UNIVERSITÉ DE STRASBOURG

igbmc institut de génétique et de biologie moléculaire et cellulaire

ÉCOLE DOCTORALE des Sciences de la Vie et de la Sante

IGBMC - CNRS UMR 7104 - Inserm U 964

THÈSE présentée par :

Amita SINGH

soutenue le : 29 Septembre 2014

pour obtenir le grade de : **Docteur de l'université de Strasbourg**

Discipline/ Spécialité : Sciences du Vivant/Aspects Moléculaires et Cellulaires de la Biologie

Involvement of TFIIH in NER factors mediated Chromatin remodeling

Contribution de TFIIH dans le remodelage de la chromatine dépendant des facteurs NER lors de la transcription

THÈSE dirigée par :

EGLY Jean-Marc	DRE INSERM, IGBMC, Strasbourg (Fr)

RAPPORTEURS :

SCHAEFFER Laurent	DR, LBMC, Lyon (Fr)
GELI Vincent	DR, CRCM, Marseille (Fr)
LAUGEL Vincent	PUPH, CHU de Strasbourg (Fr)

We are what our thoughts have made us; so take care about what you think. Words are secondary. Thoughts live; they travel far.

….Swami Vivekananda

Acknowledgements

Today, when I look back and remember my first day in institute, I feel as if it was just a matter of few months and not 4 years. When I came, I had a feeling of being in a completely different world. I did not realize when this new world became home for me. My thesis is not only my work at bench and computer, in fact there are several people behind it whom I am grateful to and would like to express my gratitude towards them.

I consider it an honor to work with Dr. Jean Marc Egly and thank him for providing me with an opportunity to work in an international environment like IGBMC. I thank him for his comments and suggestions throughout my PhD.

I would thank Dr. Laurent Schaeffer, Dr. Vincent Geli and Dr. Vincent Laugel for accepting my invitation to be the part of my PhD defense jury and for analyzing this work.

I also express my gratitude towards Dr. Frederic Coin for his valuable scientific suggestions, support and signatures for attestations all through my PhD. I also thank and appreciated Dr. Emmanuelle Compe for helping me with the footprinting assays and calmly answering my most stupid questions.

I am indebted to Cathy Braun, who helped me in performing all the in-vitro experiments. My thesis could have not been possible without her and I specially thank her for this support and contribution to my work.

I thank to all those lab members also who departed, especially Izarn and Salim for helping me with experimental suggestions and of course with several occasions of French translations. I thank my present lab member Zita, Sergey, Alexey, Charlotte, Annabelle, Philippe, and Baptiste for their support and scientific inputs. I cannot forget to thank Marc and Federico for converting our experimental box into a comedy show stage. I also thank my friends in the lab, Lise Marie and Jitika for their love and care. I thank Carlos for his support during my thesis writing.

I thank my friends Avisek, Shilpy and Poonam who made my life easy in Strasbourg with their affection and care. I also thank my friend Pallavi, Priyanka, Baby, Vishnu and Vandana who always cheered me and showed their faith to move forward.

I express my gratitude towards my teachers Ragini Madame, Kavita Madame, Rakesh Sir. Your teachings were invaluable and always helped me to become a better person. I am extremely grateful to my teacher from Poorvanchal University, Prof. DD Dubey, Dr. Vanadana Rai, and

Dr. Pradeep Kumar for their esteemed guidance. I thank Dr. Mercy Raman, my supervisor during my master dissertation, who motivated me to actually opt for research as a career.

It is with immense gratitude that I acknowledge the support and help of my mentor, Dr. Le May Nicolas. This thesis would have not been possible without his guidance. From the very first day in the lab, he essentially taught me everything. Being tough from outside, he was always soft from the core. Thank you for being with me, your support has been a true enthusiasm to me.

My accomplishment would never be possible without the support of my family. I thank my almighty for giving me such encouraging in-laws who always motivated me for moving ahead in my career. All through my journey until now, my brothers, my parents have big role in my life. I could never do anything without your love and care. Especially, my father who inspired me at every point of my life. He gave me the greatest gift of "believing in me more than his own sons". This belief was the biggest strength which always helped me to pull myself through all the tough times.

I never thank him but always thanks to God for sending my soul mate, Deepankar, in my life. He has been an amazing friend and guide. He stands behind me, supports me in all my decisions and helps me with all my problems.

At the end, I would like to dedicate my thesis to my first mentor of life, my uncle Jai Prakash Singh, who was no less than a father to me. He left this world when I was in my 12th standard. As a human being, you influenced me even when I was just starting to understand words. May be you are not with me physically, but wherever you are I could feel you deep in my heart and I owe you and dedicate you all my accomplishments.

Amita Singh

Table of content

TABLE OF CONTENT 5

LIST OF FIGURES 7

LIST OF TABLES 7

LIST OF ABBREVIATIONS 8

RESUME OF THE THESIS IN ENGLISH 11

RESUME DE LA THESE DE DOCTORANT EN FRANÇAIS 16

INTRODUCTION 25

I. THE TRANSCRIPTION MECHANISM 25

A. THE PROMOTER ELEMENTS AND TRANSCRIPTION COMPONENTS 25

A.1 The core promoter 25

1.1 The TATA box 27

1.2 The Initiator element (Inr) 28

1.3 The TFIIB recognition elements (BREu and BREd) 28

1.4 The Downstream promoter element (DPE) 28

1.5 The Motif ten element (MTE) 29

1.6 The downstream core element (DCE) 29

1.7 The TCT motif (polypyrimidine initiator motif) 29

1.8 The X core promoter element 1 (XCPE 1) 29

1.9 TATA- less promoters 31

A.2 RNA polymerase II and Transcription Machinery 31

2.1 RNA polymerase II and CTD modifications 31

2.2 TFIIA 35

2.3 TFIIB 37

2.4 TFIID 38

2.5 TFIIE 39

2.6 TFIIF 39

2.7 TFIIH 40

2.8 Mediator 40

B. THE TRANSCRIPTION CYCLE 42

B.1 Preinitiation complex assembly, open complex formation and initiation. 42

B.2 Promoter clearance 44

B.3 Proximal pausing and elongation 44

B.4 Termination and reinitiation 45

C. THE ACTIVATED TRANSCRIPTION 47

C.1 Overview of the nuclear receptor families 47

C.2 Retinoic acid receptor 48

C.3 Transcription of retinoic acid receptor target genes 52

D. CHROMATIN AND GENE REGULATION 57
D.1 Epigenetic modification of chromatin *59*
1.1 Post-translational modification of histones 59
1.1.a. Regulation by histone lysine acetylation and deacetylation. 61
1.1.b. Regulation by histone lysine/arginine methylation and demethylation 62
1.1.c. Regulation by other histone modifications 63
1.2 DNA methylation 64
D.2 Chromatin rearrangement *68*

II. THE TFIIH AND NER FACTORS IN TRANSCRIPTION 71

A. TFIIH: A MULTIFUNCTIONAL COMPLEX 71
A.1 Composition of TFIIH *71*
1.1 XPB 74
1.2 XPD 77
1.3 p62 79
1.4 p52 80
1.5 p44 80
1.6 p34 81
1.7 Cdk7 82
1.8 Cylin H 83
1.9 Mat1 84
1.10 p8/TTD-A 84
B. NER FACTORS 86
C. TFIIH: BRIDGING TRANSCRIPTION AND REPAIR 89

III. HUMAN DISORDERS ASSOCIATED WITH MUTATION IN TFIIH AND NER FACTORS 92

A. XERODERMA PIGMENTOSUM 94
B. COCKYANE SYNDROME 95
C. XERODERMA PIGMENTOSUM/COCKYANE SYNDROME 96
D. TRICHOTHIODYSTROPHY 97

RESULTS 101

A. INVOLVEMENT OF TFIIH IN NER FACTORS DEPENDENT CHROMATIN REMODELING 102
B. ENZYMATIC ACTIVITY OF TFIIH 142
B.1 Subunit composition of TFIIH *142*
B.2 ATPase and helicase activity of TFIIH *144*
B.3 Experimental procedures: *148*

DISCUSSION 150

A. TFIIH, NER FACTORS AND CHROMATIN REMODELING 151
A.1 Biochemistry of TFIIH and the overlapping genetics of diseases *151*
A.2 TFIIH deploys the recruitment of NER factors and chromatin remodeling events *153*
B. FUTURE DIRECTION FOR TFIIH AND NER FACTORS 157

CONCLUSION 160

BIBLIOGRAPHY 162

List of Figures

Figure 1. Core promoter motifs for transcription by RNA polymerase II. 26
Figure 2. General transcription factors of RNA polymerase II 30
Figure 3. RNA polymerase II transcription cycle. 43
Figure 4. Structure of retinoic acid receptor 49
Figure 5. Coregulator exchange at RXR/RAR heterodimers. 53
Figure 6. Crosstalk between the RA-activated p38MAPK pathway and the expression of RAR target genes. 55
Figure 7. Nucleosome organization. 58
Figure 8. Schematic cartoon of core histone PTMs and their potential function.................................... 60
Figure 9. DNA demethylation Pathway. 66
Figure 10. Structure of TFIIH 72
Figure 11. Linear schematics of XPB and XPD helicases. 76
Figure 12. XPD and its interactions 78
Figure 13. Relationship of genetic disorders to molecular defects. 93
Figure 14. TFIIH subunits in XPBwt and XPB mutant cell lines. 143
Figure 15. TFIIH subunits in XPDwt, XPD mutant, p8/TTD-Awt and p8/TTD-A mutant cell line............. 143
Figure 16. ATPase activity of XPB.. 145
Figure 17. ATPase activity of XPD 147
Figure 18. Helicase activity of XPD.. 147

List of Tables

Table 1. Consensus sequence and position of the different core promoter elements. 26
Table 2. Different subunits of RNA polymerase II in eukaryotes, archaea, bacteria and their features. .. 32
Table 3. CTD kinases and Phosphatase in mammals and in yeast. 34
Table 4. General transcription factors of RNA polymerase II. 36
Table 5. TFIIH components and their functions.. 73
Table 6. NER factors and their interaction partners 88
Table 7. Mutation in XPB, XPD and TTD-A gene studied during PhD.. 93
Table 8. Comparison of features of XP, XP-ND, TTD, CS and XP/CS.. 99
Table 9. Functions of TFIIH... 156

List of abbreviations

ATM	Ataxia telangiectasia mutated
ATR	ATM and Rad3 related
BER	Base excision repair
bp	Base pair
BRCA	Breast cancer protein
BRE	TFIIB recognition element
BTF2	Basal transcription factor 2
CAF-1	Chromatin assembly factor 1
CAK	Cdk-Activated Kinase
CDC	Cell division cycle
CDK	Cyclin dependent kinase
ChIP	Chromatin immuno-precipitation
CHO	Chinese hamster ovary
COFS	Cerebro-oculo-facio-skeletal syndrome
CPD	Cyclobutane pyrimidine dimer
CTD	C-terminal Domain
CS	Cokayne’s syndrome
CSA/B	Cockayne syndrome protein group A/B
DNA	Deoxyribonucleic acid
DBD	DNA-binding domain
DCE	Downstream core element
DDB	DNA damage binding protein
DNA-PK	DNA protein kinase
DPE	Downstream promoter element
DSB	Double strand break
ERCC1	Excision repair cross-complementing rodent repair
FANC	Fanconi anemia complementation group
GFP	Green fluorescent protein
GGR	Global genome repair
GTF	General Transcription factor
GTF2H	General Transcription factor IIH
HAT	Histone acetyl-transferase
HD	Helicase domain
HDAC	Histone deacetylase
HP1	Heterochromatin protein 1

HR23B	Homologue Rad23B
Inr	Initiator
kB	Kilo-base
kDa	Kilo-Dalton
LBD	Ligand binding domain
Lig	Ligase
MAT1	Ménage à trois
MLL	Mixed lineage leukemia
MSH	Mut S Homologues
NER	Nucleotide excision repair
NR	Nuclear receptors
PARP	Poly (ADP-ribose) polymerase
PARG	Poly-ADP-ribose-glycohydrolase
PCNA	Proliferating cell nuclear antigen
PHD	Pleckstrin homology domain
PIC	Preinitiation complex
PPAR	Peroxisome Proliferator -Activated Receptor
Pol	Polymerase
RAD	Homologous recombination factor
RAR	Retinoic Acid Receptor
t-RA	Trans-Retinoic Acid
RE	Response element
RFC	Replication factor C
RING	Really interesting new gene (a type of zinc finger motif)
RPA	Replication protein A
RXR	Retinoid X Receptor
SAM	S-Adenosyl methionine
ssDNA	Single stranded DNA
SUMO	Sumoylation N-methylpurine DNA glycosylase
SWI/SNF	Switch/Sucrose non fermentable
TBP	TATA box binding protein
TCR	Transcription coupled repair
TFIIA-H	Transcription factor II A-H
Tfb	Transcription factor b
TTD	Trichothiodystrophy
TRF	Telomeric repeat binding factor

TSS	Transcription stat site
UV	Ultra-violet
VDR	Vitamin D Receptor
XP	Xeroderma Pigmentosum
XAB	XPA- binding protein
XPA-V	XP complementation group A-V
XRCC	X-ray Cross complementation

Resume of the thesis in English

Involvement of TFIIH in NER factor mediated chromatin remodeling

TFIIH: a dynamic complex

TFIIH is a multi-subunit complex, first characterized and purified as a general transcription factor of RNA polymerase II (RNA pol II). Originally TFIIH was thought to be exclusively a basal transcriptional factor but later it was found to be involved in DNA repair and possibly in cell cycle regulation as one of the major component essential for the life of the cell. It consists of two sub-complexes: core complex and cyclin-activating kinase (CAK). The core complex consists of six subunits: XPB, p62, p52, p44, p34 and p8/TTD-A), (Hoeijmakers 2001). CAK is composed of the three subunits: CDK7, cyclin H and MAT1. The core and the CAK are held together by the XPD, subunit. TFIIH possesses several enzymatic activities which are indispensable for nucleotide excision repair (NER) and transcription.

TFIIH during NER

Living organisms are continuously exposed to damaging agents such as UV-rays, X-rays which creates DNA damage, CPD- photoproducts in the genome. To maintain genomic stability these DNA lesions are removed by several DNA repair pathways. NER pathway is one of them and subdivided into global genome repair (GGR) pathway and the transcription-coupled repair (TCR) pathway. TFIIH functions after a DNA lesion has been recognized by either the GGR pathway or the TCR pathway of NER. In GGR pathways the DNA damage is recognized by XPC/HR23B complex while in TCR the stalled RNA pol II stalled in front of the DNA lesions on the transcribed DNA strand initiates the recruitment of the CSA and CSB. TFIIH then unwinds the DNA around the lesions via its 3′-5′ helicase/ATPase activity of XPB and 5′- 3′ helicase activity by XPD. Opening of the bubble around the damage is maintained by the arrival of XPA and RPA. Finally the damaged DNA is removed by XPG and XPF endonucleases which creates cut at 3′ and 5′ respectively.

TFIIH in transcription

Eukaryotic gene expression is driven by a complex series of events that starts with the recruitment of basal transcription factors (TFIIF, TFIIB, TFIID and TFIIE) to the promoter region. Following addition of TFIIH, promoter melting and the open complex formation take place, leading to promoter escape, a transition away from promoter into elongation complex. In particular, the XPB enzymatic activity plays an important role in the promoter escape during transcription initiation by RNA pol II (Bradsher, Coin et al. 2000). The XPB activity inside of TFIIH is highly regulated by others TFIIH subunits, such as directly by p52(Lee, Park et al. 2004), or indirectly by p8/TTD-A through p52 (Kim, Patel et al. 1995). Another subunit cdk7 mediates the phosphorylation of the carboxyl terminal domain (CTD) of the largest subunit of the RNA pol II. It also mediates the phosphorylation of some other transcription factors and nuclear receptors and regulates basal as well as activated transcription. It has been recently revealed that NER factors (XPC, CSB, XPA, XPG and XPF) along with TFIIH shares the dual activity in DNA repair and transcription. Indeed, they are recruited at active promoters and participate in the regulation of gene expression in the absence of genotoxic attack(Le May, Mota-Fernandes et al. 2010). These NER factors contribute to achieve chromatin remodeling including histones post-translational modifications (PTMs), DNA breaks, DNA demethylation and gene looping which allows accurate and optimal transcription (Le May, Fradin et al. 2012).

Mutation in XPB, XPD and p8/TTD-A subunit of TFIIH and NER factors (XPA-G, ERCC1, TTD-A, CSA, and CSB), have been associated with the human genetic disorders such as *Xeroderma pigmentosum* (XP), trichothiodystrophy (TTD), cockayne syndrome (CS), and cerebro-oculo-facio-skeletal syndrome (COFS). These patients show a perplexing clinical heterogeneity ranging from 1,000-fold increased frequency of skin cancers (Kraemer, Lee et al. 1987; Kraemer, Levy et al. 1994), short stature to neurological problems and premature aging. The clinical features are now being attributed to the combined effect of DNA repair deficiency and transcriptional dysregulation. Studies of TFIIH have demonstrated a connection between transcription and DNA repair and have opened a new field of transcription diseases. It have been shown that mutations in the C-terminal domain of the XPD subunit disturbs the

architecture of TFIIH and its molecular communication with the retinoic acid receptor α (RARα) (Keriel, Stary et al. 2002), PPARs (Compe, Drane et al. 2005), the estrogen receptor (Chen, Riedl et al. 2000), the thyroid hormone receptors(Compe, Malerba et al. 2007) and the androgen receptor (Chymkowitch, Le May et al. 2011) leading to the dysregulation of cdk7-related phosphorylation and consequently the transactivation of nuclear receptor- targeted genes.

During my thesis I investigated deeper molecular intricacies of TFIIH particularly, the roles of XPB, XPD, and p8/TTD-A subunit of TFIIH during transcription that could finally lead to better comprehension of the etiology of the clinical features and the underlying molecular mechanisms associated with distinct mutations in TFIIH found in XP, XP/CS, or TTD patients. Using fibroblasts cell lines derived from XP, XP/CS and TTD patients harboring the corresponding mutations in respective subunits of TFIIH, I studied different mutations in XPB (F99S and T119P), XPD (R112H, G602D, R683W, R722W) and p8/TTD-A (L21P, R56stop) subunit of TFIIH. As a model, one of the nuclear receptor responsive genes RARβ2 which is targeted and activated by the retinoic acid receptor (RAR) upon treatment with all-trans retinoic acid (RA) was used. Transactivation of RARβ2 was thus analyzed during time course, which showed significant disturbance in mRNA expression profile of RARβ2 in XPB, XPD and p8/TTD-A mutant cell lines compared to their respective wild type cell line. Further, we investigated the recruitment of the basal transcriptional machinery and NER factors along the promoter and the terminator region of RARβ2 by Chromatin Immunoprecipitation (ChIP) assay. In wild type (WT) cell lines we observed the recruitment of the basal transcriptional machinery (RNA pol II, TFIIB and TFIIH) together with the NER factors (XPA, CSB, XPG, and XPF) at a specific time point which coincides with the RARβ2 mRNA expression peak. Although each XPB and XPD mutations led to different and specific dysregulation in the formation of the transactivation complex, but taken together it suggests a disturbed recruitment of the transcriptional machinery, TFIIH complex and the NER factors at the active promoter. Such observations suggested a link between the formation of TFIIH complex upon initiation and the concomitant recruitment of the NER factors. The recruitment of the NER factors are shown to be prerequisite for the chromatin remodeling events which includes histone PTMs, DNA breaks and active DNA demethylation at the promoter. We here showed that histone

PTMs in XPBwt, XPDwt, p8/TTD-Awt displayed a clear enrichment in the di-methyl of histone H3K4 and acetylation of H3K9 with a simultaneous decrease in di-methyl of H3K9 marks which coincides with mRNA expression of RARβ2, transcription machinery, NER factors. The histone landmarks linked to the active transcription were significantly disturbed in all the XPB, XPD and p8/TTD-A mutant cell lines.

Considering the role of XPG and XPF endonucleases in DNA break formation and DNA demethylation which eventually stabilizes the chromatin looping between the promoter and terminator region of RARβ2, we analyzed association of these chromatin remodeling events with TFIIH. We thus showed the presence of DNA breaks both at the promoter and the terminator of XPBwt, XPDwt and p8/TTD-Awt cells upon ATRA treatment with concomitant recruitment of XPG and XPF at respective regions. In all the mutant cells, we observed a disturbed correlation between the XPG/ XPF endonucleases with the DNA breaks at both promoter and terminator regions.

Such DNA break formations by XPG at promoter and by XPF at terminator have eventually been correlated with the achievement of active DNA demethylation at promoter. We thus next analyzed the implication of TFIIH in DNA demethylation process. Using Un-methyl IP approach, we observed that active DNA demethylation is significantly disturbed at the promoter of RARβ2 in all the XPB, XPD and p8/TTD-A mutant cell lines. Presence of all the components, the basal transcriptional machinery, NER factors, CTCF, DNA break and DNA demethylation together have been correlated for the presence of chromatin looping between promoter and terminator. Furthermore using 3C (Chromosome confirmation capture) technique we demonstrated that TFIIH mutations disturb the CTCF dependent gene looping between promoter and terminator region of RARβ2 in the XPB, XPD and p8/TTD-A cells.

Conclusion

As a general conclusion, my results strongly supported an involvement of TFIIH in the recruitment of the NER factors at active promoter and consequently in their roles in the chromatin remodeling including PTMs of histones, active DNA demethylation, DNA breaks induction and CTCF-dependent gene looping of *RARβ2*. Significantly my work has contributed to unveil key roles of TFIIH in transcription, thereby providing a step forward towards the understanding of transcriptional diseases: XP, XP/CS and TTD.

Résume de la thèse de doctorant en français

Contribution de TFIIH dans le remodelage de la chromatine dépendant des facteurs NER lors de la transcription

TFIIH: un complexe dynamique

TFIIH est un complexe à plusieurs sous-unités, caractérisé et purifié en tant que facteur de transcription générale de l'ARN polymérase II (Conaway and Conaway 1989; Gerard, Fischer et al. 1991). Ultérieurement il a été démontré qu'il s'agissait d'un facteur multifonctionnel impliqué dans plusieurs processus cellulaires tels que la réparation de l'ADN par excision de nucléotides (NER) et la régulation du cycle cellulaire. Il comprend deux sous-complexes, le core constitué de six sous-unités : XPB, p62, p52, p44, p34 et p8/TTDA(Keeney, Chang et al. 1993; Scrima, Konickova et al. 2008) et le CAK (Cyclin-activating kinase) composé de trois sous-unités: CDK7, cycline H and MAT1. La sous unité XPD permet de relier le core et le CAK. TFIIH possède plusieurs activités enzymatiques indispensables pour la réparation par excision de nucléotides (NER) et pour la transcription (Figure a).

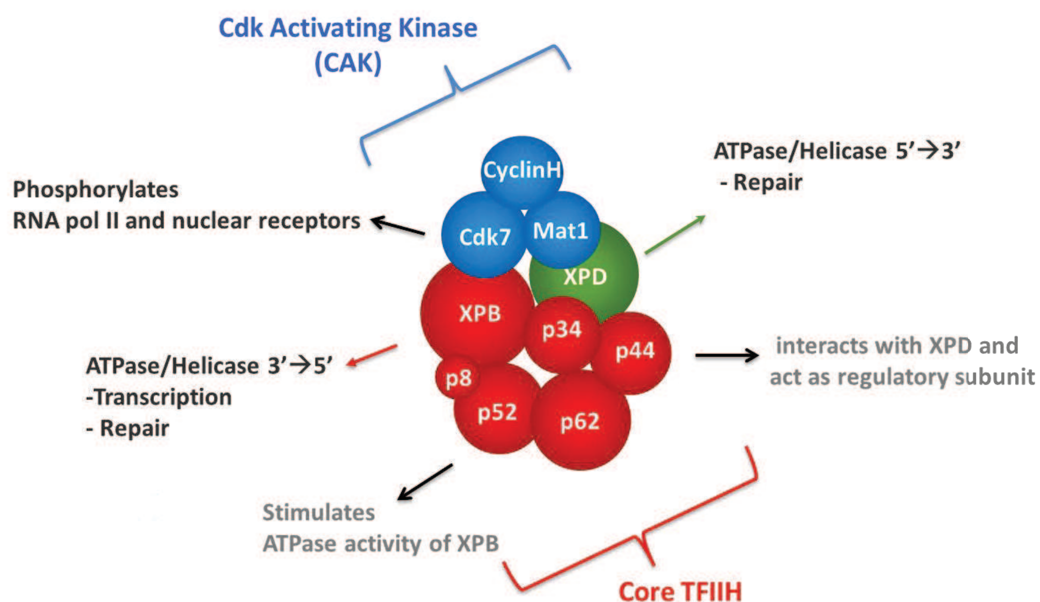


Figure a. Sous-unités de TFIIH

TFIIH dans la réparation NER

Les organismes vivants sont constamment exposés à des agents endommageant l'ADN, provenant de l'environnement tels que les rayons UV ou résultant du métabolisme cellulaire tel que le stress oxydatif. Afin de maintenir l'intégrité du génome il existe plusieurs systèmes de réparation qui permettent d'éliminer ces dommages de l'ADN. Parmi ces systèmes de réparation, la voie NER cible les lésions induisant une distorsion de la double hélice de l'ADN; elle est subdivisée en deux voies la GGR (global génome repair) et la TCR (transcription coupled repair). Dans la voie GGR, la lésion est reconnue par le complexe XPC/hHR23B alors que dans la voie TCR c'est l'ARN polymérase II bloquée en cours de transcription par la lésion qui permet d'initier la réparation en recrutant les facteurs CSA et CSB. Cette étape de reconnaissance est suivie par le recrutement de TFIIH dont le rôle est l'ouverture de l'ADN autour de la lésion grâce à l'activité ATPase de la sous unité XPB et l'activité hélicase de XPD. Ensuite il y a recrutement des autres facteurs XPA, RPA, les endonucléases XPF/ERCC1 et XPG qui induisent une double incision du brin contenant la lésion, et enfin la machinerie de resynthèse et ligation.

TFIIH dans la transcription

La transcription des gènes de classe II implique une série d'évènements et est initiée par le recrutement des facteurs généraux de transcription (TFIID, TFIIB, TFIIE TFIIF et TFIIH) sur le promoteur; l'arrivée de TFIIH permet l'ouverture de l'ADN, l'échappée du promoteur et l'initiation de l'élongation. L'activité enzymatique de XPB joue un rôle important dans l'échappée du promoteur durant l'initiation de la transcription par l'ARN pol II (Bradsher, Coin et al. 2000). Cette activité est régulée par d'autres sous unités de TFIIH, directement par p52 (Takedachi, Saijo et al. 2010), ou indirectement par p8/TTDA à travers son interaction avec p52 (Sugasawa, Okuda et al. 2005). Cdk7 une autre sous unité de TFIIH assure la phosphorylation du domaine C-terminal de la plus grande sous unité de l'ARN pol II ; elle permet aussi la phosphorylation d'autres facteurs de transcription incluant notamment la famille des récepteurs nucléaires.

Récemment, il a été démontré que les facteurs NER (XPC, CSB, XPA, XPG and XPF) tout comme TFIIH partagent une double activité dans la réparation de l'ADN et la transcription(Le May, Mota-Fernandes et al. 2010). En effet, ils sont recrutés au niveau des promoteurs de gènes induits ciblés et activés par des récepteurs nucléaires et participent à la régulation de la transcription en absence de dommages de l'ADN. Ces facteurs NER contribuent au remodelage de la chromatine constitué par des modifications post traductionnelles (PTMs) d'histones, des cassures de l'ADN au niveau du promoteur, la demethylation active de l'ADN et la formation de réarrangement chromatinien, permettent une transcription précise et optimale (Le May, Fradin et al. 2012).

Les mutations dans les sous-unités XPB, XPD et p8/TTDA de TFIIH, et des autres facteurs NER, ont été associées avec des maladies génétiques rares telles que le *Xeroderma pigmentosum* (XP), la trichothiodystrophie (TTD), le syndrome de cockayne (CS), et le syndrome cérébro-oculo-facio-squeletique (COFS). Ces patients montrent une hétérogénéité clinique allant d'une très forte prédisposition à développer des cancers de la peau(Kraemer, Lee et al. 1987; Kraemer, Levy et al. 1994) à des problèmes neurologiques ou un vieillissement prématuré. Ces symptômes cliniques sont attribués à l'effet combiné d'un défaut de la réparation NER et d'une dérégulation de la transcription.

L'étude du complexe TFIIH a démontré la présence d'un lien entre la transcription et la réparation de l'ADN et a ouvert un nouveau concept de maladies de la transcription. Il a été démontré que les mutations au domaine C-terminal de la sous unité XPD perturbaient l'architecture du TFIIH et son interaction moléculaire avec plusieurs récepteur nucléaires, comme le récepteur de l'acide rétinoïque (RAR) (Keriel, Stary et al. 2002), les récepteurs peroxisome proliferator-activated (PPARs) (Compe, Drane et al. 2005), le récepteur d'œstrogène (ER) (Chen, Riedl et al. 2000), le récepteurs de l'hormone thyroïdienne (TRs) (Compe, Malerba et al. 2007), les récepteurs de l'hormone androgène (Chymkowitch, Le May et al. 2011), résultant souvent d'une dérégulation dans le mécanisme de phosphorylation par cdk7 ; il en résulte un défaut de transactivation des gènes ciblés par ces récepteurs nucléaires.

Durant ma thèse, je me suis intéressée aux interactions moléculaires de TFIIH, et particulièrement aux rôles des sous unités XPB, XPD et p8/TTDA durant la transcription, afin de mieux comprendre les mécanismes moléculaires sous-jacents associés aux mutations concernant ces protéines, ce qui permettrait de mieux expliquer les symptômes cliniques observés chez les patients XP, XP/CS, et TTD.

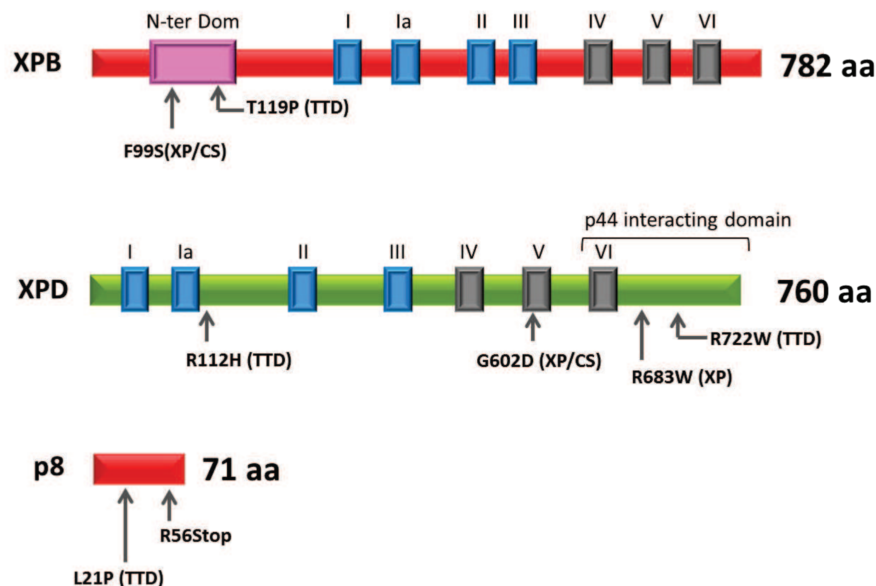


Figure b. Mutations étudiées dans XPB, XPD et p8 / TTD-A sous-unités de TFIIH

Nous avons utilisé des fibroblastes provenant des patients cites ci-dessus, pour étudier l'effet de différentes mutations des sous unités concernées de TFIIH: XPB (F99S and T119P), XPD (R112H, G602D, R683W, R722W) et p8/TTD-A (L21P, R56stop). Notre modèle consistait en l'évaluation de la transactivation d'un gène cible du récepteur nucléaire RARβ2 après traitement des cellules en culture par l'acide rétinoïque. Nous avons observé une perturbation significative du profil d'expression de l'ARNm de RARβ2 dans les cellules mutantes pour XPB, XPD et p8/TTDA en comparaison avec la lignée sauvage. Ensuite nous avons étudié par ChIP (Immunoprécipiation de la chromatine) le recrutement de la machinerie de transcription et des facteurs NER sur le promoteur et le terminateur du gène RARβ2. Dans la lignée sauvage, nous avons observé que le temps de recrutement de l'ARN pol II, des facteurs de transcription TFIIB et TFIIH ainsi que les facteurs NER (XPA, CSB, XPG et XPF) coïncidait avec le pic d'expression de l'ARNm de RARβ2. Au contraire dans les lignées mutantes pour XPB et XPD, il y avait un défaut de recrutement de

ces différents facteurs sur le promoteur actif qui variait selon le type de gêne et la position de la mutation. Ces observations suggèrent un lien entre TFIIH et le recrutement concomitant des facteurs NER lors de l'initiation de la transcription. D'autant plus qu'il a été démontré que ce recrutement des facteurs NER lors de la transcription en absence de dommages de l'ADN était une condition préalable aux événements de remodelage chromatinien incluant les modifications d'histones, les cassures d'ADN et la déméthylation active au niveau du promoteur. Nous avons ici montré que dans les cellules sauvages il y avait une augmentation des modifications d'histones H3K4me2 et H3K9ac avec une réduction de H3K9me simultanément au pic d'expression de l'ARNm de RARβ2, du recrutement des facteurs de transcription et des facteurs NER. Ces différentes modifications d'histones caractéristiques de la transcription active ont été significativement perturbées dans les lignées de cellules mutantes pour XPB, XPD et p8/TTDA.

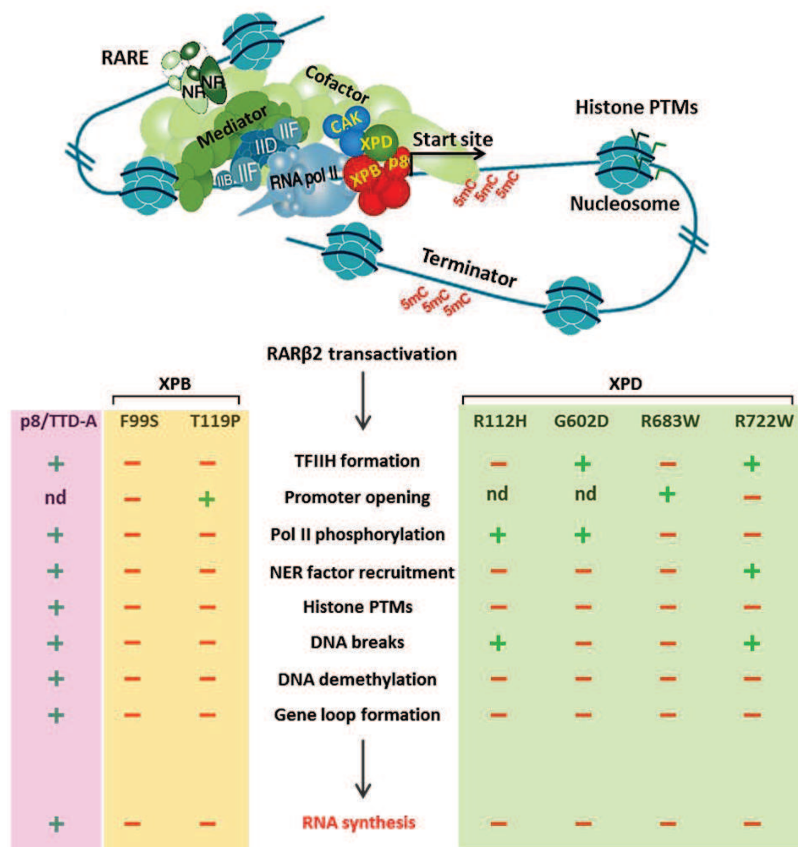
Prenant en considération le rôle des endonucléases XPG et XPF dans la formation des cassures d'ADN et la déméthylation qui permettrait de stabiliser la formation de structures en boucles (gène looping) entre le promoteur et le terminateur du gène RARβ2, nous avons analysé l'association entre ces événements de remodelage chromatinien et l'état de TFIIH. Nous avons ainsi observé en plus du recrutement de XPG et XPF, la présence de cassures d'ADN autant tant au niveau du promoteur qu'au niveau du terminateur suite au traitement par l'acide rétinoïque dans les cellules WT, alors que dans les cellules mutantes, cette corrélation entre les deux endonucléases et les cassures d'ADN était perturbée.

Il a été montré que les cassures de l'ADN induites par XPG au niveau du promoteur, et par XPF au niveau du terminateur, étaient corrélées au processus de déméthylation active, ces évènements sont hautement coordonnés afin de permettre une transcription optimale. Nous avons analysé l'implication de TFIIH dans ce processus en utilisant l'approche Unmethyl-IP. Nous avons ainsi observé que dans les cellules mutantes, la déméthylation active au niveau du promoteur de RARβ2 était significativement perturbée. Nous avons aussi démontré par la technique 3C (Chromosome conformation capture) que ces mutations de TFIIH altéraient la formation des boucles de chromatine entre le promoteur et le terminateur du RARβ2 aidée par l'organisateur de la chromatine CTCF.

Conclusion

Mes résultats suggèrent fortement une implication de TFIIH dans le recrutement des facteurs NER sur le promoteur actif ainsi que pour leurs rôles lors du remodelage de la chromatine incluant les modifications des histones, la déméthylation active de l'ADN, l'induction des cassures d'ADN et la formation des boucles dépendante de CTCF.

Chaque mutation dans TFIIH ici étudiée présente une dérégulation spécifique illustrant la complexité des défauts transcriptionnels observés chez les patients XP, TTD et CS issus de la combinaison des fonctions enzymatiques de TFIIH et les étapes de remodelage de la chromatine dans la cascade des évènements qui constituent la régulation de l'expression des gènes. .



p8/TTD-A	F99S	T119P		R112H	G602D	R683W	R722W
+	−	−	TFIIH formation	−	+	−	+
nd	−	+	Promoter opening	nd	nd	+	−
+	−	−	Pol II phosphorylation	+	+	−	−
+	−	−	NER factor recruitment	−	−	−	+
+	−	−	Histone PTMs	−	−	−	−
+	−	−	DNA breaks	+	−	−	+
+	−	−	DNA demethylation	−	−	−	−
+	−	−	Gene loop formation	−	−	−	−
+	−	−	RNA synthesis	−	−	−	−

Figure c. Chaque mutation dans les sous-unités de TFIIH représente spécifique dérégulation de transcription

References:

1 Conaway, J. W. & Conaway, R. C. A multisubunit transcription factor essential for accurate initiation by RNA polymerase II. *J. Biol. Chem.* 264, 2357-2362 (1989).

2 Gerard, M. *et al.* Purification and interaction properties of the human RNA polymerase B(II) general transcription factor BTF2. *The Journal of biological chemistry* 266, 20940-20945. (1991).

3 Giglia-Mari, G. *et al.* A new, tenth subunit of TFIIH is responsible for the DNA repair syndrome trichothiodystrophy group A. *Nature genetics* 36, 714-719 (2004).

4 Coin, F. *et al.* p8/TTD-A as a Repair-Specific TFIIH Subunit. *Molecular cell* 21, 215-226 (2006).

5 Bradsher, J., Coin, F. & Egly, J. M. Distinct roles for the helicases of TFIIH in transcript initiation and promoter escape. *J Biol Chem* 275, 2532-2538 (2000).

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

7 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. *Molecular cell* 38, 54-66, doi:S1097-2765(10)00211-X [pii] 10.1016/j.molcel.2010.03.004 (2010).

8 Le May, N., Fradin, D., Iltis, I., Bougneres, P. & Egly, J. M. XPG and XPF Endonucleases Trigger Chromatin Looping and DNA Demethylation for Accurate Expression of Activated Genes. *Mol Cell* 47, 622-632, doi:S1097-2765(12)00500-X [pii] 10.1016/j.molcel.2012.05.050 (2012).

9 Kraemer, K. H., Lee, M. M. & Scotto, J. Xeroderma pigmentosum. Cutaneous, ocular, and neurologic abnormalities in 830 published cases. *Arch Dermatol* 123, 241-250 (1987).

10 Kraemer, K. H. *et al.* Xeroderma pigmentosum and related disorders: examining the linkage between defective DNA repair and cancer. *J. Invest. Dermatol.* 103, 96S-101S (1994).

11 Keriel, 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, 125-135 (2002).

12 Compe, E. *et al.* Dysregulation of the Peroxisome Proliferator-Activated Receptor Target Genes by XPD Mutations. *Mol Cell Biol* 25, 6065-6076 (2005).

13 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, 127-137 (2000).

14 Compe, E. *et al.* Neurological defects in trichothiodystrophy reveal a coactivator function of TFIIH. *Nature neuroscience* 10, 1414-1422 (2007).

15 Chymkowitch, P., N. Le May, et al. (2011). "The phosphorylation of the androgen receptor by TFIIH directs the ubiquitin/proteasome process." EMBO J 30(3): 468-479.

Introduction

INTRODUCTION

Transcription is a fundamental process that plays an important role in development and cellular responses. It is defined as the synthesis of RNA from a DNA template. In bacteria and archaea, transcription of all genes is catalyzed by a single RNA polymerase, but in eukaryotic genome the different RNA polymerases pol I, pol II, and pol III transcribe different classes of genes which were discovered during 1965-1970 (Roeder and Rutter 1970). RNA pol I synthesize the 25S rRNA precursors and pol III synthesize the 5S rRNA and transfer RNAs (tRNAs). By contrast RNA pol II is responsible for the transcription of mRNAs and several small nuclear RNAs and does so using protein machinery comprising approximately 60 polypeptides. In this thesis detailed description of RNA pol II machinery is particularly discussed.

I. The transcription mechanism

The regulation of transcription is intricate, often involving interplay between promoters and various regulatory elements. A prerequisite for understanding the mechanisms and principles of this process requires identification of these elements and their relationships with several general transcription factors (GTFs). These elements include the core promoter, additional cis-acting DNA sequences such as proximal promoter elements, enhancers, silencers, and insulators.

A. The promoter elements and transcription components

A.1 The core promoter

The RNA pol II core promoter lies at the center of the transcription process. The core promoter serves as the base for the assembly of RNA pol II and all the stimulatory and repressive protein factors that are involved in the regulation of transcriptional activity. It is defined as the DNA sequence which directs the initiation of the transcription by

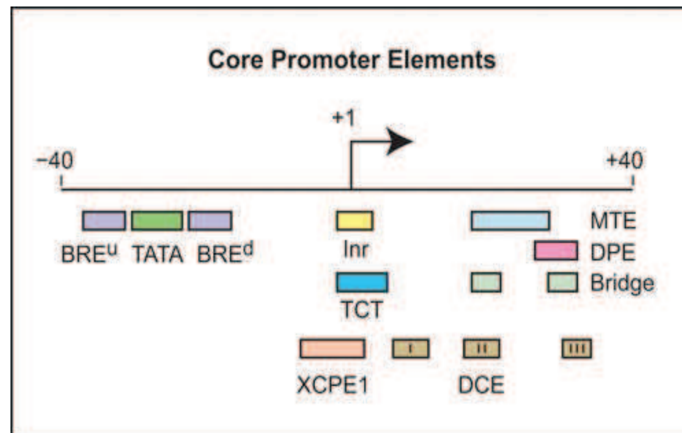


Figure 1. Core promoter motifs for transcription by RNA polymerase II.

