

12-2016

NUCLEOTIDE EXCISION REPAIR, CROSSLINK REPAIR AND TRANSCRIPTIONAL FUNCTION OF XPA IN HUMAN CELLS

Mandira Mananadhar

Follow this and additional works at: https://digitalcommons.library.tmc.edu/utgsbs_dissertations

 Part of the [Medicine and Health Sciences Commons](#)

Recommended Citation

Mananadhar, Mandira, "NUCLEOTIDE EXCISION REPAIR, CROSSLINK REPAIR AND TRANSCRIPTIONAL FUNCTION OF XPA IN HUMAN CELLS" (2016). *The University of Texas MD Anderson Cancer Center UTHealth Graduate School of Biomedical Sciences Dissertations and Theses (Open Access)*. 726.
https://digitalcommons.library.tmc.edu/utgsbs_dissertations/726

This Dissertation (PhD) is brought to you for free and open access by the The University of Texas MD Anderson Cancer Center UTHealth Graduate School of Biomedical Sciences at DigitalCommons@TMC. It has been accepted for inclusion in The University of Texas MD Anderson Cancer Center UTHealth Graduate School of Biomedical Sciences Dissertations and Theses (Open Access) by an authorized administrator of DigitalCommons@TMC. For more information, please contact digitalcommons@library.tmc.edu.

NUCLEOTIDE EXCISION REPAIR, CROSSLINK REPAIR AND
TRANSCRIPTIONAL FUNCTION OF XPA IN HUMAN CELLS

BY

MANDIRA MANANDHAR, M.S.

APPROVED:

Richard D. Wood, PhD
Major Advisor

Karen M. Vasquez, PhD

Mark T. Bedford, PhD

Shawn Bratton, PhD

Rick A. Finch, PhD, DABT

APPROVED:

Dean, The University of Texas
Graduate School of Biomedical Sciences at Houston

NUCLEOTIDE EXCISION REPAIR, CROSSLINK REPAIR AND
TRANSCRIPTIONAL FUNCTION OF XPA IN HUMAN CELLS

A DESSERTATION

Presented to the faculty of

The University of Texas

Health Science Center at Houston

and

The University of Texas

MD Anderson Cancer Center

Graduate School of Biomedical Sciences

in partial fulfillment of

the requirements

for the degree of

DOCTOR OF PHILOSOPHY

by

Mandira Manandhar, M.S

Houston, Texas

December, 2016

DEDICATION

This dissertation work is dedicated to my parents, my husband and my siblings.

ACKNOWLEDGEMENTS

First and foremost, I would like to thank Dr. Richard Wood for his all kinds of strategic guidance, discussion and support to help me accomplish my scientific goals during six and half years of working in his laboratory. His visionary thoughts and ideas during research and general discussions have always helped me to enjoy the research work throughout my being here. Rick is a real teacher and I will be always thankful to him for raising me as a sustainable learner.

I would like to thank Dr. Vasquez for allowing me to learn, work and be a part in her laboratory from the very beginning of my doctoral research. I thank her and her lab members for their valuable time, discussions and hands-on-training necessary for my projects to keep it moving despite of many difficulties.

I would like to thank my past and present committee members; Drs. Rodney Nairn, Sharon Dent, Karen Vasquez, Mark Bedford, Shawn Bratton and Rick Finch. Valuable suggestions from the committee were very important to shape up my research projects. Their critical comments have been key in giving right direction to my research projects and for my overall performance as a graduate student.

It would be never enough to thank my lab members for supporting me and helping me in every possible way they could. Karen Boulware, Shelley Reh and Megan Lowery are excellent critics for lab works, lab activities and friendly nature, that I can always rely on. Megan's expertise in cell culture was very important for the accomplishments with all the cell works. Dr. Kei-ichi Takata and Dr. Junya Tomida are great support system in the lab and I would like to thank them for being great trainers and for discussing every confusion I had, with wide big smiles even during hot discussions. I would like to thank Kei-ichi for his meticulous training during my lab-rotation in Wood lab that helped me to get exposed to DNA repair research.

No words will be enough to express my gratitude towards my family for supporting me everyday during my high and low moments during graduate school. My mom (Bina Manandhar) who started her school after all her kids went to college is the biggest inspiration for me to do aspiring work for others and myself. I thank my dad (Shreekrishna Manandhar) for always challenging me to do things that seem to be impossible. My siblings (Indira, Bikash and Bikram) are my best friends and support to keep me on track towards progress. I thank my husband (Surendra Prajapati) for being my biggest critic and the best friend who has been by my side through all the joys, weeps, amusements and frustrations.

NUCLEOTIDE EXCISION REPAIR, CROSSLINK REPAIR AND TRANSCRIPTIONAL FUNCTION OF XPA IN HUMAN CELLS

Mandira Manandhar, PhD

Supervisory Professor: Richard D. Wood, PhD

Nucleotide excision repair (NER) in mammalian cells includes xeroderma pigmentosum group A protein (XPA) as a core factor. XPA and other NER proteins have been detected previously at some active promoters, and NER deficiency is reported to decrease activated transcription of selected genes. To determine the global extent of XPA influence on transcription, we analyzed the human transcriptome by RNA sequencing. We first confirmed that XPA is confined to the cell nucleus even in the absence of external DNA damage, in contrast to previous reports that XPA is normally resident in the cytoplasm and is imported following DNA damage. We then analyzed four genetically matched human cell line pairs deficient or proficient in XPA. At a false discovery rate of 0.05, 325 genes were common in all four pairs with a significant XPA-dependent directional change in gene expression. These genes were highly represented in pathways for the maintenance of mitochondria, metabolism and neurological system. Only 27 genes were regulated by more than 1.5 fold change. The most significant hits were *AKR1C1* and *AKR1C2*, involved in steroid hormone metabolism, and the corresponding proteins were lower in XPA-deficient cells. Transactivation by retinoic acid caused a modest enrichment of genes involved in

transcription-related functions in XPA proficient cells. The results show that XPA status significantly influences a small subset of human genes that are important for mitochondrial and metabolic functions. The results may help explain defects in neurological function and sterility in individuals with xeroderma pigmentosum (XP).

An NER deficiency enhances sensitivity of mammalian cells to DNA interstrand crosslinks (ICL)-generating agents. I found that XPA is retained on damaged DNA following exposure to UVA-activated psoralen, and investigated repair of a triplex forming oligonucleotide (TFO)-directed psoralen ICL. A TFO-directed psoralen DNA ICL was constructed in closed-circular DNA. In NER proficient human cell extracts, incisions were detected on both strands of the damaged DNA 3' to the psoralen ICL. Incision sites on the TFO bound strand were flanked by incision sites 40-42 nucleotides away from the ICL, with incisions 10-12 nucleotides away on the other strand.

Table of Contents

Table of Contents	viii
List of Illustrations	xiii
List of Tables	xvi
Abbreviations	xvii
Chapter 1. Introduction	1
1.1. DNA-interstrand crosslink in mammalian cells	1
1.2. Nucleotide excision repair pathway and XPA protein.....	6
1.3. Transcriptional role of XPA, a major NER protein	13
1.4. Localization of XPA protein	17
Chapter 2. Methods and Materials	20
2.1. Preparation of samples for the next generation sequencing of RNA from XPA deficient and proficient human cell lines	20
2.1.1. Cell lines	20
2.1.2. Immunoblotting	23
2.1.3. UVC clonogenic survival assay	24
2.1.4. NER extracts and assays	25
2.1.5. Immunofluorescence assay	26
2.1.6. XPA-extraction immunofluorescence assay	26
2.1.7. UVC exposure	27
2.1.8. PUVA treatment.....	27
2.2. Transactivation and NGS	28

2.3. RNA Sequencing	29
2.4. Analysis of NGS data	30
2.5. Preparation of components needed to establish the NER dual incision control assay	30
2.5.1. M13mp18GTG single stranded DNA	30
2.5.2. Whole cell extracts.....	33
2.5.3. Double stranded circular DNA with intrastrand cisplatin adduct	33
2.6. DNA incision and repair assays with cisplatin intrastrand substrate	37
2.6.1. NER dual incision control assay	37
2.6.2. Primer extension assay to a 1,3 (dGpTpG) cisplatin adduct	38
2.6.3. DNA repair assay.....	39
2.7. Construction of TFO-directed psoralen ICL on a a closed circular double stranded DNA (psoralen-TFO ICL).....	40
2.7.1. Vector with BbsI ends	40
2.7.2. 50 bp psoralen-TFO ICL	41
2.7.3. Ligation reaction	43
2.7.4. Confirmation of the ICL in the final product	43
2.7.5. Plasmid relaxation assay in whole cell extract.....	44
2.8. DNA incision and repair assays with psoralen-TFO ICL	45
2.8.1. Detection of 3' incision on the upper and lower strand	45
2.8.2. psoralen-TFO ICL repair assays.....	46
Chapter 3. Results - Consequences of XPA disruption in human cells	47
3.1. Validation of pairs of XPA-deficient and proficient cell lines.....	47

3.2. XPA is localized in the nucleus of cells and does not require UV irradiation for import	56
3.3. Influence of XPA status on overall gene transcription.....	59
3.4. XPA status influences expression of genes affecting mitochondria and mitophagy.....	78
3.5. Classification of the most differentially expressed genes common among all cell pairs.....	78
3.6. More commonalities within fibroblast and HeLa cell pairs.....	83
3.7. Retinoic acid transactivation affects XPA-dependent transcription related functions	83
Chapter 4. Results - Processing of a TFO-directed interstrand crosslink in human cell extracts	95
4.1. Validation of the cisplatin intrastrand substrate by primer extension assay	95
4.2. NER dual incision control assay was established	98
4.2.1. Major incision products and minor incision products were observed as shown by previous studies.....	98
4.2.2. A 1,3 d[GpTpG] cisplatin intrastrand crosslink is repaired by HeLa whole cell extract as observed by previous studies.....	101
4.2.3. Incisions on the 1,3 d[GpTpG] cisplatin intrastrand crosslink were dependent on NER protein	101
4.3. TFO-directed Psoralen interstrand crosslink was constructed and purified	104

4.3.1. 50 bp psoralen-TFO ICL insert was constructed and purified	104
4.3.2. Vector with BbsI sites was prepared.....	108
4.3.3. Vector and psoralen-TFO ICL insert was successfully ligated in small and then in large scale.....	108
4.3.4. Confirmation of the ICL in the final product	109
4.4. Plasmid relaxation assay.....	114
4.5. Incisions were observed on the 3' side of psoralen-TFO ICL on both strands.....	119
4.6. Psoralen-TFO ICL was not repaired by HeLa whole cell extract.....	120
Chapter 5. Discussion	124
5.1. Consequences of XPA disruption in human cells.....	124
5.1.1. Expression of a subset of genes is influenced by XPA status in cell lines.....	124
5.1.2. Defective mitophagy associated with XPA deficiency and relation to neurological deficits	127
5.1.3. The most significant XPA-regulated genes may explain steroid and sexual degeneration issues in xeroderma pigmentosum group A	128
5.1.4. Multiple non-isogenic cell lines improve the analysis of RNA-Seq data	130
5.1.5. Possible additional role of XPA may not be related to NER	130
5.1.6. Cause of XPA instability in KO142	131
5.1.7. Future directions	133
5.1.7.1. Validation of RNA-Seq results by immunoblotting and CHIP	133

5.1.7.2. Analyzing differentially expressed genes in XPA deficient mice	134
5.1.7.3. Biochemical assays for mitochondrial and metabolic functions in XPA deficient and proficient cells	134
5.1.7.4. Analyzing brain tissues of XPA deficient mice and patients with neurological diseases.....	141
5.2. Processing of TFO-directed psoralen interstrand crosslink in human cell extracts.....	141
5.2.1. Psoralen-TFO ICL or TFO as an adduct for NER.....	141
5.2.3. TFO in targeted DNA damage and gene alteration	145
5.2.4. ICL substrate in a closed circular double stranded DNA	146
5.2.5. Future directions	149
5.2.5.1. Incision assay with NER deficient cells complementation with purified XPA protein	149
5.2.5.2. Detection of 5' side incisions.....	149
5.2.5.3. Incision assays with or without TFO.....	150
5.2.5.4. Involvement of NEIL3 in the processing of psoralen-TFO ICL..	151
References.....	153
Vita	178

List of Illustrations

FIGURE 1. CHEMICAL STRUCTURE OF PSORALEN DNA ICL AND SPACE FILLING MODEL OF A TFO-PSORALEN DNA ICL.....	5
FIGURE 2. DIAGRAM OF NER IN MAMMALIAN CELLS SHOWING STEPWISE EVENTS. ...	11
FIGURE 3: DIAGRAM OF XPA WITH THE INTERACTION WITH OTHER PROTEINS	19
FIGURE 4. XPA DISRUPTION IN PROFICIENT AND DEFICIENT CELL LINES USED IN THE STUDY.	22
FIGURE 5. SINGLE STRANDED AND DOUBLE STRANDED CLOSED CIRCULAR DNA DERIVED FROM M13 BACTERIOPHAGE	32
FIGURE 6. VERIFICATION OF 1,3 GTG CISPLATIN PRIMER AND THE CISPLATIN ADDUCT IN M13MP18GTG DOUBLE STRANDED CIRCULAR DNA.	36
FIGURE 7. XPA EXPRESSION IN THE CELL LINES USED IN THIS STUDY.	51
FIGURE 8. XPA EXPRESSION IN THE CELL LINES USED IN THIS STUDY BY IMMUNOBLOTTING.	52
FIGURE 9. UVC SENSITIVITY AND NER COMPLEMENTATION IN THE CELL LINES USED IN THIS STUDY.	53
FIGURE 10. PROTEASOME INHIBITION ASSAY	55
FIGURE 11. XPA IS A NUCLEAR PROTEIN AND IS RESISTANT TO EXTRACTION FROM THE NUCLEUS DURING DNA REPAIR.	57
FIGURE 12. XPA DEPENDENT CHANGE IN GENE EXPRESSION PATTERN AND BIOLOGICAL PATHWAYS INFLUENCED BY XPA STATUS.	63
FIGURE 13. LOCATION OF <i>AKR1C1</i> , <i>AKR1C2</i> AND <i>AKR1C3</i> ON HUMAN CHROMOSOME 10.	80
FIGURE 14. ANALYSIS OF THE MOST DRASTICALLY CHANGED GENE EXPRESSIONS IN XPA PROFICIENT COMPARED TO THE DEFICIENT CELL LINES.	81

FIGURE 15. RETINOIC ACID TRANSACTIVATION AFFECTS XPA-DEPENDENT TRANSCRIPTION RELATED FUNCTIONS.	93
FIGURE 16. PRIMER EXTENSION ASSAY USING A FRAGMENT OF M13MP18GTGX FROM THE 5' END OF THE RADIOLABELLED PRIMER TO THE 3' END CUT BY PvuI ENZYME.	97
FIGURE 17. NER DUAL INCISION CONTROL-ASSAY.	100
FIGURE 18. REPAIR SYNTHESIS ASSAY.....	103
FIGURE 19. STRATEGY FOR THE CONSTRUCTION OF TFO- DIRECTED PSORALEN ICL IN CLOSED CIRCULAR DOUBLE STRANDED DNA.	106
FIGURE 20. 50 BP TFO DIRECTED PSORALEN ICL.	107
FIGURE 21. VECTOR WITH <i>Bbs</i> I SITES.	110
FIGURE 22. SMALL SCALE LIGATION REACTION WITH BBSI-VECTOR AND 50 BP PSO- TFO ICL.	111
FIGURE 23. LARGE-SCALE LIGATION REACTION OF BBSI-VECTOR AND 50 BP PSORALEN-TFO ICL INSERT.	112
FIGURE 24. CONFIRMATION OF THE FINAL PURIFIED TFO- DIRECTED PSORALEN ICL IN A CLOSED CIRCULAR DUPLEX DNA.....	113
FIGURE 25. PLASMID RELAXATION ASSAY.	115
FIGURE 26. 3' INCISION ASSAY ON PSORALEN-TFO ICL SUBSTRATE.	116
FIGURE 27. SEQUENCE-MAP OF THE INCISION SITES.	117
FIGURE 28. REPAIR SYNTHESIS ASSAY 1.....	122
FIGURE 29. REPAIR SYNTHESIS ASSAY 2.....	123
FIGURE 30. MODEL: POSSIBLE MECHANISM BY WHICH XPA IS INVOLVED IN TRANSCRIPTION.....	132
FIGURE 31. MITOCHONDRIAL ASSAYS.	139
FIGURE 32. MODEL: POSSIBLE MECHANISMS OF PSORALEN-TFO ICL REPAIR IN HUMAN CELLS.....	148

FIGURE 33. STRATEGY TO DETECT INCISIONS AT THE 5' SIDE OF THE ICL ON THE	
UPPER STRAND.....	152

List of Tables

TABLE 1. SHORT TANDEM REPEAT (STR) FINGERPRINTING REPORT OF CELL LINES USED IN THE STUDY	50
TABLE 2. SUMMARY OF GENES ANALYZED BY HIGH-THROUGHPUT RNA SEQUENCING (RNA-SEQ) AND SIGNIFICANT XPA-RELATED DIFFERENCES.	61
TABLE 3. PANTHER PATHWAYS CREATED USING GENE ONTOLOGY (GO) DATABASE FOR COMMON GENES WITH SIMILAR EXPRESSION PATTERN AT FDR0.05 IN XPA- VS. XPA+ COMPARISON IN ALL FOUR PAIRS OF CELL LINES USED IN THE STUDY. ...	70
TABLE 4. GENE ONTOLOGY (GO) TERMS FOR GENES FOR 'MITOPHAGY' OR 'MITOCHONDRIA' OVERLAPPING WITH 2 FOLD OR MORE CHANGED (FC2) GENES IN XPA- VS. XPA+ COMPARISON IN ALL FOUR CELL LINES.	73
TABLE 5. COMMON GENES THAT ARE THE MOST DIFFERENTIALLY EXPRESSED IN ALL FOUR PAIRS OF CELL LINES.	77
TABLE 6. MORE COMMONALITIES WHEN FIBROBLASTS AND HeLA CELLS ARE CONSIDERED INDEPENDENTLY.	82
TABLE 7. CELL SPECIFIC ANALYSIS OF BIOLOGICAL PATHWAYS WITH MORE NUMBER OF GENES REPRESENTED IN EACH PATHWAY COMPARED TO ANALYSIS WITH ALL CELL LINES.	88
TABLE 8. CELL SPECIFIC ANALYSIS OF BIOLOGICAL PATHWAYS SHOWING DIFFERENCE IN HeLA AND FIBROBLASTS.	91

Abbreviations

AKR1C1: Aldo-Keto Reductase Family 1, Member C1

AKR1C2: Aldo-Keto Reductase Family 1, Member C2

AKR1C3: Aldo-Keto Reductase Family 1, Member C3

APIM: AlkB Homolog PCNA Interacting Motif

ATR: Ataxia Telangiectasia and Rad3 related

CAK: cdk Activating Kinase

CPD: Cyclobutane Pyrimidine Dimers

CSA: Cockayne Syndrome complementation group A

CSB: Cockayne Syndrome complementation group B

DMEM: Dulbecco's Modified Eagle Medium

DMSO: Dimethyl Sulfoxide

EME1: Essential Meiotic structure-specific Endonuclease 1

ERCC1: Excision Repair Cross-Complementing group 1

FA: Fanconi Anemia

FAAP24: Fanconi Anemia core complex Associated Protein 24

FAN1: FANCD2/FANCI-Associated Nuclease 1

FANCI: Fanconi Anemia Complementation group I

FANCM: Fanconi Anemia Complementation group M

FC: Fold Change

FDR: False Discovery Rate

GG-NER: Global Genomic -NER

GO: Gene Ontology

GSEA: Gene Set Enrichment Analysis

HBSS: Hank's Balanced Salt Solution

HEPES: 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid

HMGCS2: 3-hydroxy-3-methylglutaryl-CoA synthase 2

HMT: 4'-Hydroxymethyl-4,5'-8-trimethylpsoralen

ICL: Interstrand Crosslink

IPA: Ingenuity Pathway Analysis

MMC: Mitomycin C

NER: Nucleotide Excision Repair

NGS: Next Generation Sequencing

NLS: Nuclear Localization Sequence

PARP: Poly(ADP-ribose) Polymerase

PBS: Phosphate Buffered Saline

PCNA: Proliferating Cell Nuclear Antigen

PPAR: Peroxisome proliferator-Activated Receptor

PVDF: Polyvinylidene Fluoride

RA: Retinoic Acid

RARB: Retinoic Acid Receptor beta

RPA: Replication Protein A

RPMI: Roswell Park Memorial Institute medium

TBB: Triplex Binding Buffer

TBE: Tris-Borate-EDTA

TBS-T: Tris buffered saline-Tween 20

TC-NER: Transcription Coupled- NER

TFO: Triplex Forming Oligonucleotide

TLS: Translesion Synthesis

XPA: xeroderma pigmentosum complementation group A

XPC: xeroderma pigmentosum complementation group C

XPF: xeroderma pigmentosum complementation group F

Chapter 1. Introduction

1.1. DNA-interstrand crosslink in mammalian cells

DNA-interstrand crosslinks (ICLs) form a bridge between the two strands of the DNA double helix, blocking transcription and DNA replication (Figure 1). ICLs are formed by exogenous sources such as mitomycin C (MMC), nitrogen mustard, cisplatin compounds and natural compounds such as psoralen. Endogenous metabolites also form ICLs, including nitric oxide and the lipid peroxidation products acrolein, crotonaldehyde and malondialdehyde (1, 2). ICL-inducing agents are among the most genotoxic and mutagenic compounds as compared to agents that introduce only monoadducts (3). ICL-causing nitrogen mustards such as cyclophosphamide and melphalan are widely used for the treatment of leukemia as well as solid tumors. Similarly mitomycin C, cisplatin and carboplatin are used for treating testicular and ovarian cancers and psoralen is used to treat cutaneous T cell lymphoma (2).

However, it is not fully understood how these ICLs are processed in mammalian cells. In replicating cells, the Fanconi anemia pathway plays major role in the ICL repair. The FA complex consists many proteins and 8 of them (FANCA, B,C,E,F,G,L and M) form a core complex. Ubiquitination of FANCD2 and FANCI by the core complex occurs upon encountering ICL blockage, which activates the repair process (4). Initially FANCM makes a complex with FAAP24

protein that causes the activation of the recruitment of other FA proteins to the damage site. These proteins form a complex and stabilize the replication fork, which activates ATR mediated cell cycle signalling (5-7). Monoubiquitination of FANCD2 is the key for the recruitment of nucleases such as FAN1 and SLX4 (8, 9). ERCC1-XPF complex and MUS81-EME is also assembled for the processing of the ends and the unhooking of the crosslink (6). The unhooked structure is cleaved by NER (10, 11) and the double strand break is repaired through homologous recombination (HR) (12, 13). Translesion synthesis (TLS) polymerases like Rev1-Rev3L are involved in NER process after unhooked structures are removed (14) and FA complex is necessary for their recruitment (15, 16).

There is also evidence that nucleotide excision repair (NER) is an important repair pathway for ICL in nonreplicating cells. NER dual incision is well studied in *E. coli*. In mammalian cells, NER deficient cells are sensitive to ICL inducing agents (10, 17, 18) . Also, ICL repair defects have been shown in NER defective cells by various researchers. ICL formed by psoralen could not be removed in XPA deficient cells in a gene specific assay (19). XPA and XPF defective cells were defective in unhooking ICL formed by psoralen (18) and formed by cisplatin in XPG defective cells (20).

Incision is an important step in the processing of ICL as it determines how ICL will be being processed. Some studies have been done focusing on incisions

around ICLs. Bessho et al (1997) found that NER in human cell extracts introduced two incisions separated by about 21 nt on the 5' side of a psoralen ICL (21). Smeaton et al. (2008) used a crosslink placed in a 150-mer linear duplex and verified the presence of these futile incisions (22). They also suggested that strong incisions were present about 1-5nt on the 5' side and 3-5nt on the 3' side of an ICL. The origin of these incisions is not known. Either they are artifacts produced by unsuitable ICL substrates for incision assays or if they are real, they should be further studied to find their role in the processing of an ICL. R. Legerski found a repair patch in a circular psoralen ICL substrate that runs from 5' side to 3' side of the ICL suggesting a dual incision (23). Thus a very controlled and precise mapping study of ICL repair in mammalian cells is necessary in order to fully understand the mechanism of ICL processing. This understanding will allow the manipulation of different steps of processing of an ICL.

ICL forming agents have been effective in treating some forms of cancer, but they are not so useful in others. In addition, resistance to ICL forming chemotherapies is a major problem for treatment of cancer. This problem could be solved to a greater extent, if we understood the details of processing of ICL. Study of the repair of a site-specific ICL is a controlled way of exactly mapping out how cells process ICL damage.

A psoralen ICL was studied here, targeted by a triplex-forming oligonucleotide (TFO). In addition to directing psoralen to a specific site of DNA, TFOs have a very important therapeutic potentiality. TFOs can be used to targeted gene alterations. For example, TFO mediated correction of AID deficient human lymphocytes and p53 mutant glioblastoma cells have been demonstrated (24). There are also studies showing that a TFO might enhance the recruitment of repair proteins. In *E. coli*, a TFO-directed psoralen ICL is an NER substrate and leads to significantly more incision products than psoralen by itself (25). NER was indicated to be involved in TFO mediated homologous exchange, and sequence specific mutagenesis was shown to be dependent on the NER protein XPA (26, 27).

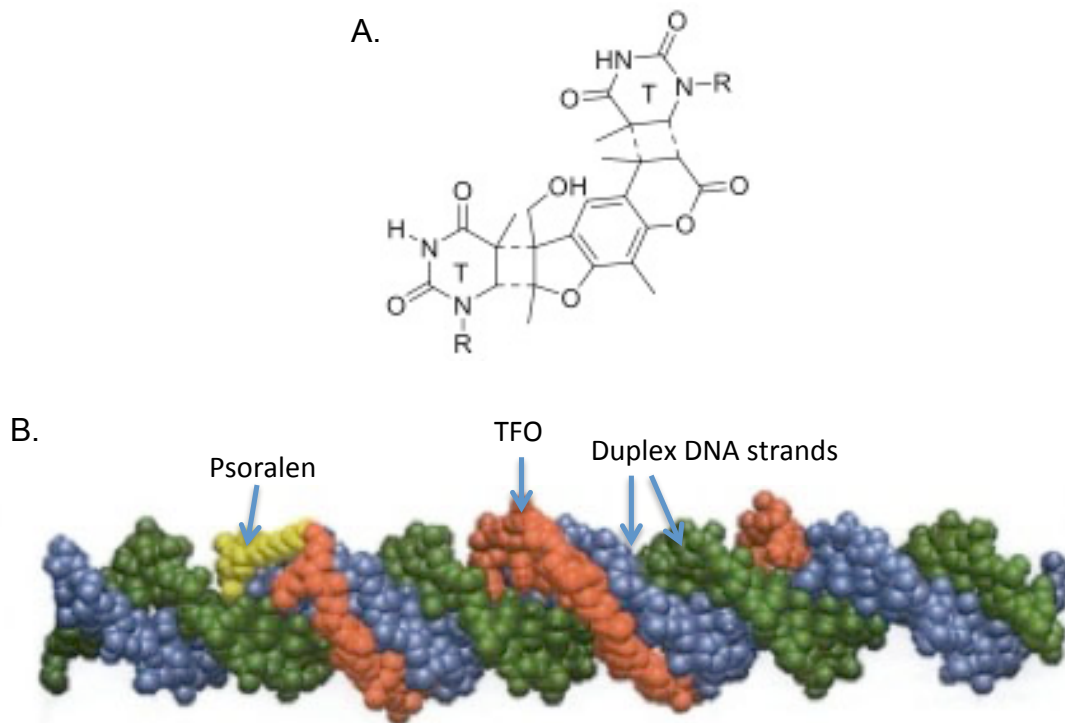


FIGURE 1. Chemical structure of psoralen DNA ICL and space filling model of a TFO-psoralen DNA ICL

A. Structure of an interstrand crosslink formed by HMT psoralen on thymines (T) on both strands of DNA. The furan ring of psoralen binds to one thymine and the pyrone ring of psoralen binds to a thymine on the complementary DNA strand. B. Space filling model of TFO binding to duplex DNA in the major groove. Yellow indicates psoralen bound at 5' end of the TFO, green and blue indicate two strands of DNA.

FIGURE 1B is adapted from Christensen LA, Wang H, Van Houten B, Vasquez KM: Efficient processing of TFO-directed psoralen DNA interstrand crosslinks by the UvrABC nuclease. Nucleic Acids Res 2008, 36(22): 7136-7145.

Permission has been acquired from the journal to use this figure.

1.2. Nucleotide excision repair pathway and XPA protein

Nucleotide excision repair is the DNA repair pathway for a wide range of DNA lesions. The lesions include pyrimidine dimers formed by UV radiation, and bulky adducts formed by many chemicals and some endogenous metabolites (28, 29). NER has two sub-pathways: Global genomic NER (GG-NER) and transcription coupled NER (TC- NER). The difference in the two pathways lies in the recognition of the damage. GG-NER recognizes the distortion in the helix of DNA in all parts the genome. TC-NER is activated when RNA polymerase II is stalled during the elongation of transcript due to the presence of DNA damage on the template strand. The overall function of mammalian NER is to precisely remove the damaged nucleotides as a 24-32 base long single-stranded oligonucleotide and then to restore the DNA with undamaged nucleotides. About 30 polypeptides are involved in the whole process of NER, which make successive complexes necessary for damage recognition, incision and restoration (28, 30, 31) (Figure 2).

In GG-NER, the initial damage recognition is done by protein complex called XPC-RAD23B (32). XPC is sequestered by damaged DNA and its binding to damaged DNA triggers the further repair process (32). RAD23B enhances the DNA binding activity of XPC, stabilizes the complex and protects it from proteasome degradation (28). It does not come in direct contact with DNA and is

not immobilized on the damaged DNA but it is released after XPC binds to the DNA (33). An N-terminal region (amino acids 154-334) of XPC binds to XPA (34).

After DNA distortion is recognized, TFIIH is recruited to the damaged site. TFIIH is a ten subunit complex consisting of proteins XPB, XPD, p62, p44, p34, p52, p8 and a cdk activating kinase (CAK) with subunits Mat1, Cdk7 and CyclinH. TFIIH is a general transcription factor but also acts as a core NER factor for DNA repair. The XPB subunit of TFIIH is a 3'→5' DNA helicase and XPD is a 5'→3' helicase. TFIIH unwinds DNA at the damaged site to form a single stranded region, in an ATP dependent manner (35, 36). RPA and XPA interact with TFIIH in the pre-incision complex. RPA is a single stranded DNA binding protein and consists of 3 subunits with multiple cellular functions (37). RPA helps to form a stable pre-incision bubble with four functional DNA binding OB (oligonucleotide/saccharide binding)-folds present in its p70 and p32 subunits. These interact with DNA in a polar manner (5'→3') (38, 39). They occupy 30 nucleotides on the undamaged strand and thus protect this strand from degradation by nucleases as well as facilitate the precise positioning of the endonucleases at 5' and 3' side of the adduct.

XPA protein has increased affinity towards DNA structures such as distortion of the helix due to damage of DNA. Thus XPA is part of the confirmation process that detects the structural DNA distortion caused by DNA damage (40). It directly interacts with damaged DNA (41, 42) at the junction of

single strand and double strand (43). XPA is recruited to the damaged site by TFIIH complex through its p8 and p52 subunits (44, 45). Interestingly XPA is also reported to interact with another transcription factor TFIIIE, however the function of the interaction is not clear (46). XPA is a small protein with 273 amino acids, does not have any enzymatic activity but interacts with other NER proteins playing role as a major scaffolding protein in NER. XPA coordinates with RPA to facilitate the functions of nucleases for NER.

XPA has a globular central domain (98-219) with a C4 type zinc finger motif in the N terminal region and a cleft in the C-terminal region. XPA has many protein binding interaction domains (47-50) (Figure 3). Interestingly *XPA* point mutations associated with xeroderma pigmentosum are principally in the central domain (50, 51). Mutations compromising or eliminating the DNA binding region are associated with the most severe XP syndromes including accelerated aging and neurodegeneration (51). XPA also interacts with XPE at XPA residues 185-226, and the mutation at this site has been reported to cause failure in the removal of cyclobutane pyrimidine dimers (CPD) (52). XPA interacts with RPA2 primarily via XPA residues 29-46 and XPA residues 98-219 secondarily or weakly interact with RPA1 (49, 53). XPA has a nuclear localization (NLS) domain at 30-42 aa region. However it has been shown that when mutation of XPA at this region affects its localization in the nucleus, these residues were not found to be essential for DNA repair (54).

ERCC1-XPF nuclease is recruited to the damage site by interaction of XPA protein amino acids 96-114 and ERCC1 residues 92-119 (48). XPA contains a PCNA binding site called an APIM (*AlkB* homolog PCNA interacting motif) sequence but the functional aspect of the interaction is not clear yet (55). XPA also interacts with other proteins such as ATR and PARP that are not involved in NER. ATR phosphorylates XPA at Ser196. XPA has a PARP binding motif at 213-237. PARP can inhibit the DNA binding activity of XPA (56).

Mutations in XPA cause XP, with a severity that varies according to the particular mutation. A common consequence of XPA mutations is sensitivity to UV radiation, leading to skin abnormalities and a predisposition to skin cancers. The most severe cases result in progressive neurodegeneration and a higher incidence of melanomas (57-59). The reason behind the variation in the severity in the XPA deficient patients is still to be studied. This could be the result of expression of different fragments of XPA protein with different functions. Truncation of Exon 1 or Exon 6 which is thought to be dispensable for NER, have been reported to yield only mild XP syndrome whereas deletion of Exon 2-5 which encodes the DNA binding domain and is essential for NER has been reported to cause a severe form of XP (51, 60, 61).

ERCC1-XPF and XPG are structure specific endonucleases that cleave on the 5' side and 3' side of an adduct on the damaged DNA strand during nucleotide excision repair. ERCC1-XPF protein has multiple functions besides a

role in NER. The endonuclease activity of this complex can act on different substrates including those formed by the Fanconi anemia pathway during repair of ICL-induced DNA damage, 3' ends during homologous recombination and hairpin structures (62), removal of non-homologous ends during non-homologous end joining (63-65) and maintenance of telomeres (62, 66). Disruption of *ERCC1* has severe consequences. Mice with disrupted *ERCC1* die before weaning from liver disorder and other organ failure (67, 68). Mice with a knock out of the *XPG* gene die due to a developmental defect in intestine formation (69) and humans with *XPG* mutations sometimes have a short life-span and developmental defects similar to Cockayne syndrome, indicating that *XPG* has a second function other than in NER (70).

CSA and CSB proteins are essential for transcription coupled NER. Mutation in CSA and or CSB genes cause a severe condition called Cockayne syndrome which is characterized by severe neurodegeneration, developmental defects, accelerated aging (progeroid syndrome) and UV radiation sensitivity (71). The mixed cases of CS/XP are even worse, with both CS and XP symptoms caused by specific mutations in the *XPB*, *XPD* or *XPG* genes. These individuals have an increased photosensitivity and more cancer-proneness as they have defects in both TC-NER and GG-NER (71).

