

Université de Montréal

Caractérisation structurale et fonctionnelle des interactions impliquant TFIIH
et la machinerie de réparation de l'ADN

par

Julien Lafrance-Vanasse

Département de Biochimie

Faculté de Médecine

Thèse présentée à la Faculté de Médecine
en vue de l'obtention du grade de Philosophiae Doctor (Ph.D.)
en Biochimie

27 septembre, 2011

© Julien Lafrance-Vanasse, 2011

Université de Montréal
Faculté des études supérieures et postdoctorales

Cette thèse intitulée:

Caractérisation structurale et fonctionnelle des interactions impliquant TFIID
et la machinerie de réparation de l'ADN

Présentée par :
Julien Lafrance-Vanasse

a été évaluée par un jury composé des personnes suivantes :

Gerardo Ferbeyre, président-rapporteur
James G. Omichinski , directeur de recherche
Jurgen Sygusch, co-directeur
Damien D'Amours, membre du jury
Sylvie Doublié, examinatrice externe
Elliot Drobetsky, représentant du doyen de la FES

Résumé

La réparation de l'ADN par excision des nucléotides (NER) est un mécanisme capable de retirer une large variété de lésions causant une distorsion de la double hélice, comme celles causées par les rayons ultraviolets (UV). Comme toutes les voies de réparation de l'ADN, la NER contribue à la prévention de la carcinogénèse en prévenant la mutation de l'ADN. Lors de ce processus, il y a d'abord reconnaissance de la lésion par la protéine XPC/Rad4 (humain/levure) qui recrute ensuite TFIIH. Ce complexe déroule l'ADN par son activité hélicase et recrute l'endonucléase XPG/Rad2 ainsi que d'autres protéines nécessaires à l'excision de l'ADN. Lors de son arrivée au site de lésion, XPG/Rad2 déplace XPC/Rad4.

TFIIH agit également lors de la transcription de l'ADN, entre autres par son activité hélicase. Outre cette similarité de la présence de TFIIH lors de la transcription et la réparation, il est possible de se demander en quoi les deux voies sont similaires. Nous nous sommes donc intéressés aux interactions impliquant TFIIH et la machinerie de réparation de l'ADN.

Nous avons donc entrepris une caractérisation structurale et fonctionnelle de ces interactions. Nous avons découvert que Rad2 et Rad4 possèdent un motif d'interaction en nous basant sur d'autres interactions de la sous-unité Tfb1 de TFIIH. Par calorimétrie à titrage isotherme, nous avons observé que les segments de ces deux protéines contenant ce motif interagissent avec une grande affinité au domaine PH de Tfb1. Le site de liaison de ces segments sur Tfb1PH est très semblable au site de liaison du domaine de transactivation de p53 et au domaine carboxy-terminal de TFIIIE α avec Tfb1PH, tel que démontré par résonance magnétique nucléaire (RMN). De plus, tous ces segments peuvent faire compétition les uns aux autres pour la liaison à Tfb1PH. Nous avons aussi démontré *in vivo* chez la levure qu'une délétion de Tfb1PH crée une sensibilité aux radiations UV. De plus, la délétion de multiples segments de Rad2 et Rad4, dont les segments d'interaction à Tfb1PH, est nécessaire pour voir une sensibilité aux rayons UV. Ainsi, de multiples interactions sont impliquées dans la liaison de Rad2 et Rad4 à TFIIH. Finalement, les structures des complexes Rad2-Tfb1PH et Rad4-Tfb1PH ont été résolues par RMN. Ces structures sont identiques entre elles et impliquent des résidus hydrophobes interagissant

avec des cavités peu profondes de Tfb1PH. Ces structures sont très semblables à la structure de TFIIIE α -p62PH.

Ces découvertes fournissent ainsi un lien important entre la transcription et la réparation de l'ADN. De plus, elles permettent d'émettre un modèle du mécanisme de déplacement de XPC/Rad4 par XPG/Rad2 au site de dommage à l'ADN. Ces connaissances aident à mieux comprendre les mécanismes de maintien de la stabilité génomique et peuvent ainsi mener à développer de nouvelles thérapies contre le cancer.

Mots clés :

Réparation de l'ADN par excision des nucléotides, XPC/Rad4, XPG/Rad2, TFIIH, interaction protéine-protéine, résonance magnétique nucléaire, titrage calorimétrique isotherme

Abstract

The nucleotide excision repair pathway (NER) is a mechanism capable of removing a wide variety of helix-distorting lesions, such as those caused by ultraviolet irradiation (UV). As all DNA repair pathways, NER contributes to the prevention of carcinogenesis by preventing DNA mutation. During this process, the lesion is first recognized by the protein XPC/Rad4 (human/yeast), which then recruits TFIIH. This complex unwinds the DNA with its helicase activity and then recruits the endonuclease XPG/Rad2 and other proteins necessary for DNA excision. Upon arrival at the lesion site, XPG/Rad2 displaces XPC/Rad4.

TFIIH also acts in DNA transcription, using its helicase activity. In addition to the similarity of the presence of TFIIH in transcription and DNA repair, it is possible to ask ourselves how the two pathways are similar. We were interested in the interactions involving TFIIH and the DNA repair machinery.

We have therefore undertaken a structural and functional characterization of these interactions. We have found that Rad2 and Rad4 have a motif of interaction based on other interactions of the Tfb1 subunit of TFIIH. Using isothermal titration calorimetry, we found that segments of these two proteins containing this motif interact with high affinity to the PH domain of Tfb1. The binding site of these segments is very similar to Tfb1PH binding site of transactivation domain of p53 and the carboxyl-terminal domain of TFIIE α with Tfb1PH, as demonstrated by nuclear magnetic resonance (NMR). In addition, these segments can compete with each other for binding to Tfb1PH. We also demonstrated *in vivo* that deletion of Tfb1PH in yeast creates a sensitivity to UV irradiation. In addition, the deletion of multiple segments of Rad2 and Rad4, including segments of interaction Tfb1PH, is required to observe a sensitivity to UV. Thus, multiple interactions are involved in the binding of TFIIH to Rad2 and Rad4. Finally, the structures of the Rad2-Tfb1PH and Rad4-Tfb1PH complexes were solved by NMR. These structures are identical to each other and involve hydrophobic residues interacting with shallow grooves on Tfb1PH. These structures are very similar to the structure of TFIIE α -p62PH.

These findings provide an important mechanistic link between transcription and DNA repair. In addition, they provide a model of the mechanism of the displacement of XPC/Rad4 by XPG/Rad2 at the damaged site. This knowledge helps to better understand the mechanisms of genomic stability and can lead to novel cancer therapies.

Keywords :

Nucleotide excision repair of DNA, XPC/Rad4, XPG/Rad2, TFIIH, protein-protein interaction, nuclear magnetic resonance, isothermal titration calorimetry

Table des matières

Résumé.....	iii
Abstract.....	v
Table des matières.....	vii
Liste des tableaux.....	x
Liste des figures.....	xi
Liste des abbréviations.....	xiv
Remerciements.....	xix
1.0 Introduction.....	1
1.1 Dommages à l'ADN et réparation.....	3
1.1.1 Types de dommages.....	3
1.1.2 Réactions cellulaires aux dommages à l'ADN.....	4
1.1.3 Mécanismes de réparation de l'ADN.....	5
1.1.4 Noms et conventions.....	6
1.2 Réparation de l'ADN par excision des nucléotides.....	8
1.2.1 Survol de la voie.....	8
Description générale.....	9
TC-NER.....	11
GG-NER.....	12
1.2.2 Détection du dommage à l'ADN par XPC/Rad4.....	13
Séquences.....	13
Fonctions biologiques.....	15
Structure et mécanisme de reconnaissance.....	16
1.2.3 TFIIH lors de la NER.....	18
Sous-unités de TFIIH et leurs fonctions.....	18
Sous-unité p62/Tfb1.....	20
1.2.4 Endonucléase XPG/Rad2.....	23
Séquence.....	23
Fonctions biologiques.....	24
Structure et catalyse.....	26

1.2.5	Syndromes et mutations associées à la NER	28
	Xeroderma pigmentosum.....	28
	Syndrome de Cockayne	28
	Trichothiodystrophie.....	29
	Mutations génétiques	29
1.3	TFIIH : pont entre la transcription et la NER	31
1.3.1	Rôles dans la transcription de l'ADN.....	31
1.3.2	Recrutement lors de la transcription basale	33
1.3.3	Recrutement lors de la transcription activée	34
1.3.4	Interactions protéine-protéine de TFIIH lors de la NER.....	37
	Interaction entre TFIIH et XPC/Rad4.....	37
	Interaction entre TFIIH et XPG/Rad2.....	39
1.4	Objectifs de la thèse	43
1.5	Techniques utilisées	44
1.5.1	RMN.....	44
	Phénomène de résonance magnétique.....	45
	Expérience 1D.....	46
	Expériences multidimensionnelles.....	47
	Expériences d'attribution.....	48
	Expériences NOESY	50
	Calculs de structures	50
1.5.2	ITC	52
2.0	Caractérisation de l'interaction entre Rad2 et Tfb1	55
	ABSTRACT.....	58
	INTRODUCTION	59
	RESULTS	62
	DISCUSSION	68
	METHODS	71
	ACKNOWLEDGEMENTS.....	75
	FIGURES.....	76
	REFERENCES.....	84

SUPPLEMENTARY METHODS	89
SUPPLEMENTARY TABLES	92
SUPPLEMENTARY FIGURES.....	94
3.0 Caractérisation de l'interaction entre Rad4 et Tfb1	99
ABSTRACT	102
INTRODUCTION.....	103
RESULTS	105
DISCUSSION	111
METHODS	114
ACKNOWLEDGEMENTS	117
FIGURES	118
REFERENCES.....	125
SUPPLEMENTARY METHODS	130
SUPPLEMENTARY TABLES	132
SUPPLEMENTARY FIGURES.....	133
4.0 Discussion et conclusion	137
Interactions multiples	139
Un modèle de transition pour la NER.....	140
Régulation des interactions	142
Comparaison des complexes formés entre p62/Tfb1 et ses partenaires.....	143
Désordre intrinsèque	143
Importance de l'étude et conclusions.....	145
Bibliographie.....	147

Liste des tableaux

Chapitre 1 :

Tableau 1. Expériences permettant l'attribution des déplacements chimiques des atomes d'une protéine.	49
--	----

Chapitre 2 :

Table 1. Structural Statistics for Rad2 in Complex with Tfb1PH.	83
Supplemental Table S1: Yeast strains used in this study.....	92
Supplemental Table S2 Primers used in this study	93

Chapitre 3 :

Table 1. Structural Statistics for Rad4 in Complex with Tfb1PH.	124
Supplemental Table S1: Yeast strains used in this study.....	132
Supplemental Table S2 Primers used in this study	132

Liste des figures

Chapitre 1 :

Figure 1-1. Résumé des dommages à l'ADN et des voies de réparations	4
Figure 1-2. Structure chimique de la cisplatine	8
Figure 1-3. Substrats de la NER.....	9
Figure 1-4. Étapes de la GG-NER.....	11
Figure 1-5. Structures primaires de Rad4, XPC, Rad23, RAD23A et Rad24B.....	14
Figure 1-6. Structure tridimensionnelle de Rad4 lié à l'ADN endommagé.....	17
Figure 1-7. Sous-unités de TFIIH chez la levure <i>S. cerevisiae</i> et chez l'humain	18
Figure 1-8. Structures globales de TFIIH	19
Figure 1-9. Structures primaires de Tfb1 et p62	21
Figure 1-10. Structures des domaines PH de p62 (A) et de Tfb1 (B).....	21
Figure 1-11. Le domaine PH de Tfb1 possède une face chargée positivement	22
Figure 1-12 Structures primaires des membres de la famille de FEN-1	24
Figure 1-13. Les nucléases de la famille FEN-1 clivent spécifiquement différents substrats	25
Figure 1-14. Structure de FEN-1 liée à l'ADN.....	27
Figure 1-15. Diagramme de Venn représentant les trois maladies liées à la NER: xeroderma pigmentosum (XP), syndrome de Cockayne (CS), trichothiodystrophie (TTD)	29
Figure 1-16. Modèle d'assemblage séquentiel du complexe de pré-initiation.....	32
Figure 1-17. Structures du domaine PH de p62/Tfb1 avec TFIIIE α (A), p53 (B), VP16 (C), EKLF (D)	34
Figure 1-18. Résumé des interactions de diverses troncations de XPC/Rad4 avec TFIIH et ses sous-unités.....	39
Figure 1-19. Résumé des interactions de diverses troncations de Rad2 et XPG avec TFIIH et ses sous-unités.....	42
Figure 1-20. Spectres RMN 1D et 2D.....	47
Figure 1-21. Spectre RMN 3D.....	48
Figure 1-22. Vue schématique d'un micro-calorimètre.....	52
Figure 1-23. Thermogrammes d'expériences d'ITC.....	54

Chapitre 2 :

Figure 1. <i>tfb1ΔPH</i> is sensitive to UV irradiation, but not bleomycin or temperature	76
Figure 2. The Rad2 spacer region contains a high affinity Tfb1PH-binding site between residues 642 and 690.....	77
Figure 3. Rad2 ₆₄₂₋₆₉₀ and p53TAD2 share a common binding site on Tfb1PH	78
Figure 4. NMR structure of the Rad2 ₆₄₂₋₆₉₀ -Tfb1PH complex.....	79
Figure 5. The Rad2 spacer region contains a second high-affinity Tfb1PH-binding site	80
Figure 6. Multiple regions of the Rad2 spacer are required for repair of UV damage	81
Figure 7. The structures of the Rad2 ₆₄₂₋₆₉₀ -Tfb1PH and TFIIE α CTD-p62PH interfaces are remarkably similar	82
Supplementary Figure 1. Tfb1 and Rad2 protein expression levels in yeast.....	94
Supplementary Figure 2. Rad2 ₆₄₂₋₇₆₀ is a disordered peptide and possesses two distinct segments binding to Tfb1PH.....	95
Supplementary Figure 3. Rad2 ₆₄₂₋₆₉₀ binds to Tfb1PH on a similar site as TFIIE α CTD.....	96
Supplementary Figure 4. Tfb1PH binds to Rad2 ₆₄₂₋₆₉₀	97
Supplementary Figure 5. Rad2 ₃₅₉₋₃₈₃ binds to Tfb1PH on a similar site as Rad2 ₆₄₂₋₆₉₀	98

Chapitre 3 :

Figure 1. The N-terminal region of Rad4 contains a high affinity Tfb1PH-binding site between residues 76 and 115..	118
Figure 2. Rad4 ₇₆₋₁₁₅ and Rad2 ₃₅₉₋₃₈₃ share a common binding site on Tfb1PH.	119
Figure 3. Rad34 contains a high affinity Tfb1PH-binding motif.....	120
Figure 4. Two regions of Rad4 are required for survival following UV irradiation.....	121
Figure 5. NMR structure of the Rad4 ₇₆₋₁₁₅ -Tfb1PH complex.....	122
Figure 6. The structures of the Rad2 ₆₄₂₋₆₉₀ -Tfb1PH and Rad4 ₇₆₋₁₁₅ -Tfb1PH interfaces are very similar	123
Supplementary Figure 1. Rad4 ₇₆₋₁₁₅ binds to a similar site of Tfb1PH as Rad2 ₆₄₂₋₆₉₀	133
Supplementary Figure 2. Rad34 ₄₁₋₆₃ binds to a similar site of Tfb1PH as Rad4 ₇₆₋₁₁₅	134

Supplementary Figure 3. Rad4 protein expression levels in yeast.....	135
Supplementary Figure 4. Tfb1PH binds to Rad4 ₇₆₋₁₁₅	136

Chapitre 4 :

Figure 4-1. Modèle « velcro » de transition pour la NER.....	141
--	-----

Liste des abréviations

Il a été choisi pour quelques abréviations d'utiliser l'abréviation de la littérature, en anglais, pour faciliter la lecture.

¹³C: Carbone 13

¹⁵N: Azote 15

6-4PP: pyrimidine (6-4) pyrimidone

A : Alanine

ADN: Acide désoxyribonucléique

Ala : Alanine

Arg : Arginine

ARN : Acide ribonucléique

Asp : Acide aspartique

ATP : Adénosine triphosphate

BER: *Base Excision Repair* (Réparation par excision des bases)

BHD: *Beta-Hairpin Domain* (Domaine en épingle bêta)

BSD: *BTF2-Like transcription factors, Synapse-associated and DOS2-like protein*

C : Cystéine

CAK: *cyclin-dependent kinase-activating kinase* (kinase activant les kinases d'une façon dépendante de la cycline)

CBP/p300 : *CREB binding protein* (protéine liant CREB)

CDK: *Cyclin-dependent kinase* (kinase dépendante de la cycline)

CPD: *Cyclobutane pyrimidine dimer* (Dimère cyclobutylique de pyrimidines)

CREB : *c-AMP response element-binding protein*

CS: Syndrome de Cockayne

CSA: Syndrome de Cockayne du groupe de complémentation A

CSB: Syndrome de Cockayne du groupe de complémentation B

CTD: *Carboxy-terminal domain* (domaine carboxy-terminal)

D: Acide aspartique

E: Acide glutamique

EKLF: *Erythroid Krüppel-like Factor*

ERCC: *Excision Repair Cross-Complementing* (complémentation croisée de la réparation par excision)

F: Phénylalanine

FEN-1: *Flap endonuclease 1*

G: Glycine

GG-NER: *global-genomic NER* (NER génomique globale)

Gln: Glutamine

Glu: Acide glutamique

HSQC: *Heteronuclear Single Quantum Coherence*

Hz : Hertz

I: Isoleucine

IDP: *Intrinsically Disordered Protein* (protéine intrinsèquement désordonnée)

Ile: Isoleucine

ITC: Isothermal titration calorimetry

J: Joule
 J/m² : Joule par mètre carré
 kDa: kilo Dalton
 L: Leucine
 Leu: Leucine
 Lys: Lysine
 μM: Micromolaire
 M: Méthionine
 Met: Méthionine
 mM: Millimolaire
 MMR: *Mismatch repair* (réparation des mésappariements)
 N: Asparagine
 NER: *Nucleotide excision repair* (réparation par excision de nucleotides)
 NHEJ: *Non-homologous end joining* (jonction de brins non-homologues)
 NLS: *Nuclear localization signal* (séquences de localisation nucléaire)
 nm: nanomètre
 NOE: *Nuclear Overhauser effect* (effet Overhauser nucléaire)
 P: Proline
 PCNA: *Proliferation cell nuclear antigen* (antigène nucléaire de prolifération cellulaire)
 PDB: *Protein data bank*
 PH: *Pleckstrin homology* (homologie à la pleckstrine)
 Phe: Phénylalanine
 PIC: *Pre-initiation complex* (complexe de pré-initiation)
 ppm: Partie par million
 Pro: Proline
 Q: Glutamine
 R: Arginine
 RFC: *Replication Factor C* (facteur de réplication C)
 RMN: Résonance Magnétique Nucléaire
 RMSD: *Root Mean Square Deviation* (racine de la moyenne des déviations au carré)
 RPA: *Replication protein A* (Protéine de réplication A)
 S: Sérine
 SAXS: *Small angle X-ray scattering* (diffusion des rayons X aux petits angles)
 Ser: Sérine
 SWI/SNF: *Switching mating type/sucrose non fermenting*
 T: Thréonine
 TAD: *Transactivation domain* (domaine de transactivation)
 TBP: *TATA box-Binding Protein* (protéine liant la boîte TATA)
 TC-NER: *transcription-coupled NER* (NER couplée à la transcription)
 TFII: *ARN polymerase II Transcription Factor* (Facteur de transcription de l'ARN polymérase II)
 Tfb1: Transcription factor b 1
 Tfb1PH: Domaine PH de Tfb1
 TGD: *Transglutaminase-like domain* (domaine semblable à la transglutaminase)
 Thr: Thréonine
 Trp: Tryptophane

TTD: Trichothiodystrophie

UV: Ultraviolet

V: Valine

Val: Valine

VP16: Virion protein 16

W: Tryptophane

XP: Xeroderma pigmentosum

XPA: Xeroderma pigmentosum de groupe de complémentation A

XPB: Xeroderma pigmentosum de groupe de complémentation B

XPC: Xeroderma pigmentosum de groupe de complémentation C

XPD: Xeroderma pigmentosum de groupe de complémentation D

XPE: Xeroderma pigmentosum de groupe de complémentation E

XPF: Xeroderma pigmentosum de groupe de complémentation F

XPG: Xeroderma pigmentosum de groupe de complémentation G

À ma famille

Remerciements

Je tiens à remercier mon directeur de recherche, Dr James G. Omichinski qui m'a encadré durant les quatre dernières années, partageant aussi ses idées, hypothèses et conseils. Avec lui, j'ai pu apprendre la technique de résonance magnétique nucléaire, en plus d'avoir une formation scientifique générale solide. Je le remercie pour sa disponibilité et ses encouragements. Je remercie également mon co-directeur, Dr Jurgen Sygusch, pour m'avoir introduit à la biologie structurale via la cristallographie rayons-X. Il a su développer en moi précision et rigueur. Je suis également reconnaissant pour l'apprentissage de la résonance magnétique nucléaire que m'a offert Dr Pascale Legault. Dans la même veine, Dr Tara Sprules a également participé à cet enseignement en plus de prodiguer d'excellents conseils lors de la collecte de données.

Les expériences *in vivo* de cette thèse ont été réalisées par Geneviève Arseneault qui a fait un travail extraordinaire dont je suis très reconnaissant. Je dois également remercier les stagiaires qui ont contribué aux travaux : Maryse Lefebvre, Thomas Morse, Aurélie Bernier, Joannie Guillemette, Antoine Chiasson, Emmeline Ruka et Mara Buruian. Merci à Laurent Cappadocia qui a apporté sa contribution à la fin du projet, tout en relisant attentivement cette thèse. Je remercie aussi plusieurs étudiants du département qui m'ont côtoyé et aidé dans les dernières années: Mathieu Coincon. Caroline Mas, Luca Raiola, Paola Di Lello, Jimmy Dikeakos ainsi que tous les autres membres des laboratoires Omichinski/Legault. Je remercie l'ensemble du personnel de soutien du département de biochimie sans qui la recherche ne serait possible. Je tiens aussi à exprimer ma gratitude pour Dr Elliott Drobetsky. Je l'ai rencontré pour la première fois en 1996 dans le cadre d'une recherche en biologie, au secondaire. Il a ensuite été jury à mon examen de synthèse doctoral. Qui aurait su que ma thèse de Doctorat porterait sur un de ses sujets de recherche?

Merci mille fois à mes amis pour leur soutien des dernières années. Philippe, Mathieu, Benjamin, Erik, Eric, Richard, Bertrand, Ariane, Jana, Rym, Olivier, les Tanguy et Helen. Finalement, merci infiniment à ma famille, mes parents et mon frère pour leur soutien constant toutes ces années.

1.0 Introduction

Le dogme central de la biologie moléculaire, consiste en une formule simple: ADN → ARN → protéines (1). Occasionnellement, des erreurs peuvent s'immiscer lors de la transcription de l'ADN en ARN, tout comme lors de la traduction de l'ARN en protéine. Toutefois, parce qu'elles sont occasionnelles, ces erreurs sont diluées dans l'ensemble des protéines produites. Par contre, si ces erreurs deviennent plus fréquentes, il peut y avoir des conséquences graves pour la cellule. Or, si le premier maillon de la chaîne, l'ADN, est muté d'une façon permanente, toutes les protéines produites par le gène muté seront affectées. Ces mutations sont causées par des dommages à l'ADN n'ayant pas été réparés: des mutations sont alors insérées par des mécanismes de translesion. Ceci peut ultimement mener à des maladies comme le cancer. La réparation de l'ADN est ainsi un mécanisme primordial de maintien de l'intégrité génomique.

Les voies de réparation et de transcription de l'ADN sont très bien conservées chez toutes les espèces vivantes et spécialement chez les eucaryotes. Les protéines et leurs fonctions sont également conservées transversalement, dans ces deux processus. La présente thèse veut apporter une nouvelle vision de ces similarités. Le but principal sera de faire le pont, non seulement fonctionnel, mais surtout au niveau du mécanisme moléculaire, entre la transcription et la réparation de l'ADN.

Après un bref survol des types de dommage à l'ADN et des voies de réparation, cette thèse approfondira une voie en particulier: la réparation par excision des nucléotides. Ensuite, nous verrons comment nos découvertes permettent d'élucider certains mécanismes d'interaction de protéines nécessaires à cette voie. Finalement, nous terminerons par une discussion générale sur les implications biologiques de ces découvertes.

1.1 Dommages à l'ADN et réparation

L'ADN des cellules est constamment endommagé par des agents chimiques et physiques. Ceci peut compromettre l'intégrité génomique et de surcroît la survie cellulaire. Heureusement, il existe des réponses cellulaires aux dommages à l'ADN, incluant sa réparation (2). Avant d'entrer au cœur du sujet de cette thèse, voici un bref survol des dommages pouvant être faits à l'ADN, des réactions cellulaires à ces dommages ainsi que des divers mécanismes de réparation de l'ADN (voir Figure 1-1 pour un résumé visuel de cette section).

1.1.1 Types de dommages

Malgré sa grande stabilité, l'ADN peut être endommagé d'une multitude de façons. Ces dommages peuvent être divisés en quatre grandes catégories. Une première catégorie consiste en des modifications effectuées à la base azotée des nucléotides. Ceci peut être le résultat des radiations retirant des groupes chimiques ou même des agents alkylants qui ajoutent des groupements alkyles à la base azotée. Les espèces oxygénés réactives telles que l'anion superoxyde et le peroxyde d'hydrogène, sous-produits de l'activité cellulaire, oxydent l'ADN (3). Les bases peuvent même être altérées d'une façon spontanée. La coupure entre du lien base-sucre affecte 18000 purines par cellule par jour (4). La déamination des bases entraîne aussi des modifications importantes, par exemple la cytosine peut être convertie en uracile (5). Une seconde catégorie consiste en les erreurs causées par une mauvaise réplification de l'ADN (6). Il peut y avoir insertion d'une mauvaise base causant un mésappariement, une insertion de nucléotides supplémentaires ou même un retrait de nucléotides. Ces erreurs sont le résultat d'un glissement de l'ADN polymérase sur l'ADN (7). Une troisième catégorie consiste en la réticulation (*cross-link*) entre les bases du même brin (sur des bases consécutives) ou entre les bases de brins complémentaires. Ceci a le même effet que des agents intercalants qui, sans nécessairement créer des liaisons covalentes, interagissent suffisamment pour créer des déformations de l'ADN et empêcher la machinerie transcriptionnelle ou réplificative d'effectuer leur travail. Ces dommages peuvent être créés par des agents chimiques (cisplatine) ou physiques. Par exemple, les rayons ultraviolets (UV) causent les types de dommages étudiés dans la

présente thèse. Finalement, les radiations ionisantes et certains agents comme la bléomycine créent des brisures double brin (8). Les liens phosphodiesters sont alors coupés sur les deux brins de l'ADN. Ceci peut également arriver sur un seul brin. Aussi, un effondrement de la fourche de réplication peut causer les brisures double brin.

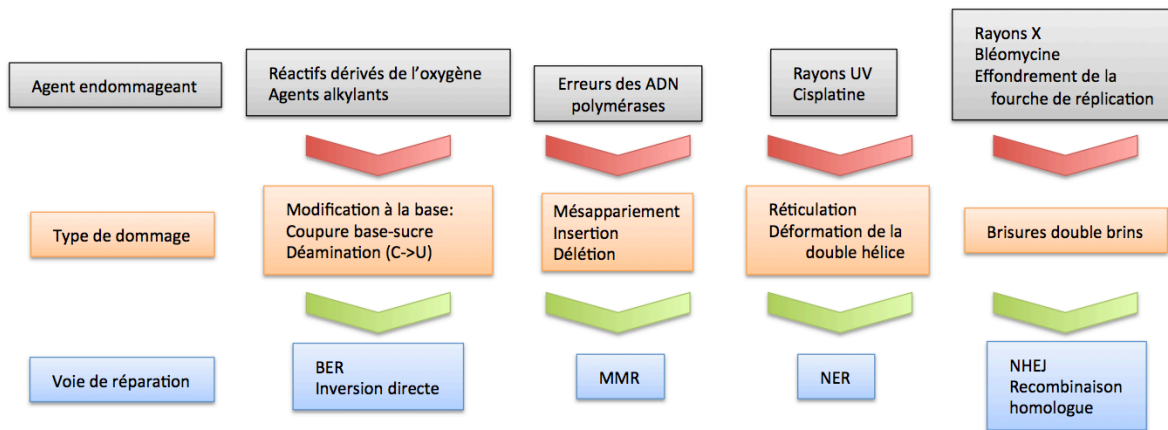


Figure 1-1. Résumé des dommages à l'ADN et des voies de réparations

1.1.2 Réactions cellulaires aux dommages à l'ADN

Les dommages à l'ADN peuvent être réparés en tout temps. Ainsi, durant la phase G1 et G2 du cycle cellulaire, certains types de réparation sont effectués. Toutefois, le tout devient plus critique à l'entrée de la phase S. En effet, comme l'ADN est répliqué durant cette phase, l'insertion d'une mutation pourrait s'y produire. Pire encore, si la machinerie de réplication de l'ADN s'arrête de répliquer suite à la rencontre d'un dommage, il peut y avoir effondrement de la fourche de réplication, ce qui crée une brisure double brin. Ces brisures sont particulièrement difficiles à réparer et peuvent mener à des aberrations chromosomiques.

Il existe des points de contrôle (*checkpoint*) contrôlant la progression du cycle cellulaire. Les dommages qui interfèrent avec les processus de transcription et de réplication de l'ADN (9) activent la réponse au dommage à l'ADN, qui est constituée : (i) d'une cascade de phosphorylation pour signaler et amplifier le signal du dommage; (ii) de

l'activation de gènes et de protéines; (iii) de l'arrêt du cycle cellulaire afin de permettre la réparation par différents mécanismes énumérés dans la prochaine section.

Si ces réparations ne se font pas à temps, des mécanismes de translésion permettront à la cellule de pouvoir continuer ses activités. Lorsque la machinerie de réplication rencontre un dommage qu'elle ne peut outrepasser, il y a recrutement de polymérases spécialisées (dites de translésion) qui peuvent continuer la polymérisation de l'ADN malgré la présence de la lésion. Les polymérases η , ι , κ , ζ et Rev1 peuvent être fidèles et insérer les bons nucléotides, mais elles créeront le plus souvent des mutations (10). Si toutefois il y a trop de dommages, les protéines du point de contrôle activeront plutôt l'apoptose, ce qui détruira la cellule, pour le bénéfice de l'organisme.

1.1.3 Mécanismes de réparation de l'ADN

Le premier mécanisme de réparation consiste en l'excision de la partie endommagée (en retirant un ou plusieurs nucléotides consécutifs) suivi de la polymérisation de l'ADN avec le brin non-endommagé comme modèle. La réparation par excision des nucléotides (*Nucleotide Excision Repair*, NER), sujet de cette thèse, est un bon exemple de ce type de réparation où plusieurs nucléotides sont retirés. Lorsqu'une seule base est retirée, on parle plutôt de réparation par excision d'une base (*Base Excision Repair*, BER) (11). Cette voie permet la réparations de bases oxydées et alkylées et comporte fréquemment l'ajout d'une étape initiale de clivage du lien base-sucre par une glycosylase, pour créer un site apurinique ou apyrimidique (sites-AP).

Une autre voie de réparation dont le mécanisme ressemble beaucoup aux voies d'excision est la réparation des bases mal appariées (*Mismatch Repair*, MMR) (6, 12). Une fois le mésappariement causé lors de la réplication détecté, la cellule doit savoir quel est le bon brin d'ADN et quel est le mauvais. Comme cette voie agit souvent en même temps que la réplication, le brin nouvellement synthétisé est dégradé par une exonucléase, puis la synthèse continue.

En troisième lieu, l'ajout d'un agent chimique sur une base peut être simplement retiré. C'est le cas de plusieurs types de réparation par inversion directe du dommage (13). Par exemple, la méthyleguanine méthyletransférase est capable de retirer le méthyle de la

méthyleguanine, pour ainsi retourner à la guanine originelle. Un autre exemple intéressant, qu'on ne retrouve pas chez tous les eucaryotes supérieurs, est celui des photolyases. Ces enzymes peuvent capter des photons de la lumière pour renverser les dommages causés par la lumière elle-même.

Pour terminer, les brisures double brin peuvent être réparées par recombinaison homologue (14). Une séquence d'ADN homologue est alors utilisée comme modèle pour réparer la lésion. La jonction de brins non-homologues (*Non-Homologous End Joining*, NHEJ) est une autre voie permettant la réparation des brisures double brin (8). Deux molécules d'ADN non-homologues (ayant subi une brisure double brin) sont « recollées » ensemble grâce à une ligase. Par contre, ceci demande parfois l'excision des extrémités de ces ADN, ce qui peut entraîner une perte d'information génétique.

1.1.4 Noms et conventions

Tout au long de cette thèse, on rencontrera très souvent des noms similaires pour certains gènes. Chez la levure *Saccharomyces cerevisiae*, ce sont les gènes *RAD* qui sont prédominants. En 1970, une convention a été établie : tous les mutants sensibles aux rayons UV portent les numéros 1 à 49 alors que les numéros 50 et plus sont attribués aux mutants sensibles aux radiations ionisantes, tels que les rayons X (15). Par convention également, le nom du gène est en lettres majuscules et italique (*RAD*), la souche mutante portant ce gène en italique (*rad*) alors que la protéine associée au gène a uniquement sa première lettre en majuscule (Rad) (16).

D'autre part, chez l'humain, on parle plutôt en terme de maladies pour définir le nom des gènes. Par exemple, les gènes *XP* sont porteurs de mutations chez des patients atteints de xeroderma pigmentosum, une maladie génétique d'hypersensibilité aux rayons UV. Des expériences de complémentation par hybridation des cellules provenant de différents patients ont mené à la création de différents groupes de complémentation, qui ont été nommés par les lettres de A à F (17). Si une souche cellulaire peut pallier aux manques d'une autre souche cellulaire lors de leur fusion, ces souches sont considérées comme étant dans des groupes de complémentation différents. Par exemple, les cellules du groupe de complémentation XP-G peuvent compléter les cellules du groupe de complémentation

XP-C. Ce dernier contient des mutations dans le gène *XPC*, qui produit la protéine XPC (par convention le nom du gène est en majuscule-italique et le nom de la protéine en majuscule).

Finalement, on verra parfois des gènes portant le nom *ERCC*. Ceci désigne des gènes qui ont été trouvés chez des rongeurs et qui compléteraient des cellules humaines déficientes en réparation par excision des nucléotides, d'où le nom *Excision Repair Cross-Complementing (ERCC)* (18). Par contre, peu de gènes sont nommés ainsi désormais. Dès que des mutations causant une maladie sont trouvés chez les gènes humains, la nomenclature humaine prédomine (19). Maintenant que nous avons une vision globale de la réparation de l'ADN, explorons la réparation par excision des nucléotides.

1.2 Réparation de l'ADN par excision des nucléotides

1.2.1 Survol de la voie

Les rayons ultraviolets (UV) peuvent créer des dimères de pyrimidine cyclobutane (CPD) et des photoproduits pyrimidine (6-4) pyrimidone ((6-4)PP) (Figure 1-3)). Ces modifications covalentes de l'ADN créent une déformation de la double hélice, ce qui empêche la progression normale des différentes polymérases agissant sur l'ADN (l'ARN polymérase II lors de la transcription et l'ADN polymérase lors de la réplication). Leur réparation se fera surtout grâce à la voie de NER. Ce type de réparation s'occupe également des adduits à l'ADN formés par la cisplatine (Figure 1-2), les carcinogènes polycycliques, les psoralènes et d'autres agents chimiques qui plient légèrement la double hélice d'ADN. Ceci est l'effet de leur présence stériquement encombrante entre les paires de base (20). En somme, il faut une modification de la conformation chimique de l'ADN, mais aussi une modification de la conformation tridimensionnelle de l'ADN pour que la NER ait lieu.

Comme son nom l'indique, la NER retire les nucléotides endommagés (en plus d'autres nucléotides autour de la lésion), puis re-synthétise la portion manquante. Ces mêmes étapes existent chez toutes les espèces vivantes mais le processus chez les eucaryotes est très différent de celui des archées et des bactéries, car il n'emploie pas exactement les mêmes enzymes que ces organismes. Par contre, chez les eucaryotes, la NER est très bien conservée, de la levure à l'humain. Voici donc une description générale de la NER eucaryote (21).

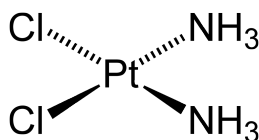


Figure 1-2. Structure chimique de la cisplatine. La cisplatine est un agent utilisé dans la chimiothérapie. Elle se lie à l'ADN, créant ainsi un adduit.

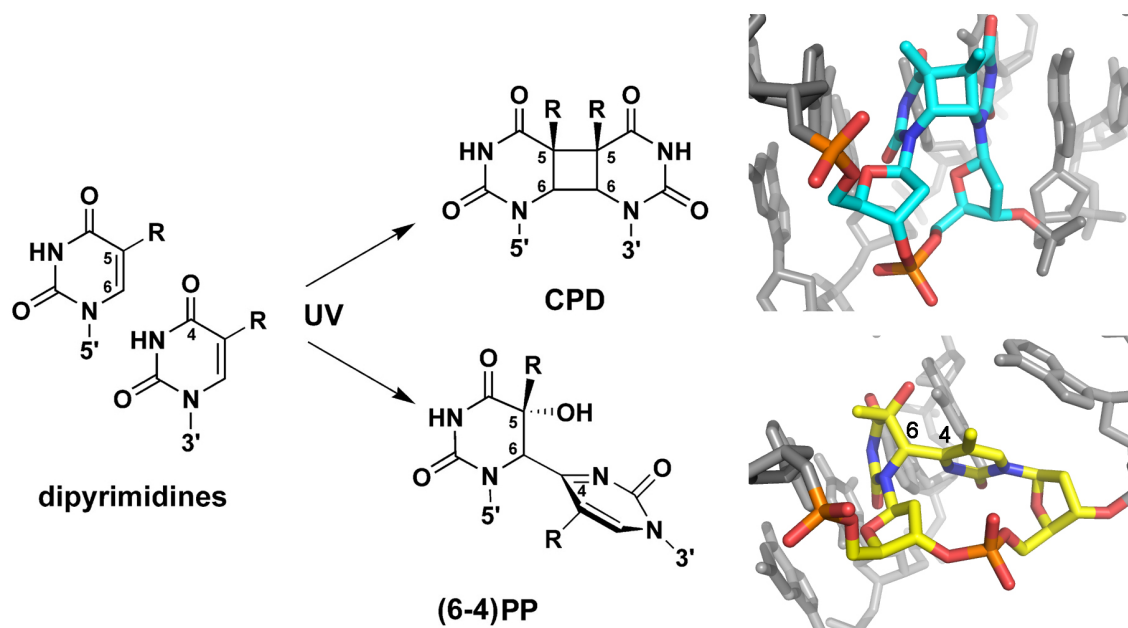


Figure 1-3. Substrats de la NER. Deux pyrimidines adjacentes irradiées par des rayons UV peuvent créer des liens covalents entre leurs bases et former des dimères de pyrimidine. Si deux liens sont créés, il y a formation d'un anneau cyclobutane, un CPD, alors que la création d'un seul lien forme un (6-4)PP. À droite, les structures tridimensionnelles d'un CPD (haut) et d'un (6,4)PP (bas). Ces images de structures tridimensionnelles et toutes les autres de cette thèse ont été créées par l'auteur grâce au logiciel Pymol (22).

Description générale

Il existe deux sous-voies à la NER dont la différence majeure réside dans la première étape de reconnaissance de la lésion. Dans la NER couplée à la transcription (TC-NER), c'est l'ARN polymérase II qui, arrêtée par une lésion lors de l'élongation, est le signal de déclenchement de la voie. Dans la NER génomique globale (GG-NER), c'est XPC/Rad4^a qui reconnaît directement la lésion (23), ce qui entraîne le recrutement des autres protéines en vue de former un complexe de pré-incision (Figure 1-4). Dans certains cas, XPC peut être aidé de DDB1 et DDB2 (*UV DNA-Damage Binding proteins*, aussi appelées XPE), des protéines liant les dommages à l'ADN créés par les rayons UV.

^a Nous emploierons le nom des protéines de levure et d'humain de cette façon, humain/levure, si l'information s'applique pour les protéines orthologues de deux organismes ayant des noms différents.

XPC/Rad4 recrute ensuite le complexe TFIIH qui possède une activité hélicase-ATPase déroulant l'ADN autour de la lésion (24). L'hélicase XPB/Ssl2 de TFIIH agit en 5' de la lésion alors que l'hélicase XPD/Rad3 agit en 3' de la lésion. XPG/Rad2 se joint alors à TFIIH, déplaçant XPC/Rad4 du complexe de pré-incision (25, 26). XPG/Rad2 est une endonucléase et est donc positionnée en 3' de la lésion où elle fera ultérieurement une coupure de l'ADN (27). Mais avant, RPA s'ajoute au complexe. Cette protéine est connue pour lier et protéger l'ADN simple brin dans plusieurs processus de réparation et de réplication de l'ADN. Son rôle dans la NER est donc de lier l'ADN simple brin non-endommagé. L'ADN endommagé est quant à lui lié par XPA/Rad14, qui s'ajoute au complexe en interagissant avec les protéines présentes (28). En plus de vérifier la lésion avant que les nucléases coupent l'ADN, XPA/Rad14 positionne les nucléases au bon endroit (29). Finalement, XPF/Rad1 et son partenaire ERCC1/Rad10 viennent s'associer en 5' de la lésion, pour compléter le complexe de pré-incision.

L'ordre des étapes subséquentes est débattu, mais un modèle actuel veut que l'incision se fasse d'abord par XPF, de 15 à 24 nucléotides en 5' de la lésion (30). Plusieurs protéines (TFIIH, XPA/Rad14, XPF/Rad1, ERCC1/Rad10) quittent alors l'ADN pour laisser la place à la polymérase. Celle-ci doit être hautement fidèle pour éviter les erreurs. Ainsi, la polymérase δ , la polymérase ϵ ou la polymérase κ sont utilisées (31, 32), accompagnées de PCNA, une protéine homo-trimérique reconnue pour entourer l'ADN afin de stabiliser les polymérases sur la double hélice. PCNA doit être chargée sur l'ADN par RFC (33). La polymérase re-synthétise l'ADN (31), déplaçant RPA du même coup. XPG fera son incision, de 2 à 8 nucléotides en 3' de la lésion, retirant ainsi un brin de 24 à 32 nucléotides (34) et se décrochant de l'ADN. Finalement, la polymérase termine son travail et la ligase I ou III vient relier l'ancien brin au nouveau (35). À l'exception de l'étape initiale de reconnaissance de l'ADN endommagé, toutes ces étapes sont les mêmes pour la GG-NER et la TC-NER.

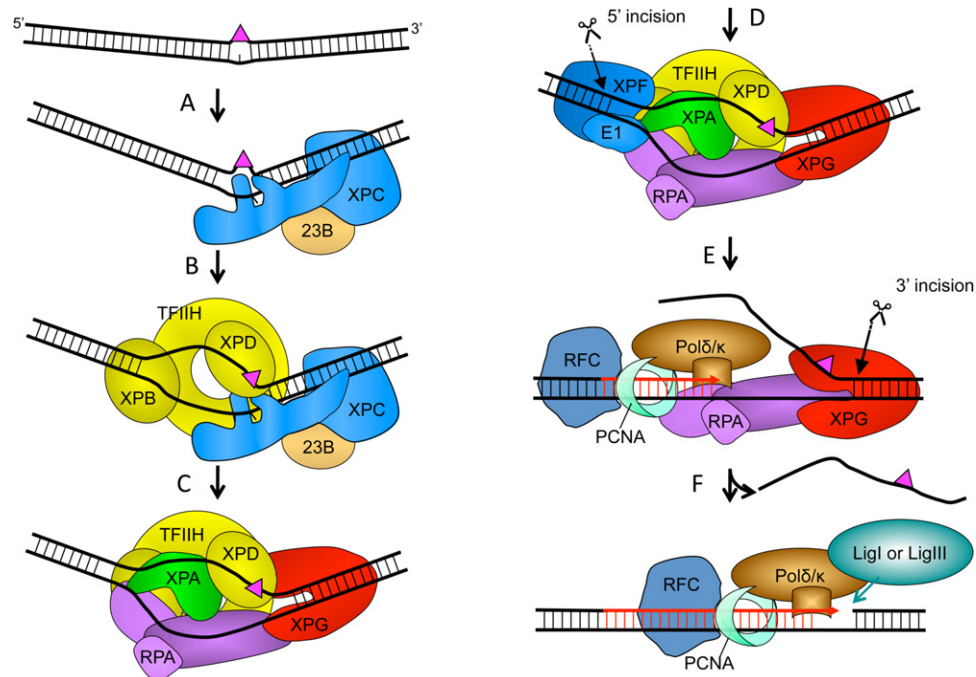


Figure 1-4. Étapes de la GG-NER. A) Reconnaissance de la lésion par XPC/Rad4 (23B représente la sous-unité RAD23B/Rad23). B) Recrutement de TFIIH et ouverture de la double hélice par les hélicases XPB/Ssl2 et XPD/Rad3. C) Recrutement de XPG/Rad2 à la lésion et déplacement de XPC/Rad4. Recrutement de XPA/Rad14 et RPA. D) Recrutement de XPF/Rad1, la nucléase en 5' et son partenaire ERCC1/Rad10. Le complexe de pré-incision est formé, une première incision (en 5') est effectuée. E) La nucléase XPF/Rad1, TFIIH et XPA laissent la place à la polymérase δ , ϵ ou κ , accompagnée de PCNA et de RFC. La seconde incision (en 3') est effectuée. F) L'ADN endommagé est retiré et la polymérase remplit le vide. La ligase I ou III vient lier l'ancien brin au nouveau. (Tirée de (36)).

TC-NER

L'ADN endommagé peut se trouver dans plusieurs états: enroulé dans la chromatine, accessible, en pleine réplication ou même en pleine transcription. Dans ce dernier cas, c'est la voie de la TC-NER qui est prépondérante. Lorsque l'ARN polymérase II transcrit un gène et qu'elle rencontre une lésion, elle est bloquée et ne peut continuer. Son arrêt est alors détecté par CSB/Rad26 qui procède au recrutement de TFIIH (37). Celui-ci déplace les protéines en place pour ensuite recruter XPG/Rad2 et continuer la voie de NER de la même façon que décrite précédemment. D'autres protéines comme CSA/Rad28 sont également nécessaires à la TC-NER, mais leur fonction précise n'est pas connue (38).

GG-NER

Dans tous les états de l'ADN autres que transcrits (enroulé dans la chromatine, libre, etc.), on parle plutôt de GG-NER. Bien que la protéine XPC, aidée par sa sous-unité RAD23B, puisse reconnaître à elle seule, elle est parfois aidée les UV-DDBs pour certains types de lésions, ce qui augmente l'efficacité de réparation. En effet, XPC/Rad4 a une affinité plus grande pour les (6-4)PP comparativement aux CPD (39, 40). Or, DDB1 et DDB2 lient fortement les CPD et aident XPC à les retrouver (41). De plus, les protéines DDB1, DDB2, culline 4A et Rbx1 forment une complexe ubiquitine ligase (E3) qui peuvent ubiquitiner XPC. Ceci ne l'envoie pas vers la dégradation par le protéasome, mais augmente plutôt son affinité pour l'ADN endommagé (42). Aucune protéine orthologue à DDB1 et DDB2 n'a été découverte chez la levure, mais les protéines Rad7 et Rad16 auraient des fonctions similaires (dont celles de E3 ligase et de liaison à l'ADN endommagé) (43, 44).

Cette fonction ubiquitine ligase est également importante lorsque les lésions sont enfouies dans la chromatine. En effet, les lésions à l'ADN sont souvent situées dans de l'ADN non-transcrit et difficilement accessible. Pour que la réparation ait lieu, il faut donc y avoir accès en remodelant la chromatine. Le complexe des protéines DDB effectue en partie cette fonction par ubiquitination des histones H2A, H3 et H4 (45, 46). Toujours chez les mammifères, des liens fonctionnels ont été découverts entre SWI/SNF et la NER (47), sans que des interactions directes soient prouvées. Ce complexe réarrange la chromatine suite à une irradiation UV et aide ainsi à augmenter l'efficacité de la NER. Chez la levure, l'existence d'une interaction directe entre Rad4-Rad23 et Snf6-Snf5, deux protéines du complexe de remodelage de la chromatine SWI/SNF, a par contre été démontrée (48).

Dans cette thèse, nous référerons à l'acronyme «NER» en parlant surtout de GG-NER, pour des raisons de simplicité, mais aussi parce que nous nous concentrerons sur la reconnaissance de la lésion avec XPC/Rad4. Toutefois, suite à cette étape, notre utilisation de l'acronyme NER pourrait s'appliquer autant à la GG-NER qu'à la TC-NER. Nous accorderons une attention particulière aux protéines XPC/Rad4, TFIIH et XPG/Rad2. Voyons maintenant plus en détails cette première étape de reconnaissance par XPC/Rad4.

1.2.2 Détection du dommage à l'ADN par XPC/Rad4

La détection d'un dommage à l'ADN est une étape cruciale dans le processus de réparation. Il faut pouvoir détecter toutes les lésions présentes, tout en évitant de réparer l'ADN non-endommagé. Ceci augmenterait le risque d'effectuer des erreurs qui pourraient devenir des mutations. Également, la détection est importante, car c'est elle qui enclenche la série d'évènements menant à la réparation.

Cette importance est démontrée par la prépondérance du groupe de complémentation C de xeroderma pigmentosum qui constitue au-delà de 25% des cas de XP mondialement (49). Ce sont des mutations dans le gène *XPC* qui causent cette version de la maladie. C'est en cherchant un gène capable de réhabiliter des cellules XP-C que *XPC* a été découvert et cloné (50, 51). Par homologie de séquence, il a alors été relié à *RAD4*, découvert chez la levure beaucoup plus tôt (52, 53). Par la suite, on a relié la reconnaissance du dommage à l'ADN à la protéine XPC/Rad4, devenant ainsi l'initiatrice de la NER (23).

Par contre, XPC/Rad4 n'agit pas seul dans la reconnaissance. La protéine Rad23 a été associée à Rad4. Puis, par homologie de séquence, on a découvert RAD23A et RAD23B chez les mammifères (51). C'est RAD23B qui est en prépondérance dans la cellule. Un troisième partenaire s'ajoute à ce dimère. Chez les mammifères, il s'agit de la centrine 2 (54, 55) alors que chez la levure il s'agit de Rad33 ou Cdc31 (56-58). Décortiquons les séquences, le fonctionnement et la structure de ce complexe protéique.

Séquences

Rad4 est une protéine constituée de 754 acides aminés alors que XPC en possède 940. Cette différence d'un peu moins de 186 acides aminés pourrait expliquer certaines fonctions différentes chez les deux protéines. Chose certaine, plusieurs domaines ont une grande similarité de séquence. En se basant sur la structure tridimensionnelle de Rad4 et sur l'alignement des séquences de Rad4 et de XPC, il est possible d'identifier et de séparer ces domaines (Figure 1-5, voir structure Figure 1-6). D'abord, un segment amino-terminal de 122 chez Rad4 et de 172 acides aminés chez XPC sont prédits comme étant désordonnés et est sensibles à la protéolyse (59). Ensuite, un domaine structurellement similaire à la

transglutaminase (TGD) existe chez les deux protéines, mais est séparé par un segment de 142 acides aminés de fonction inconnue chez XPC. Viennent ensuite trois domaines de liaison à l'ADN qui présentent un repliement en épingle- β (*Beta-Hairpin Domain*, BHD). Ceux-ci sont très bien conservés chez les deux organismes. XPC possède également une séquence de liaison pour la centrine 2 d'une quinzaine d'acides aminés dont la structure hélicale a été résolue (60-62). Pour Rad4, ce même site, dont les acides aminés sont conservés, lie Rad33 (57). Les segments carboxy-terminaux des deux protéines sont prédits comme étant désordonnés, sont sensibles à la protéolyse et ont des séquences différentes (59, 62). Finalement, Rad23 et ses orthologues interagissent avec XPC/Rad4 via le TGD.

