

**Functional relevance of spontaneous alternative splice variants of
xeroderma pigmentosum genes: Prognostic marker for skin
cancer risk and disease outcome?**

Doctoral Thesis

In partial fulfillment of the requirements for the degree

“Doctor rerum naturalium (Dr. rer. nat.)“

in the Molecular Medicine Study Program

at the Georg-August University Goettingen

submitted by

Janin Lehmann

born in Westerstede, Germany

Goettingen, March 9, 2017

Members of the thesis committee

Supervisor:

Prof. Dr. med. Steffen Emmert

University Medical Center, Clinic and Policlinic for Dermatology and Venereology

University of Rostock

First member of the thesis committee:

Prof. Dr. med. Michael P. Schön

University Medical Center, Dept. of Dermatology, Venereology and Allergology

Georg-August University Goettingen

Second member of the thesis committee:

Prof. Dr. rer. nat. Steven Johnsen

University Medical Center, Clinic for General, Visceral and Pediatric Surgery

Georg-August University Goettingen

Third member of the thesis committee:

Prof. Dr. rer. nat. Dieter Kube

University Medical Center, Dept. of Hematology and Oncology

Georg-August University Goettingen

Date of Disputation: May 4, 2017

AFFIDAVIT

Here I declare that my doctoral thesis entitled

“Functional relevance of spontaneous alternative splice variants of xeroderma pigmentosum genes: Prognostic marker for skin cancer risk and disease outcome?”

has been written independently with no other sources and aids than quoted.

Date

Signature (Janin Lehmann)

Firstly, I would like to thank my supervisor Prof. Dr. med. Steffen Emmert for giving me the opportunity to write this thesis in his group and for all the successful publications during this time. I am very thankful for his inspiring guidance and his constructive and critical support during this project.

I greatly want to acknowledge the members of my thesis committee, Prof. Dr. med. Michael P. Schön, Prof. Dr. rer. nat. Steven Johnsen and Prof. Dr. rer. nat. Dieter Kube for their advice regarding research and academic life.

I am especially indebted to Dr. rer. nat. Christina Seebode for her enormous encouragement and support as my friend and colleague.

I would like to thank all my friends and colleagues in Goettingen and Rostock for their support during the development of this thesis

Especially, I would like to express my gratitude to Dr. rer. nat. Steffen Schubert, Antje Apel and Dr. rer. nat. Andreas Ohlenbusch.

I wish to acknowledge the German Cancer Aid (Deutsche Krebshilfe e.V., grant number 111377) for funding of this project.

Finally, I want to thank my parents and friends for their encouragement and support during the entire course of my PhD.

List of Publications

Stojadinovic O, Yin N, Lehmann J, Pastar I, Kirsner RS, Tomic- Canic M. “Increased Number of Langerhans Cells in the Epidermis of Diabetic Foot Ulcers Correlates with Healing Outcome.” *Immunol Res.* 2013; 57(1-3):222-8.

Schubert S, Lehmann J, Kalfon L, Slor H, Falik-Zaccai T, Emmert S. “CUGC (clinical utility gene card) for Xeroderma Pigmentosum.” *Eur J Hum Genet*, 2014; 22, doi: 10.1038/ejhg.2013.233.

Lehmann J, Schubert S, Schäfer A, Gratchev A, Apel A, Laspe P, Schiller S, Ohlenbusch A, Emmert S (2013). “An unusual mutation in the XPG gene leads to an internal in-frame deletion and a XP/CS complex phenotype.” *Br J Dermatol.* 2014; 171(4):903-5.

Lehmann J, Schubert S, Emmert S. “Xeroderma pigmentosum (XP): Diagnostic procedures, interdisciplinary patient care, and novel therapeutic approach.” *J Dtsch Dermatol Ges.* 2014; 12(10):867-72.

Lehmann J, Schubert S, Schäfer A, Laspe P, Haenssle HA, Ohlenbusch A, Gratchev A, Emmert S. “A novel mutation in the XPA gene results in two truncated protein variants and leads to a severe XP/neurological symptoms phenotype.” *J Eur Acad Dermatol Venereol.* 2015; 29(12):2479-82.

Schiller S, Schubert S, Lehmann J, Seebode C, Smolorz S, Tiede R, Apel A, Laspe P, Emmert S. “Von seltenen genetischen Erkrankungen lernen: Hautkrebs und DNA Reparatur, Ichthyosen und epidermale Differenzierung sowie kaltes Atmosphärendruckplasma als neue Therapiemodalität.“ *Spitzenforschung in der Dermatologie. Innovationen und Auszeichnungen 2014*, hrsg. von der ALPHA Informations-GmbH, Lampertheim 2014, S. 62 - 70.

Brauns B, Schubert S, Lehmann J, Laspe P, Körner A, Brockmann K, Schön MP, Emmert S. “Photosensitive form of trichothiodystrophy associated with a novel mutation in the XPD gene.” *Photodermatology, Photoimmunology & Photomedicine*, 2016, 32(2):110-112.

Seebode C, Lehmann J, Emmert S. “Photocarcinogenesis versus skin cancer prevention strategies.“ *Anticancer Research*, 2016, 36:1371-8.

Rump A, Benet-Pages A, Schubert S, Kuhlmann JD, Janavičius R, Macháčková E, Foretová L, Kleibl Z, Lhota F, Zemankova P, Betcheva-Krajcir E, Mackenroth L, Hackmann K, Lehmann J, Nissen A, DiDonato N, Opitz R, Thiele H, Kast K, Wimberger P, Holinski-Feder E, Emmert S, Schröck E, Klink B. “Identification and Functional Testing of ERCC2 Mutations in a Multi-national Cohort of Patients with Familial Breast- and Ovarian Cancer.” *PLoS Genet.* 2016 Aug 9;12(8):e1006248.

Schubert S, Rieper P, Ohlenbusch A, Seebode C, Lehmann J, Gratchev A, Emmert S. “A unique chromosomal in-frame deletion identified among seven XP-C patients.” *Photodermatol Photoimmunol Photomed* 2016 Sep 32(5-6):276-283.

Lehmann J, Seebode C, Smolorz S, Schubert S, Emmert S. “*XPF* knockout via CRISPR/Cas9 reveals that ERCC1 is retained in the cytoplasm without its heterodimer partner XPF.” *Cell Mol Life Sci*. 2017 Jan 27. doi: 10.1007/s00018-017-2455-7.

Lehmann J, Seebode C, Martens MC, Emmert S. “Xeroderma pigmentosum – Facts and Perspectives.” *Aktuelle Dermatologie* 2017, in press.

Martens MC, Seebode C, Lehmann J, Emmert S. “Molecular mechanisms of cutaneous photocarcinogenesis: an update.” *Aktuelle Dermatologie* 2017, in press.

Lehmann J, Seebode C. “Research on genodermatoses using novel Genome Editing Tools.” *J Dtsch Dermatol Ges* 2017, submitted.

Schubert S, Lehmann J, Ohlenbusch A, Platten J, Apel A, Kramer W, Emmert S. “Functional characterisation of XPG and its spontaneous splice variants during nucleotide excision repair.” *DNA repair* 2017, submitted.

Lehmann J, Seebode C, Schubert S, Emmert S. “Splice variants of the endonuclease XPF present a valuable tool for personalized medicine.” 2017, in preparation.

Table of contents

List of Publications.....	III
List of Tables.....	IX
List of Figures	X
Abstract	XII
List of Abbreviations.....	XIV
1 Introduction.....	1
1.1 UV irradiation and the skin	1
1.1.1 Penetration.....	1
1.1.2 UV induced DNA lesions.....	2
1.1.3 The multistep carcinogenesis model	3
1.2 DNA repair	4
1.2.1 DNA damage response (DDR).....	4
1.2.2 DNA repair mechanisms	5
1.2.2.1 Nucleotide excision repair (NER).....	5
1.2.2.2 Interstrand Crosslink Repair	8
1.2.2.3 DNA double strand repair pathways.....	10
1.3 DNA repair deficiency disorders	11
1.3.1 Xeroderma pigmentosum	11
1.3.2 Other associated disorders.....	14
1.3.3 Genotype-phenotype-correlation.....	15
1.4 Multiple functions of the XPG and XPF/ERCC1 endonucleases.....	16
1.4.1 XPG.....	17
1.4.2 XPF/ERCC1	19
1.4.3 XP-G and XP-F patients.....	21
1.5 Genome editing techniques	22
1.5.1 TALEN, Zinc finger and meganucleases	22
1.5.2 CRISPR/Cas9	23
1.5.3 Limitations and precautions	25
1.6 Alternative mRNA Splicing	25
1.6.1 Mechanism of mRNA splicing.....	26
1.6.2 Splice site mutations and polymorphisms.....	27
1.7 Aim of the study	29

2	Materials	32
2.1	Biological material	32
2.1.1	Cell lines.....	32
2.1.2	Bacteria.....	32
2.2	Consumable supplies	32
2.3	Equipment.....	33
2.4	Chemicals	34
2.5	Buffers, solutions and media	36
2.6	Oligonucleotides	37
2.7	Ready to use reaction systems	41
2.8	Enzymes.....	42
2.9	Plasmids.....	42
2.10	Antibodies.....	42
2.11	Software and online tools	43
3	Methods.....	45
3.1	Molecular Biology	45
3.1.1	Ethanol precipitation	45
3.1.2	Nucleic acid quantitation.....	45
3.1.3	Extraction of nucleic acids	45
3.1.3.1	Isolation of genomic DNA.....	45
3.1.3.2	Isolation of mRNA and cDNA synthesis.....	45
3.1.3.3	Alkaline lysis plasmid extraction.....	46
3.1.3.4	Agarose gel electrophoresis (AGE)	47
3.1.3.5	Extraction of DNA from agarose gel.....	47
3.1.4	Enzymatic manipulation of DNA.....	48
3.1.4.1	Polymerase chain reaction (PCR)	48
3.1.4.2	Quantitative Real-time PCR	49
3.1.4.3	Site-directed Mutagenesis (Quickchange PCR).....	50
3.1.4.4	Plasmid vector dephosphorylation, oligonucleotide phosphorylation	50
3.1.4.5	Restriction digestion of DNA	50
3.1.4.6	Ligation	51
3.1.4.7	DNA sequencing and sequence analysis.....	51
3.1.4.8	Enzyme mismatch cleavage	52
3.1.4.9	CRISPR construct generation	52

3.2	Microbiology	53
3.2.1	Sterilization and autoclavation	53
3.2.2	Generation of chemically competent <i>Escherichia coli</i> (<i>E. coli</i>) DH5 α	53
3.2.3	Transformation of <i>E. coli</i> DH5 α	53
3.3	Cell Biology	54
3.3.1	Cell culture techniques	54
3.3.1.1	Culture of primary human fibroblasts and immortalized cell lines	54
3.3.1.2	Cell counting	54
3.3.1.3	Freezing and thawing cells	55
3.3.1.4	Transient Transfection	55
3.3.1.5	CRISPR/Cas9 transfection and single clone expansion	55
3.3.2	Functional assays	56
3.3.2.1	Host Cell Reactivation Assay (HCR)	56
3.3.2.2	DSB Assay	57
3.3.2.3	Determination of post-toxin cell survival	58
3.3.2.4	Fluorescence microscopy	59
3.4	Biochemical methods	59
3.4.1	Preparation of protein lysates	59
3.4.2	Bradford protein quantification	60
3.4.3	Polyacrylamide gel electrophoresis (SDS-PAGE)	60
3.4.4	Western blotting	61
3.4.5	Membrane stripping	61
3.5	Statistics analyses	61
4	Results	62
4.1	A novel mutation in the <i>XPA</i> gene results in two truncated protein variants and leads to a severe XP/neurological symptoms phenotype	62
4.2	Establishment of a complete <i>XPF</i> and <i>ERCCI</i> CRISPR/Cas9 knockout in MRC5Vi cells	65
4.2.1	Characterization of an <i>XPF</i> CRISPR/Cas9 KO single clone	65
4.2.2	A viable complete <i>ERCCI</i> KO cannot be generated	70
4.2.3	<i>XPF</i> KO cells show an increased sensitivity to several DNA damaging toxins	71
4.2.4	Loss of <i>XPF</i> reduces the cellular repair capability for NER, ICL, and HRR	73
4.3	Amplification and characterization of <i>XPF</i> and <i>ERCCI</i> splice variants	75
4.3.1	Protein expression and subcellular localization of <i>XPF</i> and <i>ERCCI</i> splice variants	76

4.4	Splice variants and their involvement in different DNA repair pathways	78
4.4.1	<i>XPG</i> isoforms V and VI significantly increase the repair capability of intra- and interstrand crosslinks	79
4.4.2	<i>XPF</i> splice variants show residual repair capabilities during NER and ICL repair	80
4.4.3	<i>XPG</i> and <i>XPF</i> splice variants do exhibit dominant negative effects during DNA repair.....	81
4.4.4	Analyses of <i>XPF</i> point mutants in the newly generated <i>XPF</i> KO cells give insights into mechanistic aspects of <i>XPF</i> 's functions in different DNA repair pathways.....	84
4.5	Functionally relevant splice variants can be implicated as prognostic markers for individual cancer risk, therapeutic success, or disease outcome	88
5	Discussion.....	91
5.1	A large deletion in the <i>XPA</i> gene results in XP with severe neurological symptoms.....	91
5.2	The <i>XPF</i> CRISPR/Cas9 KO cells present a great tool to model XP.....	93
5.3	ERCC1 is retained in the cytosol without its heterodimeric partner <i>XPF</i>	95
5.4	<i>XPF</i> is markedly involved in HRR but dispensable for NHEJ.....	95
5.5	<i>XPF</i> and ERCC1 splice variants could successfully be cloned from wildtype fibroblasts, show stable expression, and localize to the nucleus	97
5.6	Splice variants of the two endonucleases <i>XPG</i> and <i>XPF</i> show residual repair capabilities in NER and ICL repair	98
5.7	<i>XPG</i> and <i>XPF</i> splice variants exert a dominant negative effect on wild type NER capacities	102
5.8	Artificially generated <i>XPF</i> point mutants behave differently in the newly generated <i>XPF</i> KO cells.....	104
5.9	Importance of <i>XPG</i> isoforms for personalized medicine and further perspectives	106
5.10	Summary and conclusions	108
6	Appendix	110
7	Bibliography	114
8	Curriculum vitae	134

List of Tables

Table 1: XP complementation groups	14
Table 2: Cell lines	32
Table 3: Consumables	32
Table 4: Equipment	33
Table 5: Chemicals.....	34
Table 6: Buffers, solutions and media.....	36
Table 7: Oligonucleotides	38
Table 8: Reaction systems.....	41
Table 9: Enzymes	42
Table 10: Plasmids	42
Table 11: Primary antibodies	43
Table 12: Secondary antibodies	43
Table 13: List of utilized software and online tools.....	43

List of Figures

Figure 1: Penetration depth and effects of UV irradiation	2
Figure 2: UV-induced DNA lesions	3
Figure 3: Schematic overview of the NER pathway	8
Figure 4: Schematic overview of interstand crosslink repair (ICL)	10
Figure 5: Picture of two XP patients with typical clinical features	13
Figure 6: 12 clinical entities and 14 molecular defects show the complex genotype- phenotype correlation of DNA repair deficiency disorders	16
Figure 7: Schematic overview of XPG and its domains	18
Figure 8: Schematic overview of XPF/ERCC1 and their domains	20
Figure 9: Schematic view of XPF/ERCC1 and mutations associated with DNA repair defective disorders	22
Figure 10: Schematic overview of the CRISPR/Cas9 system of <i>Streptococcus pyogenes</i>	24
Figure 11: Major spliceosomal assembly	27
Figure 12: Overview of the <i>XPG</i> spontaneous mRNA splice variants	29
Figure 13: Overview of the <i>XPF</i> and <i>ERCC1</i> spontaneous mRNA splice variants	31
Figure 14: Simplified scheme of the HCR	57
Figure 15: Schematic illustration of the DSB assay principle	58
Figure 16: Characterization of four new XP-A patients	64
Figure 17: Structural analyses of <i>XPF</i> CRISPR/Cas9 KO and WT MRC5Vi cells	68
Figure 18: Decrease in <i>XPF</i> expression in <i>XPF</i> KO cells assessed by quantitative real- time PCR	70
Figure 19: Generation of a heterozygous <i>ERCC1</i> KO via CRISPR/Cas9 genome editing	71
Figure 20: WT MRC5Vi and <i>XPF</i> KO post-toxin cell survival analyses	73
Figure 21: Reactivation of a reporter gene after treatment with UVC, cisplatin or trimethylpsoralen activated by UVA light in <i>XPF</i> KO and WT MRC5Vi cells	74
Figure 22: Analyses of HRR und NHEJ repair pathways in <i>XPF</i> KO and WT MRC5Vi cells.	75
Figure 23: Immunoblot results for protein levels of <i>XPF</i> splice variants over time	77
Figure 24: Subcellular localization of eGFP-tagged <i>XPF</i> isoforms and patient alleles	78
Figure 25: Reactivation of a reporter gene after treatment with UVC, cisplatin or trimethylpsoralen activated by UVA light in XP20BE patient or primary WT fibroblasts	80
Figure 26: Reactivation of a reporter gene after treatment with UVC, cisplatin or trimethylpsoralen activated by UVA light in <i>XPF</i> KO cells complemented with <i>XPF</i> splice variants	81
Figure 27: Immunoblot analyses and reactivation of a reporter gene after treatment with UVC in MRC5Vi WT cells and single clones overexpressing <i>XPG</i> , <i>XPF</i> or respective splice variants	83
Figure 28: Overview of <i>XPF</i> point mutants	85
Figure 29: Reactivation of a reporter gene after treatment with UVC, cisplatin or trimethylpsoralen activated by UVA light in <i>XPF</i> KO cells complemented with <i>XPF</i> point mutants	86

Figure 30: Immunoblot results for XPF point mutants generated by site-directed mutagenesis	87
Figure 31: Immunoblot results for cytosolic and nuclear fractions of XPF point mutants generated by site-directed mutagenesis.....	88
Figure 32: Comparative analysis of <i>XPG</i> expression in human tissues.....	89
Figure 33: Analysis of <i>XPG</i> expression in human blood samples	90
Figure A1: Immunoblot results for wildtype XPG, the seven isoforms and two patient alleles after overexpression in HeLa cells.....	110
Figure A2: Subcellular localization of eGFP-tagged <i>XPG</i> isoforms and patient alleles	111
Figure A3: Immunoblot results for protein levels of <i>ERCC1</i> splice variants over time	112
Figure A4: Subcellular localization of eGFP-tagged <i>ERCC1</i> isoforms and patient alleles ...	113

Abstract

The nucleotide excision repair (NER) pathway is a central DNA repair mechanism to repair a variety of bulky DNA lesions. Accumulation of these types of damage all over the genome results in the development of a cancer prone mutator phenotype, as it can be seen in patients with the autosomal recessive disease xeroderma pigmentosum, and their high frequency of ultraviolet induced skin tumors. It is known that decreased NER level are a risk factor for several cancer entities. Several components of the NER pathway already serve as biomarkers for cancer risk and treatment success. The endonucleases XPF/ERCC1 and XPG are the core components of the incision complex of NER. XPF/ERCC1 is also involved in the repair of DNA interstrand crosslinks (ICL). Due to the essential roles of this complex, patient cells retain at least one full-length allele and residual repair capabilities, rendering them unsuitable for XPF variant analyses.

In the course of this thesis, the CRISPR/Cas9 technology was established in the laboratory and applied to generate an *XPF* knockout in a fetal lung fibroblast cell line (MRC5Vi cells), to analyze the unknown functional relevance of physiologically occurring, spontaneous *XPF* mRNA splice variants. Furthermore, functional roles of XPF point mutants in NER and ICL repair were investigated.

The successfully generated *XPF* knockout cells were markedly sensitive to UVC, cisplatin, and PUVA (psoralen activated by UVA) and had reduced repair capabilities for NER and ICL repair as assessed by reporter gene assays. Using the knockout cells it was shown that human XPF is predominately involved in homologous recombination repair but dispensable for non-homologous end-joining. Notably, while ERCC1 was stably expressed in the cytosol, it was not detectable in the nucleus without its heterodimeric partner XPF, implicating the necessity of functional XPF to retain ERCC1 in the nucleus. Overexpression of wildtype XPF reversed these effects. Functional analyses revealed two *XPF* splice variants with residual repair capabilities (XPF-201 and XPF-003) in NER, as well as ICL repair. XPF-201 lacks the first 12 amino acids of the protein, while XPF-003 is severely C-terminally truncated. Interestingly, another variant, XPF-202, which differs to XPF-003 in the first 12 amino acids only, had no repair capability whatsoever, suggesting an important role of this protein region. It might be involved in interacting with other proteins of the DNA repair machinery.

Splice variants of *XPF* and *XPG*, already characterized during my master thesis, were identified to exert dominant negative effects on NER, when stably overexpressed in wildtype cells. Additionally, the newly generated KO cells represent a highly promising tool for

mechanistic studies. In this cellular background without XPF expression, point mutants showed different catalytic activities compared to reconstituted *in vitro* systems, which are limited by the artificial combination of recombinant proteins, or patient cell lines retaining at least one full-length allele.

Finally, it was shown that the *XPF* and *XPG* splice variants varied in their inter-individual expression in healthy donors, as well as in various tissues. Together with their residual repair capability, dominant-negative effects, and different expression levels, functionally relevant spontaneous splice variants of *XPF* and *XPG* present promising prognostic marker candidates for individual cancer risk, disease outcome, and therapeutic success. Association studies and translational research within clinical trials will have to confirm this assumption in the future.

List of Abbreviations

6-4PP	pyrimidine (6-4) pyrimidone photoproduct
°C	degree Celsius
µg	micro gram
A	adenine
aa	amino acid
ATM	ataxia-telangiectasia mutated
ATP	adenosine triphosphate
ATPase	adenosine triphosphatase
ATR	Ataxia telangiectasia and Rad3-related protein
bp	base pair
BCC	basal cell carcinoma
BD	Big Dye
bidest	double distilled water
BER	base excision repair
C	cytosine
CAK	cdk-activating kinase
CaCl ₂	calcium chloride
cdk	cyclin-dependent kinase
cDNA	complementary DNA
CMV	cytomegalo virus
COFS	cerebro-oculo-facio-skeletal
CP	cisplatin
CPD	cyclobutane pyrimidine dimer
CS	Cockayne syndrome
DDB1/2	DNA damage-binding protein 1/2
DDR	DNA damage response
dH ₂ O	distilled water
DMEM	Dulbecco's modified Eagle's medium
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
dNTP (N = A, C, T or G)	desoxyribonucleotide
ddNTP (N = A, C, T or G)	didesoxyribonucleotide
ds	double strand
DSB	double strand break
<i>E. coli</i>	<i>Escherichia coli</i>
EDTA	ethylenediaminetetraacetic acid
e.g.	for example
eGFP	enhanced green-fluorescent protein

List of Abbreviations

ERCC	excision repair cross-complementing
et al.	Lat.: et alteri
EtOH	ethanol
FA	Fanconi anemia
FBS	fetal bovine serum
FEN1	flap structure-specific endonuclease
fs	femto second
for	forward
gDNA	genomic DNA
G	guanine
GER	Germany
GGR	global genome repair
HCl	hydrochlorid acid
HR	homologous recombination
HRR	homologous recombination repair
hr(s)	hour(s)
ICL	interstrand crosslink
Ig	Immunoglobulin
IR	ionizing radiation
JPN	Japan
KO	knockout
kb	kilo bases
kDa	kilo dalton
LB	Lysogeny Broth
MAPK	mitogen-activated protein kinase
mcs	multiple cloning site
min	minute
mol	molar
ml	milliliter
mRNA	messenger RNA
MTT	3-(4,5-Dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide
MW	molecular weight
mer	repeat unit
MgCl ₂	magnesium chloride
min	minute(s)
mM	milli molar
mm	millimeters
MMEJ	microhomology-mediated end joining
NaCl	sodium chloride

List of Abbreviations

NaHCO ₃	sodium bicarbonate
NaOH	sodium hydroxide
NER	nucleotide excision repair
NHEJ	non-homologous end joining
NMSC	non-melanoma skin cancer
nt	nucleotide
o/n	over night
OD _{xxnm}	optical density at xxnm
oligo	oligonucleotide
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate-buffered saline
PCR	polymerase chain reaction
PCNA	proliferating cell nuclear antigen
PFA	paraformaldehyd
pH	potentium hydrogenii
PH	pleckstrin homology
PIP	PCNA interacting protein
PMSF	phenylmethanesulfonylfluoride
PTB	phosphotyrosine binding
P/S	penicillin – streptavidin
qRT-PCR	quantitative real-time PCR
RB	retinoblastoma
rev	reverse
RLT buffer	RNeasy Lysis buffer Qiagen
RNA	ribonucleic acid
ROS	reactive oxygen species
RPA	replication protein A
rpm	rounds per minute
RT	room temperature
RLU	relative light units
RNA PolII	RNA Polymerase II
RPA	replication protein A
SSA	single strand annealing
SCC	squamous cell carcinoma
SDS	sodium dodecyl sulfate
sec	seconds
SNP	single nucleotide polymorphism
snRNP	small nuclear ribonuclear protein
snRNA	small nuclear RNA

List of Abbreviations

ss	single strand
SEM	standard error of the mean
T	thymidine
TBE	Tris-Borat-EDTA
TCR	transcription coupled repair
ter	premature stop codon
<i>Taq</i>	<i>Thermus aquaticus</i>
TFIIH	transcription factor II H
TMP	trimethylpsoralen
TTD	Trichothiodystrophy
TEMED	N,N,N,N-tetramethyl-ethane-1,2-diamine
Tris	Tris(hydroxymethyl)-aminomethane
UK	United Kingdom
UMG	University Medical Center Goettingen
USA	United states of America
UV	ultraviolet
V	volts
v/v	volume/volume
w/v	weight/volume
WT	wild type
XP A-G	Xeroderma Pigmentosum complementation group A-G
XRCC1	X-ray repair cross-complementing protein 1
μl	micro liter

1 Introduction

The genetic information necessary for growth, development, functioning, and reproduction of all known living organisms, except RNA viruses, is coded in the deoxyribonucleic acid (DNA). According to a central dogma, DNA is transcribed into messenger RNA (mRNA), and mRNA is then translated into proteins (protein biosynthesis), which fulfill all sorts of different functions in the cell (Crick, 1970). A living organism is constantly exposed to different exogenous and endogenous DNA damaging agents that can alter the genomic sequence, causing impairment of protein function, resulting in accumulation of defective proteins, and finally leading to a mutator phenotype (reviewed in Bertram, 2000). On the one hand, this is an evolutionary engine, but on the other hand can cause carcinogenesis in case of mutated tumor suppressor genes (Goh *et al.*, 2011; Hoeijmakers, 2009; Hollstein *et al.*, 1991).

1.1 UV irradiation and the skin

Electromagnetic radiation with a range from 100-400nm, UV light, is a commonly known, very potent mutagen, and can be subdivided into UVA, UVB and UVC (reviewed in Seebode *et al.*, 2016).

1.1.1 Penetration

UVC irradiation (100-280nm) comprises the shortest wavelength, but the highest energy and lies in the absorption maximum of DNA (245nm). Anyhow, the largest portion of UVC irradiation is blocked by the ozone-containing stratosphere and would be absorbed by the corneal barrier of the skin (*stratum corneum*) if it would reach the earth's surface. UVB (280-315nm) and UVA (315-400nm) irradiation are able to penetrate the atmosphere and cause damage to the DNA (Rastogi *et al.*, 2010). UVA irradiation is commonly known to be the factor that influences skin aging as it reaches deeper layers of the skin, the dermis, where elastic and collagen fibers are located, whereas UVB can only penetrate into the deepest layers of the epidermis (*stratum basale*) (see Figure 1).

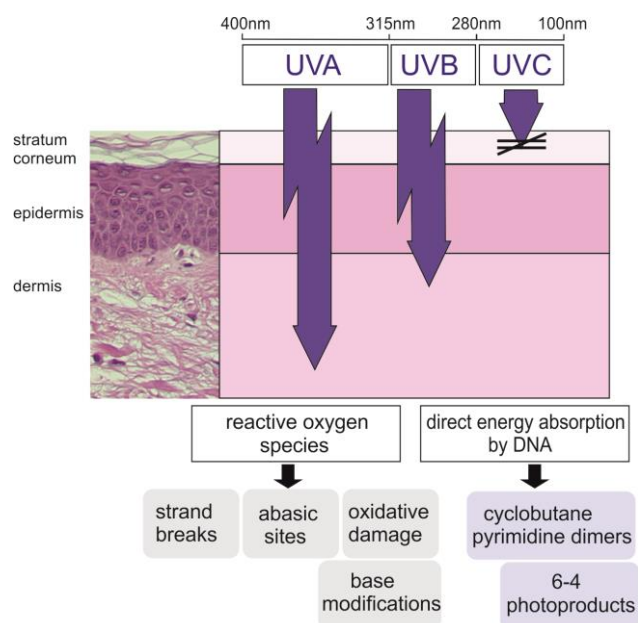


Figure 1: Penetration depth and effects of UV irradiation

The left panel shows a histological cut of human skin (hematoxylin eosin staining, x40), while the right panel shows a schematic view of the penetration depth of UV and visible light through the skin. UVC irradiation (100-280nm) is blocked by the stratosphere and does not reach the earth's surface. UVB irradiation (280-315nm) penetrates into the *stratum basale* and can directly be absorbed by the DNA leading to the generation of bulky DNA lesions (6-4PPs and CPDs). UVA irradiation (315-400nm) reaches the dermis mainly causing the formation of reactive oxygen species, which then result in strand breaks, abasic sites, oxidative damage or base modifications. Illustration by Dr. rer. nat. Christina Seebode.

1.1.2 UV induced DNA lesions

While UVA only causes indirect DNA damage through the creation of free radicals, UVB and UVC directly lead to DNA lesions by crosslinking of adjacent pyrimidine bases and the formation of bulky adducts. This results in two lesions of different types, the cyclobutane pyrimidine dimers (CPDs) or 6,4- pyrimidine pyrimidone photoproducts (6-4PPs) (Lippke *et al.*, 1981; Mitchell & Nairn, 1989) (see Figure 2). CPDs make up 75% of the lesions and contain a four membered ring arising from the coupling of the C=C double bonds of pyrimidines, while 6-4PPs (25%) lead to a crosslink of C₆ of one pyrimidine and C₄ of the other pyrimidine resulting in a stronger distortion of the DNA backbone (Vink & Roza, 2001; Yokoyama *et al.*, 2012). 6-4PPs are more rapidly repaired in the cell (within six hours) than CPDs that still persist after 12 hours (50%) (Kobayashi *et al.*, 2001). Therefore, CPDs are the major source of lasting UV-induced mutations (You *et al.*, 2001).

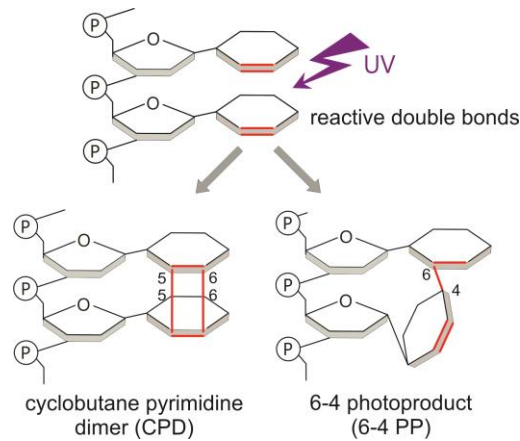


Figure 2: UV-induced DNA lesions

UVB irradiation can directly be absorbed by the DNA and leads to the formation of DNA photoproducts. Adjacent pyrimidine bases (thymidines (T) and cytosines (C)) are crosslinked and form CPDs and 6-4PPs with a ratio of 3:1 (from (Seebode *et al.*, 2016)).

In the course of replication, one possibility for cells to cope with these DNA damages, are translesion polymerases, if the damage has not been repaired before the cell enters the S-phase of the cell cycle. A typical UV signature mutation is the C to T or CC to TT transition that has been found in *TP53* genes of squamous cell carcinomas (SCCs) (Brash *et al.*, 1991). These signature mutations follow the A-rule, meaning that the translesion polymerase η , which lacks a proof-reading function, complements the crosslinked pyrimidine bases with two adenines on the opposite strand (reviewed in Matsumura & Ananthaswamy, 2002). That is why evolutionary other pathways developed to repair DNA damage (see 1.2).

1.1.3 The multistep carcinogenesis model

As described by Hanahan and Weinberg, the accumulation of mutations is one of the hallmarks of cancer (reviewed in Hanahan & Weinberg, 2000). In the multistep carcinogenesis process, a key step is the mutational activation of oncogenes or inhibition of tumor suppressor genes, resulting in the loss of cell cycle control and apoptosis while leading to uncontrolled cell proliferation (reviewed in Soehnge *et al.*, 1997).

Different molecular mechanisms are involved in skin cancer development. Among other effects, UVB irradiation can promote cell proliferation through the MAPK (mitogen-activated protein kinase) pathway, via activating mutations of the *RAS* oncogene (reviewed in Melnikova & Ananthaswamy, 2005). Furthermore, activation of the sonic hedgehog (SHH) pathway components *PTCH1* and *SMO* is a key feature of basal cell carcinoma (BCCs) driving cell proliferation and tumor growth (reviewed in Athar *et al.*, 2014; Emmert *et al.*, 2014). UV signature mutations occurring in early stages of oncogenic transformations in the

skin can also be found in the *TP53* gene in SCCs and BCCs, resulting in a defective p53 protein, which is known to be the “guardian of the genome” (Ehrhart *et al.*, 2003; Inciarte *et al.*, 1976). The p53 protein is involved in multiple cellular processes and fulfills various functions amongst which are the activation of cell cycle checkpoints, DNA repair, and initiation of apoptosis (Fridman & Lowe, 2003; Greenblatt *et al.*, 2003; Oren, 1999). The *CDKN2A* gene encodes for two strong tumor suppressors involved in cell cycle regulation. p16INK4A prevents phosphorylation of the retinoblastoma (RB) protein by CDK4/CDK6 and therefore progression from G₁- to S-phase, while p14ARF stabilizes the cellular p53 level, thereby preventing oncogenic cell transformations (Emmert *et al.*, 2014; Saridaki *et al.*, 2003; Sharpless & Chin, 2003).

Xeroderma pigmentosum (XP) patients, who have a high predisposition for skin cancer development, present a unique model disease to analyze the effects of unrepaired DNA lesions in skin carcinogenesis and accelerated skin cancer development (reviewed in Daya-Grosjean, 2008) (see 1.3.1).

1.2 DNA repair

As previously mentioned, UV light was of great importance during the earliest steps of evolution as the ability of DNA to absorb UV light led to the development of long chain RNA molecules generating more complex organisms (Mulkidjanian *et al.*, 2012). Therefore, UV induced mutations can be seen as an evolutionary engine in the development and progression of life. However, forced by evolutionary pressure, at some point all living organisms developed a complex network of DNA repair pathways to alleviate harmful effects of DNA damage and maintain genome integrity, the DNA damage response (DDR) (reviewed in Giglia-Mari *et al.*, 2011).

1.2.1 DNA damage response (DDR)

The DDR is a complex network of cellular pathways that sense, signal and repair DNA lesions through activating cell cycle checkpoints and DNA repair pathways in response to DNA damage by surveillance proteins that monitor DNA integrity. The DDR consists of a phosphorylation cascade that is initiated by stalled replication forks or polymerases (reviewed in Giglia-Mari *et al.*, 2011).

The protein kinases ATM (PI3K-like kinases ataxia-telangiectasia mutated) and ATR (ataxia-telangiectasia and Rad3-related), which are recruited and activated by double strand breaks

(DSBs) and replication protein A (RPA)-coated single stranded DNA (ssDNA), are key components of the DDR-signaling cascade in mammalian cells (reviewed in Bartek & Lukas, 2007; Cimprich & Cortez, 2008; Shiloh, 2003). Subsequently, ATM and ATR phosphorylate the target protein kinases CHK1 and CHK2 and together reduce cyclin-dependent kinase (CDK) activity by various mechanisms, e.g. activation of p53 and p21 resulting in G₁/S and G₂/M cell cycle arrest (reviewed in Kastan & Bartek, 2004; Riley *et al.*, 2008). Thereby, the time for DNA repair before replication and mitosis is prolonged. Furthermore, ATM and ATR also increase DNA repair protein activity by transcriptional or post-transcriptional modifications like phosphorylation, acetylation, ubiquitylation, or sumoylation (reviewed in Huen & Chen, 2008). If the damage cannot be removed, continuous activation of the DDR response triggers cell death via apoptosis or cellular senescence to prevent tumor formation (Campisi & d'Adda di Fagagna, 2007; Halazonetis *et al.*, 2008).

The nucleotide excision repair (NER), a special case of the DDR response, is a particularly important mechanism in the removal of mutations resulting from UV-induced DNA damage (see 1.2.2).

1.2.2 DNA repair mechanisms

Cells have developed a number of repair mechanisms to ensure genome integrity and cope with different sorts of DNA damage. This thesis mainly focusses on the repair of UV-induced DNA lesions, interstrand crosslinks, and DNA DSBs (reviewed in Lombard *et al.*, 2005).

1.2.2.1 Nucleotide excision repair (NER)

In bacteria pyrimidine dimers formed by UV irradiation are reversed using a special light-dependent process, called photoreactivation. A photolyase binds to the pyrimidine dimer and catalyzes a second photochemical reaction using visible light breaking the cyclobutane ring and reforming the two adjacent thymidylates (reviewed in Thoma, 1999).

Prokaryotic NER only involves three proteins, UvrA, UvrB and UvrC, which carry out the complete process of damage recognition and excision. At the beginning, there is an energy-independent distortion recognition factor (UvrA), followed by energy-dependent recognition of DNA damage using a DNA helicase (UvrB). This helicase creates an open preincision complex and subsequently an oligonucleotide is released by dual incision (Kisker *et al.*, 2013; Sancar & Rupp, 1983). The new DNA strand is synthesized and sealed by DNA polymerase II. A direct reversal process performed by photolyases is present in prokaryotic organisms as well (reviewed in Zhang *et al.*, 2013a). Photolyases can be found in prokaryotes, eukaryotes

and archaea, but are not present in higher mammals like humans. Hence, other highly conserved DNA repair processes developed among both prokaryotes and eukaryotes.

The eukaryotic NER is one of the most versatile repair systems as it can recognize various different types of lesions, e.g. UV-induced lesions or intrastrand crosslinks and involves numerous factors (>30) (reviewed in Truglio *et al.*, 2006). The DNA damage is sensed through a distortion present in the DNA structure (bulky lesions). Thus, different lesions can all be repaired by a common set of enzymes. It is increasingly evident that the overall strategy for NER in eukaryotes has many similarities to the process initiated by the UvrABC nuclease in prokaryotes.

On the molecular level, the repair cascade is made up of several subsequent steps: lesion sensing, opening of a denaturation bubble, incision of the damaged strand, displacement of the lesion-containing oligonucleotide, gap filling (re-synthesis according to the complementary strand) and closing ligation of a nick (reviewed in Nospikel, 2009; Scharer, 2013) (see Figure 3). The damage recognition step is divided into two subpathways: transcription coupled repair (TCR) and global genome repair (GGR). TCR only works on actively transcribed genes (Mellon *et al.*, 1987), while the GGR removes DNA lesions throughout the whole genome (Bohr *et al.*, 1985).

XPC preferably binds to DNA backbone distorting structures and its binding activity can further be stimulated by HR23B and Centrin2 (Krasikova *et al.*, 2012). 6-4 PPs, lead to a strong distortion of the DNA backbone and can directly be recognized by the XPC-HR23B-Cen2 complex (Araki *et al.*, 2001). Additionally, HR23B protects XPC from proteolytic degradation. The formation of an XPC-HR23B complex with DNA is enhanced by the single strand DNA (ssDNA) binding protein RPA (Krasikova *et al.*, 2008). As CPDs only slightly distort the DNA backbone, they are poorly recognized by XPC, although their removal depends on XPC recognition (Sugasawa *et al.*, 1998). To further increase XPC's affinity to the damaged DNA it can be polyubiquitinated and depends on the UV-DDB-ubiquitin ligase complex consisting of DDB1 and DDB2 (UV-damaged-DNA-binding protein 2 = XPE) (Sugasawa *et al.*, 2005). In addition, damage recognition is facilitated by a stronger distortion of the DNA backbone due to binding of UV-DDB (Fujiwara *et al.*, 1999). A stalled RNA polymerase II itself is the damage recognition factor in TCR independent of XPC. Together with the Cockayne syndrome (CS) proteins, CSA and CSB, it initiates the repair (Mu & Sancar, 1997).

During GGR as well as TCR, the basal transcription factor TFIIH is recruited via direct interaction with XPC-HR23B or by XPA, respectively (Park *et al.*, 1995; Riedl *et al.*, 2003). TFIIH consists of the CAK (CDK7; Cyclin H and MAT1) and core complex (XPD, XPB, p62, p52, p44, p34 and p8/TTDA) (see Figure 3). XPG is thought to be the eleventh subunit of TFIIH, while XPD connects the two subcomplexes (Chen *et al.*, 2003) (reviewed in Compe & Egly, 2012; Egly & Coin, 2011). TFIIH performs the opening of a denaturation bubble (24-30 nucleotides (nts)) around the lesion. In NER the ATPase activity of XPB and the helicase activity of XPD are needed for the opening of the denaturation bubble (Coin *et al.*, 2007). TFIIH is supported by XPA and RPA, two factors that have a high affinity for ssDNA, in displacing the XPC complex (Overmeer *et al.*, 2011). Furthermore, RPA also assists TFIIH to open the DNA helix around the damage and protects the undamaged strand opposite the lesion (de Laat *et al.*, 1998b; Lee *et al.*, 2003). Together with XPA it is important for verifying the lesion and damage demarcation (reviewed in Fadda, 2016).

After anchoring TFIIH to the site of DNA damage by XPA and RPA, the two endonucleases XPF/ERCC1 (5') and XPG (3') are recruited. TFIIH is responsible for XPG recruitment via its pleckstrin homology/phosphotyrosine-binding (PH/PTB) domain of subunit p62 (Dunand-Sauthier *et al.*, 2005; Gervais *et al.*, 2004), while XPF/ERCC1 is recruited to the damage site through a direct interaction between the central domain of ERCC1 and XPA (Orelli *et al.*, 2010). XPF/ERCC1 is a heterodimeric endonuclease complex that cleaves upstream of the lesion, whereas XPG cleaves downstream (Mu *et al.*, 1996; O'Donovan *et al.*, 1994). Strand incision and repair synthesis of NER are highly coordinated with several subsequent steps. 5' incision by XPF/ERCC1 is necessary and sufficient for the initiation of repair synthesis as it generates a free 3' OH group and a branched flap structure with a free 5' end. On the other hand, 3' incision by XPG is not necessarily needed to initiate polymerization by DNA polymerase δ or ϵ (Staresincic *et al.*, 2009). Actually, efficient 3' incision requires the presence and catalytic activity of XPF/ERCC1 (Tapias *et al.*, 2004). Hence, 5' incision has to precede the 3' incision. Regardless, 3' incision is needed for completion of repair synthesis.

Depending on the lesion, the excised fragment is about 24-32nts in length (Evans *et al.*, 1997b). After displacement of the lesion-spanning oligonucleotide the gap is accurately filled by either of the replicative polymerases δ , ϵ , or κ according to the undamaged complementary strand (Ogi *et al.*, 2010). The nick is then closed by ligase I and III together with XRCC1 (Moser *et al.*, 2007).

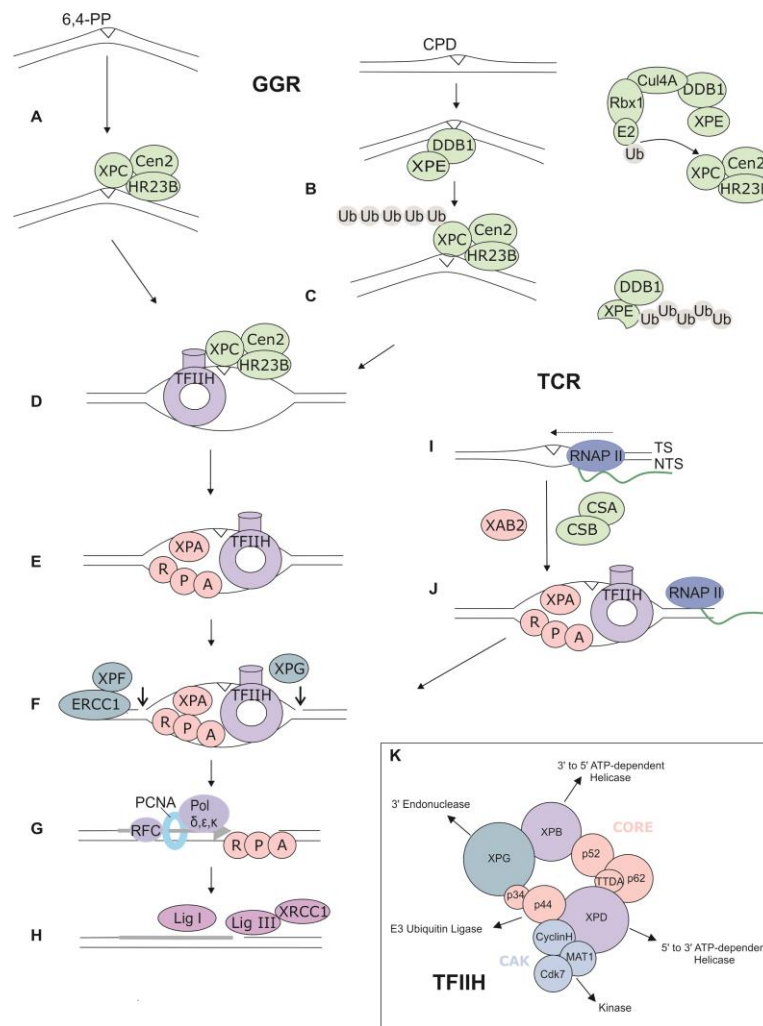


Figure 3: Schematic overview of the NER pathway

(A) Global genome repair: 6-4PPs lead to a strong DNA distortion and are directly recognized by the XPC-HR23B-Centrin2 complex (green), while CPDs can only be identified by XPC together with the DDB complex (DDB1, XPE) (B). (C) Lesion affinity of XPC and DDB is increased by ubiquitylation. (I) and (J). In TCR, repair is initiated by a stalling of the RNA polymerase II at the site of a lesion on the transcribed strand as well as interaction with CSA, CSB, and XAB2. (D) The basal transcription factor TFIIH, harbouring ATPase XPB and helicase XPD, is responsible for unwinding the dsDNA around the lesion. In case of TCR, RNAPII is replaced by TFIIH, XPA, and RPA. (E) XPC is replaced by XPA and RPA, while TFIIH is recruited. (F) The two endonucleases XPF/ERCC1 (5') and XPG (3') are recruited and incise the damaged DNA strand. (G) After recruitment of PCNA the gap is filled by DNA polymerase δ , ϵ , or κ . (H) Ligase I or III seal the nick between the newly synthesized and free 5' phosphate at the XPG restriction site. (K) The ten/eleven-subunit complex TFIIH is composed of a core associated to the CAK through the XPD subunit (purple). The core (rose) is made up of XPB, p62, p52, p44, p34, and p8/TTDA, while the CAK (blue) contains Cdk7, Cyclin H, and MAT1. XPG is proposed to present the eleventh subunit and is a component of TFIIH. Illustration by Dr. rer. nat. Christina Seebode, adapted from (Egly & Coin, 2011; Nouspikel, 2009).

1.2.2.2 Interstrand Crosslink Repair

DNA interstrand crosslinks (ICLs) can be caused by various endogenous metabolites, environmental exposures, and cancer chemotherapeutic agents with at least two reactive chemical groups (reviewed in Clauson *et al.*, 2013). Thereby, two nts of opposite DNA

strands can be covalently bound preventing strand separation during replication or transcription by blocking the replication fork or RNA polymerase. This results in highly cytotoxic ICLs and until now repair pathways are not completely defined. In principle, damage recognition can either be replication-bound (blocked replication fork) or non-replication bound (blocked RNA polymerase). The basic mechanism is shown in a schematic overview below (see Figure 4) (Bhagwat *et al.*, 2009; Kottemann & Smogorzewska, 2013; Moldovan & D'Andrea, 2009; Niedernhofer *et al.*, 2004; Sengerova *et al.*, 2011). After damage recognition, the endonucleases XPF/ERCC1 and Mus81/Eme1 are thought to cut the lesion-containing (parental) strand, thereby unhooking one daughter duplex from the damage, forming a DNA DSB. Subsequently, translesion synthesis by error-prone damage-tolerating polymerases takes place. Afterwards, the NER pathway, especially the excision complex consisting of XPF/ERCC1 and XPG repairs the overhanging strand. During S-phase the sister chromatid is synthesized via homologous recombination (HR). Especially the exact mechanisms of non-replication-bound repair, which is mainly coordinated by NER proteins, needs to be further investigated and a lot of open questions, still remain. For example there is an ongoing discussion whether XPF/ERCC1 is the only required endonuclease for ICL repair (Fisher *et al.*, 2008; Kuraoka *et al.*, 2000) or whether other proteins are needed for processing of intermediate states of ICL repair, e.g. other endo- or exonucleases (Clauson *et al.*, 2013; Giannattasio *et al.*, 2010; Zhang & Walter, 2014).

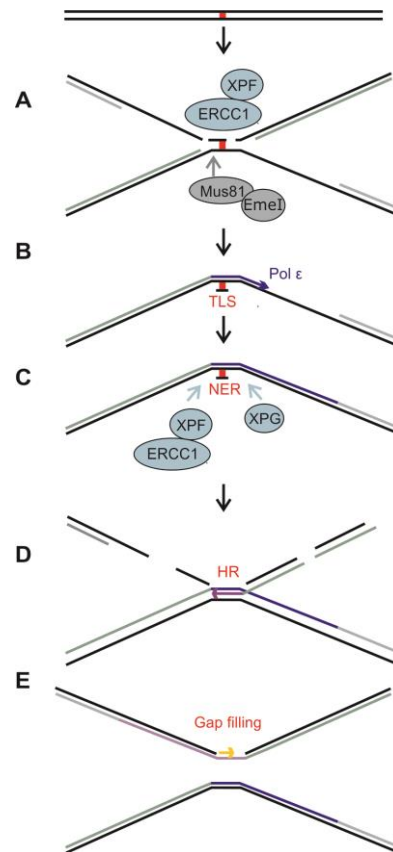


Figure 4: Schematic overview of interstrand crosslink repair (ICL)

The interstrand crosslink (ICL) of two DNA strands is repaired by a multistep process, coordinated by the Fanconi anemia complex (not shown) by a yet unclear mechanism. An ICL can either be detected in a replication bound (stalled replication fork) or replication-independent context (stalled RNA polymerase). **(A)** The endonucleases XPF/ERCC1 and Mus81/Eme1 are thought to cut the lesion-containing strand, thereby unhooking the damage. **(B)** A translesion polymerase is recruited and fills the resulting gap around the ICL (TLS = translesion synthesis). **(C)** The overhanging ICL is repaired by the NER excision complex consisting of the XPF/ERCC1 and XPG endonucleases. **(D)** During S-phase the sister chromatid is synthesized via homologous recombination (HR). As the endonuclease XPF/ERCC1 is involved in two steps of ICL repair it is of special importance **(A and C)**. Illustration by Dr. rer. nat. Christina Seebode, adapted from (Moldovan & D'Andrea, 2009).

1.2.2.3 DNA double strand repair pathways

This thesis focusses on the endonucleases XPF/ERCC1 and XPG, which have originally been detected in NER. Mutations in ERCC1 or XPF can cause very severe phenotypes as well in humans as in mice, which cannot solely be explained by defects in NER or ICL repair often showing accelerated aging. Orthologs of XPF/ERCC1 like DmERCC1-MEI9 in *Drosophila melanogaster* or Rad1-Rad10 in *Saccharomyces cerevisiae* have been implicated in DSB repair (Baker *et al.*, 1978; Fishman-Lobell & Haber, 1992; Ivanov & Haber, 1995) due to their ability to remove non-homologous 3' single-stranded flaps (Al-Minawi *et al.*, 2008; Niedernhofer *et al.*, 2001; Sargent *et al.*, 1997). Furthermore, different studies reported about the moderate sensitivity of XPF/ERCC1 deficient mammalian cells to DSB inducing agents

and ionizing radiation, suggesting an important function of XPF/ERCC1 in one or more sub-pathways of DSB repair (Ahmad *et al.*, 2008; Mogi & Oh, 2006; Murray *et al.*, 1996; Wood *et al.*, 1983).