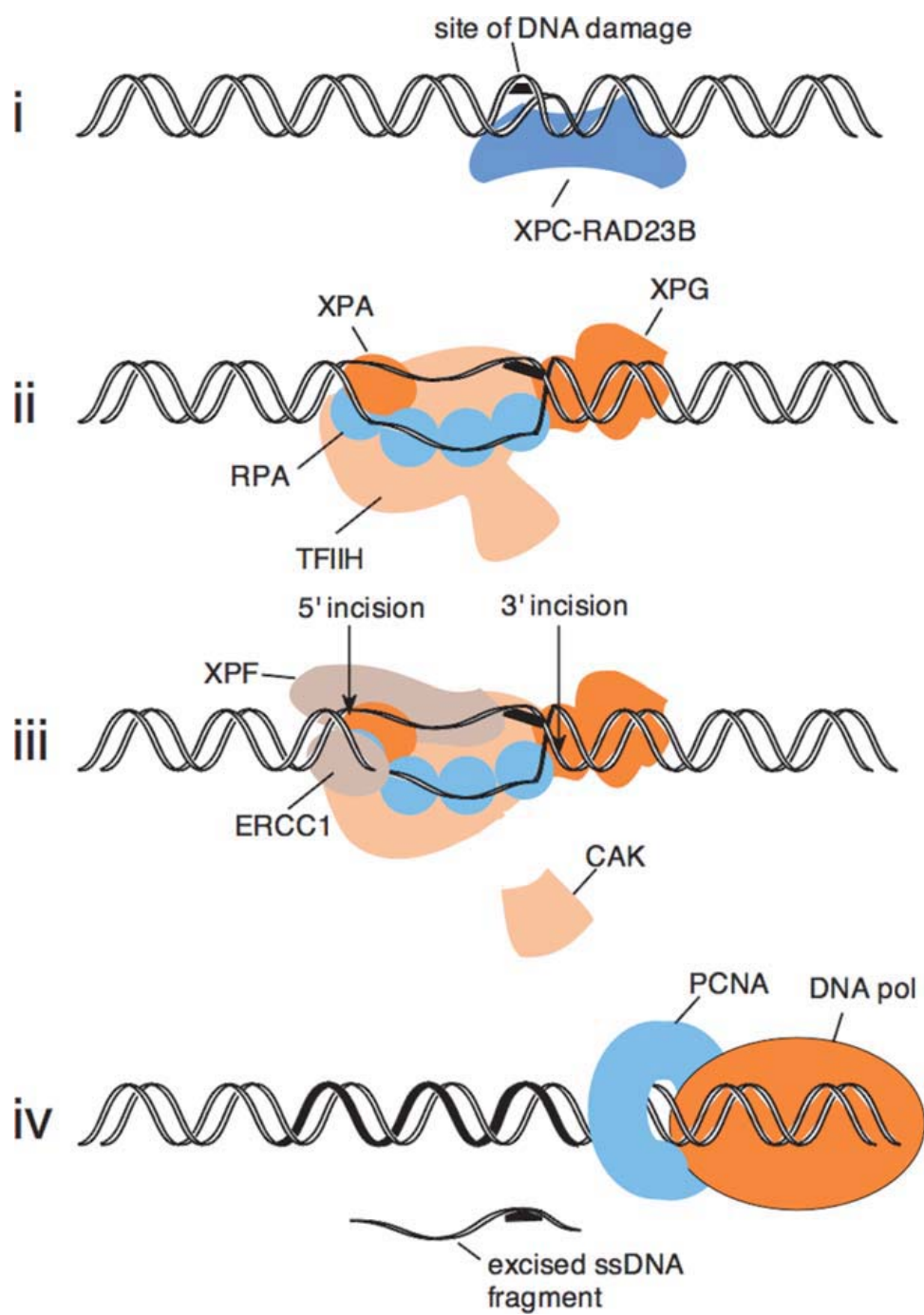


FIGURE 2. Diagram of NER in mammalian cells showing stepwise events.

This diagram shows the GG-NER pathway. (i) XPC-RAD23B protein complex recognizes the distortion of DNA helical structure caused by the lesion. TFIIH is recruited by XPC-RAD23B complex at a site of damage. (ii) The NER proteins XPA, RPA, XPG and ERCC1-XPF are recruited at the damaged site to form a preincision complex together with TFIIH complex. The two DNA strands are separated in an ATP dependent manner to form a bubble (72). (iii) The ERCC1-XPF nuclease makes an incision at the 5' side of the lesion and the XPG nuclease produces the incision at the 3' side. This process of producing incision at both sides of the lesion simultaneously is referred as dual incision. The CDK7-activating (CAK) subunits of TFIIH are detached from the precision complex after it is formed (73). Dual incision releases a fragment of damaged DNA about 27-32 nucleotides long. This is a characteristic of NER in mammalian cells. (iv) A DNA polymerase is loaded by PCNA and fills the gap made by the incisions.

FIGURE 2 is adapted from Wood, R. D. (2010): Mammalian nucleotide excision repair proteins and interstrand crosslink repair. Environ Mol Mutagen 51(6): 520-526

Permission has been acquired from the journal to use this figure.

1.3. Transcriptional role of XPA, a major NER protein

Nucleotide excision repair (NER) is the only pathway that mammalian cells have available for the removal of DNA lesions arising from ultraviolet (UV) radiation damage. NER also repairs helix-distorting lesions produced by other types of radiation and some chemicals. In eukaryotes, most of the components of the NER machinery have major additional biological functions that are essential for normal viability. This has profound consequences for cells or organisms with mutations in NER genes. The ten subunits of TFIIH form a core initiation factor for basal transcription of all mRNAs, for example, and XPG also has a transcription-related function (74). ERCC1-XPF participates in some homologous recombination reactions and in crosslink repair (62). As a consequence, complete disruption of some NER components is incompatible with cellular survival or embryonic development (for example, TFIIH subunits or RPA), while perinatal lethality occurs following disruption of other components (ERCC1, XPF, XPG). Mutations that partially disable these factors can lead to severe diseases in human beings.

Complete disruption of two major NER factors is tolerated, however. One of these is XPC, part of a distortion recognition complex. XPC disruption causes xeroderma pigmentosum in humans, but is relatively less severe than other forms because a transcription-coupled form of NER remains intact. Nevertheless,

even XPC is reported to play a dual function as a component of a transcription complex for specific genes (75).

The other NER factor that can be completely inactivated without impairing cell viability is XPA. This is notable, as XPA is absolutely required for NER. It is a scaffold protein that contacts many of the other factors in NER and binds to DNA (Figure 3). As a consequence, xeroderma pigmentosum (XP) group A patients or mice without XPA function have no NER activity. Because each of the other NER factors has an identified additional function in cells, it is important to consider whether XPA also has a significant biological function other than NER. Broadly, a complete XPA defect has no major effect on mammalian development, growth or cellular function. Patients with mutations that completely ablate XPA function are born, develop normally and live for several decades. Similarly, mice with complete XPA defects are born, develop normally, and have a near-normal lifespan, with a low background of spontaneous tumors occurring later in life.

XPA patients often succumb to metastatic UV radiation-induced skin tumors. This can be delayed by protection of individuals from sun exposure. In addition, many XPA patients with loss of XPA function show accelerated neurological deterioration over decades, initially characterized as de Sanctis-Cacchione syndrome (57). Retention of a small amount of XPA function results in much milder symptoms (76). A likely explanation for the origin of the neurological impairment is the accumulation of genomic DNA lesions over decades in non-

replicating neural cells that can only be repaired by NER (77). Prime candidates are the cyclopurines (78) induced by reactive oxygen species, which are repaired by NER and cannot be removed by other repair systems (78, 79). XPA-defective mice, with a much shorter lifespan, do not appear to exhibit these neurological deficits (80).

Nevertheless, a few observations have suggested that XPA might have additional functions beyond NER. At several tested promoters, all NER proteins including XPA can be detected by chromatin immunoprecipitation. The basal transcription factor TFIIH is expected to be present at promoters, and XPA may be detectable because it binds to TFIIH and associated proteins. However at least one result indicated that XPA depletion affects retinoic acid (RA) – induced transcription of some genes (81). Indeed, depletion of XPA reduced RA-activated transcription of *RARB*, *PPAR*, and *HMGCS2*. Moreover, a study comparing XPA proficient and deficient cells by microarray analysis found changes in gene expression associated with XPA status, and indicated that XPA-deficient cells have mitochondrial dysfunction with defects in pathways of mitophagy (82), a condition that could impact neural health. Decline of mitochondrial activity is the hallmark of many neurological diseases, and neurons are the most sensitive to mitochondrial dysfunctioning (83). These studies indicate the possibility of mitochondria being affected when NER proteins are defective. The syndromes observed in NER-deficient cases in humans and mice are mainly due to the inability to repair DNA damage but more studies are necessary to understand the

mechanisms of all the syndromes. An RNA- sequencing study on XPD/CS patient cells, after UV exposure, suggests that the syndromes are not solely due to DNA repair defects but also due to other transcriptional dysfunctioning (84). It is very intriguing to see if a defect in NER proteins causes significant changes in expression of genes in human cells.

Nucleotide excision repair (NER) proteins XPC, XPA, ERCC1-XPF, XPG, TFIIH and CSB have also been reported to play an important role in the regulation of nuclear receptor-dependent genes (81, 84-88). Down regulation of any of these NER factors have shown to decrease the transcription of several genes as well as impeding of the recruitment of the other NER factors at the transcription sites (85). It is interesting that NER proteins play a significant role in modulating the transcription of some genes and raises the question of whether the whole genome is affected.

However, no study has been done in order to pursue genome wide changes in the transcriptome in absence of specific NER proteins in absence of any genotoxic challenge. To specifically investigate the extent of a possible transcriptional defect in XPA-deficient cells, we examined genome wide expression of transcribed genes by RNA-Seq analysis.

Based on previous studies (81, 82, 85-88), we expected that an XPA defect might result in down regulation of genes responsible for mitochondrial,

metabolic and neurological functions. We also transactivated cells with retinoic acid (RA), which is the ligand for receptors including *PPAR* and *RAR*. The products of the genes expressed by the activation of these receptors are involved in many cellular functions such as cell differentiation, cell proliferation, apoptosis and metabolism.

1.4. Localization of XPA protein

As described above, XPA has multiple functions, including a novel transcriptional role, which we are actively pursuing. In order to have a direct role as a transcription factor, XPA would have to be present in the nucleus where transcription takes place. Although XPA is localized in the nucleus under standard fixation conditions, we were concerned about reports that XPA is normally resident in the cytoplasm, and is only recruited to nuclei following UV radiation exposure (89, 90). Notably, in those studies, cells were first briefly exposed to wash buffers containing Triton X-100 detergent or the milder fixation agent methanol (89, 91), or with 0.05% NP-40 detergent (90). It is possible that previous observations suggesting nuclear import of XPA following UV radiation may have been misinterpreted because of nuclear leakage.

Nuclear pores are complex structures composed of 50 -100 different proteins in vertebrates, making large pores through which small proteins less than 50 kDa can passively pass. Nuclear pores have a diameter of about 120 nm

and a molecular weight of about 125 million dalton (Da) (92). Detergents used in different cellular studies can form micelles in the nuclear membrane and make it even easier for the small proteins to leak out of the nucleus. The permeabilization with detergents like NP-40 or Triton-X-100, which are non-selective in nature could lead to false results regarding the localization of the proteins (93). Studies have shown leaking of unassembled proteins from the nucleus to the cytoplasm by detergent extraction (94).

UVC radiation exposure to cells leads to two major kinds of DNA photoproducts (6-4) photoproducts and cyclobutane pyrimidine dimers (CPD) (95, 96). 80% of the adducts are CPDs and 20% are 6-4 Photoproducts (6-4PP) (29, 97). XPA is involved strongly in repair of these photoproducts as NER pathway is the major DNA repair pathway for both kinds of DNA damage. We hypothesized that during the repair process, XPA may be strongly recruited to sites of DNA damage, inhibiting leakage even following the use of detergent. This would give a different interpretation than hypothesizing that XPA is imported to the nuclei after exposure to UV radiation.

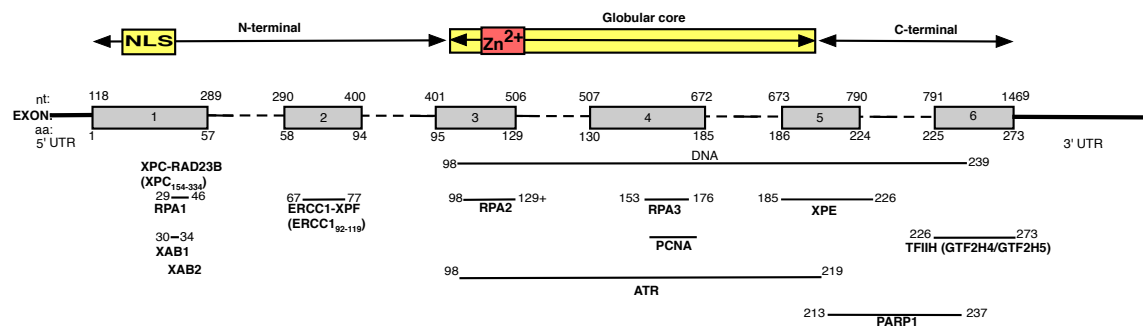


FIGURE 3: Diagram of XPA with the interaction with other proteins

At the top, the diagram shows human XPA protein with location of the central globular core in relation to the six coding exons (gray). The locations of the nuclear localization signal (NLS) and zinc finger (Zn²⁺) are shown. The bottom part of the panel shows mapped regions of XPA that interact with other proteins.

FIGURE 3 is adapted from Sugitani N, Sivley RM, Perry KE, Capra JA, Chazin WJ: XPA key scaffold for human nucleotide excision repair. DNA Repair (Amst) Permission has been acquired from the journal to use this figure.

Chapter 2. Methods and Materials

2.1. Preparation of samples for the next generation sequencing of RNA from XPA deficient and proficient human cell lines

2.1.1. Cell lines

Four human cell lines deficient in XPA and their complementary XPA-proficient cell lines were used. SV40-immortalized human fibroblast cells from patients deficient in XPA (XP12RO and XP2OS) (98) and XPA-proficient (XPA/XP12RO and XPA/XP2OS) cell lines complemented with *XPA* cDNA were kindly provided by Dr. Masafumi Saijo, Osaka University. The mutations in XP12RO and XP2OS are also shown in Figure 4A. Only 10% of our initial culture of XPA/XP12RO cells expressed XPA as shown by an immunofluorescent assay with 12F5 anti-XPA antibody. We recloned XPA/XP12RO obtaining clone 2, which uniformly expresses XPA. All experiments reported here used clone 2 of XPA/XP12RO as the XPA-proficient cell line paired with XP12RO. HeLa S3 cell line was from laboratory stock. *XPA* disrupted HeLa S3 cell lines KO142 and KO38 were derived by CRISPR-Cas9 targeting technology by the Gene Editing and Cellular Model Core Facility at the University of Texas MD Anderson Cancer Center (Figure 4B). XP12RO and XP2OS and the complemented XPA cell lines were grown in DMEM-high glucose (Sigma D5796) and HeLa S3, KO142 and KO38 in RPMI-1640 25 mM HEPES (Sigma R4130) medium. All cells were

grown in medium containing 10% FBS, 1% streptomycin/penicillin solution at 37 °C and 5% CO₂.

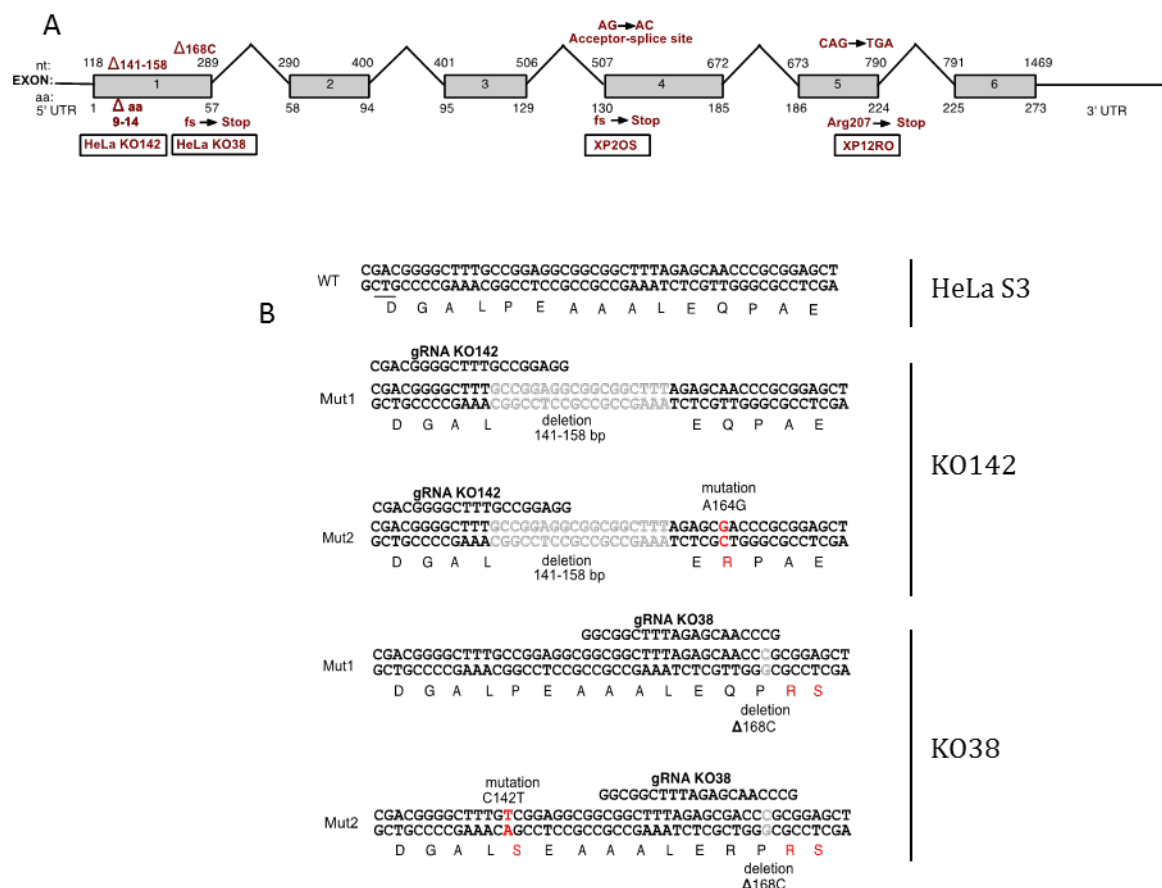


FIGURE 4. XPA disruption in proficient and deficient cell lines used in the study.

A. The exon-intron structure of the human *XPA* gene showing the sites of the causative *XPA* mutation in each *XPA*-deficient cell line used here. B. Location of CRISPR-Cas9 mutations generated in the *XPA* gene of HeLa S3 cells. Two mutated HeLa cell lines were obtained using two guide RNAs (gRNA). KO142 had the indicated 18 bp deletion in both alleles, and an A164G mutation leading to a Q16R amino acid change in one allele. KO38 had a deletion of C168 (ΔC168) leading to early termination in both alleles and a C142T mutation leading to a P9S amino acid change in one allele.

2.1.2. Immunoblotting

Anti-XPA antibodies used were rabbit polyclonal CJ1 (99), mouse monoclonal 12F5, and rabbit polyclonal GTX103168 (Cat. No. GTX103168, GeneTex). Other antibodies used were mouse monoclonal to AKR1C1 (Cat. No. GTX53684, GeneTex), rabbit polyclonal to AKR1C2 (Cat. No. 13035, Cell Signalling Technologies Inc.), rabbit polyclonal to AKR1C3 (Cat. No. ABS1172, EMD Millipore), and mouse or rabbit α -tubulin (Sigma Aldrich). Mouse monoclonal 1E4 (Trevigen) to XPF was a gift of Dr. Laura Niedernhofer. Cell lysates were prepared by lysing 10^6 or 10^7 cells in lysis buffer (500 mM KCl, 20 mM Tris-HCl [pH 8.0], 5 mM $MgCl_2$, 10% glycerol, 1 mM PMSF, 0.1% Tween 20, 10 mM β -mercaptoethanol) by pipetting up and down and then sonicated on ice (45 sec with 15 sec intervals, for 2 times at 1 amplitude). Lysates were then mixed with SDS loading buffer for 3 min at 95 °C before loading onto a 4-20% gradient polyacrylamide gel. Proteins were transferred to a PVDF (polyvinylidene fluoride membrane) for 1 hr in transfer-buffer (Tris-glycine with 20% methanol) cooled on ice. The blots were blocked with 10% non-fat dairy milk for 2 hr at room temperature and washed 3 x 10 min with TBS-T (Tris buffered saline-Tween 20). Membranes were incubation with primary antibodies diluted as follows (CJ1; 1: 10000, 12F5; 1: 5000, GTX103168; 1: 1000, XPF 1E4; 1:2000, AKR1C1; 1:1000, AKR1C2; 1: 500, AKR1C3; 1: 1000, α -tubulin; 1:8000). Membranes were incubated overnight at 4 °C except for tubulin, which was for 20 minutes at room temperature. Then they were washed 3 x 10 min and incubated

with mouse or rabbit IgG secondary antibodies (1:10,000) at room temperature for 2 hr, washed and treated with Pierce Western blot signal enhancer (Thermo Scientific Inc.) and then exposed to x-ray film.

2.1.3. UVC clonogenic survival assay

The UVC clonogenic survival experiments were performed using two 15 W germicidal lamps emitting predominantly 254 nm light adjusted to a fluence of 0.3 J/m²/s. Cells were plated onto 10 cm dishes at densities of 1-5x10⁵ cells per plate and incubated overnight prior to treatment. The plates were swabbed around their perimeters with sterile cotton swabs, growth media aspirated, the attached cells rinsed once with sterile phosphate buffered saline without magnesium or calcium (PBS) and irradiated with lids off through 4 ml of PBS. Following treatment, the PBS was aspirated and cells were detached by trypsin-EDTA treatment, passed through a 23 gauge needle, diluted, counted using a Beckman Coulter Particle Counter Z1 Single, and plated at the appropriate cell number for the given dose, in triplicate, into 10 cm plates containing 10 mL of growth medium. Following incubation for 12-17 days, colonies were rinsed with PBS, fixed with methanol, stained with 0.5% crystal violet in ethanol and counted manually. The original plating counts and colony counts were used to determine the plating efficiency and the % survival.

2.1.4. NER extracts and assays

Whole cell extracts of HeLa S3 and KO38 were prepared as described (100, 101). A protein concentration of at least 20 mg/ml was obtained in each extract. A single 1,3 (dGpTpG) cisplatin intrastrand adduct substrate was prepared and tested in an NER dual incision assay using the end-labelling method (102, 103). Briefly, the reaction mixture (10 μ l) consisted of 50 ng of Pt-GTG or control substrate, 40-45 μ g of whole cell extract protein, in the NER buffer with final concentration of 45 mM HEPES-KOH (pH 7.8), 70 mM KCl, 7.5 mM MgCl_2 , 0.9 mM DTT, 0.4 mM EDTA, 2 mM ATP, 2.5 μ g of creatine phosphokinase, 3.4 % glycerol and 18 μ g of bovine serum albumin. An oligonucleotide containing G-overhang was used to detect the dual incision product by incorporating complementary [α - ^{32}P] dCTP. NER dual incision complementation was done using purified XPA protein. For XPA purification, cell growth, induction, and lysis was performed as described (104). The sonicate was spun down at 15,000 x g, 30 min at 4 $^{\circ}\text{C}$. Supernatant was incubated with HIS Select Nickel Affinity Gel (Sigma) for 3 hr at 4 $^{\circ}\text{C}$. The protein was eluted by gravity flow with buffer containing 50 mM sodium phosphate (pH 8.0), 100 mM KCl, 0.01% NP 40, 1 mM EDTA, 10% glycerol, and 250 mM imidazole. Fractions were pooled and the buffer was exchanged to 25 mM HEPES KOH (pH 7.8), 10% glycerol, 0.4 mM EDTA, 1 mM DTT, and 0.5 M KCl.

2.1.5. Immunofluorescence assay

Cells were washed 3 x 3 min with phosphate buffer saline (PBS) followed by fixation with 4% formaldehyde for 10 min at room temperature. The fixed cells were washed twice with 0.2% Triton-X 100 in PBS at room temperature for 3 min followed by washing twice with PBS for 3 min. The cells were blocked with 5% normal goat serum for 30 min at room temperature followed by two washes with PBS for 3 min each. Incubation with 1:10,000 dilution of anti-XPA antibody 12F5 was done overnight at 4 °C on a shaker. Followed by a wash with PBS and incubation with goat anti-mouse Alexa-Fluor 488 labelled IgG at a 1: 2000 dilution at room temperature for 2 hr, washed with PBS, incubated 10 min with 1 mg/ml DAPI (4',6-Diamidino-2-Phenylindole, Dihydrochloride from Molecular Probes) and washed with PBS. The slides were covered with mounting media and then coverslips.

2.1.6. XPA-extraction immunofluorescence assay

In experiments observing extraction of XPA from the nucleus into the cytoplasm, the cells were exposed to 0.05% NP-40 detergent in PBS for 10 sec twice and fixed immediately with 4% formaldehyde for 10 min at room temperature followed by two washes of 3 min each with PBS. The fixed cells were blocked with 5% goat serum and staining was accomplished as described above for normal immunofluorescence.

2.1.7. UVC exposure

40,000 cells were plated on 18 mm x 18 mm coverslips in 35 mm diameter petri dishes. After overnight incubation, the plates were washed twice with PBS for 3 min each time and exposed to UVC for 5 J/m², 20 J/m² or 60 J/m². Medium was added back to the cells and incubated for 1 hr at 37 °C. Immunofluorescent analysis was carried out as described above.

2.1.8. PUVA treatment

40,000 cells were plated on 18 mm x 18 mm coverslips in 35 mm diameter petri dishes at 37 °C overnight. Cells were washed twice with Hank's balanced salt solution (HBSS) and incubated with 10 µM of 4'-Hydroxymethyl-4,5'-8-trimethylpsoralen (HMT) in HBSS or HBSS alone for 30 min at 37 °C. The cells were exposed to UVA for 30 min at 10 J/m²/s and washed with HBSS to remove free psoralen. A second UVA dose of 18 kJ/m², 36 kJ/m², 54 kJ/m² or 10.8 kJ/m² was delivered to the cells on plates on ice. Media was placed back on the plates and incubation continued for 1 hr at 37 °C. Cells treated with only psoralen were kept in HBSS for the same time as for the corresponding PUVA treatments. Immunofluorescent analysis was performed as described above using the flourochromes A4 FLUO (blue) for DAPI and GFP FLUO (green) for XPA. The images were acquired with a Leica DMI6000 microscope and objective lenses

(magnification/numerical aperture) HCPLAN APO 10X/0.03 or HCX PL FLUOTAR 20X/0.07. Dry imaging medium was used at room temperature. A Leica DFC360 FX camera captured the images and Leica Application Suite X (LAS X) 1.1.0.12420 1997-2014 version of software was used.

2.2. Transactivation and NGS

10^6 fibroblast cells were cultured 24 to 48 hr or until 90% confluent. 2×10^6 cells were treated with either 10 μ M all-trans retinoic acid (RA) or an equal volume of dimethyl sulfoxide (DMSO) for different periods. HeLa S3 were cultured in RPMI (Thermo Fisher Scientific) media with 10% FBS and 1% penicillin/streptomycin solution. HeLa cell lines were treated with 1 μ M RA or an equal volume of DMSO. The final concentration of DMSO was 0.01% in the cell culture. The fibroblast cells were harvested with a cell scraper. Each sample was centrifuged and washed with phosphate buffer saline (PBS) once, centrifuging at 1500 rpm (485 x g) for 5 min at 4 °C. The pellets were quick-frozen and stored at -80 °C. RNA was extracted from pellets using a QIAGEN RNA isolation kit for a qPCR assay. The qPCR data were analyzed for fold-change in mRNA level of the human *RARB* gene at different time points. A time point of maximum response was determined to prepare the samples for RNA-Seq analysis. Cells were treated as described above and cell pellets were submitted for Next Generation sequencing (NGS).

2.3. RNA Sequencing

Sequencing: The libraries were prepared using the Illumina TruSeq stranded mRNA kit according to the manufacturer's protocol, except that the number of PCR amplification cycles was reduced to 8. The libraries were loaded on cBot (Illumina, San Diego, CA) at a final concentration of 10 pM to perform cluster generation, followed by 2 x 76 bp paired end sequencing on a HiSeq 2500 instrument (Illumina). A total of 48 libraries (three biological replicates per condition) were sequenced in 6 lanes, generating 16-41 million pairs of reads per sample. Each pair of reads represents a cDNA fragment from the library.

Mapping: The reads were mapped to the human genome (hg19) using TopHat (version 2.0.10) (105). The overall mapping rate of reads was 87-96%. 76-93% fragments, depending on the experiment, had both ends mapped to human genome.

Differential Expression: The number of fragments in each gene from the RefSeq database (106) (downloaded from UCSC Genome Browser on July 17, 2015) was enumerated using HTSeq-count from the HTSeq software package (version 0.6.0) (107). Genes with less than 10 fragments in all samples were removed before differential expression analysis. The differential expression between conditions was statistically assessed using the R/Bioconductor package DESeq (version 1.18.0, for DMSO vs. DMSO) (108) and edgeR (version 3.8.6,

for DMSO vs. RA) (109). Differential expression of genes with FDR (false discovery rate) ≤ 0.05 was considered significant and a fold change (FC) ≥ 1.5 was used for more rigorous analyses.

2.4. Analysis of NGS data

NGS data were curated to identify the significant (FDR 0.05) differentially expressed genes in XPA-proficient cells compared to their respective deficient cells, and genes with fold change levels >1.5 . Gene set enrichment analysis, pathway analysis and statistical gene enrichment analysis were performed using Gene Ontology (GO) Consortium (110), Gene Set Enrichment Analysis (GSEA) (111) and QIAGEN's IPA programs using GO as the data source. Bonferroni's correction was used wherever possible when using GO consortium for pathway and gene enrichment or over-representation test analysis.

2.5. Preparation of components needed to establish the NER dual incision control assay

2.5.1. M13mp18GTG single stranded DNA

Single stranded circular DNA of M13 bacteriophage (Mp13m18GTG) was needed to make the double stranded circular DNA containing a cisplatin adduct, to use as substrate for nucleotide excision repair assays (102). M13mp18GTG

single stranded circular DNA from M13 bacteriophage was isolated by infecting a culture of *E. coli* strain JM109, with the bacteriophage, grown on M9 minimal selecting media supplemented with 0.1% thiamine. 0.5 ml of bacteriophage stock (10^{12} pfu/ml) was incubated at room temperature for 5 minutes in 2.5 ml of fresh culture of JM109. The mixture was added to 500 ml of Luria Broth (LB) with 5 mM $MgCl_2$ and incubated for 5 hr with vigorous shaking at 37 °C. The culture broth was centrifuged at 700 x g for 15 minutes to pellet the bacteriophage-infected JM109. Replicative form (RF) of M13mp18GTG, which is a double stranded circular DNA, was extracted from the pellet. The phage from the supernatant was precipitated using polyethylene glycol (PEG) and the single stranded circular DNA was extracted by the phenol chloroform extraction method. Single stranded DNA was obtained with a concentration of 1.6-2 $\mu g/\mu l$. The purity of single and double stranded DNA was analyzed on 0.8% agarose gel run in presence of ethidium bromide at 20 mA for 5 hours (Figure 5).



FIGURE 5. Single stranded and double stranded closed circular DNA derived from M13 bacteriophage

0.8% agarose gel with ethidium bromide (40 µg/ 100 mL) run in Tris-Borate-EDTA (TBE) buffer with ethidium bromide at 20 mA for 5 hours. Single stranded (SSDNA) and double stranded (DSDNA) derived from M13 single stranded DNA are used as controls to confirm the newly purified single stranded DNA (M13mp18GTG) as indicated at the top of the gel.

2.5.2. Whole cell extracts

Whole cell extracts of HeLa S3 and KO38 were prepared following Biggerstaff and Wood's protocol (100, 112). A high protein concentration of about 15-20 mg/ml is necessary for the repair assays to take place efficiently in vitro. The cell extracts are made from 10^9 cells in 2L or 3L cultures to assure a high protein concentration. The cells were pelleted and incubated in hypotonic lysis buffer for 30 min followed by homogenization. Protease inhibitors were added to inhibit proteolysis. Protein was precipitated and concentrated using sucrose-glycerol buffer and ammonium sulfate in 4°C. High-speed centrifugation at 213000 x g at 4°C for 3 hr was done to separate DNA and other impurities from protein. Supernatant was separated carefully and dialyzed for 8 hr and 2 hr in dialysis buffer consisting of 0.1 M KCl, 12 mM MgCl₂, 1 mM EDTA, 17% glycerol, 2 mM DTT and 25 mM HEPES at pH 7.8. The whole cell extracts were quick frozen and stored at -80°C in aliquots to avoid repeated freezing and thawing.

2.5.3. Double stranded circular DNA with intrastrand cisplatin adduct

A primer containing a cisplatin adduct for synthesis of double stranded circular DNA was prepared as described (102). Platination of the 24-mer oligonucleotide (TCT TCT TCT GTG CAC TCT TCT TCT) was achieved by incubating a 3:1 molar ratio of oligonucleotide and cisplatin in platination buffer (3 mM NaCl, 0.5 mM Na₂HPO₄ and 0.5 mM NaH₂PO₄) for 16 hr. The reaction

mixture was run on 20% polyacrylamide sequencing gel and gel-purified by the crush and soak method. In this method, an excised piece of gel is crushed and soaked in nuclease free water and incubated at 37 °C in a high-speed shaker (250 rpm). DNA is recovered from the water by removing the gel pieces and then concentrating DNA in a speed vacuum concentrator. To check if the DNA obtained was modified or unmodified, DNA samples obtained from the gel purification were radiolabeled with [γ -³²P] deoxyadenosine triphosphates, run on a 20% polyacrylamide sequencing gel and exposed to X-ray film for 5 minutes. The unmodified and modified DNA was observed differing in migration by about 1 nucleotide (Figure 6A).

The 24-mer oligonucleotide with a cisplatin adduct was used as primer to make double stranded circular DNA. The primer (100 μ M) was annealed with single stranded circular DNA derived from M13 bacteriophage (25 μ g) in NEB Buffer 2 in a thermocycler at 65 °C for 5 minutes and gradually cooled to 4°C over the period of 75 minutes. The complementary strand synthesis was done using annealed DNA with 8 μ g/ml RB69gp43 DNA polymerase in a reaction mixture containing 3.75 μ l NEB buffer 2, 5 mg/ml BSA, 2 mM ATP, 0.6 mM each dNTP and 30 units of T4 DNA ligase (NEB) in 50 μ l volume was done at 37°C for 3 hours. The double stranded DNA so formed was purified by mixing the DNA with cesium chloride and ethidium bromide and centrifuging for 24 hours at 60,000 rpm in a Type Ti 70 rotor to create a density gradient of cesium chloride to separate pure double stranded DNA from nicked and linear DNA. The DNA

bands intercalated with ethidium bromide were identified under UVA light. Double stranded circular DNA was withdrawn with a syringe and purified from salts and other impurities by using Amicon 100K centrifugation columns (EMD Millipore Inc.). The substrates were analyzed by restriction digestion with *ApaLI* enzyme. Cisplatinated double stranded DNA has a unique digestion site for *ApaLI* at the GTG site where cisplatin reacts with the DNA. The cisplatin containing DNA is resistant to *ApaLI* while *ApaLI* cuts the undamaged double stranded DNA (Figure 6B). RB69gp43 was purified by Karen Boulware.