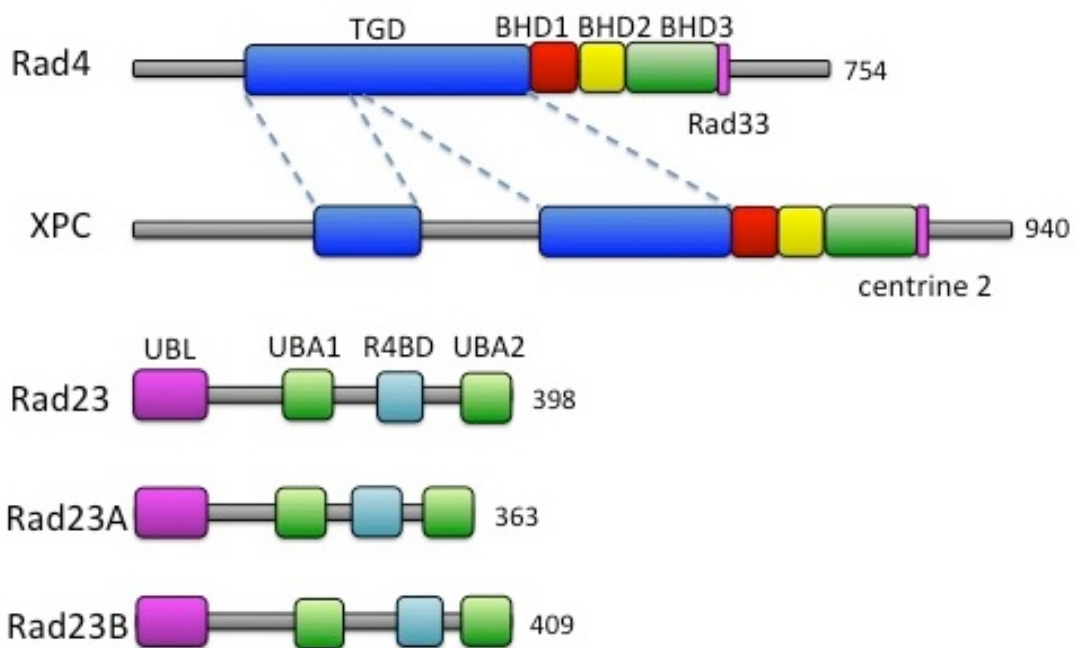


Figure 1-5. Structures primaires de Rad4, XPC, Rad23, RAD23A et Rad24B. Les segments n'ayant pas de domaine précis sont dessinés d'un trait gris. Rad4 et XPC ont des séquences différentes, mais il est possible de reconnaître un domaine ressemblant à la transglutaminase (TGD, bleu) et trois domaines de liaison à l'ADN ayant la forme d'épingles bêta (en rouge, jaune, vert pour BHD1, 2 et 3). Les résidus conservés du TGD de Rad4 sont séparés en deux parties chez XPC, tel que montré par les lignes pointillées. Le segment d'XPC liant la centrine 2 et le segment homologue de Rad4 liant Rad33 sont en magenta. Rad23 et ses orthologues RAD23A et RAD23B ont un domaine semblable à l'ubiquitine (UBL, magenta), deux domaines associés à l'ubiquitine (UBA1 et UBA2, vert) ainsi qu'un domaine liant Rad4 (R4BD, cyan).

Contrairement à XPC/Rad4, les protéines Rad23, RAD23A et RAD23B sont très semblables et possèdent la même organisation de leurs domaines. Elles débutent par un domaine similaire à l'ubiquitine (*Ubiquitin-Like Domain*, UBL) interagissant avec le protéasome 26S (63, 64). Ces protéines peuvent lier les chaînes d'ubiquitines via deux domaines de liaison à l'ubiquitine (*Ubiquitin-Associated Domain*, UBA) (65). Ce sont les segments variables entre les domaines qui créent une différence dans la longueur de ces protéines, allant de 363 à 409 acides aminés.

Fonctions biologiques

Le complexe de XPC/Rad4 reconnaît l'ADN plié et non la lésion en tant que telle (66, 67). Ceci explique sa faible spécificité et la diversité des lésions reconnues, détectant les lésions créées par les UV (CPD et (6,4)PP) ainsi que les gros adduits à l'ADN (68). Après la liaison à l'ADN, XPC/Rad4 recrute TFIIH au site du dommage à l'ADN (69, 70). Malgré la simplicité de ce modèle, plusieurs subtilités existent.

Le rôle précis de Rad23 et de ses orthologues n'est pas bien connu, mais ces protéines favorisent la stabilité de XPC/Rad4 (71) et stimulent la NER, sans être nécessaire à ce processus (72). Chez l'humain, RAD23A et RAD23B ont des fonctions identiques dans la NER (73). Étant donné qu'il y a de 10 à 20 fois plus de RAD23A/B que de XPC (74), il est possible de supposer que les protéines qui ne sont pas en dimère avec XPC jouent d'autres rôles dans la cellule. En effet, l'interaction du domaine UBL avec le protéasome 26S lie Rad23 à la dégradation des protéines, sans qu'on n'en connaisse le mécanisme précis (63, 64). Paradoxalement, Rad23 protège XPC/Rad4 de la protéolyse (75). Les paradoxes se multiplient lorsqu'on tente de déterminer le rôle précis des domaines UBA de Rad23, capables de lier les chaînes de poly-ubiquitine liées par leur lysine 48 (76-80). Tout ceci laisse croire que Rad23 fait partie d'un mécanisme complexe de contrôle du niveau cellulaire de XPC/Rad4, mais probablement aussi d'autres protéines.

Tout comme Rad23, la centrine 2 augmentent grandement l'efficacité de la NER, sans lui être essentielle (81). Son interaction avec XPC a été bien caractérisée et son segment d'interaction sur XPC chevauche le segment d'interaction de XPC avec TFIIH. Ceci suggère qu'il existe d'autres sites d'interaction de TFIIH sur XPC/Rad4. Aucun

orthologue à la centrine 2 n'a été trouvé à ce jour chez la levure. Par contre, la protéine Rad33 joue un rôle similaire et interagit avec Rad4 sur la même séquence d'acides aminés (conservée) que la centrine 2 sur XPC, ce qui laisse croire à des fonctions semblables, mais sans homologie de séquence d'acides aminés (57).

Bien que la voie de NER soit bien conservée chez les eucaryotes, il existe des différences importantes entre le fonctionnement de Rad4 et de XPC. D'abord, Rad4 agit à la fois dans l'ADN non-transcrit et dans l'ADN transcrit (82). Ceci contraste avec XPC qui n'est nécessaire que dans les segments non-transcrits (83). De plus, il existe un équivalent de Rad4, ne fonctionnant que pour la réparation de l'ADN ribosomal, Rad34 (84). L'ADN ribosomal étant presque toujours transcrit, Rad34 agirait surtout dans la NER couplée à la transcription (85). Il est à noter que l'ADN ribosomal est transcrit par l'ARN polymérase I, ce qui permet d'expliquer en partie pourquoi il existe une autre protéine pour détecter spécifiquement cet ADN. Il se peut qu'un équivalent de Rad34 existe également chez l'humain, mais qu'il n'ait pas été découvert à ce jour.

Structure et mécanisme de reconnaissance

La structure de Rad4-Rad23 a été résolue par cristallographie rayons-X, dans sa forme libre et liée à l'ADN (59) (Figure 1-6). Il est important de noter que cette structure omet plusieurs segments ayant été prédits comme désordonnés et étant sensibles à une protéolyse limitée. Les 100 résidus amino-terminaux ainsi que les 122 résidus carboxy-terminaux ne font ainsi pas partie de la construction de la protéine ayant été cristallisée. Rad23 ne possède également pas ses 229 résidus amino-terminaux.

Dans sa forme libre d'ADN, la structure montre le domaine similaire à la transglutaminase (TGD) et les domaines épingle- β (BHD1, BHD2, BHD3) tous alignés, les uns à la suite des autres, dans cet ordre. Rad23 interagit uniquement avec le domaine TGD. Les domaines en épingle- β sont très semblables structurellement les uns avec les autres, avec des différences dans les boucles reliant les brins- β .

Dans sa forme liée à l'ADN, tous les domaines se repositionnent légèrement les uns par rapport aux autres. Seul Rad23 bouge peu, n'étant pas impliqué dans la liaison à l'ADN. Les domaines semblent s'adapter à la forme légèrement pliée de l'ADN, typique des lésions

reconnues par XPC/Rad4. Ce sont les domaines BHD2 et BHD3 qui interagissent le plus près de la lésion, via leurs boucles flexibles. Ces domaines n'interagissent pas directement avec la lésion, mais plutôt avec les bases opposées à la lésion qui ne sont pas appariées. Les domaines BHD1 et TGD, quant à eux, interagissent avec l'ADN double brin non-endommagé, en 3' de la lésion. Cette structure explique en partie le large éventail de substrats que XPC/Rad4 reconnaît, puisque ce n'est qu'indirectement que XPC/Rad4 interagit avec la lésion.

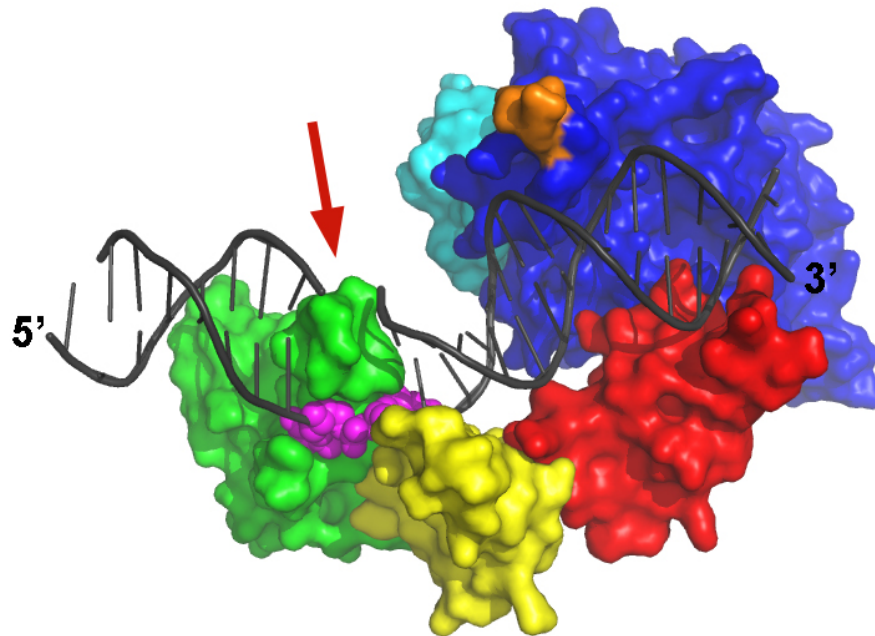


Figure 1-6. Structure tridimensionnelle de Rad4 lié à l'ADN endommagé. Le code-couleur est le même qu'à la Figure 1-5 : domaine TGD (bleu), domaines BHD1 (rouge), BHD2 (jaune), BHD3 (vert), ADN (gris), nucléotides non-pairés, opposés à la lésion (représentation en balles, magenta). Le segment amino-terminal est coloré en orange, montrant la proximité du côté de la lésion. La lésion n'est pas visible dans la structure cristallographique, puisqu'elle adopte probablement plusieurs conformations. Son emplacement est pointé par une flèche.

1.2.3 TFIIH lors de la NER

Après la détection du dommage à l'ADN par XPC/Rad4 et ses partenaires, il y a assemblage d'un complexe protéique en vue de l'incision. Un membre très important est le complexe protéique TFIIH, qui ouvrira la double hélice d'ADN, testera le dommage et recrutera le reste de la machinerie. C'est par son rôle dans la transcription que TFIIH a été découvert (86, 87). Ce n'est que par la suite qu'il a été relié à la réparation de l'ADN, en remarquant qu'une de ses sous-unité hélicase était le produit du gène *ERCC3* (renommé par la suite *XPB*) (88). Nous étudierons les fonctions de TFIIH dans la transcription dans une section subséquente (voir section 1.3). Survolons maintenant la composition générale de ce complexe et son rôle dans la NER.

Sous-unités de TFIIH et leurs fonctions

TFIIH est très bien conservé chez les eucaryotes. Le complexe chez la levure est très semblable au complexe humain, possédant les mêmes 10 sous-unités (Figure 1-7). La structure générale de TFIIH a été résolue par cryomicroscopie électronique (Figure 1-8) et montre également une similarité structurale entre l'humain et la levure (89, 90). En forme d'anneau avec une protubérance, la structure peut être séparée en deux sous-complexes : le « noyau » (*core*) et le sous-complexe CAK.

	<i>S. cerevisiae</i>	Humain
hélicases	<u>Ssl2</u>	<u>XPB</u>
	<u>Rad3</u>	<u>XPB</u>
Sous-complexe « noyau »	Tfb4	p34
	Ssl1	p44
	Tfb2	p52
	Tfb1	p62
	Tfb5	p8 (TTDA)
Sous-complexe CAK	<u>Kin28</u>	<u>CDK7</u>
	Ccl1	CCNH
	Tfb3	Mat1

Figure 1-7. Sous-unités de TFIIH chez la levure *S. cerevisiae* et chez l'humain. Les sous-unités catalytiques sont surlignées et les deux sous-complexes ont été divisés.

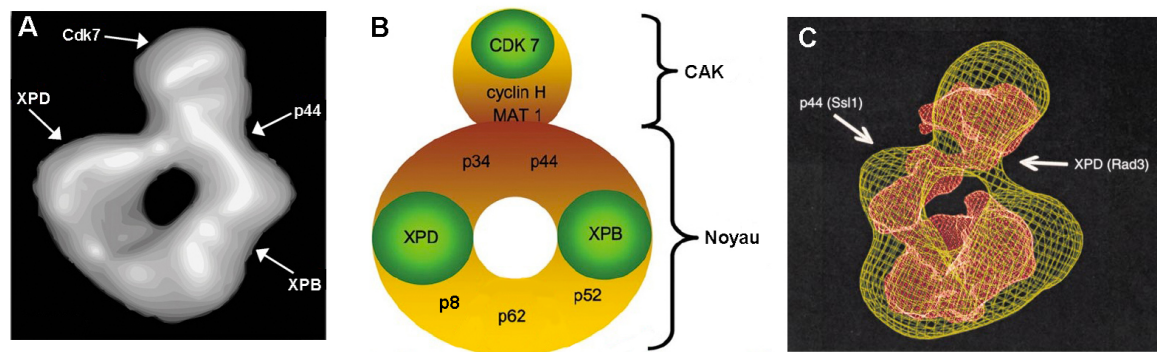


Figure 1-8. Structures globales de TFIIH. Les structures de TFIIH d'humain (A, adapté de (90)) et de levure (C, adapté de (89)) ont été résolues par cryomicroscopie électronique. Les positions des sous-unités identifiées sont pointées par des flèches. B) Vue schématique de TFIIH, montrant le noyau en anneau et le sous-complexe CAK (adapté de (91)). Les sous-unités catalytiques sont colorées en vert. Pour simplifier le schéma, seuls les noms des molécules humaines ont été écrits, sauf pour C où les noms chez la levure sont entre parenthèses.

Sous-complexe CAK

Comme son nom l'indique, le sous-complexe CAK (cyclin-dependent kinase-activating kinase) possède une activité kinase via sa sous-unité CDK7/Kin28 (92). Il possède également deux autres sous-unités. La Cycline H (CCNH/Ccl1) régule CDK7 tandis que Mat1/Tfb3 ancre le sous-complexe CAK au noyau de TFIIH (93). Rien ne prouve que cette activité kinase ait un rôle à jouer lors de la NER.

Sous-unités hélicases

Le sous-complexe noyau possède les deux sous-unités cruciales pour la NER, puisque des mutations chez celles-ci causent le XP. XPB/Ssl2 et XPD/Rad3 sont les sous-unités catalytiques effectuant l'action de déroulement de l'ADN (24). XPB/Ssl2 est l'hélicase agissant en 5' de la lésion et déroulant l'ADN de 3' vers 5' (88), alors que XPD/Rad3 est l'hélicase déroulant l'ADN en 3' de la lésion en agissant de 5' vers 3'. Puisqu'elles ont toutes deux besoin de l'ATP pour effectuer leur action hélicase, elles sont donc aussi des ATPases. Par contre, la fonction hélicase de XPB n'est pas absolument nécessaire à la NER. Dans ce processus, XPB arrimerait TFIIH à l'ADN, tout en utilisant de l'ATP (94, 95).

Sous-unités architecturales

Outre XPD et XPB, p8/Tfb5 (TTDA) est la seule autre sous-unité de TFIIH dont les mutations chez l'humain sont reconnues pour causer une maladie, la trichothiodystrophie (TTD) (96). Cette sous-unité n'a été découverte que récemment, et elle est la seule sous-unité de TFIIH à ne pas être essentielle à la viabilité de la cellule, ni à la transcription de l'ADN (97). Son rôle est principalement de stabiliser l'échafaudage de TFIIH par son interaction avec p52/Tfb2, facilitant ainsi la NER (98). La sous-unité p52/Tfb2, quant à elle, régule XPB et l'aiderait dans l'arrimage de TFIIH à la lésion (99). Une fonction similaire est attribuée à p44/Ssl1 qui interagit avec XPD/Rad3 pour stimuler son activité hélicase. Toutes les molécules précédentes sont considérées comme ayant un rôle dans l'échafaudage et l'architecture de TFIIH, tout comme p34/Tfb4 qui interagit avec les autres sous-unités du noyau de TFIIH, sans avoir de rôle catalytique (100). Ce rôle architectural est aussi partagé par p62/Tfb1 qui interagit avec Ssl1/p44. Cette interaction est essentielle pour la NER (101). p62/Tfb1 a une fonction supplémentaire de régulation à travers des interactions protéine-protéine. Cette thèse étant centrée sur cette protéine, attardons-nous à sa séquence, à sa structure et à son rôle dans la NER.

Sous-unité p62/Tfb1

Les protéines Tfb1 et p62 ont été clonées environ au même moment et ont été immédiatement désignées comme orthologues étant donné la similarité de leurs séquences (102, 103). En effet, les séquences s'alignent très bien à plusieurs endroits, malgré le fait que Tfb1 possède 642 acides aminés alors que p62 en possède 548.

Cette similarité s'observe entre autres dans les deux domaines BSD (*BTF2-Like transcription factors, Synapse-associated and DOS2-like protein*) que possèdent Tfb1 et p62 (104) (Figure 1-9). La fonction de ces domaines est inconnue, mais leur structure a été résolue et montre un petit ballot de trois hélices- α (données non-publiées, code PDB 2DII). Le second BSD serait peut-être important pour l'interaction avec XPD/Rad3 à l'intérieur du complexe TFIIH, puisque cette interaction a été cartographiée dans cette région de p62 (105).

Chez la levure, les premiers mutants de Tfb1 à avoir été étudiés sont des délétions de 34 à 114 acides aminés à l'extrémité carboxy-terminale (101, 106). Tous ces mutants sont sensibles aux UV, ce qui supporte l'hypothèse que Tfb1 pourrait jouer un rôle lors de la NER. Les protéines porteuses de ces délétions n'interagissent plus avec Ssl1, ce qui compromet la stabilité du complexe TFIIH. L'activité de TFIIH est diminuée, mais pas éliminée. Même si les mêmes expériences n'ont pas été faites chez p62, on peut supposer que p44 se lie au même endroit.

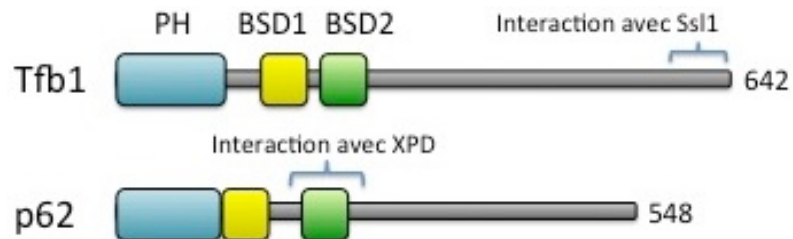


Figure 1-9. Structures primaires de Tfb1 et p62. Les domaines PH (bleu) et BSD (jaune et vert) se trouvent dans la portion amino-terminale. Dans la portion carboxy-terminale se trouve le site d'interaction avec Ssl1, uniquement mis en évidence chez Tfb1. L'interaction avec XPD a uniquement été démontrée chez p62 et coïncide avec le second BSD.

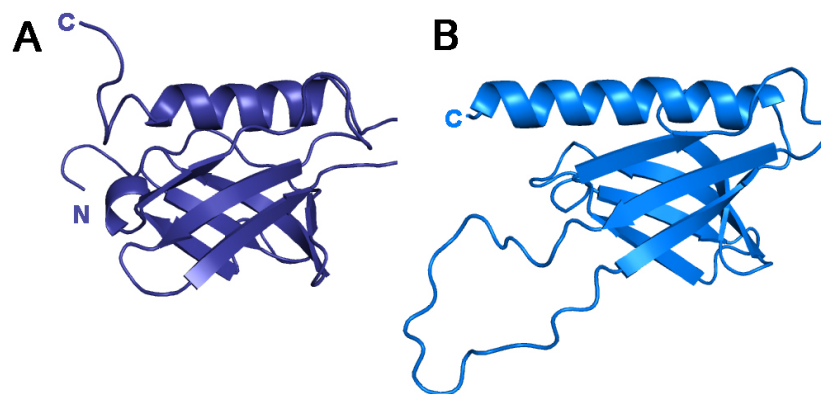


Figure 1-10. Structures des domaines PH de p62 (A) et de Tfb1 (B).

C'est dans la portion amino-terminale qu'on retrouve le domaine ayant suscité le plus d'intérêt à présent. Celui-ci n'est pas responsable de l'assemblage ni de l'architecture de TFIIH. La structure de ce segment a été résolue pour p62 (107), puis Tfb1 (108). Les

deux structures, très similaires, sont composées de deux feuillets- β (de 3 et 4 brins- β) couronnées d'une hélice- α (Figure 1-11). Ce type de repliement est caractéristique de la famille des domaines homologues à la pleckstrine (*Pleckstrin Homology*, PH). Cette famille fait partie des 20 domaines les plus communs du protéome humain (109). Le fait que ce repliement soit très stable et qu'il permette de présenter plusieurs types de sites de liaison est probablement à l'origine de cette prépondérance (110). Les domaines PH sont reconnus pour lier les phosphoinositides et les protéines phosphorylées (111). Bien que Tfb1 semble pouvoir lier les phosphatidyl inositols via une de ses surfaces chargée positivement, aucune fonction biologique liée à cette caractéristique n'a été découverte. Par contre, plusieurs interactions protéine-protéine ont été démontrées.

L'interaction la mieux documentée du domaine PH de p62/Tfb1 dans la NER est le recrutement de XPG/Rad2 au site de dommage à l'ADN, afin que TFIIH en modifie l'activité catalytique (112). Cette interaction sera détaillée dans une prochaine section, mais avant penchons-nous plus précisément sur la structure et la fonction de l'enzyme XPG.

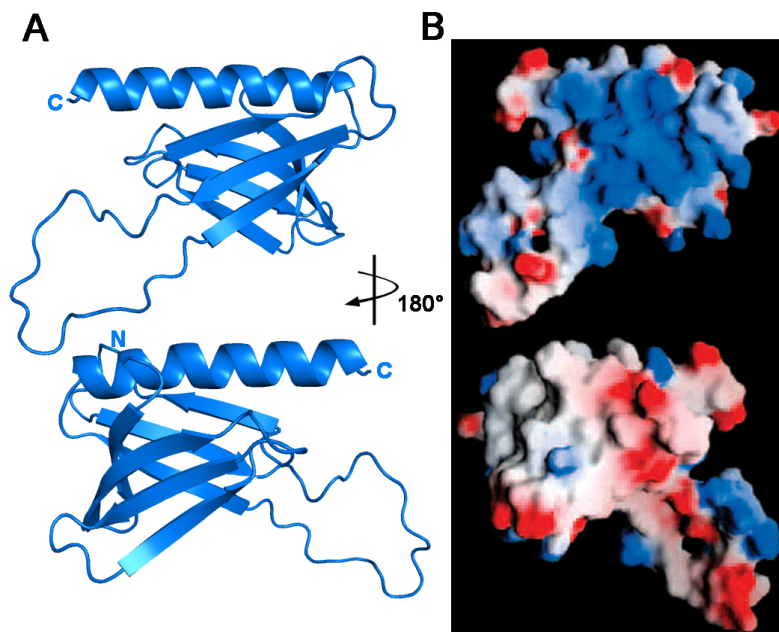


Figure 1-11. Le domaine PH de Tfb1 possède une face chargée positivement. A) Structure du domaine PH de Tfb1. Deux faces sont montrées, à 180° l'une de l'autre. B) Surface électrostatique de Tfb1 (Tiré de (108)), dans les mêmes orientations qu'en A.

1.2.4 Endonucléase XPG/Rad2

Le gène *RAD2* de la levure *S. cerevisiae* a été cloné au milieu des années 80 grâce à des expériences de complémentation du mutant *rad2* (113, 114). Ce n'est que dix ans plus tard que la protéine Rad2 a été identifiée premièrement comme étant une endonucléase (115, 116) et ensuite comme actrice de la NER (117). Le gène *XPG*, orthologue humain de *RAD2*, a été quant à lui cloné et fonctionnellement identifié environ au même moment (118-120). Initialement appelé *ERCC5* par sa capacité à corriger une carence en NER chez des cellules de rongeurs (121), il a été renommé *XPG* lorsque relié au groupe de complémentation XP-G chez l'humain (122). Depuis lors, nous en avons beaucoup appris sur la séquence, la fonction et la structure de cette enzyme, présentées dans cette section.

Séquence

XPG/Rad2 est une protéine de 1186/1031 acides aminés dont la séquence est très similaire à celle d'autres nucléases. On fait souvent référence à la famille des nucléase de FEN-1 (123, 124), puisque toutes ces nucléases coupent une structure spécifique de l'ADN. Parmi les membres de cette famille, on retrouve plusieurs protéines avec des rôles dans des voies cellulaires différentes, comme FEN-1 (flap endonucléase 1, réplication de l'ADN), EXO-1 (exonucléase 1, réparation des mésappariements) ou GEN-1 (gap endonucléase 1, recombinaison homologue). En se basant sur l'alignement des séquences de tous les membres de la famille, XPG/Rad2 a été séparée en quatre domaines (Figure 1-12).

Comme on le verra plus loin, ces domaines ne sont pas des domaines structuraux à proprement parler. Les domaines N (95 acides aminés) et I (129 acides aminés) sont les domaines catalytiques de l'enzyme qui effectuent le clivage. Ils sont conservés chez tous les membres de la famille. Entre ces deux domaines, une région de longueur variable appelée segment espaceur (spacer region) est d'une longueur très variable, allant de 24 acides aminés (FEN-1) à plus de 600 acides aminés (XPG/Rad2). Chez XPG/Rad2, la séquence est très acide et il est prédit qu'elle n'adopte aucune structure secondaire (125). Finalement, le domaine carboxy-terminal est lui aussi variable en longueur et en séquence. Contrairement au domaine catalytique, les fonctions des segments espaceur et carboxy-terminal ne sont pas bien définies.

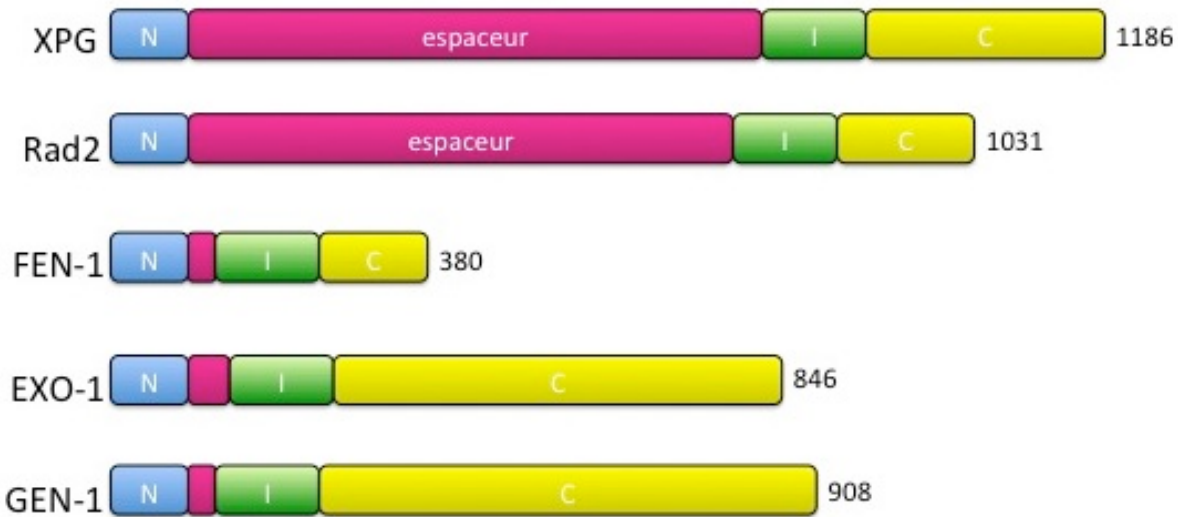


Figure 1-12 Structures primaires des membres de la famille de FEN-1. Les domaines catalytiques N (bleu) et I (vert) ne changent pas, alors que le segment espaceur (magenta) et le segment carboxy-terminal (jaune) sont de longueur variables.

Fonctions biologiques

Par définition, toutes les nucléases effectuent le même travail: couper l'ADN. Par contre, les nucléases de la famille FEN-1 ont toutes des substrats spécifiques, mais différents entre les membres (Figure 1-13). De façon générale, elles clivent l'ADN simple brin flottant en 3', près d'une jonction d'ADN double brin (124). Plus spécifiquement, XPG/Rad2 clive les bulles, les boucles et les formations en « bras ouverts » (27, 126-128). Ceci fait en sorte qu'elle clive l'ADN en 3' de la lésion lors de la NER. Un autre exemple est celui de FEN-1 qui, par son implication dans le processus de traitement des fragments d'Okazaki lors de la réplication de l'ADN, clive uniquement l'ADN simple brin des jonctions qui ont un 5' OH libre (129). L'ADN simple brin clivé par FEN-1 doit être libre en 5' et ne clive donc pas les bulles comme XPG.

Cette spécificité a été attribuée d'une part au site actif des nucléases, qui est légèrement différent et permet d'accueillir des substrats ayant une structure tridimensionnelle précise. D'autre part, le segment espaceur différent des nucléases serait aussi responsable de la spécificité (125). En effet, lorsqu'on échange le segment espaceur

de FEN-1 par celui de XPG, FEN-1 acquiert alors des propriétés semblables à XPG (130). Une autre fonction biologique du segment espaceur de XPG est d'interagir avec plusieurs autres protéines de la NER (125, 131-133). Toutes ces interactions sont spécifiques à XPG et ne sont pas présentes chez les autres membres de la famille FEN-1. Certaines d'entre elles sont même assez fortes pour permettre la co-purification de XPG/Rad2 avec TFIIF (134, 135).

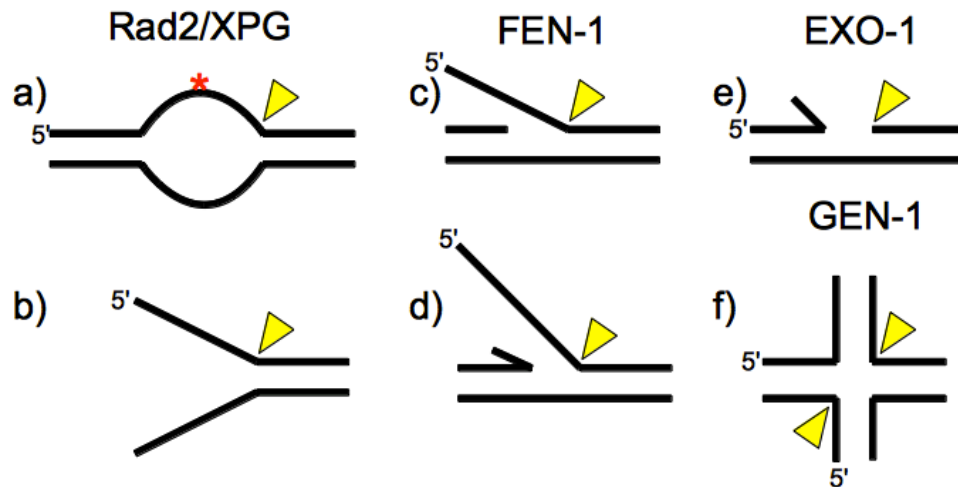


Figure 1-13. Les nucléases de la famille FEN-1 clivent spécifiquement différents substrats. La flèche montre le lieu de clivage de l'ADN, toujours en 5' d'une jonction entre l'ADN double brin et l'ADN simple brin. XPG/Rad2 clive les bulles causées par un dommage à l'ADN (montré par une étoile) (a), mais aussi les formations en « bras ouverts » (b). Les structures en volet (*flap*) (c) ou en double volet (d) sont la spécialité de FEN-1. L'exonucléase EXO-1 clive les nucléotides un à un tel que montré en (e), alors que GEN-1 clive des jonctions de Holliday (f).

D'autres interactions protéine-protéine sont permises grâce au segment carboxy-terminal de XPG/Rad2, notamment avec PCNA et CSB (136). Cette dernière interaction aide à la formation du complexe d'incision lors de la TC-NER (137). Le segment carboxy-terminal possède aussi des séquences de localisation nucléaire (NLS) postulées (138). Aucune donnée n'est malheureusement disponible sur le mécanisme associé à l'import au noyau de XPG/Rad2 ou si des mécanismes de rétention au cytoplasme existent.

Une dernière fonction de XPG dans la NER est la stabilisation du complexe de pré-*incision*, ce qui permet à XPF-ERCC1 de se positionner et d'effectuer le clivage de l'ADN

en 5' de la lésion. La présence de XPG, mais pas son activité catalytique est nécessaire au clivage par XPF-ERCC1, tel que démontré par l'utilisation d'un mutant de XPG inactif pour le clivage, mais actif pour la liaison à l'ADN (139, 140). En utilisant ce mutant, l'ADN est clivé par XPF-ERCC1, mais pas par XPG.

Outre son rôle dans la NER, XPG/Rad2 a été relié à la transcription (141, 142), bien que son rôle exact dans ce processus soit inconnu. Des mutations chez XPG empêchent l'association entre XPG et TFIIH, ce qui provoque la dissociation du domaine CAK et de XPD du reste du complexe. De plus, une surexpression de Rad2 cause une catastrophe mitotique (143). Finalement, Rad2 pourrait avoir des fonctions dans la maturation des fragments d'Okazaki (144). Tous ces exemples démontrent l'importance de XPG/Rad2 pour les cellules.

En conclusion, certaines mutations de XPG créent des problèmes autres que dans la réparation de l'ADN, notamment dans le développement. Ceci nous indique qu'il y aurait d'autres fonctions encore inconnues associées à cette protéine. Une fonction certaine, par contre, est l'activité catalytique de XPG pour cliver l'ADN.

Structure et catalyse

Malgré des fonctions biologiques différentes chez les membres de la famille de FEN-1, leur mécanisme catalytique est semblable. La forte conservation des résidus des domaines catalytiques nous permet aussi d'affirmer que la structure est également conservée. De plus, la structure d'un chimère XPG-FEN-1 est identique à la structure de FEN-1 (145). Ainsi, plusieurs informations de cette section seront tirées des études sur FEN-1, que nous pourrons extrapoler à XPG/Rad2.

On pense souvent, à tort, qu'un domaine dans une séquence d'acide aminé égale domaine structural. Les nucléases de la famille de FEN-1 en sont un excellent exemple. La structure de FEN-1 nous montre à quel point les régions N et I sont imbriquées, ne formant qu'un seul domaine catalytique (Figure 1-14). La majeure partie des interactions avec l'ADN se font par ce domaine catalytique. Le domaine carboxy-terminal et le segment espaceur de 24 acides aminés jouent aussi un rôle dans l'interaction avec l'ADN (146). Il est

possible d'imaginer qu'un segment espaceur de 700 acides aminés, comme dans le cas de XPG/Rad2, pourrait s'échapper du domaine catalytique pour interagir avec d'autres protéines.

Lié à la structure, l'ADN est plié à 90°, ce qui permet aux acides aminés participant à la réaction de clivage d'avoir accès au lien phosphodiester. On sait depuis longtemps que XPG/Rad2 dépend du magnésium ou du manganèse divalent (Mg^{2+} , Mn^{2+}) (147). Or, c'est par un mécanisme dit de « two-metal ion » que la catalyse des nucléases de la famille de FEN-1 s'effectue. Ces ions sont retenus par des résidus acides, dont Asp-77, Glu-791 et Asp-812 chez XPG (148). Mutés, ces résidus rendent l'enzyme inactive. L'importance de ces résidus dans la catalyse explique des maladies telles que le xeroderma pigmentosum, sujet traité dans la prochaine section.

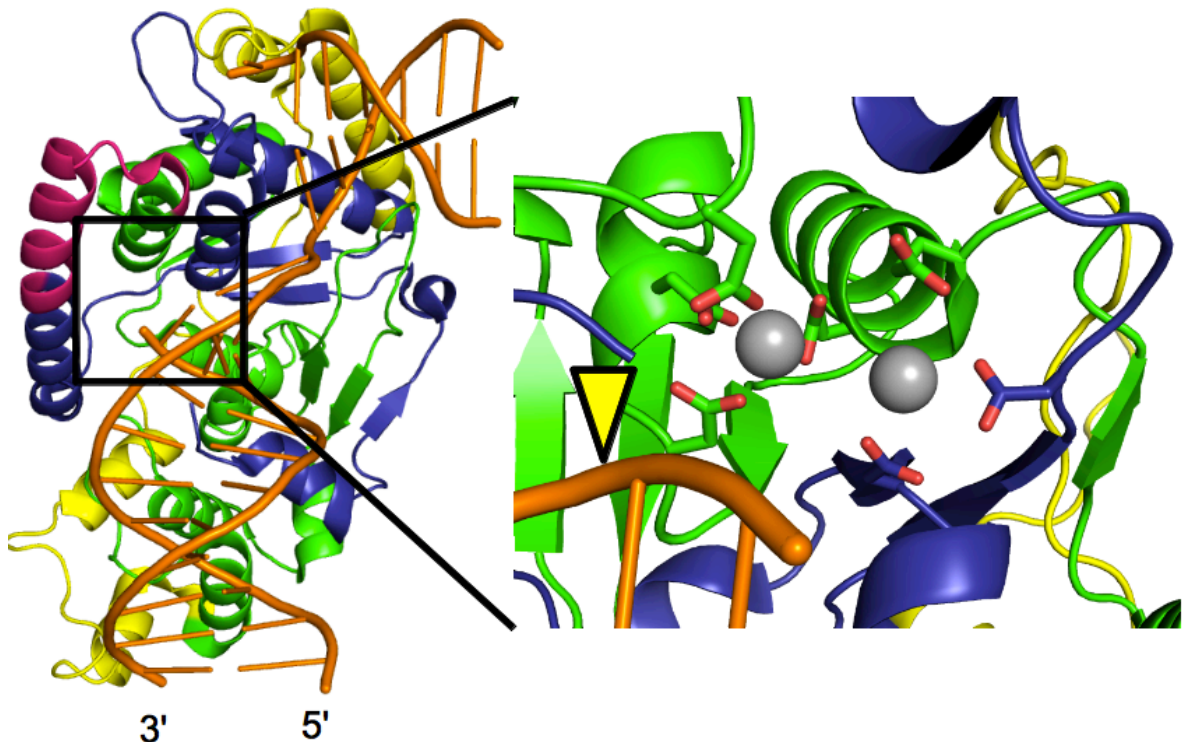


Figure 1-14. Structure de FEN-1 liée à l'ADN. Les couleurs sont les mêmes qu'à la Figure 1-12 (domaine N (bleu), domaine I (vert), segment espaceur (magenta), segment carboxy-terminal (jaune), ADN (orange), ions divalents (gris)). À gauche, la structure globale de FEN-1 démontre que tous les domaines contribuent à la liaison à l'ADN. À droite, gros plan sur le site actif montrant les résidus acides coordonnant les ions divalents, très près de l'endroit où le clivage s'effectuera sur l'ADN (flèche jaune). Une rotation de 90 degrés par rapport à la figure de la structure globale a été effectuée pour créer la figure du site actif.

1.2.5 Syndromes et mutations associées à la NER

Trois syndromes sont reliés à la NER: xeroderma pigmentosum (XP), le syndrome de Cockayne (CS) et la Trichothiodystrophie (TTD). Ces maladies génétiques sont toutes autosomiques récessives. Les personnes porteuses d'un gène défectueux (hétérozygotes) n'ont pas de symptômes, bien que certaines études affirment que les personnes portant un gène XP défectueux ont plus de chance de développer un cancer (149).

Xeroderma pigmentosum

La première maladie à avoir été reliée à la NER est le XP (150). La principale caractéristique des patients atteints de XP est leur photo-sensibilité. Ils ont mille fois plus de risques de développer un cancer de la peau et développent donc des carcinomes à un très jeune âge (151). En relation à ce problème de sensibilité aux rayons UV, le XP cause des problèmes oculaires comme des anomalies de la cornée et une faible vision (152). Finalement, les problèmes neurologiques sont aussi une caractéristique importante. Les patients atteints de XP ont souvent des retards mentaux ou une microcéphalie. Ces symptômes varient selon les patients. Certains peuvent être absents, d'autres de gravité différente. La plupart du temps, à l'intérieur d'un même groupe de complémentation, les symptômes seront semblables. Par exemple, les patients atteints de XP-C ne présentent généralement pas de détérioration neurologique. D'autres groupes de complémentation ont des niveaux de gravité moindres, comme par exemple XP-E qui est l'un des plus rares et des moins nocifs de tous les groupes de complémentation.

Syndrome de Cockayne

Les patients atteints de CS sont moyennement photo-sensibles, mais vieillissent plus rapidement, ont une petite taille (nanisme), des malformations faciales particulières et des retards de développement neuronal, ce qui entraîne la surdité (153). Les patients atteints de CS sévère meurent avant l'âge de 7 ans. Les majorité des cas rapportés sont du groupe CS-B.

Trichothiodystrophie

Les patients atteints de TTD ont un retard mental en plus d'avoir des problèmes de développement physique (154). Une caractéristique particulière est que leurs cheveux et leurs ongles sont fragiles et cassants. Leur peau est également fragile, créant des « écailles ». Ceci est dû à une déficience dans la synthèse de certaines protéines de structure à haute teneur en soufre, comme la kératine (154).

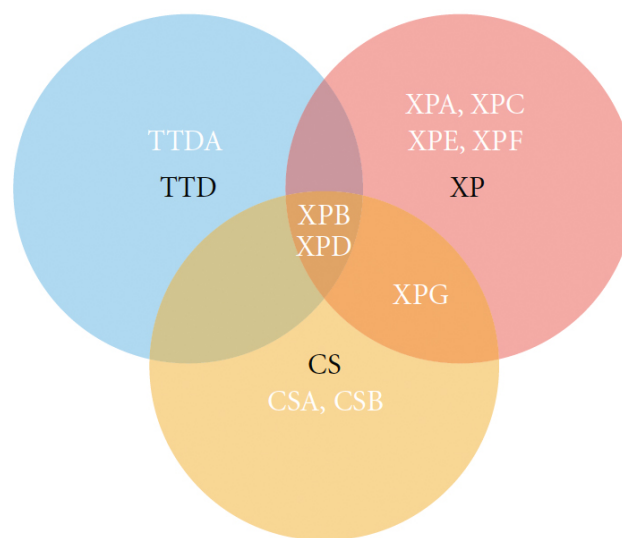


Figure 1-15. Diagramme de Venn représentant les trois maladies liées à la NER: xeroderma pigmentosum (XP), syndrome de Cockayne (CS), trichothiodystrophie (TTD). À l'intérieur des cercles, en blanc, les protéines possédant des mutations causant ces maladies. Par exemple, des mutations chez XPG peut causer le XP, mais aussi, une combinaison XP-CS (Tiré de (155)).

Mutations génétiques

Tel que le démontre la Figure 1-15, il n'existe pas nécessairement de lien direct entre les mutations d'une protéine et un syndrome. Des mutations chez XPD peuvent causer le XP, d'autres le CS, le TTD ou même une combinaison de ces syndromes. Une pléthore de mutations ont été identifiées et nous n'en survolerons que quelques-unes dans cette section.

Des mutations de la sous-unité p8 de TFIIH sont majoritairement responsables de la TTD. Comme p8 est reconnue pour stimuler la NER en stabilisant TFIIH (156), ce sont les bas niveaux de TFIIH stables et effectifs dans les noyaux qui seraient la cause de cette maladie. D'autres mutations de XPB et XPD entraînant une déstabilisation TFIIH, sont également responsables de ce syndrome. En effet, 80% des mutations de XPD empêchent l'association avec p44 (157), causant non seulement le TTD, mais aussi le CS et le XP. Cette interaction est nécessaire pour la stimulation de l'activité de XPD. Ainsi ce n'est pas ses activités hélicase ni ATPase qui sont touchées directement. Même chose pour XPB où une mutation affecte son association avec p52, ce qui a pour effet de réduire (sans anéantir) son activité hélicase (158).

D'autre part, chez XPG, les mutations sont souvent directement reliées à une dysfonction directe de l'activité catalytique. Par exemple, la mutation A792V perturbe l'acide glutamique 791, résidu conservé et nécessaire à la catalyse (148, 159). D'autres perturbations causant les pires déficiences sont des mutations causant des troncations de XPG, rendant cette enzyme complètement inactive.

Les troncations sont également une catégorie importante de mutations chez XPC. Les troncations aux résidus 474, 579 et 814 empêchent toutes la liaison à l'ADN endommagé (160). D'autre part, des mutations ponctuelles (W690S ou l'insertion d'une valine à la position 697) rendent la protéine instable (160, 161). Certains émettent l'hypothèse qu'une diminution de l'interaction avec RAD23B serait la cause de cette instabilité. Finalement, la mutation P334H empêche la pleine activité hélicase de XPB, par un mécanisme inconnu (160).

En conclusion, toutes ces maladies ont des caractéristiques ne s'expliquant pas uniquement par une déficience en réparation de dommages à l'ADN, ce qui suggère que les mutations reliées à ces maladies affectent des protéines ayant un rôle dans d'autres processus. C'est le cas de TFIIH, primordial pour la NER, mais aussi pour la transcription de l'ADN.

1.3 TFIIH : pont entre la transcription et la NER

TFIIH a d'abord été reconnu pour son rôle dans la transcription (86, 87) avant de savoir qu'il agissait aussi dans la réparation de l'ADN (88). Or, c'est son rôle d'hélicase qui est commun aux deux voies. Pour ces processus, l'ADN doit être déroulé pour que les simples brins soient accessibles. De ce fait, il agit en véritable pont entre la transcription de l'ADN et la NER. Nous avons déjà vu son rôle dans la NER, voyons maintenant en détails son rôle dans la transcription de l'ADN.

1.3.1 Rôles dans la transcription de l'ADN

TFIIH est un des six complexes protéiques nécessaires à la formation d'un complexe de pré-initiation lors du processus de transcription de l'ADN en ARN par l'ARN polymérase II. Ces complexes sont appelés facteurs de transcription généraux par opposition aux facteurs de transcription qui sont spécifiques et qui peuvent varier d'un gène à l'autre. Selon un modèle d'assemblage séquentiel, les facteurs de transcription généraux sont recrutés les uns après les autres au promoteur d'un gène. Brièvement, TFIID se lie à une séquence d'ADN nommée boîte TATA via sa sous-unité TBP (TATA Binding Protein) (162) (Figure 1-16). Ceci crée une courbure dans l'ADN (163), permettant à TFIIA et TFIIB d'être ensuite recrutés. Puis, l'ARN polymérase II est elle aussi recrutée, liée à TFIIF. TFIIE arrive au site d'initiation pour finalement recruter TFIIH ce qui débutera la transcription par une phase appelée élongation (164).

Dans ce processus, le premier rôle de TFIIH est celui d'hélicase: il ouvre l'ADN autour de la région d'initiation de la transcription, d'une façon dépendante de l'ATP (165, 166) et grâce à la sous-unité XPB principalement (167). Cette sous-unité serait aussi importante dans la sortie de l'ARN polymérase II du promoteur, en empêchant l'arrêt prématuré de la polymérisation de l'ARN au tout début de l'élongation (168, 169). Un second rôle catalytique consiste en la phosphorylation du domaine carboxy-terminal (CTD) de la sous-unité Rpb1 de la RNA polymérase II grâce à la sous-unité Cdk7 (166). Cette phosphorylation est le déclencheur de la polymérisation de l'ARN par l'ARN polymérase II (170).

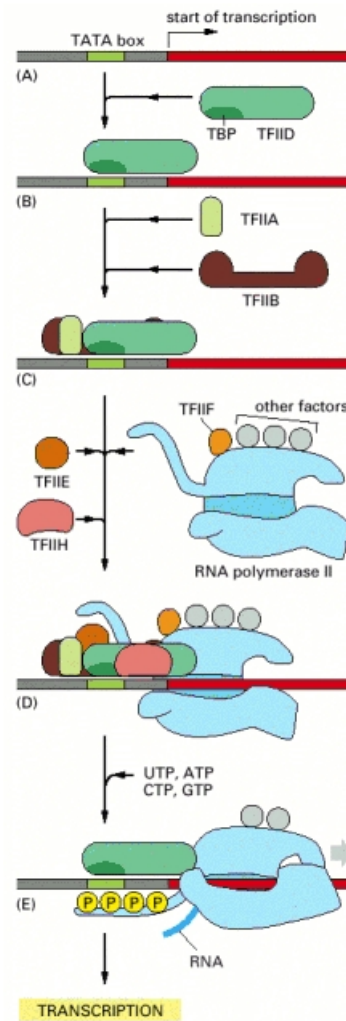


Figure 1-16. Modèle d'assemblage séquentiel du complexe de pré-initiation (Tiré de (171)).

Toujours grâce à sa sous-unité Cdk7, TFIIH phosphoryle également d'autres protéines, comme les récepteurs nucléaires. C'est le cas du récepteur de l'acide rétinoïque γ (172) et du récepteur de l'œstrogène α (173). Cette phosphorylation modifie et régule leur activité.

TFIIH aurait aussi un rôle dans la transcription de l'ADN ribosomal, ADN transcrit par l'ARN polymérase I en ARN ribosomal, constituant des ribosomes (174, 175). En effet, on retrouve TFIIH dans le nucléole et aux sites de transcription de l'ADN ribosomal. De plus, l'ARN polymérase I est inactive lorsque TFIIH est absent de la cellule. Par contre, contrairement à la transcription par l'ARN polymérase II, TFIIH ne serait pas nécessaire à

l'initiation de la transcription, mais plutôt à d'autres étapes subséquentes, telles le départ du promoteur (*promoter clearance*) et à l'élongation.

Finalement, les rôles de TFIID ne se limitent pas à ses activités catalytiques lors de la transcription. Les interactions protéine-protéine qu'il peut effectuer sont également très importantes. Voyons plus en détail son recrutement à l'ADN lors de la transcription basale, la transcription activée, pour finalement faire le pont avec les interactions lors de la NER.

1.3.2 Recrutement lors de la transcription basale

L'interaction entre TFIID et TFIIE est très importante, puisque TFIIE recrute TFIID au promoteur, complétant ainsi le complexe de pré-initiation. Puis, TFIIE stimule l'activité ATPase de XPB, initiant ainsi l'ouverture de l'ADN (176). TFIIE est un hétérodimère possédant deux sous-unités, TFIIE α et TFIIE β (177), mais l'interaction avec TFIID n'a lieu qu'avec la partie carboxy-terminale de TFIIE α (178). Celle-ci lie la sous-unité p62/Tfb1 de TFIID (179, 180). Tel que vu précédemment, cette sous-unité joue un rôle structural dans le complexe TFIID et n'a pas d'activité catalytique. Par contre, elle possède un domaine PH qui est responsable de certaines interactions protéine-protéine, dont celle avec le segment carboxy-terminal TFIIE α .

Ce segment possède une structure repliée composée de trois hélices- α et de deux brins- β (Figure 1-17A) (181). Cette structure vient s'asseoir sur l'hélice- α du domaine PH de p62 (p62PH) (182). La portion amino-terminale de ce segment, composée principalement de résidus aspartates et glutamates, interagit avec un des feuillets- β du domaine PH qui est chargé positivement. Les rares résidus hydrophobes (Phe387 et Val390) viennent se loger dans des cavités hydrophobes de p62PH. Outre une petite portion formant un brin- β prolongeant le feuillet- β de p62PH, ce segment n'a pas de structure secondaire particulière.

Fait intéressant, cette surface d'interaction est la même que pour des protéines impliquées dans la transcription activée. Ces activateurs sont donc en compétition avec TFIIE pour la liaison avec p62/Tfb1.