There are two major pathways of DSB repair in eukaryotes: the correct HR mediated repair (HRR) and the error-prone nonhomologous end joining (NHEJ) (Brugmans *et al.*, 2007). As HRR recovers lost sequence information from a sister chromatid it can only take place during S or G₂ phases of the cell cycle. During NHEJ, two broken ends are connected by ligation, not restricting it to proliferating cells (reviewed in Karran, 2000). In HRR, the XPF/ERCC1 heterodimeric complex is thought to function through the error-prone single-strand annealing (SSA) sub-pathway, gene conversion, and homologous gene targeting in yeast and mammals (Adair *et al.*, 2000; Fishman-Lobell & Haber, 1992; Ivanov & Haber, 1995; Niedernhofer *et al.*, 2001; Sargent *et al.*, 2000), while for NHEJ the complex is only involved in the Rad52- and Ku70/Ku86-independent microhomology-mediated end-joining (MMEJ) sub-pathway (Ahmad *et al.*, 2008; Bennardo *et al.*, 2008; Ma *et al.*, 2003; McVey & Lee, 2008; Yan *et al.*, 2007). Unfortunately, these studies primarily focus on mice, yeast, and hamster cells, and until now human XPF/ERCC1 has only been assumed to be involved.

1.3 DNA repair deficiency disorders

Mutations in components of the different DNA repair pathways can result in various clinical entities and diseases. A rare genetic disorder arising from defects in one of the NER components is XP (reviewed in Kraemer *et al.*, 1987). There is no prevalence for ethnic groups and it appears all over the world.

1.3.1 Xeroderma pigmentosum

In 1874, Hebra and Kaposi were the first to describe XP as patients with “parchment skin” (Hebra & Kaposi, 1874). XP is a rare autosomal recessive disorder that affects the repair of DNA damage caused by UV light (see Figure 3). Worldwide the prevalence of XP is very low and varies in different regions (North America/Northern Europe 1:1.000.000, North Africa/Middle East 1:50.000, Japan 1:22.000) (Ben Rekaya *et al.*, 2009; Hiraï *et al.*, 2006; Kleijer *et al.*, 2008; Kraemer & Slor, 1985; Messaoud *et al.*, 2010; Soufir *et al.*, 2010). The DNA repair mechanism impaired in this disease is the NER, affecting the genes *XPA to XPG* or the translesion polymerase η (reviewed in Lehmann *et al.*, 2011).

XP patients exhibit high sun sensitivity and freckling from birth on leading to a strongly increased risk for skin cancers (reviewed in Kraemer *et al.*, 1987) (see Figure 5). Development of non-melanoma skin cancer (NMSC) is 10.000-fold increased and occurs at a median age of nine years compared to 67 years in the general population (Bradford *et al.*, 2011). Tumors are preferentially located at sun-exposed areas (face, head, neck, or the back of the hands). The median age for melanoma is 22 years in contrast to 55 in the general population and the risk is 2.000-fold increased. Melanoma preferentially occur on the extremities (Kraemer *et al.*, 1994). XP is an interesting model disease for fast-forward skin carcinogenesis as well as accelerated aging. In the healthy population, skin cancer is rare before the age of 20. Typically, melanomas develop between the age of 30-50, while non-melanoma skin cancer is most prominent in elderly people (reviewed in de Vries & Coebergh, 2004; Madan *et al.*, 2010). Therefore, studies of XP patients provide a molecular foundation to demonstrate the UV induced origin of mutations of cancer suppressing genes in NMSCs as well as melanoma (Couve-Privat *et al.*, 2004; Giglia *et al.*, 1998; Wang *et al.*, 2009a).

Interestingly, XP cells show normal killing after exposure to X-rays, while they are hypersensitive to UV irradiation. X irradiation is also used as a therapy for tumors in XP patients (DiGiovanna *et al.*, 1998; Giannelli *et al.*, 1981), suggesting that the repair of x-ray induced damage is independent from NER.

Kraemer *et al.* could show that experiments fusing different strains of XP patient fibroblasts lead to a higher DNA repair rate than of either strain's unfused cells, indicating complementary corrections in the fused cells. Thereby, they identified five complementation groups, meaning that there are at least five mutation causing genes decreasing DNA repair among these fibroblast strains (Kraemer *et al.*, 1975a; Kraemer *et al.*, 1975b). To date, seven XP complementation groups, XP-A to XP-G as well as a variant form (defect in translesion polymerase η) were identified.

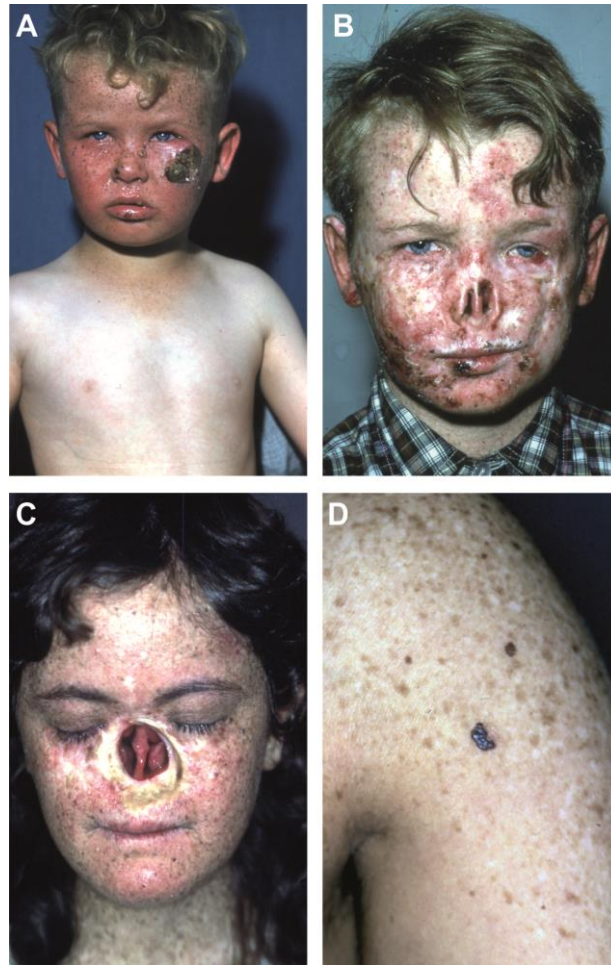


Figure 5: Picture of two XP patients with typical clinical features

(A) This patient shows the typical freckling in sunexposed areas together with a severe sun burn of the face and several non-melanoma skin cancers of the face (nose, lips, cheek). (B) This picture shows the same patient a few years later. Due to tumor resection he has lost his nose, shows scarring of the face, and new precancerous lesions. (C) This female patient has also lost her nose due to an invasive tumor, (D) but additionally also shows melanoma skin cancer in sunexposed areas, like the arms.

Among the different complementation groups there is a large difference in frequency with XP-C and XP-A being the most prominent ones, and XP-B being quite rare. Furthermore, they differ in the severity of the symptoms like the number of skin cancers or neurological involvement (see Table 1). The general belief has been that heterozygous carriers of XP show no clinical symptoms and have a normal DNA repair capability. Anyhow, XP heterozygous mice have a higher susceptibility for skin and internal organ cancers (Cheo *et al.*, 2000), which also suggests an effect in humans. A clinical study (NCT00046189) addressing this issue is currently going on at the National Institute of Health (NIH).

Table 1: XP complementation groups

The table shows the frequency of the different XP complementation groups and gives information on skin cancer number, neurological involvement, repair capability, the defective gene and chromosomal location (from (Lehmann *et al.*, 2014a)).

Complementation Group	Frequency (%)	Skin cancer	Neurological involvement	Cellular repair capability	Defective Gene	Chromosome
XPA	30	++	+++	<10%	XPA	9q22.3
XPB	0.5	+	+	3~7%	XPB/ERCC3	2q21
XPC	27	++	+	10~20%	XPC	3p25
XPB	15	++	+++	25~50%	XPB/ERCC2	19q13.2-q13.3
XPE	1	+	-	40~50%	XPE-DDB2	11p12-p11
XPF	2	+	-	10~20%	XPF/ERCCC4	16p13.3
XPG	1	+	++	<5%; 25%	XPG/ERCC5	13q33
Variant	23.5	+	-	100%	Pol H	6p21.1-p12

1.3.2 Other associated disorders

Mutations in the *CSB/ERCC6* and *CSA/ERCC8* genes encoding proteins involved in the TCR DNA repair pathway can cause another rare autosomal recessive congenital disorder, Cockayne syndrome (CS) (reviewed in Cleaver *et al.*, 2009). The disease was first described in 1936 by Edward Cockayne (Cockayne, 1936) and has an incidence of less than one case per 250.000 live births in Northern Europe. In 65% of the cases the mutations lie in *CSB*, while *CSA* is affected in 35% of the cases (reviewed in Laugel, 1993). Clinical hallmarks of the disease are microcephaly and growth failure, photosensitivity, hearing loss, cataracts, retinal dystrophy, developmental delay, and premature aging (Hoeijmakers, 2009).

Cerebrooculofacioskeletal (COFS) syndrome represents the prenatal extreme form of CS and is clinically characterized by microcephaly, cataract and/or microphthalmia, arthrogryposis, severe psychomotor developmental delay, height-weight growth delay and facial dysmorphism (prominent metopic suture, micrognathism) (Pena & Shokeir, 1974). Additionally, cutaneous photosensitivity, peripheral neuropathy, sensorineural hearing loss and pigmentary retinopathy can be observed. Mutations can mainly be found in the *CSB/ERCC6* gene; however, one case has been linked to the *ERCC1* gene (Faridounnia *et al.*, 2015; Jaakkola *et al.*, 2010). To date, fewer than 20 cases have been confirmed.

Another extremely rare disease is the XPF-ERCC1 (XFE) progeroid syndrome (Niedernhofer *et al.*, 2006). The patient presented with dwarfism, cachexia, and microcephaly, sun-

sensitivity, learning disabilities, hearing loss, and visual impairment. The progeroid features of accelerated aging led to an early death at the age of 16.

1.3.3 Genotype-phenotype-correlation

The genotype-phenotype correlation of DNA repair deficiency disorder patients is highly complex. Patients can show combined symptoms of different diseases, e.g. XP and CS. These patients exhibit XP symptoms like photosensitivity and increased risk for cutaneous malignancies in combination with CS symptoms like neurologic abnormalities (Emmert *et al.*, 2006; Kraemer *et al.*, 2007; Lehmann *et al.*, 2014a), and belong to XP complementation groups B, D, F, or G. All these genes encode for proteins important for interactions with TFIIH (see Figure 3). Currently, it is believed that a transcription defect may be the reason for neurologic impairment (reviewed in van Gool *et al.*, 1997). Furthermore, oxidative processes involving mitochondria caused by the generation of free radicals could be the basis of neurodegeneration in XP patients (reviewed in Bohr *et al.*, 2002; Jeppesen *et al.*, 2011).

Figure 6 shows the complex genotype-phenotype interactions of the mutations in the different XP genes (A-G plus the variant form), *CSA* and *CSB* as well as *TTD-A* (part of the TFIIH complex) and *TTDNI* (unknown function). COFS syndrome, XFE, Fanconi anemia (FA), and trichothiodystrophy (TTD) can also derive from defects in NER proteins. For example, XP-G patients can be subdivided into three categories showing either typical XP symptoms, XP with late onset CS symptoms, or XP with severe CS symptoms. Patients who express at least one full-length allele of *XPG* can be found in the XP only group. On the other hand, XP-G patients with late-onset CS symptoms show several different types of mutations; however, all of these patients either express a mutant full-length protein or splice variants (in the following also referred to as isoforms) with residual functions in NER or the ability to interact with TFIIH (reviewed in Scharer, 2008).

Likewise, different mutations in the *XPF* gene can result in distinct clinical outcomes: either cancer, as in XP, or progeroid symptoms, as in XFE syndrome. This may be explained by effects of the mutation on NER as well as ICL repair, thereby primarily resulting in cell death and senescence in response to DNA damage, leaving fewer possibilities for accelerated carcinogenesis but enhanced aging. Classic XP-F patients that mainly show mild symptoms only suffer from a defect in NER causing less cell death, but allowing mutation accumulation and consequently cancer (Niedernhofer *et al.*, 2006).

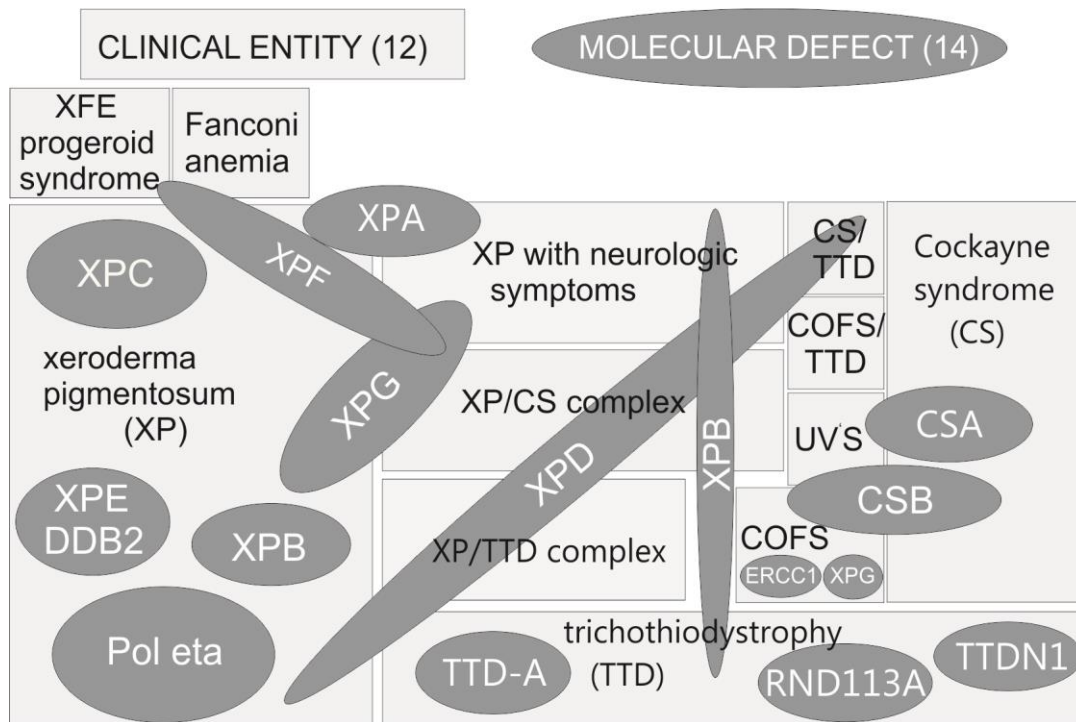


Figure 6: 12 clinical entities and 14 molecular defects show the complex genotype-phenotype correlation of DNA repair deficiency disorders

Different mutations in various genetic regions of DNA repair proteins can lead to 12 different clinical entities. In turn, mutations in different genes can cause the same clinical entity. TTD = trichothiodystrophy, COFS = cerebro-oculo-facial-skeletal syndrome, UV'S = UV sensitive syndrome. Illustration by Dr. rer. nat. Christina Seebode, modified from (Schubert & Emmert, 2016).

1.4 Multiple functions of the XPG and XPF/ERCC1 endonucleases

The specialty of the two endonucleases XPG and XPF/ERCC1 is not only founded in their coordinated interplay during the dual incision step of NER, but also in the fact that they have many other functions besides excision during NER. For example, XPG as well as XPF are involved in basal transcription via interaction with TFIIF (Ito *et al.*, 2007; Le May *et al.*, 2012), and also in ICL repair (reviewed in Clauson *et al.*, 2013). As shown in XPF/ERCC1 - deficient patient cells, neither the components of the FA signaling pathway, nor components of homologous recombination (HR) can be recruited (Bhagwat *et al.*, 2009). Furthermore, genome-wide association studies of single nucleotide polymorphisms (SNPs) and protein expression analysis implicate the *XPF* and *XPG* genes as potential marker for skin cancer risk as well as for disease outcome (Li *et al.*, 2013a). This thesis focuses on spontaneous mRNA splice variants of the XPG and XPF/ERCC1 endonucleases.

1.4.1 XPG

The human gene encoding for xeroderma pigmentosum group G was identified as *ERCC5* (Cleaver *et al.*, 1999; MacInnes & Mudgett, 1990). The *XPG* gene is located on chromosome *13q32.3-q33.1* and encodes for a 136kDa protein, organized into 15 exons and 14 introns. So far, several spontaneous mRNA splice variants, due to alternative splicing, have been identified in the healthy population, whose amounts vary in different tissues (Emmert *et al.*, 2001).

The XPG endonuclease is involved in multiple processes from DNA repair to basal transcription due to its interplay with TFIIF. XPG's function during NER is highly conserved among different species. The yeast homolog of human XPG, Rad2, displays an evolutionary conserved endonuclease activity (Habraken *et al.*, 1993).

XPG belongs to the FEN 1 (flap endonuclease 1) protein family of structure-specific endonucleases and harbors two highly conserved nuclease domains (Scherly *et al.*, 1993). Members of the FEN 1 family cleave a variety of substrates containing ss/dsDNA junctions including 5' single-stranded overhangs and bubble structures. The XPG protein sequence shows two conserved N and I nuclease regions that XPG shares with other nucleases, e.g. Rad2. Furthermore, XPG has unique domains that are responsible for its specific functions during NER. The spacer region (about 600 amino acids (aa)) includes a conserved, highly acidic patch and separates the N and I region (see Figure 7). In the ternary structure, after protein folding, these two regions come close together and form the active center for XPG's endonuclease function. The spacer is also partly involved in the interaction with other proteins, e.g. TFIIF and RPA (Dunand-Sauthier *et al.*, 2005; Thorel *et al.*, 2004). The C-terminal region, beyond the I region is engaged in protein-protein interactions, e.g. with TFIIF, while the PIP-box motif mediates interaction with PCNA (Gary *et al.*, 1997). Two distinct interaction patches between XPG and TFIIF suggest a strong functional interaction. Furthermore, the C-terminus contains two strong nuclear localization signals (NLSs) (Knauf *et al.*, 1996).

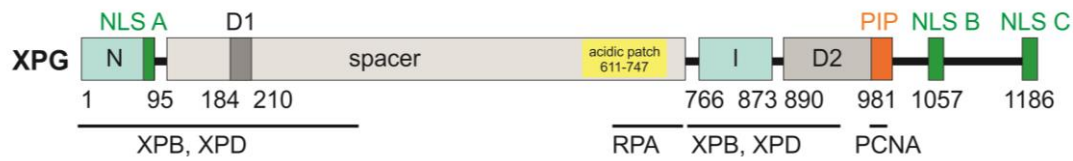


Figure 7: Schematic overview of XPG and its domains

The N and I domains (light blue-green) form the catalytic center of the XPG endonuclease. They are separated by a spacer region in the primary sequence, but come close together in the ternary structure after protein folding. The PIP-box motif (PCNA binding motif) (orange) and nuclear localization signals (NLS-A N-terminal, NLS-B and NLS-C) (green) are located at the C-terminal region. Interaction regions with TFIIH (XPB, XPD), RPA and PCNA are indicated by black bars. The D1 and D2 boxes are conserved in higher eukaryotes (grey). Modified from (Scharer, 2008).

XPG plays a key role during GGR as well as TCR, by displaying a structural endonuclease activity. On the one hand, XPG performs the 3' downstream cut and on the other hand its presence is necessary for the binding of the endonuclease XPF/ERCC1 and its previous 5' cut (Staresincic *et al.*, 2009). Furthermore, it stabilizes TFIIH by forming a complex with TFIIH, XPA and RPA leading to a constitutive recruitment of XPF to NER complexes (Ito *et al.*, 2007; Wakasugi *et al.*, 1997). XPG can only reveal its catalytic function once XPF-ERCC1 has made the 5' incision, as the 3' incision is triggered by a conformational change of the damaged structure (Hohl *et al.*, 2003; Staresincic *et al.*, 2009).

Indicated by a complex genotype-phenotype relationship in patients with defects in XPG, resulting in different clinical entities like XP, CS and XP/CS, XPG might have important roles outside of NER. Besides its role during NER, XPG is also involved in basal transcription via its interaction with TFIIH. The architecture of TFIIH was found to be highly dependent on interaction with XPG (see Figure 3K), therefore impaired interaction due to mutations in XPG could result in dissociation of CAK and core, suggesting an XPG function independent of the nuclease activity (Arab *et al.*, 2010; Ito *et al.*, 2007) (see Figure 3K). Moreover, XPG as well as XPF are present at the promoter and terminator forming loop structures that are demethylated and DNA is relaxed by small cuts produced by these endonucleases (Le May *et al.*, 2012).

Additionally, XPG as well as other NER factors are involved in the repair of oxidative DNA damage like ROS that lead to damages like 8-oxo-guanine, a non-bulky lesion, or bulky lesions like 8,5'-cyclopurine-2'-deoxynucleoside (Slupphaug *et al.*, 2003). These lesions are primarily repaired by the short patch base excision repair (BER) (reviewed in Berquist & Wilson, 2012). Interestingly, NER seems to provide a backup repair mechanism for BER as

XPG is involved in TCR of oxidative lesions via interaction with TFIIH (Klungland *et al.*, 1999) stimulating the initial step of BER (Melis *et al.*, 2013). Additionally, XPG is not only involved in transcription, but has also been reported to interact with the replisome (Gilljam *et al.*, 2012) and to be involved in chromatin organization (Le May *et al.*, 2012).

1.4.2 XPF/ERCC1

The human *XPF (ERCC4)* gene (OMIM: 278760) encodes for a 916 aa protein that is part of the heterodimer XPF/ERCC1 cleaving 5' of UV induced lesions with the chromosomal location *16p13.2-p13.1* (Brookman *et al.*, 1996; Sijbers *et al.*, 1996a). The *ERCC1* gene (OMIM: 126380) encodes for a 297 aa protein and is located on chromosome *19q13.32* (Sijbers *et al.*, 1996b). These two proteins are paralogs probably arisen by gene duplication of the conserved Helix-hairpin-Helix (HhH₂) protein-interaction and nuclease domains (Gaillard & Wood, 2001). Interestingly, also for *XPF* and *ERCC1* several spontaneously occurring physiological mRNA splice variants have been postulated (www.ensembl.org (Friboulet *et al.*, 2013; Gerhard *et al.*, 2004; Matlin *et al.*, 2005)).

In eukaryotes, the XPF/ERCC1 heterodimeric protein complex is evolutionary conserved. Orthologs comprise Rad1/Rad10 in *Saccharomyces cerevisiae*, Xpf/ercc1 in mice, and Mei-9/Ercc1 in *Drosophila* (Baker *et al.*, 1978; Fishman-Lobell & Haber, 1992; reviewed in de Buendia, 1998). These nucleases belong to a family of structure specific nucleases also including MUS81/Eme1 and SLX4-SLX1, which are related through structural similarities (Hanada *et al.*, 2006; reviewed in Heyer, 2004). Even XPF and ERCC1 show sequence similarities with each other in their HhH₂ domains (Gaillard & Wood, 2001). This domain is involved in dimerization as well as DNA binding (de Laat *et al.*, 1998a) (Doherty *et al.*, 1996; Su *et al.*, 2012).

XPF and ERCC1 function together as a heterodimer, where the catalytic activity lies in the XPF subunit, while ERCC1 presents the DNA binding unit. The nuclease specifically cuts junctions between ss/dsDNA, where the single strand departs 5' to 3' from the junction (Bessho *et al.*, 1997; de Laat *et al.*, 1998a; Evans *et al.*, 1997b). The endonuclease complex XPF/ERCC1 has a core function in NER, but also plays a key role in ICL repair (De Silva *et al.*, 2002), in which context it is also involved in HR at DNA replication forks (Al-Minawi *et al.*, 2009). Furthermore, there have been extensive studies on the complex functions of XPF/ERCC1 -besides NER and ICL repair- for example to assess its role in sub-pathways of homology-directed double strand repair via SSA, gene conversion, and homologous gene targeting (Adair *et al.*, 2000; Niedernhofer *et al.*, 2001; Sargent *et al.*, 1997).

For XPF there is an even more complex genotype-phenotype correlation than for XPG. Defects in XPF have been associated with several disorders including XP, XP/CS complex phenotype, CS, FA, and XFE (see Figure 7 and Figure 9) (Ahmad *et al.*, 2010; Bogliolo *et al.*, 2013; Hayakawa *et al.*, 1981; Kashiwama *et al.*, 2013; Niedernhofer *et al.*, 2006; Schubert *et al.*, 2014). This highlights the importance of the XPF protein in different cellular pathways and may extend beyond DNA repair. Defects in *ERCC1* have been associated with XP, CS, and even the severe COFS syndrome (see Figure 9) (Jaspers *et al.*, 2007; Kashiwama *et al.*, 2013).

XPF/ERCC1 is also involved in the removal of reactive oxygen species (ROS), which explains why XPF/ERCC1 deficient patients often display an accelerated aging phenotype through damage accumulation (Fisher *et al.*, 2011b).

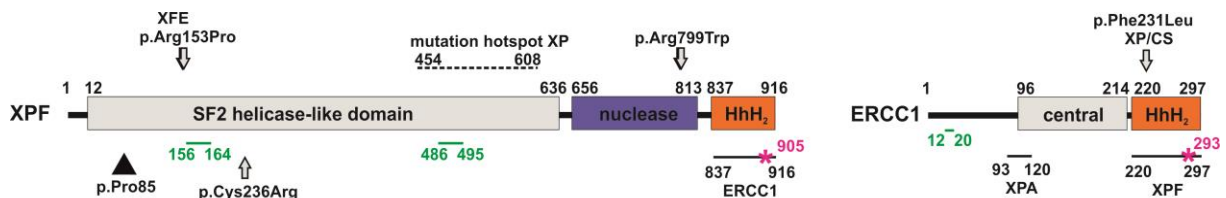


Figure 8: Schematic overview of XPF/ERCC1 and their domains

The colored bars depict functional protein domains (e.g. nuclease domain, Helix-hairpin-Helix (HhH₂) motif). Green lines present putative NLS. The grey arrows indicate mutations leading to different disease entities, as well as a mutation hotspot (dotted line), while black arrows and lines highlight essential residues and protein interaction domains. Pink asterisk mark the essential residues for the interaction between XPF and ERCC1. Illustration by Dr. rer. nat. Christina Seebode, modified from (de Laat *et al.*, 1998c; Fisher *et al.*, 2011a; Li *et al.*, 1995; McNeil & Melton, 2012; van Duin *et al.*, 1986).

Indeed, XPF as well as XPG have been implicated in different other cellular processes. They are involved in transcription, even without genotoxic stress, by gene regulation through chromatin looping and triggering DNA methylation. While XPG induces DNA breaks and demethylation around the promoter region to catalyze the recruitment of the transcriptional repressor CTCF to promote gene looping, XPF only functions in stabilizing the structures (Le May *et al.*, 2012). Moreover, XPF/ERCC1 also seems to be involved in telomere maintenance through the interaction with TRF2 and cleaving of 3' overhangs of uncapped telomeres. This makes the complex a critical factor for genome stability (Munoz *et al.*, 2005; Wan *et al.*, 2013; Wu *et al.*, 2008; Zhu *et al.*, 2003).

1.4.3 XP-G and XP-F patients

In order to study DNA repair, patient cell lines with defective components of repair pathways can be of great help to understand the function of these proteins, the basic mechanism and the influence of protein variants, e.g. splice variants. For the rare complementation group G less than 30 patients have been reported so far (Europe, Japan, North- and South America) (Lehmann *et al.*, 2014b; Schafer *et al.*, 2013; Soltys *et al.*, 2013). The patient XP20BE belongs to the XP-G complementation group and showed clinical features of the XP/CS complex phenotype. The symptoms included sun sensitivity, microcephaly, hyperpigmentation, and mental impairment. The patient died at the age of six years (Moriwaki *et al.*, 1996). He inherited a mutation from his father in exon one of the *XPG* gene that resulted in a premature termination codon. Furthermore, the maternal allele produces an unstable or poorly expressed mRNA. Moreover, the cDNA contained an mRNA species with a large splicing defect that encompassed a deletion from exon one to exon fourteen (Okinaka *et al.*, 1997). Therefore, the largest XP20BE allele leads to synthesis of an extremely shortened protein (only 138 aa) and presents a usable model cell line to study variants of the XPG endonuclease.

So far, a key limitation to study the functions of the XPF/ERCC1 complex has been the lack of an appropriate human XPF- or ERCC1-defective cell line. Up until now, patient cell lines were utilized to study the function of these proteins. However, due to the essential role of these genes, all mutations characterized in patients so far retained at least one full-length allele with residual functional capabilities (see Figure 9) (Schubert *et al.*, 2014). Most XP-F patients show relatively mild symptoms of photosensitivity with occasional neurological abnormalities and there are only a few severely affected patients (Niedernhofer *et al.*, 2006). Therefore, a suitable host cell line with a complete knockout is of great interest to investigate the multiple roles of the XPF/ERCC1 complex in different repair pathways.

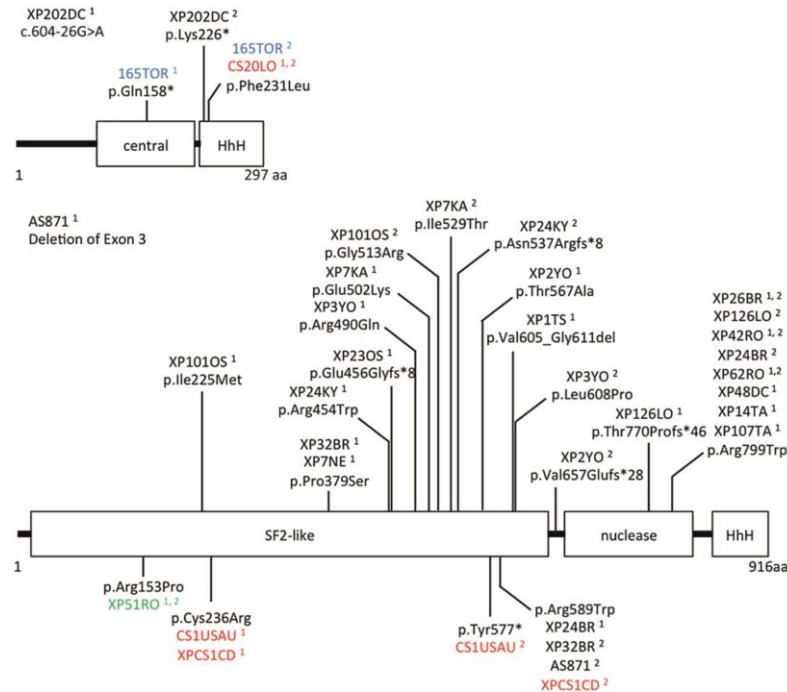


Figure 9: Schematic view of XPF/ERCC1 and mutations associated with DNA repair defective disorders

Mutations in the genes coding for *XPF* or *ERCC1* can lead to different clinical entities. Black font = XP, blue font = COFS, red font = CS, or XP/CS complex phenotype, green font = XFE. From (Kashiyama *et al.*, 2013)

1.5 Genome editing techniques

The demand for targeted, fast, and customizable gene knockout has rapidly grown over the last decade. The molecular basis for gene editing relies on the cellular repair mechanisms of DSBs. These lesions can be repaired by HRR or NHEJ as mentioned above (Takata *et al.*, 1998) (see 1.2.2.3). HRR is an error-free process, but the cell depends on a template strand and therefore it only takes place in S, G₂, or M phase of the cell cycle (reviewed in Szostak *et al.*, 1983). DSBs are re-ligated through the NHEJ pathway in the absence of a repair template, resulting in insertion/deletion mutations leading to frameshift mutations and premature stop codons (Deltcheva *et al.*, 2011). For genome editing it is critical to introduce a directed DSB in the target gene. Recently, notable attention has been paid to genome editing tools like zinc-finger nucleases (ZFNs), meganucleases or bacterial transcription activator-like type III effector nucleases (TALENs). Those techniques have been used for targeted gene knockout (KO) in multiple fields (Wood *et al.*, 2011).

1.5.1 TALEN, Zinc finger and meganucleases

ZFNs or TALENs function through tethering of endonuclease catalytic domains to modular DNA-binding proteins to induce targeted DSBs at specific genomic loci (Wood *et al.*, 2011).

ZFNs are artificial restriction enzymes in which a zinc-finger DNA-binding domain is fused to a DNA-cleavage domain (Kim *et al.*, 1996). In order to cleave DNA, the cleavage domain has to dimerize and thus one pair of ZFNs is required to target non-palindromic DNA sites (Bitinaite *et al.*, 1998).

TAL effectors are naturally occurring proteins from *Xanthomonas*, a plant bacterial pathogen. The central targeting domain contains a series of 33–35 aa repeats each recognizing a single base pair (bp). Specificity is determined by two hypervariable aa, the repeat-variable diresidues (RVDs) (Deng *et al.*, 2012; Mak *et al.*, 2012). Modular TALE repeats are linked together to recognize contiguous DNA sequences, similar to ZFNs. In contrast to ZFN proteins, no re-engineering of the linkage between repeats is necessary to construct long arrays of TALEs with the ability to basically target single sites in the genome. DNA binding domains can be coupled to various effectors, most importantly including cleaving domains of nucleases like *FokI* to induce targeted DSBs (Cermak *et al.*, 2011; reviewed in (Gaj *et al.*, 2013). On the other hand, meganuclease technology utilizes naturally occurring homing endonucleases, e.g. I-*CreI* and I-*SceI* enzymes, to re-engineer their DNA-binding specificity to target novel sequences (reviewed in Chevalier & Stoddard, 2001).

Meganuclease and TALEN based gene editing approaches have already been used in preclinical trials for XP gene therapy. For example, Dupuy *et al.* (Dupuy *et al.*, 2013) applied engineered meganuclease and TALEN to perform a targeted correction of an *XPC* mutation in a patient cell line (XP4PA). After successful *XPC* gene correction, the full-length *XPC* protein was re-expressed resulting in full recovery of WT UV resistance and functional repair capabilities in the patient cell line.

1.5.2 CRISPR/Cas9

Notwithstanding, the above mentioned techniques are often inefficient, time consuming, laborious, and expensive (Wei *et al.*, 2013). The Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR)/ CRISPR associated (Cas) nuclease 9 system is adapted from the adaptive bacterial immune system of *Streptococcus pyogenes* that evolved in bacteria to defend against invading plasmids and viruses. In this setting, a nuclease is guided by small RNAs to highly efficiently target specific DNA sequences (Cong *et al.*, 2013; Li *et al.*, 2013b). In the CRISPR/Cas9 strategy, a DSB is induced by Cas9 cleavage and the target locus is typically repaired by the error-prone NHEJ or the error-free HRR. In the absence of a repair template, the NHEJ pathway re-ligates DSBs, resulting in insertion/deletion mutations

and therefore in frameshift mutations and premature stop codons as mentioned above (Deltcheva *et al.*, 2011) (see 1.2.2.3).

In the natural context, invading DNA from viruses or plasmids is cut into small fragments and incorporated into the CRISPR locus amidst a series of short repeats (around 20 bps). When the loci are transcribed, and transcripts are processed, small RNAs (crRNA – CRISPR RNA) are generated guiding the Cas9 effector endonuclease together with a tracrRNA to target invading DNA based on sequence complementarity. Upon another encounter, the Cas9 complexed with a crRNA and separate tracrRNA then cleaves the foreign DNA containing the 20 nt crRNA complementary sequence adjacent to the PAM sequence. (Barrangou *et al.*, 2007; Jinek *et al.*, 2012) (see Figure 10).

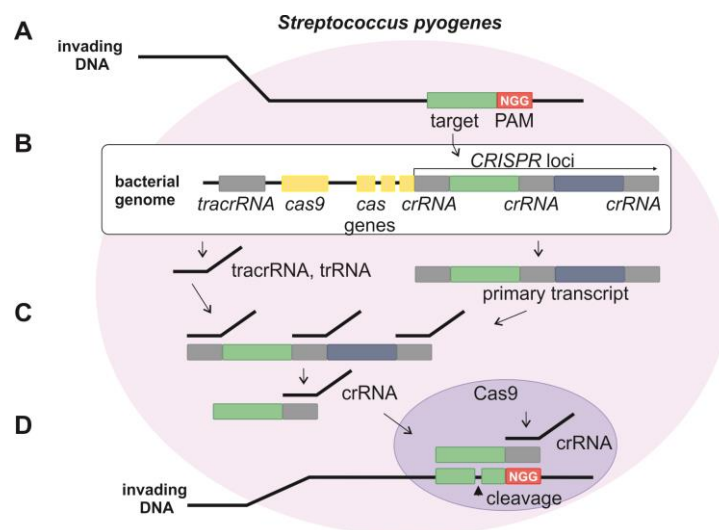


Figure 10: Schematic overview of the CRISPR/Cas9 system of *Streptococcus pyogenes*

Foreign invading DNA (A) is incorporated into the bacterial genome at the CRISPR loci during the acquisition phase (B). The CRISPR loci is transcribed, and processed into crRNA during crRNA biogenesis (C). If the cell is challenged with the same foreign invading DNA again, the Cas9 endonuclease complexed with a crRNA and separate tracrRNA cleaves foreign DNA containing a 20 nt crRNA complementary sequence adjacent to the PAM sequence (D). Illustration by Dr. rer. nat. Christina Seebode.

For genome editing, synthetic guide RNAs have been simplified by fusing together the crRNA and tracrRNA of the CRISPR/Cas9 system (Jinek *et al.*, 2012). This system can be applied to all tissues and cell types containing the mentioned DSB repair pathways. The advantage of the CRISPR/Cas9 system over ZFNs and TALENs lies in the markedly easier design, high specificity, efficiency, and applicability for high-throughput and multiplexed gene editing in a variety of cell types and organisms bringing the CRISPR/Cas9 technology to the fore of the genome editing field (reviewed in Boettcher & McManus, 2015).

1.5.3 Limitations and precautions

Special attention has to be paid to nuclease off-target activity, which can be predicted by several online tools (Hsu *et al.*, 2013; Pellagatti *et al.*, 2015). However, a precise prediction of off-target sites due to tolerable mismatches, epigenomic effects, DNA methylation and chromatin structure is not possible yet. However, off-target mutations are less common in CRISPR/Cas9 systems compared to ZFNs, TALENs, and meganucleases analyzed by whole-genome sequencing studies (Suzuki *et al.*, 2014; Veres *et al.*, 2014). In regard to cleavage efficiency, it has to be mentioned that ZFNs show relatively weak cleavage of chromosomal DNA, in comparison to 100% cleavage efficiency of TALENs in mammalian cell lines (Kim *et al.*, 2010). Concededly, mismatched dimer formation of TALENs can cause high mutation rates (Kim *et al.*, 2013). Interestingly, ZFNs are nearly equally effective in creating both deletions and insertions, while TALENs preferably induce deletions (Cornu *et al.*, 2008).

1.6 Alternative mRNA Splicing

In prokaryotes the process of splicing cannot be observed, as their genes are not organized into exons and introns. Consequently, many bacteria contain polycistronic mRNA, meaning one mRNA that encodes for several proteins (reviewed in Kozak, 1983). In eukaryotic cells, all classes of RNA are co- or post-transcriptionally processed including mRNA splicing, 5' capping, and 3' polyadenylation. All these processes occur in the nucleus before the mRNA is transported into the cytoplasm for protein biosynthesis (reviewed in Bentley, 2002). Pre-mRNA splicing is defined as the joining together of exons while removing the introns and is an essential part of eukaryotic gene expression as it presents a regulatory tool of the cell to produce multiple mRNA molecules from a single gene. The process of alternative splicing is highly coordinated, regulated and has an influence on the amount of information encoded in the transcriptome of the proteome (reviewed in McManus & Graveley, 2011). Usually, alternative splicing events can be grouped into four categories:

- 1) alternative 5' splicing sites
- 2) alternative 3' splicing sites
- 3) cassette exons
- 4) retained introns

It is possible that two or more of these events come together or take place at multiple locations in a gene, thereby evolving extremely complex splicing patterns that give rise to different isoforms of a gene. Moreover, alternative splicing can function in quantitative gene

control by targeting RNAs for nonsense-mediated decay (reviewed in Matlin *et al.*, 2005). In humans, 95% of all genes that consist of more than two exons are alternatively spliced to increase the diversity of the cell's protein composition (Pan *et al.*, 2008). Interestingly, especially for XP genes a number of physiologically occurring spontaneous mRNA variants could be identified. In particular, this applies for splice variants of the two endonuclease complexes XPF/ERCC1 and XPG (see www.ensembl.org).

Hitherto, there have been several reports of intron-retaining splice variants of functional importance (Busse *et al.*, 2009; Honda *et al.*, 2012; Whiley *et al.*, 2011) being a good rationale to analyze the residual function of XPF/ERCC1 and XPG spontaneous mRNA splice variants.

1.6.1 Mechanism of mRNA splicing

Pre-mRNA splicing is a highly coordinated process involving different *cis*-acting elements. The 5' splice site is characterized by the AG/GURAGU sequence (‘/’ defining the splice site), while the 3' splice site contains a polypyrimidine patch followed by an AG dinucleotide at the actual splice site. Furthermore, the so-called branchpoint can be found upstream the 3' splice site (reviewed in Sperling *et al.*, 2008). 95% of splicing reactions are catalyzed by the major spliceosome consisting of five small nuclear ribonucleoprotein particles (snRNPs; U1, U2, U4, U5 and U6), each containing a small nuclear RNA (snRNA) and different proteins. Moreover, the spliceosome also contains 100-200 non-snRNP proteins (reviewed in Wahl *et al.*, 2009). A number of other RNA binding proteins are involved in the splicing process, e.g. SR proteins (proteins with long stretches of serine and arginine residues) and hnRNPs (reviewed in Wang & Burge, 2008). The minor spliceosome processes a rare class of pre-mRNA introns, denoted U12-type, consisting of less abundant snRNAs U11, U12, U4atac, and U6atac, together with U5. The minor spliceosome is also located in the nucleus like its major counterpart (Pessa *et al.*, 2008).

Figure 11 depicts the mechanism of major spliceosomal assembly. Firstly, U1 interacts with the 5' splice site, while U2 binds to the 3' polypyrimidine tract. The U4-U6-U5 tri-snRNP is pre-assembled from the U5 and U4/U6 snRNPs. Subsequently, it is recruited, generating the pre-catalytic B complex. Large rearrangements in RNA–RNA and RNA–protein interactions, leading to the destabilization of the U1 and U4 snRNPs, give rise to the activated spliceosome (reviewed in Will & Luhrmann, 2011).

U1 snRNP binding to the 5' splice site and other non-snRNP associated factors is essential for first recognition and formation of the early (E) complex. U2 snRNP is recruited to the branch

region through interactions with the E complex component U2AF (U2 snRNP auxiliary factor). U2 snRNP becomes tightly associated with the branch point sequence to form complex A in an ATP-dependent reaction (Jamison *et al.*, 1992). A duplex formed between U2 snRNP and the hnRNA branch region bulges out the branch adenosine specifying it as the nucleophile for the first transesterification (Query *et al.*, 1994). The U2 snRNA is placed nearly opposite of the branch site, which results in an altered conformation of the RNA-RNA duplex. In detail, the altered structure of the duplex induced places the 2' OH of the bulged adenosine in a favorable position for the first step of splicing (Newby & Greenbaum, 2002). Afterwards, the U4/U5/U6 tri-snRNP is recruited to the assembling spliceosome to form complex B. Following additional conformational changes, complex C is activated for catalysis. U5 snRNP interacts with sequences at the 5' and 3' splice sites and U5 protein components interact with the 3' splice site region (Chiara *et al.*, 1997). Finally, the intron is spliced out and ligation of the 5' and 3' exons leads to formation of a mature mRNA.

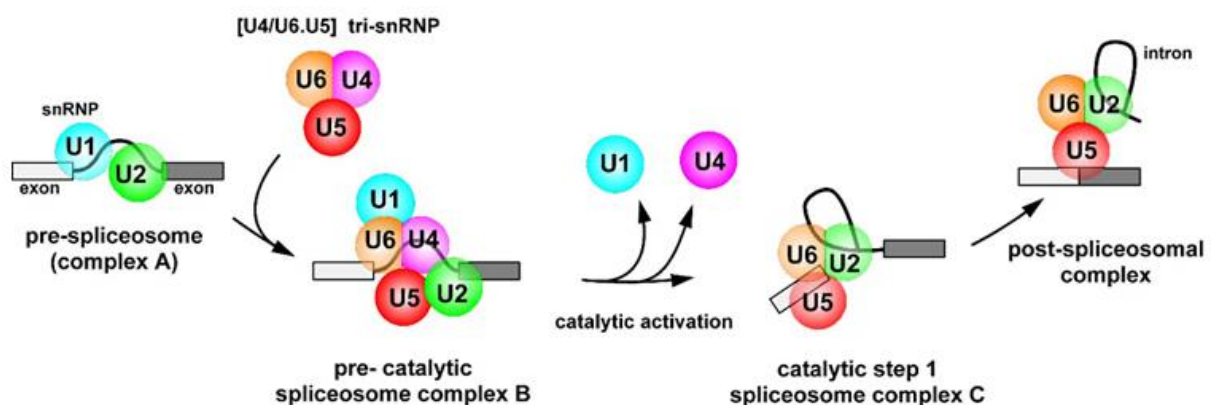


Figure 11: Major spliceosomal assembly

The spliceosomal complex is assembled on each pre-mRNA intron in a multi-step process. Thereto, the snRNPs U1, U2, U4, U5, and U6, together with various additional splicing factors, interact with the pre-mRNA. This results in the formation of "mature" spliceosomes (spliceosomal B complexes) which are then catalytically activated by subsequent steps involving complex structural and compositional rearrangements. In a first step the 5' end of the intron is cleaved, generating the so-called spliceosomal C complex. In the following, this complex catalyzes cleavage at the 3' end of the intron, and ligation of the 5' and 3' exons to form a mature mRNA. From <http://www.mpibpc.mpg.de/luehrmann> [01.03.2017].

1.6.2 Splice site mutations and polymorphisms

In approximately 15% of all diseases, point mutations lead to defective pre-mRNA splicing (Krawczak *et al.*, 1992). These splice site mutations can have different consequences, e.g. decreased recognition of the adjacent exon consequently inhibiting splicing of the adjacent intron, exon skipping, and activation of cryptic splice sites or intron retention (Nakai & Sakamoto, 1994). As already mentioned, deep sequencing has shown that over 95% of human genes undergo alternative splicing, thereby requiring exon-intron boundary recognition. In the

case that SNPs are located in the splice sites this can influence exon configuration. Furthermore, these splice site SNPs can alter translation efficiency of the mRNA and lead to important changes in disease susceptibility (Field *et al.*, 2005; Iwao *et al.*, 2004). Splice junction sequences are highly important because changes in splice sites alter recognition efficiency of splicing factors, then leading to altered exon recognition events and changes in the composition of amino acids and protein domains. Changes of nucleotide sequences in splice sites weaken exon-intron junction strength (splice site strength score) and frequently lead to the breaking of consensus nts at splice sites (Fox-Walsh *et al.*, 2005; Roca *et al.*, 2008). To simultaneously profile both gene expression levels and the types of isoforms that are being expressed next-generation RNA sequencing is an effective tool (reviewed in Marioni *et al.*, 2008; Wang & Burge, 2008; Wang *et al.*, 2009b).

Despite their existence and sequence not much is known about the function of spontaneous splice variants of XP genes. In the context of residual catalytic activity and individual expression levels those variants could be able to modify a person's individual NER and ICL repair capabilities. Additionally, these physiologically occurring variants, together with artificially generated point mutants, display interesting tools to investigate the relevance of functional protein domains. It can be suggested that individual expression levels of specific functionally relevant splice variants influence cancer risk, disease progression and therapeutic success. Interestingly, the XPC polyA-T polymorphism (PAT) has been associated with an increased risk for head and neck cancer, SCC, as well as melanoma (Blankenburg *et al.*, 2005; Marioni *et al.*, 2008; Shen *et al.*, 2001; Wang & Burge, 2008; Wang *et al.*, 2009b). Zhang *et al.* could show that the expression level of splice variants in different tissues is more suitable to distinguish between oncogene and non-oncogene samples than the primary gene transcript itself (Zhang *et al.*, 2013b). Additionally, tumor-specific splice variants, in the case of BRCA1/2 or p53-inhibitor MDM2, are often overexpressed (Brinkman, 2004; Yi & Tang, 2011). In human melanoma cell lines splice variants of the p53 protein could be observed to be augmented as well (reviewed in Wei *et al.*, 2012).

In view of these results, it is of great interest to characterize physiologically occurring spontaneous mRNA splice variants of the *XPF/ERCC1* and *XPG* genes, making it the central aim of this thesis.

1.7 Aim of the study

The NER pathway is a central DNA repair mechanism to repair a variety of bulky DNA lesions (e.g. pyrimidine photo products) independent of the cell cycle. Accumulation of these types of damage all over the genome results in the development of a cancer prone mutator phenotype as it can be seen in patients with the autosomal recessive model disease XP, their high frequency of UV-induced skin tumors (~1000 fold increase compared to the healthy population), and reduced life span (approximately 37 years). Furthermore, it is known that a decreased NER level is a risk factor for several cancer and tumor entities in the normal population, components of the NER pathway already serve as biomarker for the etiopathology of tumor diseases (e.g. XPG in melanoma), and the chemotherapeutic success of the treatment. The endonucleases XPF/ERCC1 and XPG are the core components of the NER incision complex and the heterodimer XPF/ERCC1 is also involved in repair of DNA ICLs. Currently, there is an ongoing discussion about ERCC1 and its ability to serve as a biomarker for personalized medicine. Our group and others already demonstrated that these three essential repair genes show a high number of physiologically occurring spontaneous alternative spliced transcripts with unclear functions and a difference in expression in several tissues (Emmert *et al.*, 2001).

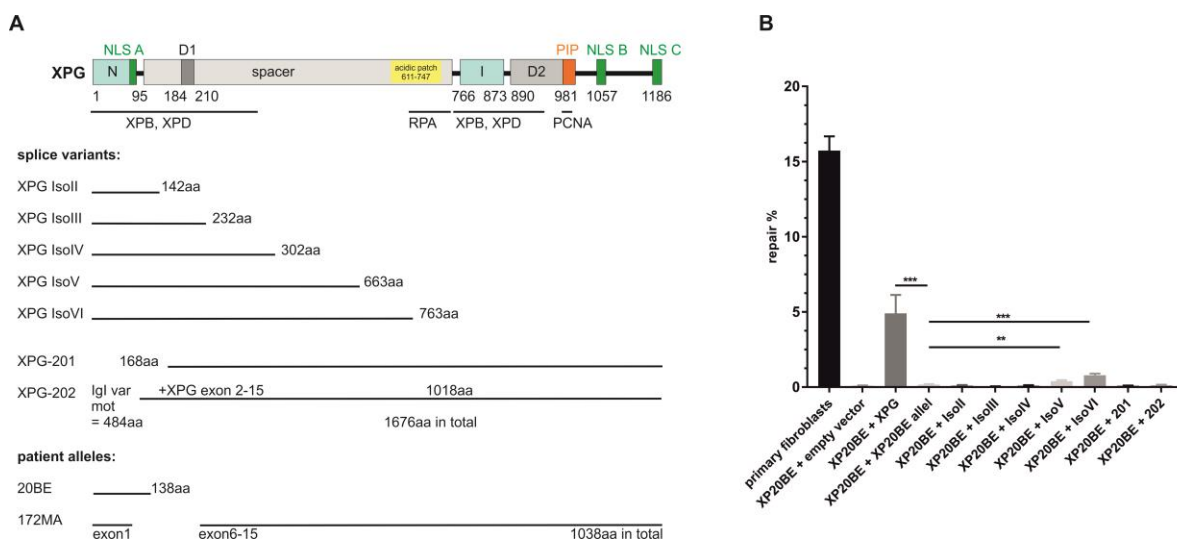


Figure 12: Overview of the XPG spontaneous mRNA splice variants

(A) Physiologically occurring spontaneous mRNA splice variants of the XPG endonuclease produce several C-terminally truncated proteins (IsoII - IsoVI) missing one or more of the functional domains. On the other hand, variant 201 misses exons one - four (in frame) including one of the nuclease domains (N) and part of an interaction patch with TFIIF. 202 only lacks exon one, but in addition contains the immunoglobulin like variant motif of 484 aa, caused by splicing defects leading to a conjoined gene (see www.ensembl.org; (Emmert *et al.*, 2001)). (B) Reactivation of a reporter gene after treatment with UVC. A firefly plasmid containing specific lesions due to treatment with UVC was transfected into XP20BE primary fibroblasts and complemented with

plasmids encoding for *XPG* or its splice variants. The relative repair capability is calculated as the percentage (repair %) of the reporter gene activity (firefly luciferase) compared to the untreated plasmid, after normalization to an internal co-transfected control (*Renilla* luciferase). Data are presented as the mean \pm SEM. The one-tailed, unpaired student's t-test was applied, ** $P < 0.01$, *** $P < 0.001$. At least four independent experiments in triplicates were performed.