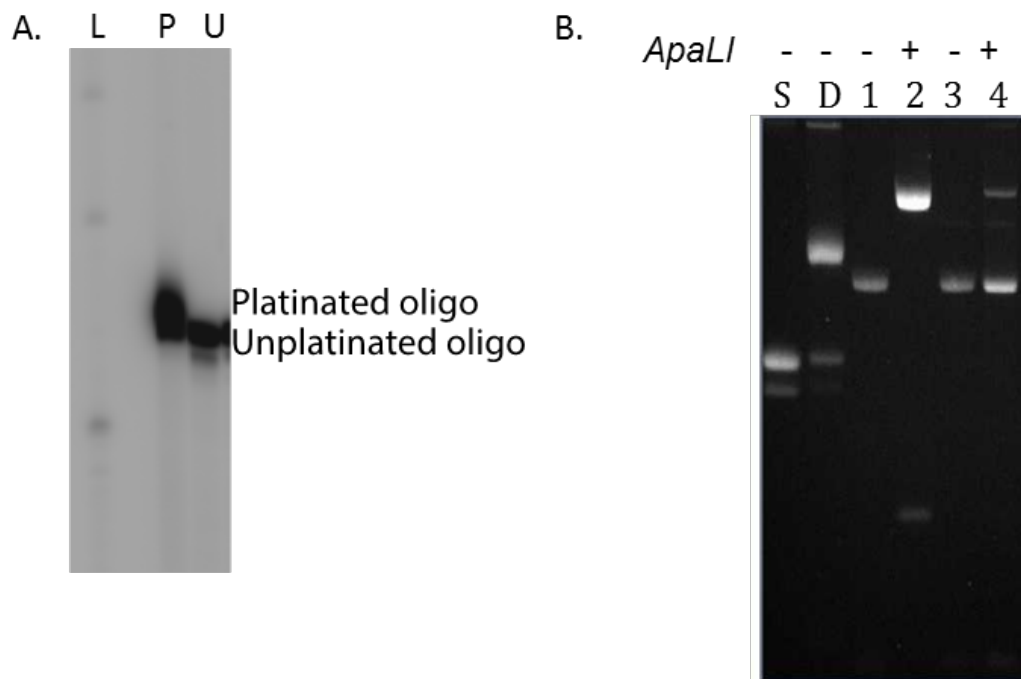


FIGURE 6. Verification of 1,3 GTG cisplatin primer and the cisplatin adduct in M13mp18GTG double stranded circular DNA.

A. Denaturing 20% polyacrylamide with radiolabeled 24-mer oligonucleotides platinated (P), unplatinated (U) and 100 bp DNA ladder (L). B. 0.8% agarose gel with ethidium bromide run in TBE with ethidium buffer at 20 mM for 8 hours. The lanes consist of following samples; S: single stranded M13mp18GTG plasmid D: Double stranded M13mp18GTG. 1 and 2: undamaged DNA, 3 and 4: cisplatin substrate. 1-4 are used in restriction digestion conditions with or without *ApaLI*.

2.6. DNA incision and repair assays with cisplatin intrastrand substrate

2.6.1. NER dual incision control assay

Since it is well known that mammalian cells process cisplatin intrastrand crosslink adduct by NER dual incision (102), we established a control assay using cisplatin adduct in a circular plasmid. We prepared mammalian whole cell extracts and a single 1,3 (dGpTpG) cisplatin adduct substrate as described above and tested in NER dual incision assay using an end-labeling method as described by Moggs et al., 1996 (102). The 1,3 (dGpTpG) cisplatin substrate was incubated with mammalian whole cell extracts. Briefly, the reaction mixture consisted of 50 ng of Pt-GTG or control substrate, 40-45 µg of whole cell extract, in the NER buffer with a final concentration of 45 mM HEPES-KOH (pH 7.8), 70 mM KCl, 7.5 mM MgCl₂, 0.9 mM DTT, 0.4mM EDTA, 2 mM ATP, 2.5 µg of creatine phosphokinase, 3.4% glycerol and 18 µg of bovine serum albumin in a 10 µl volume. The reaction is incubated at 30 °C for 5 min without DNA and incubated further for 30 minutes after DNA substrate was added and mixed. An oligonucleotide with a G-overhang at the 5' end (GGG GGA AGA GTG CAC AGA AGA AGA GGC CTG GTC G) was added to the mixture and heated at 95 °C for 5 minutes and cooled gradually at room temperature for 30 minutes. During this process, the G-overhang oligonucleotide was annealed to a complementary oligonucleotide excised by dual incision NER. Sequenase polymerase (0.13 U) was used to add [α -³²P] deoxycytidine triphosphates opposite the G-overhang

using a final concentration of 2 μCi of [α - ^{32}P] dCTP and incubated at 37 °C for 3 min followed by addition of 10 μM each of dATP, dTTP and dGTP, and 5 μM dCTP in the mixture and incubated for 12 more minutes. The reaction was stopped by adding 90% formamide with bromophenol blue dye and denatured by heating at 95 °C for 4 minutes. The reaction was run 40 cm on a denaturing 14% polyacrylamide gel or the sequencing gel at 50°C for about 2 hours or until the dye ran at the bottom of the gel. The gel was dried for 1-2 hr in gel drier and exposed to phosphoscreen for 1-2 days.

2.6.2. Primer extension assay to a 1,3 (dGpTpG) cisplatin adduct

A primer extension assay was performed in order to confirm the 1,3 (dGpTpG) cisplatin adduct at a specific site of one strand of the closed circular double stranded DNA M13mp18GTGX. 25 pmoles of primer (5'-CAGGAAACAGCTATGAC-3') was radiolabeled with 17 pmoles of [γ - ^{32}P] ATP at its 5' end by T4 Polynucleotide kinase (T4 PNK) in 1X T4 PNK buffer for 1 hour at 37 °C. M13mp18GTGX was restriction digested by *PvuI* restriction enzyme at 37 °C for 1 hour, heat inactivated, purified by phenol-chloroform-isoamyl alcohol and precipitated by 100% ethanol and 3M sodium acetate. 300 ng of purified DNA was incubated with 20 fold more radiolabeled primer at 95 °C for 4 min, 48°C for 4 min and 37 °C for 1 min in the mixture of Sequenase V2.0 polymerase (USB Inc.) 1.3 U, 1X Sequenase buffer and 80 μM dNTPs mixture, in a thermocycler. The amount of Sequenase V 2.0, a T7 DNA polymerase, which

lacks exonuclease activity, was titrated for efficient primer extension. As a control we used RB69gp43 polymerase, which had worked well in the past for primer extension of a 30 mer oligonucleotide with its specific primer. Sequenase V 2.0 was found to be working well and we continued to use this polymerase for further primer extension assays. The reaction was stopped with 90% formamide buffer and heated at 95 °C for 4 minutes and run in a denaturing 14% polyacrylamide gel. The gel was dried in heated vacuum gel drier for 30 min to 2 hr. The dried gel was exposed to phosphoscreen for 1-2 days and scanned in GE Typhoon FLA 9500.

2.6.3. DNA repair assay

The DNA repair assay was done based on a DNA synthesis reaction as described in Moggs et al (102). DNA repair assay was performed in similar reaction conditions as NER dual incision assay in 50 µl volume and it included radiolabeled with [α -³²P] dNTPs (20 mM of each dATP, dTTP, dCTP and dGTP, 2 µCi of 3000 Ci/mmol). 300 ng of DNA was incubated in this reaction for 30 min at 30 °C after pre-incubating for 5 min without DNA. The repair intermediates were purified by phenol chloroform isoamyl extraction and ethanol precipitation as described earlier. The purified DNA is restriction digested with 15 units of BstN1 in 30 µl volume at 60 °C for 4 hours. The fragments were analysed by electrophoresis in a denaturing 14% polyacrylamide gel for 1.5 to 2 hr at 50 °C as described earlier.

2.7. Construction of TFO-directed psoralen ICL on a a closed circular double stranded DNA (psoralen-TFO ICL)

2.7.1. Vector with *BbsI* ends

A pUC19 vector was mutated at T685A by using QuickChange Site-Directed Mutagenesis Kit (Stratagene), to mutate the *Nt.BspQI* site. A fragment was cloned into pCR4-TOPO TA cloning vector (Invitrogen). The mutation was confirmed by sequencing. The mutated plasmid was transformed into SCS110 Competent Cells (Agilent Technologies) and colonies containing the mutated plasmid were selected for in LB broth with carbenicillin. DNA was extracted using a Wizard SV miniprep kit (Promega). Purified mutated-plasmid was restriction digested with *XbaI* and *KpnI* to remove the fragment and an insert consisting of a TFO binding site and *BbsI* restriction digestion sites at the ends was ligated with it. The purified ligated product was transformed in DH5- α cells. Colony PCR was done to confirm the insertion. Colonies with the correct insertion (result from PCR) were cultured in LB broth with carbenicillin and DNA was extracted. The sequences were confirmed for the whole plasmid. Thus this modified pUC19 (pUC19N) consisted of two *BbsI* sites that brackets a TFO binding site together with 5'-AT-3' site for crosslinking with psoralen (113).

pUC19N (500 µg or more) was restriction digested with *Bbs*I (2 Units/ µg DNA) at 37 °C for 16-20 hours. DNA was purified by passing through DNA Clean & Concentrator centrifugation columns (Zymo Research), run in denaturing 6% polyacrylamide gel, stained with SYBR Safe (Invitrogen) for 30 min in dark. The gel was scanned in Typhoon 9500 at 532 nm. The vector bands were marked and excised with a clean razor blade. The bands were electroeluted in a dialysis bag in 0.5 X TBE buffer by running the gel at 200 V/cm for 45 minutes. DNA was precipitated from the electroeluate with 3 M sodium acetate pH 5.2 and ethanol at -20 °C overnight, centrifuged at 20,000 x g at 4 °C for 30 minutes, washed with 2 volumes of 70% ethanol, air dried and resuspended in TE buffer pH 8.0. The concentration and the purity were measured on a NanoDrop 2000 (Thermo Scientific).

2.7.2. 50 bp psoralen-TFO ICL

Equal molarity of a 50 bp upper strand oligonucleotide
(5' Phos-
CGATGCTCTTCATCCTTCCCCCCCCACCACCCCCTCCCCCTCGAAGACGC
3')
and lower strand oligonucleotide (5'Phos-
GCACGCGTCTTCGAGGGGGAGGGGGTGGTGGGGGGGAAGGATGAAGAG
C 3') were annealed together to form a duplex DNA, in TE buffer by heating at 95 °C for 2 minutes, and gradually letting it cool to room temperature for 3-4 hours

and stored at 4°C. 50 fold more psoralen (4' Hydroxymethyltrioxsalen)-TFO (5' HMT-C6-AGGAAGGGGGGGGTGGTGGGGGAGGGGGA-dGTP-NH₂ 3') was incubated with 50 bp duplex DNA in 5X TBB (triplex binding buffer), which consisted of 10 mM MgCl₂ and 10 mM Tris-Cl pH 7.6 for 16 hours at 37 °C. The psoralen-TFO bound to DNA is exposed to UVA (365 nm) at 10 J/m²/s for 30 minutes or 18 kJ/m². Small amounts of the UVA exposed products were confirmed for the interstrand crosslink by photo-reversal with UVC (0.7 kJ/m²), radiolabeled and run in a denaturing 6% polyacrylamide gel. The rest of the UVA-exposed products were separated in a denaturing 6% polyacrylamide gel and stained with SYBR safe (Invitrogen). The interstrand crosslink bands were excised from the gel making sure not to contaminate them with any monoadducts or duplex DNA. The gel pieces were crushed in microfuge tubes and soaked in nuclease free water in a shaker at 37 °C for 16-20 hours. The water separated from the gel pieces was collected and filtered through 0.22 µm pore cellulose centrifugation tubes (Corning® Coaster® SpinX®). It was kept at -80 °C for 20 minutes and concentrated in a vacuum centrifuge with 37 °C heat until the volume reduced to 100-200 µl. The product was purified through a DNA Clean & Concentration column (Zymo-Research). The purity and the concentration of the insert were measured on a NanoDrop 2000.

2.7.3. Ligation reaction

A small-scale ligation reaction was performed with 1:1.5, 1:3, and 1:5 ratio of vector to insert, 2 units of T4 DNA ligase, 1X T4 ligase buffer, 4 µg of BSA in a 40 µl reaction volume at 16°C for 2 hours in dark. The reaction was run in a 1% agarose gel with ethidium bromide. A 1:5 vector to insert reaction was chosen to scale up the reaction. A large-scale reaction was done in a 20 mL volume (500 X). The reaction was concentrated in Amicon 30K 4 mL centrifugation tubes (EMD Millipore) at 3452 x g at 4 °C until the volume was reduced to 1 mL. Further, the double stranded closed circular DNA was separated from nicked and single stranded DNA by cesium chloride / ethidium bromide gradient centrifugation. The double stranded circular DNA is collected from the lower band. The upper band consists of nicked and single stranded DNA. Ethidium bromide was removed by several extractions with water-saturated butanol and butanol is removed by centrifuging it several times with TE in Amicon 10K 4 mL centrifugation tube (EMD Millipore) at 4 °C at 800 x g. The solution was concentrated to a final volume of 250-300 µl. 5X TBB was added to the purified DNA and incubated at 37 °C for overnight and stored at 4 °C.

2.7.4. Confirmation of the ICL in the final product

The final TFO-directed psoralen ICL on a closed circular double stranded DNA is referred here as psoralen-TFO ICL unless specified. Small amounts of

the final ICL product (psoralen-TFO ICL) or pUC19N were restriction digested with *Bbs*I for an hour at 37 °C and run in a denaturing 6% polyacrylamide gel along with purified ICL insert, stained with SYBR gold for 30 min in the dark and scanned in the Typhoon 9500. The restriction digestion of psoralen-TFO ICL released a 50 bp ICL that runs similar to 50 bp ICL insert whereas the 50 bp fragment released from pUC19N runs as a normal 50 bp DNA. This confirms that the final product is TFO directed psoralen interstrand crosslink in closed circular double stranded DNA.

2.7.5. Plasmid relaxation assay in whole cell extract

500 ng of pUC19N was incubated for 5 minutes to 40 minutes in 100 µg of HeLa S3 whole cell extract protein or without whole cell extract in ATP regenerating buffer (NER buffer) as described in the NER incision reactions above. The reaction was stopped by using 20 mM EDTA and extracted using RNase, proteinase K, SDS and phenol-chloroform-isoamyl alcohol, and again with chloroform, precipitated with 100% ethanol and sodium acetate. The DNA pellet was obtained by high speed centrifugation at 4°C for 30 minutes, washed with 70% ethanol, air dried and dissolved in 20 µl of TE buffer. The samples were divided into halves and run in two 1% agarose gels with or ethidium bromide for either 20 mA for 8 hours or 80V for 2 hours. Pictures were taken in a Bio-Rad ChemiDoc imager.

2.8. DNA incision and repair assays with psoralen-TFO ICL

2.8.1. Detection of 3' incision on the upper and lower strand

50 ng of pso-TFO ICL substrate was incubated in HeLa S3 whole cell extract in 50 µl reaction volumes in the buffer condition as described in the NER dual incision control assay. The reaction was stopped by adding 20 mM of EDTA and purified by equal volume of phenol-chloroform-isoamyl extraction and one more extraction with chloroform followed by precipitation of intermediate DNA products with absolute ethanol and sodium acetate by incubating at -20°C for 1 hour. The precipitate was obtained by centrifugation at maximum speed in a table top centrifuge at 4°C for 30 minutes. The precipitate was washed with 70% ethanol, air-dried and dissolved in 10 µl TE buffer. The purified DNA products were restriction digested by 20 units of *SbfI* (NEB) enzyme for 2 hours at 37 °C and purified as before and dissolved in 10 µl of TE buffer. A G-overhang labelling oligonucleotide was designed to detect the incision products on the upper strand. 6 ng of 3' incision labelling oligonucleotide (5' GGGGTGCAGGTCGACTCTAGAGCGTCTTCG 3') was added and denatured at 95 °C for 3 minutes and allowed to anneal with the incision products by gradually cooling at room temperature over 40 minutes. Incorporation of [α -³²P]dCTP complementary to the G-overhangs on the labelling oligonucleotides was done as described in the NER dual incision control assay. Running the samples in the

sequencing gel and further processes were done as in the NER dual incision control assay.

Another G-overhang oligonucleotide (5'GGGGGAGCACGCGTCTTCGAGGGGGAGGGGGT 3') was used to detect the incision products at the site within the TFO binding site following the method described above. For this detection process, the purified DNA product after incision reaction was restriction digested with 20 units of *BfaI* (NEB). Similarly to detect the incision products at the 3' side on the lower strand, we used a G-overhang oligonucleotide (5' GGGGGTACCGAAGACATCGATGCTCTTCAT 3').

2.8.2. psoralen-TFO ICL repair assays

Repair assays were performed to determine if the psoralen-TFO ICL could be repaired by human cell extract. The DNA repair synthesis reaction was based on that described by Moggs *et al.* (102) and purified as described for the DNA repair assay above. The purified DNA is restriction digested with 20 units of each *SapI* (NEB) and *AluI* (NEB) at 37 °C in 30 µl volume for 2 hr. The fragments are analysed by electrophoresis in a denaturing 14 % polyacrylamide gel for 1.5 to 2 hr at 50 °C as described earlier.

A similar repair synthesis assay was also performed where the purified DNA after the repair reaction was restriction digested by *BbsI* (NEB).

Chapter 3. Results - Consequences of XPA disruption in human cells

3.1. Validation of pairs of XPA-deficient and proficient cell lines

The purpose of this study was to determine the extent to which XPA expression status influences overall gene expression in cultured cells. We considered it important to use independent, genetically matched pairs of cell lines where one cell line was completely XPA-deficient, and the other was XPA-proficient. Four pairs of cell lines were investigated. Figure 4A indicates the sites of XPA mutation and deletion in each of the cell lines. Two pairs include widely used and characterized XPA-deficient cell lines derived from human skin fibroblasts of individuals with xeroderma pigmentosum group A, XP2OS and XP12RO. These were compared to the same cell lines complemented with a plasmid expressing *XPA* cDNA (Table 1). In both cases, there is ample evidence that XPA expression fully corrects the UVC radiation sensitivity and NER defect in these fibroblasts. As another comparison pair of cell lines, two CRISPR-Cas9 mediated XPA-disrupted HeLa S3 cell lines were generated (Figure 4B). We reasoned that if any XPA-associated gene expression changes were found in common across several cell lines, they would represent the most biologically significant consequences of XPA expression.

In three of the *XPA* mutant cell lines, XPA protein was undetectable by immunofluorescent staining of cells (Figure 7) or by immunoblotting of cell

extracts (Figure 8A, B). This is consistent with the known mRNA destabilizing mutations in XP2OS (114) and XP12RO (115) (Figure 4). The HeLa KO142 cells encoded an XPA protein with a deletion of six amino acids (residues 9-14) near the N-terminus (Figure 4A and 4B). This, perhaps surprisingly, destabilizes XPA expression, reducing it to ~10% of the level in the HeLa cells (Figure 8A and 8B). The GTX103168 antibody recognizes the globular core domain of XPA and also detected a reduced level of XPA in HeLa KO142 cells (Figure 8B). The HeLa KO38 cells had a frameshift mutation causing early termination and complete loss of XPA protein. UV irradiation of paired cell lines confirmed the NER deficiency in the *XPA* mutant cell lines showing that there is a reduced level of NER in KO142 (Figure 9A). This is consistent with previous data showing that XPA levels are rate-limiting for NER when reduced sufficiently (116, 117). HeLa KO38 cell extracts were completely deficient in NER and could be complemented by XPA protein (Figure 9B). The common genetic origin of the paired cell lines was confirmed by short tandem repeat analysis (Table 1). KO38 was a complete knockout. UV irradiation of paired cell lines confirmed the NER deficiency in the *XPA* mutant cell lines showing that there is a reduced level of NER in KO142 (Figure 9A). This is consistent with previous data showing that XPA levels are rate-limiting for NER when reduced sufficiently (118). We also tested if XPA protein in KO142 is affected by proteosomal degradation and thus is observed as low expression. MG132 proteasome inhibitor was used in the cell cultures for 24 hours and cells were subjected to immunoblotting. No difference was observed in XPA protein level in presence or absence of MG132 (Figure 10). HeLa KO38

cell extracts were completely deficient in NER and could be complemented by XPA protein (Figure 9B).

Sample Name	AMEL	CSF1PO	D16S539	D18S51	D21S11	D3S1358	D5S818	D7S820
XP12RO	X,Y	12	9,11	13,22	30.2	15,16	11,12	12
XPAXP12RO	X,Y	12	9,11	13,22	30.2	15,16	11,12	12
XP20S	X	10,11	12	13,16	29,30	16	11	8,12
XPAXP20S	X	10,11	12	13,16	29,30	16	11	8,12
KO142	X	9,10	9,10	16	27,28	15,18	11,12	8,12
KO38	X	9,10	9,10	16	27,28	15,18	11,12	8,12
HeLa S3	X	9,10	9,10	16	27,28	15,18	11,12	8,12

TABLE 1. Short Tandem Repeat (STR) fingerprinting report of cell lines used in the study

Color-coding is used to denote the XPA proficient cell lines with their deficient pairs. The scores for repeats of 8 STR fingerprinting markers are given in the table to show that the cell line pairs matched each other. A total of 14 markers were used for the STR fingerprinting. All had matching scores for the pairs of cell lines, showing that the pairs are isogenic.

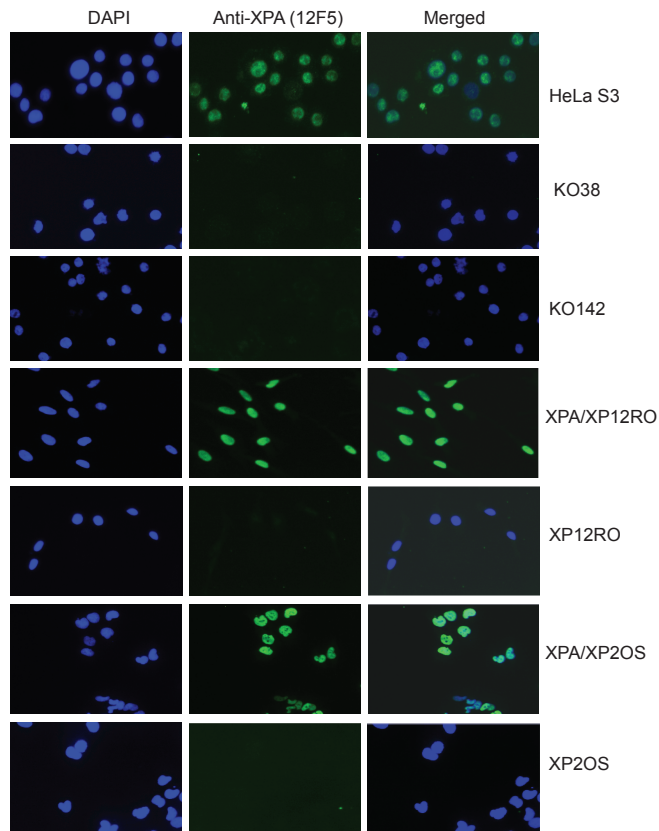


FIGURE 7. XPA expression in the cell lines used in this study.

Immunofluorescence images using anti-XPA antibody (12F5) and DAPI showing the presence of XPA in all XPA-proficient cell lines and little or no XPA in the deficient cell lines. DAPI (blue) staining of nuclei included for reference.

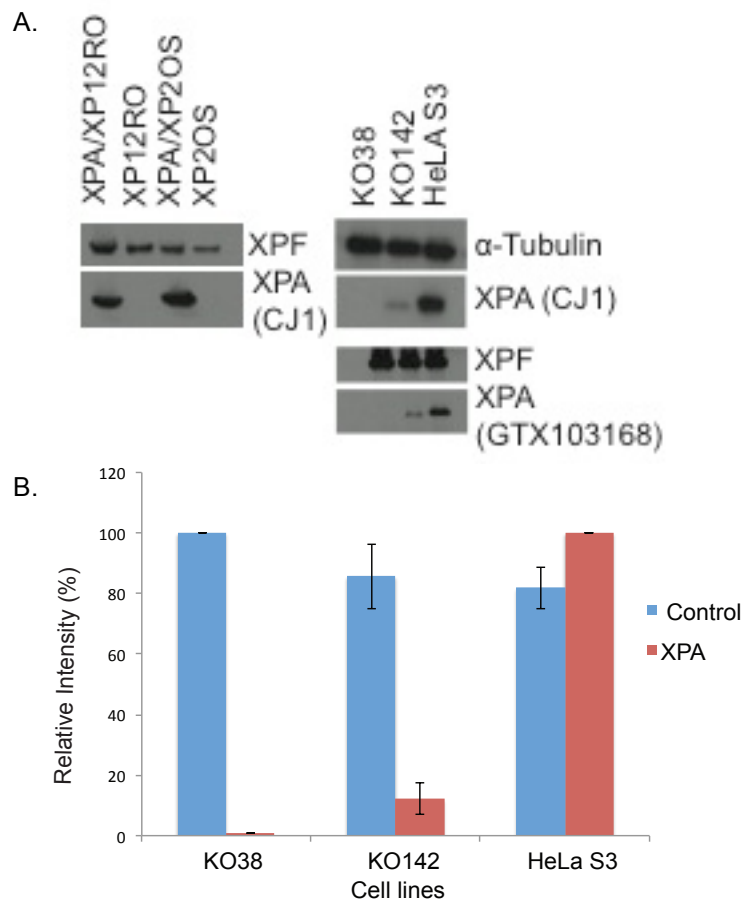


FIGURE 8. XPA expression in the cell lines used in this study by immunoblotting.

A. Immunoblot of the cell lines with anti-XPA antibodies CJ1 and GTX103168, and alpha-tubulin or XPF antibodies as loading controls. B. Quantification of the immunoblots for the HeLa S3 cell lines shown in (A). The blue bars represent the alpha-tubulin or XPF control. The red bars represent the anti-XPA antibody.

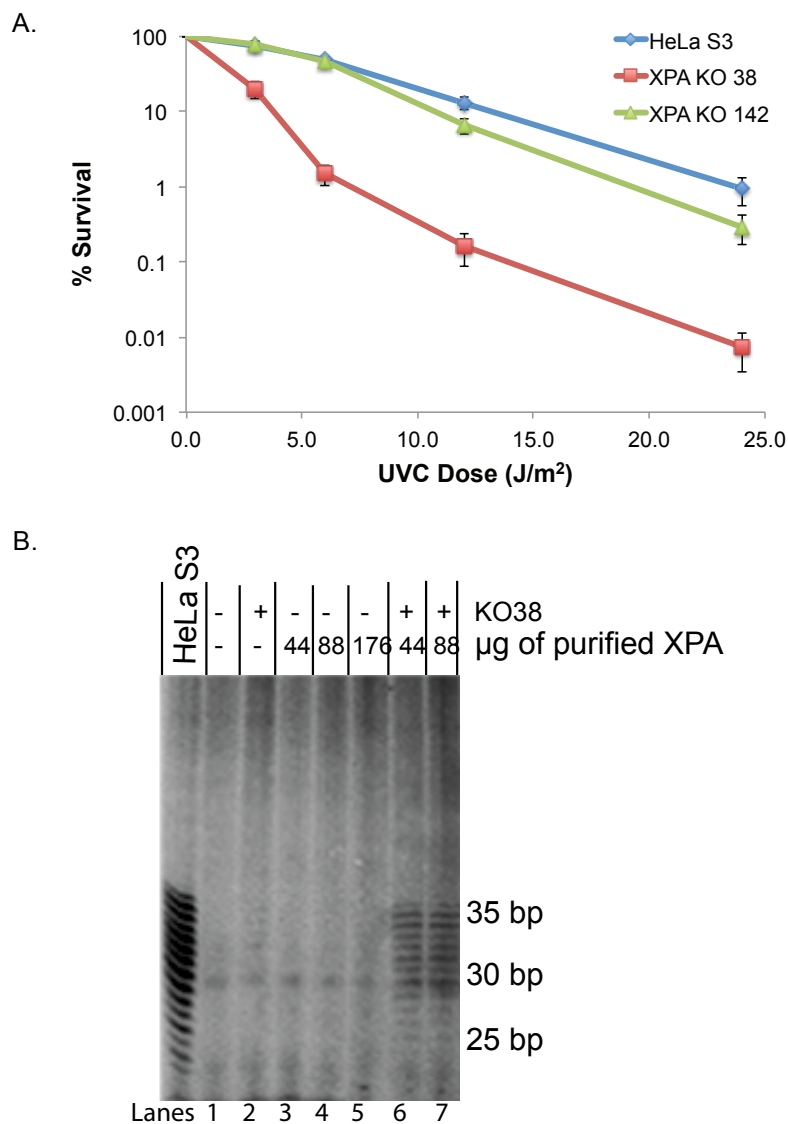


FIGURE 9. UVC sensitivity and NER complementation in the cell lines used in this study.

A. UVC clonogenic survival for XPA-disrupted cell lines KO38 and KO142 compared to wild type HeLa S3. Each point on the survival curves represents the mean of 3 independent experiments. KO38 is more sensitive to UVC compared to HeLa S3. KO142 has more moderate sensitivity compared to HeLa S3. B.

NER complementation assay. KO38 whole cell extracts complemented with purified XPA protein to produce incision products. 45 µg 1,3-intrastrand d(GpTpG)-cisplatin substrate was incubated with 45 µg whole cell extract. Various amounts of purified XPA protein were added to KO38 extracts. NER activity is detected with 44 and 88 µg of purified XPA combined with KO38 extracts. Incision products were observed as dark bands between 25 bp to 35 bp where complementation was successful (lanes 6 and 7), similar to the incision products with NER proficient HeLa S3 extract (lane 1).

UVC sensitivity clonogenic assay was performed by Megan Lowery and Karen Boulware purified XPA for NER assay.

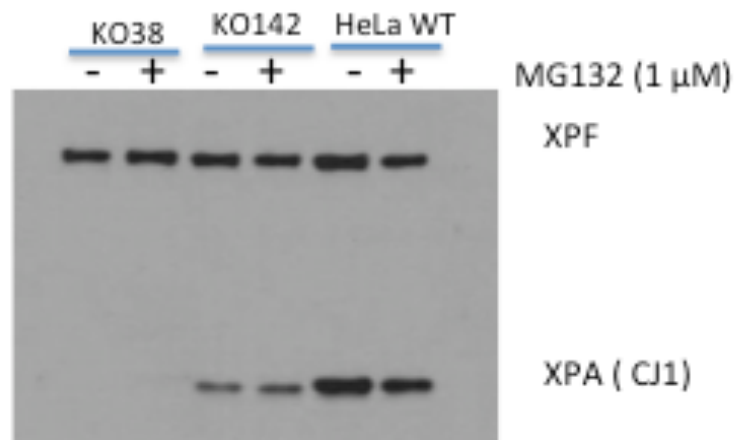


Figure 10. **Proteasome Inhibition assay**

Cells were treated with proteasome inhibitor (MG132) or DMSO for 24 hours. Cell lysates from the treated cells were immunoblotted to determine if expression of XPA in KO142 increases in absence of proteasomal degradation. No difference was observed with and without MG132. XPF was used as a loading control.

3.2. XPA is localized in the nucleus of cells and does not require UV irradiation for import

In order to have a direct role as a transcription factor, XPA would have to be present in the cell nucleus, where transcription takes place. Although XPA is localized in the nucleus under standard fixation conditions (Figure 7, 11A), we were concerned about reports that XPA is normally resident in the cytoplasm, and is only recruited to nuclei following UV radiation exposure (89, 91). We note that in those studies, cells were first briefly exposed to wash buffers containing Triton X-100 detergent or the milder fixation agent methanol, or with 0.05% NP-40 detergent. As shown in Figure 11A (top row), XPA is entirely resident in the nucleus as visualized after paraformaldehyde fixation. However, because XPA is a relatively small protein (31 kDa), it easily leaks out of nuclei into the cytoplasm after only a few seconds' exposure to a buffer containing 0.05% NP-40 (Figure 11A, third row).

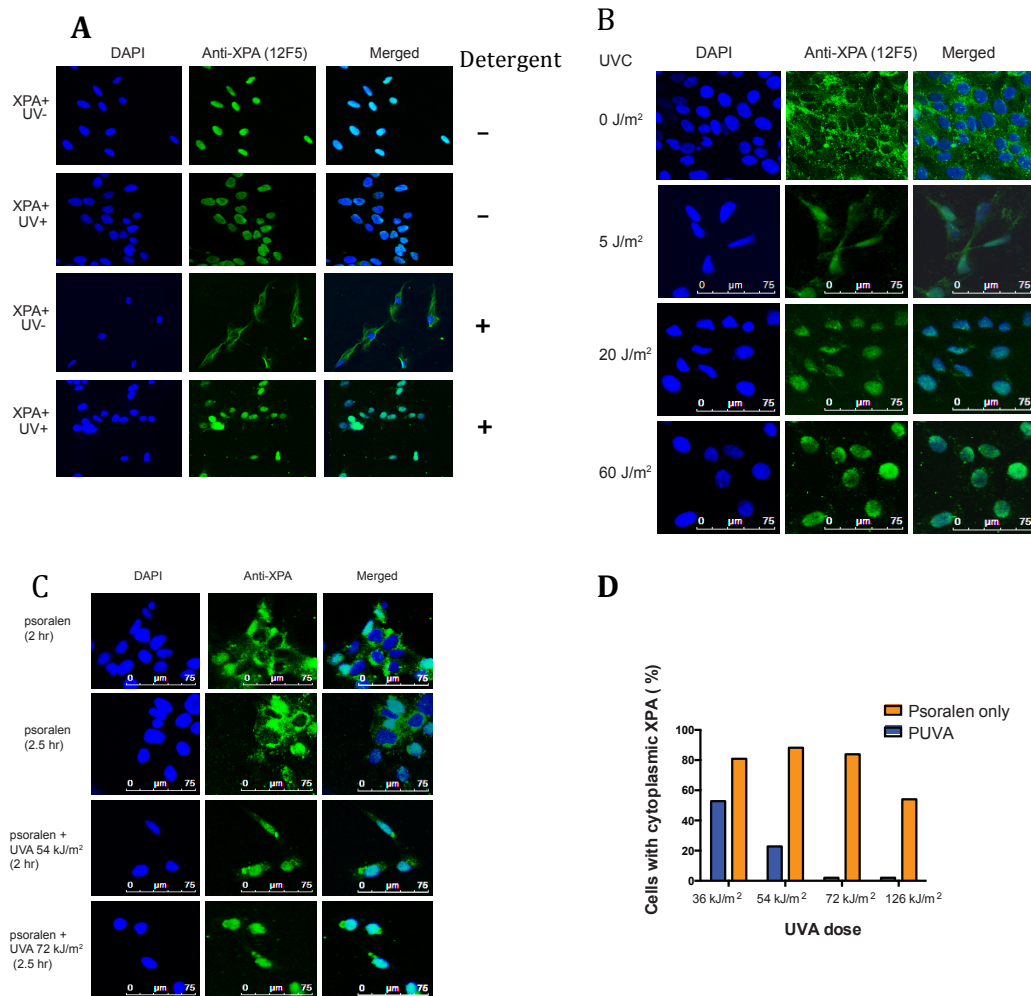


FIGURE 11. XPA is a nuclear protein and is resistant to extraction from the nucleus during DNA repair.