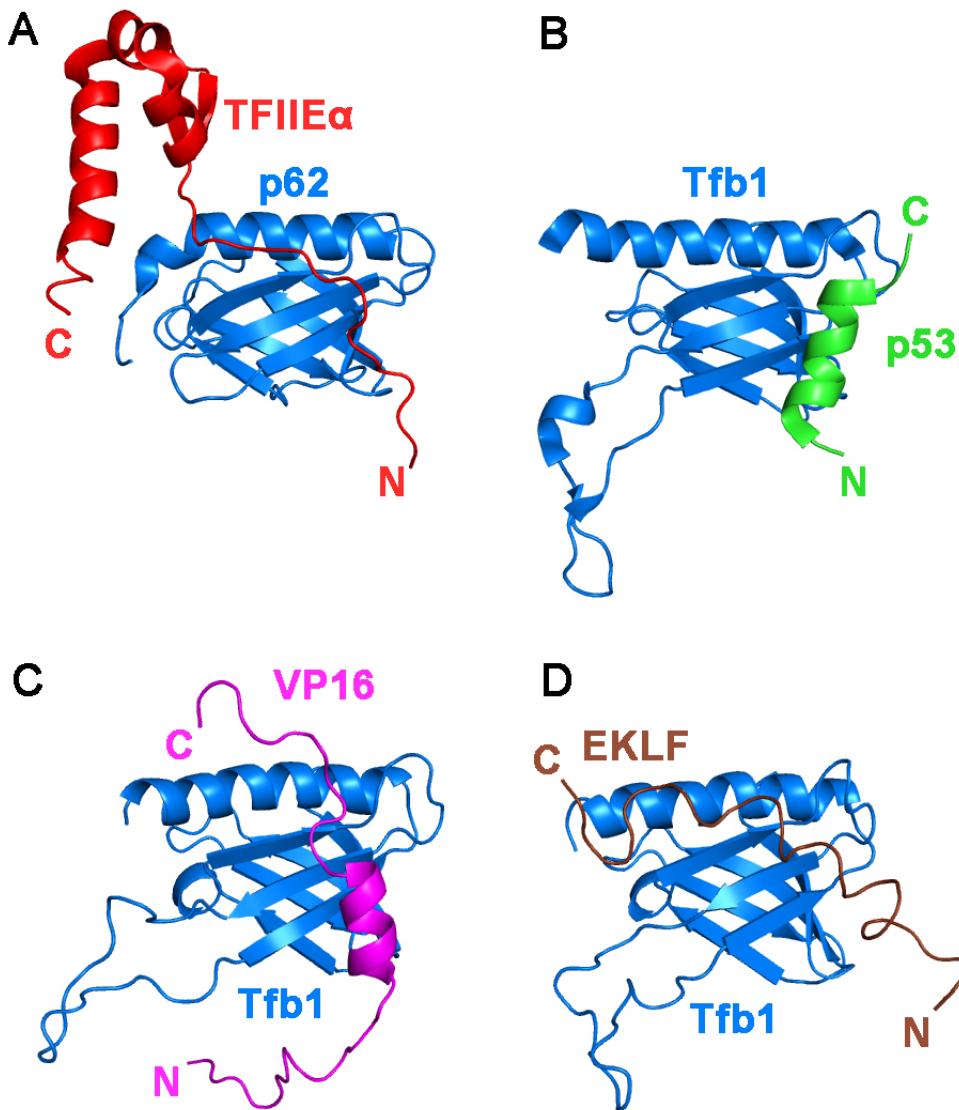


Figure 1-17. Structures du domaine PH de p62/Tfb1 avec TFIIE α (A), p53 (B), VP16 (C), EKLF (D). Les extrémités de chaque peptide sont annotées d'un C (carboxy-terminal) et d'un N (amino-terminal).

1.3.3 Recrutement lors de la transcription activée

Contrairement aux facteurs généraux de transcription, les facteurs de transcription ne sont impliqués que dans la transcription de gènes précis. Il existe une pléthore de facteurs de transcription ayant des fonctionnements différents, mais on peut les diviser en deux grandes catégories : les activateurs et les répresseurs. Les activateurs agissent afin d'activer la transcription d'un gène précis. L'activateur-type de base possède deux domaines : un domaine de liaison à l'ADN, qui lie une séquence d'ADN précise au

promoteur d'un gène, et un domaine de transactivation (*Transactivation Domain*, TAD), qui est responsable des interactions protéine-protéine menant à l'activation d'un gène (183). L'activateur recrute donc les protéines responsables du remodelage de la chromatine, les co-régulateurs, les facteurs généraux de transcription et le complexe médiateur (184). Tout ceci a pour effet de stimuler la transcription du gène. Or, plusieurs activateurs sont reconnus pour interagir avec TFIID (185) grâce à leur TAD riche en résidus acides (aspartate, glutamate). C'est le cas de E2F-1, p53, EBNA2, VP16 (186-189).

E2F-1 est un facteur de transcription régulant l'expression de gènes lors du cycle cellulaire, plus précisément lors de la transition G1/S. Son TAD lie le suppresseur de tumeur du rétinoblastome (Rb), mais interagit aussi avec TFIID à travers sa sous-unité p62/Tfb1 (188). Ces deux interactions sont en compétition, ce qui pourrait permettre un rôle d'interrupteur lors du cycle cellulaire. De plus, TFIID phosphoryle E2F-1 durant la phase S, ce qui déclenche la dégradation de ce facteur de transcription (190). Alors que l'interaction avec TFIID favorise la progression du cycle cellulaire menant ultimement à la division, Rb inhibe plutôt cette progression.

Un autre facteur de transcription impliqué dans la régulation du cycle cellulaire, p53, est probablement l'activateur le plus connu, puisqu'on le retrouve muté dans 50% des cancers (191). Sans entrer dans les détails, il peut, entre autres, induire un arrêt du cycle cellulaire afin de permettre la réparation de l'ADN (192). Si cette réparation n'est pas suffisante, il induit la mort cellulaire, l'apoptose. La protéine p53 possède deux TADs : TAD1 (résidus 1-40) et TAD2 (résidus 40-83) (193). TAD1 interagit, entre autres, avec MDM2 (194), TBP (195) et CBP/p300 (194, 196) alors que le TAD2 interagit avec RPA70 (197) et TFIID (189), parmi d'autres. Dans ce dernier cas, c'est spécifiquement avec le domaine PH de la sous-unité p62/Tfb1 que TAD2 se lie (186), tout comme dans le cas de TFIIE α . La structure du complexe entre p53TAD2 et le domaine PH de Tfb1 (Tfb1PH) est par contre différente (Figure 1-17B) (198). Contrairement à TFIIE α qui possède une forme allongée au contact de p62PH, p53 adopte une conformation en hélice- α sur 9 résidus du TAD2, passant ainsi d'un état désordonné à un état ordonné. Les chaînes latérales aliphatiques de résidus hydrophobes (Ile50, Trp53, Phe54) sont responsables de l'ancrage du peptide à Tfb1, alors que des résidus acides positionnent le peptide à la surface de Tfb1,

chargée négativement. Les résidus pouvant potentiellement être phosphorylés, Ser46 et Thr55, améliorent la liaison du peptide. Ceci pourrait être une explication à la régulation de l'interaction.

Les cellules eucaryotes ne sont pas les seules à utiliser des TAD. En effet, certains virus les utilisent dans les phases d'infection des cellules. C'est le cas du virus Epstein Barr dont l'antigène nucléaire 2 (EBNA2) interagit avec p62 (187). La protéine VP16 du virus Herpes Simplex lie également TFIIH via la sous-unité p62/Tfb1 (186). La structure de ce complexe a d'ailleurs été résolue et montre des similarités importantes avec le complexe Tfb1-p53 (199) (Figure 1-17C). En effet, le TAD de VP16 se replie au contact de Tfb1, formant une hélice- α de 9 acides aminés. Des résidus hydrophobes sont responsables de l'ancrage du peptide (Phe475, Met478, Phe479). Le TAD est également acide, certains résidus formant des ponts ioniques avec des résidus basiques de Tfb1. Même si ce mécanisme de formation d'une hélice amphipathique est semblable à ce qu'on observe avec Tfb1-p53, aucun site de phosphorylation potentiel permettant d'améliorer la liaison à Tfb1 n'est observé chez VP16.

EKLF est un facteur de transcription impliqué dans l'érythropoïèse. Tout comme p53 et VP16, son TAD peut être divisé entre deux domaines distincts pouvant indépendamment interagir avec des protéines. Le TAD2 interagit avec CBP/p300 et TFIIH via la sous-unité p62/Tfb1 (200). Ce TAD est acide et sa structure en complexe avec le domaine PH de Tfb1 montre que des résidus hydrophobes sont responsables de l'ancrage (Figure 1-17D). La similarité avec les structures de p53-Tfb1 et de VP16-Tfb1 s'arrête toutefois là, puisque le TAD2 d'EKLF ne forme pas d'hélice. Il possède une conformation ordonnée, mais sans structure secondaire particulière. De ce fait, la structure de ce complexe ressemble plus à la structure de TFIIE α avec p62.

Tous ces activateurs ont le même site de liaison sur p62/Tfb1 et peuvent donc entrer en compétition avec TFIIE α (181, 182). Sachant que l'interaction entre TFIIH et TFIIE α est nécessaire à la transcription, il est possible de se demander quel est le rôle de ces autres interactions protéine-protéine entre TFIIH et les activateurs? Le modèle actuel veut que toutes ces protéines diffusent librement dans le noyau de la cellule. La capacité de recruter

une protéine permet d'augmenter sa concentration localement et de favoriser les interactions essentielles à un processus. Ainsi, les activateurs jouent ce rôle de recrutement en vue de préparer l'interaction entre TFIIH et TFIIIE α . Comme leur nom l'indique, ils activent le processus, sans en être absolument essentiels.

Toutes les structures décrites dans cette section démontrent qu'il peut exister une diversité de modes de liaison entre les TAD des facteurs de transcription et leurs effecteurs. Sachant que TFIIH n'agit pas uniquement dans la transcription, il est d'intérêt de se demander si d'autres processus, comme la réparation de l'ADN, utilisent ces mêmes mécanismes.

1.3.4 Interactions protéine-protéine de TFIIH lors de la NER

Les interactions entre TFIIH et les protéines de la voie de NER sont multiples. TFIIH interagit avec XPC/Rad4, CSB/Rad26, XPG/Rad2 et XPA/Rad14 (201-203). Par des expériences de co-immunoprécipitation avec les protéines humaines à différentes concentrations salines, il a été mesuré que les interactions les plus fortes sont avec XPC et XPG, l'interaction XPC-TFIIH étant la plus forte (69).

Interaction entre TFIIH et XPC/Rad4

Plusieurs efforts de cartographie de cette interaction ont été effectués tant chez la levure que chez l'humain (voir Figure 1-18 pour un résumé de la section). La première fois que l'interaction a été rapportée chez la levure (202), seules deux troncations de Rad4 ont été testées par co-immunoprécipitation. Les segments 1-167 et 1-502 interagissent avec TFIIH, ce qui pourrait laisser croire que seule la portion amino-terminale de 167 résidus est importante. Par contre, il faut se montrer critique, car ces troncations s'arrêtent au milieu de domaines importants, comme nous le montre la structure de Rad4. Cette étude ne précise pas avec quelle sous-unité de TFIIH est responsable de l'interaction avec Rad4.

Ce n'est que chez l'humain que l'interaction de XPC a été ciblée à XPB et p62, encore par co-immunoprécipitation (70). Grâce à cette technique, une autre étude démontre que les résidus 816-940 de XPC sont nécessaires et suffisants pour lier TFIIH (204). De plus, les résidus 1-117 de XPC n'interagissent pas avec TFIIH. Par une autre technique,

XPC tronquée des résidus 816-940 ne rétablit que 25% de l'activité NER d'extraits cellulaires XP-C (déficients en XPC). Encore une fois, il faut être vigilants avec ces résultats, puisqu'à la lumière de nouvelles interactions, la centrine 2 lie XPC aux résidus 847-866 (60), qui est en plein milieu de cette troncation. Bien qu'elle ne soit pas absolument nécessaire, la centrine 2 augmente l'activité NER. De plus, un alignement de séquence entre Rad4 et XPC nous informe que de tronquer au résidu 816 est l'équivalent de tronquer Rad4 au résidu 617 qui se trouve à l'intérieur du domaine BHD3, ce qui pourrait compromettre sa structure tridimensionnelle. Ce domaine est responsable de la liaison au brin d'ADN opposé à la lésion.

Il existe encore des incertitudes et des paradoxes quant aux domaines de liaisons exacts, tant du côté de XPC/Rad4 que de celui de TFIIH. De nouveaux indices pourraient aider dans cette quête. Les segments amino- et carboxy-terminaux de XPC/Rad4 ont été prédits comme étant désordonnés et sont sensibles aux protéases (59). De plus, la structure de Rad4 lié à l'ADN montre que le segment amino-terminal pourrait se situer près de la lésion (59). Qui plus est, les segments amino-terminaux de Rad4 et de XPC sont hautement acides tout comme les domaines de transactivation acides des activateurs de la transcription. Idem pour Rad34, protéine effectuant les fonctions de Rad4 pour la réparation de l'ADN ribosomal. En somme, il se pourrait qu'une interaction ait lieu au segment amino-terminal de XPC/Rad4. L'acidité est également une caractéristique importante du segment de XPG/Rad2 liant TFIIH.

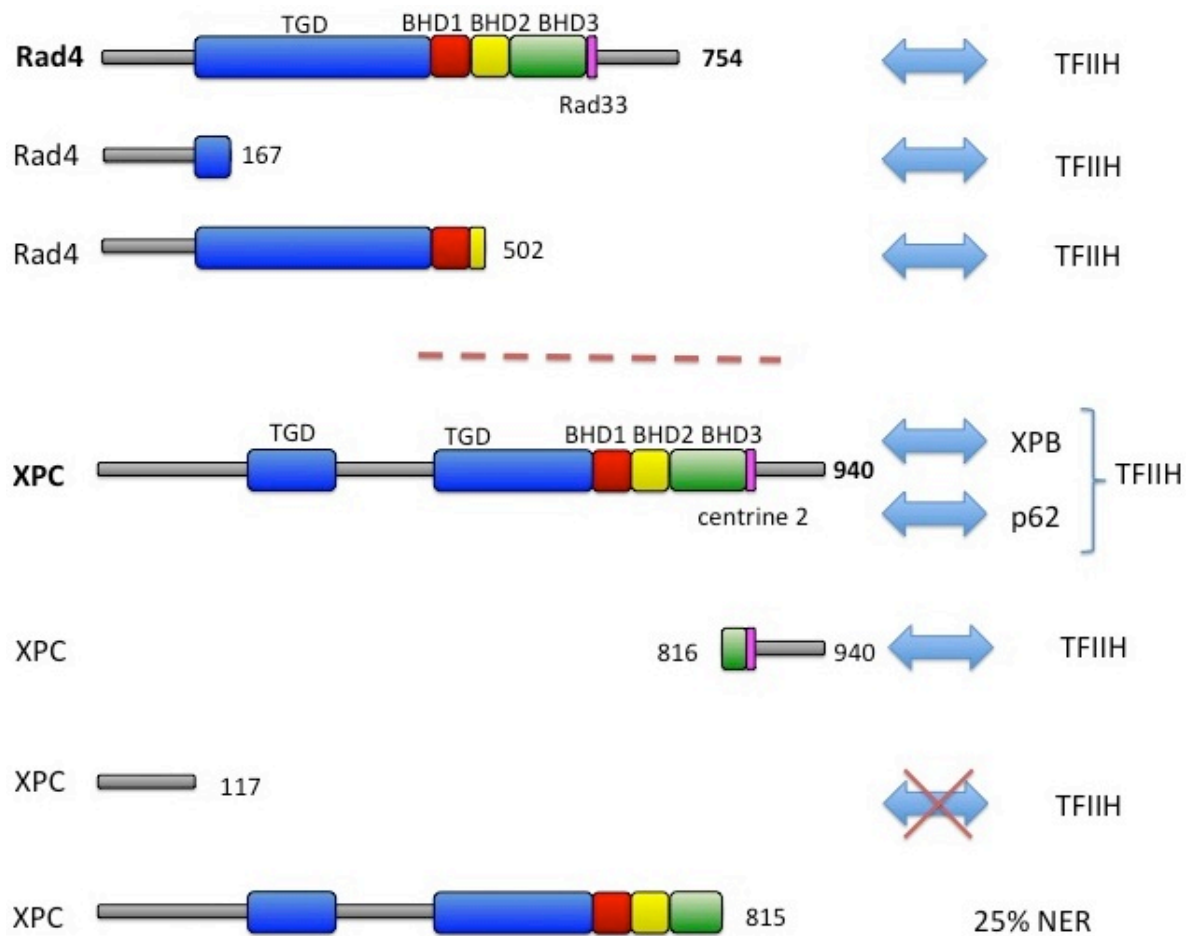


Figure 1-18. Résumé des interactions de diverses troncations de XPC/Rad4 avec TFIIH et ses sous-unités. Les flèches bleues représentent une interaction avec un partenaire, souvent TFIIH. Les segments ont été colorés comme à la Figure 1-5. La troncation 1-815 rétablit 25% de l'activité NER d'extraits déficients en XPC.

Interaction entre TFIIH et XPG/Rad2

Chez la levure, c'est le segment 642-900 de Rad2 qui interagit avec la sous-unité Tfb1 de TFIIH (202) (voir Figure 1-19 pour un résumé de la section). Ce segment comprend le domaine I (catalytique), mais aussi une partie du segment espaceur, hautement acide et désordonné. On peut donc penser que l'interaction ne se fera qu'avec la partie provenant du domaine espaceur, soit de 642 à 756. La même étude montre que les segments 1-327 et 758-1031 n'interagissent pas avec Tfb1. Malheureusement, les segments précis de Tfb1 responsables de l'interaction avec Rad2 n'ont pas été précisés.

Par contre, il a été possible de localiser le domaine d'interaction à XPG au niveau du domaine PH de p62, chez l'humain (107). Le tout se complexifie lorsque l'on cherche à localiser le segment de XPG se liant au domaine PH de p62. Comme les séquences des segments espaceurs ne sont pas conservés entre l'humain et la levure, il est difficile de faire des alignements de séquences entre XPG et Rad2. Outre la prépondérance en résidus acides, il y a peu de rapprochement à faire entre la séquence 642-756 de Rad2 et son équivalent chez XPG.

Une première étude a toutefois tenté de cartographier cette interaction (132) et plusieurs autres études s'y réfèrent. Par contre, un problème majeur est que la portion 1-65 de p62 a été utilisée avec les protéines partenaires. Selon la structure de p62, ce segment coupe en deux le domaine PH et l'empêche fort probablement de se replier correctement, ce qui peut induire des interactions non-spécifiques.

Plus tard, d'autres se sont intéressés au rôle du segment espaceur de XPG (125). Par co-immunoprécipitation, des protéines ne possédant pas les résidus 111-730 et 111-550 ne semblaient pas interagir avec TFIIH. D'autres mutants plus restreints, sans les régions 184-210 ou 554-730, avaient perdu plus de 50% de leur pouvoir de liaison à TFIIH. Dans tous les cas, ces mutants n'étaient pas présents *in vivo* au site de dommage à l'ADN (par microscopie, en utilisant des anticorps fluorescents). Le segment 184-210 pourrait contenir un motif de liaison à l'ubiquitine (*Ubiquitin Binding Motif*, UBM) conservé aussi chez Rad2 (205). Aucune fonction biologique ne lui a cependant été attribué à ce jour. Une autre étude utilisant une délétion de 7 acides aminés, de 225 à 231, provenant d'un patient atteint de XP a démontré que les protéines XPG mutantes pouvaient être recrutées au site de dommage à l'ADN, mais n'y étaient pas stables, ce qui empêche la NER (131).

Toutefois, lorsque le domaine espaceur de XPG est mis dans FEN-1, cette chimère est capable d'effectuer les activités catalytiques de XPG *in vitro*, mais est incapable de réparer complètement les lésions causées par les UV *in vivo* (130). Des études par microscopie révèlent qu'elle n'est pas non plus présente au site de dommage. Ceci indique que le domaine espaceur joue un rôle dans la catalyse de XPG, mais il n'est pas possible d'expliquer complètement le recrutement ou la stabilité de l'enzyme au site de dommage à

l'ADN. La même étude stipule brièvement qu'un tel mutant dont le segment carboxy-terminal de FEN-1 a été changé pour celui de XPG est recruté normalement au site de dommage.

Ce même segment carboxy-terminal de XPG, tronqué à partir du résidu 926, n'est plus capable de s'associer à TFIIH (142). Les auteurs de cette étude ne mettent toutefois pas l'emphase sur le recrutement de XPG au dommage à l'ADN, mais plutôt sur la stabilité du complexe TFIIH. Le mutant de XPG entraîne une dissociation de la sous-unité CAK de TFIIH. Aucune donnée n'est toutefois disponible quant à son rôle dans la NER. Par contre, cette étude ne tient pas en compte que les NLS sont probablement dans cette partie de XPG. Il y aurait un NLS entre les résidus 1051 et 1084 et un autre, dont la séquence est conservée chez l'humain et la levure, entre les résidus 1169 à 1186 (138, 206). De plus, une autre étude a démontré que le segment 981-1009 pouvait lier PCNA, sans toutefois démontrer le rôle biologique de cette interaction (136).

Une conclusion émerge de toutes ces études : l'interaction entre XPG/Rad2 et TFIIH ne peut être réduite à un seul segment de chaque côté. Ceci pose un défi considérable, puisqu'il est difficile de trouver tous les domaines d'interaction, sans compromettre la fonction de certains domaines par les techniques classiques d'interaction protéine-protéine.

Ces interactions protéine-protéine entre XPG/Rad2 et TFIIH ainsi qu'entre XPC/Rad4 et TFIIH n'étant pas bien caractérisées, nous avons entrepris de les explorer, ce qui a mené à cette thèse.

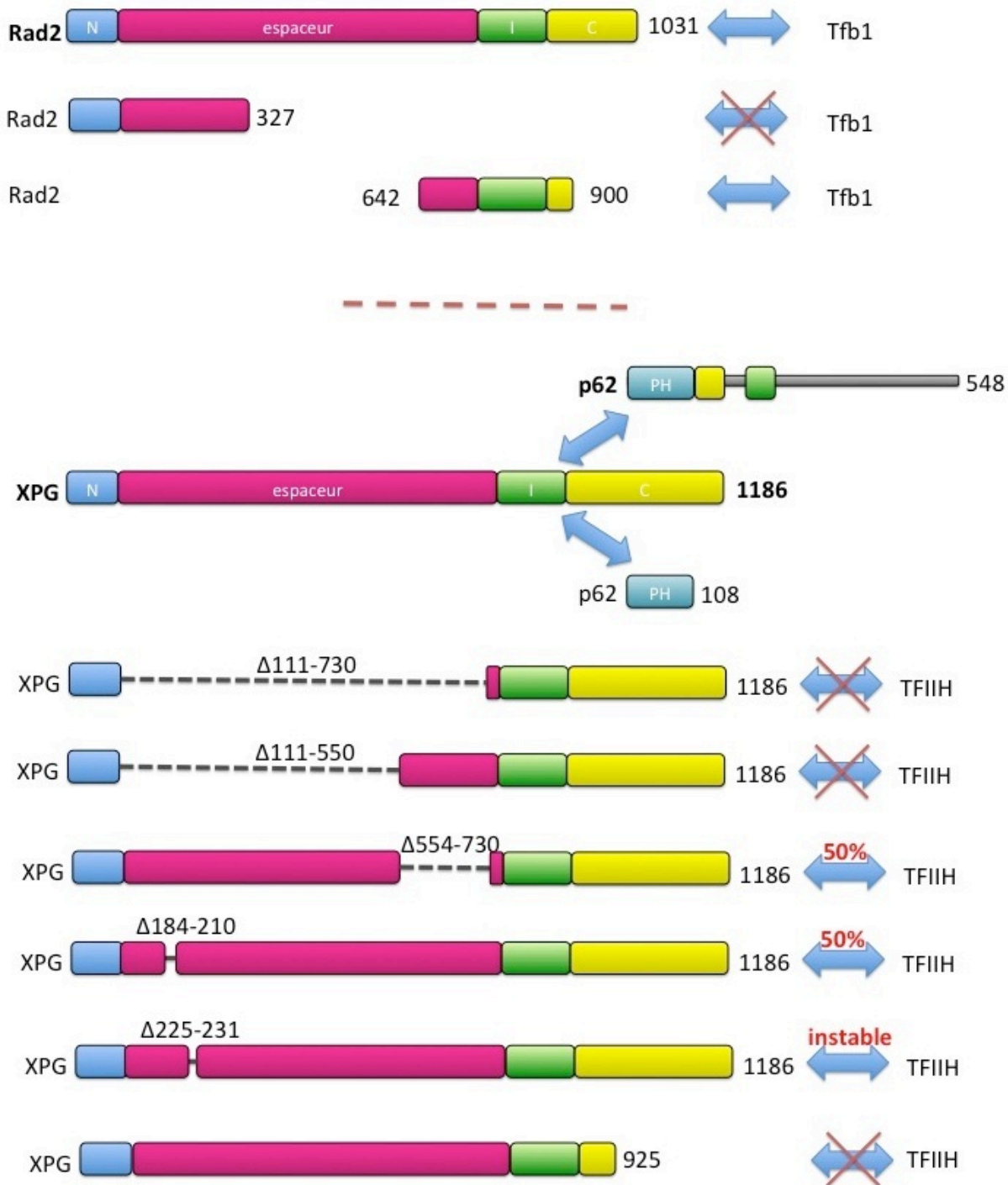


Figure 1-19. Résumé des interactions de diverses troncations de Rad2 et XPG avec TFIIH et ses sous-unités. Les segments ont été colorés comme à la Figure 1-9. Les flèches bleues représentent une interaction avec un partenaire, avec parfois une caractéristique particulière (interagissent à 50% par rapport à l'interaction avec le type sauvage, interaction mais instabilité).

1.4 Objectifs de la thèse

Plusieurs protéines responsables de la réparation de l'ADN lient l'ADN via un domaine et recrutent d'autres protéines via d'autres domaines. Par ces deux fonctions fondamentales, elles ressemblent à des facteurs de transcription activateurs. En effet, ceux-ci lient l'ADN à des sites spécifiques, puis recrutent la machinerie responsable de la transcription grâce à leur domaine de transactivation.

Tel que décrit précédemment, il existe un point commun important entre la transcription et la réparation de l'ADN: le complexe TFIIH. Or, outre sa fonction d'hélicase commune aux deux processus, existe-t-il d'autres parallèles? Est-ce que les mécanismes de recrutement à l'ADN ayant une lésion ou l'ADN devant être transcrit sont les mêmes? Notre hypothèse principale est que certaines protéines de la réparation de l'ADN possèdent des « domaines de transactivation » semblables à ceux des activateurs leur permettant de recruter leurs partenaires au site de dommage à l'ADN.

Le domaine PH de la sous-unité p62/Tfb1 de TFIIH est impliqué dans les interactions avec des activateurs lors de la transcription. Il est aussi responsable de l'interaction avec XPG/Rad2 et plus globalement, p62/Tfb1 interagit avec XPC/Rad4. XPG/Rad2 possède également des portions acides désordonnées dans son segment espaceur. Les objectifs de la thèse sont donc de :

- Caractériser les interactions entre p62/Tfb1 et XPG/Rad2
- Caractériser les interactions entre p62/Tfb1 et XPC/Rad4.
- Faire un parallèle entre ces interactions et celles qui ont lieu durant la transcription afin de dégager quel est le langage moléculaire utilisé.

C'est dans le but global de mieux comprendre le fonctionnement de TFIIH dans la réparation et la transcription que cette recherche a été effectuée. Pour ce faire, nous utiliserons une caractérisation fonctionnelle chez la levure ainsi qu'une caractérisation structurale et biophysique, par résonance magnétique nucléaire (RMN) et par calorimétrie isotherme (ITC).

1.5 Techniques utilisées

1.5.1 RMN

La biologie structurale offre plusieurs outils permettant une analyse de la structure tridimensionnelle des molécules de la vie. Ces techniques possèdent toutes des avantages et des inconvénients et fournissent des informations différentes et souvent complémentaires. Pour certaines techniques, il est possible d'observer les atomes et les acides aminés individuellement. On parle alors de résolution atomique. Pour d'autres, on ne peut distinguer que des formes générales ou au mieux des structures secondaires, à basse résolution.

La cristallographie aux rayons X est la technique la plus utilisée dans le domaine de la biologie structurale, puisqu'elle fournit une résolution atomique. Cette technique permet de résoudre la structure de molécules de toutes tailles, à condition qu'elles cristallisent, ce qui est le défi majeur de cette technique. Malheureusement, les protéines ayant des segments désordonnés sont très difficiles à cristalliser. Même si on réussit à obtenir des cristaux et à résoudre la structure générale de la protéine, ces segments sont peu ou pas visibles.

La diffusion des rayons X aux petits angles (*Small Angle X-ray Scattering*, SAXS) nécessite plutôt des macromolécules en solution, donc nul besoin de trouver des conditions de cristallisation. Par contre, la résolution est faible et ne permet pas de visualiser la structure à une résolution atomique. On obtient donc la forme générale de la molécule. La cryo-microscopie électronique est une autre technique ne nécessitant pas la cristallisation des molécules, celles-ci étant déposées sur une grille et observées par un microscope électronique. Il est parfois possible de distinguer la structure secondaire.

Finalement, la résonance magnétique nucléaire (RMN) utilise le magnétisme des noyaux afin d'effectuer des expériences diverses, comme mesurer des distances interatomiques. L'avantage majeur de cette technique est d'atteindre une résolution atomique, tout en étant en solution. Bien qu'il ne faille pas cristalliser la molécule d'intérêt, la limitation de la taille de la molécule étudiée représente un désavantage par rapport à la

cristallographie. En général, seules les molécules de moins de 30 kDa peuvent être étudiées par cette technique.

Pour cette thèse, la RMN est la technique privilégiée puisqu'une résolution atomique est nécessaire et que les molécules étudiées sont de faible taille. Également, la flexibilité et le désordre intrinsèque des protéines étudiées les rendent difficiles à cristalliser. De plus, la faible taille des complexes que nous étudions et le besoin d'une haute résolution rendent les analyses de SAXS et de cryo-microscopie électronique peu utiles. Survolons maintenant les bases de la RMN.

Phénomène de résonance magnétique

Le noyau des atomes possède une masse, une charge, mais aussi une propriété quantique appelée spin (207). Selon la nature de l'atome, les spin I peuvent prendre une valeur nulle ($I=0$), entière ($I = 1, 2, 3$, etc) ou fractionnaire ($I= \frac{1}{2}, \frac{3}{2}, \frac{5}{2}$, etc). Les atomes ayant un spin nul ne peuvent être détectés alors que la plupart des autres types de spin créent des signaux complexes à analyser. C'est majoritairement le spin $I=\frac{1}{2}$ qui nous intéresse en RMN des protéines.

L'atome le plus utilisé en RMN est l'hydrogène (proton), puisqu'il possède un spin $I=\frac{1}{2}$. Parmi les atomes composant les molécules de la vie, peu d'entre eux ont un spin $I=\frac{1}{2}$. Ceci n'est vrai que pour leurs isotopes ayant la plus grande abondance naturelle. Par exemple, l'azote 14 (^{14}N) a une abondance naturelle de 99,63% et possède un spin $I=1$ tandis que l'azote 15 (^{15}N) est peu abondant dans la nature, mais possède un spin $I=\frac{1}{2}$. Le carbone 12 (^{12}C) possède lui aussi une abondance naturelle importante (98,89%), mais à cause de son spin $I=0$, il ne peut être détecté. Le carbone 13 (^{13}C) possède quant à lui un spin $I=\frac{1}{2}$. Ainsi, pour détecter ces atomes dans les molécules d'intérêt, il faudra procéder à un enrichissement isotopique.

Pour expliquer le phénomène de résonance magnétique, prenons l'exemple du proton, puisqu'il s'agit de l'atome le plus détecté en RMN. Si on place une molécule dans un champ magnétique externe, les spins possèdent deux valeurs : dans le sens du champ magnétique ($+\frac{1}{2}$) et à l'inverse du champ ($-\frac{1}{2}$). En stimulant les noyaux au moyen de

radiofréquences, il est possible, grâce au phénomène de résonance, de faire changer les spins d'un état de basse à haute énergie. Après l'émission de ces fréquences, le retour à l'équilibre est mesuré. La différence d'énergie entre les deux états est proportionnelle au champ magnétique : plus le champ est grand, plus il est facile de détecter cette différence. C'est la raison pour laquelle les spécialistes de la RMN utilisent des spectromètres de plus en plus puissants (un spectromètre de 800MHz est plus puissant qu'un 500MHz).

La fréquence de résonance de chaque proton dépend de son environnement chimique. Mesurée en Hertz (Hz), elle est convertie en partie par million (ppm) en prenant en considération la puissance du spectromètre. Ainsi, d'un spectromètre à un autre, la mesure est la même. On parle alors de déplacement chimique (δ , *chemical shift*).

Expérience 1D

L'expérience RMN la plus simple est l'enregistrement d'un spectre de protons d'une dimension (1D). Tous les protons sont excités simultanément et on mesure l'ensemble de leur relaxation que l'on traite grâce à une transformée de Fourier. Ceci donne un spectre 1D où les déplacements chimiques sont en abscisse tandis que l'ordonnée représente l'intensité du signal (Figure 1-20A). L'intégration de ces signaux permet de déterminer la quantité relative de protons contribuant aux signaux. Ceci nous permet de voir l'éventail des déplacements chimiques de tous les protons. Or, pour une protéine, les protons ont des déplacements chimiques caractéristiques. Par exemple, les protons α ont un déplacement chimique autour de 4 à 6 ppm, alors que pour les protons amides, la valeur est autour de 7 à 10 ppm. Sachant qu'une protéine peut contenir autant protons amides que d'acides aminés, une protéine-type possède plusieurs dizaines de signaux autour de 7 à 10 ppm. La résolution des spectres n'étant pas très élevée, il est difficile de séparer les signaux, ce qui donne lieu à des chevauchements, d'où l'intérêt de la spectroscopie multidimensionnelle.

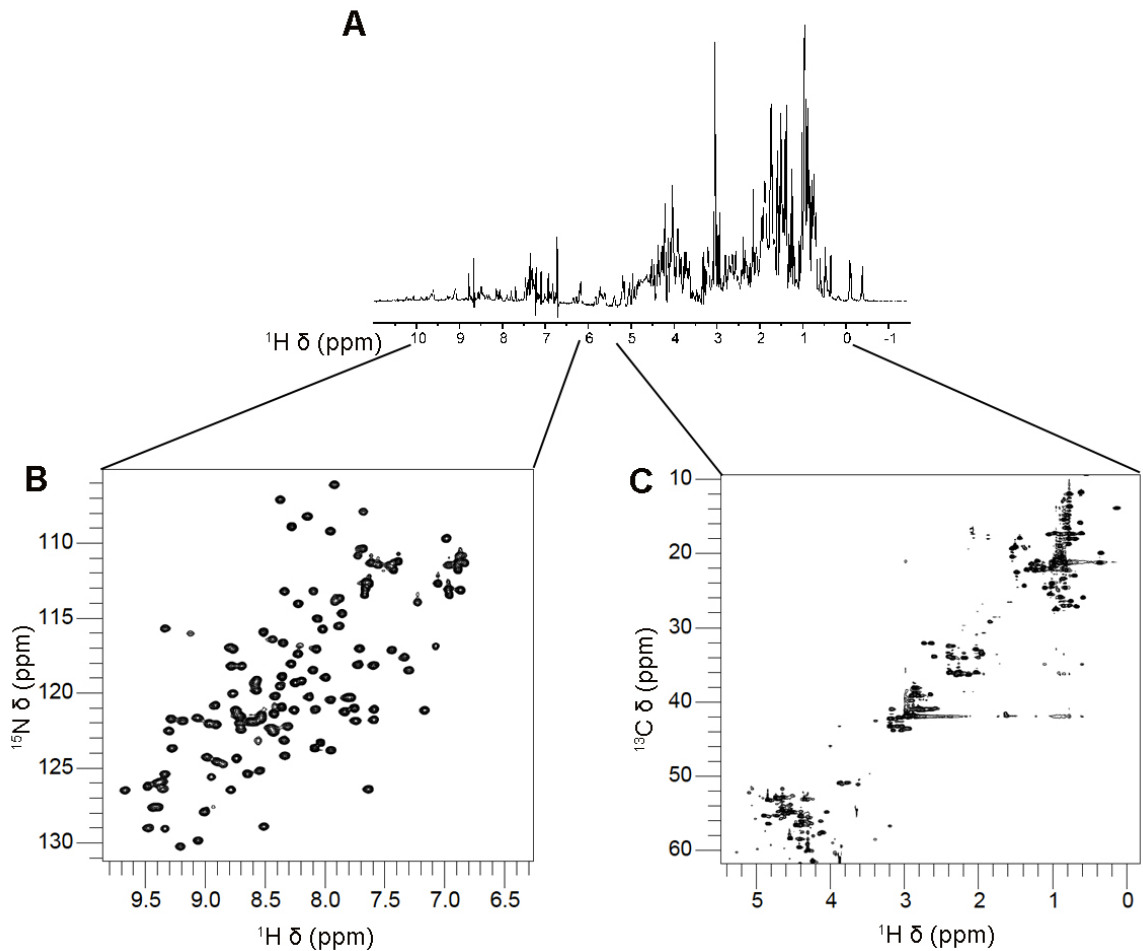


Figure 1-20. Spectres RMN 1D et 2D. A) Spectre 1D des protons. B) Spectre 2D ^1H - ^{15}N HSQC. C) Spectre 2D ^1H - ^{13}C HSQC.

Expériences multidimensionnelles

Par des séquences complexes de pulses, il est possible de transférer le magnétisme d'un atome à un autre à travers les liens chimiques ou à travers l'espace. Cette propriété permet de réaliser des expériences plus complexes ayant plusieurs dimensions. Par exemple, une expérience très utile en RMN de protéines est le ^1H - ^{15}N HSQC (*Heteronuclear Single Quantum Coherence*) (Figure 1-20B). Par cette expérience, un signal est observé pour chacune des paires d'atomes H-N liés par un lien covalent. Le déplacement chimique des protons est en abscisse alors que celui des azotes est en ordonnée. L'intensité du signal est schématisée comme une carte topographique. Dans un spectre ^1H - ^{15}N HSQC, on observe tous les protons amides (un signal pour chaque résidu,

sauf les prolines). Également, une expérience similaire est possible pour les protons aliphatiques où un signal est observé pour chacune des paires d'atomes H-C liés par un lien covalent. Le spectre ^1H - ^{13}C HSQC possède beaucoup plus de signaux qu'un spectre ^1H - ^{15}N HSQC (Figure 1-20C). Ces expériences nécessitent que la protéine soit marquée en ^{15}N ou doublement marquée en ^{15}N et ^{13}C .

Ces HSQC nous permettent d'effectuer des titrages. Suite à l'ajout d'un ligand non-marqué à une protéine marquée, il est possible de voir les signaux d'un HSQC se déplacer par rapport au HSQC de la protéine marquée libre. En effet, la liaison provoque un changement d'environnement chimique des atomes impliqués dans la liaison, ce qui crée une différence dans leurs déplacements chimiques. Ceci permet d'établir si une protéine et un ligand interagissent, mais surtout, ceci permet de connaître le site de liaison du ligand. Toutefois, il faudra effectuer des expériences d'attribution pour savoir à quels atomes chaque signal appartient.

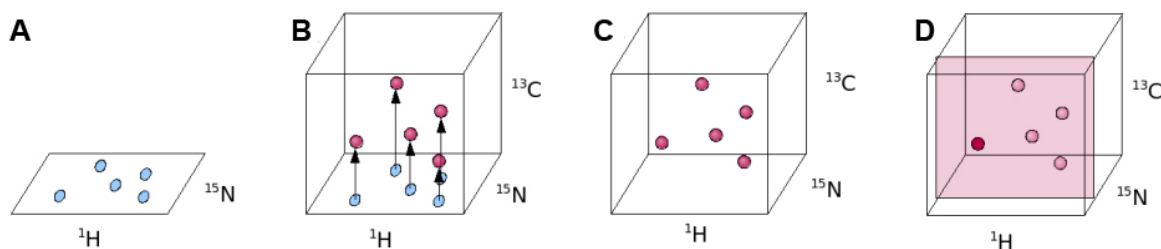


Figure 1-21. Spectre RMN 3D. A) La base de toute expérience 3D est une expérience 2D (ici un ^1H - ^{15}N HSQC). B et C) Le spectre 2D est développé dans une 3^e dimension. D) On observe ce type de spectre en regardant un plan coupant l'ensemble du spectre. Ainsi, c'est une série de spectres 2D qui est observée. (Adaptation du site <http://www.protein-nmr.org.uk/spectra.htm>)

Expériences d'attribution

Les expériences d'attribution sont multidimensionnelles, le plus souvent à trois dimensions. Toutes les expériences présentées ici ont pour base le ^1H - ^{15}N HSQC ou le ^1H - ^{13}C HSQC, qui se développent dans une troisième dimension (Figure 1-21). Les stratégies d'attribution diffèrent d'un laboratoire à un autre. La stratégie générale présentée ici est

celle employée au laboratoire Omichinski (voir le Tableau 1 pour une description sommaire des spectres utilisés).

Tableau 1. Expériences permettant l'attribution des déplacements chimiques des atomes d'une protéine.

<i>Expérience</i>	<i>Corrélation</i>	<i>Atome attribué</i>
3D HNCACB	H, N (amides) du résidu <i>i</i> avec le C α et le C β des résidus <i>i</i> et <i>i-1</i>	C α et C β pour l'attribution séquentielle (H et N amides)
3D CBCA(CO)NH	H, N du résidu <i>i</i> avec le C α et le C β du <i>i-1</i>	C α et C β pour l'attribution séquentielle (H et N amides)
3D HNCO	H, N du résidu <i>i</i> avec le C carbonyle du résidu <i>i-1</i>	C carbonyle
3D (H)C(CO)NH	H, N du résidu <i>i</i> avec tous les carbones de la chaîne latérale du résidu <i>i-1</i>	C aliphatiques (chaîne latérale)
3D H(CCO)NH	H, N du résidu <i>i</i> avec tous les protons de la chaîne latérale du résidu <i>i-1</i>	H aliphatiques (chaîne latérale)
2D (HB)CB(CGCD)HD & (HB)CB(CGCDCE)HE	H δ et H ϵ avec le C β d'un même résidu	H aromatiques

1. D'abord, nous cherchons à déterminer à quoi correspondent chacun des signaux du ^1H - ^{15}N HSQC grâce au HNCACB (208) et au CBCA(CO)NH (209). Ceux-ci nous permettent d'abord de connaître l'ordre des signaux, puis on peut assigner les signaux si on connaît la séquence et quelques indices. Par exemple, les glycines n'ont qu'un signal C α , les sérines et thréonines ont un déplacement chimique de leur C β particulier, etc.

2. Par la suite, le carbone carbonyle peut être attribué grâce au spectre HNCO (210) et les carbones des chaînes latérales grâce au 3D (H)C(CO)NH (211).

3. Finalement, on déterminera le déplacement chimique des hydrogènes des chaînes latérales avec le H(CCO)NH (211). Les déplacements chimiques des protons des chaînes latérales des résidus aromatiques sont trouvés à l'aide des (HB)CB(CGCD)HD et (HB)CB(CGCDCE)HE (212). Toutes ces données sont vérifiées grâce à une expérience

appelée spectre de corrélation (HCCH-COSY) (213). Basé sur le ^1H - ^{13}C HSQC, ce spectre corréle une paire H-C avec un autre proton situé à 2 liens covalents du carbone.

Toutes ces expériences ont comme but ultime de déterminer le déplacement chimique de chaque atome d'une protéine dans le but de mesurer des distances par l'effet nucléaire Overhauser (*Nuclear Overhauser Effect*, NOE).

Expériences NOESY

Le NOE est un transfert de magnétisation entre deux atomes dans l'espace. L'intensité d'un NOE est inversement proportionnel à la distance entre ces deux atomes, à la puissance 6 ($1/r^6$ où r est la distance). Il est donc possible de déterminer des milliers de distances dans une protéine. Deux spectres sont enregistrés pour la résolution d'une protéine. Le ^{15}N -edited NOESY-HSQC a pour base le ^1H - ^{15}N HSQC et tous les NOE mesurés le sont à partir du proton amide seulement. Le ^{13}C -edited HMQC-NOESY a pour base un équivalent du ^1H - ^{13}C HSQC et mesure les NOE à partir des proton liés à des carbones (214). En vue du calcul de structure, on sépare les intensités en grandes catégories, puis on leur attribue une fourchette de distance (ex : 1.8 Å à 2.8Å).

Pour un complexe composé de deux protéines, on mesure les NOE des deux protéines séparément (A marquée, protéine B non-marquée et vice-versa), puis on mesurera des NOE intermoléculaires grâce au spectre 3D $^{15}\text{N}/^{13}\text{C}$ (F1)-filtered, (F3)-edited NOESY (215, 216). On doit donc faire toutes les expériences en double avant de passer au calcul des structures.

Calculs de structures

Résoudre une structure par RMN est une tâche colossale. Les NOE sont la majeure partie des contraintes servant au calcul, mais il est possible d'utiliser d'autres contraintes. Premièrement, les déplacements chimiques servent à prédire les angles diédraux phi et psi dans les parties de la protéine ayant des structures secondaires précises (hélice- α et feuillet- β) (217). Les ponts-hydrogène peuvent également être déterminés grâce au ^1H - ^{15}N HSQC. Suite au remplacement de l'eau par du D_2O comme solvant, plusieurs signaux du ^1H - ^{15}N HSQC disparaissent, puisque les protons amides échangent avec le solvant et le deutérium

n'est pas détecté en RMN. Certains protons prennent plus de temps à échanger avec le solvant, puisqu'ils forment un pont-hydrogène, les protégeant de l'échange. Finalement, d'autres types de contraintes peuvent être ajoutés, tels que des couplages dipolaires résiduels.

Toutes ces contraintes sont données à un logiciel de calcul de structures qui tentera de minimiser l'énergie du système. Cette énergie est calculée à partir des contraintes fournies, mais aussi à partir des contraintes physiques des acides aminés et de la chimie des protéines (longueur des liens, volumes de van der Waals, angles formés par les liens, etc). Pour minimiser l'énergie, la plupart des logiciels utilisent le recuit simulé (*simulated annealing*), qui est une méthode d'exploration de l'espace conformationnel (218). Ces calculs donnent lieu à une centaine de structures répondant à toutes les contraintes. Ces structures sont triées en ordre d'énergie et celles possédant la plus basse énergie sont choisies pour représenter les données. Nous sommes alors devant un ensemble de structures répondant aux contraintes fournies au logiciel.

Contrairement à la cristallographie aux rayons X où les facteurs R et R_{free} sont utilisés pour valider le modèle par rapport aux données expérimentales, il n'existe pas de tels critères qui font l'unanimité dans la communauté RMN. Par contre, certains critères sont utilisés pour valider la qualité du modèle, sans égard aux données enregistrées. D'abord, la géométrie de la chaîne principale est vérifiée par le diagramme de Ramachandran. Il se peut parfois que les segments désordonnés divergent des valeurs idéales. Les longueurs des liens et les angles entre les atomes sont également vérifiés par rapport à leur valeur idéale. Finalement, la précision des structures du modèle est mesurée par la racine de la moyenne des déviations au carré (*Root Mean Square Deviation, RMSD*). Ceci nous permet de savoir à quel point les structures sont semblables entre elles. Ici encore, des structures plutôt désordonnées auront un RMSD plus élevé que les structures ordonnées, ce qui est normal étant donné la diversité des structures.

1.5.2 ITC

Toute réaction chimique absorbe (endothermique) ou dégage (exothermique) de la chaleur. Il en va de même pour les réactions d'interaction entre les macromolécules (liaisons non-covalentes). Toutefois, la chaleur transférée lors de la formation ou le bris telles liaisons est habituellement dans un ordre de grandeur inférieur à celle transférée au cours de réactions chimiques. C'est pourquoi sa mesure nécessite l'emploi d'un micro-calorimètre, pouvant détecter un signal de l'ordre des micro-calories.

De plus, les protéines sont des molécules chimiques très sensibles aux changements de température. L'utilisation d'un micro-calorimètre isotherme est donc de mise (Figure 1-22). Contrairement aux calorimètres conventionnels qui mesurent un changement de température, le micro-calorimètre isotherme mesure l'énergie nécessaire pour garder la température du système constante (219).

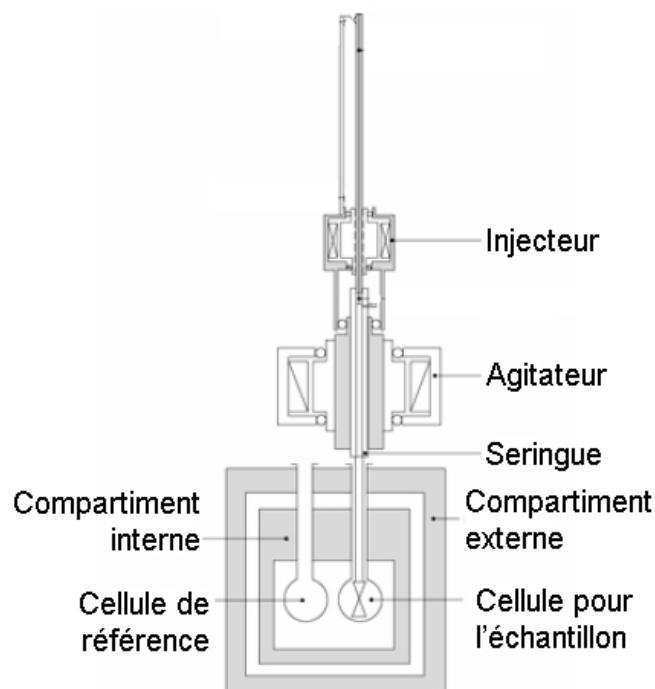


Figure 1-22. Vue schématique d'un micro-calorimètre. Les compartiments externe et interne servent à réguler la température des cellules de référence (blanc) et la cellule pour l'échantillon. La seringue est agitée constamment au cours de l'expérience, afin d'avoir une répartition uniforme des molécules injectées.

Plusieurs types d'expériences peuvent être ainsi faites, mais celle qui nous intéresse est la calorimétrie à titrage isotherme (*Isothermal Titration Calorimetry*). Lors d'une telle expérience, les deux molécules dont la liaison est à étudier sont séparées, l'une dans une cuvette maintenant la température constante, l'autre dans une seringue d'injection. L'injection de petites quantités de cette protéine dans la cuvette à des intervalles réguliers permet de mesurer la chaleur transférée à chaque ajout. Au fur et à mesure que les sites de liaison de la molécule contenue dans la cuvette se saturent, la chaleur transférée diminue, jusqu'à ce que plus aucune chaleur ne soit transférée. Le système est alors à l'équilibre.

Il est donc possible de tracer un tel thermogramme et sa courbe de titrage (Figure 1-23). Par la suite, une équation complexe est utilisée afin d'extraire trois paramètres : l'enthalpie (ΔH), le nombre de sites de liaison (N) ainsi que la constante d'association K_a . Finalement, il est possible de déterminer la valeur d'entropie (ΔS) grâce aux équations thermodynamiques suivantes : $\Delta G = \Delta H - T\Delta S$ $\Delta G = -R T \ln K_a$ où ΔG est l'énergie libre de Gibbs, T est la température et R est la constante des gaz parfaits. Pour la présente étude, c'est surtout la constante d'association K_a qui nous intéresse, puisqu'on pourra y dériver la constante de dissociation K_D par la simple formule $K_D = 1/K_a$.

Les limites de détection de cette technique sont des valeurs de K_D se situant entre le nanomolaire et le micromolaire. De plus, contrairement à la résonance des plasmons de surface (*Surface Plasmon Resonance*, SPR) où il faut fixer le récepteur sur une surface, l'expérience ITC se déroule en solution, ce qui constitue un avantage majeur de la technique (220).

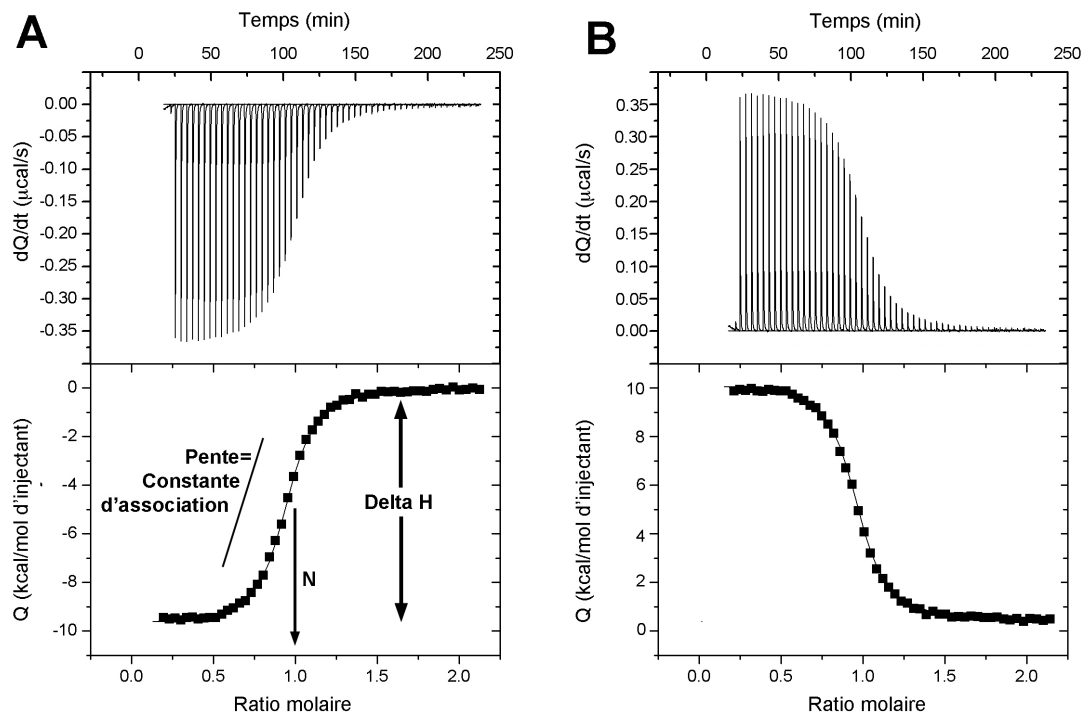


Figure 1-23. Thermogrammes d'expériences d'ITC A) Exemple de réaction exothermique. Dans ce schéma, l'enthalpie (ΔH), le nombre de sites (N) et la pente (représentant la constante d'association K_a) sont annotés. B) Exemple de réaction endothermique.