During my master thesis, I characterized six spontaneous *XPG* splice variants and was able to identify two splice variants with residual repair capabilities (see Figure 12). These results suggest that other endonucleases may also take part in NER and could be reaction partners of the described splice variants. Furthermore, splice variants of the second endonuclease (*XPF/ERCC1*) involved during NER are of special interest.

Therefore, the goal of my Ph.D. project was to elucidate the catalytic functions or negative influence of the alternative splice variants of the three genes *XPF*, *ERCC1* and *XPG* during NER and ICL removal. So far, a key limitation to study the functions of the *XPF/ERCC1* complex has been the lack of an appropriate human *XPF*- or *ERCC1*-defective cell line. With regard thereto, generation of *XPF* and *ERCC1* KO cell lines using the CRISPR/Cas9 tool was the first part of the project. Afterwards, this cell lines had to be characterized in regard to sensitivity against different genotoxins and repair capabilities. In a second step, the goal was to amplify and clone alternative splice variants of *XPF* and *ERCC1* from WT human skin fibroblasts to elucidate catalytic and possible dominant negative effects, by functional testing using different repair assays (see Figure 13).

The last step was the evaluation if and how the different transcript levels of the relevant spontaneous isoforms of the three genes are able to serve as prognostic factors for the individual UV-induced skin cancer risk. With these investigations the project provides an important contribution to the understanding of the molecular mechanisms in NER and ICL repair as well as the establishment of splice variants as new prognostic biomarkers regarding personalized medicine in terms of carcinogenesis of the skin. It can be summarized by the following bullet points:

- 1) generation of an *XPF* and *ERCC1* KO cell line using the CRISPR/Cas9 tool
- 2) characterization of KO cells as a useful model cell line
- 3) amplification and cloning of postulated alternative splice variants of *XPF* and *ERCC1* from normal human skin fibroblasts
- 4) functional analysis of alternative splice variants of *XPF* and *XPG*
- 5) implication of functionally relevant splice variants as prognostic marker for individual cancer risks or therapeutic success/disease course

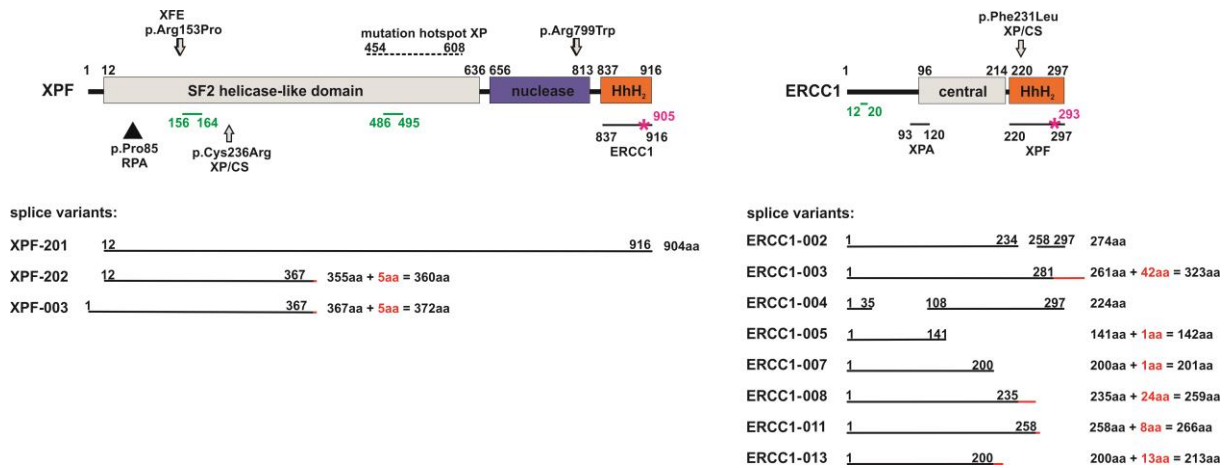


Figure 13: Overview of the *XPF* and *ERCC1* spontaneous mRNA splice variants

The *XPF* gene can be alternatively spliced into C- or N-terminally truncated physiologically occurring variants. While XPF-201 only lacks the first 12 amino acids, XPF-202 and XPF-003 are severely C-terminally truncated, lacking functional domains, e.g. the nuclease domain. ERCC1 also comprises C-terminally truncated splice variants (ERCC1-003, ERCC1-005, -007, -008, -011, and -013). Furthermore, ERCC1-002 and ERCC1-004 have internal in frame deletions. Red lines and numbers indicate the retention of intronic sequences.

2 Materials

2.1 Biological material

2.1.1 Cell lines

Cell lines that were used in this project are listed in Table 2. Primary human fibroblasts have been isolated from skin punch biopsies of healthy donors. They function as a source for physiological mRNA splice variant preparations.

Table 2: Cell lines

Cell line	Source	Distributor
WT fibroblasts	primary fibroblasts	derived from healthy donor punch biopsies
XP20BE	XP patient fibroblasts	Coriell Cell Repository, Camden New Jersey, USA
MRC5Vi	immortalized fetal lung fibroblasts	gift from Sarah Sertic (Sertic <i>et al.</i> , 2011)
HeLa	immortal cervical cancer cells	ATCC Manassas Virginia, USA

2.1.2 Bacteria

Escherichia coli (*E. coli*) DH5 α [F-, endA1, glnV44, thi-1, recA1, relA1, gyrA96, deoR, nupG, Φ 80dlacZ Δ M15, Δ (lacZYA-argF)U169, hsdR17(rK- mK+), λ -] (Hanahan, 1983) were used for transformation and amplification of plasmid DNA.

2.2 Consumable supplies

Table 3: Consumables

Consumables	Manufacturer
96 well Glomax™ 96 Microplate	Promega, Mannheim GER
Blotting paper extra thick	Bio-Rad Laboratories, Munich GER
Cell culture flasks (25cm ³ , 75cm ³ , 175cm ³)	Greiner Bio-One, Frickenhausen GER
CELLSTAR® CELLreactor™, Polypropylene Filter Top Tube	Greiner Bio-One, Frickenhausen GER
Cell strainer (40 μ m)	VWR, Darmstadt GER
Cryo boxes	Nunc, Wiesbaden GER
Cryo tubes 2ml	Greiner Bio-One, Frickenhausen GER
Dispenser multipipette® plus	Eppendorf, Hamburg GER
Dispenser tips (50 μ l, 5ml)	Eppendorf, Hamburg GER

Erlenmeyer flask (200ml, 500ml)	VWR, Darmstadt GER
Falcon tubes 15ml, 50ml	Sarstedt, Nuembrecht GER
Glass cover slips, round	Roth, Karlsruhe GER
Kimtech wipes	Kimberly-Clarke Professional, Roswell USA
Glass pipettes (5ml, 10ml, 25ml)	Brand, Wertheim GER
Neubauer cell counting chamber	Brand, Wertheim GER
Microscope slides	Thermo Fisher Scientific, Braunschweig GER
Mini TGX Gels	Biorad, Munich GER
Nitrocellulose, 0.45 μ M Protran BA85	Amersham Bioscience, Piscataway USA
Parafilm [®]	Brand, Wertheim GER
Pasteurpipettes (230nm)	Brand, Wertheim GER
PCR tubes	Sarstedt, Nuembrecht GER
Petri dishes	Greiner Bio-One, Frickenhausen GER
Pipettes (10 μ l, 100 μ l, 200 μ l, 1000 μ l)	Eppendorf, Hamburg GER
Pipette tips (10 μ l, 100 μ l, 1000 μ l)	Sarstedt, Nuembrecht GER
Polystyrene round bottom tube	BD Falcon, Heidelberg GER
Reaction tubes (0,5ml, 1,5ml, 2,0ml, 5,0ml)	Eppendorf, Hamburg GER
Scalpel, disposable	Feather, Osaka JPN
Tissue culture multiwell plates (6-well, 24-well, 96-well)	Greiner Bio-One, Frickenhausen GER
Tissue culture dish (10cm)	Greiner Bio-One, Frickenhausen GER
Toothpick	Fackelmann, Hersbruck GER

2.3 Equipment

Table 4: Equipment

Instrument	Manufacturer
3100-Avant Genetic Analyzer	Applied Biosystems, Foster City USA
7900HT Fast Real-Time PCR System	Applied Biosystems, Foster City USA
Analytical balance BP2100; MC1	Sartorius, Goettingen GER
Autoclave DE-65	Systec, Wettenberg GER
Benchtop centrifuge 5415C	Eppendorf, Hamburg GER
ChemoCam Imager 6.0	Intas, Goettingen GER
CO ₂ -Incubator	Sanyo, Munich GER
Double Distilled Water System Arium [®] 611VF	Sartorius, Goettingen GER
DU 640 [®] Spectrophotometer	Beckmann, Munich GER
Freezer -20, +4 combination	Liebherr, Rostock GER
Gel iX 20 Imager	Intas, Goettingen GER
Glomax [®] Discover System	Promega, Mannheim GER

Hera freeze -80°C freezer	Heraeus Instruments, Hanau GER
Ice machine ZBE 30-10	Ziegra, Isernhagen GER
Incubator model 200	Memmert, Buechenbach GER
Microscope Axiovert A1	Carl Zeiss, Oberkochen GER
Microscope Axioskop 2	Carl Zeiss, Oberkochen GER
Microscope Axiovert 100	Carl Zeiss, Oberkochen GER
Microscope Axiovert Imager M1	Carl Zeiss, Oberkochen GER
Microwave	Panasonic, Hamburg GER
Mini-PROTEAN Tetra Cell	Bio-Rad Laboratories, Munich GER
Mini-PROTEAN® TGX™ Precast Protein Gels 4-15%	Bio-Rad Laboratories, Munich GER
Mini Rocking Platform	Biometra, Goettingen GER
NanoVuePlus®	GE Healthcare, Buckinghamshire UK
pH meter	Schuett, Goettingen GER
Pipetboy acu	Integra Biosciences, Fernwald GER
Power Supply Ease 500	Invitrogen, Karlsruhe GER
Shaker	Infors, Bottmingen SWZ
Hera safe sterile bench	Thermo Fisher, Waltham MA, USA
Thermomixer 5436	Eppendorf, Hamburg GER
T Gradient Thermo Block	Biometra, Goettingen GER
ULT7150-9-M Cryogenic freezer	Thermo Fisher, Waltham MA, USA
UVA Crosslinker BIO-Link BLX 365	Vilber Lourmat, Eberhardzell GER
Trans-Blot® Turbo™ Transfer System	Bio-Rad Laboratories, Munich GER
UVC 500 Ultraviolet Crosslinker with five 8W 254nm UV lamps	Amersham Bioscience, Piscataway USA
Vortexer Vibrofix VF1 Electronic	IKA Labortechnik, Staufen GER
Water bath	GFL, Großburgwedel GER

2.4 Chemicals

Table 5: Chemicals

Chemical	Manufacturer
5x HOT FIREpol® EvaGreen® qPCR Mix Plus (ROX)	Solis Biodyne, Tartu EST
Agarose Saekem®	Fluka Chemie, Neu-Ulm GER
Ammonium persulfate (APS)	Sigma-Aldrich, Taufkirchen GER
Ampicillin	Sigma-Aldrich, Munich GER
Aqua ad iniectabilia	Braun, Melsungen GER
Boric acid	Sigma-Aldrich, Taufkirchen GER
Bradford Mix Roti®	Roth, Karlsruhe GER

Bromphenol blue	Sigma-Aldrich, Taufkirchen GER
Calcium chloride (CaCl ₂)	Merck, Darmstadt GER
Complete ULTRA Tablets Mini EDTA	Roche, Mannheim GER
Dithiothreitol (DTT)	Sigma-Aldrich, Munich GER
Dimethylsulfoxide (DMSO)	Sigma-Aldrich, Taufkirchen GER
DNase	Qiagen, Hilden GER
dNTP Mix (dATP, dCTP, dGTP, dTTP)	Thermo Fisher Scientific, Braunschweig GER
Ethanol absolute, 99,8%	Merck, Darmstadt GER
Ethylenediaminetetraacetic acid (EDTA)	Sigma-Aldrich, Taufkirchen GER
EGTA (ethylene glycol-bis(β-aminoethyl ether)-N,N,N',N'-tetraacetic acid)	Sigma-Aldrich, Taufkirchen GER
Fluorescence mounting medium	Agilent Technologies Inc., Santa Clara USA
Formaldehyde 4%	Serva, Heidelberg GER
G418 disulfate salt	Sigma-Aldrich, Taufkirchen GER
Gene Ruler (100bp #SM0241, 1kb SM0311) DNA ladder	Thermo Fisher Scientific, Braunschweig GER
Glycine (C ₂ H ₅ NO ₂)	Sigma-Aldrich, Taufkirchen GER
Glycerol (C ₃ H ₈ O ₃)	Merck, Darmstadt GER
HD Green	Intas, Goettingen GER
Hepes	PAA, Coelbe GER
Hi-Di Formamide	Applied Biosystems, Foster City USA
Hoechst33342	Life Technologies, Eggenstein GER
Isopropanol (C ₃ H ₈ O)	Merck, Darmstadt GER
LB Agar	Life technologies, Darmstadt GER
LB Broth Base	Life technologies, Darmstadt GER
Magnesium Chloride (MgCl ₂)	Merck, Darmstadt GER
Marker VI, protein ladder	Applichem, Chicago IL, USA
Methanol (CH ₃ OH)	Mallinckrodt Baker, Griesheim GER
N,N,N,N-tetramethyl-ethane-1,2-diamine (TEMED)	Sigma-Aldrich, Taufkirchen GER
Nonidet P40	Sigma-Aldrich, Taufkirchen GER
Paraformaldehyde (PFA)	Merck, Darmstadt GER
PhosSTOP Phosphatase Inhibitor Cocktail	Roche, Mannheim GER
Phenylmethanesulfonylfluoride (PMSF)	Sigma-Aldrich, Taufkirchen GER
Ponceau S	Sigma-Aldrich, Taufkirchen GER
Potassium Chloride (KCl)	Merck, Darmstadt GER
Puromycin	Invivogen, San Diego, CA USA
Roti [®] -Free Stripping Buffer	Roth, Karlsruhe GER
Sodium chloride (NaCl)	Merck, Darmstadt GER
Sodium dodecyl sulfate (SDS)	Roth, Karlsruhe GER

Sodium hydroxide (NaOH)	Merck, Darmstadt GER
Superfibronectin	Sigma-Aldrich, Taufkirchen GER
Trifluoroacetic acid (TFA)	Merck, Darmstadt GER
Tris-Base	Merck, Darmstadt GER
Triton X-100	Merck, Darmstadt GER
Vectashield Mounting Medium for Fluorescence with DAPI 1:1200	Vector Laboratories, Burlingham CA/USA
β -mercaptoethanol (C ₂ H ₆ OS)	Merck, Darmstadt GER

2.5 Buffers, solutions and media

Frequently used buffers, solutions and media were prepared as listed below.

Table 6: Buffers, solutions and media

	Receipt/ Manufacturer
Cell culture	
Dulbecco's Modified Eagle Medium (DMEM) high glucose	PAA, Coelbe GER
FibroLife [®] fibroblast medium	Lifeline Cell Technology, Frederick MD, USA
Freezing medium (fibroblasts)	40% DMEM 40% (v/v) FBS 20% (v/v) DMSO
Fetal Bovine Serume (FBS)	Biochrom AG, Berlin GER
Penicillin-Streptomycin (100x)	PAA, Coelbe GER
Trypanblue solution 0,4%	Sigma-Aldrich, Taufkirchen GER
Trypsin/EDTA	Biochrom AG, Berlin GER
Gel electrophoresis	
TBE	89mM Tris-Base 89mM Boric acid 2mM EDTA
Protein biochemistry	
Buffer A	10mM Hepes 10mM KCl 0.1mM EGTA 0.1mM EDTA 1mM DTT 0.5mM PMSF
Buffer B	20mM Hepes 400mM NaCl 1mM EGTA

	1mM EDTA 1mM DTT 1mM PMSF
Immunoblot transfer buffer, pH 8.3	0.192M Glycine 0.025M Tris-Base 20% MeOH (v/v) 0.01% SDS
Ponceau S solution	0.2% (v/v) Ponceau S 3% (v/v) TFA
4% paraformaldehyde (PFA)	3.6g PFA 8ml PBS 8 drops 1M NaOH → dissolve stirring at 60°C → fill up to 50ml with PBS
SDS running buffer, pH 8.3	0.192M Glycine 0.025M Tris-Base 0.1% SDS
SDS splitting buffer, pH 4,7	30mM Tris-Base 3% SDS 15% Glycerine 0,04% Brompehnol blue before use add 10% β-mercaptoethanol (v/v)
Additional commonly used buffers	
10x PBS, pH 7.2	1.5M NaCl 30mM KCl 80mM Na ₂ HPO ₄ x 2H ₂ O 10mM KH ₂ PO ₄
6x DNA loading buffer	0.5 M EDTA 50% (v/v) Glycerol 0.01% Brompehnol blue

2.6 Oligonucleotides

The following oligonucleotides were utilized in this work. For cloning into pcDNA3.1(+) (XPF KpnI (5'- ttaGGTACC -3'), XbaI (5'- ttaTCTAGA -3')); ERCC1 HindIII (5'- ttaAAGCTT -3'), XhoI (5'- ttaCTCGAG -3')), pcDNA3.1(-)mycHisA2 (XPF XbaI, KpnI; ERCC1 XhoI, HindIII), and pcDNA3.1(+)eGFP (XPF KpnI, XbaI; ERCC1 HindIII, XhoI) respective restriction recognition sequences were added.

Table 7: Oligonucleotides

Gene	Oligonucleotide
Amplification primer XPF/ERCC1 and splice variants	
XPF_fwd	5'- ATGGAGTCAGGGCAGCC -3'
XPF_rev	5'- TCACTTTTTCCCTTTTCCTTTTGA -3'
XPF_rev_w/oStop	5'- CTTTTTCCCTTTTCCTTTTGATAC -3'
XPF201_fwd	5'- ATGGCGCCGCTGCTGGA -3'
XPF003_rev	5'- TTAACCCCAACAAGATACCTTCCC -3'
XPF003_rev_w/oStop	5'- ACCCCACAAGATACCTTCCCCT -3'
ERCC1_fwd	5'- ATGGACCCTGGGAAGGACAAA -3'
ERCC1_rev	5'- TCAGGGTACTTTCAAGAAGGGCT -3'
ERCC1_rev_w/oStop	5'- GGGTACTTTCAAGAAGGGCTCGT -3'
ERCC1-003_rev	5'- TTACAGGCGGAAGCCACTGT -3'
ERCC1-003_rev_w/oStop	5'- CAGGCGGAAGCCACTGTGT -3'
ERCC1-005_rev	5'- TCACCTGAGGAACAGGGCA -3'
ERCC1-005_rev_w/oStop	5'- CCTGAGGAACAGGGCACA -3'
ERCC1-007_rev	5'- TTACAGATGAGGAAACTGAAGGC -3'
ERCC1-007_rev_w/oStop	5'- CAGATGAGGAAACTGAAGGCCA -3'
ERCC1-008_rev	5'- TTAAAATTGGAAGTGAAGCTCAAC -3'
ERCC1-008_rev_w/oStop	5'- AAATTGGAAGTGAAGCTCAACCAC -3'
ERCC1-011_rev	5'- GGCAGGGAGATGGAAGGAA -3'
ERCC1-011_rev_w/oStop	5'- TGGGGCAGGGGAGCC -3'
ERCC1-013_rev	5'- CAAGAAGGGCTCGTGCAG -3'
ERCC1-013_rev_w/oStop	5'- AGAAGGGCTCGTGCAGGAC -3'
Quickchange mutagenesis XPF	
QC_XPF_D668A_fwd	5'- CATCTGCAGATGTTTCCACTGCCACTCGGAA -3'
QC_XPF_D668A_rev	5'- CCGGCTTCCGAGTGGCAGTGGAA -3'
QC_XPF_P85S_fwd	5'- GAAGGAGTTGAACACCTCTCTCGCCGTGTAA -3'
QC_XPF_P85S_rev	5'- GTGATTTTCATTTGTTACACGGCGAGAGAGGTGTT -3'
QC_XPF_F905A_fwd	5'- GATTTTCATTCACACCTCTGCTGCAGAAGTCGTATCAAA -3'
QC_XPF_F905A_rev	5'- CCTTTTGATACGACTTCTGCAGCAGAGGTGTGAA -3'
QC_XPF_F905P_fwd	5'- GATTTTCATTCACACCTCTCCTGCAGAAGTCGTATCAAA -3'
QC_XPF_F905P_rev	5'- CCTTTTGATACGACTTCTGCAGGAGAGGTGTGAA -3'
QC_XPF_D731A_fwd	5'- CGTGGAGCGCAAGAGTATCAGTGCTTTAATCGGCT -3'
QC_XPF_D731A_rev	5'- TAAAGAGCCGATTAAGCACTGATACTCTTGCGCT -3'
QC_XPF_R153P_fwd	5'- CATTTCATCTTGCGCCTCTTTCCCCAGAAAAACAA -3'
QC_XPF_R153P_rev	5'- TTTGTTTTTCTGGGGAAAGAGGGCGCAAGATGAAT -3'
QC_XPF_C236R_fwd	5'- CAGACTGCTATACTGGACATTTTAAATGCACGTCTAAAG -3'

QC_XPF_C236R_rev	5'- ATGGGTTATGGCATTTTAGTTCCTTTAGACGTGCAT -3'
QC_XPF_R689S_fwd	5'- CACAGCAAAGCATAGTTGTGGATATGAGTGAATTTTCGAA -3'
QC_XPF_R689S_rev	5'- AGATGGAAGCTCACTTCGAAATTCACTCATATCCACAA -3'
QC_XPF_R490Q_fwd	5'- CCAAAGAAAGAACCCTCAAAAAGAAAAACAGAAGTTGACCTT -3'
QC_XPF_R490Q_rev	5'- GGTTCCTACCATTTGAGTTAAGGTCAACTTCTGTTTTTCTT -3'
Sequencing primer	
M13	5'- TAGAAGGCACAGTCGAG -3'
T7	5'- TAATACGACTCACTATAGGG -3'
U6fwd	5'- ACTATCATATGCTTACCGTAAC -3'
XPFseq_exon1_fwd	5'- ATGGAGTCAGGGCAGC -3'
XPFseq_exon4_fwd	5'- GTTCCATGTAGCAGTAAACT -3'
XPFseq_exon7_fwd	5'- CAAGTGATGACCGAACATGT -3'
XPFseq_exon9_rev	5'- TCACTTTTTCCCTTTTCCT -3'
XPFseq_exon3_rev	5'- CTTGGCCACAGATACAGTT -3'
XPFseq_exon8_rev	5'- GCTTGGCCACAGATACAGT -3'
CMVpromoter_long_fwd	5'- CTGCTTAGGGTTAGGCGTTTTGCGCT -3'
Genomic DNA primer XPF/ERCC1	
XPFexon1_fwd	5'- CACGATCATCTCAGTCTCAGCTC -3'
XPFexon1_rev	5'- CCTAGCGACCCCTTACATACGTC -3'
XPFexon2_fwd	5'- TGTAGACTGGTTGGCTGAAGTTAC -3'
XPFexon2_rev	5'- TGATGTAGGGAGCTGAGTCCTC -3'
XPFexon3_fwd	5'- CTCTGTTCTGTGCGTGGCTATATG -3'
XPFexon3_rev	5'- GAAAAGCAACCATCAAATTGCTCTC -3'
XPFexon4_fwd	5'- GCTTTTCGTGTTGTTTGTAGCA -3'
XPFexon4_rev	5'- GTGATGCTTATATGCCAATCCACAT -3'
XPFexon5_fwd	5'- TAGCCACTTCCTTGAATAATGCTTG -3'
XPFexon5_rev	5'- ACGTTAAGTAGGCGGAAACATTAGC -3'
XPFexon6_fwd	5'- AAGACTTGCCATGCTGTATACTTCG -3'
XPFexon6_rev	5'- GCCAGTTACGTATGTAGGTCATGTG -3'
XPFexon7_fwd	5'- TTCAGGAAAGCAGATTCCATCTAAC -3'
XPFexon7_rev	5'- CACTAGGATCTCAGTGTTTCATTTGC -3'
XPFexon8_fwd	5'- GATTTAAGTAATCTTGCCAGAGAGG -3'
XPFexon8_rev	5'- AGCAGCATCGTAACGGATATTAAG -3'
XPFexon9_fwd	5'- CCTGTGTGGTAACGAGACCTTTAAC -3'
XPFexon9_rev	5'- GGACAATTCAGACCACAGGTTTATC -3'
XPFexon10_fwd	5'- CTTCCTTTACCCATCATTGTCTTG -3'
XPFexon10_rev	5'- CTGGAACATAACCCATTCTAAGCTG -3'
XPFexon11_fwd	5'- CTTTCCTATTAGCTCGGTTTCCTTC -3'
XPFexon11_rev	5'- CTGAAAAGTACAGGCATGGGATAA -3'

ERCC1exon1_fwd	5'- GATGGGACTTGTGGACCTGTA -3'
ERCC1exon1_rev	5'- CTGCCTTAGCCCTCTCTTAGAA -3'
ERCC1exon2_fwd	5'- CCTCAGATGTCCTCTGCTCA -3'
ERCC1exon2_rev	5'- GGAGAACAAAGTGGCTGGAA -3'
CRISPR/Cas9 guideRNA oligonucleotides XPF/ERCC1	
XPF_Cas9-1fwd	5'- CACCgagtcagggcagccggctcgacgg -3'
XPF_Cas9-1rev	5'- AAACccgctcgagccggctgccctgactc -3'
XPF_Cas9-2fwd	5'- CACCtgctggagtacgagcgacagctgg -3'
XPF_Cas9-2rev	5'- AAACccagctgctcctctactccagca -3'
XPF_Cas9-3fwd	5'- CACCctttcgccagaaaaacaaactgg -3'
XPF_Cas9-3rev	5'- AAACccacgtttgtttctggcgaag -3'
XPF_Cas9-4fwd	5'- CACCctgctcctctactccagcagcgg -3'
XPF_Cas9-4rev	5'- AAACccgctcctggagtacgagcgacag -3'
XPF_Cas9-5fwd	5'- CACCtcagctggagaaagtggtagagg -3'
XPF_Cas9-5rev	5'- AAACcctctaccactttctccagctgca -3'
XPF_Cas9-6fwd	5'- CACCatttcattgttacacggcgaggg -3'
XPF_Cas9-6rev	5'- AAACccctcggcgttaacaatgaaat -3'
ERCC1_Cas9-1fwd	5'- CACCaatgtgatcccctcgacgagg -3'
ERCC1_Cas9-1rev	5'- AAACcctcgtcgaggggtatcacaatt -3'
ERCC1_Cas9-2fwd	5'- CACCcatattcgctaggtctgagggg -3'
ERCC1_Cas9-2rev	5'- AAACcccctcagacctacgccgaatag -3'
ERCC1_Cas9-3fwd	5'- CACCcctgtctaccactccaggaggg -3'
ERCC1_Cas9-3rev	5'- AAACcctcctggagtggtagacaagg -3'
ERCC1_Cas9-4fwd	5'- CACCaggacctcatcctcgtcgagggg -3'
ERCC1_Cas9-4rev	5'- AAACcccctcgacgaggatgaggtccct -3'
ERCC1_Cas9-5fwd	5'- CACCcacaatttctctgtggtgggc -3'
ERCC1_Cas9-5rev	5'- AAACgccccagcaaggagaattgtg -3'
ERCC1_Cas9-6fwd	5'- CACCgctctgttagatcggaataaggg -3'
ERCC1_Cas9-6rev	5'- AAACccttattccgatctacacagagc -3'
Oligonucleotides for quantitative real-time PCR purchased from Qiagen, Hilden GER	
Gene	Order number
ERCC1	QT01000195
GAPDH	QT00079247
XPF	QT00029246
XPG	QT00029246
Oligonucleotides for quantitative real-time PCR	
XPG_fwd	5'- GGATCTTCAAGTGAACATGCTGAA -3'
XPG_rev	5'- TGCGAATCTGAAGCACTGGT -3'
XP20BE_fwd	5'- CCAGTGACTCCAGGAAAACGA -3'

XP20BE_rev	5'- GGTACCAACTTGGGTAAGACT -3'
XPGIsoII_fwd	5'- CGTCCAGTGA CTCCAGGAAA -3'
XPGIsoII_rev	5'- CTTTTTTTAAACTTCATCTCTAACACGACT -3'
XPGIsoIII_fwd	5'- GAGTTCACCAAGCGCAGAAG -3'
XPGIsoIII_rev	5'- GAAGAGGGCAAGGGAATGTG -3'
XPGIsoIV_fwd	5'- CCAAAGGAAATGAATCAGCAA -3'
XPGIsoIV_rev	5'- CCCTCTAGCTACGTCAAAGACATG -3'
XPGIsoV_fwd	5'- GAATCTGCAGGCCAGGATTT -3'
XPGIsoV_rev	5'- CCTACCGTTCAGATTTCTACAAA -3'
XPGIsoVI_fwd	5'- GAGCCACAGGAAGCTGAGAAA -3'
XPGIsoVI_rev	5'- CAGCAAGAAGTCGAAACACAATG -3'
XPG201_fwd	5'- CCCAAGCTGGCTAGCGTTTA -3'
XPG201_rev	5'- TGAGGATTATGAAAGAACTCTTCCTGTA -3'
XPG202_fwd	5'- CATCAGGACTCGGCATGGAA -3'
XPG202_rev	5'- CCCGGACTCCTTTAAGTGCT -3'

2.7 Ready to use reaction systems

The following ready to use reaction systems were utilized in this work.

Table 8: Reaction systems

Description	Manufacturer
Attractene Transfection Reagent	Qiagen, Hilden GER
BigDye Terminator v3.1 Cycle Sequencing Kit	Applied Biosystems, Foster City USA
CellTiter 96 [®] Non-Radioactive Cell Proliferation	Promega, Mannheim GER
CloneJET PCR Cloning Kit	Thermo Fisher, Waltham MA, USA
Dual-Luciferase Reporter Assay System	Promega, Mannheim GER
Gel extraction and PCR Clean Up	Machery + Nagel, Düren GER
NucleoBond [®] Xtra MiDi/Maxi	Machery + Nagel, Düren GER
NucleoSpin [®] Plasmid	Machery + Nagel, Düren GER
RNA from 20 different healthy human tissues	Ambion [®] , Huntington, UK (discontinued)
RNase free DNase Set	Qiagen, Hilden GER
RNeasy Mini Kit	Qiagen, Hilden GER
RevertAid H Minus First Strand cDNA synthesis Kit	Thermo Fisher Scientific, Braunschweig GER
Surveyor Mutation Detection Kit	IDT, Coralville IW, USA
QIAamp DNA Blood Kit	Qiagen, Hilden GER
WesternBreeze Chemiluminescent Immunodetection Systems anti mouse/rabbit	Applied Biosystems, Foster City USA

2.8 Enzymes

Table 9: Enzymes

Description	Manufacturer
BbsI	Thermo Fisher, Waltham MA, USA
DpnI 10u/μl	Thermo Fisher, Waltham MA, USA
ExoSAP	Affymetrix, Santa Clara CA, USA
HindIII 10u/μl	Thermo Fisher, Waltham MA, USA
KpnI 10u/μl	New England Biolabs, Ipswich MA, USA
Phusion DNA Polymerase	New England Biolabs, Ipswich MA, USA
PvuI	New England Biolabs, Ipswich MA, USA
T4 DNA Ligase 1u/μl	Thermo Fisher, Waltham MA, USA
T4 Polynukleotid Kinase	New England Biolabs, Ipswich MA, USA
T7 Endonuclease I	New England Biolabs, Ipswich MA, USA
XbaI 10u/μl	New England Biolabs, Ipswich MA, USA
XhoI 10u/μl	Thermo Fisher, Waltham MA, USA
CIAP (calf intestinal alkaline phosphatase) 1u/μl	New England Biolabs, Ipswich MA, USA

2.9 Plasmids

Table 10: Plasmids

Description	Manufacturer
pcDNA3.1(+)	Life technologies, Darmstadt GER,
pcDNA3.1(-)mycHisA2	Life technologies, Darmstadt GER,
pCMVluc	Promega, Mannheim GER
pRL-CMV	Promega, Mannheim GER
pX330	Addgene, Cambridge MA, USA
pX458	Addgene, Cambridge MA, USA
pX459	Addgene, Cambridge MA, USA
pX462	Addgene, Cambridge MA, USA

2.10 Antibodies

Primary and secondary antibodies were diluted in 10% FCS in PBS and concentrations are listed below.

Table 11: Primary antibodies

Antigen	Antibody	Host	Manufacturer	Dilution
β -actin	anti-human β -actin clone AC-74	mouse	Sigma-Aldrich, Taufkirchen GER	1:2500
Cas9	anti- Cas9 from <i>Streptococcus pyogenes</i> clone 7A9-3A3	mouse	active motif, La Hulpe, BEL	0.5 μ g/ml
ERCC1	anti-human ERCC1 clone: FL-297	rabbit	Santa Cruz Biotechnology, Heidelberg GER	1:250
XPA	anti-human XPA	mouse	Kamiya Biomedical, Seattle, Washington, USA	1 μ g/ml
XPF	anti-human XPF Clone: 3F2/3	mouse	Santa Cruz Biotechnology, Heidelberg GER	1:250
XPG	anti-human XPG A301-484A	rabbit	Bethyl, Montgomery USA	1:1000

Table 12: Secondary antibodies

Antibody	Host	Manufacturer	Dilution
Alexa-Fluor [®] 488 goat anti-rabbit	goat	Dianova, Hamburg GER	1:800
Alexa-Fluor [®] 594 goat anti-mouse	goat	Dianova, Hamburg GER	1:800

2.11 Software and online tools

The following software and online tools were used during this project:

Table 13: List of utilized software and online tools

Label	Supplier/web address	application
Chromas Lite 2.1	Technelysium, South Brisbane AUS	sequence analysis
CorelDraw Graphics Suits X5	Corel Corporation, Ottawa CA, USA	image editing
Double Digest Finder	New England Biolabs, Ipswich MA, USA	restriction digestion
Elensembl	http://www.ensembl.org	genome browser
Endnote X1.01	Thomson Reuters, New York USA	literature management
Glomax [®] Software	Promega, Mannheim GER	luciferase and cell survival assay
ImageJ	National Institutes of Health, Bethesda USA	western blot quantification by densitometry

LabImage 1D	Intas, Goettingen GER	Imaging System, western blot
Ligation Calculator	http://www.insilico.uni-duesseldorf.de/Lig_Input.html	ligation
Microsoft Office	Microsoft, Unterschleissheim GER	Word, PowerPoint, Excel
MultAlin	http://multalin.toulouse.inra.fr/multalin	alignments
Pubmed	http://www.ncbi.nlm.nih.gov/pubmed	literature request
Reverse Complement	http://www.bioinformatics.org/sms/rev_comp.html	<i>in silico</i> sequence translation
GraphPad	GraphPad Software, Inc., La Jolla CA, USA	statistical analysis
T _m calculator	http://www6.appliedbiosystems.com/support/techtools/calc/	primer design
Translation tool	http://web.expasy.org/translate	DNA to protein sequence translation

3 Methods

3.1 Molecular Biology

3.1.1 Ethanol precipitation

High salt sodium acetate (a tenth part of the sample volume) and 96% ethanol (2.5 times sample volume) was added to the sample. After 30min of high speed centrifugation (14 000rpm) the supernatant was discarded and 300µl 70% ethanol was added to wash the pellet followed by 10min 14000rpm of centrifugation. Finally, the remaining ethanol was removed and the pellet was resuspended in water or distinct buffer.

3.1.2 Nucleic acid quantitation

Nucleic acid quantitation by spectrophotometry was used to determine the average concentration of DNA or RNA as well as their purity. Nucleic acids absorb UV light in a specific pattern. The absorbance at 230, 260, and 280nm was measured using the NanoVuePlus[®] according to the Lambert-Beer-Law. The ratio of the absorbance at 260 and 280 nm ($A_{260/280}$) was used to assess the purity in relation to contamination by proteins. For pure DNA a ration around 1.8 and for RNA a ration around 2 should be determined. Absorbance at 230nm was the result of other contaminations. Expected values for $A_{260/230}$ are commonly in the range of 2.0-2.2.

3.1.3 Extraction of nucleic acids

3.1.3.1 Isolation of genomic DNA

Genomic DNA was extracted from cells using the Qiagen QIAamp Blood and Tissue Kit according to manufacturer's specifications.

3.1.3.2 Isolation of mRNA and cDNA synthesis

Total RNA was isolated from MRC5Vi cells, primary WT or XP20BE fibroblasts using the RNeasy Mini Kit by Qiagen. Cells were harvested as described in 3.3.1.1 and disrupted using RLT buffer. Depending on the number of pelleted cells, either 350µl or 600µl RLT buffer was added to the pellet, and further steps were carried out according to the manufacturer's specifications. In order to remove any residual contaminating DNA, DNase treatment was performed. Finally, RNA quantification and its assessment of purity were determined by spectrophotometry using the NanoVuePlus[®] (see 3.1.2) and the RNA was stored at -80°C

until further use. Subsequently, cDNA was synthesized from the isolated RNA with the help of the Revert Aid H Minus First Strand cDNA Synthesis Kit. This kit included a reverse transcriptase, an enzyme that uses RNA as a template to synthesize complementary DNA (cDNA) lacking a ribonuclease H activity and therefore, it does not degrade RNA in RNA-DNA hybrids during synthesis of the first strand cDNA. To synthesize mRNA anchored Oligo (dT)₁₈ primer annealing to the poly-A tail of mRNA molecules was used. In order to generate cDNA, 1µg of total RNA was mixed with 1µl Oligo d(T) primer, *aqua bideest* was added to 12µl and incubated at 70°C for 5min. In the meantime, a premix consisting of reaction buffer, RNase inhibitor and dNTPs was prepared. The premix was added to the reaction mix and incubated at 37°C for 5min. Thereafter, 1µl reverse transcriptase was added and the cDNA was synthesized at 42°C for 60min. The reaction was stopped by incubation at 70°C for 10min and the synthesized cDNA was stored at -20°C until further use.

3.1.3.3 Alkaline lysis plasmid extraction

Alkaline lysis was first described by Birnboim and Doly in 1979 and has, with a few modifications, been the preferred method for plasmid DNA extraction from bacteria ever since. The procedure starts with the growth of the bacterial cell culture harbouring the plasmid under selective conditions (ampicillin in LB-medium). Cells were pelleted by centrifugation to remove them from the growth medium. The pellet was then resuspended in a solution containing Tris, EDTA, glucose, and RNase A. Bivalent cations (Mg^{2+} , Ca^{2+}) are essential for DNase activity and the integrity of the bacterial cell wall. EDTA chelates bivalent cations in the solution preventing DNases from damaging the plasmid and also destabilizes the cell wall. Glucose maintains the osmotic pressure so the cells do not burst and RNase A is included to degrade cellular RNA when the cells are lysed. The lysis buffer contains sodium hydroxide (NaOH) and the detergent SDS. SDS is there to solubilize the cell membrane. NaOH helps to break down the cell wall, but more importantly it disrupts the hydrogen bonding between the DNA bases, converting the double-stranded DNA (dsDNA) in the cell, including the genomic DNA (gDNA) and the plasmid, to single stranded DNA (ssDNA). This denaturation is the central part of the procedure. Addition of potassium acetate (neutralization) decreases the alkalinity of the mixture. Under these conditions the hydrogen bonding between the bases of the single stranded DNA can be re-established, so the ssDNA can re-nature to dsDNA. While it is easy for the small circular plasmid DNA to re-nature it is impossible to properly anneal huge gDNA stretches. As the double-stranded plasmid can dissolve easily in solution, the single stranded genomic DNA, the SDS and the denatured

cellular proteins stick together through hydrophobic interactions to form a white precipitate. The precipitate can easily be separated from the plasmid DNA solution by centrifugation. The last step is to clean up the solution and concentrate the plasmid DNA. This was performed by binding to glass-silica membranes and ethanol precipitation (see 3.1.1).

For isolation of smaller amounts of plasmid DNA 4ml LB medium containing ampicillin (100µg/ml), according to the selection marker of the plasmid were inoculated shaking overnight at 37°C with a single bacterial colony. According to manufacturer's instructions of the NucleoSpin® Plasmid Kit by Machery and Nagel plasmid DNA was isolated. To isolate larger DNA amounts the NucleoBond® Xtra MiDi/Maxi Kit by Machery and Nagel was used according to the user's manual. 100ml LB medium containing ampicillin (100µg/ml) were inoculated with bacteria and incubated shaking at 37°C overnight.

3.1.3.4 Agarose gel electrophoresis (AGE)

For analysis or preparative purposes DNA fragments or plasmids, generated by PCR or restriction digestion were subjected to agarose gel electrophoresis. In an electric field negatively charged DNA fragments move towards the anode ((+)-pole). Depending on fragment size DNA fragments exhibit different mobilities in the agarose gel matrix. Smaller fragments move faster than bigger ones so that they can be separated by size and visualized by fluorescence dyes (e.g. HD Green® Safe DNA Dye) that intercalate into the double helix, and UV-light (254nm). Agarose gels were generated by diluting a suitable amount of agarose powder (e.g. 1%) in 1x TBE buffer and dissolving it by boiling using the microwave. For visualization, a nucleic acid stain (HD Green® Safe DNA Dye) was added (5µl per 100ml agarose). 1x TBE buffer was used as running buffer at a current between 60-80 V. Samples were prepared with 6x loading buffer and compared to a molecular weight size standard. Visualization of fragments was performed using the gel documentation system by Intas.

3.1.3.5 Extraction of DNA from agarose gel

Extraction of specific DNA fragments from agarose gels was performed according to the instructions of the NucleoSpin® Gel extraction and PCR clean up Kit by Machery and Nagel. Previously, DNA fragments were separated by agarose gel electrophoresis (see 3.1.3.4) and the fragment of interest was excised using a scalpel.

3.1.4 Enzymatic manipulation of DNA

3.1.4.1 Polymerase chain reaction (PCR)

Developed in 1983 by Kary Mullis, PCR is a technique to amplify a specific DNA region of a template DNA strand. First template DNA is denatured into single strands. Then a pair of specific oligonucleotides (primer), complementary to the target region anneals to the single stranded DNA creating free 3' hydroxyl ends so that the DNA polymerase can start the synthesis. Finally, strands are elongated. These three steps are repeated 30-40 times leading to an exponential amplification of a single DNA molecule. For cloning, restriction site sequences of specific restriction enzymes can be fused to the 5' ends of the oligonucleotides resulting in a PCR product, which can be inserted into a vector after restriction digestion with the corresponding enzymes. The utilized Phusion DNA polymerase contains a proofreading activity and was used to amplify template DNA for cloning purposes. In colony PCR, single colonies serve as template for the amplification reaction. In this thesis, the Phusion polymerase was used for amplification, because it has a lower error rate as it is very important to generate fragments without mutations. For colony PCR and single clone screening this was not as important, because it was only performed to check the size of the integrated fragments. Therefore, in that case the *Taq* polymerase was used. For standard PCR a GC content of the primers between 40-60% and a melting temperature around 50-60°C was preferred.

Reaction mix:

- 0.5µl Phusion/ 0,6 µl *Taq* polymerase
- 1.5µl DMSO/ 1.5µl DMSO
- 10µl high fidelity (HF) buffer/ 10µl *Taq* buffer + 10µl MgCl₂ (25mM)
- 1µl dNTPs/ 2µl dNTPs
- 2.5µl forward primer/ 4µl forward primer
- 2.5µl reverse primer/ 4µl forward primer
- cDNA template (100-200ng)/ genomic DNA 400ng
- H₂O up to 50µl/ 100µl

Cycler protocol Phusion polymerase:

- 98°C, 30s initial denaturation
 - 98°C, 20s denaturation
 - 50-60°C, 20s primer annealing
 - 72°C, (30sec/kb) elongation
 - 72°C, 5min remaining elongation
- } 35x

Cycler protocol *Taq* polymerase:

- 95°C, 5min
 - 95°C, 30s
 - 50-60°C, 30s
 - 72°C, 1min/kb
 - 72°C, 5min
- } 35x

3.1.4.2 Quantitative Real-time PCR

For quantitative real-time PCR (qRT-PCR), total RNA was extracted from primary human fibroblasts using the RNeasy Mini Kit according to manufacturer's instructions. Afterwards, cDNA was synthesized from 1µg of total RNA using the RevertAid H Minus First Strand cDNA synthesis Kit (see 3.1.3.2). Analysis of gene expression was carried out using EvaGreen-based real-time PCR detection. For each reaction, 50ng of cDNA was used in a final reaction mixture of 10µl, containing 1pmol oligonucleotides and 2µl EvaGreen® qPCR Mix Plus. A passive reference dye (ROX) is included in this mix to normalize the fluorescent reporter signal. Each sample was analyzed in duplicates. Gene expression levels were normalized against the *GAPDH* housekeeping gene. For quantitative analyses and quality control of the applied primer pair standard curves were established. Therefore, the PCR product was amplified from cDNA of WT fibroblasts and serially diluted (500amol – 5ymol, in 1:10 steps). To ensure high quality standard curves only standards with a dissociation curve containing one peak, meaning no unspecific byproducts or primer dimers were applied. For a PCR efficiency of 100% the standard curve was supposed to have a slope of -3.32 and an $R^2 > 0.99$. The non-template control (H₂O) should not be detectable. For the analyses of this thesis slopes between -3.58 and -3.10, corresponding to a PCR efficiency of 90 – 110%, were accepted. Quantitative real-time PCR analyses were performed using the Applied Biosystems 7900HT Fast Real-Time PCR System. The sequences of the PCR primer pairs are shown in Table 7.

Cycling protocol:

- 95°C, 15min initial denaturation
 - 95°C, 15sec denaturation
 - 60°C, 20sec primer annealing
 - 72°C, 20sec elongation
- } 39x
- and a final dissociation curve

3.1.4.3 Site-directed Mutagenesis (Quickchange PCR)

Quickchange PCR is a method of site-directed mutagenesis that is used to make specific and intentional changes to the DNA sequence of a gene or even a plasmid. This procedure requires the synthesis of a DNA primer pair, which contains the desired mutation and is complementary to the template DNA around the mutation site (15-20bp upstream and downstream) in order to hybridize with the DNA at the position of interest. The single-stranded primer is then extended using a DNA polymerase, which copies the rest of the gene or whole plasmid. The same cycle program as described above can be used when adjusting the annealing temperature of the primer and the number of cycles. Subsequently, a DpnI digestion leads to removal of the template plasmid, as DpnI is an enzyme that digests methylated DNA. 1µl enzyme was added to the PCR-mix and incubated at 37°C for 2h. Afterwards, the enzyme was inactivated at 80°C for 20min. The gene or construct thus copied contained the mutated site, and could then be introduced into a host cell by transformation (see 3.2.3). Finally, mutants were selected.

3.1.4.4 Plasmid vector dephosphorylation, oligonucleotide phosphorylation

A plasmid vector dephosphorylation was performed to prevent re-ligation of the vector after restriction digestion. Therefore, 20µg of the plasmid was mixed with 1µl of calf intestinal alkaline phosphatase (CIAP) and 5µl of the respective buffer. The volume was adjusted to 50µl with water. Then the mixture was incubated at 37°C for 1 hour. Enzyme inactivation was performed at 65°C for 15min. Afterwards, ethanol precipitation was performed to concentrate the DNA. Oligonucleotide phosphorylation was performed using the T4 Polynucleotide Kinase from NEB according to manufacturer's instructions.

3.1.4.5 Restriction digestion of DNA

The desired sequence to be cloned into an expression vector (insert) was amplified by PCR (see 3.1.4.1). Restriction enzymes (KpnI/XbaI or HindIII/XhoI) were used to cut as well insert as plasmid vector DNA. Thereby, complementary sticky ends were generated so that insert and vector could be ligated and not vector and vector or insert and insert. Restriction enzymes were chosen according to their presence in the multiple cloning site (MCS) of the respective plasmid. Furthermore, it was important that there were no additional restriction sites in the plasmid or insert DNA. Otherwise this would have led to undesired fragments. A double digestion was performed according to manufacturer's instructions over night at 37°C. Afterwards an ethanol precipitation (see 3.1.1) was performed, to prevent buffer components

from disturbing the following ligation step. For plasmid linearization PvuI was utilized according to manufacturer's instructions.

3.1.4.6 Ligation

DNA ligases are enzymes catalyzing the formation of covalent phosphodiester bonds between free 3' hydroxyl and 5' phosphate ends of DNA molecules, fusing two fragments together. A DNA T4 ligase was used to insert DNA fragments into a plasmid vector according to their sticky ends. A vector insert ratio of 1:3 (mol) according to 100ng of vector was used. Insert, vector, 1µl ligase, 2µl ligase buffer and H₂O (up to 20µl) were incubated for 2h at room temperature or overnight at 16°C.

3.1.4.7 DNA sequencing and sequence analysis

Sequencing reactions according to Sanger's chain-termination method (1977) were performed. This method requires a ssDNA template, a DNA primer, a DNA polymerase, normal deoxynucleotidetriphosphates (dNTPs), and modified nucleotides (dideoxynucleotides) that terminate DNA strand elongation. Today, these ddNTPs (in this case dye-terminating nucleotides) can be fluorescently labelled for detection (dye-terminator sequencing). This permits sequencing in a single reaction as each of the four ddNTPs is labelled differently, emitting light at different wavelengths. In this case, the products were separated by capillary gel electrophoresis by the Paediatrics Department of the University Medical Center Goettingen or send to Eurofins genomics. The resulting chromatogram displays the sequence. For this project the Big Dye Terminator v3.1 Cycle Sequencing Kit was used according to manufacturer's instructions except the amount of BDs.

Sequencing Mix:

- 300-500ng of DNA
- 0.5µl primer
- 1µl BD
- 2µl BD buffer
- H₂O ad 10µl

Cycling protocol:

- 96°C, 30sec
 - 55°C, 15sec
 - 60°C, 4min
- } 26x

The resulting product was cleaned by ethanol precipitation (see 3.1.1) and resuspended in 10µl Hi-Di Formamide for analysis with a 3100-Avant Genetic Analyzer. The Chromas Lite software (Technelysium Pty Ltd, Brisbane Australia) was utilized for analysis of the resulting sequence.

For single clone analysis, genomic DNA was isolated and the exonic region of interest was amplified using specific primers. PCR products were purified using an ExoSAP digestion and sequenced with the BigDye[®] Terminator v3.1 Cycle Sequencing Kit. For allele-specific sequencing, exonic regions were amplified and cloned into the pJET1.2/blunt plasmid according to manufacturer's instruction of the CloneJET PCR Cloning Kit, transformed in DH5α cells and single colonies were sequenced as described above.

3.1.4.8 Enzyme mismatch cleavage

The T7 endonuclease I (T7EI) and surveyor nuclease assay were performed according to the suggested manufacturer's protocol for genome targeting and surveyor mutation detection (NEB, IDT). These enzyme mismatch cleavage (EMC) methods make use of nucleases that cleave heteroduplex DNA at mismatches and extrahelical loops (single or multiple nucleotides). The bacteriophage T7EI cleaves cruciform DNA structures, Holliday structures or junctions, heteroduplex DNA and more slowly, nicked double-stranded DNA. T7EI is proficient at recognizing insertions and deletions of ≥ 2 bases that are generated by NHEJ activity. The surveyor nuclease is a member of the celery endonuclease (CEL) family, which recognizes and cleaves mismatches due to the presence of single nucleotide polymorphisms (SNPs) or small insertions or deletions.

3.1.4.9 CRISPR construct generation

The pX462 vector containing the cDNA encoding *Streptococcus pyogenes* Cas9 (hSpCas9) and a puromycin resistance cassette was purchased from Addgene (plasmid #48141) (Ran *et al.*, 2013). The guide sequence oligonucleotides (Table 7) including the BbsI restriction site overhangs and the PAM sequence targeting exon two within *XPF* were annealed (95°C, 5min, cool down overnight) and cloned into the pX462 vector. The plasmid was digested with BbsI according to manufacturer's instruction. The oligonucleotides were phosphorylated by T4 Kinase (see 3.1.4.4) and ligated into the vector using the T4 Ligase according to manufacturer's instructions. Constructs were transformed in *Escherichia coli* strain DH5α cells and selected using ampicillin (see 3.2.3). To validate the correct sequence the BigDye[®] Terminator v3.1 Cycle Sequencing Kit was utilized (see 3.1.4.7) applying the U6 fwd primer

(Table 7). The guide sequence was preselected from six different XPF guideRNA pairs using *in silico* on- and off-target predictions (<http://crispr.mit.edu/>, <http://portals.broadinstitute.org/gpp/public/analysis-tools/sgrna-design>, <http://crispr.cos.uni-heidelberg.de/index.html>) and T7EI analysis of polyclonal populations in respect of *XPF* gene targeting in MRC5Vi cells.

