through GAL11 Regions Essential for Its in Vivo Function." *Proceedings of the National Academy of Sciences of the United States of America* 93 (18): 9488–92.

Sakurai, H., and T. Fukasawa. 2003. "Artificial Recruitment of Certain Mediator Components Affects Requirement of Basal Transcription Factor IIE." *Genes to Cells: Devoted to Molecular & Cellular Mechanisms* 8 (1): 41–50.

Salghetti, S. E., A. A. Caudy, J. G. Chenoweth, and W. P. Tansey. 2001. "Regulation of Transcriptional Activation Domain Function by Ubiquitin." *Science (New York, N.Y.)* 293 (5535): 1651–53. doi:10.1126/science.1062079.

Samuels, M., A. Fire, and P. A. Sharp. 1982. "Separation and Characterization of Factors Mediating Accurate Transcription by RNA Polymerase II." *Journal of Biological Chemistry* 257 (23): 14419–27.

Sarker, A. H., S. E. Tsutakawa, S. Kostek, C. Ng, D. S. Shin, M. Peris, E. Campeau, J. A. Tainer, E. Nogales, and P. K. Cooper. 2005. "Recognition of RNA Polymerase II and Transcription Bubbles by XPG, CSB, and TFIIH: Insights for Transcription-Coupled Repair and Cockayne Syndrome." *Molecular Cell* 20 (2): 187–98. doi:10.1016/j.molcel.2005.09.022.

Sato, S., C. Tomomori-Sato, T. J. Parmely, L. Florens, B. Zybaylov, S. K. Swanson, C. A. S. Banks, et al. 2004. "A Set of Consensus Mammalian Mediator Subunits Identified by Multidimensional Protein Identification Technology." *Molecular Cell* 14 (5): 685–91. doi:10.1016/j.molcel.2004.05.006.

Sawadogo, M., and R. G. Roeder. 1985. "Interaction of a Gene-Specific Transcription Factor with the Adenovirus Major Late Promoter Upstream of the TATA Box Region." *Cell* 43 (1): 165–75. doi:10.1016/0092-8674(85)90021-2.

Schiano, C., A. Casamassimi, M. Rienzo, F. de Nigris, L. Sommese, and C. Napoli. 2014. "Involvement of Mediator Complex in Malignancy." *Biochimica Et Biophysica Acta* 1845 (1): 66–83. doi:10.1016/j.bbcan.2013.12.001.

Schones, D. E., K. Cui, S. Cuddapah, T.-Y. Roh, A. Barski, Z. Wang, G. Wei, and K. Zhao. 2008. "Dynamic Regulation of Nucleosome Positioning in the Human Genome." *Cell* 132 (5): 887–98. doi:10.1016/j.cell.2008.02.022.

Schultz, P., S. Fribourg, A. Poterszman, V. Mallouh, D. Moras, and J.M. Egly. 2000. "Molecular Structure of Human TFIIH." *Cell* 102 (5): 599–607. doi:10.1016/S0092-8674(00)00082-9.

Schulz, D., N. Pirkl, E. Lehmann, and P. Cramer. 2014. "Rpb4 Subunit Functions Mainly in mRNA Synthesis by RNA Polymerase II." *Journal of Biological Chemistry* 289 (25): 17446–52. doi:10.1074/jbc.M114.568014.

Schulz, D., B. Schwalb, A. Kiesel, C. Baejen, P. Torkler, J. Gagneur, J. Soeding, and P. Cramer. 2013. "Transcriptome Surveillance by Selective Termination of Noncoding RNA Synthesis." *Cell* 155 (5): 1075–87. doi:10.1016/j.cell.2013.10.024.

Schwartz, C. E., P. S. Tarpey, H. A. Lubs, A. Verloes, M. M. May, H. Risheg, M. J. Friez, et al. 2007. "The Original Lujan Syndrome Family Has a Novel Missense Mutation (p.N1007S) in the MED12 Gene." *Journal of Medical Genetics* 44 (7): 472–77. doi:10.1136/jmg.2006.048637.

Seila, A. C., J. M. Calabrese, S. S. Levine, G. W. Yeo, P. B. Rahl, R. A. Flynn, R. A. Young, and P. A. Sharp. 2008. "Divergent Transcription from Active Promoters." *Science (New York, N.Y.)* 322 (5909): 1849–51. doi:10.1126/science.1162253.

Seizl, M., L. Larivière, T. Pfaffeneder, L. Wenzek, and P. Cramer. 2011. "Mediator Head Subcomplex Med11/22 Contains a Common Helix Bundle Building Block with a Specific Function in Transcription Initiation Complex Stabilization." *Nucleic Acids Research* 39 (14): 6291–6304. doi:10.1093/nar/gkr229.

Sekinger, E. A., Z. Moqtaderi, and K. Struhl. 2005. "Intrinsic Histone-DNA Interactions and Low Nucleosome Density Are Important for Preferential Accessibility of Promoter Regions in Yeast." *Molecular Cell* 18 (6): 735–48. doi:10.1016/j.molcel.2005.05.003.

Serizawa, H., J. W. Conaway, and R. C. Conaway. 1994. "An Oligomeric Form of the Large Subunit of Transcription Factor (TF) IIE Activates Phosphorylation of the RNA Polymerase II Carboxyl-Terminal Domain by TFIIH." *Journal of Biological Chemistry* 269 (32): 20750–56.

Sermwittayawong, D., and S. Tan. 2006. "SAGA Binds TBP via Its Spt8 Subunit in Competition with DNA: Implications for TBP Recruitment." *The EMBO Journal* 25 (16): 3791–3800. doi:10.1038/sj.emboj.7601265.

Shen, X., G. Mizuguchi, A. Hamiche, and C. Wu. 2000. "A Chromatin Remodelling Complex Involved in Transcription and DNA Processing." *Nature* 406 (6795): 541–44. doi:10.1038/35020123.

Shin, C. H., W-S. Chung, S-K. Hong, E. A. Ober, H. Verkade, H. A. Field, J. Huisken, and D. Y. R. Stainier. 2008. "Multiple Roles for Med12 in Vertebrate Endoderm Development." *Developmental Biology* 317 (2): 467–79. doi:10.1016/j.ydbio.2008.02.031.

Shpakovski, G. V., J. Acker, M. Wintzerith, J. F. Lacroix, P. Thuriaux, and M. Vigneron. 1995. "Four Subunits That Are Shared by the Three Classes of RNA Polymerase Are Functionally Interchangeable between Homo Sapiens and Saccharomyces Cerevisiae." *Molecular and Cellular Biology* 15 (9): 4702–10. doi:10.1128/MCB.15.9.4702.

Siaut, M., C. Zaros, E. Levivier, M-L. Ferri, M. Court, M. Werner, I. Callebaut, P. Thuriaux, A. Sentenac, and C. Conesa. 2003. "An Rpb4/Rpb7-Like Complex in Yeast RNA Polymerase III Contains the Orthologue of Mammalian CGRP-RCP." *Molecular and Cellular Biology* 23 (1): 195–205. doi:10.1128/MCB.23.1.195-205.2003.

Sil, A. K., S. Alam, P. Xin, L. Ma, M. Morgan, C. M. Lebo, M. P. Woods, and J. E. Hopper. 1999. "The Gal3p-Gal80p-Gal4p Transcription Switch of Yeast: Gal3p Destabilizes the Gal80p-Gal4p Complex in Response to Galactose and ATP." *Molecular and Cellular Biology* 19 (11): 7828–40.

Sims, R. J., L. A. Rojas, D. Beck, R. Bonasio, R. Schüller, W. J. Drury, D. Eick, and D. Reinberg. 2011. "The C-Terminal Domain of RNA Polymerase II Is Modified by Site-Specific Methylation." *Science* 332 (6025): 99–103. doi:10.1126/science.1202663.

Smale, S. T., and D. Baltimore. 1989. "The 'initiator' as a Transcription Control Element." *Cell* 57 (1): 103–13. doi:10.1016/0092-8674(89)90176-1.

Smale, S. T., and J. T. Kadonaga. 2003. "The RNA Polymerase II Core Promoter." *Annual Review of Biochemistry* 72 (1): 449–79. doi:10.1146/annurev.biochem.72.121801.161520.

Smith, R. L., and A. D. Johnson. 2000. "Turning Genes off by Ssn6-Tup1: A Conserved System of Transcriptional Repression in Eukaryotes." *Trends in Biochemical Sciences* 25 (7): 325–30.

Soutourina, J., S. Wydau, Y. Ambroise, C. Boschiero, and M. Werner. 2011. "Direct Interaction of RNA Polymerase II and Mediator Required for Transcription in Vivo." *Science* 331 (6023): 1451–54. doi:10.1126/science.1200188.

Spaeth, J. M., N-H. Kim, and T. G. Boyer. 2011. "Mediator and Human Disease." *Seminars in Cell & Developmental Biology* 22 (7): 776–87. doi:10.1016/j.semcd.2011.07.024.

Spåhr, H., G. Calero, D. A. Bushnell, and R. D. Kornberg. 2009. "Schizosaccharomyces Pombe RNA Polymerase II at 3.6-Å Resolution." *Proceedings of the National Academy of Sciences* 106 (23): 9185–90. doi:10.1073/pnas.0903361106.

Sprague, B. L., R. L. Pego, D. A. Stavreva, and J. G. McNally. 2004. "Analysis of Binding Reactions by Fluorescence Recovery after Photobleaching." *Biophysical Journal* 86 (6): 3473–95. doi:10.1529/biophysj.103.026765.

Sprouse, R. O., T. S. Karpova, F. Mueller, A. Dasgupta, J. G. McNally, and D. T. Auble. 2008. "Regulation of TATA-Binding Protein Dynamics in Living Yeast Cells." *Proceedings of the National Academy of Sciences of the United States of America* 105 (36): 13304–8. doi:10.1073/pnas.0801901105.

Starick, S. R., J. Ibn-Salem, M. Jurk, C. Hernandez, M. I. Love, H-R. Chung, M. Vingron, M. Thomas-Chollier, and S. H. Meijsing. 2015. "ChIP-Exo Signal Associated with DNA-Binding Motifs Provides Insight into the Genomic Binding of the Glucocorticoid Receptor and Cooperating Transcription Factors." *Genome Research* 25 (6): 825–35. doi:10.1101/gr.185157.114.

Steinmetz, E. J., C. L. Warren, J. N. Kuehner, B. Panbehi, A. Z. Ansari, and D. A. Brow. 2006. "Genome-Wide Distribution of Yeast RNA Polymerase II and Its Control by Sen1 Helicase." *Molecular Cell* 24 (5): 735–46. doi:10.1016/j.molcel.2006.10.023.

Stenoien, D. L., K. Patel, M. G. Mancini, M. Dutertre, C. L. Smith, B. W. O'Malley, and M. A. Mancini. 2001. "FRAP Reveals That Mobility of Oestrogen Receptor-Alpha Is Ligand- and Proteasome-Dependent." *Nature Cell Biology* 3 (1): 15–23. doi:10.1038/35050515.

Sterner, D. E., R. Belotserkovskaya, and S. L. Berger. 2002. "SALSA, a Variant of Yeast SAGA, Contains Truncated Spt7, Which Correlates with Activated Transcription." *Proceedings of the National Academy of Sciences of the United States of America* 99 (18): 11622–27. doi:10.1073/pnas.182021199.

Stiller, J. W., and B. D. Hall. 2002. "Evolution of the RNA Polymerase II C-Terminal Domain." *Proceedings of the National Academy of Sciences* 99 (9): 6091–96. doi:10.1073/pnas.082646199.

Stolc, V., Z. Gauhar, C. Mason, G. Halasz, M. F. van Batenburg, S. A. Rifkin, S. Hua, et al. 2004. "A Gene Expression Map for the Euchromatic Genome of *Drosophila Melanogaster*." *Science (New York, N.Y.)* 306 (5696): 655–60. doi:10.1126/science.1101312.

Stumpf, M., C. Waskow, M. Krötschel, D. van Essen, P. Rodriguez, X. Zhang, B. Guyot, R. G. Roeder, and T. Borggrefe. 2006. "The Mediator Complex Functions as a Coactivator for GATA-1 in Erythropoiesis via Subunit Med1/TRAP220." *Proceedings of the National Academy of Sciences of the United States of America* 103 (49): 18504–9. doi:10.1073/pnas.0604494103.

Sun, X., Y. Zhang, H. Cho, P. Rickert, E. Lees, W. Lane, and D. Reinberg. 1998. "NAT, a Human Complex Containing Srb Polypeptides That Functions as a Negative Regulator of Activated Transcription." *Molecular Cell* 2 (2): 213–22.

Sun, Z. W., and M. Hampsey. 1995. "Identification of the Gene (SSU71/TFG1) Encoding the Largest Subunit of Transcription Factor TFIIF as a Suppressor of a TFIIB Mutation in *Saccharomyces Cerevisiae*." *Proceedings of the National Academy of Sciences of the United States of America* 92 (8): 3127–31.

Syring, I., N. Klümper, A. Offermann, M. Braun, M. Deng, D. Boehm, A. Queisser, et al. 2016. "Comprehensive Analysis of the Transcriptional Profile of the Mediator Complex across Human Cancer Types." *Oncotarget* 7 (17): 23043–55.

Taatjes, D. J., A. M. Näär, F. Andel, E. Nogales, and R. Tjian. 2002. "Structure, Function, and Activator-Induced Conformations of the CRSP Coactivator." *Science (New York, N.Y.)* 295 (5557): 1058–62. doi:10.1126/science.1065249.

Takagi, Y., G. Calero, H. Komori, J. A. Brown, A. H. Ehrensberger, A. Hudmon, F. Asturias, and R. D. Kornberg. 2006. "Head Module Control of Mediator Interactions." *Molecular Cell* 23 (3): 355–64. doi:10.1016/j.molcel.2006.06.007.

Takagi, Y., and R. D. Kornberg. 2006. "Mediator as a General Transcription Factor." *The Journal of Biological Chemistry* 281 (1): 80–89. doi:10.1074/jbc.M508253200.

Takahashi, H., T. J. Parmely, S. Sato, C. Tomomori-Sato, C. A. S. Banks, S. E. Kong, H. Szutorisz, et al. 2011. "Human Mediator Subunit MED26 Functions as a Docking Site for Transcription Elongation Factors." *Cell* 146 (1): 92–104. doi:10.1016/j.cell.2011.06.005.

Takahashi, H., K. Kasahara, and T. Kokubo. 2009. "*Saccharomyces Cerevisiae* Med9 Comprises Two Functionally Distinct Domains That Play Different Roles in Transcriptional Regulation." *Genes to Cells: Devoted to Molecular & Cellular Mechanisms* 14 (1): 53–67. doi:10.1111/j.1365-2443.2008.01250.x.

Tan, Q., K. L. Linask, R. H. Ebright, and N. A. Woychik. 2000. "Activation Mutants in Yeast RNA Polymerase II Subunit RPB3 Provide Evidence for a Structurally Conserved Surface Required for Activation in Eukaryotes and Bacteria." *Genes & Development* 14 (3): 339–48. doi:10.1101/gad.14.3.339.

Tan, S., Y. Hunziker, D. F. Sargent, and T. J. Richmond. 1996. "Crystal Structure of a Yeast TFIIA/TBP/DNA Complex." *Nature* 381 (6578): 127–34. doi:10.1038/381127a0.

Tardiff, D. F., K. C. Abruzzi, and M. Rosbash. 2007. "Protein Characterization of *Saccharomyces Cerevisiae* RNA Polymerase II after in Vivo Cross-Linking." *Proceedings of the National Academy of Sciences of the United States of America* 104 (50): 19948–53. doi:10.1073/pnas.0710179104.

Teixeira, M. C., P. T. Monteiro, J. Fernandes Guerreiro, J. Pinho Gonçalves, N. Pereira Mira, S. Costa dos Santos, T. Rodrigues Cabrito, et al. 2014. "The YEASTRACT Database: An Upgraded Information System for the Analysis of Gene and Genomic Transcription Regulation in *Saccharomyces Cerevisiae*." *Nucleic Acids Research* 42 (D1): D161–D166. doi:10.1093/nar/gkt1015.

Terriente-Félix, A., A. López-Varea, and J. F. de Celis. 2010. "Identification of Genes Affecting Wing Patterning through a Loss-of-Function Mutagenesis Screen and Characterization of med15 Function during Wing Development." *Genetics* 185 (2): 671–84. doi:10.1534/genetics.109.113670.

Teytelman, L., D. M. Thurtle, J. Rine, and A. van Oudenaarden. 2013. "Highly Expressed Loci Are Vulnerable to Misleading ChIP Localization of Multiple Unrelated Proteins." *Proceedings of the National Academy of Sciences of the United States of America* 110 (46): 18602–7. doi:10.1073/pnas.1316064110.

Thakur, J. K., H. Arthanari, F. Yang, K. H. Chau, G. Wagner, and A. M. Näär. 2009. "Mediator Subunit Gal11p/MED15 Is Required for Fatty Acid-Dependent Gene Activation by Yeast Transcription Factor Oaf1p." *The Journal of Biological Chemistry* 284 (7): 4422–28. doi:10.1074/jbc.M808263200.

Theisen, J. W. M., C. Y. Lim, and J. T. Kadonaga. 2010. "Three Key Subregions Contribute to the Function of the Downstream RNA Polymerase II Core Promoter." *Molecular and Cellular Biology* 30 (14): 3471–79. doi:10.1128/MCB.00053-10.

Thomas, M. C., and C-M. Chiang. 2006. "The General Transcription Machinery and General Cofactors." *Critical Reviews in Biochemistry and Molecular Biology* 41 (3): 105–78. doi:10.1080/10409230600648736.

Thompson, C. M., and R. A. Young. 1995. "General Requirement for RNA Polymerase II Holoenzymes in Vivo." *Proceedings of the National Academy of Sciences of the United States of America* 92 (10): 4587–90.

Thompson, C. M., A. J. Koleske, D. M. Chao, and R. A. Young. 1993. "A Multisubunit Complex Associated with the RNA Polymerase II CTD and TATA-Binding Protein in Yeast." *Cell* 73 (7): 1361–75. doi:10.1016/0092-8674(93)90362-T.

Thuault, S., Y.-Gangloff, J. Kirchner, S. Sanders, S. Werten, C. Romier, P. A. Weil, and I. Davidson. 2002. "Functional Analysis of the TFIID-Specific Yeast TAF4 (yTAF(II)48) Reveals an Unexpected Organization of Its Histone-Fold Domain." *The Journal of Biological Chemistry* 277 (47): 45510–17. doi:10.1074/jbc.M206556200.

Tietjen, J. R., D. W. Zhang, J. B. Rodríguez-Molina, B. E. White, M. S. Akhtar, M. Heidemann, X. Li, et al. 2010. "Chemical-Genomic Dissection of the CTD Code." *Nature Structural & Molecular Biology* 17 (99): 1154–61. doi:10.1038/nsmb.1900.

Tora, L. 2002. "A Unified Nomenclature for TATA Box Binding Protein (TBP)-Associated Factors (TAFs) Involved in RNA Polymerase II Transcription." *Genes & Development* 16 (6): 673–75.

Triezenberg, S. J. 1995. "Structure and Function of Transcriptional Activation Domains." *Current Opinion in Genetics & Development* 5 (2): 190–96.

Tsai, K-L., C. Tomomori-Sato, S. Sato, R. C. Conaway, J. W. Conaway, and F. J. Asturias. 2014. "Subunit Architecture and Functional Modular Rearrangements of the Transcriptional Mediator Complex." *Cell* 157 (6): 1430–44. doi:10.1016/j.cell.2014.05.015.

Tsukiyama, T., C. Daniel, J. Tamkun, and C. Wu. 1995. "ISWI, a Member of the SWI2/SNF2 ATPase Family, Encodes the 140 kDa Subunit of the Nucleosome Remodeling Factor." *Cell* 83 (6): 1021–26.

Tudek, A., T. Candelli, and D. Libri. 2015. "Non-Coding Transcription by RNA Polymerase II in Yeast: Hasard or Nécessité?" *Biochimie* 117 (October): 28–36. doi:10.1016/j.biochi.2015.04.020.

Tudor, M., P. J. Murray, C. Onufryk, R. Jaenisch, and R. A. Young. 1999. "Ubiquitous Expression and Embryonic Requirement for RNA Polymerase II Coactivator Subunit Srb7 in Mice." *Genes & Development* 13 (18): 2365–68.

Uetz, P., L. Giot, G. Cagney, T. A. Mansfield, R. S. Judson, J. R. Knight, D. Lockshon, et al. 2000. "A Comprehensive Analysis of Protein–protein Interactions in *Saccharomyces Cerevisiae*." *Nature* 403 (6770): 623–27. doi:10.1038/35001009.

Ulmasov, T., R. M. Larkin, and T. J. Guilfoyle. 1996. "Association between 36- and 13.6-kDa -Like Subunits of *Arabidopsis Thaliana* RNA Polymerase II." *Journal of Biological Chemistry* 271 (9): 5085–94. doi:10.1074/jbc.271.9.5085.

Urahama, N., M. Ito, A. Sada, K. Yakushijin, K. Yamamoto, A. Okamura, K. Minagawa, et al. 2005. "The Role of Transcriptional Coactivator TRAP220 in Myelomonocytic Differentiation." *Genes to Cells: Devoted to Molecular & Cellular Mechanisms* 10 (12): 1127–37. doi:10.1111/j.1365-2443.2005.00906.x.

Usheva, A., E. Maldonado, A. Goldring, H. Lu, C. Houbavi, D. Reinberg, and Y. Aloni. 1992. "Specific Interaction between the Nonphosphorylated Form of RNA Polymerase II and the TATA-Binding Protein." *Cell* 69 (5): 871–81. doi:10.1016/0092-8674(92)90297-P.

Vakoc, C. R., D. L. Letting, N. Gheldof, T. Sawado, M. A. Bender, M. Groudine, M. J. Weiss, J. Dekker, and G. A. Blobel. 2005. "Proximity among Distant Regulatory Elements at the Beta-Globin Locus Requires GATA-1 and FOG-1." *Molecular Cell* 17 (3): 453–62. doi:10.1016/j.molcel.2004.12.028.

Valay, J. G., M. F. Dubois, O. Bensaude, and G. Faye. 1996. "Ccl1, a Cyclin Associated with Protein Kinase Kin28, Controls the Phosphorylation of RNA Polymerase II Largest Subunit and mRNA Transcription." *Comptes Rendus de l'Académie Des Sciences. Série III, Sciences de La Vie* 319 (3): 183–89.

Valouev, A, S. M. Johnson, S. D. Boyd, C. L. Smith, A. Z. Fire, and A. Sidow. 2011. "Determinants of Nucleosome Organization in Primary Human Cells." *Nature* 474 (7352): 516–20. doi:10.1038/nature10002.

Van de Peppel, J., N. Kettelarij, H. van Bakel, T. T. J. P. Kockelkorn, D. van Leenen, and F. C. P. Holstege. 2005. "Mediator Expression Profiling Epistasis Reveals a Signal Transduction Pathway with Antagonistic Submodules and Highly Specific Downstream Targets." *Molecular Cell* 19 (4): 511–22. doi:10.1016/j.molcel.2005.06.033.

Van Dijk, E. L., C. L. Chen, Y. d'Aubenton-Carafa, S. Gourvennec, M. Kwapisz, V. Roche, C. Bertrand, et al. 2011. "XUTs Are a Class of Xrn1-Sensitive Antisense Regulatory Non-Coding RNA in Yeast." *Nature* 475 (7354): 114–17. doi:10.1038/nature10118.

Verma-Gaur, J., S. Narayana Rao, T. Taya, and P. Sadhale. 2008. "Genomewide Recruitment Analysis of Rpb4, a Subunit of Polymerase II in *Saccharomyces Cerevisiae*, Reveals Its Involvement in Transcription Elongation." *Eukaryotic Cell* 7 (6): 1009–18. doi:10.1128/EC.00057-08.

Vignali, M., A. H. Hassan, K. E. Neely, and J. L. Workman. 2000. "ATP-Dependent Chromatin-Remodeling Complexes." *Molecular and Cellular Biology* 20 (6): 1899–1910.

Vijayvargia, R., M. S. May, and J. D. Fondell. 2007. "A Coregulatory Role for the Mediator Complex in Prostate Cancer Cell Proliferation and Gene Expression." *Cancer Research* 67 (9): 4034–41. doi:10.1158/0008-5472.CAN-06-3039.

Wada, T., T. Takagi, Y. Yamaguchi, D. Watanabe, and H. Handa. 1998. "Evidence That P-TEFb Alleviates the Negative Effect of DSIF on RNA Polymerase II-Dependent Transcription in Vitro." *The EMBO Journal* 17 (24): 7395–7403. doi:10.1093/emboj/17.24.7395.

Wada, T., T. Takagi, Y. Yamaguchi, A. Ferdous, T. Imai, S. Hirose, S. Sugimoto, et al. 1998. "DSIF, a Novel Transcription Elongation Factor That Regulates RNA Polymerase II Processivity, Is Composed of Human Spt4 and Spt5 Homologs." *Genes & Development* 12 (3): 343–56.

Wallberg, A. E., K. E. Neely, A. H. Hassan, J. A. Gustafsson, J. L. Workman, and A. P. Wright. 2000. "Recruitment of the SWI-SNF Chromatin Remodeling Complex as a Mechanism of Gene Activation by the Glucocorticoid Receptor tau1 Activation Domain." *Molecular and Cellular Biology* 20 (6): 2004–13.

Walmacq, C., M.L. Kireeva, J. Irvin, Y. Nedialkov, L. Lubkowska, F. Malagon, J. N. Strathern, and M. Kashlev. 2009. "Rpb9 Subunit Controls Transcription Fidelity by Delaying NTP Sequestration in RNA Polymerase II." *Journal of Biological Chemistry* 284 (29): 19601–12. doi:10.1074/jbc.M109.006908.

Walmsley, R. W., C. S. Chan, B. K. Tye, and T. D. Petes. 1984. "Unusual DNA Sequences Associated with the Ends of Yeast Chromosomes." *Nature* 310 (5973): 157–60.

Wang, Q., D. Sharma, Y. Ren, and J. D. Fondell. 2002. "A Coregulatory Role for the TRAP-Mediator Complex in Androgen Receptor-Mediated Gene Expression." *The Journal of Biological Chemistry* 277 (45): 42852–58. doi:10.1074/jbc.M206061200.

Wang, W., J. D. Gralla, and M. Carey. 1992. "The Acidic Activator GAL4-AH Can Stimulate Polymerase II Transcription by Promoting Assembly of a Closed Complex Requiring TFIID and TFIIA." *Genes & Development* 6 (9): 1716–27.

Wang, W., X. Yao, Y. Huang, X. Hu, R. Liu, D. Hou, R. Chen, and G. Wang. 2013. "Mediator MED23 Regulates Basal Transcription in Vivo via an Interaction with P-TEFb." *Transcription* 4 (1): 39–51. doi:10.4161/trns.22874.

Wang, X., Q. Sun, Z. Ding, J. Ji, J. Wang, X. Kong, J. Yang, and G. Cai. 2014. "Redefining the Modular Organization of the Core Mediator Complex." *Cell Research* 24 (7): 796–808. doi:10.1038/cr.2014.64.

Wang, Z., and R. G. Roeder. 1997. "Three Human RNA Polymerase III-Specific Subunits Form a Subcomplex with a Selective Function in Specific Transcription Initiation." *Genes & Development* 11 (10): 1315–26. doi:10.1101/gad.11.10.1315.

Washburn, M. P., D. Wolters, and J. R. Yates. 2001. "Large-Scale Analysis of the Yeast Proteome by Multidimensional Protein Identification Technology." *Nature Biotechnology* 19 (3): 242–47. doi:10.1038/85686.

Wasylyk, B., R. Derbyshire, A. Guy, D. Molko, A. Roget, R. Téoule, and P. Chambon. 1980. "Specific in Vitro Transcription of Conalbumin Gene Is Drastically Decreased by Single-Point Mutation in T-A-T-A Box Homology Sequence." *Proceedings of the National Academy of Sciences* 77 (12): 7024–28.

Weil, P. A., D. S. Luse, J. Segall, and R. G. Roeder. 1979. "Selective and Accurate Initiation of Transcription at the ad2 Major Late Promotor in a Soluble System Dependent on Purified Rna Polymerase Ii and Dna." *Cell* 18 (2): 469–84. doi:10.1016/0092-8674(79)90065-5.

Weinmann, R., and R. G. Roeder. 1974. "Role of DNA-Dependent RNA Polymerase III in the Transcription of the tRNA and 5S RNA Genes." *Proceedings of the National Academy of Sciences of the United States of America* 71 (5): 1790–94.

Weis, L., and D. Reinberg. 1997. "Accurate Positioning of RNA Polymerase II on a Natural TATA-Less Promoter Is Independent of TATA-Binding-Protein-Associated Factors and Initiator-Binding Proteins." *Molecular and Cellular Biology* 17 (6): 2973–84. doi:10.1128/MCB.17.6.2973.

