



**HAL**  
open science

## Rôles transcriptionnels des facteurs NER

Izarn Iltis

► **To cite this version:**

Izarn Iltis. Rôles transcriptionnels des facteurs NER. Biochimie, Biologie Moléculaire. Université de Strasbourg, 2012. Français. NNT : 2012STRAJ079 . tel-00849966

**HAL Id: tel-00849966**

**<https://theses.hal.science/tel-00849966>**

Submitted on 2 Aug 2013

**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.

**ÉCOLE DOCTORALE Sciences de la vie et de la santé**  
**IGBMC - Inserm U 964**

**THÈSE** présentée par :  
**Izarn ILTIS**

soutenue le : **07 Décembre 2012**

pour obtenir le grade de : **Docteur de l'université de Strasbourg**  
Discipline/ Spécialité : **Sciences du vivant**

**Rôles transcriptionnels des facteurs NER**

**THÈSE dirigée par :**  
**Dr EGLY Jean-Marc**

DRE INSERM, Strasbourg (Fr)

**RAPPORTEURS :**  
**Dr JALINOT Pierre**  
**Dr ROMEO Paul-Henry**

DR CNRS, Lyon (Fr)  
DR INSERM, Paris (Fr)

---

**AUTRES MEMBRES DU JURY :**

**Dr DAVIDSON Irwin**  
**Pr LINDAHL Tomas**

DR INSERM, Strasbourg (Fr)  
Professeur, Londres (GB)



## THESE

Présentée par

**Izarn ILTIS**

Pour l'obtention du grade de

**Docteur en Sciences de l'Université de Strasbourg (UDS)**

Discipline : Sciences du Vivant

Spécialité : Aspects Moléculaires et Cellulaires de la Biologie

# Rôles transcriptionnels des facteurs NER

Soutenue publiquement le 07 décembre 2012 devant la commission d'examen :

Directeur de thèse :	Jean-Marc EGLY, DRE INSERM, Strasbourg (Fr)
Rapporteur externe :	Pierre JALINOT, DR CNRS, Lyon (Fr)
Rapporteur externe :	Paul-Henry ROMEO, DR INSERM, Paris (Fr)
Rapporteur interne :	Irwin DAVIDSON, DR INSERM, Strasbourg (Fr)
Examineur :	Tomas LINDAHL, Professeur, Londres (GB)

*« Savoir s'étonner à propos est le premier pas fait sur la route de la découverte »*

Louis Pasteur

## REMERCIEMENTS

*Je tiens à remercier vivement Messieurs Pierre Jalinot, Paul-Henry Romeo, Irwin Davidson et Tomas Lindahl de m'avoir fait l'honneur d'accepter d'être membres de mon jury.*

*Je remercie tout particulièrement Jean-Marc Egly, pour m'avoir accueilli dans son laboratoire et pour avoir été mon directeur de thèse. Merci pour sa confiance tout au long de ces quatre années, pour ses encouragements et pour son entrain afin que je donne à tout moments le meilleur de moi-même.*

*Je remercie également l'ensemble des membres de l'équipe, qui ont participé, de près ou de loin à ce travail.*

*Merci à Annabel, Cathy et Charlotte pour leur disponibilité et leur capacité à nous supporter.*

*Merci à Philippe, qui attend toute l'année les chutes de neiges pour aller rider.*

*Merci à Manu, qui a toujours une histoire rocambolique à nous raconter.*

*Merci à Fred, pour sa solidité défensive lors de nos matchs de football indoor.*

*Merci également à Christophe pour les bavardages intempestifs sur les technologies en tous genres, les joujoux Apple et les fixies.*

*Je tiens également à remercier tous les thésards, post-doc du laboratoire et autres compagnons de galère : merci à Amita, Lyse-Marie, Salim, Alex, Marc, Jitka, Alexei, Serena... je vous souhaite à tous bon courage pour la suite.*

*Des remerciements très particuliers et très chaleureux à mon grand frère de labo Nico, qui m'a encadré tout au long de ma thèse et qui m'a appris tant de techniques (excepté les westerns). Merci aussi pour avoir servi de paratonnerre lors des orages « eglyiens ».*

*Je tiens également à remercier la fondation ARC et leurs nombreux donateurs pour m'avoir permis de financer ma thèse durant les deux dernières années.*

*Dans la rubrique sportive, j'ai une pensée toute particulière pour Boul, qui a été mon coéquipier de handball et mon compagnon de footing. A bientôt sur un marathon !*

*Merci également à Christophe pour les parties de tennis endiablées, qui permettaient d'évacuer la pression.*

*Je n'oublie pas tous mes potes du hand, qui m'ont permis sur le terrain et en soirée, d'oublier le quotidien du labo.*

*Merci également aux gars du vélo, pour m'avoir initié à la pédale.*

*Merci à Docteur Romain, rencontré sur les bancs de la fac et avec qui nous avons écumés les soirées étudiantes.*

*Merci à Morgan et Cédric, qui à l'heure actuelle, doivent être entrain de reluquer les bikinis sur les plages australiennes. Merci à eux pour les week-ends de folie, les fêtes du vin arrosées et les divertissements en tous genres qui durent depuis des années.*

*Je dédie ce travail à mes parents qui m'ont soutenu et encouragé tout au long de mes études et plus particulièrement pendant les quatre années de recherche pour ma thèse. Je remercie aussi mon frère, Loan, et lui souhaite bon courage pour sa soutenance à venir. J'ai également une pensée particulière pour mes grands-parents, qui malheureusement n'ont pas pu voir l'aboutissement de mon travail. Un grand merci à tous les membres de ma famille pour leurs soutiens et leurs encouragements, même si le sujet de ma thèse restera peut-être un mystère pour eux.*

*Enfin des remerciements très chaleureux, très particuliers et très amoureux à Lucile. Grâce à son soutien permanent, à sa patience de tous les jours et à son sourire quotidien, j'ai pu mener à bien l'ensemble de ma thèse. Mille excuses pour le stress des derniers jours et surtout pour toutes ces grasses matinées du samedi matin qu'elle a passée toute seule.*

## TABLES DES MATIERES

FIGURES ET TABLEAUX .....	8
ABREVIATIONS .....	9
PREFACE .....	11
INTRODUCTION .....	12
I. La réparation de l'ADN .....	13
1. La réparation par excision de nucléotides (NER) .....	14
A. Les dommages .....	14
B. Le mécanisme de NER .....	15
B.1 Le mécanisme de GG-NER .....	15
B.2 Le mécanisme de réparation par TC-NER .....	19
2. Les autres mécanismes de réparation de l'ADN .....	22
A. La BER .....	22
B. La réparation des mésappariements (MMR) .....	25
C. La réparation des cassures doubles brins .....	27
II. Les maladies liées à un défaut de la NER .....	29
1. Xeroderma Pigmentosum (XP) .....	30
2. Le syndrome de Cockayne (CS) .....	33
3. La trichothiodystrophie (TTD) .....	34
III. La transcription des gènes de classe II .....	36
1. Les séquences régulatrices de l'ADN .....	37
A. La structure du promoteur minimal .....	38
A.1 La boîte TATA .....	38
A.2 L'élément initiateur .....	38
A.3 Les éléments en aval du promoteur .....	39
A.4 Le site de reconnaissance de TFIIB .....	39
A.5 Les autres éléments du promoteur minimal .....	40
A.6 Les promoteurs à îlots CpG .....	40
B. Les éléments de régulation .....	40
2. L'ARN polymérase II .....	41
3. Les facteurs généraux de transcription .....	45
4. Le cycle de transcription .....	48
5. La régulation de la transcription .....	51
A. La régulation de la transcription par les récepteurs nucléaires .....	51
A.1 Le récepteur à l'acide rétinoïque .....	52
A.2 La transcription des gènes cibles à l'acide rétinoïque .....	54

<i>B. La régulation de la chromatine</i> .....	56
<i>B.1 Organisation de la chromatine</i> .....	57
<i>B.2 Les variants d'histones</i> .....	60
<i>B.3 Les modifications post-traductionnelles des histones</i> .....	61
<i>B.4 La méthylation de l'ADN</i> .....	69
<i>B.5 Remodelage de la chromatine</i> .....	71
<b>IV. Les liens entre transcription et réparation de l'ADN</b> .....	74
1. TFIIH .....	74
<i>A. Composition de TFIIH</i> .....	74
<i>B. TFIIH et la réparation par excision de nucléotide</i> .....	76
<i>C. TFIIH : un facteur général de transcription</i> .....	77
2. Les autres facteurs .....	78
<b>RESULTATS</b> .....	80
Publication 1 .....	82
Publication 2 .....	84
Publication 3 .....	85
<b>DISCUSSION</b> .....	86
<b>REFERENCES</b> .....	95
<b>ANNEXES</b> .....	118
Publication 4 .....	119
Publication 5 .....	121



## FIGURES ET TABLEAUX

<b>Figure 1</b> : Les différents mécanismes de réparations.....	13
<b>Figure 2</b> : Lésions réparées par la NER.....	14
<b>Figure 3</b> : Mécanisme de réparation GG-NER.....	18
<b>Figure 4</b> : Mécanisme de réparation TC-NER.....	21
<b>Figure 5</b> : Mécanisme de réparation BER.....	24
<b>Figure 6</b> : Mécanisme de réparation MMR.....	26
<b>Figure 7</b> : Mécanisme de réparation des cassures double brin.....	29
<b>Figure 8</b> : Différentes maladies induites par des mutations dans les gènes NER.....	31
<b>Figure 9</b> : Le promoteur minimal eucaryote.....	37
<b>Figure 10</b> : Structure de l'ARN polymérase II.....	43
<b>Figure 11</b> : Cycle de phosphorylation du CTD de Rpb1 lors de la transcription.....	44
<b>Figure 12</b> : Les différentes étapes de l'initiation de la transcription.....	49
<b>Figure 13</b> : Organisation du récepteur à l'acide rétinoïque.....	53
<b>Figure 14</b> : Modèle classique d'activation des gènes cibles de l'acide rétinoïque.....	55
<b>Figure 15</b> : Une nouvelle voie de signalisation de RAR.....	56
<b>Figure 16</b> : Structure cristallographique du nucléosome.....	57
<b>Figure 17</b> : Les différents degrés d'organisation de la chromatine.....	59
<b>Figure 18</b> : Réaction d'acétylation des résidus lysines des histones.....	64
<b>Figure 19</b> : Les différentes formes de méthylation des résidus lysines des histones.....	67
<b>Figure 20</b> : Mécanisme de déméthylation active.....	71
<b>Figure 21</b> : Schéma de boucle de chromatine.....	73
<b>Figure 22</b> : Structure de TFIIH.....	74
<b>Tableau 1</b> : Tableau clinique des symptômes associés aux XP, TTD, CS et XP/CS.....	35
<b>Tableau 2</b> : Composition des différentes sous unités de l'ARN polymérase II.....	42
<b>Tableau 3</b> : Les différentes modifications post-traductionnelles des histones.....	62
<b>Tableau 4</b> : Composition du complexe humaine TFIIH.....	75

## ABREVIATIONS

<b>8-oxo-G</b>	8-oxo-guanine
<b>6-4PP</b>	(6-4) pyrimidine pyrimidone
<b>AP</b>	Apurique/Apyrimidique
<b>ATM</b>	Ataxia telengienstasia mutated
<b>ATR</b>	ATM and Rad3 related
<b>ARN</b>	Acide ribonucléique
<b>ARN Pol II</b>	ARN polymérase II
<b>BER</b>	Réparation par excision de base
<b>bp</b>	Paire de bases
<b>BRE</b>	TFIIB recognition element
<b>CAF-1</b>	Chromatin assembly factor 1
<b>CAK</b>	Cdk-Activated Kinase
<b>CDK</b>	Cyclin dependante kinase
<b>ChIP</b>	Chromatine immuno-précipitation
<b>COFS</b>	Syndrome Oculo-Facio-Skeletal
<b>CPD</b>	Dimère de cyclobutane pyrimidine
<b>CTD</b>	C-terminal domain
<b>CS</b>	Syndrome de Cokayne
<b>CSA/B</b>	Protéine du syndrome de Cockayne groupe A/B
<b>DNA</b>	Acide désoxyribonucléique
<b>DBD</b>	DNA-binding domain
<b>DCE</b>	Downstream core element
<b>DDB1/2</b>	Protéine damaged DNA binding 1/2
<b>DNA-PK</b>	DNA protein kinase
<b>DPE</b>	Downstream promoter element
<b>DRB</b>	5,6-Dichloro-1-beta-D-Ribofuranosylbenzimidazole
<b>DSB</b>	Double strand break
<b>ERCC1</b>	Excision repair cross-complementing rodent repair
<b>GGR</b>	Réparation globale du génome
<b>GTF</b>	Facteurs généraux de transcription
<b>HAT</b>	Histone acetyl-transférase
<b>HDAC</b>	Histone déacétyl-transférase
<b>HP1</b>	Heterochromatine protein 1
<b>HR</b>	Recombinaison homologue

<b>HR23B</b>	Homologue Rad23B
<b>Inr</b>	Initiateur
<b>kB</b>	Kilo-base
<b>kDa</b>	Kilo-dalton
<b>LBD</b>	Ligand binding domain
<b>Lig</b>	Ligase
<b>MAT1</b>	Protéine ménage à trois 1
<b>MMR</b>	Réparation des mésappariement
<b>MLL</b>	Mixed Lineage Leukemia
<b>MSH2</b>	Mut S Homologue 2 factor
<b>MPTs</b>	Modifications post-traductionnelles
<b>NER</b>	Réparation par excision de nucléotides
<b>NR</b>	Récepteur nucléaire
<b>Ogg1</b>	OxoG DNA glycosylase 1
<b>PCNA</b>	Proliferating cell nuclear antigen
<b>PIC</b>	Complexe de pré-initiation
<b>PPAR</b>	Peroxisomal Proliferator -Activated Receptor
<b>Pol</b>	Polymérase
<b>RAR</b>	Retinoic Acid Receptor
<b>t-RA</b>	trans-Retinoic Acid
<b>RE</b>	Elément de réponse
<b>RFC</b>	Facteur de réplication C
<b>RPA</b>	Facteur de réplication A
<b>RXR</b>	Retinoid X Receptor
<b>SWI/SNF</b>	Switch/Sucrose non fermentable
<b>TBP</b>	TATA box binding protein
<b>TCR</b>	Réparation couplée à la transcription
<b>TFIIA-H</b>	Facteur de transcription II A-V
<b>TTD</b>	Trichothiodystrophie
<b>UV</b>	Ultra-violet
<b>UV-DDB</b>	UV damaged DNA binding protein
<b>VDR</b>	Vitamin D Receptor
<b>XP</b>	Xeroderma Pigmentosum
<b>XPA-V</b>	XP complementation group A-V protein
<b>XRCC</b>	X ray cross complementation

## PREFACE

Il y a près d'un demi-siècle, l'ADN a été décrit comme étant le support de l'information génétique de nos cellules (Watson et Crick). Selon le dogme central de la biologie moléculaire, l'ADN est transcrit en ARN, qui est lui-même traduit en protéines. Les protéines permettent d'assurer les fonctions indispensables au développement et à la vie de la cellule. La synthèse de l'ARN résulte d'une cascade d'évènements, faisant intervenir de nombreuses protéines. Au moment de la transcription, une multitude de facteurs tels que l'ARN polymérase II, les facteurs généraux de transcription, des co-activateurs ainsi que des facteurs de remodelage de la chromatine sont assemblés au niveau de la région promotrice des gènes. Ces protéines contribuent à la formation de complexes permettant l'expression des gènes.

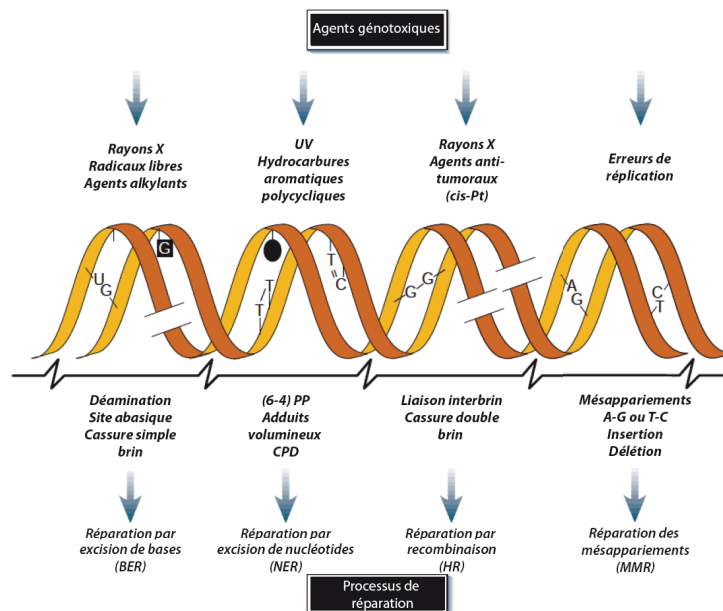
L'ADN est constamment altéré, aussi bien par le métabolisme cellulaire, que par des sources exogènes. Cet ADN endommagé, doit être réparé par la cellule, afin d'éviter la persistance de mutations et pour conserver la stabilité génétique de l'ADN. En effet, les mutations peuvent avoir de nombreuses conséquences phénotypiques, comme l'apparition de cancers ou le vieillissement cellulaire. Certaines mutations sont également à l'origine de maladies génétiques. Au sein du laboratoire, nous nous intéressons plus particulièrement à trois maladies génétiques humaines : *Xeroderma pigmentosum*, syndrome de Cockayne et Trichotiodystrophie. Ces pathologies sont issues de la mutation de gènes codant pour les protéines de réparation par excision de nucléotides (NER).

Au cours de ma thèse, je me suis intéressé aux différents rôles des facteurs NER (XPC, XPA, XPG et XPF). Nous avons mis en évidence que ces protéines, connues pour leurs actions dans la réparation de l'ADN, étaient recrutées lors de la transcription. Nous avons également montré qu'elles permettaient la mise en place d'un contexte de chromatine favorable à la transcription. Dans l'introduction de ma thèse, je décrirai d'une part les différents mécanismes de réparation et leurs principaux acteurs. D'une autre part, je développerai le processus de transcription, et les moyens mis en œuvre par la cellule pour le réguler.

## INTRODUCTION

## I. La réparation de l'ADN

Au cours de la vie, l'organisme est continuellement exposé à de nombreux agents génotoxiques susceptibles d'altérer l'ADN (Figure 1). Ainsi, l'exposition au soleil (rayonnement UV), la fumée de cigarettes ou encore le contact avec de nombreux composés chimiques (médicaments anticancéreux comme le cisplatine) peuvent créer des dommages sur l'ADN. Si ces lésions ne sont pas prises en charge par des processus de réparation, elles peuvent perturber les mécanismes nucléaires de transcription ou de réparation, et ainsi favoriser le développement de cancers. Afin de préserver l'intégrité du génome, la cellule possède au moins 4 mécanismes de réparation, qui impliquent plus de 130 protéines chez l'homme (Hoeijmakers, 2001). Ces divers processus sont : la réparation des mésappariements (MMR), qui corrige les erreurs faites par les ADN polymérases lors de la réplication ; la recombinaison homologue (HR) et la ligation des extrémités non homologues (NHEJ), qui prennent en charge les cassures double brin ; le processus d'excision de base (BER), qui élimine les bases chimiquement modifiées et les cassures simple brin ; la réparation par excision de nucléotides (NER), qui élimine la majorité des dommages induits par les UV ou les adduits volumineux (Figure 1).

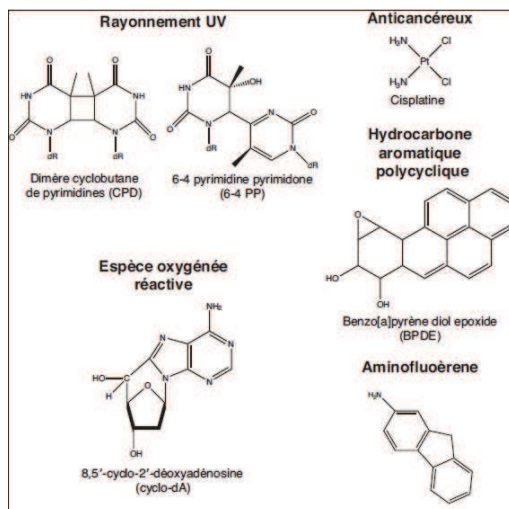


**Figure 1** Les différents mécanismes de réparations. Différents agents génotoxiques peuvent engendrer plusieurs types de lésions, qui seront pris en charges par différents mécanismes de réparation (adapté de (Hoeijmakers, 2001)).

## 1. La réparation par excision de nucléotides (NER)

### A. Les dommages

La NER est un mécanisme de réparation, qui va prendre en charge une grande variété de dommages de l'ADN. En effet, les agents génotoxiques activant la NER peuvent aller des rayonnements UV aux composés anticancéreux (Gillet and Scharer, 2006). Concernant les lésions causées par les rayonnements UV, 70% d'entre elles sont des dimères cyclobutanes de pyrimidines (CPD) formées par une lésion covalente entre 2 thymines, 2 cytosines ou une thymine et une cytosine en fonction de la longueur d'onde. Les 30% restantes sont composées par les photoproduits 6-4 pyrimidine pyrimidone (6-4 PP) (Figure 2). Ces lésions provoquent une distorsion de l'ADN, qui sera reconnue par la machinerie de réparation. Les CPDs induisent une légère courbure de l'ADN ( $\sim 15^\circ$ ), avec quasiment aucune altération de l'appariement entre bases (Lee et al., 2004). Au contraire, les 6-4PPs produisent une distorsion plus prononcée du squelette de l'ADN (avec une courbure allant jusqu'à  $44^\circ$ ), provoquant une rupture de l'appariement entre bases (Kim et al., 1995).



**Figure 2** Lésions réparées par la NER. Les lésions induites par le rayonnement UV (tels que les CPDs et les 6-4 PPs), par le cisplatine, par des composés aromatiques polycycliques, par le stress oxydatif et les aminofluorènes modifient la structure de l'ADN et sont prises en charge par le mécanisme de réparation par NER.

La NER va également prendre en charge et éliminer des composés chimiques, capables de se fixer à l'ADN et d'en modifier la conformation. Ces composés peuvent être des anticancéreux, comme le cisplatine ou ses dérivés, des hydrocarbures polycycliques aromatiques, comme les Benzo[a]pyrenes que l'on retrouve dans les gaz d'échappement, la fumée de cigarette ou encore des amines aromatiques (aminofluorènes) (Figure 2). Le mécanisme de NER est aussi impliqué dans l'élimination de dommages oxydatifs tels que les

8,5'-cyclopurine-2'-deoxyadenosines (cyclo-dA) ou 8,5'-cyclopurine-2'-deoxyguanosines (cyclo-dG) (Mocquet et al., 2007). Ces derniers sont formés suite à l'attaque spécifique d'un radical hydroxyle, ciblant la position 2' d'une déoxyadénosine et d'une déoxyguanosine.

Compte tenu de la grande diversité de lésions réparées par la NER, il semble que ce ne soit pas le dommage en temps que tel, qui soit reconnu. Il s'agit plutôt de certaines caractéristiques conformationnelles de l'ADN, causées par la présence du dommage. L'état de distorsion de l'hélice jouerait notamment un rôle primordial, puisqu'il est proportionnel au taux de réparation (Hess et al., 1997). Ceci expliquerait pourquoi les 6-4PP sont réparés 5 fois plus vite que les CPD, pourtant plus nombreux mais moins distordant (Reardon and Sancar, 2003). Qu'ils soient des adduits volumineux s'intercalant entre les bases azotées (comme les Benzo[a]pyrènes) ou des liaisons intra-brins causées par les UV et/ou les produits chimiques, tous les dommages provoquant d'importantes distorsions de l'ADN sont reconnus puis réparés par la NER.

## B. Le mécanisme de NER

Le processus de NER se divise en 2 voies : la réparation globale du génome (GG-NER) qui prend en charge les dommages présents sur l'intégralité du génome (Figure 3) et la réparation couplée à la transcription (TC-NER), ne s'occupant que des lésions présentes sur le brin transcrit des gènes codants (Figure 4). Ces deux mécanismes se divisent en 5 étapes : la reconnaissance de la lésion, l'ouverture de l'ADN, la double incision, la resynthèse du brin puis la ligation du brin néo-synthétisé. Ces mécanismes diffèrent principalement par leurs étapes de reconnaissance de la lésion. En effet, la TC-NER implique la présence d'une ARN polymérase bloquée face à un dommage.

### B.1 Le mécanisme de GG-NER

La GG-NER est un processus se divisant en 5 étapes (Figure 3).

- **Reconnaissance de la lésion** : La reconnaissance de la lésion s'effectue dans la GG-NER par la protéine XPC, qui est capable de détecter plusieurs types de lésions. Dans la cellule, XPC forme un hétéro-trimère composé de XPC, HR23B (Sugasawa et al., 1996) et de la centrine 2



(Masutani et al., 1994). Ce complexe est essentiel pour le recrutement des autres composants de la machinerie NER, pour former le complexe de pré-incision de l'ADN (Araujo et al., 2000). La centrine 2 et HR23B stabilisent le complexe XPC, ils inhibent la poly-ubiquitination de XPC et la protègent de la dégradation par le protéasome (Nishi et al., 2005).

Le mécanisme de reconnaissance des sites lésés par XPC n'est toujours pas bien définie : il semblerait que le complexe XPC reconnaisse une structure secondaire de l'ADN, et non la lésion en elle-même (Sugasawa et al., 2001) (Sugasawa et al., 2002). Effectivement XPC pourrait scanner l'ADN, afin d'en détecter les distorsions (Hoogstraten et al., 2008). Lorsqu'elle rencontre une lésion, il y aurait alors un changement de conformation de la structure de XPC, rendant possible l'interaction avec les autres facteurs NER.

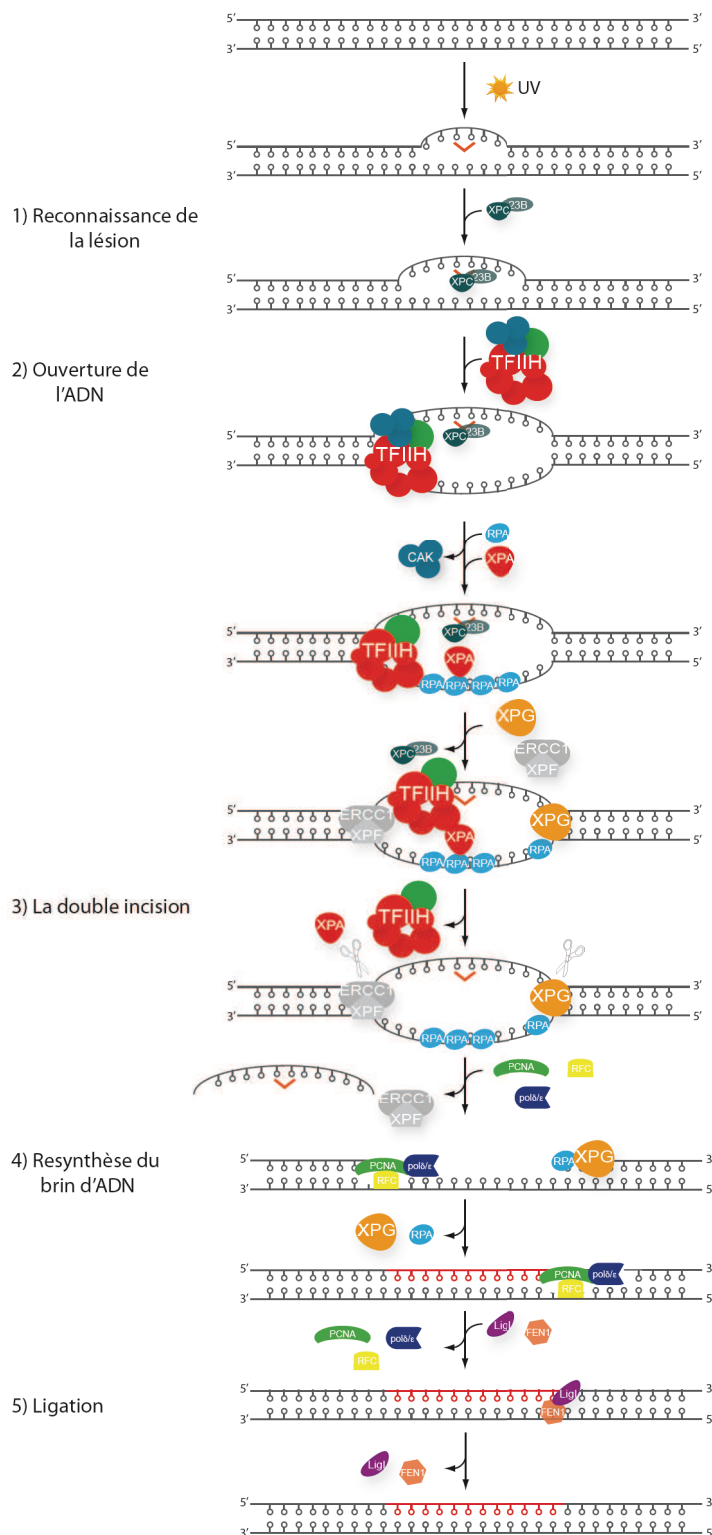
L'affinité de XPC est variable en fonction du type de lésion. En effet, les 6-4 PP qui provoquent une forte distorsion de l'ADN, seront facilement détectés par XPC, contrairement aux CPD. Des études récentes ont montré qu'un complexe appelé UV-DDB, impliqué dans la NER, favoriserait la reconnaissance des lésions de types CPD (Fitch et al., 2003). Ce complexe UV-DDB est un hétéro-dimère, composé des protéines DDB1 et DDB2 (Keeney et al., 1993). Il reconnaît les lésions CPD du brin non-transcrit (Scrima et al., 2008). La protéine DDB1 peut également interagir avec la protéine cullin CUL4A, afin de former un complexe E3 ligase, capable d'ubiquitiner de nombreuses protéines cibles (Higa et al., 2006). Les lésions CPD du brin non-transcrit sont reconnues par DDB2, qui va recruter le complexe E3 ligase DDB1-CUL4A. Ce complexe DDB1-CUL4A-DDB2 nouvellement formé, peut alors ubiquitiner XPC et auto-ubiquitiner DDB2 (Takedachi et al., 2010). L'ubiquitination de DDB2 réduirait son affinité pour le dommage et provoquerait sa dégradation par le protéasome (Sugasawa et al., 2005). L'ubiquitination de XPC stabiliserait son interaction avec HR23B et augmenterait l'affinité de XPC pour l'ADN lésé (Ng et al., 2003), permettant ainsi la mise en place de la réparation de l'ADN. Il semblerait également que le complexe DDB1-CUL4A-DDB2 joue un rôle dans les phénomènes de remodelage de la chromatine, notamment en interagissant avec l'histone H2A (Kapetanaki et al., 2006) (Yeh et al., 2012).

- **Ouverture de l'ADN** : Une fois la lésion détectée par le complexe XPC, ce dernier ouvre partiellement la double hélice et sert de plateforme de recrutement pour TFIIH ((Volker et

al., 2001) (Figure 3). En effet, la protéine XPC présente un domaine de liaison au complexe TFIIH (Sugasawa, 2008), lui permettant d'être recruté sur le dommage.

TFIIH est un facteur multifonctionnel, décrit dans un premier temps pour son rôle dans la transcription. Il est composé du « cœur » (XPB, p63, p52, p44, p34 et P8/TTD-A) et du « CAK » (*cdk-activating kinase*) (CDK7, MAT1, cyclin H). Ces deux sous-complexes sont maintenus ensemble par la sous-unité XPD (Roy et al., 1994) (Compe and Egly, 2012a). TFIIH peut ouvrir l'ADN par l'intermédiaire de ces deux hélicases ATP-dépendantes XPB et XPD. L'activité ATPase de XPB permet l'ancrage de TFIIH sur l'ADN et l'activité hélicase 5'-3' de XPD provoque l'ouverture de la double hélice d'ADN (Coin et al., 2007) (Oksenych et al., 2009). Suite à cette ouverture, la protéine RPA est recrutée sur l'ADN simple brin. Elle protège le brin non endommagé de potentielles attaques d'enzymes nucléasiques. Cela permet le maintien de la configuration ouverte de la double hélice d'ADN. On observe par la suite, le recrutement de XPA qui est favorisé par la sous-unité p8 de TFIIH. Ce recrutement provoque la dissociation du CAK de TFIIH (Coin et al., 2008) et sera suivi par l'arrivée du facteur d'excision XPG. Le complexe de pré-incision composé de RPA, XPA et XPG est alors assemblé autour du dommage (Zotter et al., 2006).

- **La double incision** : Elle commence par le recrutement de la 3' endonucléase XPG au niveau de la lésion, qui provoque le relargage de XPC (Riedl et al., 2003). Cette étape est suivie du recrutement de la 5' endonucléase XPF/ERCC1, via une interaction possible avec XPA (Orelli et al., 2010). XPA et RPA positionneraient et stimuleraient l'activité des 2 endonucléases, alors que la mise en place de XPF/ERCC1 induirait le relargage de TFIIH (Riedl et al., 2003). L'action conjointe de XPG et de XPF/ERCC1 va mener à l'incision d'un fragment de 24 à 32 nucléotides contenant le dommage (Hess et al., 1997). L'endonucléase XPG est responsable de l'incision du brin endommagé en 3' et est nécessaire pour l'incision en 5' par XPF/ERCC1 (Wakasugi et al., 1997). Néanmoins, il semblerait que l'incision en 5' par XPF/ERCC1 précède l'incision en 3' par XPG (Staresinic et al., 2009). Après cette double incision, le facteur XPA est relargué de l'ADN pour être recyclé dans un nouveau complexe d'incision. Le facteur RPA est alors le seul facteur du complexe de pré-incision avec les deux endonucléases à rester sur l'ADN, afin d'assurer la protection du brin non-endommagé d'attaques nucléasiques (Riedl et al., 2003) (Overmeer et al., 2011).



**Figure 3** Schéma du mécanisme de la réparation globale du génome par excision de nucléotides (GG-NER). (1) Reconnaissance de la lésion par le complexe XPC et recrutement de TFIIH. (2) TFIIH ouvre l'ADN autour de la lésion, cette ouverture sera maintenue par XPA et RPA. (3) Double incision de l'ADN en 5' par XPF/ERCC1 et en 3' par XPG. (4) Après l'élimination de la lésion, la machinerie de réplication re-synthétise le brin d'ADN. (5) Ligation de l'ADN par l'ADN ligase I et FEN1. (Adapté de (Mocquet et al., 2007)).

- **Resynthèse de l'ADN** : Le départ du fragment endommagé fait place au recrutement de la machinerie de réplication, afin de re-synthétiser un nouvel ADN dans la région simple brin (Mocquet et al., 2008) (Figure 3). Cette étape doit être étroitement coordonnée avec la double incision de XPG et XPF/ERCC1, afin d'éviter la multiplication des régions simples brins potentiellement mutagènes. Il semblerait que certains facteurs de la machinerie de réplication soient recrutés sur le brin d'ADN avant l'exclusion de l'oligonucléotide endommagé. L'incision en 5' par XPF/ERCC1 doit avoir eu lieu pour pouvoir permettre le recrutement de PCNA via XPG et de RFC par l'intermédiaire de RPA. Cette étape permettrait ainsi le relargage de l'endonucléase XPF/ERCC1. Les facteurs XPG, RPA, RFC et PCNA servent de plateforme pour le recrutement des ADN polymérase (ADN pol  $\delta$ ,  $\epsilon$  ou  $\kappa$ ) (Aboussekhra et al., 1995). D'après une étude récente, le recrutement de la pol  $\kappa$  requière l'ubiquitination de PCNA et de la protéine XRCC1 (*X-ray repair cross-complementing protein 1*), intervenant dans la ligation. Le recrutement de la pol  $\delta$  nécessite la présence des protéines RFC et PCNA, alors que celui de la pol  $\epsilon$  a besoin d'une variante de RFC (Ogi et al., 2010). L'exclusion de XPG et RPA initie alors la phase de resynthèse par l'ADN polymérase, qui utilise le brin non endommagé comme matrice.

- **La ligation** : La dernière étape consiste à lier le brin néosynthétisé et l'ADN préexistant. Des études récentes s'accordent à dire que c'est l'ADN ligase I, connue pour son rôle lors de la réplication de l'ADN, qui est recrutée au niveau du brin néosynthétisé en compagnie de FEN1. Les ADN ligases III et IV sont connues pour leurs rôles dans les mécanismes de BER et de recombinaison non-homologue. Néanmoins, l'ADN ligase III et son partenaire XRCC1 pourraient jouer un rôle dans la ligation, en étant recrutés avec la polymérase  $\delta$  sur l'ADN lors de la NER *in vivo* (Moser et al., 2007). Une fois la ligation effectuée, FEN1 et l'ADN ligase I sont relargués et des facteurs de remodelage de la chromatine, tels que CAF-1, vont être recrutés pour modifier l'environnement chromatinien.

## B.2 Le mécanisme de TC-NER

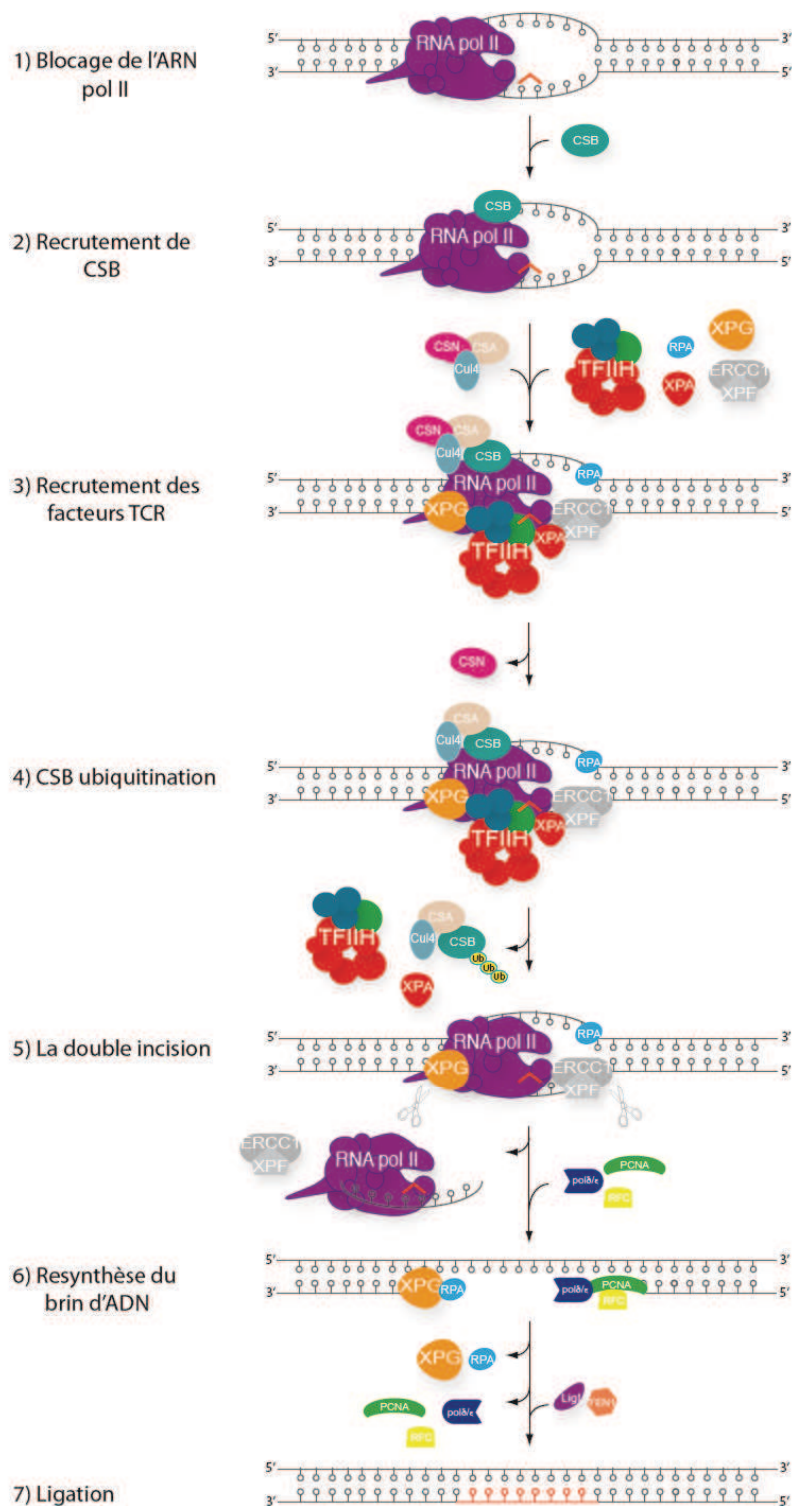
La réparation par excision de nucléotides couplée à la transcription (TC-NER) est un mécanisme de réparation rapide, qui va spécifiquement réparer les lésions d'ADN présentes sur le brin transcrit. Ces lésions ont pour effet de bloquer la progression de l'ARN

polymérase II. Ce mécanisme diffère de la GG-NER par son processus de reconnaissance de la lésion au niveau de la chromatine. En effet, lors de la transcription, la machinerie d'élongation peut être bloquée par des dommages au niveau du brin transcrit, ayant pour conséquence la perturbation du développement, de la croissance ou de la survie cellulaire. La TC-NER, décrite dans un premier temps par Hanawalt et son équipe (Bohr et al., 1985) (Mellon et al., 1987), est un mécanisme fortement conservé au cours de l'évolution, que l'on retrouve chez une grande variété d'organismes vivants comme les bactéries, les levures et les mammifères.

L'arrêt de l'ARN polymérase II permet de localiser le dommage et donc de déclencher le mécanisme de réparation TC-NER (Mellon et al., 1986). Les lésions bloquant l'ARN polymérase II peuvent être de différents types, tels que les CPD ou les 6-4PP produits par les UV-C, les composés anticancéreux comme le cisplatine (Zhen et al., 1993) ou encore certains adduits chimiques (Figure 2). Lorsque l'ARN pol II rencontre une lésion, plusieurs hypothèses concernant son comportement ont été décrites. En effet, il a été démontré que l'ARN pol II peut être poly-ubiquitinée et dégradée par le protéasome (Selby and Sancar, 1994) ou encore arrêtée au niveau de la lésion pendant la réparation (Treméau-Bravard et al., 2004) (Laine and Egly, 2006).

Le processus de TC-NER est enclenché par l'arrêt de l'ARN pol II et nécessite le recrutement de facteurs spécifiques à la TCR tels que CSB et CSA. Le rôle de CSB a été démontré *in vitro* et *in vivo*. Il s'agit d'un facteur clé, essentiel pour l'assemblage fonctionnel du complexe ARN pol II/TCR facteurs (Fousteri et al., 2006) (Laine and Egly, 2006). La formation du complexe ARN pol II/CSB permet alors le recrutement des facteurs TFIIH, RPA, XPA et des 2 endonucléases XPG et XPF/ERCC1 au niveau du dommage. Ce complexe de pré-incision alors formé, ouvre l'ADN autour de la lésion, pour permettre la double incision.

CSB est un facteur faisant partie de la famille des facteurs SWI/SNF, remodelant la chromatine de façon ATP dépendante (Troelstra et al., 1992). Il présente *in vitro*, une activité de remodelage de la chromatine, en interagissant directement avec le cœur des histones (Citterio et al., 1998). De plus, lors de la transcription de gènes par l'ARN pol I, il a été démontré que CSB interagissait avec la méthyltransférase G9a afin de permettre la méthylation de H3K9 (Yuan et al., 2007).



**Figure 4** Schéma de la réparation par excision de nucléotide couplée à la transcription (TC-NER). (1) Arrêt de l'ARN polymérase II par le dommage. (2) Suivi par le recrutement de la protéine CSB. (3) CSB permet le recrutement des facteurs spécifiques de la TC-NER dont le complexe CSA/Cul4/CSN et les facteurs NER. Ce la forme le complexe de pré-incision. (4) CSB est ubiquitiné et relargé. (5) Double incision du brin endommagé par XPG et XPF/ERCC1. (6-7) Resynthèse et ligation du brin transcrit.

CSB interviendrait dans l'initiation de la transcription de certains gènes après irradiation aux UV (Proietti-De-Santis et al., 2006). Il est également nécessaire pour le recrutement de l'acétyl-transférase p300 et de la protéine CSA au niveau du dommage.

CSA fait partie d'un complexe E3 ubiquitine ligase contenant les protéines cul4, DDB1 et Roc1, régulé par CSN (*COP9 signalosome complex*) (Groisman et al., 2003). Le complexe CSA, en interagissant avec CSN, n'a pas d'activité ubiquitine ligase propre. Ce n'est qu'après le relargage de CSN, que CSA pourra ubiquitiner CSB, afin de permettre sa dégradation et le retour de la synthèse d'ARN après réparation (Groisman et al., 2006). Le complexe CSA permet le recrutement de 3 autres facteurs additionnels de la TC-NER : la protéine d'interaction nucléosomale HMG1, le facteur de transcription TFIIIS et le facteur de splicing XAB2 (XPA binding protein 2) (Fousteri et al., 2006). Il a été suggéré que le recrutement de p300 et de HMG1 faciliterait le remodelage de la chromatine et permettrait la translocation de l'ARN pol II pour favoriser l'élimination du dommage et le redémarrage de la transcription (Hanawalt and Spivak, 2008). TFIIIS stimulerait l'activité de clivage de l'ARN pol II dans la TC-NER, afin de provoquer son arrêt. Cela permettrait le redémarrage de l'élongation après élimination de la lésion.

Une fois la lésion reconnue et le complexe de pré-incision formé, CSB est ubiquitiné et relargué avec XPA et TFIIH. La double incision peut alors être effectuée. 3h après la génération de dommages par les UV, CSB sera totalement dégradé (Groisman et al., 2006). Après la double incision par XPF/ERCC1 en 5' et XPG en 3' et le relargage du brin endommagé, PCNA est recrutée sur l'ADN pour former avec RFC et les DNA polymérases ( $\delta, \epsilon, \kappa$ ) le complexe de resynthèse. Finalement la ligation est effectuée par l'ADN ligase I et FEN1, permettant ainsi le redémarrage de la transcription.

## 2. Les autres mécanismes de réparation de l'ADN

### A. La BER

La BER (*base excision repair*) est le principal mécanisme de réparation des bases endommagées ou modifiées, induisant de faibles changements structuraux de l'ADN (Figure 1). Ces lésions peuvent être de différents types, tels que les dommages oxydatifs (Lindahl, 1993), les déaminations des bases (Kavli et al., 2007) ou l'alkylation des bases. Les origines de ces lésions sont multiples : elles peuvent être à la fois endogènes, en provenant du

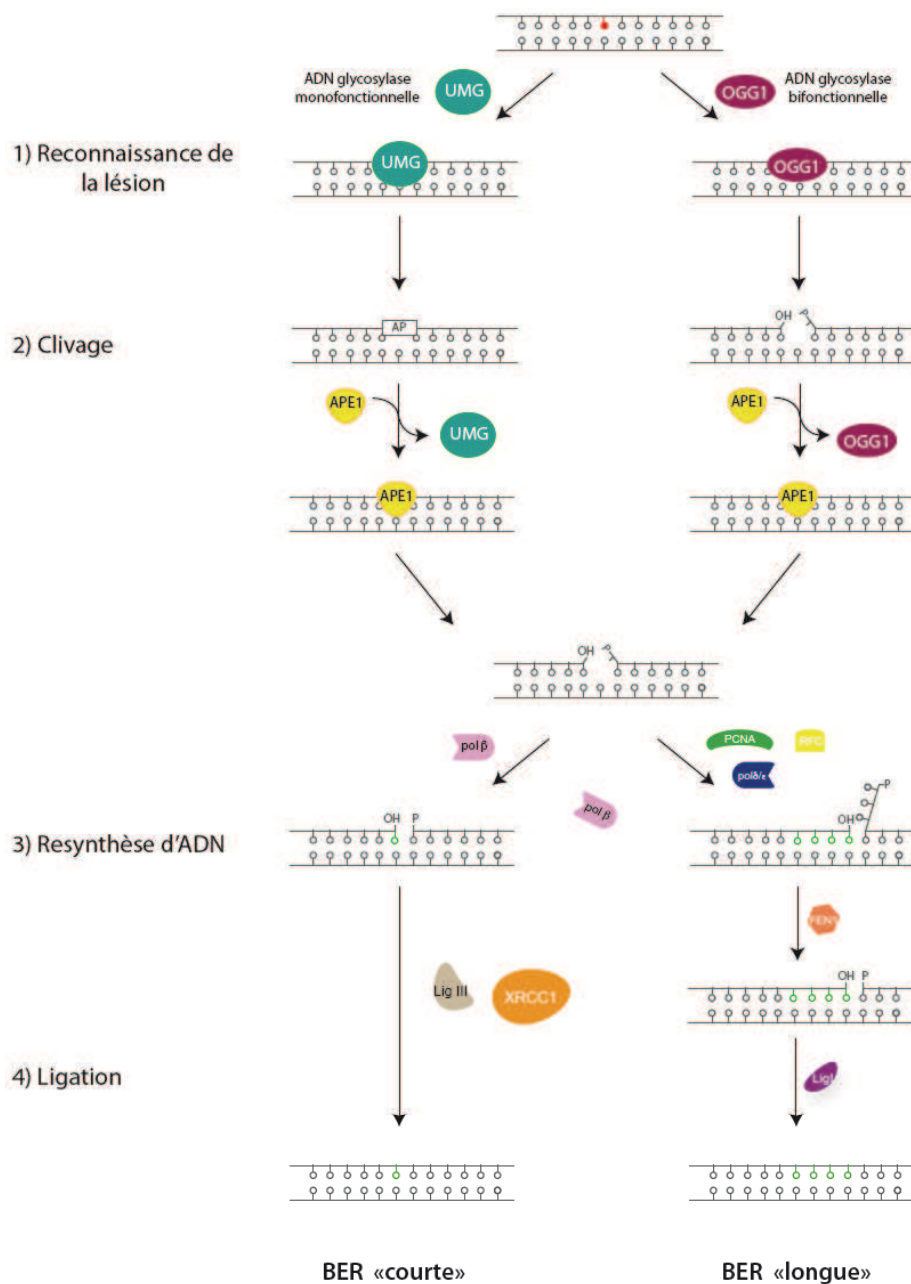


métabolisme cellulaire ou exogènes dues à l'environnement. Une grande partie des lésions endogènes sont générées par les ROS (*reactive oxygen species*) provenant du processus de respiration mitochondriale (Zorov et al., 2006). Néanmoins des ROS exogènes peuvent provenir de différents stimuli, tels que la fumée de cigarettes, les UV-A ou les radiations ionisantes. Ces agents induisent une modification des bases, tels que l'apparition de la 8-hydroxyguanine (8 oxoG). Ces lésions sont faiblement cytotoxiques pour la cellule, mais peuvent avoir un effet mutagène (Neeley and Essigmann, 2006). L'oxydation de la 2'déoxyadénosine peut également générer 2 autres types de lésions très répandues : le 8 oxodA (8-hydroxyadénosine) et le FapydA (4,6 diamino-5-formamidopyrimidine), qui sont également mutagènes. Etant donné la diversité de lésions, différents mécanismes de BER ont été mis en place par la cellule. Ils diffèrent par leur étape de reconnaissance de la lésion et présentent ensuite le même modèle, avec l'élimination de la base modifiée, le clivage du déoxyribose, la resynthèse de l'ADN suivi de la ligation (Figure 5).

L'étape de reconnaissance sera effectuée par une ADN glycosylase, qui va reconnaître spécifiquement la lésion. 11 ADN glycosylases ont été identifiées chez l'homme, leur fonction primaire est de reconnaître la lésion mais également de couper la liaison N-glycosidique, relarguant ainsi une base libre et créant un site abasique (AP) (Lindahl, 1974).

Deux types d'ADN glycosylases ont été décrites : les monofonctionnelles et les bifonctionnelles. Ces dernières sont capables, en plus de leur activité de base, d'hydrolyser le lien phosphodiester en 3' du site AP, via une activité AP lyase. Par exemple, l'ADN glycosylase UNG (*uracil DNA glycosylase*) est la première glycosylase monofonctionnelle décrite dans la littérature, alors que OGG1 (*8 oxoguanine DNA glycosylase*) est une glycosylase bi-fonctionnelle fréquemment étudiée (Klungland and Bjelland, 2007). Quelque soit la glycosylase, l'AP endonucléase APE1 est ensuite recrutée sur le site abasique de l'ADN, en 5' de la lésion, provoquant l'élimination de l'ADN glycosylase. Elle clive le squelette sucre-phosphate, produisant une extrémité 3' OH. L'étape suivante consiste à re-synthétiser la base manquante, cela est effectué par l'ADN polymérase  $\beta$ , qui remplace le nucléotide manquant par le nucléotide correct (Matsumoto and Kim, 1995). Finalement le complexe XRCC1/Ligase III complète le processus en assurant la liaison phosphodiester des deux nucléotides adjacents, permettant ainsi le retour d'un brin d'ADN non endommagé (Wei et al., 1995).





**Figure 5** Schéma du mécanisme de réparation par excision de bases (BER). (1) Reconnaissance de la lésion par l'ADN glycosylase mono ou bifonctionnelle. (2) Clivage de la liaison N-glycosidique générant un site abasique AP. Recrutement de l'AP endonucléase APE1 afin de générer une extrémité 3' OH. (3) Etape de resynthèse via les ADN polymérase  $\beta$  ou  $\delta/\epsilon$  et ligation (4) par les ligases III ou I.

XRCC1 a été identifié comme un acteur majeur de la voie BER courte. En interagissant avec la pol  $\beta$  et la ligase III, il permet une ligation efficace (Caldecott et al., 1994). Ce mécanisme de remplacement d'une seule base (short patch BER) est mis en place dans 80-90 % des cas (Bennett et al., 1997).

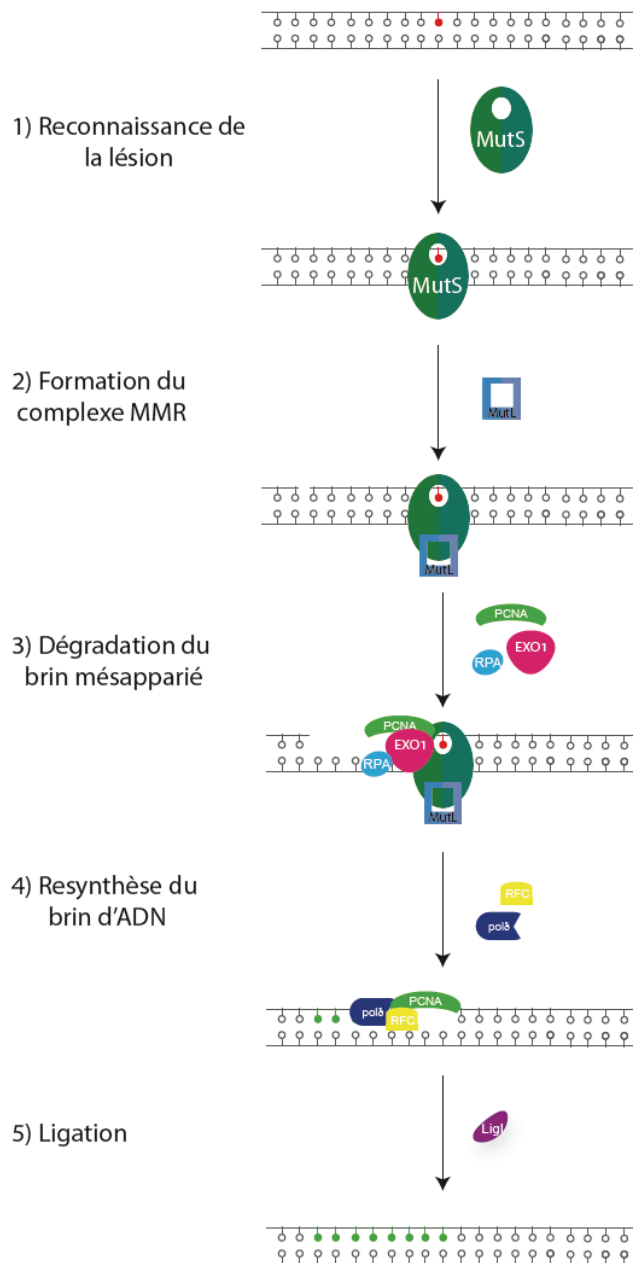
Néanmoins, un deuxième mécanisme de resynthèse/ligation peut être mis en place par la cellule. Celui-ci est plus long et nécessite l'intervention de plus de protéines. Le choix entre ces 2 mécanismes par la cellule est encore peu documenté ; il semblerait que la forte concentration en ATP, au niveau du site abasique, favorise la BER courte (Petermann et al., 2003). La BER longue nécessite le recrutement abondant de la protéine PCNA, qui s'associe au recrutement de RFC et de l'ADN polymérase  $\delta/\epsilon$ . Ce complexe remplace alors 2 à 6 nucléotides, en déplaçant le brin d'ADN. La protéine FEN1 est ensuite recrutée via son interaction avec PCNA et élimine le brin déplacé. La ligation peut s'effectuer par l'intermédiaire de l'ADN ligase I (Tomkinson et al., 2001).

### **B. La réparation des mésappariement (MMR)**

Lors de la réplication de l'ADN, l'ADN polymérase, responsable de la copie des nucléotides, peut commettre des erreurs telles que des mauvais appariements de bases, ou encore des insertions ou délétions de nucléotides. On dénombre environ 1 erreur tous les  $10^7$  nucléotides synthétisés. Ce mécanisme a été reconstitué *in vitro* (Lahue et al., 1989) et sa compréhension chez *E.Coli*, permet de mieux expliquer le mécanisme de la MMR chez les eucaryotes.

Chez *E.Coli*, le mauvais appariement de base est reconnu par l'homodimère MutS, de façon ATP dépendante (Modrich and Lahue, 1996). En parallèle, un homodimère de MutL va interagir avec MutS, ce qui accentue la reconnaissance de la lésion par ce dernier. MutL permet le recrutement et l'activation de l'endonucléase MutH, afin de former un complexe MMR fonctionnel (Sancar and Hearst, 1993). MutH interagit ensuite avec un site GATC hémiméthylé au niveau du brin mal apparié et va l'inciser (Junop et al., 2001). Cette coupure permet le recrutement de l'exonucléase Exo1 (3'-5' chez *E.Coli*) afin de dégrader le brin endommagé. Pour finir, l'étape de resynthèse est effectuée par l'ADN polymérase III.

Chez les eucaryotes, le processus de MMR fait intervenir des homologues de MutS et MutL de *E.Coli* (Figure 6). Trois homologues de MutS sont impliqués dans la détection des mésappariements : MSH2, MSH3 et MSH6. Ces protéines forment 2 hétérodimères appelés MutSa (MSH2/MSH6) et MutSb (MSH2/MSH3).



**Figure 6** Schéma du mécanisme de réparation des mésappariements (MMR). (1) Reconnaissance du mésappariement par l'hétérodimère MutS. (2) Le recrutement de MutL permet la formation du complexe MMR. (3) Interaction de PCNA avec le complexe MMR et recrutement de Exo1, qui dégrade le brin endommagé. (4) Resynthèse du brin d'ADN par l'ADN polymérase  $\delta$ . (5) Ligation de l'ADN par la ligase I.

MutSa reconnaît les mésappariements et les insertions/délétions d'1 ou 2 nucléotides alors que MutSb interagit avec les insertions/délétions plus larges (jusqu'à 16 nucléotides). Une fois l'erreur détectée par les hétérodimères MSH, le complexe MMR se forme avec le recrutement d'un hétérodimère MSL (homologue de MutL). Trois hétérodimères de MSL ont été décrits : MutLa (MLH1/PMH2), MutLb (MLH1/PMS1) et MutLg (MLH1/MLH3).

Concernant la protéine MthH, aucun homologue n'a été décrit chez les eucaryotes, ainsi qu'aucun site GATC hémiméthylé. Après la reconnaissance du dommage, le facteur PCNA est recruté sur le brin d'ADN, via son interaction avec MLH et MSH. Cela permet l'incision du brin en 5' du dommage, par MutL. En effet, des études récentes montrent que MutLa possède une activité endonucléase dépendante de PCNA/RFC (Kadyrov et al., 2006).

L'exonucléase 5'-3' Exo1 est ensuite recrutée, afin de dégrader le brin jusqu'à la lésion, tandis que l'autre brin est protégé par RPA (Genschel and Modrich, 2003). PCNA ré-initie la resynthèse en recrutant l'ADN polymérase  $\delta$  et la ligase I. L'altération des gènes de la MMR est à l'origine du cancer colorectal familial non associé à une polypose colique, du syndrome HNPCC (*Hereditary non Polyposis Colorectal Cancer*), ou encore du syndrome de Lynch qui représente une des premières cause de cancer colorectal héréditaire touchant exclusivement l'adulte (Peltomaki, 2003).

### C. La réparation des cassures double brin

Les cassures doubles brins (DSB : *double strand breaks*) sont des lésions létales majeures, induites par des radiations ionisantes ou par des stress cellulaires endogènes ; elles peuvent être sources d'instabilité génétique. Si elles ne sont pas les plus fréquentes, les DSB sont les plus toxiques. En effet, elles peuvent générer de nombreux réarrangements, ayant comme conséquences physiologiques le développement de tumeurs ou encore de malformations congénitales. Deux grands mécanismes entrent en compétition pour la réparation de ces cassures doubles brins : la recombinaison homologue (HR pour *homologous recombination*) et la ligature d'extrémités non homologues (NHEJ pour *non-homologous end joining*). La signalisation du dommage aboutit à l'arrêt transitoire du cycle cellulaire, cette stratégie permet d'éviter la propagation de cellules porteuses de dommages. La recombinaison homologue est qualifiée de mécanisme fidèle, car elle utilise une séquence homologue intacte pour réparer la cassure. Alors que la NHEJ est dite infidèle, car ne nécessite pas d'homologie de séquence et permet la ligature des extrémités d'ADN (Figure 7).

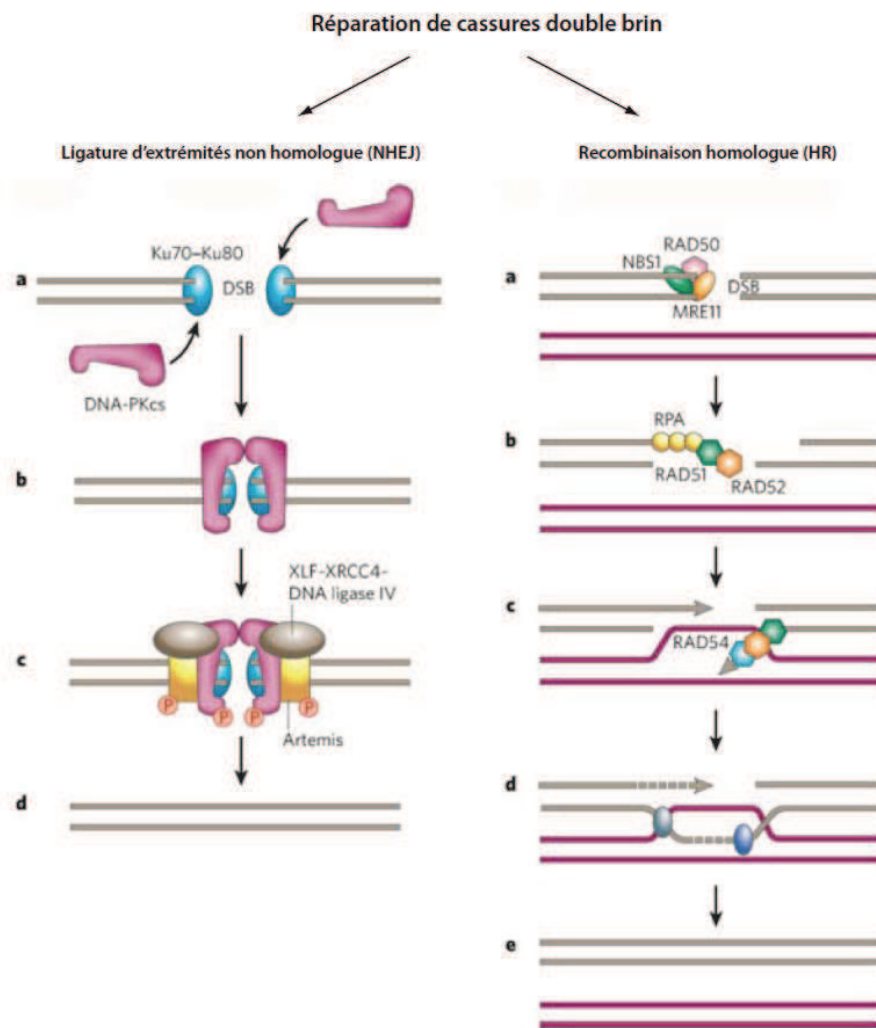
- **La recombinaison homologue (HR)** : La HR est un mécanisme de réparation nécessitant une séquence d'ADN homologue, initiée par une étape de résection de l'ADN par le complexe MRN (Mre11, Rad50, Nsb1). Elle se déroule lors de la phase S ou G2 du cycle

cellulaire. Cela consiste à éliminer l'extrémité de la cassure double brin de quelques dizaines de nucléotides en 5'-3'. Ce complexe MRN interagit avec la protéine CtIP, qui stimule l'activité endonucléasique de Mre11. Cette première étape de résection est suivie par l'action de l'exonucléase Exo1 (Bolderson et al., 2010), qui génère ainsi un ADN simple brin 3'. Cet ADN simple brin permet le recrutement de RPA et de Rad51 à son extrémité. Rad 51 peut alors déplacer RPA avec l'aide de BRCA2 (*breast cancer susceptibility type 2*). BRCA2 va également stimuler et faciliter la fixation de plusieurs protéines Rad51 sur l'extrémité 3' (Jensen et al., 2010).

Rad51 forme un nucléofilament en recrutant les protéines Rad54 et Rad52. Cela permet l'invasion du brin lésé au niveau de la séquence homologue intacte, formant ainsi une structure appelée la « D loop ». Il a été démontré, que Rad54 peut permettre la mise en place de l'étape de resynthèse, en éliminant Rad51 et en permettant le recrutement des ADN polymérases (Heyer et al., 2006). Les ADN polymérases vont ensuite synthétiser un nouvel ADN à partir des extrémités 3'.

La formation de 2 croisements d'ADN permet par l'intermédiaire de jonction hollidays, de restaurer deux molécules d'ADN intactes, avant la ligation par la ligase I (Hoeijmakers, 2001).

- **La ligature d'extrémités non homologues (NHEJ)** : La NHEJ est un processus ne nécessitant pas de séquences homologues pour être réalisé. Les extrémités d'ADN de la coupure double brin sont reconnues par un hétérodimère Ku70/Ku80 (Walker et al., 2001). Le Ku-ADN complexe recrute ensuite la sous-unité catalytique de la DNA-PK, afin de former un hétérotrimère Ku70/Ku80/DNA-PK (Uematsu et al., 2007). La DNA-PK peut alors s'autophosphoryler, induisant un changement de conformation laissant les extrémités libres d'accès pour l'enzyme de maturation Artemis. Artemis permet le clivage des extrémités en formant une structure *hairpin* (Weterings et al., 2009). Un deuxième complexe est ensuite recruté par le complexe Ku/DNA-PK/ADN, composé de la ligase IV et de XRCC4, afin d'effectuer la ligature des deux brins. Récemment, un nouveau facteur de la NHEJ a été identifié : il s'agit de la protéine XLF. Elle interagirait avec le complexe ligase IV/XRCC4 et stimulerait la capacité de ligation de la ligase (Yano and Chen, 2008). Une fois la cassure réparée, le complexe Ligase IV/XRCC4 et la DNA-PK se détachent de l'ADN et le complexe Ku migre le long de la double hélice.



**Figure 7** Schéma des mécanismes de réparation des cassures double brin par ligature d'extrémités non homologues (NHEJ) ou par recombinaison homologue (HR).

## II. Les maladies liées à un défaut de la NER

Des mutations dans les gènes codant pour les protéines jouant un rôle dans le mécanisme de réparation par excision de nucléotides (NER), peuvent être à l'origine de 3 maladies génétiques autosomales rares : le *Xeroderma Pigmentosum* (XP), le syndrome de Cockayne (CS) et la Trichothiodystrophie (TTD). Les mutations de 13 gènes NER sont associées à 11 syndromes cliniques bien distincts (Figure 8). Ces 3 maladies sont divisées en au moins 8 groupes de complémentations <sup>1</sup> pour les patients XP (XP-A à XP-G et XP-V), 2 groupes pour le syndrome de Cockayne (CS-A et CS-B) et 4 groupes pour les patients atteints

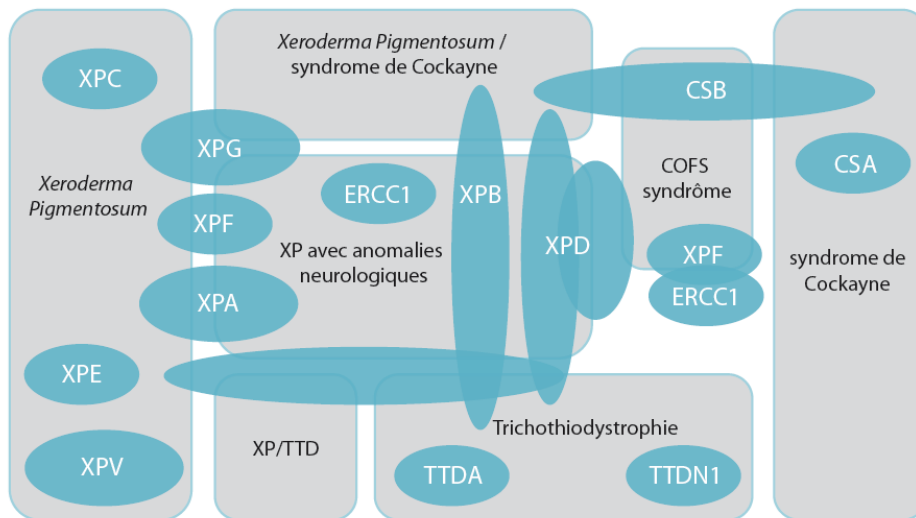
de Trichothiodystrophie (TTD-A, XP-B, XP-D et TTDN1). Chacun de ces groupes traduit une ou plusieurs mutations dans le gène correspondant. Il est toutefois notable, que des mutations dans un même gène, par exemple XPB, XPD ou XPG, peuvent être à l'origine de maladies différentes. En effet, des mutations dans XPB et XPD peuvent générer des patients XP, XP/CS ou TTD (Kraemer et al., 2007).

### 1. Xeroderma Pigmentosum (XP)

Le *Xeroderma Pigmentosum* est une maladie génétique autosomale récessive dont l'incidence est variable en fonction de l'origine géographique. En effet elle varie entre 1/20000 pour le Japon et l'Afrique du nord, et 1/1 000 000 aux USA et en Europe de l'ouest (Kleijer et al., 2008). Les patients XP sont caractérisés par une peau à l'aspect parcheminé « *xeroderma* » et présentant une pigmentation anormale « *pigmentosum* ». Ces patients présentent une forte sensibilité aux UV, induisant une hyperpigmentation de la peau et sont caractérisés par une augmentation des risques de cancers de la peau (2000 fois supérieur pour les cancers à mélanomes et 10 000 fois supérieur pour les cancers non-mélanomiens avant l'âge de 20 ans) (Bradford et al., 2011). Les premières manifestations apparaissent avant l'âge de 2 ans avec une forte augmentation du nombre de lentigines (taches brunes) sur les zones cutanées exposées au soleil (Kraemer et al., 1994). Il a également été montré, que ces patients peuvent développer des problèmes oculaires strictement localisés sur les parties antérieures de l'œil (paupières, cornées, conjonctives) exposées aux UV (Ramkumar et al., 2011) (Dollfus et al., 2003). On note également pour 20-30% des cas des troubles neurologiques et des déficiences intellectuelles (Bradford et al., 2011) (Tableau 1). L'espérance de vie moyenne des patients XP est de 32 ans.

---

<sup>1</sup> Test de complémentation : Test permettant de déterminer si 2 mutants récessifs sont causés par des mutations sur le même gène ou dans 2 gènes différents. Pour cela, on fusionne des cellules de 2 patients afin d'obtenir une cellule hybride, qui aura en commun les gènes des deux cellules parentales. On vérifie ensuite le phénotype obtenu. Si le phénotype est mutant, cela signifie qu'il n'y a pas eu complémentation, donc les 2 allèles pourraient être mutés sur le même gène. Inversement, si il y a eu complémentation, alors la mutation est sur 2 gènes différents (De Weerd-Kastelein et al., 1972).



**Figure 8. Différentes maladies induites par des mutations dans les gènes NER.** Les différentes pathologies sont représentées par des rectangles gris et les gènes mutés impliqués dans les différentes pathologies sont schématisés en bleu (DiGiovanna and Kraemer, 2012).

XP est une pathologie résultant de mutations sur 8 gènes impliqués dans différentes étapes de la NER. Elle peut être classée en 7 groupes de complémentation (XP-A à XP-G) et un groupe supplémentaire dit XP variant (XP-V). La gravité de ces phénotypes varie en fonction de la protéine mutée. En effet, selon la mutation, différentes étapes de la NER seront perturbées, générant ainsi des phénotypes de sévérités variables.

Par exemple, les patients XP-C et XP-E (DDB2) présentent une faible sensibilité aux UV et ne développent pas de troubles neurologiques. Cela s'explique par le fait que XPC et DDB2, ne jouent aucun rôle dans la TC-NER, permettant ainsi à cette voie de réparation de réparer les différentes lésions (Venema et al., 1990). Les individus XP-A présentent des symptômes plus sévères, avec des anomalies neurologiques importantes, s'expliquant sûrement par le rôle de XPA dans la GG-NER et la TC-NER. Concernant les patients XP-F, ils ne développent pas de problèmes neurologiques et sont moins sensibles aux UV. Cela s'explique par le fait qu'ils ne présentent qu'un seul allèle muté pour XPF (Matsumura et al., 1998).

Parfois une même mutation peut donner lieu à plusieurs phénotypes. Par exemple, des mutations de XPG peuvent générer des formes de patients XP ou XP/CS, patients présentant alors les symptômes combinés des XP et des CS (Ito et al., 2007) .



Il existe également un groupe supplémentaire, les patients « variants » (XP-V), n'étant pas spécialement sensibles aux UV. Ils ne développent aucun problèmes neurologiques mais ont une forte susceptibilité à développer des cancers de la peau (Lehmann et al., 1975). Ces patients présentent une mutation dans le gène POLH, codant pour l'ADN polymérase h, qui est capable de répliquer l'ADN sans être arrêtée par une lésion. Cette mutation a pour conséquence un défaut de la réparation post-réplivative. En effet, la translésion des CPD ne s'effectue pas correctement et est alors réalisée par des polymérases moins fidèles (Pol z, k ...) (Masutani et al., 1999).

Dans le cadre de ma thèse, je me suis intéressé plus particulièrement à deux mutations du gène XPC, présentes chez des patients XP. La première est une mutation stop remplaçant une cytosine par une thymine sur la base 1840. Cette mutation génère un codon stop et une protéine XPC tronquée au niveau de l'acide aminé 579 (R579St) (Gozukara et al., 2001). Cette mutation a été détectée chez un enfant de Malatya en Turquie (XP67TMA), qui présentait une hypersensibilité à la lumière et de nombreux cancers de la peau. Ce patient présentait à l'âge d'un an des lésions cutanées anormales, mais ne développait aucun trouble neurologique. Le premier cancer de la peau a été détecté à l'âge de 6 ans, puis il est décédé à l'âge de 10 ans, des complications de nombreux carcinomes. Au niveau moléculaire, les cellules de ce patient (GM14867) présentent un niveau d'expression d'ARNm très faible pour XPC. La seconde mutation est la substitution d'une proline en position 334 en une histidine (P334H) dans un domaine non conservé en N-terminal. Ce patient (XP1MI) est un des rares patients XP-C présentant des troubles neurologiques (autisme et hyperactivité). De plus, les cellules de ce patient (GM02096) présentent une expression normale de la protéine XPC (Rivera-Begeman et al., 2007) (Khan et al., 1998). Au niveau moléculaire, ces deux mutations ont des conséquences différentes. En effet, la mutation R579St provoque une perturbation du mécanisme de réparation par excision de nucléotides, en empêchant le recrutement de XPC au niveau des dommages. Cette mutation perturbe également l'interaction de XPC avec la machinerie de réparation (TFIIH, XPB, HR23B). Les effets de la mutation P334H sur la réparation NER sont moins importants. La protéine XPC peut se lier normalement aux dommages, néanmoins on observe un retard dans le recrutement de XPA et une perturbation de la fixation de TFIIH (lien entre de cette

région de XPC avec p62) (Bernardes de Jesus et al., 2008). De plus, cette mutation prévient la stimulation de la glycosylase Ogg1 dans le mécanisme de BER.

## 2. Le syndrome de Cockayne (CS)

CS est une maladie génétique rare autosomale récessive, résultant de la mutation du gène CSB pour 80% des cas ou du gène CSA pour 20% des cas (Licht et al., 2003). Cette pathologie est caractérisée par des retards de croissance (petite taille), ainsi que des troubles neurologiques. Ces patients présentent également une sensibilité aux UV, mais n'ont pas de prédisposition à développer des cancers de la peau. L'espérance de vie moyenne est de 12,5 ans (Andressoo et al., 2005). Les retards de croissance se caractérisent par une apparence « âgée » ainsi que des défauts squelettiques « tête en forme d'oiseau » associé à des troubles neurologique (retard mental, démyélinisation...) (Tableau 1) et des troubles du développement des organes génitaux (Nance and Berry, 1992). Les troubles neurologiques, importants chez les CS, se caractérisent par une forte perte neuronale, observée principalement dans le cortex cérébrale. Il se caractérise également par une distorsion des cellules de Purkinje, associée à une perte de matière blanche (Koob et al., 2010). Les CS peuvent être classés en fonction de leurs degrés de sévérité. La classe I (CS I) comprend des patients dont les premiers symptômes apparaissent dans la première année de vie et dont l'espérance de vie est inférieure à 20 ans. La classe II (CS II) est composée de quelques rares patients présentant des symptômes plus sévères. Pour la classe III (CS III), les symptômes des patients sont moins importants. La quatrième classe appelée COFS (*cerebro oculo facio skeletal*) constitue une forme prénatale très sévère de CS, caractérisée par une photosensibilité cutanée (Laugel et al., 2010). Les patients CS (CB-B et CS-A) présentent un défaut spécifique de la TCR, permettant l'explication de certains de leurs symptômes. Néanmoins, certains phénotypes CS ne peuvent être compris uniquement par le seul défaut de cette voie de réparation. En effet, des travaux ont montré que CSB était impliqué dans d'autres processus cellulaires comme la transcription, et joue un rôle dans l'initiation de la transcription et le maintien de la chromatine (Newman et al., 2006) (Proietti-De-Santis et al., 2006).

Certains patients appelés XP/CS présentent des symptômes combinés des patients XP et CS, comme des altérations cutanées, des troubles neurologiques, une hypersensibilité aux UV, ainsi qu'une immaturité sexuelle. Ces individus présentent des mutations sur 3 gènes XP : XPB, XPD et XPG (Lindenbaum et al., 2001).

Le syndrome cérébro-oculo-facio-squelettique (COFS) est une maladie génétique rare, caractérisée par une atteinte neurosensorielle sévère. A ce jour, moins de vingt cas ont été confirmés sur le plan cellulaire et moléculaire, comme étant réellement similaires aux cas princeps décrits par Lowry et Pena dans la population indigène du Manitoba (province du Canada) (Lowry et al., 1971) (Pena and Shokeir, 1974). Le syndrome COFS constitue la forme extrême et prénatale du syndrome de Cockayne. Ces patients présentent les symptômes suivants : microcéphalie congénitale, cataracte congénitale et/ou microphthalmie, retard de développement psychomoteur sévère, retard de croissance staturo-pondéral (principalement postnatal), dysmorphie faciale. Une photosensibilité cutanée, une neuropathie périphérique, une surdité de perception peuvent également être observées (Laugel et al., 2008). La mutation du facteur CSB est le plus souvent associée à cette pathologie. Néanmoins, la mutation du gène ERCC1 a été présentée pour un cas COFS (Jaspers et al., 2007). Des formes cliniques particulières de COFS sont aussi associées à la mutation du gène XPD (Graham et al., 2001) et XPG (Nouspikel et al., 1997). Le diagnostic prénatal peut être suspecté par la présence d'une cataracte, d'une arthrogrypose et d'une microcéphalie. Le syndrome COFS est une maladie sévère, entraînant le décès dans les premières années de vie, notamment par infection respiratoire.

### **3. Trichothiodystrophie (TTD)**

TTD est une maladie génétique autosomale récessive rare, se caractérisant principalement par un déficit de protéines soufrées. Cela a pour effet de provoquer des cheveux et des ongles cassants ; ces cheveux présentent des bandes claires et sombres (*tiger tail*) vues au microscope. Ces patients développent également de l'ichtyose, se traduisant par un dessèchement de la peau et la présence de squames (écaille de peau). On détecte également des troubles neurologiques sévères, ainsi que des anomalies du développement s'associant avec des retards de croissance et des retards mentaux (microcéphalie) (Tableau 1) semblables à ceux des CS (Stefanini et al., 2010). La moitié des patients TTD, appelés aussi

PIBIDS (*Photosensibility, Ichthyosis, Brittle hair, Infertility, Decrease intelligency and Short stature*) présentent une photosensibilité. Néanmoins, les TTD ne développent pas d'anomalies pigmentaires, ni de cancers de la peau. La fréquence de cette pathologie est de 1,2/1 000 000 en Europe de l'ouest (Kleijer et al., 2008).

	Xeroderma Pigmentosum (XP)	XP avec anomalies neurologiques	Trichothio- dystrophie (TTD)	Syndrome de Cockayne (CS)	XP/CS
<i>Cutanés</i>					
Photosensibilité (UV)	+	++	+/-	+	+
Pigmentation accentuée	+	+	-	-	+
Cancers de la peau	+	+	-	-	+
<i>Yeux</i>					
Photophobie	+	+	+/-	+	+
Cancer	+	+	-	-	nd
Dégénérescence des pigments rétinien	-	-	-	+	-
<i>Développement</i>					
Retard de croissance	-	+/-	+	+	+
Immaturité sexuelle	-	-	+	+	+
<i>Neurologiques</i>					
Surdité	-	+	-	+	+
Retard du développement	-	+	+	+	+
Dégénération progressive des neurones	-	+	nd	+	+
Dégénération primaire des neurones	-	+	-	-	-
Démyélinisation	-	-	+	+	+
Microcéphalie	-	+	+/-	+	+
<i>Mécanismes défectueux</i>					
Défaut de NER	+	+	+	+	+
Réaction aux agents endommageant exogènes ou endogènes	++	++	-	+	+
Défauts développementaux	-	+	++	++	++

**Tableau 1** Tableau clinique des symptômes associés aux XP, TTD, CS et XP/CS.

Trois gènes présentant des mutations ont été identifiés comme responsables de la forme photosensible des TTD : XPB, XPD, p8/TTD-A. Ces gènes codent pour différentes sous-unités du facteur TFIIH. Son implication dans plusieurs processus cellulaires peut expliquer les différentes formes connues de TTD. En effet, il a été démontré que les cellules de patients TTD présentaient un faible niveau de TFIIH, s'expliquant par le fait que les différentes mutations semblent affecter la stabilité de TFIIH.

Les mutations de XPB sont rarement viables, du fait du rôle important de XPB dans la transcription. En effet, XPB par son activité 3'-5' hélicase, est essentiel pour l'ouverture de l'ADN alors que la protéine XPD n'est pas nécessaire pour la transcription basale *in vitro* (Tirode et al., 1999). Le 3<sup>ème</sup> gène touché est p8/TTD-A, identifié comme la 10<sup>ème</sup> sous-unité

de TFIIH (Giglia-Mari et al., 2004). Les phénotypes p8/TTD-A sont moins sévères, car p8 est une sous-unité dont l'absence complète n'est pas létale. En effet, p8 permet de réguler la concentration nucléaire de TFIIH en améliorant la stabilité de ce complexe lors de la transcription. p8 joue également un rôle dans la réparation de l'ADN, en stimulant l'activité ATPase de XPB, permettant l'ouverture de l'ADN (Coin et al., 2006).

Une forme non photosensible de TTD a été décrite : ces patients appelés TTDN1 (TTD non photosensible 1) présentent des mutations du gène C7orf11. Les cellules de ces patients ont une réponse normale aux UV et un niveau constant de TFIIH (Nakabayashi et al., 2005) (Botta et al., 2007). Le rôle fonctionnel de C7orf11 est peu connu ; il a été montré que TTDN1 intervient dans la mitose, en interagissant avec la kinase Plk1 (Zhang et al., 2007). De plus, aux vues des phénotypes des patients TTDN1, il a été suggéré que cette protéine pourrait jouer un rôle dans la régulation de la transcription, mais pas dans la réparation de l'ADN.