These motifs are typically found in focused core promoters. This diagram is roughly to scale and adapted from (Tokusumi, Ma et al. 2007; Juven-Gershon and Kadonaga 2010)

Elements	Position	Consensus sequence	Binding factors
BREu	-38 to -32	SSGRCGCC	TFIIB
TATA	-31/30 to -24	TATAWAWR	TBP
BREd	-23 to -17	RTDKKKK	TFIIB
XPCE1	-8 to +2	DSGYGGRASM (human)	Unknown
Inr	-2 to +4	TCAKTY (*Drosophila*) YYANWYY (human)	TAF1/TAF2
TCT	-2 to +6	YYCTTTYY	Unknown
MTE	+18 to +27	CSARCSSAACGS	TFIID
DPE	+28 to +33	RGWCGTG	TAF6/TAF9
DCE SI SII SIII	+6 to +11 +16 to +21 +30 to +34	CTTC CTGT AGC	TAF1

Table 1. Consensus sequence and position of the different core promoter elements.

(S=C/G; W = A/ T; Y=C/ T; R=A/G; M=C/ A; K=T/G; D=T/ G/ A; N=A/ C/ G/ T). Adapted from (Juven-Gershon and Kadonaga 2010).

RNA pol II. It comprises of the transcription start site (TSS) and can extend ~ 40 bp upstream and /or downstream of the +1 position. Previously it was thought that the core promoter is a simple and a single motif but now it is clear that there is considerable diversity in the core promoter and structure. There are two kind of promoters; ***focused*** and ***dispersed*** (Smale and Kadonaga 2003). **Focused** promoters contains either single TSS or distinct clusters of short region over several nucleotides, whereas **dispersed** promoters holds number of TSS distributed over a broad region of 50 to 100 bp and typically resides in the CpG islands in vertebrates. The focused core promoters are more ancient and widespread throughout the nature than the dispersed promoter. However in vertebrates, dispersed promoters are more common than the focused promoters. The core promoter elements such as the TATA box, BREu (the *u*pstream TFII*B* *r*ecognition *e*lement), INr (*in*itiato*r*), DPE (downstream *p*romoter *e*lement), MTE (*m*otif *t*en *e*lement), DCE (downstream *c*ore *e*lement), TCT motif (polypyrimidine initiator motif) and XCPE 1 (*X* *c*ore *p*romoter *e*lement 1) are all found to be the part of focused promoter. In contrast, dispersed core promoter generally lacks the above mentioned elements (Carninci, Sandelin et al. 2006). There are however no universal core promoter elements. The elements described below are in context of focused core promoter [Figure 1 and Table 1, adapted from (Juven-Gershon and Kadonaga 2010)]

1.1 The TATA box

This sequence element was the first core promoter motif that was discovered and called as the Hogness box in old literatures. The motif has a consensus sequence TATATAA and is located 25- 30 bp upstream of the TSS. The TATA box is recognized and bound by the TATA-binding protein (TBP), subunit of the TFIID complex. Although TATA box was the first known core promoter element, it is present only in 10-15 % of the mammalian core promoters (Carninci, Sandelin et al. 2006). Most of the housekeeping genes, oncogenes, growth factors, and TFs are often TATA-less.

1.2 The Initiator element (Inr)

The initiator is the most commonly occurring motif in the focused core promoter and covers the TSS (+1). The function of Inr as different core promoter element was first observed by Smale *et al*. (Smale and Baltimore 1989). TFIID has been found to interact best with the Inr via its TBP- associated factors (TAFs). The functional analyses of mammalian promoters have shown a broader mammalian consensus YR (where R is +1) (Carninci, Sandelin et al. 2006) which is different from the classical Inr consensus sequence YYANWYY. In *Drosophila* the Inr consensus is TCAKTY, while in rice and *Arabidopsis,* a YR Inr motif (where R is +1) is observed (Yamamoto, Ichida et al. 2007).

1.3 The TFIIB recognition elements (BREu and BREd)

The TFIIB recognition element was identified as TFIIB binding sequence by Elbright and colleagues (Lagrange, Kapanidis et al. 1998). This region was initially identified as upstream of the TATA box and hence named BREu. The other sequence identified as downstream TFIIB recognition element is named as BREd. Both BREu and BREd regulate the level of basal transcription in conjunction with the TATA-box. However, BREu and BREd can have positive and negative effects on the transcription in a promoter context dependent manner (Deng, Malecová et al. 2009).

1.4 The Downstream promoter element (DPE)

The DPE is a core promoter element located downstream (+28 to +33) to A+1 in the Inr (Kadonaga 2002). It is recognized by TAF6 and TAF9 subunit of TFIID complex and is conserved from *Drosophila* to humans. The DPE consensus in *Drosophila* is RGWYVT, while in humans it has yet to be determined. DPE functions in cooperation with the Inr and the spacing between the DPE and Inr is important for the optimal transcription.

1.5 The Motif ten element (MTE)

The MTE is another functionally active core promoter element. It is located upstream of the DPE from +18 to +27 relative to the TSS and is conserved from *Drosophila* to humans. Like DPE, MTE is also a recognition site for TFIID. MTE and DPE can work in synergy but MTE can also function in cooperation with Inr independent of DPE and TATA box.

1.6 The downstream core element (DCE)

The DCE was first found in human β-globin promoter (Lewis, Kim et al. 2000). It occurs mutually exclusive of DPE. The DCE consists of three short sub elements namely SI (CTTC) from +6 to +11, SII (CTGT) from +16 to +21 and SIII (AGC) from +30 to +34.

1.7 The TCT motif (polypyrimidine initiator motif)

The TCT motif encompasses the transcription start site of nearly all ribosomal protein gene promoters in *Drosophila* and mammals. It is similar to the initiator (Inr), but is not recognized by the canonical TFIID complex (which binds to the Inr) and cannot function in lieu of an Inr. However a single T-to-A nucleotide substitution converts the TCT element into a functionally active Inr (Kadonaga 2012). TCT motif spans from -2 to +6 relative to TSS, pyrimidine nucleotide encompass the C+1 start site which is different from the canonical A/G + 1 start sites.

1.8 The X core promoter element 1 (XCPE 1)

The XCPE 1 is a rare element of promoter located from -8 to +2 relative to TSS. The sequence of XPCE1 is DSGYGGRASM (Tokusumi, Ma et al. 2007) and is present in only 1% of the human core promoters, most of which are devoid of TATA-box. It does not function by itself, instead requires the sequence specific activators such as NRF1, NF-1 and Sp1.

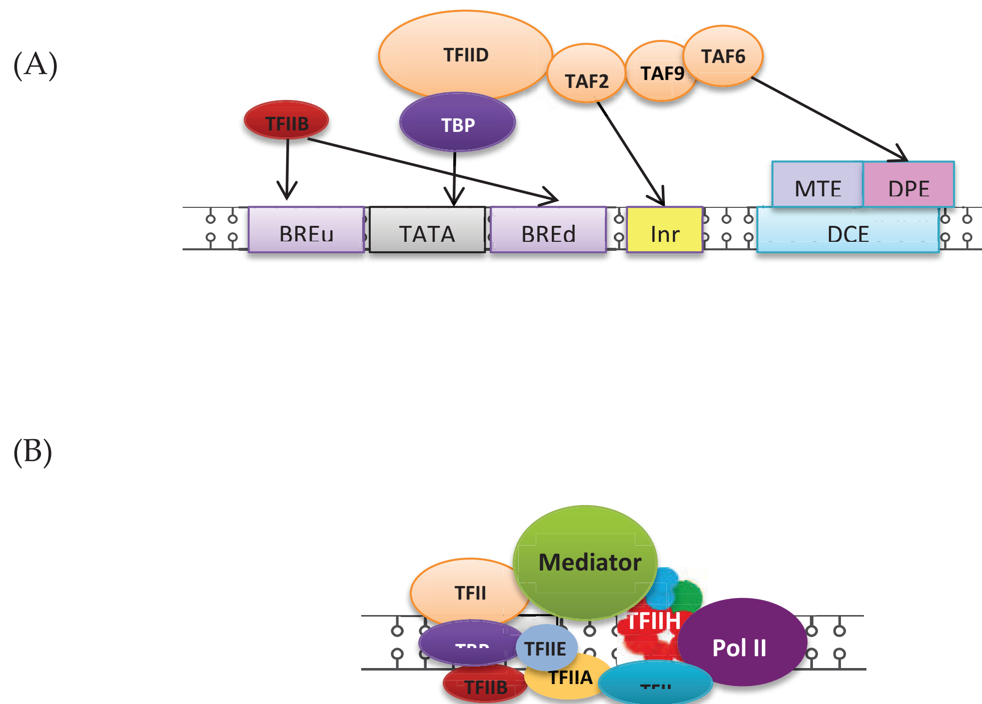


Figure 2. General transcription factors of RNA polymerase II

A) Binding of GTFs at promoter

B) The GTFs of RNA polymerase II

1.9 TATA- less promoters

Although early work suggested that TATA-box is a ubiquitous feature of RNA Pol II promoters, recent genome wide studies shows that indeed more than 80% of the mammalian promoters are TATA-less (Sandelin, Carninci et al. 2007). The majority of the eukaryotic housekeeping genes contain the promoters characterized by the CpG islands. These promoters typically lack the canonical TATA-boxes, DPEs and the Inr motifs. Bioinformatics analysis suggests that BRE elements are frequently found in the CpG+ DNA than in the CpG- DNA (Gershenzon and Ioshikhes 2005). These CpG islands have multiple binding sites for the transcription factor Sp1 and its binding further recruits the TFIID, TFIIB, RNA pol II and other basal transcription factors.

A.2 RNA polymerase II and Transcription Machinery

Transcription by RNA pol II is highly complex and tightly regulated process. RNA pol II requires several transcription factors to recognize, bind and clear the core promoters of the genes (Figure 2). These transcription factors (TFIIA, TFIIB, TFIID, TFIIE, TFIIF and TFIIH) are called general transcription factors (GTFs) and were named according to the subsequent protein fractions obtained during nuclear extract (Sawadogo and Sentenac 1990; Conaway, Hanley et al. 1991; Gerard, Fischer et al. 1991; Flores, Lu et al. 1992). The TF represents the Transcription factors, the Roman numeral indicates the RNA pol II driven transcription and the letters represents the chromatographic fraction of the specific GTFs.

2.1 RNA polymerase II and CTD modifications

Out of the three RNA polymerases, RNA pol II is the most extensively studied. RNA pol II consists of a 10-polypeptide catalytic core and the heterodimeric Rpb4/7 complex, total of 513 KDa listed in Table 2. The subunits are called Rpb1 to Rpb12, Rpb stands for RNA polymerase B, as another nomenclature system referred RNA polymerase by A, B

Sub structures		Eu-karyotes	Archaeal	Bacterial	Features	Size(KDa)
Core	Rpb1	Rpb1	Rpo1	β’	Phosphorylation site (Unique to pol II)	191.6
	Subassembly	Rpb5	Rpo5	-	Common in pol I, pol II, pol III	25.1
		Rpb6	Rpo6	Ω	Common in pol I, pol II, pol III	17.9
		Rpb8	Rpo8	-	Common in pol I, pol II, pol III	16.5
	Rpb2 Subassembly	Rpb2	Rpo2	Β	NTP binding site	138.8
		Rpb9	-	-		14.3
	Rpb3 Subassembly	Rpb3	Rpo3	Α	Promoter recognition	35.3
		Rpb10	Rpo10	Α	Common in pol I, pol II, pol III	8.3
		Rpb11	Rpo11	-		13.6
		Rpb12	Rpo12	-	Common in pol I, pol II, pol III	7.7
Stalk		Rpb4	Rpo4	-		25.4
		Rpb7	Rpo7	-	Unique to pol II	19.1

Table 2. Different subunits of RNA polymerase II in eukaryotes, archaea, bacteria and their features.

Adapted from (Thomas and Chiang 2006; Wild and Cramer 2012)

and C instead of I, II and III. Five subunits of RNA pol II are common to all the RNA polymerases. Furthermore Rpb1, 2, 3/11, 4 and 6 in yeast are conserved as subunits β′, β, $\alpha\alpha$ homodimer, σ and ω respectively in bacteria.

The assembly of RNA pol II occurs in a specific order which has been demonstrated in yeast and bacteria (Cramer, Armache et al. 2008; Decker and Hinton 2013). These studies show that the general architecture of the RNA pol II is based on the core subunits. The two large subunits Rpb1/ β′ and Rpb2/ β, form the central mass of the enzyme and lie on the opposite side of a positively charged cleft. These large subunits are stabilized by the Rpb3- Rpb11 heterodimer in yeast and $\alpha\alpha$ homodimer in bacteria. Rpb6/ω, Rpb10 and Rpb12 further stabilize the large subunits (Armache, Kettenberger et al. 2003; Wild and Cramer 2012). The assembly of RNA pol II is done fully in cytoplasm before it gets imported to the nucleus (Boulon, Pradet-Balade et al. 2010). Later dissociation and recycling of the factors occurs as their function is completed in nucleus.

The C-terminal domain (CTD) of Rpb1 is a unique feature of RNA pol II (Corden, Cadena et al. 1985) which is not present in any other RNA polymerase. The CTD consists of tandem heptad repeats with a consensus sequence of Tyr-Ser-Pro-Thr-Ser-Pro-Ser ($Y^1S^2P^3T^4S^5P^6S^7$). The number of repeats is species-dependent and generally reflects the complexity of the organism ranging from 26 in yeast, 45 in flies to 52 in vertebrates. This domain is important for the processing of mRNA (Proudfoot, Furger et al. 2002), for response of RNA pol II with the enhancers (Scafe, Chao et al. 1990; Gerber, Hagmann et al. 1995) and for the organization of transcription foci in the nucleus (Misteli 2000). Accordingly, CTD modifications can undergo dramatic changes during transcription to recruit factors needed in different phases of the transcription cycle. Out of seven amino acids, five of the hydroxylated amino acids are the potential target for the phosphorylation by several CTD kinases and CTD phosphatases [Listed in

CTD modifications	Function	Mammals	Yeast
Ser2	Kinases Phosphatase	Cdk9(P-TEFb), Cdk12/13 for subset of genes Fcp1, Cdc14	Cdk9, Lsk1, Bur1, Ctk1 Fcp1
Ser5	Kinase Phosphatase	Cdk7, Cdk8 RPAP2,Scp1,Ssu72	Kin28, Cdk8, Mcs6 Rtr1, Ssu72
Ser 7	Kinase Phosphatase	Cdk7, Cdk9 Ssu72	Kin28 Ssu72, Bur1
Tyr1	Kinase	cAbl	Unknown kinase in yeast
Thr4	Kinase	Polo-like kinase, Cdk9(not very well defined)	Not defined

Table 3. CTD kinases and Phosphatase in mammals and in yeast.

Adapted from (Egloff, Dienstbier et al. 2012; Hsin, Xiang et al. 2014)

Table 3 (Egloff, Dienstbier et al. 2012; Hsin, Xiang et al. 2014)]. Based on the phosphorylation state, RNA pol II can exist in a highly phosphorylated CTD (IIo) form and in a non-phosphorylated CTD (IIa) form. This phosphorylation occurs principally on serine 2, 5 and 7 of the CTD of the RNA pol II which is established by intricate interplay between CTD kinases and CTD phosphatases as RNA pol II moves along the genes. Analyses from a number of protein-coding genes have indicated that the CTD phosphorylation pattern changes as the transcription progresses. In fact, phosphorylated Ser5 on RNA pol II is found near the 5′ end of genes and it accumulates phosphorylated Ser2/7 towards the end of the transcription unit (Buratowski 2009). Ser5 and Ser2 phosphorylation appears to be required for all gene types, while the Ser7 phosphorylation displays gene-specific function. Other modifications include glycosylation of the serines and phosphorylation of Tyr1, however exact role of these modifications is not yet known.

2.2 TFIIA

TFIIA was originally referred to as AB or STF and was discovered as an interaction partner of TFIID (Lindahl 1974; Samuels, Fire et al. 1982). Yeast TFIIA consists of two subunits, the large subunit TOA1 and the small subunit TOA2, whereas in humans there are three subunits, TFIIAα, TFIIAβ and TFIIAγ. Previous studies shows that TFIIA was essential for *in-vitro* transcription (Reinberg and Roeder 1987), while later *in-vitro* studies shows that TFIIA was dispensable for basal level transcription (Van Dyke, Roeder et al. 1988). Some studies also show that TFIIA stimulates basal and activated transcription *in-vitro*, generally by substituting TBP with intact TFIID complex (Caldecott, McKeown et al. 1994; Hampsey 1998). Later some studies showed that TFIIA was not required for either basal or activated transcription (Petermann, Ziegler et al. 2003).

Factors	Proteins composition	Functions in RNA pol II transcription
TFIIA	TFIIAα, TFIIAβ, TFIIAγ	Antirepression, stabilization of TATA-TBP complex, coactivation
TFIIB		Selection of TSS, stabilization of TATA-TBP complex, recruitment of RNA pol II/TFIIF
TFIID	TBP, TAF1-TAF14	Core promoter binding, coactivator, protein kinase, ubiquitin activating/conjugation activity, histone acetyl-transferase activity
TFIIE	TFIIEα, TFIIEβ	Recruitment of TFIIH to the promoter, promoter escape and clearance
TFIIF	RAP30, RAP74	Selection of TSS, Recruitment of RNA pol II, TFIIE, TFIIH to the promoter, promoter escape, elongation
TFIIH	XPB, XPD, p62, p52, p44, p34, p8, Cdk7, Cyclin H, Mat1	ATPase and helicase activity for promoter opening and clearance. Kinase activity for CTD of RNA pol II and nuclear receptors
Mediator	Med1-Med31, Cdk8, cyclin C (Variable composition)	Cooperative binding with RNA pol II, kinase and acetyl-transferase activity, controls basal and activated transcription

Table 4. General transcription factors of RNA polymerase II.

Adapted from(Thomas and Chiang 2006)

Nevertheless TFIIA indeed become essential for the transcription in a reconstituted system which contains partially purified fractions. Together, these studies suggest that TFIIA has stimulatory effects probably by reversing the inhibitory effects of the negative co-factors like NC1, Dr1/NC2, Dr2/NC2, and HMG1. TFIIA stimulates transcription by stabilizing TBP binding to TATA box and is essential for the pre-initiation complex (PIC) assembly (Buratowski, Hahn et al. 1989). It becomes essential in case of TATA-less promoters in-vitro. Depletion of TFIIA has been shown to decrease the RNA pol II transcription both in TATA-containing and TATA-less promoters (Kang, Auble et al. 1995). TFIIA also communicates with several factors including TFIID, SAGA components and coactivators.

2.3 TFIIB

TFIIB was first characterized by Reinberg and Roeder in 1987. In humans, TFIIB is composed of 316 amino acids and exists as a single 33-KDa polypeptide which shares sequence homology with *Drosophila* and yeast. TFIIB consists of two domains, C-terminal domain, referred also as the core domain (TFIIBc) and N-terminal domain (TFIIBn). The C-terminal domain interacts with both the TBP and DNA encompassing the major grove upstream and the minor groove downstream to the TATA box. The N-terminal domain contains a zinc ribbon motif which interacts with the Rpb1 and Rpb2 subunits of RNA pol II and with RAP30 (RNA polymerase II associating protein 30) subunit of TFIIF. The C and N- terminal domain are connected with flexible loop which is highly conserved region called charged cluster domain (CCD) or B-finger. The B-finger is believed to regulate the conformational change of TFIIB when it interacts with DNA or activators and hence modulate its function. B-finger regulates the function of TFIIB in promoter recognition by interacting with TBP, plays an important role in recruiting RNA pol II/TFIIF at the TSS in transcription activation and hence has a

crucial role in the assembly of PIC (Orphanides, Lagrange et al. 1996; Hampsey 1998; Hahn 2004; Thomas and Chiang 2006). TFIIB can also auto acetylate itself and this modification stabilizes interaction of TFIIB and TFIIF and activates transcription both *in-vitro* and *in vivo*. TFIIB also modulates transcription after the PIC assembly through its B-finger by blocking the extension of newly synthesized RNA transcript (Bushnell, Westover et al. 2004).

2.4 TFIID

Transcription factor II D (TFIID) complex has been studied over for more than 20 years. In early literatures it was also called as DB, BTF1 and D (Matsui, Segall et al. 1980; Samuels, Fire et al. 1982). It consists of TATA-binding protein (TBP) (Hahn, Buratowski et al. 1989; Horikoshi, Yamamoto et al. 1990) and around 14 TBP-associated factors (TAFs). The TFIID complex is well conserved between different species ranging from human to yeast. It consists of three lobes, forming a horse-shoe structure, with TFIIA and TFIIB on the opposite sides of the central cavity. The TBP and some TAF components of TFIID bind different core promoter elements which classifies TFIID as a core promoter-binding factor (mentioned in Table1). The function of TBP subunit of TFIID is to contact the TATA box allowing the TFIID to recognize TATA-containing promoters. Furthermore, the interaction between TAFs and different core promoter elements (TAF-Inr, TAF-DPE, and TAF-DCE) also confer TFIID the ability to recognize TATA-less promoters. TAFs play a vital role in the selection of the promoter and co-activate the basal transcription process through recognition and binding to core promoter motifs(Burley and Roeder 1996). TAF1 and TAF2 that can bind Inr directly (Chalkley and Verrijzer 1999) and other two TAFs; TAF6 and TAF9 were reported to interact with DPE. Some complexes other than TFIID containing TAFs such as TFTC, SAGA, STAGA and PCAF have also been reported (Brand, Leurent et al. 1999). Moreover, the structure of TAF 6, 9 and 12 has many similarities with that of histone

H3, H2B and H4 suggesting the existence of an octamer structure which is similar to nucleosomes [Table 4, (Thomas and Chiang 2006)].

2.5 TFIIE

TFIIE structurally consists of two subunits, α and β, which form a α2β2 heterotetramer (Ohkuma et al., 1991). From the functional point of view TFIIE plays an essential role in the initiation of transcription by interacting directly with TFIIF, TFIIB, RNA pol II, promoter DNA and helps further to recruit the TFIIH (Maxon, Goodrich et al. 1994). TFIIE can stimulate the ATPase, the kinase and the helicase activities of the TFIIH facilitating the formation of initiation-competent RNA pol II complex (Ohkuma and Roeder 1994). TFIIE binds from -10 to +10 regions on the promoter DNA, thus initiating the promoter melting around the TSS to create a transcription bubble. Specifically, the N-terminal half of TFIIEα interacts with the TFIIEβ and RNA pol II via non-overlapping regions which results in CTD phosphorylation thus facilitating transition from initiation to elongation in basal transcription. The C-terminal region of TFIIEα interacts with p62 subunit of TFIIH (Okuda, Tanaka et al. 2004) and facilitates the entry of TFIIH in PIC assembly. Similar to TFIIEα, the N-terminal of TFIIEβ enhances the TFIIH- mediated CTD phosphorylation while the C-terminal domain is involved in the transition from transcription initiation to elongation by RNA pol II.

2.6 TFIIF

TFIIF was identified based on its strong, physical interaction with RNA pol II(Burton, Killeen et al. 1988). TFIIF is hetero-tetramer consisting of the repetition of two subunits: RAP30 (26 kDa) and RAP74 (58kDa). RAP30 subunit of human TFIIF is capable of interacting with the RNA pol II, DNA and subunit RAP74. RAP74 has three functional domains capable of interacting with RAP30, the TAF1 subunit of TFIID, TFIIB, RNA pol II and FCP1 phosphatase (Ruppert and Tjian 1995). In particular, RAP74 interact with the Rpb9 subunit of RNA pol II, which facilitates the recruitment of RNA pol II to the

promoter complex/TFIID/TFIIB (Flores, Lu et al. 1991). Furthermore, TFIIF contributes to open complex formation by enabling entry of TFIIE and TFIIH to the scaffold (Maxon, Goodrich et al. 1994; Tirode, Busso et al. 1999). TFIIF remains bound to RNA pol II during RNA elongation, enhancing its processivity and the polymerization rate (Lei, Ren et al. 1999; Yan, Moreland et al. 1999). Besides, TFIIF can be phosphorylated by protein kinase CK2 and stimulate elongation (Újvári, Pal et al. 2011). TFIIF is also implicated in the process of RNA pol II reinitiation by recruiting and enhancing the activity of phosphatase FCP1, which is required for dephosphorylation of the CTD of RNA pol II and is a prerequisite for the next round of transcription (Archambault, Pan et al. 1998).

2.7 TFIIH

TFIIH will be extensively presented in Chapter II, since the TFIIH is a major subject of the thesis and requires further discussion.

2.8 Mediator

Mediator is a large complex composed of 25-30 protein arranged in structural modules that is thought to act as a molecular bridge between DNA binding transcription factor and RNA pol II. The Mediator was discovered and first purified from *Saccharomyces cerevisiae* and is organized in four subcomplexes, the head, middle, tail, and Cdk8 modules (Poss, Ebmeier et al. 2013). Mediator also shows significant evolutionary conservation ranging from yeast to human (Malik and Roeder, 2000), but the subunit composition of mediator from different species can vary depending on the organism and the stage of cell cycle. Association of mediator with RNA pol II is inhibited strongly *in-vitro* by phosphorylation of the CTD by Kin28 which suggests that mediator is a component of the PIC and that its phosphorylation is linked to promoter escape (Guidi, Bjornsdottir et al. 2004). The mediator not only binds to RNA pol II but also binds to the transactivation domain of the large number of the transcription factors and to many

GTFs. Through these interactions, mediator is capable of promoting transcription initiation by RNA pol II and facilitates the assembly of functional PIC. The Cdk8 kinase module of mediator has also been shown to regulate transcription by targeting the CDK7/cyclin H subunits of TFIIH, thus modulating the transition from initiation to elongation by stimulating the CTD kinase activity of TFIIH (Kim et al., 1994).

Whole genome ChIP analysis has shown that mediator can bind to the enhancer region and at the promoter region, where it colocalizes with the RNA pol II (Andrau, van de Pasch et al. 2006). Mediator was also shown to be implicated in transactivation of RNA pol II (G) which is a distinct form of RNA pol II that contains the tightly associated Gdown1 polypeptide (encoded by POLR2M) (Jishage, Malik et al. 2012). Mediator is also involved in DNA looping resulting from the mediator- cohesin interaction which provides the chromosomal architecture required for the gene transcription (Kagey, Newman et al. 2010; Poss, Ebmeier et al. 2013).

Mutation in mediator lead to disorders which shows clinical features very similar to those found in TFIIH-related disease like XP, TTD and CS (Mention in Chapter III in detail)

B. The transcription cycle

Eukaryotic transcription is a precisely timed event, which can be divided into number of distinct steps: Promoter binding and preinitiation complex assembly, open complex formation, initiation, promoter clearance, elongation and termination (Hahn 2004; Laugel, Dalloz et al. 2010; Shandilya and Roberts 2012), (Figure 3).

B.1 Preinitiation complex assembly, open complex formation and initiation.

The first step in the RNA pol II mediated transcription is the binding of the gene specific regulatory factors around the transcription initiation site. These factors either can act directly by interacting with components of the transcription machinery or indirectly by recruiting chromatin modifying proteins on the transcription machinery. This assembly of the factors occurs at the core promoter via two different mechanisms; by sequential assembly pathway/the stepwise model or by the holoenzyme pathway. In stepwise model, pre- initiation complex (PIC) assembly formation occurs in sequential manner. It starts with the recognition of the TATA-flanking region by TBP and its associated factors; TFIID and then the factors such as TFIIA, TFIIB contribute to the ternary complex formation at the promoter. TFIIF, RNA pol II, further joins this ternary complex which allows the recruitment of TFIIE and TFIIH on the promoter to form a complex which is competent for transcription. Whereas, in holoenzyme pathway, the assembly of a preinitiation complex might happen in a single concerted 'recruitment step' which brings a large preassembled RNA pol II holoenzyme containing the mediator complex and most, if not all, of the GTFs to the promoter.

In the complete PIC, RNA pol II is intertwined by many interactions. TFIIB interact via its B-finger/reader segment with the RNA pol II that will ultimately provide the exit path of the nascent RNA. TFIIB also contacts other portions of RNA pol II as well as the upstream promoter DNA (Kostrewa, Zeller et al. 2009; Liu, Bushnell et al. 2010). TFIIF

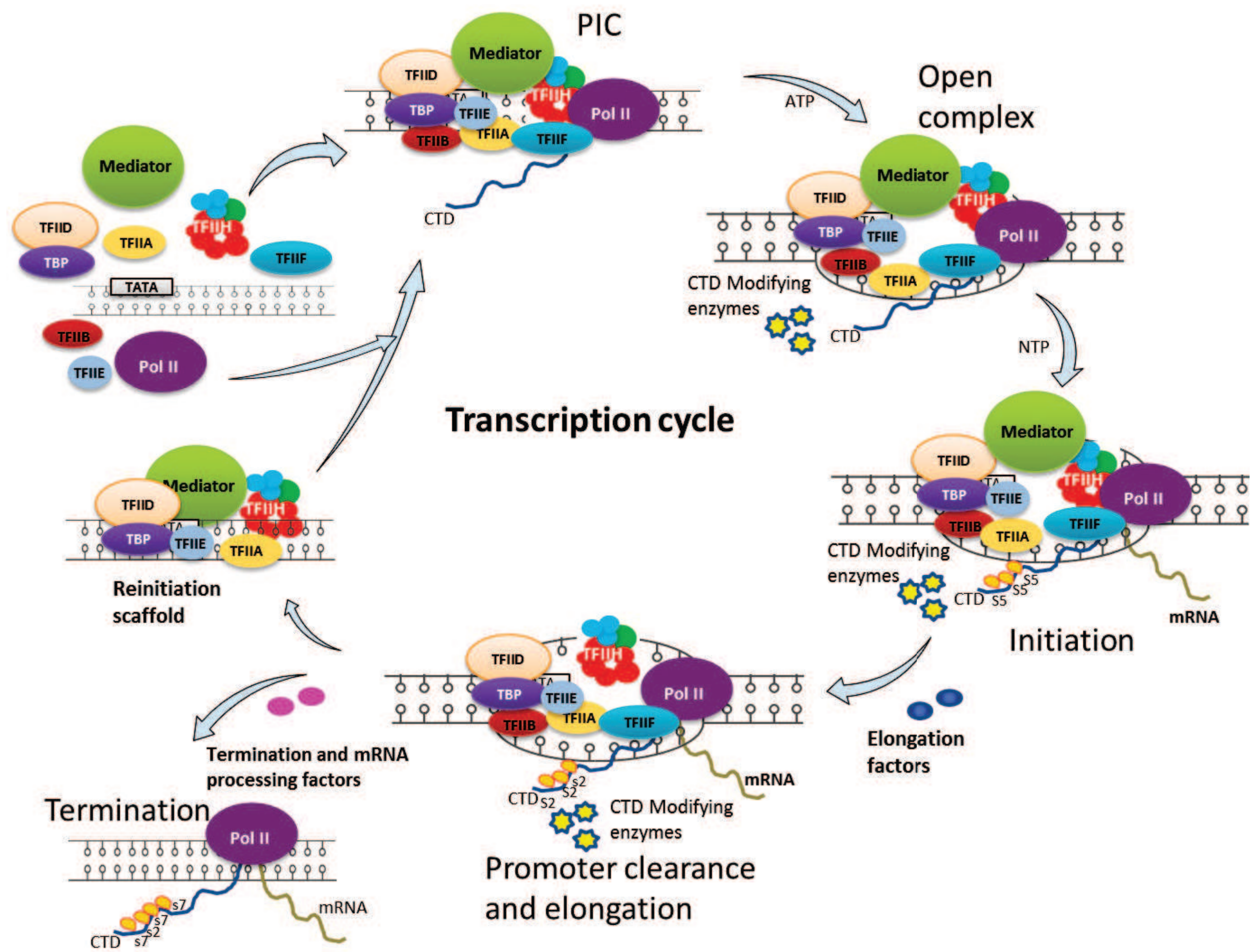


Figure 3. RNA polymerase II transcription cycle.

Adapted from (Hahn 2004; Shandilya and Roberts 2012)

and TFIIE further bind to polymerase on either side of the central cleft in which the template will ultimately reside (Lu, Zawel et al. 1992) (Ohkuma and Roeder 1994). TFIIH interact with DNA downstream of transcription start (Wang, Spangler et al. 2003). Before the promoter clearance these interactions must be lost to allow the smooth release of RNA pol II.

B.2 Promoter clearance

Once the PIC assembly is complete, TFIIH controls the ATP-dependent transition of RNA pol II from closed to open PIC in which the melted single-stranded DNA is inserted into the active site. Recent study from Eva Nogales labs show that this transition is catalyzed probably by the XPB translocase activity of TFIIH (He, Fang et al. 2013). According to this study, XPB docked directly on the downstream DNA path, between +10 and +20-bp position relative to the TSS resulting in the formation of 11-15 bp bubble around the TSS which is required for the productive transcription initiation.

B.3 Proximal pausing and elongation

Though formation of PIC is foremost importance for successful initiation, assembly of PIC at the promoter does not necessarily assure about the productive transcription. Transcripts of less than 5 nucleotides are unstable and hence results in a high abortive initiation. However, a 10 nucleotide transcript favors the promoter escape over abortive initiation and if transcript length reaches around 25 nucleotides successful initiation is achieved. Once this 25 nucleotide of nascent mRNA is synthesized, the cap structure is attached to its 5′end. The phosphorylated Ser5 of CTD of RNA pol II is further recognized by the capping enzyme which then catalyzes the addition of a methylguanosine cap to 5′ end of the nascent mRNA and productive initiation is accomplished. The CTD modification of RNA pol II and the N-terminal histone tails of the nucleosome are crucial for an active chromatin environment of a productive

initiation and mediate the elongation. CTD is phosphorylated primarily at Ser5 by Cdk7 subunit of TFIIH during transcription initiation. Before entering the elongation, RNA pol II is paused by DRB Sensitivity Inducing Factor (DSIF) and negative elongation factor (NELF), which binds to RNA pol II and inhibits its function. This negative effect can be relieved by the phosphorylation of Spt5 (largest subunit of DSIF) and NELF by Cdk9 associated with p-TEFb (Yamada, Yamaguchi et al. 2006). As the RNA pol II proceeds towards the 3′ of the gene, dephosphorylation of Ser5 occurs by Rtr1 with the simultaneous increase in Ser2 phosphorylation by Cdk9. RNA pol II is further assisted by elongation factors TFIIS, the ELL phosphatase, elongin, histone chaperone complex (FACT), histone deacetylases (HDACs). These factors remains associated with the elongating RNA pol II and maintain the 8 - 9 nucleotides RNA: DNA hybrid which is critical determinant of the processivity of the RNA pol II during the elongation.

B.4 Termination and reinitiation

Termination is the last step in the series of events of transcription cycle. This event also serves as a junction for the reloading of RNA pol II to the promoter for another round of transcription. There are two known pathways for the transcription termination; the poly (A) dependent pathway and the Sen1-dependent pathway (Kuehner, Pearson et al. 2011). The choice of the pathways depends on the RNA 3′-end processing signals and the termination factors that are present at the end of a gene. In eukaryotes, most of the protein coding mRNA precursors contain poly (A) signal, 5′-AAUAAA-3′ which is then followed by A/U rich sequence at the 3′-end of the gene. During this process, CTD acts as a scaffold for the recruitment of the several factors required for the termination. Poly (A) signal is recognized by cleavage and polyadenylation specificity factor (CPSF) which interacts with the CTD and induces pausing of elongating RNA pol II. Another factor called cleavage stimulatory factor (CstF) binds to the downstream GU-rich region and interaction between these two factors brings the endoribonucleolytic cleavage of

the nascent transcript. This cleavage is then followed by the polyadenylation of the upstream cleavage product, degradation of the downstream cleavage product by XRN2 and release of the paused RNA pol II.

Sen1-dependent pathway was first discovered in the yeast *S. cerevisiae* and interestingly now it has been shown that the termination of the several long non-coding RNAs which are classified as cryptic unstable transcripts (CUTs) and stable uncharacterized transcripts (SUTs) which do not possess poly (A) utilizes the Sen1-dependent pathway. The transduction and termination of the transcript is achieved by a distinct set of factors which include RNA-binding proteins Nrd1, nuclear polyadenylated RNA-binding3 protein (Nab3) and the putative RNA and DNA helicase Sen1.

It would not be wrong to say that *"the end is a new beginning"*. Being at two distal positions from one another the transcriptional phenomenon that takes place at the beginning and the end of the gene are often studied as two separate processes. However several studies have shown that the RNA pol II machinery at promoter and terminator are entangled together (Hampsey, Singh et al. 2011) . Indeed, the release of the RNA pol II requires the reversal of the associated covalent marks on the CTD repeats which will bring RNA pol II to its original hypo phosphorylated state for a subsequent round of reinitiation. The released RNA pol II is competent for a fresh round of transcription. However, sometimes several GTFs remain associated to promoter and act as the scaffold for the reinitiation. The activator dependent interaction of the promoter and terminator region has been shown to be involved in transcription re-initiation via gene looping (El Kaderi, Medler et al. 2009). TFIIB directs the assembly of such reinitiation scaffold to the promoter by interacting with CPSF and Cstf complexes. This interaction is conserved in mammalian system and is regulated by the phosphorylation of the TFIIB (Wang, Fairley et al. 2010). An essential role for the TFIIH kinase subunit, Cdk7 in the formation of gene loops has also emerged. Cdk7 is known

to regulate phosphorylation of Ser5 and Ser7 of CTD of RNA pol II. While phospho-Ser5 CTD is required for transcription initiation, phospho-Ser7 has been linked with termination and 3′ processing and also with promoter-bound paused RNA pol II. Inhibition of Ser7 phosphorylation resulted in the loss of RNA pol II pausing, both at the promoter and termination sites suggesting the implication of Cdk7 in gene looping (Glover-Cutter, Larochelle et al. 2009). In addition to TFIIB, TFIIH and PC4 (positive cofactor 4) are the some other factor known to promote reinitiation via gene looping through its interaction with 3'end of the gene (Shandilya and Roberts 2012).

C. The activated transcription

During the development of an organism there are several proteins which are required for the basic cell maintenance and thus they are expected to be expressed in all cells of an organism under normal conditions, irrespective of tissue type, developmental stage, cell cycle state and external signal. These genes are called as housekeeping genes. However, there are large numbers of genes whose expression is required only at specific time and hence they are regulated by certain endogenous or exogenous stimulus (such as cellular differentiation and stress response). Hormones are one such example of a stimulus which interact with the specific nuclear receptor and further regulate the expression of nuclear receptor targeted genes.

C.1 Overview of the nuclear receptor families

Nuclear receptors (NRs) are the most abundant class of transcription factors that regulate diverse biological functions such as homeostasis, reproduction, or development. They function as a ligand- activated transcription factors, and thus provide a link between the signaling molecules which control these processes and the transcriptional responses. NRs share a common structural organization, usually consists of five functional regions: the A/B region that contains an N-terminal activation

function-1 domain (NTD), the central C region that contains a DNA-binding domain (DBD), the C-terminal E region that contains a ligand-binding domain (LBD), and the D hinge region that connects the DBD and the LBD. NTD is a highly variable region which contains at least one constitutionally active transactivation region (AF-1) and several autonomous transactivation domains (AD). DBD is highly conserved region containing the two zinc fingers that recognizes specific NR- responsive elements on the regulatory region of the target genes. Whereas, LBD plays a crucial role in ligand-mediated nuclear receptor activity. Besides, its role in ligand recognition, the LBD also contains an activation function-2 (AF-2) domain, whose action is highly dependent on the bound ligand.(Mangelsdorf, Thummel et al. 1995). The hinge D, together with C-terminal E region, is less conserved and show distinct structural features among different nuclear receptors.

Depending on the nature of the ligand they bind, the nuclear receptor superfamily is divided into three classes (i) Class I include receptors for steroid hormones in the DNA-binding homo-dimeric form and often located in cytosol. Once they bind to the ligand they are translocated to the nucleus. It includes the glucocorticoid receptor (GR), estrogen (ER) or progesterone (PR). (ii) Class II receptor binds mostly as heterodimers with RXR to DNA and are usually located in the nucleus. They include mainly RXR receptors (X receptor with retinoic acid), the thyroid hormone receptor (TR), retinoic acid (RAR) and vitamin D (VDR). (iii) Receptors of class III includes family of orphan receptors for which ligands are still not very well known. The members of this group include the liver receptor homologue-1 (LRH1), reverse ERBA-α/-β, chicken ovalbumin upstream transcription factors (COUP-TFs), receptors SF-1.

C.2 Retinoic acid receptor

Nuclear retinoic acid receptors (RARs) are transcriptional regulators which controls the expression of specific subsets of genes in a ligand-dependent manner. The compounds

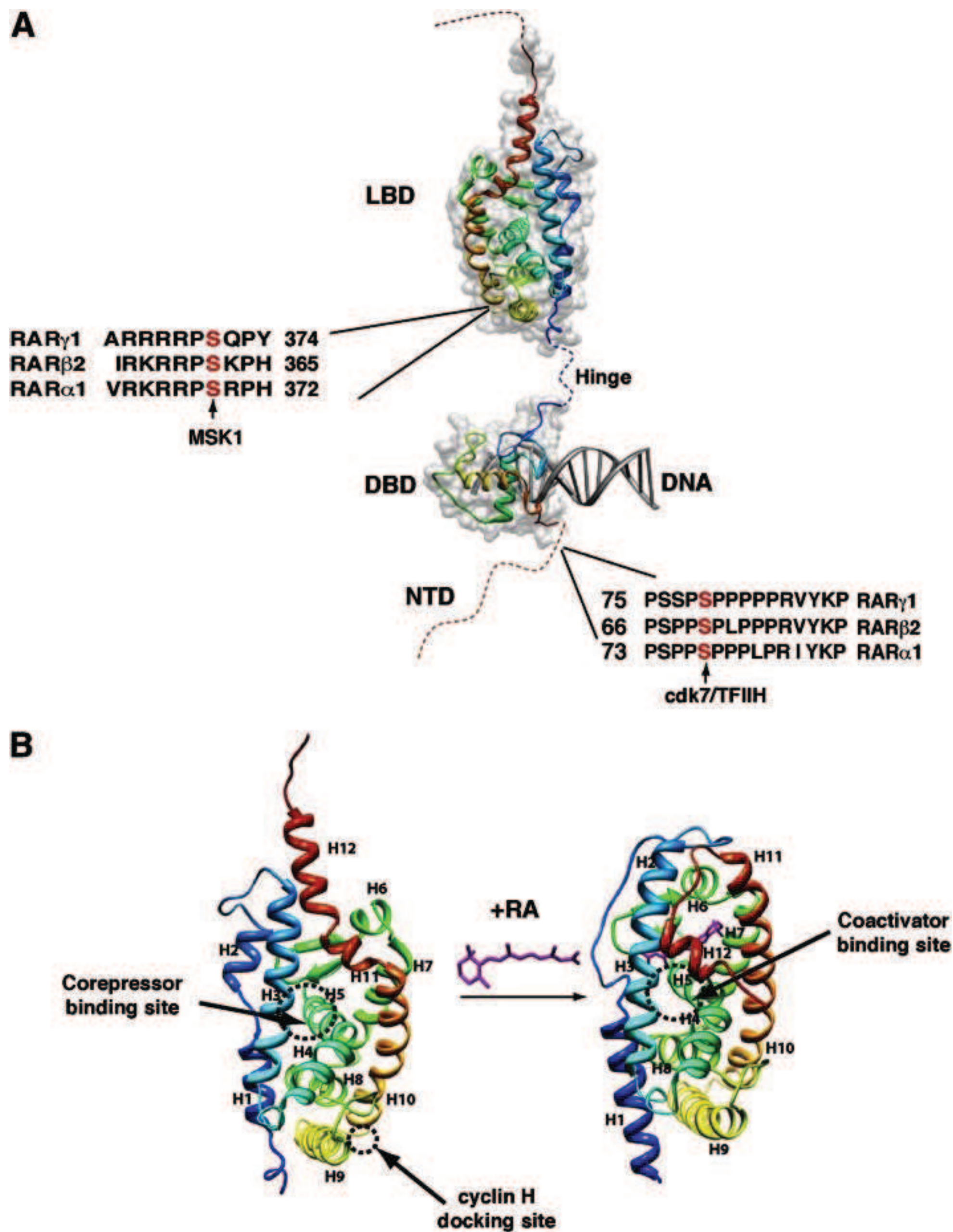


Figure 4. Structure of retinoic acid receptor

Adapted from (Al Tanoury, Piskunov et al. 2013)

(A) RARs depict a domain organization with an unstructured NTD and two well-structured domains: a central DBD and a C-terminal LBD. The phosphorylation sites located in the NTD and the LBD are shown. (B) Structural changes induced upon RA binding. The crystal structures of the unliganded RXRα and liganded RARγ LBDs are shown. Helices are represented as ribbons and labeled from H1 to H12. The binding domains for corepressors/coactivators and for cyclin H are shown.

that activate the RARs are referred to as retinoids. Retinoids are small size lipid soluble, hydrophobic molecules that can easily cross the lipid- bilayer of the cell membranes. Retinoids includes both compounds which are structurally related to vitamin A, as well as compounds that exhibit biological vitamin A activity. Vitamin A is an essential component for human life cycle, starting from embryonic development, organogenesis to immune functions and reproduction. Indeed deficiency of Vitamin A can lead to neonatal growth retardation and a large number of congenital malformations (Samarut and Rochette-Egly 2012). Natural retinoids, such as *all-trans* retinoic acid (All t-RA), are produced *in vivo* from oxidation of vitamin A (Chambon 2005). This newly synthesized retinoic acid then binds to cytosolic cellular retinoic acid-binding proteins; CRABP-I and CRABP-II to enter the nucleus (Delva, Bastie et al. 1999; Budhu and Noy 2002). Subsequently it regulates the gene expression by activating the nuclear receptor RAR which exists in three subtypes : α (NR1B1), β(NR1B2) and ϒ(NR1B3) (Chambon 1996) encoded by three separate genes. For each subtypes there are at least two isoforms which are generated by differential promoter usage and alternative splicing leading to the difference only in their N-terminal ends (Germain, Staels et al. 2006) . *In vivo*, RAR transduce RA signal as heterodimer together with RXR which also exist as three subtypes: α (NR2B1), β (NR2B2) and ϒ (NR2B3) and are involved in regulating genes implicated in cell differentiation, proliferation and in apoptosis.