Weiss, S. B., and L. Gladstone. 1959. "A MAMMALIAN SYSTEM FOR THE INCORPORATION OF CYTIDINE TRIPHOSPHATE INTO RIBONUCLEIC ACID1." *Journal of the American Chemical Society* 81 (15): 4118–19. doi:10.1021/ja01524a087.

Wellinger, R. J., A. J. Wolf, and V. A. Zakian. 1993. "Saccharomyces Telomeres Acquire Single-Strand TG1-3 Tails Late in S Phase." *Cell* 72 (1): 51–60.

Werner, F., and D. Grohmann. 2011. "Evolution of Multisubunit RNA Polymerases in the Three Domains of Life." *Nature Reviews Microbiology* 9 (2): 85–98. doi:10.1038/nrmicro2507.

Werner, F., and R. O. J. Weinzierl. 2002. "A Recombinant RNA Polymerase II-like Enzyme Capable of Promoter-Specific Transcription." *Molecular Cell* 10 (3): 635–46. doi:10.1016/S1097-2765(02)00629-9.

Werner, M., P. Thuriaux, and J. Soutourina. 2009. "Structure–function Analysis of RNA Polymerases I and III." *Current Opinion in Structural Biology, Catalysis and regulation / Proteins*, 19 (6): 740–45. doi:10.1016/j.sbi.2009.10.005.

Wery, M., E. Shematorova, B. Van Driessche, J. Vandenhautte, P. Thuriaux, and V. Van Mullem. 2004. "Members of the SAGA and Mediator Complexes Are Partners of the Transcription Elongation Factor TFIIS." *The EMBO Journal* 23 (21): 4232–42. doi:10.1038/sj.emboj.7600326.

West, M. L., and J. L. Corden. 1995. "Construction and Analysis of Yeast RNA Polymerase II CTD Deletion and Substitution Mutations." *Genetics* 140 (4): 1223–33.

Westerling, T., E. Kuuluvainen, and T. P. Mäkelä. 2007. "Cdk8 Is Essential for Preimplantation Mouse Development." *Molecular and Cellular Biology* 27 (17): 6177–82. doi:10.1128/MCB.01302-06.

Whyte, W. A., D. A. Orlando, D. Hnisz, B. J. Abraham, C. Y. Lin, M. H. Kagey, P. B. Rahl, T. I. Lee, and R. A. Young. 2013. "Master Transcription Factors and Mediator Establish Super-Enhancers at Key Cell Identity Genes." *Cell* 153 (2): 307–19. doi:10.1016/j.cell.2013.03.035.

Wong, K. H., Y. Jin, and K. Struhl. 2014. "TFIIH Phosphorylation of the Pol II CTD Stimulates Mediator Dissociation from the Preinitiation Complex and Promoter Escape." *Molecular Cell* 54 (4): 601–12. doi:10.1016/j.molcel.2014.03.024.

Wu, C-C., F. Herzog, S. Jennebach, Y-C. Lin, C-Y. Pai, R. Aebersold, P. Cramer, and H-T Chen. 2012. "RNA Polymerase III Subunit Architecture and Implications for Open Promoter Complex Formation." *Proceedings of the National Academy of Sciences* 109 (47): 19232–37. doi:10.1073/pnas.1211665109.

Wu, S. Y. 1998. "TAFII-Independent Activation Mediated by Human TBP in the Presence of the Positive Cofactor PC4." *The EMBO Journal* 17 (15): 4478–90. doi:10.1093/emboj/17.15.4478.

Wu, W. H., I. Pinto, B. S. Chen, and M. Hampsey. 1999. "Mutational Analysis of Yeast TFIIB. A Functional Relationship between Ssu72 and Sub1/Tsp1 Defined by Allele-Specific Interactions with TFIIB." *Genetics* 153 (2): 643–52.

Wyers, F., M. Rougemaille, G. Badis, J-C. Rousselle, M-E. Dufour, J. Boulay, B. Régnault, et al. 2005. "Cryptic Pol II Transcripts Are Degraded by a Nuclear Quality Control Pathway Involving a New poly(A) Polymerase." *Cell* 121 (5): 725–37. doi:10.1016/j.cell.2005.04.030.

Xu, Z., W. Wei, J. Gagneur, F. Perocchi, S. Clauder-Münster, J. Camblong, E. Guffanti, F. Stutz, W. Huber, and L. M. Steinmetz. 2009. "Bidirectional Promoters Generate Pervasive Transcription in Yeast." *Nature* 457 (7232): 1033–37. doi:10.1038/nature07728.

- Xue, W., J. Wang, Z. Shen, and H. Zhu. 2004. "Enrichment of Transcriptional Regulatory Sites in Non-Coding Genomic Region." *Bioinformatics (Oxford, England)* 20 (4): 569–75. doi:10.1093/bioinformatics/btg450.
- Yamaguchi, Y., T. Takagi, T. Wada, K. Yano, A. Furuya, S. Sugimoto, J. Hasegawa, and H. Handa. 1999. "NELF, a Multisubunit Complex Containing RD, Cooperates with DSIF to Repress RNA Polymerase II Elongation." *Cell* 97 (1): 41–51. doi:10.1016/S0092-8674(00)80713-8.
- Yan, Q., R. J. Moreland, J. W. Conaway, and R. C. Conaway. 1999. "Dual Roles for Transcription Factor IIF in Promoter Escape by RNA Polymerase II." *Journal of Biological Chemistry* 274 (50): 35668–75. doi:10.1074/jbc.274.50.35668.
- Yang, F. B. W. Vought, J. S. Satterlee, A. K. Walker, Z.-Y. Jim Sun, J. L. Watts, R. DeBeaumont, et al. 2006. "An ARC/Mediator Subunit Required for SREBP Control of Cholesterol and Lipid Homeostasis." *Nature* 442 (7103): 700–704. doi:10.1038/nature04942.
- Yeo, M., P. S. Lin, M. E. Dahmus, and G. N. Gill. 2003. "A Novel RNA Polymerase II C-Terminal Domain Phosphatase That Preferentially Dephosphorylates Serine 5." *The Journal of Biological Chemistry* 278 (28): 26078–85. doi:10.1074/jbc.M301791200.
- Yin, J.-W., and G. Wang. 2014. "The Mediator Complex: A Master Coordinator of Transcription and Cell Lineage Development." *Development (Cambridge, England)* 141 (5): 977–87. doi:10.1242/dev.098392.
- Yoon, S., H. Qiu, M. J. Swanson, and A. G. Hinnebusch. 2003. "Recruitment of SWI/SNF by Gcn4p Does Not Require Snf2p or Gcn5p but Depends Strongly on SWI/SNF Integrity, SRB Mediator, and SAGA." *Molecular and Cellular Biology* 23 (23): 8829–45.
- Young, R. A. 1991. "RNA Polymerase II." *Annual Review of Biochemistry* 60 (1): 689–715. doi:10.1146/annurev.bi.60.070191.003353.
- Yuan, G.-C., Y.-J. Liu, M. F. Dion, M. D. Slack, L. F. Wu, S. J. Altschuler, and O. J. Rando. 2005. "Genome-Scale Identification of Nucleosome Positions in *S. Cerevisiae*." *Science (New York, N.Y.)* 309 (5734): 626–30. doi:10.1126/science.1112178.
- Yudkovsky, N., C. Logie, S. Hahn, and C. L. Peterson. 1999. "Recruitment of the SWI/SNF Chromatin Remodeling Complex by Transcriptional Activators." *Genes & Development* 13 (18): 2369–74.
- Yudkovsky, N., J. A. Ranish, and S. Hahn. 2000. "A Transcription Reinitiation Intermediate That Is Stabilized by Activator." *Nature* 408 (6809): 225–29. doi:10.1038/35041603.
- Yuryev, A., M. Patturajan, Y. Litingtung, R. V. Joshi, C. Gentile, M. Gebara, and J. L. Corden. 1996. "The C-Terminal Domain of the Largest Subunit of RNA Polymerase II Interacts with a Novel Set of Serine/arginine-Rich Proteins." *Proceedings of the National Academy of Sciences of the United States of America* 93 (14): 6975–80.
- Zeitlinger, J., A. Stark, M. Kellis, J.-W. Hong, S. Nechaev, K. Adelman, M. Levine, and R. A. Young. 2007. "RNA Polymerase Stalling at Developmental Control Genes in the *Drosophila Melanogaster* Embryo." *Nature Genetics* 39 (12): 1512–16. doi:10.1038/ng.2007.26.
- Zhang, C., and Z. F. Burton. 2004. "Transcription Factors IIF and IIS and Nucleoside Triphosphate Substrates as Dynamic Probes of the Human RNA Polymerase II Mechanism." *Journal of Molecular Biology* 342 (4): 1085–99. doi:10.1016/j.jmb.2004.07.070.
- Zhang, C., K. L. Zobeck, and Z. F. Burton. 2005. "Human RNA Polymerase II Elongation in Slow Motion: Role of the TFIIF RAP74 α 1 Helix in Nucleoside Triphosphate-Driven Translocation." *Molecular and Cellular Biology* 25 (9): 3583–95. doi:10.1128/MCB.25.9.3583-3595.2005.
- Zhu, C., K. J. R. P. Byers, R. Patton McCord, Z. Shi, M. F. Berger, D. E. Newburger, K. Saulrieta, et al. 2009. "High-Resolution DNA-Binding Specificity Analysis of Yeast Transcription Factors." *Genome Research* 19 (4): 556–66. doi:10.1101/gr.090233.108.
- Zhu, X., B. Liu, J. O. P. Carlsten, J. Beve, T. Nyström, L. C. Myers, and C. M. Gustafsson. 2011. "Mediator Influences Telomeric Silencing and Cellular Life Span." *Molecular and Cellular Biology* 31 (12): 2413–21. doi:10.1128/MCB.05242-11.
- Zhu, Y., C. Qi, S. Jain, M. M. Le Beau, R. Espinosa, G. B. Atkins, M. A. Lazar, A. V. Yeldandi, M. S. Rao, and J. K. Reddy. 1999. "Amplification and Overexpression of Peroxisome Proliferator-Activated Receptor Binding Protein (PBP/PPARBP) Gene in Breast Cancer." *Proceedings of the National Academy of Sciences of the United States of America* 96 (19): 10848–53.

- Ziegler, Lynn M., Denys A. Khapersky, Michelle L. Ammerman, and Alfred S. Ponticelli. 2003a. "Yeast RNA Polymerase II Lacking the Rpb9 Subunit Is Impaired for Interaction with Transcription Factor IIF." *Journal of Biological Chemistry* 278 (49): 48950–56. doi:10.1074/jbc.M309656200.
- Zou, J., Y. Guo, T. Guettouche, D. F. Smith, and R. Voellmy. 1998. "Repression of Heat Shock Transcription Factor HSF1 Activation by HSP90 (HSP90 Complex) That Forms a Stress-Sensitive Complex with HSF1." *Cell* 94 (4): 471–80.
- Zurita, M., and C. Merino. 2003. "The Transcriptional Complexity of the TFIID Complex." *Trends in Genetics* 19 (10): 578–84. doi:10.1016/j.tig.2003.08.005.

Annexes

Publications

Durant ma thèse, j'ai été impliqué dans d'autres projets que ceux faisant l'objet de ce manuscrit. J'ai ainsi participé au projet dans lequel nous avons montré le lien entre le Médiateur et la réparation de l'ADN (Eyboulet et al., 2013). J'ai également participé à l'étude des mutants thermosensibles de la sous-unité Med17, nous ayant permis de mettre en évidence le rôle du Médiateur et du module de tête dans la formation des PICs selon différentes voies d'assemblage *in vivo* (Eyboulet et al., 2015). Ces deux projets ont fait l'objet de publications dont je suis auteur. Dans l'article concernant le lien entre le Médiateur et la réparation de l'ADN, j'ai apporté ma contribution en réalisant les expériences de CHIP-seq pour analyser la distribution de Med17 sur le génome. J'ai également démontré la conservation de l'interaction entre Med17 et Rad2/XPG chez l'Homme en utilisant la méthode de co-IP dans des extraits de protéines de cellules HeLa. Concernant l'étude des mutants de Med17, j'ai contribué à la publication des résultats en réalisant les expériences de purification et d'analyse du Médiateur par spectrométrie de masse pour vérifier l'intégrité des complexes issus des souches mutantes de Med17.

I) Article 1: Mediator links transcription and DNA repair by facilitating Rad2/XPG recruitment

Mediator links transcription and DNA repair by facilitating Rad2/XPG recruitment

Fanny Eyboulet,^{1,2,3} Camille Cibot,^{1,2,3,5} Thomas Eychenne,^{1,2,3} Helen Neil,^{1,2,3,6} Olivier Alibert,⁴ Michel Werner,^{1,2,3} and Julie Soutourina^{1,2,3,7}

¹FRE3377, Institut de Biologie et de Technologies de Saclay (iBiTec-S), Commissariat à l'Énergie Atomique et aux Énergies Alternatives (CEA), F-91191 Gif-sur-Yvette cedex, France; ²FRE3377, Centre National de la Recherche Scientifique (CNRS), F-91191 Gif-sur-Yvette Cedex, France; ³FRE3377, Université Paris-Sud, F-91191 Gif-sur-Yvette Cedex, France; ⁴Institut de Radiobiologie Cellulaire et Moléculaire (iRCM), CEA, F-91057 Evry Cedex, France

Mediator is a large multiprotein complex conserved in all eukaryotes. The crucial function of Mediator in transcription is now largely established. However, we found that this complex also plays an important role by connecting transcription with DNA repair. We identified a functional contact between the Med17 Mediator subunit and Rad2/XPG, the 3' endonuclease involved in nucleotide excision DNA repair. Genome-wide location analyses revealed that Rad2 is associated with RNA polymerase II (Pol II)- and Pol III-transcribed genes and telomeric regions in the absence of exogenous genotoxic stress. Rad2 occupancy of Pol II-transcribed genes is transcription-dependent. Genome-wide Rad2 occupancy of class II gene promoters is well correlated with that of Mediator. Furthermore, UV sensitivity of *med17* mutants is correlated with reduced Rad2 occupancy of class II genes and concomitant decrease of Mediator interaction with Rad2 protein. Our results suggest that Mediator is involved in DNA repair by facilitating Rad2 recruitment to transcribed genes.

[*Keywords:* Mediator; transcription; DNA repair; Rad2/XPG; RNA polymerase II; *Saccharomyces cerevisiae*]

Supplemental material is available for this article.

Received July 24, 2013; revised version accepted October 23, 2013.

Mediator of transcription regulation is a large multi-protein complex conserved in all eukaryotes. It interacts with RNA polymerase II (Pol II) to transmit signals from specific regulators to the Pol II transcription machinery. Together with the general transcription factors (GTFs; TFIIA, TFIIB, TFIID, TFII E, TFII F, and TFII H), this coactivator stimulates the assembly of the preinitiation complex (PIC) and, ultimately, transcription initiation (Kornberg 2005; Ries and Meisterernst 2011). Mediator is generally required at Pol II-transcribed gene promoters for regulated transcription (Holstege et al. 1998; Andrau et al. 2006; Zhu et al. 2006). As expected given its central role in transcription activation, it has been implicated in numerous developmental processes. Mutations that affect Mediator subunits lead to a number of human pathologies (Risheg et al. 2007; Schwartz et al. 2007). Recently, mutations in the Med17 Mediator subunit have been

associated with infantile cerebral atrophy (Kaufmann et al. 2010), and a mutation in the Med23 subunit cosegregated with intellectual disability (Hashimoto et al. 2011). Since oncogenesis results from gene dysregulation, it is not unexpected that Mediator is involved in several cancers (Zhang et al. 2005; Vijayvargia et al. 2007; Firestein et al. 2008; Gade et al. 2009; Li et al. 2010; Kuuselo et al. 2011). Mediator has at least 25 subunits organized into four modules: head, middle, tail, and Cdk8 kinase modules. The complete crystallographic structure of the Mediator complex has not been determined. However, recent structural data have been obtained for the Mediator head module (Lariviere et al. 2012; Robinson et al. 2012). Mediator subunits are engaged in numerous contacts within the complex (Guglielmi et al. 2004) and probably with partners belonging to other components of the transcription machinery. For example, a cooperation of Mediator with TFIID has been reported in

Present addresses: ⁵Centre de Génétique Moléculaire du CNRS, F-91198 Gif-sur-Yvette, France; ⁶Institut National de la Santé et de la Recherche Médicale (Inserm) U1052, Centre de Recherche en Cancérologie de Lyon, F-69008 Lyon, France.

⁷Corresponding author

E-mail julie.soutourina@cea.fr

Article is online at <http://www.genesdev.org/cgi/doi/10.1101/gad.225813.113>.

© 2013 Eyboulet et al. This article is distributed exclusively by Cold Spring Harbor Laboratory Press for the first six months after the full-issue publication date (see <http://genesdev.cshlp.org/site/misc/terms.xhtml>). After six months, it is available under a Creative Commons License (Attribution-NonCommercial 3.0 Unported), as described at <http://creativecommons.org/licenses/by-nc/3.0/>.

yeast and humans (Johnson et al. 2002; Lariviere et al. 2006; Cai et al. 2010; Takahashi et al. 2011). Recently, we identified a direct interaction between the Med17 Mediator subunit and the Rpb3 Pol II subunit required for global Pol II transcription in vivo (Soutourina et al. 2011). Previously, we showed that a direct contact between the Med11 subunit of Mediator and the Rad3 subunit of TFIIH is essential for the recruitment of the GTF to the PIC independently of Pol II (Esnault et al. 2008). In addition to the crucial role of Mediator as a link between activators and the basal transcription machinery, recent studies suggest broader Mediator transcriptional functions, notably in the post-recruitment steps of transcription (Wang et al. 2005; Takahashi et al. 2011) and in DNA looping between the enhancers and promoters through a functional connection with cohesin (Kagey et al. 2010).

Transcription is coupled with DNA repair, ensuring the continuity of Pol II progression. DNA lesions impede the elongating transcription machinery and can potentially affect gene expression, leading to problems in development, cell growth, and survival. Proteins first identified as components of transcription or DNA repair machineries may be involved in both processes, operating beyond their initially described pathway. The TFIIH complex is the best-characterized dual factor, playing an essential role in Pol II transcription as a GTF and in nucleotide excision DNA repair (NER) (Compe and Egly 2012). NER is a major and evolutionarily conserved DNA repair pathway that removes DNA lesions such as cyclobutane pyrimidine dimers (CPD), the most abundant photoproduct arising upon UV irradiation. During NER, >30 proteins cooperate to recognize, incise, and excise a damaged oligonucleotide from the genomic DNA to protect the cell from the deleterious effects of DNA damages (Lagerwerf et al. 2011). There are two distinct NER pathways: global genome repair (GGR), which removes DNA lesions in the genome overall, and transcription-coupled repair (TCR), which removes DNA lesions that interfere with the progression of RNA polymerase through actively transcribed genes (Svejstrup 2002, 2007; Hanawalt and Spivak 2008). A single irreversibly stalled Pol II in an essential gene could be enough to lead to cell death. This situation explains why a large number of mechanisms and factors have evolved to protect the cell from stalled Pol II. While the NER reaction and the components required for efficient DNA lesion recognition and repair are fairly well understood, many questions on the active coupling of DNA repair with transcription remain unanswered.

Recent reports suggest that an interplay between transcription and DNA repair is likely to be more complex than previously thought. PAF and Ccr4–Not complexes, known for their role in transcription, have been proposed to be required for efficient TCR in yeast (Gaillard et al. 2009). In addition to TFIIH, other NER factors have been suggested to play a dual role in DNA repair and transcription. Rad2/XPG, the 3' endonuclease implicated in NER, is one of such factor. Mutations in the human XPG gene give rise to a xeroderma pigmentosum (XP) sometimes associated with Cockayne syndrome (CS)

(Clarkson 2003; Scharer 2008). Rad2/XPG, together with the Rad1–Rad10/XPF–ERCC1 5' endonuclease, cleaves the damaged DNA strand in both NER pathways, GGR and TCR, resulting in the release of a DNA fragment containing the DNA lesion. Moreover, a specific non-enzymatic function of XPG in recognition of stalled transcription in TCR initiation has been also proposed (Sarker et al. 2005). In addition, a transcriptional role for the NER factor Rad2/XPG has been suggested in the yeast *Saccharomyces cerevisiae* (Lee et al. 2002). In humans, mutations in XPG result in impaired transactivation (Ito et al. 2007), and XPG binds to the promoters and distal regions of several nuclear receptor (NR)-dependent genes in the absence of any genotoxic stress (Le May et al. 2010). Recently, XPG and XPF endonucleases have been found to be involved in chromatin looping between the promoter and the terminator of the activated RAR β 2 human gene (Le May et al. 2012).

In this study, we identified a functionally important interaction between the essential Med17 Mediator subunit and the Rad2/XPG DNA repair protein, suggesting a functional link between Mediator of transcription regulation and DNA repair. Genome-wide location analyses revealed that Rad2 was associated with Pol II-transcribed genes but also with Pol III-transcribed genes and telomeric regions in the absence of exogenous genotoxic stress. We showed that Rad2 occupancy of Pol II-transcribed genes was transcription-dependent. Moreover, Rad2 occupancy of class II gene promoters was highly correlated with that of Mediator. Several Mediator *med17* mutants were UV-sensitive in a GGR-deficient background, and their UV sensitivity was correlated with reduced Rad2 occupancy of class II genes and a concomitant decrease of the interaction between Mediator and Rad2 protein. We thus uncovered a previously unsuspected role of Mediator in DNA repair via Rad2/XPG recruitment.

Results

Physical interaction between the Med17 Mediator subunit and the Rad2/XPG protein

Previously, we characterized Mediator complex organization by screening the Mediator subunits with a yeast genomic library in a two-hybrid system. Protein–protein interactions within the Mediator complex have been published (Guglielmi et al. 2004). Many interactions that we uncovered in the screen were confirmed crystallographically and served as a basis for other Mediator studies (for example, Lariviere et al. 2012; Robinson et al. 2012). In addition, our two-hybrid screening revealed a large number of interactions between *S. cerevisiae* Mediator and other nuclear proteins (B Guglielmi, NL van Berkum, C Boschiero, FCP Holstege, and M Werner, unpubl.). One particularly interesting interaction is the contact between the Med17 Mediator head module subunit and the Rad2 endonuclease, the yeast homolog of human XPG protein involved in DNA repair. Med17 fused to the Gal4 DNA-binding domain interacted in the two-hybrid assay with the Rad2 fragment (549–857) fused to the Gal4 activation

domain (Fig. 1A). The contact between Mediator and Rad2 was confirmed by coimmunoprecipitation (co-IP) experiments with crude extracts of a yeast strain expressing Med17-Myc and HA-Rad2 from their native promoters (Fig. 1B; Supplemental Fig. S1). Our results show that Mediator coimmunoprecipitates with Rad2 in crude extracts and vice versa. Rad2 was detected by Western blotting when Med17-Myc was used to immunoprecipitate Mediator complex, and the Med17-Myc Mediator subunit was coimmunoprecipitated when Rad2 was immunoprecipitated via the HA tag.

Pol II is the main component of the transcription machinery and the first complex in TCR that recognizes the DNA damage (Svejstrup 2002, 2007; Hanawalt and Spivak 2008). The XPG protein was reported to coimmunoprecipitate with Pol II in crude extracts from undamaged HeLa cells (Sarker et al. 2005). In line with these results, our co-IP experiments also indicated that Rad2 coimmunoprecipitated with Pol II in yeast crude extracts

(Fig. 1C). Coimmunoprecipitated Pol II forms phosphorylated on Ser5 and Ser2 of the Rpb1 C-terminal repeat domain (CTD) were also detected by Western blotting (data not shown).

Genome-wide location analysis of Rad2

Recently, an association of XPG, the human homolog of Rad2, with several NR-dependent genes was shown in the absence of genotoxic stress (Le May et al. 2010, 2012). However, to date, no genomic data are available on the Rad2/XPG occupancy in humans or yeast. To determine the targets of Rad2 in the yeast genome in the absence of exogenous genotoxic stress, we performed Rad2 chromatin immunoprecipitation (ChIP)-chip experiments. Two yeast strains carrying N-terminal or C-terminal HA-tagged version of Rad2 were constructed and used in conventional ChIP experiments. No increase in the UV sensitivity of these strains was observed, suggesting that both tagged versions of Rad2 were functional (data not shown). We found that N-terminally and C-terminally tagged Rad2 were significantly enriched inside three constitutively expressed Pol II-transcribed genes tested: *ADH1*, *PYK1*, and *PMA1* (Supplemental Fig. S2A). The strain carrying an N-terminal tagged version of Rad2 was used for ChIP-chip analysis, since the ChIP signal to noise ratios were higher. We found 2471 oligonucleotides significantly bound by Rad2 (P -value < 0.005; out of 40,000 oligonucleotide probes present on the DNA arrays). The genomic profile of Rad2 occupancy was complex, spanning from Pol II-transcribed genes (promoter and transcribed regions) to, unexpectedly, Pol III-transcribed genes and telomeric/subtelomeric regions. Examples of Rad2 ChIP-chip enrichment profiles illustrate these different genomic region categories (Fig. 2A–C).

To gain in resolution and determine precisely the yeast genomic regions enriched by Rad2 protein, we performed ChIP-seq experiments. In agreement with our Rad2 ChIP-chip results, we confirmed the presence of this protein inside and on the promoter regions of Pol II-transcribed genes as well as on Pol III-transcribed genes. Telomeric regions are highly repetitive and therefore are poorly represented in ChIP-seq analysis. We identified 2600 significantly enriched Rad2 peaks using MACS2 peak calling software; 1203 of these peaks were located on promoter regions of Pol II-transcribed genes, 33 were on snRNA genes, and 268 were on Pol III-transcribed genes. This bioinformatics tool, MACS2, also allowed us to identify some Rad2 enrichment peaks inside the transcribed regions of class II genes but was not particularly adapted for detection of large Rad2 distributions in these regions, missing many Rad2-enriched ORFs. To identify precisely the Rad2 enrichment signals inside ORFs and thus complete the Rad2 ChIP-seq distribution analysis, we compared Rad2 mean densities for each ORF with density distributions of two negative controls: untagged strain and input DNA samples. We observed that 818 regions inside the protein-coding genes were significantly enriched by Rad2.

The results obtained by our genome-wide analyses were confirmed by conventional ChIP on a set of selected

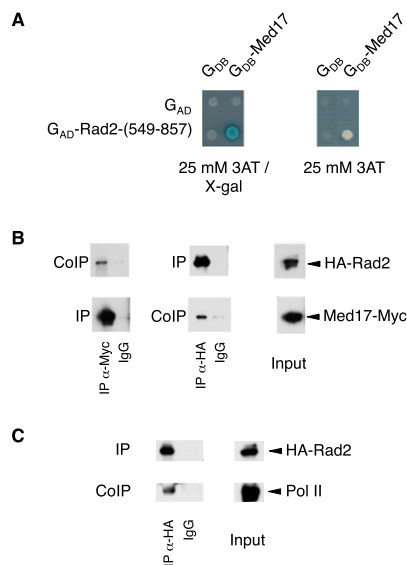


Figure 1. The Med17 Mediator subunit interacts with the Rad2/XPG DNA repair protein. (A) Two-hybrid interaction between the Med17 Mediator subunit and Rad2. Two-hybrid assays were performed in the Y190 strain, and the activation of the *lacZ* or *HIS3* reporters was tested using an X-Gal overlay plate assay or 3AT plate assay (Werner et al. 1993). Med17 was fused to the Gal4 DNA-binding domain (G_{DB} -Med17), and the Rad2-(549–857) fragment was fused to the Gal4 activation domain [G_{AD} -Rad2-(549–857)]. (B) Co-IP of Mediator and Rad2. Mediator was immunoprecipitated through the Med17-Myc subunit with α -Myc antibody from crude yeast extracts (Input; right panel) and analyzed by Western blotting with α -HA antibody (co-IP) against Rad2 (left panel). (Middle panel) Vice versa, Rad2 was immunoprecipitated with α -HA antibody from crude extracts and analyzed by Western blotting with α -Myc antibody (co-IP) against Med17. (IgG) Control immunoprecipitation with IgG magnetic beads only. (C) Co-IP of Rad2 and Pol II. Rad2 was immunoprecipitated with α -HA antibody from crude extracts (Input). Coimmunoprecipitated Pol II was detected by Western blotting with α -Rpb1 antibody (co-IP). (IgG) Control immunoprecipitation with IgG magnetic beads only.