2.0 Caractérisation de l'interaction entre Rad2 et Tfb1

Nous nous sommes d'abord intéressés à l'interaction entre TFIIH et XPG/Rad2. Notre but était de trouver précisément la paire de segments d'interaction, pour ensuite la caractériser. Nous avons choisi d'utiliser les protéines de levure, car nous pouvions ainsi effectuer plus facilement des expériences *in vivo*. Chez la levure, le segment de Rad2 interagissant avec TFIIH a déjà été délimité à la région 642-900 (202). Cette région fait partie du segment espaceur qui est désordonné et acide, ce qui est une bonne indication de la présence d'un segment d'interaction. Finalement, les segments d'interaction de p53, VP16 et TFIIE α à p62/Tfb1 pouvaient nous aider encore plus dans cette quête.

Par contre, il n'avait pas encore été prouvé que chez la levure, le domaine PH de Tfb1 était responsable de l'interaction avec Rad2. Chez l'humain, l'interaction du domaine PH de p62 avec XPG et son importance pour la NER ont été démontrées. Dans le présent chapitre, nous avons donc démontré que le domaine PH de Tfb1 est important pour la NER, pour ensuite préciser le segment d'interaction avec Rad2. Finalement, nous avons résolu la structure du complexe, par RMN.

Pour cet article, j'ai effectué les expériences *in vitro* suivantes : purifications de protéines, expériences d'ITC, enregistrement des spectres RMN, interprétation de ces spectres, extraction des NOE et calcul des structures. Geneviève Arseneault, assistante de recherche au laboratoire, a effectué les expériences *in vivo* chez la levure alors que Laurent Cappadocia, stagiaire post-doctoral, a purifié le segment 359-383 de Rad2 et m'a assisté dans les expériences ITC associées à ce segment. Pascale Legault m'a aidé dans l'enregistrement des spectres RMN pour l'attribution séquentielle et les spectres NOE. Les expériences ont été planifiées par moi-même, Geneviève Arseneault et le Dr James Omichinski, directeur du laboratoire. Nous avons également tous les trois rédigé l'article qui sera soumis sous peu à *Nature Structural and Molecular Biology*.

**Structural and functional characterization of
interactions involving the Tfb1 subunit of TFIID and the
NER factor Rad2**

Julien Lafrance-Vanasse^{1}, Geneviève Arseneault^{1*}, Laurent Cappadocia¹, Hung-Ta Chen²,
Pascale Legault¹ and James G. Omichinski^{1*}*

¹Département de Biochimie, Université de Montréal, Montréal, Québec, Canada

²Institute of Molecular Biology, Academia Sinica, Taipei, Taiwan, China

Running Title: Characterization of Rad2 interaction with Tfb1.

ABSTRACT

The general transcription factor IIIH (TFIIH) plays crucial roles in transcription as part of the pre-initiation complex (PIC) and in DNA repair as part of the nucleotide excision repair (NER) machinery. During NER, TFIIH recruits the 3'-endonuclease Rad2 to damaged DNA. In this manuscript, we functionally and structurally characterized the interaction between the Tfb1 subunit of TFIIH and Rad2. We show that deletion of either the PH domain of Tfb1 (Tfb1PH) or several segments of the Rad2 spacer region yield yeast with enhanced sensitivity to UV irradiation. Isothermal titration calorimetry studies demonstrate that two acidic segments of the Rad2 spacer bind to Tfb1PH with nanomolar affinity. Structure determination of a Rad2-Tfb1PH complex indicates that Rad2 binds to TFIIH using a similar motif as TFIIIE α uses to bind TFIIH in the PIC. Together, these results provide a mechanistic bridge between the role of TFIIH in transcription and DNA repair. **(148 Words)**

INTRODUCTION

The general transcription factor IIIH (TFIIH) plays crucial roles in both transcription as part of the pre-initiation complex (PIC) and in DNA repair as part of the nucleotide excision repair (NER) machinery¹. The human TFIIH complex and the highly homologous budding yeast (*Saccharomyces cerevisiae*) counterpart are composed of 10 subunits that can be divided in two sub-complexes, the core TFIIH complex [XPB/Ssl2, XPD/Rad3, p62/Tfb1, p52/Tfb2, p44/Ssl1, p34/Tfb4 and TTDA/Tfb5 (human/yeast)] and the CAK complex (cdk7/Kin28, cyclin H/Ccl1 and MAT1/Tfb3). As a component of the PIC, TFIIH is the only general transcription factor (GTF) to possess enzymatic activity. TFIIH helicase activities (XPB/Ssl2 and XPD/Rad3) are essential to the formation of the open complex during initiation², whereas its kinase activity (cdk7/Kin28) is required for the phosphorylation of the carboxyl-terminal domain (CTD) of RNA polymerase II (RNAP II)³. As part of the NER machinery, TFIIH helps recruit other repair factors to the damaged DNA⁴ and through its helicase activity assists in the elimination of helix-distorted DNA typically caused by UV-induced modifications of bases⁵.

NER can be divided into two distinct pathways that differ in their mechanism of initial recognition of the damaged DNA⁶. In transcription-coupled NER (TC-NER), RNAP II stalls on the damaged DNA site and recruits the Cockayne Syndrome group B protein (CSB/Rad26), whereas in global genome NER (GG-NER), the damaged site is recognized by the XPC-RAD23B/Rad4-Rad23 complex in combination with UV-damaged DNA-binding proteins (UV-DDBs). Following recognition of the damaged DNA site, the two NER pathways employ a series of common steps that include recruitment of TFIIH to the lesion to unwind the DNA, displacement of either the XPC-RAD23B or RNAPII-CSB complex by the 3'-endonuclease XPG/Rad2, recruitment of XPA/Rad14 and recruitment of the 5' endonuclease complex ERCC1-XPF/Rad10-Rad1. In both NER pathways, TFIIH functions by recruiting and stabilizing XPG/Rad2 on the damaged site⁷.

XPG/Rad2 is a member of the flap-endonuclease (FEN) family of single-stranded DNA endonucleases that includes FEN-1 (Flap endonuclease-1), EXO-1 (Exonuclease-1) and GEN-1 (gap endonuclease-1)⁸. Based on sequence alignment, the N (N-terminal) and I (internal) regions are highly conserved in all family members, and a variable-length spacer region (usually 20-70 residues) separates these two conserved regions. The crystal structure

of a FEN-1-DNA complex demonstrated that the N and I region come together to form a single structural domain that serves as the catalytic core of the enzyme⁹. XPG/Rad2 is an atypical member of the FEN family due to the fact that it contains an extended-spacer of over 600 amino acids. This extended spacer of XPG/Rad2 is highly acidic and is mostly disordered based on secondary structure predictions¹⁰. This intrinsically disordered nature of the spacer region of Rad2 enables it to participate in protein-protein interactions with several different proteins including multiple subunits of TFIIH¹⁰⁻¹³.

The interaction between TFIIH and XPG/Rad2 is essential in NER for both the recruitment of XPG/Rad2 to the repair complex and for the stabilization of the repair complex on the damaged DNA¹⁴. Mutations in both XPG and several TFIIH subunits are associated with DNA-repair associated diseases such as Xeroderma Pigmentosum (XP) and Cockayne Syndrome (CS)¹⁵. Formation of the TFIIH-XPG/Rad2 complex involves multiple regions of XPG/Rad2 and several subunits of TFIIH including p62/Tfb1, XPB/Ssl2, XPD/Rad3 and p44/Ssl1^{12,16}. In the case of XPG, residues between 184-210, 225-231, 554-730 in the spacer region and residues 1012-1186 in the C terminus have all been shown to play a role stabilizing the interaction with TFIIH and for optimal repair of DNA damage^{10,11}. Based on these results, it was postulated that different subunits of TFIIH form a series of interactions with XPG/Rad2¹⁰.

Several studies have shown that the p62/Tfb1 subunit of TFIIH plays a role in DNA-damage repair. Deletion of the extreme carboxyl-terminal (C-terminal) region (residues 532-642) of Tfb1 leads to a yeast mutant (*tfb1-1*) with decreased resistance to both temperature and UV irradiation¹⁷. Other studies showed that Tfb1 directly interacts with Rad2, and that this interaction requires an acid-rich segment within the extended spacer region¹³. In addition, the Pleckstrin homology (PH) domain of human p62 (residues 1-108) directly binds to XPG, and deletion of this domain decreases the activity of XPG in an *in vitro* repair assay¹⁸. The PH domains of p62 and Tfb1 (p62PH/Tfb1PH) are highly homologous and have been shown to bind acidic-rich domains present in several transcriptional regulatory proteins, including the large subunit of the general transcription factor IIE (TFIIE α)¹⁹, the tumor suppressor protein p53²⁰ and the Herpes Simplex Virion protein 16 (VP16)²¹.

In this paper, we evaluate the functional and structural roles of the PH domain of Tfb1 (Tfb1PH) in UV-damage repair. We demonstrate that deletion of Tfb1PH (residues 1-115) yields a yeast phenotype with a decreased resistance to UV irradiation. We then identify two acidic stretches within the spacer region of Rad2 (Rad2₃₅₉₋₃₈₃ and Rad2₆₄₂₋₆₉₀) that bind to Tfb1PH with nanomolar affinity. In addition, deletion of these two acidic segments in combination with a deletion corresponding to residues 225-231 of XPG (Rad2₂₂₈₋₂₃₇) enhances the photosensitivity of yeast¹¹. We also determine the NMR structure of a complex between Tfb1PH and Rad2₆₄₂₋₆₉₀ and show that Rad2 binds to Tfb1PH in a very similar manner as TFIIE α binds to the p62PH²². These results indicate that the recruitment of Rad2 to the TFIIH complex requires multiple interactions within the spacer region, and that Rad2 binds to TFIIH by a similar mechanism as TFIIE α binds TFIIH in the PIC.

RESULTS

***tfb1-ΔPH* yeast display enhanced sensitivity to UV damage.**

To investigate the *in vivo* role of the PH domain of the Tfb1 (Tfb1PH) in repair of damaged DNA, we constructed a *tfb1-ΔPH* mutant strain (deletion of residues 1-115). First, we tested the sensitivity of the *tfb1-ΔPH* yeast to UV irradiation and compared it to the wild-type *TFB1* and the *tfb1-1* yeast strains. The *tfb1-1* mutant yeast strain serves as our positive control as it has been shown to have decreased resistance to UV irradiation¹⁷. The survival curves show that both *tfb1-ΔPH* and *tfb1-1* are significantly more sensitive to UV irradiation than the *TFB1* wild-type strain (**Fig. 1a**). The photosensitivity of the *tfb1-ΔPH* yeast to UV irradiation is not due to a decrease in protein levels as the Tfb1-ΔPH and Tfb1 proteins are expressed at similar levels (**Supplementary Fig. 1a**). Next, we performed drop tests to evaluate the ability of *tfb1-ΔPH* to repair DNA damage induced by bleomycin. The *tfb1* wild-type strain, the *tfb1-ΔPH* mutant strain and a *rad2* strain (*RAD2* knockout) all grow similarly in either the presence or absence of bleomycin and the only strain displaying enhanced sensitivity to bleomycin is the *rad52* (*RAD52* knockout) positive control (**Fig. 1b**). Taken together, these results suggest that Tfb1PH plays a specific role in NER induced by UV irradiation.

***tfb1-ΔPH* yeast show normal growth at 37°C.**

In addition to displaying a UV sensitivity phenotype, the *tfb1-1* yeast are also sensitive to growth at high temperatures¹⁷. Therefore, we tested whether or not *tfb1-ΔPH* yeast also display a similar growth phenotype. The *TFB1*, *tfb1-ΔPH* and *tfb1-1* yeast were all grown at both 30°C and 37°C, and the drop test shows that *tfb1-1* is the only strain sensitive to growth at 37°C (**Fig. 1c**). The *tfb1-ΔPH* and the *TFB1* strains grow similarly when incubated at 37°C, indicating that the PH domain of the Tfb1 protein is not required for growth at higher temperatures. This further supports that the sensitivity to UV irradiation observed with the *tfb1-ΔPH* strain is due to a loss of function associated with

removal of Tfb1PH and not to instability associated with the protein as seen with the truncated protein expressed by *tfb1-1* yeast at 37°C.

An acid-rich segment of the Rad2 spacer region binds Tfb1PH with high affinity.

Previous studies have shown that Rad2 interacts with Tfb1 and residues 642 to 900 of Rad2 are sufficient for binding¹³. This Tfb1-binding region of Rad2 includes an acid-rich segment from the spacer region (residues 642-760) that is required for the interaction¹³. Based on these results and the fact that p62PH was shown to be required for interaction with XPG¹⁸, we were interested to determine if the acid-rich segment between residues 642-760 of Rad2 binds directly to Tfb1PH. To test this, the apparent dissociation constant (K_d) for the interaction between Tfb1PH and Rad2₆₄₂₋₇₆₀ was determined by isothermal titration calorimetry (ITC) experiments. The ITC experiments show there is an interaction between these two protein segments, but the stoichiometry of the binding ($N= 0.52$) suggests that Rad2₆₄₂₋₇₆₀ contains two distinct Tfb1PH-binding sites (**Supplementary Fig. 2a**).

Based on sequence comparison with known Tfb1PH/p62PH-binding sites from the C-terminal domain of TFIIIE α (TFIIIE α CTD)²² and the transactivation domains (TADs) of p53 (p53TAD2)²⁰ and VP16 (VP16C)²¹, we identify three segments within Rad2₆₄₂₋₇₆₀ (residues 661-681, 708-728 and 718-738) that could potentially contain a Tfb1PH-binding site (**Fig. 2**). Preliminary NMR studies show that Rad2₆₄₂₋₇₆₀ is disordered in the unbound form as predicted (**Supplementary Fig. 2b**). Given the fact that Rad2₆₄₂₋₇₆₀ is disordered and that two of the potential Tfb1PH-binding sites overlapped (650-670 and 661-681), we chose to partition Rad2₆₄₂₋₇₆₀ into two segments (Rad2₆₄₂₋₆₉₀ and Rad2₆₉₂₋₇₆₀) to determine the K_d values of the individual segments (**Fig. 2b**). By ITC, we determine that Rad2₆₄₂₋₆₉₀ ($K_d = 190$ nM) binds with much higher affinity to Tfb1PH than Rad2₆₉₂₋₇₆₀ ($K_d = 4.6$ μ M). These results support the presence of two binding sites and indicate that Tfb1PH binds preferentially to the segment Rad2₆₄₂₋₆₉₀.

The Tfb1PH/p62PH binding sites from TFII α CTD, p53TAD2 and VP16C all contain hydrophobic residues that are crucial for forming the interaction interface¹⁹⁻²¹. Based on sequence alignment, we postulate that Phe670 and Val673 of Rad2₆₄₂₋₆₉₀ would play an important role in forming the interface between Rad2₆₄₂₋₆₉₀ and Tfb1PH. To test the importance of these two hydrophobic residues, we prepared proline mutants (F670P and V673P) and measured their binding to Tfb1PH by ITC. The ITC studies show that neither the F670P nor V673P mutant of Rad2₆₄₂₋₆₉₀ bind with appreciable affinity.

Rad2₆₄₂₋₆₉₀, p53TAD2 and TFII α CTD bind a common site on Tfb1PH

To identify the binding site for Rad2₆₄₂₋₆₉₀ on Tfb1PH, NMR chemical shift perturbation studies were performed. In these experiments, addition of unlabeled Rad2₆₄₂₋₆₉₀ to ¹⁵N-labeled Tfb1PH cause significant changes in the ¹H and ¹⁵N chemical shifts for several Tfb1PH signals in the ¹H-¹⁵N HSQC spectra (**Supplementary Fig. 3a-b**). When mapped onto the structure of Tfb1PH, the residues exhibiting significant chemical shift changes are located in strands β 5, β 6, β 7 and the helix H1 (**Fig. 3a**) and the changes are very similar to those observed when p53TAD2 (**Fig. 3b**) and TFII α CTD (**Supplementary Fig. 3c**) bind to Tfb1PH.

To confirm that Rad2₆₄₂₋₆₉₀ shares a common binding site on Tfb1PH with p53TAD2 and TFII α CTD¹⁹, NMR competition experiments were performed. In the first experiment, we add a substoichiometric concentration of unlabeled Tfb1PH (0.4 mM) to a sample containing ¹⁵N-labeled p53TAD2 (0.5 mM) and, as expected, we observe significant changes in ¹H and ¹⁵N chemical shifts in the ¹H-¹⁵N HSQC spectra of p53TAD2 (**Fig. 3c**). We subsequently add an equimolar amount of unlabeled Rad2₆₄₂₋₆₉₀ (0.5 mM) to the ¹⁵N-p53TAD2-Tfb1PH sample and observe that the ¹H and ¹⁵N resonances of p53TAD2 which shift upon formation of the p53TAD2-Tfb1PH complex returned to the values of the free form of p53TAD2 (**Fig. 3d**). Taken together with previous results showing that TFII α CTD and p53TAD2 compete for binding to Tfb1PH¹⁹, these results demonstrate that TFII α CTD, p53TAD2, and Rad2₆₄₂₋₆₉₀ all share for a common binding site on Tfb1PH.

NMR structure determination of the Rad2₆₄₂₋₆₉₀-Tfb1PH complex

To structurally compare complexes of Tfb1PH involved in transcription and in DNA repair, we determined the structure of the Rad2₆₄₂₋₆₉₀-Tfb1PH complex. The structure of the Rad2₆₄₂₋₆₉₀-Tfb1PH complex (PDB code XXXX) is well defined by the NMR data (**Table 1**). The 20 lowest-energy structures (**Fig. 4a**) are characterized by good backbone geometry, no significant restraint violation and low pair wise rmsd values (**Table 1**). In complex with Rad2₆₄₂₋₆₉₀, the Tfb1PH structure is virtually identical to its free form showing a typical PH domain fold consisting of a seven-stranded β sandwich (β 1- β 7) flanked on one side by a long α helix (H1)²³. Rad2₆₄₂₋₆₉₀ exists in an extended conformation devoid of any regular secondary structural element with residues 664-678 forming the interface with Tfb1PH. This is consistent with the ¹H-¹⁵N HSQC spectra of the titration of ¹⁵N-labeled Rad2₆₄₂₋₆₉₀ with Tfb1PH as these residues undergo significant changes in their ¹H and ¹⁵N chemical shifts (**Supplementary Fig. 4**).

Rad2₆₄₂₋₆₉₀-Tfb1PH binding interface

In the complex, Rad2₆₄₂₋₆₉₀ binds in an extended form to two adjacent shallow grooves on the surface of Tfb1PH. The first groove is formed by residues Gln49, Ala50, Thr51, Pro52, Met59, Leu60, Arg61 and Met88 from strands β 5, β 6 and β 7 of Tfb1PH (**Fig. 4b**). Phe670 of Rad2 inserts into this groove where it is in position to form a cation- π interaction with Arg61. In addition, Leu665 and Leu669 of Rad2 make van der Waals interactions with Met59 and Lys57 of Tfb1PH. The second groove is composed of Leu48, Ala50, Lys101 and Gln105, Ile108, Lys112 of Tfb1PH and accommodates Val673 and Thr675 of Rad2 (**Fig. 4c**). Val673 is anchored on one side of this pocket through van der Waals interactions of its two-methyl groups and the side chains of Leu48, Ala50, Lys101 and Gln105 of Tfb1PH. Thr675 is anchored on the other side of the pocket where its methyl group interacts with the side chains of Gln105, Ile108 and Lys112.

Although the majority of the interactions within the two grooves are van der Waals contacts, an extensive series of positively charged residues on the surface of Tfb1PH

(Lys47, Lys57, Arg61, Arg86, Lys97, Lys101 and Lys112) surround the two grooves, where they function to position the negatively charged Rad2₆₄₂₋₆₉₀. The NMR structures support the formation of two potential salt bridges between acidic residues of Rad2₆₄₂₋₆₉₀ and basic residues of Tfb1PH. The first one is between Glu667 of Rad2₆₄₂₋₆₉₀ and either Arg61 or Arg86 of Tfb1PH, and the second one is between Asp672 of Rad2₆₄₂₋₆₉₀ and Lys47 of Tfb1PH.

A second acid-rich segment of the Rad2 spacer region binds to Tfb1PH

Previous studies have shown that several segments within the Rad2/XPG spacer region interact with numerous subunits of TFIIH¹⁰⁻¹³. Therefore, the remaining residues of the Rad2 spacer region (residues 100-641) were analyzed for additional Tfb1PH-binding sites based on their sequence similarity to TFIIH α CTD, p53TAD2 or Rad2₆₄₂₋₆₉₀. Through this search, one potential site was identified between residues 359-383 (Rad2₃₅₉₋₃₈₃; **Fig. 5a**), and ITC studies show that Rad2₃₅₉₋₃₈₃ binds to Tfb1PH ($K_d = 130$ nM) with a similar affinity as Rad2₆₄₂₋₆₉₀ (**Fig. 5a**). Interestingly, the predicted Tfb1PH-binding site within Rad2₃₅₉₋₃₈₃ is very similar to the Tfb1PH-binding site in Rad2₆₄₂₋₆₉₀.

To identify the mode of binding for Rad2₃₅₉₋₃₈₃ to Tfb1PH, NMR chemical shift perturbation and competition studies were performed. Addition of unlabeled Rad2₃₅₉₋₃₈₃ to ¹⁵N-labeled Tfb1PH causes significant changes in the ¹H and ¹⁵N chemical shifts for several Tfb1PH signals in the ¹H-¹⁵N HSQC spectra (**Supplementary Fig. 5**). When mapped onto the structure of Tfb1PH, the residues exhibiting significant chemical shift changes are located in strands β 5, β 6, β 7 and the helix H1 (**Fig. 5b**) and the changes are very similar to those observed with Rad2₆₄₂₋₆₉₀ (**Fig. 3a**). NMR competition experiments further demonstrate that Rad2₃₅₉₋₃₈₃ and Rad2₆₄₂₋₆₉₀ compete for binding to Tfb1PH (**Fig. 5c-d**).

TFIIH binding sites of Rad2 enhance resistance to UV irradiation

To examine the *in vivo* role of the two Tfb1PH-binding sites of Rad2 following exposure to UV irradiation, yeast mutants were created in which the key segments were deleted either alone or in combination. Initially, either residues 367-378 (*rad2-ΔD2*) or residues 642-760 (*rad2-ΔD3*) of *RAD2* were deleted and the resulting mutant strains were tested for survival to increasing doses of UV irradiation. In comparison to the wild-type *RAD2* strain, neither of the two mutant strains shows an enhancement in sensitivity to UV radiation (**Fig. 6a**). Given that both single deletions fail to induce a UV-sensitive phenotype, a third mutant was constructed in which both residues 367-378 and residues 642-760 of Rad2 (*rad2-ΔD2D3*) were deleted. When compared to the *RAD2* strain, survival curves again indicate that the *rad2-ΔD2D3* fails to induce an enhancement in sensitivity to UV irradiation (**Fig. 6a**).

Previous studies have shown that residues 225-231 within the spacer region of XPG (*XPGΔ225-231*) play an important role in binding to TFIIH, and these residues are deleted in certain patients with XP/CS¹¹. To determine if the corresponding residues of Rad2 (residues 228-237) play an important role in yeast survival following exposure to UV irradiation, a mutant in which residues 228-237 of Rad2 were deleted (*rad2-ΔD1*) was tested for its survival following exposure to UV irradiation. Consistent with what has been observed with *XPGΔ225-231* patients, the *rad2-ΔD1* yeast display a significant enhancement in sensitivity to UV irradiation in comparison to the *RAD2* strain (**Fig. 6b**). However, the *rad2-ΔD1* strain is significantly less sensitive to UV irradiation than *rad2* (**Fig. 6b**).

Since multiple segments of the Rad2 spacer region participate in binding to TFIIH, we next tested if mutations of key residues within the Tfb1PH-binding sites of Rad2 could enhance the photosensitivity of *rad2-ΔD1* strain. To do this, a yeast strain deleted of residues 228-237 in combination with proline mutations of the four key hydrophobic residues (W372P, V375P, F670P and V673P) within the Tfb1PH-binding sites of Rad2 (*rad2-ΔD1PPPP*) was prepared and tested for survival following exposure to UV irradiation. Interestingly, the *rad2-ΔD1PPPP* yeast display a significant enhancement in

sensitivity to UV irradiation in comparison to *rad2-AD1* (**Fig. 6b**). These results are consistent with the hypothesis that multiple segments of the Rad2 spacer region participate in a series of interactions with TFIIH and it is not the result of decreased levels of Rad2 as the Rad2 mutant proteins are expressed at similar levels as the wild-type Rad2 (**Supplementary Fig. 1b**).

DISCUSSION

TFIIH is unique amongst the general transcription factors in that it also plays an important role in DNA repair as a key component of the NER pathway. In NER, TFIIH serves several functions through the helicase activity of its XPD/Rad3 and XPB/Ssl2 subunits as well as through protein-protein interactions with other repair factors. It has been previously shown that multiple subunits of TFIIH interact with several other DNA repair factors including XPC/Rad4, XPG/Rad2, and CSB/Rad26, and these interactions help to stabilize the repair complex^{13,24}. Despite the importance of these protein-protein interactions, prior to this work there were no high-resolution structures reported of a complex involving one of the subunits of TFIIH and a repair factor. In this manuscript, we have functionally and structurally examined the interaction between the Tfb1 subunit of TFIIH and the repair factor Rad2 from budding yeast. We demonstrate that deletion of either the PH domain of Tfb1 or several different segments of the spacer region of Rad2 yield yeast mutants that display an enhanced sensitivity to UV irradiation. By ITC analysis, we show that two acid-rich segments of the Rad2 spacer region bind to Tfb1PH with high affinity. NMR chemical shift perturbation and competition studies indicate that the two segments of Rad2 (Rad2₃₅₉₋₃₈₃ and Rad2₆₄₂₋₆₉₀) compete for a common binding site on Tfb1PH and that this is the same site required for interaction with TFIIIE α CTD and p53TAD2. The three-dimensional structure of a complex formed by Tfb1PH and one of the acid-rich segments of Rad2 (Rad2₆₄₂₋₆₉₀) reveals that Rad2 binds to Tfb1PH in an extended form much like TFIIIE α CTD, but not in a helical structure as observed with p53TAD2 and VP16C²⁰⁻²².

Rad2 and XPG are unique members of the FEN-1 nuclease family by virtue of the fact that they directly interact with TFIIH and contain an extended spacer region (>600 amino acids) between their highly conserved N and I regions. A direct comparison of the spacer region of Rad2 and XPG is difficult since their sequences are not as highly conserved as their N and I regions, and this is consistent with the fact that both spacer regions are predicted to be intrinsically unstructured. However, the XPG and Rad2 spacer regions do share common features including a high percentage of acidic amino acids and the ability to interact with multiple subunits of TFIIH including p62/Tfb1. Based on binding studies, it has been proposed that TFIIH recruits XPG/Rad2 to the repair complex through a series of weak interactions and that the spacer region plays a key role in this recruitment¹⁰. Our results demonstrating that two segments within the spacer region of Rad2 bind the PH domain of Tfb1 are consistent with the idea of multiple interactions between Rad2 and TFIIH. Although deletion of either or both of the Tfb1PH binding sites does not directly result in a UV phenotype in yeast, mutation of the key hydrophobic residues within these binding sites enhance the sensitivity when combined with deletion of residues 228-237 that are homologous to residues 225-231 of XPG. This region of XPG has been shown to be important for binding to TFIIH and it is deleted in patients with XP-CS syndrome¹¹. It is also clear that this domain is important for repair of UV-induced DNA damage in yeast, but the exact mechanism by which this region interacts with TFIIH is currently unknown. However, our *in vivo* results with the mutations within these three segments of Rad2 in yeast are consistent with the critical role of residues 225-231 in UV repair in humans and the hypothesis that the Rad2 spacer forms multiple interactions with TFIIH that are required for NER.

It is interesting to compare our structure of the Rad2₆₄₂₋₆₉₀-Tfb1PH complex with the other structures of complexes involving Tfb1PH and p62PH. Structures have been solved with Tfb1PH bound to three acidic TAD (p53TAD2²⁰, VP16C²¹ and EKLFTAD2²⁵) and p62PH bound to TFIIIE α CTD²². p53TAD2 and VP16C form 9-residue α -helices upon binding to Tfb1PH in a coupled folding and binding mechanism. Comparison of the p53TAD2-Tfb1PH structure with the Rad2₆₄₂₋₆₉₀-Tfb1PH structure indicates that p53TAD2 and Rad2₆₄₂₋₆₉₀ bind along slightly different grooves on Tfb1PH, but share a common

anchor point involving Phe54 on p53TAD2 and Phe670 on Rad2 (**Fig. 7 a-b-d-e**). In contrast, EKLFTAD2 and TFII α CTD bind in more elongated conformation as seen with Rad2₆₄₂₋₆₉₀ and follow very similar pathways. In particular, there are a number of similarities between the interface of the Rad2₆₄₂₋₆₉₀-Tfb1PH complex and the interface of the TFII α CTD-p62PH complex (**Fig. 7a-c-d-f**). The N-terminal region of TFII α CTD binds to p62PH in an extended form and interacts with strands β 5, β 6 and β 7. Rad2₆₄₂₋₆₉₀ binds to Tfb1PH in a very similar extended conformation and the interface it forms with Tfb1PH is almost identical to the one observed for the N-terminal region TFII α CTD in complex with p62PH (**Fig. 7a-c**). In particular, Phe670 and Val673 of Rad2 are making similar van der Waals contacts as Phe387 and Val390 of TFII α CTD in the TFII α CTD-p62PH complex (**Fig. 7d-f**). In addition, there are similar ionic interactions in both complexes between positively charged residues of the PH domains and negatively charged residues of either Rad2 or TFII α . The main difference between the two structures is that TFII α CTD contains a structured region that separates two unstructured acidic regions (**Fig. 7c**). NMR studies with a longer segment of the Rad2 spacer region (Rad2₆₄₂₋₇₆₀) indicate that this segment does not contain a folded domain in the free form (**Supplementary Fig. 2**), as observed with the free form of TFII α CTD.

As seen with many signaling proteins, the PH domain of Tfb1/p62 provides an excellent scaffold for protein-protein interactions that are important for the regulation of both transcription and NER in the nucleus. The similarity between the Rad2₆₄₂₋₆₉₀-Tfb1PH and the TFII α CTD-p62PH interfaces allows us to define a Tfb1PH-binding motif that consists of an aromatic residue (W or F) followed by two acidic residues and a valine residue locating within a highly acidic segment. The reason for the two Tfb1PH binding motifs within the spacer region of Rad2 is not completely clear at this point in time, but this may reflect the highly dynamic nature of the TFIIH-Rad2 complex during NER. In addition, the remarkable similarity between the interfaces of the Rad2₆₄₂₋₆₉₀-Tfb1PH and the TFII α CTD-p62PH complexes provides a clear mechanistic link for the role of the Tfb1/p62 subunit of TFIIH in both transcription and NER.

METHODS

Strains, media and vectors

All *S. cerevisiae* strains used are listed in **Supplementary Table 1**. The *rad52* strain was a gift from Dr. Pascal Chartrand (Université de Montréal) and the SHY186 strain was a gift from Dr. Steve Hahn (Fred Hutchinson Cancer Research Center). All strains were grown either in a synthetic complete media (SC; 0.67% yeast nitrogen base w/o amino acids, 2% glucose and a mixture of amino acids and vitamins) lacking either tryptophan and leucine (SC-LW) or uracil (SC-U) to be selective or in a rich media (YPD; 1% yeast extract, 2% tryptone and 2% glucose). All yeast transformations were done using the modified lithium acetate protocol²⁶.

Plasmid preparation

The pRS314TFB1-6His plasmid (*TFB1*) was kindly provided by the laboratory of Dr. S. Hahn. From this plasmid were constructed the N-terminal and C-terminal mutants, pRS314TFB1(Δ 1-115)-6His (*tfb1- Δ PH*) and pRS314TFB1(Δ 532-642)-6His (*tfb1-1*), respectively. The pRS316RAD2cmyc plasmid (*RAD2*) was generated by amplification of the *RAD2* Open Reading Frame (ORF) complemented by 400 bp upstream and 366 bp downstream on genomic DNA and insertion into pRS316. Mutant plasmids were obtained by overlapped PCR or QuikChange Site-Directed Mutagenesis procedure (Stratagene). For details of the plasmid preparation, see **Supplementary Methods**.

Plasmid shuffling

Since *TFB1* is an essential gene, plasmid shuffling was done in SHY186 to express the mutants in the *tfb1* (*TFB1* knockout) background. pRS314TFB1WT-6His, pRS314TFB1(Δ 1-115)-6His and pRS314TFB1(Δ 532-642)-6His were individually transformed into SHY186 and selected on SC-LWU plates to obtain the wild-type *TFB1*,

tfb1-ΔPH mutant and *tfb1-1* mutant strains, respectively. The strains bearing the two plasmids were then grown four days in liquid SC-LW to remove the uracil pYCP50/TFB1-6His plasmid. During this period, the cultures were diluted (1/50) each day with fresh media. In the final dilution, 5-Fluoroorotic Acid (5-FOA; 1 mg/mL, Zymo Research) was added. The cells were then plated on SC-LW media and auxotrophy for uracil and/or tryptophan was analyzed on selective plates.

Sensitivity assays

Yeast strains were grown overnight at 30°C in complete, selective or YPD media, as indicated. The next day they were diluted to obtain an $OD_{595} = 0.5-1$. The cells were then harvested by centrifugation, washed and resuspended in sterile water to obtain an $OD_{595} = 0.5$. For UV-sensitivity assays, dilutions were plated on selective media (SC-LW for Tfb1 and SC-U for Rad2) and irradiated with UV light (XL-1000 UV crosslinker, SpectroLinker) at varying energy levels. The surviving colonies were counted after 3 days in the dark at 30°C. For bleomycin and temperature sensitivity assays, serial 10-fold dilutions (10^{-1} – 10^{-4}) were made and 8 μ L of each dilution was dropped on solid media: YPD \pm 250 ng/mL bleomycin and SC-LW respectively. For the temperature sensitivity assays, the plates were incubated for 3 days at 30°C and 37°C. For the bleomycin assay, the plates were incubated for 4 days at room temperature.

Cloning and purification of proteins

The GST-Tfb1PH (residues 1-115 of Tfb1) and GST-p53TAD2 (residues 20-73 of p53) were prepared as described²³. GST-Rad2₃₅₉₋₃₈₃, GST-Rad2₆₄₂₋₇₆₀ and GST-Rad2₆₉₂₋₇₆₀ were prepared by inserting the appropriate region of Rad2 (Open Biosystems) into the pGEX-2T expression vector. The GST-Rad2₆₄₂₋₆₉₀ was created from GST-Rad2₆₄₂₋₇₆₀ by inserting a stop codon and adding a tyrosine at the C terminus (for spectrophotometry A_{280} quantification). All point mutants were made using the QuikChange II site-directed

mutagenesis procedure (Stratagene). All coding sequences were verified by DNA sequencing. Tfb1PH and p53TAD2 were purified as described²³. Rad2 fragments and mutants were expressed as GST-fusion proteins in *E. coli* host strain TOPP2 purified over GSH resin (GE Healthcare) and cleaved with thrombin (Calbiochem) as previously described for Tfb1PH²³. Following cleavage with thrombin, the Rad2 proteins were purified over a Q-Sepharose High Performance (GE Healthcare) column and dialyzed into appropriate buffers for ITC and NMR studies. ¹⁵N-labeled and ¹⁵N/¹³C-labeled proteins were prepared in M9-minimal media containing ¹⁵NH₄Cl (Sigma) and/or ¹³C₆-glucose (Sigma) as the sole nitrogen and carbon sources, respectively. For all experiments, the concentrations of proteins were determined from A₂₈₀.

ITC experiments

ITC titrations were performed as described²⁷, at 25°C in 20 mM sodium phosphate buffer (pH 7.5). All titrations fit a single-binding site mechanism with 1:1 stoichiometry and values are the average of two or more separate experiments.

NMR experiments

The NMR chemical shift perturbation and competition experiments were performed as previously described (for sample details please see **Supplementary Methods**). For the NMR structural studies of the Rad2₆₄₂₋₆₉₀-Tfb1PH complex, four different samples containing 1.0 mM of the complex in a 1:1.25 ratio were used (¹⁵N-Tfb1PH-Rad2₆₄₂₋₆₉₀, ¹⁵N/¹³C-Tfb1PH-Rad2₆₄₂₋₆₉₀, ¹⁵N-Rad2₆₄₂₋₆₉₀-Tfb1PH, and ¹⁵N/¹³C-Rad2₆₄₂₋₆₉₀-Tfb1PH, respectively). All NMR experiments were carried out in 20 mM sodium phosphate (pH 6.5), 1 mM EDTA, 1 mM DTT and 90 %H₂O/10% D₂O or 100% D₂O, at 300 K on Varian Unity Inova 500, 600 and 800 MHz spectrometers equipped with z pulsed-field gradient units and triple resonance probes. All of the ¹H, ¹⁵N, and ¹³C resonances for Rad2₆₄₂₋₆₉₀ and Tfb1PH were assigned as reported for free Tfb1PH²⁸. Briefly, 3D HNCOC²⁹, 3D HNCACB³⁰, 3D CBCACONH³¹, 3D (H)C(CO)NH³², 3D H(CCO)NH³² and 3D HCCH-

COSY³³ were used to assign the backbone and aliphatic side chains resonances. The aromatic side chains ¹H, ¹³C and ¹⁵N resonances were assigned using a combination of 2D (HB)CB(CGCD)HD and 2D (HB)CB(CGCDCE)HE spectra³⁴. Interproton distance restraints were measured from 3D ¹⁵N-edited NOESY-HSQC and ¹³C-edited HMQC-NOESY spectra ($\tau_m=90$ ms) and intermolecular distance restraints from 3D ¹⁵N/¹³C {F1}-filtered, {F3}-edited NOESY experiment ($\tau_m=90$ ms)^{35,36}. The NMR data were processed with NMRPipe/NMRDraw³⁷ and analyzed with NMRView³⁸ and Analysis from the CCPNMR suite³⁹.

Structure calculations

The NOE-derived distance restraints were divided into four classes defined as strong (1.8-2.8 Å), medium (1.8-4.0 Å), weak (1.8-5.0 Å) and very weak (3.3-6.0 Å). Backbone dihedral angles were derived with the program TALOS+⁴⁰. The structure of the Rad2₆₄₂₋₆₉₀-Tfb1PH complex was calculated using the program CNS⁴¹. The quality of the structures was analyzed by the programs PROCHECK-NMR⁴² and MOLMOL⁴³. The figures were generated with the program PyMol⁴⁴.

ACKNOWLEDGEMENTS

We would like to thank Dr. Steve Hahn and Dr. Pascal Chartrand for strains and clones, Dr. Lawrence Myers for the anti-Tfb1 antibody, Aurélie Bernier for help with protein purification and Dr Tara Sprules for assistance with several NMR experiments. 800 MHz NMR experiments were recorded at the Québec/Eastern Canada High Field NMR Facility, supported by the Natural Sciences and Engineering Research Council of Canada. This work was supported by a grant from the Canadian Cancer Society (J.G.O.). J.L-V. is a Vanier Canada Graduate Scholar from the CIHR. L.C. is a postdoctoral fellow of the NSERC CREATE program. P.L. is a Canadian Research Chair in Structural Biology and Engineering of RNA.

FIGURES

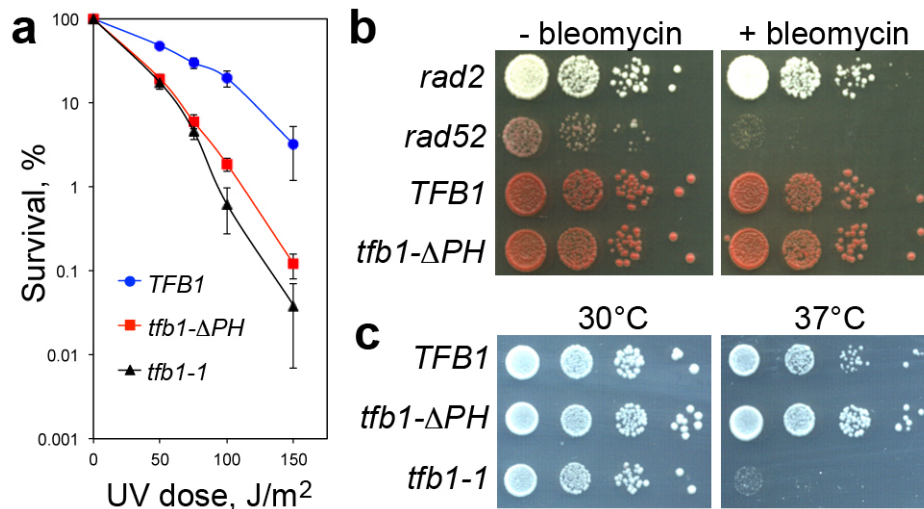


Figure 1. *tfb1-ΔPH* is sensitive to UV irradiation, but not bleomycin or temperature. **a.** The survival of *TFB1* (blue), *tfb1-ΔPH* (red) and *tfb1-1* (+ control; black) yeast were determined following increasing doses of UV irradiation. The y axis represents the percentage of surviving cells (normalized to the number of viable cells not exposed to UV light) and the x axis shows the energy levels of the UV irradiation applied (J/m²). The results are the mean ± SEM of three independent experiments. **b.** The survival of *TFB1*, *tfb1-ΔPH*, *rad2* (- control) and *rad52* (+ control) yeast were determined before (left) or after (right) exposure to bleomycin. The yeast strains are incubated on plates containing YPD ± 250 ng/mL bleomycin and survivors are determined by spotting 8 μL of cells in serial dilutions (10⁻¹-10⁻⁴). The assay is repeated three times, and a typical set of results is shown. **c.** The survival of *TFB1* (- control), *tfb1-ΔPH* and *tfb1-1* (+ control) yeast were determined following growth at either 30°C (left) or 37°C (right). The survivors are assayed by spotting 8 μL of cells in serial dilutions (10⁻¹-10⁻⁴) on SC-LW plates. This assay is repeated three times and a typical set of results is shown.

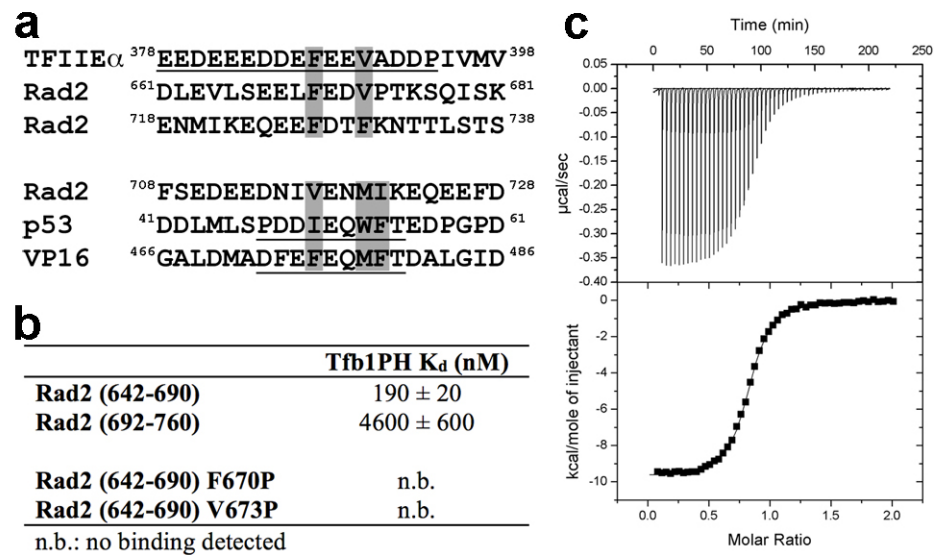


Figure 2. The Rad2 spacer region contains a high affinity Tfb1PH-binding site between residues 642 and 690. **a.** Identification of amino acid segments located between residues 642 to 760 from the Rad2 spacer region that align with the Tfb1PH-binding sites from TFIIE α CTD, p53TAD2 and VP16C. In the alignments, the residues of TFIIE α CTD, p53TAD2 and VP16C that form the binding interface with p62PH/Tfb1PH are underlined and crucial hydrophobic residues are shaded in grey. **b.** Comparison of the dissociation constant (K_d) values for the binding of different Rad2 segments located between residues 642 and 760 with Tfb1PH. **c.** Thermogram of the Tfb1PH titration with successive additions of Rad2₆₄₂₋₆₉₀. Experiments are performed at 25°C, in 20 mM NaPO₄ pH 7.5, and the results fit to a single-binding site model with 1:1 stoichiometry.

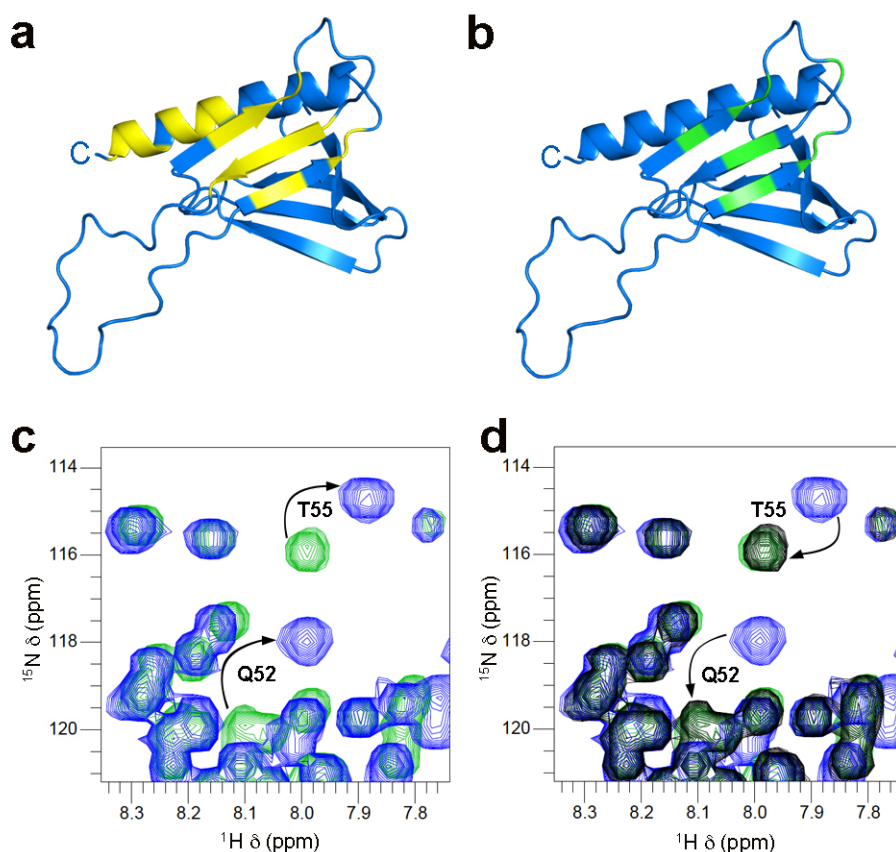


Figure 3. Rad2₆₄₂₋₆₉₀ and p53TAD2 share a common binding site on Tfb1PH. **a-b.** Ribbon models of the three dimensional structure of Tfb1PH (blue; PDB code 1Y5O). The amino acids of ¹⁵N-labeled Tfb1PH showing a significant chemical shift change $\{\Delta\delta(\text{ppm}) > 0.15; \Delta\delta = [(0.17\Delta N_H)^2 + (\Delta H_N)^2]^{1/2}\}$ upon formation of a complex with either Rad2₆₄₂₋₆₉₀ (**a**) or p53TAD2 (**b**) are highlighted in yellow and green, respectively. **c.** Overlay of a selected region from the ¹H-¹⁵N HSQC spectra of ¹⁵N-labeled p53TAD2 (0.5 mM) in the free form (green) and in the presence of unlabeled Tfb1PH (0.4 mM; blue). **d.** Same overlay as in **c**, but after the addition of unlabeled Rad2₆₄₂₋₆₉₀ (0.5 mM; black). Signals of p53TAD2 that undergo significant changes in ¹H and ¹⁵N chemical shifts upon formation of the complex with Tfb1PH (**c**), and that return towards their original position following the addition of Rad2₆₄₂₋₆₉₀ (**d**) are indicated by arrows.

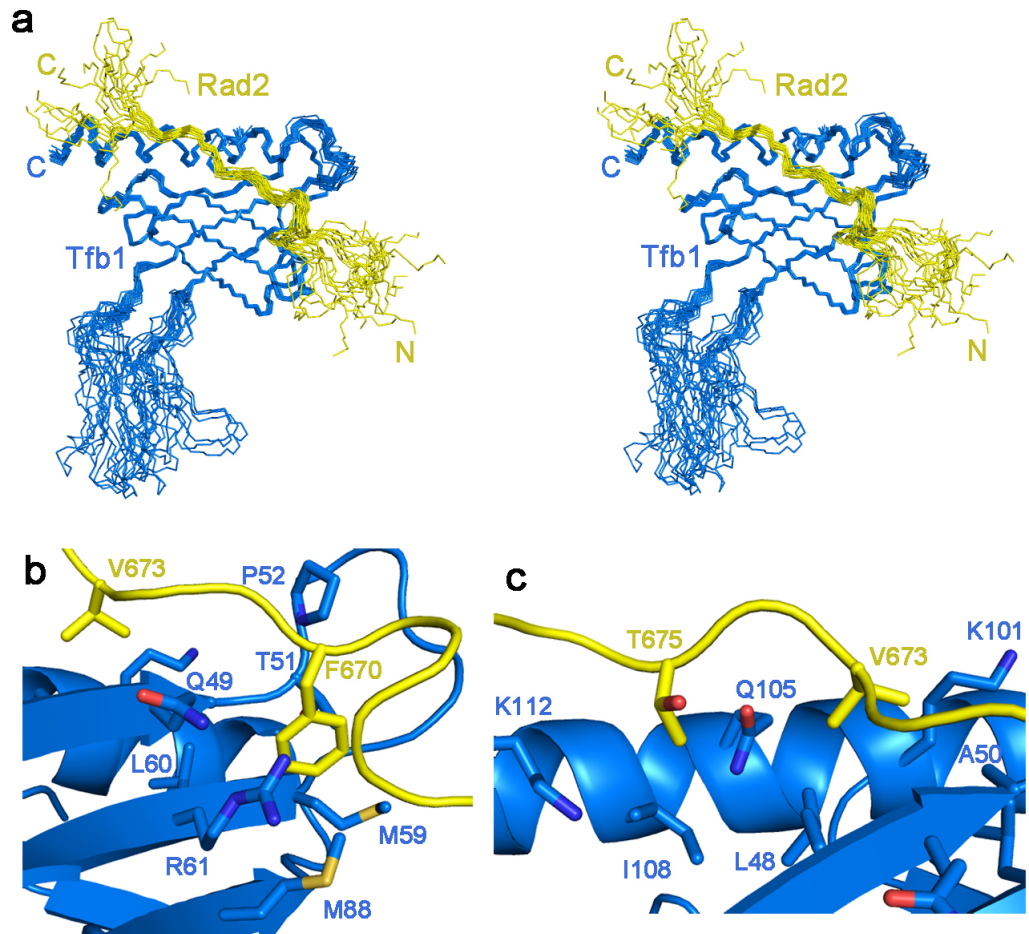


Figure 4. NMR structure of the Rad2₆₄₂₋₆₉₀-Tfb1PH complex. **a.** Stereo view of the 20 lowest-energy structures of the complex between Tfb1PH (blue) and Rad2₆₄₂₋₆₉₀ (yellow; PDB code XXXX). The structures were superimposed using the backbone atoms C', C^α, and N of residues 4–65 and 85–112 of Tfb1PH and residues 661–680 of Rad2₆₄₂₋₆₉₀. **b.** Ribbon representation of Tfb1PH (blue) and backbone trace of the region of Rad2₆₄₂₋₆₉₀ (yellow) interacting in the first binding pocket. In this pocket, Phe670 of Rad2 forms a cation- π interaction with Arg61 of Tfb1 and van der Waals interaction with Met59 and Lys57. **c.** Ribbon representation of Tfb1PH (blue) and backbone trace of the region of Rad2₆₄₂₋₆₉₀ (yellow) interacting in the second binding pocket. On one side of the pocket Val673 of Rad2 is inserted where it interacts with Leu48, Ala50, Lys101 and Gln105 of Tfb1. On the other side of the pocket Thr675 of Rad2 interacts with Gln105, Ile108 and Lys112 of Tfb1.