RARs and RXRs have a well-defined domain organization and structure. It consist mainly of a variable N-terminal domain (NTD) and two highly conserved domains, a central DNA-binding domain (DBD) and a C-terminal ligand-binding domain (LBD) bridged by a flexible hinge peptide [Figure 4, (Al Tanoury, Piskunov et al. 2013)].

DBD binds to the DNA in a sequence specific manner and is composed of two zinc-finger domains, two α-helices and a COOH-terminal extension (CTE)(Zechel, Shen et al. 1994; Smale and Kadonaga 2003). DBD includes several highly-conserved sequence

elements, referred to as P, D, T and A boxes which contribute to the specificity of the response elements and to a dimerization interface which is involved in binding with the DNA backbone (Germain, Staels et al. 2006). RARs and RXRs form an asymmetric heterodimer which binds to the RA response elements (RAREs).The RAREs are located in the regulatory region of the target gene and are composed of two direct repeats of core hexameric motif PuG (G/T) TCA(Mangelsdorf and Evans 1995). A classical RARE is a 5bp-spaced direct repeat (referred to as DR5), but RAREs with direct repeats separated by 1bp (DR1) or 2bp (DR2) are also found. Nevertheless, RXR homodimers can also bind to DR1.

The LBD is composed of 12 conserved α-helices and a β-turn which is separated by the loops and folded into a three layered, parallel helical sandwich. The LBD is functionally very complex as it contains ligand binding pocket which binds to the ligands and is involved in dimerization and interaction with several other co-regulators. The ligand binding pocket contains hydrophobic residues mainly in helices H3, H5, H11 and β-hairpin (Bourguet, Andry et al. 2000). The LBD also contains a flexible C-terminal helix H12 which can change its conformation after the ligand binding depending on one to other subtypes. Moreover, cyclin H, a subunit of CAK subcomplex of TFIIH is also known to bind to the AF-2 domain of LBD and directs the phosphorylation of AF-1 domain (Yamamoto, Ichida et al. 2007), (Figure 4).

Unlike DBD and LBD, NTD is highly variable and is not conserved even between the different subtypes and the forms of RARs and RXRs. Structural algorithms prediction suggest that the NTDs of RARs and from other NRs family have naturally disordered structure(Warnmark, Treuter et al. 2003) which provides flexibility perhaps required for the kinases and ubiquitin-ligases (Dyson and Wright 2005). In addition, NTDs of RARs and RXRs also contain phosphorylation sites which are conserved between the RARs.

C.3 Transcription of retinoic acid receptor target genes

The mechanism of transcriptional regulation of RA-target genes is a sequential process and completely relies on the binding of RARs to the RAREs. RARs can control transcription both by repression and activation which depends on the association and dissociation of certain corepressors or coactivators interacting with the hydrophobic surface of the LBD generated by the H3, H4 and H12 helices. According to the canonical transactivation model for RA- targeted genes (Dilworth and Chambon 2001), the RAREs are occupied by a RAR/RXR heterodimer in an unliganded state, which maintains the chromatin in a condensed or a repressed state (Figure 5). In an unliganded state the hydrophobic surface of RAR interacts with the corepressors [such as nuclear receptor corepressor (NCoR) and silencing mediator of retinoic acid and thyroid hormone (SMRT)]. The RAR and corepressor SMRT interaction occurs between the LBD of RAR and the receptor interaction domains (RID) of SMRT. These corepressors serve as a platform for the recruitment of the histone deacetylases (HDAC) which deacetylate the lysine of histones, thus maintaining the repressed chromatin state. Polycomb group of proteins are also known to interact with RAR/RXR heterodimers and act as a corpressor (Gillespie and Gudas 2007). Transcription is turned on upon ligand addition which induces the release of the corepressor and subsequently facilitate the binding of the coactivator to the RAR/RXR heterodimer which in turn allows the recruitment of several other complexes such as histone methyl transferases (HMT), histone acetyl transferases (HAT), histone demethylases and DNA-dependent ATPases and other chromatin remodelers (Perissi, Jepsen et al. 2010). These complexes further allow the histone modifications leading to the chromatin decompaction. These activated RARs then recruit the transcriptional machinery which includes mediator, RNA pol II, GTFs and also the NER factors to achieve chromatin remodeling for the optimal RNA synthesis (Le May, Mota-Fernandes et al. 2010; Le May, Fradin et al. 2012).

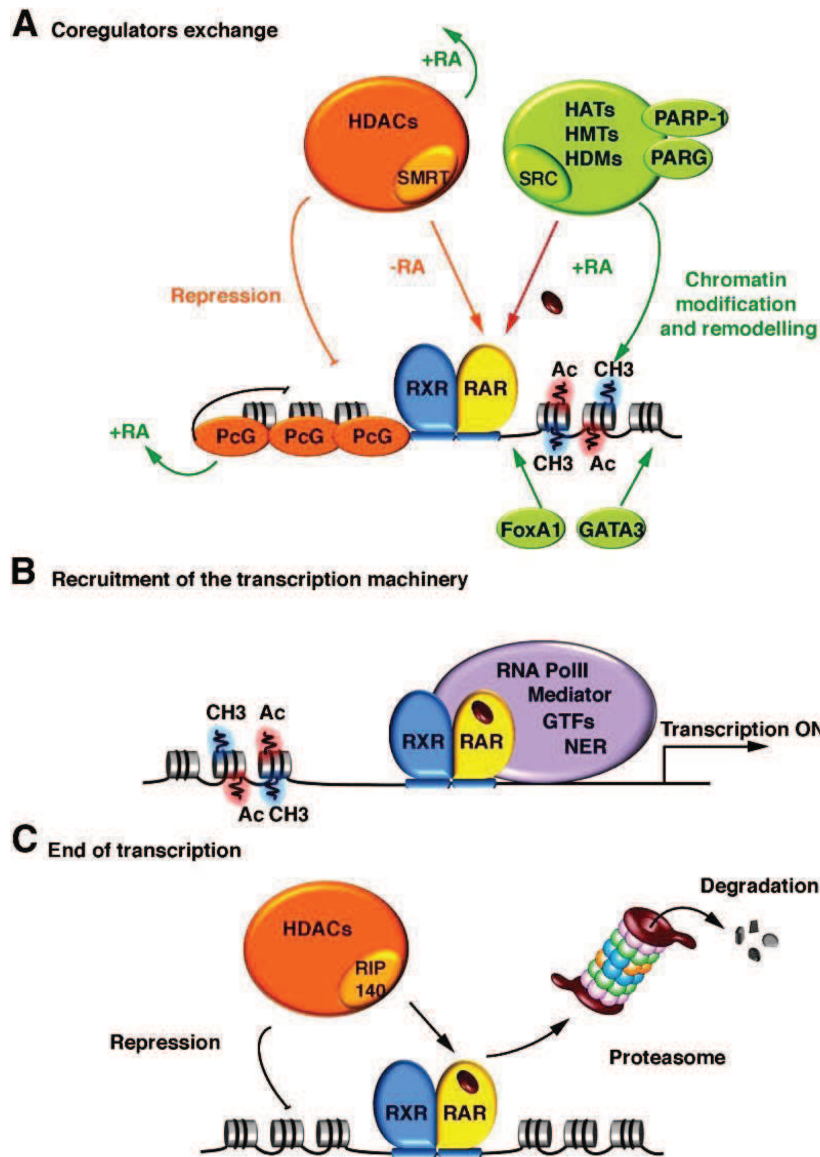


Figure 5. Coregulator exchange at RXR/RAR heterodimers.

(A) In the absence of ligand, RARα/RXR heterodimers bound to DNA are associated with corepressor complexes. Upon ligand binding, the corepressors dissociate, allowing the recruitment of coactivators and large complexes with enzymatic activities that decompact repressive chromatin. (B) When chromatin is decompacted, the transcriptional machinery, consisting of the Mediator, RNA pol II, the general transcription factors (GTF), and the nuclear excision repair (NER) factors, is recruited to the promoter, resulting in the initiation of transcription. (C) Transcription ends with the recruitment of nonconventional coactivators, such as RIP140, associated to large complexes with chromatin-repressing activity and/or through the degradation of RARs by the ubiquitin-proteasome system. Adapted from (Al Tanoury, Piskunov et al. 2013).

However, the importance of NER factors at promoter is still not very well understood. It remains to be investigated that how these NER factor contribute to chromatin remodeling, do they interacts with the GTFs, how they regulate the histone modifications or the DNA methylation on NRs targeted genes.

At the end of transcription, the liganded RARs recruit unconventional coregulators such as RIP60 (receptor interacting protein) and PRAME (preferentially expressed antigen in melanoma). These coregulators inhibit the transcriptional activity by further allowing the recruitment of HDACs and PcG proteins and hence limit the RAR activity through a functional feedback mechanism (Figure 5). In addition to direct effects, RA also targets number of signaling cascades. RA activates several kinase cascades such as p38 mitogen-activated kinase in fibroblasts, mouse embryo carcinoma cells, mammary breast tumor cells, and leukemia cells. It also activates p42/44 extracellular signal regulates kinases (Erk) in neurons, Sertoli cells and in embryonic stem cells. It also activates phosphoinositide 3-kinase (PI3K) and/or protein kinase B (PKB)/Akt pathway. RARs and RXRs are also known as phosphoproteins (Al Tanoury, Piskunov et al. 2013). RARs are phosphorylated at serine 369 in LBD probably by c-AMP dependent kinase or MSK1 and at serine 77 in NTD by cdk7 subunit of TFIIH.

Besides, phosphorylation of RARs and RXRs, signaling pathways activated by the RA are known to phosphorylate other factors involved in the gene regulation, which may not be the RAR targets [very well reviewed in (Keriel, Stary et al. 2002; Rochette-Egly 2014)], (Figure 6). As an example, upon recruitment to RARα target promoters MSK1 and phosphorylates histones H3. Based on the studies of several reports, histone phosphorylation could contribute to transcription as a chromatin mark responsible for, in cooperation with other histone modifications, chromatin remodeling and promoter recruitment of RXR/RAR heterodimers and the transcriptional machinery (Figure 6). At the end of the RA signal, phosphorylation of RARs controls the recruitment of

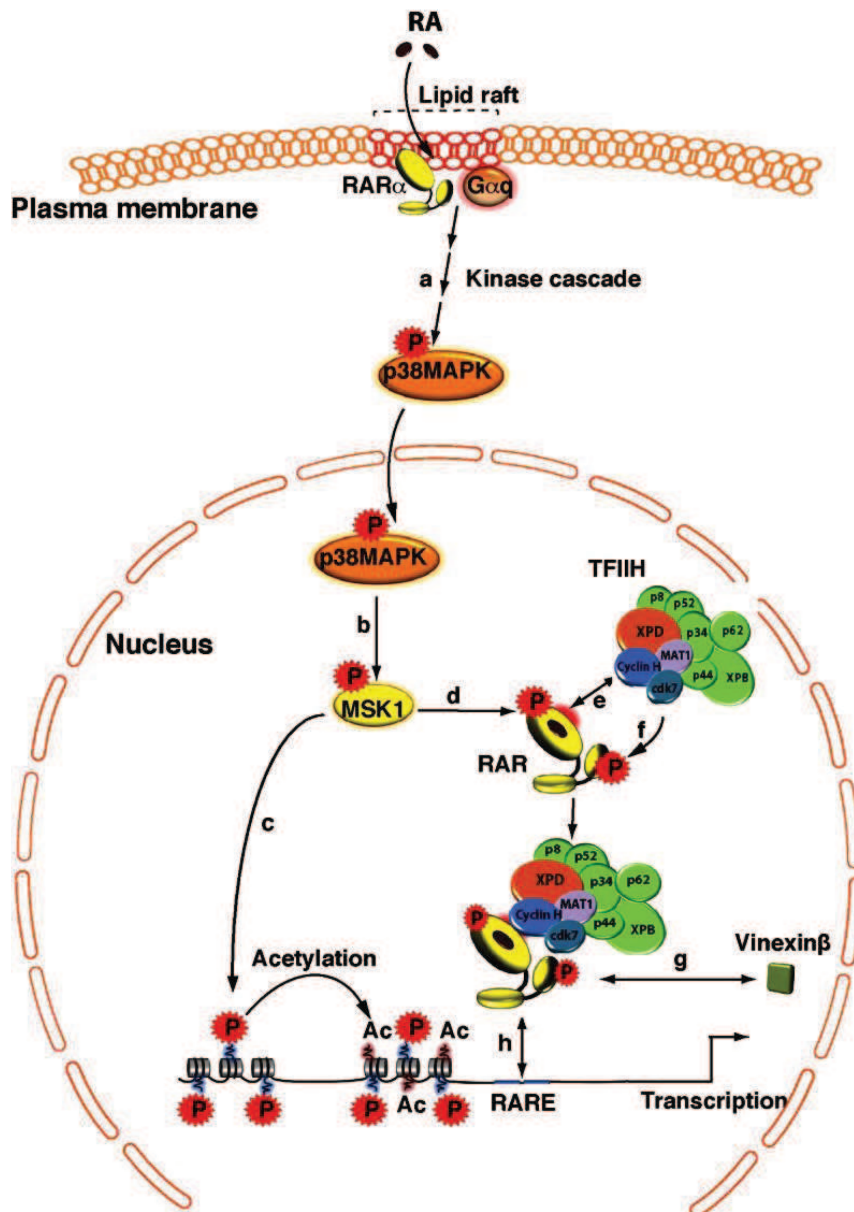


Figure 6. Crosstalk between the RA-activated p38MAPK pathway and the expression of RAR target genes.

In response to RA, p38MAPK is activated (a), and then translocates into the nucleus and phosphorylates MSK1 (b). Activated MSK1 phosphorylates histones (c) and RARα at a serine located in the LBD (d). Subsequent to conformational changes, the cyclin H subunit of the CAK subcomplex of TFIIH is recruited to an adjacent domain (e), allowing the formation of a RARα/TFIIH complex and the phosphorylation of the NTD by the cdk7 kinase (f). In the case of the RARγ subtype, phosphorylation of the NTD promotes the dissociation of coregulators, such as vinexinß (g). Finally, phosphorylated RARα is recruited to response elements located in the promoter of target genes (h)

Adapted from (Al Tanoury, Piskunov et al. 2013)

ubiquitin-proteasome system. It can be mediated through several factors such as such *Pin1*, SRC-3. At the end RA can also activate PPARβ/δ. There are evidences which also retinols activate JAKs-STAT pathway to regulate the expression of target genes such as the suppressor of cytokine signaling 3 (SOCS3) and PPAR ϒ.

D. Chromatin and gene regulation

In a cell genetic information is encoded by DNA, which is packaged into hierarchically organized complex structure called chromatin. Chromatin is composed of DNA, histone and non-histone proteins. The genome undergoes several level of compaction to fold two meters of DNA into a nucleus with a diameter of around six micrometers. It may seem contradictory that proteins are added to DNA to make it more compact but chromatin can be packaged into a much smaller volume than DNA alone. This is achieved as the histones are positively charged and DNA is negatively charged thus this electrostatic interaction provides energy to fold the DNA which can then fit into the nucleus. The first level of compaction consists of DNA coiled around a core of histone proteins creating a fundamental unit called nucleosome. Nucleosome is composed of an octamer of the four core histones (H3, H4, H2A, H2B) around which 147 base pairs of DNA are wrapped. The packaging of DNA into nucleosomes shortens the fiber length about sevenfold which resembles string of beads. The DNA is then subjected to a second degree of folding which further compact the chromatin into 30 nm solenoid structure (Woodcock and Dimitrov 2001). Histone H1 which forms the linker is very important for stabilizing the higher order chromatin structure (Figure 7). Fundamental processes such as transcription, repair and replication occur with the separation of two DNA strands and hence hindered by nucleosomes, chromatin folding and compaction. Therefore, it becomes essential for the cell to modify or remove these nucleosomes transiently to allow transcription and replication to progress. Depending on the compaction level chromatin can be categorized as euchromatin and heterochromatin. The euchromatin is typically enriched in acetylated histones (for example and H3K9Ac, H4K16Ac) and also shows the histone H3K4 methylation (H3K4Me) (Grunstein 1997; Bannister and Kouzarides 2011). Instead, heterochromatin is defined as a more condensed structure, inaccessible to transcription factors. It is characterized by hypo-

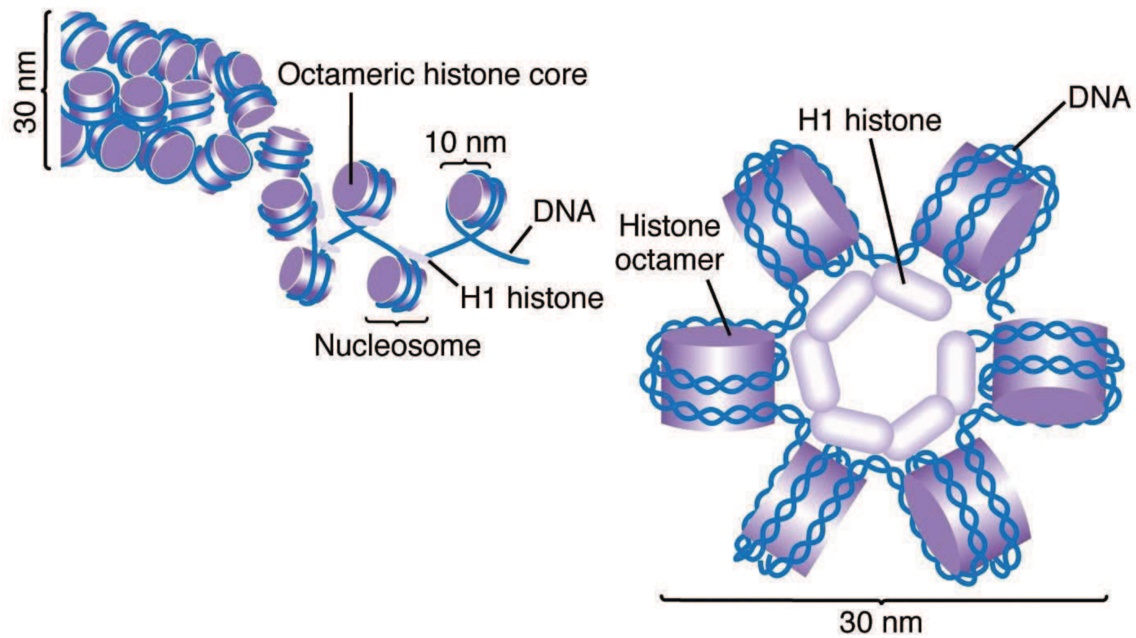


Figure 7. Nucleosome organization.

The nucleosome is the basic unit of chromatin, and consists of a protein octamer containing two molecules each of histones H2A, H2B, H3, and H4, around which 147 base pairs of nuclear DNA is wrapped. Courtesy http://bscb.org

histone acetylation and methylation of histones some specific (eg H3K9Me) (Rea, Eisenhaber et al. 2000; Jenuwein and Allis 2001)

D.1 Epigenetic modification of chromatin

Epigenetics is defined as heritable changes in gene expression that occur without a change in DNA sequence. Epigenetics is recognized as key determining factor in normal development and differentiation, and epigenetic abnormalities are relevant in many diseases, including various types of neurological disorders and cancers. Key components in the processes of epigenetic transcription regulation are DNA methylation, histone modifications and variants, non-histone chromatin proteins, small interfering RNA (siRNA) and micro RNA (miRNA). Post-translational modifications (PTMs) of histones and DNA methylation are discussed in details in this manuscript.

1.1 Post-translational modification of histones

The transition from transcriptionally silent heterochromatin to gene expressing euchromatin is mediated by distinct histone variants and the posttranslational modifications of histones at their N-terminal tails. These covalent modifications of histones are referred to as "histone code". These modifications include acetylation of lysines, methylation of lysines and arginines, phosphorylation of serines and threonines, ADP-ribosylation of glutamic acids, and ubiquitinylation and SUMOylation of lysine residues (Some important histone PTMs are shown in figure 8). These modifications are controlled by balance of the enzymatic activities of various proteins which add ("writers") or remove ("erasers") a specific PTM (Ruthenburg, Allis et al. 2007; Thompson, Guppy et al. 2013), while readers recognizes specific modifications resulting in the recruitment of the transcriptional machinery. Regulation of transcription by each histone modification is discussed in more detail below.

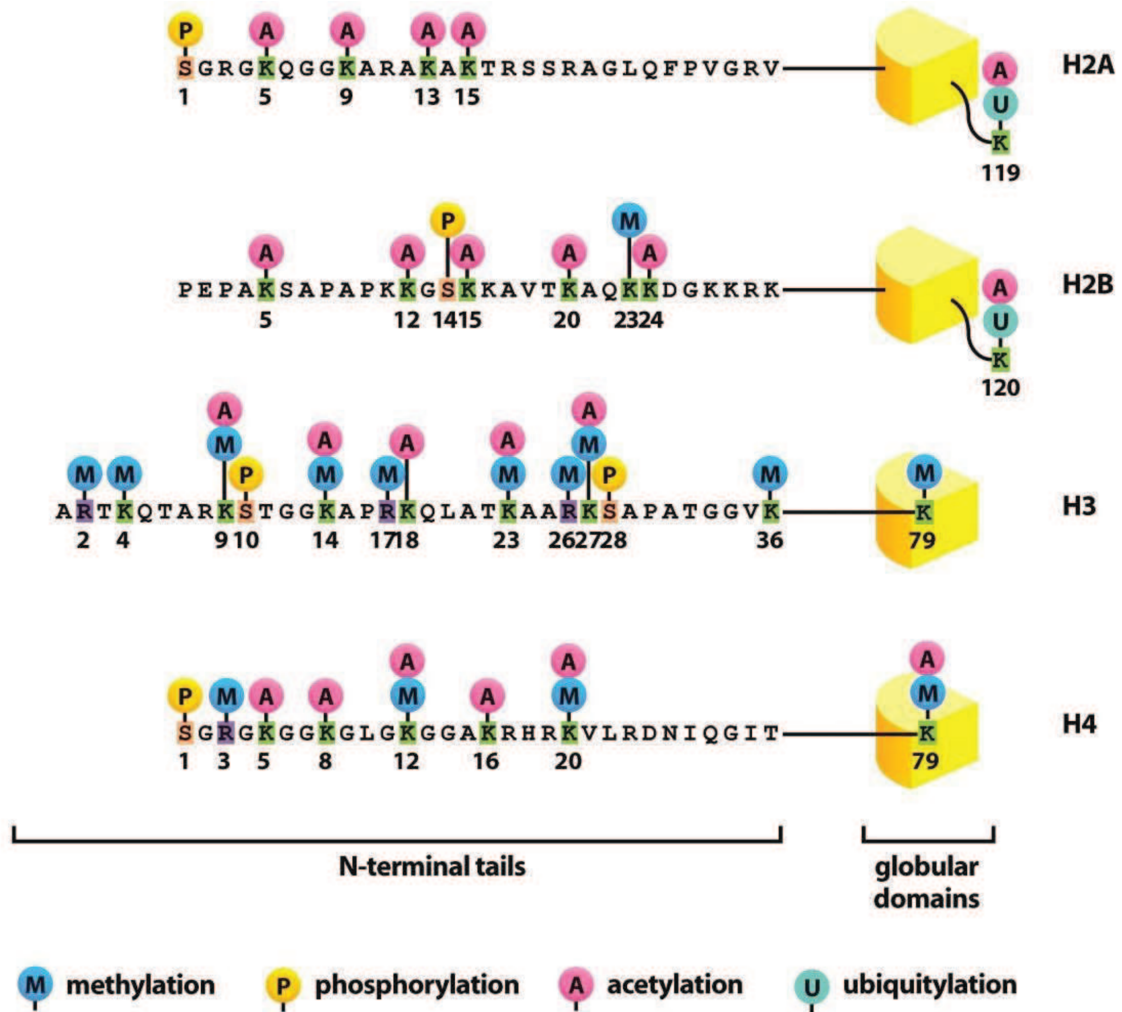


Figure 8. Schematic cartoon of core histone PTMs and their potential function

(Courtesy http://bscb.org)

1.1.a. Regulation by histone lysine acetylation and deacetylation.

The first covalent modification discovered was the acetylation of lysine(lys or K) (Allfrey and Mirsky 1964), regulated by writers of acetylation, the histone acetyletransferases (HATs) and the erasers of acetylation, the histone deacetylases (HDACs). The HATs utilize acetyl CoA as cofactor and catalyzes the transfer of an acetyl group to the ε-amino group of lysine side chains and hence neutralizes the positive charge of Lys (K), leading to the weakening of the DNA-histone interaction and subsequent activation of transcription. In addition, acetyl lysine recruits other chromatin modifiers containing a bromodomain that recognizes an acetyl lysine to activate transcription (Dhalluin, Carlson et al. 1999). Major acetylation sites on histone H3 include K9, K14, K18, K23 and K27, and on H4 include K5, K8, and K12 (Kouzarides 2007; Bannister and Kouzarides 2011). All acetylation marks are essentially correlated with transcriptional activation which is localized to TSS and/or enhancers of potentially actively transcribed genes. Indeed, the transcriptionally active regions are rich in acetylated histones, while their hypo-acetylation is found in heterochromatin. The HATs are usually classified into two groups: Type A and type B. These include enzymes such as GCN5, PCAF or HAT1, CBP/p300 etc. These enzymes are often found as a part of multi-protein complex interacting with several transcription factors and hence affect the gene regulation. For example, GCN5/PCAF HATs are responsible H3K9 acetylation, and CPB/p300 are responsible for H3K18/27 in nuclear receptors mediated transcription (Jin, Yu et al. 2011).

The HDACs works in contrast to the HATs and reverse the acetylation of histones by restoring back the positive charge on the lysine residues. HDACs are also classified into 4 classes designated as HDAC I, II, II and IV. Class I, includes HDAC1, 2, 3 and 8; class II, includes HDAC 4, 5, 6, 7, 9 and 10; class IV as HDAC11 and class III contains the NAD+ dependent Sir2-like deacetylases or sirtuins. Except HDAC 8, rest all HDACs

exist in high molecular weight multi-protein complexes. HDACs activity is tightly regulated with the other protein partners and histone PTMs. Dysfunction of these HATs and HDACs enzymes is often associated with diseases, ranging from neurodegenerative disorders to cancer (Rothbart and Strahl 2014).

1.1.b. Regulation by histone lysine/arginine methylation and demethylation

Similar to acetylation several lysine residues of histones can also be mono, di or tri-methylated. The level and the location of histone modification further add complexity in gene regulation during different processes such as transcription, DNA repair, replication and recombination. Unlike the acetylation and phosphorylation, histone methylation does not alter the charge of the histone proteins; however it does affect the hydrophobic and steric properties. Histone methylation can be implicated both in transcriptional activation and repression, in contrast to the acetylation which is associated with activation only. Histone H3 can be methylated at various positions such as methylation of K4 and K36 which is associated with activation (H3K4me and H3K36me) and at K9 and K27 which is linked with repression (H3K9 and H3K27) of transcription (Zentner and Henikoff 2013). However, some of the developmental genes resides in bivalent chromatin regions and contains both repressive H3K27me3 and active H3K4me3 marks. The enzymes that catalyze the transfer of methyl group from S-adenosylmethionine (SAM) to a lysine's ε-amino group are known as histone methyltransferases. The first histone lysine methyltransferase (HKMT) identified was SUV39H1 and it targets H3K9 methylation (Rea, Eisenhaber et al. 2000). Methylation of H3K4 is catalyzed by SET domain of KMT2A (K-specific methyltransferase 2A, commonly called MLL), whereas H3K27 is methylated by PRC2 (polycomb repressive complex 2). Regulation of transcriptional activation or repression by H3K4 or H3K27 is still obscure. It has been shown that H3K4 methylation recruits the BAF chromatin remodeling complex *via* its chromodomain to activate transcription, similarly

trimethylated H3K27 recruits PRC- complex *via* the chromodomain-containing protein CBX1 (chromobox homolog 1) and induces the compaction of chromatin, resulting in transcriptional repression (Wysocka, Swigut et al. 2006; Simon and Kingston 2013). Like addition, removal of histones methylation is also a highly regulated event and catalyzed by several histone Lys-specific demethylases (KDMs). H3K4 is demethylated by KDM1 (commonly known as LSD1) and KDM5B (JARID1), while H3K27 is demethylated by KDM6A (UTX) and KDM6B (JMJD3)(Black, Van Rechem et al. 2012).

In addition to lysine, arginine residues of histones can also be methylated. Arginine methylation is catalyzed by a group of proteins known as Arginine methyltransferases (PRMTs). The PRMTs transfer a methyl group from SAM to the ω-guanidino group of arginine within a variety of substrates.

1.1.c. Regulation by other histone modifications

The dictionary of the histone PTMs has expanded vastly since it was first identified by Vincent Allfrey. Besides, well known acetylation and methylation of the histones there are several other modifications which include phosphorylation, ubiquitylation, glycosylation, ADP-ribosylation, carbonylation, SUMOylation, citrullination and many others. The later mentioned PTMs are low in abundance as compared to the more well-studies modifications suggesting that their roles in physical and functional cross-talks may be subtle. Histone phosphorylation which occurs on serine (H3S10) and threonine residues can influence transcription, chromosome condensation, DNA repair and apoptosis. The level of the modification is controlled by the kinases and the phosphatases that add and remove the phosphate group from ATP to the hydroxyl group of the target amino-acid side chain. Ubiquitination is another histone modification which results in a much larger covalent modifications compared to the relatively small molecular change to amino acids side chains. Ubiquitin functions by attaching itself to the lysine via sequential action of the three enzymes, i.e. E1-

activating, E2-conjugating and E3-ligating enzymes. Ubiquitination is highly dynamic and the degree of ubiquitination can vary depending on the substrate specificity. Recent advances have defined critical roles of histone ubiquitination in transcriptional regulation and DNA repair. The writers, erasers, and readers of histone ubiquitination have also been linked to cancer development. Mono-ubiquitination of K119 on H2A is known to be involved in gene silencing while mono- ubiquitination of K123 on H2B plays a crucial role in transcriptional initiation and elongation. In addition to H2A and H2B, core histones H3, H4, and linker histone H1 have also been reported to be modified by ubiquitin. For example, H3 and H4 were polyubiquitinated in vivo by CUL4–DDB–RBX1 ubiquitin ligase complex after UV irradiation (Cao and Yan 2012). SUMOylation is a modification similar to ubiquitination and involves the histones lysines via E1, E2 and E3 enzymes. It can be detected in all the four core histones and functions seemingly in opposite to the acetylation and ubiquitination. Glutamate and the arginine residues of histones are also subjected to either mono or poly-ADP ribosylation. It is mediated by poly-ADP ribose polymerase (PARP) family of proteins and reversed by the poly-ADP-ribose-glycohydrolase (PARG) family of proteins. With the knowledge of all these histones PTMs it has become increasingly apparent that the chromatin- associated factors harbors multiple histone binding domains. This further provides a varied and exciting probability for multivalent histone engagements which adds a layer of specificity to histone PTM recognition.

1.2 DNA methylation

DNA methylation is a common epigenetic mark which is conserved in many eukaryotes and prokaryotes. In the lower forms and plants, methylation can occur in the context of adenine and cytosine, however in eukaryotes it occurs primarily at the 5-position of the cytosine residues (5mC) in the context of CpG dinucleotides. The main function of DNA methylation in bacteria is to provide a protection mechanism, through which it

discriminates between the endogenous and foreign DNA introduction (Hemavathy and Nagaraja 1995). In mammals the DNA methylation pattern is established during embryonic development by de novo DNA methyltransferase (Dnmt), Dnmt3a and Dnmt3b (Okano, Bell et al. 1999; Li and Zhang 2014). Further, this methylation pattern is maintained through the Dnmt1-mediated copying mechanism during the cell division. The regions of the genome with higher number of methylated cytosine usually lead to the reduction of gene expression, so called gene silencing. Methylation serves primarily as a platform for the recruitment of many enzymes. Indeed, the change is recognized by MBP (methyl binding protein), such as MeCP2, MBP2 MBP3 or which will bind at the methylated CpG, and thus cause the recruitment of histone modifying enzymes, such as histone deacetylases (HDAC) (Nan, Ng et al. 1998) or methyltransferases (Fuks, Hurd et al. 2003). These enzymes can alter the structure of chromatin (heterochromatin) to consolidate the repression of transcription.

The absence of DNA methylation is prerequisite for the active transcription. CpG dinucleotides are distributed unevenly throughout the genome. In general, CpGs are under-represented in mammals, likely due to the mutagenic properties of 5mC. However, some regions of genome have high density of CpG dinucleotides which are referred as CpG islands and these islands are DNA methylation free. Certain transcription factors such as Cfp1 and Kdm2b which contains CXXC domain binds specifically to the unmethylated CpGs within these islands (Deaton and Bird 2011) providing an understanding about the contribution of unmethylated DNA in transcriptional regulation. In fact, 60–70% of annotated gene promoters are associated with a CpG island, including most of the housekeeping genes, developmental genes as well as tissue-specific genes.

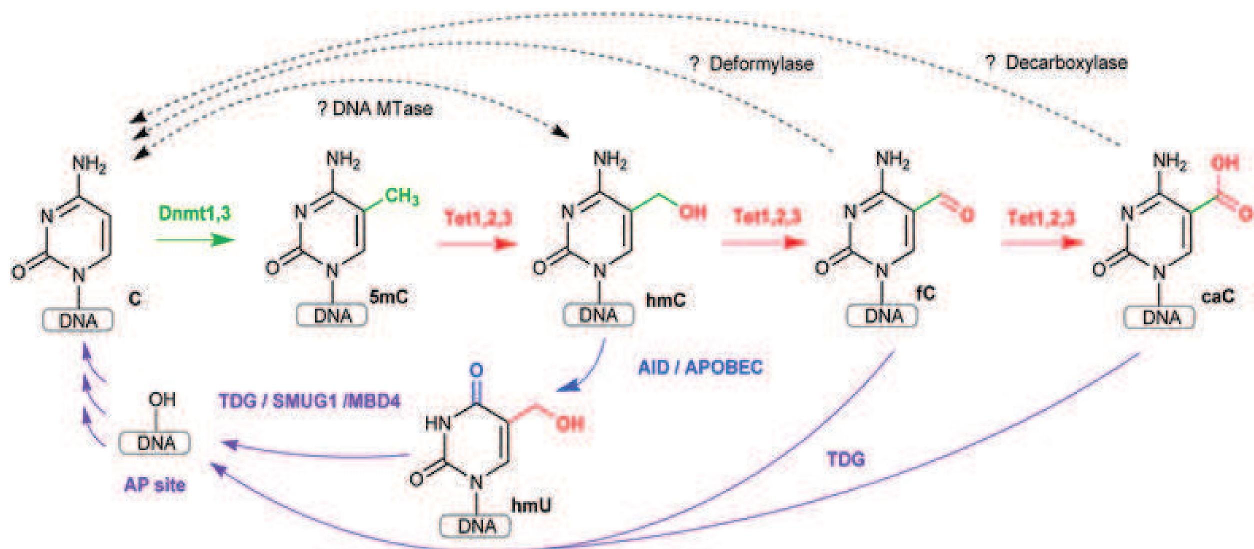


Figure 9. DNA demethylation Pathway.

Cytosine (C) is converted to 5-methylcytosine (5mC) by action of endogenous DNA methyltransferases (green pathway). Several mechanisms for DNA demethylation have been proposed. Horizontal arrows represent oxidation-based pathways performed by Tet proteins: methyl group of 5mC is consecutively oxidized to hydroxymethyl, formyl and carboxyl groups forming 5-hydroxymethylcytosine (hmC), 5-formylcytosine (fC) and 5-carboxycytosine (caC), respectively. Bent plain arrows show deamination-based pathways where hmC is deaminated to 5-hydroxymethyluracil (hmU) in the presence of AID/APOBEC family deaminases, and direct base excision repair (BER) pathways involving TDG, MBD4 and SMUG1 glycosylases, which all lead to transient formation of apyrimidinic (AP) sites in DNA. Dashed arrows denote the newly discovered hydroxymethylation and dehydroxymethylation reactions performed by cytosine-5 methyltransferases in-vitro and putative enzymes (deformylase and decarboxylase) which could directly remove the formyl and carboxyl groups from fC and caC, respectively. Adapted from (Kriukiene, Liutkeviciute et al. 2012)

Although this epigenetic mark can be transmitted to several generations, but they are not permanent. Indeed, studies have shown that methylation pattern can be altered throughout the life in response to the environmental changes or due to pathological processes such as oncogenic transformation or cellular ageing. DNA methylation marks can be removed or reprogrammed during early development by two ways. It can be either removed passively as during cell division by inhibition of the maintenance methyltransferase, Dnmt1 or by an active demethylation mechanism which involves removal of the methyl-group from 5mC by a family of DNA hydroxylases called Tet proteins (ten eleven translocation), that exist in three isoforms: TET 1, 2 and 3. Passive methylation is well understood and accepted while the molecular mechanism behind active DNA demethylation is beginning to be understood and there appears multiple modes through which it can occur. The mechanisms by which active DNA demethylation occurs can be broadly classified under two categories i.e. oxidative and repair based DNA demethylation [Figure9, (Kriukiene, Liutkeviciute et al. 2012)]. The oxidative mechanism includes the enzymatic oxidation 5mC to the 5-hydroxymethyl-2'-deoxycytidine (5hmC), which can be further oxidized to 5-formycytocise (5fC) or 5-carboxy-cytosine (5caC) (Tahiliani, Koh et al. 2009; Ito, Shen et al. 2011; Maiti and Drohat 2011). The transition from 5fC and 5caC to cytosine is done by thymidine DNA glycosylase, TDG. In particular TET1 and TET3 exhibit a CXXC domain, a binding domain CpG, which they utilize for the conversion of 5mC to the alternative form of cytosine. The other mechanisms which involve DNA repair has been proposed as an alternative to the oxidative demethylation. These includes base excision repair (BER) through direct excision of 5mC, deamination of 5mC to T followed by BER and through nucleotide excision repair (NER). The deamination based pathway, in particular uses AID cytosine deaminases (activation induced cytidine deaminase)/ APOBEC which deaminate hmC to 5-hydroxymethyluracil (hmU). This step is followed by the TDG,

MBD4 and SMUG1 glycosylases, which all lead to transient formation of apyrimidinic (AP) sites in DNA and direct base excision repair (BER) pathways (Figure 9).

The DNA damage–associated protein Gadd45a was also described to participate in selective demethylation at the promoters of target genes through recruitment of the DNA repair machinery (Barreto, Schafer et al. 2007; Schmitz, Schmitt et al. 2009), but how it was targeted to specific sites remained unknown. Recently, Schäfer et al. showed the implication of the tumor suppressor Ing1—which contains a PHD finger domain that specifically binds H3K4me3—in DNA demethylation by recruiting Gadd45a to H3K4me3 sites (Schäfer, Karaulanov et al. 2013). In addition NER factors; XPG and XPF have also been shown to be involved in DNA demethylation (Le May, Fradin et al. 2012).

D.2 Chromatin rearrangement

Regulation of gene expression can occur locally or over a large genomic distance via various regulatory elements located far upstream or downstream of the gene and can also control expression of the genes which are present on other chromosomes. Such long-range control of gene expression occurs through a process referred as chromatin looping. The first example of the chromatin loops was shown by the interaction of locus control region (LCR) of the β-globin gene with the embryonic and the adult type β-globin promoters at the appropriate time of development (Carter, Chakalova et al. 2002). Chromatin loops have been also described in numerous genes, as a general organization of the chromatin fiber, associated with the activation of transcription (Spilianakis and Flavell 2004) (Vernimmen, De Gobbi et al. 2007). Different regions of the genes have been shown to be involved in forming chromatin loops. It could be between enhancers and the promoter (Krivega and Dean 2012) or between the promoter

and terminator genes (Laine, Singh et al. 2009). Dynamic promoter–terminator loops have also been described for the breast cancer *BRCA1* gene (Tan-Wong, French et al. 2008) and at the gene encoding the immune-histological marker CD68 in mammalian cells (O'Reilly and Greaves 2007). The above two studies, by Liane and Ta-Wang shows that the promoter and terminator interaction persist following a cycle of transcriptional activation and repression and this phenomenon is referred to as transcription memory. Hence, higher order chromatin organization can contribute to physiological roles such transcription memory facilitating the rapid re-expression of the gene.