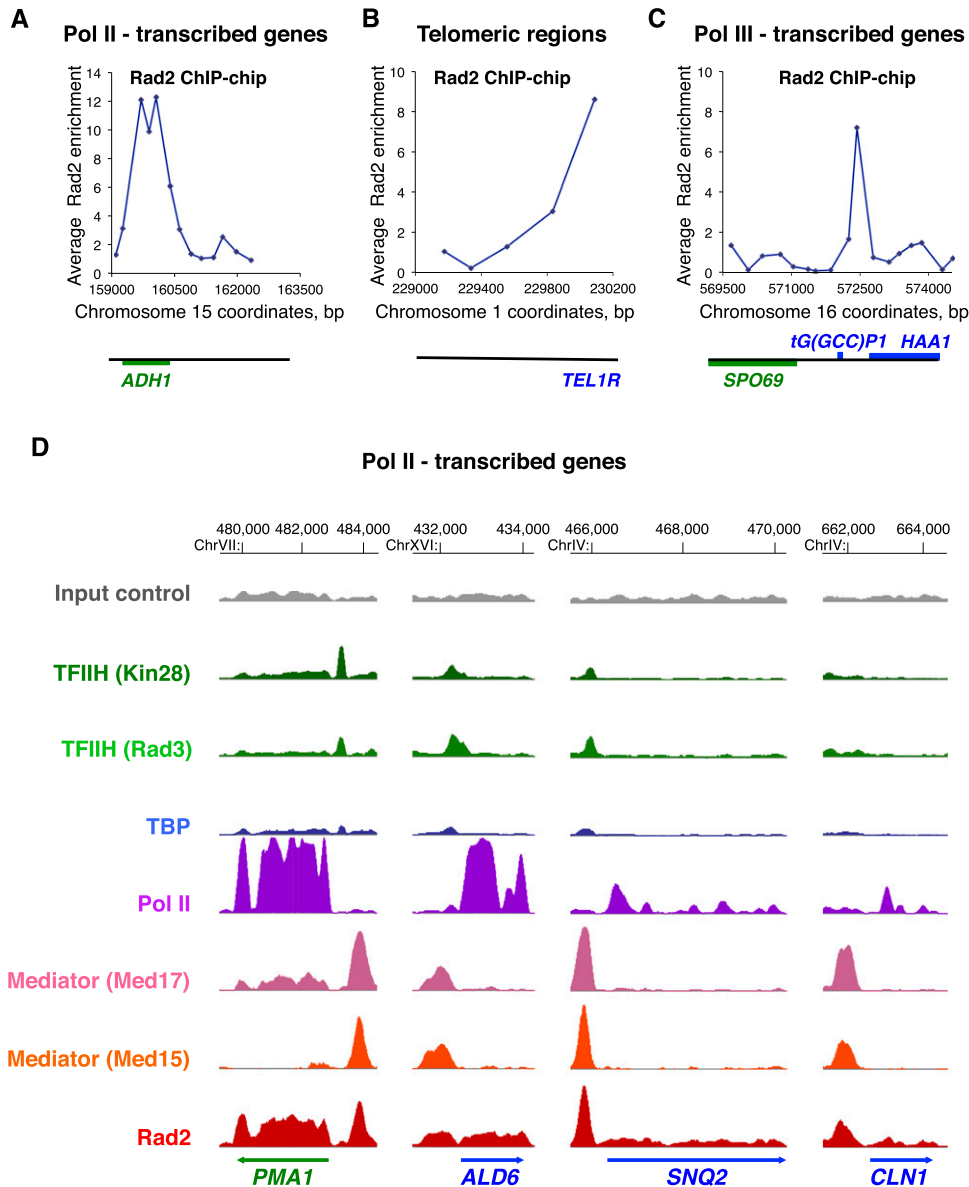


Figure 2. Genome-wide analysis of Rad2, Mediator, Pol II, TBP, and TFIIH. (A–C) Examples of Rad2 ChIP–chip enrichment profiles on Pol II-transcribed genes (A), on a telomeric region (B), and on a Pol III-transcribed gene (C). Rad2 enrichment profiles were assessed from ChIP–chip experiments with HA-Rad2 strain grown at 30°C in YPD medium. Watson strand-transcribed genes are colored in blue, and Crick strand-transcribed genes are colored in green. The genomic positions of probe regions on corresponding chromosomes are indicated along the X-axis. The enrichment ratio is indicated along the Y-axis. (D) Examples of Rad2, Mediator, Pol II, TBP, and TFIIH ChIP-seq enrichment profiles on selected Pol II-transcribed genes. Densities of sequence tags were assessed from ChIP-seq experiments performed with HA-Rad2-tagged, Med15-HA-tagged (Mediator), Med17-HA-tagged (Mediator), HA-TBP-tagged, Rad3-HA-tagged (core TFIIH), and Kin28-HA-tagged (TFIIH kinase module) strains with α -HA antibody. Pol II was immunoprecipitated using α -Rpb1 antibody. ChIP-seq density profiles are displayed using IGB yeast genome browser. Watson strand-transcribed genes are colored in blue, and Crick strand-transcribed genes are colored in green. Input DNA and DNA from ChIP with an untagged strain were used as negative controls. Densities of sequence tags were displayed after subtraction of normalized control of untagged strain as described in the Materials and Methods.

regions representing the categories of Rad2 genomic targets that we identified. Supplemental Figure S2B illustrates significant enrichment of Rad2 on several promoter and transcribed regions of selected class II genes compared with background levels measured on the coding region of the nontranscribed *GAL1* gene. All selected Pol

III-transcribed genes displayed significant Rad2 enrichment (Supplemental Fig. S2C). The spatial Rad2 distribution on the longest class III gene, *SCR1*, was similar to that of Pol III (Supplemental Fig. S2D; Ghavi-Helm et al. 2008). We also confirmed the presence of Rad2 on telomeric regions (Supplemental Fig. S2E). In addition,

a significant Rad2 enrichment was detected on the Pol I-transcribed 35S rDNA region compared with the intergenic *NTS2* region (Supplemental Fig. S2F). Taken together, these results indicated a complex Rad2 distribution on the yeast genome in the absence of exogenous genotoxic stress, spanning the genes of all three transcription classes and telomeric/subtelomeric regions. For further study, we focused particularly on Pol II-transcribed genes.

Genome-wide Rad2 occupancy correlates with Mediator occupancy on promoter regions of Pol II-transcribed genes

To compare Rad2 ChIP-seq profiles with those of Mediator, Pol II, and GTFs, we performed ChIP-seq experiments for Mediator (Med15 and Med17 subunits), Pol II, and GTFs (TBP and two TFIID subunits: the Rad3/XPD subunit of core TFIID and the Kin28/Cdk7 kinase subunit of TFIID). Input DNA and DNA from ChIP with an untagged strain were used as negative controls. Figure 2D and Supplemental Figure S3 show examples of ChIP-seq density distributions using the Integrated Genome Browser (IGB) for all of the proteins on selected class II (Fig. 2D) or class III (Supplemental Fig. S3) genes. Transcription by all three RNA polymerases requires TBP protein. As expected, density peaks of this protein were also detected on the Pol III-transcribed genes. These TBP peaks on class III genes coincided with Rad2 peaks (Supplemental Fig. S3). On the selected promoter regions of Pol II-transcribed genes, there was a particular correlation between Rad2 and Mediator for the location and intensity of peaks (Fig. 2D). Furthermore, genome-wide Mediator and Rad2 occupancies were well correlated. Rad2 and Mediator ChIP-seq densities (25-base-pair [bp] bin) on the yeast genome were computed and compared (Fig. 3A; Supplemental Fig. S4A). As expected, the Rad2 enrichment of several genomic categories, including class II promoter and transcribed regions, made the correlation between Rad2 and Mediator densities complex. Considering only class II gene promoter regions, regression analysis showed a linear correlation with a correlation coefficient of 0.75 for the Med15 Mediator subunit (Fig. 3B) or 0.73 for the Med17 Mediator subunit (Supplemental Fig. S4B). Intergenic regions encompassing Pol III-transcribed genes were excluded from this analysis. The global correlation between Rad2 and Mediator occupancies on promoter regions is consistent with a functional link between Rad2 and the Mediator complex. We determined whether Rad2 association with Pol II-transcribed regions was correlated with Pol II occupancy (Fig. 3C). Considering only Pol II-transcribed regions inside ORFs, encompassing most of the yeast Pol II ChIP signals, no strong correlation could be observed between Rad2 and Pol II ChIP-seq densities (Fig. 3D). Regression analysis showed a lower correlation coefficient of 0.54. In comparison, regression analysis indicated, as expected, a linear correlation between two subunits of the Mediator complex (Med15 and Med17) or TFIID complex (Rad3 and Kin28), with a high correlation coefficient of 0.89 (Sup-

plemental Fig. S4C,D) or 0.88 (Supplemental Fig. S5A,B), respectively, and between TBP and Rad3 or Kin28 TFIID subunits with a high correlation coefficient of 0.75 or 0.81, respectively (Fig. 3E). Pairwise correlations between ChIP-seq densities of Rad2, Mediator, Kin28, Rad3, and TBP showed that the strongest correlation for Rad2 protein occupancy on promoter regions was with Mediator (Fig. 3E). On the other hand, no strong correlation was observed between Rad2 and Kin28 ChIP-seq occupancies, with a correlation coefficient of 0.53 (Supplemental Fig. S5C,D).

The location of Rad2 peaks on the promoter regions of Pol II-transcribed genes was compared with those of Mediator and GTFs. Mean ChIP-seq read densities were calculated for each protein on promoter regions of Pol II-transcribed genes centered on Mediator enrichment peaks (Fig. 3F). To consider only Pol II-transcribed gene promoters, intergenic regions encompassing Pol III-transcribed genes were excluded. These analyses showed that Rad2 peak distribution coincides with that of Mediator (Med15 and Med17 subunits) but not with those of TFIID core (Rad3) or kinase (Kin28) modules. Taken together, our results revealed that genome-wide Rad2 occupancy was highly correlated with that of Mediator on promoter regions of Pol II-transcribed genes and that Rad2 enrichment distribution in these regions superposed with that of Mediator.

Rad2 occupancy on Pol II-transcribed genes is transcription-dependent

An essential question arising from Rad2 recruitment in the absence of exogenous genotoxic stress is whether Rad2 occupancy on class II genes depends on Pol II transcription. To address this question, we examined Rad2 occupancy on the *GAL1* gene by ChIP experiments upon galactose induction. When yeast cells were grown in raffinose-supplemented medium, Rad2 was absent from the *GAL1* promoter and ORF (Fig. 4A). After 60 min of galactose induction, Rad2 was recruited to the promoter and transcribed region of the *GAL1* gene. The constitutively expressed *ADH1* gene was always bound by the Rad2 protein. As expected, the *GAL1* mRNA level was highly induced after galactose induction, while the *ADH1* mRNA level was comparable in raffinose- and galactose-supplemented medium (Fig. 4B). These results indicated that Rad2 recruitment to *GAL1* gene was transcription-dependent.

To address more generally the question of the transcriptional dependency of Rad2 occupancy on Pol II-transcribed genes, we used a Pol II mutant, *rpb1-1*, that rapidly stops transcription after a shift to nonpermissive temperature (Nonet et al. 1987). When cells were incubated for 90 min at 37°C, as expected, Pol II occupancy on Pol II-transcribed genes was strongly reduced in the *rpb1-1* mutant compared with the wild-type strain (Fig. 4C), whereas Pol III occupancy on tested class III genes remained unchanged (Fig. 4D). We also checked that, as expected, RNA levels were strongly decreased in the *rpb1-1* mutant for class II genes (Fig. 4E) but not class III genes (Fig. 4F). Rad2 occupancy on Pol II-transcribed genes followed that of Pol II and was also greatly reduced after incubation at 37°C (Fig. 4G; Supplemental S6A,B).

Eyboulet et al.

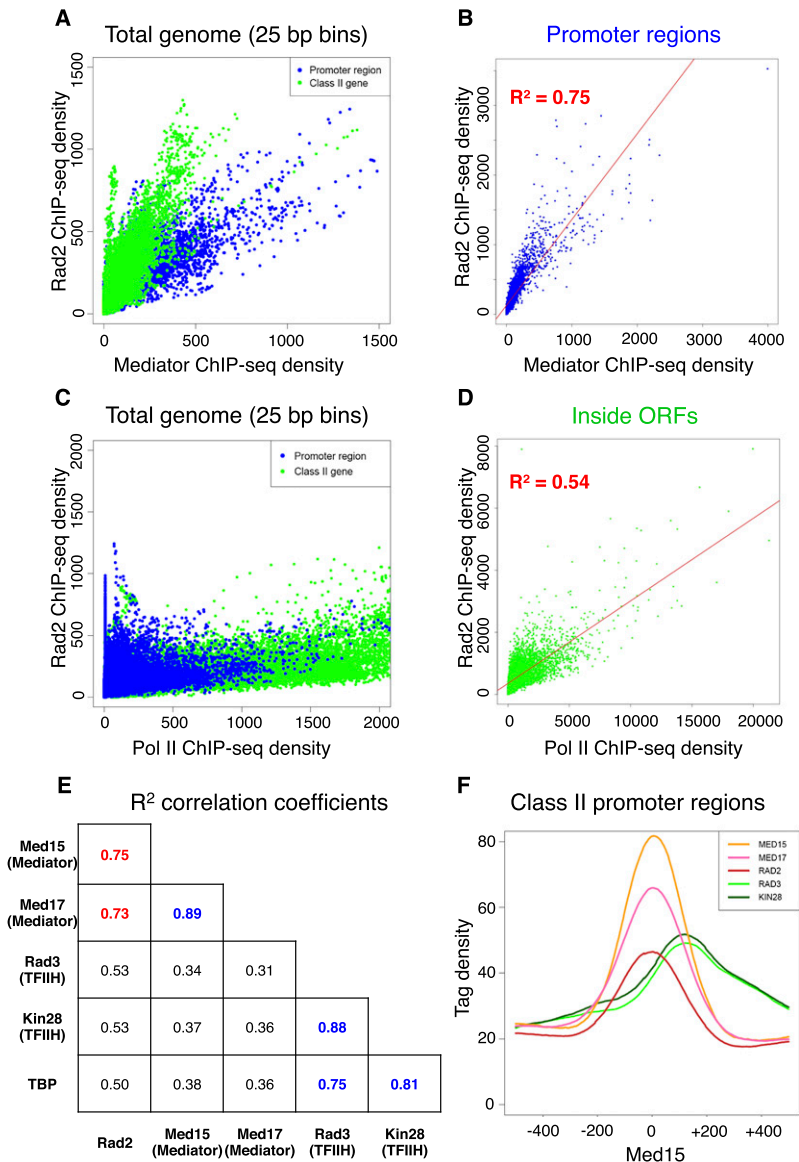


Figure 3. Correlation of Rad2 ChIP-seq density with those of Mediator and Pol II. (A) Rad2 ChIP-seq density versus Mediator (Med15 subunit) ChIP-seq density on total yeast genome. Density of sequence tags in Mediator and Rad2 ChIP-seq experiments has been calculated for each 25-bp bin of yeast genome. Each point on the plot corresponds to one genomic bin. The genomic bins located inside ORFs or class II promoter regions are indicated in green and blue, respectively. Pol III-transcribed genes were excluded from this analysis. (B) Rad2 ChIP-seq density versus Mediator (Med15 subunit) ChIP-seq density on class II promoter regions. The density of sequence tags in Mediator and Rad2 ChIP-seq experiments has been calculated for promoter region of each Pol II-transcribed gene. Each point on the plot corresponds to one promoter region. Promoter regions correspond to intergenic regions for Pol II-transcribed genes in tandem or in divergent orientation, excluding intergenic regions encompassing Pol III-transcribed genes. A linear regression for Rad2 ChIP-seq density versus Mediator ChIP-seq density (red line) and a R^2 correlation coefficient are indicated. (C) Rad2 ChIP-seq density versus Pol II ChIP-seq density on total yeast genome, as described in A. (D) Rad2 ChIP-seq density versus Pol II ChIP-seq density inside ORFs. The density of sequence tags in Rad2 and Pol II ChIP-seq experiments has been calculated for each ORF. Each point on the plot corresponds to one ORF. A linear regression for Rad2 ChIP-seq density versus Pol II ChIP-seq density (red line) and a R^2 correlation coefficient are indicated. (E) Pairwise correlation coefficients between ChIP-seq densities of Rad2, Med15 (Mediator), Med17 (Mediator), Rad3 (TFIIH core module), Kin28 (TFIIH kinase module), and TBP on promoter regions. R^2 correlation coefficients for each pairwise combination are shown. A correlation coefficient for Mediator versus Rad2 is highlighted in red, and pairwise correlation coefficients for different GTF (Rad3, Kin28, and TBP) or Mediator subunit combinations (Med15 and Med17) are shown in blue. (F) Distribution of Mediator, Rad2/XPG in transcription, since such a role has been previously suggested (Lee et al. 2002; Le May et al. 2010, 2012). In *S. cerevisiae*, the *RAD2* gene is not essential for viability. We started by examining whether, under standard growth conditions (rich glucose medium, YPD), any effect of *rad2* deletion on Mediator or Pol II occupancy could be identified. We did not observe any difference between Mediator occupancy of the *rad2Δ* strain compared with the wild-type strain on constitutively expressed gene promoters (Fig. 5A; Supplemental Fig. S7A,B). In addition, Pol II occupancy on selected genes and the corresponding mRNA level in the *rad2Δ* strain were similar to those of the wild-type strain (Fig. 5B,C; Supplemental Fig. S7C). Similarly, no effect of *rad2* deletion on Mediator or Pol II

Rad2, and TFIIH ChIP-seq densities on Pol II-transcribed gene promoters. Tag densities are shown relative to the localization of the maximum of Mediator enrichment peaks. Intergenic regions encompassing Pol III-transcribed genes were excluded. The tag density was determined for each protein in a 1000-bp window centered on the maximum of Mediator enrichment peaks per each nucleotide position. Mean tag density was then calculated and plotted over the window.

Rad2 occupancy decrease was observed for Pol II-transcribed gene promoters and ORFs. This decrease of Rad2 occupancy was not related to Mediator occupancy, since it did not change in the *rpb1-1* mutant (Supplemental Fig. S6C–E). Conversely, Rad2 occupancy on class III genes or telomeric regions remained unchanged under the same conditions (Fig. 4G). Thus, the presence of Rad2 on class II genes is specifically dependent on Pol II transcription.

UV irradiation sensitivity is the only observed phenotype of *rad2* deletion

One hypothesis for the functional interplay between Mediator and Rad2 consists of a possible implication of

Rad2/XPG in transcription, since such a role has been previously suggested (Lee et al. 2002; Le May et al. 2010, 2012). In *S. cerevisiae*, the *RAD2* gene is not essential for viability. We started by examining whether, under standard growth conditions (rich glucose medium, YPD), any effect of *rad2* deletion on Mediator or Pol II occupancy could be identified. We did not observe any difference between Mediator occupancy of the *rad2Δ* strain compared with the wild-type strain on constitutively expressed gene promoters (Fig. 5A; Supplemental Fig. S7A,B). In addition, Pol II occupancy on selected genes and the corresponding mRNA level in the *rad2Δ* strain were similar to those of the wild-type strain (Fig. 5B,C; Supplemental Fig. S7C). Similarly, no effect of *rad2* deletion on Mediator or Pol II

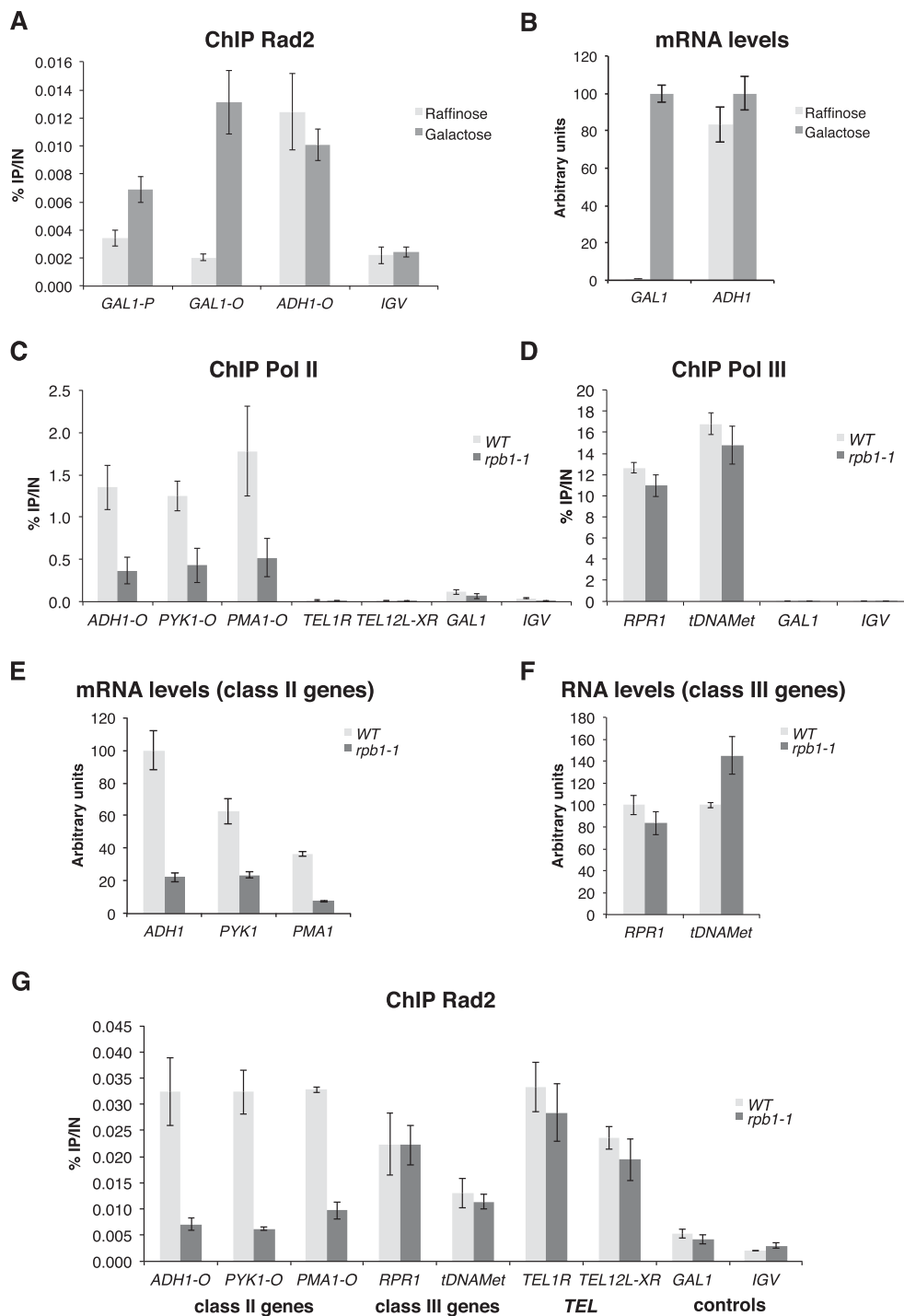


Figure 4. Transcription dependency of Rad2 occupancy on Pol II-transcribed genes. (A) Quantitative ChIP analysis of Rad2 occupancy on the *GAL1* gene before and after galactose induction. The HA-Rad2 tagged strain was grown in a raffinose-supplemented medium, and then galactose was added for 60 min. Quantitative ChIP experiments were performed with α -HA antibody against HA-Rad2. Mean values and standard deviation (indicated by error bars) of three independent experiments are shown. Immunoprecipitated fragments from ChIP experiments were amplified with primers corresponding to the *GAL1* gene promoter (*GAL1 P*) or ORF (*GAL1 O*). *ADH1* ORF represents a constitutively expressed gene control. A nontranscribed region on chromosome V (*IGV*) was used as a negative control. (B) mRNA level of *GAL1* and *ADH1* genes before and after galactose induction. Yeast cells were grown in a raffinose-supplemented medium, and then galactose was added for 60 min. RNA levels for *GAL1* and *ADH1* genes were quantified by RT-qPCR and normalized to 25S rRNA, 18S rRNA, and *SCR1* levels. Values are expressed in arbitrary units representing a percentage of the value in galactose-supplemented medium. Mean values and standard deviation (indicated by error bars) of three independent experiments are shown. (C,D,G) Effect of *rpb1-1* mutation on Rad2, Pol II, and Pol III occupancies on selected regions. Quantitative ChIP assays were performed using α -Rpb1 antibody (Pol II) (C), α -Myc antibody against Rpc160-Myc (Pol III) (D), and α -HA antibody against HA-Rad2 (G). Cells were grown in selective SD medium complemented with amino acids at 25°C and then shifted for 90 min at 37°C. A *GAL1* ORF and a nontranscribed region on chromosome V (*IGV*) were used as negative controls. (E,F) Effect of *rpb1-1* mutation on RNA level of class II (E) or class III (F) genes. Cells were grown in selective SD medium complemented with amino acids at 25°C and then shifted for 90 min at 37°C. RNA levels for selected genes were quantified by RT-qPCR and normalized to 25S rRNA, 18S rRNA, and *SCR1* levels. Values are expressed in arbitrary units representing a percentage of the value in the wild-type strain for the *ADH1* gene (E) or *RPR1* or *tDNAMet* genes (F). Mean values and standard deviation (indicated by error bars) of three independent experiments are shown.

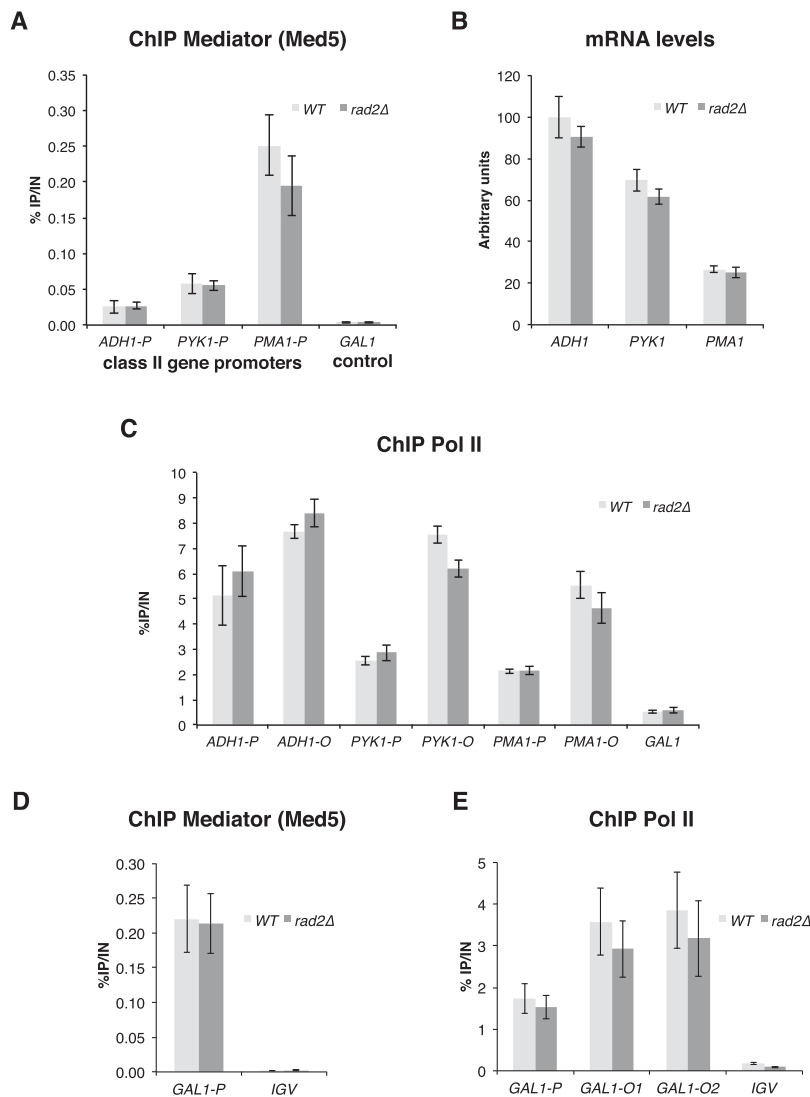


Figure 5. Mediator and Pol II occupancies in the *rad2* deletion strain on selected regions. (A,C) Quantitative ChIP analysis of Mediator and Pol II occupancies under standard growth conditions. *rad2Δ* and wild-type strains were grown in YPD medium at 30°C. Immunoprecipitations were performed using α -HA antibody against Med5-HA (Mediator) (A) and α -Rpb1 antibody (Pol II) (C). (A) Immunoprecipitated fragments from Mediator ChIP experiments were amplified with primers corresponding to selected class II gene promoters. (C) Immunoprecipitated fragments from Pol II ChIP experiments were amplified with primers corresponding to selected class II gene promoters (P) or ORFs (O). Mean values and standard deviation (indicated by error bars) of three independent experiments are shown. A *GAL1* ORF was used as a negative control. (B) Effect of *rad2Δ* mutation on mRNA levels of class II genes. *rad2Δ* and wild-type strains were grown in YPD medium at 30°C. RNA levels for selected genes were quantified by RT-qPCR and normalized to 25S rRNA, 18S rRNA, and *SCR1* levels. Values are expressed in arbitrary units representing a percentage of the value in the wild-type strain for *ADH1* gene. Mean values and standard deviation (indicated by error bars) of three independent experiments are shown. (D,E) Quantitative ChIP analysis of Mediator and Pol II occupancies upon galactose induction. *rad2Δ* and wild-type strains were grown in a raffinose-supplemented medium, and then galactose was added for 60 min. Quantitative ChIP experiments were performed with α -HA antibody against Med5-HA (Mediator) (D) and α -Rpb1 antibody (Pol II) (E). Mean values and standard deviation (indicated by error bars) of three independent experiments are shown. Immunoprecipitated fragments from ChIP experiments were amplified with primers corresponding to *GAL1* gene promoter (*GAL1-P*) or ORF (*GAL1-O1* and *GAL1-O2*). A nontranscribed region on chromosome V (*IGV*) was used as a negative control.

occupancy on the *GAL1* gene was observed in galactose-inducing medium (Fig. 5D,E). Previously, Lee et al. (2002) have stated that *rad2* deletion slowed down the kinetics of galactose-inducible genes. To check this possibility, the Pol II recruitment to the *GAL1* gene promoter and transcribed region and the *GAL1* mRNA level were examined at different times after galactose induction by ChIP and quantitative RT-PCR (qRT-PCR) experiments. However, there was no difference between the *rad2Δ* mutant and the wild-type strain (Supplemental Fig. S8).