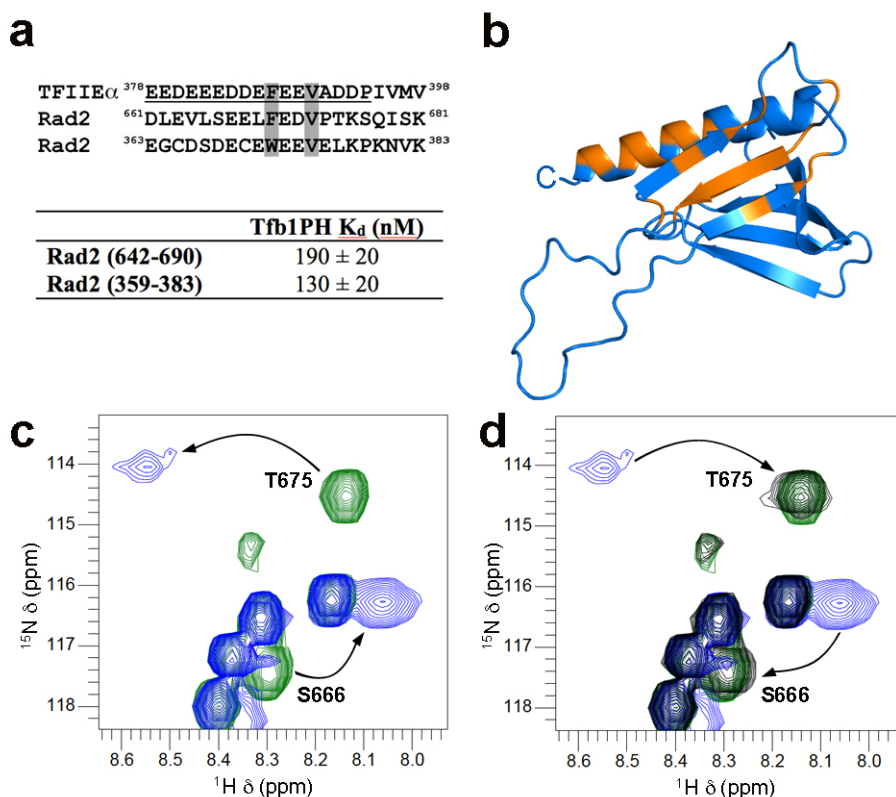


Figure 5. The Rad2 spacer region contains a second high-affinity Tfb1PH-binding site. **a. (Top)** Identification of amino acid segments located between residues 363-382 from the Rad2 spacer region that align with the Tfb1PH-binding sites from TFIIE α CTD and Rad2₆₄₂₋₆₉₀. In the alignments, the residues of TFIIE α CTD and Rad2₆₄₂₋₆₉₀ that form the binding interface with p62PH/Tfb1PH are underlined and crucial hydrophobic residues are shaded in grey. **(Bottom)** Dissociation constant (K_d) values for the binding of Rad2₃₅₉₋₃₈₃ and Rad2₆₄₂₋₆₉₀ with Tfb1PH as determined by ITC analysis. **b.** Ribbon models of the three dimensional structure of Tfb1PH (blue; PDB code 1Y5O). The amino acids of ¹⁵N-labeled Tfb1PH showing a significant chemical shift change $\{\Delta\delta(\text{ppm}) > 0.15; \Delta\delta = [((0.17\Delta N_H)^2 + (\Delta H_N)^2)]^{1/2}\}$ upon formation of a complex with Rad2₃₅₉₋₃₈₃ are highlighted in orange. **c.** Overlay of a selected region from the ¹H -¹⁵N HSQC spectra of ¹⁵N-labeled Rad2₆₄₂₋₆₉₀ (0.5 mM) in the free form (green) and in the presence of unlabeled Tfb1PH (0.4 mM; blue). **d.** Same overlay as in c, but after the addition of unlabeled Rad2₃₅₉₋₃₈₃ (1.0 mM; black). Signals of Rad2₆₄₂₋₆₉₀ that undergo significant chemical shift changes in ¹H and ¹⁵N chemical shifts upon formation of the complex with Tfb1PH (**c**), and that return towards their original position following the addition of Rad2₃₅₉₋₃₈₃ (**d**) are indicated by arrows.

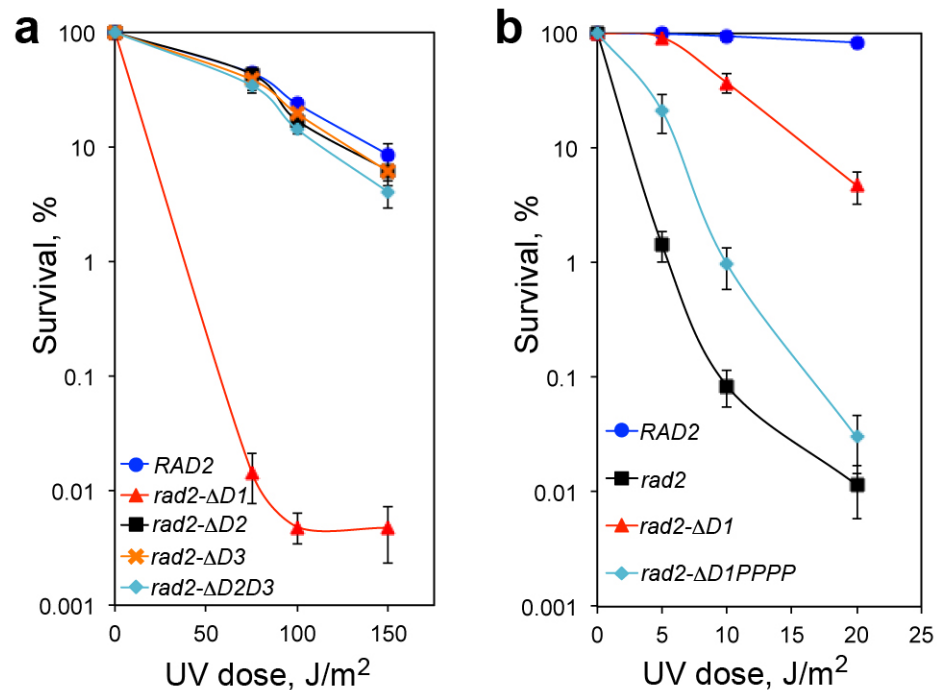


Figure 6. Multiple regions of the Rad2 spacer are required for repair of UV damage. **a.** The survival of *RAD2* (blue), *rad2-ΔD1* (red), *rad2-ΔD2* (black), *rad2-ΔD3* (orange), and *rad2-ΔD2D3* (aqua) yeast were determined following increasing doses of UV irradiation. The y axis represents the percentage of surviving cells (normalized to the number of viable cells not exposed to UV light) and the x axis shows the energy levels of the UV irradiation applied (J/m²). The results are the mean ± SEM of three independent experiments. **b.** The survival of *RAD2* (blue), *rad2* (black), *rad2-ΔD1* (red), and *rad2-ΔD1PPPP* (aqua) yeast were determined following increasing doses of UV irradiation. The y axis represents the percentage of surviving cells (normalized to the number of viable cells not exposed to UV light) and the x axis shows the energy levels of the UV irradiation applied (J/m²). The results are the mean ± SEM of four independent experiments.

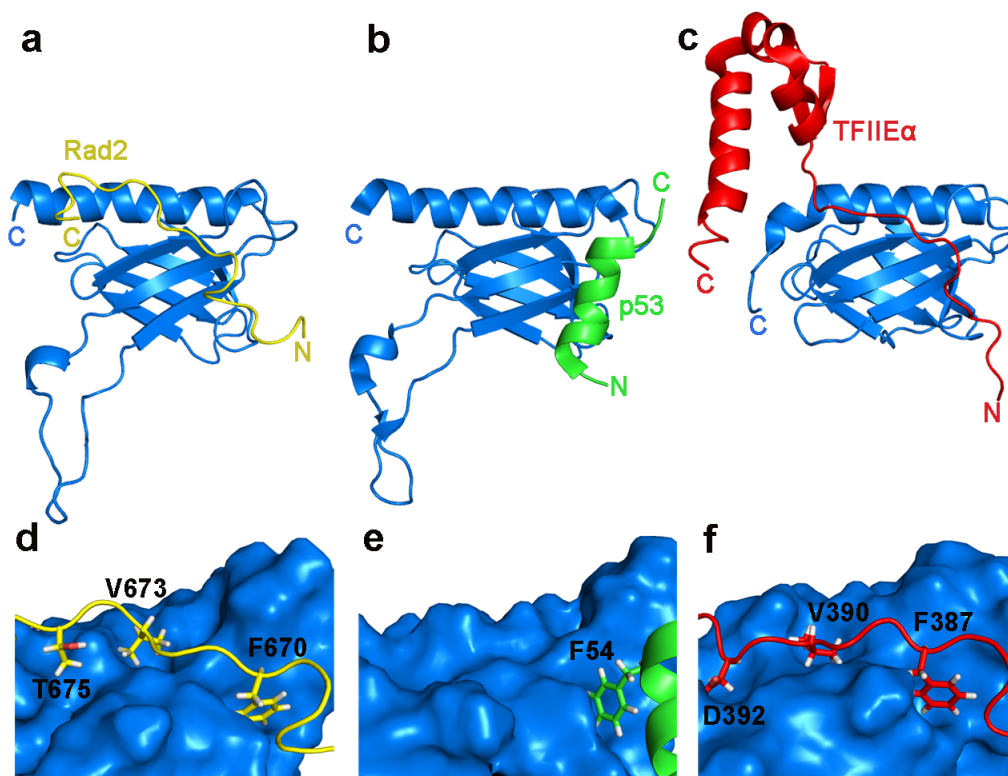


Figure 7. The structures of the Rad2₆₄₂₋₆₉₀-Tfb1PH and TFIIIE α CTD-p62PH interfaces are remarkably similar. a-c. Ribbon diagrams of the lowest energy structures of the Rad2₆₄₂₋₆₉₀-Tfb1PH (a; PDB code XXXX), the p53TAD2-Tfb1PH (b; PDB code 2GS0) and the TFIIIE α CTD-p62PH (c; PDB code 2RNR) complexes. Tfb1PH (a-b) and p62PH (c) are shown in blue, Rad2₆₄₂₋₆₉₀ (a) in yellow, p53TAD2 (b) in green and TFIIIE α CTD (c) in red. In panels d-f, the 3D structure of Tfb1PH (d-e) and p62PH (f) are shown as molecular surfaces (blue) and Rad2₆₄₂₋₆₉₀ (d), p53TAD2 (e) and TFIIIE α CTD (f) are shown as ribbons in yellow, green and red, respectively. Select residues of Rad2₆₄₂₋₆₉₀ (d), p53TAD2 (e) and TFIIIE α CTD (f) at the binding interface are also shown.

Table 1. Structural Statistics for Rad2 in Complex with Tfb1PH^a

Restrains used for the structure calculations	
Total number of NOE distances restraints	1593
Short-range (intraresidue)	377
Medium-range ($ i-j \leq 4$)	643
Long-range	480
Intermolecular	93
Hydrogen bond	36
Number of dihedral angle restraints (φ , ψ)	160
Structural statistics	
Rms deviations from idealized geometry	
Bonds (Å)	0.00260 ± 0.00007
Angles (deg)	0.423 ± 0.004
Improper (deg)	0.289 ± 0.008
Rms deviations from distance restraints (Å)	0.0197 ± 0.0004
Rms deviations from dihedral restraints (deg)	0.237 ± 0.019
Ramachandran statistics (%) ^b	
Residues in most favored regions	72.5
Residues in additional allowed regions	25.6
Residues in generously allowed regions	1.5
Residues in disallowed regions	0.4
Coordinate precision ^c	
Atomic pair wise rmsd (Å)	
Rad2 – Tfb1PH complex	
Backbone atoms (C', C ^α , N)	0.42 ± 0.08
All heavy atoms	1.09 ± 0.10
Tfb1PH alone	
Backbone atoms (C', C ^α , N)	0.38 ± 0.09
All heavy atoms	1.06 ± 0.10
Rad2 alone	
Backbone atoms (C', C ^α , N)	0.41 ± 0.14
All heavy atoms	1.3 ± 0.3

^a The 20 conformers with the lowest energies were selected for statistical analysis.

^b Based on PROCHECK-NMR analysis.

^c Only residues 5–63 and 86–112 of Tfb1PH and residues 669–675 of Rad2 were used for the rmsd calculations. Residues at the N terminus (1–4), at the C terminus (113–115), and in the flexible loop (64–85) of Tfb1PH, as well as residues at the N terminus (642–668) and at the C terminus (676–690) of Rad2 were not included in the calculation.

REFERENCES

1. Le May, N., Egly, J.M. & Coin, F. True lies: the double life of the nucleotide excision repair factors in transcription and DNA repair. *J Nucleic Acids* **2010**(2010).
2. Tirode, F., Busso, D., Coin, F. & Egly, J.M. Reconstitution of the transcription factor TFIIH: assignment of functions for the three enzymatic subunits, XPB, XPD and cdk7. *Mol. Cell* **3**, 87-95 (1999).
3. Lu, H., Zawel, L., Fisher, L., Egly, J.M. & Reinberg, D. Human general transcription factor IIH phosphorylates the C-terminal domain of RNA polymerase II. *Nature* **358**, 641-645 (1992).
4. Araujo, S.J., Nigg, E.A. & Wood, R.D. Strong functional interactions of TFIIH with XPC and XPG in human DNA nucleotide excision repair, without a preassembled repairosome. *Mol. Cell. Biol.* **21**, 2281-2291 (2001).
5. Sung, P., Guzder, S.N., Prakash, L. & Prakash, S. Reconstitution of TFIIH and requirement of its DNA helicase subunits, Rad3 and Rad25, in the incision step of nucleotide excision repair. *J. Biol. Chem.* **271**, 10821-10826 (1996).
6. Friedberg, E.C. *et al. DNA Repair and Mutagenesis*, (ASM Press, Washington, DC, 2005).
7. Schärer, O.D. The molecular basis for different disease states caused by mutations in TFIIH and XPG. *DNA Repair* **7**, 339-344 (2008).
8. Lieber, M.R. The FEN-1 family of structure-specific nucleases in eukaryotic DNA replication, recombination and repair. *Bioessays* **19**, 233-240 (1997).
9. Tsutakawa, S.E. *et al.* Human flap endonuclease structures, DNA double-base flipping, and a unified understanding of the FEN1 superfamily. *Cell* **145**, 198-211 (2011).
10. Dunand-Sauthier, I. *et al.* The spacer region of XPG mediates recruitment to nucleotide excision repair complexes and determines substrate specificity. *J. Biol. Chem.* **280**, 7030-7037 (2005).

11. Thorel, F. *et al.* Definition of a short region of XPG necessary for TFIIH interaction and stable recruitment to sites of UV damage. *Mol. Cell. Biol.* **24**, 10670-10680 (2004).
12. Iyer, N., Reagan, M.S., Wu, K.J., Canagarajah, B. & Friedberg, E.C. 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**, 2157-2167 (1996).
13. Bardwell, A.J. *et al.* Yeast nucleotide excision repair proteins Rad2 and Rad4 interact with RNA polymerase II basal transcription factor b (TFIIH). *Mol. Cell. Biol.* **14**, 3569-3576 (1994).
14. Zotter, A. *et al.* Recruitment of the nucleotide excision repair endonuclease XPG to sites of UV-induced DNA damage depends on functional TFIIH. *Mol. Cell. Biol.* **26**, 8868-8879 (2006).
15. Lehmann, A.R. DNA repair-deficient diseases, xeroderma pigmentosum, Cockayne syndrome and trichothiodystrophy. *Biochimie* **85**, 1101-1111 (2003).
16. Hohl, M. *et al.* Domain swapping between FEN-1 and XPG defines regions in XPG that mediate nucleotide excision repair activity and substrate specificity. *Nucleic Acids Res.* **35**, 3053-3063 (2007).
17. Matsui, P., DePaulo, J. & Buratowski, S. An interaction between the Tfb1 and Ssl1 subunits of yeast TFIIH correlates with DNA repair activity. *Nucleic Acids Res.* **23**, 767-772 (1995).
18. Gervais, V. *et al.* TFIIH contains a PH domain involved in DNA nucleotide excision repair. *Nat. Struct. Mol. Biol.* **11**, 616-622 (2004).
19. Di Lello, P. *et al.* p53 and TFIIH α share a common binding site on the Tfb1/p62 subunit of TFIIH. *Proc. Natl. Acad. Sci. USA* **105**, 106-111 (2008).
20. Di Lello, P. *et al.* Structure of the Tfb1/p53 complex: Insights into the interaction between the p62/Tfb1 subunit of TFIIH and the activation domain of p53. *Mol. Cell* **22**, 731-740 (2006).

21. Langlois, C. *et al.* NMR structure of the complex between the Tfb1 subunit of TFIID and the activation domain of VP16: structural similarities between VP16 and p53. *J. Am. Chem. Soc.* **130**, 10596-10604 (2008).
22. Okuda, M. *et al.* Structural insight into the TFIIE-TFIID interaction: TFIIE and p53 share the binding region on TFIID. *EMBO J.* **27**, 1161-1171 (2008).
23. Di Lello, P. *et al.* NMR structure of the amino-terminal domain from the Tfb1 subunit of TFIID and characterization of its phosphoinositide and VP16 binding sites. *Biochemistry* **44**, 7678-7686 (2005).
24. Tantin, D. RNA polymerase II elongation complexes containing the Cockayne syndrome group B protein interact with a molecular complex containing the transcription factor IID components xeroderma pigmentosum B and p62. *J. Biol. Chem.* **273**, 27794-27799 (1998).
25. Mas, C. *et al.* Structural and functional characterization of an atypical activation domain in erythroid Kruppel-like factor (EKLF). *Proc. Natl. Acad. Sci. USA* (2011).
26. Kaiser, C., Michaelis, S. & Mitchell, A. *Methods in yeast genetics*, 202 (Cold Spring Harbor Laborator Press, Cold Spring Harbor, NY, 1994).
27. Houtman, J.C. *et al.* Binding specificity of multiprotein signaling complexes is determined by both cooperative interactions and affinity preferences. *Biochemistry* **43**, 4170-4178 (2004).
28. Nguyen, B.D., Di Lello, P., Legault, P. & Omichinski, J.G. ¹H, ¹⁵N, and ¹³C resonance assignment of the amino-terminal domain of the Tfb1 subunit of yeast TFIID. *J. Biomol. NMR* **31**, 173-174 (2005).
29. Kay, L.E., Xu, G.Y. & Yamazaki, T. Enhanced-sensitivity triple-resonance spectroscopy with minimal H₂O saturation. *J. Magn. Reson. A* **109**, 129-133 (1994).
30. Wittekind, M. & Mueller, L. HNCACB, a high-sensitivity 3D NMR experiment to correlate amide-proton and nitrogen resonances with the alpha- and beta-carbon resonances in proteins. *J. Magn. Reson. B* **101**, 201-205 (1993).

31. Grzesiek, S. & Bax, A. Correlating backbone amide and side chain resonances in larger proteins by multiple relayed triple resonance NMR. *J. Am. Chem. Soc.* **114**, 6291-6293 (1992).
32. Logan, T.M., Olejniczak, E.T., Xu, R.X. & Fesik, S.W. Side chain and backbone assignments in isotopically labeled proteins from two heteronuclear triple resonance experiments. *FEBS Lett.* **314**, 413-418 (1992).
33. Ikura, M., Kay, L.E. & Bax, A. Improved three-dimensional ¹H-¹³C-¹H correlation spectroscopy of a ¹³C- labeled protein using constant-time evolution. *J. Biomol. NMR* **1**, 299-304 (1991).
34. Yamazaki, T., Forman-Kay, J.D. & Kay, L.E. Two-dimensional NMR experiments for correlating carbon-13 beta and proton delta/epsilon chemical shifts of aromatic residues in ¹³C-labeled proteins via scalar couplings. *J. Am. Chem. Soc.* **115**, 11054-11055 (1993).
35. Zhang, O., Kay, L.E., Olivier, J.P. & Forman-Kay, J.D. Backbone ¹H and ¹⁵N resonance assignments of the N-terminal SH3 domain of drk in folded and unfolded states using enhanced-sensitivity pulsed field gradient NMR techniques. *J. Biomol. NMR* **4**, 845-858 (1994).
36. Pascal, S.M., Muhandiram, D.R., Yamazaki, T., Forman-Kay, J.D. & Kay, L.E. Simultaneous acquisition of ¹⁵N- and ¹³C-edited NOE spectra of proteins dissolved in H₂O. *J. Magn. Reson.* **103**, 197-201 (1994).
37. Delaglio, F. *et al.* NMRPipe: A multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **6**, 277-293 (1995).
38. Johnson, B.A. Using NMRView to visualize and analyze the NMR spectra of macromolecules. *Methods Mol. Biol.* **278**, 313-352 (2004).
39. Vranken, W.F. *et al.* The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins* **59**, 687-696 (2005).
40. Shen, Y., Delaglio, F., Cornilescu, G. & Bax, A. TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *J. Biomol. NMR* **44**, 213-223 (2009).

41. Brunger, A.T. *et al.* Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr. D Biol. Crystallogr.* **54**, 905-921 (1998).
42. Laskowski, R.A., Rullmannn, J.A., MacArthur, M.W., Kaptein, R. & Thornton, J.M. AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J. Biomol. NMR* **8**, 477-486 (1996).
43. Koradi, R., Billeter, M. & Wuthrich, K. MOLMOL: a program for display and analysis of macromolecular structures. *J. Mol. Graph.* **14**, 51-55 (1996).
44. DeLano, W.L. The PyMOL Molecular Graphics System. (Delano Scientific, Palo Alto, CA, 2002).

SUPPLEMENTARY METHODS

Plasmid preparations

The Tfb1-6His/pRS314 plasmid contains the *TFB1* ORF and its endogenous 400 bp upstream promoter region and a 6His coding tag on the C-terminal part of the protein, all of which were inserted into the SpeI cloning site of pRS314. A small correction was made to this plasmid to respect the *TFB1 Saccharomyces cerevisiae* Genome Database sequence. A QuikChange (QuikChange Site-Directed Mutagenesis Kit from Stratagene) mutagenesis procedure was done to switch the amino acid 102 from a threonine to a methionine with the primers TFB1-T102m-F and TFB1-T102m-R (see **Supplementary Table 2** for primers used in the study). This corrected version was used throughout the paper and referred to as pRS314Tfb1WT-6His. The *tfb1- Δ PH* mutant was built by overlapped PCR with TFB1(1-115).1F, TFB1(1-115).1R, TFB1(1-115).2F and TFB1(1-115).2R, using the Tfb1WT plasmid as template. The PCR product was inserted SpeI/SpeI in pRS314. The *tfb1-I* mutant was also obtained by an overlapped PCR inserting a STOP codon followed by a 6His tag to replace the glutamine 532. The primers TFB1GFP-BlpI-F, pRSTFB1(1-531)6His-R, pRSTFB1(1-531)6His-F and T7+ were used on the Tfb1WT plasmid. The PCR product was inserted BlpI/SacII in the Tfb1WT plasmid.

The pRS316RAD2cmyc plasmid (for the *RAD2* strain) was generated by amplification of the *RAD2* ORF and 400pb upstream of the coding sequence on genomic DNA with the primers pRS-RAD2-XhoI-F(c) and pRS-RAD2-R. The PCR product was inserted XhoI/BamHI into pRS316. Then was added the 366 bp downstream of the endogenous ORF by amplification on genomic DNA with the primers pRS-RAD2(aval)-BamHI-F and pRS-RAD2(aval)-SpeI-R. The insertion was done by digesting BamHI/SpeI. Finally, the myc tag was inserted in between MscI/SpeI right before the STOP codon of *RAD2* by overlapped PCR using RAD2-2905-MscI-F, pRS-RAD2(cmyc)-R, pRS-RAD2(cmyc)-F and pRS-RAD2(aval)-SpeI-R. The previous intermediate plasmid was used as template. The mutant pRS316Rad2(Δ 367-378)cmyc plasmid (for *rad2- Δ D2*) was constructed by overlapped PCR with the primers RAD2-SnaBI-F, RAD2(m367-378)I-R, RAD2(m367-378)II-F and RAD2-BlpI-R. The template used was the pRS316RAD2cmyc plasmid and the product was inserted SnaBI/BlpI into that same plasmid. The

pRS316Rad2(Δ 642-760)cmyc plasmid (for *rad2- Δ D3*) was obtained by overlapped PCR using RAD2-1413-BlpI-F, RAD2I-R, RAD2II(m642-760)-F, RAD2II-MscI-R and the pRS316RAD2cmyc plasmid as template. The product was inserted BlpI/MscI into the pRS316RAD2cmyc plasmid. The pRS316Rad2(Δ 228-237)cmyc plasmid (for *rad2- Δ D1*) was also obtained by overlapped PCR using RAD2-SnaBI-F, RAD2(m228-237)I-R, RAD2(m228-237)II-F and RAD2-BlpI-R. The insertion SnaBI/BlpI was done into the pRS316RAD2cmyc plasmid. To obtain the double mutant, the plasmid pRS316Rad2(Δ 367-378 Δ 642-760)cmyc (for *rad2- Δ D2D3*) was built in the same way as our first mutant except the insertion was done in the pRS316RAD2(Δ 642-760)cmyc plasmid. Finally, the pRS316RAD2(Δ 228-237/PPPP)cmyc four proline mutant plasmid (*rad2- Δ D1PPPP*) was obtained by two QuikChange Site-Directed Mutagenesis procedure (Stratagene) on template pRS316RAD2(Δ 228-237)cmyc using first RAD2(W372P+V375P)-F and RAD2(W372P+V375P)-R, followed by RAD2(F670P+V673P)-F and RAD2(F670P+V673P)-R.

NMR samples

For the NMR chemical shift titration studies of Rad2 segments with Tfb1PH, either unlabeled Rad2₃₅₉₋₃₈₃ or unlabeled Rad2₆₄₂₋₆₉₀ was added to a sample containing 0.5 mM of ¹⁵N-labeled Tfb1PH in NMR buffer to a final ratio of 1:1. These studies were performed in 20 mM sodium phosphate buffer pH 6.5, 1mM EDTA, 1mM DTT (NMR buffer) with 90%H₂O/10%D₂O.

For the NMR competition experiment of p53TAD2 and Rad2₆₄₂₋₆₉₀, a sample containing 0.5 mM ¹⁵N-labeled p53TAD2 (residues 20-73) in NMR buffer was used. To this sample, unlabeled Tfb1PH was added to a final concentration of 0.4 mM. In a second addition, unlabeled Rad2₆₄₂₋₆₉₀ was added to a final concentration of 0.5 mM.

For the NMR competition experiment of Rad2₃₅₉₋₃₈₃ and Rad2₆₄₂₋₆₉₀, a sample containing 0.5 mM ¹⁵N-labeled Rad2₆₄₂₋₆₉₀ in NMR buffer was used. To this sample, unlabeled Tfb1PH was added to a final concentration of 0.4 mM. In a second addition, unlabeled Rad2₃₅₉₋₃₈₃ was added to a final concentration of 1.0 mM.

For the ^1H - ^{15}N HSQC experiment of Rad₆₄₂₋₇₆₀, the sample consisted of 0.5mM ^{15}N -labeled Rad₆₄₂₋₇₆₀ in NMR buffer. The sample was recorded at 25°C on a Varian Inova 600 MHz spectrometer.

Protein expression levels in yeast

To determine the expression level of Tfb1 and the Tfb1- ΔPH mutant proteins, yeast cells were grown in 15 mL of minimal media to $\text{OD}_{595} = 0.8\text{-}1.0$ and centrifuged. The pellets were resuspended and total protein extractions were carried out as previously described (Elagoz et al.,1999). The protein extracts (20 μg) were then migrated on a 10% (wt/vol) SDS-PAGE gel and transferred onto a nitrocellulose membrane according to the manufacturer's instructions. For detection of the Tfb1 protein and its Tfb1- ΔPH mutant, we used an anti-Tfb1 rabbit polyclonal antibody, a gift from Dr. Larry Myers (Dartmouth), at a dilution of 1/1,000. The secondary antibody used was the commercially available Peroxidase AffiniPure Goat Anti-Rabbit IgG (H+L) (Jackson ImmunoResearch Laboratories) at a dilution of 1/5,000.

For the Rad2 protein and its mutants, normal protein levels were too low to detect when expressed from their endogenous promoters. Therefore, Rad2 and all mutants were cloned into the overexpression vector pVT-101u, which is under the control of the ADH promoter. The genes were inserted XhoI/BamHI and a myc-tag was fused on their C-terminus. The Rad2 and mutants proteins levels were determined with the anti-cmyc mouse mAb 9E10 at a dilution of 1/1,000. The Peroxidase AffiniPure Goat Anti-Mouse IgG + IgM (H+L) (Jackson ImmunoResearch Laboratories) served as the secondary antibody, used at a dilution of 1/1,000. Loading was verified with the anti-actin mouse monoclonal antibody (Abcam) at a dilution of 1/1,000.

Reference :

Elagoz A., Callejo M., Armstrong J., Rokeach L. A. (1999) Although calnexin is essential in *S. pombe*, its highly conserved central domain is dispensable for viability. *J. Cell Sci.* 112, 4449–4460.

SUPPLEMENTARY TABLES

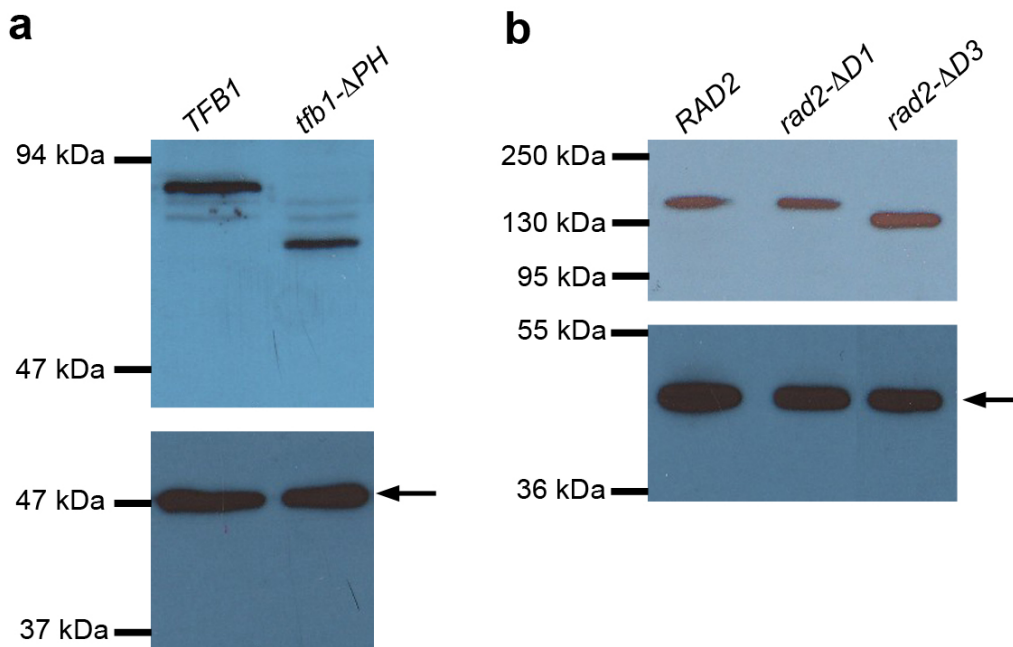
Supplementary Table 1: Yeast strains used in this study

Name	Genotype	Source
SHY186	MATa <i>ade2- ura3- lys2- trp1- his3- leu2- tfb1Δ::LEU2</i> + pYCP50/TFB1-6His	S. Hahn's Lab
<i>TFB1</i>	MATa <i>ade2- ura3- lys2- trp1- his3- leu2- tfb1Δ::LEU2</i> + pRS314TFB1-6His	This Lab
<i>tfb1-ΔPH</i>	MATa <i>ade2- ura3- lys2- trp1- his3- leu2- tfb1Δ::LEU2</i> + pRS314TFB1(Δ1-115)-6His	This Lab
<i>tfb1-1</i>	MATa <i>ade2- ura3- lys2- trp1- his3- leu2- tfb1Δ::LEU2</i> + pRS314TFB1(Δ532-642)-6His	This Lab
<i>rad52</i>	Mata <i>ade2- his3- leu2- trp1- ura3- rad52Δ::TRP</i>	P. Chartrand's Lab
<i>rad2(7289)</i>	Mata <i>his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 rad2Δ::KanMX</i>	Open Biosystems
<i>rad2</i>	<i>rad2(7289)</i> + pRS316	This Lab
<i>RAD2</i>	<i>rad2(7289)</i> + pRS316RAD2cmyc	This Lab
<i>rad2-ΔD1</i>	<i>rad2(7289)</i> + pRS316RAD2(Δ228-237)cmyc	This Lab
<i>rad2-ΔD2</i>	<i>rad2(7289)</i> + pRS316RAD2(Δ367-378)cmyc	This Lab
<i>rad2-ΔD3</i>	<i>rad2(7289)</i> + pRS316RAD2(Δ642-760)cmyc	This Lab
<i>rad2-ΔD2D3</i>	<i>rad2(7289)</i> + pRS316RAD2(Δ367-378Δ642-760)cmyc	This Lab
<i>rad2-ΔD1PPPP</i>	<i>rad2(7289)</i> + pRS316RAD2(Δ228-237/W372P/V375P/F670P/V673P)cmyc	This Lab

Supplementary Table 2: Primers used in this study

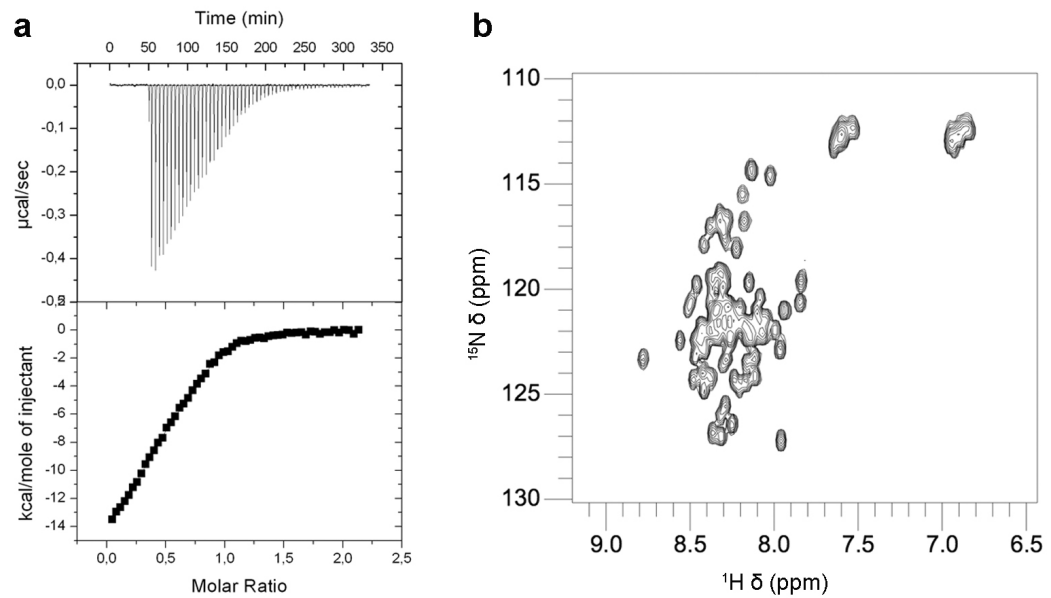
Name	Sequence
TFB1-T102m-F	5' cag tta tgg ata ata tca aga tga ccc ttc aac aaa 3'
TFB1-T102m-R	5' gat ttg ttg aag ggt cat ctt gat att atc cat aac tg 3'
TFB1(1-115).1F	5' aaa act agt cta gta atg ata 3'
TFB1(1-115).1R	5' ttc ttc gta gat cat tat ggt tta ctt tta ttt 3'
TFB1(1-115).2F	5' atc tac gaa gaa aag aga aga 3'
TFB1(1-115).2R	5' aaa act agt cta gtg atg gtg 3'
TFB1-GFP-BlpI-F	5' cag gct aag cat aac aat g 3'
pRSTFB1(1-531)6His-R	5' tct ttt gtc gac aac act tct cat cac cat cac cat cac taa gca aat gaa tta gag gtg 3'
pRSTFB1(1-531)6His-F	5' cac ctc taa ttc att tgc tta gtg atg gtg atg gtg atg aga agt gtt gtc gac aaa aga 3'
T7+	5' taa tac gac tca cta tag ggc 3'
pRS-RAD2-XhoI-F(c)	5' ccg ctc gag gga gct tag tag ctg ctg 3'
pRS-RAD2-R	5' cgg gat cct tac atc ttt ctt ttc ttt agt tta cct gtt gcg g 3'
pRS-RAD2(aval)-BamHI-F	5' cgg gat cct aac ata tac tta tgt atg 3'
pRS-RAD2(aval)-SpeI-R	5' gac tag tat cgt att atc aac acc tcg 3'
RAD2-2905-MscI-F	5' ggt tgg cca cac gaa aag tct g 3'
pRS-RAD2(cmyc)-R	5' gta tat gtt att aca ggt cct cct cc gaga tca gct tct gct cca tct ttc ttt tc 3'
pRS-RAD2(cmyc)-F	5' gaa aag aaa gat gga gca gaa gct gat ctc gga gga gga cct gta ata aca tat ac 3'
RAD2-SnaBI-F	5' ctg tac gta aat cat tga agg 3'
RAD2(m367-378)I-R	5' gtc aca tcc ctc atc ttc tga 3'
RAD2(m367-378)II-F	5' gag gga tgt gac ccc aag aac gtc aag ttt gtg g 3'
RAD2-BlpI-R	5' ccg gct tag cga ttt ttg att tgt cg 3'
RAD2(m228-237)I-R	5' acc atc ttc cca gac atg gtg aaa agg agg aat 3'
RAD2(m228-237)II-F	5' gtc aca tcc ctc atc ttc tga 3'
RAD2-1413-BlpI-F	5' cgc taa gcc ggt aac gtc aaa agg 3'
RAD2I-R	5' cgt agc att taa gcc tgt caa 3'
RAD2II(m642-760)-F	5' tta aat gct acg aaa aga gat tcg gat gag gta ac 3'
RAD2II-MscI-R	5' cgt gtg gcc aac cta gtt gag 3'
RAD2(W372P+V375P)-F	5' gat gaa tgt gag ccg gaa gaa cct gaa ctc aag ccc 3'
RAD2(W372P+V375P)-R	5' ggg ctt gag ttc agg ttc ttc cgg ctc aca ttc atc 3'
RAD2(F670P+V673P)-F	5' gaa gaa cta cct gag gac cct ccc act aaa tc 3'
RAD2(F670P+V673P)-R	5' gat tta gtg gga ggg tcc tca ggt agt tct tc 3'

SUPPLEMENTARY FIGURES

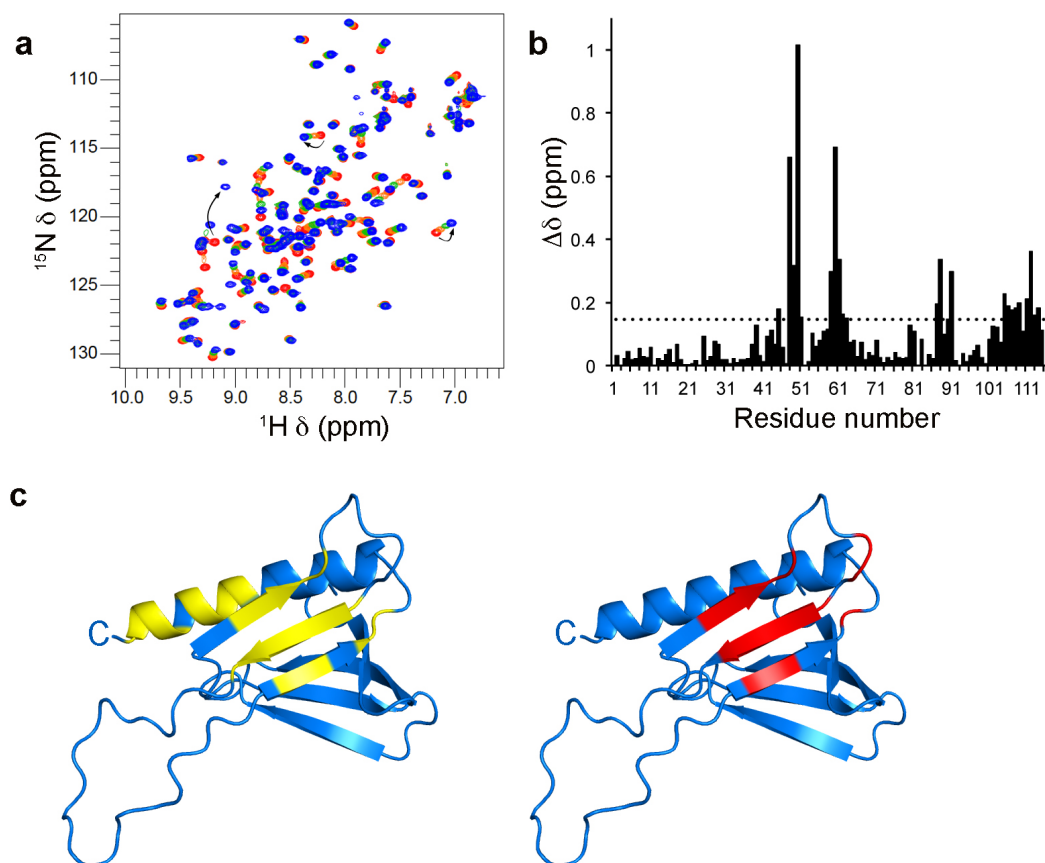


Supplementary Figure 1. Tfb1 and Rad2 protein expression levels in yeast. (a)

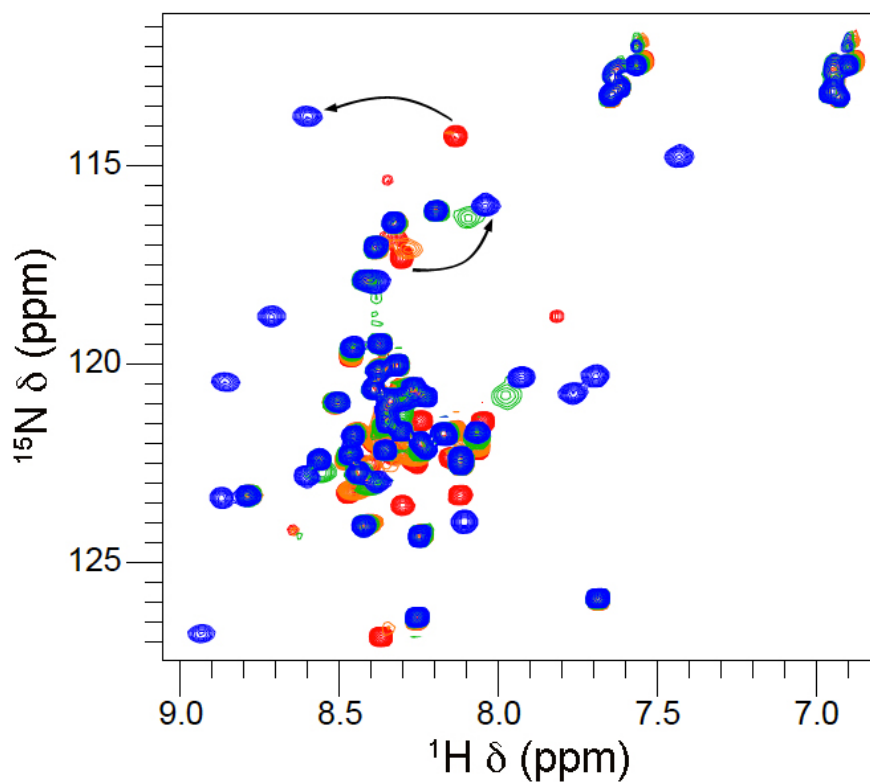
Level of expression of the Tfb1 and Tfb1- Δ PH proteins in *TFB1* and *tfb1- Δ PH* yeast. The protein levels were monitored by western blot with an antibody against Tfb1. Actin levels are shown as a loading control (arrow). **(b)** Level of expression of the Rad2, Rad2- Δ D1 and Rad2- Δ D3 proteins in *RAD2*, *rad2- Δ D1* and *rad2- Δ D3* yeast. The protein levels were monitored by western blot with an antibody against the myc tag. Actin levels are shown as a loading control (arrow).



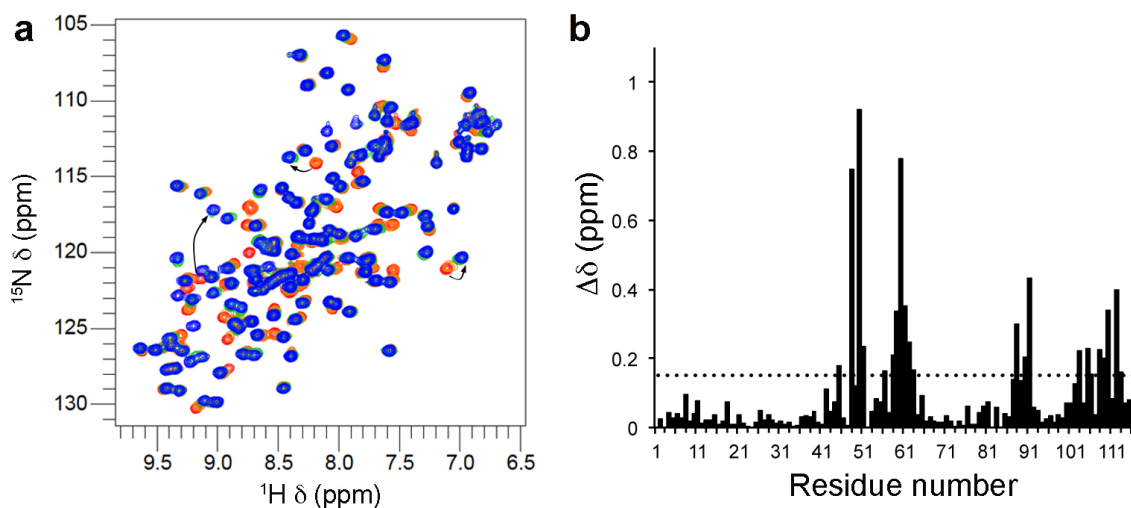
Supplementary Figure 2. Rad2₆₄₂₋₇₆₀ is a disordered peptide and possesses two distinct segments binding to Tfb1PH. (a) Thermogram of the Tfb1PH titration with successive additions of Rad2₆₄₂₋₇₆₀ at 25°C, in 20mM PO₄ pH 7.5 buffer. Fitting with a one-site binding model gave an N value of 0.52 suggesting two binding sites. (b) ¹H-¹⁵N HSQC of ¹⁵N-labeled Rad2₆₄₂₋₇₆₀. The narrow chemical shift dispersion of the signals indicates that there is no folded domain present.



Supplementary Figure 3. Rad₂₆₄₂₋₆₉₀ binds to Tfb1PH on a similar site as TFIIIE α CTD. (a) Overlay of the two dimensional ^1H - ^{15}N HSQC spectra for ^{15}N -labeled Tfb1PH in its free form (red), in the presence of 0.25 equivalents (orange), 0.5 equivalent (green) and 1 equivalent (blue) of unlabeled Rad₂₆₄₂₋₆₉₀. Arrows highlight several signals that undergo significant changes in ^1H and ^{15}N chemical shifts upon formation of the Rad₂₆₄₂₋₆₉₀-Tfb1PH complex. (b) Histogram of the variation in chemical shift [$\Delta\delta_{(\text{ppm})}$]; $\Delta\delta = [(0.17\Delta N_{\text{H}})^2 + (\Delta H_{\text{N}})^2]^{1/2}$] observed in the ^1H - ^{15}N HSQC spectrum of the ^{15}N -labeled Tfb1PH upon formation of the Rad₂₆₄₂₋₆₉₀-Tfb1PH complex. (c) Ribbon models of the three dimensional structure of Tfb1PH. The amino acids showing a significant chemical shift change ($\Delta\delta_{(\text{ppm})} > 0.15$) upon formation of a complex with Rad₂₆₄₂₋₆₉₀ are shown in yellow (left) and upon formation of a complex with TFIIIE α CTD are shown in red (right).



Supplementary Figure 4. Tfb1PH binds to Rad2₆₄₂₋₆₉₀. Overlay of the ^1H - ^{15}N HSQC spectra for ^{15}N -labeled Rad2₆₄₂₋₆₉₀ in its free form (red), in the presence of 0.25 equivalents (orange), 0.5 equivalent (green) and 1 equivalent (blue) of unlabeled Tfb1PH. Arrows highlight several signals that undergo significant changes in ^1H and ^{15}N chemical shifts upon formation of the Rad2₆₄₂₋₆₉₀-Tfb1PH complex.



Supplementary Figure 5. Rad2₃₅₉₋₃₈₃ binds to Tfb1PH on a similar site as Rad2₆₄₂₋₆₉₀. (a) Overlay of the ¹H-¹⁵N HSQC spectra for ¹⁵N-labeled Tfb1PH in its free form (red), in the presence of 0.25 equivalents (orange), 0.5 equivalent (green) and 1 equivalent (blue) of unlabeled Rad2₃₅₉₋₃₈₃. Arrows highlight several signals that undergo significant changes in ¹H and ¹⁵N chemical shifts upon formation of the Rad2₃₅₉₋₃₈₃-Tfb1PH complex. (b) Histogram of the variation in chemical shift [$\Delta\delta_{(\text{ppm})}$; $\Delta\delta = [(0.17\Delta N_{\text{H}})^2 + (\Delta H_{\text{N}})^2]^{1/2}$] observed in the ¹H-¹⁵N HSQC spectrum of the ¹⁵N-labeled Tfb1PH upon formation of the Rad2₃₅₉₋₃₈₃-Tfb1PH complex.

3.0 Caractérisation de l'interaction entre Rad4 et Tfb1

Suite à la découverte d'un motif d'interaction entre Rad2 et Tfb1 menant à un parallèle intéressant entre la réparation et la transcription de l'ADN, d'autres preuves étaient nécessaires pour soutenir notre hypothèse. Nous avons donc recherché un motif d'interaction similaire dans les séquences protéiques des partenaires d'interaction de TFIIH. Sachant que XPC/Rad4 est un de ces partenaires, nous nous sommes concentrés sur celui-ci.

Avant notre étude, aucun segment d'interaction précis entre Rad4 et TFIIH n'avait été identifié. Tout comme pour Rad2, des études précédentes nous ont fourni de bons indices. Le segment 1-167 de Rad4 interagit avec TFIIH (202). De plus, ce segment aminoterminal est désordonné et acide. Chez l'humain, XPC interagit avec XPB et p62 (70). Par contre, un paradoxe existe par le fait que le segment carboxy-terminal de XPC est responsable de cette interaction (204). Cependant, rien ne prouve que le reste de XPC n'est pas impliqué dans l'interaction, outre les résidus 1-117 qui n'interagissent pas avec TFIIH (204)

Nous avons retrouvé le motif de Rad2 interagissant avec Tfb1 chez Rad4, XPC, mais aussi chez Rad34 (équivalent de Rad4 réparant l'ADN ribosomal). Nous avons donc entrepris des études dans cette direction, dans le prochain article.

Pour cet article, j'ai effectué toutes les expériences *in vitro* (purifications de protéines, expériences ITC, enregistrement des spectres RMN, interprétation de ces spectres, extraction des NOE et calculs des structures). Geneviève Arseneault, a effectué les expériences *in vivo* chez la levure et Laurent Cappadocia a purifié le segment 359-383 de Rad2. Les expériences ont été planifiées par moi-même, Geneviève Arseneault et le Dr James Omichinski, directeur du laboratoire. J'ai rédigé l'article avec l'aide du Dr Omichinski. L'article sera soumis sous peu à *Nature Structural and Molecular Biology*.

**Structural and functional evidence that Rad4
competes with Rad2 for binding to the Tfb1 subunit of
TFIIH in NER.**

*Julien Lafrance-Vanasse, Geneviève Arseneault, Laurent Cappadocia and James G.
Omichinski*

Département de Biochimie, Université de Montréal, Montréal, Québec, Canada

Running Title: Rad4 and Rad2 compete for Tfb1.

ABSTRACT

XPC/Rad4 (human/yeast) recruits TFIIH to the nucleotide excision repair (NER) complex through interactions with its p62/Tfb1 and XPB/Ssl2 subunits. TFIIH then recruits XPG/Rad2 through interactions with similar subunits and the two repair factors appear to be mutually exclusive within the NER complex. Here, we show that Rad4 binds the PH domain of the Tfb1 (Tfb1PH) with high affinity. Structural characterization of a Rad4-Tfb1PH complex demonstrates the Rad4 binding interface is formed using a motif similar to one used by Rad2 to bind Tfb1PH. Furthermore, yeast with mutations in the Tfb1-binding motif of Rad4 display increased photosensitivity. The results support the hypothesis that XPG/Rad2 displaces XPC/Rad4 from the repair complex in part through interactions with the Tfb1/p62 subunit of TFIIH. The Rad4-Tfb1PH structure also provides detailed information regarding, not only, the interplay of TFIIH recruitment to the NER, but also links the role of TFIIH in NER and transcription. **(148 Words)**

INTRODUCTION

Exposure to ultraviolet (UV) irradiation can threaten the survival of organisms due to its ability to damage DNA by inducing modifications such as cyclobutane pyrimidine dimers (CPD) and (6,4) photoproducts (6,4PP). Due to the constant threat that UV irradiation poses, organisms have developed pathways like nucleotide excision repair (NER) to repair the damaged DNA¹. Failure to repair the damage ultimately threatens the integrity of the organism's genome, and it is now clear that defects in factors that function in the NER repair pathway results in a number of debilitating human diseases including Xeroderma Pigmentosum (XP)². XP is an autosomal recessive genetic disorder characterized by a hypersensitivity to UV irradiation, and the most common human form of XP is associated with mutations of the complementation group XP-C³. XP patients are defective in their ability to perform NER following exposure to UV irradiation, which results in an increased probability of acquiring a number of different forms of skin cancers⁴.