3.2 Microbiology

3.2.1 Sterilization and autoclavation

Liquids as well as S1-waste were autoclaved at 120°C and 3bar for 20min. Plastic pipette tips were autoclaved at 120°C and 3bar and subsequently dried at 70°C for 4h. Glass pipettes used for cell culture were washed in a dish washer at 40°C. Afterwards, glass pipettes were sealed with cotton wool and sterilized at 180°C for at least 4h.

3.2.2 Generation of chemically competent *Escherichia coli* (*E. coli*) DH5 α

To amplify recombinant plasmid DNA, plasmids have to be inserted into bacterial cells. However, under normal conditions the uptake of foreign DNA by bacterial cells, like *E. coli*, is very low. Hence, bacterial cells are manipulated using physical or chemical methods to change their cell membrane so that foreign DNA can attach and be taken up more easily. Pores are opened due to the high salt conditions. This process is called production of competent cells.

Competent bacterial cells were plated on LB without antibiotics. A single colony was used to inoculate a 10ml overnight culture (LB medium) at 37°C. On the next day the whole culture and 190ml fresh medium were incubated at 37°C until the optical density reached OD₆₀₀= 0.5. Very importantly, from then on cells were kept on ice to keep them in the beginning of the logarithmic growth phase. Cells were centrifuged at 4°C (4000rpm, 10min). Afterwards the pellet was resuspended in 25ml ice-cold 100mM MgCl₂ and incubated on ice for 5min. Cells were centrifuged again and then resuspended in 5ml ice-cold 100mM CaCl₂, incubated on ice for 20min and centrifuged again. Finally, the pellet was resuspended in 1ml (85%CaCl₂, 15% glycerol), aliquoted à 50 μ l, immediately quick-frozen in liquid nitrogen and stored at -80°C.

3.2.3 Transformation of *E. coli* DH5 α

In molecular biology transformation is defined as the genetic alteration of a cell resulting from the direct uptake and its subsequent expression of exogenous genetic material from its

surroundings through the cell membrane. This means uptake of recombinant plasmid-DNA into competent *E. coli* (DH5 α cells, 50 μ l aliquots see 3.2.2). Therefore, 5 μ l of insert DNA was added to one aliquot of competent cells (thawed on ice) and incubated on ice for 30min. This led to the attachment of plasmid DNA to the cell membrane. Then a heat shock was performed (42°C for 45s, directly put the tube on ice afterwards for 2min) to improve the uptake of the foreign DNA. This heat shock was followed by a regeneration of the cells using 450 μ l of liquid LB medium for 1h at 37°C. Subsequently, cells were plated onto an agar plate containing ampicillin. This led to a selection for bacterial cells carrying the plasmid with the resistance gene. 40% transformation efficiency was determined. Transformation efficiency is the number of transformed cells (transformants) generated by 1 μ g of plasmid DNA.

3.3 Cell Biology

3.3.1 Cell culture techniques

3.3.1.1 Culture of primary human fibroblasts and immortalized cell lines

All cell types were cultivated in 75cm³ culture flasks with 20ml DMEM high glucose culture media supplemented with 10% FBS (v/v) and 1% Penicillin and Streptomycin (v/v) in a humidified atmosphere at 37°C and 5% CO₂. Primary fibroblasts were passaged 1:1 or 1:2 when they had grown to confluence, while immortalized cells were passaged 1:10. Therefore, these adherent cells were rinsed with 10ml PBS to remove dead cells and residual FBS, and then dissociated from the culture flask using 4ml trypsin/EDTA and 5min incubation at 37°C. The reaction was stopped by addition of 10ml complete culture medium. Then cells were centrifuged at 1000rpm for 5min and the supernatant was discarded. Afterwards, the cells were distributed to an appropriate number of culture flasks to reach 10 – 50% confluence for immortalized and primary cells, respectively.

3.3.1.2 Cell counting

In order to determine the number of cells/ml, adherent cells were detached with trypsin/EDTA and resuspended in fresh culture medium (complete) as described before (see 3.3.1.1). 10 μ l cell suspension was mixed with 90 μ l trypan blue and 10 μ l of this suspension was applied to a Neubauer counting chamber. Four large squares (each 1mm²) were counted under a microscope and the number of cells was the sum of all the counted cells in all squares counted. The cell number per ml is the sum of all counted cells x 2500 x the dilution factor 10.

3.3.1.3 Freezing and thawing cells

For cryo-preservation of fibroblasts, cells were sedimented at 1000rpm for 5min, the supernatant was discarded and the pellet was resuspended in 500µl culture medium (complete). Briefly, 500µl freezing medium was added and the suspension was transferred to cryo tubes. A cell freezing container was used to ensure a controlled cooling rate (-1°C/min) at -80°C. The following day, the cryo tubes were stored at -150°C in a cryogenic freezer. In order to cultivate frozen cells, cells were thawed under running tap water and transferred to 10ml culture medium (complete) and centrifuged at 1000rpm for 5min. The supernatant was discarded and the pellet was resuspended in an appropriate volume of culture medium and re-seeded in a culture flask. The culture medium was renewed the following day to remove any residual DMSO and dead cells.

3.3.1.4 Transient Transfection

All cell lines were transfected utilizing the Attractene Reagent, a nonliposomal lipid that enables transfection of not only adherent cells. It is able to form complexes with DNA resulting in micelles that are then absorbed by cells via endocytosis. Thus, cells were either seeded in 100mm tissue culture dishes at a density of 9×10^5 for primary or 6×10^5 for immortalized cells per dish, or in 24-well-plates at a density of 20000cells/well (or 15000cells for immortalized cells). On the next day, cells were transfected according to manufacturer's instructions. DNA (4µg/ 250ng) was filled up with 300µl/60µl of non-supplemented cell culture medium, 15µl/1.8µl/0.6µl Attractene Transfection Reagent was added and the mixture was vortexed. This was followed by 15min incubation at room temperature, while the old medium of the dishes or plates was removed and 10ml/500µl fresh medium was added. After that, the plasmid-Attractene-mix was added dropwise and cells were further incubated for 6h until the medium was removed and replaced again. For fast forward transfection, cells were directly transfected after seeding.

3.3.1.5 CRISPR/Cas9 transfection and single clone expansion

MRC5Vi cells were seeded in 100mm tissue culture dishes at a density of 750 000 cells and transfected using Attractene and the fast-forward transfection protocol according to manufacturer's instructions (4µg DNA (pX462 containing the guide sequence, and a puromycin resistance gene, or no DNA for the control), filled up to 300µl DMEM without supplements and 15µl Attractene). On the next day puromycin was added to the dishes and cells were cultured in puromycin containing medium (0.25µg/ml puromycin) until the control

cells had died. For single clone selection, cells were separated using the serial dilution method in a 96-well plate after coating with superfibronectin and cells were further cultured in FibroLife[®] fibroblast medium without puromycin. After two days the plate was evaluated under the microscope (Axiovert A1, Zeiss), single cells were marked, and expanded for two weeks to form colonies. Then, the colonies were transferred into 6-well plates and further expanded for genomic DNA isolation.

3.3.2 Functional assays

3.3.2.1 Host Cell Reactivation Assay (HCR)

In general the HCR can be used to analyze NER repair capacity and complementation capabilities. In this assay a non-replicating reporter gene plasmid is either irradiated with UVC light ($750\text{J}/\text{m}^2$ for primary cells and $1000\text{J}/\text{m}^2$ for immortalized cells), treated with cis-diammineplatinum(II) dichloride (CP) (in a molecular ratio of 1:20 (primary cells) or 1:40 (immortalized cells) (vector:CP)) or with 4,5',8-trimethylpsoralen (TMP) (1:25 (primary cells) or 1:50 (immortalized cells) (vector:TMP)) followed by irradiation with $1\text{ J}/\text{cm}^2$ UVA light. Thereby, different DNA 6-4PPs and CPDs, as well as intrastrand and interstrand crosslinks were generated. After transfection (see 3.3.1.4) of these damaged plasmids into a host cell, enzyme expression can only be detected if the DNA lesions of the transcribed strand of the reporter genes in the plasmid are removed. Hence, activity of the enzyme coded on the damaged reporter gene plasmid correlates with NER activity (see Figure 14). Cells were seeded in 24-well plates the day before transfection at a density of 20000cells/well. On the following day, cells were transiently transfected with either untreated or treated pCMVluc reporter gene plasmid (100ng) coding for firefly luciferase together with pCMVluc reporter gene plasmid (50ng) coding for *Renilla* luciferase for normalization plus complementation plasmid (100ng). The empty expression vector pcDNA3.1(+) was used as a negative complementation control. After 48/72h (immortalized/primary cells) of culturing, cells were washed with PBS and lysed with 80 μl of lysis buffer for 45min at room temperature. Cell lysates were transferred into a 96-well Glomax[™] 96 Microplate and the enzyme expression was measured using the Promega's Dual-Luciferase Reporter Assay System with the Glomax[®] Discover System. Specific substrate solutions for firefly (beetle luciferase) and *Renilla* luciferase (coelenterazine) were added to the 96-well Glomax[™] 96 Microplate. The enzyme activity was measured as relative light units (RLU). The mean value of a triplicate was determined and repair percentage was calculated. To estimate the repair efficacy firefly luciferase RLU were divided by *Renilla* luciferase using the following formula:

$$\text{repair}(\%) = \frac{\text{mean} \left(\frac{\text{treated firefly}}{\text{Renilla}} \right)}{\text{mean} \left(\frac{\text{untreated firefly}}{\text{Renilla}} \right)} \times 100$$

The assay was repeated at least four times in triplicates.

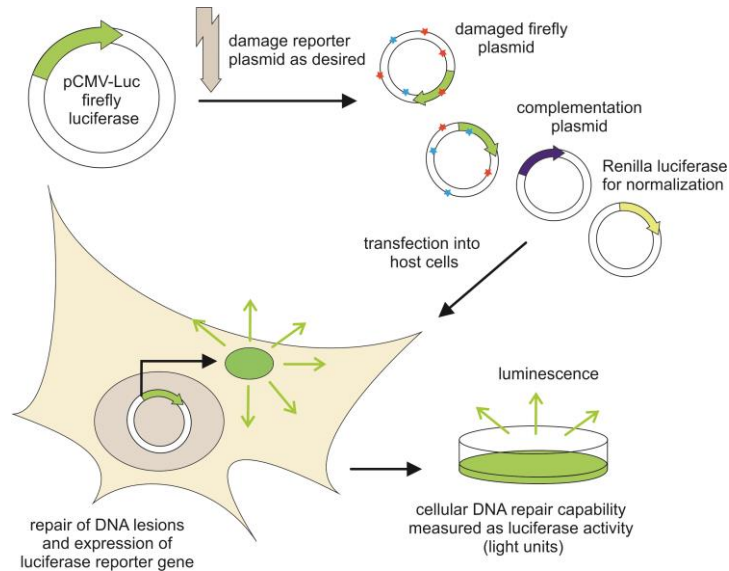


Figure 14: Simplified scheme of the HCR

Cells were transiently transfected with either a treated or untreated firefly reporter gene plasmid together with a *Renilla* reporter gene plasmid for normalization. For complementation, a plasmids coding for, e.g. full-length protein or one of the splice variants, was transfected. After 48/72h of incubation (37°C, 5%CO₂) enzyme activity was measured as relative light units (RLU).

3.3.2.2 DSB Assay

For the measurement of HRR and NHEJ capabilities the DRGFP and pEGFP-Pem1-Ad2 reporter gene assay were utilized (Seluanov *et al.*, 2010). For a schematic illustration of the assay principle see Figure 15. Cells were seeded on glass coverslips in 24-well plates and transfected with 100ng DR-GFP or pEGFP-Pem1-Ad2 plasmid, 100ng pcDNA3.1(+) (used as empty control) or pCBAI-SceI plasmid and 50ng pcDNA3.1(+)mCherry for normalization using Attractene transfection reagent according to the manufacturer's instructions. After 48h cells were washed with PBS three times and fixed with 4% paraformaldehyde (PFA) in PBS for 15min at room temperature. After fixation, the slides were washed again and then stained using Hoechst33342, diluted 1:1000 in PBS, for 20min at room temperature covered in aluminum foil. Thereafter, the slides were washed and mounted with fluoromount fluorescent mounting medium. Results were documented using an Axiovert A1 microscope with a 200x/400x magnification. Subsequently, approximately 100 mCherry positive cells per condition were assessed (blinded) for additional GFP positivity. Hence, the repair capability was calculated as the percentage of GFP positive cells compared to mCherry positive cells.

The assay was repeated at least five times. The plasmids were a kind gift from Prof. Matthias Döbelstein (Department of Molecular Oncology, Georg-August-University, Göttingen).

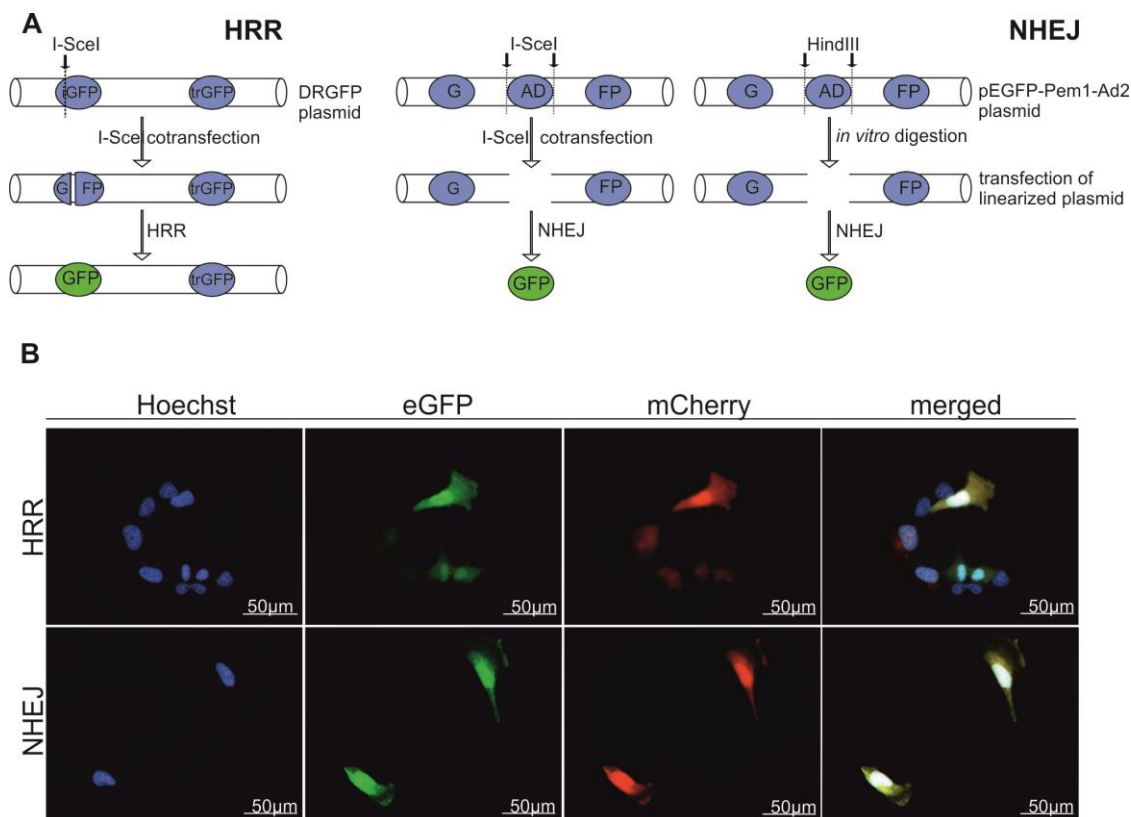


Figure 15: Schematic illustration of the DSB assay principle

(A) The left panel depicts the HRR assay. The GFP plasmid contains two GFP cassettes. One cassette (iGFP) with a deletion in the first exon of GFP combined with the insertion of I-SceI restriction sites is followed by a promoter-less/ATG-less first exon and intron of GFP (trGFP). The DRGFP plasmid was co-transfected with pCBASceI and pcDNA3.1(+)mCherry for internal normalization. Upon expression of I-SceI (from pCBASceI) a DSB is produced in the DRGFP plasmid. If the cells repair the DSB using the trGFP as template, a functional GFP cassette is formed and HRR proficient cells can express GFP, while deficient cells do not. The middle panel shows the way NHEJ was assessed using the pEGFP-Pem1-Ad2 plasmid. It contains a GFP gene with an engineered 3kb intron from the Pem1 gene with an adenoviral (AD) exon flanked by recognition sequences for HindIII and I-SceI to induce DSBs. The intact NHEJ cassette is GFP negative as the adenoviral exon disrupts the GFP ORF. We co-transfected the plasmid with pCBASceI and pcDNA3.1(+)mCherry for normalization. In NHEJ proficient cells, the I-SceI-induced DSBs eliminating the intron are repaired by NHEJ, so that the cell can transcribe and express functional GFP. Furthermore, as an additional analyses of NHEJ, we performed an *in vitro* digestion of the plasmid using HindIII and then transfected the linearized plasmid (NHEJlin) together with pcDNA3.1(+) (to maintain the same DNA ratios) and pcDNA3.1(+)mCherry for normalization (right panel). (B) Representative fluorescent images acquired for the HRR (upper panel) and NHEJ (lower panel) assay in WT cells. For both assays approximately 100 mCherry positive cells per condition were counted (blinded) for GFP positive cells after 48h. Adapted from (Lehmann *et al.*, 2017).

3.3.2.3 Determination of post-toxin cell survival

The colorimetric CellTiter96[®] Non-Radioactive Cell Proliferation Assay, also called MTT-assay, from Promega is a tool to measure cell viability. MTT (3-(4,5-Dimethylthiazole-2-yl)-

2,5-diphenyl-tetrazoliumbromid) is a membrane passing dye, which is metabolized in living cells by mitochondrial dehydrogenases to its blue colored formazan salt (Mosmann, 1983). The insoluble product is dissolved into a colored solution after addition of a solubilization solution (=MTT stop solution), e.g. containing dimethyl sulfoxide or sodium dodecyl sulfate diluted in hydrochloric acid. Absorbance can be measured at 570nm using spectrometry.

For analysis of post-toxin cell survival 2000 cells were seeded in 96-well plates and after 24h irradiated with UVC doses of 0 - 160J/m² (UVC 500 Crosslinker at 254nm), treated with CP 0 - 8µg/ml, TMP 0 - 364.5ng/ml (30min preincubation) followed by irradiation with 1J/cm² UVA light (Biolink BLX UVA crosslinker at 365nm), camptothecin (CPT) 0 - 512nM, etoposide 0 - 16µM or bleomycin 0 - 16µg/ml. After treatment cells were cultured for 48h until the substrate was added. As the compounds were solved in DMSO, the DMSO volume of the maximal toxin amount applied was used as a control and survival was set to 100%. The assay was repeated at least four times in quadruplicates. We analyzed the post-toxin cell survival method to determine the lethal dose 50 (LD₅₀), which indicates the dose of a toxin that kills 50% of the cells. Therefore, this value gives an impression of the cellular sensitivity towards a special toxin.

3.3.2.4 Fluorescence microscopy

To analyze subcellular localization, HeLa cells were seeded on round glass cover slips in 24-well-plates and transfected with the constructs containing C-terminal eGFP-tagged splice variants of *XPF* and *ERCC1* on the next day (see 3.3.1.4). After 48h the wells were washed with PBS three times and then fixed with 4% PFA for 15min followed by three times washing with PBS again. Afterwards, cells were stained using DAPI or Hoechst33342 and mounted with fluoromount (see 3.3.2.2). Pictures were taken using the Microscope Axiovert Imager A1.

3.4 Biochemical methods

3.4.1 Preparation of protein lysates

For whole protein lysates, cells were harvested as described in 3.3.1.1 and resulting cell pellets were washed twice using 10ml PBS and cells were resuspended in PBS containing PMSF and complete proteinase stop. Afterwards, cells were disrupted by rotational freezing in liquid nitrogen followed by thawing on ice for three times. Cells were sedimented by centrifugation (10min, 14 000rpm, 4°C). Then, the supernatant was transferred into a new

reaction tube to get rid of cell debris. Protein lysates were adjusted to the desired concentration, mixed with 3% SDS splitting buffer and boiled at 95°C for 10min. Afterwards they were stored at -20°C.

For preparation of cytosolic and nuclear protein extracts cells were prepared using buffer A and B (see Table 1). Cell pellets were resuspended in an appropriate volume of buffer A and incubated on ice for 15min. A tenth of the volume of 1% Nonidet NP40 in buffer A was added to the suspension and vortexed for 20sec. Afterwards, cells were centrifuged for 10min at 14000U/min. The supernatant was transferred into a new reaction tube (cytosolic fraction) and kept on ice. The nuclear pellet was washed three times with 500µl buffer A. Then 20-50µl buffer B were added and the reaction tubes were put in an ultrasonic bath with ice for 30min. Thereupon, the mixture was pelleted at 14000U/min for 30min. The supernatant was transferred into a new reaction tube (nuclear fraction). Protein concentration was examined with a ready to use Bradford solution (see 3.4.2).

3.4.2 Bradford protein quantification

Protein concentration was examined by Bradford assay (Bradford, 1976) using a ready to use Bradford solution according to the manufacturer's instructions. Protein concentration was photometrically measured (OD₅₉₅) and quantification was determined using a BSA calibration curve (standard serial dilutions, 0-1500µg/ml). Each sample was diluted 1:20 and 1:40 in the respective buffer and 150µl Bradford solution (Roti-Quant 1:5 in *aqua bidest*) and incubated shaking for 5min at room temperature. The OD₅₉₅ was measured with a GloMax[®] Discover system.

3.4.3 Polyacrylamide gel electrophoresis (SDS-PAGE)

With the help of the SDS-Page (polyacrylamide gel electrophoresis) according to Laemmli it is possible to separate proteins by size in an electric field. The binding of the aliphatic tail of SDS to the polypeptide chain imparts an even distribution of charge per unit mass, thereby resulting in a fractionation by approximate size during electrophoresis. The SDS-treated proteins have similar charge-to-mass ratios in an electric field; hence they can be separated according to their molecular weight. SDS is a negatively charged strong detergent agent used to denature native proteins to unfolded, individual polypeptides. It binds to polypeptides in a constant weight ratio of 1.4g SDS/g of polypeptide. The intrinsic charges of polypeptides become negligible and movement in the gel depends on the size. Therefore, protein lysates were mixed with 3% SDS splitting buffer, containing β-mercaptoethanol, and heated for 10min at 95°C. SDS as a detergent, along with the reducing agent (to break down protein-

protein disulphide bonds), disrupts the tertiary structure and coats the protein with an overall negative charge. Separation of proteins via gel electrophoresis was performed using 4–15% Mini-PROTEAN[®] TGX[™] Precast Protein Gels. Gels were loaded with 50-100µg of protein lysates together with a prestained protein marker.

3.4.4 Western blotting

Western blot technique allows protein visualization of one specific protein within a protein mixture by antibody detection. For semi-dry western blotting, the proteins were transferred onto a nitrocellulose membrane after horizontal SDS-PAGE. Blot paper as well as the membrane were pre-equilibrated with transfer buffer. The transfer was carried out using the Bio-Rad Transblot Turbo System at a current of 0.6A and maximal voltage of 12V for 30-45min, depending on the size of the respective proteins. Successful protein transfer was visualized by staining with Ponceau S solution, followed by destaining of the membrane with *aqua bidest*. In order to saturate free protein binding sites, membranes were incubated with blocking buffer for at least 30min according to manufacturer's instructions. Incubation with the respective primary antibody was performed rocking overnight at 4°C (see Table 11). The next day, the membranes were incubated with an appropriate secondary antibody for 45min at room temperature. Membranes were washed four times with washing solution for 5min and rinsed two times with *aqua bidest*. Antibody binding was detected using the WesternBreeze Chemiluminescent Immunodetection system, according to the manufacturer's instructions. Chemiluminescence was developed and quantified using the Chemo Cam Imager 3.2 and the LabImage 1D software. Therefore, each band was normalized to β-actin.

3.4.5 Membrane stripping

Prior to reprobing, membranes were stripped in Roti[®]-Free Stripping Buffer at 56°C for 30min, washed in PBS, and blocked in blocking buffer for 30min at room temperature. For normalization, membranes were reprobed with β-actin antibody as described in 3.4.4.

3.5 Statistics analyses

Statistics analyses were performed using the GraphPad Prism software with at least three independent experiments. Data are presented as the mean ± SEM. The one-tailed, unpaired student's t-test was applied, * $P < 0.05$, *** $P < 0.01$ ***, $P < 0.001$ were considered significant.

4 Results

During the course of my master thesis physiologically occurring splice variants of the *XPG* gene were identified and characterized (see Figure 12A). These variants encoding for the XPG endonuclease were amplified from healthy donor mRNA and cloned into a pcDNA3.1(+) plasmid vector, a pcDNA3.1(-) plasmid vector with a C-terminal myc/HisA2-tag, or a pcDNA3.1(+) plasmid vector with a C-terminal eGFP-tag. Protein expression over time was assessed as well as subcellular localization (see Figure A1, Figure A2). As previously shown, the XPG endonuclease does not travel according to its molecular weight (136kDa) (Aracil *et al.*, 2013; Dunand-Sauthier *et al.*, 2005; Tillhon *et al.*, 2012), due to an acidic patch in the spacer region. Therefore, endogenous full-length XPG presents with a band at 180kDa. However, all XPG isoforms were stably expressed over the observed time course. Assessment of subcellular localization showed that all isoforms localize to the nucleus and could therefore theoretically take part in NER. Notably, two isoforms (IsoV and IsoVI) led to a highly significant increase of cellular repair capability (see Figure 12B).

The goal of my Ph.D. project was to analyze catalytic functions or negative influences of alternative splice variants of the three genes *XPF*, *ERCC1* and *XPG* during NER and ICL removal. I established the CRISPR/Cas9 technique in the laboratory and utilized it to generate *XPF* and *ERCC1* KO cells in order to overcome the limitations of studying the function of the XPF/ERCC1 complex due to the lack of an appropriate human XPF- or ERCC1-defective cell line.

4.1 A novel mutation in the *XPA* gene results in two truncated protein variants and leads to a severe XP/neurological symptoms phenotype

In addition to characterizing and testing spontaneous mRNA splice variants of XP genes, I also worked on the characterization of XP patients and identified a novel mutation in the *XPA* gene. The human *XPA* gene (OMIM: 278700) is located on chromosome nine (9q34.1), encodes for a 273 aa protein, and plays a major role in DNA damage demarcation as well as stabilization of NER factors (reviewed in Berneburg & Lehmann, 2001). In geographically isolated areas, like Japan, the XP frequency is about 1:22,000 with a preference for XP-A cases and commonly occurring founder mutations result in a severe XP plus neurological symptoms phenotype (see Figure 6) (Moriwaki *et al.*, 2012; Satokata *et al.*, 1990).

Two 21 and 28 year old brothers (XP31MA and XP32MA), a 17 year old Moroccan adolescent from a consanguineous family (XP31GO), and a 6 year old boy (XP118MA) were characterized. XP31MA, XP32MA, and XP31GO clinically presented with only mild XP symptoms together with mild mental retardation. On the other hand, XP118MA showed a very severe mental retardation as well as skin cancer development during early childhood. Primary skin fibroblasts of all patients showed strongly reduced post UV-cell survival and high sensitivity against UVC irradiation, assessed by MTT-assay (see Figure 16A). LD₅₀ values were 5 J/m² UVC for all four patients compared to 80 J/m² UVC for WT primary fibroblasts. Cellular NER capacity was analyzed using the HCR assay (see 3.3.2.1). Compared to WT fibroblasts (14.05%), XP-A patient cells exhibited a decreased repair capacity of only 0.28% (XP31MA), 0.24% (XP32MA), 0.14% (XP31GO), and 0.49% (XP118MA) (see Figure 16B). This effect could be rescued by co-transfection of a WT *XPA* containing expression vector (9.24% (XP31MA), 8.73% (XP32MA), 7.46% (XP31GO), and 7.82% (XP118MA)).

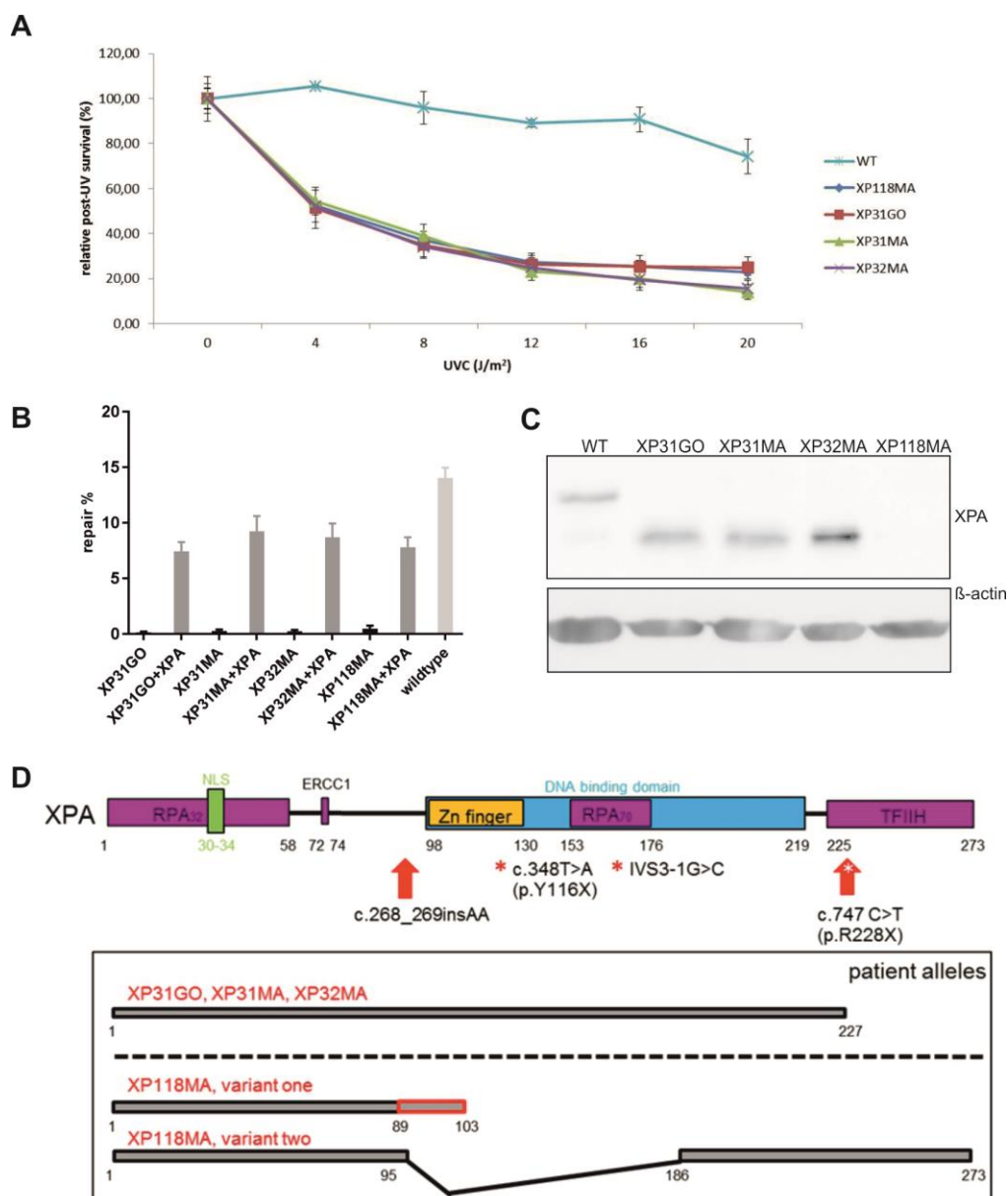


Figure 16: Characterization of four new XP-A patients

(A) Post-UV cell survival of XP-A patient fibroblasts was analyzed after cells were irradiated with increasing UVC doses at least three times in quadruplicates. (B) Reporter gene reactivation after treatment with UVC determined a decrease in repair capability of all four patients (black bars) that could be rescued to wildtype levels (light grey bar) after complementation with an *XPA* encoding plasmid (dark grey bars). Data are presented as mean \pm SEM of at least three independent experiments in triplicates. (C) XPA protein levels were assessed by immunoblotting of whole cell protein extracts from WT and patient primary fibroblasts. In extracts from XP31GO, XP31MA and XP32MA a truncated protein could be detected, while no protein was detectable for XP118MA. (D) Functional domains of the XPA protein (upper part) as well as the mutational spectrum found in the four new XP-A patients (lower part) are shown. Purple boxes: functional protein-protein interaction domains, yellow and blue boxes: DNA interaction domains, and green box: NLS; red arrow: XP-A patient mutation; asterisk: known founder mutations; grey bars: newly characterized patient alleles. Adapted from (Lehmann *et al.*, 2015).

In order to pinpoint the genetic defects, the entire *XPA* gene including exon-intron boundaries was sequenced on the genomic and cDNA level (see 3.1.3.1, 3.1.3.2, and 3.1.4.7). XP31MA, XP32MA, and XP31GO showed the already known founder mutation c.747C>T, a homozygous C>T transversion leading to a premature stop in exon six (p.R228X). In XP118MA, I identified a homozygous two nucleotide insertion mutation in exon two (c.268_269insAA) leading to a frameshift and a termination signal 40 nts downstream, resulting in a truncated protein of 103 aa with 14 non-XPA amino acids at the C-terminus (variant one) (see Figure 16D). Additionally, a second variant of this mutation could be identified. Due to in-frame skipping of exons three-four, it is missing aa 96-185 and containing a stretch of six foreign amino acids. Patient fibroblasts of XP31MA, XP32MA, and XP31GO exhibited a detectable endogenous expression of the truncated XPA protein as assessed by western blot analyses (see 3.4.4). The two variants of XP118MA could not be detected on protein level (see Figure 16C). However, mRNA decay was excluded as cell lines XP31MA, XP32MA, and XP31GO exhibited 83%, 140% and 45% *XPA* expression and even 172% in XP118MA mRNA compared to WT (100%) (data not shown) (Lehmann *et al.*, 2015).

4.2 Establishment of a complete *XPF* and *ERCC1* CRISPR/Cas9 knockout in MRC5Vi cells

To elucidate the catalytic or inhibitory function of *XPF*, *ERCC1* and *XPG* alternative splice variants during NER and ICL removal, establishing a suitable cell line was of essential importance. As described before (see 1.4.3), patient cells from XP20BE represent a good basis for mechanistic studies of *XPG* and its isoforms. In contrast, a key limitation to study the functions of the *XPF/ERCC1* complex has been the lack of an appropriate human *XPF*- or *ERCC1*-defective cell line. This is due to the essentialness of these genes and relatively mild phenotypes in XP-F patients or in the case of *ERCC1* the overall limited number of known patients. To generate targeted *XPF* and *ERCC1* KO cell lines with early mutations, the CRISPR/Cas9 technique is a great tool to study the truncated splice variants of these genes.

4.2.1 Characterization of an *XPF* CRISPR/Cas9 KO single clone

Fibroblasts from XP-F patients still retain relatively high repair activities which are reflected predominantly in the mild phenotype due to residual repair capabilities in full-length alleles with point mutations (reviewed in Matsumura *et al.*, 1998). The aim was to create an *XPF* KO cell line from immortalized fetal lung fibroblasts (MRC5Vi), already used in DNA repair

studies (Ogi *et al.*, 2010; Sertic *et al.*, 2011), by employing the CRISPR/Cas9 nuclease system targeting one of the first exons (one-three) of *XPF*. The feasibility for a complete *XPF* KO in human fibroblasts was assumed based on a previously established *Xpf*-deficient mouse model (Tian *et al.*, 2004).

A transient transfection approach with a puromycin resistance cassette containing a plasmid for selection of positive clones was applied to circumvent long-term expression of artificial nucleases in the cells in order to minimize off-target effects (see 3.1.4.9 and 3.3.1.4). Different guide RNAs were analyzed in respect to on and off-target effects to identify suitable target sequences in exon two of *XPF*. Single clone expansion was followed by amplifying and sequencing exon two to verify the successful KO (see 3.1.4.7). At this point, cells were sensitive to puromycin again. Figure 17A shows a part of the *XPF* exon two sequencing results from a positive clone, visualizing the generation of a compound heterozygous, complete, and still viable *XPF* KO cell line. In this section, the binding of the guide RNA (black arrow) in *XPF* exon two is indicated. As expected the Cas9 nuclease induces a DSB ~3bp upstream of the PAM sequence (Ran *et al.*, 2013). Part of the guide sequence was deleted during the NHEJ process. One allele contains a seven nt deletion resulting in a premature stop codon nine nts downstream. In addition to this, the second allele shows an insertion of three nts and a premature stop after twelve nts, respectively. This is probably a result of a second cutting event by the Cas9 nuclease at the same position. Complete sequencing of the *XPF* gene did not reveal any additional mutations.

Furthermore, the *XPF* KO clone containing the mutation, was subject to standard enzyme mismatch cleavage experiments (EML), the T7EI and surveyor nuclease assay (see 3.1.4.8). For the T7EI assay, PCR products were generated from genomic DNA of WT and *XPF* KO MRC5Vi cells, using primers for *XPF* exon two as described above (see 3.1.4.1) (see Table 7). Subsequently, PCR products were denatured, followed by re-annealing, leading to a population of double strand fragments, of which some contained mismatches due to the afore mentioned compound heterozygous nature of the KO. These mismatches could be detected by T7EI and visualized on an agarose gel (see 3.1.3.4). WT cells did not yield amplicons susceptible to site-specific cleavage by T7EI (see Figure 17B), while *XPF* KO cells showed efficient cleavage of the endogenous chromosomal target sequence, evidenced by T7EI digestion products (arrows). This confirmed the presence of indels at the locus. Surprisingly, there was an additional cleavage product, also in the WT cells (arrow head) that can be explained by a physiologically appearing already described three nt deletion in exon

two (polymorphism rs771203473 (NCBI), deletion AGA/-, chromosome 16:13922057-13922059). PCR amplicons from mutant (KO) and WT DNA were likewise prepared for the surveyor nuclease assay. Afterwards, equal amounts of KO and WT DNA were mixed and hybridized by heating and cooling the mixture to form hetero- and homoduplexes (see 3.1.4.8). In accordance with the T7EI assay, the WT cells did not show cleavage products, whereas in the WT+KO mixture, as well as in the KO alone, cleavage products could be detected. This was in concurrence with the compound heterozygous state of the KO (see Figure 17C, arrows). As observed in the assay before, an additional cleavage product was detected (arrow head).

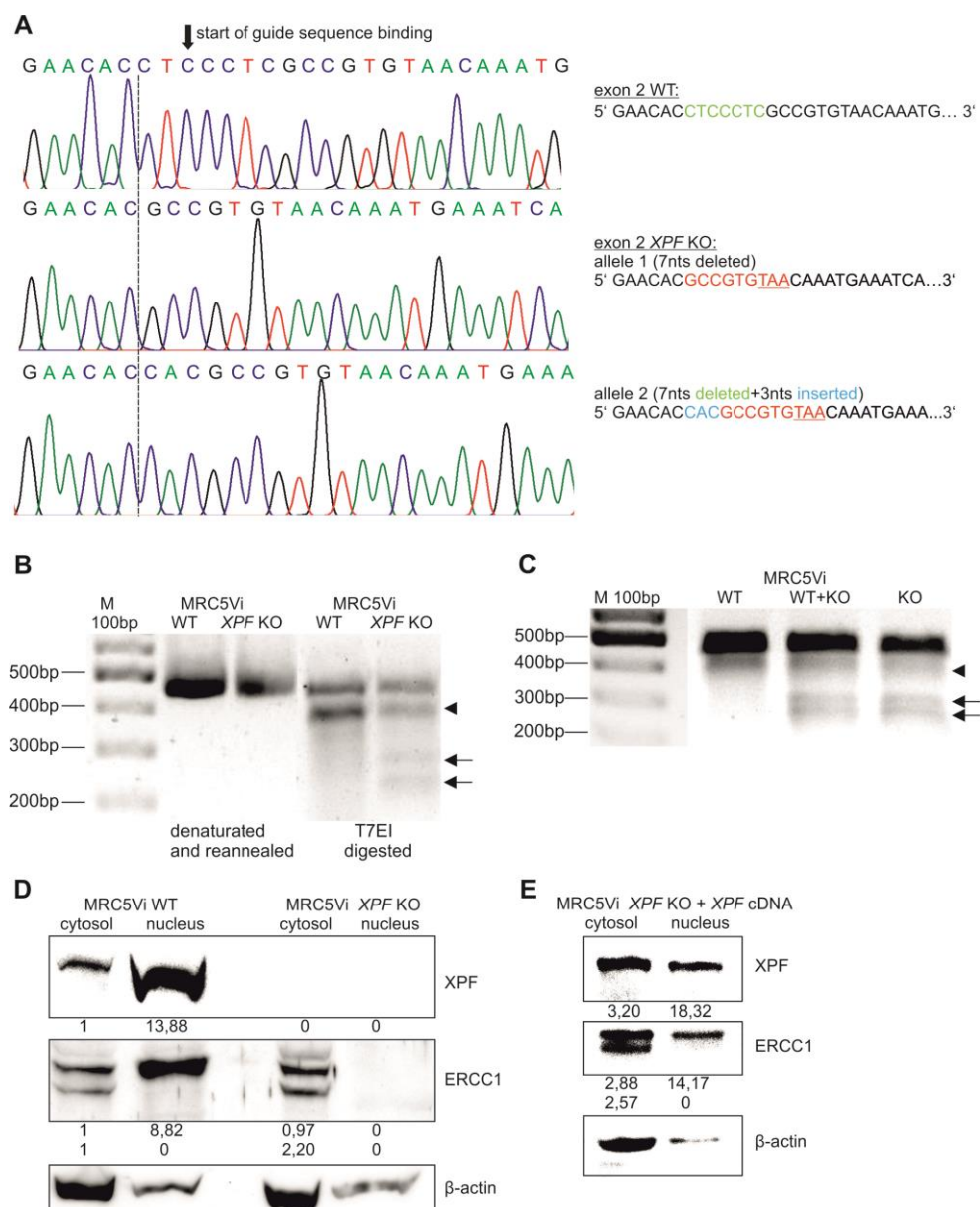


Figure 17: Structural analyses of XPF CRISPR/Cas9 KO and WT MRC5Vi cells

(A) Sanger sequencing of *XPF* exon two from WT (upper panel) and KO (middle and lower panel) cells. The start of the guide RNA binding sequence is illustrated by a black arrow, while the dashed line marks the start of the sequence alterations in the *XPF* KO cells. The green sequence represents the deletion in the KO cells, while the blue nts are inserted. Red nts indicate the continuation of the WT sequence after the CRISPR event up until the stop codon (underlined), resulting from the frameshift. Two different alleles of the KO cells originating from the CRISPR/Cas9 genome editing are depicted in the middle and lower panel. (B) *XPF* exon two was PCR amplified from genomic DNA of WT and KO cells, denatured, reannealed and T7EI-digested or (C) subject to a surveyor nuclease assay. The positions of digested DNA are illustrated by black arrows, while the arrow head marks an additional digestion product that can be seen in the WT cells as well (polymorphism). (D) Horizontal SDS-PAGE followed by western blot analyses was used to assess protein expression. Therefore, equal amounts of cytosolic and nuclear extracts of WT and KO cells were loaded onto ready-to-use SDS-gels and stained with an anti-XPF or anti-ERCC1 antibody. The membrane was stripped and reprobed for anti- β -actin to serve as loading control. Representatively, one of three independent experiments is shown. (E) Subsequent to complementation of the *XPF* KO cells with a plasmid coding for *XPF* cDNA, the same procedure was performed. Adapted from (Lehmann *et al.*, 2017).

Moreover, generation of a complete *XPF* KO on protein level was confirmed. Therefore, cytosolic and nuclear protein fractions of WT and KO cells were isolated (see 3.4.1). Full-length XPF molecular weight is predicted to be 112kDa. The KO induced a premature stop codon after 86/87 aa, resulting in a theoretic molecular weight of approximately 10kDa. Protein expression was assessed using horizontal SDS-Page followed by western blotting as described in 3.4.3, 3.4.4, and 3.4.5. WT cells showed an accumulation of nuclear XPF (13.88), while in *XPF* KO cells XPF was below limits of detection in the cytosol as well as in the nucleus, verifying the total lack of the protein (see Figure 17D). This was also true for lower sections of the gel, where a 10kDa protein would be visible (section of the gel not shown). After overexpression of a plasmid containing *XPF* cDNA (see 3.3.1.4), XPF protein expression could be restored in the cytosol (3.20) as well as in the nucleus (18.32) of *XPF* KO cells (see Figure 17E). In addition, protein expression of ERCC1, XPF's functional heterodimer partner necessary for the catalytic activity in DNA repair, was assessed. In the cytosol, ERCC1 presented as a double band at 38kDa, which is probably caused by post-translational modifications (Nowotny & Gaur, 2016; Perez-Oliva *et al.*, 2015). While WT cells displayed an accumulation of the modified form of ERCC1 in the nucleus (8.82) (Figure 17D), ERCC1 could not be detected in the nucleus of *XPF* KO cells. This suggests an inability of ERCC1 to enter the nucleus without its binding partner XPF. Transfection of *XPF* KO cells with a plasmid containing *XPF* cDNA (see 3.3.1.4) rescued nuclear localization of ERCC1 (see Figure 17E).

Quantitative real-time PCR was applied to analyze differences in *XPF* mRNA expression between WT and KO cells (see 3.1.3.2 and 3.1.4.2). Interestingly, there was a reduction of about 50% in the KO cells (WT *XPF* normalized to *GAPDH* $100 \pm 4.71\%$, KO $54.76 \pm 13.32\%$), probably due to nonsense-mediated messenger decay (see Figure 18) (reviewed in Chang *et al.*, 2007).

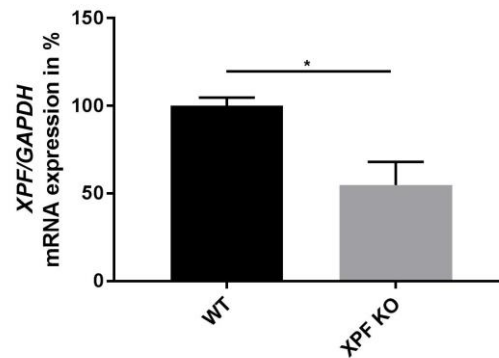


Figure 18: Decrease in *XPF* expression in *XPF* KO cells assessed by quantitative real-time PCR

Comparison of *XPF* mRNA expression levels from MRC5Vi WT and *XPF* KO cells. Expression levels are shown as relative values and *GAPDH* was used as the internal control gene to normalize the expression of the target gene. Results represent mean values from three independent experiments in duplicates. Data are presented as mean \pm SEM. Statistical significances were calculated by applying the one-tailed, unpaired student's t-test (* $P < 0.05$).

4.2.2 A viable complete *ERCC1* KO cannot be generated

As described for *XPF*, a transient transfection approach with a puromycin resistance cassette containing plasmid for selection of positive clones, was applied for *ERCC1* as well (see 3.1.4.9 and 3.3.1.4). In contrast to *Xpf*, several different mouse models with a disruption of the C-terminal part of *Ercc1* have been described. In one model, the last four exons are missing due to a mutation in exon five, disrupting the interaction between XPF and ERCC1. Whereas the other model contains a truncation in the HhH₂ motif, as a result of an insertion of a neomycin cassette into exon seven of the *Ercc1* gene (McWhir *et al.*, 1993; Weeda *et al.*, 1997) (reviewed in Gregg *et al.*, 2011). This implicates an essential role of ERCC1, as complete KO mice are not available, and challenges the creation of a complete cellular KO.

It was possible to generate a heterozygous *ERCC1* KO with a one nt insertion in exon two. This led to a premature stop codon only after about 400nts later, resulting in a residual protein of 21kDa (see Figure 19, lower band) additional to the WT protein (double band). As expected for a recessive disease, one functional allele was sufficient to constitute normal NER levels (data not shown). However, additional attempts to perform further CRISPR/Cas9 genome editing after generation of Cas9 stable MRC5Vi WT cells, as well as the aforementioned heterozygous *ERCC1* KO cells, only yielded further heterozygous single clones, but no full KO in the first exons of *ERCC1*. In regard to these results, it is most likely that homozygous complete *ERCC1* KO cells are not viable due to the essential role of the protein in many cellular processes. Therefore, exons further downstream (three-ten) should be

targeted. Anyhow, such KO cells would not be of value to analyze the splice variants shown in Figure 13 as they comprise structural changes already in early parts of the protein sequence.

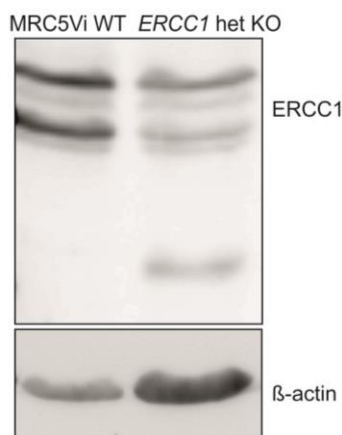


Figure 19: Generation of a heterozygous *ERCC1* KO via CRISPR/Cas9 genome editing

To assess ERCC1 protein expression, horizontal SDS-PAGE was followed by western blot analyses. Equal amounts of whole protein lysates from MRC5Vi WT and heterozygous *ERCC1* KO cells were loaded onto a ready-to-use SDS-gel and stained with an anti-ERCC1 antibody. The membrane was stripped and reprobred for anti-β-actin to serve as loading control. One of three independent experiments is shown.

4.2.3 *XPF* KO cells show an increased sensitivity to several DNA damaging toxins

The newly generated *XPF* KO cells were further characterized by investigation of cell survival after treatment with UVC irradiation, the main trigger of NER, as well as treatment with ICL inducing toxins (CP and TMP plus UVA irradiation) (see 3.3.2.3). Cisplatin is a frequently used chemotherapeutic and predominantly produces intrastrand cross-links between guanine residues or adenine and guanine (90%), which in turn are also mostly repaired by the NER pathway. Interstrand cross-links are only formed between guanines to a minor extent of 2–5% (Jones *et al.*, 1991). Psoralen and its derivatives, for example TMP, mainly form interstrand cross-links with the DNA after activation by irradiation with long-wavelength UV light (UVA) (McHugh *et al.*, 2001). The MTT method was applied to analyze post-toxin cell survival (Mosmann, 1983) and determine the LD₅₀, as a means to quantify cellular sensitivity towards a special toxin. *XPF* KO cells were significantly more sensitive to UVC irradiation in comparison to WT cells (LD₅₀WT = 50J/m², LD₅₀KO < 1J/m²) (***) $P < 0.001$, $n = 4$, one-tailed, unpaired student's t-test) (see Figure 20A). Similarly, treatment with intra- and interstrand crosslink inducing agents like CP or TMP activated by UVA irradiation provoked a significantly reduced survival of *XPF* KO cells (CP: LD₅₀WT = 1.5μg/ml, LD₅₀KO = 0.125μg/ml; TMP: LD₅₀WT = 13.5ng/ml, LD₅₀KO =

0.5ng/ml) (***) $P < 0.001$, $n = 4$) (see Figure 20B, C). Cell survival was not markedly affected by treatment with TMP or UVA alone (data not shown).

Moreover, the XPF/ERCC1 complex is involved in different sub-pathways of mammalian DSB repair, therefore effects of different DSB creating toxins on cellular survival was subject to further analyses (Ahmad *et al.*, 2008; McVey & Lee, 2008). Even though *XPF* KO cells were slightly more sensitive to bleomycin treatment, no significant reduction in survival in comparison to WT cells could be measured ($LD_{50}WT = 3\mu\text{g/ml}$, $LD_{50}KO = 2.5\mu\text{g/ml}$) (n.s. $P > 0.05$, $n = 4$) (see Figure 20D). Besides, the effects of CPT, a topoisomerase I (TOP1) inhibitor, as well as etoposide, a topoisomerase II (TOP2) inhibitor, were analyzed. In concurrence with the previously tested toxin, only a slight, but non-significant, decrease in cell survival was measurable after drug treatment (CPT: $LD_{50}WT = 50\text{nM}$, $LD_{50}KO = 50\text{nM}$; etoposide: $LD_{50}WT = 6\mu\text{M}$, $LD_{50}KO = 3.5\mu\text{M}$) (n.s. $P > 0.05$, $n = 4$) (see Figure 20E, F). Taken together, the results indicate that XPF is not as essential for coping with DSB inducing drugs, as it is for protecting cells against UVC irradiation or ICL inducing agents.