Chromatin loops are stabilized by number of proteins such as tissue specific transcription factors EKLF, GAT-1 and Ldb1 at the murine β-globin locus. One of the most studied proteins in chromatin loop formation is CCCTC- binding factor, CTCF. It is ubiquitous and has very diverse function, including enhancer-blocking, X-chromosome inactivation, genome imprinting, gene activation or repression. Genome wide studies shows that CTCF binds to tens of thousands of genomic sites and almost 14000 genomic regions are flanked by CTCF on both sides referred as CTCF-pair-defined domains (CPD) of average size 210 kb. The CTCF has been found to be linked to the methylated DNA sequence (Wang, Maurano et al. 2012). It can bind to methylated DNA sequences *in-vitro* but preferentially binds to the unmethylated sequences as observed in the H19–Igf2 locus. CTCF sites are associated with frequently flanking the lamina-associated domains (LADs) and found enriched between the active chromatin (high in H2K5Ac) and inactive chromatin domains (high in H3K27me3). Furthermore, CTCF can form complex with PARP1 and DNMT1 and thus influence the DNA methylation. CTCF activates PARP1, which then can add ADP–ribose groups to DNMT1 to inactivate this enzyme, with maintenance of methyl-free CpGs as the result (Yu, Ginjala et al. 2004). Another well studied model for gene expression which is mediated by chromatin looping involves the Polycomb groups (PcG) and the trithorax

group (trxG) regulatory system. They were initially discovered in *Drosophila* and are known to maintain the repressed state of the targeted gene. Polycomb repressive elements (PREs) containing Fab7 elements controls the bithorax gene cluster in *Drosophila* by contacting its promoter.

The most widely used method to determine the genomic loops is chromosome conformation capture (3C), originally developed by Job Dekker in 2002 (Dekker, Rippe et al. 2002). Several 3C based approaches which include 4C, 5C, Hi-C and ChIA-PET allows genome-wide identification and analysis of sequences involved in the three dimensional organization of the full genome (Montavon and Duboule 2012).

II. The TFIIH and NER factors in transcription

A. TFIIH: A multifunctional Complex

TFIIH was first characterized as a general transcription factor of RNA pol II in 1989. It was initially purified from the rat liver (Conaway and Conaway 1989), from HeLa cells as basic transcription factor 2 (Gerard, Fischer et al. 1991). From yeast, it was isolated as RNA pol II transcription factor b (Feaver, Gileadi et al. 1991) by conventional column chromatography and later purified by a single step immunoaffinity purification method (Kershnar, Wu et al. 1998). After the homology of all these factors was identified, the universal nomenclature for transcription factors was proposed and it was designated as TFIIH. Originally TFIIH was thought to be exclusively a basal transcriptional factor but later it was found to be involved in DNA repair and possibly in cell cycle regulation.

A.1 Composition of TFIIH

Transcription factor II H is a multifunctional complex of proteins which consists of two sub-complexes: a core complex and a Cdk activating kinase (CAK) (See Figures 8, Table 4). The core complex consists of six subunits: *Xeroderma pigmentosum* B (XPB; p89), p62, p52, p44, p34 and trichothiodystrophy A (TTD-A; p8) (Serizawa, Conaway et al. 1993). CAK is composed of the three subunits: CDK7 (p40), cyclin (p34) and “ménage a trios” (MAT1; p32). The core and the CAK are held together by the X*eroderma pigmentosum* D (XPD; p80), subunit. Electron microscopy shows that TFIIH is organized into a ring like core structure from which globular CAK sub-complex protrudes out (Schultz, Fribourg et al. 2000) (Figure 10, Table 5). TFIIH has importantly three enzymatic activities: DNA-dependent ATPase (XPB and XPD), ATP dependent helicase (XPD and XPB), and CTD kinase (Cdk7). In addition, it also contains Ubiquitin ligase activity displayed by p44 and MAT1 (Compe and Egly 2012). All the components of TFIIH are not found in complex form. CAK can be found independently (Rossignol, Kolb-Cheynel et al. 1997).

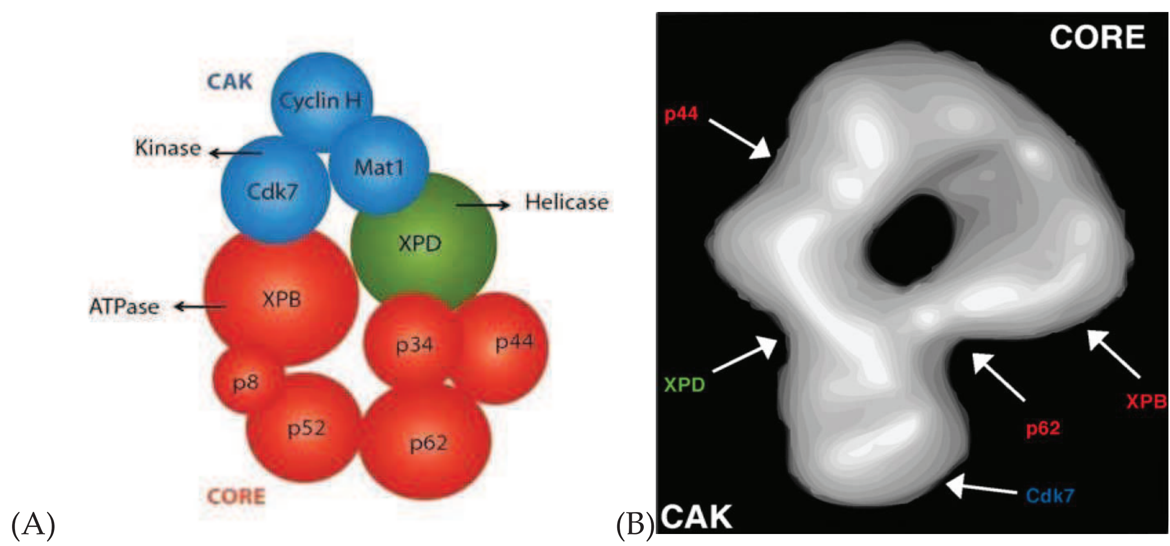


Figure 10. Structure of TFIIH

(A) Schematic representation with core in red and CAK in blow, XPD in green. (B) Electron microscopy view (Adapted from (Compe and Egly 2012)

Sub-complex TFIIH	Human	Yeast	Function	Associated genetic diseases
Core	XPB	Ssl2	3'-5' ATPase	TTD and XPC/CS
	p62	Tfb1	Structural function, interacts with transcription and NER factors	
	p52	Tfb2	Regulates the ATPase activity of XPB	
	p44	Ssl2	E3 ubiquitin ligase (in yeast)	
	p34	Tfb4	Structural function and interacts with p44	
	p8 /TTD-A	Tfb5	Regulates the ATPase activity of XPB	TTD
XPD	XPD	Rad3	ATP dependent 5'-3' helicase activity and links the core to CAK	TTD, XP and XP/CS
CAK	CDK7	Kin28	Kinase	
	Cyclin H	Ccl1	Modulates the CDK7 kinase activity	
	MAT1		Stabilize the CAK	

Table 5. TFIIH components and their functions

Adapted from (Compe and Egly 2012)

Remarkably, the free CAK and the CAK interacting with the core TFIIH have significantly different substrate specificity. Free CAK acts as Cdk-activating kinase and can phosphorylates number of substrates which include Cdk1, Cdk2, Cdk4, and Cdk6 involved in the cell cycle progression. When it is in complex with TFIIH it preferentially phosphorylates the CTD of RNA pol II. In addition to that, TFIIH also phosphorylates the NRs during transcription. The CAK is released by the core during the NER and reappears after DNA repair for the resumption of the transcription (Coin, Oksenych et al. 2008). p8/TTD-A subunit of the core also exists in the cell in dimeric form separately from TFIIH.

1.1 XPB

XPB (ERCC3; Ssl2; Rad25) is the largest member of TFIIH. The ERCC3 gene (excision repair cross-complementing group 3) was initially characterized as a gene correcting a DNA repair deficiency in XPB cells from patient suffering from a severe genetic syndrome (*Xeroderma pigmentosum*) (Weeda, van Ham et al. 1990) (Weeda, van Ham et al. 1990). The gene was cloned after transfection of human chromosomal DNA to the UV-sensitive, incision-defective Chinese hamster ovary (CHO) mutant 27-1. It was observed, that ERCC3 specifically corrects the defects in an early step of DNA excision repair pathway of UV-sensitive rodent mutants of complementation group 3. Structurally, ERCC3 is 782 amino acids protein possessing putative nucleotide binding domain, chromatin binding domain, helix-turn-helix DNA binding domain and seven consecutive motifs. XPB belongs to the SF2 superfamily of ATP-dependent DNA or RNA helicases (Weeda, van Ham et al. 1990). It is essential for both initiation of transcription by RNA polymerase II and nucleotide excision repair. Its function is evolutionarily conserved among eukaryotes.

Crystal structure of XPB homolog from *Archaeoglobus fulgidus* (AfXPB) revealed that XPB consists of two Rec-A like helicase domains (HD1and HD2) connected by a flexible hinge, a DNA damage recognition domain (DRD), a unique RED motif and a flexible thumb motif (ThM), (Fan, Arvai et al. 2006). The positively charged Thumb-like domain and an RED motif have been shown to be implicated in XPB specific activities (Oksenych, Bernardes de Jesus et al. 2009). HD1 and HD2 are packaged opposite to each other and form a cleft that brings I, II, V and VI (Figure 9a) to form a composite ATP-binding site [Figure 11, (Fuss and Tainer 2011)]. Homologue of human XPB can be found in a number of bacteria i.e., *Mycobacterium tuberculosis* and *Kineococcus radiotolerans* (Biswas, Pero et al. 2009). *Sulfolobus solfataricus* and related archaea have two homologues of the XPB protein, XPB1 (Sso0959) and XPB2 (Sso0473) which binds to single stranded DNA and have DNA-dependent ATPase activity, but neither of them work as helicase against a range of damaged and undamaged DNA substrates *in-vitro*. The enzymatic activity of XPB activity is highly regulated by other TFIIH subunits. p52 interacts with XPB directly and stimulates its ATPase activity (Merino, Madden et al. 1993; Jawhari, Laine et al. 2002) while, p8/TTD-A stimulates indirectly by interacting with p52. In addition, the NER damage recognition factor XPC-HR23B also stimulates the ATPase activity of XPB and facilitates DNA opening around the lesion (Ge and Roeder 1994).

Mutations in XPB gene lead to the rare human genetic disorders; *Xeroderma pigmentosum* (XP), cokayne syndrome (CS) and trichothiodystrohy (TTD) (Phenotypic characters are discussed in Chapter III). Up to now, only six mutations in six families are reported (Shykind, Kim et al. 1995). The two mutations in N-terminal domain, F99S and T119P are the most studied (See Result section, Manuscript Figure 1). These two mutations have been shown to show less RNA synthesis in *in-vitro* transcription assay

A) XPB

B) XPD

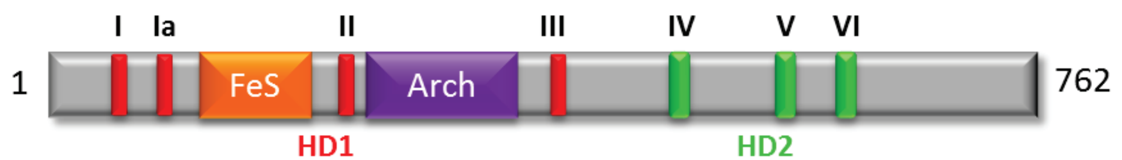


Figure 11. Linear schematics of XPB and XPD helicases.

The seven helicase motifs HD1 and HD2 are common to both XPB and XPD. The for helicase motifs in HD1 (I, Ia, II, III) are indicated in red and three (IV, V, VI) in HD2 are indicated in green, A) Damage recognition domain, DRD domain N-terminal close to HD1 is colored pink, dark red and blue represents the R-E-D and thumb domain respectively, B) FeS domain in XPD in orange and Arch is in purple. Adapted from (Fuss and Tainer 2011).

while two other nonsense mutations reported by Oh. SK et.al lead to premature termination of the protein (Kang, Auble et al. 1995).

1.2 XPD

XPD/Rad3/ERCC2 is a member of SF2 helicase family and the second largest member of TFIIH. It is ATP dependent 5′-3′ helicase of TFIIH. XPD is highly conserved in archea and eukaryotes. The closest homologue of XPD is the bacterial damage-inducible G (Din G). In addition it has several paralogues such as regular telomere length (RTEl1), Fanconi's anemia complementation group (FancJ) and Chl1. All of these proteins shares conserved Fe-S cluster binding domain and functions as helicases in different recombination and repair processes (Liu, Rudolf et al. 2008). Crystal structure of archael XPD has revealed that apart from two canonical helicase domains; HD1 and HD2, XPD is composed of 4 Fe-S cluster domains involved in DNA damage recognition and an Arch domain (Figure 11). Together, the Arch and iron-sulfur domains form a channel through which single-strand DNA is dragged by the action of the motor domains in a cyclical ATP-dependent reaction. This Arch domain has been recently shown to act as a platform to recruit the CAK, both in transcription and DNA repair (Abdulrahman, Iltis et al. 2013). Furthermore Qi et.al, has shown that monomeric XPD unwinds the duplex DNA in 1-bp steps, yet it exhibits frequent backsteps and undergoes conformational transitions manifested in 5-bp backward and forwards steps. These all mechanisms depends on the availability of ATP and finally the single-base pair stepping in forward direction give rise to the unwinding of the duplex DNA (Qi 2013).

XPD has very important role in the structural organization of TFIIH, acting as a bridge between the core and the CAK. It interacts with Mat1 of the CAK sub-complex (Lagrange, Kapanidis et al. 1998). The helicase activity of XPD which is dispensable for transcription is absolutely required for DNA unwinding during the NER. XPD helicase activity is further stimulated by p44 through its association with carboxyl-terminal

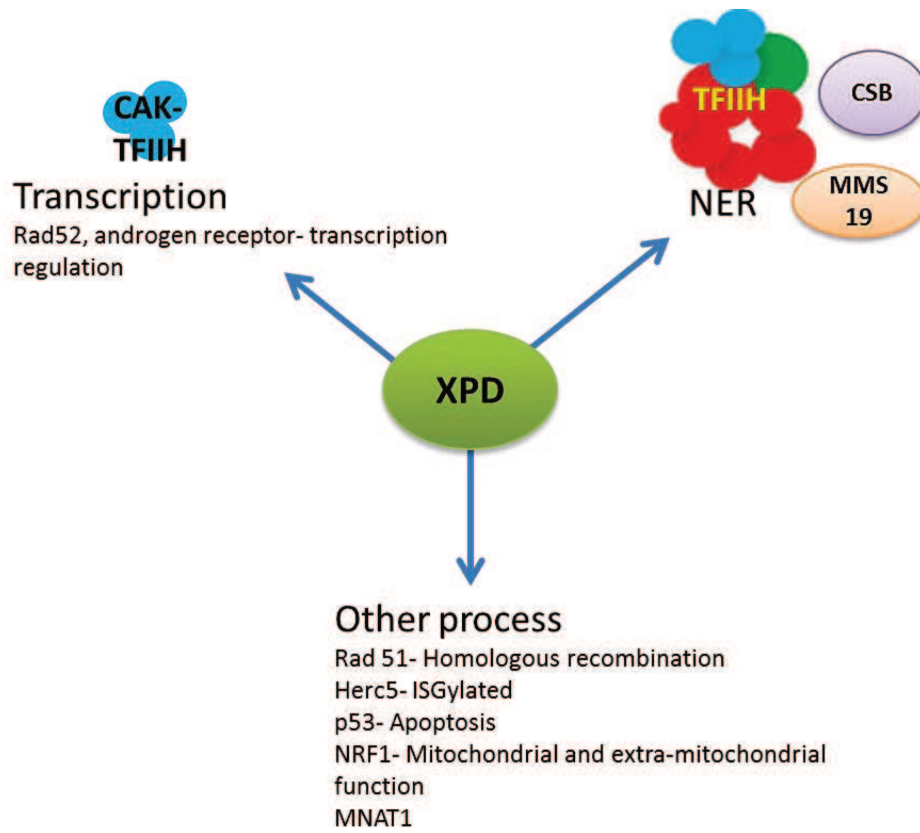


Figure 12. XPD and its interactions

Adapted from PubMed Gene ID: 2068

domain of XPD (Coin, Oksenych et al. 2007). XPD also interact with several other proteins and can be found in complex with other protein (Figure 12). Several mutations in the XPD gene are reported till date which gives rise to XP, TTD, CS and XP/CS phenotype (Mutation studied in present study are discussed in detail in Chapter III)

1.3 p62

p62/Tfb1 is the structural component of TFIIH core, highly conserved during evolution and first cloned in 1992 by Fisher et.al (Laurent Fischer 1992). p62/Tfb1 facilitates the recruitment of TFIIH to the transcription preinitiation complex through an interaction between its amino-terminal PH domain and the acidic carboxyl terminus of TFIIEα (TFIIEα336–439) (Xiao, Pearson et al. 1994). This interaction markedly stimulates TFIIH-dependent phosphorylation of the C-terminal domain of RNA pol II. In addition, the p62/Tfb1 subunit has been shown to interact with p53 through its acidic transactivation domains 2 (TAD2). The interaction between p53 TAD2 and p62/Tfb1 is enhanced by phosphorylation of Ser-46 and Thr-55 of p53. This interaction is thought to be important for regulation of selected p53-regulated genes that require p53 TAD2, such as MDM2, PUMA, WAF1, and BAX1 (Zurita and Merino 2003). NMR studies have shown that the acidic TAD of EBNA2 (Epstein–Barr virus nuclear antigen 2) (residues 431-487) interact with the Tfb1/p62 subunit of TFIIH and is required for the transactivation (Chabot, Raiola et al. 2014). Point mutations in the acidic TAD of VP16 that reduce its transactivation activity also weaken its binding to TFIIH while mutation in TAD of E2F-1 leads to the loss of TFIIH binding via p62 and hence results in 60-65% reduction in transactivation. In addition, p62/Tfb1 directly interacts with thyroid hormone receptor, enhancing T3-mediated transcription (Burley and Roeder 1996). In DNA repair, p62/Tfb1 interacts with damage recognition complex XPC/Rad4, as well as with structure-specific endonuclease XPG/Rad2 that cleaves DNA at the 3′ end of the lesion, (Iyer, Reagan et al. 1996; Yokoi, Masutani et al. 2000). In particular, Rad4 binds with the

PH domain of p62/Tfb1 to form Rad4-Tfb1PH complex which has interface similar as to one used by Rad2 to bind Tfb1PH. So, Rad4 competes with Rad2 for binding to the Tfb1 subunit of TFIIH in NER (Lafrance-Vanasse, Arseneault et al. 2013). There are no described mutations in human p62 that leads to any DNA repair disorders, however the correlation between p62 single nucleotide polymorphisms (SNP) and the lung cancer risk have been reported (Wu, Liu et al. 2009).

1.4 p52

p52/Tfb2 is a non-enzymatic component of TFIIH and first co-purified and co-precipitated as fifth subunit with eight other TFIIH subunits (Xie, Kokubo et al. 1996). p52/Tfb2 protein physically interacts with XPB and stimulates its ATPase activity during transcription and DNA repair. The complex p34/p44/p62 (Tfb4/Ssl1/Tfb1) does not interact with XPB in the absence of p52 (Tfb2) and the depletion of p52 leads to defect in DNA opening during transcription initiation (Orphanides, Lagrange et al. 1996; Feaver, Henry et al. 1997; Jawhari, Laine et al. 2002). Till date, no human genetic disorders are known due to the mutation in p52 gene. However, mutation in p52/*Dmp52* in *Drosophila* is reported which is known to cause neurological defects, UV-sensitivity and features similar to TTD patients (Fregoso, Laine et al. 2007). These mutation more likely impaired p52-p8/TTD-A interaction, as overexpression of p8/TTD-A enhances UV-irradiation resistance and suppresses TFIIH mutations in a *Drosophila* TTD model (Aguilar-Fuentes, Fregoso et al. 2008). Furthermore, p52 has been shown to be a transient component of the catalytic site of human mitochondrial RNA polymerase (mitoRNAP) and also implicated in promoter melting (Sologub, Litonin et al. 2009).

1.5 p44

The gene encoding p44 is a human counterpart of Ssl1 in yeast and has a RING finger domain which is highly conserved and interacts with XPD upon DNA binding. This binding stimulates the helicase activity of XPD *in-vitro* (Flores, Ha et al. 1990; Tanaka,

Watanabe et al. 2009). The NMR structure of p44 RING finger domain (residues 321 to 395) shows two zinc binding sites coordinated by eight cysteine residues (Kellenberger, Dominguez et al. 2005). Interestingly these conserved cysteine residues are also found in the RING finger domain of Ssl1 and are critical for the E3 ligase activity in yeast. Mutation in the first two cysteine residues in yeast (C403A and C406A) abolishes Ssl1 enzymatic activity in an *in-vitro* polyubiquitination assay performed in the presence of E1 and E2 (Ubc4) enzymes and further reduced the yeast survival rate following UV irradiation or methyl methanesulfonate (MMS) treatment. The E3 ubiquitin ligase activity is further enhanced by another subunit p34 of TFIIH. Amino acid residues from 66 to 200 are involved in the interaction with XPD, since mutations between this region have been shown to decrease the helicase activity and further leads to the defect in the first step of the transcription reaction, i.e. the first phosphodiester bond formation and promoter clearance (Seroz, Perez et al. 2000). The p44 mRNA has been shown to be regulated by microRNA, miR-27a. Moreover, miR-27a destabilizes the p44 subunit of the TFIIH complex during the G2-M phase, thus modulating the transcriptional shutdown observed during this transition (Portal 2011). p44 gene has been described in context to the spinal cord muscular atrophies. p44 gene is part of the 500 kb inverted duplication on chromosome 5q13. One copy of p44 gene is within the telomeric region while the other resides in the centromeric region. Deletion of this gene in telomeric region sometimes accompanies deletion of the neighboring SMN1 gene resulting in the spinal muscular atrophy (SMA) phenotype but it is unclear if deletion of p44 gene contributes to the SMA phenotype (Archambault, Pan et al. 1998).

1.6 p34

p34 of human TFIIH is homologous to Tfb4 of yeast and contains a zinc finger domain similar to p44 (Yonaha, Tsuchiya et al. 1997). It has been shown to be important for both transcription and DNA repair, though it is not known for any enzymatic activity

(Feaver, Henry et al. 1997). Recently, the crystal structure of the p34 subunit of the TFIIH complex from the eukaryotic thermophilic fungus *Chaetomium thermophilum* has been revealed (Schmitt, Kuper et al. 2014). It shows that p34 contains a *von Willebrand Factor A* (vWA) like fold which is generally known to be involved in protein-protein interactions. Within TFIIH p34 strongly interacts with p44, a positive regulator of the helicase XPD.

1.7 Cdk7

Cdk/7MO15/Kin 28 is a catalytic subunit of CAK and is a member of Serine/threonine protein kinase family. CDK 7 can execute dual function by recognizing different classes of substrate. In metazoans, CDK7 was initially identified as the CDK-activating kinase CAK, which phosphorylates cell-cycle CDKs within the activation segment (T-loop). It forms a heterotrimeric complex with the ring finger protein Mat1 within the multi-complex TFIIH and also phosphorylates the CTD of the largest subunit of RNA pol II. The difference between CTD-like and T-loop substrates is also seen in the ability of Cdk7 to phosphorylate peptides in solution. CTD peptides are effective substrates for Cdk7, whereas T-loop substrates are not. Human Cdk7 can phosphorylates the Ser 5 and it was first shown by Roy et al. that mammalian Cdk7 can phosphorylates a peptide with Alanine2 substitution, while it cannot phosphorylates a similar peptide with Alanine (Roy, Adamczewski et al. 1994). Cdk7 kinase activity is required for the transcription initiation (Akoulitchev, Makela et al. 1995) and this kinase activity is stimulated by the mediator complex that binds to the unphosphorylated form of the CTD and helps to recruit RNA pol II to the preinitiation complex (PIC). Phosphorylation of Ser5 and Ser7 residues in the CTD disrupts this interaction permitting promoter escape and entry into the elongation phase of the transcription cycle (Kim, Suh et al. 2009). However, recently it has been shown in yeast that the Kin28 (CdK7 homologue in yeast), is important for promoter escape but not for elongation.

Phosphorylation of CTD causes dissociation of mediator and subsequently promoter escape, whereas its depletion dramatically increases mediator occupancy at the core promoter (Wong, Jin et al. 2014). Although Cdk7 is known to control the organismal viability, dissecting its physiological role in transcription and cell cycle control seems to be tough. For example, loss of Cdk7 effects both transcription and mitosis in *C. elegans* (Luse, Spangler et al. 2011), while adults *Drosophila* with homozygous mutation in Cdk7 are viable for over 40 days at the non-permissive temperature arguing against an essential role in transcription. It has been shown that the effect on RNA pol II transcription is miniscule even though mice homozygous for a Cdk7 deletion die early in embryogenesis (Ganuza 2012). Tissue specific knockouts of Cdk7 in adult mice cause the loss of proliferating cells but little change in non-proliferating cells. This result argues that TFIIH kinase is not required for transcription in post-mitotic cells. All together these studies indicate that Cdk7 is not essential for transcription in all the cells is observed. In support of this hypothesis we can assume that there might be other CDKs which can provide the missing CTD kinase activity.

1.8 Cylin H

Cyclins are the conserved proteins and abundant during the cell cycle. They function as the regulators of CDKs and functions during two processes, transcription and cell cycle. Cyclin H forms a complex with CDK7 kinase and ring finger protein MAT1. Cyclin-CDK association enables substrate phosphorylation. Furthermore, cyclin H has an essential function in promoting the self-renewal of the pluripotent stem cells of blastocyst stage embryos suggesting the role of cyclin H in maintaining ES cell identity and in early embryonic development (Patel and Simon 2010). Also the TFIIH kinase activity is affected by the interaction of the U1 snRNA and the cyclin H (Buratowski, Hahn et al. 1989). Cyclin H is also phosphorylated by the Cdk8 subunit of the mediator

complex (Akoulitchev, Chuikov et al. 2000). This phosphorylation represses both the ability of TFIIH to activate transcription and its CTD kinase activity.

1.9 Mat1

The MAT1/Tfb3 protein stands for "ménage à trois" is a third component of CAK. This is a RING finger containing protein and plays a structural role in stabilizing the complex. The core of the molecule consists of two repeats containing five helices each and forming the canonical cyclin folds similar to TFIIB. Out of these five two helices define specificity of the cyclin H molecule. These two long helices extends the cyclin fold at its N- and C-termini and pack together against the first repeat on the side opposite to the kinase. Deletion mutants show that these terminal helices are required for a functionally active cyclin H (Kraemer, Ranallo et al. 2001). The coiled-coil and the hydrophobic domains of MAT1 interact with the N-terminal domain of XPD and prevent its helicase activity, although this inhibition in the helicase activity is overcome by XPD when p44 binds to it (Sandrock and Egly 2001).

1.10 p8/TTD-A

p8/TTD-A/ Tfb5 is a 71 amino acids small, αβ protein built around an antiparallel β-sheet that forms a homodimer with an extended interface (Roeder 1996). It presents unusual stretches of conserved and hydrophobic residues and is almost conserved among all eukaryotic genomes from human to yeast. Initially, it was shown to be a *bona fide* component of RNA pol II preinitiation complex and a component of TFIIH which is required for efficient transcription both *in vitro* and *in vivo* (Ranish, Hahn et al. 2004). Later it was shown to be a DNA repair specific factor (Chao, Gadbois et al. 1996). Further, it was shown to be a NER-dedicated subunit of TFIIH by Giglia-Mari et al. in fibroblast cells (Giglia-Mari, Miquel et al. 2006). It stimulates the ATPase activity of XPB together with the p52 and NER factor XPC-HR23B. It interacts with the hydrophobic surface of p52 and seemingly protects p52 from solvents. Structural studies on yeast

p8/TTD-A and p52 have shown that the C-terminal of p52 anchors p8/TTD-A and forms a heterodimer (Coin, Bergmann et al. 1999; Kainov, Cura et al. 2010). The p8/TTDA–p52 interaction was further confirmed in-vivo (Nonnekens, Cabantous et al. 2013). This interaction subsequently allows the NER factors to localize to the damage (Theil, Nonnekens et al. 2011).

Mutation in p8/TTD-A subunit gives rise to rare form of TTD phenotype. The first report of a TTD patients due to mutation in TTD was made long back in 1993 and it was thought to be TTD phenotype due to a new excision-repair complementation group (Stefanini, Vermeulen et al. 1993). After several years it was shown that the protein causing the rare form of TTD is indeed p8/TTD-A (Shilatifard, Conaway et al. 2003).

B. NER Factors

The NER factors are the group of proteins which are known to be involved not only in NER pathways but also in several DNA repair pathways. Mutation in 13 genes out of 30, which carry out the nucleotide excision repair process, causes human genetic disorders like *Xeroderma pigmentosum*, Cockayne syndrome and trichothiodystrophy. These proteins include the xeroderma pigmentosum complementation group proteins; the XP proteins (XPA – XPG), XP protein variant XPV, p8/TTD-A, CSA and CSB. Based on their functional aspects in the NER process, XP proteins implicated in NER pathways can be broadly divided into three groups. XPA, XPC, XPE, CSA and CSB are the initial proteins that are required for sensing DNA damage and initiating the repair process. XPB and XPD are responsible for the opening of the DNA strand surrounding DNA lesions during NER. XPG and XPF are the endonucleases that perform the dual incisions to release the damaged strand and allow resynthesis using the nondamaged strand as a template (their role in NER pathways is described in detail in next section). The sequential assembly of these proteins on the damaged DNA was very well studied by Riedl et al (Riedl 2003). Their roles in NER pathway have been very well accepted. However, these proteins are not only required in DNA repair, they have indeed many overlapping roles in DNA metabolic processes, cell cycle regulation and transcription (Table 6). Their existence in transcription was put forward precisely in 2009 (Barreto, Schafer et al. 2007; Schmitz, Schmitt et al. 2009; Le May, Mota-Fernandes et al. 2010) and their advent in transcription was shown to be a paradigm shift in the understanding of the human genetic disorders like XP, CS and TTD. These studies have clearly shown that the NER factors are recruited sequentially on the promoter of active gene. XPC seems to be the first NER factor to bind to active promoter similar to its recruitment in NER and then allows the recruitment of other NER factors such as CSB, RPA, XPA, XPG and XPF during transcription. During NER, XPC function as the damage sensor

and allows TFIIH to come at the damage site once it recognizes the lesion, followed by other NER factors which complete the DNA repair process. During transcription, XPG and XPF are shown to be required for the DNA break formation at the active promoter, further allowing chromatin remodeling to occur and ensure accurate transcription. Other NER proteins such as CSB and CSA are also known to function as chromatin remodelers (Fousteri, Vermeulen et al. 2006). In particular CSB regulates the transcriptional program after UV irradiation (Proietti-De-Santis, Drane et al. 2006). However, there is a need to be explored more in order to fully understand the role of all the NER factors during transcription. Table 6, represents the interacting partners of all the NER factors in many processes other than transcription and DNA repair.

Protein	Functions	Interacting partners	Process of Involvement
XPA	Damage sensor	ATR	Nuclear import of XPA in response to UV
		RPA, XPF, TFIIH	NER
		XAB1	Modulate the nuclear import of XPA
		XAB2	Scaffold protein in transcription and cell cycle
		DDB2	DDB-mediated NER
XPB	3'-5' ATP dependent helicase	PP2A-related kinase	Phosphorylates XPB
		P53	Reduce the rate of apoptosis
		SUG1	Modulates the degradation of RAD4
		TFIIE	Transcription
		XPC, XPG p62, p52, p8	Transcription and NER
XPC	Damage sensor	ATM	Cell cycle control and remove lesions
		Centrin2	Damage recognition
		SUMO, XPB, p62, XPE, XPG	NER
		TDG	BER
		S5a-26 proteasome complex	Degradation by the 26S proteasome post UV
XPD	5'- 3'ATP dependent helicase	hMMS19, p44, MAT1	NER
		p53	Reduce the rate of apoptosis
		MMXD	Chromosome segregation
XPE	Damage sensor	c-ABL tyrosine kinase	Degradation by the 26S proteasome post UV
		Cul4-COP9	Transcriptional repression
		XPC	NER
		STAGA , CBP/p300	Chromatin remodeling for DNA repair
		E2F1	Cell cycle arrest
XPF	Endo-nuclease	αSPIIε ,FANC-A,Msh2 RAD51/52	Mismatch repair
		RPA	Inter-cross linked repair
		XPA/ERCC4	NER
		TRF	Forms a ternary complex during NER
		TFIID	Regulates telomere integrity
			Transcription (a)
XPG	Endo-nuclease	Nth1	BER
		PCNA	DNA replication
		RNA pol II, RPA/XPA XPB, p62	Transcription, NER
XPV	Y-family DNA polymerase	PCNA, Rad18, Rev1	Allow trans-lesion synthesis past the damage
		Rad51	Recombination and double strand break repair
CSA	Ubiquitin ligase	XAB2	TCR and transcription
		RNA pol II, CSB	TCR
		pP44	Transcription by RNA pol I and II
		DDB1, Cul4A, Roc1,COP9	Ubiquitin ligase complex component, responds to UV
		Topoisomerase I	Assist in dsDNA unwinding by cutting one strand and

			later reannealing
CSB	DNA-dependent ATPase of SWI/SNF family	CSA XPB and XPD p44 and p62 RNA pol II RNA pol I, III OGG1 PARP-1 p53 XPG XPA	TCR Transcription by RNA pol I, II and NER, helicase activity Transcription by RNA pol I, II and NER Transcription of mRNA Transcription of rRNA In BER Maintain heterochromatin and repair oxidative damage Cell cycle control, G1-arrest, NER and apoptosis NER, BER NER
p8/TTD-A	Implicated in NER	XPB, p52	NER

Table 6. NER factors and their interaction partners [Adapted from book chapters by Shell M et al. 2008 and Sun G et al. 2009, (Hirose and Manley 2000)]

C. TFIIH: Bridging transcription and repair

TFIIH has been first identified and characterized as an essential component of RNA pol II mediated transcription machinery, (Conaway and Conaway 1989; Gerard, Fischer et al. 1991). Later in 2002, it was shown to be involved in the transcription of rRNA by RNA pol I (Hoogstraten, Nigg et al. 2002) and subsequently in the transcription of noncoding RNAs by RNA pol III (Wrange, Okret et al. 1984). Nonetheless, most of the studies on the role of TFIIH in transcription were focused on RNA pol II system. TFIIH is the last factor recruited in the sequential assembly of PIC at a core promoter. Once the PIC is established, the ATPase activity of XPB is necessary to open the DNA around the transcription start site. In this process, XPB probably walks on the DNA helix which generates a supertwist, downstream of a fixed RNA pol II-promoter complex promoting the melting of DNA. The activity of XPB is regulated by p52 in transcription and NER but by p8/TTD-A only in NER. XPB ATPase activity is also regulated by other

transcription factors. During the transcription of MYC gene, XPB ATPase activity is stimulated by FBP (FUSE-binding protein) and inhibited by FIR (FBP interacting repressor). Open complex formation subsequently initiates the transcription which is dependent on the phosphorylation of CTD on Ser5 by Cdk7. The CTD Ser 5 phosphorylation is also modulated by the MAT1 and Cyclin H subunits within TFIIH. Whereas, Cyclin H is phosphorylated by Cdk8 kinase of the mediator which negatively regulate the TFIIH by inhibiting the Cdk7 activity (Akoulitchev, Chuikov et al. 2000). Depending on its phosphorylation state, CTD attracts mRNA and histone modifying enzymes. At the completion of the transcription cycle Cdk7 also phosphorylates the Ser7 of the RNA pol II CTD (Described in detail in section I.A.2.1). Besides its role in the basal transcription, TFIIH can also controls the regulation of several transcription factors, such as p53, Epstein Barr nuclear antigen 2 (EBN2), hepatitis B virus (HBV) X Protein (HPX), Herpes simplex virion protein VP16 and nuclear receptors (NRs) (Mangelsdorf, Thummel et al. 1995). Phosphorylation of NRs by Cdk7 of TFIIH can occur within their A/B domain. This phosphorylation can occur either without the ligand as in RARα1, RARϒ, androgen receptors (AR), the peroxisome proliferator receptors-α, PPARϒ1, PPARϒ2 and thyroid hormone receptor-α1 or in response to hormone like oestrogen receptor-α and thyroid hormone receptor-β.

In addition to its role in transcription, TFIIH participates in the Nucleotide excision repair (NER) process. On the basis of the recognition procedure, NER pathway can be subdivided into two subpathways: the global genome repair (GGR) and the transcription coupled repair (TCR). In GGR subpathway, the damage is sensed by XPC-RAD23 complex and centrin 2 (CETN2). Upon correct binding of the XPC to the damaged DNA, the C-terminal domain of XPC forms the three-dimensional structure. In TCR subpathway, damage is indirectly recognized during transcript elongation by the stalling of RNA pol II at a lesion by the action of Cockayne syndrome WD repeat

protein CSA and CSB. After damage recognition, TFIIH is recruited to the lesion in both GGR and TCR pathways. In GGR, p62 subunit of TFIIH interacts both with the C-terminal and the amino terminal domain of XPC, whereas XPB interacts only with the C-terminal region of XPC. In TCR, the CSB binds to the stalled RNA pol II and allows the recruitment of TFIIH. Upon binding of TFIIH to the damaged DNA, the CAK subcomplex dissociates from the core TFIIH complex and this dissociation of CAK from core is essential for the unwinding of the DNA (Mangelsdorf, Thummel et al. 1995). The helicase activity of TFIIH further opens the double helix around the lesion which is further stabilized by XPA and the single-stranded DNA binding replication protein A (RPA). RPA binds to the single-stranded, chemically altered nucleotides and coats the undamaged strand. XPA recruits structure specific endonuclease, XPF–ERCC1 heterodimer, which create an incision 5′ to the lesion. This concomitantly allows the recruitments of XPG which cuts the damaged strand at 3′ to the lesion allowing the excision of the 22–30 nucleotide‐long damaged strand. The trimeric proliferating cell nuclear antigen (PCNA) ring, which is directly loaded after the 5′ incision by XPF–ERCC1, recruits DNA Pol δ, DNA Pol κ or DNA Pol ε for gap-filling DNA synthesis. The NER reaction is completed once the final nick is filled by DNA ligase 1(FEN1) or DNA ligase III (XRCC1) (Marteijn, Lans et al. 2014). Besides TFIIH acting both in transcription and NER, the involvement of NER factors in transcription have also been debated since long back when DNA-repair proteins; XPG, XPF and XPC were found to interact with RNA pol II (Maldonado, Shiekhattar et al. 1996). Later several studies came which suggested the involvement of these NER factors in transcription.

III. Human disorders associated with mutation in TFIIH and NER factors

Xeroderma Pigmentosum (XP), Trichothiodystrophy (TTD) and Cockayne Syndrome (CS) are rare, recessive disorders caused by mutation in TFIIH and NER factors. To date, a multitude of mutations in the TFIIH and NER factors have been described which give rise to transcriptional dysregulation and defective NER pathway. According to the early cell fusion studies the NER diseases are categorized into multiple complementation groups and each group subsequently corresponds to patients with mutations in individual genes (De Weerd-Kastelein, Keijzer et al. 1972). For instance eight complementation groups comes under the XP phenotype (XP-A to G and XP-V), two in CS (CS-A and CS-B) and four in TTD (XP-B, XP-D, TTD-A, TTDN1) phenotype. Mutations in the genes of these complementation groups cause a wide range of clinical symptoms, from mild photosensitivity to severe skin cancers, developmental disorders and neurodegeneration. Sometimes mutations in same gene can give rise to different diseases or a variation in the degrees of severity. Genetic studies have shown that mutation in XPB, XPD and XPG give rise to XP or a combination of XP and CS (referred as XP/CS in the rest of the text), TTD is usually caused by mutation of XPD, rarely of XPB, p8/TTD-A and TTDN1. **Figure 13** shows the overlapping genotype-phenotype complexities among different disorders (Kraemer, Patronas et al. 2007; DiGiovanna and Kraemer 2012; Sarasin 2012).

During my PhD, I particularly studied the XP, XP/CS and TTD phenotype caused by different mutations in XPB, XPD and p8/TTD-A gene (Table 7). Each of the mutation and the phenotype is explained in detail below.

Phenotypes	XPB	XPD	TTD-A
XP		R683W	
XP/CS	F99S	G602D	
TTD	T119P	R112H R722W	L21P/R56stop

Table 7. Mutation in XPB, XPD and TTD-A gene studied during PhD

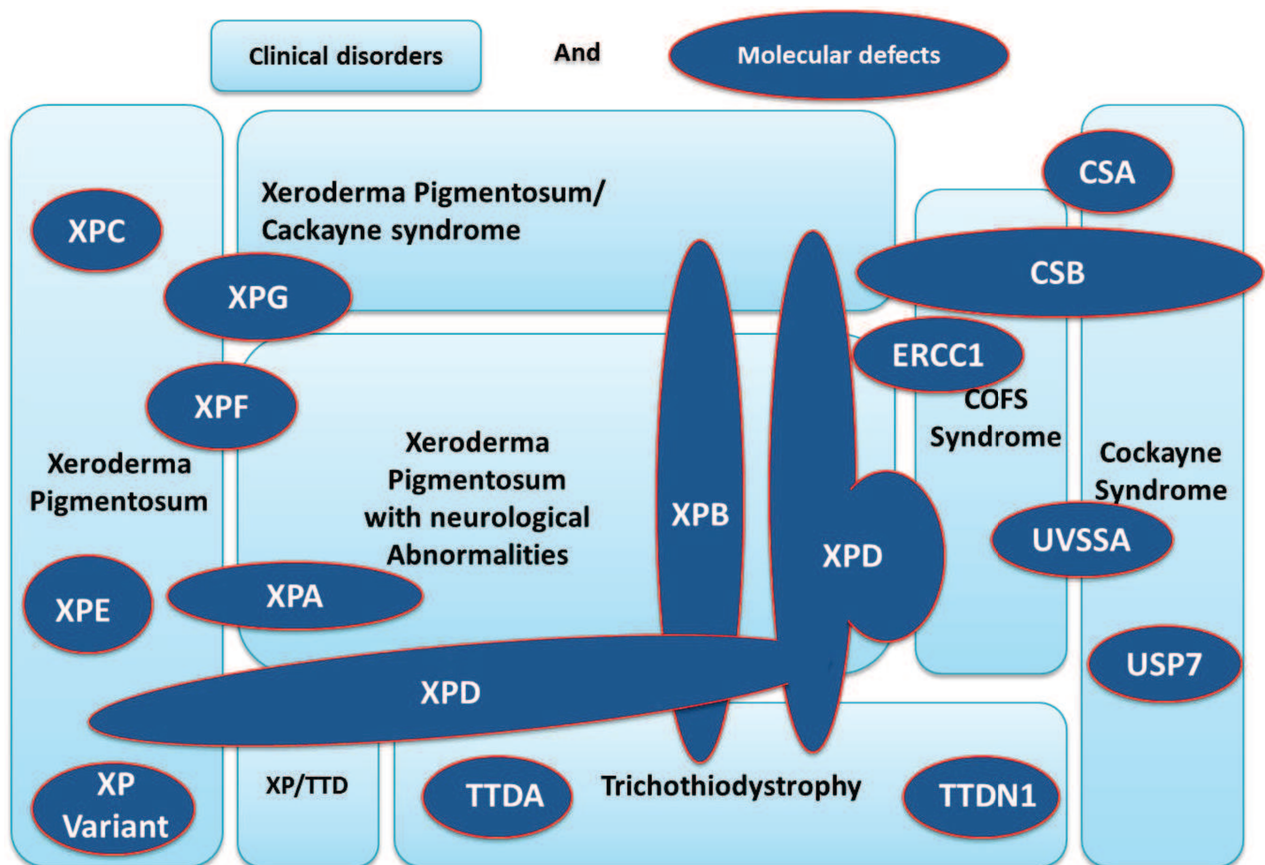


Figure 13. Relationship of genetic disorders to molecular defects.