### III. La transcription des gènes de classe II

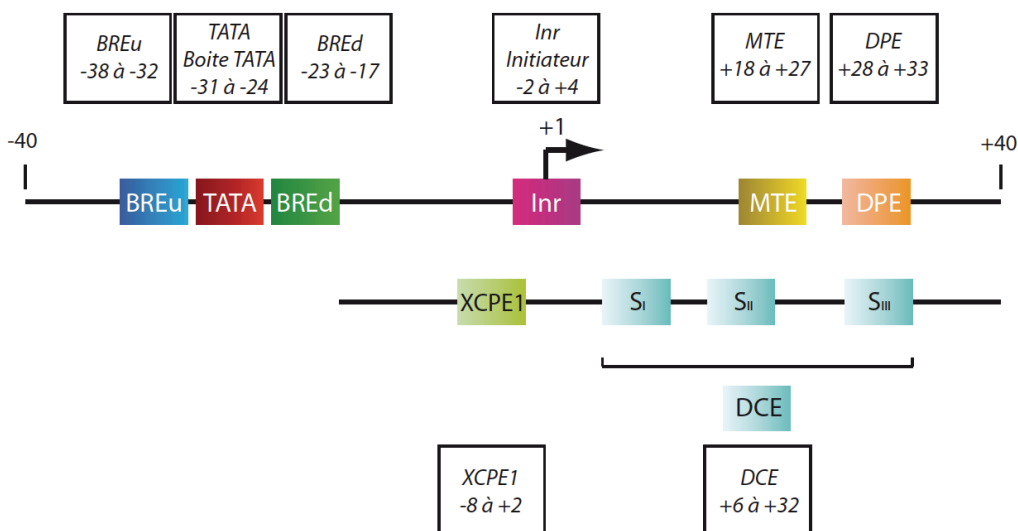
Au cours de sa vie, la cellule doit être capable de réguler de façon précise son profil d'expression génétique. Pour cela, elle utilise son ADN comme matrice au cours de l'étape de transcription, qui est essentielle pour de nombreux processus biologiques comme la prolifération, la différenciation ou encore la survie cellulaire. La transcription permet la synthèse d'ARN, qui sont ensuite maturés et transportés dans le cytoplasme pour subir l'étape de traduction permettant la synthèse des protéines (Orphanides et al., 1996) (Orphanides and Reinberg, 2002). La transcription est le premier point de contrôle de l'expression génétique, elle fait intervenir des enzymes appelées ARN polymérases, étant capables de décoder l'information génétique. Chez les eucaryotes, quatre ARN polymérases ont été décrites. L'ARN polymérase I, qui intervient dans la transcription des ARN ribosomiques (ARNr excepté l'ARNr 5S) et l'ARN polymérase III, qui est impliquée dans la synthèse des ARN de transfert (ARNt) et de l'ARN ribosomique 5S (Roeder and Rutter, 1970). Récemment, l'ARN polymérase IV a été découverte chez les plantes, pour son rôle dans la transcription de petits ARN impliqués dans le phénomène de « silencing » par la formation

d'hétérochromatine (Wierzbicki et al., 2008). Enfin, l'ARN polymérase II catalyse la formation d'ARN messager (ARNm) permettant la synthèse des protéines.

La transcription des gènes de classe II est un processus se divisant en trois grandes étapes principales qui fait intervenir des facteurs généraux de transcription (GTF : *general transcription factors*) au nombre de six (TFIIA, B, D, E, F et H). La première étape appelée l'initiation, consiste en une ouverture de la double hélice d'ADN. Elle est suivie par la synthèse des 1<sup>ers</sup> nucléotides et la libération de l'ARN polymérase II. La 2<sup>nd</sup> étape appelé l'élongation, est composée de la synthèse d'ARNm par l'ARN polymérase II. Viens ensuite l'étape de terminaison, pendant laquelle on assiste au détachement de la machinerie de transcription et à la libération de l'ARN.

### 1. Les séquences régulatrices de l'ADN

Tout le long du génome, la machinerie de transcription va reconnaître des régions spécifiques d'ADN, afin de pouvoir démarrer le processus de transcription. Ces éléments spécifiques, appelés promoteurs, vont servir de plateforme de recrutement pour la machinerie transcriptionnelle, qui est composée de l'ARN polymérase II et des facteurs généraux de transcription.



**Figure 9 Schéma du promoteur minimal eucaryote.** Le promoteur minimal est composé de la boîte TATA (TATA), de l'initiateur (Inr), de l'élément de reconnaissance de TFIIIB en amont de la boîte TATA (BREu) et en aval de la boîte TATA (BREd). Il peut également être composé du MTE, du DCE ou du DPE en aval du site d'initiation de la transcription.

## A. La structure du promoteur minimal

Le promoteur minimal a été défini comme la plus petite séquence d'ADN, nécessaire à l'initiation de la transcription. Il permet l'assemblage du complexe de pré-initiation (PIC), composé de l'ARN polymérase II et des facteurs généraux de transcription (TFIIA, B, D, E, F et H) (Butler and Kadonaga, 2002) (Smale and Kadonaga, 2003). Le promoteur minimal s'étend sur une quarantaine de nucléotides, en amont et en aval du site d'initiation de la transcription (+1). Il est composé d'un arrangement de petits motifs communs d'ADN : la boîte TATA, l'initiateur (Inr), l'élément de reconnaissance de TFIIIB (BRE), les éléments en aval du promoteur (DPE et DCE) ou encore l'élément à 10 motifs (MTE) (Figure 9). Il faut néanmoins préciser, que tous les promoteurs eucaryotes ne contiennent pas tous les éléments décrits ci-dessus.

### A.1 La boîte TATA

La boîte TATA est le premier motif du promoteur minimal à avoir été découvert en 1978 (Lifton et al., 1978) ; elle est également appelée la boîte Goldberg-Hogness. Il s'agit de l'élément du promoteur minimal le plus connu. Elle est composée de la séquence consensus TATAWAAR et est située au niveau de la position -31 ou -30 du site d'initiation de la transcription (Carninci et al., 2006). La boîte TATA est reconnue et liée par le facteur TBP (*TATA binding protein*), qui est une sous-unité du complexe TFIID (Burley and Roeder, 1996). TFIID est le complexe responsable du positionnement des autres facteurs généraux de transcription sur le promoteur. La boîte TATA est un élément conservé des archaebactéries aux humains (Reeve, 2003) et qui est également présente chez les plantes (Yamamoto et al., 2007). Néanmoins, la boîte TATA n'est pas présente dans tous les promoteurs. En effet, seul 10 à 15% des promoteurs minimaux humains présentent une boîte TATA (Carninci et al., 2006) (Cooper et al., 2006).

### A.2 L'élément initiateur (Inr)

Le motif initiateur (Inr) est situé autour du site d'initiation de la transcription (+1) entre les nucléotides -2 et +4. Il a été identifié chez les mammifères, chez la drosophile et chez la levure (Struhl, 1987). Chez les mammifères, la séquence consensus est YYANWYY

(Corden et al., 1980) et le « A » situé au milieu de la séquence, est communément considéré comme le nucléotide « +1 » du site d'initiation de la transcription. La séquence Inr est suffisante pour initier la transcription *in vivo* et *in vitro* (Javahery et al., 1994) de manière indépendante de la boîte TATA. De nombreux facteurs peuvent interagir avec l'Inr, dont le complexe TFIID via ses sous-unités TAF1 et TAF2 (Chalkley and Verrijzer, 1999) ou l'ARN polymérase II (Aso et al., 1994).

### A.3 Les éléments en aval du promoteur (DPE et DCE)

Le premier élément en aval du promoteur est le DPE (*downstream promoter element*). Il présente la séquence consensus suivante : RGWYV. Le DPE est précisément localisé entre les nucléotides +28 et +32 par rapport au site d'initiation de transcription « +1 » (Kutach and Kadonaga, 2000). C'est un élément qui est reconnu par les sous-unités TAF6 et TAF9 du facteur de transcription TFIID (Burke and Kadonaga, 1997). Pour être fonctionnel, le DPE nécessite la présence de l'élément initiateur (Kadonaga, 2002) et sa mutation provoque la perte de la fixation de TFIID (Burke and Kadonaga, 1996).

Un second élément en aval du promoteur a été décrit : le DCE (*downstream core element*). Il a été identifié sur le promoteur de la  $\beta$ -globine humaine (Lewis et al., 2000). Le DCE est constitué de trois sous-unités distinctes, localisées entre les nucléotides +6 et +32. Elles peuvent être liées par la sous-unité TAF1 du complexe TFIID (Lee et al., 2005).

### A.4 Le site de reconnaissance de TFIIB (BRE)

Le BRE (*TFIIB recognition element*) est le seul élément du promoteur minimal qui n'interagit pas avec le complexe TFIID. En effet, le BRE, situé au niveau de la boîte TATA, présente une séquence consensus G/C- G/C- G/A-CGCC permettant la liaison du complexe TFIIB (Lagrange et al., 1998). Un premier élément BRE, nommé BREu (*upstream*) est situé en amont de la boîte TATA entre les nucléotides -38 et -32. Puis un second BRE, appelé BREd (*downstream*) se trouve en aval de la boîte TATA entre les nucléotides -23 et -17 (Deng and Roberts, 2005). Ces deux éléments sont capables d'interagir avec TFIIB, ils fonctionnent en coopération avec la boîte TATA et permettent de réguler le niveau de base de la transcription (Lagrange et al., 1998) (Evans et al., 2001).



#### A.5 Les autres éléments du promoteur minimal

Le MTE (*motif ten element*) est un élément formé de la répétition d'un même motif (dix fois). Il a été identifié chez la drosophile (Ohler et al., 2002) et est conservé chez l'homme. Il est localisé en amont du DCE, entre les nucléotides +18 et +27 et peut interagir avec TFIID.

Un autre élément du promoteur minimal, l'élément XCPE1 (*X core promoter element 1*), a été découvert dans le promoteur du gène X du virus de l'hépatite B (Tokusumi et al., 2007). Cet élément, situé entre les nucléotides -8 et +2, présente une activité transcriptionnelle de base, amplifiée lorsqu'il est associé au niveau de sites de liaison activateurs. XCPE1 a été trouvé uniquement dans le promoteur minimal de 1% des gènes humains, et plus particulièrement dans les gènes dépourvus de boîte TATA (Tokusumi et al., 2007).

#### A.6 Les promoteurs à îlots CpG

Les îlots CpG (détaillés ultérieurement) peuvent être considérés comme une seconde classe de promoteur. En effet, à l'état naturel, la majorité des îlots CpG sont méthylés chez les mammifères, pour former un 5-méthylcytosine (5mC). Ils sont présents au niveau du site d'initiation de la transcription et leurs niveaux de méthylation permet de réguler le niveau d'expression des gènes (Cross and Bird, 1995). Lorsque ceux-ci sont dans un état déméthylé, il a été démontré que le promoteur est dans un état actif. Alors que les promoteurs dépendants de la boîte TATA ne représentent que 10 à 15% des promoteurs humains, on estime que les promoteurs à îlots CpG pourraient représenter environ 80% des promoteurs humains (Kim et al., 2005). Le mécanisme d'action de ces promoteurs est encore peu connu. Il semblerait que les îlots CpG présentent de nombreux sites de liaisons de la protéine Sp1. Sa fixation contribuerait au maintien de l'état non méthylé des CpG, afin de permettre la bonne fixation de la machinerie de transcription (Brandeis et al., 1994).

#### B. Les éléments de régulation

D'autres séquences d'ADN spécifiques, sont impliquées dans le contrôle de l'expression génétique, en permettant une régulation plus fine de la transcription. Il existe

deux types d'éléments de régulation, classés selon leur éloignement par rapport au site d'initiation de la transcription : les séquences dites proximales et celles dites distales.

**Les séquences proximales** sont situées au maximum à 200 paires de bases en amont du site d'initiation de la transcription. Il existe 2 types de séquences proximales, celles qui activent la transcription, appelées UAS chez la levure (*upstream activating sequence*) et celles qui ont un effet répresseur, appelées URS (*upstream repressing sequence*).

**Les éléments dits distaux**, capables de réguler la transcription, sont situés à plusieurs milliers de paires de bases en amont ou en aval du site d'initiation (Ptashne and Gann, 1998). Ils sont appelés « enhancer » lorsqu'ils activent la transcription ou « silencer » lorsqu'il la réprime. Ces éléments vont agir avec le promoteur en formant une boucle d'ADN, permettant le rapprochement de ces deux régions dans l'espace (Cook, 2003) (Saiz et al., 2005). Il a été montré qu'une proximité physique entre un enhancer et son promoteur cible, stimulait la formation du complexe de pré-initiation et activait la transcription (Deng and Blobel, 2010).

## 2. L'ARN polymérase II

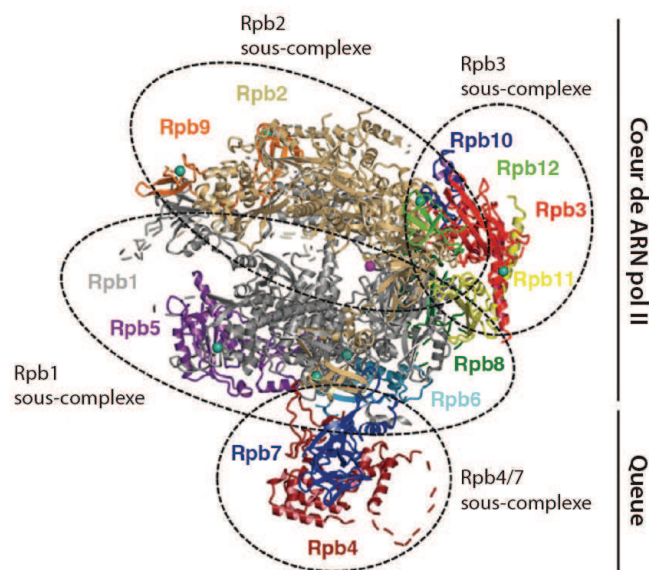
Afin de mieux comprendre le mécanisme de transcription, il a fallu dans un premier temps décrire les facteurs mis en jeu dans ce processus de production d'ARN. En effet, différentes ARN polymérases ont été identifiées chez les eucaryotes (ARN pol I, II, III) par Roeder et Rutter en 1969 (Roeder and Rutter, 1970). Nous nous intéressons plus particulièrement à l'ARN polymérase II, car il s'agit de la polymérase responsable de la synthèse des ARN messagers codant pour les protéines. L'ARN polymérase II est l'enzyme catalytique clé du complexe de pré-initiation de la transcription. Il s'agit d'un complexe multi-protéique, formé de 12 sous-unités, appelées Rpb1 à Rpb12 (Young, 1991), qui présente un poids total de 515kDa. Sa structure est très conservée au cours de l'évolution : on retrouve de nombreuses homologies avec les autres ARN polymérases humaines, des bactéries ou des levures. En effet, 7 sous-unités de la pol II humaine se retrouvent partiellement (Rpb 4/7/9) ou complètement (Rpb 6/8/10/12) chez la levure (McKune et al., 1995) (Khazak et al., 1998), et 5 sous-unités sont communes aux trois ARN polymérases humaines (Rpb 5/6/8/10/12) (Woychik and Young, 1990) (Hampsey, 1998). De plus Rpb1, 2,

3 et 6 ont des séquences similaires aux sous-unités  $\beta'$ ,  $\beta$ ,  $\alpha$  et  $\omega$  de la bactérie (Tan et al., 2000) (Tableau 2).

	Sous-unités	Fonction
Cœur	Rpb 1	Contient le CTD Impliqué dans la sélection du site d'initiation et dans l'élongation Orthologue de $\beta'$
	Rpb 2	Contient le site actif Orthologue de $\beta$ Impliquée dans la sélection du site d'initiation et dans l'élongation
	Rpb 3	Avec Rpb11, orthologue du dimère $\alpha 2$ chez la bactérie
	Rpb 5	Commune à l'ARN pol I et III, activateur transcriptionnel
	Rpb 6	Commune à l'ARN pol I et III
	Rpb 8	Commune à l'ARN pol I et III
	Rpb 9	Impliquée dans la sélection du site d'initiation
	Rpb 10	Commune à l'ARN pol I et III
	Rpb 11	Avec Rpb3, orthologue du dimère $\alpha 2$ chez la bactérie
	Rpb 12	Commune à l'ARN pol I et III
Queue	Rpb 4	Impliqué dans l'initiation, forme un sous-complexe avec Rpb7
	Rpb 7	Impliqué dans l'initiation, forme un sous-complexe avec Rpb4

**Tableau 2** Composition des différentes sous-unités de l'ARN polymérase II et leurs propriétés. Présentation des 12 sous-unités de l'ARN pol II et de leurs fonctions.

La résolution de la structure de l'ARN polymérase II par cristallographie aux rayons X, a permis de mieux comprendre son organisation (Armache et al., 2003) (Bushnell and Kornberg, 2003) (Figure 10). Cette découverte a permis à RD Kornberg d'obtenir le prix Nobel de chimie en 2006. On peut distinguer deux grandes parties : le cœur composé de 10 sous-unités (Rpb1, 2, 3, 5, 6, 8, 9, 10, 11 et 12) et la queue formée de Rpb4 et 7. Les deux plus grandes sous-unités Rpb1 et 2 forment le cœur du complexe, en créant un sillon de 25 Å chargé positivement, dans lequel vient se loger l'ADN, qui sera transcrit. Les autres sous-unités Rpb3, 6, 10, 11 et 12 vont être assemblées autour pour maintenir la structure (Figure 10). Le côté du sillon constitué par Rpb1, forme une pince mobile, qui intervient dans l'initiation de la transcription pour empêcher l'ADN de se dissocier (Gnatt et al., 2001). Rpb2 forme ce que l'on appelle « le mur », il délimite l'extrémité du sillon. Le module formé par Rpb4 et 7 est ancré à la base de la pince Rpb1 et serait impliqué dans l'initiation de la transcription, en permettant la fermeture de la pince mobile (Armache et al., 2003).

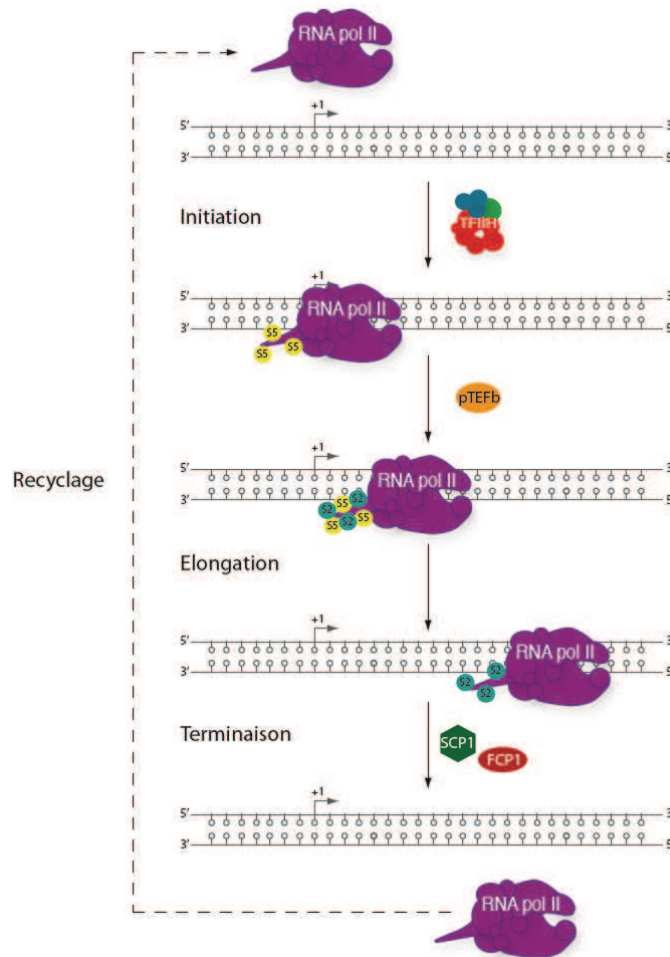


**Figure 10 Structure de l'ARN polymérase II** (adapté de (Wild and Cramer, 2012)). Structure complète de l'ARN pol II par cristallographie aux rayons X.

L'assemblage de l'ARN pol II se fait selon un ordre bien précis, qui a été démontré chez la levure (Zhang et al., 1999a). En effet, cela commence par la formation d'un complexe Rpb2, 3 et 11, qui interagit avec Rpb1. Pour maintenir la structure, Rpb10 et 12 sont nécessaires et il semblerait que Rpb6 joue un rôle dans l'assemblage de Rpb1 (Wild and Cramer, 2012). Cet assemblage s'effectue dans le cytoplasme, avant que l'ARN polymérase II complète soit importée dans le noyau. Après son action, elle est dissociée et les différentes sous-unités sont recyclées, afin de permettre l'assemblage d'une nouvelle ARN pol II.

L'ARN polymérase II se différencie des autres polymérases humaines par la présence du domaine carboxy-terminal (**CTD**) de la sous-unité Rpb1. Le CTD se caractérise par une répétition d'un heptapeptide : Tyr-Ser-Pro-Thr-Ser-Pro-Ser (YSPTSPS). Il s'agit d'un motif très conservé au cours de l'évolution, que l'on retrouve aussi bien chez la levure, les plantes ou encore les humains. La seule différence est le nombre de répétition de ce motif entre les espèces. Chez la levure, il se répète 26 à 29 fois, 42 fois chez la drosophile et 52 fois chez l'homme (Dahmus, 1995) (Hampsey, 1998) (Lee and Young, 2000). Le CTD est un élément essentiel à la vie. En effet, les cellules contenant des ARN polymérase II avec moins de répétition que la normale ne sont pas viables (Nonet and Young, 1987) (Zehring et al., 1988). Cependant, le CTD n'est pas nécessaire pour la transcription *in vitro* (Akoulitchev et al., 1995). L'ARN pol II peut exister dans une forme hyper-phosphorylée ou dans un état non phosphorylé du CTD. Ces phosphorylations impliquent principalement les sérines 2 et 5 des

différentes répétitions (Dahmus, 1995). Cette dynamique de phosphorylation va permettre la progression de l'ARN pol II durant le processus de transcription.



**Figure 11** Cycle de phosphorylation du CTD de Rpb1 lors de la transcription.

Le cycle de phosphorylation du CTD (Figure 11) commence par la fixation de l'ARN polymérase II non phosphorylée sur le promoteur (Cadena and Dahmus, 1987) (O'Brien et al., 1994). Différentes kinases interviennent ensuite pour phosphoryler le CTD. Chez l'homme CDK7, qui fait partie du facteur de transcription TFIIB, phosphoryle la sérine 5 du CTD, permettant ainsi l'assemblage du PIC, menant à l'initiation de la transcription (Serizawa et al., 1993). La seconde phosphorylation s'effectue lors de l'élongation : la kinase CDK9 du facteur d'élongation b (pTEFb), phosphoryle la sérine 2 du CTD (Cho et al., 2001) (Zhou et al., 2000). L'ARN polymérase II se trouve alors dans un état hyper-phosphorylé, permettant le

bon déroulement de l'élongation, avec l'ajout de la coiffe de l'ARN messenger. La grande majorité des sérines 5-P sont ensuite déphosphorylées par la phosphatase Rtr1. Cela favorise la balance de phosphorylation en faveur de la sérine 2 et enclenche la phase de terminaison de la transcription (Mosley et al., 2009). Puis, deux nouvelles phosphatases sont recrutées : SCP1 (*small CTD phosphatase 1*) et FCP1 (*TFIIF associated CTD phosphatase 1*), qui vont respectivement déphosphoryler la sérine 5 et la sérine 2 (Cho et al., 2001) du CTD, pour que l'ARN polymérase II soit recyclée.

### 3. Les facteurs généraux de transcription

L'ARN polymérase II, qui est l'enzyme catalysant la réaction de transcription, est incapable d'initier par elle-même la transcription *in vivo*. C'est pourquoi, d'autres facteurs généraux de transcription vont suppléer la polymérase, afin de lancer l'initiation de la transcription. Ils sont au nombre de six : TFIIA, B, D, E, F et H (Samuels et al., 1982) (Orphanides et al., 1996).

- **TFIIA** : TFIIA humain est un facteur composé de 3 sous-unités :  $\alpha$ ,  $\beta$  et  $\gamma$  respectivement de 35, 19 et 12 kDa (Ma et al., 1993). Ces trois sous-unités sont encodées par 2 gènes : TFIIA $\alpha\beta$  et TFIIA $\gamma$  chez les eucaryotes ou TOA1 et 2 chez la levure (Dejong and Roeder, 1993). TFIIA est essentiel à la transcription (Reinberg and Roeder, 1987), il stabilise la fixation de TBP à la boîte TATA (Buratowski et al., 1989). TFIIA a un effet stimulateur sur la transcription, il inhibe l'effet négatif de co-facteurs comme NC1, Topo1 et HMG1 (Inostroza et al., 1992) (Merino et al., 1993). TFIIA a également des fonctions de co-activateur, car il potentialise l'effet d'autres co-activateurs comme PC4 ou HMG2 (Ge and Roeder, 1994) (Shykind et al., 1995). La déplétion de TFIIA provoque un arrêt du cycle cellulaire en phase G2, générant alors une forte diminution de la transcription des gènes avec ou sans boîte TATA (Kang et al., 1995). Tous cela démontre que TFIIA joue un rôle important dans les mécanismes de régulation de la transcription et du cycle cellulaire.

- **TFIIB** : TFIIB a été initialement purifié à partir de cellules HeLa, comme un facteur capable de se lier au promoteur, via une interaction avec TBP. Il est décrit comme étant nécessaire pour le recrutement de l'ARN polymérase II (Sawadogo and Roeder, 1985) (Reinberg and

Roeder, 1987). TFIIIB est une protéine de 33kDa, dont la structure primaire contient deux répétitions de 5 hélices  $\alpha$  du coté C-terminal. Le domaine N-terminal est formé d'un motif « zinc ribbon » et d'une région appelée « B-Finger ». Ces deux motifs sont très conservés au cours de l'évolution (Orphanides et al., 1996). TFIIIB interagit aussi bien avec le domaine TBP du complexe TFIID, qu'avec des séquences d'ADN promotrices. Il peut se fixer sur le promoteur, en interagissant avec le motif BRE<sup>u</sup> en amont de la boîte TATA et ainsi favoriser la formation du complexe TFIIIB/TBP/promoteur, ayant un effet positif sur la transcription (Lagrange et al., 1998). Mais la majeure partie des protéines TFIIIB vont interagir avec l'ADN en aval de la boîte TATA, via le motif BRE<sup>d</sup>, potentialisant ainsi la formation du complexe TFIIIB/TBP/promoteur. En plus de son interaction directe avec TBP, TFIIIB peut interagir avec l'ARN polymérase II et le facteur de transcription TFIIF, via sa partie N-terminal. Il entre en contact avec les sous-unités Rpb1/2 et 9 de l'ARN pol II (Sun et al., 1996) et les sous-unités RAP30 et 74 de TFIIF (Ha et al., 1993). TFIIIB joue également un rôle majeur dans la sélection du site d'initiation de la transcription aussi bien chez la levure que chez l'homme (Li et al., 1994). Il est aussi impliqué dans la phosphorylation du CTD de l'ARN polymérase II, en interagissant avec la phosphatase Fcp1 (Kobor et al., 2000), réprimant ainsi son activité.

- **TFIID** : TFIID est le premier facteur général de transcription à se fixer sur le promoteur. Il s'agit d'un complexe composé de TBP (*TATA binding protein*) et de 13 autres facteurs associés à TBP appelés TAF (*TBP associated factor*). TBP reconnaît la boîte TATA du promoteur minimal alors que les TAF peuvent interagir avec le promoteur en se fixant sur les motifs Inr, DCE ou DPE. TBP permet donc l'interaction avec les promoteurs TATA dépendant et les TAF permettent d'interagir avec les promoteurs sans boîte TATA. TBP et les TAF sont très conservés chez les eucaryotes. Cependant, la composition en TAF du complexe TFIID varie en fonction du type de promoteur (Jacq et al., 1994). Les TAF jouent donc un rôle essentiel dans la sélection du promoteur et pourraient intervenir comme des co-activateurs de la transcription (Burley and Roeder, 1996) (Lee and Young, 2000). La molécule TBP est composée de 4 hélices  $\alpha$  et de 10 feuillets  $\beta$  (Nikolov et al., 1992). Lorsqu'elle se lie au promoteur, elle induit une courbure de l'ADN de 90°, qui provoque le déroulement partiel de la double hélice d'ADN (Nikolov et al., 1996). *In vitro*, TBP est suffisant pour permettre la transcription d'un promoteur avec une boîte TATA.



Certains TAF se retrouvent également dans le complexe SAGA, connu pour être capable de réguler la transcription (Baker and Grant, 2007). De plus, la structure des TAF 6, 9 et 12 présente de nombreuses similitudes avec celle des histones H3, H4 et H2B. Cela suggère l'existence d'une structure en octamère, comme pour les nucléosomes (Xie et al., 1996).

- **TFIIE** : TFIIE est un hétéro-dimère composé de 2 sous-unités d'un poids moléculaire de 56 kDa (TFIIE $\alpha$ ) et de 34 kDa (TFIIE $\beta$ ) (Ohkuma et al., 1990). TFIIE joue un rôle essentiel dans l'initiation de la transcription et dans l'échappée du promoteur (Hirose and Ohkuma, 2007). Il entre dans la formation du PIC après l'arrivée de la pol II, en interagissant directement avec la forme non phosphorylée de l'ARN polymérase II, avec TFIIB et TFIIF (Orphanides et al., 1996). TFIIE recrute ensuite TFIIH sur le promoteur et régule ses activités kinase et hélicase (Lu et al., 1992) (Ohkuma and Roeder, 1994). Le domaine C-term de TFIIE $\alpha$  va interagir avec la sous-unité p62 de TFIIH (Okuda et al., 2004). Enfin, il se fixe au niveau du site d'initiation entre la position -10 et +10, via sa sous-unité TFIIE $\beta$  (Tanaka et al., 2009).

- **TFIIF** : TFIIF humain est un hétéro-tétramère composé de la répétition de 2 sous-unités: RAP30 et RAP74 (Flores et al., 1990). La sous-unité RAP30 (26kDa) est capable d'interagir avec la pol II, l'ADN et la sous-unité RAP74 (Sopta et al., 1989) (Garrett et al., 1992). Cette dernière (58kDa) possède 3 domaines fonctionnels, capables d'interagir avec RAP30, la sous-unité TAF1 de TFIID, TFIIB, l'ARN pol II et la phosphatase Fcp1 (Ruppert and Tjian, 1995) (Archambault et al., 1998). RAP74 peut être phosphorylée par la sous-unité TAF1 et la kinase CK7 (Yonaha et al., 1997) (Rossignol et al., 1999). TFIIF joue de nombreux rôles durant l'assemblage du PIC, en étant associé à l'ARN polymérase II (Sopta et al., 1985). En effet, la sous-unité RAP74 interagit avec la sous-unité Rpb9 de la pol II (Petermann et al., 2003). Cette interaction facilite le recrutement de la pol II sur le complexe promoteur/TFIID/TFIIB (Flores et al., 1991) et la fixation de TFIIF augmente la stabilité de la pol II pour ce complexe. TFIIF est également nécessaire pour le recrutement des facteurs TFIIE et TFIIH (Orphanides et al., 1996). Il joue un rôle dans la sélection du site d'initiation de la transcription et est impliqué dans l'échappée du promoteur (Yan et al., 1999). TFIIF intervient aussi dans l'étape d'élongation de la transcription en interagissant avec la protéine d'élongation TFIS (Luse et al., 2011). De plus, il peut être phosphorylé par la protéine kinase CK2, pour stimuler l'élongation (Ujvari et al., 2011).

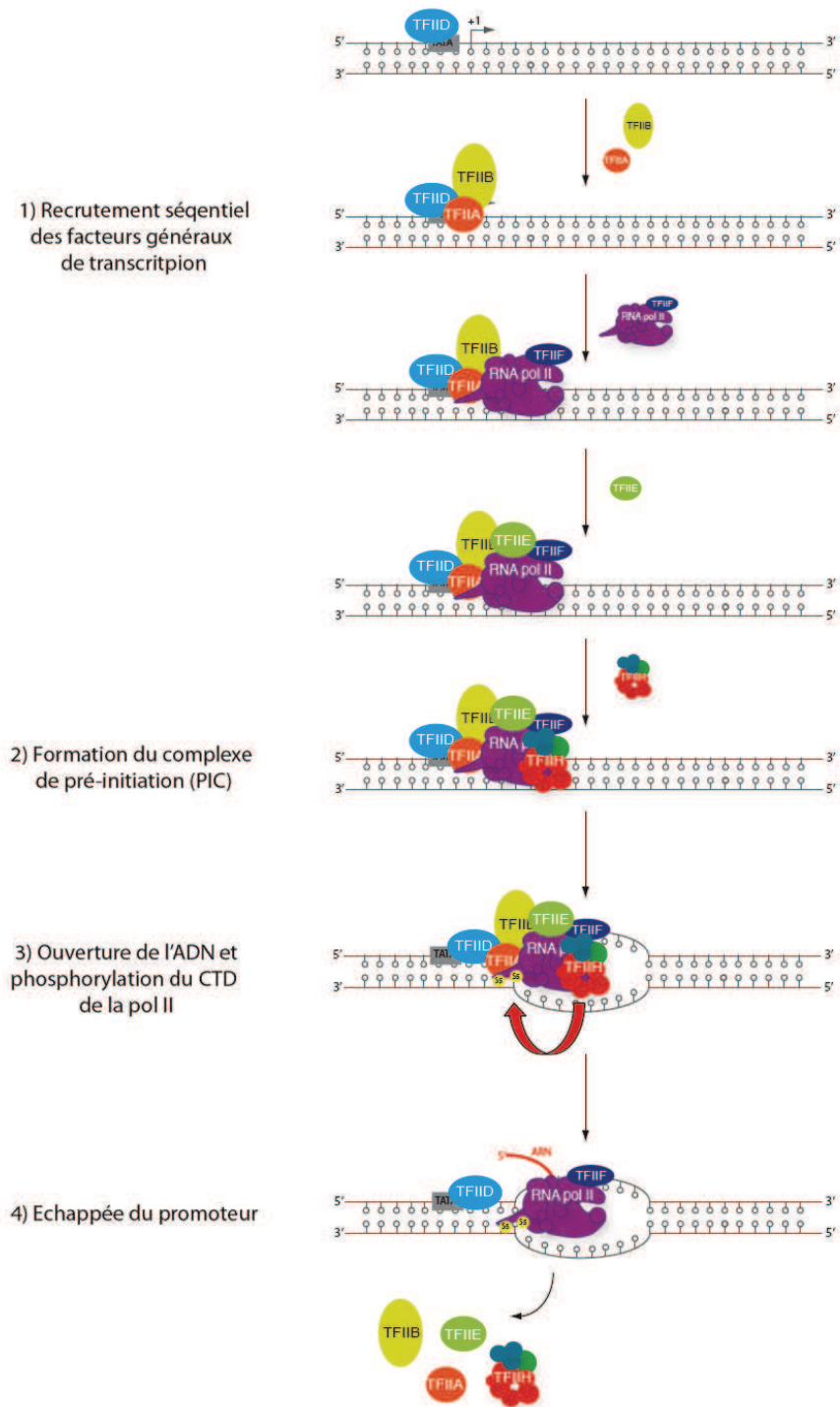


- **TFIIH** : TFIIH sera décrit plus en détail ultérieurement (partie IV) dans cette introduction. En résumé, TFIIH est un complexe formé de 2 sous-complexes : CAK (CDK7, cyclinH et MAT1) et cœur (XPB, p62, p52, p44, p34 et p8) liés par XPD. Il joue un rôle dans l'initiation de la transcription, en permettant l'ouverture de l'ADN et la phosphorylation du CTD de l'ARN polymérase II. Il est également impliqué dans les processus de réparation d'ADN, comme je l'ai décrit précédemment (Compe and Egly, 2012b).

#### 4. Le cycle de transcription

Le processus de transcription est divisé en 3 grandes étapes : l'initiation, qui consiste en la mise en place du complexe de transcription et la synthèse des premiers nucléotides. L'élongation, composée de l'allongement de l'ARNm et la terminaison, qui est effectuée lors du départ de la machinerie de transcription (Figure 12).

- **L'initiation** : L'initiation constitue la première étape de la transcription, elle commence par le recrutement séquentiel des différents facteurs généraux de transcription et par le recrutement de l'ARN polymérase II sur le promoteur du gène transcrit. Cela consiste à former le complexe de pré-initiation de la transcription (PIC) (Buratowski et al., 1989). TFIID est le premier facteur à être recruté sur le promoteur, il interagit via TBP avec la boîte TATA. Cela provoque une distorsion de l'ADN, qui facilite le recrutement des autres facteurs. Les facteurs TFIIA et TFIIB sont ensuite recrutés. TFIIB va reconnaître les motifs BRE<sup>u</sup> et BRE<sup>d</sup>, alors que TFIIA permet de stabiliser l'interaction TBP/TFIID/ADN (Kraemer et al., 2001). La structure TBP/TFIID/ADN est ensuite reconnue par l'ARN polymérase II, accompagnée de TFIIF (Figure 12.1). TFIIB est indispensable à la formation du PIC et est impliqué dans l'échappée du promoteur (Kostrewa et al., 2009). Une fois ce nouveau complexe stabilisé, TFIIIE est recruté en interagissant avec la pol II non phosphorylée, TFIIB et TFIIF (Roeder, 1996). Enfin, le recrutement de TFIIH permet la formation du complexe de pré-initiation au complet (Figure 12.2). Une autre alternative de formation du PIC a été décrite : la voie de l'holoenzyme, qui consiste en un pré-assemblage des facteurs de transcription (excepté TFIID), qui se fixeront en une étape sur le promoteur (Ossipow et al., 1995) (Chao et al., 1996).



**Figure 12** Les différentes étapes de l'initiation de la transcription. L'initiation de la transcription est divisée 4 parties, elle commence par le recrutement des facteurs généraux de transcription (TFIID, TFIIA, TFIIB ...) (1) permettant la formation du complexe de pré-initiation (2). Cela permet l'ouverture de l'ADN et la phosphorylation du CTD de l'ARN polymérase II (3), provoquant l'échappée du promoteur (4).

Néanmoins, la formation du PIC n'est pas suffisante pour démarrer la transcription, il faut une ouverture de la double hélice d'ADN, suivie de l'échappée du promoteur de la pol II. L'activité hélicase de la protéine XPB de TFIIH, permet l'ouverture de l'ADN de façon ATP dépendante au niveau du site d'initiation (-9/+2). Cela permet la formation du complexe ouvert de pré-initiation (POC) (Holstege et al., 1996) (Coin et al., 1999) (Figure 12.3). Une fois la chaîne d'ARN formée d'une dizaine de nucléotides, l'échappée du promoteur peut être effectuée. Cette étape dépend de l'état de phosphorylation du CTD de l'ARN polymérase II. En effet, la sous-unité CDK7 de TFIIH phosphoryle la sérine 5 du CTD de la pol II (voir cycle de phosphorylation du CTD – partie III.2). Par ailleurs, l'échappée du promoteur nécessite la désolidarisation des interactions de la pol II et des différents facteurs d'initiation (Figure 12.4). TFIIIF est le seul facteur accompagnant l'ARN polymérase II (Bengal et al., 1991) (Lei et al., 1999). Des facteurs comme TBP ou TFIIIE, vont se détacher de la pol II, lorsque celle-ci est phosphorylée (Maxon et al., 1994). TFIIID, reste quant à lui, fixé sur le promoteur, pour permettre une ré-initiation de la transcription.

- **L'élongation** : L'élongation est l'étape de production du transcrit d'ADN par l'ARN polymérase II. Elle débute lorsque le transcrit est constitué de 11 à 15 nucléotides. L'ADN est alors ouvert sur 12 à 15 paires de bases et l'on estime que l'ARN pol II synthétise *in vivo* entre 1500 et 2000 nucléotides par minute (Gnatt et al., 2001). La vitesse d'élongation est dépendante du nombre de facteurs régulant positivement ou négativement l'élongation (Shilatifard et al., 2003). L'ARN polymérase II peut exister dans un état actif ou inactif, ceci régule le phénomène de « pausing » de l'élongation du transcrit. Ainsi, l'ARN polymérase reste associée à l'ARN naissant et à l'ADN matrice, jusqu'à ce que les conditions soient favorables à la transcription. Par exemple, la pause de la pol II peut être induite par les facteurs répresseurs DSIF ou NELF. Dans ce cas, afin de permettre le redémarrage de la transcription, le facteur P-TEFb (CDK9/cycline T), qui est un facteur régulant positivement l'élongation, est recruté. Il phosphoryle la sérine 2 du CTD de l'ARN polymérase II, ayant pour effet de stimuler l'élongation (Peterlin and Price, 2006). D'autres facteurs comme TFIIIS ou les élongines sont connus pour réguler positivement l'élongation. Certains facteurs peuvent également potentialiser l'élongation de manière indirecte. C'est le cas du facteur FACT, qui a une activité de chaperonne des histones, modifiant ainsi l'environnement chromatinien afin de favoriser le passage de l'ARN polymérase II.

- **La terminaison** : La terminaison est l'étape finale de la transcription. Il s'agit d'une étape fondamentale, permettant le relargage du transcrit et la dissociation de l'ARN polymérase II. Cette dernière peut alors être recyclée. La terminaison est encore un mécanisme peu connu, il semblerait qu'elle soit dépendante du processus de maturation des ARN en 3' (Proudfoot, 1989) et de la présence du signal de polyadénylation (poly(A)) (Birise et al., 1998). La maturation des ARN en 3' est une étape essentielle de la terminaison (Hirose and Manley, 2000), elle permet le relargage de l'ARNm. Le site de poly(A) (AAUAAA) est reconnu par le facteur de clivage/polyadénylation CPSF, qui est accompagné par le facteur de stimulation de clivage CstF50 et de nombreux autres facteurs (le complexe de maturation fait 1MDa). Le CTD de l'ARN pol II joue, lui aussi, un rôle important en interagissant avec de nombreux facteurs comme CPSF et CstF50 (McCracken et al., 1997). En effet, la sérine 2 phosphorylée permet le recrutement de facteurs de maturation comme Pcf11 (Ahn et al., 2004). L'ARN est alors clivée au niveau du site poly(A), et la déphosphorylation du CTD de l'ARN pol II par FCP1 permet son relargage (Lian et al., 2008). L'ARN subit ensuite l'ajout de la queue poly(A) par la poly(A) polymérase. Enfin, il subit l'opération d'épissage par le spliceosome, afin de pouvoir être traduit en protéine.

## 5. La régulation de la transcription

### A. La régulation de la transcription par les récepteurs nucléaires

Au cours du développement et de la vie d'un organisme, certains gènes constitutifs sont requis de manière permanente, pour les processus basiques de maintenance de la fonction cellulaire. Ces gènes, appelés « gènes de ménages » (*housekeeping gene*), sont exprimés de manière constantes dans la cellule (exemple : GAPDH, HSP90 ...). En parallèle, on retrouve un grand nombre de gènes, régulés par l'action de divers stimuli endogènes ou exogènes. Les hormones sont un exemple de stimulus, elles vont interagir avec des récepteurs spécifiques : les récepteurs nucléaires, afin de réguler de manière précise l'expression de certains gènes dans le noyau.

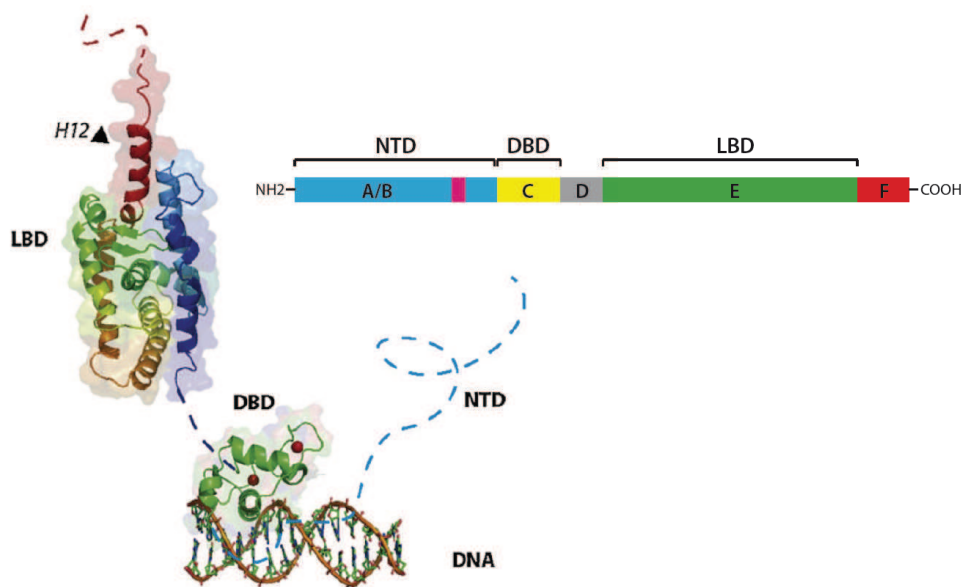
Les récepteurs nucléaires forment une superfamille de facteurs de transcription, qui régulent une grande variété de fonctions biologiques tels que la croissance, la reproduction ou encore le développement. Ces protéines exercent un contrôle direct sur l'expression de

gènes cibles, en réponse à divers stimuli. Les récepteurs nucléaires possèdent une structure constituée de différents domaines fonctionnels. Ils sont formés d'un domaine de liaison à l'ADN (DBD) et d'un domaine de liaison au ligand (LDB) (Wrangé et al., 1984). Outre ces deux séquences essentielles, on retrouve en N-terminal, un domaine peu conservé dans les différents récepteurs nucléaires appelé domaine NTD. Ces récepteurs sont classés en fonction de la nature de leurs ligands et de leurs modes de fixation à l'ADN (Kastner et al., 1995) (Mangelsdorf et al., 1995). La superfamille des récepteurs nucléaires est divisée en trois classes. La classe I comprend les récepteurs aux hormones stéroïdiennes fixant l'ADN sous forme d'homo-dimères. Elle comprend les récepteurs aux glucocorticoïdes (GR), aux estrogènes (ER) ou encore à la progestérone (PR). La classe II est formée de récepteurs liant l'ADN sous forme d'hétéro-dimères avec RXR (récepteur X à l'acide rétinoïque), tels que les récepteurs aux hormones thyroïdiennes (RT), à l'acide rétinoïque (RAR) et à la vitamine D (VDR). La troisième classe est formée par le récepteur X à l'acide rétinoïque lui-même, ainsi que certains récepteurs orphelins, capables de se lier sous forme d'homo-dimère avec l'ADN. Enfin le quatrième et dernier groupe est celui des récepteurs orphelins qui lient l'ADN sous la forme de monomères comme les récepteurs SF-1, LHR1 et ERR (Mangelsdorf et al., 1995).

#### A.1 Le récepteur à l'acide rétinoïque (RAR)

Le récepteur à l'acide rétinoïque est activé par les rétinoïdes, des dérivés de la vitamine A. La vitamine A est essentielle pour le développement embryonnaire, l'homéostasie cellulaire ou encore l'organogénèse ; la seule source de vitamine A provient de l'alimentation. Après une cascade de réaction chimique, la vitamine A donne naissance à l'acide rétinoïque, pouvant interagir avec des protéines cytosoliques (CRABP-1 et CRABP-2) afin d'entrer dans le noyau cellulaire (Budhu and Noy, 2002) (Delva et al., 1999). L'acide rétinoïque agit ensuite sur l'expression des gènes, en activant le récepteur nucléaire RAR qui existe sous trois isoformes :  $\alpha$  (NR1B1),  $\beta$  (NR1B2) et  $\gamma$  (NR1B3) (Chambon, 1996). Ces isoformes diffèrent par leurs extrémités N-terminal (Germain et al., 2006) et interviennent dans la régulation de gènes impliqués dans la différenciation cellulaire, dans la prolifération et dans l'apoptose.

Le récepteur à l'acide rétinoïque présente une organisation bien définie, sa structure est composée de trois domaines majeurs (Figure 13) : un domaine central de liaison à l'ADN (DBD), un domaine de liaison au ligand (LDB) en C-terminal, et un domaine N-terminal peu conservé (NTD) (Rochette-Egly and Germain, 2009).



**Figure 13 Organisation du récepteur à l'acide rétinoïque (RAR).** Schéma présentant les trois principaux domaines de RAR : le NTD (NH2-terminal domain), le DBD (DNA binding domain) et le LBD (ligand binding domain) auxquels s'ajoute la région D ou région charnière et la région F.

**Le DBD** est le domaine le plus conservé dans les RAR. Il s'agit d'un domaine globulaire, formé de deux motifs en doigt de zinc reliés par une boucle, suivie d'une extension COOH-terminal (Zechel et al., 1994a) (Zechel et al., 1994b) (Figure 13). Ces motifs forment deux hélices  $\alpha$  perpendiculaires, dont la première s'insère dans le grand sillon de l'ADN et la seconde maintient la stabilité du domaine (Lee et al., 1993). Le DBD confère aux RAR, la capacité de se lier à de courtes séquences d'ADN, nommées RARE (*retinoic acid response element*), localisés dans les promoteurs des gènes cibles de l'acide rétinoïque. La reconnaissance de ces RARE nécessite la formation de l'hétéro-dimère RAR/RXR. Ces éléments de réponse sont formés par la répétition d'un motif hexamère (A/G) G (G/T) TCA (Leid et al., 1992) (Mangelsdorf and Evans, 1995). L'extension COOH-terminal du DBD semble de plus en plus importante. En effet, elle jouerait un rôle dans la reconnaissance à l'ADN et/ou la dimérisation, via un changement conformationnel (Hsu et al., 1998) (Khorasanizadeh and

Rastinejad, 2001). Enfin, le DBD fait également l'objet de nombreuses modifications post-traductionnelles tels que des phosphorylations pouvant moduler son interaction avec l'ADN (Duong and Rochette-Egly, 2011).

**Le LDB** est situé en position C-terminal. Il s'agit d'un domaine majeur de la reconnaissance du ligand. C'est un domaine très conservé dans les différents RAR, qui est composé de 12 hélices  $\alpha$  formant un enchevêtrement de trois feuillets d'hélices antiparallèles (Renaud and Moras, 2000) (Figure 13). Il contient une poche de liaison au ligand (LBP), une surface de dimérisation avec le LDB du récepteur RXR et une surface de liaison pour des co-régulateurs. L'analyse par cristallographie a également permis de mettre en évidence l'importance de l'hélice H12, qui renferme la fonction de transactivation dépendante du ligand appelée AF-2 (le Maire et al., 2010).

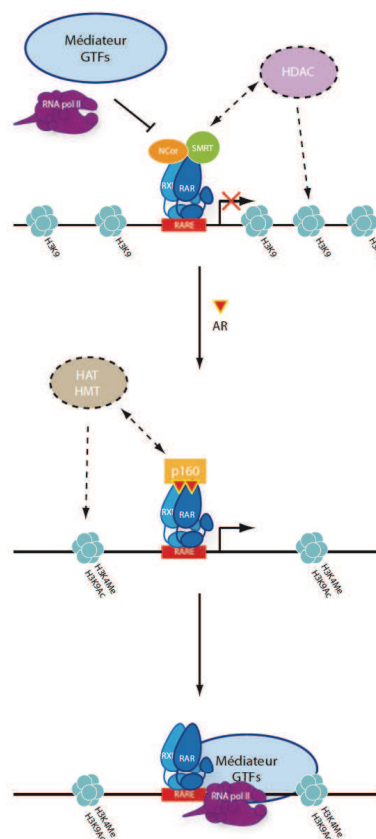
**Le NTD** est le troisième domaine du RAR (NH2 terminal domain). Il est formé des régions A et B, et inclut la fonction de transactivation indépendante du ligand appelée AF-1 (Nagpal et al., 1993). En effet, le NTD est capable à lui seul, d'activer la transactivation de gènes cibles, même en absence de ligand. Cependant, cette fonction est très minoritaire comparée à celle du LDB.

#### A.2 La transcription des gènes cibles de l'acide rétinoïque

Le mécanisme d'activation des gènes cibles de l'acide rétinoïque est un modèle séquentiel permettant de passer d'un état de répression à un état d'activation des gènes. Ce mécanisme est basé sur les changements conformationnels de l'hélice H12 de RAR et sur le recrutement de complexes multi-protéiques par le LBD (Dilworth and Chambon, 2001). Le modèle présenté par Dilworth et Chambon a été démontré pour le récepteur RAR $\alpha$  (Figure 14).

En absence de ligand, l'hétéro-dimère RAR/RXR lie l'ADN au niveau des éléments de réponses RARE. Il est associé à un complexe co-répresseur, via l'hélice H12 du LBD de RAR, réprimant ainsi l'expression du gène cible. Des co-répresseurs, comme SMRT (*silencing mediator of retinoic and thyroid receptor*) et NCoR (*nuclear receptor corepressor*), n'ayant pas d'action enzymatique propre, servent de plateforme pour le recrutement d'histone déacétylase (HDAC) comme HDAC I, II ou III. Ces déacétylases maintiennent l'histone H3K9, générant ainsi, au niveau du promoteur, un état chromatinien non favorable à la

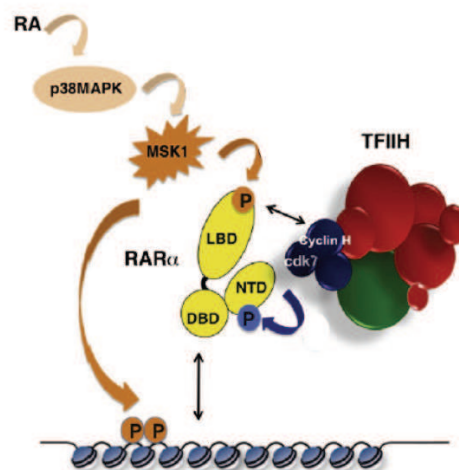
transcription, appelé hétérochromatine (décrit ultérieurement). Dès que l'acide rétinoïque se fixe sur le RAR, l'hélice H12 du LBD subit des changements conformationnels, générant une nouvelle surface d'interaction pour des co-activateurs, tels que les protéines de la famille p160. Ce changement de conformation permet également le relargage des corépresseurs. Le recrutement de ces co-activateurs permet le recrutement séquentiel de complexes, impliqués dans le remodelage de la chromatine, tels que les méthyltransférases et les HAT (histone acétyltransférase). Ces enzymes permettent respectivement la méthylation de H3K4 et l'acétylation de H3K9, créant ainsi un état de chromatine décompacté au niveau du promoteur du gène cible : euchromatine. Une fois activé, RAR recrute le médiateur, les facteurs généraux de transcription et l'ARN polymérase II, afin d'enclencher la transcription du gène cible (Bastien and Rochette-Egly, 2004) (Rosenfeld et al., 2006).



**Figure 14** Modèle classique d'activation des gènes cibles de l'acide rétinoïque. En absence de ligand, l'hétérodimère RAR/RXR est présent sur l'ADN associé aux complexes corépresseurs, réprimant l'expression du gène. La liaison de l'AR induit la dissociation des corépresseurs et le recrutement de co-activateurs. La chromatine décompactée (euchromatine), le médiateur, les facteurs généraux de transcription (GTFs) et l'ARN pol II sont recrutés, permettant ainsi l'initiation de la transcription.



A ce modèle peut s'ajouter un autre plus récent, qui présente le RAR comme intervenant dans une réponse non génomique, plus rapide, faisant intervenir une cascade de kinases. En effet, en réponse à l'AR, RAR activerait la MAPK p38 (Alsayed et al., 2001) (Gianni et al., 2002) (Figure 15). P38 active ensuite MSK1, qui phosphoryle une sérine de RAR (la sérine S369 du LBD). Cette phosphorylation provoque un changement de conformation, permettant le recrutement de TFIH via sa sous-unité CDK7 (Keriel et al., 2002). Celle-ci phosphoryle ensuite la sérine 77 du NTD de RAR, permettant l'interaction entre la cycline H de TFIH et le LBD de RAR (Bour et al., 2005). Cette cascade de phosphorylation est suivie par la phosphorylation de la sérine 10 de l'histone H3 par MSK1. Cela provoque un réarrangement de la chromatine, rendant accessible les éléments de réponses pour la fixation de l'hétéro-dimère RAR/RXR et ainsi enclencher la transcription du gène (Bruck et al., 2009).



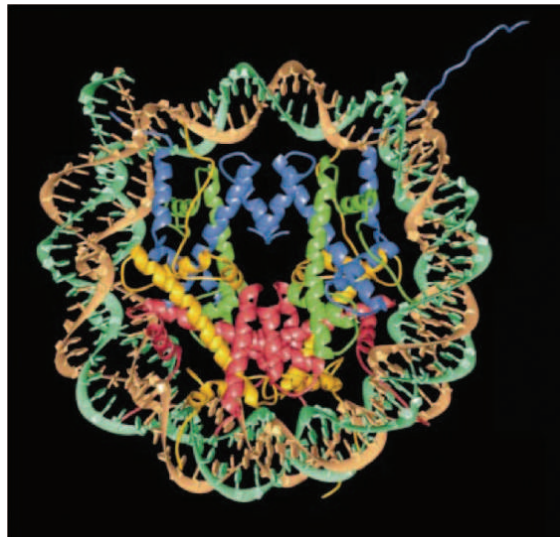
**Figure 15 Schéma d'une nouvelle voie de signalisation de RAR.** En réponse à l'AR, p38MAPK et MSK1 sont activés. MSK1 va phosphoryler la sérine du LBD permettant le recrutement de TFIH via la cycline H. La sous-unité CDK7 va alors phosphoryler le NTD de RAR qui est suivie de la phosphorylation de H3S10 par MSK1. Cela va provoquer la décompaction de l'ADN et rendre accessible les éléments de réponse pour la fixation de RAR sur le promoteur (Duong and Rochette-Egly, 2011).

## B. La régulation de la chromatine

Lors de la transcription, l'ADN polymérase II va rencontrer de nombreux obstacles, dont la chromatine, la méthylation de l'ADN ou encore les lésions d'ADN. En effet, la structure chromatinienne est un obstacle pour la transcription, car elle limite l'accessibilité de l'ADN aux divers complexes souhaitant se fixer sur l'ADN.

## B.1 Organisation de la chromatine

Dans les cellules eucaryotes, le matériel génétique n'est pas à l'état nu, il est organisé en une structure plus complexe, appelée chromatine, qui est formée d'ADN et de protéines. Cette structure mesure environ 2m de long et est contenue dans le noyau cellulaire faisant quelques micromètres de diamètres (Kornberg, 1977). Pour pouvoir tenir dans le noyau, la chromatine subit plusieurs degrés de compaction, que je décrirai ultérieurement (Figure 17). L'unité élémentaire de la chromatine est le nucléosome. Un nucléosome est constitué d'une partie centrale, appelée le cœur, et d'une région inter-nucléosomique. Le cœur est composé de 146 paires de bases, enroulées sur 1,7 tour autour d'un octamère d'histones. Cet octamère est composé de deux copies de chaque type d'histone : H2A, H2B, H3 et H4 (Luger et al., 1997) (Figure 16).



**Figure 16** Structure cristallographique du nucléosome : 146 paires de bases d'ADN (brun et vert) sont enroulées autour de huit protéines histones (H2A en jaune, H2B en rouge, H3 en bleu et H4 en vert) (Luger et al., 1997).

L'assemblage des histones commence par l'hétéro-dimérisation de H3 et H4, puis est suivie par la formation du tétramère (H3-H4)<sub>2</sub>. Viens ensuite s'ajouter deux dimères de H2A-H2B, afin de former un octamère complet (Arents and Moudrianakis, 1995). Les histones sont de petites molécules (11 à 22kDa) très conservées au cours de l'évolution ; ils sont formés d'un domaine globulaire hydrophobe, chargé positivement, qui est composé à 20% d'acides

aminés positifs (Arg, Lys), ce qui en fait une protéine basique. Cette charge positive permet l'interaction avec l'ADN, chargé négativement.

L'enroulement de l'ADN autour de plusieurs nucléosomes forme ensuite une structure dite en « collier de perle » de 11 nm de diamètre, correspondant au premier degré de compaction (Figure 17). La partie inter-nucléosomique, qui relie le cœur de deux nucléosomes, peut supporter une protéine histone supplémentaire, appelée histone H1. Il s'agit d'une protéine très riche en acides aminés basiques (riche en Lysine), qui va permettre le maintien de la compaction de l'ADN autour du nucléosome (Figure 17). Le nucléosome constitue une première barrière à la transcription. En effet, le nucléosome réprime la transcription aussi bien *in vivo* (Han and Grunstein, 1988) que *in vitro* (Knezetic and Luse, 1986) (Lorch et al., 1987).

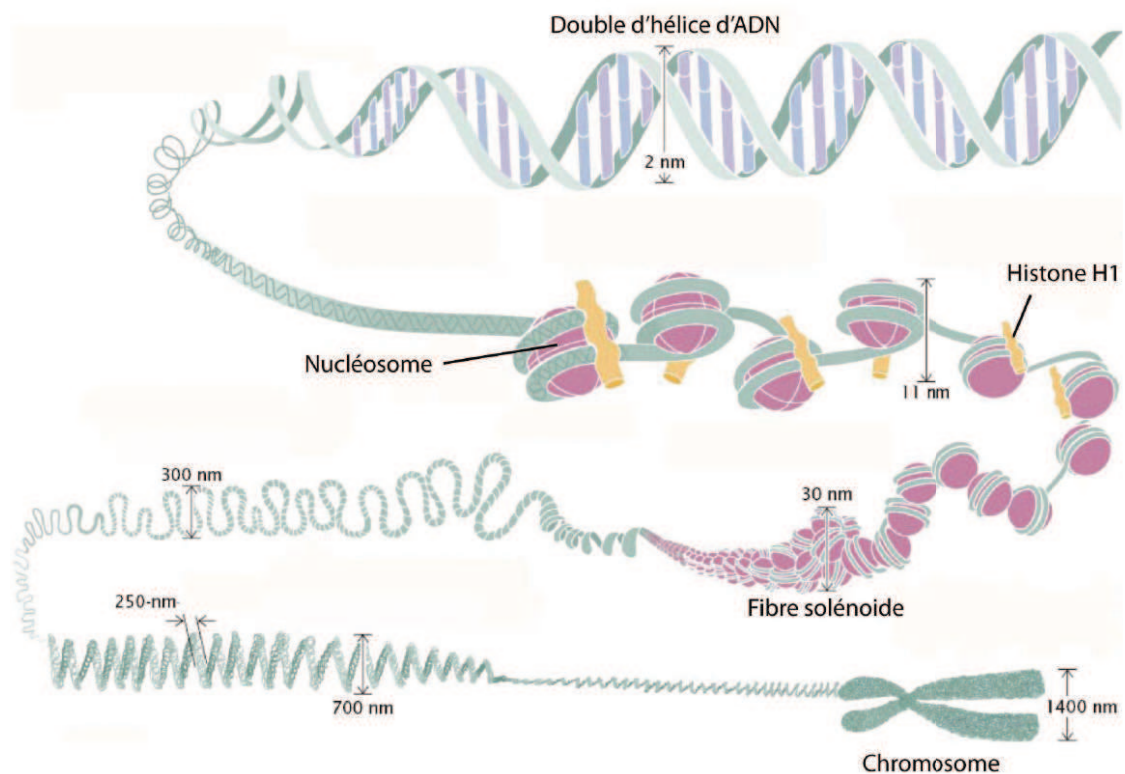
L'ADN subit ensuite un deuxième degré de compaction, par la formation d'une structure de 30 micromètres de diamètre que l'on appelle la fibre solénoïde (Finch and Klug, 1976). La présence de l'histone H1 est très importante pour la stabilité de cette structure, la fibre solénoïde se met mieux en place en présence de H1 (Fan et al., 2003). Le modèle de la fibre solénoïde est formé par l'enroulement de 6 à 8 nucléosomes par tour (Widom et al., 1985). Cependant, un deuxième modèle d'enroulement a été décrit plus récemment, il est appelé « zig-zag » du fait de son arrangement en zig-zag, et présente des contacts entre les nucléosomes 1 et 3 puis 2 et 4 (Woodcock and Ghosh, 2010).

Des états d'organisation plus condensés peuvent ensuite être mis en place avec une condensation maximale de la chromatine. Ces structures dites tertiaires, vont de la formation d'une fibre d'environ 300 nm de diamètre, jusqu'à l'état le plus condensé qui est le chromosome métaphasique de 1400 nm de diamètre (Figure 17).

Lors de la transcription, une corrélation très importante existe entre le niveau de compaction de la chromatine et le potentiel de mise en place de la transcription des gènes. En effet, en 1928, Heitz a décrit deux états fonctionnels de la chromatine possédant un degré de compaction différent. Ces deux états, appelés euchromatine et hétérochromatine ont été par la suite associés respectivement à un état favorable ou non à la transcription.

L'euchromatine est un état décondensé de la chromatine, permettant l'accès à l'ADN des différents facteurs de la machinerie de transcription. L'euchromatine est typiquement enrichie en histones acétylés (par exemple H3K9Ac et H4K16Ac) et présente également la

méthylation de l'histone H3K4 (H3K4Me) (Grunstein et al., 1995) (Litt et al., 2001) (Noma et al., 2001). Au contraire, l'hétérochromatine est définie comme une structure beaucoup plus condensée, inaccessible aux facteurs de transcription. Elle se caractérise par une hypoacétylation des histones et la méthylation de certains histones bien précis ( par exemple H3K9Me) (Tschiersch et al., 1994) (Rea et al., 2000) (Jenuwein and Allis, 2001). Elle est également associée à la présence de la protéine d'hétérochromatine HP1 (Eisenberg et al., 1990) (Bannister et al., 2001) et à la méthylation de l'ADN (Keshet et al., 1986) (Hennig, 1999). Deux types d'hétérochromatine ont été décrits : une forme dite « constitutive » qui correspond à une chromatine silencieuse de manière permanente, et une forme « facultative », qui contient des régions pouvant adopter une conformation dite « ouverte - fermée » en fonction du contexte (Trojer and Reinberg, 2007). Ces deux types d'hétérochromatine assurent le maintien de la stabilité du génome.



**Figure 17** Schéma des différents degrés d'organisation de la chromatine (Annunziato, 2012). La double hélice d'ADN s'enroule autour d'un octamère d'histones (H2A, H2B, H3 et H4) pour former le nucléosome. Ces derniers interagissent entre eux par le biais de l'histone H1, ce qui accroît le niveau de compaction de la chromatine en formant les fibres de 30 nm. Le niveau ultime de condensation est celui du chromosome, surtout pendant la métaphase.

Ces changements de conformation sont régulés dans le noyau cellulaire par différents processus tels que : les variants d'histones, les modifications post-traductionnelles des histones, la méthylation de l'ADN et le remodelage de la chromatine.

## B.2 Les variants d'histones

Les variants d'histones sont un premier moyen de modifier l'état de la chromatine. En effet, en plus des histones dits « majeurs » (H2A, H2B, H3 et H4), il existe des variants d'histones encodés par des gènes différents, qui sont synthétisés à un niveau plus faible. Chaque histone de l'octamère, excepté H4, existe sous plusieurs isoformes, codés par des gènes différents (Trojer and Reinberg, 2007). Ces variants jouent un rôle dans la dynamique chromatinienne et sont impliqués dans de nombreux processus tels que : la réparation de l'ADN, la ségrégation des chromosomes, l'initiation et la terminaison de la transcription (Pusarla and Bhargava, 2005). Les histones majeurs sont synthétisés lors de la phase S du cycle cellulaire et sont dépendants de la réplication, alors que les variants d'histones peuvent être produits tout le long du cycle cellulaire. Ces variants sont très conservés au cours de l'évolution.

Parmi les histones du cœur nucléosomique, l'histone H2A est celui qui possède le plus grand nombre de variants. Les histones majeurs H2A1 et H2A2 sont codés respectivement par dix gènes pour H2A1 et par un gène pour H2A2. Chez l'homme, H2A présente cinq variants : H2AX, H2AZ, macroH2A (1 et 2) et H2ABbd.

**Le variant H2AX**, décrit en 1980 (West and Bonner, 1980), intervient dans la signalisation cellulaire et la réparation. En effet, la sérine 139 de l'extrémité C-term peut être phosphorylée par les kinases ATM/ATR en réponse à un coupure double brin de l'ADN (Rogakou and Sekeri-Pataryas, 1999). Cela déstabiliserait le nucléosome, pour permettre une meilleure ouverture de la chromatine. Une étude récente montre que H2AX pourrait également jouer un rôle dans le remodelage de la chromatine, lors du développement de la souris (Nashun et al., 2010).

**Le variant H2AZ** est, quant à lui, impliqué dans le mécanisme de ségrégation des chromosomes et dans la régulation de la transcription (Gevry et al., 2009) (Rangasamy et al., 2004). L'acétylation de H2AZ serait associée à l'activation de gène, mais son rôle n'est pas encore bien défini (Thambirajah et al., 2009). De plus, l'accumulation de H2AZ affecte la

mobilité et le positionnement des nucléosomes. Le variant H2AZ est-il incorporé par l'action du complexe de remodelage de la chromatine ATP dépendante SWR-1 (Zhang et al., 2005). Cela permettrait le recrutement de facteurs de transcription, de la pol II, et provoquerait l'éviction du nucléosome.

**Le variant macroH2A** semble jouer un rôle dans la régulation de la transcription via la présence en C-term d'un motif « Leucine Zipper » commun à plusieurs facteurs de transcription (Pehrson and Fried, 1992). Concernant le variant H2A.Bbd, découvert il y a peu de temps (Chadwick and Willard, 2001), il semblerait que les nucléosomes le contenant soient plus ouverts et moins stables que ceux formés de H2A (Angelov et al., 2004).

Les histones H2B, H2B1 et H2B2 sont les isoformes majeurs et un seul variant est connu : hTSH2B (*human testis/sperm-specific histone*) qui est présent uniquement dans les testicules et les spermatozoïdes.

Concernant l'**histone H3**, on distingue actuellement deux isoformes majeurs : H3.1 et H3.2, et trois variants : H3.3, H3.t et CenH3S. Le variant H3.3 est le plus étudié, il est synthétisé tout au long du cycle cellulaire et est incorporé dans la chromatine de manière indépendante de la réplication, d'où son nom « d'histone de remplacement ». Alors que les histones majeurs H3.1 et H3.2 font intervenir le complexe CAF-1 (*chromatin assembly factor 1*) pour leur mise en place au sein du nucléosome, l'incorporation de H3.3 nécessite l'intervention de la chaperonne d'histone HIRA (Ray-Gallet et al., 2002). H3.3 est recruté au niveau de loci transcriptionnellement actifs (Loyola et al., 2006) et est souvent associé à des marques actives de la transcription. D'autres études suggèrent même que H3.3 pourrait agir comme un marqueur positif actif de transcription (Chow et al., 2005).

Le variant CenH3S est un variant exclusivement localisé au niveau des centromères, alors que H3.t se retrouve uniquement dans les testicules et jouerait un rôle dans la réorganisation de la chromatine, lors de la spermatogénèse (Tachiwana et al., 2010).

### B.3 Les modifications post-traductionnelles des histones (MPTs)

Les modifications post-traductionnelles des histones (MPTs) sont une autre source majeure de variation de la structure de la chromatine (Kouzarides, 2007). En effet, les premières descriptions de ces modifications décrites par Vincent Allfrey, datent des années

1960 (Allfrey and Mirsky, 1964). Les histones peuvent être ornées de nombreuses MPTs au niveau de leurs extrémités N-term et C-term. En modifiant la structure chromatinienne, ces modifications permettent l'accès ou non au génome à différents complexes, pour la mise en place de nombreux processus tels que la transcription, la réparation de l'ADN, la réplication ou la recombinaison. Néanmoins, il a été récemment mis en évidence des modifications ayant lieu sur la partie globulaire des histones : par exemple la méthylation de H3K56 qui joue un rôle dans la réplication, via une interaction avec PCNA (Lee et al., 2012).

Au total, il existe plus de 60 résidus, présentant des modifications, ce qui en fait un énorme potentiel pour la régulation de la chromatine (Kouzarides, 2007). Ces modifications peuvent être de différents types. En effet, on observe au moins neuf types de MPTs : l'acétylation, la méthylation, la phosphorylation, l'ubiquitination, la sumoylation, l'ADP-ribosylation, la biotinylation, la déimination ou encore la glycoxylation (Tableau 3). Les plus étudiées étant les quatre premières, elles feront l'objet d'une présentation plus détaillée ultérieurement (Bannister and Kouzarides, 2011).

Modifications	Résidus modifiés	Enzymes
Acétylation	K - Ac	Histone acétyltransférase / Histone déacétyltransférase
Méthylation	K - Me1 / K - Me2 R - Me1 / R - Me2 / R - Me3	Méthyltransférase / Déméthyltransférase
Phosphorylation	S - P T - P	Protéine kinase / Phosphatase
Ubiquitination	K - Ub	Complexe E1, E2, E3 / Isopeptidase
Sumoylation	K - Su	Complexe E1, E2, E3 / Isopeptidase
ADP-ribosylation	K - Ar S - Ar G - Ar	Mono ou poly ADP-ribosyltransférase / Sirtuines / Glycohydrolase
Biotinylation	K - B	Holocarboxylase / Biotinidase
Déimination	R	Peptidylarginine déiminase 4
Glycoxylation	K	ADP ribose

**Tableau 3** Liste des différentes modifications post-traductionnelles des histones (MPTs) connues, des différents résidus modifiés et des enzymes impliquées dans ces modifications.



Ces modifications dynamiquement font intervenir de nombreuses enzymes, qui ont été en partie définies. Ces MPTs modulent la chromatine, via deux principaux modes d'actions :

- Le 1<sup>er</sup> consiste en une action directe sur la structure des nucléosomes. En effet, l'acétylation ou la phosphorylation des histones vont réduire la charge positive des histones, ayant pour effet de perturber certaines interactions électrostatiques avec les nucléosomes adjacents, ou entre les histones et l'ADN. Les sites d'acétylation sont nombreux et incluent notamment H3K9, H3K14, H3K18, H4K5, H4K8 et H4K12 (Kouzarides, 2007). L'hyperacétylation des histones engendre une chromatine moins compacte, permettant l'accès de l'ADN aux protéines de transcription ou de réparation (Wang et al., 2008). Par exemple, la phosphorylation de H3S10 inhiberait la fixation du facteur HP1 sur l'hétérochromatine lors de la métaphase, facilitant le remodelage des chromosomes.

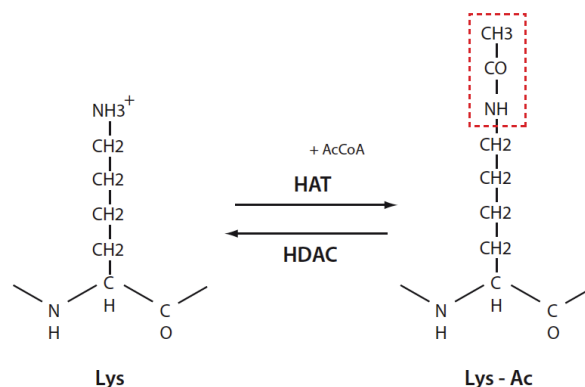
- Le 2<sup>ème</sup> mécanisme consiste à réguler la fixation des protéines interagissant avec la chromatine. En effet, de nombreux facteurs, associés à la chromatine interagissent spécifiquement avec les histones, via différents domaines. Par exemple, il existe plusieurs domaines de reconnaissance des lysines méthylées : les domaines « PHD fingers », les chromodomaines, les Tudor, PWWP ou encore les MBT domaines (Champagne and Kutateladze, 2009) (Maurer-Stroh et al., 2003). La méthylation de H3K4 est liée par le chromodomaine de CHD1, une enzyme de remodelage ATP dépendante qui repositionne les nucléosomes (Sims et al., 2005). La protéine d'hétérochromatine HP1 se lie également à H3K9me3, via son chromodomaine, lors de la compaction de la chromatine (Bannister et al., 2001).

Les Lysines acétylées sont également reconnues par un domaine spécifique, appelé bromodomaines. Par exemple, le complexe de remodelage ATP dépendant SWI2/SNF2 reconnaît les histones acétylés, via son bromodomaine, permettant ainsi l'ouverture de la chromatine (Hassan et al., 2002). Récemment, il a été démontré que le domaine « PHD fingers » était aussi capable de reconnaître spécifiquement des histones acétylés. C'est effectivement le cas de DPF3b, responsable du recrutement des complexes de remodelage de la chromatine BAF sur les histones acétylés (Zeng et al., 2010).



En plus de ces modifications, il existe un niveau supérieur de complexité, avec une interrelation entre les différentes modifications, permettant un contrôle plus fin de la chromatine (Lee et al., 2010). En effet, il arrive que la modification d'un résidu puisse altérer une seconde MPTs. C'est le cas de la méthylation de H3K4, qui nécessite la monoubiquitination de H2B pour être méthylée par le complexe COMPASS (Wood et al., 2007).

**L'acétylation** : il s'agit de la première modification post-traductionnelle décrite en 1964 par Allfrey (Allfrey and Mirsky, 1964). Depuis, il a été démontré à de nombreuses reprises, que l'acétylation des histones était corrélée avec l'activation de la transcription (Hebbes et al., 1988) (Grunstein, 1997). En effet, les régions transcriptionnellement actives sont riches en histones acétylés, alors que leur hypo-acétylation est retrouvée dans l'hétérochromatine (Spotswood and Turner, 2002). Cette modification est catalysée uniquement sur des lysines par des enzymes appelées histones acétyltransférases (HATs). Ces enzymes sont capables de provoquer le transfert d'un groupement acétylé du cofacteur acétyl-coenzyme A (AcetylCoA), vers le groupement amine d'une lysine de la queue N-term des histones (Hodawadekar and Marmorstein, 2007) (Figure 18). La réaction inverse de déacétylation, qui est corrélée à une répression de la transcription, est catalysée par des histones déacétyltransférases (HDACs).



**Figure 18** Réaction d'acétylation des résidus lysines des histones.

Les HATs sont des enzymes spécifiques de certaines lysines. Elles se divisent en deux classes principales : celles dites de type-A et celles de type-B. Les HATs de type-A sont des enzymes nucléaires, qui modifient les histones de la chromatine. Il existe trois familles principales : les

GNAT (*GCN5 related N-acetyl*) comprenant les acétyltransférases GCN5, PCAF ou HAT1 (Dyda et al., 2000), les MYST dont le nom est formé par un acronyme des enzymes qui le compose (MOF – Ybf2/Sas3 – Sas2 – Tip60) (Utley and Cote, 2003), puis les CBP/p300 qui comprennent les enzymes CBP et p300 (Bannister and Kouzarides, 1996). Ces différents HATs sont souvent retrouvés au sein de complexes multi protéiques, recrutés au niveau de la chromatine, leur permettant d'interagir avec des facteurs de transcription et pouvant, dans certains cas modifier leurs spécificités. C'est en effet le cas pour GCN5, qui seul présente une activité d'acétylation de H3K14. Cependant, lorsqu'il se retrouve dans le complexe SAGA (*Spt-Ada-Gcn5-Acetyltransferase*), il va pouvoir acétyler H3K9, K14, K18 et K23 (Rodriguez-Navarro, 2009). L'acétylation des histones est généralement réalisée au niveau de lysines de la queue N-term des histones. Mais elle peut aussi avoir lieu au niveau globulaire, comme c'est le cas pour l'acétylation de H3K56 réalisée chez l'homme par GCN5 (Tjeertes et al., 2009). Des études récentes de « genome wide » ont montrées que les HATs agissaient comme des co-activateurs de la transcription (Wang et al., 2009). En effet, les HATs GCN5/PCAF sont responsables de l'acétylation de H3K9, et CPB/p300 sont responsables de celles de H3K18/27, lors de la transcription de gènes, dépendant des récepteurs nucléaires (Jin et al., 2011).

Les HDACs sont des enzymes présentant peu de spécificité propre, étant capable de déacétyler de multiples sites. Il existe quatre classes de HDACs. La classe I composée de HDAC 1, 2, 3, 8, qui sont des homologues de Rpb3 de la levure et la classe II formée de HDAC 4, 5, 6, 7, 9 et 10, qui sont des homologues de Hda1 de la levure. La classe IV n'est formée que de HDAC 11. Enfin, la classe III aussi appelée « sirtuin » est composée des HADCS SIRT1 à 7. Ces dernières nécessitent le co-facteur NAD<sup>+</sup> pour avoir une activité de déacétylation. Elles interviennent également dans la déacétylation de protéines non-histones, tels que le facteur p53 (Ito et al., 2002). Les HDACs sont des régulateurs de plusieurs processus cellulaires tels que le contrôle du cycle cellulaire, la différenciation et l'apoptose. Leur dérégulation est impliquée dans la carcinogénèse, avec souvent une surexpression (Minucci and Pelicci, 2006) ; c'est pourquoi de nombreux inhibiteurs de HDACs sont utilisés pour leurs potentiels anticancéreux (Gallinari et al., 2007).

**La méthylation** : Contrairement à l'acétylation, la méthylation n'est pas une modification uniquement activatrice ou répressive pour la transcription. Cela varie en fonction du résidu méthylé. La méthylation peut être effectuée sur deux types de résidus : les lysines ou les arginines. C'est une modification qui n'altère pas la charge de l'histone et ne perturbe pas sa structure. Elle permettrait principalement la constitution d'un site de recrutement de protéines régulatrices (Kouzarides, 2007). Il existe différents types de méthylation : pour les lysines, une mono, di ou tri-méthylation et pour les arginines, une mono et une diméthylation symétrique ou asymétrique (Bedford and Clarke, 2009) (Ng et al., 2009) (Figure 19). Cette modification est catalysée par des méthyltransférases (HMTs), qui permettent le transfert d'un groupement méthyl du co-facteur SAM (*S-adenosylmethionine*) sur une lysine ou une arginine. Le degré de méthylation varie en fonction des propriétés intrinsèques des enzymes : par exemple DIM-5 triméthyle H3K9 (Tamaru et al., 2003) et SET7/9 monométhyle H3K4 (Xiao et al., 2003). Il existe différents types de HMTs : les PRMT, qui sont les méthyltransférases des arginines (*protein arginine methyltransferase*) comprenant en outre les PRMT 1, 4 et 5 (Wolf, 2009) et les PHMT (*protein lysine methyltransferase*), qui catalysent la méthylation des lysines. La première méthyltransférase décrite a été Suv39H1, qui catalyse la méthylation de H3K9 (Rea et al., 2000). Par la suite, il a été démontré que les PHMTs présentaient tous un domaine enzymatique en commun, appelé domaine SET. La seule exception est la protéine DOT1, qui ne présente pas ce domaine et est capable de méthyler H3K79 (Martin and Zhang, 2005). Les histones méthylées sont variables en fonction de l'état de chromatine. En effet, les formes méthylées de H3K9, K27 et H4K20 sont essentiellement associées à l'hétérochromatine, alors que les méthylation de H3K4, K36 et K79 sont associées à la mise en place de l'euchromatine (Bannister and Kouzarides, 2011). La conséquence fonctionnelle de la méthylation permet le recrutement de protéines associées à la chromatine. C'est le cas de HP1 (*heterochromatin protein 1*), qui possède un chromodomaine, lui permettant de reconnaître la méthylation répressive de H3K9 et H3K27, alors que Chd1 (*chromodomain helicase DNA binding protein 1*) se fixe sur la forme activatrice de H3K4 (Sims et al., 2005).

Pendant de nombreuses années, on pensait que la méthylation était une marque stable et irréversible. Néanmoins, il a été découvert une enzyme, la PADI4 (*peptidyl-arginine deiminase 4*) capable de réaliser la déméthylation d'une arginine, en convertissant une

méthylarginine en citrulline (Cuthbert et al., 2004). Plus récemment, d'autres déméthylases des arginines ont été mis en évidence, avec notamment l'enzyme JMJD6, qui déméthyle les histones H3R2 et H3R3 (Chang et al., 2007). Concernant la déméthylation des lysines, il a été montré que l'enzyme LSD1 (*lysine specific demethylase 1*) était capable d'utiliser FAD comme co-facteur, pour catalyser la déméthylation de H3K4 (Shi et al., 2004). Deux classes de déméthylases ont alors été décrites : celles présentant un domaine LSD1 et celles ayant un domaine catalytique JMJD (*JMJC jumonji domain*), dont fait partie JMJD2, pouvant déméthyle la tri-méthylation de H3K9 et H3K36 (Whetstine et al., 2006).

Concernant la spécificité de déméthylation de LSD1, il semblerait qu'elle soit dépendante des protéines qui lui sont associées. En effet, lorsque LSD1 est associé au co-répresseur CoREST, elle déméthyle H3K4 (Shi et al., 2005), alors qu'associé aux récepteurs des androgènes, LSD1 déméthyle H3K9 (Metzger et al., 2005). Cela fait donc de LSD1, à la fois un co-activateur (déméthylation H3K9) et un co-répresseur (déméthylation H3K4) de la transcription.

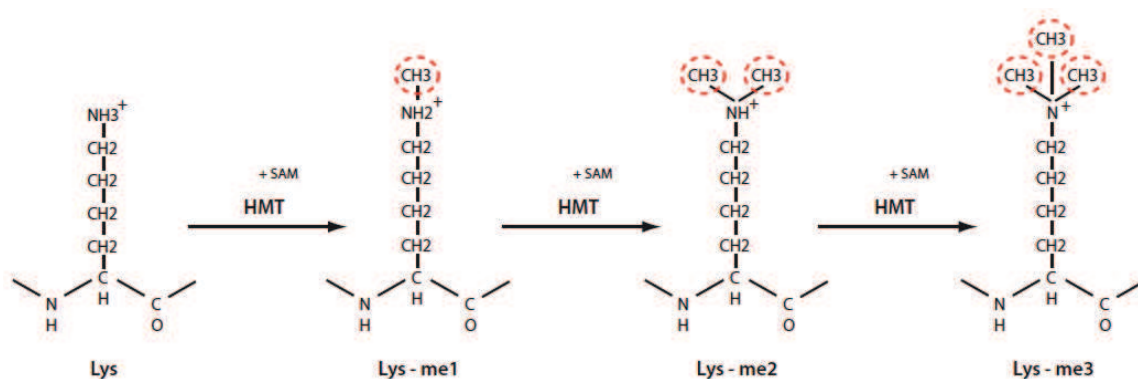


Figure 19 Schéma des différentes formes de méthylation des résidus lysines des histones.