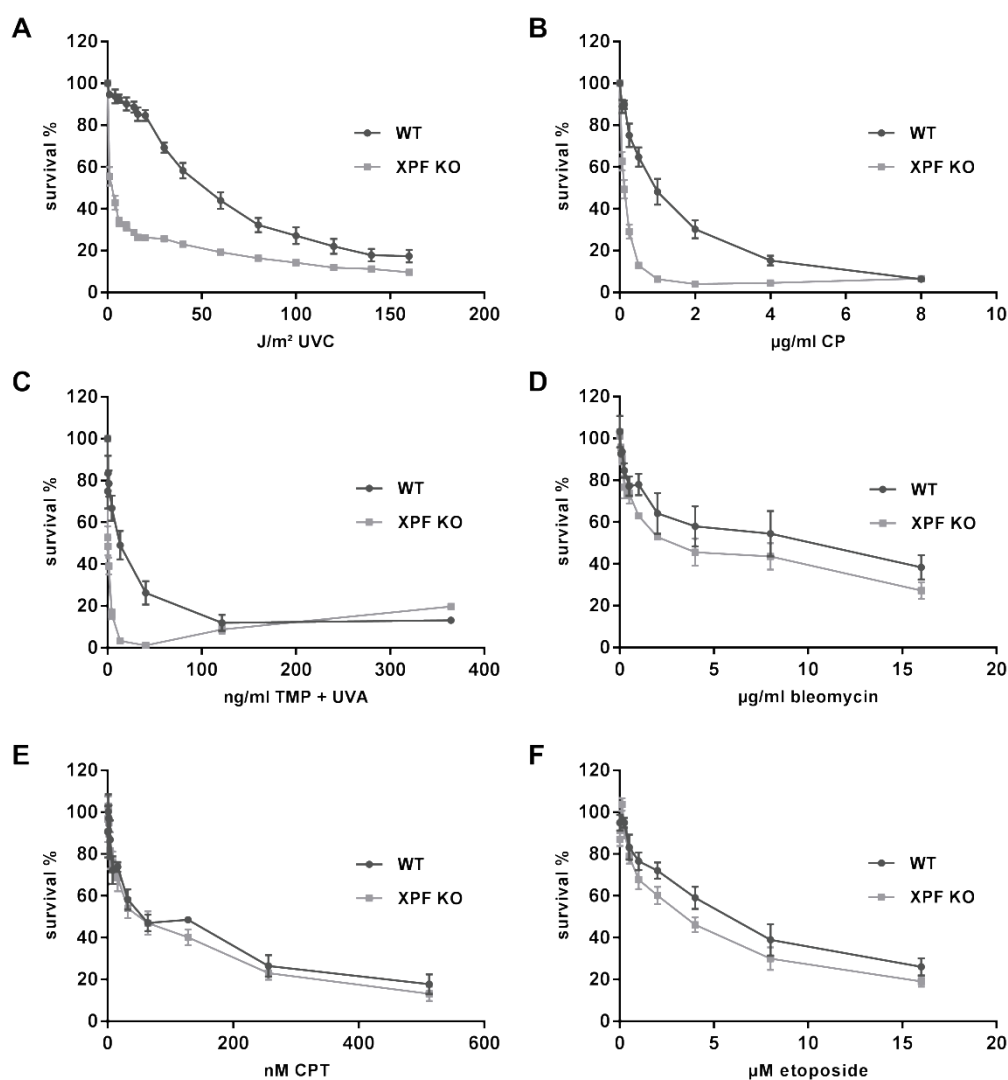


Figure 20: WT MRC5Vi and *XPF* KO post-toxin cell survival analyses

Cells were seeded in 96-well plates at a density of 2000 cells/well and (A) irradiated with increasing doses of UVC (0 – 160J/m²), (B) treated with increasing doses of cisplatin (0 - 8µg/ml), (C) trimethylpsoralen (0 - 364.5ng/ml) activated by 1J/cm² UVA light, (D) bleomycin (0 - 16µg/ml), (E) camptothecin (0 - 512nM), or (F) etoposide (0 - 16µM). After 48h, cell survival was assessed using the MTT assay, setting survival of the DMSO control cells to 100%. Data are presented as the mean \pm SEM. At least four independent experiments in quadruplicates were performed. Adapted from (Lehmann *et al.*, 2017).

4.2.4 Loss of XPF reduces the cellular repair capability for NER, ICL, and HRR

As the XPF/ERCC1 complex is involved in multiple repair pathways and cellular processes, effects of the *XPF* KO on NER, ICL, and DSB repair were studied in more detail by utilizing reporter gene assays. Therefore, the HCR assay was utilized to investigate correct repair of UVC induced DNA lesions (see 3.3.2.1). Expectedly, *XPF* KO MRC5Vi cells exhibited a significant reduction in NER capability ($0.15 \pm 0.02\%$) in comparison to WT cells ($12.45 \pm 0.86\%$) (***) ($P < 0.001$, $n = 4$). Complementation with a plasmid containing the

cDNA of full-length *XPF* rescued this effect ($6.87 \pm 0.13\%$) (** $P < 0.001$, $n = 4$), but neither *XPG* ($0.35 \pm 0.01\%$) nor *ERCC1* ($0.13 \pm 0.21\%$) used as controls (n.s. $P > 0.05$, $n = 4$) (see Figure 21A). ICLs are more challenging to repair as they obstruct DNA replication and transcription (Hodkinson *et al.*, 2014; Klein Douwel *et al.*, 2014). Moreover, it is known that *XPF*-deficient cells are sensitive to agents inducing ICLs, e.g. CP or TMP plus UVA light (reviewed in Wood, 2010). To further investigate this, the HCR assay was adapted for plasmids treated with CP or TMP activated by UVA irradiation (see 3.3.2.1). In the *XPF* KO cells, the repair of intrastrand crosslinks caused by CP, as well as interstrand crosslinks produced by TMP treatment of the firefly luciferase plasmid followed by irradiation with UVA, was impaired, shown by a significant reduction in ICL repair capability (CP $0.78 \pm 0.08\%$, TMP + UVA $2.50 \pm 0.15\%$) compared to WT cells (CP $15.43 \pm 0.035\%$, TMP + UVA $13.90 \pm 0.81\%$) (** $P < 0.001$, $n = 4$). This effect was rescued by co-transfection of a full-length *XPF* containing plasmid (CP $17.65 \pm 1.13\%$, TMP + UVA $12.65 \pm 0.37\%$) (** $P < 0.001$, $n = 4$), but neither by *XPG* (CP $1.34 \pm 0.02\%$, TMP + UVA $2.65 \pm 0.07\%$) nor *ERCC1* (CP $0.56 \pm 0.51\%$ TMP + UVA $2.04 \pm 0.22\%$) (n.s. $P > 0.05$, $n = 4$) (see Figure 21B, C).

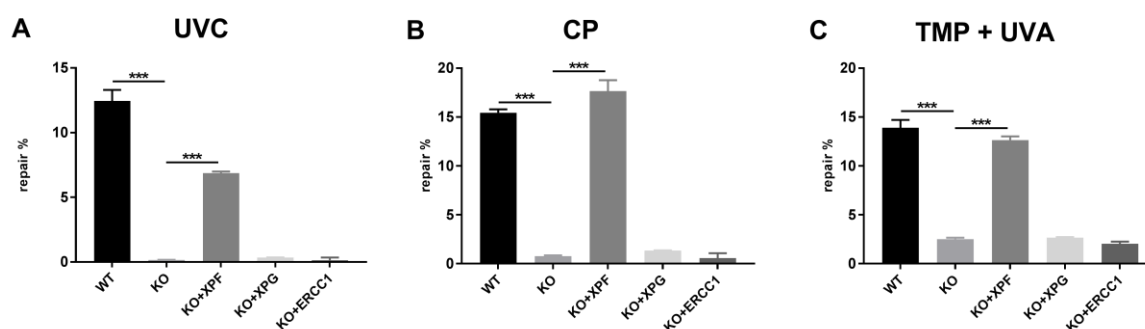


Figure 21: Reactivation of a reporter gene after treatment with UVC, cisplatin or trimethylpsoralen activated by UVA light in *XPF* KO and WT MRC5Vi cells

For the HCR assays, firefly plasmids were treated with (A) UVC irradiation (B) cisplatin (intrastrand crosslinks) or (C) trimethylpsoralen activated by $1\text{J}/\text{cm}^2$ UVA irradiation (interstrand crosslinks) to induce specific lesions, transfected into MRC5Vi WT and *XPF* KO cells, and complemented with plasmids coding for *XPF*, *XPG* or *ERCC1*. The relative repair capability was calculated as the percentage (repair %) of the reporter gene activity (firefly luciferase) compared to the untreated plasmid, after normalization to an internal co-transfected control (*Renilla* luciferase). Data are presented as the mean \pm SEM. The one-tailed, unpaired student's t-test was applied, ** $P < 0.001$. At least four independent experiments in triplicates were performed. Adapted from (Lehmann *et al.*, 2017).

To investigate DNA DSB repair, two eGFP based reporter assays were utilized. The DRGFP reporter assay was applied to investigate HRR capability, while a reporter gene assay using the pEGFP-Pem1-Ad2 plasmid was implemented to assess NHEJ (see Figure 15A, B)

(Seluanov *et al.*, 2010) (see 3.3.2.2). Compared to WT cells ($33.27 \pm 5.52\%$), *XPF* KO cells showed a significant reduction in HRR capability ($13.31 \pm 3.07\%$) (* $P < 0.05$, $n = 5$) (see Figure 22A). On the other hand, no reduction in NHEJ activity, neither for the I-SceI (WT $98.29 \pm 0.36\%$, KO $94.49 \pm 2.70\%$) nor the *in vitro* HindIII digestion, used as an additional control, could be detected (WT $92.25 \pm 3.81\%$, KO $93.94 \pm 1.12\%$) (n.s. $P > 0.05$, $n = 4$) (see Figure 22B).

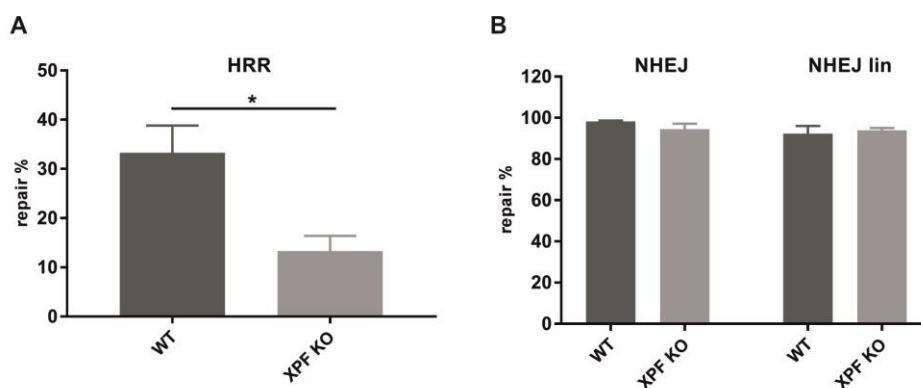


Figure 22: Analyses of HRR und NHEJ repair pathways in *XPF* KO and WT MRC5Vi cells.

For both assays, a pCBASceI and a pcDNA3.1(+)-mCherry plasmid for normalization were co-transfected with the eGFP reporter plasmid. Furthermore, as an additional analysis of NHEJ, an *in vitro* digestion of the plasmid using HindIII was performed prior to transfection. (NHEJlin). Approximately 100 mCherry positive cells per condition were counted (blinded) for GFP positive cells after (A) HRR or (B) NHEJ assay. The repair capability was calculated as percentage (repair %) of GFP compared to mCherry positive cells. Data are presented as mean \pm SEM. The one-tailed, unpaired student's t-test was applied, * $P < 0.05$. At least five independent experiments were performed. Adapted from (Lehmann *et al.*, 2017).

4.3 Amplification and characterization of *XPF* and *ERCC1* splice variants

As previously mentioned, the expression level of splice variants in different tissues is more suitable to distinguish between oncogene and non-oncogene samples than the primary gene transcript (Zhang *et al.*, 2013b) and could be predictive of cancer risk, disease progression and therapeutic success (see 1.6.2). However, not much is known about the function of spontaneous splice variants of XP genes. Therefore, part of my PhD thesis was to further characterize these variants. All until now, physiologically occurring spontaneous mRNA splice variants (see Figure 13) of the *XPF* (*XPF*-201, *XPF*-003, *XPF*-202) and *ERCC1* (*ERCC1*-002 – *ERCC1*-008, *ERCC1*-11, *ERCC1*-013) genes were amplified from WT MRC5Vi mRNA and cloned into different expression plasmids: pcDNA3.1(+), pcDNA3.1(-)-mycHisA2, and pcDNA3.1(+)-eGFP (C-terminal). In order to amplify these splice variants, mRNA was isolated from WT MRC5Vi and transcribed into cDNA (see

3.1.3.2). Specifically designed primer pairs were utilized for amplification of the different splice variants and are listed in Table 7. The cloning procedure was conducted as described in my master thesis for *XPG* and its splice variants (Lehmann, 2013).

4.3.1 Protein expression and subcellular localization of *XPF* and *ERCC1* splice variants

To verify whether the generated plasmid vectors containing the different *XPF* and *ERCC1* mRNA splice variants result in stable expression of a protein, horizontal SDS Page (see 3.4.3) followed by immunoblotting was performed (see 3.4.4). For this purpose, HeLa cells were transiently transfected (see 3.3.1.4) with the different constructs and harvested after 24h, 48h and 72h. Protein extracts were isolated from total cell lysate (see 3.4.1) and separated by SDS-Page. For immunoblot analyses anti-XPF, anti-ERCC1, or an antibody directed against the myc-tag were used (see Table 11). Myc-tagged constructs were used in case of severely truncated variants, probably lacking the antibody epitope. Expression was normalized to anti- β -actin as a housekeeping gene and quantified by densitometric analysis using the LabImage 1D software from Intas (Gassmann *et al.*, 2009).

Figure 23 shows expression patterns of full-length XPF and its isoforms. Wildtype XPF levels peak at 72h, while its isoform XPF-201 shows highest protein levels after 48h. On the other hand, protein levels of XPF-202 and XPF-003 decrease over time. Taken together, all *XPF* isoforms were successfully overexpressed over time and resulted in proteins of the expected size (XPF and XPF-201 ~ 112kDa, XPF-202 and XPF-003 ~ 48kDa). The same procedure was performed for ERCC1 and its respective splice variants (see Figure A3). Again, all isoforms were detectable over time. As mentioned before, ERCC1 presents as a double band due to posttranslational modifications (see 4.2.2). Notably, ERCC1-002 and ERCC1-008 levels increase over time, while full-length ERCC1, ERCC1-003, ERCC1-005, ERCC1-007 and ERCC1-013 decrease.

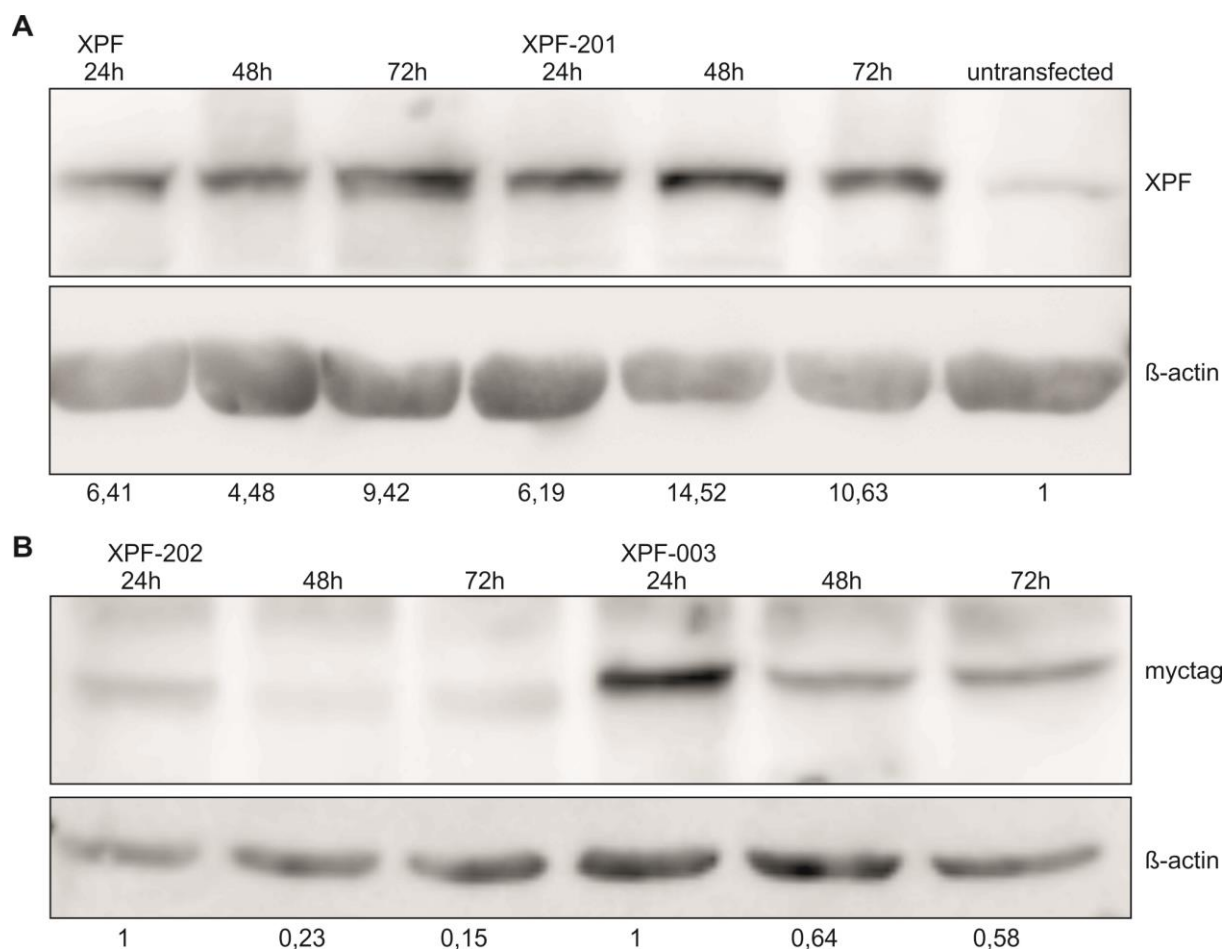


Figure 23: Immunoblot results for protein levels of *XPF* splice variants over time

In order to analyze protein levels after overexpression of the splice variants in HeLa cells, horizontal SDS Page followed by immunoblotting was performed. Therefore, HeLa cells were transiently transfected with the different constructs, harvested after 24h, 48h and 72h, and stained with an anti-XPF or an antibody directed against the myc-tag. Anti- β -actin staining was used for normalization. Protein levels were quantified in regard to untransfected control cells for full-length XPF and XPF-201, or in the case of XPF-202 and XPF-003 to the 24h value as there is no endogenous myc-tag protein. One of three representative experiments is shown.

Figure 13 illustrates an overview of all *XPF* and *ERCC1* splice variants and gives an impression on which domains are lost due to truncating mutations. Subcellular localization of the proteins was of special interest, as DNA repair takes place in the nucleus and requires active or passive nuclear import of the respective proteins. This was analyzed by cloning the isoforms into an eGFP expression vector. Thereby, proteins with a C-terminal eGFP tag were overexpressed in HeLa cells and subcellular localization was studied by fluorescence microscopy (see 3.3.1.4 and 3.3.2.4). 48h after overexpression, all isoforms as well as the WT protein, localized in the nucleus, whereas XPF-202 showed an equal distribution throughout the cell (see Figure 24 and Figure A4). Therefore, all isoforms could theoretically take part in different DNA repair pathways if they had residual enzymatic or structural function.

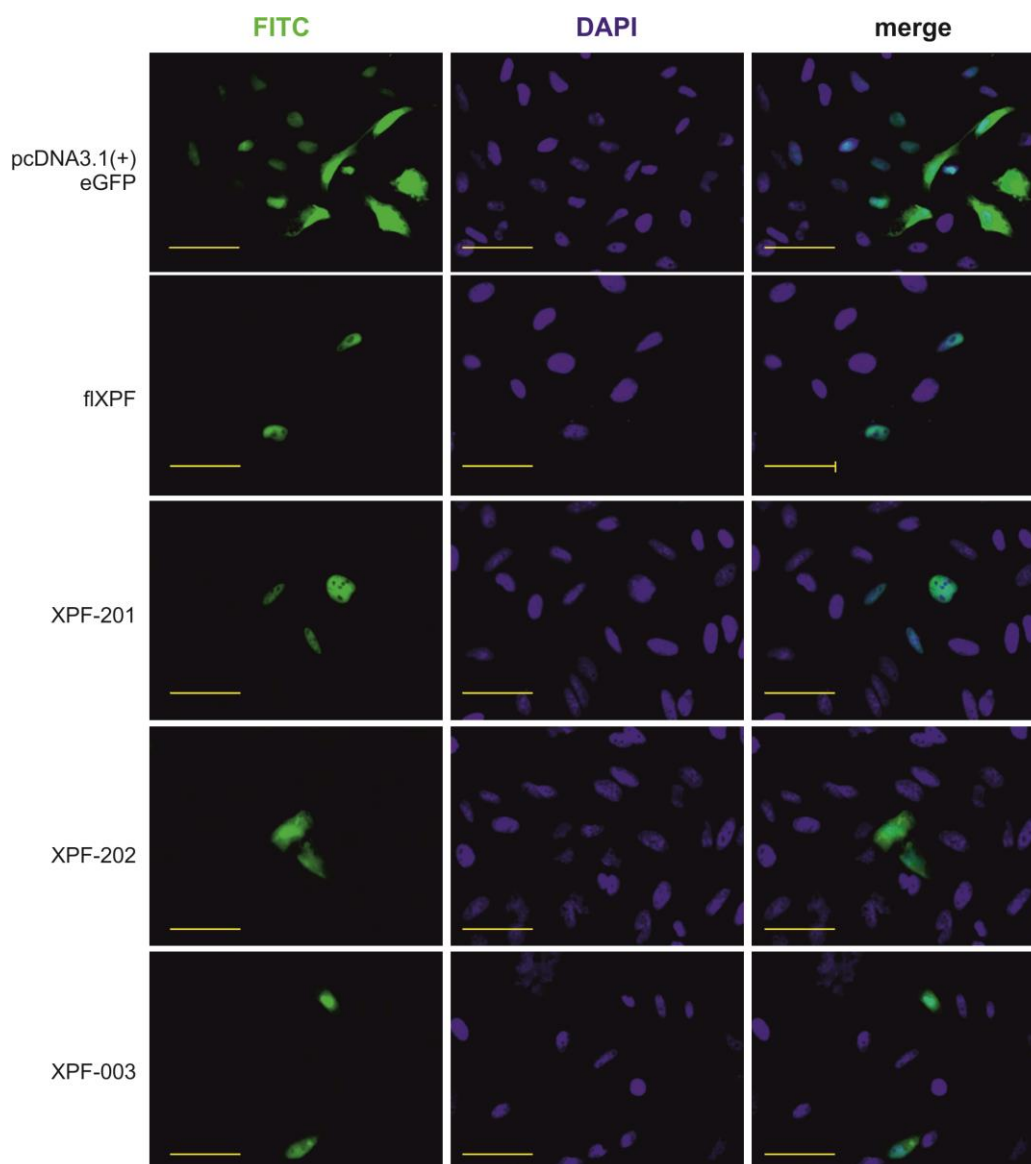


Figure 24: Subcellular localization of eGFP-tagged *XPF* isoforms and patient alleles

XPF isoforms and patient alleles were cloned into an pcDNA3.1(+)/eGFP expression vector and overexpressed in HeLa cells for 48h to analyze subcellular localization. Additionally, DAPI staining was performed to visualize the nucleus. Scale bar = 50 μ m.

4.4 Splice variants and their involvement in different DNA repair pathways

As analyzed during my master thesis, *XPG* isoforms either show C- or N-terminal truncations, leading to full or partial loss of protein domains important for enzymatic activity or protein-protein interactions. In order to investigate consequences on repair capability, functional assays were applied. Due to the lack of an *ERCC1* KO cell line, no functional analyses of the splice variants could be performed at this time, but will be subject to further investigations. This thesis therefore focusses on *XPG* and *XPF* splice variants.

4.4.1 *XPG* isoforms V and VI significantly increase the repair capability of intra- and interstrand crosslinks

The residual repair activity of splice variants was determined by HCR (see 3.3.2.1). Originally, this assay was applied to study the repair of UV induced DNA lesions (6-4PPs and CPDs) that are repaired by the NER. Therefore, a non-replicating reporter gene plasmid (firefly luciferase) was irradiated with UVC and transfected into an *XPG*-deficient host cell line (XP20BE). Under this circumstance, enzyme expression can only be detected if the DNA lesions of the reporter gene's transcribed strand are removed due to activity of full-length *XPG* or residual repair capabilities of the isoform coded on the complementation plasmid (see Figure 14). Hence, enzyme activity correlates with NER activity and is measured as repair percentage. Repair activity of different *XPG* isoforms was already determined by HCR assay during my master thesis (see Figure 12) (Lehmann, 2013). Primary WT fibroblasts exhibited a NER repair capability of approximately 15% ($15.72 \pm 0.96\%$) (see Figure 25A). Only co-transfection with *XPG* isoforms V and VI led to a significant increase of repair capability in comparison to the empty vector control (*XPG*: $4.9 \pm 1.22\%$, IsoV: $0.40 \pm 0.07\%$, IsoVI: $0.78 \pm 0.12\%$) (** $P < 0.01$, *** $P < 0.001$, or * $P < 0.05$ $n = 4$), showing a 5-fold (XP20BE) increase in NER-activity. The other isoforms analyzed displayed no residual NER repair function.

The two endonucleases XPF/ERCC1 and *XPG* represent the excision complex of NER. Both proteins are also involved in other important cellular processes like basal transcription and ICL repair (reviewed in Wood, 2010). In regard to this, the HCR assay was adapted to study the repair of intrastrand and interstrand crosslinks by treating the firefly reporter plasmid with CP and TMP plus UVA light (see 3.3.2.1 and 4.2.4). WT repair levels of about 15% (CP: $13.92 \pm 1.71\%$, TMP + UVA: $17.25 \pm 3.34\%$) were measured. Complementation with full-length *XPG* was closer to WT levels as in the case for UVC induced lesions (UVC: $4.9 \pm 1.22\%$, CP: $11.61 \pm 2.75\%$, TMP + UVA: $12.71 \pm 1.55\%$). Again, isoform V and VI showed a significant increase in repair capability (IsoV: CP $5.98 \pm 0.98\%$, TMP + UVA 8.55 ± 1.54 and IsoVI: CP $4.57 \pm 0.76\%$, TMP + UVA 9.94 ± 2.32) (** $P < 0.01$, *** $P < 0.001$, $n = 4$), while the other isoforms did not have a significant effect (n.s. $P > 0.05$, $n = 4$). However, the overall basic levels of repair were higher than for UVC (CP $1.98\% \pm 0.19$, TMP + UVA 5.12 ± 0.43), meaning that also the difference between isoform V and VI and the other isoforms were not as prominent as seen in NER (see Figure 25B and C). This suggests that the involvement of *XPG* in the repair of ICL repair is not as essential as for NER. In contrast, the MRC5Vi *XPF*

KO cells showed a nearly complete ablation of ICL repair (see Figure 21). As previously mentioned, this implicates an essential role of the XPF/ERCC1 endonuclease in ICL repair.

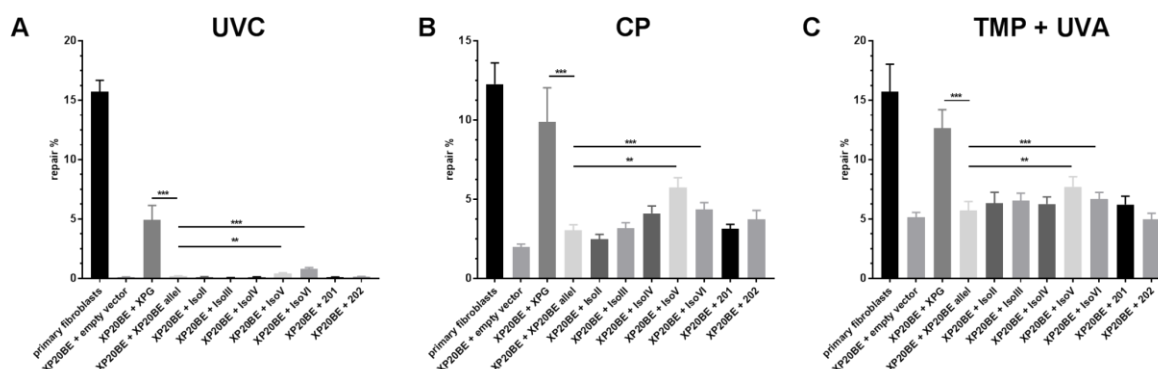


Figure 25: Reactivation of a reporter gene after treatment with UVC, cisplatin or trimethylpsoralen activated by UVA light in XP20BE patient or primary WT fibroblasts

For the HCR assays, firefly plasmids were treated with (A) UVC irradiation (B) cisplatin (intrastrand crosslinks) or (C) trimethylpsoralen activated by $1\text{J}/\text{cm}^2$ UVA irradiation (interstrand crosslinks) to induce specific lesions, transfected into XP20BE patient or primary WT fibroblasts, and complemented with plasmids coding for the empty expression plasmid, full-length *XPG*, or the *XPG* splice variants. The relative repair capability is calculated as the percentage (repair %) of the reporter gene activity (firefly luciferase) compared to the untreated plasmid, after normalization to an internal co-transfected control (*Renilla* luciferase). Data are presented as the mean \pm SEM. The one-tailed, unpaired student's t-test was applied, *** $P < 0.001$ or ** $P < 0.01$. At least four independent experiments in triplicates were performed.

Nevertheless, *XPG* isoform V and VI contained residual repair capabilities, but are both lacking essential parts of the endonuclease domains. Hence, the two isoforms seem to be able to fulfill at least structural functions of *XPG* during NER as well as ICL repair, and catalyze the subsequent excision performed by XPF/ERCC1 (Gary *et al.*, 1999).

4.4.2 XPF splice variants show residual repair capabilities during NER and ICL repair

As the XPF/ERCC1 complex is involved in multiple repair pathways (NER, ICL, DSB repair) and cellular processes (see 1.4.2), residual repair capabilities of *XPF* splice variants (see Figure 13) were studied in more detail by utilizing reporter gene assays in the complete *XPF* KO cells compared to MRC5Vi WT cells as well. Therefore, the HCR assay was applied to investigate correct repair of UVC, CP and TMP plus UVA irradiation induced DNA lesions (see 3.3.2.1). The variants as well as full-length *XPF* or an empty expression plasmid were transiently overexpressed in the KO cells (see 3.3.1.4).

As described above, *XPF* KO cells show a nearly complete loss of repair capabilities, but repair can be rescued by complementation with full-length *XPF* (see 4.2.4). Interestingly, two *XPF* splice variants exhibited significant residual repair capabilities in NER as well as ICL

repair (XPF-201: UVC 6.50 ± 0.22 , CP $16.28 \pm 0.51\%$, TMP + UVA $12.06 \pm 0.22\%$ and XPF-003: UVC $1.15 \pm 0.02\%$, CP $3.17 \pm 0.10\%$, TMP + UVA $3.89 \pm 0.03\%$) (** $P < 0.001$ or ** $P < 0.01$, $n = 4$). XPF-201 only lacks the first 12 aa of the protein, whereas XPF-003 is severely C-terminally truncated, with an intronic retention of five aa lacking the endonuclease domain. Surprisingly, XPF-202, another variant, that only differs from XPF-003 in the first 12 aa did not display any significant repair capability at all, neither in NER nor ICL repair (UVC $0.15 \pm 0.02\%$, CP $0.58 \pm 0.03\%$, TMP + UVA $2.17 \pm 0.06\%$) ($P > 0.05$, $n = 4$) (see Figure 26).

This implicates a high importance of those 12 N-terminal aa of the XPF protein. It could point towards an undescribed essential interaction domain of the protein, which may be necessary to communicate with other structural or catalytically active components of the DNA repair machinery.

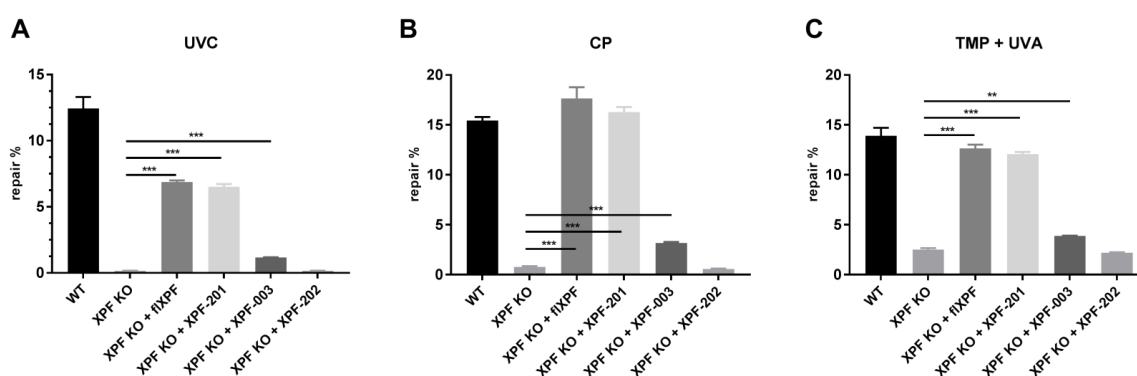


Figure 26: Reactivation of a reporter gene after treatment with UVC, cisplatin or trimethylpsoralen activated by UVA light in XPF KO cells complemented with XPF splice variants

For the HCR assays, firefly plasmids were treated with (A) UVC irradiation (B) cisplatin (intrastrand crosslinks) or (C) trimethylpsoralen activated by $1\text{J}/\text{cm}^2$ UVA irradiation (interstrand crosslinks) to induce specific lesions, transfected into MRC5Vi WT and XPF KO cells, and complemented with plasmids coding for full-length XPF and the different splice variants. The relative repair capability is calculated as the percentage (repair %) of the reporter gene activity (firefly luciferase) compared to the untreated plasmid, after normalization to an internal co-transfected control (*Renilla* luciferase). Data are presented as the mean \pm SEM. The one-tailed, unpaired student's t-test was applied, ** $P < 0.001$ or * $P < 0.01$. At least four independent experiments in triplicates were performed.

4.4.3 XPG and XPF splice variants do exhibit dominant negative effects during DNA repair

The interference of XPG and XPF spontaneous mRNA splice variants with endogenous DNA repair, exemplarily NER, in MRC5Vi WT cells was subject to further investigations. For this purpose, cell lines stably overexpressing these variants were generated by spontaneous

integration of a linearized expression plasmid into MRC5Vi WT cells. An expression plasmid coding for the full-length protein or one of the splice variants, e.g. pcDNA3.1(+)XPF, was linearized as described in 3.1.4.5, analyzed on an agarose gel (see 3.1.3.4), and purified by ethanol precipitation (see 3.1.1). Linearized plasmids were transiently transfected into MRC5Vi WT cells seeded on 10mm cell culture dishes and positive clones were selected using G418 (100mg/ml). Afterwards, single cell clones were generated by serial dilution and checked for plasmid integration by PCR using specific primer pairs for the CMV (cytomegalo virus)-promoter and the integrated gene (see 3.1.4.1, Table 7), as well as immunoblot analysis (see 3.4.4).

In order to test the effect of an overexpression on NER capability of MRC5Vi WT cells, for each variant or full-length protein, three positive clones with a similar level of overexpression were selected and repair capability was measured using the HCR assay (see 3.3.2.1). Repair capabilities were averaged for statistical purposes and compared to unmodified and non-manipulated WT cells. Figure 27A and B exemplary show, that the full-length proteins (XPG 180kDa, XPF 112kDa) as well as their splice variants (XPG IsoV 100kDa, XPF-201 112kDa) can be stably overexpressed in MRC5Vi WT cells, reaching overexpression ratios of 1.13 – 9.65 in comparison to WT cells. Regarding the impact on NER capacity of MRC5Vi WT cells ($14.17 \pm 1.20\%$), XPG-201, a variant that N-terminally lacks the first four exons in frame and therefore misses one of XPG's endonuclease domains (N-domain), showed a significant dominant negative effect on repair ($10.75 \pm 0.50\%$) ($* P < 0.05$, $n=4$) (see Figure 27C). This effect could also be observed for XPG IsoVI, the variant with the highest residual repair capability in XPG deficient cell, lacking the I-domain ($7.85 \pm 0.72\%$) ($*** P < 0.001$, $n=4$). Unexpectedly, all three *XPF* splice variants, as well as the full-length protein, significantly reduced repair capabilities of MRC5Vi WT cells (XPF: $9.19 \pm 0.83\%$, XPF-201: $5.26 \pm 0.80\%$, XPF-003: $6.90 \pm 0.06\%$, XPF-202: $8.08 \pm 0.73\%$) ($*** P < 0.001$, $n=4$).

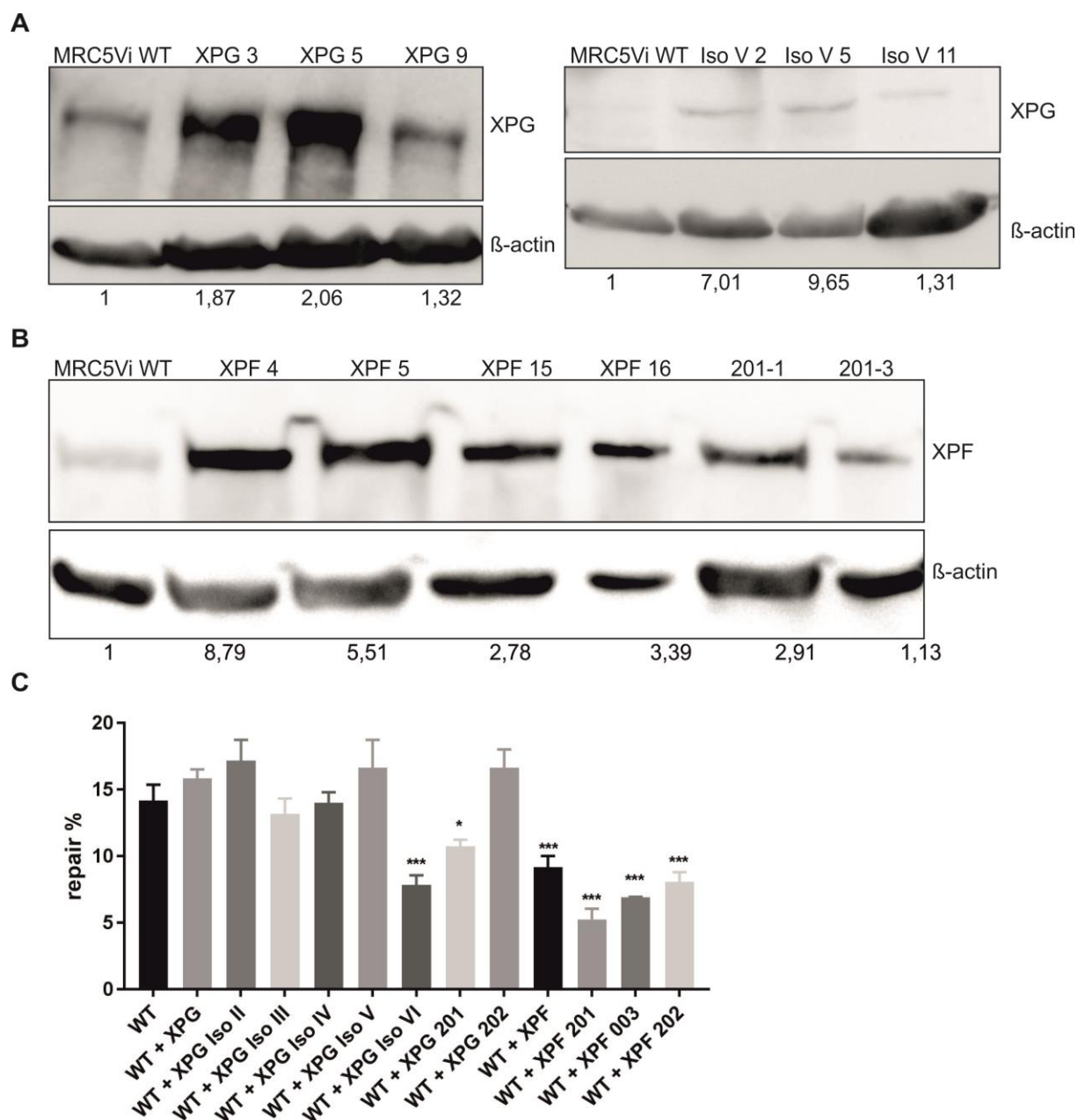


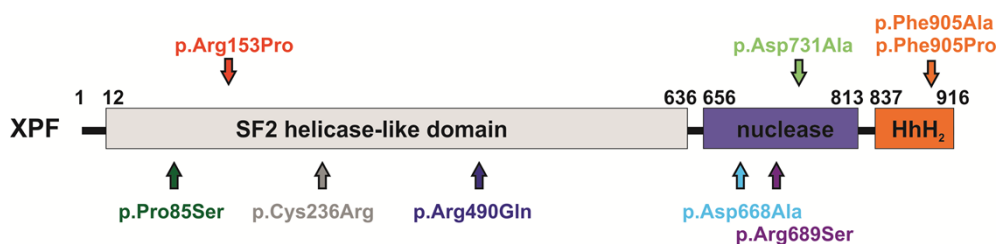
Figure 27: Immunoblot analyses and reactivation of a reporter gene after treatment with UVC in MRC5Vi WT cells and single clones overexpressing *XPG*, *XPF* or respective splice variants

Protein expression of MRC5Vi WT cells as well as single clones overexpressing the WT protein or respective splice variants was assessed by horizontal SDS Page followed by immunoblotting. Antibodies for XPG (**A**) or XPF (**B**) were used and β -actin staining was applied for normalization after reprobing. One of three independent experiments is shown. Expression was quantified by densitometric analysis. Protein levels in MRC5Vi cells were set to one. (**C**) For the HCR assays, firefly plasmids were treated with UVC irradiation to induce specific lesions, transfected into MRC5Vi WT or single clones overexpressing *XPG*, *XPF* or respective splice variants. The relative repair capability is calculated as the percentage (repair %) of the reporter gene activity (firefly luciferase) compared to the untreated plasmid, after normalization to an internal co-transfected control (*Renilla* luciferase). Data are presented as the mean \pm SEM. The one-tailed, unpaired student's t-test was applied; significances are displayed with regard to WT cell repair capability, * $P < 0.05$ or *** $P < 0.001$. At least four independent experiments in triplicates were performed.

Taken together, two *XPG* splice variants (*XPG* IsoVI and *XPG*-201) showed a dominant negative effect on MRC5Vi WT NER capabilities, as well as WT *XPF* and all three splice variants (*XPF*-201, *XPF*-003, *XPF*-202).

4.4.4 Analyses of XPF point mutants in the newly generated XPF KO cells give insights into mechanistic aspects of XPF's functions in different DNA repair pathways

The newly generated viable human fibroblasts cell line, lacking the essential endonuclease XPF, represents a great tool for mechanistic analyses of XPF's function in different DNA repair pathways. Up to this point, this has been limited to experiments in patient cells with extensive residual alleles, or *in vitro* studies. In these *in vitro* studies, NER is reconstructed using a set of 30 recombinant proteins, constituting necessary NER components. This system is highly artificial. Thus, in the literature, different point mutants have been described to change essential residues for different protein functions, according to these experiments. For further analyses, these variants, as well as patient alleles of diseases associated with XPF have been generated using site-directed mutagenesis (see 3.1.4.3). The patient alleles included an XP-F allele (XP3YO, p.R490Q), an XP-F/CS allele (XPCS1CD, p.C236R), a FA allele (FA104, p.R689S), and the severely affected XFE patient allele of XP51RO (p.R153P) (see Figure 28) (Bogliolo *et al.*, 2013; Kashiyaama *et al.*, 2013). In addition, an endonuclease defective mutant (p.D668A), postulated to be essential for NER (Sertic *et al.*, 2011; Staresincic *et al.*, 2009) as well as a mutant affecting the residue essential for ICL unhooking (p.D731A) (Fisher *et al.*, 2008; reviewed in Enzlin & Scharer, 2002), a point mutant of the essential residue for interaction with RPA and nuclear import (p.P85S) (Fisher *et al.*, 2008), and the residues for interaction with the heterodimeric partner ERCC1 (p.F905A/P) (reviewed in McNeil & Melton, 2012) were included into the experiment as well (see Figure 28).



Mutante	Funktion	Paper
D668A	endonuclease-defective mutant	Staresincic 2009, Sertic 2011
P85S	essential residue for RPA interaction and nuclear import	Fisher 2008
F905A F905P	essential residue for ERCC1 interaction	McNeil 2012
D731A	essential residue for crosslink unhooking in ICL repair	Enzlin und Schärer 2002, Fisher 2008 (psoralen)
R153P	XFE patient allele = XP51RO	Kashiyama 2013
C236R	XP-F/CS patient allele = CS1USAU = XPCS1CD	Kashiyama 2013
R689S	FA patient allele = FA104	Bogliolo 2013
R490Q	XP-F patient allele = XP3YO	Kashiyama 2013

Figure 28: Overview of XPF point mutants

The *XPF* gene was artificially manipulated by site-directed mutagenesis to produce different point mutants. Some changes were introduced on the basis of essential residues reported in the literature, while others resemble pathologically occurring mutations in patient alleles to study structural and functional domains of XPF.

The residual NER and ICL repair capability of the XPF point mutants was assessed using the HCR assay (see 3.3.2.1). In contrast to previously described studies (Sertic *et al.*, 2011; Staresincic *et al.*, 2009), the D668A mutant significantly complemented the *XPF* KO cells in NER and ICL repair up to WT *XPF* levels (UVC: $6.71 \pm 0.50\%$, CP: 16.73 ± 1.88 , TMP + UVA: $12.31 \pm 1.17\%$) (** $P < 0.01$ or *** $P < 0.001$, $n = 4$) (see Figure 29A-C). On the other hand, the D731A mutant, which was reported to be essential for ICL unhooking, did not complement the *XPF* KO cells in NER as well (UVC: $0.19 \pm 0.01\%$) (n.s. $P > 0.05$, $n = 4$) (see Figure 29A). Interestingly, the F905A/P mutant affecting the residue implicated for interaction with ERCC1, did complement the *XPF* KO cells (F905A UVC: $5.97 \pm 0.29\%$, CP: $14.19 \pm 0.85\%$, TMP + UVA: $11.02 \pm 1.05\%$, F905P UVC: $6.10 \pm 0.39\%$, CP: $14.07 \pm 0.82\%$, TMP + UVA: $12.50 \pm 1.63\%$) (** $P < 0.01$ or *** $P < 0.001$, $n = 4$). The P85S mutant, as well as mutants R689S and R490Q, resembling the FA and the XP-F patient alleles, did fully complement the *XPF* KO cells in regard to both NER and ICL repair. The mutant, resembling the XP-F/CS patient allele C236R, as well as the XFE mutant R153P did show complementation in NER and ICL repair (C236R UVC: $1.76 \pm 0.22\%$, CP: $3.76 \pm 0.32\%$, TMP + UVA: $5.09 \pm 0.66\%$, R153P UVC: $3.65 \pm 0.64\%$, CP: $4.94 \pm 0.43\%$, TMP + UVA: $6.29 \pm 0.54\%$) (** $P < 0.01$ or *** $P < 0.001$, $n = 4$), but clearly to a lesser extent than

WT *XPF* or the other complementing mutants. For both mutations, problems with misfolding, nuclear import, protein stability and interaction with ERCC1 or SLX4 have been postulated (Kashiyama *et al.*, 2013).

Clearly, these results highlight the differences between cellular backgrounds and artificial systems, emphasizing the necessity of suitable model cell lines.

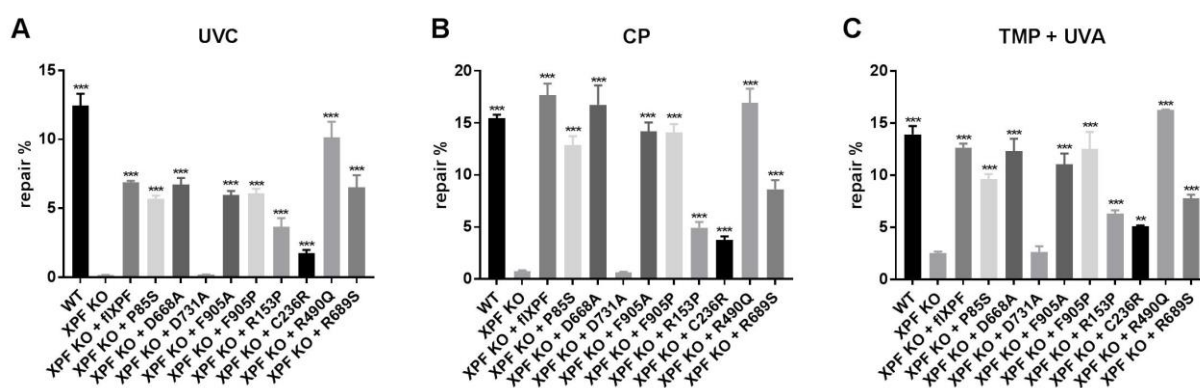


Figure 29: Reactivation of a reporter gene after treatment with UVC, cisplatin or trimethylpsoralen activated by UVA light in *XPF* KO cells complemented with *XPF* point mutants

For the HCR assays, firefly plasmids were treated with (A) UVC irradiation (B) cisplatin (intrastrand crosslinks) or (C) trimethylpsoralen activated by $1\text{J}/\text{cm}^2$ UVA irradiation (interstrand crosslinks) to induce specific lesions, transfected into MRC5Vi WT and *XPF* KO cells, and complemented with plasmids coding for full-length *XPF* and the different point mutants generated by site-directed mutagenesis. The relative repair capability was calculated as the percentage (repair %) of the reporter gene activity (firefly luciferase) compared to the untreated plasmid, after normalization to an internal co-transfected control (*Renilla* luciferase). Data are presented as the mean \pm SEM. The one-tailed, unpaired student's t-test was applied, ** $P < 0.01$ or *** $P < 0.001$. At least four independent experiments in triplicates were performed. Significances are displayed in regard to the *XPF* KO cells.

In order to assess stable protein expression of the generated point mutants, whole cell protein lysates were prepared (see 3.4.1) after 48h of overexpression in MRC5Vi *XPF* KO cells (see 3.3.1.4). Equal amounts of these extracts and an extract from MRC5Vi WT cells were loaded onto precast SDS gels, separated by horizontal SDS-Page and analyzed by immunoblotting for *XPF* and β -actin (see 3.4.3 and 3.4.4). Taken together, all *XPF* point mutants were (over)expressed over time (0.53 – 4.69) and resulted in proteins of the expected size, which could potentially function in NER and ICL repair (see Figure 30).

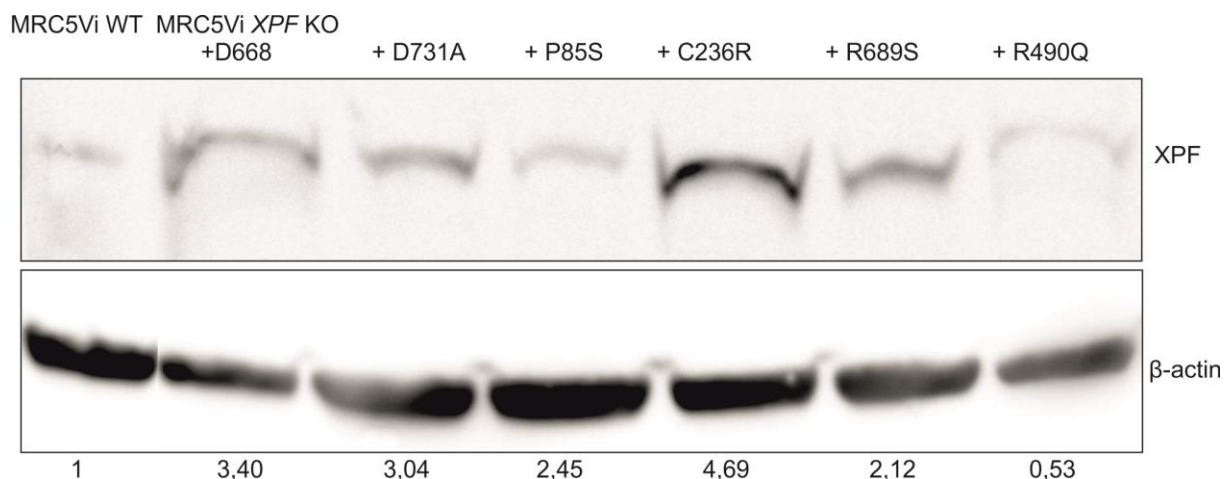


Figure 30: Immunoblot results for XPF point mutants generated by site-directed mutagenesis

In order to analyze protein levels, XPF point mutants were overexpressed in MRC5Vi XPF KO cells for 48h, whole protein extracts were prepared, separated by horizontal SDS Page, analyzed by immunoblotting, and quantified in comparison to WT expression. Blots were stained with an anti-XPF or an antibody directed against β -actin for normalization. Equal amounts were loaded onto the gel and one of three representative blots is shown.