The genetic diseases (represented in cyan rectangles) induced by different TFIIH and NER factors (represented in dark blue). Different mutations in one gene may result in several different clinical features (Adapted from (Kraemer, Patronas et al. 2007; DiGiovanna and Kraemer 2012; Sarasin 2012)

A. Xeroderma pigmentosum

Xeroderma pigmentosum (XP) was first described in 1874 by Hebra and Kaposi (Hebra and Kaposi 1874). In 1882, Kaposi coined the term *xeroderma pigmentosum* for the condition, referring to its characteristic dry, pigmented skin or **"parchment skin"**. XP is a rare autosomal recessive disorder, characterized by photosensitivity, premature skin aging, pigmentary changes, and malignant tumor development. Compared to the general population, XP patients can develop hundreds of skin cancers. XP patients under age of 20 years have a 10,000- fold increase in the frequency of non- melanoma skin cancer (NMSC), 2000- fold increase in melanomas. They also have 1000-fold high frequency of getting cancer on sun exposed tissues of the eye and 100,000-fold increase on the tongue (Kraemer, Lee et al. 1994; Bradford, Goldstein et al. 2011). XP patients under age of 20 years have an approximately 50-fold increase in cancers of the brain and other central nervous system (Mangelsdorf and Evans 1995) i.e almost 18% of these patients under age 20 are reported to have neurological symptomatology (referred as XP-ND in rest of the text) (Rapin 2013). XP patients with defects in complementation groups A, B, D and G are very likely to have blistering burns on minimal sun exposure; while those having defect in complementation groups C, E and variant-V do not show such sun burns. In XP-V it is not the NER but the post-replication repair is defective because of the Pol H gene. XPA and XPD are the main genes responsible for XP-ND; XPG is the infrequent cause of the XP-ND but the phenotype might be severe. However, all are at high risk to develop early onset freckling, lentigines and skin cancers.

R683W is one of the mutations in XPD gene which I studied. It gives rise to the XP phenotype and is located in C-terminal region of the XPD gene. It has been described extensively in so many studies as this mutation is a hotspot for the XP phenotype (Table 7) and is found as a heterozygous mutation in >80% of XP-D patients (Taylor, Broughton et al. 1997; Renaud and Moras 2000; Kobayashi, Uchiyama et al. 2002).

Curiously, the clinical manifestations of patients are compound heterozygotes for XPD/ R683W, where a second mutation includes patients with or without skin cancers and patients with or without severe neurological impairments. The patient XP135LO is 16-year-old male with R683W mutation and exhibited photosensitivity, multiple cancers and mild neurological abnormalities. The patient cell line for the first time was used to analyze the effect of mutation on the somatic hypermutation which was thought to be affected by impaired DNA repair process (le Maire, Teyssier et al. 2010). Later in 1997, Taylor et al. performed genotype phenotype analyses and showed that it belongs to XP-D complementation group. (Taylor, Broughton et al. 1997).

B. Cockyane Syndrome

Cockayne syndrome was first described by a London physician, Edward Alfred Cockayne in 1936 in two siblings as a syndrome of **"dwarfism, retinal atrophy and deafness".** Since then, more than 180 cases of Cockayne syndrome (CS) have been reported in the literature. CS, an autosomal recessive disorder, has no apparent gender or ethnic predilection. The clinical symptoms may be present at birth and typically develop by age two. Patients usually do not survive to adulthood. Besides, growth and developmental delay, affected children have a typical faces, ophthalmic and auditory disorders and sun-sensitivity. CS severely affects the nervous system, resulting in profound neurodevelopmental impairment such as microcephaly and mental retardation and developmental delay. There are five known complementation groups in CS: CSA, CSB, CS/XPB, CS/XPD and CS/XPG. Mutations in the CSA gene are responsible for the clinical phenotype of CS Type I, the classic form and are present in approximately 25% of CS patients. In contrast mutations in CSB manifest in variable phenotypic expressions. For example, Cerebro-Oculo-Facial-Skeletal syndrome (COFS) or Pena-Shokeir-II syndromes are recognized as a form of CS Type II. CS Type III symptoms are milder than Type I, manifesting later and with slower disease

progression. Another variant of CS is *Xeroderma Pigmentosum*-Cockayne syndrome which contains mixed symptoms of both the phenotypes (XP-CS which is explained in a detail in next section).

CSA and CSB proteins indeed interact with several other proteins and are also implicated in a number of cellular processes other than DNA repair (Table5). There are now several evidences which shows that CSB is implicated in transcription and chromatin remodeling (Vélez-Cruz and Egly 2013). One recent study in fact shows that CSB functions in co-operation with c-Jun and regulate the transcription and chromatin structure (Lake, Boetefuer et al. 2014).

C. Xeroderma Pigmentosum/Cockyane Syndrome

Xeroderma pigmentosum/cockayne syndrome (XP/CS) is a rare phenotype with classical clinical features of XP and CS phenotypes (Lindenbaum, Dickson et al. 2001). In addition to mental retardation, spasticity, hypogonadism and short stature, affected patients have the typical facial freckling and skin cancer predisposition characteristic of patients with XP. Their genotype is XPB, -D, or -G, not CS.

In the course of my thesis, I analyzed the two different XP/CS mutations, one in XPB and other in XPD. Mutation in XPB, **F99S** was reported by Scott et al. (Scott, Itin et al. 1993). This mutation was detected in one of the two sibling patients with XP/CS phenotypes showing the CS neurologic disease and not the XP neurologic disease. At the age of 6 weeks he developed severe sunburn. He walked at 1 year and spoke his first word at 2 years of age. He developed some hearing difficulties at the age of four years. His skin was dry and parchment-like with numerous freckle-like hyperpigmented macules, especially in sun-exposed areas however had no evidence of any malignancy. With age the patient developed several neurological anomalies such as increase in hydrocephalus with pronounced atrophy of cerebellum and an enlarged

fourth ventricle, an increase in the density of the spongiosa, sclerosis of sutures, demyelinating neuropathy and many more. He also developed several sexual abnormalities with increasing age.

Another analyzed mutation with XP/CS phenotype was **G602D,** a substitution in XPD gene and the patient was called XPCS-2. This mutation was first reported by Moshell et al., under XP-H complementation group. Later it was reassigned in XP-D complementation group (Robbins 1991; Vermeulen, Stefanini et al. 1991). Patient XPCS-2 showed symptoms of both CS and XP, included neurological symptoms which was associated with CS rather than with XP. He had early acute sun sensitivity, was freckled, and developed skin cancer at age 2. He also had mental and growth retardation, and died at age 13 from cancer (Takayama, Salazar et al. 1995).

D. Trichothiodystrophy

TTD is a rare autosomal recessive disorder first described by Pollitt et al (Pollitt, Jenner et al. 1968) and termed by Price et al. (Price, Odom et al. 1980). Clinical features of TTD are highly variable in expression, described by the acronyms, PIBIDS, IBIDS and BIDS which represent the Photosensitivity, Icthyosis, Brittle hair and nails, Intellectual impairment, Decreased fertility and Short stature (Itin, Sarasin et al. 2001). Photosensitive form of TTD is caused by the mutation in XPB, XPD and p8/TTD-A subunit of TFIIH. Cellular studies have shown that the photosensitive form normally occurs because of defective NER similar to XP and CS, however non-photosensitive TTD display a normal NER capacity (Stefanini, Lagomarsini et al. 1986). Ten percentage of all the non-photosensitive form of TTD patients are known to have mutation in TTDN1 gene, function of which is not very clear yet. According to a new Clinico-genetic classification, TTD can be classified into three categories i.e. (i) a group with DNA repair anomalies, (ii) a group without DNA repair defect and with TTDN1 mutation,

and (iii) a group without DNA repair defect and without identified genetic basis (Morice-Picard, Cario-André et al. 2009). Study done on the basis of this new classification shows that the frequency of congenital ichthyosis, colloidion baby type was significantly higher in TFIIH mutated group.

I analyzed four different mutations in XPB, XPD and p8/TTD-A subunits leading to TTD phenotype (Table 7). **T119P** mutation in XPB gene with TTD phenotype was first reported by Weeda et al (Weeda, Eveno et al. 1997). The patient was male and was examined first at birth with congenital ichthyosis (collodion baby). The skin condition improved within 3 weeks, leaving a mild ichthyosis of the trunk. TTD was suspected at the age of 3, on the basis of mild ichthyosis of the trunk, which involves the scalp, palms, and soles, mild photosensitivity, lack of second upper incisor, and hair growing normally but coarse with a tiger-tail pattern under polarized light. A diagnosis of TTD was further confirmed by hair microscopy and biochemical analysis showing low cysteine content.

Mutations **R112H** and **R722W** are located in N and C-terminal of XPD gene, respectively. The R112H was reported in an Italian male patient TTD8PV by Stefanini et al. (Han and Grunstein 1988). He was diagnosed at the age of one. At preschool level he showed poor mental development, axial hypotonia and reduced motor coordination. He showed delayed puberty, has short stature and showed moderate UV photosensitivity (Knezetic and Luse 1986). Mutation R722W was diagnosed in a 3 year old male child TTD1BEL from United Kingdom, who died at the age of 3 years. Clinical symptoms of the patient was reported by Stefanini (Stefanini, Lagomarsini et al. 1993), while the detailed biochemical analysis on the patient cells was performed by Broughton et al (Broughton, Steingrimsdottir et al. 1994) which showed a much reduced level of UV- sensitivity.

Features	Xeroderma Pigmentosum (XP)	XP with neurological abnormalities (XP-ND)	Trichothio-dystrophy (TTD)	Cockayne Syndrome (CS)	XP/CS
Skin					
Skin sun sensitivity	yes	Severe	Yes/no	yes	Yes
Lentiginous skin pigmentation	yes	yes	no	no	Yes
Sunlight induced skin cancer	yes	yes	no	no	Yes
Eyes					
Photophobia	yes	yes	yes/no	yes	Yes
Cancer(anterior eye/lids)	yes	yes	no	no	Not reported
Congenital cataracts	no	no	yes	yes	No
Pigmentary Retinal degeneration	no	no	no	yes	Yes
Somatic					
Short stature	no	yes/no	yes	yes	Yes
Immature sexual development	no	no	yes/no	yes	Yes
Nervous system					
Progressive sensorineural deafness	no	yes	no	yes	Yes
Developmental delay	no	yes	yes	yes	Yes
Dysmyelination of brain	no	no	yes	yes	Yes
Progressive neurological degeneration	no	yes	unknown	yes	Yes
Primary neuronal degeneration	no	yes	no	no	No
Atrophy	no	yes	yes/no	yes	Yes
Calcification	no	no	yes/no	yes	Yes
Disease mechanism					
NER defect	yes	yes	yes	yes	Yes
Reaction to exogenous and endogenous DNA damaging agents	Yes-severe	Yes-severe	no	yes	Yes
Molecular Defects	XP A-G XPV (PolH)	XPA, XPB, XPD, XPF, XPG	XPB, XPD, TTD, TTDN1	CSA, CSB	XPB, XPD, XPG

Table 8. Comparison of features of XP, XP-ND, TTD, CS and XP/CS

Adapted from (DiGiovanna and Kraemer 2012; Rapin 2013)

Another patient **TTD1Br**, a 16-year-old English boy with a severe deficiency in excision repair, represented a new excision-repair complementation group which showed mutation in p8/TTD-A and gives rise to photosensitive form of TTD (Stefanini, Vermeulen et al. 1993; Grunstein, Hecht et al. 1995). Table 8, represents the summary of the clinical features associated to XP, XP-ND, TTTD, CS and XP/CS phenotypes.

These cell lines obtained from the patients has benefited us immensely in knowing the molecular mechanism behind the phenotype they represent. Several studies on TFIIH and NER has shown that in fact, these factors play a crucial role in transcription and the clinical features manifested by the patients are more related to transcriptional dyregulation. During my Phd, several questions came to mind. I particularly got interested in understanding the role of TFIIH in transcription and how it affects the interplay between TFIIH and NER factors during transcription. Does TFIIH promote the positioning of the downstream NER factors? Are NER factors also parts of PIC? Does the function of XPG and XPF endonucleases are influenced by TFIIH? How the DNA demethylation which also involves XPG and XPF does get affected by TFIIH? Does the DNA demethyation and histone PTMs regulated by NER factors? How is the CTCF dependent chromatin looping is affected by TFIIH and NER factors? Does DNA demethylation is required for the promoter opening? Answers to these questions could help us to define a hallmark associated for each phenotype and I tried to answer them during my PhD research.

Results

Results

A. Involvement of TFIIH in NER factors dependent chromatin remodeling

Manuscript in preparation

Involvement of TFIIH in the NER factors mediated chromatin remodeling.

Singh Amita, Le May Nicolas*, Jean-Marc Egly*.

Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC)

Department of Functional Genomics and Cancer
CNRS/Inserm/Université de Strasbourg
BP 163, 67404 Illkirch
Cedex, CU Strasbourg, Fr.

*To whom correspondence should be addressed:

Le May Nicolas, Jean-Marc Egly.

Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC)
1, rue Laurent Fries
67404 Illkirch, France
Tel: +33 (0)3 88 65 34 47
email: nlemay@igbmc.fr , egly@igbmc.fr

Abstract:

Fidelity in transcription of the gene requires a congregation of set of proteins around the promoter, upon gene activation. The TFIIH complex is central among these proteins and plays a key role through its enzymatic subunits. Mutations in TFIIH subunits XPB, XPD and p8/TTD-A leads to three distinct autosomal recessive disorders: xeroderma pigmentosum (XP) sometimes associated with Cockayne's syndrome (XP/CS) and trichothiodystrophy (TTD). By studying the different mutation in these three subunits of TFIIH from mentioned genetic disease models, we have shown that each mutation analyzed led to a specific transcriptional dysregulation of the RAR-target gene RAR□2. The architectural and enzymatic integrity of TFIIH condition the appropriate recruitment of TFIIH complex and further the arrival of the Nucleotide Excision Repair (NER) factors. By disturbing their recruitment, mutated TFIIH consequently compromised the chromatin remodeling mediated by NER factors such as histones post-translational modifications (PTMs), DNA breaks induction, DNA demethylation and gene looping. Hence it can be concluded that in addition to its enzymatic activities, TFIIH provide a platform to recruit the NER factors and orchestrates the related functions in transcription. Such varying penetrance among mutants gives rise to a phenotype gradient as observed in TTD, XP or XP/CS patients.

Introduction

TFIIH, a well-conserved transcription factor during evolution plays a key role in maintaining the genome stability as well as in transferring the genetic information. Mammalian TFIIH includes a core, (containing the six subunits XPB, p62, p52, p44, p34, and p8/TTD-A) bridged by XPD to the cdk-activating kinase module (CAK, composed of the three subunits cdk7, cyclin H, and MAT1) (Figure 1A). In addition to its function in the RNA polymerases transcription, TFIIH is also involved in nucleotide excision repair (NER) pathway, thus illustrating the important interplay existing between both so disparate processes (Feaver, Svejstrup et al. 1993; Schaeffer, Roy et al. 1993; Drapkin, Reardon et al. 1994). NER, that is responsible for the removal of a variety of bulky DNA adducts, such as those induced by UV irradiation, is subdivided into two sub pathways: the Global Genome Repair (GGR) is responsible for the removal of DNA lesions from the whole genome and Transcription Coupled Repair (TCR) is responsible for the accelerated removal of lesions located on the transcribed strand of active genes (Mellon, Spivak et al. 1987). In GGR, the XPC-hHR23B complex recognizes the damage-induced DNA distortion, while in TCR the RNA polymerase II (pol II) stalled in front of a lesion promotes the recruitment of the TCR-specific proteins CSB and CSA (Kamiuchi, Saijo et al. 2002). Both NER sub-pathways then funnel through the action of TFIIH that unwinds the DNA via the ATPases/helicases activities of XPB and XPD regulated by p52/p8-TTD-a and p44 respectively. This event favors the recruitment of XPA and RPA assisting in the expansion of the DNA bubble around the damage and the arrival of XPG and XPF endonucleases. XPG and XPF then generate cuts in the 3' and 5' sides of the lesion respectively, thereby removing the damaged oligonucleotide before the re-synthesis machinery fills the DNA gap. In protein coding genes transcription, XPB is involved in the opening of the promoter while the cdk7 kinase of TFIIH phosphorylates the serine 5 of the carboxy-terminal domain (CTD) of pol II as well as the activators. Interestingly, NER factors were found to be associated and participate to the regulation of gene expression (Barreto, Schafer et al. 2007; Schmitz, Schmitt et al. 2009; Le May,

Mota-Fernandes et al. 2010). It remains to be established whether TFIIH influences the recruitment and the function of the NER factors at the promoter of activated genes to achieve chromatin remodeling required for accurate transcription.

The importance of TFIIH and its relationship with the other NER factors is put forward by the existence of human genetic disorders with a broad range of clinical features (Table1). Indeed, mutations in the XPB, XPD and p8/TTD-A subunits of TFIIH originate three distinct autosomal recessive disorders: xeroderma pigmentosum (XP), sometimes associated with Cockayne's syndrome (XP/CS) and trichothiodystrophy (TTD) (de Boer and Hoeijmakers 2000; Kraemer, Patronas et al. 2007; Cleaver, Lam et al. 2009). XP is characterized by numerous skin abnormalities ranging from excessive freckling, ichtyosis to multiple skin cancers, the frequency of which is about 2000-fold greater than in normal individuals. In addition to a hyper-photosensitivity to sunlight, XP patients display a progressive neurological degeneration (Kraemer, Lee et al. 1987; Lehmann, McGibbon et al. 2011). Although XP can be discriminated from CS phenotype there are rare cases of individuals with the combined symptoms of XP and CS (XP/CS). XP/CS patients display a combination of the cutaneous abnormalities from XP with the severe neonatal later onset of neurological and developmental anomalies typical of CS. The typical hallmark of TTD is sulphur-deficient brittle hair, caused by a greatly reduced content of cysteine-rich matrix proteins in the hair shafts. Mental retardation and ichthyosis also characterize TTD patients (Sarasin A 1992; Itin, Sarasin et al. 2001).Some TTD patients are sensitive to sunlight without any unusual pigmentation changes and severe skin lesions or cancer (Stefanini, Lagomarsini et al. 1993).

Accumulating evidences suggest that the clinical features of these three disorders due to mutations in all the NER factors including TFIIH cannot be exclusively explained on the basis of DNA repair defects but also involve transcription deficiencies (Evans, Moggs et al. 1997; Keriel, Stary et al. 2002; Dubaele, Proietti De Santis et al. 2003; Coin, Oksenych et al. 2007) (Table 1). In this study, we intend to understand how mutations in XPB, XPD and p8/TTD-A result in impairment of gene expression. Knowing the close connections between TFIIH and NER factors in the removal of DNA damage

(Zurita and Merino 2003; Schärer 2008; Compe and Egly 2012), we have investigated the consequences of mutations found on three of the TFIIH subunits on the recruitment of the NER factors and the roles of the later in the various steps that lead to accurate RNA synthesis including histones post-translational modifications (PTMs), DNA breaks, DNA demethylation and gene looping. In addition to improve our understanding of gene expression regulation, such systematic approach could help to determine transcriptional default hallmarks to molecularly define the different genetic disorders.

Results:

Mutations in TFIIH subunits compromise the formation of the transactivation complex

To further determine the transcriptional defects due to the mutations in XPB, XPD and p8/TTD-A subunits of TFIIH, we analyzed cells derived from XP, XP/CS and TTD patients bearing mutations as indicated (Table 1 and figure 1). To investigate the transcriptional process, we focused on the Retinoic Acid Receptor gene *RARβ2* as a model. Few hours post all-trans retinoic acid (t-RA) treatment of the TFIIH mutated cells (Materials and Methods), we observed different patterns in the *RARβ2* mRNA synthesis compared to their respective XPBwt, XPDwt and the rescued p8/TTD-Awt cells (Figure 2A1, B1 to J1). The synthesis of *RARβ2* mRNA was significantly lower for XPB/F99S and XPB/T119P compared to XPBwt that peaked at 8 h post-treatment (Figure 2A1, B1 and C1). In the four XPD mutated cells, the inductions of *RARβ2* were similarly reduced compared, to the cells rescued with an XPDwt (Figure 2F1 to J1). We did not notice any reduction in the *RARβ2* mRNA level between p8/TTD-A and its rescued (Figure 2D1 and E1).

We next evaluated the dynamic recruitment of pol II partners at the *RARβ2* promoter, using chromatin immunoprecipitation (ChIP) overtime. ChIP analyzed by quantitative PCR showed a concomitant recruitment of the retinoic acid receptor (RAR), pol II and the general transcription factor TFIIB at 6/8h post t-RA treatment in XPBwt, XPDwt and p8/TTD-Awt (Figure 2A2, D2 and F2 respectively). At this time, TFIIH was

also recruited as visualized by the presence of its XPB, XPD, p44 and cdk7 subunits (Figure 2A3, D3 and F3). We also detected the simultaneous presence of the XPA, XPG and XPF NER factors (Figure 2A4, D4 and F4).

Our analysis next indicated that each mutation in the XPB, XPD and TTD-A/p8 subunit of TFIIH led to different and specific deregulations in the recruitment of the components of the transactivation complex. In XPB/F99S RAR was detected at early time, whereas pol II and TFIIB accumulated at the promoter after 1h post t-RA induction (Figure 2B2). We also noticed a non-concomitant recruitment of the TFIIH subunits and the NER factors. In XPB/F99S cells, while XPG was recruited at 3h, XPA and XPF were not detected until 12h (Figure 2B4). The XPB/F99S mutation prevents the accurate opening of the promoter and consequently the recruitment of the NER factors within the transcription complex (Coin, Bergmann et al. 1999; Coin, Oksenych et al. 2007). In XPB/T119P cells, RAR, pol II and TFIIB were detected at the promoter around 3h (Figure 2C2). The XPB subunit was found at the promoter at 1h while XPD, p44 and cdk7 subunits were detected later (Figure 2C3). Similarly the XPA arrival at 1h precedes XPG (at 6h), whereas XPF is not detected (Figure 2C4). It therefore appeared that the recruitment and the formation of TFIIH complex could be initiated at the promoter upon gene induction.

In p8/TTDA deficient cells, we repeatedly observed a synchronized recruitment of TFIIH subunits and NER factors with the transcriptional machinery at 8h post t-RA treatment that is similar to what was observed with the corresponding rescued wild type cell (Figure 2E2 to E4 and D2 to D4).

We next focused on the XPD mutations resulting in three different patient phenotypes (Table 1). In XPD/R112H cells, the transcription and NER factors are recruited in a cyclic and synchronized manner. All these factors were found in a second cycle of recruitment that peaked at 8h, remained present until 12h post t-RA treatment and paralleled the *RARβ2* mRNA induction (Figure 2G1 and G2-G4). In the XPD/G602D cells, the recruitment pattern at 1 h and 6 h post tRA-treatment of pol II, TFIIB, TFIIH as well as the NER factors at the promoter was similar to XPDwt cells

(Figure 2H2-H4). The XPD/G602D mutation that affects the helicase motif V does not seem to disturb the TFIIH architecture (Dubaele, Proietti De Santis et al. 2003). In XPD/R683W, the recruitment of TFIIH and the NER factors was highly deregulated at early time post t-RA treatment (Figure 2I2-I4). At 8h only pol II, TFIIB and XPA were detected (Figure 2I2 and 2I4). In XPD/R722W in which the mutation is located in the C-terminal unfolded domain, we observed the recruitment of both transcription and NER factors at 8h (Figure 2J2-J4); a continuous accumulation of pol II, TFIIB as well as some of TFIIH subunits was also visible at 12h. In these cells, the R683W and R722W mutations in XPD are known to weaken the interaction with the p44 subunit and consequently destabilize the architecture of TFIIH (Botta, Nardo et al. 2002; Keriel, Stary et al. 2002), see discussion.

Altogether, our data suggested that XPB and XPD mutations disturbed the *RARβ2*-gene activation by impeding the formation of the pre-initiation complex at the promoter of activated genes. Although each dysregulation is mutation-specific, we noticed a compromised integrity of TFIIH complex and in some cases the absence of NER factors at the promoter. Such observations suggested a link between the presence of an accurate TFIIH complex at the promoter of activated genes and the concomitant recruitment of the NER factors.

Impaired chromatin remodeling reveals the implication of TFIIH

Previous works have underlined the sequential recruitment of the transcription and the NER components at the promoter of activated gene and their role in conditioning the chromatin modifications e.g. histone post-translational modifications (PTMs), DNA breaks and DNA demethylation (Le May, 2010, 2012, Schmitz, 2009, Barreto, 2006): Euchromatin is characterized by acetylated (H3K9-Ac and H4K16-Ac) and methylated (H3K4me3, and H3K79me2) histone H3 and H4 and allows transcription while heterochromatin that inhibits RNA synthesis is characterized by chromatin marks such as di- and trimethylated H3K9 (H3K9me2-3) and H3K27 (H3K27me2) (Bannister and Kouzarides 2011)

In t-RA treated XPBwt and XPDwt cells, we observed upon ChIP analysis, the characteristic histones PTMs signature (increase of H3K9ac and H3K4me2 concomitantly to a decrease H3K9me2, Figure 3A1 and F1, coincided with the *RARβ2* mRNA induction peak) at 8h. In p8/TTD-A wt cells, the H3K9me2 is not detectable (Figure 3D1). The typical histone signature observed in wild type cells at the promoter of activated *RARβ2* is not observed in all the mutated cells tested so far (Figures 3B1, C1, E1, G1, H1, I1 and J1) indicating a persistence of active histones PTMs around the *RARβ2*-promoter. As an example, in XPD/R112H cells in which the transcription and the NER factors were shown to be recruited at the promoter 8h post-treatment, H3K9 was not demethylated, while we noticed a high acetylation of H3K9 and methylation of H3K4 (Figure 3G1).

Knowing how crucial was the presence of NER factors, and particularly of XPG and XPF endonucleases in *RARβ2* expression, we next evaluated the formation of transient DNA breaks in its promoter surrounding (Ju, Lunyak et al. 2006; Le May, Fradin et al. 2012). Using a BioChIP assay that measures the incorporation of biotinylated dUTP within broken DNA; we observed a concomitant increase of DNA cleavage at both promoter and terminator of all the wild type cells upon t-RA activation (Figure 3A2, D2, F2 and supplemental figure S2). In these cells, we noticed a parallel between the presence of XPG and XPF at both the promoter and the terminator and the DNA breaks. Except in the XPD/R112H and XPD/R722W in which we noticed the recruitment of XPG and XPF together with the presence of some DNA breaks (Figure 3G2 and 3J2), in all the other mutated cell lines we did not observed any significant formation of DNA breaks nearby the *RARβ2* promoter (Figure 3B2, E2, H2 and I2). We also noticed some DNA breaks (although to a much lower extent) at the *RARβ2* promoters in XPB/T119P (Figure 3C2). Surprisingly, there was no detectable presence of DNA breaks in p8/TTD-A cells (Figure 3E2), whereas all the NER factors are present together with the transcriptional machinery both at promoter and at terminator (Figure3 E2-E4 and Supplemental figure S2). It has to be noticed that the DNA breaks were observed around the terminator of *RARβ2* in wild type cells, whereas in all the

other mutated cell lines, no DNA breaks were detected (Supplemental figure S2).

Several studies have documented a relationship between XPG and DNA demethylation upon transcription (Barreto, Schafer et al. 2007; Le May, Fradin et al. 2012) Using an unmethylated DNA immuno-precipitation (unMeDIP) approach, we measured the removal of the 5-methylcytosine (5mC) at *RARβ2* promoter. We found that DNA demethylation occurred at the promoter at 8 h post t-RA in the three wild type cells and perfectly paralleled the recruitment of the entire transcription machinery (Figures 3A3, D3 and F3). On the contrary, there was a complete lack of DNA demethylation in all the XPB and XPD cells (Figures 3B3, C3, and G3-J3) except in the p8/TTD-A cells (Figure 3E3 and Supplemental figure 2).

The above data strongly supported an involvement of TFIIH in the recruitment of the entire transcription machinery, in the chromatin remodeling including the histone PTMs, the formation of DNA breaks and active DNA demethylation.

Chromatin looping is perturbed without proper assembly of TFIIH and NER factors

The detection of the basal transcription machinery together with the NER factors at both *RARβ2* promoter and terminator is correlated with the formation of a chromatin loop mediated by the CCCTC- binding factor (CTCF) chromatin organizer (Splinter, Heath et al. 2006; Le May, Fradin et al. 2012). Such gene looping was shown to parallel DNA demethylation and DNA breaks at both regions. Quantitative chromatin conformation capture assays (q3C) were performed to analyze the interactions between -65 kb, promoter (Pro), terminator (Ter), or +323 kb as well as an intronic region (M1) of the *RARβ2* locus (Figure 4, upper scheme). Using Ter and M1 as a bait, we observed that Pro could specifically and significantly interact with Ter at 8 h in t-RA-treated wild type cells, (Figure 4A, D and F), paralleling the recruitment of the entire transcriptional apparatus on both regions (Figure 1 and Supplemental Figure S1). By contrast, in all the mutated TFIIH cells, no increase in the interaction/ frequency between Ter and Pro was revealed from t=0 to T= 6/8h post t-RA induction (Figure 4B, C, E, G-J). It is worthwhile to notice that the lack in loop formation between Pro and Ter could be related to defect

in the recruitment of some of the transcription and NER factors at Pro and Ter in most of the mutated cells (Figure 2 and Supplemental figure S1). In addition, we also observed that in these cells CTCF was either undetectable or delayed in its recruitment at both Pro and Ter of RARβ2 (Figure 2B4, C4, G4, H4, J4). 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 (Supplemental figure S3).

By impeding the accurate recruitment of NER factors at the activated RAR*β*2 promoter, TFIIH further disturbed the gene looping conformation required for optimal gene expression.

Mutations in TFIIH impair some of its enzymatic activities

We next addressed the contribution of XPB and XPD activities to the formation of an accurate transcription initiation complex, a prerequisite for optimal RNA synthesis. We first generated recombinant rIIH6 (the core TFIIH containing p62, p52, p44, p34, p8 and either XPBwt or XPBmut), XPDwt or XPD mut and CAK. These rIIH6 sub-complexes were added to an *in vitro* transcription assay containing, in addition to the adenoviral major late promoter (MLP, run-off of 309nt long), all the basal transcription factors and pol II (Gerard, Fischer et al. 1991) and when indicated, CAK and XPD either alone or in combination.

When added to the transcription assay that contains all the factors including the XPD/CAK sub-complex, rIIH6-XPB/F99S exhibits a much weaker basal transcription activity than the wild type and rIIH/XPB/T119P (Figure 5A). Addition of CAK together with XPDwt and rIIH6-XPB/F99S did not improve its transcriptional activity (lanes 5-7) contrary to what occurs with rIIH6-XPBwt and -XPB/F119P that requires absolutely its CAK sub-complex for optimal RNA synthesis (lanes 1-3 and 8-10). Since XPB/F99S mutation weakens the contact with the p52 regulatory subunit within TFIIH (Coin, 2007), it results in a defect in the unwinding of the *RARβ2* promoter by XPB and a defect in RNA synthesis. Interestingly, we also noticed that the absence of XPDwt

results in a very weak RNA synthesis (Figure 5A, lane 4 and Figure 5B, lanes 1 and2). Knowing that the XPD-p44 interaction, conditions the anchoring of CAK to the core TFIIH (Coin, Oksenych et al. 2007), we next investigated the transcription activity of XPD mutated TFIIHs. Addition of CAK and increasing amount of XPDwt to our transcription assay that already contained rIIH6wt stimulated RNA synthesis (lanes 3-4) as also observed in the presence of XPD/R112H or XPD/G602D (lanes 5-8). On the contrary, when either XPD/R683W or XPD/R722W was added instead, there was no significant increase of RNA synthesis (lanes 9-12). In these latter cases, XPD/R683W and XPD/R722W mutations prevented a contact with p44 and consequently impeded the anchoring of CAK to the core TFIIH to exploit its kinase activity towards pol II (Coin, Oksenych et al. 2007). CAK was shown to phosphorylate the serine 5 of the carboxy- terminal domain (CTD) of the largest subunit of pol II, and allow promoter escape (Lu, Zawel et al. 1992; Bensaude, Bonnet et al. 1999).

We then investigate the impact of CAK on the phosphorylation status of pol II. pol II phosphorylation status was carried out following a classical run-off transcription experiment (Materials and Methods). Using antibody directed against the CTD, we observed that the hyper-phosphorylated form of pol II (IIo) was prevalent in the presence of XPD wt, XPD/R112H or XPD/G602D (Figure 5C, lanes 2-7) and paralleled the increase in RNA synthesis (Figure 5B, lanes 3-8). On the contrary, in the presence of XPD/R683W and XPD/R722W which were deficient in stimulating RNA synthesis (Figure 5A, lanes 9-12), pol II was not hyper-phosphorylated (Figure 5C, lanes 8-11). Moreover ChIP experiments further demonstrated the presence of phosphorylation of Serine5 of pol II CTD in XPB/T119P, XPD/R112H, XPD/G602D as in the wild type cells, at RARβ2 promoter (Figure 5D). In all the other XPD, XPB cell lines, we did not detect phosphorylated pol II (Figure 4E).

Altogether the above data suggest that the gene expression dysregulation observed in some XPB and XPD cells might results at least in part from a defect in phosphorylating pol II at the transcription initiation level.

Discussion

The present study underline the key role of TFIIH in (i) the synchronized recruitment of the NER factors, (ii) the histones modifications, (iii) the DNA breaks and DNA demethylation as well as (iv) gene looping, four crucial steps in transcription. Each mutation that affects one of the TFIIH subunits specifically disturbs the transcription process of activated genes.

TFIIH integrity and enzymatic activities

Although in wild type cells, we observed upon t-RA induction a concomitant and synchronized recruitment of all the TFIIH subunits at the RARβ2 promoter (Figure 2), in XPB and XPD cells, TFIIH subunits recruitment is rather dispersed. Indeed, in XPB/F99S cells, XPD subunit is found late at the *RARβ2* promoter than the other TFIIH subunits (Figure 2B3), while in XPB/T119P cells, XPB is recruited at 1h while the other TFIIH subunits are detected at 6h (Figure 2C3). Similarly in XPD/R683W only subunits XPD and Cdk7 are the detected at early time, whereas in XPD/R722W cells, the arrival of TFIIH subunits at the promoter occurred over a much larger period of time in an unsynchronized way (Figure 2I3, J3).

In Rift Valley Fever Virus infected cells, p44 as well as XPB subunits of TFIIH were found in the nucleus, while XPD was maintained in the cytoplasm (Le May et al., 2004; Kainulainen et al., 2014). As a consequence we observed a transcriptional defect not only due to the trapping of some of the subunits as a filament but also for the absence of others such as XPD that could not enter in the nucleus. All these observations suggest that the formation of TFIIH occurs at the promoter upon gene induction to initiate RNA synthesis.

Any mutation on XPB, XPD and p8/TTD-A that would prevent the accurate recruitment of and/or the interaction network between the TFIIH subunits, would disturb transcription. The XPB/F99S mutation weakens its interaction with p52; consequently p52 could no more up-regulates XPB to unwind the activated

promoter(Coin, Oksenych et al. 2007). Similarly the XPD/R683W and XPD/R722W mutations that weaken the contact with the p44 subunit and consequently the anchoring of CAK to the core TFIIH are detrimental for both in vivo and in vitro transcription (Figures 2 and 5). In absence of accurate phosphorylation by CAK, pol II cannot be used as a docking site for enzymes required for PTMs of histones such as H3K4me3 and the un-phosphorylated nuclear receptors will be defective in regulating their responsive genes. In these cases, the transcription inhibition results from defect in phosphorylation by CAK with some consequences in chromatin modification (Figure 5) rather than defect in XPD helicase regulation by p44. This demonstrates once more that the XPD helicase activity is not crucial for transcription. It however remains that the gene expression defects could not be exclusively explained by a lack of the enzymatic activities of the mutated TFIIH. Indeed, we observed that the XPB/T119P, XPD/R112H and XPD/G602D mutations did not inhibit the *in vitro* transcription activity of the corresponding TFIIHs whereas mRNA induction of RARβ2 is defective in *vivo* (Figure 5 and Figure 2).

TFIIH a platform for recruiting NER factors

It seems that the TFIIH mutations might disturb either directly or indirectly the accurate positioning of the various components around the transcription initiation. Indeed close physical connections between TFIIH and transcription and NER factors are required for accurate RNA synthesis. We show here that the XPD/R7222W mutation impedes the recruitment of TFIIB, TFIIE and TFIIF at *RARβ2* promoter (Figure 2 & supplemental figure). By stabilizing TFIIH architecture through interaction with XPB and XPD, XPG regulates the expression of certain nuclear receptor responsive genes (Ito, Kuraoka et al. 2007). It is also likely that mutations in TFIIH could disturb its interaction with RPA, XPA (Coin, Oksenych et al. 2007), or XPF (Coin, Auriol et al. 2004) as already highlighted for NER. Here we observed that each TFIIH mutation specifically disturbs the sequential recruitment of the NER factors. The XPB/F99S mutation affects the recruitment of XPA and XPG (Figure 1B4), a point that was also

observed in NER (Coin, Oksenych et al. 2008) . In XPB/T119P, XPA and XPF recruitment does not parallel the one for TFIIH and XPG (Figure C3 and C4). In XPD/R602D cells, XPF is not detected at the *RARβ2* promoter. Altogether, our results strongly suggest a role of TFIIH as a platform to recruit the NER factors and organize the transactivation complex.

TFIIH orchestrates the chromatin remodeling

One of the crucial question concerns the interplay between TFIIH and the NER factors in the cascade of events that lead to the formation of a chromatin environment suitable for gene expression. For all the TFIIH mutations we observed a misregulation of H3K4me2, H3K9ac and H3K9me2 histone modification state around the *RARβ2* promoter that could consequently prevent the establishment of a permissive environment for transcription. The defective histone modifications would therefore disturb the DNA methylation (D'Alessio, Weaver et al. 2007) and probably DNA breaks formation (Figure 3). Several studies have documented a relationship between DNA demethylation and DNA repair mechanisms upon transcription (Barreto, Schafer et al. 2007; Wossidlo 2010; Le May, Fradin et al. 2012). The DNA breaks surrounding the RARβ2 promoter that are visible in all the wild type cells, were not detected in mutated cells except in the XPD/R112H, XPD/R722W and to a lower extend in XPB/T119P cells (Figure 3). In these cells as well as in wild type ones, XPG and XPF are present at both promoter and terminator of *RAR*β2 (Figure3) suggesting a relationship between the presence of those endonucleases and DNA breaks formation (Ju, Lunyak et al. 2006; Le May, Fradin et al. 2012). How DNA breaks and DNA demethylation are connected is unclear since in XPD/R112H and XPD/R722W cells there are DNA breaks and no detectable DNA demethylation while in p8/TT-DA we observed the inverse situation (Figure 4).

In the p8/TTD-A cells, we did not observe any default in the recruitment of the transcriptional machinery, TFIIH subunits and NER factors, the typical histones PTMs

signature for active transcription, DNA demethylation. Surprisingly, no DNA breaks induction and gene looping were detected. This allows us to suggest that the DNA methylation is unrelated to DNA breaks formation and such DNA modification is important to avoid temporal aberrant chromatin rearrangement but is not obligatory for mRNA synthesis. Another hypothesis would propose that p8/TDD-A besides mediating the DNA unwinding in NER might also regulate either directly or indirectly some DNA breaks.

Altogether our data underline the key role of TFIIH in the transcription process, in which it will combine its primary enzymatic activities together with the recruited NER factors and to further orchestrate the downstream events such as histones PTMs, DNA breaks and DNA methylation. All the cells bearing the mutations on XPB, XPD and p8/TTD-A that derived from XP, XP/CS, TTD patients, present a common transcriptional defect but a specific (and unique) pattern of dysregulation with hallmarks that are not suitable for the gene expression. As a function of the location of the mutation, we observed that the heterochromatin to euchromatin transition is rather incomplete. These subtle transcriptional differences that varies among mutants seems gives rise to a phenotype gradient as observed on these XP, XP/CS, TTD patients, in addition to some specific DNA repair deficiencies that could also have indirect effect in gene expression.

Acknowledgments

We thank F. Coin for fruitfull discussion, F. Costanzo for reading the manuscript, C. Braun for helping in in vitro experiments and the Institute of Genetics and Molecular and Cellular Biology Cell Culture Facility. This study was supported by grants from the European Research Council Advanced Scientists, l'Association de la Recherche contre le Cancer (ARC-SL2201006011335 &30607082315 to JME), l'Agence Nationale de Recherche (ANR12.BSV80017-01) La ligue nationale contre le Cancer « Equipe labellisée Ligue »and l'Association nationale des membres de l'Ordre National du

Mérite ANMONM). A.S. was supported by a JME (Jean-Marc Egly) Advanced European Research Council grant and by an ARC fellowship for young scientists.

Material and Methods:

Cell Lines, culture conditions and transfection.

XPB wt, XPB/F99S and XPB/T119P are SV40-transformed human fibroblasts (XPCS2BASV) expressing His-ERCC3-HA(Winkler, Vermeulen et al. 1998), XPB-T296C (Cl. 14, XPB/F99S) and XPB-A355C (Cl. 5, XPB/T119P) respectively (Riou, Zeng et al. 1999). p8TTDA cells are SV40-transformed human fibroblasts (TTD1Br-SV) bearing the TTDA L21P and R56Stp mutations(Stefanini, Vermeulen et al. 1993). The p8TTDA wt (TTD1Br-SV + TTDA-GFP) cells are rescued TTDBr1 with pEGFP-N1-TTDA vector stably expressing TTDA-GFP(Giglia-Mari, Miquel et al. 2006). XPD wt (GM637) cells are SV40-transformed human fibroblasts from a normal 18-year-old female. XPD/G602D (XPCS2) (Takayama, Salazar et al. 1995), XPD/R112H (TTD8PV) (Stefanini, Giliani et al. 1992), XPD/R683W (XP135LO) (Taylor, Broughton et al. 1997) and XPD/R722W (TTD1BEL) (Broughton, Steingrimsdottir et al. 1994) are human primary fibroblasts. All cells were cultured in appropriate medium. Cells were incubated with red phenol-free medium containing charcoal-treated fetal calf serum (FCS) and 40mg/ml gentamycin. Cells were treated with 10µM of all-trans retinoic acid (t-RA, MP biomedicals) into the same medium. XPD/R112H, XPD/G602D, XPD/R683W and XPD/R722W cells were transiently transfected 48 h before the t-RA treatment with pEGFP-XPD WT using the X-tremeGENE HP DNA transfection reagent (Roche).

Antibodies.

Mouse monoclonal antibodies towards XPB (1B3), XPD (2F6), p44 (1H5), RAR (9A6), XPA (1E11), XPG (1B5), RNA pol II (7C2) were from IGBMC antibody facility. Cdk7 (C-19), TFIIB (C-18), XPF (H-300), Biotin (33) antibodies were obtained from Santa-Cruz Biotechnology. CTCF (ab70303) and RNA pol II ser5P (61085) antibodies were obtained

from Abcam and Active Motif respectively. H3K4me2 (#9726), H3K9me2 (#9753), H3K9Ac (#9671) antibodies were purchased from Cell signaling technology.