**La phosphorylation** : la phosphorylation est un phénomène dynamique réversible, impliqué lors de l'activation transcriptionnelle, de l'apoptose, de la réparation de l'ADN ou de la condensation des chromosomes. Elle peut être réalisée sur les résidus sérines, thréonines ou tyrosines. Le niveau de phosphorylation est contrôlé par des kinases et des phosphatases, qui vont respectivement ajouter ou enlever ces modifications (Oki et al., 2007). Les kinases peuvent catalyser l'ajout d'un groupement phosphate de l'ATP au groupement hydroxyl d'un acide aminé cible. Cet ajout, majoritairement retrouvé au niveau de la queue N-term, a pour

effet, l'addition de charges négatives sur l'histone, ce qui influence la structure chromatinienne décondensée. Néanmoins, il a été récemment décrit l'existence de phosphorylations au niveau de la région globulaire des histones ; par exemple la phosphorylation de H3Y41 par la kinase JAK2 (Dawson et al., 2009).

La phosphorylation de H3S10 est impliquée dans des processus opposés, tels que la condensation de la chromatine, où H3S10 est phosphorylé par la kinase Aurora B (Clayton and Mahadevan, 2003). Cette phosphorylation a également été décrite lors de l'activation transcriptionnelle de gènes régulés par NF-KB et des gènes « immediate early », comme c-jun et c-fos via l'activation de cascades MAP kinase : PKA (*protein kinase A*), RSK2 (*ribosomal SG kinase 2*), MSK1 (*mitogen and stress activated kinase*). La phosphorylation de la sérine 139 de H2AX est impliquée dans la réparation de dommages double brin, en permettant l'ouverture de la chromatine sur quelques kilobases autour de la lésion (Sharma et al., 2012). Concernant les phosphatases, parmi les huit décrites chez l'homme, trois seulement ont été présentées comme jouant un rôle dans la déphosphorylation des histones (PP1, PP2A et PP4). Par exemple, H2AX est immédiatement déphosphorylée après la réparation de l'ADN par PP2A et PP4 (Nakada et al., 2008).

**L'ubiquitination** : il s'agit d'une modification covalente, réversible, mise en place au niveau des lysines des histones H2A, H2B, H3 et H2AZ. L'ubiquitin est un polypeptide volumineux de 76 acides aminés, ajouté de manière ATP dépendante, qui peut créer un encombrement stérique, influençant les autres MPTs. L'ubiquitination est impliquée dans l'activation transcriptionnelle et la réparation de l'ADN (Cao and Yan, 2012). Contrairement aux protéines qui sont polyubiquitinées, afin d'être dégradée par le protéasome, les histones sont généralement mono-ubiquitinés. L'ubiquitination est effectuée en trois étapes par l'action conjugué et successive de trois familles d'ubiquitinases : E1 pour l'activation, E2 pour la conjugaison et E3 pour la ligation. La déubiquitination est réalisée par des isopeptidases. H2A et H2B sont les histones les plus fréquemment ubiquitinés, et leurs ubiquitinations semblent être impliquées, aussi bien dans l'initiation et l'élongation de la transcription, que dans la réparation de l'ADN (Weake and Workman, 2008).

Strahl and Allis ont suggéré que ces différentes modifications post-traductionnelles des histones pouvaient constituer un « code » compréhensible par la cellule : le code histone

(Strahl and Allis, 2000). Ceci suppose alors l'existence de protéines capables d'établir le code ("writer"), de lire ce code ("reader"), et de l'effacer au besoin ("eraser"). Pour la plupart des modifications post-traductionnelles d'histones, des complexes protéiques sont connus dans les 3 catégories (Kouzarides, 2007) (Allis et al., 2007).

#### B.4 La méthylation de l'ADN

La méthylation de l'ADN est une modification, retrouvée chez la plupart des organismes vivants, allant des bactéries aux mammifères (Bird, 2002). Il s'agit d'un processus majeur de la régulation de la transcription, qui joue un rôle central dans les modifications de conformation de la chromatine. La méthylation de l'ADN est une marque épigénétique, principalement liée à la répression de la transcription. Chez les vertébrés, cette modification ne touche que les cytosines suivies d'une guanine appelées CpG. Elle consiste en l'ajout d'un groupement méthylé au niveau du carbone 5' d'une cytosine d'un dinucléotides CpG : 5 méthylcytosine (5mC). Cette réaction est catalysée par des enzymes appelées DNMT (*DNA methyltransferase*), présentant un domaine catalytique en C-term. Il existe trois types de DNMT : DNMT1, dont le rôle principal est la maintenance d'un profil de méthylation en phase S lors de la réplication. DNMT2, dont le rôle est très peu connu chez les mammifères et DNMT3, qui existe sous deux isoformes (DNMT3a/b) impliqués dans le processus de méthylation de l'ADN *de novo*. La méthylation *de novo* consiste en l'établissement de la méthylation lors de l'embryogénèse (Okano et al., 1999).

Les dinucléotides CpG sont répartis sur l'ensemble du génome (28 millions) mais ne sont pas distribués de manière homogène. On les retrouve en grand nombre au niveau de courtes régions appelées îlots CpG (Deaton and Bird, 2011). En général, ces îlots CpG, faisant environ 200 paires de bases de long, sont localisés au niveau des promoteurs et/ou du premier exon de plus de 70% des gènes (Klose and Bird, 2006) (Saxonov et al., 2006). On dénombre environ 25000 îlots CpG sur l'ensemble du génome (Illingworth and Bird, 2009). La méthylation sert principalement de plateforme pour le recrutement de nombreuses enzymes. En effet, cette modification est reconnue par des MBP (*methyl binding protein*), tels que MeCP2, MBP2 ou MBP3, qui vont se fixer au niveau des CpG méthylés, et ainsi provoquer le recrutement d'enzymes de modification d'histones telles que des histones déacétylases HDAC (Nan et al., 1998) ou des méthyltransférases (Fuks et al., 2003). Ces

enzymes peuvent modifier la structure de la chromatine (hétérochromatine), afin de consolider la répression de la transcription.

La méthylation a longtemps été considérée comme une modification stable et irréversible. Néanmoins, ce n'est que récemment, que le phénomène dynamique de la déméthylation a été accepté, avec la description d'une forme résultant du métabolisme du 5mC : le 5-hydroxyméthylcytosine (5hmC) (Kriaucionis and Heintz, 2009) (Tahiliani et al., 2009) (Figure 20). Cette réaction consiste à éliminer le groupement méthyl présent en 5' de cytosines méthylées. Elle est catalysée par des enzymes appelées TET (*ten eleven translocation*), qui existent sous trois isoformes : TET 1, 2 et 3 (Williams et al., 2012). Ces enzymes ont la capacité de convertir les 5mC en 5hmC *in vitro* et *in vivo* (Ito et al., 2010). TET1 et TET3 présentent un domaine CXXC, qui est un domaine de liaison au CpG. Le niveau de 5hmC est variable dans les tissus, ce qui laisse à penser qu'il s'agit d'un état transitoire dans le processus de déméthylation active. En effet, deux voies ont été présentées pour le passage de la forme 5hmC à une cytosine déméthylée (Figure 20). La première cascade enzymatique consisterait dans un premier temps en une étape de déamination, réalisée par les cytosines déaminases AID (*activation induced cytidine deaminase*)/APOBEC, convertissant le 5hmC en 5-hydroxyméthyluracil (5hmU). Cette étape serait suivie d'une glycosylation par Mbd4 ou TDG, générant un acide apyrimidique, remplacé en une cytosine par le mécanisme de BER (Guo et al., 2011a). La seconde voie consisterait en une oxydation successive du résidu 5hmC en 5-formylcytosine (5fC), puis en 5-carboxylcytosine (5caC) par les protéines TET (He et al., 2011) (Ito et al., 2011). Le passage de l'étape de 5caC à une cytosine déméthylée serait effectué par TDG (thymidine DNA glycosylase).

La protéine Gadd45 $\alpha$  (*growth arrest and DNA damage inducible protein 45  $\alpha$* ), connu pour son rôle dans la réparation de l'ADN, a été récemment décrite comme intervenant dans le processus de déméthylation (Barreto et al., 2007) (Schmitz et al., 2009). Il a été démontré que Gadd45 $\alpha$  interagissait avec les déaminases et les glycosylases, lors de la conversion du 5hmU en cytosine déméthylée (Figure 20) (Rai et al., 2008).

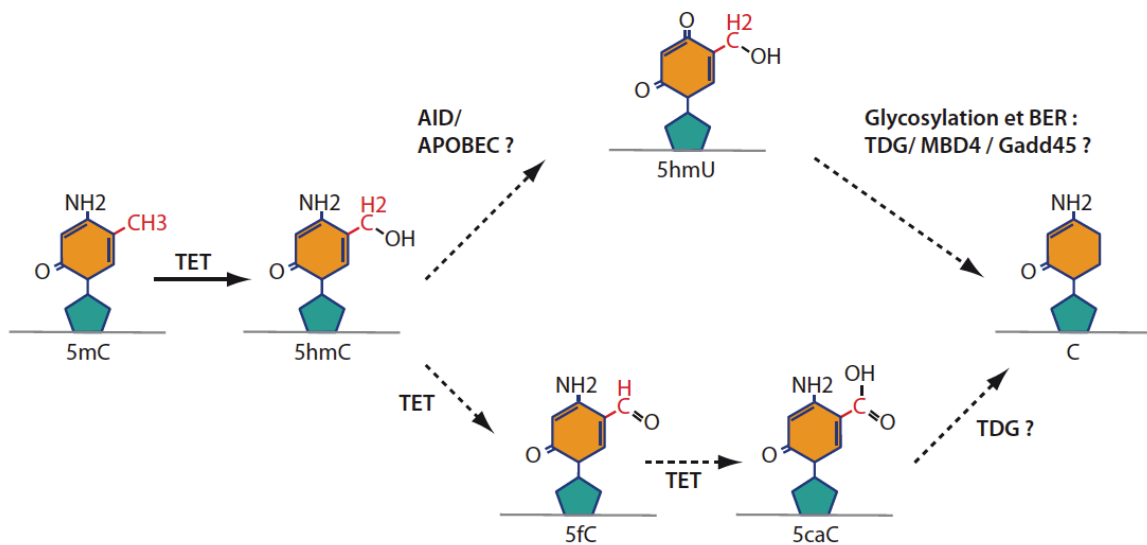


Figure 20 Mécanisme de déméthylation active (adapté de (Cedar and Bergman, 2012)).

### B.5 Remodelage de la chromatine

La chromatine peut être régulée à un niveau supérieur, avec la mise en place d'interactions tridimensionnelles au niveau du génome. Ces interactions, effectuées entre des séquences régulatrices bien définies du gène, sont appelées boucle de chromatine. En effet, les éléments régulateurs de certains gènes peuvent être distants de plusieurs kilobases. C'est pourquoi, afin de réguler la transcription, des rapprochements spatiaux sont mis en place pour permettre une interaction directe entre les différentes séquences d'un même gène (Dekker et al., 2002). Ces boucles de chromatine ont été décrites dans de nombreux gènes, comme une organisation généralisée de la fibre de chromatine, associée à l'activation de la transcription (Spilianakis and Flavell, 2004) (Vernimmen et al., 2007). Différentes régions des gènes ont été présentées comme formant des boucles de chromatine : c'est en effet le cas entre des enhanceurs et le promoteur ou encore entre le promoteur et le terminateur de gènes (Krivega and Dean, 2012) (Figure 21).

Le premier exemple de boucle de chromatine a été décrit au niveau du locus  $\beta$  globine de la souris, entre le LCR (*locus control region*) et le promoteur du gène  $\beta$  globine, lors de l'activation de la transcription (Carter et al., 2002) (Tolhuis et al., 2002). Depuis, des interactions entre le LCR et le promoteur des gènes suivants ont été décrites, afin de permettre leurs activations par un enhanceur distal :  $\alpha$ -globine, TH2 (*cytokine T helper type 2*), IFN $\gamma$  (*interferon gamma*) (Kadauke and Blobel, 2009). De nombreuses boucles d'ADN

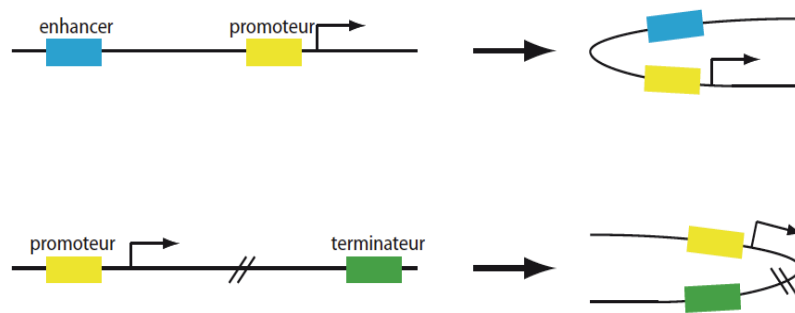


entre le promoteur et le terminateur de gènes transcrits par l'ARN polymérase II, ont été identifiées chez la levure et chez les mammifères (Laine et al., 2009). Les premières études démontrant une interaction entre le promoteur et le terminateur ont été réalisées chez la levure (*Saccharomyces cerevisiae*), pour les gènes SEN1 et BUD3 (O'Sullivan et al., 2004) (Ansari and Hampsey, 2005). De plus, il a récemment été démontré, que la formation de ce type de boucles était dépendante du recrutement du facteur de transcription TFIIB (Singh and Hampsey, 2007) (Figure 21).

De nombreux autres facteurs sont impliqués dans la formation de ces boucles. En effet, des facteurs comme GATA1, EKLF ou CTCF ont été identifiés dans la mise en place et la stabilité de la boucle. Les facteurs GATA1, FOG1 et EKLF (*erythroid kruppel-life factor*) ont été décrits comme nécessaires pour l'activation du gène  $\beta$ -globine et la formation de boucles, entre le LCR et le promoteur de ce gène (Vakoc et al., 2005) (Drissen et al., 2004). Le facteur CTCF (*CCCTC-binding factor*) est décrit comme un organisateur de chromatine, impliqué dans la mise en place d'interactions longues distances, chez les vertébrés et les insectes (Handoko et al., 2011). CTCF joue un rôle dans la stabilité de la boucle de chromatine du gène  $\beta$ -globine (Splinter et al., 2006) mais est également associé à la boucle enhancer-promoteur du locus Igf2-H19 (*insulin like growth factor 2/H19 noncoding RNA*) (Murrell et al., 2004). CTCF contribue aussi à la régulation de gènes seuls comme le locus CFTR (*cystic fibrosis transmembrane conductance regulator*) et IFN $\gamma$  (Gillen and Harris, 2011).

Aujourd'hui, le rôle des boucles de chromatine est encore méconnu. Plusieurs hypothèses ont été proposées, le premier rôle de ce réarrangement spatial serait de faciliter l'activation de la transcription comme cela a été décrit pour les locus  $\beta$ -globine, TH2, SEN1 ou BUD3. De plus, les boucles de chromatine ne sont pas limitées à l'activation des gènes. En effet, la présence de boucles a été décrite lors de la répression du gène Kit (Jing et al., 2008). Néanmoins, la question selon laquelle les boucles de chromatine sont une cause ou une conséquence de l'initiation de la transcription reste toujours ouverte. La seconde hypothèse serait l'implication des boucles de chromatine dans le phénomène de mémoire transcriptionnelle. Ce phénomène consisterait en une réactivation rapide de la transcription après un court arrêt, sans avoir besoin d'un stimulus (Deng and Blobel, 2010). Ce mécanisme a été décrit chez la levure, avec la formation d'une boucle de chromatine entre le promoteur et le terminateur des gènes HXK1 et GAL1FMP27, permettant la réactivation rapide de la

transcription (Laine et al., 2009) (Tan-Wong et al., 2009). La formation de boucle de chromatine a également été associée à l'orientation bidirectionnelle de la transcription (Tan-Wong et al., 2012). Lors de la transcription de promoteur bidirectionnel par l'ARN pol II, des ARN messagers sont synthétisés lorsque le gène est transcrit dans le bon sens. En cas de transcription en sens inverse, des ARN non codant (ncARN) peuvent alors être synthétisés. Il a été démontré, que la présence de boucle de chromatine permettait l'orientation de la transcription, afin de synthétiser des ARNm. Au contraire, lors de la production d'ARNnc, les boucles de chromatine sont absentes pour le gène FMP27 de la levure (Tan-Wong et al., 2012).



**Figure 21** Schéma de boucle de chromatine entre promoteur-enhancer et promoteur-terminateur d'un même gène.

En plus de ces mécanismes de boucle de chromatine, d'autres facteurs, appelés facteurs de remodelage ATP-dépendant, peuvent être impliqués dans le remodelage de la chromatine. Ces facteurs utilisent l'énergie de l'ATP pour réarranger la chromatine, en modifiant les liens histones/ADN. Ils font partie de complexes macromoléculaires se divisant en 4 familles : SWI/SNF (*switching defective / sucrose non fermenting*), ISWI (*imitation switch*), CHD (*chromodomain helicase DNA binding*) et INO80 (*inositol requiring 80*) (Clapier and Cairns, 2009). Ces complexes présentent une affinité pour le nucléosome et permettent son déplacement, son éjection ou la modification de sa composition. Ces facteurs sont impliqués dans l'assemblage de la chromatine, la réplication, la transcription et la réparation de l'ADN. Par exemple, le facteur NuRD (*nucleosome remodeling*), qui fait partie de la famille des CHD, renforce la répression de la transcription au niveau de l'ADN méthylé (Zhang et al., 1999b). Le complexe SWI/SNF est quant à lui impliqué dans l'activation de la transcription, en enlevant des nucléosomes, afin de rendre l'ADN accessible (Mohrmann and Verrijzer, 2005).

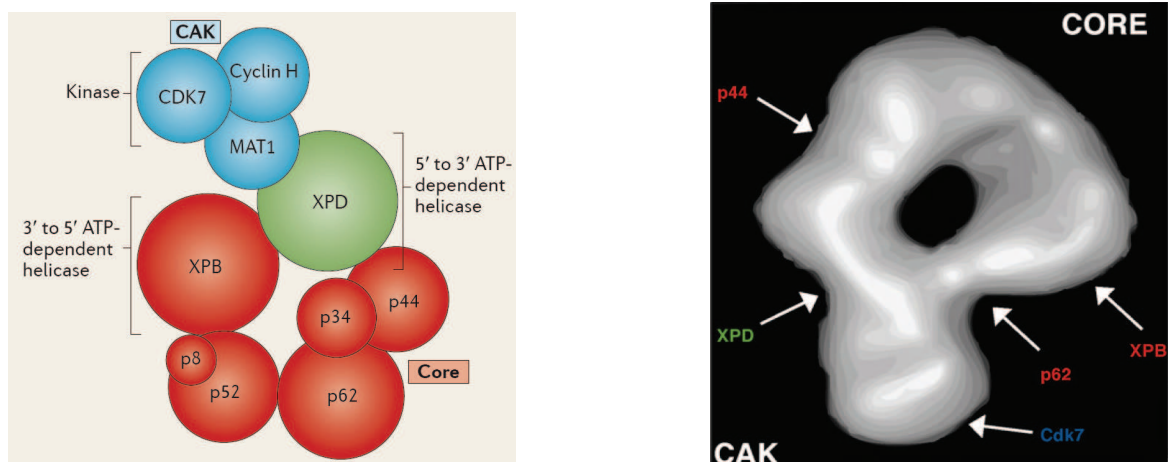
Pour résumer, tous ces facteurs pouvant modifier la structure de la chromatine agissent simultanément, pour permettre une régulation fine et dynamique de la structure chromatinienne. Cela permet le contrôle de multiples fonctions cellulaires essentielles à la vie de la cellule.

## IV. Les liens entre transcription et réparation de l'ADN

### 1. TFIIH

#### A. Composition de TFIIH

TFIIH est un complexe multi-protéique, qui fut découvert en 1989, sous le nom de facteur général de transcription d (chez le rat) (Conaway and Conaway, 1989) ou du facteur de transcription basique 2 BTF2 (cellules HeLa) (Gerard et al., 1991) comme indispensable à la transcription *in vitro*. Puis il fut isolé dans la levure en 1992, et fut nommé TFIIH (Flores et al., 1992). Il s'agit d'un des plus gros et plus complexe facteur de transcription. En effet, il est composé de 10 sous-unités, organisées en deux sous-complexes : le cœur (*core*), qui est composé de six sous-unités (XPB, p62, p52, p44, p34 et p8) et le CAK (*cdk-activating kinase*), qui lui est formé de trois sous-unités (CDK7, cyclin H et MAT1) (Roy et al., 1994) (Adamczewski et al., 1996) (Rossignol et al., 1997) (Figure 21). Ces deux sous-complexes sont reliés par une dixième sous-unités : XPD, qui interagit avec p44 et MAT1 (Coin and Egly, 1998).



**Figure 22 (A)** Représentation schématique de la structure de TFIIH (Compe and Egly, 2012a) avec le cœur (en rouge) et le CAK (en bleu). **(B)** Structure de TFIIH en microscopie électronique (Egly and Coin, 2011).

La microscopie électronique et l'imagerie de TFIIH, ont mis en évidence une forme en anneau (*ring like structure*), avec une cavité centrale permettant des interactions avec l'ADN (Schultz et al., 2000) (Figure 21). TFIIH est un facteur qui présente plusieurs activités enzymatiques, par l'intermédiaire des deux hélicases XPB et XPD, et de la kinase CDK7. Ces enzymes ont des rôles multiples, dans divers processus cellulaires (Tableau 4).

Le cœur de TFIIH est composé de XPB, un polypeptide de 89 kDa présentant une activité hélicase ATP dépendante 3'-5' (Ma et al., 1994), jouant un rôle dans la réparation et la transcription. Ce cœur est également composé des sous-unités dites « structurales » : p62, p52, p44 et p34. Des études récentes ont mis en évidence de nouvelles fonctions pour ces protéines. En effet, p52 interagit directement avec XPB et potentialise son activité ATPase (Fregoso et al., 2007). La dernière sous-unité du cœur de TFIIH, appelée TTDA ou p8, a été découverte dans la levure sous le nom de Tfb5 (Ranish et al., 2004), puis chez l'homme (Giglia-Mari et al., 2004). Cette sous unité stimule l'activité ATPase de XPB et favorise le recrutement de XPA (Coin et al., 2006). Il semble également, que p8 agit comme un stabilisateur de TFIIH, car on observe une forte diminution de la concentration cellulaire de TFIIH lorsqu'il est muté (Vermeulen et al., 2000).

Sous-complexe TFIIH	Humain	Levure	Fonction	Maladie génétique humaine
Cœur	XPB	Ssl2	Hélicase ATP dépendante 3'-5'	TTD et XPC/CS
	p62	Tfb1	Fonction structurale et interaction avec facteurs NER et de transcription	
	p52	Tfb2	Régule l'activité ATPase de XPB	
	p44	Ssl2	E3 ubiquitine ligase (dans levure)	
	p34	Tfb4	Fonction structurale et forte interaction avec p44	
	p8 (TTDA)	Tfb5	Régule l'activité ATPase de XPB	TTD
XPD	XPD	Rad3	Hélicase ATP dépendante 5'-3' et lien entre CAK et cœur	TTD, XP et XPC/CS
CAK	CDK7	Kin28	Kinase	
	Cycline H	Ccl1	Module l'activité kinase de CDK7	
	MAT1		Stabilise le CAK	

**Tableau 4** Composition du complexe humain TFIIH (adapté de (Compe and Egly, 2012a)).

Le CAK est composé de la kinase CDK7, qui intervient dans le cycle cellulaire, en phosphorylant les autres CDKs impliquées dans ce cycle. CDK7 phosphoryle aussi le CTD (*carboxy-terminal domain*) de l'ARN polymérase II lors de la transcription (Buratowski, 2009) (Lu et al., 1992). La seconde sous-unité est MAT1 (ménage à trois 1). Elle assure la stabilité du CAK, en interagissant avec CDK7 et cycline H, et stimule l'activité kinase de CDK7 (Busso et al., 2000). La dernière sous-unité est la cycline H, qui peut être phosphorylée par CDK8 du médiateur, engendrant une répression de l'activité de CDK7 (Akoulitchev et al., 2000). XPD est une sous-unité de TFIIH, qui possède une activité hélicase ATP dépendante 5'-3'. Elle permet d'assurer la stabilité entre le cœur et le CAK (Sandrock and Egly, 2001). L'étude de son activité dans la NER a été facilitée par le fait que, de nombreux patients XP et TTD présentent des mutations de XPD. Des mutations du domaine hélicase de XPD génèrent une diminution de la capacité de XPD à ouvrir l'ADN, lors de la NER (Coin et al., 1998). Il a été également suggéré, que XPD participait via son activité ATPase, à la reconnaissance du dommage et à sa vérification (Mathieu et al., 2010).

### **B. TFIIH et la réparation par excision de nucléotide**

Le rôle principale de TFIIH dans la NER est l'ouverture de l'ADN autour de la lésion, pour pouvoir permettre l'excision du dommage et son remplacement par un nouveau brin d'ADN. TFIIH est impliqué dans différentes étapes de la NER, de part l'activité de différentes sous-unités qui le composent.

TFIIH est recruté au niveau de la lésion, par le facteur NER XPC, qui couvre le côté 3' de la lésion, et laisse ainsi le 5' complètement libre (Min and Pavletich, 2007). XPC déstabilise la double hélice, provoquant une distorsion, permettant le recrutement de TFIIH. Ce recrutement est rendu possible par l'interaction entre deux sous-unités de TFIIH (p62-XPB) et XPC (Bernardes de Jesus et al., 2008). Une fois recruté au niveau du dommage, TFIIH va jouer un rôle via ses deux sous-unités hélicases XPB et XPD. Néanmoins, il a été démontré que l'hélicase XPD est indispensable à la NER, alors que XPB ne l'était pas (Coin et al., 1998) (Dubaele et al., 2003). En effet, lors de la NER, l'activité ATPase de XPB va intervenir en permettant l'ancrage de TFIIH sur la chromatine (Oksenysh et al., 2009). XPB va réguler l'ouverture de l'ADN, en assurant la bonne fixation de XPD. Celui-ci pourra alors, via son activité hélicase 5'-3', vérifier l'ADN et assurer l'ouverture de la double hélice.

Une troisième sous-unité de TFIIH, la sous-unité p8, est aussi impliquée dans ce processus. p8 stimule l'activité ATPase de XPB et favorise le recrutement de XPA sur le dommage (Coin et al., 2006). p8 agit également comme un stabilisateur de TFIIH (Giglia-Mari et al., 2004). Suite au recrutement de XPA, le CAK de TFIIH est relargué de l'ADN, puis après incisions par XPF en 5' et XPG en 3', le core de TFIIH est éliminé de la double hélice pour être recyclé.

### C. TFIIH : un facteur général de transcription

TFIIH est un facteur basal de transcription, qui intervient à différentes étapes de ce processus. En effet, il est impliqué dans l'initiation de la transcription, dans l'échappée du promoteur et dans l'élongation (Dvir et al., 2001), mais aussi dans la ré-initiation de la transcription, après le pausing de l'ARN polymérase II. L'initiation de la transcription est une étape qui résulte du recrutement séquentiel de différents facteurs (TFIID, TFIIA, TFIIIB ...) parmi lesquels TFIIH. Une fois le pic formé, deux sous-unités de TFIIH vont être nécessaires à la transcription par l'ARN polymérase II.

Dans un premier temps, l'activité ATPase de XPB est nécessaire pour ouvrir l'ADN autour du site d'initiation (Kim et al., 2000). Dans un second temps, CDK7 phosphoryle la sérine 5 du CTD (*carboxy-terminal deoxynucléotide*) de Rpb1, une sous-unité de l'ARN polymérase II dans le but de contrôler l'initiation de la transcription (Lu et al., 1992) (Serizawa et al., 1993). Cette phosphorylation est présente lors de l'initiation de la transcription, et est éliminée par la phosphatase Rtr1, lors de l'élongation (Mosley et al., 2009). La sous-unité CDK7 participe également à la transactivation de gènes dépendant des hormones, par la phosphorylation des récepteurs nucléaires comme le récepteur à l'acide rétinoïque  $\alpha, \beta$  ... (Rochette-Egly et al., 1997), ou encore le récepteur thyroïde (Compe et al., 2007).

En plus de son rôle dans la transcription par l'ARN polymérase II, TFIIH participe également à la transcription ribosomale par l'ARN polymérase I (Iben et al., 2002), mais son rôle exact reste encore flou.

## 2. Les autres facteurs

D'autres facteurs présentent cette dualité de fonction et ont été décrits comme jouant un rôle dans la transcription et la réparation de l'ADN (Le May et al., 2010a) (Kamileri et al., 2012). La protéine CSB est l'un de ces facteurs. En effet, CSB a d'abord été décrite pour son rôle dans la réparation couplée à la transcription (TC-NER). Outre cette implication dans la réparation de l'ADN, différents travaux de recherche ont démontré l'implication de CSB dans la transcription. Le facteur CSB est une ATPase dépendante de l'ADN, de la famille SWI/SNF, étant capable de remodeler la chromatine en interagissant avec le cœur du nucléosome (Citterio et al., 2000). CSB a également été décrit, comme jouant un rôle dans le maintien et le remodelage de la chromatine (Newman et al., 2006). De plus, CSB interviendrait dans la transcription, en interagissant avec l'ADN, la pol II et l'ARN naissant (Tantin et al., 1997), puis stimulerait l'élongation de la transcription *in vitro* (Selby and Sancar, 1997) (Proietti-De-Santis et al., 2006). En plus de son rôle dans la transcription dépendante de l'ARN pol II, il semble que CSB permette l'activation de la transcription dépendante de pol I, en interagissant avec la polymérase I et la méthyltransférase G9a. La conséquence de cette coopération serait la méthylation de H3K9, couplée à l'association de HP1, permettant la transcription efficace de l'ADN ribosomique (Yuan et al., 2007). CSB serait aussi impliquée dans la transcription pol III dépendante. En effet, son absence provoquerait la fragilité des chromosomes métaphasiques des gènes U1, U2 et 5S (Yu et al., 2000). Tous ces résultats ont permis de montrer que CSB n'est pas uniquement impliqué dans la réparation, mais intervient aussi dans la régulation de la transcription dépendante des trois classes de polymérases. Ce rôle transcriptionnel de CSB permettrait de mieux comprendre certains symptômes de patients CS.

L'endonucléase XPG est un exemple majeur de facteur NER, jouant un rôle dans différents processus cellulaires. Il s'agit d'un facteur décrit dans un premier temps, pour son rôle dans la réparation par excision de nucléotides (NER) dans laquelle il induit l'incision de l'ADN en 3'. La mutation du gène XPG peut être à l'origine de patients CS ou XP/CS (Clarkson, 2003). XPG est impliqué dans d'autres processus que la réparation. En effet, un homologue de XPG chez la levure, Rad2, est décrit comme intervenant dans la transcription des gènes GAL7 et GAL10 (Lee et al., 2002). Il a ensuite été montré que XPG, pouvait former un complexe stable

avec TFIIH via la sous-unité XPD ; ce complexe est actif en transcription et en réparation (Ito et al., 2007). En effet, des mutations de XPG, retrouvées chez des patients XP-G/CS, perturbent l'architecture de TFIIH. Cela provoque une absence de phosphorylation des NRs, et une dérégulation de l'expression de gènes contrôlés par les NRs. Ces démonstrations de l'intervention de XPG dans la transcription, permettraient de mieux comprendre les phénotypes de certains patients XP/CS.

Récemment, le recrutement de facteurs de réparation au niveau du promoteur de gènes pol I dépendants a été présenté comme aidant l'incision et le remplacement de meC, par des nucléotides déméthylés (Schmitz et al., 2009). En effet, le facteur *Gadd45α* (*growth arrest and DNA damaged inducible protein 45 alpha*) serait recruté sur le promoteur de gènes pol I dépendants, via une interaction avec TAF12, afin de permettre la déméthylation de l'ADN ribosomique. *Gadd45α* recruterait alors les facteurs NER (XPA, XPG et XPF), pour éliminer les cytosines méthylées au niveau du promoteur (Schmitz et al., 2009).

A la vue de ces nombreux exemples de facteurs intervenant aussi bien dans la réparation que dans la transcription, je me suis interrogé sur l'implication de tous les facteurs NER dans le mécanisme de transcription. J'ai également étudié le lien qu'il peut exister, entre des défauts de la transcription par des mutations des facteurs NER, et une meilleure compréhension des phénotypes de patients XP, CS, XP/CS ou TTD.



## RESULTATS

## LISTE DES PUBLICATIONS

1. **NER factors are recruited to active promoters and facilitate chromatin modification for transcription in the absence of exogenous genotoxic attack.**

*Molecular Cell, 2010*

Le May N, Mota-Fernandes D, Vélez-Cruz R, Iltis I, Biard D, Egly JM.

2. **XPG and XPF endonucleases trigger chromatin looping and DNA demethylation for accurate expression of activated genes.**

*Molecular Cell, 2012*

Le May N, Fradin D, Iltis I, Bognères P, Egly JM.

3. **The NER factor XPC regulates histone modifications on promoter of activated genes**

*En préparation*

Iltis I, Le May N, Egly JM.

### En annexe

4. **Poly (ADP-ribose) glycohydrolase regulates retinoic acid receptor-mediated gene transcription.**

*Molecular Cell, 2012*

Le May N, Iltis I, Amé JC, Zhovmer A, Biard D, Egly JM, Schreiber V, Coin F.

5. **The ARCH domain of XPD, an anchoring platform for CAK that conditions TFIIH DNA repair and Transcription Activities.**

*Soumis à PNAS*

Abdulrahman W, Iltis I, Radu L, Braun C, Giraudon C, Egly JM, Poterszman A.

## Publication 1

### **NER factors are recruited to active promoters and facilitate chromatin modification for transcription in the absence of exogenous genotoxic attack.**

Le May N., Mota-Fernandes D., Vélez-Cruz R., [Iltis I.](#), Biard D., Egly JM.

**Molecular Cell** (Avril 2010)

#### **Résumé**

Certains facteurs NER, initialement caractérisés dans la réparation de l'ADN, ont été décrits dans d'autres processus cellulaires. En utilisant un système de gènes sous le contrôle de récepteurs nucléaires (RAR $\beta$ 2 et PPAR), et par CHIP (Immunoprécipitation de la chromatine), nous avons observé lors de la transcription, le recrutement de l'ARN polymérase II, de ces partenaires, de TFIIH ainsi que des facteurs NER au niveau du promoteur. Ce recrutement observé en absence d'attaques génotoxiques, est sensible aux inhibiteurs de la transcription (par exemple le DRB).

La transactivation de ces gènes est associée au recrutement séquentiel des facteurs NER : XPC, XPA, RPA, CSB, XPG et XPF de manière concomitante avec le recrutement de la machinerie transcriptionnelle. Le recrutement de tous les facteurs NER est dépendant de la présence de la protéine XPC au promoteur.

Nous avons également montré que le co-recrutement des facteurs NER avec la machinerie transcriptionnelle, influence l'environnement chromatinien du promoteur actif, en induisant une déméthylation active de l'ADN autour du promoteur et des modifications post-traductionnelles des histones (H3K4me, H3K9ac). Ces modifications induisent une conformation dite « ouverte » de la chromatine (euchromatine) et permettent une initiation efficace de la transcription. De telles observations ont également été montrées chez les patients XP et pourraient expliquer certains de leurs symptômes.

# Molecular Cell

Volume 38  
Number 1

April 3, 2010

[www.cellpress.com](http://www.cellpress.com)

NUCLEOTIDE EXCISION REPAIR

CSB

TFIIH

XPF

XPA

XPG

DNA

CHROMATIN

RNA POL II TRANSCRIPTION

METHYLATION

NUCLEAR RECEPTORS

Nucleotide Excision Repair  
Factors in Transactivation

# NER Factors Are Recruited to Active Promoters and Facilitate Chromatin Modification for Transcription in the Absence of Exogenous Genotoxic Attack

Nicolas Le May,<sup>1,3</sup> David Mota-Fernandes,<sup>1,3</sup> Renier Vélez-Cruz,<sup>1</sup> Izarn Ittis,<sup>1</sup> Denis Biard,<sup>2</sup> and Jean Marc Egly<sup>1,\*</sup>

<sup>1</sup>Institut de Génétique et de Biologie Moléculaire et Cellulaire, CNRS/INSERM/ULP, BP 163, 67404 Illkirch Cedex, C. U. Strasbourg, France  
<sup>2</sup>INSERM U602/CEA-IRCM, Hôpital Paul Brousse 12-16 Avenue Paul Vaillant Couturier 94807 Villejuif Cedex, France

<sup>3</sup>These authors contributed equally to this work

\*Correspondence: [egly@igbmc.fr](mailto:egly@igbmc.fr)

DOI 10.1016/j.molcel.2010.03.004

## SUMMARY

Upon gene activation, we found that RNA polymerase II transcription machinery assembles sequentially with the nucleotide excision repair (NER) factors at the promoter. This recruitment occurs in absence of exogenous genotoxic attack, is sensitive to transcription inhibitors, and depends on the XPC protein. The presence of these repair proteins at the promoter of activated genes is necessary in order to achieve optimal DNA demethylation and histone posttranslational modifications (H3K4/H3K9 methylation, H3K9/14 acetylation) and thus efficient RNA synthesis. Deficiencies in some NER factors impede the recruitment of others and affect nuclear receptor transactivation. Our data suggest that there is a functional difference between the presence of the NER factors at the promoters (which requires XPC) and the NER factors at the distal regions of the gene (which requires CSB). While the latter may be a repair function, the former is a function with respect to transcription unveiled in the current study.

## INTRODUCTION

Protein-coding gene expression requires the presence of a battery of proteins, which includes RNA polymerase II (RNA pol II), the transcription factors (such as nuclear receptors, NR), coactivators, and mediator and histone-modifying enzymes at the promoter of activated genes. The formation of the transcription preinitiation complex (PIC) is accompanied by an important chromatin-remodeling phase including histones posttranslational modifications (PTMs) and active DNA demethylation of the promoter region. These processes produce an euchromatin environment necessary to begin the synthesis of the primary transcript (Cedar and Bergman, 2009; Li et al., 2007). If the DNA is not properly demethylated or is damaged by genotoxic agents, or if one of the members of the transcriptional machinery is missing, RNA synthesis becomes defective (Barreto et al., 2007; Gramantieri et al., 2005).

To maintain genome integrity and ensure the continuation of transcription, DNA lesions originated by UV irradiation, anti-tumor drugs, and other environmental products that modify the DNA structure, are eliminated through two subpathways of NER (Lindahl and Wood, 1999). Global genome repair (GGR) removes DNA damage from the entire genome, and the transcription-coupled repair (TCR) corrects DNA lesions located on the actively transcribed genes (Fousteri and Mullenders, 2008; Hanawalt and Spivak, 2008). In GGR, the removal of lesions requires their recognition by the repair factor XPC/HR23b and the subsequent opening of the DNA duplex by TFIIH (Riedl et al., 2003; Sugasawa et al., 2001). The resulting single-stranded structure is stabilized by XPA and RPA. XPG is recruited through its interaction with TFIIH on the 3' side of the lesion and its positioning on the cut site requires RPA. The interaction between XPA and ERCC1 stimulates the recruitment of ERCC1-XPF on the 5' side of the DNA lesion. In such condition the damaged oligonucleotide can be removed following the double incision by XPG and ERCC1-XPF endonucleases, thus allowing the DNA resynthesis (Wakasugi et al., 1997; Constantinou et al., 1999). In TCR, these factors (except XPC/HR23B) are recruited by the stalled RNA pol II in front of the damage with the help of the CSB and CSA proteins (Laine and Egly, 2006).

Mutations in 11 genes (XPA-G, ERCC1, TTD-A, CSA, and CSB), among the 30 genes involved in NER, have been associated with the human genetic disorders xeroderma pigmentosum (XP), trichothiodystrophy (TTD), cockayne syndrome (CS), and cerebro-oculo-facio-skeletal syndrome (COFS). Patients of these syndromes display a wide variety of clinical features (Lehmann, 2003; Jaspers et al., 2007). XP is characterized by sun sensitivity and greater than 1000-fold increased risk of cutaneous neoplasms. Approximately 30% of affected individuals have neurologic symptoms, including acquired microcephaly, diminished deep tendon stretch reflexes, and progressive cognitive impairment. TTD is characterized by abnormally sulfur-deficient brittle hair and accompanied by ichthyosis, in addition to neurological abnormalities. CS patients are sensitive to sunlight, have short stature, and display traits of premature aging. The majority of these clinical features cannot be explained by a DNA-repair deficiency and argue for these repair factors performing others functions beyond their role in the repair process. Besides TFIIH, whose role, in transcription and DNA repair, is well accepted (Schaeffer et al., 1993), others factors

involved in NER are also linked to the transcription process. For instance, mutations in XPG disturb the architecture of TFIIH and as a consequence affect NR-dependent transactivation (Ito et al., 2007). A similar scenario was described for XPD mutations (Compe et al., 2005, 2007; Keriél et al., 2002).

We therefore questioned whether others NER factors play a role in the transcriptional process, which can be distinguished from their functional role in DNA repair. In the present study, we found that PIC formation preceded the sequential recruitment of the NER factors XPC, XPA, RPA, XPG, and XPF/ERCC1 at the promoters of inducible genes, in the absence of exogenous genotoxic attack. All these NER factors (except CSB) were required to allow histone PTMs and active DNA demethylation necessary for efficient transcription. Finally, a transcriptional dysregulation was detected in cells with silenced NER factors and cells derived from XP patients.

## RESULTS

### NER Factors Are Recruited on Active Promoters

Employing systems where the nuclear receptors (such as the retinoic acid receptor [RAR] and peroxisome proliferator-activated receptor [PPAR]) were the key regulatory elements, and a chromatin immunoprecipitation (ChIP)-based methodology, we studied the dynamic recruitment of RNA pol II known partners and NER factors during transcription in the absence of exogenous genotoxic attack. Eight hours posttreatment of HeLa cells with all-trans retinoic acid (t-RA), we observed the accumulation of *RARβ2* mRNA (Figure 1A). ChIP followed by quantitative PCR showed at 8 hr a concomitant recruitment of RAR, the retinoid X receptor (RXR) partner, and RNA pol II (Figures 1F and 1G), that correlates with the peak of mRNA synthesis. At this time, TFIIH was also recruited as visualized by the presence of its subunits XPB, XPD and p44 (Figure 1G). While t-RA treatment resulted in a cyclical recruitment of various components of the transactivation complex at the target *RARβ2* promoter, productive RNA synthesis only occurred at 8 hr, which coincided with the concomitant arrival of the transcription intermediary factor 2 (TIF2) coactivator, the mediator subunit Med6 (Figure 1F), Med1 and the p300 coactivator (data not shown), all involved in the formation of RAR transactivation complex. On the *RARβ2* promoter, we also detected the presence of the XPG endonuclease, which interacts with TFIIH and plays a role in transactivation (Ito et al., 2007). Unexpectedly, we also detected XPF, the other NER endonuclease (Figure 1H). Moreover, we observed the presence of XPA and RPA, two NER factors with no known function in the transcription process (Figure 1H). Surprisingly, we also detected on the promoter, the CSB protein, required for TCR, and XPC, the damage-sensing protein of GGR (Figure 1I).

We next questioned whether the NER factors that we observed on the *RARβ2* promoter were also part of the elongating transcription machinery. ChIP analysis showed the presence of XPA, RPA, XPG, XPF, and CSB together with the elongating RNA pol II and the transcription elongation factor TFIIF at 8 hr after t-RA treatment, on exon 3, 4 (Figures S1B–S1G), and exon 6 (Figures 1K–1M). Neither RAR nor RXR were detected at the elongation regions of the *RARβ2* gene

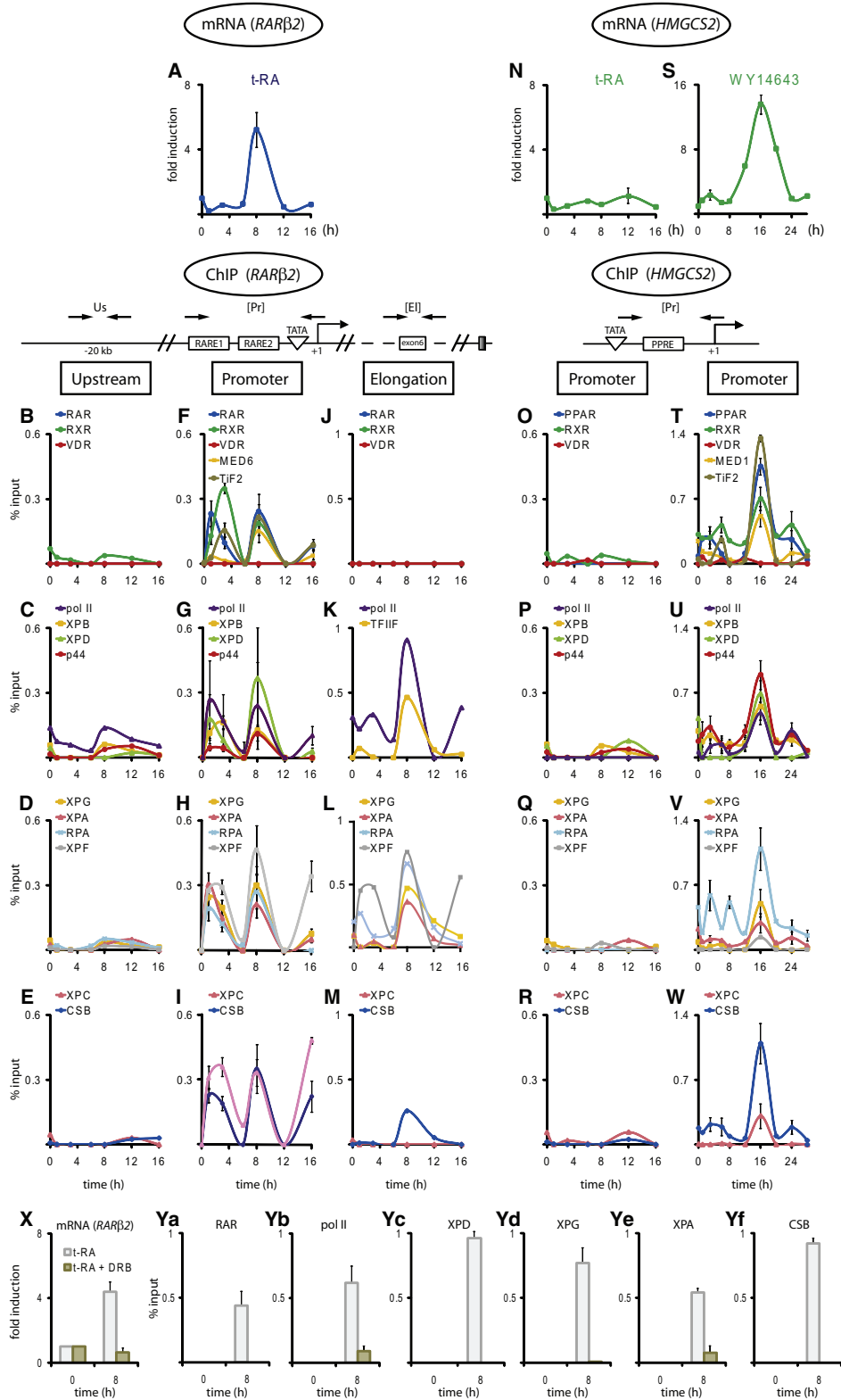
(Figure 1J). It must be noticed that under our experimental conditions, we only detected background levels of XPC at the elongation regions (Figure 1M, S1D, and S1G). To ensure that the recruitment of the repair factors was specific to the transactivation process, we analyzed the promoter of an inactive gene. In t-RA-treated HeLa cells (in which PPAR $\alpha$  is not expressed), neither the transcription machinery nor the NER factors were found associated to the promoter of 3-hydroxy-3-methylglutaryl-coenzyme A synthase 2 (*HMGCS2*), a PPAR-responsive gene (Figures 1O–1R), and no accumulation of mRNA synthesis was observed (Figure 1N). On the contrary, when cells were transfected with a PPAR $\alpha$ -expressing vector and treated with WY14643 (PPAR $\alpha$ -specific ligand), *HMGCS2* mRNA was highly expressed at 16 hr posttreatment (Figure 1S). Concomitantly, we observed the recruitment of the transcription and all NER factors, together with PPAR/RXR on the promoter (Figures 1T–1W) and on the elongation regions (Figures S1I–S1L). To further demonstrate that the presence of the NER factors at the promoter did reflect a real association with the transcription machinery, we designed several control experiments. To test our ChIP protocol, we analyzed the presence of two unrelated proteins on the promoters of *RARβ2/HMGCS2*. Neither cyclin A (data not shown) nor the vitamin D receptor (VDR) (endogenously expressed in HeLa cells) were present either at the promoter or elongating regions (Figures 1F, 1J, 1O, 1T, S1B, S1E, and S1I). Second, although we detected low amounts of RNA pol II far upstream from the *RARβ2* initiation site (–20 kb), none of the factors required either for the formation of the RAR/RXR-dependent transactivation complex or for the NER reaction were detected at the *RARβ2* upstream promoter region (Figures 1B–1E). We cannot exclude the possibility that the low amounts of RNA pol II detected upstream the *RARβ2* gene represent another transcription complex of a gene not currently identified.

Simultaneous treatments of HeLa cells with the transcription inhibitor 5,6-Dichloro-1-beta-D-ribofuranosylbenzimidazole (DRB) and t-RA, resulted in the inhibition of the *RARβ2* mRNA synthesis (Figure 1X). Under those conditions, neither RNA pol II nor NER factors were detected on *RARβ2* promoter (Figures 1Ya–1Yf) and exon 6 (Figure S1N). Similar results were obtained when HeLa cells were treated with  $\alpha$ -amanitin, another transcription inhibitor, and simultaneously the t-RA ligand (Figures S1O and S1P).

Altogether, our data show that NER factors were recruited with RNA pol II at the promoters of activated genes and that this recruitment is specific to the transactivation process and sensitive to transcription inhibitors. In addition to *RARβ2* and *HMGCS2* genes, it is important to note that all these repair factors together with the transcription apparatus were observed upon activation of all the genes so far tested such as *RARα2*, the Vitamin D responsive *Cyp24* as well as the PPAR responsive liver fatty acid binding protein *LFABP* (data not shown).

### The NER Factors Associated to RNA Pol II Are Discriminated from a Repair Complex

We next questioned whether the NER factors that assembled on promoters of inducible genes could be distinguished from a “DNA repair complex” formed upon genotoxic attack. First,





HeLa cells were UV irradiated and treated with t-RA simultaneously. Under these conditions, expression of the growth arrest DNA damage-inducible 45 alpha (*GADD45 $\alpha$* ), as well as the human double minute 2 (*hDM2*), two UV-inducible genes (Adimoolam and Ford, 2002), was elevated at 4 hr posttreatment (Figure 2A) (data not shown). The *RAR $\beta$ 2* gene expression peak, originally found at 8 hr post-t-RA treatment, was delayed to 16 hr (Figure 2E, compared blue dotted line with blue line respectively; see also Figure 1A). This delay may be due (at least in part) to the priority of the DNA repair process (Smith and Hanawalt, 1978; Friedberg et al., 1995). ChIP analysis performed over time further showed that RNA pol II, TFIIH (XPB, XPD, and cdk7), as well as all XPC, XPA, RPA, XPG, and CSB were recruited at 4 hr at the *GADD45 $\alpha$*  promoter together with p53 (Figures 2B–2D), and at 16 hr to the *RAR $\beta$ 2* gene promoter together with RAR (Figures 2F–2H).

At this stage of our work, we could speculate that NER factors recruitment reflected repair event at the analyzed promoters. However, it seems hard to conceive that *GADD45 $\alpha$*  as well as all the promoters so far tested were damaged under our experimental conditions and consequently subjected to either a GGR and/or a TCR process. Indeed, we have also observed the corecruitment of NER factors with RNA pol II machinery at the dihydrofolate reductase (*DHFR*) housekeeping gene (data not shown). Additionally, it is worthwhile to notice that none of the PPAR, TFIIH, RNA pol II transcriptional components and NER factors was detected at the promoter of the nonactivated *HMGCS2* gene (Figures 2I–2L).

ChIP/reChIP analysis using specific antibodies showed that TFIIH and RNA pol II (ChIP-cdk7/reChIP-pol II), XPA and RNA pol II (ChIP-XPA/reChIP-pol II), as well as TFIIH and XPA (ChIP-XPB/reChIP-XPA and ChIP-cdk7/reChIP-XPA) co-occupied both the *GADD45 $\alpha$*  and *RAR $\beta$ 2* promoters at 4 hr and 16 hr post-UV irradiation and ligand treatment (Figures 2M and 2O). To validate our ChIP and ChIP/reChIP data, we used a combination of antibodies against an unrelated protein (VDR, see also Figures 2B, 2F, and 2J) and either XPA, RNA pol II, XPB, or cdk7. We were not able to detect any co-occupancy between VDR and either XPA, TFIIH (XPB and cdk7) or RNA pol II (Figures 2N and 2P). It is worthwhile to notice that the presence of XPA together with CAK/cdk7 (which is crucial for transcription) and the transcriptional machinery at the activated *RAR $\beta$ 2* promoter, argue for a complex unrelated to a GGR complex (see also below). Indeed, in GGR, the presence of XPA and cdk7 is mutually exclusive: the arrival of XPA at the damage sites leads to the removal of CAK from TFIIH core (Coin et al., 2008).

To further test the idea that NER factors containing transcription complex is different to the “repair complexes,” we designed a second set of experiments. MRC5 cells were UV irradiated and treated with the transcription inhibitor DRB. Under these conditions, and as previously observed for the *RAR $\beta$ 2* gene (Figures 1X–1Yf), transcription of *GADD45 $\alpha$*  gene was inhibited contrary to what occurred in cells only treated with UV in which the *GADD45 $\alpha$*  mRNA synthesis peaks at 2 hr (Figures 3A and 3B, blue dotted histograms). At that time, in the UV-irradiated cells, our ChIP and ChIP/reChIP analysis demonstrated that RNA pol II, NER factors as well as p53 transcription factor, were corecruited and co-occupied *GADD45 $\alpha$*  promoter (Figure 3A, colored histograms; Figures S2C–S2F) as similarly observed in UV-irradiated and t-RA treated HeLa cells (Figures 2B–2D). This corecruitment of RNA pol II/NER factors and RNA pol II/TFIIH was lost in UV-irradiated MRC5 simultaneously treated with DRB (Figures 3B and S2G–S2J), thus explaining the absence of mRNA synthesis. ChIP using antibodies against XPB, coupled to Western blot (Fousteri et al., 2006; Coin et al., 2008), allowed us to observe the formation and the composition of the NER repair complexes between 0.25 and 2/4 hr post UV irradiation on the chromatin extract (Figure 3C, lanes 2–6 and curves). These repair complexes were observed even when the cells were simultaneously treated with the transcription inhibitor DRB and UV (Figure 3D, lanes 2–6 and curves). Both in the presence and absence of DRB, we clearly observed the arrival of the NER factors such as RPA and XPF concomitantly to the release of the CAK subcomplex from TFIIH (Figures 3C and 3D compared with Figures S2K and S2L for the specificity of the immunoprecipitation). To further support our results suggesting that NER complexes at promoter are different than “repair complexes,” we used local UV irradiation technology combined with fluorescent immunostaining (Volker et al., 2001). Analysis of locally UV-irradiated MRC5 cells by confocal microscopy showed that XPC, XPB/TFIIH, and XPA colocalized at sites of UV damage even in the presence of DRB (Figures 3E and 3F).

Altogether, our data demonstrate that a GGR complex is resistant to transcription inhibitors whereas the transcription complex with NER factors we describe here is sensitive to DRB. Our results suggest that the NER factors-containing transcription complex is not a NER repair complex.

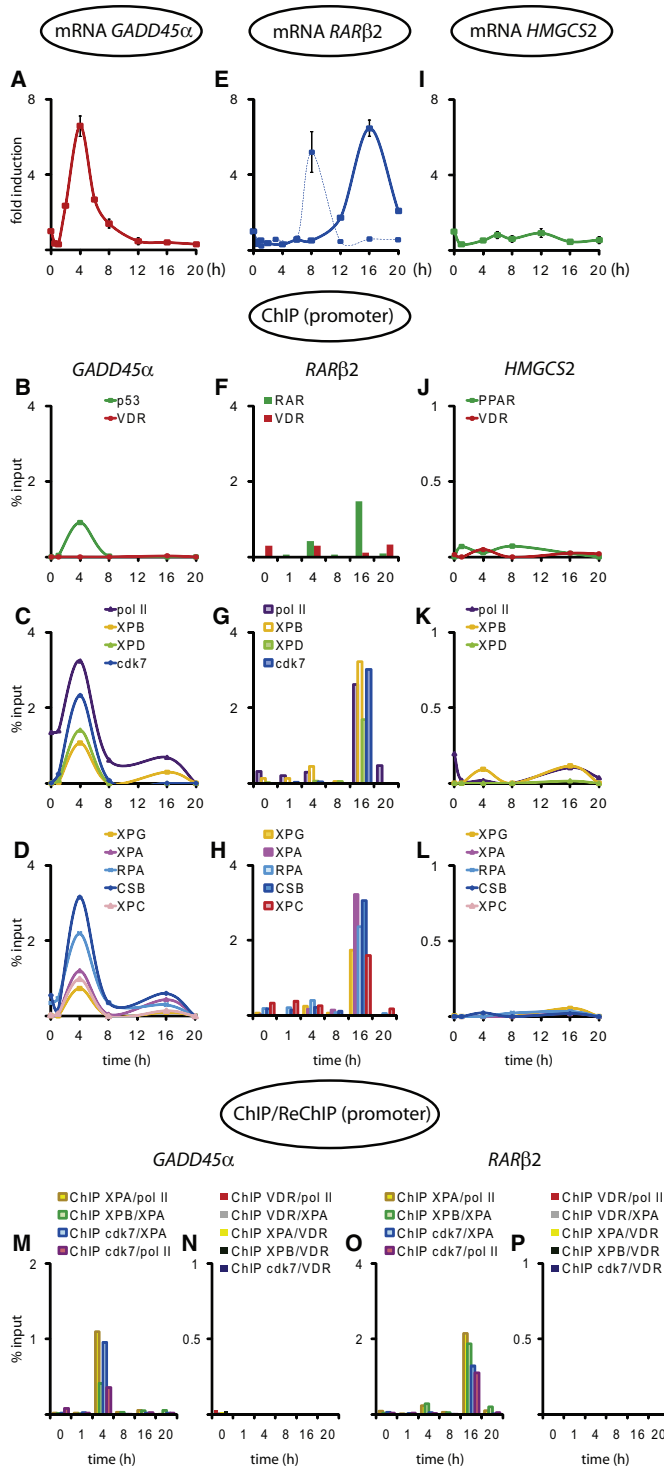
### Deficiencies in Some NER Factors Impede the Recruitment of Others

We next investigated how either the silencing of XPC, XPA, ERCC1 or XPG (using the corresponding “silencing HeLa” cells

### Figure 1. NER Factors Form Part of the Initiating and Elongating Transcription Machinery

(A–M) Relative mRNA expression of *RAR $\beta$ 2* gene from HeLa cells treated with t-RA (1  $\mu$ M) (A). Error bars represent the standard deviation of three independent experiments. Schematic representation of the *RAR $\beta$ 2* gene with the indicated amplicons designed at the upstream (Us), promoter (Pr), and elongation (El, exon 6) regions. ChIP monitoring the t-RA-dependent occupancy of: RAR, RXR, VDR, TIF2, and Med 6 (B, F, and J); RNA pol II, TFIIH subunits (XPD, XPB, and p44) and TFIIH (C, G, and K); XPG, XPA, RPA and XPF (D, H, and L); XPC and CSB (E, I, and M), on the different amplicons of *RAR $\beta$ 2* gene. (N–Y) Relative mRNA expression of *HMGCS2* gene from HeLa cells treated either with t-RA (N) or with WY14643 (1  $\mu$ M) and transfected with the pSG5-PPAR $\alpha$  (S). Error bars represent the standard deviation of three independent experiments. ChIP monitoring the t-RA- and WY14643-dependent occupancy of PPAR, RXR, VDR, TIF2, and Med1 (O and T); RNA pol II and TFIIH subunits (XPD, XPB, and p44) (P and U); XPG, XPA, RPA, and XPF (Q and V); XPC and CSB (R and W), on the *HMGCS2* promoter (Pr). Each series of ChIP is representative of at least two independent experiments. Values are expressed as percent input, which are the average of at least two qPCR reactions and error is within 10%. Relative *RAR $\beta$ 2* mRNA expression (X) and ChIP monitoring the occupancy of RAR, RNA pol II, XPD, XPG, XPA, and CSB (Ya–Yf) on the *RAR $\beta$ 2* promoter in t-RA treated HeLa cells in absence (gray histogram) or presence (green histogram) of DRB (100  $\mu$ M).





**Figure 2. NER Factors Are Corecruited with RNA pol II Machinery after UV Irradiation**

(A–L) Relative mRNA expression of *GADD45α* (A), *RARβ2* (E), and *HMGC52* (I) genes monitored by qPCR from UV-irradiated ( $20 \text{ J/m}^2$ ) HeLa cells treated with t-RA ( $1 \mu\text{M}$ ). Relative mRNA expression of *RARβ2* gene measured by qPCR from t-RA-treated cells (E, dotted blue line). A dose of  $20 \text{ J/m}^2$  generates around two photolesions per 10 kb of genomic DNA (van Hoffen et al., 1995). ChIP monitoring the UV- and t-RA-dependent recruitment of p53, RAR, PPAR, and VDR (B, F, and J) as indicated; RNA pol II and TFIIH subunits (XPB, XPD, cdk7) (C, G, and K); XPG, XPA, RPA, CSB and XPC (D, H, and L), on the *GADD45α* (curves), *RARβ2* (histograms), and *HMGC52* (curves) promoters. Each series of ChIPs is representative of two independent experiments. Values are expressed as percent input, as previously. (M–P) For ChIP/ReChIP experiments on *GADD45α* and *RARβ2* promoters, samples were subjected to either (M and O) a first IP against XPA, XPB, or cdk7 and then purified complexes were further subjected to a second IP using antibodies against either pol II or XPA or (N and P) a first IP using an antibody against an unrelated protein (VDR) and then purified complexes were subjected to a second antibody against either XPA, RNA pol II, XPB, or cdk7 or vice versa as indicated.

First, in SiXPC, -XPA, and -ERCC1 cells (Figure 4A) as well as in XPC/579st and XPA/R207st cells (Figures 5Aa and 5Ba), we observed a significant defect in the *RARβ2* mRNA expression compared to either the SiControl or the rescued cells, respectively. ChIP analysis from each of the defective cells, treated with t-RA, showed that, similarly to what was observed in SiControl, RAR, RXR, TFIIH (XPB and cdk7), and RNA pol II were concomitantly recruited to the *RARβ2* promoter (Figures 4B–4I, S4B, and S4C).

Contrary to the coordinated recruitment of XPC, XPA, XPG, XPF, and DDB1 (known to facilitate the recruitment of XPC on damaged chromatin [Fitch et al., 2003]), with the RNA pol II machinery observed in the SiControl cells, none of the NER factors were observed at the promoter in SiXPC cells (compare Figures 4J, 4N, and 4R with 4K, 4O, and 4S).

In absence of XPA (SiXPA), only XPC and CSB were detected at the promoter but neither XPF nor XPG (Figures 4L and 4P; see also Figures 4D and 4H). Remarkably, in patients fibroblasts bearing mutations on either XPC or XPA compared to XPA rescued cells, we observed a similar defect in the recruitment of NER factors on the *RARβ2* promoter while the transcriptional machinery was present (Figures 5Ab–5Ae, 5Bb–5Be, and 5Cb–5Ce). Interestingly, in SiERCC1 cells, where expression of ERCC1 and its partner XPF were silenced (Figure S3) (Gaillard and Wood, 2001), we detected XPC, XPA, and XPG together with RAR, RXR, RNA pol II, and TFIIH (Figures 4E, 4I, 4M, and 4Q). On both SiXPG and XP-G fibroblasts derived from XPCS1RO patient, all the NER factors (except XPG) were recruited to the corresponding activated promoters (Figures S4F–S4N). We also analyzed the consequences of CSB mutations in CS1AN fibroblasts in which the Q336st mutation abolished CSB expression (Figure S3). When compared to the CSB rescued cells, CSB-deficient cells

that stably express siRNA and abolish their expression, see Figure S3), or mutations in XPC, XPA, XPG, and CSB (using fibroblasts derived from XP or CS patients) would affect the t-RA dependent *RARβ2* transactivation.

displayed a similar *RARβ2* mRNA expression pattern (Figure 5Da). Moreover, the absence of CSB protein did not prevent the recruitment of XPC, XPA, XPG, and XPF together with the transcriptional machinery on the *RARβ2* promoter (compare Figures 5Db–5De with Figures 5Eb–5Ee).

Altogether, our data showed first that the recruitment of the basal transcription machinery occurred even in the absence of the NER factors and second that the NER factors were assembled on the promoter in the following order: XPC, XPA, followed by XPG and XPF. Moreover, it seems that the absence of CSB (required for TCR) was not required for the recruitment of the other NER factors and the *RARβ2* mRNA synthesis. This suggests, once again, that the transcription complex associated to NER factors at the promoter of activated gene can be discriminated from a DNA repair complex, including a TCR complex.

#### NER Factors and Gadd45 $\alpha$ Facilitate DNA Demethylation and Histones PTMs

Recent studies revealed that the Gadd45 $\alpha$  protein required XPG during active DNA demethylation, an essential step for efficient transcription (Barreto et al., 2007; Gramantieri et al., 2005). We thus investigated whether a defect in the recruitment of the NER factors on the *RARβ2* promoter could affect the Gadd45 $\alpha$  recruitment. ChIP analysis showed a recruitment of Gadd45 $\alpha$  concomitant with the NER factors on *RARβ2* promoter in SiControl cells (Figures 4N and S4E). In SiXPC, SiXPA, as well as in XPC/579st and XPA/R207st cells, Gadd45 $\alpha$  was not detected together with the other components at the promoter (Figures 4O, 4P, 5Ae, 5Be). Surprisingly, we observed its presence together either with XPG in SiERCC1 cells or with XPF in SiXPG and in XPG-deficient cells (Figures 4Q, S4J, and S4N). In CSB-deficient as well in CSB-rescued cells, Gadd45 $\alpha$  was recruited together with all the NER factors on the active *RARβ2* promoter (Figures 5De and 5Ee).

We next investigated whether the simultaneous presence of Gadd45 $\alpha$  together with the NER factors was necessary to replace the methylated cytosines by unmethylated ones required for efficient transcription initiation (Appanah et al., 2007). Since hypo- and hypermethylation of CpG dinucleotides across the genome contribute to changes in gene expression, we used the Imprint Methylated DNA quantification technique to measure global DNA methylation shifts. We first found that the global methylation level is much higher in genomic DNA from SiXPC, SiXPA, SiERCC1, and SiXPG HeLa cells compared to the SiControl cells, whereas no differences were observed between CSB-mutated patient fibroblasts and the corresponding rescued cells (Figures S5A and S5B, respectively).

We then used the methylation-sensitive enzymes HpaII and *Dpn I* to evaluate the methylated status of CpGs islands localized in the *RARβ2* promoter surroundings (Figure 4, lower panel) (Appanah et al., 2007; Santoro and Grummt, 2001). In wild-type cells (SiControl) as well in CS1AN, CS1AN/CSB, and XP12RO/XPA-rescued cells, the presence of all NER factors together with the RNA pol II machinery and Gadd45 $\alpha$ , at 6/8 hr post t-RA induction, was accompanied by a significant increase in the *Dpn I* sensitivity (thus DNA demethylation), while no variation was obtained with Hpa II (Figures 4V, 5Cg, 5Dg, and 5Eg). This reflected a decrease of methylation in CpG islands located in the 3'-proximal promoter region. On the contrary in SiXPC,

SiXPA, as well as in XPC/579st and XPA/R207st cells, in which we observed a defect in the recruitment of XPG, XPF and also the Gadd45 $\alpha$  protein, the CpG islands were not digested by *Dpn I* (Figures 4W, 4X, 5Ag, and 5Bg). Interestingly, in SiERCC1, in which XPG together with Gadd45 $\alpha$  were recruited, no significant changes in the CpG islands methylation were detected (Figure 4Y). This also occurs in SiXPG, where XPF together with Gadd45 $\alpha$  were recruited at the *RARβ2* promoter (data not shown).

DNA methylation and histones PTMs are dependent on one another. Hypermethylated CpG islands have been associated with a heterochromatin landmark such as di-/trimethylation of histone H3K9 (H3K9me); conversely, di-/trimethylation of histone H3K4 (H3K4me), acetylation of H3K9/14 (H3K9/14ac), related to active transcription, are accompanied by DNA hypomethylation (Cedar and Bergman, 2009; Hashimshony et al., 2003). We thus questioned whether a defect in the active demethylation (in absence of recruitment of NER factors) would also disturb histones PTMs. ChIP analysis revealed that in WT (SiControl) cells, the recruitment of the NER factors, Gadd45 $\alpha$ , and RNA pol II paralleled an increase of H3K4me, H3K9/14ac concomitantly to a decrease of H3K9me (Figure 4R). This profile was repeatedly observed at 8 hr in both CSB, XPA rescued, and CSB-deficient cells (Figures 5Cf, 5Df, and 5Ef). These “signatures” in histone H3 modifications were not observed in SiXPC, SiXPA, and SiERCC1 cells or in XPC/579st and XPA/R207st cells (Figures 4S, 4T, 4U, 5Af, and 5Bf). In those cells, we observed an increase of H3K9me and a decrease of H3K4me that paralleled the recruitment of transcriptional machinery, contrary to what occurs in wild-type cells.

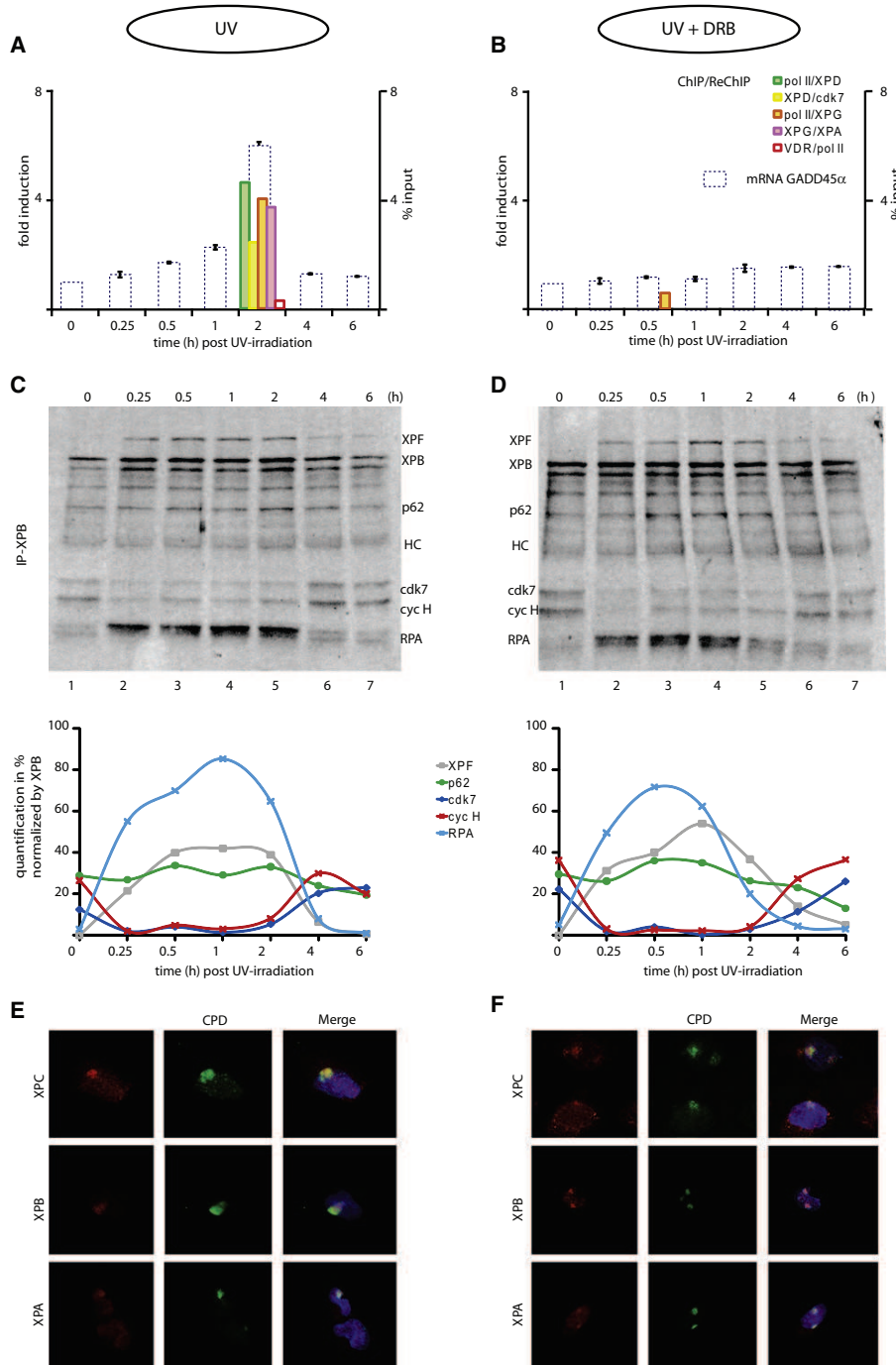
The above results strongly suggest that in addition to XPG, other NER factors (XPC, XPA, XPF/ERCC1, except CSB) are necessary for the Gadd45 $\alpha$  recruitment to the promoter of activated genes and the concomitant occurrence of DNA demethylation and histones PTMs.

#### DISCUSSION

Synthesis of mRNA is the result of a cascade of events that requires more than ~200 proteins that constitute the transcriptional machinery (Brivanlou and Darnell, 2002; Kornberg, 2007). How are each of these proteins connected in this intricate network to initiate transcription, at the right time and in the proper cell, would help to further our understanding of gene expression regulation.

#### The NER Factors Are Sequentially Recruited to the Activated Genes

In the present study we demonstrate that all the NER factors are associated with the transcription machinery on the promoters of active genes (in the absence of exogenous genotoxic attack) and escort the elongating RNA polymerase (Figure 1). The recruitment of the NER factors and their association with RNA pol II is abolished in the presence of the transcription inhibitor DRB (Figure 1). The presence of these factors is specific to factors that are involved in NER, since the base excision repair and mismatch repair proteins Ogg1 glycosylase and MutS homolog 2 factor (MSH2), respectively, were not detected (Figure S1Q).



**Figure 3. NER Factors-Containing Transcriptional Complex Can Be Discriminated from a DNA Repair Complex**

(A and B) Relative GADD45 $\alpha$  mRNA expression (blue dotted histograms) and ChIP/ReChIP (colored histograms) monitoring the coimmunoprecipitation of the corresponding promoter using antibodies combinations against RNA pol II/ XPD, XPD/cdk7, RNA pol II/XPG, XPG/XPA, and VDR/pol II from UV-irradiated (20 J/m<sup>2</sup>) MRC5 fibroblasts in absence (A) or presence (B) of DRB (100  $\mu$ M), harvested at indicated times.

(C and D) Western-blotting analysis of Ab-XPB ChIP samples from chromatin extracts of MRC5 fibroblasts incubated overtime after UV irradiation, in absence (C) or pretreated with DRB (100  $\mu$ M) during 6 hr (D). The WB signals for XPF, XPB, p62, cdk7, cyc H, and RPA were quantified using Genetool and plotted on the graphs. For each single lane, XPB was used as reference. HC indicates the heavy chain of the antibody.

Moreover, silencing or mutations in these factors revealed that the NER factors are sequentially recruited to these promoters in the following order: XPC, XPA, and XPG, XPF/ERCC1/Gadd45 $\alpha$ , once the transcription machinery is positioned into place (Figures 4, and 5; see also model Figure 6).

One thus can wonder whether this “entourage” of the RNA pol II would render the transcription machinery equipped to deal with potential threats (such as DNA lesions) or whether these factors participate in a mechanism that renders these promoters ready for efficient RNA synthesis. Importantly, this transcriptional complex equipped with NER factors is distinguishable from a repair complex, whether it would be TCR or GGR, in several ways. First, this transcription complex is sensitive to the transcription inhibitor DRB, whereas GGR is not (Figure 3). Second, this transcription complex does not require the CSB protein, whereas this protein is absolutely required for TCR (Figure 5). Third, this transcription complex requires XPC, whereas this protein is dispensable for TCR (Figures 4 and 5). Therefore, our data indicates that this transcription complex loaded with NER factors is different to a repair complex based on its composition and its sensitivity to transcription inhibitors.

Despite these results, two lines of evidence are surprising; in the absence of the CSB protein, we did not observe the presence of the NER factors at the distal regions of the genes (i.e., escorting the elongating RNA pol II) and in the absence of XPC, we did not observe the recruitment of the NER factors at the promoters of activated genes (Figure S6). Although these results may suggest that we are observing a TCR complex, analysis of the transactivation capacity of XP-C cells showed disrupted transcription, whereas the same analysis of CS-B cells (TCR deficient) did not show a transcriptional deficiency. Altogether, these data suggest that there is a functional difference between the presence of the NER factors at the promoters (which requires XPC) and the NER factors at the distal regions of the gene (which requires CSB). While the latter may be a repair function, the former is a function with respect to transcription unveiled in the current study.

#### The NER Factors Are Part of the Transcription Process

We then raised the question of the potential role of the NER factors in transcription at the promoters of active genes. The formation of the PIC is accompanied by important chromatin remodeling resulting from histone PTMs and DNA demethylation (Li et al., 2007). In the present study we observed the concomitant recruitment of the transcription machinery and the NER factors, Gadd45 $\alpha$  with the 3'-proximal promoter DNA demethylation, histone H3K4 di-/trimethylation, H3K9 demethylation, and H3K9/K14 acetylation (Figures 4 and 5) as well as H4K16 acetylation and H2B monoubiquitination (Figure S7) at active promoters. These chromatin “signatures” were not observed when one of the NER factors was absent (Figures 4 and 5). Indeed, in NER-deficient cells we observed unchanged DNA methylation levels and histones PTMs representative of a hetero-

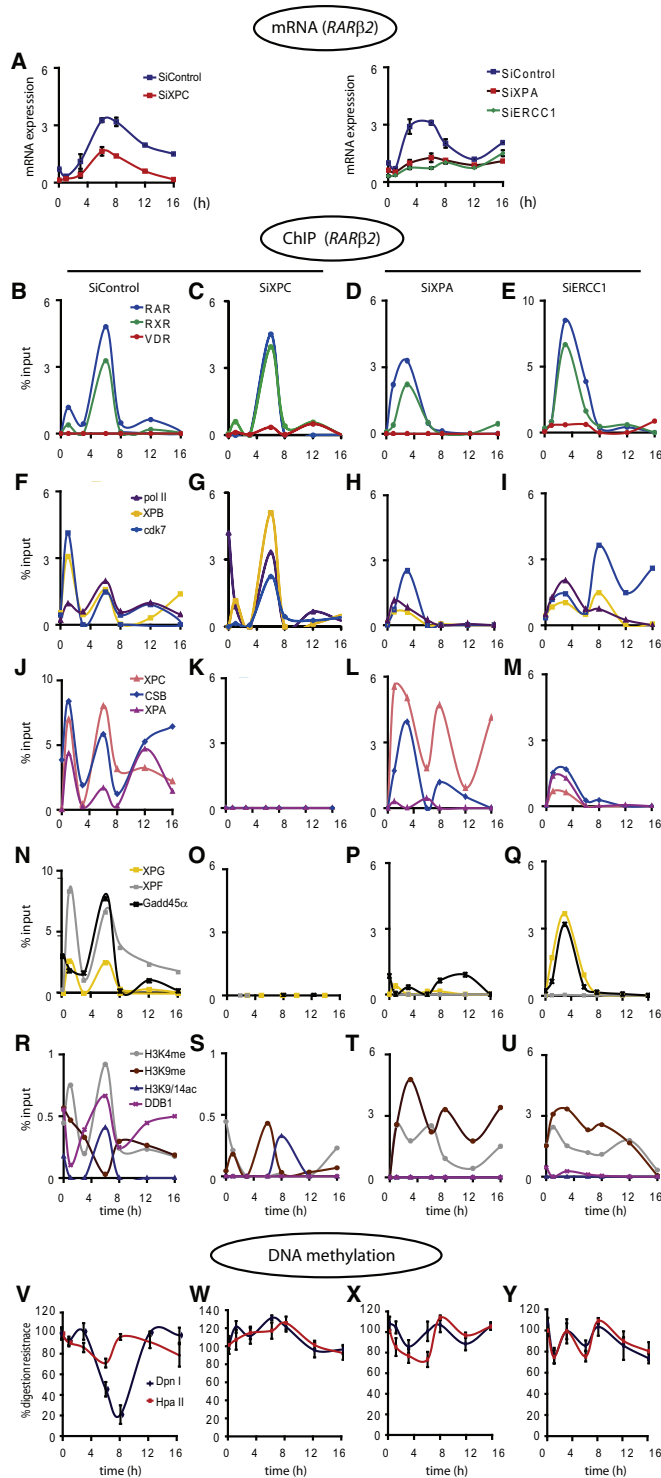
chromatin environment surrounding the promoters of activated genes, which could explain the dysregulation in mRNA synthesis despite the presence of the transcriptional machinery. It is possible that the various other enzymatic activities of the NER factors, such as the ATPase, helicase, kinase, E3-ubiquitin ligase, and/or endonuclease may participate in some local chromatin modifications as well.

#### The Transcriptional Phenotype of the XP Patients

This study demonstrates that others NER factors play a role in the transcription process. Although these factors may not be essential for the PIC formation, they clearly fine-tune the transactivation to an optimal level. This optimization may be very important during development and can potentially explain the clinical symptoms observed in XP patients that cannot be explained by a DNA-repair deficiency. For instance, a significant fraction of XP-A patients display neurological abnormalities and developmental problems. Why do these patients not display a more severe phenotype, like CS or XP/CS patients? We know XPB and XPD (thus TFIIH) are essential for the transcription process (Coin et al., 1999). TFIIH phosphorylates transcriptional activators and RNA pol II (Keriel et al., 2002; Lu et al., 1992), all of which are crucial for the formation of the PIC, and thus for RNA synthesis. On the other hand, we observed in the current work that XPA, XPC, and XPF are involved in the transcription process in a more indirect way and are thus less critical for this process (since the PIC formation is not affected). The stage at which these factors function during transcription may explain the differences in spectrum and severity between patients from different XP groups. There is also the possibility that others factors can perform redundant functions to optimize the transcription process. Finally, we should not ignore the fact that the phenotypes of XP patients are broad and although we categorize patients for a single mutation, recent work has shown that the secondary allele may also have an effect in the clinical symptoms (Ueda et al., 2009).

To summarize, our work allows us to propose a model in which, upon gene activation and in the absence of exogenous genotoxic attack, RNA pol II machinery, including TFIIH (XPB/XPD), is recruited at the promoter followed by the sequential arrival of the NER factors, concomitant histone modifications, and promoter DNA demethylation required for optimal RNA synthesis (Figure 6). Interestingly, this work points out once more the similarities between RNA pol II and RNA pol I transcription, which share in addition to some basal transcription factors, NER factors such as CSB, TFIIH, and XPG as well as Gadd45 $\alpha$  (Cavallini et al., 1988; Eberhard et al., 1993; Iben et al., 2002). This may force us to question whether the role of XPC, XPA, XPG, and XPF/ERCC1 would be first transcriptional and then upon genotoxic attack could also be required for elimination of DNA lesions. Further studies have to be engaged to dissect the specific function of the previously known “repair factors” in the various steps of the transcriptional process.

(E and F) MRC5 cells were treated or not by DRB (100  $\mu$ M), as indicated (E and F), during 6 hr and UV irradiated with 70 J/m<sup>2</sup> through a 6 mm pore filter and fixed 30 min later. Immunofluorescent labeling was performed using rabbit polyclonal anti-XPC, anti-XPB, anti-XPA, and mouse monoclonal anti-cyclobutane pyrimidine dimers (CPD) antibodies. Nuclei were counterstained with DAPI and slides were merged.