A. Immunofluorescence (IF) staining with 12F5 antibody (green) and DAPI (blue). Cells in the top two panels were stained without exposure to buffer containing detergent. This shows that XPA is localized in the nuclei of complemented XPA/XP12RO cells and is absent in XP12RO cells, regardless of exposure to

UVC. The lower two panels show that brief exposure to a buffer containing 0.05 % NP-40 causes leakage of XPA into the cytoplasm in untreated cells but not in cells exposed to 20 J/m² UVC. B. Exposure of XPA/XP12RO cells to increasing doses of UVC (from 0 J/m² to 60 J/m²) restricts the leakage of XPA to the cytoplasm. C. XPA/XP12RO cells were exposed to psoralen alone or to psoralen and then UVA radiation (PUVA). The 2 hr sample received 54 kJ/m² (first dose 18 kJ/m², second dose 36 kJ/m²) and the 2.5 hr sample received 72 kJ/m² (first dose 18 kJ/m², second dose 54 kJ/m²). All cells were exposed briefly to a buffer containing 0.05 % NP-40 before fixation for IF. XPA leaked to the cytoplasm in the cells exposed to psoralen alone, but not in the PUVA treated cells. D. Quantification showing the % of XPA/XP12RO cells with cytoplasmic XPA following PUVA treatment with increasing total doses of UVA and a constant concentration of psoralen. All cells were exposed briefly to a buffer containing 0.05 % NP-40 before fixation for IF. The orange bar shows XPA leakage in cells exposed only to psoralen, and the blue bar shows a reduction in leakage following DNA damage induction with increasing UVA doses.

Cells exposed to sufficient DNA damage retained XPA in the nucleus for at least 30 minutes even after a brief wash in buffer containing NP-40 detergent. A UVC radiation dose of 20 J/m² was sufficient to retain all detectable XPA in the nucleus (Figure 11A, fourth row and Figure 11B). This result is consistent with previous observations that XPA remains chromatin-bound after UV irradiation, but is extractable from chromatin in non-damaged cells (119). Similarly, we found that photoactivated psoralen lesions could also immobilize XPA in the nucleus, even in cells washed with buffer containing NP-40 (Figure 11C). Consequently, previous observations suggesting nuclear import of XPA following UV radiation appear to have been misinterpreted (91). Instead, XPA is always present in nuclei. Fixation in the presence of detergent facilitates nuclear membrane permeabilization, and XPA leaks out readily; DNA damage in the nucleus temporarily restricts such leakage while XPA is engaged in repair. Other examples of misinterpretation of the subcellular localization of proteins arising from the presence of detergent have been noted previously (93, 94).

3.3. Influence of *XPA* status on overall gene transcription

To analyze gene expression, RNA was extracted from a culture of 10⁶ actively growing cells of each pair, and used for next-generation RNA-Seq. To gauge reproducibility, three replicate cultures were grown independently for each cell line and RNA was extracted from cell pellets for analysis. With an overall mapping rate of 87-96%, about 14,000 genes were analyzed for each pair of cell

lines. With a False Discovery Rate (FDR) less than or equal to 0.05, a very high proportion of genes (~9000 for each pair) were initially identified as differentially expressed between paired XPA-proficient and deficient cell lines (Table 2).

However, the expression patterns of these genes for the four pairs of cell lines were very distinct from one another. The ratio of gene expression was calculated for XPA+/XPA- for each cell pair; positive log 2 values were termed “high” and negative log 2 values were termed “low”. We determined the genes that were uniformly “high” or “low” for all four cell pairs. This narrowed the set to 325 genes with expression that is uniformly influenced by XPA status at an FDR < 0.05 (Table 2).

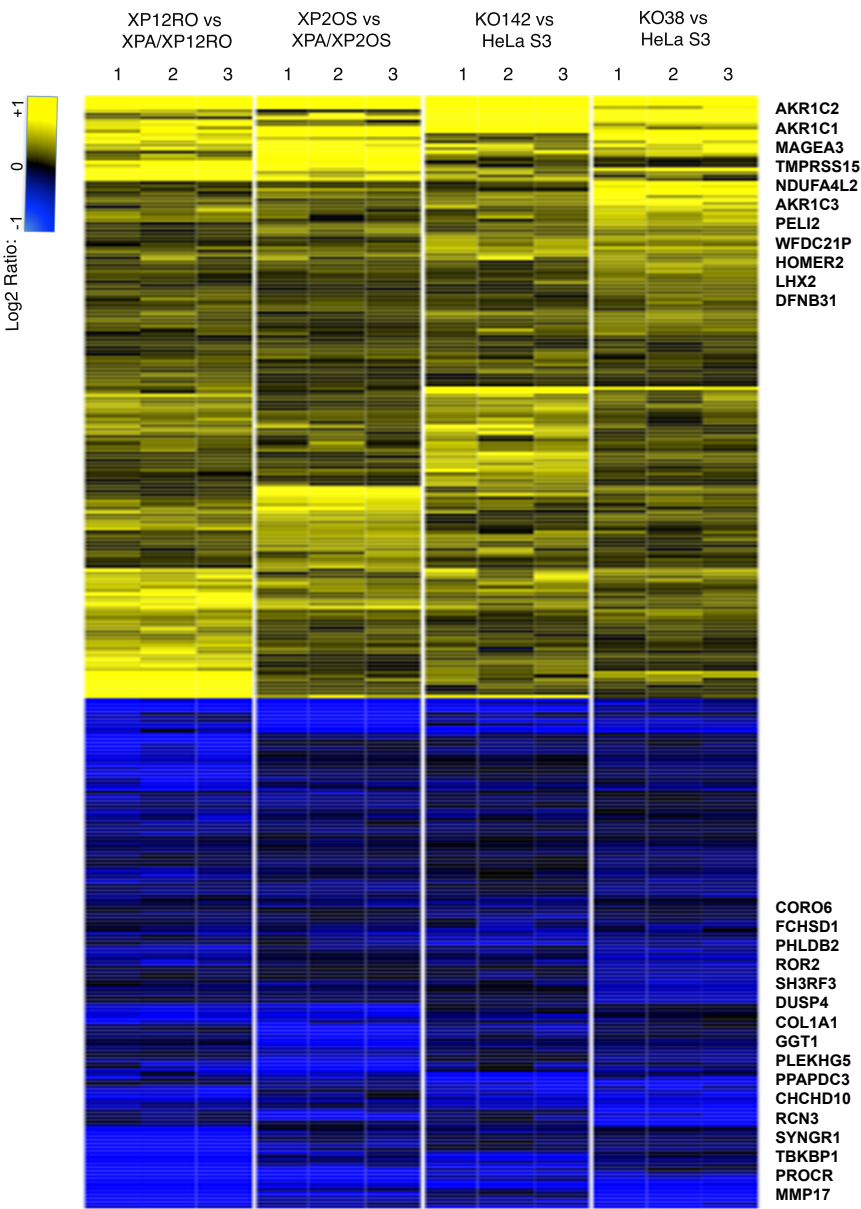
This set of 325 genes has the same trend of high or low expression in all the XPA-proficient cell lines compared to their deficient counterparts (Figure 12A). We used Gene Ontology (GO) consortium software to analyze the PANTHER pathways of these genes (Table 3). Only 32 of the 325 genes were assignable by this analysis with p-value less than 0.01, but they represented a diversity of biological functions (Figure 12B). Notably, a substantial number of pathways referred to neurological disorders including Alzheimer’s disease, Huntington’s disease and Parkinson disease.

Cell lines	Genes Analyzed	FDR0.05 (High + Low)	FC1.5 (High + Low)
XP2OS vs. XPA/XP2OS	14117	9526	1721 + 2206
XP12RO vs. XPA/XP12RO	13669	10283	2750 + 2625
KO38 vs. HeLa S3	14564	9520	1769 + 1701
KO142 vs. HeLa S3	13824	8364	1642 + 1422
Common genes	12022	176 + 149	11 + 16

TABLE 2. Summary of genes analyzed by high-throughput RNA sequencing (RNA-Seq) and significant XPA-related differences.

The first column shows the number of genes analyzable by RNA-Seq for each XPA proficient / deficient cell pair. The second column shows the number of genes with differential expression at a False Discovery Rate (FDR) of 0.05. The ratio of gene expression was calculated for XPA-/XPA+ for each cell pair; positive $\ln 2$ values are termed “high” and negative $\ln 2$ values are termed “low”. The third column shows for each cell pair the number of “high” or “low” genes with fold change (FC) of 1.5 or more.

A



B

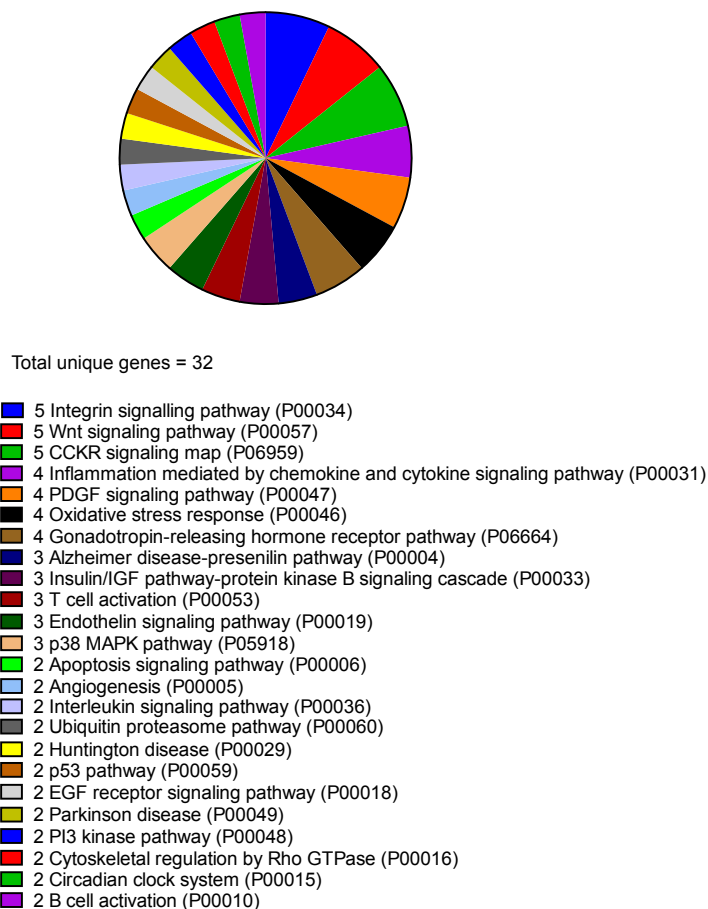


FIGURE 12. XPA dependent change in gene expression pattern and biological pathways influenced by XPA status.

A. Heatmap of the 325 genes (Table 2) having the same expression pattern, either high (yellow) or low (blue) in XPA proficient compared to deficient cell lines, at a cut off level of FDR 0.05. Differential expression of mRNA was obtained from high throughput RNA-Seq data. Experiments for each cell line pair were performed in triplicate and are numbered on the top of the heatmap. The ratio of gene expression was calculated for XPA-/XPA+ for each cell pair; positive

log 2 values are termed “high” and negative log 2 values are termed “low”. The most differentially over- or under- expressed genes are listed to the right. B. PANTHER pathway analysis was performed using Gene Ontology (GO) consortium software for the 325 genes shown in part A. Most genes were unassigned to a pathway by this analysis. The figure shows the pathway assignment for the subset of 32 genes that had two or more hits to a PANTHER pathway. The number of genes represented in each pathway is given before the name of the pathway with PANTHER pathway terms in parenthesis. Some genes were assigned to multiple pathway terms. The full data and assigned genes are given in Table 3.

Panther Pathways	XPA - vs. XPA+		
	High	Low	# of genes hits
Integrin signalling pathway (P00034)	BCAR1, PIK3R3, RRAS	COL1A1, FLNB	5
Wnt signaling pathway (P00057)	EDN1, FZD1	PPP3CA, ITPR1, CHD1L	5
CCKR signaling map (P06959)	BCAR1	PPP3CA, RPS6, ITPR1, PDK1	5
Inflammation mediated by chemokine and cytokine signaling pathway (P00031)	RRAS	CAMK2D, ITPR1, PDK1	4
PDGF signaling pathway (P00047)	MKNK2, PIK3R3	ITPR1, PDK1	4
Oxidative stress response (P00046)	MKNK2	EEF2K, DUSP4,	4

		DUSP1	
Gonadotropin-releasing hormone receptor pathway (P06664)	LHX2	PPP3CA, ITPR1, DUSP1	4
Alzheimer disease-presenilin pathway (P00004)	FZD1	MMP17, LRP3	3
Insulin/IGF pathway-protein kinase B signaling cascade (P00033)	TSC1, PIK3R3	PDK1	3
T cell activation (P00053)	PIK3R3	PPP3CA, ITPR1	3
Endothelin signaling pathway (P00019)	EDN1, PIK3R3	ITPR1	3
p38 MAPK pathway (P05918)	MKNK2	EEF2K, DUSP1	3
Apoptosis signaling pathway (P00006)	RIPK1, TM6SF1		2
Angiogenesis (P00005)	FZD1, PIK3R3		2
Interleukin signaling pathway (P00036)	MKNK2	PDK1	2
Ubiquitin proteasome pathway (P00060)	PSMC3	UBE2L6	2
Huntington disease (P00029)	CAPN2, TUBB2A		2
p53 pathway (P00059)	PIK3R3	PDK1	2

p53 pathway feedback loops 2 (P04398)	PIK3R3	PDK1	2
EGF receptor signaling pathway (P00018)	RRAS	PHLDB2	2
Parkinson disease (P00049)	PSMA4	UBE2L6	2
PI3 kinase pathway (P00048)	PIK3R3	PDK1	2
Cytoskeletal regulation by Rho GTPase (P00016)	TUBB2A, CFL1		2
Circadian clock system (P00015)	PER2, PER3		2
B cell activation (P00010)	PIK3R3	PPP3CA, ITPR1	2
Axon guidance mediated by netrin (P00009)	PIK3R3		1
Ionotropic glutamate receptor pathway (P00037)		CAMK2D	1
De novo purine biosynthesis (P02738)		IMPDH2	1
Coenzyme A biosynthesis (P02736)	PPCDC		1
Alpha adrenergic receptor signaling pathway (P00002)		ITPR1	1
Hypoxia response via HIF activation (P00030)	PIK3R3		1
Nicotine pharmacodynamics pathway (P06587)	PPP1CA		1
Ascorbate degradation (P02729)		RPE	1

Heterotrimeric G-protein signaling pathway- Gq alpha and Go alpha mediated pathway (P00027)		ITPR1	1
p53 pathway by glucose deprivation (P04397)	TSC1		1
mRNA splicing (P00058)	PRPF19		1
Heterotrimeric G-protein signaling pathway- Gi alpha and Gs alpha mediated pathway (P00026)		PHKA1	1
Vitamin D metabolism and pathway (P04396)	RXRA		1
VEGF signaling pathway (P00056)	PIK3R3		1
Transcription regulation by bZIP transcription factor (P00055)	TTF1		1
General transcription regulation (P00023)	TTF1		1
Ras Pathway (P04393)		PDK1	1
General transcription by RNA polymerase I (P00022)	TTF1		1
FGF signaling pathway (P00021)	PPP4R1		1
TGF-beta signaling pathway (P00052)	RRAS		1
FAS signaling pathway (P00020)	LMNA		1
Thiamin metabolism (P02780)	ITPA		1
Histidine biosynthesis (P02747)		AADAT	1

Histamine H1 receptor mediated signaling pathway (P04385)		ITPR1	1
Heme biosynthesis (P02746)		RSAD1	1
Cadherin signaling pathway (P00012)	FZD1		1
Blood coagulation (P00011)		PROCR	1
Dopamine receptor mediated signaling pathway (P05912)	PPP1CA		1
Muscarinic acetylcholine receptor 1 and 3 signaling pathway (P00042)		ITPR1	1
Formyltetrahydroformate biosynthesis (P02743)	DHFR		1
Angiotensin II-stimulated signaling through G proteins and beta-arrestin (P05911)		ITPR1	1
Tetrahydrofolate biosynthesis (P02742)	DHFR		1
Metabotropic glutamate receptor group I pathway (P00041)		ITPR1	1
De novo pyrimidine ribonucleotides biosynthesis (P02740)		CTPS2	1

TABLE 3. Panther pathways created using Gene Ontology (GO) database for common genes with similar expression pattern at FDR0.05 in XPA- vs. XPA+ comparison in all four pairs of cell lines used in the study.

The first column consists of the Panther pathways for common differentially expressed genes. Columns on the right consist of the common higher or lower expressed genes for each pathway, and the total number of gene hits observed for each pathway. The pathways for neurological functions are presented in bold.

GO terms for FC2 genes overlapping with 'Mitochondria' (A) or 'Mitophagy' (B) GO genes	Fold enrichment							
	KO142 vs. HeLa S3		KO38 vs. HeLa S3		XP12RO vs. XPA/XP12RO		XP2OS vs. XPA/XP2OS	
	A	B	A	B	A	B	A	B
response to mitochondrial depolarisation (GO:0098780)	27.01	> 100	21.43	> 100	24.6	> 100	28.43	> 100
mitophagy in response to mitochondrial depolarization (GO:0098779)	27.01	> 100	21.43	> 100	24.6	> 100	28.43	> 100
macromitophagy (GO:0000423)	27.01	> 100	21.43	> 100	24.6	> 100	28.43	> 100
mitophagy (GO:0000422)	21.77	> 100	17.27	> 100	19.83	98.69	24.67	> 100
mitochondrion disassembly (GO:0061726)	21.77	> 100	17.27	> 100	19.83	98.69	24.67	> 100
organelle disassembly (GO:1903008)	19.48	> 100	15.45	95.33	17.74	88.3	22.08	94.61
mitochondrion organization	13.66	33.18	11.28	28.66	11.26	26.55	15.17	28.44

(GO:0007005)								
macroautophagy (GO:0016236)	15.53	70.41	10.56	68.58	12.12	60.35	15.09	64.66
autophagy (GO:0006914)	12.43	52.83	8.45	48.02	8.96	42.26	12.07	45.28
regulation of cellular protein localization (GO:1903827)	8.66	-	8.09	-	6.57	-	6.83	-
regulation of establishment of protein localization (GO:0070201)	6.57	-	5.46	-	5.1	-	6.03	-
regulation of cellular localization (GO:0060341)	6.55	-	5.45	-	5.31	-	5.31	-
regulation of establishment of protein localization to mitochondrion (GO:1903747)	23.91	-	29.26	-	21.77	-	25.55	-
regulation of mitochondrion organization (GO:0010821)	19.11	-	20.42	-	21.54	-	23.86	-
xenophagy (GO:0098792)	-	> 100	-	68.79	-	86.48	-	92.66
regulation of immune	-	26.13	-	15.59	-	24.5	-	28

effector process (GO:0002697)								
single-organism organelle organization (GO:1902589)	-	14.02	-	12.11	-	11.21	-	12.02
cellular response to stress (GO:0033554)	-	12.09	-	12.36	-	10.36	-	11.1

TABLE 4. Gene Ontology (GO) terms for genes for ‘Mitophagy’ or ‘Mitochondria’ overlapping with 2 fold or more changed (FC2) genes in XPA- vs. XPA+ comparison in all four cell lines.

The first column consists of the GO terms related to ‘Mitochondria’ or ‘Mitophagy’ from GO database. In the table, ‘A’ represents ‘Mitochondria’ and ‘B’ represents ‘Mitophagy’. Each column ‘A’ and ‘B’ shows the fold enrichment of FC2 genes for each cell line pairs in the GO terms listed in the first column.

Biological Functions	Genes with FC >1.5	Gene Name	XP12RO vs. XPA/XPA 12RO	XP2OS vs. XPA/XP 2OS	KO38 vs. HeLa S3	KO142 vs. HeLa S3
			Fold Change	Fold Change	Fold Change	Fold Change
Metabolism	AKR1C1	Aldo-Keto Reductase Family 1, Member C1	16	48	7	7.5
Metabolism	AKR1C2	Aldo-Keto Reductase Family 1, Member C2	21	20	17	7.5
Metabolism	TMPRSS15	Transmembrane Protease, Serine 15	12	14	2	955
Unknown in normal cells	MAGEA3	Melanoma-Associated Antigen 3	16	11	4	52
Metabolism	AKR1C3	Aldo-Keto Reductase Family 1, Member C3	4	6.9	1.9	2
Mitochondrial function (ATP generation)	NDUFA4L2	NADH Dehydrogenase (Ubiquinone) 1 Alpha Subcomplex, 4-	1.5	2.63	7.5	7

		Like 2				
Immunity	PELI2	Pellino E3 Ubiquitin Protein Ligase Family Member 2	1.7	2.6	1.5	5.5
Neurological development	WFDC21P	WAP Four- Disulfide Core Domain 21, Pseudogene	1.8	2.5	1.8	1.5
Circadian entrainment	HOMER2	Homer, Neuronal Immediate Early Gene, 2	6.2	2	1.9	11
Neural crest development	LHX2	LIM Homeobox Protein 2	9	1.9	1.9	1.4
Auditory maintenance	DFNB31	Deafness, Autosomal Recessive 31	1.5	1.9	1.7	1.6
GPCR pathway	ROR2	Receptor Tyrosine Kinase Like Orphan Receptor 2	-21	-1782	-6.5	-2
Not clear	SH3RF3	SH3 Domain Containing Ring Finger 3	-2.4	-84.4	-1.6	-6.4
Mitochondrial organization	CHCHD10	Coiled-Coil-Helix- Coiled-Coil-Helix Domain Containing 10	-9.8	-42	-1.5	-1.7

Actin binding	CORO6	Coronin 6	-1.5	-32	-1.7	-1.7
Calcium ion binding	RCN3	Reticulocalbin 3	-3.2	-21	-1.5	-1.5
Negative regulation of myoblast	PPAPDC3	Phospholipid Phosphatase 7	-274	-16	-3	-3.6
Negative regulation of MAPK	DUSP4	Dual Specificity Phosphatase 4	-4.3	-6.5	-1.5	-4
Not clear	FCHSD1	FCH And Double SH3 Domains 1	-2.5	-3.5	-1.7	-1.5
Collagen	COL1A1	Collagen Type I Alpha 1	-1.5	-2.5	-1.5	-2
Not clear	GGT1	Gamma-Glutamyltransferase 1	-1.9	-1.6	-1.4	-1.9
TNF pathway	TBKBP1	TBK1 Binding Protein 1	-3.7	-1.6	-2.1	-2.1
Not clear	SYNGR1	Synaptogyrin 1	-1.9	-1.6	-2	-2.5
Matrix metalloproteinase activity	MMP17	Matrix Metalloproteinase 17	-1.7	-1.6	-9	-2.3
Breast cancer	BCAS4	Breast Carcinoma Amplified Sequence 4	-2	-1.5	-2.8	-2
Platelet activation	PROCR	Protein C Receptor	-2	-1.5	-3.4	-1.5

		Pleckstrin Homology Like Domain Family B				
Cytoskeleton	PHLDB2	Member 2	-1.5	-1.5	-1.5	-1.5

Table 5. Common genes that are the most differentially expressed in all four pairs of cell lines.

Table includes the major biological functions in the first column for the gene list given in the second column. The names of the genes are given in the third column. The left four columns are the fold change of the differential expression in four pairs of cell lines in XPA- vs. XPA+ comparison. The positive fold change indicates the higher expression whereas the negative fold change indicates the lower expression in XPA+ compared to XPA-. Genes with positive fold change are given in bold letters.

3.4. XPA status influences expression of genes affecting mitochondria and mitophagy

Because XPA status has been reported to affect mitochondrial function and mitophagy, we examined gene expression specifically in relevant pathways. To narrow down the analysis, we focused on genes with a fold change (FC) of 1.5 or more in each pair of cell lines (Table 2). Genes listed on the GO consortium for 'mitophagy' or 'mitochondria' were overlapped with FC2 genes from our experiment. Significantly, the FC2 genes with XPA-dependent changes common to all four cell lines were highly enriched for GO terms related to mitophagy or mitochondria (Table 4). This confirms and extends the evidence for an influence of XPA status on gene expression affecting mitochondrial maintenance and regulation of mitophagy.

3.5. Classification of the most differentially expressed genes common among all cell pairs

To determine which genes have expression most reproducibly affected by XPA status, we began with the group with at least a 1.5 fold change (FC1.5) in each cell pair. The overlaps in the gene list between these four sets were determined. This narrowed the set to only 27 genes that were differentially expressed at FC1.5 and have a uniform trend of high or low expression among all four pairs of cell lines (Table 5). Among the most differentially expressed

genes were *AKR1C1*, *AKR1C2*, and *AKR1C3*. These genes are adjacent on human chromosome 10p15-p14 and encode (Figure 13) members of the aldo/keto reductase superfamily, discussed further below. There were other AKR family genes also close to these three genes (*AKR1E2* and *AKR1C4*), which were not differentially expressed in XPA- vs. XPA+ analysis. It is possible that there are common regulatory factors for these three genes. Another gene more highly expressed in XPA+ cells was *NDUFA4L2*, which encodes a subunit of an NADH dehydrogenase localized in the mitochondria and important for ATP generation. Other significant “high” genes were *WFDC21P*, *HOMER2*, *LHX2* and *DFNB31*, which have neurological functions (Table 5).

We analyzed the AKR1C proteins in the cell pairs by immunoblotting. Both AKR1C1 and AKR1C2 protein levels were significantly lower in all XPA- cells compared to XPA+ cells (Figure 14A and 14B). AKR1C3 protein levels were slightly higher in HeLa S3 than in the XPA mutants, and absent in some of the fibroblast XPA- and XPA+ cell lines (Figure 14B).

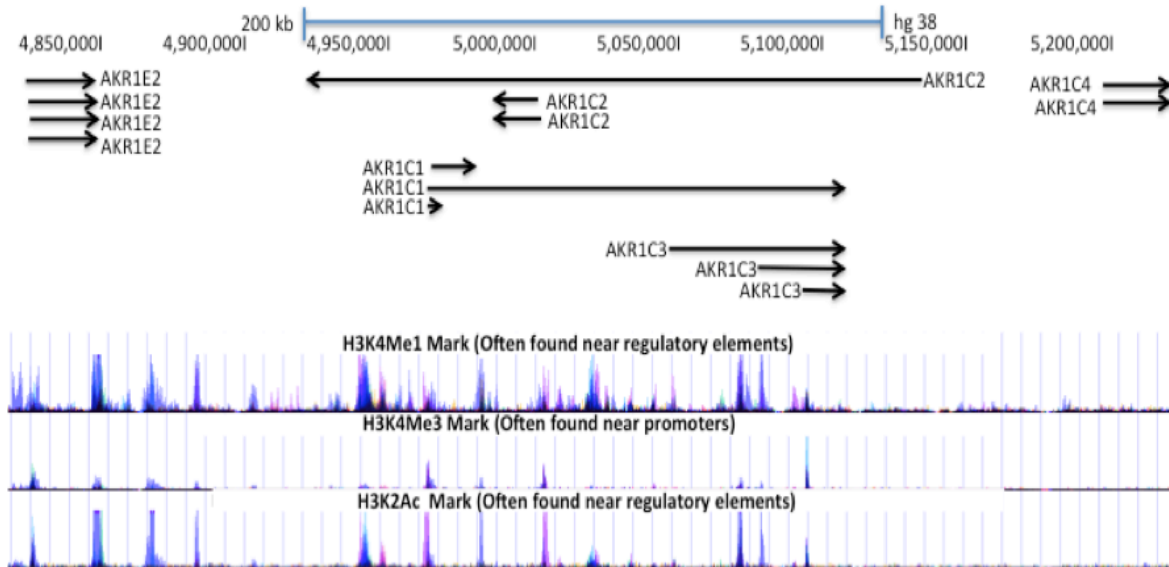


FIGURE 13. Location of *AKR1C1*, *AKR1C2* and *AKR1C3* on human chromosome 10.

AKR1C1, *AKR1C2* and *AKR1C3* are located adjacent to one another and with other AKR family genes; *AKR1E2* and *AKR1C4*. *AKR1C1* and *AKR1C3* have same direction of expression whereas *AKR1C2* has opposite direction. Histone modifications indicate there may be common regulatory factors for *AKR1C1*, *AKR1C2* and *AKR1C3* genes. Different length of the genes represents different variants.

The figure is adapted from the UCSC Genome Browser on Human (Dec. 2013 hg38 Assembly) or Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, Haussler D. The human genome browser at UCSC. Genome Res. 2002 Jun;12 (6):996-1006.

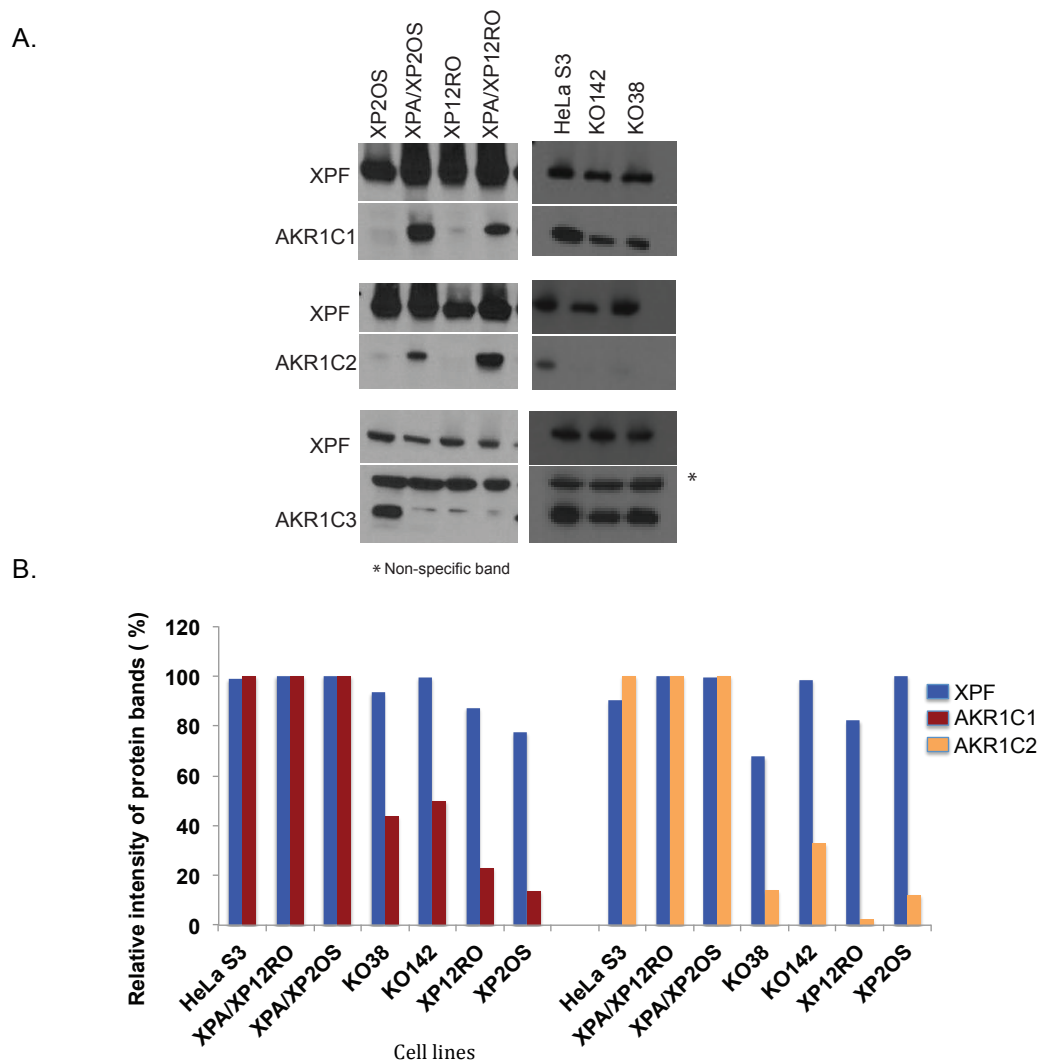


FIGURE 14. Analysis of the most drastically changed gene expressions in XPA proficient compared to the deficient cell lines.

A. Immunoblot for the most differentially high expressed genes (*AKR1C1*, *AKR1C2* and *AKR1C3*) in XPA proficient compared to deficient cell lines. *AKR1C1* and *AKR1C2* protein levels were lower in XPA deficient cells compared to proficient cells whereas *AKR1C3* protein level was inconclusive. B. Quantification of band intensity of *AKR1C1* (red) and *AKR1C2* (orange) and XPF (blue) as loading control.

	KO38 vs. HeLa S3	KO142 vs. HeLa S3	common in both HeLa pairs	XP2OS vs. XPA/XP2OS	XP12RO vs. XPA/XP12RO	common in both fibroblast pairs
High	1311	1184	458	1380	2409	341
Low	1355	1076	346	1744	2163	462
Total	2666	2260	804	3124	4572	803

TABLE 6. More commonalities when fibroblasts and HeLa cells are considered independently.

Number of genes with FC1.5 or more in XPA- vs. XPA+ comparison in fibroblast cell lines or HeLa cell lines shows more common high or low expression patterns compared to commonality among all four cell lines.

3.6. More commonalities within fibroblast and HeLa cell pairs

We also analyzed the data to compare the two skin fibroblast pairs to one another (XP12RO and XP2OS) and the two sets of data derived from HeLa cell (KO38 pair and KO142 pair). Within these pairs, there were many more genes in common that were differentially expressed in an *XPA*-dependent manner (Table 6). For FC1.5 genes, 803 (341 high + 462 low) genes in fibroblast cell lines and 804 (458 high + 346 low) genes in HeLa knockout cell lines had a similar expression pattern, in contrast to the only 27 genes common between all four cell lines at this expression level. These genes were overrepresented for pathways similar to those observed with all four pairs of cell lines with more number of genes (Table 7) and some pathways with less number of genes and are different in fibroblasts and HeLa are listed in Table 8.