To identify growth phenotypes of *rad2* deletion strains and therefore a possible transcriptional role for Rad2, we tested different growth conditions in two backgrounds: BY4741 and YPH499. No growth difference between *rad2Δ* and wild-type strains was observed on YPD medium under different temperature conditions (16°C, 25°C, 30°C, or 37°C), on different carbon sources (YP medium supplemented with galactose, lactate, ethanol, or glycerol), and on drugs to test for multidrug resistance transcrip-

tional response (cycloheximide and ketoconazole) or for NTP depletion conditions (6-azauracil and mycophenolic acid) (Supplemental Fig. S9; data not shown). The only phenotype of the *rad2* deletion strains that we observed is a high UV sensitivity, as expected for a DNA repair protein. Taken together, our results strongly suggest that Rad2 does not play a major role in the transcriptional process in yeast, at least in the two backgrounds examined here.

Physiological significance of Mediator–Rad2 contact in DNA repair

We could not identify a major implication of Rad2 in yeast Pol II transcription. This led to the hypothesis that Mediator might play a role in DNA repair via its interaction with the Rad2 protein. To investigate the functional relationship between Mediator and Rad2, we used our collection of 29 conditional mutants for the Med17 essential Mediator subunit (Soutourina et al.

2011). The integrity of Mediator in *med17* mutants has been verified by co-IP experiments (Soutourina et al. 2011). To directly address a possible role of Mediator in DNA repair, a UV sensitivity phenotype was systematically investigated for *med17* mutants. No UV sensitivity was found for *med17* mutants in the context where both NER pathways, GGR and TCR, were functional (Supplemental Fig. S10). Next, we used a GGR-deficient context (*rad7Δ* strains) to analyze specifically the potential implication of Med17 in the TCR pathway by UV sensitivity spot assays. Several Mediator mutants among our collection of *med17-ts* were UV-sensitive in a *rad7Δ* GGR-deficient context (Fig. 6A). To refine the UV sensitivity analysis, we determined UV survival curves for several *med17* mutants (three UV-sensitive *med17-140*, *med17-233*, and *med17-257* mutants and one UV-insensitive *med17-L546P* mutant) alone or in combination with *rad7Δ* mutation (Fig. 6B). The UV dose response analysis further confirmed UV sensitivity spot assays and clearly indicated that the combination of *rad7* deletion with several *med17* mutations leads to an increased UV sensitivity. An extremely UV-sensitive *rad2* deletion mutant was also included for comparison. Interestingly, the UV sensitivity of *med17-ts rad7Δ* mutants was comparable with that of the *rad26Δ rad7Δ* double mutant

deficient in both GGR (*rad7Δ*) and TCR (*rad26Δ*) NER pathways (Fig. 6B). Rad26 is the yeast homolog of the TCR-specific human CSB protein. In yeast, *RAD26* deletion does not lead to UV sensitivity but significantly increases the UV sensitivity of GGR-deficient mutants. Since *med17* mutants were UV-sensitive in a GGR-deficient context, we wondered whether these mutants could be epistatic with the TCR-deficient *rad26* deletion mutant. No UV sensitivity increase was found for *med17* mutants in the *rad26Δ* context (Supplemental Fig. S11), demonstrating that, indeed, *med17* mutants are epistatic with the *rad26Δ* mutant. Taken together, these results are consistent with a specific implication of Mediator in the TCR pathway. Our results also indicated that the UV sensitivity of *med17-ts* mutants was not a consequence of a transcriptional effect, since this phenotype was observed only in a GGR-deficient context and not in a context where both NER pathways, GGR and TCR, were functional or in a TCR-deficient context. Three of the UV-sensitive *med17-ts* mutants were selected for further molecular study (*med17-257*, *med17-140*, and *med17-233*) (the results on the two latter mutants are shown in Supplemental Fig. S12). Rad2 occupancy in these mutants was determined by ChIP experiments for selected Pol II-transcribed genes. Figure 7A and Supplemental Figure

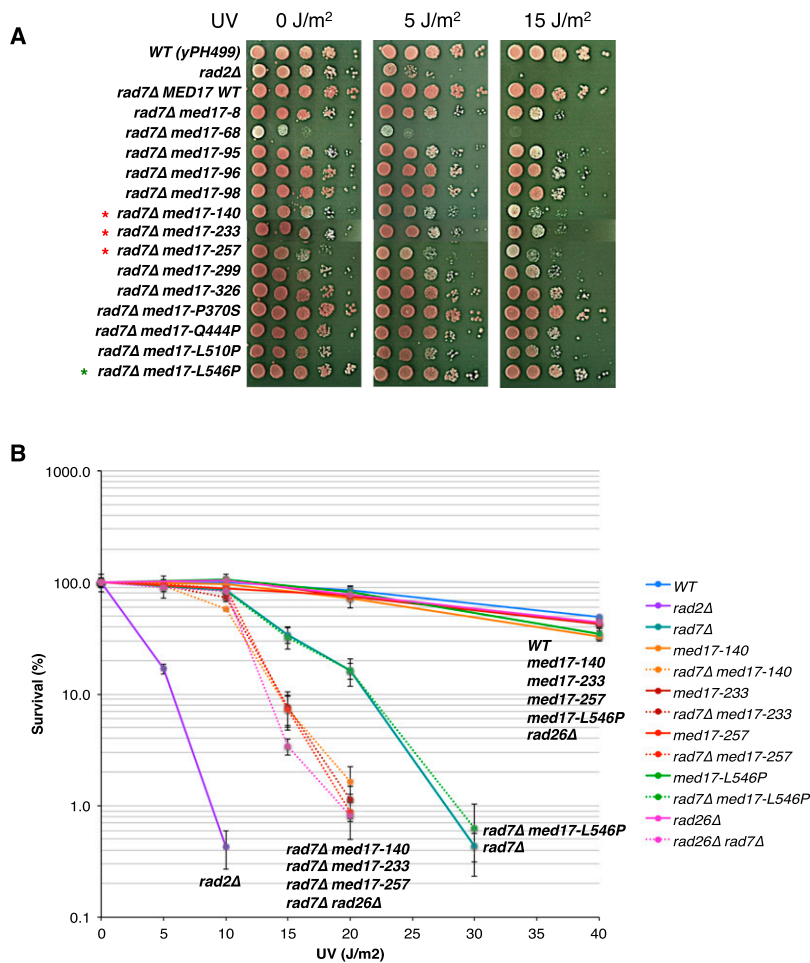


Figure 6. UV sensitivity of *med17* mutants. (A) UV sensitivity of *med17* mutants in a GGR-deficient *rad7Δ* context by spot assays. Cells were serially diluted, spotted on YPD agar plates, treated or not with the indicated doses of UV (UV Stratalinker 1800), and incubated for 3 d at 30°C. Red asterisks mark selected UV-sensitive *med17* mutants, and a green asterisk marks a selected UV-insensitive *med17* mutant. (B) UV survival curves of *med17* and *rad26Δ* mutants alone and in a GGR-deficient *rad7Δ* context. Cells were grown in YPD medium to a mid-exponential phase, diluted, plated on YPD agar, and irradiated with the different doses of UV (UV Stratalinker 1800). Plates were incubated for 3 d at 30°C in the dark before colonies were counted and survival was calculated. Mean values and standard deviation (indicated by error bars) of at least three independent experiments are shown.

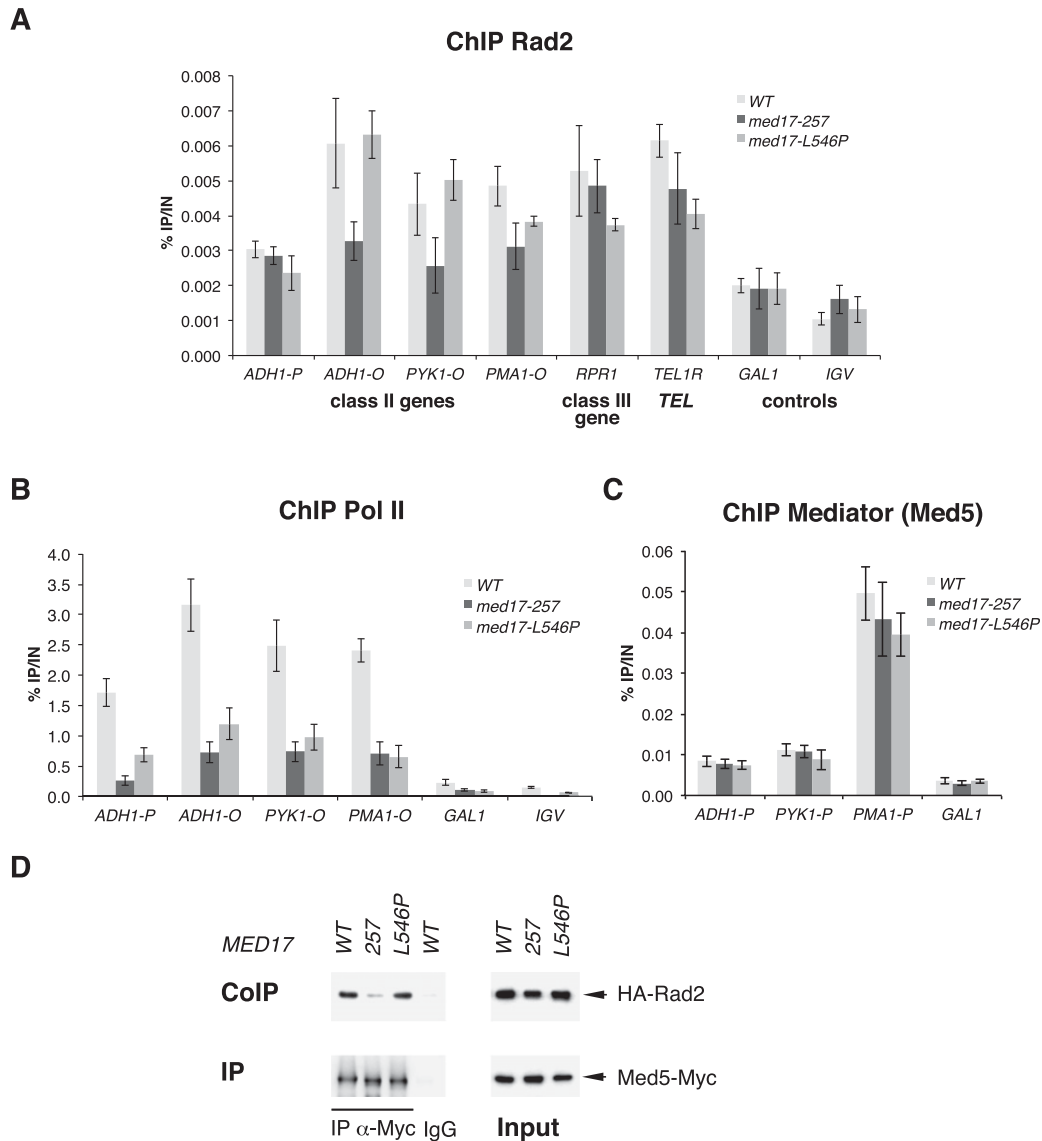


Figure 7. Effects of *med17* mutants on Rad2, Pol II, and Mediator occupancies and on Rad2–Mediator interaction. (A–C) Quantitative ChIP analysis of Rad2, Pol II, and Mediator occupancies on selected regions. ChIP assays were performed using α -HA antibody against HA-Rad2 (A), α -Rpb1 antibody (Pol II) (B), and α -HA antibody against Med5-HA (Mediator) (C). Cells were grown in YPD medium at 30°C and then shifted for 45 min at 37°C. Immunoprecipitated fragments from ChIP experiments were amplified with primers corresponding to selected class II gene promoters (P) or ORFs (O), a selected class III gene, and a telomeric region. Mean values and standard deviation (indicated by error bars) of three independent experiments are shown. A *GAL1* ORF and a nontranscribed region on chromosome V (*IGV*) were used as negative controls. (D) Co-IP of Mediator and Rad2. Mediator was immunoprecipitated through a Med5-Myc subunit with a α -Myc antibody from crude extracts (Input; right panel) and analyzed by Western blotting with α -HA antibody (co-IP) against Rad2 (left panel). (IgG) Control immunoprecipitation with IgG magnetic beads only.

S12A show a significant decrease in Rad2 association in UV-sensitive Mediator mutants compared with the wild-type strain and *med17-L546P* mutant, one of the UV-insensitive *med17-ts* mutants. Rad2 occupancy to a class III gene and a telomeric region was not affected in these mutants. All tested *med17* mutations, including *med17-L546P*, which is UV-insensitive, led to a marked decrease in Pol II association with Pol II-transcribed genes but did not modify the Mediator association to corresponding gene promoters (Fig. 7B,C; Supplemental Fig. S12B,C). A

decrease in Pol II occupancy was not responsible for Rad2 occupancy defects of UV-sensitive mutants (*med17-257*, *med17-140*, and *med17-233*) or their UV sensitivity, since a reduction of Pol II occupancy in the *med17-L546P* mutant, similar to that of the UV-sensitive *med17-ts* mutants, was not accompanied by Rad2 occupancy decrease and UV sensitivity. To determine the effects of *med17-ts* mutations on Mediator contact with the Rad2 protein, we performed co-IP experiments and showed that Mediator–Rad2 interaction was reduced in UV-sensitive

mutants *med17-257*, *med17-140*, and *med17-233* compared with the wild-type strain and the *med17-L546P* UV-insensitive mutant (Fig. 7D; Supplemental Fig. S12D). These results indicated that Mediator is involved in Rad2 recruitment and therefore suggest that it is involved in DNA repair.

Discussion

Substantial evidence has been obtained on the essential function of the Mediator complex in transcription activation. However, our experiments indicate that this complex plays an important role in linking transcription with other nuclear processes. In this study, we identified a functionally important contact between the Med17 Mediator subunit and the Rad2 DNA repair protein, the *S. cerevisiae* homolog of human XPG. Mediator might play a role in DNA repair by loading Rad2/XPG to active genes, or Rad2/XPG could be implicated in transcription, influencing PIC assembly or stability on several gene promoters or both. Our results suggest that Mediator is involved in DNA repair through a functional link with the Rad2/XPG protein. (1) The essential Med17 Mediator subunit interacts with Rad2/XPG DNA repair protein. (2) We showed that Rad2 is located on Pol II-transcribed genes (promoter and transcribed regions) in the absence of exogenous genotoxic stress. (3) Genome-wide Rad2 occupancy of class II gene promoters is well correlated with that of Mediator. (4) Moreover, Rad2 distribution on these regions overlapped with that of Mediator but not of TFIIF, further highlighting a functional cooperation between Mediator and Rad2/XPG. It should be noted that Mediator and Rad2 ChIP-seq peaks on class II gene promoters are generally located on the upstream regulating regions, whereas TFIIF and other GTFs peaks are located on core promoter regions close to transcription start sites. (5) Rad2 occupancy of Pol II-transcribed genes is transcription-dependent, since *rpb1-1* Pol II mutation reduces Rad2 occupancy after a shift to nonpermissive temperature when transcription is rapidly inhibited, and since Rad2 is recruited to *GAL1* gene upon galactose induction of this gene. (6) However, no growth phenotypes (except UV sensitivity) or transcriptional effects were observed in the *rad2Δ* context, suggesting that Rad2 does not play a major role in the transcriptional process in yeast, at least in our two backgrounds. (7) On the contrary, Mediator *med17-ts* mutants are UV-sensitive in a GGR-deficient background (*rad7Δ*) and are epistatic with a TCR-deficient mutant (*rad26Δ*). (8) This UV sensitivity of several *med17* mutants is correlated with reduced Rad2 occupancy of class II genes and concomitant decrease of interaction with Rad2 protein.

We suggest that the contact between Med17 and Rad2 is direct, since our two-hybrid screening permitted us to reveal only specific interactions between different subunits within the Mediator complex (Guglielmi et al. 2004) that were further confirmed as direct by structural analysis (for example, Lariviere et al. 2012; Robinson et al. 2012). Moreover, we demonstrated that several mutations in the Med17 Mediator subunit resulted in a decreased

interaction between Mediator and Rad2, an observation in favor of a direct Mediator–Rad2 contact.

Therefore, Mediator is likely to facilitate Rad2 recruitment to transcribed genes, setting the stage for rapid DNA lesion removal. We propose that Mediator could function in transcription-coupled DNA repair, revealing a previously unknown role of the Mediator complex as a link between transcription and DNA repair. Whether Rad2/XPG is the only NER factor that is recruited through an interaction with Mediator or assembles with other NER factors to remove lesions on transcribed genes remains to be addressed. Further investigations will determine whether Mediator is engaged in functional interplay with other NER factors.

In the light of Mediator and Rad2/XPG conservation from yeast to humans, the molecular events governing the Mediator link with DNA repair are likely to exist in all eukaryotes. In support of this idea, we observed the Mediator–XPG contact by co-IP experiments with crude extracts from HeLa cells (data not shown). The conservation of Mediator connection to DNA repair might give insights into our understanding of human diseases like XP/CS.

Interestingly, both Mediator and Rad2/XPG interact with Pol II, the main component of transcription machinery and the first complex in TCR that recognizes the DNA damage. The mechanisms of functional orchestration of these contacts remain to be investigated. Our genome-wide analyses show that Mediator and Rad2 co-occupy class II gene promoter regions and, moreover, that a high correlation is observed between the two profiles on promoter regions. Rad2 is also bound to Pol II-transcribed regions that are occupied by Pol II. However, a correlation between Rad2 and Pol II on these regions was modest, suggesting a complex relationship between Rad2 and Pol II. At this stage, we propose that Mediator facilitates Rad2 recruitment on class II genes promoters. While we do not know what the precise mechanisms of Rad2 loading on ORFs are, our results suggest that they could entail both the loading of Rad2 on Pol II and another independent mechanism. A modest genome-wide correlation between Rad2 and Pol II contrasts with the good genome-wide correlation between TFIIS and Pol II or Pol III occupancies on Pol II- or Pol III-transcribed genes, respectively (Ghavi-Helm et al. 2008). These observations suggested that TFIIS associates at the 5' end of genes and accompanies Pol II and Pol III but cannot be loaded on traveling enzymes. It should be noted that one of the *med17* mutants used in this study (*med17-257*) that showed a UV-sensitive phenotype was lethal in combination with the *rpb3-2* Pol II mutation (Soutourina et al. 2011) and was affected in Pol II–Mediator interaction (data not shown). However, other *med17* mutations with UV-sensitive phenotypes did not lead to *rpb3-2* lethality phenotypes. Our results showed that Pol II occupancy was reduced in UV-sensitive and -insensitive *med17* mutants. Thus, the lower occupancy of Pol II cannot explain the UV sensitivity of the *med17* mutants that are defective in Rad2 interaction. We also showed that the UV sensitivity of *med17* mutants was not a consequence

of a transcriptional defect, since this phenotype was observed only in a GGR-deficient context and not in a context where both NER pathways were functional or a TCR-deficient context. Taken together, our results suggest that UV sensitivity of Mediator mutants was not an indirect consequence of their transcriptional defects.

A transcriptional role for the NER factor Rad2/XPG was first proposed in the yeast *S. cerevisiae* (Lee et al. 2002). Rad2 was suggested to promote efficient Pol II transcription. However, we did not observe the mild growth phenotypes of *rad2* deletion strains or slower galactose induction kinetics in the absence of the *RAD2* gene reported by Lee et al. (2002). Moreover, we did not find any genetic interaction between *rad2Δ* and *rad26Δ*. Rad26 is the yeast homolog of the TCR-specific human CSB protein. This result might be due to the different genetic backgrounds used by us and Lee et al. (2002). In spite of our considerable efforts, we did not obtain any evidence for a direct implication of Rad2 in transcription. In human cells, it has been demonstrated that mutations in XPG prevent its association with TFIIH, resulting in the dissociation of CAK and XPD/Rad3 from the core TFIIH and in impaired transactivation (Ito et al. 2007). Recently, XPG and other NER factors were shown to be associated with the promoters and distal regions of several NR-dependent genes in the absence of any exogenous genotoxic stress (Le May et al. 2010). The presence of these repair proteins at the promoters of activated genes is sensitive to transcription inhibitors and necessary for optimal DNA demethylation and histone post-translational modifications (H3K4/H3K9 methylation and H3K9/14 acetylation) and thus for efficient transcription. Deficiencies in some NER factors impede the recruitment of others and affect NR transactivation. Recently, XPG and XPF endonucleases were found to be involved in chromatin looping between the promoter and the terminator of the activated *RARβ2* human gene (Le May et al. 2012). It should be noted that NRs do not exist in *S. cerevisiae*, which moreover has no DNA methylation system. It remains to be determined whether XPG and other NER factors are generally important for transcription activation in mammalian cells or whether their action is limited to specific pathways such as NR-mediated transcription.

We showed that in addition to Pol II-transcribed genes, Rad2 is also located on Pol III-transcribed genes and telomeric regions in the absence of exogenous genotoxic stress, indicating a complex Rad2 distribution on the yeast genome and opening interesting perspectives of possible novel Rad2 functions. Recently, a role of Mediator on telomeres was proposed, since this complex influences telomeric silencing, cellular life span, and telomere heterochromatin maintenance (Zhu et al. 2011; Peng and Zhou 2012). In yeast, Mediator is bound to the telomeric regions. Our results revealed that Rad2 is also enriched in these regions. The colocalization of Mediator and Rad2 on telomeres raises the possibility that Mediator function on the telomeres could be related to Rad2.

In conclusion, our results suggest an intriguing model that Mediator might play more roles in nuclear processes

than previously assumed by serving as an assembly platform or a regulatory element linking transcription with DNA repair and possibly other chromatin-related processes.

Materials and methods

Strains and plasmids

All *S. cerevisiae* strains are described in Supplemental Table S1. All plasmids are listed in Supplemental Table S2. The oligonucleotides used in this study can be found in Supplemental Table S3.

ChIP, ChIP-chip, and ChIP-seq

ChIP and ChIP-chip experiments were performed as described (Ghavi-Helm et al. 2008). In ChIP experiments, we considered a significant enrichment compared with negative controls or a significant difference between ChIP occupancies when *P*-value was <0.05 in a Student's *t*-test. Chromatin preparation for ChIP-seq experiments was performed as described for conventional ChIP, except that an additional sonication step with Bioruptor (Diagenode; six cycles of 30 sec with medium intensity setting) was included to generate DNA fragments of ~200-bp mean size. DNA sequencing of 40-nucleotide (nt) tags was performed on Solexa genome analyzer GA-IIx using the procedures recommended by the manufacturer (Illumina). Input DNA and DNA from ChIP with an untagged strain were used as negative controls. The ChIP-chip and ChIP-seq data have been deposited to the Array Express under accession numbers E-MEXP-3875 and E-MTAB-1595.

Co-IP experiments

Whole-yeast extract preparation and immunoprecipitation in immunoprecipitation buffer (50 mM HEPES at pH 7.5, 100 mM NaCl, 20% glycerol, 1 mM dithiothreitol, 0.5 mM EDTA, 0.05% NP-40) supplemented with a protease inhibitor cocktail (Complete, Roche) and 1 mM PMSF and Western blotting were performed as described previously (Soutourina et al. 2006). The 12CA5 anti-HA and 9E10 anti-Myc antibodies were used against HA- or Myc-tagged proteins, and the 8WG16 anti-Rpb1 antibody was used against Pol II.

qRT-PCR analysis

RNA was extracted with hot acidic phenol following a protocol derived from Schmitt et al. (1990). Reverse transcription of 0.5-μg RNA samples was performed using iScript cDNA synthesis kit (Bio-Rad) with a mix of oligo(dT) and random hexamers for priming. qPCR results were normalized using 25S rRNA, 18S rRNA, and *SCR1* RNA as internal controls. Values represent the average of three independent experiments, and error bars indicate standard deviations.

Data analysis

ChIP-seq data were analyzed using the following bioinformatics tools. The sequences were aligned on *S. cerevisiae* genome (University of California at Santa Cruz [UCSC] version sacCer3) with Bowtie version 0.12.7 (Langmead et al. 2009). Conversions to different file formats were performed using Samtools version 0.1.16. Only uniquely mapped tags were used, and a maximum of two mismatches was allowed. To avoid possible sequencing

artifacts, reads beyond five repetitions at the same position of the genome were removed, and reads with the best-quality scores were conserved for each position. The number of mappable tags for each ChIP-seq experiment is indicated in Supplemental Table S4. To calculate ChIP-seq density, reads were extended to an assumed fragment length of 180 nt, and a count of reads was determined per 1-bp bin using Bedtools version 2.15.0. ChIP-seq density profiles were displayed using UCSC or IGB yeast genome browsers.

We used input DNA and DNA from ChIP with untagged strain as negative controls. The ChIP sample from the untagged strain generally shows very low tag density on the genome, except some regions representing <1% of the genome that display an apparent enrichment in this negative control. It should be noted that the majority of these regions were located inside the highly transcribed class II genes. To correct for this nonrandom background distribution, we subtracted the normalized signal of untagged strain sample from the ChIP samples for each protein. The ChIP signals of untagged strain have different impacts on the ChIP signals for different proteins. For this reason, the subtraction step was preceded by a normalization of the ChIP signal of an untagged strain compared with the corresponding ChIP sample based on the qPCR analyses of nontranscribed control regions.

The significantly enriched regions were identified using peak calling MACS2 software version 2.0.10.20120703 (Zhang et al. 2008) with a minimum false discovery rate (q -value) cutoff of 0.01 for peak detection followed by PeakSplitter step of PeakAnalyzer software for subdivision of ChIP-seq regions into discrete peaks (Salmon-Divon et al. 2010). Significantly enriched Rad2 peaks were annotated on the yeast genome using ChIPpeakAnno R script from the Bioconductor project (Zhu et al. 2010). To identify the Rad2 enrichment signals inside protein-coding genes, Rad2 mean densities were computed for each ORF and compared with mean densities of two negative controls: the untagged strain sample, normalized as described above, and the input DNA sample, normalized using NCIS R script version 1.0.1 (Liang and Keles 2012). We required that enriched ORFs have the mean density in Rad2 sample higher than that in the untagged strain control plus two standard deviations and, at the same time, in the input DNA control plus two standard deviations.

Acknowledgments

We thank N. van Berkum for two-hybrid screening with the Med17 Mediator subunit, S. Wydau for yeast strains, the SPI (CEA/Saclay) for monoclonal antibodies, A. Polesskaya for advice in human cell culture, and C. Mann and S. Un for critical reading of the manuscript. This work has benefited from the facilities and expertise of the high-throughput sequencing platform of IMAGIF (Centre de Recherche de Gif). This work was supported by the Agence Nationale de la Recherche (grant no. ANR_11_BSV8_020_01). F.E. was supported by a grant from the Fondation ARC (grant no. DOC20130606697).

References

Andrau JC, van de Pasch L, Lijnzaad P, Bijma T, Koerkamp MG, van de Peppel J, Werner M, Holstege FC. 2006. Genome-wide location of the coactivator mediator: Binding without activation and transient Cdk8 interaction on DNA. *Mol Cell* **22**: 179–192.

Cai G, Imasaki T, Yamada K, Cardelli F, Takagi Y, Asturias FJ. 2010. Mediator head module structure and functional interactions. *Nat Struct Mol Biol* **17**: 273–279.

Clarkson SG. 2003. The XPG story. *Biochimie* **85**: 1113–1121.

Compe E, Egly JM. 2012. TFIIF: When transcription met DNA repair. *Nat Rev Mol Cell Biol* **13**: 343–354.

Esnault C, Ghavi-Helm Y, Brun S, Soutourina J, Van Berkum N, Boschiero C, Holstege F, Werner M. 2008. Mediator-dependent recruitment of TFIIF modules in preinitiation complex. *Mol Cell* **31**: 337–346.

Firestein R, Bass AJ, Kim SY, Dunn IF, Silver SJ, Guney I, Freed E, Ligon AH, Vena N, Ogino S, et al. 2008. CDK8 is a colorectal cancer oncogene that regulates β -catenin activity. *Nature* **455**: 547–551.