Selected point mutants were investigated in regard to their subcellular localization by preparing cytosolic and nuclear protein extracts (see 3.4.1), followed by horizontal SDS-Page and immunoblotting as described above. As shown in 4.2.1, in the newly generated XPF KO cells, ERCC1 remains in the cytosol and cannot be detected in the nucleus without its heterodimeric partner XPF. Therefore, subcellular localization of the point mutants F905A/P, comprising the essential residues for interaction with ERCC1, was analyzed. Notably, F905A did not conclusively lead to reduced nuclear localization of ERCC1 as well as XPF (see Figure 31A and B), whereas F905P showed a reduction in nuclear levels of both proteins (see Figure 31D and E) in comparison to WT cells, that showed a strong induction of nuclear XPF as well as ERCC1. Interestingly, as described in previous studies (Niedernhofer *et al.*, 2006), the XFE patient allele R153P seems to result in a full-length XPF protein that is only stable in the cytosol but undetectable in the nucleus. In turn, also ERCC1 could not be observed in the nucleus (see Figure 31C).

Therefore, the F905A mutation seems to have no influence on ERCC1 and XPF localization, while a stronger distortion in protein structure due to insertion of a proline (strand breaker) in F905P leads to a slight nuclear reduction of ERCC1 and XPF. Moreover, the artificial mutant generated in parallel with the XFE patient allele R153P, showed that instable XPF cannot be transported into the nucleus. As observed before, without nuclear XPF nuclear ERCC1

protein could not be detected, confirming the effect seen in the MRC5Vi *XPF* KO cells, even further aggravating the necessity to a functional and stable XPF protein.

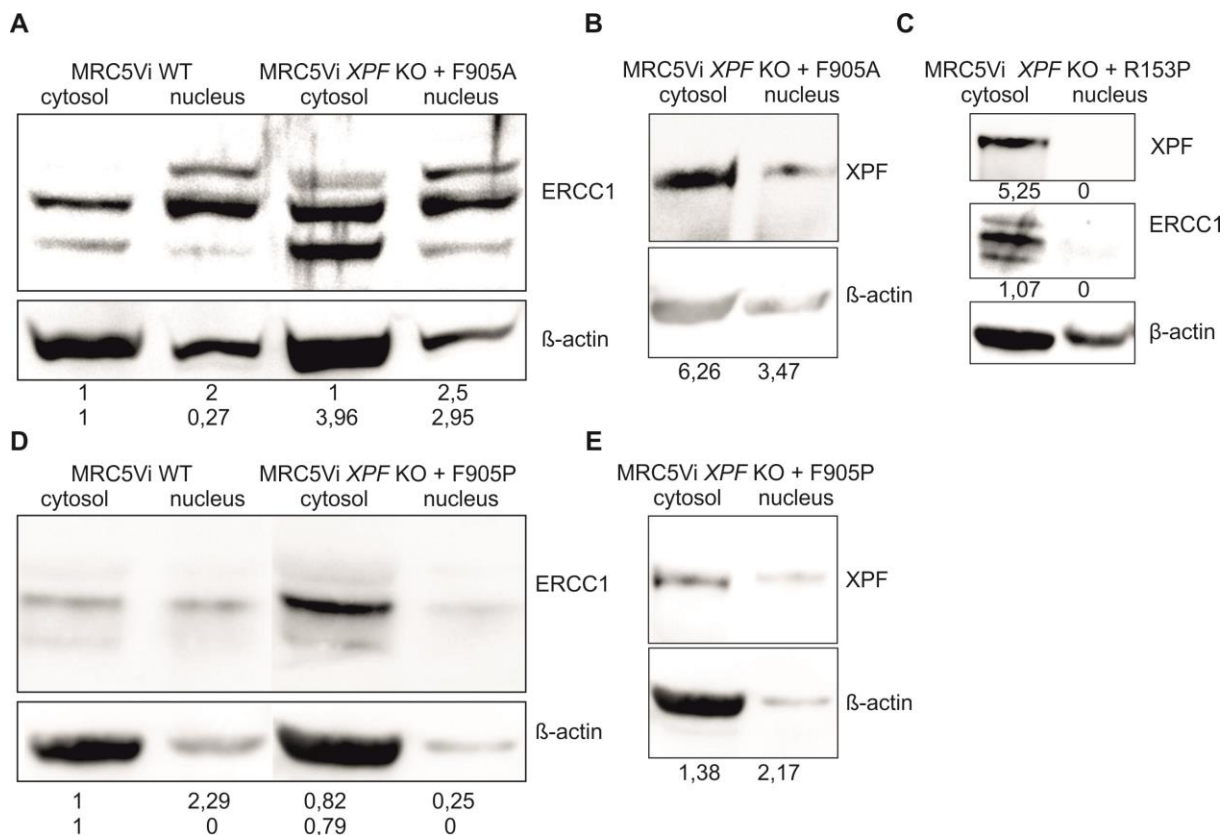


Figure 31: Immunoblot results for cytosolic and nuclear fractions of XPF point mutants generated by site-directed mutagenesis

In order to analyze subcellular protein localization and levels, XPF point mutants (**A + B**) F905A, (**C**) R153P, or (**D + E**) F905P were overexpressed in MRC5Vi *XPF* KO cells for 48h, cytosolic and nuclear fragments were prepared, equal amounts were loaded onto an SDS-Page gel, separated, analyzed by immunoblotting and quantified in comparison to WT levels. For ERCC1 the upper and lower ERCC1 band was set to one. Blots were stained with an anti-XPF or an antibody directed against β-actin for normalization. One exemplary blot of three independent experiments is shown.

4.5 Functionally relevant splice variants can be implicated as prognostic markers for individual cancer risk, therapeutic success, or disease outcome

In order to assess the physiological relevance of *XPG* isoforms with repair catalyzing functions (isoform V and VI), specific primer pairs for both isoforms as well as for WT *XPG* were utilized to analyze RNA expression in 20 different healthy human tissues obtained from Ambion® (see Table 8). Therefore, cDNA was prepared from RNA samples and quantified as described in 3.1.3.2 and 3.1.4.2. Specific primer pairs binding to the intronic regions of the isoforms were applied (see Table 7). Unfortunately, the material was only sufficient to

analyze *XPG*, but not *XPF*, variants and further material could not be reordered due to discontinuation of the Ambion[®] product.

Figure 32 illustrates the comparison of the three *XPG* species in several different tissues. All 20 tissues expressed WT *XPG*, as well as both functionally relevant isoforms. Whereas heart, liver and skeletal muscle exhibited very low expression levels of all three *XPG* transcripts (light blue arrows), the expression of both isoforms, V and VI, was relatively high in spleen, testes, and thymus compared to the primary transcript (dark blue arrows).

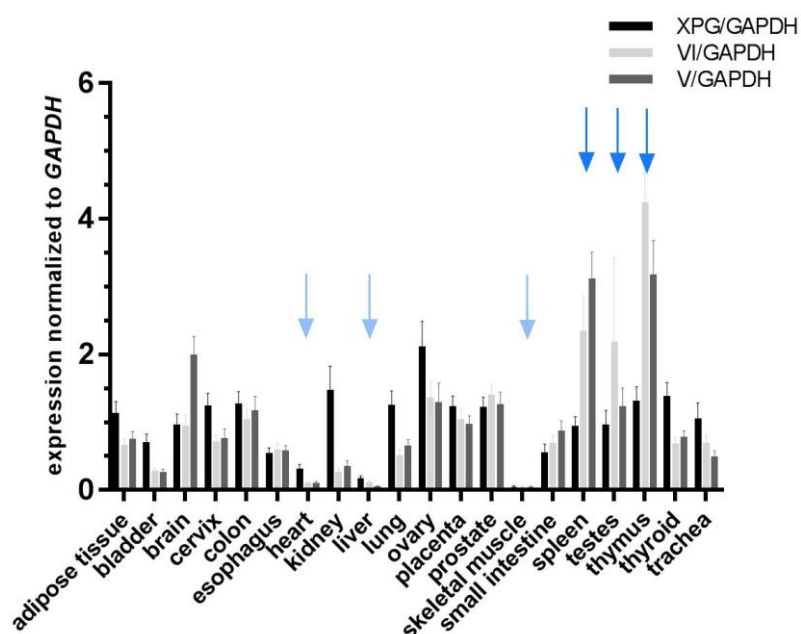


Figure 32: Comparative analysis of *XPG* expression in human tissues

RNA samples from 20 tissues were transcribed into cDNA and subject to real-time PCR measurements. Expression levels are shown as absolute values with *GAPDH* used as the internal control gene to normalize the expression of the target gene or isoform. Results represent mean values from three independent experiments in duplicates. Data are presented as mean \pm SEM. Colored arrows indicate groups with similar expression patterns.

To compare the expression levels of the three *XPG* transcripts between human individuals, blood samples from 20 healthy donors (age between 22 and 61, 7 males and 13 females) were analyzed as described above (see Figure 33). The comparison of these 20 unrelated individuals displayed a unique pattern for each of them. The expression profile either displayed high or low expression of all three transcripts (middle blue arrows), expression of a predominant primary transcript (dark blue arrows), or predominance of the two isoforms V and VI (light blue arrows). Despite a few exceptions, isoform VI, with the highest residual catalytic activity in the functional experiments, was the predominantly expressed *XPG* transcript in blood samples in 80% of the investigated subjects. No relations of expression

between age and sex were evident in this small collective. Unfortunately, the material was only sufficient to analyze the *XPG*, but not *XPF*, variants.

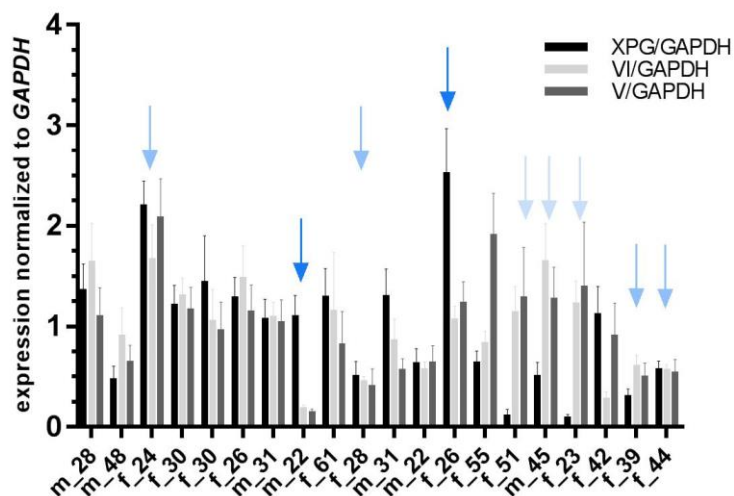


Figure 33: Analysis of *XPG* expression in human blood samples

RNA samples from 20 healthy donors were transcribed into cDNA and subject to real-time PCR measurements. Expression levels are shown as absolute values using *GAPDH* as the internal control gene to normalize the expression of the target gene or isoform. Results represent mean values from three independent experiments in duplicates. Data are presented as mean \pm SEM. Colored arrows indicate groups with similar expression patterns.

5 Discussion

The endonucleases XPG and XPF/ERCC1 are central players in NER and ICL repair, but also in basal transcription by encompassing both, catalytic and structural roles. Mutations in both genes are associated with complex genotype-phenotype correlations, for instance XPG splice variants have already been implicated in having residual functions that lead to a milder form of the XP/CS phenotype (reviewed in Scharer, 2008). In 2004 Thorel *et al.* reported about a 28-year-old patient with a splice variant of only minimal residual function, which resulted in a meaningful prolongation of the patient's life. In contrast, a functionally dominant negative splice variant that reduced NER by approximately 50% has been identified in an XP-C patient (Khan *et al.*, 2002). This splice variant was eventually identified in about 40% of normal individuals and associated with a two-fold increased melanoma risk (Blankenburg *et al.*, 2005).

Thus far, little is known about the functionality, expression patterns and structural functions of physiologically occurring, spontaneous mRNA splice variants of XPG and XPF/ERCC1, as well as their potential implication as prognostic markers in skin cancer prognosis, progression and therapy outcome. This was the main subject of this PhD thesis. Additionally, four new XP-A patients were characterized and a new mutation was detected, leading to an alternatively spliced variant in the XPA gene

5.1 A large deletion in the XPA gene results in XP with severe neurological symptoms

A new disease-causing mutation was identified in the XP-A complementation group. This novel and unusually large deletion mutation (genomic deletion of exons three-four) leads to a severely truncated XPA protein with no functional repair activity (Lehmann *et al.*, 2015).

Due to geographical isolation, a higher frequency of XP-A patients can be observed in Japan and Northern Africa compared to the world-wide distribution. So far, three common founder mutations (IVS3-1G>C; p.R228X; p.Y116X) have been described (Moriwaki *et al.*, 2012; Nishigori *et al.*, 1993) (see Figure 16F). In addition, a specific mutation in the XPA gene (c.747 C>T) has been reported in 12% of 66 unrelated families in the Maghreb region (Algeria, Morocco, Tunisia) (Soufir *et al.*, 2010; Tamura *et al.*, 2010). As part of this thesis, four XP-A patients (XP31MA, XP32MA, XP31GO, and XP118MA) were characterized. All patient cell lines showed high sensitivity against UVC irradiation (Figure 16A). However,

post-UV survival only reflects the cell's overall capability to cope with UV-induced DNA damages, which can involve multiple cellular mechanisms (e.g. NER, DDR, and apoptosis). Low viability after UV irradiation was consistent in patients with the XP plus neurological involvement phenotype. In order to specifically estimate the NER capacity, the HCR assay was applied. All four cell lines showed very low NER capability (Figure 16B), indicating a pivotal role of XPA during NER. For XP31MA, XP32MA, and XP31GO, stable protein expression was detectable, but not for XP118MA (Figure 16C). The well-known p.R228X mutation could be detected in XP31MA, XP32MA, and XP31GO in consistency with the mild XP plus neurological symptoms phenotype. As determined by epidemiological studies, consanguineous marriage and Northern African ancestry are promoting factors for this phenotype, which were evident for the siblings XP31MA and XP31GO, respectively.

A new homozygous mutation in the *XPA* gene of XP118MA was identified (c.268_269insAA), which resulted in two different XPA protein variants, one with a large in-frame deletion (missing exons three-four, p.P96_Q185del) and the other with a premature stop in exon two (p.104X). Protein expression could not be detected by western blotting, probably due to the missing antibody epitope within the truncated proteins (Figure 16C). However, *XPA* mRNA expression was measured excluding nonsense-mediated messenger decay, as well as proteasomal degradation, as no additional bands could be observed on protein level. Notably, this new mutation (c.268_269insAA) leads to two variants, which is very similar to the most common founder mutation in *XPA* (IVS3-1G>C). Both mutational events result in one truncated variant and another variant with exon skipping. The severe clinical phenotype of XP with severe neurological symptoms might be explained by the loss of important protein-protein (RPA70, TFIIH) and protein-DNA interaction domains (DNA-binding domain) (Figure 16D), provoking an extreme loss of function during NER (reviewed in Berneburg & Lehmann, 2001; Scharer, 2013).

This study in patients with DNA repair defective syndromes elucidates that not only the WT mRNA sequence undergoes spontaneous alternative splicing, but also inherited gene mutations can lead to two different mRNA variants of the mutated gene – in this case the *XPA* gene. On second note, the WT *XPA* gene is also predicted to result in several alternatively spliced variants with unknown functional relevance awaiting investigation.

5.2 The *XPF* CRISPR/Cas9 KO cells present a great tool to model XP

Over the past years, different site-specific techniques modifying genomes of various species have emerged using the site-directed ZFNs or TALENs. Recently, using the new and customizable CRISPR/Cas9 system was preferred, especially because of its simple approach to program specific gene disruptions or generate knockout cells, but also its easy design and efficient functionality. The structure-specific endonuclease XPF is involved in numerous DNA repair pathways (NER, ICL, and DSB repair) by comprising multiple protein functions. This is also evidenced by the variety of phenotypes and diseases arising from different mutations in the *XPF* gene. Mostly, mild clinical features, such as minor sun sensitivity and no neurological abnormalities, can be observed in XP-F patients (Hayakawa *et al.*, 1981). This is due to missense mutations in at least one allele not affecting the nuclease activity and therefore not completely abolishing XPF function, resulting in a residual repair capability of UV damage (up to 20%) (Ahmad *et al.*, 2010; Gregg *et al.*, 2011). On the other hand, a few more severe phenotypes like CS, XP/CS complex phenotype, XFE progeroid syndrome, or FA exist, that are caused by mutations leading to premature stop codons or missense mutations in essential residues (Kashiyama *et al.*, 2013; Niedernhofer *et al.*, 2006). Unfortunately, no suitable patient cells with a complete loss of the XPF protein have been available so far for extensive molecular and mechanistic studies. However, an *Xpf*-deficient mouse model with defective postnatal growth, a shortened life span, and hypersensitivity against UV irradiation and mitomycin C treatment has been described (Tian *et al.*, 2004).

During my PhD thesis I was able to generate the first viable human *XPF* KO cell line while applying the CRISPR/Cas9 technique targeting exon two of *XPF*. These cells are compound heterozygous and contain a premature stop codon in exon two of *XPF*, resulting in a massively truncated XPF protein with no residual activity (Figure 21). No further mutations in the *XPF* gene were detected. Protein expression, assessed by immunoblotting, was completely ablated, while mRNA levels were only reduced by 50% (Figure 18) in consistency with a subset of XP-F patients, suggesting instability of the mutant XPF protein (Matsumura *et al.*, 1998). Another explanation could be the loss of the antibody epitope mapping aa 629-905. Apparently, it is possible to generate a viable cell line of human fibroblasts lacking the essential endonuclease XPF, while germline mutations to this extent seem to be incompatible with life in humans and mice, limiting knockout strategies from studying genes critical for embryogenesis. As described in 4.2.2, it was not possible to generate a viable cell line with a full *ERCC1* KO in the first exons as the cells were probably not viable due to the essential

role of the protein in many cellular processes. In following approaches, later exons should be targeted (three-ten), but these cells would not be of value to analyze the splice variants (Figure 13) comprising structural changes already in early parts of the protein sequence. Another possibility to generate an *ERCC1* KO cell line would be to utilize a commercial approach from e.g. GE Dharmacon, offering transient as well as lentiviral systems to increase efficiency. A strategy to circumvent lethality could also be conditional gene inactivation, which has been performed using CRISPR/Cas9 in mice to generate conditional alleles before (Yang *et al.*, 2013). Providing that, the CRISPR/Cas9 technique is also a useful tool to identify sets of essential and non-essential genes (Evers *et al.*, 2016; Hart *et al.*, 2014). In case of failure to generate a complete *ERCC1* KO cell line, studies would have to focus on *in vitro* assays using recombinant proteins.

As a structure-specific endonuclease, XPF/ERCC1 is essential for the repair of helix-distorting DNA lesions, like UV-photoproducts, in replicating as well as non-replicating cells by making incisions on the damaged DNA strand on the 5' side of the open "denaturation-bubble" intermediate (Sijbers *et al.*, 1996a; reviewed in Mu *et al.*, 1996). *XPF* KO cells showed no NER capability in functional analyses and were also highly sensitive against UVC irradiation and exhibited reduced host cell reactivation of an UVC-treated reporter gene plasmid (Figure 20A, Figure 21A). This effect could be rescued up to WT repair levels by complementation with full-length *XPF*, but neither *XPG* nor *ERCC1*. Furthermore, *XPF* KO cells showed a diminished repair activity of CP induced lesions of only 0.78%, making them an exquisite model system for mechanistic analysis of ICL repair (Figure 21B). Whereas the model previously used for this purpose, GM08437 (XP-F) cells, still containing one full-length allele with a missense mutation, show 20% residual repair of CP crosslinks (Enoiu *et al.*, 2012). Moreover, *XPF* KO cells exhibited a high sensitivity against TMP plus UVA irradiation induced lesions (Figure 20B-C, Figure 21C), that are repaired by the ICL repair pathway. Also, these compounds are used in the therapy of psoriasis, vitiligo, and cutaneous T-cell lymphoma (Cimino *et al.*, 1985; Gupta & Anderson, 1987; Hearst, 1981). XPF is essential for the incision, or "unhooking" step, of ICL in an NER-independent manner (Fisher *et al.*, 2008; Klein Douwel *et al.*, 2014; Kuraoka *et al.*, 2000; Rahn *et al.*, 2010; reviewed in Wood, 2010). Whether other endonucleases like Mus81/Eme1, SLX4/SLX1, or FAN1 can perform incisions normally conducted by XPF, but in a much less efficient manner, merits further investigation (Hanada *et al.*, 2006; Klein Douwel *et al.*, 2014; Pizzolato *et al.*, 2015; Takahashi *et al.*, 2015; Yamamoto *et al.*, 2011; reviewed in Zhang & Walter, 2014).

5.3 ERCC1 is retained in the cytosol without its heterodimeric partner XPF

Analyses of protein levels in the cytosol and nucleus of the newly generated *XPF* KO cells compared to MRC5Vi WT cells, showed that ERCC1 was not detectable in the nuclear fraction of the KO cells (Figure 19D). However, after complementation with full-length *XPF*, ERCC1 did localize to the nucleus (Figure 19F). In previous studies, in the absence of the heterodimeric partner, ERCC1 and XPF have been reported to be unstable (Biggerstaff *et al.*, 1993; Sijbers *et al.*, 1996b; van Vuuren *et al.*, 1993). Moreover, low levels of ERCC1 protein are a common feature of total protein extracts from XP-F patients (Yagi *et al.*, 1997). Furthermore, extracts from XPF- and ERCC1- defective cells do not complement one another *in vitro* (Biggerstaff *et al.*, 1993; van Vuuren *et al.*, 1993), highlighting the necessity of a functional XPF/ERCC1 complex to maintain protein stability. In the absence of either binding partner, the other forms aggregates respectively and will either be subject to proteolytic degradation, or nuclear export (Gaillard & Wood, 2001). The results of this thesis show contradictory results thereto. Actually, the ERCC1 protein is stable in the cytosol of the *XPF* KO cells, but could not be detected in the nucleus without its heterodimer partner. As assumed earlier, this effect could be explained by the lack of an NLS in ERCC1 (Boulikas, 1997), which prohibits it from entering the nucleus by itself (Ahmad *et al.*, 2010). In later studies however, the presence of an NLS was shown (reviewed in McNeil & Melton, 2012). Apart from that, rapid nuclear export of single ERCC1, outside the prematurely formed XPF/ERCC1 complex, might be the likeliest explanation. Importantly, overexpression of XPF in the *XPF* KO cells could restore the localization of ERCC1 in the nucleus. Notably, the previous results have been performed with whole cell protein extracts; hence it remains unclear, whether ERCC1 can be stabilized in the cytosol by other interaction partners than XPF, and requires further investigations. A study showed that application of a compound inhibiting XPF-ERCC1 protein-protein interaction to two different cancer cell lines sensitized those cells to CP, mitomycin C and UV irradiation. Thereby, removal of DSBs was decreased confirming the findings in the *XPF* KO cells (Jordheim *et al.*, 2013).

5.4 XPF is markedly involved in HRR but dispensable for NHEJ

XPF/ERCC1 deficient mammalian cells exhibit a moderate sensitivity towards DSB inducing agents and ionizing radiation suggesting a role of XPF/ERCC1 in one or more sub-pathways of DSB repair in an NER-independent manner (Ahmad *et al.*, 2008; Mogi & Oh, 2006; Murray *et al.*, 1996; Wood *et al.*, 1983). The yeast homolog Rad1/Rad10 of XPF/ERCC1 is

known to be involved in HRR as well as NHEJ by removing non-homologous 3' single-stranded flaps (Al-Minawi *et al.*, 2008; Niedernhofer *et al.*, 2001; Sargent *et al.*, 1997). The heterodimeric complex has been accounted for the error-prone single-strand annealing (SSA) sub-pathway, gene conversion, and homologous gene targeting in HRR of yeast and mammals (Adair *et al.*, 2000; Fishman-Lobell & Haber, 1992; Ivanov & Haber, 1995; Niedernhofer *et al.*, 2001; Sargent *et al.*, 2000), whereas in NHEJ the complex is only known to take part in the Rad52- and Ku70/Ku86-independent microhomology-mediated end-joining (MMEJ) sub-pathway (Ahmad *et al.*, 2008; Bennardo *et al.*, 2008; Ma *et al.*, 2003; McVey & Lee, 2008; Yan *et al.*, 2007). Notably, these studies primarily focus on mice, yeast, and hamster cells, but only hypothesize an involvement of the human XPF/ERCC1 complex during HRR, never testing a complete human knockout cell line. As shown in Figure 22A, in the newly generated *XPF* KO cells, HRR capability is reduced to 60%, while no effect on NHEJ could be observed (Figure 22B). This strongly implicates human XPF for error-prone single-strand annealing, gene conversion, and homologous gene targeting, rather than microhomology-mediated end-joining strengthening the importance of human XPF in the HRR subpathway of DSB repair, already indicated by the yeast based studies. Anyhow, the *XPF* KO cells only displayed a mild sensitivity against DSB inducing agents (Figure 20D-F), because XPF/ERCC1 is only involved in sub-pathways of HRR as well as NHEJ, loss of the protein may be compensated in DSB repair, unlike during NER or ICL repair, where the endonuclease is essential for functional repair. XPF is involved in three different HR related repair mechanisms and plays a key role in ICL repair not only via its endonuclease function, but also in the late stage of HR during ICL repair (Al-Minawi *et al.*, 2009). Although no significant effect on cell survival of the *XPF* KO cells could be shown after treatment with DSB inducing drugs, these cells did show a slight decrease in survival. This could be due to adding effects, rendering the KO cells less capable of repair in terms of HRR substrates that cannot be removed by NHEJ. For bleomycin the mode of action is still unsolved, but it is hypothesized that it chelates metal ions forming a pseudoenzyme that reacts with oxygen resulting in superoxide and hydroxide free radicals cleaving DNA (reviewed in Hecht, 2000). CPT and etoposide are topoisomerase inhibitors. While TOP1 is irreversibly trapped to the DNA by CPT, resulting in replication or transcription induced DSBs (reviewed in Tomicic & Kaina, 2013), etoposide forms a ternary complex with TOP2 and the DNA, thus preventing re-ligation of the DNA strands and resulting in DSBs (reviewed in Pommier *et al.*, 2010). XPF/ERCC1 is also implicated in a pathway repairing Top1-attached DNA nicks in an NER dependent manner (Takahata *et al.*, 2015).

Interestingly, the overall level of HRR capability was lower (up to 33.27%) than the NHEJ activity (up to 98.92%) in the WT cells. Although, the repair template is encoded on the same plasmid, cell-cycle dependent expression regulation of DSB repair genes could be an explanation for this. Different DSB repair pathways compete for DSBs, and the balance between different subpathways varies among species, cell types and cell cycle phases. Expression of DSB repair genes of HRR and NHEJ can be regulated by phosphorylation of proteins, chromatin modulation of repair factor accessibility, or the availability of homologous repair templates. Cells actively shift from NHEJ towards HRR during cell cycle progression from G₁ to S/G₂ phase in lower and higher eukaryotes. Studies showed that *RAD51* and *RAD52* expression is low in G₀/G₁, and strongly increases in S and G₂/M phase (Chen *et al.*, 1997).

5.5 *XPF* and *ERCC1* splice variants could successfully be cloned from wildtype fibroblasts, show stable expression, and localize to the nucleus

The major aim of this project was to functionally characterize and test physiologically relevant, spontaneous mRNA splice variants of the two endonucleases XPG and XPF/ERCC1. Therefore, it was a prerequisite to generate artificial constructs of these splice variants for further functional *in vitro* experiments. This had already been done for XPG during my master thesis (Lehmann, 2013), but so far except their existence (see www.ensembl.org) nothing more was known about *XPF/ERCC1* splice variants. All three/eight selected splice variants of *XPF/ERCC1* have successfully been amplified from WT cells and cloned into different expression plasmids (pcDNA3.1(+), pcDNA3.1(+)eGFP, pcDNA3.1(+)myc/HisA2), presenting useful tools for further functional analyses.

Overexpression of the generated constructs lead to stable expression of proteins of the expected size (see Figure 23, Figure A4). Although, the *XPF* splice variants show different peak protein level over time, all three splice variants as well as the full-length protein are stable over the time of the functional assays (48h). In contrast, *ERCC1* splice variants, as well as the full-length protein, generally decreased over time, except ERCC1-002 and ERCC1-008. Differences in expression could be explained by proteolytic degradation (Fischer *et al.*, 2004) due to toxic effects of these shorter isoforms as they may constitute dominant negative effects on DNA repair, which will be discussed later (see 5.7). Anyhow, densitometric analysis is not the most resilient method of protein quantification and should only be regarded as semi-quantitative. Many experimental factors influence the final results, as well as certain subjective evaluation artifacts (Gassmann *et al.*, 2009).

DNA repair takes place in the nucleus; therefore it is necessary that all proteins involved are either passively or actively transported into the nucleus. *XPF* splice variants XPF-003 and XPF-202 both lack the C-terminal NLS (aa486-495) (see Figure 13). All *ERCC1* isoforms still maintain the NLS (aa12-20). In general, proteins of less than 40kDa (around 350aa) are able to pass the nuclear pores by passive diffusion (Marfori *et al.*, 2011). Therefore, ERCC1 should passively diffuse into the nucleus, whereas XPF-003 and XPF-202 depend on a NLS as they lead to proteins of roughly 50kDa, suggesting a more inefficient transport for these two variants as they only contain one of two NLS. Localization was analyzed by generating C-terminally eGFP-tagged constructs. eGFP itself does not contain an NLS as it can passively be transported into the nucleus due to its small size (26.9 kDa) (Seibel *et al.*, 2007). However, fusion proteins of eGFP and the different isoforms would depend on an NLS. Fluorescent analysis (see Figure 24, Figure A4) showed that after 48 hours all isoforms localized in the nucleus (see 4.3.1). This suggests that the activity of one of the two XPF NLS is sufficient for nuclear import of the proteins and presents a prerequisite for catalytic activity of the splice variants due to residual enzymatic or structural function. Interestingly, only XPF-202 displayed a rather uniform cellular distribution with an accumulation in the nucleus. As already reported in the literature and shown in 4.4.4 for the XFE patient, dysfunctional or unstable XPF protein shows an abnormal subcellular localization (Ahmad *et al.*, 2010; Kashiyama *et al.*, 2013) (see Figure 31) suggesting that XPF-202 may not be involved in functional DNA repair.

5.6 Splice variants of the two endonucleases XPG and XPF show residual repair capabilities in NER and ICL repair

It is of special importance to analyze the residual catalytic or putative structural functions of physiologically occurring, spontaneous mRNA splice variants of DNA repair genes (see 1.6.2). *XPG* splice variants II-VI lack the endonuclease activity containing I region (see Figure 12). During my master thesis, it was shown that isoform V and VI contain residual repair capabilities. In the course of this PhD project, two splice variants of *XPF* (XPF-201 and XPF-003) with residual repair capabilities were identified, but the isoform XPF-202 that only differs from XPF-003 in the first 12 aa did not show any complementation.

Complementation of XPG- or XPF-deficient cells with full-length proteins only leads to partial restoration of repair capability and does not reach WT levels. This discrepancy has several reasons, e.g. transfection efficiency, synthesis and nuclear transport of the transfected expression plasmids, resulting in a difference in actual repair time compared to WT cells

containing the endogenous protein, and last but not least, posttranslational modifications that may influence catalytic functions. Moreover, XPG and XPF are also involved in transcription making the HCR assay rather useful for qualitative but not so much quantitative analyses. The complementation effect of splice variants is even lower than observed for full-length proteins. XPG isoform V and VI as well as XPF-003 only lead to a minor relative increase in NER and ICL repair capability, yet, clearly statistically significant in comparison to mock vector transfectants (see Figure 25, Figure 26). In comparison to the full-length proteins, this appears to be a marginal increase, but it may have a strong impact on a patient's survival or skin cancer free life period. All these isoforms only contain one NLS, which results in a decreased activity of nuclear import, partially explaining the low increase in repair capability. The effect might be stronger if these variants could be transported into the nucleus more efficiently. Emmert *et al.* examined a 14 year old Caucasian girl (XP65BE) with sun sensitivity, but no neurological abnormalities that had not suffered of skin cancer so far (Emmert *et al.*, 2002). XPG mRNA expression was nearly normal. She had inherited one allele with an early stop codon mutation from her father, while the maternal allele only showed a single base missense mutation with a residual repair activity of 10%. Therefore, this residual functional activity in one allele could be accounted for the mild clinical features without neurological abnormalities, emphasizing that a rise in repair activity up to 10% can greatly influence the severity of disease progression and skin cancer free survival time up to 20 years. This indicates that also minor repair activities of splice variants may have a huge influence on the cell's repair capabilities. Furthermore, Thorel *et al.* (Thorel *et al.*, 2004) reported about a 28 year old patient with advanced XP/CS symptoms that had two XPG alleles producing a severely truncated protein. Interestingly, an XPG protein lacking seven internal amino acids, leading to residual endonuclease activity *in vitro* and weak TFIIH interaction ability, was generated by alternative splicing. The residual effects of this protein prolonged the patient's life for several years.

Commonly, all three splice variants of XPF and XPG evolve from alternative splicing and intron retention (five to eleven aa) (Emmert *et al.*, 2001). XPG isoform V shows a motif implicating the involvement of disulfide bonds in the protein-protein interaction surface (T-C-L-C-F-C-R). Numerous functions of intron retention have been explored over the last years, ranging from enhancing proteome diversity and gene expression to regulatory functions (reviewed in Graveley, 2001; Le Hir *et al.*, 2003). Intron gain or loss has been an important evolutionary engine over time (reviewed in Rogers, 1985). Large-scale analyses regarding intron retention in a set of 21,106 human genes, revealed at least one intron retention in

14.8% of the investigated genes, suggesting a biological function (Galante *et al.*, 2004). The question remains, whether there is evidence for functional introns influencing enzyme activity, for example in DNA repair gene splice variants. This has not been reported in the literature so far. However, intron-retaining splice variants of transcription factors have been postulated to be risk factors for different types of cancer, e.g. breast cancer, lymphoma or melanoma (Busse *et al.*, 2009; Honda *et al.*, 2012; Whiley *et al.*, 2011).

Notably, loss of ICL repair capability was not as prominent in XPG-deficient (XP20BE) as it was in XPF-deficient cells (see Figure 25B, C and Figure 26B, C). Thus, complementation effects of functional splice variants were not as strong for XPG IsoV and VI as they were for XPF-201 and XPF-003 compared to non-functional splice variants. There is a long and ongoing discussion about the importance of XPG during ICL repair, while it is essential for NER. Anyhow, discussion of ICL repair presents a major challenge. Up until now, understanding of the process is still poor. Indeed, there are many inconsistencies, contradictions, and uncertainties regarding the literature due to the numerous crosslinking agents with differing properties and products. Dual incision of an ICL is much more difficult to investigate. In prokaryotes, like *E. coli*, the UvrABC system, using a mechanism involving an ATP-dependent strand separation step, is able to cleave on both sides of an ICL (reviewed in Batty & Wood, 2000). In mammalian cells the mechanism is still unclear. Unhooking is a key step during ICL repair; therefore identification of the critical factors is clearly important. In regard to the NER pathway, roles of XPG (3' incision) and XPF/ERCC1 (5' incision), are of special interest. Unwinding could take place on both or only one side of an ICL, creating an open structure for endonucleases to access the lesion. Depending on the position of an ICL, XPF/ERCC1 is able to cleave on either side (Kuraoka *et al.*, 2000), but also XPG may be involved in cutting 3' to a junction between duplex and single-stranded DNA (Evans *et al.*, 1997a). Moreover, additional nucleases could be implicated in ICL unhooking depending on the situation. Cells deficient in XPF or ERCC1 are hypersensitive to crosslinking agents, but cells with defective XPG or other NER genes, generally only display modest sensitivity (Andersson *et al.*, 1996; Clingen *et al.*, 2007; Damia *et al.*, 1996; De Silva *et al.*, 2002; Hoy *et al.*, 1985). Furthermore, as shown in Figure 25 and Figure 26 reduction in repair capability compared to WT cells was lower in XPG (CP 6-fold, TMP + UVA 3-fold) compared to XPF-deficient cells (CP 20 fold, TMP + UVA 6 fold). This emphasizes the essential role of the XPF/ERCC1 endonuclease in ICL repair due to dual involvement in ICL unhooking.

Additionally, it seems as if there is a more important role for both endonucleases in removal of CP induced crosslinks, which are known to be rather intra- than interstrand crosslinks. For the experiments the HCR was adapted to CP and TMP + UVA induced lesions and only represent replication-independent removal of ICLs, as there is no origin of replication located on the pCMVluc plasmid (see Table 10). Previous studies showed that human cells deficient in XPG nearly showed WT unhooking kinetics of psoralen + UVA induced ICLs (De Silva *et al.*, 2000; Rothfuss & Grompe, 2004), while XPF/ERCC1 deficient cells were unable to unhook these lesions clearly indicating a necessity of the XPF/ERCC1 complex (De Silva *et al.*, 2000; De Silva *et al.*, 2002). Trimethylpsoralen that was used during this thesis is a psoralen derivative, but seems to be more efficient in producing ICLs. It was able to reduce repair levels of XPG- and XPF-deficient cells (5.12%, 2.50%) (see Figure 21 and Figure 25), but much less efficient than UVC or CP. Interestingly, CP induced ICLs rendered cells much more sensitive and also impaired unhooking in XPG-deficient cells in other experiments as well (De Silva *et al.*, 2002). This again raises the question, how to handle compounds minorly producing ICLs, but rather single strand adducts which are NER substrates, like CP, changing the dynamics of ICL repair.

However, current models of ICL repair propose a fundamental role of the XPF/ERCC1 complex during the unhooking step of ICL repair, which is in concurrence with the results of this thesis where both lesions result in nearly no residual repair capability in our *XPF* KO cells. Furthermore, there are studies implicating XPF/ERCC1 in an NER-independent ICL unhooking at stalled replication forks, which could not be examined using the experimental set up of this thesis (Fisher *et al.*, 2008). As proposed before, XPF/ERCC1 may associate with the scaffold protein SLX4 and two other structure specific endonucleases Mus81/Eme1 and SLX1, when SLX4 enhances their nuclease activity (Andersen *et al.*, 2009; Fekairi *et al.*, 2009; Munoz *et al.*, 2009). An siRNA knockdown of SLX4 rendered cells sensitive to crosslinking agents, but not UV irradiation, implicating an interaction of XPF/ERCC1 and SLX4 outside of NER (Munoz *et al.*, 2009). In regard to these results, a double knockout of *XPF* and *SLX4* is of special importance to test the complementation ability of the XPF-003 splice variant. The above mentioned results suggested an important protein interaction domain of the very N-terminal part of the XPF protein. SLX4 is supposed as one interaction candidate, therefore, a double knockout should diminish XPF-003's complementation ability at least in regard to CP and TMP + UVA induced lesions. Moreover, XPF-202 that only differs from XPF-003 in the first 12 aa did not show any complementation and subcellular localization showed a less distinct accumulation in the nucleus, but rather a distribution all

over the cell (see Figure 24). This could also be explained by the disrupted interaction with SLX4 that is an important platform and scaffold protein to organize DNA repair factors. Apart from this, XPF/ERCC1 is also involved in recombination and DSB repair (Adair *et al.*, 2000; Ahmad *et al.*, 2008; Niedernhofer *et al.*, 2001; Sargent *et al.*, 1997), so it remains to be elucidated whether the interaction between XPF and SLX4 is important for unhooking or other crosslink repair associated pathways. However, there is strong evidence that XPF/ERCC1 and SLX4 simultaneously load onto the ICL (Klein Douwel *et al.*, 2014). Additionally, SLX4 functions as a SUMO E3 ligase and sumoylates itself, as well as XPF (Guervilly *et al.*, 2015).

It remains to be elucidated, by which mechanisms, apart from protein-protein interactions with other repair factors, splice variants are able to confer residual repair capabilities. As mentioned before, *XPG* IsoV and VI as well as XPF-003 lack at least parts of the endonuclease domains (see Figure 12 and Figure 13). The existence of a cellular backup mechanism for the *XPG* endonuclease was subject to the PhD thesis of Dr. rer. nat Steffen Schubert, where he demonstrated that endonuclease defective *XPG* was able to perform accurate NER in living cells. He proposed that severely truncated *XPG* splice variants can structurally complement an *XPG* defect and lead to functional NER by recruitment of other structure-specific endonucleases like Fen1. These endonucleases can then cleave the DNA and repair can proceed (Schubert, 2014). A similar mechanism could be implicated for the XPF/ERCC1 endonuclease complex, but was not subject of this thesis and will therefore not be further discussed.

5.7 *XPG* and *XPF* splice variants exert a dominant negative effect on wild type NER capacities

Dominant negative inhibition describes the manifestation of an impaired WT gene product function by co-expression of a mutant variant, originating from the same gene product (reviewed in Herskowitz, 1987). Interference of splice variants with endogenous DNA repair in regard to dominant negative effects has been investigated in 4.4.3. Therefore, exemplarily effects on NER have been analyzed after generation of cell lines stably overexpressing *XPG* and *XPF* splice variants as well as full-length proteins. Overall, two *XPG* splice variants (*XPG* IsoVI and *XPG*-201), as well as WT *XPF* and all three splice variants (*XPF*-201, *XPF*-003, *XPF*-202), exhibited a dominant negative effect on MRC5Vi WT NER capabilities (see Figure 27).

In the human genome, the majority of protein-encoding genes are subject to physiological splicing. Interestingly, pathological splicing has been reported in cancer tissue (reviewed in (Srebrow & Kornblihtt, 2006; Zhao *et al.*, 2016)). For example, a negative effect of Chk2 splice variants and their effect on downstream substrates like p53, and Cdc25A/C that are involved in cell cycle arrest and apoptosis have been reported, suggesting alternative splicing as a mechanism for repression of WT protein function (Berge *et al.*, 2010). Interestingly, a polymorphism of the splice acceptor site in intron 11 of the *XPC* gene, a central player in NER, results in the expression of a functionally dominant negative splice variant that reduces NER by about 50% (Khan *et al.*, 2002). This polymorphism occurs as a haplotype with another polymorphism in the *XPC* gene (Khan *et al.*, 2000) and could be associated with an increased risk for SCCs of head and neck, as well as melanoma development (Blankenburg *et al.*, 2005; Shen *et al.*, 2001). Dominant negative effects have already been observed for other enzymes besides XPG and XPF/ERCC1, like the telomerase catalytic subunit (hTERT). Although lacking telomerase activity, low levels of *hTERT* mRNA has been shown in different human tissues. Furthermore, six splice variants, among which one deletion variant missing parts of the catalytic core, have been described (Kilian *et al.*, 1997). This particular splice variant inhibits endogenous telomerase activity in developing tissues, thereby leading to telomere shortening and chromosome end-to-end fusions (Colgin *et al.*, 2000). This is in concurrence with XPF-202, XPF-003, and XPG-201, which also lack parts of the catalytic protein domains. Moreover, dominant negative inhibition is often seen in splice variants of multisubunit proteins, e.g. isoleucyl-tRNA synthetase. If the variant is co-expressed with the WT protein, assembly of a functional complex might be impaired due to formation of dysfunctional complexes as the variant and the WT protein compete for binding partners (Michaels *et al.*, 1996). This could also be the case for XPF splice variants in view of heterodimer formation with ERCC1. Anyhow, this would only apply for XPF-201, as XPF-202 and XPF-003 lack the HhH₂ motif important for ERCC1 interaction.

The observed effects could also be due to simple protein overload in the cell, due to the strong and artificial overexpression, leading to unproductive protein interactions of splice variants with a repair capability lower than the WT protein. Furthermore, protein ratios between XPG and XPF could be disturbed. As mentioned before the excision process is highly coordinated between XPF/ERCC1 and XPG following subsequent binding and cutting steps. However, it is possible that the results only represent gene dose effects. Interestingly, effects were much more significant and extensive for XPF than for XPG, suggesting that XPF protein ratios, interactions and functions are more sensitive to changes than XPG. Milne and Weaver already

reported about the disruption of DSB repair through nonfunctional interactions of *Rad52* splice variants and proposed the formation of a higher order repair/recombination complex, potentially containing other yet unidentified components (Milne & Weaver, 1993). Taken together, interpretation of dominant negative effects is difficult due to experimental challenges and merit further investigations.

5.8 Artificially generated XPF point mutants behave differently in the newly generated XPF KO cells

In vitro reconstituted NER using purified recombinant proteins (~30 proteins), as well as dual incision assays with constructs containing defined lesions, have been applied over the past decade to analyze interactions between the endonucleases XPF/ERCC1 and XPG during the excision step of NER and ICL unhooking (Lindsey-Boltz *et al.*, 2014; Moggs *et al.*, 1996; Tapias *et al.*, 2004). In course of these experiments, several essential residues of those endonucleases have been postulated and were subject to further *in vivo* analyses using the newly generated XPF KO cells in this thesis (see 4.4.4).

In the fully reconstituted *in vitro* NER system, the mutant XPF D668A (designated D676A in older publications) was devoid any NER activity and only lead to a marginal increase of 3' incision in the dual incision assay. Therefore, this variant is often used as an endonuclease defective XPF control (Sertic *et al.*, 2011). Addition of XPF D668A to XP-F patient cells only resulted in an UDS (unscheduled DNA synthesis), another readout for functional NER, of 8% of WT cells (Staresincic *et al.*, 2009; Tapias *et al.*, 2004). In conflict with these results, the D668A mutant showed complementation like WT cells or full-length XPF protein using the HCR assay, whilst D731A did show neither NER nor ICL repair capabilities. In the literature, this variant, also known as D720A, was only implicated to be essential for ICL unhooking. Fisher *et al.* (Fisher *et al.*, 2008) performed ICL specific *in vitro* studies and postulated a role of this residue in a 3'-5' exonuclease activity of XPF, which is also present in the yeast homolog Rad1/Rad10 (Guzder *et al.*, 2004; Mu *et al.*, 2000). Other *in vitro* studies also reported about residual NER capabilities of D720A (Tapias *et al.*, 2004), while Staresincic *et al.* could not confirm this in their dual incision assays (Staresincic *et al.*, 2009). This clearly illustrates the differences between artificially reconstituted and physiological test systems and the need for suitable model cell lines with a full protein reservoir.

The HhH₂ domain is highly conserved in XPF and ERCC1 (Gaillard & Wood, 2001; Sgouros *et al.*, 1999) and is essential for the dimerization of the hydrophobic C-terminal parts of the

proteins in order to function as a stable heterodimer (Tripsianes *et al.*, 2005; Tsodikov *et al.*, 2005). XPF Phe905 and ERCC1 Phe293 are thought to be residues essential for interaction, anchoring the proteins together (Figure 8). This has been confirmed by mutational studies for ERCC1 Phe293, which resulted in abolition of dimerization and enzyme activity (de Laat *et al.*, 1998c; Sijbers *et al.*, 1996b). However, despite identical biochemical features this has not been shown for XPF. A general dogma implicates that neither protein is stable without its heterodimeric partner and will rapidly be degraded due to aggregation (Choi *et al.*, 2005; Tripsianes *et al.*, 2005). Functional assays determined complementation abilities of two artificially generated XPF Phe905 point mutants, F905A/P, like full-length XPF during NER and ICL repair (see Figure 29). In this new setting, these residues do not seem to be essential for repair, and interaction with ERCC1. Furthermore, XPF and ERCC1 proteins could be detected after overexpression of the mutant variants in the XPF *KO* cells. Anyhow, protein stability seemed to be affected in the F905P variant, because nuclear levels of ERCC1 appeared to be lower compared to WT cells (Figure 31). This suggests that XPF Phe905 is involved in XPF ERCC1 protein-protein interaction, but not essentially, and that interaction between XPF and ERCC1 might be established by a larger interaction patch including several residues (Choi *et al.*, 2005).

Moreover, protein stability and mislocalization play an important role in regard to physiological DNA repair activity. The P85S mutant was described to retain normal ERCC1 binding in dual incision assays and an intact nuclease function comparable to the WT protein *in vitro*. Anyhow, this residue is important for the interaction with RPA70 and critically associated with nuclear import as it compromises the NLS. Although defective interaction with RPA showed little impact and seemed to be dispensable *in vitro*, CHO cells with this mutation were sensitive to UV irradiation and showed a decrease in 6-4PP removal (Fisher *et al.*, 2011a). The picture does not seem this black and white, as the P85S mutant contains a high complementation capability in the new experimental setting, though not as high as WT or full-length protein levels (Figure 29). It is possible that the expression of one functional NLS is enough for a full-length protein to reach sufficient nuclear protein levels, and could be investigated further using the XPF *KO* cells. Furthermore, RPA is also involved in positioning and stimulating the catalytic activities of XPF (Matsunaga *et al.*, 1996). In contrast, the R153P allele of an XFE patient showed a decrease in complementation ability of about 5% for UVC induced lesions (also reported in (Niedernhofer *et al.*, 2006)). *In vitro* this mutation lead to a catalytically active recombinant protein, while *in vivo* the XPF/ERCC1 heterodimer could only be observed in the cytoplasm. Microinjection could rescue nuclear

localization (Ahmad *et al.*, 2010). Analyses of protein localization performed in the course of this thesis, showed no nuclear XPF or ERCC1 in the *XPF* KO cells after overexpression of the R153P variant (Figure 31). This corroborated the finding that ERCC1 is not stable in the cytosol without its heterodimeric partner (Lehmann *et al.*, 2017), and further aggravates to a functional XPF protein. Indeed, this emphasized that correct sub-compartmentalization of DNA repair and mislocalization of the XPF protein to the cytoplasm, as well as protein misfolding, seem to regulate NER as well as ICL repair. This could lead to severe symptoms like XP/CS or XFE, rather than only a catalytic impairment of the protein. In addition, studies of the XFE patient with the p.R153 mutation exhibited a global attenuation of the somatotrophic axis with an upregulation of Ppar α/γ to rescue energy storage. This implicates a link between genome maintenance and systemic dampening of the growth hormone/IGF1 hormonal axis, connecting DNA damage and aging (Niedernhofer *et al.*, 2006). Therefore, the *XPF* KO cells might also present an interesting model to study accumulated aging processes.

In line with the literature, the patient allele C236R showed a lower complementation capability than the full-length protein, confirming *in vitro* results that reported reduced endonuclease activities. This leads to an XP/CS complex phenotype (Kashiyama *et al.*, 2013). Full complementation could also be observed for the XP patient allele R490Q, as expected for a mildly affected patient. On the other hand, the FA patient allele R689S only showed complete NER complementation abilities, while ICL repair was decreased. This is confirmed by the literature, as patient cells are sensitive to the ICL inducing agent mitomycin C, but not UV irradiation and show nearly normal UDS levels. Moreover, *in vitro* studies showed that cells are proficient in NER excision, but impaired in ICL unhooking (Bogliolo *et al.*, 2013).

To summarize this, it becomes clear that this new cellular background is a tantalizing tool for further mechanistic analyses. Reconstituted systems present good models, but are limited by the artificial combination of recombinant proteins, while *in vivo* even until now unknown components could be involved and discovered.