3.7. Retinoic acid transactivation affects XPA-dependent transcription related functions

Because retinoic acid transactivation of some genes has been shown to depend on NER components in some instances, we treated cell pairs with all-trans retinoic acid (RA) or a DMSO control. Quantitative PCR was used to assess the response of retinoic acid treatment to the cell lines. All cell lines, both *XPA*-proficient and deficient, responded to RA as shown by an increase in the mRNA level of *RARB*, a direct target gene of RA (Figure 15A). HeLa cell lines

had the lowest transcriptional response of *RARB*. RNA-sequencing data from cells treated with retinoic acid or DMSO for 7 hr were analyzed for biological functions by Ingenuity Pathway Analysis (IPA). IPA analysis was done for the significantly differentially expressed genes (FDR0.05) for DMSO vs. RA, for genes with higher expression in XPA-proficient cell lines compared to their respective XPA-deficient cell lines. This showed that the XPA-dependent RA response is enriched for transcription related biological functions (Figure 15B). Transcription related functions had the highest rank on the list of the $-\log(p\text{-value})$ compared to other biological functions. Thus for retinoic acid transactivation, XPA proficiency might be functionally important. However, we found no common gene expression pattern among all four pairs of cell lines. There were 18 common genes that were FC1.5 or more changed within RA treated HeLa S3 compared to KO38 or KO142. However, no common genes were found in skin fibroblasts in similar conditions. As each cell line responded to retinoic acid treatment in a different manner, it appears that the gene expression changes are also different in different cell lines. Nevertheless, although specific gene expression patterns were different in all four sets of cell lines, common biological functions were represented by the gene expression changes in each pair of cell lines, with transcriptional functions at the top of the list (Figure 15B).

PANTHER pathways	GO term	XPA - vs. XPA+	
		# gene hits in HeLa cells	# of gene hits in Fibroblasts
Wnt signaling pathway	(P00057)	21	9
Huntington disease	(P00029)	14	6
Cadherin signaling pathway	(P00012)	14	9
Integrin signaling pathway	(P00034)	13	12
Inflammation mediated by chemokine and cytokine signaling pathway	(P00031)	13	8
CCKR signaling map	(P06959)	12	9
Gonadotropin-releasing hormone receptor pathway	(P06664)	11	17
PDGF signaling pathway	(P00047)	10	6
EGF receptor signaling pathway	(P00018)	10	10
Angiogenesis	(P00005)	9	7
Alzheimer disease-presenilin pathway	(P00004)	9	10
Nicotinic acetylcholine receptor signaling pathway	(P00044)	9	2
Cytoskeletal regulation by	(P00016)	9	3

Rho GTPase			
Apoptosis signaling pathway	(P00006)	7	4
Parkinson disease	(P00049)	7	2
Interleukin signaling pathway	(P00036)	7	6
FGF signaling pathway	(P00021)	7	6
Blood coagulation	(P00011)	7	2
Heterotrimeric G-protein signaling pathway-Gi alpha and Gs alpha mediated pathway	(P00026)	6	11
Endothelin signaling pathway	(P00019)	6	7
Alzheimer disease-amyloid secretase pathway	(P00003)	5	5
p53 pathway	(P00059)	5	4
T cell activation	(P00053)	5	3
TGF-beta signaling pathway	(P00052)	5	7
Oxidative stress response	(P00046)	5	5
Hypoxia response via HIF activation	(P00030)	4	1
Ras Pathway	(P04393)	4	6
Heterotrimeric G-protein signaling pathway-Gq alpha and Go alpha mediated	(P00027)	4	6

pathway			
B cell activation	(P00010)	4	3
5HT2 type receptor mediated signaling pathway	(P04374)	4	1
Axon guidance mediated by netrin	(P00009)	3	1
VEGF signaling pathway	(P00056)	3	3
PI3 kinase pathway	(P00048)	3	1
Insulin/IGF pathway-protein kinase B signaling cascade	(P00033)	3	1
Oxytocin receptor mediated signaling pathway	(P04391)	3	1
p38 MAPK pathway	(P05918)	3	2
Toll receptor signaling pathway	(P00054)	2	6
Notch signaling pathway	(P00045)	2	3
5HT1 type receptor mediated signaling pathway	(P04373)	2	3
Muscarinic acetylcholine receptor 2 and 4 signaling pathway	(P00043)	1	3
Opioid proopiomelanocortin pathway	(P05917)	1	3

TABLE 7. Cell specific analysis of biological pathways with more number of genes represented in each pathway compared to analysis with all cell lines.

PANTHER pathways significantly represented by common differentially expressed genes ($FC \geq 1.5$) in XPA proficient cells compared to XPA deficient cells among HeLa cells or fibroblast cells. Table consists of the list of the pathways that are common in both among HeLa cells (KO38 vs. HeLa S3 and KO142 vs. HeLa S3) and fibroblasts (XP12RO vs. XPA/XP12RO and XP2OS vs. XPA/XP2OS) excluding the pathways that consists of only one or two genes in both groups.

PANTHER pathways	GO term	XPA- vs. XPA+	
		# Gene hits in Fibroblasts	# Gene hits in HeLa cells
GABA-B receptor II signaling	(P05731)	4	0
Opioid prodynorphin pathway	(P05916)	3	0
FAS signaling pathway	(P00020)	3	0
Alpha adrenergic receptor signaling pathway	(P00002)	2	0
Enkephalin release	(P05913)	2	0
Beta3 adrenergic receptor signaling pathway	(P04379)	2	0
5-Hydroxytryptamine degradation	(P04372)	2	0
Axon guidance mediated by semaphorins	(P00007)	1	0
O-antigen biosynthesis	(P02757)	1	0
N-acetylglucosamine metabolism	(P02756)	1	0
Adrenaline and noradrenaline biosynthesis	(P00001)	1	0
Methionine biosynthesis	(P02753)	1	0
Lipoate_biosynthesis	(P02750)	1	0
Glutamine glutamate conversion	(P02745)	1	0
Coenzyme A biosynthesis	(P02736)	1	0

Synaptic vesicle trafficking	(P05734)	1	0
Endogenous cannabinoid signaling	(P05730)	1	0
Histamine H2 receptor mediated signaling pathway	(P04386)	1	0
Corticotrophin releasing factor receptor signaling pathway	(P04380)	1	0
Proline biosynthesis	(P02768)	1	0
Glycolysis	(P00024)	0	3
Pentose phosphate pathway	(P02762)	0	3
mRNA splicing	(P00058)	0	2
Transcription regulation by bZIP transcription factor	(P00055)	0	2
Nicotine pharmacodynamics pathway	(P06587)	0	2
Salvage pyrimidine ribonucleotides	(P02775)	0	2
Cell cycle	(P00013)	0	2
Androgen/estrogen/progesterone biosynthesis	(P02727)	0	1
JAK/STAT signaling pathway	(P00038)	0	1
Interferon-gamma signaling pathway	(P00035)	0	1

p53 pathway by glucose deprivation	(P04397)	0	1
Heterotrimeric G-protein signaling pathway-rod outer segment phototransduction	(P00028)	0	1
General transcription by RNA polymerase I	(P00022)	0	1
Salvage pyrimidine deoxyribonucleotides	(P02774)	0	1
Angiotensin II-stimulated signaling through G proteins and beta-arrestin	(P05911)	0	1
Histamine H1 receptor mediated signaling pathway	(P04385)	0	1
Pyrimidine Metabolism	(P02771)	0	1
Purine metabolism	(P02769)	0	1

TABLE 8. Cell specific analysis of biological pathways showing difference in HeLa and fibroblasts.

PANTHER pathways significantly represented by common differentially expressed genes ($FC \geq 1.5$) in XPA proficient cells compared to XPA deficient cells among HeLa cells or fibroblast cells. Table consists of the list of the pathways that are different in in HeLa cells (KO38 vs. HeLa S3 and KO142 vs.

HeLa S3) and fibroblasts (XP12RO vs. XPA/XP12RO and XP2OS vs. XPA/XP2OS).

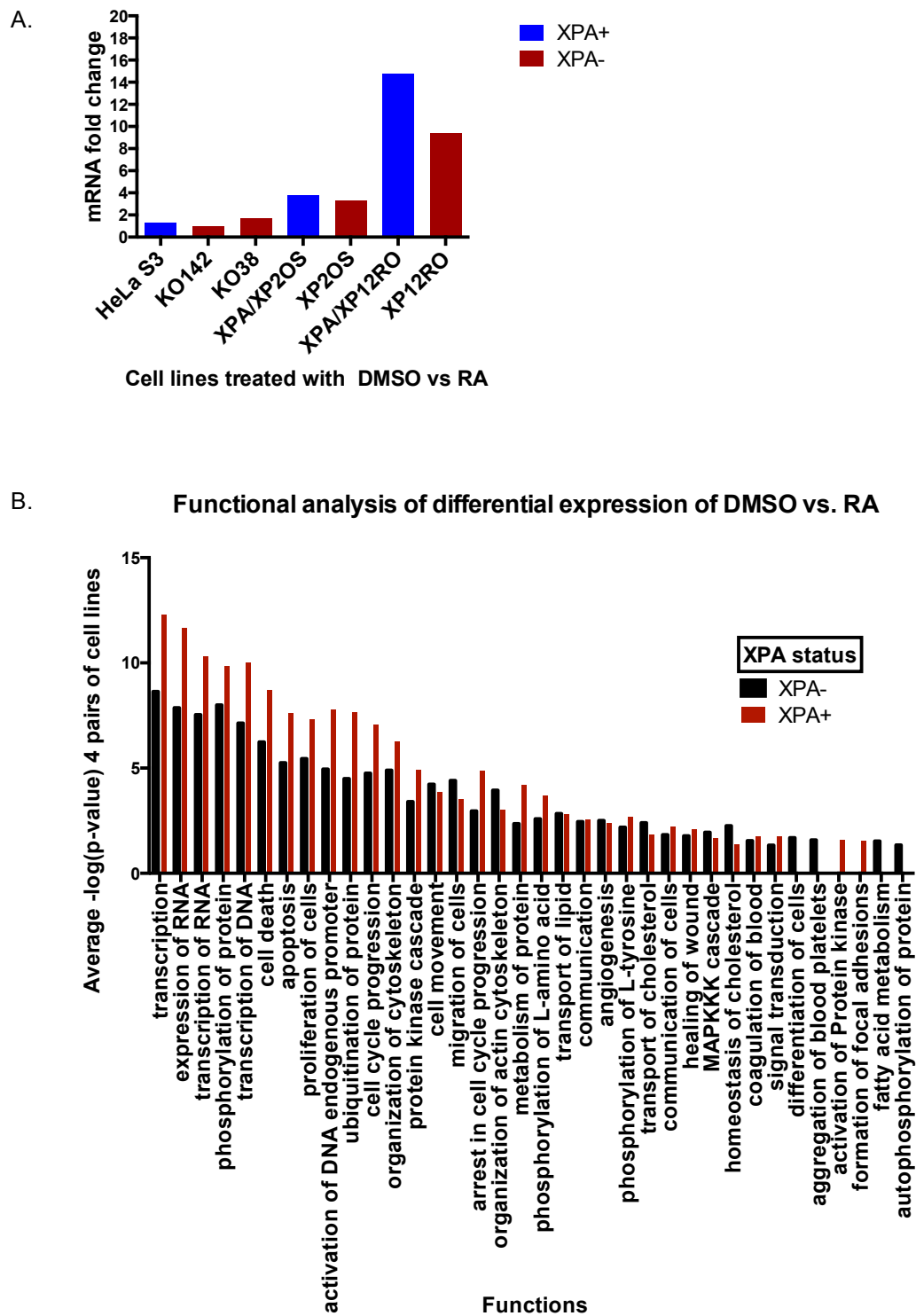


FIGURE 15. Retinoic acid transactivation affects XPA-dependent transcription related functions.

A. qPCR results for all the XPA proficient and deficient cell lines showing responses to RA treatment at 7 hours. *RARB* gene expression was analyzed for RA or DMSO treated cell lines. GAPDH was used as control. The mRNA of *RARB* in RA treated condition was normalized to the DMSO- only controls. B. IPA analysis comparing the RA treated to untreated cells. Genes that were differentially expressed at FDR0.05. XPA proficient cell lines (red) compared to the XPA deficient cell lines (black) were analyzed. The scores are presented as negative log values of p-values, showing that XPA+ cells have higher significant differences than in XPA- cells, comparing the response to RA treatment with control treatment for a given function shown on the x-axis.

Chapter 4. Results - Processing of a TFO-directed interstrand crosslink in human cell extracts

ICL causing agents are used in chemotherapies to kill cancer cells but the exact mechanism for the processing DNA ICLs in human cells is not known. This project explored the mechanism of processing DNA-ICLs in human cells. For this purpose we first established a control NER dual incision assay with a 1,3 d[GpTpG] cisplatin intrastrand crosslink built into a closed circular double stranded M13mp18GTGX DNA (102). The substrate was confirmed, and cell extracts and the reaction conditions were established to show the control assay works well for the incision and repair assays. To map out the incision sites on the psoralen-TFO ICL substrate we constructed a site-specific psoralen ICL in a closed circular DNA with a psoralen conjugated to a TFO. We used this substrate in the incision and repair assays to study the processing in human cell extracts.

4.1. Validation of the cisplatin intrastrand substrate by primer extension assay

The purified cisplatin intrastrand crosslink was used in a primer extension assay to detect blockage of the extension of DNA polymerase. The lesion successfully blocked the extension process of the Sequenase polymerase (Figure 16). A *PvuI* restriction digested fragment encompasses the cisplatin intrastrand adduct. The fragment is 201 bp long with a primer-annealing site at

one end. The adduct is 39 nucleotides away from the 5' end of the primer. As expected, we observed that the restriction digested fragment extended to 201 bp when no adduct is present and the adduct-obstructed fragment extended to only 39 bp, confirming the presence of the cisplatin adduct at the specific site.

We used two kinds of controls. One is the replicative form of the M13 single stranded plasmid, which is a double stranded circular DNA. The other control plasmid was made similar to the cisplatin-damaged plasmid, except an unplatinated oligonucleotide was used instead of cisplatin containing oligonucleotide. In both controls there was full extension of the radiolabelled primer to 201 bp. Some extended 201 bp fragments were also observed with damaged plasmid. This may be due the presence of small amount of unplatinated plasmid.

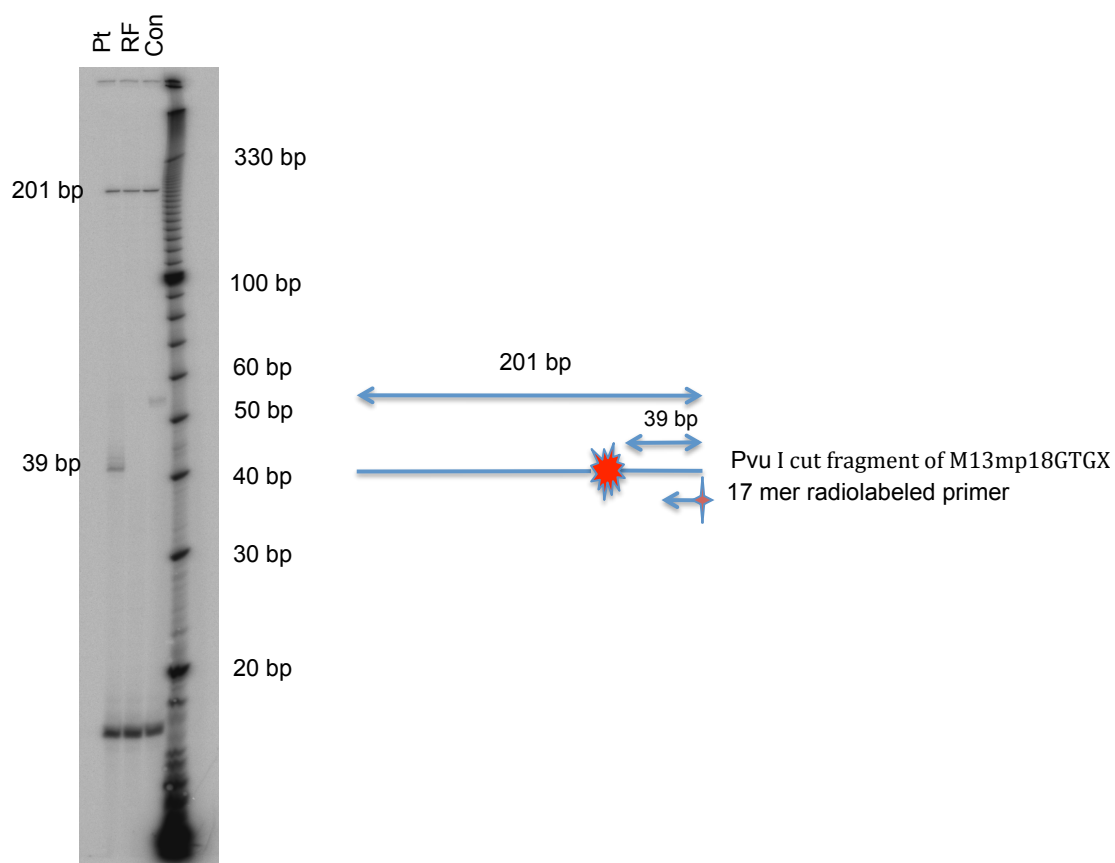


FIGURE 16. Primer extension assay using a fragment of M13mp18GTGX from the 5' end of the radiolabelled primer to the 3' end cut by PvuI enzyme.

This fragment is 201 bp long. Denaturing 14 % polyacrylamide gel showing control (con) and replicative form of M13mp18GTG (RF) extended to 201 bp and cisplatin substrate (Pt) blocked by the adduct at 39 bp.

4.2. NER dual incision control assay was established

4.2.1. Major incision products and minor incision products were observed as shown by previous studies

A 1,3d[GpTpG] cisplatin intrastrand crosslink built into a closed circular double strand M13mp18GTGX DNA has been shown to be processed by human cell extracts in NER dependent dual incision (102). It has been reported that human cell extracts make major incisions between the 8th and 9th nucleotide at the 3' side and between the 15th and 16th nucleotide at the 5' side, resulting in the major incision products of this cisplatin adducts in human cells of 27-32 nucleotide long (102). This substrate was used as a control for the psoralen interstrand crosslink substrate. Using the similar conditions of reactions that were used in Moggs et al, 1998 (102), we performed a dual incision assay and mapped the incision product by using the G-overhang oligonucleotide labelling method. We successfully detected incision products that were produced due to dual incision on either side of the cisplatin adduct using human cell extracts. The incision products produced were between 25 to 35 nucleotides long (Figure 17). The major incision products were observed between 27-29 nucleotides long. The dual incision result in Mogg's et al, 1998 (102), where the incisions were 15 nucleotides away from the adduct at the 5' side and 9 nucleotides away from the adduct on the 3' side was reproduced in our dual incision assay. This result

confirmed that the cell extract, the DNA substrate, and the reaction conditions we used are all functional for the incision reactions.

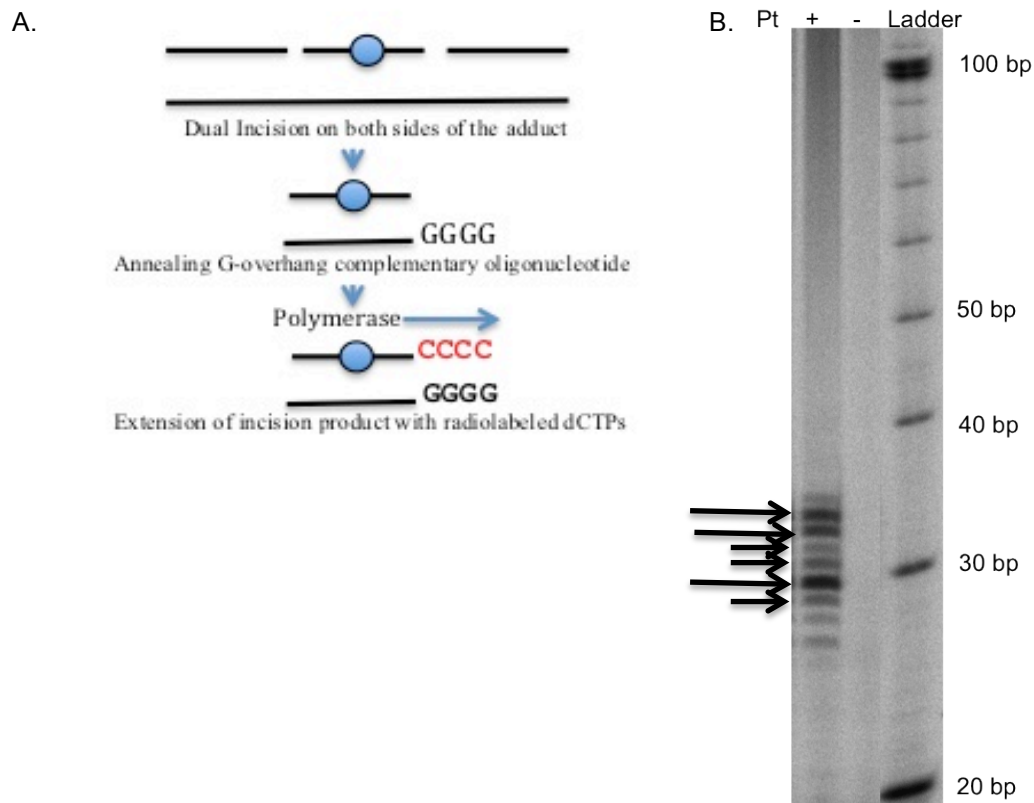


FIGURE 17. **NER dual incision control-assay.**

A. Detection of incision products by the G-overhang end labeling method, which adds 4 nt to the primary incision product length as described in the Chapter 2. B. Denaturing 14% polyacrylamide sequencing gel with major and minor incision products plus 4 radiolabelled CTPs, indicated by long and short arrows respectively, from NER dual incision control assay. The major labeled products of 30, 33 and 34 nt correspond to the major primary incision products of 26, 29 and 30 nt as observed previously (102).

4.2.2. A 1,3 d[GpTpG] cisplatin intrastrand crosslink is repaired by HeLa whole cell extract as observed by previous studies

We wanted to confirm that the reaction conditions that we have established can be used for repair synthesis with the cisplatin substrate, as it is known that this substrate can be repaired well in human cell extracts provided with the supply of necessary dNTPs. We performed the repair synthesis assay with similar reaction conditions as in the Moggs et al., 1996 paper (102). HeLa S3 cell extracts can repair the cisplatin substrate by filling the gap created by dual incision with radiolabelled dNTPs. Then, the plasmid was digested with *BstN1*, which creates 8 different size fragments in M13mp18GTGX DNA, among which a 33 bp fragment is where the repair synthesis occurs. We observed this fragment with the highest intensity by autoradiography (Figure 18). Some neighbouring fragments were also labelled. These fragments were fainter than 33 bp fragments indicating that the more radiolabelled dNTPs preferentially were incorporated into the 33 bp region. This result suggests that the repair process has occurred, which assured us that we could use this assay to study the repair of TFO-directed ICL substrate.

4.2.3. Incisions on the 1,3 d[GpTpG] cisplatin intrastrand crosslink were dependent on NER protein

Dual incision of the cisplatin substrate is mediated by NER (102). An XPA deficient HeLa S3 (KO38) was made by CRISPR-Cas9 technology. We prepared

the KO38 whole cell extracts and used purified XPA protein to complement the deficiency in the dual incision assay. We observed the similar dual incision products as with XPA proficient HeLa S3 whole cell extract in the complemented assay. No products were detected in KO38 whole cell extract without complementation, showing that the incisions are dependent on NER. The result is presented in Section 3.1 and Figure 9B.

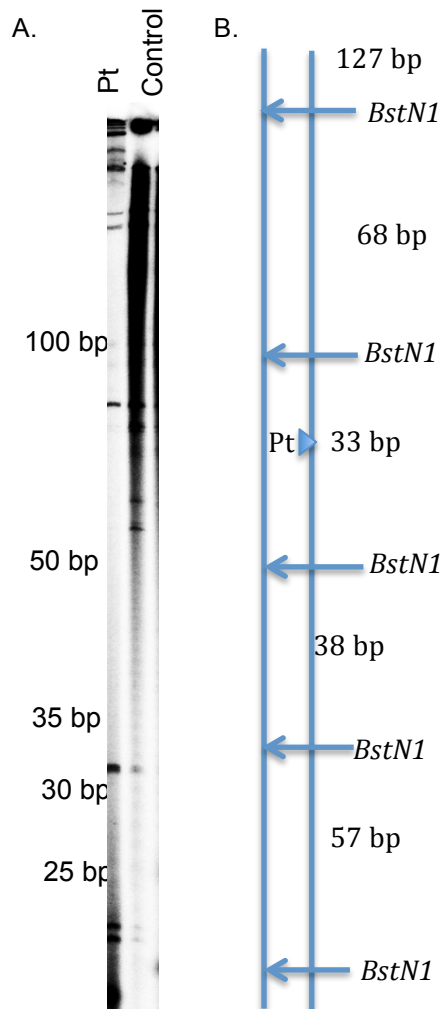


FIGURE 18. **Repair synthesis assay.**

A. Denaturing 14% polyacrylamide gel showing *BstN1* digested fragments of cisplatin and control substrate after repair synthesis. A 33 bp long band encompassing the cisplatin substrate was most intense compared to the control, showing incorporation of radiolabelled dNTPs at the repair site. B. Schematic diagram of restriction digestion sites around cisplatin adduct.

4.3. TFO-directed Psoralen interstrand crosslink was constructed and purified

The TFO –directed psoralen interstrand crosslink in a closed circular double stranded DNA was constructed, purified and used as an ICL substrate in the experiments to determine how it is processed by mammalian cells. Construction of pure and high yield ICL substrate has been difficult task for ICL repair studies. The 50 bp short duplex TFO-directed psoralen ICL substrate was made using technology and reagents developed by Dr. Karen Vasquez's laboratory (25). Then we developed the methodology to construct TFO- directed psoralen ICL in closed circular DNA (Figure 19, see methods in Chapter 2 for details). Confirmation of the components of the TFO-psoralen ICL is presented in the following sections.

4.3.1. 50 bp psoralen-TFO ICL insert was constructed and purified

A 50 bp TFO-directed psoralen ICL migrates slowly on a denaturing 6% polyacrylamide gel because it represents two strands linked together with a triplex forming oligonucleotide and psoralen. We observed the psoralen-TFO ICL at around the 150 bp (Figure 20). After purifying the UVA mediated psoralen-TFO ICL in the 50 bp duplex DNA, we confirmed the presence of interstrand crosslink, by using UVC photoreversal. UVC exposed ICL sample was observed to reverse

back to monoadduct and linear 50 bp DNA (Figure 20). This is the confirmation of the formation of the ICL.

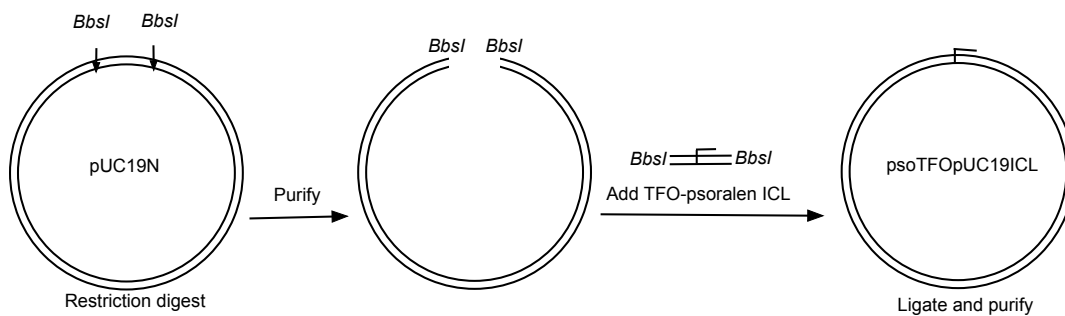


FIGURE 19. Strategy for the construction of TFO- directed psoralen ICL in closed circular double stranded DNA.

pUC19 plasmid was modified to include a *BbsI* restriction digestion site. pUC19N plasmid was constructed and purified. By restriction digestion with *BbsI* enzyme, vector was created and purified. A 50 bp duplex DNA was annealed with TFO conjugated with HMT psoralen at its 5' end and was crosslinked by UVA photoactivation to make ICL insert. The vector and insert were ligated in appropriate condition to get a circular double stranded DNA with psoralen-TFO ICL.

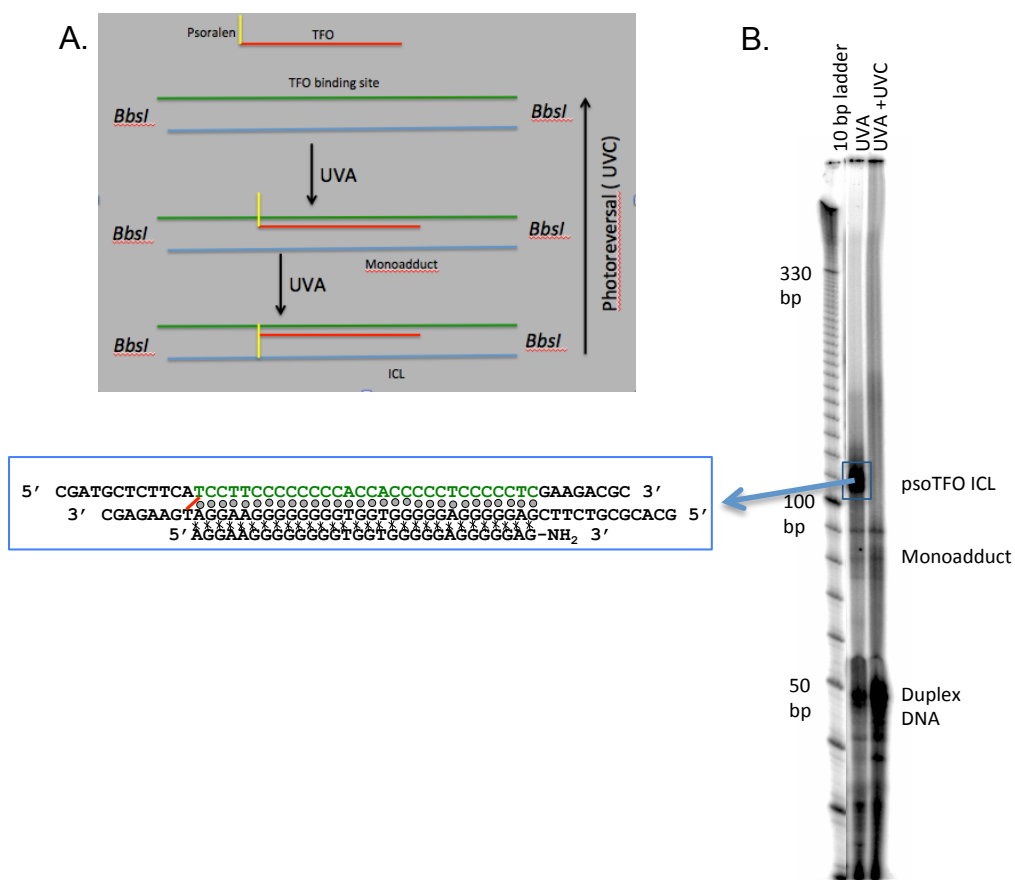


FIGURE 20. 50 bp TFO directed psoralen ICL.

A. Process of making a TFO-directed psoralen on a short duplex DNA. B. 6% polyacrylamide gel showing the formation of the ICL, monoadduct, and remaining duplex DNA. The conversion of ICL to monoadduct and duplex DNA by UVC photoreversal is the confirmation of ICL formation as shown in the right lane. The nucleotide make-up of the psoralen-TFO ICL is shown in the inset. Green represents the nucleotides of TFO binding site. The third strand is the TFO bound to the first two strands (duplex DNA) with Hoogsteen hydrogen bonding represented by the asterisk (*). Gray dots represent Watson and Crick hydrogen bonding.

4.3.2. Vector with *BbsI* sites was prepared

BbsI restriction digestion sites were inserted into the pUC19N plasmid. *BbsI* has unique sequences at both 5' side and 3' side which reduces the chances of defective ligation. After purifying modified pUC19N, we tested if the *BbsI* site is incorporated at the correct site in pUC19N plasmid by running the *BbsI* digested sample on denaturing 6% polyacrylamide gel (Figure 21). A 50 bp band was released from the pUC19N plasmid. The remaining linear pUC19N was observed near the well.

4.3.3. Vector and psoralen-TFO ICL insert was successfully ligated in small and then in large scale

We did several small-scale ligation reactions in order to find the best condition that gives the highest yield of the final ligated product. Vector to insert ratio of 1:3 to 1:5 gave high yield of ligated final product in small-scale ligation reaction (Figure 22). A small amount of unligated or linear products and some multimers were observed. A 1:5 vector to insert ratio was utilized, and a large-scale ligation reaction was performed (Figure 23). The substrate was purified by cesium chloride gradient to separate and purify the closed circular double stranded DNA (Figure 23). A 35 % yield of final ligated product was obtained.

4.3.4. Confirmation of the ICL in the final product

After the final product was purified, we tested if it retained the TFO directed psoralen ICL. Small volume of the final product was restriction digested with *BbsI* and ran on a sequencing gel. The fragment released from the restriction digestion was the ICL fragment (Figure 24). 50 bp linear duplex DNA was used as a control marker. This confirmed that the final ICL product is correct and pure TFO directed psoralen interstrand crosslink in closed circular double stranded DNA.

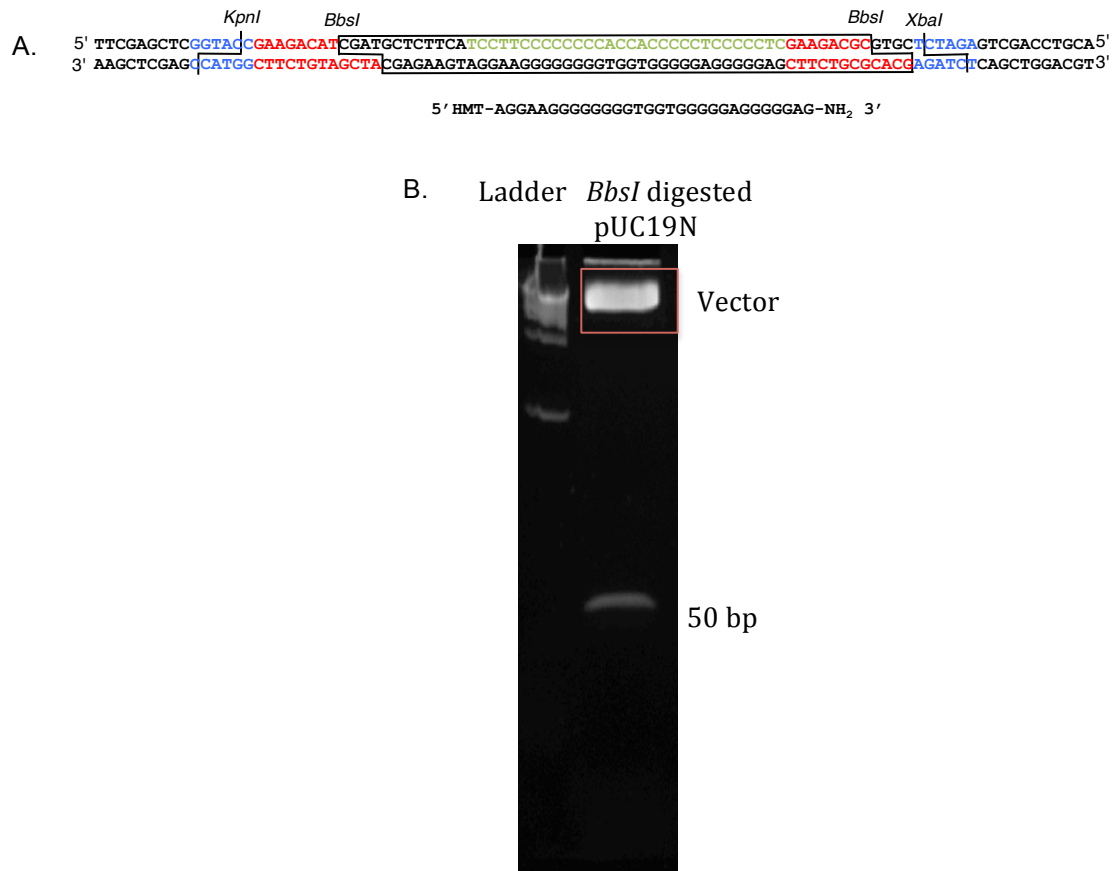


FIGURE 21. **Vector with *BbsI* sites.**

A. Fragment of duplex DNA inserted between *KpnI* and *XbaI* sites in pUC19 in order to insert *BbsI* restriction digestion site. Fragment with *BbsI* site is shown in the inset. Blue and Red indicate the nucleotides for respective restriction digestion sites mentioned in the diagram. Green indicates the TFO binding site. Modified plasmid after the new insertion is called pUC19N. B. Denaturing 6% polyacrylamide gel with restriction digested pUC19N. Linear vector is shown in the inset and the lower band is the 50 bp fragment released for the restriction digestion.