**Figure 4. Sequential Recruitment of NER Factors on *RARβ2* Promoter Necessary for Histones PTMs and Active DNA Demethylation**

(A–Y) Relative *RARβ2* mRNA expression monitored by qPCR from stable SiControl, SiXPC, SiXPA, and SiERCC1 HeLa cell lines treated with t-RA (1  $\mu$ M). ChIP monitoring the t-RA-dependent occupancy of RAR, RXR, VDR (B–E), RNA pol II, XPB, cdk7 (F–I), XPC, CSB, XPA (J–M), XPG, XPF, Gadd45 $\alpha$  (N–Q) and di-/trimethylated histones H3K9, H3K4, acetylated H3K9/14, and DDB1 (R–U) on *RARβ2* promoter from stable SiControl, SiXPC, SiXPA, and SiERCC1 HeLa chromatin extracts as indicated. Schematic representation of the *RARβ2* promoter with the indicated CpG islands, restriction enzymes, and primers used to evaluate the methylation status (lower panel). Methylation of CpG islands of *RARβ2* promoter was determined by measuring the ratio of PCR products obtained after digestion by BamHI/Hpa II, BamHI/*DpnI*, or just BamHI of genomic DNA of t-RA treated stable SiControl (V), SiXPC (W), SiXPA (X), and SiERCC1 (Y) HeLa cells.



## EXPERIMENTAL PROCEDURES

### Cell Culture and Reagents

HeLa, HeLa SiXPG (gift from R. Tanaka), XPC HeLa Silencix (SiXPC), XPA HeLa Silencix (SiXPA), ERCC1 HeLa Silencix (SiERCC1), Control HeLa Silencix (SiControl), XP12RO (XPA/R207st) (Satokata et al., 1992), GM14867 (XP-C/R579st) (Chavanne et al., 2000), XPCS1RO (XP-G/frameshift 925) (Hamel et al., 1996), CS1AN (CSB/Q336st) (Troelstra et al., 1992) fibroblasts derived from XP and CS patients and the corresponding rescued cells were cultured in the appropriate medium.

Cells, treated with 1 or 10  $\mu$ M all-trans retinoic acid (Biomol), were incubated with red phenol-free medium containing 10% charcoal treated Fetal Calf Serum (FCS) and 40  $\mu$ g/ml gentamycin during 12 hr prior to the treatment.

HeLa cells were rinsed with PBS, UV-irradiated (20 J/m<sup>2</sup>) and treated with t-RA (1  $\mu$ M). HeLa cells were transfected by pSG5-PPAR $\alpha$  with JetPei (PolyPlus) and treated 24 hr later with 1  $\mu$ M of WY-14643 ligand (Calbiochem) in a red phenol-free medium containing 10% delipidated FCS, 40  $\mu$ g/ml gentamycin. To inhibit transcription, 5,6-Dichloro-1-beta-D-Ribofuranosylbenzimidazole (DRB) (100  $\mu$ M) or  $\alpha$ -amanitin (10  $\mu$ g/ $\mu$ l) (Calbiochem) were added to media during 6 hr in addition to t-RA treatment.

### Antibodies

Monoclonal antibodies against the TFIIH subunits XPB (1B3), XPD (2F6), p62 (3C9), p44 (1H5), cdk7 (2F8), cyclin H (2D4), RNA polymerase II (7C2), TBP (3G3), TFIIIF (2A3), RXR (3A2), RAR (9A6), XPG (1B5), CSB (1A11/3H8), XPA (1E11), RPA (1E9), MED1 (1A10), TIF2 (1D12), and methylated H3K4 (2A12) acetylated H3K9 (IGBMC). Polyclonal antibodies against TFIIH subunit cdk7 (C-19), XPB (S-19), PPAR $\alpha$  (H-98), VDR (H-81), MED6 (C-16), Gadd45 $\alpha$  (C-20), and DDB1 (V-17), XPF (H-300), XPC (D-18), MSH2 (N-20), Ogg1 (N-20) were from Santa-Cruz Biotechnology. Antibodies against XPF (H5), monoubiquitinated H2B (uH2B), di-methyl H3K9, acetylated H3K9/14 and acetylated H4K16 were obtained from ThermoScientific, Medimabs, Cell Signaling Technology and Epigentek, respectively.

### Reverse Transcription and Quantitative PCR

Total RNA was isolated using a GenElute Mammalian Total RNA Miniprep kit (Sigma) and reverse transcribed with SuperScript II reverse transcriptase (Invitrogen). The quantitative PCR was done using the QuantiTect SYBR Green PCR kit (QIAGEN) and the Lightcycler 480 (Roche). The primer sequences for *RAR* $\beta$ 2, *GADD45* $\alpha$ , *HMGCS2*, and Glyceraldehyde 3-phosphate dehydrogenase (*GAPDH*) genes used in real-time PCR are available upon request. *RAR* $\beta$ 2, *GADD45* $\alpha$ , and *HMGCS2* mRNA levels represent the ratio between values obtained from treated and untreated cells normalized against the housekeeping *GAPDH* mRNA.

### Chromatin Immunoprecipitation (ChIP)

Cells were crosslinked at room temperature for 10 min with 1% formaldehyde. Chromatin was prepared (Drane et al., 2004) and sonicated on ice 20 min using a Bioruptor (Diagenode, Liège, Belgium). Samples were immunoprecipitated (IP) with antibodies at 4°C overnight and protein G-Sepharose beads (Upstate, Bellerica, MA) were added, incubated 4 hr at 4°C and sequentially washed. For ChIP/ReChIP experiments, after the first immunoprecipitation and washes, protein-DNA complexes were eluted with a 10 mM DTT solution and diluted before addition of antibodies and protein G-sepharose beads for the second immunoprecipitation. The complexes were eluted and the cross-linking was heat reversed. DNA fragments were purified using QIAquick PCR purification kit (QIAGEN) and analyzed by real-time PCR using sets of primers, available upon request, targeting different regions of *RAR* $\beta$ 2, *HMGCS2*, and *GADD45* $\alpha$  genes.

### ChIP/Western Blot on UV-Irradiated Cells

MRC5 treated with or without DRB (100  $\mu$ M) during 6 hr were rinsed with PBS, UV irradiated (20 J/m<sup>2</sup>). Cells were crosslinked at room temperature for 30 min with 1% formaldehyde at indicated times post-UV irradiation and chromatin was prepared (Fousteri et al., 2006). ChIP/Western blot, using XPB antibodies, was performed as previously described (Coin et al., 2008). Briefly, the chromatin suspension was sonicated in buffer S (10 mM Tris-HCl [pH 8.0],

140 mM NaCl, 1% Triton X-100, 0.1% Na-deoxycholate, 0.1% SDS). Samples were spun down (13,000 rpm, 15 min). In each assay, 600  $\mu$ g of protein from crosslinked chromatin were immunoprecipitated with 1  $\mu$ g of antibody in buffer S, 4 hr at 4°C. The immunocomplexes were collected by adsorption to protein G-Sepharose beads overnight at 4°C. The beads were next washed and resuspended in 1 X Laemmli SDS Buffer. Samples were incubated at 95°C for 90 min for crosslinking reversal prior electrophoresis. After Western blotting following XPB-immunoprecipitation, the intensity of XPB, XPF, p62, RPA, cyclin H, and cdk7 corresponding bands was quantified by densitometry using Genetool (Syngene).

### DNA Methylation Assay and Global Methylation Level Measurement

Genomic DNA was extracted using GenElute Mammalian Genomic DNA Mini-prep Kit (Sigma). Methylation of CpG islands localized in and downstream *RAR* $\beta$ 2 promoter was investigated by digesting 1  $\mu$ g genomic DNA with 5 units of HpaII or BamHI followed or not by a second digestion with 5 units of DpnI (Fermentas). Methylation status was evaluated by measuring the ratio between BamHI/DpnI or BamHI/HpaII and BamHI PCR amplified products using specific primers for the *RAR* $\beta$ 2 promoter (sequences are available upon request). The quantification of PCR products was performed using Genetool.

### Local UV Irradiation, Fluorescence, and Confocal Microscopy

The cells were rinsed with PBS and were covered with an isopore polycarbonate filter with pores of 6 mm diameter (Millipore, Bedford, MA). Cells were then exposed to UV irradiation with a Philips TUV lamp (predominantly 254 nm) at a dose of 70 J/m<sup>2</sup>. Subsequently, the filter was removed, the medium was added back to the cells, and cells were returned to culture conditions for 30 min.

Fibroblasts were fixed in 3% paraformaldehyde for 10 min at room temperature and permeabilized with PBS/0.5% Triton for 5 min. After washing with PBS-Tween (0.05%), the slides were incubated for 1 hr with the indicated antibodies. After extensive washing with PBS-Tween, they were incubated for 1 hr with Cy3-conjugated goat anti-rabbit IgG (Jackson Laboratory) or with anti-mouse Alexa 488 IgG (Jackson Laboratories) diluted 1:400 in PBS-Tween (0.5%). The slides were counterstained for DNA with DAPI prepared in Vectashield mounting medium (Vector lab). All images were collected using a Leica Confocal TCS 4D microscope equipped with both UV laser and an Argon/Krypton laser and standard filters to allow collection of the data at 488 and 568 nm. The software TCSTK was used for three-color reconstructions, and figures were generated using the ImageJ software.

## SUPPLEMENTAL INFORMATION

Supplemental Information includes seven figures and Supplemental Experimental Procedures and can be found with this article online at doi:10.1016/j.molcel.2010.03.004.

## ACKNOWLEDGMENTS

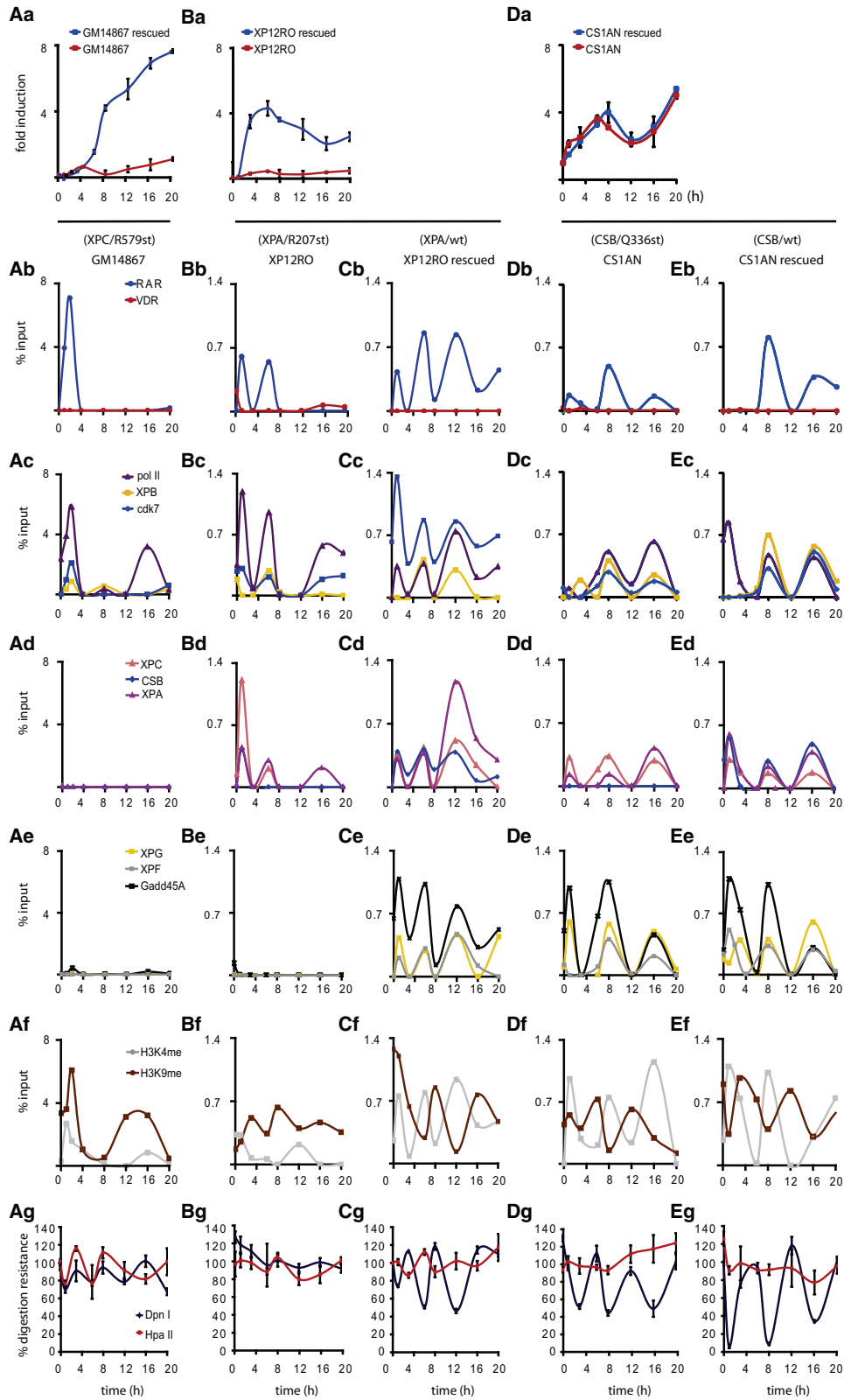
We thank F. Coin and E. Compe for fruitful discussions and critical reading of the manuscript, C. Braun for her technical expertise. We thank M. Oulad for providing antibodies and the IGBMC cell culture facilities. This study has been supported by l'Agence Nationale de la Recherche (N° ANR-05-PCOD-032-03; N° ANR-06-BLAN-0141-01; N° ANR-05-MIIM-033-03, ANR-08-MIEN-02203), l'Association de la Recherche contre le Cancer (ARC n° 3153), and an ERC Advanced grant (to J.M.E.). N.L.M. is recipient from the ARC and a young investigator contract from the Institut National de la Santé et de la Recherche Médicale; D.M.F. was supported by the Fundação para a Ciência e a Tecnologia (F.C.T.) of Portugal PhD Fellowship; R.V.C. was supported by the Chateaubriand and the ARC Postdoctoral Fellowships; Il was supported by the Centre Européen de Recherche en Biologie et Médecine fellowship.

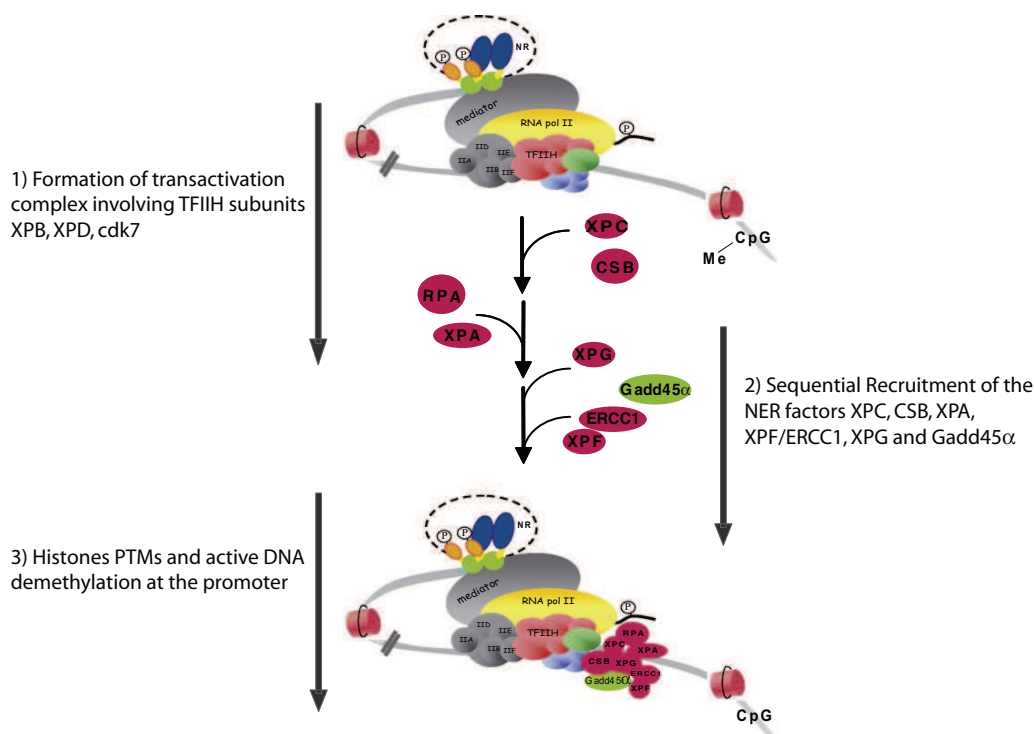
Received: July 24, 2009

Revised: December 1, 2009

Accepted: March 2, 2010

Published: April 8, 2010





**Figure 6. Sequential Recruitment of the NER Factors at the Promoter of Activated Genes Facilitates Chromatin Remodeling**

Upon the t-RA ligand induction, the transactivation complex is formed once RAR/RXR has targeted its responsive element; coactivators and mediator are assembled at RAR $\beta$ 2 promoter together with the transcriptional machinery. XPC is next recruited and allows the sequential arrival of the others NER factors in the following order CSB, XPA/RPA, XPG/XPF/Gadd45 $\alpha$ . This association of NER factors and Gadd45 $\alpha$  with transcription machinery leads to a cascade of histones PTMs. Concomitantly, an active demethylation of 5' CpG islands of RAR $\beta$ 2 gene occurs. Altogether, these remodeling chromatin events are crucial for accurate RNA synthesis. We can mention that XPG has, at least, a double action of stabilization of TFIID and chromatin remodeling in transcription process.

## REFERENCES

- Adimoolam, S., and Ford, J.M. (2002). p53 and DNA damage-inducible expression of the xeroderma pigmentosum group C gene. *Proc. Natl. Acad. Sci. USA* 99, 12985–12990.
- Appanah, R., Dickerson, D.R., Goyal, P., Groudine, M., and Lorincz, M.C. (2007). An unmethylated 3' promoter-proximal region is required for efficient transcription initiation. *PLoS Genet.* 3, e27.
- Barreto, G., Schafer, A., Marhold, J., Stach, D., Swaminathan, S.K., Handa, V., Doderlein, G., Maltry, N., Wu, W., Lyko, F., and Niehrs, C. (2007). Gadd45 $\alpha$  promotes epigenetic gene activation by repair-mediated DNA demethylation. *Nature* 445, 671–675.
- Brivanlou, A.H., and Darnell, J.E., Jr. (2002). Signal transduction and the control of gene expression. *Science* 295, 813–818.
- Cavallini, B., Huet, J., Plassat, J.L., Sentenac, A., Egly, J.M., and Chambon, P. (1988). A yeast activity can substitute for the HeLa cell TATA box factor. *Nature* 334, 77–80.
- Cedar, H., and Bergman, Y. (2009). Linking DNA methylation and histone modification: patterns and paradigms. *Nat. Rev. Genet.* 10, 295–304.
- Chavanne, F., Broughton, B.C., Pietra, D., Nardo, T., Browitt, A., Lehmann, A.R., and Stefanini, M. (2000). Mutations in the XPC gene in families with xeroderma pigmentosum and consequences at the cell, protein, and transcript levels. *Cancer Res.* 60, 1974–1982.
- Coin, F., Bergmann, E., Tremeau-Bravard, A., and Egly, J.M. (1999). Mutations in XPB and XPD helicases found in xeroderma pigmentosum patients impair the transcription function of TFIID. *EMBO J.* 18, 1357–1366.
- Coin, F., Oksenysh, V., Mocquet, V., Groh, S., Blattner, C., and Egly, J.M. (2008). Nucleotide excision repair driven by the dissociation of CAK from TFIID. *Mol. Cell* 31, 9–20.

**Figure 5. RAR $\beta$ 2 Induction, NER Factors Recruitment, Histones PTMs, and DNA Demethylation on Active Promoter in Fibroblasts Derived from XP and CS Patients**

(A–E) Relative RAR $\beta$ 2 mRNA expression monitored by qPCR from GM14867 (XPC/R579st) (Aa), XP12RO (XPA/R207st) (Ba), and CS1AN (CSB/Q336st) (Da) fibroblasts treated with t-RA (10  $\mu$ M) and compared to the corresponding rescued fibroblasts as indicated. Error bars represent the standard deviation of three independent experiments. ChIP monitoring the t-RA-dependent occupancy of RAR, VDR (Ab, Bb, Cb, Db, and Eb), RNA pol II, XPB, cdk7 (Ac, Bc, Cc, Dc, and Ec), XPC, CSB, XPA (Ad, Bd, Cd, Dd, and Ed), XPG, XPF, Gadd45 $\alpha$  (Ae, Be, Ce, De, and Ee) and di-/trimethylated histones H3K9, H3K4 (Af, Bf, Cf, Df, and Ef) on RAR $\beta$ 2 promoter from GM14867, XP12RO, XP12RO rescued, CS1AN and CS1AN rescued chromatin extracts as indicated. Methylation of CpG islands of RAR $\beta$ 2 promoter determined by the ratio of PCR products obtained after digestion by HpaII, BamHI/ *DpnI* or just BamHI of genomic DNA from t-RA treated, GM14867 (Ag), XP12RO (Bg), XP12RO rescued (Cg) CS1AN (Dg), and CS1AN rescued (Eg) fibroblasts. Error bars represent the standard deviation of three independent experiments.



- Compe, E., Drane, P., Laurent, C., Diderich, K., Braun, C., Hoeijmakers, J.H., and Egly, J.M. (2005). Dysregulation of the Peroxisome Proliferator-Activated Receptor Target Genes by XPD Mutations. *Mol. Cell Biol.* 25, 6065–6076.
- Compe, E., Malerba, M., Soler, L., Marescaux, J., Borrelli, E., and Egly, J.M. (2007). Neurological defects in trichothiodystrophy reveal a coactivator function of TFIIH. *Nat. Neurosci.* 10, 1414–1422.
- Constantinou, A., Gunz, D., Evans, E., Lalle, P., Bates, P.A., Wood, R.D., and Clarkson, S.G. (1999). Conserved residues of human XPG protein important for nuclease activity and function in nucleotide excision repair. *J. Biol. Chem.* 274, 5637–5648.
- Drane, P., Compe, E., Catez, P., Chymkowitz, P., and Egly, J.M. (2004). Selective regulation of vitamin D receptor-responsive genes by TFIIH. *Mol. Cell* 16, 187–197.
- Eberhard, D., Tora, L., Egly, J.M., and Grummt, I. (1993). A TBP-containing multiprotein complex (TIF-IB) mediates transcription specificity of murine RNA polymerase I. *Nucleic Acids Res.* 21, 4180–4186.
- Fitch, M.E., Cross, I.V., and Ford, J.M. (2003). p53 responsive nucleotide excision repair gene products p48 and XPC, but not p53, localize to sites of UV-irradiation-induced DNA damage, in vivo. *Carcinogenesis* 24, 843–850.
- Fousteri, M., and Mullenders, L.H. (2008). Transcription-coupled nucleotide excision repair in mammalian cells: molecular mechanisms and biological effects. *Cell Res.* 18, 73–84.
- Fousteri, M., Vermeulen, W., van Zeeland, A.A., and Mullenders, L.H. (2006). Cockayne syndrome A and B proteins differentially regulate recruitment of chromatin remodeling and repair factors to stalled RNA polymerase II in vivo. *Mol. Cell* 23, 471–482.
- Friedberg, E.C., Walker, G.C., and Siede, W. (1995). DNA repair and mutagenesis (Washington, DC: ASM Press).
- Gaillard, P.H., and Wood, R.D. (2001). Activity of individual ERCC1 and XPF subunits in DNA nucleotide excision repair. *Nucleic Acids Res.* 29, 872–879.
- Gramantieri, L., Chieco, P., Giovannini, C., Lacchini, M., Trere, D., Grazi, G.L., Venturi, A., and Bolondi, L. (2005). GADD45- $\alpha$  expression in cirrhosis and hepatocellular carcinoma: relationship with DNA repair and proliferation. *Hum. Pathol.* 36, 1154–1162.
- Hamel, B.C., Raams, A., Schuitema-Dijkstra, A.R., Simons, P., van der Burgt, I., Jaspers, N.G., and Kleijer, W.J. (1996). Xeroderma pigmentosum–Cockayne syndrome complex: a further case. *J. Med. Genet.* 33, 607–610.
- Hanawalt, P.C., and Spivak, G. (2008). Transcription-coupled DNA repair: two decades of progress and surprises. *Nat. Rev. Mol. Cell Biol.* 9, 958–970.
- Hashimshony, T., Zhang, J., Keshet, I., Bustin, M., and Cedar, H. (2003). The role of DNA methylation in setting up chromatin structure during development. *Nat. Genet.* 34, 187–192.
- Iben, S., Tschochner, H., Bier, M., Hoogstraten, D., Hozak, P., Egly, J.M., and Grummt, I. (2002). TFIIH plays an essential role in RNA polymerase I transcription. *Cell* 109, 297–306.
- Ito, S., Kuraoka, I., Chymkowitz, P., Compe, E., Takedachi, A., Ishigami, C., Coin, F., Egly, J.M., and Tanaka, K. (2007). XPG Stabilizes TFIIH (Allowing Transactivation of Nuclear Receptors: Implications for Cockayne Syndrome in XP-G/CS Patients. *Mol. Cell* 26, 231–243.
- Jaspers, N.G., Raams, A., Silengo, M.C., Wijgers, N., Niedernhofer, L.J., Robinson, A.R., Giglia-Mari, G., Hoogstraten, D., Kleijer, W.J., Hoeijmakers, J.H., and Vermeulen, W. (2007). First reported patient with human ERCC1 deficiency has cerebello-oculo-facio-skeletal syndrome with a mild defect in nucleotide excision repair and severe developmental failure. *Am. J. Hum. Genet.* 80, 457–466.
- Keriel, A., Stary, A., Sarasin, A., Rochette-Egly, C., and Egly, J.M. (2002). XPD mutations prevent TFIIH-dependent transactivation by nuclear receptors and phosphorylation of RAR $\alpha$ . *Cell* 109, 125–135.
- Kornberg, R.D. (2007). The molecular basis of eukaryotic transcription. *Proc. Natl. Acad. Sci. USA* 104, 12955–12961.
- Laine, J.P., and Egly, J.M. (2006). Initiation of DNA repair mediated by a stalled RNA polymerase II. *EMBO J.* 25, 387–397.
- Lehmann, A.R. (2003). DNA repair-deficient diseases, xeroderma pigmentosum, Cockayne syndrome and trichothiodystrophy. *Biochimie* 85, 1101–1111.
- Li, B., Carey, M., and Workman, J.L. (2007). The role of chromatin during transcription. *Cell* 128, 707–719.
- Lindahl, T., and Wood, R.D. (1999). Quality control by DNA repair. *Science* 286, 1897–1905.
- Lu, H., Zavel, L., Fisher, L., Egly, J.M., and Reinberg, D. (1992). Human general transcription factor IIH phosphorylates the C-terminal domain of RNA polymerase II. *Nature* 358, 641–645.
- Riedl, T., Hanaoka, F., and Egly, J.M. (2003). The comings and goings of nucleotide excision repair factors on damaged DNA. *EMBO J.* 22, 5293–5303.
- Santoro, R., and Grummt, I. (2001). Molecular mechanisms mediating methylation-dependent silencing of ribosomal gene transcription. *Mol. Cell* 8, 719–725.
- Satokata, I., Tanaka, K., and Okada, Y. (1992). Molecular basis of group A xeroderma pigmentosum: a missense mutation and two deletions located in a zinc finger consensus sequence of the XPAC gene. *Hum. Genet.* 88, 603–607.
- Schaeffer, L., Roy, R., Humbert, S., Moncollin, V., Vermeulen, W., Hoeijmakers, J.H., Chambon, P., and Egly, J.M. (1993). DNA repair helicase: a component of BTF2 (TFIIH) basic transcription factor. *Science* 260, 58–63.
- Smith, C.A., and Hanawalt, P.C. (1978). Phage T4 endonuclease V stimulates DNA repair replication in isolated nuclei from ultraviolet-irradiated human cells, including xeroderma pigmentosum fibroblasts. *Proc. Natl. Acad. Sci. USA* 75, 2598–2602.
- Sugasawa, K., Okamoto, T., Shimizu, Y., Masutani, C., Iwai, S., and Hanaoka, F. (2001). A multistep damage recognition mechanism for global genomic nucleotide excision repair. *Genes Dev.* 15, 507–521.
- Troelstra, C., van Gool, A., de Wit, J., Vermeulen, W., Bootsma, D., and Hoeijmakers, J.H. (1992). ERCC6, a member of a subfamily of putative helicases, is involved in Cockayne's syndrome and preferential repair of active genes. *Cell* 71, 939–953.
- Ueda, T., Compe, E., Catez, P., Kraemer, K.H., and Egly, J.M. (2009). Both XPD alleles contribute to the phenotype of compound heterozygote xeroderma pigmentosum patients. *J. Exp. Med.* 206, 3031–3046.
- van Hoffen, A., Venema, J., Meschini, R., van Zeeland, A.A., and Mullenders, L.H. (1995). Transcription-coupled repair removes both cyclobutane pyrimidine dimers and 6-4 photoproducts with equal efficiency and in a sequential way from transcribed DNA in xeroderma pigmentosum group C fibroblasts. *EMBO J.* 14, 360–367.
- Volker, M., Mone, M.J., Karmakar, P., van Hoffen, A., Schul, W., Vermeulen, W., Hoeijmakers, J.H., van Driel, R., van Zeeland, A.A., and Mullenders, L.H. (2001). Sequential assembly of the nucleotide excision repair factors in vivo. *Mol. Cell* 8, 213–224.
- Wakasugi, M., Reardon, J.T., and Sancar, A. (1997). The non-catalytic function of XPG protein during dual incision in human nucleotide excision repair. *J. Biol. Chem.* 272, 16030–16034.

## Publication 2

### **XPG and XPF endonucleases trigger chromatin looping and DNA demethylation for accurate expression of activated genes.**

Le May N, Fradin D, Iltis I, Bougnères P, Egly JM.

**Molecular Cell** (Aout 2012)

#### **Résumé**

Les facteurs NER, initialement décrits dans les processus de réparation de l'ADN, ont été caractérisés au niveau du promoteur de gènes activés lors de la transcription. Le rôle de ces facteurs, recrutés avec la machinerie basale de transcription, est encore méconnu. Dans cette publication, nous avons montré par CHIP que les facteurs NER associés à la machinerie de transcription ainsi que le facteur CTCF (organisateur de la chromatine), étaient recrutés aussi bien au niveau du promoteur qu'au niveau du terminateur du gène RAR $\beta$ 2.

Par la technique de 3C (chromosome conformation capture), nous avons démontré que le recrutement des facteurs XPG, XPF et/ou CTCF était nécessaire à la formation d'une boucle de chromatine entre le promoteur et le terminateur des gènes activés. Les facteurs XPG et XPF sont des enzymes présentant une activité endonucléase lors de la NER. En effet, le silencing des endonucléases XPG et/ou XPF ou la mutation de leurs sites catalytiques, empêchent la formation de la boucle de chromatine et l'expression optimale du gène RAR $\beta$ 2. Nous avons également mis en évidence que l'endonucléase XPG générait un coupure simple brin et une déméthylation de l'ADN au promoteur, permettant ainsi la mise en place de la boucle d'ADN. Cette boucle de chromatine est dépendante de CTCF et est stabilisée par XPF, qui est capable de générer une coupure et une déméthylation du terminateur.

Nos résultats ont mis en évidence le rôle des facteurs NER XPG et XPF dans la formation de boucle de chromatine contrôlant l'expression des gènes activés.

# XPG and XPF Endonucleases Trigger Chromatin Looping and DNA Demethylation for Accurate Expression of Activated Genes

Nicolas Le May,<sup>1</sup> Delphine Fradin,<sup>2</sup> Izarn Ittis,<sup>1</sup> Pierre Bougnères,<sup>2</sup> and Jean-Marc Egly<sup>1,\*</sup>

<sup>1</sup>Institut de Génétique et de Biologie Moléculaire et Cellulaire, CNRS/Inserm/ULP, BP163, 67404 Illkirch Cedex, C.U. Strasbourg, France  
<sup>2</sup>Department of Paediatric Endocrinology and U986 Inserm, Hôpital Bicêtre, Paris 11 University, 94000 Kremlin-Bicêtre, France

\*Correspondence: [egly@igbmc.fr](mailto:egly@igbmc.fr)

<http://dx.doi.org/10.1016/j.molcel.2012.05.050>

## SUMMARY

Nucleotide excision repair factors, initially characterized as part of DNA repair, have been shown to participate in the transcriptional process in the absence of genotoxic attack. However, their molecular function when recruited at the promoters of activated genes together with the transcription machinery remained obscure. Here we show that the NER factors XPG and XPF are essential for establishing CTCF-dependent chromatin looping between the promoter and terminator of the activated *RARβ2* gene. Silencing XPG and/or XPF endonucleases, or mutations in their catalytic sites, prevents CTCF recruitment, chromatin loop formation, and optimal transcription of *RARβ2*. We demonstrated that XPG endonuclease promotes DNA breaks and DNA demethylation at promoters allowing the recruitment of CTCF and gene looping, which is further stabilized by XPF. Our results highlight a timely orchestrated activity of the NER factors XPG and XPF in the formation of the active chromatin hub that controls gene expression.

## INTRODUCTION

RNA synthesis is the result of a cascade of chronologically orchestrated events that requires several hundreds of proteins. Upon gene activation, a host of proteins including RNA polymerase II (pol II), the general transcription factors (GTFs), coactivators, corepressors, and chromatin remodelers are assembled around the promoter, and their enzymatic activities contribute to protein coding gene expression (Brivanlou and Darnell, 2002; Kornberg, 2007). Among these proteins necessary to initiate RNA synthesis are the nucleotide excision repair (NER) factors. These factors are sequentially recruited and required for optimal chromatin remodeling including histone posttranslational modifications (PTMs) as well as DNA demethylation at the activated genes (Le May et al., 2010; Schmitz et al., 2009; Barreto et al., 2007). These NER factors (XPC, CSB, TFIIH, XPA, XPG, XPF-ERCC1) were first characterized as part of the DNA repair machinery to eliminate lesions originated by exogenous or endogenous genotoxic attacks (for review, see Nouspi-

kel, 2009). Mutations in the genes coding for NER factors have been associated with the human genetic disorders xeroderma pigmentosum (XP), trichothiodystrophy (TTD), Cockayne syndrome (CS), XFE progeroid syndrome, and cerebro-oculo-facio-skeletal syndrome (COFS) characterized by a combined dysregulation of DNA repair and transcription processes (Jaspers et al., 2007; Kraemer et al., 2007; Niedernhofer et al., 2006).

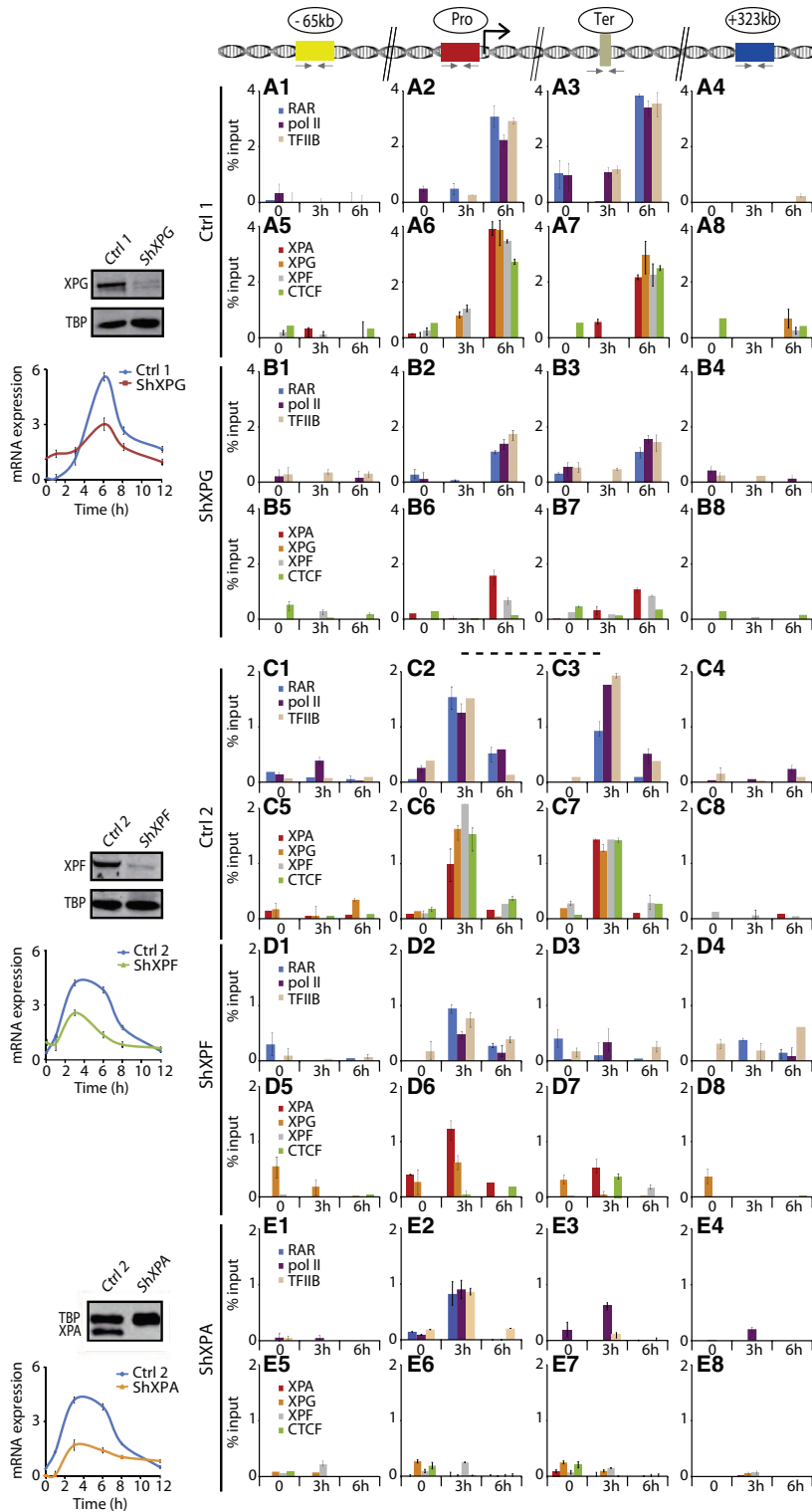
In the present study, we investigated the implication of XPG and XPF endonucleases in the transactivation of nuclear receptor (NR) target genes. We found that these two factors, detected at the promoter and terminator of the activated *RARβ2* gene, were required for DNA breaks and DNA demethylation. These two steps are crucial for the recruitment of the CCCTC-binding factor (CTCF) chromatin organizer and consequently the formation of gene looping between promoter and terminator. Altogether, these events optimize RNA synthesis.

## RESULTS

### XPG and XPF Are Recruited at the Promoter and Terminator of *RARβ2* upon Activation

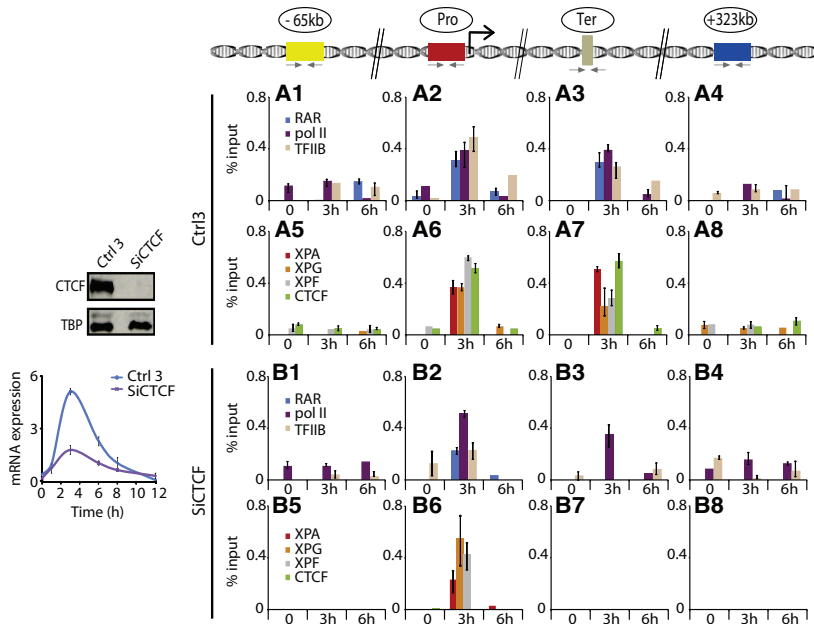
We first attempted to investigate the presence of the transcription and NER factors along the *RARβ2*-activated gene. Analysis of the promoter region (Pro), including two RAR responsive elements (RAREs) and the TATA box as well as a region defined as the terminator (Ter) including the Poly(A) site, was performed using chromatin immunoprecipitation (ChIP) followed by quantitative PCR at various times as indicated (Figure 1). Cells that either stably express targeted shRNA silencing XPG (ShXPG) and ERCC1 (named ShXPF), XPA (ShXPA), or transiently express siRNA against CTCF (SiCTCF) as well as their respective corresponding controls (named Ctrl1-3, see the Experimental Procedures) were treated with all-trans retinoic acid (t-RA). Silencing ERCC1 also silenced XPF (Figure 1, left panel; see also Gaillard and Wood, 2001).

In each silenced cell line, we observed a significant defect in *RARβ2* mRNA expression compared to the Ctrl1-3 cells (Figure 1 and Figure 2, left panels). We noticed that as a function of the cell line and the set of experiments, *RARβ2* mRNA synthesis peaks either at 3 or at 6 hr. Therefore ChIPs were performed either at 0, at 3, or at 6 hr posttreatment. In Ctrl1 cells, we observed at 6 hr posttreatment the concomitant recruitment of RAR, pol II, and TFIIIB together with the NER factors indicated by the



**Figure 1. XPG and XPF Are Necessary for Concomitant Recruitment of Transcription Machinery at Promoter and Terminator upon *RARβ2* Induction**

Schematic representation of the *RARβ2*; the amplicons at the -65 kb, Pro, Ter, and +323 kb regions are indicated (upper panel). Relative *RARβ2* mRNA expression monitored over time by qPCR from t-RA-treated ShXPG, ShXPF, ShXPA HeLa cells, and the corresponding Ctrl1 and Ctrl2 control cells (left panels). Error bars represent the standard deviation of three independent experiments. Western blotting analyses of XPG, XPF, XPA, and TBP were performed from chromatin extracts of ShXPG, ShXPF, ShXPA HeLa cells and their corresponding controls (left panels). TBP is used as a positive nuclear marker. ChIP monitoring the t-RA-dependent occupancy of RAR, pol II, TFIIIB (A1-4, B1-4, C1-4, D1-4, E1-4), XPA, XPG, XPF, and CTCF (A5-8, B5-8, C5-8, D5-8, E5-8) on *RARβ2* locus from ShXPG (B1-8), ShXPF (D1-8), ShXPA (E1-8), and the corresponding control (Ctrl1, Ctrl2) chromatin extracts at 0, 3, and 6 hr upon t-RA induction as indicated. A1-4, A5-8—E5-E8 are aligned under the corresponding region of the *RARβ2* gene being probed. Each series of ChIP is representative of at least two independent experiments as indicated by standard deviation, and values are expressed as percentage of the input.



**Figure 2. CTCF Is Necessary for Concomitant Recruitment of Transcription Machinery at Promoter and Terminator upon *RARβ2* Induction**

Schematic representation of the *RARβ2*; the amplicons at the -65 kb, Pro, Ter, and +323 kb regions are indicated (upper panel). Relative *RARβ2* mRNA expression monitored over time by qPCR from t-RA-treated SiCTCF HeLa cells, and the corresponding Ctrl3 control cells (left panel). Error bars represent the standard deviation of three independent experiments. Western blotting analyses of CTCF and TBP were performed from chromatin extracts of SiCTCF HeLa cells and the corresponding control (left panel). TBP is used as a positive nuclear marker. The t-RA-dependent occupancy of RAR, pol II, TFIIIB (A1-4, B1-4), XPA, XPG, XPF, and CTCF (A5-8, B5-8) was monitored by ChIP on *RARβ2* locus using SiCTCF (B1-8) and Ctrl3 (A1-8) chromatin extracts at 0, 3, and 6 hrs upon t-RA induction. A1-4, A5-8–B5-8 are aligned under the corresponding region of the *RARβ2* gene being probed. Each series of ChIP is representative of at least two independent experiments as indicated by standard deviation, and values are expressed as percentage of the input.

presence of XPA, XPG, and XPF (Figure 1, panels A2 and A6) at Pro, paralleling *RARβ2* mRNA synthesis. All of these proteins were also observed at Ter (panels A3 and A7). Such a recruitment pattern was not detected at position 65 kb (-65 kb) and 323 kb (+323 kb), upstream and downstream to the transcription start site (TSS), respectively (panels A1, A5 and A4, A8). The recruitment pattern at Pro/Ter is different from the one found at exon 6 that lacks the RAR as well as the TFIIIB transcription initiation factor (see Figure S1 available online). In Ctrl2 and Ctrl3, the transcription and the NER factors were similarly detected at Pro and Ter 3 hr post-t-RA treatment (Figure 1, panels C1–C8, and Figure 2, panels A1–A8). Moreover, we also noticed the absence of these proteins at the -65 kb and +323 kb regions (Figure 1 and Figure 2).

We demonstrated that upon gene activation there is formation of a large transcription complex containing the transcription and NER machineries, using as a target both the promoter and the terminator of the *RARβ2* gene. We next wondered whether the absence of NER factors would prevent complex formation.

In ShXPG cells, XPA together with XPF, and in ShXPF cells, XPA together with XPG were detected at Pro, although to a lower extent compared with their respective Ctrl1-2 controls (Figure 1, compare panels B2 and B6, with panels A2 and A6, and panels D2 and D6 with panels C2 and C6). In ShXPG cells, transcription and NER factors were still recruited at Ter, but to a lower extent in comparison to Ctrl1 (compare panels B3 and B7 with panels A3 and A7). However, in ShXPF cells, the corecruitment of transcription machinery with XPA and XPG, obvious at Pro, was less significant at Ter (compare panels D2, D6 and D3, D7 with panels B2, B6 and B3, B7).

Previous work has described the sequential arrival of the NER factors following the preinitiation complex (PIC) formation, in which XPC, RPA, and XPA precede XPG and XPF at the *RARβ2*

promoter (Le May et al., 2010). As expected, in ShXPA cells, neither XPG nor XPF was detected at Pro and Ter (Figure 1, compare panels E6, E7 and C6, C7). Moreover, as in ShXPG and shXPF cells, the transcriptional machinery was recruited to a lower extent at Ter than at Pro in ShXPA cells (panels E3 and E2).

Our data indicated that the presence of the later-recruited NER factors XPG and XPF was crucial in the formation of a transactivation complex that encompassed both the promoter and the terminator regions of the *RARβ2*-activated gene, suggesting a long-range chromatin rearrangement.

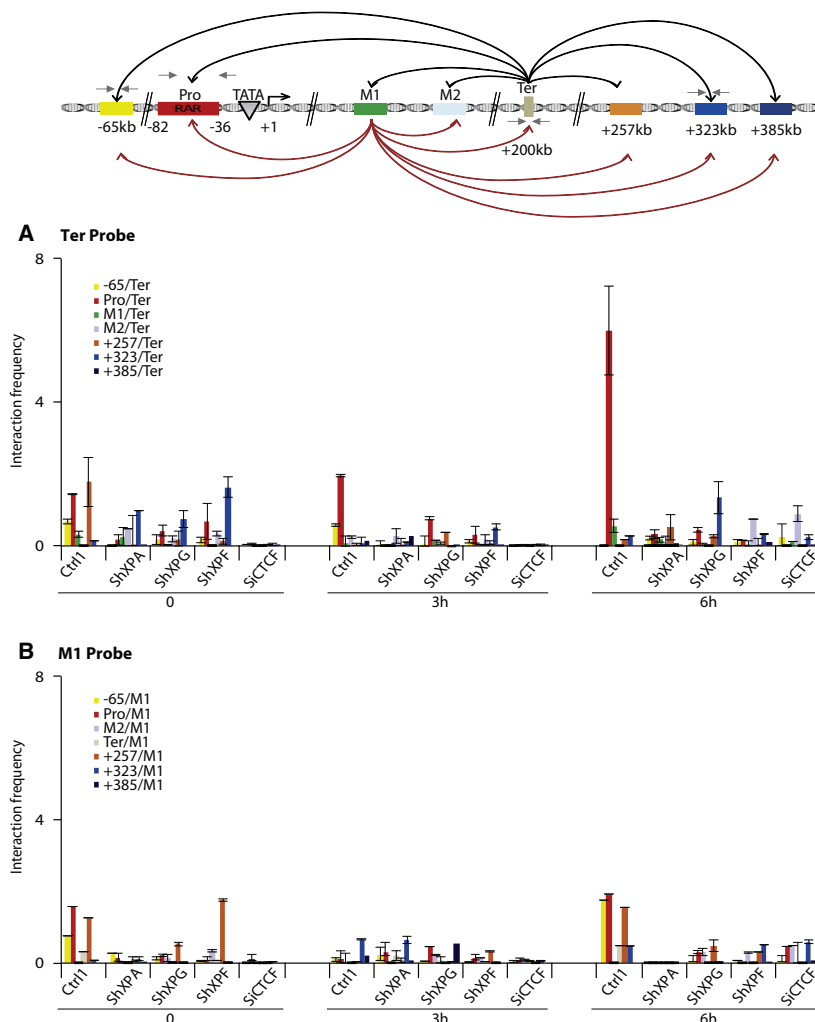
#### XPG, XPF, and CTCF Are Required for Gene Looping

Several studies have underlined the influence of DNA binding proteins in specific loop formation and distinct gene expression as demonstrated for the imprinted *Igf2/H19* and the role of the chromatin organizer CTCF (Murrell et al., 2004). In Ctrl1-3 cells, CTCF was recruited at Pro and Ter upon t-RA induction (Figure 1, panels A6, A7, C6, C7, and Figure 2, panels A6, A7). However in ShXPG, ShXPF, and ShXPA cells in which the concomitant presence of XPG and XPF was abolished, CTCF was not detected at Pro and Ter (Figure 1, panels B6, B7, D6, D7, and E6, E7, respectively) when compared with Ctrl1-2 (panels A6, A7 and C6, C7).

In SiCTCF cells, *RARβ2* expression was significantly reduced compared to Ctrl3 but not totally abolished (Figure 2, left panel). *RARβ2* mRNA synthesis still occurred, as indicated by the presence of pol II at exon 6 and at Ter (Figure S1 and Figure 2, panel B3). The ChIP assay showed a concomitant detection of the NER factors with the transcriptional machinery at Pro, but not at Ter (Figure 2, compare panels B2, B6 with panels A2, A6 and panels B3, B7 with panels A3, A7).

The above data indicate that the presence of both XPG and XPF is required for the recruitment of CTCF. The correlated





recruitment to distant regions along the *RARβ2* locus suggested the formation of gene looping. We thus performed quantitative chromatin conformation capture assays (q3C) and analyzed interactions between several regions along the *RARβ2* gene (Veriminen et al., 2007). Digestion of crosslinked chromatin by *HindIII* resulted in restriction fragments equidistantly separated by 60 kb and containing either -65 kb, Pro, Ter, or +323 kb as well as an intronic region (M1) of the *RARβ2* locus (Figure 3, upper scheme). Ter and M1 of *RARβ2* gene were used as baits. In t-RA-treated Ctrl1 cells, we observed that Pro could specifically and significantly interact with Ter at 6 hr (Figure 3A), paralleling the *RARβ2* mRNA synthesis (Figure 1, left panel) and the recruitment of the transcriptional apparatus (Figure 1, panels A2, A3, A6, A7). Similar observations were made at 3 hr in Ctrl2-3 (data not shown). By contrast, in ShXPG, ShXPF, and ShXPA cells as well as in siCTCF, no spatial proximity between Ter and Pro was revealed. As controls, no specific interactions were observed between the intronic M1 bait and Pro or between all the other analyzed fragments upon t-RA treatment (Figure 3B).

### Figure 3. XPG and XPF Are Involved in CTCF-Induced *RARβ2* Gene Looping

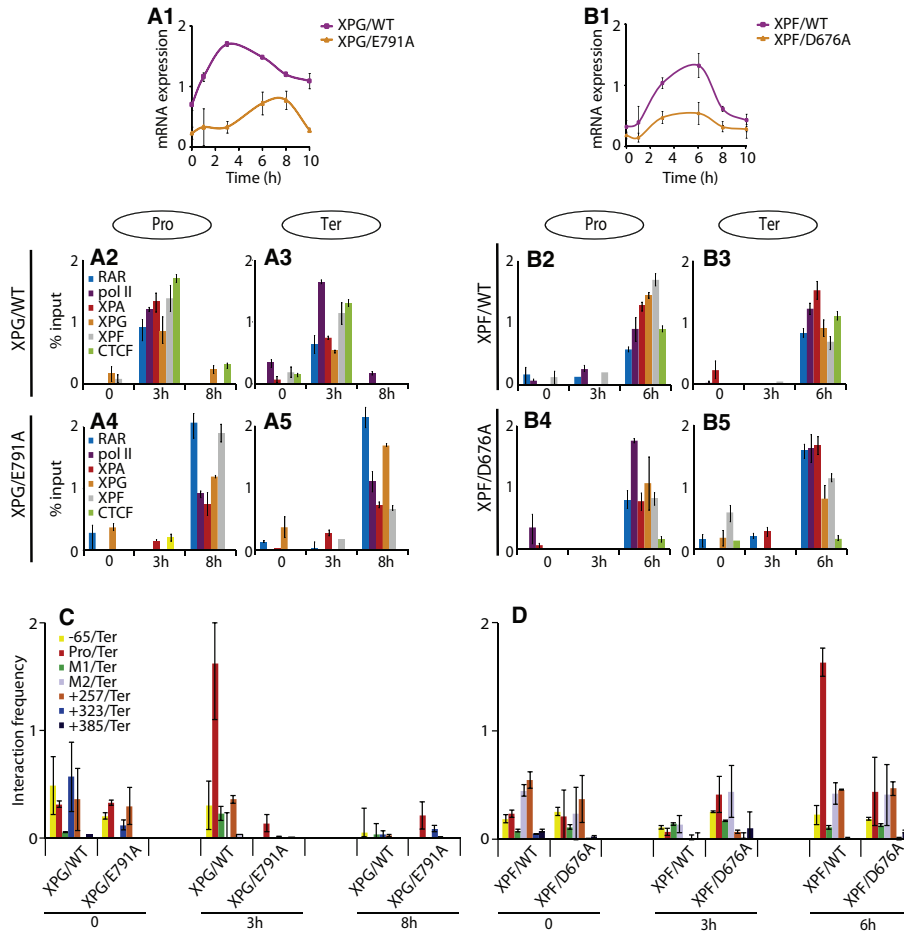
Schematic representation of the quantitative chromatin conformation capture (q3C) (upper panel). Two probes were designed at Ter (gray square) and M1 (green square) of *RARβ2* gene. These probes were used to investigate the associations between the different elements including upstream (-65 kb, Pro), intronic (M1, M2), and downstream (Ter, +257 kb, +323 kb, +385 kb) regions as indicated by the black (Ter probe) and red (M1 probe) arrows. q3C assays were performed using crosslinked and *HindIII*-digested chromatin from Ctrl1, ShXPG, ShXPA, ShXPF, and siCTCF HeLa cells at 0, 3, and 6 hr post-t-RA treatment (10 μM). The bar chart (y axis) shows the enrichment of PCR product (%) normalized to the enrichment within the human *xpb* gene (=100%), as illustrated in Figure S2G. Each PCR was performed at least three times and averaged as indicated by standard deviation. Signals were normalized to the total amount of DNA used, estimated with an amplicon located within a *HindIII* fragment in *RARβ2* gene (see the Experimental Procedures).

Altogether our data showed that the t-RA induction could initiate a CTCF-related long-range interaction between Pro and Ter of *RARβ2* gene that was abolished in the absence of either XPG or XPF, underlining the role of these NER factors in chromatin remodeling.

### The Catalytic Activity of Both XPG and XPF Is Required for Gene Looping

Consistent with the role of XPG and XPF-ERCC1 in NER (Gillet and Scharer, 2006; O'Donovan et al., 1994; Sijbers et al.,

1996), we wondered about the requirement of their endonuclease activity in chromatin looping upon transactivation. The SV40-immortalized XP-G (XPCS1RO) and XP-F (XP2YO) fibroblasts that were derived from a XP/CS and XP patient, respectively (Ellison et al., 1998; Yagi and Takebe, 1983), were stably transfected either with XPG/WT, XPG/E791A or with XPF/WT, XPF/D676A (Figure S2A). XPG/E791A and XPF/D676A mutations were found to abolish the catalytic activity of these two endonucleases and consequently their ability to eliminate DNA damages when added in an in vitro NER assay (Lalle et al., 2002; Staresinic et al., 2009; and Figure S2B). *RARβ2* mRNA synthesis was significantly inhibited in XPG/E791A and XPF/D676A cells as well as in the corresponding parental cells compared to the cells expressing either XPG/WT or XPF/WT, respectively (Figure 4, panels A1, B1, and Figures S2C and S2D). In XPG/WT and XPG/E791A cells, we observed a conserved recruitment of RAR, pol II, XPA, XPG, and XPF both at Pro and Ter at 3 and 8 hr, respectively (Figure 4, compare panels A2, A3 with panels A4, A5). In XPG/E791A cells, the optimal ChIP recruitment



**Figure 4. Endonuclease Activity of XPG and XPF Is Required for Gene Looping**

Relative *RAR* $\beta$  mRNA expression monitored by quantitative PCR from XPG/WT (A1), XPG/E791A (A1), XPF/WT (B1), XPF/D676A (B1) cells at 0, 1, 3, 6, 8, and 10 hr post-t-Ra treatment. Error bars represent the standard deviation of three independent experiments. ChIP monitoring of the t-Ra-dependent occupancy of RAR, pol II, XPA, XPG, XPF, and CTCF at Pro and Ter regions of *RAR* $\beta$  gene from XPG/WT (A2-3), XPG/E791A (A4-5), and XPF/WT (B2-3), XPF/D676A (B4-5) chromatin extracts as indicated. Each series of ChIP is representative of two independent experiments. Intrachromosomal interactions between Ter and upstream/downstream elements of *RAR* $\beta$  gene are as follows: q3C assays were performed using crosslinked, HindIII-digested chromatin from t-Ra-treated XPG/WT or XPG/E791A (C) and XPF/WT or XPF/D676A (D) fibroblasts at 0, 3, and 6 hr. These independent graphs represent a measure of the association between the upstream -65 kb; Pro; the intronic M1 and M2 elements; and the downstream +257 kb, +323 kb, and +385 kb regions to either the terminal region (Ter Probe) or the intronic region (M1 Probe, see Figures S2E and S2F) as a negative control. Each PCR was performed at least three times and averaged as indicated by standard deviation. Signals were normalized to the total amount of DNA used, estimated with an amplicon located within a HindIII fragment in *RAR* $\beta$  gene.

reflected the *RAR* $\beta$  mRNA synthesis peak at 8 hr (panel A1). In XPF/D676A and XPF/WT cells, the recruitment patterns of transcription and NER factors were similar both at Pro and Ter (compare panels B4, B5 with panels B2, B3), contrary to what was observed at Ter in ShXPF cells (Figure 1, panels D3, D7). We also noticed that both XPG/E791A and XPF/D676A proteins were recruited with the other NER factors. However, in spite of the presence of both endonucleases in XPG/E791A and XPF/D676 cells, the q3C assay did not show any stable interaction between Pro and Ter, while in both XPG/WT and XPF/WT cells, such interaction was observed at 3 and 6 hr upon t-Ra induction, respectively (Figures 4C and 4D). In XPG/WT and XPF/WT cells, CTCF was detected at both Pro and Ter concomitantly with

gene looping (Figure 4, panels A2, A3 and B2, B3). In XPG/E791A or XPF/D676 cells, the inhibition of gene looping was coincident with the absence of CTCF (Figure 4, panels A4, A5 and B4, B5).

These data indicated that the endonuclease activities of XPG and XPF were necessary to recruit CTCF and to allow stable chromatin looping between Pro and Ter of the *RAR* $\beta$ -activated gene.

**XPG Endonuclease Induces DNA Breaks and DNA Demethylation at the Promoter Region**

Previous works have revealed the formation of transient DNA breaks upon gene activation (Ju et al., 2006). This raises

questions regarding the interplay between XPG and XPF endonucleases and the transactivation process. Using a Bio-ChIP assay which measures the incorporation of biotinylated dUTP within broken DNA, we observed a concomitant increase of DNA cleavage specifically both at Pro and Ter in all five of the t-RA-treated Ctrl1, Ctrl2, XPG/WT, XPF/WT, and Ctrl3 cells (Figure 5, panels A1, B1, C1, and D1, respectively). In ShXPG and XPG/E791A cells, DNA breaks were hardly detected at Pro (Figure 5, panels A1, C1), whereas in ShXPF and XPF/D676 cells, DNA cleavage could only be observed at Pro (panels B1 and C1). Nevertheless, we noticed that in XPG-deficient cells there was a slight incorporation of bio-UTP at Ter (panel A1).

Several studies have documented a relationship between XPG and DNA demethylation upon transcription (Barreto et al., 2007; Jin et al., 2008). We thus investigated whether or not the endonuclease activity XPG and XPF was linked to DNA demethylation upon *RARβ2* gene activation. Using an unmethylated DNA immunoprecipitation (unMeDIP) approach, we measured the removal of the 5'-methylcytosine (5mC) along the different regions of activated *RARβ2*. In all of the five control cells, Pro and Ter were found unmethylated (Figure 5, panels A2, B2, C2, and D2). In ShXPG and XPG/E791A cells, Pro remained methylated, while a slight DNA demethylation was observed at Ter when compared to their corresponding Ctrl1 and XPG/WT cells (panels A2, C2). In ShXPF and XPF/D676 cells, Pro but not Ter was unmethylated (panels B2 and C2). Using the methylated DNA immunoprecipitation (MeDIP) approach, no significant decrease was observed in DNA methylation at either Pro or Ter in t-RA-treated XPG/E791A and XPF/D676A cells (Figure S3), which was found to be consistent with the UnMEDIP (see above) assays and pyrosequencing data (see below, panel E); in XPG/WT and XPF/WT, we observed a weak but significant decrease of DNA methylation. Interestingly, in ShXPA, neither DNA breaks nor DNA demethylation was detected at Pro and Ter (panels B1 and B2, respectively), likely due to the absence of XPG and XPF (Figure 1, panels E6 and E7).

Having demonstrated the connection between CTCF and the XPG and XPF endonucleases for chromatin looping, we then investigated whether the presence of CTCF was a prerequisite for DNA modifications. In SiCTCF, we repeatedly observed both DNA breaks and DNA demethylation at Pro but not at Ter when compared to Ctrl3 (Figure 5, panels D1 and D2). These data strongly suggest a role for CTCF in gene looping and Ter modification following DNA breaks and DNA demethylation at Pro. However, our results showed that these DNA modifications followed the dynamic recruitment of NER factors and the transcriptional machinery upon transactivation (Figure 5). To further localize these DNA modifications, we thus pyrosequenced the promoter of *RARβ2* from t-RA-treated cells.

Pyrosequencing analysis of Bio-ChIP DNA samples from t-RA-treated XPG/WT and XPF/WT cells indicated that a high frequency of DNA breaks on both strands localized downstream the *RARβ2* TSS (Figure 5E, red arrows); most of them were detected on the transcribed strand (TS). In XPG/E791A cells, the frequency of DNA cleavage strongly decreased and the detected breaks were differently localized (Figure 5E, light blue

arrows). Interestingly, in XPF/D676A, the cleavage pattern observed around the *RARβ2* promoter was similar (Figure 5E, dark blue arrow), demonstrating that XPF endonuclease inhibition did not prevent Bio-UTP incorporation and thus DNA breaks at Pro (see also panels B1 and C1).

Genomic DNA was next bisulfate converted and pyrosequenced to localize the demethylated 5mC at Pro in t-RA-treated cells. Seventeen CpG dinucleotides, localized inside the proximal promoter and downstream of the TSS, were analyzed (Figure 5E). In XPG/WT, XPF/WT, and XPF/D676A cells, CG1, CG3, CG15, and CG16 (closed dark blue circles) were significantly demethylated, whereas only CG1 remained demethylated in XPG/E791A (Figure 5E). The above data underlined the role of XPG and XPF endonucleases in both the formation of DNA breaks and active DNA demethylation at Pro and Ter, respectively. Our data also showed that XPG-dependent DNA breaks and DNA demethylation at Pro preceded the recruitment of CTCF.

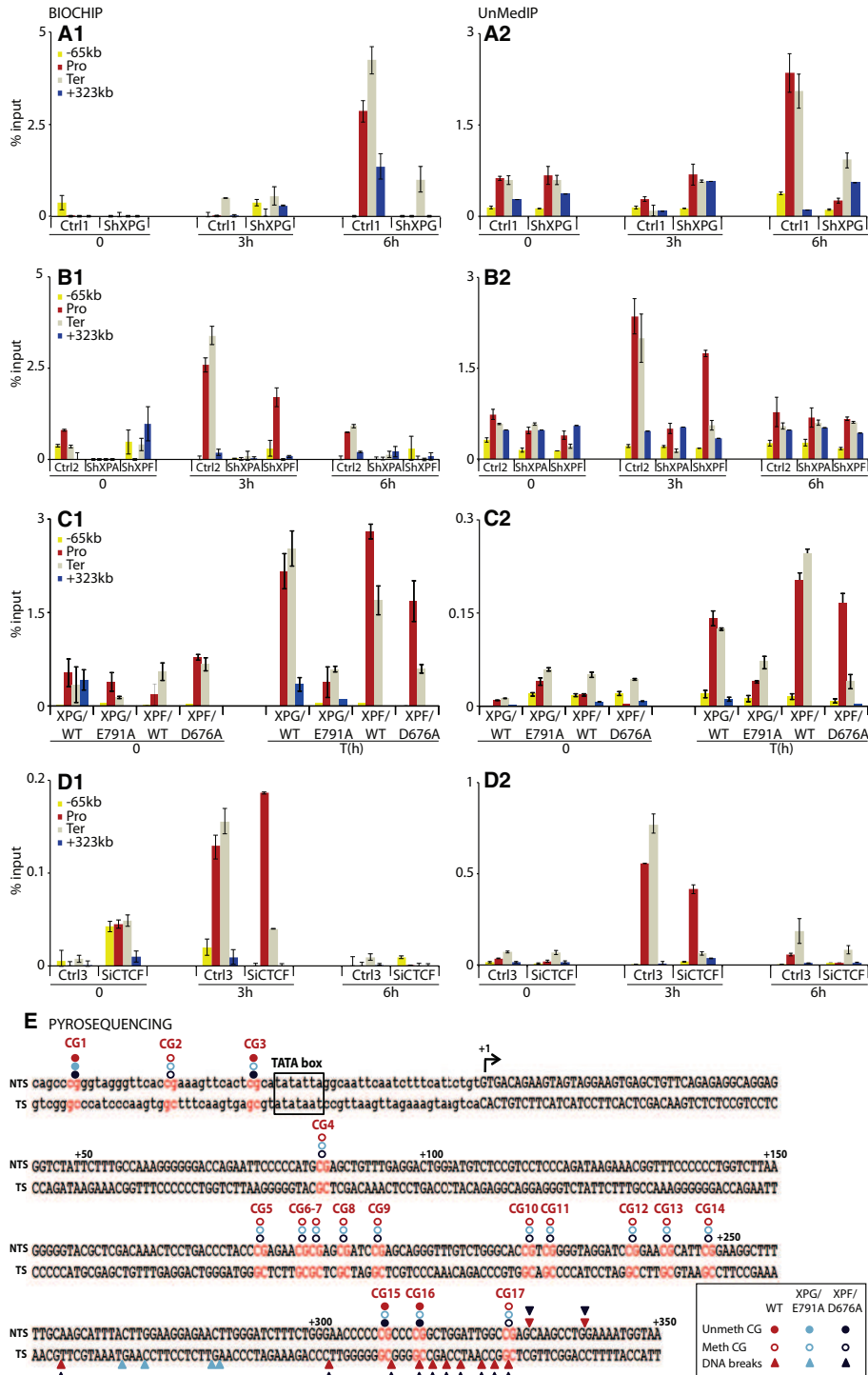
## DISCUSSION

Discovering the components of the eukaryotic transcription apparatus has been a major theme of research over the past decades. In the present study, we describe how XPG and XPF endonucleases, which are sequentially recruited following the arrival of the transcription machinery at the promoter of activated genes, are involved in their optimal expression.

We especially found that upon induction by t-RA, RAR as well as GTFs and pol II were found at both the promoter and the terminator of *RARβ2* gene. NER factors, where XPA recruitment precedes XPG and XPF, are established once RNA pol II and transcription factor are recruited (Figure 1, see also Le May et al., 2010). We next demonstrate that both XPG and XPF endonuclease activity are crucial to recruit the chromatin organizer CTCF and to promote gene looping between the promoter and the terminator of the activated *RARβ2* gene.

The formation of such an “active chromatin hub” resulting from inter- and intrachromosomal interactions between specific subsets of transcription units of activated genes as described for NR-target genes (Hu et al., 2008), *β-Globin* (de Laat and Grosfeld, 2003), *T helper type 2* (Spilianakis and Flavell, 2004), *BRCA1* (Tan-Wong et al., 2008), *GATA-4* loci (Tiwari et al., 2008), and HIV-1 provirus (Perkins et al., 2008) might provide additional control mechanisms for regulated gene expression (Deng and Blobel, 2010). We here show that the presence of XPG and XPF together with CTCF in the long-range interactions can be considered as part of the regulatory biological program of the *RARβ2* gene. Indeed, in ShXPG, ShXPF, and ShXPA cells, in which the concomitant presence of XPG and XPF is abolished, CTCF is absent, and, as well as in SiCTCF cells, there is no gene looping and *RARβ2* RNA synthesis is downregulated (Figure 1, Figure 2, and Figure 3). Moreover, it is interesting to notice the crucial role of XPA in recruiting XPG and XPF, knowing that XPA recruitment occurs before XPG and XPF (Le May et al., 2010). It is not surprising to notice the presence of XPA together with RNA pol II at exon 6 and Ter (although at a very low level) in ShXPG and ShXPF cells (Figure S1 and Figure 1).





**Figure 5. XPG and XPF Are Required for DNA Breaks and DNA Demethylation at Pro and Ter, Respectively**

Detection of DNA break at -65 kb, Pro, Ter, and +323 kb of the *RARβ2* locus from Ctrl1, ShXPG (A1), Ctrl2, ShXPA, ShXPF (B1), XPG/WT, XPG/E791A and XPF/WT, XPF/D676A (C1), Ctrl3, SiCTCF (D1) t-RA-treated cells. DNA breaks are detected first through the incorporation of Biotin-dUTP via a terminal deoxynucleotidyl transferase (TdT) reaction and second via a ChIP approach. DNA fragments containing DNA breaks are immunoselected by Biotin antibodies and analyzed by qPCR. Each series of Bio-ChIP is representative of three independent experiments as indicated by standard deviation, and values are expressed as

In addition, mutations that affect the catalytic activity of XPG and XPF (XPG/E791A and XPF/D676A, respectively) were found to prevent CTCF recruitment. More interestingly, we discover that both gene looping and CTCF recruitment occur in parallel with the formation of DNA breaks (Figure 5). Having demonstrated (1) that XPG/E791A mutation prevented DNA breaks formation at Pro while XPF catalytic mutation does not (Figure 4) and (2) that in SiCTCF there are DNA breaks at Pro but not at Ter around which no NER factors were detected (Figure 2 and Figure 5), it is tempting to propose the following model: upon t-RA transactivation and transcription and NER factors recruitment (Figure 6, step 1), XPG induces some DNA breaks at Pro (step 2) independently of both the status of XPF and the presence of CTCF. Once XPG has fulfilled its role at Pro, CTCF recruitment is initiated with the help of XPF (step 3).

Such a scenario raises questions regarding the interplay between molecular machineries that are involved in the formation of DNA breaks, DNA demethylation, the recruitment of CTCF, and the activation of the gene transcription. Using different approaches (UnMEDIP, MEDIP, and pyrosequencing), we observed a relationship between the DNA demethylation at Pro and Ter and the presence of CTCF necessary for the loop formation of the *RARβ2*-activated gene. In particular, we have shown that demethylation at Pro is not sufficient to attract CTCF when Ter is not demethylated (Figure 5). Indeed, XPG and XPF defects abolish DNA demethylation at Pro and Ter, respectively, and consequently CTCF recruitment (Figure 5). Accordingly, DNA breaks and DNA demethylation at Pro are not sufficient to attract CTCF in the absence of Ter domain modification. We thus can speculate that the DNA breaks induced by XPG and the consequent recruitment of CTCF at demethylated promoter could be responsible for important torsional tensions of DNA that allow the promoter-bound XPF to be brought in proximity to the terminator (Figure 6, step 3).

Mechanisms underlying the removal of methyl groups from genomic DNA are very controversial (Barreto et al., 2007; Jin et al., 2008) and have implicated DNA repair processes either through deamination followed by excision of the 5mC or by direct removal of the methyl moiety from the base (Gehring et al., 2009; Ma et al., 2009). Our results do not allow the identification of the precise causal relationships between DNA demethylation and the effects of XPG and XPF endonucleases. However, our data indicated that the DNA breaks induced by XPG were located near the demethylated CG dinucleotides in the vicinity of *RARβ2* TSS and almost exclusively on the TS (Figure 5E). Whether or not DNA cuts are a prerequisite for DNA demethyla-

tion remains to be shown. However, the pattern of DNA cuts and demethylated 5mC was strongly disturbed when XPG and XPF endonucleases were inactivated. Such localized roles of XPG at Pro and XPF at Ter, as well as their cooperation in promoting loop formation and consequently optimal *RARβ2* gene expression, are reminiscent to the NER process in which XPF and XPG excise at the 3' and 5' of the DNA lesion, respectively, in a cooperative manner (Tapias et al., 2004b; Araujo et al., 2000; Staresincic et al., 2009). In NER, XPG mutations completely abolished the incision, while XPF mutations still allow XPG cuts at the 5' side of the DNA damage (Lalle et al., 2002).

The present study underlines the essential role of XPG, XPF, and CTCF recruitment in the chromatin loop organization required for optimal expression of activated genes. Further genome-wide investigations are needed to question the general role of these endonucleases. Any events as presently documented that distort the interaction network involving both protein factors and the subsets of transcription units upon gene activation might result in inaccurate transactivation complex formation and altered gene expression. It would thus be legitimate to raise the possibility of some other interconnecting links in this cascade of events followed by the action of XPG, XPF, and CTCF. For example, defects in either the basal transcription factor TFIIB (Singh and Hampsey, 2007), the cdk7 kinase of TFIIF that abolishes RNA pol II (O'Sullivan et al., 2004), and/or NR phosphorylation (Compe et al., 2007), as well as mutations in the mediator that weaken its interaction with activator (Hashimoto et al., 2011; Wang et al., 2005), might indirectly disturb the formation of the transactivation complex by preventing, in the end, the recruitment and/or the function of XPG/XPF endonucleases and/or the recruitment of CTCF that is essential in chromatin rearrangement and, further, in optimal gene expression.

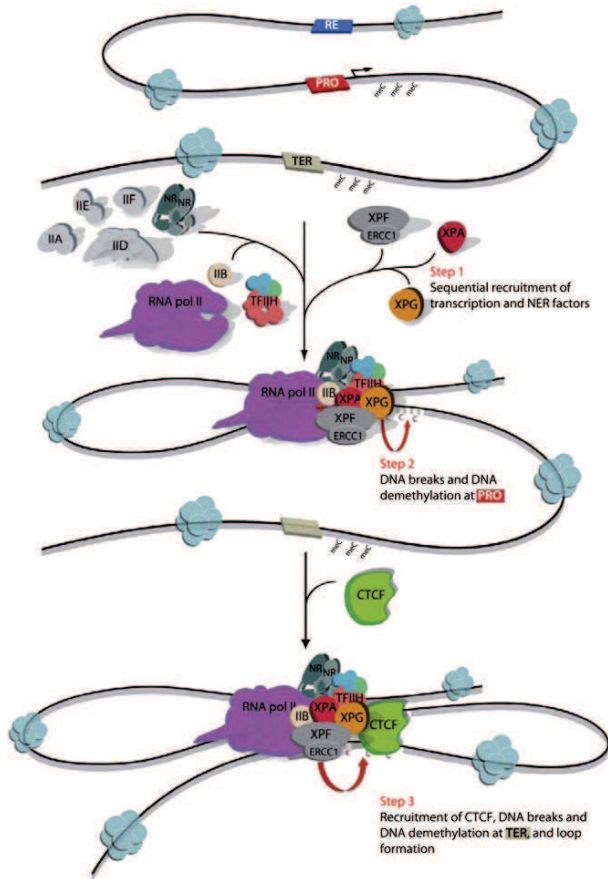
## EXPERIMENTAL PROCEDURES

### Cell Culture

HeLa Silencix cells (Tebu-Bio, provided by D. Biard) were used, including Ctrl (Ctrl2, BD690), ShXPF, and ShXPA cells. ShXPG and the corresponding control (Ctrl1, SiLuc) cells were provided by K. Tanaka. SiCTCF and Ctrl3 cells were obtained by respectively transfecting HeLa cells with SMART pool siRNA targeting *CTCF* or scrambled siRNA (Dharmacon) at a final concentration of 100 nM. XPG/WT and XPG/E791A cells, as well as XPF/WT and XPF/D676A cells (gift from O. Scharer and W. Vermeulen), were obtained by the stable transfection of the corresponding wild-type or mutated XPG and XPF constructs (Staresincic et al., 2009) in SV40-immortalized XP-G (XPCS1RO) and XP-F (XP2YO) (Ellison et al., 1998) (Yagi and Takebe, 1983).

All cells were cultured in appropriate medium. Cells were incubated with red phenol-free medium containing 10% charcoal-treated fetal calf serum (FCS)

percentage of the input. Analysis of the unmethylation of DNA (UnMedIP) at -65 kb, Pro, Ter, and +323 kb of the *RARβ2* locus for Ctrl1, ShXPG (A2), Ctrl2, ShXPA, ShXPF (B2), XPG/WT, XPG/E791A XPF/WT, XPF/D676A (C2), Ctrl3, SiCTCF (D2) t-RA-treated cells. UnMedIP was performed using the UnMethylCollector kit (Active Motif). MseI-digested genomic DNA was selected using magnetic beads conjugated with CXXC domains, and the resulted samples containing unmethylated DNA were analyzed by qPCR. Each series of UnMedIP is representative of two independent experiments as indicated by standard deviation, and values are expressed as percentage of the input. BioChIP and UnMedip were performed at the indicated times except for XPG/WT, XPG/E791A, XPF/WT, and XPF/D676A (T[h] corresponds to the peaks of factors recruitment in Figures 3A2-3B5). (E) Schematic representation of the methylated (open circle) or unmethylated (closed circle) CpG dinucleotides (in red) and DNA breaks (closed triangle) along the *RARβ2* promoter for WT (red), XPG/E791A (light blue), and XPF/D676A (dark blue) cells. DNA breaks were localized at Pro by pyrosequencing the samples generated by Bio-ChIP. The methylated status of CpG dinucleotides at Pro was measured using bisulfite converted genomic DNA from XPG/WT, XPG/E791A and XPF/WT, XPF/D676A fibroblasts and pyrosequencing.



**Figure 6. A Role for XPG and XPF in the CTCF-Dependent Gene Looping for Optimal Gene Expression**

Upon ligand induction, nuclear receptor (NR) targets its responsive element (RE). The transactivation complex is then formed at the promoter (PRO) including the arrival of GTFs (IIA, IIB, IIE, IIF, IIH) and pol II. After the formation of the PIC, NER factors (XPA, XPG, and XPF/ERCC1) are sequentially recruited (step 1). XPG (red arrow) initiates DNA break (dotted DNA) and demethylation (meC-C) at PRO (step 2), thus favoring the CTCF recruitment and gene looping between PRO and the terminator TER (step 3); concomitantly, XPF (red arrow) induces DNA break (dotted DNA) and DNA demethylation at TER.

and 40  $\mu\text{g/ml}$  gentamycin. Cells were treated with 10  $\mu\text{M}$  t-RA into the same medium.

#### Antibodies

RNA pol II (7C2), RAR (9A6), XPA (1E9), TBP (3G3), and XPG (1B5) antibodies were from IGBMC antibody facilities. TFIIB (C-18), XPF (H-300), and Biotin (33) antibodies were obtained from Santa-Cruz Biotechnology,  $\beta$ -tubulin from Millipore, and CTCF (ab70303) from Abcam.

#### Reverse Transcription and Quantitative PCR

Total RNA was isolated from several cell lines using a GenElute Mammalian Total RNA Miniprep kit (Sigma) and reverse transcribed with SuperScript II reverse transcriptase (Invitrogen). The quantitative PCR was done using the Lightcycler 480 (Roche). The primer sequences for *RAR $\beta$ 2* and glyceraldehyde 3-phosphate dehydrogenase (*GAPDH*) genes used in qPCR are available upon request. The *RAR $\beta$ 2* mRNA expression represents the ratio between

values obtained from treated and untreated cells normalized against the housekeeping *GAPDH* mRNA.

#### Chromatin Immunoprecipitation

Cells were crosslinked at room temperature (RT) for 10 min with 1% formaldehyde. Chromatin was prepared and sonicated on ice for 30 min using a Bioruptor (Diagenode) as previously described (Le May et al., 2010). Samples were immunoprecipitated with antibodies at 4°C overnight, and protein G Sepharose beads (Upstate) were added, incubated for 4 hr at 4°C, and sequentially washed. Protein-DNA complexes were eluted, and DNA fragments were purified using QIAquick PCR purification kit (QIAGEN) and analyzed by quantitative PCR.

#### Biotin-ChIP

Crosslinked cells following 1% formaldehyde treatment for 10 min at RT were permeabilized with cytonin (Active Motif) for 30 min at RT. After extensive washes with phosphate buffer salt (PBS), terminal deoxynucleotidyl transferase (TdT) reaction was performed using Biotin-16-dUTP (Roche) and 60 units of recombinant enzyme rTdT (Promega). TdT reaction was stopped with specific stop buffer (Chemicon International) for 15 min at RT. After extensive washes with PBS, the resulting samples were sonicated on ice for 20 min (40 cycles, pulse 10 s, pause 20 s) using a Bioruptor (Diagenode) and immunoprecipitated using anti-Biotin antibodies and treated as described in the ChIP protocol. DNA fragments were purified using QIAquick PCR purification kit (QIAGEN) and analyzed by quantitative PCR using sets of primers.

#### Quantitative Chromosome Conformation Capture

The quantitative chromosome conformation capture (q3C) assay was performed as previously described (Vernimmen et al., 2007). Briefly, cells were crosslinked at RT for 10 min with 2% formaldehyde. Crosslinked chromatin was digested in the appropriate restriction buffer by 400 units of enzyme HindIII. The restriction enzyme mixture was stopped, diluted in ligation buffer, and incubated with the highly concentrated T4 DNA ligase (Roche) for 4 hr at 16°C. The crosslinking was heat reversed, and DNA fragments were purified. Undigested DNA or digested DNA without ligation step was used as negative control. Moreover, we used the endogenous *xpb* locus, which has been reported to adopt the same spatial conformation in different tissues (Vernimmen et al., 2007), as an internal positive control. All q3C results were normalized by data from *xpb* analysis (see Figure S2G), controlling for changes in nuclear size, chromatin density, and crosslinking efficiency. Primers and probes were designed as follows: a universal sequence-specific Taqman probe and corresponding reverse primer on a fixed restriction fragment (Ter or M1) were used in combination with different forward primers specific for the other restriction fragments (see upper panel, Figure 3). Quantitative 3C templates (200 ng) were used for Taqman/PCR reaction using the universal PCR Master Mix and the Lightcycler 480 from Roche.

#### Unmethylated DNA Immunoprecipitation

Genomic DNA was extracted using GenElute Mammalian Genomic DNA Mini-prep Kit (Sigma). Unmethylation of 5mC on the *RAR $\beta$ 2* locus was measured by digesting 2  $\mu\text{g}$  genomic DNA with 10 units of MseI (Fermentas) and by using the UnMethylcollector kit (Active Motif). Unmethylated DNA Immunoprecipitation Kit (UnMeDIP) is based on the affinity of the three zinc-coordinating CXXC domains, localized in chromatin-associated proteins such as DNA methyltransferase 1 (DNMT1) or mixed lineage leukemia (MLL), that specifically bind nonmethylated CpG sites. The resulting samples were selected using magnetic beads conjugated with CXXC domains, extensively washed, and analyzed by quantitative PCR.

#### Pyrosequencing

##### Sequencing Analyses

Pyrosequencing was performed on a Pyromark Q96 ID Platform using the PSQ Gold SQA reagent kit (QIAGEN). Briefly, 20  $\mu\text{l}$  amplified DNA products were mixed with 2  $\mu\text{l}$  streptavidin Sepharose beads (Amersham Biosciences AB), 38  $\mu\text{l}$  binding buffer (10 mM Tris-HCl [pH 7.6], 2 M NaCl, 1 mM EDTA, and 0.1% Tween 20), and 20  $\mu\text{l}$  H<sub>2</sub>O, followed by shaking for 10 min. Then

the immobilized biotinylated PCR products-streptavidin Sepharose beads complex was captured using the QIAGEN Pyromark Q96 Work Station. Single-strand DNA purification was achieved by sequential washes with 75% ethanol, 0.2 M NaOH, and 10 mM Tris-acetate, pH 7.6 buffer. The unbiotinylated strand was dissociated and discarded. The immobilized single biotinylated strands were released to a 96-well microtiter plate, which was preadded with 38  $\mu$ l annealing buffer and 2  $\mu$ l complementary sequencing primer. The plate was incubated at 80°C for 2 min, followed by slow cooling to RT. The processed mixture was loaded onto the PyroMark ID system equipped with PyroMark ID software. The resulting pyrograms and associated sequences were generated and analyzed automatically using PSQ 96 SQA software (QIAGEN).