5.9 Importance of *XPG* isoforms for personalized medicine and further perspectives

Personalized medicine has been in the focus of research over the past decade. Differential expression of XP splice variants between individuals and even different tissues may be determining patient specific overall repair capacity, and thereby resistance to different therapies could be clarified. Resistance, e.g. in chemotherapy against cancer, arises from

different levels of DNA repair and has been shown to be a target for small molecule inhibitors (Gentile *et al.*, 2016). NER and ICL repair eliminate various types of DNA lesions including UV irradiation and platinum-based therapy induced lesions. Furthermore, personal expression levels of DNA repair genes are of special interest, as it was shown that resistances to platinum-based therapy correlate with high expression of ERCC1 (Barakat *et al.*, 2012). Interaction between ERCC1 and XPA is essential for NER regulation, and XPF/ERCC1 activity can be inhibited by blocking this interaction, sensitizing cancer cells to UV irradiation. Additionally, *in silico* drug screens have identified inhibitors of the XPF-ERCC1 interaction domain, disrupting the complex (Jordheim *et al.*, 2013). Moreover, high ERCC1 expression has been correlated with poor response in different cancer entities like e.g. non-small cell lung cancer, esophageal cancer, breast cancer, colorectal cancer, as well as head and neck cancer (Bilen *et al.*, 2014; Choueiri *et al.*, 2015; Gerhard *et al.*, 2013; Hayes *et al.*, 2011). Notably, testis tumors with reduced protein levels of XPF, ERCC1, and XPA could be cured using cisplatin-based chemotherapy even in advanced metastasis stages (Welsh *et al.*, 2004). Expression levels of splice variants in different tissues have been shown to be more suitable to distinguish between oncogene and non-oncogene samples than the primary gene transcript itself (Zhang *et al.*, 2013b). Besides, tumor-specific splice variants are often overexpressed (Blair & Zi, 2011; Brinkman, 2004; Yi & Tang, 2011). In regard to this, the abundance of splice variants with residual repair functions could be used as prognostic factors for therapy outcome, as well as for personalized and targeted medicine. Therefore, repair increasing functions and duration of skin-cancer free survival time in patients with a lack of the primary transcript (see 4.4.1) (Thorel *et al.*, 2004) as well as dominant negative functions (see 4.4.3) in healthy individuals have to be weighed and evaluated in larger collectives.

Due to a limited amount of material inter-individual differences in a small collective of 20 healthy donors as well as expression differences in 20 tissues have been assessed by qRT-PCR only for *XPG* and its two splice variants with residual repair function (IsoV and VI), whereas IsoVI also displayed dominant negative effects in wildtype cells (see 4.5). Interestingly, individuals with a high expression of isoform VI showed a low expression level of the primary transcript, implicating a dominant negative effect of this variant on full-length *XPG* expression, which could also be seen in functional repair assays (see Figure 27). In the future, *XPG* mRNA levels of splice variants could be quantified in special, e.g. high-risk cancer populations (patients with multiple dysplastic nevi and high melanoma risk, or organ transplant recipients with high SCC risk). Clinically, as well as epidemiologically, this would be of high importance in order to investigate beneficial or negative influences of these

variants on tumor development, disease outcome, therapeutic efficacies, and side effects. Inter-individual variations of DNA repair levels and association with cancer is well documented. Therefore, implication of differences in expression levels of isoforms displays a great prognostic value as cancer biomarker (Zhang *et al.*, 2013b). Taken together, the uniqueness of personal and tissue-specific expression levels, as seen in Figure 32 and Figure 33, of functionally active, alternatively spliced mRNA variants may thus be implicated in cancer susceptibility, responses to chemotherapeutics and therapeutic success (Colmegna *et al.*, 2015; Hu *et al.*, 2014; Li *et al.*, 2013a; Sun *et al.*, 2015).

5.10 Summary and conclusions

In conclusion, the novel and innovative CRISPR/Cas9 technique was established in the laboratory and applied to successfully generate a viable human *XPF* KO cell line. A markedly increased sensitivity to UVC, cisplatin, and psoralen activated by UVA, as well as reduced repair capabilities for NER and ICL repair were determined. This cell line can now be used as a valuable tool for further analyses regarding XPF's various functions in different cellular processes. It was shown that XPF is dispensable for ERCC1 protein stability but essential for ERCC1 nuclear localization. Moreover, besides its essential role in NER and ICL repair, XPF's endonuclease activity appears to be partially essential in the HRR subpathway of DSB repair. In the future, it might be possible to further investigate more general questions in DNA repair, like the ongoing discussion about the essential nuclease for incising the DNA around an ICL. This cell line could be utilized to clarify whether or not it might exclusively be XPF. Other detailed mechanistic questions of repair pathways like NER can be studied in a setting without residual XPF activity, as it has already been done for XPF point mutants in this thesis. Moreover, the proposed method could be expanded to other genes of the DNA repair machinery and *in vitro* as well as *in vivo* studies.

Furthermore, the residual repair capabilities lead to following questions regarding mechanistic studies, for which the *XPF* KO cells present an excellent model cell line. The 12 N-terminal missing aa of the inactive superfamily2 helicase-like domain of XPF may interact with the SLX4 platform for DNA repair, as it is already described for Mus81 (Nowotny & Gaur, 2016; Wyatt *et al.*, 2013). In a follow-up research grant the interaction between XPF and SLX4 could be investigated by creating *SLX4* CRISPR KO cells and even *XPF/SLX4* double mutants for functional repair assays. Moreover, additional interaction partner could be identified applying proteomics and BRET analyses.

Additionally, splice variants of *XPF* and *XPG* with residual repair capabilities, but also dominant negative effects in NER and ICL repair, were identified. These variants show inter-individual expression differences as well as expression differences in various tissues making them promising prognostic marker for individual repair capability, disease outcome, and therapy success. As these are preliminary results from a small collective of healthy donors, analyses should be expanded to screening of a larger collective to identify actual prognostic marker. As the same splice variants showed residual repair activity in a KO background, but also dominant negative effects if overexpressed in a WT background, one would expect a high expression of such splice variants to be associated with rather unfavorable disease outcomes. This has to be further elucidated in large translational studies.

6 Appendix

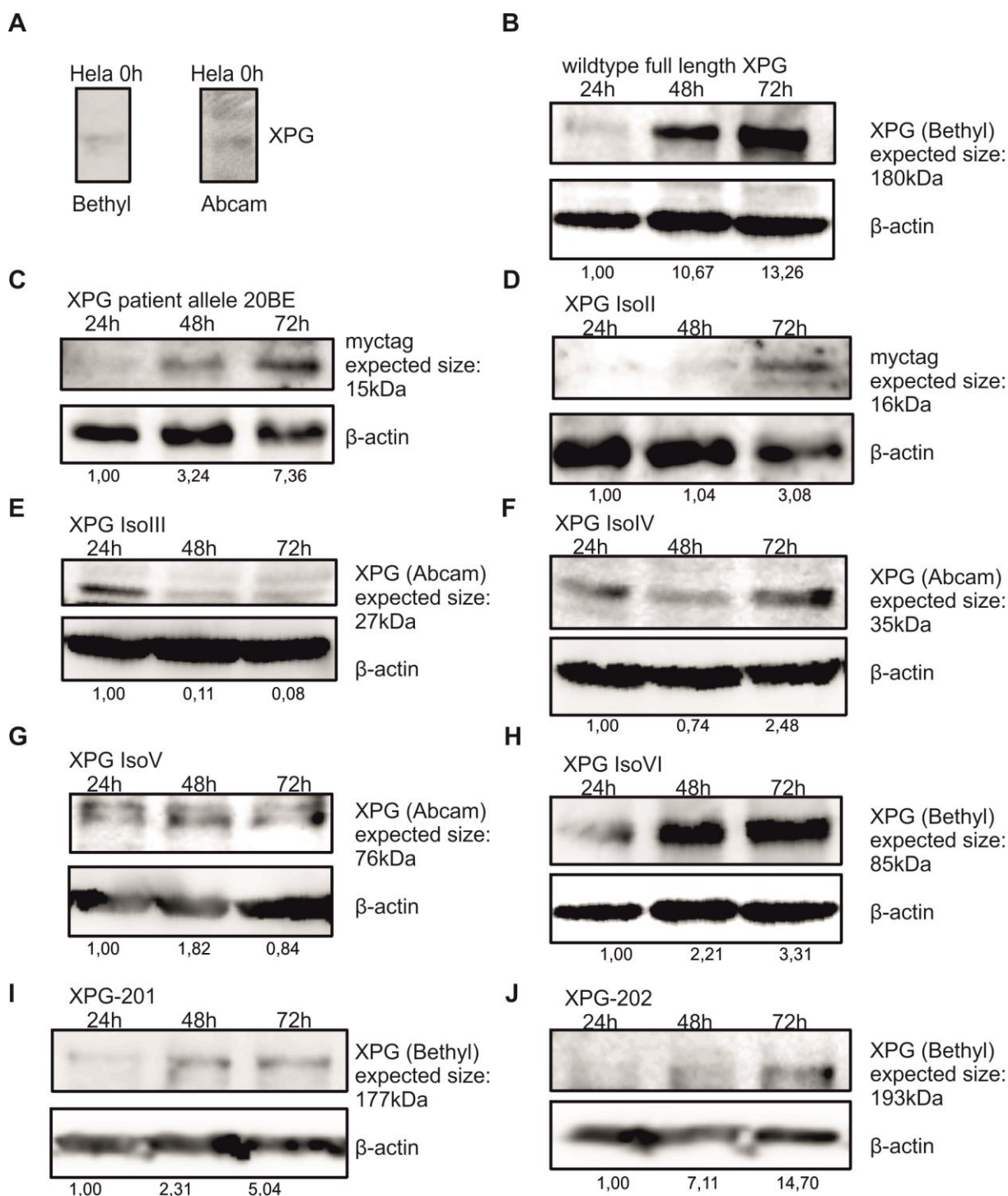


Figure A1: Immunoblot results for wildtype XPG, the seven isoforms and two patient alleles after overexpression in HeLa cells

Protein expression was assessed by horizontal SDS Page followed by immunoblotting after transient transfection of HeLa cells with the different constructs (24-72h). Two different antibodies for XPG, one spanning the epitope from amino acid 151-190 (Abcam) and the other from 650-700 (Bethyl), or an antibody directed against the myc-tag were used. β-actin staining was used for normalization. Expression was quantified by densitometric analysis.

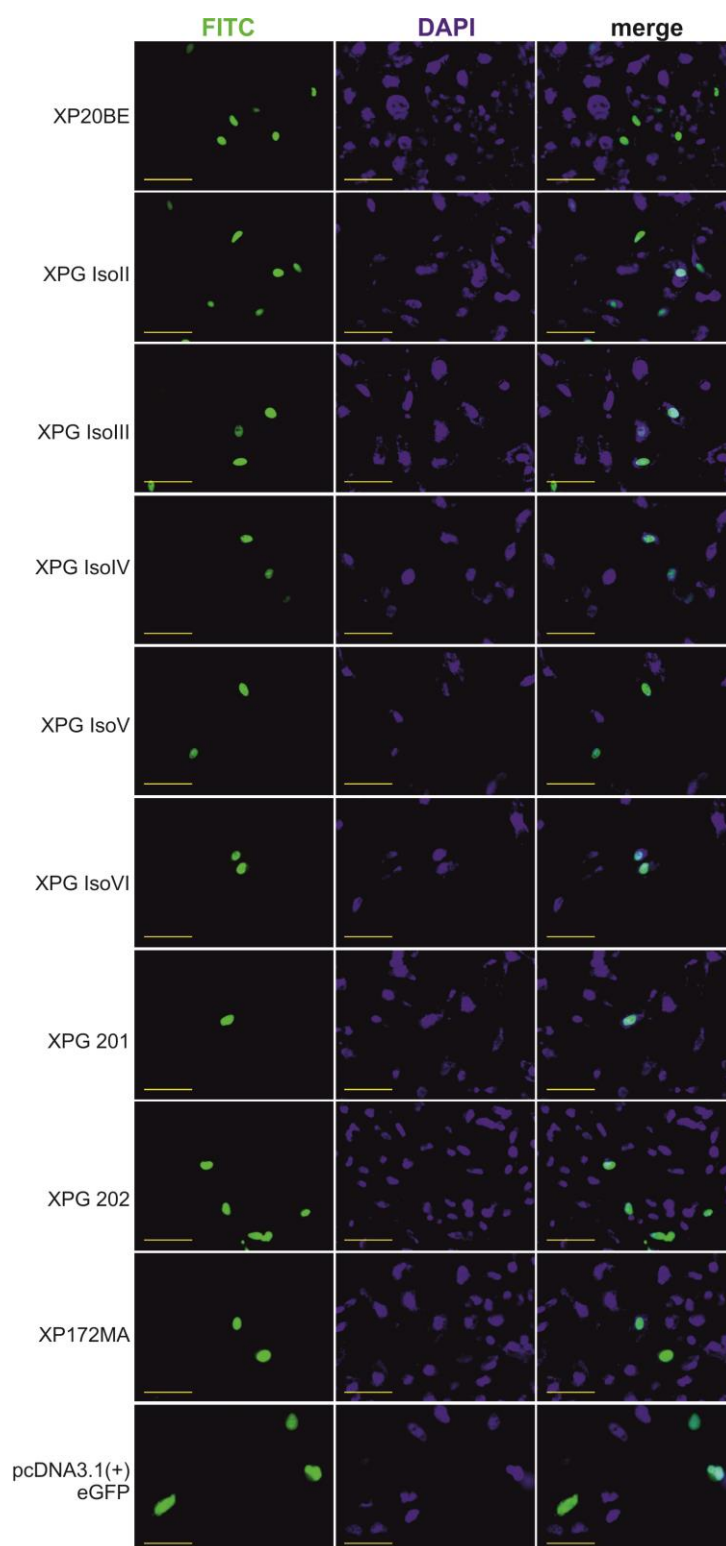


Figure A2: Subcellular localization of eGFP-tagged XPG isoforms and patient alleles

XPG isoforms and patient alleles were cloned into an pcDNA3.1(+)-eGFP expression vector and overexpressed in HeLa cells for 48h to analyze subcellular localization. Additionally, DAPI staining was performed to visualize the nucleus. Scale bar = 50 μ m.

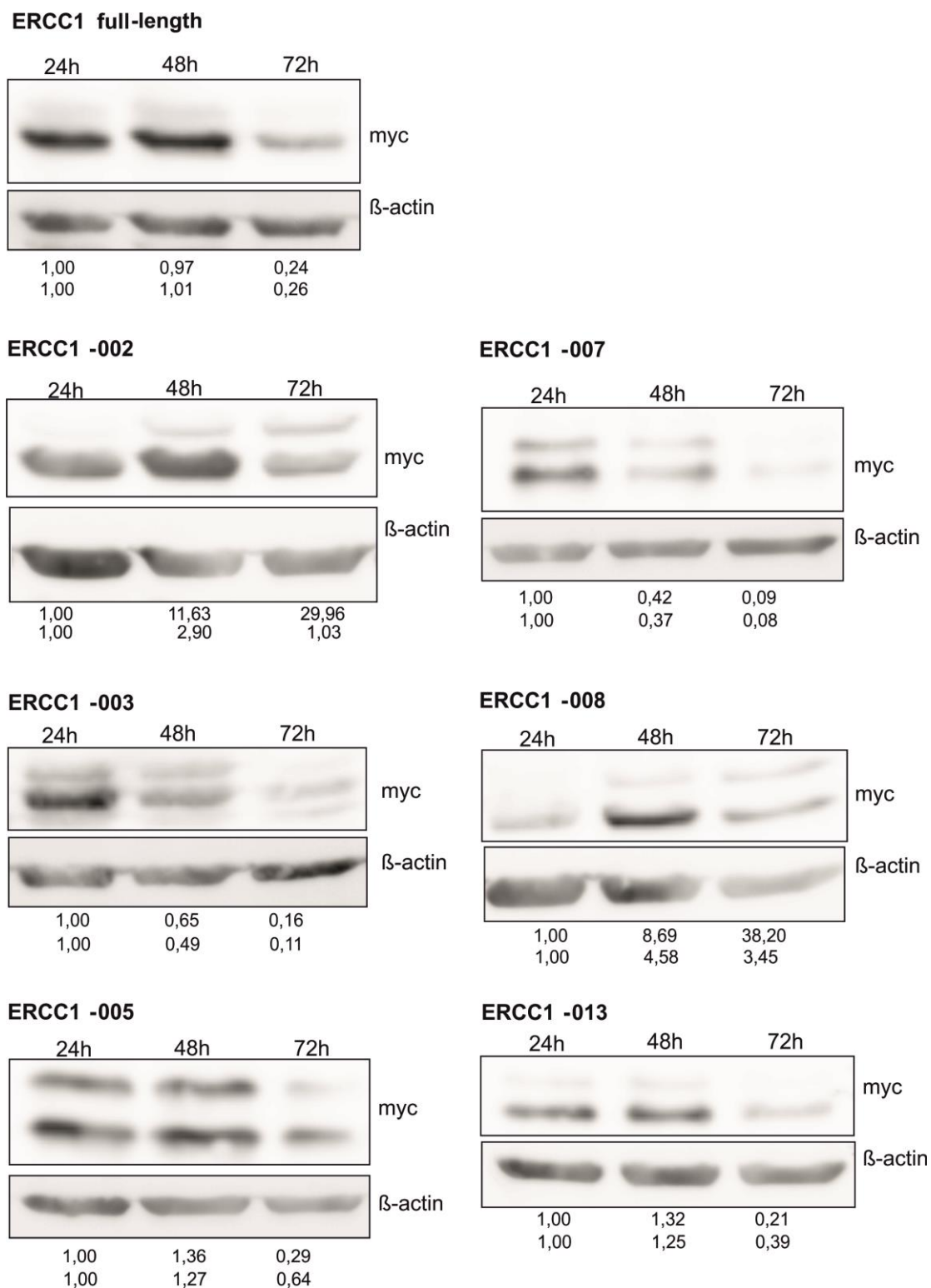


Figure A3: Immunoblot results for protein levels of *ERCC1* splice variants over time

In order to analyze protein levels after overexpression of the splice variants in HeLa cells, horizontal SDS Page followed by Immunoblotting was performed. Therefore, HeLa cells were transiently transfected with the different constructs and harvested after 24h, 48h and 72h and stained with an anti- myc-tag antibody. Anti- β -actin staining was used for normalization. Protein levels were quantified in regard to the 24h value as there is no endogenous myc-tag protein.

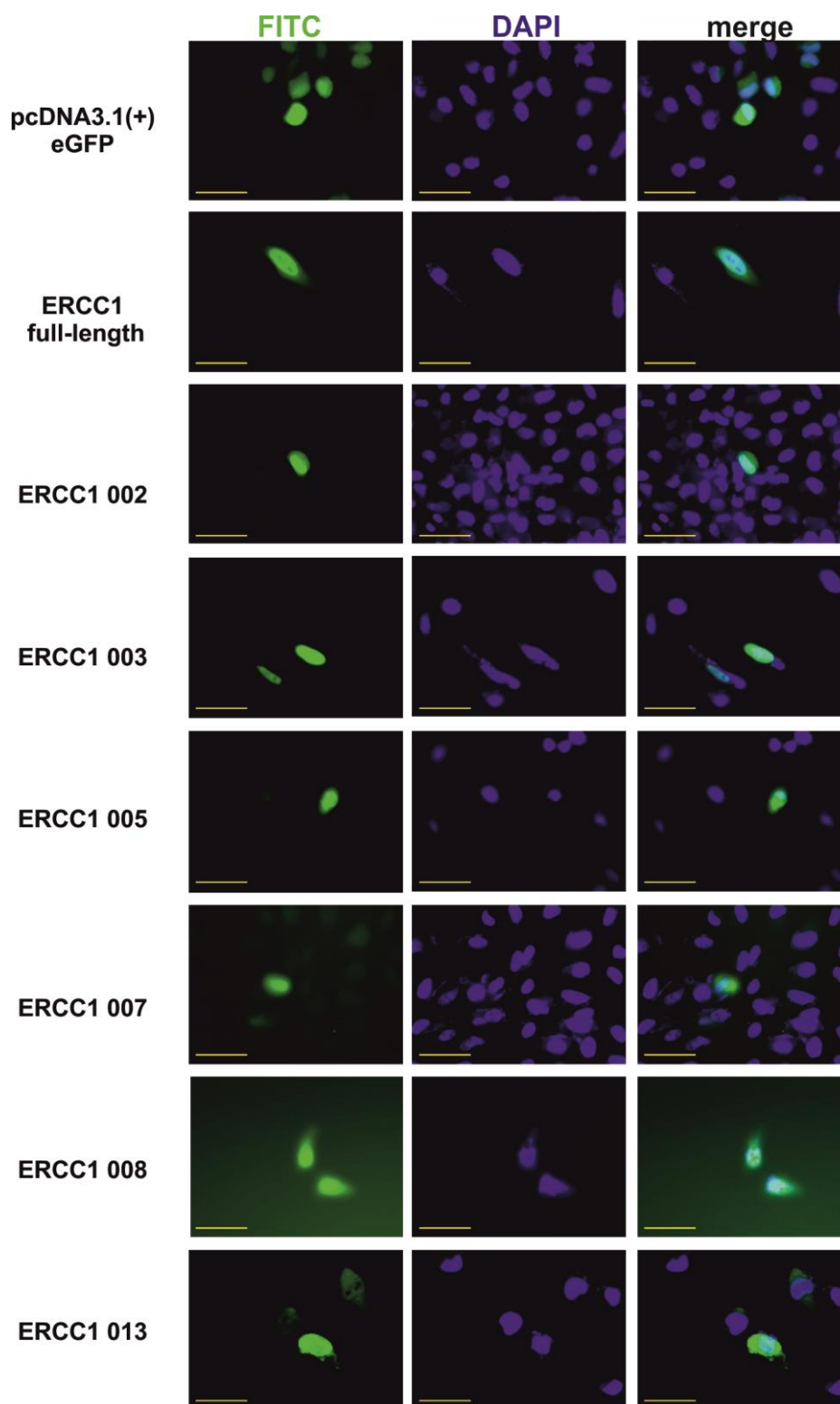


Figure A4: Subcellular localization of eGFP-tagged *ERCC1* isoforms and patient alleles

ERCC1 isoforms were cloned into an pcDNA3.1(+)-eGFP expression vector and overexpressed in HeLa cells for 48h to analyze subcellular localization. Additionally, DAPI staining was performed to visualize the nucleus. Scale bar = 50 μ m.

7 Bibliography

- Adair GM, Rolig RL, Moore-Faver D, Zabelshansky M, Wilson JH, Nairn RS (2000) Role of ERCC1 in removal of long non-homologous tails during targeted homologous recombination. *The EMBO journal* **19**: 5552-5561
- Ahmad A, Enzlin JH, Bhagwat NR, Wijgers N, Raams A, Appeldoorn E, Theil AF, JH JH, Vermeulen W, NG JJ, Scharer OD, Niedernhofer LJ (2010) Mislocalization of XPF-ERCC1 nuclease contributes to reduced DNA repair in XP-F patients. *PLoS genetics* **6**: e1000871
- 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. *Molecular and cellular biology* **28**: 5082-5092
- Al-Minawi AZ, Lee YF, Hakansson D, Johansson F, Lundin C, Saleh-Gohari N, Schultz N, Jenssen D, Bryant HE, Meuth M, Hinz JM, Helleday T (2009) The ERCC1/XPF endonuclease is required for completion of homologous recombination at DNA replication forks stalled by inter-strand cross-links. *Nucleic acids research* **37**: 6400-6413
- Al-Minawi AZ, Saleh-Gohari N, Helleday T (2008) The ERCC1/XPF endonuclease is required for efficient single-strand annealing and gene conversion in mammalian cells. *Nucleic acids research* **36**: 1-9
- Andersen SL, Bergstrahl DT, Kohl KP, LaRocque JR, Moore CB, Sekelsky J (2009) Drosophila MUS312 and the vertebrate ortholog BTBD12 interact with DNA structure-specific endonucleases in DNA repair and recombination. *Molecular cell* **35**: 128-135
- Andersson BS, Sadeghi T, Siciliano MJ, Legerski R, Murray D (1996) Nucleotide excision repair genes as determinants of cellular sensitivity to cyclophosphamide analogs. *Cancer chemotherapy and pharmacology* **38**: 406-416
- Arab HH, Wani G, Ray A, Shah ZI, Zhu Q, Wani AA (2010) Dissociation of CAK from core TFIIH reveals a functional link between XP-G/CS and the TFIIH disassembly state. *PLoS one* **5**: e11007
- Aracil M, Dauffenbach LM, Diez MM, Richeh R, Moneo V, Leal JF, Fernandez LF, Kerfoot CA, Galmarini CM (2013) Expression of XPG protein in human normal and tumor tissues. *International journal of clinical and experimental pathology* **6**: 199-211
- Araki M, Masutani C, Takemura M, Uchida A, Sugasawa K, Kondoh J, Ohkuma Y, Hanaoka F (2001) Centrosome protein centrin 2/caltractin 1 is part of the xeroderma pigmentosum group C complex that initiates global genome nucleotide excision repair. *The Journal of biological chemistry* **276**: 18665-18672
- Athar M, Li C, Kim AL, Spiegelman VS, Bickers DR (2014) Sonic hedgehog signaling in Basal cell nevus syndrome. *Cancer research* **74**: 4967-4975
- Baker BS, Carpenter AT, Ripoll P (1978) The Utilization during Mitotic Cell Division of Loci Controlling Meiotic Recombination and Disjunction in DROSOPHILA MELANOGASTER. *Genetics* **90**: 531-578
- Barakat KH, Jordheim LP, Perez-Pineiro R, Wishart D, Dumontet C, Tuszynski JA (2012) Virtual screening and biological evaluation of inhibitors targeting the XPA-ERCC1 interaction. *PLoS one* **7**: e51329
- Barrangou R, Fremaux C, Deveau H, Richards M, Boyaval P, Moineau S, Romero DA, Horvath P (2007) CRISPR provides acquired resistance against viruses in prokaryotes. *Science (New York, NY)* **315**: 1709-1712
- Bartek J, Lukas J (2007) DNA damage checkpoints: from initiation to recovery or adaptation. *Current opinion in cell biology* **19**: 238-245
- Batty DP, Wood RD (2000) Damage recognition in nucleotide excision repair of DNA. *Gene* **241**: 193-204
- Ben Rekaya M, Messaoud O, Talmoudi F, Nouria S, Ouragini H, Amouri A, Boussen H, Boubaker S, Mokni M, Mokhtar I, Abdelhak S, Zghal M (2009) High frequency of the V548A fs X572 XPC mutation in Tunisia: implication for molecular diagnosis. *Journal of human genetics* **54**: 426-429
- Bennardo N, Cheng A, Huang N, Stark JM (2008) Alternative-NHEJ is a mechanistically distinct pathway of mammalian chromosome break repair. *PLoS genetics* **4**: e1000110

- Bentley D (2002) The mRNA assembly line: transcription and processing machines in the same factory. *Current opinion in cell biology* **14**: 336-342
- Berge EO, Staalesen V, Straume AH, Lillehaug JR, Lonning PE (2010) Chk2 splice variants express a dominant-negative effect on the wild-type Chk2 kinase activity. *Biochimica et biophysica acta* **1803**: 386-395
- Berneburg M, Lehmann AR (2001) Xeroderma pigmentosum and related disorders: defects in DNA repair and transcription. *Advances in genetics* **43**: 71-102
- Berquist BR, Wilson DM, 3rd (2012) Pathways for repairing and tolerating the spectrum of oxidative DNA lesions. *Cancer letters* **327**: 61-72
- Bertram JS (2000) The molecular biology of cancer. *Molecular aspects of medicine* **21**: 167-223
- Bessho T, Sancar A, Thompson LH, Thelen MP (1997) Reconstitution of human excision nuclease with recombinant XPF-ERCC1 complex. *The Journal of biological chemistry* **272**: 3833-3837
- Bhagwat N, Olsen AL, Wang AT, Hanada K, Stuckert P, Kanaar R, D'Andrea A, Niedernhofer LJ, McHugh PJ (2009) XPF-ERCC1 participates in the Fanconi anemia pathway of cross-link repair. *Molecular and cellular biology* **29**: 6427-6437
- Biggerstaff M, Szymkowski DE, Wood RD (1993) Co-correction of the ERCC1, ERCC4 and xeroderma pigmentosum group F DNA repair defects in vitro. *The EMBO journal* **12**: 3685-3692
- Bilen N, Tekin SB, Topdagi O (2014) ERCC1 Expression in Non-Small Cell Lung and Esophageal Cancer. *The Eurasian journal of medicine* **46**: 84-88
- Bitinaite J, Wah DA, Aggarwal AK, Schildkraut I (1998) FokI dimerization is required for DNA cleavage. *Proceedings of the National Academy of Sciences of the United States of America* **95**: 10570-10575
- Blair CA, Zi X (2011) Potential molecular targeting of splice variants for cancer treatment. *Indian journal of experimental biology* **49**: 836-839
- Blankenburg S, Konig IR, Moessner R, Laspe P, Thoms KM, Krueger U, Khan SG, Westphal G, Berking C, Volkenandt M, Reich K, Neumann C, Ziegler A, Kraemer KH, Emmert S (2005) Assessment of 3 xeroderma pigmentosum group C gene polymorphisms and risk of cutaneous melanoma: a case-control study. *Carcinogenesis* **26**: 1085-1090
- Boettcher M, McManus MT (2015) Choosing the Right Tool for the Job: RNAi, TALEN, or CRISPR. *Molecular cell* **58**: 575-585
- Bogliolo M, Schuster B, Stoepker C, Derkunt B, Su Y, Raams A, Trujillo JP, Minguillon J, Ramirez MJ, Pujol R, Casado JA, Banos R, Rio P, Knies K, Zuniga S, Benitez J, Bueren JA, Jaspers NG, Scharer OD, de Winter JP, Schindler D, Surralles J (2013) Mutations in ERCC4, encoding the DNA-repair endonuclease XPF, cause Fanconi anemia. *American journal of human genetics* **92**: 800-806
- Bohr VA, Smith CA, Okumoto DS, Hanawalt PC (1985) DNA repair in an active gene: removal of pyrimidine dimers from the DHFR gene of CHO cells is much more efficient than in the genome overall. *Cell* **40**: 359-369
- Bohr VA, Stevnsner T, de Souza-Pinto NC (2002) Mitochondrial DNA repair of oxidative damage in mammalian cells. *Gene* **286**: 127-134
- Boulikas T (1997) Nuclear import of DNA repair proteins. *Anticancer research* **17**: 843-863
- Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical biochemistry* **72**: 248-254
- 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. *Journal of medical genetics* **48**: 168-176

- Brash DE, Rudolph JA, Simon JA, Lin A, McKenna GJ, Baden HP, Halperin AJ, Ponten J (1991) A role for sunlight in skin cancer: UV-induced p53 mutations in squamous cell carcinoma. *Proceedings of the National Academy of Sciences of the United States of America* **88**: 10124-10128
- Brinkman BM (2004) Splice variants as cancer biomarkers. *Clinical biochemistry* **37**: 584-594
- Brookman KW, Lamerdin JE, Thelen MP, Hwang M, Reardon JT, Sancar A, Zhou ZQ, Walter CA, Parris CN, Thompson LH (1996) ERCC4 (XPF) encodes a human nucleotide excision repair protein with eukaryotic recombination homologs. *Molecular and cellular biology* **16**: 6553-6562
- Brugmans L, Kanaar R, Essers J (2007) Analysis of DNA double-strand break repair pathways in mice. *Mutation research* **614**: 95-108
- Busse A, Rietz A, Schwartz S, Thiel E, Keilholz U (2009) An intron 9 containing splice variant of PAX2. *Journal of translational medicine* **7**: 36
- Campisi J, d'Adda di Fagagna F (2007) Cellular senescence: when bad things happen to good cells. *Nature reviews Molecular cell biology* **8**: 729-740
- Cermak T, Doyle EL, Christian M, Wang L, Zhang Y, Schmidt C, Baller JA, Somia NV, Bogdanove AJ, Voytas DF (2011) Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. *Nucleic acids research* **39**: e82
- Chang YF, Imam JS, Wilkinson MF (2007) The nonsense-mediated decay RNA surveillance pathway. *Annual review of biochemistry* **76**: 51-74
- Chen F, Nastasi A, Shen Z, Brenneman M, Crissman H, Chen DJ (1997) Cell cycle-dependent protein expression of mammalian homologs of yeast DNA double-strand break repair genes Rad51 and Rad52. *Mutation research* **384**: 205-211
- Chen J, Larochelle S, Li X, Suter B (2003) Xpd/Ercc2 regulates CAK activity and mitotic progression. *Nature* **424**: 228-232
- Cheo DL, Meira LB, Burns DK, Reis AM, Issac T, Friedberg EC (2000) Ultraviolet B radiation-induced skin cancer in mice defective in the Xpc, Trp53, and Apex (HAP1) genes: genotype-specific effects on cancer predisposition and pathology of tumors. *Cancer research* **60**: 1580-1584
- Chevalier BS, Stoddard BL (2001) Homing endonucleases: structural and functional insight into the catalysts of intron/intein mobility. *Nucleic acids research* **29**: 3757-3774
- Chiara MD, Palandjian L, Feld Kramer R, Reed R (1997) Evidence that U5 snRNP recognizes the 3' splice site for catalytic step II in mammals. *The EMBO journal* **16**: 4746-4759
- Choi YJ, Ryu KS, Ko YM, Chae YK, Pelton JG, Wemmer DE, Choi BS (2005) Biophysical characterization of the interaction domains and mapping of the contact residues in the XPF-ERCC1 complex. *The Journal of biological chemistry* **280**: 28644-28652
- Choueiri MB, Shen JP, Gross AM, Huang JK, Ideker T, Fanta P (2015) ERCC1 and TS Expression as Prognostic and Predictive Biomarkers in Metastatic Colon Cancer. *PloS one* **10**: e0126898
- Cimino GD, Gamper HB, Isaacs ST, Hearst JE (1985) Psoralens as photoactive probes of nucleic acid structure and function: organic chemistry, photochemistry, and biochemistry. *Annual review of biochemistry* **54**: 1151-1193
- Cimprich KA, Cortez D (2008) ATR: an essential regulator of genome integrity. *Nature reviews Molecular cell biology* **9**: 616-627
- Clauson C, Scharer OD, Niedernhofer L (2013) Advances in understanding the complex mechanisms of DNA interstrand cross-link repair. *Cold Spring Harbor perspectives in biology* **5**: a012732
- Cleaver JE, Lam ET, Revet I (2009) Disorders of nucleotide excision repair: the genetic and molecular basis of heterogeneity. *Nature reviews Genetics* **10**: 756-768

- Cleaver JE, Thompson LH, Richardson AS, States JC (1999) A summary of mutations in the UV-sensitive disorders: xeroderma pigmentosum, Cockayne syndrome, and trichothiodystrophy. *Human mutation* **14**: 9-22
- Clingen PH, Arlett CF, Hartley JA, Parris CN (2007) Chemosensitivity of primary human fibroblasts with defective unhooking of DNA interstrand cross-links. *Experimental cell research* **313**: 753-760
- Cockayne EA (1936) Dwarfism with retinal atrophy and deafness. *Archives of disease in childhood* **11**: 1-8
- Coin F, Oksenysh V, Egly JM (2007) Distinct roles for the XPB/p52 and XPD/p44 subcomplexes of TFIIH in damaged DNA opening during nucleotide excision repair. *Molecular cell* **26**: 245-256
- Colgin LM, Wilkinson C, Englezou A, Kilian A, Robinson MO, Reddel RR (2000) The hTERTalpha splice variant is a dominant negative inhibitor of telomerase activity. *Neoplasia (New York, NY)* **2**: 426-432
- Colmegna B, Ubaldi S, Frapolli R, Licandro SA, Panini N, Galmarini CM, Badri N, Spanswick VJ, Bingham JP, Kiakos K, Erba E, Hartley JA, D'Incalci M (2015) Increased sensitivity to platinum drugs of cancer cells with acquired resistance to trabectedin. *British journal of cancer* **113**: 1687-1693
- Compe E, Egly JM (2012) TFIIH: when transcription met DNA repair. *Nature reviews Molecular cell biology* **13**: 343-354
- Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, Hsu PD, Wu X, Jiang W, Marraffini LA, Zhang F (2013) Multiplex genome engineering using CRISPR/Cas systems. *Science (New York, NY)* **339**: 819-823
- Cornu TI, Thibodeau-Beganny S, Guhl E, Alwin S, Eichtinger M, Joung JK, Cathomen T (2008) DNA-binding specificity is a major determinant of the activity and toxicity of zinc-finger nucleases. *Molecular therapy : the journal of the American Society of Gene Therapy* **16**: 352-358
- Couve-Privat S, Le Bret M, Traiffort E, Queille S, Coulombe J, Bouadjar B, Avril MF, Ruat M, Sarasin A, Daya-Grosjean L (2004) Functional analysis of novel sonic hedgehog gene mutations identified in basal cell carcinomas from xeroderma pigmentosum patients. *Cancer research* **64**: 3559-3565
- Crick F (1970) Central dogma of molecular biology. *Nature* **227**: 561-563
- Damia G, Imperatori L, Stefanini M, D'Incalci M (1996) Sensitivity of CHO mutant cell lines with specific defects in nucleotide excision repair to different anti-cancer agents. *International journal of cancer* **66**: 779-783
- Daya-Grosjean L (2008) Xeroderma pigmentosum and skin cancer. *Advances in experimental medicine and biology* **637**: 19-27
- de Buendia PG (1998) Search for DNA repair pathways in *Drosophila melanogaster*. *Mutation research* **407**: 67-84
- de Laat WL, Appeldoorn E, Jaspers NG, Hoeijmakers JH (1998a) DNA structural elements required for ERCC1-XPF endonuclease activity. *The Journal of biological chemistry* **273**: 7835-7842
- de Laat WL, Appeldoorn E, Sugasawa K, Weterings E, Jaspers NG, Hoeijmakers JH (1998b) DNA-binding polarity of human replication protein A positions nucleases in nucleotide excision repair. *Genes & development* **12**: 2598-2609
- de Laat WL, Sijbers AM, Odijk H, Jaspers NG, Hoeijmakers JH (1998c) Mapping of interaction domains between human repair proteins ERCC1 and XPF. *Nucleic acids research* **26**: 4146-4152
- 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. *Molecular and cellular biology* **20**: 7980-7990
- 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 research* **30**: 3848-3856
- de Vries E, Coebergh JW (2004) Cutaneous malignant melanoma in Europe. *European journal of cancer (Oxford, England : 1990)* **40**: 2355-2366

- Deltcheva E, Chylinski K, Sharma CM, Gonzales K, Chao Y, Pirzada ZA, Eckert MR, Vogel J, Charpentier E (2011) CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature* **471**: 602-607
- Deng D, Yan C, Pan X, Mahfouz M, Wang J, Zhu JK, Shi Y, Yan N (2012) Structural basis for sequence-specific recognition of DNA by TAL effectors. *Science (New York, NY)* **335**: 720-723
- DiGiovanna JJ, Patronas N, Katz D, Abangan D, Kraemer KH (1998) Xeroderma pigmentosum: spinal cord astrocytoma with 9-year survival after radiation and isotretinoin therapy. *Journal of cutaneous medicine and surgery* **2**: 153-158
- Doherty AJ, Serpell LC, Ponting CP (1996) The helix-hairpin-helix DNA-binding motif: a structural basis for non-sequence-specific recognition of DNA. *Nucleic acids research* **24**: 2488-2497
- Dunand-Sauthier I, Hohl M, Thorel F, Jaquier-Gubler P, Clarkson SG, Scharer OD (2005) The spacer region of XPG mediates recruitment to nucleotide excision repair complexes and determines substrate specificity. *The Journal of biological chemistry* **280**: 7030-7037
- Dupuy A, Valton J, Leduc S, Armier J, Galetto R, Gouble A, Lebuhotel C, Stary A, Paques F, Duchateau P, Sarasin A, Daboussi F (2013) Targeted gene therapy of xeroderma pigmentosum cells using meganuclease and TALEN. *PloS one* **8**: e78678
- Egly JM, Coin F (2011) A history of TFIIH: two decades of molecular biology on a pivotal transcription/repair factor. *DNA repair* **10**: 714-721
- Ehrhart JC, Gosselet FP, Culerrier RM, Sarasin A (2003) UVB-induced mutations in human key gatekeeper genes governing signalling pathways and consequences for skin tumourigenesis. *Photochemical & photobiological sciences : Official journal of the European Photochemistry Association and the European Society for Photobiology* **2**: 825-834
- Emmert S, Leibel D, Runger TM (2006) Syndromes with genetic instability: model diseases for (skin) cancerogenesis. *Journal der Deutschen Dermatologischen Gesellschaft = Journal of the German Society of Dermatology : JDDG* **4**: 721-731
- Emmert S, Schneider TD, Khan SG, Kraemer KH (2001) The human XPG gene: gene architecture, alternative splicing and single nucleotide polymorphisms. *Nucleic acids research* **29**: 1443-1452
- Emmert S, Schon MP, Haenssle HA (2014) Molecular biology of basal and squamous cell carcinomas. *Advances in experimental medicine and biology* **810**: 234-252
- Emmert S, Slor H, Busch DB, Batko S, Albert RB, Coleman D, Khan SG, Abu-Libdeh B, DiGiovanna JJ, Cunningham BB, Lee MM, Crollick J, Inui H, Ueda T, Hedayati M, Grossman L, Shahlavi T, Cleaver JE, Kraemer KH (2002) Relationship of neurologic degeneration to genotype in three xeroderma pigmentosum group G patients. *The Journal of investigative dermatology* **118**: 972-982
- Enoiu M, Jiricny J, Scharer OD (2012) Repair of cisplatin-induced DNA interstrand crosslinks by a replication-independent pathway involving transcription-coupled repair and translesion synthesis. *Nucleic acids research* **40**: 8953-8964
- Enzlin JH, Scharer OD (2002) The active site of the DNA repair endonuclease XPF-ERCC1 forms a highly conserved nuclease motif. *The EMBO journal* **21**: 2045-2053
- Evans E, Fellows J, Coffey A, Wood RD (1997a) Open complex formation around a lesion during nucleotide excision repair provides a structure for cleavage by human XPG protein. *The EMBO journal* **16**: 625-638
- Evans E, Moggs JG, Hwang JR, Egly JM, Wood RD (1997b) Mechanism of open complex and dual incision formation by human nucleotide excision repair factors. *The EMBO journal* **16**: 6559-6573
- Evers B, Jastrzebski K, Heijmans JP, Grenrum W, Beijersbergen RL, Bernards R (2016) CRISPR knockout screening outperforms shRNA and CRISPRi in identifying essential genes. **34**: 631-633
- Fadda E (2016) Role of the XPA protein in the NER pathway: A perspective on the function of structural disorder in macromolecular assembly. *Computational and structural biotechnology journal* **14**: 78-85

- Faridounnia M, Wienk H, Kovacic L, Folkers GE, Jaspers NG, Kaptein R, Hoeijmakers JH, Boelens R (2015) The Cerebro-oculo-facio-skeletal Syndrome Point Mutation F231L in the ERCC1 DNA Repair Protein Causes Dissociation of the ERCC1-XPF Complex. *The Journal of biological chemistry* **290**: 20541-20555
- Fekairi S, Scaglione S, Chahwan C, Taylor ER, Tissier A, Coulon S, Dong MQ, Ruse C, Yates JR, 3rd, Russell P, Fuchs RP, McGowan CH, Gaillard PH (2009) Human SLX4 is a Holliday junction resolvase subunit that binds multiple DNA repair/recombination endonucleases. *Cell* **138**: 78-89
- Field LL, Bonnevie-Nielsen V, Pociot F, Lu S, Nielsen TB, Beck-Nielsen H (2005) OAS1 splice site polymorphism controlling antiviral enzyme activity influences susceptibility to type 1 diabetes. *Diabetes* **54**: 1588-1591
- Fischer JA, Muller-Weeks S, Caradonna S (2004) Proteolytic degradation of the nuclear isoform of uracil-DNA glycosylase occurs during the S phase of the cell cycle. *DNA repair* **3**: 505-513
- Fisher LA, Bessho M, Bessho T (2008) Processing of a psoralen DNA interstrand cross-link by XPF-ERCC1 complex in vitro. *The Journal of biological chemistry* **283**: 1275-1281
- Fisher LA, Bessho M, Wakasugi M, Matsunaga T, Bessho T (2011a) Role of interaction of XPF with RPA in nucleotide excision repair. *Journal of molecular biology* **413**: 337-346
- Fisher LA, Samson L, Bessho T (2011b) Removal of reactive oxygen species-induced 3'-blocked ends by XPF-ERCC1. *Chemical research in toxicology* **24**: 1876-1881
- Fishman-Lobell J, Haber JE (1992) Removal of nonhomologous DNA ends in double-strand break recombination: the role of the yeast ultraviolet repair gene RAD1. *Science (New York, NY)* **258**: 480-484
- Fox-Walsh KL, Dou Y, Lam BJ, Hung SP, Baldi PF, Hertel KJ (2005) The architecture of pre-mRNAs affects mechanisms of splice-site pairing. *Proceedings of the National Academy of Sciences of the United States of America* **102**: 16176-16181
- Friboulet L, Postel-Vinay S, Sourisseau T, Adam J, Stoclin A, Ponsonnailles F, Dorvault N, Commo F, Saulnier P, Salome-Desmoulez S, Pottier G, Andre F, Kroemer G, Soria JC, Olaussen KA (2013) ERCC1 function in nuclear excision and interstrand crosslink repair pathways is mediated exclusively by the ERCC1-202 isoform. *Cell cycle (Georgetown, Tex)* **12**: 3298-3306
- Fridman JS, Lowe SW (2003) Control of apoptosis by p53. *Oncogene* **22**: 9030-9040
- Fujiwara Y, Masutani C, Mizukoshi T, Kondo J, Hanaoka F, Iwai S (1999) Characterization of DNA recognition by the human UV-damaged DNA-binding protein. *The Journal of biological chemistry* **274**: 20027-20033
- Gaillard PH, Wood RD (2001) Activity of individual ERCC1 and XPF subunits in DNA nucleotide excision repair. *Nucleic acids research* **29**: 872-879
- Gaj T, Gersbach CA, Barbas CF, 3rd (2013) ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. *Trends in biotechnology* **31**: 397-405
- Galante PA, Sakabe NJ, Kirschbaum-Slager N, de Souza SJ (2004) Detection and evaluation of intron retention events in the human transcriptome. *RNA (New York, NY)* **10**: 757-765
- Gary R, Kim K, Cornelius HL, Park MS, Matsumoto Y (1999) Proliferating cell nuclear antigen facilitates excision in long-patch base excision repair. *The Journal of biological chemistry* **274**: 4354-4363
- Gary R, Ludwig DL, Cornelius HL, MacInnes MA, Park MS (1997) The DNA repair endonuclease XPG binds to proliferating cell nuclear antigen (PCNA) and shares sequence elements with the PCNA-binding regions of FEN-1 and cyclin-dependent kinase inhibitor p21. *The Journal of biological chemistry* **272**: 24522-24529
- Gassmann M, Grenacher B, Rohde B, Vogel J (2009) Quantifying Western blots: pitfalls of densitometry. *Electrophoresis* **30**: 1845-1855
- Gentile F, Tuszyński JA, Barakat KH (2016) New design of nucleotide excision repair (NER) inhibitors for combination cancer therapy. *Journal of molecular graphics & modelling* **65**: 71-82

- Gerhard DS, Wagner L, Feingold EA, Shenmen CM, Grouse LH, Schuler G, Klein SL, Old S, Rasooly R, Good P, Guyer M, Peck AM, Derge JG, Lipman D, Collins FS, Jang W, Sherry S, Feolo M, Misquitta L, Lee E, Rotmistrovsky K, Greenhut SF, Schaefer CF, Buetow K, Bonner TI, Haussler D, Kent J, Kiekhaus M, Furey T, Brent M, Prange C, Schreiber K, Shapiro N, Bhat NK, Hopkins RF, Hsie F, Driscoll T, Soares MB, Casavant TL, Scheetz TE, Brown-stein MJ, Usdin TB, Toshiyuki S, Carninci P, Piao Y, Dudekula DB, Ko MS, Kawakami K, Suzuki Y, Sugano S, Gruber CE, Smith MR, Simmons B, Moore T, Waterman R, Johnson SL, Ruan Y, Wei CL, Mathavan S, Gunaratne PH, Wu J, Garcia AM, Hulyk SW, Fuh E, Yuan Y, Sneed A, Kowis C, Hodgson A, Muzny DM, McPherson J, Gibbs RA, Fahey J, Helton E, Kettelman M, Madan A, Rodrigues S, Sanchez A, Whiting M, Madari A, Young AC, Wetherby KD, Granite SJ, Kwong PN, Brinkley CP, Pearson RL, Bouffard GG, Blakesly RW, Green ED, Dickson MC, Rodriguez AC, Grimwood J, Schmutz J, Myers RM, Butterfield YS, Griffith M, Griffith OL, Krzywinski MI, Liao N, Morin R, Palmquist D, Petrescu AS, Skalska U, Smailus DE, Stott JM, Schnerch A, Schein JE, Jones SJ, Holt RA, Baross A, Marra MA, Clifton S, Makowski KA, Bosak S, Malek J (2004) The status, quality, and expansion of the NIH full-length cDNA project: the Mammalian Gene Collection (MGC). *Genome research* **14**: 2121-2127
- Gerhard R, Carvalho A, Carneiro V, Bento RS, Uemura G, Gomes M, Albergaria A, Schmitt F (2013) Clinicopathological significance of ERCC1 expression in breast cancer. *Pathology, research and practice* **209**: 331-336
- Gervais V, Lamour V, Jawhari A, Frindel F, Wasielewski E, Dubaele S, Egly JM, Thierry JC, Kieffer B, Poterszman A (2004) TFIIF contains a PH domain involved in DNA nucleotide excision repair. *Nature structural & molecular biology* **11**: 616-622
- Giannattasio M, Follonier C, Tourriere H, Puddu F, Lazzaro F, Pasero P, Lopes M, Plevani P, Muzi-Falconi M (2010) Exo1 competes with repair synthesis, converts NER intermediates to long ssDNA gaps, and promotes checkpoint activation. *Molecular cell* **40**: 50-62
- Giannelli F, Avery J, Polani PE, Terrell C, Giammusso V (1981) Xeroderma Pigmentosum and medulloblastoma: chromosomal damage to lymphocytes during radiotherapy. *Radiation research* **88**: 194-208
- Giglia-Mari G, Zotter A, Vermeulen W (2011) DNA damage response. *Cold Spring Harbor perspectives in biology* **3**: a000745
- Giglia G, Dumaz N, Drougard C, Avril MF, Daya-Grosjean L, Sarasin A (1998) p53 mutations in skin and internal tumors of xeroderma pigmentosum patients belonging to the complementation group C. *Cancer research* **58**: 4402-4409
- Gilljam KM, Muller R, Liabakk NB, Otterlei M (2012) Nucleotide excision repair is associated with the replisome and its efficiency depends on a direct interaction between XPA and PCNA. *PloS one* **7**: e49199
- Goh AM, Coffill CR, Lane DP (2011) The role of mutant p53 in human cancer. *The Journal of pathology* **223**: 116-126
- Graveley BR (2001) Alternative splicing: increasing diversity in the proteomic world. *Trends in genetics : TIG* **17**: 100-107
- Greenblatt MS, Beaudet JG, Gump JR, Godin KS, Trombley L, Koh J, Bond JP (2003) Detailed computational study of p53 and p16: using evolutionary sequence analysis and disease-associated mutations to predict the functional consequences of allelic variants. *Oncogene* **22**: 1150-1163
- Gregg SQ, Robinson AR, Niedernhofer LJ (2011) Physiological consequences of defects in ERCC1-XPF DNA repair endonuclease. *DNA repair* **10**: 781-791
- Guervilly JH, Takedachi A, Naim V, Scaglione S, Chawhan C, Lovera Y, Despras E, Kuraoka I, Kannouche P, Rosselli F, Gaillard PH (2015) The SLX4 complex is a SUMO E3 ligase that impacts on replication stress outcome and genome stability. *Molecular cell* **57**: 123-137
- Gupta AK, Anderson TF (1987) Psoralen photochemotherapy. *Journal of the American Academy of Dermatology* **17**: 703-734
- Guzder SN, Torres-Ramos C, Johnson RE, Haracska L, Prakash L, Prakash S (2004) Requirement of yeast Rad1-Rad10 nuclease for the removal of 3'-blocked termini from DNA strand breaks induced by reactive oxygen species. *Genes & development* **18**: 2283-2291