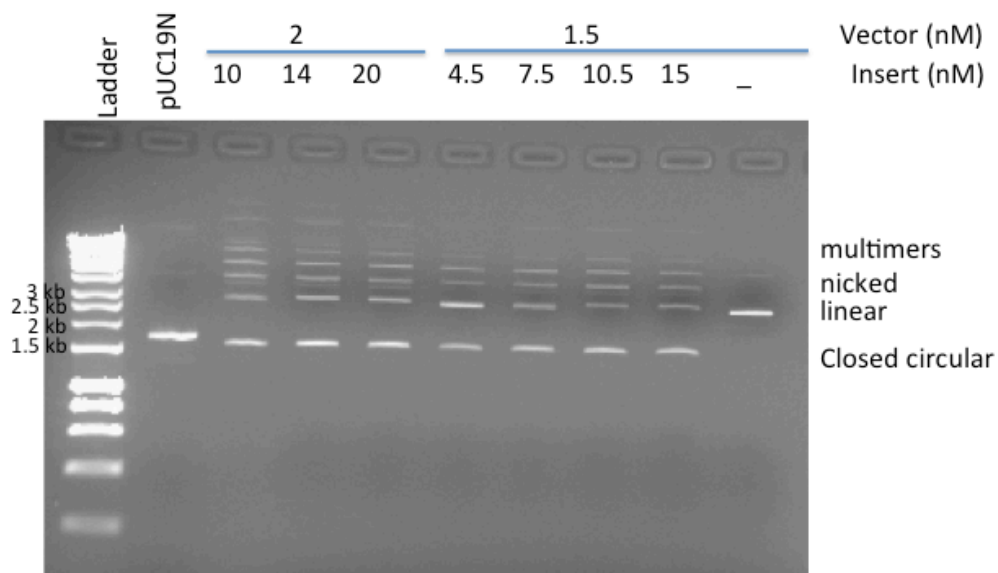


FIGURE 22. Small scale ligation reaction with BbsI-vector and 50 bp pso-TFO ICL.

Different ratios of vector to insert were tested to get the highest yield of the ligated product. The final ligated product is a closed circular double-stranded DNA with a TFO directed psoralen interstrand crosslink. Purified pUC19N and purified vector were used as controls. The final product migrated in 1% agarose gel similar to pUC19N (2739 bp). Some unligated linear vector and some multimers were also observed in lesser quantity compared to the ligated closed circular DNA.

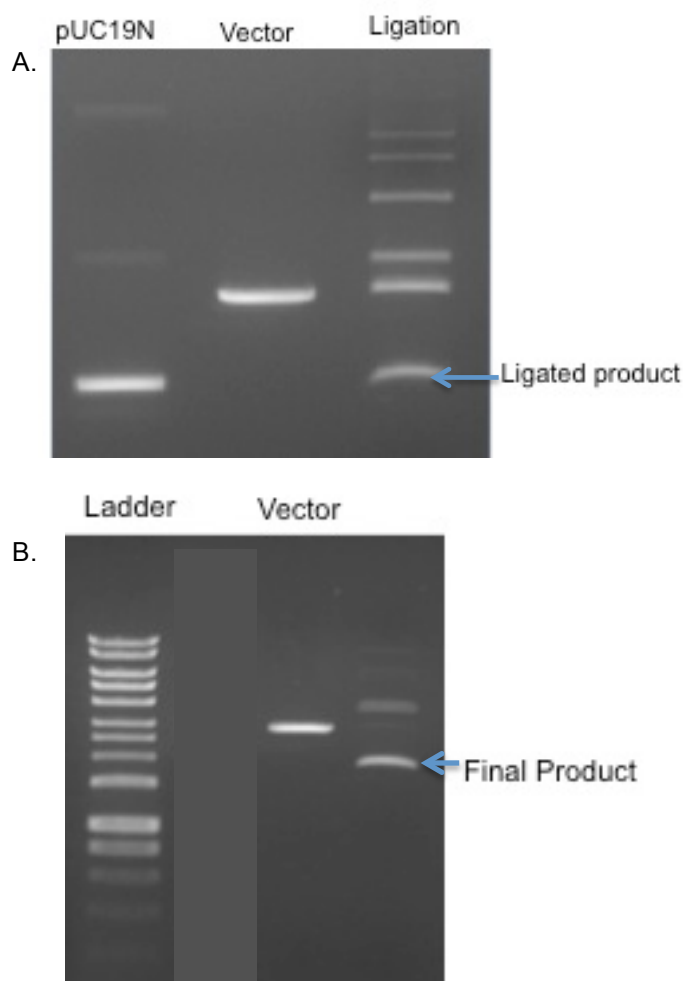


FIGURE 23. Large-scale ligation reaction of BbsI-vector and 50 bp psoralen-TFO ICL insert.

A. Ligation reaction on a 1% agarose gel right after the ligation reaction. Most of the products were closed circular DNA, indicated by the arrow. The bands above the final product are linear and multimer DNA. B. Large scale ligation reaction on a 1% agarose gel after purification by cesium chloride gradient. The linear and multimer DNA were reduced largely after purification.

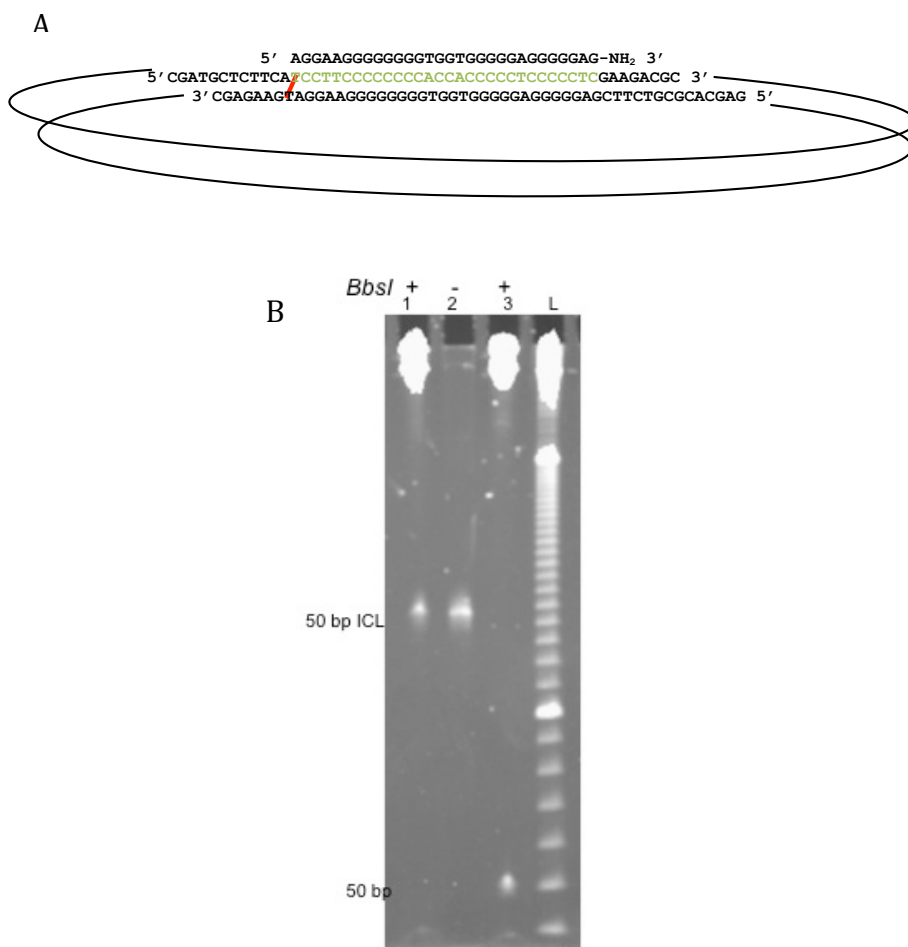


FIGURE 24. Confirmation of the final purified TFO- directed psoralen ICL in a closed circular duplex DNA.

A. Diagram of the final ICL substrate with the nucleotide arrangement of TFO and duplex DNA. Green represents the nucleotides of the TFO binding site. TFO bound to the first two strands (duplex DNA) with Hoogsteen hydrogen bonding represented by the asterisk (*) marks. Gray dots represent Watson and Crick hydrogen bonding. Red represents the psoralen bridge between two thymines. B. Restriction digested final product was run on a denaturing 6% polyacrylamide gel. The released product ran as 50 bp psoralen-TFO products. 50 bp duplex DNA was also used as control.

4.4. Plasmid relaxation assay

The closed double stranded DNA is the relaxed form of the plasmid. Experiments were performed to determine if the whole cell extracts process the supercoiled or relaxed form of DNA substrates differently. We incubated the plasmids in the presence and absence of whole cell extracts, and in NER reaction buffer and ran in gels with and without ethidium bromide. We found that the whole cell extract relaxes the plasmid within 5 minutes of incubation (Figure 25). Whole cell extracts convert the supercoiled plasmid to the relaxed plasmid, which runs slightly faster than supercoiled plasmid in agarose gel without ethidium bromide (Figure 25A). When gel is run with ethidium bromide, ethidium bromide gets intercalated between the two strands of the relaxed plasmid, which makes it run slower than the supercoiled plasmid (Figure 25B).

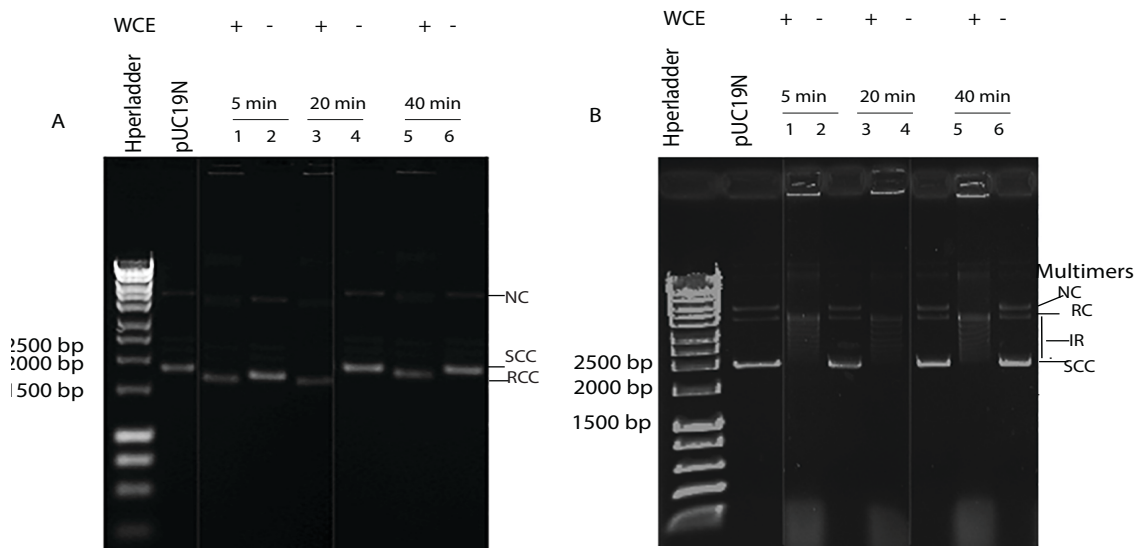


FIGURE 25. Plasmid relaxation assay.

pUC19N plasmid was incubated with or without whole cell extract in an ATP regenerating system, purified, and ran in 1% agarose gels with or without ethidium bromide. A. 40 $\mu\text{g}/100\text{ mL}$ ethidium bromide was run for 2 hours at 80 V. B. without ethidium bromide; run for 2 hours at 80 V, and stained with 40 $\mu\text{g}/100\text{ mL}$ ethidium bromide for 20 minutes. pUC19N and Hyperladder were ran as markers. SCC indicates supercoiled closed circular, RCC: relaxed closed circular, NC: nicked circular, IR: intermediate relaxed DNA.

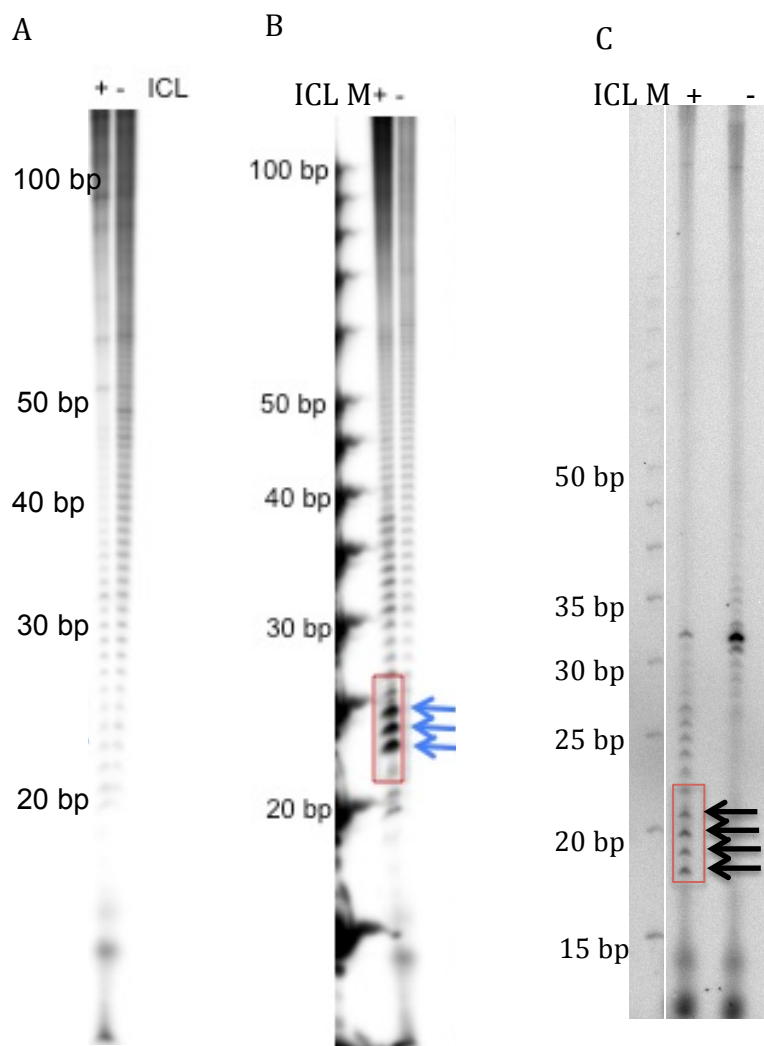


FIGURE 26. 3' Incision assay on psoralen-TFO ICL substrate.

14% sequencing gel showing 3' incision products on upper strand with G-overhang labelling oligonucleotides to detect incision products at TFO binding site (A), beyond TFO binding site (B) and at 3' side on the lower strand (C). Sequence mapping of the incision sites is shown in Figure 13. Darker bands marked with arrows are the major incision products. M denotes marker. Lanes with ICL substrate or with undamaged plasmid are labelled.

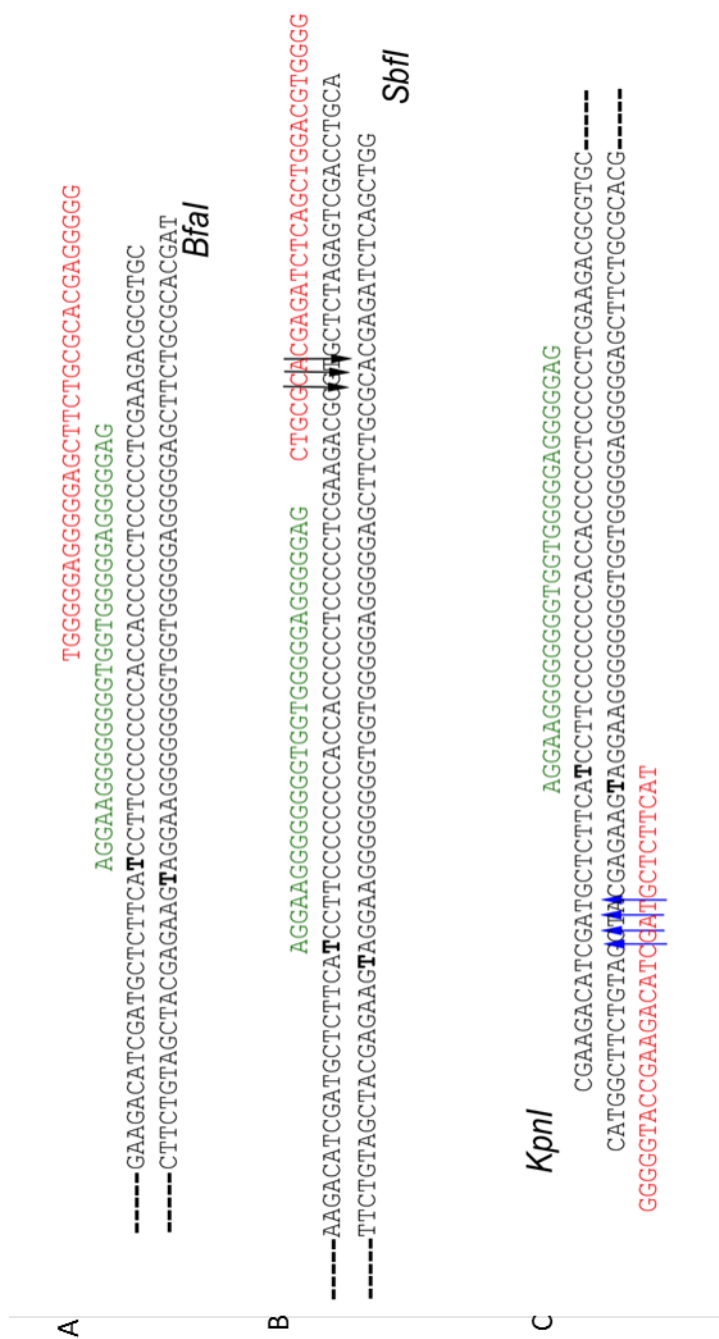


FIGURE 27. Sequence-map of the incision sites.

Positions of incision sites are marked by the arrows and are corresponding to and color coded with the incision products shown in Figure 12. 3' incision sites on the upper strand within TFO binding site are shown in (A) and beyond TFO binding site in (B), and lower strand in (C). Green sequences are TFO and red sequences are G-overhang oligonucleotides. Restriction digestion enzymes used in each experiment to prepare the ends for G-overhang complementary labelling with radiolabelled dCTPs are mentioned at respective DNA ends. Thymines involved in psoralen ICL formation are presented as bold 'T'. Dashes denote the continuity of the sequences.

4.5. Incisions were observed on the 3' side of psoralen-TFO ICL on both strands

Incision assays were performed using a G-overhang oligonucleotide to detect the incision products at the 3' side of the ICL crosslink. G-overhang oligonucleotides were used to detect the incisions at two different sites. One covered the site inclusive of TFO binding site and the other covered the area outside of TFO binding site. Incision products were not detected with the oligonucleotide at the TFO binding site, which suggests that the incisions did not occur at the TFO binding site (Figure 26A and 27A).

G-overhang oligonucleotide end labelling detected incision products were represented by three dark bands and several faint bands on the upper strand. The bands were 22, 23 and 24 oligonucleotides long, and were mapped 10, 11, and 12 nucleotides away from the TFO, and 44, 45, and 46 nucleotides away from the psoralen ICL (Figure 26B and 27B). These dark bands were absent on the undamaged DNA substrate indicating that these incision products were specific to psoralen-TFO ICL on closed circular DNA. Similarly, 3' incision assays were performed targeting possible incision products on the lower strand of the psoralen-TFO ICL substrate. Several dark bands appeared in the ICL substrate lane, which were not present in the undamaged substrate lanes. The bands were observed as 18, 19, 20 and 21 oligonucleotides long and were mapped at 8, 9,

and 10 nucleotides away from the ICL site (Figure 26C and 27C). The repetition of the assays also showed several dark bands around the same sites.

4.6. Psoralen-TFO ICL was not repaired by HeLa whole cell extract

Repair assay based on the translesion synthesis at the damaged site by the repair polymerases was performed in order to see if the incisions observed in above experiments were for the repair-purpose. In other words, by repair assays, we wanted to see if the human cell extract can repair TFO-directed psoralen ICL in closed circular double stranded DNA. Repair assays were done in similar reaction conditions as incision assays except that the reactions included radiolabelled dNTPs to be incorporated by the polymerases for the repair. The repaired products were restriction digested, and fragments were analysed if they incorporated radiolabelled dNTPs. After the repair reaction, restriction digestion enzymes (*AluI* and *SapI*) were used to create different sized fragments to make it easier to be analysed on sequencing gel. 29 bp and a 69 bp fragments from the restriction digestion were supposed to have incorporated more radiolabelled dNTPs compared to neighbouring fragments and compared to undamaged DNA. We did not observe any 29 bp bands in ICL substrate however, it was observed in undamaged plasmid (Figure 28). *SapI* sequence was present at the site of psoralen ICL. Thus it is possible that the ICL substrate was not repaired and was resistant to *SapI*. Since a 69 bp band was not resolved well in a 14 % sequencing gel, it could not be analysed.

We performed another repair assay with the same principal as the above repair assay except that the restriction digestion enzymes were different. This time, we used *BbsI* and *MluCI*. A 50 bp band was supposed to have been repaired, and have incorporated radiolabelled dNTPs as it included the site of damage. Unlike our expectation, a 50 bp band was missing from ICL substrate whereas other neighbouring bands; 25 bp and 68 bp bands were observed (Figure 29). These fragments were possible only if both *BbsI* and *MluCI* cut at their respective sites. This indicated that *BbsI* could cut at the designated site and released the fragment consisting of ICL but no new dNTPs were incorporated for repair at this fragment. Repair synthesis assay for cisplatin substrate in HeLa whole cell extract as described earlier was used as control assay. Control assay showed a 33 bp high intensity fragment indicating the successful repair synthesis happening on cisplatin substrate as described earlier. This result was reproduced in repeated experiments indicating that TFO-directed psoralen ICL was not repaired by HeLa whole cell extract.

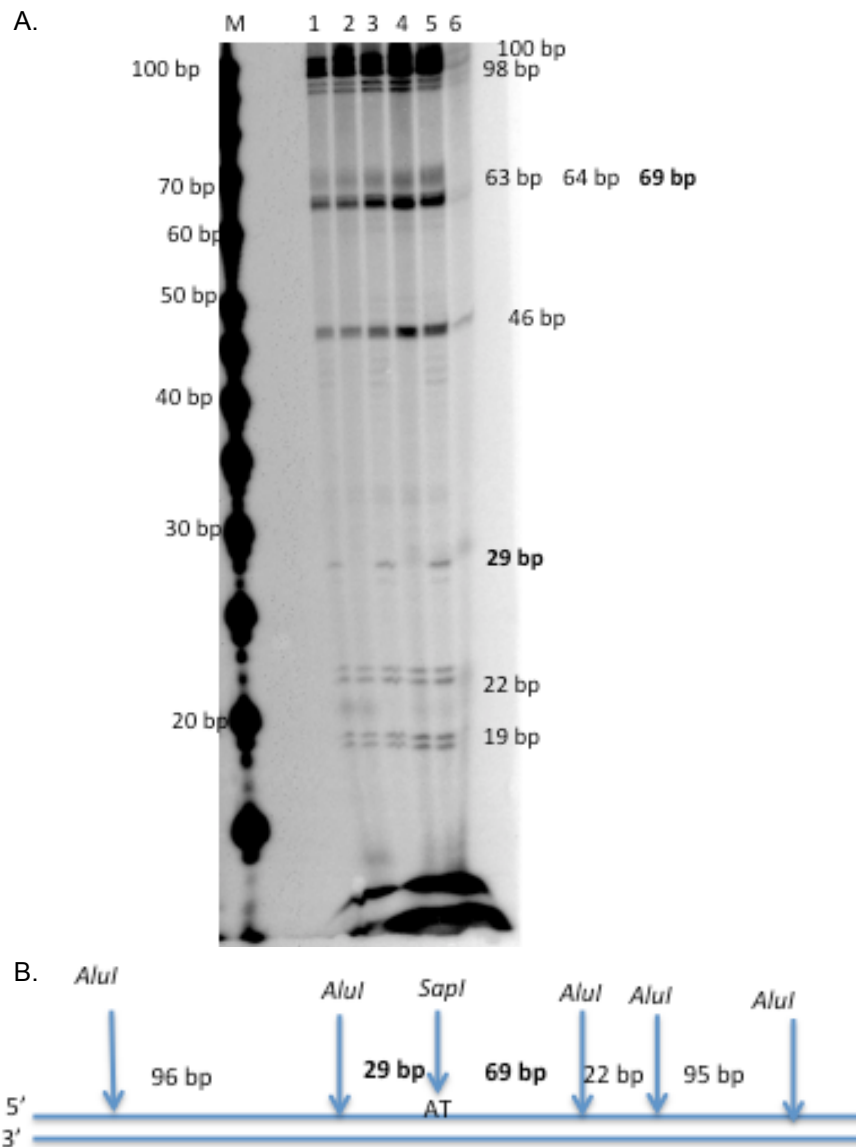


FIGURE 28. **Repair synthesis assay 1.**

A. 14 % sequencing gel showing restriction digested fragments after repair synthesis assay on psoralen-TFO ICL or pUC19N plasmid. Lanes 2, 4 and 6 are repair reactions with ICL substrate for 30 min, 1 hr and 3 hr, and lanes 1, 3 and 5 are with undamaged DNA for 30 min, 1 hr and 3 hr respectively. B. Map of fragments made by *AluI* and *SspI* enzymes. AT is the site of ICL formation where *SspI* sequence coincides.

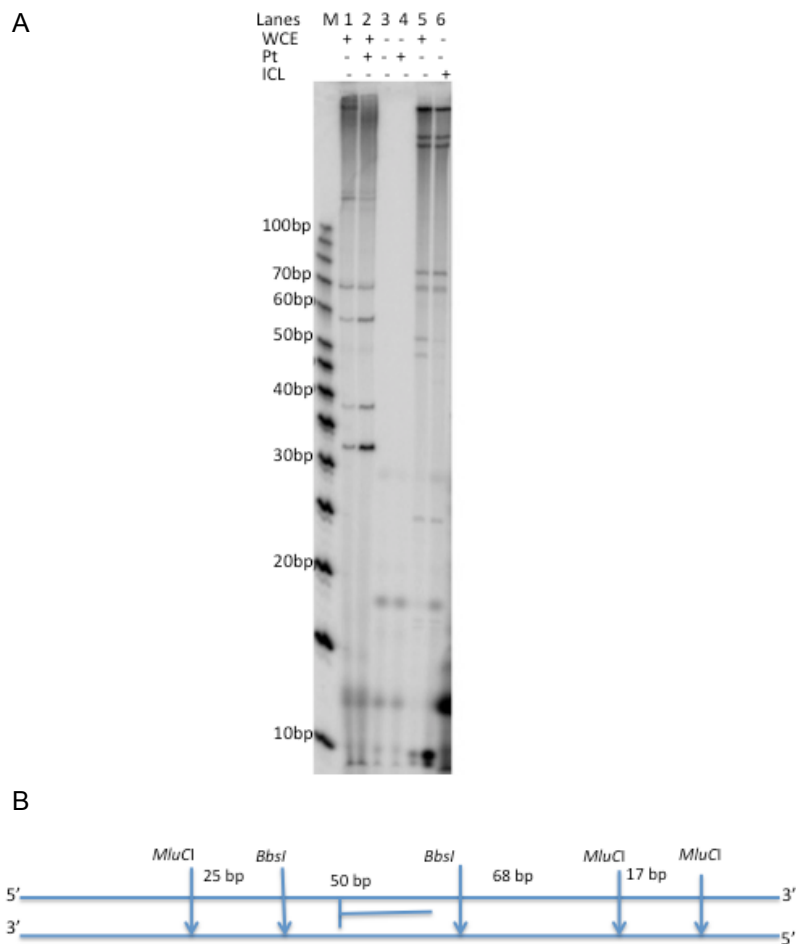


FIGURE 29. **Repair synthesis assay 2.**

A. 14 % sequencing gel showing restriction digested fragments after repair synthesis assay on psoralen-TFO ICL or pUC19N plasmid (Lanes 5,6) or cisplatin intrastrand crosslink substrate (Pt) or undamaged DNA as control for Pt substrate (Lanes 1-4) as control. Lanes 3 and 4 were performed without whole cell extract for control. Purified DNA from repair assay with or without Pt samples were restriction digested with *BstN1* (lanes 1-4) and with or without ICL substrate was restriction digested with *BbsI* and *MluCI* (Lanes 5 and 6). B. Map of fragments created by *BbsI* and *MluCI* on ICL substrate.

Chapter 5. Discussion

5.1. Consequences of XPA disruption in human cells

5.1.1. Expression of a subset of genes is influenced by XPA status in cell lines

The major purpose of this study was to determine the extent to which XPA status may influence transcriptional activity in cells. As we show here, expression of a restricted subset of genes is reproducibly affected by XPA status. These include genes that are necessary for steroid metabolism and for optimal mitochondrial function and integrity.

A plausible explanation for a positive effect of XPA on gene expression is that XPA acts as part of a transcription factor complex, for this subset of genes. For example, by binding to DNA and proteins in concert with the transcription initiation factor TFIIH and other NER proteins, it may modulate TFIIH activity for a subset of genes. As XPA is a DNA binding protein, it might also serve as a component of an enhancer or mediator complex for some genes.

Another, not necessarily exclusive possibility is that the NER defect caused by XPA deficiency gives rise to constitutive DNA damage stress that interferes with transcription of some genes. Persistent sites of DNA damage on

templates will interfere with transcription. For individual genes that are more susceptible to DNA damage, this could account for the higher transcription of some genes in *XPA*-proficient cells. Graphic explanation of some possible mechanisms by which XPA may be involved in transcription is presented as a model in Figure 30.

What are possible sources of DNA damage that would result in constitutive stress for XPA defective cells? A low yield of UV radiation-induced pyrimidine dimers could potentially accumulate in XPA-defective cells from UVA radiation that can penetrate plastic dish lids. However, we take care to culture cells in low ambient light or with yellow fluorescent lights, and this seems unlikely to be a significant source of stress to explain our data. A ubiquitous and constitutive source of damage is the lesions caused by reactive oxygen species, constantly produced in cells. Reactive oxygen species induce many lesions that are largely repaired by XPA-independent BER. However, one type of oxidative damage, the 8,5'-cyclopurine-2'-deoxynucleosides (cyclopurines), are repaired by NER but not by BER (78, 79). These lesions are known to interfere with transcription.

Oxidative damage, particularly cyclopurine lesions, has been previously suggested as an explanation for the neurodegeneration in XP patients (76, 78, 120), which includes progressive cerebral atrophy and hearing loss. Most

neurons are non-replicating and metabolically very active, depending on mitochondria to remain functional.

On the other hand, there are contradictions about oxidative damage accumulation in XPA mutant causing of neurodegeneration and accelerated aging as shown by studies in *C. elegans* (121). Deficiency of XPA did show accumulation of oxidised base like formamidopyrimidines (122) however, *xpa-1* mutants were found to have normal life span (121, 123). This study suggests that there may be mechanism other than defect in DNA repair in XP patients that causes neurodegeneration and accelerated aging.

A more direct connection between an increased load of DNA damage and mitochondrial dysfunction is also possible. There are several examples of this. In mice, a telomerase defect eventually gives rise to exposed telomeres. This genomic stress is clearly associated with mitochondrial dysfunction, and it is possible that exposed DNA ends mimicking a DSB induce a signal for disruption of key mitochondrial genes (124). An siRNA screen for factors that reduce sensitivity to a mitochondria-targeted DNA damaging agent provides another example (125). A diversity of different DNA repair factors, including XPA were detected in this screen.

With an increased load of DNA damage, transcription may be delayed because template damage interferes with transcript elongation. A second

consideration is that transcription and NER are competing processes. For example, transcription initiation efficiencies in yeast are reduced when NER is ongoing (126). In mammalian cells, there is lower NER activity at promoters when core transcription factors are bound to them (127, 128). After UV radiation damage to DNA, NER proteins are delayed in recruitment to promoter sites by several hours (81). This suggests that NER proteins are prioritized to repair damaged DNA first before resuming any activity in transcription.

5.1.2. Defective mitophagy associated with XPA deficiency and relation to neurological deficits

It has been observed that some of the symptoms of xeroderma pigmentosum group A have commonalities with mitochondrial disorders (82). Their study showed in human cells defective of XPA protein, mitochondrial functions like mitochondrial membrane potential, membrane organization and oxidoreductase activity were abnormal. XPA was reported to play a role in the regulation of mitochondrial autophagy and mitochondrial functions through NAD⁺/SIRT1 and PARP (82). In the present study, we found that genes 1.5 fold or more changed in XPA⁺ compared to XPA⁻ were indeed significantly represented for many mitochondrial and mitophagy-related GO terms (Table 4).

Mitochondrial defects and mitophagy are directly related to neurological functions. About 25 % of xeroderma pigmentosum patients develop

neurodegeneration in later life. The progressive neurodegeneration includes intellectual defects, hearing impairment, abnormal speech, neuropathy and muscular discoordination for movements of hands and legs (129). In a study of XP patients from 1971-2009, 6 out of 10 patients with XP-A and 16 out of 23 patients with XP-D had neurological degeneration (129). The cause of the death of many XP-A patients was also neurological degeneration (129-131).

One possible reason for this phenotype is the accumulated DNA damage in metabolically active but non-dividing neurons (57, 129), whereas some study in *C. elegans* show that *xpa-1* mutant, although accumulate oxidative damage, do have normal life span (122, 123). Other recent study shows that the *xpa-1* mutant old worms have higher amount of NAD⁺ production due to increase in PAMPylation, which may due to the mitochondrial dysfunction (82). This result suggests that there is a relation of old-age with mitochondrial dysfunction, which is more prominent in XPA deficient condition. Overall, these studies suggest that XPA may have role in maintenance of mitochondrial functions and consequently neurological functions.