#### Methylation Analyses

Pyrosequencing was performed using a PyroMark Q96 ID Pyrosequencing instrument (QIAGEN). Pyrosequencing assays were designed using MethPrimer (<http://www.urogene.org/methprimer/index1.html>). Genomic DNA (200 ng) was treated with EZ DNA Methylation-Gold Kit (Zymo Research Corporation) and amplified using 1.25 U Platinum Taq DNA Polymerase (Invitrogen) and 1  $\mu$ M each forward and reverse primers in a 50  $\mu$ l reaction volume with 0.2 mM dNTP and 1.5 mM MgCl<sub>2</sub>. PCR conditions were 94°C for 3 min, then 40 cycles of 94°C for 30 s, Tm for 30 s, and 72°C for 30 s, followed by a 9 min extension at 72°C. Biotin-labeled single-stranded amplicons were isolated using the QIAGEN Pyromark Q96 Work Station and underwent pyrosequencing with 0.5  $\mu$ M primer. The percent methylation for each of the CpGs within the target sequence was calculated using PyroQ cpG Software (QIAGEN). All methylation analyses were performed in duplicate.

All the sets of primers and probes targeting different regions of *RAR $\beta$ 2* gene and used for the qPCR analysis are available upon request.

#### SUPPLEMENTAL INFORMATION

Supplemental Information includes four figures, Supplemental Experimental Procedures, and Supplemental References and can be found with this article at <http://dx.doi.org/10.1016/j.molcel.2012.05.050>.

#### ACKNOWLEDGMENTS

We thank C. Braun for her technical expertise; D. Vernimmen for his help in designing the q3C experiment; and D. Biard, O. Scharer, and W. Vermeulen for providing the different cells and the IGBMC cell culture facilities. We also are grateful to F. Coin, E. Compe, A. Singh, and L. Dubs for fruitful discussions and to G. Richards for a critical reading of the manuscript. This study was supported by an ERC Advanced grant (to J.-M.E.), l'Agence Nationale de la Recherche (N#ANR1 08MIEN-022-03), l'Association de la Recherche contre le Cancer, the Institut National du Cancer (INCA-2008-041), l'Association de Recherche sur le Diabète (ARD grant 2009- 01 to P.B.), and research grants from NovoNordisk France ("PharmacoEpigenetics of Growth" 2009-2011 to P.B.) and Pfizer France ("Candidate Gene Search for Shortness Gene Variations" GETC-2010-2028326 to P.B.). N.L.M. is a recipient from the INSERM; I.I. is granted by the l'Association de la Recherche contre le Cancer; D.F. is the recipient of a Young Investigator support grant from INSERM.

Received: September 29, 2011

Revised: April 30, 2012

Accepted: May 22, 2012

Published online: July 5, 2012

#### REFERENCES

Araujo, S.J., Tirode, F., Coin, F., Pospiech, H., Syvaoja, J.E., Stucki, M., Hubscher, U., Egly, J.M., and Wood, R.D. (2000). Nucleotide excision repair of DNA with recombinant human proteins: definition of the minimal set of factors, active forms of TFIIH, and modulation by CAK. *Genes Dev.* **14**, 349–359.

Barreto, G., Schafer, A., Marhold, J., Stach, D., Swaminathan, S.K., Handa, V., Doderlein, G., Maltry, N., Wu, W., Lyko, F., et al. (2007). Gadd45a promotes

epigenetic gene activation by repair-mediated DNA demethylation. *Nature* **445**, 671–675.

Brivanlou, A.H., and Darnell, J.E., Jr. (2002). Signal transduction and the control of gene expression. *Science* **295**, 813–818.

Compe, E., Malerba, M., Soler, L., Marescaux, J., Borrelli, E., and Egly, J.M. (2007). Neurological defects in trichothiodystrophy reveal a coactivator function of TFIIH. *Nat. Neurosci.* **10**, 1414–1422.

de Laat, W., and Grosfeld, F. (2003). Spatial organization of gene expression: the active chromatin hub. *Chromosome Res.* **11**, 447–459.

Deng, W., and Blobel, G.A. (2010). Do chromatin loops provide epigenetic gene expression states? *Curr. Opin. Genet. Dev.* **20**, 548–554.

Ellison, A.R., Nouspikel, T., Jaspers, N.G., Clarkson, S.G., and Gruenert, D.C. (1998). Complementation of transformed fibroblasts from patients with combined xeroderma pigmentosum-Cockayne syndrome. *Exp. Cell Res.* **243**, 22–28.

Gaillard, P.H., and Wood, R.D. (2001). Activity of individual ERCC1 and XPF subunits in DNA nucleotide excision repair. *Nucleic Acids Res.* **29**, 872–879.

Gehring, M., Reik, W., and Henikoff, S. (2009). DNA demethylation by DNA repair. *Trends Genet.* **25**, 82–90.

Gillet, L.C., and Scharer, O.D. (2006). Molecular mechanisms of mammalian global genome nucleotide excision repair. *Chem. Rev.* **106**, 253–276.

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.

Hu, Q., Kwon, Y.S., Nunez, E., Cardamone, M.D., Hutt, K.R., Ohgi, K.A., Garcia-Bassets, I., Rose, D.W., Glass, C.K., Rosenfeld, M.G., et al. (2008). Enhancing nuclear receptor-induced transcription requires nuclear motor and LSD1-dependent gene networking in interchromatin granules. *Proc. Natl. Acad. Sci. USA* **105**, 19199–19204.

Jaspers, N.G., Raams, A., Silengo, M.C., Wijgers, N., Niedernhofer, L.J., Robinson, A.R., Giglia-Mari, G., Hoogstraten, D., Kleijer, W.J., Hoeijmakers, J.H., et al. (2007). First reported patient with human ERCC1 deficiency has cerebello-oculo-facio-skeletal syndrome with a mild defect in nucleotide excision repair and severe developmental failure. *Am. J. Hum. Genet.* **80**, 457–466.

Jin, S.G., Guo, C., and Pfeifer, G.P. (2008). GADD45A does not promote DNA demethylation. *PLoS Genet.* **4**, e1000013. <http://dx.doi.org/10.1371/journal.pgen.1000013>.

Ju, B.G., Lunyak, V.V., Perissi, V., Garcia-Bassets, I., Rose, D.W., Glass, C.K., and Rosenfeld, M.G. (2006). A topoisomerase II $\beta$ -mediated dsDNA break required for regulated transcription. *Science* **312**, 1798–1802.

Kornberg, R.D. (2007). The molecular basis of eukaryotic transcription. *Proc. Natl. Acad. Sci. USA* **104**, 12955–12961.

Kraemer, K.H., Patronas, N.J., Schiffmann, R., Brooks, B.P., Tamura, D., and DiGiovanna, J.J. (2007). Xeroderma pigmentosum, trichothiodystrophy and Cockayne syndrome: a complex genotype-phenotype relationship. *Neuroscience* **145**, 1388–1396.

Lalle, P., Nouspikel, T., Constantinou, A., Thorel, F., and Clarkson, S.G. (2002). The founding members of xeroderma pigmentosum group G produce XPG protein with severely impaired endonuclease activity. *J. Invest. Dermatol.* **118**, 344–351.

Le May, N., Mota-Fernandes, D., Velez-Cruz, R., Iltis, I., Biard, D., and Egly, J.M. (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.

Ma, D.K., Jang, M.H., Guo, J.U., Kitabatake, Y., Chang, M.L., Pow-Anpongkul, N., Flavell, R.A., Lu, B., Ming, G.L., and Song, H. (2009). Neuronal activity-induced Gadd45b promotes epigenetic DNA demethylation and adult neurogenesis. *Science* **323**, 1074–1077.

Murrell, A., Heeson, S., and Reik, W. (2004). Interaction between differentially methylated regions partitions the imprinted genes Igf2 and H19 into parent-specific chromatin loops. *Nat. Genet.* **36**, 889–893.



- Niedernhofer, L.J., Garinis, G.A., Raams, A., Lalai, A.S., Robinson, A.R., Appeldoorn, E., Odijk, H., Oostendorp, R., Ahmad, A., van Leeuwen, W., et al. (2006). A new progeroid syndrome reveals that genotoxic stress suppresses the somatotroph axis. *Nature* **444**, 1038–1043.
- Nouspikel, T. (2009). DNA repair in mammalian cells: nucleotide excision repair: variations on versatility. *Cell. Mol. Life Sci.* **66**, 994–1009.
- O'Donovan, A., Davies, A.A., Moggs, J.G., West, S.C., and Wood, R.D. (1994). XPG endonuclease makes the 3' incision in human DNA nucleotide excision repair. *Nature* **371**, 432–435.
- O'Sullivan, J.M., Tan-Wong, S.M., Morillon, A., Lee, B., Coles, J., Mellor, J., and Proudfoot, N.J. (2004). Gene loops juxtapose promoters and terminators in yeast. *Nat. Genet.* **36**, 1014–1018.
- Perkins, K.J., Lusic, M., Mitar, I., Giacca, M., and Proudfoot, N.J. (2008). Transcription-dependent gene looping of the HIV-1 provirus is dictated by recognition of pre-mRNA processing signals. *Mol. Cell* **29**, 56–68.
- Schmitz, K.M., Schmitt, N., Hoffmann-Rohrer, U., Schafer, A., Grummt, I., and Mayer, C. (2009). TAF12 recruits Gadd45a and the nucleotide excision repair complex to the promoter of rRNA genes leading to active DNA demethylation. *Mol. Cell* **33**, 344–353.
- Sijbers, A.M., de Laat, W.L., Ariza, R.R., Biggerstaff, M., Wei, Y.F., Moggs, J.G., Carter, K.C., Shell, B.K., Evans, E., de Jong, M.C., et al. (1996). Xeroderma pigmentosum group F caused by a defect in a structure-specific DNA repair endonuclease. *Cell* **86**, 811–822.
- Singh, B.N., and Hampsey, M. (2007). A transcription-independent role for TFIIIB in gene looping. *Mol. Cell* **27**, 806–816.
- Spilianakis, C.G., and Flavell, R.A. (2004). Long-range intrachromosomal interactions in the T helper type 2 cytokine locus. *Nat. Immunol.* **5**, 1017–1027.
- Staresinic, L., Fagbemi, A.F., Enzlin, J.H., Gourdin, A.M., Wijgers, N., Dunand-Sauthier, I., Giglia-Mari, G., Clarkson, S.G., Vermeulen, W., and Scharer, O.D. (2009). Coordination of dual incision and repair synthesis in human nucleotide excision repair. *EMBO J.* **28**, 1111–1120.
- Tan-Wong, S.M., French, J.D., Proudfoot, N.J., and Brown, M.A. (2008). Dynamic interactions between the promoter and terminator regions of the mammalian BRCA1 gene. *Proc. Natl. Acad. Sci. USA* **105**, 5160–5165.
- Tapias, A., Auriol, J., Forget, D., Enzlin, J.H., Scharer, O.D., Coin, F., Coulombe, B., and Egly, J.M. (2004b). Ordered conformational changes in damaged DNA induced by nucleotide excision repair factors. *J. Biol. Chem.* **279**, 19074–19083.
- Tiwari, V.K., McGarvey, K.M., Licchesi, J.D., Ohm, J.E., Herman, J.G., Schubeler, D., and Baylin, S.B. (2008). PcG proteins, DNA methylation, and gene repression by chromatin looping. *PLoS Biol.* **6**, 2911–2927. <http://dx.doi.org/10.1371/journal.pbio.0060306>.
- Vernimmen, D., De Gobbi, M., Sloane-Stanley, J.A., Wood, W.G., and Higgs, D.R. (2007). Long-range chromosomal interactions regulate the timing of the transition between poised and active gene expression. *EMBO J.* **26**, 2041–2051.
- Wang, G., Balamotis, M.A., Stevens, J.L., Yamaguchi, Y., Handa, H., and Berk, A.J. (2005). Mediator requirement for both recruitment and postrecruitment steps in transcription initiation. *Mol. Cell* **17**, 683–694.
- Yagi, T., and Takebe, H. (1983). Establishment by SV40 transformation and characteristics of a cell line of xeroderma pigmentosum belonging to complementation group F. *Mutat. Res.* **112**, 59–66.

## Publication 3

### The NER factor XPC regulates histone modifications on promoter of activated genes

Ilitis I, Le May N, Egly JM.

### En préparation

#### Résumé

La protéine XPC, connue pour son rôle dans la reconnaissance des lésions de l'ADN lors de la GG-NER, a été décrite dans la transcription comme étant le 1<sup>er</sup> facteur NER recruté avec la machinerie transcriptionnelle, au niveau du promoteur de gènes activés. Nous avons montré que son absence ou sa mutation pouvait perturber l'expression du gène RAR $\beta$ 2.

En absence de XPC et pour la mutation R579St du gène XPC, présente chez des patients *XP*, nous avons montrés qu'au promoteur et au terminateur de gènes activés, certaines modifications post-traductionnelles des histones (PTMs) nécessaires au remodelage de la chromatine sont absentes. En effet, un état chromatinien non favorable à la transcription (hétérochromatine) est observé avec une forte méthylation de H3K9 et l'absence de H3K4me et H3K9ac.

Nous avons également étudié le recrutement d'enzymes responsables de ces PTMs au niveau du promoteur de RAR $\beta$ 2. Nous avons montré que le silencing ou la mutation de XPC, provoquait une absence du recrutement de l'acétyltransférase GCN5, de la déméthylase KDM5b et de la méthyltransférase Suv39H1. Ces perturbations sont à mettre en corrélation avec des perturbations de la méthylation de H3K4 et de H3K9 et la déacétylation de H3K9. Nous avons donc montré que XPC intervient dans l'établissement d'un contexte d'euchromatine au promoteur, afin de potentialiser la transcription de gènes activés.

# The NER factor XPC regulates histone modifications on promoter of activated genes

Izarn Iltis<sup>1</sup>, Nicolas Le May<sup>1</sup> and Jean-Marc Egly<sup>1</sup>.

<sup>1</sup>IGBMC, Department of Functional Genomics and Cancer biology, CNRS/INSERM/Université de Strasbourg, BP 163, 67404 Illkirch Cedex, C. U. Strasbourg, France.

\*Corresponding authors.

Egly Jean-Marc: [jean-marc.egly@igbmc.fr](mailto:jean-marc.egly@igbmc.fr)

Key words: XPC, Histone post-translational modification, Gene expression, NER factors, *Xeroderma Pigmentosum*.

## Introduction

Upon gene activation, a host of proteins including RNA polymerase II (RNA pol II), general transcription factors (GTFs), cofactors, and chromatin-remodelling factors assemble around the promoter and contribute to accurate expression of protein-coding genes. Gene expression is compromised by genotoxic stress that challenges genome integrity and requires the function of several DNA repair pathways to remove DNA lesions. This implies that there must be connections between the seemingly disparate events of transcription and DNA repair to orchestrate the expression and repair of genes. A link between DNA repair and transcription was first established, with the discovery of the multi-protein complex TFIIH, which was both a basal transcription and nucleotide-excision repair (NER) factor (Schaeffer et al., 1993) (Feaver et al., 1993). The relationship between NER and transcription is even tighter since recent papers have revealed additional and unexpected roles for the other NER factors (XPC, XPA, XPG, XPF/ERCC1) upon transcription (Barreto et al., 2007) (Schmitz et al., 2009) (Le May et al., 2010b).

Nucleotide excision repair (NER) is one of the DNA repair pathways that can eliminate a wide variety of DNA lesions, including cyclobutane pyrimidine dimers (CPDs) and pyrimidine-pyrimidone (6-4) photoproducts (6-4PPs) induced by UV irradiation (Friedberg et al., 2006). The general sensor of DNA lesions is *xeroderma pigmentosum C* (XPC) protein, which recognizes a wide variety of DNA damage. By detecting the damaged DNA, XPC initiates the nucleotide excision repair pathway (Sugasawa, 2010). XP-C was also implicated in chromatin modification, during DNA repair. In yeast, the ATP-chromatin remodelling complex SWI/SNF interacts with the homologs of XPC-hHR23B after UV irradiation (Gong et al., 2006). Moreover, certain NER factors, including XPC, are stimulated by GCN5 and E2F1, allowing acetylation of H3K9 (H3K9ac) in the vicinity of DNA lesion (Guo et al., 2011). In addition, XPC was shown to regulate the DDB2-complex, which mediated ubiquitylation around DNA damage to promote chromatin decondensation (Takedachi et al., 2010) (Luijsterburg et al., 2012).

NER factors have been shown to be associated with the RNA polymerase transcription machinery and are sequentially recruited (XPC/CSB followed by RPA/XPA and XPG/XPF) at the promoter of activated genes (Le May et al., 2010b). The recruitment of the NER factors at the promoter is necessary in order to achieve optimal chromatin remodelling, including histone post-translational modifications (PTMs) as well as DNA demethylation (Barreto et al.,



2007) (Schmitz et al., 2009) (Le May et al., 2010b). For instance, XPG and XPF are essential for establishing CTCF-dependant chromatin loop between promoter and terminator of the *RARb2* gene upon transactivation (Le May et al., 2012). Recently, a DNA repair complex containing XPC has been identified to function as an Oct4/Sox2 coactivator in ES and iPS cells. Indeed, both ES and iPS cells depend on transcription frameworks governed by specific transcription factors namely Oct4, Sox2 and Nanog (Jaenisch and Young, 2008). Depletion of XPC in these cells compromised the pluripotency in these cells and promoted differentiation (Fong et al., 2011). However, the functions of the NER factors and especially XPC upon transcription remain elusive.

The understanding of the roles played by these NER factors is particularly important since mutations on these proteins originate with several human autosomal recessive disorders, including *Xeroderma pigmentosum* (XP), cockayne syndrome (CS) and trichothiodystrophy (TTD). XP is clinically characterized by photosensitivity and cancer predisposition for skin exposed to sunlight (Bootsma et al., 2002). The biochemical defect in XP-C patients is the most prevalent NER defective groups comprising approximately 90% of characterized classical XP patients. XP-C patients are highly susceptible to develop tumors on sunlight-exposed areas of the skin (1000 fold) (Cleaver, 2005). However, they rarely exhibit neurological disorders or developmental defects.

In the present study, using *in vivo* experiments, we investigated the role of XPC in the establishment of euchromatin marks during transactivation of nuclear target genes.

## Results

### **XPC deficiency disturbs RA responsive genes expression.**

We have previously demonstrated the involvement of NER factors in the transactivation of *RARβ2* regulated by the retinoic acid receptor (RAR) (Le May et al., 2010b). We measured the implication of XPC in the RAR-mediated transcriptional activity using stable HeLa cell lines, constitutively expressing either a shRNA directed against XPC (shXPC), or scramble shRNA (shCtrl). In parallel, XP-C patients cells (XPC/R579St) bearing a point homozygous mutation leading to the truncated XPC/R579St and known to be deficient for the NER pathway were also used (Bernardes de Jesus et al., 2008).

We first observed a peak of mRNA *RARβ2* synthesis, 6 hours after treatment with 10mM trans retinoic acid (t-RA), in all three cells lines by RT-PCR although the transactivation was lower in shXPC and XPC/R579St cells (Figure 1). We thus determined the transcriptome profiles of these cells compared to shCtrl using microarrays at 6 hours post-treatment. In shCtrl cells, 77 genes show a fold induction > 2 (supplemental Table 1) including several well-characterized genes, such as *RARβ2* or *CYP26*. Surprisingly, ~90% of genes activated in shCtrl showed a lower induction in shXPC (Figure 1 A). Interestingly, all these genes are also lower transactivated in XPC/R579St cells. We confirmed such XPC-dependency by analysing *RARβ2* (Figure 1 B, C) as well as the *CYP26* (Figure S1 A) induction by RT-qPCR. Their induction is lower up regulated (~50%) in shXPC and XPC/R579St cells compared to respective control (Figure 1B, 1C and S1A, S1B). In contrast, 10% of genes, such as the immediate early gene *GADD45* (Figure S1 B), activated in shCtrl cells showed a similar induction even in absence of XPC.

In shCtrl and XPC/WT cells, we observed a correlation between *RARβ2* induction and recruitment of the transcriptional machinery (including RAR, pol II and TFIIH subunit Cdk7) together with the NER factors (XPC, XPA, XPG) at promoter and terminator regions by ChIP assays (Figure 1D1, D3, D5, D6 and Figure 1E3, E6, E9, E12). We have recently demonstrated that the presence of NER factors XPG and XPF was necessary for the recruitment of the chromatin organizer CTCF that participate to the formation of gene looping between promoter and terminator (Le May et al., 2012). Our ChIP assay showed such recruitment of CTCF at the promoter and the terminator of *RARβ2* (Figure 1 D3, D7, E6 and E12). In shXPC as well as in XPC/R579St cells, only the basal transcription machinery is recruited at the *RARβ2* promoter, while XPC, XPA, XPG as well as CTCF were not significantly detected (Figure 1D2, D4, E3 and E6). In those cells, no factors were observed after t-RA induction at the terminator (Figure 1D6, D8, E7 and E10).

We thus next were wondering whether the absence and/or mutation of XPC would prevent chromatin looping. Quantitative chromatin conformation captures assay (q3C) (Vernimmen et al., 2007), along the *RARβ2* locus allowed the detection of long-range interaction between promoter and terminator in control cells (shCtrl and XPC/WT) (Figure 1 H). In shXPC and XPC/R579St, in accordance with the absence of XPC and CTCF recruitment detected by ChIP, no spatial interaction between promoter and terminator was revealed (Figure 1 H).

In parallel, a second XP-C patient cell (XPC/P334H), bearing a single amino acid change at position 334 (P334H) (Gozukara et al., 2001) was analysed showing an unaffected *RARβ2* and *CYP26* mRNA expression compared to the control (Figure 1C, S1A). Indeed, we noticed that the XPC/P334H mutation did not prevent transcription and NER machineries recruitment at promoter and terminator (Figure 1 E2, E5, E8 and E11). In accordance with ChIP assays, q3C assays allowed the detection of chromatin rearrangement between promoter and terminator in XPC/P334H fibroblast (Figure 1 H).

Altogether, our data indicated that the presence of XPC is necessary for the transactivation of a subset of RAR-target genes. Its presence at the promoter and the terminator of *RARβ2* seems important for the recruitment of CTCF and consequently the formation of chromatin loop, that optimizes the RNA synthesis.

#### **XPC and histones post-translational modifications (PTMs).**

We previously observed a correlation between NER factors recruitment and histones post-translational modifications (PTMs) during the active transcription (Schmitz et al., 2009) (Le May et al., 2010a). We thus investigated the establishment of the euchromatin upon transcription at the *RARβ2* promoter and terminator. In shCtrl and XPC/WT, we observed transcriptional active marks including an increase of H3K4me and H3K9ac concomitantly with the decrease of H3K9me on the promoter and the terminator at 6h (Figure 2 B1, B4, C3 and C6). Interestingly, in shXPC and XPC/R579St cells, the absence of XPC recruitment resulted in a complete dysregulation of euchromatin marks. We observed an increase of H3K9me and a decrease of H3K4me and H3K9ac on the promoter and the terminator (Figure 2 B2, B5, C1 and C4). As expected, in XPC/P334H cells, the histone active landmarks remain conserved (Figure 2 C2, C5).

To further investigate, if XPC is responsible for these histones PTMs, we used silenced XPA cells (shXPA), in which XPC was normally expressed and recruited on the promoter (Figure S2) (Le May et al., 2010b). Upon t-RA activation, all the basal transcriptional machinery, as well as XPC, are recruited at the promoter of *RARβ2* gene at 3 hours (Figure S2 B1 and B2). We also observed the re-establishment of euchromatin marks (H3K4me, H3K9ac and decrease of H3K9me) only at the promoter (Figure 2 B3). For the terminator, maintenance of the heterochromatin landscape is observed with exclusively H3K9me in accordance with the absence of XPC at the terminator (Figure 2 B6, S2B4).

These results suggest that XPC is involved in establishment of active chromatin marks at the promoter of activated gene.

**XPC is involved in the recruitment of histones modifying enzymes responsible for histone PTMs.**

We next investigated, whether histones modifying enzymes (HME) responsible of such histones modifications could be regulated by the presence of XPC. SET1/MLL1 di/tri-methylates H3K4 (Wysocka et al., 2003) (Yokoyama et al., 2004) that can be demethylated by KDM5B (Seward et al., 2007). For H3K9 methylation, many H3K9 methyltransferases have been characterized, including Suv39H1 and G9a (Fritsch et al., 2010). The demethylation of H3K9 is catalysed by KDM4D (Shin and Janknecht, 2007). GCN5 mediated the H3K9 acetylation (Brownell et al., 1996) (Kuo et al., 1996) whereas HDAC1 deacetylates it.

We observed in shCtrl cell line at 6h, the recruitment of MLL1/SET1 in parallel with H3K4me peak (Figure 3 A1 and A4). Recruitment of KDM4D correlated with decrease of H3K9me (Figure 3 A2 and A8). We also observed concomitantly to H3K9ac the presence of GCN5 at the *RARβ2* promoter (Figure 3 A3 and A6).

In shXPC and XPC/R579St cells, GCN5, Suv39H1 and KDM5B were not detected by ChIP at the promoter (Figure 3 D6, D5, D7 and E6, E5, E6). However, we observed the presence of the heterochromatin HME G9a (Figure 3 D5 and E5). In contrast, in shXPA and in XPC/P334H cells, no significant interference in the recruitment of the different HME was observed (Figure 3 C and F). Indeed, as in controls (shCtrl and XPC/WT), GCN5, Suv39H1 and KDM5B were present at the *RARβ2* promoter, correlated with the different related histones PTMs (H3K4me and H3K9ac).

Altogether, the above data strongly suggest an involvement of XPC in the recruitment of HME related to transcriptional active histones PTMs at the *RARβ2* promoter.

## Experimental Procedure

### Cell Culture:

HeLa Silencix cells (Tebu-Bio, provided by D. Biard) were used including shCtrl (BD690), shXPC cells and shXPA cells. Wild-type FB789 (XPC/WT), GM14867 (XPC/R579St) and GM02096 (XPC/P334H) derived from XP patients. All cells were cultured in appropriate medium and were treated with 10 $\mu$ M all-trans-retinoic acid (t-RA) into the corresponding medium.

### Antibodies:

RNA pol II (7C2), RAR (9A6), XPA (1E9) and XPG (1B5) antibodies were from IGBMC antibody facilities. CDK7 (C-19), XPC (D-18) and HDAC1 (C-19) antibodies were obtained from Santa-Cruz Biotechnology and CTCF (ab70303), KDM5b (ab56759), Suv39H1 (ab12405), GCN5 (ab18381) and KDM4D (ab93694) from Abcam. Antibodies MLL1 and SET1 were obtained from Bethyl and antibodies against H3K9me2/3, H3K9Ac, H3K4me3 and G9a (C6H3) were obtained from Cell signalling.

### Reverse Transcription and Quantitative PCR

Total RNA was isolated from several cell lines using a GenElute Mammalian Total RNA Miniprep kit (Sigma) and reverse transcribed with SuperScript II reverse transcriptase (Invitrogen). The quantitative PCR was done using the Lightcycler 480 (Roche). The primer sequences for *RAR $\beta$ 2* and *Glyceraldehyde 3-phosphate deshydrogenase (GAPDH)* genes used in qPCR are available upon request. The *RAR $\beta$ 2* mRNA expression represents the ratio between values obtained from treated and untreated cells normalized against the housekeeping *GAPDH* mRNA.

### Chromatin Immunoprecipitation (ChIP):

Cells were cross-linked at room temperature (RT) for 10 min with 1% formaldehyde. Chromatin was prepared and sonicated on ice 40min using a Bioruptor (Diagenode) as previously described (Le May et al., 2010b). Samples were immunoprecipitated with antibodies at 4°C overnight and protein G-sepharose beads (Upstate) were added, incubated

4 hours at 4°C and sequentially washed. Protein-DNA complexes were eluted and DNA fragments were purified using QIAquick PCR purification kit (QIAGEN) and analyzed by quantitative PCR.

### **Quantitative Chromosome Conformation Capture (q3C):**

The quantitative chromosome conformation capture assay was performed as previously described (Vernimmen et al., 2007). Briefly, cells were cross-linked at RT for 10 min with 2% formaldehyde. Cross-linked chromatin was digested in the appropriate restriction buffer by 400 units of enzyme *HindIII*. The restriction enzyme mixture was stopped, diluted in ligation buffer and incubated with the highly concentrated T4 DNA ligase (Roche) for 4 hours at 16°C. The cross-linking was heat-reversed and DNA fragments were purified. Undigested DNA or digested DNA without ligation step was used as negative controls. Moreover, we used the endogenous *xpb* locus that has been reported to adopt the same spatial conformation in different tissues as an internal positive control. All q3C results were normalized by data from *xpb* analysis, controlling for changes in nuclear size, chromatin density and cross-linking efficiency. Primers and probes were designed as follows: a universal sequence-specific Taqman probe and corresponding reverse primer on a fixed restriction fragment (Ter) were used in combination with different forward primers specific for the other restriction fragments. Quantitative 3C templates (200ng) were used for Taqman/PCR reaction using the universal PCR Master Mix and the Lightcycler 480 from Roche.

### **Microarray analysis**

Total RNA from shXPC, XPC/R579St and shCtrl cells were extracted before or 6 hours after t-RA treatment (10mM) using GenElute Mammalian Total RNA Miniprep kit (Sigma). Two independent experiments were performed. RNA quality was verified by analysis on the 2100 Bioanalyzer (Agilent). Biotinylated single strand cDNA targets were prepared, starting from 250 ng of total RNA, using the Ambion WT Expression Kit (Cat # 4411974) and the Affymetrix GeneChip® WT Terminal Labeling Kit (Cat # 900671), according to Affymetrix recommendations.

Following fragmentation and end-labeling, 1.9 µg of cDNAs were hybridized for 16 hours at 45°C on GeneChip® Human Gene 1.0 ST arrays (Affymetrix) interrogating 28,853 genes represented by approximately 27 probes spread across the full length of the gene. The chips

were washed and stained in the GeneChip® Fluidics Station 450 (Affymetrix) and scanned with the GeneChip® Scanner 3000 7G (Affymetrix). Finally, raw data (.CEL Intensity files) were extracted from the scanned images using the Affymetrix GeneChip® Command Console (AGCC) version 3.1. CEL files were further processed with Affymetrix Expression Console software version 1.1 to calculate probeset signal intensities using Robust Multi-array Average (RMA) algorithms with default settings.

## **Acknowledgments**

We thank D. Vernimmen for his help in designing the q3C experiment; D. Biard for providing the different cells and the IGBMC cell culture facilities. We also thank F. Coin, E. Compe and A. Singh for fruitful discussions. This study was supported by an ERC Advanced grant (to J.M.E.), l'Agence Nationale de la Recherche (N#ANR1 08MIEN-022-03), l'Association de la Recherche contre le Cancer, the Institut National du Cancer (INCA-2008-041) and l'Association de Recherche sur le Diabète (ARD grant 2009- 01 to PB) and research grants from NovoNordisk France ("Pharmacogenetics of Growth" 2009-2011 to PB) and Pfizer France ("Candidate Gene Search for Shortness Gene Variations » GETC-2010-2028326 to PB). I.I. is granted by the l'Association de la Recherche contre le Cancer; N.L.M. is a recipient from the INSERM.

## References

- Barreto, G., Schafer, A., Marhold, J., Stach, D., Swaminathan, S.K., Handa, V., Doderlein, G., Maltry, N., Wu, W., Lyko, F., *et al.* (2007). Gadd45a promotes epigenetic gene activation by repair-mediated DNA demethylation. *Nature* 445, 671-675.
- Bernardes de Jesus, B.M., Bjoras, M., Coin, F., and Egly, J.M. (2008). Dissection of the Molecular Defects Caused by Pathogenic Mutations in the DNA Repair Factor XPC. *Mol Cell Biol* 28, 7225-7235.
- Bootsma, D., Kraemer, K.H., Cleaver, J.E., and Hoeijmakers, J.H.J. (2002). Nucleotide excision repair syndromes: xeroderma pigmentosum, Cockayne syndrome, and trichothiodystrophy. *The Genetic Basis of Human Cancer* *The Genetic Basis of Human Cancer In: (second ed.), B. Vogelstein and K.W. Kinzler, Editors.,* 211–237.
- Brownell, J.E., Zhou, J., Ranalli, T., Kobayashi, R., Edmondson, D.G., Roth, S.Y., and Allis, C.D. (1996). Tetrahymena histone acetyltransferase A: a homolog to yeast Gcn5p linking histone acetylation to gene activation. *Cell* 84, 843-851.
- Cleaver, J.E. (2005). Cancer in xeroderma pigmentosum and related disorders of DNA repair. *Nat Rev Cancer* 5, 564-573.
- Feaver, W.J., Svejstrup, J.Q., Bardwell, L., Bardwell, A.J., Buratowski, S., Gulyas, K.D., Donahue, T.F., Friedberg, E.C., and Kornberg, R.D. (1993). Dual roles of a multiprotein complex from *S. cerevisiae* in transcription and DNA repair. *Cell* 75, 1379-1387.
- Fong, Y.W., Inouye, C., Yamaguchi, T., Cattoglio, C., Grubisic, I., and Tjian, R. (2011). A DNA repair complex functions as an oct4/sox2 coactivator in embryonic stem cells. *Cell* 147, 120-131.
- Friedberg, E.C., Aguilera, A., Gellert, M., Hanawalt, P.C., Hays, J.B., Lehmann, A.R., Lindahl, T., Lowndes, N., Sarasin, A., and Wood, R.D. (2006). DNA repair: from molecular mechanism to human disease. *DNA Repair (Amst)* 5, 986-996.
- Fritsch, L., Robin, P., Mathieu, J.R., Souidi, M., Hinaux, H., Rougeulle, C., Harel-Bellan, A., Ameyar-Zazoua, M., and Ait-Si-Ali, S. (2010). A subset of the histone H3 lysine 9 methyltransferases Suv39h1, G9a, GLP, and SETDB1 participate in a multimeric complex. *Mol Cell* 37, 46-56.
- Gong, F., Fahy, D., and Smerdon, M.J. (2006). Rad4-Rad23 interaction with SWI/SNF links ATP-dependent chromatin remodeling with nucleotide excision repair. *Nat Struct Mol Biol* 13, 902-907.
- Gozukara, E.M., Khan, S.G., Metin, A., Emmert, S., Busch, D.B., Shahlavi, T., Coleman, D.M., Miller, M., Chinsomboon, N., Stefanini, M., *et al.* (2001). A stop codon in xeroderma pigmentosum group C families in Turkey and Italy: molecular genetic evidence for a common ancestor. *J Invest Dermatol* 117, 197-204.
- Guo, R., Chen, J., Mitchell, D.L., and Johnson, D.G. (2011). GCN5 and E2F1 stimulate nucleotide excision repair by promoting H3K9 acetylation at sites of damage. *Nucleic Acids Res* 39, 1390-1397.
- Jaenisch, R., and Young, R. (2008). Stem cells, the molecular circuitry of pluripotency and nuclear reprogramming. *Cell* 132, 567-582.
- Kuo, M.H., Brownell, J.E., Sobel, R.E., Ranalli, T.A., Cook, R.G., Edmondson, D.G., Roth, S.Y., and Allis, C.D. (1996). Transcription-linked acetylation by Gcn5p of histones H3 and H4 at specific lysines. *Nature* 383, 269-272.
- Le May, N., Egly, J.M., and Coin, F. (2010a). True lies: the double life of the nucleotide excision repair factors in transcription and DNA repair. *J Nucleic Acids* 2010.
- Le May, N., Fradin, D., Iltis, I., Bougneres, P., and Egly, J.M. (2012). XPG and XPF endonucleases trigger chromatin looping and DNA demethylation for accurate expression of activated genes. *Mol Cell* 47, 622-632.



Le May, N., Mota-Fernandes, D., Velez-Cruz, R., Iltis, I., Biard, D., and Egly, J.M. (2010b). 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.

Luijsterburg, M.S., Lindh, M., Acs, K., Vrouwe, M.G., Pines, A., van Attikum, H., Mullenders, L.H., and Dantuma, N.P. (2012). DDB2 promotes chromatin decondensation at UV-induced DNA damage. *J Cell Biol* 197, 267-281.

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

Schmitz, K.M., Schmitt, N., Hoffmann-Rohrer, U., Schafer, A., Grummt, I., and Mayer, C. (2009). TAF12 recruits Gadd45a and the nucleotide excision repair complex to the promoter of rRNA genes leading to active DNA demethylation. *Mol Cell* 33, 344-353.

Seward, D.J., Cubberley, G., Kim, S., Schonewald, M., Zhang, L., Tripet, B., and Bentley, D.L. (2007). Demethylation of trimethylated histone H3 Lys4 in vivo by JARID1 JmjC proteins. *Nat Struct Mol Biol* 14, 240-242.

Shin, S., and Janknecht, R. (2007). Diversity within the JMJD2 histone demethylase family. *Biochem Biophys Res Commun* 353, 973-977.

Sugasawa, K. (2010). Regulation of damage recognition in mammalian global genomic nucleotide excision repair. *Mutat Res* 685, 29-37.

Takedachi, A., Saijo, M., and Tanaka, K. (2010). DDB2 complex-mediated ubiquitylation around DNA damage is oppositely regulated by XPC and Ku and contributes to the recruitment of XPA. *Mol Cell Biol* 30, 2708-2723.

Vernimmen, D., De Gobbi, M., Sloane-Stanley, J.A., Wood, W.G., and Higgs, D.R. (2007). Long-range chromosomal interactions regulate the timing of the transition between poised and active gene expression. *Embo J* 26, 2041-2051.

Wysocka, J., Myers, M.P., Laherty, C.D., Eisenman, R.N., and Herr, W. (2003). Human Sin3 deacetylase and trithorax-related Set1/Ash2 histone H3-K4 methyltransferase are tethered together selectively by the cell-proliferation factor HCF-1. *Genes Dev* 17, 896-911.

Yokoyama, A., Wang, Z., Wysocka, J., Sanyal, M., Aufiero, D.J., Kitabayashi, I., Herr, W., and Cleary, M.L. (2004). Leukemia proto-oncoprotein MLL forms a SET1-like histone methyltransferase complex with menin to regulate Hox gene expression. *Mol Cell Biol* 24, 5639-5649.

## Figure Legend

### Figure 1

**(A)** Transcriptome profiling of shCtrl (blue), shXPC (red) and XPC/R579St (green) cells treated with t- RA (10mM; 6 hours) using the whole transcript coverage Affymetrix Human Gene 1.0 ST arrays. Based on two independent experiments, these genes show an up-regulation > to 2 ( $\pm$ SEM) compared to t=0 hour (F-test,  $P < 0.005$ ). **(B, C)** Relative mRNA expression of *RAR $\beta$ 2* in shCtrl and shXPC cells (B) or XPC/WT, P334H and R579St fibroblasts (C) measured at different time points after treatment with t-RA (10 $\mu$ M). Error bars represent the standard deviation (SD) of three independent experiments. **(D-E)** ChIP monitoring the t-RA-dependant occupancy of RAR, POL II, CDK7, XPC, XPA, XPG and CTCF on *RAR $\beta$ 2* promoter and terminator from shCtrl and shXPC (D1-D8) and from XPC/WT, R579St and P334H fibroblasts (E1-E12) chromatin extracts. Each series of ChIP is representative of at least two independent experiments as indicated by standard deviation, and values are expressed as percentage of the input. **(H)** Intra-chromosomal interactions between terminal region of *RAR $\beta$ 2* gene (Ter) and upstream or downstream elements. q3C assays were performed using *HindIII*-digested, fixed chromatin from stable shCtrl and shXPC HeLa cells and XPC/WT, R579St and P334H fibroblast cells untreated (T0) or treated (T(h)) with t-RA (10  $\mu$ M), as indicated. The bar chart (Y axis) shows the enrichment of PCR product (%) normalised to the enrichment within the human *xpb* gene (=100%). These independent graphs represent a measure of the association between the upstream -65kb, promoter (Pro) regions, the downstream M1 and M2, +257kb, +323kb, +385kb regions to the terminal (Ter) regions. Data shown represent the average of a least three independent experiments using Taqman/Real-Time PCR. Each PCR was performed at least three times and averaged as indicated by standard deviation. Signals were normalised to the total amount of DNA used, estimated with an amplicon located within a *HindIII* fragment in *RAR $\beta$ 2* gene.

### Figure 2

**(A)** Schematic representation of the XPC protein with the mutations and the interacting domains. **(B-C)** ChIP monitoring the t-RA-dependent occupancy of di-/trimethylated and acetylated histones H3K9, trimethylated histone H3K4 at promoter and terminator regions of *RAR $\beta$ 2* gene from shCtrl (B1-B4), shXPC (B2-B5) and shXPA (B3-B6) and from XPC/WT (C3-

C6), R579St (C1-C4), and P334H (C2-C5) chromatin extracts as indicated. Each series of CHIP is representative of two independent experiments.

**Figure 3**

**(A-F)** CHIP monitoring the t-RA-dependent occupancy of di-/trimethylated and acetylated histones H3K9, trimethylated histone H3K4, G9a, Suv39H1, KDM4D (JMJD2D), GCN5, HDAC1, SET1, MLL1 and KDM5B (Jarid1b) on *RAR $\beta$ 2* promoter from stable shCtrl (A), shXPC (D), shXPA (F) HeLa and XPC/WT (B), P334H (C) and R579St (E) fibroblasts chromatin extracts as indicated. Each series of CHIP is representative of two independent experiments.

## Supplemental information

### Figure S1

**(A)** Relative mRNA expression of *CYP26* in shCtrl and shXPC cells or XPC/WT, P334H and R579St fibroblasts measured at different time points after treatment with t-RA (10 $\mu$ M). Error bars represent the standard deviation (SD) of three independent experiments. **(B)** Relative mRNA expression of *Gadd45 $\alpha$*  in shCtrl and shXPC cells or XPC/WT, P334H and R579St fibroblasts measured at different time points after treatment with t-RA (10 $\mu$ M). Error bars represent the standard deviation (SD) of three independent experiments. **(C)** Western blotting analyses of XPC and TBP, were performed from chromatin extracts of shCtrl, shXPC, shXPA HeLa cells and XPC/WT, P334H and R579St fibroblasts. TBP is used as positive nuclear marker.

### Figure S2

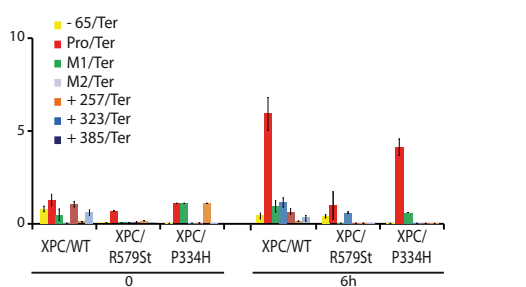
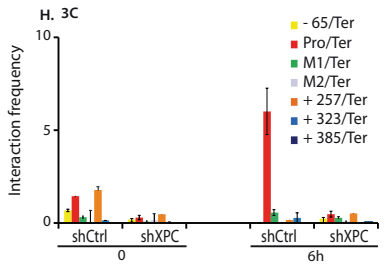
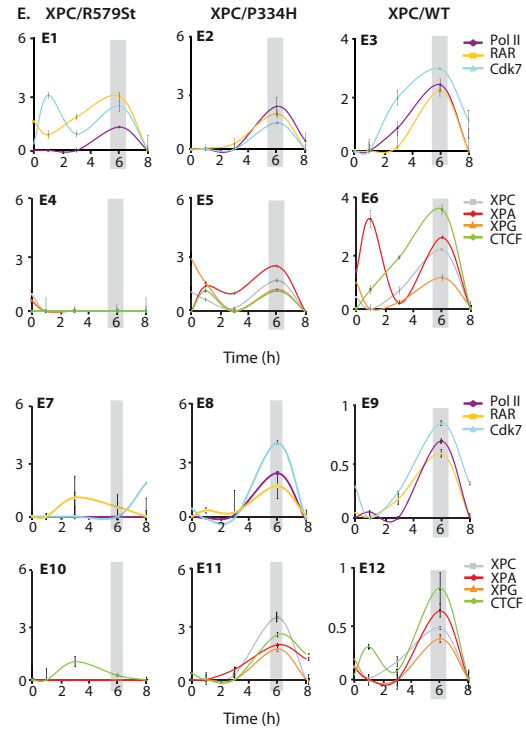
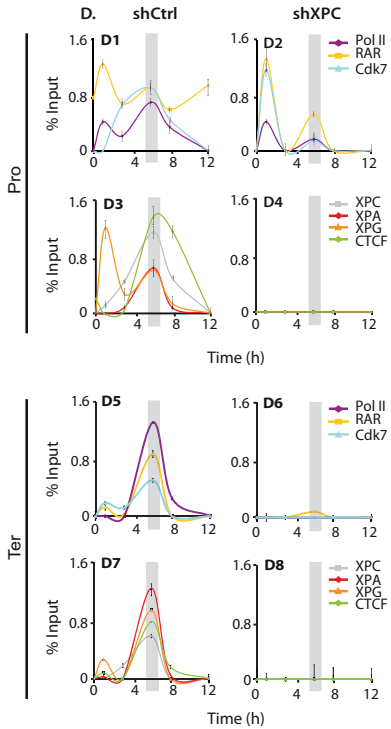
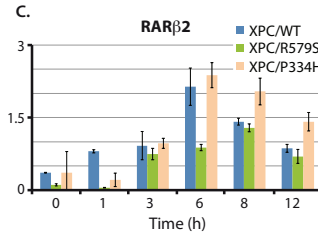
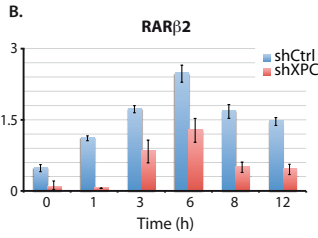
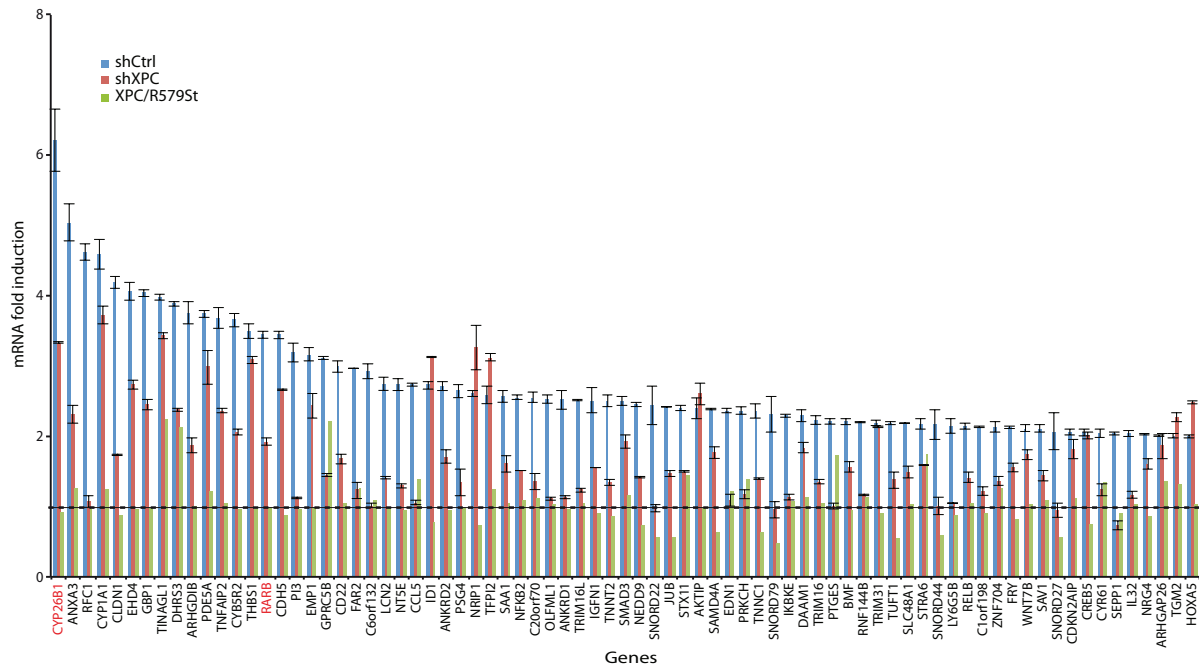
**(A)** Relative mRNA expression of *RAR $\beta$ 2*, *CYP26* and *Gadd45 $\alpha$*  in shXPA cells measured at different time points after treatment with t-RA (10 $\mu$ M). Error bars represent the standard deviation (SD) of three independent experiments. **(B)** ChIP monitoring the t-RA-dependant occupancy of RAR, POL II, CDK7, XPC, XPA, XPG and CTCF on *RAR $\beta$ 2* promoter and terminator from shXPA cells chromatin extracts. Each series of ChIP is representative of at least two independent experiments as indicated by standard deviation, and values are expressed as percentage of the input.

### Table S1

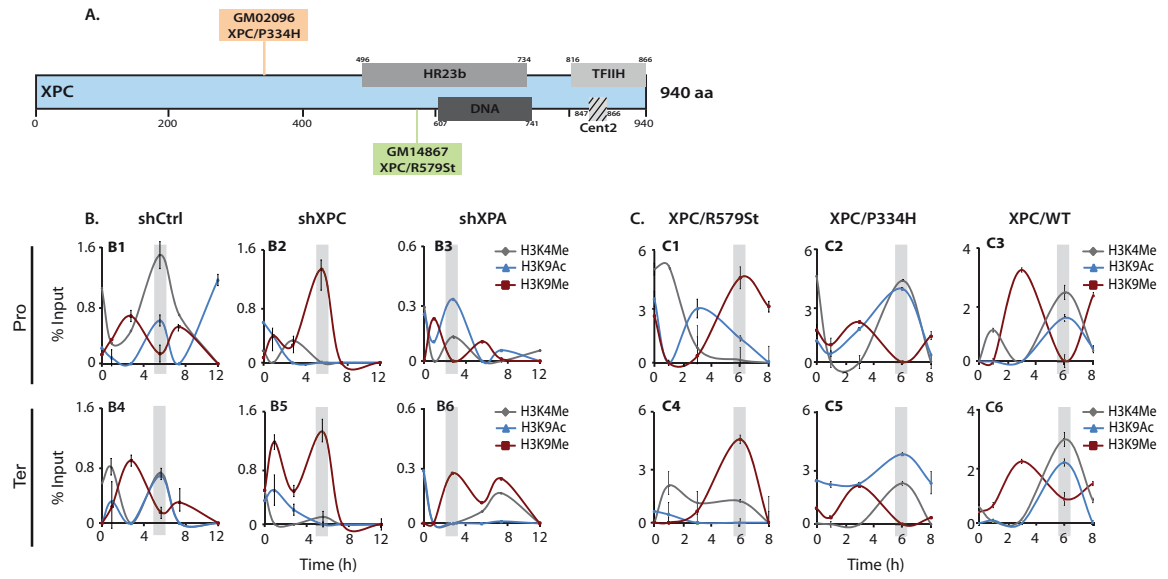
Summary of RA-induced genes in shCtrl, shXPC and XPC/R579St cells analyzed by microarrays, 6 hours after treatment with t-RA. We show arbitrary genes with an up-regulation > to 2. These results are the means of two independent experiments ( $\pm$ SEM).

**Figure 1**

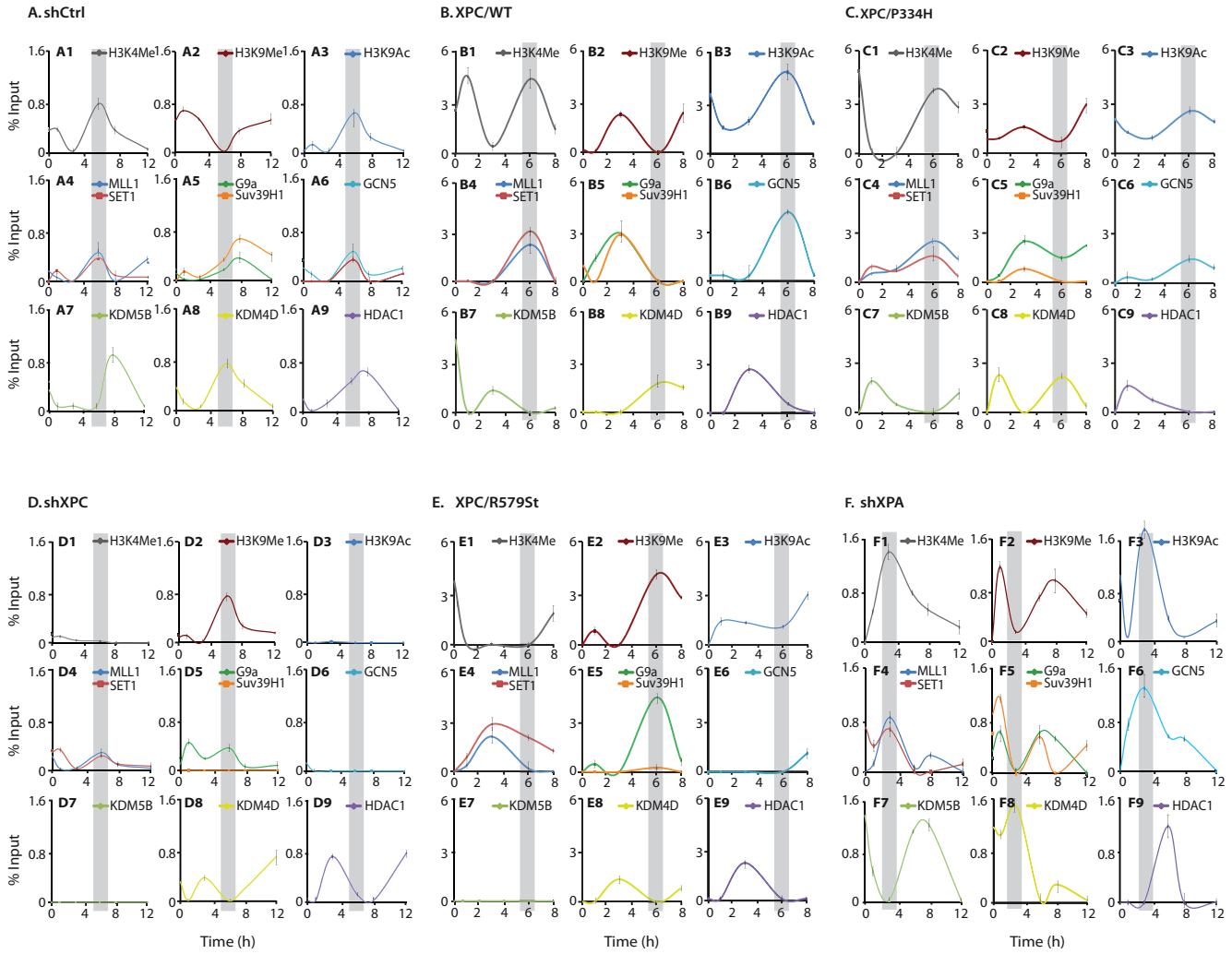
**A. Microarray**



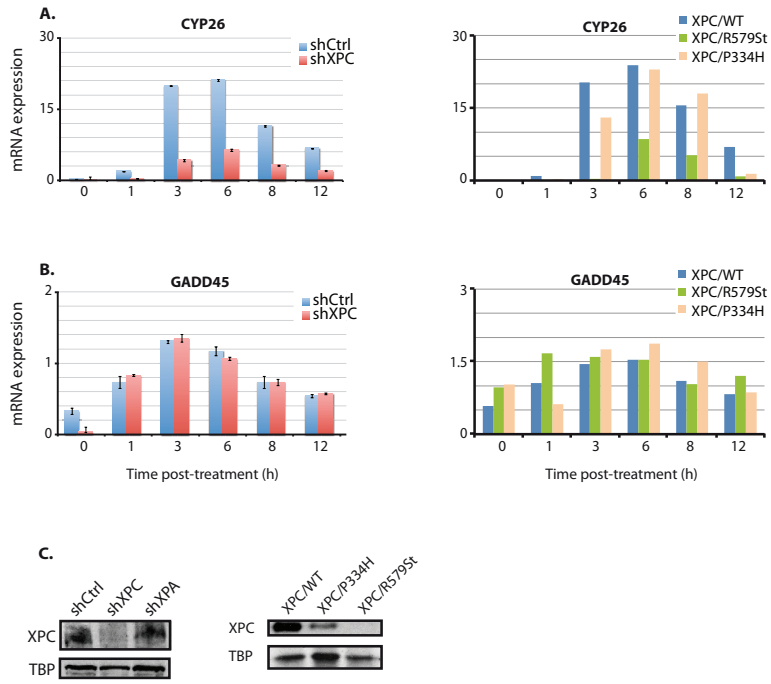
**Figure 2**



**Figure 3**



**Figure S1**





**Figure S2**

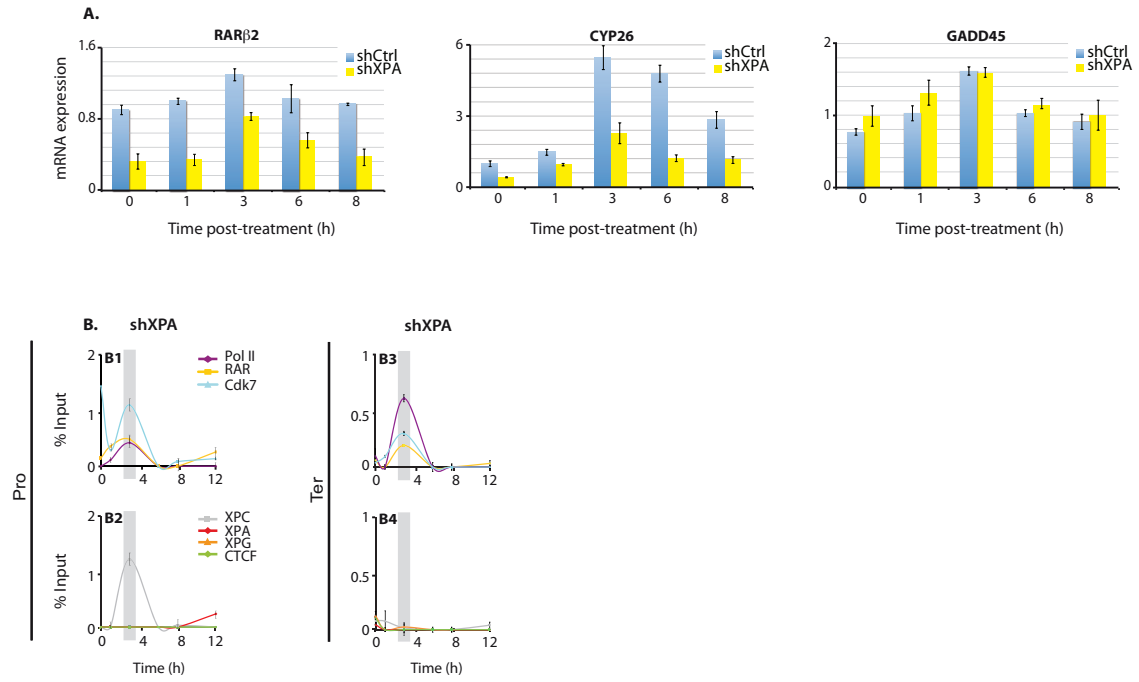


Table S1

Gene Symbol	ShCtrl		ShXPC		XPC/R579St
	Fold induction Mean	(+/-) SEM	Fold induction Mean	(+/-) SEM	Fold induction
CYP26B1	6,220	0,447	3,336	0,016	0,932
ANXA3	5,046	0,264	2,315	0,135	1,258
RFC1	4,618	0,114	1,067	0,077	0,976
CYP1A1	4,592	0,220	3,731	0,132	1,233
CLDN1	4,189	0,083	1,738	0,010	0,872
EHD4	4,063	0,128	2,733	0,061	0,975
GBP1	4,042	0,050	2,455	0,078	0,978
TINAGL1	3,980	0,039	3,431	0,046	1,829
DHRS3	3,890	0,036	2,382	0,026	2,129
ARHGDI1	3,759	0,160	1,868	0,108	1,014
PDE5A	3,739	0,045	2,986	0,244	1,218
TNFAIP2	3,681	0,150	2,361	0,034	1,056
CYB5R2	3,656	0,097	2,057	0,040	0,963
THBS1	3,498	0,103	3,092	0,055	0,962
RARB	3,450	0,053	1,923	0,048	0,984
CDH5	3,446	0,059	2,666	0,011	0,887
PI3	3,198	0,139	1,126	0,018	0,962
EMP1	3,169	0,103	2,441	0,180	0,977
GPRC5B	3,107	0,020	1,450	0,027	2,209
CD22	2,996	0,083	1,673	0,065	1,055
FAR2	2,968	0,005	1,231	0,116	1,261
C6orf132	2,924	0,110	1,012	0,031	1,097
LCN2	2,746	0,091	1,407	0,012	0,982
NT5E	2,739	0,087	1,296	0,031	0,938
CCL5	2,738	0,025	1,055	0,029	1,395
ID1	2,729	0,060	3,132	0,010	0,779
ANKRD2	2,716	0,070	1,708	0,092	0,941
PSG4	2,648	0,096	1,349	0,194	0,982
NRIP1	2,604	0,043	3,256	0,314	0,739
TFPI2	2,590	0,124	3,118	0,054	1,232
SAA1	2,572	0,086	1,613	0,121	1,040
NFKB2	2,552	0,037	1,510	0,004	1,090
C20orf70	2,551	0,072	1,360	0,118	1,120
OLFML1	2,535	0,063	1,115	0,023	1,032
ANKRD1	2,517	0,130	1,125	0,022	0,959
TRIM16L	2,515	0,008	1,232	0,025	1,056
IGFN1	2,513	0,185	1,548	0,001	0,905
TNNT2	2,503	0,080	1,343	0,041	0,864
SMAD3	2,498	0,064	1,927	0,094	1,159
NEDD9	2,458	0,034	1,421	0,016	0,736
SNORD22	2,438	0,278	0,982	0,057	0,560
JUB	2,423	0,003	1,464	0,039	0,573

Table S1 bis

Gene Symbol	ShCtrl		ShXPC		XPC/R579St
	Fold induction Mean	(+/-) SEM	Fold induction Mean	(+/-) SEM	Fold induction
STX11	2,402	0,034	1,500	0,018	1,458
AKTIP	2,392	0,145	2,602	0,152	0,976
SAMD4A	2,387	0,017	1,771	0,085	0,619
EDN1	2,362	0,030	1,086	0,083	1,208
PRKCH	2,360	0,053	1,174	0,063	1,395
TNNC1	2,357	0,096	1,397	0,009	0,634
SNORD79	2,309	0,255	0,957	0,122	0,473
IKBKE	2,296	0,028	1,130	0,036	1,075
DAAM1	2,293	0,089	1,836	0,077	1,136
TRIM16	2,238	0,066	1,355	0,032	1,056
PTGES	2,212	0,037	1,008	0,042	1,727
BMF	2,207	0,045	1,561	0,076	0,935
RNF144B	2,204	0,010	1,166	0,009	1,129
TRIM31	2,195	0,041	2,131	0,010	0,901
TUFT1	2,188	0,022	1,379	0,119	0,541
SLC48A1	2,183	0,004	1,488	0,085	1,037
STRA6	2,175	0,074	1,588	0,003	1,736
SNORD44	2,159	0,209	1,010	0,124	0,584
LY6G5B	2,151	0,106	1,052	0,002	0,887
RELB	2,150	0,046	1,415	0,074	1,041
C1orf198	2,135	0,018	1,224	0,063	0,907
ZNF704	2,132	0,072	1,367	0,065	1,260
FRY	2,130	0,019	1,562	0,067	0,814
WNT7B	2,114	0,044	1,735	0,066	1,038
SAV1	2,110	0,048	1,441	0,078	1,087
SNORD27	2,069	0,258	0,946	0,099	0,555
CDKN2AIP	2,063	0,045	1,819	0,137	1,104
CREB5	2,060	0,053	2,009	0,045	0,754
CYR61	2,046	0,062	1,238	0,091	1,347
SEPP1	2,041	0,021	0,726	0,063	0,896
IL32	2,033	0,041	1,165	0,045	1,020
NRG4	2,033	0,012	1,606	0,074	0,870
ARHGAP26	2,006	0,006	1,863	0,181	1,358
TGM2	2,006	0,036	2,270	0,065	1,326
HOXA5	2,000	0,027	2,480	0,025	1,023

## DISCUSSION

## DISCUSSION

La synthèse de l'ADN en ARN messenger est le résultat d'une cascade d'évènements, nécessitant l'intervention de plus de 200 protéines. Tous ces facteurs interagissent entre eux, afin de former la machinerie de transcription, permettant le bon déroulement de cette étape (Brivanlou and Darnell, 2002) (Kornberg, 2007). Une meilleure compréhension de la composition de ces complexes et de leurs fonctions, permettrait de mieux étudier le processus d'expression des gènes. Néanmoins, étant donnée la grande diversité de familles de gènes, il est très difficile de généraliser une découverte à l'ensemble du génome. Dans ce travail de thèse, nous nous sommes intéressés plus particulièrement aux gènes de classe II régulés par les récepteurs nucléaires, en utilisant comme modèle le gène du récepteur à l'acide rétinoïque RAR $\beta$ 2.

### **Recrutement séquentiel des facteurs NER sur les gènes activés.**

Lors de ce travail, nous avons démontré dans un premier temps, que tous les facteurs NER sont recrutés avec la machinerie transcriptionnelle au niveau du promoteur, lors de la transcription de gènes activés (Le May et al., 2010a). Ce recrutement, observé en l'absence d'attaques génotoxiques, est abolit lors de la présence d'inhibiteur de la transcription (DRB). Cette inhibition démontre clairement, que le complexe observé est spécifique de la transcription et n'est pas le résultat du mécanisme de réparation. De plus, le recrutement que l'on observe est caractéristique des facteurs NER. En effet, ni des protéines de la BER (Ogg1), ni des protéines de la réparation des mésappariements (MutS homologue 2), ne sont détectées au niveau des promoteurs de gènes activés.

Nous avons également montré que ce recrutement est séquentiel. Par l'utilisation de cellules n'exprimant pas les différents facteurs NER, nous avons mis en évidence un ordre séquentiel dans le recrutement de ces facteurs au promoteur. Le premier facteur NER recruté est XPC ; puis il est suivi des facteurs XPA, XPG et enfin XPF/ERCC1/Gadd45 $\alpha$ . Ces facteurs sont recrutés après la machinerie transcriptionnelle (RAR, ARN pol II, Médiateur, GTFs...), afin de permettre une synthèse efficace des ARN messenger.

Le complexe de transcription contenant les facteurs NER, est différent de ceux observés lors de la réparation TCR ou GGR. Premièrement, ce complexe est sensible aux inhibiteurs de

transcription (DRB), alors que celui de la GGR ne l'est pas. Deuxièmement, ce complexe ne nécessite pas le recrutement du facteur CSB pour être mis en place, ce qui le différencie de celui de la TCR. Troisièmement, notre complexe de transcription nécessite le facteur XPC, qui est le premier recruté sur le promoteur ; alors que pour la TCR, XPC n'est pas impliqué. Ces différents arguments démontrent que ce complexe est spécifique de la transcription et est différent de ceux décrits lors de la réparation.

Nous avons également remarqué une spécificité de recrutement en fonction de la région du gène. En effet, en absence du facteur CSB, les facteurs NER sont normalement recrutés sur le promoteur, mais aucun d'entre eux n'est détecté au niveau de régions distales (par exemple en élongation). Cette observation peut être mise en relation avec le fait que CSB stimulerait l'élongation de la transcription *in vitro* (Selby and Sancar, 1997) (Proietti-De-Santis et al., 2006). Lorsque XPC est absent, aucun recrutement des facteurs NER n'a lieu sur le promoteur. De plus, le facteur XPC n'est pas recruté lors de l'élongation. Cela suggère qu'il existe une spécificité dans la localisation du recrutement des facteurs NER, avec XPC au niveau du promoteur lors de l'initiation de la transcription, et CSB au niveau de régions distales lors de l'élongation.

Nous avons montré que l'absence des facteurs NER ne dérégule pas totalement la transcription des gènes activés. Ainsi, nous pouvons nous interroger sur le rôle des facteurs NER lors de la transcription. Effectivement, il subsiste toujours une transcription que l'on peut qualifier « de base » en l'absence des facteurs NER. Nous pensons donc que les facteurs NER sont impliqués dans un processus de potentialisation de la transcription. Lors de la formation du PIC, un remodelage de la chromatine et une déméthylation active de l'ADN ont été observés (Li et al., 2007). Des travaux récents ont présenté CSB comme un facteur de maintenance et de remodelage du nucléosome. C'est pourquoi, nous nous sommes interrogés sur l'implication des facteurs NER dans les processus de régulation épigénétique de la chromatine (Newman et al., 2006). Nous avons montré que le recrutement de la machinerie transcriptionnelle, des facteurs NER et de Gadd45 $\alpha$  était observé en parallèle de la mise en place d'un contexte d'euchromatine. En effet, nous avons démontré la déméthylation active de l'ADN et la mise en place de modifications post-traductionnelles des histones (H3K4me et H3K9/K14ac) au niveau du promoteur de gènes activés. En l'absence des facteurs NER, nous avons mis en évidence qu'un contexte d'hétérochromatine subsiste

au promoteur, avec la méthylation de H3K9 et une méthylation de l'ADN. D'après ces résultats, il est possible d'établir un lien entre modifications épigénétiques, facteurs NER et potentialisation de la transcription. Néanmoins, pour le moment, le lien entre facteurs NER et les différentes modifications de la chromatine reste indéfini.

L'implication des facteurs NER dans le mécanisme de la transcription a également été observée dans des lignées cellulaires de patients *Xeroderma pigmentosum*. Comme pour les cellules HeLa, le recrutement séquentiels des facteurs NER n'est pas indispensable pour la formation du PIC. Néanmoins, on observe qu'en l'absence de ces facteurs (lignées mutantes), la transactivation des gènes activés n'est pas optimale.

En fonction de la mutation, les phénotypes des patients peuvent être plus ou moins graves. Il a été démontré que les patients XP-C présentent des symptômes moins importants que les patients XP-A. Cela peut s'expliquer par le rôle du facteur XPC. En effet, lors de la transcription, XPC est le premier facteur recruté au niveau du promoteur. Il semble également que le recrutement des autres facteurs soit dépendant de sa présence sur le promoteur (Le May et al., 2010b). Néanmoins, les mutations de XPC ne perturbent pas autant les voies de réparation que les mutations de XPA. Effectivement, seul la réparation par GG-NER est perturbée par une mutation du facteur XPC. Concernant le facteur XPA, de nombreuses études ont démontré qu'il joue un rôle dans les deux voies de réparation par excision de nucléotides (TC-NER et GG-NER). Ses mutations affectent donc les deux voies de réparation. Les conséquences pour les patients XP-A, sont l'absence de mécanismes de compensation, capables de réparer les lésions de types dimères cyclobutanes de pyrimidines (CPD) ou photoproduits 6-4 pyrimidine pyrimidone (6-4 PP).

De plus, il a été récemment établi que le facteur XPC intervenait dans le développement embryonnaire, en stimulant les activateurs transcriptionnels Oct4/Sox2 dans la transcription du gène *Nanog* (Fong et al., 2011). Le gène *Nanog* est exprimé dans les cellules souches embryonnaires, et est connu pour être indispensable à l'acquisition de l'état de pluripotence. On peut donc imaginer, que des mutations de XPC peuvent affecter les premiers stades du développement cellulaire et donc être létales. Aucune forme sévère de patients XP-C n'a été décrite. On peut donc supposer que certaines mutations de XPC ne peuvent être viables.



### **Rôles des facteurs NER lors de la transcription de gènes activés.**

La découverte de nouveaux composants de la transcription fut un thème majeur de la recherche en biologie moléculaire ces dernières années. C'est pourquoi, après avoir mis en évidence l'implication des facteurs NER dans le mécanisme de transcription des gènes activés, nous nous sommes intéressés aux rôles que peuvent jouer ces facteurs.

Les facteurs XPG et XPF, sont connus dans la réparation comme ayant une activité endonucléase. Compte tenu de cela, nous avons décidé d'étudier comment ces facteurs sont impliqués dans l'expression optimale de gènes activés. Nous avons montré qu'après induction à l'acide rétinoïque, les facteurs généraux de transcription et l'ARN polymérase II étaient recrutés de manière concomitante avec les facteurs NER (XPA, XPG et XPF), au niveau du promoteur et du terminateur du gène  $RAR\beta 2$ . Le recrutement des facteurs NER au promoteur est séquentiel. Nous avons ainsi démontré que l'ordre de recrutement des facteurs est identique au terminateur, avec XPA qui précède XPG et XPF. Nous avons également mis en évidence, que l'activité endonucléase de XPG et XPF est cruciale pour le recrutement du facteur CTCF, au niveau du promoteur et du terminateur.

La présence des facteurs NER et de CTCF, aussi bien au promoteur qu'au terminateur, nous a amené à nous interroger sur la formation d'une boucle. Les boucles de chromatine résultent de l'interaction inter ou intra-chromosomal, entre différentes régions spécifiques des gènes (Hu et al., 2008). Plusieurs exemples ont été décrits, comme celui de la  $\beta$ -globine (de Laat and Grosveld, 2003), celui du gène T-helper de type 2 (TH2) (Spilianakis and Flavell, 2004), celui de BRCA1 (Tan-Wong et al., 2008) ou encore celui de GATA-4 (Tiwari et al., 2008). Nous avons montré que l'activité endonucléase des facteurs XPG et XPF, est nécessaire à la formation d'une boucle de chromatine, entre le promoteur et le terminateur des gènes activés. Nous avons également mis en évidence, que l'endonucléase XPG génère une coupure simple brin et une déméthylation de l'ADN au promoteur, permettant ainsi la mise en place de cette boucle d'ADN. La boucle de chromatine est dépendante de CTCF et est stabilisée par XPF, qui est capable de générer une coupure et une déméthylation du terminateur.

Comme mentionné précédemment, nous avons prouvé que les facteurs XPG et XPF sont indispensables pour le recrutement du facteur CTCF. CTCF est une protéine à doigts de zinc, impliquée dans la régulation de la chromatine, assurant la structuration des boucles de chromatine (Handoko et al., 2011). L'absence des différents facteurs NER perturbe le

recrutement du facteur CTCF. De plus, des mutations affectant l'activité catalytique de XPG et/ou XPF, empêchent également son recrutement et perturbent la formation de boucle de chromatine. On peut donc penser que ce n'est pas la présence des facteurs NER, mais bien leur activité, qui est nécessaire au recrutement de CTCF. De plus, le recrutement du facteur CTCF est indispensable au maintien de la boucle. Dans les cellules SiCTCF, on observe un recrutement des facteurs NER au promoteur du gène RAR $\beta$ 2 ; cependant l'absence de CTCF ne permet pas la formation de boucle.

Il semblerait que la coupure de l'ADN, provoquée par XPG, soit un facteur clé pour le recrutement de CTCF. De manière intéressante, on observe que la mutation de l'activité catalytique de XPG, prévient la formation de coupure au promoteur, la formation de boucle de chromatine ainsi que le recrutement de CTCF. Au contraire, la coupure de XPF au niveau du terminateur semble être nécessaire, non pas pour le recrutement de CTCF, mais pour en assurer sa stabilité. En effet, dans les cellules mutées sur le site catalytique de XPF, un recrutement très faible de CTCF est observé.

Il faut également noter le rôle crucial de XPA, sans qui, les facteurs XPG et XPF ne pourraient être recrutés, aussi bien au promoteur qu'au terminateur de gènes activés. On sait effectivement, qu'il arrive sur le promoteur avant ces deux facteurs (Le May et al., 2010b). De plus, nous avons montré qu'en son absence (ShXPA), ni XPG, ni XPF ne sont recrutés et aucune formation de boucle de chromatine n'est observée.

A partir de ces résultats, nous avons proposé le modèle suivant, pour l'expression des gènes activés. Après induction, les récepteurs nucléaires se fixent sur leurs éléments de réponse (RE). Le complexe de transactivation se forme alors au niveau du promoteur, avec l'arrivée des GTFs et de l'ARN pol II. Après la formation du PIC, les facteurs NER sont recrutés séquentiellement au promoteur. XPG initie alors les coupures et la déméthylation de l'ADN au promoteur, favorisant le recrutement de CTCF et la formation de boucle de chromatine entre le promoteur et le terminateur. De manière concomitante, XPF induit la coupure et la déméthylation de l'ADN au terminateur, permettant la stabilisation de CTCF et de la boucle de chromatine. Ce réarrangement de la chromatine potentialise alors la transcription des gènes activés.

Un tel scénario soulève de nombreuses questions sur les liens pouvant exister, entre la machinerie moléculaire, la déméthylation de l'ADN, les coupures et les boucles de

chromatines. En effet, le lien entre déméthylation et coupure, n'est pas encore défini. Nos résultats indiquent que la coupure et la déméthylation sont localisées dans les mêmes régions du brin transcrit (promoteur et terminateur). Néanmoins, nous avons montrés que la coupure est effectuée avant le processus de déméthylation des CG. En effet, dans les cellules mutées, au niveau du site catalytique de XPG, l'absence de coupure au promoteur est à mettre en parallèle avec l'absence de déméthylation des îlots CpG du promoteur et du terminateur. Savoir si les coupures sont un prérequis pour la déméthylation reste encore à étudier.

La question du rôle des facteurs NER dans la déméthylation est encore ouverte. En effet, le mécanisme de déméthylation est controversé, et il semblerait que deux voies aient été décrites (Cedar and Bergman, 2012). L'une d'entre elle ferait intervenir le mécanisme de réparation par BER, avec l'intervention de Gadd45 $\alpha$  (Barreto et al., 2007) (Jin et al., 2008), afin de resynthétiser une cytosine non modifiée (Gehring et al., 2009). Etant donné l'implication de facteurs de réparation dans le processus de déméthylation, la question du rôle des facteurs NER dans ce processus peut être posée. Néanmoins, nos résultats ne permettent pas pour le moment, une identification précise du lien entre XPG, XPF et la déméthylation de l'ADN.

### **XPC, un facteur clé dans l'établissement de l'euchromatine au niveau du promoteur des gènes activés.**

Après avoir démontré l'un des rôles des facteurs XPG et XPF dans la transcription de gènes activés, nous nous sommes intéressés plus particulièrement à celui du facteur XPC. Il s'agit d'un facteur déterminant, car comme nous l'avons souligné précédemment, il est le premier facteur NER à être recruté sur le promoteur lors de la transcription. De plus, son absence empêche la mise en place des autres facteurs NER et perturbe la transcription des gènes activés.

Nous avons montré, dans les cellules shXPC, que l'expression d'ARN des gènes dépendant de l'acide rétinoïque, est diminuée d'environ 50% en absence de XPC. Néanmoins, lorsque l'on étudie la population des gènes activés, nous avons constaté que, seule une petite centaine de gènes semble être sur-régulés par l'action de l'acide rétinoïque. Sur l'ensemble de ces gènes, l'absence de XPC, affecte la transactivation d'environ 90% d'entre eux. La même étude a été réalisée sur des cellules de patients XP-C, n'exprimant pas cette protéine. Des

résultats similaires ont été observés, avec une dérégulation des gènes activés par l'acide rétinoïque, en absence de XPC.

Nous avons démontré que XPC est un facteur recruté lors du pic d'expression d'ARN, au niveau du promoteur et du terminateur des gènes activés. Néanmoins, son absence empêche la formation de la boucle de chromatine entre le promoteur et le terminateur de RAR $\beta$ 2. Cela peut s'expliquer par une absence du recrutement des autres facteurs NER (XPA, XPG, XPF) sur le promoteur et sur le terminateur des gènes activés. Etant donné le caractère séquentiel du recrutement des facteurs NER, l'absence de XPC provoque l'absence des autres facteurs. On observe également la perturbation du recrutement de CTCF. Les absences de recrutements des facteurs NER empêchent alors la mise en place des processus de coupure et de déméthylation d'ADN, perturbant ainsi la formation de la boucle de chromatine. Ces résultats permettent de définir un premier rôle pour XPC. En effet, ce facteur sert de plateforme de recrutement pour les autres facteurs NER, afin de mettre en place une cascade de réactions, aboutissant à la formation de boucles de chromatine entre le promoteur et le terminateur de gènes activés.

En plus de son rôle dans la détection des lésions, la protéine XPC a été décrite dans le mécanisme de réparation par NER, comme intervenant dans des phénomènes de modifications des histones. Chez la levure, XPC interagit avec le complexe de remodelage de la chromatine ATP-dépendante SWI/SNF, après irradiation aux UV (Gong et al., 2006). Il est également stimulé par l'acétyltransférase GNC5 et E2F1, via l'acétylation de H3K9 (Guo et al., 2011b). De plus, il régule le complexe DDB2, qui en ubiquitinant l'histone H2, permet une décondensation de la chromatine (Takedachi et al., 2010) (Luijsterburg et al., 2012). Partant de cette hypothèse, nous avons étudié comment, lors de la transcription, le facteur XPC peut être impliqué dans les mécanismes de modifications post-traductionnelles des histones. Nous avons montré que certaines modifications post-traductionnelles des histones (PTMs), nécessaires à la mise en place d'un contexte d'euchromatine favorable à la transcription, sont absentes lors de la suppression de l'expression de XPC. En effet, lors de la régulation de la transcription, trois modifications post-traductionnelles majeures des histones sont mise en place: *la méthylation de H3K4, l'acétylation et la déméthylation de H3K9*. En l'absence de XPC, nous avons démontré qu'au niveau du promoteur et du terminateur de gènes activés, un état chromatinien non favorable à la transcription (hétérochromatine) est mis en place avec une forte méthylation de H3K9. Lorsque XPA n'est pas exprimé (cellules shXPA), on

observe le recrutement de XPC et un état d'euchromatine, retrouvé uniquement au promoteur. Cela laisse à penser, que l'action de XPC est principalement localisée au promoteur des gènes activés.

Nous avons enfin cherché à savoir si XPC joue un rôle sur les enzymes responsables de ces modifications. Nous avons montré qu'il est nécessaire au recrutement de 3 enzymes : l'acétyltransférase GCN5, la déméthylase KDM5b et la méthyltransférase Suv39H1. En effet, en l'absence de XPC, aucune de ces trois enzymes n'est recrutée sur le promoteur. Dans les cellules shXPA, en parallèle du recrutement de XPC, le recrutement de GCN5, KDM5b et Suv39H1 est restauré. De plus, l'absence de ces enzymes est à mettre en relation avec le contexte d'hétérochromatine observé, en l'absence de XPC. Nous avons donc démontré que XPC intervient dans la mise en place d'un contexte d'euchromatine, en agissant sur le recrutement de différentes enzymes. Aux vues des résultats sur XPC, nous pourrions à présent étudier les différents interactants de XPC lors de la transcription. Nous avons montré que XPC semble être spécifique du promoteur, c'est pourquoi il serait intéressant de voir quels peuvent être les facteurs impliqués dans son recrutement lors de la transcription. Il faudrait également approfondir l'étude du rôle de XPC sur les modifications post-traductionnelles des histones et étudier les interactions possibles entre XPC et les différentes enzymes.

Nous avons mis en évidence que les facteurs NER interviennent dans la transcription des gènes activés. Cette famille de gènes n'est pourtant qu'une petite partie des gènes du génome. Il semblerait donc intéressant d'élargir les recherches, en analysant l'implication des facteurs NER dans d'autres familles de gènes. Pour cela, il faudrait utiliser d'autres modèles que les gènes aux récepteurs nucléaires. Compte tenu des résultats de l'équipe du professeur Tjian, il serait intéressant d'approfondir le rôle des facteurs NER sur les mécanismes de différenciation cellulaire et de maintien de la pluripotence. Pour ce faire, l'utilisation de cellules ES (cellules souches embryonnaires) ou IPS (cellules souches pluripotentes induites), dérivées de fibroblastes de souris ou de patients XP, CS ou TTD, pourrait être un modèle concluant.