Reverse transcriptase 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 performed using the QuantiTect SYBR Green PCR kit (Qiagen) and the Lightcycler 480 (Roche). The primer sequences for Retinoic Acid Receptor isoform β2 (*RARβ2*) and glyceraldehyde 3-phosphate dehydrogenase (GAPDH) genes used in real-time qPCR are listed (see Table in supplemental data). *RARβ2* mRNA levels were normalized against the *GAPDH* mRNA.

Chromatin Immunoprecipitation (ChIP) and Biotin-ChIP

Cells were cross-linked 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 using a set of primers targeting the promoter and terminator regions of *RARβ2* (see Table 2 in supplemental data).

Cross-linked cells prepared as described above were permeabilized with cytonin (Gentaur biologique) 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 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 assay. The purified DNA fragments were analyzed by quantitative PCR using a similar set of primers described above.

Unmethylated DNA immunoprecipitation

Genomic DNA was extracted using GenElute Mammalian Genomic DNA Mini-prep Kit (Sigma). Unmethylation of 5mC on the promoter and terminator regions of *RARβ2* was measured by digesting 2 mg 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 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.

Quantitative chromosome conformation capture (q3C).

The quantitative chromosome conformation capture (q3C) assay was performed as previously described (Vernimmen, De Gobbi et al. 2007; Le May, Fradin et al. 2012).

Construction of baculoviruses and purification of complexes

Baculoviruses over-expressing XPB, XPD, p62, p52, p44, p34, cdk7, cyclin H, MAT1, and p8 in sf21 cells were produced as previously described (Dubaele et al., 2003; Coin et al., 2007). For expression of recombinant mutated Flag-XPB and Flag-XPD, the cDNA encoding XPB and Flag-XPD cloned in pSK278 and PVL1392 vectors respectively were used for PCR site-directed mutagenesis. The resulting vectors were recombined with baculoviruses DNA (BaculoGold; PharMingen). The recombinant viruses were purified from isolated plaques and viral stocks were prepared by three-step growth amplification.

Sf21 insect cells were infected with WT or mutated Flag-XPB, WT or mutated Flag-XPD, p62, p52, p44, p34, Flag-cdk7, cyclin H, Mat1 and p8 baculoviruses allowing the preparation of core-IIH, CAK, and XPD separately. The whole-cell extracts were

initially incubated 4 h, with agarose beads bound to anti–M2-Flag antibody at 4°C. After washing with a buffer containing, 50 mM Tris/HCl (pH 7.9), 20% glycerol, and 50 mM KCl, rIIH9 complexes were eluted for 8–12 h with the same buffer containing 0.2 mg/ml of the epitope peptide (Jawhari et al., 2002). For the core-IIH, the eluted fraction was then incubated with anti-p44 (1H5) antibody in a buffer containing 50 mM Tris/HCl (pH 7.9), 20% glycerol, and either 50 or 150 mM KCl. The immunoprecipitated fraction was then either boiled or eluted with a synthetic peptide recognized by Ab-p44. The recombinant TFIIH was made by mixing purified core-IIH, CAK, and XPD allowing the preparation of the different XPD and XPB variants.

***In-vitro* assays**

Run-off transcription assays were performed using recombinant TFIIB, TFIIE, TFIIF, TBP, endogenous RNA pol II, and the different TFIIH (rIIHs), as previously described (Gerard et al., 1991).

RNA Pol II phosphorylation was carried out as a classical run-off transcription assay as previously described (Ueda et al., 2009). Hypophosphorylated (IIA) and hyper phosphorylated (IIO) forms of RNA Pol II were resolved by SDS/PAGE and detected using the monoclonal antibody (7C2) directed against the CTD.

Figure 1

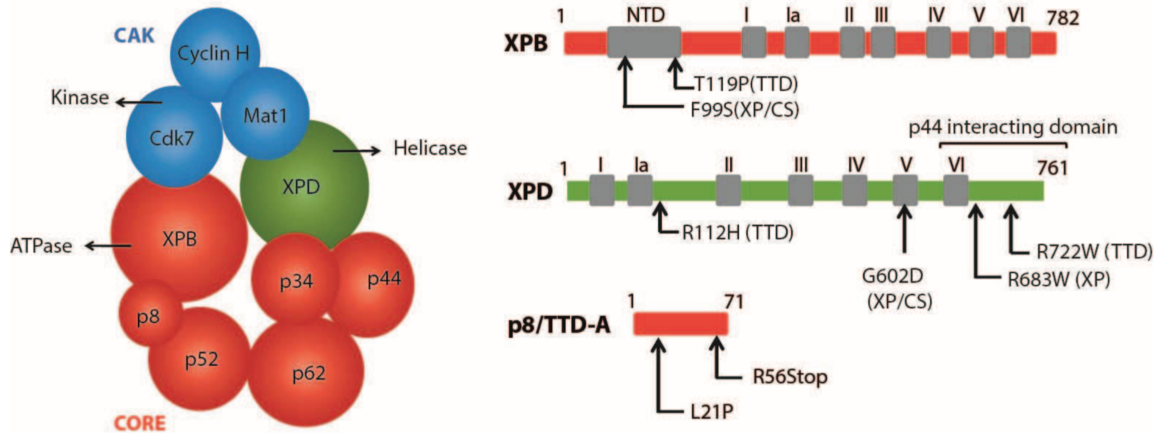


Figure 1. Schematic representation of TFIIH: The CAK sub-complex (in blue) is bridged to the core TFIIH (in red) by the XPD helicase (in green). Mutations in the XPB, XPD and p8/TTD-A subunits as well as the XP, XP/CS and/or the TTD are indicated. Black squares indicate the helicase motifs (I, Ia, II, II, IV, V, and VI); NTD: N-Terminal domain.

Figure 2

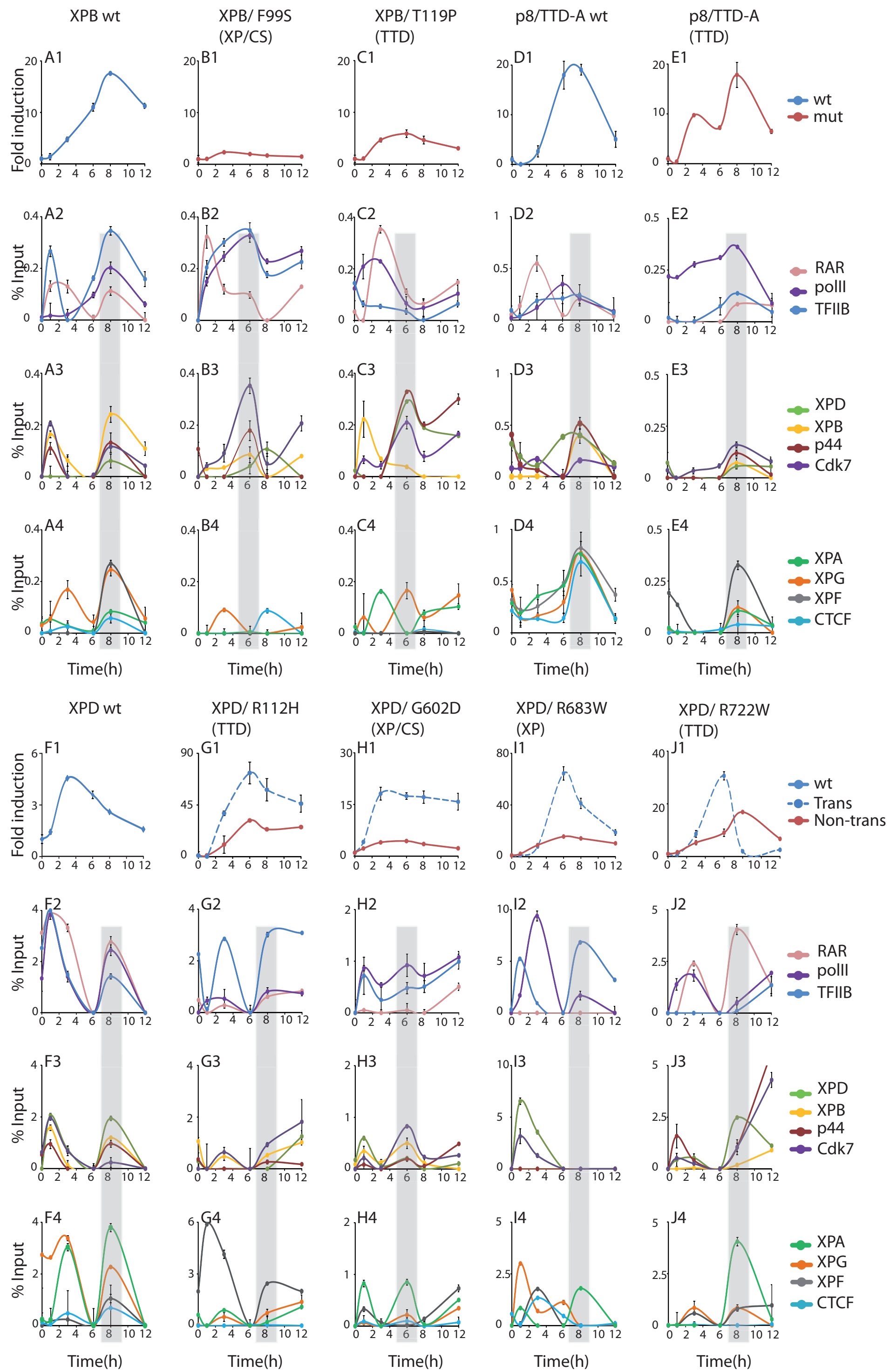


Figure 2. ***RARβ2*** **mRNA expression profile and ChIP on the** ***RARβ2*** **promoter**

Relative *RARβ2* mRNA expression monitored by qPCR overtime from t-RA-treated XPBwt, XPB/F99S, XPB/T119P, p8/TTD-A wt, p8/TTD-A, XPDwt, XPD/R112H, XPD/G602D, XPD/R683W and XPD/R722W cells (A1-J1). Red curves show the mRNA expression of RARβ2 in the relative XPD mutant cell lines; blue and dotted blue curves show mRNA expression of RARβ2 in wild type cells and transfected cells respectively as indicated in each panel. Error bars represent the standard deviation of three independent experiments. ChIP monitoring the t-RA-dependent recruitment of RAR, pol II, TFIIB (panels A2-J2), XPB, XPD, p44, Cdk7 subunits of TFIIH (panels A3-J3) and XPA, XPG, XPF, CTCF (panels A4-J4) on the *RARβ2* promoter; Each series of ChIP is representative of at least two independent experiments as indicated by standard deviation. Values are expressed as percentage of the input.

Figure 3

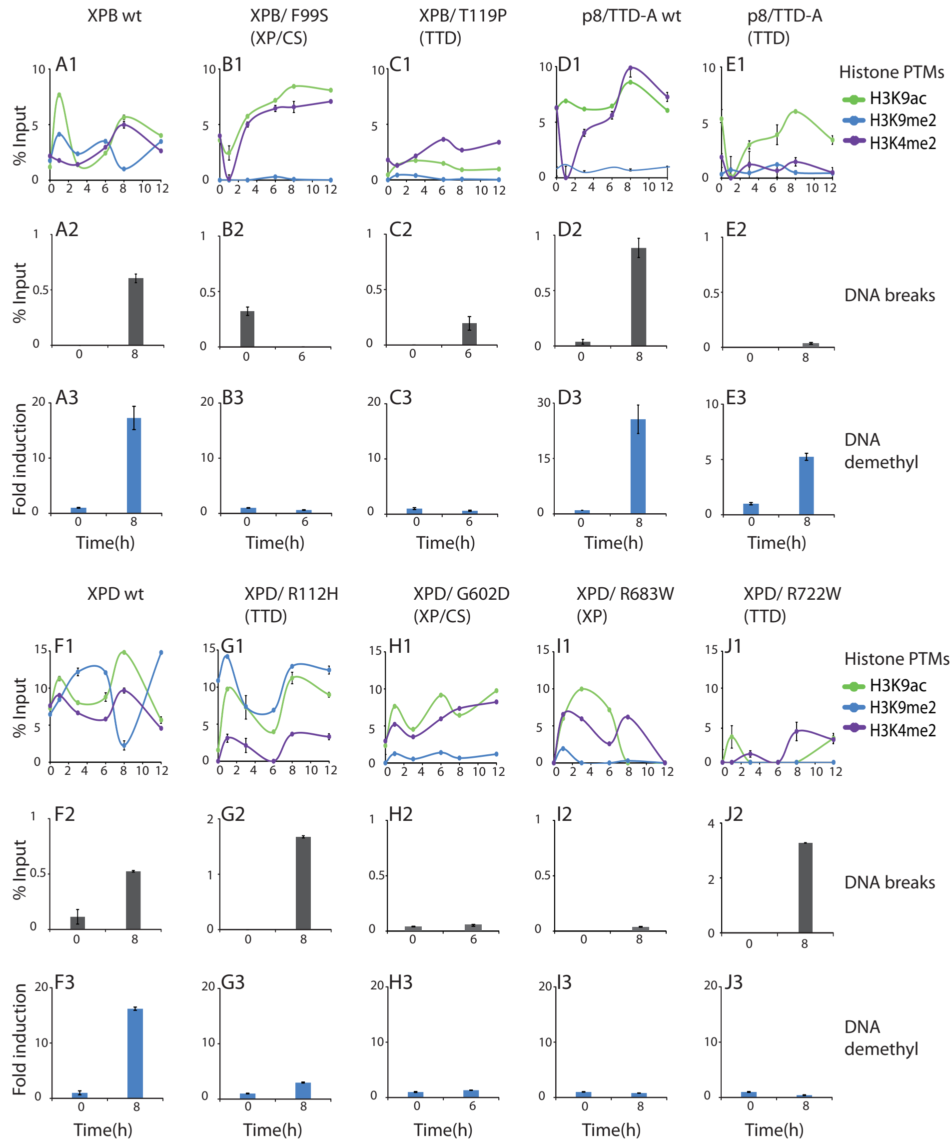


Figure 3. Histone PTMs, DNA breaks and active DNA demethylation

ChIP monitoring the t-RA dependent occupancy of H3K4me2, H3K9me2 and H3K9ac (panels A1-J1) on the *RARβ2* promoter overtime in all the cell lines as mentioned above and figure 2. Each series of ChIP is representative of at least two independent experiments. Detection of DNA breaks at *RARβ2* promoter at 0, 3 and 6/8 hour post t-RA treatment depending on the peak formation corresponding to RNA expression profile (A2-J2). DNA breaks are detected through the incorporation of Biotin-dUTP via a terminal deoxynucleotidyl-transferase (TdT) reaction. DNA fragments containing DNA breaks are immuno selected by Biotin antibodies (BioChIP) and analyzed by qPCR. Each series of BioChIP is representative of three independent experiments as indicated by standard deviation, and values are expressed as percentage of the input.

UnMedIP was performed using the UnMethyl-Collector kit (Active Motif) at the indicated time (panels A3-J3). Samples containing unmethylated DNA were analyzed by qPCR on the *RARβ2* promoter. Each series of UnMedIP is representative of two independent experiments as indicated by standard deviation, and values are expressed as percentage of the input.

Figure 4

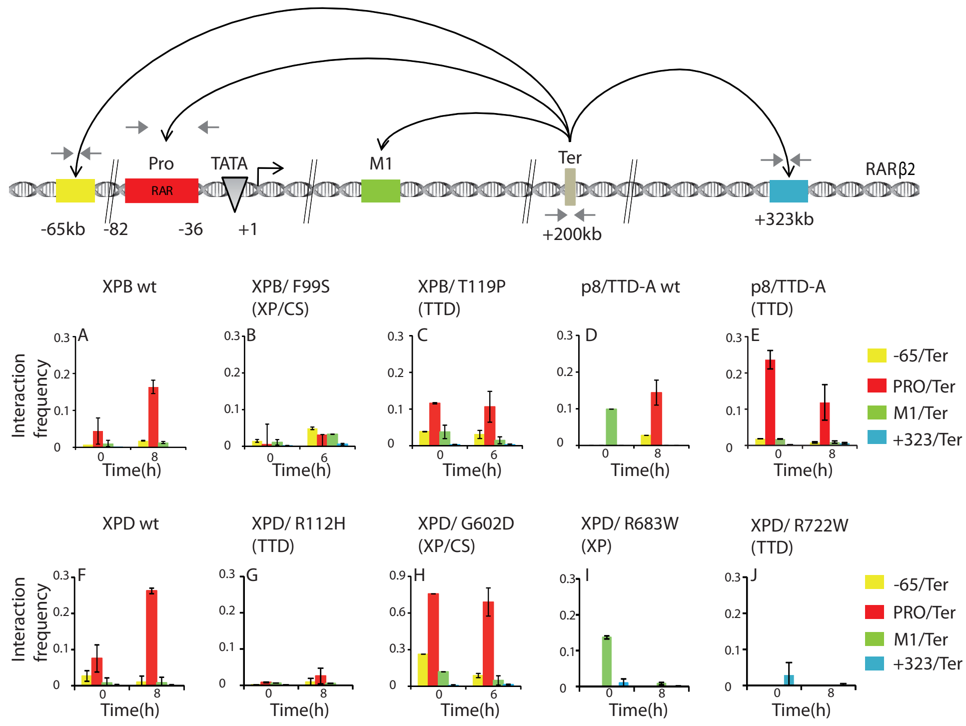


Figure.4. TFIIH is involved in *RARβ2* looping

Schematic representation of the quantitative chromatin conformation capture (q3C) (Upper panel). One probe was designed at Ter of *RARβ2* to investigate the associations between the different elements including upstream (-65 kb), Pro, intronic (M1), and downstream (+323 kb) regions as indicated by the black arrows. q3C assays were performed using crosslinked and HindIII-digested chromatin from all the cells as indicated above the A-J panels at 0, 3, and 6/8 hr post-t-RA treatment. The bar chart (y axis) shows the enrichment of PCR product (%) normalized to the enrichment within the human *XPB* (=100%). 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* (see the Experimental Procedures).

Figure 5

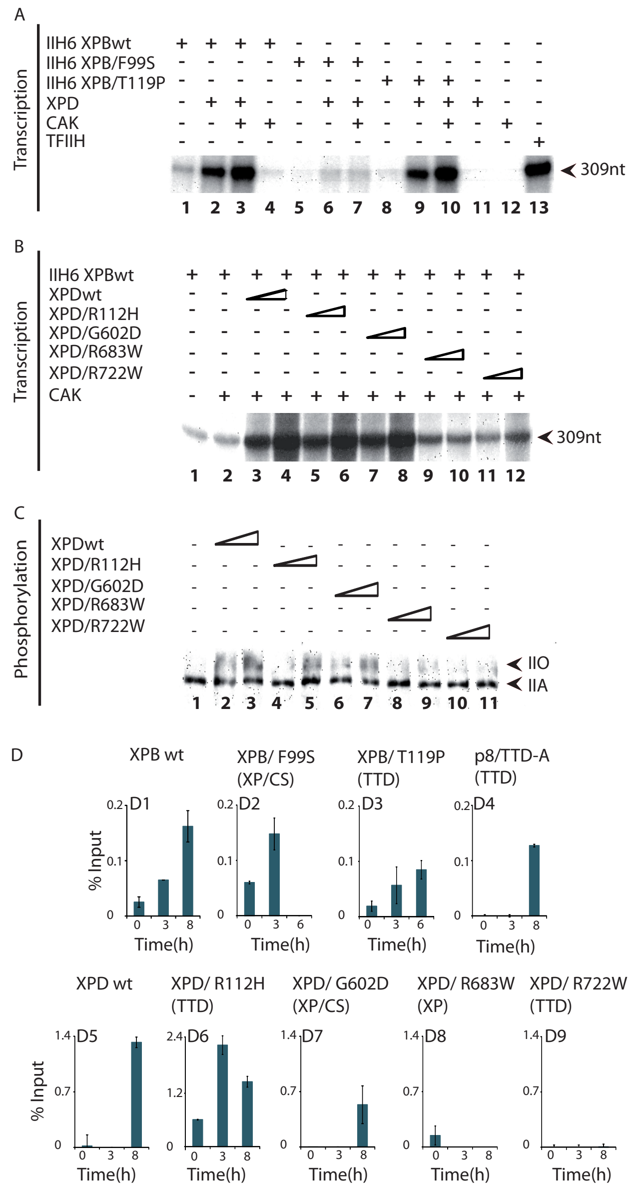


Figure 5. In vitro RNA synthesis and RNA pol II phosphorylation in TFIIH deficient cells

In vitro transcription activity of recombinant IIH6 XPBwt, XPB/F99S or XPB/T119P either alone or in combination with XPD and/or CAK (A) and of rIIH6 in addition to either XPD wt, XPD/R112H, XPD/G602D, XPD/R683W or XPD/R722W (B) in the presence or absence of CAK as indicated. Size (309 nt) of the transcript is indicated on the right side of each panel. (C) *In vitro* phosphorylation of pol II when added in the reconstituted transcription assay as described in panel B. Arrows indicate hypo (IIa) and hyper (IIo) phosphorylated forms of RNA pol II. (D) ChIP monitoring the t-RA dependent occupancy of the serine 5 phosphorylated pol II on the *RARβ2* promoter from wild type and mutated cells

Figure S1

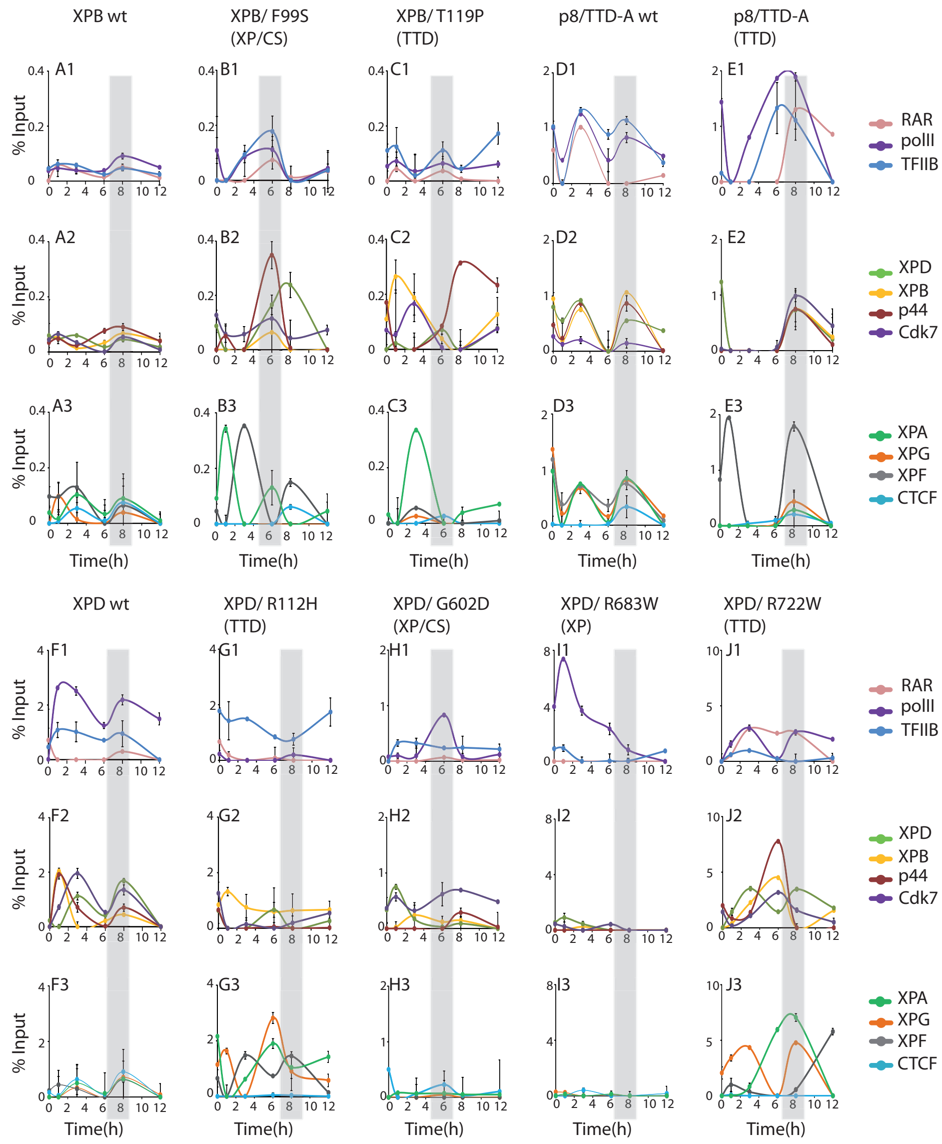


Figure S1. ChIP on the terminator region of RARβ2

ChIP monitoring the t-RA-dependent occupancy of RAR, pol II, TFIIB on the terminator region or RARβ2 during a 12 hour time course in XPBwt, XPB/F99S, XPB/T119P p8/TTD-Awt, p8/TTD-A, XPDwt, XPD/R112H, XPD/G602D, XPD/R683W and XPD/R722W cells (A1- J1) and and TFIIH (XPB, XPD, p44, CDK7) (A2- J2). Presence of NER factors XPA, XPG, XPF and chromatin organizer CTCF on *RARβ2* terminator region in XPBwt, XPB/F99S, XPB/T119P p8/TTD-Awt, p8/TTD-A, XPDwt, XPD/R112H, XPD/G602D, XPD/R683W and XPD/R722W cells (A3- J3). 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 S2

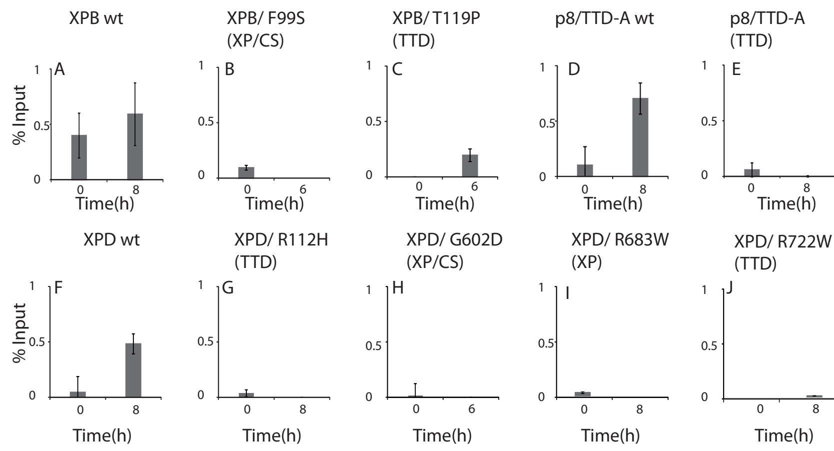
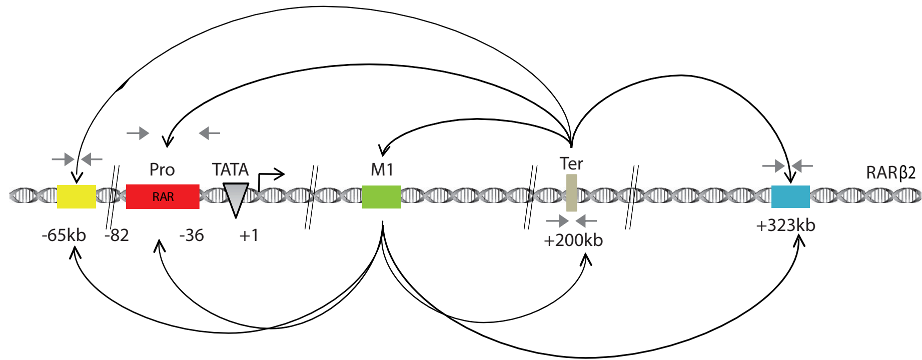


Figure S2: DNA Breaks at terminator region of RARβ2

BioChIP shows the presence of DNA breaks on the terminator region of *RARβ2* without t-RA and with t-RA treatment either at 6 hour or at 8 hour depending on the peak formation corresponding to RNA expression profile in XPBwt, XPB/F99S, XPB/T119P p8/TTD-Awt, p8/TTD-A (A to E) and in XPDwt, XPD/R112H, XPD/G602D, XPD/R683W and XPD/R722W cells (F to J).

Figure S3



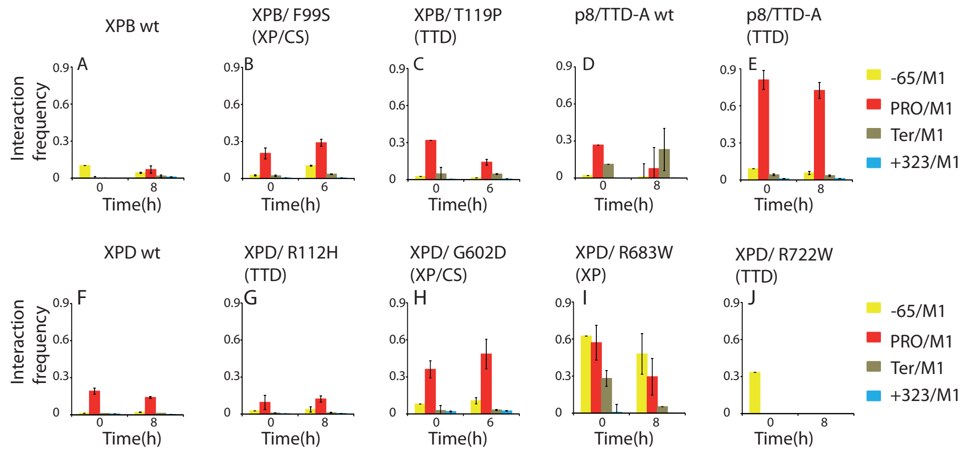


Figure S3: Chromatin looping between intronic region M1 and other regions of *RARβ2*

Chromatin looping was analyzed at upstream/downstream elements of *RARβ2* using intronic region M1 bait. q3C assays were performed using crosslinked, HindIII-digested chromatin without t-RA and with t-RA treatment either at 6 hour or at 8 hour in XPBwt, XPB/F99S, XPB/T119P p8/TTD-Awt, p8/TTD-A (A to E) and in XPDwt, XPD/R112H, XPD/G602D, XPD/R683W and XPD/R722W cells (F to J). These independent graphs represent a measure of the association between the upstream −65 kb; Pro; ter and the downstream +323 kb regions along the gene using M1 as bait (M1 Probe). 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*β*2* gene.

Table 1

Mutations	Syndrome	Clinical features	References
XPB/F99S	XP/CS	- Moderate skin photosensitivity, skin cancer - Moderate physical and mental retardation	Scott et al., (1993) Vermeulen et al.(1994)
XPB/T119P	TTD	- Moderate skin photosensitivity, skin cancer - Mild learning disability	Weeda et al., (1997)
XPD/R112H	TTD	- High skin photosensitivity, no skin cancer - Sulfur-deficient brittle hair and nail - Moderate physical and mental retardation	Stefanini et al., (1992)
XPD/G602D	XP/CS	- high skin photosensitivity, skin cancer - Progeroid features	Takayama et al., (1995)
XPD/R683W	XP	- High skin photosensitivity, skin cancer - Mental retardation	Tyalor et al., (1997)
XPD/R722W	TTD	- High skin photosensitivity, no skin cancer - Sulfur deficient brittle hair and nail - Severe physical and mental retardation	Stefanini et al., (1993) Broughton et al., (1994)
p8/L21P R56Stp	TTD	- Moderate skin photosensitivity, no skin cancer - Developmental delay, short stature - Mental retardation	Stefanini et al., (1993) Giglia-Mari et al., (2004)

Table1. Mutations in XPB, XPD and p8-TTDA subunits and the clinical phenotypes associated with the patients bearing the same mutation.

Table 2

Primers	Forward	Reverse
Primers for mRNA		
GAPDH	AGCTCACTGGCATGGCCTTC	ACGCCTGCTTCACCACCTTC
RARβ2	CCAGCAAGCCTCCATGTTC	TACACGCTCTGCACCTTTAGC
Primers for RARβ2 for ChIP/BioChIP/Unmethyl		
Promoter (Pro)	TGGTGATGTCAGACTAGTTGGGTC	GCTCACTTCCTACTACTTCTGTCAC
Terminator (Ter)	TGTTTGTGCTCTTTGGGCACT	CGGTCGGGCTAGGAAACAAGTAAA
3C primers		
-65	CCTGGCAATTGAAACATGAAAGT	
Pro	TCCAAAGATGCCTATTAAGTTGTAAGAG	
M1	AGCAGCAAAATGCAGGCTTTA	TGACACCAGTGAAAAGGAAGCA
Ter	AAGATGCAGTTTGAGAGCATC	CTGGGCAACATGAAATAAAAGATG
323	CCAAACAATTTTCTTCATGGTCATT	
RARβ2 promoter	CAGACTAGTTGGGTCATTTGAAGGT	TTGAATTGCCTAATATATGCGAGTGA
XPB	CGGTGAGGTGAGTTTGTGGAAT	AGGATCTCTGTTTAATGGAAAAGCTT
3C Probes		
Ter probe	6[FAM]TTGCTCTTTCTGATGCTCTCAAA[TAM]	
M1 probe	6[FAM]CAGTACAGTCAAGGTGGCCCGTCT[TAM]	
RARβ2 promoter probe	6[FAM]AGCCCGGGTAGGGTTCACCGAAAG[TAM]	
XPB probe	6[FAM]AAGGATGAAGGCGTGATCCGACTCTG[TAM]	

Table 2. List of Primers

References

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

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.

Bensaude, O., Bonnet, F., Casse, C., Dubois, M.F., Nguyen, V.T., and Palancade, B. (1999). Regulated phosphorylation of the RNA polymerase II C-terminal domain (CTD). Biochem Cell Biol *77*, 249-255.

Botta, E., Nardo, T., Lehmann, A.R., Egly, J.M., Pedrini, A.M., and Stefanini, M. (2002). Reduced level of the repair/transcription factor TFIIH in trichothiodystrophy. Hum Mol Genet *11*, 2919-2928.

Broughton, B.C., Steingrimsdottir, H., Weber, C.A., and Lehman, A.R. (1994). Mutations in the xeroderma pigmentosum group D DNA repair/transcription gene in patients with trichothiodystrophy. Nature gnetics *7*, 189-194.

Cleaver, J.E., Lam, E.T., and Revet, I. (2009). Disorders of nucleotide excision repair: the genetic and molecular basis of heterogeneity. Nat Rev Genet *10*, 756-768.

Coin, F., Auriol, J., Tapias, A., Clivio, P., Vermeulen, W., and Egly, J.M. (2004). Phosphorylation of XPB helicase regulates TFIIH nucleotide excision repair activity. Embo J *23*, 4835-4846.

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., Oksenych, 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., Oksenych, 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. (2012). TFIIH: when transcription met DNA repair. Nat Rev Mol Cell Biol *13*, 343-354.

D'Alessio, A.C., Weaver, I.C.G., and Szyf, M. (2007). Acetylation-Induced Transcription Is Required for Active DNA Demethylation in Methylation-Silenced Genes. Molecular and Cellular Biology *27*, 7462-7474.

de Boer, J., and Hoeijmakers, J.H. (2000). Nucleotide excision repair and human syndromes. Carcinogenesis *21*, 453-460.

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

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.

Evans, E., Moggs, J.G., Hwang, J.R., Egly, J.M., and Wood, R.D. (1997). Mechanism of open complex and dual incision formation by human nucleotide excision repair factors. Embo J *16*, 6559-6573.

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.

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.

Giglia-Mari, G., Miquel, C., Theil, A.F., Mari, P.O., Hoogstraten, D., Ng, J.M., Dinant, C., Hoeijmakers, J.H., and Vermeulen, W. (2006). Dynamic Interaction of TTDA with TFIIH Is Stabilized by Nucleotide Excision Repair in Living Cells. PLoS Biol *4*, e156.
Itin, P.H., Sarasin, A., and Pittelkow, M.R. (2001). Trichothiodystrophy: update on the sulfur-deficient brittle hair syndromes. J Am Acad Dermatol *44*, 891-920; quiz 921-894.
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.
Ju, B.-G., Lunyak, V.V., Perissi, V., Garcia-Bassets, I., Rose, D.W., Glass, C.K., and Rosenfeld, M.G. (2006). A Topoisomerase IIß-Mediated dsDNA Break Required for Regulated Transcription. Science *312*, 1798-1802.
Kamiuchi, S., Saijo, M., Citterio, E., de Jager, M., Hoeijmakers, J.H., and Tanaka, K. (2002). Translocation of Cockayne syndrome group A protein to the nuclear matrix: possible relevance to transcription-coupled DNA repair. Proc Natl Acad Sci U S A *99*, 201-206.
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.
Kraemer, K.H., Lee, M.M., and Scotto, J. (1987). Xeroderma pigmentosum. Cutaneous, ocular, and neurologic abnormalities in 830 published cases. Arch Dermatol *123*, 241-250.
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.
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. (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.
Lehmann, A.R., McGibbon, D., and Stefanini, M. (2011). Xeroderma pigmentosum. Orphanet J Rare Dis *6*, 70.
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.
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.
Riou, L., Zeng, L., Chevallier-Lagente, O., Stary, A., Nikaido, O., Taieb, A., Weeda, G., Mezzina, M., and Sarasin, A. (1999). The relative expression of mutated XPB genes results in xeroderma pigmentosum/Cockayne's syndrome or trichothiodystrophy cellular phenotypes. Hum Mol Genet *8*, 1125-1133.
Sarasin A, B.-B.C., Renault G, Lehmann A, Arlett C, Dumez Y. (1992). Prenatal diagnosis in a subset of trichothiodystrophy patients defective in DNA repair. Br J Dermatol *127*, 485-491.
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.
Schärer, O.D. (2008). The molecular basis for different disease states caused by mutations in TFIIH and XPG. DNA Repair (Amst) *7*, 339-344.
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.

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 β-globin locus. Genes & Development *20*, 2349-2354.
Stefanini, M., Giliani, S., Nardo, T., Marinoni, S., Nazzaro, V., Rizzo, R., and Trevisan, G. (1992). DNA repair investigations in nine Italian patients affected by trichothiodystrophy. Mutat Res *273*, 119-125.
Stefanini, M., Lagomarsini, P., Giliani, S., Nardo, T., Botta, E., Peserico, A., Kleijer, W.J., Lehmann, A.R., and Sarasin, A. (1993a). Genetic heterogeneity of the excision repair defect associated with trichothiodystrophy. Carcinogenesis *14*, 1101-1105.
Stefanini, M., Vermeulen, W., Weeda, G., Giliani, S., Nardo, T., Mezzina, M., Sarasin, A., Harper, J.I., Arlett, C.F., Hoeijmakers, J.H.J.*, et al.* (1993b). A new nucleotide-excision repair gene associated with the disorder trichothiodystrophy. Am J Hum Genet *53*, 817-821.
Takayama, K., Salazar, E.P., Lehmann, A., Stefanini, M., Thompson, L.H., and Weber, C.A. (1995). Defects in the DNA repair and transcription gene ERCC2 in the cancer-prone disorder xeroderma pigmentosum group D. Cancer Res *55*, 5656-5663.
Taylor, E., Broughton, B., Botta, E., Stefanini, M., Sarasin, A., Jaspers, N., Fawcett, H., Harcourt, S., Arlett, C., and Lehmann, A. (1997). Xeroderma pigmentosum and trichothiodystrophy are associated with different mutations in the XPD (ERCC2) repair/transcription gene. Proc Natl Acad Sci USA *94*, 8658-8663.
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. The EMBO journal *26*, 2041-2051.
Winkler, G.S., Vermeulen, W., Coin, F., Egly, J.M., Hoeijmakers, J.H., and Weeda, G. (1998). Affinity purification of human DNA repair/transcription factor TFIIH using epitope-tagged xeroderma pigmentosum B protein. J Biol Chem *273*, 1092-1098.
Wossidlo, M.A.J.S.V.L.K.B.M.R.R.S.H.W.J. (2010). Dynamic link of DNA demethylation, DNA strand breaks and repair in mouse zygotes. The EMBO journal *29*, 1877-1888.
Zurita, M., and Merino, C. (2003). The transcriptional complexity of the TFIIH complex. Trends in genetics : TIG *19*, 578-584.

B. Enzymatic activity of TFIIH

Thus far role of TFIIH in transcription has been shown in several studies. The previous section of the results added some more information to our knowledge, showing that mutation in one or the other subunit of TFIIH influenced different events in chromatin remodeling during transcription. To provide an explanation of some clinical features observed within the XP, XP/CS and TTD patients due to mutation in XPB, XPD and p8/TTD-A subunits and to further define their role in transcription, I have examined two defined enzymatic activities (ATPase and helicase) and have characterized these pathogenic mutations with respect to their deficiencies in each. The Cdk7 kinase activity was already shown in the previous section. Since, p8/TTD-A subunit has not been known to have any enzymatic activity, only TFIIH composition analysis in p8/TTD-Awt and mutant cell line was performed.

B.1 Subunit composition of TFIIH

To provide further insights into the role of XPB and XPD, whole cell extract (WCE) was prepared from the XP-B and XP-D patients carrying either XPB or XPD mutation (as shown in Table 7 and in the manuscript of the paper). Western blot and Co-immunoprecipiation (Co-IP) analysis reveals a similar concentration and subunit composition of the immunopurified TFIIH in XPBwt, XPB/F99S, and XPB/T119P complexes (Figure 14, western; lane 2-4, Co-IP; lane 5-7). Similarly, western and Co-IP was performed using WCE from XPDwt, XPD/R112H, XPD/G602D, XPD/R683W, XPD/R722W, p8/TTD-Awt and p8/L21P-R56Stp (Figure 15, western; lane 1-7, Co-IP; lane 9-15). TFIIH concentration was observed to be significantly low in XPD/R683W and XPD/R722W, but in XPD/R112H and XPD/G602D it was comparable to XPDwt. Moreover there was a remarkable disturbance in the integrity of TFIIH in XPD/R683W and XPD/R722W (Figure 15, lane 12-13). These mutations are located in C-terminal domain of XPD and lead to the dissociation of the CAK with Core TFIIH. However

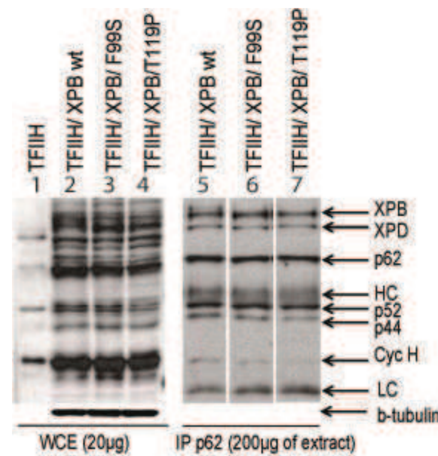


Figure 14. TFIIH subunits in XPBwt and XPB mutant cell lines.

20 µg of WCE from Hela cells, XPBwt, XPB/F99S, and XPB/T119P were used to analyze the TFIIH subunits (lane 1-4 respectively). Co-IP was performed using the 200 µg of WCE in XPBwt, XP/F99S, and XPB/T119P (lane 5-7 respectively).