Gade P, Singh AK, Roy SK, Reddy SP, Kalvakolanu DV. 2009. Down-regulation of the transcriptional mediator subunit Med1 contributes to the loss of expression of metastasis-associated dapk1 in human cancers and cancer cells. *Int J Cancer* **125**: 1566–1574.

Gaillard H, Tous C, Botet J, Gonzalez-Aguilera C, Quintero MJ, Viladevall L, Garcia-Rubio ML, Rodriguez-Gil A, Marin A, Arino J, et al. 2009. Genome-wide analysis of factors affecting transcription elongation and DNA repair: A new role for PAF and Ccr4–not in transcription-coupled repair. *PLoS Genet* **5**: e1000364.

Ghavi-Helm Y, Michaut M, Acker J, Aude JC, Thuriaux P, Werner M, Soutourina J. 2008. Genome-wide location analysis reveals a role of TFIIS in RNA polymerase III transcription. *Genes Dev* **22**: 1934–1947.

Guglielmi B, van Berkum NL, Klapholz B, Bijma T, Boube M, Boschiero C, Bourbon HM, Holstege FC, Werner M. 2004. A high resolution protein interaction map of the yeast Mediator complex. *Nucleic Acids Res* **32**: 5379–5391.

Hanawalt PC, Spivak G. 2008. Transcription-coupled DNA repair: Two decades of progress and surprises. *Nat Rev Mol Cell Biol* **9**: 958–970.

Hashimoto S, Boissel S, Zarhrate M, Rio M, Munnich A, Egly JM, Colleaux L. 2011. MED23 mutation links intellectual disability to dysregulation of immediate early gene expression. *Science* **333**: 1161–1163.

Holstege FC, Jennings EG, Wyrick JJ, Lee TI, Hengartner CJ, Green MR, Golub TR, Lander ES, Young RA. 1998. Dissecting the regulatory circuitry of a eukaryotic genome. *Cell* **95**: 717–728.

Ito S, Kuraoka I, Chymkowitz P, Compe E, Takedachi A, Ishigami C, Coin F, Egly JM, Tanaka K. 2007. XPG stabilizes TFIIF, allowing transactivation of nuclear receptors: Implications for Cockayne syndrome in XP-G/CS patients. *Mol Cell* **26**: 231–243.

Johnson KM, Wang J, Smallwood A, Arayata C, Carey M. 2002. TFIID and human mediator coactivator complexes assemble cooperatively on promoter DNA. *Genes Dev* **16**: 1852–1863.

Kagey MH, Newman JJ, Bilodeau S, Zhan Y, Orlando DA, van Berkum NL, Ebmeier CC, Goossens J, Rahl PB, Levine SS, et al. 2010. Mediator and cohesin connect gene expression and chromatin architecture. *Nature* **467**: 430–435.

Kaufmann R, Straussberg R, Mandel H, Fattal-Valevski A, Ben-Zeev B, Naamati A, Shaag A, Zenvirt S, Konen O, Mimouni-Bloch A, et al. 2010. Infantile cerebral and cerebellar atrophy is associated with a mutation in the MED17 subunit of the transcription preinitiation mediator complex. *Am J Hum Genet* **87**: 667–670.

Kornberg RD. 2005. Mediator and the mechanism of transcriptional activation. *Trends Biochem Sci* **30**: 235–239.

Kuuselo R, Savinainen K, Sandstrom S, Autio R, Kallioniemi A. 2011. MED29, a component of the mediator complex, possesses both oncogenic and tumor suppressive characteristics in pancreatic cancer. *Int J Cancer* **129**: 2553–2565.

Lagerwerf S, Vrouwe MG, Overmeer RM, Fousteri MI, Mullenders LH. 2011. DNA damage response and transcription. *DNA Repair (Amst)* **10**: 743–750.

Eyboullet et al.

- Langmead B, Trapnell C, Pop M, Salzberg SL. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* **10**: R25.
- Lariviere L, Geiger S, Hoepfner S, Rother S, Strasser K, Cramer P. 2006. Structure and TBP binding of the Mediator head subcomplex Med8-Med18-Med20. *Nat Struct Mol Biol* **13**: 895–901.
- Lariviere L, Plaschka C, Seizl M, Wenzek L, Kurth F, Cramer P. 2012. Structure of the Mediator head module. *Nature* **492**: 448–451.
- Lee SK, Yu SL, Prakash L, Prakash S. 2002. Requirement of yeast RAD2, a homolog of human XPG gene, for efficient RNA polymerase II transcription: Implications for Cockayne syndrome. *Cell* **109**: 823–834.
- Le May N, Mota-Fernandes D, Velez-Cruz R, Iltis I, Biard D, Egly JM. 2010. NER factors are recruited to active promoters and facilitate chromatin modification for transcription in the absence of exogenous genotoxic attack. *Mol Cell* **38**: 54–66.
- Le May N, Fradin D, Iltis I, Bougneres P, Egly JM. 2012. XPG and XPF endonucleases trigger chromatin looping and DNA demethylation for accurate expression of activated genes. *Mol Cell* **47**: 622–632.
- Li LH, He J, Hua D, Guo ZJ, Gao Q. 2010. Lentivirus-mediated inhibition of Med19 suppresses growth of breast cancer cells in vitro. *Cancer Chemother Pharmacol* **68**: 207–215.
- Liang K, Keles S. 2012. Normalization of ChIP-seq data with control. *BMC Bioinformatics* **13**: 199.
- Nonet M, Scafe C, Sexton J, Young R. 1987. Eucaryotic RNA polymerase conditional mutant that rapidly ceases mRNA synthesis. *Mol Cell Biol* **7**: 1602–1611.
- Peng J, Zhou JQ. 2012. The tail-module of yeast Mediator complex is required for telomere heterochromatin maintenance. *Nucleic Acids Res* **40**: 581–593.
- Ries D, Meisterernst M. 2011. Control of gene transcription by Mediator in chromatin. *Semin Cell Dev Biol* **22**: 735–740.
- Risheg H, Graham JM Jr, Clark RD, Rogers RC, Opitz JM, Moeschler JB, Peiffer AP, May M, Joseph SM, Jones JR, et al. 2007. A recurrent mutation in MED12 leading to R961W causes Opitz-Kaveggia syndrome. *Nat Genet* **39**: 451–453.
- Robinson PJ, Bushnell DA, Trnka MJ, Burlingame AL, Kornberg RD. 2012. Structure of the Mediator Head module bound to the carboxy-terminal domain of RNA polymerase II. *Proc Natl Acad Sci* **109**: 17931–17935.
- Salmon-Divon M, Dvinge H, Tammoja K, Bertone P. 2010. PeakAnalyzer: Genome-wide annotation of chromatin binding and modification loci. *BMC Bioinformatics* **11**: 415.
- Sarker AH, Tsutakawa SE, Kostek S, Ng C, Shin DS, Peris M, Campeau E, Tainer JA, Nogales E, Cooper PK. 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.
- Scharer OD. 2008. XPG: Its products and biological roles. *Adv Exp Med Biol* **637**: 83–92.
- Schmitt ME, Brown TA, Trumpower BL. 1990. A rapid and simple method for preparation of RNA from *Saccharomyces cerevisiae*. *Nucleic Acids Res* **18**: 3091–3092.
- Schwartz CE, Tarpey PS, Lubs HA, Verloes A, May MM, Risheg H, Friez MJ, Futreal PA, Edkins S, Teague J, et al. 2007. The original Lujan syndrome family has a novel missense mutation (p.N1007S) in the MED12 gene. *J Med Genet* **44**: 472–477.
- Soutourina J, Bordas-Le Floch V, Gendrel G, Flores A, Ducrot C, Dumay-Odelot H, Soularue P, Navarro F, Cairns BR, Lefebvre O, et al. 2006. Rsc4 connects the chromatin remodeler RSC to RNA polymerases. *Mol Cell Biol* **26**: 4920–4933.
- Soutourina J, Wydau S, Ambroise Y, Boschiero C, Werner M. 2011. Direct interaction of RNA polymerase II and mediator required for transcription in vivo. *Science* **331**: 1451–1454.
- Svejstrup JQ. 2002. Mechanisms of transcription-coupled DNA repair. *Nat Rev Mol Cell Biol* **3**: 21–29.
- Svejstrup JQ. 2007. Contending with transcriptional arrest during RNAPII transcript elongation. *Trends Biochem Sci* **32**: 165–171.
- Takahashi H, Parmely TJ, Sato S, Tomomori-Sato C, Banks CA, Kong SE, Szutorisz H, Swanson SK, Martin-Brown S, Washburn MP, et al. 2011. Human mediator subunit MED26 functions as a docking site for transcription elongation factors. *Cell* **146**: 92–104.
- Vijayvargia R, May MS, Fondell JD. 2007. A coregulatory role for the mediator complex in prostate cancer cell proliferation and gene expression. *Cancer Res* **67**: 4034–4041.
- Wang G, Balamotis MA, Stevens JL, Yamaguchi Y, Handa H, Berk AJ. 2005. Mediator requirement for both recruitment and postrecruitment steps in transcription initiation. *Mol Cell* **17**: 683–694.
- Werner M, Chaussivert N, Willis IM, Sentenac A. 1993. Interaction between a complex of RNA polymerase III subunits and the 70-kDa component of transcription factor IIIB. *J Biol Chem* **268**: 20721–20724.
- Zhang X, Krutchinsky A, Fukuda A, Chen W, Yamamura S, Chait BT, Roeder RG. 2005. MED1/TRAP220 exists predominantly in a TRAP/Mediator subpopulation enriched in RNA polymerase II and is required for ER-mediated transcription. *Mol Cell* **19**: 89–100.
- Zhang Y, Liu T, Meyer CA, Eeckhoutte J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W, et al. 2008. Model-based analysis of ChIP-Seq (MACS). *Genome Biol* **9**: R137.
- Zhu X, Wiren M, Sinha I, Rasmussen NN, Linder T, Holmberg S, Ekwall K, Gustafsson CM. 2006. Genome-wide occupancy profile of mediator and the Srb8-11 module reveals interactions with coding regions. *Mol Cell* **22**: 169–178.
- Zhu LJ, Gazin C, Lawson ND, Pages H, Lin SM, Lapointe DS, Green MR. 2010. ChIPpeakAnno: A Bioconductor package to annotate ChIP-seq and ChIP-chip data. *BMC Bioinformatics* **11**: 237.
- Zhu X, Liu B, Carlsten JO, Beve J, Nystrom T, Myers LC, Gustafsson CM. 2011. Mediator influences telomeric silencing and cellular life span. *Mol Cell Biol* **31**: 2413–2421.



Mediator links transcription and DNA repair by facilitating Rad2/XPG recruitment

Fanny Eyboulet, Camille Cibot, Thomas Eychenne, et al.

Genes Dev. 2013 27: 2549-2562

Access the most recent version at doi:[10.1101/gad.225813.113](https://doi.org/10.1101/gad.225813.113)

Supplemental Material

<http://genesdev.cshlp.org/content/suppl/2013/12/02/27.23.2549.DC1.html>

References

This article cites 51 articles, 14 of which can be accessed free at:
<http://genesdev.cshlp.org/content/27/23/2549.full.html#ref-list-1>

Creative Commons License

This article is distributed exclusively by Cold Spring Harbor Laboratory Press for the first six months after the full-issue publication date (see <http://genesdev.cshlp.org/site/misc/terms.xhtml>). After six months, it is available under a Creative Commons License (Attribution-NonCommercial 3.0 Unported), as described at <http://creativecommons.org/licenses/by-nc/3.0/>.

Email Alerting Service

Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](#).

To subscribe to *Genes & Development* go to:
<http://genesdev.cshlp.org/subscriptions>

II) Article 2 : Mediator independently orchestrates multiple steps of preinitiation complex assembly in vivo.

Mediator independently orchestrates multiple steps of preinitiation complex assembly *in vivo*

Fanny Eyboulet¹, Sandra Wydau-Dematteis¹, Thomas Eychenne¹, Olivier Alibert², Helen Neil¹, Claire Boschiero¹, Marie-Claire Nevers³, Hervé Volland³, David Cornu⁴, Virginie Redeker⁴, Michel Werner¹ and Julie Soutourina^{1,*}

¹Institute for Integrative Biology of the Cell (I2BC), Institut de Biologie et de Technologies de Saclay (iBiTec-S), CEA, CNRS, Université Paris Sud, F-91191 Gif-sur-Yvette cedex, France, ²CEA, iRCM, LEFG, Genopole G2, F-91057 Evry cedex, France, ³CEA, iBiTec-S, Service de Pharmacologie et d'Immunoanalyse, F-91191 Gif sur Yvette cedex, France and ⁴CNRS, Centre de Recherche de Gif, SICaPS, F-91198 Gif-sur-Yvette cedex, France

Received January 21, 2015; Revised July 17, 2015; Accepted July 21, 2015

ABSTRACT

Mediator is a large multiprotein complex conserved in all eukaryotes, which has a crucial coregulator function in transcription by RNA polymerase II (Pol II). However, the molecular mechanisms of its action *in vivo* remain to be understood. Med17 is an essential and central component of the Mediator head module. In this work, we utilised our large collection of conditional temperature-sensitive *med17* mutants to investigate Mediator's role in coordinating preinitiation complex (PIC) formation *in vivo* at the genome level after a transfer to a non-permissive temperature for 45 minutes. The effect of a yeast mutation proposed to be equivalent to the human Med17-L371P responsible for infantile cerebral atrophy was also analyzed. The ChIP-seq results demonstrate that *med17* mutations differentially affected the global presence of several PIC components including Mediator, TBP, TFIID modules and Pol II. Our data show that Mediator stabilizes TFIID kinase and TFIID core modules independently, suggesting that the recruitment or the stability of TFIID modules is regulated independently on yeast genome. We demonstrate that Mediator selectively contributes to TBP recruitment or stabilization to chromatin. This study provides an extensive genome-wide view of Mediator's role in PIC formation, suggesting that Mediator coordinates multiple steps of a PIC assembly pathway.

INTRODUCTION

In eukaryotes, the synthesis of mRNAs and a large number of small non-coding RNAs requires RNA polymerase II (Pol II) and the general transcription factors (GTFs) TFIIA, B, D, E, F and H that assemble into a large complex on the promoter DNA. This transcription preinitiation complex (PIC) is a key intermediate in Pol II transcription. *In-vitro* reconstitution studies of transcription initiation suggested a model in which PIC components assemble in a linear sequence (1,2). The first GTF that binds to the promoter is TFIID. TFIIA and TFIIB are then recruited followed by Pol II in association with TFIIF. Finally, TFIIE and TFIIH complete the formation of a PIC that is sufficient for basal *in-vitro* transcription but unable to drive activated transcription in response to specific activators. Important insights have been made on PIC architecture in yeast and human systems by biochemical and structural studies (3). A precise map of PIC locations across the yeast genome, including Pol II and GTFs, has been recently obtained, enabling the identification of TATA-like elements bound by TBP on TATA-less promoters and distinct PICs for divergent transcription (4,5).

Pol II transcriptional regulation requires additional multiprotein complexes, coactivators and corepressors, which modify the chromatin structure or directly contribute to PIC formation. Mediator of transcription regulation is one of these coregulators. While Mediator has been studied intensively, its complexity has precluded a detailed understanding of the molecular mechanisms of its action *in vivo*. Mediator is a huge complex (1.5 MDa) composed of 25–30 subunits depending on the organism that are organized into four structural modules: head, middle, tail and Cdk8 kinase modules. Recent electron microscopy studies redefined the modular organization of Mediator complex (6,7).

*To whom correspondence should be addressed. Tel: +33 1 69 08 54 13; Fax: +33 1 69 08 47 12; Email: julie.soutourina@cea.fr
Present addresses:

Sandra Wydau-Dematteis, EA 4065, CRP2, DHU Risks in pregnancy, Faculté de Pharmacie de Paris, Université Paris Descartes, Sorbonne Paris Cité, France.
Fanny Eyboulet, Institut de recherches cliniques de Montréal, 110 avenue des Pins Ouest, Montréal, Québec H2W 1R7, Canada.

Mediator subunits are engaged in numerous contacts within the complex (8–10). The large size, complexity and extreme flexibility have precluded the determination of a complete crystallographic structure of Mediator. Recently, structural data have been obtained for the yeast Mediator head module, the largest Mediator subcomplex characterized to date at high resolution (10,11).

Mediator transmits regulatory signals from specific transcription factors to the Pol II transcription machinery (12,13). This coregulator is generally required for Pol II transcription (14–18) and is present at most Pol II-transcribed gene promoters (19,20). Together with the GTFs, it promotes PIC formation and transcription initiation. Mediator also stimulates the phosphorylation of the Pol II carboxy-terminal domain (CTD) by TFIIF (21,22). Mediator cooperates with TFIID in yeast and human cells to stabilize the PIC (23–26). Previously, we have identified a direct contact between the Med11 Mediator head subunit and the Rad3 TFIIF subunit and showed that this contact is essential for the recruitment of the GTF to the PIC independently of Pol II (27). Recently, we showed that the Rpb3 Pol II subunit and the Med17 Mediator subunit are in direct contact to position Pol II on promoters and activate transcription, which strongly supports the notion that the Mediator–Pol II interaction is generally required for Pol II recruitment and transcription of class II genes in eukaryotes (28). Mediator was found to play an essential role in PIC assembly at the level of Pol II, TFIIF and TFIIE recruitment (27,28). Based on our results, we proposed a model of PIC assembly through multiple pathways (27).

Mediator is conserved in all eukaryotes (29) supporting a very ancient eukaryotic origin for this four-module complex. Given its central role in transcription regulation, it is not surprising that Mediator has been implicated in numerous developmental processes in animals, and mutations in Mediator subunits are involved in a number of pathologies (30,31). For example, a mutation in the Med23 subunit cosegregates with intellectual disability (32). In line with the fact that oncogenesis results from gene dysregulation, Mediator has been found to be involved in several cancers (33–38).

Med17 is one of the ten essential yeast Mediator subunits. *MED17* was initially isolated as *SRB4*, one of the *SRB* genes whose mutations suppressed the growth phenotype of truncations of the Pol II Rpb1 CTD (39). The general requirement of Mediator for Pol II transcription has been shown using the *med17-138* mutant that generally affects Pol II transcription in a manner comparable with that of the *rpb1-1* Pol II mutant (14). This classical *med17* temperature-sensitive (ts) allele causes dissociation of the head module from the rest of Mediator leading to a loss of Mediator function at the restrictive temperature (40–42). The central role of the Med17 subunit in Mediator's function is also consistent with a central and extended positioning of this subunit within the Mediator head structural model (10). The importance of the Med17 subunit's role has been highlighted by the fact that a mutation in this subunit has been associated with infantile cerebral atrophy (43) and because of this subunit's involvement in cancer (37).

We aimed to enhance our understanding of the key mechanisms that allow Mediator to stimulate PIC formation

in vivo at the genomic scale. We chose a strategy using temperature-sensitive mutants that allow the study of essential Mediator subunits providing specific changes in Mediator function without a complete loss or disassembly of Mediator. In this work, we have obtained a detailed genome-wide view of Mediator's role in PIC assembly *in vivo* by characterizing our large collection of conditional ts mutants in the Med17 Mediator head subunit. A mutant of yeast Med17 proposed to be equivalent to the human Med17-L371P responsible for infantile cerebral atrophy, which has a severe ts phenotype, was also included in our analysis. We show that *med17* mutations, which do not have a major effect on Mediator stability, differentially affect the genome-wide occupancy of PIC components. This work suggests that Mediator independently orchestrates multiple steps of the PIC assembly pathway.

MATERIALS AND METHODS

Strains and plasmids

All *S. cerevisiae* strains are described in Supplementary Table S1. All plasmids are listed in Supplementary Table S2. The oligonucleotides used in this study can be found in Supplementary Table S3.

ChIP and ChIP-seq

ChIP experiments were performed as described previously (44). Cell cultures (100 ml) were grown to exponential phase in YPD medium at 30°C, transferred for 45 minutes to a shaking incubator set at 37°C to allow gradual warming and cross-linked with 1% formaldehyde for 10 min. The 3HA-tagged proteins were immunoprecipitated with 12CA5 antibody and Pol II with 8WG16 anti-CTD antibody (Covance), bound to IgG magnetic beads (Dynabead). ChIP-seq experiments were performed as described previously (45). Chromatin preparation for ChIP-seq experiments was performed as described previously for conventional ChIP, except that an additional sonication step with Bioruptor (Diagenode, 6 cycles of 30s with medium intensity setting) was included to generate DNA fragments of ~200 bp mean size. DNA sequencing of 40 nt-tags was performed on a Solexa genome analyzer GA-IIx or Hi-Seq using the procedures recommended by the manufacturer (Illumina). Input DNA and DNA from ChIP with an untagged strain were used as negative controls. Mediator ChIP-seq analysis was performed only for the Med15 subunit. Low ChIP signals were obtained for the Med8-HA subunit. High background levels in the ChIP-seq experiments for Med5-HA and Med6-HA subunits precluded accurate enrichment detection. It should be noted that Mediator enrichment levels in ChIP experiments are relatively low, depending on the Mediator subunit used, possibly due to the low efficiency of formaldehyde crosslinking and/or the dynamic association of Mediator to chromatin.

The ChIP-seq data have been deposited to the Array Express under accession numbers E-MTAB-2961 for the mutant strains and E-MTAB-1595 for the wild-type (WT) strain.

Coimmunoprecipitation experiments

Whole yeast extracts were prepared from 100 ml of cells grown to exponential phase in YPD medium at 30°C. When indicated, the cells were then transferred for 45 min to 37°C. Whole yeast extract preparation, immunoprecipitation (IP) in IP buffer (50 mM HEPES, pH 7.5, 100 mM NaCl, 20% glycerol, 1 mM dithiothreitol, 0.5 mM EDTA, 0.05% NP-40) supplemented with a protease inhibitor cocktail (Complete, Roche) and 1 mM PMSF and western blotting were performed as described previously (46). The 12CA5 anti-HA antibodies were used against HA-tagged proteins; rabbit polyclonal anti-Med14 and anti-Med17 antibodies were used against corresponding N-terminal peptides to detect Med14 and Med17 Mediator subunits, respectively.

Formaldehyde cross-linking experiments were performed as described previously (28). Briefly, cells were cross-linked with 1% formaldehyde for 10 min, as described for ChIP. Lysis was performed in 0.5 ml FA/SDS buffer in the presence of glass beads (0.75 ml, 425 to 600 μ m) by vortexing for 45 min at 4°C. Crude extract preparation and IP were performed as described for conventional CoIP except that the immune complexes were washed as described for ChIP. A sonication step was added to recover chromatin-associated proteins. Denaturing conditions (8 M urea for 20 min at room temperature, cooled to 4°C and diluted with lysis buffer to a final urea concentration of 2 M) were used before IP to dissociate the subunits of Pol II and Mediator.

Mediator purification

Mediator purification was performed following a procedure adapted from Cai *et al.* (47) with some modifications. Med17 mutant and WT strains derived from CA001 protease-deficient background were grown at 30°C in 1 liter of 2 \times YPD medium overnight to late exponential phase. Cells were collected by centrifugation, washed twice in distilled water, resuspended in a buffer (1 ml/g of cells) containing 0.25 M Tris-HCl, pH7.6, 0.3 M ammonium sulfate, 2 mM EDTA, 30% glycerol, 5 mM 2-mercapthoethanol supplemented with 2 \times protease inhibitor cocktail (Complete, Roche) and frozen with liquid nitrogen. To prepare whole cell extract, cells were disrupted in an Eaton-Press. After thawing, the extract was clarified by ultracentrifugation (4°C, 1 h, 40 000g). The clear extract was submitted to a selective ammonium sulfate precipitation to isolate the 30–55% fraction. All buffers used for purification were supplemented with a protease inhibitor cocktail except for TEV digestion buffer. The precipitate was resuspended in IgG binding buffer (25 mM sodium phosphate buffer, pH7.5, 200 mM potassium acetate, 1 mM EDTA, 10% glycerol, 5 mM 2-mercapthoethanol) and incubated for 3 h with 100 μ l of IgG-coated beads (GE Healthcare) in poly-prep columns (BioRad) at 4°C on a rotator. Beads were drained by gravity and washed with 10 ml of IgG wash buffer I (50 mM sodium phosphate buffer, pH 7.5, 500 mM ammonium sulfate, 1 mM EDTA, 10% glycerol, 5 mM 2-mercapthoethanol), followed by a wash with 10 ml of IgG wash buffer II (50 mM sodium phosphate buffer, pH 7.5, 100 mM ammonium sulfate, 1 mM EDTA, 10% glycerol, 5 mM 2-mercapthoethanol) and 10 ml of TEV digestion buffer (50 mM sodium phosphate buffer, pH 7.5, 100 mM

potassium acetate, 1 mM EDTA, 10% glycerol, 5 mM 2-mercapthoethanol). Beads were incubated overnight with 20 U of TEV protease (Life Technologies) in 400 μ l of TEV digestion buffer. The eluate was recovered, diluted 6-fold in Ni-NTA binding buffer (50 mM sodium phosphate buffer, pH 7.5, 0.01% NP-40, 10% glycerol, 500 mM potassium acetate, 10 mM imidazole) and incubated for 3 h on a rotator at 4°C with 100 μ l of Ni-NTA sepharose beads (GE Healthcare) on poly-prep columns. Beads were drained by gravity and washed with 10 ml of Ni-NTA wash buffer (50 mM sodium phosphate buffer, pH 7.5, 0.01% NP-40, 10% glycerol, 500 mM potassium acetate, 40 mM imidazole). The Mediator complex was eluted with 200 μ l of Ni-NTA elution buffer (50 mM sodium phosphate buffer, pH 7.5, 0.01% NP-40, 10% glycerol, 500 mM potassium acetate, 200 mM imidazole), and flash-frozen after addition of 10% glycerol to a final glycerol concentration of 20%.

The presence of selected Mediator subunits (Med8-HA, Med14, Med17 and Med21) was verified by western blotting with anti-HA antibody (12CA5), polyclonal rabbit anti-Med14, anti-Med17 or anti-Med21 antibodies. Purified Mediator complex was then analyzed using a mass spectrometry approach.

Mass spectrometry analysis

To analyze Mediator integrity in *med17* mutants, Mediator was immunoprecipitated from crude extracts of Med5-HA tagged strains as described for CoIP experiments, except that protein G magnetic beads were used instead of IgG magnetic beads (Dynabead) and the elution was performed for 20 min at 65°C instead of 2 min at 85°C. The *MED17* strain carrying the non-tagged Mediator subunit was used as a negative control. Proteins from each eluate were separated by SDS-PAGE to fractionate the protein samples into three fractions and exclude the regions corresponding to heavy and light IgG chains. The gels were stained with Coomassie Blue and each fraction was cut into bands of ~2 mm and subjected to in-gel trypsin digestion with the Progest robot (Genomic Solutions, Chelmsford, MA) using standard conditions including reduction and alkylation. To analyse the purified Mediator complex, a short migration in NuPage 4–12% gel (Life Technologies) was performed before mass spectrometry analysis. After overnight tryptic digestion, peptides were extracted with 60% acetonitrile and 0.1% (v/v) formic acid. Tryptic peptides extracted from different bands of three gel regions were pooled, vacuum dried and resuspended in 0.1% (v/v) formic acid prior to nanoLC-MS/MS mass spectrometry analyses. The same cutting pattern of the SDS-PAGE lane was performed for each eluate.

The nanoLC-MS/MS analyses were performed with the Triple-TOF 4600 mass spectrometer (AB Sciex, Framingham, MA, USA) coupled to the nanoRSLC ultra performance liquid chromatography (UPLC) system (Thermo Scientific) equipped with a trap column (Acclaim PepMap100C18, 75 μ m i.d. \times 2 cm, 3 μ m) and an analytical column (Acclaim PepMapRSLCC18, 75 μ m i.d. \times 15 cm, 2 μ m, 100 Å). Peptide separation was performed with a 5–35% solvent B gradient for 40 min. Solvent A was 0.1% formic acid in water, and solvent B was 0.1%

formic acid in 100% acetonitrile. The nanoLC–MS/MS experiments were conducted using the data-dependent acquisition method by selecting the 20 most intense precursors for CID fragmentation with Q1 quadrupole set at low resolution for better sensitivity. Raw data were processed using MS Data Converter software (AB Sciex) for generating .mgf data files and protein identification was performed using the MASCOT search engine (Matrix science, London, UK) against the Swissprot database (release 2014.08) with carbamidomethylation of cysteines and oxidation of methionines set as fixed and variable modifications, respectively. Peptide and fragment tolerance were respectively set at 20 ppm and 0.05 Da. Results were analyzed using Scaffold 3.6.5 software (Proteome Software). Only proteins with at least two unique peptides and a 95% probability for both peptide and protein identifications were considered.

Quantitative real-time PCR (qRT-PCR) analysis

RNA was extracted with hot acidic phenol following a protocol derived from Schmitt *et al.* (48). Reverse transcription of 0.5 μ g RNA samples was performed using an iScript cDNA synthesis kit (BIO-RAD) with a mixture of oligo(dT) and random hexamers for priming. qRT-PCR results were normalized using 25S rRNA, 18S rRNA and *SCR1* Pol III-synthesized RNA as internal controls. Values represent the average of three independent experiments, and error bars indicate standard deviations.