- Habraken Y, Sung P, Prakash L, Prakash S (1993) Yeast excision repair gene RAD2 encodes a single-stranded DNA endonuclease. *Nature* **366**: 365-368
- Halazonetis TD, Gorgoulis VG, Bartek J (2008) An oncogene-induced DNA damage model for cancer development. *Science (New York, NY)* **319**: 1352-1355
- Hanada K, Budzowska M, Modesti M, Maas A, Wyman C, Essers J, Kanaar R (2006) The structure-specific endonuclease Mus81-Eme1 promotes conversion of interstrand DNA crosslinks into double-strand breaks. *The EMBO journal* **25**: 4921-4932
- Hanahan D, Weinberg RA (2000) The hallmarks of cancer. *Cell* **100**: 57-70
- Hart T, Brown KR, Sircoulomb F, Rottapel R, Moffat J (2014) Measuring error rates in genomic perturbation screens: gold standards for human functional genomics. *Molecular systems biology* **10**: 733
- Hayakawa H, Ishizaki K, Inoue M, Yagi T, Sekiguchi M, Takebe H (1981) Repair of ultraviolet radiation damage in xeroderma pigmentosum cells belonging to complementation group F. *Mutation research* **80**: 381-388
- Hayes M, Lan C, Yan J, Xie Y, Gray T, Amirkhan RH, Dowell JE (2011) ERCC1 expression and outcomes in head and neck cancer treated with concurrent cisplatin and radiation. *Anticancer research* **31**: 4135-4139
- Hearst JE (1981) Psoralen photochemistry and nucleic acid structure. *The Journal of investigative dermatology* **77**: 39-44
- Hebra F, Kaposi M (1874) On diseases of the skin including exanthemata. *The New Sydenham Society* **61**: 252-258
- Hecht SM (2000) Bleomycin: new perspectives on the mechanism of action. *Journal of natural products* **63**: 158-168
- Herskowitz I (1987) Functional inactivation of genes by dominant negative mutations. *Nature* **329**: 219-222
- Heyer WD (2004) Recombination: Holliday junction resolution and crossover formation. *Current biology : CB* **14**: R56-58
- Hirai Y, Kodama Y, Moriwaki S, Noda A, Cullings HM, Macphee DG, Kodama K, Mabuchi K, Kraemer KH, Land CE, Nakamura N (2006) Heterozygous individuals bearing a founder mutation in the XPA DNA repair gene comprise nearly 1% of the Japanese population. *Mutation research* **601**: 171-178
- Hodskinson MR, Silhan J, Crossan GP, Garaycochea JI, Mukherjee S, Johnson CM, Scharer OD, Patel KJ (2014) Mouse SLX4 is a tumor suppressor that stimulates the activity of the nuclease XPF-ERCC1 in DNA crosslink repair. *Molecular cell* **54**: 472-484
- Hoeijmakers JH (2009) DNA damage, aging, and cancer. *The New England journal of medicine* **361**: 1475-1485
- Hohl M, Thorel F, Clarkson SG, Scharer OD (2003) Structural determinants for substrate binding and catalysis by the structure-specific endonuclease XPG. *The Journal of biological chemistry* **278**: 19500-19508
- Hollstein M, Sidransky D, Vogelstein B, Harris CC (1991) p53 mutations in human cancers. *Science (New York, NY)* **253**: 49-53
- Honda A, Valogne Y, Bou Nader M, Brechot C, Faivre J (2012) An intron-retaining splice variant of human cyclin A2, expressed in adult differentiated tissues, induces a G1/S cell cycle arrest in vitro. *PloS one* **7**: e39249
- 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 research* **45**: 1737-1743
- Hsu PD, Scott DA, Weinstein JA, Ran FA, Konermann S, Agarwala V, Li Y, Fine EJ, Wu X, Shalem O, Cradick TJ, Marraffini LA, Bao G, Zhang F (2013) DNA targeting specificity of RNA-guided Cas9 nucleases. *Nature biotechnology* **31**: 827-832

- Hu W, Pan J, Zhao P, Yang G, Yang S (2014) Genetic polymorphisms in XPG could predict clinical outcome of platinum-based chemotherapy for advanced non-small cell lung cancer. *Tumour biology : the journal of the International Society for Oncodevelopmental Biology and Medicine* **35**: 5561-5567
- Huen MS, Chen J (2008) The DNA damage response pathways: at the crossroad of protein modifications. *Cell research* **18**: 8-16
- Inciarte MR, Lazaro JM, Salas M, Vinuela E (1976) Physical map of bacteriophage phi29 DNA. *Virology* **74**: 314-323
- Ito S, Kuraoka I, Chymkowitch P, Compe E, Takedachi A, Ishigami C, Coin F, Egly JM, Tanaka K (2007) XPG stabilizes TFIIH, allowing transactivation of nuclear receptors: implications for Cockayne syndrome in XP-G/CS patients. *Molecular cell* **26**: 231-243
- Ivanov EL, Haber JE (1995) RAD1 and RAD10, but not other excision repair genes, are required for double-strand break-induced recombination in *Saccharomyces cerevisiae*. *Molecular and cellular biology* **15**: 2245-2251
- Iwao M, Morisaki H, Morisaki T (2004) Single-nucleotide polymorphism g.1548G > A (E469K) in human ICAM-1 gene affects mRNA splicing pattern and TPA-induced apoptosis. *Biochemical and biophysical research communications* **317**: 729-735
- Jaakkola E, Mustonen A, Olsen P, Miettinen S, Savuoja T, Raams A, Jaspers NG, Shao H, Wu BL, Ignatius J (2010) ERCC6 founder mutation identified in Finnish patients with COFS syndrome. *Clinical genetics* **78**: 541-547
- Jamison SF, Crow A, Garcia-Blanco MA (1992) The spliceosome assembly pathway in mammalian extracts. *Molecular and cellular biology* **12**: 4279-4287
- Jaspers NG, Raams A, Silengo MC, Wijgers N, Niedernhofer LJ, Robinson AR, Giglia-Mari G, Hoogstraten D, Kleijer WJ, Hoeijmakers JH, Vermeulen W (2007) First reported patient with human ERCC1 deficiency has cerebro-oculo-facio-skeletal syndrome with a mild defect in nucleotide excision repair and severe developmental failure. *American journal of human genetics* **80**: 457-466
- Jeppesen DK, Bohr VA, Stevnsner T (2011) DNA repair deficiency in neurodegeneration. *Progress in neurobiology* **94**: 166-200
- Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E (2012) A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science (New York, NY)* **337**: 816-821
- Jones JC, Zhen WP, Reed E, Parker RJ, Sancar A, Bohr VA (1991) Gene-specific formation and repair of cisplatin intrastrand adducts and interstrand cross-links in Chinese hamster ovary cells. *The Journal of biological chemistry* **266**: 7101-7107
- Jordheim LP, Barakat KH, Heinrich-Balard L, Matera EL, Cros-Perrial E, Bouledrak K, El Sabeh R, Perez-Pineiro R, Wishart DS, Cohen R, Tuszynski J, Dumontet C (2013) Small molecule inhibitors of ERCC1-XPF protein-protein interaction synergize alkylating agents in cancer cells. *Molecular pharmacology* **84**: 12-24
- Karran P (2000) DNA double strand break repair in mammalian cells. *Current opinion in genetics & development* **10**: 144-150
- Kashiyama K, Nakazawa Y, Pilz DT, Guo C, Shimada M, Sasaki K, Fawcett H, Wing JF, Lewin SO, Carr L, Li TS, Yoshiura K, Utani A, Hirano A, Yamashita S, Greenblatt D, Nardo T, Stefanini M, McGibbon D, Sarkany R, Fasshi H, Takahashi Y, Nagayama Y, Mitsutake N, Lehmann AR, Ogi T (2013) Malfunction of nuclease ERCC1-XPF results in diverse clinical manifestations and causes Cockayne syndrome, xeroderma pigmentosum, and Fanconi anemia. *American journal of human genetics* **92**: 807-819
- Kastan MB, Bartek J (2004) Cell-cycle checkpoints and cancer. *Nature* **432**: 316-323
- Khan SG, Metter EJ, Tarone RE, Bohr VA, Grossman L, Hedayati M, Bale SJ, Emmert S, Kraemer KH (2000) A new xeroderma pigmentosum group C poly(AT) insertion/deletion polymorphism. *Carcinogenesis* **21**: 1821-1825

- Khan SG, Muniz-Medina V, Shahlavi T, Baker CC, Inui H, Ueda T, Emmert S, Schneider TD, Kraemer KH (2002) The human XPC DNA repair gene: arrangement, splice site information content and influence of a single nucleotide polymorphism in a splice acceptor site on alternative splicing and function. *Nucleic acids research* **30**: 3624-3631
- Kilian A, Bowtell DD, Abud HE, Hime GR, Venter DJ, Keese PK, Duncan EL, Reddel RR, Jefferson RA (1997) Isolation of a candidate human telomerase catalytic subunit gene, which reveals complex splicing patterns in different cell types. *Human molecular genetics* **6**: 2011-2019
- Kim JS, Lee HJ, Carroll D (2010) Genome editing with modularly assembled zinc-finger nucleases. *Nature methods* **7**: 91; author reply 91-92
- Kim Y, Kweon J, Kim A, Chon JK, Yoo JY, Kim HJ, Kim S, Lee C, Jeong E, Chung E, Kim D, Lee MS, Go EM, Song HJ, Kim H, Cho N, Bang D, Kim S, Kim JS (2013) A library of TAL effector nucleases spanning the human genome. *Nature biotechnology* **31**: 251-258
- Kim YG, Cha J, Chandrasegaran S (1996) Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. *Proceedings of the National Academy of Sciences of the United States of America* **93**: 1156-1160
- Kisker C, Kuper J, Van Houten B (2013) Prokaryotic nucleotide excision repair. *Cold Spring Harbor perspectives in biology* **5**: a012591
- Kleijer WJ, Laugel V, Berneburg M, Nardo T, Fawcett H, Gratchev A, Jaspers NG, Sarasin A, Stefanini M, Lehmann AR (2008) Incidence of DNA repair deficiency disorders in western Europe: Xeroderma pigmentosum, Cockayne syndrome and trichothiodystrophy. *DNA repair* **7**: 744-750
- Klein Douwel D, Boonen RA, Long DT, Szybowska AA, Raschle M, Walter JC, Knipscheer P (2014) XPF-ERCC1 acts in Unhooking DNA interstrand crosslinks in cooperation with FANCD2 and FANCP/SLX4. *Molecular cell* **54**: 460-471
- Klungland A, Hoss M, Gunz D, Constantinou A, Clarkson SG, Doetsch PW, Bolton PH, Wood RD, Lindahl T (1999) Base excision repair of oxidative DNA damage activated by XPG protein. *Molecular cell* **3**: 33-42
- Knauf JA, Pendergrass SH, Marrone BL, Strniste GF, MacInnes MA, Park MS (1996) Multiple nuclear localization signals in XPG nuclease. *Mutation research* **363**: 67-75
- Kobayashi N, Katsumi S, Imoto K, Nakagawa A, Miyagawa S, Furumura M, Mori T (2001) Quantitation and visualization of ultraviolet-induced DNA damage using specific antibodies: application to pigment cell biology. *Pigment cell research* **14**: 94-102
- Kottemann MC, Smogorzewska A (2013) Fanconi anaemia and the repair of Watson and Crick DNA crosslinks. *Nature* **493**: 356-363
- Kozak M (1983) Comparison of initiation of protein synthesis in procaryotes, eucaryotes, and organelles. *Microbiological reviews* **47**: 1-45
- Kraemer KH, Coon HG, Petinga RA, Barrett SF, Rahe AE, Robbins JH (1975a) Genetic heterogeneity in xeroderma pigmentosum: complementation groups and their relationship to DNA repair rates. *Proceedings of the National Academy of Sciences of the United States of America* **72**: 59-63
- Kraemer KH, De Weerd-Kastelein EA, Robbins JH, Keijzer W, Barrett SF, Petinga RA, Bootsma D (1975b) Five complementation groups in xeroderma pigmentosum. *Mutation research* **33**: 327-340
- Kraemer KH, Lee MM, Andrews AD, Lambert WC (1994) The role of sunlight and DNA repair in melanoma and nonmelanoma skin cancer. The xeroderma pigmentosum paradigm. *Archives of dermatology* **130**: 1018-1021
- Kraemer KH, Lee MM, Scotto J (1987) Xeroderma pigmentosum. Cutaneous, ocular, and neurologic abnormalities in 830 published cases. *Archives of dermatology* **123**: 241-250
- Kraemer KH, Patronas NJ, Schiffmann R, Brooks BP, Tamura D, DiGiovanna JJ (2007) Xeroderma pigmentosum, trichothiodystrophy and Cockayne syndrome: a complex genotype-phenotype relationship. *Neuroscience* **145**: 1388-1396

- Kraemer KH, Slor H (1985) Xeroderma pigmentosum. *Clinics in dermatology* **3**: 33-69
- Krasikova YS, Rechkunova NI, Maltseva EA, Craescu CT, Petrusseva IO, Lavrik OI (2012) Influence of centrin 2 on the interaction of nucleotide excision repair factors with damaged DNA. *Biochemistry Biokhimiia* **77**: 346-353
- Krasikova YS, Rechkunova NI, Maltseva EA, Petrusseva IO, Silnikov VN, Zatsepin TS, Oretskaya TS, Scharer OD, Lavrik OI (2008) Interaction of nucleotide excision repair factors XPC-HR23B, XPA, and RPA with damaged DNA. *Biochemistry Biokhimiia* **73**: 886-896
- Krawczak M, Reiss J, Cooper DN (1992) The mutational spectrum of single base-pair substitutions in mRNA splice junctions of human genes: causes and consequences. *Human genetics* **90**: 41-54
- Kuraoka I, Kobertz WR, Ariza RR, Biggerstaff M, Essigmann JM, Wood RD (2000) Repair of an interstrand DNA cross-link initiated by ERCC1-XPF repair/recombination nuclease. *The Journal of biological chemistry* **275**: 26632-26636
- Laugel V (1993) Cockayne Syndrome. In *GeneReviews(R)*, Pagon RA, Adam MP, Ardinger HH, Wallace SE, Amemiya A, Bean LJH, Bird TD, Ledbetter N, Mefford HC, Smith RJH, Stephens K (eds). Seattle (WA): University of Washington, Seattle University of Washington, Seattle. GeneReviews is a registered trademark of the University of Washington, Seattle.
- Le Hir H, Nott A, Moore MJ (2003) How introns influence and enhance eukaryotic gene expression. *Trends in biochemical sciences* **28**: 215-220
- 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**: 622-632
- Lee JH, Park CJ, Arunkumar AI, Chazin WJ, Choi BS (2003) NMR study on the interaction between RPA and DNA decamer containing cis-syn cyclobutane pyrimidine dimer in the presence of XPA: implication for damage verification and strand-specific dual incision in nucleotide excision repair. *Nucleic acids research* **31**: 4747-4754
- Lehmann AR, McGibbon D, Stefanini M (2011) Xeroderma pigmentosum. *Orphanet journal of rare diseases* **6**: 70
- Lehmann J (2013) Characterization of novel Xeroderma pigmentosum Group G spontaneous mRNA splicevariants. [Master thesis]; Department of Dermatology, Venereology and Allergology, Georg-August-University Goettingen.
- Lehmann J, Schubert S, Emmert S (2014a) Xeroderma pigmentosum: diagnostic procedures, interdisciplinary patient care, and novel therapeutic approaches. *Journal der Deutschen Dermatologischen Gesellschaft = Journal of the German Society of Dermatology : JDDG* **12**: 867-872
- Lehmann J, Schubert S, Schafer A, Apel A, Laspe P, Schiller S, Ohlenbusch A, Gratchev A, Emmert S (2014b) An unusual mutation in the XPG gene leads to an internal in-frame deletion and a XP/CS complex phenotype. *The British journal of dermatology* **171**: 903-905
- Lehmann J, Schubert S, Schafer A, Laspe P, Haenssle HA, Ohlenbusch A, Gratchev A, Emmert S (2015) A novel mutation in the XPA gene results in two truncated protein variants and leads to a severe XP/neurological symptoms phenotype. *Journal of the European Academy of Dermatology and Venereology : JEADV* **29**: 2479-2482
- Lehmann J, Seebode C, Smolorz S, Schubert S, Emmert S (2017) XPF knockout via CRISPR/Cas9 reveals that ERCC1 is retained in the cytoplasm without its heterodimer partner XPF. *Cellular and Molecular Life Sciences* doi: 10.1007/s00018-017-2455-7
- Li C, Yin M, Wang LE, Amos CI, Zhu D, Lee JE, Gershenwald JE, Grimm EA, Wei Q (2013a) Polymorphisms of nucleotide excision repair genes predict melanoma survival. *The Journal of investigative dermatology* **133**: 1813-1821
- Li J, Zhang Y, Chen KL, Shan QW, Wang YP, Liang Z, Gao CX (2013b) [CRISPR/Cas: a novel way of RNA-guided genome editing]. *Yi chuan = Hereditas* **35**: 1265-1273

- Li L, Peterson CA, Lu X, Legerski RJ (1995) Mutations in XPA that prevent association with ERCC1 are defective in nucleotide excision repair. *Molecular and cellular biology* **15**: 1993-1998
- Lindsey-Boltz LA, Kemp MG, Reardon JT, DeRocco V, Iyer RR, Modrich P, Sancar A (2014) Coupling of human DNA excision repair and the DNA damage checkpoint in a defined in vitro system. *The Journal of biological chemistry* **289**: 5074-5082
- Lippke JA, Gordon LK, Brash DE, Haseltine WA (1981) Distribution of UV light-induced damage in a defined sequence of human DNA: detection of alkaline-sensitive lesions at pyrimidine nucleoside-cytidine sequences. *Proceedings of the National Academy of Sciences of the United States of America* **78**: 3388-3392
- Lombard DB, Chua KF, Mostoslavsky R, Franco S, Gostissa M, Alt FW (2005) DNA repair, genome stability, and aging. *Cell* **120**: 497-512
- Ma JL, Kim EM, Haber JE, Lee SE (2003) Yeast Mre11 and Rad1 proteins define a Ku-independent mechanism to repair double-strand breaks lacking overlapping end sequences. *Molecular and cellular biology* **23**: 8820-8828
- MacInnes MA, Mudgett JS (1990) Cloning of the functional human excision repair gene ERCC-5: potential gene regulatory features conserved with other human repair genes. *Progress in clinical and biological research* **340a**: 265-274
- Madan V, Lear JT, Szeimies RM (2010) Non-melanoma skin cancer. *Lancet (London, England)* **375**: 673-685
- Mak AN, Bradley P, Cernadas RA, Bogdanove AJ, Stoddard BL (2012) The crystal structure of TAL effector PthXo1 bound to its DNA target. *Science (New York, NY)* **335**: 716-719
- Marfori M, Mynott A, Ellis JJ, Mehdi AM, Saunders NF, Curmi PM, Forwood JK, Boden M, Kobe B (2011) Molecular basis for specificity of nuclear import and prediction of nuclear localization. *Biochimica et biophysica acta* **1813**: 1562-1577
- Marioni JC, Mason CE, Mane SM, Stephens M, Gilad Y (2008) RNA-seq: an assessment of technical reproducibility and comparison with gene expression arrays. *Genome research* **18**: 1509-1517
- Matlin AJ, Clark F, Smith CW (2005) Understanding alternative splicing: towards a cellular code. *Nature reviews Molecular cell biology* **6**: 386-398
- Matsumura Y, Ananthaswamy HN (2002) Short-term and long-term cellular and molecular events following UV irradiation of skin: implications for molecular medicine. *Expert reviews in molecular medicine* **4**: 1-22
- Matsumura Y, Nishigori C, Yagi T, Imamura S, Takebe H (1998) Characterization of molecular defects in xeroderma pigmentosum group F in relation to its clinically mild symptoms. *Human molecular genetics* **7**: 969-974
- Matsunaga T, Park CH, Bessho T, Mu D, Sancar A (1996) Replication protein A confers structure-specific endonuclease activities to the XPF-ERCC1 and XPG subunits of human DNA repair excision nuclease. *The Journal of biological chemistry* **271**: 11047-11050
- McHugh PJ, Spanswick VJ, Hartley JA (2001) Repair of DNA interstrand crosslinks: molecular mechanisms and clinical relevance. *The Lancet Oncology* **2**: 483-490
- McManus CJ, Graveley BR (2011) RNA structure and the mechanisms of alternative splicing. *Current opinion in genetics & development* **21**: 373-379
- McNeil EM, Melton DW (2012) DNA repair endonuclease ERCC1-XPF as a novel therapeutic target to overcome chemoresistance in cancer therapy. *Nucleic acids research* **40**: 9990-10004
- McVey M, Lee SE (2008) MMEJ repair of double-strand breaks (director's cut): deleted sequences and alternative endings. *Trends in genetics : TIG* **24**: 529-538
- 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. *Nature genetics* **5**: 217-224

- Melis JP, van Steeg H, Luijten M (2013) Oxidative DNA damage and nucleotide excision repair. *Antioxidants & redox signaling* **18**: 2409-2419
- Mellon I, Spivak G, Hanawalt PC (1987) Selective removal of transcription-blocking DNA damage from the transcribed strand of the mammalian DHFR gene. *Cell* **51**: 241-249
- Melnikova VO, Ananthaswamy HN (2005) Cellular and molecular events leading to the development of skin cancer. *Mutation research* **571**: 91-106
- Messaoud O, Ben Rekaya M, Kefi R, Chebel S, Boughammoura-Bouatay A, Bel Hadj Ali H, Gouider-Khouja N, Zili J, Frih-Ayed M, Mokhtar I, Abdelhak S, Zghal M (2010) Identification of a primarily neurological phenotypic expression of xeroderma pigmentosum complementation group A in a Tunisian family. *The British journal of dermatology* **162**: 883-886
- Michaels JE, Schimmel P, Shiba K, Miller WT (1996) Dominant negative inhibition by fragments of a monomeric enzyme. *Proceedings of the National Academy of Sciences of the United States of America* **93**: 14452-14455
- Milne GT, Weaver DT (1993) Dominant negative alleles of RAD52 reveal a DNA repair/recombination complex including Rad51 and Rad52. *Genes & development* **7**: 1755-1765
- Mitchell DL, Nairn RS (1989) The biology of the (6-4) photoproduct. *Photochemistry and photobiology* **49**: 805-819
- 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**: 7177-7186
- Mogi S, Oh DH (2006) gamma-H2AX formation in response to interstrand crosslinks requires XPF in human cells. *DNA repair* **5**: 731-740
- Moldovan GL, D'Andrea AD (2009) How the fanconi anemia pathway guards the genome. *Annual review of genetics* **43**: 223-249
- Moriwaki S, Stefanini M, Lehmann AR, Hoeijmakers JH, Robbins JH, Rapin I, Botta E, Tanganelli B, Vermeulen W, Broughton BC, Kraemer KH (1996) DNA repair and ultraviolet mutagenesis in cells from a new patient with xeroderma pigmentosum group G and cockayne syndrome resemble xeroderma pigmentosum cells. *The Journal of investigative dermatology* **107**: 647-653
- Moriwaki S, Yamashita Y, Nakamura S, Fujita D, Kohyama J, Takigawa M, Ohmichi M (2012) Prenatal diagnosis of xeroderma pigmentosum group A in Japan. *The Journal of dermatology* **39**: 516-519
- Moser J, Kool H, Giakzidis I, Caldecott K, Mullenders LH, Foustari MI (2007) Sealing of chromosomal DNA nicks during nucleotide excision repair requires XRCC1 and DNA ligase III alpha in a cell-cycle-specific manner. *Molecular cell* **27**: 311-323
- Mosmann T (1983) Rapid colorimetric assay for cellular growth and survival: application to proliferation and cytotoxicity assays. *Journal of immunological methods* **65**: 55-63
- Mu D, Bessho T, Nechev LV, Chen DJ, Harris TM, Hearst JE, Sancar A (2000) DNA interstrand cross-links induce futile repair synthesis in mammalian cell extracts. *Molecular and cellular biology* **20**: 2446-2454
- Mu D, Hsu DS, Sancar A (1996) Reaction mechanism of human DNA repair excision nuclease. *The Journal of biological chemistry* **271**: 8285-8294
- Mu D, Sancar A (1997) Model for XPC-independent transcription-coupled repair of pyrimidine dimers in humans. *The Journal of biological chemistry* **272**: 7570-7573
- Mulkidjanian AY, Bychkov AY, Dibrova DV, Galperin MY, Koonin EV (2012) Origin of first cells at terrestrial, anoxic geothermal fields. *Proceedings of the National Academy of Sciences of the United States of America* **109**: E821-830

- Munoz IM, Hain K, Declais AC, Gardiner M, Toh GW, Sanchez-Pulido L, Heuckmann JM, Toth R, Macartney T, Eppink B, Kanaar R, Ponting CP, Lilley DM, Rouse J (2009) Coordination of structure-specific nucleases by human SLX4/BTBD12 is required for DNA repair. *Molecular cell* **35**: 116-127
- 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. *Nature genetics* **37**: 1063-1071
- Murray D, Macann A, Hanson J, Rosenberg E (1996) ERCC1/ERCC4 5'-endonuclease activity as a determinant of hypoxic cell radiosensitivity. *International journal of radiation biology* **69**: 319-327
- Nakai K, Sakamoto H (1994) Construction of a novel database containing aberrant splicing mutations of mammalian genes. *Gene* **141**: 171-177
- Newby MI, Greenbaum NL (2002) Sculpting of the spliceosomal branch site recognition motif by a conserved pseudouridine. *Nature structural biology* **9**: 958-965
- Niedernhofer LJ, Essers J, Weeda G, Beverloo B, de Wit J, Muijtjens M, Odijk H, Hoeijmakers JH, Kanaar R (2001) The structure-specific endonuclease Ercc1-Xpf is required for targeted gene replacement in embryonic stem cells. *The EMBO journal* **20**: 6540-6549
- Niedernhofer LJ, Garinis GA, Raams A, Lalai AS, Robinson AR, Appeldoorn E, Odijk H, Oostendorp R, Ahmad A, van Leeuwen W, Theil AF, Vermeulen W, van der Horst GT, Meinecke P, Kleijer WJ, Vijg J, Jaspers NG, Hoeijmakers JH (2006) A new progeroid syndrome reveals that genotoxic stress suppresses the somatotroph axis. *Nature* **444**: 1038-1043
- Niedernhofer LJ, Odijk H, Budzowska M, van Drunen E, Maas A, Theil AF, de Wit J, Jaspers NG, Beverloo HB, Hoeijmakers JH, Kanaar R (2004) The structure-specific endonuclease Ercc1-Xpf is required to resolve DNA interstrand cross-link-induced double-strand breaks. *Molecular and cellular biology* **24**: 5776-5787
- Nishigori C, Zghal M, Yagi T, Imamura S, Komoun MR, Takebe H (1993) High prevalence of the point mutation in exon 6 of the xeroderma pigmentosum group A-complementing (XPAC) gene in xeroderma pigmentosum group A patients in Tunisia. *American journal of human genetics* **53**: 1001-1006
- Nospikel T (2009) DNA repair in mammalian cells : Nucleotide excision repair: variations on versatility. *Cellular and molecular life sciences : CMLS* **66**: 994-1009
- Nowotny M, Gaur V (2016) Structure and mechanism of nucleases regulated by SLX4. *Current opinion in structural biology* **36**: 97-105
- O'Donovan A, Davies AA, Moggs JG, West SC, Wood RD (1994) XPG endonuclease makes the 3' incision in human DNA nucleotide excision repair. *Nature* **371**: 432-435
- Ogi T, Limsirichaikul S, Overmeer RM, Volker M, Takenaka K, Cloney R, Nakazawa Y, Niimi A, Miki Y, Jaspers NG, Mullenders LH, Yamashita S, Fousteri MI, Lehmann AR (2010) Three DNA polymerases, recruited by different mechanisms, carry out NER repair synthesis in human cells. *Molecular cell* **37**: 714-727
- Okinaka RT, Perez-Castro AV, Sena A, Laubscher K, Strniste GF, Park MS, Hernandez R, MacInnes MA, Kraemer KH (1997) Heritable genetic alterations in a xeroderma pigmentosum group G/Cockayne syndrome pedigree. *Mutation research* **385**: 107-114
- Orelli B, McClendon TB, Tsodikov OV, Ellenberger T, Niedernhofer LJ, Scharer OD (2010) The XPA-binding domain of ERCC1 is required for nucleotide excision repair but not other DNA repair pathways. *The Journal of biological chemistry* **285**: 3705-3712
- Oren M (1999) Regulation of the p53 tumor suppressor protein. *The Journal of biological chemistry* **274**: 36031-36034
- Overmeer RM, Moser J, Volker M, Kool H, Tomkinson AE, van Zeeland AA, Mullenders LH, Fousteri M (2011) Replication protein A safeguards genome integrity by controlling NER incision events. *The Journal of cell biology* **192**: 401-415
- Pan Q, Shai O, Lee LJ, Frey BJ, Blencowe BJ (2008) Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. *Nature genetics* **40**: 1413-1415

- 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. *The Journal of biological chemistry* **270**: 4896-4902
- Pellagatti A, Dolatshad H, Valletta S, Boulwood J (2015) Application of CRISPR/Cas9 genome editing to the study and treatment of disease. *Archives of toxicology* **89**: 1023-1034
- Pena SD, Shokeir MH (1974) Syndrome of camptodactyly, multiple ankyloses, facial anomalies, and pulmonary hypoplasia: a lethal condition. *The Journal of pediatrics* **85**: 373-375
- Perez-Oliva AB, Lachaud C, Szyniarowski P, Munoz I, Macartney T, Hickson I, Rouse J, Alessi DR (2015) USP45 deubiquitylase controls ERCC1-XPF endonuclease-mediated DNA damage responses. *The EMBO journal* **34**: 326-343
- Pessa HK, Will CL, Meng X, Schneider C, Watkins NJ, Perala N, Nymark M, Turunen JJ, Luhrmann R, Frilander MJ (2008) Minor spliceosome components are predominantly localized in the nucleus. *Proceedings of the National Academy of Sciences of the United States of America* **105**: 8655-8660
- Pizzolato J, Mukherjee S, Scharer OD, Jiricny J (2015) FANCD2-associated nuclease 1, but not exonuclease 1 or flap endonuclease 1, is able to unhook DNA interstrand cross-links in vitro. *The Journal of biological chemistry* **290**: 22602-22611
- Pommier Y, Leo E, Zhang H, Marchand C (2010) DNA topoisomerases and their poisoning by anticancer and antibacterial drugs. *Chemistry & biology* **17**: 421-433
- Query CC, Moore MJ, Sharp PA (1994) Branch nucleophile selection in pre-mRNA splicing: evidence for the bulged duplex model. *Genes & development* **8**: 587-597
- Rahn JJ, Adair GM, Nairn RS (2010) Multiple roles of ERCC1-XPF in mammalian interstrand crosslink repair. *Environmental and molecular mutagenesis* **51**: 567-581
- Ran FA, Hsu PD, Wright J, Agarwala V, Scott DA, Zhang F (2013) Genome engineering using the CRISPR-Cas9 system. *Nature protocols* **8**: 2281-2308
- Rastogi RP, Richa, Kumar A, Tyagi MB, Sinha RP (2010) Molecular mechanisms of ultraviolet radiation-induced DNA damage and repair. *Journal of nucleic acids* **2010**: 592980
- Riedl T, Hanaoka F, Egly JM (2003) The comings and goings of nucleotide excision repair factors on damaged DNA. *The EMBO journal* **22**: 5293-5303
- Riley T, Sontag E, Chen P, Levine A (2008) Transcriptional control of human p53-regulated genes. *Nature reviews Molecular cell biology* **9**: 402-412
- Roca X, Olson AJ, Rao AR, Enerly E, Kristensen VN, Borresen-Dale AL, Andresen BS, Krainer AR, Sachidanandam R (2008) Features of 5'-splice-site efficiency derived from disease-causing mutations and comparative genomics. *Genome research* **18**: 77-87
- Rogers J (1985) Exon shuffling and intron insertion in serine protease genes. *Nature* **315**: 458-459
- Rothfuss A, Grompe M (2004) Repair kinetics of genomic interstrand DNA cross-links: evidence for DNA double-strand break-dependent activation of the Fanconi anemia/BRCA pathway. *Molecular and cellular biology* **24**: 123-134
- Sancar A, Rupp WD (1983) A novel repair enzyme: UVRABC excision nuclease of Escherichia coli cuts a DNA strand on both sides of the damaged region. *Cell* **33**: 249-260
- Sargent RG, Meservy JL, Perkins BD, Kilburn AE, Intody Z, Adair GM, Nairn RS, Wilson JH (2000) Role of the nucleotide excision repair gene ERCC1 in formation of recombination-dependent rearrangements in mammalian cells. *Nucleic acids research* **28**: 3771-3778
- 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. *Proceedings of the National Academy of Sciences of the United States of America* **94**: 13122-13127

- Saridaki Z, Liloglou T, Zafiroopoulos A, Koumantaki E, Zoras O, Spandidos DA (2003) Mutational analysis of CDKN2A genes in patients with squamous cell carcinoma of the skin. *The British journal of dermatology* **148**: 638-648
- 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. *Proceedings of the National Academy of Sciences of the United States of America* **87**: 9908-9912
- Schafer A, Schubert S, Gratchev A, Seebode C, Apel A, Laspe P, Hofmann L, Ohlenbusch A, Mori T, Kobayashi N, Schurer A, Schon MP, Emmert S (2013) Characterization of three XPG-defective patients identifies three missense mutations that impair repair and transcription. *The Journal of investigative dermatology* **133**: 1841-1849
- Scharer OD (2008) XPG: its products and biological roles. *Advances in experimental medicine and biology* **637**: 83-92
- Scharer OD (2013) Nucleotide excision repair in eukaryotes. *Cold Spring Harbor perspectives in biology* **5**: a012609
- Scherly D, Nospikel T, Corlet J, Ucla C, Bairoch A, Clarkson SG (1993) Complementation of the DNA repair defect in xeroderma pigmentosum group G cells by a human cDNA related to yeast RAD2. *Nature* **363**: 182-185
- Schubert S (2014) Characterization of the multifunctional XPG protein during Nucleotide-excision-repair. [Doctoral thesis]; Department of Dermatology, Venereology and Allergology, Georg-August-University Goettingen
- Schubert S, Emmert S (2016) Xeroderma Pigmentosum and related diseases. In: *Harper's Textbook of Pediatric Dermatology*: Chapter 29.1
- Schubert S, Lehmann J, Kalfon L, Slor H, Falik-Zaccari TC, Emmert S (2014) Clinical utility gene card for: Xeroderma pigmentosum. *European journal of human genetics : EJHG* **22** doi: 10.1038/ejhg.2013.233
- Seebode C, Lehmann J, Emmert S (2016) Photocarcinogenesis and Skin Cancer Prevention Strategies. *Anticancer research* **36**: 1371-1378
- Seibel NM, Eljouni J, Nalaskowski MM, Hampe W (2007) Nuclear localization of enhanced green fluorescent protein homomultimers. *Analytical biochemistry* **368**: 95-99
- Seluanov A, Mao Z, Gorbunova V (2010) Analysis of DNA double-strand break (DSB) repair in mammalian cells. *Journal of visualized experiments : JoVE*
- Sengerova B, Wang AT, McHugh PJ (2011) Orchestrating the nucleases involved in DNA interstrand cross-link (ICL) repair. *Cell cycle (Georgetown, Tex)* **10**: 3999-4008
- Sertic S, Pizzi S, Cloney R, Lehmann AR, Marini F, Plevani P, Muzi-Falconi M (2011) Human exonuclease 1 connects nucleotide excision repair (NER) processing with checkpoint activation in response to UV irradiation. *Proceedings of the National Academy of Sciences of the United States of America* **108**: 13647-13652
- Sgouros J, Gaillard PH, Wood RD (1999) A relationship between a DNA-repair/recombination nuclease family and archaeal helicases. *Trends in biochemical sciences* **24**: 95-97
- Sharpless E, Chin L (2003) The INK4a/ARF locus and melanoma. *Oncogene* **22**: 3092-3098
- Shen H, Sturgis EM, Khan SG, Qiao Y, Shahnavi T, Eicher SA, Xu Y, Wang X, Strom SS, Spitz MR, Kraemer KH, Wei Q (2001) An intronic poly (AT) polymorphism of the DNA repair gene XPC and risk of squamous cell carcinoma of the head and neck: a case-control study. *Cancer research* **61**: 3321-3325
- Shiloh Y (2003) ATM and related protein kinases: safeguarding genome integrity. *Nature reviews Cancer* **3**: 155-168
- Sijbers AM, de Laat WL, Ariza RR, Biggerstaff M, Wei YF, Moggs JG, Carter KC, Shell BK, Evans E, de Jong MC, Rademakers S, de Rooij J, Jaspers NG, Hoeijmakers JH, Wood RD (1996a) Xeroderma pigmentosum group F caused by a defect in a structure-specific DNA repair endonuclease. *Cell* **86**: 811-822

- Sijbers AM, van der Spek PJ, Odijk H, van den Berg J, van Duin M, Westerveld A, Jaspers NG, Bootsma D, Hoeijmakers JH (1996b) Mutational analysis of the human nucleotide excision repair gene ERCC1. *Nucleic acids research* **24**: 3370-3380
- Slupphaug G, Kavli B, Krokan HE (2003) The interacting pathways for prevention and repair of oxidative DNA damage. *Mutation research* **531**: 231-251
- Soehnge H, Ouhtit A, Ananthaswamy ON (1997) Mechanisms of induction of skin cancer by UV radiation. *Frontiers in bioscience : a journal and virtual library* **2**: d538-551
- Soltys DT, Rocha CR, Lerner LK, de Souza TA, Munford V, Cabral F, Nardo T, Stefanini M, Sarasin A, Cabral-Neto JB, Menck CF (2013) Novel XPG (ERCC5) mutations affect DNA repair and cell survival after ultraviolet but not oxidative stress. *Human mutation* **34**: 481-489
- Soufir N, Ged C, Bourillon A, Austerlitz F, Chemin C, Stary A, Armier J, Pham D, Khadir K, Roume J, Hadj-Rabia S, Bouadjar B, Taieb A, de Verneuil H, Benchiki H, Grandchamp B, Sarasin A (2010) A prevalent mutation with founder effect in xeroderma pigmentosum group C from north Africa. *The Journal of investigative dermatology* **130**: 1537-1542
- Sperling J, Azubel M, Sperling R (2008) Structure and function of the Pre-mRNA splicing machine. *Structure (London, England : 1993)* **16**: 1605-1615
- Srebrow A, Kornbliht AR (2006) The connection between splicing and cancer. *Journal of cell science* **119**: 2635-2641
- Staresinic L, Fagbemi AF, Enzlin JH, Gourdin AM, Wijgers N, Dunand-Sauthier I, Giglia-Mari G, Clarkson SG, Vermeulen W, Schärer OD (2009) Coordination of dual incision and repair synthesis in human nucleotide excision repair. *The EMBO journal* **28**: 1111-1120
- Su Y, Orelli B, Madireddy A, Niedernhofer LJ, Schärer OD (2012) Multiple DNA binding domains mediate the function of the ERCC1-XPF protein in nucleotide excision repair. *The Journal of biological chemistry* **287**: 21846-21855
- 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. *Molecular cell* **2**: 223-232
- Sugasawa K, Okuda Y, Saijo M, Nishi R, Matsuda N, Chu G, Mori T, Iwai S, Tanaka K, Tanaka K, Hanaoka F (2005) UV-induced ubiquitylation of XPC protein mediated by UV-DDB-ubiquitin ligase complex. *Cell* **121**: 387-400
- Sun K, Gong A, Liang P (2015) Predictive impact of genetic polymorphisms in DNA repair genes on susceptibility and therapeutic outcomes to colorectal cancer patients. *Tumour biology : the journal of the International Society for Oncodevelopmental Biology and Medicine* **36**: 1549-1559
- Suzuki K, Yu C, Qu J, Li M, Yao X, Yuan T, Goebel A, Tang S, Ren R, Aizawa E, Zhang F, Xu X, Soligalla RD, Chen F, Kim J, Kim NY, Liao HK, Benner C, Esteban CR, Jin Y, Liu GH, Li Y, Izpisua Belmonte JC (2014) Targeted gene correction minimally impacts whole-genome mutational load in human-disease-specific induced pluripotent stem cell clones. *Cell stem cell* **15**: 31-36
- Szostak JW, Orr-Weaver TL, Rothstein RJ, Stahl FW (1983) The double-strand-break repair model for recombination. *Cell* **33**: 25-35
- Takahashi D, Sato K, Hirayama E, Takata M, Kurumizaka H (2015) Human FAN1 promotes strand incision in 5'-flapped DNA complexed with RPA. *Journal of biochemistry* **158**: 263-270
- Takahata C, Masuda Y, Takedachi A, Tanaka K, Iwai S, Kuraoka I (2015) Repair synthesis step involving ERCC1-XPF participates in DNA repair of the Top1-DNA damage complex. *Carcinogenesis* **36**: 841-851
- Takata M, Sasaki MS, Sonoda E, Morrison C, Hashimoto M, Utsumi H, Yamaguchi-Iwai Y, Shinohara A, Takeda S (1998) Homologous recombination and non-homologous end-joining pathways of DNA double-strand break repair have overlapping roles in the maintenance of chromosomal integrity in vertebrate cells. *The EMBO journal* **17**: 5497-5508

- Tamura D, DiGiovanna JJ, Kraemer KH (2010) Founder mutations in xeroderma pigmentosum. *The Journal of investigative dermatology* **130**: 1491-1493
- Tapias A, Auriol J, Forget D, Enzlin JH, Scharer OD, Coin F, Coulombe B, Egly JM (2004) Ordered conformational changes in damaged DNA induced by nucleotide excision repair factors. *The Journal of biological chemistry* **279**: 19074-19083
- Thoma F (1999) Light and dark in chromatin repair: repair of UV-induced DNA lesions by photolyase and nucleotide excision repair. *The EMBO journal* **18**: 6585-6598
- Thorel F, Constantinou A, Dunand-Sauthier I, Nospikel T, Lalle P, Raams A, Jaspers NG, Vermeulen W, Shivji MK, Wood RD, Clarkson SG (2004) Definition of a short region of XPG necessary for TFIIH interaction and stable recruitment to sites of UV damage. *Molecular and cellular biology* **24**: 10670-10680
- Tian M, Shinkura R, Shinkura N, Alt FW (2004) Growth retardation, early death, and DNA repair defects in mice deficient for the nucleotide excision repair enzyme XPF. *Molecular and cellular biology* **24**: 1200-1205
- Tillhon M, Cazzalini O, Nardo T, Necchi D, Sommatis S, Stivala LA, Scovassi AI, Prosperi E (2012) p300/CBP acetyl transferases interact with and acetylate the nucleotide excision repair factor XPG. *DNA repair* **11**: 844-852
- Tomicic MT, Kaina B (2013) Topoisomerase degradation, DSB repair, p53 and IAPs in cancer cell resistance to camptothecin-like topoisomerase I inhibitors. *Biochimica et biophysica acta* **1835**: 11-27
- Tripsianes K, Folkers G, Ab E, Das D, Odijk H, Jaspers NG, Hoeijmakers JH, Kaptein R, Boelens R (2005) The structure of the human ERCC1/XPF interaction domains reveals a complementary role for the two proteins in nucleotide excision repair. *Structure (London, England : 1993)* **13**: 1849-1858
- Truglio JJ, Croteau DL, Van Houten B, Kisker C (2006) Prokaryotic nucleotide excision repair: the UvrABC system. *Chemical reviews* **106**: 233-252
- Tsodikov OV, Enzlin JH, Scharer OD, Ellenberger T (2005) Crystal structure and DNA binding functions of ERCC1, a subunit of the DNA structure-specific endonuclease XPF-ERCC1. *Proceedings of the National Academy of Sciences of the United States of America* **102**: 11236-11241
- van Duin M, de Wit J, Odijk H, Westerveld A, Yasui A, Koken MH, Hoeijmakers JH, Bootsma D (1986) Molecular characterization of the human excision repair gene ERCC-1: cDNA cloning and amino acid homology with the yeast DNA repair gene RAD10. *Cell* **44**: 913-923
- van Gool AJ, van der Horst GT, Citterio E, Hoeijmakers JH (1997) Cockayne syndrome: defective repair of transcription? *The EMBO journal* **16**: 4155-4162
- van Vuuren AJ, Appeldoorn E, Odijk H, Yasui A, Jaspers NG, Bootsma D, Hoeijmakers JH (1993) Evidence for a repair enzyme complex involving ERCC1 and complementing activities of ERCC4, ERCC11 and xeroderma pigmentosum group F. *The EMBO journal* **12**: 3693-3701
- Veres A, Gosis BS, Ding Q, Collins R, Ragavendran A, Brand H, Erdin S, Cowan CA, Talkowski ME, Musunuru K (2014) Low incidence of off-target mutations in individual CRISPR-Cas9 and TALEN targeted human stem cell clones detected by whole-genome sequencing. *Cell stem cell* **15**: 27-30
- Vink AA, Roza L (2001) Biological consequences of cyclobutane pyrimidine dimers. *Journal of photochemistry and photobiology B, Biology* **65**: 101-104
- Wahl MC, Will CL, Luhrmann R (2009) The spliceosome: design principles of a dynamic RNP machine. *Cell* **136**: 701-718
- Wakasugi M, Reardon JT, Sancar A (1997) The non-catalytic function of XPG protein during dual incision in human nucleotide excision repair. *The Journal of biological chemistry* **272**: 16030-16034
- Wan B, Yin J, Horvath K, Sarkar J, Chen Y, Wu J, Wan K, Lu J, Gu P, Yu EY, Lue NF, Chang S, Liu Y, Lei M (2013) SLX4 assembles a telomere maintenance toolkit by bridging multiple endonucleases with telomeres. *Cell reports* **4**: 861-869

- Wang Y, Digiovanna JJ, Stern JB, Hornyak TJ, Raffeld M, Khan SG, Oh KS, Hollander MC, Dennis PA, Kraemer KH (2009a) Evidence of ultraviolet type mutations in xeroderma pigmentosum melanomas. *Proceedings of the National Academy of Sciences of the United States of America* **106**: 6279-6284
- Wang Z, Burge CB (2008) Splicing regulation: from a parts list of regulatory elements to an integrated splicing code. *RNA (New York, NY)* **14**: 802-813
- Wang Z, Gerstein M, Snyder M (2009b) RNA-Seq: a revolutionary tool for transcriptomics. *Nature reviews Genetics* **10**: 57-63
- Weeda G, Donker I, de Wit J, Morreau H, Janssens R, Vissers CJ, Nigg A, van Steeg H, Bootsma D, Hoeijmakers JH (1997) Disruption of mouse ERCC1 results in a novel repair syndrome with growth failure, nuclear abnormalities and senescence. *Current biology : CB* **7**: 427-439
- Wei C, Liu J, Yu Z, Zhang B, Gao G, Jiao R (2013) TALEN or Cas9 - rapid, efficient and specific choices for genome modifications. *Journal of genetics and genomics = Yi chuan xue bao* **40**: 281-289
- Wei J, Zaika E, Zaika A (2012) p53 Family: Role of Protein Isoforms in Human Cancer. *Journal of nucleic acids* **2012**: 687359
- 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. *International journal of cancer* **110**: 352-361
- Whiley PJ, Guidugli L, Walker LC, Healey S, Thompson BA, Lakhani SR, Da Silva LM, Tavtigian SV, Goldgar DE, Brown MA, Couch FJ, Spurdle AB (2011) Splicing and multifactorial analysis of intronic BRCA1 and BRCA2 sequence variants identifies clinically significant splicing aberrations up to 12 nucleotides from the intron/exon boundary. *Human mutation* **32**: 678-687
- Will CL, Luhrmann R (2011) Spliceosome structure and function. *Cold Spring Harbor perspectives in biology* **3**
- Wood AJ, Lo TW, Zeitler B, Pickle CS, Ralston EJ, Lee AH, Amora R, Miller JC, Leung E, Meng X, Zhang L, Rebar EJ, Gregory PD, Urnov FD, Meyer BJ (2011) Targeted genome editing across species using ZFNs and TALENs. *Science (New York, NY)* **333**: 307
- Wood RD (2010) Mammalian nucleotide excision repair proteins and interstrand crosslink repair. *Environmental and molecular mutagenesis* **51**: 520-526
- Wood RD, Burki HJ, Hughes M, Poley A (1983) Radiation-induced lethality and mutation in a repair-deficient CHO cell line. *International journal of radiation biology and related studies in physics, chemistry, and medicine* **43**: 207-213
- Wu Y, Mitchell TR, Zhu XD (2008) Human XPF controls TRF2 and telomere length maintenance through distinctive mechanisms. *Mechanisms of ageing and development* **129**: 602-610
- Wyatt HD, Sarbajna S, Matos J, West SC (2013) Coordinated actions of SLX1-SLX4 and MUS81-EME1 for Holliday junction resolution in human cells. *Molecular cell* **52**: 234-247
- Yagi T, Wood RD, Takebe H (1997) A low content of ERCC1 and a 120 kDa protein is a frequent feature of group F xeroderma pigmentosum fibroblast cells. *Mutagenesis* **12**: 41-44
- Yamamoto KN, Kobayashi S, Tsuda M, Kurumizaka H, Takata M, Kono K, Jiricny J, Takeda S, Hirota K (2011) Involvement of SLX4 in interstrand cross-link repair is regulated by the Fanconi anemia pathway. *Proceedings of the National Academy of Sciences of the United States of America* **108**: 6492-6496
- Yan CT, Boboila C, Souza EK, Franco S, Hickernell TR, Murphy M, Gumaste S, Geyer M, Zarrin AA, Manis JP, Rajewsky K, Alt FW (2007) IgH class switching and translocations use a robust non-classical end-joining pathway. *Nature* **449**: 478-482
- Yang H, Wang H, Shivalila CS, Cheng AW, Shi L, Jaenisch R (2013) One-step generation of mice carrying reporter and conditional alleles by CRISPR/Cas-mediated genome engineering. *Cell* **154**: 1370-1379
- Yi Q, Tang L (2011) Alternative spliced variants as biomarkers of colorectal cancer. *Current drug metabolism* **12**: 966-974

- Yokoyama H, Mizutani R, Satow Y, Sato K, Komatsu Y, Ohtsuka E, Nikaido O (2012) Structure of the DNA (6-4) photoproduct dTT(6-4)TT in complex with the 64M-2 antibody Fab fragment implies increased antibody-binding affinity by the flanking nucleotides. *Acta crystallographica Section D, Biological crystallography* **68**: 232-238
- You YH, Lee DH, Yoon JH, Nakajima S, Yasui A, Pfeifer GP (2001) Cyclobutane pyrimidine dimers are responsible for the vast majority of mutations induced by UVB irradiation in mammalian cells. *The Journal of biological chemistry* **276**: 44688-44694
- Zhang F, Scheerer P, Oberpichler I, Lamparter T, Krauss N (2013a) Crystal structure of a prokaryotic (6-4) photolyase with an Fe-S cluster and a 6,7-dimethyl-8-ribityllumazine antenna chromophore. *Proceedings of the National Academy of Sciences of the United States of America* **110**: 7217-7222
- Zhang J, Walter JC (2014) Mechanism and regulation of incisions during DNA interstrand cross-link repair. *DNA repair* **19**: 135-142
- Zhang Z, Pal S, Bi Y, Tchou J, Davuluri RV (2013b) Isoform level expression profiles provide better cancer signatures than gene level expression profiles. *Genome medicine* **5**: 33
- Zhao S, Liu W, Li Y, Liu P, Li S, Dou D, Wang Y, Yang R, Xiang R, Liu F (2016) Alternative Splice Variants Modulates Dominant-Negative Function of Helios in T-Cell Leukemia. *PloS one* **11**: e0163328
- Zhu XD, Niedernhofer L, Kuster B, Mann M, Hoeijmakers JH, de Lange T (2003) ERCC1/XPF removes the 3' overhang from uncapped telomeres and represses formation of telomeric DNA-containing double minute chromosomes. *Molecular cell* **12**: 1489-1498

8 Curriculum vitae

Personal details

Name Janin Lehmann
 born 21st March 1989 in Westerstede, Germany

Address Clinic and Policlinic for Dermatology and Venereology
 University Medical Center Rostock
 Stempelstrasse 13
 18057 Rostock

Education

02/2016 – today Continuation of PhD thesis after laboratory move: “Functional relevance of spontaneous alternative splice variants of xeroderma pigmentosum genes: prognostic markers for skin cancer risk and disease outcome?”; Clinic and Policlinic for Dermatology and Venereology, University Medical Center Rostock

12/2015 – 01/2016 Organization and execution of laboratory move

03/2014 – 11/2015 PhD thesis at the Department of Dermatology, Venereology and Allergology, University Medical Center Goettingen, supervised by Prof. Dr. med. Steffen Emmert with the title: "Functional relevance of spontaneous alternative splice variants of xeroderma pigmentosum genes: Prognostic marker for skin cancer risk and disease outcome?".

05/2013 – 02/2014 Research assistant at the Department of Dermatology, Venereology and Allergology, University Medical Center Goettingen working on “Characterization of xeroderma pigmentosum patients”.

10/2012 – 04/2013 Master thesis at the Department of Dermatology, Venereology and Allergology, University Medical Center Goettingen, supervised by Prof. Dr. med. Steffen Emmert with the title: "Characterization of novel xeroderma pigmentosum group G spontaneous mRNA splice variants."

10/2011 – 04/2013 Master of Science in Molecular Medicine, Georg-August-University Goettingen, Germany

03/2011 – 07/2011 Bachelor thesis at the Department of Dermatology, University of Luebeck, supervised by Prof. Dr. med. Ralf Paus with the title: "The influence of nestin+ progenitor cells on wound healing in human skin."

10/2008 – 09/2011 Bachelor of Science in Molecular Life Science, University of Luebeck, Germany

05/2008 High School Diploma at the Europaschule-Gymnasium Westerstede, Germany