## REFERENCES

## A

- Aboussekhra, A., Biggerstaff, M., Shivji, M.K., Vilpo, J.A., Moncollin, V., Podust, V.N., Protic, M., Hubscher, U., Egly, J.M., and Wood, R.D. (1995). Mammalian DNA nucleotide excision repair reconstituted with purified protein components. *Cell* **80**, 859-868.
- Adamczewski, J.P., Rossignol, M., Tassan, J.P., Nigg, E.A., Moncollin, V., and Egly, J.M. (1996). MAT1, cdk7 and cyclin H form a kinase complex which is UV light-sensitive upon association with TFIIH. *Embo J* **15**, 1877-1884.
- Ahn, S.H., Kim, M., and Buratowski, S. (2004). Phosphorylation of serine 2 within the RNA polymerase II C-terminal domain couples transcription and 3' end processing. *Mol Cell* **13**, 67-76.
- Akoulitchev, S., Chuikov, S., and Reinberg, D. (2000). TFIIH is negatively regulated by cdk8-containing mediator complexes. *Nature* **407**, 102-106.
- Akoulitchev, S., Mäkelä, T.P., Weinberg, R.A., and Reinberg, D. (1995). Requirement for TFIIH kinase activity in transcription by RNA polymerase II. *Nature* **377**, 557-560.
- Allfrey, V.G., and Mirsky, A.E. (1964). Structural Modifications of Histones and their Possible Role in the Regulation of RNA Synthesis. *Science* **144**, 559.
- Allis, C.D., Berger, S.L., Cote, J., Dent, S., Jenuwien, T., Kouzarides, T., Pillus, L., Reinberg, D., Shi, Y., Shiekhattar, R., *et al.* (2007). New nomenclature for chromatin-modifying enzymes. *Cell* **131**, 633-636.
- Alsayed, Y., Uddin, S., Mahmud, N., Lekmine, F., Kalvakolanu, D.V., Minucci, S., Bokoch, G., and Plataniias, L.C. (2001). Activation of Rac1 and the p38 mitogen-activated protein kinase pathway in response to all-trans-retinoic acid. *J Biol Chem* **276**, 4012-4019.
- Andressoo, J.O., Hoeijmakers, J.H., and de Waard, H. (2005). Nucleotide excision repair and its connection with cancer and ageing. *Adv Exp Med Biol* **570**, 45-83.
- Angelov, D., Verdel, A., An, W., Bondarenko, V., Hans, F., Doyen, C.M., Studitsky, V.M., Hamiche, A., Roeder, R.G., Bouvet, P., *et al.* (2004). SWI/SNF remodeling and p300-dependent transcription of histone variant H2ABbd nucleosomal arrays. *Embo J* **23**, 3815-3824.
- Annunziato, A.T. (2012). Assembling chromatin: The long and winding road. *Biochim Biophys Acta* **1819**, 196-210.
- Ansari, A., and Hampsey, M. (2005). A role for the CPF 3'-end processing machinery in RNAP II-dependent gene looping. *Genes Dev* **19**, 2969-2978.
- Araujo, S.J., Tirode, F., Coin, F., Pospiech, H., Syvaoja, J.E., Stucki, M., Hubscher, U., Egly, J.M., and Wood, R.D. (2000). Nucleotide excision repair of DNA with recombinant human proteins: definition of the minimal set of factors, active forms of TFIIH, and modulation by CAK. *Genes Dev* **14**, 349-359.
- Archambault, J., Pan, G., Dahmus, G.K., Cartier, M., Marshall, N., Zhang, S., Dahmus, M.E., and Greenblatt, J. (1998). FCP1, the RAP74-interacting subunit of a human protein phosphatase that dephosphorylates the carboxyl-terminal domain of RNA polymerase II. *J Biol Chem* **273**, 27593-27601.
- Arents, G., and Moudrianakis, E.N. (1995). The histone fold: a ubiquitous architectural motif utilized in DNA compaction and protein dimerization. *Proc Natl Acad Sci U S A* **92**, 11170-11174.
- Armache, K.J., Kettenberger, H., and Cramer, P. (2003). Architecture of initiation-competent 12-subunit RNA polymerase II. *Proc Natl Acad Sci U S A* **100**, 6964-6968.
- Aso, T., Conaway, J.W., and Conaway, R.C. (1994). Role of core promoter structure in assembly of the RNA polymerase II preinitiation complex. *J Biol Chem* **269**, 26575-26583.

## B

- Baker, S.P., and Grant, P.A. (2007). The SAGA continues: expanding the cellular role of a transcriptional co-activator complex. *Oncogene* **26**, 5329-5340.
- Bannister, A.J., and Kouzarides, T. (1996). The CBP co-activator is a histone acetyltransferase. *Nature* **384**, 641-643.

Bannister, A.J., and Kouzarides, T. (2011). Regulation of chromatin by histone modifications. *Cell Res* 21, 381-395.

Bannister, A.J., Zegerman, P., Partridge, J.F., Miska, E.A., Thomas, J.O., Allshire, R.C., and Kouzarides, T. (2001). Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. *Nature* 410, 120-124.

Barreto, G., Schafer, A., Marhold, J., Stach, D., Swaminathan, S.K., Handa, V., Doderlein, G., Maltry, N., Wu, W., Lyko, F., *et al.* (2007). Gadd45a promotes epigenetic gene activation by repair-mediated DNA demethylation. *Nature* 445, 671-675.

Bastien, J., and Rochette-Egly, C. (2004). Nuclear retinoid receptors and the transcription of retinoid-target genes. *Gene* 328, 1-16.

Bedford, M.T., and Clarke, S.G. (2009). Protein arginine methylation in mammals: who, what, and why. *Mol Cell* 33, 1-13.

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

Bennett, R.A., Wilson, D.M., 3rd, Wong, D., and Demple, B. (1997). Interaction of human apurinic endonuclease and DNA polymerase beta in the base excision repair pathway. *Proc Natl Acad Sci U S A* 94, 7166-7169.

Bernardes de Jesus, B.M., BJORAS, M., COIN, F., and Egly, J.M. (2008). Dissection of the Molecular Defects Caused by Pathogenic Mutations in the DNA Repair Factor XPC. *Mol Cell Biol* 28, 7225-7235.

Bird, A. (2002). DNA methylation patterns and epigenetic memory. *Genes Dev* 16, 6-21.

Birse, C.E., Minvielle-Sebastia, L., Lee, B.A., Keller, W., and Proudfoot, N.J. (1998). Coupling termination of transcription to messenger RNA maturation in yeast. *Science* 280, 298-301.

Bohr, V.A., Smith, C.A., Okumoto, D.S., and Hanawalt, P.C. (1985). DNA repair in an active gene: removal of pyrimidine dimers from the DHFR gene of CHO cells is much more efficient than in the genome overall. *Cell* 40, 359-369.

Bolderson, E., Tomimatsu, N., Richard, D.J., Boucher, D., Kumar, R., Pandita, T.K., Burma, S., and Khanna, K.K. (2010). Phosphorylation of Exo1 modulates homologous recombination repair of DNA double-strand breaks. *Nucleic Acids Res* 38, 1821-1831.

Botta, E., Offman, J., Nardo, T., Ricotti, R., Zambruno, G., Sansone, D., Balestri, P., Raams, A., Kleijer, W.J., Jaspers, N.G., *et al.* (2007). Mutations in the C7orf11 (TTDN1) gene in six nonphotosensitive trichothiodystrophy patients: no obvious genotype-phenotype relationships. *Hum Mutat* 28, 92-96.

Bour, G., Gaillard, E., Bruck, N., Lalevee, S., Plassat, J.L., Busso, D., Samama, J.P., and Rochette-Egly, C. (2005). Cyclin H binding to the RARalpha activation function (AF)-2 domain directs phosphorylation of the AF-1 domain by cyclin-dependent kinase 7. *Proc Natl Acad Sci U S A* 102, 16608-16613.

Bradford, P.T., Goldstein, A.M., Tamura, D., Khan, S.G., Ueda, T., Boyle, J., Oh, K.S., Imoto, K., Inui, H., Moriwaki, S., *et al.* (2011). Cancer and neurologic degeneration in xeroderma pigmentosum: long term follow-up characterises the role of DNA repair. *J Med Genet* 48, 168-176.

Brandeis, M., Frank, D., Keshet, I., Siegfried, Z., Mendelsohn, M., Nemes, A., Temper, V., Razin, A., and Cedar, H. (1994). Sp1 elements protect a CpG island from de novo methylation. *Nature* 371, 435-438.

Brivanlou, A.H., and Darnell, J.E., Jr. (2002). Signal transduction and the control of gene expression. *Science* 295, 813-818.

Bruck, N., Vitoux, D., Ferry, C., Duong, V., Bauer, A., de The, H., and Rochette-Egly, C. (2009). A coordinated phosphorylation cascade initiated by p38MAPK/MSK1 directs RARalpha to target promoters. *Embo J* 28, 34-47.

Budhu, A.S., and Noy, N. (2002). Direct channeling of retinoic acid between cellular retinoic acid-binding protein II and retinoic acid receptor sensitizes mammary carcinoma cells to retinoic acid-induced growth arrest. *Mol Cell Biol* 22, 2632-2641.

Buratowski, S. (2009). Progression through the RNA polymerase II CTD cycle. *Mol Cell* 36, 541-546.

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.



Burke, T.W., and Kadonaga, J.T. (1996). *Drosophila* TFIID binds to a conserved downstream basal promoter element that is present in many TATA-box deficient promoters. *Genes Dev* *10*, 711-724.

Burke, T.W., and Kadonaga, J.T. (1997). The downstream core promoter element, DPE, is conserved from *Drosophila* to humans and is recognized by TAFII60 of *Drosophila*. *Genes Dev* *11*, 3020-3031.

Burley, S.K., and Roeder, R.G. (1996). Biochemistry and structural biology of transcription factor IID (TFIID). *Annu Rev Biochem* *65*, 769-799.

Bushnell, D.A., and Kornberg, R.D. (2003). Complete, 12-subunit RNA polymerase II at 4.1-Å resolution: implications for the initiation of transcription. *Proc Natl Acad Sci U S A* *100*, 6969-6973.

Busso, D., Keriel, A., Sandrock, B., Poterszman, A., Gileadi, O., and Egly, J.M. (2000). Distinct regions of MAT1 regulate cdk7 kinase and TFIIF transcription activities. *J Biol Chem* *275*, 22815-22823.

Butler, J.E., and Kadonaga, J.T. (2002). The RNA polymerase II core promoter: a key component in the regulation of gene expression. *Genes Dev* *16*, 2583-2592.

## C

Cadena, D.L., and Dahmus, M.E. (1987). Messenger RNA synthesis in mammalian cells is catalyzed by the phosphorylated form of RNA polymerase II. *J Biol Chem* *262*, 12468-12474.

Caldecott, K.W., McKeown, C.K., Tucker, J.D., Ljungquist, S., and Thompson, L.H. (1994). An interaction between the mammalian DNA repair protein XRCC1 and DNA ligase III. *Mol Cell Biol* *14*, 68-76.

Cao, J., and Yan, Q. (2012). Histone ubiquitination and deubiquitination in transcription, DNA damage response, and cancer. *Front Oncol* *2*, 26.

Carninci, P., Sandelin, A., Lenhard, B., Katayama, S., Shimokawa, K., Ponjavic, J., Semple, C.A., Taylor, M.S., Engstrom, P.G., Frith, M.C., *et al.* (2006). Genome-wide analysis of mammalian promoter architecture and evolution. *Nat Genet* *38*, 626-635.

Carter, D., Chakalova, L., Osborne, C.S., Dai, Y.F., and Fraser, P. (2002). Long-range chromatin regulatory interactions in vivo. *Nat Genet* *32*, 623-626.

Cedar, H., and Bergman, Y. (2012). Programming of DNA methylation patterns. *Annu Rev Biochem* *81*, 97-117.

Chadwick, B.P., and Willard, H.F. (2001). Histone H2A variants and the inactive X chromosome: identification of a second macroH2A variant. *Hum Mol Genet* *10*, 1101-1113.

Chalkley, G.E., and Verrijzer, C.P. (1999). DNA binding site selection by RNA polymerase II TAFs: a TAF(II)250-TAF(II)150 complex recognizes the initiator. *EMBO Journal* *18*, 4835-4845.

Chambon, P. (1996). A decade of molecular biology of retinoic acid receptors. *FASEB J* *10*, 940-954.

Champagne, K.S., and Kutateladze, T.G. (2009). Structural insight into histone recognition by the ING PHD fingers. *Curr Drug Targets* *10*, 432-441.

Chang, B., Chen, Y., Zhao, Y., and Bruick, R.K. (2007). JMJD6 is a histone arginine demethylase. *Science* *318*, 444-447.

Chao, D.M., Gadbois, E.L., Murray, P.J., Anderson, S.F., Sonu, M.S., Parvin, J.D., and Young, R.A. (1996). A mammalian SRB protein associated with an RNA polymerase II holoenzyme. *Nature* *380*, 82-85.

Cho, E.J., Kobar, M.S., Kim, M., Greenblatt, J., and Buratowski, S. (2001). Opposing effects of Ctk1 kinase and Fcp1 phosphatase at Ser 2 of the RNA polymerase II C-terminal domain. *Genes Dev* *15*, 3319-3329.

Chow, C.M., Georgiou, A., Szutorisz, H., Maia e Silva, A., Pombo, A., Barahona, I., Dargelos, E., Canzonetta, C., and Dillon, N. (2005). Variant histone H3.3 marks promoters of transcriptionally active genes during mammalian cell division. *EMBO Rep* *6*, 354-360.

Citterio, E., Rademakers, S., van der Horst, G.T., van Gool, A.J., Hoeijmakers, J.H., and Vermeulen, W. (1998). Biochemical and biological characterization of wild-type and ATPase-deficient Cockayne syndrome B repair protein. *J Biol Chem* *273*, 11844-11851.

Citterio, E., Van Den Boom, V., Schnitzler, G., Kanaar, R., Bonte, E., Kingston, R.E., Hoeijmakers, J.H., and Vermeulen, W. (2000). ATP-Dependent chromatin remodeling by the cockayne syndrome B DNA repair-transcription-coupling factor. *Mol Cell Biol* *20*, 7643-7653.

Clapier, C.R., and Cairns, B.R. (2009). The biology of chromatin remodeling complexes. *Annu Rev Biochem* *78*, 273-304.

Clarkson, S.G. (2003). The XPG story. *Biochimie* *85*, 1113-1121.

Clayton, A.L., and Mahadevan, L.C. (2003). MAP kinase-mediated phosphoacetylation of histone H3 and inducible gene regulation. *FEBS Lett* *546*, 51-58.

Coin, F., Bergmann, E., Tremeau-Bravard, A., and Egly, J.M. (1999). Mutations in XPB and XPD helicases found in xeroderma pigmentosum patients impair the transcription function of TFIIH. *Embo J* *18*, 1357-1366.

Coin, F., De Santis, L.P., Nardo, T., Zlobinskaya, O., Stefanini, M., and Egly, J.M. (2006). p8/TTD-A as a Repair-Specific TFIIH Subunit. *Mol Cell* *21*, 215-226.

Coin, F., and Egly, J.M. (1998). Ten years of TFIIH. *Cold Spring Harb Symp Quant Biol* *63*, 105-110.

Coin, F., Marinoni, J.C., Rodolfo, C., Fribourg, S., Pedrini, A.M., and Egly, J.M. (1998). Mutations in the XPD helicase gene result in XP and TTD phenotypes, preventing interaction between XPD and the p44 subunit of TFIIH. *Nat Genet* *20*, 184-188.

Coin, F., Oksenysh, V., and Egly, J.M. (2007). Distinct Roles for the XPB/p52 and XPD/p44 Subcomplexes of TFIIH in Damaged DNA Opening during Nucleotide Excision Repair. *Mol Cell* *26*, 245-256.

Coin, F., Oksenysh, V., Mocquet, V., Groh, S., Blattner, C., and Egly, J.M. (2008). Nucleotide excision repair driven by the dissociation of CAK from TFIIH. *Mol Cell* *31*, 9-20.

Compe, E., and Egly, J.M. (2012a). TFIIH: when transcription met DNA repair. *Nat Rev Mol Cell Biol* *13*, 343-354.

Compe, E., and Egly, J.M. (2012b). TFIIH: when transcription met DNA repair. *Nat Rev Mol Cell Biol* *13*, 343-354.

Compe, E., Malerba, M., Soler, L., Marescaux, J., Borrelli, E., and Egly, J.M. (2007). Neurological defects in trichothiodystrophy reveal a coactivator function of TFIIH. *Nat Neurosci* *10*, 1414-1422.

Conaway, R.C., and Conaway, J.W. (1989). An RNA polymerase II transcription factor has an associated DNA-dependent ATPase (dATPase) activity strongly stimulated by TATA region of promoters. *Proc Natl Acad Sci USA* *86*, 7356-7360.

Cook, P.R. (2003). Nongenic transcription, gene regulation and action at a distance. *J Cell Sci* *116*, 4483-4491.

Cooper, S.J., Trinklein, N.D., Anton, E.D., Nguyen, L., and Myers, R.M. (2006). Comprehensive analysis of transcriptional promoter structure and function in 1% of the human genome. *Genome Res* *16*, 1-10.

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

Cross, S.H., and Bird, A.P. (1995). CpG islands and genes. *Curr Opin Genet Dev* *5*, 309-314.

Cuthbert, G.L., Daujat, S., Snowden, A.W., Erdjument-Bromage, H., Hagiwara, T., Yamada, M., Schneider, R., Gregory, P.D., Tempst, P., Bannister, A.J., *et al.* (2004). Histone deimination antagonizes arginine methylation. *Cell* *118*, 545-553.

## D

Dahmus, M.E. (1995). Phosphorylation of the C-terminal domain of RNA polymerase II. *Biochim Biophys Acta* *1261*, 170-182.

Dawson, M.A., Bannister, A.J., Gottgens, B., Foster, S.D., Bartke, T., Green, A.R., and Kouzarides, T. (2009). JAK2 phosphorylates histone H3Y41 and excludes HP1alpha from chromatin. *Nature* *461*, 819-822.

de Laat, W., and Grosveld, F. (2003). Spatial organization of gene expression: the active chromatin hub. *Chromosome Res* *11*, 447-459.

- De Weerd-Kastelein, E.A., Keijzer, W., and Bootsma, D. (1972). Genetic heterogeneity of xeroderma pigmentosum demonstrated by somatic cell hybridization. *Nat New Biol* **238**, 80-83.
- Deaton, A.M., and Bird, A. (2011). CpG islands and the regulation of transcription. *Genes Dev* **25**, 1010-1022.
- Dejong, J.R., and Roeder, R.G. (1993). A single cDNA, hTFIIA/a encodes both the p35 and p19 subunits of human TFIIA. *Genes Dev* **7**, 2220-2234.
- Dekker, J., Rippe, K., Dekker, M., and Kleckner, N. (2002). Capturing chromosome conformation. *Science* **295**, 1306-1311.
- Delva, L., Bastie, J.N., Rochette-Egly, C., Kraiba, R., Balitrand, N., Despouy, G., Chambon, P., and Chomienne, C. (1999). Physical and functional interactions between cellular retinoic acid binding protein II and the retinoic acid-dependent nuclear complex. *Mol Cell Biol* **19**, 7158-7167.
- Deng, W., and Blobel, G.A. (2010). Do chromatin loops provide epigenetic gene expression states? *Curr Opin Genet Dev* **20**, 548-554.
- Deng, W., and Roberts, S.G. (2005). A core promoter element downstream of the TATA box that is recognized by TFIIIB. *Genes Dev* **19**, 2418-2423.
- DiGiovanna, J.J., and Kraemer, K.H. (2012). Shining a light on xeroderma pigmentosum. *J Invest Dermatol* **132**, 785-796.
- Dilworth, F.J., and Chambon, P. (2001). Nuclear receptors coordinate the activities of chromatin remodeling complexes and coactivators to facilitate initiation of transcription. *Oncogene* **20**, 3047-3054.
- Dollfus, H., Porto, F., Caussade, P., Speeg-Schatz, C., Sahel, J., Grosshans, E., Flament, J., and Sarasin, A. (2003). Ocular manifestations in the inherited DNA repair disorders. *Surv Ophthalmol* **48**, 107-122.
- Drissen, R., Palstra, R.J., Gillemans, N., Splinter, E., Grosveld, F., Philipsen, S., and de Laat, W. (2004). The active spatial organization of the beta-globin locus requires the transcription factor EKLF. *Genes Dev* **18**, 2485-2490.
- Dubaele, S., Proietti De Santis, L., Bienstock, R.J., Keriel, A., Stefanini, M., Van Houten, B., and Egly, J.M. (2003). Basal transcription defect discriminates between xeroderma pigmentosum and trichothiodystrophy in XPD patients. *Mol Cell* **11**, 1635-1646.
- Duong, V., and Rochette-Egly, C. (2011). The molecular physiology of nuclear retinoic acid receptors. From health to disease. *Biochim Biophys Acta* **1812**, 1023-1031.
- Dvir, A., Conaway, J.W., and Conaway, R.C. (2001). Mechanism of transcription initiation and promoter escape by RNA polymerase II. *Curr Opin Genet Dev* **11**, 209-214.
- Dyda, F., Klein, D.C., and Hickman, A.B. (2000). GCN5-related N-acetyltransferases: a structural overview. *Annu Rev Biophys Biomol Struct* **29**, 81-103.

## E

- Egly, J.M., and Coin, F. (2011). A history of TFIIH: two decades of molecular biology on a pivotal transcription/repair factor. *DNA Repair (Amst)* **10**, 714-721.
- Eissenberg, J.C., James, T.C., Foster-Hartnett, D.M., Hartnett, T., Ngan, V., and Elgin, S.C. (1990). Mutation in a heterochromatin-specific chromosomal protein is associated with suppression of position-effect variegation in *Drosophila melanogaster*. *Proc Natl Acad Sci U S A* **87**, 9923-9927.
- Evans, R., Fairley, J.A., and Roberts, S.G. (2001). Activator-mediated disruption of sequence-specific DNA contacts by the general transcription factor TFIIIB. *Genes Dev* **15**, 2945-2949.

## F

- Fan, Y., Nikitina, T., Morin-Kensicki, E.M., Zhao, J., Magnuson, T.R., Woodcock, C.L., and Skoultchi, A.I. (2003). H1 linker histones are essential for mouse development and affect nucleosome spacing in vivo. *Mol Cell Biol* **23**, 4559-4572.
- Finch, J.T., and Klug, A. (1976). Solenoidal model for superstructure in chromatin. *Proc Natl Acad Sci U S A* **73**, 1897-1901.

Fitch, M.E., Nakajima, S., Yasui, A., and Ford, J.M. (2003). In vivo recruitment of XPC to UV-induced cyclobutane pyrimidine dimers by the DDB2 gene product. *J Biol Chem* 278, 46906-46910.

Flores, O., Ha, I., and Reinberg, D. (1990). Factors involved in specific transcription by mammalian RNA polymerase II. Purification and subunit composition of transcription factor IIF. *J Biol Chem* 265, 5629-5634.

Flores, O., Lu, H., Killeen, M., Greenblatt, J., Burton, Z.F., and Reinberg, D. (1991). The small subunit of transcription factor IIF recruits RNA polymerase II into the preinitiation complex. *Proc Natl Acad Sci USA* 88, 9999-10003.

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

Fong, Y.W., Inouye, C., Yamaguchi, T., Cattoglio, C., Grubisic, I., and Tjian, R. (2011). A DNA repair complex functions as an oct4/sox2 coactivator in embryonic stem cells. *Cell* 147, 120-131.

Fousteri, M., Vermeulen, W., van Zeeland, A.A., and Mullenders, L.H. (2006). Cockayne syndrome A and B proteins differentially regulate recruitment of chromatin remodeling and repair factors to stalled RNA polymerase II in vivo. *Mol Cell* 23, 471-482.

Fregoso, M., Laine, J.P., Aguilar-Fuentes, J., Mocquet, V., Reynaud, E., Coin, F., Egly, J.M., and Zurita, M. (2007). DNA Repair and Transcriptional Deficiencies Caused by Mutations in the Drosophila p52 Subunit of TFIIH Generate Developmental Defects and Chromosome Fragility. *Mol Cell Biol* 27, 3640-3650.

Fuks, F., Hurd, P.J., Wolf, D., Nan, X., Bird, A.P., and Kouzarides, T. (2003). The methyl-CpG-binding protein MeCP2 links DNA methylation to histone methylation. *J Biol Chem* 278, 4035-4040.

## G

Gallinari, P., Di Marco, S., Jones, P., Pallaoro, M., and Steinkuhler, C. (2007). HDACs, histone deacetylation and gene transcription: from molecular biology to cancer therapeutics. *Cell Res* 17, 195-211.

Garrett, K.P., Serizawa, H., Hanley, J.P., Bradsher, J.N., Tsuboi, A., Arai, N., Yokota, T., Arai, K., Conaway, R.C., and Conaway, J.W. (1992). The carboxyl terminus of RAP30 is similar in sequence to region 4 of bacterial sigma factors and is required for function. *J Biol Chem* 267, 23942-23949.

Ge, H., and Roeder, R.G. (1994). The high mobility group protein HMG1 can reversibly inhibit class II gene transcription by interaction with TATA-binding protein. *J Biol Chem* 269, 17136-17140.

Gehring, M., Reik, W., and Henikoff, S. (2009). DNA demethylation by DNA repair. *Trends Genet* 25, 82-90.

Genschel, J., and Modrich, P. (2003). Mechanism of 5'-directed excision in human mismatch repair. *Mol Cell* 12, 1077-1086.

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

Germain, P., Staels, B., Dacquet, C., Spedding, M., and Laudet, V. (2006). Overview of nomenclature of nuclear receptors. *Pharmacol Rev* 58, 685-704.

Gevry, N., Hardy, S., Jacques, P.E., Laflamme, L., Svtelis, A., Robert, F., and Gaudreau, L. (2009). Histone H2A.Z is essential for estrogen receptor signaling. *Genes Dev* 23, 1522-1533.

Gianni, M., Bauer, A., Garattini, E., Chambon, P., and Rochette-Egly, C. (2002). Phosphorylation by p38MAPK and recruitment of SUG-1 are required for RA-induced RAR gamma degradation and transactivation. *Embo J* 21, 3760-3769.

Giglia-Mari, G., Coin, F., Ranish, J.A., Hoogstraten, D., Theil, A., Wijgers, N., Jaspers, N.G., Raams, A., Argentini, M., van der Spek, P.J., *et al.* (2004). A new, tenth subunit of TFIIH is responsible for the DNA repair syndrome trichothiodystrophy group A. *Nat Genet* 36, 714-719.

Gillen, A.E., and Harris, A. (2011). The role of CTCF in coordinating the expression of single gene loci. *Biochem Cell Biol* 89, 489-494.

Gillet, L.C., and Scharer, O.D. (2006). Molecular mechanisms of mammalian global genome nucleotide excision repair. *Chem Rev* 106, 253-276.

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

Gong, F., Fahy, D., and Smerdon, M.J. (2006). Rad4-Rad23 interaction with SWI/SNF links ATP-dependent chromatin remodeling with nucleotide excision repair. *Nat Struct Mol Biol* 13, 902-907.

Gozukara, E.M., Khan, S.G., Metin, A., Emmert, S., Busch, D.B., Shahlavi, T., Coleman, D.M., Miller, M., Chinsomboon, N., Stefanini, M., *et al.* (2001). A stop codon in xeroderma pigmentosum group C families in Turkey and Italy: molecular genetic evidence for a common ancestor. *J Invest Dermatol* 117, 197-204.

Graham, J.M., Jr., Anyane-Yeboah, K., Raams, A., Appeldoorn, E., Kleijer, W.J., Garritsen, V.H., Busch, D., Edersheim, T.G., and Jaspers, N.G. (2001). Cerebro-oculo-facio-skeletal syndrome with a nucleotide excision-repair defect and a mutated XPD gene, with prenatal diagnosis in a triplet pregnancy. *Am J Hum Genet* 69, 291-300.

Groisman, R., Kuraoka, I., Chevallier, O., Gaye, N., Magnaldo, T., Tanaka, K., Kisselev, A.F., Harel-Bellan, A., and Nakatani, Y. (2006). CSA-dependent degradation of CSB by the ubiquitin-proteasome pathway establishes a link between complementation factors of the Cockayne syndrome. *Genes Dev* 20, 1429-1434.

Groisman, R., Polanowska, J., Kuraoka, I., Sawada, J., Saijo, M., Drapkin, R., Kisselev, A.F., Tanaka, K., and Nakatani, Y. (2003). The ubiquitin ligase activity in the DDB2 and CSA complexes is differentially regulated by the COP9 signalosome in response to DNA damage. *Cell* 113, 357-367.

Grunstein, M. (1997). Histone acetylation in chromatin structure and transcription. *Nature* 389, 349-352.

Grunstein, M., Hecht, A., Fisher-Adams, G., Wan, J., Mann, R.K., Strahl-Bolsinger, S., Laroche, T., and Gasser, S. (1995). The regulation of euchromatin and heterochromatin by histones in yeast. *J Cell Sci Suppl* 19, 29-36.

Guo, J.U., Su, Y., Zhong, C., Ming, G.L., and Song, H. (2011a). Hydroxylation of 5-methylcytosine by TET1 promotes active DNA demethylation in the adult brain. *Cell* 145, 423-434.

Guo, R., Chen, J., Mitchell, D.L., and Johnson, D.G. (2011b). GCN5 and E2F1 stimulate nucleotide excision repair by promoting H3K9 acetylation at sites of damage. *Nucleic Acids Res* 39, 1390-1397.

## H

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

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

Han, M., and Grunstein, M. (1988). Nucleosome loss activates yeast downstream promoters in vivo. *Cell* 55, 1137-1145.

Hanawalt, P.C., and Spivak, G. (2008). Transcription-coupled DNA repair: two decades of progress and surprises. *Nat Rev Mol Cell Biol* 9, 958-970.

Handoko, L., Xu, H., Li, G., Ngan, C.Y., Chew, E., Schnapp, M., Lee, C.W., Ye, C., Ping, J.L., Mulawadi, F., *et al.* (2011). CTCF-mediated functional chromatin interactome in pluripotent cells. *Nat Genet* 43, 630-638.

Hassan, A.H., Prochasson, P., Neely, K.E., Galasinski, S.C., Chandy, M., Carrozza, M.J., and Workman, J.L. (2002). Function and selectivity of bromodomains in anchoring chromatin-modifying complexes to promoter nucleosomes. *Cell* 111, 369-379.

He, Y.F., Li, B.Z., Li, Z., Liu, P., Wang, Y., Tang, Q., Ding, J., Jia, Y., Chen, Z., Li, L., *et al.* (2011). Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. *Science* 333, 1303-1307.



Hebbes, T.R., Thorne, A.W., and Crane-Robinson, C. (1988). A direct link between core histone acetylation and transcriptionally active chromatin. *EMBO J* 7, 1395-1402.

Hennig, W. (1999). Heterochromatin. *Chromosoma* 108, 1-9.

Hess, M.T., Schwitter, U., Petretta, M., Giese, B., and Naegeli, H. (1997). Bipartite substrate discrimination by human nucleotide excision repair. *Proc Natl Acad Sci U S A* 94, 6664-6669.

Heyer, W.D., Li, X., Rolfsmeier, M., and Zhang, X.P. (2006). Rad54: the Swiss Army knife of homologous recombination? *Nucleic Acids Res* 34, 4115-4125.

Higa, L.A., Wu, M., Ye, T., Kobayashi, R., Sun, H., and Zhang, H. (2006). CUL4-DDB1 ubiquitin ligase interacts with multiple WD40-repeat proteins and regulates histone methylation. *Nat Cell Biol* 8, 1277-1283.

Hirose, Y., and Manley, J.L. (2000). RNA polymerase II and the integration of nuclear events. *Genes Dev* 14, 1415-1429.

Hirose, Y., and Ohkuma, Y. (2007). Phosphorylation of the C-terminal domain of RNA polymerase II plays central roles in the integrated events of eucaryotic gene expression. *J Biochem* 141, 601-608.

Hodawadekar, S.C., and Marmorstein, R. (2007). Chemistry of acetyl transfer by histone modifying enzymes: structure, mechanism and implications for effector design. *Oncogene* 26, 5528-5540.

Hoeijmakers, J.H. (2001). Genome maintenance mechanisms for preventing cancer. *Nature* 411, 366-374.

Holstege, F.C., van der Vliet, P.C., and Timmers, H.T. (1996). Opening of an RNA polymerase II promoter occurs in two distinct steps and requires the basal transcription factors IIE and IIH. *Embo J* 15, 1666-1677.

Hoogstraten, D., Bergink, S., Ng, J.M., Verbiest, V.H., Luijsterburg, M.S., Geverts, B., Raams, A., Dinant, C., Hoeijmakers, J.H., Vermeulen, W., *et al.* (2008). Versatile DNA damage detection by the global genome nucleotide excision repair protein XPC. *J Cell Sci* 121, 2850-2859.

Hsu, M.H., Palmer, C.N., Song, W., Griffin, K.J., and Johnson, E.F. (1998). A carboxyl-terminal extension of the zinc finger domain contributes to the specificity and polarity of peroxisome proliferator-activated receptor DNA binding. *J Biol Chem* 273, 27988-27997.

Hu, Q., Kwon, Y.S., Nunez, E., Cardamone, M.D., Hutt, K.R., Ohgi, K.A., Garcia-Bassets, I., Rose, D.W., Glass, C.K., Rosenfeld, M.G., *et al.* (2008). Enhancing nuclear receptor-induced transcription requires nuclear motor and LSD1-dependent gene networking in interchromatin granules. *Proc Natl Acad Sci U S A* 105, 19199-19204.

## I

Iben, S., Tschochner, H., Bier, M., Hoogstraten, D., Hozak, P., Egly, J.M., and Grummt, I. (2002). TFIIH plays an essential role in RNA polymerase I transcription. *Cell* 109, 297-306.

Illingworth, R.S., and Bird, A.P. (2009). CpG islands--'a rough guide'. *FEBS Lett* 583, 1713-1720.

Inostroza, J.A., Mermelstein, f., Ha, I., Lane, W., and Reinberg, D. (1992). Dr1, a TATA-binding protein-associated phosphoprotein and inhibitor of class II gene transcription. *Cell* 70, 477-489.

Ito, A., Kawaguchi, Y., Lai, C.H., Kovacs, J.J., Higashimoto, Y., Appella, E., and Yao, T.P. (2002). MDM2-HDAC1-mediated deacetylation of p53 is required for its degradation. *Embo J* 21, 6236-6245.

Ito, S., D'Alessio, A.C., Taranova, O.V., Hong, K., Sowers, L.C., and Zhang, Y. (2010). Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. *Nature* 466, 1129-1133.

Ito, S., Kuraoka, I., Chymkowitch, P., Compe, E., Takedachi, A., Ishigami, C., Coin, F., Egly, J.M., and Tanaka, K. (2007). XPG Stabilizes TFIIH, Allowing Transactivation of Nuclear Receptors: Implications for Cockayne Syndrome in XP-G/CS Patients. *Mol Cell* 26, 231-243.

Ito, S., Shen, L., Dai, Q., Wu, S.C., Collins, L.B., Swenberg, J.A., He, C., and Zhang, Y. (2011). Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. *Science* 333, 1300-1303.

## J

- Jacq, X., Brou, C., Lutz, Y., Davidson, I., Chambon, P., and Tora, L. (1994). Human TAFII30 is present in a distinct TFIID complex and is required for transcriptional activation by estrogen receptor. *Cell* **79**, 107-117.
- Jaspers, N.G., Raams, A., Silengo, M.C., Wijgers, N., Niedernhofer, L.J., Robinson, A.R., Giglia-Mari, G., Hoogstraten, D., Kleijer, W.J., Hoeijmakers, J.H., *et al.* (2007). First reported patient with human ERCC1 deficiency has cerebro-oculo-facio-skeletal syndrome with a mild defect in nucleotide excision repair and severe developmental failure. *Am J Hum Genet* **80**, 457-466.
- Javahery, R., Khachi, A., Lo, K., Zenzie-Gregory, B., and Smale, S.T. (1994). DNA sequence requirements for transcriptional initiator activity in mammalian cells. *Mol Cell Biol* **14**, 116-127.
- Jensen, R.B., Carreira, A., and Kowalczykowski, S.C. (2010). Purified human BRCA2 stimulates RAD51-mediated recombination. *Nature* **467**, 678-683.
- Jenuwein, T., and Allis, C.D. (2001). Translating the histone code. *Science* **293**, 1074-1080.
- Jin, Q., Yu, L.R., Wang, L., Zhang, Z., Kasper, L.H., Lee, J.E., Wang, C., Brindle, P.K., Dent, S.Y., and Ge, K. (2011). Distinct roles of GCN5/PCAF-mediated H3K9ac and CBP/p300-mediated H3K18/27ac in nuclear receptor transactivation. *Embo J* **30**, 249-262.
- Jin, S.G., Guo, C., and Pfeifer, G.P. (2008). GADD45A does not promote DNA demethylation. *PLoS Genet* **4**, e1000013.
- Jing, H., Vakoc, C.R., Ying, L., Mandat, S., Wang, H., Zheng, X., and Blobel, G.A. (2008). Exchange of GATA factors mediates transitions in looped chromatin organization at a developmentally regulated gene locus. *Mol Cell* **29**, 232-242.
- Junop, M.S., Obmolova, G., Rausch, K., Hsieh, P., and Yang, W. (2001). Composite active site of an ABC ATPase: MutS uses ATP to verify mismatch recognition and authorize DNA repair. *Mol Cell* **7**, 1-12.

## K

- Kadauke, S., and Blobel, G.A. (2009). Chromatin loops in gene regulation. *Biochim Biophys Acta* **1789**, 17-25.
- Kadonaga, J.T. (2002). The DPE, a core promoter element for transcription by RNA polymerase II. *Exp Mol Med* **34**, 259-264.
- Kadyrov, F.A., Dzantiev, L., Constantin, N., and Modrich, P. (2006). Endonucleolytic function of MutLalpha in human mismatch repair. *Cell* **126**, 297-308.
- Kamileri, I., Karakaslioti, I., and Garinis, G.A. (2012). Nucleotide excision repair: new tricks with old bricks. *Trends Genet.*
- Kang, J.J., Auble, D.T., Ranish, J.A., and Hahn, S. (1995). Analysis of the yeast transcription factor TFIIA: distinct functional regions and a polymerase II-specific role in basal and activated transcription. *Mol Cell Biol* **15**, 1234-1243.
- Kapetanaki, M.G., Guerrero-Santoro, J., Bisi, D.C., Hsieh, C.L., Rapic-Otrin, V., and Levine, A.S. (2006). The DDB1-CUL4ADDB2 ubiquitin ligase is deficient in xeroderma pigmentosum group E and targets histone H2A at UV-damaged DNA sites. *Proc Natl Acad Sci U S A* **103**, 2588-2593.
- Kastner, P., Mark, M., and Chambon, P. (1995). Nonsteroid nuclear receptors: what are genetic studies telling us about their role in real life? *Cell* **83**, 859-869.
- Kavli, B., Otterlei, M., Slupphaug, G., and Krokan, H.E. (2007). Uracil in DNA--general mutagen, but normal intermediate in acquired immunity. *DNA Repair (Amst)* **6**, 505-516.
- Keeney, S., Chang, G.J., and Linn, S. (1993). Characterization of a human DNA damage binding protein implicated in xeroderma pigmentosum E. *J Biol Chem* **268**, 21293-21300.
- Keriel, A., Sary, A., Sarasin, A., Rochette-Egly, C., and Egly, J.M. (2002). XPD mutations prevent TFIIH-dependent transactivation by nuclear receptors and phosphorylation of RARalpha. *Cell* **109**, 125-135.
- Keshet, I., Lieman-Hurwitz, J., and Cedar, H. (1986). DNA methylation affects the formation of active chromatin. *Cell* **44**, 535-543.

Khan, S.G., Levy, H.L., Legerski, R., Quackenbush, E., Reardon, J.T., Emmert, S., Sancar, A., Li, L., Schneider, T.D., Cleaver, J.E., *et al.* (1998). Xeroderma pigmentosum group C splice mutation associated with autism and hypoglycinemia. *J Invest Dermatol* *111*, 791-796.

Khazak, V., Estojak, J., Cho, H., Majors, J., Sonoda, G., Testa, J.R., and Golemis, E.A. (1998). Analysis of the interaction of the novel RNA polymerase II (pol II) subunit hsRPB4 with its partner hsRPB7 and with pol II. *Mol Cell Biol* *18*, 1935-1945.

Khorasanizadeh, S., and Rastinejad, F. (2001). Nuclear-receptor interactions on DNA-response elements. *Trends Biochem Sci* *26*, 384-390.

Kim, J.K., Patel, D., and Choi, B.S. (1995). Contrasting structural impacts induced by cis-syn cyclobutane dimer and (6-4) adduct in DNA duplex decamers: implication in mutagenesis and repair activity. *Photochem Photobiol* *62*, 44-50.

Kim, T.H., Barrera, L.O., Zheng, M., Qu, C., Singer, M.A., Richmond, T.A., Wu, Y., Green, R.D., and Ren, B. (2005). A high-resolution map of active promoters in the human genome. *Nature* *436*, 876-880.

Kim, T.K., Ebright, R.H., and Reinberg, D. (2000). Mechanism of ATP-dependent promoter melting by transcription factor IIH. *Science* *288*, 1418-1422.

Kleijer, W.J., Laugel, V., Berneburg, M., Nardo, T., Fawcett, H., Gratchev, A., Jaspers, N.G., Sarasin, A., Stefanini, M., and Lehmann, A.R. (2008). Incidence of DNA repair deficiency disorders in western Europe: Xeroderma pigmentosum, Cockayne syndrome and trichothiodystrophy. *DNA Repair (Amst)* *7*, 744-750.

Klose, R.J., and Bird, A.P. (2006). Genomic DNA methylation: the mark and its mediators. *Trends Biochem Sci* *31*, 89-97.

Klungland, A., and Bjelland, S. (2007). Oxidative damage to purines in DNA: role of mammalian Ogg1. *DNA Repair (Amst)* *6*, 481-488.

Knezetic, J.A., and Luse, D.S. (1986). The presence of nucleosomes on a DNA template prevents initiation by RNA polymerase II in vitro. *Cell* *45*, 95-104.

Kobor, M.S., Simon, L.D., Omichinski, J., Zhong, G., Archambault, J., and Greenblatt, J. (2000). A motif shared by TFIIF and TFIIB mediates their interaction with the RNA polymerase II carboxy-terminal domain phosphatase Fcp1p in *Saccharomyces cerevisiae*. *Mol Cell Biol* *20*, 7438-7449.

Koob, M., Laugel, V., Durand, M., Fothergill, H., Dalloz, C., Sauvanaud, F., Dollfus, H., Namer, I.J., and Dietemann, J.L. (2010). Neuroimaging in Cockayne syndrome. *AJNR Am J Neuroradiol* *31*, 1623-1630.

Kornberg, R.D. (1977). Structure of chromatin. *Annu Rev Biochem* *46*, 931-954.

Kornberg, R.D. (2007). The molecular basis of eukaryotic transcription. *Proc Natl Acad Sci U S A* *104*, 12955-12961.

Kostrewa, D., Zeller, M.E., Armache, K.J., Seizl, M., Leike, K., Thomm, M., and Cramer, P. (2009). RNA polymerase II-TFIIB structure and mechanism of transcription initiation. *Nature* *462*, 323-330.

Kouzarides, T. (2007). Chromatin modifications and their function. *Cell* *128*, 693-705.

Kraemer, K.H., Lee, M.M., Andrews, A.D., and Lambert, W.C. (1994). The role of sunlight and DNA repair in melanoma and nonmelanoma skin cancer. The xeroderma pigmentosum paradigm. *Arch Dermatol* *130*, 1018-1021.

Kraemer, K.H., Patronas, N.J., Schiffmann, R., Brooks, B.P., Tamura, D., and DiGiovanna, J.J. (2007). Xeroderma pigmentosum, trichothiodystrophy and Cockayne syndrome: a complex genotype-phenotype relationship. *Neuroscience* *145*, 1388-1396.

Kraemer, S.M., Ranallo, R.T., Ogg, R.C., and Stargell, L.A. (2001). TFIIA interacts with TFIID via association with TATA-binding protein and TAF40. *Mol Cell Biol* *21*, 1737-1746.

Kriaucionis, S., and Heintz, N. (2009). The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. *Science* *324*, 929-930.

Krivega, I., and Dean, A. (2012). Enhancer and promoter interactions-long distance calls. *Curr Opin Genet Dev* *22*, 79-85.

Kutach, A.K., and Kadonaga, J.T. (2000). The downstream promoter element DPE appears to be as widely used as the TATA box in *Drosophila* core promoters. *Mol Cell Biol* *20*, 4754-4764.



## L

- Lagrange, T., Kapanidis, A.N., Tang, H., Reinberg, D., and Ebricht, R. (1998). New core promoter element in RNA polymerase-II dependent transcription: sequence-specific DNA binding by transcription factor IIB. *Genes Dev* *12*, 34-44.
- Lahue, R.S., Au, K.G., and Modrich, P. (1989). DNA mismatch correction in a defined system. *Science* *245*, 160-164.
- Laine, J.P., and Egly, J.M. (2006). Initiation of DNA repair mediated by a stalled RNA polymerase IIO. *Embo J*.
- Laine, J.P., Singh, B.N., Krishnamurthy, S., and Hampsey, M. (2009). A physiological role for gene loops in yeast. *Genes Dev* *23*, 2604-2609.
- Laugel, V., Dalloz, C., Durand, M., Sauvanaud, F., Kristensen, U., Vincent, M.C., Pasquier, L., Odent, S., Cormier-Daire, V., Gener, B., *et al.* (2010). Mutation update for the CSB/ERCC6 and CSA/ERCC8 genes involved in Cockayne syndrome. *Hum Mutat* *31*, 113-126.
- Laugel, V., Dalloz, C., Stary, A., Cormier-Daire, V., Desguerre, I., Renouil, M., Fourmaintraux, A., Velez-Cruz, R., Egly, J.M., Sarasin, A., *et al.* (2008). Deletion of 5' sequences of the CSB gene provides insight into the pathophysiology of Cockayne syndrome. *Eur J Hum Genet* *16*, 320-327.
- le Maire, A., Teyssier, C., Erb, C., Grimaldi, M., Alvarez, S., de Lera, A.R., Balaguer, P., Gronemeyer, H., Royer, C.A., Germain, P., *et al.* (2010). A unique secondary-structure switch controls constitutive gene repression by retinoic acid receptor. *Nat Struct Mol Biol* *17*, 801-807.
- Le May, N., Egly, J.M., and Coin, F. (2010a). True lies: the double life of the nucleotide excision repair factors in transcription and DNA repair. *J Nucleic Acids* *2010*.
- Le May, N., Mota-Fernandes, D., Velez-Cruz, R., Iltis, I., Biard, D., and Egly, J.M. (2010b). 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.
- Lee, D.H., Gershenzon, N., Gupta, M., Ioshikhes, I.P., Reinberg, D., and Lewis, B.A. (2005). Functional characterization of core promoter elements: the downstream core element is recognized by TAF1. *Mol Cell Biol* *25*, 9674-9686.
- Lee, J.H., Park, C.J., Shin, J.S., Ikegami, T., Akutsu, H., and Choi, B.S. (2004). NMR structure of the DNA decamer duplex containing double T\*G mismatches of cis-syn cyclobutane pyrimidine dimer: implications for DNA damage recognition by the XPC-hHR23B complex. *Nucleic Acids Res* *32*, 2474-2481.
- Lee, J.S., Smith, E., and Shilatifard, A. (2010). The language of histone crosstalk. *Cell* *142*, 682-685.
- Lee, M.S., Kliewer, S.A., Provencal, J., Wright, P.E., and Evans, R.M. (1993). Structure of the retinoid X receptor alpha DNA binding domain: a helix required for homodimeric DNA binding. *Science* *260*, 1117-1121.
- Lee, S.B., Jasencakova, Z., and Groth, A. (2012). H3K56me1 marks a spot for PCNA. *Mol Cell* *46*, 1-2.
- Lee, S.K., Yu, S.L., Prakash, L., and 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.
- Lee, T.I., and Young, R.A. (2000). Transcription of eukaryotic protein-coding genes. *Annu Rev Genet* *34*, 77-137.
- Lehmann, A.R., Kirk-Bell, S., Arlett, C.F., Paterson, M.C., Lohman, P.H., de Weerd-Kastelein, E.A., and Bootsma, D. (1975). Xeroderma pigmentosum cells with normal levels of excision repair have a defect in DNA synthesis after UV-irradiation. *Proc Natl Acad Sci U S A* *72*, 219-223.
- Lei, L., Ren, D., and Burton, Z.F. (1999). The RAP74 subunit of human transcription factor IIF has similar roles in initiation and elongation. *Molecular & Cellular Biology* *19*, 8372-8382.
- Leid, M., Kastner, P., and Chambon, P. (1992). Multiplicity generates diversity in the retinoic acid signalling pathways. *Trends Biochem Sci* *17*, 427-433.
- Lewis, B.A., Kim, T.K., and Orkin, S.H. (2000). A downstream element in the human beta-globin promoter: evidence of extended sequence-specific transcription factor IID contacts. *Proc Natl Acad Sci U S A* *97*, 7172-7177.

Li, B., Carey, M., and Workman, J.L. (2007). The role of chromatin during transcription. *Cell* 128, 707-719.

Li, Y., Flanagan, P.M., Tschochner, H., and Kornberg, R.D. (1994). RNA polymerase II initiation factor interactions and transcription start site selection. *Science* 263, 805-807.

Lian, Z., Karpikov, A., Lian, J., Mahajan, M.C., Hartman, S., Gerstein, M., Snyder, M., and Weissman, S.M. (2008). A genomic analysis of RNA polymerase II modification and chromatin architecture related to 3' end RNA polyadenylation. *Genome Res* 18, 1224-1237.

Licht, C.L., Stevensner, T., and Bohr, V.A. (2003). Cockayne syndrome group B cellular and biochemical functions. *Am J Hum Genet* 73, 1217-1239.

Lifton, R.P., Goldberg, M.L., Karp, R.W., and Hogness, D.S. (1978). The organization of the histone genes in *Drosophila melanogaster*: functional and evolutionary implications. *Cold Spring Harb Symp Quant Biol* 42 Pt 2, 1047-1051.

Lindahl, T. (1974). An N-glycosidase from *Escherichia coli* that releases free uracil from DNA containing deaminated cytosine residues. *Proc Natl Acad Sci U S A* 71, 3649-3653.

Lindahl, T. (1993). Instability and decay of the primary structure of DNA. *Nature* 362, 709-715.

Lindenbaum, Y., Dickson, D., Rosenbaum, P., Kraemer, K., Robbins, I., and Rapin, I. (2001). Xeroderma pigmentosum/cockayne syndrome complex: first neuropathological study and review of eight other cases. *Eur J Paediatr Neurol* 5, 225-242.

Litt, M.D., Simpson, M., Gaszner, M., Allis, C.D., and Felsenfeld, G. (2001). Correlation between histone lysine methylation and developmental changes at the chicken beta-globin locus. *Science* 293, 2453-2455.

Lorch, Y., LaPointe, J.W., and Kornberg, R.D. (1987). Nucleosomes inhibit the initiation of transcription but allow chain elongation with the displacement of histones. *Cell* 49, 203-210.

Lowry, R.B., MacLean, R., McLean, D.M., and Tischler, B. (1971). Cataracts, microcephaly, kyphosis, and limited joint movement in two siblings: a new syndrome. *J Pediatr* 79, 282-284.

Loyola, A., Bonaldi, T., Roche, D., Imhof, A., and Almouzni, G. (2006). PTMs on H3 variants before chromatin assembly potentiate their final epigenetic state. *Mol Cell* 24, 309-316.

Lu, H., Zawel, L., Fisher, L., Egly, J.M., and Reinberg, D. (1992). Human general transcription factor IIH phosphorylates the C-terminal domain of RNA polymerase II. *Nature* 358, 641-645.

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

Luijsterburg, M.S., Lindh, M., Acs, K., Vrouwe, M.G., Pines, A., van Attikum, H., Mullenders, L.H., and Dantuma, N.P. (2012). DDB2 promotes chromatin decondensation at UV-induced DNA damage. *J Cell Biol* 197, 267-281.

Luse, D.S., Spangler, L.C., and Ujvari, A. (2011). Efficient and rapid nucleosome traversal by RNA polymerase II depends on a combination of transcript elongation factors. *J Biol Chem* 286, 6040-6048.

## M

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

Ma, L., Siemssen, E.D., Noteborn, H.M., and van der Eb, A.J. (1994). The xeroderma pigmentosum group B protein ERCC3 produced in the baculovirus system exhibits DNA helicase activity. *Nucleic Acids Res* 22, 4095-4102.

Mangelsdorf, D.J., and Evans, R.M. (1995). The RXR heterodimers and orphan receptors. *Cell* 83, 841-850.

Mangelsdorf, D.J., Thummel, C., Beato, M., Herrlich, P., Schutz, G., Umesono, K., Blumberg, B., Kastner, P., Mark, M., Chambon, P., *et al.* (1995). The nuclear receptor superfamily: the second decade. *Cell* 83, 835-839.

Martin, C., and Zhang, Y. (2005). The diverse functions of histone lysine methylation. *Nat Rev Mol Cell Biol* 6, 838-849.

Masutani, C., Kusumoto, R., Yamada, A., Dohmae, N., Yokoi, M., Yuasa, M., Araki, M., Iwai, S., Takio, K., and Hanaoka, F. (1999). The XPV (xeroderma pigmentosum variant) gene encodes human DNA polymerase eta [see comments]. *Nature* 399, 700-704.

Masutani, C., Sugasawa, K., Yanagisawa, J., Sonoyama, T., Ui, M., Enomoto, T., Takio, K., Tanaka, K., van der Spek, P.J., Bootsma, D., *et al.* (1994). Purification and cloning of a nucleotide excision repair complex involving the xeroderma pigmentosum group C protein and a human homologue of yeast RAD23. *Embo J* 13, 1831-1843.

Mathieu, N., Kaczmarek, N., and Naegeli, H. (2010). Strand- and site-specific DNA lesion demarcation by the xeroderma pigmentosum group D helicase. *Proc Natl Acad Sci U S A* 107, 17545-17550.

Matsumoto, Y., and Kim, K. (1995). Excision of deoxyribose phosphate residues by DNA polymerase beta during DNA repair. *Science* 269, 699-702.

Matsumura, Y., Nishigori, C., Yagi, T., Imamura, S., and Takebe, H. (1998). Characterization of molecular defects in xeroderma pigmentosum group F in relation to its clinically mild symptoms. *Hum Mol Genet* 7, 969-974.

Maurer-Stroh, S., Dickens, N.J., Hughes-Davies, L., Kouzarides, T., Eisenhaber, F., and Ponting, C.P. (2003). The Tudor domain 'Royal Family': Tudor, plant Agenet, Chromo, PWWP and MBT domains. *Trends Biochem Sci* 28, 69-74.

Maxon, M.E., Goodrich, J.A., and Tjian, R. (1994). Transcription factor IIE binds preferentially to RNA polymerase IIa and recruits TFIIH: a model for promoter clearance. *Genes Dev* 8, 515-524.

McCracken, S., Fong, N., Yankulov, K., Ballantyne, S., Pan, G., Greenblatt, J., Patterson, S.D., Wickens, M., and Bentley, D.L. (1997). The C-terminal domain of RNA polymerase II couples mRNA processing to transcription. *Nature* 365, 357-361.

McKune, K., Moore, P.A., Hull, M.W., and Woychik, N.A. (1995). Six human RNA polymerase subunits functionally substitute for their yeast counterparts. *Mol Cell Biol* 15, 6895-6900.

Mellon, I., Bohr, V.A., Smith, C.A., and Hanawalt, P.C. (1986). Preferential DNA repair of an active gene in human cells. *Proc Natl Acad Sci U S A* 83, 8878-8882.

Mellon, I., Spivak, G., and Hanawalt, P.C. (1987). Selective removal of transcription-blocking DNA damage from the transcribed strand of the mammalian DHFR gene. *Cell* 51, 241-249.

Merino, A., Madden, K.R., and Lanes, W.S. (1993). DNA topoisomerase I is involved in both repression and activation of transcription. *Nature* 365, 227-232.

Metzger, E., Wissmann, M., Yin, N., Muller, J.M., Schneider, R., Peters, A.H., Gunther, T., Buettner, R., and Schule, R. (2005). LSD1 demethylates repressive histone marks to promote androgen-receptor-dependent transcription. *Nature* 437, 436-439.

Min, J.H., and Pavletich, N.P. (2007). Recognition of DNA damage by the Rad4 nucleotide excision repair protein. *Nature* 449, 570-575.

Minucci, S., and Pelicci, P.G. (2006). Histone deacetylase inhibitors and the promise of epigenetic (and more) treatments for cancer. *Nat Rev Cancer* 6, 38-51.

Mocquet, V., Kropachev, K., Kolbanovskiy, M., Kolbanovskiy, A., Tapias, A., Cai, Y., Broyde, S., Geacintov, N.E., and Egly, J.M. (2007). The human DNA repair factor XPC-HR23B distinguishes stereoisomeric benzo[a]pyrenyl-DNA lesions. *Embo J* 26, 2923-2932.

Mocquet, V., Laine, J.P., Riedl, T., Yajin, Z., Lee, M.Y., and Egly, J.M. (2008). Sequential recruitment of the repair factors during NER: the role of XPG in initiating the resynthesis step. *Embo J* 27, 155-167.

Modrich, P., and Lahue, R. (1996). Mismatch repair in replication fidelity, genetic recombination and cancer biology. *Annu Rev Biochem* 65, 101-133.

Mohrmann, L., and Verrijzer, C.P. (2005). Composition and functional specificity of SWI2/SNF2 class chromatin remodeling complexes. *Biochim Biophys Acta* 1681, 59-73.

Moser, J., Kool, H., Giakzidis, I., Caldecott, K., Mullenders, L.H., and Foustieri, M.I. (2007). Sealing of chromosomal DNA nicks during nucleotide excision repair requires XRCC1 and DNA ligase III alpha in a cell-cycle-specific manner. *Mol Cell* 27, 311-323.

Mosley, A.L., Pattenden, S.G., Carey, M., Venkatesh, S., Gilmore, J.M., Florens, L., Workman, J.L., and Washburn, M.P. (2009). Rtr1 is a CTD phosphatase that regulates RNA polymerase II during the transition from serine 5 to serine 2 phosphorylation. *Mol Cell* 34, 168-178.

Murrell, A., Heeson, S., and Reik, W. (2004). Interaction between differentially methylated regions partitions the imprinted genes *Igf2* and *H19* into parent-specific chromatin loops. *Nat Genet* 36, 889-893.

## N

Nagpal, S., Friant, S., Nakshatri, H., and Chambon, P. (1993). RARs and RXRs: evidence for two autonomous transactivation functions (AF-1 and AF-2) and heterodimerization in vivo. *Embo J* 12, 2349-2360.

Nakabayashi, K., Amann, D., Ren, Y., Saarialho-Kere, U., Avidan, N., Gentles, S., MacDonald, J.R., Puffenberger, E.G., Christiano, A.M., Martinez-Mir, A., *et al.* (2005). Identification of *C7orf11* (TTDN1) gene mutations and genetic heterogeneity in nonphotosensitive trichothiodystrophy. *Am J Hum Genet* 76, 510-516.

Nakada, S., Chen, G.I., Gingras, A.C., and Durocher, D. (2008). PP4 is a gamma H2AX phosphatase required for recovery from the DNA damage checkpoint. *EMBO Rep* 9, 1019-1026.

Nan, X., Ng, H.H., Johnson, C.A., Laherty, C.D., Turner, B.M., Eisenman, R.N., and Bird, A. (1998). Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. *Nature* 393, 386-389.

Nance, M.A., and Berry, S.A. (1992). Cockayne syndrome: review of 140 cases. *Am J Med Genet* 84, 42-68.

Nashun, B., Yukawa, M., Liu, H., Akiyama, T., and Aoki, F. (2010). Changes in the nuclear deposition of histone H2A variants during pre-implantation development in mice. *Development* 137, 3785-3794.

Neeley, W.L., and Essigmann, J.M. (2006). Mechanisms of formation, genotoxicity, and mutation of guanine oxidation products. *Chem Res Toxicol* 19, 491-505.

Newman, J.C., Bailey, A.D., and Weiner, A.M. (2006). Cockayne syndrome group B protein (CSB) plays a general role in chromatin maintenance and remodeling. *Proc Natl Acad Sci U S A* 103, 9613-9618.

Ng, J.M., Vermeulen, W., van der Horst, G.T., Bergink, S., Sugawara, K., Vrieling, H., and Hoeijmakers, J.H. (2003). A novel regulation mechanism of DNA repair by damage-induced and RAD23-dependent stabilization of xeroderma pigmentosum group C protein. *Genes Dev* 17, 1630-1645.

Ng, S.S., Yue, W.W., Oppermann, U., and Klose, R.J. (2009). Dynamic protein methylation in chromatin biology. *Cell Mol Life Sci* 66, 407-422.

Nikolov, D.B., Chen, H., Halay, E.D., Hoffmann, A., Roeder, R.G., and Burley, S.K. (1996). Crystal structure of a human TATA box-binding protein/TATA element complex. *Proc Natl Acad Sci USA* 93, 4862-4867.

Nikolov, D.B., Hu, S.H., Lin, J., Gash, A., Hoffmann, A., Horikoshi, M., Chua, N.H., Roeder, R.G., and Burley, S.K. (1992). Crystal structure of TFIID TATA-box binding protein. *Nature* 360, 40-46.

Nishi, R., Okuda, Y., Watanabe, E., Mori, T., Iwai, S., Masutani, C., Sugawara, K., and Hanaoka, F. (2005). Centrin 2 stimulates nucleotide excision repair by interacting with xeroderma pigmentosum group C protein. *Mol Cell Biol* 25, 5664-5674.

Noma, K., Allis, C.D., and Grewal, S.I. (2001). Transitions in distinct histone H3 methylation patterns at the heterochromatin domain boundaries. *Science* 293, 1150-1155.

Nonet, M.L., and Young, R.A. (1987). Intragenic and extragenic suppressors of mutations in the heptapeptide repeat domain of *Saccharomyces cerevisiae* RNA polymerase II. *Gnetics* 123, 715-724.

Nouspikel, T., Lalle, P., Leadon, S.A., Cooper, P.K., and Clarkson, S.G. (1997). A common mutational pattern in Cockayne syndrome patients from xeroderma pigmentosum group G: implications for a second XPG function. *Proc Natl Acad Sci USA* 94, 3116-3121.

## O

- O'Brien, T., Hardin, S.E., Greenleaf, A., and Lis, J.T. (1994). Phosphorylation of RNA polymerase II C-terminal domain and transcriptional elongation. *Nature* 370, 75-77.
- O'Sullivan, J.M., Tan-Wong, S.M., Morillon, A., Lee, B., Coles, J., Mellor, J., and Proudfoot, N.J. (2004). Gene loops juxtapose promoters and terminators in yeast. *Nat Genet* 36, 1014-1018.
- Ogi, T., Limsirichaikul, S., Overmeer, R.M., Volker, M., Takenaka, K., Cloney, R., Nakazawa, Y., Niimi, A., Miki, Y., Jaspers, N.G., *et al.* (2010). Three DNA polymerases, recruited by different mechanisms, carry out NER repair synthesis in human cells. *Mol Cell* 37, 714-727.
- Ohkuma, Y., and Roeder, R.G. (1994). Regulation of TFIIH ATPase and kinase activities by TFIIE during active initiation complex formation. *Nature* 368, 160-163.
- Ohkuma, Y., Sumimoto, M., Horikoshi, M., and Roeder, R.G. (1990). Factors involved in specific transcription by mammalian RNA polymerase II: purification and characterization of general transcription factor TFIIE. *Proc Natl Acad Sci USA* 87, 9163-9167.
- Ohler, U., Liao, G.C., Niemann, H., and Rubin, G.M. (2002). Computational analysis of core promoters in the *Drosophila* genome. *Genome Biol* 3, RESEARCH0087.
- Okano, M., Bell, D.W., Haber, D.A., and Li, E. (1999). DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell* 99, 247-257.
- Oki, M., Aihara, H., and Ito, T. (2007). Role of histone phosphorylation in chromatin dynamics and its implications in diseases. *Subcell Biochem* 41, 319-336.
- Oksenych, V., de Jesus, B.B., Zhovmer, A., Egly, J.M., and Coin, F. (2009). Molecular insights into the recruitment of TFIIH to sites of DNA damage. *Embo J* 28, 2971-2980.
- Okuda, M., Tanaka, A., Arai, Y., Satoh, M., Okamura, H., Nagadoi, A., Hanaoka, F., Ohkuma, Y., and Nishimura, Y. (2004). A novel zinc finger structure in the large subunit of human general transcription factor TFIIE. *J Biol Chem* 279, 51395-51403.
- Orelli, B., McClendon, T.B., Tsodikov, O.V., Ellenberger, T., Niedernhofer, L.J., and Scharer, O.D. (2010). The XPA-binding domain of ERCC1 is required for nucleotide excision repair but not other DNA repair pathways. *J Biol Chem* 285, 3705-3712.
- Orphanides, G., Lagrange, T., and Reinberg, D. (1996). The general transcription factors of RNA polymerase II. *Genes Dev* 10, 2657-2683.
- Orphanides, G., and Reinberg, D. (2002). A unified theory of gene expression. *Cell* 108, 439-451.
- Ossipow, V., Tassan, J.P., Nigg, E.A., and Schibler, U. (1995). A mammalian RNA polymerase II holoenzyme containing all components required for promoter-specific transcription initiation. *Cell* 83, 137-146.
- Overmeer, R.M., Moser, J., Volker, M., Kool, H., Tomkinson, A.E., van Zeeland, A.A., Mullenders, L.H., and Foustieri, M. (2011). Replication protein A safeguards genome integrity by controlling NER incision events. *J Cell Biol* 192, 401-415.

## P

- Pehrson, J.R., and Fried, V.A. (1992). MacroH2A, a core histone containing a large nonhistone region. *Science* 257, 1398-1400.
- Peltomaki, P. (2003). Role of DNA mismatch repair defects in the pathogenesis of human cancer. *J Clin Oncol* 21, 1174-1179.
- Pena, S.D., and Shokeir, M.H. (1974). Autosomal recessive cerebro-oculo-facio-skeletal (COFS) syndrome. *Clin Genet* 5, 285-293.
- Peterlin, B.M., and Price, D.H. (2006). Controlling the elongation phase of transcription with P-TEFb. *Mol Cell* 23, 297-305.
- Petermann, E., Ziegler, M., and Oei, S.L. (2003). ATP-dependent selection between single nucleotide and long patch base excision repair. *DNA Repair (Amst)* 2, 1101-1114.
- Proietti-De-Santis, L., Drane, P., and Egly, J.M. (2006). Cockayne syndrome B protein regulates the transcriptional program after UV irradiation. *Embo J* 25, 1915-1923.



Proudfoot, N.J. (1989). How RNA polymerase II terminates transcription in higher eukaryotes. *Trends Biochem Sci* 14, 105-110.

Ptashne, M., and Gann, A. (1998). Imposing specificity by localization: mechanism and evolvability. *Curr Biol* 8, R812-822.

Pusarla, R.H., and Bhargava, P. (2005). Histones in functional diversification. Core histone variants. *Febs J* 272, 5149-5168.

## R

Rai, K., Huggins, I.J., James, S.R., Karpf, A.R., Jones, D.A., and Cairns, B.R. (2008). DNA demethylation in zebrafish involves the coupling of a deaminase, a glycosylase, and gadd45. *Cell* 135, 1201-1212.

Ramkumar, H.L., Brooks, B.P., Cao, X., Tamura, D., Digiovanna, J.J., Kraemer, K.H., and Chan, C.C. (2011). Ophthalmic manifestations and histopathology of xeroderma pigmentosum: two clinicopathological cases and a review of the literature. *Surv Ophthalmol* 56, 348-361.

Rangasamy, D., Greaves, I., and Tremethick, D.J. (2004). RNA interference demonstrates a novel role for H2A.Z in chromosome segregation. *Nat Struct Mol Biol* 11, 650-655.

Ranish, J.A., Hahn, S., Lu, Y., Yi, E.C., Li, X.J., Eng, J., and Aebersold, R. (2004). Identification of TFB5, a new component of general transcription and DNA repair factor IIH. *Nat Genet* 36, 707-713.

Ray-Gallet, D., Quivy, J.P., Scamps, C., Martini, E.M., Lipinski, M., and Almouzni, G. (2002). HIRA is critical for a nucleosome assembly pathway independent of DNA synthesis. *Mol Cell* 9, 1091-1100.

Rea, S., Eisenhaber, F., O'Carroll, D., Strahl, B.D., Sun, Z.W., Schmid, M., Opravil, S., Mechtler, K., Ponting, C.P., Allis, C.D., *et al.* (2000). Regulation of chromatin structure by site-specific histone H3 methyltransferases. *Nature* 406, 593-599.

Reardon, J.T., and Sancar, A. (2003). Recognition and repair of the cyclobutane thymine dimer, a major cause of skin cancers, by the human excision nuclease. *Genes Dev* 17, 2539-2551.

Reeve, J.N. (2003). Archaeal chromatin and transcription. *Mol Microbiol* 48, 587-598.

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

Renaud, J.P., and Moras, D. (2000). Structural studies on nuclear receptors. *Cell Mol Life Sci* 57, 1748-1769.

Riedl, T., Hanaoka, F., and Egly, J.M. (2003). The comings and goings of nucleotide excision repair factors on damaged DNA. *Embo J* 22, 5293-5303.

Rivera-Begeman, A., McDaniel, L.D., Schultz, R.A., and Friedberg, E.C. (2007). A novel XPC pathogenic variant detected in archival material from a patient diagnosed with Xeroderma Pigmentosum: a case report and review of the genetic variants reported in XPC. *DNA Repair (Amst)* 6, 100-114.

Rochette-Egly, C., Adam, S., Rossignol, M., Egly, J.M., and Chambon, P. (1997). Stimulation of RAR alpha activation function AF-1 through binding to the general transcription factor TFIID and phosphorylation by CDK7. *Cell* 90, 97-107.

Rochette-Egly, C., and Germain, P. (2009). Dynamic and combinatorial control of gene expression by nuclear retinoic acid receptors (RARs). *Nucl Recept Signal* 7, e005.

Rodriguez-Navarro, S. (2009). Insights into SAGA function during gene expression. *EMBO Rep* 10, 843-850.

Roeder, R.G. (1996). Nuclear RNA polymerases: role of general initiation factors and cofactors in eukaryotic transcription. *Methods Enzymol* 273, 165-171.

Roeder, R.G., and Rutter, W.J. (1970). Specific nucleolar and nucleoplasmic RNA polymerases. *Proc Natl Acad Sci U S A* 65, 675-682.

Rogakou, E.P., and Sekeri-Pataryas, K.E. (1999). Histone variants of H2A and H3 families are regulated during in vitro aging in the same manner as during differentiation. *Exp Gerontol* 34, 741-754.

Rosenfeld, M.G., Lunyak, V.V., and Glass, C.K. (2006). Sensors and signals: a coactivator/corepressor/epigenetic code for integrating signal-dependent programs of transcriptional response. *Genes Dev* 20, 1405-1428.

Rossignol, M., Keriél, A., Staub, A., and Egly, J.M. (1999). Kinase activity and phosphorylation of the largest subunit of TFIIF transcription factor. *J Biol Chem* 274, 22387-22392.

Rossignol, M., Kolb-Cheynel, I., and Egly, J.M. (1997). Substrate specificity of the cdk-activating kinase (CAK) is altered upon association with TFIIF. *Embo J* 16, 1628-1637.

Roy, R., Schaeffer, L., Humbert, S., Vermeulen, W., Weeda, G., and Egly, J.M. (1994). The DNA-dependent ATPase activity associated with the class II basic transcription factor BTF2/TFIIF. *J Biol Chem* 269, 9826-9832.

Ruppert, S., and Tjian, R. (1995). Human TAFII250 interacts with RAP74: implications for RNA polymerase II initiation. *Genes Dev* 9, 2747-2755.

## S

Saiz, L., Rubi, J.M., and Vilar, J.M. (2005). Inferring the in vivo looping properties of DNA. *Proc Natl Acad Sci U S A* 102, 17642-17645.

Samuels, M., Fire, A., and Sharp, P.A. (1982). Separation and characterization of factors mediating accurate transcription by RNA polymerase II. *J Biol Chem* 257, 14419-14427.

Sancar, A., and Hearst, J.E. (1993). Molecular matchmakers. *Science* 259, 1415-1420.

Sandrock, B., and Egly, J.M. (2001). A yeast four-hybrid system identifies Cdk-activating kinase as a regulator of the XPD helicase, a subunit of transcription factor IIF. *J Biol Chem* 276, 35328-35333.

Sawadogo, M., and Roeder, R. (1985). Interaction of a gene-specific transcription factor with the adenovirus major late promoter upstream of the TATA box region. *Cell* 43(1), 165-175.

Saxonov, S., Berg, P., and Brutlag, D.L. (2006). A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters. *Proc Natl Acad Sci U S A* 103, 1412-1417.

Schmitz, K.M., Schmitt, N., Hoffmann-Rohrer, U., Schafer, A., Grummt, I., and Mayer, C. (2009). TAF12 recruits Gadd45a and the nucleotide excision repair complex to the promoter of rRNA genes leading to active DNA demethylation. *Mol Cell* 33, 344-353.

Schultz, P., Fribourg, S., Poterszman, A., Mallouh, V., Moras, D., and Egly, J.M. (2000). Molecular structure of human TFIIF. *Cell* 102, 599-607.

Scrima, A., Konickova, R., Czyzewski, B.K., Kawasaki, Y., Jeffrey, P.D., Groisman, R., Nakatani, Y., Iwai, S., Pavletich, N.P., and Thoma, N.H. (2008). Structural basis of UV DNA-damage recognition by the DDB1-DDB2 complex. *Cell* 135, 1213-1223.

Selby, C.P., and Sancar, A. (1994). Mechanisms of transcription-repair coupling and mutation frequency decline. *Microbiol Rev* 58, 317-329.

Selby, C.P., and Sancar, A. (1997). Human transcription-repair coupling factor CSB/ERCC6 is a DNA-stimulated ATPase but is not a helicase and does not disrupt the ternary transcription complex of stalled RNA polymerase II. *J Biol Chem* 272, 1885-1890.

Serizawa, H., Conaway, J.W., and Conaway, R.C. (1993). Phosphorylation of C-terminal domain of RNA polymerase II is not required in basal transcription. *Nature* 363, 371-374.

Sharma, A., Singh, K., and Almasan, A. (2012). Histone H2AX Phosphorylation: A Marker for DNA Damage. *Methods Mol Biol* 920, 613-626.

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

Shi, Y.J., Matson, C., Lan, F., Iwase, S., Baba, T., and Shi, Y. (2005). Regulation of LSD1 histone demethylase activity by its associated factors. *Mol Cell* 19, 857-864.

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

Shykind, B.M., Kim, J., and Sharp, P.A. (1995). Activation of the TFIID-TFIIA complex with HMG-2. *Genes Dev* 9, 1354-1365.

Sims, R.J., 3rd, Chen, C.F., Santos-Rosa, H., Kouzarides, T., Patel, S.S., and Reinberg, D. (2005). Human but not yeast CHD1 binds directly and selectively to histone H3 methylated at lysine 4 via its tandem chromodomains. *J Biol Chem* 280, 41789-41792.

Singh, B.N., and Hampsey, M. (2007). A transcription-independent role for TFIIIB in gene looping. *Mol Cell* 27, 806-816.

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

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

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

Spilianakis, C.G., and Flavell, R.A. (2004). Long-range intrachromosomal interactions in the T helper type 2 cytokine locus. *Nat Immunol* 5, 1017-1027.

Splinter, E., Heath, H., Kooren, J., Palstra, R.J., Klous, P., Grosveld, F., Galjart, N., and de Laat, W. (2006). CTCF mediates long-range chromatin looping and local histone modification in the beta-globin locus. *Genes Dev* 20, 2349-2354.

Spotswood, H.T., and Turner, B.M. (2002). An increasingly complex code. *J Clin Invest* 110, 577-582.

Staresinic, L., Fagbemi, A.F., Enzlin, J.H., Gourdin, A.M., Wijgers, N., Dunand-Sauthier, I., Giglia-Mari, G., Clarkson, S.G., Vermeulen, W., and Scharer, O.D. (2009). Coordination of dual incision and repair synthesis in human nucleotide excision repair. *Embo J* 28, 1111-1120.

Stefanini, M., Botta, E., Lanzafame, M., and Orioli, D. (2010). Trichothiodystrophy: from basic mechanisms to clinical implications. *DNA Repair (Amst)* 9, 2-10.

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

Struhl, K. (1987). Promoters, activator proteins, and the mechanism of transcriptional initiation in yeast. *Cell* 49, 295-297.

Sugasawa, K. (2008). XPC: its product and biological roles. *Adv Exp Med Biol* 637, 47-56.

Sugasawa, K., Masutani, C., Uchida, A., Maekawa, T., van der Spek, P.J., Bootsma, D., Hoeijmakers, J.H., and Hanaoka, F. (1996). HHR23B, a human Rad23 homolog, stimulates XPC protein in nucleotide excision repair *in vitro*. *Mol Cell Biol* 16, 4852-4861.

Sugasawa, K., Okamoto, T., Shimizu, Y., Masutani, C., Iwai, S., and Hanaoka, F. (2001). A multistep damage recognition mechanism for global genomic nucleotide excision repair. *Genes Dev* 15, 507-521.

Sugasawa, K., Okuda, Y., Saijo, M., Nishi, R., Matsuda, N., Chu, G., Mori, T., Iwai, S., Tanaka, K., and Hanaoka, F. (2005). UV-induced ubiquitylation of XPC protein mediated by UV-DDB-ubiquitin ligase complex. *Cell* 121, 387-400.

Sugasawa, K., Shimizu, Y., Iwai, S., and Hanaoka, F. (2002). A molecular mechanism for DNA damage recognition by the xeroderma pigmentosum group C protein complex. *DNA Repair (Amst)* 1, 95-107.

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

## T

Tachiwana, H., Kagawa, W., Osakabe, A., Kawaguchi, K., Shiga, T., Hayashi-Takanaka, Y., Kimura, H., and Kurumizaka, H. (2010). Structural basis of instability of the nucleosome containing a testis-specific histone variant, human H3T. *Proc Natl Acad Sci U S A* 107, 10454-10459.

Tahiliani, M., Koh, K.P., Shen, Y., Pastor, W.A., Bandukwala, H., Brudno, Y., Agarwal, S., Iyer, L.M., Liu, D.R., Aravind, L., *et al.* (2009). Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science* 324, 930-935.

Takedachi, A., Saijo, M., and Tanaka, K. (2010). DDB2 complex-mediated ubiquitylation around DNA damage is oppositely regulated by XPC and Ku and contributes to the recruitment of XPA. *Mol Cell Biol* 30, 2708-2723.

Tamaru, H., Zhang, X., McMillen, D., Singh, P.B., Nakayama, J., Grewal, S.I., Allis, C.D., Cheng, X., and Selker, E.U. (2003). Trimethylated lysine 9 of histone H3 is a mark for DNA methylation in *Neurospora crassa*. *Nat Genet* 34, 75-79.



Tan-Wong, S.M., French, J.D., Proudfoot, N.J., and Brown, M.A. (2008). Dynamic interactions between the promoter and terminator regions of the mammalian BRCA1 gene. *Proc Natl Acad Sci U S A* *105*, 5160-5165.

Tan-Wong, S.M., Wijayatilake, H.D., and Proudfoot, N.J. (2009). Gene loops function to maintain transcriptional memory through interaction with the nuclear pore complex. *Genes Dev* *23*, 2610-2624.

Tan-Wong, S.M., Zaugg, J.B., Camblong, J., Xu, Z., Zhang, D.W., Mischo, H.E., Ansari, A.Z., Luscombe, N.M., Steinmetz, L.M., and Proudfoot, N.J. (2012). Gene Loops Enhance Transcriptional Directionality. *Science*.

Tanaka, A., Watanabe, T., Iida, Y., Hanaoka, F., and Ohkuma, Y. (2009). Central forkhead domain of human TFIIIE beta plays a primary role in binding double-stranded DNA at transcription initiation. *Genes Cells* *14*, 395-405.

Tantin, D., Kansal, A., and Carey, M. (1997). Recruitment of the putative transcription-repair coupling factor CSB/ERCC6 to RNA polymerase II elongation complexes. *Mol Cell Biol* *17*, 6803-6814.

Thambirajah, A.A., Li, A., Ishibashi, T., and Ausio, J. (2009). New developments in post-translational modifications and functions of histone H2A variants. *Biochem Cell Biol* *87*, 7-17.

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

Tiwari, V.K., McGarvey, K.M., Licchesi, J.D., Ohm, J.E., Herman, J.G., Schubeler, D., and Baylin, S.B. (2008). PcG proteins, DNA methylation, and gene repression by chromatin looping. *PLoS Biol* *6*, 2911-2927.

Tjeertes, J.V., Miller, K.M., and Jackson, S.P. (2009). Screen for DNA-damage-responsive histone modifications identifies H3K9Ac and H3K56Ac in human cells. *Embo J* *28*, 1878-1889.

Tokusumi, Y., Ma, Y., Song, X., Jacobson, R.H., and Takada, S. (2007). The new core promoter element XCPE1 (X Core Promoter Element 1) directs activator-, mediator-, and TATA-binding protein-dependent but TFIIID-independent RNA polymerase II transcription from TATA-less promoters. *Mol Cell Biol* *27*, 1844-1858.

Tolhuis, B., Palstra, R.J., Splinter, E., Grosveld, F., and de Laat, W. (2002). Looping and interaction between hypersensitive sites in the active beta-globin locus. *Mol Cell* *10*, 1453-1465.

Tomkinson, A.E., Chen, L., Dong, Z., Leppard, J.B., Levin, D.S., Mackey, Z.B., and Motycka, T.A. (2001). Completion of base excision repair by mammalian DNA ligases. *Prog Nucleic Acid Res Mol Biol* *68*, 151-164.

Tremeau-Bravard, A., Riedl, T., Egly, J.M., and Dahmus, M.E. (2004). Fate of RNA polymerase II stalled at a cisplatin lesion. *J Biol Chem* *279*, 7751-7759.

Troelstra, C., van Gool, A., de Wit, J., Vermeulen, W., Bootsma, D., and Hoeijmakers, J.H. (1992). ERCC6, a member of a subfamily of putative helicases, is involved in Cockayne's syndrome and preferential repair of active genes. *Cell* *71*, 939-953.

Trojer, P., and Reinberg, D. (2007). Facultative heterochromatin: is there a distinctive molecular signature? *Mol Cell* *28*, 1-13.

Tschiersch, B., Hofmann, A., Krauss, V., Dorn, R., Korge, G., and Reuter, G. (1994). The protein encoded by the *Drosophila* position-effect variegation suppressor gene *Su(var)3-9* combines domains of antagonistic regulators of homeotic gene complexes. *Embo J* *13*, 3822-3831.

## U

Uematsu, N., Weterings, E., Yano, K., Morotomi-Yano, K., Jakob, B., Taucher-Scholz, G., Mari, P.O., van Gent, D.C., Chen, B.P., and Chen, D.J. (2007). Autophosphorylation of DNA-PKCS regulates its dynamics at DNA double-strand breaks. *J Cell Biol* *177*, 219-229.

Ujvari, A., Pal, M., and Luse, D.S. (2011). The functions of TFIIIF during initiation and transcript elongation are differentially affected by phosphorylation by casein kinase 2. *J Biol Chem* *286*, 23160-23167.

Utley, R.T., and Cote, J. (2003). The MYST family of histone acetyltransferases. *Curr Top Microbiol Immunol* 274, 203-236.

## V

Vakoc, C.R., Letting, D.L., Gheldof, N., Sawado, T., Bender, M.A., Groudine, M., Weiss, M.J., Dekker, J., and Blobel, G.A. (2005). Proximity among distant regulatory elements at the beta-globin locus requires GATA-1 and FOG-1. *Mol Cell* 17, 453-462.

Venema, J., Mullenders, L.H., Natarajan, A.T., van Zeeland, A.A., and Mayne, L.V. (1990). The genetic defect in Cockayne syndrome is associated with a defect in repair of UV-induced DNA damage in transcriptionally active DNA. *Proc Natl Acad Sci USA* 87, 4707-4711.

Vermeulen, W., Bergmann, E., Auriol, J., Rademakers, S., Frit, P., Appeldoorn, E., Hoeijmakers, J.H., and Egly, J.M. (2000). Sublimiting concentration of TFIIH transcription/DNA repair factor causes TTD-A trichothiodystrophy disorder. *Nat Genet* 26, 307-313.

Vernimmen, D., De Gobbi, M., Sloane-Stanley, J.A., Wood, W.G., and Higgs, D.R. (2007). Long-range chromosomal interactions regulate the timing of the transition between poised and active gene expression. *Embo J* 26, 2041-2051.

Volker, M., Mone, M.J., Karmakar, P., van Hoffen, A., Schul, W., Vermeulen, W., Hoeijmakers, J.H., van Driel, R., van Zeeland, A.A., and Mullenders, L.H. (2001). Sequential assembly of the nucleotide excision repair factors in vivo. *Mol Cell* 8, 213-224.

## W

Wakasugi, M., Reardon, J.T., and Sancar, A. (1997). The non-catalytic function of XPG protein during dual incision in human nucleotide excision repair. *J Biol Chem* 272, 16030-16034.

Walker, J.R., Corpina, R.A., and Goldberg, J. (2001). Structure of the Ku heterodimer bound to DNA and its implications for double-strand break repair. *Nature* 412, 607-614.

Wang, Z., Zang, C., Cui, K., Schones, D.E., Barski, A., Peng, W., and Zhao, K. (2009). Genome-wide mapping of HATs and HDACs reveals distinct functions in active and inactive genes. *Cell* 138, 1019-1031.

Wang, Z., Zang, C., Rosenfeld, J.A., Schones, D.E., Barski, A., Cuddapah, S., Cui, K., Roh, T.Y., Peng, W., Zhang, M.Q., *et al.* (2008). Combinatorial patterns of histone acetylations and methylations in the human genome. *Nat Genet* 40, 897-903.