Data analysis

ChIP-seq data were analyzed as described previously (45). The sequences were aligned on the *S. cerevisiae* genome (UCSC version sacCer3) using Bowtie version 0.12.7 (49). Only uniquely mapped tags were used and a maximum of two mismatches was allowed. Conversions to different file formats were performed using Samtools version 0.1.16. To avoid possible sequencing artifacts, reads with the best-quality scores were conserved for each position. The number of mappable tags for each ChIP-seq experiment is indicated in Supplementary Table S4.

To calculate ChIP-seq density, reads were extended to an assumed fragment length of 150–180 nt, and a count of reads was determined per 1-bp bin using Bedtools version 2.15.0. ChIP-seq density profiles were displayed using the Integrated Genome Browser (IGB) yeast genome browser. Input DNA and DNA from ChIP with an untagged strain were used as negative controls. The ChIP sample from the untagged strain generally shows very low tag density on the genome, except for some regions representing <1% of the genome that displays an apparent enrichment. It should be noted that the majority of these regions were located inside the highly transcribed class II genes. To correct for this non-random background distribution, we subtracted the normalized signal of the untagged strain sample from the ChIP samples for each protein. The ChIP signals of the untagged strain have different impacts on the ChIP signals for different proteins. For this reason, the subtraction step was preceded by a normalization of the ChIP signal of an untagged strain compared with the corresponding ChIP sample based on the qPCR analyses of nontranscribed control

regions. The normalization values for untagged strain sample before subtraction are indicated in Supplementary Table S5.

The significantly enriched regions were identified using peak calling MACS2 software version 2.0.10.20120703 (50). To avoid possible sequencing artifacts, reads beyond five repetitions at the same position of the genome were removed. A minimum false discovery rate (*q*-value) cutoff peak detection of 0.01 was used for Med15 and 0.05 for Pol II, Rad3, Kin28 and TBP. This step was followed using the PeakSplitter step of PeakAnalyzer software for the subdivision of ChIP-seq regions into discrete peaks (51). The significantly enriched regions were also identified in the control ChIP sample with an untagged strain using Bayespeak version 1.10.0 (52) with a threshold of 0.9. Around 200 peaks were identified by this method in the control sample and they were used to correct peakcalling data from Med15, Rad3, Kin28 and TBP ChIP-seq.

GTFs, Mediator and Pol II significantly enriched peaks were annotated relative to the intergenic region or the ORF in which they were detected using BedTools version 2.15.0. The nomenclature used for the intergenic regions was as follows: tandem regions were named relative to the downstream-located gene, divergent and convergent regions were named relative to the two genes surrounding the region.

To compare read counts in WT and *med17* mutant ChIP-seq data, a count of reads was determined on promoter regions of Pol II-transcribed genes for Rad3, Kin28, TBP and Med15, and on Pol II-transcribed gene ORFs for Pol II, using Bedtools. For ChIPseq data analysis, promoter regions were defined as corresponding intergenic regions on the yeast genome. For this step, all mapped reads were conserved to allow quantitative comparison between the different samples. Read numbers were normalized relative to qPCR data on a set of selected regions (Supplementary Table S6) using the ratio between WT and mutant strains. The ratio of ChIP-qPCR values (*V*) between WT and mutant strains was calculated for each selected region ($R_{\text{qPCR}} = V_{\text{qPCRmut}}/V_{\text{qPCRwt}}$). In parallel, the ratio of ChIP-seq reads (*N*) between WT and mutant strains was determined for the same region ($R_{\text{seq}} = N_{\text{seqmut}}/N_{\text{seqwt}}$). The normalization coefficient between ChIP-qPCR and read ratios was then calculated for each region ($Q = R_{\text{qPCR}}/R_{\text{seq}}$). The median of normalization coefficients between ChIP-qPCR and read ratios was used for ChIPseq data normalization (Supplementary Table S7).

RESULTS

Collection of conditional *med17* Mediator mutants

Previously, to investigate the *in-vivo* role of Med17 in Pol II recruitment, we selected a large collection of conditional *med17* mutants (28). The *med17* mutants with different thermosensitive phenotypes, from slow growth to lethal at non-permissive temperature (37°C), were generated by random mutagenesis (Supplementary Figure S1, Table S8). A number of multiple mutations were separated to obtain 5 additional point mutants by site-directed mutagenesis. An L371P mutation in the Med17 subunit has been

associated with infantile cerebral atrophy (43). A corresponding M504P mutation in the homologous *S. cerevisiae* Med17 (Supplementary Figure S2A) localized within one of the conserved domain based on multiple sequence alignments and secondary structure features (29) inactivated the protein at 37°C. To study the molecular consequences of this *med17* mutation, a *med17-504* yeast strain carrying a *med17-M504P* mutation was constructed and the growth phenotype of this mutant was investigated. Figure 1A shows that *med17-504* leads to a slow growth phenotype at 30°C and a strong thermosensitive phenotype at 37°C. In total, a collection of 30 conditional mutants in the Med17 subunit was obtained. The integrity of the main Mediator modules in all *med17* mutants has been verified by CoIP experiments (Supplementary Table S9). We decided to use this collection of Med17 Mediator mutants to investigate the *in-vivo* action of Mediator in general, and of the Med17 subunit in particular in PIC formation.

Effects of *med17* mutations on Mediator stability, interaction with Pol II and mRNA level

Four mutants were selected for further study based on their growth phenotypes (normal or slightly slower growth at 30°C combined with pronounced ts phenotype at 37°C), Mediator integrity and low mutation number: *med17-98* (L86Q, E186G, E448Q), *med17-444* (Q444P), *med17-504* (M504P) and *med17-670* (V670E) (Figure 1A). These four *med17* mutants all have mutations localized within or near the conserved domains of this subunit defined by a primary sequence comparison and predicted secondary structure (Supplementary Figure S2B) (29). The mutated amino acid residues of *med17* mutants were mapped on a partial structural model of the Mediator head module if the corresponding residues were present on the model (10) (Supplementary Figure S2C) and were found to often be located close to the module surface.

Since Mediator mutations can lead to dissociation of Mediator modules and a loss of Mediator function, the effect of *med17* mutations on the integrity of the main Mediator modules was first verified by CoIP experiments (Supplementary Figure S3A and Table S9). The HA-tagged Med5 Mediator subunit was immunoprecipitated with an anti-HA antibody. The coimmunoprecipitated Med17 Mediator head subunit and Med14 Mediator subunit bridging tail and middle modules were detected by western blotting with antibodies produced in house. Mediator head, middle and tail modules coimmunoprecipitated with the Med5 tail module subunit in all four mutants without any significant difference compared to the WT strain. We also performed CoIP experiments with whole yeast extracts prepared from cells grown to exponential phase in YPD medium at 30°C and then transferred for 45 min to 37°C (Supplementary Figure S3A). Similar to cells grown at 30°C, no significant differences were observed in *med17* mutants compared to the WT strain.

To investigate Mediator integrity in the selected *med17* mutants in more detail, Mediator was immunoprecipitated from crude extracts of Med5-HA tagged strains and immunoprecipitated proteins were analyzed using a mass spectrometry approach (Figure 1B, immunoprecipitation). To

improve Mediator subunit identification using mass spectrometry, the immunoprecipitated proteins were separated on gel and the bands corresponding to IgG heavy and light chains were cut out. This procedure excluded some Mediator subunits with molecular weights close to those of the IgG chains. A strain carrying an untagged Med5 subunit was used as a negative IP control. Up to 19 Mediator subunits were identified belonging to all four Mediator modules. Thirteen Mediator subunits of head, middle, tail and Cdk8 modules (Med12, 13, 17, 6, 20, 1, 4, 7, 5, 14, 15, 16, 3) were present for all *med17* mutants and the WT strain. Mediator subunits with low molecular weight (Med9, 10, 11, 21, 22) were not detected in some samples. Mass spectrometry analysis thus showed the presence of Mediator head, middle, tail and Cdk8 modules for all *med17* mutants, indicating that mutations did not have a major effect on Mediator stability.

To improve our analysis of Mediator integrity in *med17* mutants, we purified Mediator complex by a two-step affinity purification method that allows to prepare core Mediator containing head, middle and tail modules but lacking Cdk8 module (47). Purified Mediator complex from the four *med17* mutants and a WT strain was assayed by SDS-PAGE (Figure 1C) and analyzed using a mass spectrometry approach. Yeast core Mediator is composed of 21 subunits since it lacks the 4 Cdk8 module subunits. The Med31 Mediator subunit escaped mass spectrometric identification (8), leading to 20 core Mediator subunits that could be detected by this approach. Our mass spectrometry analysis showed that all 20 Mediator subunits were present in Mediator complex purified from *med17* mutants and a WT strain (Figure 1B, Mediator purification). Even though no peptides were initially detected for Med21 subunit (a small Mediator subunit of 16 kDa) in *med17-98* mutant, the presence of this subunit in *med17-98* Mediator complex was confirmed by a second data acquisition identifying 2 unique peptides and our western blotting analysis with anti-Med21 antibody (Supplementary Figure S3B). Taken together, our data strongly suggest that *med17-98*, 444, 504 and 670 mutations did not compromise Mediator integrity.

The Med17 Mediator subunit directly interacts *in vivo* with the Rpb3 Pol II subunit (28). As described previously (28), CoIP experiments for Mediator–Pol II interaction are not appropriate to investigate the effect of mutations since their effects depend on the buffer conditions. Hence the need to determine how Mediator Med17 mutations affect the specific Med17–Rpb3 contact using a formaldehyde-crosslinking approach with IP under denaturing conditions. This method was used in our previous work to show an impaired Med17–Rpb3 contact in an *rpb3-2* mutant (28). We examined whether this contact was modified in *med17* mutants. For these experiments, yeast strains were constructed carrying an EGFP-tagged Med17 subunit and an Myc-tagged Rpb3 subunit in WT and *med17* mutant contexts. Proteins were cross-linked with formaldehyde, the Med17-EGFP-tagged subunit was immunoprecipitated under denaturing conditions and the cross-linked Rpb3-Med17 band was detected with an anti-Myc antibody (Figure 1D, red arrow). The results demonstrated that Rpb3–Med17 contact was strongly reduced in the *med17-504* mutant (Figure 1D), but remained largely unchanged

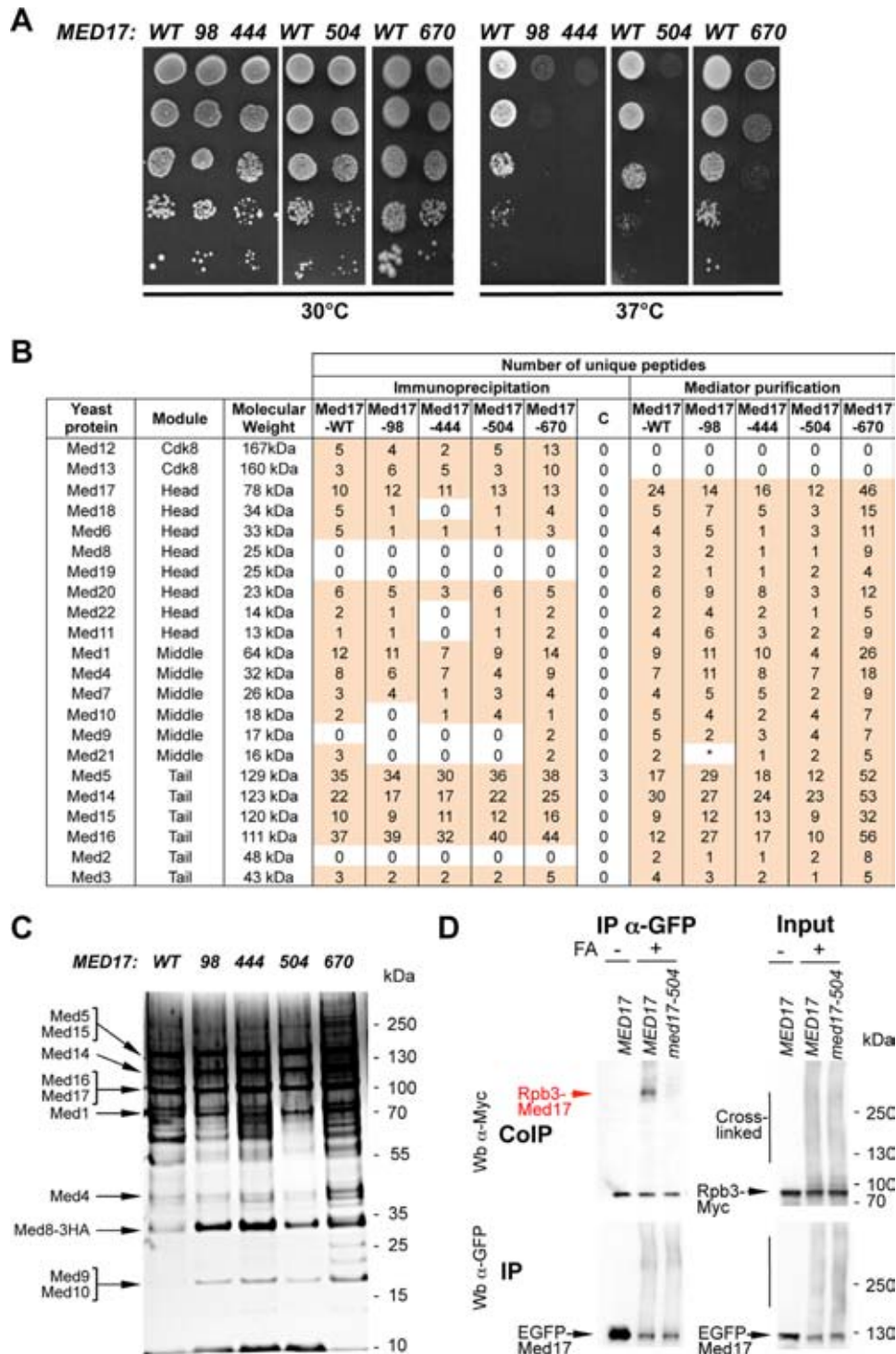


Figure 1. *med17* thermosensitive mutants. (A) Med17 mutant phenotypes. Cultures of WT and mutant *med17* yeast strains were serially diluted, spotted on YPD agar plates and incubated for 3 days at permissive (30°C) or non-permissive (37°C) temperatures. (B) Mass spectrometry analysis of Mediator integrity in *med17* mutants. Mediator subunits identified by mass spectrometry analysis in WT and *med17* mutants are indicated. (Immunoprecipitation) Mediator was immunoprecipitated through Med5-HA from crude extracts using magnetic protein G beads coupled to anti-HA antibodies. The *MED17* strain carrying a non-tagged Mediator subunit was used as a negative control (C for Control IP). (Mediator purification) Core Mediator complex containing head, middle and tail modules was purified from *med17* mutants and a wild-type strain. An asterisk marks Med21 subunit in *med17-98* mutant that was detected in a second data acquisition identifying two unique peptides and also by western blotting with anti-Med21 antibody (Supplementary Figure S3B). (C) Silver-stain SDS-PAGE analysis of purified Mediator complex from the WT strain and *med17* mutants. (D) Interaction between Pol II and Mediator in *med17-504* mutant. Rpb3-Myc Med17-EGFP strains with WT *MED17* or a *med17* mutation were grown at 30°C in YPD medium and cross-linked or not with formaldehyde (FA), as indicated. Med17-EGFP was immunoprecipitated (IP) with anti-EGFP antibody from crude extracts (Input) and analyzed by western blotting with anti-Myc antibody (CoIP) against Rpb3. The cross-linked Rpb3-Med17 band is indicated in red. The position of unidentified cross-linked proteins with the tagged Med17 or Rpb3 subunits is indicated by a vertical bar.

in *med17-98*, *444* and *670* compared to the wild type (Supplementary Figure S4).

The effect of *med17* mutations after a transfer to 37°C on mRNA abundance of *ADHI*, *PYK1* and *PMA1* genes was investigated (Supplementary Figure S5). The *med17-98*, *444* and *670* mutants were found to affect the level of these mRNAs. The effect of *med17-98* was accentuated at 37°C and that of other mutations was essentially similar at 30°C and 37°C. It should be noted that *med17-504* shows little, if any, effect on the steady-state levels of the three mRNAs. As the measurement of mRNA levels is a cumulative result of RNA synthesis and RNA turnover, mRNA levels do not fully reflect the effects of *med17* mutations on transcription. Moreover, it has been shown recently that cells can compensate for global transcription changes by RNA degradation changes to buffer cellular mRNA levels (53,54). We measured Pol II occupancy on transcribed regions by ChIP since it indicates transcriptional activity in yeast (55,56).

Mutations in Med17 differentially affect chromatin association of PIC components on selected constitutively expressed genes

To analyse the molecular effect of *med17* mutations on PIC formation, occupancy of Pol II, GTFs and Mediator was determined by ChIP on selected constitutively expressed genes. Yeast strains grown in YPD-rich medium and transferred to a non-permissive temperature (37°C) for 45 min were used for chromatin preparation. The association of Mediator was measured by ChIP experiments with three different Mediator subunits: Med6-HA (head), Med15-HA (tail) and Med5-HA (tail) (Figure 2A–C). The Mediator occupancy of the *ADHI*, *PYK1* and *PMA1* promoter regions remained largely unchanged except for Med15 Mediator occupancy that was reduced in the *med17-444* mutant and to some extent for other mutants on the *PMA1* promoter region. In contrast, all examined mutations led to a decrease in Pol II association with the promoter and transcribed regions of class II genes compared to the WT (Figure 2D), with *med17-444* and *med17-504* showing the most pronounced effect, reflecting the effects on Pol II transcription.

With the Mediator head module directly interacting with TFIID and TBP, the ability of *med17* mutations to influence the occupancy of these GTFs was assessed *in vivo*. Two TFIID modules, TFIID core (Rad3-HA subunit) and TFIID kinase module (Kin28-HA) were ChIPed (Figure 2E and F). Interestingly, mutations in the Med17 Mediator subunit were found to differentially affect the presence of the two TFIID modules on the *ADHI*, *PYK1* and *PMA1* promoters. In the *med17-98* and *med17-504* mutants, Med17 mutations led to a decrease in the occupancy of both TFIID modules. The *med17-444* mutation specifically affected Kin28 occupancy. More unexpectedly, a decrease in Rad3 occupancy without any effect on Kin28 occupancy was observed in *med17-670*, suggesting that TFIID modules can behave independently *in vivo*.

TBP occupancy in *med17* mutants was also analyzed (Figure 2G). As previously observed for *med11* Mediator mutants, TBP occupancy was similar to that of the WT strain in *med17-98* and *670* mutants. In contrast, TBP as-

sociation to the *ADHI*, *PYK1* and *PMA1* gene promoters was reduced in *med17-444* and *med17-504* indicating that Mediator could affect the recruitment and/or stability of this GTF *in vivo*.

Effects of *med17* mutations on chromatin occupancy of PIC components on the inducible *GALI* gene

To address the recruitment of PIC components to an inducible yeast gene, ChIP experiments were performed with Mediator, Pol II and GTFs in *med17* mutants on the *GALI* gene. The association of Mediator (Med5 subunit), Pol II, TFIID core (Rad3 subunit), TFIID kinase module (Kin28 subunit) and TBP were analyzed by ChIP at 45 min of galactose induction (data not shown). The association of PIC components was generally affected in the same manner as for the three constitutively expressed genes, with *med17-504* being an exception with reduced Mediator recruitment and unmodified Rad3 and TBP recruitment to the *GALI* promoter region at 45 min of galactose induction.

Two *med17* mutants were chosen for further experiments to address the kinetics of the *GALI* gene induction: *med17-444*, since it is the only mutant that affected TBP recruitment to the *GALI* promoter and *med17-504*, since its effects on the *GALI* gene differ from those on the *ADHI*, *PYK1* and *PMA1* genes for several PIC components. Yeast cells were grown in YP raffinose medium at 30°C, then galactose was added at the 0 time point and cultures were transferred to 37°C. Samples were collected at 0, 20, 40 and 60 min of galactose induction and transfer to 37°C. The kinetics of Mediator, Pol II and TBP recruitment upon galactose induction were analyzed in *med17-444* and *-504* mutants compared to the wild type (Figure 3). Mediator recruitment was decreased in *med17-504* for all kinetics points, showing a particular Mediator behavior on the *GALI* gene in this mutant (Figure 3A). In the *med17-444* mutant, Mediator occupancy of the *GALI* gene promoter reached a level similar to that in the wild type, even though the kinetics of Mediator recruitment were slower (20-min point). Pol II occupancy generally followed that of Mediator (Figure 3B). The *med17-444* mutation was found to slow TBP recruitment during galactose induction and *med17-504* showed TBP recruitment kinetics similar to that of the WT strain (Figure 3C).

Taken together, the ChIP results demonstrate that mutations in the Med17 Mediator subunit could differentially affect the recruitment of PIC components and also highlight some gene-specific effects suggesting that regulatory mechanisms might differ depending on particular Pol II-transcribed genes.

Genome-wide location analysis of Mediator, Pol II, TBP and TFIID modules

To extend the ChIP analysis of PIC assembly to the whole yeast genome, ChIP-seq experiments were performed for Mediator (Med15 subunit), Pol II and GTFs (TBP and two TFIID subunits: Kin28 and Rad3). Input DNA and DNA from ChIP with an untagged strain were used as negative controls. Figure 4A shows examples of ChIP-seq density distributions using the IGB for all proteins on selected

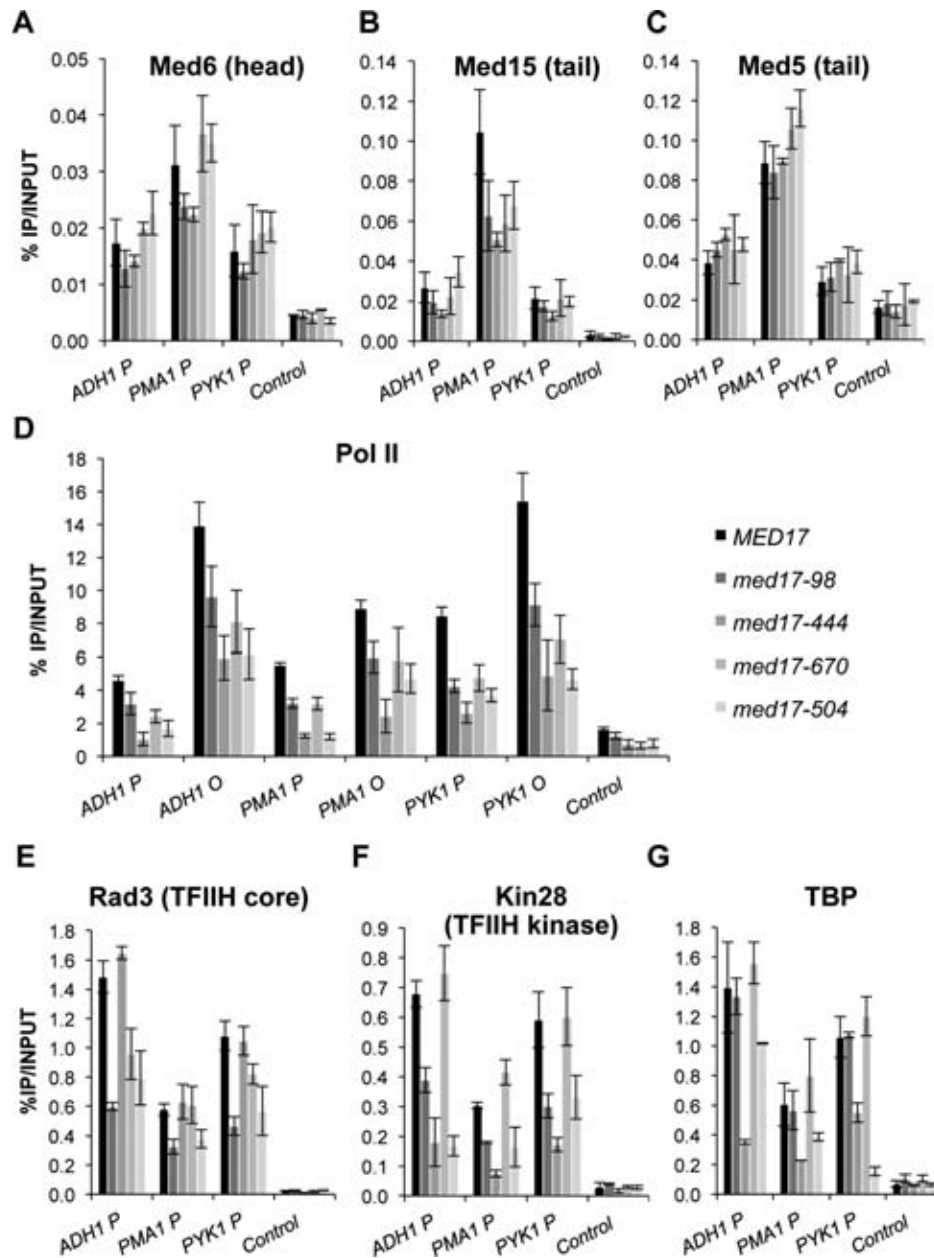


Figure 2. Effect of *med17* mutations on Mediator, Pol II, TFIIH and TBP occupancy. Cells were grown to exponential phase at 30°C on YPD medium and then transferred for 45 min to 37°C. Quantitative ChIP experiments were performed using anti-HA antibody (12CA5) against Mediator subunits Med6-HA (A), Med15-HA (B) or Med5-HA (C), Rad3-HA (E), Kin28-HA (F), TBP-HA (G) or anti-Rpb1 Pol II antibody (8WG16) (D). Immunoprecipitated DNA was amplified with primers corresponding to *ADH1*, *PMA1* and *PYK1* ORF (O) or promoters (P). P1 primers listed in Supplementary Table S3 located close to upstream regulatory regions were used for Mediator ChIP experiments and P2 primers located close to core promoters were used for Pol II, Rad3, Kin28 and TBP ChIP experiments. *GAL1* ORF was used as a negative control. The mean values and standard deviation (indicated by error bars) of three independent experiments are shown.

class II genes in WT strains. To correct for chromatin-state bias corresponding to the highly expressed genomic loci observed for control samples (15,45,57–59), the normalized signals of an untagged strain sample were subtracted from the ChIP samples for each protein as described previously (45) and in the ‘Materials and Methods’ section. The ChIP signals of Mediator inside ORFs most likely resulted from chromatin-state bias and are mostly corrected by untagged strain signal subtraction.

The location of Mediator, Pol II and GTF peaks were compared around transcription start sites (TSS) in WT strains in a metagenome analysis (Figure 4B). To consider only Pol II transcribed gene promoters, intergenic regions encompassing Pol III-transcribed genes were excluded. Divergent genes with double peaks for GTFs were also excluded from the analysis. Mean tag density (1 bp bin) was determined for each protein in a 1600-bp window centred on the TSS. These analyses show that the ChIP-seq distributions of

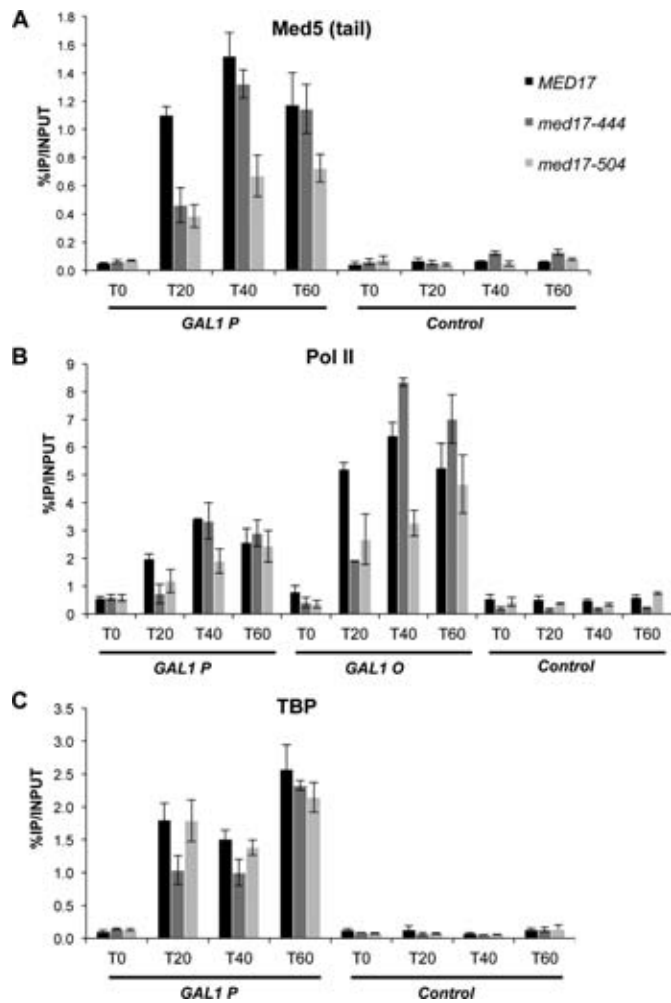


Figure 3. Kinetics of galactose induction in *med17-444*, *-504* and WT strains. Yeast strains were grown in raffinose-supplemented medium at 30°C, then galactose was added, cultures were transferred at the same time to 37°C and samples were collected for ChIP experiments at indicated time points upon galactose induction (T0, T20, T40 and T60 min). Quantitative ChIP experiments were performed using anti-HA antibody (12CA5) against Mediator subunit Med5-HA (A), TBP-HA (C) or anti-Rpb1 Pol II antibody (8WG16) (B). The mean values and standard deviation (indicated by error bars) of three independent experiments are shown. Immunoprecipitated fragments from ChIP experiments were amplified with primers corresponding to the *GAL1 P* gene promoter (*GAL1 P*) or ORF (*GAL1 O*). A nontranscribed region on chromosome V was used as a negative control.