NER can be divided into two pathways depending on whether the damaged DNA that needs to be repaired is being actively transcribed (transcription-coupled NER; TC-NER) or not (global genome NER; GG-NER). Although these two pathways of NER share many similarities, they differ considerably in how the damaged DNA is initially recognized. In GG-NER, the XPC-RAD23B/Rad4-Rad23 (human/yeast) complex recognizes distorted DNA, mainly due to CPDs and 6,4PPs⁵. In contrast, during TC-NER, the RNA polymerase II (RNAP II) complex stalls on the damaged site, which leads to the recruitment of the Cockayne Syndrome B (CSB/Rad26) protein^{6,7}. Following this initial recognition of the damaged DNA site, the second step of both NER pathways involves the recruitment of the general transcription factor IIIH (TFIIH) due to the presence of either CSB/Rad26 (TC-NER) or XPC/Rad4 (GG-NER)⁶⁻¹⁰.

As a crucial component of the GG-NER pathway, XPC/Rad4 participates in a number of protein-protein and protein-DNA interactions. Studies with the full-length XPC/Rad4 demonstrated that the central region of Rad4 (residues 120-630) was sufficient for both specific binding to the CPD-damaged DNA and heterodimerization with Rad23¹¹. These studies also concluded that the amino-terminal (N-terminal; residues 1-120) and carboxyl-terminal (C-terminal; residues 630-754) regions of Rad4 are most likely

disordered based on their sensitivity to proteolysis and this is consistent with NMR studies showing that the C-terminal 125 amino acids of XPC (residues 816-940) are also disordered¹². Despite being disordered, several studies have shown that the N- and C-terminal regions of XPC/Rad4 play important roles in interactions with Centrin-2, Cdc31, Rad33 and TFIIH^{8,13-16}.

TFIIH is a 10-subunit complex that is highly conserved between humans and budding yeast (*Saccharomyces cerevisiae*). Once recruited to the damaged DNA site by XPC/Rad4, TFIIH plays several important functions. First, through the helicase activities of its XPB/Ssl2 and XPD/Rad3 subunits, TFIIH assists in unwinding the DNA in and around the damaged site¹⁷. Next, TFIIH recruits the 3'-endonuclease XPG/Rad2, which initiates the cutting of the DNA and the subsequent excision/repair processes¹⁸. In GG-NER, the recruitment of XPG/Rad2 to the damaged DNA site coincides with the release of the XPC-RAD23B/Rad4-Rad23 complex, and this exchange between XPC/Rad4 and XPG/Rad2 appears to be mediated through interactions involving TFIIH^{9,19}. As a result, a number of studies have examined the interaction between TFIIH and XPC/Rad4 to understand its role in NER. Co-immunoprecipitation (Co-IP) studies with Rad4 indicated that residues 1-167 were sufficient for binding to TFIIH⁸. In addition, *in vitro* and *in vivo* studies showed that both the XPB and p62 subunits of TFIIH directly interacted with XPC and that the C-terminal region of XPC played an important role in recruiting the TFIIH complex^{16,20}. Interestingly, the two subunits of TFIIH that interact with XPC/Rad4 (p62/Tfb1 and XPB/Ssl2) are also involved in recruiting XPG/Rad2^{21,22}, and this suggests these two repair factors may form similar types of interactions with TFIIH that leads to their exchange during NER.

Despite the important role that the XPC/Rad4-TFIIH complex plays in regulating GG-NER, there is currently no detailed information describing the regions of XPC/Rad4 that interact with either the p62/Tfb1 or XPB/Ssl2 subunits of TFIIH. In this manuscript, we characterize the interaction between Rad4 and the Tfb1 subunit of TFIIH. We identify an acidic segment from the N-terminal region of Rad4 (Rad4₇₆₋₁₁₅; residues 76-115) that binds with high affinity to the PH domain located at the N terminus of Tfb1 (Tfb1PH; residues 1-115). NMR spectroscopy studies show that Rad4₇₆₋₁₁₅ binds to the same region

of Tfb1PH as two homologous acidic segments from Rad2, and Rad4₇₆₋₁₁₅ competes with these Rad2 segments for binding to Tfb1PH. Structural characterization of a Rad4₇₆₋₁₁₅-Tfb1PH complex indicates that Rad4 binds to Tfb1PH using a very similar Tfb1PH-binding motif as Rad2²³. In addition, *in vivo* studies show that yeast with mutations of key residues of Rad4 required for interaction with Tfb1 display enhanced sensitivity to UV irradiation. Together these results provide a rationale for the role that TFIIH plays in the exchange between XPC/Rad4 and XPG/Rad2 during GG-NER.

RESULTS

The N-terminal segment of Rad4 binds Tfb1PH with high affinity

We previously identified two acidic segments within the spacer region of Rad2 (Rad2₃₅₉₋₃₈₃ and Rad2₆₄₂₋₆₉₀) that contain a high affinity Tfb1PH-binding motif²³. Based on the hypothesis that Rad4 and Rad2 form similar interactions with TFIIH, we analyzed the sequences of Rad4 and XPC to see if they also contain a region similar to the Tfb1PH-binding motif found in Rad2. The requirements for the two Tfb1PH-binding motifs of Rad2 are that they are located within highly acidic segments, and they contain an aromatic residue (W or F) followed by two acidic residues and a valine residue. As shown in **Figure 1a**, sequences similar to the Tfb1PH-binding motifs of Rad2 are located near the N termini of both Rad4 (residues 86-106) and XPC (residues 124-144). This observation is consistent with the first 167 amino acids of Rad4 being sufficient for binding to TFIIH and the fact that the p62/Tfb1 subunit of TFIIH has been shown to directly interact with XPC/Rad4^{8,16,20}.

To determine if the N-terminal region of Rad4 does in fact bind to Tfb1PH, the apparent dissociation constant (K_d) between Tfb1PH and Rad4₇₆₋₁₁₅ was measured by isothermal titration calorimetry (ITC; **Fig. 1b-c**). As we observed with Rad2₃₅₉₋₃₈₃ and Rad2₆₄₂₋₆₉₀ ($K_d = 130$ nM and $K_d = 190$ nM, respectively), Rad4₇₆₋₁₁₅ binds to Tfb1PH with high affinity ($K_d = 50$ nM). Next, we wanted to verify that Phe95 and Val98 of Rad4 were crucial residues in the Tfb1PH-binding motif of Rad4. These residues correspond to Phe670 and Val673 of Rad2₆₄₂₋₆₉₀ and are crucial for Rad2 binding to Tfb1PH²³. Mutation of either Phe95 or Val98 to proline (F95P and V98P) reduces the affinity of Rad4₇₆₋₁₁₅ for

Tfb1PH by 50- fold and 24-fold, respectively (**Fig. 1b**), supporting that the Tfb1PH-binding site of Rad4 is very similar to the two sites found in Rad2.

Rad4₇₆₋₁₁₅, Rad2₃₅₉₋₃₈₃ and Rad2₆₄₂₋₆₉₀ bind a common site on Tfb1PH

To identify the binding site of Rad4₇₆₋₁₁₅ on Tfb1PH, NMR chemical shift perturbation studies were performed. In these experiments, addition of unlabeled Rad4₇₆₋₁₁₅ to ¹⁵N-labeled Tfb1PH causes significant changes in the ¹H and ¹⁵N chemical shifts for several Tfb1PH signals in the ¹H-¹⁵N HSQC spectra (see **Supplementary Fig. S1a-b**). When mapped onto the structure of Tfb1PH, the residues exhibiting significant changes are located in strands β5, β6, β7 and helix H1 (**Fig. 2a**). The residues experiencing significant changes are almost identical to those observed when either of the two acidic segments from Rad2 (Rad2₃₅₉₋₃₈₃ and Rad2₆₄₂₋₆₉₀) bind to Tfb1PH (**Fig. 2b and Fig. S1c**).

To confirm that Rad4₇₆₋₁₁₅ shares a common binding site on Tfb1PH with Rad2₃₅₉₋₃₈₃, NMR competition experiments were performed. In these experiments, we first add a substoichiometric concentration of unlabeled Tfb1PH (0.4 mM) to a sample containing ¹⁵N-labeled Rad4₇₆₋₁₁₅ (0.5 mM) and, as expected, we observe significant changes for several ¹H and ¹⁵N chemical shifts in the ¹H-¹⁵N HSQC spectra of Rad4₇₆₋₁₁₅ (**Fig. 2c**). We then add an equimolar amount of unlabeled Rad2₃₅₉₋₃₈₃ (0.5 mM) to the ¹⁵N- Rad4₇₆₋₁₁₅ -Tfb1PH complex and observe that the ¹H and ¹⁵N resonances of Rad4₇₆₋₁₁₅, which shift upon formation of the Rad4₇₆₋₁₁₅-Tfb1PH complex return to values characteristic of the free form of Rad4₇₆₋₁₁₅ (**Fig. 2d**). Taken together with previous results showing that Rad2₃₅₉₋₃₈₃ and Rad2₆₄₂₋₆₉₀ compete for binding to Tfb1PH²³, these results demonstrate that Rad4₇₆₋₁₁₅, Rad2₃₅₉₋₃₈₃ and Rad2₆₄₂₋₆₉₀ all compete for a common binding site on Tfb1PH.

The Rad4 homolog Rad34 also contains a Tfb1PH-binding motif

Rad34 is a repair factor found in yeast that plays a role similar to Rad4 in TC-NER of ribosomal DNA (rDNA)^{24,25}. Given the fact that TFIIF also plays a critical role in RNA

polymerase I (RNAP I)-dependent transcription and TC-NER of rDNA²⁶, we were interested to determine if Rad34 contained a Tfb1PH-binding motif. The Rad34 protein is homologous to the Rad4 protein, and it contains both a DNA-binding domain that recognizes damaged DNA and a Centrin-2/Rad33 binding motif within its C-terminal region²⁷. Although there is currently no evidence that Rad34 interacts with TFIIH, we postulated based on its homology to Rad4, that it also contains a Tfb1PH-binding motif near the N terminus of the protein. Based on sequence homology with Rad4₇₅₋₁₁₆, we identified a segment between residues 41 and 63 of Rad34 (Rad34₄₁₋₆₃; **Fig. 3a**) as a potential Tfb1PH-binding motif, since it contains a highly acidic stretch that includes an aromatic residue (Trp54) followed by two acidic residues and a valine (Val57). ITC studies show that Rad34₄₁₋₆₃ also binds to Tfb1PH with high affinity ($K_d = 11.0$ nM; **Fig. 3b**) and that mutation of Trp54 to a serine (W54S mutant) significantly reduces binding to Tfb1PH (over 3 orders of magnitude). To further define the mode of binding of Rad34₄₁₋₆₃ to Tfb1PH, an NMR chemical shift perturbation study was performed. Like for Rad4₇₆₋₁₁₅, addition of Rad34₄₁₋₆₃ to ¹⁵N-Tfb1PH causes significant changes in the ¹H and ¹⁵N chemical shifts for several Tfb1PH signals in the ¹H-¹⁵N HSQC spectra (**Supplementary Fig. 2**) and the residues exhibiting significant changes are located in strands β 5, β 6, β 7 and the H1 helix (**Fig. 3c**).

Mutations in both the N- and C-terminal TFIIH-binding regions of Rad4 leads to enhanced photosensitivity

Previous studies have shown that both the N- and C-terminal regions of XPC/Rad4 interact with TFIIH^{8,13-16}. To examine the functional role of the two TFIIH-binding regions of Rad4 in yeast survival following exposure to UV irradiation, we first prepared a yeast strain in which the two key hydrophobic residues (Phe95 and Val98) in the Tfb1PH-binding motif were mutated to proline (*rad4-PP*) and we tested it for survival following exposure to UV irradiation. The *rad4-PP* strain shows only a slightly higher photosensitivity compared to the wild-type *RAD4* strain (**Fig. 4a**). Next, we tested the role of the C terminus of Rad4. Previous studies have shown that mutating three key

hydrophobic residues (Trp649, Leu652 and Leu656) in this region to alanine leads to yeast strains (*rad4-AAA*) with photosensitivity as well as decreased binding to Centrin-2, Rad33 and TFIIH¹⁴. Consistent with previous studies, the *rad4-AAA* mutant strain is more photosensitive than the *RAD4* strain and slightly more sensitive than the *rad4-PP* strain (**Fig. 4a**).

To further examine the *in vivo* role of the two TFIIH-binding regions of Rad4, the N-terminal mutation (*rad4-PP*) and the C-terminal mutation (*rad4-AAA*) were combined to produce an additional mutant strain (*rad4-PPAAA*). The survival curves show that the *rad4-PPAAA* strain is considerably more photosensitive than either the *rad4-PP* or *rad4-AAA* strains and displays similar photosensitivity to what is observed with the *rad4* deletion strain (**Fig. 4b**). This increased photosensitivity is not the result of decreased levels of Rad4 as the Rad4-PPAAA mutant protein is expressed at similar levels as the wild-type Rad4 (**Supplementary Fig. S3**). Taken together, the *in vivo* studies in yeast indicate that the Tfb1PH-binding motif of Rad4 plays a key role in yeast survival following exposure to UV irradiation. In addition, like Rad2, mutating both of the known TFIIH-binding regions of Rad4 produces an additive effect on yeast photosensitivity, suggesting that multiple interactions between Rad4 and TFIIH are required for formation of a stable repair complex.

NMR structure determination of the Rad4₇₆₋₁₁₅-Tfb1PH complex

To compare Rad2 and Rad4 in complex with Tfb1PH, we determined the NMR solution structure of a Rad4₇₆₋₁₁₅-Tfb1PH complex. The structure of the Rad4₇₆₋₁₁₅-Tfb1PH complex (**PDB code XXXX**) is well defined by the NMR data (**Table 1**). The 20 lowest-energy structures (**Fig. 5a**) are characterized by good backbone geometry, no significant restraint violation and low pair-wise rmsd values (**Table 1**). As is the case for the Rad2₆₄₂₋₆₉₀-Tfb1PH complex, the structure of Tfb1PH in complex with Rad4₇₆₋₁₁₅ is virtually identical to its free form showing a typical PH domain fold consisting of a seven-stranded β sandwich (β 1- β 7) flanked on one side by a long α helix (H1)²⁸. In complex with Tfb1PH, Rad4₇₆₋₁₁₅ binds in an extended conformation devoid of any regular secondary structural element with residues 95-98 forming a crucial portion of the interface with Tfb1PH. This is

consistent with the ^1H - ^{15}N HSQC spectra of the titration of ^{15}N -labeled Rad4₇₆₋₁₁₅ with Tfb1PH as residues 95-98 undergo the most significant changes in both their ^1H and ^{15}N chemical shifts (**Supplementary Fig. S4**).

Rad4₇₆₋₁₁₅-Tfb1PH binding interface

In the complex, Rad4₇₆₋₁₁₅ binds in an extended form to two adjacent shallow grooves on the surface of Tfb1PH. The first groove is formed by residues Gln49, Ala50, Thr51, Pro52, Met59, Leu60, Arg61 and Met88 from strands β 5, β 6 and β 7 of Tfb1PH (**Fig. 5b**). Phe95 of Rad4 inserts into this groove where it is in position to form a cation- π interaction with Arg61. The second groove is composed of Leu48, Ala50, Lys101 and Gln105, Ile108, Lys112 of Tfb1PH and accommodates Val98 and Thr99 of Rad4 (**Fig. 5c**). Val98 is anchored on one side of this pocket through van der Waals interactions with both its methyl groups. Thr99 is anchored on the other side of the pocket through van der Waals interactions with its methyl group. Although the majority of the interactions within the two grooves are van der Waals contacts, an extensive series of positively charged residues on the surface of Tfb1PH (Lys47, Lys57, Arg61, Arg86, Lys97, Lys101 and Lys112) surround the two grooves, where they function to position the negatively charged Rad4₇₆₋₁₁₅. The NMR structures support the formation of one potential salt bridge between Asp97 of Rad4₇₆₋₁₁₅ and Lys47 of Tfb1PH.

Comparison of the Rad4₇₆₋₁₁₅-Tfb1PH complex with the Rad2₆₄₂₋₆₉₀-Tfb1PH complex

The structure of the Rad4₇₆₋₁₁₅-Tfb1PH interface is virtually identical to the previous structure of the Rad2₆₄₂₋₆₉₀-Tfb1PH interface (**Fig. 6**). In both structures, a phenylalanine (Phe95 in Rad4 and Phe670 in Rad2) binds in a first pocket, where it is in position to make a cation- π interaction with Arg61 of Tfb1PH. In addition, a valine (Val98 in Rad4 and Val673 in Rad2) inserts into a second pocket, where it makes van der Waals contacts with a

series of residues in Tfb1PH. Rad4 and Rad2 also contain a threonine (Thr99 in Rad4 and Thr675 in Rad2) that forms van der Waals contacts in the second binding pocket. In Rad2₆₄₂₋₆₉₀, Val673 and Thr675 are separated by Pro674, whereas in Rad4₇₆₋₁₁₅, Val98 and Thr99 are consecutive residues. In Rad2₆₄₂₋₆₉₀, Pro674 creates a slight bend in the backbone so that the side-chains of Val673 and Thr675 can both be oriented towards the surface of Tfb1PH. An additional commonality between the two structures is that Rad4₇₆₋₁₁₅ and Rad2₆₄₂₋₆₉₀ contain several acidic residues that participate in electrostatic interactions with basic amino acids that surround the perimeter of the two binding pockets of Tfb1PH. Overall, the similarities between two interfaces provide a structural framework to understand how Rad4 and Rad2 can compete for binding to Tfb1.

DISCUSSION

Following DNA damage recognition, the recruitment of TFIIH by XPC/Rad4 to the repair complex occurs through direct interactions with its p62/Tfb1 and XPB/Ssl2 subunits^{16,20}. Once at the damaged DNA, TFIIH recruits XPG/Rad2 through direct interactions with multiple subunits including both its p62/Tfb1 and the XPB/Ssl2 subunits^{21,22}. Given that XPC/Rad4 and XPG/Rad2 interact with the same subunits of TFIIH and that these two factors appear to be mutually exclusive within the repair complex^{9,19}, XPG/Rad2 likely displaces XPC/Rad4 from the pre-incision complex in part by competing for common binding sites on the various subunits of TFIIH. Thus, the competition observed for regions of Rad4 and Rad2 binding to Tfb1PH by NMR and the remarkable similarity between the structures of the Rad4₇₆₋₁₁₅-Tfb1PH and Rad2₆₄₂₋₆₉₀-Tfb1PH interfaces provides a molecular mechanism to rationalize the mutual exclusivity observed between XPC/Rad4 and XPG/Rad2 at the repair complex *in vivo*. Although a hand-off or displacement mechanism has previously been suggested between XPC/Rad4 and XPG/Rad2 in NER²⁹, this is the first structural evidence to support such a displacement mechanism.

This displacement of XPC/Rad4 by XPG/Rad2 could serve at least two possible purposes for maintaining the NER complex in a repair competent form. First, it assures that TFIIH is always stabilized on the DNA and available for repair as opposed to being recruited to perform alternative functions such as transcription. Second, it helps to establish the correct orientation of XPG/Rad2 on the DNA lesion. XPG/Rad2 cleaves the phosphate backbone of DNA on the 3' side of the lesion, and it is crucial that this is done in a precise fashion for NER to be carried out in an accurate manner. Since XPC/Rad4 specifically interacts on the same side of the damaged DNA¹¹, a sequential displacement mechanism would serve to place XPG/Rad2 in the correct orientation for precise repair of the DNA lesion without the need for releasing TFIIH. This model is similar to the “passing the baton” mechanism, seen in base excision repair (BER)³⁰. In the BER model, the substrate is channeled from one protein to another as opposed to a pre-formed protein complex that possesses all the activities necessary for repair.

The fact that both the N-terminal Tfb1PH-binding motif and the C-terminal TFIIH/Centrin-2 binding region of Rad4 play a role in yeast survival following exposure to UV irradiation is consistent with the sequential displacement mechanism between XPC/Rad4 and XPG/Rad2 in NER. In the case of XPG/Rad2, it has also been shown to interact with multiple subunits of TFIIH, including p62 and XPB, by participating in a series of weak interactions with its extended spacer region^{21,31,32}. In the case of XPC/Rad4, it has been shown to directly interact with both the p62/Tfb1 and XPB/Ssl2 subunits of TFIIH through regions in both termini of the protein^{9,19} and the *in vivo* results in yeast support the importance of both these regions. The fact that a mutation of the C-terminal region of Rad4 has a slightly more dramatic effect on yeast photosensitivity than a mutation of the N-terminal region, is consistent with previous results showing that the C-terminal region of XPC appeared more critical for TFIIH binding¹⁶. Taken together with our current results, this indicates that although displacement of both the N-terminal and the C-terminal region of XPC/Rad4 by XGP/Rad2 is important for NER, the displacement of the C-terminal region may be the rate-limiting step. However, the details of the interaction between the C-terminal region of Rad4 and TFIIH are currently unknown and it is still not clear in which order the displacement of the two regions of XPC/Rad4 occurs.

The sequential displacement mechanism clearly presents advantages for maintaining the integrity and accuracy of the repair complex, however, it may also be important for sequestering TFIIH. We have previously shown that the Rad2₆₄₂₋₆₉₀-Tfb1PH interface is similar to the interface formed in a complex between the C-terminal domain of TFIIIE α (TFIIIE α CTD) and p62PH^{23,33}. Like Rad4 and Rad2, TFIIIE α CTD contains a Tfb1PH-binding motif that consists of an aromatic residue followed by two acidic residues and a valine³⁴. We previously suggested that the similarity of the Tfb1PH-binding motif in Rad2 and TFIIIE α , provides a mechanistic link between the role of TFIIH in transcription and repair. Our current work further strengthens this concept since a highly similar Tfb1PH-binding motif is present in XPC, Rad4 and Rad34. TFIIH plays an important role both in RNAP I-dependent transcription as well as in NER of rDNA²⁶, and Rad34 likely forms a similar interaction with Tfb1 that may regulate RNAP I-associated NER. Based on the structural similarities of the Rad4₇₆₋₁₁₅-Tfb1PH, Rad2₆₄₂₋₆₉₀-Tfb1PH and the TFIIIE α CTD-

p62PH interfaces, it is tempting to propose that the interplay of TFIIIE α , Rad4 and Rad2 interacting with p62/Tfb1 helps to determine whether TFIIH is functioning in transcription or repair. In RNAP II-dependent transcription, TFIIIE α CTD recruits TFIIH to the pre-initiation complex (PIC) through interactions with the p62/Tfb1 subunit of TFIIH³⁴. Once at the PIC, TFIIH serves multiple functions during transcription by virtue of its helicase (XPB/Ssl2 and XPD/Rad3) and kinase (Cdk7/Kin28) activities^{35,36}. During NER, TFIIH's role in transcription would be limited by its recruitment to the DNA lesion through interactions with Rad4 that involves in part the same p62/Tfb1 subunit. As repair progresses past the initial recognition of the DNA lesion to the incision steps, TFIIH remains associated with the repair complex through interactions with XPG/Rad2, and this would make the process more efficient since it would help to prevent TFIIH being recruited for transcription.

In conclusion, we provide the first structural evidence for the displacement mechanism between XPC/Rad4 and XPG/Rad2 during NER. Our data demonstrates that XPC/Rad4 and XPG/Rad2 form similar interactions with the p62/Tfb1 subunit of TFIIH, and this supports the hypothesis that XPG/Rad2 displaces XPC/Rad4 from the repair complex through a series of interactions with different subunits of TFIIH. In addition, the structure of the Rad4₇₆₋₁₁₅-Tfb1PH complex provides information regarding the interplay of TFIIH recruitment to the PIC and the NER complex. Future studies are needed to determine if other subunits of TFIIH, in particular XPB/Ssl2, also form similar interactions with XPC/Rad4 and XPG/Rad2, and what roles they play in the displacement of XPC/Rad4 by XPG/Rad2.

METHODS

Strains, media and vectors

S. cerevisiae strains used for these studies are listed in **Supplementary Table S1**. The strains were grown a synthetic complete media (SC: 0.67% yeast nitrogen base without amino acids, 2% glucose and a mixture of amino acids and vitamins) lacking uracil (SC-U) for selection. All yeast transformations were performed using the modified lithium acetate protocol³⁷.

Plasmid preparations

The pRS316RAD4cmyc plasmid (*RAD4*) was generated by amplification of the *RAD4* Open Reading Frame (ORF) complemented by 190 bp upstream and 282 bp downstream on genomic DNA and insertion into pRS316. The pRS316RAD4(F95P/V98P)cmyc plasmid (*rad4-PP*) was obtained by the QuikChange. The pRS316RAD4(W649A/L652A/L656A)cmyc (*rad4-AAA*) and pRS316RAD4(F95P/V98P/W649A/L652A/L656A)cmyc (*rad4-PPAAA*) plasmids were both obtained by overlapped PCR using either pRS316RAD4cmyc or pRS316RAD4(F95P/V98P)cmyc respectively as the cloning vector. (For details see **Supplementary Methods**).

Sensitivity assays

Yeast strains were grown overnight at 30°C in selective media. The next day they were diluted in order to obtain an OD₅₉₅ = 0.5-1 the following morning. The cells were then harvested by centrifugation, washed and resuspended in sterile water to obtain an OD₅₉₅ = 0.5. For UV-sensitivity assays, dilutions were plated on selective media (SC-U) and irradiated with UV light (XL-1000 UV crosslinker, SpectroLinker) at varying energy levels. The surviving colonies were counted after 3 days growth at 30°C, in the dark.

Cloning and purification of proteins

The GST-Tfb1PH (residues 1-115 of Tfb1) was prepared as described²⁸. GST-Rad4₇₆₋₁₁₅ was prepared by inserting the appropriate region of Rad4 (Open Biosystems) into the pGEX-2T expression vector. GST-Rad34₄₁₋₆₃ was prepared by inserting the appropriate region of Rad34 (Open Biosystems) into the pGEX-2T expression vector. All point mutants were made using QuickChange. All coding sequences were verified by DNA sequencing. Tfb1PH was purified as described²⁸. Rad4₇₆₋₁₁₅, GST-Rad34₄₁₋₆₃ and their mutants were expressed as GST-fusion proteins in *E. coli* host strain TOPP2, purified over GSH resin (GE Healthcare) and cleaved with thrombin (Calbiochem), as previously described for Tfb1PH²⁸. Following cleavage with thrombin, the proteins were purified over a Q-Sepharose High Performance (GE Healthcare) column and dialyzed into appropriate buffers for ITC and NMR studies. ¹⁵N-labeled and ¹⁵N/¹³C-labeled proteins were prepared in M9-minimal media containing ¹⁵NH₄Cl (Sigma) and/or ¹³C₆-glucose (Sigma), as the sole nitrogen and carbon sources. For all experiments, the concentrations of proteins were determined from A₂₈₀.

ITC experiments

ITC titrations were performed as described³⁸ at 25°C in 20 mM sodium phosphate buffer (pH 7.5). All titrations fit a single-binding site mechanism with 1:1 stoichiometry and values are the average of two or more separate experiments.

NMR experiments

The NMR chemical shift perturbation and competition experiments were performed as previously described (for sample details please see **Supplementary Methods**). For the NMR structural studies of the Rad4₇₆₋₁₁₅-Tfb1PH complex, four different samples

containing 1.0 mM of the complex in a 1:1.25 ratio were used (^{15}N -Tfb1PH-Rad4₇₆₋₁₁₅, $^{15}\text{N}/^{13}\text{C}$ -Tfb1PH-Rad4₇₆₋₁₁₅, ^{15}N -Rad4₇₆₋₁₁₅-Tfb1PH and $^{15}\text{N}/^{13}\text{C}$ -Rad4₇₆₋₁₁₅-Tfb1PH, respectively). All NMR experiments were carried out in 20 mM sodium phosphate (pH 6.5), 1 mM EDTA, 1 mM DTT and 90% H_2O /10% D_2O or 100% D_2O , at 300 K on Varian Unity Inova 500, 600 and 800 MHz spectrometers equipped with z pulsed-field gradient units and triple resonance probes. All of the ^1H , ^{15}N , and ^{13}C resonances for Rad4₇₆₋₁₁₅ and Tfb1PH were assigned as reported for free Tfb1PH³⁹. Briefly, 3D HNC O ⁴⁰, 3D HNCACB⁴¹, 3D CBCACONH⁴², 3D (H)C(CO)NH⁴³, 3D H(CCO)NH⁴³ and 3D HCCH-COSY⁴⁴ spectra were used to assign the backbone and aliphatic side chains resonances. The aromatic side chains ^1H , ^{13}C and ^{15}N resonances were assigned using a combination of 2D (HB)CB(CGCD)HD and 2D (HB)CB(CGCDCE)HE spectra⁴⁵. Interproton distance restraints were measured from 3D ^{15}N -edited NOESY-HSQC, ^{13}C -edited HMQC-NOESY ($\tau_m=90$ ms) and 3D $^{15}\text{N}/^{13}\text{C}$ {F1}-filtered, {F3}-edited NOESY spectra ($\tau_m=90$ ms)^{46,47}. The NMR data were processed with NMRPipe/NMRDraw⁴⁸ and analyzed with NMRView⁴⁹ and Analysis from the CCPNMR suite⁵⁰.

Structure calculations

The NOE-derived distance restraints were divided into four classes defined as strong (1.8-2.8 Å), medium (1.8-4.0 Å), weak (1.8-5.0 Å) and very weak (3.3-6.0 Å). Backbone dihedral angles were derived with the program TALOS+⁵¹. The structure of the Rad4₇₆₋₁₁₅-Tfb1PH complex was calculated using the program CNS⁵². The quality of the structures was analyzed with the programs PROCHECK-NMR⁵³ and MOLMOL⁵⁴. The figures were generated with the program PyMol⁵⁵.

ACKNOWLEDGEMENTS

We would like to thank Aurélie Bernier and Antoine Chiasson for help with protein purification and Dr Tara Sprules for assistance with several NMR experiments. 800 MHz NMR experiments were recorded at the Québec/Eastern Canada High Field NMR Facility, supported by the Natural Sciences and Engineering Research Council of Canada. This work was supported by a grant from the Canadian Cancer Society (J.G.O.). J.L-V. is a Vanier Canada Graduate Scholar from the CIHR.

FIGURES

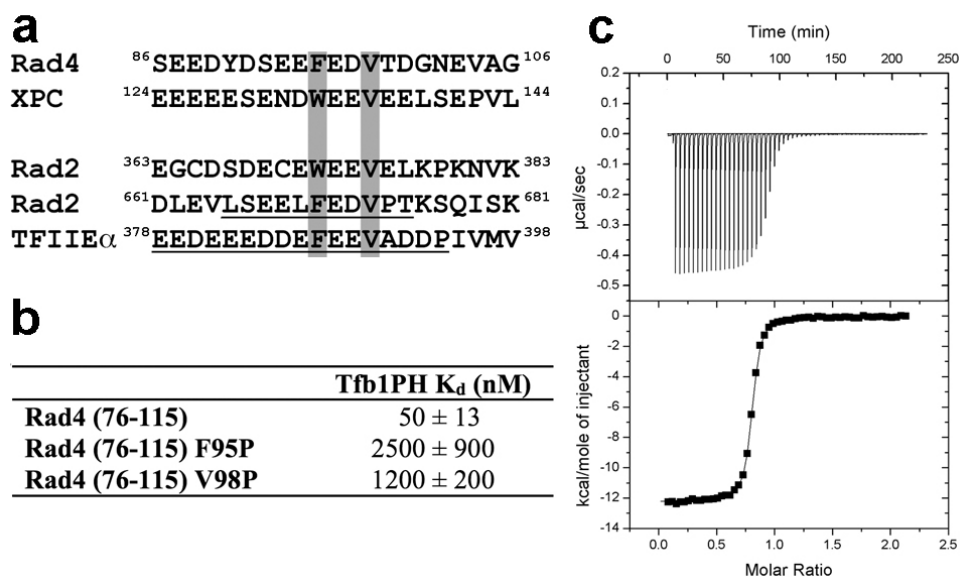


Figure 1. The N-terminal region of Rad4 contains a high affinity Tfb1PH-binding site between residues 76 and 115.

a. Identification of an amino acid segment located between residues 86 to 106 of Rad4 that aligns with the Tfb1PH-binding motif found in Rad2 and TFIIE α CTD. A similar motif is also located between residues 124 to 144 of XPC. In these alignments, the residues of TFIIE α CTD and Rad2₆₄₂₋₆₉₀ that form the binding interface with p62PH/Tfb1PH are underlined and crucial hydrophobic residues are shaded in grey. **b.** Comparison of the dissociation constant (K_d) values for the binding of Rad4₇₆₋₁₁₅ and its mutants (F95P and V98P) to Tfb1PH. **c.** Thermogram of the Tfb1PH titration with successive additions of Rad4₇₆₋₁₁₅. Experiments are performed at 25°C, in 20 mM NaPO₄ pH 7.5 buffer, and the results fit to a single-binding site model with 1:1 stoichiometry.

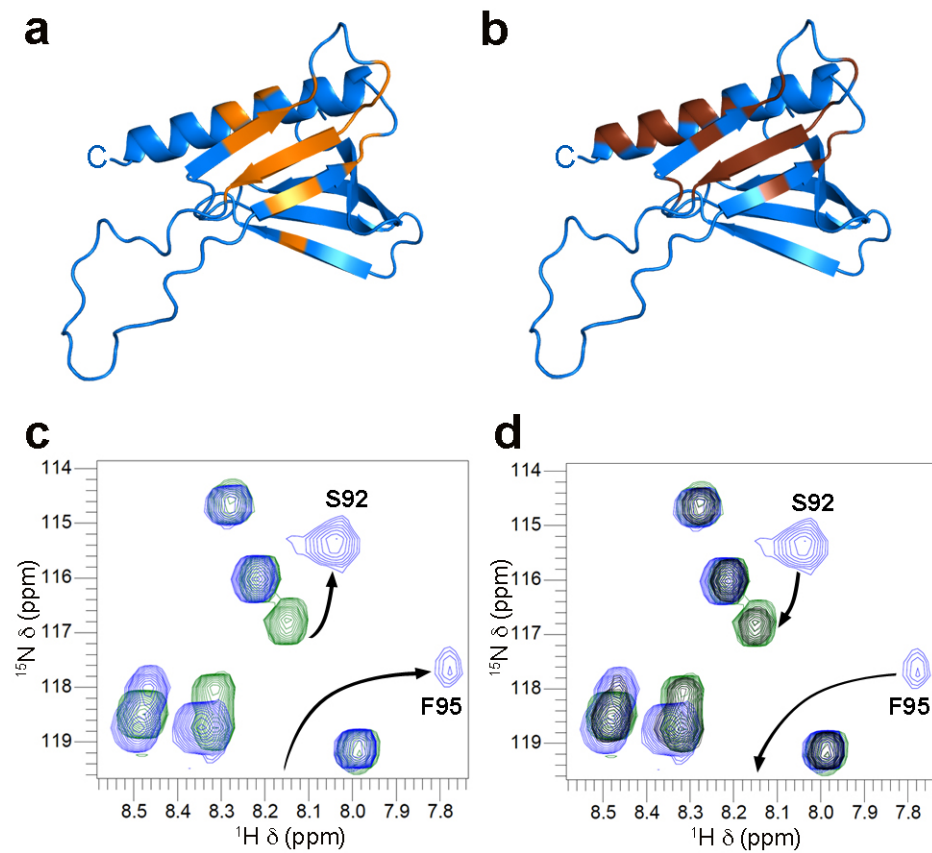


Figure 2. Rad4₇₆₋₁₁₅ and Rad2₃₅₉₋₃₈₃ share a common binding site on Tfb1PH. a-b. Ribbon model of the three-dimensional structure of Tfb1PH (blue; PDB code 1Y5O). The amino acids of Tfb1PH showing a significant chemical shift change $\{\Delta\delta(\text{ppm}) > 0.15; \Delta\delta = [((0.17\Delta N_H)^2 + (\Delta H_N)^2)]^{1/2}\}$ upon formation of a complex with either Rad4₇₆₋₁₁₅ (**a**) or Rad2₃₅₉₋₃₈₃ (**b**) are highlighted in orange and brown, respectively. **c.** Overlay of a selected region from the ¹H-¹⁵N HSQC spectra of ¹⁵N-labeled Rad4₇₆₋₁₁₅ (0.5 mM) in the free form (green) and in the presence of unlabeled Tfb1PH (0.4 mM; blue). **d.** Same overlay as in c, but after the addition of unlabeled Rad2₃₅₉₋₃₈₃ (1.5 mM; black). Rad4₇₆₋₁₁₅ signals that undergo significant changes in ¹H and ¹⁵N chemical shifts upon formation of the complex with Tfb1PH (**c**), and return towards their original position following the addition of Rad2₃₅₉₋₃₈₃ (**d**) are indicated by arrows.

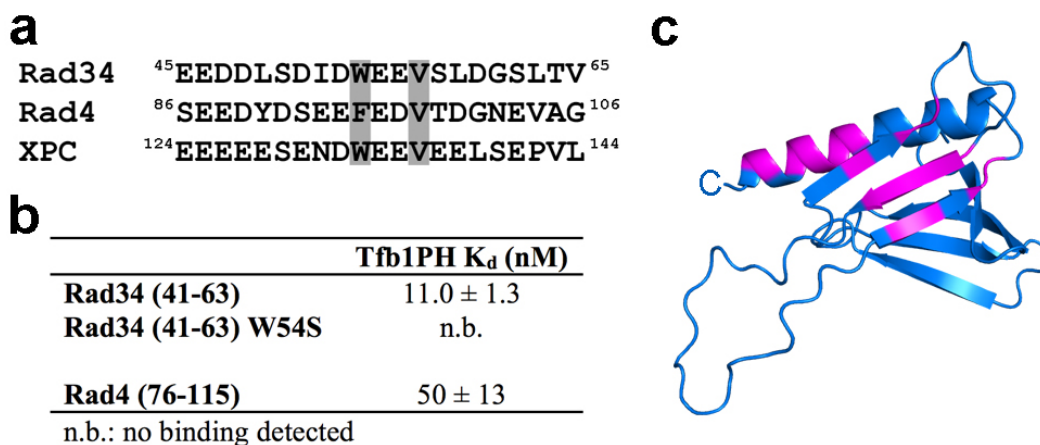


Figure 3. Rad34 contains a Tfb1PH-binding motif. **a.** Identification of an amino acid segment located between residues 41 to 63 from Rad34 that aligns with the Tfb1PH-binding motif from Rad4 and XPC. The two crucial hydrophobic residues in the motif are shaded in grey. **b.** Comparison of the dissociation constant (K_d) values for the binding of Rad4₇₆₋₁₁₅ and Rad34₄₁₋₆₃ with Tfb1PH. No binding is observed with the W54S mutant of Rad34₄₁₋₆₃ under the experimental conditions indicating a $K_d > 10 \mu\text{M}$. **c.** Ribbon model of the three dimensional structure of Tfb1PH (blue). The amino acids of ¹⁵N-labeled Tfb1PH showing a significant chemical shift change $\{\Delta\delta(\text{ppm}) > 0.15; \Delta\delta = [(0.17\Delta N_H)^2 + (\Delta H_N)^2]^{1/2}\}$ upon formation of a complex with Rad34₄₁₋₆₃ are highlighted in magenta.

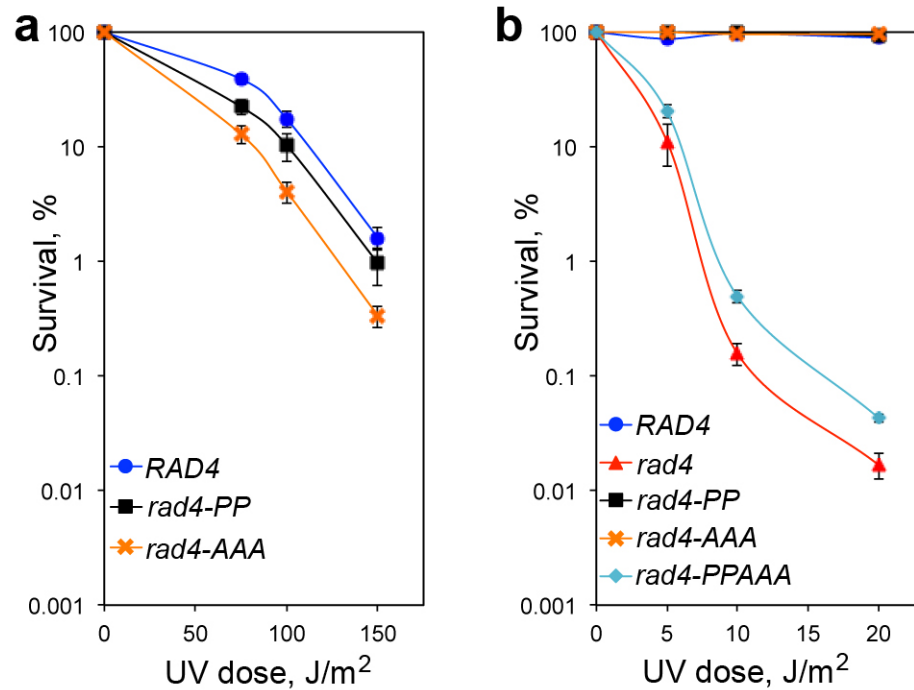


Figure 4. The two TFIIH-binding regions of Rad4 are required for survival following UV irradiation. a. The survival of *RAD4* (blue), *rad4-PP* (black) and *rad4-AAA* yeast were determined following increasing doses of UV irradiation. **b.** The survival of *RAD4* (blue), *rad4* (red), *rad4-PP* (black), *rad4-AAA* (orange) and *rad4-PPAAA* (aqua) yeast were determined following increasing doses of UV irradiation. In both **a** and **b**, the y axis represents the percentage of surviving cells (normalized to the number of viable cells not exposed to UV light) and the x axis shows the energy levels of the UV irradiation applied (J/m²). The results are the mean \pm SEM of three independent experiments.

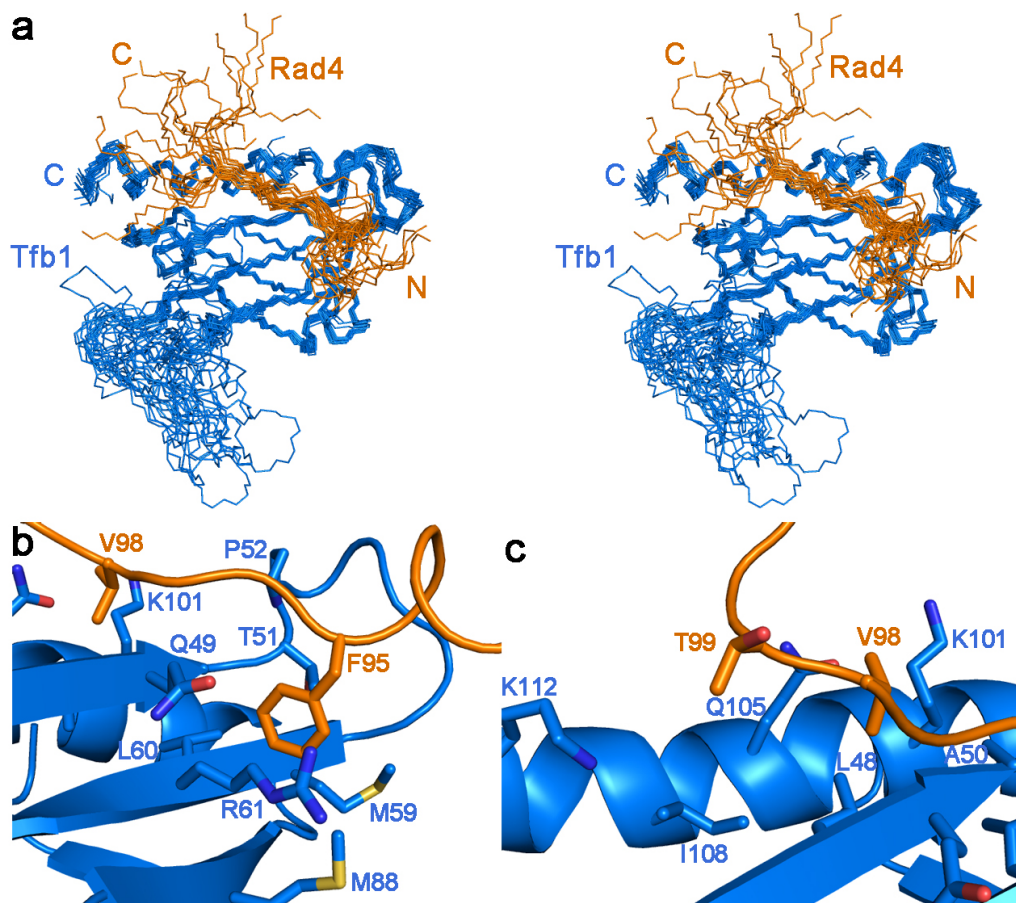


Figure 5. NMR structure of the Rad4₇₆₋₁₁₅-Tfb1PH complex. **a.** Stereo view of the 20 lowest-energy structures of the complex between Tfb1PH (blue) and Rad4₇₆₋₁₁₅ (orange; PDB code XXXX). The structures were superimposed using the backbone atoms C', C^α, and N of residues 4–65 and 85–112 of Tfb1PH and residues 90-104 of Rad4₇₆₋₁₁₅. **b.** Ribbon representation of Tfb1PH (blue) and backbone trace of the region of Rad4₇₆₋₁₁₅ (orange) interacting in the first binding pocket. In this pocket, Phe95 of Rad4 forms a cation- π interaction with Arg61 of Tfb1 and van der Waals interactions with Met59. **c.** Ribbon representation of Tfb1PH (blue) and backbone trace of the region of Rad4₇₆₋₁₁₅ (orange) interacting in the second binding pocket. On one side of the pocket Val98 of Rad4 interacts with Leu48, Ala50, Lys101 and Gln105 of Tfb1. On the other side of the pocket Thr99 of Rad4 interacts with Gln105, Ile108 and Lys112 of Tfb1.

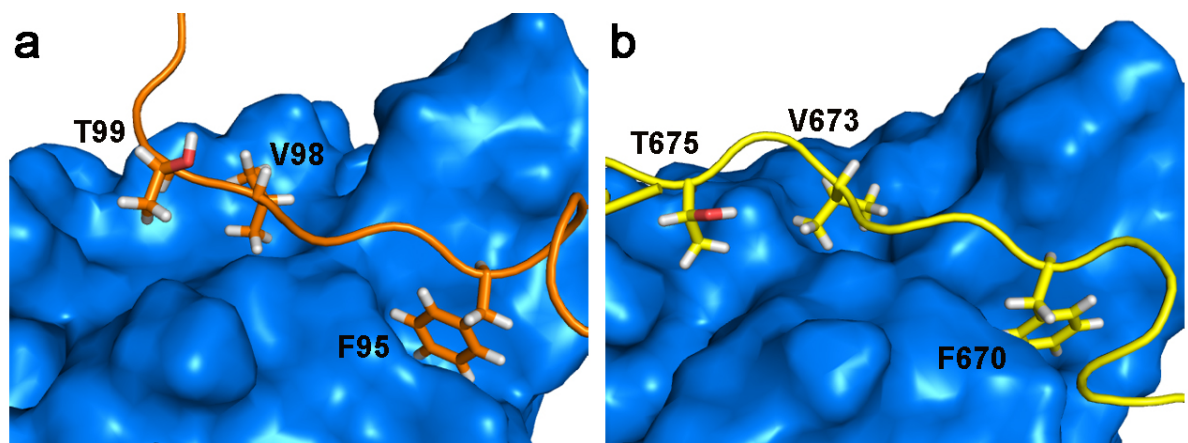


Figure 6. The structures of the Rad2₆₄₂₋₆₉₀-Tfb1PH and Rad4₇₆₋₁₁₅-Tfb1PH interfaces are very similar. The 3D structures of Tfb1PH are shown as molecular surfaces (blue) and Rad4₇₆₋₁₁₅ (a; orange), Rad2₆₄₂₋₆₉₀ (b; yellow) are shown as ribbons. Selected residues in the Tfb1PH-binding motif of Rad4₇₆₋₁₁₅ and Rad2₆₄₂₋₆₉₀ are also shown to demonstrate the similarity between their binding interfaces.

Table 1. Structural Statistics for Rad4 in Complex with Tfb1PH^a

Restrains used for the structure calculations	
Total number of NOE distances restraints	1565
Short-range (intraresidue)	420
Medium-range ($ i-j \leq 4$)	622
Long-range	495
Intermolecular	28
Hydrogen bond	36
Number of dihedral angle restraints (φ, ψ)	148
Structural statistics	
Rms deviations from idealized geometry	
Bonds (Å)	0.00210 ± 0.00007
Angles (deg)	0.376 ± 0.004
Impropers (deg)	0.209 ± 0.008
Rms deviations from distance restraints (Å)	0.0179 ± 0.0006
Rms deviations from dihedral restraints (deg)	0.18 ± 0.03
Ramachandran statistics (%) ^b	
Residues in most favored regions	76.9
Residues in additional allowed regions	20.7
Residues in generously allowed regions	1.5
Residues in disallowed regions	0.9
Coordinate precision ^c	
Atomic pair wise rmsd (Å)	
Rad2 – Tfb1PH complex	
Backbone atoms (C', C ^α , N)	0.67 ± 0.14
All heavy atoms	1.31 ± 0.12
Tfb1PH alone	
Backbone atoms (C', C ^α , N)	0.60 ± 0.14
All heavy atoms	1.22 ± 0.10
Rad2 alone	
Backbone atoms (C', C ^α , N)	0.60 ± 0.14
All heavy atoms	1.80 ± 0.35

^a The 20 conformers with the lowest energies were selected for statistical analysis.

^b Based on PROCHECK-NMR analysis.

^c Only residues 5–63 and 86–112 of Tfb1PH and residues 95–99 of Rad4 were used for the rmsd calculations. Residues at the N terminus (1–4), at the C terminus (113–115), and in the flexible loop (64–85) of Tfb1PH, as well as residues at the N terminus (76–94) and at the C terminus (100–115) of Rad4 were not included in the calculation.

REFERENCES

1. Friedberg, E.C. *et al.* *DNA Repair and Mutagenesis*, (ASM Press, Washington, DC, 2005).
2. Lehmann, A.R. DNA repair-deficient diseases, xeroderma pigmentosum, Cockayne syndrome and trichothiodystrophy. *Biochimie* **85**, 1101-1111 (2003).
3. Cartault, F. *et al.* A new XPC gene splicing mutation has lead to the highest worldwide prevalence of xeroderma pigmentosum in black Mahori patients. *DNA Repair* **10**, 577-585 (2011).
4. Hoeijmakers, J.H. Genome maintenance mechanisms for preventing cancer. *Nature* **411**, 366-374 (2001).
5. Sugasawa, K. *et al.* Xeroderma pigmentosum group C protein complex is the initiator of global genome nucleotide excision repair. *Mol Cell* **2**, 223-232 (1998).
6. Tantin, D., Kansal, A. & Carey, M. Recruitment of the putative transcription-repair coupling factor CSB/ERCC6 to RNA polymerase II elongation complexes. *Mol. Cell. Biol.* **17**, 6803-6814 (1997).
7. Sarker, A.H. *et al.* Recognition of RNA polymerase II and transcription bubbles by XPG, CSB, and TFIIH: insights for transcription-coupled repair and Cockayne syndrome. *Mol. Cell* **20**, 187-198 (2005).
8. Bardwell, A.J. *et al.* Yeast nucleotide excision repair proteins Rad2 and Rad4 interact with RNA polymerase II basal transcription factor b (TFIIH). *Mol. Cell. Biol.* **14**, 3569-3576 (1994).
9. Araujo, S.J., Nigg, E.A. & Wood, R.D. Strong functional interactions of TFIIH with XPC and XPG in human DNA nucleotide excision repair, without a preassembled repairsome. *Mol. Cell. Biol.* **21**, 2281-2291 (2001).
10. Yokoi, M. *et al.* The xeroderma pigmentosum group C protein complex XPC-HR23B plays an important role in the recruitment of transcription factor IIIH to damaged DNA. *J. Biol. Chem.* **275**, 9870-9875 (2000).
11. Min, J.-H. & Pavletich, N.P. Recognition of DNA damage by the Rad4 nucleotide excision repair protein. *Nature* **449**, 570-575 (2007).