Weake, V.M., and Workman, J.L. (2008). Histone ubiquitination: triggering gene activity. *Mol Cell* 29, 653-663.

Wei, Y.F., Robins, P., Carter, K., Caldecott, K., Pappin, D.J., Yu, G.L., Wang, R.P., Shell, B.K., Nash, R.A., Schar, P., *et al.* (1995). Molecular cloning and expression of human cDNAs encoding a novel DNA ligase IV and DNA ligase III, an enzyme active in DNA repair and recombination. *Mol Cell Biol* 15, 3206-3216.

West, M.H., and Bonner, W.M. (1980). Histone 2A, a heteromorphous family of eight protein species. *Biochemistry* 19, 3238-3245.

Weterings, E., Verkaik, N.S., Keijzers, G., Florea, B.I., Wang, S.Y., Ortega, L.G., Uematsu, N., Chen, D.J., and van Gent, D.C. (2009). The Ku80 carboxy terminus stimulates joining and artemis-mediated processing of DNA ends. *Mol Cell Biol* 29, 1134-1142.

Whetstine, J.R., Nottke, A., Lan, F., Huarte, M., Smolikov, S., Chen, Z., Spooner, E., Li, E., Zhang, G., Colaiacovo, M., *et al.* (2006). Reversal of histone lysine trimethylation by the JMJD2 family of histone demethylases. *Cell* 125, 467-481.

Widom, J., Finch, J.T., and Thomas, J.O. (1985). Higher-order structure of long repeat chromatin. *Embo J* 4, 3189-3194.

Wierzbicki, A.T., Haag, J.R., and Pikaard, C.S. (2008). Noncoding transcription by RNA polymerase Pol IVb/Pol V mediates transcriptional silencing of overlapping and adjacent genes. *Cell* 135, 635-648.

- Wild, T., and Cramer, P. (2012). Biogenesis of multisubunit RNA polymerases. *Trends Biochem Sci* 37, 99-105.
- Williams, K., Christensen, J., and Helin, K. (2012). DNA methylation: TET proteins-guardians of CpG islands? *EMBO Rep* 13, 28-35.
- Wolf, S.S. (2009). The protein arginine methyltransferase family: an update about function, new perspectives and the physiological role in humans. *Cell Mol Life Sci* 66, 2109-2121.
- Wood, A., Shukla, A., Schneider, J., Lee, J.S., Stanton, J.D., Dzuiba, T., Swanson, S.K., Florens, L., Washburn, M.P., Wyrick, J., *et al.* (2007). Ctk complex-mediated regulation of histone methylation by COMPASS. *Mol Cell Biol* 27, 709-720.
- Woodcock, C.L., and Ghosh, R.P. (2010). Chromatin higher-order structure and dynamics. *Cold Spring Harb Perspect Biol* 2, a000596.
- Woychik, N.A., and Young, R.A. (1990). RNA polymerase II: subunit structure and function. *Trends Biochem Sci* 15, 347-351.
- Wrange, O., Okret, S., Radojic, M., Carlstedt-Duke, J., and Gustafsson, J.A. (1984). Characterization of the purified activated glucocorticoid receptor from rat liver cytosol. *J Biol Chem* 259, 4534-4541.

## X

- Xiao, B., Wilson, J.R., and Gamblin, S.J. (2003). SET domains and histone methylation. *Curr Opin Struct Biol* 13, 699-705.
- Xie, X., Kokubo, T., Cohen, S.L., Mirza, U.A., Hoffmann, A., Chait, B.T., Roeder, R.G., Nakatani, Y., and Burley, S.K. (1996). Structural similarity between TAFs and the heterotetrameric core of the histone octamer. *Nature* 380, 316-322.

## Y

- Yamamoto, Y.Y., Ichida, H., Abe, T., Suzuki, Y., Sugano, S., and Obokata, J. (2007). Differentiation of core promoter architecture between plants and mammals revealed by LDSS analysis. *Nucleic Acids Res* 35, 6219-6226.
- Yan, Q., Moreland, R.J., Conaway, J.W., and Conaway, R.C. (1999). Dual roles for transcription factor IIF in promoter escape by RNA polymerase II. *J Biol Chem* 274, 35668-35675.
- Yano, K., and Chen, D.J. (2008). Live cell imaging of XLF and XRCC4 reveals a novel view of protein assembly in the non-homologous end-joining pathway. *Cell Cycle* 7, 1321-1325.
- Yeh, J.I., Levine, A.S., Du, S., Chinte, U., Ghodke, H., Wang, H., Shi, H., Hsieh, C.L., Conway, J.F., Van Houten, B., *et al.* (2012). Damaged DNA induced UV-damaged DNA-binding protein (UV-DDB) dimerization and its roles in chromatinized DNA repair. *Proc Natl Acad Sci U S A*.
- Yonaha, M., Tsuchiya, T., and Yasukochi, Y. (1997). Cell-cycle-dependent phosphorylation of the basal transcription factor RAP74. *FEBS Lett* 410, 477-480.
- Young, R.A. (1991). RNA polymerase II. *Annu Rev Biochem* 60, 689-715.
- Yu, A., Fan, H.Y., Liao, D., Bailey, A.D., and Weiner, A.M. (2000). Activation of p53 or loss of the Cockayne syndrome group B repair protein causes metaphase fragility of human U1, U2, and 5S genes. *Mol Cell* 5, 801-810.
- Yuan, X., Feng, W., Imhof, A., Grummt, I., and Zhou, Y. (2007). Activation of RNA polymerase I transcription by cockayne syndrome group B protein and histone methyltransferase G9a. *Mol Cell* 27, 585-595.

## Z

- Zechel, C., Shen, X.Q., Chambon, P., and Gronemeyer, H. (1994a). Dimerization interfaces formed between the DNA binding domains determine the cooperative binding of RXR/RAR and RXR/TR heterodimers to DR5 and DR4 elements. *Embo J* 13, 1414-1424.

Zechel, C., Shen, X.Q., Chen, J.Y., Chen, Z.P., Chambon, P., and Gronemeyer, H. (1994b). The dimerization interfaces formed between the DNA binding domains of RXR, RAR and TR determine the binding specificity and polarity of the full-length receptors to direct repeats. *Embo J* 13, 1425-1433.

Zehring, W.A., Lee, J.M., Weeks, J.R., Jokerst, R.S., and Greenleaf, A.L. (1988). The C-terminal repeat domain of RNA polymerase II largest subunit is essential *in vivo* but is not required for accurate transcription initiation *in vitro*. *Proc Natl Acad Sci USA* 85, 3698-3702.

Zeng, L., Zhang, Q., Li, S., Plotnikov, A.N., Walsh, M.J., and Zhou, M.M. (2010). Mechanism and regulation of acetylated histone binding by the tandem PHD finger of DPF3b. *Nature* 466, 258-262.

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

Zhang, H., Roberts, D.N., and Cairns, B.R. (2005). Genome-wide dynamics of Htz1, a histone H2A variant that poises repressed/basal promoters for activation through histone loss. *Cell* 123, 219-231.

Zhang, Y., Ng, H.H., Erdjument-Bromage, H., Tempst, P., Bird, A., and Reinberg, D. (1999b). Analysis of the NuRD subunits reveals a histone deacetylase core complex and a connection with DNA methylation. *Genes Dev* 13, 1924-1935.

Zhang, Y., Tian, Y., Chen, Q., Chen, D., Zhai, Z., and Shu, H.B. (2007). TTDN1 is a Plk1-interacting protein involved in maintenance of cell cycle integrity. *Cell Mol Life Sci* 64, 632-640.

Zhen, W., Evans, M.K., Haggerty, C.M., and Bohr, V.A. (1993). Deficient gene specific repair of cisplatin-induced lesions in Xeroderma pigmentosum and Fanconi's anemia cell lines. *Carcinogenesis* 14, 919-924.

Zhou, M., Halanski, M.A., Radonovich, M.F., Kashanchi, F., Peng, J., Price, D.H., and Brady, J.N. (2000). Tat modifies the activity of CDK9 to phosphorylate serine 5 of the RNA polymerase II carboxyl-terminal domain during human immunodeficiency virus type 1 transcription. *Mol Cell Biol* 20, 5077-5086.

Zorov, D.B., Juhaszova, M., and Sollott, S.J. (2006). Mitochondrial ROS-induced ROS release: an update and review. *Biochim Biophys Acta* 1757, 509-517.

Zotter, A., Luijsterburg, M.S., Warmerdam, D.O., Ibrahim, S., Nigg, A., van Cappellen, W.A., Hoeijmakers, J.H., van Driel, R., Vermeulen, W., and Houtsmuller, A.B. (2006). Recruitment of the nucleotide excision repair endonuclease XPG to sites of UV-induced dna damage depends on functional TFIIH. *Mol Cell Biol* 26, 8868-8879.

## ANNEXES

## Publication 4

### **Poly (ADP-ribose) glycohydrolase regulates retinoic acid receptor-mediated gene expression**

Le May N., Iltis I., Amé JC., Zhovmer A., Biard D., Egly JM., Schreiber V. and Coin F.

**Molecular Cell** (Décembre 2012)

#### **Résumé**

PARG est une enzyme capable de cliver les polymères ADP-ribose (PAR) synthétisés par les protéines de la famille PARP (poly (ADP-ribose) polymérase). Tout comme les facteurs NER, PARG a été décrite dans le processus de réparation de l'ADN. Nous avons montré que l'activité enzymatique de PARG est nécessaire à la transcription de gènes activés. PARG est précocement détectée au niveau du promoteur, lors de l'activation du gène et permet la formation d'un environnement permissif indispensable pour le recrutement de la machinerie transcriptionnelle.

Nous avons identifié une des cibles de PARG, l'histone déméthylase KDM4D qui intervient directement dans l'établissement des modifications chromatinienne associées à l'expression des gènes. KDM4D existe sous une forme inactive, lorsqu'elle est PARsylés. Le recrutement précoce de PARG au niveau du promoteur va permettre le clivage du PAR de KDM4D, pouvant alors à nouveau assurer son activité de déméthylase de l'histone H3K9. Ces résultats définissent PARG comme un co-activateur régulant le remodelage de la chromatine pendant l'activation d'un gène.



# Molecular Cell

Volume 48  
Number 5

December 14, 2012

[www.cellpress.com](http://www.cellpress.com)



**Placing PARG in the Transcription Puzzle**



# Poly (ADP-Ribose) Glycohydrolase Regulates Retinoic Acid Receptor-Mediated Gene Expression

Nicolas Le May,<sup>1</sup> Izarn Ittis,<sup>1</sup> Jean-Christophe Amé,<sup>2</sup> Alexander Zhovmer,<sup>1</sup> Denis Biard,<sup>3</sup> Jean-Marc Egly,<sup>1</sup> Valérie Schreiber,<sup>2,\*</sup> and Frédéric Coin<sup>1,\*</sup>

<sup>1</sup>IGBMC, Department of Functional Genomics and Cancer, CNRS/INSERM/Université de Strasbourg, BP 163, 67404 Illkirch Cedex, C.U. Strasbourg, France

<sup>2</sup>UMR7242 Biotechnology and Cell Signaling, Université de Strasbourg, CNRS, Laboratory of Excellence MEDALIS, ESBS, Bd S. Brant, BP 10413, 67412 Illkirch, France

<sup>3</sup>CEA-DSV-iRCM, INSERM U935. Institut A. Lwoff-CNRS 7 rue Guy Moquet, 94800 Villejuif, France

\*Correspondence: [valerie.schreiber@unistra.fr](mailto:valerie.schreiber@unistra.fr) (V.S.), [fred@igbmc.fr](mailto:fred@igbmc.fr) (F.C.)

<http://dx.doi.org/10.1016/j.molcel.2012.09.021>

## SUMMARY

Poly-(ADP-ribose) glycohydrolase (PARG) is a catabolic enzyme that cleaves ADP-ribose polymers synthesized by poly-(ADP-ribose) polymerases. Here, transcriptome profiling and differentiation assay revealed a requirement of PARG for retinoic acid receptor (RAR)-mediated transcription. Mechanistically, PARG accumulates early at promoters of RAR-responsive genes upon retinoic acid treatment to promote the formation of an appropriate chromatin environment suitable for transcription. Silencing of PARG or knockout of its enzymatic activity maintains the H3K9me2 mark at the promoter of the RAR-dependent genes, leading to the absence of preinitiation complex formation. In the absence of PARG, we found that the H3K9 demethylase KDM4D/JMJD2D became PARsylated. Mutation of two glutamic acids located in the Jumonji N domain of KDM4D inhibited PARsylation. PARG becomes dispensable for ligand-dependent transcription when either a PARG inhibitor or a non-PARsylable KDM4D/JMJD2D mutant is used. Our results define PARG as a coactivator regulating chromatin remodeling during RA-dependent gene expression.

## INTRODUCTION

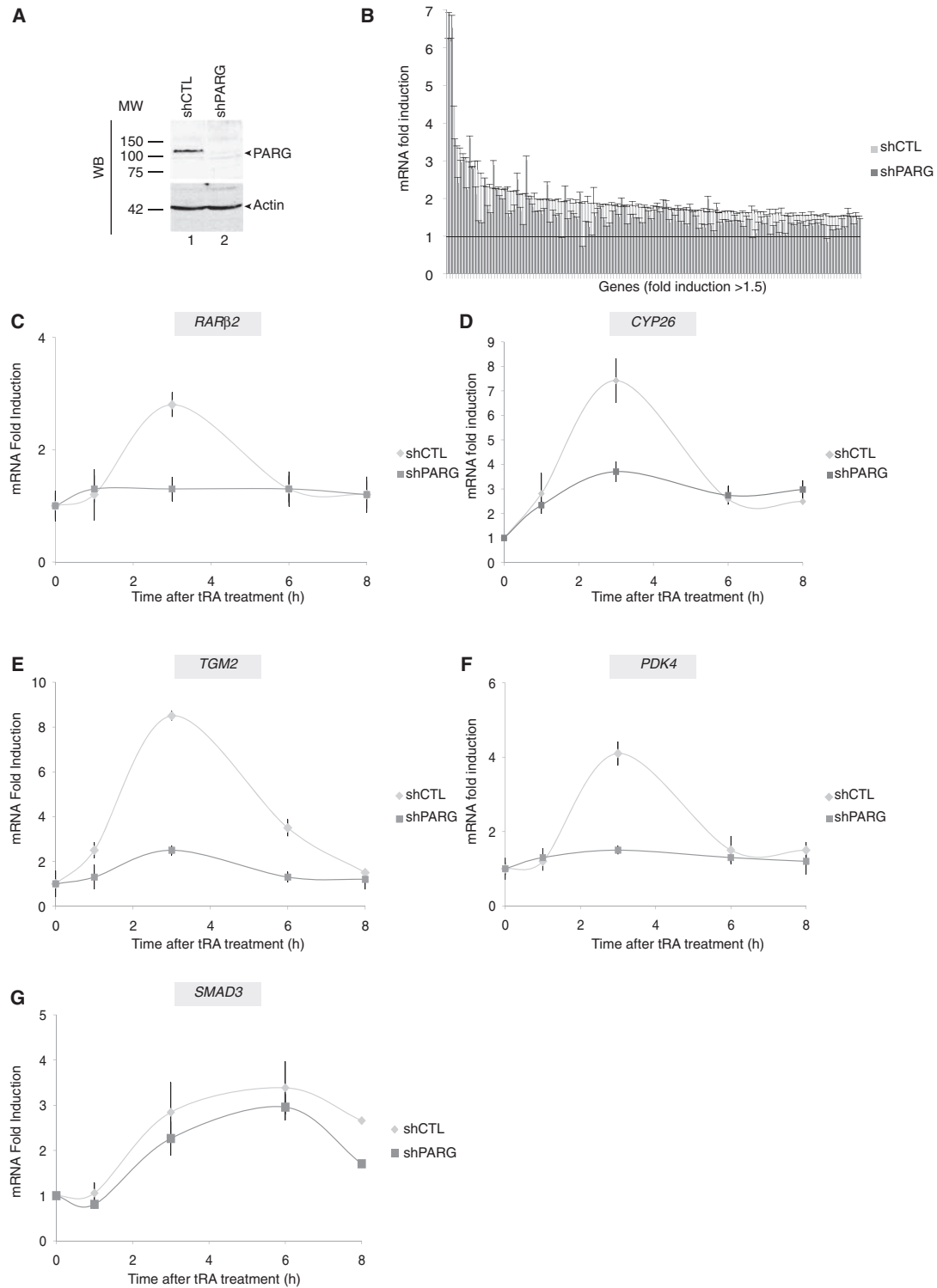
Poly-(ADP-ribosyl)ation (PARsylation) of proteins is a NAD(+)-dependent posttranslational modification catalyzed by poly-(ADP-ribose) polymerases (PARPs) (Krishnakumar and Kraus, 2010b; Rouleau et al., 2010). PARP-1 is responsible for about 90% of the total PARsylation activity in the cell (Amé et al., 2004) and is involved in various cellular processes such as single-strand/base excision repair, alternative end-joining during immunoglobulin class switch recombination and mitotic segregation (Robert et al., 2009; Schreiber et al., 2006).

Increasing evidence supports the active involvement of PARP-1 in gene expression and repression (Ji and Tulin, 2010;

Kraus, 2008). Studies of the *Parp-1*<sup>-/-</sup> mouse and PARP-1-silenced human cells have shown that an absence of PARP-1 alters the expression of many genes involved in both cell-cycle control and the stress response (Frizzell et al., 2009; Krishnakumar and Kraus, 2010a; Ogino et al., 2007; Simbulan-Rosenthal et al., 2000). In addition, it has been demonstrated that PARP-1 is also required for nuclear receptor (NR)-mediated transcription processes, including those of the retinoic acid (RAR) and estrogen (ER) receptors (Ju et al., 2006; Pavri et al., 2005). The involvement of PARP-1 enzymatic activity in transcription depends on the nature of the gene and stimuli studied, for reasons that are not yet fully understood (Hassa et al., 2003; Ju et al., 2006; Pavri et al., 2005; Simbulan-Rosenthal et al., 2003; Tulin and Spradling, 2003).

The metabolism of PAR is tightly regulated not only through the regulation of PARP activity, but also through the activity of poly(ADP-ribose) glycohydrolase (PARG), an enzyme with both exo- and endoglycosidase activities that generates large amounts of free ADP-ribose (Heeres and Hergenrother, 2007). PARG is encoded by a single gene in humans, which produces several isoforms with various subcellular localizations (Meyer et al., 2007; Meyer-Ficca et al., 2004; Niere et al., 2008). The existence of these multiple PARG isoforms and the embryonic lethality of mice deficient in all of them (Koh et al., 2004) complicate the study of this enzyme in the context of fundamental cellular processes. However, the use of human knockdown cells deficient in all PARG isoforms revealed that PARG is involved in the repair of single- and double-strand breaks and in the mitotic spindle checkpoint (Amé et al., 2009; Erdélyi et al., 2009; Fisher et al., 2007). The structure of a bacterial PARG has been recently solved and has revealed that the PARG catalytic domain is a distant member of the ubiquitous ADP-ribose binding macrodomain family (Slade et al., 2011). However, our understanding of the cellular function of PARG remains poor compared to the PARP family members.

Given the importance of PAR in NR-dependent transcription and the increasing evidence that inhibitors of PARG are potent anticancer drug candidates (Miwa and Masutani, 2007), we sought to investigate the potential function of PARG in NR-dependent gene transcription. The NR superfamily consists of a variety of DNA binding transcription factors that control the expression of genes involved in many cellular processes



**Figure 1. PARG Stimulates RAR-Dependent Gene Expression**

(A) Proteins from whole-cell extracts (50  $\mu$ g) of either shCTL or shPARG cells were resolved by SDS-PAGE followed by western blotting using anti-PARG and anti-actin antibodies. Molecular weights are indicated.

## Molecular Cell

### PARG Is a Transcriptional Coactivator

(Kastner et al., 1995). NRs are mainly classified as either steroid receptors, such as ER, or nonsteroid receptors, such as RAR. These receptors target promoters by binding specific recognition sequences and activate transcription by recruiting coactivators (Glass and Rosenfeld, 2000). The all-trans retinoic acid (t-RA), the ligand of RAR, is a pleiotropic signaling molecule derived from vitamin A that regulates critical genetic programs, thereby controlling cell proliferation and differentiation, as well as cell death or survival (Claggett-Dame and Knutson, 2011).

In the present study, we demonstrate the participation of the PAR-degrading activity of PARG in the transactivation of genes mediated by RAR. We have shown that t-RA-induced F9 embryonic cell differentiation depends on PARG. Global transcriptome analysis revealed that the induction of a significant number of RA-induced genes was stimulated by PARG. Mechanistically, we show that PARG is recruited to the promoters of RAR-dependent genes in a ligand-dependent manner to promote the formation of a permissive chromatin environment suitable for transcription. In particular, PARG targeted the demethylase KDM4D/JMJD2D to favor the removal of the repressive H3K9me2 mark around the promoters of RAR-dependent genes and the subsequent formation of the transcription preinitiation complex (PIC).

## RESULTS

### Absence of PARG Impairs RAR-Dependent Gene Transactivation

To evaluate the influence of PARG in transcription, we assayed RAR-mediated transcriptional activity of a HeLa cell line constitutively expressing shRNA directed against the catalytic domain of all PARG isoforms (shPARG) (Amé et al., 2009). The HeLa control cell line (shCTL) expresses a nonfunctional shRNA. A robust knockdown of PARG (Figure 1A) and an absence of PAR-degrading activity (see Figure S1A online) were observed in shPARG cells.

We analyzed the transcriptome profiles of shCTL and shPARG cells 3 hr after t-RA treatment (10  $\mu$ M). Treatment of shCTL cells with t-RA resulted in the induction of several immediate-early responsive genes, such as *RAR $\beta$ 2*, *CYP26*, or *TGM2* (Table S1 and Table S2). Interestingly, ~70% of genes showing an induction  $\geq 1.5$  in shCTL cells had a lower induction level in shPARG cells (Table S3 and Figure 1B). Using RT-qPCR, we confirmed that *RAR $\beta$ 2* (Figure 1C), *CYP26* (Figure 1D), *TGM2* (Figure 1E), *PDK4* (Figure 1F), *NRIP1* (Figure S1B), and *TNFAIP2* (Figure S1B) were less induced in shPARG cells over time after t-RA treatment. In contrast, another group of genes including *SMAD3* was induced similarly in shCTL and shPARG cells (Table S1 and Figure 1G). Note that no dramatic changes in expression were observed in the absence of t-RA between shCTL and shPARG cells (Table S4).

Since PAR activity has been implicated in mRNA splicing (Mangla et al., 2008), we measured the accumulation of pre-mRNA to determine if the defect in transactivation could be explained by a deficiency in splicing (Figure S1C). We observed a defect in *RAR $\beta$ 2* expression in shPARG cells whether we measured pre-mRNA or total mRNA (Figures S1D and S1E), consistent with the participation of PARG in an early step of the RAR-dependent transcription process, before mRNA splicing.

### RAR-Dependent Gene Transactivation Depends on the PAR-Degrading Activity of PARG

We next explored whether the PAR-degrading activity of PARG was required for RAR-dependent gene transcription. We used GFP-tagged shRNA-resistant constructs expressing either a wild-type (PARG<sup>WT</sup>-GFP) or a catalytically inactive version of PARG, targeting the two consecutive glutamates of the active site (PARG<sup>E755/756A</sup>-GFP) (Patel et al., 2005; Mortusewicz et al., 2011; Slade et al., 2011). Transfection of PARG<sup>WT</sup>-GFP into shPARG cells treated with H<sub>2</sub>O<sub>2</sub> resulted in the depletion of the nuclear PAR (Amé et al., 2009), which was not achieved with PARG<sup>E755/756A</sup>-GFP (compare Figures 2Aa–2Ac with Figures 2Ad–2Af). Furthermore, expression of PARG<sup>WT</sup>-GFP reduced the genotoxic-stress-independent PARylation of PARP-1 observed in shPARG cells (Amé et al., 2009) that was not achieved with GFP-PARG<sup>E755/756A</sup> (Figure 2B). These results indicate that the E755/756A mutations abolished the PAR-degrading activity of PARG.

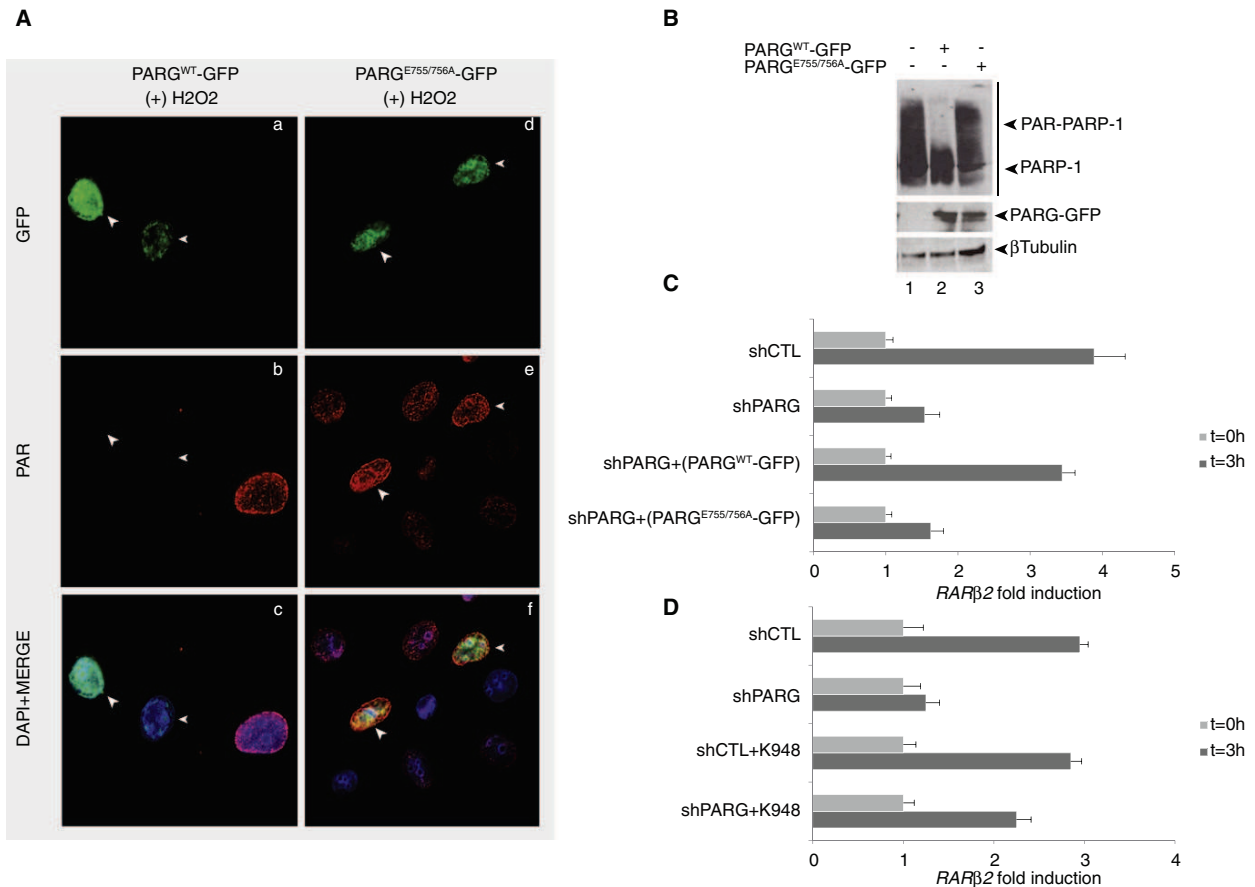
When expressed in shPARG cells, PARG<sup>WT</sup>-GFP rescued the ligand-dependent expression of *RAR $\beta$ 2* (Figure 2C), *CYP26*, *PDK4*, *NRIP1*, and *TNFAIP2* (Figure S2A). The rescue was greatly compromised following expression of PARG<sup>E755/756A</sup>-GFP (Figure 2C and Figure S2A) despite producing similar protein amounts as PARG<sup>WT</sup>-GFP (Figure 2B). These data indicate that the PAR-degrading activity of PARG stimulates RAR-dependent gene transcription and imply that PAR polymerase activity may thwart gene activation in the absence of an active PARG. Consequently, we pretreated shPARG cells with the PARP-specific inhibitor KU0058948 (K948) (McCabe et al., 2005) for 12 hr prior to t-RA addition and observed a rescue of the transactivation of *RAR $\beta$ 2* (Figure 2D), *CYP26*, *PDK4*, *NRIP1*, and *TNFAIP2* (Figure S2B). The transactivation of *RAR $\beta$ 2* was not affected by the inhibition of PARP in shCTL cells. Altogether, these results suggest that the catalytic activity of PARG is involved in NR-dependent gene transactivation as a countermeasure to PAR polymerase activity.

### Absence of PARG Inhibits RA-Induced F9 Embryonic Cell Differentiation

Treatment with t-RA induces proliferation arrest, differentiation, and apoptosis in many cells (Strickland and Mahdavi, 1978). To study the role of PARG in a physiological context, we transfected F9 mouse embryonic carcinoma cells with siRNA against

(B) Transcriptome profiling of shCTL (light gray) or shPARG (dark gray) cells treated with t-RA (10  $\mu$ M; 3 hr) using the whole transcript coverage Affymetrix Human Gene 1.0 ST arrays. Based on two independent experiments, these genes show an upregulation  $>1.5$  ( $\pm$ SEM) compared to t = 0 hr (F test,  $p < 0.005$ ).

(C–G) Relative mRNA expression of *RAR $\beta$ 2* (C), *CYP26* (D), *TGM2* (E), *PDK4* (F), and *SMAD3* (G) in either shCTL or shPARG cells measured at different time points after treatment with t-RA (10  $\mu$ M). Error bars represent the standard deviation (SD) of three independent experiments. The values are plotted relative to the expression level of the no treatment control that is set to 1 in all experiments.



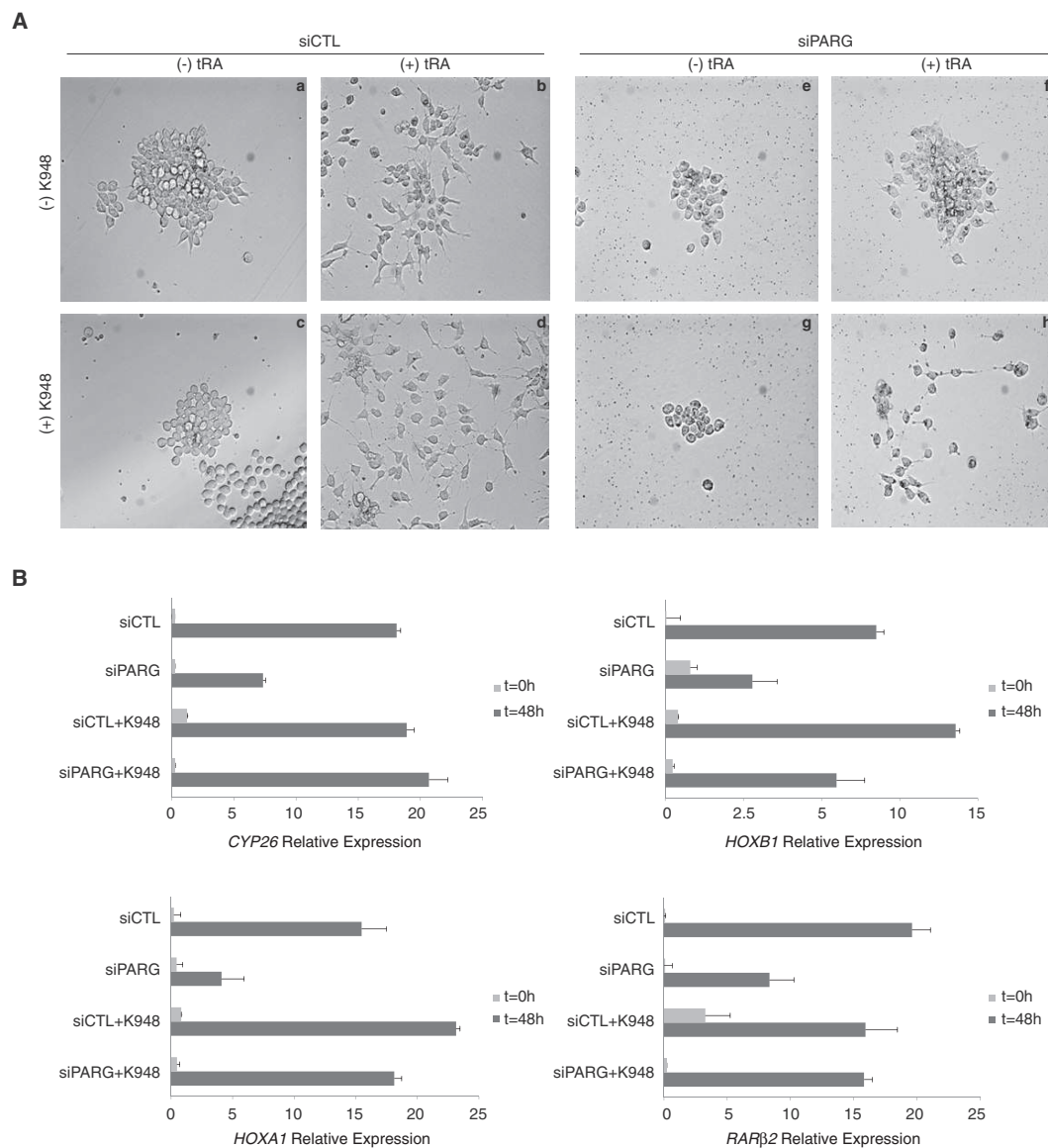
**Figure 2. The PAR-Degrading Activity of PARG Is Required for RAR-Dependent Gene Expression**

(A) Immunodetection of PAR in shPARG cells treated with 1 mM H<sub>2</sub>O<sub>2</sub>, during 10 min (to stimulate PAR synthesis). Cells were transfected 36 hr before treatment with either PARG<sup>WT</sup>-GFP (Aa–Ac) or PARG<sup>E755/756A</sup>-GFP (Ad–Af). Transfected cells (indicated with an arrow) were detected with GFP. PAR was detected with a mouse monoclonal anti-PAR antibody (10H). DNA was counterstained with DAPI. Immunofluorescence was performed as described (Amé et al., 2009). (B) Proteins from whole-cell extracts (50 μg) of shPARG cells expressing either PARG<sup>WT</sup>-GFP or PARG<sup>E755/756A</sup>-GFP were resolved by SDS-PAGE followed by western blotting using anti-GFP, anti-PARP-1, and anti-β-tubulin antibodies. (C) Relative mRNA expression (±SD, three independent experiments) of *RARβ2* measured at the transactivation peak, 3 hr post-t-RA treatment, in the indicated cell lines. The values are expressed relative to the expression level of the no treatment control that is set to 1 in all experiments. (D) Relative mRNA expression (±SD, three independent experiments) of *RARβ2* measured 3 hr after t-RA treatment in shCTL and shPARG cells incubated with K948 (100 nM) for 12 hr before addition of t-RA. The values are plotted relative to the expression level of the no treatment control that is set to 1 in all experiments.

PARG (siPARG versus siCTL; Figure S3) and followed the process of differentiation after t-RA treatment based on the morphological features characteristic of pre-endoderm-like cells. Seventy-two hours after t-RA treatment, siCTL-transfected F9 cells became more adhesive and grew separately (compare Figures 3Aa and 3Ab). At the same time point, the morphology of the siPARG-transfected F9 cells treated with t-RA remained unchanged, revealing a weak adhesive capacity and the formation of clusters (compare Figures 3Ae and 3Af). In the presence of the PARP inhibitor K948, the morphology and individual growth capacities of the siPARG-transfected F9 cells were similar to those of the siCTL-transfected cells treated with t-RA alone (compare Figures 3Ac and 3Ad with Figures 3Ag and 3Ah). The treatment of siCTL-transfected cells with K948 only slightly inter-

fered with the differentiation process (compare Figures 3Aa and 3Ab with Figures 3Ac and 3Ad), in agreement with previous data (Quénet et al., 2008).

In parallel to the morphological study, we followed the relative expression of genes associated with t-RA-induced differentiation of F9 cells in pre-endoderm-like cells (Eifert et al., 2006; Lalevé et al., 2011). *CYP26*, *HOXA1*, *HOXB1*, and *RARβ2* showed increased expression in siCTL-transfected F9 cells 48 hr after t-RA treatment (Figure 3B). The expression of all of these genes was reduced in siPARG-transfected F9 cells treated under the same conditions (Figure 3B), while the maintenance of K948 treatment during differentiation restored their expression (Figure 3B). Altogether, these studies suggest a physiological role for PARG in RA-induced F9 embryonic cell differentiation.



**Figure 3. PARG Is Required for RA-Induced Differentiation of Pluripotent F9 Cells**

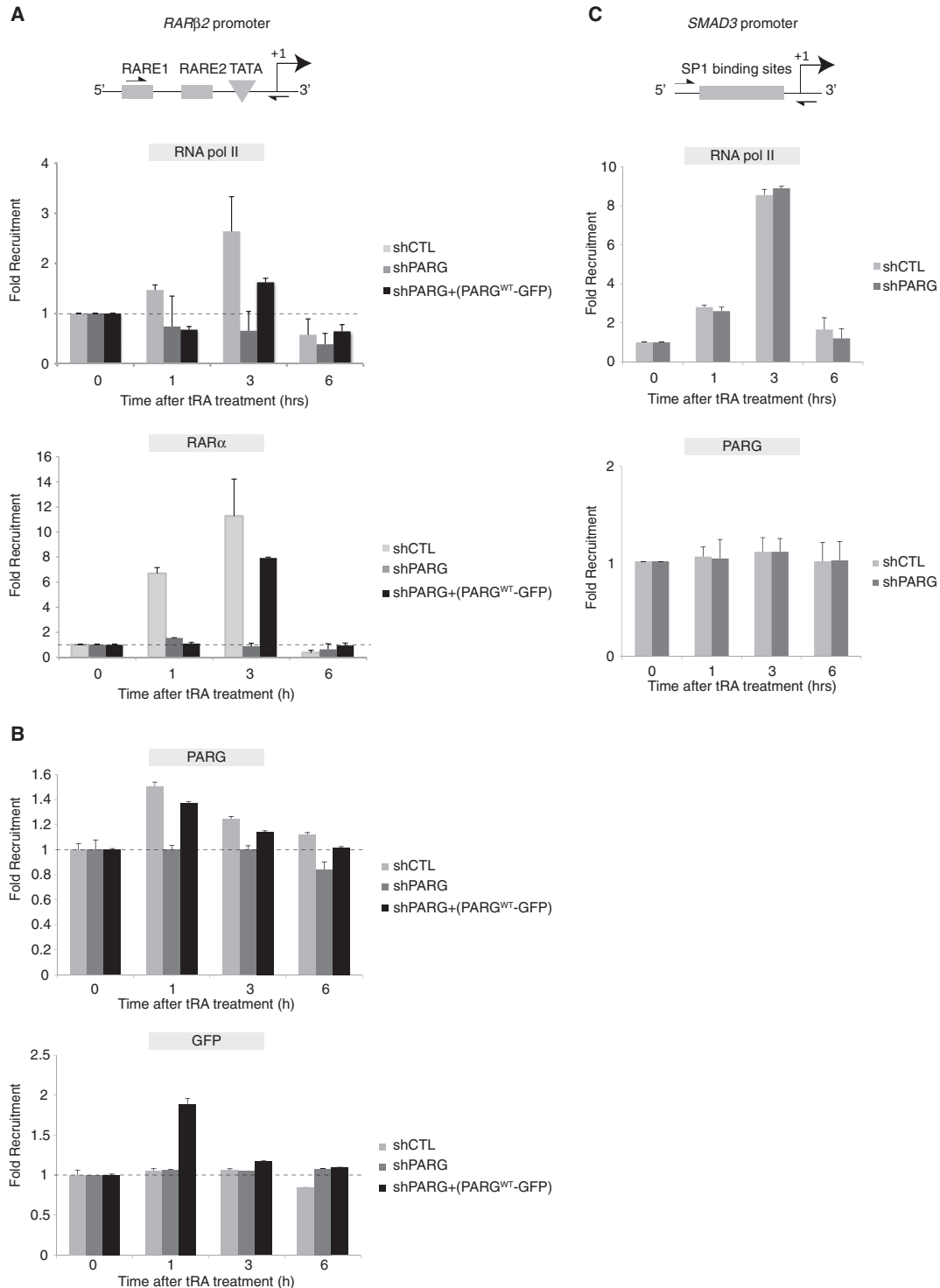
(A) Effects of PARG silencing on the differentiation of F9 embryonic carcinoma cells. Thirty-six hours after siRNA transfection, cells were treated with t-RA (1  $\mu$ M) in the presence or absence of the PARP inhibitor K948 (100 nM). The morphology of the F9 cells was analyzed 72 hr later by phase-contrast microscopy.

(B) mRNA level ( $\pm$ SD, three independent experiments) of upregulated RAR target genes in differentiated F9 cells, measured 48 hr after t-RA treatment. F9 cells were incubated with the PARP-1 inhibitor K948 during differentiation when indicated. Values are expressed as the percentage of mRNA level compared with *GAPDH*.

### Absence of PARG Leads to Defective Preinitiation Complex Formation

To address the precise role of PARG in RAR-dependent transcription, we used chromatin immunoprecipitation (ChIP) assay and monitored the recruitment of transcription factors to several *RAR*-dependent promoters in shCTL versus shPARG HeLa cells. In shCTL cells, we observed the ligand-dependent recruitment of RNA Pol II and *RAR $\alpha$  to the *RAR $\beta$ 2 (Figure 4A and Figure S4A)**

and *CYP26* promoters (Figure S5A), which was optimal at the mRNA transactivation peak, 3 hr after t-RA treatment. In contrast, PIC formation was deficient in PARG-depleted cells throughout the time course (Figure 4A and Figures S4B and S5B) and partially restored following the expression of PARG<sup>WT</sup>-GFP (Figure 4A and Figure S4G). The recruitment of PARP-1 parallels that of the PIC in shCTL cells (Figures S4A and S5A). In shPARG cells, the recruitment of PARP-1 was still



**Figure 4. Defect in PIC Assembly at the *RARβ2* Promoter in the Absence of PARG**

(A) (Upper panel) Diagram of the endogenous human *RARβ2* promoter. Positions of the primers for amplification of promoter region in ChIP are indicated with arrows. (Lower panels) ChIP monitoring the t-RA-dependent occupancy ( $\pm$ SEM) of RNA Pol II and RAR $\alpha$  on the *RARβ2* promoter in chromatin extracts from either shCTL (light gray), shPARG (dark gray), or shPARG cells transfected with shRNA-resistant PARG<sup>WT</sup>-GFP construct (black).



detected, although to a lesser extent than in shCTL cells (Figures S4B and S5B).

We also observed the recruitment of PARG to the *RAR*-dependent promoters in a ligand-dependent manner either in shCTL cells or in shPARG cells transiently expressing shRNA resistant PARG<sup>WT</sup>-GFP (Figure 4B and Figures S4A, S4G, and S5A). In contrast to PIC, the optimal accumulation of both endogenous and transiently expressed PARG occurred before the transactivation peak, 1 hr after t-RA treatment. The formation of the PIC on the *SMAD3* promoter was not altered in shPARG cells, and PARG was not recruited to this promoter in shCTL cells (Figure 4C and Figures S6A and S6B).

### PARG Controls Chromatin Derepression at *RAR*-Dependent Gene Promoters

We next assayed chromatin remodeling around the *RAR*-dependent promoters. We found a parallel between the recruitment of the PIC and the increase in both dimethylation of histone H3 lysine 4 (H3K4me2) and acetylation of histone H3 lysine 9 (H3K9ac), two marks of active transcription, on the promoters of *RARβ2* (Figures 5A and 5B and Figure S4C), *CYP26* (Figure S5C), and *SMAD3* (Figure S6C) in shCTL cells. A decrease in the dimethylation of histone H3 lysine 9 (H3K9me2), a mark of repressive transcription, took place before PIC formation, 1 hr after t-RA treatment, on both *RARβ2* (Figure 5C, Figure S4E) and *CYP26* promoters (Figure S5E). In addition, the exclusion of the repressive histone H1 from these promoters was observed soon after t-RA treatment, along with H3K9me2 demethylation (Figure 5D and Figures S4E and S5E).

In shPARG cells, H3K4me2 and H3K9ac appeared on *RARβ2*, *CYP26*, and *SMAD3* promoters, 3 hr after treatment, similarly to shCTL cells (Figures 5A and 5B and Figures S4D, S5D, and S6D). In contrast to shCTL cells, a dramatic increase in H3K9me2 and Histone H1 (for *RARβ2*) occurred at the *RARβ2* and *CYP26* promoters, 1 hr posttreatment (Figures 5C and 5D, Figures S4F and S5F). Transfection of PARG<sup>WT</sup>-GFP in shPARG restored a profile of H3K9me2 and H1 occupancy at the promoter of *RARβ2* similar to that of shCTL cells (Figures 5C and 5D and Figure S4I). Note that H3K9me2 and H1 were released similarly from the promoter of *SMAD3* in both cell lines (Figures S6E and S6F).

Recent finding demonstrated that PARP-1 plays a role in the regulation of histone methylation (Krishnakumar and Kraus, 2010a). Since the absence of PARG led to an increase in the repressive H3K9me2 histone mark, we next focused on a histone methyltransferase and a demethylase whose activities regulate the methylation state of H3K9. The methyltransferase KMT1C (also called G9a) (Allis et al., 2007) is responsible for the majority of H3K9me2 marks in cells (Patnaik et al., 2004). The level of KMT1C decreased at the *RARβ2* promoter soon after t-RA-treatment in shCTL cells, while it remained stable in shPARG cells (Figure 5E). The H3K9me3/2 demethylase KDM4D (Klose et al.,

2006; Whetstine et al., 2006; Shin and Janknecht, 2007) accumulated at the *RARβ2* promoter in shCTL cells soon after t-RA treatment but was not recruited throughout the time course in shPARG cells (Figure 5F). However, transfection of PARG<sup>WT</sup>-GFP construct restored the accumulation of KDM4D and the release of KMT1C (Figures 5E and 5F).

Using siRNA against KDM4D, we next demonstrated the involvement of KDM4D in *RARβ2*, *CYP26*, and *PDK4* gene activation in HeLa cells (Figure 5G and Figure S7). Interestingly, KDM4D was dispensable for the transactivation of the PARG-independent gene *SMAD3* (Figure S7). Taken together, these results suggest that PARG controls chromatin derepression at *RAR*-dependent gene promoters.

### Mutations of Two Residues Located in the JmjN Domain of KDM4D Prevent PARsylation after t-RA Treatment

The above data suggest that the interplay between PARP and PARG may control the activity of KDM4D or KMT1C. Since any substrate of PARG is first a substrate of PARP enzymes, we asked whether KDM4D or KMT1C was modified by PARP-1. Recombinant KDM4D (His-KDM4D[1–350] [Weiss et al., 2010]) was modified following the addition of recombinant affinity-purified human PARP-1, in vitro (Figure 6A, compare lanes 6 and 7). By comparison, KMT1C was only weakly modified by PARP-1 (Figure 6A, compare lanes 4 and 5). Histone H1, a well-known PARP target, was used as a positive control (Figure 6A, lanes 2 and 3). Furthermore, a polymer blot assay (Quénet et al., 2008) showed that KDM4D was not able to bind PAR noncovalently (data not shown).

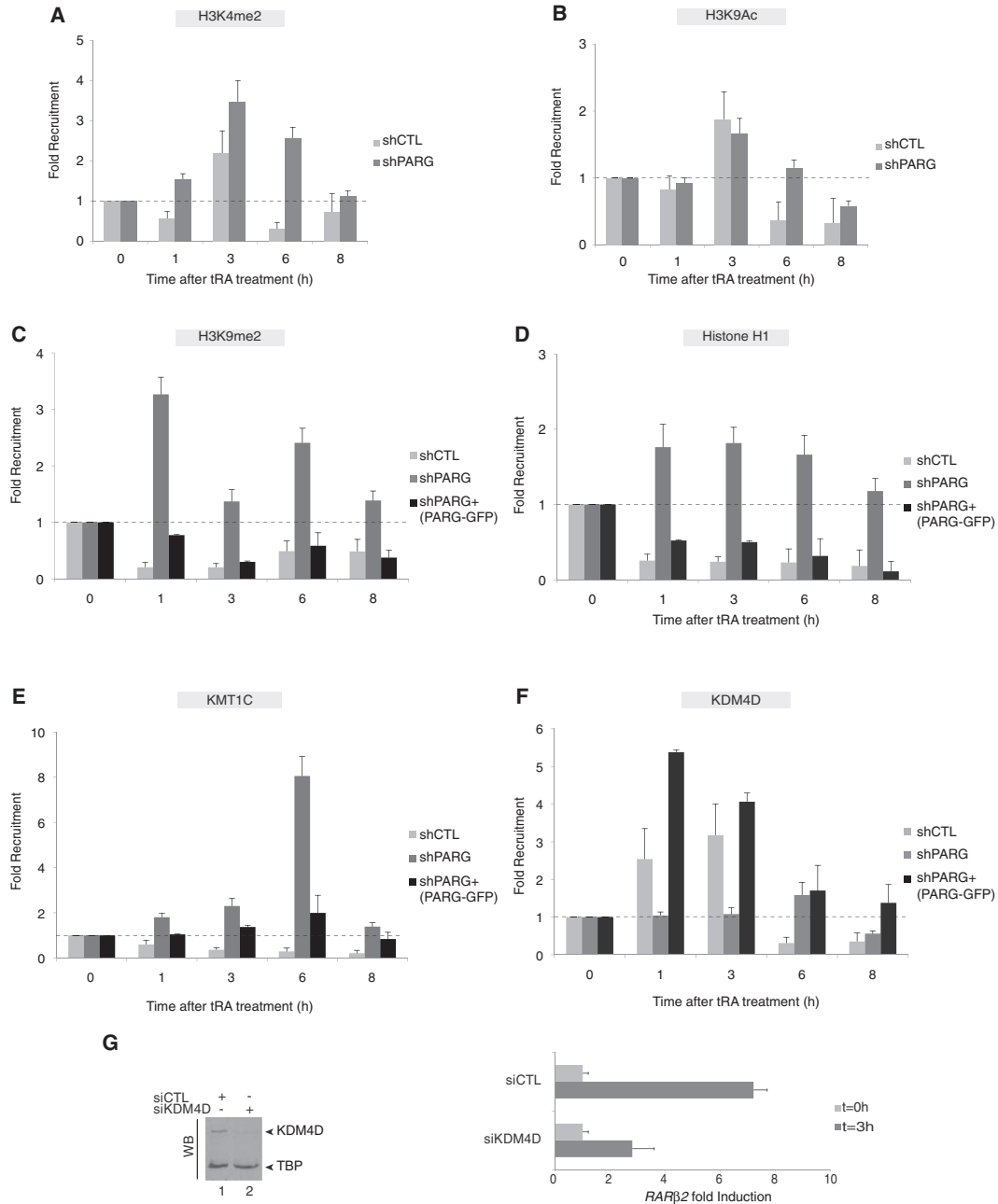
The catalytic Jumonji C domain (JmjC) characterizes the KDM4/JMJD2 family members (Tsukada et al., 2006). In addition, a conserved N-terminal domain called JmjN, rich in amino acids (aa) that can be substrates for PARP-1, is present in KDM4 proteins including KDM4D (from aa 15 to 58) (Balcianas and Ronne, 2000; Chen et al., 2006; Katoh and Katoh, 2004). After deletion of its JmjN domain, GST-KDM4D<sup>Del(15–58)</sup> was no longer modified by PARP-1 (Figure 6B, compare lanes 5 and 6 with lanes 3 and 4). Two potential substrates for PARP-1 (E26 and E27) present in the JmjN domain and conserved in the KDM4 demethylase family (Figure 6C, upper panel) were mutated to alanine, and the resulting GST-KDM4D<sup>E26/27A</sup> was assayed for PARsylation. The double mutant showed a complete lack of modification by PARP-1 (Figure 6C, compare lanes 8 and 9 with lanes 2–7), demonstrating that mutation of these two residues prevents PARsylation. In a Histone H3K9me2 demethylation assay, the mutations E26A/E27A did not affect the activity of the demethylase (Figure 6D, compare lanes 5–7 with lanes 2–4), suggesting that they do not alter the structure of KDM4D.

We next analyzed the PARsylation of KDM4D in vivo. We expressed GFP-KDM4D<sup>WT</sup> and GFP-KDM4D<sup>E26/27A</sup> in shPARG

(B) ChIP monitoring the t-RA-dependent occupancy ( $\pm$ SEM, two independent experiments) of PARG on the *RARβ2* promoter in chromatin extracts from either shCTL (light gray), shPARG (dark gray), or shPARG cells transfected with shRNA-resistant PARG<sup>WT</sup>-GFP construct (black). Chromatin extracts were immunoprecipitated with an anti-PARG in the upper panel and with an anti-GFP in the lower panel.

(C) (Upper panel) Diagram of the endogenous human *SMAD3* promoter. Positions of the primers for amplification of promoter region in ChIP are indicated with arrows. (Lower panels) ChIP monitoring the t-RA-dependent occupancy ( $\pm$ SEM, two independent experiments) of RNA Pol II and PARG on the *SMAD3* promoter in chromatin extracts from either shCTL or shPARG cells.



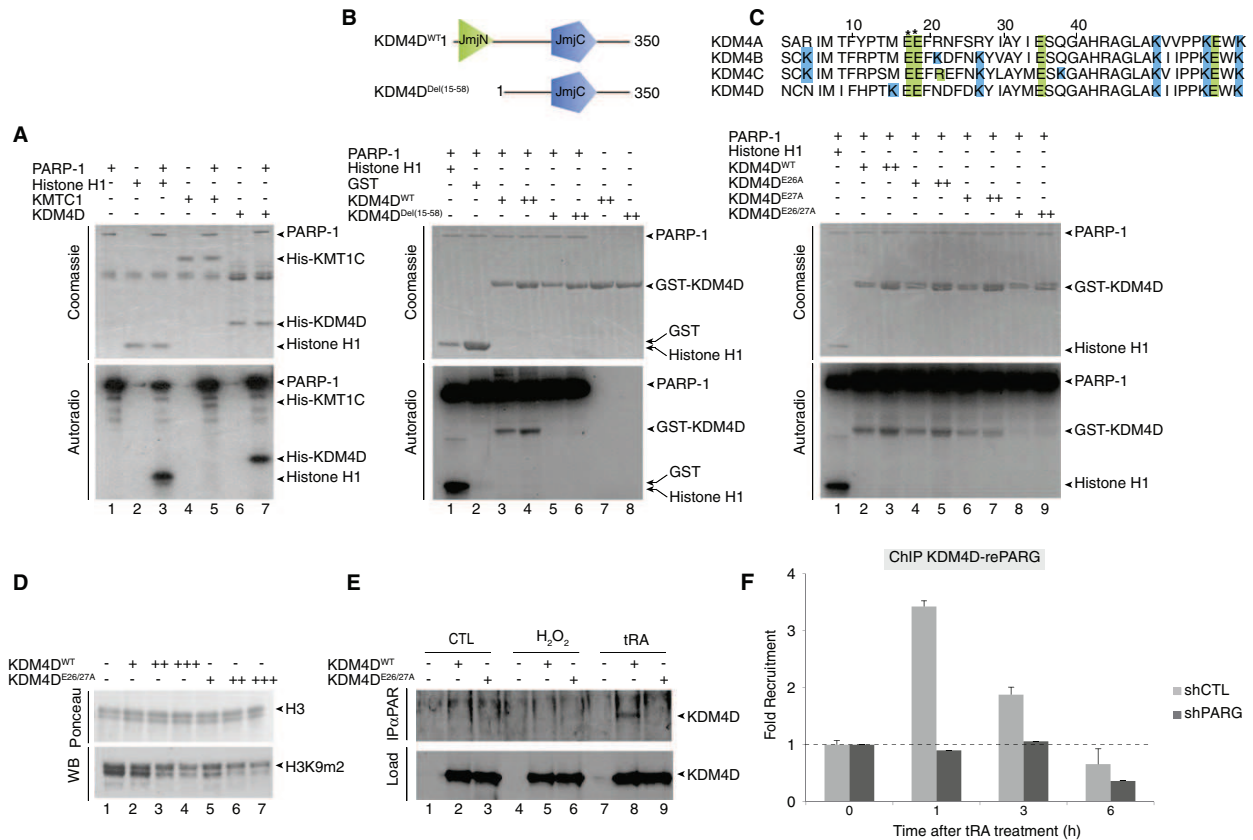


**Figure 5. Chromatin Remodeling at the *RARβ2* Promoter**

(A and B) ChIP monitoring the t-RA-dependent occupancy ( $\pm$ SEM, two independent experiments) of H3K4me2 (A), H3K9Ac (B) on the promoter of *RARβ2* gene in chromatin extracts from either shCTL cells or shPARG cells.

(C–F) ChIP monitoring the t-RA-dependent occupancy ( $\pm$ SEM, two independent experiments) of H3K9me2 (C), Histone H1 (D), KMT1C (E), KDM4D (F) on the promoter of *RARβ2* gene in chromatin extracts from either shCTL cells, shPARG cells, or shPARG cells transfected with (PARG<sup>WT</sup>-GFP).

(G) (Left panel) HeLa cells were transfected either with siRNA control or with siRNA against KDM4D. Forty-eight hours later, proteins from whole-cell extracts (50  $\mu$ g) were resolved by SDS-PAGE followed by western blotting using anti-KDM4D and anti-TBP antibodies. (Right panel) Relative mRNA expression ( $\pm$ SD, three independent experiments) of *RARβ2* in HeLa cells transfected either with siRNA control or with siRNA against KDM4D, measured 3 hr after treatment with t-RA (10  $\mu$ M). The values are expressed relative to the expression level of the no treatment control that is set to 1 in all experiments.



**Figure 6. Two Glutamic Acid Residues in the JmJN Domain of KDM4D Are Required for PARylation**

(A) Two hundred and fifty nanograms of recombinant Histone H1 (lanes 2 and 3), His-KMT1C (lanes 4 and 5), and His-KDM4D(1–350) (lanes 6 and 7) were mixed with 100 ng of recombinant PARP-1 (lanes 3, 5, and 7) in the presence of <sup>32</sup>P-NAD<sup>+</sup> and 100 nM of cold NAD<sup>+</sup>. These assay conditions used favor synthesis of short polymers due to limiting amounts of NAD. Samples were run on a SDS-PAGE gel followed by Coomassie staining (Coomassie) and autoradiography (Autoradio).

(B) (Upper panel) Schematic representation of KDM4D(1–350). JmJN and JmJC domains are represented in green and blue, respectively. (Lower panel) One hundred and fifty nanograms of recombinant Histone H1 (lane 1), 300 ng of GST (lane 2), 150 and 300 ng of GST-KDM4D<sup>WT</sup> (lanes 3 and 4), or GST-KDM4D<sup>Del(15–58)</sup> (lanes 5 and 6) was mixed with 100 ng of recombinant PARP-1 and <sup>32</sup>P-NAD<sup>+</sup>. Samples were run on a SDS-PAGE gel followed by Coomassie staining (Coomassie) and autoradiography (Autoradio). Three hundred and fifty nanograms of GST-KDM4D<sup>WT</sup> (lane 7) or GST-KDM4D<sup>Del(15–58)</sup> (lane 8) was incubated alone as controls.

(C) (Top panel) Sequence alignment of the JmJN domains of the KDM4 family members performed with the ClustalX multiple sequence alignment software. PARP-1 potential targets (lysine and glutamic acid residues) are indicated in blue and green, respectively. Lysine and glutamic acid residues are particularly enriched in the JmJN domain of KDM4D in which they represent ~30% of the total amino acids content compared with 6% in the total protein. The two conserved glutamic acids present in all KDM4 family members are marked with a star. (Lower panel) GST-KDM4D<sup>WT</sup> (lanes 2 and 3), GST-KDM4D<sup>E26A</sup> (lanes 4 and 5), GST-KDM4D<sup>E27A</sup> (lanes 6 and 7), and GST-KDM4D<sup>E26/27A</sup> (lanes 8 and 9) were treated as in (B).

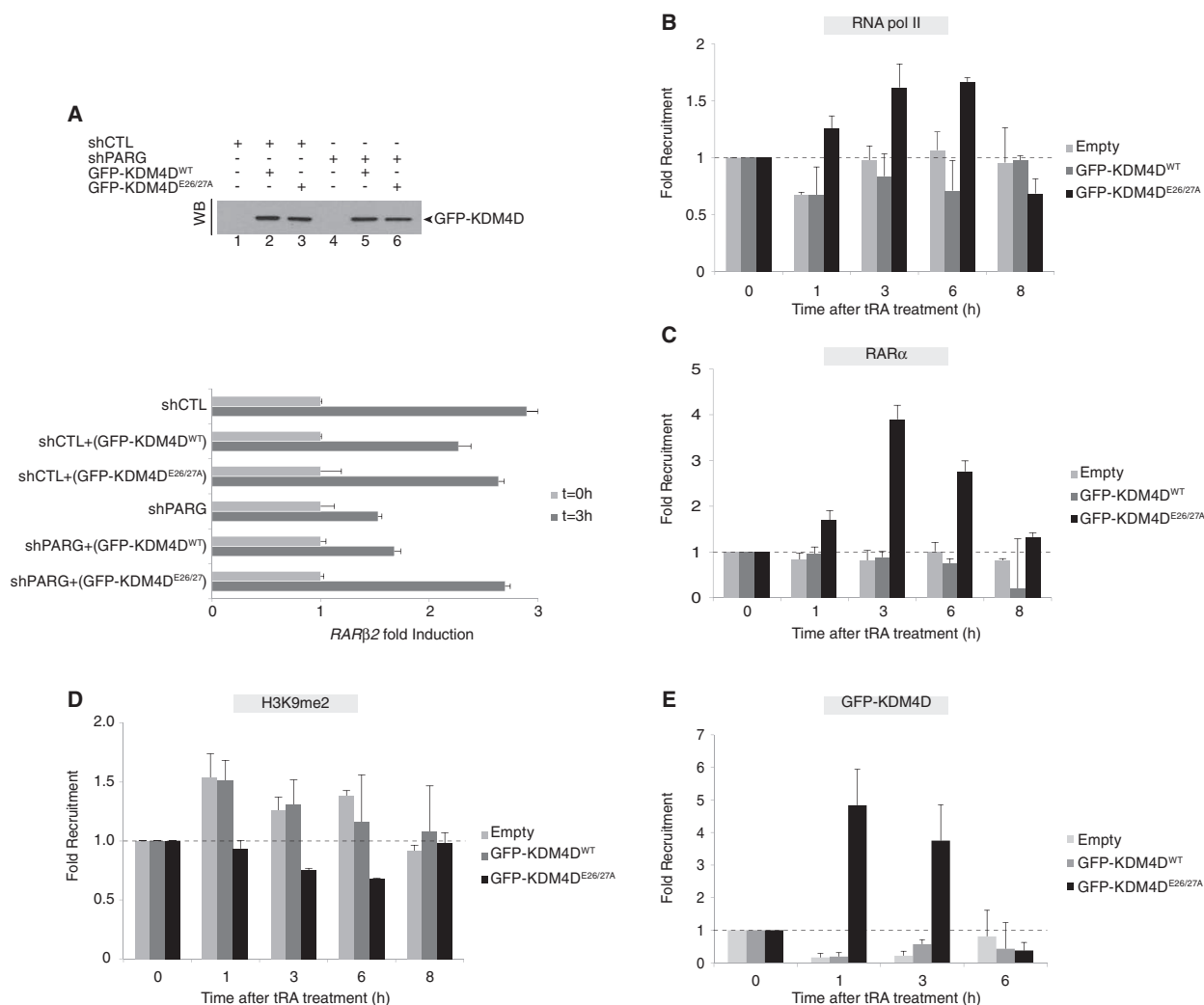
(D) GST-KDM4D<sup>WT</sup> and GST-KDM4D<sup>E26/27A</sup> were used in a demethylation assay containing core histones from HeLa. The reactions were then subjected to either western blotting using anti-Histone H3K9me2 antibody (WB) or Red Ponceau staining (Ponceau).

(E) GFP-KDM4D<sup>WT</sup> or GFP-KDM4D<sup>E26/27A</sup> was transiently expressed in shPARG cells. Immunoprecipitation was performed from 200 μg of whole-cell extracts 10 min after treatment with H<sub>2</sub>O<sub>2</sub> (1 mM) or 1 hr after treatment with t-RA (10 μM), using anti-PARG antibody. Following SDS-PAGE, western blotting analysis was performed with an anti-GFP antibody on the immunoprecipitated material (upper panel) or on 20 μg of whole-cell extracts (lower panel).

(F) ChIP-reChIP monitoring the t-RA-dependent coaccumulation (±SEM, three independent experiments) of PARG and KDM4D on the *RARβ2* promoter in chromatin extracts of shCTL and shPARG cells. A first IP performed against KDM4D was followed by a second IP using anti-PARG antibody on the eluted complexes.

cells, treated the cells with t-RA for 1 hr, immunoprecipitated the material with an anti-PARG antibody, and subjected the precipitate to western blotting with an anti-GFP antibody. GFP-KDM4D<sup>WT</sup> was pulled down by anti-PARG antibody after t-RA treatment (Figure 6E, compare lanes 2 and 8), while GFP-

KDM4D<sup>E26/27A</sup> was not (Figure 6E, compare lanes 3 and 9). We also noticed that GFP-KDM4D<sup>WT</sup> was not pulled down following H<sub>2</sub>O<sub>2</sub> treatment (Figure 6E, compare lanes 2, 5, and 8), suggesting that PARylation of KDM4D(1–350) is specific to t-RA treatment.



**Figure 7. KDM4D<sup>E26/27A</sup> Restores RAR-Dependent Transactivation in the Absence of PARG**

(A) (Upper panel) shCTL (lanes 1–3) or shPARG (lanes 4–6) cells were transfected either with GFP-KDM4D<sup>WT</sup> (lanes 2 and 5) or GFP-KDM4D<sup>E26/27A</sup> (lanes 3 and 6). Forty hours later, proteins from whole-cell extracts (50  $\mu$ g) were resolved by SDS-PAGE followed by western blot using an anti-GFP antibody. (Lower panel) Relative mRNA expression of *RAR $\beta$ 2* in the indicated cells ( $\pm$ SD, three independent experiments), measured 3 hr after treatment with t-RA (10  $\mu$ M). The values are expressed relative to the expression level of the no treatment control that is set to 1 in all experiments.

(B–D) ChIP monitoring the occupancy ( $\pm$ SEM, two independent experiments) of RNA Pol II (B), RAR $\alpha$  (C), and H3K9me2 (D) on the *RAR $\beta$ 2* promoter in chromatin extracts from shPARG cells transfected either with an empty vector (light gray), with GFP-KDM4D<sup>WT</sup> (dark gray), or with GFP-KDM4D<sup>E26/27A</sup> (black) expression vectors.

(E) ChIP monitoring the occupancy ( $\pm$ SEM, two independent experiments) of either GFP-KDM4D<sup>WT</sup> or GFP-KDM4D<sup>E26/27A</sup> on the *RAR $\beta$ 2* promoter in chromatin extracts from shPARG cells after t-RA treatment.

Next, we used a ChIP-reChIP assay to test whether PARG and KDM4D were in close contact at the *RAR $\beta$ 2* promoter. Our results revealed a coaccumulation of PARG and KDM4D, which occurred 1 hr after t-RA treatment in shCTL cells (Figure 6F and Figure S8). This coaccumulation was absent in shPARG cells (Figure 6F and Figure S8). No signal was detected in the ChIP-reChIP experiment with KMT1C (data not shown). Altogether, these data suggest that KDM4D is PARsylated following t-RA treatment and that PARG counteracts this modification.

#### KDM4D<sup>E26/27A</sup> Restores RAR-Dependent Transactivation in the Absence of PARG

To test the above hypothesis, we sought to assess the need of KDM4D PARsylation for RAR-mediated gene expression. Expression of the unPARsylable GFP-KDM4D<sup>E26/27A</sup> in shPARG cells rescued *RAR $\beta$ 2*, *CYP26*, *PDK4*, and *NR1P1* expression, which was not achieved with GFP-KDM4D<sup>WT</sup> (Figure 7A and Figure S9). In agreement with the rescue of mRNA transactivation, expression of GFP-KDM4D<sup>E26/27A</sup>, but not GFP-KDM4D<sup>WT</sup>,

## Molecular Cell

### PARG Is a Transcriptional Coactivator

partially restored the recruitment of both RNA Pol II and RAR $\alpha$  to the RAR $\beta$ 2 promoter (Figures 7B and 7C and Figures S10A and S10B). This recruitment was accompanied by a decrease in H3K9me2 modification until 6 hr posttreatment (Figure 7D and Figure S10C). We also monitored the presence of KDM4D at the RAR $\beta$ 2 and CYP26 promoters in shPARG and found a recruitment of GFP-KDM4D<sup>E26/27A</sup> early after t-RA treatment that did not take place with GFP-KDM4D<sup>WT</sup> (Figures 7E, Figures S10D and S10E), suggesting that PARsylation regulates the binding of the histone demethylase to the chromatin (Krishnakumar and Kraus, 2010a). These results demonstrate that an unPARsylatable KDM4D mutant can compensate for the absence of PARG and restores a permissive chromatin environment suitable for RAR-dependent gene transcription.

## DISCUSSION

### PARG Is a Coactivator of RAR-Mediated Gene Expression

PARG has been extensively studied for its role in the DNA damage response and in cell death, but some studies have also suggested a role in transcription (Frizzell et al., 2009; Tulin et al., 2006). Here, we demonstrate that PARG functions as a transcriptional coactivator that modulates RAR-mediated gene expression. Global transcription profiling and individual gene studies show that the transactivation of several RAR-dependent genes was impaired in the absence of PARG, which localized to the promoter of these genes in a ligand-dependent manner. The recruitment of PARG to these promoters occurred 1 hr posttreatment and preceded the recruitment of the transcription machinery. This observation suggests that PARG may be required for the formation of the PIC. Indeed, we observed that the recruitment of RNA Pol II or RAR to the promoters of RAR-dependent genes was dependent on PARG. An shRNA-resistant PARG<sup>WT</sup> construct rescued the NR-dependent transactivation in shPARG cells, excluding any off-target effects of the shRNA. We also identified genes whose transactivation by t-RA was independent of PARG. In agreement with this observation, PARG does not localize to the promoter of these genes, indicating that transactivation by t-RA can occur following different molecular pathways.

### PAR-Degrading Activity Is Required for RAR-Mediated Transcription Process

The role of the PAR-degrading activity of PARG in transcription has been examined but has led to contradictory results (Frizzell et al., 2009; Rapizzi et al., 2004). In our well-defined system, we observed that RAR-mediated transcription required the PARG catalytic activity. Our data imply that PARG targets a substrate that must be dePARsylated before transactivation takes place. In agreement with this hypothesis, a rescue in the transactivation of the RAR-dependent genes in PARG-depleted cells was obtained by treatment of these cells with a PARP-specific inhibitor. We noticed that treatment of shCTL cells with this inhibitor had no effect on RAR-dependent gene expression, in agreement with previous results indicating that RAR-mediated gene transcription was dependent on the presence of the PARP-1 protein but not on its activity (Pavri et al., 2005).

Our data support these observations but further demonstrate that PARP activity may even have a repressive effect on RAR-dependent gene expression, in the absence of PARG.

### PARG Promotes a Permissive Chromatin Environment around the Promoter of NR-Dependent Genes

The absence of PIC formation in PARG-depleted cells and the involvement of PAR in the regulation of chromatin structure (Krishnakumar and Kraus, 2010b) led us to analyze histone modifications around the promoters of RAR-dependent genes. H3K9ac or H3K4me2, active marks of transcription, were present after t-RA treatment even in the absence of PARG. Surprisingly, we detected a peak of H3K9me2 and the persistence of Histone H1, two strong marks of transcriptional repression, soon after t-RA treatment of shPARG cells. The peak of H3K9me2 appeared at the promoter of the RAR-dependent genes in shPARG between 1 and 3 hr after t-RA-treatment, a time when PARG was recruited to these promoters in shCTL cells. This indicates that PARG may work to alleviate the transcriptional repression induced by H3K9me2, thereby triggering the formation of the PIC.

In line with this hypothesis, the H3K9me2 demethylase KDM4D was efficiently PARsylated by PARP-1 and coaccumulates with PARG on these promoters after t-RA treatment. Besides, KDM4D was required for the transactivation of PARG-dependent genes but dispensable for PARG-independent ones. What exactly dictates a gene's PARG dependency remains to be further investigated. However, it appears that SMAD3, a PARG-independent gene, contains a promoter with multiple SP1 binding sites that are supposed to recruit RAR. On the other hand, RAR $\beta$ 2, CYP26, PDK4, and TNFAIP2 show the classical RXR/RAR binding sites (Cheng et al., 2011). Therefore, the nature of the NR binding element may induce a variation in the cofactors required to transactivate various genes with the same NR.

### Interplay between PARP-1 and PARG Controls the Activity of KDM4D on the Promoter of NR-Dependent Genes

The JmjC domain-containing histone demethylases (JHDM proteins) consist of several subfamilies that can remove methyl groups from the modified H3K4, H3K9, H3K27, or H3K36 residues (Hou and Yu, 2010; Varier and Timmers, 2011). KDM2A is the founding member of this large family, which is characterized by the presence of a JmjC catalytic domain (Tsukada et al., 2006). Some JHDM proteins also possess a conserved N-terminal motif (JmjN domain), strictly associated with the JmjC domain (Balciunas and Ronne, 2000). The function of the JmjN domain is still unclear, but it has been shown to have extensive interactions with JmjC through the formation of hydrogen bonds and hydrophobic contacts (Chen et al., 2006). Furthermore, removal of the JmjN domain impairs the stability and activity of JmjC (Chen et al., 2006). The JmjN domain typically spans 40 amino acids, as depicted in Figure 6. Our results show that two glutamic residues belonging to the JmjN domain are required for the modification of KDM4D by PARP-1.

The fact that unPARsylatable mutant can compensate for the absence of PARG and rescued RAR-dependent transactivation

in shPARG cells, while overexpression of KDM4D<sup>WT</sup> did not, suggests that PARsylation of KDM4D was responsible for the inhibition of RAR-dependent gene transactivation. These data suggest that PARsylation of KDM4D negatively regulates its activity at the promoter of RAR-dependent genes and, thereby, the transactivation of these genes. PARsylation may lead to the destabilization of the JmjN/JmjC interaction that is required for KDM4D activity. Alternatively, the PARsylation of KDM4D may destabilize the interaction between KDM4D and the chromatin. ChIP results show a higher occupancy of the GFP-KDM4D<sup>E26/27A</sup> mutant at RAR-dependant promoters following t-RA treatment that may suggest that PARsylation of KDM4D regulates its binding to chromatin, as demonstrated for the H3K4 demethylase KDM5B (Krishnakumar and Kraus, 2010a). Due to the low level of PARsylation of KDM4D that we were able to achieve *in vitro*, it was not possible to test the consequences of PARsylation on KDM4D activity in a demethylase assay. Altogether, our data allow us to propose a model in which the activity of PARP-1 regulates RAR-dependent gene transcription through the depletion of an active KDM4D from the chromatin, thereby establishing a heterochromatin landscape. In this model, PARG counteracts the action of PARP-1 to induce an open chromatin structure and an active transcription process. Whether or not the amount of PARG, PARP-1, or their subcellular localization in different cell types would lead to a cell-specific modulation of RAR-dependant gene transcription remains to be established.

As stated above, it is also likely that other KDMs are regulated in a similar manner. Krishnakumar and Kraus demonstrated recently that the PARsylation of KDM5B regulates the genome-wide methylation of histone H3K4 in basal transcription by impacting its interaction with the nucleosomes (Krishnakumar and Kraus, 2010a). It is interesting to note that KDM5B contains a JmjN domain (Yamane et al., 2007) with two conserved successive glutamic acid residues that, in light of our results, may be crucial for its PARsylation. Future work in this area will broaden our knowledge of the role of PARG in transcription, which could prove as important as that of PARP-1.

## EXPERIMENTAL PROCEDURES

### Cell Lines

shCTL (BD650) and shPARG (PARG<sup>KD</sup>) cells were characterized previously (Amé et al., 2009). Cells were maintained in Dulbecco's modified Eagle's medium (DMEM; GIBCO-BRL, Invitrogen) supplemented with 10% fetal calf serum (FCS), 40 µg/ml gentamicin, and 125 µg/ml hygromycin B. NR-dependent gene activation was performed as described (Keriel et al., 2002) (also see the Supplemental Information). F9 mouse embryonic carcinoma cells were cultured in DMEM supplemented with 10% FCS and 40 µg/ml gentamicin. ON-TARGET plus smart pool siRNA control or targeting the mouse PARG was transiently transfected in F9 cells with Lipofectamine 2000 (Invitrogen). After 24 hr of transfection, the cells were first treated with KU0058948 (100 nM) to avoid the lethal accumulation of PAR and 24 hr later with t-RA (10<sup>-7</sup>M) for differentiation.

### <sup>32</sup>P-NAD<sup>+</sup> PARsylation Assay

PAR reaction was carried out as described (Amé et al., 1999). Briefly, 100 ng of recombinant PARP-1 (Enzo Life Sciences) was incubated with 250 ng of recombinant histone H1.2 (Enzo Life Sciences), KMT1C (active motif), or KDM4D (Weiss et al., 2010) or purified GST-tagged proteins as indicated in

the Figure 6 legend, in PAR reaction buffer (50 mM Tris-HCl [pH 8.0], 4 mM MgCl<sub>2</sub>, 300 mM NaCl, 1 mM DTT) supplemented with 0.1 µg/ml BSA, 6 pmol of DNase I-activated DNA, 1 µCi<sup>32</sup>P-NAD<sup>+</sup>, and 100 nM cold NAD<sup>+</sup>. Reaction was allowed for 20 min at 25°C, stopped by addition of SDS-PAGE loading buffer and boiling for 5 min. Samples were subjected to SDS-PAGE. Detection of automodification was carried out by autoradiography and detection of proteins by Coomassie staining.

### Demethylation Assay

Demethylation of core histones by KDM4D was performed in 20 mM Tris (pH 7.3), 150 mM NaCl, 50 µM [NH<sub>4</sub>]<sub>2</sub>Fe[SO<sub>4</sub>]<sub>2</sub>·6H<sub>2</sub>O, 1 mM α-ketoglutarate, 2 mM ascorbic acid using 5 µg of purified HeLa histones. The demethylation reaction was incubated for 12 hr at 37°C, and the samples were subsequently analyzed by western blotting.

### Antibodies

Mouse monoclonal antibodies against the RNA polymerase II (7C2), RARα (9A6), and TBP (3G3) were produced at the IGBMC facility.

Rabbit polyclonal anti-Nter-PARG, mouse monoclonal anti-PARP1 (C2-10), and anti poly(ADP-ribose) (10H) were previously described (Amé et al., 2009).

Rabbit polyclonal anti-GFP (TP401) was produced at Torrey Pines Biolabs. Rabbit polyclonal anti-KMT1C/G9a (3306), anti-dimethyl histone H3K9 (9753), anti-dimethyl histone H3K4 (9726), and anti-acetylhistone H3K9 (9671) were produced at Cell Signaling. Rabbit polyclonal anti-KDM4D (AB93694) was produced at Abcam. Mouse monoclonal anti-H1 antibody (clone AE4) was produced at Millipore. Mouse monoclonal anti-β-tubulin (MAB3408) was produced at Chemicon.

### ACCESSION NUMBERS

The data in this paper have been deposited in the GEO Database (<http://www.ncbi.nlm.nih.gov/geo/>) under accession number GSE40883.

### SUPPLEMENTAL INFORMATION

Supplemental Information includes four tables, ten figures, Supplemental Experimental Procedures, and Supplemental References and can be found with this article at <http://dx.doi.org/10.1016/j.molcel.2012.09.021>.

### ACKNOWLEDGMENTS

We are grateful to Zita Nagy and Emmanuel Compe for their critical reading. We thank Steve Brooks for English proofreading. We thank Robert Schneider for the His-KDM4D construct. This study was supported by funds from the Ligue Contre le Cancer (V.S. Equipe Labellisée and CCIR-GE); from the Laboratory of Excellence MEDALIS, Initiative of Excellence (IdEx), Strasbourg University, France; from Electricité de France; from the French National Research Agency (ANR-08-GENOPAT-042); and from the Institut National du Cancer (INCA-2008-041). The group of J.-M.E and F.C. is supported by the ERC (ERC-2008-TRANSREACT). N.L. is supported by the INSERM.

Received: February 16, 2012

Revised: June 29, 2012

Accepted: September 18, 2012

Published online: October 25, 2012

### REFERENCES

- Allis, C.D., Berger, S.L., Cote, J., Dent, S., Jenuwien, T., Kouzarides, T., Pillus, L., Reinberg, D., Shi, Y., Shiekhhattar, R., et al. (2007). New nomenclature for chromatin-modifying enzymes. *Cell* 131, 633–636.
- Amé, J.C., Apiou, F., Jacobson, E.L., and Jacobson, M.K. (1999). Assignment of the poly(ADP-ribose) glycohydrolase gene (PARG) to human chromosome 10q11.23 and mouse chromosome 14B by *in situ* hybridization. *Cytogenet. Cell Genet.* 85, 269–270.