5.1.3. The most significant XPA-regulated genes may explain steroid and sexual degeneration issues in xeroderma pigmentosum group A

The genes that are uniformly altered with the highest magnitude and significance are *AKR1C1* and *AKR1C2*. These genes are critical for steroid

metabolism and some bile metabolism (132-134). AKR1C1 converts progesterone to its inactive form, 20-alpha-dihydroxyprogesterone. AKR1C2 catalyzes the inactivation of the most potent androgen, 5-alpha-dihydrotestosterone, to 5-alpha-androstane-3-alpha,17-beta-diol (3-alpha-diol). Therefore, an imbalance of steroid hormones may be expected in XPA-defective cells. Both genes are indeed associated with obesity, sex reversal, hyperphasia and developmental delays (135, 136). A previous study using microarray analysis to study gene expression changes in XPA-defective cells (82). We analyzed the most XPA-dependent genes found in this microarray analysis (NCBI GEO dataset GSE55486) and compared it to our list of most changed genes. *AKR1C2* was the only gene that coincides with our list of most changed genes by RNA-Seq analysis. We confirmed that XPA-deficient cells have exceptionally low levels of AKR1C1 and AKR1C2 at the protein level, compared to proficient cells. This was true in both fibroblasts and HeLa cells.

Our finding of XPA-dependent expression of the steroid regulator AKR1C1 and AKR1C2 gene products may be highly relevant to the sterility and sexual developmental issues observed in XPA-deficient mice (137). Degenerating seminiferous tubules and no spermatozoa were detected in 24 month old *XPA*^{-/-} mice (137). Secondary sexual development is delayed in some XP patients (138). Consequently, our finding of XPA-dependent expression of the steroid regulators AKR1C1 and AKR1C2 may be highly relevant to the sterility and sexual developmental issues observed in mammals.

5.1.4. Multiple non-isogenic cell lines improve the analysis of RNA-Seq data

Despite these correlations, there are many unanswered questions concerning the mechanism of how XPA affects mitochondrial function. One puzzle is that although mitochondrial functions are affected, the specific gene transcripts that are affected are different in each cell line pair. Our current study shows how important it is to not only have biological replicates for RNA-Seq, but also to investigate cell lines of different genetic background in order to uncover the most significant effects of a single gene disruption.

5.1.5. Possible additional role of XPA may not be related to NER

It is significant that the XPA-dependent transcriptional profiles in KO142 and KO48 cells are similar, even though KO142 cells retain about 10% of the XPA protein level found in HeLa cells. The amount of XPA present in the KO142 line is enough to confer near-normal UV sensitivity, consistent with previous findings that XPA is in excess in cells for NER activity (116). Because complete knockout of XPA or knockdown to 10% has a similar effect on transcription, this suggests that at least part of the influence of XPA on transcription is not directly related to the NER function.

5.1.6. Cause of XPA instability in KO142

We checked if the loss of XPA protein in KO142 is due to the proteasomal degradation. Our result did not indicate such possibility although this study needs to be repeated with proper control for proteosomal degradation. Another cause of less XPA in KO142 may be due to the alternative splicing. The six amino acid deletion may have created an exonic splicing enhancer or disrupt exonic splicing silencer that could lead to an activation of alternative splicing site causing some exon skipping. This can be tested by PCR using one primer at Exon 1 and other one at different positions on exons beyond hypothesised splice sites. The products can be run in gel to compare the products from different pairs of primers.

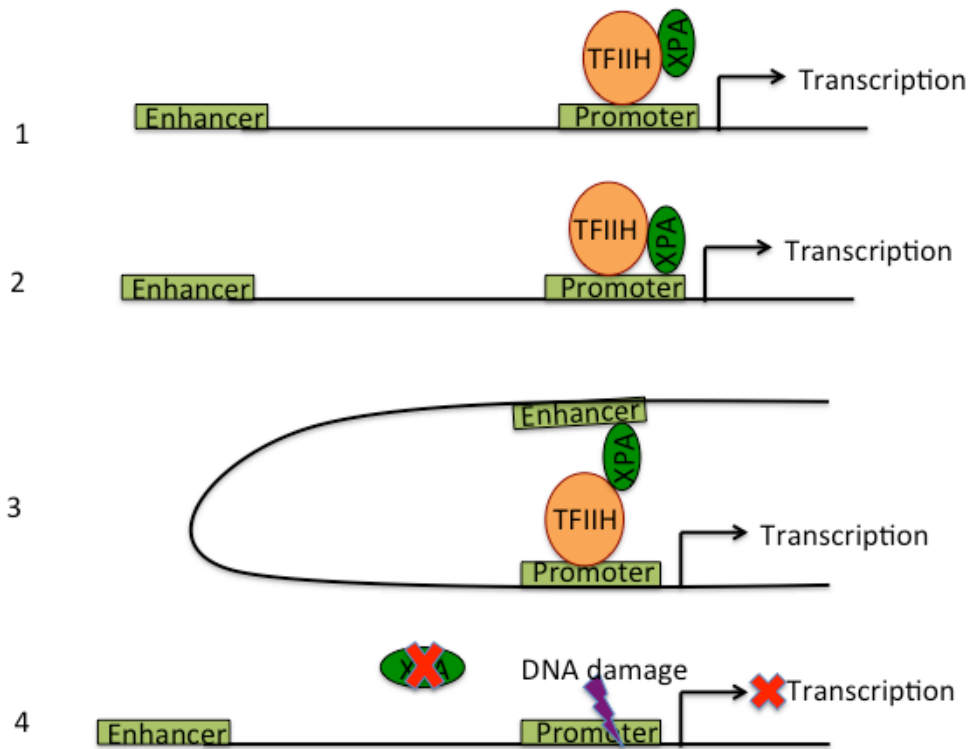


FIGURE 30. Model: Possible mechanism by which XPA is involved in transcription.

1. XPA may be involved in transcription by its interaction with TFIID. 2. XPA may directly bind to the promoter to directly play role in transcription. 3. XPA may bind to the enhancer region and be involved in transcription as a co-activator. 4. Due to unavailability of XPA for the DNA repair at the promoter site, transcription may be obstructed or delayed.

5.1.7. Future directions

5.1.7.1. Validation of RNA-Seq results by immunoblotting and CHIP

This study is the initial effort to get an idea of the extent of transcriptome changes in XPA-defective human cells. Our result shows that a small percentage of genes is affected by the XPA status. We found by RNA-Seq analysis that a high percentage of affected genes represented specific functions related to mitochondrial regulation and pathways related to neurological functions. We validated some of the important genes from RNA-Seq result at protein level and verified 2 of the three tested targets. However, this is only a small percentage of genes out of 27 highly differentially expressed genes in XPA- vs. XPA+. In the future, it will be valuable to check all of the 27 genes at the protein level.

Another way of validating the results from RNA-Seq will be to perform a chromatin immunoprecipitation (ChIP) assay at the promoter sites of the highly differentially expressed genes like *AKR1C1*, *AKR1C2*, and more genes of interest, to see if XPA is present at the promoter sites of these genes. This will give an idea about the mechanism how XPA is involved in the transcription. Since it has also been observed that in retinoic acid transactivated conditions, not only XPA but many other NER proteins were also found at the promoter

regions of few genes (81), it will be worth to look at if other NER proteins can also be pulled down.

The cell lines we used in our study are immortalized or cancer cell lines. It is important to get similar results in the primary cell lines as well to make sure the change in the genes are not related to the immortalization. Validations of the highly differential genes by similar methods as mentioned above can be done in the XPA deficient primary cell (skin fibroblasts) lines and normal human skin fibroblast as control. Comparison with primary cells from other XP groups would confirm that the changes are XPA-specific.

5.1.7.2. Analyzing differentially expressed genes in XPA deficient mice

In XPA-deficient mouse models, male sterility was one of the characteristics noted (139). However it not known whether the XPA-dependent genes that we found by RNA-Seq are involved in phenotypes of such mice. As AKR1C2 deficiency is found to be one of the causes for the reversal of sexual orientation and sterility (140), it will be valuable information to look for AKR1C2 and other genes of our interest in tissues of XPA deficient mouse models.

5.1.7.3. Biochemical assays for mitochondrial and metabolic functions in XPA deficient and proficient cells

Since we found metabolic and mitochondrial pathways being represented by significant differentially expressed genes in our analysis, XPA may have a role in metabolism or mitochondrial maintenance. It is worth looking into more details of XPA being involved in mitochondrial functions by performing biochemical assays. Monitoring ATP generation, mitochondrial membrane potential, ROS formation and biochemical metabolic reactions such as reduction of aldehyde-ketone compounds in XPA deficient cells compared to the proficient, can be done for the further insight in the study. Immunofluorescent assays can also be done to compare the expression of these genes like *NDUFA4L2* (NADH Dehydrogenase (Ubiquinone) 1 Alpha Subcomplex, 4-Like 2) compared to XPA status.

Although it needs more study, we obtained some preliminary results for the influence of XPA status on mitochondrial functions. We performed biochemical assays to detect the mitochondrial mass, mitochondrial membrane potential defect and amount of ROS levels in XPA proficient and deficient cell lines. Following assays were performed in collaboration with Dr. Miao-der Chen in Bratton laboratory.

Mitochondrial mass:

To determine whether the mitochondrial dysfunction obtained suggested by expression pathway analysis is due to a low mitochondrial mass in XPA deficient cells, we looked at the total mitochondrial mass in all cells. Cells were treated with TMRE (tetramethylrhodamine, ethyl ester) dye for 5 min at 37 °C,

washed with PBS and subjected to Flow cytometer (Accuri C6 Flowcytometer) using FL3 fluorescence channel and 8000 events (141). TMRE is a positively charged red dye that is accumulated easily in mitochondria due to its negative charge. Depolarized or inactive mitochondria have low membrane potential and are unable to pump in the dye (142). We observed no difference in the total fluorescence count (total populations) in XPA proficient and deficient cells except slight lower count in XP12RO compared to XPA/XP12RO cells (Figure 31A). This result indicated that there is no difference in number of mitochondria in XPA proficient and deficient cells. Cells not treated with TMRE were used as negative control and they were observed as population without fluorescent counts. This assay needs to be repeated to confirm the result.

Mitochondrial membrane potential defect:

To confirm the result about defect in mitochondrial membrane potential in XPA deficient cells from the pathway analysis, we performed similar assay as described above and looked at the population that could not pump in TMRE in the mitochondria. This population represents the defect in membrane potential (143). We observed that XPA deficient cells have higher percentage of population that has defect in mitochondrial membrane potential compared to the proficient cells. This population has low mitochondrial membrane potential (Figure 31B). Similar negative control was used as mentioned above. This assay needs to be repeated for the confirmation of the result.

ROS levels:

Higher ROS production is one of the factors observed when there is defect (high) in mitochondrial membrane potential. We performed an assay where a reagent (CM-H₂DCFDA) for indicator of ROS was used in cell culture at 37 °C for 15 min. It turns to fluorescent compound DCF after oxidation, which was measured by flow cytometer at fluorescent channel FL1 and 8000 events. Cells not treated with CM-H₂DCFDA were used as negative control. We did not observe difference or only little lower ROS in XPA deficient cells compared to the proficient cells (Figure 31C) which is relevant to the lower mitochondrial membrane potential observed in Figure 31B. This result indicated that the cells are not going through the oxidative stress. This assay needs to be repeated for the confirmation of the result.

Our preliminary results mentioned above contradict with the recent study in XPA deficient human cell lines showing the increased mitochondrial potential and ROS levels in XPA deficient cells (82). If our result is true about lower or no difference in ROS level in XPA deficient cells, it could be because the XPA deficiency is not causing problem for repairing of the oxidative damages, as BER is efficient enough to repair most of the oxidative damages.

Both low and high mitochondrial potential are defective for cells as they have abnormal physiological consequences (144, 145). If our result about lower mitochondrial membrane potential in XPA deficient cells compared to XPA

proficient cells is true, it may indicate that the XPA deficient cells are going through low generation of ATP or low oxidative phosphorylation. This explanation also concurs with the high fold change (1.5 - 7) in *NDUFA4L2* gene in XPA proficient cells compared to the deficient. This result suggests that XPA may play role in metabolism by maintaining mitochondrial membrane potential.

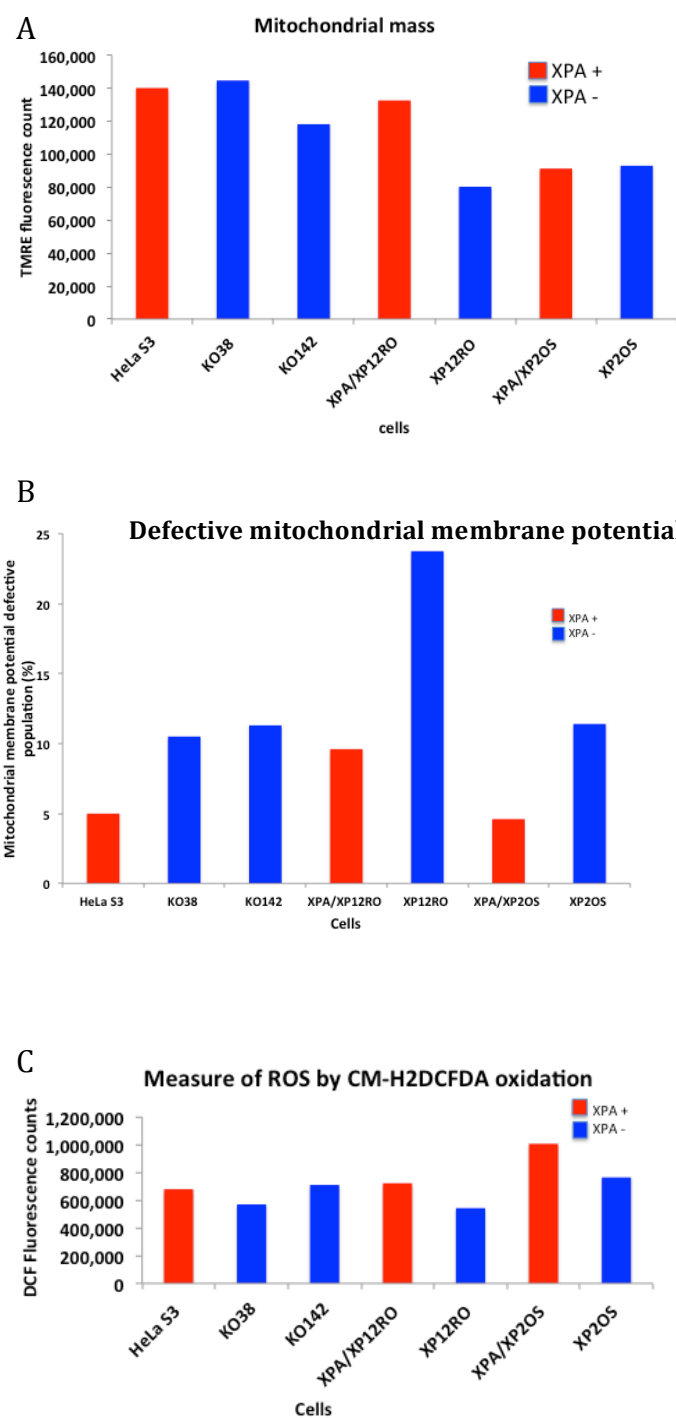


FIGURE 31. Mitochondrial assays.

A. Mitochondrial mass was measured by the fluorescent count produced by the uptake of TMRE dye in the mitochondria. B. Defect in mitochondrial membrane potential was measured as indicated by the population in the flow cytometer read that were unable to sequester TMRE dye in the mitochondria by the membrane potential pump. C. Level of ROS in the cells was measured as indicated by the fluorescent count produced due to the oxidation of CM-H₂DCFDA in to DCF.

5.1.7.4. Analyzing brain tissues of XPA deficient mice and patients with neurological diseases

Similarly, we observed many genes and pathways responsible for neurological functions. It is also important to explore XPA *in vivo* analysis. This can be studied in brain tissues of XPA deficient mice in absence of any external DNA damage and also observe these changes regarding some of the neurological genes such as *WFDC21P*, *LHX2*, *HOMER2* and others we observe as highly expressed in XPA proficient cells. Immunofluorescent assays on brain tissues for the expression of these genes compared to XPA will be helpful to understand the effect of XPA status in neurological functions. Alternatively, tissue samples from patients with Alzheimer's or Parkinson or Huntington disease can be analyzed for the expression of XPA. The result will be helpful in further discussing on the cause of neurological defects in XPA patients.

5.2. Processing of TFO-directed psoralen interstrand crosslink in human cell extracts

5.2.1. Psoralen-TFO ICL or TFO as an adduct for NER

The enzymatic processing of a TFO-directed psoralen ICL in closed circular double stranded DNA has not been studied previously human cells. It is highly possible that it is a substrate for NER in human cells as the mutation caused by TFO was found to be dependent on NER (146). In non-replicating

conditions, damage in DNA is recognized by NER as distortion to DNA structure. Psoralen causes enough distortion in duplex DNA, which makes it a substrate of nucleotide excision repair, even though the distortion is less than other ICL causing agents such as cisplatin (22). It makes a large 55.8° buckle in the d (T.A) base pair with thymine on the pyrone strand. This results in a steric collision of the O4 of thymine and pyrone ring, resulting in destruction of Watson and Crick hydrogen bond between the base pairs (147). We used the TFO conjugated psoralen primarily to construct a site specific ICL. However, since both psoralen and TFOs may be substrates for NER, we expected the processing of TFO directed psoralen ICL substrate may be dependent on NER. The addition of TFO to psoralen may cause even more degree of distortion compared to that caused by TFO or psoralen alone. This may be the reason why we observed incisions only beyond the TFO binding site on the strand (upper), where TFO is bound to the DNA. On this strand, incision products were observed 44 nucleotides away from the ICL site, which were beyond TFO binding site. These incisions were 9 nucleotides or farther from the 3' end of the TFO, which is relevant to the distance of 3' side incisions observed with cisplatin intrastrand adduct in human cells (102). We also tried to detect incision products within TFO binding site, but we were unable to find any incisions within TFO binding site. This result may be the indication that TFOs are substrates for the nucleotide excision repair pathway.

5.2.2. TFO-directed psoralen ICL was not repaired by HeLa whole cell extract

We observed incisions at the 3' side of TFO-directed psoralen ICLs but we were unable to detect repair in human cell extract. We did not find incorporation of radiolabelled dNTPs at the ICL site, which was indicated by the disappearance of the fragment containing the ICL in the damaged DNA substrate compared to the undamaged DNA substrate. In one of our repair assays, we used a restriction site that coincides with the ICL formation site, which makes the ICL substrate resistant to the restriction digestion by *SapI*. 29 bp and 69 bp fragments were supposed to be released if *SapI* and *AluI* cut at their sites. We found the disappearance of the 29 bp on ICL substrate whereas we could not confirm similar result for the 69 bp fragment, as it appeared to be coinciding with other bands (64 and 67 bp) on the gel.

We designed another repair assay with different restriction digestion sites. Since we were confident that *BbsI* cuts and releases the 50 bp ICL fragment, we used *BbsI* and one other site, *MluCI*, to create multiple bands to make the comparison of intensity of the fragments more efficient. In this assay also, we found the ICL containing fragment (50 bp) was not observed on the ICL substrate, whereas it was observed on the undamaged plasmid. Since all other neighboring bands were also observed in both damaged and undamaged DNA, the restriction digestion was successful with both enzymes but the 50 bp ICL fragment was not replaced with the radiolabeled dNTPs. This result suggests an unsuccessful repair of the psoralen-TFO ICL in HeLa whole cell extracts.

This turns our attention towards the question whether the incisions we observed in human whole cell extracts are repair-related. If they are not for the purpose of repair, why are the incisions made? It could be that the cell extracts cannot proceed beyond incision to complete the repair on psoralen-TFO ICL. However in mammalian cells, previous studies have indicated the repair of TFO or psoralen ICL or psoralen-TFO ICL. TFO-induced mutation was dependent on NER (XPA) (146), and thus TFOs are possibly repaired by human cells. Psoralen ICLs are also repaired in human cells as shown by reporter assays (148). TFO-directed psoralen in mammalian cells has shown to cause a higher degree of mutation compared to spontaneous mutation (149), indicating its recovery. However, no study has been done so far to study the repair of psoralen-TFO ICL in human cell extract. If TFO-directed psoralen ICLs are not repaired by the whole cell extracts, but are repaired in human cells, it may be because they are processed beyond incision by other DNA repair enzymes besides NER proteins. Recently it has been reported that NEIL3 glycosylase is involved in unhooking of psoralen ICL in *xenopus* egg extract in incision independent way (150). When N-glycosyl bond cleavage is inhibited, the incisions were produced in a FANCI-D2 dependent manner (replicating system), indicating that the NEIL3 unhooking is the preferred way of processing psoralen ICLs in vertebrate cell free system. Thus it may be possible that psoralen-TFO ICLs are also processed by NEIL3 at the unhooking step. In absence of this step in processing, repair may not have

occurred in HeLa whole cell extract. The possible mechanism by which psoralen-TFO ICLs are processed in human cells is represented as a model in Figure 32.

5.2.3. TFO in targeted DNA damage and gene alteration

We have not tested the difference between incision products with psoralen only and psoralen-TFO ICL. If it is true as observed in *E.coli* where TFO resulted in more incision products than psoralen ICL only (25), psoralen-TFO ICL may be more effective in chemotherapy to kill cancer cells through DNA damage. Thus, determining the mechanism of TFO-directed psoralen-DNA ICL is an important step towards improving chemotherapy.

Currently, due to lack of specific targets, chemotherapies are not being effective. TFOs can be used to target specific genes, like oncogenes, to disrupt them or to deliver targeted drugs like antimetabolites to specific sites to cause cytotoxic effects in cancer cells (151). Since TFOs bind to specific sequences of the genome and with high affinity, it has been put forward as an option for altering targeted sequences (152-155). It can lead to improving the cure and treatment of different genetic diseases. There is an abundance of TFO-binding homopurine or homopyrimidine rich sequences present in human genome (156), giving a good platform for TFO-targeting treatment. TFOs can be used in conjugation with DNA damaging agents, such as ICL causing agents or photo induced endonucleases (157) for the targeted effects. The concept of TFO

conjugated drug therapy is still not implemented in humans. In mice, cells transfected with TFO conjugated with a cytotoxic drugs was delivered to treat tumors, and was successful in reducing the tumor load (151). Similar strategy can be used to deliver TFOs, TFO conjugated drugs, or DNA damaging agents in cancer patients. However, since it is not exactly known about the mechanism of processing of TFO or TFO conjugated psoralen, the results from TFO-targeting therapy will be highly unpredictable.

5.2.4. ICL substrate in a closed circular double stranded DNA

Most of the efforts to study the processing of ICL were done with ICL in short duplex DNA. Since the duplex DNA has open ends, it is possible to be processed by unspecific nucleases available in the cell extract. To avoid the artifacts and confusions of processing of ICL due to the accessibility of the free ends, we constructed ICL in closed circular double stranded DNA and purified high-yield of the final product. It was also important to purify ICL products from the monoadducts, which are formed during the photoactivation of psoralen. It was even more challenging to do this purification with closed circular double stranded DNA. Thus we first made ICL in short duplex DNA, purified in denaturing condition as covalent bonding in ICL cannot be separated by denaturing while DNA with monoadducts can be separated to single stranded DNA. Short duplex DNA with ICL purified by this strategy was efficiently ligated to

the vector with unique DNA ends (*Bbs*I) to confirm the correct orientation of the ligation.

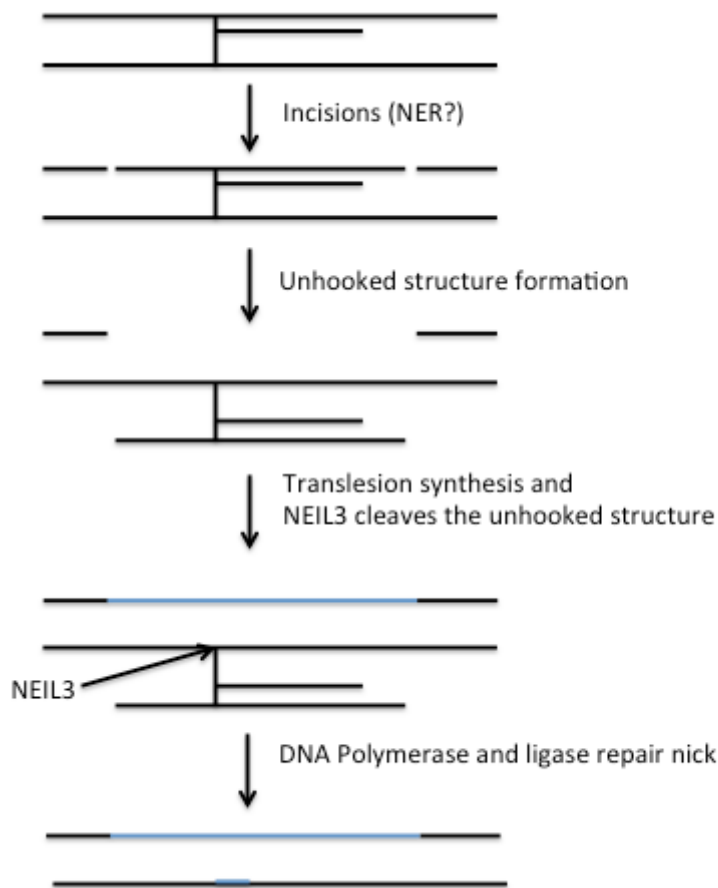


FIGURE 32. Model: Possible mechanisms of psoralen-TFO ICL repair in human cells.

The damage is recognized and incisions are made around the psoralen-TFO ICL, which may be dependent on NER. The gap created by incisions is repaired by translesion synthesis. In the later step of repair, unhooked structure is formed which may be cleaved by NEIL3 followed by repair by DNA polymerase and ligase.

5.2.5. Future directions

5.2.5.1. Incision assay with NER deficient cells complementation with purified XPA protein

There are some indications from the previous studies that the incisions we observed with the psoralen-TFO ICLs may be dependent on NER. However, we have not ruled this out for certain. The important study to do in the future will be to confirm this with NER deficient whole cell extracts and complement with the purified NER protein. XPA deficient whole cell extracts from KO38 cells will be made, and the incision reactions will be performed as described earlier, with or without addition of XPA purified protein.

5.2.5.2. Detection of 5' side incisions

We looked for the incision products only at the 3' side of the adduct, on both strands. Detection of the incisions at the 5' side of the psoralen-TFO ICL, in similar process as we detected 3' incision sites, is a challenge due to the direction of the polymerase activity and the limited restriction digestion sites available on the plasmid. An effective plan should be developed to detect the incision products at the 5' side as well. Only after confirming that the incisions occur on this ICL substrate at both sides of the adduct, we could be more close to understanding if the processing of the TFO-directed psoralen ICL in closed

circular double stranded DNA is by NER. Some strategies are presented as follows: 1). Design a G-overhang oligonucleotide complementing incision product around between ICL and *ClaI*. This is only 10 bp long distance, which may be too short to cover the incision site at 5' side of the ICL. Another G-overhang can be designed around *EcoRI*, which is 37 bp away from the ICL, and may be too far to cover the incision site. However we have not tried this and it may be worth trying. 2). In another strategy an oligonucleotide can be designed in which its 3' end coincides with the 5' end of *EcoRI*, a restriction digestion site 37 bp away at the 5' side of the ICL. The coinciding of the ends is important to exactly map the size of the incision product. The oligo should be long enough to cover the incisions close to 5' side of the ICL. This oligo will be radiolabelled at its 3' end. After the incision reaction of the DNA substrates, they will be purified, denatured, and annealed with radiolabelled oligonucleotide. A single strand specific nuclease will be used to digest the 3' overhang resulting in the duplex DNA of same size as the incision product and will be detected in sequencing gel (Figure 33).

5.2.5.3. Incision assays with or without TFO

Our results from the incision assay on psoralen-TFO indicated that TFO may be substrate for the incision reaction, but to differentiate if psoralen ICL and TFO-psoralen ICL are processed differently, TFO has to be removed after interstrand crosslink with psoralen is formed and the similar incision assays need to be performed and analyzed. A strategy as described in Christensen et al, 2008

can be used for this purpose (25). TFO and psoralen can be bound to a disulfide bond that can be easily removed after ICL is formed, by treating with DTT at 60°C for few hours (158).

5.2.5.4. Involvement of NEIL3 in the processing of psoralen-TFO ICL

NER proficient whole cell extract could not repair psoralen-TFO ICL in our experiments. It may be possible that other repair pathways are involved in the processing of the ICL substrate that the whole cell extract we prepared are not proficient in. Recently it has been reported that NEIL3 cleaves the unhooked structure of psoralen ICL in incision independent manner (150). It is worth to check if psoralen TFO is also unhooked by NEIL3 after initial incisions are made. The reaction can be done following the protocol in Liu. M, et al, 2009 (159).

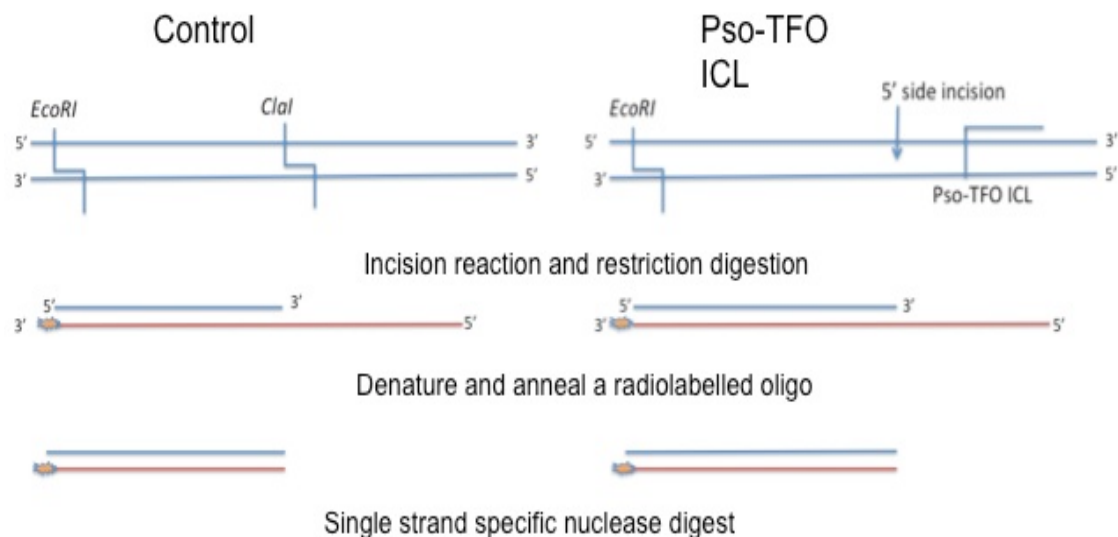


FIGURE 33. Strategy to detect incisions at the 5' side of the ICL on the upper strand.

After the incision reaction, a 3' radiolabelled oligo will be annealed to the purified and denatured DNA intermediate products of control substrate and ICL substrate indicated in the figure. A single strand specific nuclease will be used to digest the overhang of the oligo at 5' side. The remaining fragment will be detected on a sequencing gel.