TFIIH subunits and TBP superposed with each other close to TSS. As expected, Pol II is distributed inside transcribed regions. Mediator distribution was located upstream of TSS on regulatory regions.

med17 mutations can globally destabilize the PIC *in vivo*

To investigate whether PIC assembly is generally affected in *med17* mutants or in a gene-specific manner, the genome-wide occupancy of Mediator, Pol II and GTFs were analyzed using ChIP-seq experiments in mutants to compare them with the WT strains. Cells were grown at 30°C in YPD medium and then transferred for 45 min to 37°C. The normalization step is essential for ChIP-seq comparative analysis, but widely used normalization methods like a quantile

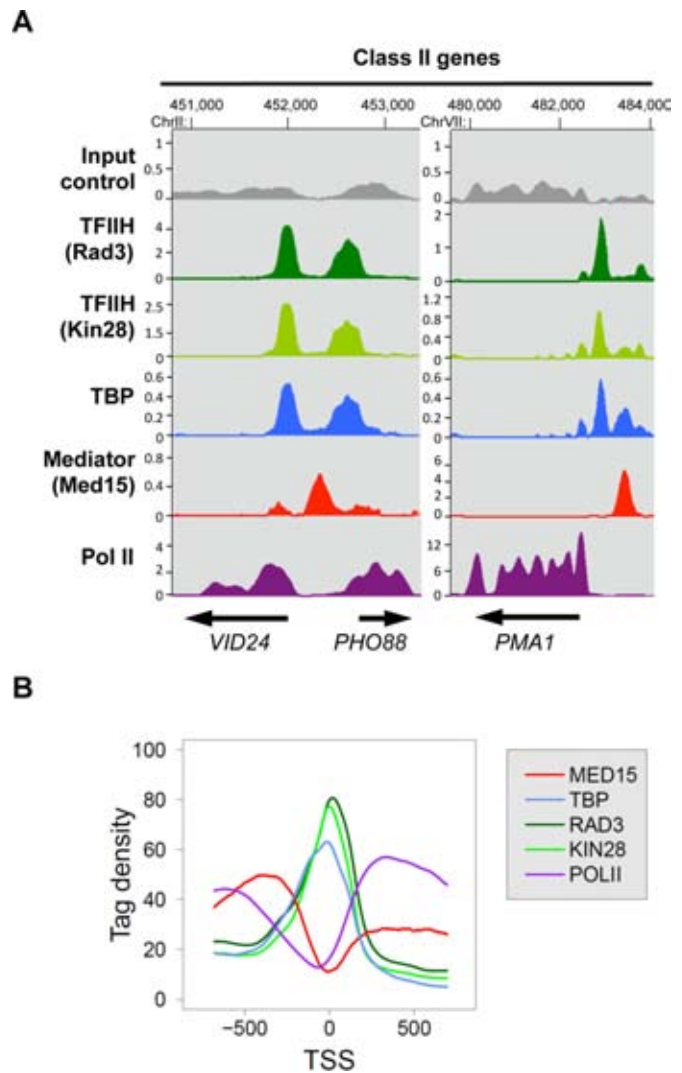


Figure 4. Enrichment profiles of Mediator, Pol II, TFIIH and TBP on yeast genome. Cells were grown at 30°C in YPD medium and then transferred for 45 minutes to 37°C. (A) Examples of Mediator, Pol II, TBP, and TFIIH ChIP-seq enrichment profiles on selected class II genes. Densities of sequence tags were assessed from ChIP-seq experiments performed with Med15-HA (Mediator), TBP-HA, Rad3-HA (TFIIH core), and Kin28-HA (TFIIK) strains using anti-HA antibody. Pol II was immunoprecipitated using anti-Rpb1 antibody. ChIP-seq density profiles are displayed using the IGB yeast genome browser. Input DNA and DNA from ChIP with an untagged strain were used as negative controls. Densities of sequence tags were displayed after subtraction of the normalized control of an untagged strain. The scales in kiloreads are indicated for each profile. (B) Distribution of Mediator, TBP, TFIIH and Pol II ChIP-seq densities around TSS. Intergenic regions encompassing Pol III-transcribed genes and divergent genes were excluded. The tag density was determined for each protein in a 1600-bp window centred on the TSS. The TSS positions were provided from previous studies (4,5). Mean tag density for each nucleotide position was then calculated and plotted over the window.

normalization step cannot be applied in situations where occupancy increases or decreases genome-wide. To detect the potential global effects of Mediator mutations on PIC component occupancy, we performed the following normalization step. A count of reads was determined on the promoter regions of Pol II-transcribed genes for Med15, Rad3, Kin28 and TBP, and on Pol II-transcribed gene ORFs for

Pol II and then normalized relative to qPCR data on a set of selected regions, as described in the 'Materials and Methods' section. Regression analysis of PIC component binding in each *med17* mutant versus WT was systematically performed.

Genome-wide Mediator occupancy was analyzed for the Med15 Mediator subunit. It was not possible to obtain good quality ChIP-seq results for the Med5 Mediator tail module subunit or the Med6 and Med8 Mediator head module subunits (see 'Materials and Methods' section). Chromatin associated with the Med17 Mediator head subunit can be ChIPed efficiently. However, the addition of tags to mutant subunits has been observed to potentially lead to an aggravation of growth phenotypes. For this reason, Med17 occupancy was not investigated in mutant strains.

The ChIP results on the selected class II genes showed the most pronounced effects on Pol II, TBP and Kin28 occupancy for the *med17-444* mutant whereas Rad3 occupancy remained unaffected in this mutant (Figures 2 and 3). Genome-wide analysis of the *med17-444* mutant revealed a global 1.8-fold decrease (the slope of regression line, shown in red, is equal to 0.55) in Med15 Mediator occupancy with a high correlation coefficient (R^2 equal to 0.93) (Figure 5A). As a consequence, the presence of other PIC components like Pol II, TFIID and TBP was also globally reduced to a different extent. A global decrease in genome-wide occupancy for Pol II and the Kin28 TFIID subunit (2.6- and 2.7-fold, respectively) was observed. The TBP association to chromatin was also globally affected with a 1.6-fold decrease. The Rad3 TFIID core subunit binding was slightly affected (1.2-fold decrease), consistent with the ChIP results on the *ADHI*, *PYK1* and *PMA1* gene promoters. Taken together, the genome-wide results on the *med17-444* mutant confirmed a central role for the Med17 Mediator subunit in PIC formation and suggested that the Mediator tail module destabilization on the chromatin could lead to a global change in the *in-vivo* binding of several PIC components but only marginally affected Rad3 occupancy.

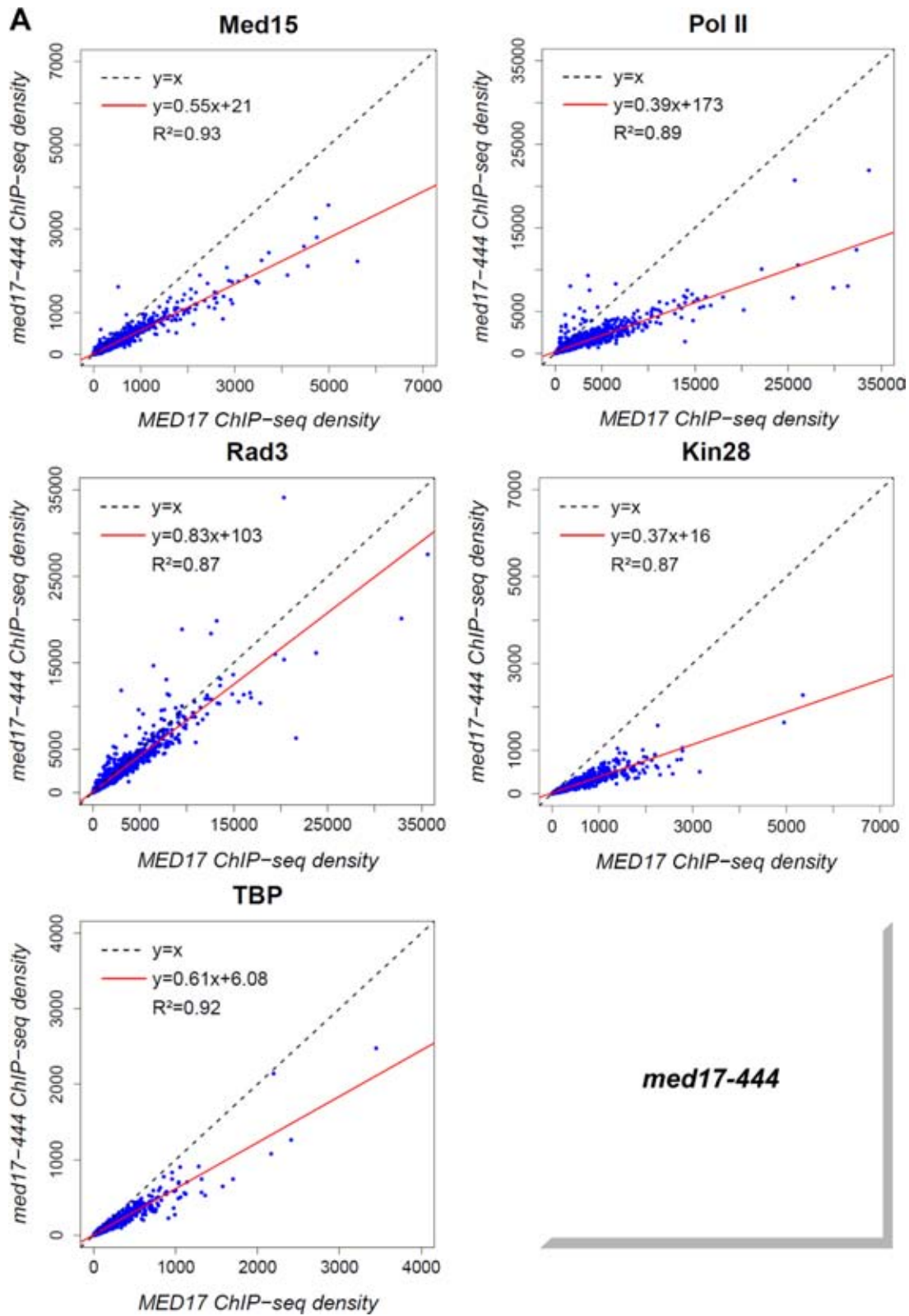
The *med17-504* mutation was found to impair Pol II-Mediator interaction (Figure 1D) and led to reduced binding of Pol II and GTFs to the *ADHI*, *PYK1* and *PMA1* genes (Figure 2) and Mediator, Pol II and some GTFs on the *GALI* gene (Figure 3 for Mediator, Pol II and TBP, and data not shown for TFIID). To determine how general the effects of *med17-504* were, the genome-wide Mediator, Pol II, Rad3, Kin28 and TBP occupancy were compared in this mutant and the WT (Figure 5B). The genome-wide binding of the Mediator Med15 subunit in *med17-504* was similar to that of the WT strain (the slope equal to 0.93 and R^2 equal to 0.98). In contrast, the association of Pol II and the GTFs was globally diminished in this mutant compared to the WT, suggesting that the impaired Pol II-Mediator interaction leads to a general PIC assembly defect. The genome-wide analysis of *med17-444* and *-504* mutants demonstrated that a destabilization of Mediator binding to chromatin or an impaired interaction between Mediator and Pol II globally destabilize the PIC assembly *in vivo*.

Mutations in the Med17 Mediator subunit specifically affect the genome-wide association of the TFIID core subcomplex

Previously, we have shown that mutations in the Med11 Mediator subunit affecting its interaction with the Rad3 TFIID core subunit can selectively impair the binding of the TFIID kinase module (27). Unexpectedly, the ChIP experiments suggested that some *med17* mutations like *med17-670* could lead to a selective dissociation of the Rad3 TFIID core subunit. TFIID core genome-wide binding in *med17* mutants was therefore investigated. The ChIP-seq results revealed that the genome-wide association of this TFIID subcomplex was affected independently of the TFIID subcomplex in *med17-670* (Figure 6A) and *-98* mutants (Figure 6B). A global 2-fold decrease of the Rad3 TFIID core subunit binding was observed in these mutants with the slope of regression line equal to 0.55 and 0.51 for *med17-670* and *-98*, respectively, with high correlation coefficients. In contrast, the association of the Kin28 TFIID subunit was similar to the WT level (the slope equal to 1.01 and 0.94, R^2 equal to 0.98 and 0.90 for *med17-670* and *-98*, respectively). The presence of the Med15 Mediator subunit was mostly unchanged or only slightly decreased (the slope equal to 0.87 and 0.80, R^2 equal to 0.93 and 0.94 for *med17-670* and *-98*, respectively). The genome-wide binding of TBP was not reduced and even showed a slight increase (the slope equal to 1.25 and 1.14, R^2 equal to 0.92 and 0.96 for *med17-670* and *-98*, respectively). In the *med17-670* mutant, a global 1.4-fold decrease in Pol II occupancy was observed when compared to the WT strain with the slope of regression line equal to 0.72 and a high correlation coefficient equal to 0.97. The Pol II association was also reduced in the *med17-98* mutant showing a 1.2-fold global decrease with a reduced correlation coefficient of 0.83 that suggested a more complex transcriptional effect for this mutation. The genome-wide analysis of PIC components of *med17-670* and *-98* mutants suggests that Mediator can selectively affect the TFIID core subcomplex binding to the yeast genome without impairing that of the TFIID kinase module or TBP.

DISCUSSION

In this study, a genome-wide analysis of PIC assembly *in vivo* was performed by taking advantage of conditional Mediator mutants of the essential Med17 head module subunit after a transfer to a non-permissive temperature. We propose that the Mediator head module in general, and Med17 in particular, independently coordinate the PIC assembly *in vivo* at different levels. (i) In accordance with previous studies, the results of this work support a central role for the Med17 Mediator subunit in Mediator organization and function. We demonstrate that *med17* mutants globally affect Pol II transcription and specific mutations in Med17 could result in a global destabilization of PIC assembly *in vivo*. (ii) The ChIP and ChIP-seq results show that Mediator selectively contributes to TBP recruitment or stabilization to chromatin *in vivo*, supporting Mediator head-TBP functional contact that was previously identified *in vitro*. (iii) We show that Mediator recruitment stabilizes TFIID kinase and TFIID core modules independently since *med17* mutations can differentially affect the genome-wide occupancy of the two TFIID modules. We therefore propose that



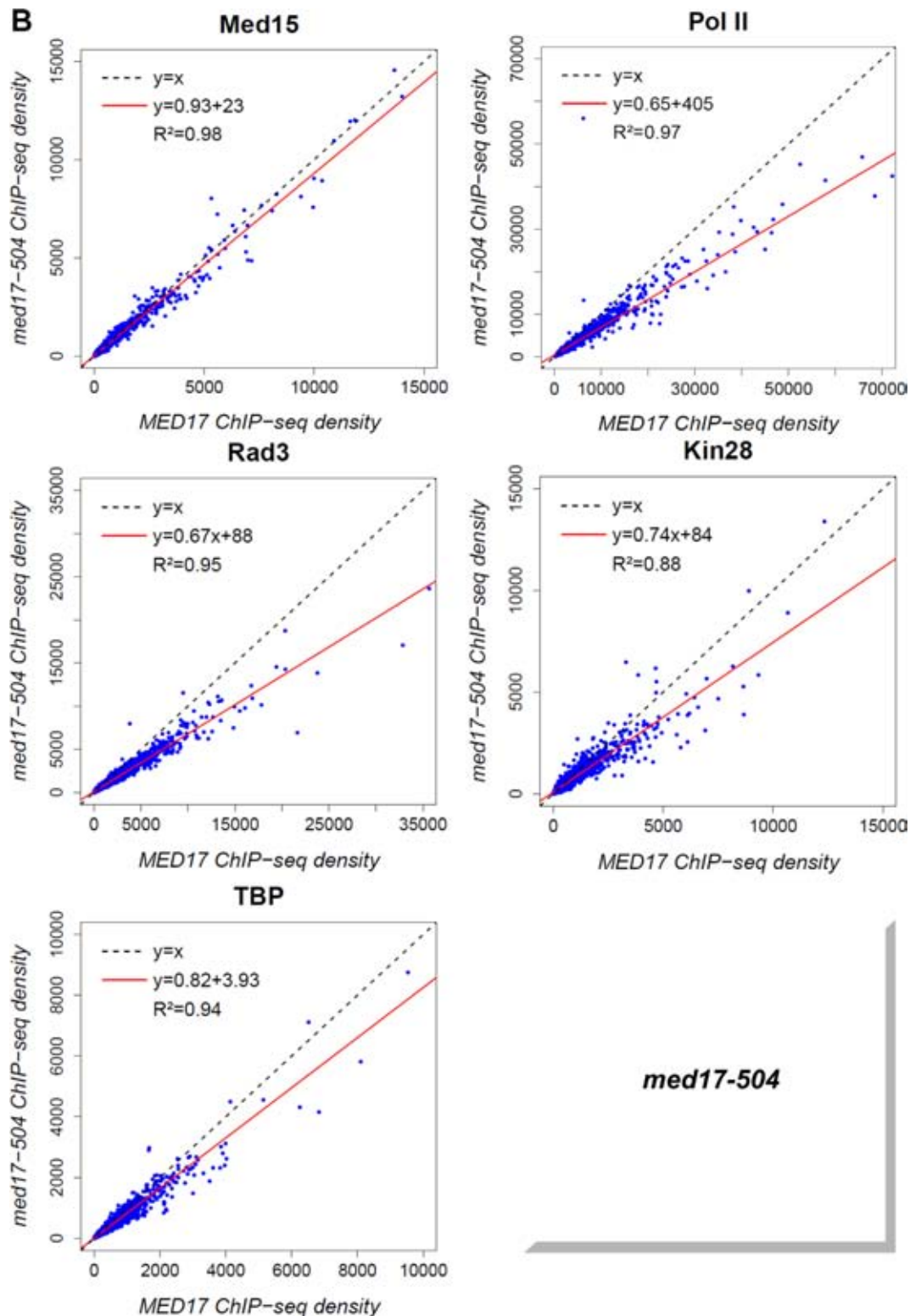
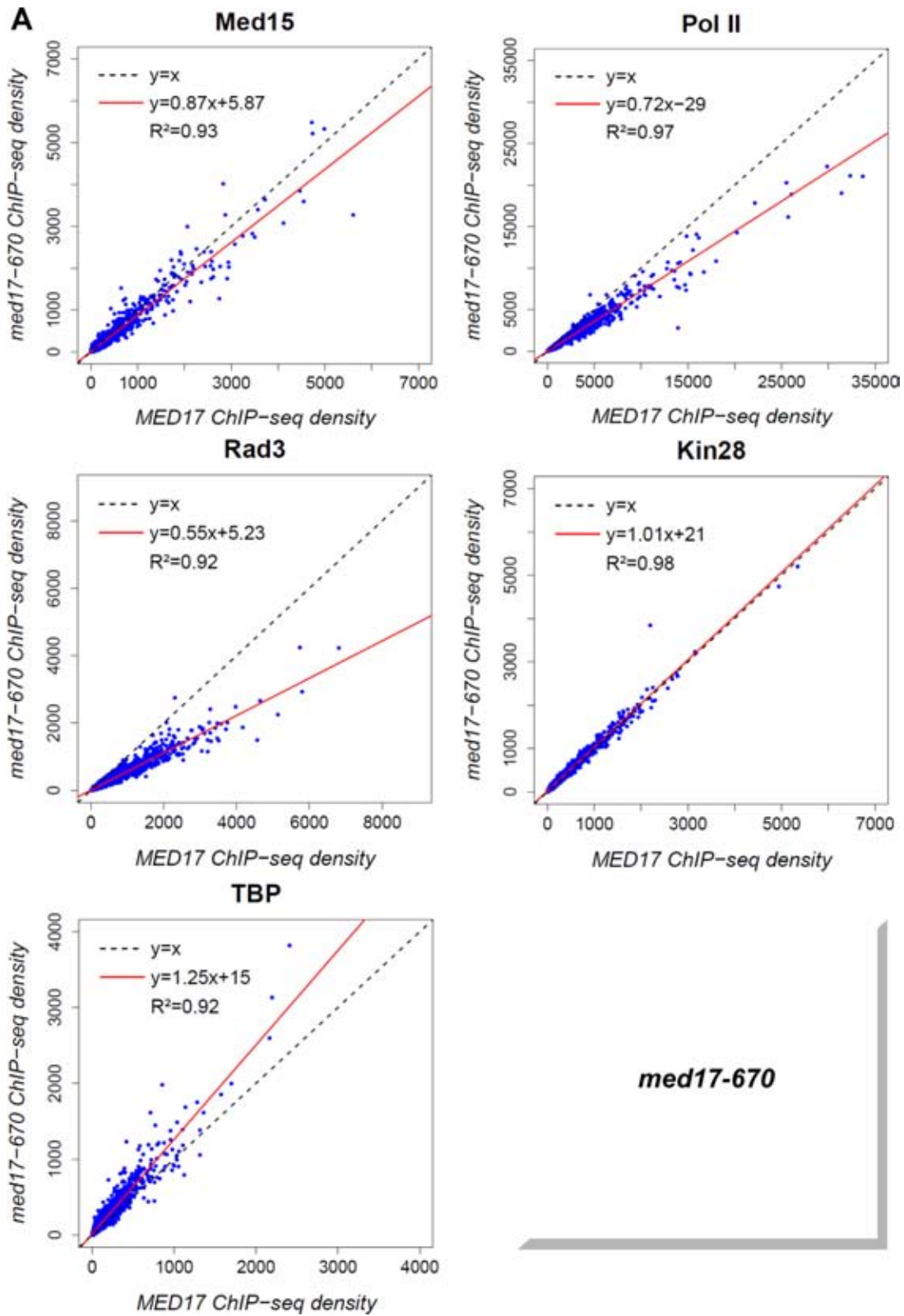


Figure 5. Global destabilization of PIC assembly in *med17-444* and *-504* mutant strains. The density of sequence tags in Med15, Rad3, Kin28 and TBP ChIP-seq experiments was calculated for promoter regions of Pol II-transcribed genes. Cells were grown at 30°C in YPD medium and then transferred for 45 minutes to 37°C. The density of sequence tags in the Pol II ChIP-seq experiments was calculated for the Pol II-transcribed genes. In all, 3852 Pol II-enriched genes were used for these analyses. Tag densities were normalized relative to qPCR data on a set of selected genes. Each point on the plot corresponds to one promoter region or one ORF. Promoter regions correspond to intergenic regions in tandem or in divergent orientation, excluding intergenic regions encompassing Pol III-transcribed genes. A linear regression (red line) for ChIP-seq density in mutant versus ChIP-seq density in WT and an R^2 correlation coefficient are indicated. (A) Med15, Pol II, Rad3, Kin28 and TBP ChIP-seq density in *med17-444* versus ChIP-seq density in WT on class II promoter regions or ORFs (for Pol II). (B) Med15, Pol II, Rad3, Kin28 and TBP ChIP-seq density in *med17-504* versus ChIP-seq density in WT on class II promoter regions or ORFs (for Pol II).



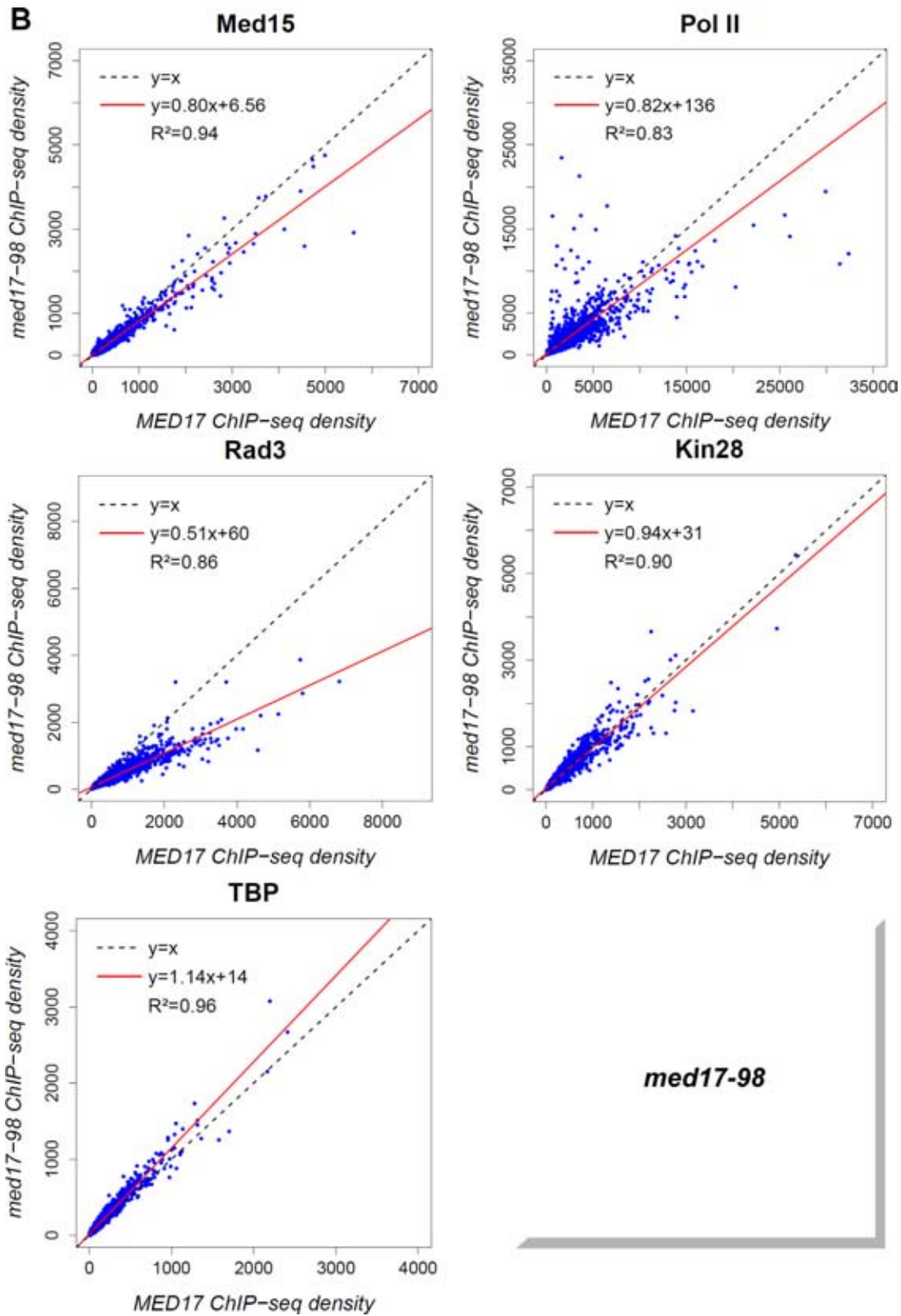


Figure 6. Global decrease of TFIIH core association independent of TFIK in *med17-670* and *-98* mutant strains. The density of sequence tags in the Med15, Pol II, Rad3, Kin28 and TBP ChIP-seq experiments was calculated as in Figure 5. (A) Med15, Pol II, Rad3, Kin28 and TBP ChIP-seq density in *med17-670* versus ChIP-seq density in WT on class II promoter regions or ORFs (for Pol II). (B) Med15, Pol II, Rad3, Kin28 and TBP ChIP-seq density in *med17-98* versus ChIP-seq density in WT on class II promoter regions or ORFs (for Pol II).

the recruitment or the stability of TFIID kinase and core modules is regulated independently on the yeast genome. (iv) This work suggests a possible mechanism for infantile cerebral atrophy caused by a mutation in Med17. We propose that the corresponding mutation in yeast Med17 affects Mediator–Pol II contact that leads to a global destabilization of PICs.