- Amé, J.C., Spenlehauer, C., and de Murcia, G. (2004). The PARP superfamily. *Bioessays* 26, 882–893.
- Amé, J.C., Fouquerel, E., Gauthier, L.R., Biard, D., Boussin, F.D., Dantzer, F., de Murcia, G., and Schreiber, V. (2009). Radiation-induced mitotic catastrophe in PARG-deficient cells. *J. Cell Sci.* 122, 1990–2002.
- Balciunas, D., and Ronne, H. (2000). Evidence of domain swapping within the jumonji family of transcription factors. *Trends Biochem. Sci.* 25, 274–276.
- Chen, Z., Zang, J., Whetstine, J., Hong, X., Davrazou, F., Kutateladze, T.G., Simpson, M., Mao, Q., Pan, C.H., Dai, S., et al. (2006). Structural insights into histone demethylation by JMJD2 family members. *Cell* 125, 691–702.
- Cheng, Y.H., Utsunomiya, H., Pavone, M.E., Yin, P., and Bulun, S.E. (2011). Retinoic acid inhibits endometrial cancer cell growth via multiple genomic mechanisms. *J. Mol. Endocrinol.* 46, 139–153.
- Clagett-Dame, M., and Knutson, D. (2011). Vitamin A in reproduction and development. *Nutrients* 3, 385–428.
- Eifert, C., Sangster-Guity, N., Yu, L.M., Chittur, S.V., Perez, A.V., Tine, J.A., and McCormick, P.J. (2006). Global gene expression profiles associated with retinoic acid-induced differentiation of embryonal carcinoma cells. *Mol. Reprod. Dev.* 73, 796–824.
- Erdélyi, K., Bai, P., Kovács, I., Szabó, E., Mocsár, G., Kakuk, A., Szabó, C., Gergely, P., and Virág, L. (2009). Dual role of poly(ADP-ribose) glycohydrolase in the regulation of cell death in oxidatively stressed A549 cells. *FASEB J.* 23, 3553–3563.
- Fisher, A.E., Hohegger, H., Takeda, S., and Caldecott, K.W. (2007). Poly(ADP-ribose) polymerase 1 accelerates single-strand break repair in concert with poly(ADP-ribose) glycohydrolase. *Mol. Cell. Biol.* 27, 5597–5605.
- Frizzell, K.M., Gamble, M.J., Berrocal, J.G., Zhang, T., Krishnakumar, R., Cen, Y., Sauve, A.A., and Kraus, W.L. (2009). Global analysis of transcriptional regulation by poly(ADP-ribose) polymerase-1 and poly(ADP-ribose) glycohydrolase in MCF-7 human breast cancer cells. *J. Biol. Chem.* 284, 33926–33938.
- Glass, C.K., and Rosenfeld, M.G. (2000). The coregulator exchange in transcriptional functions of nuclear receptors. *Genes Dev.* 14, 121–141.
- Hassa, P.O., Buerki, C., Lombardi, C., Imhof, R., and Hottiger, M.O. (2003). Transcriptional coactivation of nuclear factor-kappaB-dependent gene expression by p300 is regulated by poly(ADP-ribose) polymerase-1. *J. Biol. Chem.* 278, 45145–45153.
- Heeres, J.T., and Hergenrother, P.J. (2007). Poly(ADP-ribose) makes a date with death. *Curr. Opin. Chem. Biol.* 11, 644–653.
- Hou, H., and Yu, H. (2010). Structural insights into histone lysine demethylation. *Curr. Opin. Struct. Biol.* 20, 739–748.
- Ji, Y., and Tulin, A.V. (2010). The roles of PARP1 in gene control and cell differentiation. *Curr. Opin. Genet. Dev.* 20, 512–518.
- Ju, B.G., Lunyak, V.V., Perissi, V., Garcia-Bassets, I., Rose, D.W., Glass, C.K., and Rosenfeld, M.G. (2006). A topoisomerase IIbeta-mediated dsDNA break required for regulated transcription. *Science* 312, 1798–1802.
- Kastner, P., Mark, M., and Chambon, P. (1995). Nonsteroid nuclear receptors: what are genetic studies telling us about their role in real life? *Cell* 83, 859–869.
- Katoh, M., and Katoh, M. (2004). Identification and characterization of JMJD2 family genes in silico. *Int. J. Oncol.* 24, 1623–1628.
- Keriel, A., Stary, A., Sarasin, A., Rochette-Egly, C., and Egly, J.M. (2002). XPD mutations prevent TFIIH-dependent transactivation by nuclear receptors and phosphorylation of RARalpha. *Cell* 109, 125–135.
- Klose, R.J., Yamane, K., Bae, Y., Zhang, D., Erdjument-Bromage, H., Tempst, P., Wong, J., and Zhang, Y. (2006). The transcriptional repressor JHDM3A demethylates trimethyl histone H3 lysine 9 and lysine 36. *Nature* 442, 312–316.
- Koh, D.W., Lawler, A.M., Poitras, M.F., Sasaki, M., Wattler, S., Nehls, M.C., Stöger, T., Poirier, G.G., Dawson, V.L., and Dawson, T.M. (2004). Failure to degrade poly(ADP-ribose) causes increased sensitivity to cytotoxicity and early embryonic lethality. *Proc. Natl. Acad. Sci. USA* 101, 17699–17704.
- Kraus, W.L. (2008). Transcriptional control by PARP-1: chromatin modulation, enhancer-binding, coregulation, and insulation. *Curr. Opin. Cell Biol.* 20, 294–302.
- Krishnakumar, R., and Kraus, W.L. (2010a). PARP-1 regulates chromatin structure and transcription through a KDM5B-dependent pathway. *Mol. Cell* 39, 736–749.
- Krishnakumar, R., and Kraus, W.L. (2010b). The PARP side of the nucleus: molecular actions, physiological outcomes, and clinical targets. *Mol. Cell* 39, 8–24.
- Lalève, S., Anno, Y.N., Chatagnon, A., Samarut, E., Poch, O., Laudet, V., Benoit, G., Lecompte, O., and Rochette-Egly, C. (2011). Genome-wide in silico identification of new conserved and functional retinoic acid receptor response elements (direct repeats separated by 5 bp). *J. Biol. Chem.* 286, 33322–33334.
- Malanga, M., Czuby, A., Girstun, A., Staron, K., and Althaus, F.R. (2008). Poly(ADP-ribose) binds to the splicing factor ASF/SF2 and regulates its phosphorylation by DNA topoisomerase I. *J. Biol. Chem.* 283, 19991–19998.
- McCabe, N., Lord, C.J., Tutt, A.N., Martin, N.M., Smith, G.C., and Ashworth, A. (2005). BRCA2-deficient CAPAN-1 cells are extremely sensitive to the inhibition of Poly (ADP-Ribose) polymerase: an issue of potency. *Cancer Biol. Ther.* 4, 934–936.
- Meyer, R.G., Meyer-Ficca, M.L., Whatcott, C.J., Jacobson, E.L., and Jacobson, M.K. (2007). Two small enzyme isoforms mediate mammalian mitochondrial poly(ADP-ribose) glycohydrolase (PARG) activity. *Exp. Cell Res.* 313, 2920–2936.
- Meyer-Ficca, M.L., Meyer, R.G., Coyle, D.L., Jacobson, E.L., and Jacobson, M.K. (2004). Human poly(ADP-ribose) glycohydrolase is expressed in alternative splice variants yielding isoforms that localize to different cell compartments. *Exp. Cell Res.* 297, 521–532.
- Miwa, M., and Masutani, M. (2007). PolyADP-ribosylation and cancer. *Cancer Sci.* 98, 1528–1535.
- Mortusewicz, O., Fouquerel, E., Amé, J.C., Leonhardt, H., and Schreiber, V. (2011). PARG is recruited to DNA damage sites through poly(ADP-ribose)- and PCNA-dependent mechanisms. *Nucleic Acids Res.* 39, 5045–5056.
- Niere, M., Kernstock, S., Koch-Nolte, F., and Ziegler, M. (2008). Functional localization of two poly(ADP-ribose)-degrading enzymes to the mitochondrial matrix. *Mol. Cell. Biol.* 28, 814–824.
- Ogino, H., Nozaki, T., Gunji, A., Maeda, M., Suzuki, H., Ohta, T., Murakami, Y., Nakagama, H., Sugimura, T., and Masutani, M. (2007). Loss of Parp-1 affects gene expression profile in a genome-wide manner in ES cells and liver cells. *BMC Genomics* 8, 41.
- Patel, C.N., Koh, D.W., Jacobson, M.K., and Oliveira, M.A. (2005). Identification of three critical acidic residues of poly(ADP-ribose) glycohydrolase involved in catalysis: determining the PARG catalytic domain. *Biochem. J.* 388, 493–500.
- Patnaik, D., Chin, H.G., Estève, P.O., Benner, J., Jacobsen, S.E., and Pradhan, S. (2004). Substrate specificity and kinetic mechanism of mammalian H9a histone H3 methyltransferase. *J. Biol. Chem.* 279, 53248–53258.
- Pavri, R., Lewis, B., Kim, T.K., Dilworth, F.J., Erdjument-Bromage, H., Tempst, P., de Murcia, G., Evans, R., Chambon, P., and Reinberg, D. (2005). PARP-1 determines specificity in a retinoid signaling pathway via direct modulation of mediator. *Mol. Cell* 18, 83–96.
- Quénet, D., Gasser, V., Fouillen, L., Cammas, F., Sanglier-Cianferani, S., Losson, R., and Dantzer, F. (2008). The histone subcode: poly(ADP-ribose) polymerase-1 (Parp-1) and Parp-2 control cell differentiation by regulating the transcriptional intermediary factor TIF1beta and the heterochromatin protein HP1alpha. *FASEB J.* 22, 3853–3865.
- Rapizzi, E., Fossati, S., Moroni, F., and Chiarugi, A. (2004). Inhibition of poly(ADP-ribose) glycohydrolase by gallotannin selectively up-regulates expression of proinflammatory genes. *Mol. Pharmacol.* 66, 890–898.
- Robert, I., Dantzer, F., and Reina-San-Martin, B. (2009). Parp1 facilitates alternative NHEJ, whereas Parp2 suppresses IgH/c-myc translocations during immunoglobulin class switch recombination. *J. Exp. Med.* 206, 1047–1056.
- Rouleau, M., Patel, A., Hendzel, M.J., Kaufmann, S.H., and Poirier, G.G. (2010). PARP inhibition: PARP1 and beyond. *Nat. Rev. Cancer* 10, 293–301.

- Schreiber, V., Dantzer, F., Ame, J.C., and de Murcia, G. (2006). Poly(ADP-ribose): novel functions for an old molecule. *Nat. Rev. Mol. Cell Biol.* 7, 517–528.
- Shin, S., and Janknecht, R. (2007). Activation of androgen receptor by histone demethylases JMJD2A and JMJD2D. *Biochem. Biophys. Res. Commun.* 359, 742–746.
- Simbulan-Rosenthal, C.M., Ly, D.H., Rosenthal, D.S., Konopka, G., Luo, R., Wang, Z.Q., Schultz, P.G., and Smulson, M.E. (2000). Misregulation of gene expression in primary fibroblasts lacking poly(ADP-ribose) polymerase. *Proc. Natl. Acad. Sci. USA* 97, 11274–11279.
- Simbulan-Rosenthal, C.M., Rosenthal, D.S., Luo, R., Samara, R., Espinoza, L.A., Hassa, P.O., Hottiger, M.O., and Smulson, M.E. (2003). PARP-1 binds E2F-1 independently of its DNA binding and catalytic domains, and acts as a novel coactivator of E2F-1-mediated transcription during re-entry of quiescent cells into S phase. *Oncogene* 22, 8460–8471.
- Slade, D., Dunstan, M.S., Barkauskaite, E., Weston, R., Lafite, P., Dixon, N., Ahel, M., Leys, D., and Ahel, I. (2011). The structure and catalytic mechanism of a poly(ADP-ribose) glycohydrolase. *Nature* 477, 616–620.
- Strickland, S., and Mahdavi, V. (1978). The induction of differentiation in teratocarcinoma stem cells by retinoic acid. *Cell* 15, 393–403.
- Tsukada, Y., Fang, J., Erdjument-Bromage, H., Warren, M.E., Borchers, C.H., Tempst, P., and Zhang, Y. (2006). Histone demethylation by a family of JmjC domain-containing proteins. *Nature* 439, 811–816.
- Tulin, A., and Spradling, A. (2003). Chromatin loosening by poly(ADP)-ribose polymerase (PARP) at *Drosophila* puff loci. *Science* 299, 560–562.
- Tulin, A., Naumova, N.M., Menon, A.K., and Spradling, A.C. (2006). *Drosophila* poly(ADP-ribose) glycohydrolase mediates chromatin structure and SIR2-dependent silencing. *Genetics* 172, 363–371.
- Variar, R.A., and Timmers, H.T. (2011). Histone lysine methylation and demethylation pathways in cancer. *Biochim. Biophys. Acta* 1815, 75–89.
- Weiss, T., Hergeth, S., Zeissler, U., Izzo, A., Tropberger, P., Zee, B.M., Dunder, M., Garcia, B.A., Daujat, S., and Schneider, R. (2010). Histone H1 variant-specific lysine methylation by G9a/KMT1C and Glp1/KMT1D. *Epigenetics Chromatin*. 3, 7.
- Whetstone, J.R., Nottke, A., Lan, F., Huarte, M., Smolnikov, S., Chen, Z., Spooner, E., Li, E., Zhang, G., Colaiacovo, M., and Shi, Y. (2006). Reversal of histone lysine trimethylation by the JMJD2 family of histone demethylases. *Cell* 125, 467–481.
- Yamane, K., Tateishi, K., Klose, R.J., Fang, J., Fabrizio, L.A., Erdjument-Bromage, H., Taylor-Papadimitriou, J., Tempst, P., and Zhang, Y. (2007). PLU-1 is an H3K4 demethylase involved in transcriptional repression and breast cancer cell proliferation. *Mol. Cell* 25, 801–812.



## Publication 5

### **The ARCH domain of XPD, an anchoring platform for CAK that conditions TFIIH DNA repair and Transcription Activities.**

Abdulrahman W., Iltis I., Radu L., Braun C., Giraudon C., Egly JM., Poterszman A.

**PNAS (soumis)**

#### **Résumé**

L'hélicase XPD est une sous-unité du facteur de réparation/transcription TFIIH, qui catalyse l'ouverture de l'ADN endommagé lors de la réparation par NER. XPD est formé d'un domaine cluster 4FeS impliqué dans la reconnaissance du dommage et d'un nouveau module appelé domaine ARCH, dont la fonction est inconnue.

Nous avons montrés que ce domaine est crucial pour le recrutement du sous-complexe CAK de TFIIH. En effet la mutation C259Y du domaine ARCH, que l'on retrouve chez des patients TTD, n'affecte pas seulement l'interaction avec MAT1 de TFIIH, mais diminue également l'activité transcriptionnelle de TFIIH *in vitro*. Cette mutation perturbe aussi le recrutement de TFIIH au niveau du promoteur du gène activé RAR $\beta$ 2.

De plus, nous avons montré qu'elle pouvait affecter l'activité hélicase de XPD et l'activité NER de TFIIH. Nos résultats montrent que le domaine ARCH de XPD joue un rôle de plateforme de recrutement du CAK, lors de la formation de TFIIH pendant la transcription. Il intervient également dans la reconnaissance du dommage et dans la régulation de l'activité hélicase de XPD.

**The ARCH domain of XPD, an anchoring platform for CAK that conditions TFIID  
DNA repair and Transcription Activities**

Wassim Abdulrahman, Izarn Iltis, Laura Radu, Cathy Braun, Anne Maglott-Roth,  
Christophe Giraudon, Jean-Marc Egly & Arnaud Poterszman

Correspondance to [Arnaud.POTERSZMAN@igbmc.fr](mailto:Arnaud.POTERSZMAN@igbmc.fr) or [egly@igbmc.fr](mailto:egly@igbmc.fr)

Institut de Génétique et de Biologie Moléculaire et Cellulaire, CNRS/INSERM/UDS,  
BP 163, 67404 Illkirch Cedex, France

## **Keywords**

trichothiodystrophy , rare disease, helicase, regulation of gene expression

## **Abstract**

The *xeroderma pigmentosum* group D helicase is a subunit of transcription/DNA repair factor TFIIH that catalyzes the unwinding of a damaged DNA duplex during nucleotide excision repair. Apart from two canonical helicase domains, XPD is composed of a 4FeS cluster domain involved in DNA damage recognition and of a novel module of uncharacterized function, termed the ARCH domain. By investigating the consequence of a mutation found in a TTD patient, we show that the ARCH domain is critical for the recruitment of the CAK kinase complex. Indeed, this mutation not only affects the interaction with the MAT1 CAK subunit, thereby decreasing the *in vitro* basal transcription activity of TFIIH itself and impeding the efficient recruitment of the transcription machinery on the promoter of an activated gene but also impairs the DNA unwinding activity of XPD and the NER activity of TFIIH. We further demonstrate the role of CAK in down-regulating the XPD helicase activity within TFIIH. Altogether, our results not only identify the ARCH domain of XPD as a platform for the recruitment of CAK but also as a potential molecular switch that might control TFIIH composition and play a key role in the conversion of TFIIH from a factor active in transcription into a DNA repair factor.

\body

## **Introduction**

The xeroderma pigmentosum group D (XPD) gene encodes a 5'-3' helicase (XPB) which harbours mutations in patients suffering from xeroderma pigmentosum (XP), trichothiodystrophy (TTD) and Cockayne syndrome (CS), three rare autosomal recessive diseases<sup>1,2</sup>: XP is characterized by a deficit of the nucleotide excision repair (NER) pathway, leading to sun sensitivity and skin cancer susceptibility; TTD is characterized by sulfur-deficient brittle hair and a variety of neuroectodermal symptoms<sup>3</sup>. XPB is the founding member of a family of DNA helicases conserved in archaea and eukaryotes. All family members share a four domain organization including a conserved (Fe-Se) cluster binding domain that is essential for the helicase activity and a novel module of uncharacterized function, named the ARCH domain<sup>4-7</sup>. While archeal XPB homologues are monomers and have no known stable interactors, eukaryotic XPB homologues are part of the general transcription/DNA repair factor TFIIH, a multi-subunit complex made up of ten subunits (see reference<sup>8</sup> for a review). Low resolution models for TFIIH have been obtained for the yeast<sup>9,10</sup> and the human complex<sup>11</sup>, showing an overall conservation of shape. Human TFIIH can be resolved into two functional and structural entities bridged by XPB: the core-TFIIH consists of XPB, p62, p52, p44, p34, and p8, whereas the CDK-activating kinase (CAK) subcomplex contains cyclin-dependent kinase 7 (CDK7), cyclin H, and MAT1. XPB interacts with the p44 core-TFIIH subunit and with MAT1, a subunit of CAK involved in the regulation of CDK 7 transcription activity<sup>12-15</sup>.

TFIIH was first identified as a basal transcription factor and was subsequently shown to play a key role in DNA repair. The XPB helicase is involved in promoter opening during transcription initiation while XPB allows strand separation around the DNA lesion in the

context of DNA repair by NER<sup>16,17</sup>. Biochemical and genetic studies have shown the need of both XPB and XPD ATPase activities to open up DNA around a damaged site<sup>18,19</sup>. Recent data showed that only the ATPase activity of XPB was required for opening and remodeling of DNA in NER and transcription, while its helicase would be devoted to promoter escape in transcription<sup>20,21</sup>. On the contrary, the XPD helicase activity plays a minor role in transcription, but is necessary for NER<sup>19,22</sup>. Once recruited to DNA damage/XPC:HR23B complex, CAK dissociates from core-TFIIH and XPD upon arrival of XPA and other NER factors, thus promoting incision/excision of the damaged oligonucleotide and repair of the DNA<sup>23</sup>. The CDK7 subunit of TFIIH phosphorylates residues Ser5 and Ser7 from the C-terminal domain (CTD) of the rpb1 RNA Polymerase II (RNA Pol II) subunit and functions in promoter-proximal pausing and termination<sup>24-26</sup>. In addition to its role in basal transcription, CDK7 also participates in the transactivation of several hormone-dependent genes by phosphorylating nuclear receptors (NR) including retinoic acid receptors<sup>27,28</sup>, the estrogen receptor<sup>29</sup>, peroxysome proliferator-activated receptor<sup>30</sup> or the androgen receptor<sup>31</sup>. The importance of NR phosphorylation by TFIIH in transcription has been highlighted in studies using cells from patients bearing mutations in the XPD subunit of TFIIH or in the XPG endonuclease that established the consequence of hormonal/transcriptional dysfunctions in XP-CS/TTD phenotypes. In patient cell lines, NRs are hypophosphorylated and the ligand-dependent response is decreased<sup>28,32</sup>.

Most mutations in TFIIH which originate XP, TTD and XP/CS both disturb regulatory interactions between TFIIH components and affect the catalytic activities of the complex. For example, mutations in p8/TTD-A weaken interactions with the p52 core-TFIIH subunit, leading to a reduced intracellular TFIIH concentration and a defect in NER, a common feature of TTD cells<sup>33,34</sup>. The F99S mutation in XPB weakens the XPB/p52 interaction and thus the resulting decrease in the XPB ATPase activity<sup>21</sup>. Similarly, R683W and R772W mutations

found in XPD patients weaken its interaction with p44, another TFIIH subunit, and consequently disrupt its helicase activation function<sup>35,36</sup>.

We have combined *in vitro* reconstituted assays with cell-based approaches to provide new insights on structure/function relations for XPD and on the architecture of TFIIH. Based on a detailed characterization of mutations identified in cell lines derived from patients suffering from TTD, we show that the previously uncharacterized ARCH domain of XPD not only plays a key role for DNA recognition and for the strand-displacement activity of the helicase but have also identified this domain as platform for the recruitment of CAK and thus the regulation of TFIIH transcription and NER activities.

## **Results**

### **Impaired *in vitro* NER of reconstituted mutant TFIIH**

To address the role of the ARCH domain of human XPD we take advantage of the C259Y mutation, which is to our knowledge the only one identified in this domain and which is found associated with the R772W allele in patients TTD12PV and TTD15PV (**Fig. 1a**). Fibroblasts from these patients show both a drastic reduction of the capacity to perform DNA-repair synthesis with UDS levels that are 20% of normal and poor survival after UV irradiation<sup>37</sup>.

We first examined the consequences of the C259Y mutation on TFIIH composition and activities. Recombinant complexes TFIIH (rIIH), resulting from coinfection by baculoviruses expressing either wild-type (XPD-wt) or mutant proteins (XPD-C259Y and XPD-R772W) were produced in insect cells. *Sf21* cells were co-infected with a first virus for expression of the six core-TFIIH subunits, a second one for the three CAK subunits and a last one for production of XPD. Recombinant complexes were immunopurified from infected cell extracts

using an antibody directed against the p44 core-TFIIH subunit under physiological salt conditions (150 mM KCl) and after extensive washing at the same ionic strength, bound complexes were eluted with a competitor peptide. Western Blot analysis of immuno-purified complexes shows that the C259Y mutation does not affect the composition and stoichiometry of rIIH XPD-C259Y, which is comparable to that of rIIH XPD-wt or rIIH XPD-R722W: Immunopurification by the p44 antibody precipitated not only the p44, XPB and p8 subunits of core-TFIIH but also the XPD as well as CDK7, a component of the CAK subcomplex (**Fig. 1b**, compare lanes 1, 5 and 6). For further insights, we also engineered an XPD variant (XPD- $\Delta$ ARCH) in which residues 248-438 that correspond to the entire ARCH domain were deleted and replaced by a short linker peptide. Deletion of the ARCH domain however has no visible effect of the composition of TFIIH (lane 4). As controls, we also analyzed XPD-Ins199 and XPD-Q452X two mutant forms located on each side of the ARCH domain<sup>36</sup>. In rIIH XPD-Q452X, where the XPD ARCH domain is present but the C-terminal domain lacking, the level of CDK7 is similar to that observed in the control (compare lanes 1 and 3). The frameshift at position 199 in XPD-Ins199 leads to a protein lacking the ARCH domain and the C-terminal part of the protein. In rIIH XPD-Ins199, XPD as well as CDK7, a component of CAK are sub- stoichiometric while XPB and p8 are not affected (compare lanes 1 and 2), suggesting that such a mutation strongly impairs association with core-TFIIH and CAK anchoring. This does not exclude the influence of the other subunits in maintaining the compact architecture of TFIIH.



### **The ARCH domain and the helicase activity of XPD**

We tested the different rIIHs recombinant complexes in a dual incision assay that contains XPC-HR23b, XPA, RPA, XPG, ERCC1-XPF factors and a closed-circular plasmid with a single 1,3-intra-strand d(GpTpG) cisplatin-DNA crosslink as template<sup>35</sup>. Both deletion of the ARCH domain and the C259Y mutation impair incision: While rIIH XPD- $\Delta$ ARCH is totally inactive (**Fig. 2a**, lanes 7-8), low but significant residual incision activity is detectable with rIIH XPD-C259Y (lanes 3-4), suggesting that the C259Y is a causative mutation. Note that rIIH XPD-R722W, the complex harboring the mutation found in the second allele in TTD12PV and TTD15PV patients, is unable to potentiate dual incision of the damaged DNA (lanes 5-6).

We then evaluated the helicase 5' 3' activity of purified XPD using a double stranded oligonucleotide substrate with a 5' single strand extension (**Fig.2b**). Experiments were performed in presence of the p44 TFIIH core subunit, known to specifically interact with and regulate the XPD helicase activity<sup>38</sup>. In contrast to XPD-wt which is able to displace the 25 bp P<sup>32</sup>-labelled oligonucleotide in the presence of p44, neither XPD-C259Y nor XPD- $\Delta$ ARCH exhibit detectable helicase activity, even upon addition of a large excess of p44 (compare lanes 4-6 and lanes 7-9 to lanes 1-3). We next investigated interactions with the substrate DNA using an electrophoretic mobility shift assay (EMSA) and tested whether mutations in the ARCH domain impair DNA binding. While the inactive XPD-R722W mutant has partially retained the capacity to form a stable complex with the duplex oligonucleotide, (**Fig. 2c**, compare lanes 6-7 to lanes 2-3), neither XPD-C259Y nor XPD- $\Delta$ ARCH are able to shift the helicase DNA substrate (compare lanes 4-5 and lanes 8-9). As TFIIH subunits are involved in an intricate protein-protein interaction network that is likely to modulate its catalytic activities, we also analyzed the 5' 3' helicase activity of XPD when associated to core-TFIIH and in the context of Holo-TFIIH (that is composed of core-TFIIH, XPD and CAK). To our

surprise, the 5' 3' helicase activity of Holo-TFIIH was reduced by an order of magnitude when compared to that of core-TFIIH (**Fig. 2d**, compare lane 1 to lane 4). However, and as shown above for the isolated proteins, none of the mutant complex exhibited detectable helicase activity (**Fig 2d**, lower panel, compare lane 1 to lanes 2-3 and lane 4 to lanes 5-6). Altogether, these results show that the ARCH domain is required for the unwinding reaction and is involved in recognition of the DNA substrate. Comparison of the helicase activity of core-TFIIH to that of Holo-TFIIH suggests that the presence of CAK represses the XPD helicase activity. In agreement with this observation, gel shift experiments performed with CAK/XPD instead of XPD do not lead to the formation of a stable nucleoprotein complex detectable under our experimental conditions (**Fig. 2c** compare lanes 11-12 to 13-14).

#### **XPD mutations impair *in vitro* transcription activity of TFIIH**

The rIIHs complexes were next tested for their transcription activity using an *in vitro* reconstituted assay in the presence of recombinant TBP, TFIIA, TFIIB, TFIIE and TFIIF in addition to purified RNA Pol II and a linearized DNA template containing the adenovirus major late promoter. rIIH XPD-C259Y and rIIH XPD-R722W significantly impair RNA synthesis relative to wild type (**Fig. 3a**, compare lane 1-3 to 4-6 and 7-9).

Phosphorylation of RNA Pol II CTD by CAK being required for the transition between initiation and elongation and further promoter escape<sup>39,40</sup>, we investigated how mutation in XPD impact on the phosphorylation status of the polymerase. When added to an *in vitro* transcription system that contains all the basal transcription factors in addition to core-TFIIH, CAK and XPD-wt results in a shift from the lower molecular weight RNA Pol IIA (the non phosphorylated form) to higher molecular weight RNA Pol IIO (the hyperphosphorylated form; **Fig. 3b**, lanes 1-3). Interestingly, no phosphorylation of RNA Pol II CTD is observed in

the presence of XPD-C259Y or XPD-722W (**Fig. 3b**, compare lanes 2-3 to 4-5 and 6-7), while TFIIH containing the XPD-R48K (that carries a mutation in the helicase motif I) allow CTD phosphorylation<sup>22</sup> (lanes 10-11). We then wondered whether an excess of CAK could enhance the transcription activity of recombinant TFIIH harboring the XPD mutations and added increasing amounts of CAK to an *in vitro* transcription system reconstituted with purified recombinant TFIIH. Addition of CAK to recombinant rIIH XPD-C259Y leads to a 2-fold stimulation and a partial restoration of transcription activity (Fig. 3c, lanes 2-4) suggesting that the XPD-C259Y disturbs the equilibrium between core-TFIIH and core-TFIIH plus CAK but does not lead to a totally inactive conformation. In contrast, the effects of the R722W point mutation, which impairs interaction with the p44 core-TFIIH subunit<sup>38</sup> as well as deletion of the entire ARCH domain, which strongly affects transcription (**Fig 3a**, lanes 7-12) and CTD phosphorylation (**Fig 3b**, lanes 5-9) could not be compensated by an excess of CAK (**Fig. 3c**, lanes 5-10). Altogether, the above results have established that the C259Y mutation impairs TFIIH basal transcription initiation.

### **The ARCH domain is a platform for CAK anchoring**

To further provide a mechanistic explanation for the rIIH XPD C259Y transcription defect, we first tested pair-wise interactions between either mutant XPDs and p44 from core-TFIIH or XPDs with MAT1 from CAK, the two TFIIH subunits that tightly interact with the helicase<sup>12,38</sup>. Proteins were co-expressed in insect cells, immunoprecipitated in 250 mM NaCl using the Flag-affinity tag fused to the N-terminal extremity of XPD, eluted with competitor peptide in native conditions and analyzed by Western Blot. Mutation C259Y in the ARCH domain as well as deletion of the entire module have no visible effect on association of XPD with p44 (**Fig. 4a**, compare lanes 6 and 8 to lane 5) but strongly affect the MAT1/XPD interaction (compare lanes 2 and 4 to lane 1), suggesting that the ARCH domain is involved in

the anchoring of CAK to XPD. Point mutation R772W in contrast impairs the p44/XPD interaction (compare lane 7 to lane 5) but does not modify of its interaction with MAT1 (compare lane 3 to lane 1). In a second set of experiments, we analyzed interactions between XPD variants and the XPG endonuclease that strongly associates with TFIIH and stabilizes the interaction between core-TFIIH, CAK and XPD<sup>32</sup>. Pull down from whole cell extracts of *Sf9* cells infected by viruses expressing XPG and XPD showed that mutations in the ARCH domain of XPD did not modify association with XPG (**Fig. 4b**, compare lane 1 to lanes 2 and 4), suggesting that the ARCH domain is not critical for XPG binding.

Next, to further characterize the CAK/XPD interface, XPD deletion fragments corresponding either to the N-terminal region (residues 1-277), the ARCH domain (residues 245-443) or the C-terminal region (residues 444-760) of XPD (**Fig. 1a**) were co-expressed with the three subunits of CAK in insect cells. Complexes were affinity purified using either the Flag-tag fused to the N-terminus of XPD constructs or the C-terminal strep-tag from the CDK7 component of CAK. Both FLAG-XPD and CDK7-strep pull down experiments followed by coomassie staining show that the XPD fragment corresponding to the ARCH domain is able to co-immunoprecipitate with the three subunits of CAK (**Fig. 4c**, lanes 3 and 6). This is not the case of fragments corresponding to the N- and C-terminal regions of XPD which do not associate with CAK (**Fig. 4c**, lanes 1-2 and 4-5), showing that this module is indeed required for stable association with CAK and can be considered as a recruitment platform.

To investigate consequences of XPDs mutations on association with CAK and core-TFIIH, recombinant complexes were co-expressed in insect cells and immunopurified using the epitope Flag fused to the N-terminus of XPD in 250 mM NaCl to eliminate weak and non specific interactions. Analysis of CDK7 kinase activity using a peptide that mimics the RNA Pol II CTD shows that the immunopurified complexes rIIH XPD-C259Y and rIIH XPD- $\Delta$ ARCH are unable to phosphorylate their substrate because they contain only limited

amounts of active CAK (**Fig. 4d**, compare lanes 1 and 3 to lanes 2 and 4). This is in perfect agreement with a critical role of the ARCH domain as a recruitment platform for CAK. Western Blot analysis shows that these immunoprecipitated complexes contain detectable amounts of p52 a core-TFIID subunit, which is not the case of the complex harboring the R722W mutation. Surprisingly, a quantitative interpretation of Western Blot experiments shows that mutations in the ARCH domain and particularly the deletion of the entire domain not only affect association with CAK but also with the p44 core-TFIID subunit (**Fig. 4d**, histogram). Together, these observations establish the role of the ARCH domain as a platform for recruitment of CAK but also highlighting the intricacy of the protein-protein interaction network within TFIID.

#### **Incidence of XPD mutations on transactivation mediated by nuclear receptors.**

Knowing that mutations in XPD can affect ligand-dependent transactivation mediated by nuclear hormone receptors<sup>28</sup>, we tested the effect of the C259Y mutation on the capacity of TFIID to transactivate the RAR $\beta$ 2 gene in response to all-*trans* retinoic acid (t-RA). The established cell line HD2 harboring the causative XPD-R683 point mutation<sup>41</sup> was transfected with a plasmid expressing either wild-type (XPD-wt), XPD-C259Y or XPD-R722W, treated t-RA and the concentration of RAR $\beta$ 2 mRNA was quantified by quantitative PCR (Q-PCR) (**Fig. 5a**). Overexpression of XPD-wt allows a significant transcriptional activation upon ligand induction compared to HD2 control. This correction is specific of XPD-wt and neither XPD-C259Y nor XPD-R722W was able to significantly increase the level of RAR $\beta$ 2 mRNA, demonstrating that although not lethal, the C259Y mutation strongly affects TFIID function *in vivo*.

Chromatin immunoprecipitation (ChIP) assays were then undertaken to investigate whether mutations affect recruitment of TFIID on activated RAR $\beta$ 2 promoter using isogenic stable

cell lines engineered by site-specific integration of transgenes into HEK293 cells using an FLP-FRT system. These cells express in addition to endogenous XPD, either a wild type (XPD-wt) or a mutant (XPD-C259Y and XPD-R722W) protein fused to an amino terminal Flag peptide (**Fig. 5b**). The genomic DNA fragments bound to XPD were immunoprecipitated with antibodies directed against either the Flag epitope or the XPD protein and were further analyzed by Q-PCR. After 6 h of treatment with t-RA and synthesis of RAR $\beta$ 2mRNA<sup>42</sup>, the recruitment of TFIID containing the Flag XPD at the RAR $\beta$ 2 promoter was evaluated. Using antibodies directed against either Flag XPD or the XPD subunit itself demonstrated that both integrate TFIID and are recruited together with RNA Pol II at the RAR $\beta$ 2 promoter in XPD-wt/HEK293 cells (**Fig. 5c and 5f**); ChIP/reChIP further confirm that the immunoprecipitated Flag XPD-wt complex recruited at the RAR $\beta$ 2 promoter also contains MAT1 and p44 TFIID subunits (**Fig. 5f**). We next observed that mutants XPD-C259Y and XPD-R722W are not efficiently recruited at the RAR $\beta$ 2 promoter (**Fig. 5d and 5e**) explaining the inability of this mutated protein to restore the RAR $\beta$ 2 expression (**Fig. 5a**). In such case, only XPD-wt is found at the RAR $\beta$ 2 promoter (**Fig. 5d and 5e**). We also should notice that the Flag XPD-C259Y and -R722W immunoprecipitated complexes are not integrated *in vivo* with the other TFIID subunits (**Fig. 5g and 5h**).

All together our data show that the XPD-C259Y mutation is detrimental for RNA synthesis; this results from its inability to integrate in optimal conditions the TFIID complex and by extension the formation of the transactivation complex.

## Discussion

Helicases are modular proteins that use ATP to bind and/or remodel nucleic acid complexes; Involved in virtually all aspects of RNA and DNA metabolism, these motor proteins are built around a conserved helicase core consisting of two similar RecA like motor domains with additional accessory domains: N- or C-flanking regions or inserts within the core. Specific regulation or complementary catalytic activities are often provided by interactions with protein partners, and several helicases are components of large macromolecular complexes.

## Structural considerations

The XPD helicase illustrates the molecular organization of these motor proteins and the key role of accessory domains for regulation. Apart from its motor domains HD1 and HD2 human XPD is composed of a 4FeS cluster domain, the ARCH domain and a specific C-terminal extension required for association with the p44 core-TFIIH subunit<sup>5-7</sup> (**Fig. 6a**). The 4FeS cluster domain contains four redox-sensitive cysteines that are thought to play a role in DNA damage recognition<sup>43</sup> and together with the ARCH domain forms a small tunnel, large enough to accommodate single stranded DNA<sup>5-7,44</sup>. We have shown that the ARCH domain of human XPD constitutes a platform for the recruitment of the kinase module of TFIIH and that this module specifically interacts with the MAT1 subunit of CAK. In this work, we largely took advantage of the C259Y mutation which impedes the MAT1/XPD interface. Residue C259 maps onto the first helix of the ARCH domain and its side chain points into the center of the helical bundle where it packs with other residues that form the core of the domain (**Fig 6a**). Mutation of C259 into a tyrosine probably destabilizes or unfolds the protein locally, which is consistent with the reduced thermal stability of a mutant XPD homologue from *Sulfolobus Tokodai*. We have shown that mutation C259Y in human XPD affects the XPD/MAT1 interface and thus the stability of mutant TFIIH. In TTD12PV fibroblasts, this mutation is



found associated with the R772W allele that leads to a destabilization of the XPD/p44 interaction<sup>38</sup> and as a consequence also impacts on the stability of TFIIH, explaining the low intracellular concentration of TFIIH measured, a common feature of TTD derived cell lines<sup>37,45,46</sup>.

### **Functional implications**

The C259Y mutation does not drastically modify the composition of recombinant TFIIH, when purified under physiological conditions, but both *in vitro* NER and transcription activities are significantly reduced. We have shown that this mutation directly impairs DNA binding and the capacity of human XPD to unwind a DNA substrate, explaining the observed low NER activity of a mutant TFIIH harboring the C259Y mutation. In favor of a direct implication of the ARCH domain in DNA recognition, mutagenesis and structural data obtained on archeobacterial XPD, suggest that the translocated DNA strand is protruding through the pore formed by HD1, the iron–sulphur cluster and the ARCH domain<sup>44</sup>. Mutation of the residue corresponding to C259 in human affects the stability of XPD from *sulfolobus Tokodai* as well as its DNA dependent ATPase activity but not the helicase activity<sup>6</sup>.

The XPD helicase activity is not required for the transcription activity and the limited capacity of TFIIH containing mutant XPD-C259Y to stimulate transcription/trans-activation probably result from the perturbed interaction of XPD with CAK. We can reasonably speculate that this mutation leads to an inactive conformation and to the mis-positioning of TFIIH within transcription/trans-activation complexes. This can explain the inability of the CDK7 subunit of CAK to phosphorylate the CTD of RNA Pol II (and thus the transcription defect observed *in vitro*) as well as the absence of TFIIH recruitment on the RAR $\beta$ 2 promoter in response to t-RA. Altogether, these data show that the ARCH domain of XPD not only

plays a direct role in catalysis but also constitute a platform for the recruitment of CAK, and is thus essential for transcription initiation and its regulation.

### **A negative regulation mechanism**

XPB, XPD and CDK7 exert their enzymatic activities in a complex network of interactions essential to ensure the correct functioning of TFIIH in transcription and repair. Fine tuning mechanisms exerted by core-TFIIH subunits have been proposed to explain regulation of XPB ATPase and XPD helicase activities (**Fig 6b**). Interactions between XPB, p52 and p8/TTD-A stimulate XPB ATPase activity and consequently favor binding of TFIIH to damaged DNA<sup>21</sup>. Interactions between p44 and XPD stimulate the helicase activity<sup>35,38</sup>, allowing unwinding of DNA around the lesion and subsequent double incision by the endonucleases XPF-ERCC1 and XPG. Our data provide evidences for a negative regulation mechanism that involve the CAK module of TFIIH and a cross-talk between core-TFIIH, XPD and CAK. Following previous observation with isolated subunits<sup>13</sup>, we have compared the 5' 3' catalytic activities of wild type core- and Holo-TFIIH complexes and have shown that the XPD helicase activity which is dispensable for transcription initiation but required for NER<sup>19,22</sup>, is repressed when CAK is associated with TFIIH. Underlining the crucial role of the ARCH platform domain in the anchoring of CAK, it was recently shown that NER is driven by dissociation of CAK from core-TFIIH, a mechanism that allows TFIIH to switch from its transcription to its repair function<sup>23</sup>. In favor of such a model where CAK negatively regulates NER, CDK7 has also been shown to inhibit repair by phosphorylation of one or more NER factor<sup>47</sup>.

### **Non-TFIIH XPD complexes**

The XPD helicase functions as a component of TFIIH but can also be found associated with non-TFIIH complexes which play roles in processes other than transcription and DNA repair. The MMXD complex is composed of XPD, MMS19, MIP18, Ciao1, and ANT2 and is involved in chromosome segregation and nuclear shape formation<sup>48</sup>. XPD also associates with CAK in absence of the other TFIIH subunits, regulating the localization of CAK on its substrate and therefore its activity<sup>49,50</sup>. The XPD/CAK complex plays a role in coordination and progression of mitosis during late division steps<sup>50,51</sup>. How do mutations in XPD impact on the functions of the different non-TFIIH complexes. Identification of XPD ARCH domain as the minimal region required for CAK binding suggests that mutations in ARCH such as the mutation C259Y which impair association between CAK and XPD should result in increased CAK activity, increased cell growth and proliferation in drosophila embryos with new insights to our understanding of XP, XP/CS and TTD.

In conclusion, the present study underlines the essential roles of regulatory domains from motor proteins in the context of transcription regulation and maintenance genome integrity. Using the human XPD helicase, a subunit of the transcription/DNA repair factor TFIIH, we have shown that its ARCH domain is not only involved in DNA binding and catalysis but also constitutes an essential component of the intricate protein-protein interaction network that enables assembly of functional transcription/trans-activation complexes. Introduction of a TTD mutation identified in a TTD patient impairs the helicase activity of XPD and the NER activity of TFIIH. We have also demonstrated that the ARCH domain is critical for the recruitment of the CAK kinase module of TFIIH and alteration of the interface with the MAT1 CAK subunit both decreases the *in vitro* basal transcription activity of TFIIH and impedes an efficient recruitment of the trans-activation complex on the promoter of the

activated RAR $\beta$ 2 gene. Interestingly, comparison of the 5' 3' catalytic activities of wild type core- and Holo-TFIIH complexes showed that XPD is repressed when CAK is associated with TFIIH, suggesting that the ARCH domain also participates in the fine tuning of TFIIH enzymatic activities.

## Methods

### Construction of baculoviruses, protein production and purification

The cDNAs encoding XPD wild type (residues 1-762), XPD variants (K48R, C259Y, R722W, ins199PP, Q452X, N-ter(1-245), C-ter(443-762), ARCH (245-443) and  $\Delta$ ARCH) and XPG were cloned into pAC8F<sup>52</sup>. Variant XPD- $\Delta$ ARCH, designed on the basis of the crystal structures of archeal XPD homologues<sup>4,7</sup> and analysis of multiple sequence alignments, results from deletion of residues 248-438 of human XPD that were replaced by the hexapeptide SGASAS. The mutations generating a frame shift after 2-aa PP insertion at position 199 (ins199PP) and the premature stop codon at position 452 (Q452X) have been described previously<sup>36</sup>. All resulting transfer vectors were recombined with baculovirus DNA (BaculoGold DNA, Pharmingen™) in *Sf9* cells to generate viruses for production of proteins in fusion with the Flag peptide (DYKDDDDK). In the case of XPG, the protein is also fused to a C-terminal c-Myc peptide (AEEQKLISEEDLLRKRREQLKHKLE). Baculoviruses overexpressing p44 and MAT1 were as previously described<sup>22</sup>. Two multigene expressing viruses were used: the first (VCAK) allows expression of CDK7 with a carboxy-terminal strep tag II together with Cyclin H and MAT1<sup>53</sup>; the second (Vcore), co-expression of the core-TFIIH subunits p8, p34, p44, p52, p62 and XPB. Vcore was generated from a transfer vector resulting from Cre/LoxP fusion of three vectors<sup>54</sup> (i) pFL in which XPB, p34 and p44

were cloned (ii) pUCDM in which p52 and p62 were cloned (iii) pSPL in which p8 was cloned.

For production of proteins and complexes, *Sf21* insect cells grown in suspension (typically 250 ml) were infected/co-infected with the appropriate viruses/combinations of viruses, collected 48 hours post infection and resuspended in 10ml of buffer A (20 mM Tris-HCl, pH 8.0, 150 mM KCl, 1mM DTT) supplied with Complete protease inhibitor cocktail (Roche™). Cells were disrupted using a Vibracell sonicator (3mm probe at 20% intensity for 30 seconds) and after centrifugation at 14,000g for 30 min at 4°C, the lysate was incubated for 2h with 200 µl of protein A sepharose beads crosslinked to the M2 anti-Flag antibody (SIGMA™) for purification of Flag-XPD, to the 1H5 anti-p44 antibody (directed against residues 1-17 of Human p44) for purification of core-IIH and rIIH or with Streptactin sepharose (IBA™) for purification of CAK. After extensive washing in buffer A and equilibration in buffer B (50 mM Tris-HCl pH 8.0, 75 mM KCl, 20% glycerol, 0.1% NP-40, 1mM DTT), proteins were eluted by competition with 2 CV of buffer B containing the appropriate synthetic peptide at 0.5 mg/ml for immunoprecipitations or 2.5 mM d-desthiobiotin for elution from Streptactin sepharose (IBA™).

Monoclonal antibodies against the TFIID subunits XPB (1B3), XPD (2F6), p52 (1D11), p44 (1H5), CDK7 (2F8), MAT1 (2D3), p8 (1D1), XPG (1B5) were from IGBMC's facilities.

### ***In vitro* transcription and phosphorylation assays**

Run-off transcription assays was performed using recombinant TFIIB, TFIIE, TFIIF, TBP, endogenous RNA Pol II, and the different rIIHs as previously described<sup>22,55</sup>. RNA Pol II phosphorylation was carried out as a classical runoff transcription except that ATP was added

to a final concentration of 5 mM, that the amount of purified RNA Pol II polymerase was adjusted (typically reduced by a factor 3) and that rIIH was substituted by a mixture purified core-IIH, CAK and XPD, which allows to prepare a pre-mix containing all components but XPD variants. Hypo (IIA) and hyper (IIO) phosphorylated forms of RNA Pol II were resolved on a 6% SDS-PAGE and detected by Western Blot using the monoclonal antibody (7C2) directed against the CTD.

To analyze the CDK7 kinase activity, purified complexes (typically 5 $\mu$ l) were added to a reaction mixture containing 1  $\mu$ g of bacterially expressed GST-fused human CTD in 20mM Tris-HCl pH 7.5, 250mM NaCl, 10mM MgCl<sub>2</sub>, 20% glycerol and 0.005  $\mu$ M 0.5  $\mu$ Ci [ $\gamma$ -<sup>32</sup>P] in a total reaction volume of 20 $\mu$ l for 30' at 25 °C. Reactions were stopped by the addition of 5 $\mu$ l concentrated Laemmli sample buffer and the amount of phosphate incorporated is analyzed by autoradiography after separation on a 12 % SDS- polyacrylamide gel.

### **Dual incision, helicase and gel shift experiments**

Nucleotide excision repair dual incision assays were performed using a plasmid with a single a single 1,3-intrastrand d(GpTpG) (50ng) in a buffer containing 50 mM Hepes-KOH (pH 7.8), 5 mM MgCl<sub>2</sub>, 1 mM DTT, 0.3 mM EDTA, 10% glycerol, 2.5  $\mu$ g BSA, 50 mM KCl, and 2 mM ATP. Reactions mixes (10  $\mu$ l) containing XPG (5 ng), XPF/ERCC1 (15 ng), XPC/hHR23B (10 ng), RPA (50 ng), XPA (25 ng), and increasing amount of rIIH complexes were incubated at 30 C for 90 minutes and analyzed as previously described<sup>35</sup>.

DNA unwinding reaction substrate was prepared by annealing two oligonucleotides: AL (TTCGAGCACCGCTGCGGCTGCACCGGC) and BC (GCCGGTGCAGCCGCAGCGGTGCTCG). Prior to annealing, BC is labeled using [ $\gamma$ -<sup>32</sup>P] ATP and T4 polynucleotide kinase (New England Biolabs™) and purified using Micro Bio-

Spin clean-up columns (Biorad™). Reaction was performed during 40min at 25°C by adding the immuno-purified helicase to the DNA probe at 10 nM 20 mM Tris/HCl (pH 8.0), 75 mM KCl, 4mM MgCl<sub>2</sub>, 1 mM DTT, 4mM ATP and 0.1 mg/ml BSA with a total reaction volume of 20 µl. Reaction was stopped by adding 20mM EDTA, 14% glycerol, 0.2% SDS and 0.028% bromophenol to the reaction mixture. Analyses were performed by migration in 14 % polyacrylamide gel (acrylamide/bis-acrylamide ratio: 33/1) and autoradiography.

The helicase assay probe described above was used for gel shift experiments. Increasing amounts (2.5 and 10 µl) of immuno-purified XPD variants or TFIIH complexes were incubated with the probe in absence of ATP during 15 min at 25°C in 20 mM Tris/HCl (pH 8.0), 75 mM KCl, 4mM MgCl<sub>2</sub>, 1 mM DTT, 0.1 mg/ml BSA. Analyses were performed by migration in 6% polyacrylamide gel (acrylamide/bis-acrylamide ratio: 29/1) and autoradiography.

### **Protein interaction assays**

Wild-type or mutated XPDs complexes were co-expressed by co-infecting *Sf9* insect cells monolayers (typically 20 10<sup>6</sup> cells) with the appropriate baculoviruses, resuspended in 1ml of buffer A and cell extracts were prepared as described above. Pull down were performed in buffer C (50 mM Tris-HCl pH 8.0, 250 mM NaCl, 0.1% NP-40, 1mM DTT) using 30 µl of Protein A sepharose cross-linked with M2 anti-Flag (SIGMA™) or 9E10 anti c-Myc antibodies or Streptactin sepharose (IBA™). Pulled down complexes were eluted by competition in non denaturing conditions, analyzed on SDS PAGE followed by coomasie staining or Western Blot experiments.



## **Cell culture and Reagents**

Human fibroblasts HD2<sup>41</sup> cells derived from patients and HEK Flp-In cells (Invitrogen™) were cultured in DMEM with 4,5g/L glucose and 862mg/L Glutamax-I supplemented with Penicillin (100UI/ml), Streptomycin (100μg/ml) and 10% Fetal Calf Serum (FCS) at 37°C in the presence of 5% CO<sub>2</sub>. HD2 cells were transfected with a pSG5 plasmid containing XPD-wt or mutants (-C259Y and -R722W) with Lipofectamine 2000 (Invitrogen™) according to manufacturer's instructions. Stable cells lines were obtained after transfection of HEK Flp-In cells with a pcDNA5FRT plasmid containing a 3Flag tagged version of XPD (wt or mutants), pOG44 and JetPEI according to manufacturer's instructions. After selection with Hygromycin (150μg/ml), cells were diluted to establish subclonal populations. For the time course, cells were incubated with red phenol-free medium containing 10% charcoal-treated fetal calf serum (FCS) and 40 mg/ml gentamycin one day before treatment with 10μM t-RA (Biomol™).

## **Reverse transcription and Quantitative PCR**

Total RNA was isolated using a GenElute Mammalian Total RNA Miniprep kit (Sigma™) and reverse transcribed with SuperScript II reverse transcriptase (Invitrogen™). The quantitative PCR was done using the Lightcycler 480 SYBR Green I Master and the Lightcycler 480 (Roche™). The primer sequences for RARβ2 and Glyceraldehyde 3-phosphate deshydrogenase (GAPDH) genes used in real-time PCR are available upon request. RARβ2 mRNA levels represent the ratio between values obtained from treated and untreated cells normalized against the housekeeping GAPDH mRNA.

## **Chromatin immunoprecipitation**

Cells were crosslinked at room temperature for 10min with 1% formaldehyde. Chromatin was prepared (Drane et al., 2004) and sonicated on ice 30min using a Bioruptor (Diagenode™). Samples were immunoprecipitated (IP) with antibodies at 4°C for 6h and protein G-Sepharose beads (Upstate™) were added, incubated overnight at 4°C and sequentially washed. For ChIP/ReChIP experiments, after the IP with Flag M2 Sepharose beads (Sigma™) and washes, protein-DNA complexes were eluted with a competitive peptide (PD157) and samples were immunoprecipitated with antibodies at 4°C overnight, and protein G Sepharose beads were added, incubated for 4 hr at 4°C, and sequentially washed. The complexes were eluted and DNA fragments were purified using QiAquick PCR purification kit (QIAGEN) and analyzed by real-time PCR using a set of primers targeting the transcription initiation start (from -80 to -30) of the RARβ2 promoter region.

## **Acknowledgements**

This work was funded by the CNRS, the INSERM, the Université de Strasbourg (UdS), the Alsace Region. This project benefited from grants from the Agence Nationale de la Recherche (ANR-08-GENO-042-02), the Institut National du Cancer (INCA-2008-041), the Association pour la Recherche sur le Cancer (ARC, subvention fixe), the European Commission (SPINE2 complexes, LSHG-CT-2006-031220 and a personal advanced grant to JME). W.A. was supported by the Ministère de l'Éducation Nationale de la Recherche et de Technologie (MENRT) and ARC, L.R. by La Ligue contre le Cancer and A.M-R by the Fondation pour la Recherche Médicale (FRM). Nathalie Troffer-Charlier and Isabelle Kolb-Cheynel (IGBMC Baculovirus Facility) are acknowledged for productions in insect cells, Jean-Marie Garnier for help with cloning, Emmanuel Compe, Anne Catherine Dock-Bregeon and Patrick Schultz for fruitful advices and constructive discussions.

## References

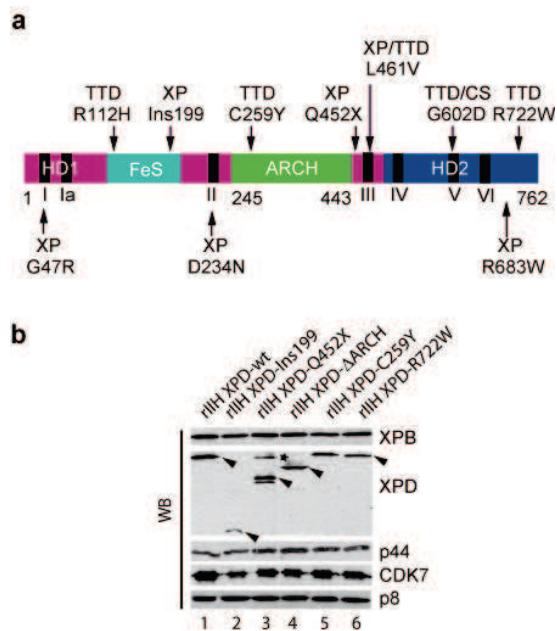
1. de Boer, J. & Hoeijmakers, J.H. Nucleotide excision repair and human syndromes. *Carcinogenesis* **21**, 453-60 (2000).
2. Lehmann, A.R. The xeroderma pigmentosum group D (XPD) gene: one gene, two functions, three diseases. *Genes Dev* **15**, 15-23 (2001).
3. Price, V.H., Odom, R.B., Ward, W.H. & Jones, F.T. Trichothiodystrophy: sulfur-deficient brittle hair as a marker for a neuroectodermal symptom complex. *Arch Dermatol* **116**, 1375-84 (1980).
4. Fairman-Williams, M.E., Guenther, U.P. & Jankowsky, E. SF1 and SF2 helicases: family matters. *Curr Opin Struct Biol* **20**, 313-24 (2010).
5. Fan, L. et al. XPD helicase structures and activities: insights into the cancer and aging phenotypes from XPD mutations. *Cell* **133**, 789-800 (2008).
6. Liu, H. et al. Structure of the DNA repair helicase XPD. *Cell* **133**, 801-12 (2008).
7. Wolski, S.C. et al. Crystal structure of the FeS cluster-containing nucleotide excision repair helicase XPD. *PLoS Biol* **6**, e149 (2008).
8. Compe, E. & Egly, J.M. TFIIH: when transcription met DNA repair. *Nat Rev Mol Cell Biol* **13**, 343-54 (2012).
9. Chang, W.H. & Kornberg, R.D. Electron crystal structure of the transcription factor and DNA repair complex, core TFIIH. *Cell* **102**, 609-13 (2000).
10. Gibbons, B.J. et al. Subunit architecture of general transcription factor TFIIH. *Proc Natl Acad Sci U S A* **109**, 1949-54 (2012).
11. Schultz, P. et al. Molecular structure of human TFIIH. *Cell* **102**, 599-607 (2000).
12. Busso, D. et al. Distinct regions of MAT1 regulate cdk7 kinase and TFIIH transcription activities. *J Biol Chem* **275**, 22815-23 (2000).
13. Sandrock, B. & Egly, J.M. A yeast four-hybrid system identifies Cdk-activating kinase as a regulator of the XPD helicase, a subunit of transcription factor IIH. *J Biol Chem* **276**, 35328-33 (2001).
14. Helenius, K. et al. Requirement of TFIIH kinase subunit Mat1 for RNA Pol II C-terminal domain Ser5 phosphorylation, transcription and mRNA turnover. *Nucleic Acids Res* **39**, 5025-35 (2011).
15. Helenius, K., Yang, Y., Alasaari, J. & Makela, T.P. Mat1 inhibits peroxisome proliferator-activated receptor gamma-mediated adipocyte differentiation. *Mol Cell Biol* **29**, 315-23 (2009).
16. Egly, J.M. & Coin, F. A history of TFIIH: two decades of molecular biology on a pivotal transcription/repair factor. *DNA Repair (Amst)* **10**, 714-21 (2011).
17. Fuss, J.O. & Tainer, J.A. XPB and XPD helicases in TFIIH orchestrate DNA duplex opening and damage verification to coordinate repair with transcription and cell cycle via CAK kinase. *DNA Repair (Amst)* **10**, 697-713 (2011).
18. Sung, P., Higgins, D., Prakash, L. & Prakash, S. Mutation of lysine-48 to arginine in the yeast RAD3 protein abolishes its ATPase and DNA helicase activities but not the ability to bind ATP. *EMBO J* **7**, 3263-9 (1988).
19. Guzder, S.N. et al. DNA repair gene RAD3 of *S. cerevisiae* is essential for transcription by RNA polymerase II. *Nature* **367**, 91-4 (1994).
20. Lin, Y.C., Choi, W.S. & Gralla, J.D. TFIIH XPB mutants suggest a unified bacterial-like mechanism for promoter opening but not escape. *Nat Struct Mol Biol* **12**, 603-7 (2005).
21. Coin, F., Oksenysh, V. & Egly, J.M. Distinct roles for the XPB/p52 and XPD/p44 subcomplexes of TFIIH in damaged DNA opening during nucleotide excision repair. *Mol Cell* **26**, 245-56 (2007).

22. Tirode, F., Busso, D., Coin, F. & Egly, J.M. Reconstitution of the transcription factor TFIIH: assignment of functions for the three enzymatic subunits, XPB, XPD, and cdk7. *Mol Cell* **3**, 87-95 (1999).
23. Coin, F. et al. Nucleotide excision repair driven by the dissociation of CAK from TFIIH. *Mol Cell* **31**, 9-20 (2008).
24. Lu, H., Zawel, L., Fisher, L., Egly, J.M. & Reinberg, D. Human general transcription factor IIH phosphorylates the C-terminal domain of RNA polymerase II. *Nature* **358**, 641-5 (1992).
25. Akhtar, M.S. et al. TFIIH kinase places bivalent marks on the carboxy-terminal domain of RNA polymerase II. *Mol Cell* **34**, 387-93 (2009).
26. Glover-Cutter, K. et al. TFIIH-associated Cdk7 kinase functions in phosphorylation of C-terminal domain Ser7 residues, promoter-proximal pausing, and termination by RNA polymerase II. *Mol Cell Biol* **29**, 5455-64 (2009).
27. Rochette-Egly, C., Adam, S., Rossignol, M., Egly, J.M. & Chambon, P. Stimulation of RAR alpha activation function AF-1 through binding to the general transcription factor TFIIH and phosphorylation by CDK7. *Cell* **90**, 97-107 (1997).
28. Keriell, A., Stary, A., Sarasin, A., Rochette-Egly, C. & Egly, J.M. XPD mutations prevent TFIIH-dependent transactivation by nuclear receptors and phosphorylation of RARalpha. *Cell* **109**, 125-35 (2002).
29. Chen, D. et al. Activation of estrogen receptor alpha by S118 phosphorylation involves a ligand-dependent interaction with TFIIH and participation of CDK7. *Mol Cell* **6**, 127-37 (2000).
30. Compe, E. et al. Dysregulation of the peroxisome proliferator-activated receptor target genes by XPD mutations. *Mol Cell Biol* **25**, 6065-76 (2005).
31. Chymkowitz, P., Le May, N., Charneau, P., Compe, E. & Egly, J.M. The phosphorylation of the androgen receptor by TFIIH directs the ubiquitin/proteasome process. *EMBO J* **30**, 468-79 (2011).
32. Ito, S. et al. XPG stabilizes TFIIH, allowing transactivation of nuclear receptors: implications for Cockayne syndrome in XP-G/CS patients. *Mol Cell* **26**, 231-43 (2007).
33. Giglia-Mari, G. et al. A new, tenth subunit of TFIIH is responsible for the DNA repair syndrome trichothiodystrophy group A. *Nat Genet* **36**, 714-9 (2004).
34. Kainov, D.E., Vitorino, M., Cavarelli, J., Poterszman, A. & Egly, J.M. Structural basis for group A trichothiodystrophy. *Nat Struct Mol Biol* **15**, 980-4 (2008).
35. Dubaele, S. et al. Basal transcription defect discriminates between xeroderma pigmentosum and trichothiodystrophy in XPD patients. *Mol Cell* **11**, 1635-46 (2003).
36. Ueda, T., Compe, E., Catez, P., Kraemer, K.H. & Egly, J.M. Both XPD alleles contribute to the phenotype of compound heterozygote xeroderma pigmentosum patients. *J Exp Med* **206**, 3031-46 (2009).
37. Botta, E. et al. Analysis of mutations in the XPD gene in Italian patients with trichothiodystrophy: site of mutation correlates with repair deficiency, but gene dosage appears to determine clinical severity. *Am J Hum Genet* **63**, 1036-48 (1998).
38. Coin, F. et al. Mutations in the XPD helicase gene result in XP and TTD phenotypes, preventing interaction between XPD and the p44 subunit of TFIIH. *Nat Genet* **20**, 184-8 (1998).
39. Dvir, A., Tan, S., Conaway, J.W. & Conaway, R.C. Promoter escape by RNA polymerase II. Formation of an escape-competent transcriptional intermediate is a prerequisite for exit of polymerase from the promoter. *J Biol Chem* **272**, 28175-8 (1997).
40. Akoulitchev, S., Makela, T.P., Weinberg, R.A. & Reinberg, D. Requirement for TFIIH kinase activity in transcription by RNA polymerase II. *Nature* **377**, 557-60 (1995).
41. Takayama, K. et al. Defects in the DNA repair and transcription gene ERCC2 in the cancer-prone disorder xeroderma pigmentosum group D. *Cancer Res* **55**, 5656-63 (1995).

42. Le May, N. et al. 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 (2010).
43. Mui, T.P., Fuss, J.O., Ishida, J.P., Tainer, J.A. & Barton, J.K. ATP-stimulated, DNA-mediated redox signaling by XPD, a DNA repair and transcription helicase. *J Am Chem Soc* **133**, 16378-81 (2011).
44. Kuper, J., Wolski, S.C., Michels, G. & Kisker, C. Functional and structural studies of the nucleotide excision repair helicase XPD suggest a polarity for DNA translocation. *EMBO J* **31**, 494-502 (2012).
45. Botta, E. et al. Reduced level of the repair/transcription factor TFIIH in trichothiodystrophy. *Hum Mol Genet* **11**, 2919-28 (2002).
46. Stefanini, M., Botta, E., Lanzafame, M. & Orioli, D. Trichothiodystrophy: from basic mechanisms to clinical implications. *DNA Repair (Amst)* **9**, 2-10 (2010).
47. Araujo, S.J. et al. Nucleotide excision repair of DNA with recombinant human proteins: definition of the minimal set of factors, active forms of TFIIH, and modulation by CAK. *Genes Dev* **14**, 349-59 (2000).
48. Ito, S. et al. MMXD, a TFIIH-independent XPD-MMS19 protein complex involved in chromosome segregation. *Mol Cell* **39**, 632-40 (2010).
49. Aguilar-Fuentes, J., Valadez-Graham, V., Reynaud, E. & Zurita, M. TFIIH trafficking and its nuclear assembly during early Drosophila embryo development. *J Cell Sci* **119**, 3866-75 (2006).
50. Chen, J., Larochelle, S., Li, X. & Suter, B. Xpd/Ercc2 regulates CAK activity and mitotic progression. *Nature* **424**, 228-32 (2003).
51. Li, X., Urwyler, O. & Suter, B. Drosophila Xpd regulates Cdk7 localization, mitotic kinase activity, spindle dynamics, and chromosome segregation. *PLoS Genet* **6**, e1000876 (2010).
52. Abdulrahman, W. et al. A set of baculovirus transfer vectors for screening of affinity tags and parallel expression strategies. *Anal Biochem* **385**, 383-5 (2009).
53. Fouillen, L. et al. Analysis of recombinant phosphoprotein complexes with complementary mass spectrometry approaches. *Anal Biochem* **407**, 34-43 (2010).
54. Fitzgerald, D.J. et al. Protein complex expression by using multigene baculoviral vectors. *Nat Methods* **3**, 1021-32 (2006).
55. Gerard, M. et al. Purification and interaction properties of the human RNA polymerase B(II) general transcription factor BTF2. *J Biol Chem* **266**, 20940-5 (1991).

## Figures

**Figure 1 TTD and XP mutations in the XPD TFIIH subunit.**

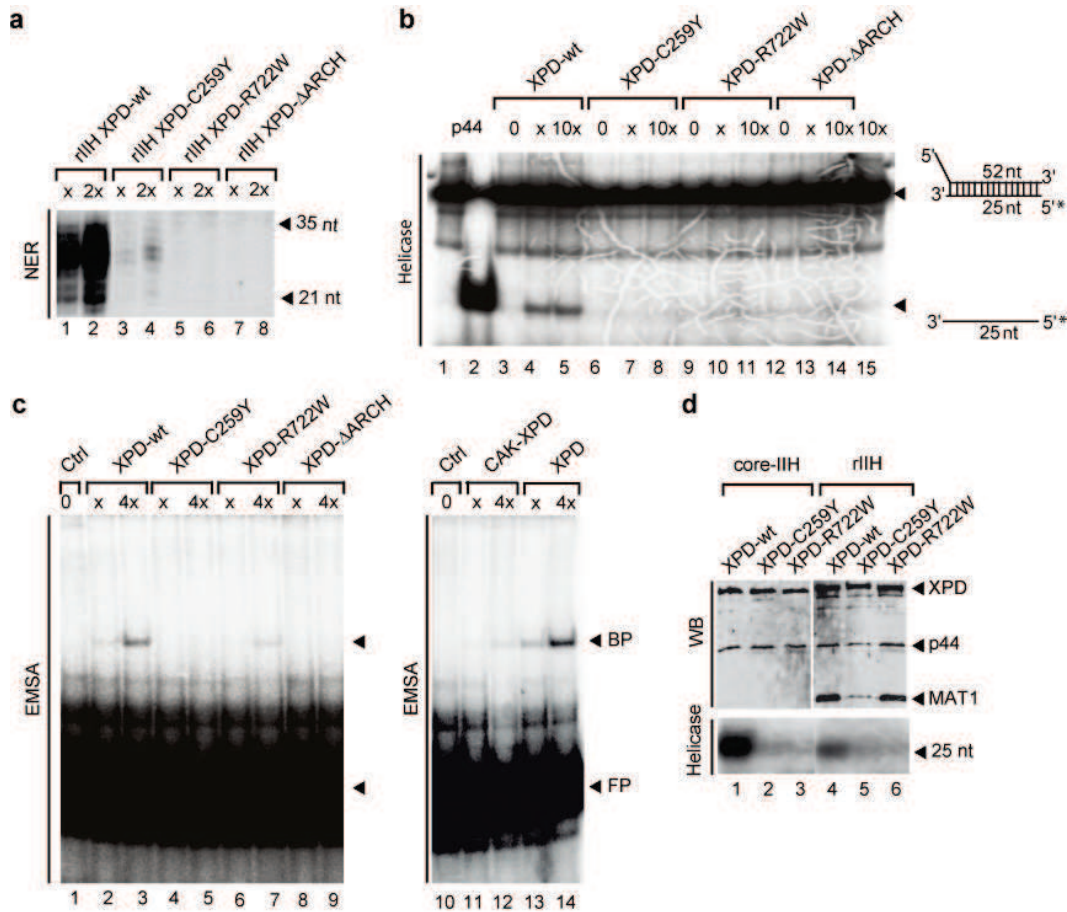


**a.** Schematic representation of XPD. Helicase motor domains HD1 and HD2 are in magenta and blue, respectively, the FeS iron sulfur containing domain in cyan and the ARCH domain in green. Black bars indicate the helicase motifs (I, Ia, II, IV, V and VI). Positions of mutations flanking the ARCH domain and associated diseases are shown. **b.** Production of recombinant TFIIH (rIIH). The 10 subunits of human TFIIH, including either wt or mutated XPD were co-expressed in insect cells using the baculovirus expression system, and complexes were immunoprecipitated using an antibody (Ab) directed toward the p44 subunit of the core TFIIH in low salt conditions (Buffer B containing 75mM KCl). After elution with a synthetic peptide recognized by Ab-p44, equal amounts of purified rIIHs were then

analyzed on a 12.5% SDS-PAGE followed by Western Blot analysis with Abs directed against either XPB, the amino-terminus of XPD, p44, CDK7 or p8. Arrows indicate the theoretical molecular weight of each XPD mutated form. \* indicate a non specific band.



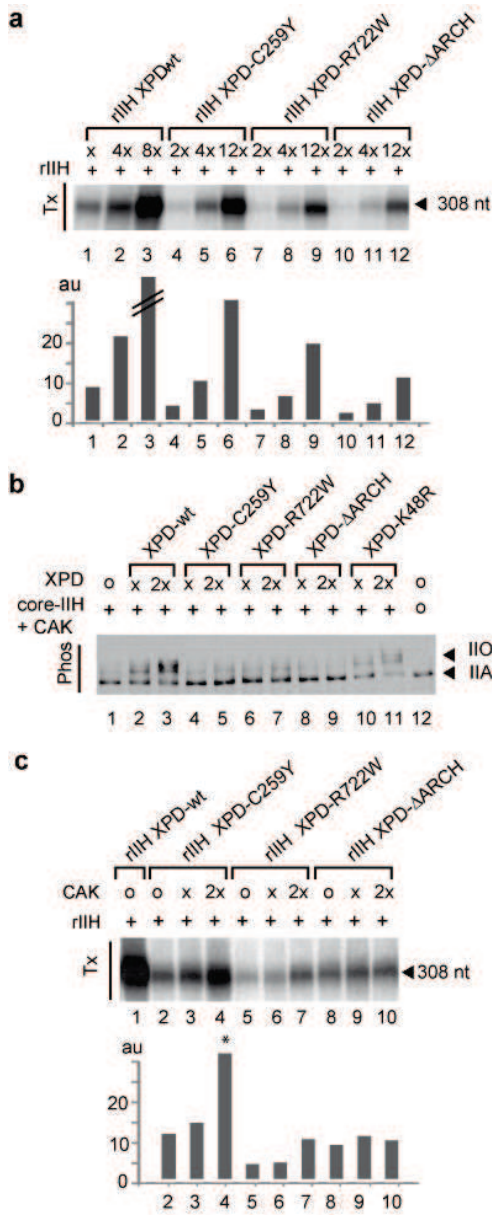
**Figure 2 DNA repair and helicase activities of rIIHs.**



**a.** *In vitro* double incision assay. Increasing amounts immuno-purified rIIHs were added to an incision/excision assay using recombinant NER factors. The reaction was analyzed by electrophoresis followed by auto-radiography. **b** 5'-3' helicase activity of XPD variants. Equivalent amounts (adjusted according to Western blot analysis) of each Flag-purified XPD variant was added to 5' strand extension probe obtained by annealing an 52 nt single strand DNA to a 5' P<sup>32</sup> labeled 25 nt single stranded DNA. Unwinding assay was performed in absence or in presence of increasing amounts of p44 (0, x or 10x). Single and double stranded DNA are separated by electrophoresis in 14% polyacrylamide gel and analyzed by autoradiography (lanes 3-15). Native and denatured probes have been loaded as controls (lanes 1 and 2, respectively) **c.** DNA binding activity of XPD variants. The capacity of XPD variants (left panel, lanes 1-9) and CAK-XPD (right panel, lanes 10-14) to bind to the 5' strand extension probe (shown in panel **b**) were compared by adding increasing equivalent amounts (adjusted according to Western blot analysis) of XPD variants and CAK-XPD to the 5' extension probe. DNA shifts were analyzed by electrophoresis in 6% poly-acrylamid gel followed by autoradiography. BP Bound probe, FB Free probe **d.** 5'-3' helicase activities of rIIHs complexes. Insect cells were infected with a set of baculoviruses overexpressing the subunits of TFIIH including either wt or mutated Flag-tagged XPD and complexes were immuno-precipitated using an antibody (Ab) directed toward the Flag epitope in buffer B. After elution with the Flag synthetic peptide, immunopurified complexes were analyzed by Western Blot analysis (WB) and their 5'-3' helicase activities (Helicase) were tested.

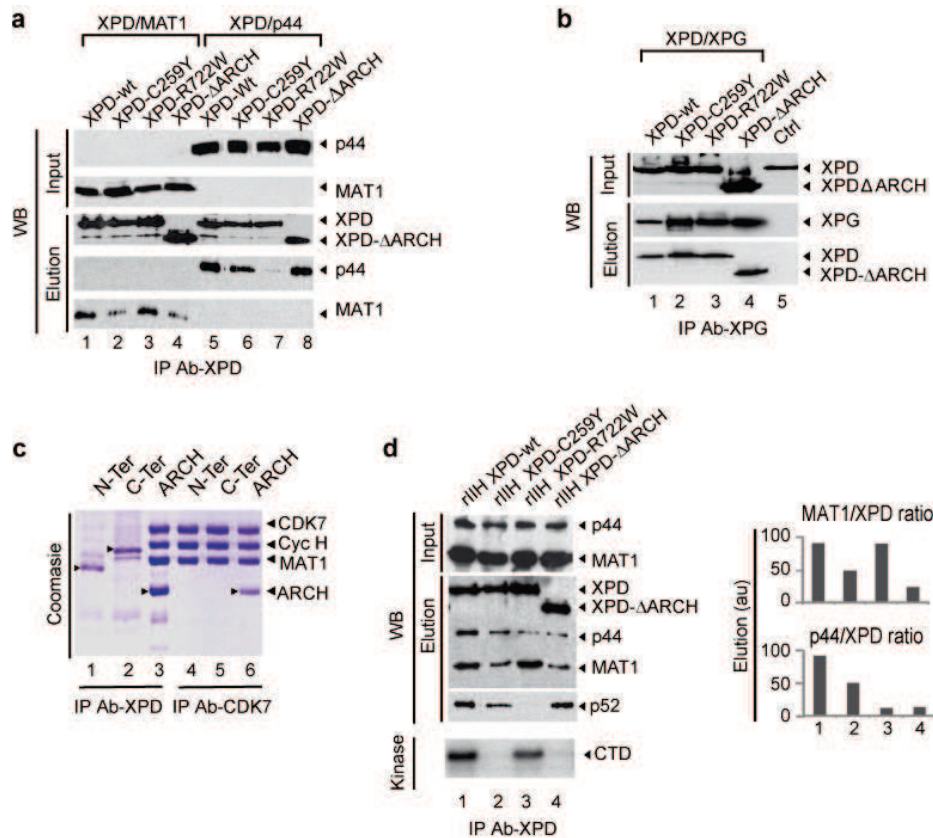


**Figure 3: Transcription and CTD kinase activity of rIIHs.**



**a.** Basal transcription activity. Increasing amounts of the purified rIIHs with different XPD variants were added to an *in vitro* reconstituted transcription system lacking TFIIF. Transcripts were analyzed by electrophoresis followed by autoradiography. The length of the corresponding transcript is indicated on the left side. The signals were quantified and plotted in arbitrary units (au). **b.** To evaluate the RNA Pol II kinase activity of the TFIIF variants, purified CAK, core-TFIIF and XPD mutants were mixed in an *in vitro* assay containing all the basal transcription factors and the AdMLP. Arrows indicate hypo (IIA) and hyper (IIO) phosphorylated forms of RNA pol II. **c.** The transcription activity of XPD variants was assessed using fixed amounts of purified rIIHs in presence of increasing amounts (0, x and 2x) of purified recombinant CAK complex. The signals were quantified and plotted in arbitrary units (au).

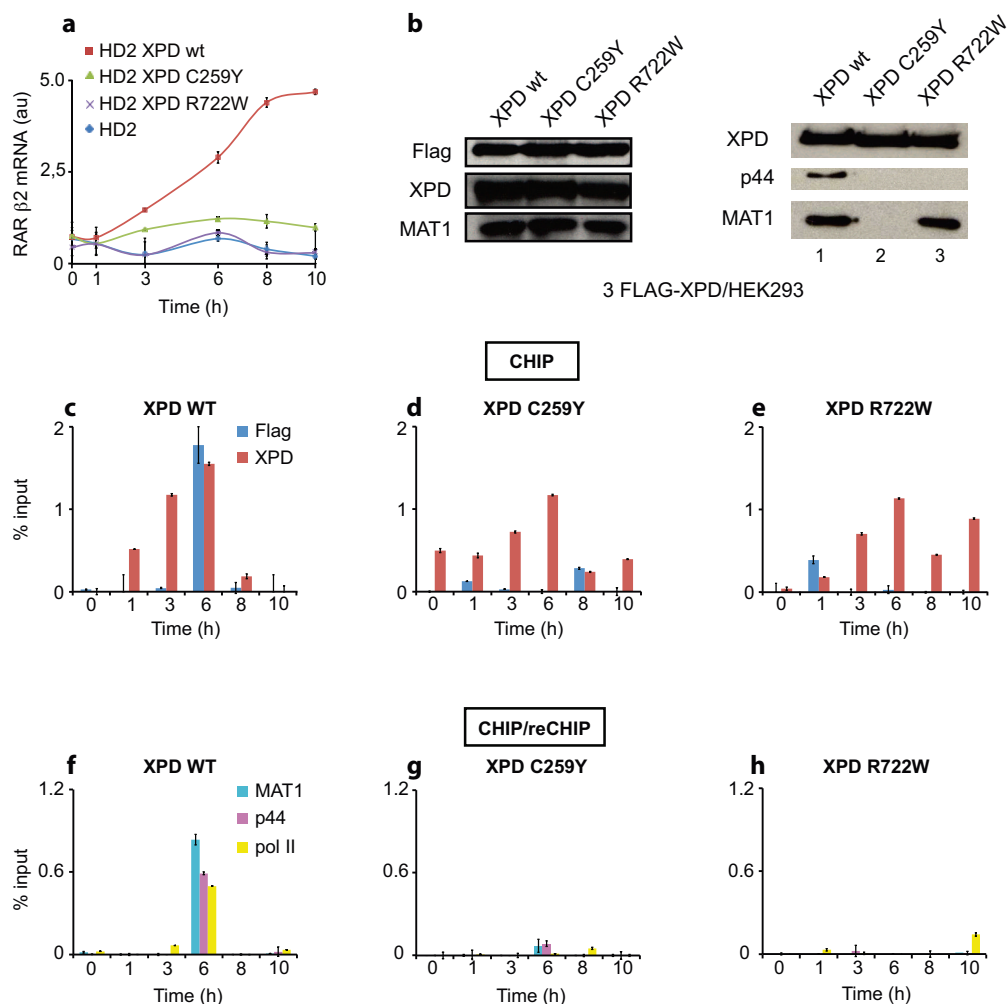
**Figure 4: The ARCH domain is a recruitment platform for CAK**



**a.** Pair wise interactions between XPD variants and MAT1 or p44. *Sf9* insect cells were co-infected with viruses which allow the expression of Flag-tagged XPD variants and a virus which allows the expression of MAT1 or p44. Proteins were immuno-purified with anti-Flag M2 resin epitope in presence of 250 mM KCl (Buffer C). Input and purified complexes were analyzed by anti-p44, anti MAT1 or anti-Flag western blot. **b.** Pair wise interactions between XPD variants and XPG. *Sf9* insect cells were co-infected with viruses which allow the expression of Flag-tagged XPD variants and a virus which allows the expression of c-myc tagged XPG. Immuno-purification was performed with anti-c-Myc resin in buffer C. Input and the purified complexes were analyzed by western blot with anti-Flag antibody. **c.** Interactions between XPD fragments and CAK. *Sf9* insect cells were co-infected with a virus which allows the expression of the CAK subcomplex with CDK7 carboxy-terminal strep tag fusion and with viruses which allow the expression of either N-ter XPD (1-245), C-ter XPD (443-762) or ARCH (245-443) with amino-terminal Flag tag. The capacity of XPD fragments to bind CAK and to co-purify was tested using the Flag tag (lane 1,2 and 3) or the strep tag II (lane 4,5 and 6) for the purification of the complexes in buffer C. Purified proteins were analyzed on 12.5 % SDS-PAGE followed by coomassie staining. Proteins are indicated with arrowhead and molecular weight markers are shown with provided sizes. **d.** Consequences of XPDs mutations on association with CAK and core-TFIIH. Insect cells were infected with a set of baculoviruses overexpressing the subunits of TFIIH including either wt or mutated Flag-tagged XPD and complexes were immuno-precipitated using an antibody (Ab) directed toward the Flag epitope in buffer C. After elution with the Flag synthetic peptide, immunopurified complexes (rIIHs) were analyzed by Western Blot analysis (WB, top) and tested for their capacity to phosphorylate the CTD of RNA Pol II (autoradio; bottom).

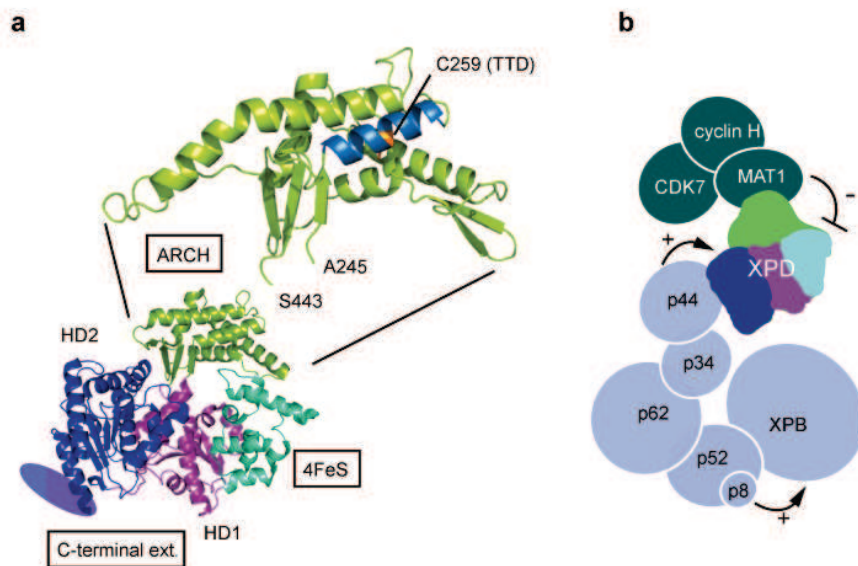
Diagrams represent estimated ratios (arbitrary units, AU) between MAT1 (a representative CAK subunit) and XPD or p44 (a representative core-TFIIF subunit).

**Figure 5: Implication of XPD in retinoic acid dependent recruitment of TFIIF on RAR $\beta$ 2 promoter.**



**a.** Relative RAR $\beta$ 2 mRNA expression monitored by qPCR from XPD transfected HD2 cells (XPD wt, XPD C259Y, XPD R722W) treated with 1 $\mu$ M *all-trans* retinoic acid (t-RA). Effect of XPD mutations on nuclear hormone receptors mediated transactivation. The XPD deficient cell line HD2 was transfected with an empty plasmid (blue diamond), with a plasmid expressing XPD-wt (red square), XPD-C259Y (green triangle) or XPD-R722W (magenta cross) and treated with *all-trans* retinoic acid. Transcription of the RAR $\beta$ 2 mRNA was quantified by Q-PCR. **b.** Western-blotting analysis for Flag, XPD and MAT1 from chromatin extracts of 3Flag-XPD transfected HEK stable cell lines (XPD wt, XPD C259Y, XPD R722W). **c-h.** CHIP/ReCHIP monitoring the coimmunoprecipitation of RAR $\beta$ 2 promoter using Flag or XPD antibodies (upper panels c, d and e) or antibodies combination against Flag/MAT1, Flag/p44 or Flag/pol II (lower panels f, g and h) from 3Flag-XPD transfected HEK stable cell lines (XPD wt, XPD C259Y, XPD R722W) treated with t-RA (1 $\mu$ M).

**Figure 6 XPD structure and TFIIH architecture**



**a.** Modular organization of XPD. *T.acidophilum* XPD (PDB: 2VSF) is composed of four structural domains: two RecA like helicase domain (HD1 in blue and HD2 in magenta), one domain which contains an iron sulfur group (in cyan) and a domain described as the ARCH domain (in green). The C-terminal extension of eukaryotic XPDs that is required for interaction with the p44 core-TFIIH subunit and for which no structural data is available is represented by an ellipse. Residue 259 mutated into a tyrosine in the TTD12PV patient cell line maps onto the first helix of the ARCH domain and its side chain points into the center of the helical bundle. **b.** Schematic of subunit architecture of TFIIH. Each subunit is represented by a circle with the radius of a sphere corresponding to its molecular weight. Interactions between XPB, p52 and p8/TTD-A stimulate XPB ATPase activity and consequently favor binding of TFIIH to damaged DNA. Interactions between p44 and XPD stimulate the helicase activity, allowing unwinding of DNA around the lesion and subsequent double incision by the endonucleases XPF-ERCC1 and XPG. XPD helicase activity, dispensable for transcription initiation but is required for NER is repressed when CAK is associated with TFIIH.

## Rôles Transcriptionnels des facteurs NER

### Résumé

Lors de la vie, des mécanismes de réparation de l'ADN sont mis en œuvre lors d'agressions, pour protéger le génome. La réparation par excision de nucléotides (NER) est l'un de ces mécanismes. Des mutations des facteurs NER sont à l'origine de 3 maladies génétiques humaines: *Xeroderma pigmentosum* (XP), la trichothiodystrophie (TTD) et le syndrome de Cockayne (CS). Certains de leurs signes cliniques ne sont pas expliqués par un défaut de réparation de l'ADN. Des études suggèrent que ces facteurs interviennent dans d'autres processus, notamment lors de l'expression des gènes.

Durant ma thèse, je me suis intéressé aux rôles des facteurs NER dans la transcription. En effet, j'ai montré que ces facteurs, dit de réparation, étaient recrutés avec la machinerie transcriptionnelle au niveau du promoteur et du terminateur de gènes activés. Ils influencent l'environnement chromatinien des gènes activés (boucles de chromatine et modifications post-traductionnelles des histones). Ma thèse apporte une meilleure compréhension du processus de transcription des gènes activés, permettant de mieux comprendre certaines anomalies associées aux syndromes XP, CS et TTD.

### Résumé en anglais

Throughout life, the mechanisms of DNA repair are implemented in attacks to protect the integrity of our DNA. The nucleotide excision repair (NER) is one of these mechanisms. Mutations targeting genes of NER factors (XPA-G, TTD-A, CSA and CSB) are responsible for three human genetic diseases: *Xeroderma pigmentosum* (XP), trichothiodystrophy (TTD) and Cockayne syndrome (CS). Some of them clinical features cannot be explained by a defect in DNA repair only. Previous studies suggest that these factors could be involved in other functions, including gene expression.

In my thesis, I am interested in the roles of NER factors during the transcription process. Indeed, we have shown that these "repair" factors, were recruited with the transcriptional machinery at the promoter and terminator of activated genes during transcription. They influence the chromatin environment of activated genes (chromatin loops and post-translational modifications of histones).

My thesis provides a better understanding of the transcription process of activated genes and allows a better understanding of some syndromes associated with XP, CS and TTD.