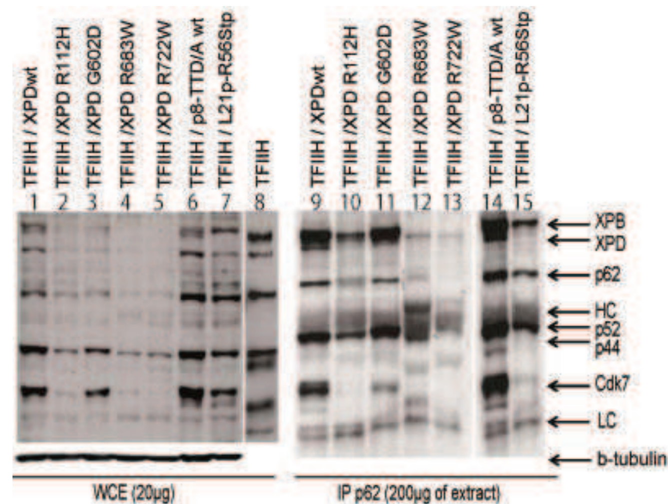


Figure 15. TFIIH subunits in XPDwt, XPD mutant, p8/TTD-Awt and p8/TTD-A mutant cell line.

20 µg of WCE from XPDwt, XPD/R112H, XPD/G602D, XPD/R683W, and XPD/R722W, p8/TTD-Awt and p8/L21P-R56Stp were used to analyze the TFIIH subunits (lane 1-7 respectively). Co-IP was performed using the 200 µg of WCE in XPDwt, XPD/R112H, XPD/G602D, XPD/R683W, and XPD/R722W, p8/TTD-Awt and p8/L21P-R56Stp (lane 9-15 respectively). WCE from hela cells shows the TFIIH control (lane 8)

p8/TTD-A mutation showed very insignificant reduction in the concentration of TFIIH (Figure 15, lane 6-7), but the integrity of TFIIH in p8/L21P-R56Stp seems to be disturbed when compared to the p8/TTD-Awt (Figure 15, lane 14-15). It is important to note that TFIIH from p8/TTD-Awt and mutant fibroblast cell lines showed similar recruitment pattern of all the transcription machinery including TFIIH subunits (Results shown in manuscript of the paper).

B.2 ATPase and helicase activity of TFIIH

To observe the importance of enzymatic activity in transcriptional dysregulation in all the XPB and XPD mutant cells, ATPase and helicase assay was performed. Purified extracts from baculovirus-infected insect cells overexpressing either wild type XPB or mutant XPB, p52 and p8/TTD-A (p8 in figure) were used in an in-vitro ATPase assay to measure the ATPase activity of XPB.

ATPase activity was estimated as a function of the hydrolyzed phosphates (Pi). XPBwt clearly shows a very high ATPase activity (Figure 16A lane 2-3) and addition of p52 and subsequently p8/TTD-A further enhanced the signal for ATPase activity (Figure 16.A lane 6-7 and lane 12-13), also shown by some previous studies (Jawhari, Laine et al. 2002; Coin, Oksenych et al. 2007). When focusing on XPB/F99S and XPB/T119P, a significant decrease in the ATPase activity was observed (Figure 16B). Another mutation XPB/K346R, in ATPase A Walker motif I, which was known to completely disrupt the ATPase activity of XPB was used as a control. The graph represents the densitometry analysis in terms of the percentage of phosphate released (Pi/ [ATP+Pi]) from two independent experiments.

Previous studies have shown that XPD is dispensable for transcription but required for NER (Coin, Oksenych et al. 2007). The removal of lesions in NER depends on the opening of the DNA around the damaged site. Mutations in XPD proteins can disable

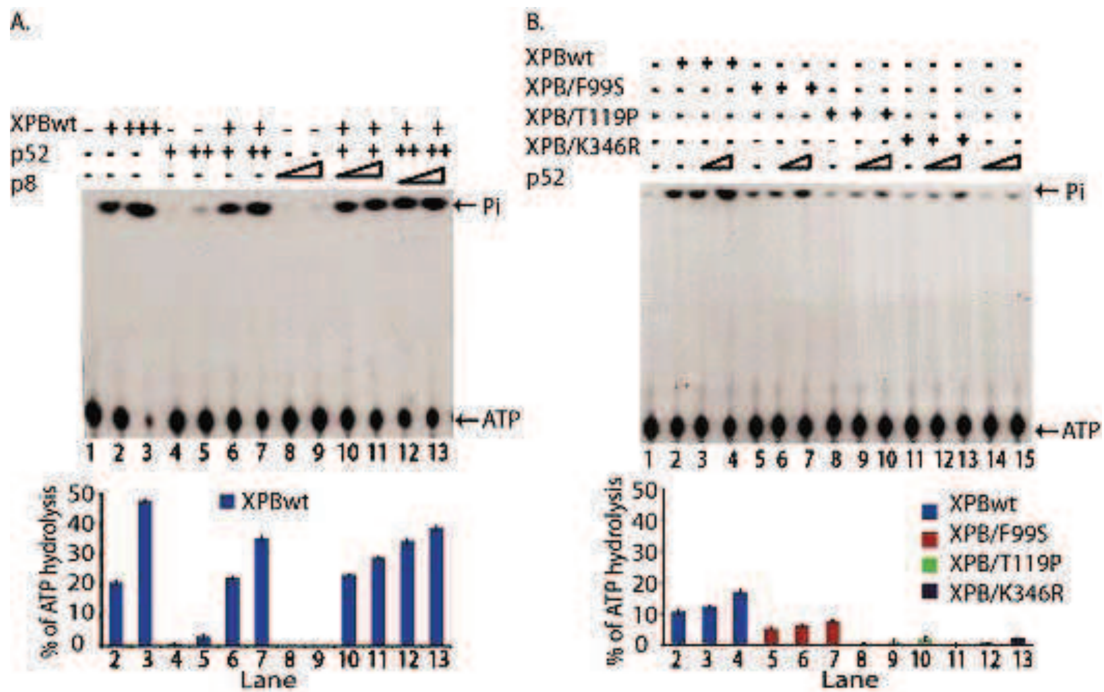


Figure 16. ATPase activity of XPB

A. *Purified XPBwt in increasing amount was tested in an ATPase assay either alone (lanes 2–3) or in the presence of 50 and 100 ng of purified p52 (lanes 6 and 7), 10 and 20 ng of purified p8/TTDA (lanes 10-13) subunits of TFIIH.*
B. *Similarly, purified XPBwt, XPB/F99S, XPB/T119P and XPB/K346R were tested either alone (lane2,5,8 and 11 respectively) or in presence of increasing p52 (lanes 3-4, 6-7, 9-10, 12-13). The graph represents the percentage of phosphate released (Pi/[ATP+Pi]) from two independent experiments.*

DNA opening by impeding both its ATPase and Helicase activity in NER. Curiously, *in-vivo* studies from patients cell lines as shown in the manuscript shows strong transcription dysregulation arguing further investigation about the involvement of XPD during transcription as well. It also raises the question if ATPase activity is required as a provider of energy for its helicase activity during promoter opening or it has distinct function. Purified extracts from baculovirus-infected insect cells overexpressing XPDwt, XPD/R112H, XPD/G602D, XPD/R683W and XPD/R722W were used for the estimation of the ATPase activity. ATPase activities of XPD/G602D and XPD/R683W mutants (Figure 17, lane 8-13) were reduced while other two; XPD/R112H and XPD/R722W are comparable to the XPDwt (Compare Figure 17 lane 5-7 and 14-16 to lane 2-4). Despite of the very strong dysregulation in the transcription and chromatin remodeling events in all the XPD mutant patient cell lines, the ATPase activity is not impeded that strongly.

Next, I analyzed the helicase activity of XPD in all the XPD mutatnts. Purified extracts obtained from the baculoviruses infected insect cells overexpressing XPDwt, XPD/R112H, XPD/G602D, XPD/R683W and XPD/R722W were used in increasing amount. p44 is the regulatory subunit of TFIIH, it interacts with XPD and hence enhance its helicase activity, also shown previously (Dubaele, Proietti De Santis et al. 2003; Coin, Oksenych et al. 2007). XPDwt alone without p44 shows the minimal signal which is nearly equivalent to the non-denatured probe (Compare Figure 18 Lane 12 to lane 13). The helicase activity increases significantly when p44 interacts with XPDwt (Figure 18 Lane 3) but there was no helicase activity observed in XPD/R112H, XPD/R683W and XPD/R722W except XPD/G602D which shows slightly reduced but noticeable signal for the displaced helicase probe.

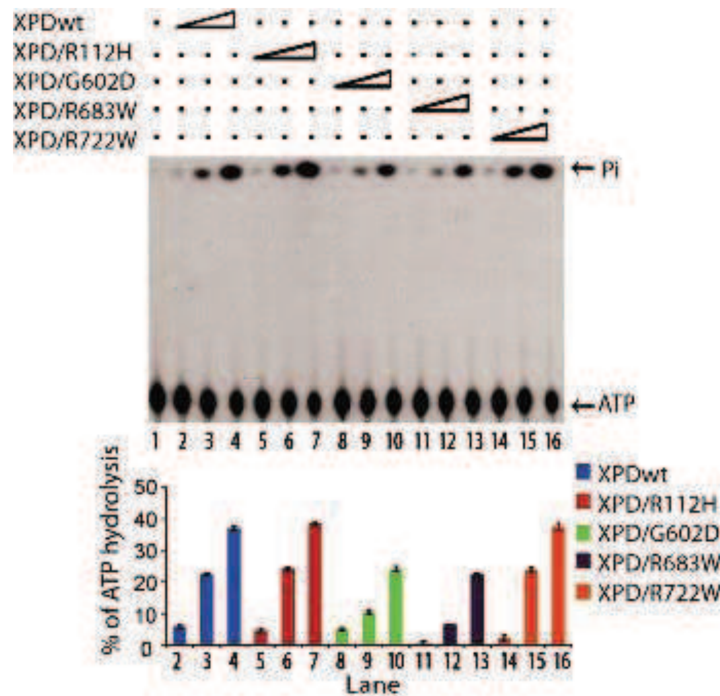


Figure 17. ATPase activity of XPD

Increasing amount of purified XPDwt, XPD/R112H, XPD/G602D, XPD/R683W and XPD/R722W extract tested in an ATPase assay. The graph represents the percentage of phosphate released (Pi/[ATP+Pi]) from two independent experiments.

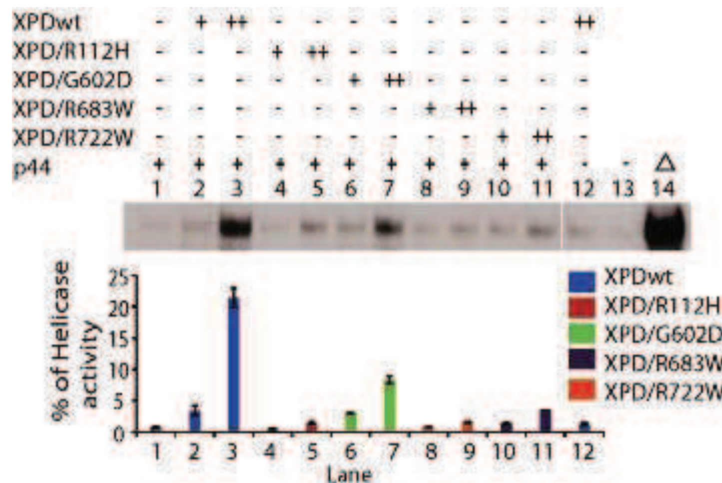


Figure 18. Helicase activity of XPD

Increasing amount of purified XPDwt, XPD/R112H, XPD/G602D, XPD/R683W and XPD/R722W extract tested in Helicase assay. The graph represents the percentage of displaced oligonucleotide by the XPD from two independent experiments.

B.3 Experimental procedures:

This part of the experimental procedure covers only the details of the experiments done in this section such as western, Co-IP, ATPase assay and Helicase assay.

Western and Co-IP:

Whole cell extract was prepared using RIPA buffer (25 mM Tris-HCl, pH 7.6, 150 mM NaCl, 1% Nonidet P-40, 1% sodium deoxycholate, 0.1% SDS, and protease inhibitor mixture). Immunoprecipitation was carried out with 200 μg of WCE, overnight using antibody against p62 (Santa Cruz Biotechnology) and dynabeads Protein-G (Invitrogen). After several washes with PBST (1X PBS+ 0.05 % tween), the protein was eluted and loaded on 10% SDS/PAGE, transferred to nitrocellulose membrane and probed with XPB (1B3), XPD (2F6), p52 (1D11), p44 (1H5), Cyc H (2D4), which were from IGBMC antibody facility and Cdk7 (C-19) (Santa Cruz Biotechnology). Tubulin was used as a loading control.

ATPase assay:

Protein fractions were incubated for 2 hr at 30°C in the presence of 1 μCi [γ-32P] ATP (7000 Ci/mmol, Hartmann Analytic) in a 20 μl reaction volume in 20 mM Tris-HCl (pH 7.9), 4 mM MgCl2, 1 mM DTT, 50 mg/ml BSA, with 120 ng of supercoiled double strand DNA (pSK). Reactions were stopped by adding EDTA to 50mM and SDS to 1% (w/w). The reactions were then diluted 5-fold, spotted onto polyethylenimine (PEI) TLC plates (Merck), run in 0.5 M LiCl/1 M formic acid, and autoradiographed.

Helicase assay:

Helicase probe is constructed from the single-stranded M13 phage which is a hybridized oligonucleotide complementary to the region 6218-6251. Once labeled the ends, the probe is placed on a column of Sepharose CL-4B gel filtration (Pharmaciaã). The opening of DNA is demonstrated by the displacement of the oligonucleotides

separated on non-denaturing gel. The fractions containing the purified factor TFIIH is incubated 45 minutes at 37°C in the presence of helicase probe (30000-40000 cpm ; 1 - 3 ng DNA) in 20 mM Tris - HCl, pH 7.9, 1 mM DTT, 4 mM MgCl 2 and 4 mM ATP , containing 100 ug/ml BSA, to a final volume of 25 µl . The reaction is stopped by adding 10 of 60 µl mM EDTA, 50% glycerol, 0.75% SDS and 0.1 % bromophenol blue. Samples are run on a 10% polyacrylamide gel. The gel was fixed for 10 minutes in 10% acetic acid, 10 % methanol and autoradiography 6 to 12 hours at -80 ° C.

Discussion

Discussion

A. TFIIH, NER factors and chromatin remodeling

TFIIH came into existence in 1989, as one of the GTFs implicated in transcription by RNA pol II (Conaway and Conaway 1989). The understanding of the TFIIH became more captivating and intricate when, several studies confirmed its role in the maintaining genome stability by its participation in DNA nucleotide excision repair (Schaeffer, Roy et al. 1993; Schaeffer, Moncollin et al. 1994). In transcription, TFIIH functions as the basal factor working at initiation, promoter scape and early elongation steps and also in transcription re-initiation after RNA pol II pausing. XPB and Cdk7 are the two enzymatic subunits which modulates these transcriptional events. Contrary to this, XPB and XPD are the two key helicases subunits which participate in NER pathways. A recent development which demonstrate that NER factors localize to the promoter of the activated gene without any DNA damage and further contribute in chromatin remodeling events have strengthen the idea of the interplay between transcription and NER (Schmitz, Schmitt et al. 2009; Le May, Mota-Fernandes et al. 2010; Le May, Fradin et al. 2012).

A.1 Biochemistry of TFIIH and the overlapping genetics of diseases

Being right at the epicenter of the transcription and DNA repair, the detailed understanding of how TFIIH functions to co-ordinate these two processes could further provide an explanation for the phenotypes observed in the patients bearing mutation in XPB, XPD and p8/TTD-A subunits of TFIIH. The overlapping clinical features of these patients (Figure 13) further hammer the notion of interrelationship of TFIIH with NER factors in transcription and NER.

The present study clearly underlines the functional relationship between TFIIH and NER factor in transcription. Here we demonstrated that presence of TFIIH is an essential criterion for the synchronized recruitment of NER factors. Any mutation in XPB and XPD subunits prevents the proper association of all the TFIIH components at the promoter which could further thwart the transcriptional network downstream. For instance XPB/F99S mutation weakens its interaction with p52, which reduces XPB ATPase activity required during transcription (Result B.2, Figure 16). However, XPB/T119P which shows *in-vitro* transcription comparable to XPBwt, also shows much reduced RNA synthesis of *RARβ2* (Manuscript Figure 2 and Figure 5). Similarly, XPD/R683W and XPD/R722W mutations impair the ATPase and helicase activity of XPD (Result B.2, Figure 17, 18). Although, till now XPD has been shown to be dispensable for transcription, knowing the complicated network of interaction of TFIIH with several proteins, it cannot be denied for indirectly regulating the transcription. Moreover, these two XPD mutations are known to disrupt the XPD-p44 interaction and consequently cause the dissociation of CAK from the core. As a result of this dissociation the phosphorylation of CTD of RNA pol II by CAK is hampered leading to impairment in RNA synthesis *in-vitro* (Manuscript Figure 5 C). In addition to phosphorylating RNA pol II, CAK also phosphorylates the nuclear receptor (NR) and mutation in C-terminal domain of XPD disturbs thus disrupt the phosphorylation resulting in the transcription dysregulation of the NR responsive genes (Rochette-Egly, Adam et al. 1997; Compe, Drane et al. 2005). Furthermore, ChIP results showed that the phosphorylation status of Ser5 of CTD of RNA pol II is weakened in these two XPD mutations compared to the XPDwt cells (Manuscript Figure 5 C and D). This shows a correlation with a much reduced RNA expression of *RARβ2* in these XPD mutant cells compared to their respective XPDwt overexpressed controls (Manuscript Figure 2). Conversely, our understanding for some mutations like XPB/T119P, XPD/R112H and XPD/G602D needed more investigation as the defect in enzymatic activity solely cannot

explain the contradiction about the perfect in-vitro transcription and much reduced gene expression in cells (Manuscript Figure 2, Figure 5).

A.2 TFIIH deploys the recruitment of NER factors and chromatin remodeling events

All the cellular process is cascade of events and disturbance in any of the events might disturb the pathway locally and on whole genome. The sequential assembly of NER factors in NER and in transcription has been illustrated by Mocquet et al., and Le May et al, respectively (Mocquet, Laine et al. 2008; Le May, Mota-Fernandes et al. 2010). Such sequential assembly of proteins could be required for the protein-protein interaction that mediates the progression of the NER pathways (Table 6). The functional link of TFIIH and NER factors in DNA repair has been quested over several years and is now well accepted but our understanding of their relations remains far from complete in transcription.

The ChIP data analysis revealed that TFIIH, in fact is acting like a platform for the recruitment of the NER factors during the transactivation of RARβ2, as mutation in XPB and XPD subunits of TFIIH significantly disturbs the recruitment of the NER factors at the promoter. Nevertheless, mutation in p8/TTD-A does not impair the recruitment of NER factors which coordinates with the previous studies showing p8/TTD-A is an essential component for NER but not for transcription (Coin, De Santis et al. 2006). The presence of the NER factors at the promoter is perhaps required for the establishment of the appropriate chromatin state around the promoter for accurate and optimal gene expression (Barreto, Schafer et al. 2007; Le May, Mota-Fernandes et al. 2010). The investigation of the chromatin status around the promoter revealed a signature mark with the enrichment of the di-methylation of H3K4, acetylation of H3K9 and decrease of di-methylation of H3K9, which creates open or active chromatin milieu around the promoter for the transcription (Manuscript Figure 3). Silencing of XPC, XPA and

ERCC1 and mutation in XPC, XPA, ERCC1, XPG and XPF showed the disturbance in the pattern of enrichment of these histone modifications(Le May, Mota-Fernandes et al. 2010). Besides playing a role in establishing histone PTMs, it has become increasingly apparent that the recruitment of the NER factors on the active promoters of RNA pol I and pol II genes can induce DNA demethylation around the promoter. It was shown that knockdown of Gadd45a, XPA, XPG, XPF, or TAF12 or treatment with drugs that inhibit NER causes hypermethylation of rDNA, establishes heterochromatic histone marks, and impairs transcription (Schmitz, Schmitt et al. 2009).

In fact, NER factors such as endonucleases XPG, XPF and TFIIH have been proposed several year before to be involved in DNA demethylation through NER pathway (Chu and Mayne 1996). Since NER pathway is preferentially involved in the repair of bulky lesions and a methyl group is not at all bulky, it is hard to conceive that the methyl cytosine is recognized as such and that the NER factors functions in transcriptional regulation irrespective of DNA demethylation. Supporting the same concept, our UnMeDIP data shows that TFIIH is possibly engaged in DNA demethylation. Mutation in XPB and XPD subunits clearly disrupts the DNA demethylation leaving the promoter in hypermethylated state, while p8/TTD-A mutation does not affect the process. A failure to demethylate the DNA and activate the gene could bring transcriptional dysregulation per se in the cell. The DNA demethylation following gene activation is normally observed in the close proximity to the regions where DNA breaks occurs. The DNA breaks are formed by the endonuclease XPG at the promoter and XPF at promoter as observed in XPBwt, XPDwt and p8/TTD-Awt (Manuscript, Figure 2). The formation of DNA breaks was not consistent with the presence of XPG and XPF at promoter and terminator respectively, which was logical in all of the mutant cell line. Enigmatically, despite of the presence of all the GTFs, TFIIH subunits and NER factors, DNA demethylation, there was no DNA break in p8/TTD-A cells. The significance of

the DNA break formation is still elusive; possibly it could be required for the nucleosome rearrangement and chromatin relaxation, required for the formation of chromatin looping. Also, it is known that TFIIH interacts physically with XPG during transcription as well as in NER during the formation of pre-incision complex. In fact, XPG was also considered to be the eleventh subunit of TFIIH (Ito, Kuraoka et al. 2007; Schärer 2008). Mutation in TFIIH might disturb the structural association of TFIIH and XPG which could possibly alter the endonuclease activity of XPG, thus hindering the formation of DNA breaks. It could also be possible that XPG function as a regulator for TFIIH, stabilizing the association between core-TFIIH, CAD and XPD.

Control of gene expression by formation of chromatin loops between multiple regulatory elements located over large genomic distances has been documented in several studies. Such physical interactions among regulatory elements are aided by chromatin organizers such as cohesion and CTCF. Since our ChIP data shows the presence of all the GTFs, TFIIH and NER factors specifically at two region; promoter and terminator (Manuscript Figure 2, Supplementary Figure 1), we analyzed the chromatin loop formation between promoter and terminator. We observed three conditions; formation of chromatin looping (as in XPBwt, XPDwt and p8/TTD-Awt), no chromatin looping (as in XPB/F99S, XPD/R112H, XPD/R683W and XPD/R722W) and a persistent chromatin looping (as in XPB/T119P, XPD/G602D and p8/TTD-A). Absence and presence of chromatin looping is very much understandable but the presence of consistent loop is a matter of further investigation. Presence of persistent chromatin looping might also serve to inhibit gene expression.

Altogether, this study has underscored the functions of TFIIH deeper, which suggest that each mutation in TFIIH is specific. Each of the mutations studied represents transcriptional dysregulation whether it is XPB and XPD mutants where it is affecting right at the level of recruitment of NER and then further downstream or p8/TTD-A

Mutations	XPB/F99S	XPB/T119P	XPD/R112H	XPD/G602D	XPD/R683W	XPD/R722W	p8/L21P R56Stp
TFIIH composition and interaction (from respective cell line)	Conserved in cells[f] Weak interaction of XPB with p52 by in-vitro binding assay [b]	Conserved in cells [f] and in in- vitro binding assay[b,]	Lower TFIIH con- centration[e, f]	Lower TFIIH con- centration[f]	Lower TFIIH concentration and disturbs Interaction of XPD with p44[b, f]	LowerTFIIH concentration and disturbs Interaction of XPD with p44[b, f]	Lower TFIIH Con- centration[e, f]
Helicase (in-vitro)	Not disturbed[b]	Not disturbed[b]	Completely disrupts helicase activity[f]	Much reduced helicase activity[f]	Completely disrupts helicase activity[f]	Completely disrupts helicase activity[f]	NA
ATPase (in-vitro)	Reduced ATPase activity[f]	Reduced ATPase activity[f]	Not disturbed[f]	Reduced ATPase activity[f]	Reduced ATPase activity[f]	Not disturbed[f]	NA
Phos- phorylation	Disturbed[f]	Not disturbed[f]	Not disturbed[f]	Not disturbed[f]	Hypo- phosphorylati on of NR[d] Not disturbed[f]	Disturbed[f]	Not disturbed[f]
RNA synthesis (in-vitro)	Reduced RNA synthesis[f]	Not disturbed[f]	Not disturbed[f]	Not disturbed[f]	Reduced RNA synthesis[f]	Reduced RNA synthesis[f]	Not distubed[a]
Promoter opening (in-vitro)	Disturbed[b]	Not disturbed[b]	NA	NA	Impede Promoter opening[b]	Impede Promoter opening[b]	NA
DNA repair (in-vitro)	CPD repair deficiency[b]	Not disturbed[b]	Complete NER deficiency[c, f]	Complete NER deficiency[c, f]	Low level of NER activity[c, f]	Low level of NER activity[c, f]	Low level of NER activity[a]

Table 9. Functions of TFIIH

NR- Nuclear receptor, NER- Nucleotide excision repair, NA- Not available

a-(Coin, De Santis et al. 2006), b-(Coin, Oksenych et al. 2007), c- (Dubaele, Proietti De Santis et al. 2003), d- (Keriel, Stary et al. 2002), e-(Botta, Nardo et al. 2002), f-(analyzed in present study)

mutant where it is disturbing more at the level of chromatin remodeling. These subtle defects at transcriptional level restrict us from giving a coherent explanation for such complex molecular manifestation of these syndromes. However, at the same time suggest that the varying range of clinical features in the XP, XP/CS and TTD patients could arise from these subtle defects in transcription. Table 9, represent the conclusion of the analysis obtained in the present study as well as from previous related studies.

B. Future direction for TFIIH and NER factors

Transcriptional regulation can occur at two interconnected levels: the first involves transcription factors and the transcription apparatus, and the second involves chromatin structure and its regulators. In former gene regulation typically occurs when transcription factors binds to cofactors such as coactivators and corepressors during initiation or elongation, while in later regulation occurs by chromatin remodelers and proteins that bind and modify the histones (Described in Chapter I.D). The transcription factor, cofactors and chromatin regulators are requirement for all the cell types and misregulation of these factors can cause a broad range of disease.

The present study gives an overview that how a subtle defect in any transcription factor can lead to the transcription dysregulation, further giving rise to diseases like XP, TTD, CS and XP/CS. Using such disease model, this study has provided information which definitely contributed to our knowledge of gene expression. Our observation suggests an implication of TFIIH in the active DNA demethylation associated to histone PTMs argues further investigation. The question arises is that; Does TFIIH affects a set of genes only or it influences the active DNA demethylation globally throughout the genome, is still a matter of investigation. TFIIH also interacts with XPG, XPF and XPC which are debated to be involved in DNA demethylation through BER and NER. Understanding the coordination of these transcription and NER factors, particularly

their link to the DNA methylation and demethylation machineries will show that if and how the chromatin directs DNA methyltransferase or demethylase. This could shed some light on the cause of several diseases including those discussed in present study.

Moreover active DNA demethylation is associated with the DNA conformation; R-loop, formed at the site of transcription bubble. One key factor which gives rise to the R-loop is the GC skew, which is measure of the distribution of guanine (G) and cytosine (C) residues across the two DNA strands. During transcription, it is thought that the newly synthesized RNA strand, upon leaving the RNA exit channel of the migrating RNA polymerase, can compete with the non-template DNA strand for re-annealing to the template DNA strand. (Ginno, Lott et al. 2012). It could be possible that R-loops may also signal the recruitment of DNA demethylating complexes to CG promoters and the unwinding of the DNA related to the TFIIH helicases is requires for the formation of R-loop and consequently the DNA demethylation.

Formation of RNA:DNA hybrids is also dependent on the stability of the resulting RNA:DNA hybrid and the exposed stretch of ssDNA. Clusters of G tracts on the non-template strand can fold into a stable G-quadruplex structure, which may help to stabilize the exposed ssDNA region of the R-loop. It is important to note that XPB and XPD helicase are shown to be recruited at the site of G-quadruplexes structure (Gray, Vallur et al. 2014). These studies again suggest that TFIIH holds some more functions and to delve deep certainly more investigations are required.

Another important aspect which contributes to the gene regulation is the structural organization of the genome. It will therefore not be surprising if the defects in the higher order chromatin structure and chromosome structure can cause diseases. However, it is often not very clear that the phenotype of disease observed is because of the changes brought up by local histone modifications resulting in altered gene-

expression or because of structural chromatin defects. There are several architectural proteins, like heterochromatin protein 1 (HP1) decorating the chromatin fibers, or CTCF and cohesin which are involved in the formation of chromatin loops, have been implicated in disease. CTCF also plays a role in stabilizing the higher order chromatin structure and has been suggested to play a role in human developmental disorders like Silver-Russell and Beckwith-Wiedeman Syndrome. Cohesin along with mediator is another protein known for its role in chromatin looping. It would not be a surprise, if these architectural proteins work together in a combinatorial manner in the gene regulation.

The TFIIH was also found to be interacting with mediator complex and the mediator complex connects the DNA binding proteins to the basal transcription machinery (Poss, Ebmeier et al. 2013). Also mutation in mediator gives rise to syndromes exhibiting very much similar clinical phenotypes. Studies based on the function and structural characterization of TFIIH and mediator during different transcription event could further provide knowledge of gene expression. Understanding the cellular molecular mechanism behind transcriptional regulation in XP, TTD, XP/CS patients could definitely help us to develop better diagnostic tools and therapeutic strategies for the complex diseases in future.

Conclusion

The present study underscores the functions of TFIIH in transcription. Understanding the relationship of TFIIH to NER factors in transcription could give an insight on the wide range of the clinical features shown by the XP, TTD, XP/CS patients.

- Mutations in XPB and XPD subunits lead to the disturbance in the enzymatic activities of TFIIH. These activities are required during the transcription and during the NER.
- Presence of functional TFIIH is essential condition for the recruitment of NER factors. TFIIH acts as the platform for the assembly of the NER factors. Mutations in XPB and XPD hampered the proper association of the TFIIH components at the promoter, which thwart the transcriptional network downstream.
- The chromatin remodeling associated with NER factors seems to be thwarted by the damaged TFIIH. The patterns of the histone modification and DNA demethylation were rather altered in all the XP, TTD and XP/CS cells.
- The TFIIH show its importance in DNA breaks associated with XPG and XPF at promoter and terminator and further aid the CTCF mediated chromatin looping between these two regions.
- p8/TTD-A cells, however shows no disturbance in NER factors recruitment and only affect DNA break and chromatin looping. This suggests a function association of p8/TTD-A with XPG or XPF directly or via some other subunit within TFIIH.

Conference:

- Role of TFIIH in recruitment of NER factors during transcription presented at doctoral school at 3ème journée des doctorants du Campus de médecine de Strasbourg, November 2012.

- Role of TFIIH linked to NER factors in the regulation of gene expression presented at the German-French DNA Repair Meeting: "Epigenetics and Genome Integrity" Illkirch, France, October 2013.

Courses:

- All roads lead to RNA: news and perspectives, January 2011
- Congenital disorders and their embryonic origins, January 2012
- Integrated structural biology, January-February 2012
- Insight from in-vivo studies, March 2012
- How to make the right impression and deliver clear messages along public speaking - Merck Millipore, March 2012
- Effective scientific writing, April 2012
- Career development workshop, June 2012
- Scientific Communication & Publishing, February 2013
- Industrial visit to F-Hoffmann- La Roche, Basel, Switzerland, January 2014

Bibliography

Abdulrahman, W., I. Iltis, et al. (2013). "ARCH domain of XPD, an anchoring platform for CAK that conditions TFIIH DNA repair and transcription activities." Proceedings of the National Academy of Sciences of the United States of America **110**(8): E633-642.

Aguilar-Fuentes, J., M. Fregoso, et al. (2008). "p8/TTDA Overexpression Enhances UV-Irradiation Resistance and Suppresses TFIIH Mutations in a <italic>Drosophila</italic> Trichothiodystrophy Model." PLoS genetics **4**(11): e1000253.

Akoulitchev, S., S. Chuikov, et al. (2000). "TFIIH is negatively regulated by cdk8-containing mediator complexes." Nature **407**(6800): 102-106.

Akoulitchev, S., T. P. Makela, et al. (1995). "Requirement for TFIIH kinase activity in transcription by RNA polymerase II." Nature **377**(6549): 557-560.

Al Tanoury, Z., A. Piskunov, et al. (2013). "Vitamin A and retinoid signaling: genomic and nongenomic effects." Journal of lipid research **54**(7): 1761-1775.

Allfrey, V. G. and A. E. Mirsky (1964). "Structural Modifications of Histones and their Possible Role in the Regulation of RNA Synthesis." Science **144**(3618): 559.

Andrau, J.-C., L. van de Pasch, et al. (2006). "Genome-Wide Location of the Coactivator Mediator: Binding without Activation and Transient Cdk8 Interaction on DNA." Molecular Cell **22**(2): 179-192.

Archambault, J., G. Pan, et al. (1998). "FCP1, the RAP74-interacting subunit of a human protein phosphatase that dephosphorylates the carboxyl-terminal domain of RNA polymerase IIO." The

Armache, K.-J., H. Kettenberger, et al. (2003). "Architecture of initiation-competent 12-subunit RNA polymerase II." Proceedings of the National Academy of Sciences **100**(12): 6964-6968.

Bannister, A. J. and T. Kouzarides (2011). "Regulation of chromatin by histone modifications." Cell research **21**(3): 381-395.

Barreto, G., A. Schafer, et al. (2007). "Gadd45a promotes epigenetic gene activation by repair-mediated DNA demethylation." Nature **445**(7128): 671-675.

Bensaude, O., F. Bonnet, et al. (1999). "Regulated phosphorylation of the RNA polymerase II C-terminal domain (CTD)." Biochem Cell Biol **77**(4): 249-255.

Biswas, T., J. M. Pero, et al. (2009). "DNA-Dependent ATPase Activity of Bacterial XPB Helicases." Biochemistry **48**(12): 2839-2848.

Black, Joshua C., C. Van Rechem, et al. (2012). "Histone Lysine Methylation Dynamics: Establishment, Regulation, and Biological Impact." Molecular Cell **48**(4): 491-507.

Botta, E., T. Nardo, et al. (2002). "Reduced level of the repair/transcription factor TFIIH in trichothiodystrophy." Hum Mol Genet **11**(23): 2919-2928.

Boulon, S., B. Pradet-Balade, et al. (2010). "HSP90 and Its R2TP/Prefoldin-like Cochaperone Are Involved in the Cytoplasmic Assembly of RNA Polymerase II." Molecular Cell **39**(6): 912-924.

Bourguet, W., V. Andry, et al. (2000). "Heterodimeric complex of RAR and RXR nuclear receptor ligand-binding domains: purification, crystallization, and preliminary X-ray diffraction analysis." Protein expression and purification **19**(2): 284-288.

Bradford, P. T., A. M. Goldstein, et al. (2011). "Cancer and neurologic degeneration in xeroderma pigmentosum: long term follow-up characterises the role of DNA repair." Journal of medical genetics **48**: 168-176.
Bradsher, J., F. Coin, et al. (2000). "Distinct roles for the helicases of TFIIH in transcript initiation and promoter escape." J Biol Chem **275**(4): 2532-2538.
Brand, M., C. Leurent, et al. (1999). "Three-dimensional structures of the TAFII-containing complexes TFIID and TFTC." Science **286**(5447): 2151-2153.
Broughton, B. C., H. Steingrimsdottir, et al. (1994). "Mutations in the xeroderma pigmentosum group D DNA repair/transcription gene in patients with trichothiodystrophy." Nature gnetics **7**: 189-194.
Budhu, A. S. and N. Noy (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." Molecular and Cellular Biology **22**(8): 2632-2641.
Buratowski, S. (2009). "Progression through the RNA polymerase II CTD cycle." Mol Cell **36**(4): 541-546.
Buratowski, S., S. Hahn, et al. (1989). "Five intermediate complexes in transcription initiation by RNA polymerase II." Cell **56**: 549-561.
Burley, S. K. and R. G. Roeder (1996). "Biochemistry and structural biology of transcription factor IID (TFIID)." Annu. Rev. Biochem. **65**: 769-799.
Burton, Z. F., M. Killeen, et al. (1988). "RAP30/74: a general initiation factor that binds to RNA polymerase II." Mol. Cell. Biol. **8**(4): 1602-1613.
Bushnell, D. A., K. D. Westover, et al. (2004). "Structural basis of transcription: an RNA polymerase II-TFIIB cocrystal at 4.5 Angstroms." Science **303**(5660): 983-988.
Caldecott, K. W., C. K. McKeown, et al. (1994). "An interaction between the mammalian DNA repair protein XRCC1 and DNA ligase III." Molecular and Cellular Biology **14**(1): 68-76.
Cao, J. and Q. Yan (2012). "Histone ubiquitination and deubiquitination in transcription, DNA damage response, and cancer." Frontiers in oncology **2**.
Carninci, P., A. Sandelin, et al. (2006). "Genome-wide analysis of mammalian promoter architecture and evolution." Nature genetics **38**(6): 626-635.
Carter, D., L. Chakalova, et al. (2002). "Long-range chromatin regulatory interactions in vivo." Nat Genet **32**(4): 623-626.
Chabot, P. R., L. Raiola, et al. (2014). "Structural and Functional Characterization of a Complex between the Acidic Transactivation Domain of EBNA2 and the Tfb1/p62 Subunit of TFIIH." PLoS Pathogens **10**(3).
Chalkley, G. E. and C. P. Verrijzer (1999). "DNA binding site selection by RNA polymerase II TAFs: a TAF(II)250-TAF(II)150 complex recognizes the initiator." EMBO Journal **18**(17): 4835-4845.
Chambon, P. (1996). "A decade of molecular biology of retinoic acid receptors." FASEB J. **10**: 940-954.
Chambon, P. (2005). "The nuclear receptor superfamily: a personal retrospect on the first two decades." Molecular endocrinology **19**(6): 1418-1428.
Chao, D. M., E. L. Gadbois, et al. (1996). "A mammalian SRB protein associated with an RNA polymerase II holoenzyme." Nature **380**(6569): 82-85.
Chen, D., T. Riedl, et al. (2000). "Activation of estrogen receptor alpha by S118 phosphorylation involves a ligand-dependent interaction with TFIIH and participation of CDK7." Mol Cell **6**(1): 127-137.
Chu, G. and L. Mayne (1996). "Xeroderma pigmentosum, Cockayne syndrome and trichothiodystrophy: do the genes explain the diseases?" Trends Genet **12**(5): 187-192.
Chymkowitch, P., N. Le May, et al. (2011). "The phosphorylation of the androgen receptor by TFIIH directs the ubiquitin/proteasome process." EMBO J **30**(3): 468-479.
Cleaver, J. E., E. T. Lam, et al. (2009). "Disorders of nucleotide excision repair: the genetic and molecular basis of heterogeneity." Nat Rev Genet **10**(11): 756-768.

Coin, F., J. Auriol, et al. (2004). "Phosphorylation of XPB helicase regulates TFIIH nucleotide excision repair activity." Embo J **23**(24): 4835-4846.

Coin, F., E. Bergmann, et al. (1999). "Mutations in XPB and XPD helicases found in xeroderma pigmentosum patients impair the transcription function of TFIIH." Embo J **18**(5): 1357-1366.

Coin, F., L. P. De Santis, et al. (2006). "p8/TTD-A as a Repair-Specific TFIIH Subunit." Mol Cell **21**(2): 215-226.

Coin, F., V. Oksenych, et al. (2007). "Distinct Roles for the XPB/p52 and XPD/p44 Subcomplexes of TFIIH in Damaged DNA Opening during Nucleotide Excision Repair." Mol Cell **26**(2): 245-256.

Coin, F., V. Oksenych, et al. (2008). "Nucleotide excision repair driven by the dissociation of CAK from TFIIH." Mol Cell **31**(1): 9-20.

Compe, E., P. Drane, et al. (2005). "Dysregulation of the Peroxisome Proliferator-Activated Receptor Target Genes by XPD Mutations." Mol Cell Biol **25**(14): 6065-6076.

Compe, E. and J. M. Egly (2012). "TFIIH: when transcription met DNA repair." Nat Rev Mol Cell Biol **13**(6): 343-354.

Compe, E., M. Malerba, et al. (2007). "Neurological defects in trichothiodystrophy reveal a coactivator function of TFIIH." Nature neuroscience **10**(11): 1414-1422.

Conaway, J. W. and R. C. Conaway (1989). "A multisubunit transcription factor essential for accurate initiation by RNA polymerase II." J. Biol. Chem. **264**(4): 2357-2362.

Conaway, J. W., J. P. Hanley, et al. (1991). "Transcription initiated by RNA polymerase II and transcription factors from liver. Structure and action of transcription factors epsilon and tau." J. Biol. Chem. **266**: 7804-7811.

Conaway, R. C. and J. W. Conaway (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. U.S.A. **86**: 7356-7360.

Corden, J., D. L. Cadena, et al. (1985). "A unique structure at the carboxyl terminus of the largest subunit of the eukaryotic RNA polymerase II." Proc. Natl. Acad. Sci. USA **82**: 7934-7938.

Cramer, P., K. J. Armache, et al. (2008). "Structure of Eukaryotic RNA Polymerases." Annual Review of Biophysics **37**(1): 337-352.

D'Alessio, A. C., I. C. G. Weaver, et al. (2007). "Acetylation-Induced Transcription Is Required for Active DNA Demethylation in Methylation-Silenced Genes." Molecular and Cellular Biology **27**(21): 7462-7474.

de Boer, J. and J. H. Hoeijmakers (2000). "Nucleotide excision repair and human syndromes." Carcinogenesis **21**(3): 453-460.

De Weerd-Kastelein, E. A., W. Keijzer, et al. (1972). "Genetic heterogeneity of xeroderma pigmentosum demonstrated by somatic cell hybridization." Nature: New biology **238**(81): 80-83.

Deaton, A. M. and A. Bird (2011). "CpG islands and the regulation of transcription." Genes & development **25**(10): 1010-1022.

Decker, K. B. and D. M. Hinton (2013). "Transcription Regulation at the Core: Similarities Among Bacterial, Archaeal, and Eukaryotic RNA Polymerases." Annual Review of Microbiology **67**(1): 113-139.

Dekker, J., K. Rippe, et al. (2002). "Capturing chromosome conformation." Science **295**(5558): 1306-1311.

Delva, L., J. N. Bastie, et al. (1999). "Physical and functional interactions between cellular retinoic acid binding protein II and the retinoic acid-dependent nuclear complex." Molecular and Cellular Biology **19**(10): 7158-7167.

Deng, W., B. Malecová, et al. (2009). "TFIIB Recognition Elements Control the TFIIA-NC2 Axis in Transcriptional Regulation." Molecular and Cellular Biology **29**(6): 1389-1400.

Dhalluin, C., J. E. Carlson, et al. (1999). "Structure and ligand of a histone acetyltransferase bromodomain." Nature **399**(6735): 491-496.

DiGiovanna, J. J. and K. H. Kraemer (2012). "Shining a light on xeroderma pigmentosum." The Journal of investigative dermatology **132**(3 Pt 2): 785-796.

Dilworth, F. J. and P. Chambon (2001). "Nuclear receptors coordinate the activities of chromatin remodeling complexes and coactivators to facilitate initiation of transcription." Oncogene **20**(24): 3047-3054.

Drapkin, R., J. T. Reardon, et al. (1994). "Dual role of TFIIH in DNA excision repair and in transcription by RNA polymerase II." Nature **368**: 769-772.