12. Miron, S., Duchambon, P., Blouquit, Y., Durand, D. & Craescu, C.T. The carboxy-terminal domain of xeroderma pigmentosum complementation group C protein, involved in TFIIH and centrin binding, is highly disordered. *Biochemistry* **47**, 1403-1413 (2008).
13. Chen, L. & Madura, K. Centrin/Cdc31 is a novel regulator of protein degradation. *Mol. Cell. Biol.* **28**, 1829-1840 (2008).
14. den Dulk, B., van Eijk, P., de Ruijter, M., Brandsma, J.A. & Brouwer, J. The NER protein Rad33 shows functional homology to human Centrin2 and is involved in modification of Rad4. *DNA Repair* **7**, 858-868 (2008).
15. Popescu, A. *et al.* Xeroderma pigmentosum group C protein possesses a high affinity binding site to human centrin 2 and calmodulin. *J. Biol. Chem.* **278**, 40252-40261 (2003).
16. Uchida, A. *et al.* The carboxy-terminal domain of the XPC protein plays a crucial role in nucleotide excision repair through interactions with transcription factor IIIH. *DNA Repair* **1**, 449-461 (2002).
17. Coin, F., Oksenysh, V. & Egly, J.-M. Distinct Roles for the XPB/p52 and XPD/p44 Subcomplexes of TFIIH in Damaged DNA Opening during Nucleotide Excision Repair. *Mol. Cell* **26**, 245-256 (2007).
18. Schärer, O.D. The molecular basis for different disease states caused by mutations in TFIIH and XPG. *DNA Repair* **7**, 339-344 (2008).
19. Riedl, T., Hanaoka, F. & Egly, J.M. The comings and goings of nucleotide excision repair factors on damaged DNA. *EMBO J.* **22**, 5293-5303 (2003).
20. Bernardes de Jesus, B.M., Bjoras, M., Coin, F. & Egly, J.M. Dissection of the molecular defects caused by pathogenic mutations in the DNA repair factor XPC. *Mol. Cell. Biol.* **28**, 7225-7235 (2008).
21. Hohl, M. *et al.* Domain swapping between FEN-1 and XPG defines regions in XPG that mediate nucleotide excision repair activity and substrate specificity. *Nucleic Acids Res.* **35**, 3053-3063 (2007).
22. Iyer, N., Reagan, M.S., Wu, K.J., Canagarajah, B. & Friedberg, E.C. 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**, 2157-2167 (1996).

23. Lafrance-Vanasse, J. *et al.* Structural and functional characterization of interactions involving the Tfb1 subunit of TFIIH and the NER factor Rad2. *previous article* (2011).
24. Tremblay, M., Teng, Y., Paquette, M., Waters, R. & Conconi, A. Complementary roles of yeast Rad4p and Rad34p in nucleotide excision repair of active and inactive rRNA gene chromatin. *Mol. Cell. Biol.* **28**, 7504-7513 (2008).
25. den Dulk, B., Brandsma, J.A. & Brouwer, J. The Rad4 homologue YDR314C is essential for strand-specific repair of RNA polymerase I-transcribed rDNA in *Saccharomyces cerevisiae*. *Mol. Microbiol.* **56**, 1518-1526 (2005).
26. Iben, S. *et al.* TFIIH plays an essential role in RNA polymerase I transcription. *Cell* **109**, 297-306 (2002).
27. den Dulk, B., Sun, S.M., de Ruijter, M., Brandsma, J.A. & Brouwer, J. Rad33, a new factor involved in nucleotide excision repair in *Saccharomyces cerevisiae*. *DNA Repair* **5**, 683-692 (2006).
28. Di Lello, P. *et al.* NMR structure of the amino-terminal domain from the Tfb1 subunit of TFIIH and characterization of its phosphoinositide and VP16 binding sites. *Biochemistry* **44**, 7678-7686 (2005).
29. Stauffer, M.E. & Chazin, W.J. Structural mechanisms of DNA replication, repair, and recombination. *J. Biol. Chem.* **279**, 30915-30918 (2004).
30. Wilson, S.H. & Kunkel, T.A. Passing the baton in base excision repair. *Nature structural biology* **7**, 176-178 (2000).
31. Dunand-Sauthier, I. *et al.* The spacer region of XPG mediates recruitment to nucleotide excision repair complexes and determines substrate specificity. *J. Biol. Chem.* **280**, 7030-7037 (2005).
32. Thorel, F. *et al.* Definition of a short region of XPG necessary for TFIIH interaction and stable recruitment to sites of UV damage. *Mol. Cell. Biol.* **24**, 10670-10680 (2004).
33. Okuda, M. *et al.* Structural insight into the TFIIIE-TFIIH interaction: TFIIIE and p53 share the binding region on TFIIH. *EMBO J.* **27**, 1161-1171 (2008).
34. Di Lello, P. *et al.* p53 and TFIIIEalpha share a common binding site on the Tfb1/p62 subunit of TFIIH. *Proc. Natl. Acad. Sci. USA* **105**, 106-111 (2008).

35. Tirode, F., Busso, D., Coin, F. & Egly, J.M. Reconstitution of the transcription factor TFIIH: assignment of functions for the three enzymatic subunits, XPB, XPD and cdk7. *Mol. Cell* **3**, 87-95 (1999).
36. Lu, H., Zawel, L., Fisher, L., Egly, J.M. & Reinberg, D. Human general transcription factor IIH phosphorylates the C-terminal domain of RNA polymerase II. *Nature* **358**, 641-645 (1992).
37. Kaiser, C., Michaelis, S. & Mitchell, A. *Methods in yeast genetics*, 202 (Cold Spring Harbor Laborator Press, Cold Spring Harbor, NY, 1994).
38. Houtman, J.C. *et al.* Binding specificity of multiprotein signaling complexes is determined by both cooperative interactions and affinity preferences. *Biochemistry* **43**, 4170-4178 (2004).
39. Nguyen, B.D., Di Lello, P., Legault, P. & Omichinski, J.G. ¹H, ¹⁵N, and ¹³C resonance assignment of the amino-terminal domain of the Tfb1 subunit of yeast TFIIH. *J. Biomol. NMR* **31**, 173-174 (2005).
40. Kay, L.E., Xu, G.Y. & Yamazaki, T. Enhanced-sensitivity triple-resonance spectroscopy with minimal H₂O saturation. *J. Magn. Reson. A* **109**, 129-133 (1994).
41. Wittekind, M. & Mueller, L. HNCACB, a high-sensitivity 3D NMR experiment to correlate amide-proton and nitrogen resonances with the alpha- and beta-carbon resonances in proteins. *J. Magn. Reson. B* **101**, 201-205 (1993).
42. Grzesiek, S. & Bax, A. Correlating backbone amide and side chain resonances in larger proteins by multiple relayed triple resonance NMR. *J. Am. Chem. Soc.* **114**, 6291-6293 (1992).
43. Logan, T.M., Olejniczak, E.T., Xu, R.X. & Fesik, S.W. Side chain and backbone assignments in isotopically labeled proteins from two heteronuclear triple resonance experiments. *FEBS Lett.* **314**, 413-418 (1992).
44. Ikura, M., Kay, L.E. & Bax, A. Improved three-dimensional ¹H-¹³C-¹H correlation spectroscopy of a ¹³C- labeled protein using constant-time evolution. *J. Biomol. NMR* **1**, 299-304 (1991).
45. Yamazaki, T., Forman-Kay, J.D. & Kay, L.E. Two-dimensional NMR experiments for correlating carbon-13 beta and proton delta/epsilon chemical shifts of aromatic residues in ¹³C-labeled proteins via scalar couplings. *J. Am. Chem. Soc.* **115**, 11054-11055 (1993).

46. Zhang, O., Kay, L.E., Olivier, J.P. & Forman-Kay, J.D. Backbone ^1H and ^{15}N resonance assignments of the N-terminal SH3 domain of drk in folded and unfolded states using enhanced-sensitivity pulsed field gradient NMR techniques. *J. Biomol. NMR* **4**, 845-858 (1994).
47. Pascal, S.M., Muhandiram, D.R., Yamazaki, T., Forman-Kay, J.D. & Kay, L.E. Simultaneous acquisition of ^{15}N - and ^{13}C -edited NOE spectra of proteins dissolved in H_2O . *J. Magn. Reson.* **103**, 197-201 (1994).
48. Delaglio, F. *et al.* NMRPipe: A multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **6**, 277-293 (1995).
49. Johnson, B.A. Using NMRView to visualize and analyze the NMR spectra of macromolecules. *Methods Mol. Biol.* **278**, 313-352 (2004).
50. Vranken, W.F. *et al.* The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins* **59**, 687-696 (2005).
51. Shen, Y., Delaglio, F., Cornilescu, G. & Bax, A. TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *J. Biomol. NMR* **44**, 213-223 (2009).
52. Brunger, A.T. *et al.* Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr. D Biol. Crystallogr.* **54**, 905-921 (1998).
53. Laskowski, R.A., Rullmann, J.A., MacArthur, M.W., Kaptein, R. & Thornton, J.M. AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J. Biomol. NMR* **8**, 477-486 (1996).
54. Koradi, R., Billeter, M. & Wuthrich, K. MOLMOL: a program for display and analysis of macromolecular structures. *J. Mol. Graph.* **14**, 51-55 (1996).
55. DeLano, W.L. The PyMOL Molecular Graphics System. (Delano Scientific, Palo Alto, CA, 2002).

SUPPLEMENTARY METHODS

Plasmid preparations

The pRS316RAD4cmyc plasmid (for the *RAD4* strain) was generated by amplification of the *RAD4* Open Reading Frame (ORF) complemented by 190bp upstream and 282bp downstream of the coding sequence on genomic DNA with the primers pRS-RAD4-NotI-F and pRS-RAD4-EcoRI-R. The PCR product was inserted NotI/EcoRI into pRS316 (see **Supplementary Table 2** for primers used in the study). Finally, the cmyc tag was inserted in between NdeI/EcoRI right before the STOP codon of *RAD4* by overlapped PCR using pRS-RAD4-2995-NdeI-F, pRS-RAD4-cmyc-R, pRS-RAD4-cmyc-F and pRS-RAD4-EcoRI-R. The previous intermediate plasmid was used as template. The pRS316RAD4(F95P/V98P)cmyc plasmid (*rad4-PP*) was obtained by the QuikChange (QuikChange Site-Directed Mutagenesis Kit from Stratagene) mutagenesis procedure using the primers RAD4(F95P+V98P)-F and RAD4(F95P+V98P)-R with the pRS316RAD4cmyc plasmid as template. The pRS316RAD4(W649A/L652A/L656A)cmyc (*rad4-AAA*) and pRS316RAD4(F95P/V98P/W649A/L652A/L656A)cmyc (*rad4-PPAAA*) plasmids were both obtained by overlapped PCR using the primers pRS-RAD4-1742-BglII-F(L), RAD4I(W649AL652AL656A)-R, RAD4(W649AL652AL656A)-F and pRS-RAD4-EcoRI-R with pRS316RAD4cmyc plasmid as template. The PCR product was inserted BglII/EcoRI into either pRS316RAD4cmyc or pRS316RAD4(F95P/V98P)cmyc respectively.

NMR samples

For the NMR chemical shift titration studies of Rad4 and Rad34 segments with Tfb1PH, either unlabeled Rad4₇₆₋₁₁₅ or unlabeled Rad34₄₁₋₆₃ was added to a sample containing 0.5 mM of ¹⁵N-labeled Tfb1PH in NMR buffer to a final ratio of 1:1. These studies were performed in 20 mM sodium phosphate buffer pH 6.5, 1 mM EDTA, 1 mM DTT (NMR buffer) with 90%H₂O/10%D₂O.

For the NMR competition experiment of Rad4₇₆₋₁₁₅ and Rad2₆₄₂₋₆₉₀, a sample containing 0.5 mM ¹⁵N-labeled Rad4₇₆₋₁₁₅ in NMR buffer was used. To this sample,

unlabeled Tfb1PH was added to a final concentration of 0.4 mM. In a second addition, unlabeled Rad2₃₅₉₋₃₈₃ was added to a final concentration of 1.5 mM.

The samples were recorded at 25°C on a Varian Inova 600 MHz spectrometer.

Protein expression levels in yeast

To determine the expression level of the Rad4 protein and its mutants, normal protein levels were too low to detect when expressed from their endogenous promoters. Therefore, Rad4 and all mutants were cloned into the overexpression vector pVT-101u, which is under the control of the ADH promoter. The genes were inserted XhoI/BamHI and a myc-tag was fused on their C-terminus. Yeast cells were grown in 15 mL of minimal media to OD₅₉₅ = 0.8-1.0 and centrifuged. The pellets were resuspended and total protein extractions were carried out as previously described (Elagoz et al.,1999). The protein extracts (20 µg) were then migrated on a 10% (wt/vol) SDS-PAGE gel and transferred onto a nitrocellulose membrane according to the manufacturer's instructions. The Rad4 and mutant proteins levels were determined with the anti-cmyc mouse mAb 9E10 at a dilution of 1/500. The Peroxidase AffiniPure Goat Anti-Mouse IgG + IgM (H+L) (Jackson ImmunoResearch Laboratories) served as the secondary antibody, used at a dilution of 1/1,000. Loading was verified with the anti-actin mouse monoclonal antibody (Abcam) at a dilution of 1/1,000.

Reference :

Elagoz A., Callejo M., Armstrong J., Rokeach L. A. (1999) Although calnexin is essential in *S. pombe*, its highly conserved central domain is dispensable for viability. *J. Cell Sci.* 112, 4449–4460.

SUPPLEMENTARY TABLES

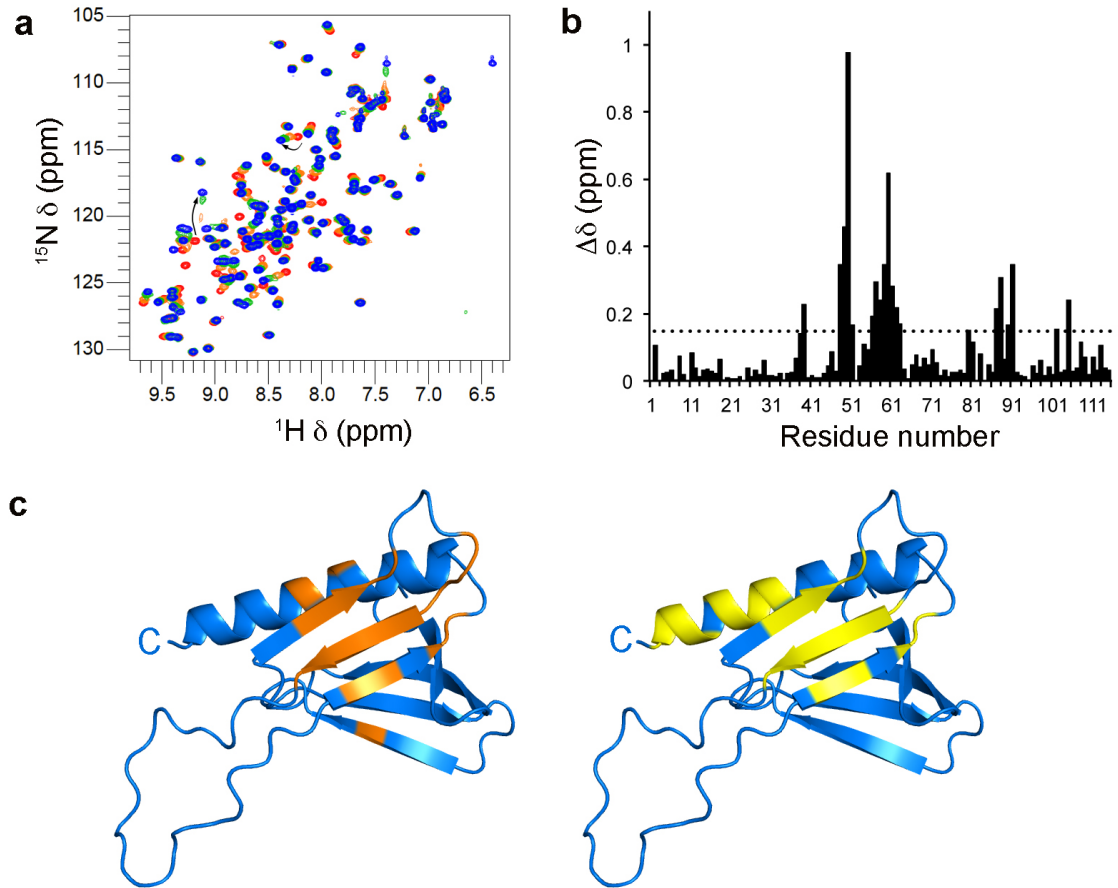
Supplementary Table 1: Yeast strains used in this study

Name	Genotype	Source
<i>rad4</i> (6158)	Mata <i>his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 rad4Δ::KanMX</i>	Open Biosystems
<i>rad4</i>	<i>rad4</i> (6158) + pRS316	This Lab
<i>RAD4</i>	<i>rad4</i> (6158) + pRS316RAD4cmyc	This Lab
<i>rad4-PP</i>	<i>rad4</i> (6158) + pRS316RAD4(F95P/V98P)cmyc	This Lab
<i>rad4-AAA</i>	<i>rad4</i> (6158) + pRS316RAD4(W649A/L652A/L656A)cmyc	This Lab
<i>rad4-PPAAA</i>	<i>rad4</i> (6158) + pRS316RAD4(F95P/V98P/W649A/L652A/L656A)cmyc	This Lab

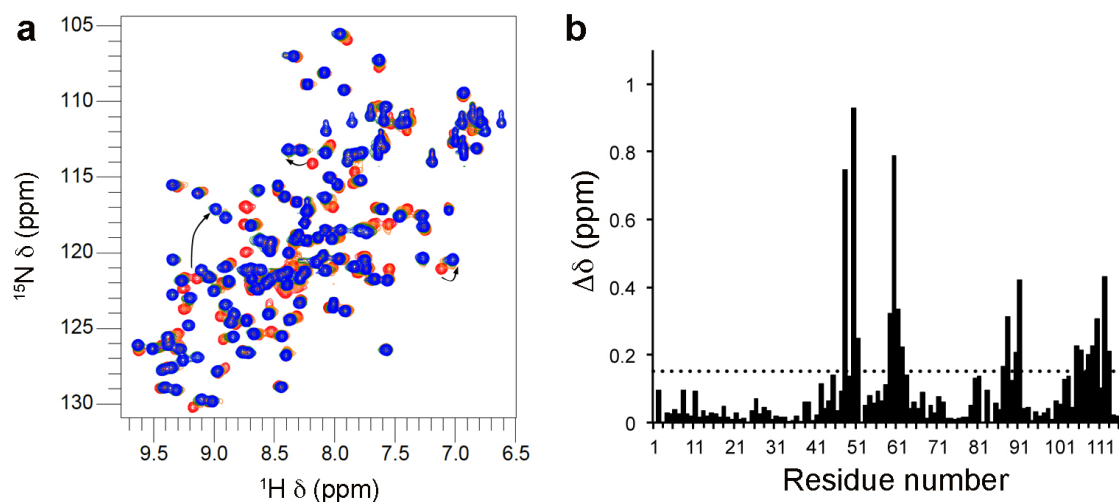
Supplementary Table 2: Primers used in this study

Name	Sequence
pRS-RAD4-NotI-F	5' aag gaa aaa agc ggc cgc atg aaa tat gta tac ata tct tga tta act atg 3'
pRS-RAD4-EcoRI-R	5' cga att caa cta aaa tat ttc act ttg 3'
pRS-RAD4-2995-NdeI-F	5' gga att cca tat ggt aaa att gcc gag g 3'
pRS-RAD4-cmyc-R	5' cag cct cat ttc aca ggt cct cct ccg aga tca get tct get cgt ctg att cct ctg ac 3'
pRS-RAD4-cmyc-F	5' gtc aga gga atc aga cga gca gaa gct gat ctc gga gga cct gtg aaa tga ggc tg 3'
RAD4(F95P+V98P)-F	5' gat tcg gag gaa cct gaa gac ccg aca gat gg 3'
RAD4(F95P+V98P)-R	5' cca tct gtc ggg tct tca ggt tcc tcc gaa tc 3'
pRS-RAD4-1742-BglII-F(L)	5' gaa gat ctg act ttt taa ggg ccg tca g 3'
RAD4I(W649AL652AL656A)-R	5' ctc tcc aaa gca cca agc 3'
RAD4(W649AL652AL656A)-F	5' ctt ggt gct ttg gag agt gcg aat act gca ctg cta aaa gcg cgt att cgt agc aag ctg 3'
pRS-RAD4-EcoRI-R	5' cga att caa cta aaa tat ttc act ttg 3'

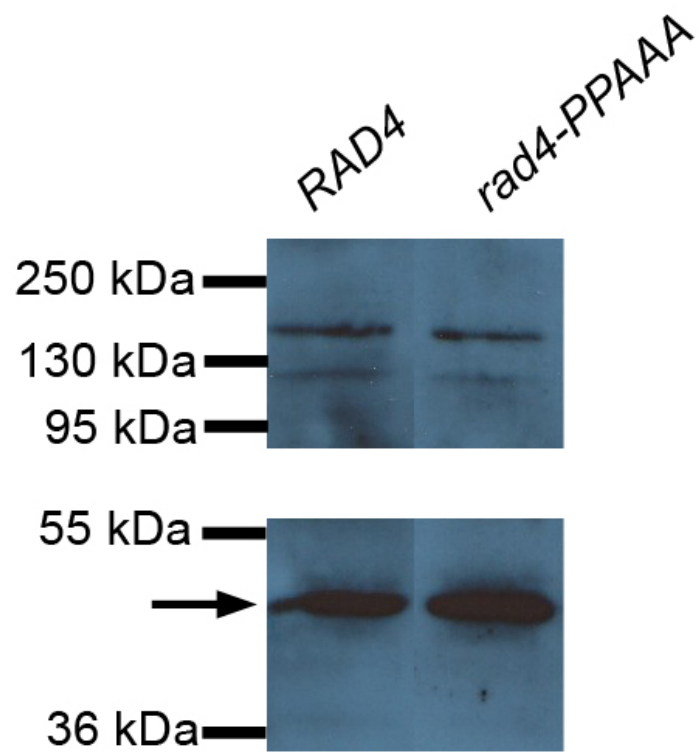
SUPPLEMENTARY FIGURES



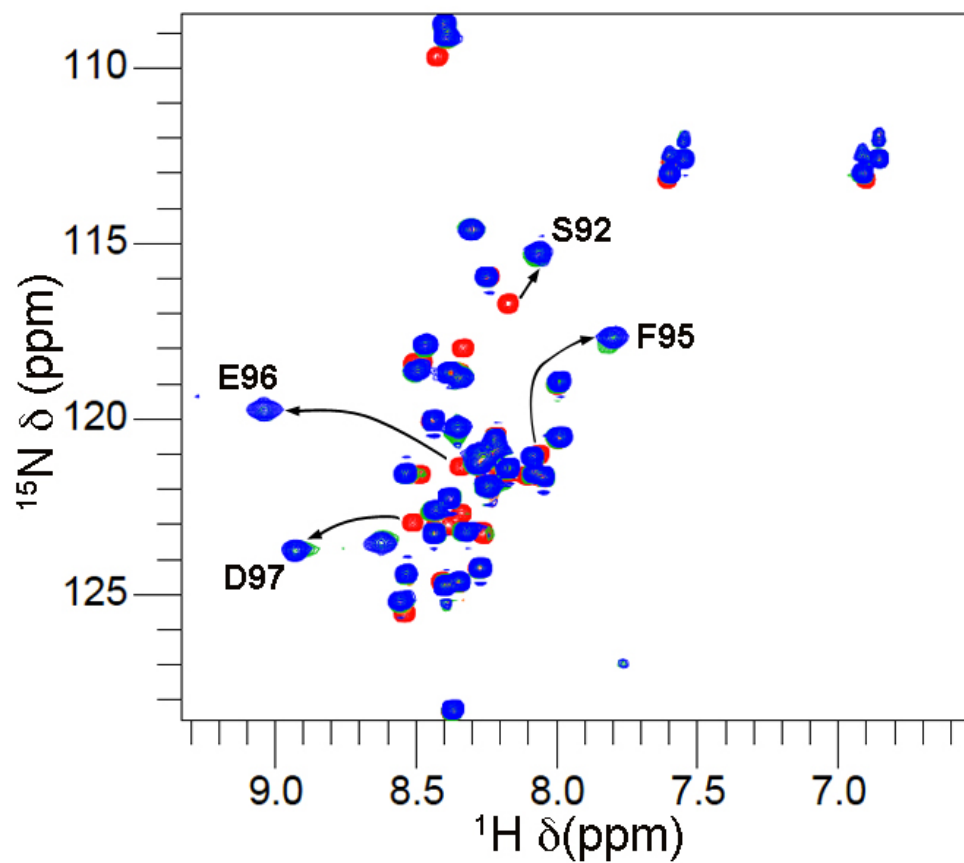
Supplementary Figure 1. Rad4₇₆₋₁₁₅ binds to a similar site of Tfb1PH as Rad2₆₄₂₋₆₉₀. (a) Overlay of the ^1H - ^{15}N HSQC spectra for ^{15}N -labeled Tfb1PH in its free form (red), in the presence of 0.25 equivalents (orange), 0.5 equivalent (green) and 1 equivalent (blue) of unlabeled Rad4₇₆₋₁₁₅. Arrows highlight two signals that undergo significant changes in ^1H and ^{15}N chemical shifts upon formation of the Rad4₇₆₋₁₁₅-Tfb1PH complex. (b) Histogram of the variation in chemical shift [$\Delta\delta_{(\text{ppm})}$; $\Delta\delta = [(0.17\Delta\text{N}_{\text{H}})^2 + (\Delta\text{H}_{\text{N}})^2]^{1/2}$] observed in the ^1H - ^{15}N HSQC spectrum of the ^{15}N -labeled Tfb1PH upon formation of the Rad4₇₆₋₁₁₅-Tfb1PH complex. (c) Ribbon models of the three dimensional structure of Tfb1PH. The amino acids showing a significant chemical shift change ($\Delta\delta_{(\text{ppm})} > 0.15$) upon formation of a complex with Rad4₇₆₋₁₁₅ are shown in orange (left) and upon formation of a complex with Rad2₆₄₂₋₆₉₀ are shown in yellow (right).



Supplementary Figure 2. Rad34₄₁₋₆₃ binds to a similar site of Tfb1PH as Rad4₇₆₋₁₁₅. **(a)** Overlay of the two dimensional ¹H-¹⁵N HSQC spectra for ¹⁵N-labeled Tfb1PH in its free form (red), in the presence of 0.5 equivalents (orange), 0.75 equivalent (green) and 1 equivalent (blue) of unlabeled Rad34₄₁₋₆₃. Arrows highlight several signals that undergo significant changes in ¹H and ¹⁵N chemical shifts upon formation of the Rad34₄₁₋₆₃-Tfb1PH complex. **(b)** Histogram of the variation in chemical shift [$\Delta\delta_{(\text{ppm})}$; $\Delta\delta = [(0.17\Delta N_{\text{H}})^2 + (\Delta H_{\text{N}})^2]^{1/2}$] observed in the ¹H-¹⁵N HSQC spectrum of the ¹⁵N-labeled Tfb1PH upon formation of the Rad34₄₁₋₆₃-Tfb1PH complex.



Supplementary Figure 3. Rad4 protein expression levels in yeast. Level of expression of the Rad4 and Rad4-PPAAA proteins in *RAD4* and *rad4-PPAAA* yeast, respectively. The protein levels were monitored by western blot with an antibody against the myc tag. Actin levels are shown as a loading control (arrow).



Supplementary Figure 4. Tfb1PH binds to Rad4₇₆₋₁₁₅. Overlay of the ^1H - ^{15}N HSQC spectra for ^{15}N -labeled Rad4₇₆₋₁₁₅ in its free form (red), in the presence of 0.5 equivalents (orange), 0.75 equivalent (green) and 1 equivalent (blue) of unlabeled Tfb1PH. Arrows highlight several signals that undergo significant changes in ^1H and ^{15}N chemical shifts upon formation of the Rad4₇₆₋₁₁₅-Tfb1PH complex.

4.0 Discussion et conclusion

L'hypothèse de départ de cette thèse est que certaines protéines de la réparation de l'ADN possèdent des segments d'interaction semblables aux domaines de transactivation des activateurs de la transcription leur permettant de recruter leurs partenaires à l'ADN endommagé. Ainsi, nous avons comme objectifs de caractériser les interactions entre Tfb1 et Rad2 ainsi qu'entre Tfb1 et Rad4 pour ensuite établir un parallèle entre ces interactions et celles qui ont lieu durant la transcription. Nous devons ainsi dégager par quel langage moléculaire tous ces processus fonctionnent.

Dans un premier article, nous avons d'abord confirmé que la délétion du domaine PH de Tfb1 augmente la sensibilité des levures aux rayons UV. Nous avons ensuite précisé la location du domaine d'interaction de Rad2 avec Tfb1 et trouvé un autre domaine similaire dans la région espaceur de Rad2. À notre surprise, nous avons découvert qu'il n'y a pas qu'un seul segment d'interaction, mais que de multiples segments qui créent chacun des interactions faibles. En combinant mutations et délétions de ces segments, la sensibilité des levures aux UV est accrue. Finalement, nous avons résolu la structure d'un complexe entre Rad2 et Tfb1PH et montré que le mode d'interaction avec Tfb1PH diffère de celui observé entre Tfb1PH et les domaines de transactivation des activateurs de la transcription tels que p53 et VP16. Malgré cela, l'interaction entre Rad2 et Tfb1PH utilise un mécanisme similaire à l'interaction entre TFIIIE α et TFIIH. Ceci crée un pont important entre la réparation et la transcription basale de l'ADN.

D'autres exemples similaires devaient être trouvés pour appuyer ce parallèle. Dans un second article, nous avons ainsi identifié un module d'interaction entre Rad4 en nous basant sur la séquence de motif trouvé chez Rad2. Nous avons aussi prouvé que la sensibilité des levures aux UV était augmentée lorsque la mutation de ce segment était combinée à une autre portion de Rad4, en carboxy-terminal de la protéine. Nous avons finalement montré que la structure du complexe Rad4-Tfb1PH est très semblable au complexe Rad2-Tfb1PH. Par ce second exemple, nous avons une fois de plus prouvé qu'un pont dans le mécanisme d'interaction existe entre la transcription et la réparation de l'ADN. Qui plus est, Rad2 et Rad4 utilisent tous deux une séquence similaire pour lier le même domaine de Tfb1, en plus d'interagir avec TFIIH par de multiples interactions. Ces similarités font la lumière sur le mécanisme de déplacement de Rad4 par Rad2.

Interactions multiples

Dans notre étude, plusieurs segments de Rad4 et de Rad2 ont dû être supprimés avant d'observer une augmentation de la sensibilité des levures aux UV. Dans les deux cas, ceci suggère que l'interaction avec TFIIH se fait grâce à de multiples interactions.

Peu d'études se sont penchées sur le rôle du segment amino-terminal de XPC/Rad4. Or, nous avons prouvé l'importance de ce segment pour l'interaction avec le domaine PH de Tfb1. Nous avons également confirmé la nécessité de la portion carboxy-terminale de Rad4 pour l'interaction avec TFIIH, tout comme ceci avait été prouvé chez XPC (160, 204). Cette portion de la protéine contient un segment conservé servant à la liaison de Rad33 chez la levure (57) et de la centrine 2 chez l'humain (61). Ce segment forme une hélice- α au contact de la centrine 2, tel que démontré par RMN et cristallographie (60-62). Ainsi, il est possible d'émettre l'hypothèse que la centrine 2 (ou Rad33) agit comme un adaptateur pour une des interactions entre TFIIH et XPC/Rad4. En conclusion, il y a minimalement deux segments de Rad4 liant TFIIH.

Rad2 implique également un long segment désordonné, le segment espaceur, dans ses interactions. Les deux segments que nous avons trouvés peuvent interagir et même faire compétition avec Tfb1PH. Chez l'humain, des études ont montré qu'un troisième segment (227-238) est nécessaire à la liaison à TFIIH, mais on ne sait pas avec quelle sous-unité de TFIIH il interagit (131). Par analogie avec XPC/Rad4 il est possible d'émettre l'hypothèse que ce segment interagit avec centrine 2/Rad33. Celle-ci pourrait être un adaptateur non seulement entre XPC/Rad4 et TFIIH, mais aussi entre XPG/Rad2 et TFIIH. La combinaison de résidus hydrophobes et hydrophiles de ce segment de XPG rappelle le segment d'interaction de XPC avec centrine 2.

Cette similarité des multiples interactions faibles entre Rad2 et Rad4 n'est pas le fruit du hasard. Nous pensons qu'elles font partie du mécanisme moléculaire de la NER. Ainsi, ces découvertes permettent la formulation d'un modèle de transition de XPC/Rad4 pour XPG/Rad2 au site de dommage de l'ADN.

Un modèle de transition pour la NER

Après la détection et la liaison à l'ADN endommagé, XPC/Rad4 recrute TFIIH et débute ainsi la formation du complexe de pré-incision. Lorsque XPG/Rad2 arrive au complexe de pré-incision de l'ADN, il déplace XPC/Rad4 (26). Or à partir de nos résultats, il est possible de proposer un modèle quant au mécanisme moléculaire par lequel ce transfert se produit. Pour expliquer ce modèle, utilisons un objet de la vie de tous les jours, le velcro. Les deux parties d'un velcro se lient ensemble d'une façon très solide. Toutefois, en s'y prenant de la bonne façon, il est facile de défaire ces deux parties. On allie donc solidité et flexibilité grâce à des dizaines de petits crochets qui individuellement s'accrochent faiblement, mais qui sont collectivement très solides. Défaire les crochets les uns après les autres et non tous en même temps est la meilleure façon pour séparer les deux parties.

Si l'interaction entre Rad4 et TFIIH ne reposait que sur une seule interface créant des liaisons très fortes, il serait très difficile de séparer les deux protéines. Évidemment, toute liaison est un équilibre entre la forme liée et libre et les interactions peuvent faire compétition entre elles. Or, plus une liaison est forte, plus la forme liée est prédominante. Par l'analogie du velcro, nous apportons une solution élégante à ce problème. De multiples interactions sont responsables de l'ancrage de TFIIH à XPC/Rad4 au site de dommage à l'ADN. Lorsque TFIIH commence à ouvrir la double hélice d'ADN, ceci crée un substrat privilégié pour XPG/Rad2. TFIIH recrute donc XPG/Rad2 au site de la lésion via les mêmes domaines d'interaction que la liaison avec XPC/Rad4. Peu à peu, chacune des interactions faibles formées entre TFIIH et XPC/Rad4 se brisent, pour laisser place à d'autres interactions faibles avec XPG/Rad2. Ce modèle correspond bien à nos données. Lorsque le domaine d'interaction à Tfb1PH de Rad4 ou de Rad2 est retiré, il n'y a que peu d'effet sur la viabilité des levures : les autres interactions sont capables de prendre le relai et assurer la liaison des molécules. Toutefois, si on combine la suppression de plusieurs domaines, il est possible d'observer un effet, puisqu'on affaiblit l'interaction globale.

Ce mécanisme moléculaire de transition offre plusieurs avantages à la cellule. Premièrement, XPC/Rad4 est positionné en 3' de la lésion, là où XPG/Rad2 clive l'ADN. Le déplacement de XPC/Rad4 permet de positionner XPG/Rad2 au bon endroit.

Deuxièmement, si XPG/Rad2 devait faire compétition avec une seule interaction forte entre Rad4 et TFIIH, il pourrait y avoir des risques d'effondrement du complexe de pré-incision (dissociation des protéines du complexe). En effet, XPC/Rad4 et XPG/Rad2 stabilisent tour à tour TFIIH au complexe. Si une seule interaction forte est en jeu, il existe donc un état de transition où ni XPC/Rad4 ni XPG/Rad2 n'interagit avec TFIIH. Or, cet état, où TFIIH est seul sur l'ADN, est fragile et peut mener à la dissociation de TFIIH de l'ADN, compromettant ainsi la formation du complexe de pré-incision. La transition doit donc se faire graduellement. Un dernier avantage d'établir de multiples interactions entre deux protéines est que de multiples sites peuvent subir des modifications post-traductionnelles. Ainsi, certains sites peuvent être modifiés alors que d'autres peuvent rester intacts. Ceci permet à la cellule de pouvoir modifier graduellement les interactions et ainsi moduler la réponse cellulaire, contrairement à d'autres processus qui n'ont que des réponses binaires (mode actif ou mode inactif).

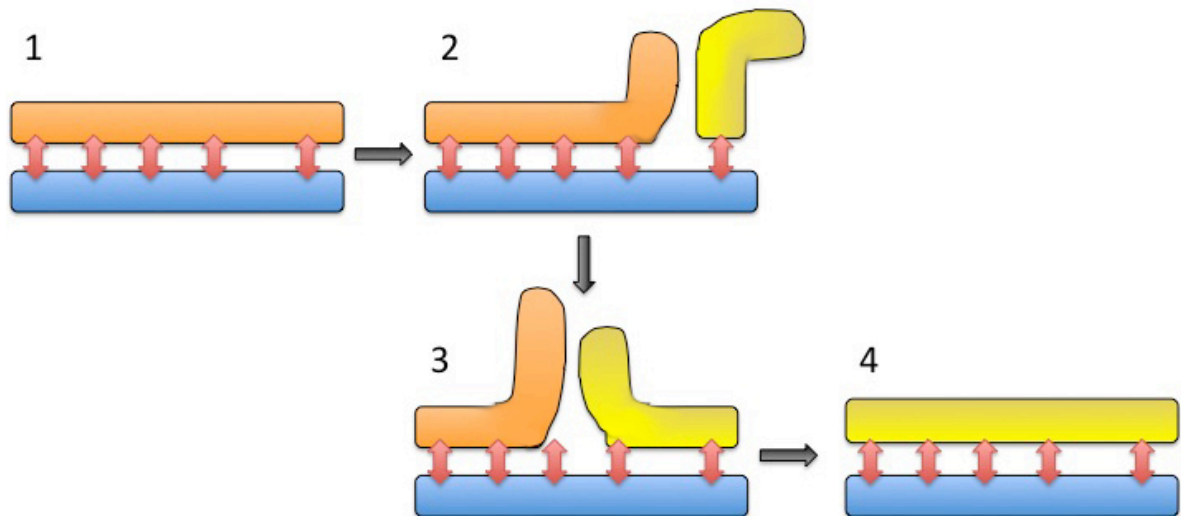


Figure 4-1. Modèle « velcro » de transition pour la NER. 1. Les multiples interactions faibles (flèches rouges) entre XPC/Rad4 (orange) et TFIIH (bleu) lient les deux facteurs très fortement. 2. Lorsque XPG/Rad2 est recruté au site de lésion, il déplace une interaction faible à la fois. 3. Peu à peu, les interactions sont déplacées. 4. Finalement XPG/Rad2 a déplacé XPC/Rad4 de TFIIH.

Régulation des interactions

La présente étude ne permet pas de comprendre comment certains signaux moléculaires peuvent agir sur la régulation des interactions étudiées. Une telle régulation est nécessaire, sinon les protéines XPC/Rad4, XPG/Rad2 et TFIIH seraient toujours liées les unes aux autres. Un bon exemple de régulation d'une interaction est la phosphorylation du domaine de transactivation de p53 qui augmente son affinité pour p62/Tfb1 (198). Nous avons donc tenté d'imiter la phosphorylation de plusieurs résidus sérine et thréonine de Rad2 et Rad4 en les mutant en acide aspartique. Aucune amélioration ou détérioration notable de l'interaction avec Tfb1PH n'a été mesurée *in vitro* (données non présentées). Comment la régulation de ces interactions pourrait-elle fonctionner?

Il est possible d'émettre plusieurs hypothèses. D'autres interactions peuvent être responsables de la régulation via des modifications post-traductionnelles. Il se peut aussi que des changements allostériques suite à la liaison à l'ADN et au complexe de pré-incision soient des signaux suffisants pour induire de nouvelles interactions. D'autres études démontrent que ce sont surtout les niveaux de ces protéines qui régulent la NER. Ainsi, au lieu d'augmenter l'affinité des interactions, on augmente plutôt la concentration des molécules, ce qui augmente la quantité absolue de protéines pouvant interagir.

La régulation des niveaux de XPC/Rad4 et XPG/Rad2 est importante pour une autre raison. Si ces protéines sont trop abondantes, il pourrait y avoir induction de réparations non-spécifiques et ainsi création de certaines erreurs. Après irradiation aux UV, certaines protéines de la NER sont exprimées plus abondamment. Il a été démontré que le gène *RAD2* est stimulé suite à une exposition aux rayons UV (221, 222), mais les résultats sont mitigés pour *XPG* (223). D'autre part, il a été prouvé clairement que les niveaux de XPC sont plus élevés après une telle exposition (224). Le gène *XPC* est même sous le contrôle de p53 (225).

En conclusion, lorsque qu'il y a peu de dommages les niveaux des protéines de réparation sont faibles et TFIIH est surtout attiré à la transcription. Immédiatement après les dommages, il y a induction des gènes XPC/Rad4 et de XPG/Rad2, notamment via p53 qui stimule la transcription spécifique de gènes, notamment grâce à son interaction avec

TFIIH. Lorsque le niveau des protéines de réparation est suffisamment élevé, ceci stimule l'interaction avec TFIIH, qui est alors attiré à la réparation. Ceci pourrait expliquer pourquoi la transcription est réduite lors de la réparation de l'ADN (226, 227).

Comparaison des complexes formés entre p62/Tfb1 et ses partenaires

Structuralement, il semble y avoir deux catégories d'interactions avec p62/Tfb1. Ainsi, les TAD acides des activateurs p53 et VP16 forment des hélices amphiphiles (198, 199) tandis que le TAD acide de EKLF ainsi que les segments acides de TFIIE α , Rad2 et Rad4 forment plutôt des segments allongés et ordonnés sur la structure (182, 200). Dans tous les cas, les résidus acides des peptides peuvent former des interactions ioniques avec des résidus basiques de p62/Tfb1 alors que des résidus hydrophobes peuvent se loger dans des cavités peu profondes de p62/Tfb1. Même s'il semble se dessiner une différence entre la transcription activée et les autres processus (à l'exception de EKLF), il n'y a pas de séparation claire dans les processus et leur mode de liaison à p62/Tfb1.

Par contre, il est important de souligner l'existence d'un pont entre la transcription basale et la réparation de l'ADN. Les interactions Rad4-Tfb1, Rad2-Tfb1 et TFIIE α -p62 sont pratiquement identiques, utilisant un motif d'acides aminés conservé. Ce motif est composé d'un résidu aromatique suivi de deux résidus acides et d'un résidu hydrophobe (la valine est très fréquente dans cette dernière position). L'ensemble du motif est entouré de résidus acides. Ces résidus interagissent de la même façon sur la surface du domaine PH de p62/Tfb1. Finalement, la caractéristique commune à tous ces partenaires de p62/Tfb1 est qu'ils sont intrinsèquement désordonnés lorsqu'ils sont libres en solution. Explorons en quoi ceci peut être important biologiquement.

Désordre intrinsèque

Depuis ses débuts, la biologie structurale s'est surtout attardée aux protéines complètement ordonnées, ayant un arrangement tridimensionnel unique ou quelques conformations distinctes. Une caractéristique importante des segments de Rad2 et Rad4 étudiés est qu'il sont désordonnés sans leur partenaire d'interaction, tel que démontré par RMN. Ils adoptent une structure tridimensionnelle précise qu'à la rencontre de Tfb1PH. Ce type de peptide est désormais classé dans la catégorie des protéines intrinsèquement

désordonnés (*Intrinsically Disordered Protein*, IDP). Il y a une vingtaine d'années, les IDP n'étaient qu'une théorie parmi tant d'autres. Maintenant, grâce à plusieurs avancées techniques, il fait de moins en moins de doute que ces segments existent et qu'ils sont très importants.

Les IDP jouent des rôles diversifiés et ont été retrouvés dans plusieurs processus cellulaires, surtout dans la transcription et dans la signalisation cellulaire (228). Rad4 et Rad2 ne sont pas les premiers exemples d'IDP dans une voie de réparation de l'ADN. En effet, dans la réparation par excision des bases (BER), NEIL1 aurait une extrémité carboxy-terminale désordonnée (229, 230). Dans la NER, les deux extrémités de XPA seraient des IDP (231). Par contre, nos structures sont parmi les premières d'IDP repliées sur leur partenaire d'interaction dans les voies de réparation. Ainsi, la liaison est couplée au repliement, ce qui est une caractéristique importante de beaucoup d'IDP (232). D'ailleurs, une spécificité de XPG/Rad2 est que l'IDP se trouve à l'intérieur de la séquence de la protéine, entre deux domaines repliés ensemble. L'IDP forme donc une longue boucle.

Ce segment de XPG/Rad2 est également le candidat parfait pour la théorie du « fly-casting mechanism » (mécanisme de pêche à la mouche) (233). La base de cette théorie est que, tout comme un pêcheur à la mouche envoie sa ligne loin de lui afin d'y attraper un poisson, les segments désordonnés possèdent un plus grand rayon d'action pour attraper la protéine avec laquelle ils interagissent. Ainsi, ils augmentent la probabilité et la portée d'une protéine à interagir avec d'autres partenaires. Il est facile d'imaginer la boucle de plus de 600 acides aminés du segment espaceur de XPG/Rad2 explorer l'espace autour de la protéine en vue d'une interaction avec TFIIH. Cette longueur donne à la protéine un rayon d'action possible de plus de 1000 Å.

Un autre avantage des IDP est que par leur flexibilité, ils peuvent lier de multiples partenaires (234). Ceci est vrai surtout pour les protéines ayant de multiples fonctions dans diverses voies. Ainsi, l'avantage pour la cellule est de limiter la taille de ces protéines (235). Au lieu de posséder plusieurs domaines interagissant avec les diverses cibles, un seul domaine désordonné et flexible est nécessaire. Malheureusement, les données ne sont pas suffisantes pour affirmer que ceci est une caractéristique des segments désordonnés de

Rad2 et Rad4. Les interactions de ces protéines dans d'autres processus devront en effet être investiguées.

Par ailleurs, le taux de mutation dans les segments désordonnés est beaucoup plus élevé que dans les domaines ordonnés (236). Des mutations peuvent avoir lieu dans les IDP sans compromettre une structure tridimensionnelle complexe. L'exemple de XPG/Rad2 est frappant puisque seule la longueur et l'acidité du segment espaceur est conservé entre l'humain et la levure. Pourtant, la fonction de ce segment est conservée au cours de l'évolution et ce malgré des séquences différentes.

Finalement, les IDP sont généralement plus susceptibles à des modifications post-traductionnelles (234). La disponibilité des résidus, non-enfouis dans une structure secondaire, facilite leur accès à des enzymes de modification telles que les kinases. Comme nous avons discuté précédemment, nous n'avons pas observé de possibilité que des modifications post-traductionnelles puissent affecter les interactions des segments de Rad2 et Rad4 étudiés. Toutefois, la susceptibilité de Rad4 à la protéolyse *in vitro* est un indicateur que ce segment pourrait aussi être sensible à la dégradation par les protéases *in vivo*. Ceci pourrait être un mécanisme intéressant de régulation de la quantité de Rad4 actifs.

Importance de l'étude et conclusions

La découverte principale de cette thèse est qu'il existe un pont entre la transcription et la réparation au niveau du mécanisme d'interaction des partenaires de TFIIH. Nos structures comptent parmi les premières structures de segments désordonnés se repliant au contact de leur partenaire d'interaction lors de la NER. Ces segments adoptent la même structure que TFIIIE α interagissant avec TFIIH, partageant également un motif d'acides aminés conservé. De plus, la similarité de l'interaction de Rad4 et Rad2 avec Tfb1PH nous permet de proposer un mécanisme de transition et de déplacement de XPC/Rad4 par XPG/Rad2 lors de la NER au site de dommage à l'ADN.

Jusqu'à maintenant, les interactions de p62/Tfb1 étaient limitées à la transcription. Par cette thèse, nous avons fourni la preuve de l'importance de la sous-unité p62/Tfb1 de TFIIH dans la réparation. Cette protéine lie ainsi des segments désordonnés de protéines

dans différents processus. Son domaine a la possibilité d'interagir avec une multitude de protéines, de différentes façons.

Nous avons également démontré que plusieurs interactions faibles étaient nécessaires à la liaison de TFIIH pour Rad4 et Rad2. Cette multitude de sites de liaison rend l'étude des interactions protéine-protéine complexe, puisqu'il faut en retirer plusieurs avant de voir un effet. Notre approche minutieuse, combinant biologie structurale, caractérisation biophysique et expériences *in vivo* prouve qu'il est toutefois possible d'en faire l'étude.

Une meilleure compréhension des mécanismes de réparation de l'ADN mènera à de nouvelles thérapies. D'ailleurs, il est possible d'augmenter l'efficacité des agents anti-cancer en diminuant l'efficacité de la NER (237). Toutefois, il faut se montrer vigilant, car le risque d'affecter plusieurs processus cellulaires est grand et nous avons fait la démonstration d'une similarité dans le mécanisme d'interaction dans la transcription et la réparation. D'autres similarités moléculaires dans des processus reliés à l'ADN seront à investiguer avant de procéder à l'ingénierie de nouvelles thérapies.

Bibliographie

1. Crick, F. (1970) Central dogma of molecular biology, *Nature* 227, 561-563.
2. Friedberg, E. C., Walker, G. C., Siede, W., Wood, R. D., Schultz, T., and Ellenberger, T. (2005) *DNA Repair and Mutagenesis*, ASM Press, Washington, DC.
3. Finkel, T., and Holbrook, N. J. (2000) Oxidants, oxidative stress and the biology of ageing, *Nature* 408, 239-247.
4. Lindahl, T., and Nyberg, B. (1972) Rate of depurination of native deoxyribonucleic acid, *Biochemistry* 11, 3610-3618.
5. Lindahl, T. (1993) Instability and decay of the primary structure of DNA, *Nature* 362, 709-715.
6. Kunkel, T. A., and Erie, D. A. (2005) DNA mismatch repair, *Annu Rev Biochem* 74, 681-710.
7. Lovett, S. T. (2004) Encoded errors: mutations and rearrangements mediated by misalignment at repetitive DNA sequences, *Mol Microbiol* 52, 1243-1253.
8. Hefferin, M. L., and Tomkinson, A. E. (2005) Mechanism of DNA double-strand break repair by non-homologous end joining, *DNA Repair* 4, 639-648.
9. Yamaizumi, M., and Sugano, T. (1994) U.v.-induced nuclear accumulation of p53 is evoked through DNA damage of actively transcribed genes independent of the cell cycle, *Oncogene* 9, 2775-2784.
10. Goodman, M. F., and Tippin, B. (2000) Sloppier copier DNA polymerases involved in genome repair, *Curr Opin Genet Dev* 10, 162-168.
11. Barnes, D. E., and Lindahl, T. (2004) Repair and genetic consequences of endogenous DNA base damage in mammalian cells, *Annu Rev Genet* 38, 445-476.
12. Jiricny, J. (1998) Eukaryotic mismatch repair: an update, *Mutat Res* 409, 107-121.
13. Eker, A. P., Quayle, C., Chaves, I., and van der Horst, G. T. (2009) DNA repair in mammalian cells: Direct DNA damage reversal: elegant solutions for nasty problems, *Cell Mol Life Sci* 66, 968-980.
14. Takata, M., Sasaki, M. S., Sonoda, E., Morrison, C., Hashimoto, M., Utsumi, H., Yamaguchi-Iwai, Y., Shinohara, A., and Takeda, S. (1998) Homologous recombination and non-homologous end-joining pathways of DNA double-strand break repair have overlapping roles in the maintenance of chromosomal integrity in vertebrate cells, *EMBO J* 17, 5497-5508.

15. Friedberg, E. C. (1988) Deoxyribonucleic acid repair in the yeast *Saccharomyces cerevisiae*, *Microbiol Rev* 52, 70-102.
16. Prakash, S., Sung, P., and Prakash, L. (1993) DNA repair genes and proteins of *Saccharomyces cerevisiae*, *Annu Rev Genet* 27, 33-70.
17. De Weerd-Kastelein, E. A., Keijzer, W., and Bootsma, D. (1972) Genetic heterogeneity of xeroderma pigmentosum demonstrated by somatic cell hybridization, *Nat New Biol* 238, 80-83.
18. Collins, A. R. (1993) Mutant rodent cell lines sensitive to ultraviolet light, ionizing radiation and cross-linking agents: a comprehensive survey of genetic and biochemical characteristics, *Mutat Res* 293, 99-118.
19. Lehmann, A. R., Bootsma, D., Clarkson, S. G., Cleaver, J. E., McAlpine, P. J., Tanaka, K., Thompson, L. H., and Wood, R. D. (1994) Nomenclature of human DNA repair genes, *Mutat Res* 315, 41-42.
20. Zamble, D. B., Mu, D., Reardon, J. T., Sancar, A., and Lippard, S. J. (1996) Repair of cisplatin--DNA adducts by the mammalian excision nuclease, *Biochemistry* 35, 10004-10013.
21. Vermeulen, W. (2011) Dynamics of mammalian NER proteins, *DNA Repair* 10, sous presse.
22. DeLano, W. L. (2002) The PyMOL Molecular Graphics System, Delano Scientific, Palo Alto, CA.
23. Sugawara, K., Ng, J. M., Masutani, C., Iwai, S., van der Spek, P. J., Eker, A. P., Hanaoka, F., Bootsma, D., and Hoeijmakers, J. H. (1998) Xeroderma pigmentosum group C protein complex is the initiator of global genome nucleotide excision repair, *Mol Cell* 2, 223-232.
24. 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.
25. Wakasugi, M., and Sancar, A. (1998) Assembly, subunit composition, and footprint of human DNA repair excision nuclease, *Proc Natl Acad Sci USA* 95, 6669-6674.
26. Riedl, T., Hanaoka, F., and Egly, J. M. (2003) The comings and goings of nucleotide excision repair factors on damaged DNA, *EMBO J* 22, 5293-5303.
27. O'Donovan, A., Davies, A. A., Moggs, J. G., West, S. C., and Wood, R. D. (1994) XPG endonuclease makes the 3' incision in human DNA nucleotide excision repair, *Nature* 371, 432-435.