Independent recruitment of TFIID modules

The Mediator head module directly interacts with the TFIID complex via Med11–Rad3 contact and this interaction is important for the recruitment of this GTF to promoters (27). Mutations in the Med11 subunit differentially affected the promoter occupancy of the TFIID core and TFIID kinase modules. A *med11* mutation that impaired Mediator–TFIID contact destabilized TFIID on a group of promoters. The genome-wide analysis of *med17* mutants performed in this study supports a functional role of Mediator in the independent recruitment of TFIID core (Rad3 subunit) and TFIID kinase (Kin28 subunit) modules. Indeed, the *med17-504* mutation reduced the global association of Rad3 and Kin28 to a similar extent. On the contrary, the *med17-444* mutation, while affecting the occupancy of both TFIID modules, displayed a particularly pronounced decrease in Kin28 occupancy. These results show that the TFIID module was further destabilized in the *med17-444* mutant, similarly to the *med11-T47A* mutant. Conversely, we also observed that the genome-wide association of the TFIID core subcomplex could be affected independently of the TFIID subcomplex in *med17-670* and *med17-98* mutants. A large decrease of Rad3 genome-wide occupancy was observed in these mutants, whereas the association of the Kin28 TFIID subunit remained unchanged. Taken together, these results suggest that Mediator differentially affects the recruitment or the stabilization of TFIID modules and that the stability of TFIID modules is regulated independently within the PICs. We show that two biochemically separable TFIID modules have independent *in-vivo* behavior. The recruitment of the core TFIID module can be selectively affected, independently of TFIID and TBP. In line with the functional links between Mediator and TFIID, one of the Mediator functions consists in stimulating the CTD phosphorylation by Kin28 TFIID kinase (22). Moreover, Kin28 was reported to phosphorylate Med4 and Med14 Mediator subunits *in vitro* (60) and Kin28 phosphorylation of Med4 occurs *in vivo* (61). Recently, the phosphorylation of the Pol II CTD by Kin28 TFIID kinase has been reported to regulate the transient association of Mediator to core promoters and a release of Mediator from the PIC and the promoter escape (15,56), highlighting reciprocal links between Mediator and the TFIID kinase module.

Mediator contribution to TBP recruitment *in vivo*

TBP directly binds to the Mediator head module via the Med8–Med18–Med20 subcomplex (23,25). Previously, human Mediator and TBP-containing TFIID complexes have been shown to assemble cooperatively on promoter DNA *in vitro* (24). In accordance with *in-vitro* data, our *in-vivo* genome-wide analysis of yeast Mediator mutants showed

that TBP chromatin binding could be affected by Mediator mutations (*med17-444* and *-504*). It should be noted that the loss of Mediator function in the *med17-138* mutant, in which the Mediator head disassembles, has been previously shown to reduce TBP occupancy of several class II promoters (62,63). TBP association depended on the Mediator mutant examined in this study. The genome-wide binding of this GTF either did not decrease or was found to slightly increase in the *med17-670* and *med17-98* mutants. In this work, we identified mutations in the Med17 subunit that affected global TBP binding to chromatin, further supporting a coordinated behavior with Mediator. Interestingly, even though the effects of *med17* mutations are global, the range of effects can depend on specific genes. For example, the *med17-504* mutation resulted in a global decrease in TBP occupancy, but did not affect TBP recruitment to the *GALI* gene promoter upon transcription activation. Conversely, the *med17-98* mutation did not reduce genome-wide TBP or TFIID occupancy. However, in this mutant the TFIID binding, together with that of Pol II and the core TFIID module, was reduced on the three model promoters, but not in the case of TBP. These results therefore suggest a selective contribution of Mediator in TBP recruitment or stabilization to chromatin.

Global stabilization of PIC by Mediator

In *med17-444* and *-504* mutants, PIC assembly was destabilized leading to a decrease of Pol II, Rad3 TFIID core subunit, Kin28 TFIID subunit and TBP presence. The following mechanistic explanations specific to each mutant can be proposed. The presence of the Mediator tail module (Med15 subunit) on chromatin was globally affected in the *med17-444* mutant. We propose that, as a consequence of this effect on tail module occupancy, overall PIC assembly is destabilized with a large decrease in Pol II, Kin28 and TBP occupancy and a moderate effect on Rad3 occupancy. In contrast, the binding of Mediator remained globally unaffected in the *med17-504* mutant. This mutation led also to a global decrease in Pol II and GTF occupancy. We propose that an impaired interaction between Mediator (the Med17 subunit) and Pol II (the Rpb3 subunit) in this mutant destabilized the global PIC assembly *in vivo*, as was observed for the Pol II *rpb3-2* mutant with reduced Mediator–Pol II interaction (28) highlighting the functional importance of Mediator–Pol II contact. Taken together, our results revealed that specific point mutations in the Med17 Mediator head subunit (*med17-Q444P*, *med17-M504P*) could result in the global destabilization of PIC assembly *in vivo* through two different mechanisms, a destabilization of Mediator binding to chromatin or an impaired interaction between Mediator and Pol II, consistent with a central role for the Med17 subunit in Mediator function.

Among the *med17* mutants examined, only the *med17-444* mutation reduced the global Mediator tail (Med15 subunit) occupancy. Other *med17* mutations either did not affect or only slightly reduced the level of genome-wide Med15 binding to chromatin. ChIP experiments on the *ADH1*, *PYK1* and *PMA1* genes did not reveal any decrease in Med6 Mediator head or Med5 Mediator tail subunit occupancy in *med17* mutants. Moreover, Mediator stability

in the *med17* mutants examined in this study is supported by mass spectrometry and CoIP experiments indicating the presence of all Mediator modules. A decrease in ChIP signal for the Med15 Mediator tail subunit in *med17-444* might reflect a conformational change in Mediator complex bound to chromatin leading to reduced cross-linking efficiency or Mediator tail destabilization on chromatin.

A possible mechanism for infantile cerebral atrophy

Mediator is conserved in all eukaryotes. The mutated residues in the *med17* mutants characterized in this study correspond to the same (human Q310 for yeast Q444P in *med17-444* and human E314 for yeast E448Q, one of the three mutations in *med17-98*) or functionally close (human L371 for yeast M504P in *med17-504*) amino acids in human Med17 protein (Supplementary Figure S6). Given the conservation of the Mediator subunits, the molecular mechanisms of the Mediator function in PIC assembly are probably conserved from yeast to human and might extend to all eukaryotes. We showed that the *med17-M504P* mutation, proposed to be equivalent to the human mutation responsible for infantile cerebral atrophy, led to a severe temperature-sensitive phenotype in yeast, impaired Pol II–Mediator interaction and resulted in a decrease of genome-wide occupancy of all PIC components tested except Mediator itself. These results suggest that infantile cerebral atrophy might be the direct consequence of a defect in Mediator–Pol II contact and a global destabilization of PICs. Recently, we have shown that Mediator, in addition to its coregulator role, can also serve as a link between transcription and DNA repair (45). Several *med17* mutants were UV-sensitive in a global-genome repair-deficient context. Furthermore, *med17-504* and the other mutants analyzed in this study do not show any increase in UV sensitivity (data not shown), suggesting that these mutations led to transcriptional defects only.

Specific versus global effects of *med17* mutations

One of the questions that we addressed in analysing the impact of conditional mutations in the Med17 Mediator subunit on genome-wide PIC assembly was its gene-specific versus global nature. A high correlation between genome-wide densities of PIC components highlights the global effects of *med17* mutations. It should be noted that the range of the effects could still be gene-specific. Consistent with a central role for the Med17 subunit in Mediator function, all *med17* mutations reduced the global Pol II chromatin binding, but to a varying extent. The global decrease of Pol II occupancy in *med17* mutants compared to the WT was associated with a high correlation coefficient, indicating that most of the genes were affected to a similar extent. The *med17-98* mutant was an exception since it showed a global decrease of Pol II presence but with a reduced correlation coefficient (R^2 equal to 0.83) suggesting a more complex and gene-specific transcriptional effect for this mutation. The occupancy of the Kin28 TFIIF subunit remained globally unchanged in the *med17-98* mutant compared to the WT, but Kin28 ChIP signals were decreased on the *ADH1*, *PYK1* and *PMA1* gene promoters. The *med17-444* mutation did

not modify the presence of the Rad3 TFIIF core subunit on the three constitutively expressed genes or the *GAL1* inducible gene, but our genome-wide analysis revealed a slight decrease in global Rad3 occupancy. The *med17-504* mutant had a particular effect on the *GAL1* gene, showing a reduced Mediator association and unmodified Rad3 and TBP recruitment, whereas the genome-wide Mediator occupancy was similar to that of the WT and the occupancy of all other PIC components was globally reduced.

In conclusion, we suggest that Mediator differentially coordinates the multiple steps of a PIC assembly pathway *in vivo* at the level of independent recruitment or stabilization of several PIC components including Pol II, TBP and two TFIIF modules. The global effects of Mediator mutations on PIC formation are consistent with a central role for the Med17 subunit in Mediator function and organization. Additional PIC components, for example, GTFs that were not examined in this study, might contribute to PIC assembly. Together with Mediator, other coactivators, such as SAGA or TFIID complexes, modify PIC assembly and activity. In future studies it would be worth considering the dynamics of PIC assembly to improve understanding of *in-vivo* transcription mechanisms at work in eukaryotes.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank F.J. Asturias for CA001 strain, C. Thermes, E. Van Dijk and Y. Jaszczyszyn for performing the high throughput-sequencing of ChIP samples, the SPI (CEA/Saclay) for monoclonal antibodies and L. Kuras for critical reading of the manuscript. This work has benefited from the facilities and expertise of the high-throughput sequencing and the SICaPS mass spectrometry platforms of IMAGIF (Centre de Recherche de Gif, www.imagif.cnrs.fr).

FUNDING

Agence Nationale de la Recherche [ANR 11 BSV8 020 01, ANR-14-CE10-0012-01]; Fondation ARC [SL 2201130607079]; Fondation ARC [DOC20130606697 to F.E.]. Funding for open access charge: Fondation ARC [SL 2201130607079].

Conflict of interest statement. None declared.

REFERENCES

- Buratowski, S., Hahn, S., Guarente, L. and Sharp, P.A. (1989) Five intermediate complexes in transcription initiation by RNA polymerase II. *Cell*, **56**, 549–561.
- Ranish, J.A. and Hahn, S. (1996) Transcription: basal factors and activation. *Curr. Opin. Genet. Dev.*, **6**, 151–158.
- Grunberg, S. and Hahn, S. (2013) Structural insights into transcription initiation by RNA polymerase II. *Trends Biochem. Sci.*, **38**, 603–611.
- Rhee, H.S. and Pugh, B.F. (2012) Genome-wide structure and organization of eukaryotic pre-initiation complexes. *Nature*, **483**, 295–301.
- Xu, Z., Wei, W., Gagneur, J., Perocchi, F., Clauder-Munster, S., Camblong, J., Guffanti, E., Stutz, F., Huber, W. and Steinmetz, L.M. (2009) Bidirectional promoters generate pervasive transcription in yeast. *Nature*, **457**, 1033–1037.

6. Tsai, K.L., Tomomori-Sato, C., Sato, S., Conaway, R.C., Conaway, J.W. and Asturias, F.J. (2014) Subunit architecture and functional modular rearrangements of the transcriptional mediator complex. *Cell*, **157**, 1430–1444.
7. Wang, X., Sun, Q., Ding, Z., Ji, J., Wang, J., Kong, X., Yang, J. and Cai, G. (2014) Redefining the modular organization of the core Mediator complex. *Cell Res.*, **24**, 796–808.
8. Guglielmi, B., van Berkum, N.L., Klapholz, B., Bijma, T., Boube, M., Boschiero, C., Bourbon, H.M., Holstege, F.C. and Werner, M. (2004) A high resolution protein interaction map of the yeast Mediator complex. *Nucleic Acids Res.*, **32**, 5379–5391.
9. Koschubs, T., Lorenzen, K., Baumli, S., Sandstrom, S., Heck, A.J. and Cramer, P. (2010) Preparation and topology of the Mediator middle module. *Nucleic Acids Res.*, **38**, 3186–3195.
10. Lariviere, L., Plaschka, C., Seizl, M., Wenzek, L., Kurth, F. and Cramer, P. (2012) Structure of the Mediator head module. *Nature*, **492**, 448–451.
11. Robinson, P.J., Bushnell, D.A., Trnka, M.J., Burlingame, A.L. and Kornberg, R.D. (2012) Structure of the Mediator Head module bound to the carboxy-terminal domain of RNA polymerase II. *Proc. Natl. Acad. Sci. U.S.A.*, **109**, 17931–17935.
12. Kornberg, R.D. (2005) Mediator and the mechanism of transcriptional activation. *Trends Biochem. Sci.*, **30**, 235–239.
13. Ries, D. and Meisterernst, M. (2011) Control of gene transcription by Mediator in chromatin. *Semin. Cell Dev. Biol.*, **22**, 735–740.
14. 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.
15. Jeronimo, C. and Robert, F. (2014) Kin28 regulates the transient association of Mediator with core promoters. *Nat. Struct. Mol. Biol.*, **21**, 449–455.
16. Lacombe, T., Poh, S.L., Barbey, R. and Kuras, L. (2013) Mediator is an intrinsic component of the basal RNA polymerase II machinery in vivo. *Nucleic Acids Res.*, **41**, 9651–9662.
17. Paul, E., Zhu, Z.I., Landsman, D. and Morse, R.H. (2015) Genome-wide association of mediator and RNA polymerase II in wild-type and mediator mutant yeast. *Mol. Cell Biol.*, **35**, 331–342.
18. Thompson, C.M. and Young, R.A. (1995) General requirement for RNA polymerase II holoenzymes in vivo. *Proc. Natl. Acad. Sci. U.S.A.*, **92**, 4587–4590.
19. Andrau, J.C., van de Pasch, L., Lijnzaad, P., Bijma, T., Koerkamp, M.G., van de Peppel, J., Werner, M. and Holstege, F.C. (2006) Genome-wide location of the coactivator mediator: Binding without activation and transient Cdk8 interaction on DNA. *Mol. Cell*, **22**, 179–192.
20. 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.
21. Jiang, Y.W., Veschambre, P., Erdjument-Bromage, H., Tempst, P., Conaway, J.W., Conaway, R.C. and Kornberg, R.D. (1998) Mammalian mediator of transcriptional regulation and its possible role as an end-point of signal transduction pathways. *Proc. Natl. Acad. Sci. U.S.A.*, **95**, 8538–8543.
22. 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.
23. Cai, G., Imasaki, T., Yamada, K., Cardelli, F., Takagi, Y. and Asturias, F.J. (2010) Mediator head module structure and functional interactions. *Nat. Struct. Mol. Biol.*, **17**, 273–279.
24. Johnson, K.M., Wang, J., Smallwood, A., Arayata, C. and Carey, M. (2002) TFIID and human mediator coactivator complexes assemble cooperatively on promoter DNA. *Genes Dev.*, **16**, 1852–1863.
25. Lariviere, L., Geiger, S., Hoepfner, S., Rother, S., Strasser, K. and Cramer, P. (2006) Structure and TBP binding of the Mediator head subcomplex Med8-Med18-Med20. *Nat. Struct. Mol. Biol.*, **13**, 895–901.
26. Takahashi, H., Parmely, T.J., Sato, S., Tomomori-Sato, C., Banks, C.A., Kong, S.E., Szutorisz, H., Swanson, S.K., Martin-Brown, S., Washburn, M.P. et al. (2011) Human mediator subunit MED26 functions as a docking site for transcription elongation factors. *Cell*, **146**, 92–104.
27. Esnault, C., Ghavi-Helm, Y., Brun, S., Soutourina, J., Van Berkum, N., Boschiero, C., Holstege, F. and Werner, M. (2008) Mediator-dependent recruitment of TFIID modules in preinitiation complex. *Mol. Cell*, **31**, 337–346.
28. Soutourina, J., Wydau, S., Ambroise, Y., Boschiero, C. and Werner, M. (2011) Direct interaction of RNA polymerase II and mediator required for transcription in vivo. *Science*, **331**, 1451–1454.
29. Bourbon, H.M. (2008) Comparative genomics supports a deep evolutionary origin for the large, four-module transcriptional mediator complex. *Nucleic Acids Res.*, **36**, 3993–4008.
30. Risheg, H., Graham, J.M. Jr, Clark, R.D., Rogers, R.C., Opitz, J.M., Moeschler, J.B., Peiffer, A.P., May, M., Joseph, S.M., Jones, J.R. et al. (2007) A recurrent mutation in MED12 leading to R961W causes Opitz-Kaveggia syndrome. *Nat. Genet.*, **39**, 451–453.
31. Schwartz, C.E., Tarpey, P.S., Lubs, H.A., Verloes, A., May, M.M., Risheg, H., Friez, M.J., Futreal, P.A., Edkins, S., Teague, J. et al. (2007) The original Lujan syndrome family has a novel missense mutation (p.N1007S) in the MED12 gene. *J. Med. Genet.*, **44**, 472–477.
32. Hashimoto, S., Boissel, S., Zarhrate, M., Rio, M., Munnich, A., Egly, J.M. and Colleaux, L. (2011) MED23 mutation links intellectual disability to dysregulation of immediate early gene expression. *Science*, **333**, 1161–1163.
33. Firestein, R., Bass, A.J., Kim, S.Y., Dunn, I.F., Silver, S.J., Guney, I., Freed, E., Ligon, A.H., Vena, N., Ogino, S. et al. (2008) CDK8 is a colorectal cancer oncogene that regulates beta-catenin activity. *Nature*, **455**, 547–551.
34. Gade, P., Singh, A.K., Roy, S.K., Reddy, S.P. and Kalvakolanu, D.V. (2009) Down-regulation of the transcriptional mediator subunit Med1 contributes to the loss of expression of metastasis-associated dapk1 in human cancers and cancer cells. *Int. J. Cancer*, **125**, 1566–1574.
35. Kuuselo, R., Savinainen, K., Sandstrom, S., Autio, R. and Kallioniemi, A. (2011) MED29, a component of the mediator complex, possesses both oncogenic and tumor suppressive characteristics in pancreatic cancer. *Int. J. Cancer*, **129**, 2553–2565.
36. Li, L.H., He, J., Hua, D., Guo, Z.J. and Gao, Q. (2010) Lentivirus-mediated inhibition of Med19 suppresses growth of breast cancer cells in vitro. *Cancer Chemother. Pharmacol.*, **68**, 207–215.
37. Vijayvargia, R., May, M.S. and Fondell, J.D. (2007) A coregulatory role for the mediator complex in prostate cancer cell proliferation and gene expression. *Cancer Res.*, **67**, 4034–4041.
38. Zhang, X., Krutchinsky, A., Fukuda, A., Chen, W., Yamamura, S., Chait, B.T. and Roeder, R.G. (2005) MED1/TRAP220 exists predominantly in a TRAP/ Mediator subpopulation enriched in RNA polymerase II and is required for ER-mediated transcription. *Mol. Cell*, **19**, 89–100.
39. Thompson, C.M., Koleske, A.J., Chao, D.M. and Young, R.A. (1993) A multisubunit complex associated with the RNA polymerase II CTD and TATA-binding protein in yeast. *Cell*, **73**, 1361–1375.
40. Linder, T., Zhu, X., Baraznenok, V. and Gustafsson, C.M. (2006) The classical srb4–138 mutant allele causes dissociation of yeast Mediator. *Biochem. Biophys. Res. Commun.*, **349**, 948–953.
41. Takagi, Y., Calero, G., Komori, H., Brown, J.A., Ehrensberger, A.H., Hudmon, A., Asturias, F. and Kornberg, R.D. (2006) Head module control of mediator interactions. *Mol. Cell*, **23**, 355–364.
42. Takagi, Y. and Kornberg, R.D. (2006) Mediator as a general transcription factor. *J. Biol. Chem.*, **281**, 80–89.
43. Kaufmann, R., Straussberg, R., Mandel, H., Fattal-Valevski, A., Ben-Zeev, B., Naamati, A., Shaag, A., Zenvirt, S., Konen, O., Mimouni-Bloch, A. et al. (2010) Infantile cerebral and cerebellar atrophy is associated with a mutation in the MED17 subunit of the transcription preinitiation mediator complex. *Am. J. Hum. Genet.*, **87**, 667–670.
44. 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.
45. Eyboullet, F., Cibot, C., Eychenne, T., Neil, H., Alibert, O., Werner, M. and Soutourina, J. (2013) Mediator links transcription and DNA repair by facilitating Rad2/XPG recruitment. *Genes Dev.*, **27**, 2549–2562.
46. Soutourina, J., Bordas-Le Floch, V., Gendrel, G., Flores, A., Ducrot, C., Dumay-Odelot, H., Soularue, P., Navarro, F., Cairns, B.R., Lefebvre, O.

- et al.* (2006) Rsc4 connects the chromatin remodeler RSC to RNA polymerases. *Mol. Cell Biol.*, **26**, 4920–4933.
47. Cai, G., Imasaki, T., Takagi, Y. and Asturias, F.J. (2009) Mediator structural conservation and implications for the regulation mechanism. *Structure*, **17**, 559–567.
 48. Schmitt, M.E., Brown, T.A. and Trumppower, B.L. (1990) A rapid and simple method for preparation of RNA from *Saccharomyces cerevisiae*. *Nucleic Acids Res.*, **18**, 3091–3092.
 49. Langmead, B., Trapnell, C., Pop, M. and Salzberg, S.L. (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.*, **10**, R25.
 50. Zhang, Y., Liu, T., Meyer, C.A., Eeckhoutte, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W. *et al.* (2008) Model-based analysis of ChIP-Seq (MACS). *Genome Biol.*, **9**, R137.
 51. Salmon-Divon, M., Dvinge, H., Tammoja, K. and Bertone, P. (2010) PeakAnalyzer: genome-wide annotation of chromatin binding and modification loci. *BMC Bioinformatics*, **11**, 415.
 52. Spyrou, C., Stark, R., Lynch, A.G. and Tavare, S. (2009) BayesPeak: Bayesian analysis of ChIP-seq data. *BMC Bioinformatics*, **10**, 299.
 53. Haimovich, G., Medina, D.A., Causse, S.Z., Garber, M., Millan-Zambrano, G., Barkai, O., Chavez, S., Perez-Ortin, J.E., Darzacq, X. and Choder, M. (2013) Gene expression is circular: factors for mRNA degradation also foster mRNA synthesis. *Cell*, **153**, 1000–1011.
 54. Sun, M., Schwab, B., Schulz, D., Pirkl, N., Etzold, S., Lariviere, L., Maier, K.C., Seizl, M., Tresch, A. and Cramer, P. (2012) Comparative dynamic transcriptome analysis (cDTA) reveals mutual feedback between mRNA synthesis and degradation. *Genome Res.*, **22**, 1350–1359.
 55. Bonnet, J., Wang, C.Y., Baptista, T., Vincent, S.D., Hsiao, W.C., Stierle, M., Kao, C.F., Tora, L. and Devys, D. (2014) The SAGA coactivator complex acts on the whole transcribed genome and is required for RNA polymerase II transcription. *Genes Dev.*, **28**, 1999–2012.
 56. Wong, K.H., Jin, Y. and Struhl, K. (2014) TFIIH Phosphorylation of the Pol II CTD Stimulates Mediator Dissociation from the Preinitiation Complex and Promoter Escape. *Mol. Cell*, **54**, 601–612.
 57. Chen, Y., Negre, N., Li, Q., Mieczkowska, J.O., Slattery, M., Liu, T., Zhang, Y., Kim, T.K., He, H.H., Zieba, J. *et al.* (2012) Systematic evaluation of factors influencing ChIP-seq fidelity. *Nat. Methods*, **9**, 609–614.
 58. Park, D., Lee, Y., Bhupindersingh, G. and Iyer, V.R. (2013) Widespread misinterpretable ChIP-seq bias in yeast. *PLoS One*, **8**, e83506.
 59. Teytelman, L., Thurtle, D.M., Rine, J. and van Oudenaarden, A. (2013) Highly expressed loci are vulnerable to misleading ChIP localization of multiple unrelated proteins. *Proc. Natl. Acad. Sci. U.S.A.*, **110**, 18602–18607.
 60. Liu, Y., Kung, C., Fishburn, J., Ansari, A.Z., Shokat, K.M. and Hahn, S. (2004) Two cyclin-dependent kinases promote RNA polymerase II transcription and formation of the scaffold complex. *Mol. Cell Biol.*, **24**, 1721–1735.
 61. Guidi, B.W., Bjornsdottir, G., Hopkins, D.C., Lacomis, L., Erdjument-Bromage, H., Tempst, P. and Myers, L.C. (2004) Mutual targeting of mediator and the TFIIH kinase Kin28. *J. Biol. Chem.*, **279**, 29114–29120.
 62. Kuras, L. and Struhl, K. (1999) Binding of TBP to promoters in vivo is stimulated by activators and requires Pol II holoenzyme. *Nature*, **399**, 609–613.
 63. Li, X.Y., Virbasius, A., Zhu, X. and Green, M.R. (1999) Enhancement of TBP binding by activators and general transcription factors. *Nature*, **399**, 605–609.

Titre : Etude intégrative du rôle de deux sous unités essentielles du Médiateur de la transcription dans la mise en place des complexes de pré-initiation.

Mots clés : Transcription, Mediateur, Génomique, Régulation, *Saccharomyces cerevisiae*, ARN polymérase II

La transcription est la première étape de l'expression des gènes. Chez les eucaryotes, la transcription par l'ARN polymérase II est un processus hautement régulé. Elle commence par la fixation d'activateurs spécifiques sur des régions régulatrices. Cela permet le recrutement de co-activateurs suivi des facteurs généraux de la transcription (GTFs) et de l'ARN polymérase II (Pol II) pour former le complexe de préinitiation (PIC). Le Médiateur est un complexe co-activateur essentiel à ce processus. Chez la levure *Saccharomyces cerevisiae*, il est composé de 25 sous-unités dont 10 sont essentielles à la viabilité. Son rôle principal est d'intégrer les signaux de régulation pour les transmettre aux composants du PIC. On connaît aujourd'hui un certain nombre de fonctions du Médiateur. Néanmoins, sa complexité et la présence de sous-unités essentielles compliquent la compréhension détaillée de son mécanisme de fonctionnement *in vivo*. Au cours de ma thèse, je me suis intéressé aux sous-unités essentielles Med10 et Med7, toutes deux appartenant au module du milieu du Médiateur, peu étudié jusqu'à présent. Nous avons construit une collection de mutants thermosensibles de ces deux sous-unités chez la levure *S. cerevisiae*.

Nous avons caractérisés ces mutants par différentes approches de biologie moléculaire, biochimie et génomique fonctionnelle. L'étude de la sous-unité Med10 nous a permis de mettre en évidence *in vivo* un lien fonctionnel entre le Médiateur et TFIIB, un GTF essentiel au recrutement de la Pol II. Nous avons ainsi identifié les sous-unités Med14 et Med10 qui sont en contact avec TFIIB. Nos analyses de ChIP-seq montrent que le module du milieu et Med10, en particulier, est requis pour la formation correcte du PIC sur l'ensemble du génome. Ces données nous ont également permis de montrer que le Médiateur influence la formation du PIC en relation avec l'architecture des promoteurs en termes de présence de boîtes TATA, d'occupation des nucléosomes et leur dynamique. Ce travail nous a permis une meilleure compréhension du rôle du Médiateur dans l'activation de la transcription et donné des informations mécanistiques sur la façon dont l'interaction entre le Médiateur et TFIIB (et les autres GTFs) ainsi que l'architecture des promoteurs mènent à une régulation gène-spécifique.

Title : Integrative study of the role of two Mediator essential subunits in transcription initiation.

Keywords : Transcription, Mediator, Genomics, Regulation, *Saccharomyces cerevisiae*, RNA polymerase II

Abstract : Transcription is the first step of gene expression. In eukaryotes, messenger RNA (mRNA) transcription is a highly regulated process. Transcription begins with the binding of a specific transcription factor on a DNA regulatory sequence. This enable the recruitment of co-activators, followed by general transcription factors (GTFs) and RNA polymerase II (Pol II) to form preinitiation complex (PIC). Mediator is a co-activator complex which is essential in this process. In yeast *Saccharomyces cerevisiae*, Mediator is composed of 25 subunits, among which 10 are essential for cell viability, organized into four distinct modules. The main role of this complex is to transmit regulatory signal to PIC components. Although Mediator has been the subject of a large numbers of studies, its complexity prevents the detailed understanding of how it acts *in vivo*. During my PhD, I focused my work on the study of the two essential subunits Med7 and Med10. Both of these subunits belong to the middle module, poorly studied so far.

We obtained a collection of temperature-sensitive mutants of Med7 and Med10 in yeast *S. cerevisiae*. We used different molecular biology and functional genomics to characterize these mutants. The work on Med10 subunit enabled us to highlight *in vivo* a functional link between Mediator and TFIIB, one of the GTFs. Notably, we have shown a new contact between Med14 subunit and TFIIB. Our ChIP-seq analysis shows that Mediator middle module, and in particular Med10 subunits, is crucial for PIC assembly genome-wide. These data also permit us to show that Mediator influence PIC formation in relation to promoter architecture. Taken together, these results indicates that Mediator in crucial to orchestrate the incorporation of the different proteins into the PIC. This work permit us to improve our understanding of how functional interplay between Mediator, TFIIB, other GTFs, and the promoter architecture leads to gene-specific transcription.