Dubaele, S., L. Proietti De Santis, et al. (2003). "Basal transcription defect discriminates between xeroderma pigmentosum and trichothiodystrophy in XPD patients." Mol Cell **11**(6): 1635-1646.

Dyson, H. J. and P. E. Wright (2005). "Intrinsically unstructured proteins and their functions." Nature reviews. Molecular cell biology **6**(3): 197-208.

Egloff, S., M. Dienstbier, et al. (2012). "Updating the RNA polymerase CTD code: adding gene-specific layers." Trends in genetics : TIG **28**(7): 333-341.

El Kaderi, B., S. Medler, et al. (2009). "Gene looping is conferred by activator-dependent interaction of transcription initiation and termination machineries." The Journal of biological chemistry **284**(37): 25015-25025.

Evans, E., J. G. Moggs, et al. (1997). "Mechanism of open complex and dual incision formation by human nucleotide excision repair factors." Embo J **16**(21): 6559-6573.

Fan, L., A. S. Arvai, et al. (2006). "Conserved XPB core structure and motifs for DNA unwinding: implications for pathway selection of transcription or excision repair." Mol Cell **22**(1): 27-37.

Feaver, W. J., O. Gileadi, et al. (1991). "Purification and characterization of yeast RNA polymerase II transcription factor b." J. Biol. Chem. **266**(28): 19000-19005.

Feaver, W. J., N. L. Henry, et al. (1997). "Genes for Tfb2, Tfb3, and Tfb4 subunits of yeast transcription/repair factor IIH. Homology to human cyclin-dependent kinase activating kinase and IIH subunits." J Biol Chem **272**(31): 19319-19327.

Feaver, W. J., J. Q. Svejstrup, et al. (1993). "Dual roles of a multiprotein complex from S. cerevisiae in transcription and DNA repair." Cell **75**(7): 1379-1387.

Flores, O., I. Ha, et al. (1990). "Factors involved in specific transcription by mammalian RNA polymerase II. Purification and subunit composition of transcription factor IIF." J. Biol. Chem. **265**(10): 5629-5634.

Flores, O., H. Lu, et al. (1991). "The small subunit of transcription factor IIF recruits RNA polymerase II into the preinitiation complex." Proc. Natl. Acad. Natl. USA **88**: 9999-10003.

Flores, O., H. Lu, et al. (1992). "Factors involved in specific transcription by mammalian RNA polymerase II. Identification and characterization of factor IIH." J Biol Chem **267**(4): 2786-2793.

Fousteri, M., W. Vermeulen, et al. (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**(4): 471-482.

Fregoso, M., J. P. Laine, et al. (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**(10): 3640-3650.

Fuks, F., P. J. Hurd, et al. (2003). "The methyl-CpG-binding protein MeCP2 links DNA methylation to histone methylation." The Journal of biological chemistry **278**(6): 4035-4040.

Fuss, J. O. and J. A. Tainer (2011). "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 **10**(7): 697-713.

Ganuza, M. S. L. C. C. M. G. G. S. R. B. M. A. P. D. P. J. M. S. D. B. M. (2012). "Genetic inactivation of Cdk7 leads to cell cycle arrest and induces premature aging due to adult stem cell exhaustion." The EMBO journal **31**(11): 2498-2510.

Ge, H. and R. G. Roeder (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.

Gerard, M., L. Fischer, et al. (1991). "Purification and interaction properties of the human RNA polymerase B(II) general transcription factor BTF2." The Journal of biological chemistry **266**(31): 20940-20945.

Gerber, H. P., M. Hagmann, et al. (1995). "RNA polymerase II C-terminal domain required for enhancer-driven transcription." Nature **374**: 660-662.

Germain, P., B. Staels, et al. (2006). "Overview of nomenclature of nuclear receptors." Pharmacological reviews **58**(4): 685-704.

Gershenzon, N. I. and I. P. Ioshikhes (2005). "Synergy of human Pol II core promoter elements revealed by statistical sequence analysis." Bioinformatics **21**(8): 1295-1300.

Giglia-Mari, G., C. Miquel, et al. (2006). "Dynamic Interaction of TTDA with TFIIH Is Stabilized by Nucleotide Excision Repair in Living Cells." PLoS Biol **4**(6): e156.

Gillespie, R. F. and L. J. Gudas (2007). "Retinoic acid receptor isotype specificity in F9 teratocarcinoma stem cells results from the differential recruitment of coregulators to retinoic response elements." The Journal of biological chemistry **282**(46): 33421-33434.

Ginno, Paul A., Paul L. Lott, et al. (2012). "R-Loop Formation Is a Distinctive Characteristic of Unmethylated Human CpG Island Promoters." Molecular Cell **45**(6): 814-825.

Glover-Cutter, K., S. Larochelle, et al. (2009). "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**(20): 5455-5464.

Gray, L. T., A. C. Vallur, et al. (2014). "G quadruplexes are genomewide targets of transcriptional helicases XPB and XPD." Nat Chem Biol **10**(4): 313-318.

Grunstein, M. (1997). "Histone acetylation in chromatin structure and transcription." Nature **389**(6649): 349-352.

Grunstein, M., A. Hecht, et al. (1995). "The regulation of euchromatin and heterochromatin by histones in yeast." Journal of cell science. Supplement **19**: 29-36.

Guidi, B. W., G. Bjornsdottir, et al. (2004). "Mutual targeting of mediator and the TFIIH kinase Kin28." J Biol Chem **279**(28): 29114-29120.

Hahn, S. (2004). "Structure and mechanism of the RNA polymerase II transcription machinery." Nature structural & molecular biology **11**(5): 394-403.

Hahn, S., S. Buratowski, et al. (1989). "Isolation of the gene encoding the yeast TATA binding protein TFIID: a gene identical to the SPT15 suppressor of Ty element insertions." Cell **58**(6): 1173-1181.

Hampsey, M. (1998). "Molecular genetics of the RNA polymerase II general transcriptional machinery." Microbiol Mol Biol Rev **62**(2): 465-503.

Hampsey, M., B. N. Singh, et al. (2011). "Control of eukaryotic gene expression: Gene loops and transcriptional memory." Advances in Enzyme Regulation **51**(1): 118-125.

Han, M. and M. Grunstein (1988). "Nucleosome loss activates yeast downstream promoters in vivo." Cell **55**(6): 1137-1145.

He, Y., J. Fang, et al. (2013). "Structural visualization of key steps in human transcription initiation." Nature **495**(7442): 481-486.
Hebra, F. and M. Kaposi (1874). "On diseases of the skin including exanthemata, volume III." New Sydenham Soc **61**: 252-258.
Hemavathy, K. C. and V. Nagaraja (1995). "DNA methylation in mycobacteria: Absence of methylation at GATC (Dam) and CCA/TGG (Dcm) sequences." FEMS Immunology & Medical Microbiology **11**(4): 291-296.
Hirose, Y. and J. L. Manley (2000). "RNA polymerase II and the integration of nuclear events." Genes & development **14**(12): 1415-1429.
Hoeijmakers, J. H. (2001). "Genome maintenance mechanisms for preventing cancer." Nature **411**(6835): 366-374.
Hoogstraten, D., A. L. Nigg, et al. (2002). "Rapid switching of TFIIH between RNA polymerase I and II transcription and DNA repair in vivo." Mol Cell **10**(5): 1163-1174.
Horikoshi, M., T. Yamamoto, et al. (1990). "Analysis of structure relationships of yeast TATA-box binding factor TFIID." Cell **61**: 1171-1178.
Hsin, J.-P., K. Xiang, et al. (2014). "Function and Control of RNA Polymerase II C-Terminal Domain Phosphorylation in Vertebrate Transcription and RNA Processing." Molecular and Cellular Biology **34**(13): 2488-2498.
Itin, P. H., A. Sarasin, et al. (2001). "Trichothiodystrophy: update on the sulfur-deficient brittle hair syndromes." J Am Acad Dermatol **44**(6): 891-920; quiz 921-894.
Ito, S., I. Kuraoka, et al. (2007). "XPG Stabilizes TFIIH, Allowing Transactivation of Nuclear Receptors: Implications for Cockayne Syndrome in XP-G/CS Patients." Mol Cell **26**(2): 231-243.
Ito, S., L. Shen, et al. (2011). "Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine." Science **333**(6047): 1300-1303.
Iyer, N., M. S. Reagan, et al. (1996). "Interactions involving the human RNA polymerase II transcription/nucleotide excision repair complex TFIIH, the nucleotide excision repair protein XPG, and Cockayne syndrome group B (CSB) protein." Biochemistry **35**(7): 2157-2167.
Jawhari, A., J. P. Laine, et al. (2002). "p52 Mediates XPB function within the transcription/repair factor TFIIH." J Biol Chem **277**(35): 31761-31767.
Jenuwein, T. and C. D. Allis (2001). "Translating the Histone Code." Science **293**(5532): 1074-1080.
Jin, Q., L. R. Yu, et al. (2011). "Distinct roles of GCN5/PCAF-mediated H3K9ac and CBP/p300-mediated H3K18/27ac in nuclear receptor transactivation." The EMBO journal **30**(2): 249-262.
Jishage, M., S. Malik, et al. (2012). "Transcriptional regulation by Pol II(G) involving mediator and competitive interactions of Gdown1 and TFIIF with Pol II." Mol Cell **45**(1): 51-63.
Ju, B.-G., V. V. Lunyak, et al. (2006). "A Topoisomerase IIß-Mediated dsDNA Break Required for Regulated Transcription." Science **312**(5781): 1798-1802.
Juven-Gershon, T. and J. T. Kadonaga (2010). "Regulation of gene expression via the core promoter and the basal transcriptional machinery." Developmental biology **339**(2): 225-229.
Kadonaga, J. T. (2002). "The DPE, a core promoter element for transcription by RNA polymerase II." Experimental & molecular medicine **34**(4): 259-264.
Kadonaga, J. T. (2012). "Perspectives on the RNA polymerase II core promoter." Wiley Interdisciplinary Reviews: Developmental Biology **1**(1): 40-51.
Kagey, M. H., J. J. Newman, et al. (2010). "Mediator and cohesin connect gene expression and chromatin architecture." Nature **467**(7314): 430-435.
Kainov, D. E., V. Cura, et al. (2010). "Structure determination of the minimal complex between Tfb5 and Tfb2, two subunits of the yeast transcription/DNA-repair factor TFIIH: a retrospective study." Acta Crystallogr D Biol Crystallogr **66**(Pt 7): 745-755.

Kamiuchi, S., M. Saijo, et al. (2002). "Translocation of Cockayne syndrome group A protein to the nuclear matrix: possible relevance to transcription-coupled DNA repair." Proc Natl Acad Sci U S A **99**(1): 201-206.

Kang, J. J., D. T. Auble, et al. (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.

Keeney, S., G. J. Chang, et al. (1993). "Characterization of a human DNA damage binding protein implicated in xeroderma pigmentosum E." J. Biol. Chem. **268**: 21293-21300.

Kellenberger, E., C. Dominguez, et al. (2005). "Solution structure of the C-terminal domain of TFIIH P44 subunit reveals a novel type of C4C4 ring domain involved in protein-protein interactions." J Biol Chem.

Keriel, A., A. Stary, et al. (2002). "XPD mutations prevent TFIIH-dependent transactivation by nuclear receptors and phosphorylation of RARalpha." Cell **109**(1): 125-135.

Kershnar, E., S. Y. Wu, et al. (1998). "Immunoaffinity purification and functional characterization of human transcription factor IIH and RNA polymerase II from clonal cell lines that conditionally express epitope-tagged subunits of the multiprotein complexes." J Biol Chem **273**(51): 34444-34453.

Kim, J. K., D. Patel, et al. (1995). "Contrasting structural impacts induced by cis-syn cyclobutane dimer and (6-4) adduct in DNA duplex decamers: implication in mutagenesis and repair activity." Photochemistry and photobiology **62**(1): 44-50.

Kim, M., H. Suh, et al. (2009). "Phosphorylation of the yeast Rpb1 C-terminal domain at serines 2, 5, and 7." J Biol Chem **284**(39): 26421-26426.

Knezetic, J. A. and D. S. Luse (1986). "The presence of nucleosomes on a DNA template prevents initiation by RNA polymerase II in vitro." Cell **45**(1): 95-104.

Kobayashi, T., M. Uchiyama, et al. (2002). "Mutations in the XPD gene in xeroderma pigmentosum group D cell strains: confirmation of genotype-phenotype correlation." Am J Med Genet **110**(3): 248-252.

Kostrewa, D., M. E. Zeller, et al. (2009). "RNA polymerase II-TFIIB structure and mechanism of transcription initiation." Nature **462**(7271): 323-330.

Kouzarides, T. (2007). "Chromatin modifications and their function." Cell **128**(4): 693-705.

Kraemer, K. H., M. M. Lee, et al. (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., M. M. Lee, et al. (1987). "Xeroderma pigmentosum. Cutaneous, ocular, and neurologic abnormalities in 830 published cases." Arch Dermatol **123**(2): 241-250.

Kraemer, K. H., D. D. Levy, et al. (1994). "Xeroderma pigmentosum and related disorders: examining the linkage between defective DNA repair and cancer." J. Invest. Dermatol. **103**(5 Suppl): 96S-101S.

Kraemer, K. H., N. J. Patronas, et al. (2007). "Xeroderma pigmentosum, trichothiodystrophy and Cockayne syndrome: a complex genotype-phenotype relationship." Neuroscience **145**(4): 1388-1396.

Kraemer, S. M., R. T. Ranallo, et al. (2001). "TFIIA interacts with TFIID via association with TATA-binding protein and TAF40." Molecular and Cellular Biology **21**(5): 1737-1746.

Kriukiene, E., Z. Liutkeviciute, et al. (2012). "5-Hydroxymethylcytosine - the elusive epigenetic mark in mammalian DNA." Chemical Society Reviews **41**(21): 6916-6930.

Krivega, I. and A. Dean (2012). "Enhancer and promoter interactions-long distance calls." Current opinion in genetics & development **22**(2): 79-85.

Kuehner, J. N., E. L. Pearson, et al. (2011). "Unravelling the means to an end: RNA polymerase II transcription termination." Nature reviews. Molecular cell biology **12**(5): 283-294.

Lafrance-Vanasse, J., G. Arseneault, et al. (2013). "Structural and functional evidence that Rad4 competes with Rad2 for binding to the Tfb1 subunit of TFIIH in NER." Nucleic acids research **41**(4): 2736-2745.

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

Laine, J. P., B. N. Singh, et al. (2009). "A physiological role for gene loops in yeast." Genes & development **23**(22): 2604-2609.

Lake, R. J., E. L. Boetefuer, et al. (2014). "The Sequence-Specific Transcription Factor c-Jun Targets Cockayne Syndrome Protein B to Regulate Transcription and Chromatin Structure." PLoS genetics **10**(4): e1004284.

Laugel, V., C. Dalloz, et al. (2010). "Mutation update for the CSB/ERCC6 and CSA/ERCC8 genes involved in Cockayne syndrome." Human mutation **31**(2): 113-126.

Laurent Fischer, M. G., Christian Chalut, Yves Lutz,Sandrine Humbert, Masamoto Kanno, Pierre Chambon,Jean-Marc Egly* (1992). "<Science-1992-Fischer-1392-5.pdf>."

le Maire, A., C. Teyssier, et al. (2010). "A unique secondary-structure switch controls constitutive gene repression by retinoic acid receptor." Nature structural & molecular biology **17**(7): 801-807.

Le May, N., D. Fradin, et al. (2012). "XPG and XPF Endonucleases Trigger Chromatin Looping and DNA Demethylation for Accurate Expression of Activated Genes." Mol Cell **47**(4): 622-632.

Le May, N., D. Mota-Fernandes, et al. (2010). "NER factors are recruited to active promoters and facilitate chromatin modification for transcription in the absence of exogenous genotoxic

Lee, J. H., C. J. Park, et al. (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 research **32**(8): 2474-2481.

Lehmann, A. R., D. McGibbon, et al. (2011). "Xeroderma pigmentosum." Orphanet J Rare Dis **6**: 70.

Lei, L., D. Ren, et al. (1999). "The RAP74 subunit of human transcription factor IIF has similar roles in initiation and elongation." Molecular & Cellular Biology **19**(12): 8372-8382.

Lewis, B. A., T. K. Kim, et al. (2000). "A downstream element in the human beta-globin promoter: evidence of extended sequence-specific transcription factor IID contacts." Proceedings of the National Academy of Sciences of the United States of America **97**(13): 7172-7177.

Li, E. and Y. Zhang (2014). "DNA Methylation in Mammals." Cold Spring Harbor perspectives in biology

Lindahl, T. (1974). "An N-glycosidase from Escherichia coli that releases free uracil from DNA containing **6**(5).

deaminated cytosine residues." Proceedings of the National Academy of Sciences of the United States of America **71**(9): 3649-3653.

Lindenbaum, Y., D. Dickson, et al. (2001). "Xeroderma pigmentosum/cockayne syndrome complex: first neuropathological study and review of eight other cases." European journal of paediatric neurology : EJPN : official journal of the European Paediatric Neurology Society **5**(6): 225-242.

Liu, H., J. Rudolf, et al. (2008). "Structure of the DNA repair helicase XPD." Cell **133**(5): 801-812.

Liu, X., D. A. Bushnell, et al. (2010). "Structure of an RNA polymerase II-TFIIB complex and the transcription initiation mechanism." Science **327**(5962): 206-209.

Lu, H., L. Zawel, et al. (1992). "Human general transcription factor IIH phosphorylates the C-terminal domain of RNA polymerase II." Nature **358**(6388): 641-645.

Luse, D. S., L. C. Spangler, et al. (2011). "Efficient and rapid nucleosome traversal by RNA polymerase II depends on a combination of transcript elongation factors." The Journal of biological chemistry **286**(8): 6040-6048.

Maiti, A. and A. C. Drohat (2011). "Thymine DNA glycosylase can rapidly excise 5-formylcytosine and 5-carboxylcytosine: Potential implications for active demethylation of CpG sites." Journal of Biological Chemistry.

Maldonado, E., R. Shiekhattar, et al. (1996). "A human RNA polymerase II complex associated with SRB and DNA-repair proteins." Nature **381**(6577): 86-89.

Mangelsdorf, D. J. and R. M. Evans (1995). "The RXR heterodimers and orphan receptors." Cell **83**(6): 841-850.

Mangelsdorf, D. J., C. Thummel, et al. (1995). "The nuclear receptor superfamily: the second decade." Cell **83**(6): 835-839.

Marteijn, J. A., H. Lans, et al. (2014). "Understanding nucleotide excision repair and its roles in cancer and ageing." Nature reviews. Molecular cell biology **15**(7): 465-481.

Matsui, T., J. Segall, et al. (1980). "Multiple factors required for accurate initiation of transcription by purified RNA polymerase II." J. Biol. Chem **255**: 11992-11996.

Maxon, M. E., J. A. Goodrich, et al. (1994). "Transcription factor IIE binds preferentially to RNA polymerase IIa and recruits TFIIH: a model for promoter clearance." Genes Dev **8**(5): 515-524.

Mellon, I., G. Spivak, et al. (1987). "Selective removal of transcription-blocking DNA damage from the transcribed strand of the mammalian DHFR gene." Cell **51**: 241-249.

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

Misteli, T. (2000). "Cell biology of transcription and pre-mRNA splicing: nuclear architecture meets nuclear function." Journal of cell science **113**(11): 1841-1849.

Mocquet, V., J. P. Laine, et al. (2008). "Sequential recruitment of the repair factors during NER: the role of XPG in initiating the resynthesis step." Embo J **27**(1): 155-167.

Montavon, T. and D. Duboule (2012). "Landscapes and archipelagos: spatial organization of gene regulation in vertebrates." Trends in Cell Biology **22**(7): 347-354.

Morice-Picard, F., M. Cario-André, et al. (2009). "New clinico-genetic classification of trichothiodystrophy." American Journal of Medical Genetics Part A **149A**(9): 2020-2030.

Nan, X., H. H. Ng, et al. (1998). "Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex." Nature **393**(6683): 386-389.

Nonnekens, J., S. Cabantous, et al. (2013). "In vivo interactions of TTDA mutant proteins within TFIIH." Journal of cell science **126**(15): 3278-3283.

O'Reilly, D. and D. R. Greaves (2007). "Cell-type-specific expression of the human CD68 gene is associated with changes in Pol II phosphorylation and short-range intrachromosomal gene looping." Genomics **90**(3): 407-415.

Ohkuma, Y. and R. G. Roeder (1994). "Regulation of TFIIH ATPase and kinase activities by TFIIE during active initiation complex formation." Nature **368**(6467): 160-163.

Okano, M., D. W. Bell, et al. (1999). "DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development." Cell **99**(3): 247-257.

Oksenych, V., B. Bernardes de Jesus, et al. (2009). "Molecular insights into the recruitment of TFIIH to sites of DNA damage." The EMBO journal **28**(19): 2971-2980.

Okuda, M., A. Tanaka, et al. (2004). "A novel zinc finger structure in the large subunit of human general transcription factor TFIIE." J Biol Chem **279**(49): 51395-51403.

Orphanides, G., T. Lagrange, et al. (1996). "The general transcription factors of RNA polymerase II." Genes Dev **10**(21): 2657-2683.

Patel, S. A. and M. C. Simon (2010). "Functional Analysis of the Cdk7·Cyclin H·Mat1 Complex in Mouse Embryonic Stem Cells and Embryos." Journal of Biological Chemistry **285**(20): 15587-15598.

Perissi, V., K. Jepsen, et al. (2010). "Deconstructing repression: evolving models of co-repressor action." Nature reviews. Genetics **11**(2): 109-123.

Petermann, E., M. Ziegler, et al. (2003). "ATP-dependent selection between single nucleotide and long patch base excision repair." DNA repair **2**(10): 1101-1114.

Pollitt, R. J., F. A. Jenner, et al. (1968). "Sibs with mental and physical retardation and trichorrhexis nodosa with abnormal amino acid composition of the hair." Arch Dis Child **43**(228): 211-216.

Portal, M. M. (2011). "MicroRNA-27a regulates basal transcription by targeting the p44 subunit of general transcription factor IIH." Proceedings of the National Academy of Sciences **108**(21): 8686-8691.

Poss, Z. C., C. C. Ebmeier, et al. (2013). "The Mediator complex and transcription regulation." Critical reviews in biochemistry and molecular biology **48**(6): 575-608.

Price, V. H., R. B. Odom, et al. (1980). "Trichothiodystrophy: sulfur-deficient brittle hair as a marker for a neuroectodermal symptom complex." Arch Dermatol **116**(12): 1375-1384.

Proietti-De-Santis, L., P. Drane, et al. (2006). "Cockayne syndrome B protein regulates the transcriptional program after UV irradiation." Embo J **25**(9): 1915-1923.

Proudfoot, N. J., A. Furger, et al. (2002). "Integrating mRNA Processing with Transcription." Cell **108**(4): 501-512.

Qi, Z. P. R. A. S. M. C. Y. R. (2013). "Sequence-dependent base pair stepping dynamics in XPD helicase unwinding." eLife **2**.

Ranish, J. A., S. Hahn, et al. (2004). "Identification of TFB5, a new component of general transcription and DNA repair factor IIH." Nature genetics **36**(7): 707-713.

Rapin, I. (2013). Chapter 167 - Disorders of nucleotide excision repair. Handbook of Clinical Neurology. M. L. Olivier Dulac and B. S. Harvey, Elsevier. **Volume 113:** 1637-1650.

Rea, S., F. Eisenhaber, et al. (2000). "Regulation of chromatin structure by site-specific histone H3 methyltransferases." Nature **406**(6796): 593-599.

Reinberg, D. and R. G. Roeder (1987). "Factors involved in specific transcription by mammalian RNA polymerase II. Purification and functional analysis of initiation factors IIB and IIE." J Biol Chem **262**(7): 3310-3321.

Renaud, J. P. and D. Moras (2000). "Structural studies on nuclear receptors." Cellular and molecular life sciences : CMLS **57**(12): 1748-1769.

Riedl, T. H. F. E. J. M. (2003). "The comings and goings of nucleotide excision repair factors on damaged DNA." The EMBO journal **22**(19): 5293-5303.

Riou, L., L. Zeng, et al. (1999). "The relative expression of mutated XPB genes results in xeroderma pigmentosum/Cockayne's syndrome or trichothiodystrophy cellular phenotypes." Hum Mol Genet **8**(6): 1125-1133.

Robbins, J. (1991). "Xeroderma pigmentosum complementation group H is withdrawn and reassigned to group D." Human genetics **88**(2): 242-242.

Rochette-Egly, C. (2014). "Retinoic acid signaling and mouse embryonic stem cell differentiation: Cross talk between genomic and non-genomic effects of RA." Biochimica et biophysica acta.

Rochette-Egly, C., S. Adam, et al. (1997). "Stimulation of RAR alpha activation function AF-1 through binding to the general transcription factor TFIIH and phosphorylation by CDK7." Cell **90**(1): 97-107.

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

Roeder, R. G. and W. J. Rutter (1970). "Specific nucleolar and nucleoplasmic RNA polymerases." Proceedings of the National Academy of Sciences of the United States of America **65**(3): 675-682.

Rossignol, M., I. Kolb-Cheynel, et al. (1997). "Substrate specificity of the cdk-activating kinase (CAK) is altered upon association with TFIIH." The EMBO journal **16**(7): 1628-1637.

Rothbart, S. B. and B. D. Strahl (2014). "Interpreting the language of histone and DNA modifications." Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms **1839**(8): 627-643.

Roy, R., J. P. Adamczewski, et al. (1994). "The MO15 cell cycle kinase is associated with the TFIIH transcription-DNA repair factor." Cell **79**(6): 1093-1101.

Ruppert, S. and R. Tjian (1995). "Human TAFII250 interacts with RAP74: implications for RNA polymerase II initiation." Genes Dev **9**: 2747-2755.

Ruthenburg, A. J., C. D. Allis, et al. (2007). "Methylation of Lysine 4 on Histone H3: Intricacy of Writing and Reading a Single Epigenetic Mark." Molecular Cell **25**(1): 15-30.

Samarut, E. and C. Rochette-Egly (2012). "Nuclear retinoic acid receptors: conductors of the retinoic acid symphony during development." Molecular and cellular endocrinology **348**(2): 348-360.

Samuels, M., A. Fire, et al. (1982). "Separation and characterization of factors mediating accurate transcription by RNA polymerase II." J Biol Chem **257**(23): 14419-14427.

Sandelin, A., P. Carninci, et al. (2007). "Mammalian RNA polymerase II core promoters: insights from genome-wide studies." Nat Rev Genet **8**(6): 424-436.

Sandrock, B. and J. M. Egly (2001). "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**(38): 35328-35333.

Sarasin, A. (2012). "UVSSA and USP7: new players regulating transcription-coupled nucleotide excision repair in human cells." Genome Medicine **4**(5): 44.

Sarasin A, B.-B. C., Renault G, Lehmann A, Arlett C, Dumez Y. (1992). "Prenatal diagnosis in a subset of trichothiodystrophy patients defective in DNA repair." Br J Dermatol. **127**(5): 485-491.

Sawadogo, M. and A. Sentenac (1990). "RNA polymerase B (II) and general transcription factors." Annu. Rev. Biochem. **59**: 711-754.

Scafe, C., D. Chao, et al. (1990). "RNA polymerase II C-terminal repeat influences response to transcriptional enhancer signals." Nature **347**: 491-494.

Schaeffer, L., V. Moncollin, et al. (1994). "The ERCC2/DNA repair protein is associated with the class II BTF2/TFIIH transcription factor." EMBO J. **13**(10): 2388-2392.

Schaeffer, L., R. Roy, et al. (1993). "DNA repair helicase: a component of BTF2 (TFIIH) basic transcription factor." Science **260**(5104): 58-63.

Schäfer, A., E. Karaulanov, et al. (2013). "Ing1 functions in DNA demethylation by directing Gadd45a to H3K4me3." Genes & Development **27**(3): 261-273.

Schärer, O. D. (2008). "The molecular basis for different disease states caused by mutations in TFIIH and XPG." DNA repair **7**(2): 339-344.

Schmitt, D. R., J. Kuper, et al. (2014). "The Structure of the TFIIH p34 Subunit Reveals a <italic>Von Willebrand Factor A</italic> Like Fold." PloS one **9**(7): e102389.

Schmitz, K. M., N. Schmitt, et al. (2009). "TAF12 recruits Gadd45a and the nucleotide excision repair complex to the promoter of rRNA genes leading to active DNA demethylation." Mol Cell **33**(3): 344-353.

Schultz, P., S. Fribourg, et al. (2000). "Molecular structure of human TFIIH." Cell **102**(5): 599-607.

Scott, R. J., P. Itin, et al. (1993). "Xeroderma pigmentosum-Cockayne syndrome complex in two, new patient: absence of skin tumors despite severe deficiency of DNA excision repair." J. Am. Acad **29**: 883-889.

Scrima, A., R. Konickova, et al. (2008). "Structural basis of UV DNA-damage recognition by the DDB1-DDB2 complex." Cell **135**(7): 1213-1223.

Serizawa, H., J. W. Conaway, et al. (1993). "Phosphorylation of C-terminal domain of RNA polymerase II is not required in basal transcription." Nature **363**(6427): 371-374.

Seroz, T., C. Perez, et al. (2000). "p44/SSL1, the regulatory subunit of the XPD/RAD3 helicase, plays a crucial role in the transcriptional activity of TFIIH." J Biol Chem **275**(43): 33260-33266.

Shandilya, J. and S. G. Roberts (2012). "The transcription cycle in eukaryotes: from productive initiation to RNA polymerase II recycling." Biochimica et biophysica acta **1819**(5): 391-400.

Shilatifard, A., R. C. Conaway, et al. (2003). "The RNA polymerase II elongation complex." Annual review of biochemistry **72**: 693-715.

Shykind, B. M., J. Kim, et al. (1995). "Activation of the TFIID-TFIIA complex with HMG-2." Genes & development **9**(11): 1354-1365.

Simon, Jeffrey A. and Robert E. Kingston (2013). "Occupying Chromatin: Polycomb Mechanisms for Getting to Genomic Targets, Stopping Transcriptional Traffic, and Staying Put." Molecular Cell **49**(5): 808-824.

Smale, S. T. and D. Baltimore (1989). "The "initiator" as transcription control element." Cell **57**: 103-113.

Smale, S. T. and J. T. Kadonaga (2003). "The RNA polymerase II core promoter." Annual review of biochemistry **72**: 449-479.

Sologub, M., D. Litonin, et al. (2009). "TFB2 Is a Transient Component of the Catalytic Site of the Human Mitochondrial RNA Polymerase." Cell **139**(5): 934-944.

Spilianakis, C. G. and R. A. Flavell (2004). "Long-range intrachromosomal interactions in the T helper type 2 cytokine locus." Nature immunology **5**(10): 1017-1027.

Splinter, E., H. Heath, et al. (2006). "CTCF mediates long-range chromatin looping and local histone modification in the β-globin locus." Genes & Development **20**(17): 2349-2354.

Stefanini, M., S. Giliani, et al. (1992). "DNA repair investigations in nine Italian patients affected by trichothiodystrophy." Mutat Res **273**(2): 119-125.

Stefanini, M., P. Lagomarsini, et al. (1986). "Xeroderma pigmentosum (complementation group D) mutation is present in patients affected with trichothiodystrophy with photosensitivity." Hum. Genet. **74**: 107-112.

Stefanini, M., P. Lagomarsini, et al. (1993). "Genetic heterogeneity of the excision repair defect associated with trichothiodystrophy." Carcinogenesis **14**(6): 1101-1105.

Stefanini, M., W. Vermeulen, et al. (1993). "A new nucleotide-excision repair gene associated with the disorder trichothiodystrophy." Am. J. Hum. Genet. **53**(4): 817-821.

Sugasawa, K., Y. Okuda, et al. (2005). "UV-induced ubiquitylation of XPC protein mediated by UV-DDB-ubiquitin ligase complex." Cell **121**(3): 387-400.

Tahiliani, M., K. P. Koh, et al. (2009). "Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1." Science **324**(5929): 930-935.

Takayama, K., E. P. Salazar, et al. (1995). "Defects in the DNA repair and transcription gene ERCC2 in the cancer-prone disorder xeroderma pigmentosum group D." Cancer Res. **55**: 5656-5663.

Takedachi, A., M. Saijo, et al. (2010). "DDB2 complex-mediated ubiquitylation around DNA damage is oppositely regulated by XPC and Ku and contributes to the recruitment of XPA." Molecular and Cellular Biology **30**(11): 2708-2723.

Tan-Wong, S. M., J. D. French, et al. (2008). "Dynamic interactions between the promoter and terminator regions of the mammalian BRCA1 gene." Proceedings of the National Academy of Sciences **105**(13): 5160-5165.

Tanaka, A., T. Watanabe, et al. (2009). "Central forkhead domain of human TFIIE beta plays a primary role in binding double-stranded DNA at transcription initiation." Genes to cells : devoted to molecular & cellular mechanisms **14**(3): 395-405.

Taylor, E., B. Broughton, et al. (1997). "Xeroderma pigmentosum and trichothiodystrophy are associated with different mutations in the XPD (ERCC2) repair/transcription gene." Proc. Natl. Acad. Sci. USA **94**(16): 8658-8663.

Theil, A. F., J. Nonnekens, et al. (2011). "Slowly Progressing Nucleotide Excision Repair in Trichothiodystrophy Group A Patient Fibroblasts." Molecular and Cellular Biology **31**(17): 3630-3638.

Thomas, M. C. and C. M. Chiang (2006). "The general transcription machinery and general cofactors." Crit Rev Biochem Mol Biol **41**(3): 105-178.

Thompson, L. L., B. J. Guppy, et al. (2013). "Regulation of chromatin structure via histone post-translational modification and the link to carcinogenesis." Cancer metastasis reviews **32**(3-4): 363-376.

Tirode, F., D. Busso, et al. (1999). "Reconstitution of the transcription factor TFIIH: assignment of functions for the three enzymatic subunits, XPB, XPD, and cdk7." Mol Cell **3**(1): 87-95.

Tokusumi, Y., Y. Ma, et al. (2007). "The New Core Promoter Element XCPE1 (X Core Promoter Element 1) Directs Activator-, Mediator-, and TATA-Binding Protein-Dependent but TFIID-Independent RNA Polymerase II Transcription from TATA-Less Promoters." Molecular and Cellular Biology **27**(5): 1844-1858.

Újvári, A., M. Pal, et al. (2011). "The Functions of TFIIF during Initiation and Transcript Elongation Are Differentially Affected by Phosphorylation by Casein Kinase 2." Journal of Biological Chemistry **286**(26): 23160-23167.

Van Dyke, M. W., R. G. Roeder, et al. (1988). "Physical analysis of transcription preinitiation complex assembly on class II gene promoter." Science **241**: 1335-1338.

Vélez-Cruz, R. and J.-M. Egly (2013). "Cockayne syndrome group B (CSB) protein: At the crossroads of transcriptional networks." Mechanisms of Ageing and Development **134**(5–6): 234-242.

Vermeulen, W., M. Stefanini, et al. (1991). "Xeroderma pigmentosum complementation group H falls into complementation group D." Mutat. Res. **255**: 201-208.

Vernimmen, D., M. De Gobbi, et al. (2007). "Long-range chromosomal interactions regulate the timing of the transition between poised and active gene expression." The EMBO journal **26**(8): 2041-2051.

Wang, H., M. T. Maurano, et al. (2012). "Widespread plasticity in CTCF occupancy linked to DNA methylation." Genome research **22**(9): 1680-1688.

Wang, X., L. Spangler, et al. (2003). "Promoter escape by RNA polymerase II. Downstream promoter DNA is required during multiple steps of early transcription." J Biol Chem **278**(12): 10250-10256.

Wang, Y., J. A. Fairley, et al. (2010). "Phosphorylation of TFIIB Links Transcription Initiation and Termination." Current Biology **20**(6): 548-553.

Warnmark, A., E. Treuter, et al. (2003). "Activation functions 1 and 2 of nuclear receptors: molecular strategies for transcriptional activation." Molecular endocrinology **17**(10): 1901-1909.

Weeda, G., E. Eveno, et al. (1997). "A mutation in the XPB/ERCC3 DNA repair transcription gene, associated with trichothiodystrophy." Am. J. Hum. Genet. **60**: 320-329.

Weeda, G., R. van Ham, et al. (1990). "Molecular cloning and biological characterization of the human excision repair gene ERCC3." Mol. Cell. Biol. **10**(6): 2570-2581.

Weeda, G., R. C. A. van Ham, et al. (1990). "A presumed DNA helicase encoded by ERCC-3 is involved in the human repair disorders Xeroderma Pigmentosum and Cockayne's syndrome." Cell **62**: 777-791.

Wild, T. and P. Cramer (2012). "Biogenesis of multisubunit RNA polymerases." Trends in Biochemical Sciences **37**(3): 99-105.

Winkler, G. S., W. Vermeulen, et al. (1998). "Affinity purification of human DNA repair/transcription factor TFIIH using epitope-tagged xeroderma pigmentosum B protein." J Biol Chem **273**(2): 1092-1098.

Wong, Koon H., Y. Jin, et al. (2014). "TFIIH Phosphorylation of the Pol II CTD Stimulates Mediator Dissociation from the Preinitiation Complex and Promoter Escape." Molecular Cell **54**(4): 601-612.

Woodcock, C. L. and S. Dimitrov (2001). "Higher-order structure of chromatin and chromosomes." Current opinion in genetics & development **11**(2): 130-135.

Wossidlo, M. A. J. S. V. L. K. B. M. R. R. S. H. W. J. (2010). "Dynamic link of DNA demethylation, DNA strand breaks and repair in mouse zygotes." The EMBO journal **29**(11): 1877-1888.

Wrange, O., S. Okret, et al. (1984). "Characterization of the purified activated glucocorticoid receptor from rat liver cytosol." The Journal of biological chemistry **259**(7): 4534-4541.

Wu, W., H. Liu, et al. (2009). "Genetic variants in GTF2H1 and risk of lung cancer: A case-control analysis in a Chinese population." Lung Cancer **63**(2): 180-186.

Wysocka, J., T. Swigut, et al. (2006). "A PHD finger of NURF couples histone H3 lysine 4 trimethylation with chromatin remodelling." Nature **442**(7098): 86-90.

Xiao, H., A. Pearson, et al. (1994). "Binding of basal transcription factor TFIIH to the acidic activation domains of VP16 and p53." Mol Cell Biol **14**(10): 7013-7024.

Xie, X., T. Kokubo, et al. (1996). "Structural similarity between TAFs and the heterotetrameric core of the histone octamer." Nature **380**(6572): 316-322.

Yamada, T., Y. Yamaguchi, et al. (2006). "P-TEFb-mediated phosphorylation of hSpt5 C-terminal repeats is critical for processive transcription elongation." Molecular Cell **21**(2): 227-237.

Yamamoto, Y. Y., H. Ichida, et al. (2007). "Differentiation of core promoter architecture between plants and mammals revealed by LDSS analysis." Nucleic acids research **35**(18): 6219-6226.

Yan, Q., R. J. Moreland, et al. (1999). "Dual roles for transcription factor IIF in promoter escape by RNA polymerase II." J Biol Chem **274**(50): 35668-35675.

Yokoi, M., C. Masutani, et al. (2000). "The xeroderma pigmentosum group C protein complex XPC-HR23B plays an important role in the recruitment of transcription factor IIH to damaged DNA [In Process Citation]." J Biol Chem **275**(13): 9870-9875.

Yonaha, M., T. Tsuchiya, et al. (1997). "Cell-cycle-dependent phosphorylation of the basal transcripiton factor RAP74." FEBS letters **410**: 477-480.

Yu, W., V. Ginjala, et al. (2004). "Poly(ADP-ribosyl)ation regulates CTCF-dependent chromatin insulation." Nature genetics **36**(10): 1105-1110.

Zechel, C., X. Q. Shen, et al. (1994). "Dimerization interfaces formed between the DNA binding domains determine the cooperative binding of RXR/RAR and RXR/TR heterodimers to DR5 and DR4 elements." The EMBO journal **13**(6): 1414-1424.

Zentner, G. E. and S. Henikoff (2013). "Regulation of nucleosome dynamics by histone modifications." Nature structural & molecular biology **20**(3): 259-266.

Zurita, M. and C. Merino (2003). "The transcriptional complexity of the TFIIH complex." Trends in genetics : TIG **19**(10): 578-584.

UNIVERSITÉ DE STRASBOURG

École Doctorale des Sciences de la Vie et de la Santé STRASBOURG

Amita SINGH

Contribution de TFIIH dans le remodelage de la chromatine dépendant des facteurs NER lors de la transcription

Résumé en Français

La transcription fidèle d'un gène lors de son activation nécessite l'assemblage d'un ensemble de protéines autour du promoteur. Parmi ces protéines, le complexe TFIIH joue un rôle central et important au travers de ses sous-unités enzymatiques. Des mutations dans les sous-unités XPB, XPD et p8/TTD-A de TFIIH conduisent à trois maladies autosomiques récessives distinctes : xeroderma pigmentosum (XP), parfois associés avec le syndrome de Cockayne (XP/CS) et la trichothiodystrophie (TTD). En étudiant différentes mutations dans ces trois sous-unités de TFIIH, nous avons montré que chaque mutation analysée conduit à une dérégulation transcriptionnelle spécifique du gène RARβ2, gène cible des RAR. L'intégrité architecturale et enzymatique de TFIIH conditionne le bon recrutement du complexe TFIIH et également des facteurs de réparation par excision de nucléotides (NER). TFIIH muté perturbe leur recrutement et par conséquence compromet le remodelage de la chromatine médiée par les facteurs NER tels que les modifications post-traductionnelles (PTMs) des histones, l'induction des cassures de l'ADN, la déméthylation de l'ADN et les boucles de chromatine. Par conséquence, en plus de ses activités enzymatiques, TFIIH forme une plate-forme afin de recruter les facteurs NER et orchestres les fonctions connexes de la transcription. Cette pénétrance variable parmi les mutants donne lieu à un gradient de phénotype observé chez les patients TTD, XP ou XP/CS. Mot-clés : TFIIH, NER, transcription, chromatine

Résumé en anglais

Fidelity in transcription of the gene requires assembly of set of proteins around the promoter, upon gene activation. The TFIIH complex is central among these proteins and plays a key role through its enzymatic subunits. Mutations in TFIIH subunits XPB, XPD and p8/TTD-A leads to three distinct autosomal recessive disorders: xeroderma pigmentosum (XP), sometimes associated with Cockayne's syndrome (XP/CS) and trichothiodystrophy (TTD). By studying the different mutation in these three subunits of TFIIH from mentioned genetic disease models, we have shown that each mutation analyzed led to a specific transcriptional dysregulation of the RAR-target gene RAR□2. The architectural and enzymatic integrity of TFIIH condition the appropriate recruitment of TFIIH complex and further the arrival of the Nucleotide Excision Repair (NER) factors. By disturbing their recruitment, mutated TFIIH consequently compromised the chromatin remodeling mediated by NER factors such as histones post-translational modifications (PTMs), DNA breaks induction, DNA demethylation and gene looping. Hence it can be concluded that in addition to its enzymatic activities, TFIIH provide a platform to recruit the NER factors and orchestrates the related functions in transcription. Such varying penetrance among mutants gives rise to a phenotype gradient as observed in TTD, XP or XP/CS patients. Keywords: TFIIH, NER, transcription, chromatin.