28. Jones, C. J., and Wood, R. D. (1993) Preferential binding of the xeroderma pigmentosum group A complementing protein to damaged DNA, *Biochemistry* 32, 12096-12104.
29. de Laat, W. L., Appeldoorn, E., Sugasawa, K., Weterings, E., Jaspers, N. G., and Hoeijmakers, J. H. (1998) DNA-binding polarity of human replication protein A positions nucleases in nucleotide excision repair, *Genes Dev* 12, 2598-2609.
30. Sijbers, A. M., de Laat, W. L., Ariza, R. R., Biggerstaff, M., Wei, Y. F., Moggs, J. G., Carter, K. C., Shell, B. K., Evans, E., de Jong, M. C., Rademakers, S., de Rooij, J., Jaspers, N. G., Hoeijmakers, J. H., and Wood, R. D. (1996) Xeroderma pigmentosum group F caused by a defect in a structure-specific DNA repair endonuclease, *Cell* 86, 811-822.
31. Shivji, M. K., Podust, V. N., Hubscher, U., and Wood, R. D. (1995) Nucleotide excision repair DNA synthesis by DNA polymerase epsilon in the presence of PCNA, RFC, and RPA, *Biochemistry* 34, 5011-5017.
32. Ogi, T., Limsirichaikul, S., Overmeer, R. M., Volker, M., Takenaka, K., Cloney, R., Nakazawa, Y., Niimi, A., Miki, Y., Jaspers, N. G., Mullenders, L. H., Yamashita, S., Fousteri, M. I., and Lehmann, A. R. (2010) Three DNA polymerases, recruited by different mechanisms, carry out NER repair synthesis in human cells, *Mol Cell* 37, 714-727.
33. Overmeer, R. M., Gourdin, A. M., Giglia-Mari, A., Kool, H., Houtsmuller, A. B., Siegal, G., Fousteri, M. I., Mullenders, L. H., and Vermeulen, W. (2010) Replication factor C recruits DNA polymerase delta to sites of nucleotide excision repair but is not required for PCNA recruitment, *Mol Cell Biol* 30, 4828-4839.
34. Staresincic, L., Fagbemi, A. F., Enzlin, J. H., Gourdin, A. M., Wijgers, N., Dunand-Sauthier, I., Giglia-Mari, G., Clarkson, S. G., Vermeulen, W., and Schärer, O. D. (2009) Coordination of dual incision and repair synthesis in human nucleotide excision repair, *EMBO J* 28, 1111-1120.
35. Moser, J., Kool, H., Giakzidis, I., Caldecott, K., Mullenders, L. H., and Fousteri, M. I. (2007) Sealing of chromosomal DNA nicks during nucleotide excision repair requires XRCC1 and DNA ligase III alpha in a cell-cycle-specific manner, *Mol Cell* 27, 311-323.
36. Fagbemi, A. F., Orelli, B., and Scharer, O. D. (2011) Regulation of endonuclease activity in human nucleotide excision repair, *DNA Repair* 10, sous presse.
37. Tantin, D., Kansal, A., and Carey, M. (1997) Recruitment of the putative transcription-repair coupling factor CSB/ERCC6 to RNA polymerase II elongation complexes, *Mol Cell Biol* 17, 6803-6814.
38. Fousteri, M., Vermeulen, W., van Zeeland, A. A., and Mullenders, L. H. (2006) Cockayne syndrome A and B proteins differentially regulate recruitment of

chromatin remodeling and repair factors to stalled RNA polymerase II in vivo, *Mol Cell* 23, 471-482.

39. Batty, D., Ropic'-Otrin, V., Levine, A. S., and Wood, R. D. (2000) Stable binding of human XPC complex to irradiated DNA confers strong discrimination for damaged sites, *J Mol Biol* 300, 275-290.
40. Hey, T., Lipps, G., Sugasawa, K., Iwai, S., Hanaoka, F., and Krauss, G. (2002) The XPC-HR23B complex displays high affinity and specificity for damaged DNA in a true-equilibrium fluorescence assay, *Biochemistry* 41, 6583-6587.
41. Fitch, M. E., Nakajima, S., Yasui, A., and Ford, J. M. (2003) In vivo recruitment of XPC to UV-induced cyclobutane pyrimidine dimers by the DDB2 gene product, *J Biol Chem* 278, 46906-46910.
42. Sugasawa, K., Okuda, Y., Saijo, M., Nishi, R., Matsuda, N., Chu, G., Mori, T., Iwai, S., Tanaka, K., and Hanaoka, F. (2005) UV-induced ubiquitylation of XPC protein mediated by UV-DDB-ubiquitin ligase complex, *Cell* 121, 387-400.
43. Gillette, T. G., Yu, S., Zhou, Z., Waters, R., Johnston, S. A., and Reed, S. H. (2006) Distinct functions of the ubiquitin-proteasome pathway influence nucleotide excision repair, *EMBO J* 25, 2529-2538.
44. Ramsey, K. L., Smith, J. J., Dasgupta, A., Maqani, N., Grant, P., and Auble, D. T. (2004) The NEF4 complex regulates Rad4 levels and utilizes Snf2/Swi2-related ATPase activity for nucleotide excision repair, *Mol Cell Biol* 24, 6362-6378.
45. Guerrero-Santoro, J., Kapetanaki, M. G., Hsieh, C. L., Gorbachinsky, I., Levine, A. S., and Ropic'-Otrin, V. (2008) The cullin 4B-based UV-damaged DNA-binding protein ligase binds to UV-damaged chromatin and ubiquitinates histone H2A, *Cancer Res* 68, 5014-5022.
46. Wang, H., Zhai, L., Xu, J., Joo, H. Y., Jackson, S., Erdjument-Bromage, H., Tempst, P., Xiong, Y., and Zhang, Y. (2006) Histone H3 and H4 ubiquitylation by the CUL4-DDB-ROC1 ubiquitin ligase facilitates cellular response to DNA damage, *Mol Cell* 22, 383-394.
47. Zhao, Q., Wang, Q. E., Ray, A., Wani, G., Han, C., Milum, K., and Wani, A. A. (2009) Modulation of nucleotide excision repair by mammalian SWI/SNF chromatin-remodeling complex, *J Biol Chem* 284, 30424-30432.
48. Gong, F., Fahy, D., and Smerdon, M. J. (2006) Rad4-Rad23 interaction with SWI/SNF links ATP-dependent chromatin remodeling with nucleotide excision repair, *Nat Struct Mol Biol* 13, 902-907.
49. Cartault, F., Nava, C., Malbrunot, A. C., Munier, P., Hebert, J. C., N'Guyen, P., Djeridi, N., Pariaud, P., Pariaud, J., Dupuy, A., Austerlitz, F., and Sarasin, A. (2011) A new XPC gene splicing mutation has lead to the highest worldwide

- prevalence of xeroderma pigmentosum in black Mahori patients, *DNA Repair* 10, 577-585.
50. Legerski, R., and Peterson, C. (1992) Expression cloning of a human DNA repair gene involved in xeroderma pigmentosum group C, *Nature* 359, 70-73.
 51. Masutani, C., Sugasawa, K., Yanagisawa, J., Sonoyama, T., Ui, M., Enomoto, T., Takio, K., Tanaka, K., van der Spek, P. J., Bootsma, D., and et al. (1994) Purification and cloning of a nucleotide excision repair complex involving the xeroderma pigmentosum group C protein and a human homologue of yeast RAD23, *EMBO J* 13, 1831-1843.
 52. Fleer, R., Nicolet, C. M., Pure, G. A., and Friedberg, E. C. (1987) RAD4 gene of *Saccharomyces cerevisiae*: molecular cloning and partial characterization of a gene that is inactivated in *Escherichia coli*, *Mol Cell Biol* 7, 1180-1192.
 53. Gietz, R. D., and Prakash, S. (1988) Cloning and nucleotide sequence analysis of the *Saccharomyces cerevisiae* RAD4 gene required for excision repair of UV-damaged DNA, *Gene* 74, 535-541.
 54. Araki, M., Masutani, C., Takemura, M., Uchida, A., Sugasawa, K., Kondoh, J., Ohkuma, Y., and Hanaoka, F. (2001) Centrosome protein centrin 2/caltractin 1 is part of the xeroderma pigmentosum group C complex that initiates global genome nucleotide excision repair, *J Biol Chem* 276, 18665-18672.
 55. Popescu, A., Miron, S., Blouquit, Y., Duchambon, P., Christova, P., and Craescu, C. T. (2003) Xeroderma pigmentosum group C protein possesses a high affinity binding site to human centrin 2 and calmodulin, *J Biol Chem* 278, 40252-40261.
 56. den Dulk, B., Sun, S. M., de Ruijter, M., Brandsma, J. A., and Brouwer, J. (2006) Rad33, a new factor involved in nucleotide excision repair in *Saccharomyces cerevisiae*, *DNA Repair* 5, 683-692.
 57. den Dulk, B., van Eijk, P., de Ruijter, M., Brandsma, J. A., and Brouwer, J. (2008) The NER protein Rad33 shows functional homology to human Centrin2 and is involved in modification of Rad4, *DNA Repair* 7, 858-868.
 58. Chen, L., and Madura, K. (2008) Centrin/Cdc31 is a novel regulator of protein degradation, *Mol Cell Biol* 28, 1829-1840.
 59. Min, J.-H., and Pavletich, N. P. (2007) Recognition of DNA damage by the Rad4 nucleotide excision repair protein, *Nature* 449, 570-575.
 60. Thompson, J. R. (2006) The Structure of the Human Centrin 2-Xeroderma Pigmentosum Group C Protein Complex, *J Biol Chem* 281, 18746-18752.
 61. Yang, A., Miron, S., Mouawad, L., Duchambon, P., Blouquit, Y., and Craescu, C. T. (2006) Flexibility and plasticity of human centrin 2 binding to the xeroderma

- pigmentosum group C protein (XPC) from nuclear excision repair, *Biochemistry* 45, 3653-3663.
62. Charbonnier, J. B., Renaud, E., Miron, S., Le Du, M. H., Blouquit, Y., Duchambon, P., Christova, P., Shosheva, A., Rose, T., Angulo, J. F., and Craescu, C. T. (2007) Structural, thermodynamic, and cellular characterization of human centrin 2 interaction with xeroderma pigmentosum group C protein, *J Mol Biol* 373, 1032-1046.
 63. Schaubert, C., Chen, L., Tongaonkar, P., Vega, I., Lambertson, D., Potts, W., and Madura, K. (1998) Rad23 links DNA repair to the ubiquitin/proteasome pathway, *Nature* 391, 715-718.
 64. Hiyama, H., Yokoi, M., Masutani, C., Sugasawa, K., Maekawa, T., Tanaka, K., Hoeijmakers, J. H., and Hanaoka, F. (1999) Interaction of hHR23 with S5a. The ubiquitin-like domain of hHR23 mediates interaction with S5a subunit of 26 S proteasome, *J Biol Chem* 274, 28019-28025.
 65. Wilkinson, C. R., Seeger, M., Hartmann-Petersen, R., Stone, M., Wallace, M., Semple, C., and Gordon, C. (2001) Proteins containing the UBA domain are able to bind to multi-ubiquitin chains, *Nat Cell Biol* 3, 939-943.
 66. Sugasawa, K., Okamoto, T., Shimizu, Y., Masutani, C., Iwai, S., and Hanaoka, F. (2001) A multistep damage recognition mechanism for global genomic nucleotide excision repair, *Genes Dev* 15, 507-521.
 67. Janicijevic, A., Sugasawa, K., Shimizu, Y., Hanaoka, F., Wijgers, N., Djurica, M., Hoeijmakers, J. H., and Wyman, C. (2003) DNA bending by the human damage recognition complex XPC-HR23B, *DNA Repair* 2, 325-336.
 68. Kusumoto, R., Masutani, C., Sugasawa, K., Iwai, S., Araki, M., Uchida, A., Mizukoshi, T., and Hanaoka, F. (2001) Diversity of the damage recognition step in the global genomic nucleotide excision repair in vitro, *Mutat Res* 485, 219-227.
 69. Araujo, S. J., Nigg, E. A., and Wood, R. D. (2001) Strong functional interactions of TFIIH with XPC and XPG in human DNA nucleotide excision repair, without a preassembled repairosome, *Mol Cell Biol* 21, 2281-2291.
 70. Yokoi, M., Masutani, C., Maekawa, T., Sugasawa, K., Ohkuma, Y., and Hanaoka, F. (2000) The xeroderma pigmentosum group C protein complex XPC-HR23B plays an important role in the recruitment of transcription factor IIIH to damaged DNA, *J Biol Chem* 275, 9870-9875.
 71. Xie, Z., Liu, S., Zhang, Y., and Wang, Z. (2004) Roles of Rad23 protein in yeast nucleotide excision repair, *Nucleic Acids Res* 32, 5981-5990.
 72. Sugasawa, K., Masutani, C., Uchida, A., Maekawa, T., van der Spek, P. J., Bootsma, D., Hoeijmakers, J. H., and Hanaoka, F. (1996) HHR23B, a human Rad23

- homolog, stimulates XPC protein in nucleotide excision repair in vitro, *Mol Cell Biol* 16, 4852-4861.
73. Sugasawa, K., Ng, J. M., Masutani, C., Maekawa, T., Uchida, A., van der Spek, P. J., Eker, A. P., Rademakers, S., Visser, C., Aboussekhra, A., Wood, R. D., Hanaoka, F., Bootsma, D., and Hoeijmakers, J. H. (1997) Two human homologs of Rad23 are functionally interchangeable in complex formation and stimulation of XPC repair activity, *Mol Cell Biol* 17, 6924-6931.
 74. van der Spek, P. J., Eker, A., Rademakers, S., Visser, C., Sugasawa, K., Masutani, C., Hanaoka, F., Bootsma, D., and Hoeijmakers, J. H. (1996) XPC and human homologs of RAD23: intracellular localization and relationship to other nucleotide excision repair complexes, *Nucleic Acids Res* 24, 2551-2559.
 75. Ng, J. M., Vermeulen, W., van der Horst, G. T., Bergink, S., Sugasawa, K., Vrieling, H., and Hoeijmakers, J. H. (2003) A novel regulation mechanism of DNA repair by damage-induced and RAD23-dependent stabilization of xeroderma pigmentosum group C protein, *Genes Dev* 17, 1630-1645.
 76. Heessen, S., Masucci, M. G., and Dantuma, N. P. (2005) The UBA2 domain functions as an intrinsic stabilization signal that protects Rad23 from proteasomal degradation, *Mol Cell* 18, 225-235.
 77. Raasi, S., Orlov, I., Fleming, K. G., and Pickart, C. M. (2004) Binding of polyubiquitin chains to ubiquitin-associated (UBA) domains of HHR23A, *J Mol Biol* 341, 1367-1379.
 78. Raasi, S., and Pickart, C. M. (2003) Rad23 ubiquitin-associated domains (UBA) inhibit 26 S proteasome-catalyzed proteolysis by sequestering lysine 48-linked polyubiquitin chains, *J Biol Chem* 278, 8951-8959.
 79. Chen, L., Shinde, U., Ortolan, T. G., and Madura, K. (2001) Ubiquitin-associated (UBA) domains in Rad23 bind ubiquitin and promote inhibition of multi-ubiquitin chain assembly, *EMBO Reports* 2, 933-938.
 80. Kang, Y., Chen, X., Lary, J. W., Cole, J. L., and Walters, K. J. (2007) Defining how ubiquitin receptors hHR23a and S5a bind polyubiquitin, *J Mol Biol* 369, 168-176.
 81. Nishi, R., Okuda, Y., Watanabe, E., Mori, T., Iwai, S., Masutani, C., Sugasawa, K., and Hanaoka, F. (2005) Centrin 2 Stimulates Nucleotide Excision Repair by Interacting with Xeroderma Pigmentosum Group C Protein, *Mol Cell Biol* 25, 5664-5674.
 82. Prakash, S., and Prakash, L. (2000) Nucleotide excision repair in yeast, *Mutat Res* 451, 13-24.
 83. Venema, J., van Hoffen, A., Karcagi, V., Natarajan, A. T., van Zeeland, A. A., and Mullenders, L. H. (1991) Xeroderma pigmentosum complementation group C cells

remove pyrimidine dimers selectively from the transcribed strand of active genes, *Mol Cell Biol* 11, 4128-4134.

84. den Dulk, B., Brandsma, J. A., and Brouwer, J. (2005) The Rad4 homologue YDR314C is essential for strand-specific repair of RNA polymerase I-transcribed rDNA in *Saccharomyces cerevisiae*, *Mol Microbiol* 56, 1518-1526.
85. Tremblay, M., Teng, Y., Paquette, M., Waters, R., and Conconi, A. (2008) Complementary Roles of Yeast Rad4p and Rad34p in Nucleotide Excision Repair of Active and Inactive rRNA Gene Chromatin, *Mol Cell Biol* 28, 7504-7513.
86. Conaway, J. W., and Conaway, R. C. (1989) A multisubunit transcription factor essential for accurate initiation by RNA polymerase II, *J Biol Chem* 264, 2357-2362.
87. 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.
88. 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.
89. Chang, W. H., and Kornberg, R. D. (2000) Electron crystal structure of the transcription factor and DNA repair complex, core TFIIH, *Cell* 102, 609-613.
90. Schultz, P., Fribourg, S., Poterszman, A., Mallouh, V., Moras, D., and Egly, J. M. (2000) Molecular structure of human TFIIH, *Cell* 102, 599-607.
91. Kwek, K. Y., O'Gorman, W., and Akoulitchev, A. (2004) Transcription meets DNA repair at a PH domain, *Nat Struct Mol Biol* 11, 588-589.
92. Shiekhatar, R., Mermelstein, F., Fisher, R. P., Drapkin, R., Dynlacht, B., Wessling, H. C., Morgan, D. O., and Reinberg, D. (1995) Cdk-activating kinase complex is a component of human transcription factor TFIIH, *Nature* 374, 283-287.
93. Adamczewski, J. P., Rossignol, M., Tassan, J. P., Nigg, E. A., Moncollin, V., and Egly, J. M. (1996) MAT1, cdk7 and cyclin H form a kinase complex which is UV light-sensitive upon association with TFIIH, *EMBO J* 15, 1877-1884.
94. Coin, F., Oksenysh, V., and Egly, J.-M. (2007) Distinct Roles for the XPB/p52 and XPD/p44 Subcomplexes of TFIIH in Damaged DNA Opening during Nucleotide Excision Repair, *Mol Cell* 26, 245-256.
95. Oksenysh, V., de Jesus, B. B., Zhovmer, A., Egly, J. M., and Coin, F. (2009) Molecular insights into the recruitment of TFIIH to sites of DNA damage, *EMBO J* 28, 2971-2980.

96. Giglia-Mari, G., Coin, F., Ranish, J. A., Hoogstraten, D., Theil, A., Wijgers, N., Jaspers, N. G., Raams, A., Argentini, M., van der Spek, P. J., Botta, E., Stefanini, M., Egly, J. M., Aebersold, R., Hoeijmakers, J. H., and Vermeulen, W. (2004) A new, tenth subunit of TFIIH is responsible for the DNA repair syndrome trichothiodystrophy group A, *Nature Genetics* 36, 714-719.
97. Ranish, J. A., Hahn, S., Lu, Y., Yi, E. C., Li, X. J., Eng, J., and Aebersold, R. (2004) Identification of TFB5, a new component of general transcription and DNA repair factor IIIH, *Nature Genetics* 36, 707-713.
98. Zhou, Y., Kou, H., and Wang, Z. (2007) Tfb5 interacts with Tfb2 and facilitates nucleotide excision repair in yeast, *Nucleic Acids Res* 35, 861-871.
99. Jawhari, A., Laine, J. P., Dubaele, S., Lamour, V., Poterszman, A., Coin, F., Moras, D., and Egly, J. M. (2002) p52 Mediates XPB function within the transcription/repair factor TFIIH, *J Biol Chem* 277, 31761-31767.
100. Feaver, W. J., Huang, W., and Friedberg, E. C. (1999) The TFB4 subunit of yeast TFIIH is required for both nucleotide excision repair and RNA polymerase II transcription, *J Biol Chem* 274, 29564-29567.
101. Matsui, P., DePaulo, J., and Buratowski, S. (1995) An interaction between the Tfb1 and Ssl1 subunits of yeast TFIIH correlates with DNA repair activity, *Nucleic Acids Res* 23, 767-772.
102. Gileadi, O., Feaver, W. J., and Kornberg, R. D. (1992) Cloning of a subunit of yeast RNA polymerase II transcription factor b and CTD kinase, *Science* 257, 1389-1392.
103. Fischer, L., Gerard, M., Chalut, C., Lutz, Y., Humbert, S., Kanno, M., Chambon, P., and Egly, J. M. (1992) Cloning of the 62-kilodalton component of basic transcription factor BTF2, *Science* 257, 1392-1395.
104. Doerks, T., Huber, S., Buchner, E., and Bork, P. (2002) BSD: a novel domain in transcription factors and synapse-associated proteins, *Trends Biochem Sci* 27, 168-170.
105. Jawhari, A., Boussert, S., Lamour, V., Atkinson, R. A., Kieffer, B., Poch, O., Potier, N., van Dorsselaer, A., Moras, D., and Poterszman, A. (2004) Domain architecture of the p62 subunit from the human transcription/repair factor TFIIH deduced by limited proteolysis and mass spectrometry analysis, *Biochemistry* 43, 14420-14430.
106. Wang, Z., Buratowski, S., Svejstrup, J. Q., Feaver, W. J., Wu, X., Kornberg, R. D., Donahue, T. F., and Friedberg, E. C. (1995) The yeast TFB1 and SSL1 genes, which encode subunits of transcription factor IIIH, are required for nucleotide excision repair and RNA polymerase II transcription, *Mol Cell Biol* 15, 2288-2293.
107. Gervais, V., Lamour, V., Jawhari, A., Frindel, F., Wasielewski, E., Dubaele, S., Egly, J. M., Thierry, J. C., Kieffer, B., and Poterszman, A. (2004) TFIIH contains a

- PH domain involved in DNA nucleotide excision repair, *Nat Struct Mol Biol* 11, 616-622.
108. Di Lello, P., Nguyen, B. D., Jones, T. N., Potempa, K., Kobor, M. S., Legault, P., and Omichinski, J. G. (2005) NMR structure of the amino-terminal domain from the Tfb1 subunit of TFIIH and characterization of its phosphoinositide and VP16 binding sites, *Biochemistry* 44, 7678-7686.
109. IHGSC. (2001) Initial sequencing and analysis of the human genome, *Nature* 409, 860-921.
110. Lemmon, M. A. (2004) Pleckstrin homology domains: not just for phosphoinositides, *Biochemical Society Transactions* 32, 707-711.
111. Blomberg, N., Baraldi, E., Nilges, M., and Saraste, M. (1999) The PH superfold: a structural scaffold for multiple functions, *Trends Biochemical Sci* 24, 441-445.
112. Winkler, G. S., Sugasawa, K., Eker, A. P., de Laat, W. L., and Hoeijmakers, J. H. (2001) Novel functional interactions between nucleotide excision DNA repair proteins influencing the enzymatic activities of TFIIH, XPG, and ERCC1-XPF, *Biochemistry* 40, 160-165.
113. Higgins, D. R., Prakash, L., Reynolds, P., and Prakash, S. (1984) Isolation and characterization of the RAD2 gene of *Saccharomyces cerevisiae*, *Gene* 30, 121-128.
114. Naumovski, L., and Friedberg, E. C. (1984) *Saccharomyces cerevisiae* RAD2 gene: isolation, subcloning, and partial characterization, *Mol Cell Biol* 4, 290-295.
115. Habraken, Y., Sung, P., Prakash, L., and Prakash, S. (1993) Yeast excision repair gene RAD2 encodes a single-stranded DNA endonuclease, *Nature* 366, 365-368.
116. Habraken, Y., Sung, P., Prakash, L., and Prakash, S. (1994) Human xeroderma pigmentosum group G gene encodes a DNA endonuclease, *Nucleic Acids Res* 22, 3312-3316.
117. Habraken, Y., Sung, P., Prakash, L., and Prakash, S. (1994) A conserved 5' to 3' exonuclease activity in the yeast and human nucleotide excision repair proteins RAD2 and XPG, *J Biol Chem* 269, 31342-31345.
118. MacInnes, M. A., Dickson, J. A., Hernandez, R. R., Learmonth, D., Lin, G. Y., Mudgett, J. S., Park, M. S., Schauer, S., Reynolds, R. J., Strniste, G. F., and et al. (1993) Human ERCC5 cDNA-cosmid complementation for excision repair and bipartite amino acid domains conserved with RAD proteins of *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe*, *Mol Cell Biol* 13, 6393-6402.
119. Shiomi, T., Harada, Y., Saito, T., Shiomi, N., Okuno, Y., and Yamaizumi, M. (1994) An ERCC5 gene with homology to yeast RAD2 is involved in group G xeroderma pigmentosum, *Mutat Res* 314, 167-175.

120. Scherly, D., Nospikel, T., Corlet, J., Ucla, C., Bairoch, A., and Clarkson, S. G. (1993) Complementation of the DNA repair defect in xeroderma pigmentosum group G cells by a human cDNA related to yeast RAD2, *Nature* 363, 182-185.
121. Mudgett, J. S., and MacInnes, M. A. (1990) Isolation of the functional human excision repair gene ERCC5 by intercosmid recombination, *Genomics* 8, 623-633.
122. O'Donovan, A., and Wood, R. D. (1993) Identical defects in DNA repair in xeroderma pigmentosum group G and rodent ERCC group 5, *Nature* 363, 185-188.
123. Harrington, J. J., and Lieber, M. R. (1994) Functional domains within FEN-1 and RAD2 define a family of structure-specific endonucleases: implications for nucleotide excision repair, *Genes Dev* 8, 1344-1355.
124. Lieber, M. R. (1997) The FEN-1 family of structure-specific nucleases in eukaryotic DNA replication, recombination and repair, *Bioessays* 19, 233-240.
125. Dunand-Sauthier, I., Hohl, M., Thorel, F., Jaquier-Gubler, P., Clarkson, S. G., and Scharer, O. D. (2005) The spacer region of XPG mediates recruitment to nucleotide excision repair complexes and determines substrate specificity, *J Biol Chem* 280, 7030-7037.
126. Cloud, K. G., Shen, B., Strniste, G. F., and Park, M. S. (1995) XPG protein has a structure-specific endonuclease activity, *Mutat Res* 347, 55-60.
127. Evans, E., Fellows, J., Coffey, A., and Wood, R. D. (1997) Open complex formation around a lesion during nucleotide excision repair provides a structure for cleavage by human XPG protein, *EMBO J* 16, 625-638.
128. Habraken, Y., Sung, P., Prakash, L., and Prakash, S. (1995) Structure-specific nuclease activity in yeast nucleotide excision repair protein Rad2, *J Biol Chem* 270, 30194-30198.
129. Liu, Y., Kao, H. I., and Bambara, R. A. (2004) Flap endonuclease 1: a central component of DNA metabolism, *Annu Rev Biochem* 73, 589-615.
130. Hohl, M., Dunand-Sauthier, I., Staresinic, L., Jaquier-Gubler, P., Thorel, F., Modesti, M., Clarkson, S. G., and Scharer, O. D. (2007) Domain swapping between FEN-1 and XPG defines regions in XPG that mediate nucleotide excision repair activity and substrate specificity, *Nucleic Acids Res* 35, 3053-3063.
131. Thorel, F., Constantinou, A., Dunand-Sauthier, I., Nospikel, T., Lalle, P., Raams, A., Jaspers, N. G., Vermeulen, W., Shivji, M. K., Wood, R. D., and Clarkson, S. G. (2004) Definition of a short region of XPG necessary for TFIIH interaction and stable recruitment to sites of UV damage, *Mol Cell Biol* 24, 10670-10680.
132. Iyer, N., Reagan, M. S., Wu, K. J., Canagarajah, B., and Friedberg, E. C. (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, 2157-2167.
133. He, Z., Henriksen, L. A., Wold, M. S., and Ingles, C. J. (1995) RPA involvement in the damage-recognition and incision steps of nucleotide excision repair, *Nature* 374, 566-569.
 134. Habraken, Y., Sung, P., Prakash, S., and Prakash, L. (1996) Transcription factor TFIIH and DNA endonuclease Rad2 constitute yeast nucleotide excision repair factor 3: implications for nucleotide excision repair and Cockayne syndrome, *Proc Natl Acad Sci USA* 93, 10718-10722.
 135. Mu, D., Park, C. H., Matsunaga, T., Hsu, D. S., Reardon, J. T., and Sancar, A. (1995) Reconstitution of human DNA repair excision nuclease in a highly defined system, *J Biol Chem* 270, 2415-2418.
 136. Gary, R., Ludwig, D. L., Cornelius, H. L., MacInnes, M. A., and Park, M. S. (1997) The DNA repair endonuclease XPG binds to proliferating cell nuclear antigen (PCNA) and shares sequence elements with the PCNA-binding regions of FEN-1 and cyclin-dependent kinase inhibitor p21, *J Biol Chem* 272, 24522-24529.
 137. Sarker, A. H., Tsutakawa, S. E., Kostek, S., Ng, C., Shin, D. S., Peris, M., Campeau, E., Tainer, J. A., Nogales, E., and Cooper, P. K. (2005) Recognition of RNA Polymerase II and Transcription Bubbles by XPG, CSB, and TFIIH: Insights for Transcription-Coupled Repair and Cockayne Syndrome, *Mol Cell* 20, 187-198.
 138. Knauf, J. A., Pendergrass, S. H., Marrone, B. L., Strniste, G. F., MacInnes, M. A., and Park, M. S. (1996) Multiple nuclear localization signals in XPG nuclease, *Mutat Res* 363, 67-75.
 139. Wakasugi, M., Reardon, J. T., and Sancar, A. (1997) The non-catalytic function of XPG protein during dual incision in human nucleotide excision repair, *J Biol Chem* 272, 16030-16034.
 140. Mu, D., Wakasugi, M., Hsu, D. S., and Sancar, A. (1997) Characterization of reaction intermediates of human excision repair nuclease, *J Biol Chem* 272, 28971-28979.
 141. Lee, S. K., Yu, S. L., Prakash, L., and Prakash, S. (2002) Requirement of yeast RAD2, a homolog of human XPG gene, for efficient RNA polymerase II transcription. implications for Cockayne syndrome, *Cell* 109, 823-834.
 142. Ito, S., Kuraoka, I., Chymkowitz, P., Compe, E., Takedachi, A., Ishigami, C., Coin, F., Egly, J.-M., and Tanaka, K. (2007) XPG stabilizes TFIIH, allowing transactivation of nuclear receptors: implications for cockayne syndrome in XP-G/CS patients, *Mol Cell* 26, 231-243.

143. Kang, M.-S., Yu, S.-L., Lim, H.-S., Choi, B., Park, C.-S., Kang, J.-H., and Lee, S.-K. (2010) Mitotic catastrophe induced by overexpression of budding yeast Rad2p, *Yeast* 27, 399-411.
144. Sun, X., Thrower, D., Qiu, J., Wu, P., Zheng, L., Zhou, M., Bachant, J., Wilson, D. M., 3rd, and Shen, B. (2003) Complementary functions of the *Saccharomyces cerevisiae* Rad2 family nucleases in Okazaki fragment maturation, mutation avoidance, and chromosome stability, *DNA Repair* 2, 925-940.
145. Cooper, P. K. (2011) Structural and functional insights into the multiple roles of XPG - a Swiss army knife for DNA repair, Gordon Research Conference on Mammalian DNA Repair, Ventura, USA, Personal communication.
146. Tsutakawa, S. E., Classen, S., Chapados, B. R., Arvai, A. S., Finger, L. D., Guenther, G., Tomlinson, C. G., Thompson, P., Sarker, A. H., Shen, B., Cooper, P. K., Grasby, J. A., and Tainer, J. A. (2011) Human flap endonuclease structures, DNA double-base flipping, and a unified understanding of the FEN1 superfamily, *Cell* 145, 198-211.
147. O'Donovan, A., Scherly, D., Clarkson, S. G., and Wood, R. D. (1994) Isolation of active recombinant XPG protein, a human DNA repair endonuclease, *J Biol Chem* 269, 15965-15968.
148. Constantinou, A., Gunz, D., Evans, E., Lalle, P., Bates, P. A., Wood, R. D., and Clarkson, S. G. (1999) Conserved residues of human XPG protein important for nuclease activity and function in nucleotide excision repair, *J Biol Chem* 274, 5637-5648.
149. Swift, M., and Chase, C. (1979) Cancer in families with xeroderma pigmentosum, *J Natl Cancer Inst* 62, 1415-1421.
150. Cleaver, J. E. (1968) Defective repair replication of DNA in xeroderma pigmentosum, *Nature* 218, 652-656.
151. Lehmann, A. R. (2003) DNA repair-deficient diseases, xeroderma pigmentosum, Cockayne syndrome and trichothiodystrophy, *Biochimie* 85, 1101-1111.
152. 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.
153. Nance, M. A., and Berry, S. A. (1992) Cockayne syndrome: review of 140 cases, *Am J Med Genet* 42, 68-84.
154. 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.

155. Le May, N., Egly, J. M., and Coin, F. (2010) True lies: the double life of the nucleotide excision repair factors in transcription and DNA repair, *J Nucleic Acids* 2010.
156. 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.
157. Coin, F., Marinoni, J. C., Rodolfo, C., Fribourg, S., Pedrini, A. M., and Egly, J. M. (1998) Mutations in the XPD helicase gene result in XP and TTD phenotypes, preventing interaction between XPD and the p44 subunit of TFIIH, *Nature Genetics* 20, 184-188.
158. Hwang, J. R., Moncollin, V., Vermeulen, W., Seroz, T., van Vuuren, H., Hoeijmakers, J. H., and Egly, J. M. (1996) A 3' --> 5' XPB helicase defect in repair/transcription factor TFIIH of xeroderma pigmentosum group B affects both DNA repair and transcription, *J Biol Chem* 271, 15898-15904.
159. Nospikel, T., and Clarkson, S. G. (1994) Mutations that disable the DNA repair gene XPG in a xeroderma pigmentosum group G patient, *Hum Mol Genet* 3, 963-967.
160. Bernardes de Jesus, B. M., Bjoras, M., Coin, F., and Egly, J. M. (2008) Dissection of the molecular defects caused by pathogenic mutations in the DNA repair factor XPC, *Mol Cell Biol* 28, 7225-7235.
161. Yasuda, G., Nishi, R., Watanabe, E., Mori, T., Iwai, S., Orioli, D., Stefanini, M., Hanaoka, F., and Sugawara, K. (2007) In vivo destabilization and functional defects of the xeroderma pigmentosum C protein caused by a pathogenic missense mutation, *Mol Cell Biol* 27, 6606-6614.
162. Burley, S. K., and Roeder, R. G. (1996) Biochemistry and structural biology of transcription factor IID (TFIID), *Annu Rev Biochem* 65, 769-799.
163. Kim, J. L., Nikolov, D. B., and Burley, S. K. (1993) Co-crystal structure of TBP recognizing the minor groove of a TATA element, *Nature* 365, 520-527.
164. Maxon, M. E., Goodrich, J. A., and Tjian, R. (1994) Transcription factor IIE binds preferentially to RNA polymerase IIa and recruits TFIIH: a model for promoter clearance, *Genes Dev* 8, 515-524.
165. Kim, T. K., Ebright, R. H., and Reinberg, D. (2000) Mechanism of ATP-dependent promoter melting by transcription factor IIH, *Science* 288, 1418-1422.
166. 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.

167. Tirode, F., Busso, D., Coin, F., and Egly, J. M. (1999) Reconstitution of the transcription factor TFIIH: assignment of functions for the three enzymatic subunits, XPB, XPD and cdk7, *Mol Cell* 3, 87-95
168. Moreland, R. J., Tirode, F., Yan, Q., Conaway, J. W., Egly, J. M., and Conaway, R. C. (1999) A role for the TFIIH XPB DNA helicase in promoter escape by RNA polymerase II, *J Biol Chem* 274, 22127-22130.
169. Fukuda, A., Nogi, Y., and Hisatake, K. (2002) The regulatory role for the ERCC3 helicase of general transcription factor TFIIH during promoter escape in transcriptional activation, *Proc Natl Acad Sci USA* 99, 1206-1211.
170. Gerber, H. P., Hagmann, M., Seipel, K., Georgiev, O., West, M. A., Litingtung, Y., Schaffner, W., and Corden, J. L. (1995) RNA polymerase II C-terminal domain required for enhancer-driven transcription, *Nature* 374, 660-662.
171. Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., and Walter, P. (2008) *Molecular Biology of the Cell*, 5th ed., Garland Science, New York, NY.
172. Bastien, J., Adam-Stitah, S., Riedl, T., Egly, J.-M., Chambon, P., and Rochette-Egly, C. (2000) TFIIH interacts with the retinoic acid receptor α and phosphorylates its AF-1-activating domain through cdk7, *J Biol Chem* 275, 21896-21904.
173. Chen, D., Riedl, T., Washbrook, E., Pace, P. E., Coombes, R. C., Egly, J. M., and Ali, S. (2000) Activation of estrogen receptor α by S118 phosphorylation involves a ligand-dependent interaction with TFIIH and participation of CDK7, *Mol Cell* 6, 127-137.
174. Iben, S., Tschochner, H., Bier, M., Hoogstraten, D., Hozak, P., Egly, J. M., and Grummt, I. (2002) TFIIH plays an essential role in RNA polymerase I transcription, *Cell* 109, 297-306.
175. Hoogstraten, D., Nigg, A. L., Heath, H., Mullenders, L. H., van Driel, R., Hoeijmakers, J. H., Vermeulen, W., and Houtsmuller, A. B. (2002) Rapid switching of TFIIH between RNA polymerase I and II transcription and DNA repair in vivo, *Mol Cell* 10, 1163-1174.
176. Ohkuma, Y., and Roeder, R. G. (1994) Regulation of TFIIH ATPase and kinase activities by TFIIIE during active initiation complex formation, *Nature* 368, 160-163.
177. Ohkuma, Y., Sumimoto, H., Horikoshi, M., and Roeder, R. G. (1990) Factors involved in specific transcription by mammalian RNA polymerase II: purification and characterization of general transcription factor TFIIIE, *Proc Natl Acad Sci USA* 87, 9163-9167.
178. Ohkuma, Y., Hashimoto, S., Wang, C. K., Horikoshi, M., and Roeder, R. G. (1995) Analysis of the role of TFIIIE in basal transcription and TFIIH-mediated carboxy-

- terminal domain phosphorylation through structure-function studies of TFIIE α ., *Mol Cell Biol* 15, 4856-4866.
179. Yamamoto, S., Watanabe, Y., van der Spek, P. J., Watanabe, T., Fujimoto, H., Hanaoka, F., and Ohkuma, Y. (2001) Studies of nematode TFIIE function reveal a link between Ser-5 phosphorylation of RNA polymerase II and the transition from transcription initiation to elongation, *Mol Cell Biol* 21, 1-15.
180. Bushnell, D. A., Bamdad, C., and Kornberg, R. D. (1996) A minimal set of RNA polymerase II transcription protein interactions, *J Biol Chem* 271, 20170-20174.
181. Di Lello, P., Miller Jenkins, L. M., Mas, C., Langlois, C., Malitskaya, E., Fradet-Turcotte, A., Archambault, J., Legault, P., and Omichinski, J. G. (2008) p53 and TFIIE α share a common binding site on the Tfb1/p62 subunit of TFIIF, *Proc Natl Acad Sci USA* 105, 106-111.
182. Okuda, M., Tanaka, A., Satoh, M., Mizuta, S., Takazawa, M., Ohkuma, Y., and Nishimura, Y. (2008) Structural insight into the TFIIE-TFIIF interaction: TFIIE and p53 share the binding region on TFIIF, *EMBO J* 27, 1161-1171.
183. Ptashne, M., and Gann, A. A. F. (1997) Transcriptional activation by recruitment., *Nature* 386, 569-577.
184. Fuda, N. J., Ardehali, M. B., and Lis, J. T. (2009) Defining mechanisms that regulate RNA polymerase II transcription in vivo, *Nature* 461, 186-192.
185. Blau, J., Xiao, H., McCracken, S., O'Hare, P., Greenblatt, J., and Bentley, D. (1996) Three functional classes of transcriptional activation domains., *Mol Cell Biol* 16, 2044-2055.
186. Xiao, H., Pearson, A., Coulombe, B., Truant, R., Zhang, S., Regier, J. L., Triezenberg, S. J., Reinberg, D., Flores, O., Ingles, C. J., and et al. (1994) Binding of basal transcription factor TFIIF to the acidic activation domains of VP16 and p53, *Mol Cell Biol* 14, 7013-7024.
187. Tong, X., Drapkin, R., Reinberg, D., and Kieff, E. (1995) The 62- and 80-kDa subunits of transcription factor IIF mediate the interaction with Epstein-Barr virus nuclear protein 2, *Proc Natl Acad Sci USA* 92, 3259-3263.
188. Pearson, A., and Greenblatt, J. (1997) Modular organization of the E2F1 activation domain and its interaction with general transcription factors TBP and TFIIF, *Oncogene* 15, 2643-2658.
189. Leveillard, T., Andera, L., Bissonnette, N., Schaeffer, L., Bracco, L., Egly, J. M., and Wasylyk, B. (1996) Functional interactions between p53 and the TFIIF complex are affected by tumour-associated mutations, *EMBO J* 15, 1615-1624.
190. Vandel, L., and Kouzarides, T. (1999) Residues phosphorylated by TFIIF are required for E2F-1 degradation during S-phase, *EMBO J* 18, 4280-4291.

191. Hollstein, M., Sidransky, D., Vogelstein, B., and Harris, C. C. (1991) p53 mutations in human cancers., *Science* 253, 49-53.
192. Levine, A. J. (1997) p53, the cellular gatekeeper for growth and division., *Cell* 88, 323-331.
193. Candau, R., Scolnick, D. M., Darpino, P., Ying, C. Y., Halazonetis, T. D., and Berger, S. L. (1997) Two tandem and independent sub-activation domains in the amino terminus of p53 require the adaptor complex for activity., *Oncogene* 15, 807-816.
194. Kussie, P. H., Gorina, S., Marechal, V., Elenbaas, B., Moreau, J., Levine, A. J., and Pavletich, N. P. (1996) Structure of the MDM2 oncoprotein bound to the p53 tumor suppressor transactivation domain, *Science* 274, 948-953.
195. Martin, D. W., Munoz, R. M., Subler, M. A., and Deb, S. (1993) p53 binds to the TATA-binding protein-TATA complex, *J Biol Chem* 268, 13062-13067.
196. Gu, W., Shi, X. L., and Roeder, R. G. (1997) Synergistic activation of transcription by CBP and p53, *Nature* 387, 819-823.
197. Bochkareva, E., Kaustov, L., Ayed, A., Yi, G. S., Lu, Y., Pineda-Lucena, A., Liao, J. C., Okorokov, A. L., Milner, J., Arrowsmith, C. H., and Bochkarev, A. (2005) Single-stranded DNA mimicry in the p53 transactivation domain interaction with replication protein A, *Proc Natl Acad Sci USA* 102, 15412-15417.
198. Di Lello, P., Jenkins, L. M., Jones, T. N., Nguyen, B. D., Hara, T., Yamaguchi, H., Dikeakos, J. D., Appella, E., Legault, P., and Omichinski, J. G. (2006) Structure of the Tfb1/p53 complex: Insights into the interaction between the p62/Tfb1 subunit of TFIIH and the activation domain of p53, *Mol Cell* 22, 731-740.
199. Langlois, C., Mas, C., Di Lello, P., Jenkins, L. M., Legault, P., and Omichinski, J. G. (2008) NMR structure of the complex between the Tfb1 subunit of TFIIH and the activation domain of VP16: structural similarities between VP16 and p53, *J Am Chem Soc* 130, 10596-10604.
200. Mas, C., Lussier-Price, M., Soni, S., Morse, T., Arseneault, G., Di Lello, P., Lafrance-Vanasse, J., Bieker, J. J., and Omichinski, J. G. (2011) Structural and functional characterization of an atypical activation domain in erythroid Kruppel-like factor (EKLF), *Proc Natl Acad Sci USA*.
201. Tantin, D. (1998) RNA polymerase II elongation complexes containing the Cockayne syndrome group B protein interact with a molecular complex containing the transcription factor IIH components xeroderma pigmentosum B and p62, *J Biol Chem* 273, 27794-27799.
202. Bardwell, A. J., Bardwell, L., Iyer, N., Svejstrup, J. Q., Feaver, W. J., Kornberg, R. D., and Friedberg, E. C. (1994) Yeast nucleotide excision repair proteins Rad2 and

- Rad4 interact with RNA polymerase II basal transcription factor b (TFIIH), *Mol Cell Biol* 14, 3569-3576.
203. Park, C. H., Mu, D., Reardon, J. T., and Sancar, A. (1995) The general transcription-repair factor TFIIH is recruited to the excision repair complex by the XPA protein independent of the TFIIIE transcription factor, *J Biol Chem* 270, 4896-4902.
204. Uchida, A., Sugasawa, K., Masutani, C., Dohmae, N., Araki, M., Yokoi, M., Ohkuma, Y., and Hanaoka, F. (2002) The carboxy-terminal domain of the XPC protein plays a crucial role in nucleotide excision repair through interactions with transcription factor IIH, *DNA Repair* 1, 449-461.
205. Hofmann, K. (2009) Ubiquitin-binding domains and their role in the DNA damage response, *DNA Repair* 8, 544-556.
206. Park, M. S., Knauf, J. A., Pendergrass, S. H., Coulon, C. H., Strniste, G. F., Marrone, B. L., and MacInnes, M. A. (1996) Ultraviolet-induced movement of the human DNA repair protein, Xeroderma pigmentosum type G, in the nucleus, *Proc Natl Acad Sci USA* 93, 8368-8373.
207. Cavanagh, J., Fairbrother, W., Palmer, A., Rance, M., and Skelton, N. (2007) *Protein NMR Spectroscopy*, Elsevier, Burlington, MA.
208. Wittekind, M., and Mueller, L. (1993) HNCACB, a high-sensitivity 3D NMR experiment to correlate amide-proton and nitrogen resonances with the alpha- and beta-carbon resonances in proteins, *J Magn Reson B* 101, 201-205.
209. Grzesiek, S., and Bax, A. (1992) Correlating backbone amide and side chain resonances in larger proteins by multiple relayed triple resonance NMR, *J Am Chem Soc* 114, 6291-6293.
210. Kay, L. E., Xu, G. Y., and Yamazaki, T. (1994) Enhanced-sensitivity triple-resonance spectroscopy with minimal H₂O saturation, *J Magn Reson A* 109, 129-133.
211. Logan, T. M., Olejniczak, E. T., Xu, R. X., and Fesik, S. W. (1992) Side chain and backbone assignments in isotopically labeled proteins from two heteronuclear triple resonance experiments, *FEBS Lett* 314, 413-418.
212. Yamazaki, T., Forman-Kay, J. D., and Kay, L. E. (1993) Two-dimensional NMR experiments for correlating carbon-13 beta and proton delta/epsilon chemical shifts of aromatic residues in ¹³C-labeled proteins via scalar couplings, *J Am Chem Soc* 115, 11054-11055.
213. Ikura, M., Kay, L. E., and Bax, A. (1991) Improved three-dimensional ¹H-¹³C-¹H correlation spectroscopy of a ¹³C-labeled protein using constant-time evolution, *J Biomol NMR* 1, 299-304.

214. Zuiderweg, E. R. P., McIntosh, L. P., Dahlquist, F. W., and Fesik, S. W. (1990) Three-dimensional carbon-13-resolved proton NOE spectroscopy of uniformly carbon-13-labeled proteins for the NMR assignment and structure determination of larger molecules, *J Magn Reson* 86, 210-216.
215. Pascal, S. M., Muhandiram, D. R., Yamazaki, T., Forman-Kay, J. D., and Kay, L. E. (1994) Simultaneous acquisition of ¹⁵N- and ¹³C-edited NOE spectra of proteins dissolved in H₂O, *J Magn Reson* 103, 197-201.
216. Zhang, O., Kay, L. E., Olivier, J. P., and Forman-Kay, J. D. (1994) Backbone ¹H and ¹⁵N resonance assignments of the N-terminal SH3 domain of drk in folded and unfolded states using enhanced-sensitivity pulsed field gradient NMR techniques, *J Biomol NMR* 4, 845-858.
217. Shen, Y., Delaglio, F., Cornilescu, G., and Bax, A. (2009) TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts, *J Biomol NMR* 44, 213-223.
218. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Crystallography & NMR System: A New Software Suite for Macromolecular Structure Determination, *Acta Crystallogr D Biol Crystallogr* 54, 905-921.
219. Weber, P. C., and Salemme, F. R. (2003) Applications of calorimetric methods to drug discovery and the study of protein interactions, *Curr Opin Struct Biol* 13, 115-121.
220. Day, Y. S., Baird, C. L., Rich, R. L., and Myszka, D. G. (2002) Direct comparison of binding equilibrium, thermodynamic, and rate constants determined by surface- and solution-based biophysical methods, *Protein Sci* 11, 1017-1025.
221. Robinson, G. W., Nicolet, C. M., Kalainov, D., and Friedberg, E. C. (1986) A yeast excision-repair gene is inducible by DNA damaging agents, *Proceedings of the National Academy of Sciences of the United States of America* 83, 1842-1846.
222. Siede, W., Robinson, G. W., Kalainov, D., Malley, T., and Friedberg, E. C. (1989) Regulation of the RAD2 gene of *Saccharomyces cerevisiae*, *Mol Microbiol* 3, 1697-1707.
223. Ludwig, D. L., Mudgett, J. S., Park, M. S., Perez-Castro, A. V., and MacInnes, M. A. (1996) Molecular cloning and structural analysis of the functional mouse genomic XPG gene, *Mamm Genome* 7, 644-649.
224. Rezvani, H. R., Mahfouf, W., Ali, N., Chemin, C., Ged, C., Kim, A. L., de Verneuil, H., Taieb, A., Bickers, D. R., and Mazurier, F. (2010) Hypoxia-inducible factor-1alpha regulates the expression of nucleotide excision repair proteins in keratinocytes, *Nucleic Acids Res* 38, 797-809.

225. Adimoolam, S., and Ford, J. M. (2002) p53 and DNA damage-inducible expression of the xeroderma pigmentosum group C gene, *Proceedings of the National Academy of Sciences of the United States of America* 99, 12985-12990.
226. Heine, G. F., Horwitz, A. A., and Parvin, J. D. (2008) Multiple mechanisms contribute to inhibit transcription in response to DNA damage, *J Biol Chem* 283, 9555-9561.
227. Rockx, D. A., Mason, R., van Hoffen, A., Barton, M. C., Citterio, E., Bregman, D. B., van Zeeland, A. A., Vrieling, H., and Mullenders, L. H. (2000) UV-induced inhibition of transcription involves repression of transcription initiation and phosphorylation of RNA polymerase II, *Proceedings of the National Academy of Sciences of the United States of America* 97, 10503-10508.
228. Dunker, A. K., Brown, C. J., Lawson, J. D., Iakoucheva, L. M., and Obradovic, Z. (2002) Intrinsic disorder and protein function, *Biochemistry* 41, 6573-6582.
229. Bandaru, V., Cooper, W., Wallace, S. S., and Doublet, S. (2004) Overproduction, crystallization and preliminary crystallographic analysis of a novel human DNA-repair enzyme that recognizes oxidative DNA damage, *Acta crystallographica Section D, Biological crystallography* 60, 1142-1144.
230. Hegde, M. L., Hazra, T. K., and Mitra, S. (2010) Functions of disordered regions in mammalian early base excision repair proteins, *Cellular and Molecular Life Sciences* 67, 3573-3587.
231. Iakoucheva, L. M., Kimzey, A. L., Masselon, C. D., Bruce, J. E., Garner, E. C., Brown, C. J., Dunker, A. K., Smith, R. D., and Ackerman, E. J. (2001) Identification of intrinsic order and disorder in the DNA repair protein XPA, *Protein Sci* 10, 560-571.
232. Dyson, H. J., and Wright, P. E. (2002) Coupling of folding and binding for unstructured proteins, *Curr Opin Struct Biol* 12, 54-60.
233. Shoemaker, B. A., Portman, J. J., and Wolynes, P. G. (2000) Speeding molecular recognition by using the folding funnel: the fly-casting mechanism, *Proceedings of the National Academy of Sciences of the United States of America* 97, 8868-8873.
234. Dyson, H. J., and Wright, P. E. (2005) Intrinsically unstructured proteins and their functions, *Nat Rev Mol Cell Biol* 6, 197-208.
235. Gunasekaran, K., Tsai, C. J., Kumar, S., Zanuy, D., and Nussinov, R. (2003) Extended disordered proteins: targeting function with less scaffold, *Trends Biochem Sci* 28, 81-85.
236. Brown, C. J., Takayama, S., Campen, A. M., Vise, P., Marshall, T. W., Oldfield, C. J., Williams, C. J., and Dunker, A. K. (2002) Evolutionary rate heterogeneity in proteins with long disordered regions, *J Mol Evol* 55, 104-110.

237. Sabatino, M. A., Marabese, M., Ganzinelli, M., Caiola, E., Geroni, C., and Broggin, M. (2010) Down-regulation of the nucleotide excision repair gene XPG as a new mechanism of drug resistance in human and murine cancer cells, *Mol Cancer* 9, 259.