References

1. Wood RD. 2010. Mammalian nucleotide excision repair proteins and interstrand crosslink repair. *Environmental and Molecular Mutagenesis* 51: 520-6
2. Deans AJ, West SC. 2011. DNA interstrand crosslink repair and cancer. *Nature reviews. Cancer* 11: 467-80
3. Vasquez KM, Legerski RJ. 2010. DNA interstrand crosslinks: repair, cell signaling, and therapeutic implications. *Environmental and Molecular Mutagenesis* 51: 491-2
4. Kim H, D'Andrea AD. 2012. Regulation of DNA cross-link repair by the Fanconi anemia/BRCA pathway. *Genes Dev* 26: 1393-408
5. Schwab RA, Blackford AN, Niedzwiedz W. 2010. ATR activation and replication fork restart are defective in FANCM-deficient cells. *EMBO J* 29: 806-18
6. Ciccia A, Ling C, Coulthard R, Yan Z, Xue Y, Meetei AR, Laghmani el H, Joenje H, McDonald N, de Winter JP, Wang W, West SC. 2007. Identification of FAAP24, a Fanconi anemia core complex protein that interacts with FANCM. *Mol Cell* 25: 331-43
7. Collis SJ, Ciccia A, Deans AJ, Horejsi Z, Martin JS, Maslen SL, Skehel JM, Elledge SJ, West SC, Boulton SJ. 2008. FANCM and FAAP24 function in ATR-mediated checkpoint signaling independently of the Fanconi anemia core complex. *Mol Cell* 32: 313-24

8. Huang M, D'Andrea AD. 2010. A new nuclease member of the FAN club. *Nat Struct Mol Biol* 17: 926-8
9. Crossan GP, Patel KJ. 2012. The Fanconi anaemia pathway orchestrates incisions at sites of crosslinked DNA. *J Pathol* 226: 326-37
10. De Silva IU, McHugh PJ, Clingen PH, Hartley JA. 2000. Defining the roles of nucleotide excision repair and recombination in the repair of DNA interstrand cross-links in mammalian cells. *Mol Cell Biol* 20: 7980-90
11. Fisher LA, Bessho M, Bessho T. 2008. Processing of a psoralen DNA interstrand cross-link by XPF-ERCC1 complex in vitro. *J Biol Chem* 283: 1275-81
12. Vaz F, Hanenberg H, Schuster B, Barker K, Wiek C, Erven V, Neveling K, Endt D, Kesterton I, Autore F, Fraternali F, Freund M, Hartmann L, Grimwade D, Roberts RG, Schaal H, Mohammed S, Rahman N, Schindler D, Mathew CG. 2010. Mutation of the RAD51C gene in a Fanconi anemia-like disorder. *Nat Genet* 42: 406-9
13. Bunting SF, Callen E, Kozak ML, Kim JM, Wong N, Lopez-Contreras AJ, Ludwig T, Baer R, Faryabi RB, Malhowski A, Chen HT, Fernandez-Capetillo O, D'Andrea A, Nussenzweig A. 2012. BRCA1 functions independently of homologous recombination in DNA interstrand crosslink repair. *Mol Cell* 46: 125-35
14. Niedzwiedz W, Mosedale G, Johnson M, Ong CY, Pace P, Patel KJ. 2004. The Fanconi anaemia gene FANCC promotes homologous recombination and error-prone DNA repair. *Mol Cell* 15: 607-20

15. Mirchandani KD, McCaffrey RM, D'Andrea AD. 2008. The Fanconi anemia core complex is required for efficient point mutagenesis and Rev1 foci assembly. *DNA Repair (Amst)* 7: 902-11
16. Hicks JK, Chute CL, Paulsen MT, Ragland RL, Howlett NG, Gueranger Q, Glover TW, Canman CE. 2010. Differential roles for DNA polymerases eta, zeta, and REV1 in lesion bypass of intrastrand versus interstrand DNA cross-links. *Mol Cell Biol* 30: 1217-30
17. Hoy CA, Thompson LH, Mooney CL, Salazar EP. 1985. Defective DNA cross-link removal in Chinese hamster cell mutants hypersensitive to bifunctional alkylating agents. *Cancer Res* 45: 1737-43
18. Clingen PH, Arlett CF, Hartley JA, Parris CN. 2007. Chemosensitivity of primary human fibroblasts with defective unhooking of DNA interstrand cross-links. *Exp Cell Res* 313: 753-60
19. Kaye J, Smith CA, Hanawalt PC. 1980. DNA repair in human cells containing photoadducts of 8-methoxypsoralen or angelicin. *Cancer Res* 40: 696-702
20. De Silva IU, McHugh PJ, Clingen PH, Hartley JA. 2002. Defects in interstrand cross-link uncoupling do not account for the extreme sensitivity of ERCC1 and XPF cells to cisplatin. *Nucleic Acids Res* 30: 3848-56
21. Bessho T, Mu D, Sancar A. 1997. Initiation of DNA interstrand cross-link repair in humans: the nucleotide excision repair system makes dual incisions 5' to the cross-linked base and removes a 22- to 28-nucleotide-long damage-free strand. *Mol Cell Biol* 17: 6822-30

22. Smeaton MB, Hlavin EM, McGregor Mason T, Noronha AM, Wilds CJ, Miller PS. 2008. Distortion-dependent unhooking of interstrand cross-links in mammalian cell extracts. *Biochemistry* 47: 9920-30
23. Zhang N, Lu X, Zhang X, Peterson CA, Legerski RJ. 2002. hMutSbeta is required for the recognition and uncoupling of psoralen interstrand cross-links in vitro. *Mol Cell Biol* 22: 2388-97
24. Culver KW, Hsieh WT, Huyen Y, Chen V, Liu J, Khripine Y, Khorlin A. 1999. Correction of chromosomal point mutations in human cells with bifunctional oligonucleotides. *Nat Biotechnol* 17: 989-93
25. Christensen LA, Wang H, Van Houten B, Vasquez KM. 2008. Efficient processing of TFO-directed psoralen DNA interstrand crosslinks by the UvrABC nuclease. *Nucleic acids research* 36: 7136-45
26. Wang G, Glazer PM. 1995. Altered repair of targeted psoralen photoadducts in the context of an oligonucleotide-mediated triple helix. *J Biol Chem* 270: 22595-601
27. Vasquez KM, Christensen J, Li L, Finch RA, Glazer PM. 2002. Human XPA and RPA DNA repair proteins participate in specific recognition of triplex-induced helical distortions. *Proceedings of the National Academy of Sciences of the United States of America* 99: 5848-53
28. Gillet LC, Scharer OD. 2006. Molecular mechanisms of mammalian global genome nucleotide excision repair. *Chem Rev* 106: 253-76
29. Friedberg EG, Walker GC, Siede W, Wood RD, Schultz RA, Ellenberger T. 2006. *DNA Repair and Mutagenesis*

- . Washington D.C: A.S.M Press. 1117 pp.
30. Sugasawa K. 2010. Regulation of damage recognition in mammalian global genomic nucleotide excision repair. *Mutat Res* 685: 29-37
 31. Volker M, Mone MJ, Karmakar P, van Hoffen A, Schul W, Vermeulen W, Hoeijmakers JH, van Driel R, van Zeeland AA, Mullenders LH. 2001. Sequential assembly of the nucleotide excision repair factors in vivo. *Mol Cell* 8: 213-24
 32. Sugasawa K, Ng JM, Masutani C, Iwai S, van der Spek PJ, Eker AP, Hanaoka F, Bootsma D, Hoeijmakers JH. 1998. Xeroderma pigmentosum group C protein complex is the initiator of global genome nucleotide excision repair. *Mol Cell* 2: 223-32
 33. Bergink S, Toussaint W, Luijsterburg MS, Dinant C, Alekseev S, Hoeijmakers JH, Dantuma NP, Houtsmuller AB, Vermeulen W. 2012. Recognition of DNA damage by XPC coincides with disruption of the XPC-RAD23 complex. *J Cell Biol* 196: 681-8
 34. Bunick CG, Miller MR, Fuller BE, Fanning E, Chazin WJ. 2006. Biochemical and structural domain analysis of xeroderma pigmentosum complementation group C protein. *Biochemistry* 45: 14965-79
 35. Araujo SJ, Nigg EA, Wood RD. 2001. Strong functional interactions of TFIIH with XPC and XPG in human DNA nucleotide excision repair, without a preassembled repairosome. *Mol Cell Biol* 21: 2281-91

36. Oksenych V, Bernardes de Jesus B, Zhovmer A, Egly JM, Coin F. 2009. Molecular insights into the recruitment of TFIIH to sites of DNA damage. *EMBO J* 28: 2971-80
37. Fanning E, Klimovich V, Nager AR. 2006. A dynamic model for replication protein A (RPA) function in DNA processing pathways. *Nucleic Acids Res* 34: 4126-37
38. de Laat WL, Appeldoorn E, Sugasawa K, Weterings E, Jaspers NG, Hoeijmakers JH. 1998. DNA-binding polarity of human replication protein A positions nucleases in nucleotide excision repair. *Genes Dev* 12: 2598-609
39. Kolpashchikov DM, Khodyreva SN, Khlimankov DY, Wold MS, Favre A, Lavrik OI. 2001. Polarity of human replication protein A binding to DNA. *Nucleic Acids Res* 29: 373-9
40. Li CL, Golebiowski FM, Onishi Y, Samara NL, Sugasawa K, Yang W. 2015. Tripartite DNA Lesion Recognition and Verification by XPC, TFIIH, and XPA in Nucleotide Excision Repair. *Mol Cell* 59: 1025-34
41. Hey T, Lipps G, Krauss G. 2001. Binding of XPA and RPA to damaged DNA investigated by fluorescence anisotropy. *Biochemistry* 40: 2901-10
42. Patrick SM, Turchi JJ. 2002. Xeroderma pigmentosum complementation group A protein (XPA) modulates RPA-DNA interactions via enhanced complex stability and inhibition of strand separation activity. *J Biol Chem* 277: 16096-101

43. Yang Z, Roginskaya M, Colis LC, Basu AK, Shell SM, Liu Y, Musich PR, Harris CM, Harris TM, Zou Y. 2006. Specific and efficient binding of xeroderma pigmentosum complementation group A to double-strand/single-strand DNA junctions with 3'- and/or 5'-ssDNA branches. *Biochemistry* 45: 15921-30
44. Giglia-Mari G, Miquel C, Theil AF, Mari PO, Hoogstraten D, Ng JM, Dinant C, Hoeijmakers JH, Vermeulen W. 2006. Dynamic interaction of TTDA with TFIIH is stabilized by nucleotide excision repair in living cells. *PLoS Biol* 4: e156
45. Ziani S, Nagy Z, Alekseev S, Soutoglou E, Egly JM, Coin F. 2014. Sequential and ordered assembly of a large DNA repair complex on undamaged chromatin. *J Cell Biol* 206: 589-98
46. Park CH, Mu D, Reardon JT, 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: 4896-902
47. Buchko GW, Ni S, Thrall BD, Kennedy MA. 1998. Structural features of the minimal DNA binding domain (M98-F219) of human nucleotide excision repair protein XPA. *Nucleic Acids Res* 26: 2779-88
48. Li L, Elledge SJ, Peterson CA, Bales ES, Legerski RJ. 1994. Specific association between the human DNA repair proteins XPA and ERCC1. *Proc Natl Acad Sci U S A* 91: 5012-6

49. Buchko GW, Daughdrill GW, de Lorimier R, Rao BK, Isern NG, Lingbeck JM, Taylor JS, Wold MS, Gochin M, Spicer LD, Lowry DF, Kennedy MA. 1999. Interactions of human nucleotide excision repair protein XPA with DNA and RPA70 Delta C327: chemical shift mapping and ¹⁵N NMR relaxation studies. *Biochemistry* 38: 15116-28
50. Sugitani N, Sivley RM, Perry KE, Capra JA, Chazin WJ. 2016. XPA: A key scaffold for human nucleotide excision repair. *DNA Repair (Amst)*
51. States JC, McDuffie ER, Myrand SP, McDowell M, Cleaver JE. 1998. Distribution of mutations in the human xeroderma pigmentosum group A gene and their relationships to the functional regions of the DNA damage recognition protein. *Hum Mutat* 12: 103-13
52. Wakasugi M, Kasashima H, Fukase Y, Imura M, Imai R, Yamada S, Cleaver JE, Matsunaga T. 2009. Physical and functional interaction between DDB and XPA in nucleotide excision repair. *Nucleic Acids Res* 37: 516-25
53. Lee SH, Kim DK, Drissi R. 1995. Human xeroderma pigmentosum group A protein interacts with human replication protein A and inhibits DNA replication. *J Biol Chem* 270: 21800-5
54. Miyamoto I, Miura N, Niwa H, Miyazaki J, Tanaka K. 1992. Mutational analysis of the structure and function of the xeroderma pigmentosum group A complementing protein. Identification of essential domains for nuclear localization and DNA excision repair. *J Biol Chem* 267: 12182-7

55. Lolli G, Lowe ED, Brown NR, Johnson LN. 2004. The crystal structure of human CDK7 and its protein recognition properties. *Structure* 12: 2067-79
56. Fischer JM, Popp O, Gebhard D, Veith S, Fischbach A, Beneke S, Leitenstorfer A, Bergemann J, Scheffner M, Ferrando-May E, Mangerich A, Burkle A. 2014. Poly(ADP-ribose)-mediated interplay of XPA and PARP1 leads to reciprocal regulation of protein function. *FEBS J* 281: 3625-41
57. DiGiovanna JJ, Kraemer KH. 2012. Shining a light on xeroderma pigmentosum. *J Invest Dermatol* 132: 785-96
58. Cleaver JE. 2005. Cancer in xeroderma pigmentosum and related disorders of DNA repair. *Nat Rev Cancer* 5: 564-73
59. Kraemer KH, Levy DD, Parris CN, Gozukara EM, Moriwaki S, Adelberg S, Seidman MM. 1994. Xeroderma pigmentosum and related disorders: examining the linkage between defective DNA repair and cancer. *J Invest Dermatol* 103: 96S-101S
60. Cleaver JE, States JC. 1997. The DNA damage-recognition problem in human and other eukaryotic cells: the XPA damage binding protein. *Biochem J* 328 (Pt 1): 1-12
61. Takahashi Y, Endo Y, Sugiyama Y, Inoue S, Iijima M, Tomita Y, Kuru S, Takigawa M, Moriwaki S. 2010. XPA gene mutations resulting in subtle truncation of protein in xeroderma pigmentosum group A patients with mild skin symptoms. *J Invest Dermatol* 130: 2481-8

62. Manandhar M, Boulware KS, Wood RD. 2015. The ERCC1 and ERCC4 (XPF) genes and gene products. *Gene* 569: 153-61
63. Ahmad A, Robinson AR, Duensing A, van Drunen E, Beverloo HB, Weisberg DB, Hasty P, Hoeijmakers JH, Niedernhofer LJ. 2008. ERCC1-XPF endonuclease facilitates DNA double-strand break repair. *Mol Cell Biol* 28: 5082-92
64. Bennardo N, Cheng A, Huang N, Stark JM. 2008. Alternative-NHEJ is a mechanistically distinct pathway of mammalian chromosome break repair. *PLoS Genet* 4: e1000110
65. Schrader CE, Vardo J, Linehan E, Twarog MZ, Niedernhofer LJ, Hoeijmakers JH, Stavnezer J. 2004. Deletion of the nucleotide excision repair gene Ercc1 reduces immunoglobulin class switching and alters mutations near switch recombination junctions. *J Exp Med* 200: 321-30
66. Munoz P, Blanco R, Flores JM, Blasco MA. 2005. XPF nuclease-dependent telomere loss and increased DNA damage in mice overexpressing TRF2 result in premature aging and cancer. *Nat Genet* 37: 1063-71
67. Sargent RG, Rolig RL, Kilburn AE, Adair GM, Wilson JH, Nairn RS. 1997. Recombination-dependent deletion formation in mammalian cells deficient in the nucleotide excision repair gene ERCC1. *Proc Natl Acad Sci U S A* 94: 13122-7

68. McWhir J, Selfridge J, Harrison DJ, Squires S, Melton DW. 1993. Mice with DNA repair gene (ERCC-1) deficiency have elevated levels of p53, liver nuclear abnormalities and die before weaning. *Nat Genet* 5: 217-24
69. Harada YN, Shiomi N, Koike M, Ikawa M, Okabe M, Hirota S, Kitamura Y, Kitagawa M, Matsunaga T, Nikaido O, Shiomi T. 1999. Postnatal growth failure, short life span, and early onset of cellular senescence and subsequent immortalization in mice lacking the xeroderma pigmentosum group G gene. *Mol Cell Biol* 19: 2366-72
70. Nospikel T, Lalle P, Leadon SA, Cooper PK, Clarkson SG. 1997. A common mutational pattern in Cockayne syndrome patients from xeroderma pigmentosum group G: implications for a second XPG function. *Proc Natl Acad Sci U S A* 94: 3116-21
71. Laugel V, Dalloz C, Durand M, Sauvanaud F, Kristensen U, Vincent MC, Pasquier L, Odent S, Cormier-Daire V, Gener B, Tobias ES, Tolmie JL, Martin-Coignard D, Drouin-Garraud V, Heron D, Journel H, Raffo E, Vigneron J, Lyonnet S, Murday V, Gubser-Mercati D, Funalot B, Brueton L, Sanchez Del Pozo J, Munoz E, Gennery AR, Salih M, Noruzinia M, Prescott K, Ramos L, Stark Z, Fieggen K, Chabrol B, Sarda P, Edery P, Bloch-Zupan A, Fawcett H, Pham D, Egly JM, Lehmann AR, Sarasin A, Dollfus H. 2010. Mutation update for the CSB/ERCC6 and CSA/ERCC8 genes involved in Cockayne syndrome. *Hum Mutat* 31: 113-26

72. Evans E, Fellows J, Coffey A, Wood RD. 1997. Open complex formation around a lesion during nucleotide excision repair provides a structure for cleavage by human XPG protein. *The EMBO journal* 16: 625-38
73. Coin F, Oksenyich V, Mocquet V, Groh S, Blattner C, Egly JM. 2008. Nucleotide excision repair driven by the dissociation of CAK from TFIIH. *Mol Cell* 31: 9-20
74. Sarker AH, Tsutakawa SE, Kostek S, Ng C, Shin DS, Peris M, Campeau E, Tainer JA, Nogales E, Cooper PK. 2005. Recognition of RNA polymerase II and transcription bubbles by XPG, CSB, and TFIIH: insights for transcription-coupled repair and Cockayne Syndrome. *Mol Cell* 20: 187-98
75. Cattoglio C, Zhang ET, Grubisic I, Chiba K, Fong YW, Tjian R. 2015. Functional and mechanistic studies of XPC DNA-repair complex as transcriptional coactivator in embryonic stem cells. *Proc Natl Acad Sci U S A* 112: E2317-26
76. Lehmann AR, McGibbon D, Stefanini M. 2011. Xeroderma pigmentosum. *Orphanet J Rare Dis* 6: 70
77. Robbins JH, Brumback RA, Polinsky RJ, Wirtschafter JD, Tarone RE, Scudiero DA, Otsuka F. 1985. Hypersensitivity to DNA-damaging agents in abiotrophies: a new explanation for degeneration of neurons, photoreceptors, and muscle in Alzheimer, Parkinson and Huntington diseases, retinitis pigmentosa, and Duchenne muscular dystrophy. *Basic Life Sci* 35: 315-44

78. Brooks PJ. 2008. The 8,5'-cyclopurine-2'-deoxynucleosides: candidate neurodegenerative DNA lesions in xeroderma pigmentosum, and unique probes of transcription and nucleotide excision repair. *DNA Repair (Amst)* 7: 1168-79
79. Kuraoka I, Bender C, Romieu A, Cadet J, Wood RD, Lindahl T. 2000. Removal of oxygen free-radical-induced 5',8-purine cyclodeoxynucleosides from DNA by the nucleotide excision-repair pathway in human cells. *Proc Natl Acad Sci U S A* 97: 3832-7
80. Wijnhoven SW, Hoogervorst EM, de Waard H, van der Horst GT, van Steeg H. 2007. Tissue specific mutagenic and carcinogenic responses in NER defective mouse models. *Mutat Res* 614: 77-94
81. Le May N, Mota-Fernandes D, Velez-Cruz R, Iltis I, Biard D, Egly JM. 2010. NER factors are recruited to active promoters and facilitate chromatin modification for transcription in the absence of exogenous genotoxic attack. *Molecular cell* 38: 54-66
82. Fang EF, Scheibye-Knudsen M, Brace LE, Kassahun H, SenGupta T, Nilsen H, Mitchell JR, Croteau DL, Bohr VA. 2014. Defective mitophagy in XPA via PARP-1 hyperactivation and NAD(+)/SIRT1 reduction. *Cell* 157: 882-96
83. Rodolfo C, Ciccocanti F, Giacomo GD, Piacentini M, Fimia GM. 2010. Proteomic analysis of mitochondrial dysfunction in neurodegenerative diseases. *Expert review of proteomics* 7: 519-42

84. Velez-Cruz R, Zadorin AS, Coin F, Egly JM. 2013. Sirt1 suppresses RNA synthesis after UV irradiation in combined xeroderma pigmentosum group D/Cockayne syndrome (XP-D/CS) cells. *Proc Natl Acad Sci U S A* 110: E212-20
85. Le May N, Fradin D, Iltis I, Bougneres P, Egly JM. 2012. XPG and XPF endonucleases trigger chromatin looping and DNA demethylation for accurate expression of activated genes. *Molecular cell* 47: 622-32
86. Velez-Cruz R, Egly JM. 2013. Cockayne syndrome group B (CSB) protein: at the crossroads of transcriptional networks. *Mechanisms of ageing and development* 134: 234-42
87. Lindahl T, Wood RD. 1999. Quality control by DNA repair. *Science* 286: 1897-905
88. Iyer N, Reagan MS, Wu KJ, Canagarajah B, Friedberg EC. 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: 2157-67
89. Wu X, Shell SM, Liu Y, Zou Y. 2007. ATR-dependent checkpoint modulates XPA nuclear import in response to UV irradiation. *Oncogene* 26: 757-64
90. Jarrett SG, Horrell EM, Christian PA, Vanover JC, Boulanger MC, Zou Y, D'Orazio JA. 2014. PKA-mediated phosphorylation of ATR promotes recruitment of XPA to UV-induced DNA damage. *Mol Cell* 54: 999-1011

91. Li Z, Musich PR, Cartwright BM, Wang H, Zou Y. 2013. UV-induced nuclear import of XPA is mediated by importin- α 4 in an ATR-dependent manner. *PLoS One* 8: e68297
92. Cooper GM. 2000. The Nuclear Envelope and Traffic between the Nucleus and Cytoplasm. In *The Cell: A Molecular Approach*. Sunderland (MA): Sinauer Associates
93. Jamur MC, Oliver C. 2010. Permeabilization of cell membranes. *Methods Mol Biol* 588: 63-6
94. Feeney RJ, Zieve GW. 1990. Nuclear exchange of the U1 and U2 snRNP-specific proteins. *J Cell Biol* 110: 871-81
95. Cadet J, Sage E, Douki T. 2005. Ultraviolet radiation-mediated damage to cellular DNA. *Mutat Res* 571: 3-17
96. Ravanat JL, Douki T, Cadet J. 2001. Direct and indirect effects of UV radiation on DNA and its components. *J Photochem Photobiol B* 63: 88-102
97. Zdzenicka MZ, Venema J, Mitchell DL, van Hoffen A, van Zeeland AA, Vrieling H, Mullenders LH, Lohman PH, Simons JW. 1992. (6-4) photoproducts and not cyclobutane pyrimidine dimers are the main UV-induced mutagenic lesions in Chinese hamster cells. *Mutat Res* 273: 73-83
98. Levy DD, Saijo M, Tanaka K, Kraemer KH. 1995. Expression of a transfected DNA repair gene (XPA) in xeroderma pigmentosum group A

- cells restores normal DNA repair and mutagenesis of UV-treated plasmids. *Carcinogenesis* 16: 1557-63
99. Welsh C, Day R, McGurk C, Masters JR, Wood RD, Koberle B. 2004. Reduced levels of XPA, ERCC1 and XPF DNA repair proteins in testis tumor cell lines. *Int J Cancer* 110: 352-61
100. Biggerstaff M, Wood RD. 2006. Repair synthesis assay for nucleotide excision repair activity using fractionated cell extracts and UV-damaged plasmid DNA. *Methods in molecular biology* 314: 417-34
101. Wood RD, Biggerstaff M, Shivji MKK. 1995. Detection and measurement of nucleotide excision repair synthesis by mammalian cell extracts *in vitro*. *Methods: A Companion to Methods in Enzymology* 7: 163-75
102. Moggs JG, Yarema KJ, Essigmann JM, Wood RD. 1996. Analysis of incision sites produced by human cell extracts and purified proteins during nucleotide excision repair of a 1,3-intrastrand d(GpTpG)-cisplatin adduct. *The Journal of biological chemistry* 271: 7177-86
103. Shivji MK, Moggs JG, Kuraoka I, Wood RD. 2006. Assaying for the dual incisions of nucleotide excision repair using DNA with a lesion at a specific site. *Methods in molecular biology* 314: 435-56
104. Jones CJ, Wood RD. 1993. Preferential binding of the xeroderma pigmentosum group A complementing protein to damaged DNA. *Biochemistry* 32: 12096-104

105. Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL. 2013. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol* 14: R36
106. Pruitt KD, Brown GR, Hiatt SM, Thibaud-Nissen F, Astashyn A, Ermolaeva O, Farrell CM, Hart J, Landrum MJ, McGarvey KM, Murphy MR, O'Leary NA, Pujar S, Rajput B, Rangwala SH, Riddick LD, Shkeda A, Sun H, Tamez P, Tully RE, Wallin C, Webb D, Weber J, Wu W, DiCuccio M, Kitts P, Maglott DR, Murphy TD, Ostell JM. 2014. RefSeq: an update on mammalian reference sequences. *Nucleic Acids Res* 42: D756-63
107. Anders S, Pyl PT, Huber W. 2015. HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics* 31: 166-9
108. Anders S, Huber W. 2010. Differential expression analysis for sequence count data. *Genome Biol* 11: R106
109. Robinson MD, McCarthy DJ, Smyth GK. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26: 139-40
110. Mi H, Poudel S, Muruganujan A, Casagrande JT, Thomas PD. 2016. PANTHER version 10: expanded protein families and functions, and analysis tools. *Nucleic Acids Res* 44: D336-42
111. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, Mesirov JP. 2005. Gene set enrichment analysis: a knowledge-based approach for

- interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* 102: 15545-50
112. Biggerstaff M, Wood RD. 1999. Assay for nucleotide excision repair protein activity using fractionated cell extracts and UV-damaged plasmid DNA. *Methods in molecular biology* 113: 357-72
 113. Enou M, Ho TV, Long DT, Walter JC, Scharer OD. 2012. Construction of plasmids containing site-specific DNA interstrand cross-links for biochemical and cell biological studies. *Methods Mol Biol* 920: 203-19
 114. Satokata I, Tanaka K, Miura N, Miyamoto I, Satoh Y, Kondo S, Okada Y. 1990. Characterization of a splicing mutation in group A xeroderma pigmentosum. *Proc Natl Acad Sci U S A* 87: 9908-12
 115. Satokata I, Tanaka K, Miura N, Narita M, Mimaki T, Satoh Y, Kondo S, Okada Y. 1992. Three nonsense mutations responsible for group-A xeroderma-pigmentosum. *Mutat. Res.* 273: 193-202
 116. Köberle B, Roginskaya V, Wood RD. 2006. XPA protein as a limiting factor for nucleotide excision repair and UV sensitivity in human cells. *DNA Repair* 5: 641-8
 117. Kang TH, Reardon JT, Sancar A. 2011. Regulation of nucleotide excision repair activity by transcriptional and post-transcriptional control of the XPA protein. *Nucleic Acids Res* 39: 3176-87
 118. Koberle B, Roginskaya V, Wood RD. 2006. XPA protein as a limiting factor for nucleotide excision repair and UV sensitivity in human cells. *DNA Repair (Amst)* 5: 641-8

119. Svetlova M, Nikiforov A, Solovjeva L, Pleskach N, Tomilin N, Hanawalt PC. 1999. Reduced extractability of the XPA DNA repair protein in ultraviolet light-irradiated mammalian cells. *FEBS Lett* 463: 49-52
120. Rass U, Ahel I, West SC. 2007. Defective DNA repair and neurodegenerative disease. *Cell* 130: 991-1004
121. Kassahun H, Nilsen H. 2013. Active transcriptomic and proteomic reprogramming in the *C. elegans* nucleotide excision repair mutant xpa-1. *Worm* 2: e27337
122. Arczewska KD, Tomazella GG, Lindvall JM, Kassahun H, Maglioni S, Torgovnick A, Henriksson J, Matilainen O, Marquis BJ, Nelson BC, Jaruga P, Babaie E, Holmberg CI, Burglin TR, Ventura N, Thiede B, Nilsen H. 2013. Active transcriptomic and proteomic reprogramming in the *C. elegans* nucleotide excision repair mutant xpa-1. *Nucleic Acids Res* 41: 5368-81
123. Hartman PS, Herman RK. 1982. Radiation-sensitive mutants of *Caenorhabditis elegans*. *Genetics* 102: 159-78
124. Sahin E, Colla S, Liesa M, Moslehi J, Muller FL, Guo M, Cooper M, Kotton D, Fabian AJ, Walkey C, Maser RS, Tonon G, Foerster F, Xiong R, Wang YA, Shukla SA, Jaskelioff M, Martin ES, Heffernan TP, Protopopov A, Ivanova E, Mahoney JE, Kost-Alimova M, Perry SR, Bronson R, Liao R, Mulligan R, Shiriha OS, Chin L, DePinho RA. 2011. Telomere dysfunction induces metabolic and mitochondrial compromise. *Nature* 470: 359-65

125. Wisnovsky S, Jean SR, Kelley SO. 2016. Mitochondrial DNA repair and replication proteins revealed by targeted chemical probes. *Nat Chem Biol* 12: 567-73
126. You ZY, Feaver WJ, Friedberg EC. 1998. Yeast RNA-polymerase-II transcription in vitro is inhibited in the presence of nucleotide excision-repair - complementation of inhibition by holo-TFIIH and requirement for Rad26. *Mol Cell Biol* 18: 2668-76
127. Sabarinathan R, Mularoni L, Deu-Pons J, Gonzalez-Perez A, Lopez-Bigas N. 2016. Nucleotide excision repair is impaired by binding of transcription factors to DNA. *Nature* 532: 264-7
128. Perera D, Poulos RC, Shah A, Beck D, Pimanda JE, Wong JW. 2016. Differential DNA repair underlies mutation hotspots at active promoters in cancer genomes. *Nature* 532: 259-63
129. Bradford PT, Goldstein AM, Tamura D, Khan SG, Ueda T, Boyle J, Oh KS, Imoto K, Inui H, Moriwaki S, Emmert S, Pike KM, Raziuddin A, Plona TM, DiGiovanna JJ, Tucker MA, Kraemer KH. 2011. Cancer and neurologic degeneration in xeroderma pigmentosum: long term follow-up characterises the role of DNA repair. *J Med Genet* 48: 168-76
130. Robbins JH, Kraemer KH, Lutzner MA, Festoff BW, Coon HG. 1974. Xeroderma pigmentosum. An inherited diseases with sun sensitivity, multiple cutaneous neoplasms, and abnormal DNA repair. *Ann Intern Med* 80: 221-48

131. Andrews AD, Barrett SF, Robbins JH. 1978. Xeroderma pigmentosum neurological abnormalities correlate with colony-forming ability after ultraviolet radiation. *Proc Natl Acad Sci U S A* 75: 1984-8
132. Jin Y, Mesaros AC, Blair IA, Penning TM. 2011. Stereospecific reduction of 5 β -reduced steroids by human ketosteroid reductases of the AKR (aldo-keto reductase) superfamily: role of AKR1C1-AKR1C4 in the metabolism of testosterone and progesterone via the 5 β -reductase pathway. *Biochem J* 437: 53-61
133. Veilleux A, Cote JA, Blouin K, Nadeau M, Pelletier M, Marceau P, Laberge PY, Luu-The V, Tchernof A. 2012. Glucocorticoid-induced androgen inactivation by aldo-keto reductase 1C2 promotes adipogenesis in human preadipocytes. *Am J Physiol Endocrinol Metab* 302: E941-9
134. Stolz A, Hammond L, Lou H, Takikawa H, Ronk M, Shively JE. 1993. cDNA cloning and expression of the human hepatic bile acid-binding protein. A member of the monomeric reductase gene family. *J Biol Chem* 268: 10448-57
135. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G. 2000. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet* 25: 25-9

136. Rappaport N, Twik M, Nativ N, Stelzer G, Bahir I, Stein TI, Safran M, Lancet D. 2014. MalaCards: A Comprehensive Automatically-Mined Database of Human Diseases. *Curr Protoc Bioinformatics* 47: 1 24 1-19
137. Nakane H, Hirota S, Brooks PJ, Nakabeppu Y, Nakatsu Y, Nishimune Y, Iino A, Tanaka K. 2008. Impaired spermatogenesis and elevated spontaneous tumorigenesis in xeroderma pigmentosum group A gene (Xpa)-deficient mice. *DNA Repair (Amst)* 7: 1938-50
138. Kraemer KH, Lee MM, Scotto J. 1987. Xeroderma pigmentosum. Cutaneous, ocular, and neurologic abnormalities in 830 published cases. *Arch Dermatol* 123: 241-50
139. van Kreijl CF, McAnulty PA, Beems RB, Vynckier A, van Steeg H, Franssón-Steen R, Alden CL, Forster R, van der Laan JW, Vandenberghe J. 2001. Xpa and Xpa/p53^{+/-} knockout mice: overview of available data. *Toxicol Pathol* 29 Suppl: 117-27
140. Zachmann M, Vollmin JA, Hamilton W, Prader A. 1972. Steroid 17,20-desmolase deficiency: a new cause of male pseudohermaphroditism. *Clin Endocrinol (Oxf)* 1: 369-85
141. Jacque N, Ronchetti AM, Larrue C, Meunier G, Birsén R, Willems L, Saland E, Decroocq J, Maciel TT, Lambert M, Poulain L, Hospital MA, Sujobert P, Joseph L, Chapuis N, Lacombe C, Moura IC, Demo S, Sarry JE, Recher C, Mayeux P, Tamburini J, Bouscary D. 2015. Targeting glutaminolysis has antileukemic activity in acute myeloid leukemia and synergizes with BCL-2 inhibition. *Blood* 126: 1346-56

142. Farrelly E, Amaral MC, Marshall L, Huang SG. 2001. A high-throughput assay for mitochondrial membrane potential in permeabilized yeast cells. *Anal Biochem* 293: 269-76
143. Condo I, Ventura N, Malisan F, Tomassini B, Testi R. 2006. A pool of extramitochondrial frataxin that promotes cell survival. *J Biol Chem* 281: 16750-6
144. Dey R, Moraes CT. 2000. Lack of oxidative phosphorylation and low mitochondrial membrane potential decrease susceptibility to apoptosis and do not modulate the protective effect of Bcl-x(L) in osteosarcoma cells. *J Biol Chem* 275: 7087-94
145. Green DR, Reed JC. 1998. Mitochondria and apoptosis. *Science* 281: 1309-12
146. Wang G, Seidman MM, Glazer PM. 1996. Mutagenesis in mammalian cells induced by triple helix formation and transcription-coupled repair. *Science* 271: 802-5
147. Eichman BF, Mooers BH, Alberti M, Hearst JE, Ho PS. 2001. The crystal structures of psoralen cross-linked DNAs: drug-dependent formation of Holliday junctions. *J Mol Biol* 308: 15-26
148. Wang X, Peterson CA, Zheng H, Nairn RS, Legerski RJ, Li L. 2001. Involvement of nucleotide excision repair in a recombination-independent and error-prone pathway of DNA interstrand cross-link repair. *Mol Cell Biol* 21: 713-20

149. Majumdar A, Khorlin A, Dyatkina N, Lin FL, Powell J, Liu J, Fei Z, Khripine Y, Watanabe KA, George J, Glazer PM, Seidman MM. 1998. Targeted gene knockout mediated by triple helix forming oligonucleotides. *Nat Genet* 20: 212-4
150. Semlow DR, Zhang J, Budzowska M, Drohat AC, Walter JC. 2016. Replication-Dependent Unhooking of DNA Interstrand Cross-Links by the NEIL3 Glycosylase. *Cell* 167: 498-511 e14
151. Boulware SB, Christensen LA, Thames H, Coghlan L, Vasquez KM, Finch RA. 2013. Triplex-forming oligonucleotides targeting c-MYC potentiate the anti-tumor activity of gemcitabine in a mouse model of human cancer. *Molecular carcinogenesis*
152. Dahmen V, Kriehuber R. 2012. Cytotoxic effects and specific gene expression alterations induced by I-125-labeled triplex-forming oligonucleotides. *Int J Radiat Biol* 88: 972-9
153. Kuznetsova S, Ait-Si-Ali S, Nagibneva I, Troalen F, Le Villain JP, Harel-Bellan A, Svinarchuk F. 1999. Gene activation by triplex-forming oligonucleotide coupled to the activating domain of protein VP16. *Nucleic Acids Res* 27: 3995-4000
154. Vekhoff P, Ceccaldi A, Polverari D, Pylouster J, Pisano C, Arimondo PB. 2008. Triplex formation on DNA targets: how to choose the oligonucleotide. *Biochemistry* 47: 12277-89
155. Mukherjee A, Vasquez KM. 2011. Triplex technology in studies of DNA damage, DNA repair, and mutagenesis. *Biochimie* 93: 1197-208

156. Goni JR, de la Cruz X, Orozco M. 2004. Triplex-forming oligonucleotide target sequences in the human genome. *Nucleic Acids Res* 32: 354-60
157. Perrouault L, Asseline U, Rivalle C, Thuong NT, Bisagni E, Giovannangeli C, Le Doan T, Helene C. 1990. Sequence-specific artificial photo-induced endonucleases based on triple helix-forming oligonucleotides. *Nature* 344: 358-60
158. Christensen LA, Wang H, Van Houten B, Vasquez KM. 2008. Efficient processing of TFO-directed psoralen DNA interstrand crosslinks by the UvrABC nuclease. *Nucleic Acids Res* 36: 7136-45
159. Liu M, Bandaru V, Bond JP, Jaruga P, Zhao X, Christov PP, Burrows CJ, Rizzo CJ, Dizdaroglu M, Wallace SS. 2010. The mouse ortholog of NEIL3 is a functional DNA glycosylase in vitro and in vivo. *Proc Natl Acad Sci U S A* 107: 4925-30

Vita

Mandira Manandhar was born in Kathmandu, Nepal on December 2nd, 1982, the daughter of Bina Manandhar and Shreekrishna Manandhar. After completing her work at Holy Vision High School, Kathmandu, Nepal in 2001, she entered Tribhuwan University in Kathmandu, Nepal. She received the degree of Bachelor of Science with a major in microbiology in 2004. She transferred to the undergraduate program in Idaho State University in Pocatello, Idaho, USA in August 2005. She received the degree of Bachelor of Science with a major in microbiology from Idaho State University in May 2006. For the next one year she worked as a laboratory investigator in Abbott Laboratories, Alameda, California. She entered Idaho State University again in August 2007. She received the degree of Master in Science with a major in microbiology from Idaho State University in August 2009. She entered the University of Texas Graduate School of Biomedical Sciences at Houston in January 2010.